frontiers | Research Topics

Digital transformation in construction: integrating metaverse, digital twin, and BIM

Edited by
Izuru Takewaki, Zheng Lu, Salman Azhar, Vagelis Plevris, Zhen Chen, Sakdirat Kaewunruen and Umberto Berardi

Published in
Frontiers in Built Environment
Frontiers in Virtual Reality
Frontiers in Environmental Science



FRONTIERS EBOOK COPYRIGHT STATEMENT

The copyright in the text of individual articles in this ebook is the property of their respective authors or their respective institutions or funders. The copyright in graphics and images within each article may be subject to copyright of other parties. In both cases this is subject to a license granted to Frontiers.

The compilation of articles constituting this ebook is the property of Frontiers.

Each article within this ebook, and the ebook itself, are published under the most recent version of the Creative Commons CC-BY licence. The version current at the date of publication of this ebook is CC-BY 4.0. If the CC-BY licence is updated, the licence granted by Frontiers is automatically updated to the new version.

When exercising any right under the CC-BY licence, Frontiers must be attributed as the original publisher of the article or ebook, as applicable.

Authors have the responsibility of ensuring that any graphics or other materials which are the property of others may be included in the CC-BY licence, but this should be checked before relying on the CC-BY licence to reproduce those materials. Any copyright notices relating to those materials must be complied with.

Copyright and source acknowledgement notices may not be removed and must be displayed in any copy, derivative work or partial copy which includes the elements in question.

All copyright, and all rights therein, are protected by national and international copyright laws. The above represents a summary only. For further information please read Frontiers' Conditions for Website Use and Copyright Statement, and the applicable CC-BY licence.

ISSN 1664-8714
ISBN 978-2-8325-8500-9
DOI 10.3389/978-2-8325-8500-9

Generative AI statement
Any alternative text (Alt text) provided alongside figures in the articles in this ebook has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

About Frontiers

Frontiers is more than just an open access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers journal series

The Frontiers journal series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the *Frontiers journal series* operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews. Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the *Frontiers journals series*: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area.

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers editorial office: frontiersin.org/about/contact

Digital transformation in construction: integrating metaverse, digital twin, and BIM

Topic editors

Izuru Takewaki — Kyoto Arts and Crafts University, Japan
Zheng Lu — Tongji University, China
Salman Azhar — Auburn University, United States
Vagelis Plevris — Qatar University, Qatar
Zhen Chen — University of Strathclyde, United Kingdom
Sakdirat Kaewunruen — University of Birmingham, United Kingdom
Umberto Berardi — Toronto Metropolitan University, Canada

Citation

Takewaki, I., Lu, Z., Azhar, S., Plevris, V., Chen, Z., Kaewunruen, S., Berardi, U., eds. (2026). *Digital transformation in construction: integrating metaverse, digital twin, and BIM*. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-8325-8500-9

Table of contents

05 Editorial: Digital transformation in construction: integrating metaverse, digital twin, and BIM
Izuru Takewaki, Zheng Lu, Salman Azhar, Vagelis Plevris, Zhen Chen, Sakdirat Kaewunruen and Umberto Berardi

09 Transitioning from BIM to Digital Twin to Metaverse
Ehsan Noroozinejad Farsangi, Alaa O. Shehata, Maria Rashidi, Nargess Ghassempour and Seyedali Mirjalili

16 Sustainable innovations in digital twin technology: a systematic review about energy efficiency and indoor environment quality in built environment
Natakarani Venkateswarlu and Mahenthiran Sathiyamoorthy

35 Opportunities for digital tracking technologies in the precast concrete sector in Sweden
Maryam Olaoti Shehu, Roma Almeida and Krushna Mahapatra

51 Semantic and ontology-based analysis of regulatory documents for construction industry digitalization
Zarina Kabzhan, Alexandr Shakhnovich, Sergey Gorshkov, Yussuf Yemenov, Fedor Gorshkov and Nazym Shogelova

66 Correction: Semantic and ontology-based analysis of regulatory documents for construction industry digitalization
Zarina Kabzhan, Alexandr Shakhnovich, Sergey Gorshkov, Yussuf Yemenov, Fedor Gorshkov and Nazym Shogelova

67 Data enrichment for semantic segmentation of point clouds for the generation of geometric-semantic road models
David Crampen and Joerg Blankenbach

86 Digital transformation of construction enterprises and carbon emission reduction: evidence from listed companies
Sanglin Zhao, Hao Deng, Jikang Cao and Måns Gustaf

109 Strategies for bridge maintenance using BIM: an analysis of methodologies and tools
Emilio Medrano-Sanchez and Erwin Martos

122 Evolution and research trends in virtual reality and augmented reality technologies for the architecture, engineering, and construction (AEC) industry: a systematic review and science mapping approach
Diego Sornoza-Parrales, Günther Oberdam Baque Vera, María Leonor Parrales Poveda and Tania Maricela Macías-Parrales

137 Digital transformation through multi-device HBIM workflow: a case study on supporting the adaptive reuse of the Odd Fellows Building in Atlanta, Georgia
Botao Li and Danielle S. Willkens

148 Integrating thermal point clouds into BIM-GIS environments: workflow proposal for multi-layer digital twins
Amanda Ramón-Constantí, Francisco Javier Castilla Pascual, Borja Garcia de Soto and Antonio Adán-Oliver

168 Artificial intelligence-driven safety assessment of scaffolding using LiDAR sensing
Sameer Prabhu, Amit Patwardhan and Ramin Karim

182 Interpretable machine learning for predicting the bearing capacity of double shear-bolted connections: a data-driven evaluation
Soheila Kookalani, Hongchen Liu, Tirtharaj Dash, Alwyn Mathew and Ioannis Brilakis

208 Integrating BIM and VR for enhanced safety communication and training in construction
Mehamooda Baikati, Arif Khan Baikati, M. P. Saravana Kumar and Prasanna Venkatesan Ramani

229 The role of digital technologies in engineering procurement: a systematic literature review
Imoleayo A. Awodele, Molusiwa S. Ramabodu, Nathaniel Ayinde Olatunde, Emmanuel C. Eze, Iruka C. Anugwo, Yesica P. Espinosa, Temitope O. Kehinde, Bamidele T. Arijeloye and Alice T. Ogunlolu

243 An efficient deep learning framework for text detection and recognition in engineering drawings
Dakan Ying, Lyu Guanghua, Zhengxu He, Rija Hasan and Syed Hadi Hussain Shah

frontiers in Built Environment

TYPE Editorial
PUBLISHED 22 July 2026
DOI 10.3389/fbuil.2026.1905914

Check for updates

OPEN ACCESS

EDITED AND REVIEWED BY
Solomon Tesfamariam,
University of Waterloo, Canada

*CORRESPONDENCE
Salman Azhar,
sza0001@auburn.edu

RECEIVED 10 June 2026
ACCEPTED 11 June 2026
PUBLISHED 22 July 2026

CITATION
Takewaki I, Lu Z, Azhar S, Plevris V, Chen Z, Kaewunruen S and Berardi U (2026) Editorial: Digital transformation in construction: integrating metaverse, digital twin, and BIM.
Front. Built Environ. 12:1905914.
doi: 10.3389/fbuil.2026.1905914

COPYRIGHT
© 2026 Takewaki, Lu, Azhar, Plevris, Chen, Kaewunruen and Berardi. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Editorial: Digital transformation in construction: integrating metaverse, digital twin, and BIM

Izuru Takewaki[1], Zheng Lu[2], Salman Azhar[3]*, Vagelis Plevris[4], Zhen Chen[5], Sakdirat Kaewunruen[6] and Umberto Berardi[7]

[1]Kyoto Arts and Crafts University, Kyoto, Japan, [2]Tongji University, Shanghai, China, [3]Auburn University, Auburn, AL, United States, [4]Qatar University, Doha, Qatar, [5]University of Strathclyde, Glasgow, United Kingdom, [6]University of Birmingham, Birmingham, United Kingdom, [7]Politecnico di Bari, Bari, Italy

KEYWORDS
artificial intelligence, building information modelling, digital twin, immersive technology, metaverse, resilience, sustainability

Editorial on the Research Topic
Digital transformation in construction: integrating metaverse, digital twin, and BIM

Introduction

The Research Topic "*Digital Transformation in Construction: Integrating Metaverse, Digital Twin, and BIM*" brings together scholarship on how data-rich, model-based, and immersive technologies are reshaping the built environment. Its aim is to examine how BIM, digital twins, and metaverse-related technologies can improve sustainability, resilience, safety, productivity, and lifecycle decision-making in construction and infrastructure. This focus is timely because immersive technology for the built world was identified among the World Economic Forum top emerging technologies of 2024 (World Economic Forum, 2024), while buildings and construction account for about 37% of global emissions in recent global reporting (UNEP and GlobalABC, 2025).

This Research Topic was initiated by Professor Izuru Takewaki and developed with the joint support of several specialty sections within *Frontiers in Built Environment*, including Earthquake Engineering, Transportation and Transit Systems, Construction Management, Structural Engineering and Design, Building Information Modelling (BIM), Sustainable Design and Construction, and Structural Sensing, Control and Asset Management. Its interdisciplinary scope was further strengthened through participation from *Frontiers in Environmental Science* and *Frontiers in Virtual Reality*, reflecting the broad relevance of digital transformation, BIM, digital twins, and immersive technologies across the built environment.

The articles in this Research Topic advance the field by moving beyond isolated digital tools toward connected cyber-physical and socio-technical ecosystems. BIM provides structured object, process, and asset information; digital twins extend this information through sensing, feedback, simulation, and lifecycle updating; and metaverse environments offer immersive spaces for visualization, training, coordination, and decision support. This progression is consistent with earlier research that framed the digital twin as a bridge

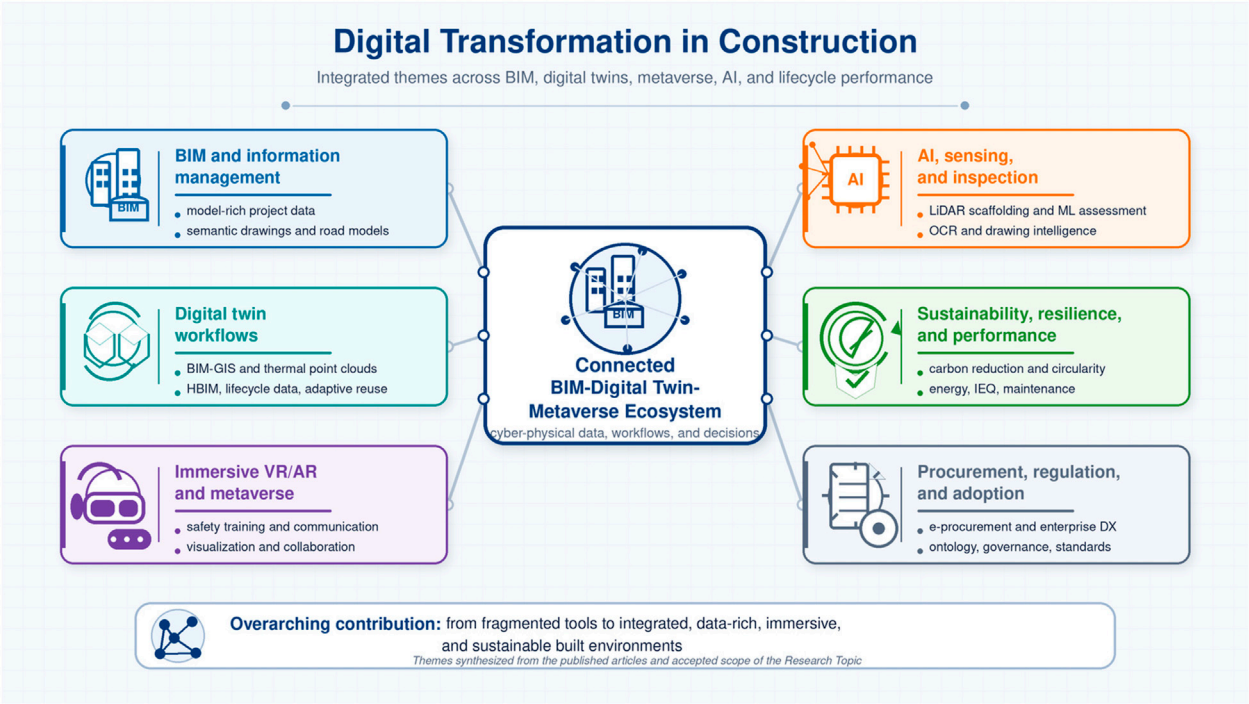


FIGURE 1
Integrated thematic framework of the Research Topic, showing how BIM, digital twins, immersive technologies, AI-enabled sensing, sustainability performance, and governance-oriented adoption collectively support digital transformation in construction and the built environment.

between physical systems and their virtual representations (Kritzinger et al., 2018), emphasized the need for semantic construction digital twins (Boje et al., 2020), and mapped the evolution from BIM to digital twins in the AEC-FM industry (Deng et al., 2021).

This Research Topic contributes to knowledge by connecting research streams that are often examined separately: BIM-based information management, AI-enabled analytics, sensing and reality capture, digital twin workflows, immersive visualization, and digital transformation for sustainability. The papers operate across project, asset, enterprise, and sector scales, addressing safety communication, inspection, drawing digitization, heritage documentation, road models, procurement, carbon reduction, circularity, and regulatory digitalization. Together, they show that digital transformation is not a single technology shift, but a coordinated change in data, workflows, skills, governance, and human interaction.

The broader contribution of the collection is its emphasis on implementation. The papers move beyond conceptual enthusiasm for BIM, digital twins, and metaverse technologies by showing how these tools can be evaluated through pilot studies, computational models, systematic reviews, case studies, and methods papers. They also identify persistent research needs, including interoperability, semantic enrichment, cybersecurity, data governance, ethical use, workforce adoption, and field validation. In doing so, the collection helps define an applied research agenda for a built environment that is more intelligent, resilient, sustainable, and inclusive. Figure 1 summarizes the thematic scope of this Research Topic, showing how BIM-based information management, digital twin workflows, immersive environments, AI-enabled sensing, sustainability performance, and governance-oriented adoption collectively support digital transformation in the built environment.

Organization of the special collection

The published contributions in this Research Topic can be organized around five major sub-themes: digital safety and AI-enabled construction intelligence; digital twins, reality capture, heritage documentation, and circularity; evidence synthesis through systematic and narrative reviews; methods for regulatory digitalization; and opinion-based conceptual framing of the BIM, digital twin, and metaverse transition.

- **Digital safety and AI-enabled construction intelligence:** Baikati et al., in *Integrating BIM and VR for enhanced safety communication and training in construction*, develop BIM-enabled VR safety training for multilingual construction environments. Ying et al., in *An efficient deep learning framework for text detection and recognition in engineering drawings*, propose a deep learning approach for extracting rebar annotations from complex engineering drawings. Kookalani et al., in *Interpretable machine learning for predicting the bearing capacity of double shear-bolted connections: a data-driven evaluation*, combine predictive modeling with interpretability to support data-driven steel design. Prabhu et al., in

Artificial intelligence-driven safety assessment of scaffolding using LiDAR sensing, present a cloud-based workflow for automated scaffolding inspection.

- **Digital twins, reality capture, heritage documentation, enterprise transformation, and circularity:** Ramón-Constantí et al., in *Integrating thermal point clouds into BIM-GIS environments: workflow proposal for multi-layer digital twins*, propose a BIM-GIS workflow for multi-layer thermal digital twins. Li and Willkens, in *Digital transformation through multi-device HBIM workflow: a case study on supporting the adaptive reuse of the Odd Fellows Building in Atlanta, Georgia*, present a scan-to-HBIM workflow for adaptive reuse. Crampen and Blankenbach, in *Data enrichment for semantic segmentation of point clouds for the generation of geometric-semantic road models*, enrich point-cloud segmentation with semantic information for road modeling. Zhao et al., in *Digital transformation of construction enterprises and carbon emission reduction: evidence from listed companies*, provide listed-company evidence on digital transformation and carbon reduction. Shehu et al., in *Opportunities for digital tracking technologies in the precast concrete sector in Sweden*, examine digital tracking technologies for circular economy practices in precast concrete.
- **Systematic and narrative review contributions:** Sornoza-Parrales et al., in *Evolution and research trends in virtual reality and augmented reality technologies for the architecture, engineering, and construction industry: a systematic review and science mapping approach*, use systematic review and science mapping methods to clarify the development of AR and VR research in AEC. Awodele et al. (2026), in *The role of digital technologies in engineering procurement: a systematic literature review*, synthesize literature on AI, blockchain, IoT, BIM, robotic process automation, and e-procurement. Medrano-Sanchez and Martos, in *Strategies for bridge maintenance using BIM: an analysis of methodologies and tools*, review BIM-based bridge maintenance methods and tools. Venkateswarlu and Sathiyamoorthy, in *Sustainable innovations in digital twin technology: a systematic review about energy efficiency and indoor environment quality in built environment*, review how digital twins can support energy efficiency, indoor environmental quality, and sustainable building operation.
- **Methods contribution:** Kabzhan et al., in *Semantic and ontology-based analysis of regulatory documents for construction industry digitalization*, develop an automated approach for analyzing construction regulatory documents using ontology modeling and natural language processing. This contribution addresses a central barrier to digital transformation: converting complex, changing, and text-heavy regulatory requirements into structured knowledge that can support digital workflows
- **Opinion contribution:** Noroozinejad Farsangi et al., in *Transitioning from BIM to Digital Twin to Metaverse*, provide a conceptual framing for the movement from static information models to dynamic digital twins and then to immersive, interconnected metaverse environments. This opinion paper helps anchor the collection by clarifying the conceptual pathway that links many of the empirical, methodological, and review contributions.

Conclusion

Collectively, the papers in this Research Topic show that digital transformation in construction is entering a more integrated and evidence-based phase. The field is moving from stand-alone BIM models and isolated visualization tools toward interoperable digital twins, intelligent sensing, AI-enabled analytics, semantic information structures, immersive interfaces, and digitally supported sustainability strategies. The collection advances knowledge by connecting technical innovation with real problems in safety, infrastructure management, heritage conservation, procurement, carbon reduction, circularity, and regulatory modernization.

Future work may extend this discussion toward generative AI, secure and ethical digital twins, metaverse-enabled collaboration, digital product passports, cyber-physical jobsites, and human-centered adoption.

Author contributions

IT: Supervision, Validation, Writing – review and editing. ZL: Writing – review and editing. SA: Conceptualization, Visualization, Writing – original draft, Writing – review and editing, Formal Analysis. VP: Writing – review and editing. ZC: Writing – review and editing. SK: Writing – review and editing. UB: Writing – review and editing.

Funding

The author(s) declared that financial support was not received for this work and/or its publication.

Conflict of interest

The author(s) declared that this work was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The authors IT, ZL, SA, VP, ZC, SK, UB declared that they were an editorial board member of Frontiers at the time of submission. This had no impact on the peer review process and the final decision.

Generative AI statement

The author(s) declared that generative AI was used in the creation of this manuscript. Figure 1 is sketched using ChatGPT.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Awodele, I. A., Ramabodu, M. S., Olatunde, N. N., Eze, E. C., Anugwo, I. C., Espinosa, Y. P., et al. (2026). The role of digital technologies in engineering procurement: a systematic literature review. *Front. Built Environ.* doi: 10.3389/fbuil.2026.1870384

Boje, C., Guerriero, A., Kubicki, S., and Rezgui, Y. (2020). Towards a semantic construction digital twin: directions for future research. *Autom. Constr.* 114, 103179. doi: 10.1016/j.autcon.2020.103179

Deng, M., Menassa, C. C., and Kamat, V. R. (2021). From BIM to digital twins: a systematic review of the evolution of intelligent building representations in the AEC-FM industry. *J. Inf. Technol. Constr.* 26, 58–83. doi: 10.36680/j.itcon.2021.005

Kritzinger, W., Karner, M., Traar, G., Henjes, J., and Sihn, W. (2018). Digital twin in manufacturing: a categorical literature review and classification. *IFAC-PapersOnLine* 51 (11), 1016–1022. doi: 10.1016/j.ifacol.2018.08.474

UNEP and GlobalABC (2025). *Global Status Report for Buildings and Construction 2024/25*. United Nations Environment Programme and Global Alliance for Buildings and Construction.

World Economic Forum (2024). *Top 10 Emerging Technologies of 2024*. Geneva: World Economic Forum.

frontiers | Frontiers in Built Environment

TYPE Opinion
PUBLISHED 05 December 2024
DOI 10.3389/fbuil.2024.1486423

Check for updates

OPEN ACCESS

EDITED BY
Salman Azhar,
Auburn University, United States

REVIEWED BY
Abiola Akanmu,
Virginia Tech, United States

*CORRESPONDENCE
Ehsan Noroozinejad Farsangi,
ehsan.noroozinejad@westernsydney.edu.au

RECEIVED 27 August 2024
ACCEPTED 22 November 2024
PUBLISHED 05 December 2024

CITATION
Noroozinejad Farsangi E, Shehata AO, Rashidi M, Ghassempour N and Mirjalili S (2024) Transitioning from BIM to Digital Twin to Metaverse.
Front. Built Environ. 10:1486423.
doi: 10.3389/fbuil.2024.1486423

COPYRIGHT
© 2024 Noroozinejad Farsangi, Shehata, Rashidi, Ghassempour and Mirjalili. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Transitioning from BIM to Digital Twin to Metaverse

Ehsan Noroozinejad Farsangi[1]*, Alaa O. Shehata[2], Maria Rashidi[1], Nargess Ghassempour[3] and Seyedali Mirjalili[4]

[1]Urban Transformations Research Centre (UTRC), Western Sydney University, Penrith, NSW, Australia, [2]Architectural Engineering and Urban Planning Department, Faculty of Engineering, Port Said University, Port Said, Egypt, [3]School of Business, Western Sydney University, Penrith, NSW, Australia, [4]Center for Artificial Intelligence Research and Optimisation, Torrens University Australia, Brisbane, QLD, Australia

KEYWORDS

BIM, Digital Twin, Metaverse, AI, IoT, AEC, AR, VR

1 From BIM to Digital Twin

In the Architecture, Engineering, and Construction (AEC) sector, Building Information Modelling (BIM) has been at the forefront of the digital revolution. BIM makes it easier to create detailed digital models that capture a building's functional and physical attributes. Throughout the project lifecycle, these models allow stakeholders to exchange insights and make better-informed decisions. But as company demands have increased, likewise has the demand for more advanced digital tools (Singh, 2019; Shahzad et al., 2022; Doherty, 2023; Azhar, 2011; Shehata et al., 2024).

Digital Twin is a real-time, dynamic digital clone of a system, process, or physical item. Digital Twins combine real-time data streams from sensors implanted in physical assets, in contrast to BIM, which primarily concentrates on static models, enabling continuous monitoring, simulation, and analysis. Digital twins have been used in a variety of ways, providing unprecedented opportunities to improve decision-making and operational effectiveness. For instance, DTs allow for real-time energy usage monitoring, maintenance requirement prediction, and space optimization in smart buildings. Facility managers may reduce downtime and increase the lifespan of building systems by using distributed testing (DT) to identify and prevent possible problems before they arise by modelling various scenarios (Zhang et al., 2020).

Digital Twins provides an entire perspective of assets in infrastructure management, enabling improved operating, maintenance, and planning methods. The quality and dependability of these digital replicas are enhanced by the integration of DTs with current BIM models, which makes them essential tools for contemporary asset management. Digital Twin adoption in the AEC sector is not without obstacles, despite its benefits. The primary hurdle is the lack of compatibility across various software platforms and systems. To be valid, DTs need to work well with other digital tools, IoT devices, and current BIM models—many of which run on different platforms.

Privacy and data security are also major issues. Since Digital Twins depend on real-time data, it is critical to preserve sensitive data. The industry now lacks standardization, which is made necessary by the complexity of digital systems growing (Stergiou et al., 2023).

2 The role of Metaverse in the future of built environment

The next frontier of digital innovation is the Metaverse, an idea that has gained a lot of traction recently. It is a 3D virtual world that combines artificial intelligence (AI), augmented

reality (AR), virtual reality (VR), and Internet of Things (IoT) technology to allow users to interact with digital items and settings in real-time. The Metaverse offers a new level of interaction, design, and management within the built environment of the AEC industry (Chang et al., 2024; Uddin et al., 2024).

The fields of facility management and construction represent some of the most exciting uses of the Metaverse in AEC. The Metaverse provides a completely immersive virtual environment for architects, engineers, and clients to work in all through the design process. This dramatically lowers the possibility of mistakes and rework by enabling real-time design reviews, virtual site inspections, and the flexibility to make changes as needed.

In order to provide managers with a risk-free virtual environment for performing inspections, monitoring systems, and training personnel, facility managers may use the Metaverse to construct virtual duplicates of actual structures. The Metaverse offers a complete platform for controlling a building's lifetime, from construction to deconstruction, by combining data from Digital Twins and BIM. The AEC sector stands to gain a great deal from Metaverse's integration, including more interaction, better stakeholder involvement, and more effective project delivery. Users may experience and engage with complicated systems in a controlled virtual environment thanks to the immersive nature of the Metaverse, which makes it an excellent tool for training and simulation.

On the other hand, there are a number of obstacles to Metaverse's broad adoption. A completely immersive Metaverse requires a significant technical infrastructure, which includes fast internet, cutting-edge VR/AR tools, and reliable data management systems. It is also essential to address concerns about privacy, data security, and the moral ramifications of developing and using virtual environments (Heruatmadja and Ramadhan, 2024).

3 Integrating BIM, Digital Twin, and the Metaverse

Although BIM, Digital Twins, and the Metaverse have different advantages, their integrated potential is most significant when used together. The AEC sector may attain a level of operational efficiency and innovation that was previously considered to be unachievable by combining the extensive modelling skills of BIM with the real-time data and simulation capability of Digital Twins and the immersive environments of the Metaverse. For example, a project may start with a BIM model for the structure's design and planning. A Digital Twin could be created while construction moves on to track and improve the building's performance. Lastly, the structure might be managed by the Metaverse once it has been constructed, offering a virtual platform for continuing training, upkeep, and upgrades.

4 Mainstream trends of relevant literature

Recently, there was a notable rise in production in the relevant literature about the evolution from BIM, DTs, to the Metaverse. The Scopus database search was used, searching about "BIM" AND "Digital Twin" AND "Metaverse" which restricted the results to final journal articles written in English and related to architectural design, turned up over 3,597 documents published in high-impact journals. This implies that while the subject has received a great deal of attention in the academic community. Figures 1, 2 generated using VOS viewer, illustrate the relationships between terms and concepts within a scientific or technological domain over time. The network visualization shows the development of BIM to Digital Twin to Metaverse in the construction industry. BIM is critical to managing digital representations of buildings, with connections to blockchain and information management. Digital Twins, on the other hand, enhance BIM capabilities by integrating real-time data from the physical world via sensors and the Internet of Things (IoT), enabling dynamic updates and real-time decisions. These interactive simulations provide insights for predictive maintenance, sustainability and performance optimization.

Also, distinguishing between Digital Twins, AR, VR and XR, the Metaverse emerges as the next major advancement in digital settings. In order to enable stakeholders to engage with models in real-time across borders and devices, Metaverse expands on Digital Twins by providing immersive, user-interactive settings. As a result of this change, there are potential wider uses outside design and construction, such as human-computer interaction, education, and sustainable development. In order to create immersive environments, smart cities, and the digital transformation of industries, it is imperative that we go beyond BIM to Digital Twins and Metaverse. Besides all of these objectives, there are some other aspects—like standardized procedures, data security and privacy, and ethical and social implications—remain largely unexplored.

There have been several attempts to discuss the transition from BIM to Digital Twin to Metaverse. The development of Digital Twins from BIM in the framework of the building life cycle had been systematically reviewed by Deng et al. (2021); this study attracted other researchers. The citation of this study until now reached 215 citations, with a total citation per year equal to 53.75. This study emphasizes the shift from BIM to Digital Twin to virtual representations of the built environment that are dynamic and real-time, providing real-time data synchronization, continuous monitoring, and operational insights. This technology might be developed further by the Metaverse, opening up interactive virtual places for stakeholder participation (Deng et al., 2021).

Another study discussed the adoption of BIM in the AEC sector, with an emphasis on project management and construction management. By integrating real-time data with IoT systems for dynamic monitoring, it establishes the foundation for Digital Twin applications (Chen et al., 2023). With the development of BIM, more complex digital environments are made possible, which opens the door to Digital Twin technologies and the Metaverse. Other researchers were drawn to this work. This study has received 112 citations, and 14.00 for an annual total citation (Olawumi et al., 2017). Additionally, Baghalzadeh Shishehgarkhaneh et al. (2022) highlighted the importance of BIM in the building industry, as its function in static digital representations. It also discusses how Digital Twins and the IoT are integrated to provide real-time data flow and link actual assets with virtual models. This study has also a high citation impact that reaches 96 citations, and 32.00 for an annual total citation (Baghalzadeh Shishehgarkhaneh et al., 2022).

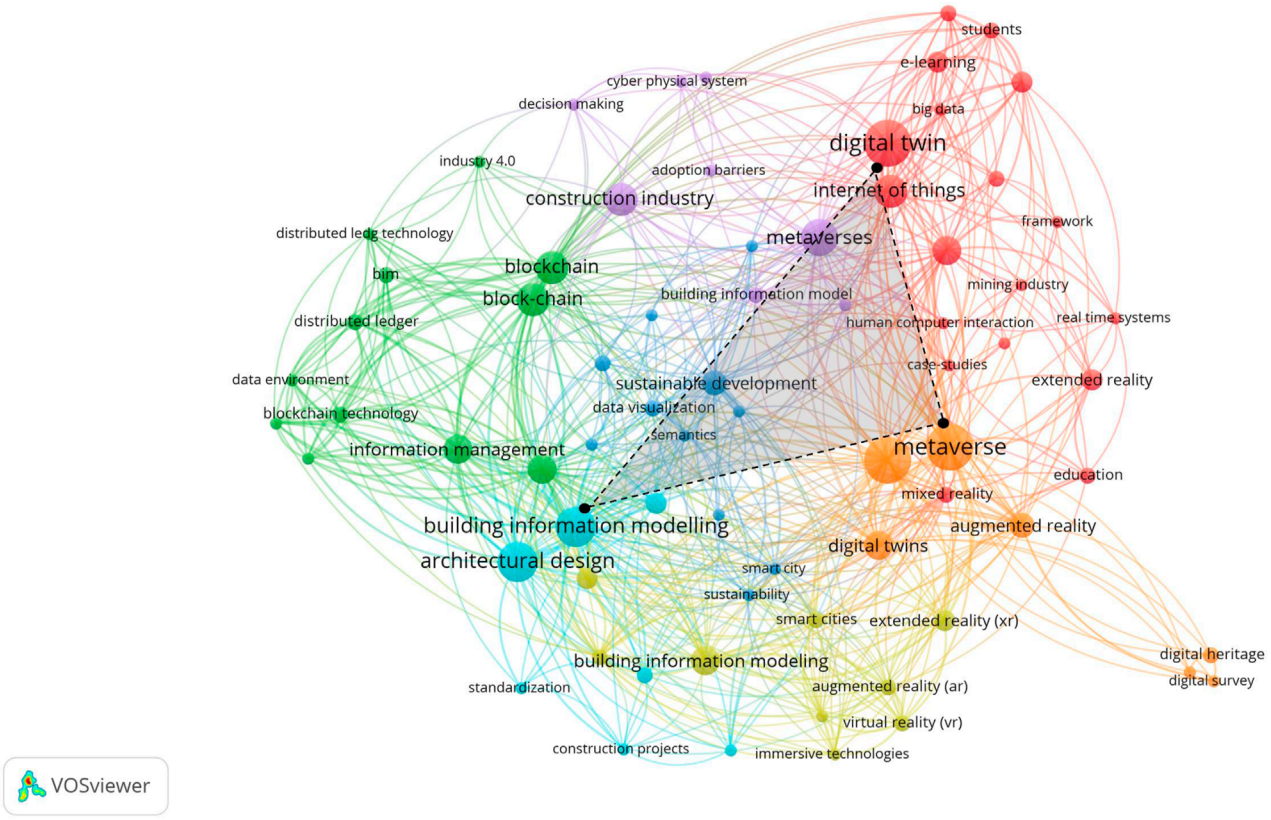


FIGURE 1
Co-occurrence keywords network. Source: the authors, by VOSviewer.

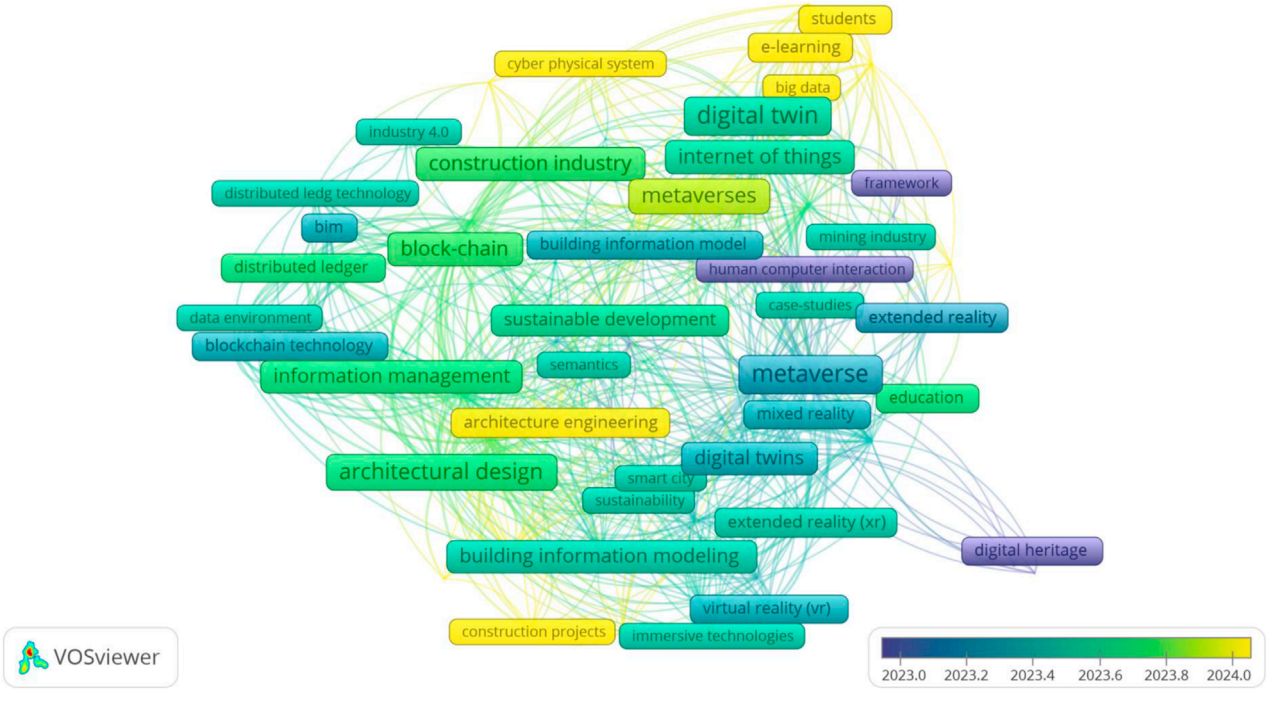


FIGURE 2
Trending topics and future research potentials. Source: the authors, by VOSviewer.

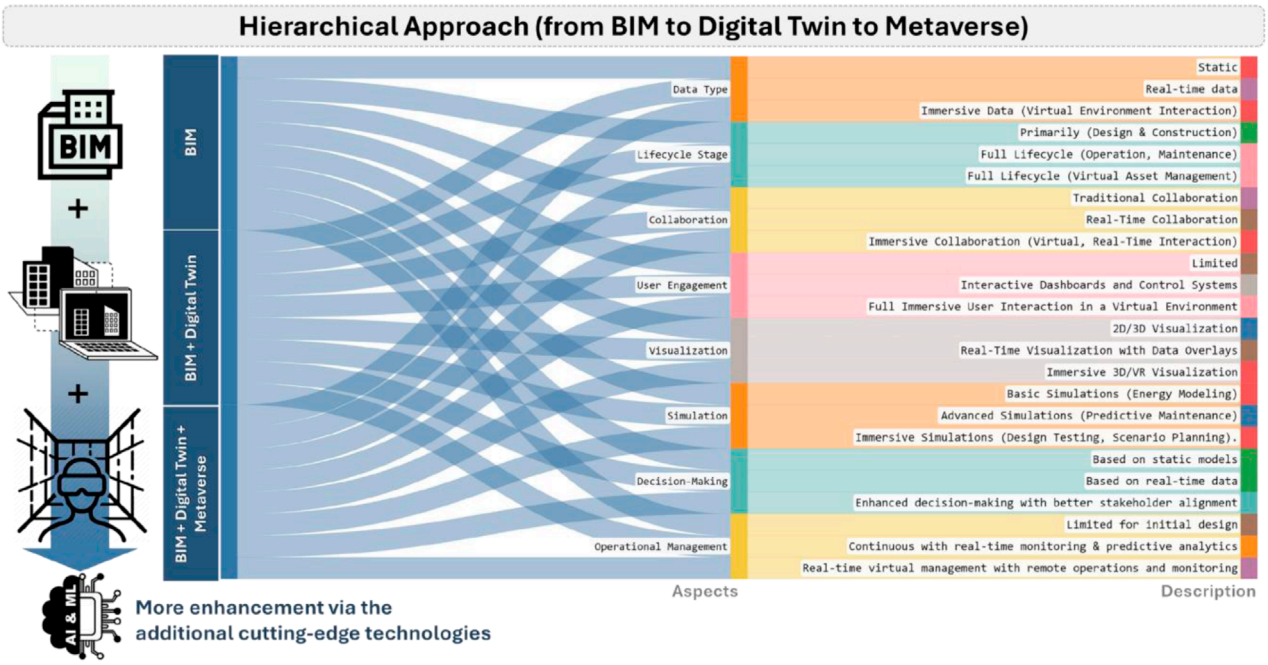


FIGURE 3
The hierarchical approach impacts several aspects of AEC. Source: the authors.

Managing the full building lifecycle was the main topic of discussion in the research of Su et al. (2023)-with 26 citations-, which included the evolution of Digital Twin technologies. Enhancing simulation, prediction, and control, it demonstrated the progression from BIM to Digital Twin. Also, mentioned the Metaverse as future research (Su et al., 2023). Furthermore, some studies extents to cover the evolution of transition the construction industry to the Metaverse based on BIM and Digital Twin as a foundation. Khan et al., (2021) demonstrated that the integration of BIM with immersive technologies such as VR, AR, and MR in the AEC sector through improving design, construction planning, and monitoring. Real-time synchronization between the digital and physical worlds is made possible by Digital Twins. Virtual worlds that are interactive are offered via the Metaverse (Khan et al., 2021).

This study has garnered attention from other researchers, accumulating 75 citations to date, with an average of 18.75 citations per year. Moreover, Markopoulos et al. (2024) elaborated the integration of Digital Twin and Metaverse into the building sector based on human-centric approach to BIM and sustainable development goals (SDGs) implementation. Mourtzis, (2023) explored the role of Digital Twin in the industrial Metaverse as a foundational technology for Industry 4.0 and 5.0, proposing Metaverse-enabled platforms to enhance interactions between operators and cyber-physical systems (Mourtzis, 2023).

In addition, some paper discussed the adoption of BIM, Digital Twin, and Metaverse, particularly in heritage sector by proposing a workflow that allows users to use virtual twins for historical building and heritage building modeling (H-BIM) (Cruz Franco et al., 2022; Guo et al., 2024). Also, Lv et al. (2022) enhanced the idea of Digital Twins to extent from the building scale to the urban scale "city", reviewing their impact and the role of digitalization in urban management and development, based on the historical progression of digital cities (DCs), explores the potential of the Metaverse in enhancing the interaction between the physical and digital worlds, ultimately providing valuable insights for improving urban governance and operational efficiency (Lv et al., 2022). Also, Domenico, (2023) discussed the urban scale to represent a virtual environment where the interaction between human and digital models occurs in architectural and urban settings (Domenico, 2023). Overall, BIM, Digital Twin, and the Metaverse all offer some important combinatory benefits they can enable in construction industry efficiency, productivity, decision-making, and collaboration. Most importantly, they become the guiding light toward more creative and sustainable approaches within a connected digital framework.

From the perspective of the hierarchical approach (from BIM to Digital Twin to Metaverse), as shown in Figure 1, BIM can be considered the starting point for construction projects. Also, the integration with BIM and Digital Twins can provide data, analytics, simulations, visualisation, automation, and predictions. Additionally, Metaverse can provide several key benefits to the AEC industry, offering enhanced results in aspects of schedule, cost, quality, and safety. Metaverse enables the efficient production, supervision, and sharing of data among stakeholders (Hadavi and Alizadehsalehi, 2024; Alizadehsalehi and Yitmen, 2021).

Looking forward, a number of developments and trends are probably going to influence how BIM, Digital Twins, and the Metaverse develop in the AEC sector. The capabilities of

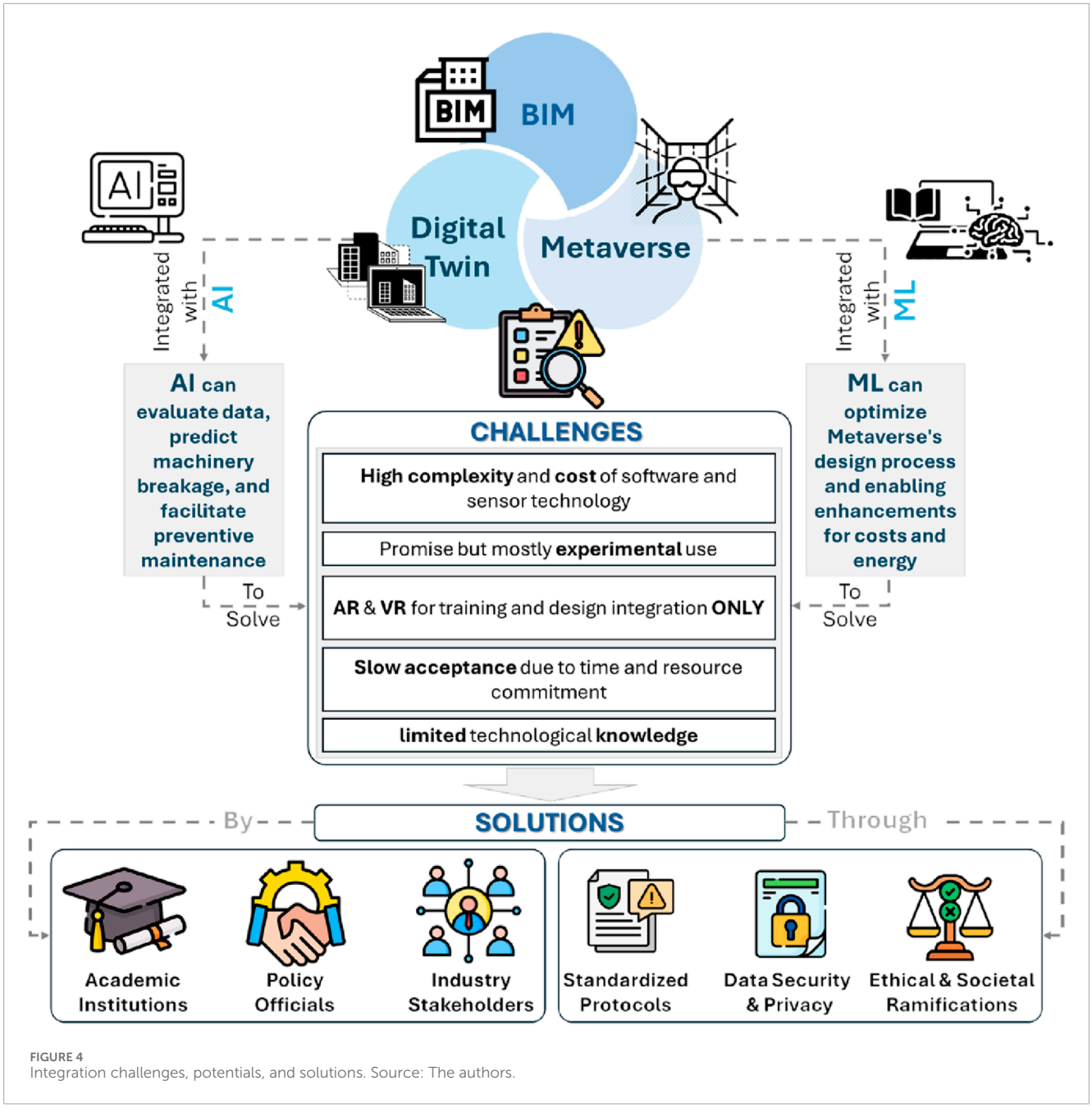


FIGURE 4
Integration challenges, potentials, and solutions. Source: The authors.

these technologies will be significantly enhanced by artificial intelligence (AI) and machine learning (ML), especially in areas like automation, decision assistance, and predictive analytics. Enhancing digital system intelligence and interoperability will also need the inclusion of semantic technologies. These technologies will aid in bridging the gap between the physical and digital worlds by allowing more intuitive interactions between people and machines, increasing the built environment's responsiveness and adaptability to changing demands. In addition to AI, other disruptive technologies, including IoT, AR, VR, blockchain, and generative design, have the potential to speed up the transformation of AEC industry, thereby moving it to unknown heights of efficacy (efficiency), sustainability, and innovation in all building, designing construction and operating stages. McKinsey and Company estimates that by these technologies, as well as improvement of productivity the global construction industry could unlock up to $1.6 trillion annually. This has the potential to redefine both our lifestyles and architecture (Filipe et al., 2017).

5 Discussion

The AEC sector is seeing a turning point with the integration of BIM, Digital Twins, and the Metaverse; nevertheless, the shift is not without its challenges, as presented in Figure 2.

Industry is now struggling to incorporate new technologies into processes that are already in place. While BIM has become commonplace, Digital Twin adoption is still in its early stages. This is sometimes due to the high price of the complex software required to run these systems and the necessary sensor technology.

Furthermore, although the Metaverse holds significant promise, its use in the AEC industry is still mostly experimental. While numerous businesses are experimenting with augmented and virtual realities (AR/VR) for training and design, a complete integration of these technologies into everyday operations is still a way off. This slow pace of acceptance may be attributed, in part, to the substantial time and resource commitment needed. Another reason is that these cutting-edge technologies come with a high learning curve, which may be a barrier for businesses with little technological know-how.

On the other hand, integrating AI and machine learning with Digital Twins and the Metaverse is one of the most promising areas of research. These innovations have the potential to significantly improve Digital Twins' predictive power, resulting in more precise simulations and improved decision-making. AI may be used, for example, to evaluate data from a Digital Twin and forecast when a piece of machinery is likely to break, enabling preventive maintenance that could save money and effort.

Algorithms for machine learning might also be used to optimize Metaverse's design process. These algorithms might provide recommendations for design enhancements that save costs or increase energy efficiency by examining data from earlier projects. In this sense, the Metaverse has the potential to develop into a potent platform for innovation and optimization rather than only a visualization tool.

There are a lot of chances for research and development as the AEC sector moves towards being entirely digital. The creation of standardized protocols for data interchange across Digital Twins, BIM, and the Metaverse is one important topic (Figure 3). At the moment, a significant obstacle to the integration of these technologies is the incompatibility of various software platforms. Creating a global standard for data sharing will make integrating these technologies more effortless and more accessible to businesses of all sizes.

Data security and privacy constitute a key study field as well. Sensitive data collection and storage will rise at an exponential rate as the AEC sector becomes more digitally integrated. This data comprises real-time information from sensors embedded in buildings as well as architectural drawings and blueprints. Adopting Digital Twin's and Metaverse's success will depend on ensuring this data is safe from cyberattacks.

However, further study is required to fully understand the ethical and societal ramifications of the Metaverse. We need to think about how developing more immersive virtual worlds can affect society as a whole and the workforce in particular. What effect, for instance, would the growing use of virtual reality in training have on employment opportunities in the AEC sector? What privacy problems result from real-time monitoring of every facet of a building's operation? As we transition to a future that is more digital, these are crucial concerns that must be answered.

Adopting these technologies will need cooperation from academic institutions, industrial players, and policy officials. To create best practices for integrating BIM, Digital Twins, and the Metaverse into practical projects, industry stakeholders must collaborate. Academic institutions may help by teaching the next generation of engineers and architects in the use of these technologies and by doing research that tackles the technical difficulties.

There is also a function for policymakers. They may hasten the adoption of these technologies by contributing funds for research and development. Additionally, they have the authority to create laws that guarantee the ethical and secure deployment of these technologies while also safeguarding people's security and privacy.

Adoption may also be significantly aided by the creation of regulations that support the use of BIM, Digital Twins, and the Metaverse in public infrastructure projects. Governments worldwide are realizing how vital digital technologies are to enhancing the sustainability and effectiveness of public infrastructure. They may assist in developing a market for these technologies and promoting their acceptance by the private sector by requiring their usage in public projects (Figure 4).

6 Conclusion

The AEC industry's approach to the design, construction, and administration of the built environment has undergone a substantial change, as seen by the transition from BIM to Digital Twin to Metaverse. When merged, these technologies provide unique chances to improve productivity, teamwork, and sustainability throughout the building lifecycle.

Stakeholders must embrace these technologies as the AEC sector develops and make the required investments in infrastructure, R&D, and training to realize these technologies' full potential. By doing this, the construction sector can provide a physical environment that is more inventive, durable, and adaptable to future problems.

Author contributions

ENF: Conceptualization, Data curation, Formal Analysis, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing–original draft, Writing–review and editing. AS: Conceptualization, Methodology, Visualization, Writing–original draft, Writing–review and editing. MR: Conceptualization, Methodology, Writing–original draft, Writing–review and editing. NG: Conceptualization, Methodology, Writing–original draft, Writing–review and editing. SM: Conceptualization, Methodology, Writing–original draft, Writing–review and editing.

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Alizadehsalehi, S., and Yitmen, I. (2021). Digital twin-based progress monitoring management model through reality capture to extended reality technologies (DRX). *Smart Sustain. Built Environ.* 12 (1), 200–236. doi:10.1108/sasbe-01-2021-0016

Azhar, S. (2011). Building information modeling (BIM): trends, benefits, risks, and challenges for the AEC industry. *Leadersh. Manag. Eng.* 11 (3), 241–252. doi:10.1061/(asce)lm.1943-5630.0000127

Baghalzadeh Shishehgarkhaneh, M., Keivani, A., Moehler, R. C., Jelodari, N., and Roshdi Laleh, S. (2022). Internet of Things (IoT), building information modeling (BIM), and digital twin (DT) in construction industry: a review, bibliometric, and network analysis. *Buildings* 12, 1503. doi:10.3390/buildings12101503

Chang, S., Chen, J., and Park, J. (2024). "Review of Metaverse technologies to broaden accessibility in civil and construction engineering education," in *Construction research congress*, 304–314.

Chen, Y., Wang, X., Liu, Z., Cui, J., Osmani, M., and Demian, P. (2023). Exploring building information modeling (BIM) and internet of Things (IoT) integration for sustainable building. *Buildings* 13 (2), 288. doi:10.3390/buildings13020288

Cruz Franco, P. A., Rueda Márquez de la Plata, A., and Gómez Bernal, E. (2022). Protocols for the graphic and constructive diffusion of digital twins of the architectural heritage that guarantee universal accessibility through AR and VR. *Appl. Sci. Switz.* 12 (17), 8785. doi:10.3390/app12178785

Deng, M., Menassa, C. C., and Kamat, V. R. (2021). From BIM to digital twins: a systematic review of the evolution of intelligent building representations in the AEC-FM industry. *J. Inf. Technol. Constr.* 26, 58–83. doi:10.36680/J.ITCON.2021.005

Doherty, P. (2023). *Unlocking the Metaverse: a strategic guide for the future of the built environment*. John Wiley and Sons.

Domenico, D. 'U. (2023). Multiscalar digital twin. Step representation towards urban multiverse in *Transizioni/transitions. FrancoAngeli srl*. doi:10.3280/oa-1016-c415

Filipe, B., Lola, W., and Jan, M. (2017). *Reinventing construction through a productivity revolution*. Available at: https://www.mckinsey.com/capabilities/operations/our-insights/reinventing-construction-through-a-productivity-revolution Accessed 1 March, 2024.

Guo, Y., Liu, L., Huang, W., Shen, M., Yi, X., Zhang, J., et al. (2024). Extending X-reality technologies to digital twin in cultural heritage risk management: a comparative evaluation from the perspective of situation awareness. *Herit. Sci.* 12 (1), 245. doi:10.1186/s40494-024-01363-6

Hadavi, A., and Alizadehsalehi, S. (2024). From BIM to Metaverse for AEC industry. *Automation Constr.* 160, 105248. doi:10.1016/j.autcon.2023.105248

Heruatmadja, C. H., and Ramadhan, A. (2024). Drivers and barriers of intention to work within Metaverse environment. *Multimedia Tools Appl.* 83 (17), 50955–50979. doi:10.1007/s11042-023-17567-0

Khan, A., Sepasgozar, S., Liu, T., and Yu, R. (2021). Integration of bim and immersive technologies for aec: a scientometrics-sot analysis and critical content review. *Buildings* 11, 126. doi:10.3390/buildings11030126

Lv, Z., Shang, W. L., and Guizani, M. (2022). Impact of digital twins and Metaverse on cities: history, current situation, and application perspectives. *Appl. Sci. Switz.* 12 (24), 12820. doi:10.3390/app122412820

Mourtzis, D. (2023). Digital twin inception in the Era of industrial Metaverse. *Front. Manuf. Technol.* 3. doi:10.3389/fmtec.2023.1155735

Olawumi, T. O., Chan, D. W. M., and Wong, J. K. W. (2017). Evolution in the intellectual structure of BIM research: a bibliometric analysis. *J. Civ. Eng. Manag.* 23 (8), 1060–1081. doi:10.3846/13923730.2017.1374301

Shahzad, M., Shafiq, M. T., Douglas, D., and Kassem, M. (2022). Digital twins in built environments: an investigation of the characteristics, applications, and challenges. *Buildings* 12 (2), 120. doi:10.3390/buildings12020120

Shehata, A. O., Megahed, N. A., Hassan, A. M., and Shahda, M. M. (2024). Holistic SWOT based matrix of BIM adoption in heritage green retrofitting processes. *Archit. Eng. Des. Manag.* 20 (3), 694–718. doi:10.1080/17452007.2024.2337933

Singh, V. (2019). Digitalization, BIM ecosystem, and the future of built environment: how widely are we exploring the different possibilities? *Eng. Constr. Archit. Manag.* doi:10.1108/ecam-01-2018-0004

Stergiou, C. L., Bompoli, E., and Psannis, K. E. (2023). Security and privacy issues in IoT-based big data cloud systems in a digital twin scenario. *Appl. Sci.* 13 (2), 758. doi:10.3390/app13020758

Su, S., Zhong, R. Y., Jiang, Y., Song, J., Fu, Y., and Cao, H. (2023). Digital twin and its potential applications in construction industry: state-of-art review and a conceptual framework. *Adv. Eng. Inf.* 57, 102030. doi:10.1016/j.aei.2023.102030

Uddin, M., Obaidat, M., Manickam, S., Laghari, S. U. A., Dandoush, A., Ullah, H., et al. (2024). *Exploring the convergence of Metaverse, Blockchain, and AI: a comprehensive survey of enabling technologies, applications, challenges, and future directions*. Wiley Interdisciplinary Reviews.

Zhang, J., Zhao, L., Ren, G., Li, H., and Li, X. (2020). Special issue "digital twin technology in the AEC industry". *Adv. Civ. Eng.* 2020 (1), 8842113. doi:10.1155/2020/8842113

frontiers | Frontiers in Built Environment

TYPE Review
PUBLISHED 13 March 2025
DOI 10.3389/fbuil.2025.1523464

Check for updates

OPEN ACCESS

EDITED BY
Vagelis Plevris,
Qatar University, Qatar

REVIEWED BY
Dimitrios Kraniotis,
Oslo Metropolitan University, Norway
Themistoklis Tsalkatidis,
Norwegian University of Life Sciences, Norway

*CORRESPONDENCE
Mahenthiran Sathiyamoorthy,
mahenthiran.s@vit.ac.in

RECEIVED 06 November 2024
ACCEPTED 18 February 2025
PUBLISHED 13 March 2025

CITATION
Venkateswarlu N and Sathiyamoorthy M (2025) Sustainable innovations in digital twin technology: a systematic review about energy efficiency and indoor environment quality in built environment.
Front. Built Environ. 11:1523464.
doi: 10.3389/fbuil.2025.1523464

COPYRIGHT
© 2025 Venkateswarlu and Sathiyamoorthy. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Sustainable innovations in digital twin technology: a systematic review about energy efficiency and indoor environment quality in built environment

Natakarani Venkateswarlu and Mahenthiran Sathiyamoorthy*

School of Civil Engineering, Vellore Institute of Technology, Vellore, Tamil Nadu, India

In the contemporary digital age, the built environment undergoes significant changes because of technological innovations that improve building management, optimize building efficiency, and enhance overall productivity. Digital Twin technology has emerged as an indispensable tool for enhancing indoor environmental quality and optimizing energy efficiency in existing buildings. This demonstrates its similarity to several SDGs, where digital twin technology is key to achieving many of them, especially those relevant to our research: 7. Affordable and clean energy; 3. Good health and wellbeing are the primary outcomes of our study; 9. Industry innovation and infrastructure are the focus of our methodology; and 11. Sustainable cities and communication, to which our research contributes. However, some challenges require further consideration. First, to assess the methods and tools used to monitor and represent environmental parameters. Second, to review previous studies on Digital Twin technology in the context of energy efficiency and indoor environmental quality. This study systematically examined 261 academic articles to address these challenges, identifying 17 relevant publications investigating Digital Twin for enhancing energy efficiency and indoor environmental quality in buildings. The research emphasizes Building Information Modeling, Internet of Things, and Big Data, which collectively improve the monitoring and management of physical assets through real-time data replication. Our research illustrates the need for a multidisciplinary framework to rigorously analyze Digital Twin applications, as a comprehensive understanding of the consequences of this technology requires the integration of different fields. The review emphasizes the confined application of sensors for monitoring the environment, the importance of residents subjective impressions, and the need for further comparative studies on energy use estimation methods. For future investigation, enhanced international collaboration is imperative to improve the scholarly exploration of Digital Twin related to this field. Finally, the built environment can benefit significantly from implementing Digital Twin technology. However, the challenges must be addressed before technology can achieve its full potential for creating sustainable and energy-efficient buildings.

KEYWORDS

digital twin, bim, scientometric review, energy efficiency, indoor environmental quality, buildings

1 Introduction

The built environment continues to evolve in the digital age, with novel concepts and technological advances accomplishing cities, homes, and construction more intelligently for the benefit of their occupants. Digital technologies have been exploited as tools to enhance projects and built environment productivity and efficiency (Manzoor et al., 2021). Table 1 represents an overview of a few applications for Digital Twin Technology. A Variety of heterogeneous digital technologies exist, encompassing including Building Information Modeling (BIM), Internet of Things (IoT), Big Data (BD), Artificial Intelligence (AI), Three-dimensional (3D) printing, Blockchain, and Digital Twin (DT) (Asif et al., 2024). The "DT" designation was initially made available to the public in the year 2010 (Conroy, 2010). DT is an emerging technology that can digitally replicate physical objects. Use DT as a foundational framework to combine different technology systems. Use real-time data obtained from sensors attached to an object to reproduce its behavior, facilitate visualization, perform monitoring, and monitor maintenance process (Author Anonymous, 2024; Madni et al., 2019). DT combines several approaches, such as IoT and BIM, to create digital entities and transfer data from a physical to a digital model. The concept of DT, which relates to optimizing energy efficiency and enhancing indoor environmental quality in buildings, is Illustrated in Figure 1. DT is still in its early stages regarding its implementation in the built environment. However, it has already been applied in many industries, including manufacturing, transportation, agriculture, aviation, and the automobile industry (Deng et al., 2021). DT has been utilized in the construction industry's particular design phase and functions as a virtual model with the aid of BIM because it has considered geometric and contextual data (Kaewunruen et al., 2019; Lin and Cheung, 2020; Lu et al., 2020a). It was also implemented during the building phase to instruct workers on on-site logistics, structural integrity, and musculoskeletal injury prevention (Akanmu et al., 2020; Angjeliu et al., 2020; Greif et al., 2020). The applications of DT improve the climate change impact parameters such as building carbon footprint, building asset performance, fault detection in systems, building occupant comfort, energy efficiency, and CO_2 monitoring (Arsiwala et al., 2023; Cespedes-Cubides and Jradi, 2024; Hosamo et al., 2022; Hosamo et al., 2023b; Jafari et al., 2020). However, energy consumption and indoor environmental quality in the built environment are important factors to consider to improve performance and productivity. People's attitudes toward these parameters vary based on activity, mood, and environmental and personal conditions.

DT experiments are created using various methods and algorithms. DT is already used for thermal comfort monitoring, visualization, tracking, energy management prediction, and optimization for existing buildings (Arowoiya et al., 2024). The scientometric study was based on a review article considering energy efficiency and thermal comfort as factors (Arowoiya et al., 2024). This academic review article reassesses the thermal comfort parameter by the basic principles of indoor environmental quality, which includes a comprehensive examination of all dimensions of indoor environmental quality. By implementing air purification systems, proper ventilation techniques, and an effective humidification process, this assessment facilitates the identification of problems and the formulation of remedial strategies.

Incorporating DT into architectural frameworks increases the energy efficiency and quality of the indoor environment and contributes significantly to achieving various Sustainability Development Goals (SDGs), thus promoting a sustainable future for urban ecosystems. SDG 7: Affordable and clean energy is given top priority. Adopting DT can meaningfully increase the energy efficiency of buildings, resulting in lower energy consumption and related costs. Another important goal is SDG 11: Sustainable Cities and Communities. Using DT can significantly enhance building efficiency, thereby facilitating the development of urban communities that prioritize inclusion, safety, resilience, and sustainability. In addition, reflecting the value of SDG 3: Good health and wellbeing, improved air quality, and amenities in buildings can significantly contribute to the overall wellbeing of residents, representing a key dimension of this goal. Finally, SDG 9: Industry, Innovation, and Infrastructure are emphasized through the integration of advanced technologies such as DT systems in the management of buildings.

The following strategies/methods will be used to achieve these objectives: i) Analyzing the tools and techniques used for parameter monitoring and visualization. ii) Reviewing previous research on DT for energy efficiency and indoor environmental quality, this study provides an iii) comprehensive analysis of the various approaches used for future research.

2 Literature review

2.1 Digital twin

Michael Grieve proposed a DT in 2000 in a course presentation on Product Life Management (PLM). The term DT was coined in 2003 when NASA's technology roadmaps provided the first description of its application, stating that it was used to simulate space conditions and conduct tests in preparation for flight (Tuegel et al., 2011). The concept of DT initially emerged in the aerospace industry and later extended to the manufacturing industry around 2012 (Sharma et al., 2022). Based on (Grieves and Vickers, 2016) DT has three types: DT Instance (DTI), DT Prototype (DTP), and DT Environment (DTE). DTI is a specific physical product that remains linked to it throughout its life; DTP is a prototypical physical artifact that produces a physical model that mirrors the digital model. DT

Abbreviations: DT, Digital Twin; IEQ, Indoor environmental quality; IAQ, Indoor air quality; AR/VR, Augmented virtuality/Virtual Reality; IoT, Internet of Things; BIM, Building information modeling; AI, Artificial Intelligence; ML, Machine learning; PLM, Product life management; GIS, Geographical information system; PRISMA, Preferred reporting items for systematic reviews and meta-analyses; NZEB, Net zero energy buildings; HVAC, Heating, Ventilation, and Air conditioning; ANN, Artificial Neural Networks; LiDAR, Light detection and ranging; IFC, Industry Foundation classes; Lod, Level of development; RIS, Research Information systems; CPS, Cyber-physical systems; BEM, Building energy management; VBM_S, Virtual modeling models; MCTS, Monte Carlo Tree Search; MPC, Model predictive control; BDT, Building Digital Twin; SBS, Sick building syndrome; YOLO, V4 You only look me; O & M, Operation & Maintenance.

TABLE 1 An overview of a few applications for Digital Twin.

Categories	Applications	References
Design and construction phase	Industry 4.0, infrastructure, net zero energy building, Building Digital twinning, visual analytics	Kaewunruen et al. (2019), Desogus et al. (2023), Liu et al. (2023), Moshood et al. (2024), Pregnolato et al. (2022), Tao et al. (2022), Ye et al., 2024; Yoon (2023)
Built environment	Buildings (Simulation, fault detection), carbon emission, predicting monitoring of CO_2 equivalent, cyber-physical systems	Hosamo et al. (2022), Arsiwala et al. (2023), Hadjidemetriou et al. (2023), Koo and Yoon (2024), Lydon et al., 2019; Yoon (2024b)
Operation and Maintenance	Anomaly detection, built asset monitoring Energy consumption and management Footprint, Asset performance	Lu et al. (2020a), Jafari et al. (2020), Agouzoul et al. (2021), 2023; Ni et al. (2023)
Real-time monitoring	Real-time analysis using sensors	Hu and Assaad (2024a), Hadjidemetriou et al. (2023), Es-haghi et al. (2024)

is operated by DTE, which offers a platform for managing and interacting with it.

2.2 Applications of digital twin

DT technology has demonstrated significant importance in various domains, including but not limited to manufacturing, construction, energy, infrastructure, healthcare, and transportation. These represent some main areas where DT technology is applied, as illustrated in Figure 2. Importance of autonomy and DT for future production and it emphasizes the necessity of DT technology for integration, accurate model building, and simulations, which play a crucial role in streamlining manufacturing procedures and preparing for unexpected events (Kritzinger et al., 2018; Rosen et al., 2015). DT in construction involves bidirectional coordination and real-time updates, which improve synchronization between virtual and physical worlds (Madubuike et al., 2022). Building operations and life cycle management may be easily controlled, monitored, and optimized because DT can create virtual models connected to physical assets. Figure 3 shows several topic areas where DT technology is used based on the Scopus database.

DT interactive virtual replicas of structures or infrastructure projects used in the construction trade. Advanced simulation

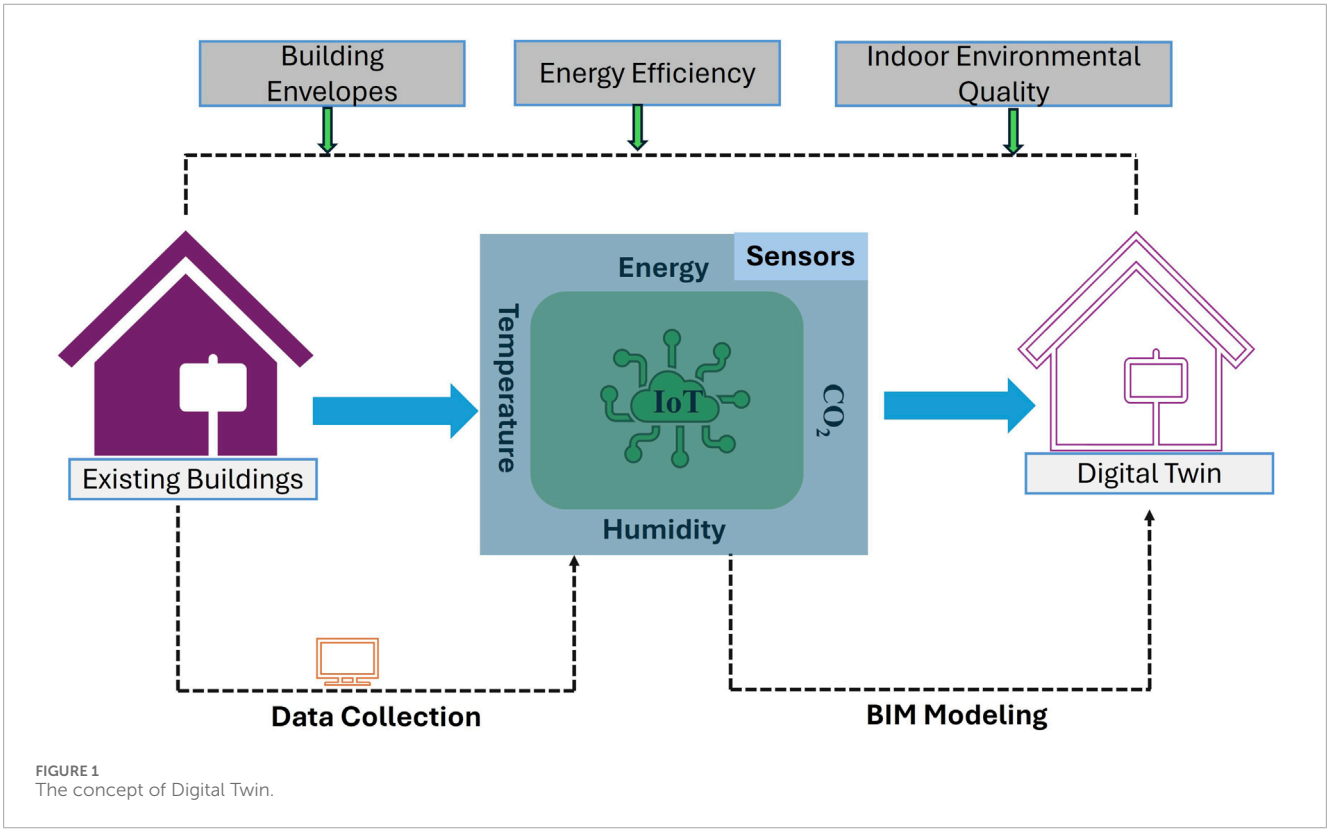


FIGURE 1
The concept of Digital Twin.

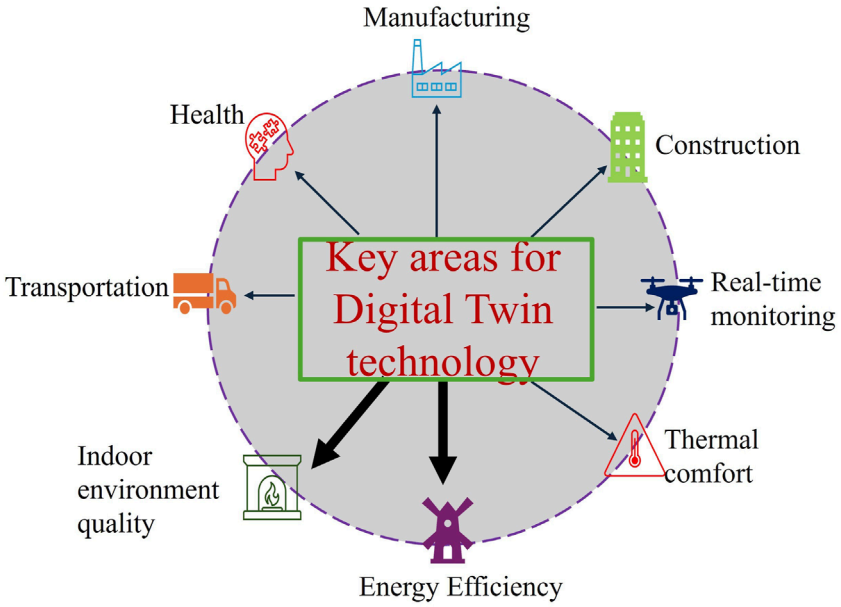


FIGURE 2
A few key areas for Digital Twin technology.

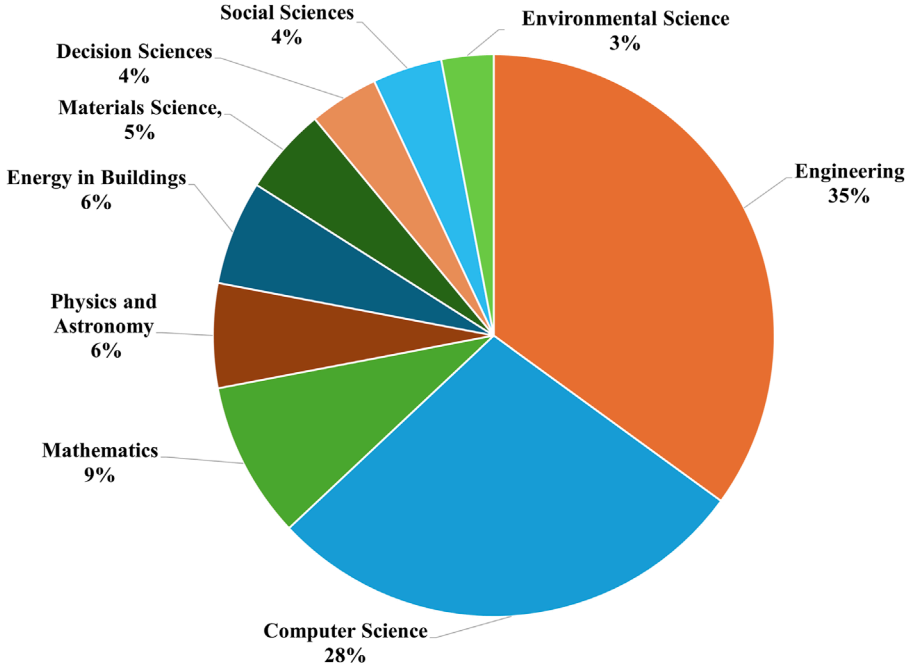


FIGURE 3
Digital Twin applications.

models are also used to construct these virtual replicas, which incorporate real-time data from several sources, including sensors, BIM models, and other IoT devices. Consequently, throughout every project phase, DT enhances decision-making, optimizes performance, and offers insights into the construction process. Development technologies such as BD, Augmented Reality (AR), and Geographic Information Systems (GIS) in the construction industry and discussed various technologies used in the manufacturing industry such as product design, simulation, product forecasting, fault diagnosis, decision-making, predictive maintenance, scheduling, and monitoring (Abanda et al., 2024; Saback et al., 2024) (Table 1).

2.3 Digital twin in the existing buildings

Energy efficiency in existing buildings to improve building energy efficiency to enhance digital technologies such as AI for predictive controls, dynamic BIM for monitoring, and DT for real-time visualization. These technologies can increase energy efficiency by up to 79% while also dramatically lowering expenses and energy usage (Zhou and Liu, 2024). A DT-based framework that evaluates university classrooms energy-saving lighting techniques by integrating occupant behavior, building design, and operating schedules. The DT model allows for quantifying prospective energy savings from methods such as better operation, schedules, and enhanced light source efficiency through the simulation of various scenarios (Seo and Yun, 2022). Simulating actual buildings and utilizing occupant behavior to control lighting and temperature using DT to optimize building energy consumption. This process eventually enables informed decision-making about energy management (Cespedes-Cubides and Jradi, 2024). With the simulation scenario, DT also aids in risk assessment, well-informed decision-making, and accident avoidance. Building system intelligence is further enhanced by integration with IoT technologies (Ghansah and Lu, 2024). DT was used to implement Net Zero Energy Buildings (NZEB) in existing buildings, and a feasibility investigation is ongoing emphasizing renewable technology, energy efficiency, and cost analysis. It accentuates the necessity of precise definitions and rules for NZEBs. It emphasizes the utmost importance of using renewable energy alternatives to achieve sustainability goals while addressing the challenges associated with existing building infrastructure. With an emphasis on improving building envelopes for energy efficiency (Kaewunruen et al., 2019). DT model design using ANN to predict energy consumption in residential buildings in Lebanon. It emphasizes the significance of DT in enhancing the building design processes, overturning a framework for architects and engineers, and improving energy performance in the context of climate change (El-Gohary et al., 2023).

DT technology for indoor environmental quality of existing buildings. This IEQ has four main parameters: thermal comfort, indoor air quality, lighting quality, acoustic comfort, etc. This review focuses on thermal comfort and indoor air quality because many researchers have used DT technology to find thermal comfort and indoor air quality among the remaining parameters. DT framework enabled by BIM that integrates indoor positioning technologies, LiDAR, IoT sensors, and autonomous robotics for real-time indoor environment monitoring (Hu and Assaad, 2024a). Point cloud datasets to address the problem of connectedness identification in building geometry. It presents a surface topology graph to depict interactions between surfaces and suggests a deep geometric neural network architecture for graph reconstruction. Improving DT capabilities for building operation and maintenance (Drobnyi et al., 2024). A methodological way to produce building information models using CAD drawings and photos. It comprises three models: IFC BIM production, building information integration, and structural geometry extraction. It also achieves a Level of Development (LOD) 300 and strongly emphasizes data processing efficiency and cost-effectiveness (Lu et al., 2020b). DT enhances building indoor environmental quality (IEQ) monitoring through three phases: real-time monitoring, visualization, and data integration. Real-time monitoring enables ongoing evaluation of indoor environmental parameters such as thermal comfort and indoor air quality, ensuring a healthy environment for occupants. Visualization makes environmental conditions easily observable, assisting in facility management decisions. For comprehensive environmental assessments, data from IoT sensors and robotics (Hu and Assaad, 2024a). In the context of the COVID-19 pandemic, DT has significant potential for establishing and preserving hygienic indoor settings (Cai et al., 2023).

3 Research method

This review employs scientometric analysis to comprehend the current use of DT in energy efficiency and indoor environment quality in existing buildings. Science mapping analyzes and visualizes scientific areas conceptual frameworks in a substantial body of literature (Cobo et al., 2011; Suleny Bojorquez-Roque et al., 2024). For 10 years (2014–2024), specific keywords such as Digital Twin, energy efficiency, indoor environmental quality, and buildings were used to collect information from the Scopus database. After cleaning duplicates, 261 papers were imported into the Mendeley Reference Manager (MRM) tool and examined with VOS Viewer software. To understand the complex network relationships and to explain existing knowledge gaps in this domain of inquiry, the analysis overlays and visualizes publications in a country and year-wise, document sources, and most cited publications. Figure 4 provides a Prismatic representation of the review paper (Salihu et al., 2022).

3.1 Database and data collection

The primary repository of information for this literature review is the Scopus database, which encompasses many scholarly articles about DT in preexisting structures and their correlation with energy efficiency and indoor environmental quality. Considering the excess of bibliometric data available about alternative databases like the Web of Science and Google Scholar, the field of buildings has been notably underrepresented in academic studies (Meho and Rogers, 2008).

3.2 Strategy for selection of keywords

Using MRM software, the systematic review used the Scopus database to retrieve estimated bibliographic information on selected keywords specific to the research domain. Scopus has established a primary database for keyword selection. It contains many publications compared with the Web of Science and Google Scholar. Keywords coded in that Scopus database TITLE-ABS KEY (((“Digital Twin” OR “Virtual Twin” OR “Virtual twinning” OR “Digital Twinning” AND “Energy efficiency” AND “Buildings” OR “Houses”))) resulted in 254 papers, and similarly coded as TITLE-ABS KEY (((“Digital Twin” OR “Virtual Twins” OR “Virtual twinning” OR “Digital Twinning” AND “Indoor environment quality” AND “Buildings” OR “Houses”))) resulted in 27 papers after applying the exclusion and inclusion

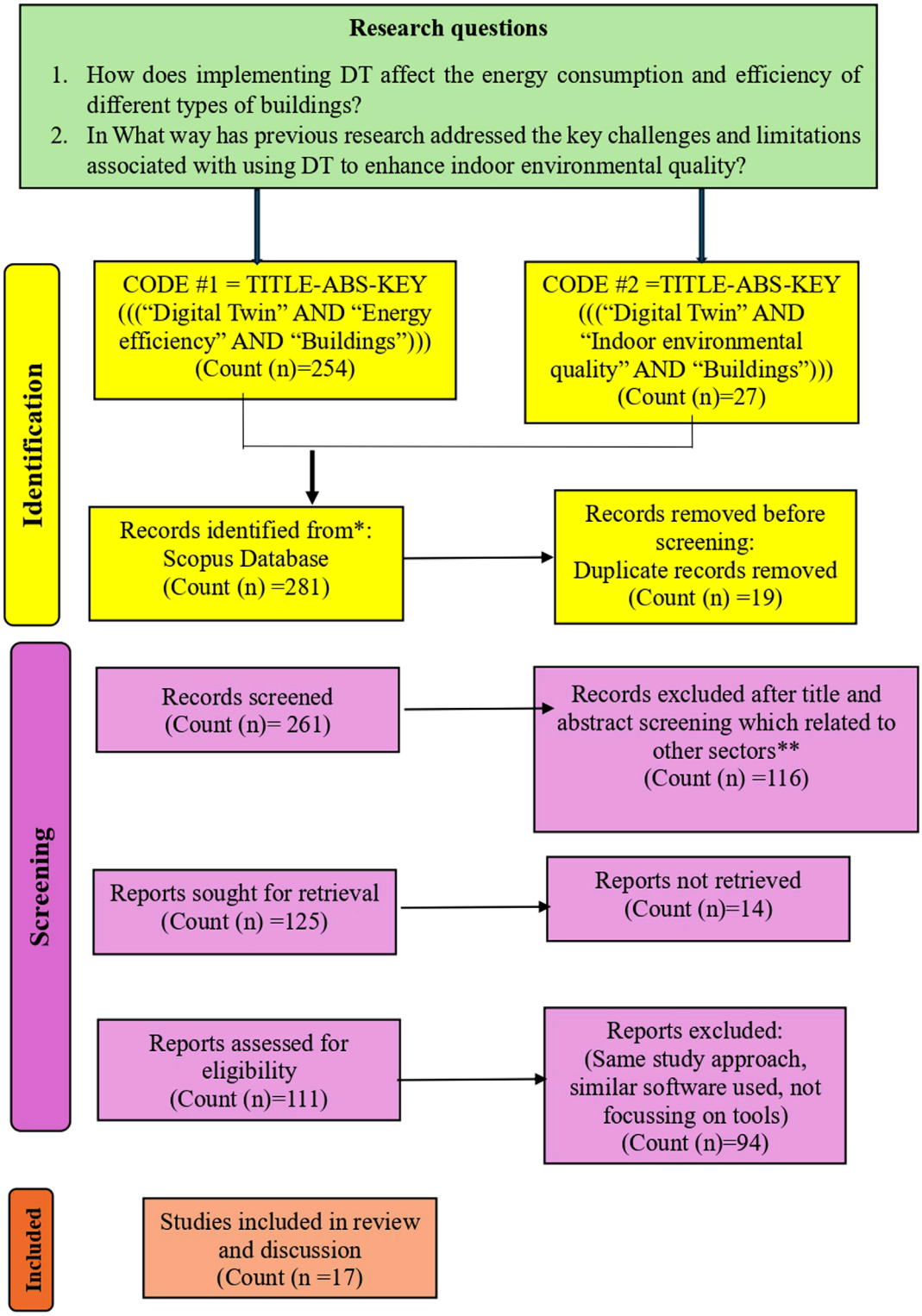


FIGURE 4
Prismatic representation of paper review.

criteria. Papers published in the last 10 years (from 2014 to 2024 in the English language in Engineering, computer science, Energy, Social sciences, Environmental sciences, etc.) met the inclusion and exclusion criteria. When imported to reference managers, 281 publications were discovered in the database.

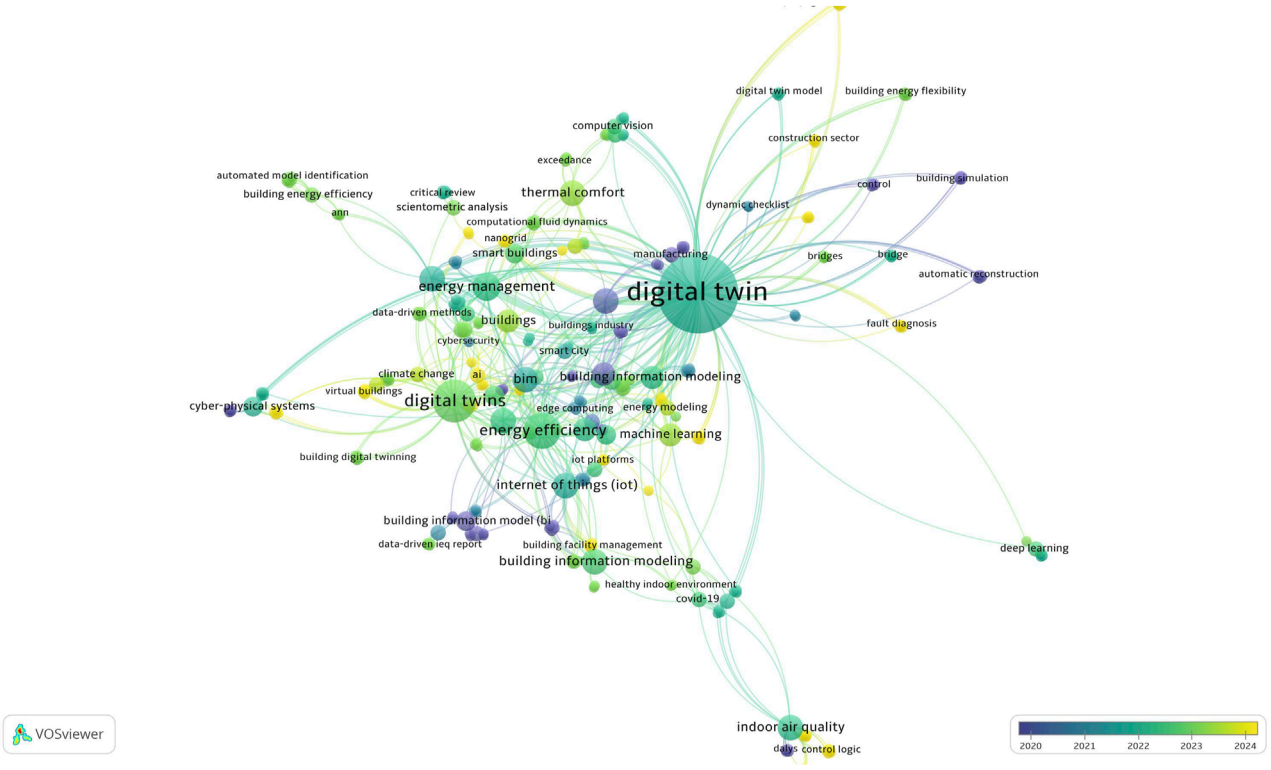


FIGURE 5
Overlay visualization of co-occurrence of keywords.

3.3 Selection of software tools for analysis and review

The scientometric evaluation uses the Scopus database and VOS Viewer software to investigate patterns and trends inherent in research activity. It displays bibliometric maps emphasizing visual aids to evaluate co-authorship and citation networks efficiently (Cobo et al., 2011). The bibliographic information underwent a meticulous review process, resulting in the elimination of duplicate entries. Subsequently, the ensuing articles were utilized to construct a comprehensive map illustrating authorship, countries, sources, and keyword co-occurrence, thereby elucidating the interconnectedness of the network. The keyword co-occurrence analysis facilitates the identification of research gaps within the discipline. It enhances the understanding of the interrelations among the subjects previously investigated by scholars in this domain. Based on the bibliographic information, the data in the Research Information Systems (RIS) file was extracted using MRM software, facilitating the generation of an overlay visualization depicting the co-occurrence analysis of keywords, as illustrated in Figure 5.

3.4 Systematic review steps

This segment elucidates the optimal reporting components for systematic reviews and meta-analyses Preferred reporting items for systematic reviews and meta-analyses (PRISMA) relevant to the article review and the overarching research design. Figure 4 illustrates the methodology employed for extracting files from the database and the procedural steps undertaken to eliminate duplicates and extraneous information from the identified or compiled records. Figure 6 shows an additional overview of the review article.

4 Results

4.1 Year-wise publications

After a comprehensive review of 261 scholarly publications, a total of 116 relevant articles were systematically abstracted and examined. The dominant category was journal articles (n = 67), with a relatively limited number of conference papers (n = 39) and books (n = 10). A yearly breakdown of publications shows that production grew at a slower pace in 2019 (3) and 2020 (6) while experiencing significant growth in 2021 (14), 2022 (23), 2023 (32) and 2024 (38) as indicated in Figure 7.

In 2021, a hybrid approach combining ML and physics-based technologies will be used to create DT in the built environment. Cyber-physical systems (CPS) and BIM highlight cyber security concerns related to DT in the building environment (Alshammari et al., 2021; Lin et al., 2021). Highlighting the uniqueness and expanding interest in this research area in

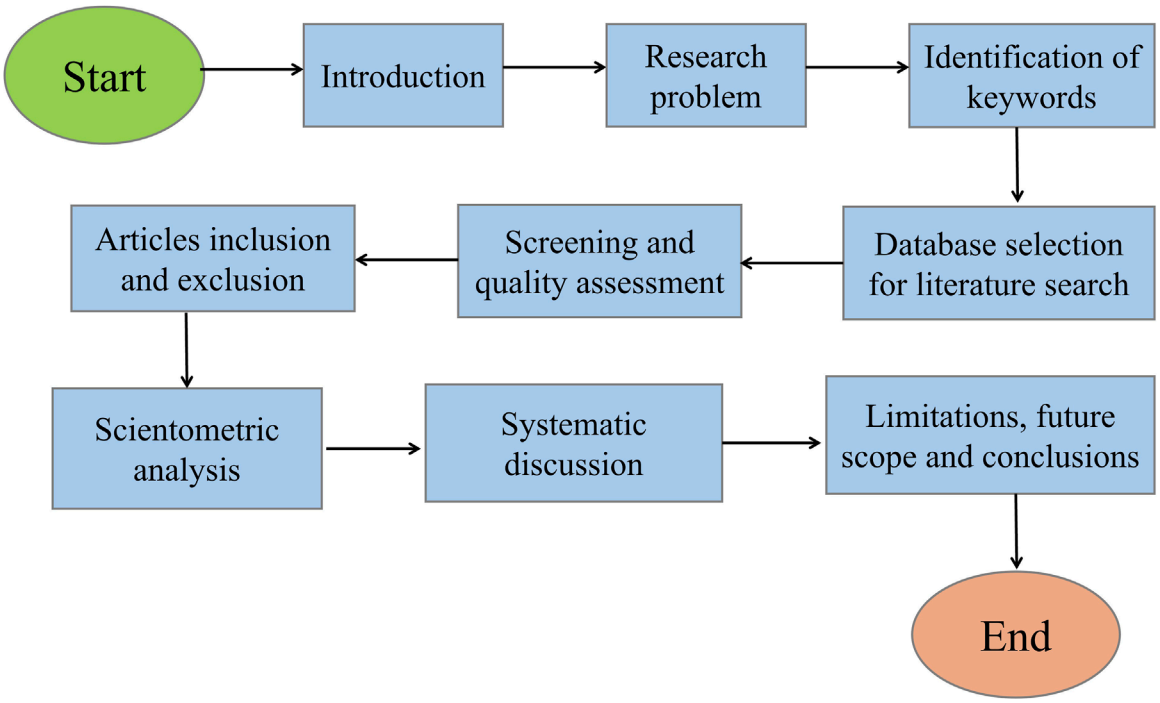


FIGURE 6
Overview of the paper.

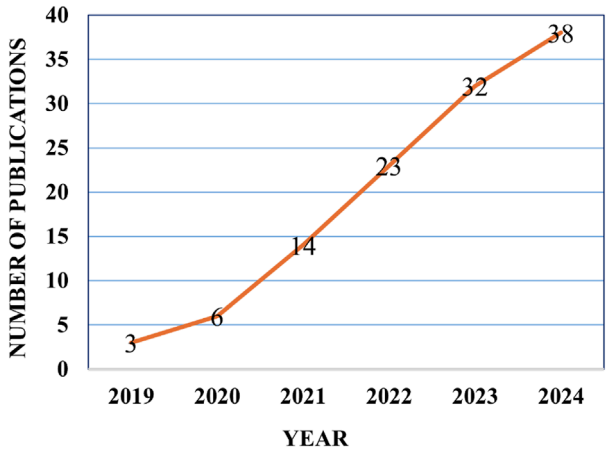


FIGURE 7
Year-wise number of publications.

2022, a Digital Twin Lighting (DTL) system that integrates lighting intelligence with safety features. It analyzes live video feeds, uses YOLOv4 for pedestrian detection, and leverages dynamic BIM to provide a graphical platform for operations and management, integrating DT and green metrics. This concept enhances sustainability planning in smaller buildings and reduces maintenance costs in cities (Corrado et al., 2022; Tan et al., 2022). In 2023, most scholars focused on using DT to enhance building energy efficiency, and they investigated the combination of gray box modeling and building energy management. Using sensor data and thermal energy assessments, it focuses on indoor temperature estimation in a specific zone of an academic facility at Griffith University, improving efficiency and occupant comfort. To develop a framework for energy efficiency assessment of buildings with DT and smart sensors and use dynamic BIM and computer vision to improve energy efficiency can save around 79% (Balali et al., 2023; Jradi and Bjornskov, 2023; Spudys et al., 2023). Virtual Building Models (VBMs) are mathematical representations of actual building behaviors throughout their lifetime, which will be introduced into the built environment in 2024 through VBMs. It has a strong focus on combining *in-situ* modeling with DT technologies to improve decision-making and sustainability in construction activities (Yoon, 2024a).

4.2 Country-wise publications

The total number of publications per country represents two distinct regions using the Scopus database. Figure 8 illustrates a) DT technology for energy efficiency and b) DT technology for indoor environmental quality. China and various European countries are leading research programs on DT aimed at increasing energy efficiency in the construction sector. As illustrated in Figure 8A, India has produced a total of 12 research publications in the last decade. Figure 8B emphasizes the importance of investigating indoor environmental quality, particularly in Europe and Asia, with China, Italy, and the United Kingdom emerging as the most active regions. In the future, India will focus efforts on

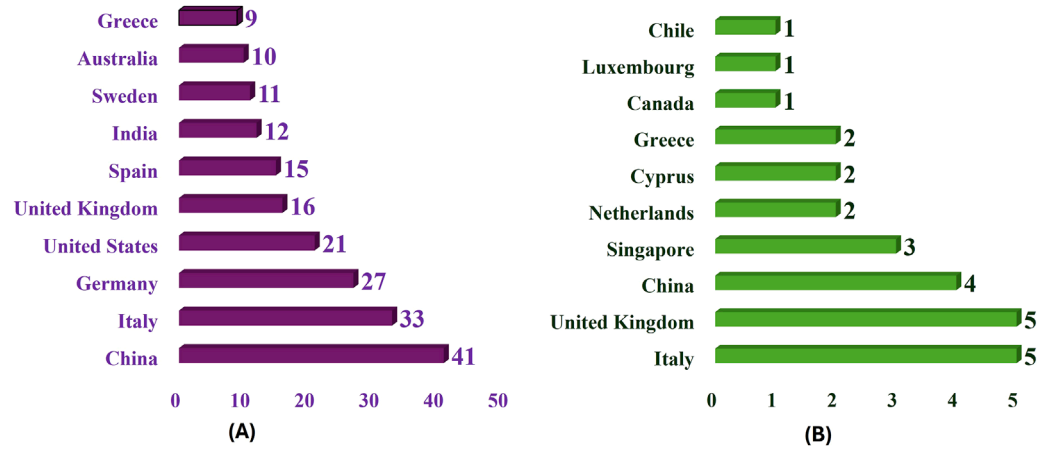


FIGURE 8
Country-wise number of publications of Digital Twin for **(A)** Energy efficiency **(B)** Indoor environmental quality in buildings.

the domain to improve the indoor air quality of its existing buildings.

4.3 Publications per document source

Key areas of study in energy, buildings, and sustainability are highlighted in Table 2, which includes a wide range of academic resources and counts of their papers. The most comprehensive source, seven papers in "Energy and Buildings," highlight the importance of energy-related matters in construction contexts. Six papers, one each on "Buildings and Environment" and "Buildings," focus on structure and environment interactions. Swiss publications with an emphasis on sustainability are particularly prominent. Additional sources address environmental science, energy technology, urban sustainability, and technological advancements. It covers several subjects, with a few papers, including energy informatics, construction automation, and civil engineering. Drawn from various scholarly areas and magazines, this distribution demonstrates a multidisciplinary approach to the subject. Information on the relationship between energy, buildings, and sustainability represents an important priority in research that pays attention to related technical and environmental issues.

TABLE 2 Summary of the number of publications according to different sources.

Sources	Number of documents
Energy and Buildings	7
Building and Environment	6
Buildings	6
Sustainability (Switzerland)	5
Energies	4
Sustainable Energy Technologies and Assessments	4
Science of the Total Environment	3
Sustainable Cities and Society	3
Developments in the Built Environment	3
Future Internet	3
Building Services Engineering Research and Technology	2
Frontiers in Built Environment	2
Journal of Building Engineering	2
Lecture Notes in Civil Engineering	2
Automation in Construction	2
Energy Informatics	2

4.4 Most cited publications with applications

The highest reference articles on DT technology were examined for this review to determine energy efficiency and indoor environmental quality in existing buildings. 17 publications out of 125 satisfied the requirements (Kaewunruen et al., 2019). most cited publications (146) use digital technologies such as BIM for energy management; the article highlights the application of net zero energy building concepts to existing buildings. Followed by (Tagliabue et al., 2021) 101 citations highlight the DT and the Internet of Things, and the framework enables real-time sustainability assessments. (Lydon et al., 2019). 85 citations mainly refer to developing thermal systems for lightweight buildings. They highlight using automation in simulation procedures to increase productivity and reduce expert user time. Table 3 represents some of the authors with the most citations with applications.

TABLE 3 Summary of most cited publications with applications.

References	Citations	Applications
Kaewunruen et al. (2019)	146	The article highlights the application of net zero energy building concepts to existing buildings using digital technologies such as BIM for energy management
Tagliabue et al. (2021)	101	Using DT and the Internet of Things, the framework enables real-time sustainability assessments
Lydon et al. (2019)	85	Development of thermal systems for lightweight buildings to highlight the use of automation in simulation procedures to increase productivity and reduce expert user time
Zhao et al. (2021b)	52	Using DT and BIM for energy efficiency evaluation, it focuses on converting existing buildings into virtually zero-energy buildings using 3D laser scanning to evaluate retrofitting plans
Ghenai et al. (2022)	49	Using DT technology to reduce fuel consumption in many industries is about 63.75% matter. Studies on industrial fuel consumption, increasing sustainability, and energy savings make up 27.5% of the total. Applications related to power generation accounted for 13.75%
Wang et al. (2022)	48	DT can potentially improve urban governance and spur technological innovation. BIM helps with energy-efficient material and design selection to facilitate simulating industrial processes and monitoring real-world systems
Arsiwala et al. (2023)	47	Using IoT, BIM, and AI technologies, the DT system can automate monitoring of comparable CO2 emissions from existing buildings. It seeks to use data to help data-driven retrofitting techniques in the built environment to meet net-zero goals
Clausen et al. (2021)	38	A DT framework manages building heating and ventilation systems that improve occupant comfort and energy efficiency. This allows switching from rule-based control to model predictive control without compromising energy efficiency or comfort
Hosamo et al. (2023a)	34	The framework increases occupant comfort by enabling real-time problem detection in HVAC systems. Examining Sensor data and occupant input assists in predictive maintenance methods
Hosamo et al. (2023b)	27	The approach estimates comfort levels using building performance characteristics and satisfaction surveys. By integrating real-time sensor data, it can identify HVAC problems and improve system efficiency and reliability. The framework achieves energy savings through performance optimization and HVAC system lifetime extension
Petri et al. (2023)	24	DT provides real-time data and insights for better decision-making, improving energy efficiency and sustainability in industrial and construction applications. Using sensors offers advanced monitoring, improves operational effectiveness, and adapts to environmental fluctuations
Spudys et al. (2023)	20	Using DT, smart sensors, and BIM data, the study suggests a framework for operational energy performance assessment that will improve the accuracy of energy assessments. The objective is to enhance the comprehensive understanding of occupant behavior regarding energy consumption and enable energy evaluations
Manfren et al. (2023)	20	This study uses gas-absorption heat pumps and smart thermostatic radiator valves to improve building energy efficiency
Arowoiya et al. (2024)	16	DT improves construction lifecycle management by emphasizing the design, operation, and maintenance phases. By using real-time monitoring and predictive analytics, they increase buildings' thermal comfort and energy efficiency. Applications in DT include decision-making, asset stability optimization, and performance estimation
Tahmasebinia et al. (2023)	15	DT technology optimizes real-time building energy use, including HVAC and lighting systems. It supports energy efficiency assessment and monitoring of indoor environments, facilitating energy control and cost reduction for sustainable buildings
Zhao et al. (2022)	11	This study improves the ventilation designs of public restrooms by increasing indoor air quality and reducing pollutant concentrations. By using computational fluid dynamics (CFD) and DT technologies, efficient ventilation methods can be designed, resulting in healthy environments and energy savings
Hu and Assaad (2024b)	4	Using this framework, the spatial characteristics of air quality at different indoor locations can be efficiently captured, and indoor air quality can be monitored and visualized in real time

5 Systematic discussion

5.1 Digital twin technology for Building energy efficiency

At the global level, energy consumption represents a significant challenge, as an expanding urban population complicates efforts to satisfy the ever-increasing energy demand. The environment is significantly affected by the depletion of energy resources, including climate change, ozone depletion, and global warming. As a result of changes in building design, energy resource transitions, and technological improvements, building energy use has changed dramatically over time.

Energy efficiency in existing buildings has developed significantly because of the need to reduce energy consumption and combat climate change. Many strategies have been used to improve the energy efficiency of ancient and modern buildings, including modern modeling techniques and retrofitting. BIM is becoming a crucial tool for designing energy-efficient buildings, enabling better energy simulations and building thermal system optimization. Research from China has demonstrated how BIM can improve design efficiency and energy standard compliance (Zhao L. et al., 2021). Simulation-based decision support systems have been developed to help investors make educated decisions on energy-efficient solutions, estimating energy consumption after retrofitting (Neves-Silva and Camarinha-Matos, 2022). Using a combination of passive strategies and advanced technology to reduce energy consumption in existing buildings is possible. According to (Pavirani et al., 2023), using demand response algorithms such as Monte Carlo Tree Search (MCTS) can achieve a 4% reduction in energy costs by optimizing heating systems while maintaining thermal comfort. Smart thermostats that detect human activity can save up to 15% in energy by dynamically adjusting set points based on occupancy and facility needs (Mata et al., 2023). Building designs that use thermal mass and passive solar energy can dramatically reduce energy consumption savings by up to 34.71%, which is seen in Mediterranean habitats (Bekele and Atakara, 2023).

Building energy efficiency is being revolutionized by DT, which allows real-time monitoring and optimization of energy use. Ultimately, this technology promotes sustainable behaviors by improving building design, operation, and retrofitting through integration with BIM. Investigates how BIM and DT technologies will be used to improve functional energy assessments of buildings. To bridge the performance gap between asset and operational energy ratings, it proposes a collection of 26 indicators for real-time monitoring and analysis. It highlights the role of digital tools and smart sensors in increasing energy efficiency (Spudys et al., 2023). According to (Kaewunruen et al., 2019), DT facilitates the assessment of NZEB solutions for existing buildings. This enables efficient integration of renewable technologies and accurate cost estimation. Energy consumption analysis can be performed automatically by tools developed using BIM plugins, which significantly helps designers optimize energy consumption during the design phase (Kurniawan et al., 2023). AI-powered simulations coupled with DT can improve building performance through consumption pattern analysis and layout optimization, increasing user comfort and energy efficiency (Almusaed and Yitmen, 2023). The fragmented structure of the construction sector and the demand for specialized personnel in BIM and DT technologies present opportunities but also constraints. To maximize the benefits of energy efficiency, it is imperative to address these gaps through focused training (Alhamami et al., 2020).To improve building occupant comfort and energy efficiency (Clausen et al., 2021), describe a DT framework that utilizes Model Predictive Control (MPC). By integrating occupancy prediction with a multi-objective optimization approach, the methodology demonstrates possible energy savings and improved comfort levels compared to existing management systems (Zhou and Liu, 2024).

A Digital Twin framework for evaluating energy savings strategies for lighting in university classrooms. By controlling occupancy and adjusting lighting, it finds potential savings of more than 60% without the need to replace the system (Seo and Yun, 2022). A computer vision and DT-based method (Tan et al., 2022) for improving interior lighting energy efficiency. The proposed approach significantly reduces energy consumption, savings up to 79.86% by combining real-time data with intelligent control mechanisms. A DT-driven architecture for IoT-based energy trading between intelligent nano grids, improving stability and efficiency. Integrating blockchain for secure transactions and the use of advanced predictive algorithms has revolutionized energy management practices, resulting in a 24% increase in renewable energy use and a 52% reduction in peak demand (Jamil et al., 2024). A cloud-based digitalization framework that utilizes DT, IoT, and AI for energy efficiency and intelligent management of historic structures. The framework uses information collected from environmental sensors to provide analytics and real-time monitoring (Ni et al., 2021).

AI methods and DT will increase energy conservation in the Roman residential area (Agostinelli et al., 2021). It strongly emphasizes improving energy efficiency by combining IoT and providing affordable IT infrastructure for reliable data collection and analysis (Agostinelli et al., 2021). An approach to building energy efficiency using DT to manage issues with thermal parameters and privacy issues raised by residents. Integrating data-driven and mechanism models to provide optimum dispatch speed guarantees fault tolerance and privacy protection. It makes building energy flexibility more effective, which helps achieve carbon neutrality goals (Song et al., 2023)Although much of the research has focused on tracking, monitoring, and optimizing energy efficiency, it has not included applications of DT in this domain recent DT applications for building energy efficiency, as shown in Table 4.

5.2 Digital twin technology for indoor environment quality

As urban growth moves away from adding new buildings towards maximizing existing ones, the environmental quality of such structures has become a significant issue. While they often lack contemporary infrastructure, historic buildings present difficulties in indoor quality management due to antiquated materials and systems that lead to unfavorable environmental conditions (Qian et al., 2024a). Research shows that various pollutants affect residents' wellbeing and health and require proper IEQ management (Qian et al., 2024a). In addition, it has been suggested that integrating cutting-edge technologies such as DT technology and adaptive HVAC systems can improve energy efficiency and indoor environmental quality in heritage buildings, thus addressing the historical context of environmental quality (Zhang et al., 2023). The potential for enhancing the environmental quality of existing buildings by adopting innovative energy management approaches is highlighted, resulting in considerable savings in energy consumption (Borja-Conde et al., 2023). As urban development moves from growth to optimization of built environments, the history of indoor environmental quality in existing buildings has garnered more attention. Studies emphasize how important it is to monitor and regulate indoor environment quality conditions, such as sick building syndrome (SBS), which is brought on by occupants emissions of carbon dioxide (CO_2) and

TABLE 4 Recent applications of DT for energy efficiency.

References	Research method	Tools used	Applications
Cespedes-Cubides and Jradi (2024)	Case study	Arup's Neuron Building twin Autodesk Tandem	Scenario modeling, predictive maintenance, operational optimization, anomaly detection, and component monitoring
Han et al. (2024)	Experimental	Mat lab Unity 3D engine	DT is used in industry, transport, healthcare, and smart cities. Using machine learning, they are developing predictive models for energy efficiency
Aguilera et al. (2024)	Case study	SciPy module for python	Monitoring and optimizing large-scale, fouling-prone heat pumps in real-time. Using a model-based approach for set point optimization improves energy performance
Koltsios et al. (2022)	Experimental	BIM AI	Energy efficiency measurement platform for building DT (BDTs). AI integration for building performance analysis, assessment, and energy efficiency applications for energy consumption correlation in the building design process
Bonomolo et al. (2024)	Case study	DIALux LabView	Thermal comfort study and visualization of buildings, HVAC system optimization, and visual comfort monitoring
Renganayagalu et al. (2024)	Case study	3DF Zephyr Leica BLK 360 Leica Cyclone Unity game engine	They are maximizing energy efficiency with DTs, AI, and IoT sensors. They use real-time data to optimize and assess building energy performance. DTs for buildings are shown in three dimensions for an overview and insights
Kaewunruen et al. (2019)	Case study	Revit Energy +	Prioritize Net Zero Energy Buildings (NZEB) for existing buildings. BIM and hierarchy flow charts are used for evaluation. A prospective framework for NZEB implementations in existing buildings using renewable technology. Feasibility of DT in NZEB buildings for renewable technology
Lu et al. (2020a)	Case study	Autodesk Revit AWS DynamoDB Autodesk Forge API	Monitoring of assets in O&M asset monitoring, Anomaly detection system, and a data integration technique based on extended IFC.
Jafari et al. (2020)	Case study	PMC	Predictive modeling control
Cespedes-Cubides and Jradi (2024)	Case study	Building Energy information systems	Monitor components, detect anomalies, optimize operations, perform predictive maintenance, and create scenarios
Tan et al. (2022)	Experimental	Autodesk Revit Tensorflow used for YOLOv4	Surveillance systems, intelligent lighting systems, and DT lighting systems use the integration of data from multiple sources
Balali et al. (2023)	Case study	OSIsoft WebCTRL server Modbus IP	Energy management, forecasting, and controlling of indoor temperature for buildings and integration of IoT devices and DT.

(Continued on the following page)

TABLE 4 (*Continued*) Recent applications of DT for energy efficiency.

References	Research method	Tools used	Applications
Jradi and Bjornskov (2023)	Experimental	FIWARE TRNSYS EnergyPlus DOE-2	Smart buildings, DT platform applications, comfort, safety, and energy efficiency
Bortolini et al. (2022)	Case study	VOSViewer Excel	Design optimization, occupant comfort, simulation of energy consumption, and building maintenance and management
Dave et al. (2018)	Case study	IoT Sensors BIM Otaniemi3D	IoT sensor for energy consumption, occupancy, and user comfort data. A campus-wide platform for connecting IoT devices and building information models. IoT integration with the built environment is demonstrated through real-world usage scenarios
Fathy et al. (2021)	Case study	IoT smart gateway	Energy efficient home production by optimizing consumption. Efficient energy use using a home-centric energy management system. Real IoT dataset trials for energy optimization were conducted on 17 households

Volatile Organic Compounds (VOCs) (Arsiwala et al., 2023; Chiesa and Vigliotti, 2024a).

DT integrates real-time data from several sensors, essential for monitoring and improving indoor environmental quality in existing buildings. To develop a user-friendly platform that monitors indoor environmental quality parameters like thermal comfort and indoor air quality and pollutants like e CO_2 and TVOC, a DT solution can integrate IoT, BIM, and ML. To create healthier indoor environments, this technology provides real-time insights into the indoor environment and enables future emissions to be predicted. It allows adequate ventilation and air purification systems to be implemented (Arsiwala et al., 2023). Furthermore, the use of DT for adaptive management of HVAC systems increases energy efficiency and guarantees adherence to indoor air quality regulations, protecting both building integrity and occupant comfort (Zhang et al., 2023). Controlling and improving the indoor environmental quality of existing buildings has advanced significantly with the use of DT.

DT technologies have positively impacted various phases of the architecture, engineering, and construction (AEC) sector, incorporating case studies into buildings indoor environments. However, the prevailing scholarly discussion focuses primarily on built infrastructure and urban settings while demonstrating comparatively reduced rigor concerning the landscape sector (Liu et al., 2023). It is crucial to design DT systems, specifically those for indoor building environmental conditions and DT with BIM technologies. In this case study, the model emphasizes using sensors and cameras for data collection to provide real-time identification and dynamic representation of building components. The framework developed is relevant to case studies of indoor settings as it is the basis for effective building management and operational efficiency (Wang et al., 2022). The authors introduce parametric DT for using ontology to model indoor environments in manufactured buildings. DT is combined with deep learning methods in a case study of historic buildings in Osterg, Otland, Sweden. It focused on building indoor temperature and relative humidity prediction models. It demonstrates the potential of a time series dense encoder model for multi-horizon predictions, underscoring the promise of DT technology for indoor environmental optimization (Ni et al., 2023). The authors do not directly cover DT for building indoor environments, but they focus on a case study of a university campus, emphasizing the layered integration of multi-functional models for infrastructure and building lifecycle management. The research demonstrates digital twinning (DODT), highlighting several applications, including condition assessment, construction management, and environmental planning (Chen et al., 2024).

Prior studies have addressed the main barriers to using DT to improve IEQ through integrating several technologies, including real-time data capturing systems, IoT, and BIM. For example, the DT framework was built to monitor and optimize indoor conditions, which addresses data integration and visualization challenges by combining BIM and IoT (Opoku et al., 2024). Furthermore, the accuracy of indoor mapping, which is essential for effective environmental monitoring, has increased with the use of LiDAR-based SLAM to generate high-fidelity point clouds (Hu and Assaad, 2024a). The use of DT optimizes energy performance and operational efficiency. It also gives facility managers insights into energy consumption trends and can enable proactive changes to improve sustainability and comfort (Renganayagalu et al., 2024). A computational framework that analyzes human-generated aerosol and CO_2 buildup in classrooms using CFD simulations and confirmed by *in-situ* observations. It investigates how individual heat sources and breathing momentum flux affect air stratification and aerosol dispersion in a stale air environment. The study highlights the significance of accurate exhaled aerosol and CO_2 modeling (Mahmoud et al., 2024a). A new classification and weighting scheme for IEQ assessment models aimed at improving the relevance and accuracy of assessments. An ensemble hierarchical clustering method is introduced to improve HVAC management practices based on occupancy variety, with

the potential to save energy. Using hidden random variables and physical process equations, thermal comfort systems are argued to maximize occupant satisfaction and incorporate lower energy consumption (Karatzas et al., 2024).

IEQ in buildings has been significantly improved by combining AI techniques with DT technologies using real-time data and predictive analytics. By providing an all-inclusive framework that integrates IoT, BIM, and ML. DT facilitates IEQ monitoring and enables efficient data integration and visualization (Qian et al., 2024a). This approach enables rapid assessment and management of air pollution, guaranteeing prompt resolution of IEQ issues. In addition, data from multiple sensors is analyzed using AI technologies such as supervised and unsupervised learning to increase energy efficiency and occupant comfort (Karatzas et al., 2024). Ultimately, managers can bridge the gap between problem identification and subsequent resolution using a DT platform by enhancing the overall welfare and functionality of internal environments. Table 5 represents recent DT applications for indoor environmental quality in buildings.

6 Discussions

The review results illuminate the application of AI and ML in improving preexisting architectural structures. The prevalence and widespread terminologies associated with DTs, AI, BIM, smart urbanism, architectural design, energy efficiency and consumption, and indoor environmental quality parameters corroborate this claim. Moreover, the results emphasize the need for additional research in underexplored domains such as smart grids, energy storage systems, augmented and virtual reality, 5G connectivity, and edge computing. Ongoing investigations in these domains are imperative to comprehensively utilize AI and ML capabilities to optimize intelligent building systems.

The significant number of citations in the seventeen research papers, as indicated in Table 3, demonstrates the depth of research on applying DT for energy efficiency and indoor environmental quality. Although IEQ has received more attention than energy efficiency, the issue of indoor air quality in existing buildings requires additional attention. Figure 8 illustrates that extensive scholarly research has been undertaken in various global regions, including China, Italy, the United States, the United Kingdom, Germany, Singapore, Spain, and India. In future efforts, enhanced international collaboration is hoped to further advance this research area. Urbanization exacerbates the energy consumption problem, which impacts resource depletion and climate change. BIM and DT technologies are essential in increasing energy efficiency in existing buildings. Building energy efficiency has evolved with design and technological advances. DT supports nearly zero-energy buildings (Kaewunruen et al., 2019), which combines real-time data and smart sensors to increase building efficiency. BIM enhances design and compliance. Using advanced simulations and energy trading, AI can further improve energy efficiency. However, to realize the full potential of these technologies, barriers such as industry fragmentation and the need for specialized training must be overcome.

Preserving the environmental integrity of historic buildings has become an essential priority as urban development shifts from constructing new buildings to renovating older buildings. These historic buildings often have IEQ issues due to old materials and systems that negatively affect people's health and wellbeing. Integrating advanced technology such as DT and adaptive HVAC systems is recommended to overcome these issues. IoT, BIM, and ML are combined in DT to use real-time sensor data to monitor and improve IEQ accurately. This technology provides information about indoor variables, including humidity, temperature, and pollution levels. At the same time, it enables predictive analytics, which can improve ventilation and air purification system management. According to empirical studies, DT significantly improves energy efficiency, operational effectiveness, and compliance with IEQ laws. Furthermore, LiDAR-based SLAM for accurate indoor mapping and CFD simulations for aerosol and CO2 modeling can further improve environmental monitoring and ventilation practices. In summary, the integration of AI and DT has the potential to significantly improve indoor environmental quality by enabling real-time building condition monitoring and optimization, which ultimately leads to improved energy efficiency and occupant comfort.

In addition, DT offers many advantages regarding enhancing IEQ and energy efficiency. It offers several benefits over conventional building management techniques. DT allows for the incorporation of real-time data from IoT devices. It allows for ongoing building performance monitoring, which enhances IEQ and energy efficiency and facilitates improved decision-making. With its emphasis on communications and sustainable cities, DT supports the SDGs and encourages inexpensive, renewable energy and health and wellbeing. DT promotes environmental science, engineering, and construction. In addition, the preventive maintenance methods made possible by DT save maintenance costs and interruptions. Finally, DT is a valuable tool for improving performance compared to traditional approaches.

7 Limitations

It is important to acknowledge several limitations when discussing a systematic review of Digital Twin technology for improving energy efficiency and indoor environmental quality. This review focused on English-language publications, potentially omitting key findings from research disseminated in alternative languages. This analysis makes it clear that the use of DT technology in different typologies and geographical settings has not been adequately examined. This study was limited to specific academic disciplines, raising concerns about the representativeness of the research. This may have overlooked significant contributions from different fields, potentially obscuring important insights and applications of DT in different contexts. Synthesis of results proved problematic due to the different methods used in this study. Since different methods may lead to different results, this discrepancy undermines the reliability of the findings derived from the review.

The prominent use of sensors in the reviewed research focuses primarily on baseline data, thereby complicating comprehensive environmental assessments. This suggests that the full potential of DT technology cannot be realized without sophisticated sensor

TABLE 5 Recent applications of DT for indoor environmental quality.

References	Research method	Tools used	Applications
Hu and Assaad (2024a)	Experimental	Arduino UNO Wi-Fi REV2, MQ7 Gas sensor, Insta 360 × 2 360 panoramic camera, ESP32 UWB	DT framework for indoor air quality monitoring and visualization integrates LiDAR, indoor location, sensing, and autonomous robots
Karatzas et al. (2024)	Experimental	ANN, SVM, RF, DT Gradient boosting, XGB, NB, KNN, and Multilayer perception	DT with real-time control and monitoring for energy-efficient buildings. AI/ML methods to optimize indoor conditions and predict thermal comfort
Nair et al. (2022)	Experimental	Air quality sensor Generalized additive model (GAM)	Techniques to improve IEQ through ventilation. Methods for purifying and filtering air to reduce the risk of viral transmission
Cony Renaud Salis et al. (2017)	Experimental	Modeling tools Field measurements	Developing an IAQ metric for residential low-energy buildings. Evaluating contaminants IAQ using a subset of pollutants
Chiesa and Vigliotti (2024b)	Experimental	Energy simulation software, Calibration signature approach	DMV application of control algorithms to improve IAQ in school buildings. In the school building, mechanical ventilation system monitoring and simulation
Yuan et al. (2024)	Case study	Genetic algorithm, Structural Similarity, CFD simulations, Tecplot	Optimizing sensor placement for IEQ monitoring over wide areas. A genetic method for sensor placement optimization in complex indoor settings
Oh et al. (2023)	Empirical	Data-driven IEQ reports, Statistical analysis, IoT-based sensor network	IEQ reports based on data on changes in office occupant behavior. A health-related real-time indoor environmental quality monitoring system
Mucha et al. (2024)	Experimental	Gravimetric method Andersen impactor	IAQ improvement strategies have been modified for use in different types of buildings. They improve IAQ by influencing building design and retrofitting processes—significant technological advances in advanced filtration systems for motorized air conditioning systems
Qian et al. (2024b)	Case study	BLE Sensors MATLAB WeChat Trimble x7 3D laser scanning	Indoor Air Quality Management with IoT-based DT platform. Rapid prediction models for real-time indoor environmental conditions. BLE sensors and BIM are combined to track the location of interior occupants' algorithms for intelligent control of indoor environments in real-time
Opoku et al. (2024)	Experimental	Autodesk Revit, TogoIO API MultiTech Conduit AP for LoRa Technology	The university library uses a DT to monitor internal conditions. They are combining building information modeling with IoT to obtain real-time data

(Continued on the following page)

TABLE 5 (*Continued*) Recent applications of DT for indoor environmental quality.

References	Research method	Tools used	Applications
Babich et al. (2023)	Case study	Mass balance equation Box (EQ-OX)	Assess IAQ and thermal comfort in school buildings based on established criteria. Analyze data on indoor environmental quality and thermal comfort
Mahmoud et al. (2024b)	Experimental	CFD, TSI sensor, BuoyantPimpleFoam solver	They assessed indoor air quality in the classroom environments with CFD. They examined aerosol dispersion, air stratification, and structure with CO_2

applications that capture a diverse range of data. As a result, the findings of this review may soon be outdated, and further research is needed to accelerate progress in this field. This limitation suggests that the research may not have universal applicability, obscuring key opportunities from different contexts. The review highlights the challenges posed by the heterogeneity of study methods. This inconsistency hinders the synthesis of results and negatively impacts the reliability of conclusions drawn from the review. The effectiveness of the DT methodology primarily depends on the efficiency and accessibility of real-time data. The DT technology domain is undergoing rapid evaluation, making the review impractical in a short time. The manuscript recognizes the need for ongoing research to address these developments, suggesting that the conclusions require urgent reevaluation.

The review highlights the need for a multidisciplinary framework to thoroughly investigate the applications of DT technology. However, achieving effective collaboration across different fields can be challenging, leading to oversimplified conclusions that fail to capture the complexity inherent in the topic.

8 Future scope

Future applications of DT technology in optimizing energy efficiency and improving indoor environmental quality in existing buildings are highly encouraging. However, this requires significant technological advances. Recent scholarly research emphasizes the need for better synthesis of real-time data obtained from different sensors to create more accurate digital models of buildings, thus enabling superior monitoring and control of energy consumption and indoor environments. Moreover, expanding sensor usage beyond basic metrics such as temperature and humidity is imperative for holistic environmental assessments. There is a need for research efforts tailored explicitly to specific architectural typologies and geographic contexts. Realizing the operational capability of DT in different climatic conditions and by varying construction regulations will significantly improve its relevance and efficiency. Although the review describes shortcomings in the existing literature, it does not provide comprehensive guidelines for correcting them. Further research should produce clear methodological frameworks for future inquiries in this dynamic domain.

LiDAR technology improves the accuracy of DT by supporting high-resolution spatial information, facilitating indoor mapping, and enabling real-time surveillance and visualization. Implementing advanced predictive algorithms, including deep learning and reinforcement learning methodologies, can significantly enhance energy management strategies. Furthermore, promoting global collaboration and cross-disciplinary studies is crucial to reduce knowledge gaps and optimize DT implementations in practical contexts. These advances ultimately contribute to more sustainable construction practices and improved occupant satisfaction in the former.

9 Conclusion

According to a systematic review, DT technology holds great promise for improving building indoor environmental quality and energy efficiency. DT can be used to create real-time simulations that optimize building efficiency and improve occupant comfort by integrating with various digital technologies, including the Internet of Things and building information modeling. The review findings suggest that integrating DT can lead to significant energy savings and improved construction performance in various building types. The assessment highlights the potential of digital transformation technology to provide analytical insights derived from real-time data to support more informed decision-making at every stage of the buildings life cycle, from design to operation management. However, the analysis also points to several difficulties and limitations of the current application of DT technology. The narrow significance of current research, mainly limited to studies published in English and specific academic subjects, is a serious cause for concern. This can show how broadly the results can be applied and obscure critical perspectives from other fields and geographies. To fill these gaps, we anticipate the need for more research into how well DT works in different building types and locations. To ensure that anyone can use the benefits of DT technology, future research will work to provide standardized procedures for evaluating its effects.

Finally, DT offers a revolutionary opportunity to improve buildings indoor environmental quality and energy efficiency. However, to fully realize its potential, additional research is mandatory to overcome existing challenges. The findings described in the article illustrate the complex interrelationships between technological processes, sustainability theories, and health outcomes, thus underscoring the need for ongoing scholarly exploration and application of DT in the built environment. Conclusions drawn from the review form a framework for further

research and advocate for a multidisciplinary perspective to examine different applications of DT in the built environment, ultimately facilitating the advancement of more sustainable and efficient construction technologies.

Author contributions

NV: Conceptualization, Data curation, Methodology, Software, Visualization, Writing–original draft. MS: Investigation, Supervision, Validation, Writing–review and editing.

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The authors declare that no Generative AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fbuil.2025.1523464/full#supplementary-material

References

Abanda, F. H., Jian, N., Adukpo, S., Tuhaise, V. V., and Manjia, M. B. (2024). "Digital twin for product versus project lifecycles' development in manufacturing and construction industries," in *Journal of intelligent manufacturing* (Springer). doi:10.1007/s10845-023-02301-2

Agostinelli, S., Cumo, F., Guidi, G., and Tomazzoli, C. (2021). Cyber-physical systems improving building energy management: digital twin and artificial intelligence. *Energies* 14 (8), 2338. doi:10.3390/en14082338

Agouzoul, A., Simeu, E., and Tabaa, M. (2023). "Enhancement of building energy consumption using a digital twin based neural network model predictive control," in International conference on control, automation and diagnosis, ICCAD 2023, Rome, Italy, 10-12 May 2023. doi:10.1109/ICCAD57653.2023.10152308

Agouzoul, A., Tabaa, M., Chegari, B., Simeu, E., Dandache, A., and Alami, K. (2021). Towards a digital twin model for building energy management: case of Morocco. *Procedia Comput. Sci.* 184, 404–410. doi:10.1016/j.procs.2021.03.051

Aguilera, J. J., Meesenburg, W., Markussen, W. B., Zühlsdorf, B., and Elmegaard, B. (2024). Real-time monitoring and optimization of a large-scale heat pump prone to fouling - towards a digital twin framework. *Appl. Energy* 365, 123274. doi:10.1016/j.apenergy.2024.123274

Akanmu, A. A., Olayiwola, J., Ogunseiju, O., and McFeeters, D. (2020). Cyber-physical postural training system for construction workers. *Automation Constr.* 117, 103272. doi:10.1016/j.autcon.2020.103272

Alhamami, A., Petri, I., Rezgui, Y., and Kubicki, S. (2020). Promoting energy efficiency in the built environment through adapted BIM training and education. *Energies* 13 (9), 2308. doi:10.3390/en13092308

Almusaed, A., and Yitmen, I. (2023). Architectural reply for smart building design concepts based on artificial intelligence simulation models and digital twins. *Sustain. Switz.* 15 (6), 4955. doi:10.3390/su15064955

Alshammari, K., Beach, T., and Rezgui, Y. (2021). Cybersecurity for digital twins in the built environment: current research and future directions. *J. Inf. Technol. Constr.* 26, 159–173. doi:10.36680/j.itcon.2021.010

Angjeliu, G., Coronelli, D., and Cardani, G. (2020). Development of the simulation model for Digital Twin applications in historical masonry buildings: the integration between numerical and experimental reality. *Comput. Struct.* 238, 106282. doi:10.1016/j.compstruc.2020.106282

Arowoiya, V. A., Moehler, R. C., and Fang, Y. (2024). Digital twin technology for thermal comfort and energy efficiency in buildings: a state-of-the-art and future directions. *Energy Built Environ.* 5 (5), 641–656. doi:10.1016/j.enbenv.2023.05.004

Arsiwala, A., Elghaish, F., and Zoher, M. (2023). Digital twin with Machine learning for predictive monitoring of CO2 equivalent from existing buildings. *Energy Build.* 284, 112851. doi:10.1016/j.enbuild.2023.112851

Asif, M., Naeem, G., and Khalid, M. (2024). Digitalization for sustainable buildings: technologies, applications, potential, and challenges. *J. Clean. Prod.* 450, 141814. doi:10.1016/j.jclepro.2024.141814

Author Anonymous (2024). Digital Twins The Convergence of Multimedia Technologies. Available at: https://www.computer.org/multimedia.

Babich, F., Torriani, G., Corona, J., and Lara-Ibeas, I. (2023). Comparison of indoor air quality and thermal comfort standards and variations in exceedance for school buildings. *J. Build. Eng.* 71, 106405. doi:10.1016/j.jobe.2023.106405

Balali, Y., Busch, A., and O'Keefe, S. (2023). "Modelling and prediction of energy efficient building climate toward digital twin integration," in *Asia-pacific power and energy engineering conference* (APPEEC). doi:10.1109/APPEEC57400.2023.10561934

Bekele, M. T., and Atakara, C. (2023). Residential building energy conservation in mediterranean climate zone by integrating passive solar and energy efficiency design strategies. *Buildings* 13 (4), 1073. doi:10.3390/buildings13041073

Bonomolo, M., Testasecca, T., Buscemi, A., Alberto Munafo, F. L., and Beccali, M. (2024). "Digital twin-based assessment framework for monitoring visual comfort," in 2024 IEEE 22nd mediterranean electrotechnical conference, MELECON 2024, Porto, Portugal, 25-27 June 2024, 1139–1145. doi:10.1109/MELECON56669.2024.10608560

Borja-Conde, J. A., Witheephanich, K., Coronel, J. F., and Limon, D. (2023). Thermal modeling of existing buildings in high-fidelity simulators: a novel, practical methodology. *Energy Build.* 292, 113127. doi:10.1016/j.enbuild.2023.113127

Bortolini, R., Rodrigues, R., Alavi, H., Vecchia, L. F. D., and Forcada, N. (2022). Digital twins' applications for building energy efficiency: a review. *Energies* 15 (Issue 19), 7002. doi:10.3390/en15197002

Cai, J., Chen, J., Hu, Y., Li, S., and He, Q. (2023). Digital twin for healthy indoor environment: a vision for the post-pandemic era. *Front. Eng. Manag.* 10 (2), 300–318. doi:10.1007/s42524-022-0244-y

Cespedes-Cubides, A. S., and Jradi, M. (2024). A review of building digital twins to improve energy efficiency in the building operational stage. *Springer Nat.* 7 (1), 11. doi:10.1186/s42162-024-00313-7

Chen, G., Alomari, I., Taffese, W. Z., Shi, Z., Afsharmovahed, M. H., Mondal, T. G., et al. (2024). Multifunctional Models in Digital and Physical Twinning of the Built Environment—A University Campus Case Study. *Smart Cities* 7 (2), 836–858. doi:10.3390/smartcities7020035

Chiesa, G., and Vigliotti, M. (2024a). Comparing mechanical ventilation control strategies for indoor air quality: monitoring and simulation results of a school

building in northern Italy. *Energy Build.* 322, 114665. doi:10.1016/j.enbuild.2024.114665

Chiesa, G., and Vigliotti, M. (2024b). Comparing mechanical ventilation control strategies for indoor air quality: monitoring and simulation results of a school building in northern Italy. *Energy Build.* 322, 114665. doi:10.1016/j.enbuild.2024.114665

Clausen, A., Arendt, K., Johansen, A., Sangogboye, F. C., Kjærgaard, M. B., Veje, C. T., et al. (2021). A digital twin framework for improving energy efficiency and occupant comfort in public and commercial buildings. *Energy Inf.* 4, 40. doi:10.1186/s42162-021-00153-9

Cobo, M. J., López-Herrera, A. G., Herrera-Viedma, E., and Herrera, F. (2011). Science mapping software tools: review, analysis, and cooperative study among tools. *J. Am. Soc. Inf. Sci. Technol.* 62 (7), 1382–1402. doi:10.1002/asi.21525

Conroy, M. (2010). Modeling, simulation, information technology and processing roadmap. Available at: https://www.researchgate.net/publication/280310295.

Cony Renaud Salis, L., Abadie, M., Wargocki, P., and Rode, C. (2017). Towards the definition of indicators for assessment of indoor air quality and energy performance in low-energy residential buildings. *Energy Build.* 152, 492–502. doi:10.1016/j.enbuild.2017.07.054

Corrado, C. R., DeLong, S. M., Holt, E. G., Hua, E. Y., and Tolk, A. (2022). Combining green metrics and digital twins for sustainability planning and governance of smart buildings and cities. *Sustain. Switz.* 14 (20), 12988. doi:10.3390/su142012988

Dave, B., Buda, A., Nurminen, A., and Främling, K. (2018). A framework for integrating BIM and IoT through open standards. *Automation Constr.* 95, 35–45. doi:10.1016/j.autcon.2018.07.022

Deng, M., Menassa, C. C., and Kamat, V. R. (2021). From BIM to digital twins: a systematic review of the evolution of intelligent building representations in the AEC-FM industry. *J. Inf. Technol. Constr.* 26, 58–83. doi:10.36680/J.ITCON.2021.005

Desogus, G., Frau, C., Quaquero, E., and Rubiu, G. (2023). From building information model to digital twin: a framework for building thermal comfort monitoring, visualizing, and assessment. *Buildings* 13 (8), 1971. doi:10.3390/buildings13081971

Drobnyi, V., Li, S., and Brilakis, I. (2024). Connectivity detection for automatic construction of building geometric digital twins. *Automation Constr.* 159, 105281. doi:10.1016/j.autcon.2024.105281

El-Gohary, M., El-Abed, R., and Omar, O. (2023). Prediction of an efficient energy-consumption model for existing residential buildings in Lebanon using an artificial neural network as a digital twin in the era of climate change. *Buildings* 13 (12), 3074. doi:10.3390/buildings13123074

Es-haghi, M. S., Anitescu, C., and Rabczuk, T. (2024). Methods for enabling real-time analysis in digital twins: a literature review. *Comput. Struct.* 297, 107342. doi:10.1016/j.compstruc.2024.107342

Fathy, Y., Jaber, M., and Nadeem, Z. (2021). Digital twin-driven decision making and planning for energy consumption. *J. Sens. Actuator Netw.* 10 (2), 37. doi:10.3390/JSAN10020037

Ghansah, F. A., and Lu, W. (2024). Major opportunities of digital twins for smart buildings: a scientometric and content analysis. *Smart Sustain. Built Environ.* 13 (Issue 1), 63–84. doi:10.1108/SASBE-09-2022-0192

Ghenai, C., Husein, L. A., Al Nahlawi, M., Hamid, A. K., and Bettayeb, M. (2022). Recent trends of digital twin technologies in the energy sector: a comprehensive review. *Sustain. Energy Technol. Assessments* 54. doi:10.1016/j.seta.2022.102837

Greif, T., Stein, N., and Flath, C. M. (2020). Peeking into the void: digital twins for construction site logistics. *Comput. Industry* 121, 103264. doi:10.1016/j.compind.2020.103264

Grieves, M., and Vickers, J. (2016). "Digital twin: mitigating unpredictable, undesirable emergent behavior in complex systems," in *Transdisciplinary perspectives on complex systems: new findings and approaches* (Springer International Publishing), 85–113. doi:10.1007/978-3-319-38756-7_4

Hadjidemetriou, L., Stylianidis, N., Englezos, D., Papadopoulos, P., Eliades, D., Timotheou, S., et al. (2023). A digital twin architecture for real-time and offline high granularity analysis in smart buildings. *Sustain. Cities Soc.* 98, 104795. doi:10.1016/j.scs.2023.104795

Han, F., Du, F., Jiao, S., and Zou, K. (2024). Predictive analysis of a building's power consumption based on digital twin platforms. *Energies* 17 (15), 3692. doi:10.3390/en17153692

Hosamo, H. H., Nielsen, H. K., Alnmr, A. N., Svennevig, P. R., and Svidt, K. (2022). A review of the Digital Twin technology for fault detection in buildings. *Front. Built Environ.* 8. doi:10.3389/fbuil.2022.1013196

Hosamo, H. H., Nielsen, H. K., Kraniotis, D., Svennevig, P. R., and Svidt, K. (2023a). Digital Twin framework for automated fault source detection and prediction for comfort performance evaluation of existing non-residential Norwegian buildings. *Energy Build.* 281, 112732. doi:10.1016/j.enbuild.2022.112732

Hosamo, H. H., Nielsen, H. K., Kraniotis, D., Svennevig, P. R., and Svidt, K. (2023b). Improving building occupant comfort through a digital twin approach: a Bayesian network model and predictive maintenance method. *Energy Build.* 288, 112992. doi:10.1016/j.enbuild.2023.112992

Hu, X., and Assaad, R. H. (2024a). A BIM-enabled digital twin framework for real-time indoor environment monitoring and visualization by integrating autonomous robotics, LiDAR-based 3D mobile mapping, IoT sensing, and indoor positioning technologies. *J. Build. Eng.* 86, 108901. doi:10.1016/j.jobe.2024.108901

Hu, X., and Assaad, R. H. (2024b). A BIM-enabled digital twin framework for real-time indoor environment monitoring and visualization by integrating autonomous robotics, LiDAR-based 3D mobile mapping, IoT sensing, and indoor positioning technologies. *J. Build. Eng.* 86, 108901. doi:10.1016/j.jobe.2024.108901

Jafari, M. A., Zaidan, E., Ghofrani, A., Mahani, K., and Farzan, F. (2020). Improving building energy footprint and asset performance using digital twin technology. *IFAC-PapersOnLine* 53 (3), 386–391. doi:10.1016/j.ifacol.2020.11.062

Jamil, H., Jian, Y., Jamil, F., Hijjawi, M., and Muthanna, A. (2024). Digital twin-driven architecture for AIoT-based energy service provision and optimal energy trading between smart nanogrids. *Energy Build.* 319, 114463. doi:10.1016/j.enbuild.2024.114463

Jradi, M., and Bjornskov, J. (2023). "A digital twin platform for energy efficient and smart buildings applications," in 2023 5th international conference on advances in computational tools for engineering applications, ACTEA 2023, Zouk Mosbeh, Lebanon, 05-07 July 2023, 1–6. doi:10.1109/ACTEA58025.2023.10194071

Kaewunruen, S., Rungskunroch, P., and Welsh, J. (2019). A digital-twin evaluation of Net Zero Energy Building for existing buildings. *Sustain. Switz.* 11 (1), 159. doi:10.3390/su11010159

Karatzas, S., Papageorgiou, G., Lazari, V., Bersimis, S., Fousteris, A., Economou, P., et al. (2024). A text analytic framework for gaining insights on the integration of digital twins and machine learning for optimizing indoor building environmental performance. *Dev. Built Environ.* 18, 100386. doi:10.1016/j.dibe.2024.100386

Koltsios, S., Katsaros, N., Mpouzianas, N., Klonis, P., Giannopoulos, G., Pastaltzidis, I., et al. (2022). Digital twin application on next-generation building energy performance certification scheme. *ISC2 2022 - 8th IEEE Int. Smart Cities Conf.*, 1–7. doi:10.1109/ISC255366.2022.9921821

Koo, J., and Yoon, S. (2024). Simultaneous *in-situ* calibration for physical and virtual sensors towards digital twin-enabled building operations. *Adv. Eng. Inf.* 59, 102239. doi:10.1016/j.aei.2023.102239

Kritzinger, W., Karner, M., Traar, G., Henjes, J., and Sihn, W. (2018). Digital Twin in manufacturing: a categorical literature review and classification. *IFAC-PapersOnLine* 51 (11), 1016–1022. doi:10.1016/j.ifacol.2018.08.474

Kurniawan, T. B., Dewi, D. A., Usman, F., and Fadly, F. (2023). Towards energy analysis and efficiency for sustainable buildings. *Emerg. Sci. J.* 7 (6), 2226–2238. doi:10.28991/ESJ-2023-07-06-022

Lin, Y.-C., and Cheung, W.-F. (2020). Developing WSN/BIM-Based environmental monitoring management system for parking garages in smart cities. *J. Manage. Eng.* 36. doi:10.1061/(asce)me.1943-5479.0000760

Lin, Y. W., Tang, T. L. E., and Spanos, C. J. (2021). "Hybrid approach for digital twins in the built environment," in *E-energy 2021 - proceedings of the 2021 12th ACM international conference on future energy systems*, 450–457. doi:10.1145/3447555.3466585

Liu, C., Zhang, P., and Xu, X. (2023). Literature review of digital twin technologies for civil infrastructure. *Elsevier B.V* 2 (Issue 3), 100050. doi:10.1016/j.iintel.2023.100050

Lu, Q., Chen, L., Li, S., and Pitt, M. (2020a). Semi-automatic geometric digital twinning for existing buildings based on images and CAD drawings. *Automation Constr.* 115, 103183. doi:10.1016/j.autcon.2020.103183

Lu, Q., Xie, X., Parlikad, A. K., and Schooling, J. M. (2020b). Digital twin-enabled anomaly detection for built asset monitoring in operation and maintenance. *Automation Constr.* 118, 103277. doi:10.1016/j.autcon.2020.103277

Lydon, G. P., Caranovic, S., Hischier, I., and Schlueter, A. (2019). Coupled simulation of thermally active building systems to support a digital twin. *Energy Build.* 202, 109298. doi:10.1016/j.enbuild.2019.07.015

Madni, A. M., Madni, C. C., and Lucero, S. D. (2019). Leveraging digital twin technology in model-based systems engineering. *Systems* 7 (1), 7. doi:10.3390/systems7010007

Madubuike, O. C., Anumba, C. J., and Khallaf, R. (2022). A review of digital twin applications in construction. *J. Inf. Technol. Constr.* 27, 145–172. doi:10.36680/j.itcon.2022.008

Mahmoud, M. M. A., Bahl, P. V. de A., Aquino, A. F., MacIntyre, C. R., Bhattacharjee, S., Green, D., et al. (2024a). A numerical framework for the analysis of indoor air quality in a classroom. *J. Build. Eng.* 92, 109659. doi:10.1016/j.jobe.2024.109659

Mahmoud, M. M. A., Bahl, P. V. de A., Aquino, A. F., MacIntyre, C. R., Bhattacharjee, S., Green, D., et al. (2024b). A numerical framework for the analysis of indoor air quality in a classroom. *J. Build. Eng.* 92, 109659. doi:10.1016/j.jobe.2024.109659

Manfren, M., James, P. A., Aragon, V., and Tronchin, L. (2023). Lean and interpretable digital twins for building energy monitoring – a case study with smart thermostatic radiator valves and gas absorption heat pumps. *Energy AI* 14, 100304. doi:10.1016/j.egyai.2023.100304

Manzoor, B., Othman, I., and Pomares, J. C. (2021). Digital technologies in the architecture, engineering and construction (Aec) industry—a bibliometric—qualitative

literature review of research activities. *Int. J. Environ. Res. Public Health* 18 (11), 6135. doi:10.3390/ijerph18116135

Mata, O., Méndez, J. I., Ponce, P., Peffer, T., Meier, A., and Molina, A. (2023). Energy savings in buildings based on image depth sensors for human activity recognition. *Energies* 16 (3), 1078. doi:10.3390/en16031078

Meho, L. I., and Rogers, Y. (2008). Citation counting, citation ranking, and h-index of human-computer interaction researchers: a comparison of scopus and web of science. *J. Am. Soc. Inf. Sci. Technol.* 59 (11), 1711–1726. doi:10.1002/asi.20874

Moshood, T. D., Rotimi, J. O., Shahzad, W., and Bamgbade, J. A. (2024). Infrastructure digital twin technology: a new paradigm for future construction industry. *Technol. Soc.* 77, 102519. doi:10.1016/j.techsoc.2024.102519

Mucha, W., Mainka, A., and Brągoszewska, E. (2024). Impact of ventilation system retrofitting on indoor air quality in a single-family building. *Build. Environ.* 262, 111830. doi:10.1016/j.buildenv.2024.111830

Nair, A. N., Anand, P., George, A., and Mondal, N. (2022). A review of strategies and their effectiveness in reducing indoor airborne transmission and improving indoor air quality. *Environ. Res.* 213, 113579. doi:10.1016/j.envres.2022.113579

Neves-Silva, R., and Camarinha-Matos, L. M. (2022). Simulation-based decision support system for energy efficiency in buildings retrofitting. *Sustain. Switz.* 14 (19), 12216. doi:10.3390/su141912216

Ni, Z., Eriksson, P., Liu, Y., Karlsson, M., and Gong, S. (2021). Improving energy efficiency while preserving historic buildings with digital twins and artificial intelligence. *IOP Conf. Ser. Earth Environ. Sci.* 863 (1), 012041. doi:10.1088/1755-1315/863/1/012041

Ni, Z., Zhang, C., Karlsson, M., and Gong, S. (2023). "Leveraging deep learning and digital twins to improve energy performance of buildings," in 2023 IEEE 3rd international conference on industrial electronics for sustainable energy systems, IESES 2023, Shanghai, China, 26-28 July 2023. doi:10.1109/IESES53571.2023.10253721

Oh, J., Wong, W., Castro-Lacouture, D., Lee, J. H., and Koo, C. (2023). Indoor environmental quality improvement in green building: occupant perception and behavioral impact. *J. Build. Eng.* 69, 106314. doi:10.1016/j.jobe.2023.106314

Opoku, D. G. J., Perera, S., Osei-Kyei, R., Rashidi, M., Bamdad, K., and Famakinwa, T. (2024). Digital twin for indoor condition monitoring in living labs: university library case study. *Automation Constr.* 157, 105188. doi:10.1016/j.autcon.2023.105188

Pavirani, F., Gokhale, G., Claessens, B., and Develder, C. (2023). Demand response for residential building heating: effective Monte Carlo Tree Search control based on physics-informed neural networks. Available at: http://arxiv.org/abs/2312.03365.

Petri, I., Rezgui, Y., Ghoroghi, A., and Alzahrani, A. (2023). Digital twins for performance management in the built environment. *J. Industrial Inf. Integration* 33, 100445. doi:10.1016/j.jii.2023.100445

Pregnolato, M., Gunner, S., Voyagaki, E., De Risi, R., Carhart, N., Gavriel, G., et al. (2022). Towards civil engineering 4.0: concept, workflow and application of digital twins for existing infrastructure. *Automation Constr.* 141, 104421. doi:10.1016/j.autcon.2022.104421

Qian, Y., Leng, J., Zhou, K., and Liu, Y. (2024a). How to measure and control indoor air quality based on intelligent digital twin platforms: a case study in China. *Build. Environ.* 253, 111349. doi:10.1016/j.buildenv.2024.111349

Qian, Y., Leng, J., Zhou, K., and Liu, Y. (2024b). How to measure and control indoor air quality based on intelligent digital twin platforms: a case study in China. *Build. Environ.* 253, 111349. doi:10.1016/j.buildenv.2024.111349

Renganayagalu, S. K., Bodal, T., Bryntesen, T. R., and Kvalvik, P. (2024). "Optimising energy performance of buildings through digital twins and machine learning: lessons learnt and future directions," in *2024 4th international conference on applied artificial intelligence, ICAPAI 2024*. doi:10.1109/ICAPAI61893.2024.10541224

Rosen, R., Von Wichert, G., Lo, G., and Bettenhausen, K. D. (2015). About the importance of autonomy and digital twins for the future of manufacturing. *IFAC-PapersOnLine* 28 (3), 567–572. doi:10.1016/j.ifacol.2015.06.141

Saback, V., Popescu, C., Blanksvärd, T., and Täljsten, B. (2024). Analysis of digital twins in the construction industry: practical applications, purpose, and parallel with other industries. *Buildings* 14 (Issue 5), 1361. doi:10.3390/buildings14051361

Salihu, C., Hussein, M., Mohandes, S. R., and Zayed, T. (2022). Towards a comprehensive review of the deterioration factors and modeling for sewer pipelines: a hybrid of bibliometric, scientometric, and meta-analysis approach. *J. Clean. Prod.* 351, 131460. doi:10.1016/j.jclepro.2022.131460

Seo, H., and Yun, W. S. (2022). Digital twin-based assessment framework for energy savings in university classroom lighting. *Buildings* 12 (5), 544. doi:10.3390/buildings12050544

Sharma, A., Kosasih, E., Zhang, J., Brintrup, A., and Calinescu, A. (2022). Digital Twins: state of the art theory and practice, challenges, and open research questions. *J. Industrial Inf. Integration* 30, 100383. doi:10.1016/j.jii.2022.100383

Song, Y., Xia, M., Chen, Q., and Chen, F. (2023). A data-model fusion dispatch strategy for the building energy flexibility based on the digital twin. *Appl. Energy* 332, 120496. doi:10.1016/j.apenergy.2022.120496

Spudys, P., Afxentiou, N., Georgali, P.-Z., Klumbyte, E., Jurelionis, A., and Fokaides, P. (2023). Classifying the operational energy performance of buildings with the use of digital twins. *Energy Build.* 290, 113106. doi:10.1016/j.enbuild.2023.113106

Suleny Bojorquez-Roque, M., Garcia-Cabot, A., Garcia-Lopez, E., and Magdiel Oliva-Cordova, L. (2024). Digital competence learning ecosystem in higher education: a mapping and systematic review of the literature. *IEEE Access* 12, 87596–87614. doi:10.1109/ACCESS.2024.3416906

Tagliabue, L. C., Cecconi, F. R., Maltese, S., Rinaldi, S., Ciribini, A. L. C., and Flammini, A. (2021). Leveraging digital twin for sustainability assessment of an educational building. *Sustain. Switz.* 13 (2), 480. doi:10.3390/su13020480

Tahmasebinia, F., Lin, L., Wu, S., Kang, Y., and Sepasgozar, S. (2023). Exploring the benefits and limitations of digital twin technology in building energy. *Appl. Sci. Switz.* 13 (Issue 15), 8814. doi:10.3390/app13158814

Tan, Y., Chen, P., Shou, W., and Sadick, A. M. (2022). Digital Twin-driven approach to improving energy efficiency of indoor lighting based on computer vision and dynamic BIM. *Energy Build.* 270, 112271. doi:10.1016/j.enbuild.2022.112271

Tao, F., Xiao, B., Qi, Q., Cheng, J., and Ji, P. (2022). Digital twin modeling. *J. Manuf. Syst.* 64, 372–389. doi:10.1016/j.jmsy.2022.06.015

Tuegel, E. J., Ingraffea, A. R., Eason, T. G., and Spottswood, S. M. (2011). Reengineering aircraft structural life prediction using a digital twin. *Int. J. Aerosp. Eng.* 2011, 1–14. doi:10.1155/2011/154798

Wang, W., Guo, H., Li, X., Tang, S., Xia, J., and Lv, Z. (2022). Deep learning for assessment of environmental satisfaction using BIM big data in energy efficient building digital twins. *Sustain. Energy Technol. Assessments* 50, 101897. doi:10.1016/j.seta.2021.101897

Ye, X., Jamonnak, S., Van Zandt, S., Newman, G., and Suermann, P. (2024). Developing campus digital twin using interactive visual analytics approach. *Front. Urban Rural Plan.* 2 (1), 9. doi:10.1007/s44243-024-00033-2

Yoon, S. (2023). Building digital twinning: data, information, and models. *J. Build. Eng.* 76, 107021. doi:10.1016/j.jobe.2023.107021

Yoon, S. (2024a). Virtual building models in built environments. *Dev. Built Environ.* 18, 100453. doi:10.1016/j.dibe.2024.100453

Yoon, S. (2024b). Virtual building models in built environments. *Dev. Built Environ.* 18, 100453. doi:10.1016/j.dibe.2024.100453

Yuan, M., Geng, Y., Lin, B., Tang, H., and Yang, Y. (2024). Optimization of indoor temperature sensor deployment in large spaces for multiple building operation scenarios using the genetic algorithm. *J. Build. Eng.* 96, 110446. doi:10.1016/j.jobe.2024.110446

Zhang, J., Chan, C. C. C., Kwok, H. H. L., and Cheng, J. C. P. (2023). Multi-indicator adaptive HVAC control system for low-energy indoor air quality management of heritage building preservation. *Build. Environ.* 246, 110910. doi:10.1016/j.buildenv.2023.110910

Zhao, L., Zhang, H., Wang, Q., Sun, B., Liu, W., Qu, K., et al. (2022). Digital twin evaluation of environment and health of public toilet ventilation design based on building information modeling. *Buildings* 12 (4), 470. doi:10.3390/buildings12040470

Zhao, L., Zhang, H., Wang, Q., and Wang, H. (2021a). Digital-twin-based evaluation of nearly zero-energy building for existing buildings based on scan-to-BIM. *Adv. Civ. Eng.* 2021. doi:10.1155/2021/6638897

Zhao, T., Qu, Z., Liu, C., and Li, K. (2021b). BIM-based analysis of energy efficiency design of building thermal system and HVAC system based on GB50189-2015 in China. *Int. J. Low-Carbon Technol.* 16 (4), 1277–1289. doi:10.1093/ijlct/ctab051

Zhou, Y., and Liu, J. (2024). Advances in emerging digital technologies for energy efficiency and energy integration in smart cities. *Energy Build.* 315, 114289. doi:10.1016/j.enbuild.2024.114289

frontiers | Frontiers in Built Environment

TYPE Original Research
PUBLISHED 24 April 2025
DOI 10.3389/fbuil.2025.1566784

Check for updates

OPEN ACCESS

EDITED BY
Vagelis Plevris,
Qatar University, Qatar

REVIEWED BY
Vanni Nicoletti,
Marche Polytechnic University, Italy
Alberto Cerezo Narváez,
University of Cádiz, Spain

*CORRESPONDENCE
Roma Almeida,
roma.almeida@lnu.se

RECEIVED 25 January 2025
ACCEPTED 31 March 2025
PUBLISHED 24 April 2025

CITATION
Shehu MO, Almeida R and Mahapatra K (2025) Opportunities for digital tracking technologies in the precast concrete sector in Sweden. *Front. Built Environ.* 11:1566784. doi: 10.3389/fbuil.2025.1566784

COPYRIGHT
© 2025 Shehu, Almeida and Mahapatra. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Opportunities for digital tracking technologies in the precast concrete sector in Sweden

Maryam Olaoti Shehu, Roma Almeida* and Krushna Mahapatra

Department of Built Environment and Energy Technology, Faculty of Technology, Linnaeus University, Växjö, Sweden

The construction sector is a major consumer of natural resources and a significant contributor to environmental degradation through resource depletion, waste generation, and greenhouse gas emissions. As the industry shifts toward sustainable and circular practices, improving material traceability and resource efficiency becomes crucial. Digital Tracking Technologies (DTT) offer a promising solution, yet their adoption in the precast concrete sector remains limited, necessitating an in-depth examination of influencing factors. This study explores DTT adoption in Sweden's precast concrete industry, addressing a critical gap in understanding how digital technologies support Circular Economy (CE) objectives. By identifying key barriers and drivers, the research provides a holistic view of the challenges and opportunities shaping this transition. A mixed-methods approach was employed, integrating bibliometric analysis, social media sentiment analysis, and stakeholder interviews. Data analysis was conducted using VOS Viewer, Orange Data Miner, and NVivo, while findings were interpreted through a sociotechnical lens. The study utilizes PESTLE and SWOT analyses within the Multi-Level Perspective (MLP) framework to examine dynamics at the niche, regime, and landscape levels. Findings highlight systemic barriers such as high investment costs, fragmented infrastructure, limited knowledge, and cultural resistance. However, opportunities exist, including enhanced operational efficiency, improved information dissemination, CE alignment, supportive policies, and targeted incentives. Innovations like RFID tags, QR codes, and Digital Product Passports can enhance material traceability and circularity. Given regime inertia and landscape pressures, cohesive policy interventions and industry collaboration are essential. This study provides actionable insights for industry stakeholders, policymakers, and sustainability advocates to advance digital innovation and sustainable construction practices.

KEYWORDS

barriers, circular economy, concrete sector, digital tracking technology, drivers, multilevel perspective

1 Introduction

The construction sector is one of the largest global consumers of natural resources, significantly contributing to environmental degradation through resource depletion, waste generation, and greenhouse gas emissions (Bonoli et al., 2021; Fei et al., 2021). This is rooted in the linear economy (LE) model of the sector, which operates on a *"take-make-dispose"* approach (Ellen MacArthur Foundation, 2015; Oberle et al., 2019). Thus, transitioning

to a Circular Economy (CE) framework, which emphasizes material reuse, repair, and recycling, has become imperative in mitigating these adverse impacts (Ellen MacArthur Foundation, n. d). This shift not only promotes sustainable resource use but also minimizes environmental impact, improves waste management practices, and advances several United Nations Sustainable Development Goals (SDGs) (Ellen MacArthur Foundation, 2015; Eurostat, 2023; European Environmental Agency, 2024).

Sweden, a global leader in climate neutrality, has made considerable strides in reducing greenhouse gas emissions and fostering sustainable urban development (RE:Source, 2022). However, the construction sector continues to pose major environmental challenges. Accounting for 125 million tonnes or ca 47% of the country's material consumption and generating approximately 14 million tonnes of waste annually, this sector is the second largest waste producer after the mining industry (Zakrisson, 2018; RE:Source, 2022). This highlights the urgent need for the Swedish construction sector to shift from a linear to a circular model to help close the nation's Circularity Gap and support its ambitious sustainability targets (RE:Source, 2022).

Concrete is a dominant construction material, which is valued for its durability and strength (Babor et al., 2009). However, its highly resource-intensive production process contributes significantly to CO_2 emissions. Cement alone is responsible for nearly 8% of global CO_2 emissions (Andrew, 2018; Habert et al., 2020; Statista, 2023). Consequently, these environmental impacts pressure the industry to explore lower-carbon, traceability, and circularity-focused alternatives (Habert et al., 2020; Nilimaa, 2023). One such alternative is the pre-casting of concrete in controlled environments, which aims to mitigate some of the environmental challenges (Wong and Loo, 2022). However, it also presents unique operational, logistical, and storage challenges as it heavily relies on modular components that require precise tracking from production to installation (Wong and Loo, 2022; Liu and Zainul Abidin, 2024).

The shift toward circularity in construction is emphasized in the European Parliament's harmonized rules for marketing construction products (European Union, 2023). This regulation stresses the importance of sustainable practices, aiming to reduce environmental impact by promoting resource efficiency and lifecycle management of construction materials. Given this, digital innovation has emerged as a crucial enabler of transformation within the construction sector (UNEP, 2023). Initial tools like Computer-Aided Design (CAD) laid the groundwork for this digital shift (Javaid et al., 2022), which has evolved to include advanced systems such as Building Information Modelling (BIM) (Biswas et al., 2024), the Internet of Things (IoT) (Giovanardi et al., 2022; Shishehgarkhan et al., 2024), and real-time tracking technologies that improve material flow and asset management (Zhao et al., 2017; Daniotti et al., 2022). A key advancement within the BIM framework is the integration of digital twins, which streamline construction management and maintenance processes, ultimately enhancing operational efficiency (Nguyen and Adhikari, 2023; Jiang et al., 2024). This evolution aligns with the Green Deal and the Circular Economy Action Plan (CEAP) of the European Union. These initiatives not only advocate for resource efficiency and emissions reduction, but also introduce the Digital Product Passport (DPP) as a transformative mechanism to ensure traceability, transparency, and material reuse across industries (European Commission, 2020; 2021).

Recent studies have highlighted the emergence of Digital Tracking Technologies (DTT) to facilitate circular construction practices by enhancing traceability, transparency, and compliance, while enabling data-driven decision-making (Camodeca and Almici, 2021; Del Rio et al., 2021; Davari et al., 2023). Despite the potential benefit of these technologies, their broader implementation globally has been slow (ECSO, 2021; Wuni, 2022; Dervishaj et al., 2023; Davari et al., 2023; Sepasgozar et al., 2023; Giovanardi, 2024) due to technical constraints, financial limitations, cultural resistance, and a shortage of skilled expertise (Gerhardsson et al., 2020; Lobo et al., 2021; Kaufman et al., 2022; Thirumal et al., 2024). Sweden has an advanced digital infrastructure to achieve various sustainable development goals (Business Sweden, 2022; RE:Source, 2022; UNEP, 2023), but still the adoption of DTT in the construction sector is slow.

This study examines the adoption of DTT in Sweden's precast concrete sector using a mixed-method approach to examine both macro-environmental factors and stakeholder perspectives. By combining bibliometric analysis, social media sentiment analysis, and expert interviews to capture the trends, barriers, and opportunities related to DTT adoption, the study offers a holistic understanding of the dynamics driving DTT adoption in the context of CE practices. To further contextualize the findings, the study applies the socio-technical transition (STT) framework. A socio-technical system involving diverse stakeholders, institutions, and technologies faces unique challenges in adapting to change due to the long lifespan of buildings and decentralized decision-making (Gerhardsson et al., 2020; Vogel, 2020; Kaufman et al., 2022). Digital transformation in this context requires more than mere technological adoption; it demands a culture of continuous change, exploration, experimentation, and the strategic use of emerging tools to drive sustainability (Herbert, 2017). Social acceptance of digital tools and shifting attitudes toward technology, influence user interactions and propel the evolution of STT (Graf-Vlachy et al., 2018).

The Multi-Level Perspective (MLP) is a widely used framework in sociotechnical transition studies, which illustrates the dynamic relationships between niche innovations (e.g., DTT), stable societal practices and regulatory structures (regimes), and the contextual factors (landscapes) (Geels, 2002; 2011). These transitions involve changes in technology, user practices, infrastructure, regulatory frameworks, and symbolic meanings (Geels, 2002; Schot and Geels, 2008). Previous research has applied MLP to sustainability transitions in energy, low-carbon initiatives, and governance (Li, 2020; Kanger, 2021; Jayaraj et al., 2024). However, studies focusing on transitions in the construction industry, particularly in the precast concrete sector, remain limited. While several frameworks and strategies exist to guide digitalization and CE transitions within the EU, including the Digital Product Passport (DPP) and various BIM-based solutions, a comprehensive analytical framework integrating both PESTLE and SWOT analyses in the Swedish precast concrete sector has yet to be established. Most existing studies on digitalization of the construction sector focus on the technical aspects of individual technologies, such as BIM, IoT, RFID, and blockchain, while overlooking the broader socio-technical dynamics that consider macro-environmental influences and stakeholder interactions (Teisserenc and Sepasgozar, 2021; Giovanardi et al., 2022; Dervishaj et al., 2023; Dervishaj and

Gudmundsson, 2024; Giovanardi, 2024). This narrow focus creates a significant knowledge gap in the role of socio-technical interactions in the transition to digitalization of the construction sector, especially the precast concrete sector. To address these research gaps, this study applies the PESTLE-SWOT approach in conjunction with the MLP framework as a foundation to provide a holistic understanding of the trends, opinions, and dynamics influencing the adoption of DTT in the Swedish precast concrete sector.

2 Materials and methods

The study adopted a deductive approach, which emphasizes the progression from theoretical foundations to empirical validation (Bryman and Bell, 2015). The MLP framework served as the theoretical framework, providing a structured approach to examining socio-technical interactions across niche, regime, and landscape levels. The dynamics of the sociotechnical interactions were analyzed through PESTLE (Political, Economic, Social, Technological, Legal, and Environmental) and SWOT (Strengths, Weaknesses, Opportunities, and Threats) frameworks. These methods have been previously applied in studies, such as assessing sustainability in adaptive reuse projects (Vardopoulos et al., 2021) and analyzing strategies for the deep renovation of detached housing markets (Mainali et al., 2021). PESTLE identified macro-environmental factors, such as regulatory, technological, and societal pressures, while SWOT examined internal and external dynamics (Gurel and Tat, 2017; Rastogi and Trivedi, 2022). This integrated, multi-layered methodology enabled a robust examination of systemic drivers and barriers within the Swedish precast concrete sector, while also capturing broader trends and challenges across the construction industry.

A mixed-method approach, integrating both quantitative and qualitative data, facilitates a comprehensive analysis of a research problem (Creswell, 2014). This method was organized into three steps: bibliometric analysis identified research trends and clusters, social media analysis provided insights into industry attitudes and public discourse, while expert interviews offered in-depth perspectives from key stakeholders.

2.1 Bibliometric analysis

The bibliometric analysis aimed to uncover existing barriers and opportunities linked to the adoption of Digital Technologies (DT) by examining academic abstracts sourced from Scopus and Web of Science databases. The following search string was utilized to capture relevant publications: "("circular economy" AND (digit* OR "industry 4.0 OR passports) AND (barrier OR challenge OR problems OR limit* OR opportunities OR benefits OR possibilities) AND (construction* OR concrete* OR "built environment"))." From the initial search, 168 articles were identified in Scopus, and 83 in Web of Science. These results were refined based on publication year (2010–2024) and relevance to identifying barriers and opportunities in adopting DT for circularity in the construction sector. After applying this criteria, 31 articles from Scopus and 10 from Web of Science were selected. The dataset was further refined by merging both sources and removing 15 duplicates, resulting in a core set of 26 unique articles. To ensure a focus on peer-reviewed academic and industry-relevant research, only journal papers and conference proceedings were included in the final selection. A snowballing technique by reviewing references from the selected articles was applied to expand the dataset, which led to the inclusion of 31 additional sources, resulting in a final dataset of 57 articles. These were then exported in RIS format and analyzed using a VOS viewer to generate a keyword co-occurrence map. This map provided visual insights into key themes and clusters, facilitating a nuanced understanding of the interaction between DT and CE in construction. Each identified cluster was interpreted and mapped to the corresponding MLP level in the STT, offering a foundational insight into trends, drivers, and barriers surrounding DT adoption in the broader construction industry.

2.2 Social media analysis

Social media posts from Facebook, X (formerly Twitter), and LinkedIn were scraped to assess sentiments toward DT adoption for CE in the construction industry. The majority of the authors of these posts were professionals and industry practitioners, all within the construction industry, with only a minimal number of contributions from academic researchers. Similar to Tsironis et al. (2024), who scraped LinkedIn posts to identify common CE practices, this study collected posts using targeted search phrases, such as *"Digital technologies + Circular Economy + Construction," "Digitalisering + Byggbranschen + Spårbarhet," "Digitalisering Byggbranschen för Cirkulär ekonomi," and "#Digitalisering #Byggbranschen."* Orange Data Miner, an open-source software created for interactive data visualization, machine learning, and data analysis (Demšar and Zupan, 2013; Ljubljana, no date) was used for the analysis. While its application for data analysis in engineering studies remains scarce, Orange has been widely utilized in computing, geoinformatics, and social sciences (Tiwari et al., 2023; Dobesova, 2024). Following a manual data cleaning, a structured text analysis workflow (illustrated in Figure 1) was applied to analyze a total of 241 social media posts. The process began with importing the text dataset into the corpus module, where the raw text was preprocessed to remove inconsistencies and improve readability. Next, the preprocessed text was passed into the Sentiment Analysis module, which evaluated the emotional tone of the posts and categorized them as positive, neutral, or negative (Dang and Ahmad, 2015; Tang and Feng, 2019; *NLTK Sentiment Analysis Tutorial: Text Mining & Analysis in Python*, no date). The model assessed the text polarity using a range from −1.0 (negative) to 1.0 (positive). To visualize key patterns, the data was further analyzed using multiple modules in the Data Miner workflow:

- A Word Cloud (Figure 3) was generated to highlight the most frequently occurring terms.
- A Sentiment heat map (Figure 4) provided a structured representation of sentiment distribution across different categories.
- A Data Table enabled further inspection of textual data, allowing for group-based and categorical analysis. While the Corpus Viewer facilitated manual exploration of the processed text data.

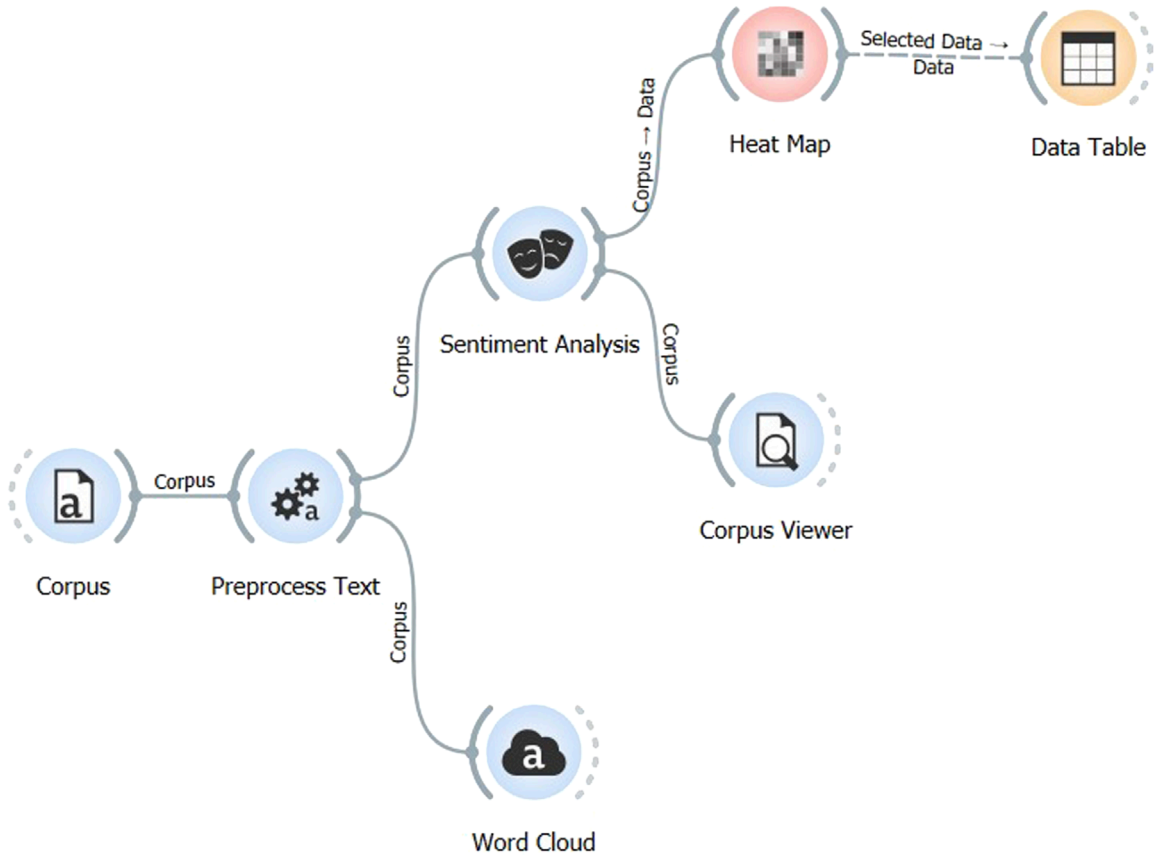


FIGURE 1
Workflow diagram illustrating the text analytics process used for social media analysis with the Orange Data Mining software.

These visualizations helped capture sentiment trends, identify discussion themes, and highlight key topics relevant to DT adoption for circularity in construction. This analysis also supported the MLP framework, linking positive sentiments to potential opportunities or drivers and negative sentiments to perceived barriers in the adoption of DTs.

2.3 Stakeholder interviews

Semi-structured interview has flexibility to explore a specific topic in detail (Adams, 2015), while purposive sampling is appropriate to obtain in-depth insights from participants with relevant expertise, ensuring a comprehensive understanding of the research problem (Bryman and Bell, 2015). Therefore, a purposive sample of 10 participants was chosen, following the guidance of Galvin (2015), who recommends a sample size of six to nine participants for key reliable results. Participants were chosen for their expertise in DT, DTT adoption, and CE practices in Sweden's precast concrete sector. This included heads of three precast concrete manufacturing companies, three technology providers, one architect, one BIM expert, and two CE and digitalization experts. All participants were briefed on research objectives, data handling, and confidentiality measures. This approach ensured the collection of rich, relevant data that complements broader analyses. A thematic analysis of the interview data was conducted using NVivo software. This step involved systematically coding the responses, identifying key themes related to the barriers and drivers of DTT adoption, and interpreting these themes in the context of CE practices. These coded themes were exported to the Python programming IDE to generate a heatmap. The analysis followed the six-phase process as outlined by Braun and Clarke (2012), ensuring a rigorous approach. The thematic analysis specifically focused on the PESTLE factors influencing DTT adoption in the precast concrete sector. This process was done to critically understand the external and internal sociotechnical dynamics influencing the adoption.

2.4 Data triangulation and merging strategy

To ensure the validity and reliability of the findings, data and methodological triangulation as established by Jick, (1979) were applied. This process integrated insights from all three steps to identify key themes (i.e., the barriers and drivers), and effectively mapped them across PESTLE, SWOT, and MLP frameworks. This triangulated approach provided a comprehensive, multi-layered understanding of the research problem, supporting well-grounded conclusions and actionable recommendations for advancing DTT adoption and CE practices.

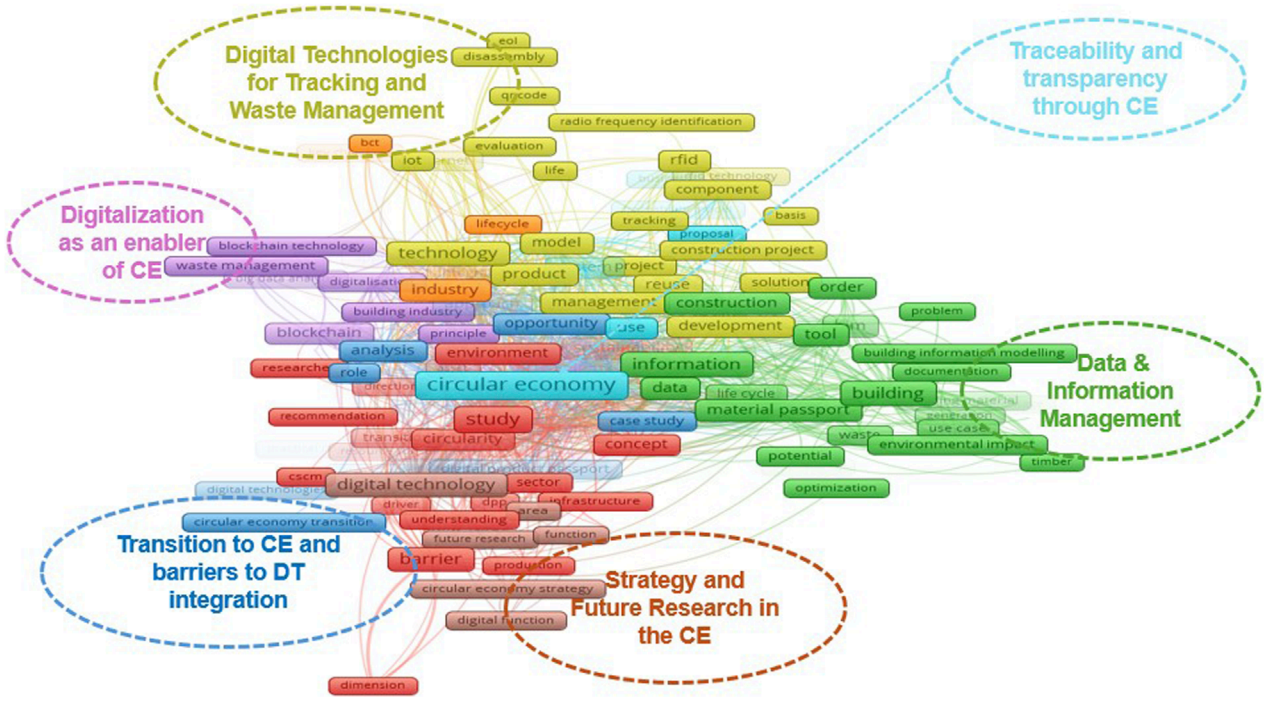


FIGURE 2
A keyword co-occurrence map for the analyzed 57 Academic literature.

3 Results

This section presents and discusses the findings from the multi-method analysis from bibliometric analysis (keyword co-occurrence map), social media analysis, and interview insights.

3.1 Key themes from bibliometric analysis

The bibliometric analysis with the aid of VOS viewer generated a keyword co-occurrence map (Figure 2), which provides a structured overview of interconnected themes related to the barriers and opportunities in adopting DT for CE in the construction sector. Given the limited literature specifically addressing DTT in this area, the analysis was expanded to include broader discussions on DT within the construction sector. The analysis identified six interconnected clusters; each assigned a title reflecting their central theme (as illustrated by circles in the figure). These themes are critically analyzed below in relation to the research objective and interpreted through the different MLP levels. The interconnectedness of these clusters reflects similarities in the work of the corresponding authors and the systemic nature of challenges and solutions, emphasizing the need for integrated approaches to address CE objectives.

3.1.1 Digitalization as an enabler of circular economy (CE)

This cluster highlights keywords such as *"digitalisation," "waste management," "building industry,"* and *"blockchain technology."* A study by Rodrigo et al. (2023) discusses how these DTs streamline resource optimization by addressing inefficiencies, while Çetin et al. (2022) identify DTs like digital twins and scanning technologies as key drivers for sustainable construction practices. Furthermore, Teisserenc and Sepasgozar (2021) emphasizes the importance of digitalization in material tracking, waste management, precise data management, and lifecycle analysis. Few studies by Trevisan et al. (2023) and Munaro and Tavares (2023) illustrate how DTs overcome systemic barriers such as inefficiencies in construction workflows, while Honic et al. (2019) argue that integrated digital solutions foster stakeholder collaboration and advance broader CE objectives. Collectively, these tools contribute to landscape, regime, and niche-level transformations. Digitalization is both a transformative force and an integral part of an evolving industry, reshaping norms, opening new avenues for CE, and supporting sustainable goals within the prefab concrete and construction sector. However, systemic incentives remain necessary to bridge the gap between potential and actual implementation (Munaro and Tavares, 2021).

3.1.2 Data and information management

Effective data management is necessary for successfully adopting DT in construction (Honic et al., 2019; Bellini and Bang, 2022). Keywords like *"data," "information," "BIM,"* and *"material passport"* emphasize the criticality of structured data systems for enabling CE (Kedir et al., 2021; Q; Liu et al., 2022). Accurate tracking of materials from production to end-of-life and material passports enhances traceability, reusability, and recyclability of construction materials (Kedir et al., 2021; Munaro and Tavares, 2021; Mêda et al., 2023).

However, the fragmented integration of such tools (Honic et al., 2019) as well as data interoperability issues and inconsistent data-sharing practices (Bellini and Bang, 2022) limit cross-stakeholder collaboration, which hinders the broader adoption of DT and seamless CE practices. Reflecting the regime and niche level of the MLP, this cluster underscores the need for industry-wide standards and consistent practices to unlock the full potential of structured data systems for CE.

3.1.3 Transition to circular economy (CE) and barriers to digital technology integration

The transition from linear to circular business models represents a paradigm shift, requiring the adoption of innovative business models, restructured workflows, and enhanced digital infrastructures. However, significant challenges including data management, limited access to reliable data (Bellini and Bang, 2022), and integrating DT into construction practices persist (Munaro and Tavares, 2021). Keywords such as *"transition," "circular economy," "digital technologies,"* and *"role"* in the transition cluster (dark blue) along with *"infrastructure," "understanding," "production,"* and *"sector"* in the challenges' cluster (red) highlight the reconfiguration of norms and processes alongside barriers such as infrastructure deficits, high costs, and interoperability issues. Overcoming the challenges of a shift from the entrenched linear models to circular models requires institutional flexibility and governance frameworks to incentivize collaboration (Munaro and Tavares, 2021). The lack of cohesive policy alignment across regions often leads to piecemeal efforts, and therefore there is a need for an integrated policy environment to overcome such regulatory barriers and drive systemic change (Adriana Hofmann Trevisan et al., 2023). While Liu et al. (2022) proposed a framework identifying critical digital functions for circularity, achieving scalability remains challenging. This cluster highlights that resistance to DT adoption stems from entrenched practices and structures at the regime, requiring significant investments in infrastructure, training, and policy integration at the landscape level to overcome these barriers.

3.1.4 Digital technologies for tracking and waste management

Focusing on material lifecycle tracking and waste management, the yellow and orange clusters align closely with the core principles of CE, with keywords like "*QR codes," "IoT,"* and *"RFID."* Dervishaj et al. (2023) explored the integration of tracking technologies with BIM to support reuse strategies for prefabricated concrete components. Additionally, Jiang et al. (2023) emphasize blockchain's potential for creating secure and traceable material flows, though technological complexity and cost barriers hinder widespread adoption. Although Uçar et al. (2020) identified IoT-enabled systems that facilitate efficient waste segregation and recycling, yet these innovations are not widely adopted across the industry for CE practices and face regulatory and industry uptake challenges. Coordinated efforts are essential to scale these technologies and integrate them into mainstream waste management practices, as they hold significant potential for supporting circular practices by improving material traceability and accountability.

3.1.5 Traceability and transparency for circular economy

With circular economy as the central focus, terms such as *"traceability," "transparency,"* and *"sustainability"* indicate the importance of robust frameworks to verify material flows and ensure accountability for achieving CE across the entire supply chain. Studies such as Rahel Kebede et al. (2023) present a framework for implementing Digital Product Passports (DPP) for improved and detailed data sharing and tracking of material flows, ultimately for transparency purposes. However, Katenbayeva et al. (2016) highlight that achieving transparency requires addressing significant data accessibility and privacy concerns. This cluster can be positioned between niche innovations and regime levels as they are emerging innovations in the digital space. This reveals a paradox: while traceability tools are critical for CE, their effectiveness depends on achieving industry-wide data-sharing agreements, often constrained by competitive industry and regulatory barriers.

3.1.6 Strategy and future research in the circular economy

This cluster highlights the importance of strategic initiatives and future research to address gaps in CE implementation. Keywords such as *"circular economy strategy," " future research,"* and *"digital function"* reflect the ongoing quest for clear pathways and actionable solutions. Benachio et al. (2020) call for localized material banks and digital inventories to support CE practices. Munaro and Tavares (2021) emphasize aligning policy and research agendas to facilitate the transition to circular models. Furthermore, Teisserenc and Sepasgozar (2021) note that cross-disciplinary collaboration is often lacking, limiting the scalability of strategic efforts. This cluster stresses the need for bridging research and practice to ensure that strategic initiatives result in tangible outcomes, aligning with the transition to circularity.

3.2 Social media analysis

3.2.1 Key themes identified from social media discussions

The word cloud visualization, as shown in Figure 3, highlights key topics frequently discussed concerning DT within the construction industry, with prominent words like *"circular economy," "sustainability," "digital technologies,"* and *"construction"* indicating a clear focus on aligning DT with circular economy principles. These terms reflect broader societal pressures and shifts toward sustainability and resource conservation, which drive DT adoption. Meanwhile, words like *"materials," "management," "efficiency,"* and *"energy"* point to specific operational areas where DT could enhance resource efficiency and traceability, aligning with established industry practices. Additionally, terms such as *"innovation," "environmental," "industry," and "future"* suggest that stakeholders see DT as a forward-looking approach to modernizing the construction sector. The presence of these terms highlights both niche innovations and the integration of DT within existing industry norms. However, the appearance of words like *"challenges"* and *"critical"* indicates that, while there is substantial interest in digital transformation, significant perceived obstacles remain, reflecting



FIGURE 3
A word cloud visualizing the key themes identified from social media analysis.

the barriers that need to be overcome for DT adoption to gain traction across the broader construction industry.

3.2.2 Sentiment distribution on social media: Insights from heatmap and bar chart

The heatmap in Figure 4 visualizes the intensity and distribution of sentiments across social media posts on DT adoption in the construction industry. The first calibrated line aggregates all posts, while the first three labeled rows break down the presence of positive (row 1), neutral (row 2), and negative (row 3) sentiments within individual posts. The final row, labeled "compound," represents the overall sentiment of each post, ranging from strongly negative (−0.8053, blue) to neutral (0, white) and highly positive (1, red). This structure allows for a nuanced understanding of how sentiment varies across discussions, with the compound score providing a conclusive measure of each post's prevailing tone. These insights are further supported by specific excerpts from social media posts, which illustrate the sentiment categories in more detail:

3.2.2.1 Positive sentiment (151 occurrences)

Most posts reflect a positive sentiment, indicating optimism about the potential of DT generally to drive circularity. For instance, one post states:

> "The potential of Building Information Modelling (BIM) and digital twin technologies in the construction industry highlights the benefits of using BIM and digital twins, such as improved collaboration, increased efficiency, enhanced stakeholder engagement, and sustainability through the application of circular economy principles."

This post is categorized as positive due to its emphasis on the benefits of DT, such as improved collaboration, increased efficiency, and sustainability. Keywords like "*improved*," "*increased*," "*enhanced*," and "*sustainability*" contribute to the positive sentiment score, illustrating that social media users see DT as an enabler of better practices and sustainable outcomes in the industry.

3.2.2.2 Neutral sentiment (77 occurrences)

A considerable portion of posts falls into the neutral category, suggesting that while there is interest in DT, many discussions are still exploratory or cautious. A representative excerpt includes:

> "I have been guilty of some sweeping reasoning about how the construction industry is far behind the manufacturing industry in digitalization. But Y has a more nuanced view, and I learned a lot during our conversation. Go in and listen."

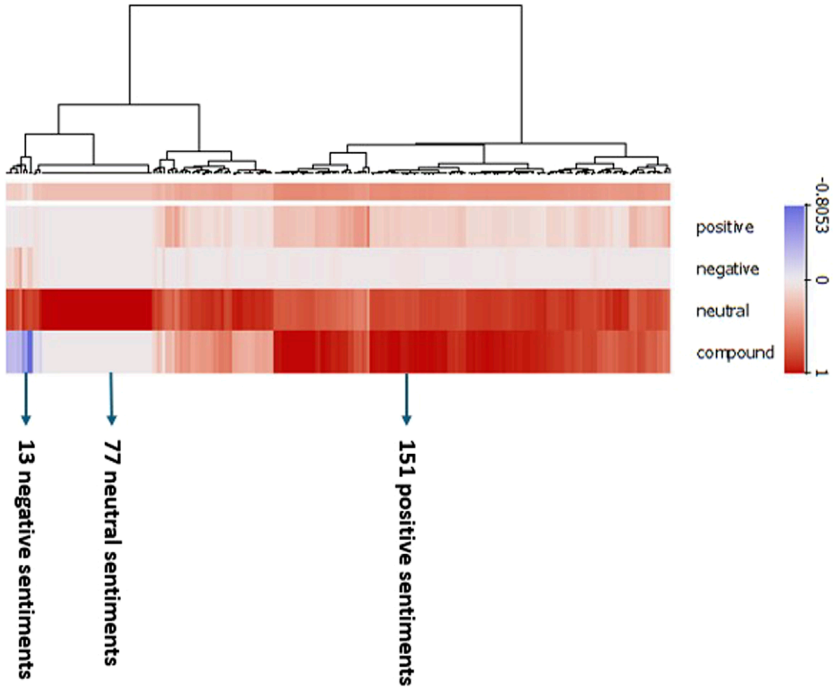


FIGURE 4
Heatmap of sentiment distribution across 241 social media posts.

This post reflects a neutral sentiment because it does not explicitly endorse or criticize DT. Instead, it suggests an evolving understanding and a balanced view. The absence of strong positive or negative words aligns with the sentiment score, indicating an objective or observational tone. This neutrality could suggest that many stakeholders are still in the information-gathering phase regarding DTT.

3.2.2.3 Negative sentiment (13 occurrences)

A smaller proportion of posts express negative sentiment, highlighting specific challenges or criticisms related to DT adoption. An example of such a post reads:

> "We identify challenges in retrofitting digital technologies to existing structures and emphasize the need for a unified digital transformation strategy and increased cross-industry collaboration."

This post is categorized as negative due to phrases like "*identify challenges*" and *"need for a unified strategy,"* which underscore the difficulties and strategic gaps in adopting DTT. Words like *"challenges"* and *"need"* contribute to negative sentiment, reflecting concerns over integration difficulties and the lack of cohesive digital strategies.

3.3 Stakeholder interview

The content analysis of the stakeholder interview is presented in the form of a heatmap in Figure 5. It visualizes the frequency of mentions (from 10 stakeholders) for both barriers and drivers of DTT within the context of circularity in Sweden's prefabricated concrete sector. A deeper colour represents a higher frequency of responses. Each barrier and driver are categorized according to the PESTLE framework, allowing for a structured analysis of the factors influencing DTT adoption. These PESTLE categories will be further examined to assess the implications of each factor, with insights triangulated against findings from previous analyses to create a comprehensive understanding of the sector's challenges and enablers for achieving a CE. Several key observations emerge from the heatmap analysis:

3.3.1 High-Frequency Barriers

Among the most frequently mentioned barriers are limited stakeholder awareness (9 mentions), reliance on established traditional practices (8 mentions), difficulties in integrating DTT (8 mentions), high investment costs (9 mentions), and lack of IT competence (6 mentions). These factors suggest that, despite recognition of DTT's potential, resistance from conventional structures and limited awareness serve as substantial impediments to progress.

3.3.2 High-Frequency Drivers

Conversely, a *positive attitude towards sustainability* (10 mentions), the *data-driven nature of DTT, knowledge/awareness,* and *rising demand for sustainable practices* stand out as major drivers. These opinions reflect a growing shift toward sustainable practices, highlighting a crucial motivational factor for adopting digital solutions within the sector.

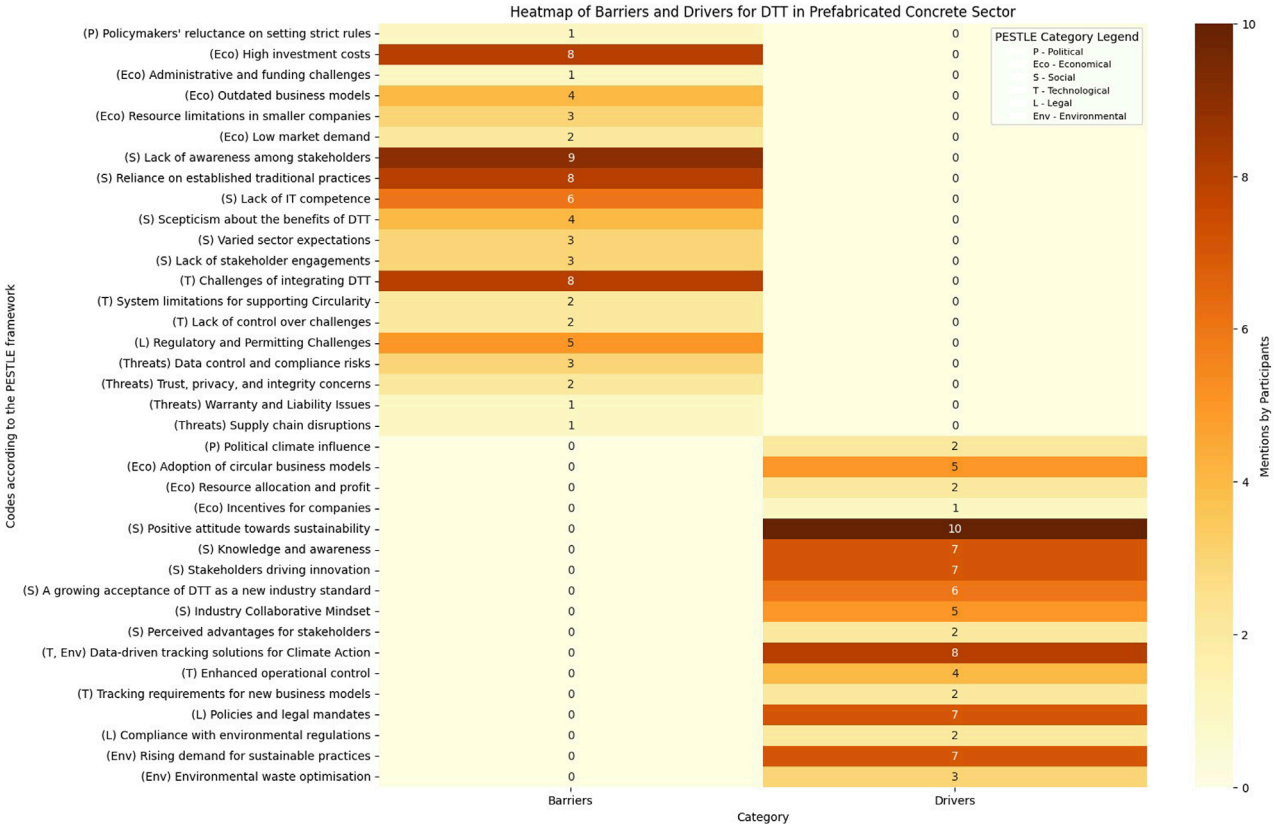


FIGURE 5
Heatmap of barriers and drivers identified through thematic analysis of stakeholder interviews.

4 Discussion

4.1 PESTLE analysis

4.1.1 Political and legal factors

The regulatory environment and policy frameworks play a critical role in driving the adoption of DTT. Interviews revealed that policymakers significantly shape this landscape, with initiatives like the EU Green Deal promoting innovation and enforcing environmental compliance standards. However, regulatory and permitting challenges remain formidable barriers to DTT adoption. Given DTT's intersection with compliance-heavy sectors, such as the precast concrete industry, companies face stringent regulatory requirements that complicate implementation. One participant noted, "*Meeting regulatory requirements often complicates the DTT process; however, we recognize that compliance is also a motivator in aligning with sustainable practices.*" This dual role of regulations indicates their complexity: while they pose hurdles, they also create structured frameworks encouraging sustainable practices. Legal compliance, therefore, serves as both a barrier and a driver, pushing companies toward responsible adoption of DTT. Additional challenges and threats include trust and privacy concerns, data control issues, warranty and liability uncertainties, and policymakers' hesitance to enforce stricter regulations. These factors, compounded by unclear guidance on material tracking, hinder progress. Studies, such as Thirumal et al. (2024) and Trevisan et al. (2023) identified from the bibliometric analysis, have highlighted a lack of cohesive policy alignment and inadequate management support as significant obstacles to adopting innovations for CE, such as DTT. These insights highlight the need for clearer regulations and streamlined permitting processes, as well as the need for institutional and governmental incentives to drive systemic change (Munaro and Tavares, 2021). Social media discussions also reflect the political dynamics at play. Terms such as *"strategy," "Sweden,"* and *"European,"* which appeared prominently in word cloud analyses, emphasize the influence of national and regional strategies on CE initiatives. While regulations and building codes like the EU Energy Performance Directive (EPBD), Swedish Building Code (BBR), and the Construction Products Regulation (CPR) encourage resource management, they fall short of mandating digital material tracking. This regulatory limitation constrains the full integration of DTT in the sector.

4.1.2 Economic factors

Economic and market-related challenges are key barriers to implementing DTT, especially for smaller companies. Interviews highlighted the significant time and financial investments required to integrate DTT into existing systems and circular practices. All the respondents (9/10) except one sustainability and digitalization

expert emphasized the high costs, and time demands of tracking building materials, with one participant noting: "*Mostly money and time, because just tracking and tracing the material is costly.*" Smaller companies face additional obstacles, including limited resources, insufficient infrastructure, and capital, which hinder their ability to adopt advanced technologies and comply with regulations. Additionally, the terms *"economic"* and *"infrastructure"* on social media word cloud and bibliometric keyword map, further emphasize the infrastructural challenges in adopting DT generally. These challenges may arise from the high costs of upgrading existing systems, limited technological readiness, and fragmented digital infrastructure across companies in the construction sector. Despite acknowledging the investment challenge, the Software/Technology Providers expressed optimism about the economic potential of DTT. One provider remarked, "*Investment in digital tools enhances operational efficiency and can yield significant Return on Investment (ROI) when applied correctly.*" The respondents also view digital transformation as essential for maintaining competitiveness, as clients increasingly expect digitally enabled services. These findings align with previous studies that identify financial constraints as critical barriers to adopting advanced digital solutions (Lobo et al., 2021; Dervishaj et al., 2023). Similarly, insights from the bibliometric analysis (Munaro and Tavares, 2021) highlight the importance of adequate infrastructural investment to facilitate the adoption of DT. Interestingly, despite widespread concerns about costs, only one respondent mentioned incentives as a driver for long-term adoption. This raises important questions about whether there are implicit regulatory restrictions limiting access to public funding such as grants, regional funding programs and tax rebates for green innovations or if the industry is generally unaware of the available financial support mechanisms. These types of financial incentives might alleviate the burden on individual companies and encourage consensus and standardization in how such work is undertaken.

4.1.3 Social and cultural factors

Industry attitudes toward sustainability and technology adoption vary widely, shaped by regime-level cultural and organizational norms. A positive attitude toward sustainability emerged as a key driver for DTT implementation. Social media analyses further reinforced this strength, revealing widespread enthusiasm and approval for CE initiatives. However, the Swedish construction industry's complexity, marked by diverse projects, processes, and actors creates unique challenges. As (Bygballe and Ingemansson, 2014; Gerhardsson et al., 2020; Vogel, 2020) note, that the heterogeneity of the sector necessitates multi-level interactions among developers, contractors, building owners, and consultants, each with varying levels of digital maturity. Traditionally, the industry is known to be resistant to change and sceptical toward innovation (Akintoye et al., 2012; Gerhardsson et al., 2020). A notable knowledge gap within the industry further exacerbates these challenges, as stakeholders often lack a clear understanding of DTT's functionality, benefits, and implementation pathways. This gap is particularly pronounced among traditional actors such as precast concrete manufacturers, who perceive DTT as disruptive. One concrete manufacturer observed, "*There's a strong reliance on traditional methods, and many stakeholders are not convinced of the value digital systems bring.*" Such views highlight not only entrenched practices but also the absence of comprehensive awareness and education about DTT's potential to transform operations and improve sustainability outcomes. On the other hand, Sustainability and Digitalization experts expressed optimism, noting growing acceptance of digital standards and a collaborative mindset among stakeholders. One expert stated, "*Stakeholders are beginning to see these tools as essential for achieving sustainability goals.*" This cultural divide shows the need for education, case studies, and collaborative approaches to bridge gaps and foster industry-wide acceptance of DTT. Themes of knowledge sharing and collaborative learning, as seen in word cloud analysis and bibliometric analysis (Bellini and Bang, 2022), and the interviews appeared as both drivers and barriers of adoption. This dynamic interplay of actors and attitudes illustrates the social complexities inherent in the transition toward digital and sustainable practices in the construction industry. These barriers further align with studies showing the construction sector as one of the least digitized industries, hampered by attitudinal and cultural resistance (ECSO, 2021; Munaro and Tavares, 2021; 2023).

4.1.4 Technological factors

The data-driven nature of DTT emerged as a key technological strength, enabling enhanced operational efficiency and circular practices. However, barriers identified in the bibliometric analysis (Honic et al., 2019; Teisserenc and Sepasgozar, 2021; Bellini and Bang, 2022) were also evident in the interview findings, particularly among concrete manufacturers. These challenges include difficulties with data integration and adapting to new technologies. One respondent remarked, "*Integrating digital systems into our traditional processes is challenging and disrupts our current operations.*" This reflects a hesitance to move away from established practices, highlighting the need for technologies that seamlessly integrate into existing workflows to minimize disruptions. In contrast, Software/Technology Providers highlighted the transformative potential of DTT. They emphasized its ability to enhance operational control, track materials, and enable innovative design solutions. As one provider stated, "*Digital systems allow us to monitor and optimize operations in real-time, which is a game-changer for our clients.*" The architect and AEC expert also echoed this, noting that new technologies enable innovative design solutions that can flexibly respond to evolving client needs. These suggest that while some stakeholders see DTT as disruptive, others perceive it as transformative, enhancing operational efficiency and flexibility. Existing studies further support these findings, emphasizing the importance of digital solutions like asset tracking and resource management in promoting CE practices (Zhao et al., 2017; Uçar et al., 2020; Daniotti et al., 2022; Thirumal et al., 2024). Additionally, the word cloud and bibliometric keyword co-occurrence map prominently featured terms such as *'digital,' 'technology,'* and *'innovation,'* alongside emerging technologies like *'RFID,' 'blockchain,' 'material passports,' 'QR codes,'* and *'IoT.'* These tools are recognized for their potential to transform conventional practices, enabling real-time monitoring, improved resource management, and the reuse of materials through DTT (Zhao et al., 2017; Camodeca and Almici, 2021; Del Rio et al., 2021; Kedir et al., 2021; Munaro and Tavares, 2021; Daniotti et al., 2022; Davari et al., 2023; Mêda et al., 2023).

4.1.5 Environmental factors

Environmental drivers, including the rising demand for sustainable practices and waste reduction, strongly support DTT adoption. Organizations increasingly view DTT as a critical tool for achieving sustainability goals amid societal and market pressures for greener practices. One participant emphasized, "*Sustainability is becoming a non-negotiable standard, and digital solutions are seen as a viable means to achieve it*." Findings of Camodeca and Almici (2021) and Del Rio et al. (2021) further highlight the environmental benefits of DTT, such as improved resource management, waste reduction, and material recycling, aligning with the 2030 Sustainable Development Goals (SDGs). The word cloud and keyword map reinforced this environmental focus, with terms like *"circular economy," "recycling,"* and *"sustainability"* dominating the industry, academic, and social media discourse. Digital solutions enable precise material tracking, quality control, and smarter production practices, fostering the transition to circular and sustainable approaches.

4.2 SWOT analysis

The PESTLE factors identified are grouped into Strengths (S), Weaknesses (W), Opportunities (O), and Threats (T) within the SWOT framework (Table 1), providing a comprehensive evaluation of the dynamics of the change process driving digitalization in the prefab concrete sector. A key insight from this analysis is the interaction between internal barriers (Weaknesses-W), external opportunities (O), and external barriers (Threats-T), which provides a clear pathway for stakeholders to address barriers. For example, high investment costs and resource limitations in smaller companies (W2 & W3) often hinder their ability to adopt innovative technologies. However, these challenges present opportunities for targeted incentives or government support (O3) to alleviate financial burdens. Similarly, regulatory and permitting challenges (T1) can be addressed by leveraging environmental policies like the EU Green Deal (S1), which may drive innovation.

To overcome these barriers, different industry stakeholders can take distinct approaches. Small concrete manufacturers, for instance, could benefit from government incentives (O3) and strategic collaborations (O8) to reduce high adoption costs and expand access to digital technologies. Larger manufacturers, with more resources, could lead industry-wide initiatives to integrate digital technologies and share knowledge across the value chain, supporting smaller companies in the process. Technology providers can capitalize on opportunities to improve stakeholder collaboration (O8), ensuring that digital solutions are adaptable to the diverse needs of the market and can be integrated more easily into existing processes. Meanwhile, policymakers can promote a supportive environment by enforcing regulations that incentivize digital adoption and circular practices while providing guidance to ease regulatory challenges (T1) and support the adoption of new business models. Addressing the complex network of actors (W11) and the need for significant mindset shifts (W12), overcoming resistance to change (T2) will require stronger stakeholder engagement and education initiatives (O6 & O8), focusing on building trust and creating a unified vision for the digital transformation of the sector. This multi-faceted approach will enable the industry to capitalize on opportunities while mitigating the challenges that hinder the adoption of digital technologies for circular economy practices.

4.3 MLP dynamics

The integrated PESTLE and SWOT analysis identified that the Swedish prefab concrete sector, like the broader construction industry, operates within a complex socio-technical system in which dynamics at the landscape, regime, and niche levels interact and shape the digital transformation trajectory for circularity (Vogel, 2020). This interaction illustrated in Figure 6 represents subcategories of the observed barriers and drivers from the study.

At the landscape level, overarching drivers such as global climate and sustainability goals, EU policies and regulations, and public and political pressures create a compelling need for change. Frameworks like the UN SDGs, the EU Green Deal, and Sweden's commitment to carbon neutrality are accelerating the shift towards circularity in the construction sector. These forces place significant pressure on the industry (regime) to adopt DT to meet sustainability targets, enhance resource efficiency, traceability, and transparency, and reduce emissions. However, external pressures (economic and environmental) revealed through themes in academic literature, social media discourse, and interviews, highlight disparities in preparedness among actors within the prefab concrete sector. While larger firms are often better positioned to respond proactively to regulatory incentives, smaller companies face considerable challenges in meeting these expectations, primarily due to limited resources and capacity.

At the regime level, various PESTLE factors interact simultaneously, shaping industry dynamics. Interviews revealed that a significant number of barriers remain, particularly the influence of established industry practices and the tension between entrenched norms and innovation. The precast concrete sector, in particular, exhibits notable stability due to its deeply rooted norms and the complex network of influential actors, as described by Geels (2002) and Vogel (2020). This stability, however, often translates into resistance to change, driven by a conservative mindset and ingrained practices. The study identified several critical weaknesses within the industry. Smaller companies face significant resource limitations, which hinder their ability to adopt new technologies. A major challenge lies in the lack of IT knowledge required for the implementation of DTT, compounded by scepticism regarding the tangible benefits DTT offers to companies. This scepticism is further worsened by the lack of technological readiness due to fragmented digital infrastructure and a heavy reliance on traditional norms. Additionally, low market demand for DTT-enabled products from customers further impedes the industry's digital transformation efforts. Tensions between traditional industry practitioners and technology developers highlight divergent perspectives on digitalization. While some stakeholders view digitalization as disruptive to established practices, others perceive it as a transformative force, paving the way for a more innovative and sustainable construction sector.

Although building codes and regulations such as the Construction Products Regulation (CPR) and Swedish Building Code (BBR) have yet to mandate digital tracking of construction

TABLE 1 Summary of PESTLE factors in the SWOT framework.

Strengths (S)	Weaknesses (W)
S1. Environmental policies like the EU Green Deal (L) S5. Compliance with environmental regulations (L) S2. Collaborative mindset among industry stakeholders (S) S3. Positive attitude and sentiment towards sustainability and technology adoption (S) S4. Data-driven tracking solutions enabling material and climate action tracking (T, Env) S6. Rising demand and enthusiasm for sustainable practices (Env)	W1. Political reluctance to set strict rules (P) W2. High investment costs (Eco) W3. Resource limitation in smaller companies (Eco) W4. Outdated business models limiting DTT integration (Eco) W5. Varied sector expectations and low market demand (Eco) W6. Lack of knowledge and awareness of DTT (S) W7. Scepticism about the benefits of DTT (S) W8. Reliance on established traditional practices (S) W9. Lack of stakeholder engagements (S) W10. Negative sentiments on adoption (S) W11. The complex and heterogeneous network of actors (S) W12. The need for a significant mindset shift (S) W13. Integration and compatibility issues (T) W14. The need for operational changes (T)
OPPORTUNITIES (O)	**THREATS (T)**
O1. Adaptive and supportive political climate (P) O2. Leadership and strategic vision supporting DTT adoption (P) O3. Targeted incentives for DTT adoption (Eco) O4. Adoption of circular business models (Eco) O5. Increased profitability and stakeholder advantages through DTT adoption (Eco) O6. Improved information dissemination to bridge knowledge gaps on DTT's benefits (S) O7. A growing acceptance of DTT as the new industry standard (S) O8. Increased stakeholder collaboration and partnerships (S) O9. Precise monitoring and material reuse through DTT adoption (T) O10. Technological and market advancements in DTT (T, Eco) O11. Technological innovation through DTT adoption (T) O12. The potential for Digital Product Passports (DPP) to drive material traceability and support regulatory compliance in the construction sector (T) O13. Enhanced operational efficiency and control (T) O14. Environmental optimisation and resource management (Env)	T1. Regulatory and permitting challenges (L) T2. Resistance to change and reliance on traditional established practices (S) T3. Risks around data control and compliance (T) T4. Market Volatility (Eco) T5. Concerns about trust, privacy, and data integrity(T) T6. Warranty and liability uncertainties (T)



FIGURE 6
A diagram illustrating the subcategories of observed barriers and drivers according to the three MLP levels.

materials, implementing Digital Products Passports (DPP) provides a pathway for future adoption. As these frameworks evolve to prioritize sustainability, they are fostering an environment that is both stable and adaptable, encouraging the gradual integration of DTT. This structured progression not only aligns with sustainability objectives but also enables the industry to incrementally reshape its standards, creating a foundation for a more circular and technologically advanced construction sector.

At the niche level, innovative technologies such as QR codes, BIM, material passports, and the IoT have emerged as powerful tools identified in the analysis for improving resource management and advancing the principles of the CE in the

Swedish prefab concrete sector. For example, the implementation of material passports provides detailed information about concrete components, including their composition, origin, and recycling potential (Dervishaj et al., 2023; Ali et al., 2024). This traceability reduces waste by ensuring that materials can be reused or recycled efficiently at the end of their lifecycle (Giovanardi et al., 2022). Findings from this study indicated that QR codes were primarily utilized for project management, logistics, and the storage of concrete components, rather than directly contributing to CE objectives. However, when viewed collectively, DT like BIM, RFID, and IoT are driving incremental advancements in data management and material traceability (Mallawaarachchi and Jayakodi, 2023; Shishehgarkhaneh et al., 2024). These innovations are gradually reshaping industry norms and laying the groundwork for a broader digital transformation. Despite their potential, these technologies remain in the early stages of development and face significant challenges. High upfront costs, integration complexities, and uncertain returns on investment present substantial barriers to widespread adoption. Furthermore, concerns persist regarding how data can be effectively retrieved and verified throughout the supply chain, as well as the security and privacy risks associated with tracking every construction component. These challenges pinpoint the experimental nature of these technologies and highlight the need for further refinement to unlock their full potential. Nonetheless, these innovations play a critical role in driving the initial stages of the Socio Technical Transition toward a CE in the prefab concrete sector. They provide a foundation for broader adoption and systemic change, enabling disruptive innovations to challenge entrenched norms at the regime level. As more actors within the industry adopt circular business models and leverage DT for sustainability, the sector moves closer to achieving systemic transformation.

5 Conclusion

This study aimed to examine the key barriers and drivers influencing the adoption of DTT for transitioning to a CE in Sweden's precast concrete sector. A triangulation approach was employed, cross-validating findings from bibliometric analysis, social media data, and interviews to ensure robustness and depth. The analysis of the drivers and barriers in the form of an integrated PESTLE and SWOT framework uncovered the systemic socio-technical dynamics shaping DTT adoption for circularity in Sweden's precast concrete sector. Key dynamics observed highlighted the interplay between stakeholders, innovation, and the challenges and opportunities inherent in systemic transformation. While landscape-level drivers, such as sustainability regulations, create pressure for change, regime-level inertia underlines the need for targeted interventions to overcome resistance. Meanwhile, niche innovations when supported by effective policies and collaborative initiatives could emerge as potential disruptors capable of reshaping entrenched practices and advancing the sector toward CE objectives. Despite its complexity, this transition represents a critical step in aligning the construction industry with broader sustainability goals.

To accelerate the adoption of DTT, a multifaceted approach involving policy initiatives, industry collaboration, and technological investments is essential. From a policy perspective, introducing targeted financial subsidies and tax incentives can significantly reduce the upfront costs of DTT implementation, particularly for small and medium-sized companies that may struggle with limited resources. Additionally, developing standardized data-sharing protocols is crucial to addressing issues of fragmentation and fostering greater interoperability among stakeholders, enabling smoother integration of DTT solutions across the supply chain. In terms of industry collaboration and social media discourse, organizing multi-stakeholder workshops and forums can help build trust and foster collaboration, ensuring that industry actors are aligned in their expectations and efforts. Such platforms also serve as opportunities to address resistance to change by highlighting the shared benefits of DTT adoption. Moreover, enhancing stakeholder engagement and awareness of the environmental and operational advantages of DTT is vital to bridge existing knowledge gaps. For example, initiatives like the Ellen MacArthur Foundation demonstrate how collaborative networks can connect industries, researchers, and policymakers to scale circular innovations. Such efforts illustrate the tangible impact of aligning strategic objectives with actionable outcomes in CE transitions. Technologically, investing in scalable and user-friendly digital tracking solutions is imperative to ensure that these technologies can be seamlessly integrated into existing workflows without disrupting daily operations. Furthermore, promoting the use of advanced tools like Digital Product Passports (DPP) in the construction sector can boost material traceability efforts and support compliance with regulatory standards, reinforcing the role of DTT in transitioning to a CE. Together, these strategic efforts can address existing barriers, leverage opportunities, and create an enabling environment for the widespread adoption of DTT in the precast concrete sector.

This study makes a theoretical contribution to the socio-technical systems and CE transition literature by adopting a triangulated approach that integrates the PESTLE, SWOT, and MLP frameworks. By applying these frameworks to bibliometric analysis, social media data, and interview findings, the study offers a replicable methodology for future research exploring similar themes. However, the research is subject to certain limitations. The scope of the study was constrained by the relatively small sample size of interview participants and its focus on the Swedish precast concrete sector. While the initial intention was to investigate this sector specifically, the responses often reflected the dynamics of the broader construction industry. As a result, some nuances unique to the precast sector may not have been fully captured.

To address these gaps, future research could expand the sample size to include a more diverse range of stakeholders relevant to DTT adoption. Similarly, engaging local policymakers and municipal regulators could reveal additional barriers, such as inconsistencies in regional compliance standards. Additionally, assessing digital readiness and benchmarking maturity levels across different sizes of precast concrete firms in Sweden and the EU could offer valuable insights. Comparative market gap analyses across regions would further help benchmark CE transition rates and identify region-specific barriers and drivers influencing DTT adoption.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

Ethical approval was not required for this study involving human participants, as per local and institutional guidelines. Verbal informed consent was obtained from all participants in accordance with GDPR. Prior to participation, each individual was thoroughly briefed on the research objectives, data handling procedures, and confidentiality measures. Written consent was not required under national legislation and institutional policy, provided that verbal consent was obtained and no sensitive personal data was collected.

Author contributions

MO: Conceptualization, Data curation, Formal Analysis, Investigation, Methodology, Software, Validation, Visualization, Writing – original draft, Writing – review and editing. RA: Supervision, Validation, Writing – review and editing, Conceptualization, Project administration. KM: Funding acquisition, Supervision, Validation, Writing – review and editing, Resources, Project administration.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This study is partially financed by the Interreg North Sea region and Region Kronoberg under the project "Circular Trust Building".

Acknowledgments

The authors acknowledge the interviewees for their time and insightful opinions. This research builds on a master's thesis (Shehu, 2024) originally published at Linnaeus University and an earlier, shorter version presented in the IEEE Conference Proceedings (Mahapatra et al., 2024).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Adams, W. C. (2015). "Conducting semi-structured interviews," in *Handbook of practical program evaluation* (John Wiley & Sons, Ltd), 492–505. doi:10.1002/9781119171386.ch19

Adriana Hofmann Trevisan, Lobo, A., Guzzo, D., Gomes, L. A. d. V., and Mascarenhas, J. (2023). Barriers to employing digital technologies for a circular economy: a multi-level perspective. *J. Environ. Manag.* 332, 117437. doi:10.1016/j.jenvman.2023.117437

Akintoye, A., Goulding, J., and Zawdie, G. (2012). "Construction innovation and process improvement," in *Construction innovation and Process improvement*. Editors A. Akintoye, J. Goulding, and G. Zawdie (Wiley-Blackwell), 3–17. doi:10.1002/9781118280294.ch1

Ali, V., Gebremariam, A. T., Di Maio, F., Meister, K., Koulaeian, T., and Rem, P. (2024). RFID-based material passport system in a recycled concrete circular chain. *J. Clean. Prod.* 442, 140973. doi:10.1016/j.jclepro.2024.140973

Andrew, R. M. (2018). Global CO_2 emissions from cement production. *Earth Syst. Sci. Data* 10 (1), 195–217. doi:10.5194/essd-10-195-2018

Babor, D., Plian, D., and Judele, L. (2009). Environmental impact of concrete. *Bull. Polytech. Inst. Jassy, Constr. Archit. Sect. Tomme LV (LIX), Fascicle* 4, 27–36.

Bellini, A., and Bang, S. (2022). Barriers for data management as an enabler of circular economy: an exploratory study of the Norwegian AEC-industry. *IOP Conf. Ser. Earth Environ. Sci.* 1122, 012047. doi:10.1088/1755-1315/1122/1/012047

Benachio, G. L. F., Freitas, M. d. C. D., and Tavares, S. F. (2020). Circular economy in the construction industry: a systematic literature review. *J. Clean. Prod.* 260, 121046. doi:10.1016/j.jclepro.2020.121046

Biswas, H. K., Sim, T. Y., and Lau, S. L. (2024). Impact of building information modelling and advanced technologies in the AEC industry: a Contemporary review and future Directions. *J. Build. Eng.* 82, 108165. doi:10.1016/j.jobe.2023.108165

Bonoli, A., Zanni, S., and Serrano-Bernardo, F. (2021). Sustainability in building and construction within the framework of circular Cities and European new green deal. The contribution of concrete recycling. *Sustainability* 13 (4), 2139. doi:10.3390/su13042139

Braun, V., and Clarke, V. (2012). '*Thematic analysis*. 57–71.

Bryman, A., and Bell, E. (2015). *Business research methods*. Oxford University Press.

Business Sweden (2022). Sweden's digital technologies Ecosystem, business Sweden. Available online at: https://www.business-sweden.com/insights/articles/swedens-digital-technologies-ecosystem/(Accessed September 15, 2024).

Bygballe, L. E., and Ingemansson, M. (2014). The logic of innovation in construction. *Ind. Mark. Manag.* 43 (3), 512–524. doi:10.1016/j.indmarman.2013.12.019

Camodeca, R., and Almici, A. (2021). Digital transformation and convergence toward the 2030 Agenda's sustainability development goals: Evidence from Italian listed firms. *Sustainability* 13 (21), 11831. doi:10.3390/su132111831

Çetin, S., Gruis, V. H., and Straub, A. (2022). Digitalization for a circular economy in the building industry: multiple-case study of Dutch social housing organizations. *Resour. conservation & Recycl. Adv.* doi:10.1016/j.rcradv.2022.200110

Creswell, J. W. (2014). *A Concise introduction to mixed methods research*. Thousand Oaks, CA: SAGE Publications.

Dang, D. S., and Ahmad, P. H. (2015). A review of text mining techniques associated with various application areas. Available online at: https://www.semanticscholar.org/paper/A-Review-of-Text-Mining-Techniques-Associated-with-Dang-Ahmad/58be9b174785c74444beea35c944b4aa57ce23f7 (Accessed April 16, 2024).

Daniotti, B., Pavan, A., Bolognesi, C., Mirarchi, C., and Signorini, M. (2022). 'Digital transformation in the construction sector: from BIM to digital twin. doi:10.5772/intechopen.103726

Davari, S., Jaberi, M., Yousfi, A., and Poirier, E. (2023). A traceability framework to enable circularity in the built environment. *Sustainability* 15 (10), 8278. doi:10.3390/su15108278

Del Rio, G., González, C., and Colsa, Á. (2021). Unleashing the convergence amid digitalization and sustainability towards pursuing the Sustainable Development Goals (SDGs): a holistic review. *J. Clean. Prod.* 280, 122204. doi:10.1016/j.jclepro.2020.122204

Demšar, J., and Zupan, B. (2013). Orange: data mining fruitful and fun - a historical perspective. *Inform. Slov.* 37, 55–60.

Dervishaj, A., and Gudmundsson, K. (2024). From LCA to circular design: a comparative study of digital tools for the built environment. *Resour. Conservation Recycl.* 200, 107291. doi:10.1016/j.resconrec.2023.107291

Dervishaj, A., Vargas, J. H., and Gudmundsson, K. (2023). 'Enabling reuse of prefabricated concrete components through multiple tracking technologies and digital twins'. doi:10.35490/ec3.2023.220

Dobesova, Z. (2024). Evaluation of Orange data mining software and examples for lecturing machine learning tasks in geoinformatics. *Comput. Appl. Eng. Educ.* 32 (4), e22735. doi:10.1002/cae.22735

ECSO (2021). Digitalisation in the construction sector - ECSO report. Available online at: https://bimaplus.org/uncategorized/digitalisation-in-the-construction-sector-ecso-report/(Accessed April 18, 2024).

Ellen MacArthur Foundation (2015). Towards a circular economy: business rationale for an accelerated transition. Available online at: https://www.ellenmacarthurfoundation.org/towards-a-circular-economy-business-rationale-for-an-accelerated-transition (Accessed April 17, 2024).

Ellen MacArthur Foundation (no date). The circular economy in detail. Available online at: https://www.ellenmacarthurfoundation.org/the-circular-economy-in-detail-deep-dive (Accessed April 18, 2024).

European Commission (2020). New circular economy action plan. *Eur. Comm. - Eur. Comm.* Available online at: https://ec.europa.eu/commission/presscorner/detail/en/ip_20_420 (Accessed April 19, 2024).

European Commission (2021). Delivering the European green deal - European Commission. Available online at: https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/european-green-deal/delivering-european-green-deal_en (Accessed March 17, 2024).

European Environmental Agency (2024). *Capturing the climate change mitigation benefits of circular economy and waste sector policies and measures*. European Environment Agency. Available online at: https://www.eea.europa.eu/publications/capturing-the-climate-change-mitigation/capturing-the-climate-change-mitigation (Accessed June 22, 2024).

European Union (2023). New regulation on construction products. Available online at: http://data.europa.eu/eli/C/2024/4018/oj/eng (Accessed November 11, 2024).

Eurostat (2023). Waste statistics, Waste statistics. Available online at: https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Waste_statistics (Accessed June 4, 2024).

Fei, W., Opoku, A., Agyekum, K., Oppon, J. A., Ahmed, V., Chen, C., et al. (2021). The critical role of the construction industry in achieving the sustainable development goals (SDGs): Delivering projects for the common Good. *Sustainability* 13 (16), 9112. doi:10.3390/su13169112

Galvin, R. (2015). How many interviews are enough? Do qualitative interviews in building energy consumption research produce reliable knowledge? *J. Build. Eng.* 1, 2–12. doi:10.1016/j.jobe.2014.12.001

Geels, F. W. (2002). Technological transitions as evolutionary reconfiguration processes: a multi-level perspective and a case-study. *Res. Policy* 31 (8), 1257–1274. doi:10.1016/S0048-7333(02)00062-8

Geels, F. W. (2011). The multi-level perspective on sustainability transitions: responses to seven criticisms. *Environ. Innovation Soc. Transitions* 1 (1), 24–40. doi:10.1016/j.eist.2011.02.002

Gerhardsson, H., Lindholm, C. L., Andersson, J., Kronberg, A., Wennesjö, M., and Shadram, F. (2020). Transitioning the Swedish building sector toward reuse and circularity. *IOP Conf. Ser. Earth Environ. Sci.* 588 (4), 042036. doi:10.1088/1755-1315/588/4/042036

Giovanardi, M. (2024). "RFId for construction sector. Technological innovation in circular economy perspective," in *Lecture notes in Mechanical Engineering*, 315–329. doi:10.1007/978-3-031-36922-3_18

Giovanardi, M., Konstantinou, T., Pollo, R., and Klein, T. (2022). Internet of Things for building façade traceability: a theoretical framework to enable circular economy through life-cycle information flows. *J. Clean. Prod.* 382, 135261. doi:10.1016/j.jclepro.2022.135261

Graf-Vlachy, L., Buhtz, K., and König, A. (2018). Social influence in technology adoption: taking stock and moving forward. *Manag. Rev. Q.* 68, 37–76. doi:10.1007/s11301-017-0133-3

Gurel, E., and Tat, M. (2017). Swot Analysis: A Theoretical Review. *J. Int. Soc. Res.* 10 (51), 994–1006. doi:10.17719/jisr.2017.1832

Habert, G., Miller, S. A., John, V. M., Provis, J. L., Favier, A., Horvath, A., et al. (2020). Environmental impacts and decarbonization strategies in the cement and concrete industries. *Nat. Rev. Earth & Environ.* 1, 559–573. doi:10.1038/s43017-020-0093-3

Herbert, L. (2017). *Digital transformation: B.U.I.L.D Your Organization's future*. London, UK: Bloomsbury Business.

Honic, M., Kovacic, I., Sibenik, G., and Rechberger, H. (2019). Data- and stakeholder management framework for the implementation of BIM-based Material Passports. *J. Build. Eng.* 23, 341–350. doi:10.1016/j.jobe.2019.01.017

Javaid, M., Haleem, A., Singh, R. P., Suman, R., and Gonzalez, E. S. (2022). Understanding the adoption of Industry 4.0 technologies in improving environmental sustainability. *Sustain. Operations Comput.* 3, 203–217. doi:10.1016/j.susoc.2022.01.008

Jayaraj, N., Klarin, A., and Ananthram, S. (2024). The transition towards solar energy storage: a multi-level perspective. *Energy Policy* 192, 114209. doi:10.1016/j.enpol.2024.114209

Jiang, P., Zhang, L., You, S., Fan, Y. V., Tan, R. R., Klemeš, J. J., et al. (2023). Blockchain technology applications in waste management: overview, challenges and opportunities. *J. Clean. Prod.* 421, 138466. doi:10.1016/j.jclepro.2023.138466

Jiang, Y., Su, S., Zhao, S., Zhong, R. Y., Qiu, W., Skibniewski, M. J., et al. (2024). Digital twin-enabled synchronized construction management: a roadmap from construction 4.0 towards future prospect. *Dev. Built Environ.* 19, 100512. doi:10.1016/j.dibe.2024.100512

Jick, T. D. (1979). Mixing qualitative and quantitative methods: triangulation in action. *Adm. Sci. Q.* 24 (4), 602. doi:10.2307/2392366

Kanger, L. (2021). Rethinking the Multi-level Perspective for energy transitions: from regime life-cycle to explanatory typology of transition pathways. *Energy Res. & Soc. Sci.* 71, 101829. doi:10.1016/j.erss.2020.101829

Katenbayeva, A., Glass, J., Anvuur, A., and Ghumra, S. (2016). Developing a theoretical framework of traceability for sustainability in the construction sector. Available online at: https://repository.lboro.ac.uk/articles/conference_contribution/Developing_a_theoretical_framework_of_traceability_for_sustainability_in_the_construction_sector/9436811/1 (Accessed April 18, 2024).

Kaufman, S., Saeri, A. K., Raven, R., Malekpour, S., and Smith, L. D. G. (2022). Behaviour in sustainability transitions: a mixed methods literature review. *OSF*. doi:10.31219/osf.io/krmcv

Kebede, R., Moscati, A., Tan, H., and Johansson, P. (2023). Circular economy in the built environment: a framework for implementing digital product passports with knowledge graphs. *Comput. Constr.* doi:10.35490/ec3.2023.245

Kedir, F., Bucher, D. F., and Hall, D. M. (2021). A proposed material passport Ontology to enable circularity for Industrialized construction. *null*. doi:10.35490/ec3.2021.159

Li, L. (2020). The governance of low-carbon transitions in a Multilevel perspective framework: how does the Concept of 'system transformation' work? *Energy Res. J.* 11, 45–53. doi:10.3844/erjsp.2020.45.53

Liu, H., and Zainul Abidin, N. (2024). A review on research of prefabricated building costs: exploring collaborations, Intellectual Basis, and research trends. *Sustainability* 16, 9823. doi:10.3390/su16229823

Liu, Q., Trevisan, A. H., Yang, M., and Mascarenhas, J. (2022). A framework of digital technologies for the circular economy: digital functions and mechanisms. *Bus. Strategy Environ.* 31, 2171–2192. doi:10.1002/bse.3015

Liu, Q., Trevisan, A. H., Yang, M., and Mascarenhas, J. (2022). A framework of digital technologies for the circular economy: digital functions and mechanisms. *Bus. Strategy Environ.* 31, 2171–2192. doi:10.1002/bse.3015

Ljubljana, B. L. University of (no date) *orange data mining, orange data mining*. Available online at: https://orangedatamining.com (Accessed December 10, 2024).

Lobo, A., Trevisan, A. H., Liu, Q., Yang, M., and Mascarenhas, J. (2021). "Barriers to transitioning towards Smart circular economy: a systematic literature review," in *Sustainable design and manufacturing*. Editors S. G. Scholz, R. J. Howlett, and R. Setchi (Singapore: Springer), 245–256. doi:10.1007/978-981-16-6128-0_24

Mahapatra, K., Shehu, M. O., and Almeida, R. (2024). "Exploring the prospect of digital tracking technology in enhancing the circularity of the Swedish precast concrete sector," in *2024 International conference on sustainable energy: energy transition and Net-Zero climate future (ICUE). 2024 International conference on sustainable energy: energy transition and Net-Zero climate future (ICUE)*, 1–7. doi:10.1109/ICUE63019.2024.10795516

Mainali, B., Mahapatra, K., and Pardalis, G. (2021). Strategies for deep renovation market of detached houses. *Renew. Sustain. Energy Rev.* 138, 110659. doi:10.1016/j.rser.2020.110659

Mallawaarachchi, B. H., and Jayakodi, S. (2023). Internet of things (iot)-enabled industrial symbiosis model for construction material sharing: bibliometric analysis, 595, 605. doi:10.31705/wcs.2023.49

Mêda, P., Munir, M., Calvetti, D., and Sousa, H. (2023). 'Information needs in digital products passport – discussing data framework and reasonability'. doi:10.35490/ec3.2023.244

Munaro, M., and Tavares, S. (2021). Materials passport's review: challenges and opportunities toward a circular economy building sector. *Built Environ. Proj. Asset Manag.* 11, 767–782. doi:10.1108/BEPAM-02-2020-0027

Nguyen, T. D., and Adhikari, S. (2023). The role of BIM in integrating digital twin in building construction: a literature review. *Sustainability* 15 (13), 10462. doi:10.3390/su151310462

Nilimaa, J. (2023). Smart materials and technologies for sustainable concrete construction. *Dev. Built Environ.* 15, 100177. doi:10.1016/j.dibe.2023.100177

NLTK Sentiment Analysis Tutorial: Text Mining & Analysis in Python (no date). Available online at: https://www.datacamp.com/tutorial/text-analytics-beginners-nltk (Accessed December 10, 2024).

Oberle, B., Bringezu, S., Hatfield-Dodds, S., Hellweg, S., Schandl, H., Clement, J., et al. (2019). UN global resources Outlook 2019: natural resources for the future We want.

Rastogi, N., and Trivedi, M. (2022). 'Pestle technique -a tool to identify external risks in construction projects'.

Regina Munaro, M., and Tavares, S. F. (2023). A review on barriers, drivers, and stakeholders towards the circular economy: the construction sector perspective. *Clean. Responsible Consum.* 8, 100107. doi:10.1016/j.clrc.2023.100107

RE:Source (2022). Circularity gap report Sweden, RE:source. Available online at: https://resource-sip.se/en/circularity-gap-report-sweden-en/(Accessed September 15, 2024).

Rodrigo, N., Omrany, H., Chang, R., and Zuo, J. (2023). Leveraging digital technologies for circular economy in construction industry: a way forward. *SMART Sustain. BUILT Environ.* 13, 85–116. doi:10.1108/SASBE-05-2023-0111

Schot, J., and Geels, F. W. (2008). Strategic niche management and sustainable innovation journeys: theory, findings, research agenda, and policy. *Technol. Analysis & Strategic Manag.* 20 (5), 537–554. doi:10.1080/09537320802292651

Sepasgozar, S. M. E., Khan, A., Smith, K., Romero, J., Shen, X., Shirowzhan, S., et al. (2023). BIM-And digital twin for developing convergence technologies as future of digital construction. *Buildings* 13 (2), 441. doi:10.3390/buildings13020441

Shehu, M. O. (2024). A multi-level analysis of the barriers 850 and opportunities in adopting digital tracking technologies for concrete circularity in Sweden. Available online at: https://urn.kb.se/resolve?urn=urn:nbn:se:lnu:diva-130716 (Accessed January 27, 2025).

Shishehgarkhaneh, M. B., Keivani, A., Moehler, R. C., Jelodari, N., and Laleh, S. R. (2024). Internet of things (IoT), building information modeling (BIM), and digital twin (DT) in construction industry: a review, bibliometric, and network analysis. Available online at: https://www.mdpi.com/2075-5309/12/10/1503 (Accessed November 10, 2024).

Statista (2023). *Global cement CO_2 emissions 1960-2022, Statista*. Available online at: https://www.statista.com/statistics/1299532/carbon-dioxide-emissions-worldwide-cement-manufacturing/(Accessed: November 9, 2024).

Tang, L., and Feng, Y. (2019). 'Sentiment analytics of social media data in the construction industry: preliminary study on Weibo', *Constructing Smart Cities: proceedings of the 22nd CIB World building Congress (CIB2019), 17-21 June 2019*, The Hong Kong Polytechnic University, Hong Kong, China. Available online at: https://researchdirect.westernsydney.edu.au/islandora/object/uws%3A55408/(Accessed: 23 December 2024).

Teisserenc, B., and Sepasgozar, S. (2021). Adoption of blockchain technology through digital twins in the construction industry 4.0: a PESTELS approach. *BUILDINGS* 11 (12), 670. doi:10.3390/buildings11120670

Thirumal, S., Udawatta, N., Karunasena, G., and Al-Ameri, R. (2024). Barriers to adopting digital technologies to implement circular economy practices in the construction industry: a systematic literature review. *Sustainability* 16 (8), 3185. doi:10.3390/su16083185

Tiwari, R., Kumar, G., and Gunjan, V. (2023). 'Effect of environment on Students performance through orange tool of data mining', in, pp. 283–292. doi:10.1007/978-981-99-2058-7_26

Tsironis, G., Daglis, T., and Tsagarakis, K. P. (2024). The 21 most practiced RE-s of circular economy from LinkedIn company profiles on a global scale. *Resour. Conservation & Recycl. Adv.* 21, 200202. doi:10.1016/j.rcradv.2024.200202

Uçar, E., Dain, M. A. L., and Joly, I. (2020). Digital technologies in circular economy transition: Evidence from case studies. *Procedia CIRP* 90, 133–136. doi:10.1016/j.procir.2020.01.058

UNEP (2023). Digital transformations | UNEP - UN environment Programme. Available online at: https://www.unep.org/topics/digital-transformations (Accessed November 10, 2024).

Vardopoulos, I., Tsilika, E., Sarantakou, E., Zorpas, A., Salvati, L., and Tsartas, P. (2021). An integrated SWOT-PESTLE-AHP model assessing sustainability in adaptive reuse projects. *Appl. Sci.* 11 (15), 7134. doi:10.3390/app11157134

Vogel, J. (2020). Incentivising innovation in the Swedish construction industry. doi:10.13140/RG.2.2.25481.47203

Wong, R. W. M., and Loo, B. P. Y. (2022). Sustainability implications of using precast concrete in construction: an in-depth project-level analysis spanning two decades. *J. Clean. Prod.* 378, 134486. doi:10.1016/j.jclepro.2022.134486

Wuni, I. Y. (2022). Mapping the barriers to circular economy adoption in the construction industry: a systematic review, Pareto analysis, and mitigation strategy map. *Build. Environ.* 223, 109453. doi:10.1016/j.buildenv.2022.109453

Zakrisson, A. (2018). Swedish step-by-step guide to construction waste. Available online at: https://iimpcoll.com/content/swedish-guide-construction-waste (Accessed September 16, 2024).

Zhao, J., Olivieri, H., Seppänen, O., Peltokorpi, A., Badihi, B., and Lundström, P. (2017). "Data analysis on applying real time tracking in production control of construction," in *2017 IEEE International conference on industrial engineering and engineering management (IEEM). 2017 IEEE International conference on industrial engineering and engineering management (IEEM)*, 573–577. doi:10.1109/IEEM.2017.8289956

frontiers | Frontiers in Built Environment

TYPE Methods
PUBLISHED 28 April 2025
DOI 10.3389/fbuil.2025.1575913

Check for updates

OPEN ACCESS

EDITED BY
Salman Azhar,
Auburn University, United States

REVIEWED BY
Rehan Masood,
Otago Polytechnic, New Zealand
Roy Lan,
University of Texas at San Antonio, United States

*CORRESPONDENCE
Nazym Shogelova,
nazymshogelova@gmail.com

RECEIVED 13 February 2025
ACCEPTED 16 April 2025
PUBLISHED 28 April 2025
CORRECTED 09 June 2025

CITATION
Kabzhan Z, Shakhnovich A, Gorshkov S, Yemenov Y, Gorshkov F and Shogelova N (2025) Semantic and ontology-based analysis of regulatory documents for construction industry digitalization.
Front. Built Environ. 11:1575913.
doi: 10.3389/fbuil.2025.1575913

COPYRIGHT
© 2025 Kabzhan, Shakhnovich, Gorshkov, Yemenov, Gorshkov and Shogelova. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Semantic and ontology-based analysis of regulatory documents for construction industry digitalization

Zarina Kabzhan[1], Alexandr Shakhnovich[1], Sergey Gorshkov[2], Yussuf Yemenov[1], Fedor Gorshkov[2] and Nazym Shogelova[1]*

[1]JSC Kazakh Research and Design Institute of Construction and Architecture, Almaty, Kazakhstan, [2]LLC Datavera, Almaty, Kazakhstan

In the context of the digitalization of the construction industry, the demand for automation in the analysis of regulatory documents is increasing. The significant volume, structural complexity, and frequent amendments of regulatory acts lead to semantic inconsistencies, duplication of provisions, and contradictions in requirements. The aim of this study is to develop a method for the automated analysis of regulatory documents in the construction sector based on the integration of ontological modeling and natural language processing (NLP) techniques. The relevance of the topic is driven by the need for the digital transformation of construction standardization, which involves eliminating redundant provisions, logical inconsistencies, and outdated references in the regulatory framework. The study proposes a methodology for constructing semantic "profiles" of regulatory statements, which include structured components: subject, predicate, object, modality, and additional conditions. A software prototype has been developed, implementing an algorithm for semantic matching of regulatory requirements using an ontological model that incorporates SKOS-based terminology, deontic logic, and domain-specific concepts of the construction sector. Experiments were conducted on a corpus of 14 regulatory documents of the Republic of Kazakhstan (≈242,000 words), demonstrating high computational efficiency (document analysis time <10 s) and acceptable quality metrics (F1-score up to 0.86). The results confirm the applicability of the proposed method for integration into automated regulatory compliance control systems and Building Information Modeling (BIM).

KEYWORDS

semantic analysis, ontology modeling, regulatory documents, digitalization of construction, NLP, automated analysis

1 Introduction

Modern construction is characterized by an extensive system of regulatory oversight throughout all stages of the asset life cycle—from design to operation. Projects and construction activities must comply with building codes, technical regulations, national, and industry standards that ensure safety, quality, and efficiency requirements. However, the large volume, complex structure, and frequent updates of the regulatory framework complicate the interpretation and practical application of requirements,

potentially hindering effective project design and implementation (Guo et al., 2021).One of the key challenges is the ambiguity of formulations resulting from discrepancies in terminology and definitions, which allows for multiple interpretations depending on the context of the document. Additionally, the duplication of provisions and logical contradictions among different regulatory acts further complicate the use of normative documents, making compliance verification more difficult and increasing the risk of legal conflicts (Guo et al., 2021). In this context, the automation of regulatory requirement analysis is viewed as a necessary component of the digital transformation of the construction industry, as it can enhance the transparency and speed of project approval processes.

Traditionally, compliance verification of construction projects with regulatory requirements is performed manually by experts, which is time-consuming, labor-intensive, and prone to subjective errors. For instance, the average review period for a commercial project in the city of Mesa (United States) is 18 working days, with analysis costs reaching up to $90 per hour (City of Mesa, 2012). Non-compliance, in turn, can lead to serious financial consequences—for example, Wal-Mart was once fined $1 million for violating environmental regulations (Salama and El-Gohary, 2013). Automating the compliance checking process could significantly reduce these costs and associated risks. Automated Compliance Checking (ACC) is currently viewed as a promising solution for accelerating expert review, reducing errors, and improving the efficiency of regulatory document management (Tang et al., 2012; Sydora and Stroulia, 2020). Over the past decades, various research approaches and software systems for ACC have been proposed, demonstrating the feasibility of partial automation. However, in practice, the level of automation in regulatory compliance within the construction domain remains low. To date, there has been no widespread adoption of digital methods in either the content of regulatory requirements or the procedures for their verification. This can be attributed to several limitations of the existing approaches (Zhang and El-Gohary, 2015).

Most modern ACC systems require manual rule authoring: experts manually extract requirements from normative texts and translate them into machine-readable formats (Zhang et al., 2022). This process is labor-intensive, prone to inaccuracies, and complicates system updates when regulations change. Many existing solutions rely on hardcoded rules or static databases, which limits their flexibility and scalability. Moreover, common methods for analyzing regulatory documents—such as keyword search and syntactic parsing—fail to capture semantic relationships and logical dependencies between provisions. This results in low accuracy of automated requirement interpretation, particularly for complex documents that contain numerous cross-references, exceptions, and conditional constructs. As experts have noted, full automation of compliance checking is not feasible without intelligent text processing—that is, transforming fragmented textual requirements into formal, machine-executable rules and linking them to digital project models (Zhong et al., 2012; Tierney, 2012). These limitations highlight a significant scientific and practical niche for developing novel methods to automate regulatory compliance control.

This study aims to develop a method for the automated semantic analysis of regulatory documents in the construction domain, based on ontological modeling and natural language processing (NLP) techniques, to overcome existing limitations in compliance automation. To achieve this goal, an analysis of current approaches to automated regulatory control is conducted to identify their weaknesses and justify the need for a new solution. The research involves the development of a multi-level ontological model designed to structure the content of regulatory documents and represent logical-semantic relationships between their provisions. In parallel, algorithms are developed to apply NLP methods for the automatic extraction of key terms and structured components of requirements from the text, as well as for generating their semantic representations ("profiles"). The proposed method enables the automatic identification of duplicated provisions, contradictions, and outdated references through the integration of the ontological model and semantic analysis. Its effectiveness is evaluated using a corpus of construction regulations, allowing for the assessment of inconsistency detection quality and demonstrating the method's potential for integration into digital platforms, including regulatory document management systems and Building Information Modeling (BIM) environments.

Achieving the stated goal and addressing the outlined objectives will enhance the degree of automation in the analysis of regulatory requirements, enable more accurate interpretation of normative documents, and provide a foundation for integrating compliance checking into BIM and other digital construction processes. The scientific contribution of this study lies in the advancement of methods for semantic analysis of regulatory texts, while its practical value is reflected in establishing the prerequisites for more efficient and reliable project expert review within the context of the industry's digital transformation.

2 Literature review

2.1 Analysis of regulatory requirements and ACC systems

The problem of Automated Compliance Checking (ACC) has been the focus of intensive research over the past several decades. As early as the 1980s and 1990s, initial attempts were made to formalize construction standards using expert systems and knowledge bases.

Existing ACC systems are based on transforming regulatory documents into logical rules suitable for machine analysis. Zhang and El-Gohary (2017) proposed a system that combines semantic text processing with logical reasoning, enabling the automatic extraction of requirements from building codes and their alignment with building BIM models (Zhang and El-Gohary 2017). Subsequently, the authors developed a method based on deep neural networks, which achieves high accuracy in extracting both syntactic and semantic elements from regulatory texts, reaching a precision of over 93% (Zhang and El-Gohary 2021).

The issue of scalability and flexibility of ACC systems when dealing with diverse regulatory documents remains a pressing challenge. In the study by Xue and Zhang, a mechanism was proposed for extending the set of rule transformation templates, which significantly improves the coverage of verifiable provisions and the accuracy of transformation through the iterative generation of logical expressions based on predefined patterns (Xue and Zhang, 2022).

Another emerging direction is the integration of large language models (LLMs), such as GPT-4, into the processes of automated regulatory compliance. Al-Turki et al. (2024) developed an LLM-based system that transforms building regulations into a structured YAML format using an active learning strategy. This approach significantly improves both the structural and semantic accuracy of interpreting regulatory requirements.

In addition to technical aspects, institutional prerequisites for the implementation of ACC systems are also being explored. Nama and Alalawi (2023) analyzed the adoption of automated compliance checking systems in Middle Eastern countries, identifying key stages in the transformation of building permit issuance processes and the integration of digital solutions into governmental workflows.

2.2 Digital technologies and BIM

Building Information Modeling (BIM) has become a key enabler in the development of Automated Compliance Checking (ACC) systems, providing a digital representation of the designed asset in the form of interconnected elements and their parameters. This data structure allows for the formalized description of design solutions, and consequently, enables programmatic verification of their compliance with regulatory requirements. In contrast to traditional checks based on drawings and textual documentation, the BIM-oriented approach improves accuracy, reduces time expenditures, and minimizes the impact of human error (Preidel and Borrmann, 2018).

Modern approaches to automated compliance checking involve the application of semantic processing to regulatory documents, followed by the alignment of extracted information with BIM model parameters. For example, in the study by Guo et al., a method was proposed in which regulatory requirements are extracted from texts using natural language processing (NLP), transformed into semantic rules, and then automatically applied to BIM models through the use of SPARQL queries (Guo et al., 2021).

More specialized applications of BIM for regulatory compliance checking have also demonstrated effectiveness. For instance, within the RegBIM project, an approach was developed to automate the verification of sustainable design solutions by extending the semantics of IFC to assess the environmental performance of projects (Kasim, 2015). In Pakistan, the implementation of a BIM-oriented automated checking system at the municipal level reduced the project documentation approval time from 1 week to 6 h, showcasing the high efficiency of digitalization in the expert review process (Aslam and Umar, 2023).

Literature reviews highlight that most modern systems rely on rule formalization based on the IFC standard. However, the interpretation of regulatory requirements remains the most labor-intensive stage: translating textual rules into machine-readable formats requires the use of either logical or ontological models, or specialized software plugins (Ismail et al., 2017).

Recent studies also propose innovative ACC system architectures. In particular, Li et al. (2024) developed a conceptual framework for automated compliance checking based on knowledge graphs, in which regulatory documents are transformed into ontologies using a Chinese NLP model and then applied to BIM models to detect violations. Another study introduced a system that leverages blockchain technology to ensure transparency and traceability of both automated and manual compliance verification stages within the BIM environment (Gao and Zhong, 2022).

2.3 Ontological modeling of regulatory knowledge

One of the most promising directions in the development of ACC systems is the use of ontological modeling for the formalization of regulatory knowledge. Ontologies enable the representation of building codes and standards as formal concepts, objects, properties, and logical relationships between them, thus making automatic interpretation and verification of regulatory requirements possible. This approach ensures a high degree of machine-readability and supports logical reasoning, which is essential for intelligent analysis of design documentation.

A number of studies focus on developing ontologies that capture the structure and semantics of building regulations. For example, Zhong et al. (2012) proposed an ontological approach to construction quality checking, based on the formalization of requirements using the OWL format and SWRL rules, which are applicable to inspection data. Jiang and colleagues developed a comprehensive multi-ontology merging method that enables the alignment of terminology between the building model and the regulatory ontology, as well as the application of logical reasoning through SPARQL queries (Jiang et al., 2022).

Of particular interest is the application of deontic logic—a logical system that describes concepts of obligation and prohibition—for the interpretation of regulatory requirements. Salama and El-Gohary (2011) proposed the use of deontological concepts in ACC systems to formalize rights and obligations within regulations, which is especially important for evaluating permissible and impermissible actions in construction.

One of the limitations of the ontological approach remains its labor intensity: the development of comprehensive ontologies requires the involvement of domain experts and significant time investment. Moreover, expressing dynamic or context-dependent requirements can be challenging within static ontological structures without the incorporation of additional logical mechanisms, such as first-order logic or conditional rules (Zhang and El-Gohary, 2017).

Nevertheless, the use of ontologies represents a powerful tool for integrating regulatory knowledge into intelligent systems, enabling a higher level of automation and reusability of formalized requirements across various tasks, including design, expert review, and compliance control.

2.4 Natural language processing (NLP) in regulatory analysis

Alongside ontological modeling, methods of Natural Language Processing (NLP) are actively evolving within the ACC domain, aiming to automate the analysis of regulatory document texts. The primary task of NLP in this context is to transform unstructured text—such as articles, clauses, and conditions of building codes—into formalized data suitable for subsequent logical processing and alignment with design model parameters.

In the early stages of such systems, template-based syntactic approaches were commonly used. Key markers (e.g., "shall," "is not permitted," "must") were employed to identify regulatory provisions and subsequently match them to elements of the construction model. However, these methods were limited in their ability to interpret complex constructions of regulatory language. In response to this, Zhang and El-Gohary (2016) proposed a semantics-oriented approach that applies a system of grammatical and semantic rules to extract objects, parameters, and logical relationships from regulatory texts.

Subsequently, hybrid algorithms were developed that combine syntactic parsing, semantic rules, and machine learning methods. In one study, the authors implemented a multi-component approach involving text classification, information extraction (IE), and the transformation of requirements into logical rules suitable for automated compliance checking. The system achieved high precision and recall in extracting quantitative requirements from international building codes (Zhang and El-Gohary 2013).

The specific characteristics of regulatory documents—such as complex syntactic structures, numerous conditions, exceptions, and cross-references—make the application of standard NLP models insufficiently effective. As a result, modern approaches combine linguistic rules with ontological information to eliminate semantic ambiguities and improve the accuracy of interpretation (Zheng et al., 2022).

A new and emerging direction involves the application of large language models (LLMs), such as GPT-4, to regulatory analysis tasks. Chen et al. proposed an architecture that combines LLMs with ontology to extract structured requirements from regulatory texts and verify them against BIM models. Pre-trained LLMs demonstrate a strong capability for interpreting complex texts and significantly reduce the need for manual data annotation (Chen et al., 2024).

However, the application of such models in the construction domain presents several challenges: the need to adapt to the specificity of professional language, limited transparency in logical reasoning, and high sensitivity to contextual nuances of regulatory phrasing. In this regard, the greatest potential lies in the integration of neural models with formalized knowledge in the form of ontologies and logical rules, which provides both flexibility and controllability in the interpretation of regulations.

2.5 Limitations of existing approaches and the research gap

Despite significant progress in the development of automated ACC systems, most existing solutions are focused on analyzing design models within the context of strictly predefined, manually formalized rules. These rules are typically encoded by domain experts based on regulatory texts and cover only limited aspects of requirements—such as geometric parameters, accessibility, fire safety clearances, and similar criteria (Jiang et al., 2022). As a result, ACC systems primarily operate as checklist-based verification modules and rarely treat the regulatory corpus itself as a subject of analysis.

Comprehensive semantic processing of the regulatory texts themselves—including the identification of internal contradictions, duplicated provisions, and outdated references—is currently addressed in only a limited number of studies. The vast majority of existing approaches assume that regulatory documents are consistent and free of contradictions, which does not always reflect real-world practice. In actual design conditions, regulations may contain ambiguous wording, overlapping or conflicting requirements, as well as references to outdated or repealed documents, all of which pose significant risks of misinterpretation (Zheng et al., 2022).

The analysis of regulatory overlaps and conflicts across multiple documents—such as different building codes, standards, and technical regulations—remains a largely unexplored area, despite its critical importance. Most modern ACC systems lack mechanisms for comparing provisions across various sources and do not track their currency, which necessitates manual verification and updating of references by domain experts (Beach et al., 2015).

Given these limitations, there is a growing need for approaches that can not only verify model compliance but also perform semantic analysis of the regulatory documents themselves. This includes automatic knowledge structuring, detection of logical inconsistencies, terminology alignment, and ensuring the relevance of regulatory provisions. One of the promising directions is the integration of ontological modeling with natural language processing (NLP) methods, enabling the extraction, formalization, and comparison of requirements from multiple sources (Zhang and El-Gohary 2016).

The approach proposed in this study is aimed precisely at addressing these challenges. Unlike conventional ACC systems focused solely on verifying BIM models, the developed method emphasizes in-depth digital processing of the regulatory texts themselves. By combining ontological knowledge representation with NLP techniques, the approach enables automation of regulation analysis at a more fundamental level, laying the groundwork for a fully digital chain: regulatory requirements–information model–compliance verification. Thus, the presented research fills the identified scientific and practical gap, advancing the construction industry toward a new stage of digitalization that enhances both the efficiency and quality of design decisions.

3 Methods and materials

For the purpose of formalization and automated analysis of regulatory requirements applied in the construction industry, this study proposes a methodology based on the use of ontological modeling, natural language processing (NLP) techniques, and semantic analysis (Chen et al., 2024). The domain of interest is the set of building codes and regulations adopted in the Republic of Kazakhstan. The study is grounded in the following key principles.

First, regulatory documents possess a complex, multi-level structure that includes nested conditions, exceptions, cross-references, and hierarchically organized requirements.

Accordingly, the proposed methodology incorporates a step-by-step analysis of the syntactic and semantic characteristics of regulatory statements, aimed at their subsequent formalization and alignment.

At the first stage, preliminary linguistic processing of the texts is performed, including tokenization, lemmatization, part-of-speech (POS) tagging, dependency parsing, and coreference resolution.

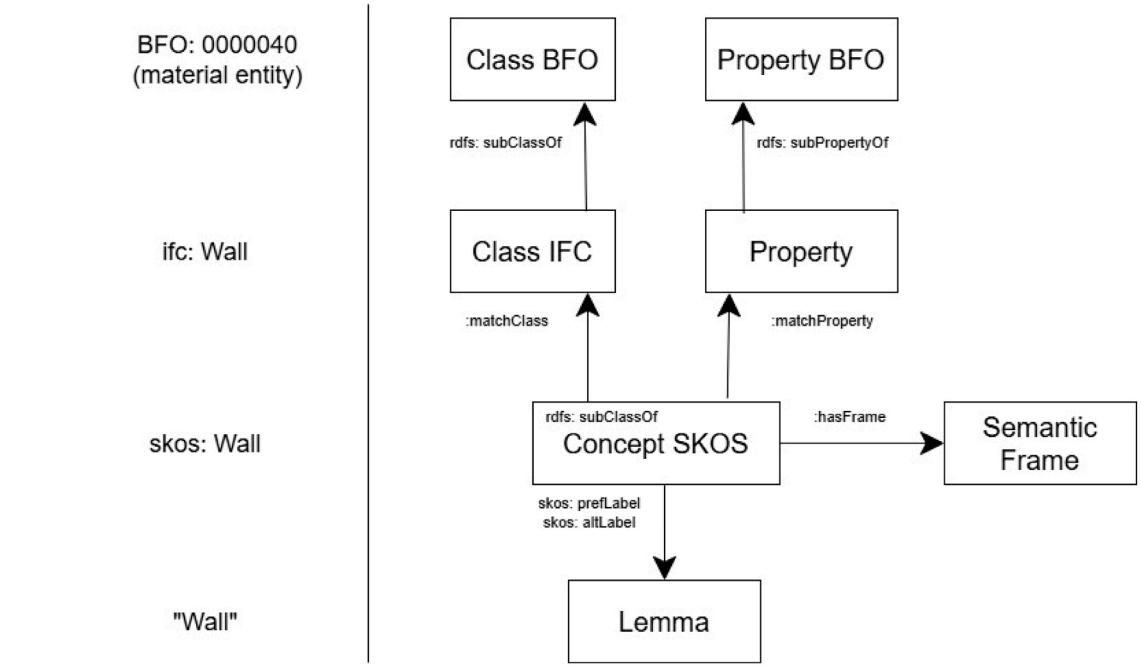


FIGURE 1
Relationship between elements of the proposed ontologies.

These procedures are carried out using the DataVera EKG Language Processing (EKG LP) software module (DataVera, 2025), which is built on the SpaCy library and adapted to the specifics of regulatory vocabulary.

At the second stage, textual fragments are aligned with the ontological model, which is represented as a set of interconnected ontologies (Figure 1):

- Upper-level ontology (based on BFO), used to represent universal categories such as objects, processes, and relationships;
- Domain ontology of the construction sector (based on IFC), covering capital construction assets, engineering systems, and life cycle processes;
- Regulatory statement ontology, based on deontic logic, describing the structure of norms (subject, modality, action, object, and applicability condition);
- Terminology ontology (SKOS model), providing linkage between the terms used in regulatory documents and the concepts of the domain ontology.

The formalized representation of regulatory provisions is carried out in the form of semantic profiles, which include the following elements: subject (addressee of the requirement), modality (obligation, possibility, prohibition), predicate (action or characteristic), object (result of the action), as well as additional attributes (conditions, exceptions, time frames, etc.).

To account for the complex structure of regulatory texts, the methodology implements mechanisms for:

- Detection of nested conditions (through the analysis of syntactic structures and conditional operators);
- Processing of exceptions, formed through negation constructs or limitations on the scope of regulations;
- Reconstruction of hierarchical relationships between regulatory provisions, using structural markers and contextual analysis of headings, articles, and subsections.

At the final stage, a comparative semantic analysis is performed, aimed at identifying:

- Duplicated provisions (when key elements of the semantic profile match);
- Contradictions (when there are discrepancies in modalities or conditions of application);
- Semantic inconsistencies (in definitions of terms and interpretations of concepts).

The comparison of semantic profiles is carried out based on a calculated similarity metric, the threshold value of which is determined empirically. In the case of significant discrepancies, the corresponding fragments are forwarded for expert review.

The developed system is designed for the automated semantic analysis of regulatory documents, identifying contradictions, duplicated provisions, and semantic inconsistencies. The architectural solution (Figure 2) is based on the use of ontological models, graph and relational databases, as well as natural language processing (NLP) methods.

The system includes several key components that ensure its functionality. A graph-based RDF triple store database (Apache Fuseki) is used for storing ontological models, enabling complex semantic queries and analysis of relationships between concepts. A relational or document-oriented storage system (PostgreSQL) is employed to store the results of the linguistic analysis of regulatory

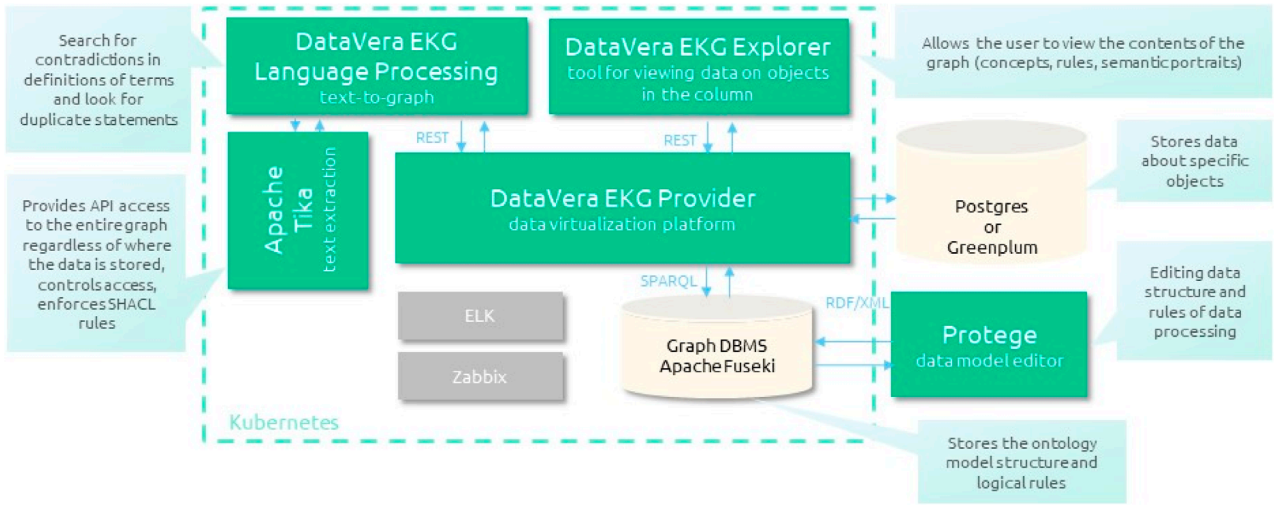


FIGURE 2
Architecture of the automated system for processing regulatory document texts.

texts (Jadala and Burugari, 2024). An important element is the data management platform (DataVera EKG Provider (DataVera, 2025)), which ensures information storage in accordance with the ontological model, supports both synchronous and asynchronous APIs, executes SPARQL queries, and performs data validation using SHACL rules (Ke et al., 2024). The system also includes application software modules, such as the linguistic analysis module for regulatory documents (DataVera EKG LP (DataVera, 2025)) and the semantic analysis module, which identifies contradictions in terminology and detects duplicated provisions. Monitoring and logging tools, such as ELK and Zabbix, are used to ensure system oversight and log collection (Bilobrovets, 2023).

The system is implemented as a set of containers deployed in a Kubernetes environment (Poniszewska-Marańda and Czechowska, 2021), which ensures its scalability and fault tolerance.

The processing of regulatory texts is performed in stages, starting with grammatical and semantic analysis (DataVera, 2025):

- Sentence structure analysis includes POS-tagging and dependency parsing, which allows for the identification of parts of speech and the establishment of grammatical dependencies between words. Coreference resolution is also performed, involving the substitution of nouns for pronouns and clarification of implied elements in the statement.
- Lemmatization ensures the conversion of word forms to their base form, simplifying subsequent processing and matching.
- Semantic matching involves identifying the concepts corresponding to the words in the sentence based on ontological models. In the absence of an exact match in the existing ontology, the system automatically generates *ad hoc* concepts limited to the specific context of the document.
- Formation of the semantic profile involves identifying subjects, predicates, modalities, objects, circumstances, and other elements necessary for the structured representation of regulatory content.

The result of the algorithm's operation is the formalized representation of each statement in the form of a set of semantic profiles, suitable for further analysis. Based on the obtained semantic profiles, a comparison of regulatory provisions is performed, allowing for the identification of contradictions, duplication, and semantic inconsistencies.

The identification of contradictions in terminology is carried out by analyzing statements that contain definitions of regulatory terms. The comparison of such statements allows for classifying the results into three groups (Liu et al., 2020):

- Semantic equivalence (the definitions are identical or close in meaning).
- Difference in scope (one definition is a specific case of the other).
- Semantic contradiction, when mutually exclusive interpretations of the same term are identified.

The search for duplicated regulatory provisions is performed by comparing the key elements of the semantic profile. If statements from different documents have matching predicates, objects, subjects, modalities, and additional parameters, the system calculates a numerical similarity metric. If the threshold value is exceeded, the statements are considered duplicated.

Similarly, contradictory statements are identified. If two statements refer to the same entity (matching subject, predicate, and object) but have different modalities, a logical contradiction is detected. In cases where additional elements of the semantic description differ, the inconsistency is evaluated quantitatively. If the discrepancy exceeds the established threshold, the divergences are forwarded for expert analysis.

The developed method for analyzing regulatory documents has a number of limitations related to the depth of semantic processing. First, the system evaluates the semantic profile of each statement in isolation, which excludes the possibility of analyzing situations where a single statement in one document corresponds to multiple statements in another. Second, the current implementation does not

account for the temporal aspect of regulatory provisions, meaning it does not analyze to which time period a particular directive applies (past, present, or likely future). Third, the system does not generate a comprehensive semantic description of the situations to which the requirements apply, but is limited to representing the regulatory directive in a structured form. While this simplifies the development and implementation of the system, such a level of formalization is insufficient for automated compliance checking and is intended solely for identifying inconsistencies and duplications in regulatory provisions.

To address the identified limitations, it is proposed to further develop the methodology across several interrelated directions. One of the key vectors is the development of a mechanism for inter-document semantic aggregation, which would enable the establishment of relationships such as equivalence, specification, logical entailment, and subordination between regulatory statements—both within a single document and across multiple sources. This would allow for the modeling of complex regulatory dependencies and improve the accuracy of contradiction detection.

Special attention is planned to be given to incorporating the temporal aspect of regulatory requirements. This involves annotating regulatory provisions with temporal markers (such as effective date, duration, and period of applicability), followed by integration with temporal ontologies. Such an approach will enable the tracking of regulatory evolution and the assessment of the applicability of provisions at a given point in time.

Another important direction is the modeling of regulatory situations through the expansion of the ontological model by incorporating concepts that describe typical scenarios for the application of requirements. This creates a foundation for shifting from the analysis of isolated provisions to a comprehensive assessment of regulatory conditions based on the context of design or operation of built assets. Such a level of detail will enhance the practical relevance of the developed system in professional practice.

To improve the completeness and validity of the analysis, it is proposed to integrate logic-based semantic reasoning using ontological rule languages such as SHACL or SWRL. This will enable not only the interpretation of individual statements, but also the formalization of logical relationships between them, thereby allowing for deductive consistency checking of regulatory requirements.

Finally, an important element of future work is the implementation of a contextual semantic disambiguation mechanism using trainable language models (e.g., BERT or GPT) adapted to a corpus of regulatory texts. The use of such models will enable accurate interpretation of terms and constructions depending on their usage context, especially in cases where the same concept may have different meanings in different sections or documents.

The implementation of the proposed directions will eliminate current limitations and significantly expand the functional capabilities of the system. This will pave the way for the development of a full-featured intelligent platform for regulatory analysis, capable of supporting tasks related to design, expert review, auditing, and legal compliance in the context of the construction industry's digital transformation.

The proposed architecture and methodology enable effective analysis of regulatory documents in the construction sector by providing their structured representation, identifying semantic inconsistencies, and supporting the development of a more coherent regulatory framework.

4 Results

To assess the applicability of the proposed approach, the study employed the EKG LP software suite, developed to address a wide range of text processing tasks. The choice of this software is justified by its ability not only to extract key entities and relationships from text, but also to generate an ontological representation of document structure, which is critically important for analyzing complex regulatory acts. Unlike many other systems, EKG LP provides built-in tools for constructing knowledge graphs and performing semantic annotation, enabling the automation of regulatory requirement interpretation, contradiction detection, and the formalization of logical relationships between provisions.

In addition, the software suite is integrated with corporate databases and electronic document management systems, making it particularly valuable in the context of digital transformation in the construction industry. Although EKG LP has not yet achieved widespread adoption among construction professionals, its potential is actively being explored within projects aimed at the digitalization of the regulatory and technical framework, including initiatives for implementing information modeling technologies and developing digital codes and standards. The present study demonstrates the applicability of this tool specifically in the context of construction regulation tasks, confirming its relevance and effectiveness within this domain.

The first stage of text processing in EKG LP involves lemmatization and grammatical structure parsing of sentences, performed using tools from the SpaCy framework (Díaz et al., 2024). During analysis, each element of the text is assigned morphological and syntactic characteristics, and the identified grammatical dependencies are structured hierarchically. These dependencies are visualized using the displacy tool and are shown in Figure 3. The output of this processing phase, printed from the EKG LP source code, is presented in Figure 4. As an example, consider the sentence: "Joint connections of prefabricated elements and multilayer structures shall be designed to withstand temperature deformations and forces arising from uneven foundation settlement and other operational impacts."

The result of this processing phase is a structured representation of the sentence, in which each word and punctuation mark is linked to its lemmatized form along with its grammatical function. On the left side of the visualized representation, the words of the original sentence are arranged according to the identified syntactic dependencies. On the right side of the table (Figure 4), each word is annotated with its part of speech (POS-tag) and the type of syntactic relation it holds with other sentence elements (Relation type), enabling further processing at the level of semantic dependencies.

Based on the data obtained, EKG LP constructs a "semantic profile" of the statement (Table 1), the structure of which is analogous to the model used in the Nòmos two framework (Mandal et al., 2015). During this process, the core semantic structure of the text is identified, including the

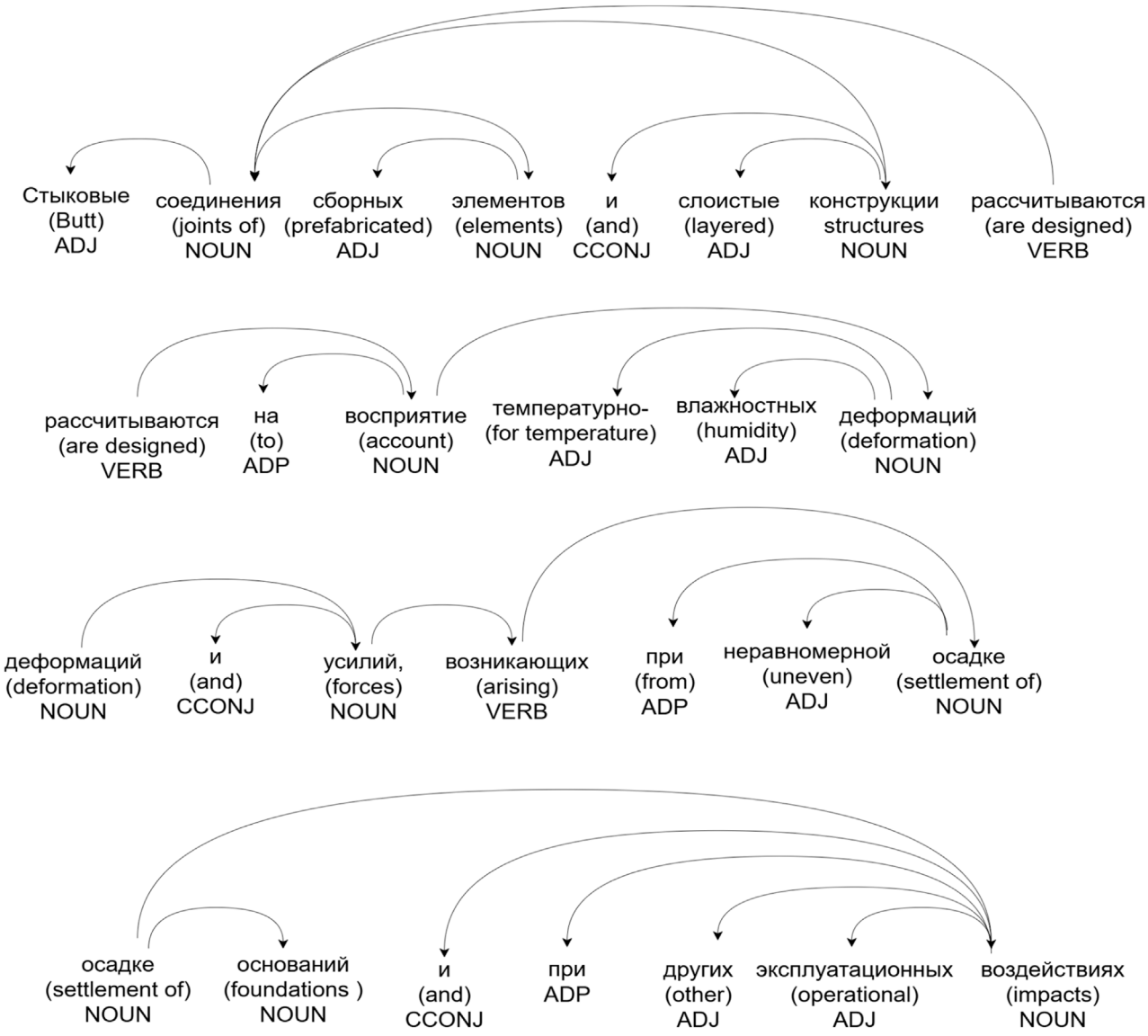


FIGURE 3
Result of lemmatization and grammatical structure parsing of the sentence.

key components of the statement: predicate, object, subject, and modality.

To illustrate, consider the analysis of a specific example, where in the phrase "connections are designed to withstand" the following semantic components are extracted: "connection" as the subject, "designed" as the predicate (normalized to the base form "design"), and "withstand" as the object.

In addition, dependency chains are generated for both the subject and the object, enabling a more detailed description of regulatory provisions and contributing to the precise identification of their semantic structure. The resulting semantic structures are subsequently used to detect contradictions, duplicated provisions, and semantic inconsistencies in regulatory documents.

To verify the proposed approach, the EKG LP software suite was used, developed for the automated analysis of regulatory documents. Its primary purpose in this study is to identify duplicated requirements, analyze the semantic similarity of phrases, and detect contradictions in regulatory provisions.

In its default configuration, EKG LP generates a "semantic profile" for each statement, consisting of seven key components: subject, predicate, object, modality, negation, definition, and complement/circumstance. The analysis revealed that this structure is sufficient for accurately representing simple sentences; however, regulatory documents in the construction sector are characterized by a high degree of syntactic complexity. As a result, the basic algorithm requires further refinement to enable more accurate modeling of complex statements. Nevertheless, even the current version of the algorithm demonstrates satisfactory performance in comparing phrases with similar semantic structures.

When analyzing two semantically similar statements (Martinez-Gil and Chaves-González, 2022), EKG LP generates their semantic profiles, which turn out to be nearly identical, with only minor

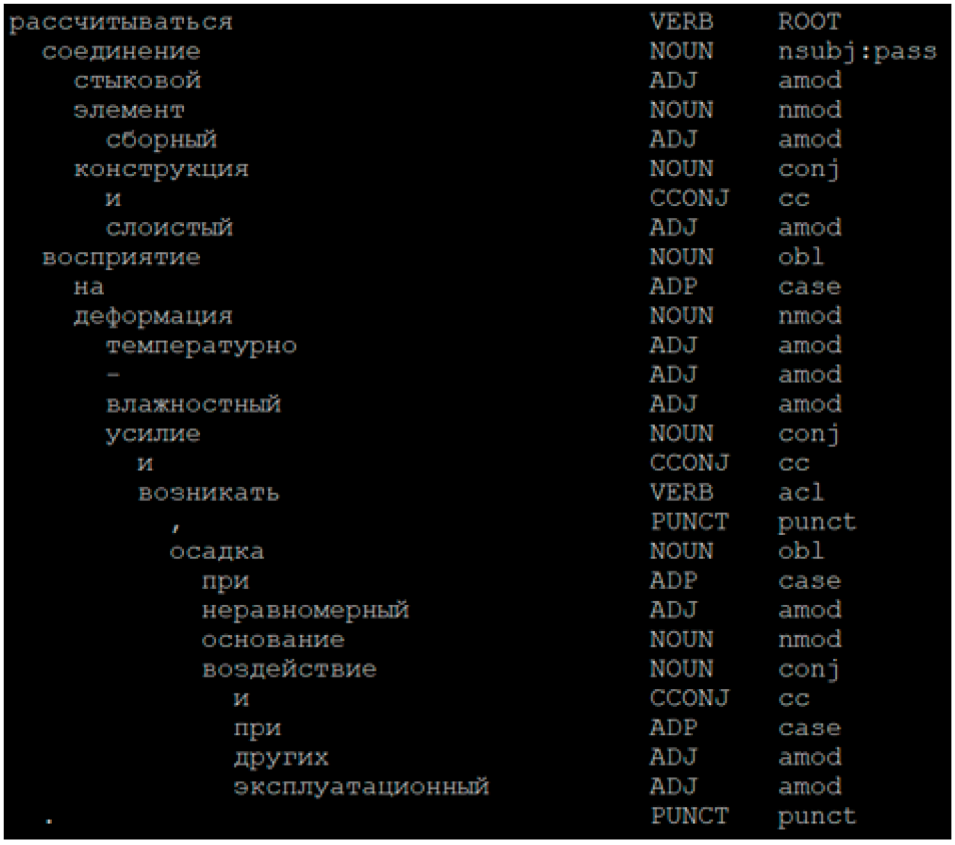


FIGURE 4
Output of lemmatization and grammatical structure parsing of the sentence.

TABLE 1 Comparison of «semantic portraits».

Subject	Connection	Connection
Predicate	Calculate	Calculate
Object	Perception	Perception
Modality	-	-
Negation	-	-
Definition	Joint, prefabricated, layered, temperature, moisture, irregular, others, operational	Joint, prefabricated, multilayer, thermal, irregular, others, operational

differences in definitions. The software calculates a semantic similarity metric ranging from −1 to 1, where −1 indicates completely opposite meanings and one indicates full equivalence. In the example considered, the metric value was 0.91, indicating a high degree of similarity between the phrases. By setting a threshold for this metric, it becomes possible to identify regulatory requirements that are duplicated either within a single document or across different regulatory sources. This confirms the applicability of the proposed methodology for the automated detection of redundant regulatory information (Colla et al., 2020).

When comparing semantic profiles, statements are considered equivalent only if they share the same predicate. Otherwise, the comparison result is set to zero. Negative metric values may occur in cases where the analyzed phrases differ in modality (e.g., *"may"* vs. *"shall"*) or when one of the statements includes predicate negation (e.g., *"is designed"* vs. *"is not designed"*).

One of the limitations of the basic algorithm is that it does not account for the semantic similarity of individual lexemes. As a result, phrases that are equivalent in meaning but differ in lexical composition may receive a semantic similarity score of zero. To address this issue, two possible approaches can be considered:

- Using vector-based models (e.g., Word2Vec, BERT), which enable the assessment of term similarity based on their contextual usage. However, for domain-specific texts, such models often demonstrate limited accuracy, as construction-related terms tend to be semantically close to each other, reducing the algorithm's discriminative capability.
- Shifting from lemmata to concepts using a SKOS-based ontological model, which makes it possible to account for hierarchical relationships between terms, such as equivalence, broader terms, and narrower terms. This approach enables more accurate computation of semantic similarity and allows for the identification of logical contradictions at a deeper level.

TABLE 2 Semantic portraits of phrases constructed from lemmas.

Elements of the semantic portrait	Proposition 1	Proposition 2
Subject	-	-
Predicate	Consider	Observe
Object	Process, purpose	Governance, interest
Modality	Need	Necessary
Negation	-	-
Definition	-	-
Complement/ Circumstance	Management, information, object, owner	Data, asset, owner

To demonstrate the advantages of using a conceptual model, let us consider two synthetic phrases:

- Sentence 1: "In the process of managing information about assets, it is necessary to consider the goals of their owners."
- Sentence 2: "During asset data management, the interests of their holders must be respected."

Despite the semantic equivalence of these statements, their lexical composition differs, which leads the lemma-based algorithm to assign them a semantic similarity score of zero (Table 2).

To overcome this limitation, a SKOS-based conceptual model was developed, incorporating the following terminological relationships:

- *"Consider"* = *"Respect"* (equivalent terms)
- *"Must"* = *"Necessary"* (equivalent terms)
- *"Holder"* = *"Owner"* (equivalent terms)
- *"Goal"* < *"Interest"* (narrower term)
- *"Information"* > *"Data"* (broader term)
- *"Asset"* < *"Object"* (narrower term)

This ontology was deployed in the Apache Fuseki system (Figure 5), which is accessed by EKG LP. During the processing of semantic profiles, the algorithm replaces lemmata with their corresponding concepts (Table 3), allowing for a more accurate calculation of similarity.

As a result of recalculation, the semantic similarity metric for the considered phrases was 0.48, reflecting their partial equivalence. In this approach, terms with broader or narrower meanings are interpreted as 75% matches, which enables the adjustment of the metric calculation algorithm accordingly.

To assess the scalability of the proposed approach, a preliminary evaluation of the lexical core of regulatory documents in the construction sector was conducted. The analysis was performed using Apache Tika to extract text from 14 regulatory documents provided by the client. Only Russian-language text was processed.

The sample used for the experimental evaluation comprised 14 regulatory documents with a total length of approximately 242,000 words, which is equivalent to an average industry-level regulatory corpus. Despite the representativeness of the content (the documents cover various aspects of construction regulation—design, operation, information modeling, etc.), this volume should be considered a pilot dataset suitable for initial testing of the proposed method's effectiveness.

From the standpoint of scalability, the evaluation conducted on this dataset made it possible to identify key characteristics of the algorithm's performance and confirm its applicability to real regulatory data. However, to ensure a high degree of generalizability and robustness of the method against variability in phrasing, structure, and lexical patterns, further expansion of the corpus is required.

Expanding the size of the training and test document sets, including a broader range of regulations (such as international standards, technical regulations, sanitary and fire safety codes), as well as covering documents with varying structural complexity, will enhance the validity of the obtained quality metrics. An extended corpus will make it possible to more accurately calibrate the parameters of semantic similarity, test the algorithms across a wider variety of contexts, and identify potential bottlenecks in the ontological model.

Thus, the effect of scaling lies not only in improving the reliability of the evaluation, but also in enhancing the ability of the developed method to adapt to new types of regulatory texts—an aspect that is critically important for its future practical application in the context of a constantly evolving regulatory landscape.

Based on lemmatization performed using Pymorphy2, the following quantitative characteristics were obtained:

- Number of unique lemmata: 9,400
- Percentage coverage by frequent lemmata: 91% of the total document text

A significant portion of regulatory documents employs a relatively limited set of key terms, which makes the task of constructing a SKOS-based conceptual model feasible.

Additionally, an analysis of the most frequently occurring words in the examined documents was conducted (Table 4). The ten most common lemmata account for 9% of the total text volume, indicating a high degree of lexical unification in regulatory documents. This observation supports the feasibility of effective conceptualization of industry-specific terminology, including the establishment of semantic frames and dependencies.

The analysis demonstrated that the use of semantic profiles in combination with a SKOS-based ontological model enables effective identification of duplicated regulatory requirements and assessment of their semantic similarity. The developed method also supports the detection of logical contradictions at the level of terms and their relationships.

The introduction of an ontological model in place of simple lemma comparison represents a fundamentally different level of text analysis. While lemmatization provides only superficial matching of word forms, the ontological model allows for the consideration of hierarchical and associative relationships between terms, their contextual roles, and their affiliation with specific concepts. This approach enables a deeper and more contextually grounded understanding of regulatory texts, which is critically important for the automated interpretation of requirements and the identification of logical relationships between document provisions.

```
12 SELECT * WHERE { skos:Information ?prop ?value }
13
```

Table | Response | 5 results in 0.163 seconds | Simple view ☐ Ellipse ☑ | Filter query results | Page size: 50

	prop	value
1	<http://www.w3.org/1999/02/22-rdf-syntax-ns#type>	<http://www.w3.org/2004/02/skos/core#Concept>
2	<http://www.w3.org/1999/02/22-rdf-syntax-ns#type>	<http://www.w3.org/2002/07/owl#NamedIndividual>
3	<http://www.w3.org/2004/02/skos/core#prefLabel>	"информация"@ru
4	<http://www.w3.org/2004/02/skos/core#narrower>	<http://www.w3.org/2004/02/skos/core#Data>
5	<http://www.w3.org/2000/01/rdf-schema#label>	"информация"@ru

FIGURE 5
Concept "Information" in the results of a SPARQL query to Apache Fuseki

TABLE 3 Substitution of lemmas for concepts in the "semantic portrait" of the EKG LP.

Elements of the semantic portrait	Proposition 1	Proposition 2
Subject	-	-
Predicate	skos:Consider	skos:Consider
Object	Process, skos:Goal	Management, skos:Interest
Modality	skos:Need	skos:Need
Negation	-	-
Definition	-	-
Complement/ Circumstance	Management, skos:Information, skos:Object, skos:Owner	skos:Data, skos:Asset, skos:Owner

TABLE 4 Ten most frequently occurring words and their frequency of use (excluding prepositions).

Terms	Quantity
Object	3 151
Construction	2,469
Work	2,455
Construction	2,441
Project	2,329
Informational	2,218
Building	2,100
Information	1,964
Model	1,787
System	1,685

The results obtained in the course of the study confirm that the construction of a conceptual (ontological) model of industry-specific terminology is a labor-intensive but feasible process that can significantly enhance the accuracy and completeness of automated analysis of regulatory documents in the construction sector.

5 Discussion

5.1 Analysis of identical and similar terminological definitions

To validate the proposed algorithm using practical examples, a series of experiments was conducted to identify duplicated and similar regulatory provisions in construction regulations.

The following documents were used as test materials:

– SP RK 1.02-120-2019 *"Application of Information Modeling in Construction Organizations"* (Zakon.kz, 2025);

– SP RK 1.02-121-2019 *"Application of Information Modeling in Operating Organizations"* (Zakon.kz, 2025);

These documents contain a significant number of identical or similar definitions and provisions, making them well-suited for analyzing the capabilities of the developed method.

Before conducting semantic analysis, the texts underwent preliminary processing aimed at removing elements that hinder the automated parsing of regulatory documents. At this stage, textual data were extracted from the original PDF files using the Apache Tika tool (Burgess and Mattmann, 2014), allowing them to be obtained in a structured format. Subsequently, auxiliary elements of the documents—including headers, page numbers, line breaks, titles, and other components not affecting the semantic content—were removed. After cleaning, the data were processed in the EKG LP environment, resulting in sets of semantic profiles of statements. The final processing step involved comparing the obtained semantic structures of the two documents at the lemma level, without applying the conceptual model.

The EKG LP algorithm successfully identifies matching definitions present in both documents. For example, during the analysis, it correctly detected a duplicated definition:

> "Stakeholder: A person, group, or organization that can affect, be affected by, or perceives itself to be affected by decisions, activities, or outcomes of a project."

However, due to the specifics of the SpaCy framework, variations in the grammatical structure analysis of the same phrase may occur. As a result, the similarity metric for comparable definitions does not always reach 100%. To improve accuracy, the implementation of an additional post-processing algorithm is proposed, which would take into account the sequence and set of words in the sentence. This would allow for more precise determination of textual identity.

5.2 Identification of similar statements in the text

In addition to terminological definitions, the algorithm also detects similar statements appearing in both documents. For example,: *"The information management function includes monitoring compliance with standards and requirements (the organizational standard for building information modeling technology, asset information requirements), monitoring the content and updating of the asset information model (AIM), and ensuring adherence to information approval and coordination procedures."*

The developed method demonstrated high computational efficiency. The following performance indicators were obtained during the experiments:

- Document preparation and grammatical structure parsing take less than 10 s per document.
- Comparison of statements between two documents is performed in less than 1 s.

Thanks to the high processing speed, it becomes feasible to implement a method for large-scale document comparison. In particular, a database of grammatical parsing results for regulatory acts can be created, enabling pairwise comparison of each document with all others to automatically identify duplicated and contradictory provisions.

The developed method for analyzing regulatory documents enables the task of automatic contradiction detection based on the comparison of semantic profiles of statements. In particular, inconsistencies may arise from differences in numerical values, mismatches in modality, or the presence of negation.

To illustrate, consider the following regulatory provision:

> "Tactile indicators serving a warning function on pedestrian pathway surfaces shall be placed no less than 0.8 m / 0.6 m from the information object or the beginning of a hazardous area, change in direction, entrance, etc."

In this case, the semantic profiles of phrases containing numerical characteristics will be nearly identical, with the only source of discrepancy being the difference in numerical values. Since the sentence structure parser marks such values as POS = NUM and dep = nummod, their comparison does not present technical difficulties and can be implemented as a specialized application module.

Another common type of contradiction is a difference in modalities. Consider the following examples:

- *"When designing the site, it is necessary to take into account the condition of natural landscape development."*
- *"When designing the site, it is recommended to take into account the condition of natural landscape development."*

Despite the similarity in the overall structure of the sentences, their semantic profiles differ due to the use of different modal operators: *"necessary"* and *"recommended."* During analysis, this results in a semantic similarity metric value of −0.98, indicating a high degree of semantic divergence between the statements. A similar result would be obtained in the case of opposing modalities (e.g., *"shall"*/ *"shall not"*).

The contradiction detection method is applicable to at least three types of discrepancies:

- Differences in numerical values within regulatory provisions;
- Mismatches in the modality of statements;
- Presence of negation that alters the meaning of the statement.

It should be noted that, despite the high level of automation, human intervention remains necessary at certain stages to improve the accuracy and reliability of the results. In particular, expert review helps to:

- Interpret the context of statements not captured by the algorithm;
- Clarify cases of terminological discrepancies related to industry-specific language;
- Determine the criticality of the identified inconsistencies.

Combining automated analysis with expert evaluation ensures more reliable and well-grounded detection of contradictions in regulatory documents. To effectively detect such inconsistencies, the algorithm uses a semantic similarity threshold calibrated on real-world data. For example, if the similarity of semantic profiles exceeds 60% but one of the listed discrepancies is detected, the statements are considered potentially contradictory and are flagged for further review.

5.3 Comparative analysis with existing methods

To assess the contribution of the developed method, a comparative analysis was conducted, evaluating its characteristics against other contemporary approaches used for automated analysis of regulatory documents. The following baseline solutions were selected:

- Text matching method based on TF-IDF and Jaccard similarity metric (Plansangket and Gan, 2015);
- Topic modeling using Latent Dirichlet allocation (LDA) (Subramanian et al., 2024);
- Sentence-BERT model, representing a modern class of transformer-based models for semantic similarity estimation (Westin, 2024);

TABLE 5 Comparative summary of methods.

No	Method	Precision	Recall	F1-score	Interpretability
1	TF-IDF + Jaccard	0,68	0,53	0,6	0,3
2	LDA (Topic modeling)	0,61	0,49	0,54	0,4
3	Sentence-BERT	0,87	0,8	0,84	0,6
4	Proposed method	0,89	0,84	0,86	0,9

– The proposed ontological method, which uses the formalization of statements in the form of semantic profiles and ontological relationships (Motta and Ladouceur, 2017).

The comparison was carried out using the following criteria (Table 5):

– Precision — the proportion of correctly classified duplicates and contradictions among all identified by the system;
– Recall — the proportion of actual duplicates and contradictions correctly detected by the system;
– F1-score — the harmonic mean of precision and recall;
– Interpretability — expert evaluation of the transparency of the algorithm's logic and the interpretability of its results (on a scale from 0 to 1).

The comparative analysis conducted showed that the proposed ontological method for semantic analysis of regulatory documents demonstrates high efficiency in identifying duplicated and contradictory provisions. According to standard quality metrics (precision = 0.89, recall = 0.84, F1-score = 0.86), the method is comparable to or outperforms existing solutions, including models based on Sentence-BERT, while having significantly higher interpretability (0.9 on the expert scale).

Unlike general-purpose text processing methods, the proposed methodology takes into account the specifics of regulatory documents: the presence of modal constructions, logical constraints, the subject structure of requirements, and the hierarchy of statements. The use of a semantically rich format for representing regulatory provisions in the form of "profiles" with ontological annotations not only enables the identification of semantic discrepancies but also provides a foundation for the automated logical verification of the consistency of regulatory requirements.

Thus, the developed approach can serve as the foundation for building intelligent systems for regulatory analysis support, providing both high-quality metrics and transparency in decision-making.

The comparison results confirm the relevance and practical significance of the proposed methodology in the context of the digitalization of the construction industry and the reform of the regulatory framework.

5.4 Practical recommendations for applying the developed method

To integrate the developed approach into regulatory analysis processes, project documentation expertise, and regulatory framework management in the construction industry, the following aspects of practical application should be considered:

– Use at the regulatory expertise stage. The method can be implemented as a supplementary tool in the regulatory document review process—to automatically identify duplicated and contradictory provisions between existing and proposed regulations. This is particularly relevant when developing new versions of standards, technical regulations, and departmental regulations.
– Support for the digital transformation of the regulatory framework. The proposed approach can be utilized within the digitalization of the regulatory and technical framework, including the construction of ontologically organized databases of building codes and their automated verification. This creates the foundation for transitioning from textual representation of requirements to their formalized, machine-readable structure.
– Integration into project documentation information systems (BIM). Integrating the developed method into software systems supporting building information modeling (BIM) technologies will allow for automatic verification of design decisions against current regulations at the early stages of design, helping to prevent regulatory conflicts. This is especially useful when generating automatic compliance reports for model requirements.
– Training and preparation of experts. To ensure effective implementation, it is recommended to develop training modules for specialists in technical standardization, expertise, and design, explaining the logic behind semantic profile construction, the principles of ontology formation, and the interpretation of analysis results.
– Support for the development of new regulations. The method can be applied during the regulatory drafting stage to compare draft documents with existing regulations, assess the consistency of provisions, and ensure the uniformity of terminology, particularly in cases where documents of different levels (state, industry, corporate) are functioning simultaneously.

The proposed methodology has a high degree of adaptability and can be integrated into the practices of various participants in the construction process—from regulatory bodies and expert centers to design and operational organizations. Implementing this approach will enhance the consistency of the regulatory framework, reduce the risks of regulatory contradictions, and support a more sustainable digital transformation of the construction industry.

6 Conclusion

This work proposes and validates a method for the automated analysis of regulatory documents in the construction industry, based on a combination of natural language processing (NLP) techniques and ontological modeling. The developed algorithm ensures the formation of semantic profiles for regulatory provisions, identification of duplicated and contradictory statements, and verification of the relevance of referenced documents.

The use of an ontocentric approach allows for the formalization of knowledge contained in regulatory documents and its integration into digital platforms for managing regulatory requirements. The developed methodology demonstrated its effectiveness through the analysis of the building codes of the Republic of Kazakhstan, showing its ability to identify logical inconsistencies, automate the classification of regulatory provisions, and compare requirements across different documents.

Experimental studies have confirmed the high computational efficiency of the developed algorithm, making it suitable for use in scalable regulatory document analysis systems. In particular, the processing speed of a single document does not exceed 10 s, and the comparison of regulatory provisions is completed in less than 1 s. This enables the implementation of a concept for large-scale comparative analysis of regulatory documents, identifying inconsistencies across large datasets of legal information.

Despite the achieved results, the algorithm requires further improvement. Key directions for future research include:

- Expanding the system's functionality to recognize the type of regulatory statements (definitions, requirements, notes, etc.);
- Automatic extraction of document structure (sections, articles, tables) to improve the processing of complex regulatory acts;
- Implementing semantic disambiguation algorithms using large language models (LLMs) to enhance text analysis accuracy;
- Integrating the system into the BIM ecosystem to ensure automated compliance control of design decisions with regulatory requirements.

The results of the study confirm that the use of ontological modeling combined with NLP methods is a promising direction for the automated analysis of regulatory documents. The developed method could serve as the foundation for creating intelligent Automated Compliance Checking (ACC) systems, supporting the digitalization of the construction industry and enhancing the efficiency of regulatory governance.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

Author contributions

ZK: Conceptualization, Validation, Writing – review and editing. AS: Conceptualization, Project administration, Writing – review and editing. SG: Investigation, Methodology, Writing – original draft. YY: Data curation, Visualization, Writing – original draft. FG: Software, Writing – original draft. NS: Formal Analysis, Resources, Writing – original draft, Writing – review and editing.

Funding

The author(s) declare that no financial support was received for the research and/or publication of this article.

Acknowledgments

The methods for creating "semantic profiles" were studied as part of the R&D project "Research on Methods for Automating Semantic Control of Term Definitions in Regulatory Documents of the Construction Industry" (JSC "KazNIISA," LLP "DataVera").

Conflict of interest

Authors SG and FG were employed by LLC Datavera.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Correction Note

A correction has been made to this article. Details can be found at: 10.3389/fbuil.2025.1624950.

Generative AI statement

The authors declare that no Generative AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Al-Turki, D., Hettiarachchi, H., Gaber, M., Abdelsamea, M., Basurra, S., Iranmanesh, S., et al. (2024). Human-in-the-Loop learning with LLMs for efficient RASE tagging in building compliance regulations. *IEEE Access* 12, 185291–185306. doi:10.1109/ACCESS.2024.3512434

Aslam, B., and Umar, T. (2023). Automated code compliance checking through building information modelling. *Proc. Institution Civ. Eng. - Struct. Build.* 177, 822–837. doi:10.1680/jstbu.22.00214

Beach, T., Rezgui, Y., Li, H., and Kasim, T. (2015). A rule-based semantic approach for automated regulatory compliance in the construction sector. *Expert Syst. Appl.* 42, 5219–5231. doi:10.1016/j.eswa.2015.02.029

Bilobrovets, I. (2023). Network threat detection technology using Zabbix software. *Mod. Inf. Secur.* 54. doi:10.31673/2409-7292.2023.020003

Burgess, A., and Mattmann, C. (2014). "Automatically classifying and interpreting polar datasets with Apache Tika," in Proceedings of the 2014 IEEE 15th International Conference on Information Reuse and Integration (IEEE IRI 2014), 13-15 August 2014, Redwood City, CA, USA. 863–867. doi:10.1109/IRI.2014.7051982

Chen, N., Lin, X., Jiang, H., and An, Y. (2024). Automated building information modeling compliance check through a large language model combined with deep learning and ontology. *Buildings* 14 (7), 1983. doi:10.3390/buildings14071983

City of Mesa (2012). Construction plan review. *Official Website City Mesa, Ariz.* Available online at: http://www.mesaaz.gov/devsustain/PlanReview.aspx (November 25, 2022).

Colla, D., Mensa, E., and Radicioni, D. (2020). Novel metrics for computing semantic similarity with sense embeddings. *Knowl. Based Syst.* 206, 106346. doi:10.1016/j.knosys.2020.106346

DataVera (2025). DataVera NLU: natural language understanding platform. Available online at: https://datavera.kz/ru/nlu.html February 11, 2025).

Díaz, H., Braumann, S., Van Der Poel, J., Gog, T., and De Bruin, A. (2024). Towards adaptive support for self-regulated learning of causal relations: evaluating four Dutch word vector models. *Br. J. Educ. Technol.* 55, 1354–1375. doi:10.1111/bjet.13431

Gao, H., and Zhong, B. (2022). A blockchain-based framework for supporting BIM-based building code compliance checking workflow. *IOP Conf. Ser. Mater. Sci. Eng.* 1218, 012016. doi:10.1088/1757-899X/1218/1/012016

Guo, D., Onstein, E., and La Rosa, A. (2021). A semantic approach for automated rule compliance checking in construction industry. *IEEE Access* 9, 129648–129660. doi:10.1109/ACCESS.2021.3108226

Ismail, A., Ali, K., and Iahad, N. (2017). "A Review on BIM-based automated code compliance checking system," in 2017 International Conference on Research and Innovation in Information Systems (ICRIIS), 16-17 July 2017, Langkawi, Malaysia. 1–6. doi:10.1109/ICRIIS.2017.8002486

Jadala, V., and Burugari, V. (2024). Artificial intelligence techniques in semantic web services. *Int. J. Innovation Multidiscip. Sci. Res.* 02, 42–49. doi:10.61239/ijimsr.2024.2220

Jiang, L., Shi, J., and Wang, C. (2022). Multi-ontology fusion and rule development to facilitate automated code compliance checking using BIM and rule-based reasoning. *Adv. Eng. Inf.* 51, 101449. doi:10.1016/j.aei.2021.101449

Kasim, T. (2015). *BIM-based smart compliance checking to enhance environmental sustainability*. PhD Thesis (Cardiff University).

Ke, J., Zacouris, Z., and Acosta, M. (2024). Efficient validation of SHACL shapes with reasoning. *Proc. VLDB Endow.* 17, 3589–3601. doi:10.14778/3681954.3682023

Li, S., Wang, J., and Xu, Z. (2024). Automated compliance checking for BIM models based on Chinese-NLP and knowledge graph: an integrative conceptual framework. *Eng. Constr. Archit. Manag.* doi:10.1108/ecam-10-2023-1037

Liu, X., Chen, Q., Wu, X., Hua, Y., Chen, J., Li, D., et al. (2020). Gated semantic difference based sentence semantic equivalence identification. *IEEE/ACM Trans. Audio, Speech, Lang. Process.* 28, 2770–2780. doi:10.1109/TASLP.2020.3030493

Mandal, S., Gandhi, R., and Siy, H. (2015). "Semantic web representations for reasoning about applicability and satisfiability of federal regulations for information security," in 2015 IEEE Eighth International Workshop on Requirements Engineering and Law (RELAW), 25-25 August 2015, Ottawa, ON, Canada. 1–9. doi:10.1109/RELAW.2015.7330205

Martinez-Gil, J., and Chaves-González, J. (2022). Sustainable semantic similarity assessment. *J. Intell. Fuzzy Syst.* 43, 6163–6174. doi:10.3233/jifs-220137

Motta, J., and Ladouceur, J. (2017). A CRF machine learning model reinforced by ontological knowledge for document summarization.

Nama, E., and Alalawi, A. (2023). "The adoption of automated building code compliance checking systems in the architecture, engineering, and construction industry," in 2023 International Conference On Cyber Management And Engineering (CyMaEn), 26-27 January 2023, Bangkok, Thailand. 289–296. doi:10.1109/CyMaEn57228.2023.10050930

Plansangket, S., and Gan, J. (2015). A query suggestion method combining TF-IDF and Jaccard Coefficient for interactive web search. *Artif. Intell. Res.* 4, 119–125. doi:10.5430/AIR.V4N2P119

Poniszewska-Marańda, A., and Czechowska, E. (2021). Kubernetes cluster for automating software production environment. *Sensors (Basel, Switz).* 21, 1910. doi:10.3390/s21051910

Preidel, C., and Borrmann, A. (2018). BIM-based code compliance checking. In *Building Information Modeling*. 367–381. doi:10.1007/978-3-319-92862-3_22

Salama, D., and El-Gohary, N. (2011). Semantic modeling for automated compliance checking. *Computing in Civil Engineering*. 641–648. doi:10.1061/41182(416)79

Salama, D., and El-Gohary, N. (2013). Automated compliance checking of construction operation plans using a deontology for the construction domain. *J. Comput. Civ. Eng.* 27, 681–698. doi:10.1061/(ASCE)CP.1943-5487.0000298

Subramanian, V., Prince, G., Venkateshwara, S., Thangakumar, J. S., and Preetha, M. (2024). "Similarities and ranking of documents using TF-IDF, LDA and WAM," in 2024 International Conference on Advances in Data Engineering and Intelligent Computing Systems (ADICS), 18-19 April 2024, Chennai, India. 01–07. doi:10.1109/ADICS58448.2024.10533526

Sydora, C., and Stroulia, E. (2020). Rule-based compliance checking and generative design for building interiors using BIM. *Automation Constr.* 120, 103368. doi:10.1016/j.autcon.2020.103368

Tang, K., Fei-Fei, L., and Koller, D. (2012). "Learning latent temporal structure for complex event detection," in 2012 IEEE Conference on Computer Vision and Pattern Recognition, 16-21 June 2012, Providence, RI, USA. 1250–1257. doi:10.1109/CVPR.2012.6247808

Tierney, P. J. (2012). A qualitative analysis framework using natural language processing and graph theory. *Int. Rev. Res. Open Distributed Learn.* 13, 173–189. doi:10.19173/IRRODL.V13I5.1240

Westin, F. (2024). Time Period categorization in fiction: a comparative analysis of machine learning techniques. *Cataloging & Classif. Q.* 62, 124–153. doi:10.1080/01639374.2024.2315548

Xue, X., and Zhang, J. (2022). Regulatory information transformation ruleset expansion to support automated building code compliance checking. *Automation Constr.* 138, 104230. doi:10.1016/j.autcon.2022.104230

Zakon.kz (2025). Online legislative database. Available online at: https://online.zakon.kz/Document/?doc_id=39721077&pos=1;-16#pos=1 February 11, 2025).

Zhang, J., and El-Gohary, N. (2013). Information transformation and automated reasoning for automated compliance checking in construction. *Computing in Civil Engineering*. 701–708. doi:10.1061/9780784413029.088

Zhang, J., and El-Gohary, N. (2015). Automated information transformation for automated regulatory compliance checking in construction. *J. Comput. Civ. Eng.* 29. doi:10.1061/(ASCE)CP.1943-5487.0000427

Zhang, J., and El-Gohary, N. (2016). Semantic NLP-based information extraction from construction regulatory documents for automated compliance checking. *J. Comput. Civ. Eng.* 30. doi:10.1061/(ASCE)CP.1943-5487.0000346

Zhang, J., and El-Gohary, N. (2017). Integrating semantic NLP and logic reasoning into a unified system for fully-automated code checking. *Automation Constr.* 73, 45–57. doi:10.1016/J.AUTCON.2016.08.027

Zhang, J., and El-Gohary, N. (2017). Semantic-based logic representation and reasoning for automated regulatory compliance checking. *J. Comput. Civ. Eng.* 31. doi:10.1061/(ASCE)CP.1943-5487.0000583

Zhang, R., and El-Gohary, N. (2021). A deep neural network-based method for deep information extraction using transfer learning strategies to support automated compliance checking. *Automation Constr.* 132, 103834. doi:10.1016/j.autcon.2021.103834

Zhang, Z., Ling, M., and Tim, B. (2022). "Towards fully-automated code compliance checking of building regulations: challenges for rule interpretation and representation," in Proceedings of the 2022 European Conference on Computing in Construction. doi:10.35490/ec3.2022.148

Zheng, Z., Zhou, Y., Lu, X., and Lin, J. (2022). Knowledge-informed semantic alignment and rule interpretation for automated compliance checking. *Automation Constr.* 142, 104524. doi:10.1016/j.autcon.2022.104524

Zhong, B., Luo, H., Hu, Y., and Sun, J. (2012). Ontology-based approach for automated quality compliance checking against regulation in metro construction project. 385–396. doi:10.1007/978-3-642-27963-8_35

Zhong, B. T., Ding, L. Y., Luo, H. B., Zhou, Y., Hu, Y. Z., and Hu, H. M. (2012). Ontology-based semantic modeling of regulation constraint for automated construction quality compliance checking. *Autom. Constr.* 28 (2012), 58–70. doi:10.1016/j.autcon.2012.06.006

frontiers | Frontiers in Built Environment

TYPE Correction
PUBLISHED 09 June 2025
DOI 10.3389/fbuil.2025.1624950

Check for updates

OPEN ACCESS

APPROVED BY
Frontiers Editorial Office,
Frontiers Media SA, Switzerland

*CORRESPONDENCE
Nazym Shogelova,
nazymshogelova@gmail.com

RECEIVED 08 May 2025
ACCEPTED 28 May 2025
PUBLISHED 09 June 2025

CITATION
Kabzhan Z, Shakhnovich A, Gorshkov S, Yemenov Y, Gorshkov F and Shogelova N (2025) Correction: Semantic and ontology-based analysis of regulatory documents for construction industry digitalization.
Front. Built Environ. 11:1624950.
doi: 10.3389/fbuil.2025.1624950

COPYRIGHT
© 2025 Kabzhan, Shakhnovich, Gorshkov, Yemenov, Gorshkov and Shogelova. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Correction: Semantic and ontology-based analysis of regulatory documents for construction industry digitalization

Zarina Kabzhan[1], Alexandr Shakhnovich[1], Sergey Gorshkov[2], Yussuf Yemenov[1], Fedor Gorshkov[2] and Nazym Shogelova[1]*

[1]JSC Kazakh Research and Design Institute of Construction and Architecture, Almaty, Kazakhstan, [2]LLC Datavera, Almaty, Kazakhstan

KEYWORDS

semantic analysis, ontology modeling, regulatory documents, digitalization of construction, NLP, automated analysis

A Correction on
Semantic and ontology-based analysis of regulatory documents for construction industry digitalization

by Kabzhan Z, Shakhnovich A, Shogelova N, Glyzno Y, Gorshkov S and Gorshkov F (2025). Front. Built Environ. 11:1575913. doi: 10.3389/fbuil.2025.1575913

In the published article, there was an error in the **Author list**, and author Yevgeniya Glyzno was erroneously included. And author Yussuf Yemenov was erroneously excluded. The corrected author list appears below.

"Zarina Kabzhan, Alexandr Shakhnovich, Sergey Gorshkov, Yussuf Yemenov, Fedor Gorshkov, and Nazym Shogelova."

The authors apologize for this error and state that this does not change the scientific conclusions of the article in any way. The original article has been updated.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Built Environment

TYPE Original Research
PUBLISHED 11 June 2025
DOI 10.3389/fbuil.2025.1607375

Check for updates

OPEN ACCESS

EDITED BY
Salman Azhar,
Auburn University, United States

REVIEWED BY
Rizwan Farooqui,
Mississippi State University, United States
Rana Muhammad Irfan Anwar,
Auburn University, United States

*CORRESPONDENCE
David Crampen,
crampen@gia.rwth-aachen.de

RECEIVED 07 April 2025
ACCEPTED 22 May 2025
PUBLISHED 11 June 2025

CITATION
Crampen D and Blankenbach J (2025) Data enrichment for semantic segmentation of point clouds for the generation of geometric-semantic road models. *Front. Built Environ.* 11:1607375. doi: 10.3389/fbuil.2025.1607375

COPYRIGHT
© 2025 Crampen and Blankenbach. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Data enrichment for semantic segmentation of point clouds for the generation of geometric-semantic road models

David Crampen* and Joerg Blankenbach

Civil Engineering Department, Geodetic Institute and Chair for Computing in Civil Engineering and GIS, RWTH Aachen University, Aachen, Germany

Digitalizing highway infrastructure is gaining interest in Germany and other countries due to the need for greater efficiency and sustainability. The maintenance of the built infrastructure accounts for nearly 30% of greenhouse gas emissions in Germany. To address this, Digital Twins are emerging as tools to optimize road systems. A Digital Twin of a built asset relies on a geometric-semantic as-is model of the area of interest, where an essential step for automated model generation is the semantic segmentation of reality capture data. While most approaches handle data without considering real-world context, our approach leverages existing geospatial data to enrich the data foundation through an adaptive feature extraction workflow. This workflow is adaptable to various model architectures, from deep learning methods like PointNet++ and PointNeXt to traditional machine learning models such as Random Forest. Our four-step workflow significantly boosts performance, improving overall accuracy by 20% and unweighted mean Intersection over Union (mIoU) by up to 43.47%. The target application is the semantic segmentation of point clouds in road environments. Additionally, the proposed modular workflow can be easily customized to fit diverse data sources and enhance semantic segmentation performance in a model-agnostic way.

KEYWORDS

point clouds, semantic segmentation, deep learning, machine learning, geometric-semantic road models, automation, digital twin, data enrichment

1 Introduction

The implementation of Digital Twins of roads has great potential in increasing the efficiency in maintenance and management of dense road infrastructure networks. A Digital Twin leverages multi-modal data to support and, in the future automize decision-making (Federal Ministry for Digital and Transport, BMDV Germany, 2021). Digital twins of roads can be extremely useful for many applications, such as predictive maintenance, traffic control or autonomous driving. However, for establishing a Digital Twin use case in the road environment, a digital geometric-semantic representation is required to act as a foundation for georeferencing sensors and relevant components in the environment and provide a container to store and interlink relevant data in the form of an as-is model (Noroozinejad Farsangi et al., 2024). To create such a representation, reality-capturing technologies are indispensable. The process from data acquisition to the final

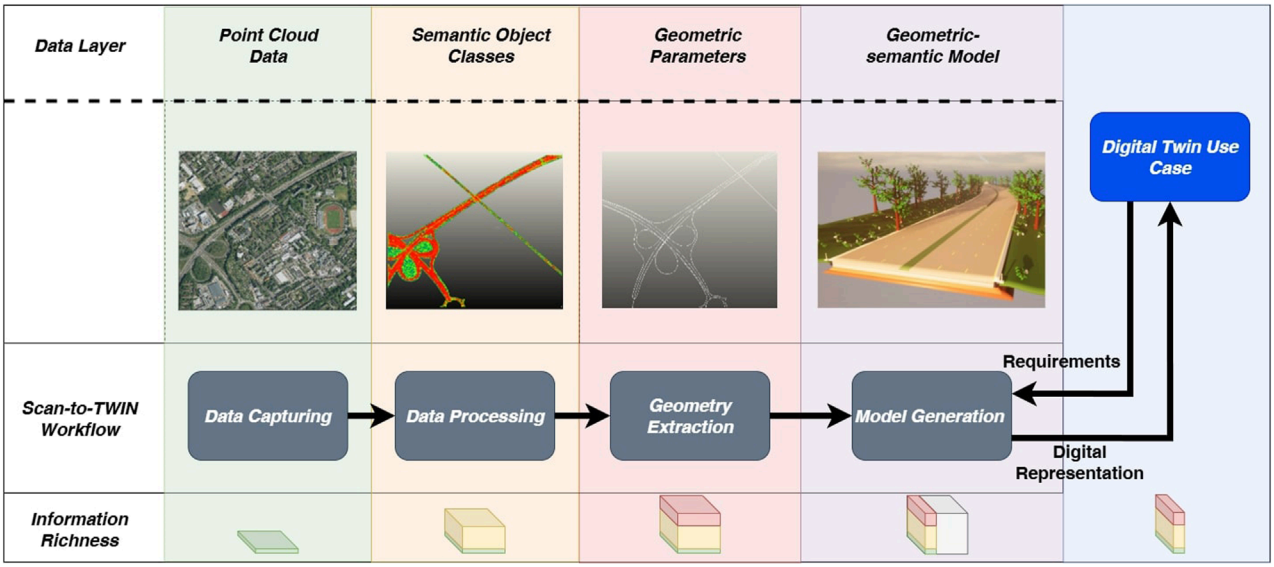


FIGURE 1
Overview scan-to-twin workflow.

digital as-is model is composed of several steps. Performing those manually is very time-consuming and linked to major modeling effort and costs. Therefore, automating this process has enormous time- and cost-saving potential. In Building Information Modeling (BIM), regardless whether automated or performed manually, the Scan-to-BIM process targets the built environment without an existing, up-to-date digital representation. While BIM aims at the aggregation of all relevant information about the asset and acts as an information container for all available data, this is not necessarily desired for a Digital Twin use case (Crampen et al., 2024). Irrelevant information is a burden for simulation and synchronization between the physical world and the virtual world. Therefore, the digital representation should only inherit the information necessary for the use case conducted. In that way, a use case defines requirements for the underlying digital representation. The resulting Scan-to-Twin workflow is schematically illustrated in Figure 1.

The workflow consists of a sequential chain of data acquisition, data processing, geometry extraction, and model generation, with the data processing step being our primary focus. This step is crucial as it introduces semantic information into the data collection, directly impacting the quality and robustness of the derived model. Since the accuracy of semantic object class segmentation greatly influences the detail and reliability of the resulting model, we leverage prior knowledge from geospatial data to enrich the data foundation and enhance the performance of various data-driven segmentation methods. Additionally, our workflow is designed with modularity in mind, making it adaptable to different data sources and supporting both georeferenced and unreferenced point cloud data. There are various techniques employed in the data processing step, which are introduced in Section 2.

The rest of this contribution is structured as follows: Section 2 reviews the state of research; Section 3 summarizes the identified research gaps and goals of our contribution; and Section 4 describes our workflow, which includes a uniform pre-segmentation step followed by an automatic feature extraction process that automatically adapts to the dataset. Section 5 presents our results. The impact of our preprocessing workflow is evaluated using different approaches by comparing model performance at each stage of the preprocessing scheme to highlight performance enhancements achieved in each consecutive step. In Section 6, we discuss our findings and possible future directions, particularly in linking our workflow to model generation for Digital Twins of roads, while we summarize our work in Section 7. The results of an ablation study conducted in the course of our analyses and visual examples of the final model predictions can be reviewed in the Supplementary Material.

2 State-of-research and related work

One major challenge for automation of the Scan-to-Twin process is its sequential nature, which leads to multiple interfaces and error propagation. This can lead to large errors in the output of the workflow, especially if inaccuracies occur in early steps (Esfahani et al., 2021). Consequently, relying on uncertain semantic objects may introduce intractable errors in the final geometric-semantic model (GSM) (Vassilev et al., 2024). Although there are several approaches to directly extract structures' geometries for indoor purposes, such as Martens and Blankenbach (2023), such approaches are difficult to apply in outdoor environments due to many natural objects and the large spatial extent, leading to extremely sparse voxel grids. Extracting the relevant objects through semantic segmentation in a first step, however, leads to a better foundation for geometry reconstruction in an outdoor environment (Grandio et al., 2022). Continuously, new deep learning architectures introduce new ways of capturing significant aspects of the point cloud data. Still, many works indicate, that most deep learning models meet a trade-off between local and

global context. This trade-off has been recognized since the first architecture working directly on point cloud data was published in Qi et al. (2017a). Since then, the complexity of new architectures has largely increased. Qian et al. (2022) indicate, that using better preprocessing techniques, the performance of less recent deep learning architectures can potentially exceed the performance of the latest and way more complex architectures. Introducing geospatial data (often also referred to as geodata or GIS data) to perform pre-segmentation as a preprocessing step for semantic segmentation could prove beneficial for the integration of prior knowledge and improve the perfomance of arbitrary models by increasing the data informativeness instead of the complexity of the model. While the available computational power has increased rapidly over the last decade, making complex deep learning architectures applicable, earlier methods did not rely on data, but on a model, formulating rules as well as mathematical filters based on heuristics and other methods to dissect data into characteristic segments. Today, methods for point cloud segmentation can be categorized into two groups: Point cloud segmentation (PCS) and point cloud semantic segmentation (PCSS). While the categorization might indicate that research is progressing from PCS to PCSS, there are several aspects currently limiting the broad application of PCSS for geometric-semantic model reconstruction, with a major factor being data availability. Therefore, significant work is conducted to develop hybrid approaches, as in Schatz and Domer (2024), to mitigate the respective issues arising in both sets of methods.

2.1 Point cloud segmentation (PCS)

In PCS, filters, rules, and static algorithms are commonly used to segment point clouds. Early approaches, were edge detection applied to range images (Bhanu et al., 1986) or region growing for surface approximation (Besl and Jain, 1988). Today similar methods are utilized for preprocessing, such as feature extraction, as demonstrated in recent studies (Grilli et al., 2017). Justo et al. (2021) used region growing with voxelized mobile laser scan (MLS) data to extract the centerlines of road segments in order to derive a digital model. Yan and Hajjar (2021) applied a region-growing algorithm to extract bridge girders from point clouds.

Various techniques, such as Hough Transform and RANSAC (Random Sample Consensus), are extensively used in fitting techniques to fit diverse shapes into point cloud data, enabling the extraction of objects with characteristic shapes. Safaie et al. (2021) utilized Hough Transform to extract tree stems near roads by vertically slicing the point cloud data and fitting circles to the vertical cross-sections of the stems. Li and Vosselman (2018) applied Hough Transform to extract pole-like objects, including road signs, traffic lights, and road lamps, from Mobile Mapping System (MMS) data. Wang et al. (2022) improved the Universal-RANSAC algorithm to perform ground filtering on MMS point clouds, simplifying the semantic segmentation of objects by robustly removing ground points.

Edge features, widely used in image segmentation, play a significant role in direct object extraction. The detection of changes in local curvature extends beyond images and applies to point clouds, as well as rough changes in normal orientation, gradients, intensity, or principal curvature within point clouds to indicate object transitions. In that way, Rato and Santos (2021) utilized edge detection in MMS point clouds for vehicle positioning in autonomous driving, while Wang et al. (2019) incorporated edge features in their artificial neural network architecture to enhance the ability of distinguishing different objects by constructing local neighborhood graphs. For extracting pole-like objects from MMS data, Li Y. et al. (2019) first applied elevation filtering and combined region growing with multiple filtering techniques and unsupervised learning to extract different object classes from an urban road environment, while Lucas et al. (2019), Soilán et al. (2017) and Ma et al. (2017) used reflectance-based features to filter vegetation road markings, respectively, leveraging information provided by most laser scanners.

2.2 Point cloud semantic segmentation (PCSS)

PCSS uses machine learning methods to classify point cloud segments into semantic classes, that involves supervised machine learning, semi-supervised as well as unsupervised machine learning approaches. Supervised machine learning can be further differentiated into conventional machine learning, where manual feature engineering has to be preliminary performed to fit a model to the feature set, and deep learning, where artificial neural networks learn abstract features themselves to allow a classification of different semantic classes. A classical machine learning model used for PCSS are Random Forests, which have been used for instance to segment pole-like structures from MMS data (Wang et al., 2021). Biçici and Zeybek (2021) applied random forest classifiers to airborne laser scan (ALS) data to extract road surfaces.

The unstructured nature of point clouds is taken into account by different deep learning approaches that attempt to structure the data or were designed to deal with the point cloud directly to form connections during training by passing groups of points to the neural network at once. Although methods employing spectral images and voxelization have been early approaches, there are several recent works dealing with outdoor point cloud segmentation, such as Xu et al. (2021). Another approach applied to outdoor point clouds is presented in Roynard et al. (2018a), where a voxel-based neural network was developed, achieving state-of-the-art performance on a terrestrial laser scanning (TLS) outdoor dataset. A popular approach leveraging spectral-images is SnapNet (Boulch et al., 2017), which mitigates the structuring issue by generating images from the meshed point cloud.

As opposed to the previously discussed methods, the deep learning architectures working directly on unstructured point clouds eliminate the problems introduced in the structuring process, retaining the information inherent to the point cloud and requiring less computational effort to form structures (Zhang et al., 2019). The first method in this domain was PointNet (Qi et al., 2017a), where a symmetric function is applied to order the point cloud and enable the use of multi-layer perceptrons to learn point-wise features that are aggregated to obtain global features for segmentation. For road segmentation, PointNet++ (Qi et al., 2017b) is often used. Ma et al. (2022) for example, used PointNet++ for road footprint segmentation, whereas in Shin et al. (2022) it was used for building extraction from ALS data. Recently, transformer-based networks

have been adapted to point clouds, using an attention mechanism, which was first introduced in the natural language processing domain, to outperform most other architectures (Zhao et al., 2021). There are many approaches being published every year that regularly set new measures in state-of-the-art performance on well-known benchmarking datasets such as ModelNet (Wu et al., 2015), S3DIS (Armeni et al., 2016) or Paris-Lille-3D (Roynard et al., 2018b) to name a few. Interestingly, in Qian et al. (2022) it is shown with their PointNext architecture that updating the training strategy and adding rather small modifications to the model architecture of PointNet++ can lead to a model outperforming most of the recent approaches, indicating that at some point the leverage of higher model complexity on actual performance becomes smaller than that of an improved training approach and potentially data preprocessing workflow.

2.3 Prior knowledge and feature extraction

Regarding preprocessing, especially encoding prior knowledge into the data directly has large potential for performance enhancement of arbitrary models. Li F. et al. (2019), for example, first split their MMS data into blocks and then removed ground points through local height variance as a first step of their workflow for road furniture extraction. A standard method is separating ground and above ground points to create two separate classification problems (Baligh Jahromi et al., 2011; Chen et al., 2021; Shi et al., 2018; Sithole and Vosselman, 2001; Yilmaz, 2021). In Murtiyoso et al. (2020), roofs were extracted using available geospatial data of building footprints. In Paden et al. (2022) ALS point clouds and geospatial data of building footprints were used to directly create CityJSON data by extruding the footprints by the height of the inlying point cloud segments. In a previous study (Murtiyoso and Grussenmeyer, 2019) geospatial data has been used to extract point cloud segments of heritage complexes and enrich the segments with the information present in the geospatial data. Aljumaily et al. (2019) used Open Street Map (OSM) data in the same way to provide segmented point clouds with richer semantic data. Lastly, in Park and Guldmann (2019) OSM data was used for building extraction using building footprint data to cut point cloud segments out of the point cloud.

Typical features used to enrich point cloud data are eigenvalue-based features (Bremer et al., 2013). Higher-order features can be derived from various transformations or preprocessing techniques that can be called descriptor-based features, like in Alexiou et al. (2024). Others involve spin images, shape distributions (Osada et al., 2002), or histograms (Rusu et al., 2009). In road environments, height gradient features can indicate the height change when reaching the road edge if it is bound by a curb, as done in Yang et al. (2020). Using the timestamps on MMS data, the road orientation can be retrieved. Voxel-based region-growing approaches use seed points as features, like in Hirt et al. (2021). The resulting segments can be used to extract other higher-order features, as done by Zhang et al. (2022).

Another significant aspect of using extracted features is the selection of the most meaningful ones. To do this heuristically, there are different solutions, such as Recursive Feature Elimination (RFE) Schlosser et al. (2020) or filter-based methods Atik and Duran (2022). Tan et al. (2007) developed a Genetic algorithm (GA) that heuristically finds the best set of features from a set of populations of features and improves the set iteratively through mechanisms of recombination, reproduction, and mutation. Compared to filter-based methods, GA is a model-agnostic wrapper-based approach. Therefore, it is applicable to arbitrary models using diverse fitness functions. In comparison to RFE, which is also wrapper-based, the advantage of GA is the capability of conducting a global search, while RFE works greedy and therefore more efficient, which however may lead to suboptimal feature sets. Road segmentation on point clouds involves various approaches from both PCS and PCSS, primarily utilizing data captured from MMS. However, MMS are designed to capture the near field of a road within the scanner system's field of view, while many use cases require more extensive coverage beyond what an MMS system can capture. Connecting multiple data sources during the capturing stage can leverage synergies and enhance connections between different domains, including urban planning, traffic simulation, and risk management. However, combining different point cloud data sources can be challenging due to heterogeneity in density, features and scan patterns that make it harder for a model to generalize well (Hu et al., 2023). Still, establishing a data source-agnostic workflow that maximizes the quality of the semantic segmentation is necessary to enable automatic model generation. Especially since models for Digital Twins have to stay up-to-date and repeated manual modeling in relatively short time intervals is simply not feasible (Mansour et al., 2024).

3 Gaps and research questions

The research gaps addressed within this contribution are as follows: (1) Development of a strategy to enrich point cloud data of roads for semantic segmentation. (2) Efficient leveraging of prior knowledge to improve the output of an arbitrary machine learning approach. (3) Generalized data enrichment scheme to automatically adapt to different point cloud data sources. Therefore, this contribution aims to answer the following research questions:

1. How can reality capture data of roads be enriched for semantic segmentation purposes without limiting the applicability to a specific data source? (RQ1)
2. What prior knowledge can be leveraged to further enhance information richness of point cloud data of roads? (RQ2)
3. What methods can be combined to improve the overall fit of point cloud data of roads to machine learning applications without limiting optimization to a specific model? (RQ3)
4. What is the best approach to make the workflow automatically adapt to changes in data source and data quality? (RQ4)

This contribution introduces a workflow specifically designed for performing pre-segmentation and preprocessing for Point Cloud Semantic Segmentation (PCSS) applications on large-scale laser scanning data. Through our automatic preprocessing workflow, we showcase improvements in various PCSS models, including standard machine learning methods such as bagging and boosting methods for decision trees, as well as deep learning-based techniques like PointNet++ (Qi et al., 2017b) and PointNeXt (Qian et al., 2022). This improvement is achieved while significantly reducing the required amount of data and training time.

By combining PCS and PCSS with prior knowledge, we leverage the abundant ALS point cloud data provided by governments as open data. Our approach combines automatic adaptation to input data with a modular structure, enabling the flexible application of our process to different point cloud related tasks. This distinguishes our method form other task specific solutions to point cloud semantic segmentation approaches which are often tailored to specific datasets. With this strategy, we aim at ensuring ease of use and a high degree of generalizability for different point cloud related tasks. We validate this claim by comparing the model performances on ALS data as well as MMS data.

4 Methodology

Our workflow consists of four steps, which can be executed sequentially or individually, based on the characteristics of the available data (Figure 2):

1. Prior knowledge in the form of existing geospatial data is utilized to perform pre-segmentation of roads from ALS data by generating horizontal boundaries for the area of interest.
2. Existing geospatial data is again used to identify road blocks along the trajectory of the roads. These blocks are then transformed to a uniform orientation, which improves the regularity of the input for feature extraction.
3. Various feature extraction methods are employed to generate an initial set of features.
4. A feature selection method based on a custom GA is applied to reduce the initial feature space to a more condensed set of features. This selection process helps to improve the efficiency and effectiveness of the dataset.

The workflow is structured into two phases: pre-segmentation and feature generation (Figure 2). This design allows evaluation of performance enhancements achieved by applying each phase independently. It also enables the use of either phase separately to accommodate data that is not georeferenced and, therefore, not suitable for Phase 1. While our workflow combines different approaches and adaptations from other domains, we want to emphasize that the focus of this contribution is not merely investigating the performance of deep learning models. However, we test the preprocessing scheme outputs on various models from deep learning and machine learning as a proof of the versatility and universal impact of our preprocessing. Specifically, we use PointNet++ and the more recent PointNeXt as proxies to show the impact of our data enrichment process for deep learning approaches, while training decision tree-based models such as Random Forest and simple multi-layer perceptrons (MLP) as a comparison.

4.1 Dataset

The data used for our workflow is open data accessible from "geoportal.nrw" (Interministerial Committee, 2025). This platform provides ALS point clouds and further geospatial data specifically for the North Rhine-Westphalia region in Germany. The ALS point cloud data has an average density of approximately 20 points per square meter. It contains several scanner features, of which intensity and number of returns are most relevant for road segmentation. The number of returns is captured by partial reflections of the same beam that are registered multiple times, indicating a transparent or semi-transparent surface, while the intensity is the strength of the reflection received, which differs for surfaces with different reflectivity characteristics. Both the ALS point cloud data and the geospatial data are georeferenced in the official ETRS89/UTM32 coordinate reference system of Germany, ensuring spatial alignment between the datasets. Since the ALS point clouds did not include RGB data initially, we employ a simple spatial interpolation technique using a photogrammetric point cloud of the same region to assign color values to it. In this work, we focused on extracting primary and secondary roads, including highways and federal roads. This selection criterion helped to narrow down the scope of the segmentation task. To demonstrate the effectiveness of Step 1, the pre-segmentation process, we observed a substantial reduction in the number of points in each segment. On average, applying Step 1 reduced the amount of data by 95%, narrowing down the scope to the area of interest. Two examples of the twelve used tiles are shown in Figure 3.

The MMS dataset used was captured using a vehicle-mounted mapping system Trimble MX9 by the german road agency Straßen.NRW and comprises small scenes of highways close to bridges and shield gantries in the State of North-Rhine-Westafalia. The average density of this dataset is approximately 700 points per square meter. It consists of localized coordinates and intensity values, while lacking color information. To create the dataset, we manually labeled 37 of those section sections, where 32 were used for training 2 for validation and 3 for testing. For the annotation of the semantic object classes of both datasets, the open source software CloudCompare was used.

4.2 Phase 1: pre-segmentation

For coarse pre-segmentation of the road space from the ALS data, we apply an approach comparable to the idea presented in Murtiyoso et al. (2020), but apply existing geospatial data of the road network to extract the road environment as a region of interest, instead of cutting using building footprints to cut out buildings. While different file formats such as GML, XML or ESRI Shapefile are applicable, we tested our approach on OSM data and German ATKIS (the Official Topographic-Cartographic Information System) data, which is generally more precise than OSM data. One advantage of using the above-mentioned geospatial data is its nationwide availability. The Shapefile from OSM and ATKIS consist of simple polylines and semantic information, which is attached to each polyline. We use the functional classification in the data to determine respective maximum road widths for different road classes in order to create polygons from the polylines by buffering, and then utilize these to automatically slice out our area of interest.

4.2.1 Step 1: cutting point clouds with geospatial data

The data of the whole road network of NRW is cut automatically by using the boundaries of the point clouds to first filter the geospatial data and then apply the buffered polygons to the point

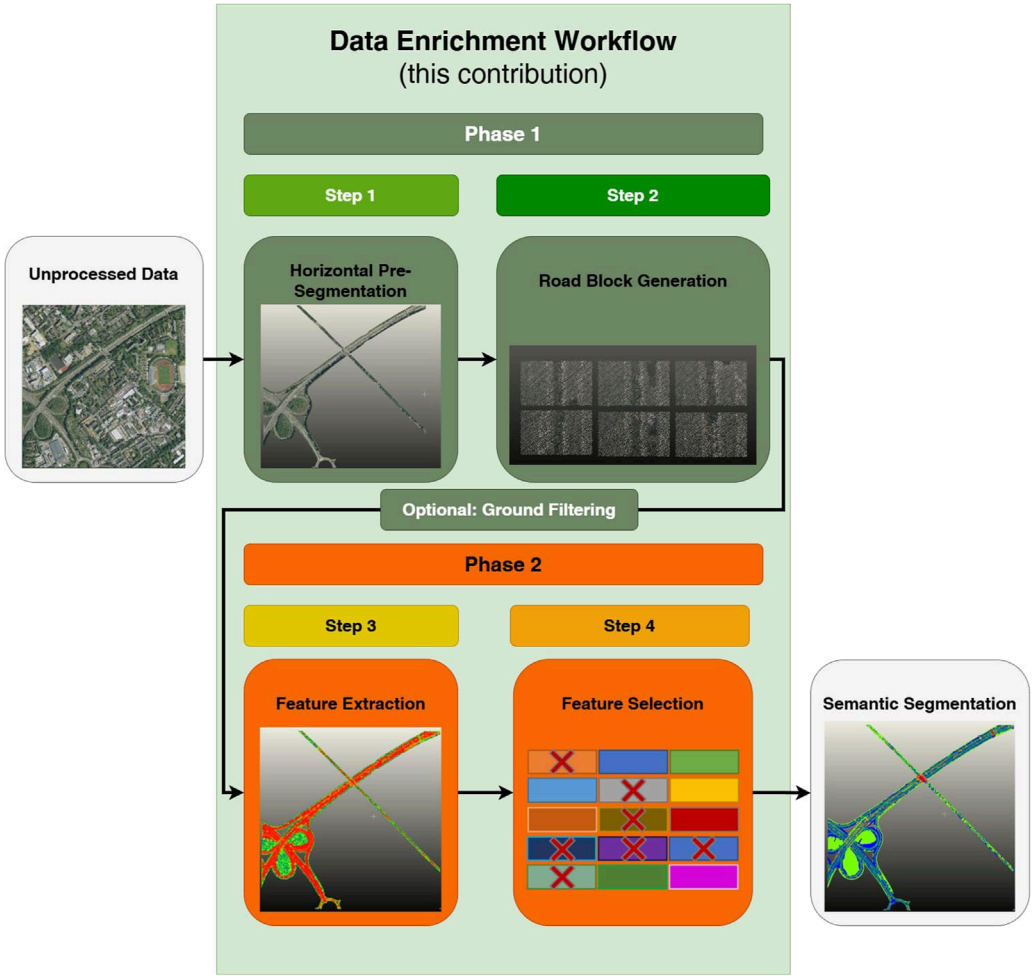


FIGURE 2
Overall workflow.

clouds, which are provided in tiles of 1 square kilometer. This is schematically shown in Figure 4 and largely reduces search time for matching geospatial data and point clouds.

The point clouds are cut using a path-cutting algorithm. After the horizontal pre-segmentation, we utilize a cloth simulation filter (CSF) to separate ground and above ground points. The ground and above-ground points are then separately treated for the rest of the workflow. For ground points, the classes for segmentation are road, non-road ground and road markings, while the above ground points are segmented into road furniture, vehicles and vegetation.

4.2.2 Step 2: road block generation

One advantage of using road-related geospatial data is the possibility to determine the road orientation, which can be used to process data quite similar to the trajectory information obtained using MMS data timestamps. We sample an arbitrary number of points on the polyline and use the predecessor-successor relation of these points to construct rectangular blocks along the road segment with sample point distance and road width as input parameters. This allows us to further extract smaller road segments and rotate them uniformly around the height axis to obtain homogenous road blocks. This procedure enables the extraction of features along the road trajectory. Figure 5 schematically shows the procedure of Step 2.

The road blocks are used in the second Phase 2 to extract additional histogram-based features. After the features have been extracted from the individual blocks, the point cloud is reconstructed by rotating the road blocks back to their original positions. When using the road block approach, we add the localized Northing and Easting UTM coordinates as features and create bins of points with uniform thickness along the principal axis as a basis for the histograms. In each bin, we compute the mean intensity and the intensity gradient over the bins as additional features. When visualizing the intensity gradients of the ground points along the road width, a characteristic pattern can be recognized, indicating a good potential for separation of lanes and road markings in one direction. Figure 6 shows three examples of histograms of the intensity gradients and their corresponding road segments.

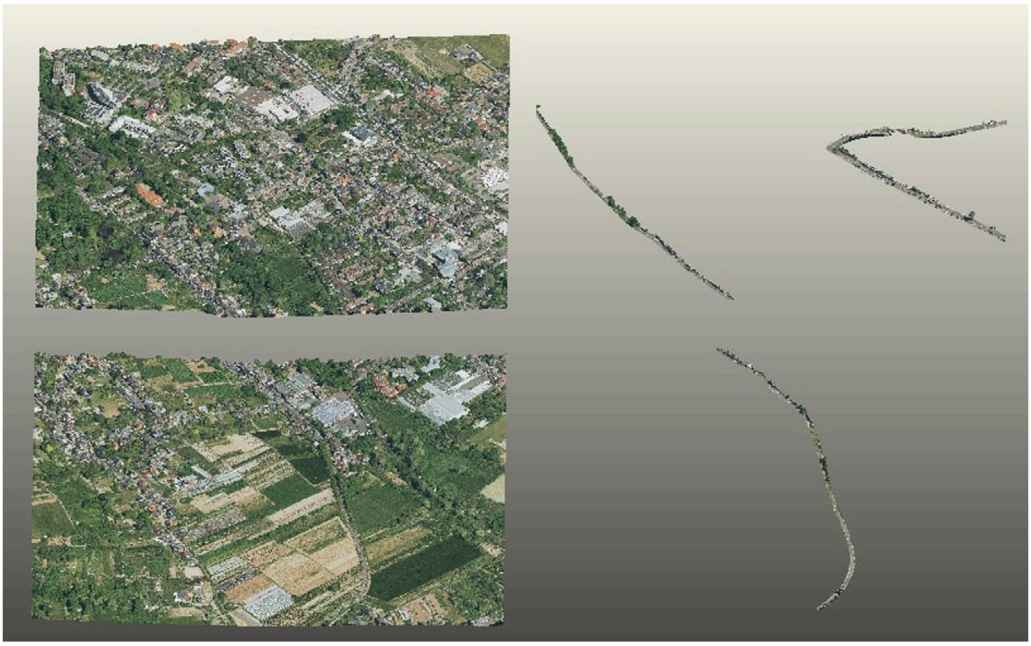


FIGURE 3
Example of two point clouds, raw point clouds on the left, pre-segmented point clouds on the right.

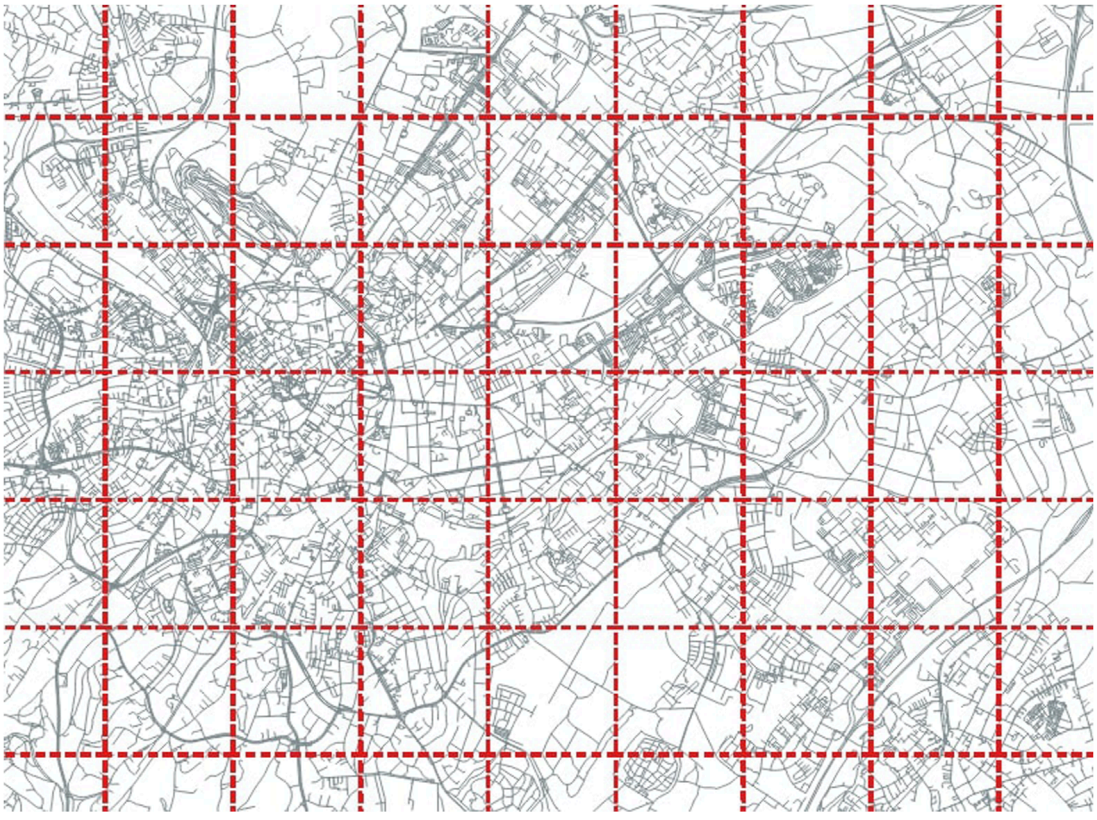


FIGURE 4
Tesselating geospatial data by point cloud boundaries to optimize matching efficiency for pre-segmentation.

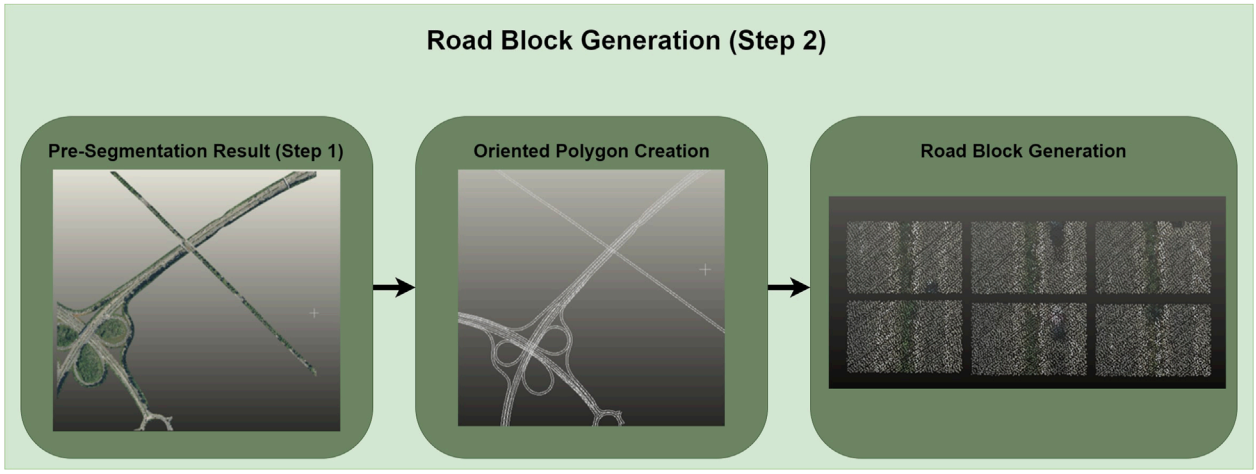


FIGURE 5
Overview of step 2 road block generation.

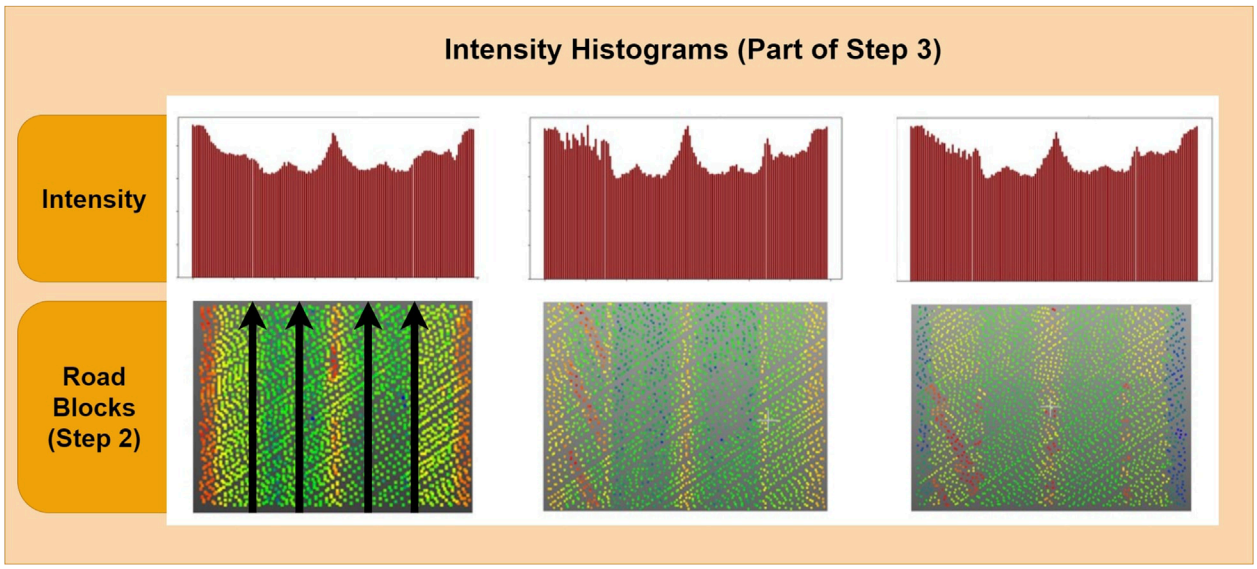


FIGURE 6
Intensity gradients along the road width axis of three different road blocks.

4.3 Phase 2: feature generation

In the second phase, we extract features from the point cloud. Preliminary analyses showed that extracting all features directly from road blocks introduces unnecessary edges to the receptive fields, and thereby results in less informative features than if applying the feature extraction to the whole point cloud. We first extract features and optimize them by varying the parameters of the applied approaches for extraction, which is validated using a simple multi-layer perceptron classifier (MLP) that is trained with ten manually labeled road segments as a training set and measuring the accuracy reached with the given features on a validation set consisting of two road segments. After this, we employ a custom GA for feature selection, thereby reducing the initially large feature set of up to 86 features depending on the used steps to a more condensed set of features. The choice of GA results from our aim towards the model agnostic design of our workflow. We validate the feature selection step by training an MLP on the new set of features using the same training split as in the previous step to verify whether model performance can be maintained using fewer features.

4.3.1 Step 3: feature extraction

The extracted features are grouped into: eigenvalue-based features, reflectance-based features, height-based features, that are computed by applying a cylindrical neighborhood search and computing the local distribution of height and other basic

TABLE 1 Initial feature set grouped by extraction module, and area of application.

Feature	Extraction module	Application area
XYZ (mult.)	Position	Global
RGB (mult.)	Colors	Global
Number of Returns	Reflectance	Global
Intensity	Reflectance	Global
Planarity	Eigenvalue	Local Neigh.
Linearity	Eigenvalue	Local Neigh.
Sphericity	Eigenvalue	Local Neigh.
Verticality	Eigenvalue	Local Neigh.
Mean Height	Local Distribution	Local Neigh.
Height Standard Dev.	Local Distribution	Local Neigh.
Mean Intensity	Local Distribution	Local Neigh.
Intensity Standard Dev.	Local Distribution	Local Neigh.
Height Range	Local Distribution	Local Neigh.
Number of Neighbors	Local Distribution	Local Neigh.
Eigenvalue Ranges (mult.)	Local Distribution	Local Neigh.
Height Difference to Ground	Ground Relational	Local Neigh.
Custom Height Features (mult.)	Ground Relational	Local Neigh.
DB Mean Height	DBSCAN	Segment
DB Standard Dev. Height	DBSCAN	Segment
DB Mean Intensity	DBSCAN	Segment
DB Standard Dev. Intensity	DBSCAN	Segment
DB Custom Shape features (mult.)	DBSCAN	Segment
Eigenvalue Segment features (mult.)	DBSCAN	Segment
Ransac Plane	RANSAC	Segment
Ransac Cube	RANSAC	Segment
Ransac Sphere	RANSAC	Segment
Ransac Cylinder	RANSAC	Segment
Ransac Line	RANSAC	Segment
Ransac Circle	RANSAC	Segment
Intensity Gradients X and Y (mult.)	Road Blocks	Bins along X and Y
Mean Intensity	Road Blocks	Bins along X and Y
Intensity Gradients X	Road Blocks	Bins along X
Intensity Gradients Y	Road Blocks	Bins along Y
Gaussian Mixture model	GMM	Point-wise
K-Means	K-Means	Point-wise

(Continued on the following page)

TABLE 1 (*Continued*) Initial feature set grouped by extraction module, and area of application.

features, segment-based features, which are obtained by using unsupervised approaches such as generalized Random Sample Consensus (RANSAC) for model fitting as well as Density-Based Spatial Clustering of Application with Noise (DBSCAN). We apply shape constraints to the resulting segments to obtain segment-based shape descriptor features. For the segment features computed with RANSAC, we simply encode the recognized shapes, which are: cube, plane, cylinder, circle, line and sphere, as binary features, for belonging or not belonging to the inliers of a fit shape. Lastly, we use a Gaussian Mixture Model and k-Means clustering to cluster points into six clusters based on the previously extracted features and add the cluster labels to the point-wise features. Table 1 shows the full list of extracted features.

All extracted features either directly or indirectly depend on the receptive field that is steered by the neighborhood radius for neighborhood-based features and the maximum distance in the case of DBSCAN-based features. Therefore, we conducted an ablation study by varying the parameters controlling the receptive field for feature extraction. We compute the features in six separate groups to find a trade-off between computational expense and flexibility in feature optimization. Optimizing the reception radius for each single feature is the upper boundary, and optimizing one uniform radius for all features is the lower boundary of the trade-off. The resulting groups are named as follows: Eigenvalue-features, LocalDist-features, DBSCAN-features, RANSAC-features, Ground-Relation-features and RoadBlock-features, where the second last group is obviously only applicable to the above ground points after ground filtering the results of the ablation study conducted are detailed in the Supplementary Material Section. For the RoadBlock-features, we create bins with a fixed width along the direction of travel, which is denoted with Y, as well as orthogonal to the direction of travel, denoted with X to assign points inside each bin to intensity gradients over consecutive bins. A schematic visualization of this procedure is shown in Figure 7.

Since varying the radius to obtain the best features on a full dataset is time consuming, we also compare the manual reception radius optimization with the use of the optimal neighborhood

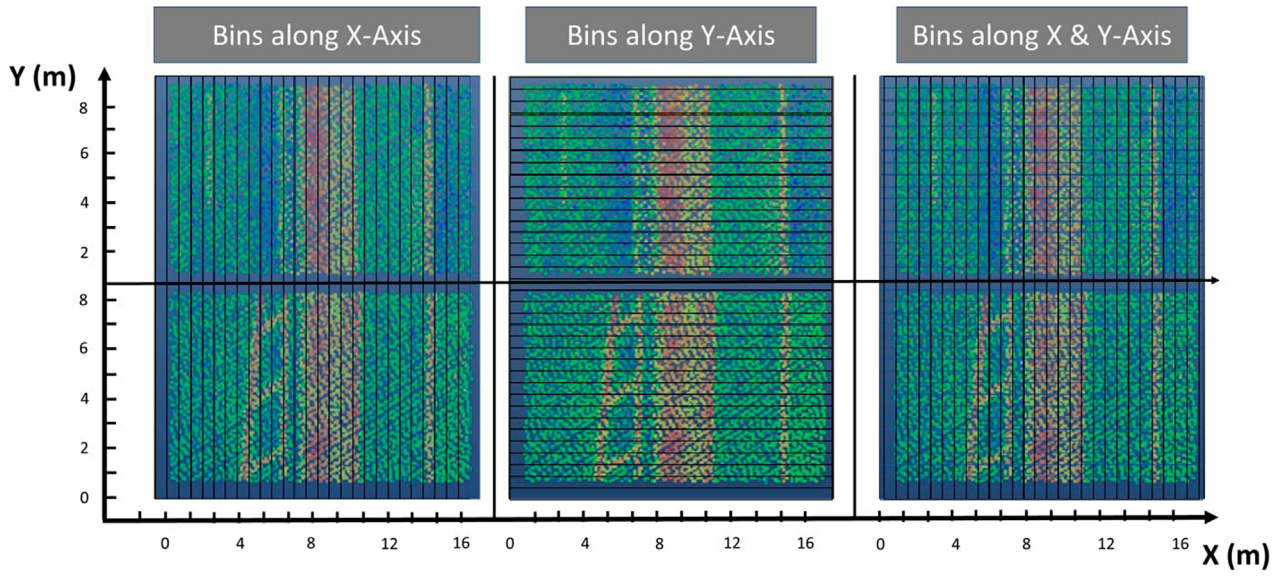


FIGURE 7
Bins along axes for feature extraction on ground points.

definition, which can be used to automatically find optimal k for k-nearest neighborhood search per-point. We compute the optimal neighborhood according to Weinmann et al. (2014) as shown in the equations below Equations 1–3:

$$E_\lambda = -e_1 ln(e_1) - e_2 ln(e_2) - e_3 ln(e_3) \tag{1}$$

$$e_i = \frac{\lambda_i}{\sum \lambda}, \; i \in \{1,2,3\} \tag{2}$$

$$\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq 0 \tag{3}$$

However, instead of using k-nearest neighborhood, we use radius neighborhood in a range between 1 m and 7 m, with step sizes of 0.5 m, to being able to pass the identified reception radius to other feature extraction modules for improved adaptiveness. The segment-based features using DBSCAN rely on a criterion that depends on point cloud density. Therefore, we choose Epsilon according to a fraction of the average number of neighbors computed in the optimal neighborhood approach. To increase variability of the receptive field, we also used the k-nearest neighborhood definition for additional feature extraction, with an inverse relation compared to radius neighborhood, increasing k for small radius neighborhoods.

4.3.2 Step 4: feature selection

For feature selection, we utilize a custom GA that iteratively searches for the best feature set. We additionally compare the performance of the reduced feature set to the performance of the initial feature set to test the impact of reducing the feature set by employing Step 4. Our custom GA, is configured to run 10 epochs, aggregating a population of 20 feature sets with a specified number of features and using 3-fold cross-validation of a Naïve Bayesian classifier (Huang and Li, 2011) on random subsamples of 100,000 points with a matching class distribution of the full dataset for each fold. The results are ranked according to their unweighted mIoU and the first five ranking feature sets are transferred to the next epoch, while ten new feature sets are generated by randomly combining the five best feature sets from the previous epoch. The number of recombining sets is reduced in each epoch to make the optimization converge. We additionally allow random mutation with a probability of 5% in all features of a new set, limiting the maximum number of mutations to 10% of the feature count, while choosing the remaining 5 sets from again randomly sampling features from the full feature set. After each epoch the best feature set is validated on the full validation set to ensure that the result is representative.

4.4 Evaluation strategy

To evaluate the impact of the stages of our workflow, we compare different configurations of our workflow. The resulting six configurations are depicted in Figure 8. The colored markers highlight the most relevant stages to clarify the overall performance boost achieved, where gray denotes the baseline dataset and green denotes the overall highest performance gain on all metrics over several different models.

For the optimizations in Phase 2, we use a simple multi-layer perceptron for initial evaluation. The final configuration of each stage is then tested on Random Forest, AdaBoost, XGBoost, MLP, PointNet++ and PointNext-b to get a robust estimation for the impact of the specific configuration. The ALS dataset consists of 12 manually labeled point clouds, where segments 0–9 are used for training and segments 10 and 11 are used for validation. Though this dataset is very small, we show that through our preprocessing regiment, the models are still able to generalize to new data. We labeled 3 additional point cloud segments that were not involved

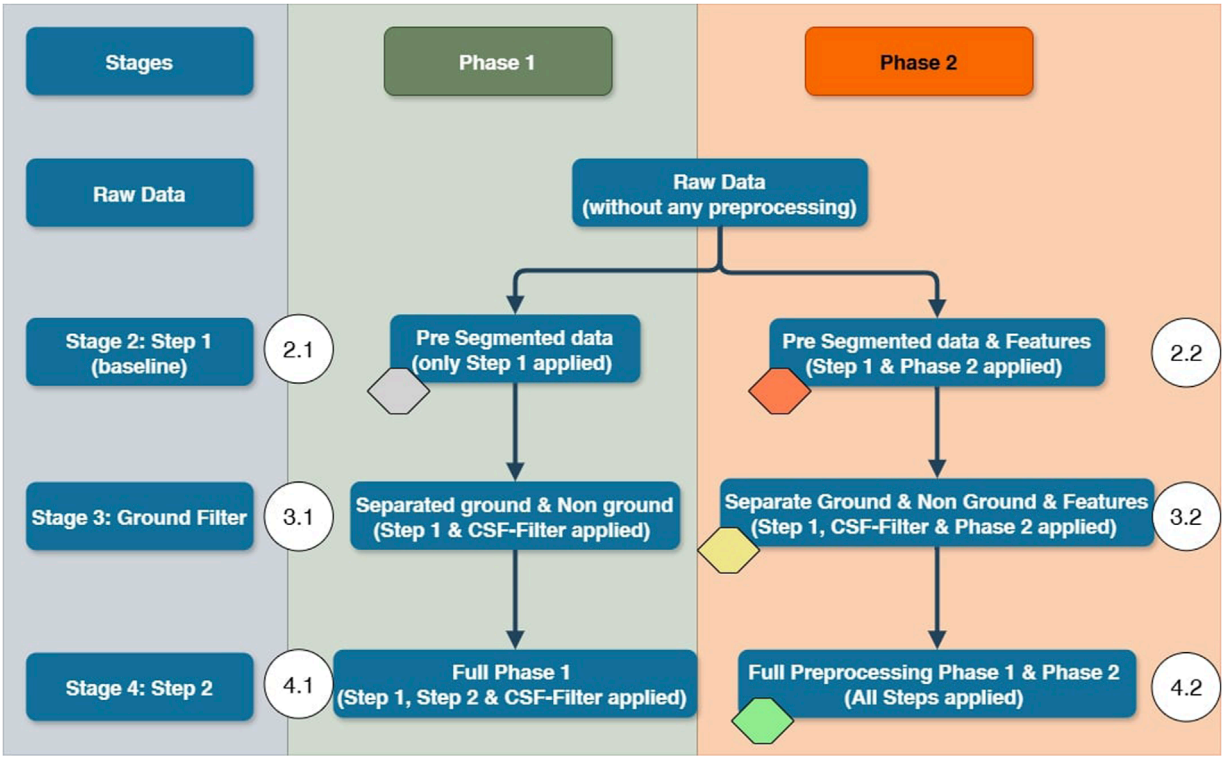


FIGURE 8
Overview of evaluation strategy.

in any of the previous steps as test data. After determining the best strategy on the ALS data, we apply the same workflow to the MMS dataset and evaluate the performance boost on three of those segments. We structured the evaluation in a way that we can compare different configurations directly to being able to interpret the results of each consecutive step. Improvements are reported with respect to the models trained on stage 2.1, since comparing different data foundations would be pointless. Configuration 2.1 only employs Step 1 and does not consider ground filtering. In configuration 3.1, we evaluate the impact of treating ground and non-ground points separately. In configuration 4.1, we use the full scheme of Phase 1. We split each stage into two leaves, employing our Phase 2, consisting of feature extraction and selection, in the second leaf while leaving it out in leaf one. This procedure allows for a detailed evaluation of the different steps, leading to a better understanding of the differences in data as well as learning methods. We train a PointNeXt-b model for validation of the performance-boosting potential on more recent architectures in Stages 2.1, 2.2, 3.2, and 4.2.

5 Results

In this section, we present the results of our comparison of the different stages of our preprocessing workflow. The results of our ablation study showed that using optimal neighborhood outperforms manual receptive field optimization while significantly reducing the computation time. Additionally, the measures taken to make the feature extraction self-adaptive with respect to the data distribution based on the optimal neighborhood approach benefit the degree of automation of the workflow. We compare the results of our workflow evaluation strategy based on overall accuracy (OA) (Equation 4), precision (pr) (Equation 5) and recall (rec) (Equation 6), weighted mean intersection over union (mIoU) (w) (Equations 7, 8) and unweighted intersection over union (unw. mIoU) (u) (Equation 9) to better measure performance changes for minority classes. Those metrics are calculated by first counting true positives (TP), true negatives (TN), false positives (FP) and false negatives (FN) and applying the following formulars, where n is the number of samples, i denotes the class and N denotes the number of classes. The difference between weighted mIoU and unweighted mIoU is that the number of samples per class affects the weighted metric; therefore, in cases where there is a large class imbalance, the unweighted IoU better captures improvements of minority classes, while the weighted mIoU better measures the overall performance.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{4}$$

$$Precision = \frac{TP}{TP + FP} \tag{5}$$

$$Recall = \frac{TP}{TP + FN} \tag{6}$$

$$IoU = \frac{TP}{TP + FN + FP} \tag{7}$$

$$mIoU = \frac{\sum_i n_i * IoU_i}{\sum_i n_i} \quad (8)$$

$$unweighted\ mIoU = \frac{\sum_i IoU_i}{N} \quad (9)$$

5.1 Results - preprocessing workflow

In this section, we discuss the results of our six workflow configurations on the six chosen models. The results for the full feature set are shown in Table 2. For PointNet++, we reduced the input point number to 2048 points due to the low point density of the ALS dataset in configurations 2.1 and 2.2, to 1,024 points for ground points in 3.1, 3.2, 4.1, and 4.2, and to just 512 points for above ground points. The block size was extended to 10 m for all configurations. 4,096 points were used for the MMS data. The remaining major hyperparameters of PointNet++ were chosen according to the default values:

- Learning rate 0.001
- Step size: 10
- Optimizer: Adam
- Learning rate decay: 0.7
- Weight decay rate: 0.0001

The model was trained for a maximum of 200 epochs and varying batch sizes due to the large inputs for full feature sets. The batch size ranged from 64 for the 2.1, 3.1, and 4.1 configuration down to 8 for the 2.2, 3.2, and 4.2 configuration.

For our future target application of model generation, the most important classes are: road, non-road ground, road markings and road furniture. Since we focus on the road structure, the classes "vehicle" and "vegetation" mainly fulfill the purpose of separating the important classes from the less important objects in the road environment. For our PointNeXt model, we chose the default configuration, except for the voxel sampling size of 0.04 and an initial radius of 0.3 compared to the default of 0.1. We used the relatively light-weight PointNeXt-b model for fast training time. The number of points per class in the training dataset are significantly imbalanced where the majority class (road) has approximately 3 million points whereas the minority class (road furniture) only comprises 63,000 points. The number of points per class for the ALS dataset are denoted beneath the class names in Table 2, whereas those for the MMS dataset are shown in 3. The majority class of the MMS dataset is the road class, while the minority class is the non road ground class. The class imbalance is representative for validation and test sets for both datasets. No additional measures for compensating class imbalance were taken during the training of any of the models.

To clarify the significant improvements over the stages, we elaborate the performance differences between stages at the example of PointNeXt. Starting with the baseline dataset (highlighted in gray) the most significant performance metrics are bold. Adding the Phase 2 preprocessing (2.2) (orange), the overall accuracy increases by +9.70%, while unweighted mIoU increases sharply by +21.49%. Adding ground separation in stage 3.2 (yellow), while only adding +0.97% in overall accuracy, stage 2.2 is strongly outperformed regarding mIoU with a performance gain of additional +11.04% unweighted mIoU. This performance is matched and partially outperformed in stage 4.2 (green) when using the road block approach, adding another +2.68% to the overall accuracy and increasing the weighted mIoU by +5.08%, while only slightly losing on unweighted mIoU (−0.13%) compared to stage 3.2. This consistently increasing performance throughout the steps of our approach is mirrored by all other models as well and totals a performance gain against the baseline of +13,35% OA, +15.16% weighted mIoU and +32.40% in unweighted mIoU for PointNeXt. Additionally, if only the stages without feature extraction are taken into comparison (2.1, 3.1, and 4.1), the models still show a consistent performance gain, in unweighted mIoU, e.g., for Random Forest with 16.90% from 2.1 to 3.1 and another 6.80% from 3.1 to 4.1 (highlighted bold). In a way, one could argue that the steps involving only the first phase of preprocessing are data cleaning and homogenization steps, which enable better feature aggregation. Limiting receptive fields in the edge areas of road blocks and adding small error portions through erroneously separated points during ground filtering corresponds to lowering the upper boundary of the achievable performance. However, these steps add a small bias, which averaged at approximately 2% of points in the ALS dataset, compared to the significant boost of model performances when applying feature extraction to their resulting point cloud subset.

It can be noted that the results for the machine learning models on stage 2.2 could be significantly improved. Specifically, the minimum performance gain over all four machine learning models (MLP, Random Forest, XGBoost, AdaBoost) exceeded 11.4% of unweighted mIoU for the XGBoost model, while Random Forest improved by 27,25% in the same metric comparing 2.1 to 2.2, even surpassing PointNet++ performance of stage 2.2 by 4.69% unweigthed mIoU. PointNet++ shows no significant change in performance until stage 3.2, where the unweighted mIoU improved by 10.87% compared to stage 2.1. It becomes obvious that on the ALS dataset, the PointNeXt model easily outperforms PointNet++ in stages with feature generation.

Generally, we assume that the low sensitivity of PointNet++ to the preprocessing workflow on the ALS dataset is due to the low point density, leading to low variability in data when using the block-wise data input approach of PointNet++. Here it becomes obvious that the change made in PointNext, directly running on the full scene similar to Randla-Net (Hu et al., 2020), is better suited for sparse data than sampling blocks with a fixed number of points in them. We validate this assumption by comparing results of the ALS dataset with those of the MMS dataset, which provides a higher point density in the next section.

5.2 Results - MMS dataset comparison

The MMS dataset was first randomly subsampled the sections to a point density of 120 points per square meter. The object classes differed slightly compared to the ALS dataset lacking the car class, since dynamic objects had been removed beforehand. The individual sections covered approximately the same lateral area as the ALS data after step 1 of our workflow and were not georeferenced. We, therefore, focus on evaluating the impact of our phase 2 for the comparison, demonstrating the flexibility of our pipeline. Since

TABLE 2 Results of our evaluation strategy for all used models on ALS data.

St.	Mod.	OA	w	u	N. R. Gr.		Road		R. Mark.		Veg.		Car		R. Frunit.	
			Num	points	2166946		3057436		387451		766310		109877		63069	
					Pr	Rec	Pr	Rec	Pr	Rec	Pr	Rec	Pr	Rec	Pr	Rec
2.1 (baseline)	MLP	71,45	55,58	31,2	0.73	0.81	0.74	0.94	0.38	0.12	0.81	0.48	0.08	0.01	0	0
	RF	79,91	66,54	**36,0**	0.78	0.8	0.86	0.93	0.64	0.28	0.75	0.83	0.34	0	0	0
	AB	74,48	59,35	31,4	0.74	0.82	0.72	0.92	0.17	0.0	0.93	0.55	0.26	0.04	1.00	0
	XGB	70,35	54,26	35,6	0.77	0.72	0.75	0.84	0.67	0.43	0.72	0.53	0.08	0.47	0.46	0.14
	PN++	88,65	76,49	58,55	0.9	0.7	0.88	1.00	0.79	0.56	0.88	0.7	0.76	0.57	0.68	0.34
	PNX	**82,53**	**76,61**	**48,40**	0.85	0.97	0.87	0.97	0.57	0.00	0.94	1.00	0.71	0.04	0.99	0.24
2.2	MLP	83,86	72,22	52,19	0.89	0.78	0.88	0.94	0.58	0.61	0.6	0.97	0.81	0.49	0.30	0.19
	RF	87,83	78.30	63,25	0.93	0.85	0.87	0.98	0.68	0.51	0.96	0.91	0.79	0.65	0.33	0.76
	AB	84,87	71,68	49,04	0.91	0.87	0.93	0.96	0.52	0.68	0.75	0.46	0.65	0.53	0.1	0.68
	XGB	85,67	74,93	47,01	0.9	0.82	0.92	0.95	0.54	0.71	0.83	0.88	0.72	0.36	0.19	0.61
	PN++	88,39	76,25	58,56	0.81	0.95	0.97	0.91	0.71	0.5	0.98	0.97	0.57	0.53	0.12	0.50
	PNX	**92,23**	**84,72**	**69,89**	0.86	0.86	0.95	0.98	0.94	0.76	0.99	0.68	0.66	0.50	0.85	0.94
3.1	MLP	78,64	64,75	45,26	0.87	0.72	0.75	0.95	0.47	0.12	0.9	0.78	0.68	0.87	0.48	0.01
	RF	81,26	68,48	**53,1**	0.83	0.77	0.92	0.92	0.35	0.51	0.84	0.87	0.76	0.74	0	0
	AB	79,94	67,63	52,82	0.81	0.88	0.92	0.91	0.66	0.31	0.76	0.64	0.67	0.67	0	0
	XGB	78,49	65,41	51,56	0.83	0.81	0.88	0.93	0.47	0.36	0.84	0.55	0.68	0.79	0.05	0.12
	PN++	88,99	70,35	54,97	0.86	0.88	0.95	0.94	0.69	0.69	0.96	0.98	0.44	0.17	0.49	0.03
3.2	MLP	84,85	73,73	65,08	0.94	0.59	0.90	0.98	0.48	0.69	0.95	0.89	0.57	0.55	0.63	0.92
	RF	92,91	86,81	81,10	0.89	0.87	0.96	0.98	0.78	0.71	0.99	0.94	0.77	0.89	0.79	0.96
	AB	89,85	81,57	70,79	0.9	0.89	0.91	0.96	0.78	0.55	0.98	0.91	0.84	0.95	0.74	0.47
	XGB	91,38	84,13	71,75	0.89	0.94	0.95	0.96	0.78	0.49	0.98	0.91	0.86	0.91	0.55	0.68
	PN++	86,96	81,48	69,42	0.89	0.84	0.96	0.93	0.60	0.81	0.81	0.99	0.9	0.38	0.88	0.93
	PNX	**93,20**	**86,69**	**80,93**	0.92	0.83	0.97	0.99	0.81	0.88	1.00	0.70	0.81	0.97	0.92	0.96
4.1	MLP	87,62	78,09	61,55	0.9	0.9	0.89	0.98	0.89	0.3	0.89	0.92	0.82	0.77	0.32	0.23
	RF	87,2	77,31	**59,9**	0.92	0.86	0.86	0.98	0.7	0.3	0.89	0.93	0.86	0.79	0.56	0.59
	AB	85,64	74,88	61,83	0.92	0.79	0.85	0.98	0.56	0.46	0.95	0.89	0.83	0.8	0.41	0.76
	XGB	86,96	76,97	62,57	0.92	0.86	0.88	0.98	0.59	0.4	0.94	0.91	0.86	0.72	0.33	0.87
	PN++	87,31	76,06	61,32	0.86	0.9	0.98	0.91	0.58	0.68	0.77	0.68	0.9	0.84	0.23	0.59

(Continued on the following page)

TABLE 2 (*Continued*) Results of our evaluation strategy for all used models on ALS data.

St.	Mod.	OA	w	u	N. R. Gr.		Road		R. Mark.		Veg.		Car		R. Frunit.	
			Num	points	2166946		3057436		387451		766310		109877		63069	
					Pr	Rec	Pr	Rec	Pr	Rec	Pr	Rec	Pr	Rec	Pr	Rec
4.2	MLP	86,55	76,35	65,15	0.88	0.69	0.92	0.98	0.54	0.63	0.98	0.89	0.50	0.70	0.58	0.62
	RF	92,35	85,82	80,38	0.86	0.89	0.95	0.98	0.86	0.65	0.99	0.93	0.71	0.9	0.84	0.96
	AB	88,37	79,17	70,57	0.88	0.86	0.91	0.96	0.63	0.48	0.96	0.89	0.82	0.94	0.89	0.62
	XGB	89,57	81,11	69,97	0.93	0.86	0.95	0.96	0.53	0.67	0.95	0.9	0.85	0.86	0.52	0.77
	PN++	85,07	71,05	61,82	0.82	0.95	0.98	0.91	0.74	0.61	0.85	0.81	0.48	0.59	0.78	0.54
	PNX	**95,88**	**91,77**	**80,80**	0.93	0.87	0.96	1.00	0.94	0.82	0.99	0.99	0.92	0.77	0.62	0.91

Random Forest, MLP, PointNet++ and PointNeXt had previously shown to be most interesting for further comparison, we tested only with these four models.

From the results in Table 3, it becomes obvious, that the higher point density has a positive effect on the performance of PointNet++ and PointNeXt. It can be noted, that in comparison to the ALS dataset, where the intensity is much noisier than in the MMS dataset, the road markings were far better extracted, while the distinction between road surface and the lateral non-road area remained the main challenge for models trained on the MMS dataset. The performance of PointNet++ strongly increases over each consecutive stage, more than doubling the unweighted mIoU between stage 1.1 and stage 3.2, adding over 43%. PointNeXt delivered strong performance even in Stage 1.1, without any preprocessing. Still, the feature extraction in Stage 2.2 could improve the unweighted mIoU by over 5%, while additionally adding 5.62% in overall accuracy. Another interesting finding being demonstrated through the MMS dataset is that the machine learning models could be strongly improved, more than doubling unweighted mIoU and adding more than 20% overall accuracy in the case of Random Forest from stage 1 to stage 2.2 and over 14% for MLP. Still, both were significantly outperformed by the deep learning models in stage 3.2 regarding unweighted mIoU.

5.3 Results - feature selection

After demonstrating the performance enhancement potential of our work, we want to briefly outline the potential for reducing the computation time demand of our workflow, focusing on training time in this subsection. It is to show how the model performance changes when using feature subsets assembled via our GA. For the sake of computation speed, we chose a simple naive Bayesian classifier for validation with the unweighted mIoU as decision metric. We compared four different numbers of features 20, 35, 55 and the full feature set for each of the five datasets from configurations 2.2, 3.2 ground, 3.2 non-ground, 4.2 ground and 4.2 non-ground. The results of the comparison are depicted in Figure 9. We, additionally, used the resulting feature subsets to train an MLP for a second comparison, which is depicted in Figure 10. We used a fixed model configuration for the MLP, specifying the following parameters:

- Hidden layer size: 500
- Activation: ReLu
- Optimizer: Adam
- Maximum iteration: 5

Comparing the two different classifiers, the Bayesian classifier seems to improve for smaller feature sets, which might be the case due to it assuming uncorrelated features while many of the extracted features are in fact correlated. Overall, it can be noted that our feature selection scheme works well in reducing feature space while maintaining performance. The comparison between the datasets of the different stages also indicates the performance boost of each consecutive stage of our workflow. The performance of the Bayesian classifier additionally shows that through our automatic workflow, even simple models become applicable to the segmentation task.

With these results, it becomes clear that, while we are extracting a high number of features in every stage, the amount of information we encode into individual features may still increase in higher stages. This might be due to improved receptive fields with reduced heterogeneity and more characteristic neighborhoods being aggregated. This interesting finding could lead to a significant reduction in precomputation effort by introducing a learning scheme for feature relevance directly during computation. The potential future research topics will be discussed in more detail in Section 6.

6 Discussion

Through our extensive evaluation of the proposed preprocessing workflow, we could show that it significantly boosts performance on versatile data foundations. Sparse as well as dense point cloud data in the road environment is consistently enriched and thereby far better suited for arbitrary machine learning and deep learning models. A notable achievement is the finding that enhancing point

TABLE 3 Results of stages 1, 2.2 and 3.2 on MMS dataset comparing PointNet++, PointNeXt-b, MLP and Random Forest.

St.	Mod.	OA	w	u	N. R. Gr.		Road		R. Mark.		Veg.		R. Frunit.	
			Num	points	2343695		40666973		1953186		3415971		5968978	
					Pr	Rec	Pr	Rec	Pr	Rec	Pr	Rec	Pr	Rec
1	MLP	78,16	64,16	33,51	0.00	0.00	0.87	0.97	0.39	0.27	0.42	0.27	0.62	0.63
	RF	74,94	59,92	28,67	0.00	0.00	0.84	0.93	0.23	0.07	0.65	0.23	0.49	0.64
	PN++	61,64	44,55	42,54	0.44	0.06	0.78	0.67	0.13	1.00	0.93	0.67	0.78	0.95
	PNX	90.66	93.33	85.12	0.87	0.84	1.00	0.99	0.97	1.00	0.74	0.85	0.94	0.96
2.2	MLP	92,20	85,52	66,96	0.19	0.63	0.99	0.94	0.76	0.84	0.86	0.84	0.90	0.93
	RF	95,65	91,67	66,72	0.98	0.05	0.97	1.00	0.99	0.68	0.84	0.90	0.94	0.92
	PN++	94,74	90,01	66,43	0.62	0.93	0.97	0.99	0.95	0.60	1.00	0.25	0.94	0.99
	PNX	96.04	97.52	90.29	0.94	0.90	0.99	1.00	0.93	0.83	0.97	0.96	0.97	0.98
3.2	MLP	86,66	76,72	59,81	0.72	0.56	0.92	0.97	0.70	0.43	0.71	0.72	0.86	0.85
	RF	88,24	79,06	60,84	0.92	0.44	0.90	1.00	0.87	0.37	0.73	0.84	0.91	0.84
	PN++	93,26	88,05	86,01	0.95	0.60	0.91	0.99	0.92	1.00	1.00	0.99	0.95	0.98
	PNX	97.23	94.42	87.59	0.96	0.97	0.97	0.99	0.97	0.57	0.98	0.96	0.96	0.98

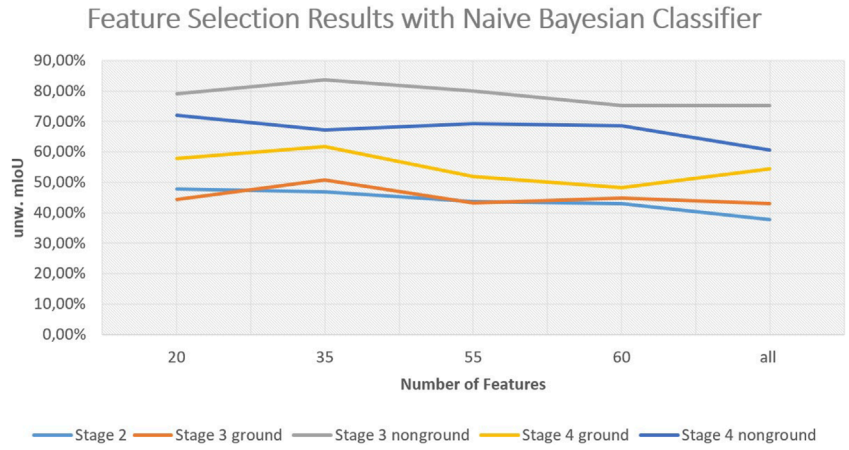


FIGURE 9
Results of GA for naive Bayesian classifier.

cloud data in this way makes models more receptive to imbalanced class distributions, which otherwise limits the applicability of unspecialized machine learning approaches. While the sparsity may still be a factor for specific models, working on fixed block sizes to sample input data, models that mitigate this issue by using either full scenes to generate input or even work with point-wise input take a large profit of our workflow. Therefore, the main limitation of our approach is the computation time needed for dataset preprocessing. We measured the computation time for the processing of Stage 2 of the ALS dataset averaging to 53 min per 1 million points, which results in a total processing time of approximately 6 h and 30 min (GPU: Nvidia Quadro 5000, CPU: Intel i9-11900K, RAM: 64 GB). It has to be mentioned that the preprocessing at the current stage is purely Python-based and will be accelerated in the future by moving to the C++ programming language. This aligns with the strategies of many point cloud processing

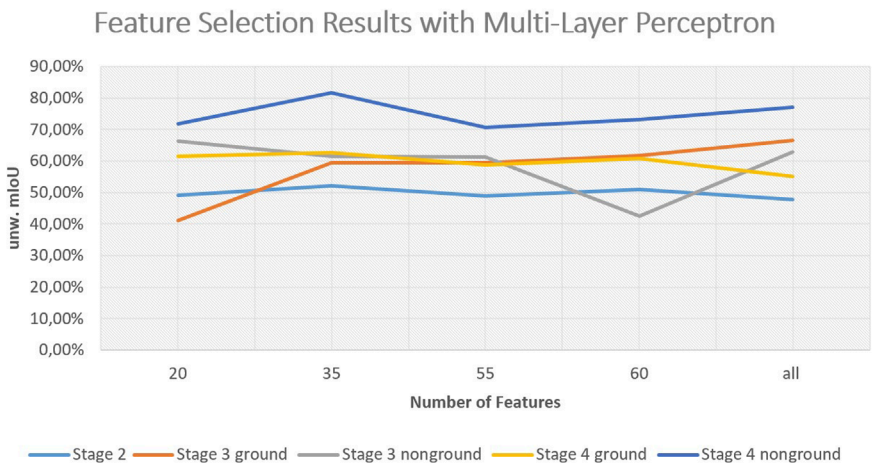


FIGURE 10
Results for feature selection with a baseline multi-layer perceptron.

neural network architectures, where preprocessing frameworks are typically implemented in C++ to accelerate the steps before model training as done in (Qian et al., 2022) with their openpoints library. However, our aim is the optimization of semantic segmentation for further use in geometric-semantic model generation. For geometric-semantic model generation from segmented point clouds, it is essential to lay a robust and as accurate as possible semantic foundation in order to enable automatic and sound object generation in the next steps. We therefore accept the longer computation time if the quality of the predicted results increases significantly, which we have proven. The possibility of performing feature extraction on arbitrary data with differing semantic classes simplifies the process of machine learning, since we can improve the models' robustness to class imbalances while also greatly improving the overall model performance. The modular approach allows for easy customization of the workflow to georeferenced data and data in local coordinate reference systems. In that way, incorporating our preprocessing workflow into deep learning models to improve segmentation results even leads to a reduction in the amount of data necessary. We have gained insights into the change in model performance when changing the receptive fields in feature extraction, and could thereby derive new research ideas towards direct integration of prior knowledge into deep neural network architectures.

7 Conclusion

With our work, we could show that our workflow improves the performance of arbitrary machine learning and deep learning models on ALS as well as on MMS datasets of the road space. Our modular and flexible preprocessing workflow is applicable to arbitrary neural network architectures and is capable of mitigating class imbalances and limited data supply. The best performing model regarding mIoU for the MMS dataset was the most recent PointNeXt model, which shows that even state-of-the-art neural networks can be boosted by data enrichment to date.

Following our findings from the immaculate evaluation of our workflow, we can answer RQ1, since we established a workflow directly applicable to arbitrary data, as we showed through evaluation of both ALS and MMS datasets. With our geospatial data integration step, we are able to automatically and efficiently limit the area of interest to the road space and further into homogenous road blocks, while data in local coordinate reference systems can still be processed with steps 3 and step 4. This, together with the finding that refining data subsets via ground filtering and road block extraction leads to an improved capability of extracting meaningful features with our approach, answers our RQ2. The combination of our four steps showed consistent improvement on all models, with the exception of PointNet++ in the ALS case, where the data sparsity negatively interacted with the data processing scheme of the deep learning architecture. Still, given this caveat, being a solvable problem as demonstrated by comparison between PointNet++ and its successor PointNeXt, where one significant change is the approach for processing input data, we consider RQ3 solved. Lastly, the optimal neighborhood approach we intertwined into the majority of our feature set, effectively allowed the process to automatically adapt to changing data. We additionally analyzed, that this approach not only automates the adaptation of receptive fields for feature extraction, but even matches the performance of tedious manual optimization of the receptive fields for several sets of features in the ablation study. We, therefore, also consider RQ4 answered.

Since in this work we focused on highway environments only, it is yet to be evaluated whether our approach can also improve the performance of point cloud semantic segmentation for different target environments, such as indoor semantic segmentation or bridge semantic segmentation. In future work, we aim to directly incorporate our preprocessing workflow into artificial neural network architectures to improve existing model architectures and to allow out-of-the-box performance boosting, without the necessity of excessive precomputation. This will be achieved by directly learning the most information-rich features currently handled in our feature selection step, which showed that the performance boost

can be maintained well by choosing the most representative features from the full feature set. This will accelerate the processing time on the one hand and make the preprocessing scheme even more adaptive to changing datasets on the other hand. Alternatively, our road block approach could potentially be deployed for block-based neural networks based on PointNet to directly utilize our road blocks to gain context information by an ordered reception over roads. Another idea is to combine the roadblock approach with image data of the road surface to incorporate more knowledge of the pavement. Generally, image data is better suited for crack detection, and mapping those detected damages into the point cloud would allow the generation of more realistic geometric-semantic as-is models for damage assessment within a digital twin use case. The workflow could also be extended to buildings or land use applications in the future, since geospatial data of the same kind we used is available for different domains and could be applied similarly. This will allow the evaluation of the performance enhancement potential for more diverse scenes, apart from the road space we focused on in this contribution. Overall, improving the semantic segmentation quality within the process chain of Scan-to-Twin is essential to creating correct and accurate digital representations for digital twins. The value of a digital twin depends on the accuracy of its predictions and implications. Therefore, the errors and uncertainties within the base model have to be minimized and, if possible, quantified to allow the system to become a reliable decision-support. In our future work, we will focus on completing the comprehensive framework to automatically derive and control the geometric-semantic models for digital twins of roads.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation. The datasets will be made available under: https://github.com/DavidCrampen/Data-enrichment-for-semantic-segmentation-of-point-clouds-of-roads.

Author contributions

DC: Visualization, Writing – original draft, Formal Analysis, Conceptualization, Investigation, Methodology, Data curation, Software, Validation. JB: Supervision, Project administration, Writing – review and editing, Funding acquisition, Resources.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This work has been conducted for and funded by the Collaborative Research Center SFB/TRR 339 “Digital Twin of the Road” (Project number 453596084). We acknowledge the funding by the German Research Foundation (DFG).

Acknowledgments

We gratefully acknowledge Straßen.NRW (Landesbetrieb Straßenbau Nordrhein-Westfalen) for providing access to their data.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Publisher’s note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fbuil.2025.1607375/full#supplementary-material

References

Alexiou, E., Zhou, X., Viola, I., and Cesar, P. (2024). Pointpca: point cloud objective quality assessment using pca-based descriptors. *EURASIP J. Image Video Process.* 20. doi:10.48550/arXiv.2111.12663

Aljumaily, H., Laefer, D. F., and Cuadra, D. (2019). Integration of lidar data and gis data for point cloud semantic enrichment at the point level. *Photogrammetric Eng. and Remote Sens.* 85, 29–42. doi:10.14358/PERS.85.1.29

Armeni, I., Sener, O., Zamir, A. R., Jiang, H., Brilakis, I., Fischer, M., et al. (2016). “3d semantic parsing of large-scale indoor spaces,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 1534–1543. doi:10.1109/CVPR.2016.170

Atik, M. E., and Duran, Z. (2022). Selection of relevant geometric features using filter-based algorithms for point cloud semantic segmentation. *Electronics* 11, 3310. doi:10.3390/electronics11203310

Baligh Jahromi, A., Zoej, M. J. V., Mohammadzadeh, A., and Sadeghian, S. (2011). A novel filtering algorithm for bare-earth extraction from airborne laser scanning data using an artificial neural network. *IEEE J. Sel. Top. Appl. Earth Observations Remote Sens.* 4, 836–843. doi:10.1109/JSTARS.2011.2132793

Besl, P. J., and Jain, R. C. (1988). Segmentation through variable-order surface fitting. *IEEE Trans. Pattern Analysis Mach. Intell.* 10, 167–192. doi:10.1109/34.3881

Bhanu, B., Lee, S., Ho, C.-C., and Henderson, T. (1986). "Range data processing: representation of surfaces by edges," in Proceedings of the eighth international conference on pattern recognition (Piscataway, NJ, USA: IEEE Computer Society Press), 238.

Biçici, S., and Zeybek, M. (2021). Effectiveness of training sample and features for random forest on road extraction from unmanned aerial vehicle-based point cloud. *Transp. Res. Rec.* 2675, 401–418. doi:10.1177/03611981211029645

Boulch, A., Saux, B. L., and Audebert, N. (2017). "Unstructured point cloud semantic labeling using deep segmentation networks," in *Proceedings of the workshop on 3D object retrieval* (Lyon, France: Goslar, DEU: Eurographics Association), 17–24. doi:10.2312/3dor.20171047

Bremer, M., Wichmann, V., and Rutzinger, M. (2013). Eigenvalue and graph-based object extraction from mobile laser scanning point clouds. *ISPRS Ann. Photogrammetry, Remote Sens. Spatial Inf. Sci.* II-5 (W2), 55–60. doi:10.5194/isprsannals-II-5-W2-55-2013

Chen, C., Guo, J., Wu, H., Li, Y., and Shi, B. (2021). Performance comparison of filtering algorithms for high-density airborne lidar point clouds over complex landscapes. *Remote Sens.* 13, 2663. doi:10.3390/rs13142663

Crampen, D., and Blankenbach, J. (2024). Semantic segmentation of road infrastructure point clouds supported by geospatial data. *arXiv*, 4951547. doi:10.2139/ssrn.4951547

Crampen, D., Hein, M., and Blankenbach, J. (2024). "A level of as-is detail concept for digital twins of roads—a case study," in *Recent advances in 3D geoinformation science*. Editors T. H. Kolbe, A. Donaubauer, and C. Beil (Cham: Springer Nature Switzerland), 499–515.

Esfahani, M. E., Rausch, C., Sharif, M. M., Chen, Q., Haas, C., and Adey, B. T. (2021). Quantitative investigation on the accuracy and precision of scan-to-bim under different modelling scenarios. *Automation Constr.* 126, 103686. doi:10.1016/j.autcon.2021.103686

Federal Ministry for Digital and Transport, BMDV Germany (2021). Masterplan bim bundesfernstraßen. Available online at: https://bmdv.bund.de/SharedDocs/DE/Anlage/StB/bim-rd-masterplan-bundesfernstrassen.pdf?__blob=publicationFile (Accessed March 26, 2025).

Grandio, J., Riveiro, B., Soilán, M., and Arias, P. (2022). Point cloud semantic segmentation of complex railway environments using deep learning. *Automation Constr.* 141, 104425. doi:10.1016/j.autcon.2022.104425

Grilli, E., Menna, F., and Remondino, F. (2017). A review of point clouds segmentation and classification algorithms. *Int. Archives Photogrammetry, Remote Sens. Spatial Inf. Sci.* XLII-2 (W3), 339–344. doi:10.5194/isprs-archives-XLII-2-W3-339-2017

Hirt, P.-R., Hoegner, L., and Stilla, U. (2021). A concept for the segmentation of individual urban trees from dense mls point clouds. *Int. Archives Photogrammetry, Remote Sens. Spatial Inf. Sci.* XLIII-B2-2021, 171–178. doi:10.5194/isprs-archives-XLIII-B2-2021-171-2021

Hu, Q., Liu, D., and Hu, W. (2023). Density-insensitive unsupervised domain adaption on 3d object detection. *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 17556–17566. doi:10.1109/cvpr52729.2023.01684

Hu, Q., Yang, B., Xie, L., Rosa, S., Guo, Y., Wang, Z., et al. (2020). Randla-net: efficient semantic segmentation of large-scale point clouds. *Proc. IEEE/CVF Conf. Comput. Vis. pattern Recognit.*, 11108–11117. doi:10.1109/CVPR42600.2020.01112

Huang, Y., and Li, L. (2011). Naive bayes classification algorithm based on small sample set. *2011 IEEE Int. Conf. Cloud Comput. Intell. Syst.*, 34–39. doi:10.1109/CCIS.2011.6045027

Interministerial Committee (2025). Interministerial committee for spatial data infrastructures in North Rhine-Westphalia. *OpenGeoPortalNRW*.

Justo, A., Soilán, M., Sánchez-Rodríguez, A., and Riveiro, B. (2021). Scan-to-bim for the infrastructure domain: generation of ifc-compliant models of road infrastructure assets and semantics using 3d point cloud data. *Automation Constr.* 127, 103703. doi:10.1016/j.autcon.2021.103703

Li, F., Lehtomäki, M., Oude Elberink, S., Vosselman, G., Kukko, A., Puttonen, E., et al. (2019a). Semantic segmentation of road furniture in mobile laser scanning data. *ISPRS J. Photogrammetry Remote Sens.* 154, 98–113. doi:10.1016/j.isprsjprs.2019.06.001

Li, F., and Vosselman, G. (2018). Pole-like road furniture detection and decomposition in mobile laser scanning data based on spatial relations. *Remote Sens.* 10, 531. doi:10.3390/rs10040531

Li, Y., Wang, W., Li, X., Xie, L., Wang, Y., Guo, R., et al. (2019b). Pole-like street furniture segmentation and classification in mobile lidar data by integrating multiple shape-descriptor constraints. *Remote Sens.* 11, 2920. doi:10.3390/rs11242920

Lucas, C., Bouten, W., Koma, Z., Kissling, W., and Seijmonsbergen, A. (2019). Identification of linear vegetation elements in a rural landscape using lidar point clouds. *Remote Sens.* 11, 292. doi:10.3390/rs11030292

Ma, H., Ma, H., Zhang, L., Liu, K., and Luo, W. (2022). Extracting urban road footprints from airborne lidar point clouds with pointnet++ and two-step post-processing. *Remote Sens.* 14, 789. doi:10.3390/rs14030789

Ma, H., Pei, Z., Wei, Z., and Zhong, R. (2017). Automatic extraction of road markings from mobile laser scanning data. *Int. Archives Photogrammetry, Remote Sens. Spatial Inf. Sci.* XLII-2 (W7), 825–830. doi:10.5194/isprs-archives-XLII-2-W7-825-2017

Mansour, M., Martens, J., and Blankenbach, J. (2024). Hierarchical svm for semantic segmentation of 3d point clouds for infrastructure scenes. *Infrastructures* 9, 83. doi:10.3390/infrastructures9050083

Martens, J., and Blankenbach, J. (2023). Vox2bim+ - a fast and robust approach for automated indoor point cloud segmentation and building model generation. *PFG – J. Photogrammetry, Remote Sens. Geoinformation Sci.* 91, 273–294. doi:10.1007/s41064-023-00243-1

Murtiyoso, A., and Grussenmeyer, P. (2019). Point cloud segmentation and semantic annotation aided by gis data for heritage complexes. *Int. Archives Photogrammetry, Remote Sens. Spatial Inf. Sci.* XLII-2/W9, 523–528. doi:10.5194/isprs-archives-XLII-2-W9-523-2019

Murtiyoso, A., Veriandi, M., Suwardhi, D., Soeksmantono, B., and Harto, A. (2020). Automatic workflow for roof extraction and generation of 3d citygml models from low-cost uav image-derived point clouds. *ISPRS Int. J. Geo-Information* 9, 743. doi:10.3390/ijgi9120743

Noroozinejad Farsangi, E., Shehata, A. O., Rashidi, M., Ghassempour, N., and Mirjalili, S. (2024). Transitioning from bim to digital twin to metaverse. *Front. Built Environ.* 10, 1486423. doi:10.3389/fbuil.2024.1486423

Osada, R., Funkhouser, T., Chazelle, B., and Dobkin, D. (2002). Shape distributions. *ACM Trans. Graph.* 21, 807–832. doi:10.1145/571647.571648

Paden, I., Garcia-Sanchez, C., and Ledoux, H. (2022). Towards automatic reconstruction of 3d city models tailored for urban flow simulations. *Front. Built Environ.* 8, 899332. doi:10.3389/fbuil.2022.899332

Park, Y., and Guldmann, J.-M. (2019). Creating 3d city models with building footprints and lidar point cloud classification: a machine learning approach. *Comput. Environ. Urban Syst.* 75, 76–89. doi:10.1016/j.compenvurbsys.2019.01.004

Qi, C. R., Su, H., Mo, K., and Guibas, L. J. (2017a). Pointnet: deep learning on point sets for 3d classification and segmentation. *Proc. IEEE Conf. Comput. Vis. pattern Recognit.*, 652–660. doi:10.1109/CVPR.2017.16

Qi, C. R., Yi, L., Su, H., and Guibas, L. J. (2017b). Pointnet++: deep hierarchical feature learning on point sets in a metric space. *Adv. neural Inf. Process. Syst.* 30, 14. doi:10.48550/arXiv.1706.02413

Qian, G., Li, Y., Peng, H., Mai, J., Hammoud, H., Elhoseiny, M., et al. (2022). Pointnext: revisiting pointnet++ with improved training and scaling strategies. *Adv. neural Inf. Process. Syst.* 35, 23192–23204. doi:10.48550/arXiv.2206.04670

Rato, D., and Santos, V. (2021). Lidar based detection of road boundaries using the density of accumulated point clouds and their gradients. *Robotics Aut. Syst.* 138, 103714. doi:10.1016/j.robot.2020.103714

Roynard, X., Deschaud, J.-E., and Goulette, F. (2018a). Classification of point cloud for road scene understanding with multiscale voxel deep network. *10th workshop Plan. Perceptionand Navigation Intelligent Veh. PPNIV'2018*, 1–11. doi:10.48550/arXiv.1804.03583

Roynard, X., Deschaud, J.-E., and Goulette, F. (2018b). Paris-lille-3d: a large and high-quality ground-truth urban point cloud dataset for automatic segmentation and classification. *Int. J. Robotics Res.* 37, 545–557. doi:10.1177/0278364918767506

Rusu, R. B., Blodow, N., and Beetz, M. (2009). Fast point feature histograms (fpfh) for 3d registration. *2009 IEEE Int. Conf. Robotics Automation (IEEE)*, 3212–3217. doi:10.1109/ROBOT.2009.5152473

Safaie, A. H., Rastiveis, H., Shams, A., Sarasua, W. A., and Li, J. (2021). Automated street tree inventory using mobile lidar point clouds based on hough transform and active contours. *ISPRS J. Photogrammetry Remote Sens.* 174, 19–34. doi:10.1016/j.isprsjprs.2021.01.026

Schatz, Y., and Domer, B. (2024). Semi-automated creation of ifc bridge models from point clouds for maintenance applications. *Front. Built Environ.* 10, 1375873. doi:10.3389/fbuil.2024.1375873

Schlosser, A. D., Szabó, G., Bertalan, L., Varga, Z., Enyedi, P., and Szabó, S. (2020). Building extraction using orthophotos and dense point cloud derived from visual band aerial imagery based on machine learning and segmentation. *Remote Sens.* 12, 2397. doi:10.3390/rs12152397

Shi, X., Ma, H., Chen, Y., Zhang, L., and Zhou, W. (2018). A parameter-free progressive tin densification filtering algorithm for lidar point clouds. *Int. J. Remote Sens.* 39, 6969–6982. doi:10.1080/01431161.2018.1468109

Shin, Y.-H., Son, K.-W., and Lee, D.-C. (2022). Semantic segmentation and building extraction from airborne lidar data with multiple return using pointnet++. *Appl. Sci.* 12, 1975. doi:10.3390/app12041975

Sithole, G., and Vosselman, G. (2001). Filtering of laser altimetry data using a slope adaptive filter. *Int. Archives Photogrammetry Remote Sens. Spatial Inf. Sci.* 34, 203–210.

Soilán, M., Riveiro, B., Martínez-Sánchez, J., and Arias, P. (2017). Segmentation and classification of road markings using mls data. *ISPRS J. Photogrammetry Remote Sens.* 123, 94–103. doi:10.1016/j.isprsjprs.2016.11.011

Tan, F., Fu, X., Zhang, Y., and Bourgeois, A. G. (2007). A genetic algorithm-based method for feature subset selection. *Soft Comput.* 12, 111–120. doi:10.1007/s00500-007-0193-8

Vassilev, H., Laska, M., and Blankenbach, J. (2024). Uncertainty-aware point cloud segmentation for infrastructure projects using bayesian deep learning. *Automation Constr.* 164, 105419. doi:10.1016/j.autcon.2024.105419

Wang, B., Lan, J., and Gao, J. (2022). Lidar filtering in 3d object detection based on improved ransac. *Remote Sens.* 14, 2110. doi:10.3390/rs14092110

Wang, Y., Sun, Y., Liu, Z., Sarma, S. E., Bronstein, M. M., and Solomon, J. M. (2019). Dynamic graph cnn for learning on point clouds. *ACM Trans. Graph. (tog)* 38, 1–12. doi:10.1145/3326362

Wang, Z., Yang, L., Sheng, Y., and Shen, M. (2021). Pole-like objects segmentation and multiscale classification-based fusion from mobile point clouds in road scenes. *Remote Sens.* 13, 4382. doi:10.3390/rs13214382

Weinmann, M., Jutzi, B., and Mallet, C. (2014). Semantic 3d scene interpretation: a framework combining optimal neighborhood size selection with relevant features. *ISPRS Ann. Photogrammetry, Remote Sens. Spatial Inf. Sci.* II-3, 181–188. doi:10.5194/isprsannals-II-3-181-2014

Wu, Z., Song, S., Khosla, A., Yu, F., Zhang, L., Tang, X., et al. (2015). 3d shapenets: a deep representation for volumetric shapes. *Proc. IEEE Conf. Comput. Vis. pattern Recognit.*, 1912–1920. doi:10.48550/arXiv.1406.5670

Xu, J., Zhang, R., Dou, J., Zhu, Y., Sun, J., and Pu, S. (2021). Rpvnet: a deep and efficient range-point-voxel fusion network for lidar point cloud segmentation. *Proc. IEEE/CVF Int. Conf. Comput. Vis.*, 16024–16033. doi:10.48550/arXiv.2103.12978

Yan, Y., and Hajjar, J. F. (2021). Automated extraction of structural elements in steel girder bridges from laser point clouds. *Automation Constr.* 125, 103582. doi:10.1016/j.autcon.2021.103582

Yang, R., Li, Q., Tan, J., Li, S., and Chen, X. (2020). Accurate road marking detection from noisy point clouds acquired by low-cost mobile lidar systems. *ISPRS Int. J. Geo-Information* 9, 608. doi:10.3390/ijgi9100608

Yilmaz, V. (2021). Automated ground filtering of lidar and uas point clouds with metaheuristics. *Opt. and Laser Technol.* 138, 106890. doi:10.1016/j.optlastec.2020.106890

Zhang, H., Duan, Z., Zheng, N., Li, Y., Zeng, Y., and Shi, W. (2022). An efficient class-constrained dbscan approach for large-scale point cloud clustering. *IEEE J. Sel. Top. Appl. Earth Observations Remote Sens.* 15, 7323–7332. doi:10.1109/JSTARS.2022.3201991

Zhang, J., Zhao, X., Chen, Z., and Lu, Z. (2019). A review of deep learning-based semantic segmentation for point cloud. *IEEE Access* 7, 179118–179133. doi:10.1109/ACCESS.2019.2958671

Zhao, H., Jiang, L., Jia, J., Torr, P. H., and Koltun, V. (2021). Point transformer. *Proc. IEEE/CVF Int. Conf. Comput. Vis.*, 16259–16268. doi:10.48550/arXiv.2012.09164

frontiers | Frontiers in Environmental Science

TYPE Original Research
PUBLISHED 25 June 2025
DOI 10.3389/fenvs.2025.1570182

Check for updates

OPEN ACCESS

EDITED BY
Izuru Takewaki,
Kyoto Arts and Crafts University, Japan

REVIEWED BY
Minzhe Du,
South China Normal University, China
Runsen Yuan,
Yanshan University, China
Guangming Yang,
Chongqing University of Technology, China

*CORRESPONDENCE
Sanglin Zhao,
13367312260@163.com
Måns Gustaf,
gustafedu@yeah.net

RECEIVED 03 February 2025
ACCEPTED 29 May 2025
PUBLISHED 25 June 2025

CITATION
Zhao S, Deng H, Cao J and Gustaf M (2025) Digital transformation of construction enterprises and carbon emission reduction: evidence from listed companies.
Front. Environ. Sci. 13:1570182.
doi: 10.3389/fenvs.2025.1570182

COPYRIGHT
© 2025 Zhao, Deng, Cao and Gustaf. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Digital transformation of construction enterprises and carbon emission reduction: evidence from listed companies

Sanglin Zhao[1]*, Hao Deng[1], Jikang Cao[1] and Måns Gustaf[2]*

[1]School of Engineering Management, Hunan University of Finance and Economics, Changsha, China, [2]School of Business, Society and Engineering, Mälardalens University, Västerås, Sweden

As a key sector for energy consumption and carbon emissions, the construction industry's carbon reduction measures have important strategic significance for achieving the "dual carbon" goals. Based on data from Chinese listed construction companies from 2000 to 2021, this study empirically explores the effects and pathways of digital transformation on carbon reduction. The results indicate that digital transformation can significantly reduce the carbon emission intensity of enterprises, mainly through promoting green technology innovation, improving total factor productivity, and optimizing production processes and business structures. Heterogeneity analysis shows that digital technology has a more significant emission reduction effect on highly competitive enterprises in the industry, and there are significant differences in carbon emission reduction between regions. This study provides a reference for carbon neutrality pathways in the field of architecture.

KEYWORDS

digital transformation of enterprises, carbon emission reduction, carbon emission of construction enterprises, listed company, dual carbon targets

1 Introduction

With the severe challenge of global climate change, reducing carbon emissions and achieving carbon neutrality have become the focus of international attention. As one of the largest carbon emitters in the world, China announced to the world the grand goal of "achieving peak carbon dioxide emissions by 2030 and striving to achieve carbon neutrality by 2060". In this context, the construction industry is one of the main fields of energy consumption and carbon emission, and its carbon emission reduction actions have an important impact on achieving carbon emission reduction targets.

In recent years, with the rapid development of digital technology, the digital transformation of construction enterprises has become an important force to promote industry reform. At the same time, the wide application of digital technology in enterprises has promoted the sustainable development of the industry. Existing research shows that the digital transformation of enterprises has a significant role in promoting carbon emission reduction. Through digital transformation, enterprises can achieve structural optimization, resource integration, efficiency improvement and benefit increase, thus reducing carbon emissions (Zhang Zack et al., 2024). It is suggested that digital transformation has a significant inhibitory effect on urban carbon emissions, and carbon emission reduction is mainly achieved through internal optimization of enterprises and industrial integration, and the carbon emission reduction effects in different regions and economic agglomeration

areas are heterogeneous (Chen and Zhang, 2024). It is pointed out that digital transformation is helpful to reduce the carbon emission intensity of enterprises, and energy conservation and emission reduction can be realized mainly by improving the effectiveness of internal control and promoting green technology innovation of enterprises. However, some scholars believe that if enterprises, governments and other subjects cannot take reasonable measures to promote digital transformation, it may be difficult to achieve sustainable development of carbon emission reduction (Zhang, 2023). It is considered that institutional pressure has a negative regulatory effect between digital transformation of construction enterprises and carbon emission of enterprises, that is, under high institutional pressure, certain economic compensation should be considered to improve the enthusiasm of construction enterprises to achieve carbon neutrality (Niu, 2023). It is believed that if the government does not encourage enterprises to actively adopt digital technology, the efficiency of digital transformation of the industry is very low, and it is not conducive to the sustainable development of the environment.

However, despite the great potential of digital transformation of construction enterprises, the empirical research on the specific impact of carbon emission reduction is not sufficient. Therefore, based on the evidence of listed companies, this paper explores the impact of digitalization of construction enterprises on carbon emission reduction through empirical analysis, robustness test, mechanism analysis and heterogeneity analysis, which not only helps to reveal the carbon emission reduction effect of digital transformation, but also provides scientific decision-making basis for the government to formulate relevant policies and enterprises to provide transformation strategies.

The marginal contributions of this paper are as follows: First, the empirical analysis of the relationship between digital transformation and carbon emission reduction of construction enterprises will help to enrich and improve the theoretical system in the field of digital transformation and environmental protection, and provide a new perspective and methodology for subsequent research. Secondly, for construction enterprises, it reveals the actual effect of digital transformation in reducing carbon emissions, and provides enterprises with the motivation and direction of transformation; At the same time, according to the research results, the government can formulate more accurate digital transformation support policies and green financial policies to promote the green and low-carbon development of construction enterprises; In addition, according to the research results, it can provide reference for other industries and promote the whole society to form a virtuous circle of digital transformation and carbon emission reduction.

2 Literature review

Related research focuses on the various impacts of digital transformation on the development of the construction industry and the operation of enterprises themselves. By promoting the information transformation of the construction industry, promoting the integration of intelligent technology and prefabricated buildings and optimizing the management efficiency of enterprises, digital technology has significantly improved the production efficiency of the industry and the core competitiveness of enterprises. At the level of industry development, digital technology has brought profound changes to the construction industry. Construction industry is one of the industries with the lowest informatization and digitalization in traditional industries. Whoever can realize digital transformation faster will occupy an advantageous position in market competition (Yang Yingnan et al., 2022). The rapid development of digital perception, analysis and decision-making service technology in recent years can transform and upgrade the construction industry to technology-intensive industries, and bring new opportunities for the development of construction projects (Wang Pujin et al., 2024). The organic integration of intelligent control technology and prefabricated building production is irreplaceable in prefabricated buildings and even the whole construction industry, and marks the birth of a new intelligent production mode of prefabricated buildings (Liu et al., 2023). The application of digital technology can improve the overall production efficiency of the construction industry and promote the efficient development of the industry. From the perspective of enterprise operation, digital transformation has brought many positive changes to construction enterprises. The digital transformation of construction enterprises should ensure the deep integration of data governance and business innovation, fully integrate business management experience and digital technology, and ensure that digital technology can be effectively applied in practical business (Cultur et al., 2024; Borozan and Pirgaip, 2024). The digital transformation of the construction industry can improve the scattered communication information and low communication efficiency of all departments of the enterprise, build a functional management and control system that is connected from top to bottom and linked internally and externally, and improve the work efficiency of employees and the management efficiency of enterprises (Pingru et al., 2023; Zhao et al., 2025; Lei, 2024; Zuo, 2023).

Carbon emission reduction has become a key measure to deal with global warming. For construction enterprises, the path and effect of carbon emission reduction are the focus of research. The construction industry accounts for a large share of global greenhouse gas emissions and waste generation, and digital transformation is expected to provide new solutions for the most urgent challenges of this industry (Kathrin and Cordula, 2023). Some scholars have found that the digital economy has a significant carbon reduction effect (Hui and Zhang, 2025; Xiao and Kong, 2025; Huang et al., 2025a; Huang et al., 2025c). Digital technology makes buildings more integrated, flexible, energy-saving, intelligent and sustainable by optimizing resource utilization, improving operational efficiency and minimizing environmental impact (Asif et al., 2024). The cost borne by traditional production methods is the key factor to determine the speed of low-carbon transformation (Luo and Tang, 2020). After a variety of robustness tests, digital transformation can significantly promote the carbon emission reduction effect of China's construction enterprises, and can promote carbon emission reduction by enhancing technological innovation and increasing corporate social attention (He et al., 2025; Li et al., 2024; Huipeng et al., 2023; Xing and Zhu, 2025). However, some studies have pointed out that the relationship between digital transformation and carbon emission reduction of construction enterprises is complicated. Some scholars believe that there is a nonlinear relationship between digital economy and

carbon emissions (Peng et al., 2024; Zhao et al., 2024; Shao et al., 2024). It is pointed out that with the development of digital economy to a higher level, the effect of emission reduction becomes more obvious (Wang et al., 2023). found that there is an inverted U-shaped relationship between digital economy and carbon emissions, and the emission reduction effect is greater in areas with scarce resources (Li et al., 2024; Huang et al., 2025b). Based on the study of this relationship, many scholars have further explored the mechanism of digital economy affecting carbon emissions. (Changqi et al., 2025). It is found that the development of digital economy promotes carbon emission reduction through technological progress, structural optimization and strengthening education. Digital economy reduces unnecessary waste of resources, changes the traditional energy use mode, and accelerates the replacement of high-carbon industries by low-carbon industries (Peak, 2025; Zhang and Yang, 2025; yi and Su, 2025-04; Cook, 2017) And reduce the intensity of carbon emissions by reducing energy demand (Zhao et al., 2024). Although digital technology provides new opportunities for carbon emission reduction, in the initial stage of digital transformation, enterprises may increase energy consumption and fluctuate carbon emissions in the short term due to equipment updating and technology investment (Wang Wenjie et al., 2024). Because the extensive use of related infrastructure may lead to higher power consumption, thus increasing emissions (Zhou and Chu, 2024). Existing literature on the impact of digital transformation on carbon emissions generally emphasizes the energy efficiency improvement and industrial structure optimization effects brought about by technological progress, but lacks attention to the potential "rebound effect" it may trigger (Mai et al., 2025; Han et al., 2025; Ying et al., 2024; Huang et al., 2025b; Huang et al., 2025c). The rebound effect theory suggests that when technological progress improves energy efficiency, the relative decrease in energy service costs may stimulate the growth of consumer demand, thereby partially or completely offsetting the expected emission reduction effect (Zhang Yuxin et al., 2024; Liu Haiying et al., 2024; Su et al., 2025). This controversy is particularly prominent in the context of digital transformation and requires theoretical analysis and empirical testing from multiple dimensions. For high carbon industries (Chen et al., 2024; Liu Bei et al., 2024; Shang Yuping et al., 2023), the initial stage of digital transformation may lead to a significant increase in carbon emissions due to equipment updates and capacity expansion (Wei et al., 2025; Zhang and Cui, 2025; Yang and Liu, 2024); But for low-carbon industries, the emission reduction effect achieved by optimizing supply chain management through digital technology is more significant (Yang et al., 2023; Yang Xiaoguang et al., 2022; Chen et al., 2022; Li et al., 2024; Wang, 2024). This industry heterogeneity requires that policy design should go beyond the unified model and adopt differentiated transformation strategies.

The existing literature mainly focuses on the relationship between digital transformation and carbon emission reduction, but the impact on low-carbon technology and how technological progress can promote low-carbon technology and sustainable economic development still need further study. First of all, the research on the influence of digital economy on technological progress mostly focuses on whether it can improve energy efficiency and provide new impetus for long-term economic growth. At present, there is a lack of research on how the digital economy affects technological progress by indirectly empowering different enterprise production factors. Secondly, many studies have discussed the impact of digital economy on carbon emissions from the perspectives of environmental governance, industrial structure optimization and resource utilization. Few scholars pay attention to the impact of digital economy on carbon emissions from the perspective of enterprise digital technology progress through the differential empowerment effect of digital economy on various factors. Most documents define technological progress as that technological progress improves the marginal productivity of one factor of production relative to another. However, technological progress is often tendentious, and not all forms of progress can bring about pollution reduction. When technological progress improves the marginal productivity of energy factors, it improves energy efficiency, thus reducing pollutant emissions. This paper studies the digital transformation to improve enterprise innovation mechanism, total factor productivity and optimization of production and operation structure to promote the reduction of enterprise carbon emissions, further reduce regional total carbon emissions, and further analyze regional heterogeneity.

3 Theoretical analysis and research hypothesis

This paper focuses on construction enterprises, deeply analyzes the impact of digital transformation on their carbon emission reduction and its internal mechanism, and puts forward the following three research hypotheses.

Digital transformation can reduce energy consumption by reconstructing the resource allocation and operation mode of construction enterprises. The deep application of digital technology not only changed the production mode of traditional construction industry, but also optimized the energy utilization efficiency through real-time data collection and analysis. For example, the Internet of Things and big data technology can accurately monitor the energy consumption in the whole life cycle of a building and reduce redundant energy consumption in the construction and operation stages. In the construction stage, the intelligent equipment scheduling system can reduce the mechanical idling rate; In the building operation stage, the intelligent management system can significantly reduce carbon emissions by dynamically adjusting the use of air conditioning, lighting and other equipment. In addition, digital supply chain management can optimize the transportation route and inventory level of building materials and reduce the implied carbon emissions (Jun-Qian and Yong-Kou, 2025). Digital transformation realizes carbon emission reduction by improving the ability of resource integration and decision-making efficiency. Therefore, this paper puts forward hypothesis H1, and thinks that digital transformation can significantly reduce the carbon emission intensity of construction enterprises.

H1. Digital transformation can significantly reduce the carbon emission intensity of construction enterprises.

The carbon emission reduction effect of digital transformation is not a single path, but is realized through the synergistic mechanism

of technological innovation, efficiency improvement and structural optimization. First, digital technology has accelerated green innovation. Tools such as building information model (BIM) and artificial intelligence not only shorten the research and development cycle of green technology, but also optimize the architectural design scheme through simulation to reduce the waste of building materials and construction energy consumption. For example, the SPCS system developed by Sany Construction Engineering Co., Ltd. Significantly reduces the energy waste caused by building materials loss and construction interruption through standardized assembly production technology. Secondly, digital tools improve the efficiency of production factors through automation and intelligent transformation. Intelligent algorithm can optimize the construction process and reduce the idle time of manpower and equipment; Blockchain technology improves the efficiency of supervision and traceability through transparent carbon emission data (Lan and Zhou, 2025). Digital transformation significantly improves the total factor productivity and reduces the operating cost rate, which proves the effectiveness of the efficiency improvement path. Finally, digital transformation promotes the transformation of enterprises from high-carbon to low-carbon mode. For example, construction enterprises increase the proportion of green building projects through digital market analysis, and optimize supply chain management by using carbon accounting platform to achieve emission reduction in the whole life cycle (Yan et al., 2023). The interaction of the above three paths shows that digital transformation provides a systematic solution for carbon emission reduction of construction enterprises through the comprehensive effects of technology empowerment, efficiency drive and structural remodeling, so hypothesis H2 is proposed.

H2. Digital transformation realizes carbon emission reduction through technological innovation, improvement of production factor efficiency and optimization of production and operation structure.

The structural differences within China's construction industry and the unbalanced regional development lead to the asymmetric carbon emission reduction effect of digital transformation. In terms of heterogeneity of property rights, the difference of institutional environment between state-owned and non-state-owned enterprises significantly affects the implementation effect of digital transformation. State-owned enterprises are subject to lengthy decision-making processes and administrative goal orientation, and tend to promote digitalization through administrative instructions, but lack flexible technical application scenarios (Han and Marco, 2020). For example, non-state-owned enterprises are more likely to adopt intelligent construction technology to optimize energy consumption (Julie et al., 2022), and state-owned enterprises may be difficult to release the digital potential because of the complex management level (Zhang and Shen, 2024; hai et al., 2025; Fu and Pan, 2024; Guangmiao, 2023). In terms of regional heterogeneity, although the eastern coastal areas have more perfect digital infrastructure and policy support, it stems from the fact that the central and western regions have accelerated the layout of digital infrastructure through "late-comer advantage" and gradually narrowed the technological gap with the east. For example, by introducing cloud computing and remote collaboration platform, enterprises in the central and western regions have overcome the shortcomings of local technology and achieved emission reduction efficiency similar to that in the east. This shows that the inclusive characteristics of digital transformation enable its carbon emission reduction effect to break through the limitation of unbalanced regional development, but its heterogeneity in property rights still needs targeted policy intervention. Therefore, this paper puts forward hypothesis H3, emphasizing the need to formulate differentiated transformation strategies based on enterprise attributes and regional characteristics.

H3. The carbon emission reduction effect of digital transformation has property right heterogeneity and regional heterogeneity.

4 Analysis of carbon emission status of construction enterprises

4.1 Status of carbon emissions

With the increasingly severe global climate change, the construction industry, as an important field of energy consumption and carbon emission, has attracted much attention. According to the Report on the Status of Global Construction Industry in 2024–2025 issued by the United Nations Environment Programme and the Global Building and Construction Alliance, although the emissions of the construction industry will stop rising after 2020, the slow progress and insufficient funds put the global climate target at risk. According to the report, the construction industry is still the main driver of the climate crisis, consuming 32% of global energy and contributing 34% of global carbon dioxide emissions.

In China, according to "Research Report on Carbon Emissions in Urban and Rural Construction in China (2024 Edition)", the total carbon emissions of buildings and construction industry in China reached 5.13 billion tons of carbon dioxide in 2022, accounting for 48.3% of the national energy-related carbon emissions. Among them, the construction industry emits 2.82 billion tons of carbon dioxide, and the construction operation emits 2.31 billion tons of carbon dioxide. This data shows that the carbon emission of China's construction industry is huge, and the task of emission reduction is very arduous (Table 1).

4.2 Main emission sources

The carbon emissions of construction enterprises mainly come from the following aspects:

(1) Production of building materials: The production of building materials is one of the main sources of carbon emissions in construction. For example, in the production process of building materials such as cement, steel and wood, a large amount of carbon dioxide emissions will be generated. Cement production is one of the main sources of carbon emissions, and its carbon emissions account for a large proportion of the total.
(2) Construction process: During the construction process, mechanical equipment operation, transportation, and

TABLE 1 Carbon emission from construction process.

Classify		Energy consumption		Carbon dioxide emission	
		108 tce (Coal consumption for power generation)	Occupy the whole country proportion	108tCQ	Occupy the national energy phase Carbon emission ratio
Architecture and construction industry construction		**24.2**	**44.7%**	**51.3**	**48.3%**
Construction industry construction		**12.3**	**22.7%**	**28.2**	**26.6%**
Classification	Building construction	7.9		18.4	
	Infrastructure construction	4.4		9.8	
Subprocess	Building materials production and transportation	11.5		27.2	
	construct	0.8		1.00	
Building operation		**11.9**	**22.0%**	**23.1**	**21.7%**
Classification	Public building	4.9		9.4	
	Urban residential buildings	4.5		8.9	
	Rural residential buildings	2.5		4.8	
Separate sources	fossil energy	2.3		4.5	
	power	7.9		14.4	
	central heating	1.7		4.2	

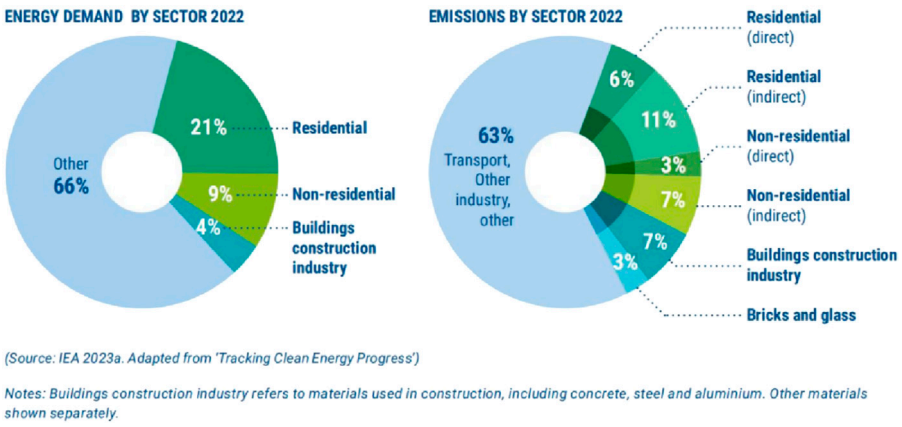


FIGURE 1
Schematic diagram of carbon emission in construction industry.

temporary facilities construction will all produce carbon emissions. In addition, waste disposal at the construction site and energy consumption during construction will also increase carbon emissions.

(3) Building operation stage: During the operation of the building, such as lighting, air conditioning, heating, etc., a lot of energy will be consumed, and then carbon emissions will be generated. The carbon emissions in the operation stage of a building account for a large proportion of the carbon emissions in the whole life cycle of the building.

4.3 Influencing factors

The carbon emissions of construction enterprises are affected by many factors, mainly including the following aspects (Figure 1):

Population and urbanization rate: the total population and urbanization rate are important factors affecting building energy consumption and carbon emissions. With the growth of population and the improvement of urbanization rate, the carbon emissions in the construction sector may increase accordingly.

Building area and type: the expansion of building area and the diversification of building types will also lead to the increase of carbon emissions. For example, large public buildings and high-rise residential buildings usually have high energy consumption and carbon emissions.

Energy intensity and structure of buildings: Changes in energy intensity and energy structure of buildings will also affect carbon emissions. For example, the use of energy-efficient building equipment and materials can reduce the energy intensity of buildings, thus reducing carbon emissions.

Urban construction land: The expansion of urban construction land will bring higher energy consumption and carbon dioxide emissions. Therefore, rational planning of urban construction land is of great significance for reducing building carbon emissions. Climate factors: Climate factors will also have an impact on building carbon emissions. For example, areas with cold weather in winter need more heating energy, thus increasing carbon emissions.

To sum up, the current situation of carbon emissions of construction enterprises is not optimistic. In order to reduce carbon emissions and achieve green and low-carbon development, construction enterprises need to start from many aspects, including optimizing the selection of building materials, improving energy utilization efficiency, and promoting the application of renewable energy. At the same time, the government and all walks of life also need to work together to promote the low-carbon transformation and sustainable development of the construction industry through policy guidance, technological innovation and social supervision.

5 Research data and model design

5.1 Model assumptions

5.1.1 Benchmark regression model

According to the above theoretical analysis, this paper designs a benchmark regression model (Equation 1):

$$\text{inten}co_{2it} = \alpha_0 + \alpha_1 Digit_{it} + \sum \alpha_k Controls_{it} + year + firm + year + \xi_{it} \tag{1}$$

In which intenCO2it represents the carbon emission intensity of enterprise I in the t year. The core explanatory variable is digitalit, which represents the degree of digital transformation of enterprise I in T year; Controls is the control variable set model, which also controls the time-fixed effect Year and the enterprise-fixed effect Firm; ξ_{it} is a random error term.

5.1.2 Interaction effect model

In order to deeply explore the impact of digital access level, this paper divides all regions in China into eastern coastal areas and other regions, and at the same time constructs an interactive term (region × digital) of regional virtual variables and digital transformation, so as to distinguish the heterogeneous impact of regions on the digital carbon emission reduction effect (Equations 2).

$$\text{inten}co_{2it} = \alpha_0 + \alpha_1 Digit_{it} + \alpha_1 (Digit_{it} \times \text{region}) + \sum \alpha_k Controls_{it} + year + firm + year + \xi_{it} \tag{2}$$

5.1.3 Intermediary variable model

To examine the mediating mechanism of the impact of enterprise digital transformation on carbon emissions, this study draws on Abbott, K. W. (2017) and Wilkinson, L. (1979), who estimated intermediate variables using the "stepwise regression method". Specifically, taking Total factor productivity (TF), innovation level (Inv), and R&D intensity (RD) as intermediate variables, the benchmark regression is tested, and the corresponding research model is constructed.

$$\text{Media}_{i,t} = \beta_0 + \beta_1 Digit_{it} + \sum \beta_x \text{Control}_{i,t} + \sum Code + \sum Year + \varepsilon_{i,t} \tag{3}$$

$$\text{inten}co_{2i,t} = \delta_0 + \delta_1 DT_{i,t} + \delta_2 \text{Media}_{i,t} + \sum \delta_x \text{Control}_{i,t} + \sum Code + \sum Year + \varepsilon_{i,t} \tag{4}$$

The mediating variable plays a mediating role when the coefficients in Equation 4 and Equation 3 are significant. When and the coefficient in Equation 4 is significant, it indicates that the mediating effect is partial, and *vice versa*. β_1, δ_2 all significant, This means that the mediation variable plays a mediating role. When and in Equation 4 δ_1 coefficients are significant when, This indicates that the mediation effect is partial mediation effect and otherwise complete mediation effect.

5.2 Data source and variable description

5.2.1 Data sources

The research sample of this paper comes from the annual data of China A-share listed construction companies from 2000 to 2021. The annual reports of relevant enterprises were obtained from official website of Shenzhen and Shanghai Stock Exchanges; Relevant data of listed companies come from CSMAR database. At the same time, the non-ratio continuous variables are truncated by 1% to reduce the influence of outliers. In this paper, the data of digital transformation that has not been disclosed in the annual report of enterprises are excluded, and the construction industry is divided into industry categories: civil engineering construction, architectural decoration, decoration and other construction industries, construction and installation industries, and housing construction industries (Table 2) (Sanglin, 2025).

The data processing process in this paper is as follows:

5.2.1.1 Logarithmic transformation

In the second step, the variables with skewed distribution are logarithmically transformed to reduce the influence of skewed data and make the data distribution closer to normal distribution, which

TABLE 2 Original sample classification.

Industry classification	Count	Proportion
Civil engineering construction industry	837	74.9%
Architectural decoration, decoration and other construction industries	226	20.2%
Construction and installation industry	28	2.5%
housing industry	27	2.4%

is beneficial to subsequent statistical analysis. Logarithmic transformation is shown in Formula 5

$$y = log(x + 1) \tag{5}$$

5.2.1.2 Missing value filling

In order to deal with missing data, interpolation method is used to fill the blank values in time series data. The linear interpolation formula of Equation 6 is adopted.

$$x_{miss} = x_{after} + \frac{x_{after} - x_{before}}{2} \tag{6}$$

"indicates missing data", and indicate data before and after missing data points, respectively. $x_{miss} x_{before} x_{after}$

5.2.2 Description of variables

Descriptive statistics of the main variables in this paper are shown in Table 3. Specific descriptions of related variables are as follows:

(1) Explained variable: enterprise carbon dioxide emission intensity (intenco2).

The carbon emissions, fossil energy consumption, electricity and heat consumption used in this article were manually collected from annual social responsibility reports, sustainable development reports, and environmental reports disclosed by companies, and further calculated according to the methods published by the National Development and Reform Commission (NDRC) to obtain partial carbon emission data. The sample year span is from 2000 to 2021. The accounting of greenhouse gases by enterprises needs to be based on the internationally recognized standards of the Greenhouse Gas Protocol, and the calculation method published by the National Development and Reform Commission is also the same. According to the greenhouse gas accounting system, a company's carbon emissions can be divided into three categories. Scope one refers to the direct greenhouse gas emissions generated by emission sources owned or controlled by enterprises, such as combustion emissions from boilers, furnaces, vehicles, etc. Owned or managed by enterprises; Emissions generated from chemical production using owned or controlled process equipment. Scope 2: Accounting for indirect greenhouse gas emissions generated by the consumption of electricity and heat purchased by the company. Scope 1 and Scope 2 carbon emissions are mandatory disclosures required by the greenhouse gas accounting system and are also the focus of this article. The total carbon emissions typically disclosed by companies refer to the sum of Scope 1 and Scope 2 emissions. Scope three emissions consider all other indirect emissions that are the result of company activities but not generated by emission sources owned or controlled by the enterprise. For example, raw materials used for mining and production procurement, fuels used for transportation procurement, the use of products and services sold, and transportation vehicles used by employees and customers. This article does not consider Scope 3 emissions. The regional carbon emissions (CO2) data used in this article is from the China Energy Statistical Yearbook.

(2) Core explanatory variable: Digital transformation degree of listed companies.

The measurement method of digital transformation of the company has been relatively mature, and the measurement method adopts text analysis. Firstly, this paper constructs a digital transformation keyword list; Then use Python software to match the glossary with the annual report text of listed companies, and use Jieba function module to count the word frequency of related keywords in the annual report file of listed companies; Finally, after adding 1 to the word frequency for logarithmic processing. Get the enterprise digital transformation index.

Note: Referring to Zhao Chenyu (2021), this paper makes statistics on 99 digital related word frequencies in four dimensions: digital technology application, Internet business model, intelligent manufacturing and modern information system.

(3) Control variables. In this paper, the enterprise-level indicators are included in the control variables: the nature of enterprise property rights (soe, state-owned enterprises take 0, private enterprises take 1), the board size (board, Logarithmic number of board members), logarithm of enterprise age (AGE), asset-liability ratio (lev), return on net assets (roe), operating cash flow (cf), sales growth rate (growth), net profit growth rate (gprofit), tangible assets ratio (tangibi), independent directors ratio (indep), shareholding ratio of the largest shareholder (top1), chairman and chairman.

Descriptive statistical analysis shows that (Table 4), the research sample covers 737 observation objects, and the core variables show multi-dimensional distribution characteristics. In terms of financial performance, the average return on equity (ROE) of enterprises is only 0.023%, but the standard deviation is 1.343%, reflecting the significant differentiation of the profitability of sample enterprises. The concentration of cash flow (Cflow) is relatively high, with the median value of 0.014 close to the average value of 0.008, but the lower limit of negative value (−3.224) reveals the financial liquidity risk of some enterprises. Finlev is distributed to the right, indicating that most

TABLE 3 Construction of enterprise digital transformation index and key words selection.

Dimension	Classified words	Text combinations with high frequency of occurrence	Word segmentation dictionary
Numeric technique app; application	Data, numbers, digitization	Data management, data mining, data network, number According to the platform, data center, data science, digital Control, digital technology, digital communication, digital network Network, digital intelligence, digital terminal, digital marketing, digitization	Data management, data mining, data network, data platform Data center, data science, digital control, digital technology Digital communication, digital network, digital intelligence, digital terminal Digital Marketing, Digitalization, Big Data, Cloud Computing, Cloud IT, Cloud Ecology, cloud services, cloud platforms, blockchain, Internet of Things, machines study
Internet business model	Internet, e-commerce	Mobile Internet, Industrial Internet and Industrial Interconnection Network, Internet solutions, Internet technology Internet thinking, Internet action, Internet industry Service, Internet mobile, Internet application and interconnection Internet marketing, Internet strategy, Internet platform Internet model, Internet business model, interconnection Network ecology, e-commerce, e-commerce	Mobile internet, industrial internet, industrial internet, interconnection Network solution, Internet technology, Internet thinking, interconnection Network action, Internet business, Internet mobile and Internet should Use, Internet marketing, Internet strategy, Internet platform Internet model, Internet business model, Internet ecology E-commerce, e-commerce, Internet. “internet plus”, online Offline, online to offline, online and offline, 020, B2B C2C、B2C、C2B
smart manufacturing	Intelligent, intelligent Automatic, numerical control, one Systematization and integration	Artificial intelligence, high-end intelligence, industrial intelligence, mobile Dynamic intelligence, intelligent control, intelligent terminal, intelligence Mobile, intelligent management, intelligent factory, intelligent things Flow, intelligent manufacturing, intelligent warehousing, intelligent technology Intelligent equipment, intelligent production, intelligent networking, intelligence Can be systematic, intelligent, automatic control and automatic monitoring Measurement, automatic monitoring, automatic detection, automatic production Numerical control, integration, integration and integration solution Case, integrated control, integrated system	Artificial intelligence, high-end intelligence, industrial intelligence, mobile intelligence Intelligent control, intelligent terminal, intelligent mobile, intelligent management Intelligent factory, intelligent logistics, intelligent manufacturing, intelligent warehousing Intelligent technology, intelligent equipment, intelligent production, intelligent networking, intelligence Can be systematic, intelligent, automatic control, automatic monitoring and automatic supervision Control, automatic detection, automatic production, numerical control, integration, integration Chemistry, integrated solutions, integrated control, integrated systems, industry Cloud, future factory, intelligent fault diagnosis, life cycle management Manufacturing execution system, virtualization, virtual manufacturing
Contemporary information system	Information, informatization, networking	Information sharing, information management, information integration and information communication Information software, information system, information network, information Terminal, information center, informationization and networking	Information sharing, information management, information integration, information software Information system, information network, information terminal, information center Informatization, networking, industrial information and industrial communication

enterprises maintain a moderate level of leverage, but there are a few cases of high leverage operation.

6 Empirical analysis

6.1 Benchmark regression

Table 4 reports the results of fitting regression according to Formula 1. Column (1) gives the results of adding only the core explanatory variables, and it can be found that the digital transformation of enterprises can significantly reduce the carbon emission intensity of enterprises; On this basis, columns (2) and (3) control the fixed effect of the industry and region where the enterprise is located, and the digital coefficient is still significantly negative; Finally, (4) regression with control variables shows that the higher the degree of digital transformation of enterprises, the better the effect of carbon emission reduction. Thus, the regression results in Table 5 prove that the digital transformation of enterprises can promote enterprises to reduce carbon emissions. That is, to verify the research hypothesis H1.

6.2 Robustness test

6.2.1 Replacing core explanatory variables

In order to further verify the robustness of the conclusion, the core explanatory variables are replaced to test the robustness. Firstly,

TABLE 4 Variable declaration.

Variable type	Symbol	Variable declaration
Explained variable	intenco2	Carbon emission intensity of enterprises
Explanatory variable	digtial	Enterprise digital transformation, ln (related word frequency of enterprise digital transformation +1)
Control variable	age	Enterprise age
	lev	Asset-liability ratio
	roe	Rate of Return on Common Stockholders' Equity
	cflow	Operating cash flow
	growth	Sales growth rate
	gprofit	net profit growth rate
	Tangibi	Proportion of tangible assets
	board	Board size
	indep	Proportion of independent directors
	Top1	Proportion of shares held by the largest shareholder
	dual1	The value of chairman and general manager is 1, otherwise it is 0
	soe	Property right nature, state-owned enterprises take 0, private enterprises take 1

TABLE 5 Descriptive statistic.

Name	Sample size	Min	Max	Mean	Std	Median
age	1,118	−1.000	28.000	8.474	6.764	7.000
Region id	1,118	0.000	1.000	0.384	0.487	0.000
roe	1,118	−13.949	29.725	0.023	1.343	0.070
cflow	1,118	−3.224	0.325	0.008	0.139	0.014
Finlev	1,118	0.000	0.908	0.359	0.232	0.357
growth	1,118	−0.747	82.792	0.511	3.742	0.108
gprofit	1,118	−1,058.374	4254.917	4.980	179.507	0.123
size	1,118	16.185	26.282	22.358	1.313	22.237
Board size	1,118	0.000	18.000	9.075	1.723	9.000
Number of independent directors	1,118	0.000	5.000	3.170	0.860	3.000
Proportion of independent directors	1,118	0.000	0.571	0.357	0.090	0.333
top1	1,118	4.420	87.460	34.285	14.331	33.670
Nature of equity	1,118	1.000	6.000	3.467	1.535	2.000
Carbon emission intensity	1,118	14.005	547,276.396	19,274.929	44,711.957	5962.537
soe	1,118	0.000	1.000	0.517	0.500	1.000
intenco2	1,118	32.248	1260150.472	44,382.164	102,953.085	13,729.249
Does the chairman concurrently serve as the general manager?	1,118	0.000	1.000	0.199	0.400	0.000
Enterprise carbon emission lg	1,118	2.449	7.152	5.158	0.673	5.120
Regional carbon emission lg	1,118	2.693	3.814	3.530	0.199	3.579
Digital transformation level	1,118	0.000	4.000	1.128	1.675	0.000
Frequency proportion of words in digital transformation (10E-5)	1,118	0.000	3.135	0.240	0.675	0.000
Digital transformation	1,118	0.000	13.000	0.474	1.689	0.000

this paper uses (Wu et al., 2021) to measure the digital transformation of enterprises, which is recorded as Dig1 and Dig2. The results are shown in columns (1) and (2) of Table 6, and Wu Fei (2021) is used as the second one (Wu et al., 2021; Junqian and Yong-kou, 2025). Others' digital transformation rating index (Dig3) is used to measure the degree of digital transformation of enterprises. As shown in column (3), the result is still remarkable. Secondly, the regression results obtained by using sulfur dioxide emissions (SO2) of fossil fuels instead of carbon emissions of enterprises are shown in column (4) (Zhu et al., 2023; Kathrin and Cordula, 2023; Cultur et al., 2024), which shows that the results are still significant, and the digital transformation has effectively reduced carbon emissions of enterprises and other pollutants.

6.2.2 Substitution sample regression

After excluding municipalities directly under the central government, this study uses the fixed effect (FE) model to empirically analyze the relationship between digital transformation of construction enterprises and carbon emissions. The overall significance of the model is extremely high, and the F statistic is 300.989 ($p = 0.000$), which shows that the model has a good fitting effect on the data and can well explain the variation of the dependent variable. The value of r is 0.118 and R^2 (within) is 0.367, which shows that the model has strong explanatory power within the group (that is, after excluding municipalities directly under the central government).

This study further verifies the negative impact of digital transformation on carbon emissions of construction enterprises through panel regression analysis after excluding municipalities directly under the central government. The results are shown in Table 7. This is consistent with the research conclusions on the relationship between digital transformation and carbon emission reduction in the existing literature. (Zhou and Chu, 2024; Zhu et al., 2023; Zhao et al., 2024; He et al., 2024; Liu et al., 2025; Qian et al., 2024; Guangqing et al., 2022). Digital transformation has significantly reduced the carbon emission intensity of construction enterprises by promoting the research and application of green technologies, improving resource utilization efficiency and optimizing production processes. In addition (Yang and Wang, 2021; Jing et al., 2018), this study also found that other variables, such as the nature of property rights, the proportion of tangible assets, and the size of the board of directors, have no significant impact on carbon emissions after excluding municipalities directly under the central government, which may be due to the particularity of municipalities in terms of economic development level and policy environment. After eliminating these particularities, the impact of these variables on carbon emissions is no longer obvious.

6.2.3 Endogenous problems

This paper adopts the following two methods to deal with endogenous problems. One is instrumental variable method. In this paper, the number of post offices in each prefecture-level city in 1984 is selected as the tool variable (iv) of enterprise digital transformation, which reflects the early informatization development level of each prefecture-level city and the active degree of folk information exchange. (Zhu et al., 2023). As shown in column (1) of Table 8, in the first stage of regression, the iv coefficient is 0.0001, which is significant at the level of 1%, indicating that instrumental variables are strongly correlated with endogenous variables, meeting the requirements of correlation. As shown in column (2) of Table 9, in the second-stage regression, the Dig coefficient is −0.001, which indicates that digital transformation significantly reduces carbon emissions and supports the H1 hypothesis. Because the main function of post office is

TABLE 6 Influence of digital transformation on carbon emissions of construction enterprises.

Variable	(1)	(2)	(3)	(4)
Digital	intenCo2	intenCo2	intenCo2	intenCo2
const	−0.0001***	−0.0001***	−0.0001***	−0.0001***
age				0.031
roe				0.011
inlev				0.110
cflow				0.658
gprofit				0.005
growth				−0.005
board				0.165
indep				1.169
Top1				0.006
Tangibi				−2.523
equity				−0.187
dual				−0.231
Time-fixed effect	YES	YES	YES	YES
Regional fixation effect	No	YES	YES	YES
Firm fixed effect	No	No	YES	YES
sample size	805	805	805	805
R2	0.184	0.183	0.182	0.083

TABLE 7 Excluding the regression results of municipalities directly under the central government.

Variable	(1)
	intenCo2
Digital	−0.0001* (0.020)
Control variable	YES
Time-fixed effect	YES
Regional fixation effect	YES
Firm fixed effect	YES
sample size	738
R2	0.118

TABLE 8 Robustness test: replacing core explanatory variables and explanatory variables.

Variable	(1)	(2)	(3)	(4)
	intenCO2	intenCO2	intenCO2	SO2
Dig1	−0.026**			
Dig2		−0.103***		−151596.814***
Dig3			−0.001***	
Time-fixed effect	YES	YES	YES	YES
Regional fixation effect	YES	YES	YES	YES
Firm fixed effect	YES	YES	YES	YES
sample size	1,118	1,118	1,118	1,118
R2	0.072	0.060	0.177	0.008

TABLE 9 Endogenous test results.

Variable	2SLS regression		Pilot policy regression	
	Dig	CO2	Regional CO2	Regional CO2
Iv Telephone * Number of Internet users in the last year	0.0001** (7.688)			
Dig		−0.001** (−2.786)		
du*Dig			161.637** (6.727)	
Lag1_du*dig				148.335** (6.080)
cons	200.672 (1.727)	−4.082** (−15.499)	4595.951** (3.438)	4277.729** (3.170)
Time-fixed effect	no	no	no	no
Regional fixation effect	no	no	no	no
sample size	478	478	292	292
R^2	0.317	0.937	0.190	0.169

information transmission, it is not directly related to the carbon emission reduction performance of construction enterprises, so it also meets the exogenous requirements. As shown in columns (3) and (4) of Table 9, in the regression analysis of the pilot policy, the coefficient of du*Dig reached 161.637, which reflected that the pilot policy may weaken the emission reduction effect of digital transformation due to the expansion of high energy-consuming industries; The coefficient of interaction term (lag1_du*dig) is 148.335, which indicates that the pilot policy has a certain persistence to this impact.

Meanwhile, based on the development history of the digital economy, Shenzhen and Hangzhou are considered the birthplaces of China's digital economy. Based on the correlation and exogenous conditions of instrumental variables, this study selects the natural logarithm of the nearest spherical geographic distance (lndis) from each enterprise to Shenzhen and Hangzhou as the instrumental variable for enterprise digital transformation, further testing endogeneity (Shang Y. et al., 2023).

In order to control for sufficient fixed effects, time-varying instrumental variables were also constructed. Construct the interaction term between lndis and the time trend term (lndis x year), and continue to use two-stage least squares estimation (2SLS) for instrumental variable regression (Table 10). The WaldF statistics are all greater than the critical value of 10% in the Stock Logo test, indicating that the instrumental variables are reasonable and reliable. The regression coefficients of the enterprise digitalization index are significantly negative, which still proves the robustness of the conclusion.

The lagged regression results show that when regressing on the explanatory variable of enterprise digital transformation lagging behind by one period, the results are also significant (Table 11). (Coefficient less than 0, $R^2 > 0.75$).

6.3 Mechanism analysis

According to the previous empirical analysis results, digital transformation can significantly reduce the carbon emissions of construction enterprises, that is, it supports the view of carbon emission reduction of digital transformation. So this part will analyze the carbon emission reduction mechanism of digital transformation from three dimensions: enterprise innovation,

TABLE 10 Endogenous test results.

Variable	IV1:lndis		IV2:lndis × Year	
	Dig	CO2	Dig	CO2
Iv1:lnDis (HangZhou)	−0.134* (−1.83)			
Dig		−0.038*** (−2.786)		
Iv2:lndis × Year			−0.002*** (−3.638)	
Cons	−11.128 (−6.572)	−15.95*** (−15.950)	315.691*** (36.426)	4277.729** (3.170)
Time-fixed effect	No	No	No	No
Regional fixation effect	YES	YES	YES	YES
sample size	1,118	1,118	1,118	1,118
Wald F statistic	Wald = 8450.151 P = 0.000***		Wald = 785.396 P = 0.000***	
R^2	0.859	0.88	0.163	0.169

TABLE 11 Lag regression results.

Variable	Lagged regression
Constant	−4.157** (−3.703)
Lag1_dig	−0.0001* (−2.405)
R^2	0.766
N	208
F	F (14,160) = 15.561,p = 0.000
Code	*YES*
Year	*YES*

efficiency improvement and optimization of production and operation structure.

Innovative mechanism of digital transformation.

6.3.1 Innovative mechanism of digital transformation

First of all, the digital transformation of construction enterprises has a positive impact on the carbon emission reduction effect of construction enterprises by promoting green technology innovation. With the wide application of digital technology, construction enterprises can integrate internal and external resources more effectively and accelerate the development and application of new technologies. As an important achievement of technological innovation of construction enterprises, the increase in the number of green patents directly reflects the investment and effectiveness of construction enterprises in environmental protection and energy-saving technologies. In this paper, the logarithm of the number of green patents (lngpat) is used to measure the green technology innovation level of enterprises. The empirical analysis shows that the degree of digital transformation is positively correlated with the number of green patents, which shows that digital transformation helps to improve the output efficiency of green patents, promote the application of green patents in actual construction production, and stimulate the green innovation ability of construction enterprises, thus reducing the intensity of carbon emissions. Second, the increase in R&D investment of construction enterprises is also an important embodiment of digital transformation to promote technological innovation. As shown in column (2) of Table 4, there is also a significant positive correlation between the degree of digital transformation and the R&D investment of construction enterprises, which means that digital transformation provides more R&D resources and motivation for construction enterprises and promotes the R&D and application of energy-saving and emission-reduction technologies. Digital transformation enables construction enterprises to more accurately grasp the construction market demand and construction technology trends, thus rationally allocating R&D resources and improving R&D efficiency. This increase in investment not only promotes the birth of new technologies, but also accelerates the optimization and upgrading of existing technologies, further promoting the realization of carbon emission reduction.

6.3.2 Efficiency improvement mechanism of digital transformation

According to the previous empirical analysis, the digital transformation of construction enterprises is not only conducive to improving the total factor productivity, but also may improve the production efficiency by optimizing the production process and reducing the production interruption of construction products, so as to help construction enterprises reduce the carbon emission effect. In this paper, the total factors measured by OP method and LP method are used as proxy variables of enterprise production efficiency. According to the results of correlation regression analysis, as shown in columns (3) and (4) of Table 4, there is a significant positive correlation between the degree of digital transformation and total factor productivity, indicating that digital transformation is helpful to improve the production efficiency of construction enterprises. The improvement of production efficiency means that under the same output, the

resources consumed by construction enterprises and the carbon emissions generated will be reduced. The SPCS system of prefabricated concrete building innovatively developed by Sany Construction Company simplifies the construction process and realizes "full prefabrication of wall column beam slab + full assembly on the ground and underground". This technical system has the overall advantages of "cavity overlapping and post-pouring, equivalent heterogeneity is fast and economical", which directly hits the pain points of building structure safety, water leakage, external wall insulation, labor shortage, low efficiency and high cost, significantly improves the production efficiency of all factors, reconstructs the ecology of construction industry, and realizes the low-carbon and high-quality development of construction enterprises (Pingru et al., 2023).

6.3.3 Optimization of production and operation structure

First of all, the wide application of digital technology can significantly improve the utilization efficiency of resources in the whole life cycle of construction enterprises, thus reducing the discharge of various polluting construction wastes. In this paper, the operating cost ratio (cost/revenue) is used to measure the negative index of the optimization degree of enterprise production and operation. Secondly, we improve the operating efficiency of construction enterprises through internal control and realize the diversified and hierarchical development strategy of enterprises. We adopt the "Internal Control Index of listed companies in Dibo China" (InnCt for short) widely recognized by academic circles as an indicator to measure the level of internal control, which comprehensively covers many key dimensions such as internal environment, risk assessment, control activities, information communication and internal supervision. The results of correlation regression analysis show that, as shown in columns (5) and (6) of Table 12, the degree of digital transformation is negatively correlated with the operating cost rate, and the coefficient is significant, indicating that digital transformation is helpful to reduce the operating cost of construction enterprises. This cost reduction may come from the improvement of production efficiency and the optimization of resource utilization, which will help to reduce carbon emissions. Combined with the above, the research hypothesis H2 can be verified. By using digital technologies such as big data and artificial intelligence, construction enterprises can realize automatic and intelligent production, reduce repetitive and low-value manual tasks, and greatly reduce labor costs and time costs, thus achieving cost reduction and efficiency improvement. Use intelligent building management system to optimize energy use and reduce energy consumption and emissions. At the same time, construction enterprises can also use digital technology to improve the accuracy of carbon accounting, enhance the compliance and transparency of carbon emission information disclosure, promote carbon market transactions, and achieve sustainable development.

6.3.4 Analysis of intermediary role

This study explores in depth the three pathways through which digital transformation affects corporate carbon emission intensity by constructing a mediation effect model, namely, the R&D innovation pathway, the total factor productivity (TFP) improvement pathway, and the operational cost optimization pathway.

1. R&D Innovation (R&D) Path

The results of the mediation effect test indicate that digital transformation significantly promotes the R&D innovation activities of enterprises (the regression coefficient of dig on R&D is 0.0009, $p < 0.01$) (Table 13) The increase in R&D intensity significantly reduces the carbon emission intensity of enterprises (the regression coefficient of R&D on intCO2 is −0.0003, $p < 0.01$)。 This indicates that digital transformation can indirectly reduce carbon emission intensity by promoting green technology research and development. Verified the theoretical logic of digital transformation achieving environmental benefits through technological empowerment.

2. Path to Total Factor Productivity (TFP) Improvement

Although the direct impact of digital transformation on total factor productivity is not significant (the regression coefficient of dig on TFP is −0.0004, $p > 0.05$), However, the significant improvement in total factor productivity has reduced carbon emission intensity (the regression coefficient of TFP for intCO2 is −0.0003, $p < 0.01$) (Table 13). This indicates that digital transformation can indirectly suppress carbon emissions by optimizing the efficiency of factor allocation. Its economic explanation is that digital transformation improves resource utilization efficiency through data-driven decision-making, intelligent production scheduling, and other functions, thereby reducing energy consumption and carbon emissions while maintaining the same output level.

3. Optimization path for operating costs

Digital transformation significantly reduces the operating costs of enterprises (the regression coefficient of dig for cost is −0.0001, $p < 0.1$), And there is a positive correlation between cost optimization and carbon emission intensity (the regression coefficient of Cost to intCO2 is 0.0013, $p < 0.01$) (Table 13). This indicates that while digital transformation reduces operating costs through mechanisms such as lean management and supply chain collaboration, it also indirectly reduces carbon emission intensity. The inherent logic is that cost optimization often accompanies the improvement of energy efficiency and the reduction of waste, such as reducing transportation energy consumption through intelligent logistics systems or reducing equipment idle through predictive maintenance.

6.4 Heterogeneity test

Under the background of different attributes of construction enterprises, the carbon emission reduction effect brought by their digital transformation may show asymmetric characteristics. In view of this situation, it is helpful to formulate more accurate and differentiated policy guidance to deeply discuss and distinguish the carbon emission reduction effects of digital transformation of

TABLE 12 Mechanism analysis results.

Variable	Enterprise innovation		Efficiency improvement		Optimization of production and operation structure	
	Lngpat	RD	TEP_OP	TFP_LP	Cost	Innct
	(1)	(2)	(3)	(4)	(5)	(6)
digital	0.001**(0.0094)	0.053*(0.0264))	0.001***(0.0001)	0.003***(0.0001)	−0.001*(0.0031)	−0.001***(0.0001)
Control variable	YES	YES	YES	YES	YES	YES
Time-fixed effect	YES	YES	YES	YES	YES	YES
Regional fixation effect	YES	YES	YES	YES	YES	YES
Sample size	1,118	1,118	1,118	1,118	1,118	1,118
R^2	0.126	0.234	0.344	0.124	0.231	0.365

TABLE 13 Intermediary model results (RD).

Variable	R&D	intenco2	intenco2
Constant	−1.4879* (−2.2040)	−5.2209** (−17.8752)	−5.0849** (−17.6881)
dig	0.0009** (3.9934)	−0.0003** (−3.4576)	−0.0004** (−4.3140)
R&D			0.0914** (4.4187)
N	434	434	434
R^2	0.7563	0.9435	0.9460
F	F (14,419) = 92.8990,p = 0.0000	F (14,419) = 499.5128,p = 0.0000	F (15,418) = 488.1262,p = 0.0000
Variable	TFP_FE	intenco2	intenco2
Constant	−12.9142** (−4.7034)	−5.2544** (−18.1002)	−5.0106** (−17.1286)
dig	−0.0004 (−0.5329)	−0.0003** (−3.5680)	−0.0003** (−3.5247)
TFP			0.0189** (3.9470)
N	488	488	488
R^2	0.4489	0.9363	0.9383
F	F (13,474) = 29.6959,p = 0.0000	F (13,474) = 535.6878,p = 0.0000	F (14,473) = 513.8367,p = 0.0000
Variable	Cost	intenco2	intenco2
constant	0.1457 (1.4562)	0.3209 (0.4198)	0.3835 (0.5008)
digi	−0.0001 (−1.8354)	0.0014** (5.3725)	0.0013** (5.2536)
Cost			−0.4295 (−1.2285)
N	488	488	488
R^2	0.1946	0.4317	0.4335
F	F (10,477) = 11.5274,p = 0.0000	F (10,477) = 36.2380,p = 0.0000	F (11,476) = 33.1160,p = 0.0000

*$p < 0.05$ **$p < 0.01$.

different construction enterprises. In order to achieve this goal, this paper will comprehensively examine the digital transformation practice of construction enterprises, and make detailed classification discussions according to regional differences, property rights and digital technologies adopted.

6.4.1 Regional heterogeneity

China has a vast territory, and the level of digital development varies among different regions, which may lead to digital "access barriers". In order to further explore the impact of digital access level, this article divides all regions of China into eastern, central,

western, and northern regions, and analyzes listed companies based on provinces. The eastern region usually refers to a series of provinces and municipalities directly under the central government in China's administrative divisions. The listed companies in the eastern region covered in this article include six provinces and cities: Anhui, Fujian, Jiangsu, Shandong, Shanghai, and Zhejiang. The eastern region is usually economically developed, with high levels of industrialization and urbanization. At the same time, it is also at the forefront of China's opening-up and economic development. Rich in natural resources, convenient transportation network, highly concentrated scientific and educational resources, and relatively complete market system, it occupies an important position in China's economic development. The results in Table 10, Figures 2, 3 indicate that there are significant differences in carbon reduction effects among different regions of China during the digitalization process of enterprises, and also demonstrate the regional heterogeneity of carbon reduction effects in H3 digital transformation. The carbon reduction effect of enterprise digital transformation is significant in the eastern region, while the coefficient is positive but not significant in the central and southern regions. Figure 1 Bubble map of digital transformation degree in 2012.

Eastern region (coefficient −0.054, $p < 0.01$): Digital transformation has a significant inhibitory effect on carbon emissions, consistent with theoretical expectations (Table 14). This may be due to the strong technological foundation of enterprises in the eastern region, where digital transformation focuses more on optimizing energy structure and iterating green technologies; Central region (coefficient 0.128, $p = 0.082$): The positive effect is close to significant, or due to the fact that central enterprises are in a critical period of industrialization, digital transformation is accompanied by capacity expansion in the short term, leading to an increase in carbon emissions; Southern region (coefficient 0.030, $p = 0.109$): The positive effect is not significant and may be related to the outward oriented economic characteristics of the southern region. Digital transformation focuses on market response rather than production side emission reduction. Further grouping regression results indicate that the digital transformation of enterprises in the eastern region has shown significant carbon emission reduction dividends, while the central and western regions need supporting structural reforms to unleash the potential of digital technology.

Enterprises in the central region are in a critical period of industrialization, and digital transformation may be accompanied by capacity expansion in the early stages to meet market demand and enhance competitiveness. During this process, equipment updates, technological investments, and expansion of production scale may temporarily increase energy consumption and carbon emissions, thereby masking the long-term potential carbon reduction effects of digital transformation. Although the central region is accelerating its digital infrastructure layout, compared to the eastern region, its digital technology application may still be in a relatively early stage. If digital transformation only stays on the surface and fails to deeply promote green technology innovation, optimize production processes and management structures, it will be difficult to achieve significant carbon reduction effects.

6.4.2 Heterogeneity of property rights

Faced with technological innovation in digital transformation, construction enterprises of different ownerships often adopt different response strategies due to different policy environments and target incentives. In order to explore the impact of ownership differences on carbon reduction in digital transformation, this paper constructs an interaction term (digital × soe) between the ownership logic variables of construction enterprises and digital transformation, to test whether there are significant differences in the impact of different property rights on carbon reduction in the process of digital transformation. Further classify state-owned

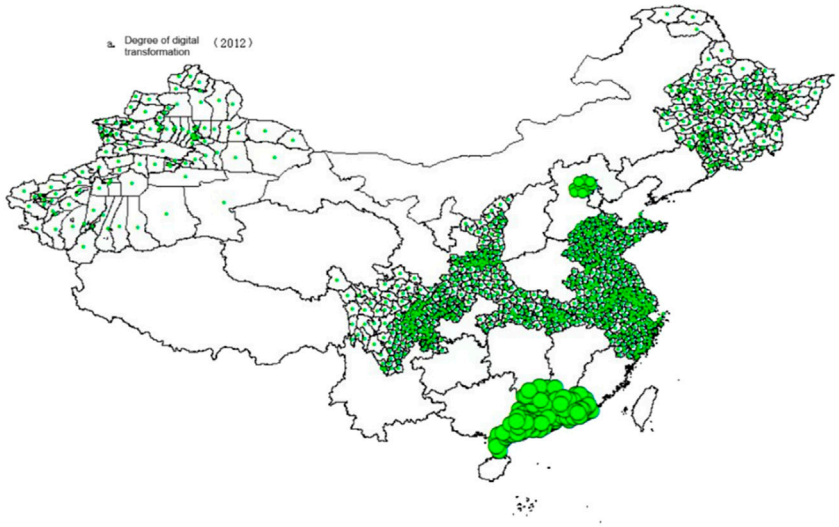


FIGURE 2
Bubble map of digital transformation degree in 2012.

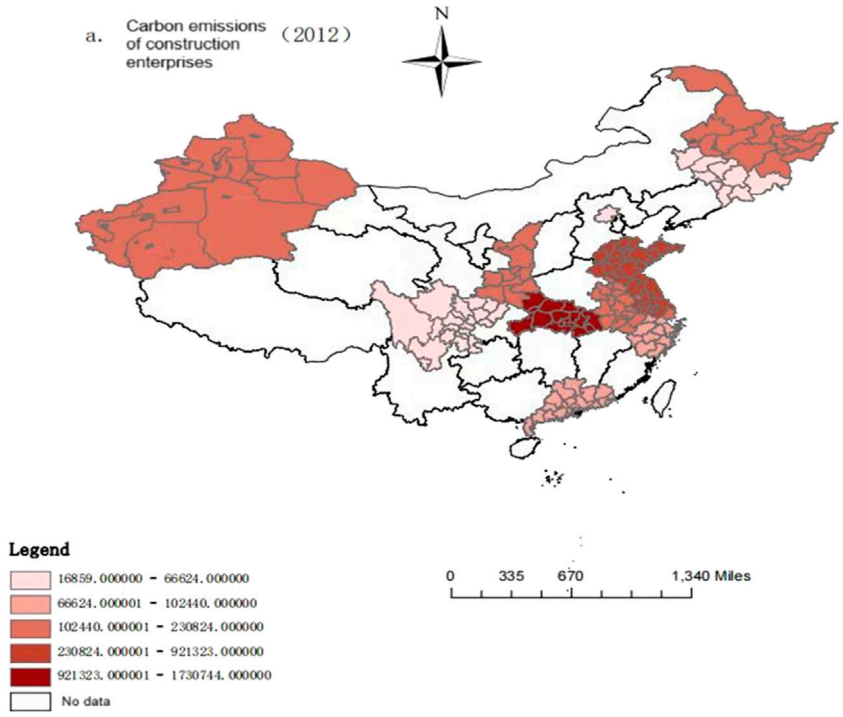


FIGURE 3
Thermal map of carbon emission of construction enterprises in 2012.

TABLE 14 Group regression results.

Variable	Whole	East	North	Central	South
Constant	16.949* (2.271)	12.030 (1.186)	42.412 (1.164)	−861.647 (−1.387)	−9.930 (−0.884)
Digital	−0.013 (−1.020)	−0.054** (−3.049)	−0.020 (−0.423)	0.128 (1.821)	0.030 (1.635)
N	1,118	473	307	71	267
R^2	0.826	0.890	0.924	0.980	0.818
F	F (16,336) = 99.415,p = 0.000	F (16,141) = 71.567,p = 0.000	F (16,43) = 32.732,p = 0.000	F (16,15) = 46.166,p = 0.000	F (16,86) = 24.205,p = 0.000
Code	YES	YES	YES	YES	YES
Year	YES	YES	YES	YES	YES

*p < 0.05 **p < 0.01.

enterprises and non-state-owned enterprises based on their controlling nature, and conduct group regression analysis. As shown in the first column of Table 15, the cross-term coefficient is significant, which indicates that the effect of digital transformation of state-owned construction enterprises on reducing carbon emission intensity may not be as significant as that of non-state-owned construction enterprises, or it may even lead to an increase in carbon emission intensity. The carbon reduction coefficient of state-owned enterprises is positive and not significant, while that of non-state-owned enterprises is negative and significant, further indicating that the carbon reduction effect of state-owned construction enterprises is not as significant as that of non-state-owned construction enterprises. The verification study assumes that the carbon emission reduction effect of H3 digital transformation has property rights heterogeneity. First of all, it is difficult for state-owned construction enterprises to make full use of the emission reduction potential brought by digital transformation due to complicated decision-making process, traditional management mode and slow improvement of operational efficiency. In contrast, non-state-owned construction enterprises may be able to adopt new technologies and modes such as BIM and intelligent construction more quickly because of their higher autonomy and flexibility, so as to achieve rapid improvement of operational efficiency and thus reduce carbon emission intensity more effectively. Secondly, the carbon emission content of non-state-owned construction enterprises is higher than that of state-

TABLE 15 Property rights grouping regression results.

Variable	Whole	State controlled	Non-state-owned holding
Constant	−5.9998** (−18.3964)	−6.9356** (−5.1078)	−6.1830** (−18.0836)
dig	−0.0004** (−2.6079)	0.0009 (1.3301)	−0.0003* (−2.2062)
N	316	39	277
R^2	0.8133	0.8772	0.8195
F	F (12,303) = 109.9699,p = 0.0000	F (12,26) = 15.4760,p = 0.0000	F (12,264) = 99.8825,p = 0.0000
Code	*YES*	*YES*	*YES*
Year	*YES*	*YES*	*YES*

*$p < 0.05$ **$p < 0.01$.

owned construction enterprises. After digital transformation, the marginal benefit of carbon emission of non-state-owned construction enterprises is higher than that of state-owned construction enterprises, so the carbon emission reduction effect of state-owned construction enterprises may not be as good as that of non-state-owned construction enterprises. Finally, state-owned construction enterprises use more ways to reduce production to achieve energy-saving goals, while non-state-owned construction enterprises use more ways to improve energy efficiency.

6.4.3 Heterogeneity of digital technology

Through Wu Fei's research (2021), we found that various digital technologies are helpful to promote carbon emission reduction. Therefore, we divide the indicators of digital transformation into two categories: "the bottom technology level" and "the practical application level", including five important sub-projects: AI, (BC) blockchain, CC, big data (DT) and (ADT). The empirical analysis results are shown in columns (3)–(7) of Table 16. By comparing the data of various indicators, it can be seen that digital technologies such as artificial intelligence, blockchain, cloud computing and big data have a significant negative impact on carbon emission intensity. These technologies not only optimize building energy consumption through intelligent control and management, but also use blockchain technology to ensure accurate tracking and management of carbon emission data. They also use cloud computing technology to improve the energy efficiency of data centers and realize remote monitoring. At the same time, big data technology digs deep into energy consumption data to formulate optimization strategies. In addition, through digital management, these technologies have played a key role in the whole life cycle of prefabricated buildings, such as design, production, transportation, installation and operation and maintenance. Especially in the field of prefabricated buildings, digital technology not only optimizes production planning and logistics transportation, reduces energy consumption and emissions, but also realizes accurate prediction and optimization of building energy consumption through intelligent control and management systems. At the same time, the application of big data technology provides rich energy consumption data support for construction enterprises to help them formulate more scientific carbon emission reduction strategies. The deep integration of these digital technologies and prefabricated buildings has jointly promoted the green transformation of the construction industry and laid a solid foundation for achieving the goal of sustainable development. To sum up, these digital technologies have jointly acted on the carbon emission reduction practice of construction enterprises in various ways, which has promoted the sustainable development of the industry, especially in key links such as prefabricated buildings and digital management.

6.4.3.1 Analysis of the effects of dual carbon policy

In September 2020, China proposed a policy of dual carbon targets, requiring carbon peaking by 2030 and carbon neutrality by 2060. To explore the actual impact of the policy on construction enterprises, this article sets 2021 as a segmented time point to study the effect of digital transformation and carbon reduction on construction enterprises before and after 2021. And perform a double difference test (Figure 4).

The control group represents that the enterprise has not undergone digital transformation. After refers to 2022–2023, and before refers to 2000–2021 (Tables 17–19).

Table 18 shows the grouping of sample enterprises before and after policy implementation ("Before" for 2000–2021 and "After" for 2022–2023). Before the implementation of the policy, the average carbon emission intensity (intCO2) of the control group and the experimental group were 3.2094 and 3.1469, respectively (Table 18), with no significant difference (t = −1.0734, p = 0.2841). This indicates that before policy intervention, the carbon emission intensity of the two groups of enterprises was at a similar level, meeting the prerequisite for the parallel trend assumption of the DID model. After the implementation of the policy, there was a significant change in the carbon emission intensity between the control group and the experimental group. The mean experimental value decreased to 0.6494, while the mean of the control group increased to 2.1852, and the inter group difference reached a significant level (t = 8.7548, p = 0.0000). This preliminary result suggests a correlation between digital transformation and changes in carbon emission intensity.

Trend item x Grouping variable Policy interaction term: The coefficient of this interaction term is −1.5708 (p = 0.000) and is significant at the 1% level. This indicates that after the implementation of the policy, the carbon emission intensity of

TABLE 16 Heterogeneity test results.

Variable	Heterogeneity of property rights	Regional heterogeneity	Heterogeneity of digital technology				
	Intenco2	Intenco2	Intenco2	Intenco2	Intenco2	Intenco2	Intenco2
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Digital	−0.006 (0.580)	−0.031***(0.002)					
Digital × soe	0.42***(0.000)						
Digital × region		0.139***(0.003)					
AI			−0.038***(0.000)				
BC				−0.011***(0.010)			
DT					−0.013***(0.001)		
CC						−0.014***(0.004)	
ADT							−0.013***(0.001)
Control variable	YES	YES	YES	YES	YES	YES	YES
Time-fixed effect	YES	YES	YES	YES	YES	YES	YES
Regional fixation effect	YES	YES	YES	YES	YES	YES	YES
Sample size	1,118	1,118	1,118	1,118	1,118	1,118	1,118
R^2	0.072	0.031	0.028	0.006	0.045	0.031	0.045

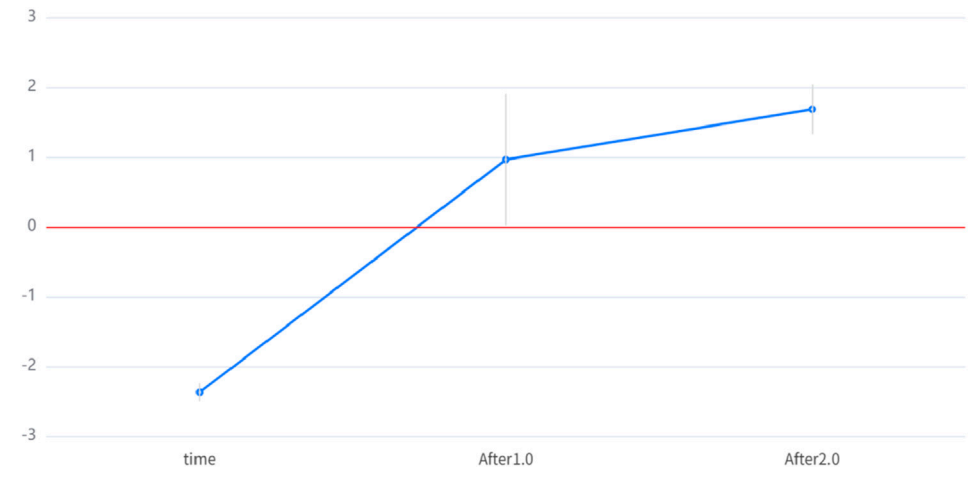


FIGURE 4
Event study method result chart.

the experimental group decreased by 1.5708 units compared to the control group. This result contradicts the expected hypothesis, and further decomposition of policy effects reveals that before policy implementation, there is no significant difference in carbon emission intensity between the experimental group and the control group (Diff = −0.0315, p = 0.6242), further validating the parallel trend hypothesis.

TABLE 17 DID model description statistics.

Variable	Before	After	All	
Control		238	31	269
Treated		35	9	44
ALL		273	40	313

TABLE 18 Comparison before and after the dual carbon policy.

T-test (Before)					
	Treated *N* = 238	Control *N* = 35	Diff	*t*	*p*
intco2	3.2094	3.1469	−0.0625	−1.0734	0.2841
T-test (After)					
	Treated *n* = 31	Control *n* = 9	Diff	*t*	*p*
intco2	0.6494	2.1852	1.5358	8.7548	0.0000**

*p < 0.05 **p < 0.01.

After policy implementation: The carbon emission intensity of the experimental group significantly increased compared to the control group (Diff = −1.5393, p = 0.0000), consistent with the results of the interaction term. Double difference estimator (Diff in Diff): The net effect is −1.5708 (p = 0.0000), indicating that after policy implementation, digital transformation resulted in an additional reduction of 1.5708 units in carbon emission intensity in the experimental group compared to the control group. Therefore, the dual carbon policy can promote enterprises to implement digital transformation and reduce carbon emissions.

7 Conclusions and suggestions

7.1 Research conclusion

Based on the results of empirical analysis, this paper draws the following conclusions:

(1) As the main source of carbon emissions, construction enterprises bear a decisive responsibility in the process of realizing the goal of "double carbon" in China. Although limited by the difficulty of data acquisition, relevant research was relatively scarce in the past, through the detailed data processing and analysis of listed companies in China from 2000 to 2021, this paper not only verified the positive impact of digital transformation of construction enterprises on carbon emission reduction, but also deeply analyzed the transmission mechanism behind it, which filled an important gap for the research in this field. The research results show that digital transformation not only significantly reduces the carbon emission intensity of construction enterprises, but also strongly supports this conclusion in various robustness tests, further proving its stability and reliability.

(2) The reason why digital transformation can effectively promote carbon emission reduction is mainly due to its multiple functions in technological innovation, production

TABLE 19 DID model results.

OLS regression analysis results					
Variable	β	95% CI			
constant	2.4669** (6.0250)	1.6611–3.2726			
Trend items (with or without digital transformation)	−0.0315 (−0.4904)	−0.1578–0.0949			
Grouping Variables - Policy	−2.5938** (−40.2377)	−2.7206 ~ −2.4669			
Trend item (with or without digital transformation) × Grouping variable - Policy	−1.5708** (−11.4118)	1.2999–1.8416			
N	313				
R^2	0.8594				
F	*F* (15,297) = 121.0542,*p* = 0.0000				
Dependent variable: intCO2					
Grouping variables - policy	Item	Efect value intco2	Standard error	*t*	*p*
Before	Control	2.4669			
	Treated	2.4354			
	Diff (T − C)	−0.0315	0.0642	−0.4904	0.6242
After	Control	−0.1269			
	Treated	1.4124			
	Diff (T − C)	−1.5393	0.1226	−12.5603	0.0000**
Diff-in-Diff		−1.5708	0.1376	−11.4118	0.0000**
R^2 = 0.8594					

*p < 0.05 **p < 0.01.

efficiency improvement and optimization of production and operation structure. Specifically, digital transformation has stimulated the green innovation ability of enterprises, promoted the research and development and application of green technologies, such as intelligent building design, energy efficiency management system, etc., significantly improved the utilization efficiency of resources and reduced unnecessary energy consumption. At the same time, digital transformation also reduces the operating cost rate of enterprises by optimizing production processes and improving total factor productivity, thus indirectly reducing carbon emissions.

(3) The effect of digital transformation on carbon emission reduction of construction enterprises with different property rights is heterogeneous. Due to the differences in resource endowment and management mechanism, state-owned enterprises and private enterprises have shown different carbon emission reduction effects in the process of digital transformation. However, from a regional perspective, the carbon emission reduction effect of digital transformation does not show obvious regional differences, which shows that the green and low-carbon effect of digital transformation has universal applicability.

7.2 Policy recommendations

This study also puts forward specific measures to promote the digital transformation of enterprises and carbon emission reduction from many angles.

(1) Construction enterprises should increase investment in research and development of digital green building technology, use digital technology to tap innovative points and improve energy efficiency. In terms of technological innovation, we should continue to increase investment in research and development of digital technology, continue to attack low-energy and high-efficiency technologies, promote the transformation and application of technological achievements, and use digital technology to tap innovative points in the fields of building energy conservation and renewable energy utilization. For example, through digital simulation technology to optimize the architectural design scheme, improve the energy efficiency of buildings; Using Internet of Things technology to monitor building energy consumption in real time and realize accurate management. Further leverage the promotion effect of digitalization of listed companies on total factor productivity. We should accelerate the upgrading and application of digital technology, deepen the application of digital technology in research and development design, production and manufacturing, warehousing and logistics, and marketing services, and achieve digitization from the production end to the sales end. Secondly, enterprises should accelerate the establishment of digital factories and smart factories, increase investment in intelligent equipment such as artificial intelligence, robots, interconnected platforms, and smart sensors, accurately control product production efficiency and quality, and gradually transform towards intelligent manufacturing. Enterprises should transform their operational thinking, enhance their IT awareness, further expand their investment in software and hardware, and increase their digital investment. Finally, we should innovate the traditional self support model, constantly explore new business models through the Internet and e-commerce platforms, such as the current emerging live broadcast sales, and keep up with consumer demand and market changes.

(2) Using digital tools and energy management system to optimize the process, and introducing intelligent technology to realize efficient use of energy. In improving operational efficiency, we should make full use of digital project management tools and energy management systems, optimize the construction process, rationally arrange resources, and reduce energy waste and carbon emissions during construction. At the same time, the introduction of intelligent equipment and technology, the transformation and upgrading of high energy-consuming equipment, to achieve efficient use of energy.

(3) Strengthen exchanges and cooperation among construction enterprises, and use digital financial support to promote the construction of green supply chain. In improving the business environment, we will share the experience and technology of energy conservation and emission reduction, and jointly promote the construction of green supply chain. In addition, we should make full use of the support of digital finance and apply for low-interest green loans to provide financial guarantee for green building projects and digital transformation.

(4) Pilot a digital carbon regulatory exemption mechanism in the free trade zone, granting a 3-year carbon emission quota exemption period to enterprises that apply new technologies such as blockchain traceability and AI energy efficiency management. Introduce smart contract technology to achieve automatic carbon trading clearing and cross chain mutual recognition, reducing transaction costs.

Establish a “data sandbox” regulatory model and pilot a blockchain carbon data certification platform in central provinces such as Hubei and Hunan. Through smart contracts, achieve hierarchical management of data access permissions between government regulatory nodes, enterprise data nodes, and third-party authentication nodes, ensuring a balance between data sharing and trade secret protection. Referring to the “regulatory chain” model in the financial field, a construction carbon chain responsibility traceability network is constructed, requiring high annual emission enterprises to access and achieve full chain penetration supervision of carbon emission data from design, construction to operation and maintenance, and implementing blockchain certification and accountability for data fraud.

7.3 Research limitations and future prospects

7.3.1 Research limitations

Although this study systematically analyzes the impact of digital transformation of construction enterprises on carbon emission

reduction and its mechanism, there are still the following limitations:

(1) Limited data coverage: The research sample only covers listed construction enterprises in China, excluding small and medium-sized unlisted enterprises and local construction companies. Due to the significant differences in management mechanism, resource endowment and technology application of unlisted enterprises, the universality of the research conclusions may be limited. In addition, carbon emission data depends on indirect measurement methods, which may not fully reflect the actual emission level of enterprises.
(2) Limitations of digital transformation measurement: Although text analysis method is used to construct digital transformation indicators, this method mainly relies on the frequency of keywords in annual reports, and it is difficult to comprehensively measure the actual depth and technology landing effect of digital transformation of enterprises. For example, some enterprises may only carry out superficial digital publicity without substantially promoting technology application.
(3) The depth of regional heterogeneity analysis is insufficient: although the regional differences are not significant, the specific path of realizing "late-comer advantage" through policy subsidies or technological advancement in the central and western regions has not been further explored, and the regulatory role of regional policy environment on the digital transformation effect has not been quantified.
(4) There are unavoidable data robustness issues in this article. It is generally believed that carbon dioxide is the main atmospheric pollutant, so it is difficult to replace the sample with carbon dioxide as the dependent variable. In this study, sulfur dioxide was used as a substitute to demonstrate the robustness of the data, which has inherent errors. Subsequent research can improve the measurement and data collection of carbon dioxide.

7.3.2 Future outlook

In view of the above limitations, future research can be expanded in the following directions:

(1) Expand the data dimension and sample range: integrate the data of unlisted enterprises and micro-project-level carbon emission data, and improve the accuracy of carbon emission measurement by combining field research or sensor data of the Internet of Things. At the same time, extend the research time span to capture the long-term dynamic relationship between digital transformation and carbon emission reduction.
(2) Deepen the measurement system of digital transformation: construct multi-dimensional indicators, covering the intensity of technology application (such as BIM utilization rate), employees' digital skills, supply chain collaboration level, etc., and verify the effectiveness of the indicators with case studies.
(3) Explore the synergistic mechanism of policy and technology: analyze the interactive effects of carbon trading market, green financial policy and digital transformation, for example, explore how the carbon price signal encourages enterprises to adopt intelligent emission reduction technology.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

Author contributions

SZ: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review and editing. HD: Conceptualization, Data curation, Formal analysis, Writing – original draft. JC: Writing – original draft, Writing – review and editing. MG: Conceptualization, Data curation, Formal analysis, Software, Writing – original draft, Visualization.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. Key Support Project of Hunan Province College Students' Innovation and Entrepreneurship Training Program (S20241,1532001). National College Students' Innovation and Entrepreneurship Training Program (S20241,1532005) (S20231,1532004). Hunan Provincial Social Science Fund (2022ZDB089).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Asif, M., Naeem, G., and Khalid, M. (2024). Digitalization for sustainable buildings: technologies, applications, potential, and challenges. *J. Clean. Prod.* 450, 450141814. doi:10.1016/j.jclepro.2024.141814

Borozan, D., and Pirgaip, B. (2024). Climate policy uncertainty and firm-level carbon dioxide emissions: assessing the impact in the U.S. market. *Bus. Strategy Environ.* 33 (6), 5920–5938. doi:10.1002/bse.3784

Changqi, T., Xie, T., and Ye, X. (2025). Synergy between green technology innovation and pollution reduction and carbon reduction-based on the perspective of joint emission reduction and environmental improvement. *Stat. Res.* 42 (04), 31–47. doi:10.19343/J.CNKI.11-1302/C.2025

Chen, J., Guo, Z., and Lei, Z. (2024). Research on the mechanisms of the digital transformation of manufacturing enterprises for carbon emissions reduction. *J. Clean. Prod.* 449, 141817. doi:10.1016/j.jclepro.2024.141817

Chen, J., and Zhang, J. (2024). Study on the influence of digital transformation of manufacturing enterprises in China on carbon emission intensity. *Stat. Theory Pract.* (04), 19–24.

Chen, X., Li, Y., Song, L., and Wang, Y. (2022). Theoretical system and research prospect of digital economy. *Manag. World* 38 (02), 208–224+13-16. doi:10.1974/j.cnki.11-1235/f.2020.0020

Cook, J. (2017). P. Research and development of carbon capture and storage technology and deployment of clean energy industry in the future-Australia's experience and lessons in the past 20 years. *Engineering* 3 (04), 103–121.

Culture, J., Xie, X., and Xiangge, Y. (2024). Research on the path and practice of promoting digital transformation in construction enterprises. *J. Railw. Eng. Soc.* 41 (01), 88–92.

Fu, C., and Pan, L. (2024). Research on the influence of blockchain technology on the total factor productivity of enterprises-from the perspective of trust construction and cost reduction and efficiency improvement. *J. Hangzhou Dianzi Univ. Soc. Sci. Ed.* 20 (04), 26–41. doi:10.13954/j.cnki.hduss2024

Guangmiao, Xu (2023). Application of SPCS technology system in a project construction. *Guangdong Civ. Archit.* 30 (06), 32–34+60. doi:10.19731/j.gdtmyjz.2023.06.009

Guangqing, X., Zhang, W., and Chen, X. (2022). Estimation of macro-energy rebound effect considering technological progress and structural changes. *Econ. Theory Econ. Manag.* 42 (11), 26–41.

hai, W., Guo, G., and Yin, J. (2025). Going green in the process of transformation: industrial structure adjustment and green innovation of enterprises. *Res. quantitative Econ. Technol. Econ.* 42 (01), 93–115. doi:10.13653/j.cnki.jqte.20016.10010101116

Han, C., and Marco, W. Z. (2020). Study on the mechanism of pollution reduction effect of enterprises under the constraint of energy-saving targets. *China Ind. Econ.* (10), 43–61. doi:10.19581/j.cnki.ciejournal.2020.10.009

Han, S., Zhang, H., Li, H., and Xun, Z. (2025). Digital transformation and carbon emission reduction: the moderating effect of external pressure and support. *J. Clean. Prod.* 500, 145108. doi:10.1016/j.jclepro.2025.145108

He, L., Haoqing, Z., and Ke, F. (2025). Digital transformation and carbon reduction of enterprises. *J. Anhui Normal Univ. Soc. Sci. Ed.* 53 (01), 85–99.

He, M., Zhang, Y., Luan, S., Gao, H., Zhang, W., and Huang, J. (2024). Analysis of carbon emission of urban and rural residents in China based on LMDI-Tapio-GA-BP. *J. South China Normal Univ. Nat. Sci. Ed.* 56 (06), 1–16.

Huang, J., Lu, H., and Du, M. (2025a). Regional differences in agricultural carbon emissions in China: measurement, decomposition, and influencing factors. *Land* 14 (4), 682. doi:10.3390/land14040682

Huang, J., Lu, H., and Du, M. (2025b). Can digital economy narrow the regional economic gap? Evidence from China. *J. Asian Econ.* 98, 101929. doi:10.1016/j.asieco.2025.101929

Huang, J., Zheng, B., and Du, M. (2025c). How digital economy mitigates urban carbon emissions: the green facilitative power of industrial coagglomeration. *Appl. Econ.*, 1–19. doi:10.1080/00036846.2025.2486785

Hui, T., and Zhang, J. (2025). *Does the regional green transformation policy help to improve the technical complexity of export —— evidence based on the construction of low-carbon cities*. Social Science Edition. Harbin, China: Journal of Harbin University of Commerce, 114–128. 02.

Huipeng, L., and Zhou, X. (2023). Study on the impact of low-carbon city construction on digital transformation of enterprises —— empirical evidence from listed companies. *Industrial Technol. Econ.* 42 (05), 105–113.

Jing, L., Jin, C., and Huaqing, W. (2018). Measurement and analysis of industrial green biased technological progress based on water resources. *China Popul. Resour. Environ.* 28 (10), 131–142.

Julie, Lu, S., Liang, J., Wu, J., and Yang, Y. (2022). Energy saving and emission reduction analysis and evaluation of prefabricated buildings based on BIM technology. *Build. Econ.*, 43 (S2), 302–306. doi:10.14181/j.cnki.1002-851x.2022S2302

Jun-qian, Z., and Yong-kou, Z. (2025). Can digital transformation reduce carbon emissions of enterprises? —— evidence from listed companies in China. *Front. Eng. Manag. Sci. Technol.*, 1–13. Available online at: http://kns.cnki.net/kcms/detail/34.1013.n.2024128.0948.006.

Kathrin, B., and Cordula, K. (2023). Building a better world? Competing promises, visions, and imaginaries-in-the-making of the digitalization of architecture and construction. *Futures*, 154.

Lan, F., and Zhou, S. (2025). Digital transformation, moderate scale of enterprises and M&A behavior [J/OL]. *World Econ.* (04), 197–228. doi:10.19985/J

Lei, G. (2024). *Research on the mechanism of carbon emission reduction of intelligent building and enabling construction enterprises [D]*. Harbin, China: Harbin university of commerce. doi:10.287/d.cnki.ghrbs.20087.100000000606

Li, S., Suo, Y., Dai, L., Wang, L., Jiang, Z., Cao, X., et al. (2024). Meta earth and digital twin: breakthrough in ideas, technological change, and paradigm shift. *Front. Earth Sci*, 31 (01), 46–63. doi:10.13745/j.esf.sf.2041.48

Li, Y., Li, M., Hou, C., Gao, Y., Wang, Q., Tian, J., et al. (2024). Construction and implementation of a digital carbon management system for steel. *Carbon Neutrality Steel*, 59 (09), 13–21. doi:10.13228/j.boyuan.issn0449-749x.20240342

Li, Z., Yuan, P., Zhao, S., Deng, H., and Yuan, B. (2024). Research on crop planting strategy problem based on optimization greedy algorithm. Transactions on environment. *Energy Earth Sci.*

Liu, B., Qiu, Z., Hu, L., Hu, D., and Nai, Y. (2024b). How digital transformation facilitate synergy for pollution and carbon reduction: evidence from China. *Environ. Res.* 251 (P2), 118639. doi:10.1016/j.envres.2024.118639

Liu, B., Zhang, Y., and Peng, J. (2025). Synergy between digital transformation of chain owners and carbon emission reduction. *Bus. Res.* 1, 79–89. doi:10.13902/j.cnki.syyj.2025.01.006

Liu, H., Han, P., Wang, D., Wang, S., and Bao, H. (2024a). Decoding enterprise digital transformation: external oversight and carbon emission reduction performance. *J. Environ. Manag.* 359, 121039. doi:10.1016/j.jenvman.2024.121039

Liu, J., Chen, J., and Mu, R. (2023). Optimal design and application of production control of prefabricated buildings under the background of intelligent construction. *Archit. Struct.* 53 (S1), 1178–1182.

Luo, F., and Tang, C. (2020). Evolutionary game study on low-carbon emission reduction strategies of government and enterprises under the supervision of the third party. *Ecol. Econ.* 36 (04), 30–34.

Mai, W., Xiong, L., Liu, B., and Liu, S. (2025). Spatial–temporal evolution, drivers, and pathways of the synergistic effects of digital transformation on pollution and carbon reduction in heavily polluting enterprises. *Sci. Rep.* 15 (1), 11963. doi:10.1038/s41598-025-96834-6

Niu, L. (2023). *Analysis of the influencing factors of digital transformation of construction industry and its influence on sustainable development*. Taiyuan, China: Shanxi university. doi:10.284/d.cnki.gsxiu.20080.100000000016

Peak (2025). Adaptive adjustment and optimization path of fiscal and taxation policies from the perspective of new quality productivity. *J. Xi' Univ. finance Econ.*, 1–17. doi:10.19331/j.cnki.jxufe.20422.0001.000000000506

Peng, W., Su, X., Kuang, C., Yang, S., Zhao, D., and Wei, X. (2024). The impact and spatial effects of digital economy on carbon emission efficiency in the Changsha Zhuzhou Xiangtan metropolitan area. *J. Geogr.* 79 (11), 2915–2928.

Pingrui, L., and Xu, Z. (2023). Research on the transformation and development of the construction industry in the era of digital economy. *Earthq. Resist. Reinf. Eng.* 45 (05), 194.

Qian, J., Bo, Z. L., and Xiao 'ang, Z. (2024). Digital transformation and enterprise energy efficiency-also on the boosting effect of ESG performance. *Financial Theory Pract.* (12), 59–71.

Sanglin, Z. (2025). *Carbon emission data, digital transformation and corporate financial data of listed construction companies from 2000 to 2023*. Beijing, China: Science Data Bank. V1. doi:10.57760/sciencedb.19398

Shang, Y., Ali, R. S., Huo, Z., Shahzad, U., and Xin, Z. (2023a). Does enterprise digital transformation contribute to the carbon emission reduction? Micro-level evidence from China. *Int. Rev. Econ. Finance* 86, 1–13.

Shang, Y., Raza, S. A., Huo, Z., Shahzad, U., and Zhao, X. (2023b). Does enterprise digital transformation contribute to the carbon emission reduction? Micro-level evidence from China. *Int. Rev. Econ. and Finance* 86, 1–13. doi:10.1016/j.iref.2023.02.019

Shao, T., Zhang, S., Cheng, J., Gao, Y., Sun, H., Dou, L., et al. (2024). Renewable energy driven discharge plasma assisted dual carbon target strategy and route Chinese. *J. Electr. Eng.*, 44 (18), 7453–7479. doi:10.13334/j.0258-8013.pcsee.240320

Su, M., Yang, W., Li, B., Ren, F., Yi, M., Yang, C., et al. (2025). Promote the deep integration of artificial intelligence and real economy and accelerate the development of new quality productivity. *Ind. Technol. Econ.* 44 (04), 32–59.

Wang, K., Guan, R., Hu, M., Li, X., Gan, C., et al. (2023). Digital economy and carbon emission performance: a case study of 276 cities in China. *Environ. Sci. Res.* 36 (09), 1824–1834. doi:10.13198/j.issn.1001-6929.2023.07.23

Wang, L. (2024). Research on the transformation of China's government supervision in the era of digital economy. *Manag. world* 40 (03), 110–126+204+127. doi:10.19744/j.cnki.11-1235/f.2024.100000000406

Wang, P., Xiao, J., Xiao, X., and Zhu, H. (2024a). The Application and prospect of digital technology in construction engineering. *J. Tongji Univ. Nat. Sci. Ed.* 52 (07), 1068–1078.

Wang, W., Bao, B., and Chunjiao, Y. (2024b). External demand, digitalization and carbon emission performance of enterprises. *World Econ. Res.* (09), 77–91+136.

Wei, L., Song, K., Qi, Y., Sun, S., Ni, G., Zhou, C., et al. (2025). Effects of straw returning after anaerobic microbial pretreatment on soil carbon sequestration and emission reduction. *Environ. Technol. and Innovation* 38, 38104124–104124. doi:10.1016/j.eti.2025.104124

Wu, F., Hu, H., Lin, H., and Ren, X. (2021). Enterprise digital transformation and capital market performance: empirical evidence from stock liquidity Manag. *the World*, 37 (07), 130–144+10. doi:10.1974/j.cnki.11-1235/f.2021.097

Xiao, Y., and Kong, Q. (2025). The influence of sustainable development policy of resource-based cities on carbon emission efficiency. *J. Nat. Resour.* 40 (03), 833–854. doi:10.31497/zrzyxb.20250316

Xing, M., and Zhu, P. (2025). Research on the effect and mechanism of digital infrastructure driving green innovation —— policy experiment based on "broadband China" strategy. *Res. Econ. Manag.* 46 (01), 90–108. doi:10.13502/J.CNKI.ISSN1000-7636.2007.2000100100006

Yan, K., Gao, H., Feng, Z., Gao, Y., Lv, Y., Liao, K., et al. (2023). Thoughts on establishing a unified and standardized carbon accounting method system for industrial parks. *China Environ. Manag.* 15 (03), 5–16. doi:10.16868/j.cnki.1674-6252.2023.005

Yang, H., and Wang, X. (2021). Is environmental regulation an effective means to restrain the rebound effect of energy? *J. Manag.* 34 (05), 74–91. doi:10.19808/j.cnki.41-1408/f.2021.0038

Yang, S., Wang, J., Shi, L., Tan, Y., and Qiao, F. (2023). Theory and method of intelligent manufacturing engineering management for high end equipment in the new generation information technology environment. *Manag. World* 39 (01), 177–190. doi:10.1974/j.cnki.11-1235/f.2023.0009

Yang, X., Gao, Z., Sheng, Z., Huang, L., Zeng, D., Chen, X., et al. (2022b). Complex system management is an important component of the management system with Chinese characteristics. *Manag. World* 38 (10), 1–24. doi:10.1974/j.cnki.11-1235/f.2021.141

Yang, Y., and Liu, Y. (2024). Environmental regulations and firms' trade-offs between innovations: empirical evidence from the quasi-experiment in China. *J. Environ. Manag.* 370, 122819. doi:10.1016/j.jenvman.2024.122819

Yang, Y., Zhang, Z., Ma, Y., and Sun, X. (2022a). Analysis of digital transformation path of construction industry from the perspective of technical logic. *Res. Sci. Technol. Manag.* 42 (24), 137–142.

Yi, Y., and Su, Y. (2025). Spatial and temporal differentiation characteristics of carbon budget and driving factors of carbon balance in Guanzhong plain urban agglomeration [J/OL]. *china Environ. Sci.* 1-23. doi:10.19674/J.CNKI.ISSN1000-6923.1000100106

Ying, X., Jia, F., Chen, L., and Yonggui, W. (2024). Does digital transformation foster carbon emissions reduction? Evidence from China's manufacturing supply chain. *Int. J. Logist. Manag.* 35 (6), 1827–1846. doi:10.1108/ijlm-07-2023-0260

Zhang, H., and Shen, Q. (2024). Mixed ownership reform of state-owned enterprises promotes green innovation of enterprises-Empirical evidence based on the ownership structure of listed state-owned enterprises. *Financ. Rev.* 16 (03), 131–154+158.

Zhang, X., and Cui, X. (2025). Impact of cutting-edge hybrid electric vehicle technological innovation on carbon emissions in China. *Sustain. Futur.* 9, 100447. doi:10.1016/j.sftr.2025.100447

Zhang, X., and Yang, S. (2025). Study on the effect and mechanism of artificial intelligence on pollution reduction and carbon reduction [J/OL]. Modern Finance and Economics. *J. Tianjin Univ. Finance Econ.* (05), 77–94.

Zhang, Y. (2023). *Study on the impact of digital transformation of construction enterprises on carbon neutralization*. Xuzhou, China: China University of Mining and Technology. doi:10.27623/d.cnki.gzkyu.20080.000000000666

Zhang, Y., Jia, K., Cen, L., and Ma, J. (2024b). Does digital transformation promote corporate carbon reduction? Evidence from ChineseA-share listed companies. *Environ. Dev. Sustainability,(prepublish)*, 1–28. doi:10.1007/s10668-024-05074-7

Zhang, Z., Lai, F., and Cheng, Y. (2024a). The digital transformation and evolution of listed enterprises at city scale and its impact on carbon emissions. *Econ. Geogr.* 44 (05), 106–116.

Zhao, S., Hao, D., and Yua, B. (2025). Digital trade and common prosperity: evidence from China Province. *Cogent Econ. and Finance* 13 (1), 2494125. doi:10.1080/23322039.2025.2494125

Zhao, S., Li, Z., Deng, H., You, X., Tong, J., Yuan, B., et al. (2024). Spatial-temporal evolution characteristics and driving factors of carbon emission prediction in China-research on ARIMA-BP neural network algori thm. *Front. Environ. Sci.* 12, 121497941–1497941. doi:10.3389/fenvs.2024.1497941

Zhou, D., and Chu, J. (2024). The carbon emission reduction effect of the digital economy from the perspective of biased technological progress. *J. Environ. Manag.* 373, 373123857. doi:10.1016/j.jenvman.2024.123857

Zhao, W., Yin, C., Zhang, J., and Fu, B. (2024). Research progress and prospect of SDGs supported by geography-also on the theoretical framework of "sustainable geography". *J. Geogr.* 79 (11), 2699–2720.

Zhu, H., Wang, L., Li, C., Philbin, S. P., Li, H., Li, H., et al. (2023). Building a digital transformation maturity evaluation model for construction enterprises based on the analytic hierarchy process and decision-ma king trial and evaluation laboratory method. *Buildings* 14 (1), 91. doi:10.3390/buildings14010091

Zuo, Y. (2023). Research on optimization of solar integrated design of assembled buildings based on BIM technology. *J. Sol. Energy* 44 (05), 490.

frontiers | Frontiers in Built Environment

TYPE Review
PUBLISHED 17 October 2025
DOI 10.3389/fbuil.2025.1693644

Check for updates

OPEN ACCESS

EDITED BY
Zhen Chen,
University of Strathclyde, United Kingdom

REVIEWED BY
Furkan Luleci,
Louisiana State University, United States
Andrei Crisan,
Polytechnic University of Timisoara, Romania

*CORRESPONDENCE
Emilio Medrano-Sanchez,
emilio.medranos@epg.usil.pe

RECEIVED 27 August 2025
ACCEPTED 09 October 2025
PUBLISHED 17 October 2025

CITATION
Medrano-Sanchez E and Martos E (2025) Strategies for bridge maintenance using BIM: an analysis of methodologies and tools. *Front. Built Environ.* 11:1693644.
doi: 10.3389/fbuil.2025.1693644

COPYRIGHT
© 2025 Medrano-Sanchez and Martos. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Strategies for bridge maintenance using BIM: an analysis of methodologies and tools

Emilio Medrano-Sanchez[1]* and Erwin Martos[2]

[1]Universidad San Ignacio de Loyola, Lima, Peru, [2]Universidad de Piura, Piura, Peru

Ageing bridge stocks and rising traffic loads in Latin America and worldwide demand cost-effective maintenance strategies. Building Information Modelling (BIM) and its convergence with digital-twin, IoT and AI techniques have shown promise, yet their adoption for bridge upkeep remains fragmentary. This review aimed to (i) synthesise current scientific evidence on BIM-based bridge maintenance, (ii) classify methodologies and tools through a domain taxonomy, and (iii) identify research gaps that hinder large-scale implementation. A PRISMA-guided systematic literature review was conducted in Scopus and Web of Science (search cut-off = 1 February 2024). Inclusion criteria targeted peer-reviewed, open-access studies (2020–2024) that applied BIM to the maintenance of existing bridges. Twenty-five articles met the criteria and were appraised with Mixed-Methods Appraisal Tool (MMAT 2018). Seven dominant research themes were identified, with damage visualization (7 studies) and 3-D geometric modelling (6) being the most frequent, followed by information exchange/management (4). Specifically, LiDAR and photogrammetry enabled sub-centimetre models; Convolutional Neural Networks (CNN) and You Only Look Once (YOLO) algorithms reached mean average precision up to 0.91 for crack detection. Digital-twin workflows reduced operating costs while requiring higher upfront investment. A seven-domain taxonomy and a cost–technology comparison table is proposed. Key barriers reported include IFC 4.3 interoperability, high LiDAR costs (>10% of annual budgets), limited visual-programming skills, and cybersecurity concerns in cloud-IoT integrations. BIM supports preventive, data-driven bridge maintenance and has been linked to lower operating costs in several studies; mainstream adoption requires IFC 4.3 based interoperability, targeted training, and open-standard workflows. Future research should focus on standardised performance metrics, edge-AI monitoring and blockchain-secured data exchange.

KEYWORDS

bridge maintenance, 3D visualization technologies, photogrammetry and LiDAR, real-time structural monitoring, bim

1 Introduction

Today, infrastructure managers and owners, including roads, bridges, and pavements, face multiple challenges such as aging structures and their exposure to increased loads, higher speeds, and extreme weather events (Hagedorn et al., 2023). In response to these challenges, the maintenance of existing structures has been prioritized, which is essential

due to the high cost associated with building new infrastructure (Byun et al., 2021). Specifically, bridges, as critical components of road infrastructure, require specialized treatment throughout their lifecycle (Isailovic et al., 2020). During this cycle, bridges incur annual costs ranging from 0.4% to 2% of construction costs, accounting for between 16% and 80% of the total cost in the phases of operation, inspection, and eventual demolition (Artus and Koch, 2020).

Recent studies have enabled a cost-profile comparison between digital twins and traditional inspections. Digital-twin implementations demanded higher upfront expenditure, covering 3-D modelling, IoT integration and cloud-computing resources (Futai et al., 2022), whereas conventional visual-inspection approaches involved lower initial outlays (Sofia et al., 2020). Nonetheless, digital twins achieved lower operating costs by supporting predictive maintenance and reducing service interruptions (Futai et al., 2022; Ai et al., 2024); traditional practices, being largely reactive, exhibited higher operational spending due to frequent inspections (Sofia et al., 2020; Ai et al., 2024). In the long term, digital-twin adoption generated substantial savings through optimised maintenance schedules and early fault detection (Futai et al., 2022; Schneider et al., 2025; Haslbeck and Braml, 2024), while traditional methods accrued greater costs because of unexpected failures and extensive repairs (Ai et al., 2024; Hossain and Chang, 2024). Deploying a digital twin also required an advanced technological stack-laser scanners, IoT devices and cloud-computing capabilities (Futai et al., 2022; Hosamo and Hosamo, 2022), contrasting with the basic manual tools and periodic inspections typical of traditional schemes (Sofia et al., 2020). As summarised in Table 1, these cost differentials reinforce the economic rationale for adopting digital-twin strategies.

It is estimated that in the next decade, 47% of these structures will exceed 30 years of age, a figure expected to rise to 80% in the following 20 years (Jeon et al., 2023). Such projections raise serious concerns about operational safety, given the significant number of bridges that will remain in service (Li et al., 2023). In this context, the maintenance and operational safety of bridges become critically important for countries reliant on these infrastructures (Isailovic et al., 2020). A promising strategy to address these challenges is the implementation of Building Information Modeling (BIM) (Ciccone et al., 2022).

Building Information Modeling (BIM) is an advanced methodology that has been applied in architecture, engineering, and construction (Van et al., 2022). However, creating BIM models for pre-existing structures presents significant challenges due to the difficulty in accurately representing the existing structure within a model (Artus et al., 2022). To tackle these challenges, various 3D modeling techniques have been developed, such as photogrammetry (Isailovic et al., 2020; Artus and Koch, 2020; Gan et al., 2023; Wójcik and Zarski, 2021), LiDAR scanning (Lee et al., 2021), semi-automatic methods from hand-drawn and computer-assisted plans (Poku-Agyemang and Reiterer, 2023), and direct structural modeling (Poku-Agyemang and Reiterer, 2023; Kaewunru et al., 2021; Zhou, 2022). These techniques provide the necessary geometric and semantic information to facilitate proper bridge maintenance (Artus and Koch, 2020). In this context, these techniques should be understood as data-acquisition and modeling enablers that feed the information model; by themselves they do not constitute BIM.

Within this framework, Building Information Modeling (BIM) is understood as an information management process across the asset life cycle rather than as a stand-alone technology. This perspective is established by the ISO 19650 series, which defines roles, responsibilities, and workflows for information delivery and management (Abanda et al., 2025a; Park, 2024; Abanda et al., 2025b). Within this scheme, key deliverables structure the production and exchange of data, namely OIR, AIR, EIR, BEP, and the PIM and AIM models (Oliveira et al., 2024; Trani et al., 2024). In addition, the Common Data Environment (CDE) provides a central repository that ensures data integrity, accessibility, and traceability throughout the asset life cycle (Park, 2024; Xiang et al., 2025; Jaskula et al., 2024; Preidel et al., 2016; Preidel et al., 2018; Klemt-Albert et al., 2018). The Level of Information Need (LOIN), in turn, specifies the scope and granularity required at each stage, so that information remains sufficient, relevant, and usable for decision-making in bridge operation and maintenance (Trani et al., 2024; Abualdenien and Borrmann, 2022; Wogan et al., 2024). This information-management perspective underpins the assessment of maintenance needs discussed next. Determining the need for bridge maintenance is a complex process that requires detailed and meticulous analysis to conduct a proper assessment (Poku-Agyemang and Reiterer, 2023). Conventional methods, reliant on visual inspections, are subjective and error-prone due to the lack of objective data

TABLE 1 Cost and Technology Comparison—Digital Twins vs. Traditional Methods.

Aspect	Digital twins	Traditional methods
Initial costs	High (3-D modelling, IoT integration, cloud computing) (Futai et al., 2022)	Lower (manual inspections, conventional tools) (Sofia et al., 2020)
Operating costs	Lower (predictive maintenance, reduced downtime) (Futai et al., 2022; Ai et al., 2024)	Higher (reactive maintenance, frequent inspections) (Sofia et al., 2020; Ai et al., 2024)
Long-term outcome	Significant savings through optimised schedules and early-fault detection (Futai et al., 2022; Schneider et al., 2025; Haslbeck and Braml, 2024)	Higher long-term costs owing to unexpected failures and extensive repairs (Ai et al., 2024; Hossain and Chang, 2024)
Technological needs	Advanced stack (laser scanners, IoT, cloud computing) (Futai et al., 2022; Hosamo and Hosamo, 2022)	Basic tools (manual equipment, periodic inspections) (Sofia et al., 2020)

supporting maintenance decisions (Zhou, 2022). In contrast, BIM-based workflows provide a more accurate and efficient alternative for damage detection and structural maintenance (Gan et al., 2023). Despite its benefits, the application of BIM in bridge management and maintenance has been limited due to the complexity associated with these tasks (Zhou, 2022). A feasible solution is to integrate BIM with Structural Health Monitoring (SHM) systems to assess the operational performance of old structures and determine maintenance needs (Fawad et al., 2023; Sakr and Sadhu, 2023). SHM is a non-destructive system that allows real-time monitoring of structural health (Scianna et al., 2022), aiding in the performance evaluation of the bridge during its operational and service phase (Sakr and Sadhu, 2023). Recent studies have shown that SHM can be employed for real-time structural monitoring to assess maintenance needs (Sakr and Sadhu, 2023). Additionally, the integration of the Internet of Things (IoT) with BIM models enables the development of a real-time connection between the physical model and its digital twin (Scianna et al., 2022), providing an advanced technological solution for this purpose (Fawad et al., 2023). This not only delivers detailed monitoring information throughout the bridge's lifecycle (Singh and Sadhu, 2020), but also assists in the implementation of preventative maintenance strategies (Ciccone et al., 2022).

In this context, it is imperative to conduct a comprehensive review of the existing literature on the use of BIM in the maintenance of existing bridges. This review is crucial as periodic BIM-based monitoring accurately identifies structural failures (Zhou, 2022; Yang and Xia, 2023), essential for planning effective maintenance strategies that enhance infrastructure safety. Additionally, such data-driven management promotes efficient lifecycle management of the infrastructure (Hagedorn et al., 2023; Singh and Sadhu, 2020).

Finally, this article systematically examines the recent scientific literature on the use of BIM in bridge maintenance. It aims to identify methodologies, tools, and reported outcomes, enabling the determination of best practices, and addressing the challenges associated with using BIM in this field. The study seeks to answer the following research question: What are the trends, key findings, and knowledge gaps in the scientific literature on the use of BIM in existing bridge maintenance, and how can these contributions inform best practices and future strategies in this area?

2 Methodology

This study was conducted as a descriptive literature review, using scientific articles as primary information sources. The focus was on the maintenance of existing bridges using BIM. A comprehensive review was performed using the Scopus and Web of Science databases, alone were selected because recent bibliometric studies showed that the two databases together covered more than 95% of citations in civil engineering and construction science, provided standardized metadata that facilitated deduplication, and allowed trend comparisons with previous reviews on BIM for infrastructure. This ensured the quality and pertinence of the reviewed articles. In accordance with the PRISMA 2020 guidelines, an ex-post protocol was registered in the Open Science Framework, which publicly documented the objectives, search strategy, and selection criteria applied. This registration aimed to reinforce the transparency of the process without altering the corpus under analysis.

Consequently, deliberately narrow scope criteria were adopted in order to preserve thematic coherence and cross-study comparability; the asset type was limited to bridges and the life-cycle phase was restricted to operation and maintenance, with contributions exclusively focused on design or construction excluded. These decisions prioritise practical applicability for bridge O&M and account for the number of studies ultimately included. Regarding the inclusion criteria, the search was conducted considering: 1) articles that included the keywords "Building Information Modeling" and "Bridge" in the title, abstract, and keywords; 2) articles published between 2020 and 2024; 3) articles addressing the topic of bridge maintenance using BIM. On the other hand, the following exclusion criteria were considered: 1) works that did not correspond to the article format were excluded; 2) articles that were not available as open access were excluded, thus ensuring the accessibility and relevance of the selected material. The 2020–2024 time window was justified because, from 2020 onward, the IFC 4.3 bridge extension, the widespread diffusion of portable LiDAR scanners, and the integration of BIM and IoT in infrastructure maintenance reached maturity. The restriction therefore ensured that the review captured state-of-the-art technologies and avoided duplicating earlier syntheses focused on preceding periods.

This rigorous selection process identified a total of 139 articles, with 89 extracted from Scopus and 50 from Web of Science. These documents can be retrieved using the following search formulas with a cutoff date of February 1, 2024: for Scopus: (TITLE-ABS-KEY ("Building Information Modeling") AND TITLE-ABS-KEY (bridge)) AND PUBYEAR > 2019 AND (LIMIT-TO (DOCTYPE,"ar")) AND (LIMIT-TO (OA,"all")) and for Web of Science: Click here, as evidenced in Figure 1.

Following a meticulous search strategy that initially identified 139 studies, Figure 1 presents the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) Diagram, which outlines the document filtration process. From the identified documents, 40 duplicates were excluded using Power BI. Additionally, the article "A BIM technology-based underwater structure damage identification and management method" was removed due to inaccessibility, and six more were excluded for not being open access, including publications from ASCE (2 articles), Emerald Insight (3 articles), and Taylor and Francis Online (1 article). Consequently, the "Screening" phase commenced with 92 articles, ready for further rigorous evaluation.

During the "Screening" phase, 63 studies were excluded. Of these, 43 were discarded for not featuring the crucial study variables "Building Information Modeling" and "Bridge" in their titles, with 22 from Scopus and 21 from Web of Science. Furthermore, after reviewing the abstracts, 20 more articles were eliminated as their content strayed from the research focus, addressing topics such as bridge design using BIM, BIM in buildings, BIM in cultural heritage contexts, BIM for seismic behavior of bridges, BIM in railway bridges, and BIM related to thermal properties, with 11 from Scopus and 9 from Web of Science. This rigorous selection process led to a subsequent phase with 29 relevant articles.

Continuing with the selection process, a detailed reading and analysis of the full text of the remaining 29 studies were conducted. In this scrutiny, four articles were excluded for addressing themes unrelated to the objectives of this research: three focused on railway

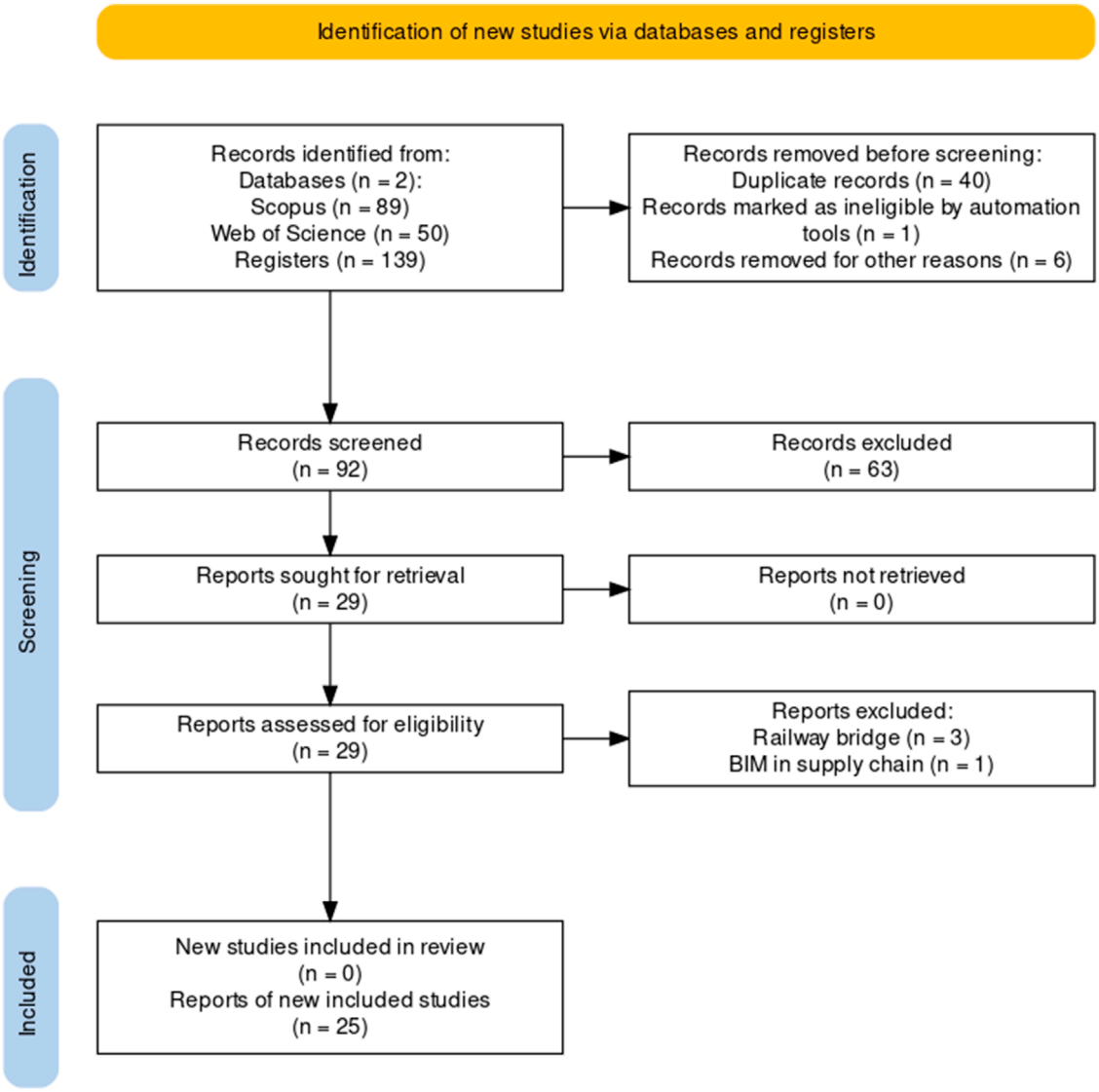


FIGURE 1
Research PRISMA diagram.

bridges and one on the implementation of BIM in the supply chain. This review refined the set to 25 articles strictly aligned with bridge maintenance using BIM, thus completing the "Screening" phase with precision and coherence with the proposed methodology.

The selected 25 articles exhibited the following geographical distribution: six from China, five from Germany, four from South Korea, two from Poland, two from Australia, one from Serbia, one from India, two from Italy, and two from Canada. In the temporally moment, 36% of these studies were from 2023, while 20% covered the years from 2020 to 2022, and 4% were from 2024. Regarding the study sample, 96% of the research focused on only one bridge, while 4% analyzed three bridges.

In the analyzed articles, various instruments were used, categorized into five main groups: The first, design and modeling tools, includes software such as Revit, Dynamo, Navisworks, and CSI Bridge, Midas CIM, RM Bridge, OpenSfM, CloudCompare, Xbim, ReviAPI, JFCOpenShel, ifcElement, Staad, Ifc Bridge, and RolAling, accounting for 46.3% of usage. The second category covers data capture tools like RGB-D cameras, Unmanned Aerial Vehicle (UAV), MPU6050, GPR, laser scanners, drones, OpenSfM, WebGL, GIS, and LiDAR, making up 19.4%. The third includes development and programming tools like Java, Arduino ACC, Excel, Matlab, WebG, SQL, C++, and Python, representing 16.4%. Finally, data management and analysis tools such as SQL and MySQL account for 13.4%, with an additional 4.5% dedicated to advanced data analysis technologies.

To effectively organize the vast amount of data collected, a specially designed matrix was used to classify the 25 selected articles into four key categories: authorship and date of study, research topic, main proposals and contributions, and the methodology used. Following this initial organization, a detailed classification of the articles by their research topics and specific

contributions was carried out, allowing for precise segmentation for future analysis.

The research topics were organized using the Excel tool, where they were appropriately categorized and grouped. For instance, the topic "Semi-automatic method to construct 3D bridge models from hand-drawn and computer-generated plans to efficiently plan bridge supervision and evaluation activities" (Poku-Agyemang and Reiterer, 2023) was classified under the category "3D geometric model of the existing structure." This classification approach ensured that all related topics were coherently grouped under this category. In total, seven distinct categories were identified, encompassing 25 specific topics Similarly, the proposals and contributions were categorized and grouped, totaling 75. These were organized based on the frequency of their appearance in the reviewed studies. When selecting proposals and contributions, priority was given to those premises and statements that enhance both theoretical and practical knowledge, essential for advancing maintenance practices of bridges using BIM.

To conclude the methodological section in an organized and accessible manner, the results were structured in two significant ways. First, a bar graph was created to clearly illustrate the distribution of the researched topics. Second, a summary table was generated detailing the frequencies of the proposals and contributions identified in the various studies analyzed. These visual tools enable a quick and effective understanding of the key data gathered during the research, which will be explained in the results section.

Finally, the methodological quality of the twenty-five studies was appraised with the Mixed-Methods Appraisal Tool (MMAT 2018); all obtained scores equal to or higher than 3 out of 5. It was nevertheless acknowledged that limiting the search to two databases and to open-access publications might have excluded relevant work published behind paywalls; Additionally, the exclusive focus on open-access articles retrieved from Scopus and Web of Science may have left out publications from publishers such as ASCE, Emerald, or Taylor and Francis, not due to deliberate exclusion but because they were not retrieved under the applied criteria. This limitation was accepted as a potential coverage bias, partly mitigated by the protocol's transparency and the availability of the extracted dataset for future replications. However, this restriction responded to the need to privilege comparability and methodological consistency over the mere expansion of document counts, so that the resulting synthesis emphasizes the robustness of the evidence rather than the absolute number of studies.

3 Results

In this chapter, the results from the previously detailed methodological application, which focused on bridge maintenance using BIM, are presented. Twenty five studies implementing various techniques and tools under this approach were exhaustively analyzed to identify emerging patterns and current trends in this research area. Two fundamental aspects are presented: first, a synthesis of the researched topics illustrated through a bar graph, and second, a compilation of the most significant proposals and contributions summarized in a detailed table. These elements form the basis for discussing the main findings, which are detailed in the following sections: the first focuses on specific research topics and the second on major contributions to bridge maintenance using BIM-based workflows.

3.1 Main research themes on bridge maintenance using BIM

The detailed analysis of the selected articles identified seven predominant research themes in the application of BIM for bridge maintenance. These themes are represented by the Figure 2, which categorizes the articles according to their thematic focus. The sequential arrangement of the topics in the figure has been carefully considered to reflect the progressive flow of maintenance processes, thus enhancing clarity, and understanding of the stages involved.

The data presentation in Figure 2 follows a logical order from left to right, which simplifies the understanding of the collected data. This arrangement is intended to facilitate an objective interpretation of the research results, without making value judgments about them. To enhance readability, a coded legend was added to identify every thematic category: T1 "3-D Geometric Model of the Existing Structure," T2 "Information Exchange and Management," T3 "Risk Inspection," T4 "Damage Identification and Detection," T5 "Damage Visualization," T6 "Real-Time Monitoring," and T7 "Bridge Rehabilitation." The x-axis also displayed the article count for each category, providing a clear quantitative reference.

The Figure 2, in its left column, features the theme "3D Geometric Model of the Existing Structure," addressed by six articles. The first of these discusses building a semi-automatic 3D model for bridges from manually drawn and computer-aided design (CAD) plans (Poku-Agyemang and Reiterer, 2023). The second study develops a 3D model in Revit from existing 2D plans using three-dimensional modeling (Kaewunru et al., 2021). These two studies demonstrate different strategies for generating 3D models: one through semi-automatic scanning of plans, and the other through modeling in Revit.

The analysis of the third article describes the creation of a master 3D digital model from existing data on suspension bridges (Dang et al., 2020). The fourth article employs Light Detection and Ranging (LiDAR) technology to model the bridge's existing structure in 3D, automatically extracting critical parameters such as height, length, and width (Lee et al., 2020). The fifth study uses a camera and artificial neural networks to capture images and segment the point cloud, refining the 3D model and removing outliers (Martens et al., 2023). Each of these contributions underscores the use of specific tools and technologies for creating accurate 3D models. Finally, the sixth article offers a panoramic view of 3D modeling methods for identifying damages in bridges, providing a comprehensive overview of the existing techniques for detecting and documenting structural deteriorations (Artus and Koch, 2020).

Continuing with the analysis in the second column of Figure 2, four studies are presented under the theme of "Information Exchange and Management." The first article proposes an integrated BIM modeling and information management platform aimed at structural monitoring and decision-making (Ciccone et al., 2022). This study explores the potential of BIM modeling to streamline data management and facilitate bridge structural monitoring. The second study uses a web platform to manage information from various

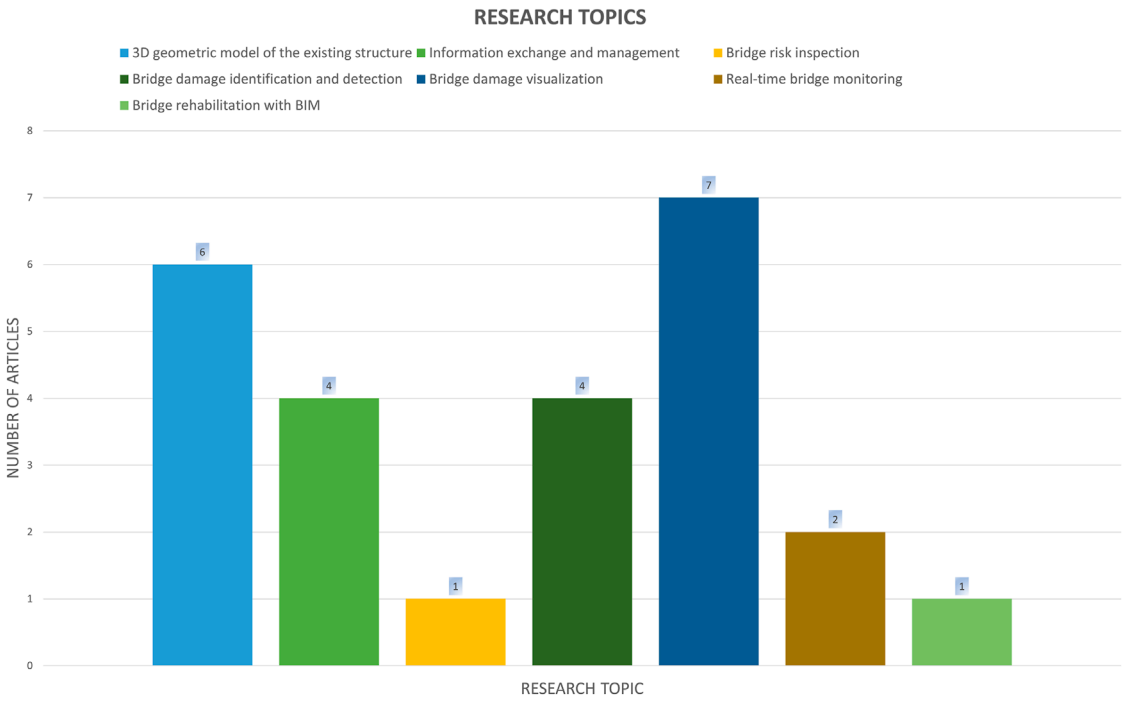


FIGURE 2
Research themes on BIM-based bridge maintenance. Legend: T1 3-D Model; T2 Information Exchange; T3 Risk Inspection; T4 Damage Detection; T5 Damage Visualization; T6 Real-Time Monitoring; T7 Rehabilitation.

bridges, optimizing damage detection and the communication of necessary repairs (Hagedorn et al., 2023). The third article describes the implementation of a Web Data Management Program (WDMP), combined with Midas CIM, to enhance efficiency in maintenance, repair, and lifespan management of the structure (Byun et al., 2021). This research establishes a management system designed to reduce maintenance costs and information loss. Lastly, the fourth study discusses the application of various BIM and Geographic Information System (GIS) technologies in managing information throughout the planning, design, construction, operation, and maintenance phases (Wei et al., 2021). This study emphasizes the need to deepen data management to optimize decision-making related to the detection of structural irregularities.

The third column of Figure 2 addresses the topic "Risk Inspection in Bridges," researched in a single study. This article explores the use of digital twins to integrate models, collect, and exchange information relevant to structural risk inspection. Through the BIM model, visualization of cracks and other signs of deterioration is facilitated, allowing for a more precise and efficient evaluation of the bridge's structural integrity (Kaewunru et al., 2021). According to the authors, the use of digital twins proves to be a powerful tool for data processing and detailed inspection, leveraging the advanced capabilities of BIM.

The fourth column of Figure 2 showcases advancements in "Identification and Damage Detection in Bridges," a central theme of four articles. The first study combines BIM and neural networks to develop a structural damage identification system (Yang and Xia, 2023). The second paper applies deep learning algorithms alongside BIM, presenting an alternative methodology to using neural networks (Gan et al., 2023). Both articles identify bridge damages, differing only in their use of BIM with neural networks or deep learning algorithms. The third article uses images and the You Only Look Once (YOLO) model to detect existing structural damage (Jeon et al., 2023), integrating computer vision technologies in infrastructure inspection. The fourth study describes how detected damages are incorporated into the BIM model in the Industry Foundation Classes (IFC) format (Isailovic et al., 2020), facilitating the update and management of the existing model.

In the fifth column of Figure 2, dedicated to "Visualizing Bridge Damage," seven articles are analyzed. The first study proposes an intelligent system for monitoring damage and cracks using three-dimensional visualization, demonstrating how BIM facilitates structural analysis (Zhou, 2022). The second article develops a maintenance system using a neural network to predict and visualize the detected structural state, highlighting the utility of predictive tools in maintenance (Shen et al., 2023). Both studies showcase the potential of modern technology to enhance the precision in identifying and visualizing structural damage.

The third article explores the integration of 3D crack models within the BIM system, enhancing their visualization and significantly improving the accuracy of infrastructure inspections (Wójcik and Zarski, 2021). Complementing this approach, the fourth study develops a three-dimensional BIM library that catalogs structural defects in bridges, allowing their incorporation into the existing BIM model for detailed visualization of defect conditions (Li et al., 2023). These investigations emphasize the effectiveness of using predefined defect models stored in a BIM library to enrich structural data visualization and management. The fifth article

focuses on visualizing and simulating bridge damage using an existing model in IFC format and processed photographs of the damaged structure, enabling detailed representation within the BIM environment (Artus et al., 2022). This approach facilitates an accurate assessment of structural integrity from images and processed data.

The sixth article focuses on representing and visualizing data obtained from real-time SHM, highlighting the importance of this technology for continuous structural integrity assessment (Singh and Sadhu, 2020). Meanwhile, the seventh article explores the integration of BIM with the IoT using an Arduino microprocessor unit, enabling the real-time tracking and visualization of bridge defects, demonstrating the benefits of combining these technologies for proactive maintenance management (Sakr and Sadhu, 2023). Both studies illustrate the capability to visualize structural defects in real time. However, a key distinction lies in the fact that while the sixth utilizes traditional SHM technologies, the seventh combines BIM with IoT through Arduino, enhancing interaction and precision in detecting structural anomalies.

In the sixth column of Figure 2, two studies under the theme "Real-time Bridge Monitoring" are discussed. The first article implements a SHM system for a bridge, using BIM-based workflows and a Common Data Environment (CDE) to automate the monitoring and analysis of the structure (Fawad et al., 2023). This research demonstrates the effectiveness of a BIM model in managing and monitoring the structural health of bridges. In contrast, the second study sets up a real-time monitoring system for a bridge beam using IoT sensors and a BIM model, allowing sensor data to be transmitted to the BIM model for continuous structural assessment (Scianna et al., 2022).

In the seventh column of Figure 2, the theme "Bridge Rehabilitation with BIM" is introduced, covered by one significant study. This article develops a BIM model specifically designed for bridge rehabilitation, demonstrating the applicability and effectiveness of these tools in structural restoration processes (Van et al., 2022). The research highlights how Staad Pro software was essential in creating BIM models that facilitate bridge rehabilitation, offering detailed and adaptive solutions to enhance existing infrastructure.

The analysis of Figure 2 showcases a variety of themes centered on the use of BIM for bridge maintenance. Damage visualization is a prominent area covered in seven articles, emphasizing the importance of precise damage detection. The 3D geometric modeling of existing structures is also extensively explored in six articles, reflecting the need to accurately represent non-digitized bridges. Additionally, topics such as information exchange and management, along with the identification of existing damages, have been addressed in four articles each. These studies lay a solid foundation for future research in monitoring, inspection, and rehabilitation of bridges using BIM, supported by advancements in 3D modeling and collaborative platforms for information management.

It is important to note that some research areas have received less attention. For instance, real-time bridge monitoring has been the subject of only two articles. Additionally, the risk inspection and rehabilitation of bridges using BIM have each been explored in only one article. These findings reveal a significant knowledge gap in the monitoring, inspection, and rehabilitation of existing bridges using BIM, highlighting the need for further research in these critical areas to enhance bridge maintenance practices. To complement the thematic analysis, Figure 3 depicts a keyword co-occurrence network generated with VOSviewer. The map revealed five main clusters: (i) a yellow cluster centred on *building information modelling* and *architectural design*, which dominated the literature and acted as the articulating hub; (ii) a blue cluster linking *point cloud*, *structural analysis* and *information management*, indicating the prominence of data capture and processing; (iii) a red cluster where *maintenance* connected with *sustainability*, *life-cycle analysis* and *information system*, reflecting the focus on maintenance strategies and lifecycle management; (iv) a green cluster associated with *engineering*, *decision making* and *bridge*, highlighting BIM-supported decision making for infrastructure; and (v) a cyan cluster grouping terms related to *bridge* and *generative design*, suggesting emerging lines in parametric optimisation. Link-density metrics confirmed the central role of BIM as an intermediary node between geometric modelling, structural analysis and maintenance, whereas the relative dispersion of the term *maintenance* underscored the need for greater convergence between conceptual research and practical applications.

As a complement to the thematic structure shown in Figure 3, and before presenting the domain synthesis, the CiteScore quartile of the journal in which each included article was published was retrieved from Scopus/SciVal. This established that the reviewed articles drew primarily on high-impact venues. Table 2 reports the distribution of the 25 articles by the quartile of their publishing journals.

The distribution confirmed a predominance of Q1 with 16 of 25 articles and Q2 with 7 of 25, so 92,% of the set came from top tiers. Q3 and Q4 contained one article each. On this basis, the subsequent thematic and methodological analysis was conducted, which strengthened the credibility of the identified trends and the practical relevance of the conclusions for maintenance management.

3.2 Domain taxonomy and key contributions

The examination of the 25 selected articles revealed seven thematic domains that encompassed the key contributions to BIM-based bridge maintenance. Consistent with the research themes identified in Table 3, these domains were: 3-D Modelling and Data Capture, Monitoring and Diagnosis, Information Management, Damage Detection, Advanced Visualisation, Parameter Quality, and Lifecycle Collaboration. Table 2 presents the resulting taxonomy and summarises each domain with a representative contribution and its frequency.

As observed in Table 3, the frequencies of the proposals and/or contributions varied between 19 and 3, reflecting the diversity of approaches present in the academic literature that was analyzed.

A particularly highlighted category was "3D Model Construction from photogrammetry, 2D plans, modeling, or laser scanning" which had a frequency of 18. This category described methods for converting an existing bridge's model into a 3D representation. Techniques included photogrammetry using Unmanned Aerial Vehicles (UAVs), LiDAR scanning, and 2D plan digitization, which generated a point cloud for creating

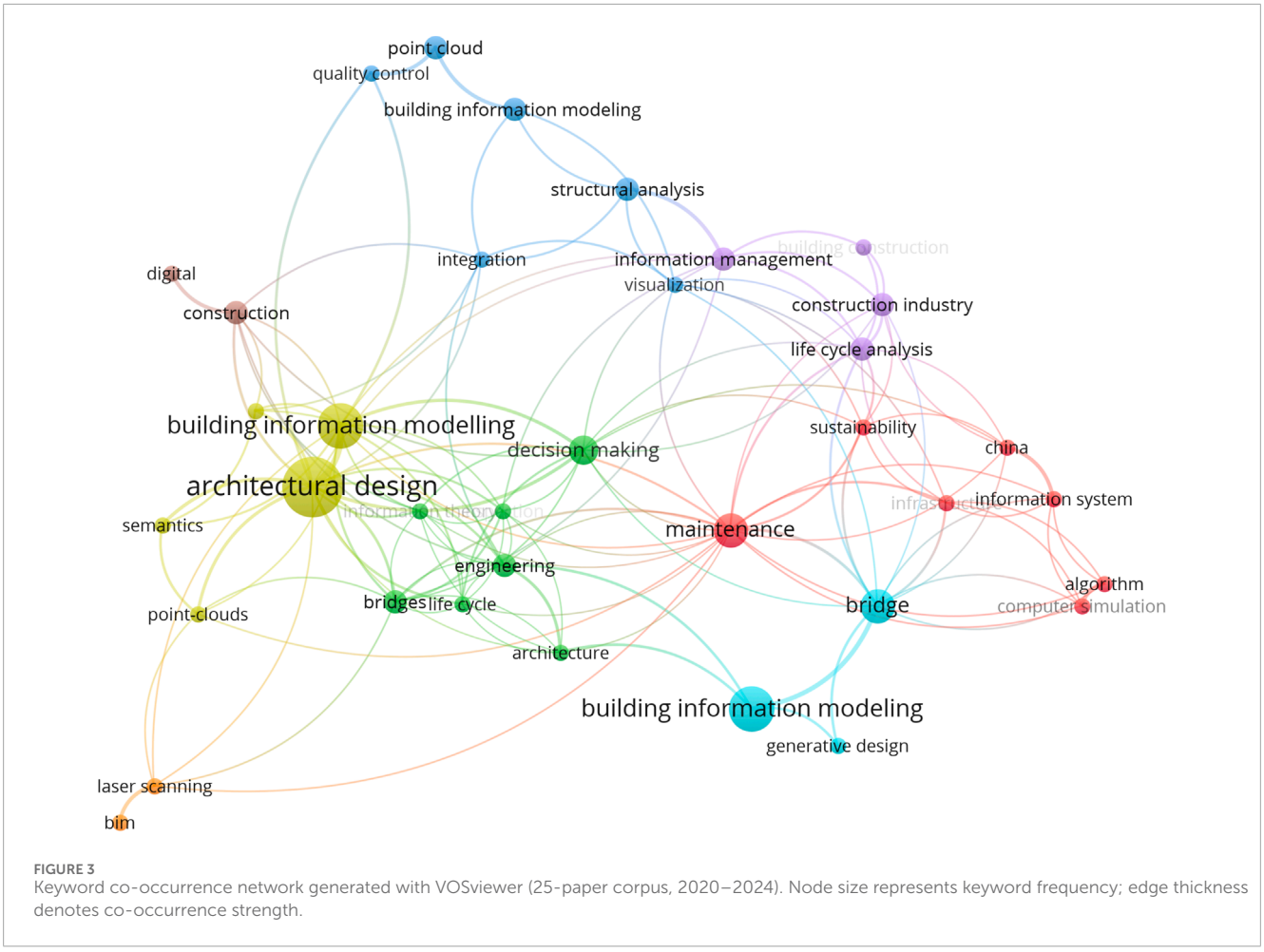


FIGURE 3
Keyword co-occurrence network generated with VOSviewer (25-paper corpus, 2020–2024). Node size represents keyword frequency; edge thickness denotes co-occurrence strength.

TABLE 2 CiteScore quartiles by year (Scopus/SciVal; snapshot: 6 August 2025; export: 18 August 2025).

Quartile	Overall n	Overall %	2020	2021	2022	2023	2024
Q1 (top 25%)	16	64%	4	2	3	7	0
Q2 (26%–50%)	7	28%	2	2	0	1	2
Q3 (51%–75%)	1	4%	0	0	0	1	0
Q4 (76%–100%)	1	4%	0	0	1	0	0
Total	25	100%	6	4	4	9	2

the 3D bridge model. Additionally, alternatives like manual 3D modeling using Dynamo scripts in Revit were optimized by creating representative families for the bridge's superstructure and substructure, which could serve as templates for future projects.

In another highlighted category, "Data Exchange According to Requirements, Processing Time, and Details," which reached a frequency of 11, a collaborative platform compatible with Ethernet and Wi-Fi was developed. This platform was designed to ensure a continuous data workflow and integrate BIM models for storage, processing, and information management. It facilitated standardized information management among project participants, thereby speeding up decision-making processes regarding bridge maintenance by managers. Additionally, this collaborative platform was integrated with BMS or ERP management systems to enhance system interconnectivity.

Among the categories with lower frequency is "Quality of the Parameters Extracted from the Model" with a frequency of three. These studies focused on evaluating the quality of parameters derived from BIM. To address this challenge, the use of LiDAR scanning, model validation through Finite Element Modeling (FEM) of bridges, and more comprehensive studies on BIM/GIS integration were recommended. These measures aim to increase the

TABLE 3 Domain taxonomy and key contributions for BIM-based bridge maintenance.

Thematic domain	Representative contribution	Frequency
3-D Modelling and Data Capture	3-D model construction via photogrammetry, LiDAR and plan digitisation	18
Monitoring and Diagnosis	Defect monitoring in 3-D models using sensors and AI (LSTM, neural networks)	19
Information Management	Data-exchange platforms compliant with IFC and BIM–GIS/ERP links	11
Damage Detection	CNN/YOLO algorithms for crack localisation and sizing	7 (+4)
Advanced Visualisation	Interactive defect visualisation in 3-D structures	4
Parameter Quality	Parameter validation through LiDAR and FEM	3
Lifecycle Collaboration	Collaborative BIM models during operation and maintenance	9

accuracy and reliability of 3D models, allowing for more precise extraction of structural parameters from bridges.

In another noteworthy category, "Increased data volume used to train the model's accuracy in crack detection" which reached a frequency of four, the use of extensive databases to train convolutional neural network (CNN) models is highlighted. This approach significantly enhances the accuracy of damage detection in bridges. Additionally, the use of the DIM extension to enrich models with additional data, such as chloride migration in bridge structures, is discussed.

Another highlighted category is "Damage Detection in Bridges Using Algorithms" with a frequency of seven. Research in this area focuses on the development of smart systems that integrate BIM and GIS technology and utilize sensors and high-resolution photography as data sources. These systems feed neural networks and the YOLO model, significantly improving the accuracy of identifying the location and extent of damage in bridges.

Within the category "Visualization of Defects in 3D Structures" which has a frequency of four, efforts to visually represent damage and cracks in three dimensions in critical areas of bridges are highlighted. This method not only enhances the understanding of structural defects but also integrates vital information for making strategic decisions about bridge rehabilitation.

The most prominent category according to Table 3 is "Monitoring of Defects in Structures through 3D Models" which has a frequency of 19. It focuses on using sensors installed on bridges to collect data periodically. This data is processed using advanced techniques such as neural networks, deep learning, and Long Short-Term Memory (LSTM). Identified defects are integrated into 3D BIM models to simulate structural behavior using tools like Stadd Pro software. This system aims to provide real-time analytical capabilities and is augmented by augmented reality to optimize decision-making in bridge maintenance, thus minimizing potential damage and deterioration.

Another category, with a frequency of nine, focuses on" Collaborative BIM Model Throughout the Project Lifecycle." These studies emphasize the importance of BIM models throughout the entire construction project lifecycle, especially during the operation and maintenance phases of bridges. BIM tools facilitate collaboration and interconnection among various project stakeholders, thereby improving decision-making in maintenance.

In conclusion, Table 3 highlights key areas in bridge maintenance using BIM. The construction of 3D models and the monitoring of defects are the most researched topics, reflecting the critical need for detailed models and defect analysis for effective management. Other important areas include data exchange and the collaborative use of BIM, which are essential for optimizing decision-making and maintenance. Less frequent studies, such as improving the accuracy of crack detection and evaluating the quality of parameters, are also crucial for advancing BIM practices.

4 Discussion

In this study, the following stages are outlined for discussing the results: 1) 3D Geometric Modeling of the existing structure, 2) Identification and management of defects in bridges using BIM, 3) Visualization and monitoring of bridge damages through BIM, and 4) Bridge rehabilitation via BIM. Subsequently, the findings from the research related to each of these stages are explored in detail.

4.1 3D geometric model of the existing structure

The 3D BIM model of the existing structure is crucial for efficient data exchange, incorporating appropriate geometric and semantic information (Artus and Koch, 2020; Martens et al., 2023). To achieve this, a detailed Scan-to-BIM process is essential, which involves collecting the necessary data to develop the BIM model (Lee et al., 2020). This process is organized into four fundamental stages: data capture, semantic segmentation, 3D modeling, and BIM model generation, each critical for the project's success (Martens et al., 2023).

In the data capture stage for 3D modeling of existing structures within a BIM environment, various techniques such as photogrammetry (Isailovic et al., 2020; Artus and Koch, 2020; Gan et al., 2023; Wójcik and Zarski, 2021) and LiDAR scanning (Lee et al., 2020). This approach demanded a balance between

resolution and field efficiency: point densities above 50 pts/m^2 increased sub-centimetre accuracy, yet doubled acquisition time and raised over-noise in heavily vegetated areas, thereby requiring additional filtering stages. Although effective, these methods can be costly and limited in access (Poku-Agyemang and Reiterer, 2023). To overcome these barriers, a semi-automatic method using hand-drawn or computer-generated plans has been developed to construct the 3D model (Poku-Agyemang and Reiterer, 2023). Alternatively, modeling in Revit from 2D drawings provides precise geometric parameters such as height, length, and width (Poku-Agyemang and Reiterer, 2023; Kaewunru et al., 2021; Zhou, 2022). This phase results in a point cloud that provides detailed information on the structural shape, making it crucial to assign semantic information to each point for subsequent use in the BIM model (Lee et al., 2021; Lee et al., 2020).

In the semantic segmentation stage, neural networks and deep learning techniques are used to assign semantic information to each point in the point cloud, simplifying the process without extensive training and with minimal data (Martens et al., 2023; Shen et al., 2023). Although deep learning is effective (Lee et al., 2021), semantic information alone is insufficient to create a complete BIM model; accurate structural parameters such as height, length, and width are required during the BIM scanning stage (Artus and Koch, 2020; Lee et al., 2021; Lee et al., 2020). To streamline this process, automatic methods have been developed to reduce noise, perform 3D transformations, and extract parameters, thereby reducing the time and resources needed to achieve a detailed BIM model (Isailovic et al., 2020; Lee et al., 2020).

4.2 Bridge defect identification and management with BIM

Bridge defects can be classified as structural, functional, and durability related (Artus and Koch, 2020). Traditionally, inspectors manually collected defect data, a method that is subjective and error-prone, complicating accurate measurement of damage characteristics like length, depth, and width (Ciccone et al., 2022; Gan et al., 2023). In response, photogrammetry produces point clouds, orthomosaics, and measurements that, when integrated into the information model within BIM-based workflows, enable objective and traceable damage detection and quantification for maintenance management (Isailovic et al., 2020; Jeon et al., 2023; Artus et al., 2022; Gan et al., 2023; Yang and Xia, 2023). Absolute accuracy hinged on proper georeferencing with ground-control points; a root-mean-square error of ±14 mm was recorded when fewer than five linearly distributed points were used, compared with ±6 mm when at least eight concentric points were deployed.

To precisely locate damage on bridge sections, intelligent algorithms based on neural networks are employed (Kaewunru et al., 2021; Yang and Xia, 2023). Beyond locating damage, it's crucial to determine its magnitude, a task efficiently performed by the R-CNN method (Artus et al., 2022; Gan et al., 2023). Alternatively, the YOLO algorithm is noted for its damage detection capabilities (Jeon et al., 2023). Model performance was constrained by the scarcity of bridge-specific training datasets; a 9-percentage-point drop in mAP was observed when the model was applied to carbonation-degraded concrete, highlighting the need for domain-adaptation strategies prior to transfer across pathologies. The data on the location and extent of cracks is integrated into the BIM model of the bridge for accurate structural assessment. Additionally, predictive crack models are generated using the LSTM method, enhancing the anticipation of structural issues (Shen et al., 2023). To manage this extensive information, collaborative platforms are essential to facilitate storage, organization, and access to data by all project stakeholders (Ciccone et al., 2022).

A digital platform optimizes the management of defects and the BIM model of the existing structure, facilitating comprehensive information handling that allows access for all project participants (Hagedorn et al., 2023; Jeon et al., 2023). These platforms support management and maintenance, enhance bridge condition assessments, and foster collaboration through interactive visualization (Wei et al., 2021). Additionally, another platform integrates BIM modeling and information management, enriched with safety diagnostics and maintenance plans (Byun et al., 2021; Ciccone et al., 2022). Finally, creating libraries that store bridge defect information streamlines the integration of these data into BIM models (Li et al., 2023).

4.3 Visualization and monitoring of bridge damages with BIM

Once the geometric and semantic information is integrated into the BIM model on the platform, it facilitates detailed visualization of defects, showing hierarchy, properties, and geometry (Artus et al., 2022; Shen et al., 2023). Before real-time monitoring, it is crucial to validate the model and maintenance system using Finite Element Analysis (FEA), applying static and dynamic loads to the bridge (Fawad et al., 2023). After validation, sensors are placed at critical points on the bridge to capture continuous data (Singh and Sadhu, 2020), while the use of IoT with Arduino allows for real-time monitoring and visualization of these data (Sakr and Sadhu, 2023). This methodology was applied to monitor the deflection of the bridge beam, contributing to the creation of a digital twin that continuously monitors the structure and detects any anomalies (Scianna et al., 2022; Kaewunruen et al., 2020). The literature likewise reported operational challenges: sensor-node energy autonomy ranged between 4 and 7 months under 2.4 GHz transmission, whereas network congestion imposed practical limits of 60 Hz on continuous sampling rates to prevent packet loss.

4.4 Bridge rehabilitation with BIM

The availability of a digital twin that allows continuous real-time monitoring lays the groundwork for moving towards effective bridge rehabilitation (Kaewunruen et al., 2020). Using Stadd Pro software, it is possible to simulate the structural behavior of the bridge and apply advanced rehabilitation techniques, optimizing intervention plans to restore and enhance the infrastructure (Van et al., 2022).

Despite the highlighted advantages, the adoption of BIM for bridge maintenance faced substantial limitations that constrained its operational feasibility. Interoperability issues remained, stemming from the coexistence of proprietary formats with the IFC 4.3 standard and forcing conversion workflows that introduced

geometric and semantic data loss. In addition, the initial cost of LiDAR scanners and 3-D modelling exceeded, in several case studies, 10% of the annual maintenance budget, thus lowering the incentive for agencies operating under financial constraints. The demand for specialised personnel, proficient in visual programming (e.g., Dynamo) and large-scale data management, also persisted, and such expertise was still scarce within public infrastructure agencies. Finally, the reviewed studies revealed gaps in governance and cybersecurity, particularly when BIM models were synchronised with IoT sensors in cloud environments, raising data-protection and continuous-monitoring requirements.

The documented experiences suggested that these obstacles could be mitigated through gradual strategies: workflows based on open standards, modular staff training, and cost-benefit analyses demonstrating medium-term positive returns. Nevertheless, the evidence indicated that a comprehensive resolution of these barriers remained an active research area, especially regarding regulatory frameworks capable of harmonising information management throughout the bridge life cycle.

The review further indicated that the incorporation of emerging technologies began to extend the conventional boundaries of BIM. The most recent studies investigated the use of digital twins to replicate the real-time condition of bridges, enabling scenario simulations of deterioration and more precise prioritisation of interventions. Initiatives integrating blockchain as an immutable ledger for lifecycle data were also identified, thereby strengthening trust among stakeholders. Finally, pilot applications of edge computing processed IoT sensor data on site, reducing latency and mitigating cloud congestion risks. Although these innovations remained in early adoption stages, the evidence suggested that their convergence with BIM could accelerate the shift toward predictive and collaborative maintenance models.

5 Conclusion

This study has cataloged various methodologies and tools from scientific literature for bridge maintenance using BIM. It emphasizes the significance of 3D geometric modeling, defect identification, and damage visualization and monitoring as vital practices.

Likewise, trends toward automation in monitoring and real-time structural assessment are evident, reflecting significant advancements in BIM applications. Nevertheless, critical gaps still constrained large-scale implementation. Chief among them was the absence of harmonised regulatory frameworks for transferring models across platforms, together with a need for official guidelines governing the integration of digital twins and IoT sensors into existing infrastructure. The evidence also revealed shortages in training programmes aimed at strengthening visual-programming and big-data skills among bridge managers. Addressing these regulatory and capacity gaps will be essential to consolidate predictive and collaborative maintenance strategies.

However, knowledge gaps are also identified, particularly in real-time data integration and collaborative information management throughout project lifecycles. These findings underscore the need for further research in specific underexplored areas, particularly real-time bridge monitoring (2 studies), risk inspection (1), and rehabilitation with BIM (1). It is also recommended to develop more robust collaborative platforms that effectively integrate BIM with emerging technologies like IoT, enhancing accuracy in structural damage detection and management.

While these findings synthesise global trends, they gain greater value when translated into policy guidelines adapted to regional contexts. In Latin America, this involves adopting IFC 4.3-based exchange profiles to reduce information loss across platforms, mitigating the upfront costs of technologies such as LiDAR, which can exceed ten percent of annual maintenance budgets, strengthening the capacities of public teams in visual programming and data management, establishing minimum cybersecurity requirements for BIM workflows connected to IoT in the cloud, and promoting standardised metrics for project evaluation. Linking global evidence with these local priorities enables predictive and collaborative workflows to be effectively integrated into bridge management.

6 Research agenda and future directions

Building on the synthesis (Table 3) and the keyword network (Figure 3), this review identified actionable gaps and outlines how future work should address them.

1. Standardized performance metrics. Future studies should define interoperable metrics to compare BIM/digital–twin workflows across contexts: end-to-end latency (ms), detection accuracy (mAP/F1), cost per inspected square metre, downtime hours avoided, and ROI over 5–10 years. Multi-site reporting templates should be shared as open protocols.
2. Public benchmark datasets for bridges. The community should curate open, bridge-specific datasets (images, point clouds, annotated defects) with clear licenses and train/validation/test splits to enable fair comparisons and domain-adaptation studies of CNN/YOLO models.
3. Operational IFC 4.3 for maintenance. Research should prototype and validate maintenance-phase information schemas (attributes, property sets, exchange requirements) and run round-trip tests to quantify semantic loss across authoring and asset-management platforms.
4. Cost–benefit models for public agencies. Future work should develop decision-oriented economic models that estimate total cost of ownership and net benefits of BIM/DT adoption under budget constraints, using before–after or matched-control evaluations on real bridge portfolios.
5. Edge-AI reliability for IoT-to-BIM streaming. Experiments should characterize accuracy–latency–energy trade-offs (e.g., mWh/day, packet-loss %) for on-device inference and propose fallback strategies under network outages and cybersecurity policies.
6. Adoption in Latin America. Empirical studies should investigate organizational readiness, governance and regulatory pathways, identifying enablers and barriers for mainstream deployment, with replicable instruments (surveys, maturity indices) and policy recommendations.

Addressing these lines of research would align technical advances with procurement, budgeting and safety decisions in real bridge-maintenance programs.

Author contributions

EM-S: Conceptualization, Formal Analysis, Investigation, Methodology, Resources, Software, Supervision, Validation, Visualization, Writing – review and editing. EM: Data curation, Formal Analysis, Investigation, Software, Writing – original draft.

Funding

The author(s) declare that no financial support was received for the research and/or publication of this article.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fbuil.2025.1693644/full#supplementary-material

References

Abanda, F. H., Balu, B., Adukpo, S. E., and Akintola, A. (2025a). Decoding ISO 19650 through process modelling for information management and stakeholder communication in BIM. *Buildings* 15, 431. doi:10.3390/buildings15030431

Abanda, F. H., Akintola, A., Tuhaise, V. V., and Tah, J. H. M. (2025b). BIM ontology for information management (BIM-OIM). *J. Build. Eng.* 107, 112762. doi:10.1016/j.jobe.2025.112762

Abualdenien, J., and Borrmann, A. (2022). Levels of detail, development, definition, and information need: a critical literature review. *J. Inf. Technol. Constr.* 27, 363–392. doi:10.36680/j.itcon.2022.018

Ai, L., C, L. K., Elbatanouny, E., Henderson, A., and Ziehl, P. (2024). Assisting load rating testing of precast reinforced concrete bridge slab through digital twins and field monitoring data. *J. Test. Eval.* 52, 1926–1938. doi:10.1520/JTE20230748

Artus, M., and Koch, C. (2020). State of the art in damage information modeling for rc bridges – a literature review. *Adv. Eng. Inf.* 46, 101171. doi:10.1016/j.aei.2020.101171

Artus, M., Alabassy, M. S. H., and Koch, C. (2022). A bim based framework for damage segmentation, modeling, and visualization using ifc. *Appl. Sci.* 12, 2772. doi:10.3390/app12062772

Byun, N., Han, W. S., Kwon, Y. W., and Kang, Y. J. (2021). Development of bim-based bridge maintenance system considering maintenance data schema and information system. *Sustain.* 13, 4858. doi:10.3390/su13094858

Ciccone, A., Suglia, P., Asprone, D., Salzano, A., and Nicolella, M. (2022). Defining a digital strategy in a bim environment to manage existing reinforced concrete bridges in the context of Italian regulation. *Sustain.* 14, 11767. doi:10.3390/su141811767

Dang, N. S., Rho, G. T., and Shim, C. S. (2020). A master digital model for suspension bridges. *Appl. Sci.* 10, 7666–22. doi:10.3390/app10217666

Fawad, M., Salamak, M., Poprawa, G., Koris, K., Jasinski, M., Lazinski, P., et al. (2023). Automation of structural health monitoring (shm) system of a bridge using bimification approach and bim-based finite element model development. *Sci. Rep.* 13, 13215. doi:10.1038/s41598-023-40355-7

Futai, M. M., Bittencourt, T. N., Santos, R. R., Araújo, C. R. R., Ribeiro, D. M., Rocha, A. R., et al. (2022). Utilization of digital twins for bridge inspection, monitoring and maintenance. In: Proceedings of the 1st conference of the european association on quality control of bridges and structures. Cham, Switzerland: Springer. p. 166–173. doi:10.1007/978-3-030-91877-4_20

Gan, L., Liu, H., Yan, Y., and Chen, A. (2023). Bridge bottom crack detection and modeling based on faster r-cnn and bim. *IET Image Process.* 18, 664–677. doi:10.1049/ipr2.12976

Hagedorn, P., Liu, L., Koenig, M., Hajdin, R., Blumenfeld, T., Stoeckner, M., et al. (2023). Bim-enabled infrastructure asset management using information containers and semantic web. *J. Comput. Civ. Eng.* 37, 04022041. doi:10.1061/(ASCE)CP.1943-5487.0001051

Haslbeck, M., and Braml, T. (2024). The role of uncertainty propagation for digital twins. In: *20th International Probabilistic Workshop*. Cham, Switzerland: Springer. p. 303–312. doi:10.1007/978-3-031-60271-9_28

Hosamo, H. H., and Hosamo, M. H. (2022). Digital twin technology for bridge maintenance using 3d laser scanning: a review. *Adv. Civ. Eng.* 2022, 2194949. doi:10.1155/2022/2194949

Hossain, A., and Chang, C. M. (2024). Modeling life expectancy and cost effectiveness for uhpc bridge retrofitting techniques. *Adv. Bridge Eng.* 5, 9. doi:10.1186/s43251-024-00120-w

Isailovic, D., Stojanovic, V., Trapp, M., Richter, R., Hajdin, R., and Doellner, J. (2020). Bridge damage: detection, ifc-based semantic enrichment and visualization. *Automation Constr.* 112, 103088. doi:10.1016/j.autcon.2020.103088

Jaskula, K., Kifokeris, D., Papadonikolaki, E., and Rovas, D. (2024). Common data environments in construction: state-of-the-art and challenges for practical implementation. In: *Construction Innovation: Information Process Management*. Leeds, England: Emerald Publishing Limited. doi:10.1108/CI-04-2023-0088

Jeon, C. H., Nguyen, D. C., Roh, G., and Shim, C. S. (2023). Development of brim-based bridge maintenance system for existing bridges. *Buildings* 13, 2332. doi:10.3390/buildings13092332

Kaewunruen, S., Sresakoolchai, J., Ma, W., and Phil-Ebosie, O. (2021). Digital twin aided vulnerability assessment and risk-based maintenance planning of bridge infrastructures exposed to extreme conditions. *Sustain.* 13, 2051–19. doi:10.3390/su13042051

Kaewunruen, S., Sresakoolchai, J., and Zhou, Z. (2020). Sustainability-based lifecycle management for bridge infrastructure using 6d bim. *Sustain.* 12, 2436. doi:10.3390/su12062436

Klemt-Albert, K., Hagedorn, P., and Pullmann, T. (2018). Utilising the potential of standardised bim models by a fundamental transformation of collaboration processes. In: 25th workshop of the European group for computing in engineering; 2018 June 10–13; Lausanne, Switzerland. Cham, Switzerland: Springer. p. 470–486. doi:10.1007/978-3-319-91638-5_26

Lee, J. H., Park, J. J., and Yoon, H. (2020). Automatic bridge design parameter extraction for scan-to-bim. *Appl. Sci.* 10, 7346–17. doi:10.3390/app10207346

Lee, J. S., Park, J., and Ryu, Y. M. (2021). Semantic segmentation of bridge components based on hierarchical point cloud model. *Automation Constr.* 130, 103847. doi:10.1016/j.autcon.2021.103847

Li, S., Zhang, Z., Lin, D., Zhang, T., and Han, L. (2023). Development of a bim-based bridge maintenance system (bms) for managing defect data. *Sci. Rep.* 13, 846. doi:10.1038/s41598-023-27924-6

Martens, J., Blut, T., and Blankenbach, J. (2023). Cross domain matching for semantic point cloud segmentation based on image segmentation and geometric reasoning. *Adv. Eng. Inf.* 57, 102076. doi:10.1016/j.aei.2023.102076

Oliveira, A., Granja, J., Bolpagni, M., Motamedi, A., and Azenha, M. (2024). Development of standard-based information requirements for the facility management of a canteen. *J. Inf. Technol. Constr.* 29, 281–307. doi:10.36680/j.itcon.2024.014

Park, J. (2024). Framework for managing multiple common data environments. In: Construction Research Congress 2024; 2024 March 20–23; Des Moines, Iowa. Reston, VA: ASCE. p. 1117–1127. doi:10.1061/9780784485262.114

Poku-Agyemang, K. N., and Reiterer, A. (2023). 3d reconstruction from 2d plans exemplified by bridge structures. *Remote Sens.* 15, 677. doi:10.3390/rs15030677

Preidel, C., Borrmann, A., Oberender, C., and Tretheway, M. (2016). Seamless integration of common data environment access into bim authoring applications: the bim integration framework In: 11th European Conference on Product and Process Modelling; 2016 September 7–9: CRC Press/Balkema. p. 119–128.

Preidel, C., Borrmann, A., Mattern, H., König, M., and Schapke, S. E. (2018). *Common data environment*. Cham, Switzerland: Springer International Publishing. p. 279–291. doi:10.1007/978-3-319-92862-3_15

Sakr, M., and Sadhu, A. (2023). Visualization of structural health monitoring information using internet-of-things integrated with building information modeling. *J. Infrastructure Intell. Resil.* 2, 100053. doi:10.1016/j.iintel.2023.100053

Schneider, R., Simon, P., Herrmann, R., Hille, F., and Baeßler, M. (2025). Bestandteile digitaler zwillinge im erhaltungsmanagement von verkehrsbrücken. *Bautechnik* 102, 355–365. doi:10.1002/bate.202400101

Scianna, A., Gaglio, G. F., and Guardia, M. L. (2022). Structure monitoring with bim and iot: the case study of a bridge beam model. *ISPRS Int. J. Geo Inform* 11, 173. doi:10.3390/ijgi11030173

Shen, C., Jiang, B., and Yue, L. (2023). Lstm combined with bim technology in the management of small and medium-sized span highway concrete beam bridges. *Results Eng.* 20, 101539. doi:10.1016/j.rineng.2023.101539

Singh, P., and Sadhu, A. (2020). System identification-enhanced visualization tool for infrastructure monitoring and maintenance. *Front. Built Environ.* 6, 76. doi:10.3389/fbuil.2020.00076

Sofia, H., Anas, E., and Faiz, O. (2020). Mobile mapping, machine learning and digital twin for road infrastructure monitoring and maintenance: case study of mohammed vi bridge in morocco. In: 2020 IEEE international conference of Moroccan geomatics (Morgeo); 2020 May 11-13: IEEE. p. 1–6. doi:10.1109/Morgeo49228.2020.9121882

Trani, M. L., Madaschi, F., and Predieri, G. (2024). Construction site information modelling applied to infrabim. In: 5th European and Mediterranean Structural Engineering and Construction Conference; 2024 May 13–18: ISEC press. doi:10.14455/ISEC.2024.11(2).CON-30

Vanjari, P. B., and Kulkarni, S. S. (2022). Building information modeling (bim) as tool to develop solution for bridge rehabilitation. *ASEAN J. Sci. Eng.* 2, 77–90. doi:10.17509/ajse.v2i1.37715

Wei, J., Chen, G., Huang, J., Xu, L., Yang, Y., Wang, J., et al. (2021). Bim and gis applications in bridge projects: a critical review. *Appl. Sci.* 11, 6207. doi:10.3390/app11136207

Wogan, M., Mohan, N., Wolf, G., Nejat, N., Menzel, K., and Gross, R. (2024). Synergising lifecycle project management for sustainability: towards a streamlined approach through different project phases. In: 9th International Conference on Civil, Structural and Transportation Engineering, ICCSTE 2024; 2024 June 13–15; Toronto, Canada. Orléans, ON: Avestia Publishing. doi:10.11159/iccste24.157

Wójcik, B., and Zarski, M. (2021). The measurements of surface defect area with an rgb-d camera for a bim-backed bridge inspection. *Bull. Pol. Acad. Sci. Tech. Sci.* 69. doi:10.24425/bpasts.2021.137123

Xiang, L., Li, G., and Li, H. (2025). A common data environment framework applied to structural life cycle assessment: coordinating multiple sources of information. *Buildings* 15, 1315. doi:10.3390/buildings15081315

Yang, H., and Xia, M. (2023). Advancing bridge construction monitoring: Ai-based building information modeling for intelligent structural damage recognition. *Appl. Artif. Intell.* 37, 2224995. doi:10.1080/08839514.2023.2224995

Zhou, Z. (2022). An intelligent bridge management and maintenance model using bim technology. In: *Mobile information systems*. Hoboken, NJ: Wiley. doi:10.1155/2022/7130546

frontiers | Frontiers in Virtual Reality

TYPE Systematic Review
PUBLISHED 18 December 2025
DOI 10.3389/frvir.2025.1725875

Check for updates

OPEN ACCESS

EDITED BY
Salman Azhar,
Auburn University, United States

REVIEWED BY
Rana Muhammad Irfan Anwar,
Auburn University, United States
Muhammad Umer,
University of West Florida, United States

*CORRESPONDENCE
Diego Sornoza-Parrales,
diego.sornoza@unesum.edu.ec

RECEIVED 15 October 2025
REVISED 23 November 2025
ACCEPTED 01 December 2025
PUBLISHED 18 December 2025

CITATION
Sornoza-Parrales D, Baque Vera GO, Parrales Poveda ML and Macías-Parrales TM (2025) Evolution and research trends in virtual reality and augmented reality technologies for the architecture, engineering, and construction (AEC) industry: a systematic review and science mapping approach.
Front. Virtual Real. 6:1725875.
doi: 10.3389/frvir.2025.1725875

COPYRIGHT
© 2025 Sornoza-Parrales, Baque Vera, Parrales Poveda and Macías-Parrales. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Evolution and research trends in virtual reality and augmented reality technologies for the architecture, engineering, and construction (AEC) industry: a systematic review and science mapping approach

Diego Sornoza-Parrales[1,2]*, Günther Oberdam Baque Vera[3], María Leonor Parrales Poveda[2,4] and Tania Maricela Macías-Parrales[3,4]

[1]Faculty of Technical Sciences, Universidad Estatal del Sur de Manabí, Jipijapa, Ecuador, [2]Institute of Postgraduate Studies, Universidad Estatal del Sur de Manabí, Jipijapa, Ecuador, [3]Admission and Leveling Unit, Universidad Estatal del Sur de Manabí, Jipijapa, Ecuador, [4]Faculty of Economic Sciences, Universidad Estatal del Sur de Manabí, Jipijapa, Ecuador

This systematic review and bibliometric analysis examine the evolution and research trends of Virtual Reality (VR) and Augmented Reality (AR) technologies within the Architecture, Engineering, and Construction (AEC) industry to identify key research themes, application domains, and future directions. A comprehensive literature search was conducted using both Scopus and Web of Science databases covering publications from 2015 to 2025 to ensure robust multi-database validation. After applying engineering subject area filters, the Scopus search yielded 1,301 articles while Web of Science returned 898 articles. The methodology combined quantitative bibliometric analysis using Bibliometrix software with qualitative thematic analysis of the most-cited and relevant studies following PRISMA guidelines. Performance analysis, science mapping, and co-occurrence analysis were employed to identify research clusters and trends across both databases. Cross-database validation confirmed convergent patterns in publication growth, thematic structure, and collaboration networks, with both databases demonstrating exponential growth and similar dominant research themes centered on digital twins, virtual reality applications, and BIM integration. This study provides a comprehensive quantitative mapping of VR/AR research in the AEC industry with multi-database validation, revealing the intellectual structure and research fronts. The integrated approach combining bibliometric analysis with thematic synthesis offers insights for researchers and practitioners to understand current applications, identify research gaps, and guide future technology adoption strategies.

KEYWORDS

virtual reality, augmented reality, construction industry, research trends, digital twin, science mapping, bibliometric analysis, building information model (BIM)

1 Introduction

1.1 Digital transformation in the AEC industry

The Architecture, Engineering, and Construction (AEC) industry is undergoing a fundamental digital transformation aligned with Industry 4.0, characterized by the integration of Cyber-Physical Systems, Internet of Things (IoT), and cloud computing technologies (Hossain and Nadeem, 2019). This transformation aims to enhance productivity, improve collaboration, and address evolving construction demands through comprehensive digital solutions (Bou Hatoum et al., 2022).

Building Information Modeling (BIM) has emerged as a cornerstone technology, fundamentally redefining information management across project lifecycles and enabling efficient design, planning, execution, and maintenance processes (Azhar, 2011; Mannino et al., 2021). The integration of digital twin technologies further augments these capabilities by providing real-time data visualization and analysis, facilitating improved decision-making throughout building lifespans (Sepasgozar, 2020). This technological shift enhances project management while improving safety standards and operational efficiencies in construction environments (Sidani, Poças Martins and Soeiro, 2022).

Construction 4.0 embodies Industry 4.0 principles, illustrating how digital and physical realm convergence drives innovation through advanced technologies including robotics, artificial intelligence, and automated systems integrated into construction workflows (Klarin and Xiao, 2024). Organizations seeking modernization must navigate these digital transformation complexities to foster industry resilience and sustainability (Hossain and Nadeem, 2019; Bou Hatoum et al., 2022).

1.2 Immersive technologies landscape

Immersive technologies exist along the reality-virtuality continuum, a conceptual framework introduced by Milgram and Kishino (1994) that spans from completely real environments to fully virtual ones. Extended Reality (XR) serves as an umbrella term encompassing Virtual Reality (VR), Augmented Reality (AR), and Mixed Reality (MR). VR provides fully immersive digital environments isolated from the physical world; AR overlays digital content onto real-world views; and MR enables interaction between physical and digital objects in real-time. These technologies have emerged as transformative tools offering innovative solutions for visualization, design collaboration, and training within the AEC sector, addressing complex challenges inherent in modern construction practices through enhanced spatial understanding and interactive experiences. While XR represents the complete spectrum of immersive technologies, VR and AR have achieved greater technological maturity and documented implementation within AEC contexts, forming the empirical focus of this review.

VR provides fully immersive environments enabling users to interact with digital representations of physical spaces, facilitating intricate detailed visualization before physical execution and supporting informed decision-making early in project lifecycles (Afzal et al., 2021). This technology enhances stakeholder communication and engagement by providing shared virtual spaces for real-time modifications and assessments, potentially reducing costly later-phase changes (Jacobsen et al., 2022).

AR complements VR by superimposing digital information onto real-world environments, proving particularly advantageous during construction phases through real-time project overlays that enhance work accuracy (Xu and Moreu, 2021). Multiple studies have highlighted AR's potential for construction safety enhancement by providing critical information in hazardous settings, enabling workers to visualize safety instructions within their operational context (Adami et al., 2022; Jacobsen et al., 2022).

Mixed Reality applications, while less prevalent in current AEC literature, demonstrate promising results in specific construction tasks. For instance, MR has been successfully applied to electrical construction design communication, showing higher productivity rates and fewer assembly errors compared to traditional methods (Chalhoub and Ayer, 2018). However, MR implementation remains limited due to higher hardware requirements and technological complexity compared to VR and AR solutions.

The integration of VR and AR with BIM significantly streamlines AEC workflows, enabling interactive BIM data visualization in immersive formats that enhance project collaboration and optimize workflow management (Chi et al., 2022; Li et al., 2024). Recent literature indicates expanding applications in civil engineering structure condition assessment, presenting paradigm shifts that enable virtual inspections and simulations previously requiring extensive time and resources (Catbas et al., 2022).

1.3 Research gaps and systematic review rationale

The rapid expansion of VR applications in civil engineering necessitates a comprehensive quantitative review to effectively map the growing knowledge base and elucidate emerging research trends. As an evolving field, VR has gained recognition for its potential to transform various civil engineering and construction aspects, from design processes to training and safety practices (Huang and Roscoe, 2021).

Bibliometric analysis provides systematic empirical methods utilizing statistical tools to quantify academic literature, offering insights into publication patterns, co-authorship networks, keyword trends, and citation dynamics (Agbo et al., 2021). Such analyses have been effectively employed across various domains to highlight trends and research gaps (Rashid et al., 2021). The noteworthy expansion of VR-related publications, coupled with increasing interdisciplinary collaboration, underscores the necessity of bibliometric approaches to efficiently synthesize this information wealth.

Current literature lacks comprehensive systematic reviews combining quantitative bibliometric analysis with qualitative thematic synthesis specifically focusing on VR and AR applications in the AEC industry context. Existing reviews tend to focus on narrow application areas or lack the temporal scope necessary to capture the recent technological advances and integration patterns.

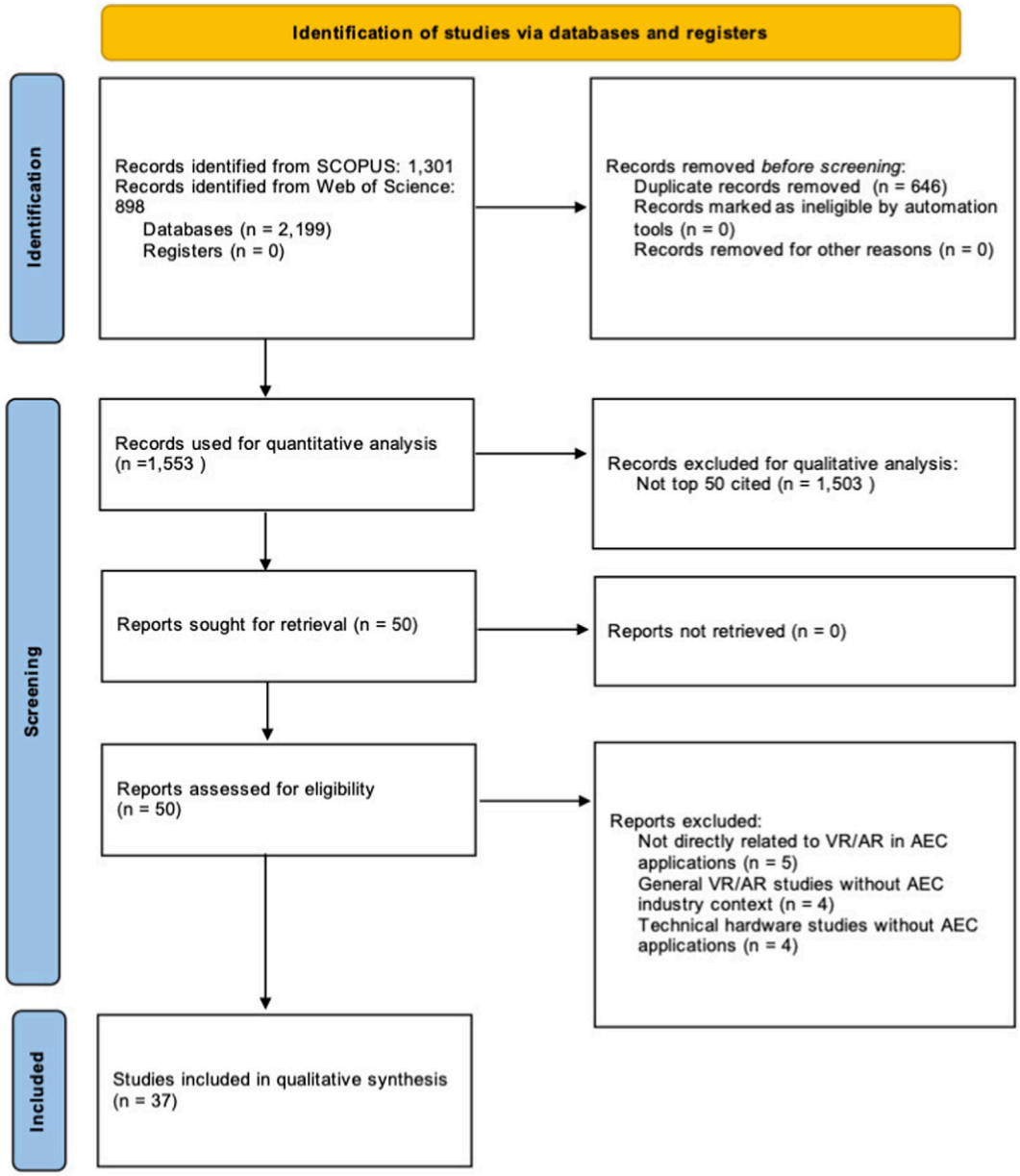


FIGURE 1
PRISMA flow diagram showing study selection.

1.4 Research objectives

This systematic review aimed to provide a comprehensive mapping of VR/AR research evolution in the AEC industry through an integrated bibliometric and thematic analysis. Specific objectives include.

1. Quantify publication trends, growth patterns, and collaboration networks in VR/AR research for AEC applications (2015–2025)
2. Identify main research clusters, thematic areas, and intellectual structure through science mapping techniques
3. Synthesize key application domains, technological integration patterns, and implementation challenges
4. Determine research gaps and propose future research directions for VR/AR technologies in the AEC industry

2 Materials and methods

2.1 Protocol and registration

This systematic review was conducted following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (Figure 1) to ensure methodological rigor and transparency (Page et al., 2021). The review protocol integrated quantitative bibliometric analysis with qualitative

thematic synthesis to provide a comprehensive mapping of Virtual Reality and Augmented Reality research in the Architecture, Engineering, and Construction industries.

2.2 Search strategy and database selection

A comprehensive literature search was performed using two major academic databases—Scopus and Web of Science—to ensure robust multi-database validation as recommended for bibliometric studies. Both databases were selected for their extensive coverage of engineering and technology publications, comprehensive citation tracking capabilities, and robust metadata for bibliometric analysis. The search strategy was designed to capture publications addressing VR/AR technologies specifically applied within the context of the AEC industry.

The inclusion of contextual terms such as 'building information model', 'BIM', and 'digital twin' alongside core AEC descriptors was a deliberate methodological choice to ensure comprehensive retrieval of VR/AR applications within established AEC digital workflows. Importantly, these terms were connected using the Boolean operator OR, meaning articles needed to match only one of these contextual terms rather than all of them, thereby expanding rather than restricting the search scope. These terms function as contextual anchors rather than thematic constraints, as preliminary searches using only VR/AR terms with general AEC descriptors yielded excessive noise from unrelated domains (e.g., medical VR, gaming applications). This approach aligns with bibliometric search strategies that recommend combining technology terms with domain-specific connectors to improve precision without sacrificing recall (Aria and Cuccurullo, 2017). Identical search queries were adapted to the syntax of each database to maintain consistency.

2.2.1 Scopus query

TITLE-ABS-KEY ((“virtual reality” OR “augmented reality” OR VR) AND (“civil engineering” OR “construction industry” OR “architecture engineering construction” OR “building information model*” OR BIM OR “digital twin*” OR “structural engineering” OR “construction management”)) AND PUBYEAR ≥ 2015 AND PUBYEAR <2026 AND (LIMIT-TO (DOCTYPE, “ar”)) AND (LIMIT-TO (LANGUAGE, “English”)) AND (LIMIT-TO (SUBJAREA, “ENGI”)).

2.2.2 Web of science query

TS=(“virtual reality” OR “augmented reality” OR VR) AND TS=(“civil engineering” OR “construction industry” OR “architecture engineering construction” OR “building information model*” OR BIM OR “digital twin*” OR “structural engineering” OR “construction management”) AND PY=(2015–2025) AND DT=(Article) AND LA=(English). Research Area Filter: Engineering.

The application of the Engineering subject area filter in both databases was a deliberate choice to maintain focus on technical implementation studies rather than purely theoretical or pedagogical works. While this filter is labeled 'Engineering,' it encompasses a broad range of AEC-related disciplines in both Scopus and Web of Science, including civil engineering, construction engineering, architectural engineering, and building technology. Analysis of the retrieved dataset confirmed that publications from construction management, architectural design, and related AEC disciplines were successfully captured, as evidenced by the diversity of source journals including Engineering, Construction and Architectural Management; Journal of Construction Engineering and Management; and Automation in Construction (Table 3). This filtering approach balanced comprehensiveness with relevance, reducing noise from non-technical domains while retaining the interdisciplinary breadth characteristic of AEC research.

The search encompassed publications from 2015 to 2025. This timeframe was selected to capture the period of significant technological advancement and increased adoption of immersive technologies in construction practices.

2.3 Eligibility criteria and selection process

This study employed strict inclusion criteria to ensure the relevance and quality of selected publications. The included studies were peer-reviewed journal articles published between 2015 and 2025 in English, focusing on VR and AR technologies applied within the context of the AEC industry. While the search strategy centered on VR and AR as the most established and documented immersive technologies in AEC literature, articles addressing MR applications were captured when indexed under VR or AR terminology, reflecting the overlapping nature of these technologies in current academic databases. Research was required to address applications in civil engineering, construction management, architectural design, or related disciplines while demonstrating the clear integration of immersive technologies within AEC workflows.

Exclusion criteria eliminated conference proceedings, book chapters, and grey literature, along with studies mentioning VR/AR technologies without substantial AEC application focus, purely technical hardware development studies without AEC implementation, and articles in languages other than English.

The selection process followed a systematic dual-database approach. The Scopus search yielded 1,301 articles after applying the engineering subject area filter, while the Web of Science search returned 898 articles, resulting in a combined total of 2,199 articles. Using the Bibliometrix mergeDbSources function with automatic duplicate detection, 646 duplicated documents were identified and removed based on DOI, title, and author matching algorithms, representing a 29.4% overlap between databases.

The final merged corpus comprised 1,553 unique articles, which formed the complete dataset for comprehensive bibliometric analysis. From this merged dataset, the top 50 most-cited articles were selected for detailed qualitative analysis. This citation-based threshold follows established bibliometric practices where “authors often exclusively use documents or journals that exceed some minimum citation threshold for the purpose of selecting only influential publications and limiting the core document set to a manageable size” (Zupic and Čater, 2015, p. 441). Accordingly, the selection of 50 highly-cited articles ensures comprehensive coverage of seminal works that have shaped the field's intellectual structure while maintaining analytical depth for rigorous qualitative synthesis.

Following rigorous title and abstract screening against inclusion criteria, 37 articles met the requirements for qualitative thematic synthesis, representing the most influential and relevant research in VR/AR applications for the AEC industry.

2.4 Data extraction and analysis methods

Comprehensive bibliometric data were extracted for all 1,553 articles to enable thorough quantitative analysis. This included publication metadata including titles, authors, journals, and publication years, alongside citation counts and patterns, author collaboration networks and institutional affiliations, keywords and subject classifications, and journal impact metrics with source distributions.

For the 37 selected articles subjected to qualitative analysis, detailed thematic data were extracted focusing on application domains within the AEC industry, technology integration approaches and methodologies, implementation challenges and limitations, future research directions and recommendations, and technological convergence patterns with other digital technologies.

2.5 Bibliometric analysis tools

Quantitative analysis was performed using the Bibliometrix R package, which provides comprehensive capabilities for bibliometric and scientometric analysis. The analysis encompassed performance analysis including annual publication trends and growth patterns, journal productivity and citation impact, author collaboration metrics and institutional analysis, and citation performance with temporal citation patterns.

Science mapping techniques included co-occurrence analysis of keywords and concepts, bibliographic coupling and co-citation analysis, thematic clustering and research front identification, and collaboration network visualization. Temporal analysis examined citation aging and knowledge diffusion patterns, whereas conceptual structure analysis using Multiple Correspondence Analysis revealed the intellectual organization of research themes and their dimensional relationships. This integrated methodology enabled comprehensive understanding of both quantitative research patterns and qualitative thematic developments within VR/AR applications in the AEC industry, providing robust foundations for identifying research gaps and future directions.

3 Results

3.1 Study selection and characteristics

The systematic dual-database search yielded 1,301 articles from Scopus and 898 articles from Web of Science, totaling 2,199 articles spanning 2015–2025. After applying the mergeDbSources function to eliminate duplicates, 646 overlapping documents were removed (29.4% duplication rate), resulting in a final merged corpus of 1,553 unique articles that formed the complete dataset for the quantitative bibliometric analysis (Table 1). The merged dataset encompassed publications from 453 different sources, with an average document age of 2.45 years. The collection included contributions from 4,707 authors, with an average of 4.42 co-authors per document. International co-authorships represented 17.41% of the total publications, indicating substantial global collaboration in VR/AR research within the AEC domain. The 29.4% overlap between Scopus and Web of Science databases demonstrates complementary coverage, with each database contributing unique articles while providing cross-validation for core influential research.

TABLE 1 Main information about merged dataset.

Description	Results
Databases	Scopus + web of science
Scopus articles	1,301
Web of science articles	898
Total before deduplication	2,199
Duplicates removed	646
Duplication rate	29.4%
Final unique documents	1,553
Timespan	2015:2025
Sources	453
Annual growth rate %	42.78
Document average age	2.45
Average citations per doc	29.75
Authors	4,707
Authors of single-authored docs	68
Single-authored docs	71
Co-authors per doc	4.42
International co-authorships %	17.41

3.2 Publication trends and growth patterns

As shown in Figure 2, temporal analysis of the merged 1,553-article dataset revealed exponential growth in VR/AR research within the AEC industry from 2015 to 2025. The number of publications increased dramatically from nine articles in 2015 to 331 in 2024, representing a nearly 37-fold increase over the last decade. The most significant growth occurred after 2018, with publications increasing from 48 in 2018 to 76 in 2019, followed by consistent year-over-year increases. The growth acceleration intensified after 2020, with publications increasing from 93 in 2020 to 147 in 2021 (58% increase), 229 in 2022 (56% increase), 263 in 2023 (15% increase), and peaking at 331 in 2024. The year 2025 recorded 317 articles, representing nearly complete annual output given the data collection timeframe. This sustained exponential growth pattern, reflected in the 42.78% annual growth rate, demonstrates the rapidly increasing research interest and maturation of VR/AR technologies for AEC applications.

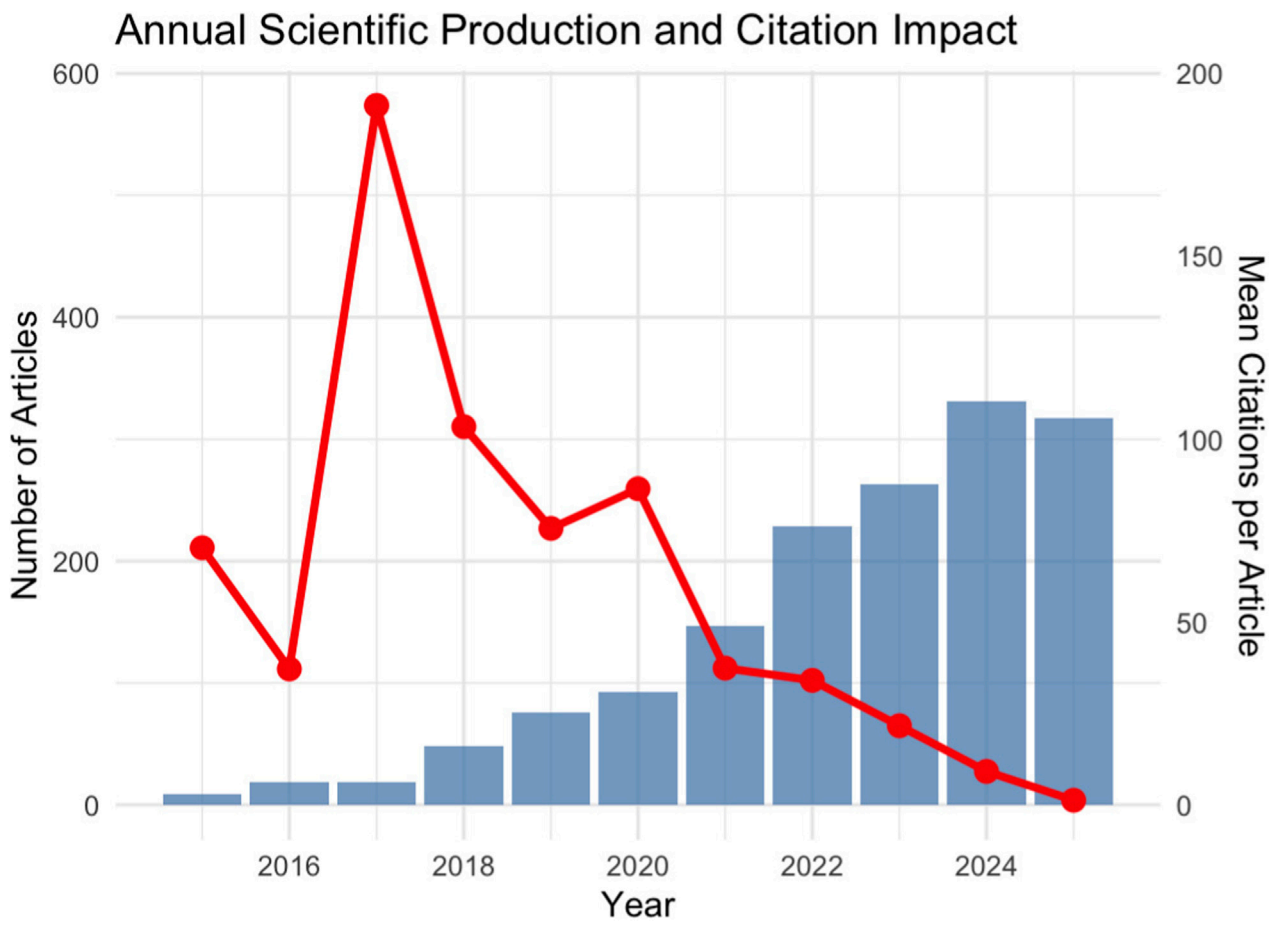


FIGURE 2
Annual publication growth (bars) and citation impact (line) of VR/AR research in AEC, 2015-2025.

TABLE 2 Citation performance by year (merged dataset).

Year	Articles	Mean citations per article	Mean citations per year	Citable years
2015	9	70.33	6.39	11
2016	19	37.16	3.72	10
2017	19	191.26	21.25	9
2018	48	103.44	12.93	8
2019	76	75.63	10.80	7
2020	93	86.49	14.41	6
2021	147	37.39	7.48	5
2022	229	34.04	8.51	4
2023	263	21.65	7.22	3
2024	331	9.15	4.58	2
2025	317	1.26	1.26	1

Citation performance analysis (Table 2) revealed interesting temporal patterns in research impact across the merged 1,553-article dataset. Earlier publications demonstrated higher average citations per article, with 2017 showing the highest mean total citations per article (191.26), followed by 2018 (103.44) and 2020 (86.49). The mean citations per year metric

TABLE 3 Top publication sources.

Sources	Articles
Automation in construction	84
IEEE access	72
Buildings	69
Applied sciences-basel	67
Journal of information technology in construction	42
Sensors	42
Advanced engineering informatics	28
Journal of construction engineering and management	28
Engineering construction and architectural management	27
International journal of advanced manufacturing technology	22
Electronics	20
Journal of manufacturing systems	20
Journal of computing in civil engineering	19
Safety science	19
Smart and sustainable built environment	17
Construction innovation	16
International journal of computer integrated manufacturing	14
International journal of construction management	14
Applied sciences (Switzerland)	13
Journal of building engineering	13
Robotics and computer-integrated manufacturing	13
IEEE communications magazine	11
IEEE transactions on consumer electronics	11
Journal of industrial information integration	11
Multimedia tools and applications	11

indicated that 2017 publications had the highest annual citation rate (21.25).

3.3 Source analysis and journal distribution

The research landscape across the merged 1,553-article dataset was dominated by several key journals demonstrating sustained focus on VR/AR technologies in AEC applications (Table 3). Automation in Construction emerged as the leading publication venue with 84 articles, followed by IEEE Access with 72 articles and Buildings with 69 articles. Applied Sciences-Basel contributed 67 articles, while the Journal of Information Technology in Construction published 42 articles, establishing its role as a specialized venue for construction technology research. Sensors also contributed 42 articles, demonstrating strong interest in IoT and sensing technologies integration with VR/AR systems.

The diversity of publication venues reflects the interdisciplinary nature of VR/AR research in AEC, spanning construction management journals, engineering technology publications, and manufacturing systems outlets. Advanced Engineering Informatics (28 articles) and the Journal of Construction Engineering and Management (28 articles) demonstrated a sustained interest in construction and engineering informatics communities. Engineering, Construction and Architectural Management (27 articles) represent traditional construction management outlets embracing immersive technologies. The presence of manufacturing-focused journals such as the International Journal of Advanced Manufacturing Technology (22 articles) and the Journal of Manufacturing Systems (20 articles) indicated significant cross-pollination between the construction and manufacturing applications of VR/AR technologies.

Emerging interdisciplinary venues included Smart and Sustainable Built Environment (17 articles), Construction Innovation (16 articles), and Safety Science (19 articles), demonstrating the technology's application across sustainability, innovation, and safety domains. Specialized computing journals such as the Journal of Computing in Civil Engineering (19 articles) and electronics-focused outlets such as Electronics (20 articles) further illustrated the technological breadth of VR/AR research.

The merged dataset's source distribution of 453 unique journals demonstrates complementary coverage between Scopus and Web of Science, which captures both mainstream construction engineering journals and emerging interdisciplinary venues, providing comprehensive coverage of the VR/AR research landscape in AEC contexts.

3.4 Author collaboration networks

The analysis of author productivity (Table 4) across the merged 1,553-article dataset revealed concentrated research activity among leading scholars in the field. Yonghao Wang and Huazhen Zhang emerged as the most prolific authors with 22 articles each, followed by Steven K. Ayer (21 articles) and Yuxuan Zhang (20 articles). The fractionalized article counts, which account for multiple authorship, showed Steven K. Ayer leading with 6.44 fractionalized articles, followed by Yonghao Wang with 5.58 and Jong-Hyun Kim with 5.23, indicating substantial individual contributions and leadership roles in research publications.

Author collaboration patterns revealed both individual productivity and extensive collaborative research networks. The top 10 authors included Jae Yeol Lee, Yang Aron Liu, and Zining Liu (19 articles each); Xuewen Wang (18 articles), Jong-Hyun Kim and Jingwen Wang (17 articles each. The presence of multiple authors with the same surnames (Wang, Zhang, Li, Liu) reflects the strong representation of Chinese researchers in the field, which is consistent with the geographic distribution observed in collaboration network analyses.

The fractionalized article counts provide insight into authorship positions and research leadership. Authors with high fractionalization ratios relative to total articles (such as Steven K. Ayer: 6.44/21 = 0.31, Zining Liu: 4.78/19 = 0.25, and Jong-Hyun Kim: 5.23/17 = 0.31) demonstrated frequent first or corresponding author positions, indicating research leadership roles. Conversely,

TABLE 4 Most productive authors (top 10).

Authors	Affiliation	Articles	Articles fractionalized
Yonghao Wang	Birmingham City University	22	5.58
Huazhen Zhang	University of Science and Technology Beijing	22	4.10
Steven K. Ayer	Arizona State University	21	6.44
Yuxuan Zhang	Guangzhou University=	20	3.83
Jae Yeol Lee	Chonnam National University	19	3.62
Yang Aron Liu	Academy of Mathematics and Systems Science	19	3.69
Zining Liu	Electric Power Research Institute	19	4.78
Xuewen Wang	Taiyuan University of Technology	18	4.05
Jong-Hyun Kim	Inha University	17	5.23
Jingwen Wang	University of Alberta	17	3.36

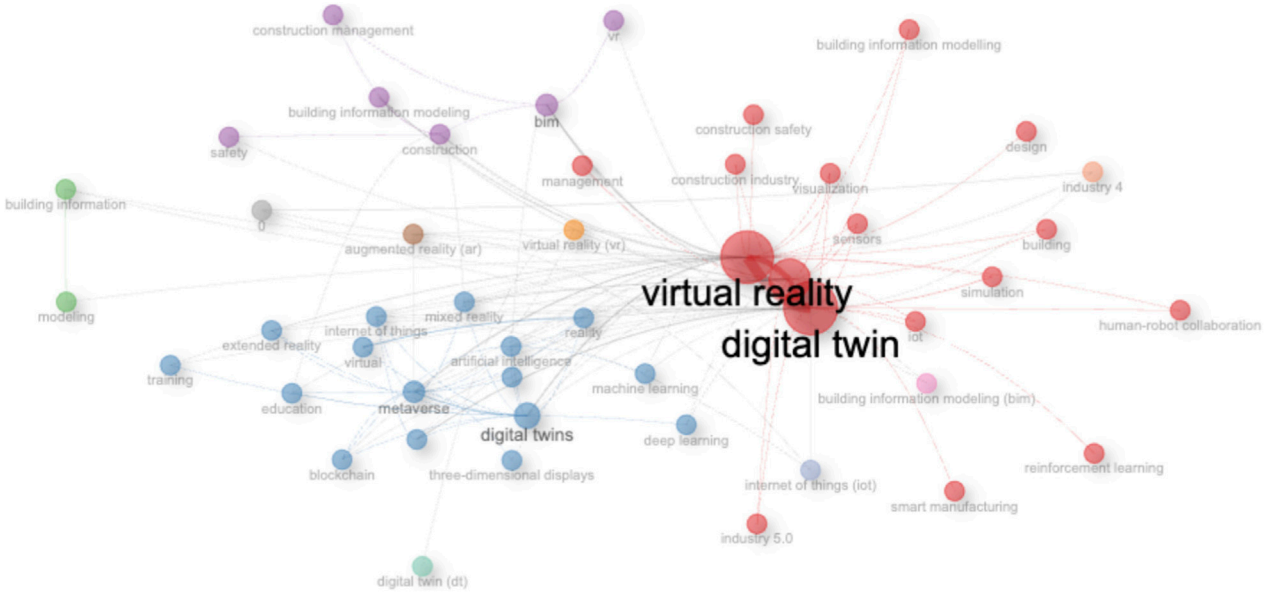


FIGURE 3
Keyword co-occurrence network showing two main clusters: digital twin applications (red) and emerging technologies (blue).

authors with lower fractionalization ratios indicated extensive collaborative partnerships. The presence of authors from diverse geographic regions, as evidenced by international co-authorship rates of 17.41%, demonstrated the global nature of VR/AR research in AEC applications, with the merged dataset capturing collaborative networks spanning multiple continents and research institutions.

3.5 Science mapping and co-occurrence analysis

The science mapping analysis of the merged 1,553-article dataset revealed distinct thematic clusters within VR/AR research for AEC applications. The keyword co-occurrence network (Figure 3) identified interconnected research areas spanning technical implementation, application domains, and integration approaches. Network Analysis revealed two primary clusters based on betweenness centrality metrics (Figure 3).

Cluster 1 (Red nodes) centered on "digital twin" forms the largest thematic group, encompassing core VR/AR technologies and their direct AEC applications. This cluster demonstrated the highest betweenness centrality values, with virtual reality (453.025), digital twin (319.240), and augmented reality (188.486) serving as the primary conceptual bridges connecting the different research sub-domains. Key nodes in this cluster included BIM (376 occurrences), building information modeling (327 occurrences), augmented reality (234 occurrences), and industry 4.0 (109 occurrences),

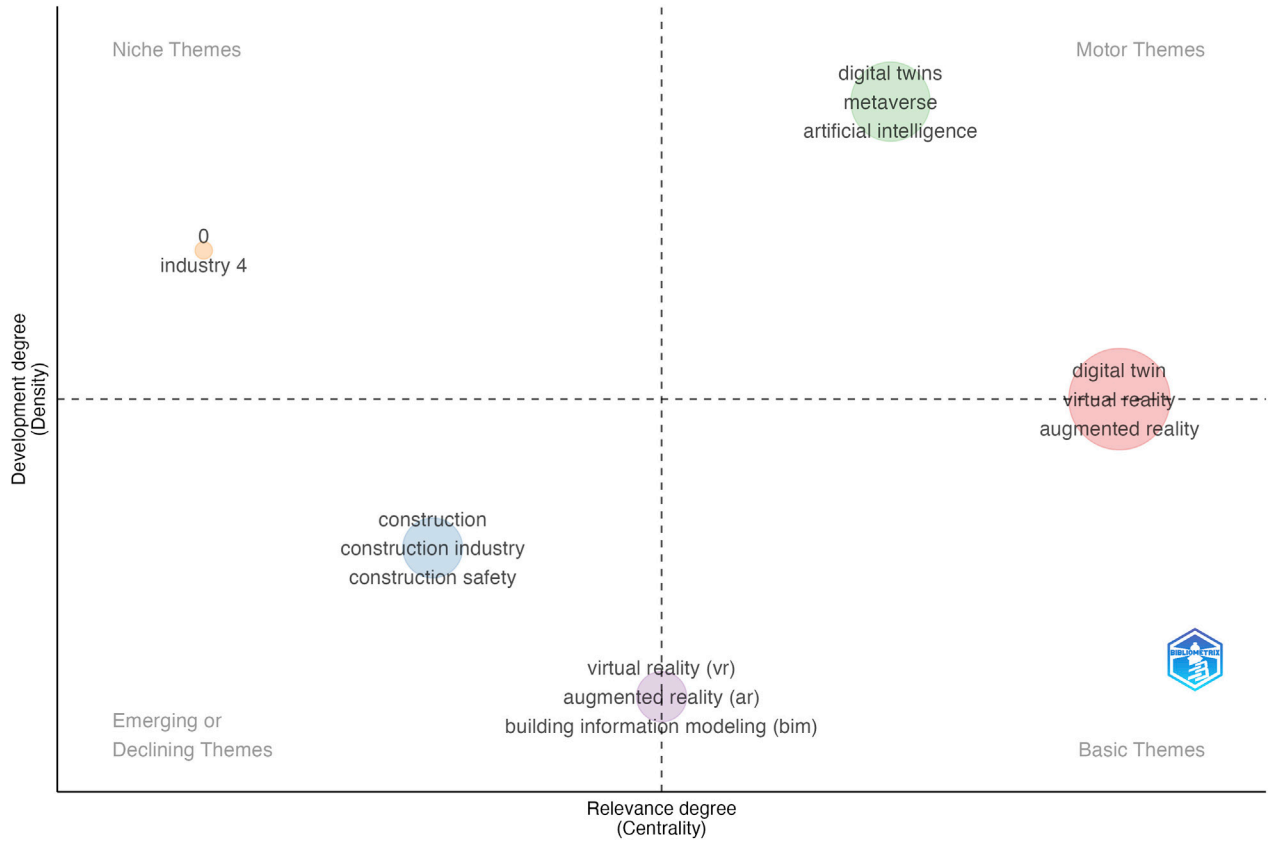


FIGURE 4
Thematic map displaying research themes by centrality and density.

indicating strong integration between immersive technologies and established AEC digital workflows. Application-focused keywords such as construction safety, simulation, visualization, and sensors demonstrated zero betweenness centrality, suggesting that they represent terminal application nodes rather than bridging concepts.

Cluster 2 (Blue nodes) centered on "digital twins" (plural form), "metaverse", and emerging technologies, represented a conceptually distinct but related research stream. This cluster showed substantial betweenness centrality for metaverse (175.723) and digital twins (19.649), indicating their role as bridges into emerging computational paradigms. Keywords in this cluster included artificial intelligence, Internet of Things, reality, and mixed reality, demonstrating the convergence of VR/AR technologies with advanced computational approaches and Industry 4.0/5.0 concepts.

Thematic map analysis (Figure 4) provided a complementary view of the field's intellectual structure organized across two dimensions: relevance degree (centrality) and development degree (density). The analysis identified four distinct quadrants revealing research maturation levels.

Motor Themes (upper-right quadrant) included "digital twins," "metaverse," and "artificial intelligence," representing well-developed and highly central research areas that drive the field forward. These themes demonstrate both internal cohesion (high density) and strong connections to other research areas (high centrality), positioning them as the strategic research fronts of the field.

Basic Themes (lower-right quadrant) contained the largest cluster with "digital twin," "virtual reality," and "augmented reality" as core concepts. Despite their lower density, these themes showed high centrality, indicating that they serve as fundamental, transversal concepts that connect multiple research streams but lack specialized development within specific niches. This positioning suggests these are mature and widely, adopted technologies serving as foundations for more specialized applications.

Niche Themes (upper-left quadrant) included "industry 4.0″as a specialized but peripheral research area. The single term in this quadrant demonstrated high internal development but limited connection to mainstream VR/AR research, suggesting industry-specific applications that remain somewhat isolated from core technological developments.

The Emerging or Declining Themes (lower-left quadrant) encompassed "construction," "construction industry," "construction safety," "virtual reality (VR)," "augmented reality (AR)," and "Building Information Modeling (BIM)." The positioning of these application-focused terms in this quadrant indicates that they represent either emerging application areas still developing conceptual coherence or declining research interests superseded by more integrated approaches such as digital twins and metaverse applications.

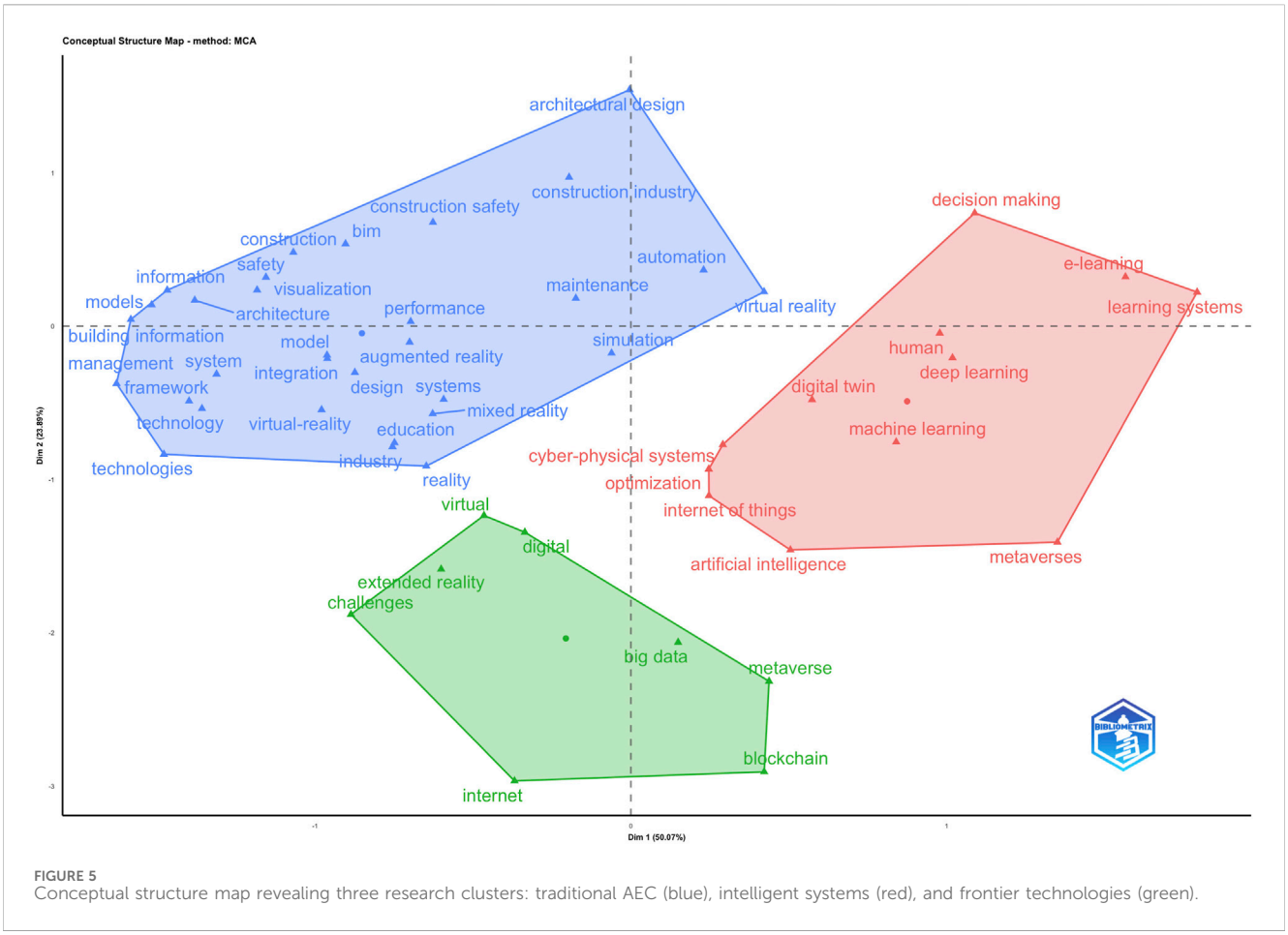


FIGURE 5
Conceptual structure map revealing three research clusters: traditional AEC (blue), intelligent systems (red), and frontier technologies (green).

Conceptual structure analysis using Multiple Correspondence Analysis (Figure 5) revealed the dimensional organization of the research field across two principal components explaining 73.96% of the total variance in keyword associations (Dim 1: 50.07%; Dim 2: 23.89%). The analysis identified three distinct thematic clusters representing different research trajectories within VR/AR applications for the AEC industry.

Cluster 1 (Blue) encompasses traditional AEC applications and core immersive technologies, including "construction," "bim," "construction safety," "visualization," "architecture," "design," "building information," "augmented reality," and "virtual-reality." This cluster occupies the upper-left region of the map, representing established application domains where VR/AR technologies have achieved substantial implementation maturity.

Cluster 2 (Red) centers on emerging intelligent systems and learning technologies, featuring "digital twin," "machine learning," "deep learning," "decision making," "e-learning," "learning systems," and "human." Positioned in the right region of the map, this cluster demonstrates the convergence of VR/AR technologies with artificial intelligence and advanced computational approaches for enhanced decision support and training applications.

Cluster 3 (Green) represents frontier computational paradigms, including "metaverse," "blockchain," "big data," "artificial intelligence," "extended reality," "internet of things," and "cyber-physical systems." Located in the lower region of the map, this cluster indicates emerging research directions exploring the integration of immersive technologies within broader digital transformation frameworks.

The spatial organization demonstrates that core VR/AR concepts serve as conceptual bridges connecting these three domains, with terms such as "simulation," "virtual reality," and "mixed reality" positioned at cluster intersections. This three-cluster structure reveals a field simultaneously advancing along complementary trajectories: Consolidating established AEC applications, integrating intelligent decision-support capabilities, and expanding toward comprehensive cyber-physical ecosystems.

3.6 Citation performance metrics

The overall citation performance of the merged 1,553-article dataset indicated a substantial research impact, with an average of 29.75 citations per document across the entire collection. Temporal citation analysis revealed that documents from 2017 to 2020 achieved the highest citation impact, with 2017 publications showing the highest mean total citations per article (191.26) and the highest annual citation rate (21.25 citations per year), suggesting that this year produced particularly influential research that shaped subsequent developments in the field. Publications from 2018 (103.44 mean citations, 12.93 per year) and 2020 (86.49 mean citations, 14.41 per year) also demonstrated strong citation

performance, establishing these years as foundational periods for VR/AR research in AEC applications.

The citation patterns reflected both the growing recognition of VR/AR technologies' potential in AEC applications and the increasing maturity of research methodologies and practical implementations. Earlier publications (2015–2017) demonstrated higher average citations per article due to longer citation windows, with 2015 showing 70.33 mean citations (6.39 per year) and 2016 showing 37.16 mean citations (3.72 per year). The consistently high citations per year for 2017–2020 publications, despite varying publication volumes, indicate sustained research quality and impact during this period of field establishment.

More recent publications naturally showed lower absolute citation counts owing to limited citation windows, with 2024 articles averaging 9.15 total citations (4.58 per year) and 2025 articles averaging 1.26 citations. However, the annual growth rate of 42.78% in publication output demonstrates the field's rapid expansion, suggesting that while individual article citations may be lower for recent works, collective knowledge production and dissemination has accelerated substantially. The multi-database validation approach, with a 29.4% overlap between Scopus and Web of Science, ensured comprehensive citation tracking across complementary indexing systems, providing robust citation metrics for the merged corpus.

The science mapping and bibliometric analysis of the 1,553-article dataset established a comprehensive quantitative foundation for understanding the research landscape of VR/AR technologies in AEC applications. The identified patterns of growth, collaboration, and thematic development, validated across two major databases, provide essential context for the qualitative thematic analysis that follows in the discussion section.

4 Discussion

4.1 Main application domains in AEC

The qualitative analysis of the 37 most influential studies reveals four primary application domains where VR/AR technologies demonstrate transformative potential within the AEC industry. These domains represent mature areas of implementation with demonstrated benefits and established research foundations.

Design and architectural visualization have emerged as the most developed application area, with VR providing superior conditions for collaborative decision-making and problem-solving during conceptual design phases (Paes et al., 2017). Users can experience materials, lighting, furniture, and architectural details from multiple perspectives on a real-world scale, thereby addressing the fundamental limitations of traditional BIM presentations (Alizadehsalehi et al., 2020). These immersive environments enhance spatial perception and understanding of 3D architectural models, enabling more suitable solutions for spatial problems through experiential interaction.

This technology has been specifically explored for specialized applications such as signage visibility analysis and optimization within BIM-enabled virtual reality environments (Dávila Delgado et al., 2020; Dávila Delgado et al., 2020). Despite these advantages, there remains a recognized scarcity of comparative, quantitative, and user-centered research demonstrating the effectiveness of immersive VR systems over traditional non-immersive approaches in architectural design (Paes et al., 2017).

Construction planning, monitoring, and assembly represent the second major application domain with demonstrated quantitative benefits. Experimental studies on Mixed Reality applications for electrical conduit construction design communication have revealed substantially higher productivity rates, reduced design comprehension time, fewer assembly errors, and increased accuracy in constructed conduits compared to traditional paper-based plans (Chalhoub and Ayer, 2018). AR visualization in piping assembly operations achieved remarkable improvements including a 50% reduction in task completion time, 50% decrease in assembly errors, 55% reduction in original time requirements, 46% decrease in rework time, and 66% reduction in error correction costs (Hou et al., 2015). Progress monitoring capabilities have been enhanced through frameworks integrating image processing, machine learning, BIM, and VR technologies for the on-demand automated simulation of construction projects (Rahimian et al., 2020).

Safety training and human-robot collaboration constitute the third critical application area, addressing workforce development and hazard mitigation challenges. VR applications in the Human-Robot Collaboration workspace design enable the comprehensive testing of workstation layouts, skill-based task distribution, virtual commissioning procedures, and safety protocol validation (Oyekan et al., 2019; Malik et al., 2020). Construction safety and health education benefit substantially from social and collaborative VR platforms that allow students to identify hazard scenarios, correct unsafe conditions through simulation tools, and acquire comprehensive safety knowledge through experiential learning (Le et al., 2015). The iCRT system captures detailed information from complex and hazardous construction activities such as underground excavation and piling operations, reproducing authentic practices in immersive virtual environments (Sepasgozar, 2020).

Facility management and urban planning applications represent the fourth domain, extending the VR/AR impact beyond construction phases into building operations and broader urban contexts. AR applications support the BIM-based inspection and maintenance of critical building systems, particularly fire safety equipment, enabling real-time access to system information and maintenance protocols (Chen et al., 2021). Urban planning and smart city development benefit from the integration of Geographic Information Systems and BIM as fundamental components for developing comprehensive 3D urban models, facilitating intelligent city planning processes, and managing complex urban infrastructure networks (Xia et al., 2022).

4.2 Technology integration patterns

The analysis revealed sophisticated integration patterns between VR/AR technologies and complementary digital systems that amplify their effectiveness within AEC workflows. These integrated approaches represent critical success factors for practical implementation and sustained adoption.

Digital Twins and Extended Reality convergence emerges as a fundamental integration pattern, with Digital Twins representing

digital environments maintaining bidirectional linkages with physical counterparts, enabling advanced simulation capabilities and data-centric decision-making processes (Shahzad et al., 2022). XR technologies serve as critical interfaces for visualizing and interacting with Digital Twin systems, creating immersive bridges between virtual models and physical assets (Mourtzis et al., 2022; Alizadehsalehi and Yitmen, 2023). The Digital Twin-Reality Capture-to-Extended Reality system represents advanced integration, combining reality capture and XR technologies to create, capture, generate, analyze, manage, and visualize construction progress data and comprehensive reports (Alizadehsalehi and Yitmen, 2023).

Building Information Modeling integration serves as the foundational technology for VR/AR implementation, acting as a comprehensive source of geometric and building information (Du and Shi, 2018; Du and Zou, 2018; Alizadehsalehi et al., 2020). While interoperability between BIM and VR/AR tools remains fundamental, it has been identified as a primary obstacle owing to manual conversion processes and information loss during data transfer (Du and Shi, 2018; Rahimian et al., 2020). Despite these challenges, BIM-VR/AR integration enables the visualization of building designs in interactive, human-scale, and immersive formats, allowing users to interact with 3D models superimposed onto physical spaces while accessing additional contextual information.

The synergies of IoT, AI, and cloud computing synergies create comprehensive intelligent systems that enhance VR/AR capabilities through real-time data integration and advanced processing. IoT sensors provide real-time data from physical environments, which are integrated into BIM models and visualized through AR and VR applications, improving capabilities including fire safety monitoring and situational awareness (Tao and Zhang, 2017; Chen et al., 2021; Mihai et al., 2022). Artificial Intelligence and Machine Learning technologies enhance VR/R and Digital Twin system capabilities through advanced algorithms and automated processes, enabling intelligent systems through advanced control algorithms and sophisticated data processing capabilities (Schluse et al., 2018).

4.3 Current challenges and limitations

Despite significant demonstrated potential, the widespread adoption of VR/AR technologies in AEC remains constrained by numerous technical, economic, and human factors that require systematic attention for technology maturation.

Technical infrastructure and performance limitations are primary obstacles to practical implementation. High processing and equipment requirements demand specialized high-performance computing equipment and sophisticated algorithms operating under substantial computational loads to deliver satisfactory immersive experiences (Li X., 2018; Dávila Delgado et al., 2020). Building Information Modeling data conversion for visualization creates exceptionally large files that challenge even powerful computing systems, resulting in significant file size limitations that constrain model complexity and detail levels (Zaker and Coloma Pico, 2018; Alizadehsalehi et al., 2020). Hardware limitations persist across display technologies, with current head-mounted displays suffering from inadequate resolution and restricted fields of view, whereas commercial AR headsets demonstrate reliability issues including low refresh rates and resolution problems (Dávila Delgado et al., 2020; Chen et al., 2021).

Industry adoption and economic barriers constitute the most significant obstacles to the widespread implementation of VR/AR. High costs associated with hardware acquisition and comprehensive training consistently rank as the primary limiting factors (Zaker and Coloma Pico, 2018; Dávila Delgado et al., 2020). Industry perceptions continue to view AR and VR as expensive and immature technologies that are unsuitable for complex engineering and construction tasks (Dávila Delgado et al., 2020).

Construction firms frequently lack comprehensive knowledge of AR and VR market dynamics and capabilities, while skill shortages and difficulties accessing qualified graduates and expert knowledge in these technologies persist throughout the industry. Resistance to technological change represents a fundamental barrier, particularly among personnel and firms demonstrating aversion to adopting innovative technologies (Zaker and Coloma Pico, 2018).

Human factors and user experience challenges significantly impact VR/AR technology effectiveness and user acceptance. Users frequently report unrealistic interactions in virtual environments, such as multiple attempts to operate virtual door handles, leading to reduced efficiency and inconsistent experiences (Chen et al., 2021). VR training systems, while beneficial, may result in loss of the real-life experience and fail to replicate authentic mental load and stress levels compared to actual environments (Li X., 2018). Cognitive overload and distraction represent ongoing concerns, as AR and VR technologies can potentially increase cognitive loads and cause user distraction (Dávila Delgado et al., 2020).

4.4 Research gaps and future directions

The analysis identified significant research gaps that require systematic investigation for technology advancement and practical implementation improvement. Current research demonstrates a notable scarcity of comparative, quantitative, user-centered research demonstrating the effectiveness of immersive VR systems compared to traditional non-immersive approaches in architectural design applications (Paes et al., 2017). Empirical investigations requiring larger, more diverse participant groups, including expert practitioners, are necessary for comprehensive technology validation (Sepasgozar, 2020; Chen et al., 2021).

Methodological and technological limitations require future research, including improved calibration accuracy between real movements and virtual animations (Shi et al., 2019). VR training systems need development of sophisticated human-environment interaction capabilities to enhance immersive experiences (Chen et al., 2021), while automatic evaluation processes and methods for VR/AR applications require development (Li Z., 2018). Research approaches predominantly focus on technical development and individual risk factors rather than adopting multi-disciplinary system-thinking approaches that consider various risk types across complete project lifecycles (Zou et al., 2017).

Digital Twin research requires investigation into appropriate and realistic fidelity levels for different use cases, balancing implementation

benefits with costs and technical complexity (Jones et al., 2020). Data ownership and integration between virtual entities represent critical but under-researched areas for Digital Twin development, while ethical considerations including privacy, transparency, and security value system encoding in socio-technical Digital Twins require further investigation (Mihai et al., 2022).

4.5 Implications for practice and research

The integrated analysis revealed both immediate opportunities and long-term strategic directions for VR/AR technology adoption in the AEC industry context. For practitioners, evidence supports focused implementation in proven application areas including design visualization, construction safety training, and progress monitoring, where quantitative benefits have been demonstrated. Organizations should prioritize BIM-VR/AR integration capabilities while addressing interoperability challenges through standardized data exchange protocols.

For researchers, this analysis indicates the critical need for user-centered evaluation methodologies, standardized performance metrics, and interdisciplinary collaboration frameworks. Future research should emphasize practical implementation studies over purely technical development, with particular attention to economic justification models and change management strategies that address industry adoption barriers.

4.6 Study limitations

This systematic review acknowledges several limitations that may have influenced the interpretation of findings. Despite the dual-database approach using both Scopus and Web of Science, the study's reliance on these two databases, while comprehensive for engineering publications, may have excluded relevant research, published in specialized construction industry venues or emerging interdisciplinary journals. The focus on English-language publications potentially limited the geographic and cultural diversity of the included studies. The citation-based selection approach for qualitative analysis inherently favors established publications with longer citation windows; consequently, although 317 articles from 2025 were included in the quantitative bibliometric analysis, none met the citation threshold for qualitative synthesis due to insufficient time for citation accumulation. This methodological trade-off prioritizes analysis of demonstrably influential research that has shaped the field's intellectual structure, while the quantitative analysis comprehensively captures recent publication trends and emerging topics. The rapid evolution of VR/AR technologies means that the technical limitations identified in the literature may not reflect current capabilities, requiring ongoing monitoring of technological developments and practical implementations.

5 Conclusion

This systematic review and science mapping analysis addresses the research objectives by providing a comprehensive quantitative assessment of Virtual Reality and Augmented Reality within the Architecture, Engineering, and Construction industry through robust multi-database validation. The dual-database approach using Scopus (1,301 articles) and Web of Science (898 articles) yielded a merged corpus of 1,553 unique publications spanning 2015–2025, with 29.4% overlap demonstrating complementary coverage and cross-validation of core influential research across both indexing systems.

The bibliometric analysis revealed exponential growth with publications increasing 37-fold from nine articles in 2015 to 331 in 2024, with a sustained annual growth rate of 42.78%, establishing VR/AR technologies as the mainstream research focus within AEC applications. The analysis identifies concentrated expertise among leading researchers, with Yonghao Wang and Huazhen Zhang (22 articles each), Steven K. Ayer (21 articles), and Yuxuan Zhang (20 articles) demonstrating sustained productivity, whereas international co-authorships (17.41%) indicate substantial global collaboration spanning multiple continents and research institutions.

The conceptual structure analysis reveals a two-dimensional intellectual organization spanning from emerging computational technologies (metaverse, artificial intelligence, and blockchain) to established AEC applications (BIM, construction safety, and visualization), with VR/AR concepts serving as conceptual bridges between technological innovation and practical implementation. The identification of the main research clusters through science mapping techniques reveals mature application domains including design visualization, construction planning and monitoring, safety training, and facility management, each demonstrating quantifiable benefits such as 50% reductions in task completion time and assembly errors in experimental studies.

The synthesis of key application domains and technological integration patterns shows sophisticated convergence with Digital Twins, BIM, IoT, and AI systems, thereby establishing comprehensive frameworks for practical deployment. The thematic map analysis identified motor themes (digital twins, metaverse, artificial intelligence), basic themes (digital twins, virtual reality, and augmented reality), and niche themes (Industry 4.0), revealing the field's maturation and strategic research directions. However, significant barriers including technical infrastructure limitations, economic constraints (hardware costs and training requirements), and human factors (user experience challenges and cognitive overload) continue to impede widespread industry adoption.

The analysis revealed critical research gaps requiring systematic attention, particularly the scarcity of comparative, quantitative, user-centered evaluation studies demonstrating immersive the effectiveness of VR systems compared to traditional approaches, and insufficient interdisciplinary collaboration frameworks addressing complete project lifecycles. Future research directions should prioritize practical implementation studies over purely technical development, standardized evaluation methodologies, economic justification models that address real-world adoption challenges, and Digital Twin research investigating appropriate fidelity levels balancing benefits with implementation costs.

For practitioners, the evidence supports strategic implementation in proven application areas (design visualization with 37-fold publication growth, construction safety training with

demonstrated 50% error reductions, and BIM-integrated progress monitoring) while developing comprehensive change management strategies. Organizations should prioritize BIM-VR/AR integration capabilities while addressing interoperability challenges through standardized data exchange protocols. The multi-database validation approach ensures the robustness of these recommendations across complementary academic indexing systems.

This integrated quantitative and qualitative analysis, validated across Scopus and Web of Science databases with 453 unique journal sources and 4,707 contributing authors, provides essential guidance for researchers and practitioners navigating immersive technology adoption within the construction industry. The established research clusters, application domains, and integration patterns offer stable frameworks for continued innovation, while the identified limitations and research gaps define clear pathways for advancing both technological capabilities and practical implementation strategies in the rapidly evolving VR/AR landscape. The 29.4% database overlap confirms that both Scopus and Web of Science capture core influential research while contributing unique perspectives, strengthening the validity and comprehensiveness of the findings to guide future research directions and technology adoption strategies in AEC applications.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Author contributions

DS-P: Conceptualization, Data curation, Formal Analysis, Investigation, Methodology, Writing – original draft, Writing – review and editing. GB: Funding acquisition, Project administration, Validation, Writing – original draft, Writing – review and editing. MP: Project administration, Resources, Supervision, Writing – original draft, Writing – review and editing. TM-P: Data curation, Methodology, Software, Validation, Visualization, Writing – original draft, Writing – review and editing.

Funding

The author(s) declared that financial support was not received for this work and/or its publication.

Conflict of interest

The author(s) declared that this work was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declared that generative AI was used in the creation of this manuscript. During the preparation of this work, the authors used Claude Sonnet 4 and Grammarly for writing assistance and for grammar, spelling, and style review. The authors declare that they have reviewed and edited the final output as needed and take full responsibility for the content of the published article.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Adami, P., Rodrigues, P. B., Woods, P. J., Becerik-Gerber, B., Soibelman, L., Copur-Gencturk, Y., et al. (2022). Impact of VR-Based training on human–robot interaction for remote operating construction robots. *J. Comput. Civ. Eng.* 36 (3), 04022006. doi:10.1061/(ASCE)CP.1943-5487.0001016

Afzal, M., Shafiq, M. T., and Jassmi, H. A. (2021). Improving construction safety with virtual-design construction technologies – a review. *J. Inf. Technol. Constr.* 26, 319–340. doi:10.36680/j.itcon.2021.018

Agbo, F. J., Sanusi, I. T., Oyelere, S. S., and Suhonen, J. (2021). Application of virtual reality in computer science education: a systemic review based on bibliometric and content analysis methods. *Educ. Sci.* 11 (3), 142. doi:10.3390/educsci11030142

Alizadehsalehi, S., and Yitmen, I. (2023). Digital twin-based progress monitoring management model through reality capture to extended reality technologies (DRX). *Smart Sustain. Built Environ.* 12, 200–236. doi:10.1108/SASBE-01-2021-0016

Alizadehsalehi, S., Hadavi, A., and Huang, J. C. (2020). From BIM to extended reality in AEC industry. *Autom. Constr.* 116, 103254. doi:10.1016/j.autcon.2020.103254

Aria, M., and Cuccurullo, C. (2017). Bibliometrix: an R-tool for comprehensive science mapping analysis. *J. Informetrics* 11 (4), 959–975. doi:10.1016/j.joi.2017.08.007

Azhar, S. (2011). Building information modeling (BIM): trends, benefits, risks, and challenges for the AEC industry. *Leadersh. Manag. Eng.* 11 (3), 241–252. doi:10.1061/(ASCE)LM.1943-5630.0000127

Bou Hatoum, M., Nassereddine, H., and Badurdeen, F. (2022). Towards a canvas for construction 4.0 implementation in AECO organizations. In: Proceedings of the creative construction e-Conference 2022. Creative construction e-Conference 2022. Budapest, Hungary: Budapest University of Technology and Economics. p. 214–220. doi:10.3311/CCC2022-028

Catbas, F. N., Luleci, F., Zakaria, M., Bagci, U., LaViola, J. J., Jr, Cruz-Neira, C., et al. (2022). Extended reality (XR) for condition assessment of civil engineering structures: a literature review. *Sensors* 22 (23), 9560. doi:10.3390/s22239560

Chalhoub, J. M., and Ayer, S. K. (2018). Using mixed reality for electrical construction design communication. *Autom. Constr.* 99, 52-58. doi:10.1016/j.autcon.2017.10.028

Chen, H., Hou, L., Zhang, G. (Kevin), and Moon, S. (2021). Development of BIM, IoT and AR/VR technologies for fire safety and upskilling. *Autom. Constr.* 125, 103631. doi:10.1016/j.autcon.2021.103631

Chi, H.-Y., Juan, Y.-K., and Lu, S. (2022). Comparing BIM-based XR and traditional design process from three perspectives: aesthetics, gaze tracking, and perceived usefulness. *Buildings* 12 (10), 1728. doi:10.3390/buildings12101728

Dávila Delgado, J. M., et al. (2020). Augmented and virtual reality in construction: drivers and limitations for industry adoption. *J. Constr. Eng. Manag.* 146. doi:10.1061/(ASCE)CO.1943-7862.0001844

Du, J., Zou, Z., Shi, Y., and Zhao, D. (2018). Zero latency: real-time synchronization of BIM data in virtual reality for collaborative decision-making. *Autom. Constr* 86, 256-276. doi:10.1016/j.autcon.2017.10.009

Du, J., Shi, Y., Zou, Z., and Zhao, D. (2018). CoVR: cloud-based multiuser virtual reality headset system for project communication of remote users. *J. Constr. Eng. Manag.* 144, 04017109. doi:10.1061/(ASCE)CO.1943-7862.0001426

Hossain, M. A., and Nadeem, A. (2019). Towards digitizing the construction industry: state of the art of construction 4.0. *Proc. Int. Struct. Eng. Constr.* 6. doi:10.14455/ISEC.res.2019.184

Hou, L., Wang, X., and Truijens, M. J. (2015). Using augmented reality to facilitate piping assembly: an experiment-based evaluation. *J. Comput. Civ. Eng.* 29, 05014007. doi:10.1061/(ASCE)CP.1943-5487.0000344

Huang, W., and Roscoe, R. D. (2021). Head-mounted display-based virtual reality systems in engineering education: a review of recent research. *Comput. Appl. Eng. Educ.* 29 (5), 1420–1435. doi:10.1002/cae.22393

Jacobsen, E. L., Solberg, A., Golovina, O., and Teizer, J. (2022). Active personalized construction safety training using run-time data collection in physical and virtual reality work environments. *Constr. Innov.* 22 (3), 531–553. doi:10.1108/CI-06-2021-0113

Jones, D. E., Snider, C., Nassehi, A., Yon, J., and Hicks, B. (2020). Characterising the digital twin: a systematic literature review. *CIRP J. Manuf. Sci. Technol.* 29, 36–52. doi:10.1016/j.cirpj.2020.02.002

Klarin, A., and Xiao, Q. (2024). Automation in architecture, engineering and construction: a scientometric analysis and implications for management. *Eng. Constr. Archit. Manag.* 31 (8), 3308–3334. doi:10.1108/ECAM-08-2022-0770

Le, Q. T., Pedro, A., and Park, C. (2015). A social virtual reality based construction safety education system for experiential learning. *J. Intell. Robot. Syst.: Theory Appl.* 79, 487–506. doi:10.1007/s10846-014-0112-z

Li, S., Wang, Q.-C., Wei, H.-H., and Chen, J.-H. (2024). Extended reality (XR) training in the construction industry: a content review. *Buildings* 14 (2), 414. doi:10.3390/buildings14020414

Li, X., Yi, W., Chi, H.-L., Wang, X., and Chan, A. P. C. (2018). A critical review of virtual and augmented reality (VR/AR) applications in construction safety. *Autom. Constr* 86, 150-162. doi:10.1016/j.autcon.2017.11.003

Li, Z., Xue, F., Li, X., Hong, J., and Shen, G. Q. (2018). An internet of things-enabled BIM platform for on-site assembly services in prefabricated construction. *Autom. Constr* 196, 96-107. doi:10.1016/j.autcon.2018.01.001

Malik, A. A., Masood, T., and Bilberg, A. (2020). Virtual reality in manufacturing: immersive and collaborative artificial-reality in design of human-robot workspace. *Int. J. Comput. Integr. Manuf* 33. doi:10.1080/0951192X.2019.1690685

Mannino, A., Dejaco, M. C., and Re Cecconi, F. (2021). Building information modelling and internet of things integration for facility management—literature review and future needs. *Appl. Sci.* 11 (7), 3062. doi:10.3390/app11073062

Mihai, S., Yaqoob, M., Hung, D., Davis, W., Towakel, P., Raza, M., et al. (2022). Digital twins: a survey on enabling technologies, challenges, trends and future prospects. *IEEE Commun Surv Tutor* 24. doi:10.1109/COMST.2022.3208773

Milgram, P., and Kishino, F. (1994). A taxonomy of mixed reality visual displays. *IEICE Trans. Inf. Syst.* 77 (12), 1321–1329.

Mourtzis, D. A., Panopoulos, N., Angelopoulos, J., Wang, B., and Wang, L. (2022). Human centric platforms for personalized value creation in metaverse. *Manuf. Syst* 62, 334-346. doi:10.1016/j.jmsy.2022.11.004

Oyekan, J. O., Hutabarat, W., Tiwari, A., Grech, R., Aung, M. H., Mariani, M. P., et al. (2019). The effectiveness of virtual environments in developing collaborative strategies between industrial robots and humans *Robotics Computer-Integrated Manuf.* 55, 41–54. doi:10.1016/j.rcim.2018.07.006

Paes, D., Arantes, E., and Irizarry, J. (2017). Immersive environment for improving the understanding of architectural 3D models: comparing user spatial perception between immersive and traditional virtual reality systems. *Autom. Constr* 20, 189-195. doi:10.1016/j.autcon.2017.09.016

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., et al. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* 372, n71. doi:10.1136/bmj.n71

Rahimian, F. P., Seyedzadeh, S. M., Oliver, S., Rodriguez, S., and Dawood, N. N. (2020). On-demand monitoring of construction projects through a game-like hybrid application of BIM and machine learning. *Autom. Constr* 110, 103012. doi:10.1016/j.autcon.2019.103012

Rashid, S., Khattak, A., Ashiq, M., Ur Rehman, S., and Rashid Rasool, M. (2021). Educational landscape of virtual reality in higher education: bibliometric evidences of publishing patterns and emerging trends. *Publications* 9 (2), 17. doi:10.3390/publications9020017

Schluse, M., Priggemeyer, M., Atorf, L., and Rossmann, J. (2018). Experimentable digital twins-streamlining simulation-based systems engineering for industry 4.0. In: IEEE transactions on industrial informatics. IEEE Computer Society. doi:10.1109/TII.2018.2804917

Sepasgozar, S. M. E. (2020). Digital twin and web-based virtual gaming technologies for online education: a case of construction management and engineering. *Appl. Sci.* 10 (13), 4678. doi:10.3390/app10134678

Shahzad, M., Shafiq, M. T., Douglas, D., and Kassem, M. (2022). Digital twins in built environments: an investigation of the characteristics, applications, and challenges. *Build. MDPI* 12, 120. doi:10.3390/buildings12020120

Shi, Y., Du, J., Ahn, C. R., and Ragan, E. D. (2019). Impact assessment of reinforced learning methods on construction workers' fall risk behavior using virtual reality. *Autom. Constr* 104, 197–214. doi:10.1016/j.autcon.2019.04.015

Sidani, A., Poças Martins, J., and Soeiro, A. (2022). BIM approaches for enhanced health and safety status in construction - protocol for a systematic review: protocol. *Int. J. Occup. Environ. Saf.* 6 (1), 1–8. doi:10.24840/2184-0954_006.001_0001

Tao, F., and Zhang, M. (2017). Digital twin shop-floor: a new shop-floor paradigm towards smart manufacturing. In: IEEE access. Institute of Electrical and Electronics Engineers Inc. doi:10.1109/ACCESS.2017.2756069

Xia, H., Liu, Z., Efremochkina, M., Liu, X., and Lin, C. (2022). Study on city digital twin technologies for sustainable smart city design: a review and bibliometric analysis of geographic information system and building information modeling integration. *Sustain. Cities Soc* 84, 104009. doi:10.1016/j.scs.2022.104009

Xu, J., and Moreu, F. (2021). A review of augmented reality applications in civil infrastructure during the 4th industrial revolution. *Front. Built Environ.* 7, 640732. doi:10.3389/fbuil.2021.640732

Zaker, R., and Coloma Pico, E. (2018). Virtual reality-integrated workflow in BIM-Enabled projects collaboration and design review: a case study. *Vis. Eng* 6. doi:10.1186/s40327-018-0065-6

Zou, Y., Kiviniemi, A., and Jones, S. W. (2017). A review of risk management through BIM and BIM-related technologies. *Saf. Sci.* 97, 88–98. doi:10.1016/j.ssci.2015.12.027

Zupic, I., and Čater, T. (2015). Bibliometric methods in management and organization. *Organ. Research Methods* 18 (3), 429–472. doi:10.1177/1094428114562629

frontiers | Frontiers in Built Environment

TYPE Original Research
PUBLISHED 06 January 2026
DOI 10.3389/fbuil.2025.1720481

Check for updates

OPEN ACCESS

EDITED BY
Salman Azhar,
Auburn University, United States

REVIEWED BY
Junshan Liu,
Auburn University, United States
Rizwan Farooqui,
Mississippi State University, United States
Wasiq Ahmad,
Exelon Corp., United States

*CORRESPONDENCE
Botao Li,
botaoli825@gatech.edu
Danielle S. Willkens,
danielle.willkens@design.gatech.edu

RECEIVED 08 October 2025
REVISED 08 December 2025
ACCEPTED 12 December 2025
PUBLISHED 06 January 2026

CITATION
Li B and Willkens DS (2026) Digital transformation through multi-device HBIM workflow: a case study on supporting the adaptive reuse of the Odd Fellows Building in Atlanta, Georgia.
Front. Built Environ. 11:1720481.
doi: 10.3389/fbuil.2025.1720481

COPYRIGHT
© 2026 Li and Willkens. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Digital transformation through multi-device HBIM workflow: a case study on supporting the adaptive reuse of the Odd Fellows Building in Atlanta, Georgia

Botao Li* and Danielle S. Willkens*

School of Architecture, Georgia Institute of Technology, Atlanta, GA, United States

This paper presents a multi-device scan-to-HBIM workflow developed for the adaptive reuse of the Odd Fellows Building (b.1911), a socially significant historic landmark with seven stories above ground and a basement, located in Atlanta's Sweet Auburn National Historic Landmark District. As a major contributing structure within the culturally significant corridor, the building's adaptive reuse demands precise documentation, rapid data collection, and efficient stakeholder communication. To meet these goals within strict time and funding limits, the research integrates Terrestrial Laser Scanning (TLS), Matterport 360-camera technology, and Unmanned Aerial Vehicles (UAVs) to create an efficient and accessible pipeline for documenting historic buildings, assisting with renovation and construction processes. TLS is employed as a benchmark capture method to establish a baseline point cloud for aligning all other data. Matterport technology enables fast interior scanning while providing high-resolution visuals and an interactive virtual tour platform. UAVs are deployed to supplement data collection in hard-to-access exterior areas, especially where ground-based reality capture tools encounter spatial or logistical barriers. This not only accelerates data acquisition under time and budget constraints but also offers stakeholders the ability to remotely revisit the space without geographic limitation. This multi-device strategy optimizes the trade-off between capture speed and spatial tolerance, accommodating a complex mix of preserved historic elements (exterior and lobby space) and altered upper floors. Besides, a full TLS capture approach could cause potential data saturation and processing inefficiencies. The multi-device approach has multiple data processed registered to provide one combined point cloud, which is integrated into a semantically rich HBIM model developed in Autodesk Revit at Level of Development (LoD) 300. Although not a full Digital Twin (DT), the resulting model establishes a foundational framework for future digital twinning and immersive applications. By demonstrating how different reality capture technologies can be effectively integrated to accommodate site limitations, stakeholder needs, and funding constraints, this case study contributes to the broader discourse on digital transformation in construction, particularly the evolving intersection of BIM, remote visualization, and the early stages of DT development in the preservation sector.

KEYWORDS

built heritage, digital documentation, heritage building information modeling (HBIM), point cloud integration, reality capture

1 Introduction

The architecture, engineering, construction, and operations (AECO) industry is undergoing a profound digital transformation, driven by the convergence of Building Information Modeling (BIM), immersive visualization platforms, and reality capture technologies. Within this transition, Heritage Building Information Modeling (HBIM) has emerged as a crucial workflow for documenting, managing, and reusing heritage structures. Unlike conventional BIM practices that are oriented toward new construction, HBIM processes must contend with the complexities of irregular geometries, incomplete historical records, and the need to balance conservation ethics with renovation imperatives. Recent advances in terrestrial laser scanning (TLS), unmanned aerial vehicles (UAVs), and photogrammetry have expanded the potential of HBIM by enabling high-fidelity data capture. Traditional on-site surveys rely on human labor measurement, which is time-consuming and labor-intensive. Current HBIM methodology applies digital reality capture tools such as TLS, UAVs, etc., to achieve more accurate and time-saving measurements (Yang et al., 2020). However, despite these technological developments, challenges remain integrating diverse capture devices into efficient, accessible, and resource-conscious workflows.

Recent studies have shown that HBIM research has expanded rapidly over the last decade, shifting from a primary focus on geometric representation to a broader analysis of information management, conservation workflows, and integration with other digital systems. Yang et al. (2020) reviewed built heritage modeling and emphasized the need to integrate HBIM with photogrammetry, GIS, and semantic enrichment techniques to support more robust documentation and analysis. Dore and Murphy (2017) emphasized that the complexity of heritage structures requires adaptive workflows that balance data density and semantic modeling accuracy. Inzerillo et al. (2025) explored AI-powered modeling for heritage bridges via semantic segmentation in point clouds to reduce modeling time and improve geometric quality. Zhang and Zou (2022) used a CiteSpace-based analysis to map research hotspots and trends in HBIM, finding a change toward preventive conservation, multi-source data integration, and lifecycle management of built heritage assets. Parente et al. (2025) reviewed most recent HBIM literature, which focuses on information management and the integration of non-geometric data, such as diagnostic, monitoring, and conservation records, into BIM environments, underscoring ongoing challenges in interoperability and data structure. Alshawabkeh and Baik (2023) demonstrated how combining laser scanning and photogrammetry can enhance scan-to-HBIM workflows by enriching point clouds with texture and damage information, while Giuliani et al. (2024) proposed an HBIM pipeline for large-scale architectural heritage that emphasizes fit-for-purpose workflows and interoperability. In recent years, the convergence of HBIM and Digital Twin (DT) frameworks has gained increasing attention as scholars and practitioners seek to transform static 3D heritage models into dynamic systems capable of monitoring and simulation. There is the potential of linking HBIM with real-time data and monitoring systems but also the complexity of doing so in practice is very challenging (Lucchi, 2023). Niccolucci and Felicetti (2024) introduced an ontology-based approach that connects HBIM models with sensor networks and cultural heritage data structures, demonstrating how real-time environmental inputs can enrich the semantic layers of a digital twin. Extending these efforts, Yu et al. (2025) provided a systematic review of digital technologies for architectural heritage risk management, identifying HBIM–DT integration as a crucial pathway toward predictive maintenance and preventive conservation. These studies underscore the research interest shift from geometric documentation toward intelligent, interoperable heritage information systems.

The reuse of historic structures places particular demands on digital documentation, as it normally needs an initial building survey to help stakeholders understand the overall structure and the building's as-builtconditions (Khalil and Stravoravdis, 2022). In addition, having digital models, especially BIM models, can accelerate design proposals and streamline construction progress. Adaptive reuse projects can face compressed timelines, constrained budgets, and site limitations that make conventional building survey approaches impractical. Moreover, heritage sites present additional layers of significance: beyond their material fabric, they embody cultural, social, and historical narratives that must be studied and safeguarded in the renovation process. In this context, there is a growing demand for utilizing different reality capture tools survey and record for historic structures. In addition, historic building practitioners often face various challenges in project management, financial loss, and project delay during the restoration process due to the lack of digital documentation and updated digital models. Traditionally, information about historic structures has been fragmented across different formats and managed independently by different specialists. HBIM offers a unified digital environment that can overcome these long-standing challenges by improving interoperability, information integration, and interdisciplinary collaboration (Khan et al., 2022). The alignment of HBIM with the early stages of DT development further underscores the importance of scalable digital pipelines that support not only documentation but also long-term stewardship, stakeholder engagement, and immersive applications (Cassar et al., 2025). Despite this growing body of research, relatively few studies documented the combination of multiple capture devices under strict time and budget constraints in adaptive reuse projects. The presented case study addresses this gap by testing a rapid, multi-device HBIM workflow that is calibrated to the availability of limited resources while remaining sensitive to both critical heritage values and stakeholder needs.

This paper responds to these challenges by proposing and testing a multi-device HBIM workflow for adaptive reuse. Specifically, it investigates how TLS, Matterport scanning, and UAV photogrammetry can be strategically combined to overcome site access limitations, mitigate data saturation, and deliver integrated outputs within severe time and funding constraints. The approach is evaluated through the case study of the Odd Fellows Building, a culturally significant landmark in Atlanta's Sweet Auburn Historic District. Once a hub of African American professional and social life, the structure now faces both physical deterioration and the opportunity of renewed community service through its adaptive reuse project. By applying a multi-device digital strategy, this study demonstrates how reality capture

TABLE 1 Correlation between project goals and suitable reality capture tools.

Project goals	Point cloud	Virtual tour	High-accuracy 3D information	High-resolution visual data	High-accuracy geo-reference
TLS	✓		✓		
UAVs	✓		~	✓	~
CRC	✓			✓	
Matterport	✓	✓		✓	
MHS	✓		✓		
TS	✓				✓

✓, Applicable; ~, May applicable.

technologies can be orchestrated to produce a semantically rich HBIM model at Level of Development (LoD) 300, serving immediate renovation needs while laying a foundation for future digital twinning.

While previous studies combined different reality capture devices for built heritage documentation, the innovation of this work lies in developing a significance-driven and resource-constrained multi-device HBIM workflow. The proposed framework prioritizes reality capture devices based on the cultural and historic importance of building components, balancing accuracy and efficiency according to project significance and accessibility. This study is also distinguished by its real-world application of an adaptive-reuse project under a 1-day scanning and 1-week modeling limit, demonstrating a pragmatic model for limited-budget preservation project. The contribution therefore extends beyond technical integration to provide a decision-making framework that can be adaptable to other built heritage facing similar financial and logistical limitations. Through this contribution, the paper situates HBIM not merely as a technical tool but as an integral part of the broader digital transformation in construction. It highlights the potential of hybrid capture workflows to support both project-specific requirements and long-term cultural heritage preservation, bridging the gap between rapid documentation and comprehensive digital integration.

2 Materials and methods

Creating a HBIM model through reality capture has been demanded urgently to assist stakeholders' renovation and demolition proposal, such as producing two-dimensional drawings and renderings. Meanwhile, under short notice and limited working-time windows, applying reality capture tools through a scan-to-BIM process significantly reduces on-site time and minimizes human error compared to manual surveys (Rocha et al., 2020). From long-term preservation perspective, establishing a digital model would also be beneficial to keep archive integrity and assist future renovation and maintenance needs. In order to achieve HBIM model creation goal, selecting appropriate reality capture tool to use becomes challenging. Different reality capture tools have different advantages and disadvantages. The selection of reality capture tools should depend on project goals, site limitation, and resource availability to make a balanced decision.

As the ultimate product goal of creating HBIM model, the reality capture devices that can produce point cloud or assist HBIM process are included in this discussion: TLS, UAVs, close-range camera (CRC), Matterport, mobile handheld scanner (MHS), and total station (TS). To align with the balanced decision goal, those devices are categorized by project goals and site limitations to demonstrate their feasibility under different scenarios.

The project goal represents diverse use cases and different LoD for the final product that the involved stakeholders try to achieve. Table 1 summarizes how different reality-capture devices align with specific project goals, including point-cloud generation, high-resolution visuals, and geo-referencing accuracy. The specialty of Matterport offering a virtual tour is valuable, which can provide remote access to stakeholders without physical geolocation restrictions. Meanwhile, it provides high-resolution visual data so that any people can easily refresh memories and use it as a visual reference. UAVs and CRC are photography capture tools that can assist in reexamining visual data. Besides, their captured photos can be used for 3D object reconstruction via photogrammetry process. In order to achieve delicate 3D information capture, scanner-based devices, such as TLS, MHS, and TS, are proven to have higher accuracy (El-Din Fawzy, 2019). TS with the feature of attaining accurate geo-reference, is usually utilized as a ground control reference that supports the scan alignment of TLS or drone photogrammetry alignment, further calibrating point cloud data. Except for conventional camera-based drones, currently, LiDAR sensor-mounted drones are in practice for archaeological excavation sites (Adamopoulos et al., 2023), city models (Setyawan et al., 2022), etc. With ideal weather conditions, the LiDAR drone can capture accurate and reliable information, but it also enormously increases the cost. Considering current UAVs market development, UAVs are still largely camera-based drones which can take photographs and assist with photogrammetry techniques to generate point cloud. Even photogrammetry techniques generally have lower tolerance than scanner-based devices, it provides great flexibility and affordable cost. In addition, UAVs' data can be upgraded by real-time kinematic (RTK) techniques and achieve much higher and centimeter-level UAVs positioning accuracy in real-time (Czyża et al., 2023), which further support photogrammetry accuracy. Different reality capture tools should be selected or integrated to tackle the aims of project.

Besides different project goals, site limitations can also restrict reality capture tools. Table 2 presents how site constraints, such as interior and exterior conditions, lighting, and accessibility, affect the

TABLE 2 Site condition limitations and applicability of reality capture tools.

Site condition limitations	Interior	Exterior	Hard-access area	Narrow area	Resistance to light	Short time window
TLS	✓	✓			~	
UAVs	~	✓	✓	~		✓
CRC	✓	✓				
Matterport	✓	✓			~	✓
MHS	✓	✓		✓	~	
TS	✓	✓			~	

✓, Applicable; ~, May applicable.

feasibility of using each tool. Conducting a preliminary on-site investigation is necessary for identifying data capture challenges such as physical access, vegetation, or lighting conditions that can dramatically impact the quality of the data capture process (Fobiri et al., 2022; Willkens, 2022). Most of the tools can be used for interior capture. Scanner-based devices using laser beams can be conducted in a dark environment and still extract good-quality 3D information. Camera-based devices usually need external light sources to collect texture information. Restricted by indoor space, flying a drone in interior space can be a challenge. UAVs for indoor utilization are being studied so far; for example, micro drones have been tested for interior mapping and demonstrated sufficient quality for exploration purposes (Karam et al., 2022). All those tools can be used for exterior capture but might be limited by location and environmental features. UAVs are usually used for hard-access area capture and are suitable for congested areas. MHS has relatively smaller sizes, which usually can be used to scan narrow areas as well but limited to human hand-reachable areas. Scanner-based devices can resist dim light situations when collecting object 3D information, but they often get impacted on collecting material texture and color information on dim environment. Camera-based devices are easily affected by lighting situations. The newest version of Matterport Pro 3 has a LiDAR sensor embedded, which makes it more capable of diverse lighting situations than the last version of Matterport. Previous research (Liu et al., 2023) revealed that a Matterport-captured point cloud would have slightly lower quality and accuracy than a TLS-captured, and Matterport also collected more noise data. However, it is still a sufficient method for assisting with HBIM model-making. Moreover, the allowable time on site is a critical factor to consider. For instance, scanning tourist areas requires avoiding crowds in the daytime, which needs a dedicated scanning plan with proper reality capture tools. Drones and Matterport are deemed as fast scanning devices since drones have flight path and automatic execution features on mapping tasks, and Matterport can do quick 360-degree photography captures. TLS can also achieve quick capture with low-quality and low-resolution parameters setting, while its essential advantage is high-quality and high-resolution capture which can take much longer time to scan. In general, evaluating the site situation is an essential step before comprising a detailed scanning plan. Table 1 and 2 allows users to weigh trade-offs among precision, accessibility, environmental conditions, resource availability, etc. By adjusting the relative weight of each constraint, users can replicate or modify the workflow across diverse built heritage scenarios, from archaeological sites to large-scale urban regions.

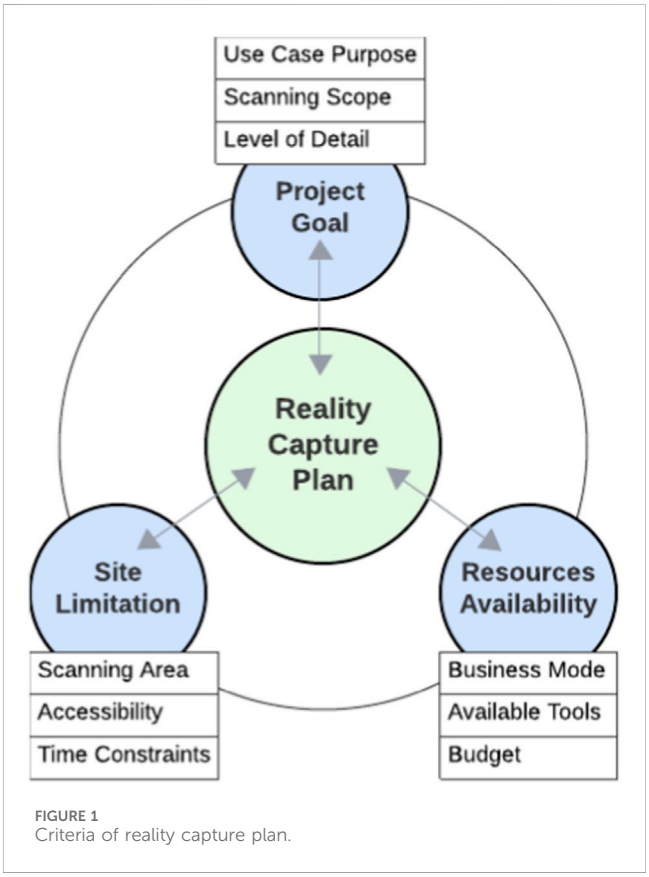


FIGURE 1
Criteria of reality capture plan.

Resource availability varies for different projects and stakeholders. Conducting a reality capture task can be burdensome on financial cost, including costs on equipment, processing software, computer station, and reality capture experts. Setting up an achievable goal with available funding resources is the key to the project's success. In most scenarios, singular reality capture device to finish the whole project scanning work is challenging. Different equipment and software can be combined to finish one project together with good coverage. For

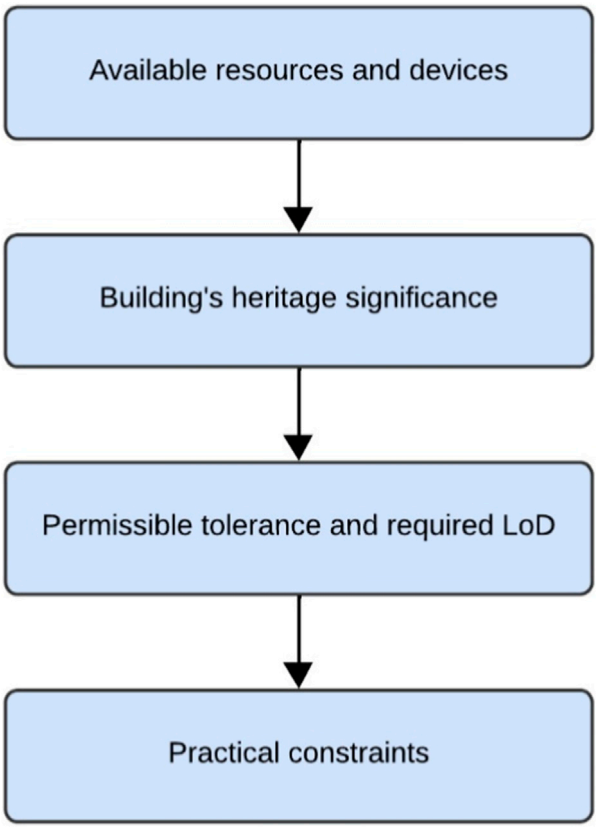


FIGURE 2
Decision flowchart with four core factors in sequence.

example, TLS collects data from the ground level, while UAV gathers data from an aerial view as supplementary information. TLS requires the scan processing software to align all scan positions into one coordinate system, while aerial photos need photogrammetry software to process images and reconstruct the object's 3D information. As the article mentioned, each reality capture tool has pros and cons and diverse date quality parameters. Usually, the higher resolution data would require more expensive tools and sometimes cost more in terms of scanning time and processing stations. The tool selection should be based on project goals and resource availability to determine.

The project goal, site limitation, and resource availability are essential factors in determining the optimal reality capture plan. At the initial stage of the project, stakeholders meeting is critical to enumerate all requirements as three essential factors can be impacted by each other (Figure 1). In addition, the site situation should be informed in order to come up with a scanning strategy that ensures good coverage rate of scan. When the situation allows, reality capture experts should inspect the site before executing the task to identify any unapparent challenges from the expert's perspective. To guide practitioners make decisions on capture strategy, Figure 2 introduces a transferable decision flowchart with four core factors in sequence that balances project goals, site constraints, and resource availability, as well as devices capabilities in Tables 1 and 2. Rather than selecting devices solely based on technical specifications, this framework evaluates four core factors in sequence: (1) Available resources and devices; (2) Heritage significance of building component; (3) Permissible tolerance and required LoD; (4) Practical constraints such as budget, time, and site accessibility. The practitioners can evaluate these four core factors in sequence when facing competing constraints. This yields a structured and repeatable method for prioritizing devices under real case conditions. For example, when heritage significance is high and the project requires geometric accuracy with low tolerance, the practitioners should prioritize TLS capture for building elements and area with higher heritage value, even under strict time or budget constraints. The rest of building space and elements capture can shift toward Matterport scanning to achieve rapid capture or shift toward UAVs photogrammetry to overcome labor safety and accessibility constraints. When the project can only get Matterport to use and heritage significance is high, applying high quality capture mode of Matterport on major important area while applying standard quality capture in transition area between major essential space. For adaptive reuse project, prioritizing TLS capture for heritage significance elements and desirable preserving area. Conversely, the space, which is going to be altered heavily and requires LoD is in medium level, can use Matterport scanning served as visual reference record and fast survey device without compromising rehabilitation planning. These scenarios demonstrate how the decision framework can be adapted with competing constraints and provide a transferable methodology that practitioners can apply to varied built heritage project contexts.

3 Results—case study

3.1 Building background

The Odd Fellows Building, located in the Sweet Auburn Historic District of Atlanta, Georgia, is selected as a case study to demonstrate and verify the HBIM reality capture plan methodology. As an established leader in the African-American community, Booker T. Washington (1856–1915), dedicated the Odd Fellows Building when it was dedicated in 1912. Initially, the building had six stories above ground and one underground floor. Over time, a penthouse floor was added. The building façade was made of red brick except for the first floor, which was made of stone. Amid the Jim Crow era, it was a vital landmark for the Atlanta community, providing office and retail space for Black professionals and Black-owned businesses, while also hosting myriad social activities and cooperative events. With its legacy in supporting the African-American community, Odd Fellows Building was added to the National Register of Historic Places in 1975. However, due to a lack of maintenance and tornado damage in 2008, the building was damaged, and its state severely hindered its legacy and ability to support community needs. Fortunately, with the new adaptive reuse proposal led by a non-profit, Georgia Works, the building will transform into a multi-use complex with office space, discounted retail space, and a transitional housing program to reduce recidivism for Georgia Works program participants. Through these rehabilitations, the building will once again serve

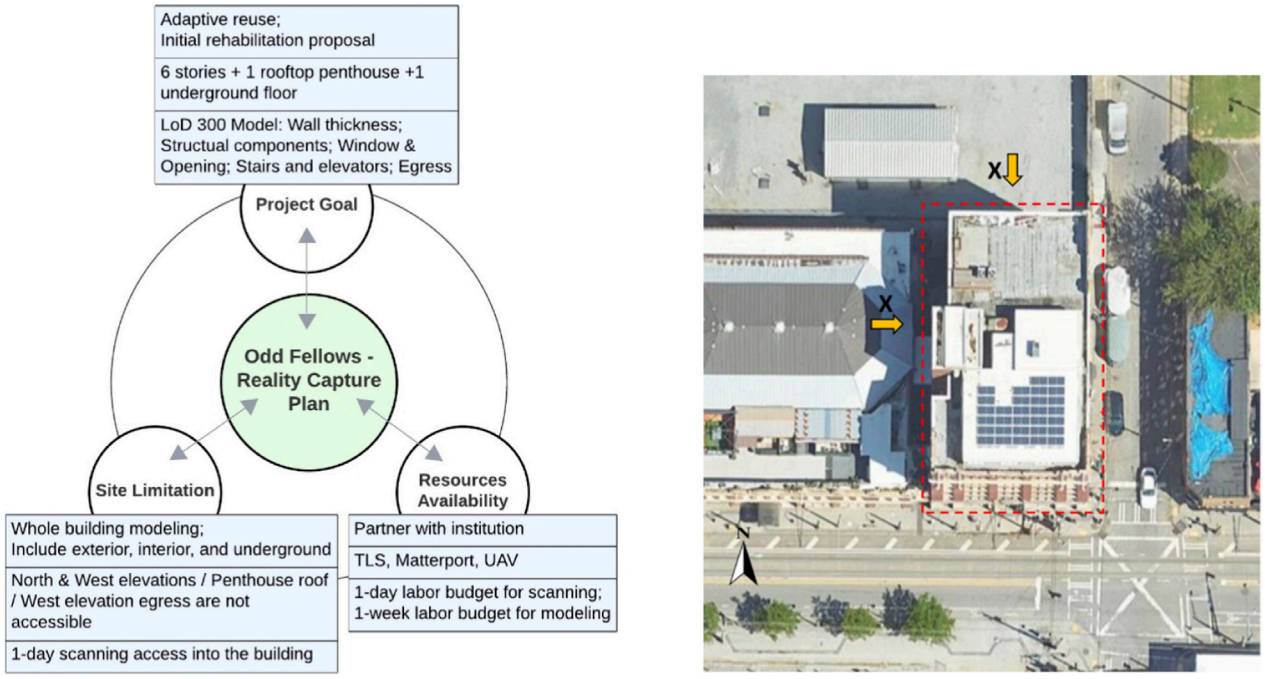


FIGURE 3
(Left) considerations for Odd Fellows Building; (right) site access limitation.

as a significant mutual aid function, while conserving the distinctive brick and terracotta of this architectural landmark.

3.2 Reality capture plan

Before handling reality capture tasks and starting digital modeling, examining the project's circumstances is needed as following the HBIM reality capture plan methodology (Figure 3). The overall goal of adaptive reuse is evident. However, when funding is limited at the initial stage, the project steward aims to obtain an initial digital model to support grant applications, indicating that the budget for scanning and modeling is restricted. Having only a 1-day scanning period and a 1-week modeling budget for this eight-story building makes the selection of suitable reality capture tools challenging. Besides, due to the site limitation of inaccessibility on the North and West sides and rooftop of penthouse, multi-devices have to be integrated to collect complete information. Both TLS and Matterport devices need physically move and set up on every scan location which limits its capture position within a physically reachable area. UAVs can go beyond physical limitation within open air space but becomes risk in congested or indoor space. From labor safety consideration, UAVs can reach out to high elevation and hazard area where historic buildings often have this type of concerns for unstable structure (Treccani et al., 2024). Combining those considerations on project goals, site limitation, and resources availability, the steward determined to collaborate with the institution, which is more accessible and affordable solution and set the desired LoD for the model is 300. Collaborating with a university research institution provided both economic and technical advantages. Unlike professional survey firms that charge commercial rates for fieldwork and modeling, academic collaborations operate under research or educational frameworks that leverage institutional equipment, software licenses, and trained personnel. This project collaborated with the School of Architecture, Georgia Institute of Technology, which possesses advanced reality capture tools, including Faro TLS, Matterport Pro 3, and UAVs. Moreover, faculties and graduate researchers possess pioneer experience with heritage documentation methodologies, data registration, and HBIM workflows, ensuring high-quality outcomes despite limited budgets. This arrangement not only reduces financial expenses but also fosters knowledge exchange between practitioners and the academic team, aligning with the broader educational mission of training future professionals.

After a comprehensive evaluation, the scanning plan and strategy of using different tools are developed to achieve the general goal (Figure 4). With short time window circumstances, although TLS is deemed the most accurate survey tool (Lachat et al., 2017), it can only be utilized for a small portion since it will consume much longer on-site time than others. Faro Focus S350, as one of the high-end TLS, is applied to capture mainly exterior data and the first-floor area, whose point cloud serves as a baseline for other point clouds to align. Matterport Pro 3 embedded LiDAR sensor is a suitable tool for executing interior scans, as it can operate fast and provide virtual tour as supplementary visual data. Visual data can assist BIM modelers for digital modeling and serve as a photography record as time goes by. The Matterport is set to scan all eight floors, including the basement and rooftop penthouse. Due to the

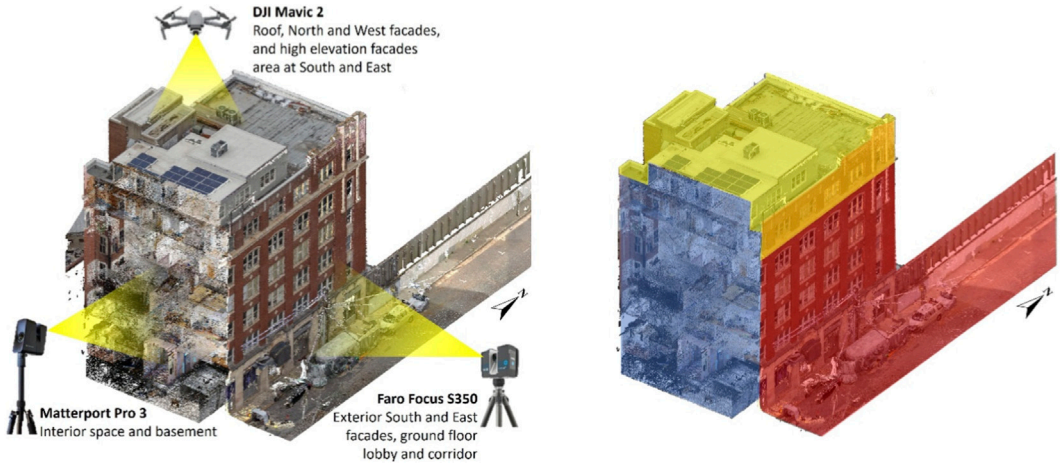


FIGURE 4
(Left) Odd Fellows Building scan plan strategy; (right) color-coded device coverage diagram showing scanning contributions from faro focus S350 (red), matterport pro 3 (blue), and DJI mavic 2 (yellow).

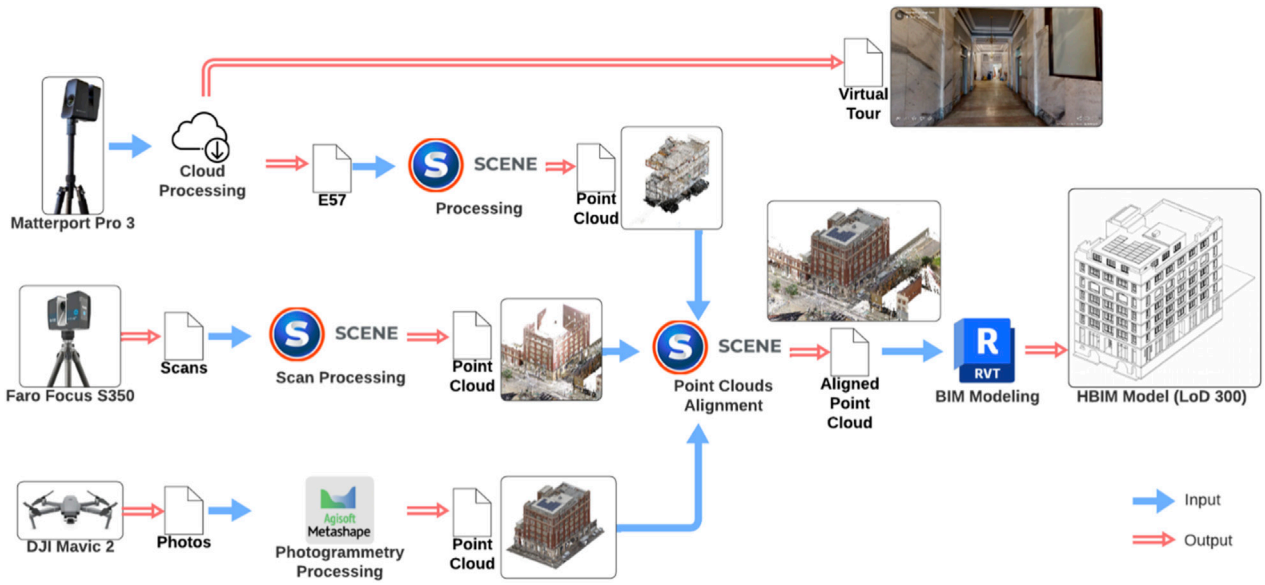


FIGURE 5
Date capture-process-modeling workflow.

inaccessibility of the North and West facades, UAV, one of the great mobile robots, can easily reach out to areas that human labor cannot access. DJI Mavic 2 takes aerial images to compensate for deficient capture by other tools.

3.3 Data capture-process-modeling workflow

The overall workflow from capture to processing and modeling is summarized in Figure 5. Aligning the digital documentation strategy with the building's heritage values, TLS was used to document the lobby corridor and two essential facades (South and East) retained original materials and design intent from the building's period of significance, as defined in its National Register listing. A total of 14 scan positions were captured by TLS. Matterport scanning was utilized for the renovated upper floors and other non-original areas, including interior space of basement level and upper seven stories where rapid capture and visualization were sufficient for rehabilitation planning but detailed heritage documentation was less critical. Total 205 scan positions were captured by Matterport Pro 3. UAV serve as supplementary reality capture device that collect aerial photos for inaccessibility places of high elevation area and North and West facades. A total of 483 aerial s were collected by DJI Mavic 2. Matterport technology has cloud processing systems that do not require external processing

TABLE 3 Quantitative summary of each device's capture setting and results.

Device	Faro Focus S350	
Capture area	Exterior	Interior
Number of scan positions	9	5
Scan resolution	1/5	1/8
Scan quality	3X	2X
Scan time/Position	2 min 54 s	1 min 3 s
Scan with color	Yes	Yes
Scan size	8,284 × 3,415 [pt]	5,174 × 2,134 [pt]
Registration report		
Maximum point error	5.5 mm	
Mean point error	2.6 mm	
Device	Matterport Pro 3	
Capture area	Interior	
Number of scan positions	205	
Scan setting	Standard	
Depth points/position	~0.8 M	
Scan time/position	<20 s	
Device	DJI Mavic 2	
Number of images	483	
Capture area	Exterior	
Flying altitude	68.4 ft	
Photogrammetry processing	Agisoft metashape	
Tie points	185,334	
RMS reprojection error	0.311763 (1.21737 pix)	
Max reprojection error	0.953245 (48.7032 pix)	
Total points	64,050,026	

stations. After a Matterport capture, the captured scans can automatically be uploaded to a cloud computing system to generate one integrated virtual tour model directly in their platform. The scan setting, registration accuracy and point density of each capture devices were summarized in Table 3. The TLS dataset achieved a mean point error of 2.6 mm, confirming its role as the prioritized geometric baseline. Since Matterport's point cloud are pre-aligned in a proprietary coordinate system without ground control referencing. Aligning these data required iterative transformations in Faro Scene and the use of shared reference features such as floor planes and doorframes to reduce geometric drift. Scanning time of different devices further highlights the efficiency of utilizing the combined strategy to balance accuracy and time limitation.

Toward the project goal of having a point cloud for assisting HBIM model making, the virtual tour data is converted as an E57 file, a vendor-neutral format containing 3D point cloud information. After registering in the Faro Scene, the Matterport scanned data serves as an interior point cloud. The data collected from Faro Focus S350 is registered and integrated by Faro Scene to produce a point cloud that will be used as a baseline reference. Aerial images collected by DJI Mavic 2 are processed by photogrammetry techniques in Agisoft Metashape software, whose point cloud is utilized as supplementary information to incomplete scan area by TLS and Matterport. The three separate point clouds have their individual coordinate systems from TLS, Matterport, and UAV (Figures 6a–c), which can cause disorder in the BIM modeling procedure if they have not been aligned. Faro Scene can assist in finding overlapping areas and combining different point clouds. In the HBIM environment, these multi-source point clouds require careful preprocessing to ensure geometric and semantic consistency. As the point cloud from TLS was deemed the accurate baseline, the interior's point cloud from Matterport was majorly aligned with points in first floor area from TLS, while the point cloud from UAV capture was mainly aligned with points in TLS's exterior area.

Once the multi-source point clouds aligned (Figure 6d), the integrated dataset was imported into Autodesk ReCap Pro for final filtering and down sampling to optimize file size and maintain manageable density for HBIM modeling. The registered point cloud was then linked into Autodesk Revit, where model development followed a hierarchical, element-based approach consistent with HBIM best practices. Major components, including structural columns, floor slabs, walls, windows, and staircases, were modeled as Revit families using the Architecture and Structure toolsets. To maintain geometric fidelity, elements were traced directly over the point cloud. Material properties were documented through parameterized attributes and labeled according to the LoD 300 definition, which ensures accurate geometric representation and basic non-geometric information (Figure 7).

4 Discussion and conclusion

4.1 Discussion

The case study of the Odd Fellows Building demonstrates the practical advantages and constraints of implementing a multi-device HBIM workflow in adaptive reuse projects. By integrating TLS, Matterport, and UAV technologies, the project achieved comprehensive coverage of interior and exterior conditions under severe time and budget constraints. This hybrid strategy illustrates a valuable trade-off: TLS provided highly accurate baseline data, Matterport delivered rapid interior scanning with the added benefit of immersive visualization, and UAVs captured inaccessible areas that would otherwise be excluded from conventional surveys. The resulting dataset, once registered and unified, offered stakeholders both the precision required for design decisions and the accessibility needed for communication and funding efforts.

These findings underscore the potential of multi-device workflows to address persistent challenges in heritage documentation. Previous studies have often emphasized either the precision of TLS or the accessibility of photogrammetry and low-cost scanning technologies. This case instead demonstrates the value of strategically combining devices to balance accuracy, speed,

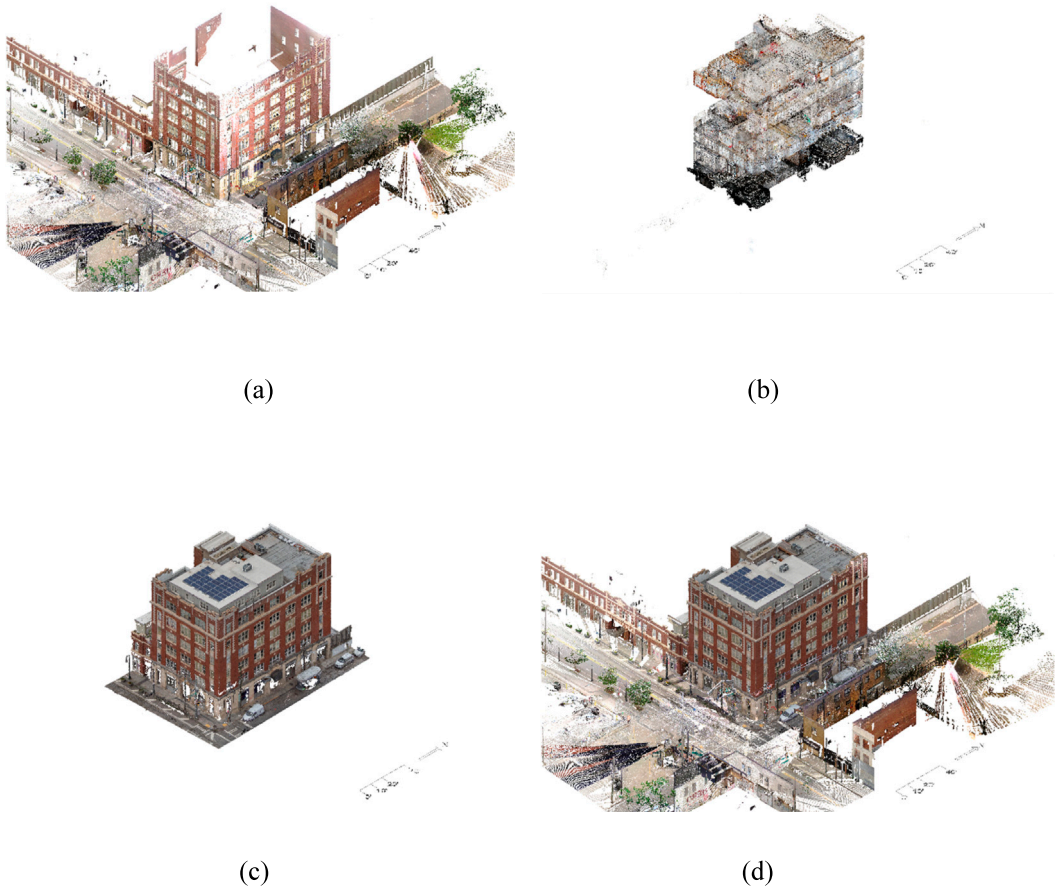


FIGURE 6
Integrated point cloud datasets captured by TLS, matterport, and UAV. **(a)** Point cloud from TLS. Capture area: South and East facades and ground floor lobby and corridor. Registration accuracy: 2.6 mm mean point error. **(b)** Point Cloud from Matterport Pro 3. Capture area: interior floors above ground and basement level. Depth points: ~0.8 M per position. **(c)** Point Cloud from UAV photogrammetry. Capture area: inaccessible roof, North and West facades, and high elevation facades area at South and East. RMS reprojection error: 0.311763 (1.21737 pix). **(d)** Aligned Point Clouds. The alignment among datasets is achieved through overlapping ground floor lobby and corridor, and façade features. Manual marked reference points are used to assist software alignment.



FIGURE 7
(Left) HBIM model with LoD 300; (right) HBIM model with point clouds on.

and resource limitations. The Atlanta case study successfully achieved comprehensive coverage of both exterior and interior spaces within the allotted 1-day scanning and 1-week modeling timeframe, delivering a unified LoD 300 HBIM model. Despite limited resources, this outcome met all project goals for documentation completeness, LoD requirement, and assisting stakeholder communication on initial rehabilitation proposal. The workflow also highlights the evolving role of HBIM as a bridge toward DT applications: while the resulting model does not achieve real-time monitoring, it establishes the foundational infrastructure necessary for future data integration and immersive exploration.

Nevertheless, the study also points to important limitations. The reliance on multiple devices requires expertise in data registration and integration, which may present barriers for projects without institutional or technical support. The integration of heterogeneous datasets also revealed significant HBIM challenges. Aligning point clouds from TLS, UAV, and Matterport required a robust data management protocol to avoid geometric distortion and ensure interoperability across software platforms, such as Faro Scene, Matterport, Agisoft Metashape, and Autodesk Revit. The accuracy assessment of aligned point cloud was primarily based on visual alignment and sample measurements. A comprehensive quantitative validation using standardized error metrics and control targets would strengthen its reliability. Another major constraint involves file size and computing resource management. Integrating multiple high density point clouds often produce large datasets, requiring progressive down sampling and segmentation to maintain Revit performance without compromising accuracy. This challenge highlights the importance of establishing standardized data resolution thresholds for HBIM projects. Besides, the HBIM model was developed to LoD 300, which is suitable for design coordination at this stage but not for detailed conservation analysis. Historically significant elements such as decorative brickwork and terra cotta would require further detail modeling and semantic data enrichment to assist preservation work. Furthermore, the workflow relies on commercial platforms such as Matterport's hosting service, raising questions about licensing costs, data interoperability, and long-term accessibility in heritage contexts. Future research should pursue higher LoD and semantically enriched models, which can support DT integration to enable dynamic monitoring and preventive maintenance applications. By incorporating the HBIM model with sensor data, environmental monitoring, and maintenance logs, users can transition from static representation to a dynamic performance model capable of real-time tracking and predictive analysis, advancing from documentation toward preventive conservation and life-cycle management of historic assets.

4.2 Conclusion

This paper presented a multi-device HBIM workflow tested through the adaptive reuse project of the Odd Fellows Building in Atlanta. By leveraging TLS, Matterport, and UAV technologies, the project developed a rapid, resource-conscious, and semantically rich HBIM model at LoD 300. The workflow demonstrates how hybrid capture strategies can overcome site access limitations, accelerate documentation, and support both stakeholder communication and renovation planning.

Importantly, decisions about which technologies to use were guided by considerations of authenticity and integrity, aligning the digital documentation strategy with the building's heritage values. Since the interior had undergone extensive alterations, only the lobby corridor and two primary facades (south and east) retained original materials and design intent from the building's period of significance, as defined in its National Register listing. These areas were documented with TLS to ensure the highest geometric accuracy and tolerance. In contrast, Matterport scanning was prioritized for the renovated upper floors and other non-original areas, where rapid capture and visualization were sufficient for rehabilitation planning, but detailed heritage documentation was less critical. This selective approach exemplifies how HBIM workflows can integrate both technical precision and preservation ethics, balancing the documentation of historically significant elements with pragmatic resource allocation.

Beyond the immediate project outcomes, the study contributes to the broader discourse on digital transformation in construction by positioning HBIM as a scalable framework that integrates diverse capture technologies while remaining sensitive to the cultural and material authenticity of historic buildings. The Odd Fellows case underscores the importance of balancing accuracy, efficiency, and significance-based decision-making in heritage contexts where time, funding, and access are constrained.

Future research should focus on refining interoperability between commercial and open-source platforms, exploring automation in point cloud registration, and testing the scalability of significance-driven multi-device workflows across different heritage typologies. By advancing such approaches, HBIM can more fully realize its role in preserving cultural heritage with integrity while supporting the digital transformation of the construction sector.

Data availability statement

The datasets presented in this article are not readily available because the data supporting this study's findings are available upon request from the corresponding authors. Due to the sensitive nature of the data and their file size, access to the data will be granted on a case-by-case basis and may require a data use agreement to be signed. Requests to access the datasets should be directed to DW, danielle.willkens@design.gatech.edu.

Author contributions

BL: Methodology, Visualization, Data curation, Software, Writing – original draft, Conceptualization, Writing – review and editing. DW: Data curation, Conceptualization, Supervision, Methodology, Project administration, Writing – review and editing, Software.

Funding

The author(s) declared that financial support was not received for this work and/or its publication.

Conflict of interest

The author(s) declared that this work was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declared that generative AI was not used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Adamopoulos, E., Papadopoulou, E. E., Mpia, M., Deligianni, E. O., Papadopoulou, G., Athanasoulis, D., et al. (2023). 3D survey and monitoring of ongoing archaeological excavations via terrestrial and drone lidar. *ISPRS Ann. Photogrammetry Remote Sens. Spatial Inf. Sci.* X-M-1-2023, 3–10. doi:10.5194/isprs-annals-x-m-1-2023-3-2023

Alshawabkeh, Y., and Baik, A. (2023). Integration of photogrammetry and laser scanning for enhancing scan-to-HBIM modeling of Al Ula heritage site. *Herit. Sci.* 11 (1), 147. doi:10.1186/s40494-023-00997-2

Cassar, A., Baker, D., and Ioannides, M. (2025). From digital twin to memory twin: a holistic framework for cultural heritage documentation, interpretation, and adaptive reuse. *Int. Arch. Photogrammetry Remote Sens. Spatial Inf. Sci.* 203–209. doi:10.5194/isprs-archives-xlviii-m-9-2025-203-2025

Czyża, S., Szuniewicz, K., Kowalczyk, K., Dumalski, A., Ogrodniczak, M., and Zieleniewicz, Ł. (2023). Assessment of accuracy in unmanned aerial vehicle (UAV) pose estimation with the real-time kinematic (RTK) method on the example of DJI matrice 300 RTK. *Sensors* 23 (4), 2092. doi:10.3390/s23042092

Dore, C., and Murphy, M. (2017). Current state of the art historic building information modelling. *Int. Arch. Photogrammetry Remote Sens. Spatial Inf. Sci.* 185–192. doi:10.5194/isprs-archives-xlii-2-w5-185-2017

El-Din Fawzy, H. (2019). 3D laser scanning and close-range photogrammetry for buildings documentation: a hybrid technique towards a better accuracy. *Alexandria Eng. J.* 58 (4), 1191–1204. doi:10.1016/j.aej.2019.10.003

Fobiri, G., Musonda, I., and Muleya, F. (2022). Reality capture in construction project management: a review of opportunities and challenges. *Buildings* 12 (9), 1381. doi:10.3390/buildings12091381

Giuliani, F., Gaglio, F., Martino, M., and De Falco, A. (2024). A HBIM pipeline for the conservation of large-scale architectural heritage: the city Walls of Pisa. *Herit. Sci.* 12 (1), 35. doi:10.1186/s40494-024-01141-4

Inzerillo, L., Pisciotta, A., Acuto, F., Mantalovas, K., and Di Mino, G. (2025). Beyond traditional H-BIM: AI-powered modeling for heritage bridges. *Int. Arch. Photogrammetry Remote Sens. Spatial Inf. Sci.* 615–622. doi:10.5194/isprs-archives-xlviii-m-9-2025-615-2025

Karam, S., Nex, F., Karlsson, O., Rydell, J., Bilock, E., Tulldahl, M., et al. (2022). Micro and macro quadcopter drones for indoor mapping to support disaster management. *ISPRS Ann. Photogrammetry Remote Sensing Spatial Inform. Sci.* V-1-2022, 203–210. doi:10.5194/isprs-annals-v-1-2022-203-2022

Khalil, A., and Stravoravdis, S. (2022). Challenges of digital building data usage with a focus on the digital documentation of heritage buildings—results from an online survey. *Heritage* 5 (4), 3220–3259. doi:10.3390/heritage5040166

Khan, M., Khan, M., Bughio, M., Talpur, B., Kim, I., and Seo, J. (2022). An integrated HBIM framework for the management of heritage buildings. *Buildings* 12 (7), 964. doi:10.3390/buildings12070964

Lachat, E., Landes, T., and Grussenmeyer, P. (2017). Investigation of a combined surveying and scanning device: the trimble SX10 scanning total station. *Sensors* 17 (4), 730. doi:10.3390/s17040730

Liu, J., Willkens, D., López, C., Cortés-Meseguer, L., García-Valldecabres, J. L., Escudero, P. A., et al. (2023). Comparative analysis of point clouds acquired from a TLS survey and a 3d virtual tour for Hbim development. *Int. Arch. Photogrammetry Remote Sens. Spatial Inf. Sci.* 959–968. doi:10.5194/isprs-archives-xlviii-m-2-2023-959-2023

Lucchi, E. (2023). Digital twins for the automation of the heritage construction sector. *Automation Constr.* 156, 105073. doi:10.1016/j.autcon.2023.105073

Niccolucci, F., and Felicetti, A. (2024). Digital Twin sensors in cultural heritage ontology applications. *Sensors* 24 (12), 3978. doi:10.3390/s24123978

Parente, M., Bruno, N., and Ottoni, F. (2025). HBIM and information management for knowledge and conservation of architectural heritage: a review. *Heritage* 8, 306. doi:10.3390/heritage8080306

Rocha, G., Mateus, L., and Fernández, J. (2020). A Scan-to-BIM methodology applied to heritage buildings. *Heritage* 3, 47–67. doi:10.3390/heritage3010004

Setyawan, A. A., Taftazani, M. I., Bahri, S., Noviana, E. D., and Faridatunnisa, M. (2022). Drone LiDAR application for 3D city model. *J. Appl. Geospatial Inf.* 6 (1), 572–576. doi:10.30871/jagi.v6i1.3462

Treccani, D., Adami, A., Fregonese, L., Treccani, D., Adami, A., and Fregonese, L. (2024). Drones and real-time kinematic base station integration for documenting inaccessible ruins: a case study approach. *Drones* 8 (6), 268. doi:10.3390/drones8060268

Willkens, D. S. (2022). The craft and care of reality capture. *Build. Landscapes J. Vernac. Archit. Forum* 29 (2), 2–14. doi:10.1353/bdl.2022.0009

Yang, X., Grussenmeyer, P., Koehl, M., Macher, H., Murtiyoso, A., and Landes, T. (2020). Review of built heritage modelling: integration of HBIM and other information techniques. *J. Cult. Herit.* 46, 350–360. doi:10.1016/j.culher.2020.05.008

Yu, Y., Raed, A. A., Peng, Y., Pottgiesser, U., Verbree, E., and Van Oosterom, P. (2025). How digital technologies have been applied for architectural heritage risk management: a systemic literature review from 2014 to 2024. *Npj Herit. Sci.* 13 (1), 45. doi:10.1038/s40494-025-01558-5

Zhang, Z., and Zou, Y. (2022). Research hotspots and trends in heritage building information modeling: a review based on CiteSpace analysis. *Humanit. Soc. Sci. Commun.* 9 (1), 394. doi:10.1057/s41599-022-01414-y

frontiers | Frontiers in Built Environment

TYPE Original Research
PUBLISHED 14 January 2026
DOI 10.3389/fbuil.2025.1715347

Check for updates

OPEN ACCESS

EDITED BY
Izuru Takewaki,
Kyoto Arts and Crafts University, Japan

REVIEWED BY
Rizwan Farooqui,
Mississippi State University, United States
Roy Lan,
University of Texas at San Antonio, United States
Rana Muhammad Irfan Anwar,
Auburn University, United States

*CORRESPONDENCE
Amanda Ramón-Constantí,
amracon@ega.upv.es

RECEIVED 29 September 2025
REVISED 09 December 2025
ACCEPTED 23 December 2025
PUBLISHED 14 January 2026

CITATION
Ramón-Constantí A, Castilla Pascual FJ, Garcia de Soto B and Adán-Oliver A (2026) Integrating thermal point clouds into BIM-GIS environments: workflow proposal for multi-layer digital twins.
Front. Built Environ. 11:1715347.
doi: 10.3389/fbuil.2025.1715347

COPYRIGHT
© 2026 Ramón-Constantí, Castilla Pascual, Garcia de Soto and Adán-Oliver. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Integrating thermal point clouds into BIM-GIS environments: workflow proposal for multi-layer digital twins

Amanda Ramón-Constantí[1]*, Francisco Javier Castilla Pascual[2], Borja Garcia de Soto[3] and Antonio Adán-Oliver[4]

[1]Departamento de Expresión Gráfica Arquitectónica, Universitat Politècnica de València, Valencia, Spain, [2]Universidad de Castilla-La Mancha, MAEE-UCLM Group, Albacete, Spain, [3]S.M.A.R.T. Construction Research Group, Division of Engineering, New York University Abu Dhabi, Abu Dhabi, United Arab Emirates, [4]Universidad de Castilla-La Mancha, 3D Visual Computing & Robotics Group, Ciudad Real, Spain

Introduction: The integration of Thermal Point Clouds (TPCs) into professional Building Information Models (BIM) and Geographic Information System (GIS) workflows is currently hampered by a lack of established methodologies and significant format interoperability challenges. This study addresses this methodological gap by developing and testing integration processes for both BIM and GIS platforms.
Methods: Two proofs of concept were developed using data from a commercial scanner at the New York University Abu Dhabi (NYUAD) campus. For BIM integration, a process was designed to generate open Industry Foundation Classes (IFC) files where average surface temperatures are embedded as native properties of architectural elements. For GIS integration, thermal data was assigned as custom attributes to a manually generated 3D geometric reference model, establishing the preliminary steps for a dedicated thermal-GIS workflow.
Results: The methodologies were successfully validated through visualization in ArcGIS Pro and ACCA Software GeoTwin. The results demonstrate a tangible path to overcoming current format limitations, enabling the creation of multi-layer thermal digital twins.
Discussion: This approach makes complex thermal data more accessible to Architecture, Engineering, Construction, and Operation (AECO) professionals. By providing a structured workflow for interoperability, the study facilitates improved building management and more accurate energy analysis through the use of integrated thermal digital models.

KEYWORDS

AECO, BIM, data integration, digital twins, education building, GIS, point clouds, thermal models

1 Introduction

The Architecture, Engineering, Construction, and Operations (AECO) industry is in a constant search for innovative solutions to understand the current state of buildings and infrastructure, monitor different variables that can improve management, seek greater energy efficiency in the built environment, and move towards zero-energy buildings and infrastructure. In this context, digital twins are already a reality in the industry, employing methodologies such as Building Information Models (BIM) and Heritage Building

Information Models (H-BIM) to create information models that can be enriched with a multitude of data. Over the last decade (2015–2025), numerous studies have also emerged proposing the contextualization of these building information models in broader urban settings, where georeferenced data can be included using GIS systems. For example, López-González and García-Valldecabres (2023), explored and determined the possibilities of interoperability between H-BIM and GIS to synchronize and manage information on heritage architecture, sustainable conservation, and territorial and cultural tourism planning, developing a novel technological system that integrates H-BIM and GIS into a single tool for cultural development and heritage preservation. A 3D-GIS prototype is created for visitor management and preventive maintenance of heritage buildings and centers using the ArcGIS Pro platform, controlling problems caused by an excessive number of visitors in the conservation of a monument through sensors (CO2, humidity, temperature) and connecting this data with the number of visitors. Bottaccioli et al. (2017) presented a cloud-based software architecture for managing and simulating the energy performance of buildings, integrating heterogeneous data (BIM, IoT, GIS, weather services) to allow for (near) real-time visualization of energy consumption, building performance evaluation through modeling and simulation, and the creation of GIS energy maps. In this way, models are created that are enriched with information from all kinds of sensors to contribute to the monitoring of buildings and infrastructure, which can also help in the analysis and study of, among other aspects, energy efficiency. Back et al. (2023) also used a GIS model to improve urban heat and thermal comfort assessment, considering the vertical and horizontal variability of wind speed and air temperature patterns through computational fluid dynamics (CFD) simulations in an urban case study model in Innsbruck, Austria. Mutani et al. (2016) compared two GIS-based methodologies for the spatial characterization of the residential built environment at an urban scale in terms of building distribution and energy consumption for heating. Additionally, they proposed a third method for non-residential buildings, with the aim of supporting local energy planning and creating urban energy maps.

Despite their clear potential, the integration of thermal data in AECO remains an under-researched area. Currently, most projects focus on pairing 2D thermal images with geometric models, which limits the management of temperature data for applications such as energy simulations or thermal comfort studies. There are different methodologies for collecting thermal data for their inclusion in urban or building digital models, such as the one proposed by Lin et al. (2019), which fuses data from terrestrial thermal images with 3D point clouds to map thermal attributes on building facades, focusing on a registration between the thermal images and the point clouds. Thermal and RGB point clouds are created from their respective images using Structure from Motion (SfM) tools. In this case, the RGB point cloud is used as a 3D reference. In Hou et al. (2021), the performance of the fusion of exterior thermal data from facades is evaluated, considering how different Unmanned Aerial Vehicle (UAV) flight configurations, such as altitude, camera angle, flight path design, and construction typologies, affect it. This study seeks to provide suggestions for the collection and better processing of this type of data for the diagnosis of pathologies and failures in energy efficiency.

Within these types of digital models, the inclusion of thermal information has gained significant relevance in the last 5 years (2020–2025), allowing for the collection of the three-dimensional geometry of a building or space along with temperature data in continuous monitoring campaigns, and the correspondence between these points and temperatures, as developed in Adán et al. (2020), and the work carried out by Ramón-Constantí et al. (2022a). This information is initially presented as an unstructured TPC, which, once processed, can be transformed into a Proprietary Thermal Model (TPM). TPCs offer valuable applications in the AECO industry by providing a detailed and accurate representation of the distribution of surface temperatures on the interior and exterior of buildings, which makes them useful tools for understanding the thermal behavior of constructive elements and monitoring thermal changes over time. Additionally, they can be segmented and tagged for its use in BIM environments, containing specific information about the thermophysical characteristics of materials, facilitating maintenance tasks such as the identification of singular regions like thermal bridges, as well as serving as input data for performing energy simulations, as developed in Ramón-Constantí et al. (2024), where an experimental methodology is developed to work with this type of thermal model and calibrate building energy simulations using CFD. The integration of thermal data with other aspects of the architectural project also promotes multidisciplinary collaboration and more efficient decision-making regarding the management and maintenance of buildings.

Despite their clear potential, the use and applicability of building TPCs in professional software nowadays remains limited. This is largely due to significant challenges related to interoperability and format compatibility, which hinder the incorporation of thermal data and the subsequent management of proprietary thermal models in BIM and GIS environments. By laying the groundwork for a broader use of thermal models in the management and conservation of buildings, this work seeks to establish the path toward the application of thermal digital twins in the construction sector through a use case using a BIM-GIS integration.

The article is organized as follows: Section 1 is an introduction to the topic that discusses the current status and limitations in the AECO industry regarding thermal models and the poor state of the art regarding the integration of TPC into BIM and GIS; Section 2 presents the objectives, general methodology and contributions of the study; Section 3 describes the case study; Section 4 describes the use and integration of TPCs in an open-BIM viewer; Section 5 describes the integration of TPCs in a GIS platform; Section 6 discusses the limitations of the methodologies; Section 7 shows the conclusions of the study, and Section 8 proposes future works regarding the acquisition, implementation and use of TPCs in BIM-GIS scenarios.

Abbreviations: AECO: Architecture, Engineering, Construction and Operation; TPM: Thermal Proprietary Models; TPC: Thermal point clouds; BIM: Building Information Models; BEM: Building Energy Model; BES: Building Energy Simulation; BMS: building management system; GIS: Geographic Information System; IR: infrared; CityGML: City Geography Markup Language; LOD: Level of Detail; OSM: OpenStreetMap; UAV: Unmanned Aerial Vehicle; CFD: Computational Fluid Dynamics; IFC: Industry Foundation Classes; TLS: terrestrial laser scanning; VPI: visual programming interface.

1.1 The irruption and potential utility of the as-is thermal models in AECO

There has been a growing interest in systems and methodologies that allow the digitalization of buildings in their current state of the art. The term "as-is" (or "as-built") refers to a record of the dimensions, components and other aspects of an existing building or infrastructure at a time after its construction has been completed. In this framework, the aim is to optimize a workflow that, until now, has depended mainly on a technician who performs many of the processes manually and handles large volumes of data and information necessary to understand the operation and improve the preventive maintenance of buildings.

Traditionally, 3D digitalization of buildings has focused on capturing their geometry. However, in the last decade, studies and proposals have opened new lines of research aimed at the analysis and study of the energy efficiency of existing buildings, as pointed out by Rahhal et al. (2019). This has led to the first works regarding thermal 3D digitalization of buildings, and as a result, digital twins can now be enriched with this new data. Desogus et al. (2023) developed a BIM-centric workflow for monitoring, visualizing, and assessing thermal comfort in buildings, integrating BIM tools and sensor technology through visual programming using Dynamo visual programming interface (VPI). The objective is to optimize building energy use while achieving good indoor environmental conditions and supporting facility managers by updating the as-built model with efficiency improvements and linking parameters for monitoring. Monitoring data is associated with the "room" object in the proposed BIM case study. In Dlesk et al. (2023) it is proposed to optimize facility management by integrating IoT infrastructure with digital building models, enabling informed decisions using spatially oriented sensor data and creating a digital twin aimed at improving users' thermal comfort and reducing energy consumption. Thermal 3D digitalization entails collecting the three-dimensional geometry of a building or space, temperature data, and matching points and temperatures. The whole information is originally presented as an unstructured TPC, which can be subsequently processed until obtaining a Thermal Proprietary Model (TPM). In this regard, recent research into the generation of 3D models with thermal information stands out, although its scope is still very limited to the academic field, as concluded in the study by Ramón-Constantí et al. (2022b).

TPCs can have different useful applications in the AECO industry as they provide a detailed and accurate representation of the distribution of surface temperatures in particular areas or elements of a building or facility. Thus, they are tools to understand the thermal behaviour of construction elements, rooms or groups of rooms, as well as for monitoring thermal changes over time, as shown in Adán et al. (2023b), where an automated robotic platform is presented for obtaining thermal maps of interiors at different times through an automatic data acquisition approach. They are also useful for inferring thermal characteristics of buildings, as demonstrated by Hou et al. (2021) which proposed a framework for RGB-thermal data fusion to create district-level thermal maps, evaluating the impact of flight configurations such as camera altitude, angle, flight path design, and building types, and the similar study by Alba et al. (2011) where a 3D acquisition and processing method that integrates Red, Green and Blue (RGB) colour, thermal infrared (IR), and near-infrared (NIR) imagery onto detailed 3D building models is presented. The approach streamlines data fusion for anomaly detection and conservation analysis, using a "bi-camera" system, photogrammetry, and terrestrial laser scanning.

One of the advantages of these models is that they can be segmented and tagged for use in BIM environments, containing specific information related to other attributes of the material thermophysical characteristics. This allows working with local or global temperature data that can be used for various maintenance tasks, from the identification of heat loss regions to the evaluation of energy performance over time of residential buildings. Furthermore, integrating thermal data with other aspects of the architectural project facilitates collaboration between multidisciplinary teams, thus enabling informed and more efficient decision-making. Professionals can also assess the thermal impact of different designs, construction and retrofitting strategies and compare it with historical records and surface temperature readings taken in real time.

Another highlight is the scalability of these models. Thermal models can be enriched by adding new measurements or data coming from local sensor networks or extending the scope of study to other nearby buildings. This means that the models can be adapted to address specific challenges at different stages of the operation phase and long-term maintenance, even if the area of the building to be analysed is expanded.

1.2 Integration problems

There is a significant lack of research regarding the integration of data from different thermal data sources in AEC. Currently, the majority of the projects and research are mainly focused on matching thermal IRF images to geometric models of the facility or building, which does not really allow the construction professionals to manage quantitative temperature data for further applications, such as energy simulations or thermal comfort studies, what it is possible using TPCs, and, if any, the subsequently obtained Thermal Proprietary Models (TPM). A TPM is a basic three-dimensional model generated by programmers after processing the original TPCs of a scene. This model incorporates the surface temperatures of key architectural elements of a building, such as floors, ceilings, walls, columns, doors, and windows, into the geometry, as described by Adán et al. (2023c).

However, the use of TPMs in the AEC industry presents a series of adaptation problems to standard formats that are still little addressed today, especially regarding the variety and inter-compatibility between them. Moreover, adapting these thermal data to standard formats entails a high technical complexity that does not facilitate their widespread use. Therefore, the creation of thermal models in formats that allow their inclusion in digital twins of buildings would lead to a significant advance in applications related to building management and architecture in their operational phase.

Although there are standards, such as IDF or gbXML and others, the lack of a universally accepted format that can include thermal data from buildings makes it difficult to transfer thermal

information between different software tools and platforms, both open source and commercial software. This leads to the need to make specific adaptations to the data structures of these models for each software, which increases the complexity and time required to integrate thermal models into existing workflows. In addition, these modifications are usually made using tools and processes that involve specific programming knowledge, so it is not a simple process and has to be carried out by specialists. The most recent references deal with the issue of interoperability between formats in an energy simulation context, as in Bracht et al. (2021), where they develop using Autodesk Revit a gbXML-based schema, but at no point do they work with thermal models as the proposed work is focused in the early design stages. In Kamel and Memari (2019), this topic is also reviewed, but from a BIM-to-BEM point of view, where it is concluded that there are unresolved aspects such as data loss between formats, the improvement of different standards and formats, the use of intermediate tools and the requirement of manual processes to avoid data loss between two tools during data exchange.

Furthermore, the integration of real temperature data into BIM may require adding specific attributes or tags to represent the thermal information natively in each software through manual processes. This is because many BIM software programs are not designed to natively visualize thermal data, such as through color-coded representations. Although there are recent studies that show the integration of IR images in different BIM software on 3D geometric models, such as in Adamopoulos et al. (2021), no studies have been found that show how to integrate TPCs.

In summary, the integration of dense thermal information in the AEC sector currently faces significant challenges in terms of format compatibility and data exchange between open and commercial software, problems related to the capacity of the available software, and the representation and management of thermal data in BIM models.

1.3 TPC in BIM-GIS

Apart from the critical integration problems previously discussed, the use of TPCs has experienced remarkable growth over the last decade (2015–2025), driven by advances in 3D digitalization technologies and the increased accuracy and speed of obtaining data from buildings and infrastructure. In particular, integrating thermal point clouds into BIM and GIS is of great interest, as it can help detect issues related to building maintenance and management in the context of smart cities and the creation of digital twins of the built environment.

Beyond the thermal component, researchers have extensively explored the joint use of BIM and GIS as complementary methodologies that support building analysis and management. Göçer et al. (2016) present an integrated BIM–GIS model for pre-renovation processes, enabling the organization and visualization of relevant data to optimize renovation decisions in economic and environmental terms. Their approach addresses limitations of previous models, such as fragmented information and insufficient analytical tools, thereby improving collaboration between stakeholders. Kamel and Memari (2019) review the application of BIM in energy simulation using tools such as Revit, OpenStudio, and EnergyPlus, identifying persistent issues related to data transfer, information loss, and interoperability constraints across standards such as IFC and gbXML. Oti et al. (2016) propose a framework for integrating building management system (BMS) data into BIM models, demonstrating how operational data can support more sustainable building design and performance evaluation.

As highlighted by Hong-mei et al. (2001), GIS offers strong analytical capabilities for geospatial information. Its integration with BIM has therefore attracted interest in diverse applications, such as traffic noise modelling (Deng et al., 2016), flood damage assessment (Amirebrahimi et al., 2016), and indoor evacuation planning (Isikdag et al., 2013). In García-Valldecabres et al. (2023), a 3D-GIS database integrating HBIM models is implemented to study visitor behavior and mobility in historic centers. Such examples show how the combination of BIM and GIS offers enriched spatial and semantic information suitable for complex building and urban analyses.

Regarding the thermal dimension, incorporating thermal information into BIM–GIS contexts has become increasingly relevant in recent years. Antón and Amaro-Mellado (2021) proposed an open-source GIS-based platform for the visualization and analysis of 3D thermal point clouds of exteriors and urban environments. Previtali et al. (2012) introduced a photogrammetric methodology for mapping infrared images onto 3D models in a GIS environment. Lagüela et al. (2014) presented a workflow for thermal mapping of exterior building models using UAV-mounted thermographic cameras, with comparison against terrestrial laser scanning as reference data. These studies underline the importance of integrating thermal information into BIM-GIS models for applications such as assessing thermal insulation systems, detecting construction defects, or monitoring energy performance. Complementarily, Tomchinskaya and Galanina (2020) address the integration of non-geometric data for infrastructure management in a campus setting, while Niu et al. (2015) develop a BIM–GIS-based system for visualizing building energy data. However, despite these advances, no studies to date describe structured methodologies for integrating full TPC datasets into BIM-GIS systems.

While previous research has focused primarily on mapping 2D infrared imagery onto 3D geometries or on limited thermal-fusion techniques, these approaches do not enable the structured propagation of point-based thermal attributes into standard BIM or GIS formats. Moreover, existing BIM–GIS interoperability studies remain centered on geometric or alphanumeric information and do not address the challenges associated with representing dense, three-dimensional thermal datasets. The workflow proposed in this study addresses this gap by enabling the incorporation of TPC-derived temperature attributes into both an open BIM schema (IFC) and a GIS environment. This dual integration ensures that thermal information can be preserved, queried, and exchanged across platforms, thereby extending current practices beyond IR-image mapping methods and providing a novel, comprehensive pathway for thermal data interoperability in BIM–GIS workflows. Surface temperatures assigned to different elements in the model.

2 Materials and methods

2.1 Objectives of the study

The main objective of this research is to develop a novel methodology for managing TPCs and their associated data within a combined BIM-GIS environment. The goal is to establish a workaround that lays the foundation for the creation of thermal digital twins, designed to enable the various building stakeholders to understand and improve the thermal and energy performance of interior spaces, as well as the comfort of their users. To achieve this goal, the following specific objectives have been defined.

1. To establish a workaround for the integration of TPC into GIS environments. Through a case study, the aim is to create a workaround that allows, for the first time, the direct incorporation of TPC into a geospatial environment, addressing a significant gap in current research.
2. To propose a methodology trying to reduce the need for advanced programming, while acknowledging that certain steps, particularly IFC generation and attribute automation, currently still require use of scripting.
3. To generate a digital replica with thermal properties. The goal is to create a digital twin of the physical structure that incorporates its thermal characteristics. This model should allow for three-dimensional visualization with multiple layers of information (TPC data, sensor readings, etc.), facilitating a detailed analysis of each component.
4. Incorporate time variables for dynamic analysis. The integration of the time dimension into the digital model will be explored, allowing managers to observe and verify the building's thermal evolution and changes throughout its use cycle.

2.2 Methodology workflow

This research addresses a notable gap in the scientific literature which is the lack of established methodologies for integrating TPCs into GIS environments and the scarcity of studies that do so in BIM environments. While previous works exists, such as that of Ramón-Constantí et al. (2024), which integrates a TPC into a model for building energy simulation (BES) using several CFD simulations of an indoors scenario, direct connection with GIS platforms at the level of three-dimensional thermal data remains an unexplored field. This research seeks to establish a workflow for the efficient generation of a thermal digital twin, a virtual replica that facilitates the integration of geometric and alphanumeric data based on spatial location for its use in building management by different stakeholders.

To develop the methodological framework, existing approaches for integrating BIM and HBIM (Heritage Building Information Modelling) models with GIS were analyzed. For example, the study by Escudero et al. (2024) demonstrates the feasibility of incorporating detailed alphanumeric data for the comprehensive visualization and analysis of 3D models in such environments using an IFC format. Complementarily, the methodology proposed by Albourae et al. (2017) for generating HBIM models offers a relevant workflow which begins with data collection using terrestrial laser scanning (TLS) and photogrammetry; continues with the extraction of a 3D model from the integrated data; and does an element identification within a BIM platform such as Autodesk Revit. A key aspect of their work is the creation of parametric object libraries for distinctive historical elements enabling a detailed study of architectural heritage. They propose integrating the HBIM model into a 3D-GIS to semantically enrich it with non-architectural information, such as materials, historical context, or repair history, overcoming the limitations of traditional 2D-GIS. The methodology used to create the TPC is not the subject of this article since it is already a topic developed in Ramón-Constantí et al. (2022b), Adán et al. (2023a), Adán et al. (2023b). Since there is no established protocol for integrating TPCs into GIS, the methodology proposed in this article is structured as follows:

(1) An exhaustive information search was conducted to determine the state of the art, using sources such as Web of Science and Journal Citation Reports, Google Scholar, and ResearchGate, with a focus on the last decade (2015–2025) and using keywords such as "BIM-GIS," "thermal point cloud integration," and "TPC-GIS."; (2) Data acquisition planning was carried out, which involved familiarization with the available hardware (fixed and mobile sensors) and (3) coordination with the stakeholders involved in the acquisition process, establishing the optimal location for the sensors; (4) The field phase consisted of on-site data collection, followed by (5) downloading and processing the information, generating two products: a non-geolocated TPC model and a geometric model from a manual Scan-to-BIM process in Revit. In parallel, an IFC file was partially-automatically generated from the TPC, containing the average thermal information for each element; (6) The manual 3D model was subsequently imported into a GIS environment, where it was modified to assign the average temperature as a GIS attribute to each element. Finally, (7) it was semantically enriched with open data and information from the available sensors. The result is an integrated model that is imported and public viewed in ArcGIS Online (sceneviewer) and connected to ACCASoftware GeoTwin (for the TPC), serving as the basis for a thermal digital twin that can be used for efficient building management and monitoring which can be used by different building stakeholders in a university or limited urban indoors/outdoors context. The workflow includes both automated and manual processes: scripting is used for vertex extraction, attribute assignment, and IFC generation, whereas segmentation and modelling remain manual. A flowchart of this process is provided in Table 1.

2.3 Contribution of the paper

This article presents a novel, practical workaround for integrating TPC data into both standard BIM formats and a GIS environment, demonstrating a significant step toward developing thermal digital twins. The contribution lies in formalizing and validating a workflow that has not been described in prior literature, particularly for TPC–GIS integration. This contribution entails a step towards the development of thermal digital twins and

TABLE 1 Diagram of the BIM-to-thermal-GIS methodology, its steps, automatic vs manual process and used software (Partially Automatic means automatic execution of pre-developed scripts).

Step	Action	Description	Automatic/partially-automatic/manual	Software/S used
Step 1	State of the art	Using sources as web of Science, Journal citation reports, google Scholar and ResearchGate. Last decade (2015–2025)	-	-
Step 2	Acquisition planning	Get familiar with available sensors, data formats and hardware	Manual	-
Step 3	Optimal locations and coordination	Establish optimal locations for sensors and measurements. Coordinate procedures with acquisition stakeholders	Manual	-
Step 4	On-site data collection	Collecting data with sensors (laser scanner, IR camera environment sensors)	Partially automatic	Matlab R2020 (to launch the TPC acquisition) + leica cyclone REGISTER 1.4.3. + thingspeak platform
Step 5	Downloading and data processing	Download: TPCs; processing: Global TPC + Scan-to-BIM models.RVT IFC	TPCs: Automatic Global TPC: Partially-automatic TPC segmentation: Manual Scan-to-BIM: Manual	TPC: Matlab R2020; CloudCompare 2.13.2 Scan-to-BIM: + Autodesk Revit 2020
Step 6	Import to gis environment	Import both models: RVT (building space), IFC (curtain wall) to ArcGIS pro. Assign average temperature as GIS attributes to each element in the geometric model of the curtain wall	Manual	ArcGIS pro 3.5
Step 7	Semantically enrich and integrate bim-gis models	Get available sensor data from Step 5 and integrate into models in ARQ GIS online connected with other online platforms	Manual	Thingspeak platform + google Drive + ACCASoftware geotwin

aims to establish a first novel workflow for integrating this type of data into BIM-GIS platforms. The novelty of this workflow against prior literature is achieved by incorporating high-resolution, complex thermal data (obtained from large-scale, multi-story glazed façades) into common AECO industry platforms, particularly addressing the current lack of direct, full-feature integration tools between TPC processing software, BIM, and GIS. The methodology for obtaining TPC data of buildings is based on that developed by Ramón-Constantí et al. (2022a), and which was also discussed in Adán et al. (2023a); Adán et al. (2023b), so it is not the object of this article.

The article presents the steps using a case study composed of an interior scene of a building together with a portion of a complex façade. This work has been structured in two parts:

1. Integration of TPCs into open BIM formats: The aim is to import TPCs into the IFC BIM format and, in this way, obtain collaborative and interdisciplinary thermal models in the BIM field.
2. Integration of TPCs into geographic information systems (GIS): The aims are to introduce a TPC into a collaborative environment beyond the building scale, and to share information obtained from different data sources.

3 Case study

In this case study, geometric and thermal data were collected over 10 days in a building located on the New York University Abu Dhabi (NYUAD) Campus. The sensors used were the Leica's BLK360 laser scanner, a network of ambient temperature and contact sensors, and a set of infrared (IR) images taken with a handheld IR camera. Table 2 specifies the technical characteristics of each of the sensors used.

The building studied is located on the west side of the campus and consists of a spacious ground floor with multiple entrances, wide corridors, and areas dedicated to classrooms, laboratories, offices, and open and closed workspaces. The work area includes a ground floor and a mezzanine. The ground floor is mainly dedicated to transit areas, with a large double-height hall that opens onto an interior courtyard built with an ellipsoidal glass curtain wall. Figure 1a shows the setup of one of the scans. Figure 1b shows the interior view of the RGB point cloud inside the building. The upper floor houses circulation areas, open offices, and workspaces that were not accessible during the study. During data collection, the entire interior space was air-conditioned.

To perform a complete thermal scan of the scene, the scanner laser sensor was placed in three different positions in the interior of the building: two on the ground floor to capture data from the half curtain wall, one on the landing of the stairs connecting the double height to the office space on the mezzanine and two on the upper floor. Three thermal point clouds (one vertical and two oblique at +45° and −45°) were obtained for each position in order to cover the visible space from the position, resulting in a total of 15 scans. Figures 2a–e provide an example of five of these scans, while Figure 2f illustrates the complete TPC after registering the collected samples. Using the scanner configuration to create TPCs does not allow them to be geolocated; this data is lost during the processing and formation of the thermal cloud.

TABLE 2 Technical characteristics of the sensors and equipment used.

Sensor	Task	Accuracy	Measuring range	Data logging frequency	Power source/ consumption
Leica BLK360 laser scanner	3D scanning and spherical imaging	3D point accuracy: 4 mm @ 10 m	Scan: 0.5–45 m	Full scan: 20 s; 680,000 pts/sec	Rechargeable battery (GEB825), up to 70 setups
Handheld IR camera PCTE-TC3 (TC3-X)	Thermal imaging and detection	Temp accuracy: +/- 3 °C (<100 °C), +/- 3% (>100 °C) @ 5 m	Temp: −20 °C–150 °C or 0 °C–550 °C	Frame rate: 8.3 fps (Slow)/ 25/30 fps (fast)	AC via housing (24/115/ 230 Vac), 10W–130 W (housing)
Environment sensor Sensirion SEN54	Air quality monitoring (PM, RH/T, VOC)	PM accuracy: ±10%; RH: 4.5 %RH; temp: 0.45 °C	PM: 0–1,000 μg/m^3; RH: 0–90 %RH; VOC: 0–1,000 ppm	RH: 20 s (τ63%); VOC: <10 s (τ63%); temp: <60 s (τ63%); VOC power-on: 1 h	4.5–5.5 V, 63 mA (average)
Environment sensor BME280	Humidity, pressure, and temperature measurement	RH: ±3% RH; pressure: ±1.0 hPa; temp: ±0.5 °C	RH: 0%–100% RH; pressure: 300–1,100 hPa; temp: −40 °C–85 °C	1 Hz (low power); pressure: 157–182 Hz (max); RH: 1 s (τ63%)	1.71–3.6 V (VDD), 1.2–3.6 V (VDDIO); 0.1 μA (sleep), 1.8–3.6 μA (active @ 1Hz)

Table 3 shows a summary of the TPCs obtained, the number of points, the time spent per scan and the resolution used. In addition, ambient temperature (Celsius degree) and humidity data (in %) were recorded for 1 week. These data were stored in a common data environment together with point clouds and a set of IR photographs, thus complementing the monitoring of the scenario. Table 4 shows relevant information about sensors, location, time periods, data and formats used. The methodology used to obtain the TPC does not allow them to be geolocated automatically.

4 Use and integration of TPCs in BIM

The main objective, in this case, is to process a thermal point cloud in order to generate an open IFC file containing information on the geometry of the curtain wall with the associated temperature (mean value) of each of its vertical and horizontal components at a certain moment.

4.1 Data collection methodology

The following steps were followed to obtain the IFC file for the curtain wall:

1. Segmenting the total thermal cloud. The total thermal cloud (TotalTCloud.txt) of the scenario was obtained following different processes (e.g., aligning, segmentation, reduction and regularization) developed in *Matlab*, which can be found in Adán et al. (2023c). However, in this case, since only the curtain wall part was of interest, TotalTCloud.txt was manually segmented in *CloudCompare* (see Figure 3). The segmentation consisted of a range filter to select only the points on the facade between 20 °C and 28 °C, eliminating outliers and noise produced by reflections, external objects, or ventilation grilles, and adjusting the false color to this range to maximize the contrast of the facade, which was at a higher temperature than the average of 19 °C shown in TotalTCloud.txt.
2. Generating the .las file. The segmented cloud was saved in *CloudCompare* in.las format, generating the file TCurtainCloud.las.
3. Generating the .rvt file. TCurtainCloud.las was read in *Autodesk Recap* and exported in.rvt format so that it can be imported into *Autodesk Revit 2020*. In this step, a change in the coordinate reference system occurred. Once TCurtainCloud.rvt was imported into *Autodesk Revit 2020*, the geometry was manually generated, structured in segments of the "panel", "mullion" (vertical segment) and "crossbeam" (horizontal segment), which are native Revit families. This model, which we call CurtainModel.rvt, contains, for the moment, only geometric information and the native properties of the curtain wall elements (see Figure 4).
4. Obtaining Coordinates.csv. In order to save the list of 3D coordinates of the vertices of the links and crossbars in an editable file, the *Dynamo* plugin was used, which generated a Coordinates.csv file to be able to automatically later assign the temperature data from the TPC.
5. Aligning and assigning temperature to curtain wall elements (mullions and crossbeams). The point cloud acquired by the terrestrial laser scanner (*TCurtainCloud.las*) and the point cloud derived from the Revit model vertices (*Coordinates.csv*) were imported into MATLAB, resulting in *TCurtainCloud.mat* and *Coordinates.mat*, respectively. To align the two datasets, a coarse registration of the curtain wall edge points to a predefined world coordinate system was first performed using rigid transformation routines. This initial step ensured a rough spatial consistency between both sources. In a second stage, a refined alignment was carried out using the well-known Iterative Closest Point (ICP) algorithm, which minimized the residual distances between corresponding surface regions of the two-point sets and yielded the final accurate registration.
6. Geometric deviations between the laser-scanner point cloud and the Revit-derived vertices were quantified from the residual point-to-surface distances after ICP convergence. The ICP iterations were stopped when either (i) the relative decrease of the root-mean-square error (RMSE) between two consecutive iterations fell below 1×10^{-4}, or (ii) a maximum of 100 iterations was reached. In our case, convergence was achieved after 27 iterations, yielding a final RMSE of 0.035 m, with a mean deviation of 0.030 m (±0.012 m) and a maximum deviation of 0.070 m over all vertices.

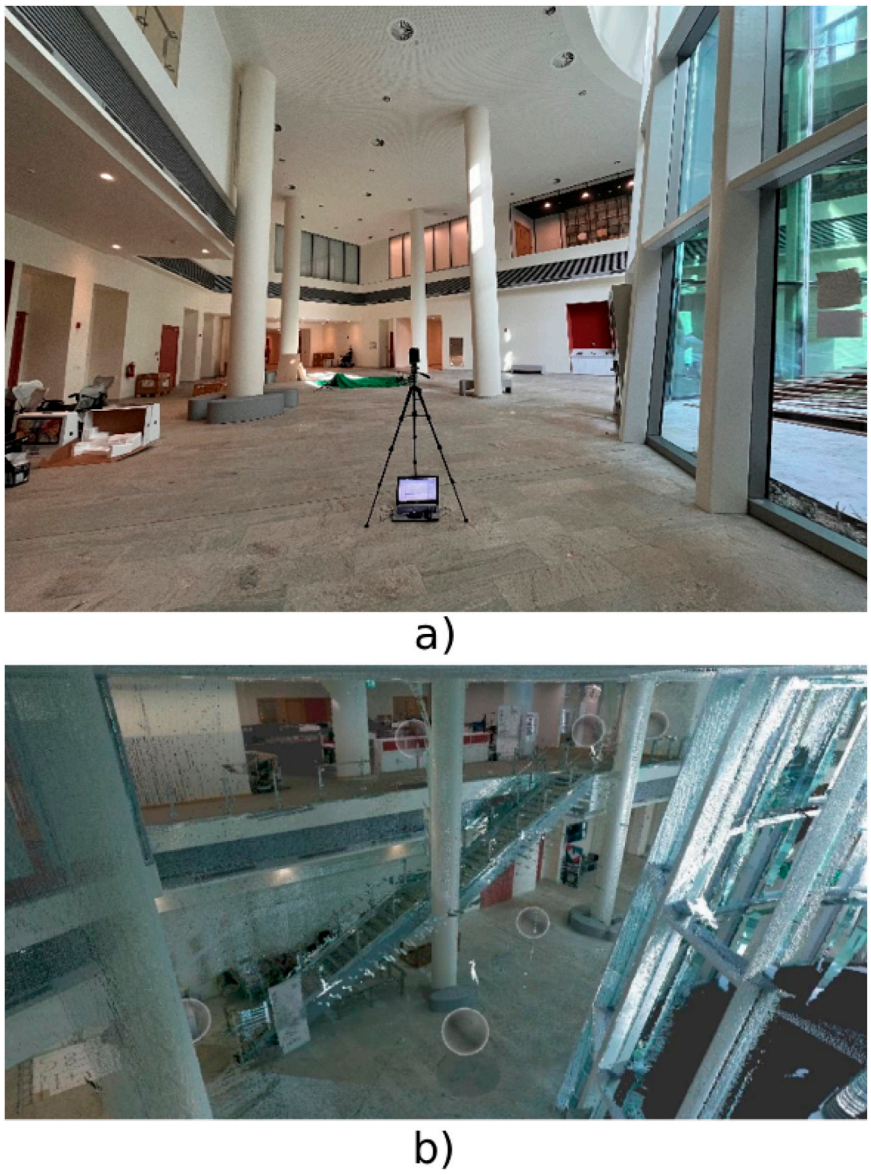


FIGURE 1
(a) Set-up of the laser scanner mounted on the tripod and computer equipment to capture data from a single position; **(b)** Interior view of the RGB point cloud inside the building.

7. Once the two datasets were aligned, the temperature values were assigned to each vertex in *Coordinates.mat* using a nearest-neighbor algorithm. Subsequently, the representative average temperature of each segment was computed by averaging the temperatures of its corresponding vertices (Figure 5a). We aggregated vertex temperatures using per-element averages because IFC schemas do not support point-level scalar attributes and because excessive data density would hinder interoperability. In the case study there is a large number of elements (Figure 5b) that require the assignment of a representative temperature, and this value can only be robustly obtained by averaging multiple temperature values. The mean values and standard deviation of temperature for each of the elements of the curtain wall (Mullion, Panel and Crossbeam) vary from 22.0 ˚C to 24.5 ˚C and 0.21 to 0.41 Std, while temperature absolute values vary from 21.5 to 25.5 in the TPC (Figure 5c). This summarization is consistent with thermal-simulation workflows that use element-level properties, as surface temperatures for CFD (Computational Fluid Dynamics) simulations.
8. Generating the .ifc file. Having obtained the representative average temperature for each segment of the curtain wall, the corresponding IFC ASCII code was generated using a MATLAB script, producing the file named *TCurtainModel.ifc*. The code was generated following the standard IFC text-based specification, using the basic parameters: *ifcName*: seg**_{panel, mullion, crossbeam}, *ifcExportAs*: IfcWallNotDefined, *ifcSpatialContainer*: Ground Floor, *Value (Temperature)*: **, and *ifcPropertySetList*: “Temperature”.

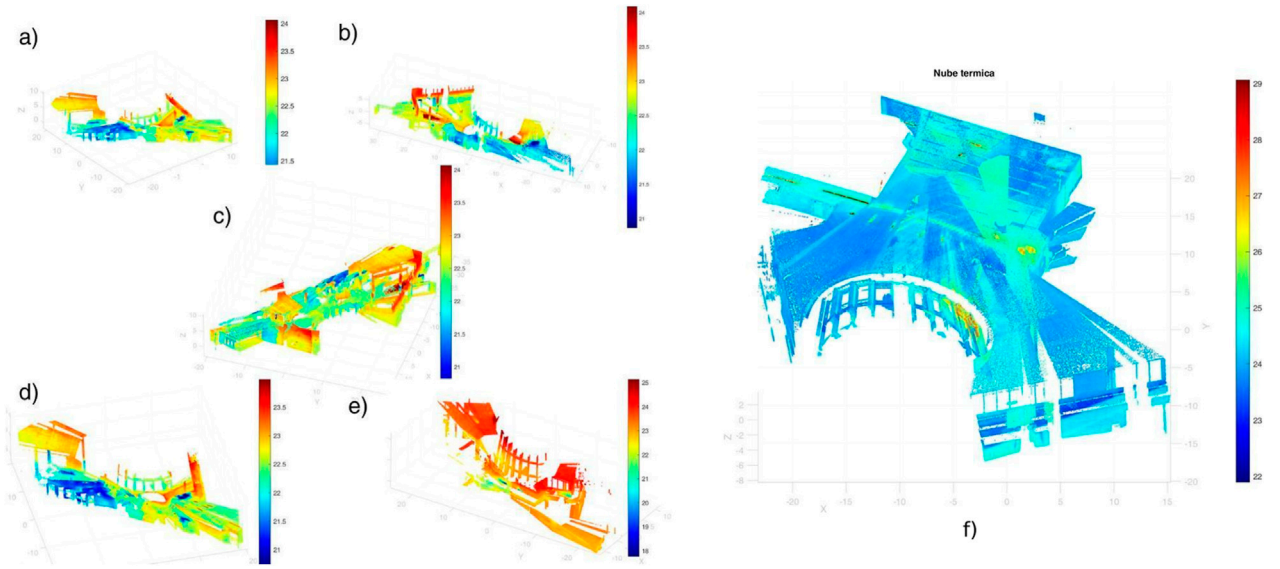


FIGURE 2
Sample of thermal point clouds from 5 scanner positions on: the interior ground floor **(a,b)**, the landing of the stairs **(c)**, and the upper floor **(d,e)**; **(f)** Total thermal cloud obtained after aligning 9 partial clouds, which covers the whole scene.

TABLE 3 Characteristics of point clouds taken on February 27, 2023.

Data times	Position number	Takes	Location	Point COUNT	Resolution
8:30 to 9 h	1	1 vertical (90°) + 2 oblique (+- 45°)	Ground floor (west part curtain wall)	23,7M	
9–10 h	2	1 vertical (90°) + 2 oblique (+- 45°)	Ground floor (east part of curtain wall)	10M	LOW MEDIUM HIGH
10:30 to 11:30 h	3	1 vertical (90°) + 2 oblique (+- 45°)	Staircase landing between floors	12M	
12–14 h	4	1 vertical (90°) + 2 oblique (+- 45°)	Upper floor landing	12M	
14–15 h	5	1 vertical (90°) + 2 oblique (+- 45°)	Upper floor corridor	11M	

TABLE 4 Summary of sensors, location, time periods, data and formats for monitoring the case study space.

Sensor type	Location	Date	Data	Data format
#1 SEN54	Bottom of curtain wall	23/02/2023 to 4/03/2023	Temperature + relative humidity + Pm2.5	Text file.csv
#2 SEN54	Interior forged front between floors	23/02/2023 to 4/03/2023	Temperature + relative humidity + Pm2.5	Text file.csv
#3 SEN54	Wall in office space	23/02/2023 to 4/03/2023	Temperature + relative humidity + Pm2.5	Text file.csv
#4 BME280	Hallway in office space	20/02/2023 to 1/03/2023	Temperature + relative humidity	Text file.csv
#5 BME280	Hallway in office space	23/02/2023 to 4/03/2023	Temperature + relative humidity	Text file.csv
#6 BME280	False ceiling office space	20/02/2023 to 1/03/2023	Temperature + relative humidity	Text file.csv
Outside sensors	Various locations outside the campus	23/02/2023 to 4/03/2023	Temperature + relative humidity	Text file.csv
BLK360 scanner	Detailed in Table 3		TPC	Point cloud file.txt

9. Checking the .ifc file. Finally, in order to check the quality of the resulting file in a BIM environment, TCurtainModel.ifc was imported into both *BIM-Vision* and *Revit* viewers. See Figure 6. As an example, the temperature assigned to a specific element of the curtain wall can be seen as a parameter or property.

4.2 Discussion about the software tools used

The software *Matlab, CloudCompare, Recap, Revit, BIMVision* and *Dynamo* were used in this process. A brief discussion about the software used can help the reader to better understand the methodology defined in the previous section.

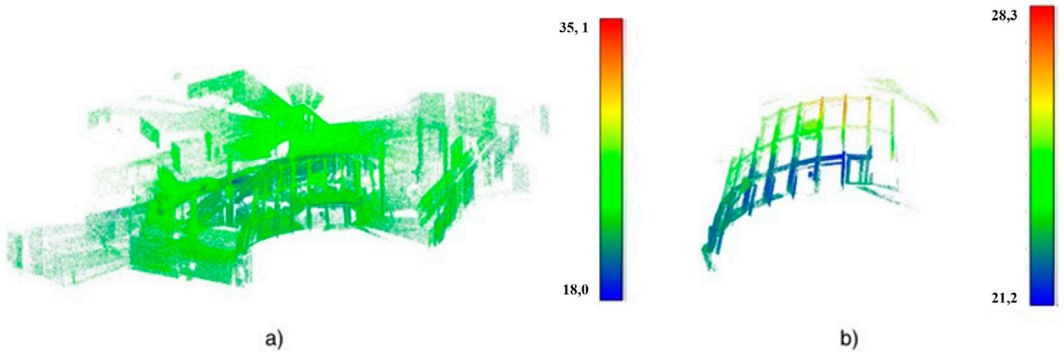


FIGURE 3
(a) Total thermal cloud **(b)** Segmented cloud in *CloudCompare*.

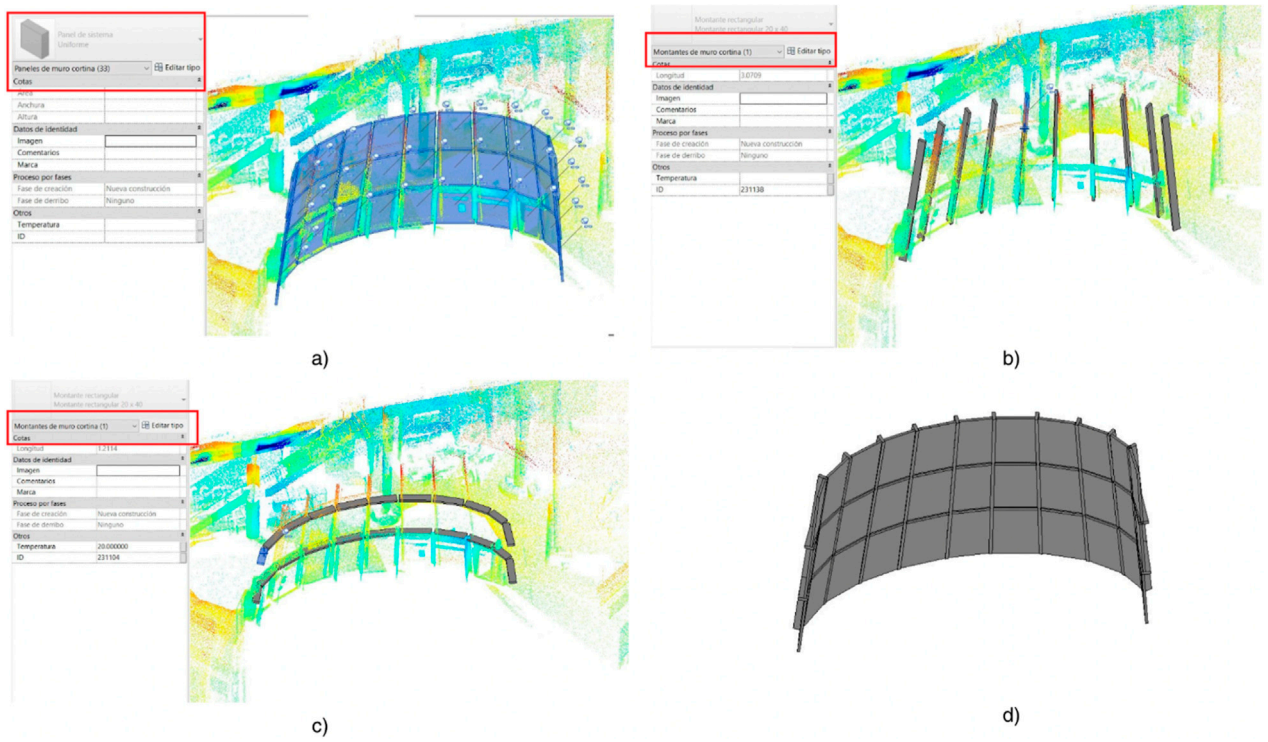


FIGURE 4
(a) TPC and detail of "panel" segment tracing; **(b)** "mullion" tracing (vertical segment); **(c)** "crossbeam" element tracing (horizontal segment); **(d)** Final curtain wall model made in Revit.

- Use of *Matlab*. Matlab was used in several phases of the process, essentially as a programming language. First, the point cloud was generated in a native.txt format so that it was accessible to other commercial software. Later, the temperature assignment to each vertex of the segments created in the curtain wall model was programmed. Finally, the.ifc format was generated in Matlab, containing the coordinates of the vertices and the assigned temperature of each segment of the curtain wall.
- Use of *CloudCompare*. CloudCompare was used to perform a manual segmentation of the total cloud to keep the required part of the curtain wall and to export it in.las format.
- Use of *Recap and Revit*: *Recap* was used as intermediate software to move from.las to.rvt formats, while *Revit* was used as a modelling tool for the curtain wall. As mentioned above, the point cloud served as a guide to manually trace the curtain wall, structured in segments (parallelepipeds) of three types: panel, mullion and crossbeams. At the end of the process, the final.ifc file containing the assigned

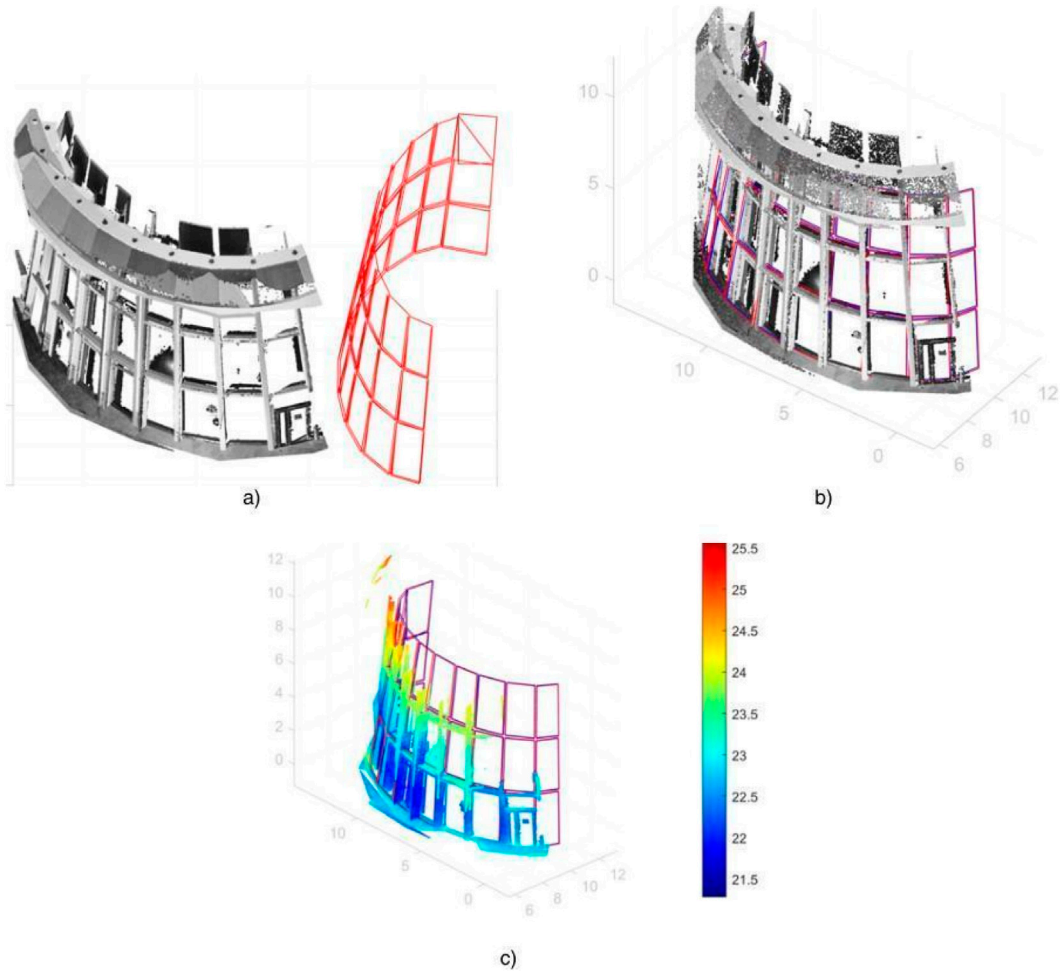


FIGURE 5
(a) Image of the original point cloud and the misaligned *Revit* model due to the change in reference system produced in the import step in *Recap*; **(b)** Original point cloud and the *Revit* model after alignment; **(c)** Visualization of the thermal cloud and the overexposed aligned *Revit* model.

temperature was imported into *Revit* to verify the validity of the file. *BIMVision* viewer was also used to check that the temperature was correctly assigned as an attribute of each segment of the curtain wall and visualized in an open format.

- Use of Dynamo: The *Dynamo* plugin was used exclusively to generate the ordered lists of the coordinates of the vertices of each segment of the curtain wall.

5 Results. Integration of TPCs and local sensor data into GIS

The objective of this section is to define a methodology to integrate TPCs and other local sensor data in a GIS environment. This methodology could be extended to other buildings on the campus during long-term monitoring periods.

The starting point of our methodology is the model in.rvt format of the building studied. This model was manually obtained in Revit from the TPC, as mentioned in Section 4.1. Step 2. This is used as a layer in the *ArcGIS Pro 3.5* desktop software, in which to add other geometric elements that represent the corresponding physical sensors used. At a later stage, this BIM model is uploaded to the *ArcGIS Online* platform along with the layers with elements created in the desktop tool. Finally, this GIS model is uploaded to *ACCA-Software's GeoTwin* platform, where the TPC is finally added as another data that can be consulted. *ArcGIS Online* platform serves as a centralized container where the rest of the data obtained are uploaded. All this will be discussed in this section.

5.1 Data integration methodology

Because direct import of non-geolocated LAS-based TPCs was not supported in ArcGIS Online, the GIS workflow adopts a workaround in which manual multipatch geometries receive thermal attributes derived from the TPC. The methodology carried out to integrate TPCs into GIS is summarized in the following steps:

1. Geo-referenced model of the *NYUAD* campus. The detailed and geo-referenced model of the *NYUAD* campus volumes was obtained from *OpenStreetMap* data and integrated into a

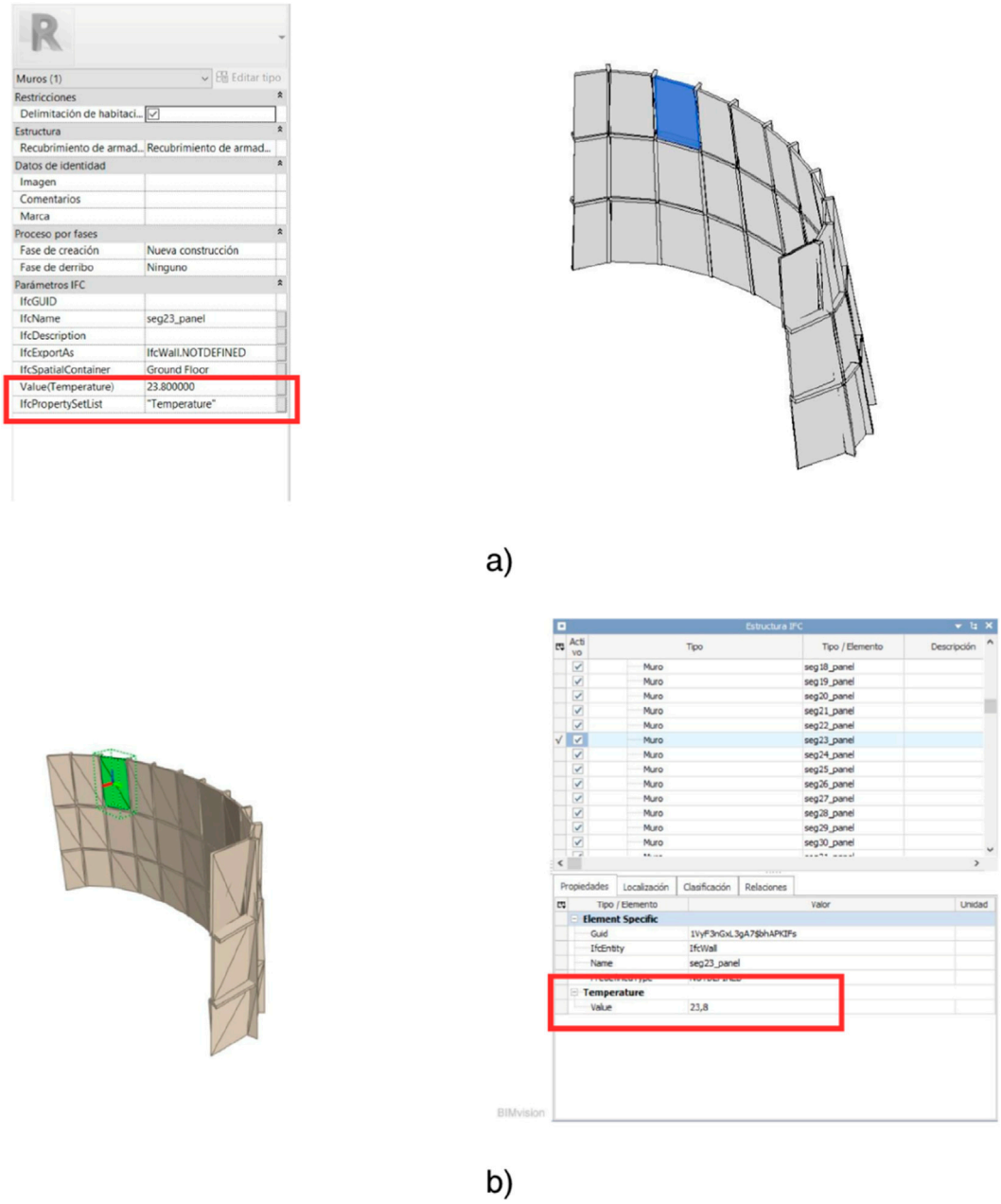


FIGURE 6
(a) TCurtainModel.ifc imported into *Revit*. The temperature assigned to the blue element is seen in the Parameters section. **(b)** TCurtainModelT.ifc imported into *BIMVision*. The temperature assigned to the earlier element is now seen in the Property section.

base map using satellite images. The metadata associated with the campus buildings was filtered, retaining only the parameters that identify the different buildings and specify their function (educational, office, residential, and services, among others).

2. 3D geometric model of the studied building. The 3D model made in Autodesk Revit of the double-height space studied was designed with the help of the original TPC, obtaining the Building.rvt file (see Section 4.1. Step 3).
3. Basic GIS model. To import the Building.rvt model into *ArcGISPro* from *Revit*, an internal tool of the software was used that facilitates the conversion of data "From BIM to geodatabase." This process allows the model to be segmented into elements categorized according to the default Revit labels (Architectural, Floors, Masses, Structural, among others). It is important to note that these categories cannot be edited manually, nor can new ones be added.
4. Inserting local sensors into the GIS model. The local sensors used were added to the space in the curtain wall and the upper floor as "3DPoint" elements in the model, each assigned an attribute "name" and "link" that corresponds to a.csv file in a cloud storage (*Google Drive*), where the data recorded from the ambient temperature (in Celsius degree) and contact sensors during the test period are stored.

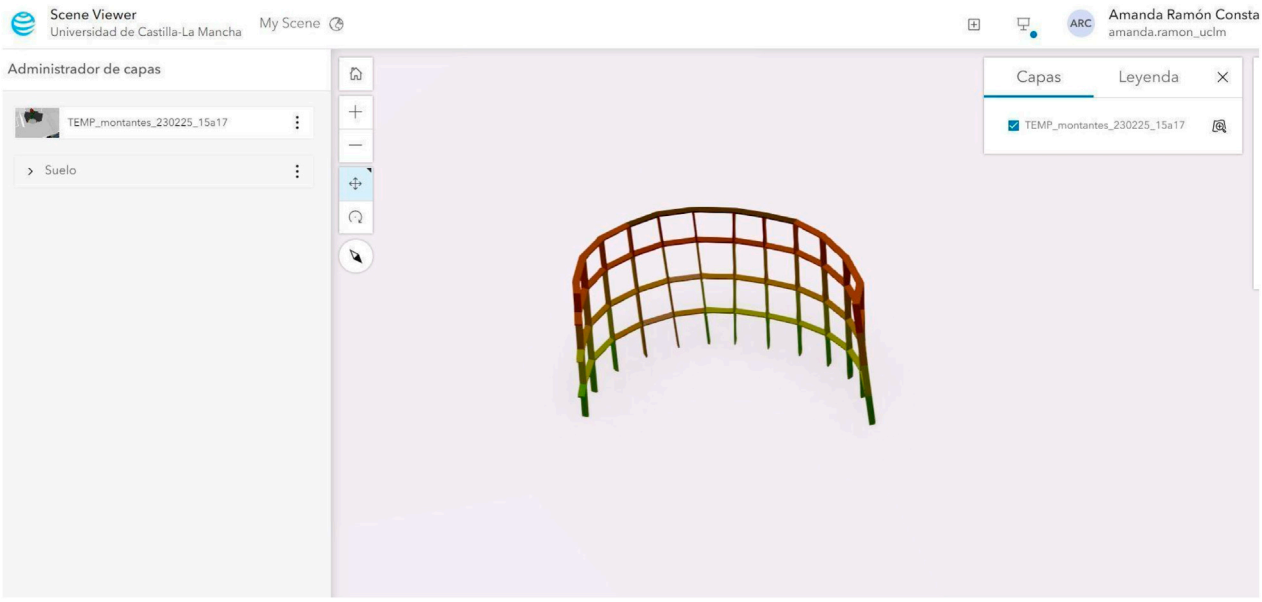


FIGURE 7
Multipatch object in ArcGIS Pro Scene Viewer, where a custom scale is added to visualize the temperature.

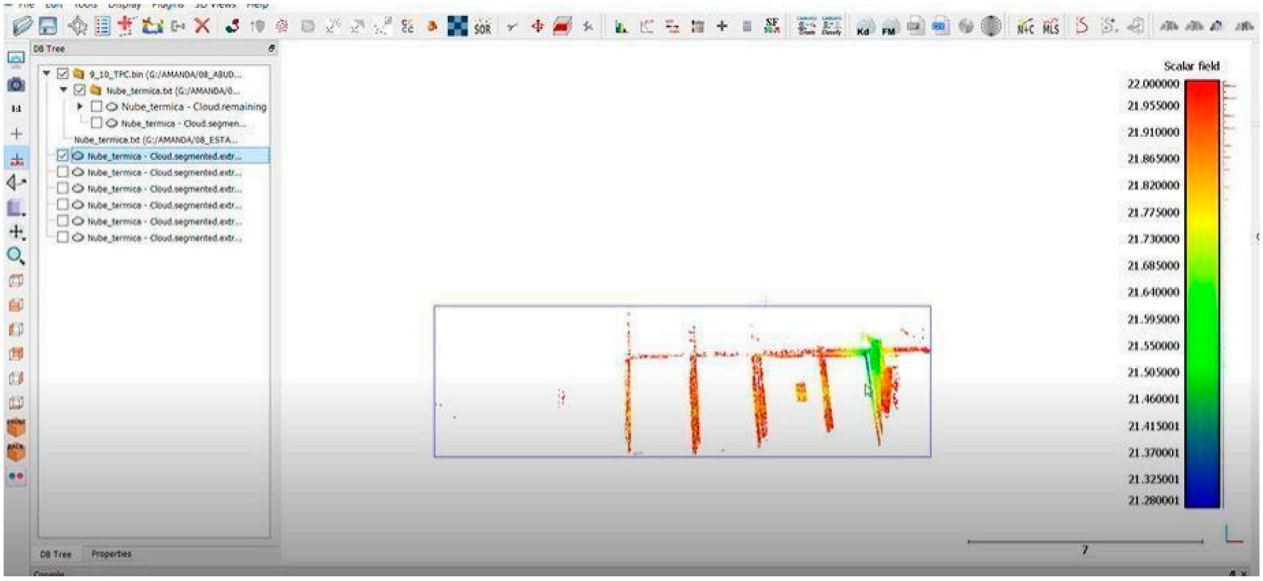


FIGURE 8
Interface of *CloudCompare* software showing manual filtering by temperature values of the TPC.

5. Inserting temperature into the GIS. Regarding the inclusion of temperature in the curtain wall façade, this was firstly converted from a commercial BIM Revit format to a three-dimensional data structure (named "multipatch") to which custom attributes can be added and then visualized using a colour scale (Figure 7). An attribute called "TEMP" was created to store the scalar values previously obtained from the TPM through a manual thermal segmentation process by values performed in *CloudCompare* (Figure 8). Two different "multipatches" were created to examine how to insert data from two different readings taken at different times. Figure 9 also provides a legend showing how these temperature values are organized in the BIM-GIS model.

The proposed TPC to GIS thermal model is available for consultation at Ramón-Constantí (2024) on the *Esri* online platform.

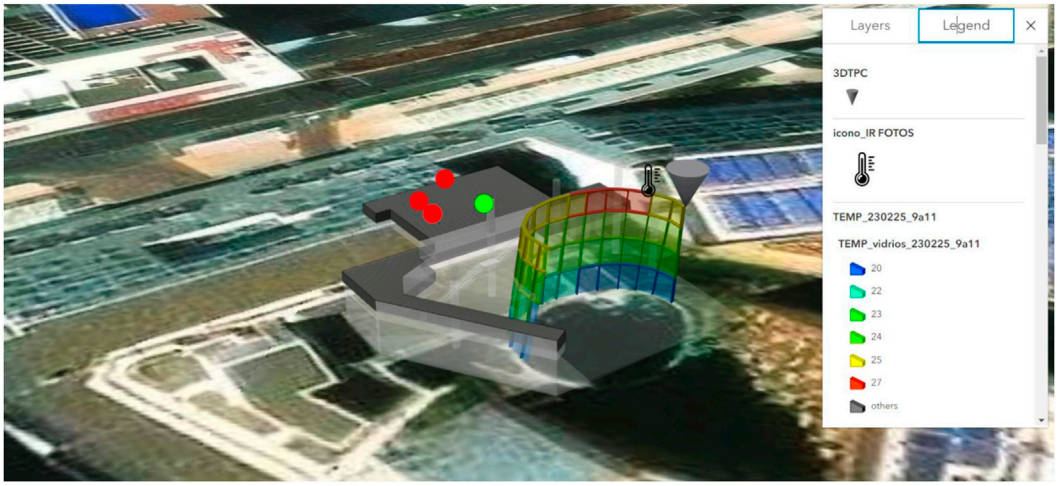


FIGURE 9
View in *ArcGIS Online* of the GIS model with the legend corresponding to the temperature values of the curtain wall and other elements on the satellite image.

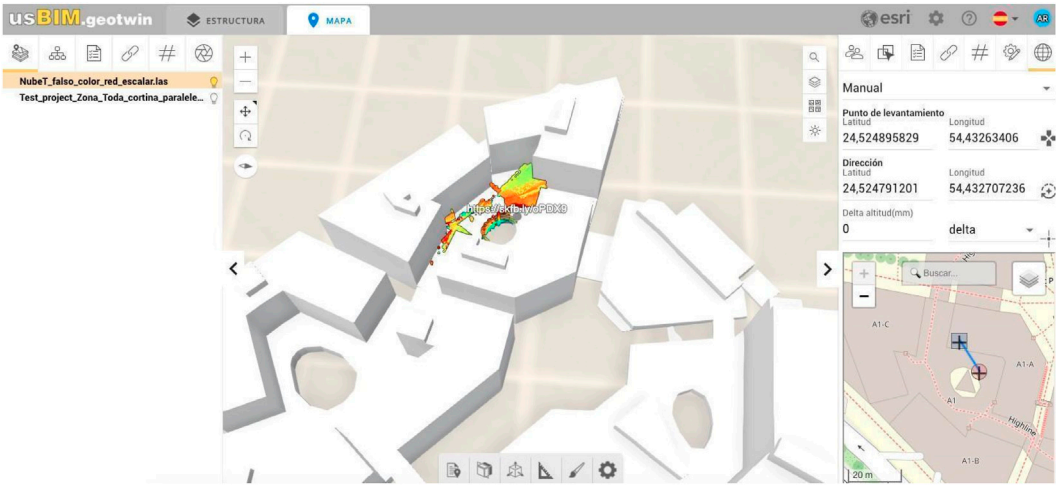


FIGURE 10
View of the model in ACCASoftware's GeoTwin tool.

1. Inserting reference to TPC into the GIS models. Complementarily, and in order to visualize the TPC in GIS, a "3D geometric element" type icon was created with a link to the *Sketchfab* platform, where one of the building's TPCs was uploaded. Since it was not possible to integrate the TPC directly using a.las file in *ArcGIS Pro* or its web version, this TPC can be viewed at the following link: https://skfb.ly/oPDX9.
2. Inserting TPC into *ACCASoftware's GeoTwin* tool. The aforementioned GIS model has been used as a model and base map in *ACCASoftware's GeoTwin* tool, where the thermal point cloud obtained from the case study has been added in.las format after being filtered for outliers. However, there are still problems in the visualization of this type of thermal models, such as the impossibility of placing scales that allow selecting a point and obtaining the value of its attribute (in our case, the temperature) natively in the application itself, without having to manually create a legend. In this case, the platform allows loading a.las model of the point cloud and geolocating it on the map (Figure 10) but does not allow interaction with it beyond being able to view it (rotate the view, zoom in and out).

5.2 Discussion about the software tools used

The tools used in this approach *include Autodesk Revit, Esri ArcGIS Pro 3.5, ArcGIS Pro Online,* and *GeoTwin from ACCASoftware.* As in the previous section, a brief discussion about the software used is included in this section. Several

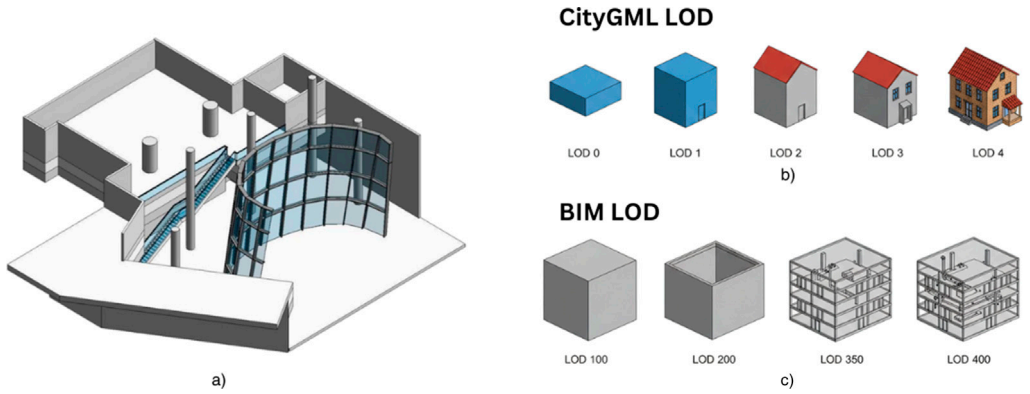


FIGURE 11
(a) 3D model made in *Autodesk Revit* of the double-height space studied and the façade of the interior courtyard.; **(b)** CityGML LOD, where LOD0 refers to plan views and LOD4 covers interior layout and furniture; **(c)** BIM LOD, where LOD200 is simple volumes and LOD400 includes the installations.

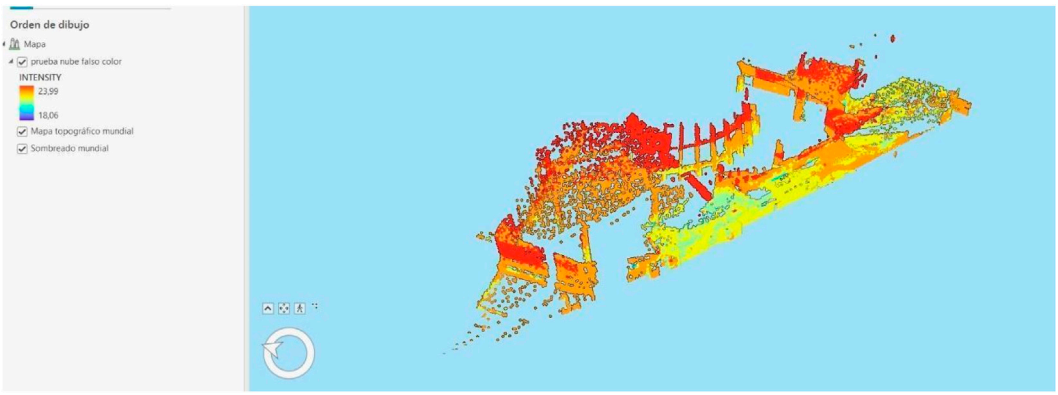


FIGURE 12
View of a.las file with "Intensity" values corresponding to the scalar value of the temperature at each point in *ArcGIS Pro* desktop.

important details and valuable extra information can help the reader to better understand the methodology defined in this process.

The tools were used thanks to the technical support provided to the university, along with their advanced features compared to the open-source alternatives available. The tools used include not only data visualization but also the availability of toolboxes with various functionalities for the implementation of digital twins and the integration of data from the different external sources that were collecting data while the thermal point clouds were being taken.

- Use of *Autodesk Revit*: *Autodesk Revit* was used to manually obtain the geometric model from the original TPC and to add location data of the building (Figure 11a). The software was used to achieve a Level of Detail (LOD) 4 of the building studied, according to the City Geography Markup Language (CityGML) standard, which includes the interior distribution of the spaces. There are differences between the GIS and BIM LOD, as shown in Figure 11b. A LOD 4 in BIM (Figure 11c) might include very specific details, such as individual wall components or fixtures, whereas a LOD 4 in CityGML includes the interior layout of the building, but often with less construction detail than a BIM model at the same level. This can lead to inconsistencies when trying to integrate or compare data between the two systems.
- Use of *Esri ArcGIS Pro*: Point clouds can be included in the desktop tools as.las files, but, so far, their display is limited, and there are important restrictions when working with this type of file in this environment (for example, colour scale values cannot be chosen, clouds cannot be clipped, different clouds cannot be superimposed), as shown in Figure 12. The resulting.rvt model was exported directly to *ArcGIS Pro* on the desktop in order to serve as a base model where the local sensors can be placed as geometric elements in *ArcGIS* as was explained in Section 4.1 Methodology.
- Use of Esri ArcGIS Pro Online: To contextualize the model within the campus, open data from *OpenStreetMap* (OSM) was incorporated into the workflow. This complementary information was essential to accurately model the campus



FIGURE 13
(a) Model view alongside data provided by *OpenStreetMap* in *ArcGIS Online*; **(b)** Model view with Esri online viewer.

layout and locate the monitored buildings within it. The integration of these tools allows for a comprehensive and accurate representation of the built environment, as seen in Figure 13a, facilitating detailed analysis within the framework of thermal digital twins for better decision-making.

- ArcGIS Pro Online software was used to integrate all the generated information and upload it to the Esri online platform, allowing data to be centralized and remote access to be facilitated. On this platform, the Building Explorer interactive viewer, included in the Indoors plugin, was used, offering tools to explore and manipulate the created BIM-GIS models. Users can change viewpoints, adjust the scale (from a global view to specific building details), activate or deactivate floors, and visualize data from the sensors. This allows for interaction with the models, as shown in Figure 13b.
- Use of ACCASoftware's GeoTwin: *ACCASoftware's GeoTwin* tool was used as an alternative to display the TPC on an online platform different from Esri's ArcGIS, where data in different formats can be included. In this case, the *Geotwin* application was used, which is designed to load other types of data onto a georeferenced map based on *ArcGIS* that this platform does not fully display, such as point clouds or IFC models that can be updated continuously. The TPC was loaded onto this platform using the model created in ArcGIS, as shown in Figure 14.

6 Limitations

Although based on a single case study, the workflow exposes recurring interoperability gaps that are likely to arise in similar architectural contexts. The implementation of methodologies for integrating thermal models in the construction sector faces several significant limitations that affect its efficiency and usability within professional settings in the AEC industry. In summary, these are:

1. Manual modelling dependency. The TPC processing requires specific programming skills, but the rest of the process remains manual and repetitive, involving the need to manually model over the point cloud to obtain the vertices of key construction

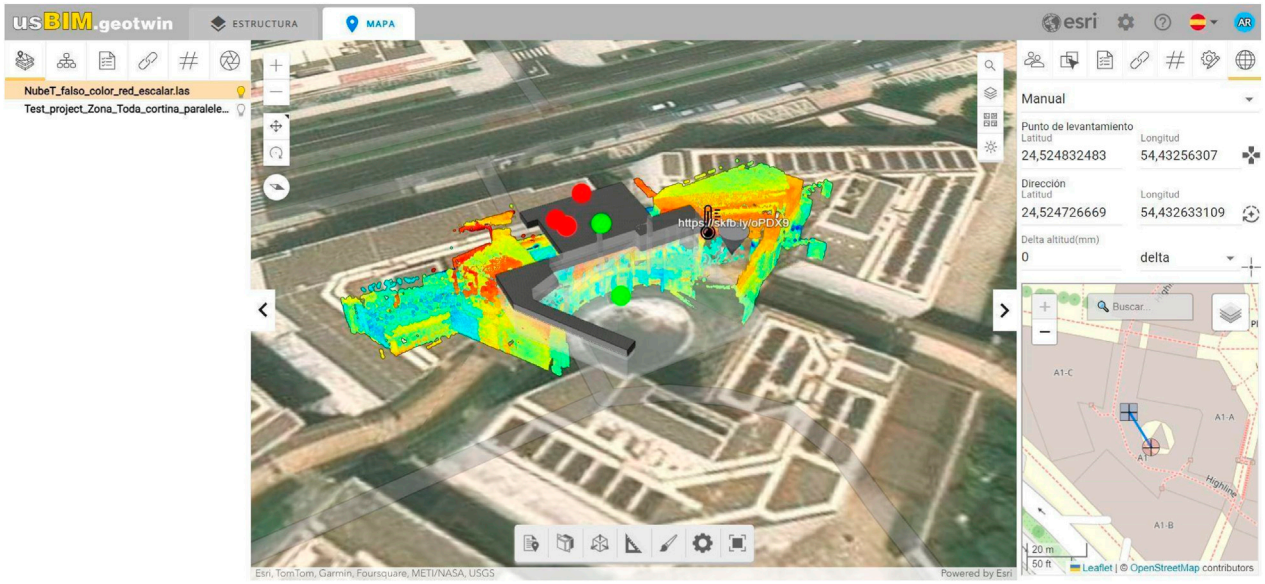


FIGURE 14
ACCASoftware Geotwin platform interface with TPC on the geolocated model.

elements and then use these coordinates within the thermal point cloud to define volumes. The average temperatures of the elements must be calculated through an automated programming process. Particularly in the case of buildings with irregular or non-Manhattan geometries, a great deal of manual work is required.

2. Absence of automatic geolocation in the TPC. The TPC loses its geolocation due to the way it is processed, so it is necessary to manually geolocate this thermal data in BIM and GIS software and platforms.

6.1 Working in complex buildings and structures

Working with such complex or large-scale buildings also continues to be a future challenge, as it requires lengthy data acquisition times, a significant deployment of computing resources, and a hard processing of the data. Additionally, numerous manual data segmentation and cleaning processes (outliers) are necessary. Although some scripting is needed for the IFC generation stage, the majority of operations remain visual and do not require advanced programming. The scripting components represent isolated bottlenecks that are not conceptually intrinsic to the methodology and are a target for future automation. The methodology is presented as a proof of concept demonstrating where automation is possible and where current tools still require manual effort.

1. Inability of GIS to handle non-standard point clouds. There is a major limitation regarding the capabilities of the platforms and software used (file size limitations, supported formats, how the models can be visualized. . .), both open-source and commercial, which are intended for the creation of digital models or the visualization of building data. In the case of software designed to work with BIM, it's possible to import and visualize point clouds with colour but not manipulate the data (add or change values, trim the cloud, add new clouds. . .). For thermal clouds, the visualization doesn't go beyond a false colour-temperature display.
2. Commercial tool reliance. In the case of GIS platforms, they are not equipped to handle non-standard, non-geolocated LIDAR point clouds without the metadata required by software or platforms such as *ArcGIS*. This workaround does not constitute full TPC integration. Its purpose is instead to demonstrate practical feasibility under current platform constraints.

7 Conclusion

The proposed proof of concept has proven to be effective addressing various aspects that had not been previously addressed in similar research, such as the use of thermal point clouds in large-scale buildings with several floors, the presence of a mostly glazed façade with complex geometry, the implementation of a sensor network in the as-is building and the subsequent integration of the collected data into a comprehensive BIM-GIS model that will serve as a basis for future developments. Temperatures have also been assigned to all the elements that make up the façade, which allows the creation of a data structure for a geometric-thermal model that can be consulted in the cloud.

This article also presents two TPC applications in the early stages of development for inclusion in architectural digital twins. TPCs are useful tools for understanding the thermal behaviour of an as-is space or infrastructure, and offer a comprehensive view for agents

involved in its preventive maintenance. Although, in a limited way, it is possible to work with these models in common formats in the architecture, engineering and construction industry, such as the IFC format. Although the use of TPC in GIS environments is interesting, there are still important limitations today in its implementation as a useful methodology for professionals in the AECO sector.

1. Regarding the sensing system: the employed method faces challenges when performing laser scans in short periods of time, which implies the need to generate TPC for long periods, especially in very large spaces that require between 1 and 3 h of scanning. Likewise, the lack of automation in the acquisition and processing methodology limits efficiency, especially in cases of buildings with multiple floors, since the lack of automated tools implies the need for manual processing in subsequent steps by a specialist. This stage involves manually obtaining average temperature values, which must then be applied to the BIM-GIS model.
2. Regarding the use of software and dependence on commercial software: most of the tools currently used are commercial, such as *Autodesk Revit* and *Esri ArcGIS Pro*, although free software tools are also used in intermediate processes, such as *CloudCompare*. There is a need to integrate more open-source tools into the entire workflow, such as *Blender* or *QGIS*, although this requires adequate programming skills that are not common for professionals engaged in preventive maintenance or architectural design. The physical deployment of media is also a challenge, as data collection in large-scale spaces demands significant resource mobilization, and the connection of sensors to power sources and wireless networks to download the data obtained can be limited, resulting in a lack of these for a complete quantitative analysis.
3. GIS limitations when handling point clouds: Regarding thermal visualization, although thermal data has been collected, its visualization and analysis in BIM and GIS environments still have limitations. The existing literature on thermal data visualization in BIM-GIS is scarce or non-existent, highlighting the need to develop and apply more advanced techniques in this field. Finally, the use of GIS has proven to be valuable in monitoring indoor spaces, especially for the precise location of elements such as sensors and the segmentation of buildings into smaller units such as structural elements or independent rooms, facilitating the creation of digital twins.
4. Lack of automatic geolocation of the TPCs: Although GIS is valuable for sensor location and building segmentation, there are still limitations in its ability to efficiently manage TPCs. Moreover, the lack of inherent geolocation data in the TPCs can limit their relevance in certain digital twin applications.

8 Future works

Progress in the integration of thermal models in the construction context requires addressing the identified limitations and improving the efficiency of several of the processes performed in the case studies. An important focus is the development of more automated methods, accessible to AECO industry stakeholders without specialized computer skills, for the conversion of TPCs into BIM models. It is important that these methods can deal with complex geometries and the diversity of elements and structures present in most current construction environments. Automating thermal Scan-to-BIM would not only streamline processes but could also make the integration of thermal models more accessible to a wider majority of professionals in the sector. In addition, the integration of thermal data from different sensor networks over time into the specific data structures of BIM formats needs to be explored. This integration will allow for continuous and detailed analysis of the thermal behaviour of buildings, which in turn will facilitate the identification of areas for energy improvement and the optimization of facility performance and comfort.

Regarding the integration of TPC into GIS, similar challenges are encountered in terms of data visualization and processing. Some of these are the precise geolocation of the positions from which the thermal point clouds have been taken, together with the inclusion of other geolocated thermal data, such as point clouds obtained from drone flights, and improvement in importing and integrating data coming from sensors of different types.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Author contributions

AR-C: Data curation, Formal Analysis, Methodology, Software, Validation, Visualization, Writing – original draft. FC: Conceptualization, Funding acquisition, Project administration, Resources, Supervision, Writing – review and editing. BGdS: Conceptualization, Funding acquisition, Project administration, Resources, Supervision, Validation, Writing – review and editing. AA-O: Conceptualization, Funding acquisition, Methodology, Project administration, Resources, Software, Supervision, Validation, Writing – review and editing.

Funding

The author(s) declared that financial support was received for this work and/or its publication. This publication is as part of the PID2022-141275OB-C21, CIN/AEI/10.13039/501100011033/FEDER, UE and SBPLY23180225000113 projects co-funded by FEDER (European Regional Development Fund), and 2022-GRIN-34261 project funded by Plan Propio de Investigación at Universidad Castilla-La Mancha. Part of this research was partially supported by various centers at NYU Abu Dhabi (NYUAD), specifically the Center for Sand Hazards and Opportunities for Resilience, Energy, and Sustainability (SHORES) and the Center for Interacting Urban Networks (CITIES). These centers are funded by Tamkeen under the NYUAD Research Institute Awards CG013 and CG001, respectively.

Conflict of interest

The author(s) declared that this work was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declared that generative AI was not used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Adamopoulos, E., Patrucco, G., Volinia, M., Girotto, M., Rinaudo, F., Giulio Tonolo, F., et al. (2021). "3D thermal mapping of architectural heritage: up-to-date workflows for the production of three-dimensional thermographic models for built heritage NDT," in *Lecture notes in computer science (including subseries lecture notes in artificial intelligence and lecture notes in bioinformatics)* (Cham: Springer), 26–37. doi:10.1007/978-3-030-73043-7_3

Adán, A., Quintana, B., García Aguilar, J., Pérez, V., and Castilla, F. J. (2020). Towards the use of 3D thermal models in constructions. *Sustain. Switz.* 12 (20), 1–13. doi:10.3390/su12208521

Adán, A., Ramón, A., Vivancos, J., Vilar, A., and Aparicio-Fernández, C. (2023c). Automatic generation of as-is BEM models of buildings. *J. Build. Eng.* 73, 106865. doi:10.1016/J.JOBE.2023.106865

Adán, A., López-Rey, A., and Ramón, A. (2023a). "Obtaining 3D dense thermal models of interiors of buildings using Mobile robots," in *ROBOT2022: fifth iberian robotics conference* (Springer International Publishing), 3–14. doi:10.1007/978-3-031-21065-5_1

Adán, A., López-Rey, A., and Ramón, A. (2023b). Robot for thermal monitoring of buildings. *Automation Constr.* 154, 105009. doi:10.1016/j.autcon.2023.105009

Alba, M. I., Barazzetti, L., Scaioni, M., Rosina, E., and Previtali, M. (2011). Mapping infrared data on terrestrial laser scanning 3D models of buildings. *Remote Sens.* 3 (9), 1847–1870. doi:10.3390/rs3091847

Albourae, A. T., Armenakis, C., and Kyan, M. (2017). Architectural heritage visualization using interactive technologies. *Int. Archives Photogrammetry, Remote Sens. Spatial Inf. Sci. - ISPRS Archives*, 7–13. doi:10.5194/isprs-archives-XLII-2-W5-7-2017

Amirebrahimi, S., Rajabifard, A., Mendis, P., and Ngo, T. (2016). A BIM-GIS integration method in support of the assessment and 3D visualisation of flood damage to a building. *J. Spatial Sci.* 61 (2), 317–350. doi:10.1080/14498596.2016.1189365

Antón, D., and Amaro-Mellado, J. L. (2021). Engineering graphics for thermal assessment: 3D thermal data visualisation based on infrared thermography, GIS and 3D point cloud processing software. *Symmetry* 13 (2), 1–20. doi:10.3390/sym13020335

Back, Y., Kumar, P., Bach, P. M., Rauch, W., and Kleidorfer, M. (2023). Integrating CFD-GIS modelling to refine urban heat and thermal comfort assessment. *Sci. Total Environ.* 858, 159729. doi:10.1016/j.scitotenv.2022.159729

Bottaccioli, L., Aliberti, A., Ugliotti, F., Patti, E., Osello, A., Macii, E., et al. (2017). "Building energy modelling and monitoring by integration of IoT devices and building information models," in *Proceedings - International computer software and applications conference* (IEEE), 914–922. doi:10.1109/COMPSAC.2017.75

Bracht, M. K., Melo, A. P., and Lamberts, R. (2021). A metamodel for building information modeling-building energy modeling integration in early design stage. *Automation Constr.* 121, 103422. doi:10.1016/j.autcon.2020.103422

Deng, Y., Cheng, J. C. P., and Anumba, C. (2016). Mapping between BIM and 3D GIS in different levels of detail using schema mediation and instance comparison. *Automation Constr.* 67, 1–21. doi:10.1016/j.autcon.2016.03.006

Desogus, G., Frau, C., Quaquero, E., and Rubiu, G. (2023). From building information model to digital twin: a framework for building thermal comfort monitoring, visualizing, and assessment. *Buildings* 13 (8), 1971. doi:10.3390/buildings13081971

Dlesk, A., Strogonov, V., Vach, K., and Pollert, J. (2023). Integration of spatially oriented data from iot to facility management and bim. *Int. Archives Photogrammetry, Remote Sens. Spatial Inf. Sci. - ISPRS Archives*, 31–36. doi:10.5194/isprs-archives-XLVIII-5-W2-2023-31-2023

Escudero, P. A., González-López, C., and García-Valldecabres, J. L. (2024). Methodology for data integration in 3D-HBIM digital models. *Case Study The Holy Chalice Chapel Valencia Cathedr.* doi:10.20365/DISEGNARECON.32.2024.11

García-Valldecabres, J. L., Liu, J., Willkens, D. S., Escudero, P. A., López-González, C., Cortés Meseguer, L., et al. (2023). Development of a virtual itinerary with hbim and gis. *Int. Archives Photogrammetry, Remote Sens. Spatial Inf. Sci.*, 645–652. doi:10.5194/isprs-archives-xlviii-m-2-2023-645-2023

Göçer, Ö., Hua, Y., and Göçer, K. (2016). A BIM-GIS integrated pre-retrofit model for building data mapping. *Build. Simul.* 9 (5), 513–527. doi:10.1007/s12273-016-0293-4

Hong-mei, Z., Cheng-hu, Z., Wei-qiang, G., and Jin-cai, D. (2001). Spatial distribution characteristics of urban thermal conditions: application of GIS and remote sensing. *J. Geogr. Sci.* 11 (3), 261–270. doi:10.1007/bf02892309

Hou, Y., Volk, R., Chen, M., and Soibelman, L. (2021). Fusing tie points' RGB and thermal information for mapping large areas based on aerial images: a study of fusion performance under different flight configurations and experimental conditions. *Automation Constr.* 124, 103554. doi:10.1016/j.autcon.2021.103554

Isikdag, U., Zlatanova, S., and Underwood, J. (2013). A BIM-oriented model for supporting indoor navigation requirements. *Comput. Environ. Urban Syst.* 41, 112–123. doi:10.1016/j.compenvurbsys.2013.05.001

Kamel, E., and Memari, A. M. (2019). Review of BIM's application in energy simulation: tools, issues, and solutions. *Automation Constr.* 97, 164–180. doi:10.1016/j.autcon.2018.11.008

Lagüela, S., Díaz-Vilariño, L., Roca, D., and Armesto, J. (2014). "Aerial oblique thermographic imagery for the generation of building 3D models to complement geographic information systems," in *Proceedings of the 2014 international conference on quantitative InfraRed thermography* (Bordeaux, France: QIRT Conference). doi:10.21611/qirt.2014.041

Lin, D., Jarzabek-Rychard, M., Tong, X., and Maas, H. G. (2019). Fusion of thermal imagery with point clouds for building façade thermal attribute mapping. *ISPRS J. Photogrammetry Remote Sens.* 151, 162–175. doi:10.1016/j.isprsjprs.2019.03.010

López-González, C., and García-Valldecabres, J. (2023). The integration of HBIM-SIG in the development of a virtual itinerary in a historical centre. *Sustainability* 15 (18), 13931. doi:10.3390/su151813931

Mutani, G., Delmastro, C., Gargiulo, M., and Corgnati, S. P. (2016). Characterization of building thermal energy consumption at the urban scale. *Energy Procedia* 101 (September), 384–391. doi:10.1016/j.egypro.2016.11.049

Niu, S., Pan, W., and Zhao, Y. (2015). A BIM-GIS integrated web-based visualization system for low energy building design. *Procedia Eng.* 121, 2184–2192. doi:10.1016/j.proeng.2015.09.091

Oti, A. H., Kurul, E., Cheung, F., and Tah, J. (2016). A framework for the utilization of building management system data in building information models for building design and operation. *Automation Constr.* 72, 195–210. doi:10.1016/j.autcon.2016.08.043

Previtali, M., Erba, S., Rosina, E., Redaelli, V., Scaioni, M., and Barazzetti, L. (2012). "Generation of a GIS-based environment for infrared thermography analysis of buildings," in *Infrared remote sensing and instrumentation*. Editors M. Strojnik and G. Paez, XX, 85110U. doi:10.1117/12.930050

Rahhal, A., Matthys, C., Ben Rajeb, S., and Leclercq, P. (2019). "Limitations of using digital BIM models to carry out thermal analysis," in *Fassi 2019: fifth international conference on fundamentals and advances in software systems integration*. [Preprint].

Ramón-Constantí, A., Adán Oliver, A., Quintana Galera, B., Castilla Pascual, F. J., and Pérez-Andréu, V. J. (2022a). "Technology and methodology for obtaining omnidirectional 3D thermal point clouds models of buildings," in *26th international congress on project management and engineering terrassa*.

Ramón-Constantí, A., Adán-Oliver, A., and Castilla-Pascual, F.-J. (2022b). Thermal point clouds of buildings: a review. *Energy Build.* 274, 112425. doi:10.1016/j.enbuild.2022.112425

Ramón-Constantí, A., Adán-Oliver, A., Castilla-Pascual, F. J., and Pérez-Andreu, V. (2024). An experimental methodology for the calibration of indoor building environment models using thermal point clouds and CFD simulation. *Adv. Build. Energy Res.* 18 (3), 261–294. doi:10.1080/17512549.2024.2358923

Tomchinskaya, T., and Galanina, M. (2020). "University management system based on bim-gis-technologies," in *CEUR workshop proceedings*. doi:10.51130/graphicon-2020-2-4-26

frontiers | Frontiers in Built Environment

TYPE Original Research
PUBLISHED 09 February 2026
DOI 10.3389/fbuil.2026.1723491

Check for updates

OPEN ACCESS

EDITED BY
Sakdirat Kaewunruen,
University of Birmingham, United Kingdom

REVIEWED BY
Jiaqi Suo,
Purdue University, United States
Rana Muhammad Irfan Anwar,
Auburn University, United States

*CORRESPONDENCE
Sameer Prabhu,
Sameer.prabhu@ltu.se

RECEIVED 12 October 2025
REVISED 26 December 2025
ACCEPTED 05 January 2026
PUBLISHED 09 February 2026

CITATION
Prabhu S, Patwardhan A and Karim R (2026) Artificial intelligence-driven safety assessment of scaffolding using LiDAR sensing.
Front. Built Environ. 12:1723491.
doi: 10.3389/fbuil.2026.1723491

COPYRIGHT
© 2026 Prabhu, Patwardhan and Karim. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Artificial intelligence-driven safety assessment of scaffolding using LiDAR sensing

Sameer Prabhu*, Amit Patwardhan and Ramin Karim

Department of Civil, Environmental and Natural Resources Engineering, Division of Operation and Maintenance Engineering, Luleå University of Technology, Luleå, Sweden

The construction industry is embracing transformation through the integration of digitization, artificial intelligence (AI), and immersive technologies. On a construction site, continuous assessment is vital for ensuring both the reliability of assets and safety of workers. Scaffolding is a key structural support asset that requires regular inspections for detection and identification of alterations from the design rules that could compromise integrity and stability. At present, such inspections to identify deviations are primarily visual and conducted by the site managers or accredited personnel. However, visual inspection is time-intensive and susceptible to human errors, which can lead to unsafe conditions. This study explores the use of AI and digital technologies to automate and enhance scaffolding inspections process to contribute toward safety improvement. A cloud-based AI platform is developed to process and analyze 3D point-cloud data of scaffolding structures to detect modifications through comparisons as well as evaluate the certified reference scan with a recent scan. The proposed workflow incorporates prognostics and health management concepts with continuous monitoring to identify structural modifications and further assist with decision-making. The results indicate that the proposed approach can limit reliance on manual visual inspections. By enabling automated monitoring of scaffolding, the proposed approach reduces the time and effort required for inspection process, while enhancing the safety on a construction site.

KEYWORDS

3D point-cloud analysis, artificial intelligence, cloud-based monitoring platform, construction safety, prognostics and health management, scaffolding inspection

1 Introduction

Ensuring workers' safety on a construction site is a crucial concern that requires regular inspection and maintenance of the structural assets to avoid accidents. In these instances, site managers play key roles in making critical decisions that impact safety and operational efficiency. According to the accident causation model (Heinrich, 1941), accidents can be attributed to unsafe conditions and unsafe actions. Monitoring these two causes can reduce the safety risks on a construction site. Given the dynamic working conditions and complex nature of a construction site, workers are often progressively exposed to hazardous environments. Identifying the causes of unsafe conditions, implementing mitigatory actions, performing equipment inspections, and conducting safety evaluations are the main activities involved in safety and health monitoring (Reese and Eidson, 2006). Accidents occurring on construction sites may frequently result in fatalities, damage to assets, injuries to workers, financial losses, and delays (Hamdan and Awang, 2015).

Accidents are caused by various factors, including construction errors, insufficient protection equipment, inadequate technical specifications, excessive loads on scaffoldings, non-compliant components, improper physical actions, and distraction (Whitaker et al., 2003; Ismail and Ghani, 2012). The factors influencing accidents on construction sites are susceptible to changes over time (Hamdan and Awang, 2015). Scaffolding is a critical asset at a construction site that concerns safety-related risks. The majority of construction on a project is performed using scaffolding that comprises steel tubes and joints, where the tube axes describe the spatial configuration. Scaffolds serve as temporary structures that facilitate access to elevated areas during and after construction. The collapse of scaffolding can result in loss of life and financial costs and therefore requires strict compliance with safety protocols (Swedish Work Environment Authority, 2016). Regular inspections and maintenance are necessary to ensure good working conditions and safety for the workers operating at heights as well as prevent accidents involving pedestrians around the scaffolding.

Site managers play essential roles in coordinating and supervising operations at construction sites, ensuring the safety of workers, and completing projects in a timely manner while adhering to the budget. Site managers are also accountable for inspecting and monitoring the scaffolding. At present, scaffolding inspections rely mainly on visual procedures (Cho et al., 2018a), which are inefficient and may lead to errors, especially at large and dynamic construction sites. During construction, the scaffolding structure often undergoes modifications, which may sometimes lead to deviations from standard guidelines and necessitate regular inspections. Such modifications may result in missing or incorrectly replaced scaffolding components that could impact stability. Continuous monitoring of scaffolding to detect modifications and verify compliance requires investment of time and effort by the site manager. As construction operations progress and particularly at complex sites, manual inspection of scaffoldings becomes challenging and inaccurate (Wang, 2019), which may pose potential threats to the safety of workers and individuals situated on and around the scaffolding. Comparison of the scaffolding structure with a reference structure created according to established design rules is required to identify modifications. Additionally, periodic maintenance is needed to restore structural integrity. Beyond traditional inspection methods, there is rapid evolution in the field of digital transformation in construction. The integration of digital twin (DT) and building information modeling (BIM) with immersive technologies is changing the process of monitoring and managing structures in real time. Farsangi et al. (2024) discussed the transition from BIM to DT and the metaverse as well as highlighted the associated challenges and opportunities. Additionally, Ammar et al. (2022) examined the role of the DT in supporting transparency and enabling real-time feedback to enhance collaboration in construction projects. Thus, embedding scaffolding inspection into a digital ecosystem enables continuous data-driven monitoring that surpasses the limitations of periodic visual assessments.

Prognostics and health management (PHM) provides a framework focusing on the degradation mechanisms of an asset to forecast its health and enhance maintenance optimization. The PHM framework ensures optimal functioning, enhanced reliability, and minimal maintenance cost of an asset (Soualhi et al., 2022). It involves steps like anomaly detection, where an alert is generated whenever the acquired data deviate from the required or desired performance. Later, diagnostics are performed to determine the probable causes for the deviations in terms of location and nature. Finally, given the current health and degree of degradation, prognostics is used to estimate the time when an asset will no longer be able to perform its intended functions (Soualhi et al., 2022). Figure 1 shows the workflow of the proposed scaffolding inspection system, which utilizes the concepts of PHM for asset monitoring and health management. The process begins with data acquisition, where 3D point-cloud data scans of the scaffolding structure are collected. When scaffolding is installed by experts, an initial scan is performed that serves as a reference or certified scan. During structural health monitoring, recently acquired data are continuously compared with the certified or reference data to highlight potential modifications in the scaffolding structure. If a structural modification is detected that could cause a potential safety concern, the system alerts the site manager. This workflow enables automated inspection of scaffolding to potentially reduce the repeated time and effort required from the site manager. Furthermore, the errors caused by manual visual inspections can be minimized, thereby assisting site managers in their decision-making.

The key contributions of this study are as follows:

1. A cloud-based artificial intelligence (AI) platform is developed to process, analyze, and visualize scaffolding structure, thereby enabling automated inspection and decision-making.
2. A method for monitoring scaffolding using 3D point-cloud data is proposed, which facilitates detection of structural changes impacting stability and enhances safety at the construction site.
3. An approach representing the scaffolding using a graph data structure is presented, which enables the development of a design rule engine for compliance checking.

The remainder of this paper is structured as follows. Section 2 presents literature reviews on traditional and current approaches for scaffolding inspection as well as highlights the research gap. Section 3 presents the methodology of the proposed approach, and Section 4 presents the results. Section 5 includes the conclusions and future directions of this work.

2 Related work on scaffolding health monitoring

The digitalization of scaffolding inspection intersects with advances in sensor technologies, AI-driven analyses, and digital transformation in construction. To overcome the limitations of manual inspection, extant research on scaffolding inspections has adopted a range of technological approaches. The related works are further grouped under two categories as follows: (i) scaffolding inspection approaches that are not based on light detection and ranging (LiDAR) sensing technologies; (ii) LiDAR-based approaches using 3D point-cloud data for geometric and structural analyses.

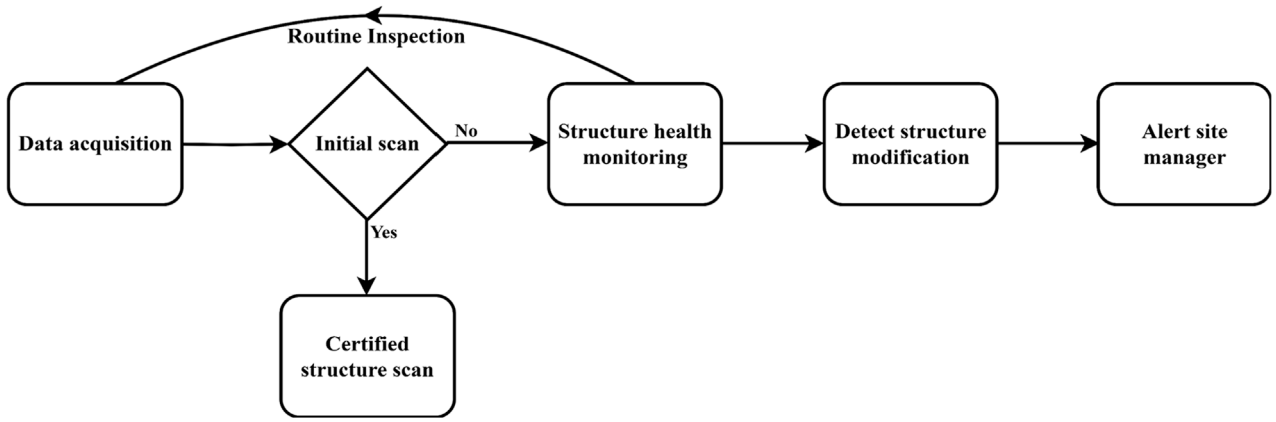


FIGURE 1
Workflow of the proposed scaffolding inspection system.

2.1 Scaffolding inspections using non-LiDAR sensors and data-driven methods

Construction workers are often exposed to safety risks as the health assessments of scaffolding structures are still in the premature stages of development (Cho et al., 2018b). The variability of the design parameters depending on the load conditions, boundary conditions, and materials makes monitoring of scaffolding difficult. Cho et al. (2018b) employed wireless strain sensors to collect data from a scaffolding prototype and transmit them to a finite element machine (FEM) model that assesses the real-time structural conditions; these authors also explored a comprehensive approach to scaffold monitoring that integrated aspects like strain data, FEM, machine learning, and scaffolding (Cho et al., 2018a). The structural state was assessed using a strain gauge, and a data-driven approach was implemented to monitor the scaffolding. The FEM model was designed to replicate the structural response for a given load condition; hence, the training data for the machine learning model were generated using random loading scenarios. Finally, the structural condition was classified using a machine learning algorithm as safe, overloaded, unevenly settled, or risk of overturning. Machine learning has been highlighted as increasingly important to structural health monitoring, especially in support of damage detection and decision-making (Malekloo et al., 2022). Magnetic-flux-leakage-based 3D defect visualization was also employed by some researchers to detect the metal loss in scaffolding steel pipes using a custom Hall-sensor prototype (Yigzew et al., 2024).

Given the rapid developments in the field of data-driven approaches, advanced deep-learning architectures and generative AI methods have been leveraged for condition assessments (Wang et al., 2024; Shao et al., 2025). Scaffolding inspections are generally labor-intensive and subjective, sometimes resulting in inconsistencies. Computer vision and machine learning have been used to develop automated inspection processes (Kim A. et al., 2022). Both real and synthetic labeled datasets have been employed to train machine learning models to detect various conditions like rust, cracks, and deformations on scaffolding components. BIM and image processing methods have been combined to estimate scaffolding progress through comparisons with site photographs (Chi et al., 2017). Safety monitoring models are frequently restricted to specific types of hazards; however, the dynamic nature of a construction site makes it impractical to detect hazards by type (Khan et al., 2021). Scaffolding inspections are intended to ensure structural stability and workers' safety. Furthermore, accidents may occur due to the workers' actions while on the scaffolding. To address worker-related risks on scaffolding, one approach proposed to detect unsafe actions by workers while situated on a mobile scaffolding by analyzing correlations drawn from safety rules. A deep-learning model was implemented for object detection and instance segmentation on manually labeled images of scaffolding with and without outriggers. The unsafe actions by workers were identified by integrating the classification and segmentation of workers' tasks with an object correlation detection module. This module determined the correlation between the predicted task and specific conditions (Khan et al., 2021). A related approach was proposed to mitigate the risk of construction-related falls from a height through a smart safety hook monitoring system; this system integrates computer vision with Internet of Things monitoring technologies and employs deep learning to identify workers and scaffolding from live camera feed. Upon detecting a worker in a risk zone, an Arduino Nano equipped with an inertial measurement unit and an altimeter sensor is activated to check the status of the safety hook (Khan et al., 2022).

The aforementioned approaches have been mostly implemented and evaluated on either scaffolding prototypes or parts of the scaffolding. Moreover, these models rely significantly on the quality of the collected image data, which can be affected by lighting conditions and image acquisition angles. Machine learning models require labeled data for training, which can be labor-intensive, and the model evaluations depend on their ground truths. All of these approaches require additional development and validation prior to real-world application.

2.2 LiDAR-based approaches

Among the sensor-based inspection technologies, LiDAR is a unique and powerful modality owing to its ability to acquire dense 3D geometric data (Wang and Kim, 2019). LiDAR or 3D laser scanning emits laser beams toward a target and detects the reflected signals to determine the distance. Scanners utilizing the time-of-flight principle emit laser pulses with predetermined velocities and measure the time elapsed before receiving the reflected pulses to estimate the target distances. LiDAR creates a 3D map of the physical environment by measuring distances and spatial interactions among the objects. 3D point-cloud data obtained by laser scanning technology can accurately and efficiently capture spatial structures (Wang et al., 2020). This technology allows high frame rates, an extended field of view, and the ability to operate under various lighting conditions, making it suitable for dynamic and complex construction environments (Teizer, 2008).

3D point-cloud data have emerged as invaluable tools in construction, offering applications to improve efficiency and accuracy through progress tracking as well as enhancing safety throughout the project lifecycle via quality inspections (Wang and Kim, 2019). Displacement mapping methods have been developed to analyze sequential point clouds of retaining structures (Zhao et al., 2021), demonstrating efficient detection of deformations and safety risks. LiDAR technology has been employed extensively for as-built inspections and building quality assessments (Styliadis, 2007; Bosché, 2010; Aydin, 2014; Yu-Fei et al., 2016). One study focused on deriving pictorial, geometric, spatial, topological, and semantic information from architectural objects like historical buildings and monuments; this work presented a practical demonstration of generating semantic models from high-resolution images to produce point-cloud data through calibration and alignment; the semantic models were subsequently integrated with computer-aided architectural design systems for advanced analysis and documentation (Styliadis, 2007). Another study employed high-resolution digital cameras and photogrammetry techniques to remodel building facades that were adversely affected by visual pollution owing to urban development (Aydin, 2014). Integration of 3D computer-aided design (CAD) models with laser scanning technologies has been proposed to address issues like progress tracking and quality control; in this approach, the CAD models are optimally registered with site scan objects to subsequently compute the as-built poses for verifying dimensional compliance with the specifications (Bosché, 2010). In a related approach, crack information was extracted from 2D images and reconstructed into 3D scenes using the structure-from-motion (SfM) algorithm by combining image processing and 3D modeling for structural health assessment (Yu-Fei et al., 2016).

One study aimed to reduce the time-consuming and error-prone manual inspections of highway retaining walls by extracting geometric features from laser scan data to identify displacements in mechanically stabilized earth (MSE) walls; here, the horizontal joints are extracted from point-cloud data of the MSE walls obtained at different time intervals that serve as the benchmarks for displacement measurements (Oskouie et al., 2016). Existing approaches for safety planning at excavation sites susceptible to cave-ins and falls are manual and prone to errors; to address these limitations, a semi-automated approach was developed that utilizes the geometric characteristics of 3D point-cloud data to detect fall and cave-in hazards; to enhance safety planning, the hazards identified from point-cloud data and necessary safety rules were integrated with BIM (Wang et al., 2015). Active safety systems capable of real-time hazard detection have been explored to improve construction site safety, including the use of 3D range cameras for accident prevention; furthermore, algorithms have been developed to enhance the abilities of the 3D range cameras for detecting and tracking objects, such that real-time feedback can be provided to operators to enhance safety during heavy machinery operation (Teizer, 2008).

Scaffolding-related issues are a primary source of injuries and accidents at construction sites, so safety features like toe boards and guard rails are provided to avoid falls and other related injuries. Wang (2019) proposed an automated inspection method to ensure compliance with safety regulations. Here, upon detecting the locations of the vertical elements of scaffolding from point-cloud data, the planar and horizontal surfaces are identified for scaffolding platforms; the four sides of the detected work platforms are then considered for establishing toe boards and guard rails. Next, safety regulation assessments based on the heights of the toe boards as well as numbers and locations of the guard rails are conducted on the extracted toe-board and guard-rail components; however, this method primarily focuses on the toe boards and guard rails rather than the complete scaffolding structure (Wang, 2019). A 3D deformation monitoring model has been proposed for scaffolding to address limitations associated with the restricted scanning range of LiDAR using point-cloud data acquired from multiple stations, which is referred to as multithread LiDAR technology. The proposed model comprises point-cloud alignment and tube-axis modeling, where the geometric relationships of the planar features in the scans are used to compute the transformation parameters for point-cloud alignment; for scaffolding tube-axis modeling, the model integrates a least-squares approach with the RANdom SAmple Consensus (RANSAC) algorithm. The scaffolding deformations are monitored by comparing the tube-axis models at various time intervals; however, this alignment strategy requires the presence of sufficient planar surfaces in the scanned data (Rui et al., 2022). Kim et al. (2022b) proposed a 3D reconstruction approach for scaffolding monitoring using a semantic segmentation method called RandLA-Net, which detects scaffolding from point-cloud data collected using a robotic dog; this deep-learning model is trained from scratch as well as through transfer learning with the Semantic3D dataset. The 3D CAD model generated from the predicted scaffolding point cloud is more focused on the identification and 3D reconstruction of scaffolding at the construction site. For a large construction site, it is computationally inefficient to directly process point clouds to identify small objects. Accordingly, a methodology that integrates 3D point-cloud data with 2D image data has been reported for detecting and locating unsafe scaffold joints. In this approach, 3D semantic segmentation was applied to isolate scaffolding from the acquired point-cloud data, after which the scaffold joints were identified using the coordinates of uprights and guard rails to

TABLE 1 Previous studies on scaffolding inspections using point-cloud data.

Study	Objective	Methodology	Data acquisition	Focus
Wang (2019)	Detection of toe boards and guard rails	Using the geometric properties of point-cloud data, the four sides of a work platform were extracted after detecting the uprights (vertical) and platform (horizontal)	Terrestrial laser scanner	Toe boards and guard rails
Rui et al. (2022)	Monitoring scaffolding for deformations	Point-cloud data were aligned to compare the tube-axis model for deformations	Multithread LiDAR	Planar surfaces for alignment
Kim et al. (2022b)	Scaffolding extraction	3D semantic segmentation model was implemented to identify scaffolding	Mobile laser scanner	Scaffolding identification
Kim et al. (2023)	Inspection of scaffold joints	Semantic segmentation was performed to extract scaffolding, followed by point-to-image translation of the joints for safety inspection	Terrestrial laser scanner	Scaffold joints

generate joint images. Then, a deep-learning model was trained to detect ledger ends and tails from images of scaffolding joints to provide insights into the safety status of these joints (Kim et al., 2023).

To the best of our knowledge, there is limited research on integrating scaffolding monitoring with point-cloud data. This was also noted by Rui et al. (2022), who identified that there was a lack of studies employing point-cloud data for deformation monitoring of temporary structures. Table 1 demonstrates that extant studies incorporating point-cloud data for scaffolding are limited to either specific parts of the scaffold or detection of scaffolding. Hence, comprehensive methods for monitoring complete scaffolding structures using point-cloud data are lacking.

The aim of the present study was to explore the use of AI and digital technologies for scaffolding inspections to assist site managers with decision-making related to the safety of construction sites. Accordingly, our objective was to develop a cloud-based AI-driven platform that automates monitoring and safety assessments of scaffolding structures using 3D point-cloud data. The proposed approach is expected to contribute toward the ongoing digital transformation of the construction industry.

3 Methodology

This work focuses on identifying modifications to the scaffolding during the construction process. As discussed in Section 1, routine manual visual inspection of scaffolding is time-consuming and may be prone to errors, especially on complex sites and owing to human fatigue. The following insights are based on semi-structured interviews conducted with site managers, whose responsibilities include ensuring the safety of workers on construction sites. According to the managers, scaffolding inspections are repeated every few weeks, and the inspectors may struggle to recall specific changes made to the reference structures. The accuracy of inspection is also influenced by the inspector's familiarity with the design principles or rules. During construction, worker–structure interactions pose additional risks as workers occasionally remove the components or braces of the scaffolding for accessibility. The most common challenges encountered by site managers during scaffolding inspections are that the structural elements may not be replaced or that the replaced parts may not be compliant with the design rules. These issues can compromise the structural integrity and eventually affect worker safety.

Our research seeks to automate the routine inspection processes necessary to maintain the structural integrity of the scaffolding and ensure safe working conditions. We further explore the roles of AI and digital technologies to assist site managers with informed decision-making on construction sites. As illustrated in Figure 1, the proposed scaffolding monitoring framework is structured into three parts, namely, acquiring the certified or reference structure and campaign-based data, performing a routine inspection for structural comparison and modifications, and assisting the site manager with decision-making. Figure 2 adapted from Karim et al. (2016) presents the comprehensive platform used in this study that encompasses the entire workflow from data acquisition to knowledge extraction and ultimately visualization. The platform follows a microservice-based architecture to enable seamless data acquisition, secure transfer, storage, processing, analytics, and visualization.

The proposed framework is designed to support various sensing devices like LiDAR scanners and robotics platforms like autonomous robot dog and drones. However, in the present work, the point-cloud data are exclusively acquired using a handheld LiDAR scanner. The acquired data are subsequently transferred to a binary large object (BLOB) storage, which is a cloud-based storage mechanism designed to efficiently handle large unstructured binary files like LiDAR point clouds within the proposed workflow. The point-cloud data represent the spatial configuration of the environment and contain extensive information that requires further processing. As shown in Figure 2, the knowledge extraction stage encompasses various operations like data preprocessing, structural comparison, and graph data structure representation. Data cleaning by eliminating irrelevant data points like outliers and isolating the object of interest, namely the scaffolding structure, ensure that only high-quality information is retained for further analysis. The scaffolding extracted from the raw data still contains a large number of points that must be represented effectively. Since scaffolding can be considered the interconnection of rods and joints, we employed a graph data structure as an effective representation method. This representation is well-suited to capture the relationships between the nodes (scaffolding joints) and edges (scaffolding rods or braces). Furthermore, graph-based approaches not only preserve the structural relationships but also enable the use of well-established

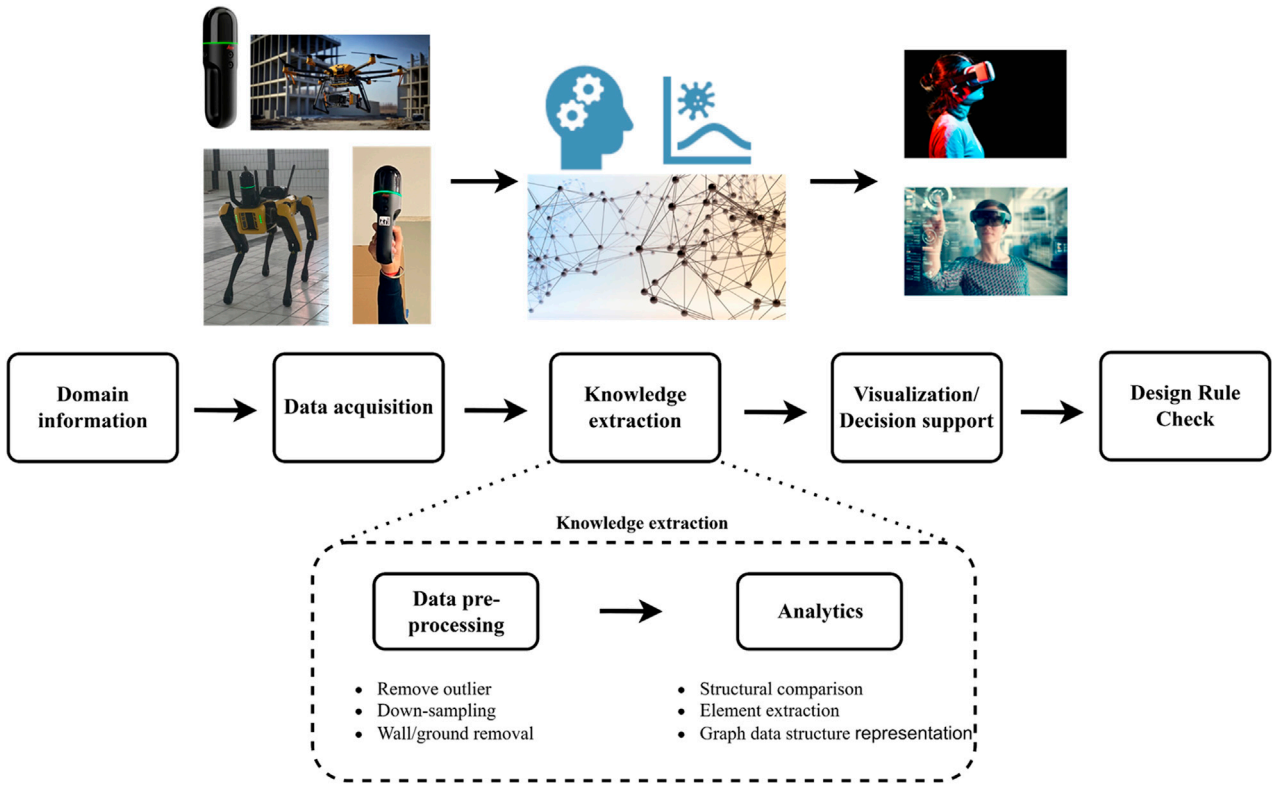


FIGURE 2
Artificial intelligence-based cloud platform for scaffolding inspection.

methodologies in network analysis to better understand the scaffolding structure. Structural changes occurring during the construction process can affect stability; therefore, the processed and extracted scaffolding point-cloud data are compared against the certified reference scaffolding design to detect structural changes or deviations. To this end, continuous monitoring of the scaffolding is essential so that any structural alterations can be highlighted and reported promptly. By drawing attention to these modified areas, the site manager can prioritize inspections to critical zones and implement targeted corrective actions, optimizing both time and resource allocations. Hence, digital technologies are valuable tools that can assist site managers in making informed decisions on the health of the scaffolding and safety of the workers. Finally, the visualization stage employs immersive technologies like augmented reality (AR) and virtual reality (VR) to present the extracted knowledge and support decision-making through intuitive interactions with the processed data. Prior to data acquisition, construction domain knowledge is incorporated into the platform to guide both data collection and processing. Although the platform developed herein focuses on data acquisition, knowledge extraction, and visualization, it can be extended to incorporate the design rule check module. Such an addition would validate the extracted scaffolding model against established design rules and regulatory requirements to strengthen compliance and enhance worker safety.

The point-cloud data uploaded to the BLOB storage serve as the primary inputs for all subsequent processing steps. All data preprocessing and analytical algorithms, including noise removal, downsampling, structural comparisons, and graph-based scaffolding representation, are executed on cloud infrastructure as part of the proposed platform. These algorithms are applied directly to the stored .e57 point-cloud files. The cloud-based processing steps are triggered upon the arrival of a new point-cloud data file to initiate the knowledge extraction stage. The processed results are subsequently reused at the visualization stage, including AR- and VR-based interfaces. Overall, this work presents a computational intelligence system that transforms raw sensory inputs into structured representations, evaluates structural integrity through algorithmic inference, and generates decision-support information for scaffold safety assessments without human intervention.

In the following section, we present detailed results corresponding to the three stages shown in Figure 2, namely, data acquisition, knowledge extraction, and visualization, to demonstrate the effectiveness of the proposed AI-based platform.

4 Results and discussion

This section demonstrates the results achieved during different stages of the developed AI-based platform.

4.1 Data acquisition

Figure 3 illustrates the raw point-cloud data of the scaffolding from a construction site that was acquired using a LiDAR device.

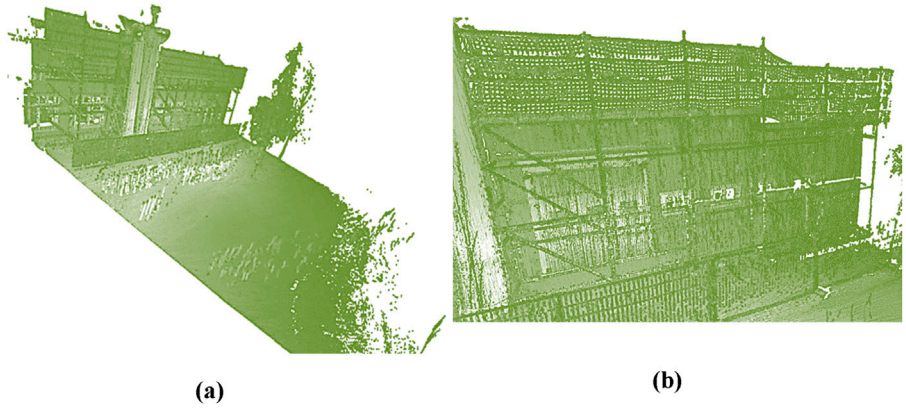


FIGURE 3
Raw point-cloud data of the scaffolding from **(a)** perspective 1 and **(b)** perspective 2.

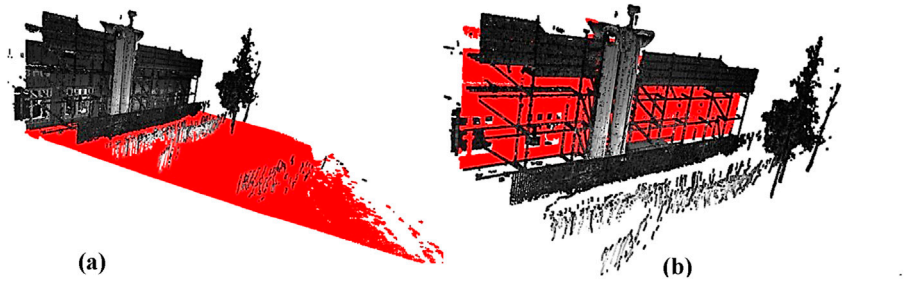


FIGURE 4
Processing of the point-cloud data to detect **(a)** the ground and **(b)** walls.

The figure presents different perspectives of the acquired data. The raw 3D point-cloud scan obtained from the construction site consists of approximately 16 million points and includes not only the scaffolding but also other objects like the nearby trees and surrounding structures, which are not of interest to this work. The raw point-cloud data are next processed to isolate the primary object of interest.

4.2 Data preprocessing

The preprocessing step discussed herein is applied to both the initial and periodic data acquired. Raw data often contain noise, outliers, and redundant points; therefore, preprocessing ensures that the dataset is cleaner and more interpretable. To extract the object of interest, preprocessing involves several operations like removal of the ground plane, elimination of outliers, and voxelization to reduce the density of data points to make the data more manageable for further processing. During the voxelization step, the point-cloud data are partitioned into uniformly sized grids such that the points within each grid can be aggregated to achieve a manageable resolution. For planar surface removal, iterative algorithms like RANSAC (Fischler and Bolles, 1981; Rui et al., 2022) are commonly used to identify and separate the planar surfaces from the point-cloud data. In this study, RANSAC was employed by repeatedly fitting planes to some random subsets of data points and

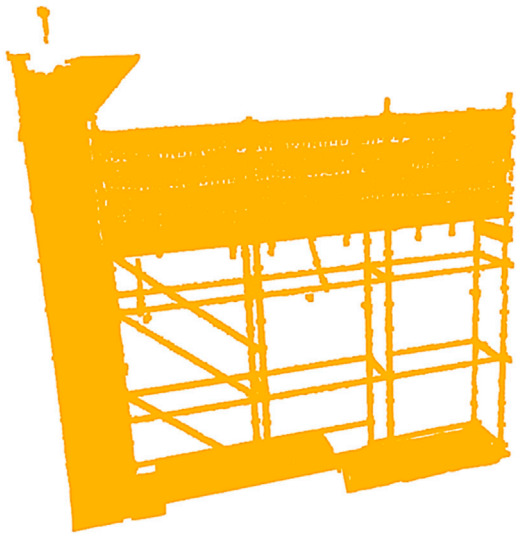


FIGURE 5
Initial point-cloud data of the scaffolding.

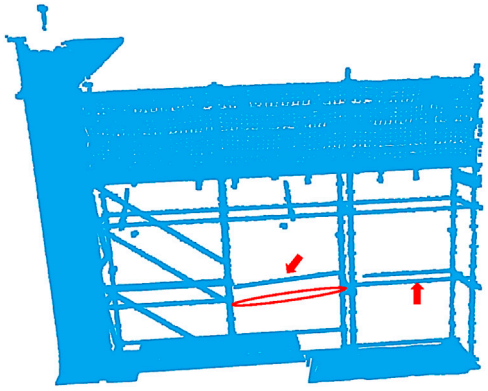


FIGURE 6
Campaign-based scan showing the structural modifications to the scaffolding.

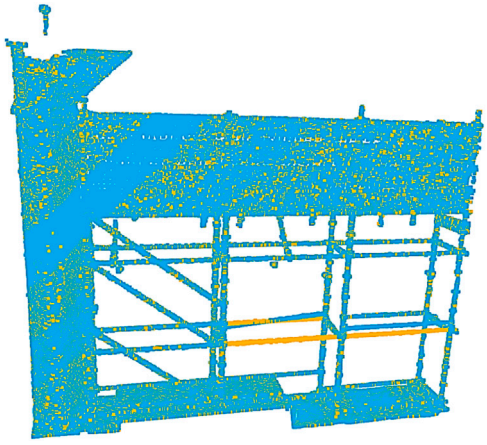


FIGURE 7
Structure comparison based on the two sets of point-cloud data.

selecting the planes with the highest number of inliers; this allowed identification and removal of the planar surfaces, leaving only the scaffolding structure as the primary object of interest. Figure 4 shows the detected ground plane and walls in red color. Since the scaffolding is installed at a predefined distance from the wall, all data points beyond this distance are eliminated to obtain the object of interest, as shown in Figure 5. In this work, the certified scaffolding structure is considered as the initial or reference scan.

4.3 Structure comparison

This step replicates the routine inspection process typically undertaken by the site manager, where a LiDAR scan of the construction site is obtained every 2 weeks or as required. For periodic monitoring, campaign-based LiDAR scans are acquired and compared with the certified reference scan to detect modifications to the scaffolding. Each newly acquired set of point-cloud data undergoes the preprocessing steps mentioned in Section 4.2. The two most common issues faced by the site manager during inspections are missing and deviated scaffolding braces or elements; such changes usually occur as the construction progresses and workers modify the temporary structure for ease of access, as shown in Figure 6. In the figure, the missing element is shown by the red circle, whereas the shifted or deviated elements are indicated by red arrows.

The recently extracted and cleaned scaffolding point-cloud data are compared with the reference data to detect modifications so that the site manager can focus on the modified sections. The iterative closest point algorithm (Li et al., 2021) is used to align and minimize the distance between the two sets of point-cloud data; this algorithm estimates the transformation matrix, which is a combination of a rotation matrix R and a translation matrix T, that minimizes the mean-squared error, as given in Equation 1. Here, I represents the initial data from the certified scan, C is the current scan, and Np denotes the total number of data points.

$$Error\ (R,T) = \frac{1}{N_p}\sum_{i=1}^{N_p}\|I_i - RC_i - T\|^2 \quad (1)$$

For each point in C, the algorithm identifies the closest point in I to establish a point correspondence; these correspondences are then used to estimate the optimal transformation matrix that best aligns the two sets of point-cloud data. To assess convergence, the algorithm evaluates the change in error and iterates until convergence or the maximum iteration limit is reached. Figure 7 shows the output from the algorithm, where the matching or identical points between the scans appear in blue color, while the modified points are highlighted in yellow color. This enables the site manager to focus on the sections requiring attention, which is particularly beneficial for a large complex construction site with scaffolding at multiple locations.

To further evaluate and visualize the alignment accuracy between the current and reference point-cloud data, the distance between the corresponding point data is computed. A threshold is applied to distinguish well-aligned from poorly aligned points, and the result is color-coded depending on the severity level. If the distance between the current and reference data points is within the threshold, the points are colored green, while those exceeding the threshold are displayed in red to signal possible misalignments for closer inspection. By tuning the threshold parameter in the point-cloud distance calculation, the percentage of deviation or modification can be adjusted. The flexibility of tuning the threshold allows site managers to adapt the severity levels to site-specific requirements and professional judgment. Figure 8 shows the points having distances less than the defined thresholds in green, whereas points exceeding the threshold distances are shown in red. Figure 8a shows the result for 10% threshold, where small deviations are acceptable and do not compromise the structural integrity; Figure 8b demonstrates the result for 5% threshold, where even small deviations are unacceptable.

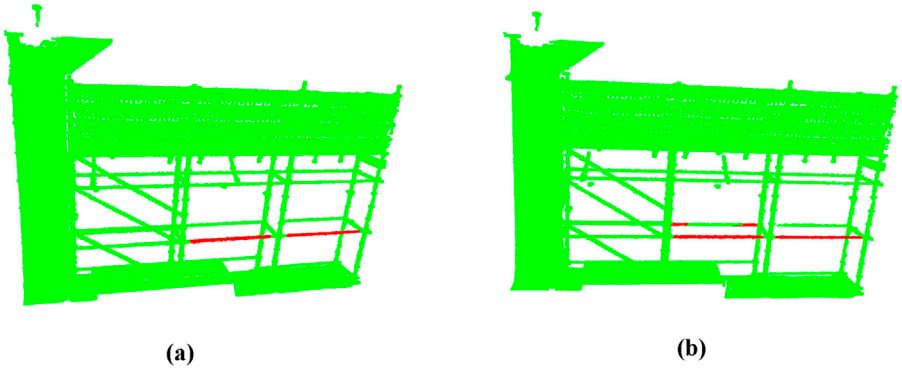


FIGURE 8
Distance comparisons based on **(a)** 10% and **(b)** 5% thresholds.

4.4 Graph representation

Scaffolding structures are composed of interconnected elements like tubes or braces and joints, which can naturally be represented using a graph data structure. The physical elements of the scaffolding like rods or braces represent the edges in the graph, whereas the joints represent the nodes. This approach allows representation of the complex point-cloud data of scaffolding efficiently and systematically.

Figure 9 illustrates the steps involved in the process of representing the graph data structure of a scaffolding from its raw point-cloud data. First, the scaffolding is extracted from the raw point-cloud data by segmentation according to local geometric characteristics; here, the KDTree algorithm (Guo et al., 2017; Lin et al., 2020) was employed to identify neighboring points within a predefined radius. Singular value decomposition was then applied to compute the principal components that were then used to determine the distribution of the selected points and classify their shapes as linear, planar, or spherical. This classification allows the braces to be identified as linear features, platform and safety sheets as planar structures, and joints as spherical features. The result of this step is shown in Figure 9 under the element extraction block, where the braces are represented in green, joints are shown in red, and planar surfaces like the safety sheet are depicted in blue. Second, the density-based spatial clustering of applications with noise (DBSCAN) (Ester et al., 1996; Kim et al., 2024) algorithm was used to cluster the point-cloud data and isolate the linear features corresponding to the scaffolding braces. To reduce the data and computational complexity, the two farthest points were determined for each cluster along with the alignment of the braces. For each of these farthest points, the corresponding close data points were identified and their mean was computed to represent a physical joint of the scaffolding. As shown in the figure, the braces are represented in green, while the joints shown in red are the mean values derived from the bounding boxes represented by the dotted lines. The last row of Figure 9 presents the color coding for each brace in the horizontal x direction, horizontal y direction, and vertical z direction of a scaffolding.

4.5 Hybrid clustering

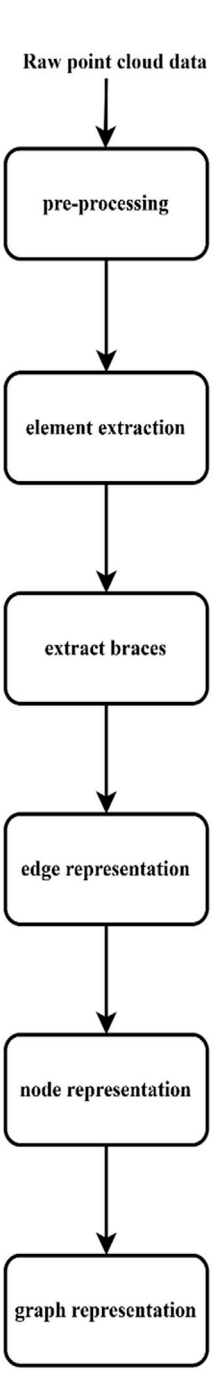
In the element extraction block, the linear elements (i.e., braces) are separated from the spherical elements like joints, after which the DBSCAN algorithm is applied to cluster the braces. However, as shown in Figure 10a, certain clusters contain both vertical and horizontal braces under one label. Such edge cases require further refinement before clustering. Accordingly, an additional normal vector-based separation is introduced to specifically target incorrect or misclassified clusters. The normal vector representation for the edge case is shown in Figure 10b. Thus, two-stage or hybrid clustering is adopted for edge cases, where the first stage is based on spatial distance and second stage employs normal vector direction. As shown in Figure 10c, the results generated from the two clustering stages are combined for more accurate differentiation between the vertical and horizontal braces of the scaffolding structure.

4.6 Visualization

Once the scaffolding is represented as a graph data structure, its layout is as shown in Figure 11. In this representation, the horizontal braces are colored green, whereas the vertical braces are shown in blue and joints are denoted by small red circles. Although Figure 11 may not visually resemble a physical scaffolding structure, it effectively captures the connectivity and relationships between the components. As discussed previously, one of the major issues during inspection is missing scaffolding elements. Figure 12 demonstrates this case by comparing the graph representations of the reference scan with the recent scan in which a brace is missing. Figure 12 shows the graph representation of a scaffolding with a missing brace highlighted in red.

VR and AR are increasingly being recognized as transformative technologies that are changing the construction industry through better visualization capabilities (Li et al., 2018). VR creates an experience by immersing the users in a fully simulated environment to effectively isolate them from the physical world; this allows the users to engage in natural interactions and navigation within the virtual environment in a manner that closely resembles reality. However, AR superimposes the digital information directly

10.3389/fbuil.2026.1723491



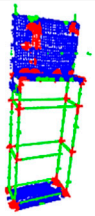
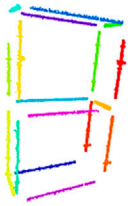
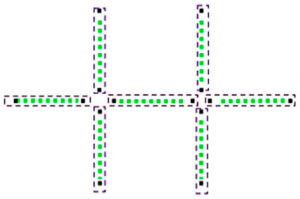
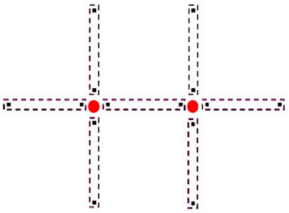
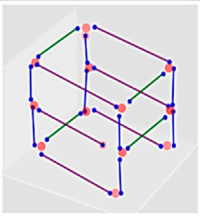
Process	Outcome	Figure
• Voxelization • RANSAC	• Reduce number of points • Object of interest (scaffolding)	
• KDTree • PCA	• Linear element - braces • Planar surface - sheets • Spherical element - joints	
• DBSCAN	• Clustering of braces • Removing small clusters	
• Bounding boxes • Find two farthest points in each bounding box	• Two points for each brace • Alignment of each brace	
• For each farthest point find all the closest points • Mean of all those points	• Joint between braces	
• Mean value, brace alignment with connected nodes • Color coding the x, y, z and diagonal oriented braces	• Reduce complexity • Easier to handle	

FIGURE 9
Steps involved in the graph data structure representation of a scaffolding.

onto the physical environment to enhance the user's interactions with their surroundings. These immersive tools enable site managers to conduct virtual inspections of the scaffolding and specifically focus on critical sections to ensure structural integrity. To enable immersive visualization, the point-cloud files (.e57 format) from the knowledge extraction stage are transformed into optimized 3D representations suitable for real-time rendering. Azure remote rendering (ARR) service was used in this work to facilitate the rendering of high-quality and interactive 3D content for seamless real-time streaming to AR and VR devices. Figure 13 illustrates the conversion pipeline used in the visualization stage; here, the processed point-cloud data are converted to AR/VR-friendly "arrAsset" format using ARR. This conversion involves exporting the .e57 format file to a standard intermediate format like .glb as the .e57 format is a vendor-neutral exchange format; however, arrAsset is a runtime asset format for cloud-based AR/VR applications.

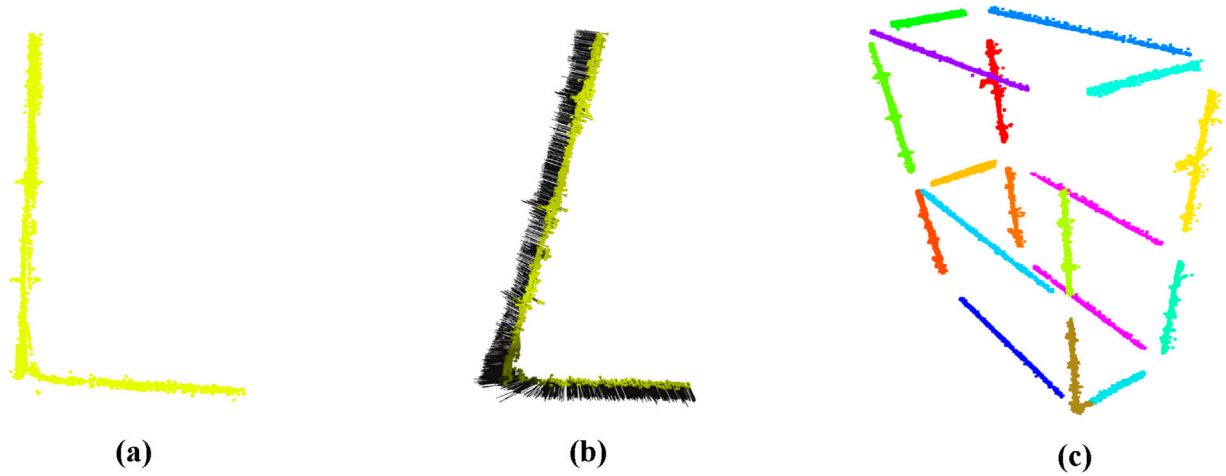


FIGURE 10
Hybrid clustering of the braces showing **(a)** an edge case, **(b)** normal vector representation of the edge case, and **(c)** labeling based on the two-stage clustering approach.

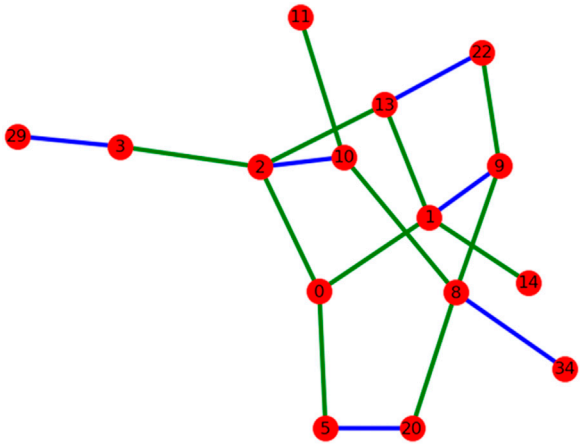


FIGURE 11
Graph data structure representation of the scaffolding, where the horizontal braces are shown in green, vertical braces are depicted in blue, and joints are presented as red circles.

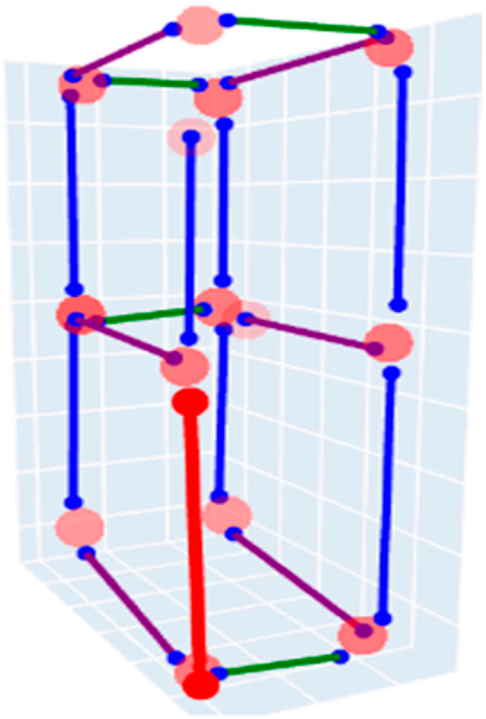


FIGURE 12
Comparison of the scaffolding structure using graph representation.

An illustrative screenshot captured using AR glasses is shown in Figure 14, where a missing brace is shown in red color and a deviated scaffolding element is shown in yellow color. This example demonstrates the conceptual visualization of the detected scaffolding anomalies and is intended to present the inspection results in an immersive environment. Since the 3D model is not rigidly aligned with the physical environment, it allows the users to freely reposition and rotate the AR scaffolding. This enables inspection from multiple viewpoints and improved perception of all spatial dimensions. Although Figure 14 illustrates the AR-based visualization using the HoloLens glasses, the ARR service also supports rendering of 3D scaffolding models in VR glasses. These technologies eliminate geographical barriers to support problem-solving and coordination as well as extend collaborative practices by diving into 3D models and performing hands-on tasks that go beyond computer screens. This, in turn, promotes real-time discussions among experts, resulting in improved decision-making.

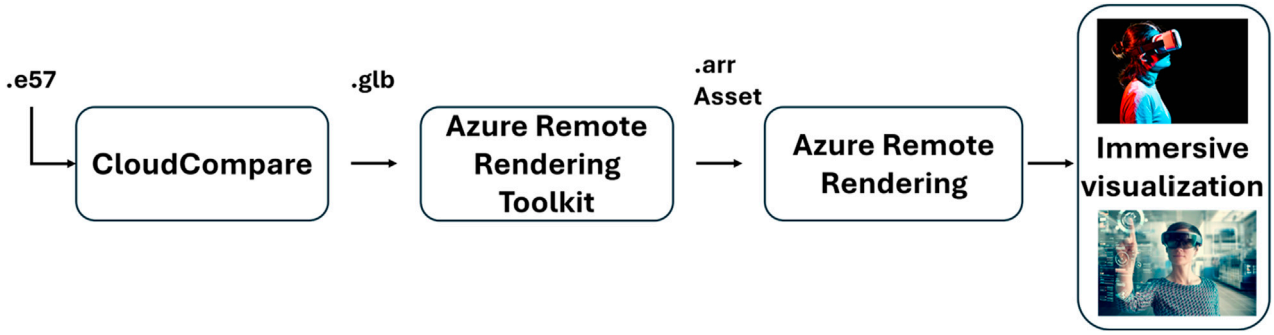


FIGURE 13
Cloud-based pipeline for immersive visualization.

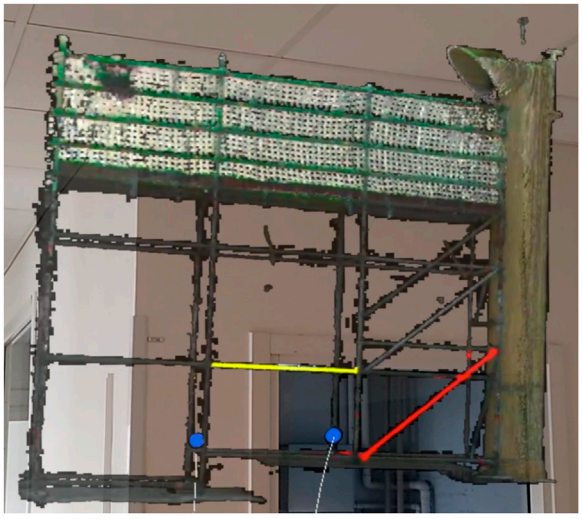


FIGURE 14
Augmented reality-based illustrative visualization that allows the site manager to remotely inspect scaffold deviations (the model is not rigidly aligned with the physical environment).

5 Conclusion and future work

This work proposes a cloud-based AI platform for monitoring scaffolding at construction sites. The proposed method processes the raw 3D point-cloud data acquired using a 360° LiDAR scan to isolate the scaffolding; then, it compares the certified reference scan of the scaffolding structure with recent scans acquired at regular intervals. Deviations or missing elements of the scaffolding are then detected using alignment algorithms. To quantify the difference between the reference and recent scans, the distance between their point-cloud data is computed for precise identification of structural modifications. By automating scaffolding monitoring on complex construction sites, the proposed approach reduces the time and effort required from the site manager. Additionally, this work also suggests an efficient method of representing scaffolding using a graph data structure, which not only captures the connectivity between the joints and braces but also allows integration of numerous well-established graph algorithms for advanced analyses. Highlighting the structural modifications to the scaffolding during construction progress can direct the site manager's attention to focus on these sections, thereby reducing human errors during inspection. By providing data-driven insights into scaffolding structural modifications, the proposed approach assists site managers with critical decision-making. Moreover, the integration of advanced visualization tools like VR and AR can further enhance remote visualization, foster real-time discussions, and facilitate informed decision-making. The proposed cloud-based platform demonstrates the use of AI and digitization technologies to automate scaffolding inspections, which currently rely heavily on manual visual inspections. Traditional visual inspections can be time-consuming, error-prone, and inconsistent, potentially compromising the safety on a construction site. The proposed approach is one method of enhancing the safety of the workers and other individuals at a construction site by reducing errors associated with visual inspections. By linking LiDAR-based data acquisition, AI-driven analysis, and immersive visualization within a connected platform, our work contributes to the broader digital transformation within the construction industry.

From the perspective of practical implementation, the proposed approach can be adapted to real-world construction environments. Although LiDAR-based sensing involves an initial cost, its ability to acquire high-resolution spatial data can be utilized to monitor multiple assets at construction sites to improve long-term cost efficiency. The cloud-based processing pipeline enables execution of computationally intensive algorithms and supports scalability. However, there may be some barriers to adoption of methods like those proposed herein, include the need for trained personnel, integration with existing site workflows, and regulatory acceptance of digital inspection methods.

Within this digital framework, continuous monitoring and predictive capabilities are essential components of proactive safety management. To enhance safety at construction sites, there is a critical need to detect wear, damage, or unintended changes at the earliest. The present work aligns with the PHM concept by enabling continuous monitoring of scaffolding to track changes or modifications against the certified reference structure, thereby supporting the prediction of potential failures or safety hazards. Continuous monitoring of the health of the scaffolding also assists with forecasting when the elements may require maintenance, thereby enhancing the overall safety and reliability. Our work is

expected to be a foundation for future advances in scaffolding condition monitoring. When scaffolding structures are installed, they must adhere to the established design rules to ensure structural stability and integrity. One promising extension of the proposed approach is the integration of design rule verification, which enables the system to automatically assess the current condition of the scaffolding against predetermined safety specifications. Furthermore, implementing an expert system framework could advance the platform toward a higher degree of autonomous decision support.

Beyond the technical contributions, the proposed approach supports broader societal goals by promoting consistent safety practices through data-driven inspection. Increasing accessibility to digital inspection tools can reduce reliance on subjective manual assessments and improve protection for construction workers. In addition, continuous digital monitoring may enable early detection of deviations in the scaffolding even under adverse environmental conditions, thereby contributing toward a more resilient system.

Data availability statement

The datasets presented in this article are not readily available because of confidentiality issues and cannot be shared at the moment. Requests to access the datasets should be directed to corresponding author at sameer.prabhu@ltu.se.

Ethics statement

Ethical approval was not required for the studies involving humans in accordance with the local legislation and institutional requirements. The Institutional Review Board of Luleå University of Technology waived the requirement of written informed consent for participation from the participants or their legal guardians/next of kin because the interactions with site managers were part of the approved project activities and no separate written consent was required from the participants. Involvement in the project inherently included the sharing of domain knowledge and professional insights.

Author contributions

SP: Formal analysis, Writing – review and editing, Visualization, Data curation, Software, Validation, Investigation, Writing – original draft, Methodology, Conceptualization. AP: Formal analysis, Software, Data curation, Conceptualization, Writing – review and editing, Validation, Investigation. RK: Writing – review and editing, Supervision, Conceptualization, Validation.

Funding

The author(s) declared that financial support was received for this work and/or its publication. This work was conducted as a part of project number 14110 funded by the Swedish Construction Industry's Development Fund (SBUF), and the corresponding report (Patwardhan et al., 2025) has been submitted.

Acknowledgements

The authors would like to acknowledge the participating industries for their continuous support and valuable contributions.

Conflict of interest

The author(s) declared that this work was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declared that generative AI was used in the creation of this manuscript. During the preparation of this work, the authors used ChatGPT, a language model developed by OpenAI, to assist with grammar refinement and text clarity. All contents have been reviewed and edited by the authors, who assume full responsibility for the published version.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence, and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Ammar, A., Nassereddine, H., AbdulBaky, N., AbouKansour, A., Tannoury, J., Urban, H., et al. (2022). Digital twins in the construction industry: a perspective of practitioners and building authority. *Front. Built Environ.* 8, 8–2022. doi:10.3389/fbuil.2022.834671

Aydin, C. C. (2014). Designing building façades for the urban rebuilt environment with integration of digital close-range photogrammetry and geographical information systems. *Autom. Constr.* 43, 38–48. doi:10.1016/j.autcon.2014.03.005

Bosché, F. (2010). Automated recognition of 3D CAD model objects in laser scans and calculation of as-built dimensions for dimensional compliance control in construction. *Adv. Eng. Inf.* 24, 107–118. doi:10.1016/j.aei.2009.08.006

Chi, H. L., Chai, J., Wu, C., Zhu, J., Liu, C., and Wang, X. (2017). "Scaffolding progress monitoring of LNG plant maintenance project using BIM and image processing technologies," in *International conference on research and innovation in information systems* ICRIIS: IEEE Computer Society. doi:10.1109/ICRIIS.2017.8002505

Cho, C., Kim, K., Park, J., and Cho, Y. K. (2018a). Data-driven monitoring system for preventing the collapse of scaffolding structures. *J. Constr. Eng. Manage.* 144, 04018077. doi:10.1061/(asce)co.1943-7862.0001535

Cho, C., Sakhakarmi, S., Kim, K., and Park, J. W. (2018b). "Scaffolding modelling for real-time monitoring using a strain sensing approach," in *Isarc 2018 - 35Th international symposium on automation and robotics in construction and international AEC/FM hackathon: the future of building things, (international association for automation and robotics in construction I.A.A.R.C)*. doi:10.22260/isarc2018/0007

Ester, M., Kriegel, H.-P., Sander, J., and Xu, X. (1996). "A density-based algorithm for discovering clusters in large spatial databases with noise," in *Proceedings of the second international conference on knowledge discovery and data mining* (Portland Oregon: AAAI Press), 226–231.

Farsangi, E. N., Shehata, A. O., Rashidi, M., Ghassempour, N., and Mirjalili, S. (2024). Transitioning from BIM to digital twin to metaverse. *Front. Built Environ.* 10, 10–2024. doi:10.3389/fbuil.2024.1486423

Fischler, M. A., and Bolles, R. C. (1981). Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Commun. ACM* 24, 381–395. doi:10.1145/358669.358692

Guo, B., Yu, Z., Zhang, N., Zhu, L., and Gao, C. (2017). 3D point cloud segmentation, classification and recognition algorithm of railway scene. *Yi Qi Yi Biao Xue Bao/Chinese J. Sci. Instrum.* 38, 2103–2111.

Hamdan, N., and Awang, H. (2015). Safety scaffolding in the construction site. *J. Teknol.* 75. doi:10.11113/jt.v75.4956

Heinrich, H. W. (1941). *Industrial accident prevention. A scientific approach*. Industrial Accident Prevention. A Scientific Approach.

Ismail, H. B., and Ghani, K. D. A. (2012). Potential hazards at the construction workplace due to temporary structures. *Procedia Soc. Behav. Sci.* 49, 168–174. doi:10.1016/j.sbspro.2012.07.015

Karim, R., Westerberg, J., Galar, D., and Kumar, U. (2016). Maintenance analytics – the new know in maintenance. *IFAC-PapersOnLine* 49, 214–219. doi:10.1016/j.ifacol.2016.11.037

Khan, N., Saleem, M. R., Lee, D., Park, M. W., and Park, C. (2021). Utilizing safety rule correlation for Mobile scaffolds monitoring leveraging deep convolution neural networks. *Comput. Ind.* 129, 103448. doi:10.1016/j.compind.2021.103448

Khan, M., Khalid, R., Anjum, S., Tran, V.-T., Park, C., et al. (2022). Fall prevention from scaffolding using computer vision and IoT-based monitoring. *J. Constr. Eng. Manage.* 148, 04022051. doi:10.1061/(asce)co.1943-7862.0002278

Kim, A., Lee, K., Lee, S., Song, J., Kwon, S., and Chung, S. (2022a). Synthetic data and computer-vision-based automated quality inspection system for reused scaffolding. *Appl. Sci. Switz.* 12, 10097. doi:10.3390/app121910097

Kim, J., Chung, D., Kim, Y., and Kim, H. (2022b). Deep learning-based 3D reconstruction of scaffolds using a robot dog. *Autom. Constr.* 134, 104092. doi:10.1016/J.AUTCON.2021.104092

Kim, J., Paik, S., Lian, Y., Kim, J., and Kim, H. (2023). "Transformation of point clouds to images for safety analysis of scaffold joints," in *Proceedings of the 40th international symposium on automation and robotics in construction. International association for automation and robotics in construction (IAARC)*. Editors B. de Soto, V. Gonzalez-Moret, and I. Brilakis (Chennai, India), 365–370. doi:10.22260/ISARC2023/0050

Kim, J., Kim, J., Koo, N., and Kim, H. (2024). Automating scaffold safety inspections using semantic analysis of 3D point clouds. *Autom. Constr.* 166, 105603. doi:10.1016/j.autcon.2024.105603

Li, X., Yi, W., Chi, H.-L., Wang, X., and Chan, A. P. C. (2018). A critical review of virtual and augmented reality (VR/AR) applications in construction safety. *Autom. Constr.* 86, 150–162. doi:10.1016/j.autcon.2017.11.003

Li, L., Wang, R., and Zhang, X. (2021). A tutorial review on point cloud registrations: principle, classification, comparison, and technology challenges. *Math. Probl. Eng.* 2021, 1–32. doi:10.1155/2021/9953910

Lin, S., Xu, C., Chen, L., Li, S., and Tu, X. (2020). LiDAR point cloud recognition of overhead catenary system with deep learning. *Sensors* 20, 2212. doi:10.3390/s20082212

Malekloo, A., Ozer, E., AlHamaydeh, M., and Girolami, M. (2022). Machine learning and structural health monitoring overview with emerging technology and high-dimensional data source highlights. *Struct. Health Monit.* 21, 1906–1955. doi:10.1177/14759217211036880

Oskouie, P., Becerik-Gerber, B., and Soibelman, L. (2016). Automated measurement of highway retaining wall displacements using terrestrial laser scanners. *Autom. Constr.* 65, 86–101. doi:10.1016/j.autcon.2015.12.023

Patwardhan, A., Prabhu, S., Karim, R., Samuelson, O., and Simu, K. (2025). AI factory/construction: slutrapport. Available online at: https://vpp.sbuf.se/Public/Documents/ProjectDocuments/923aa3f4-d40b-446b-b7e6-86fb41259660/FinalReport/SBUF%2014110%20Sltrp%20AI%20Factory%20Construction.pdf.

Reese, C. D., and Eidson, J. V. (2006). *Handbook of OSHA construction safety and health*. Boca Raton: CRC Press.

Rui, L., Zhixiang, Z., Xi, C., Wei, M., and Junhao, M. (2022). 3D deformation monitoring method for temporary structures based on multi-thread LiDAR point cloud. *Measurement* 200, 111545. doi:10.1016/j.measurement.2022.111545

Shao, Y., Li, L., Li, J., Yao, X., Li, Q., and Hao, H. (2025). Advancing crack detection with generative AI for structural health monitoring. *Struct. Health Monit.* 0, 14759217251369000. doi:10.1177/14759217251369000

Soualhi, A., Lamraoui, M., Elyousfi, B., and Razik, H. (2022). PHM SURVEY: implementation of prognostic methods for monitoring industrial systems. *Energies* 15, 6909–6915. doi:10.3390/EN15196909

Styliadis, A. (2007). Digital documentation of historical buildings with 3D modeling functionality. *Autom. Constr.* 16, 498–510. doi:10.1016/j.autcon.2006.09.003

Swedish Work Environment Authority (2016). Scaffolding (AFS 2013:4Eng), provisions. Available online at: https://www.av.se/en/work-environment-work-and-inspections/publications/provisions/stallningar-afs-20134-provisions/.

Teizer, J. (2008). 3D range image sensing for active safety in construction. *J. Inf. Technol. Constr.* 13, 103–117. Available online at: https://www/itcon.org/2008/8.

Wang, Q. (2019). Automatic checks from 3D point cloud data for safety regulation compliance for scaffold work platforms. *Autom. Constr.* 104, 38–51. doi:10.1016/J.AUTCON.2019.04.008

Wang, Q., and Kim, M. K. (2019). Applications of 3D point cloud data in the construction industry: a fifteen-year review from 2004 to 2018. *Adv. Eng. Inf.* 39, 306–319. doi:10.1016/J.AEI.2019.02.007

Wang, J., Zhang, S., and Teizer, J. (2015). Geotechnical and safety protective equipment planning using range point cloud data and rule checking in building information modeling. *Autom. Constr.* 49, 250–261. doi:10.1016/j.autcon.2014.09.002

Wang, Q., Tan, Y., and Mei, Z. (2020). Computational methods of acquisition and processing of 3D point cloud data for construction applications. *Archives Comput. Methods Eng.* 27, 479–499. doi:10.1007/s11831-019-09320-4

Wang, R., Shao, Y., Li, Q., Li, L., Li, J., and Hao, H. (2024). A novel transformer-based semantic segmentation framework for structural condition assessment. *Struct. Health Monit.* 23, 1170–1183. doi:10.1177/14759217231182303

Whitaker, S. M., Graves, R. J., James, M., and McCann, P. (2003). Safety with access scaffolds: development of a prototype decision aid based on accident analysis. *J. Saf. Res.* 34, 249–261. doi:10.1016/S0022-4375(03)00025-2

Yigzew, F., Kim, H., Mun, S., and Park, S. (2024). Non-destructive damage detection for steel pipe scaffolds using MFL-based 3D defect visualization. *J. Civ. Struct. Health Monit.* 14, 501–509. 10.1007/s13349-023-00726-0.

Yu-Fei, L., Soojin, C., Spencer, B. F., and Jian-Sheng, F. (2016). Concrete crack assessment using digital image processing and 3D scene reconstruction. *J. Comput. Civ. Eng.* 30, 04014124. doi:10.1061/(ASCE)CP.1943-5487.0000446

Zhao, Y., Seo, H., and Cheng, C. (2021). Displacement mapping of point clouds: application of retaining structures composed of sheet piles. *J. Civ. Struct. Health Monit.* 11, 1–16. doi:10.1007/s13349-021-00491-y

frontiers | Frontiers in Built Environment

TYPE Original Research
PUBLISHED 17 February 2026
DOI 10.3389/fbuil.2026.1753382

Check for updates

OPEN ACCESS

EDITED BY
Vagelis Plevris,
Qatar University, Qatar

REVIEWED BY
Carlos Couto,
University of Aveiro, Portugal
Yanping Zhu,
Montana Technological University,
United States

*CORRESPONDENCE
Soheila Kookalani,
sk2268@cam.ac.uk

RECEIVED 24 November 2025
REVISED 09 January 2026
ACCEPTED 14 January 2026
PUBLISHED 17 February 2026

CITATION
Kookalani S, Liu H, Dash T, Mathew A and Brilakis I (2026) Interpretable machine learning for predicting the bearing capacity of double shear-bolted connections: a data-driven evaluation.
Front. Built Environ. 12:1753382.
doi: 10.3389/fbuil.2026.1753382

COPYRIGHT
© 2026 Kookalani, Liu, Dash, Mathew and Brilakis. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Interpretable machine learning for predicting the bearing capacity of double shear-bolted connections: a data-driven evaluation

Soheila Kookalani[1]*, Hongchen Liu[1], Tirtharaj Dash[2], Alwyn Mathew[1] and Ioannis Brilakis[1]

[1]Department of Engineering, University of Cambridge, Cambridge, United Kingdom, [2]Department of Biochemistry, University of Cambridge, Cambridge, United Kingdom

Introduction: Accurate prediction of the bearing capacity of double shear-bolted connections in structural steel is essential for ensuring safety and efficiency in structural design. This study explores the application of ten machine learning algorithms to enhance prediction accuracy while addressing the interpretability challenges often associated with such models.
Methods: Models were tuned with 10-fold crossvalidation and assessed using RMSE, R^2 and a20 accuracy index. A comprehensive sensitivity analysis evaluates the influence of input parameters, while advanced interpretability techniques, such as partial dependence plots, accumulated local effects, and Shapley additive explanations, are employed alongside parametric studies to elucidate the decision-making processes of the models.
Results: These methods facilitate the identification of critical variables that influence bearing capacity predictions at both local and global scales.
Discussion: The study demonstrates that machine learning can be a trustworthy and data-driven complement to conventional mechanics-based approaches, when coupled with rigorous interpretability, advancing both safety and efficiency in steelconnection design. The findings highlight the potential of interpretable machine learning approaches to not only improve predictive precision but also provide actionable insights into complex model behaviours, ultimately advancing structural engineering practices and promoting data-driven design methodologies.

KEYWORDS

bearing capacity prediction, double shear-bolted connections, interpretable AI, machine learning, sensitivity analysis, structural steel joints

1 Introduction

Bolted and welded connections are widely regarded as the backbone of steel structures, ensuring load transfer and global stability across a broad spectrum of engineering applications. In comparison to welded joints, bolted connections offer several advantages, including expedited assembly and reduced cost by obviating specialized labour and on-site welding procedures, while still achieving reliable structural performance when bolt holes and net-section effects are properly accounted for in design (Zakir et al., 2022). Within the family of bolted connections, bearing-type joints

have garnered particular attention due to the primary load transfer mechanism arising from the bearing action between the bolt shank and the plate hole surface. Over the years, significant research has been dedicated to clarifying the behaviour of bearing-type bolted connections, often focusing on parameters such as bolt diameter, edge and end distances, plate thickness, and material characteristics.

Recent investigations have expanded our understanding of bolted connections under a variety of geometric configurations and load conditions. For instance, Lyu et al. (2019) provided a comprehensive review of how edge and end distances influence the ultimate capacity of bolted connections. Earlier work also identified the presence of an "active shear plane" between the gross and net shear planes, lending crucial insights into design practices governed by structural steel specifications (Clements and Teh, 2013; Teh and Clements, 2012). Building upon these findings, The and Uz (Teh and Uz, 2015a; Teh and Uz, 2015b) introduced a formula to predict the ultimate shear-out capacity, which was subsequently refined to integrate a more accurate bearing coefficient that better captures combined bearing and shear-out failures (Teh and Uz, 2016).

In recent years, the integration of Artificial Intelligence (AI) into civil engineering has led to transformative advances across structural, geotechnical, and materials domains (Liu et al., 2025; Torres et al., 2025; Kookalani et al., 2025). For instance, Zhang et al. (2022) utilized deep neural networks enhanced with Harris Hawks Optimization (HHO) to generalize friction angle predictions of clays for slope stability analysis. Ghanizadeh et al. (2023) developed a hybrid Multivariate Adaptive Regression Spline- Evolutionary-Based Search (MARS-EBS) model for estimating the bearing capacity of geogrid-reinforced stone columns, emphasizing the value of hybrid soft computing techniques. Similarly, Barkho et al. (2023) employed ensemble Convolutional Neural Networks (CNNs) to achieve high-accuracy structural damage classification. Further, Zhao et al. (2022) presented a Whale Optimization Algorithm (WOA)-optimized neural network to accurately predict RC beam deflection, outperforming conventional Artificial Neural Network (ANN) models. In the materials domain, Ashrafian et al. (2023) demonstrated the application of MARS and Extreme Learning Machine (ELM) to predict the strength of sustainable concrete mixes. Similarly, Mahmood et al. (2022) combined various mathematical models to assess the compressive strength of cement-grouted sands, illustrating hybrid Machine Learning (ML) model effectiveness in complex materials characterization. These studies collectively illustrate the growing sophistication and diversity of ML applications in civil and structural engineering.

Despite these advances, predicting the bearing capacity of multi-bolt configurations, particularly in double shear arrangements, continues to pose significant challenges. The majority of existing design equations and predictive models have their roots in single-bolted connection tests, which do not fully encapsulate the intricate load distribution that arises in multi-bolt layouts (Lyu et al., 2020a). Lyu et al. (2020a) demonstrated that the bearing strength of individual fasteners in multi-bolt connections exhibits a non-linear relationship with the edge distance-to-hole diameter ratio, underscoring the pitfalls of extrapolating single-bolt findings to multi-bolt scenarios. This discrepancy highlights the urgent need for more robust, data-driven approaches that can account for the complexity of multi-bolt load sharing and failure mechanisms. In particular, this research is motivated by the lack of accurate and interpretable predictive tools for multi-bolt double shear, bearing-type bolted connections, where current design equations, largely extrapolated from single-bolt tests and simplified interaction rules, may fail to capture the combined influence of end and edge distances, pitch, and bolt-row configuration on bearing capacity.

While traditional methods such as Eurocode and American Institute of Steel Construction (AISC) analytical formulations, or numerical models like Finite Element Analysis (FEA), have been validated over decades and remain essential in design practice, they rely on simplifying assumptions such as linear elasticity, idealized boundary conditions, and uniform material behaviour. These constraints can hinder their effectiveness in capturing the complex, nonlinear behaviours present in double shear-bolted connections, especially in multi-bolt scenarios where interaction effects dominate. In contrast, ML models offer a flexible, data-driven complement to traditional mechanics-based and numerical approaches, as they can learn intricate patterns from large experimental datasets without being restricted to a predefined functional form or the simplifying assumptions embedded in existing design equations, provided that the training data are sufficiently rich and representative (Cabrera et al., 2023; Sadrossadat et al., 2022; Sarir et al., 2021a). This study demonstrates that gradient-boosting models improve predictive accuracy compared with traditional regression methods and other ML techniques, achieving error levels that are compatible with use as complementary, decision-support tools in connection design.

Furthermore, unlike conventional ML applications often criticized for their opacity, this study incorporates explainability tools, which transparently reveal how key features such as normalized end and edge distances and the number of bolt rows influence predictions. This not only enhances trust in ML outputs but also bridges the gap between black-box models and engineering insight. When paired with Building Information Modelling (BIM) in future implementations, such interpretable ML frameworks could enable real-time design evaluations and optimization workflows that are often impractical using FEA or static code-based checks alone (Mahdavipour et al., 2026). Therefore, ML should not be viewed as a replacement for traditional methods, but rather as a robust and scalable complement, particularly suited for design scenarios that demand generalization across multiple configurations, rapid feedback, and insight into high-dimensional parameter interactions.

Furthermore, the implementation of BIM in the context of bolted connections introduces new opportunities for enhanced structural analysis and decision-making. BIM enables the integration of ML-based predictive models within digital twins, allowing for real-time assessment of connection performance and facilitating data-driven optimization of bolted connections in steel structures. BIM-driven workflows can improve the efficiency and reliability of bolted connection design by incorporating as-built data, geometric constraints, and material properties, reducing material waste and construction errors (Alavi et al., 2024a; Alavi et al., 2024b; Alavi et al., 2024c; Alavi et al., 2024d; Alavi et al., 2024e; Alavi et al., 2024f). The synergy between BIM and ML-driven predictions fosters a more comprehensive understanding of structural behaviour, ultimately contributing to more sustainable and efficient construction practices.

In response to these challenges, the present study harnesses the capabilities of ML to predict the bearing capacity of double shear-bolted connections in structural steel. The primary objective of this study is to develop and interpret ML-based models that can accurately predict the bearing capacity of double shear-bolted connections, thereby contributing to safer and more data-informed structural engineering practices. Leveraging a sizable dataset of 443 experimental results, 10 ML algorithms, including linear and ridge regression, support vector machines, tree-based methods, and boosting ensembles, are systematically trained, and tuned using grid search and *K*-fold Cross-Validation (CV) to ensure reliable performance. The choice of these algorithms reflects the growing recognition that ML techniques can capture complex, nonlinear relationships in engineering problems more effectively than many traditional predictive models.

However, from an engineering standpoint, predictive performance alone is insufficient (Li et al., 2023; Lim and Chi, 2019; Ma et al., 2024); explainability and interpretability are equally critical for effective design and decision-making (Geyer et al., 2024; Love et al., 2023). To this end, the study employs advanced interpretability frameworks such as Shapley Additive Explanations (SHAP), Accumulated Local Effects (ALE), and Partial Dependence Plots (PDPs). Sensitivity analyses, including first-order and total-effect measures, further elucidate the relative significance of input parameters, while a parametric study sheds light on how variations in key variables influence the bearing capacity. This work not only delivers highly accurate predictions but also offers insights into how different factors drive the ML model's behaviour by integrating these techniques.

The remainder of this paper is structured as follows: Section 2 introduces the 10 ML algorithms, including Linear Regression (LR), Ridge Regression (RR), Support Vector Regression (SVR), K-Nearest Neighbors (KNN), Decision Trees (DT), Random Forest (RF), Adaptive Boosting (AdaBoost), Extreme Gradient Boosting (XGBoost), Category Boosting (CatBoost), and Light Gradient Boosting Machine (LightGBM), and details the performance metrics, hyperparameter tuning strategies, and Sensitivity Analysis (SA) approaches. Section 3 focuses on interpretable ML methodologies, illustrating their respective contributions to model transparency. Section 4 presents a numerical example that demonstrates the practical application of these ML models in predicting bearing capacity. Section 5 then discusses the interpretability results, highlighting the implications for both design and research. Finally, Section 6 summarizes the key findings, outlines current limitations, and proposes potential directions for future work.

2 Digital twin integration

The rapid digital transformation of the construction sector has established digital twins as a central paradigm for data-driven design, monitoring, and lifecycle management of structural systems. Digital twins require analytical engines capable of providing continuous, reliable, and explainable predictions using both as-designed and as-built data. The interpretable ML framework developed in this study aligns with these requirements by offering a transparent, computationally efficient, and easily deployable predictive model for assessing the bearing capacity of double shear-bolted connections.

In a digital twin environment, geometric and material parameters, such as end and edge distances, pitch dimensions, bolt configurations, and strength ratios, are inherently embedded within BIM data structures. These parameters can be automatically extracted from Industry Foundation Classes (IFC)-compliant models or connected databases and fed directly into the ML-based prediction engine. The system can deliver real-time assessments of connection performance as the model geometry or material conditions evolve by integrating the ML model and its interpretability layers into a digital twin data pipeline. This capability enables faster and more adaptive structural evaluations compared to conventional FEA, which can be computationally intensive and time-prohibitive for iterative or real-time use.

The interpretability tools incorporated in this study reinforce the suitability of the model for digital twin adoption. SHAP values provide detailed, instance-level explanations of how geometric and material features influence predicted bearing capacity, ensuring transparency and auditing capability, essential requirements for engineering decision-making in digital twin ecosystems. Moreover, sensitivity analysis supports prioritisation of key parameters, guiding both model refinement and potential sensor placement strategies in operational digital twins.

Beyond design-phase applications, the proposed ML framework can be embedded into digital twins for structural health monitoring and predictive maintenance. When paired with inspection data, periodic updates, or real-time sensor readings, the digital twin can autonomously evaluate deviations from expected behaviour, identify performance degradation, and support proactive decision-making. Additionally, in emerging metaverse or immersive simulation environments, this model can serve as a computational engine for interactive scenario testing, virtual training, and construction rehearsal.

Overall, the integration of the interpretable ML model into digital twin workflows enhances the analytical capabilities of digital construction ecosystems by enabling fast, transparent, and data-driven assessments of bolted steel connections. This contribution supports ongoing efforts to establish intelligent, adaptive, and reliable digital twins for sustainable and efficient construction practice. It should be noted that, in this study, the ML model is developed and validated against experimental data only; the BIM and digital twin workflows discussed in this section are presented as a conceptual integration framework and as directions for future implementation, rather than as a fully realised software integration demonstrated in this paper.

3 Machine learning model development

This study examines the performance of 10 ML algorithms, LR, RR, SVR, KNN, DT, RF, AdaBoost, XGBoost, CatBoost, and LightGBM, in order to identify the most suitable approach for accurately predicting the bearing capacity of double shear-bolted connections. Additionally, it introduces relevant performance metrics, an efficient hyperparameter tuning methodology, and a SA framework, all of which are designed to facilitate a

comprehensive evaluation and optimization of the selected models. Optimization plays a crucial role in enhancing the predictive accuracy and generalisability of these models, ensuring practical applicability in engineering design and decision-making (Kookalani et al., 2022a).

3.1 Machine learning models

3.1.1 Linear regression (LR)

LR is a supervised ML technique used to model the relationship between dependent and independent variables through a linear equation (Liang and Song, 2009). It is among the simplest and most commonly employed methods for analysing how output variables are influenced by input features. LR operates by minimizing the least squares error between the observed and predicted values, enabling the determination of an optimal set of coefficients for the parameters. The result of LR is represented as a linear combination of input variables, defined by Equation 1:

$$f(X) = \beta_0 + \sum_{i=1}^{M} X_i \beta_i \tag{1}$$

where β_0 represents the intercept and β_i denotes the regression coefficient, with M being the total number of parameters. The regression coefficients are determined by minimizing the sum of squared residuals, as expressed in Equation 2:

$$\beta_{LR} = \underset{\beta}{\operatorname{argmin}} \left[\sum_{j=1}^{N} \left(y_j = f\left(x_j\right) \right)^2 \right] \tag{2}$$

where N represents the total number of instances and y_j signifies the corresponding target output.

3.1.2 Ridge regression (RR)

RR is an advanced form of LR designed to mitigate the high variance often associated with LR [10]. In many ML methods, there is a balance to be struck between bias and variance. RR addresses this by incorporating a regularization term, which reduces the magnitude of regression coefficients. This helps to prevent overfitting and enhances the ability of the model to generalize (Hoerl and Kennard, 1970). RR achieves this by minimizing a penalized cost function, as expressed in Equation 3:

$$\beta_{Ridge} = \underset{\beta}{\operatorname{argmin}} \sum_{j=1}^{N} \left(y_j - \beta_0 - \sum_{i=1}^{M} X_{ji} \beta_i \right)^2 + \lambda \sum_{i=1}^{M} \beta_i^2 \tag{3}$$

where λ controls the degree of coefficient shrinkage, with the RR coefficients tending toward zero as λ increases.

3.1.3 Support vector regression (SVR)

SVR applies the foundational concepts of Support Vector Machines (SVMs) to address regression problems. In SVM, hyperplanes are designed to divide the data space with maximum margin, thereby enhancing the robustness of the predictive model. The SVM predictions are formulated as follows expressed in Equation 4:

$$f(x) = \phi(x)^T \omega + b \tag{4}$$

where $\phi(x)$ represents a mapping function that transforms the input into a higher-dimensional space, while b denotes the bias of the model. The model achieves flatness in $f(x)$ by minimizing a convex optimization problem, ensuring the smallest possible values for ω as expressed in Equation 5 (Smola and Schölkopf, 2004):

$$Minimise \frac{1}{2}\|\omega\|^2 + C \sum_{j=1}^{N} \left(\xi_j + \xi_j^* \right) \tag{5}$$

where C represents the box constraint, and ξj and $\xi^*{}_j$ are the slack variables. Consequently, the SVM prediction is expressed in Equation 6:

$$f\left(x, a_i, a_i^*\right) = \sum_{j=1}^{N} \left(a_i - a_i^*\right) K\left(x_i, x\right) + b \tag{6}$$

where α_i and α_i^* are Lagrange multipliers, and $K(x_i, x)$ is the kernel function. These equations indicate that the kernel function, along with the parameters ε and C, can be adjusted to optimize SVM predictions.

3.1.4 K-nearest neighbours (KNN)

KNN predicts the output variable by taking the mean of several nearby values, where k represents the number of neighbours considered (Cover and Hart, 1967). The fundamental idea behind KNN is that it assigns greater importance to the k closest samples in the training dataset that are near the new data point x. The conditional probability of x is determined using the Equation 7:

$$P(Y = m \mid X = x) = \frac{1}{K} \sum_{i \in N_k} I\left(y_i = m\right) \tag{7}$$

where $I(y_i = m)$ is an indicator function that returns 1 if the given observation belongs to the m-th class, and 0 otherwise; N_k denotes the number of instances.

3.1.5 Decision tree (DT)

DT is a supervised ML technique that can be utilised for regression models. This method constructs a predictive model in a tree-like structure based on training data (Dietterich, 2000). The tree comprises input variables from the dataset represented as internal nodes, decision rules forming the branches, and outputs appearing as leaf nodes. Regression trees fall under a category of ML methods that create estimation models by segmenting the feature space into multiple high-dimensional regions. By partitioning the feature space into D distinct regions, denoted as, $R_1,\cdots,R_D$, the regression problem can be formulated as expressed in Equation 8:

$$f(x) = \sum_{d=1}^{D} c_d I\left(x \in R_d\right) \tag{8}$$

where c_d represents the average of the observations.

3.1.6 Random forest (RF)

RF is a ML technique that leverages an ensemble of decision trees. RF integrates the bagging approach with a strategy that selects random subsets of input parameters. In the bagging approach, each decision tree is constructed using a bootstrap sample, which is generated by randomly sampling the training dataset with replacement. This method mitigates overfitting by reducing the

dependency on individual trees. Moreover, instead of considering all input parameters, RF uses a randomly chosen subset of them for each tree. As an advanced form of the bagging method, RF combines these features to improve performance. The final output of the model is determined by averaging the predictions of all individual decision trees, as shown in Equation 9:

$$Y = \frac{1}{B}\sum_{j=1}^{b} Y_b(X') \tag{9}$$

where B represents the total number of decision trees, Y_b denotes the output of each decision tree, and X' refers to the unknown instances.

3.1.7 Adaptive boosting (AdaBoost)

The AdaBoost algorithm enhances the predictive accuracy of a model by combining multiple weak learners to create a strong learner (Freund and Schapire, 1997; Zhang and Ma, 2012). The structure of the learner is expressed in Equation 10:

$$F_T(x) = \sum_{t=1}^{T} f_t(x) \tag{10}$$

In this approach, samples that were misclassified in a previous step are assigned higher weights in subsequent iterations (Schapire and Singer, 1999; Schapire, 2013). The performance of the model is progressively improved by reducing errors at each step. Initially, all samples are assigned equal weights, and the mean square error of the prediction is calculated. Higher weights are then given to samples with larger errors, and the process is repeated until the results converge, as shown in Equation 11:

$$E_t = \sum_i E[F_{t-1}(x_i) + a_t h(x_i)] \tag{11}$$

where $E(.)$ represents the error function; $F_{t\text{-}1}(x)$ denotes the learner generated in the preceding step, and $f_t(x) = a_t h(x)$ represents the weak learner contributing to the strong learner. Ultimately, AdaBoost combines these weak learners to construct a robust, unified learner.

3.1.8 Extreme gradient boosting (XGBoost)

XGBoost is a boosting algorithm that improves predictive accuracy by iteratively assigning higher weights to weak learners. The approach combines these weak learners to create a more robust and accurate model. Originally introduced by Chen et al. (Chen et al., 2018), gradient boosting employs the gradient vector of the misfit function to build a regression model, akin to gradient-descent methods. XGBoost represents a parallelized tree-based implementation of gradient boosting, where parameters are iteratively adjusted to minimize the residuals from the previous step (Chen and Guestrin, 2016). XGBoost incorporates regularization into both its objective function and loss function to prevent overfitting. The objective function for XGBoost is expressed in Equation 12:

$$Obj = \sum_{i=1}^{n} L\left(\tilde{y}_i, y_i\right) + \sum_{t=1}^{k} \omega\left(f_t\right) \tag{12}$$

where L represents the loss function that measures model bias, while ω denotes the regularization term designed to reduce model complexity.

3.1.9 Category boosting (CatBoost)

CatBoost (Dorogush Veronika et al., 2018) introduces a novel gradient boosting technique that specifically addresses categorical input parameters. This approach employs symmetric decision trees to improve the efficiency of the inference process when using pre-trained weak learning models. CatBoost excels in delivering high performance, especially when handling highly noisy data with diverse characteristics and intricate relationships. To accomplish this, CatBoost randomly orders all instances and assigns values to categorical features. Priority weight coefficients and a priority factor are utilised to reduce the influence of low-frequency category instances and noise on data distribution. This process can be represented as expressed in Equation 13:

$$\hat{x}_k^i = \frac{\sum_{j=1}^{n} I_{\left\{x_j^i = x_k^i\right\}} \cdot y_i + \beta p}{\sum_{j=1}^{n} I_{\left\{x_j = x_k^i\right\}} + \beta} \tag{13}$$

where p represents a prior value, and β denotes its corresponding weight.

3.1.10 Light gradient boosting machine (LightGBM)

LightGBM (Ke et al., 2017) is a decision tree-based algorithm that employs a leaf-wise growth strategy rather than the traditional depth-wise approach, leading to improved accuracy in more complex tree structures. It introduces two novel techniques, Exclusive Feature Bundling and Gradient-based One-Side Sampling, which enhance its efficiency and distinguish it from other gradient-boosted decision tree methods. The fundamental principle of LightGBM is to integrate M weak regression trees into a single, more powerful model, mathematically expressed in Equation 14:

$$F(x) = \sum_{m=1}^{M} f_m(x) \tag{14}$$

where $f_m(x)$ and $F(x)$ denote the output of the mth weak regression tree and the final model output, respectively. The leaf-wise strategy with depth constraints and the histogram-based technique are key advancements that improve the efficiency and accuracy of the LightGBM algorithm.

3.2 Performance indexes

This study utilizes three performance metrics: the root mean square error (*RMSE*), the coefficient of determination (R^2), and the a_{20} *index*. The equations for *RMSE*, R^2, and a_{20} are expressed in Equations 15–17, respectively:

$$RMSE = \sqrt{\frac{1}{N_t}\sum_{i=1}^{N_t}\left(Y_{pi} - Y_i\right)^2} \tag{15}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{N_t}\left(Y_{pi} - Y_i\right)^2}{\sum_{i=1}^{N_t}\left(Y_{pi} - \bar{Y}_{pi}\right)^2} \tag{16}$$

$$a_{20} - index = \frac{m20}{M} \tag{17}$$

where Y_i and Y_{pi} represent the actual and predicted values for the ith observation, respectively; N_t denotes the number of testing models,

and $\bar{Y}_{pi}$ is the mean of the predicted values. The optimal ML model is characterized by the lowest *RMSE* and the highest R^2 and a_{20} index values. The parameter M denotes the total number of samples in the dataset, while $m20$ refers to the number of samples for which the ratio of the observed value to the predicted value lies within the range of 0.80–1.20. This range indicates a ±20% tolerance, commonly accepted in engineering practices as a benchmark for acceptable prediction accuracy. The a_{20}, defined as the ratio $m20/M$, therefore provides an intuitive and practical measure of a model's predictive performance (Asteris et al., 2024a; Asteris et al., 2024b). In the case of a perfectly accurate model, the a_{20} would equal 1, indicating that all predictions fall within the acceptable range. The value of the proposed a_{20} lies in its interpretability and relevance to engineering applications, as it directly reflects the proportion of predictions that closely match the corresponding experimental results within a tolerable margin of error.

3.3 Hyperparameters tuning

Once the dataset has been prepared and the ML technique selected, the next crucial step is defining the parameters of the model, which significantly influence its performance. In this study, hyperparameter optimization is carried out using a combination of grid search and K-fold CV to minimise the risk of overfitting. Initially, potential parameter ranges are defined as grids. The model is then trained iteratively, testing all possible parameter combinations, with performance evaluated using K-fold CV. This approach ensures a reliable assessment of predictive accuracy while reducing bias from the random division of training and testing data. The dataset is split into K equally sized subsets, and the model undergoes K iterations, where K–1 subsets are used for training and the remaining subset for validation. The final model performance is determined by averaging the results across all K iterations. In this study, a 10-fold CV approach is implemented, as it is a widely recognised standard that partitions the data into ten groups, effectively reducing the risk of overfitting.

3.4 Sensitivity analysis (SA)

Saltelli et al. (2008) proposed a variance-based SA method to evaluate how changes in model input values affect the corresponding output. This technique examines interactions between input variables and the output factor by keeping all input parameters constant except for one, which is systematically varied (Liu et al., 2020).

3.4.1 First-order sensitivity indices

The measurement vector in a multivariate k-input model can be expressed as $y = f(x_1, x_2, \ldots, x_k)$. The first-order index is determined using the Equation 18:

$$S_i = \frac{V_{x_i}\left[E_{x_{\sim i}}\left(y \mid x_i\right)\right]}{V\left(y\right)} \tag{18}$$

where $V_{x_i}[E_{x_{\sim i}}(y \mid x_i)]$ quantifies the influence of the variable x_i on the output. When x_i is held constant, $E_{x_{\sim i}}(y \mid x_i)$ represents the variance of the mean value $E(y)$, while $V(y)$ denotes the unconditional variance of y.

3.4.2 Total-effect sensitivity indices

To capture the full variance of the output, higher-order indices for coupling terms must be considered, as the first-order index only accounts for the portion of output variation attributable to the variance of the input variable x_i. Consequently, the total-effect index S_T is employed to quantify the overall influence of x_i on the output variance. The total-effect index is defined as expressed in Equation 19:

$$S_T = 1 - \frac{V_{x_{\sim i}}\left[E_{xi}\left(y \mid x_{\sim i}\right)\right]}{V\left(y\right)} \tag{19}$$

where $E_{x_i}(y \mid x_{\sim i})$ and $V_{x_{\sim i}}[E_{x_i}(y \mid x_{\sim i})]$ represent the mean value of y when all variables except x_i are fixed and its variance, respectively. It is important to note that the difference between S_i and S_T reflects the interactions of x_i with other input variables.

4 Interpretable ML approaches

ML has transformed several disciplines, including structural engineering, by providing predictive models capable of addressing complex relationships and analysing large datasets (Kookalani et al., 2021; Kookalani et al., 2022b; Kookalani et al., 2022c; Kookalani and Cheng, 2021; Xiang et al., 2020a; Xiang et al., 2020b; Xiang et al., 2021; Kookalani et al., 2024). While these ML models can deliver accurate predictions, they often function as "black boxes," lacking transparency in explaining the underlying mechanical or physical principles. This limitation can undermine the credibility of the ML models. To address this challenge, the current study employs three interpretability techniques, including PDP, ALE, and SHAP. Feature significance is assessed by evaluating the increase in prediction error resulting from alterations to factor values. A feature is deemed significant if the error rises substantially after modification, whereas minimal error change indicates low significance.

4.1 Partial dependence plot (PDP)

Friedman (2001) introduced the PDP to examine the marginal effect of a specific parameter on the output by depicting the average outcome values across different parameter settings. The PDP is useful for identifying the relationship between a feature and the target variable. The partial dependence, represented as f_S, for a subset of features x_S, is defined as expressed in Equation 20:

$$f_S(x_S) = E_{X_C}[f(x_S, x_C)] = \int f(x_S, x_C) dP(x_C) \tag{20}$$

where x_S represents the factors considered for the PDP, while x_C denotes the complementary factors. A PDP can be generated for a dataset $\{X_i,\ i = 1, \ldots, n\}$, as expressed in Equation 21:

$$\bar{f}_S(x_S) = \frac{1}{n}\sum_{i=1}^{n} f(x_S, x_{iC}) \tag{21}$$

TABLE 1 Comparative summary of interpretability techniques.

Technique	Advantages	Limitations
PDP	Simple and intuitive; shows global trends	Assumes feature independence; can mislead with correlated variables
ALE	Handles correlated features; less biased than PDP	Sensitive to interval size; offers less intuitive visuals than PDP
SHAP	Provides local and global insights; accounts for feature interactions; consistent	Computationally intensive; complex to implement for large datasets

A critical assumption of PDP is the independence of input parameters. When features are highly correlated, the analysis may involve artificial data samples that are unrealistic, resulting in significant bias in the estimated effect of the feature.

4.2 Accumulated local effects (ALE)

ALE is an unbiased alternative to PDP that represents the average effect of a feature within a ML algorithm (Apley and Zhu, 2020). Unlike PDP, which averages f_S (x_S) over all features and can produce biased results when features are highly correlated, ALE mitigates this issue by focusing on changes in predictions and constraining data to specific grids. ALE averages the changes in predictions by isolating the influence of correlated features, simplifying complex models by considering only one or two factors at a time. ALE can be expressed as Equation 22:

$$f_{xs}, ALE(x_S) = \int_{z_{0,1}}^{x_S} E_{X_C/X_S}\left[\hat{f}^S(x_S, x_C) \mid x_S = z_S\right] dz_S - c$$
$$= \int_{z_{0,1}}^{x_S} \int_{x_C} \hat{f}^S(z_S, x_C) P(x_C \mid z_S) dx_C dz_S - c \quad (22)$$

where c is a fixed constant, $\hat{f}^S(x_S, x_C) = \frac{\delta \hat{f}(x_S, x_C)}{\delta x_S}$ denotes the local effects of x_S on $\hat{f}$ (.) at (x_1, x_S); $z_{0,1}$ represents a value smaller than the minimum observation, and P ($x_C|x_S$) indicates the density. ALE plots are centred around zero, making them particularly effective for correlated data and enhancing visualization. However, the choice of grid or interval can influence the plots and obscure data variability.

4.3 Shapley additive explanations (SHAP)

Lundberg and Lee (2017) introduced the SHAP method, a technique grounded in conditional expectations and game theory, for assessing model predictions. SHAP is used to analyse the influence of individual input features on each output. Broadly, SHAP helps rank features based on their contribution to interaction effects. SHAP employs an additive feature attribution approach to create an interpretable model. The resulting model, represented as a linear function, is the sum of the actual contributions associated with each parameter. The interpretable framework is formulated as expressed in Equation 23:

$$f(x) = g(x') = \phi_0 + \sum_{i=1}^{M} \phi_i x_i', \quad (23)$$

where $x = (x_1, x_2, \ldots, x_p)$ represents the M input variables; and x'_1 denotes the simplified inputs. Using a mapping function $x = hx$ (x'), the parameter x' is transformed into x. Additionally, ϕ_0 and ϕ_i represent a fixed value and the contribution of each parameter, respectively.

4.4 Comparative evaluation of interpretability techniques

PDP, ALE, and SHAP each offer distinct strengths in interpreting complex ML models, as presented in Table 1. PDP is easy to understand and effective for identifying overall trends but may misrepresent effects when features are correlated. ALE addresses this by providing unbiased local effects even with correlated inputs, though its accuracy depends on appropriate interval selection. SHAP offers the most comprehensive insights, attributing both global and local feature importance while capturing interaction effects, albeit at a higher computational cost. Their combined use provides a robust interpretability framework that balances simplicity, accuracy, and depth of insight.

5 Numerical examples

Shear connections in steel structures are generally designed to transfer only shear forces, with no bending moment. However, while this is generally true, there are instances where a bending moment is also present in the shear connection, which can become a critical factor in determining its dimensions. Figure 1 shows two samples of this type of connection.

This study involved the careful assembly of a comprehensive dataset comprising 443 samples extracted from existing literature (Lyu et al., 2020a; Može and Beg, 2014; Može and Beg, 2010; Wang et al., 2017; Ahmed and Teh, 2019; Rex and Easterling, 2003; Kim and Yura, 1999; Yang et al., 2013; Hai et al., 2019; Moe and Beg, 2011; Kim and Lee, 2020; Lyu et al., 2020b; Freitas et al., 2005; Guo et al., 2020; Puthli and Fleischer, 2001; Može, 2018). The input data used in the analysis comprise six distinct features that, according to established design standards and experimental studies, govern bearing and shear-out behavior in bolted connections. Each feature provides critical insight into the mechanical response and characteristics of the system. These features include end distance (e_1), edge distance (e_2), pitch parallel to loading (p_1), pitch perpendicular to loading (p_2), the number of bolt rows (N_r), and stress ratio (fu/fy). Consistent with conventional design formulations, geometric quantities are expressed as normalized ratios, e_1/d_0, e_2/d_0, p_1/d_0 and p_2/d_0, where d_0 is the bolt-hole diameter, while fu/fy is already dimensionless. This choice ensures that the model learns from physically meaningful, scale-independent variables that govern bearing capacity across different

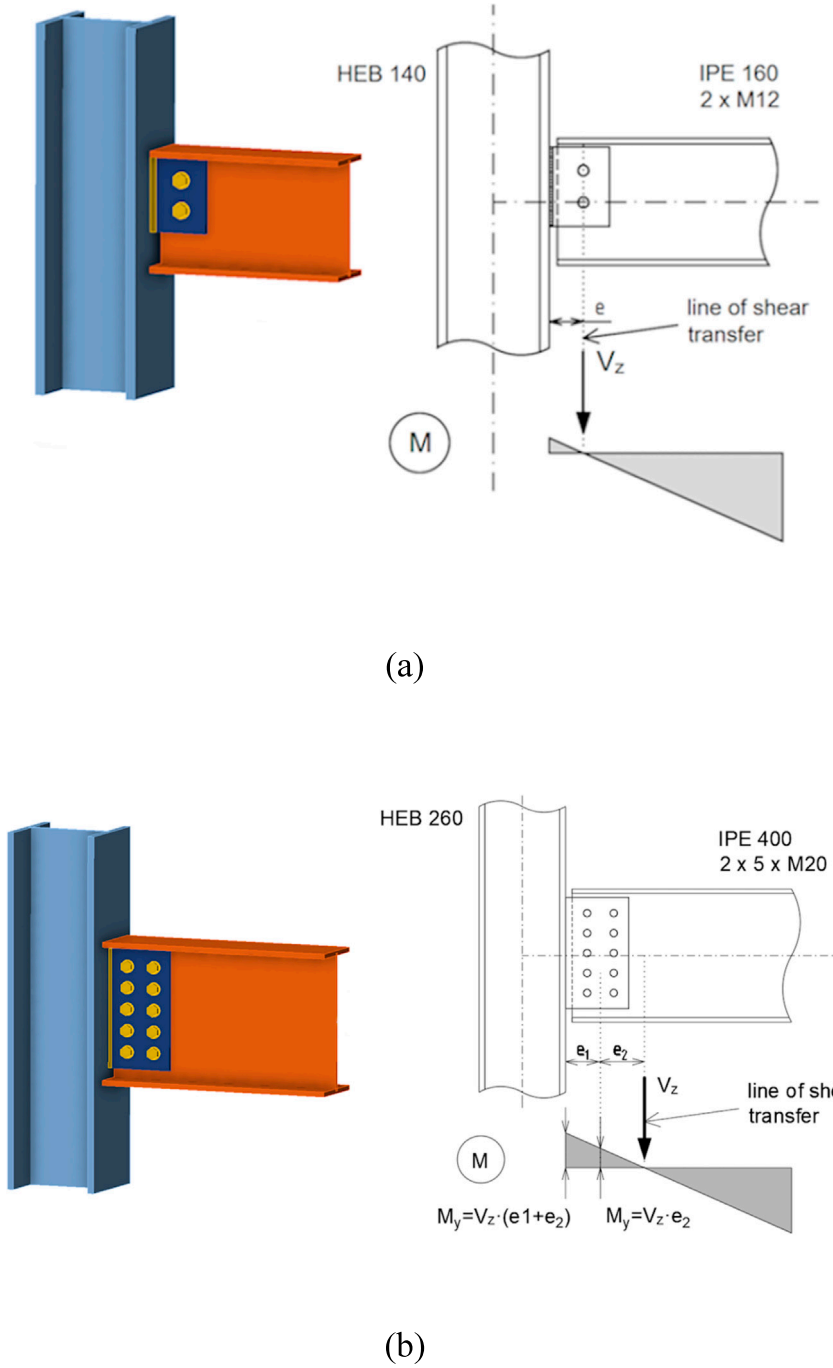


FIGURE 1
Shear connection in steel structures: **(a)** Simple shear connection; **(b)** large plate shear connection of the I-beam to the column (Kalab, 2023).

bolt diameters. A visual representation of these features is provided in Figure 2.

The key output parameter, represented as $f(x) = F_{max}/f_u ndt$, corresponds to the normalized bearing capacity, F_{max} is the maximum measured load in the test, f_u is the ultimate tensile strength of the connected plate, n is the total number of bolts in the connection, d is the bolt-hole diameter, and t is the thickness of the critical plate. This normalization enables standardized comparisons across various scenarios and conditions, ensuring more consistent and reliable analyses. This study utilizes three performance metrics: the *RMSE*, expressed in terms of the normalized bearing capacity ($F_{max}/f_u ndt$), R^2, and a_{20} index. As $F_{max}/f_u ndt$ is dimensionless, the *RMSE* values reported in this study are also dimensionless. Table 2 outlines the detailed ranges for each input parameter, ensuring that the data utilized for the analysis adheres to the specified limits and criteria.

The distribution and coverage of the dataset were examined using the scatterplot matrix shown in Figure 3, which reports

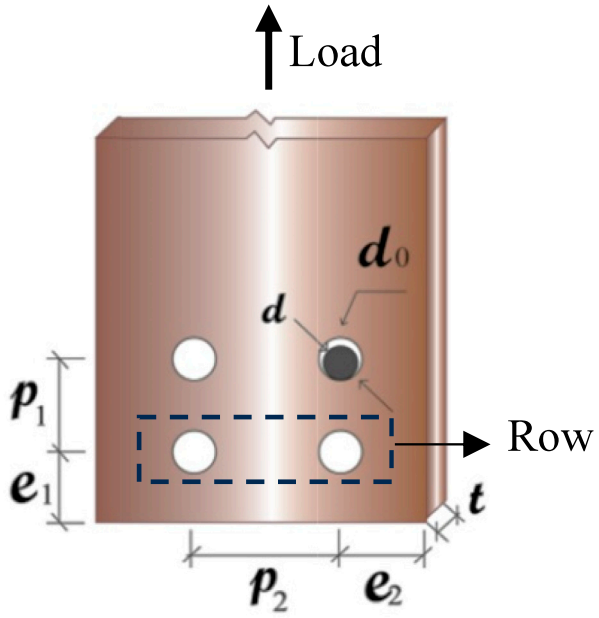


FIGURE 2
Input parameters (Zakir et al., 2022).

histograms of each variable on the diagonal and pairwise scatter plots off the diagonal. The plots indicate that all input features and the normalized bearing capacity span a broad portion of their admissible ranges, with no single narrow interval dominating the sample. While some clustering is visible at practically preferred detailing values (e.g., discrete levels of pitch and bolt rows), the observations are well spread across the joint feature space. In addition, all models are trained and assessed using 10-fold cross-validation with random shuffling, ensuring that each fold contains a representative mix of the feature ranges and reducing the risk of biased performance due to data imbalance.

The range and distribution spread of each variable were examined to assess their variability across the dataset and to critically evaluate the coverage of the input space. This analysis confirmed that the data span a broad and continuous range of values for each input parameter. To further validate the representativeness of the dataset, a sensitivity analysis was conducted to identify and evaluate the most influential input parameters (Sarir et al., 2021b). Our variance-based global sensitivity analysis revealed balanced contributions from all input variables, supporting the conclusion of comprehensive data coverage. Moreover, interpretability techniques such as PDPs and ALEs demonstrated smooth and interpretable relationships across the entire domain of each variable, indicating that the model had sufficient data to learn from throughout the feature space. These insights confirm the robustness of the dataset and its suitability for training predictive ML models.

A correlation matrix was created to gain a deeper understanding of the relationships and interactions among the input parameters, as shown in Figure 4. Each correlation coefficient within the matrix represents the strength of the interaction between two parameters, providing insight into their interdependencies. The analysis revealed several noteworthy associations. A strong correlation coefficient of 0.84 was identified between the number of bolt rows and the normalized pitch parallel to loading, indicating that an increase in the number of bolt rows tends to correspond with an increase in the pitch parallel to loading. A moderate correlation of 0.46 was found between the number of bolt rows and the normalized edge distance, suggesting a moderate relationship where changes in the number of bolt rows can influence the edge distance to some extent. A correlation of 0.44 was observed between the normalized pitch parallel to loading and the normalized edge distance, indicating a moderate relationship where changes in one may affect the other. No significant correlations were found for the other parameters, suggesting that they exhibit independent behaviour and make distinct contributions to the system being studied.

5.1 Preprocessing and hyperparameters fine-tuning

Effective ML model development requires rigorous data preprocessing and systematic hyperparameter tuning to ensure robust and generalizable performance. In this study, all input features were normalized with respect to the bolt diameter d_0 or relevant physical properties to ensure comparability across different test cases and avoid scale-related bias in models sensitive to input magnitude. Specifically, geometric inputs such as end distance, edge distance, and pitch dimensions were divided by d_0 to create dimensionless variables. This representation aligns with established bearing and shear-out design formulations and

TABLE 2 Statistical attributes of dataset.

Attribute	Attribute	Minimum	Maximum	Average
Normalized end distance	e_1/d_0	0.8	5	2.9
Normalized edge distance	e_2/d_0	0.8	7.5	4.15
Normalized pitch parallel to loading	p_1/d_0	0	4	2
Normalized pitch perpendicular to loading	p_2/d_0	0	8.4	4.2
Number of bolt rows	N_r	1	4	2.5
Ultimate to yield stress ratio of the critical steel plate	fu/fy	1.04	1.7	1.37
Normalized bearing capacity	$F_{max}/f_u ndt$	0.28	5.44	2.86

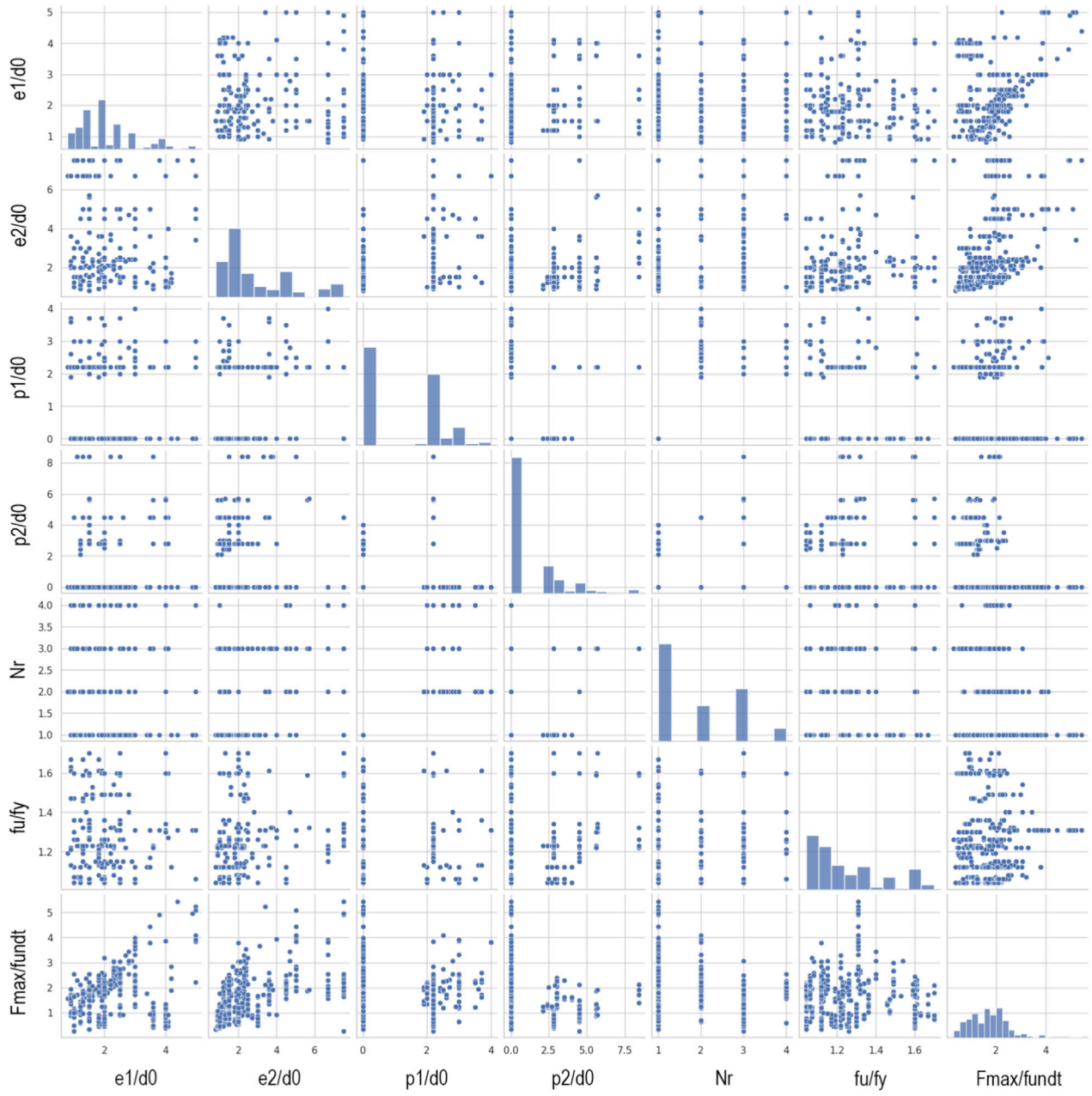


FIGURE 3
Scatterplot matrix showing empirical distributions.

enables the model to generalise across connections with different bolt sizes. From a numerical standpoint, using normalized features mitigates scale-related bias in algorithms that are sensitive to feature magnitude, improving optimization stability and reducing the risk that variables with larger numerical ranges dominate the learning process. If raw, non-normalized geometric dimensions were used instead, the model would be forced to infer the relevant geometric ratios implicitly, potentially reducing generalisation quality across bolt diameters and degrading performance for scale-sensitive algorithms, without offering a clear advantage in predictive accuracy. This standardization improves convergence and stability across ML algorithms. The dataset compiled from 443 experimental results was reviewed for completeness. Since the source data was manually extracted from published studies, any entries with incomplete feature sets were excluded from analysis to avoid bias introduced by imputation. As a result, the dataset used in this study contained no missing values.

Tuning hyperparameters is a crucial process in refining ML models to achieve optimal performance. It entails modifying the settings of the model to obtain the best results for a given dataset. A grid search method is often used for this purpose, where a range of hyperparameter values is systematically explored to find the combination that provides the highest performance. 10-fold CV is incorporated to ensure the selected hyperparameters are reliable



FIGURE 4
Correlation matrix for input variables.

and to prevent overfitting. In this technique, the dataset is divided into 10 roughly equal-sized subsets. The model is trained and evaluated 10 times, with each iteration selecting a different subset as the validation set while the remaining subsets are used for training. This approach systematically explored combinations of hyperparameters and assessed their impact on model performance across multiple data splits. The robustness of this methodology lies not only in its breadth of parameter search but also in the multi-metric evaluation strategy used to determine the best-performing models. This approach guarantees that the performance of the model is assessed across various data splits, offering a more dependable measure of its ability to generalize. For each model and parameter combination, two primary metrics were computed: *RMSE* and R^2. Both metrics were computed for each fold, and their average values across the 10 folds were used to evaluate each hyperparameter set. The following criteria were applied:

- Primary selection metric: The optimal configuration minimized the average *RMSE* across the 10 folds, ensuring minimal prediction error.
- Secondary validation metric: The selected configuration was also required to yield a high average R^2, ensuring that the model not only minimized error but also captured underlying data structure.
- Stability criterion: Configurations that exhibited low variance in *RMSE* across folds were favored, indicating model consistency and robustness.

Different models were tuned with specific parameter ranges, and selections reflected trade-offs between bias and variance. Table 3 presents the optimal hyperparameter values for each ML algorithm, which were selected based on their ability to maximize predictive accuracy while balancing bias and variance.

5.2 Sensitivity analysis

The sensitivity indices for the first-order (S_i) and total effect (S_T) of the input variables are determined using the regression models, as shown in Table 4 and Figure 5. These indices illustrate the contributions of the input variables to the variance of the model output. It can be seen that the variables e_1/d_0 and e_2/d_0 consistently show high contributions to the variance of the model, especially for S_i, across most models. The variable *fu/fy* generally has the lowest sensitivity indices, indicating a negligible influence on the variance of the model. The sensitivity of N_r varies across models, being significant in some cases (e.g., for SVR and RF) but less so in others. The LR and RR models exhibit similar sensitivity patterns, with e_1/d_0

10.3389/fbuil.2026.1753382

TABLE 3 Optimal hyperparameters for $f(x)$.

Model	Optimal configuration
LR	N/A
RR	Alpha = 1
SVR	Kernel = RBF, C = 5, degree = 1, epsilon = 0.1
KNN	Leaf_size = 20, n_neighbors = 2, p = 1
DT	Max_depth = 6, min_samples_leaf = 1, min_samples_split = 3, random_state = 2
RF	Max_depth = 9, max_features = 3, n_estimators = 100
AdaBoost	Learning_rate = 1, n_estimators = 1,000, random_state = 0
XGBoost	Colsample_bytree = 0.5, learning_rate = 0.1, max_depth = 5, n_estimators = 500
CatBoost	Depth = 7, iterations = 500, learning_rate = 0.1
LightGBM	Colsample_bytree = 0.9, learning_rate = 0.1, max_depth = 8, n_estimators = 1,000

TABLE 4 First-order (S_i) and total effect (S_T) sensitivity indices for $f(x)$.

Model	Sensitivity index	e_1/d_0	e_2/d_0	p_1/d_0	p_2/d_0	N_r	fu/fy	$\sum$
LR	S_i	0.355365	0.393051	0.002714	0.017332	0.228118	0.007624	1.004204
	S_T	0.354044	0.394288	0.00286	0.017706	0.227382	0.007668	1.003948
RR	S_i	0.356894	0.395383	0.002182	0.018554	0.224645	0.006471	1.004129
	S_T	0.355674	0.39668	0.002312	0.018931	0.223881	0.006512	1.00399
SVR	S_i	0.059165	0.154455	0.014107	0.149582	0.178913	0.00158	0.557802
	S_T	0.307411	0.288294	0.100557	0.495602	0.282168	0.004543	1.478575
KNN	S_i	0.013174	0.229624	0.038107	0.131202	0.025796	0.002347	0.44025
	S_T	0.350931	0.526861	0.219344	0.543754	0.271965	0.023632	1.936487
DT	S_i	0.104547	0.306555	0.001471	0.000618	0.214867	0.069672	0.69773
	S_T	0.27974	0.510827	0.012023	0.024494	0.378748	0.200154	1.405986
RF	S_i	0.321323	0.31375	0.060056	0.118243	0.095438	0.002225	0.911035
	S_T	0.379783	0.348437	0.068097	0.151227	0.119505	0.025398	1.092447
AdaBoost	S_i	0.235724	0.327846	0.026983	0.046607	0.127748	0.00376	0.768668
	S_T	0.429271	0.407525	0.046982	0.07014	0.275602	0.026335	1.255855
XGBoost	S_i	0.336854	0.358612	0.047908	0.012907	0.070186	0.02074	0.847207
	S_T	0.407268	0.421556	0.087578	0.021108	0.111595	0.047171	1.096276
CatBoost	S_i	0.166621	0.273095	0.037458	0.168462	0.062072	0.006087	0.713795
	S_T	0.355354	0.339181	0.10215	0.314361	0.138233	0.059549	1.308828
LightGBM	S_i	0.20592	0.506169	0.015883	0.002926	0.094909	0.010239	0.836046
	S_T	0.333635	0.600722	0.031425	0.020936	0.143146	0.030892	1.160756

and e_2/d_0 contributing the most to the output variance ($S_i \approx$ 0.35–0.39). For the SVR model, sensitivity indices are more evenly distributed among variables, with p_2/d_0 and N_r showing significant contributions. In KNN, S_T indicates high combined interactions (S_T values exceed 1.9). DT shows relatively even sensitivity, with moderate values for e_1/d_0 and e_2/d_0. RF highlights strong contributions from e_1/d_0, e_2/d_0, and interactions among other variables ($S_T \approx 1.1$). For the AdaBoost model, variables e_1/d_0 and e_2/d_0 dominate both S_i and S_T. XGBoost, CatBoost, and LightGBM generally show high interaction effects, with S_T often exceeding 1. The e_2/d_0 variable consistently has the highest impact. The total variance contribution for first-order indices ($\sum S_i$) is less

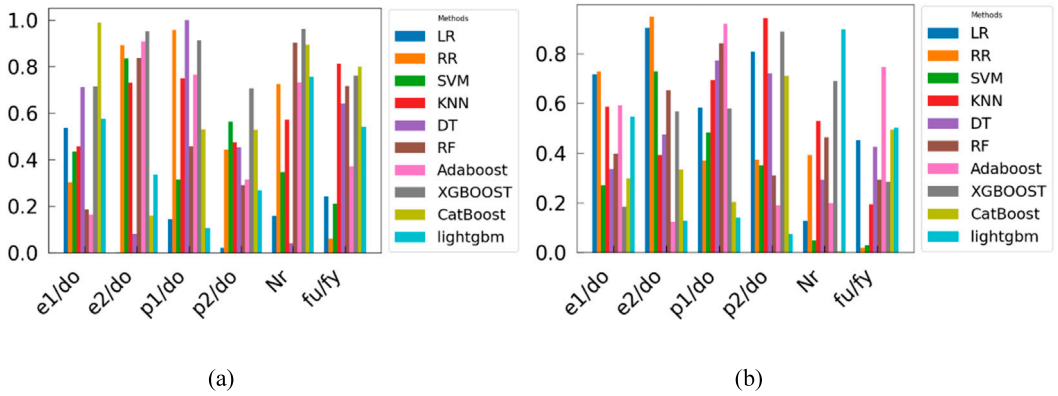


FIGURE 5
Sensitivity analysis of $f(x)$: **(a)** First-order; **(b)** Total-effect.

than or close to 1 in most cases, indicating that individual variables largely explain the model variance. For total effects ($\sum S_T$), values exceed 1 for certain models (e.g., KNN, LightGBM), suggesting significant interaction effects among the input variables.

In general, this SA is essential for identifying critical inputs and their interactions, guiding feature selection or model improvement strategies. The consistently high indices of e_1/d_0 and e_2/d_0 confirm them as primary design drivers, which can therefore be prioritised in the analyses. By contrast, the systematically low first-order and total-effect indices of p_1/d_0 indicate that it is a secondary descriptor; it can be fixed or omitted in simplified surrogate models with limited expected impact on predictive performance, although it is retained here for physical completeness. From a model-development standpoint, the pronounced gaps between S_i and S_T for KNN, RF, XGBoost, CatBoost, and LightGBM confirm the presence of strong interaction effects, justifying the adoption of tree-based ensemble models that can capture higher-order interactions and guiding both the selection of CatBoost as the final surrogate and the design of the two-variable contour analyses.

5.3 Regression model

Figure 6 presents a set of regression plots comparing predicted and measured values across different ML models. Each subfigure contains four plots:

a. Histogram of Residuals: Displays the distribution of residuals (measured-predicted), showing the error spread;
b. Residuals vs. Predicted Values: A scatterplot highlighting how residuals vary across the predicted range. Red and blue points distinguish between two datasets or conditions;
c. Predicted vs. Measured Scatterplot: A scatterplot comparing predicted values against actual measured values, with a diagonal line ($y = x$) as a reference for perfect predictions;
d. Histogram of Predictions: Displays the distribution of predicted values in relation to the actual measured range.

The diagonal line in the predicted vs. measured scatterplot serves as a reference to evaluate prediction accuracy. Points closer to the line indicate better performance. The spread in the residuals vs. predicted values plot reveals whether the model has systematic bias (e.g., over- or under-prediction for certain ranges). Differences in the residual and prediction histograms reflect variations in model accuracy and distribution alignment. The figure highlights how different models behave in terms of prediction accuracy, bias, and variance. Color coding (e.g., red and blue points) differentiates between subsets of data or categories, indicating performance consistency across different data types.

From visual inspection, XGBoost, CatBoost, and LightGBM appear to perform better. The residual histograms for these models are narrower and more centred around 0 compared to others. In the residual vs. predicted plots, no clear patterns or trends are visible, indicating reduced systematic error. The scatterplot of predicted vs. measured values shows that points are closer to the diagonal line, reflecting better prediction accuracy. LR and KNN seem to perform the worst, as they have wider residual distributions and systematic patterns in the residuals vs. predicted plots, suggesting poorer performance.

In general, XGBoost, CatBoost, and LightGBM leverage gradient boosting techniques, which combine the strengths of multiple weak learners (e.g., decision trees) to iteratively minimize error. When properly tuned, these models are robust to overfitting and effectively handle complex relationships and interactions within the data. In contrast, simpler models like LR or KNN may fail to capture non-linear relationships or interactions, leading to poorer performance.

In conclusion, the gradient boosting models (XGBoost, CatBoost, and LightGBM) perform better due to their superior ability to capture complex patterns, resulting in tighter residual distributions, minimal bias, and high alignment between predicted and measured values.

Figure 7 shows Taylor Diagrams used to evaluate and compare the performance of different ML models during (a) the training phase and (b) the testing phase. Taylor diagrams provide a concise summary of how well a model matches observations by plotting three statistics simultaneously:

- Standard Deviation (horizontal axis): This indicates the variability of the predicted data relative to the observed

10.3389/fbuil.2026.1753382

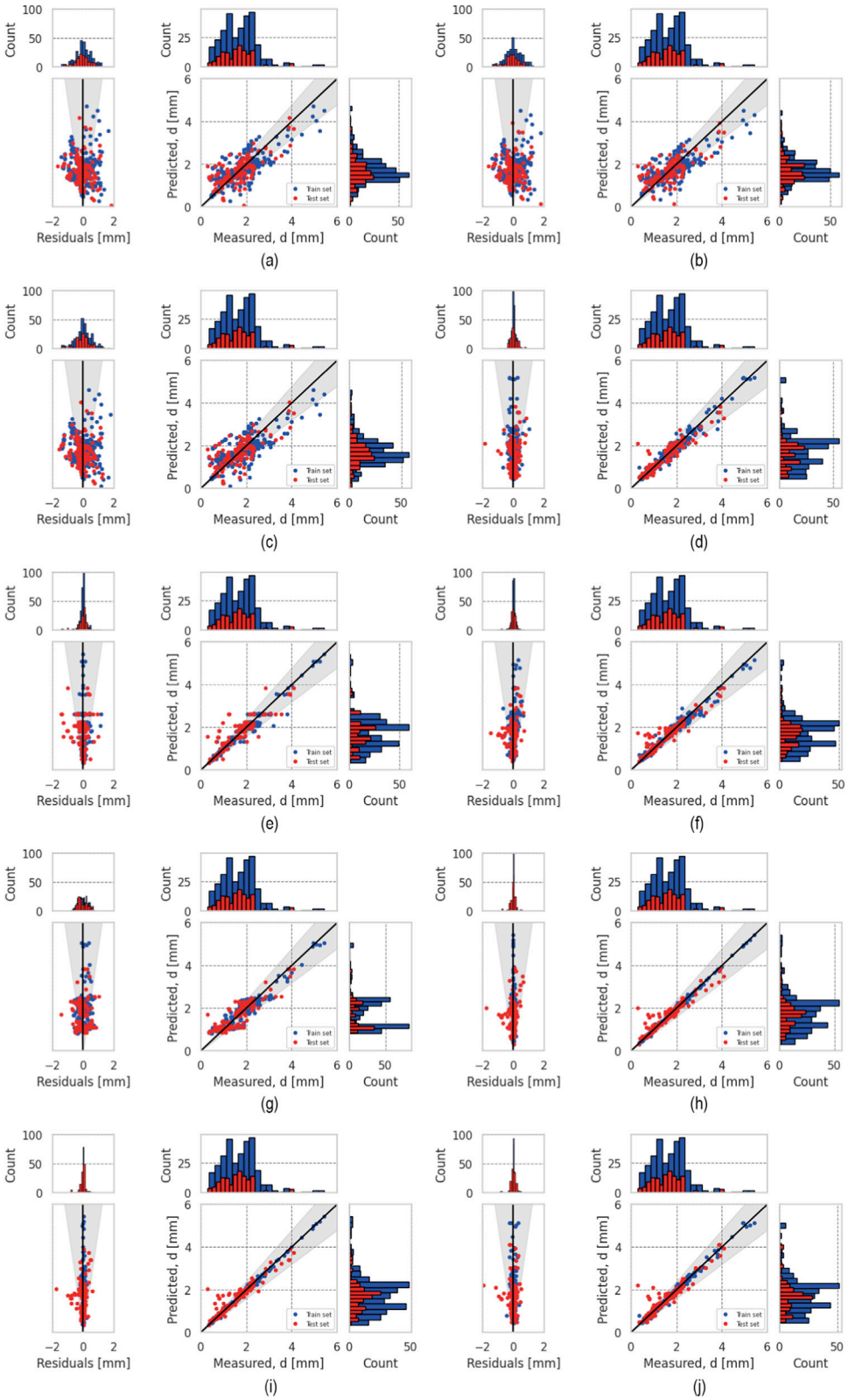


FIGURE 6
Regression plot of *f*(*x*): **(a)** LR; **(b)** RR; **(c)** SVR; **(d)** KNN; **(e)** DT; **(f)** RF; **(g)** AdaBoost; **(h)** XGBoost; **(i)** CatBoost; **(j)** LightGBM.

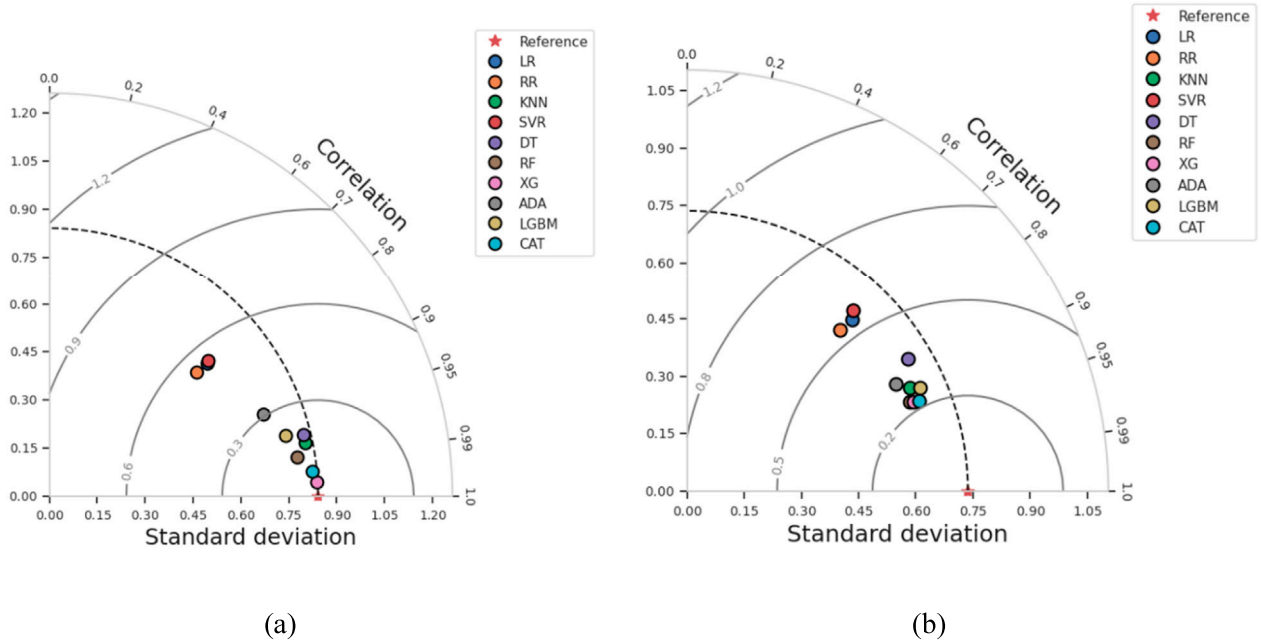


FIGURE 7
Taylor Diagram of ML models: **(a)** training; **(b)** testing.

data. Ideally, the standard deviation of the model (represented by the markers) should be close to that of the reference (red star).

- Correlation Coefficient (curved axis): This measures the linear relationship between predicted and observed data. Values closer to 1 indicate a strong correlation (good agreement between predicted and observed).
- Cantered Root Mean Square Error (*CRMSE*) (distance from the red star): This represents the overall error magnitude between predicted and observed data, excluding bias. Smaller distances imply better performance.

Figure 7a shows the training phase. Most models cluster near the reference, indicating high correlation and standard deviation close to the reference during training. Certain models, such as XGBoost and LightGBM, performed better by showing higher correlation and smaller *CRMSE* compared to others. Figure 7b presents the testing phase. The models show more variation in their correlation and standard deviation compared to the training phase. The scatter suggests that some models generalize better to unseen data (higher correlation and smaller *CRMSE*), while others may overfit the training data. It can be seen that the CatBoost model outperforms the others, as it maintains relatively high correlation coefficients compared to the other models, indicating good generalization on unseen data. Its standard deviation remains closer to the reference compared to the other models, and its *CRMSE* value (distance to the red star) is smaller than those of the other models, suggesting lower prediction errors on the testing data.

Table 5 evaluates the performance of various ML-based regression models for predicting $f(x)$. The metrics used to compare the models are Average R^2, Average *RMSE*, and a_{20} index. Both LR and RR have low R^2 values (~0.286–0.320), indicating poor explanatory power. The *RMSE* is relatively high (0.554–0.556), suggesting less accurate predictions. Both SVR and KNN show significantly higher R^2 values (~0.685–0.686), indicating better predictive accuracy. KNN has the lowest *RMSE* (0.334), suggesting it provides the most precise predictions among the models. DT presents moderate performance with R^2 = 0.320, *RMSE* = 0.526, and a_{20} = 0.805. RF shows better performance (R^2 = 0.703, *RMSE* = 0.363, and a_{20} = 0.827). CatBoost has the highest R^2 (0.711), showing the best explanatory power among all models. XGBoost and CatBoost have relatively low *RMSE* values

TABLE 5 Performance of ML based regression models for $f(x)$ prediction.

Regression method	Average R^2	Average *RMSE*	a_{20}
LR	0.286	0.554	0.511
RR	0.320	0.556	0.511
SVR	0.686	0.362	0.496
KNN	0.685	0.360	0.759
DT	0.320	0.526	0.805
RF	0.703	0.363	0.827
AdaBoost	0.522	0.454	0.617
XGBoost	0.688	0.359	0.857
CatBoost	**0.711**	**0.359**	**0.872**
LightGBM	0.573	0.393	0.835

Bold values indicate the best performance.

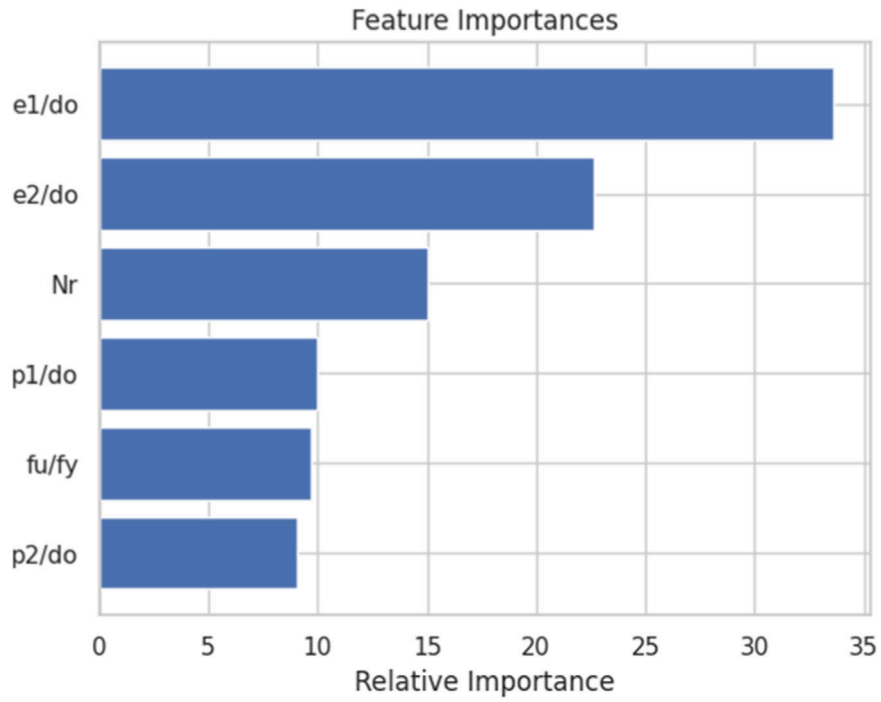


FIGURE 8
CatBoost feature importance.

(0.359), indicating good predictive precision. AdaBoost and LightGBM perform moderately, with lower R^2 values and higher *RMSE* compared to CatBoost and XGBoost. As a result, CatBoost has the highest R^2 (0.711), indicating it explains the target variable variance the best.

Among all models, the CatBoost model achieves the highest average R^2 and one of the lowest *RMSE* values. Given that the normalised bearing capacity ranges from 0.28 to 5.44 with an average of 2.86, this *RMSE* corresponds to an error of approximately 12%–13% of the mean response and about 7% of the full range. Furthermore, the a_{20} index of 0.872 indicates that around 87% of predictions lie within ±20% of the experimental values. Although this does not eliminate variability, it represents a substantial improvement over simple linear models and is comparable to the scatter typically observed between design-code predictions and test results for bolted connections. On this basis, CatBoost is selected as the reference surrogate model for the interpretability and parametric analyses in the following sections, with the understanding that it is intended as a complementary, not stand-alone, design tool.

6 Interpretable methods

Figure 8 displays the feature importance derived from a CatBoost model. The bar chart shows the relative importance of features, ranked in descending order of their contribution to model predictions. The features and their corresponding relative importance values are as follows:

- e_1/d_0: The most important feature, with the highest relative importance of approximately 35%.
- e_2/d_0: The second most important feature, slightly lower than e_1/d_0, with a relative importance of around 30%.
- N_r: The third feature, with a relative importance close to 20%.
- p_1/d_0: A moderately significant feature, with a relative importance of about 10%.
- *fu/fy*: A less significant feature, with a relative importance around 9%.
- p_2/d_0: The least significant feature, with a relative importance of about 8%.

The chart effectively highlights the varying contributions of these features, helping to identify which factors most strongly influence the performance of the model. This insight can guide further analysis or refinement of the predictive model.

6.1 Partial dependence plot

PDPs illustrate the relationship between a specific feature and the predicted outcome of a ML model, while keeping all other features fixed, as shown in Figure 9. Each subplot corresponds to a different input variable affecting the output of the model. The *x*-axis represents the values of the specific input variables, and the *y*-axis represents the Partial Dependence, or how the predictions of the model vary with changes in that specific variable. The black ticks indicate the distribution of data points (or unique values) for that variable in the training dataset. Denser areas suggest more frequent values in the data.

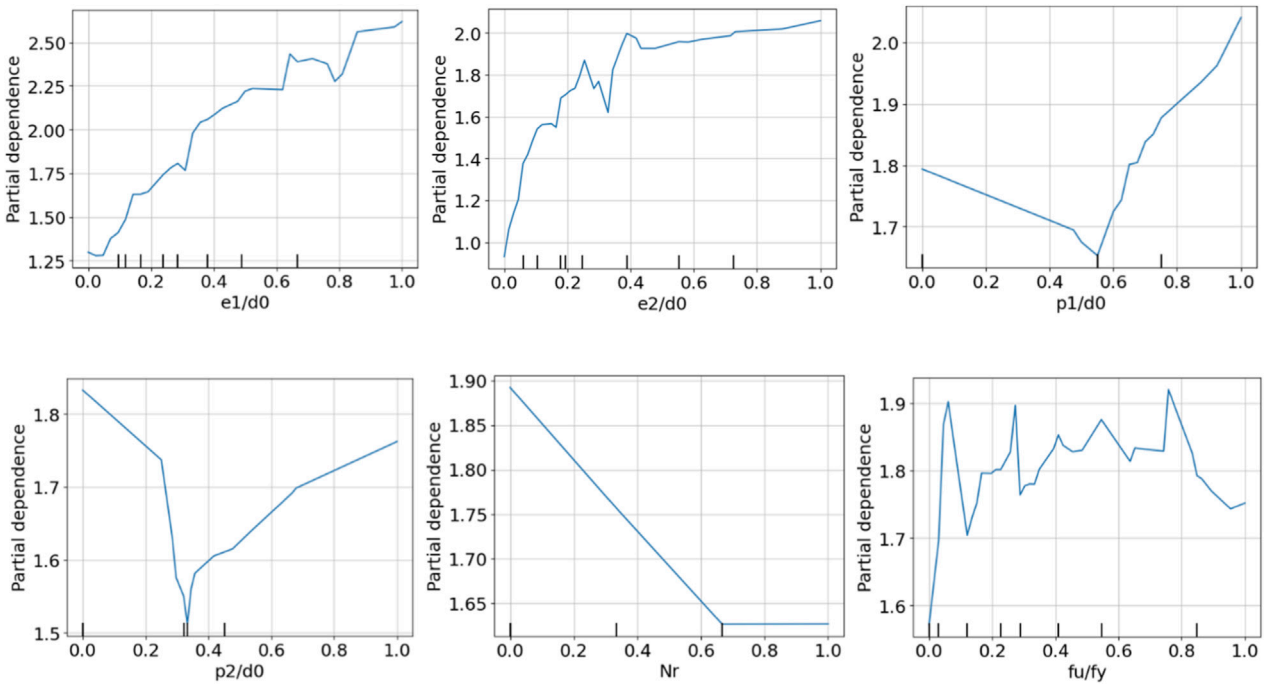


FIGURE 9
PDP of input variables for $f(x)$.

The observations for each variable are as follows:

- e_1/d_0: Partial dependence increases almost monotonically, indicating a positive relationship between e_1/d_0 and the model output.
- e_2/d_0: A general upward trend is observed, although there are some fluctuations, suggesting a positive but slightly nonlinear relationship.
- p_1/d_0: A U-shaped relationship appears, indicating that intermediate values might negatively impact the prediction, while both higher and lower values increase the model output.
- p_2/d_0: A similar U-shape is observed, but with less fluctuation compared to p_1/d_0, as indicated by the smaller vertical range of the partial dependence values.
- N_r: The dependence decreases linearly, suggesting a strong negative relationship with the output.
- *fu/fy*: The relationship is irregular, with fluctuations indicating a complex interaction with the output variable.

Variables like N_r and e_1/d_0 seem to have clearer trends, while *fu/fy* might require further investigation due to its variability.

6.2 Accumulated local effects

ALE represents the average effect of the input variable on the prediction of the model, accounting for interactions with other variables, as shown in Figure 10. Positive ALE values suggest that the variable contributes positively to the prediction, while negative ALE values indicate a negative contribution. The *x*-axis represents the range of the respective input variable, normalized between 0 and 1.

Each plot corresponds to the effect of a specific input variable on the predictions of the model:

- e_1/d_0: Shows a positive correlation; as e_1/d_0 increases, the ALE increases significantly, indicating a strong positive effect.
- e_2/d_0: Exhibits non-linear fluctuations, suggesting complex behaviour.
- p_1/d_0: Displays a moderate upward trend with a dip around 0.5.
- p_2/d_0: Shows a U-shaped curve, indicating a decrease in ALE up to a certain point, followed by an increase.
- N_r: Displays a negative slope, indicating that N_r contributes negatively to the prediction.
- *fu/fy*: The ALE increases rapidly at lower values and then stabilizes, suggesting a diminishing positive effect.

Tick marks on the *x*-axis represent the distribution of data points for the respective variable, indicating where the model is well-supported by data.

6.3 Shapley additive explanations

Figure 11a presents the SHAP summary plot, where each point corresponds to a Shapley value associated with a specific parameter. The plot arranges samples into rows, each containing an equal number. The Shapley values are plotted along the *x*-axis, while the input variables are listed on the *y*-axis in order of importance, with the most significant ones at the top. Samples with identical SHAP values for a factor are spread horizontally. The variable values are color-coded, with high values in red and low values in blue. Red

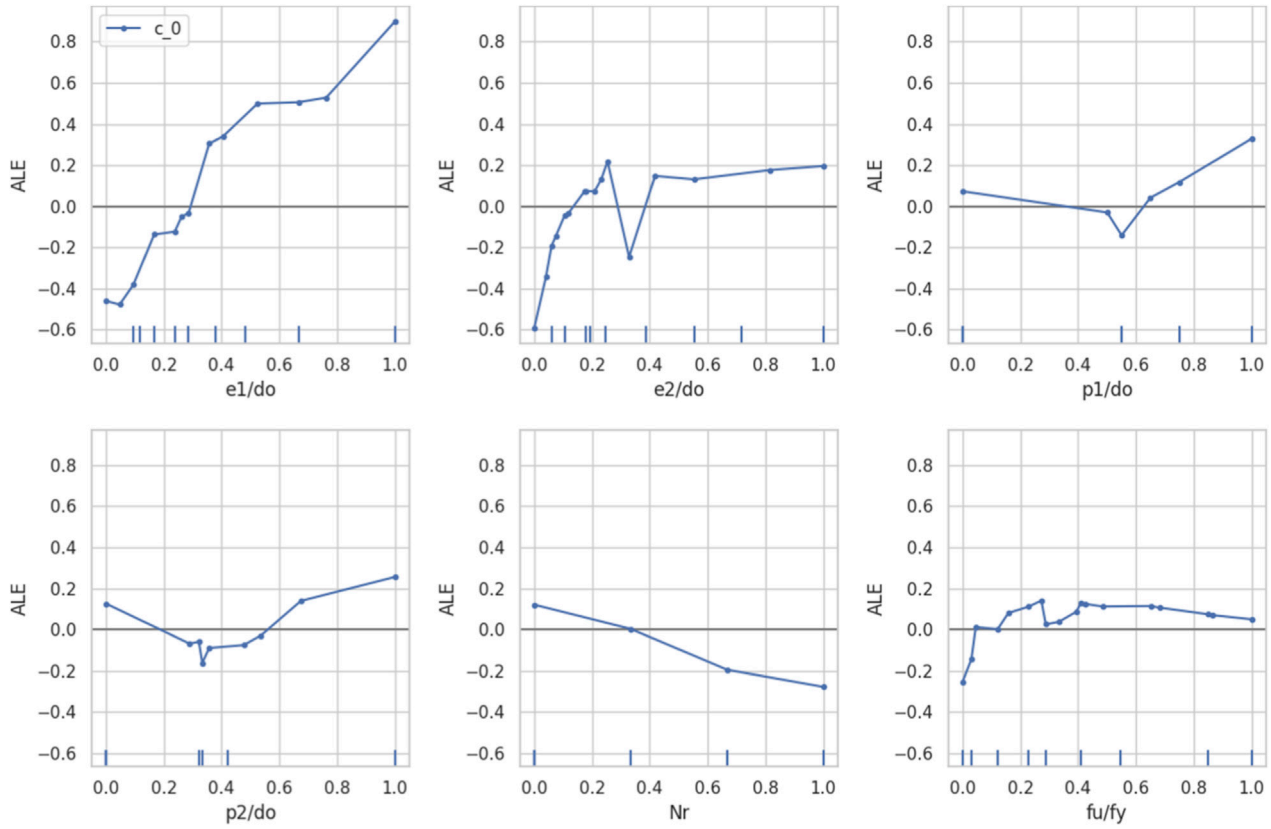


FIGURE 10
ALE of input variables for $f(x)$.

indicates values that increase the SHAP score and the associated estimate.

The data suggests that increasing the normalized pitch perpendicular to loading and the number of bolt rows leads to a decrease in the SHAP value, and thus a reduction in the normalized bearing capacity. Conversely, increasing the normalized end distance, normalized edge distance, or the ultimate-to-yield stress ratio of the critical steel plate results in an increase in the normalized bearing capacity. Figure 11a ensures an even distribution of points across the rows. In Figure 11b, the global significance factor is represented by the average absolute SHAP value for each factor. SHAP analysis identifies normalized end distance and edge distance as the most significant factors, aligning with the variable importance results from the CatBoost model.

Figure 12 presents the SHAP dependence graph through scatter plots that show the SHAP value of one parameter against others. The colours in the graph signify how interactions with different variables affect the values on the horizontal axis, with many of these interactions being non-linear. In Figure 12a, the SHAP values increase as e_1/d_0 increases, indicating that higher values of e_1/d_0 have a more positive contribution to the predictions of the model. The colour gradient represents N_r, where higher N_r values tend to amplify the impact of e_1/d_0 on the predictions. In Figure 12b, the SHAP values generally increase as e_2/d_0 increases, although the relationship appears to level off or saturate at higher values. The colour gradient corresponds to e_1/d_0, showing that lower values of e_1/d_0 are associated with smaller SHAP contributions, while higher e_1/d_0 amplifies the SHAP values for e_2/d_0. In Figure 12c, for most of the range of p_1/d_0, the SHAP values are clustered near zero, but there is an increasing trend at higher p_1/d_0 values, indicating a positive contribution to predictions. The colour gradient for e_1/d_0 shows that higher e_1/d_0 values correspond to more positive SHAP values, suggesting an interaction between p_1/d_0 and e_1/d_0.

In Figure 12d, the SHAP values are distributed across a wide range, with noticeable clusters, suggesting non-linear relationships. There are regions where p_2/d_0 has little to no effect (SHAP values near zero) and others where it contributes positively or negatively. The feature e_1/d_0 interacts with p_2/d_0, with higher e_1/d_0 amplifying the SHAP values. In Figure 12e, the SHAP values exhibit a slight overall decreasing trend with N_r, indicating that larger N_r values generally contribute negatively to the predicted normalized bearing capacity. The colour gradient for e_1/d_0 further shows that this negative influence is more pronounced for lower e_1/d_0 and can be partially mitigated when e_1/d_0 is large. In Figure 12f, the SHAP values are scattered without a strong linear trend, indicating a complex or weak relationship between *fu*/*fy* and the predictions of the model. Certain clusters show positive or negative contributions. The colour bar for p_2/d_0 suggests that the interaction between *fu*/*fy* and p_2/d_0 is significant, with higher p_2/d_0 values leading to more pronounced positive or negative SHAP contributions.

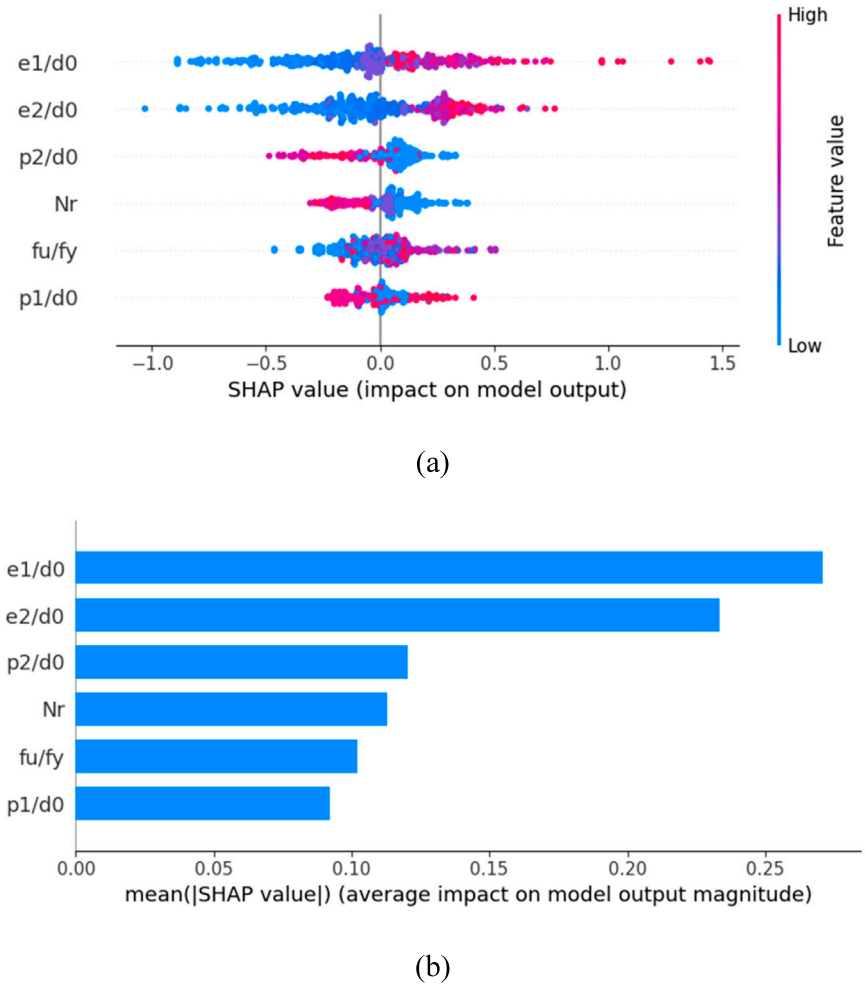


FIGURE 11
SHAP plots: **(a)** Shapley value; **(b)** Global importance factor.

In general, the behaviour of SHAP values indicates that e_1/d_0 and e_2/d_0 have the strongest and most consistent positive impact on predictions. Features like p_2/d_0 and *fu/fy* exhibit non-linear or scattered relationships, implying complex interactions within the model. The colour gradients (interaction effects) reveal secondary relationships between the displayed feature and auxiliary features, emphasizing the multi-feature dynamics of the model.

Figure 13 shows a SHAP force plot, which visualizes how individual feature values influence the prediction of the model. The plot represents how different input variables contribute to pushing the prediction away from its base value (the expected value of the model output across the dataset). The red section represents features that push the prediction higher. The blue section represents features that push the prediction lower. The grey vertical line in the middle indicates the base value of the prediction. 1.73 is the expected value of the model output before considering the influence of individual feature values. Features shown in red, including *fu/fy*, p_2/d_0, N_r, e_2/d_0, increase the output from the base value. Features shown in blue, including e1/d0 decrease the output from the base value. In general, the feature interpretations are as follows:

- *fu/fy* = 0.8485: A feature that pushes the output upwards.
- N_r = 0.3333: Also contributes positively.
- p_2/d_0 = 0: Has a small or neutral effect.
- P_1/d_0 = 0.55: provides a positive contribution.
- e_2/d_0 = 1: Significantly increases the output.
- e_1/d_0 = 0.04762: The only feature that pulls the model output lower.

Figure 14 presents a SHAP decision plot. The order of features on the *y*-axis suggests their relative importance. Higher-ranked features, such as e_1/d_0, e_2/d_0, have a larger overall impact on the output. Features lower on the axis, such as p_1/d_0, contribute less overall, indicating they may be less influential. The blue-to-red gradient shows how feature values relate to their SHAP impact. For e_1/d_0, higher values (red) mostly result in positive SHAP values, meaning high e_1/d_0 pushes the output of the model to increase. For e_2/d_0, similarly, higher values tend to have a positive SHAP impact, but some variability suggests a non-linear relationship. Low feature values (blue) generally have a negative or neutral SHAP impact, pulling the output lower. Features like p_2/d_0 and N_r show tightly clustered SHAP values

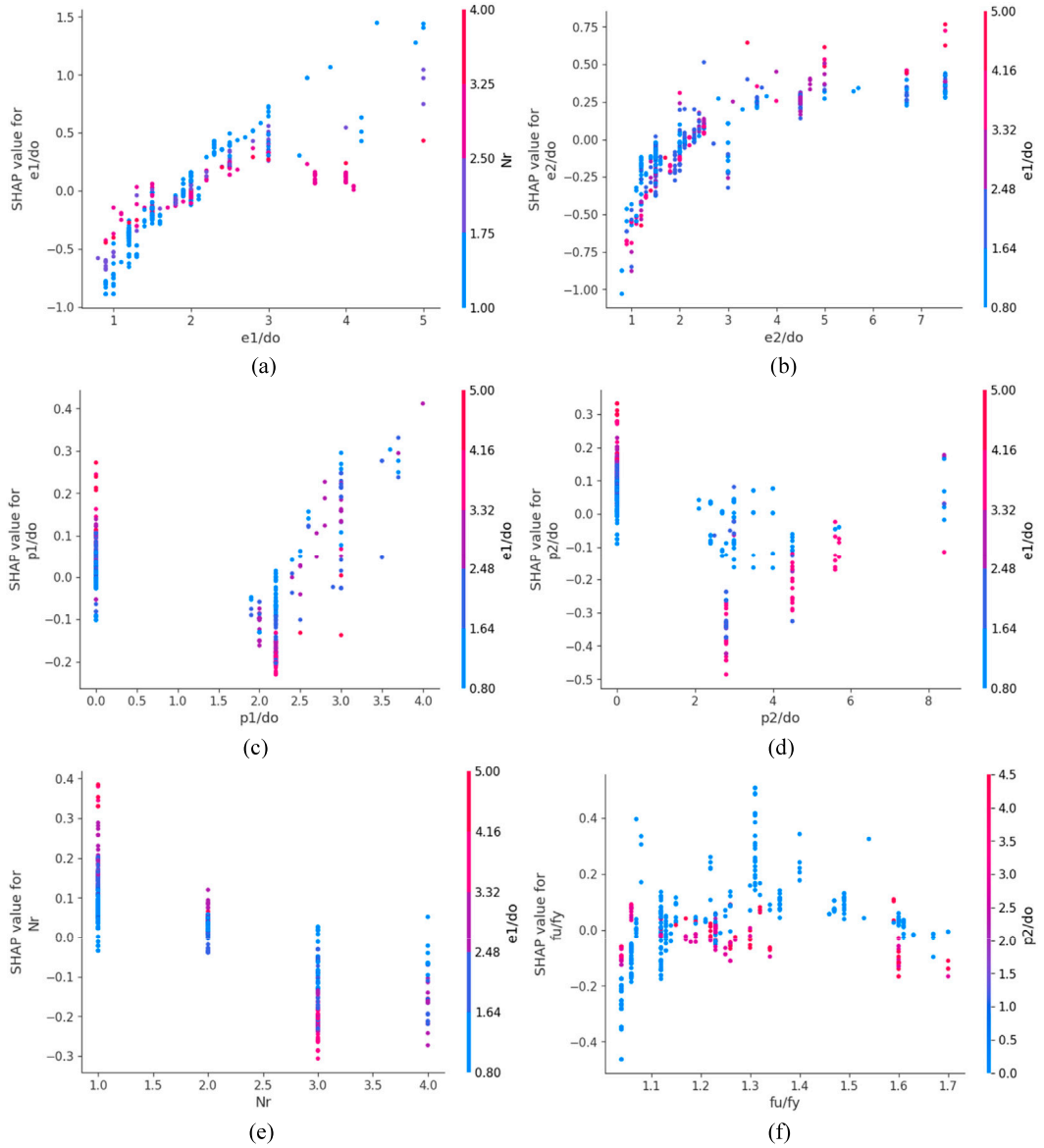


FIGURE 12
SHAP partial dependence plots: **(a)** e1/d0-Nr; **(b)** e2/d0-e1/d0; **(c)** p1/d0-e1/d0; **(d)** p2/d0-e1/d0; **(e)** Nr-e1/d0; **(f)** fu/fy-p2/d0.



FIGURE 13
SHAP force plot.

around 0. This could mean their contributions are smaller and more uniform across the dataset. Conversely, features like e_1/d_0 have a wider spread, indicating a more variable contribution to the output. In SHAP plots, overlapping colours and values suggest potential feature interactions. For example, the spread in N_r and *fu/fy* might indicate interactions with other features, where their SHAP values depend on the presence of other variables. The model output appears to centre near 2, suggesting that for most samples, predictions are not strongly influenced (in either positive or negative directions) by individual features. Features like e_1/d_0 and e_2/d_0 have notable impacts, pulling the output above or below this centre.

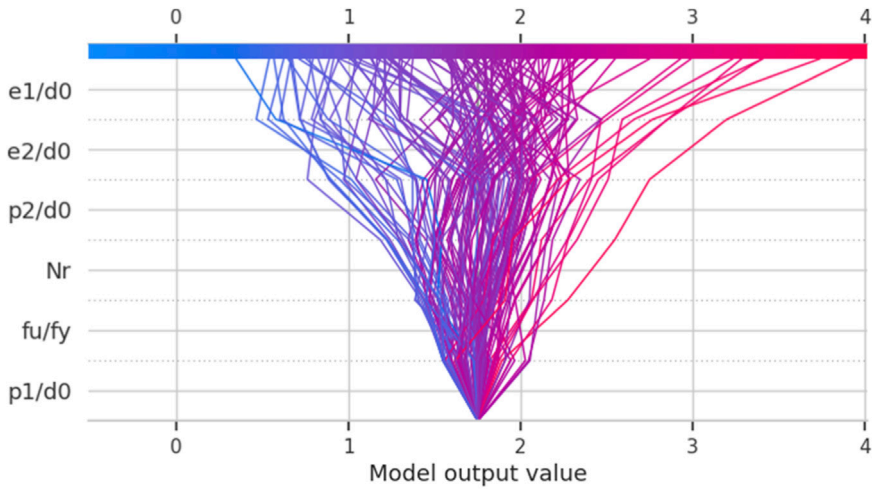


FIGURE 14
SHAP decision plot.

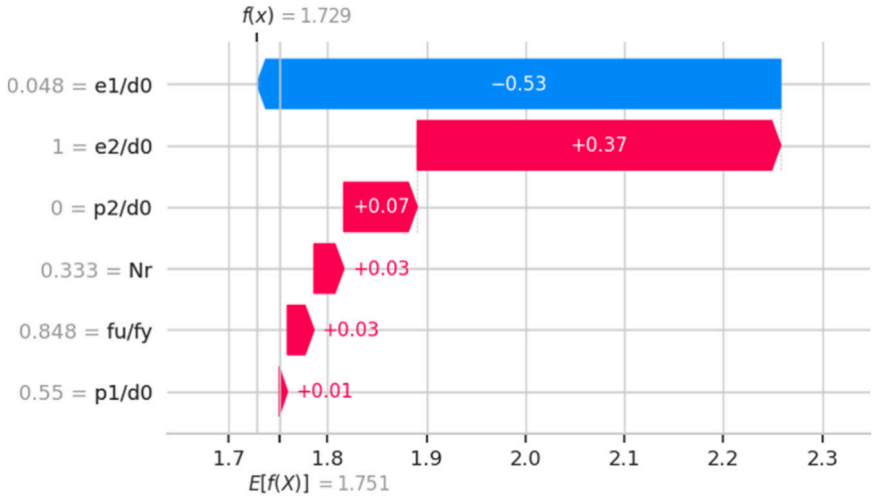


FIGURE 15
SHAP waterfall plot.

Figure 15 shows a SHAP waterfall plot, which is commonly used to explain the contribution of different input variables to a ML prediction for a single instance. The top-left part of the plot states $f(x) = 1.729$, which is the predicted output for the given instance. The bottom label shows $E[f(X)] = 1.751$, which represents the expected prediction across all instances in the dataset. The largest impact comes from e_1/d_0, which decreases the prediction by 0.53. The second major factor is e_2/d_0, which increases the prediction by 0.37. Smaller contributions are made by p_2/d_0 (+0.07), N_r (+0.03), *fu/fy* (+0.03), and p_1/d_0 (+0.01).

6.4 Parametric study

Figure 16 shows a series of contour plots that represent the output of a two-variable exponential decay function. Each subplot displays the relationship between two input variables (on the axis) and the corresponding output of the decay function (indicated by the contour levels and the colour gradient). The *x*- and *y*-axis in each subplot represent different parameter pairs or input variables. Each subplot investigates how these variable pairs interact and influence the output of the function. The contour lines connect points of equal output values, providing a visual representation of the decay behaviour. The colour map indicates the magnitude of the output, where typically:

- Red shades signify higher output values.
- Blue shades signify lower output values.

The objective is likely to analyse and fit a two-variable exponential decay model to some data. The plots help visualize how different parameter combinations affect the output, which may

10.3389/fbuil.2026.1753382

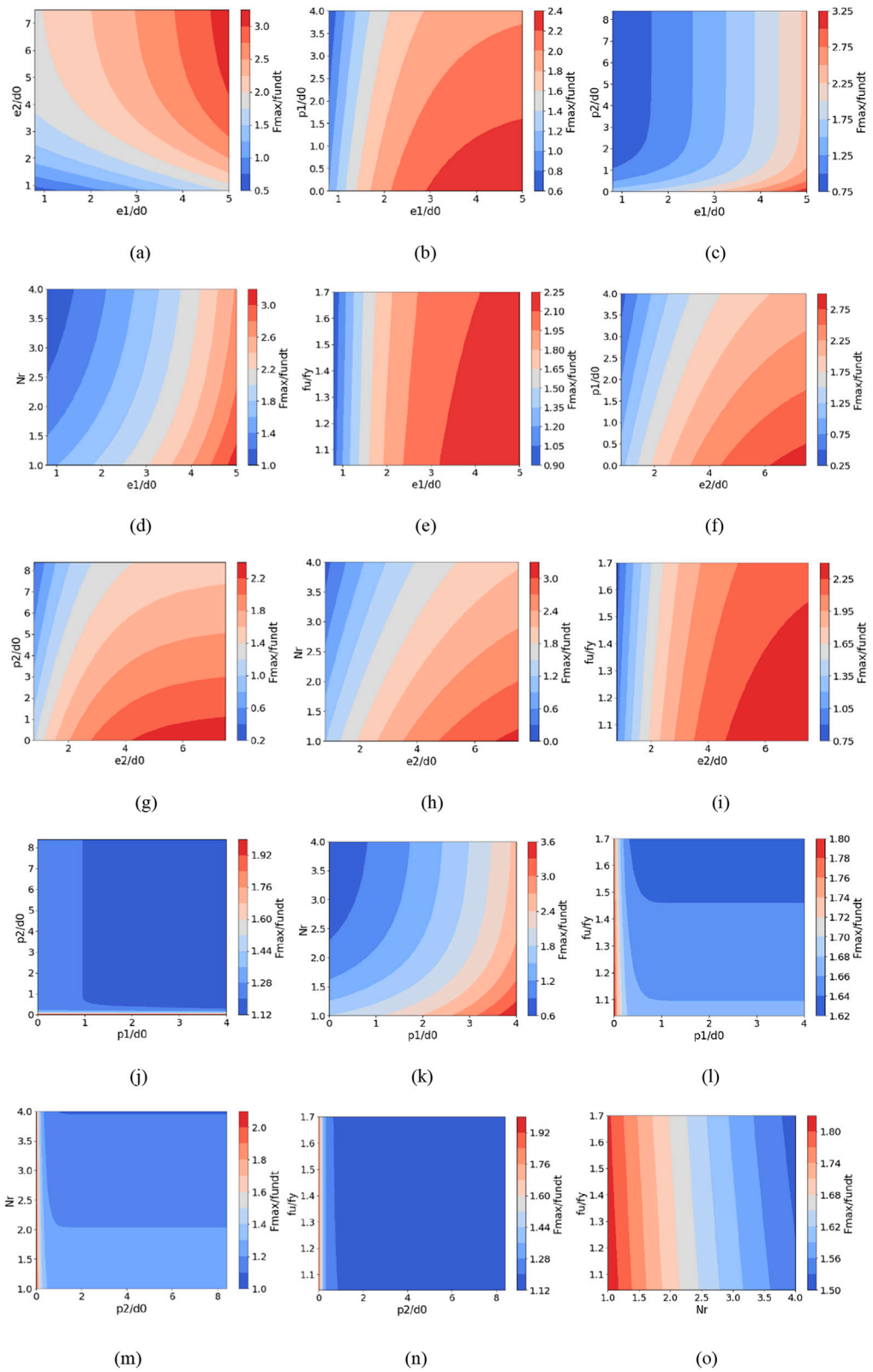


FIGURE 16
Two-variable exponential decay fit of output function: **(a)** e1/d0-e2/d0; **(b)** e1/d0-p1/d0; **(c)** e1/d0-p2/d0; **(d)** e1/d0-Nr; **(e)** e1/d0-fu/fy; **(f)** e2/d0-p1/d0; **(g)** e2/d0-p2/d0; **(h)** e2/d0-Nr; **(i)** e2/d0-fu/fy; **(j)** p1/d0-p2/d0; **(k)** p1/d0-Nr; **(l)** p1/p0-fu/fy; **(m)** p2/d0-Nr; **(n)** p2/d0-fu/fy; **(o)** Nr-fu/fy.

be useful for identifying trends, dependencies, or sensitivity of the function to specific variables. Peaks (red areas) and troughs (blue areas) in the contour plots highlight regions of high and low output values, respectively. Smooth transitions or sharp gradients in colour and contour lines indicate the rate of change in the output with respect to changes in the input variables.

Figure 16a illustrates the relationship between the parameters e_1/d_0, e_2/d_0, and the normalized variable $F_{max}/f_u ndt$. As e_1/d_0 and e_2/d_0 increase, $F_{max}/f_u ndt$ also increases, indicating a direct correlation between these variables. The gradient of the contour lines suggests a nonlinear interaction, with a steeper increase in $F_{max}/f_u ndt$ for higher values of e_2/d_0. Figure 16b represents the variation of the parameter $F_{max}/f_u ndt$ with respect to e_1/d_0 and p_1/d_0. The contours show that $F_{max}/f_u ndt$ increases as both e_1/d_0 and p_1/d_0 increase, with a steeper gradient in regions of higher e_1/d_0. The behaviour suggests a nonlinear relationship, where changes in p_1/d_0 influence $F_{max}/f_u ndt$ more strongly for larger e_1/d_0 values. Figure 16c illustrates the relationship between two normalized parameters, e_1/d_0 and p_2/d_0, and their influence on the response parameter $F_{max}/f_u ndt$. As e_1/d_0 increases, the response parameter $F_{max}/f_u ndt$ grows significantly, particularly for higher values of p_2/d_0. The behaviour suggests a nonlinear dependency, with a steep gradient for lower e_1/d_0 transitioning to a smoother variation at higher values.

Figure 16d shows that the variable $F_{max}/f_u ndt$ increases with both e_1/d_0 and N_r, indicating a direct relationship between these parameters and the maximum normalized force. The gradient is steeper along the e_1/d_0 axis compared to the N_r axis, suggesting that e_1/d_0 has a more significant influence on $F_{max}/f_u ndt$. Regions with higher $F_{max}/f_u ndt$ values are concentrated towards the top-right corner, emphasizing the combined effect of high e_1/d_0 and N_r. Figure 16e displays the relationship between the parameters e_1/d_0, fu/fy, and the normalized value $F_{max}/f_u ndt$. As e_1/d_0 increases, $F_{max}/f_u ndt$ generally rises, indicating a strong dependency on this ratio. Similarly, higher values of fu/fy correspond to increased $F_{max}/f_u ndt$, showcasing its influence on the performance of the system. Figure 16f shows the parametric relationship between the variables p_1/d_0, e_2/d_0, and $F_{max}/f_u ndt$, represented as contour levels. The behaviour indicates that as e_2/d_0 and p_1/d_0 increase, $F_{max}/f_u ndt$ rises steadily, transitioning from lower values to higher values. The gradient suggests a nonlinear correlation, where $F_{max}/f_u ndt$ is more sensitive to changes in e_2/d_0 at higher values of p_1/d_0.

Figure 16g shows that as e_2/d_0 and p_2/d_0 increase, $F_{max}/f_u ndt$ generally increases, suggesting a direct influence of these parameters on the force ratio. The steep gradients at lower values of e_2/d_0 and p_2/d_0 indicate higher sensitivity of $F_{max}/f_u ndt$ to changes in these regions. Figure 16h illustrates that as e_2/d_0 and N_r increase, the value of $F_{max}/f_u ndt$ rises, transitioning from blue (low values) to red (high values), indicating a strong positive correlation. The plot reveals that $F_{max}/f_u ndt$ grows more rapidly at higher values of both e_2/d_0 and N_r, suggesting nonlinear behaviour in this region. Figure 16i shows that as e_2/d_0 increases, the normalized $F_{max}/f_u ndt$ also increases significantly, indicating a stronger structural response with higher ratios of eccentricity to diameter. Similarly, higher values of fu/fy correlate with increased $F_{max}/f_u ndt$, suggesting that higher ultimate-to-yield strength ratios lead to improved maximum force response.

Figure 16j indicates a sharp transition from higher values to lower values as p_1/d_0 increases beyond a threshold, particularly at lower p_2/d_0 values. The behaviour stabilizes at consistently lower values for larger p_1/d_0 and p_2/d_0, suggesting diminishing sensitivity in these regions. Figure 16k shows that as p_1/d_0 and N_r increase, the value of $F_{max}/f_u ndt$ generally increases, transitioning from blue to red regions. The gradient patterns suggest a nonlinear interaction where both parameters significantly influence the outcome, with higher values concentrated in the upper-right corner. Figure 16l shows that the highest values ($F_{max}/f_u ndt \approx 1.8$) occur at very low values of p_1/d_0 and slightly higher fu/fy values, suggesting sensitivity to these parameters. As p_1/d_0 increases, $F_{max}/f_u ndt$ decreases significantly, indicating reduced structural performance in these regions.

Figure 16m indicates that the feature $F_{max}/f_u ndt$ shows a significant gradient along the p_2/d_0 axis near zero, with values exceeding 2, indicating sensitivity to changes in p_2/d_0 at lower values. Beyond a certain threshold of p_2/d_0, $F_{max}/f_u ndt$ stabilizes around 1 for most of the N_r range, suggesting a reduced influence of p_2/d_0 at higher values. Figure 16n shows that the strong gradient near $p_2/d_0 = 0$ suggests significant sensitivity in this region, with the ratio $F_{max}/f_u ndt$ increasing rapidly. For larger values of p_2/d_0, the ratio stabilizes, and fu/fy appears to have a weaker influence overall, with the region dominated by a uniform lower value. Figure 16o shows that as N_r increases, $F_{max}/f_u ndt$ decreases, indicating an inverse relationship. Similarly, as fu/fy increases, $F_{max}/f_u ndt$ also increases, suggesting a direct relationship between these parameters.

6.5 Design implications

From a practical design standpoint, the interpretable CatBoost model and the associated interpretability methods provide several qualitative recommendations for detailing double-shear, bearing-type bolted connections within the investigated domain. First, the contour plots confirm that increasing the normalized end and edge distances systematically enhances the normalized bearing capacity, with the highest values concentrated in regions where both ratios are simultaneously large. This observation highlights the benefit of providing end and edge distances that exceed minimum code requirements rather than merely satisfying them. Second, the transverse pitch ratio should not be selected too small: the response surfaces show a steep increase in normalized bearing capacity as the transverse pitch ratio increases from very low values, followed by a plateau at moderate spacings. This indicates that excessively tight transverse spacing should be avoided, while very large transverse pitch ratios offer limited additional benefit. Third, the longitudinal pitch ratio exhibits a non-monotonic influence. Relatively small longitudinal pitch ratios, when combined with suitable material properties, yield the highest normalized capacities, whereas excessively large longitudinal pitch ratios lead to a marked reduction in normalized bearing capacity. Finally, the effect of the number of bolt rows is secondary to, and interacts with, the geometric parameters: increasing the number of bolt rows can be beneficial when accompanied by sufficient end and edge distances and appropriate pitch values, but it does not guarantee higher normalized capacity on its own. These insights are intended as qualitative guidance to support detailing decisions and rapid design exploration using the proposed ML model. They remain constrained by the range of geometries and material properties

represented in the experimental dataset and should be applied in conjunction with existing code provisions and checks for other governing failure modes.

7 Conclusion

This study investigated the application of ML techniques to predict the bearing capacity of double shear-bolted connections in structural steel, an area of critical importance for ensuring safety and efficiency in engineering design. Leveraging a dataset of 443 experimental results, 10 ML algorithms, including LR, RR, SVR, KNN, DT, RF, and advanced gradient boosting models, were systematically trained, validated, and assessed. Among these, CatBoost emerged as the most robust performer, achieving high predictive accuracy while maintaining interpretability, a key requirement for practical engineering applications.

Key findings from this research emphasize the pivotal roles of normalized end distance (e_1/d_0) and edge distance (e_2/d_0) as primary factors influencing the normalized bearing capacity ($F_{max}/f_u ndt$). This conclusion was consistently supported across various interpretability frameworks, such as PDP, ALE, and SHAP. These methods provided a transparent and nuanced understanding of the ML models, highlighting their decision-making processes and the complex, nonlinear interactions between critical input variables. The findings also revealed significant contributions from the number of bolt rows (N_r) and pitch distances (p_1/d_0, p_2/d_0), reinforcing their importance in multi-bolt connection design.

The research not only demonstrated the effectiveness of ML in addressing the complexity of bearing capacity predictions but also showcased the value of integrating interpretable models into structural engineering workflows. This approach bridges the gap between black-box ML models and the engineering need for actionable and explainable results by offering clear insights into how key parameters influence predictions. This transparency is vital for fostering trust and adoption of ML-driven design methodologies in structural engineering.

Despite these advances, the study has limitations that present opportunities for future research. It is important to emphasise that, despite the relatively good performance of the CatBoost model, the remaining scatter is non-negligible for safety-critical applications. The present model is therefore not proposed as a direct replacement for codified design equations, but rather as a complementary tool for parametric exploration, preliminary assessment, and insight generation. Any future deployment of such models as part of formal design checks would require a dedicated reliability-based calibration, including the derivation of suitable resistance factors or safety margins that explicitly account for model uncertainty.

In addition, the reliance on experimental datasets, while providing valuable insights, may introduce biases due to variability in test setups and conditions. Expanding the dataset with more diverse scenarios and material properties, including high-strength and composite steels, could enhance the generalizability of the models. Furthermore, future exploration involves the integration of physics-informed constraints or hybrid modeling approaches that combine data-driven learning with mechanics-based principles. However, implementing such models presents notable challenges. Accurately embedding physical laws, such as equilibrium conditions or constitutive relationships, into ML architectures requires careful formulation and often domain-specific customization. Additionally, while this study focused on double shear-bolted connections, future work could explore ML-driven predictions for other connection configurations, such as single shear, staggered, or multi-row bolt arrangements, under varying loading conditions.

Moreover, high-fidelity FEA calibrated against the existing tests could be employed to generate synthetic data and enrich the coverage of the parameter space (e.g., larger bolt patterns, alternative plate thicknesses, or rarely tested combinations of end distance, edge distance, and pitch). Such a hybrid experimental-numerical dataset would enable more robust training, facilitate reliability-based calibration of the surrogate model, and ultimately support safer and more comprehensive design recommendations. Furthermore, exploring dynamic loading and fatigue conditions could extend the applicability of ML frameworks to broader engineering challenges. A natural next step will be to embed the proposed interpretable ML model into BIM- and IFC-based design environments and operational digital twins, to empirically validate the end-to-end workflow from automated geometry extraction to real-time connection assessment in practical projects.

In conclusion, this research underscores the transformative potential of ML in advancing structural design and analysis. ML enables more data-driven, efficient, and reliable engineering practices by enhancing prediction accuracy and interpretability. The integration of these methods into industrial workflows not only optimizes design processes but also promotes sustainability, safety, and innovation in structural engineering. This study serves as a foundation for further exploration into the intersection of ML and engineering, aiming to develop adaptable, intelligent tools for the next-generation of structural systems.

Data availability statement

Publicly available datasets were analyzed in this study. This data can be found here: The dataset was compiled from the published studies cited in the paper: https://doi.org/10.1016/J.JCSR.2013.12.002, https://doi.org/10.1016/J.JCSR.2010.03.009, https://doi.org/10.1016/J.JCSR.2017.06.001, https://doi.org/10.1016/J.JCSR.2019.05.023, https://doi.org/10.1061/(ASCE)0733-9445(2003)129:6(792), https://doi.org/10.1016/S0143-974X(98)00220-X, https://doi.org/10.1007/S13296-013-4005-Y/METRICS, https://doi.org/10.1016/J.ENGSTRUCT.2019.03.060, https://doi.org/10.1016/J.JCSR.2010.10.007, https://doi.org/10.1061, https://doi.org/10.1061, https://research.tudelft.nl/en/publications/experimental-research-on-single-bolt-connections-for-high-strengt, https://doi.org/10.1016/J.ISTRUC.2019.12.002, https://doi.org/10.1016/S0143-974X(00)00017-1, https://doi.org/10.1016/J.JCSR.2018.04.006.

Author contributions

SK: Writing – original draft, Methodology, Software, Conceptualization. HL: Writing – review and editing. TD: Visualization, Writing – review and editing. AM: Writing – review and editing, Visualization. IB: Writing – review and editing.

Funding

The author(s) declared that financial support was not received for this work and/or its publication.

Conflict of interest

The author(s) declared that this work was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declared that generative AI was not used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Ahmed, A., and Teh, L. H. (2019). Thread effects on the stiffness of bolted shear connections. *J. Constr. Steel Res.* 160, 77–88. doi:10.1016/j.jcsr.2019.05.023

Alavi, H., Kookalani, S., Rahimian, F., and Forcada, N. (2024a). Introduction of methodology for BIM and DSS. *Integr. Build. Intell.* doi:10.1007/978-3-031-68865-2_3

Alavi, H., Kookalani, S., Rahimian, F., and Forcada, N. (2024b). BIM-based augmented reality for facility maintenance management. *Integr. Build. Intell.* doi:10.1007/978-3-031-68865-2_7

Alavi, H., Kookalani, S., Rahimian, F., and Forcada, N. (2024c). BIM-based DSS for enhancing occupants' comfort. *Integr. Build. Intell.*, 79–99. doi:10.1007/978-3-031-68865-2_6

Alavi, H., Kookalani, S., Rahimian, F., and Forcada, N. (2024d). BIM-based DSS for building condition assessment. *Integr. Build. Intell.*, 59–78. doi:10.1007/978-3-031-68865-2_5

Alavi, H., Kookalani, S., Rahimian, F., and Forcada, N. (2024e). BIM-based DSS for HVAC root-cause detection. *Integr. Build. Intell.*, 43–57. doi:10.1007/978-3-031-68865-2_4

Alavi, H., Kookalani, S., Rahimian, F., and Forcada, N. (2024f). Integrated building intelligence. doi:10.1007/978-3-031-68865-2

Apley, D. W., and Zhu, J. (2020). Visualizing the effects of predictor variables in black box supervised learning models. *J. R. Stat. Soc. Ser. B Stat. Methodol.* 82 (4), 1059–1086. doi:10.1111/rssb.12377

Ashrafian, A., Panahi, E., Salehi, S., Karoglou, M., and Asteris, P. G. (2023). Mapping the strength of agro-ecological lightweight concrete containing oil palm by-product using artificial intelligence techniques. *Structures* 48, 1209–1229. doi:10.1016/j.istruc.2022.12.108

Asteris, P. G., Tsavdaridis, K. D., Lemonis, M. E., Ferreira, F. P. V., Le, T. T., Gantes, C. J., et al. (2024a). AI-powered GUI for prediction of axial compression capacity in concrete-filled steel tube columns. *Neural Comput. Appl.* doi:10.1007/s00521-024-10405-w

Asteris, P. G., Karoglou, M., Skentou, A. D., Vasconcelos, G., He, M., Bakolas, A., et al. (2024b). Predicting uniaxial compressive strength of rocks using ANN models: incorporating porosity, compressional wave velocity, and schmidt hammer data. *Ultrasonics* 141, 107347. doi:10.1016/j.ultras.2024.107347

Barkhordari, M. S., Armaghani, D. J., and Asteris, P. G. (2023). Structural damage identification using ensemble deep convolutional neural network models. *CMES - Comput. Model. Eng. Sci.* 134 (2), 835–855. doi:10.32604/cmes.2022.020840

Cabrera, M., Ninic, J., and Tizani, W. (2023). Fusion of experimental and synthetic data for reliable prediction of steel connection behaviour using machine learning. *Eng. Comput.* 39 (6), 3993–4011. doi:10.1007/s00366-023-01864-1

Chen, T., and Guestrin, C. (2016). "XGBoost: a scalable tree boosting system," in *Proceedings of the ACM SIGKDD international conference on knowledge discovery and data mining*, 785–794.

Chen, T., He, T., Benesty, M., Khotilovich, V., Tang, Y., Cho, H., et al. (2018). XGBoost: extreme Gradient boosting. *R. Package Version 071-2* 1 (4), 1–4.

Clements, D. D. A., and Teh, L. H. (2013). Active shear planes of bolted connections failing in block shear. *J. Struct. Eng. (United States)* 139 (3), 320–327. doi:10.1061/(asce)st.1943-541x.0000626

Cover, T. M., and Hart, P. E. (1967). Nearest neighbor pattern classification. *IEEE Trans. Inf. Theory* 13 (1), 21–27. doi:10.1109/tit.1967.1053964

Dietterich, T. G. (2000). "Ensemble methods in machine learning," in *Lecture notes in computer science (including subseries lecture notes in artificial intelligence and lecture notes in bioinformatics)*.

Dorogush Veronika, A., Ershov, V., and Gulin, A. (2018). CatBoost: gradient boosting with categorical features support. arXiv preprint arXiv:181011363.

Freitas, S. T. de, Vries, P. de, and Bijlaard, F. (2005). Experimental research on single bolt connections for high strength steel S690. *cmm - Assoc. Port. construcao Met. mista.* Available online at: https://research.tudelft.nl/en/publications/experimental-research-on-single-bolt-connections-for-high-strengt.

Freund, Y., and Schapire, R. E. (1997). A decision-theoretic generalization of On-Line learning and an application to boosting. *J. Comput. Syst. Sci.* 55 (1), 119–139. doi:10.1006/jcss.1997.1504

Friedman, J. H. (2001). Greedy function approximation: a gradient boosting machine. *Ann. Stat.*, 1189–1232. doi:10.1214/aos/1013203451

Geyer, P., Singh, M. M., and Chen, X. (2024). Explainable AI for engineering design: a unified approach of systems engineering and component-based deep learning demonstrated by energy-efficient building design. *Adv. Eng. Inf.* 62, 102843. doi:10.1016/j.aei.2024.102843

Ghanizadeh, A. R., Ghanizadeh, A., Asteris, P. G., Fakharian, P., and Armaghani, D. J. (2023). Developing bearing capacity model for geogrid-reinforced stone columns improved soft clay utilizing MARS-EBS hybrid method. *Transp. Geotech.* 38, 100906. doi:10.1016/j.trgeo.2022.100906

Guo, H., Xie, Y., Liu, Y., and Yang, D. (2020). Study on mechanical behavior of Q690D high strength steel bearing-type bolted connections. *Structures* 23, 588–601. doi:10.1016/j.istruc.2019.12.002

Hai, L. T., Li, G. Q., Wang, Y. B., Sun, F. F., and Jin, H. J. (2019). Experimental investigation on cyclic behavior of Q690D high strength steel H-section beam-columns about strong axis. *Eng. Struct.* 189, 157–173. doi:10.1016/j.engstruct.2019.03.060

Hoerl, A. E., and Kennard, R. W. (1970). Ridge regression: biased estimation for nonorthogonal problems. *Technometrics* 12 (1), 55–67. doi:10.1080/00401706.1970.10488634

Kalab, P. (2023). When a shear connection transmits a bending moment. IDEA StatiCa. Available online at: https://www.ideastatica.com/blog/when-a-shear-connection-transmits-a-bending-moment.

Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., et al. (2017). "LightGBM: a highly efficient gradient boosting decision tree," in *Advances in neural information processing systems*.

Kim, D. K., and Lee, C. H. (2020). Generalized load deformation relationship for bearing-type single-bolted connections. *J. Struct. Eng.* 146 (7), 04020116. doi:10.1061/(asce)st.1943-541x.0002640

Kim, H. J., and Yura, J. A. (1999). The effect of ultimate-to-yield ratio on the bearing strength of bolted connections. *J. Constr. Steel Res.* 49 (3), 255–269. doi:10.1016/s0143-974x(98)00220-x

Kookalani, S., and Cheng, B. (2021). Structural analysis of GFRP elastic gridshell structures by particle swarm optimization and least square support vector machine

algorithms. *J. Civ. Eng. Mater. Appl.* 5 (3), 139–150. doi:10.22034/jcema.2021.304981.1064

Kookalani, S., Cheng, B., and Xiang, S. (2021). Shape optimization of GFRP elastic gridshells by the weighted lagrange ε-twin support vector machine and multi-objective particle swarm optimization algorithm considering structural weight. *Structures* 33, 2066–2084. doi:10.1016/j.istruc.2021.05.077

Kookalani, S., Shen, D., Zhu, L. L., Zhu, L., and Lindsey, M. (2022a). An overview of optimal damper placement methods in structures. *Trans. Civ. Eng.* 46, 1785–1804. doi:10.1007/s40996-021-00752-2

Kookalani, S., Nyunn, S., and Xiang, S. (2022b). Form-finding of lifting self-forming GFRP elastic gridshells based on machine learning interpretability methods. *Struct. Eng. Mech.* 84 (5), 605–618. doi:10.12989/sem.2022.84.5.605

Kookalani, S., Cheng, B., and Torres, J. L. C. (2022c). Structural performance assessment of GFRP elastic gridshells by machine learning interpretability methods. *Front. Struct. Civ. Eng.* 16 (10), 1249–1266. doi:10.1007/s11709-022-0858-5

Kookalani, S., Parn, E., Brilakis, I., Dirar, S., Theofanous, M., Faramarzi, A., et al. (2024). Trajectory of building and structural design automation from generative design towards the integration of deep generative models and optimization: a review. *J. Build. Eng.* 97, 110972. doi:10.1016/j.jobe.2024.110972

Kookalani, S., Alavi, H., and Rahimian, F. P. (2025). *Structural Design and Optimization of Lifting Self-forming GFRP Elastic Gridshells based on Machine Learning*. London, Newyork: Routledge. doi:10.1201/9781003565055

Li, L., Liu, Z., Shen, J., Wang, F., Qi, W., and Jeon, S. (2023). A LightGBM-based strategy to predict tunnel rockmass class from TBM construction data for building control. *Adv. Eng. Inf.* 58, 102130. doi:10.1016/j.aei.2023.102130

Liang, H., and Song, W. (2009). Improved estimation in multiple linear regression models with measurement error and general constraint. *J. Multivar. Anal.* 100 (4), 726–741. doi:10.1016/j.jmva.2008.08.003

Lim, S., and Chi, S. (2019). Xgboost application on bridge management systems for proactive damage estimation. *Adv. Eng. Inf.*, 41. doi:10.1016/j.aei.2019.100922

Liu, B., Vu-Bac, N., Zhuang, X., and Rabczuk, T. (2020). Stochastic multiscale modeling of heat conductivity of polymeric clay nanocomposites. *Mech. Mater.* doi:10.1016/j.mechmat.2019.103280

Liu, H., Su, H., Kookalani, S., Cao, M., Lei, Y., and Wu, W. (2025). Attention-enhanced spatiotemporal deep learning-based automatic warning model with uncertainty estimation in dam safety. *Struct. Health Monit.* doi:10.1177/14759217251358537

Love, P. E. D., Fang, W., Matthews, J., Porter, S., Luo, H., and Ding, L. (2023). Explainable artificial intelligence (XAI): precepts, models, and opportunities for research in construction. *Adv. Eng. Inf.* 57 102024. doi:10.1016/j.aei.2023.102024

Lundberg, S. M., and Lee, S. I. (2017). A unified approach to interpreting model predictions. *Adv. Neural Inf. Process Syst.* 30. doi:10.48550/arXiv.1705.07874

Lyu, Y. F., Wang, Y. B., Li, G. Q., and Jiang, J. (2019). Numerical analysis on the ultimate bearing resistance of single-bolt connection with high strength steels. *J. Constr. Steel Res.* 153, 118–129. doi:10.1016/j.jcsr.2018.10.006

Lyu, Y. F., Li, G. Q., Wang, Y. B., Li, H., and Wang, Y. Z. (2020a). Bearing behavior of multi-bolt high strength steel connections. *Eng. Struct.* 212, 110510. doi:10.1016/j.engstruct.2020.110510

Lyu, Y. F., Li, G. Q., and Wang, Y. B. (2020b). Behavior-based resistance model for bearing-type connection in high-strength steels. *J. Struct. Eng.* 146 (7), 04020109. doi:10.1061/(asce)st.1943-541x.0002639

Ma, S., Flanigan, K. A., and Bergés, M. (2024). State-of-the-art review and synthesis: a requirement-based roadmap for standardized predictive maintenance automation using digital twin technologies. *Adv. Eng. Inf.*, 62. doi:10.1016/j.aei.2024.102800

Mahdavipour, M. A., Faramarzi, A., Dirar, S., Theofanous, M., Jowhari Moghadam, S., Feng, Q., et al. (2026). Optimised steel frame design using reclaimed steel: logistics impact on reuse efficiency. *J. Constr. Steel Res.* 236, 110046. doi:10.1016/j.jcsr.2025.110046

Mahmood, W., Mohammed, A. S., Sihag, P., Asteris, P. G., and Ahmed, H. (2022). Interpreting the experimental results of compressive strength of hand-mixed cement-grouted sands using various mathematical approaches. *Archives Civ. Mech. Eng.* 22 (1), 19. doi:10.1007/s43452-021-00341-0

Moe, P., and Beg, D. (2011). Investigation of high strength steel connections with several bolts in double shear. *J. Constr. Steel Res.* 67 (3), 333–347. doi:10.1016/j.jcsr.2010.10.007

Može, P. (2018). Bearing strength at bolt holes in connections with large end distance and bolt pitch. *J. Constr. Steel Res.* 147, 132–144. doi:10.1016/j.jcsr.2018.04.006

Može, P., and Beg, D. (2010). High strength steel tension splices with one or two bolts. *J. Constr. Steel Res.* 66 (8–9), 1000–1010. doi:10.1016/j.jcsr.2010.03.009

Može, P., and Beg, D. (2014). A complete study of bearing stress in single bolt connections. *J. Constr. Steel Res.* 95, 126–140. doi:10.1016/j.jcsr.2013.12.002

Puthli, R., and Fleischer, O. (2001). Investigations on bolted connections for high strength steel members. *J. Constr. Steel Res.* 57 (3), 313–326. doi:10.1016/s0143-974x(00)00017-1

Rex, C. O., and Easterling, W. S. (2003). Behavior and modeling of a bolt bearing on a single plate. *J. Struct. Eng.* 129 (6), 792–800. doi:10.1061/(asce)0733-9445(2003)129:6(792)

Sadrossadat, E., Basarir, H., Karrech, A., and Elchalakani, M. (2022). Multi-objective mixture design and optimisation of steel fiber reinforced UHPC using machine learning algorithms and metaheuristics. *Eng. Comput.* 38 (3), 2569–2582. doi:10.1007/s00366-021-01403-w

Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., et al. (2008). Global sensitivity analysis: the primer. *Glob. Sensit. Analysis Primer*. doi:10.1002/9780470725184

Sarir, P., Shen, S. L., Wang, Z. F., Chen, J., Horpibulsuk, S., and Pham, B. T. (2021a). Optimum model for bearing capacity of concrete-steel columns with AI technology *via* incorporating the algorithms of IWO and ABC. *Eng. Comput.* 37 (2), 797–807. doi:10.1007/s00366-019-00855-5

Sarir, P., Chen, J., Asteris, P. G., Armaghani, D. J., and Tahir, M. M. (2021b). Developing GEP tree-based, neuro-swarm, and whale optimization models for evaluation of bearing capacity of concrete-filled steel tube columns. *Eng. Comput.* 37 (1), 1–19. doi:10.1007/s00366-019-00808-y

Schapire, R. E. (2013). "Explaining adaboost," in *Empirical inference: festschrift in honor of Vladimir N vapnik* (Berlin, Heidelberg: Springer), 37–52.

Schapire, R. E., and Singer, Y. (1999). Improved boosting algorithms using confidence-rated predictions. *Mach. Learn* 37 (3), 297–336. doi:10.1023/a:1007614523901

Smola, A. J., and Schölkopf, B. (2004). A tutorial on support vector regression. *Statistics Comput.* doi:10.1023/B:STCO.0000035301.49549.88

Teh, L. H., and Clements, D. D. A. (2012). Block shear capacity of bolted connections in cold-reduced steel sheets. *J. Struct. Eng.* 138 (4), 459–467. doi:10.1061/(asce)st.1943-541x.0000478

Teh, L. H., and Uz, M. E. (2015a). Block shear failure planes of bolted connections — direct experimental verifications. *J. Constr. Steel Res.* 111, 70–74. doi:10.1016/j.jcsr.2015.04.006

Teh, L. H., and Uz, M. E. (2015b). Ultimate shear-out capacities of structural-steel bolted connections. *J. Struct. Eng.* 141 (6), 04014152. doi:10.1061/(ASCE)ST.1943-541X.0001105

Teh, L. H., and Uz, M. E. (2016). Combined bearing and shear-out capacity of structural steel bolted connections. *J. Struct. Eng.* 142 (11), 04016098. doi:10.1061/(ASCE)ST.1943-541X.0001573

Torres, J. L. C., Feng, T., Zhang, K., and Kookalani, S. (2025). Ecological engineering contributions to slope stability through root reinforcement in diverse soils. *Appl. Sci.* 15, 11810. doi:10.3390/app152111810

Wang, Y. B., Lyu, Y. F., Li, G. Q., and Liew, J. Y. R. (2017). Behavior of single bolt bearing on high strength steel plate. *J. Constr. Steel Res.* 137, 19–30. doi:10.1016/j.jcsr.2017.06.001

Xiang, S., Cheng, B., and Kookalani, S. (2020a). An analytic solution for form finding of GFRP elastic gridshells during lifting construction. *Compos Struct.* 244, 244. doi:10.1016/j.compstruct.2020.112290

Xiang, S., Cheng, B., Zou, L., and Kookalani, S. (2020b). An integrated approach of form finding and construction simulation for glass fiber-reinforced polymer elastic gridshells. *Struct. Des. Tall Special Build.* 29 (5), e1698. doi:10.1002/tal.1698

Xiang, S., Cheng, B., Kookalani, S., and Zhao, J. (2021). An analytic approach to predict the shape and internal forces of barrel vault elastic gridshells during lifting construction. *Structures* 29, 628–637. doi:10.1016/j.istruc.2020.11.032

Yang, K. C., Hsu, R. J., and Hsu, C. F. (2013). Effect of end distance and bolt number on bearing strength of bolted connections at elevated temperature. *Int. J. Steel Struct.* 13 (4), 635–644. doi:10.1007/s13296-013-4005-y

Zakir, S. S., Sakil Ahmed, K., Imtiaz Khan, N., Ahmed, A., and Nehdi, M. L. (2022). Predicting bearing capacity of double shear bolted connections using machine learning. *Eng. Struct.* 251, 113497. doi:10.1016/j.engstruct.2021.113497

Zhang, C., and Ma, Y. (2012). *Ensemble machine learning: methods and applications. Ensemble machine learning: methods and applications*. Springer Science & Business Media.

Zhang, H., Nguyen, H., Bui, X. N., Pradhan, B., Asteris, P. G., Costache, R., et al. (2022). A generalized artificial intelligence model for estimating the friction angle of clays in evaluating slope stability using a deep neural network and harris hawks optimization algorithm. *Eng. Comput.* 38, 3901–3914. doi:10.1007/s00366-020-01272-9

Zhao, J., Nguyen, H., Nguyen-Thoi, T., Asteris, P. G., and Zhou, J. (2022). Improved levenberg–marquardt backpropagation neural network by particle swarm and whale optimization algorithms to predict the deflection of RC beams. *Eng. Comput.* 38, 3847–3869. doi:10.1007/s00366-020-01267-6

frontiers in Virtual Reality

TYPE Original Research
PUBLISHED 29 May 2026
DOI 10.3389/frvir.2026.1768058

Check for updates

OPEN ACCESS

EDITED BY
Domna Banakou,
Ecole Nationale Superieure d'Arts et Metiers - Campus d'Angers, France

REVIEWED BY
Salman Azhar,
Auburn University, United States
Mohamed Assaf,
University of Alberta, Canada

*CORRESPONDENCE
Prasanna Venkatesan Ramani,
prasanna.venkatesan@vit.ac.in

RECEIVED 15 December 2025
REVISED 07 May 2026
ACCEPTED 08 May 2026
PUBLISHED 29 May 2026

CITATION
Baikati M, Baikati AK, Saravana Kumar MP and Ramani PV (2026) Integrating BIM and VR for enhanced safety communication and training in construction.
Front. Virtual Real. 7:1768058.
doi: 10.3389/frvir.2026.1768058

COPYRIGHT
© 2026 Baikati, Baikati, Saravana Kumar and Ramani. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Integrating BIM and VR for enhanced safety communication and training in construction

Mehamooda Baikati, Arif Khan Baikati, M. P. Saravana Kumar and Prasanna Venkatesan Ramani*

School of Civil Engineering, Vellore Institute of Technology, Vellore, India

Conventional job-site safety training methods often face challenges in effectively communicating hazards, particularly on large and complex construction projects involving multicultural and multilingual workforces. This study explores the application of an integrated Building Information Modelling (BIM)–Virtual Reality (VR)–based safety training approach as an interactive and time-efficient method to support hazard communication and learning. Site-specific BIM models were developed in Autodesk Revit and visualized in an immersive VR environment using Enscape, and the system was implemented in an indoor training setting near the active construction site, with hazards modeled from the actual project environment. Thirty participants, including construction workers, site supervisors, and safety personnel from diverse language backgrounds, took part in short, scenario-based training sessions covering hazards such as work at height, falling objects, scaffold safety, and edge protection. Multilingual support was provided through facilitator-assisted explanations in local languages (Telugu and Hindi), together with visual-first immersive cues. Training outcomes were evaluated using structured hazard-recognition tasks, rating scales, knowledge checks, and qualitative feedback. Results indicated improved hazard recognition performance following VR-based training, with a mean participant-level post-training score of 91.8% (SD = 5.6), compared with the same participants' pre-training baseline scores collected using the same structured hazard-recognition task before the VR sessions, with a mean participant-level score of 68.2% (SD = 8.3). Descriptive analysis also indicated high levels of engagement, clarity, and perceived usefulness. The findings suggest that BIM–VR–based training can support hazard awareness and safety communication in multilingual construction environments; however, as a pilot study without a controlled comparison group, the results should be interpreted as indicative rather than conclusive evidence of training effectiveness or risk reduction.

KEYWORDS
building information modelling (BIM), construction safety training, hazard identification, immersive learning, multilingual workforce, pilot study, safety communication, virtual reality (VR)

1 Introduction

Effective safety management is widely recognized as an important contributor to successful construction project delivery, particularly in today's complex, geographically dispersed projects and dynamic workforce environments. Workplaces in the construction sector often comprise workers from several cultures and languages, thereby posing challenges to effective safety communication (Alsamadani et al., 2013). Conventional

safety programs primarily rely on static documentation, such as handbooks, two-dimensional drawings, and oral briefings, with limited interaction and a high risk of misunderstanding in multicultural and multilingual workplaces (Carter and Smith, 2006; Gao et al., 2019). These limitations have prompted the exploration of alternative image-based safety communication and training methods.

Building Information Modelling (BIM) enables the creation of detailed three-dimensional digital representations of construction projects, supporting the visualization of spatial configurations and potential hazard conditions. Several studies have shown that BIM can assist in proactive safety planning by enabling the visualization of potential hazards and spatial relationships within construction environments prior to on-site execution (Malekitabar et al., 2016; Zhang et al., 2013). When integrated with virtual reality (VR), BIM models can be experienced within an immersive environment that simulates job-site conditions, allowing users to explore construction scenarios and interact with virtual elements in three dimensions. These visual and experiential training approaches can reduce reliance on text-based instruction and, when combined with facilitator-led explanations, may support safety comprehension in multilingual workforces. Recent studies also indicate the growing application of BIM- and VR-based approaches in construction safety training. For example, Getuli et al. (2020) reported improved hazard awareness among participants exposed to BIM–VR simulations, while Zhang et al. (2022) identified perceived ease of use and usefulness as key factors influencing VR adoption. Collectively, these studies suggest a shift in digital safety training research toward more applied implementations and user-centered evaluations.

While there is growing interest in BIM- and VR–based safety applications, a substantial portion of existing research has focused on technological development and controlled or laboratory-based evaluations, often with limited integration of site-specific conditions, structured field-based assessment, and multilingual communication considerations involving active construction workers (Gao et al., 2019). In addition, many studies give limited attention to safety communication issues arising from language diversity and hierarchical relationships among workers, frontline supervisors, and occupational safety personnel, which can affect practical applicability in construction settings. Furthermore, the role of multilingual communication in influencing training effectiveness has been explored in prior research but remains comparatively underexamined in terms of field-based implementation, particularly regarding how immersive and visual environments may assist safety communication across diverse language groups. This indicates the need for further field-oriented investigations into BIM-integrated VR as a safety communication tool in multilingual construction environments. More recently, systematic reviews such as that of Scorgie et al. (2024) report an increasing number of VR-based safety training studies across multiple industries, but caution about the relatively limited validation based on real field data from ongoing construction projects. In the same way, recent research since 2021 increasingly emphasizes the need for real-site implementation studies that use realistic methods to assess communication effectiveness and evaluate extrapolated long-term knowledge retention rather than short-term laboratory performance metrics. These discussions further reinforce the practical relevance of applied BIM–VR investigations conducted in training settings linked to active construction project environments.

In this paper, BIM-embedded VR simulations for a residential construction project with a multilingual workforce are examined to assess their applicability. First, it aims to explore the feasibility and acceptability of immersive, site-specific VR training units for safety communication and hazard awareness across various role groups and languages. An architectural BIM model representing typical site conditions was exported to a VR environment, where common site hazards, such as falling from height, falling objects, scaffolding safety, stair handrails, and emergency evacuation conditions, were modelled. Training modules were intended to be brief, visually appealing, and consistent with average toolbox training times. To address workforce diversity, training delivery incorporated multilingual facilitation alongside immersive visual scenarios, supported by facilitator-led explanations to aid comprehension across different language groups.

Safety employees, site managers, and workers across various trades used HMDs (Head-Mounted Displays) and handheld controllers in VR-based training sessions. The VR exercises were designed to facilitate hazard identification and discussion, enabling participants to identify risks and collaboratively consider preventive solutions. Rather than imposing quantifiable risk reduction, the research presented here addresses training-relevant matters such as hazard perception, safety message clarity, and user engagement. The findings are presented as preliminary evidence from this pilot study to inform further longitudinal and extended field-based research on BIM-integrated VR safety training.

2 Background

2.1 Shortcomings of conventional construction safety management

Traditionally, construction safety management has largely relied on paper-based systems, 2D drawings, and experience-based hazard lists, established in response to safety risks in relatively static project environments. However, increasing site complexity and dynamic workflows mean that such static documentation may not adequately keep pace with evolving task sequences, spatial configurations, and workforce interactions. Carter and Smith (2006) stress that construction safety hazards are dynamic and must be actively verified through job-site oversight rather than based solely on predetermined paper exercises. Current first-generation safety management systems operate primarily in two dimensions through method statements and checklists, which fail to capture rapidly changing site conditions in a timely and worker-intuitive manner.

In terms of training, construction safety education remains largely instructor-led, relying on classroom lectures, workshops, toolbox talks, and slide presentations. Gao et al. (2019) note that, despite their popularity, traditional, non-embodied modes of learning may result in comparatively lower engagement and retention when conveying knowledge about how to stay safe outside their natural context. These limitations are increasingly pronounced on large construction sites, where safety monitoring relies heavily on human observation and control. Safety managers

commonly conduct periodic exercises using checklists. However, given the scale and complexity of construction projects, it can be challenging to consistently perform standardized audits that comprehensively identify potential hazards on a job site.

Previous research suggests that traditional safety procedures often place greater emphasis on visible, easily auditable aspects of safety, such as PPE, signage, and harness use. In contrast, hazards that are less visible but equally, if not more, hazardous, such as temporary edge conditions, formwork instability, or moving materials, may receive less attention (Zhang et al., 2013). Time constraints and infrequent manual inspections can reduce the effectiveness of safety inspections, potentially creating gaps between identified risks and actual on-site conditions. Therefore, traditional safety management methods are often characterized as reactive rather than proactive.

Safety management on construction projects is often reactive, with lessons from incidents or near-misses recorded only after they occur. Martínez-Rojas et al. (2020) point out that construction safety reporting systems are often based on unstructured or generalized data, which may not adequately reflect conditions at a specific site for proactive safety planning. Incident reports are commonly stored in archives to satisfy regulatory obligations; however, they are not always used as dynamic inputs or references for adaptive, continuously updated safety strategies. This gap between evidence-based safety insights and prescriptive planning underscores the need for more proactive, site-responsive safety management systems.

2.2 Communication breakdown in the multilingual workforce

On larger residential and infrastructure projects, contractors increasingly rely on multinational, multilingual workforces. Although managerial staff often use English for communication, many workers primarily communicate in languages such as Hindi, Telugu, or Tamil. This language difference poses risks to safety communication, as key messages on hazard awareness, safe work procedures, and emergency response may be interpreted inconsistently across workers.

Safety performance and incident occurrence can be influenced by language skills. Alsamadani et al. (2013) found that low language comprehension among some American construction workers resulted in miscommunication of safety directions and noncompliance with protective measures. Likewise, Lim and Alum (1995) found that communication barriers were among the significant factors influencing project performance and accident rates on multicultural construction sites. Together, these studies suggest that language barriers are an important factor affecting the effectiveness of safety management systems.

Standard mitigation methods – translated manuals, multilingual signs, or interpreter-facilitated toolbox talks – face recognized constraints. Translations can be resource-intensive, may become outdated, and are not always effectively implemented, particularly for technical terms and regional dialects. Additionally, written communication presupposes literacy, attention, and an understanding of context, which are not equally available across different worker populations. As a result, safety messages may sometimes be communicated but not fully internalized at the task level.

Thus, effective multilingual safety communication must rely not only on verbal instruction but also on non-linguistic presentation methods. These can include demonstration/outreach, scenario/case-based training, and interactive simulations, which help improve understanding and memory of concepts by allowing workers to view hazards and controls firsthand. Malekitabar et al. (2016) argue that the framing of safety knowledge during design, particularly through BIM-assisted processes, can be translated to field personnel in a visual, practice-oriented manner through a structured transmission process from designers and safety managers to on-site workers. Nonetheless, current BIM-based safety research provides limited discussion on how such visual materials are integrated and used for multilingual workers in their training and daily practice.

Recent studies have increasingly explored the integration of BIM and virtual reality for safety communication in construction environments. Afzal and Shafiq (2021) demonstrate that BIM-integrated VR environments can enhance safety communication and improve accessibility of training content for multilingual construction workers by enabling immersive visualization of hazards and safety procedures. However, while their study highlights the communication advantages of VR-based approaches, it provides limited focus on structured hazard identification frameworks and does not address systematic evaluation of learning performance across multiple time points.

This limitation is further supported by a recent systematic review by Wang et al. (2025), which reports that VR applications in construction safety management are expanding in areas such as training, hazard recognition, and risk communication. However, the review also identifies persistent challenges related to the lack of standardized assessment methodologies, limited integration of VR within formal safety management systems, and insufficient evidence on knowledge retention and longitudinal performance evaluation across diverse workforce groups. Collectively, these studies indicate that although VR-enhanced BIM approaches show strong potential for improving safety communication, there remains a clear need for structured, repeatable assessment frameworks that evaluate hazard recognition performance over time in multilingual construction settings.

In this context, the present study addresses these gaps by developing and implementing a structured BIM–VR-based safety training framework that incorporates systematic hazard identification tasks, standardized scoring criteria, and multi-timepoint evaluation (baseline, immediate post-training, and follow-up assessment). This enables a more consistent and measurable approach to evaluating hazard recognition performance within a multilingual construction workforce, while also aligning immersive training with structured safety assessment methodologies.

This gap underscores the need for integrated solutions that combine BIM with immersive visualization to support safety communication across language barriers. The immersive component enables workers to encounter hazards, observe potential consequences, and practice appropriate responses in a virtual environment where visual and spatial cues are prioritized over verbal instruction. Accordingly, this study aims to address this gap by developing and applying VR-based safety modules within a BIM environment to support multilingual construction workers

through visual-first safety communication and shared situational awareness.

2.3 Technological advancement in digital design for safety planning and training

Recent developments in the digital design and construction sector have supported more preventive, visually based safety planning approaches beyond traditional paper-based documentation. Virtual Design and Construction (VDC) tools, such as Building Information Modelling (BIM), Virtual Reality (VR), Augmented Reality (AR), and simulation-based systems, enable the visualization of construction processes and risks before they are carried out on-site. In this way, they support early-stage hazard identification by integrating safety information with spatial construction data, rather than relying solely on experience-based checklists.

The penetration of ICT (Information and Communication Technologies) into safety management has contributed to a shift from predominantly reactive, monitoring-based approaches toward those supported by design, planning tools, and training. Rajendran and Clarke-Well (2011) have shown that digitally supported safety plans can help clarify safety planning procedures, improve stakeholder involvement in this process, and provide a more organized way to deliver training. There is increasing evidence that the use of digital and interactive visualization can enhance worker engagement and comprehension when working with hazard scenarios, especially in work involving temporary works, work at height, and dynamic site conditions (Zhang et al., 2013).

However, while the proliferation of digital tools is observed, especially in construction safety training, their use remains relatively limited in practice, with issues including implementation costs, interoperability across software platforms, and user acceptance. Moreover, many currently available BIM- and VR-based safety studies are based on conceptual or laboratory work, with comparatively fewer demonstrations on operational construction sites involving workers from diverse nationalities and languages. This gap underscores the need for applied, project-linked studies that examine how digital design tools can facilitate safety communication and training using real project conditions, which is the focus of this study.

2.4 Building information modelling applications

Building Information Modelling (BIM) extends beyond traditional 3D modelling by enabling the structured representation of building components and their spatial relationships within a digital environment. By organizing model elements within a coordinated digital framework, BIM allows safety planners to visualize construction layouts and examine potential temporary site conditions that could develop into hazards. Such visualisation provides a better understanding of space relations and task interactions than static method statements or conventional 2D drawings, thereby enabling more informed safety-related decision-making (Azhar, 2017).

The capacity to associate model units, such as structural elements, scaffolding, access paths, and temporary works, within a coordinated digital environment enables improved visualization of site organization, thereby allowing safety managers to better anticipate potential hazards prior to physical execution. This capability is especially pertinent for identifying risks in work-at-height and material handling activities, where information derived from paper documentation can be difficult to interpret. When combined with visualization tools such as VR, BIM-based models also improve transparency in safety planning and communication on building sites.

Choe and Leite (2017) showed that when safety information is integrated with BIM-based representations of spatial conditions, it can facilitate proactive safety planning for early identification of unsafe conditions. In addition to hazard identification and training, BIM can support improved coordination and monitoring in construction environments. When project conditions change due to factors such as weather delays, scheduling conflicts, or design modifications, BIM models can be updated to reflect revised site layouts, allowing potential safety implications to be examined more systematically. This flexibility makes BIM a useful framework for supporting adaptive, site-specific safety planning and training workflows.

2.5 Immersive VR in construction safety training

Virtual Reality (VR) has been identified as a promising alternative training medium for construction safety because it can immerse users in interactive, three-dimensional (3D) representations of job-site environments, allowing them to experience hazards without physical exposure. VR-based safety training enables users to view and interact with hazardous situations, supporting experiential learning rather than relying solely on reading or video-based instruction. Goulding et al. (2012) demonstrated the feasibility of using interactive VR environments to simulate construction hazards and illustrate preventive measures through direct user experience in controlled virtual scenarios.

Previous research has discussed the potential of VR for procedural learning, such as learning safety-oriented behaviours through repeated action and experience rather than rote memorization (Zhao and Lucas, 2015). For instance, VR has been applied in electrical safety training by providing a high level of presence in situations that are difficult or unsafe to replicate at operating sites. In such scenarios, workers can walk around and inspect virtual scaffolding systems, steel structures, or equipment layouts while receiving real-time safety guidance within the virtual environment.

Immersive VR has also been used in collaborative safety-planning contexts. Sacks et al. (2015) demonstrated the potential application of Cave Automatic Virtual Environment (CAVE) systems to support communication between design and construction teams in identifying and preventing safety hazards during the design and planning phases. Similarly, Pedro et al. (2016) reported increased motivation and improved hazard recognition in VR-enhanced learning settings by comparing conventional safety training with a Virtual Safety Education System (VSES) as an alternative form of safety instruction. However, despite these benefits, many VR safety studies remain confined to academic

environments or controlled demonstrations, with comparatively fewer studies conducted in project-linked training settings involving mixed, multilingual workforces. This limitation highlights the need for on-site applied research to explore how immersive VR can enable safety communication and training in actual project settings, which is addressed in the present work.

Recent studies published after 2020 further reinforce the growing role of immersive VR in construction safety training. For example, Shringi et al. (2022) report that immersive VR environments can enhance hazard recognition performance in construction-related tasks compared with traditional approaches. In addition, Xiong et al. (2026) demonstrate that VR-based safety training supports improved transfer of hazard recognition skills in controlled experimental settings. These findings collectively indicate the continued relevance and advancement of VR-based safety training approaches within the construction sector. In this context, the present study extends existing work by implementing a structured BIM–VR training framework that integrates standardized hazard identification tasks with multi-timepoint assessment (baseline, immediate post-training, and follow-up using the same scoring framework), specifically targeting multilingual construction workers to evaluate hazard recognition performance within a consistent evaluation structure.

2.6 Contemporary research on BIM-VR integration

Combining Building Information Modelling (BIM) and Virtual Reality (VR) to enhance visualization, collaboration, and decision-making in construction projects has been gaining increased attention in recent years. Johansson and Roupé (2024) showed that VR interfaces streaming full-scale BIM models can support shared understanding among construction workers by allowing users to view and interact with spatial information in a common virtual environment. Such shared visual contexts have been associated with reduced misinterpretations and coordination challenges, particularly when complex construction activities or spatial constraints are involved.

Several works have explored BIM–VR solutions for safety, focusing on enhancing the understanding of hazards and the spatial environment. Getuli et al. (2020) employed mixed-methods research based on user feedback and interaction analytics. They reported that the integration of VR into workspace planning was associated with increased attention among participants, supporting safety awareness and improved recognition of hazard zones. Similarly, Pedro et al. (2016) found greater engagement and hazard perception in VR-based safety education compared with traditional training approaches. However, many such studies have been conducted in controlled or academic settings, with applications often involving students rather than on-site construction workers, which may limit their direct applicability to actual project environments.

Moreover, systematic reviews highlight methodological shortcomings in current VR safety studies. Scorgie et al. (2024), who reviewed 52 VR-based safety training studies across various industries, found that fewer than 10% explicitly claimed a basis in learning or behavioural theories, and only 36% had evidence of knowledge retention over time. This suggests that although short-term learning gains are often reported, the evaluation of learning outcomes and retention over time remains comparatively limited, particularly in construction-specific contexts. Adoption-based studies, such as those by Zhang et al. (2022), also indicate that perceived usefulness, ease of use, and user confidence are important factors influencing VR acceptance, accounting for a substantial proportion of the variance in behavioural intention and highlighting the importance of usability and contextual relevance.

However, persistent obstacles to BIM–VR implementation remain, despite significant technological development. Alizadehsalehi et al. (2020) found that more than 65% of construction companies rely on external collaboration for XR development due to limited in-house experience. Additionally, interoperability and model-optimization challenges continue to affect performance and usability (Johansson et al., 2015). These limitations are significant for applications of safety training that require real-time interaction, stable performance, and effective visual communication. In addition, comparative analyses of communication effectiveness suggest that while immersive VR supports distributed collaboration, decision-making accuracy in face-to-face interaction may be influenced by limitations in conveying nonverbal cues (Abbas Ali et al., 2019).

Overall, the current literature suggests the potential benefits of BIM–VR integration in safety planning and training, while also highlighting several unresolved issues: comparatively limited field-based empirical evidence on operational construction projects; limited consideration of multilingual workforces; inconsistent outcome measurements; and relatively few studies examining safety communication effectiveness beyond technological performance.

2.7 Research gap and study objectives

While interest in integrating BIM and VR for construction safety is growing, and several applications have been developed, a number of issues remain insufficiently addressed in the current literature. A considerable portion of existing work emphasizes prototyping and controlled evaluations, with comparatively limited focus on integrating project-specific conditions, structured field-based assessment, and communication challenges that arise during implementation in training contexts linked to active construction projects. Moreover, while existing research has shown the advantages of enhanced hazard visualization and increased user engagement (e.g., Getuli et al., 2020; Pedro et al., 2016), comparatively limited studies directly investigate safety communication effectiveness among multilingual workforces or examine how immersive and visual training environments may assist hazard communication and hazard recognition across diverse language groups. For example, systematic reviews like Scorgie et al. (2024) also point out that outcome measurement is inconsistent, that reporting on participant characteristics is limited, and that there is often little information on procedures for assessing retention. These methodological limitations may affect comparability, reproducibility, and transferability to the operational context of construction projects. Moreover, existing studies commonly emphasize technological capability or user engagement over in-depth accounts of experimental protocols and outcome measurement procedures, thereby restricting

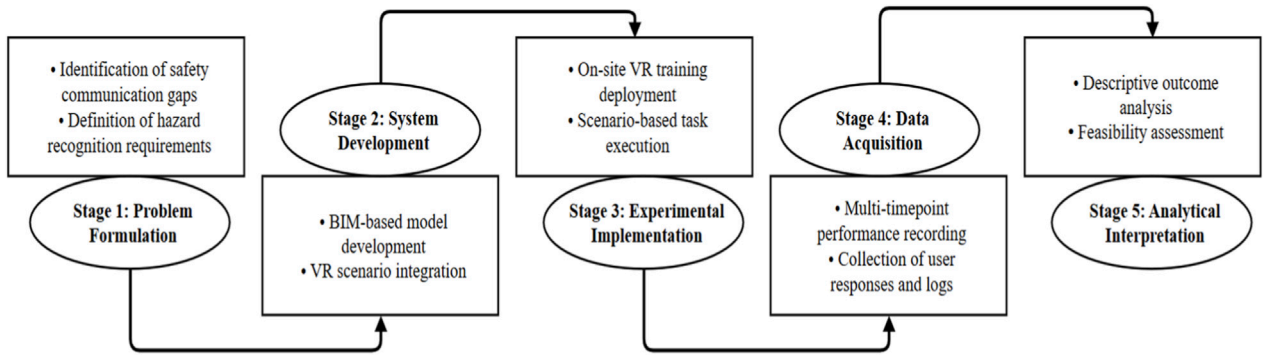


FIGURE 1
Stepwise framework of the BIM–VR safety training approach.

reproducibility and real-world relevance. In response to these identified gaps, the current study examines a BIM-integrated VR safety training process implemented for a residential construction project, focusing on hazard recognition, clarity of safety communication, and user engagement within a multilingual workforce using a pilot quasi-experimental mixed-methods approach. The study also documents participant characteristics, hazard identification tasks, scoring procedures, and retention assessment methods to improve methodological transparency and support reproducibility in future applications.

3 Materials and methods

Figure 1 presents the methodological workflow followed in this pilot study, structured as a sequence of five stages: problem formulation, system development, experimental implementation, data acquisition, and analytical interpretation.

The process begins with the identification of safety communication gaps and definition of hazard recognition requirements based on literature review and site-informed observations. These inputs guide the development of BIM-based models and VR training scenarios that represent selected construction activities and hazard conditions.

The developed scenarios are then implemented through indoor VR training sessions conducted near the project site, where participants perform scenario-based hazard identification tasks. During these sessions, participant interactions, responses, and task performance are recorded.

Data acquisition includes multi-timepoint performance recording, structured participant responses, and system-generated interaction logs. These data sources are used to capture hazard identification actions, task completion patterns, and participant feedback.

The final stage involves descriptive analysis of the collected data to examine training outcomes and feasibility within the study context. The framework is presented as a linear sequence to clearly document the methodological steps followed in this pilot study, without representing an iterative or continuous deployment process.

3.1 Pre-study assessment of safety training and communication challenges

Figure 2 presents the structure of the pre-study assessment conducted prior to the development of BIM–VR training scenarios. The assessment is organized into three sequential phases: assessment, development, and deployment, with a focus on examining existing safety training practices and communication conditions at the project site.

The first phase involves a pre-study safety assessment and communication analysis. This phase focuses on identifying training-related and communication-related gaps through site observations and structured assessment tools. The evaluation includes review of safety procedures, observation of training practices, and examination of how safety information is communicated among site personnel. Safety experts and site staff are involved in this process, and activities include site walkthroughs and focused discussions to document current practices and potential limitations.

The second phase involves BIM–VR scenario development. Based on the identified gaps, hazards are selected and translated into scenario-based training content. BIM models are developed using Autodesk Revit and integrated with VR simulations using Enscape to represent site conditions and hazard situations. This phase involves BIM engineers and safety experts, with activities focused on scenario modelling and system setup to prepare training content aligned with observed site conditions.

The third phase involves VR-based training implementation and pilot testing. The developed scenarios are deployed in VR-based training sessions, where workers, supervisors, and safety personnel participate in hazard identification tasks. This phase uses immersive visualization tools to present hazards and safety instructions. Training activities include structured sessions in which participants interact with the VR environment and perform hazard recognition tasks.

The framework shown in Figure 2 outlines the flow from initial assessment to scenario development and training implementation. It summarizes the roles, technologies, and activities involved at each stage, providing a structured representation of how site observations and communication analysis are translated into BIM–VR training scenarios.

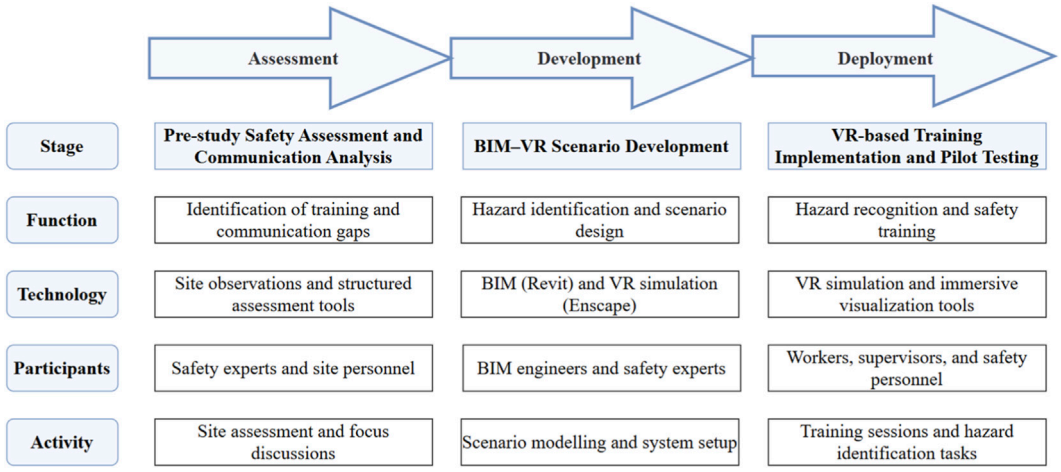


FIGURE 2
Pre-study assessment framework for safety training practices and communication analysis.

3.2 Study design

The present work employed a pilot, exploratory mixed-methods design to assess the feasibility and perceived usefulness of BIM-integrated VR safety training implemented in an indoor training setting linked to an active construction project. The design combined quantitative and qualitative approaches to descriptively document participants' engagement, experiences, and feedback during VR training. Participants completed the VR-based safety training as a single-group intervention. Hazard-recognition performance was assessed within the same participants using a pre-training baseline assessment, an immediate post-training assessment, and a 4-week follow-up assessment using the same structured task and scoring criteria. The same VR-based hazard-recognition task, scenarios, and scoring criteria were consistently applied across all three assessment time points. No separate traditional-training group was used, and references to conventional training are provided only for contextual discussion rather than as a formal comparator group.

The assessment descriptively documented participants' hazard recognition, clarity of safety communication, and engagement following the VR sessions Observations were recorded using structured rating scales and task-based hazard-recognition exercises (scoring rubrics provided in Supplementary Appendix C and observer checklists provided in Supplementary Appendix D), along with qualitative feedback collected immediately after training, along with qualitative feedback collected immediately after training. Given the exploratory nature of the study in an active construction environment, randomization and control groups were not used; this design aimed to capture real-world training experiences and inform potential future controlled studies.

3.3 Participants and recruitment

A total of 30 participants employed on the ongoing residential building project volunteered to take part in the study. Participants were recruited in coordination with site management to include a range of roles involved in safety-critical activities, without formal randomization or stratification. The final participant group included construction workers (N = 20), site supervisors (N = 6), and safety officers (N = 4), reflecting voluntary participation across these roles.

The participants consisted of 26 men and 4 women, ages 22–54 years (mean ≈34 years). The workforce was multilingual, with Hindi, Telugu, Tamil, Bengali, and English as the predominant spoken languages. Although English was the primary language for official safety documentation, many workers received instructions in their native languages with facilitator assistance and translated forms, as the VR system itself was primarily in English. The majority of participants (N = 23) had not used VR as a training method before, and the remaining participants (N = 7) had limited prior exposure to VR, mostly brief introductory experiences with the technology. Written informed consent was obtained from all subjects.

To support participants from diverse linguistic backgrounds, the feedback questionnaire was made available in multiple languages, including English, Telugu, and Hindi. Participants were allowed to respond in their preferred language, and facilitators provided clarification where necessary to ensure consistent interpretation of questionnaire items. The complete questionnaire and its translated versions are provided in Supplementary Appendix A.

3.4 Overall research workflow

This project followed a phased, project-linked process to structure problem diagnosis, system development, participant training, and outcome observation. Pre-study site assessment: The workflow was introduced based on pre-study site visits to explore potential challenges in existing safety training and communication within a multilingual workforce. Insights from this phase informed the development of training goals and guided the selection of site- or activity-specific hazard scenarios for inclusion in the VR environment.

Firstly, BIM-based site models and immersive VR training environments were created to represent typical construction scenarios, with structural hazards identified for training purposes. Secondly, VR safety training sessions conducted in an indoor setting

near the project site were iteratively developed with input from workers, supervisors, and safety personnel. Finally, training outcomes were descriptively summarized using quantitative observations and qualitative feedback collected during and after the sessions.

We followed this phased approach to maintain a clear structure between technical development and outcome observation. The format supports repeatability by documenting what was designed, how it was implemented, and how outcomes were observed, which is feasible within a pilot study under real-world construction constraints.

3.5 Outcome measures and data sources

As this was a pilot and exploratory study, outcome measures were selected to descriptively assess feasibility, usability, and participants' perceived training experience, rather than causal efficacy or long-term risk reduction. Three descriptive sources of data were collected: structured participant questionnaires (see Supplementary Appendix A); task-based hazard-recognition exercises administered at pre-training baseline, immediately after VR training, and at the 4-week follow-up (see Supplementary Appendix C); and feedback, including open-ended discussions, from workers, supervisors, and safety personnel.

Quantitative outcome measures included participant-rated engagement, clarity of safety communication, ease of following the training content, and self-reported perceived usefulness of the VR training, captured using Likert-scale items. To explore hazard perception, participants completed a structured hazard-recognition task within the VR scenarios (see Supplementary Appendix C, featuring six predefined site-specific hazards (e.g., unguarded edges, scaffold instability, missing handrails, unsafe material handling, falling-object exposure, and unsafe stair access). Responses were recorded as correct when participants identified the hazard and indicated an appropriate safe response, based on the pilot scoring rubric provided. The same hazard-recognition task content and scoring criteria were applied at baseline, immediately after training, and during the 4-week retention assessment using the VR environment to maintain internal consistency within this pilot study.

To improve consistency in observational scoring, all observers were briefed using the same checklist definitions and assessment criteria before data collection. A subset of participant assessments was independently reviewed by two trained observers, and any scoring discrepancies were discussed and resolved prior to final data entry. This process served as a practical inter-rater consistency check within the pilot study.

Each participant completed a single VR session in which all six predefined hazard scenarios were embedded within one integrated BIM-based virtual construction environment (see Supplementary Appendix B for detailed workflow). Participants were allowed to freely navigate the environment using joystick-based locomotion in a standing position, without a fixed sequence, time restriction, or predefined path, and were required to independently locate hazards distributed across the virtual site. Prior to the assessment, a brief familiarization phase (approximately 1–2 min) was provided to introduce basic navigation and interaction controls. During the task, participants identified hazards by visually locating them in the environment, pointing toward them within the VR scene, and verbally reporting the hazard to the observer. Following identification, participants verbally explained the appropriate safety response in an open-ended format without predefined answer options. The average duration of each VR session was approximately 20 min, although completion time varied depending on individual navigation and familiarity with the environment.

Hazard-identification responses were recorded using a combination of observer-based documentation and system-recorded interaction timing (limited to session duration and approximate hazard identification time), including total session duration and approximate time taken to identify hazards. All hazard identification outcomes, correctness of responses, and participant behaviour (e.g., confusion or assistance required) were manually documented using the structured observer checklist (see Supplementary Appendix D). Scoring was not performed in real time within the VR system; instead, all responses were evaluated after session completion using the standardized scoring rubric provided in Supplementary Appendix C. No automated feedback on correctness was provided during the task to avoid influencing participant responses; however, general experiential feedback was collected after session completion through the questionnaire. The same VR environment, hazard scenarios, facilitator involvement, and assessment procedure were consistently applied across pre-training, post-training, and 4-week follow-up stages. As the same hazards were reused, potential practice effects are acknowledged as a limitation, partially mitigated by the 4-week interval between assessments.

The collected data were summarized using descriptive statistics for quantitative measures and thematic clustering for qualitative feedback, providing an overview of pilot observations. Results are presented as preliminary, exploratory findings intended to inform the design of future controlled studies, rather than to claim demonstrated efficacy of the training.

The instruments used for data collection are provided in the Supplementary Materials and are referenced at relevant points within the manuscript. The post-training participant feedback questionnaire (see Supplementary Appendix A) includes multilingual versions (English, Telugu, and Hindi). The hazard-recognition assessment instrument and scoring rubric used for performance evaluation are provided in Supplementary Appendix C, and the observer checklist used during training sessions is presented in Supplementary Appendix D.

The detailed workflow of the VR-based hazard-identification task, including scenario structure, participant navigation, hazard reporting procedure, safe-response elicitation, and scoring implementation, is provided in Supplementary Appendix B. While a summary of the procedure is described in the main text, Supplementary Appendix B presents a structured and fully replicable account of participant interaction within the VR environment, including locomotion method, session timing, observer roles, and data recording procedures. The hazard scenarios and scoring framework referenced in this workflow correspond to the assessment instrument presented in Supplementary Appendix C, while observer-based monitoring procedures align with the checklist provided in Supplementary Appendix D.

4 Development of VR-based BIM simulations

This stage produced VR-BIM simulations developed for pilot safety training on a residential construction site. The simulations were based on a BIM model of the site, which primarily provided static representations of key areas and hazards relevant to training scenarios. Development was incremental and iterative, with site visits used to refine identified hazards and guide BIM visualization to create representative scenarios while considering practical usability constraints.

The key purpose of this phase was to depict representative site practices and frequent safety hazards in immersive VR training scenarios, with multilingual comprehension supported through facilitator assistance and translated forms. Instead of overly complex or experimental systems, the focus was on clarity, realism, and ease of use to make the simulations feasible and understandable for pilot training purposes.

4.1 Data collection and analysis at site

Data were collected during multiple site visits to gather observational insights on routine work and safety practices. During these visits, ongoing activities and observed unsafe conditions were documented, sometimes with photographic evidence, to inform VR scenario development. Attention was focused on representative areas involving scaffolds, temporary structures, open edges, and material-handling zones.

Data were also gathered from informal conversations with site managers and safety staff, focusing on frequently mentioned issues during daily safety briefings. These discussions helped identify areas for potential improvement in training tools, particularly regarding workers' attention, comprehension of on-site hazards, and cross-language communication.

The collected site data were processed to develop simplified BIM models and VR environments for pilot training scenarios. The aim was not to replicate the site with full theoretical accuracy, but to create representative space arrangements, job sequences, and hazardous areas that were feasible and illustrative for safety learning.

4.2 Hazard analysis and identification of training focus areas

A pilot hazard review was conducted, including a site visit and discussions with safety personnel, to identify key safety issues for inclusion in VR training scenarios. Risks were categorized based on approximate occurrence, severity, and relevance to daily tasks, using an ordinal and descriptive approach rather than a detailed quantitative risk model.

The investigation identified several frequently observed hazard types: falls from height, falling objects, unsafe access to scaffolding, open window or slab edges, unprotected stair edges, and emergency response incidents. These hazards were selected for inclusion in the pilot VR training scenarios because they were observed across multiple trades on site.

A simplified, descriptive risk-ranking approach was applied to guide selection of scenarios for the pilot training design, without making claims about direct risk-reduction effects. Safety standards from OSHA and relevant local construction regulations were referenced to guide scenario realism, ensuring the recreated scenes reflected typical site practices rather than demonstrating regulatory compliance outcomes.

Notably, the hazard analysis considered language and communication factors, with multilingual comprehension supported through facilitator assistance and translated forms, rather than fully language-independent delivery. Training scenarios were designed to visually depict unsafe conditions and proper practices, aiming to reduce reliance on written instruction while still requiring facilitator guidance and translated support for comprehension across a multilingual workforce.

4.3 3D BIM model development

A three-dimensional BIM model of the site was generated in Autodesk Revit to develop pilot VR training scenarios. The model focused on architectural configuration and primary structural components, using simplified representations to illustrate spatial relations, work areas, and hazardous locations relevant to the pilot training.

As shown in Figure 3, the BIM model includes the general site structure, partially built structures, open levels, vertical circulation areas, and the surrounding environment, which were potentially relevant to observed hazards. The Level of Development (LOD) used for the BIM model was basic, suitable for pilot VR training and spatial orientation. Building massing, floor levels, staircases, slab edges, scaffolding zones, and surrounding site context were modelled to illustrate representative scenarios from site observations. Mechanical, electrical, and plumbing components were not modelled in detail and were included only as needed for illustrative hazards.

As shown in Figure 4, the BIM model represents the building structure along with key safety features such as scaffolding and handrails. The model includes floor levels, columns, staircases, and slab edges, highlighting areas associated with work at height and edge protection. A basic level of development was used, focusing on spatial layout and hazard-prone zones relevant to training. The model provides a static visualization of typical site conditions and was used to support the development of VR-based safety training scenarios.

To better reflect observed site conditions, site observations and photographs were used to verify key work areas and risk-prone zones, with informal feedback from site staff. The BIM model did not include sequence- or time-based simulations; it provided static visualizations consistent with a generic construction sequence for pilot training purposes.

4.4 VR environment creation

The completed BIM model was converted into an immersive virtual environment using Enscape, a real-time visualization extension for the BIM platform. This method provided a fast way to construct the virtual site while maintaining the general geometric and spatial conditions defined in the BIM model, as shown in Figure 5.

The VR environment was designed to represent a typical operational construction site, addressing spatial arrangement,

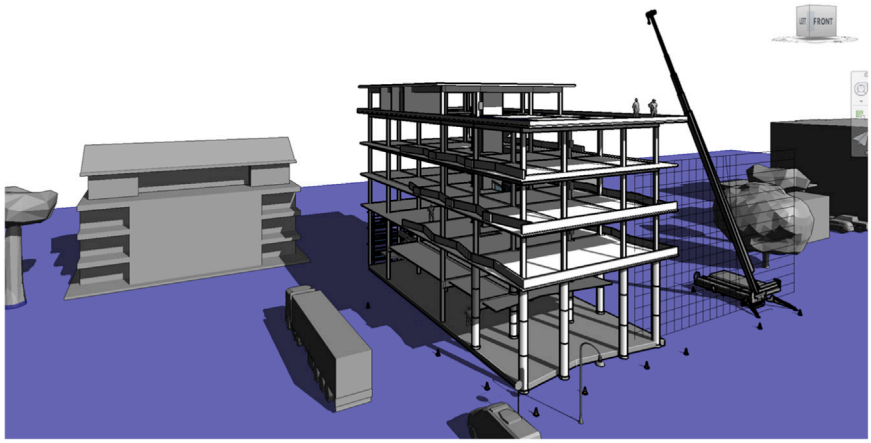


FIGURE 3
Autodesk Revit screenshot of the BIM-based 3D construction site model used for developing VR safety training scenarios.

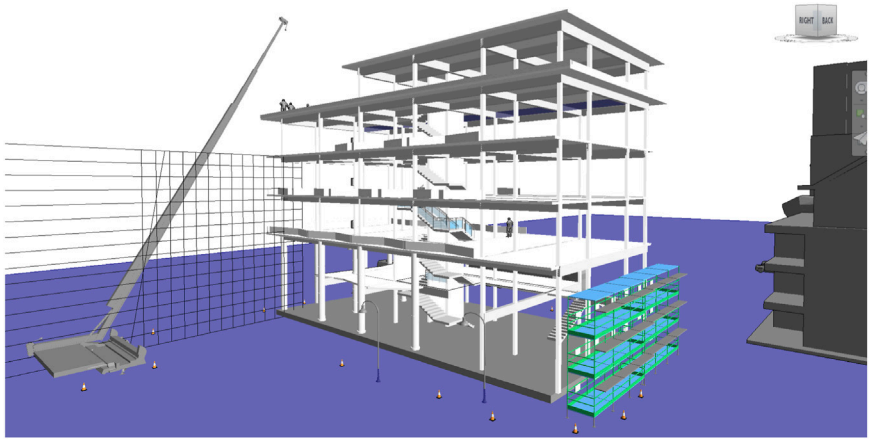


FIGURE 4
Autodesk Revit screenshot of the BIM-based construction model illustrating scaffolding and handrail features used for VR scenario development.

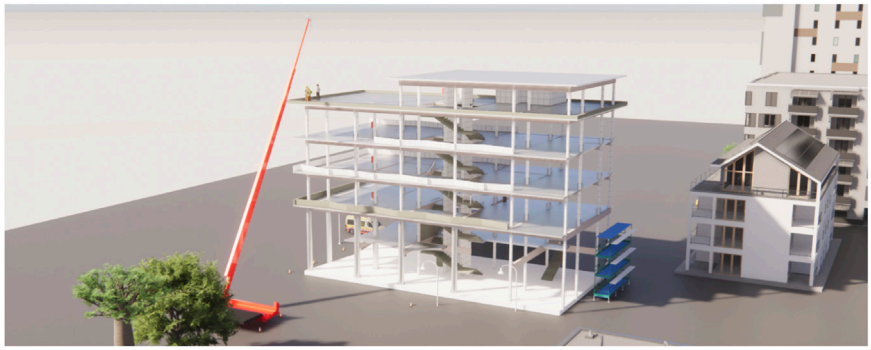


FIGURE 5
Screenshot of the VR-based construction site environment rendered from the BIM model using Enscape (in-application view).

work areas, temporary facilities, and equipment locations for safety training. Visual realism was lower than that of high-fidelity simulation environments (e.g., simplified depth cues, clearly visible open edges, and proximate equipment), but was sufficient for hazard perception.

We optimized performance by balancing model complexity and Enscape rendering settings to ensure stable visualization and comfortable navigation during training sessions. This ensured a consistent training experience across the hardware configuration used during the training sessions without compromising usability.

The interaction in the VR application was restricted to essential operations, such as moving the user, manipulating the viewpoint, and visually identifying hazardous conditions. The training scenarios were designed to allow participants to explore the site at their own pace, supporting experiential learning and hazard recognition while minimizing cognitive overload.

In general, the Enscape-based 3D VR environment provided a consistent interface for visual safety training grounded in established safety training principles, supporting the transition from static instruction toward more context-based job-site visualization, which was particularly relevant for workers with diverse language proficiency and educational backgrounds, with facilitation playing a key role in comprehension.

Although the VR scenarios were developed from observations on an active construction site, the training sessions were conducted in an indoor setting to ensure safety, equipment stability, and consistent delivery.



FIGURE 6
In-application VR interface showing participant view and controller instructions during training.

5 Implementation and testing

The test phase aimed to implement the developed VR-based BIM training system in an indoor training setting near the active construction site and to examine its usability for construction safety education. Rather than quantifying real-world risk reduction, the objective was to explore training feasibility, user interaction, and hazard-recognition performance in a simulation setting. The evaluation focused on participants' ability to identify predefined hazards embedded within the VR scenarios.

Data collection during implementation combined structured observer-based recording with manual documentation. During training, user navigation in the VR environment, task performance, hazard identification responses, and completion time were documented using structured observation checklists (see Supplementary Appendix D) and facilitator records. Ease of use, clarity, and participant engagement were documented using predefined observation templates, and user feedback was collected. This approach was aligned with the exploratory and applied nature of the investigation. Observations were supplemented by structured hazard identification responses and brief knowledge checks to support consistent assessment of training outcomes.

5.1 Training session administration

The VR safety training sessions were implemented in an indoor training setting near the project site using scenario-based modules generated from the BIM model of an actual construction project site, as shown in Figure 6. An initial introduction to VR technology, head navigation, and basic interaction features, such as object selection and interface interaction, was provided at each session.

The immersive training environment was delivered using an Oculus Rift S–type head-mounted display (HMD) with handheld motion controllers. The system provided stereoscopic visualisation, positional tracking, and six-degrees-of-freedom (6DoF) interaction, enabling participants to navigate the BIM-based virtual environment and perform controller-based hazard reporting tasks in real time. The configuration was selected to ensure stable performance and responsive interaction throughout the training sessions. Detailed technical specifications of the VR system used in this study are provided in Supplementary Appendix E.

During the orientation, participants were guided through simulated examples of site conditions commonly associated with high-risk work at height, open edges, scaffolding areas, and areas where a falling object could pose a danger. Through the training scenarios, participants were encouraged to inspect the site and visually identify unsafe conditions. Each scenario contained predefined hazards, and participants were asked to recognize and report them based on their observations rather than being explicitly directed toward specific hazards.

The VR interface in Figure 6 provided simplified interaction through visual cues and low user control complexity, allowing users to focus more on hazard recognition than on system operation. Small groups were used so participants could interact with both researchers, who facilitated and guided the sessions as part of the training process. This structured approach supported consistency across sessions and enabled observation of participants' responses to simulated safety conditions.

As shown in Figure 6, the VR simulation interface presents the user view during training, with safety instructions and interaction cues in English due to software constraints. To address the multilingual aspect of the training, language support was



FIGURE 7
Participants engaged in VR-based safety training sessions.

incorporated through facilitator-assisted translation during VR sessions. A trained facilitator provided real-time explanations and guidance in local languages, including Hindi and Telugu, to support participant comprehension throughout the immersive experience. Prior to the VR session, participants were introduced to the interface and overall training procedure in both English and their respective local languages. During the session, key safety instructions and hazard explanations were verbally communicated in local languages alongside the immersive visuals, supporting participant understanding among workers with diverse linguistic backgrounds. Additionally, all feedback instruments, including questionnaires and evaluation forms, were translated into Hindi and Telugu to facilitate participant responses. This combined approach supported more inclusive communication while addressing limitations in modifying in-system language settings.

5.2 Multi-stakeholder participation

Construction workers, site supervisors, and safety personnel were invited to participate in VR-based safety training sessions to incorporate both operational and supervisory perspectives. Participants were trained using an Oculus Rift S–type head-mounted display (HMD) with handheld controllers, enabling consistent interaction across users during the training sessions. Figure 7 illustrates participant engagement during the indoor VR training sessions.

Workers interacted with predefined hazard scenarios, including work-at-height conditions, open edges, and material-handling situations, by identifying visible hazards and selecting appropriate safe-response actions. Supervisors and safety personnel used a structured observation checklist to record participant responses, including identified hazards and selected safety actions.

Participants were provided with a brief orientation to VR controls prior to the session to ensure familiarity with system interaction. Hazard identification and response tasks were primarily conveyed through visual and spatial cues, with facilitator-supported verbal guidance used where necessary to support consistent understanding across participants with varying literacy levels. Facilitators provided verbal instructions in local languages where necessary to support comprehension during the training sessions.

Figure 8 shows a participant engaged in the VR-based safety training using a head-mounted display, while the simulation view is projected on a screen for real-time monitoring. The facilitator provides guidance and translates safety instructions into the participant's preferred language, ensuring clear understanding of hazard scenarios and training tasks. This setup illustrates the combined use of immersive interaction and supervised instruction during the training sessions.

Was voluntary and coordinated with site management, which may have introduced selection bias toward individuals more motivated or more comfortable engaging with new training methods.



FIGURE 8
Participant-session photograph showing VR training with projected display and facilitator support.

5.3 Hazard analysis, data collection, and performance assessment

Data were collected from the BIM–VR safety training through structured observations, system-recorded interaction timing, and risk-scoring matrices to document hazard-identification responses using a predefined checklist of site hazards and standardized scoring criteria. Hazard-identification data were collected at three time points: a pre-training baseline assessment conducted before the VR session, an immediate post-session assessment conducted after completion of the VR training, and a 4-week follow-up assessment. All three assessments were conducted using the same VR environment and scoring criteria to ensure consistency across time points within the study context.

Primary data were collected from (i) human observations by trained safety personnel during VR sessions, (ii) system-recorded interaction timing data (e.g., total session duration and approximate time taken to identify hazards), and (iii) manual records of post-session feedback. All hazard identification outcomes, correctness of responses, and participant behaviour (e.g., confusion or assistance required) were documented using standardized observation procedures. These data were collected from all 30 participants and compiled to track participants' ability to identify and respond to six predefined hazard scenarios, with performance scored based on correct identification and appropriate safe responses.

For standardized risk assessment, we used the Probability–Impact Matrix presented in Table 1. Likelihood of risk was classified into five levels: Very Unlikely (1) to Almost Certain (5), while impact severity varied from Trivial (1) to Severe (5). Each identified hazard was assigned probability and severity scores, and based on these scores, a risk level was determined, assigning qualitative risk categories from Low to Extreme, as described in Table 1. For instance, a hazard classified as Almost Certain–Severe was categorized as Extreme risk, whereas Very Unlikely–Trivial was categorized as Low and thus considered lower priority.

To relate hazard identification to corresponding safety practices, a second evaluation step matched each identified hazard with appropriate control measures (e.g., harnesses, guardrails, and access control), adapted from established safety standards such as OSHAD and comparable regulations. The risk assessment matrix was applied as a design-stage analytical tool to support hazard classification and scenario development. The "residual risk" levels estimated using the matrix represent hypothetical conditions assuming the application of appropriate safety controls and are included for illustrative purposes only. These values do not represent measured outcomes and should not be interpreted as evidence of risk reduction resulting from the VR training intervention.

The matrix framework allows comparison of qualitative risk classifications across scenarios based on applied control measures and hazard characteristics. The matrix-based approach was used solely as a structured hazard-analysis and training-design tool to support scenario selection and prioritization; it does not represent participant outcome data or independent evidence of training effectiveness, as summarized in Table 2.

To ensure transparency and reproducibility of the intervention, the VR task protocol was implemented using a consistent procedure across all assessment stages (pre-training, post-training, and follow-up). Participants navigated a single BIM-based virtual environment using joystick locomotion in a standing position, freely exploring and identifying hazards without a fixed sequence. Hazard identification and corresponding safety responses were reported verbally and recorded by observers, with performance scoring applied post-session using standardized criteria. A detailed description of this workflow, including session structure, interaction mechanics, and data capture methods, is provided in Supplementary Appendix B.

To ensure clarity and consistency in data collection, all instruments were standardized and documented. The structured hazard-recognition task and scoring criteria were consistently applied across all assessment stages, while observer-based session monitoring was conducted using a predefined checklist. The detailed formats of these instruments, including questionnaire items (Supplementary Appendix A), hazard scenarios and scoring rubrics (Supplementary Appendix C), and observation checklist criteria, are provided in Supplementary Appendix D.

The matrix represents an ordinal risk classification framework in which probability and severity categories are combined for qualitative risk prioritization. The resulting levels are intended for comparative classification and should not be interpreted as precise quantitative measurements.

Risk assessment was conducted using the NCEMA framework, as referenced in Table 2. Each hazard was evaluated using a risk scoring approach in which probability (Pos) and severity (Ser) categories were combined to determine an overall risk level, as described in the Probability–Impact framework presented in Table 1. The risk score for each hazard was calculated using the formula:

TABLE 1 Probability impact matrix.

Probability (Prob)/ Severity (Sev)	Trivial (Carter and Smith, 2006)	Minor (Gao et al., 2019)	Moderate (Zhang et al., 2013)	Major (Martínez-Rojas et al., 2020)	Severe (Alsamadani et al., 2013)
Very Unlikely (Carter and Smith, 2006)	Low	Low	Low	Medium	Medium
Unlikely (Gao et al., 2019)	Low	Low	Medium	Medium	High
Possible (Zhang et al., 2013)	Low	Medium	Medium	High	High
Likely (Martínez-Rojas et al., 2020)	Medium	Medium	High	High	Extreme
Almost Certain (Alsamadani et al., 2013)	Medium	High	High	Extreme	Extreme

$$\text{Risk Score} = \text{Severity (Sev)} \times \text{Probability (Prob)}$$

Severity and probability were scored on ordinal 5-point scales using standardized rating guidance to ensure consistency across assessments. The same scoring procedure was applied across all hazard scenarios within the matrix framework for comparative risk classification.

Each hazard scenario was mapped to corresponding control measures based on established occupational safety standards, including engineering and administrative controls. Table 2 presents a structured hazard–control mapping framework used for risk classification within the study design. The matrix is intended as a structured analytical tool for hazard categorization and does not represent participant outcome data or direct evidence of training effectiveness.

The risk-matrix framework is conceptually informed by established national and occupational safety guidelines, including NCEMA and OSHAD standards. It provides a consistent structure for categorizing hazards based on severity and probability to support systematic hazard analysis within the study framework.

For work at unguarded floor perimeters and door/window openings, the risk score was calculated as $4 \times 4 = 16$ based on assigned probability and severity categories, categorizing these hazards as High. These hazards are typical of fall-from-height situations and are associated with high likelihood and high severity consequences. Control measures such as edge protection, safe access practices, harness use, barricading, and supervision were mapped to this hazard scenario within the risk assessment framework.

Falling material hazards were initially scored as $4 \times 3 = 12$ based on assigned probability and severity categories, indicating a high-risk classification. Control measures such as safety nets, material handling, and supervision were associated with this hazard scenario within the risk assessment framework.

Regarding handrails, the inherent risk score was $4 \times 3 = 12$, mainly due to the potential for slips, trips, and falls on stairs. VR training modules included content on proper handrail use, identification of missing or insufficient handrails, and safe movement practices. This hazard scenario was included within the structured risk assessment framework using the predefined severity and probability scoring system.

The riskiest hazard was scaffolding collapse caused by poor erection, initially scored as $5 \times 3 = 15$, corresponding to high severity and moderate likelihood. This hazard was addressed through VR-based spatial training covering scaffold erection, inspection procedures, anchorage awareness, certification requirements, and inspection guidelines, in combination with standard safety practices. This scenario was included within the structured risk assessment framework using predefined severity and probability criteria.

The risk-matrix presented in Table 2 is an illustrative hazard–control framework used for systematic risk classification. The matrix is conceptually informed by the National Emergency Crisis and Disaster Management Authority (NCEMA) risk assessment model, which applies ordinal categories of probability and severity.

To assess knowledge retention and hazard identification over time, a follow-up assessment was conducted 4 weeks post-training using the same hazard identification instrument and scoring criteria as the baseline assessment. Participants were required to navigate the VR scenarios and identify hazards along with appropriate safe responses. The same assessment procedure was applied consistently across all time points to ensure uniform data collection within the study.

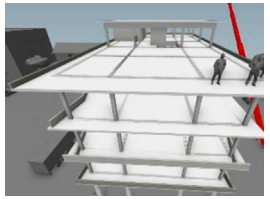
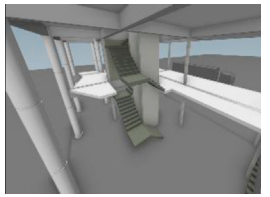
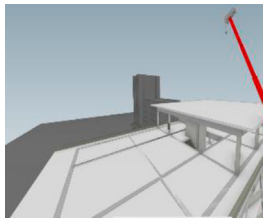
6 Effectiveness evaluation

6.1 Analysis of participant experience through simulated survey

The participant responses presented in Figure 9 were obtained from a structured post-training questionnaire administered via Google Forms to all participants (N = 30) immediately after completion of the VR safety training sessions. The survey consisted of seven closed-ended items rated on a five-point Likert scale (1 = very low, 5 = very high). The items evaluated clarity of instructions, engagement during training, ease of VR device use, physical comfort during immersion, perceived usefulness for hazard recognition, understanding of safety procedures, and overall satisfaction and willingness to recommend VR-based training.

Participant responses were summarized using mean scores for each item, which are presented in Figure 9. Higher mean scores were

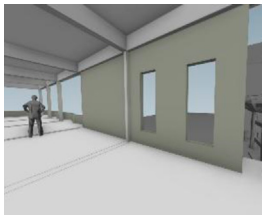
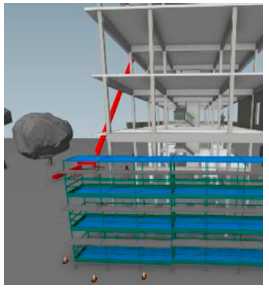
TABLE 2 Illustrative hazard–control and qualitative risk classification framework for scenario design.

S.No	Hazard scenario	Potential consequence	Affected personnel	Initial qualitative risk rating (Sev × Prob)	Integrated control measures (Including VR support)	References standards	Illustrative residual risk rating (Design assumption) (Sev × Prob)	VR Scenario
1	Working at an unguarded floor edge	Slips, trips, and falls from height	Workers, supervisors, site personnel	4 × 4 (High)	Guardrails/barricades, safety harness use, supervision, and VR-supported fall hazard awareness	OSHAD CoP 22, 17	1 × 4 (Medium)	
2	Working near door or floor openings at height	Falls from openings	Workers, supervisors, site personnel	4 × 4 (High)	Temporary guardrails, controlled access, supervision, and VR-supported safe movement awareness	OSHAD CoP 22, 17	1 × 4 (Medium)	
3	Falling materials from height	Injury from falling objects	Workers in lower zones	4 × 3 (High)	Safety nets, controlled material handling, exclusion zones, and VR-supported overhead hazard awareness	OSHAD CoP 26	1 × 3 (Low)	
4	Inadequate or missing handrails	Slips, trips, and falls on stairs	Workers and site personnel	4 × 3 (High)	Compliant handrails, safe access practices, and VR-supported navigation awareness	OSHAD CoP 22, 17	1 × 3 (Low)	

(*Continued*)

10.3389/frvir.2026.1768058

TABLE 2 Continued

S.No	Hazard scenario	Potential consequence	Affected personnel	Initial qualitative risk rating (Sev × Prob)	Integrated control measures (Including VR support)	References standards	Illustrative residual risk rating (Design assumption) (Sev × Prob)	VR Scenario
5	Window or edge openings	Falls from unprotected openings	Workers, supervisors	4 × 4 (High)	Barricading, restricted access, supervision, and VR-supported spatial hazard recognition	OSHAD CoP 22, 17	1 × 4 (Medium)	
6	Scaffolding collapse due to improper setup	Falls and structural collapse injuries	Workers using scaffolding	5 × 3 (High)	Certified scaffold erection, inspection protocols, supervision, and VR-supported safety awareness training	OSHAD CoP 26	2 × 3 (Medium)	

Residual risk values are hypothetical and provided for illustrative purposes within the scenario design framework; they do not represent measured outcomes or evidence of risk reduction due to the VR, training intervention.

10.3389/frvir.2026.1768058

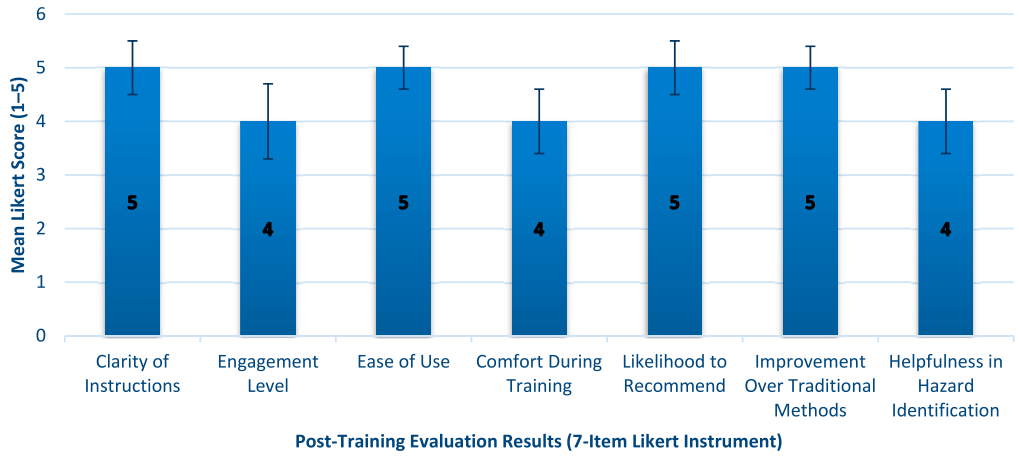


FIGURE 9
Mean scores from a seven-item post-training Likert-scale questionnaire. Error bars represent standard deviation.

recorded for items related to engagement, clarity of instructions, and perceived usefulness, as shown in the survey results.

Open-ended responses included comments on scenario duration and visual/auditory cue strength. Suggestions included shorter scenario durations, improved transition cues, and reduced orientation time for first-time users.

The mean score (μ) was computed for each survey item by:

$$\mu = \frac{\sum_{i=1}^{N} x_i}{N}$$

Where x_i represents the individual participant's rating and N is the total number of respondents. The standard deviation (**σ**) was also to describe the variability of responses as:

$$\boldsymbol{\sigma} = \sqrt{\frac{\sum (x_i - \mu)^2}{N - 1}}$$

The results showed mean values of 5.0 for Clarity of Instructions, Ease of Use, and Likelihood to Recommend, with standard deviations approximately between 0.8 and 1.0, indicating some variability despite high central tendency. Engagement Level, Comfort During Training, and Helpfulness in Hazard Identification recorded mean values of 4.0 with standard deviations ranging between approximately 0.8–1.0.

To examine the internal consistency of the survey items, the seven Likert-scale items (k = 7) were evaluated using Cronbach's alpha (α). The coefficient was calculated using the standard formula:

$$\alpha = \left(k \,/\, (k-1)\right) \times \left[1 - \left(\Sigma\sigma^2 i \,/\, \sigma^2{}_t\right)\right]$$

Where k = 7 represents the number of Likert-scale items, $\Sigma\sigma^2$ is the sum of individual item variances, and $\sigma^2{}_t$ is the variance of the total summed scores.

For each item, variance was calculated as:

$$\sigma^2 i = \sum (x - \bar{x})^2 \Big/ (n-1)$$

Where x represents individual responses, $\bar{x}$ is the mean of that item, and n = 30 participants.

The variance of total scores was computed as:

$$\sigma^2 t = \sum (T - \bar{T})^2 \Big/ (n-1)$$

Where T represents each participant's total score (sum of seven items), and $\bar{T}$ represents the mean total score of participants.

From the dataset of 30 responses, the seven item variances were 0.82, 0.74, 0.91, 0.65, 0.88, 0.77, and 0.71, giving:

$$\Sigma\sigma^2 \mathrm{i} = 5.48$$

The variance of total summed scores was:

$$\sigma_t^2 = 28.35$$

Substituting into the Cronbach's alpha formula yielded:

$$\alpha = (7/6) \times [1 - (5.48/28.35)]$$

$$\alpha = 0.94$$

Since α = 0.94 (>0.70), the Cronbach's alpha value indicates excellent internal consistency of the seven Likert-scale items used in the survey instrument.

A percentage distribution was calculated based on the proportion of responses in each Likert-scale category (Very low to Very high). Across all survey categories, more than 80% of responses were recorded in the "High" or "Very high" categories. The results are presented as descriptive frequency distributions of participant responses across the seven survey items.

The statistical results include measures of central tendency, percentage distribution, and internal consistency for the BIM-integrated VR training survey. The instrument showed a Cronbach's alpha value of 0.94, and more than 80% of responses were recorded in the "High" or "Very high" categories across all items. The results are reported descriptively within the study sample (N = 30).

6.2 Contextual comparison with traditional training approaches

A contextual comparison is presented between VR-based training in the present study and characteristics reported in literature for traditional safety training methods, such as

TABLE 3 Contextual overview of VR-based training and conventional training approaches (literature-informed).

Aspect	VR-Based training (This study)	Traditional training (Literature context)
Engagement Level	Mean score: 4.8/5 (study participants)	Generally reported as moderate in prior studies
Hazard Recognition	92% accuracy in VR scenarios	Reported to vary depending on delivery method
Knowledge Retention	84% retention after 4 weeks	Often reported to decline over time
Training Time	15–20 min per session	Typically longer session durations
Multilingual Effectiveness	Facilitator-assisted with visual cues and translated support	Often language-dependent
Cost-effectiveness	Requires VR setup and facilitation	Typically implemented through classroom-based or presentation-based formats

classroom lectures and slide-based instruction. The comparison is structured across parameters including engagement, hazard recognition performance, knowledge retention, training duration, multilingual support, and implementation characteristics, as summarized in Table 3.

Participant responses in the exit survey were measured using a five-point Likert-type scale (1 = very low, 5 = very high). Mean engagement scores were calculated using:

$$\mu = \frac{\sum_{i=1}^{N} x_i}{N}$$

Hazard-detection performance was measured using a structured identification task integrated into the training scenarios. Six standardized hazards associated with work-at-height and falling-object risks were presented to participants (e.g., unguarded edges, insecure scaffolds, lack of handrails, and poor material handling). Participant responses were recorded as correct when the hazard was accurately identified and an appropriate safe response was selected. Accuracy was calculated for each participant as the proportion of correctly identified hazards out of the total number of hazards presented. A pre-training assessment was conducted to establish baseline performance, resulting in a mean score of 68.2% (SD = 8.3). A post-training assessment recorded a mean accuracy score of 91.8% (SD = 5.6).

Knowledge retention was assessed 4 weeks after training using the same hazard recall and application instrument administered immediately post-training. Participants completed a short assessment involving hazard identification within the VR scenarios and selection of appropriate preventive actions. Retention percentage was calculated as the ratio of delayed-test scores to immediate post-training scores. The mean retention score at the 4-week follow-up was 84.0% (SD = 6.4).

Comparison of total training time was used to describe training duration within the study context. Sessions conducted in VR took approximately 15–20 min per participant.

Sacks et al. (2013) found that VR-based construction safety training was associated with higher hazard identification performance compared to lecture-based instruction. A meta-analysis by Scorgie et al. (2024) further reports that VR-based training is generally associated with behavioural and skill-related outcomes across multiple domains, with standardized effect sizes ranging from 0.432 to 0.777. Similar patterns are reported by Man et al. (2024) in construction safety training contexts, while Zhang et al. (2022) highlights higher engagement and perceived learning outcomes in VR-based approaches compared to conventional methods.

In the present study, participants recorded a mean hazard identification accuracy of 91.8% and a mean engagement score of 4.8. These values are reported for contextual comparison with findings from previous studies.

As shown in Table 3, VR-based training is descriptively compared across engagement, hazard recognition, knowledge retention, training duration, and inclusiveness based on study data and literature evidence.

6.3 Thematic insights on VR-Based safety training

Table 4 presents thematic results from participants' and experts' responses on the BIM-integrated VR safety training framework. Quantitative data were collected using the same structured post-training questionnaire described in Section 6.1, administered via Google Forms, in which respondents rated each item on a five-point Likert scale (1 = very low, 5 = very high). The survey was administered to site workers, supervisors, and safety personnel shortly after training sessions. Mean scores were calculated for each safety dimension to summarize responses.

The results presented in Table 4 do not represent direct one-to-one questionnaire items from Supplementary Appendix A. Instead, the seven Likert-scale items (Items A1–A7 in Supplementary Appendix A) were interpreted and grouped into broader thematic dimensions (e.g., engagement, hazard detection, communication clarity, and perceived usefulness) to provide a higher-level summary of participant perceptions. Each thematic category reflects one or more related questionnaire items, and in some cases a combination of multiple items, supported by qualitative feedback collected during the sessions. For example, engagement-related scores correspond directly to responses from Item A2 (training engagement), while hazard detection ability reflects Item A5 (usefulness in identifying hazards). Other categories, such as multilingual participation and safety capacity building, were derived by combining responses related to clarity of instructions (A1), understanding of safety protocols (A6), and supporting qualitative feedback.

As indicated in Table 4, the mean score for hazard detection was 4.7. Safety engagement recorded the highest mean score (M = 4.8), followed by safety capacity building (M = 4.6) and multilingual team participation (M = 4.5). These values represent descriptive mean scores across the evaluated dimensions.

TABLE 4 Mixed-method research findings overview.

Sr No.	Issues	Mean score	Results
1	Improve overall safety planning process	4.0	High
2	Improve hazard detection ability	4.7	Very High
3	Enhance engagement levels during safety training	4.8	Very High
4	Facilitate multilingual team participation	4.5	Very High
5	Increase safety-related capacity building	4.6	Very High
6	Enhance communication of safety instructions	4.2	High
7	Improve access to safety information	4.0	High

Moderate-to-high mean scores were recorded for communication flow (M = 4.2) and access to safety information (M = 4.0). The safety planning process recorded a mean score of 4.0. These values represent descriptive responses across the evaluated dimensions. Mean scores across all categories were ≥4.0. Standard deviations across *items* were generally low, indicating limited variability in responses across the measured dimensions. Descriptive statistical results are reported for all survey items (N = 30).

The findings are based on self-reported responses collected immediately after completion of the VR training sessions. Responses were obtained using a structured questionnaire administered via Google Forms to participants (N = 30), covering engagement, clarity of instructions, ease of understanding safety procedures, and perceived usefulness of training. The dataset represents immediate post-training feedback captured during the training sessions.

The results reflect participants' ratings of the VR-based training system across the evaluated dimensions. Responses were recorded on a five-point Likert scale and summarized using descriptive statistics including mean values and standard deviations. The evaluation captured participants' immediate perceptions following exposure to BIM-integrated VR safety scenarios within the training environment.

All responses were collected immediately after training completion to ensure consistency in timing across participants. The dataset includes responses from workers, supervisors, and safety personnel who participated in the VR training sessions within the study site context (N = 30).

7 Discussion

This study presents exploratory evidence on the integration of VR with BIM for supporting safety communication and hazard identification training in multilingual construction environments. The observed outcomes in participant engagement, hazard recognition performance, and short-term retention should be interpreted within the constraints of a pilot quasi-experimental design, an indoor training environment near the project site, and a volunteer-based sample. Because participation was voluntary and coordinated through site management without randomization or stratification, the sample may have included workers who were more motivated, more comfortable with technology, or more engaged than the broader workforce, potentially limiting generalizability. While performance outcomes were consistently high across measured variables, several contextual and methodological factors may have influenced these patterns. Accordingly, the high observed scores should be interpreted cautiously as context-dependent pilot findings rather than direct indicators of generalized training effectiveness.

A key factor contributing to elevated performance outcomes is the novelty and immersion effect associated with VR-based training. Participants engaged with an interactive and visually rich environment that differs significantly from conventional instructional methods, which may have increased attention, motivation, and task focus during both training and assessment phases. Such effects are typically more pronounced in early-stage exposure studies and may not necessarily reflect sustained learning outcomes over time.

In addition, the structured and predefined nature of hazard identification tasks may have contributed to ceiling effects in performance measurement. Participants were required to identify a limited set of clearly defined hazards within predefined VR scenarios, reducing ambiguity and potentially resulting in clustered high scores. This limits the ability to distinguish between surface-level recognition and deeper conceptual understanding of hazard conditions.

The training design also incorporated facilitator-assisted translation and verbal clarification to support multilingual participants. While necessary for ensuring comprehension, this introduces a co-intervention effect, making it difficult to isolate the independent contribution of the VR system. The observed outcomes therefore reflect an interaction between immersive simulation, human facilitation, and structured task design. Direct differences in comprehension or learning outcomes between language groups were not formally measured in this study.

Furthermore, the immediate post-training assessment likely captured short-term recall effects rather than stable behavioural change. The indoor simulation environment also differs from real construction settings in terms of time pressure, multitasking demands, environmental complexity, and safety risk variability, all of which may influence hazard recognition performance in practice.

Additional methodological considerations include the absence of formal inter-rater reliability testing for observer-based

assessments, which may introduce variability in scoring despite the use of structured checklists. Self-reported engagement measures may also be influenced by response bias due to the immediate post-session evaluation context and participant awareness of being part of a training study.

Collectively, these factors indicate that the observed performance levels likely reflect a combination of immersive novelty, structured task design, facilitator support, and short-term cognitive effects within the study setting. Accordingly, the findings primarily support the feasibility of BIM–VR integration for structured safety training rather than evidence of sustained behavioural change or real-world risk reduction.

Despite these constraints, the study demonstrates a structured BIM–VR framework capable of representing construction hazards in an immersive and structured format with multilingual facilitation support. This highlights the potential role of visual-spatial training tools as a supplementary component within existing safety management systems, particularly in contexts where language diversity may present communication challenges, with facilitator-supported delivery playing an important role in conventional training approaches.

Future research should extend this work through multi-site experimental designs with control groups, larger and more diverse participant samples, and longitudinal assessment of safety behaviour. The inclusion of objective workplace safety indicators such as near-miss events, behavioural observations, and incident tracking would strengthen evidence of real-world transfer. Additionally, integrating adaptive scenario complexity, biometric measures (e.g., eye-tracking and response latency), and standardized inter-rater reliability protocols would improve methodological rigor and interpretability of outcomes.

8 Conclusion

This study investigated the feasibility of integrating BIM and VR within a multilingual construction project to support safety communication and hazard-identification training. Location-specific BIM-based VR scenarios were developed and implemented with construction workers, supervisors, and safety personnel in an indoor training setting near the project site, demonstrating the practical applicability of immersive visualization tools for site-relevant safety training.

The findings should be interpreted as indicative outcomes from this pilot implementation rather than evidence of causal effectiveness. The observed performance patterns were likely influenced by multiple interacting factors, including immersive novelty effects, structured task design, facilitator-assisted translation, and immediate post-training assessment conditions. In addition, the presence of existing engineering and administrative controls within the site environment limits the ability to attribute outcomes solely to the VR-based intervention.

This work contributes a structured BIM–VR framework that supports safety training delivery in multilingual environments, highlighting its potential role as a supplementary tool within existing safety management systems. However, the results do not confirm long-term behavioural change or direct reductions in real-world safety incidents, and therefore should be interpreted within the constraints of this pilot study.

Future research should focus on multi-site experimental validation, larger and more diverse participant samples, and longitudinal study designs to evaluate sustained learning outcomes and real-world safety performance. Further development may include adaptive scenario complexity, expanded hazard libraries, collaborative multi-user environments, and integration with real-time construction site data to enhance realism and training effectiveness.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author.

Ethics statement

The studies involving humans were approved by Institutional Ethics Committee on Human Studies (IECH) VIT. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

MB: Conceptualization, Data curation, Formal Analysis, Investigation, Methodology, Visualization, Writing – original draft. AB: Conceptualization, Data curation, Formal Analysis, Investigation, Methodology, Writing – original draft. MS: Conceptualization, Data curation, Methodology, Supervision, Writing – review and editing. PR: Conceptualization, Data curation, Supervision, Writing – review and editing, Methodology, Validation.

Funding

The author(s) declared that financial support was not received for this work and/or its publication.

Conflict of interest

The author(s) declared that this work was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declared that generative AI was not used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/frvir.2026.1768058/full#supplementary-material

References

Abbas Ali, A., Choi, M., Seo, J. O., Cha, S. H., and Li, H. (2019). Effectiveness of immersive virtual reality-based communication for construction projects. *KSCE J. Civ. Eng.* 23 (12), 4972–4983. doi:10.1007/s12205-019-0898-0

Afzal, S., and Shafiq, M. T. (2021). Evaluating 4D-BIM and VR for effective safety communication and training: a case study of multilingual construction job-site crew. *Saf. Sci.* 134, 105047. doi:10.3390/buildings11080319

Alizadehsalehi, S., Hadavi, A., and Huang, J. C. (2020). From BIM to extended reality in AEC industry. *Automation Constr.* 116, 103254. doi:10.1016/j.autcon.2020.103254

Alsamadani, R., Hallowell, M. R., Javernick-Will, A., and Cabello, J. (2013). Relationships among language proficiency, communication patterns, and safety performance in small work crews in the United States. *J. Constr. Eng. Manag.* 139 (9), 1125–1134. doi:10.1061/(asce)co.1943-7862.0000724

Azhar, S. (2017). Role of visualization technologies in safety planning and management at construction jobsites. *Procedia Eng.* 171, 215–226. doi:10.1016/j.proeng.2017.01.329

Carter, G., and Smith, S. D. (2006). Safety hazard identification on construction projects. *J. Constr. Eng. Manag.* 132 (2), 197–205. doi:10.1061/(asce)0733-9364(2006)132:2(197)

Choe, S., and Leite, F. (2017). Construction safety planning: site-Specific temporal and spatial information integration. *Automation Constr.* 84, 335–344. doi:10.1016/j.autcon.2017.09.007

Du, J., Zou, Z., Shi, Y., and Zhao, D. (2018). Zero latency: real-Time synchronization of BIM data in virtual reality for collaborative decision-making. *Automation Constr.* 85, 51–64. doi:10.1016/j.autcon.2017.10.009

Gao, Y., Gonzalez, V. A., and Yiu, T. W. (2019). The effectiveness of traditional tools and computer-aided technologies for health and safety training in the construction sector: a systematic review. *Comput. and Educ.* 138, 101–115. doi:10.1016/j.compedu.2019.05.003

Getuli, V., Capone, P., Bruttini, A., and Isaac, S. (2020). BIM-Based immersive virtual reality for construction workspace planning: a safety-oriented approach. *Automation Constr.* 114, 103160. doi:10.1016/j.autcon.2020.103160

Goulding, J., Nadim, W., Petridis, P., and Alshawi, M. (2012). Construction industry offsite production: a virtual reality interactive training environment prototype. *Adv. Eng. Inf.* 26 (1), 103–116. doi:10.1016/j.aei.2011.09.004

Johansson, M., and Roupé, M. (2024). Real-world applications of BIM and immersive VR in construction. *Automation Constr.* 158, 105233. doi:10.1016/j.autcon.2023.105233

Johansson, M., Roupé, M., and Bosch-Sijtsema, P. (2015). Real-time visualization of building information models (BIM). *Automation Constr.* 54, 69–82. doi:10.1016/j.autcon.2015.03.018

Li, X., Yi, W., Chi, H.-L., Wang, X., and Chan, A. P. (2018). A critical review of virtual and augmented reality (VR/AR) applications in construction safety. *Automation Constr.* 86, 150–162. doi:10.1016/j.autcon.2017.11.003

Lim, E. C., and Alum, J. (1995). Construction productivity: issues encountered by contractors in Singapore. *Int. J. Proj. Manag.* 13 (1), 51–58. doi:10.1016/0263-7863(95)95704-H

Malekitabar, H., Ardeshir, A., Sebt, M. H., and Stouffs, R. (2016). Construction safety risk drivers: a BIM approach. *Saf. Sci.* 82, 445–455. doi:10.1016/j.ssci.2015.11.002

Man, S. S., Wen, H., and So, B. C. L. (2024). Are virtual reality applications effective for construction safety training and education? A systematic review and meta-analysis. *J. Saf. Res.* 88, 230–243. doi:10.1016/j.jsr.2023.11.011

Martínez-Rojas, M., Martín Antolín, R., Salguero-Caparrós, F., and Rubio-Romero, J. C. (2020). Management of construction safety and health plans based on automated content analysis. *Automation Constr.* 120, 103362. doi:10.1016/j.autcon.2020.103362

Pedro, A., Le, Q. T., and Park, C. S. (2016). Framework for integrating safety into construction methods education through interactive virtual reality. *J. Prof. Issues Eng. Educ. Pract.* 142 (1), 04015011. doi:10.1061/(ASCE)EI.1943-5541.0000261

Prabhakaran, A., Mahamadu, A. M., and Mahdjoubi, L. (2022). Understanding the challenges of immersive technology use in the architecture and construction industry: a systematic review. *Automation Constr.* 137, 104228. doi:10.1016/j.autcon.2022.104228

Rajendran, S., and Clarke, B. (2011). Building information modeling: safety benefits and opportunities. *Prof. Saf.* 56 (10), 44–51.

Sacks, R., Perlman, A., and Barak, R. (2013). Construction safety training using immersive virtual reality. *Constr. Manag. Econ.* 31 (9), 1005–1017. doi:10.1080/01446193.2013.828844

Sacks, R., Whyte, J., Swissa, D., Raviv, G., Zhou, W., and Shapira, A. (2015). Safety by design: dialogues between designers and builders using virtual reality. *Constr. Manag. Econ.* 33 (1), 55–72. doi:10.1080/01446193.2015.1029504

Scorgie, D., Feng, Z., Paes, D., Parisi, F., Yiu, T. W., and Lovreglio, R. (2024). Virtual reality for safety training: a systematic literature review and meta-analysis. *Saf. Sci.* 171, 106372. doi:10.1016/j.ssci.2023.106372

Shringi, A., Arashpour, M., Golafshani, E. M., Rajabifard, A., Dwyer, T., and Li, H. (2022). Efficiency of VR-based safety training for construction equipment: hazard recognition in heavy machinery operations. *Buildings* 12 (12), 2084. doi:10.3390/buildings12122084

Wang, K., Antwi-Afari, M. F., Ossei-Gudom, B. M., Yi, W., Li, J., and Banihashemi, S. (2025). Recent advances in virtual reality applications for construction safety management. *Constr. Innov. Inf. Process Manag.* doi:10.1108/CI-11-2024-0357

Xiong, Z., Ding, Z., Li, Y., and Zhang, J. (2026). Investigating VR safety training transfer in construction hazard recognition: a neurocognitive perspective. *Adv. Eng. Inf.* 71, 104412. doi:10.1016/j.aei.2026.104412

Zhang, S., Teizer, J., Lee, J., Eastman, C. M., and Venugopal, M. (2013). Building information modeling (BIM) and safety: automatic safety checking of construction models and schedules. *Automation Constr.* 29, 183–195. doi:10.1016/j.autcon.2012.05.006

Zhang, M., Shu, L., Luo, X., Yuan, M., and Zheng, X. (2022). Virtual reality technology in construction safety training: extended technology acceptance model. *Automation Constr.* 135, 104113. doi:10.1016/j.autcon.2021.104113

Zhao, D., and Lucas, J. (2015). Virtual reality simulation for construction safety promotion. *Int. J. Inj. Control Saf. Promot.* 22 (1), 57–67. doi:10.1080/17457300.2013.861853

frontiers in Built Environment

TYPE Review
PUBLISHED 15 July 2026
DOI 10.3389/fbuil.2026.1870384

Check for updates

OPEN ACCESS

EDITED BY
Vagelis Plevris,
Qatar University, Qatar

REVIEWED BY
Aydin Teymourifar,
Istanbul Sabahattin Zaim University, Türkiye
Evi Susilawati,
Islamic University of North Sumatra, Indonesia

*CORRESPONDENCE
Imoleayo A. Awodele,
a.imole@yahoo.com

RECEIVED 01 May 2026
REVISED 26 May 2026
ACCEPTED 31 May 2026
PUBLISHED 15 July 2026

CITATION
Awodele IA, Ramabodu MS, Olatunde NA, Eze EC, Anugwo IC, Espinosa YP, Kehinde TO, Arijeloye BT and Ogunlolu AT (2026) The role of digital technologies in engineering procurement: a systematic literature review.
Front. Built Environ. 12:1870384.
doi: 10.3389/fbuil.2026.1870384

COPYRIGHT
© 2026 Awodele, Ramabodu, Olatunde, Eze, Anugwo, Espinosa, Kehinde, Arijeloye and Ogunlolu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

The role of digital technologies in engineering procurement: a systematic literature review

Imoleayo A. Awodele[1]*, Molusiwa S. Ramabodu[1], Nathaniel Ayinde Olatunde[1], Emmanuel C. Eze[2], Iruka C. Anugwo[1], Yesica P. Espinosa[3], Temitope O. Kehinde[4], Bamidele T. Arijeloye[1] and Alice T. Ogunlolu[5]

[1]Construction Management and Quantity Surveying, Durban University of Technology, Durban, South Africa, [2]Teesside University School of Computing Engineering and Digital Technologies, Middlesbrough, United Kingdom, [3]Dpto. Organización de Empresas y CIM E., Universidad de Valladolid, Valladolid, Spain, [4]Department of Industrial and Systems Engineering, The Hong Kong Polytechnic University, Hong Kong, Hong Kong SAR, China, [5]Department of Quantity Surveying and Commercial Management, Coventry University, Coventry, United Kingdom

The increasing complexity of engineering procurement and the growing demand for efficiency, transparency, and resilience in supply chains have intensified the need for digital transformation in procurement systems. Despite the rapid advancement of Industry 4.0 technologies, existing studies on digital procurement remain fragmented, often focusing on individual technologies or general supply chain contexts rather than engineering procurement environments. This study examines the role of digital technologies in engineering procurement and provides an integrated understanding of their impact, associated challenges, and transformation pathways. A systematic literature review (SLR) approach was adopted, guided by PRISMA principles, using peer-reviewed studies published between 2015 and 2024 from major academic databases. A thematic and cross-study synthesis was employed to analyse the conceptual evolution of digital procurement, the role of key technologies, and barriers to adoption. The findings reveal that technologies such as artificial intelligence (AI), blockchain, the Internet of Things (IoT), Building Information Modelling (BIM), robotic process automation (RPA), and e-procurement systems are transforming procurement from a transactional function into a strategic, data-driven capability. These technologies enhance procurement performance by improving operational efficiency, cost optimization, transparency, and supply chain resilience. However, their adoption is constrained by key challenges, including system integration complexities, organizational resistance, skills gaps, and regulatory uncertainties. The study further demonstrates that the effectiveness of digital procurement depends on the integration and complementarity of technologies within interconnected ecosystems, supported by organizational readiness and technological capabilities. This study contributes to Procurement 4.0 literature by providing an integrated synthesis, developing a conceptual framework, and proposing a digital procurement maturity perspective to guide transformation in engineering procurement.

KEYWORDS
digital procurement, engineering procurement, industry 4.0 technologies, procurement 4.0, systematic literature review

1 Introduction

In recent years, the issue of digital transformation in procurement has garnered significant attention due to its implications for supply chain efficiency, organizational performance, and strategic decision-making. As global challenges in supply chain management intensify particularly due to disruptions such as the COVID-19 pandemic and increasing market volatility the need for focused research on digital procurement becomes increasingly evident (Ivanov and Dolgui, 2020). Within engineering sectors such as construction, manufacturing, and infrastructure development, procurement plays a pivotal role, often accounting for a substantial proportion of total project costs. Traditionally viewed as an administrative and transactional function, procurement is increasingly evolving into a strategic, data-driven process driven by advances in digital technologies (Bienhaus and Haddud, 2018). The emergence of Industry 4.0 has accelerated this transformation by integrating cyber-physical systems, data analytics, and automation into organizational processes. In this context, the concept of Procurement 4.0 has gained prominence, emphasizing the use of technologies such as artificial intelligence (AI), blockchain, the Internet of Things (IoT), and e-procurement platforms to enhance procurement efficiency, transparency, and decision-making capabilities (Bienhaus and Haddud, 2018; Seyedghorban et al., 2020; Olatunde et al., 2023; Awodele et al., 2025). Empirical and review-based studies have demonstrated that these technologies enable predictive analytics, improved supplier evaluation, cost optimization, and enhanced supply chain resilience (Klünder et al., 2019; Althabatah et al., 2023). As a result, digital procurement is increasingly recognized as a critical enabler of competitive advantage in modern engineering and supply chain systems.

At the global level, research has extensively examined the role of digital technologies in transforming procurement and supply chain functions. However, the adoption and implementation of these technologies vary significantly across regions. In developing and emerging economies, particularly within the African context, digital procurement adoption is influenced by infrastructural limitations, institutional frameworks, and technological readiness (Ibem et al., 2018; Awodele et al., 2024). In South Africa, for instance, digital procurement is increasingly acknowledged as a value-adding function that enhances supply chain resilience and operational performance, although its implementation is hindered by challenges such as skills shortages, lack of system interoperability, and cybersecurity risks (Motaung and Sifolo, 2023). Furthermore, within engineering and construction sectors, technologies such as Building Information Modelling (BIM) are being integrated into procurement processes to improve coordination, reduce information asymmetry, and enhance project delivery outcomes (Matos et al., 2024). Despite these advancements, procurement systems often remain fragmented, with limited integration across the procurement lifecycle.

A substantial body of literature exists on digital procurement and Industry 4.0 technologies, with scholars emphasizing their impact on operational efficiency, cost performance, and decision-making (Nicoletti, 2020; Seyedghorban et al., 2020). Empirical evidence suggests that organizations adopting digital procurement systems achieve improved efficiency, reduced cycle times, and enhanced supplier collaboration (Senaratne et al., 2023). In addition, technologies such as AI, blockchain, and IoT have been shown to significantly improve transparency, traceability, and risk management within procurement systems (De Rejeb et al., 2018). However, despite these contributions, existing research remains fragmented. Many studies focus on individual technologies or general supply chain contexts, with limited attention to the integrated role of multiple digital technologies within engineering procurement environments. Consequently, there remains a lack of comprehensive understanding of how digital technologies collectively transform engineering procurement processes. Existing studies often overlook the complexity of engineering procurement, which involves technical specifications, multi-stakeholder coordination, and long project lifecycles. Moreover, limited research has provided a holistic synthesis of how different digital technologies interact to form integrated procurement ecosystems. This gap highlights the need for a systematic and comprehensive analysis of digital technologies within the specific context of engineering procurement.

This study addresses this gap by examining the role of digital technologies in engineering procurement through a systematic literature review. The primary objective is to identify key digital technologies and analyze their impact on procurement performance, including efficiency, cost optimization, transparency, and supply chain resilience. In addition, the study seeks to examine the challenges and barriers associated with the adoption of digital procurement technologies and to propose a digital procurement maturity framework that illustrates the progression from traditional procurement systems to intelligent, data-driven ecosystems.

Specifically, this study makes four key contributions to the literature and distinguishes itself from previous Procurement 4.0 reviews through its integrated engineering procurement perspective. First, it provides an integrated synthesis of digital technologies in engineering procurement, addressing the fragmentation in existing studies that predominantly focus on individual technologies or general supply chain contexts (Althabatah et al., 2023). Second, the study critically synthesizes empirical, conceptual, and review-based literature to examine how technologies such as artificial intelligence, blockchain, IoT, BIM, and e-procurement systems collectively influence procurement transformation and operational performance within engineering environments. Third, it develops a conceptual framework that explains the relationships between Industry 4.0 technologies, digital procurement capabilities, and procurement performance outcomes, while incorporating moderating barriers such as organizational resistance and technological integration challenges (Bienhaus and Haddud, 2018; Motaung and Sifolo, 2023). Fourth, it proposes a digital procurement maturity framework and integrated ecosystem perspective that outline the progression from traditional procurement systems toward intelligent and data-driven procurement ecosystems (Seyedghorban et al., 2020).

2 Literature review

2.1 Conceptual foundations and evolution of digital procurement

The concept of procurement has undergone a fundamental transformation from a transactional, administrative function to a

strategic component of supply chain management. Traditionally, procurement activities were centered on cost minimization, supplier selection, and contract administration, often supported by manual processes and fragmented information systems. However, the emergence of digital transformation has significantly redefined procurement practices by embedding advanced technologies into operational and decision-making processes (Aduwo, 2018; Bienhaus and Haddud, 2018; Awodele et al., 2024).

This transformation is closely aligned with the broader paradigm of Industry 4.0, which emphasizes the integration of cyber-physical systems, automation, and data-driven analytics across industrial operations. Within this context, the concept of Procurement 4.0 has emerged as a critical framework, highlighting how digital technologies enable real-time data processing, predictive analytics, and enhanced supply chain coordination (Klünder et al., 2019). Empirical evidence demonstrates that increased levels of digitalization are significantly associated with reductions in procurement costs, thereby reinforcing the economic rationale for digital transformation. The transformation of procurement through digital technologies can also be interpreted through the lens of the Technology–Organization–Environment (TOE) framework, which explains how technological adoption is influenced by technological readiness, organizational capabilities, and external environmental conditions. Within engineering procurement contexts, the TOE framework provides a useful theoretical perspective for understanding how organizations adopt and integrate Industry 4.0 technologies such as AI, blockchain, BIM, IoT, and e-procurement systems. Technological factors such as system interoperability and data infrastructure, organizational factors including digital competencies and resistance to change, and environmental factors such as regulatory conditions and market pressures collectively shape the success of digital procurement transformation. The TOE perspective therefore complements the Procurement 4.0 discourse by providing a broader theoretical explanation for variations in digital procurement adoption and maturity across organizations and regions.

The evolution of digital procurement has occurred progressively, beginning with the adoption of early digital tools such as Electronic Data Interchange (EDI) and Enterprise Resource Planning (ERP) systems. These technologies facilitated the digitization of procurement documentation and improved communication efficiency. Subsequently, the development of e-procurement platforms marked a critical milestone, enabling electronic tendering, supplier management, and transaction processing. More recently, the integration of advanced technologies such as artificial intelligence (AI), blockchain, and the Internet of Things (IoT) has enabled a transition toward intelligent and autonomous procurement systems (Seyedghorban et al., 2020).

A key insight emerging from the literature is the role of digital procurement capabilities as mediating mechanisms that translate data into actionable insights. These capabilities enhance supplier evaluation, demand forecasting, and risk management, thereby improving supply chain resilience and operational efficiency (Hallikas et al., 2021). Furthermore, the increasing convergence of procurement with other organizational functions such as engineering design, logistics, and project management underscores its growing strategic importance.

In engineering and construction contexts, procurement processes are inherently complex due to technical specifications, multi-stakeholder interactions, and extended project lifecycles. Recent studies emphasize the integration of Building Information Modelling (BIM) into procurement systems as a critical development. BIM facilitates the alignment of design and procurement activities, reduces information asymmetry, and enhances coordination across project stakeholders (Matos et al., 2024). However, despite these advancements, the transition toward fully integrated digital procurement ecosystems remains uneven, particularly in developing regions where infrastructural limitations and institutional constraints persist.

3 Digital technologies and their impact on engineering procurement

The rapid advancement of digital technologies has significantly transformed procurement processes across engineering industries. A substantial body of literature highlights the role of key technologies including e-procurement systems, artificial intelligence, blockchain, the Internet of Things, and robotic process automation in enhancing procurement performance (Dubolazov et al., 2019).

E-procurement systems represent the foundational layer of digital procurement, enabling organizations to conduct procurement activities electronically. These systems improve process efficiency, reduce administrative workload, and enhance transparency in procurement operations. Empirical evidence demonstrates that organizations adopting e-procurement systems experience improved procurement efficiency, faster transaction cycles, and enhanced supplier collaboration (Senaratne et al., 2023).

Beyond e-procurement, artificial intelligence and machine learning technologies have emerged as critical enablers of predictive and data-driven procurement. These technologies facilitate demand forecasting, supplier performance evaluation, and risk identification, thereby supporting more informed and proactive decision-making (Seyedghorban et al., 2020). Similarly, blockchain technology enhances transparency and trust by providing secure and immutable transaction records, reducing the risk of fraud and improving traceability across supply chains (Kouhizadeh and Sarkis, 2018).

The Internet of Things contributes to procurement transformation by enabling real-time monitoring of physical assets and supply chain activities. IoT-based systems provide continuous data streams that support proactive decision-making and enhance supply chain resilience (De Rejeb et al., 2018). In parallel, robotic process automation (RPA) automates repetitive procurement tasks, such as invoice processing and order management, thereby reducing human error and improving operational efficiency.

In engineering procurement contexts, the integration of these technologies with project management systems further enhances coordination and performance. For instance, BIM-based procurement systems enable the integration of design data with

TABLE 1 Key digital technologies and their roles in engineering procurement.

Technology	Key functions	Procurement impact	Supporting Studies
E-procurement	Electronic sourcing, tendering	Efficiency, transparency	Senaratne et al. (2023)
AI and ML	Forecasting, analytics	Decision-making, risk management	Seyedghorban et al. (2020)
Blockchain	Secure transactions	Transparency, trust	Kouhizadeh and Sarkis (2018)
IoT	Real-time monitoring	Resilience, visibility	De Rejeb et al. (2018)
BIM	Design–procurement integration	Coordination, accuracy	Matos et al. (2024)
RPA	Process automation	Efficiency, error reduction	De Rejeb et al. (2018)

procurement planning, improving material estimation accuracy and reducing project delays (Matos et al., 2024).

Importantly, cross-study evidence indicates that digital technologies collectively contribute to multiple dimensions of procurement performance, including efficiency, cost reduction, transparency, and resilience. This is supported by studies demonstrating a strong correlation between digitalization and procurement cost reduction (Klündera et al., 2019), as well as improved operational coordination and decision-making. Key digital technologies and their respective roles in engineering procurement are summarised in Table 1.

As shown in Table 1, digital technologies play complementary roles in enhancing procurement performance, with AI and machine learning supporting data-driven decision-making, while blockchain and IoT improve transparency and real-time visibility across supply chains.

4 Challenges, barriers, and future directions

Despite the significant benefits associated with digital procurement, the adoption of digital technologies remains constrained by several technological, organizational, and institutional challenges. One of the primary barriers is the integration of new digital technologies with existing legacy systems. Many organizations continue to rely on outdated information systems that lack compatibility with modern digital platforms, making integration complex and costly (Bienhaus and Haddud, 2018).

Organizational resistance to change also represents a critical challenge. Employees may be reluctant to adopt new technologies due to concerns about job displacement, lack of digital competencies, and uncertainty regarding process changes. Empirical evidence from South Africa highlights that digital procurement adoption is often hindered by skills shortages, lack of training, and organizational inertia (Motaung and Sifolo, 2023). Cybersecurity and data privacy concerns further complicate the adoption of digital procurement systems. As procurement processes become increasingly digitalized, organizations face heightened risks of cyberattacks and data breaches. Additionally, legal and regulatory uncertainties particularly regarding blockchain-based smart contracts pose significant challenges for organizations seeking to adopt advanced technologies.

Despite these challenges, the future of digital procurement is characterized by increasing technological integration and innovation. Emerging technologies such as digital twins, advanced analytics, and generative artificial intelligence are expected to further enhance procurement processes and enable more intelligent and autonomous systems. Recent studies emphasize the importance of digital procurement maturity frameworks in guiding organizations through progressive stages of digital transformation (Althabatah et al., 2023). Overall, while digital technologies offer substantial opportunities for improving procurement performance, their successful adoption requires a holistic approach that addresses technological, organizational, and regulatory challenges. Future research should focus on empirical validation of digital procurement frameworks and explore long-term transformation dynamics within engineering procurement systems. Table 2 summarises the barrier categories, key issues, and supporting evidence related to the adoption of digital technologies in engineering procurement.

As shown in Table 2, the adoption of digital technologies is constrained by a combination of technological, organizational, security, and regulatory barriers, highlighting the multifaceted nature of challenges in engineering procurement.

Table 3 presents a comprehensive synthesis of key studies on digital technologies in engineering procurement, highlighting their methodological approaches, technological focus, and key contributions.

Table 4 highlights that existing studies predominantly focus on individual technologies or general supply chain contexts, with limited attention to integrated digital procurement systems within engineering environments. Furthermore, there is a lack of comprehensive frameworks that capture the interaction of multiple technologies and their collective impact on procurement performance. This study addresses these gaps by providing an integrated synthesis and proposing a digital procurement maturity framework.

While digital technologies offer significant benefits to procurement systems, their adoption is often constrained by various technological, organizational, and regulatory barriers. Table 5 presents a synthesized Technology–Impact–Barrier matrix, highlighting the dual role of digital technologies in enhancing procurement performance while simultaneously introducing implementation challenges.

Table 5 demonstrates that while previous studies have primarily focused on individual technologies, conceptual discussions, or general supply chain contexts, the present study provides a more integrated and

TABLE 2 Major barriers to digital procurement adoption.

Barrier category	Key issues	Supporting evidence
Technological	Legacy systems, interoperability	Bienhaus and Haddud (2018)
Organizational	Resistance, skills gaps	Motaung and Sifolo (2023)
Security	Cyber threats, data privacy	Motaung and Sifolo (2023)
Regulatory	Smart contract uncertainty	De Rejeb et al. (2018)

engineering-focused synthesis of digital procurement transformation. The study further distinguishes itself through the development of a conceptual framework, technology–impact–barrier matrix in Table 6, comparative cross-study synthesis, and digital procurement maturity perspective.

The matrix reveals that although technologies such as artificial intelligence and blockchain provide substantial improvements in decision-making, transparency, and efficiency, their implementation is hindered by critical challenges including skills shortages, regulatory uncertainty, and system integration issues. Notably, the findings suggest that the benefits of digital procurement are closely linked to organizational readiness and technological infrastructure. This highlights the importance of adopting a holistic approach to digital transformation that balances technological investment with human and institutional capacity development.

5 Conceptual framework

Figure 1 presents a conceptual framework illustrating how Industry 4.0 technologies enable the development of digital procurement capabilities, which in turn influence procurement performance outcomes. The framework further highlights the moderating role of technological, organizational, and regulatory barriers, as well as the progression toward digital procurement maturity. Beyond illustrating the sequential progression of digital procurement transformation, Figure 1 also demonstrates the increasing interconnectivity and convergence of digital technologies within engineering procurement systems. The framework highlights that procurement transformation is not merely a linear technological transition but rather an evolving ecosystem characterized by integration, data exchange, automation, and intelligent decision-making capabilities. Earlier procurement systems were primarily transactional and operational in nature, whereas advanced Procurement 4.0 environments increasingly emphasize predictive analytics, interoperability, real-time visibility, and autonomous coordination across procurement functions. The figure further suggests that digital procurement maturity is influenced not only by technological advancement but also by organizational readiness, integration capability, and the ability to align digital systems with strategic procurement objectives. Consequently, the transformation pathway represented in Figure 1 should be interpreted as a dynamic and interconnected progression rather than a strictly linear sequence of technological adoption.

6 Research methodology

This study adopts a systematic literature review (SLR) approach to examine the role of digital technologies in engineering procurement. The choice of SLR is informed by its ability to provide a structured, transparent, and reproducible synthesis of existing knowledge, particularly in emerging and multidisciplinary fields such as digital procurement and Industry 4.0 (Seyedghorban et al., 2020). The study is guided by the objective of identifying key digital technologies, analysing their impact on procurement performance, and examining the challenges associated with their adoption within engineering contexts.

To ensure methodological rigor, the review follows the principles of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) framework, which provides a standardized procedure for identifying, screening, and selecting relevant literature. The PRISMA approach enhances transparency and minimizes bias by systematically documenting each stage of the review process. A comprehensive search strategy was developed to capture relevant studies across major academic databases, including Scopus, Web of Science, and ScienceDirect. These databases were selected due to their extensive coverage of peer-reviewed publications in engineering management, procurement, and supply chain research. The search strategy employed a combination of keywords and Boolean operators to ensure broad yet focused retrieval of relevant literature. Key search terms included “digital procurement”, “Procurement 4.0”, “engineering procurement”, “AI in procurement”, “blockchain procurement”, “BIM procurement”, and “digital supply chain technologies”. Boolean operators such as AND and OR were used to combine search terms and expand the scope of the search. To reflect the rapid development of digital technologies, the search was limited to peer-reviewed journal articles published between 2015 and 2024.

The study selection process followed a multi-stage procedure consistent with PRISMA guidelines. Initially, all retrieved records were screened to remove duplicates. This was followed by a preliminary assessment of titles and abstracts to exclude studies that were not directly related to procurement or digital technologies. The remaining articles were subjected to full-text evaluation to determine their eligibility based on predefined inclusion and exclusion criteria. Studies were included if they addressed digital technologies within procurement or supply chain contexts, focused on engineering-related sectors such as construction and manufacturing, and provided empirical, conceptual, or systematic analyses. Articles lacking methodological clarity or substantive discussion of procurement processes were excluded. To further

TABLE 3 Comprehensive literature review table on digital technologies in engineering procurement.

Author(s)	Study type	Technology focus	Context/ Sector	Key findings	Contribution to this Study
Bienhaus and Haddud (2018)	Conceptual	Procurement 4.0	Supply chain	Identified key drivers and barriers of procurement digitalization	Provides theoretical foundation for digital procurement evolution
Klündera et al., 2019)	Empirical	Industry 4.0 technologies	Manufacturing	Found strong correlation between digitalization and cost reduction	Supports economic benefits of digital procurement
Seyedghorban et al. (2020)	Survey	Digital procurement capability	Supply chain	Digital capabilities improve decision-making and performance	Links digital maturity with procurement efficiency
Nicoletti (2020)	Empirical	Procurement 4.0	Supply chain	Digital procurement improves operational and supply chain performance	Validates performance impact
Ivanov and Dolgui (2020)	Analytical	Digital supply chains	Supply chain	Digitalization enhances supply chain resilience	Supports resilience argument
Hallikas et al. (2021)	Case study	Data-driven procurement	Industrial procurement	Analytics improves supplier risk management	Shows role of data in procurement decisions
Bigliardi et al. (2022)	Literature review	Industry 4.0	Supply chain	Digital technologies reshape supply chain structures	Provides broader context
Karttunen et al. (2023)	Empirical	Digital procurement tools	Supply chain	Improves agility and visibility	Supports operational benefits
Senaratne et al. (2023)	Empirical	E-procurement	Firms (Denmark)	Improves efficiency, transparency, and collaboration	Strong real-world validation
Motaung and Sifolo (2023)	Case study	Digital procurement systems	South Africa (airport)	Identified barriers: skills gap, cybersecurity, interoperability	Supports barrier analysis
Althabatah et al. (2023)	Systematic review	Industry 4.0 technologies	Procurement	AI, blockchain, e-procurement dominate literature	Supports technology dominance
Matos et al. (2024)	Mixed (review + Framework)	BIM	Construction	BIM improves coordination and procurement integration	Supports engineering procurement linkage
Ibem et al., 2018)	Survey	Digital technologies	Construction (Nigeria)	Adoption uneven; influenced by stakeholders	Supports regional context
De Rejeb et al. (2018)	Conceptual	IoT, Blockchain, robotics	Logistics	Technologies enhance automation and predictability	Supports technology classification
Dubolazov et al. (2020)	Analytical	Digital logistics systems	Industrial enterprises	Digitalization changes procurement structure and functions	Supports organizational transformation
Mashabane and Van der Lingen (2024)	Review	Digital procurement technologies	South Africa	AI, blockchain, big data most impactful	Reinforces key technology trends
Chopra (2018)	Conceptual	IT tools in procurement	Supply chain	IT improves cost efficiency and supplier management	Supports early digital tools
Karmenhag and Löfnertz (2018)	Thesis (qualitative)	AI, IoT, Blockchain	Industrial procurement	Digitalization may automate procurement functions	Supports future direction
Ibem and Laryea (2014)	Literature review	Digital tools	Construction	Identified 36 procurement technologies	Supports historical evolution

enhance methodological rigor and transparency, the study incorporated a structured quality assessment procedure for all eligible studies included in the review. Following the full-text screening stage, the methodological quality and relevance of each study were critically appraised using predefined evaluation criteria adapted from established systematic review practices in management and engineering research. The appraisal focused on four key dimensions: methodological clarity, relevance to engineering procurement, adequacy of data or conceptual discussion, and contribution to digital procurement knowledge.

TABLE 4 Gap analysis of digital technologies in engineering procurement literature.

Study/ Author(s)	Focus area	Key contributions	Identified limitations	Research gap	How this Study addresses it
Bienhaus and Haddud (2018)	Procurement 4.0	Identified drivers and barriers of digital procurement	Conceptual; lacks empirical validation	Limited empirical evidence on real-world implementation	Integrates empirical and conceptual studies to validate findings
Klünder et al. (2019)	Digitalization and cost	Demonstrated cost reduction through digital technologies	Focused mainly on cost dimension	Ignores broader performance impacts (e.g., resilience, transparency)	Expands analysis to multi-dimensional procurement performance
Seyedghorban et al. (2020)	Digital capability	Established link between digital capability and performance	General supply chain focus	Limited application to engineering procurement context	Focuses specifically on engineering procurement systems
Nicoletti (2020)	Procurement 4.0 performance	Showed improvement in supply chain performance	Limited sector-specific analysis	Lack of engineering-specific insights	Provides sector-specific synthesis (engineering/ construction)
Ivanov and Dolgui (2020)	Supply chain resilience	Highlighted role of digitalization in resilience	Focus on supply chain, not procurement	Procurement-specific mechanisms not explored	Links resilience directly to procurement technologies
Hallikas et al. (2021)	Data-driven procurement	Demonstrated value of analytics in risk management	Case-based; limited generalization	Lack of integration with other technologies	Synthesizes multiple technologies (AI, IoT, blockchain)
Bigliardi et al. (2022)	Industry 4.0 supply chains	Reviewed impact of digital technologies broadly	Broad scope; lacks procurement focus	Procurement function underexplored	Focuses specifically on procurement transformation
Althabatah et al. (2023)	Procurement 4.0 technologies	Identified dominant technologies (AI, blockchain)	Technology-centric analysis	Lack of integration across technologies	Provides integrated ecosystem perspective
Senaratne et al. (2023)	E-procurement	Empirical evidence of efficiency gains	Focus limited to e-procurement	Ignores advanced technologies integration	Expands scope to multiple digital technologies
Motaung and Sifolo (2023)	Adoption barriers	Identified skills, cybersecurity, interoperability issues	Case-specific (South Africa)	Limited cross-context generalization	Provides cross-study synthesis of barriers
Ibem et al. (2018)	Construction procurement	Identified adoption levels and influencing factors	Regional focus (Nigeria)	Lack of global comparative analysis	Integrates global and regional perspectives
Matos et al. (2024)	BIM integration	Developed BIM-based procurement framework	Focus on single technology (BIM)	Lack of multi-technology integration	Combines BIM with AI, IoT, blockchain insights
De Rejeb et al. (2018)	Emerging technologies	Identified role of IoT, blockchain, robotics	Conceptual; lacks empirical depth	Limited validation in procurement systems	Includes empirical and applied studies
Dubolazov et al. (2020)	Digital logistics systems	Showed transformation of procurement structures	Focus on logistics integration	Procurement-specific frameworks limited	Proposes procurement-focused maturity framework
Mashabane and Van der Lingen (2024)	Technology selection	Identified dominant technologies in procurement	Lacks integrated framework	No maturity or adoption pathway	Introduces digital procurement maturity framework

Studies that lacked sufficient methodological transparency, presented weak conceptual grounding, or demonstrated limited relevance to procurement processes were excluded from the final synthesis. This quality assessment process ensured that only robust and credible studies informed the analysis and interpretation of findings.

In addition, the review process adhered closely to the PRISMA 2020 guidelines to ensure systematic identification, screening, eligibility assessment, and inclusion of studies. The search and screening procedures were explicitly documented to improve reproducibility and reduce selection bias. A PRISMA checklist was also incorporated as supplementary material to further enhance methodological transparency and compliance with systematic review reporting standards. Furthermore, the study adopted a comparative and cross-study synthesis approach rather than a purely descriptive narrative review. This enabled the identification of converging and diverging patterns across studies, critical comparison of technological capabilities, and deeper interpretation of relationships between digital technologies, procurement capabilities, and procurement performance

TABLE 5 Comparative positioning of this study against previous reviews.

Study	Primary focus	Scope limitation	Distinct contribution of this Study
Bienhaus and Haddud (2018)	Procurement 4.0 drivers and barriers	Primarily conceptual and supply-chain focused	Provides integrated engineering procurement perspective and cross-technology synthesis
Althabatah et al. (2023)	Industry 4.0 technologies in procurement	Technology-centric review	Develops integrated ecosystem and maturity perspective
Bigliardi et al. (2022)	Industry 4.0 in supply chains	Broad supply chain orientation	Focuses specifically on engineering procurement transformation
Seyedghorban et al. (2020)	Digital procurement capabilities	General procurement maturity focus	Extends analysis to technology complementarity and procurement ecosystems
Motaung and Sifolo (2023)	Digital procurement barriers	Context-specific case study	Provides broader comparative synthesis across technologies and contexts
Matos et al. (2024)	BIM-based procurement integration	Single-technology orientation	Integrates BIM with AI, IoT, blockchain, and procurement 4.0 systems
This study	Integrated digital technologies in engineering procurement	Multi-dimensional systematic synthesis	Develops conceptual framework, comparative synthesis, technology–impact–barrier matrix, and digital procurement maturity perspective

TABLE 6 Technology–impact–barrier matrix for digital procurement.

Technology	Key procurement impacts	Supporting evidence	Major barriers	Supporting evidence
E-procurement systems	Improved efficiency, reduced cycle time, enhanced transparency	Senaratne et al. (2023)	Resistance to adoption, system integration issues	Motaung and Sifolo (2023)
Artificial Intelligence (AI)	Predictive analytics, supplier evaluation, demand forecasting	Seyedghorban et al. (2020)	Skills gap, data dependency, high implementation cost	Motaung and Sifolo (2023)
Blockchain	Transparency, traceability, trust, fraud reduction	Kouhizadeh and Sarkis (2018)	Regulatory uncertainty, legal issues, scalability	De Rejeb et al. (2018)
Internet of things (IoT)	Real-time monitoring, supply chain visibility, resilience	De Rejeb et al. (2018)	Infrastructure requirements, data security concerns	Motaung and Sifolo (2023)
Building Information modelling (BIM)	Integration of design and procurement, improved coordination, reduced delays	Matos et al. (2024)	Lack of standards, interoperability challenges	Matos et al. (2024)
Robotic process automation (RPA)	Automation of repetitive tasks, error reduction, process efficiency	De Rejeb et al. (2018)	Workforce resistance, limited flexibility	Motaung and Sifolo (2023)
Integrated procurement 4.0 systems	Cost reduction, strategic decision-making, supply chain resilience	Klünder et al. (2019); Althabatah et al. (2023)	System complexity, integration challenges, organizational readiness	Bienhaus and Haddud (2018)

outcomes. Through this approach, the review extends beyond summarizing prior studies by providing analytical synthesis and theoretical integration across the reviewed literature.

The inclusion and exclusion criteria presented in Table 7 were established to ensure consistency, relevance, and methodological rigor during the study selection process. These criteria enabled the systematic filtering of studies to ensure that only high-quality and contextually relevant literature informed the final synthesis. Following the selection process, relevant data were systematically extracted from the included studies. The extracted information included authorship, publication year, research methodology, technology focus, sectoral context, and key findings. The data were organized into structured synthesis tables to facilitate comparison and integration across studies. A thematic analysis approach was then employed to identify recurring patterns and relationships within the literature. This analysis was structured around three core dimensions: the conceptual evolution of digital procurement, the role and impact of digital technologies, and the barriers to their adoption.

To enhance analytical depth, the study adopted a cross-study synthesis approach, integrating findings from empirical, conceptual, and review-based studies. This enabled the identification of

10.3389/fbuil.2026.1870384

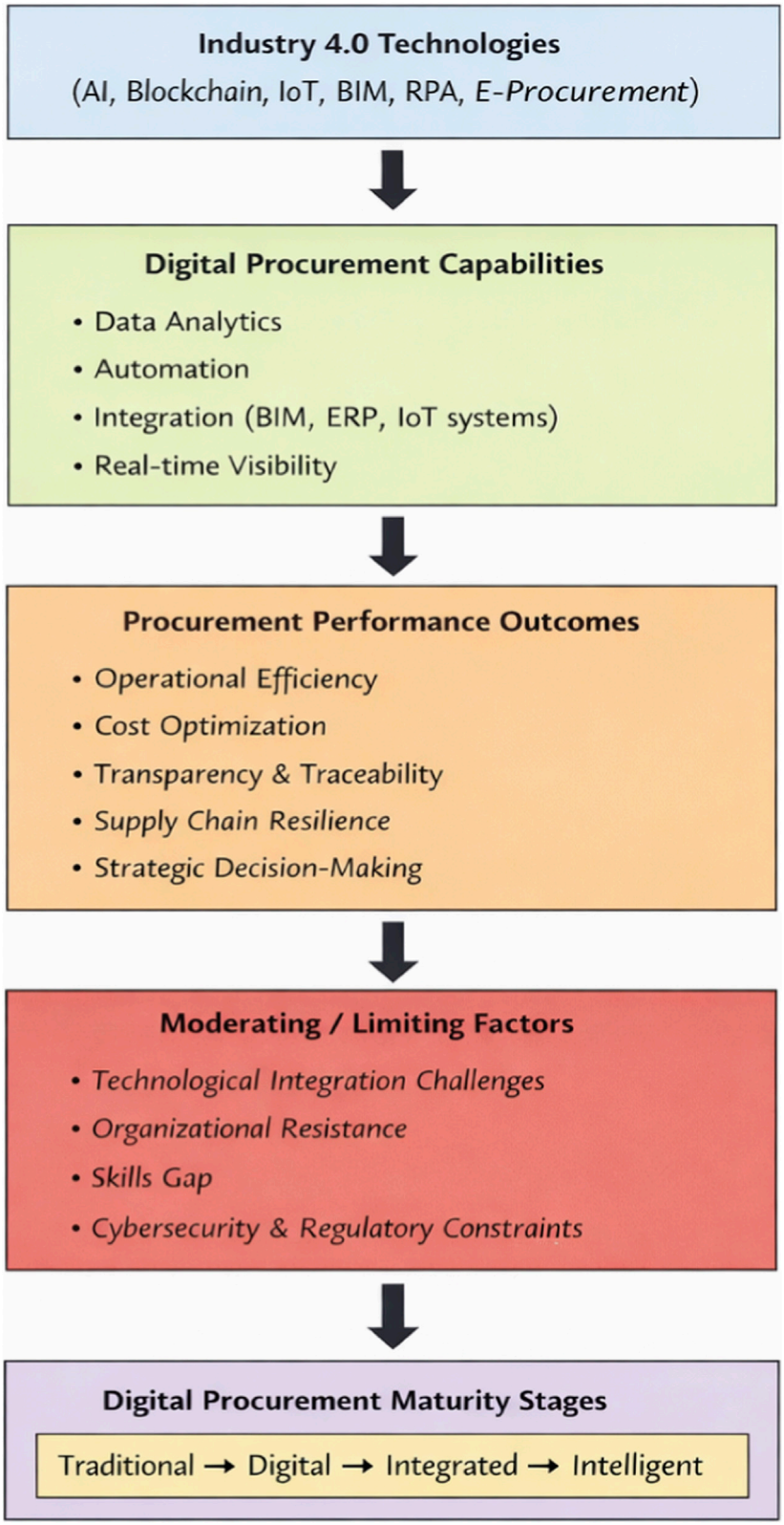


FIGURE 1
Conceptual framework for digital technologies in engineering procurement.

TABLE 7 Inclusion and exclusion criteria for study selection.

Criteria	Inclusion Criteria	Exclusion Criteria
Publication Type	Peer-reviewed journal articles and review papers	Conference abstracts, editorials, dissertations, non-peer-reviewed sources
Language	Studies published in English	Non-English publications
Publication period	Studies published between 2015 and 2024	Studies published before 2015
Subject Focus	Studies addressing digital technologies in procurement, supply chains, or engineering procurement	Studies unrelated to procurement or digital transformation
Sectoral context	Engineering, construction, manufacturing, and supply chain sectors	Studies outside relevant industrial or engineering contexts
Content relevance	Studies discussing procurement technologies, adoption, impacts, barriers, or transformation	Studies lacking substantive discussion of procurement processes
Methodological Quality	Studies with clear methodology, conceptual rigor, or empirical evidence	Studies lacking methodological clarity or analytical depth

converging evidence across different contexts, as well as areas of divergence and inconsistency. The use of multiple data sources and methodological triangulation strengthens the validity of the findings, while the consistency of observed patterns across studies provides internal reliability. Furthermore, the inclusion of diverse study types, including surveys, case studies, and systematic reviews, enhances the generalizability of the conclusions. The reliability and reproducibility of the study are ensured through the explicit definition of the search strategy, inclusion criteria, and analytical procedures. By adhering to established systematic review protocols and documenting each stage of the review process, the study minimizes potential bias and enhances transparency. In addition, all sources are drawn from peer-reviewed literature and are appropriately cited, ensuring compliance with academic integrity and ethical research standards. Overall, the adopted methodology provides a robust and comprehensive framework for analysing digital technologies in engineering procurement. It ensures that the findings are grounded in a systematic and reproducible process, thereby supporting the credibility and scholarly contribution of the study. To ensure transparency and methodological rigor, the study selection process is illustrated in Figure 2.

The research selection procedure, as depicted in Figure 2, adhered to the PRISMA 2020 paradigm, facilitating a systematic and transparent identification, screening, eligibility evaluation, and incorporation of pertinent studies. The procedure commenced with an exhaustive database search across Scopus, Web of Science, and ScienceDirect, subsequently followed by the elimination of duplicate entries. Subsequently, titles and abstracts were evaluated to exclude irrelevant studies, after which full-text papers were examined against established inclusion and exclusion criteria. This meticulous multi-phase process guaranteed the inclusion of only pertinent and high-caliber research in the final synthesis, hence augmenting the reliability and validity of the review outcomes.

7 Discussion

The findings of this study reveal a clear and consistent shift in procurement from a traditionally transactional function toward a strategic, technology-driven capability within engineering environments. Across the reviewed literature, digital technologies particularly artificial intelligence (AI), blockchain, the Internet of Things (IoT), Building Information Modelling (BIM), and e-procurement systems emerge as key enablers of procurement transformation. This transformation is not merely incremental but represents a structural reconfiguration of procurement processes, decision-making mechanisms, and inter-organizational relationships. A central insight from the analysis is that digital technologies do not operate in isolation; rather, their value is realized through integration and complementarity. While individual technologies such as AI or blockchain offer specific functional advantages, the literature consistently indicates that the most significant improvements in procurement performance occur when these technologies are combined within integrated digital ecosystems. For example, the integration of BIM with procurement systems enhances coordination between design and sourcing activities, while the combination of IoT and AI enables real-time monitoring and predictive decision-making (Matos et al., 2024; Bigliardi et al., 2020; De Rejeb et al., 2018). This supports the argument that digital procurement should be conceptualized as a system-level transformation rather than a technology-specific intervention.

A more critical interpretation of the reviewed literature further reveals important differences in the maturity, applicability, and strategic value of digital procurement technologies. The findings suggest that technologies such as e-procurement systems and BIM demonstrate relatively higher levels of maturity because they have evolved over longer implementation periods and are supported by more established operational standards and industry acceptance. In contrast, technologies such as blockchain and generative artificial intelligence remain comparatively less mature within procurement environments due to regulatory uncertainty, integration complexity, scalability concerns, and evolving governance structures. The review also indicates that the procurement value delivered by digital technologies varies depending on organizational context and environmental uncertainty. Under highly uncertain and disruption-prone conditions, technologies such as artificial intelligence and IoT appear to provide greater strategic value

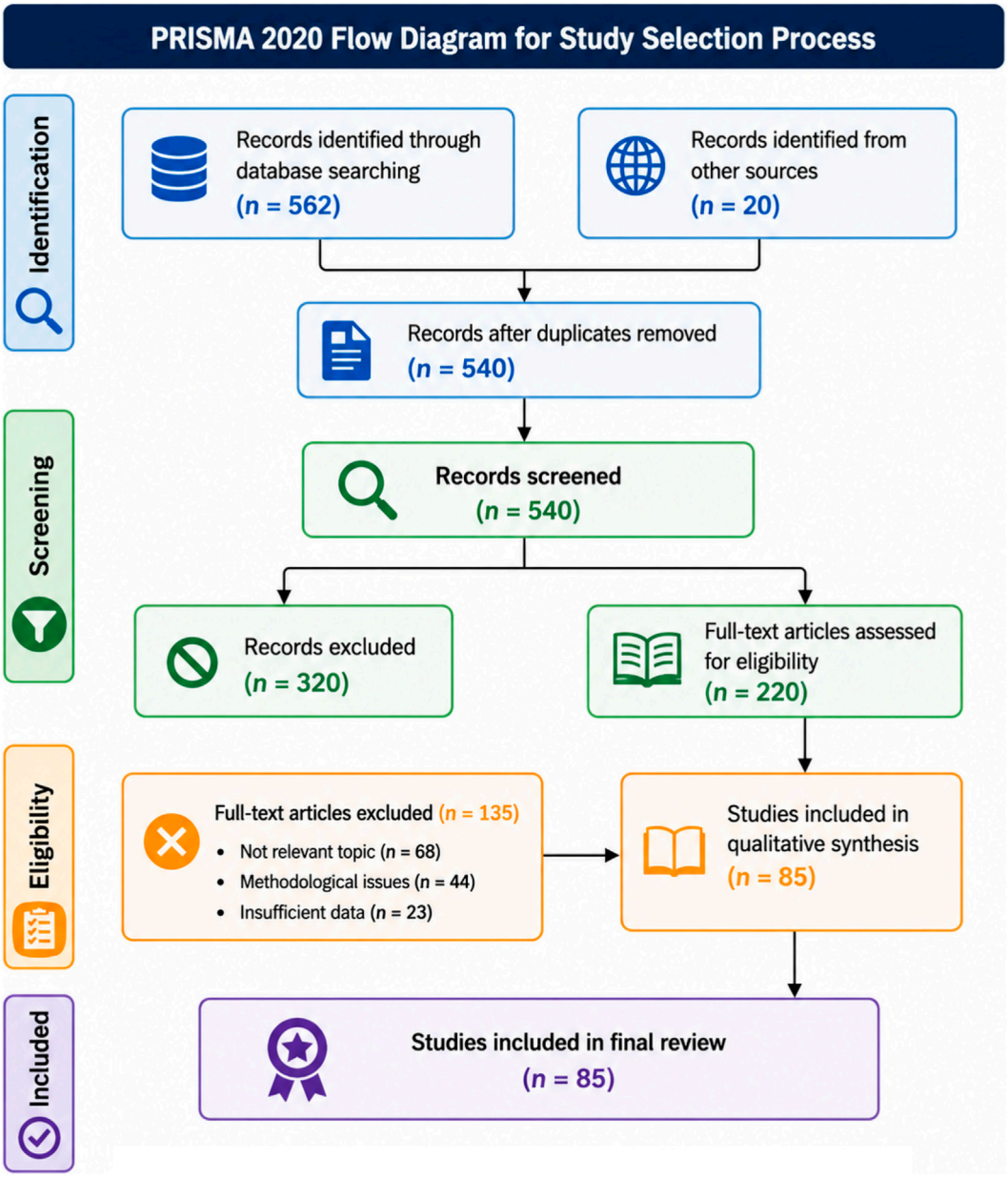


FIGURE 2
Prisma 2020 flow diagram for study selection process.

because of their predictive, real-time monitoring, and data-driven decision-making capabilities. Conversely, blockchain technologies demonstrate stronger value in contexts requiring transparency, traceability, trust, and transaction security.

In addition, the analysis highlights that many digital procurement technologies are complementary rather than substitutive. For example, AI capabilities are significantly enhanced when integrated with IoT-generated real-time data, while BIM-based procurement systems become more effective when combined with e-procurement platforms and analytics tools. This suggests that procurement transformation is driven less by isolated technological adoption and more by the development of integrated digital ecosystems. Consequently, organizations that adopt interconnected digital procurement architectures are more likely to achieve higher levels of procurement maturity, operational resilience, and strategic performance.

The evidence further demonstrates that digital procurement capabilities significantly influence multiple dimensions of procurement performance. Consistent with prior studies, improvements are observed in operational efficiency, cost optimization, transparency, and supply chain resilience (Senaratne et al., 2023; Klünder et al., 2019). Notably, the impact on decision-making quality emerges as a critical outcome, driven by the ability of digital technologies to process large volumes of data and generate actionable insights. This finding aligns with the

broader literature on data-driven supply chains, which emphasizes the role of analytics in enhancing organizational responsiveness and competitiveness. However, the analysis also reveals that the benefits of digital procurement are not uniformly realized across contexts. A recurring theme in the literature is the presence of significant implementation barriers, which mediate the relationship between digital technologies and procurement performance. Technological challenges, particularly system interoperability and integration with legacy systems, remain a major constraint (Bienhaus and Haddud, 2018). At the organizational level, resistance to change and skills shortages limit the effective adoption of digital tools, especially in developing regions (Motaung and Sifolo, 2023). These findings suggest that digital transformation in procurement is not solely a technological issue but also a socio-technical challenge requiring alignment between technology, people, and organizational processes.

The Technology–Impact–Barrier matrix developed in this study further highlights the dual nature of digital technologies, illustrating how each technology simultaneously introduces both opportunities and constraints. For instance, while blockchain enhances transparency and trust, its adoption is hindered by regulatory uncertainty and scalability concerns. Similarly, AI enables predictive analytics but requires significant data infrastructure and specialized expertise. This reinforces the need for a balanced and context-sensitive approach to digital procurement implementation.

Another important insight relates to the uneven adoption of digital procurement technologies across regions and sectors. While developed economies demonstrate higher levels of digital maturity, developing regions continue to face infrastructural and institutional limitations that constrain adoption (Ibem et al., 2018; Awodele et al., 2025). This disparity underscores the importance of contextual factors in shaping digital transformation outcomes and suggests that a "one-size-fits-all" approach to digital procurement may be ineffective. The proposed conceptual framework and maturity perspective provide a structured lens through which these findings can be interpreted. The progression from traditional procurement systems to intelligent, data-driven ecosystems reflects not only technological advancement but also organizational learning and capability development. The literature suggests that organizations must first establish foundational digital systems before advancing toward integrated and autonomous procurement models. This staged progression highlights the importance of digital maturity as a critical determinant of successful transformation.

From a theoretical perspective, this study contributes to the growing body of knowledge on Procurement 4.0 by extending existing frameworks to explicitly incorporate engineering procurement contexts. Unlike general supply chain studies, this research emphasizes the complexity of engineering procurement, including its integration with design processes, technical specifications, and project-based workflows. From a practical perspective, the findings provide actionable insights for organizations seeking to implement digital procurement systems, particularly in identifying key technologies, anticipating barriers, and adopting a phased transformation strategy. Despite these contributions, the study also identifies areas for future research. The majority of existing studies remain either conceptual or limited in empirical scope, highlighting the need for more large-scale empirical investigations to validate the long-term impacts of digital procurement. In addition, emerging technologies such as digital twins and generative artificial intelligence remain underexplored within procurement contexts, presenting opportunities for further research. Finally, there is a need to examine the interplay between digital procurement and broader organizational strategies, including sustainability and circular economy initiatives. In summary, the findings demonstrate that digital technologies have the potential to fundamentally transform engineering procurement, but their successful implementation depends on the integration of technologies, organizational readiness, and contextual alignment. This reinforces the importance of adopting a holistic and strategic approach to digital procurement transformation.

8 Conclusion

This study examined the role of digital technologies in engineering procurement through a systematic and comprehensive synthesis of existing literature. The findings demonstrate that technologies such as artificial intelligence, blockchain, the Internet of Things, Building Information Modelling, and e-procurement systems are fundamentally transforming procurement from a transactional function into a strategic, data-driven capability. These technologies collectively enhance procurement performance by improving operational efficiency, cost optimization, transparency, and supply chain resilience. A key contribution of this study lies in its integrated perspective, which moves beyond the isolated analysis of individual technologies to emphasize the importance of technology complementarity and system integration. The findings indicate that the value of digital procurement is maximized when technologies are implemented within interconnected ecosystems supported by data analytics and real-time information flows. At the same time, the study highlights critical barriers including technological integration challenges, organizational resistance, skills gaps, and regulatory uncertainties that must be addressed to fully realize the benefits of digital transformation. This underscores the need to conceptualize digital procurement as a socio-technical transformation involving both technological advancement and organizational adaptation. The study further contributes by proposing a conceptual framework and a digital procurement maturity perspective that provide a structured pathway for transitioning from traditional procurement systems to intelligent, data-driven ecosystems. These contributions extend existing Procurement 4.0 literature by explicitly situating digital transformation within the complex and multidisciplinary context of engineering procurement. From a practical perspective, the findings offer actionable insights for practitioners and policymakers by identifying key technologies, potential challenges, and strategic pathways for implementation. Organizations are encouraged to adopt a phased and integrated approach to digital transformation, supported by investments in skills development, system interoperability, and governance mechanisms. Overall, digital technologies present a transformative opportunity for engineering procurement; however, their successful adoption depends on the alignment of technological capabilities, organizational readiness, and contextual factors. By providing a comprehensive synthesis and structured framework, this study advances both scholarly understanding and practical application of digital procurement in engineering systems.

Future research directions and study limitations are discussed in the subsequent section.

9 Limitations of the study and areas for future research

Despite its contributions, this study is subject to several limitations that should be acknowledged. First, the study is based entirely on secondary data derived from published literature. While the systematic literature review approach ensures methodological rigor and breadth of coverage, it inherently depends on the availability, scope, and quality of existing studies. Consequently, the findings may reflect prevailing research trends and biases within the literature rather than fully capturing the diversity of real-world procurement practices. Second, the body of literature on digital procurement remains unevenly distributed across regions and sectors. A significant proportion of studies are concentrated in developed economies, with relatively limited empirical evidence from developing and emerging contexts. This geographical imbalance may affect the generalizability of the findings, particularly in relation to infrastructural constraints, institutional environments, and technological readiness in different regions. Third, although this study provides an integrated synthesis of multiple digital technologies, the analysis is primarily conceptual and comparative in nature. The absence of primary empirical validation limits the ability to quantitatively assess the magnitude of the relationships identified, particularly with respect to the impact of digital technologies on procurement performance outcomes. In addition, the rapid evolution of digital technologies implies that some findings may become outdated as new tools and applications emerge. Given these limitations, several avenues for future research are identified. There is a need for large-scale empirical studies, including surveys, longitudinal analyses, and case studies, to validate the proposed conceptual framework and digital procurement maturity perspective. Such studies would provide deeper insights into causal relationships and enable the quantification of performance impacts across different contexts. Future research should also explore the role of emerging technologies, such as digital twins, generative artificial intelligence, and advanced analytics, which remain underexplored within procurement literature. In particular, the integration of these technologies with existing procurement systems presents opportunities for developing more intelligent and autonomous procurement models. Furthermore, additional research is required to examine the long-term implications of digital procurement, particularly in relation to sustainability, circular economy practices, and supply chain resilience. Investigating how digital technologies influence environmental performance, resource efficiency, and risk management would provide valuable insights for both academia and practice. Finally, there is a need for more context-specific and comparative studies that examine digital procurement adoption across different regions, industries, and organizational scales. Such research would enhance the generalizability of findings and support the development of tailored strategies for digital transformation in diverse engineering procurement environments.

Author contributions

IA: Methodology, Project administration, Writing – original draft, Writing – review and editing. MR: Supervision, Writing – review and editing. NO: Methodology, Supervision, Writing – review and editing. EE: Formal Analysis, Investigation, Methodology, Writing – review and editing. IA: Supervision, Writing – review and editing. YE: Methodology, Writing – original draft. TK: Formal Analysis, Methodology, Visualization, Writing – original draft. BA: Resources, Visualization, Writing – original draft. AO: Formal Analysis, Methodology, Software, Writing – original draft.

Funding

The author(s) declared that financial support was received for this work and/or its publication. Funding was obtained from Durban University of Technology in accordance with the Post-Doctoral Fellowship agreement.

Conflict of interest

The author(s) declared that this work was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declared that generative AI was used in the creation of this manuscript. During the preparation of this work, the author(s) used (OpenAI, GPT-5.1, 2025 version) to rephrase, check grammar and spelling. After using this tool, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the publication. All conceptual frameworks, interpretations, and ultimate intellectual contributions were executed by the human authors. In accordance with publishing regulations, the AI system is not designated as an author and assumes no responsibility for any portion of the manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Aduwo, E. B. (2018). A survey of digital technologies used in the procurement of building projects: empirical evidence from Nigeria. *Asian J. Sci. Res.* 11 (4). doi:10.3923/ajsr.2018.456.465

Althabatah, A., Yaqot, M., Menezes, B., and Kerbache, L. (2023). Transformative procurement trends: integrating industry 4.0 technologies for enhanced procurement processes. *Logistics* 7 (3), 63. doi:10.3390/logistics7030063

Awodele, I. A., Mewomo, M. C., Municio, A. M. G., Chan, A. P. C., Darko, A., Taiwo, R., et al. (2024). Awareness, adoption readiness and challenges of railway 4.0 technologies in a developing economy. *Heliyon* 10 (4), e25934. doi:10.1016/j.heliyon.2024.e25934

Awodele, I. A., Mewomo, M. C., Chan, A. P. C., Darko, A., Municio, A. M. G., Olatunde, N. A., et al. (2025). "Benefits and drivers of implementing sustainable procurement in a developing economy,". *Creating Capacity and Capability: Embracing Advanced Technologies and Innovations for Sustainable Future in Building Education and Practice. Lecture Notes in Civil Engineering*. Editors M. Sutrisna, M. B. Jelodar, N. Domingo, A. Le, and R. Kahandawa (Singapore: Springer), 563, 359–371. doi:10.1007/978-981-96-2904-6_28

Bigliardi, B., Filippelli, S., Petroni, A., and Tagliente, L. (2022). The digitalization of supply chain: a review. *Procedia Comput. Sci.* 200, 1806–1815.

Bienhaus, F., and Haddud, A. (2018). Procurement 4.0: factors influencing the digitisation of procurement and supply chains. *Bus. Process Manag. J.* 24 (4), 965–984. doi:10.1108/bpmj-06-2017-0139

Bigliardi, B., Bottani, E., and Casella, G. (2020). Enabling technologies, application areas and impact of industry 4.0: a bibliographic analysis. *Procedia Manufacturing* 42, 322–326. doi:10.1016/j.promfg.2020.02.086

Chopra, A. (2018). "Technology in procurement and supply as prevalent today and scope for future," in Proceedings of the 2018 International Conference on Automation and Computational Engineering (ICACE).

De Rejeb, A., Sűle, E., and Keogh, J. G. (2018). Exploring new technologies in procurement. *Transp. and Logist. Int. J.* 18 (45), 76–86.

Dubolazov, V., Simakova, Z., and Iusma, I. (2019). "The impact of digital technologies on procurement infrastructure and functions at industrial enterprises," in International Conference on Digital Technologies in Logistics and Infrastructure (ICDTLI 2019), 241–245.

Dubolazov, V., Simakova, Z., Leicht, O., and Shchelkonogov, A. (2020). "The impact of digitalization on a production structures and management in industrial enterprises and complexes," in *Technological Transformation: A New Role For Human, Machines And Management. TT 2020. Lecture Notes in Networks and Systems*. Editors H. Schaumburg, V. Korablev, and L. Ungvari (Cham: Springer), 157. doi:10.1007/978-3-030-64430-7_4

Hallikas, J., Immonen, M., and Brax, S. (2021). Digitalizing procurement: the impact of data analytics on supply chain performance. *Supply Chain Manag. Int. J.*, Vol. 26 No. 5 pp. 629–646. doi:10.1108/SCM-05-2020-0201

Ibem, E. O., and Laryea, S. (2014). Survey of digital technologies in procurement of construction projects. *Automation Constr.* 46, 11–21. doi:10.1016/j.autcon.2014.07.003

Ibem, E. O., Aduwo, E. B., Afolabi, A. O., Oluwunmi, A. O., Tunji-Olayeni, P. F., Ayo-Vaughan, E. A., et al. (2018). Electronic (e-) procurement adoption and users' experience in the Nigerian construction sector. *Int. J. Constr. Educ. Res.*, 17 (3), 258–276. doi:10.1080/15578771.2020.1730527

Ivanov, D., and Dolgui, A. (2020). Viability of supply chain systems in the context of COVID-19. *Int. J. Prod. Res.* 58 (10), 2904–2915. doi:10.1080/00207543.2020.1750727

Karmenhag, J., and Löfnertz, E. (2018). *How Digital Technologies Will Impact the Procurement Process and Organisation*. Sweden: Chalmers University of Technology. Master's thesis.

Karttunen, E., Pynnönen, M., Treves, L., and Hallikas, J. (2023). Capabilities for the internet of things enabled product-service system business models. *Technol. Anal. Strateg. Mgmt.* 35 (12), 1618–1634.

Klünder, T., Dörseln, J. N., and Steven, M. (2019). Procurement 4.0: how digital disruption supports cost reduction in procurement. *Production* 29, e20180104. doi:10.1590/0103-6513.20180104

Kouhizadeh, M., and Sarkis, J. (2018). Blockchain practices, potentials, and perspectives in greening supply chains. *Sustainability* 10 (10), 3652. doi:10.3390/su10103652

Mashabane, S. T., and Van der Lingen, E. (2024). "The role of digital technologies in procurement," in *Proceedings of the SAIIE Conference*. South Africa.

Matos, B. C., Cruz, C. O., and Branco, F. B. (2024). Digitalization and procurement in construction projects: an integrated BIM-Based approach. *J. Inf. Technol. Constr.* 29, 400–423. doi:10.36680/j.itcon.2024.019

Motaung, J. R., and Sifolo, P. P. S. (2023). Benefits and barriers of digital procurement: lessons from an airport company. *Sustainability* 15, 4610. doi:10.3390/su15054610

Nicoletti, B. (2020). "Procurement 4.0 and the fourth industrial revolution," in *The Opportunities and Challenges of a Digital World*. Switzerland Palgrave Macmillan. doi:10.1007/978-3-030-35979-9

Olatunde, N. A., Gento, A. M., Okorie, V. N., Oyewo, O. W., Mewomo, M. C., and Awodele, I. A. (2023). Construction 4.0 technologies in a developing economy: awareness, adoption readiness and challenges. *Front. Eng. Built Environ.* 3 (2), 108–121. doi:10.1108/FEBE-08-2022-0037

Senaratne, L. D., Kaufmann, B. A., and Haller, L. S. (2023). Role of technology in enhancing procurement processes and efficiency in firms in odense, Denmark. *J. Procure. Supply Chain Manag.* 7 (1), 12–21. doi:10.53819/81018102t5175

Seyedghorban, Z., Samson, D., and Tahernejad, H. (2020). Digitalization opportunities for the procurement function: pathways to maturity. *Int. J. Operations and Prod. Manag.* 40 (11), 1685–1693. doi:10.1108/ijopm-04-2020-0214

frontiers in Built Environment

TYPE Original Research
PUBLISHED 23 July 2026
DOI 10.3389/fbuil.2026.1839808

Check for updates

OPEN ACCESS

EDITED BY
Salman Azhar,
Auburn University, United States

REVIEWED BY
Chaojie Zhang,
Anhui University of Technology, China
Manxu Zhou,
Xinjiang Agricultural University, China

*CORRESPONDENCE
Syed Hadi Hussain Shah,
hadihussain010@gmail.com

RECEIVED 26 March 2026
REVISED 17 May 2026
ACCEPTED 18 May 2026
PUBLISHED 23 July 2026

CITATION
Ying D, Guanghua L, He Z, Hasan R and Shah SHH (2026) An efficient deep learning framework for text detection and recognition in engineering drawings. *Front. Built Environ.* 12:1839808. doi: 10.3389/fbuil.2026.1839808

COPYRIGHT
© 2026 Ying, Guanghua, He, Hasan and Shah. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

An efficient deep learning framework for text detection and recognition in engineering drawings

Dakan Ying[1], Lyu Guanghua[1], Zhengxu He[1], Rija Hasan[2] and Syed Hadi Hussain Shah[1]*

[1]Power China Huadong Engineering Corporation Limited, Hangzhou, China, [2]Mohammad Ali Jinnah University, Karachi, Pakistan

The efficient digitization of engineering drawings is critical for automating construction workflows, including quantity takeoff, procurement, fabrication, inspection, BIM modeling, digital twin creation, and lifecycle information management. Manual extraction of steel reinforcement annotations is labor-intensive, error-prone, and impedes digital project delivery. To address this, we propose a deep learning-based text detection and recognition framework that accurately extracts rebar annotations from complex construction drawings. Our detection network uses an enhanced Differentiable Binarization Network (DBNet) with reduced computational complexity (3.8M parameters vs. 12.3M for standard DBNet detection pipeline), representing a 69% reduction in detection parameters. Combined with a Transformer-based recognition model (15.56M parameters), the system achieves high detection (94.2%) and recognition (96.3%) accuracy, even under challenging conditions such as noise, occlusion, and varied styles. By enabling automatic, accurate conversion of technical drawings into structured digital information, our method significantly advances construction automation. This contribution supports the broader goals of reducing manual processes, enhancing information reliability, and accelerating digital transformation in construction projects.

KEYWORDS
automated quality inspection, digital twin development, engineering drawings digitization, optical character recognition, steel rebar annotation digitization

1 Introduction

In line with the growing emphasis on digital transformation in the construction industry, the automation of engineering drawing interpretation—particularly steel bar sample annotations—represents a crucial step toward improving data integration, reducing human workload, and enhancing lifecycle efficiency of infrastructure projects. The extraction of structured information from unstructured drawing content directly supports the development of digital twins, Building Information Modeling (BIM), and intelligent construction management systems, which are central themes in contemporary construction informatics. Automated rebar annotation extraction plays a vital role across multiple construction workflows. In quantity takeoff, accurate annotation extraction enables precise material estimation, reducing waste and over-ordering. For procurement, structured digital data facilitates automated purchasing systems. In fabrication, rebar bending schedules derived from annotations directly feed into CNC machines. During inspection, digital annotations enable automated compliance verification against design

specifications. For BIM and digital twins, structured annotation data provides the semantic layer necessary for lifecycle information management, from design through maintenance to eventual decommissioning. Currently, rebar annotations are typically extracted manually by engineers or technicians. This process is slow—a single complex drawing can require 30–60 min of manual effort. It is error-prone, with human transcription errors ranging from 5%–15% depending on drawing complexity and fatigue levels. Furthermore, it does not scale: large infrastructure projects with thousands of drawings require hundreds of person-hours, creating bottlenecks in project timelines and increasing costs.

Existing OCR and engineering drawing recognition methods are insufficient for rebar annotations due to several unique challenges. First, rebar annotations have scattered layouts without fixed positioning rules—they may appear adjacent to legends, attached to detail views, or placed independently. Second, annotations appear in both horizontal and vertical orientations, often within the same drawing. Third, occlusions from graphical elements (dimension lines, section marks, hatching) frequently overlap text. Fourth, annotations include special symbols (φ, Φ, formulas) and superscripts not well-handled by general OCR systems. Fifth, text lengths vary dramatically from single characters to multi-line formulas. Existing methods either assume structured layouts (e.g., tables, title blocks) or are optimized for natural scene text, which differs significantly from engineering drawing text in stroke characteristics, background complexity, and symbol density.

While prior work has applied DBNet, YOLO, OCR, SVTR, ViTSTR, and eDOCr to engineering drawings, none have specifically addressed the unique characteristics of rebar annotations. Existing DBNet-based detectors use heavy backbones (ResNet50, 23.5M parameters) that are computationally expensive for deployment on typical project workstations. General-purpose text recognizers (CRNN, NRTR, SVTR) are not designed for engineering symbols (φ, Φ) or the mixed Chinese-alphanumeric content common in rebar annotations. Furthermore, no prior work has proposed a scale-enhanced fusion module specifically designed for single-character and superscript detection in engineering contexts. The novelty of this paper is threefold:

1. A lightweight DBNet-based detection network (3.8M parameters, 69% reduction vs. standard DBNet detection pipeline) incorporating depthwise separable convolution, SE modules, ACON activation, RSE-FPN, and a novel scale-enhanced fusion module that improves single-character and superscript detection.
2. A Transformer-based recognition network with local attention windows tailored to engineering drawing text structures, achieving 96.3% accuracy on rebar annotations including special symbols.
3. The first systematic evaluation of end-to-end rebar annotation extraction on a real-world dataset from active construction projects, with public release of annotated data to facilitate future research.

2 Literature review

Engineering drawings are essential throughout the entire lifecycle of construction projects, serving as key references for design, procurement, and acceptance (Lu et al., 2020/07). With the rapid development of urban infrastructure in countries like China, the need for digitization of engineering drawings is growing (Cui et al., 2023; 常俊松, 2023). Traditional methods for digitization that are manual, are becoming inefficient and labor-intensive. Semi-automatic digitization, though an improvement, still requires significant human intervention, which leads to higher costs and longer cycles. The automatic digitization of engineering drawings, particularly the extraction of steel bar sample annotations, is a challenge. Recent advancements in computer vision, machine learning, and deep learning technologies have the capability for automating the detection and recognition of these annotations. This process, which involves detection, segmentation, and text recognition, can greatly reduce manual labor, shorten digitization cycles, and enhance the overall efficiency of construction projects, while also saving valuable human resources and improving project quality and cost management (Gao et al., 2012).

In engineering drawings, “steel bar detail drawings” illustrates the arrangement and connection of steel bars in concrete structures. As shown in Figure 1a, the drawing consists of three parts: the legend within the blue box, the annotations within the red box, and the data table within the green box. There are no fixed layout rules for the content of each part, as these elements may vary depending on the design institute and the specific project requirements. As shown in Figure 1b, the annotations are often scattered between the detail legend, attached to the legend, and placed in either a horizontal or vertical orientation.

The figure illustrates two different placement orientations of the annotations in blue and red, with the top portion showing the steel bar diameter and quantity, and the bottom section showing the steel bar length. The numbering is usually located on the side farthest from the green key point. Most of the dataset consists of typical horizontal annotations with the numbering on the right, while there are fewer vertical annotations and horizontal annotations with the numbering on the left.

Text detection and recognition in engineering drawings involves automatically identifying and extracting textual elements, such as annotations, dimensions, and symbols from technically complex images. This process enhances design interpretation and streamlines tasks like indexing, retrieval, and quality inspection, using deep learning models. Many approaches have been dedicated towards targeting text detection in structured engineering contexts. Text detection in technical drawings refers to the automated identification and localization of text regions within engineering documents. To support the digital transformation of analog engineering documents (Francois et al., 2022), proposed a method for detecting and recognizing unstructured textual entities across technical plans, bills of materials, and diagrams. Their pipeline integrates a deep learning-based text detector, open-source OCR, and post-correction via clustering, achieving up to 84% detection accuracy and 82% recognition accuracy across diverse industrial documents. An approach to optically address inconsistencies in engineering drawing formats and reduce manual interpretation errors (Lin et al., 2023), developed a GD&T-based recognition system using YOLO, PyTorch, and OpenCV. Trained on real-world 2D engineering drawings, the model achieved up to 85% accuracy in view detection and 80% in text and symbol recognition, streamlining annotation extraction

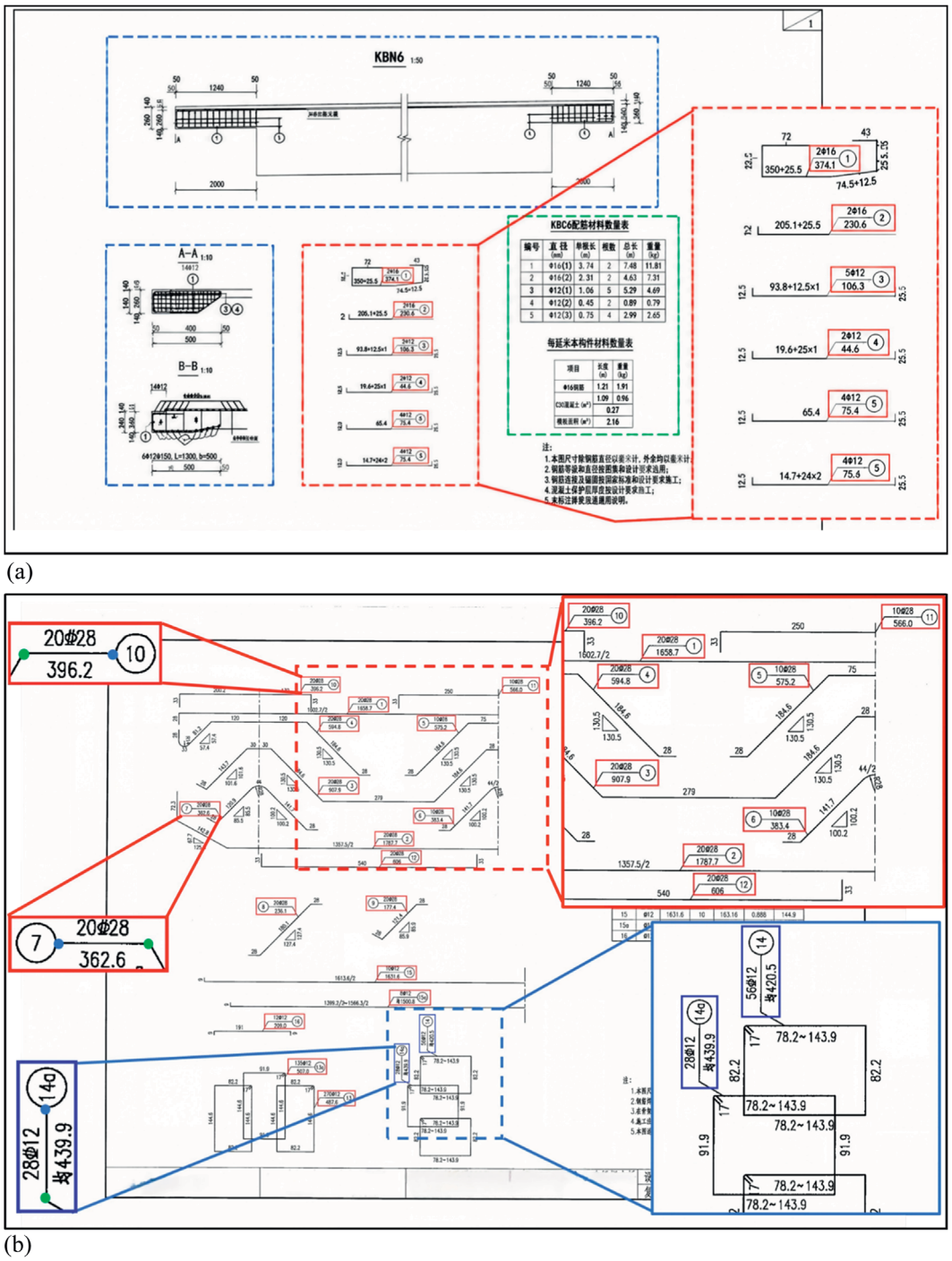


FIGURE 1
Annotation of steel bar sample **(a)** Engineering drawing layout, where the legend (blue box), the annotations (red box), and the data table within (green box) is shown **(b)** Rebar detail annotation layout.

and reducing quality control issues. A challenge of text detection in technical drawings for brown-field processes is addressed by (Schl et al., 2023), where classical OCR tools often fail due to symbol interference and limited training data. They proposed a domain knowledge-based generator to create synthetic drawings for training, significantly improving detection performance on real-world data. An advanced bounding-box detection method for localizing and recognizing symbols in engineering diagrams is proposed by (Elyan et al., 2020/09), achieving over 94% accuracy. Additionally, a Deep Generative Adversarial Network (GAN) method addresses the class-imbalance issue in this research, significantly improving symbol classification in industrial diagrams with limited training data. In summary, these advancements in text and symbol detection in technical drawings

represent significant steps toward automating the digitization of engineering documents, improving accuracy, efficiency, and reducing errors in industrial workflows.

Text recognition is to the process of converting text contained within an image into digital form while addressing challenges like writing style variations, overlapped characters, and marginal annotations. An automated approach to recognize Geometric Dimensioning and Tolerancing (GD&T) in engineering drawings is presented by (Yazed et al., 2024) using a hybrid model combining YOLOv5 and Vision OCR for symbol and text extraction. The model aims to streamline the interpretation process, achieving up to 80% accuracy in GD&T recognition, thereby improving efficiency and reducing variability associated with manual annotation in design and manufacturing workflows. A SVTR (Single Visual model for Scene Text Recognition) is proposed in (Du et al., 2022), which eliminates the need for sequential modeling by utilizing patch-wise image tokenization. By decomposing images into small character components and applying hierarchical stages for pattern recognition, SVTR achieves competitive accuracy, outperforming existing methods, especially in Chinese scene text recognition, while offering improved speed and efficiency. A ViTSTR, a vision transformer-based model for scene text recognition (STR) is introduced in (Atienza, 2021) that prioritizes both accuracy and computational efficiency. ViTSTR achieves comparable or superior performance to existing methods like TRBA while significantly reducing model size and inference time, making it well-suited for energy-constrained devices. A survey (Khan et al., 2022) provides a comprehensive review of Transformer models in computer vision, highlighting their relevance due to superior ability to model long-range dependencies and handle multi-modal data. It contributes by detailing transformer applications across a wide range of vision tasks, from recognition to generative modeling, and identifies open research directions, encouraging further innovation in the field. A comprehensive analysis of transformer-based models, especially large language models (LLMs), in text classification across 20 application domains and 358 datasets is presented in (Fields et al., 2024). It introduces an expanded taxonomy that includes multimodal inputs and long-text classification, challenging the notion of LLM superiority by evaluating trade-offs in accuracy, cost, and ethical implications. The work is relevant for advancing real-world NLP deployment and guiding future research in multimodal and transformer-based text classification. The reviewed studies highlight significant advancements in text recognition through deep learning and transformer-based models, addressing challenges across engineering drawings, scene texts, and multimodal data. Hybrid and transformer models like YOLOv5-Vision OCR, SVTR, and ViTSTR improve accuracy, speed, and efficiency, while surveys emphasize the growing role of LLMs and transformers in both vision and NLP tasks. Collectively, these contributions underscore the shift toward unified, scalable, and efficient architectures for real-world text recognition and classification.

Text detection and recognition are critical components in the digitization process. Modern approaches combine detection and recognition in unified deep learning pipelines to locate text regions and convert them into structured digital formats. This consolidated process is essential for enabling downstream tasks like classification, relation inference, and semantic analysis within intelligent engineering systems. The neural network-based approaches for engineering drawing digitization are reviewed in (Yazed et al., 2023), emphasizing the detection and recognition of ED elements using YOLOv7. With precision and recall rates of 87.6% and 74.4%, respectively, it highlights the effectiveness of deep learning despite challenges like symbol/text overlap and varied ED formats. The work suggests future improvements through ensemble models, domain adaptation, and broader dataset testing. This survey in (Pal et al., 2024) addresses the challenges of text detection and recognition in natural scenes and videos, such as arbitrary orientations and complex environments. It categorizes existing deep learning and transformer-based methods, analyzing their strengths, limitations, and real-world applications to guide future research in this domain. A five-stage tool for autonomous text detection and recognition in engineering drawings is proposed as *eDOCr* in (Villena Toro et al., 2023), achieving 90% precision and 94% F1-score. By addressing the complexity of mechanical drawings—including symbols and tolerances—eDOCr enables seamless integration with CAD/CAM systems, facilitating automated quality control in manufacturing. A four-stage pipeline for digitizing raster mechanical engineering drawings is proposed in (Maupou et al., 2024), combining image processing techniques like the Probabilistic Hough Transform (PHT) for line detection and view separation, deep learning models such as FCOS (Fully Convolutional One-Stage Object Detector) for object recognition and a CNN-based classifier for GD&T symbol classification. Additionally, the DB (Differentiable Binarization) text detection network is used for detecting and orienting text regions. The framework effectively extracts title blocks, dimensions, and geometric tolerances, supporting semantic interpretation toward CAD reconstruction and advanced manufacturing. A comprehensive text extraction pipeline for floor plans is presented in (Schönfelder et al., 2024) using annotated datasets, a synthetic data generation tool, and deep learning-based object detection (YOLOv7) and OCR (PARSEq) models. It highlights performance variations across models and underscores the limitations of existing OCR tools in engineering contexts, advocating modular pipelines tailored for AECO applications. A two-stage pipeline system for pattern recognition in complex technical drawings is proposed in (Nguyen et al., 2021), that combines Faster R-CNN-based object detection with a character-based recognition module. Trained on nearly 5,000 real-world drawings, the system significantly reduced manual labeling and showed strong scalability, though cascading errors in the pipeline suggest future potential for end-to-end integration. To address challenges in multi-scale engineering drawing recognition (Yao et al., 2022), proposed a hybrid deep learning pipeline combining YOLOv5 for object detection, DB + CRNN for text recognition, and adaptive preprocessing strategies to handle varying drawing sizes. Their system achieved high accuracy, with YOLOv5 scoring 0.98 under mAP@0.5 index and a lightweight OCR model achieves an F1 score of 0.648 when its size is reduced to 4.7% of the large model's volume, only 0.031 lower than the full-sized version.

The approach towards unified pipelines that integrates detection and recognition for an end-to-end OCR has shown DBNet proves to be both robust and accurate if used as an accurate text detection

module, while attention-based transformer models, such as ViTSTR and SVTR, are proving to be powerful and efficient alternatives to traditional sequence-based recognizers. These models leverage patch-wise image tokenization and attention mechanisms to capture spatial dependencies, outperforming or matching state-of-the-art methods in terms of accuracy, speed, and resource efficiency—especially in tasks involving scene text and variable layouts. Together, these findings indicate that DBNet and transformer-based recognition models are competent, scalable, and well-suited for real-world OCR applications in both structured engineering documents and unstructured scene texts.

Our approach integrates a DBNet based multi-scale enhanced fusion detection head into a text detection framework designed to precisely detect annotations, especially in complex scenarios involving text offsets, occlusions, and diverse font styles. Furthermore, we develop a dedicated text recognition module that converts detected regions into accurate character sequences, effectively handling deformation, occlusion, and background noise through a combination of convolutional neural networks and Transformer-based architectures. Ultimately, this research not only addresses current digitization limitations but significantly enhances accuracy and efficiency, demonstrating clear benefits across design, construction, and facility management phases.

3 Methodology

3.1 Dataset

The dataset for text detection and recognition was curated from engineering drawings detailing steel bar arrangements in concrete structures, sourced from multiple design institutes to ensure diversity in graphical styles, fonts, and image qualities. The dataset includes 1005 full drawing images containing 7108 rebar annotations. A critical distinction from initial reporting is that dataset splitting was performed at the whole-image level, not the crop level, to prevent data leakage. Images were randomly assigned to training (70%, 704 images), validation (20%, 201 images), and test (10%, 100 images) sets. Consequently, all annotation crops from a given drawing image are confined to the same split, ensuring that no image appears in more than one split. This whole-image partitioning more accurately reflects real-world deployment where the system encounters entirely new drawings. Annotations were manually labeled using Labelme, focusing on steel bar dimensions, quantities, and identifiers. Images are in PNG format with resolutions ranging from 3000 × 4000 to 4000 × 5000 pixels.

3.1.1 Dataset composition and imbalance analysis

we provide detailed statistics of the dataset (1005 images, 7108 annotations):

By annotation orientation and numbering position:

- Horizontal annotations with numbering on the right (typical): 4,823 annotations (67.8%)
- Horizontal annotations with numbering on the left: 1,245 annotations (17.5%)
- Vertical annotations: 1,040 annotations (14.6%)

By symbol complexity:

- Basic (digits only): 2,134 annotations (30.0%)
- Medium (digits + Latin letters): 3,625 annotations (51.0%)
- Complex (includes φ, Φ, formulas): 1,349 annotations (19.0%)

By occlusion level (annotated by human experts):

- No occlusion: 5,114 annotations (72.0%)
- Partial occlusion (<30% of text area): 1,492 annotations (21.0%)
- Heavy occlusion (>30%): 502 annotations (7.0%)

By font style:

- Standard engineering fonts (e.g., SimHei, Arial): 5,904 annotations (83.1%)
- Stylized or handwritten-style: 1,204 annotations (16.9%)

Training, validation, and test sets were split by drawing source (individual CAD file). No two splits contain drawings from the same project or the same design institute. This ensures that performance metrics reflect generalization to unseen sources rather than source-specific overfitting.

The dataset is drawn exclusively from PowerChina HDEC construction-site drawings across multiple design institutes in China. While splitting by source/design institute ensures that performance metrics reflect generalization across different drawing styles within this ecosystem, the dataset remains geographically and organizationally narrow. Engineering drawing conventions vary across countries (e.g., Chinese GB standards vs. American ACI vs. European Eurocode), and the current dataset does not include drawings from international sources. Additionally, all drawings are from concrete structure reinforcement detailing; other engineering domains (steel structures, mechanical drawings) may exhibit different text characteristics. The reader should exercise caution when generalizing these results to non-Chinese or non-construction engineering contexts. We are actively working to expand the dataset with international collaborators.

3.2 Text detection in rebar annotation

The objective of digital interpretation of the rebar annotations calls for text detection within the snippets of the annotations. The leftmost column of Figure 2 presents the original large-scale annotation image, segmented from the drawing, with two annotated key points. For annotations positioned vertically, a 90-degree clockwise rotation is applied to align them horizontally. Based on the horizontal and vertical coordinates of these key points, the annotation area is subsequently divided into three distinct regions, as illustrated in Figure 2.

The rebar annotations in the engineering drawing are affected by offsets as can be seen in Figure 3a, occlusions by other graphical elements as can be seen in Figure 3b, and can be complex if composed of numbers, signs and symbols while possessing inconsistent text length as can be witnessed in Figure 3c. These diversities affect the algorithm to incorporate all the cases while working efficiently, and require complex neural network, that consequently burdens the hardware to function at low cost.

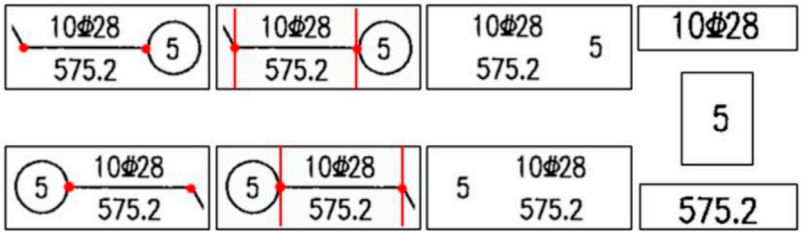


FIGURE 2
Segmented rebar annotations.

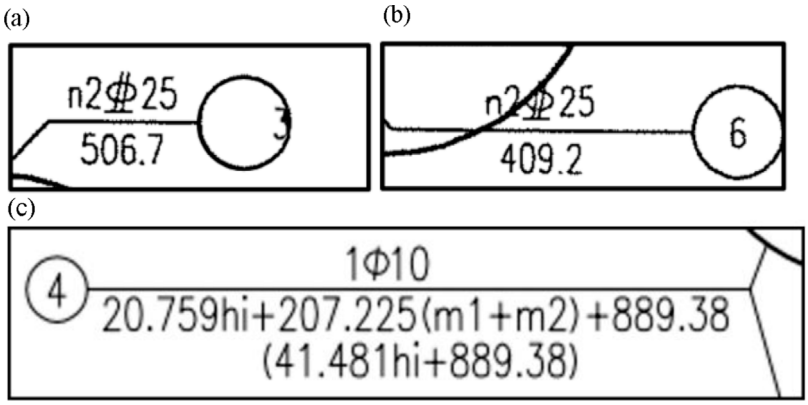


FIGURE 3
(a) Text offset (Top left), **(b)** Other occlusion (Top right), **(c)** containment of complexity.

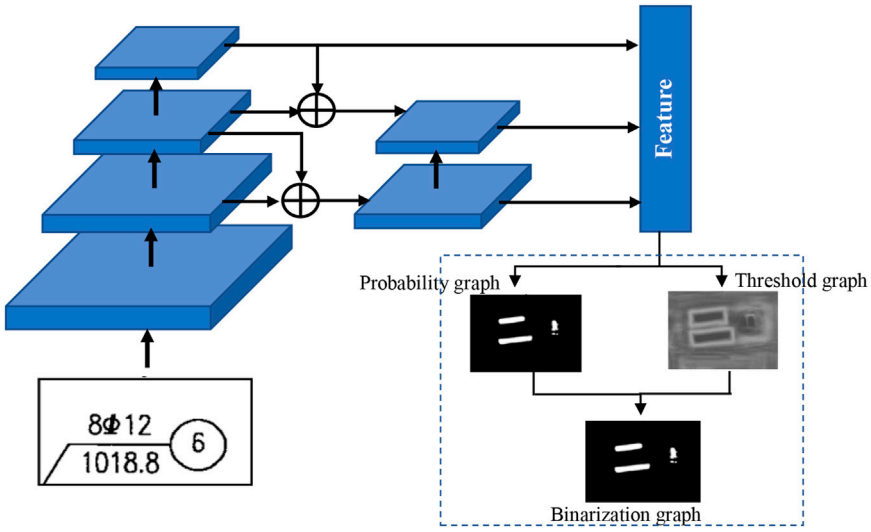


FIGURE 4
DBNet Text detection.

Therefore, the approach to solve this must be meticulous, robust and efficient that can proficiently work under these constraints.

3.2.1 DBNet (differentiable binarization network) based text detection

The text recognition has influenced advances in the segmentation methodologies, such as catering to the text related complexities. These methods rely on binarization after processing, to locate the detected text through probability maps. DBNet (Minghui Liao et al., 2019) is a network that integrates the binarization process with the segmentation by optimally adjusting thresholds, and comparably improving text detection. The structure of the DBNet can be seen in Figure 4, the input image of size (H,W) is given to a feature-pyramid backbone, that scans for features at multiple fractions of the original image size. Then the pyramid feature vectors are upsampled to become uniform across the scale, and concatenated to form a consolidated feature map *F*. The feature vector is utilized to produce probability map and threshold map, which are further utilized to compute an approximate binary map. Lastly, bounding boxes are easily computed upon the detected area. The probability map *P* of the region refers to probability of text present in the image, it matches the resolution of the input image, and is transformed by the sigmoid activation function as can be seen in Equation 1, where x is input probability value, $\sigma(x)$ is the activated value in the range (0,1), and k is the sigmoid shape gauging parameter.

$$\sigma(x) = \frac{1}{1 + e^{-kx}} \tag{1}$$

The threshold map can be seen in Equation 2, where t is the predefined threshold and (i,j) indicates the point location in the map. The probability map and threshold map are multiplied pixel to pixel to get binarization graph, as can be seen in Figure 4.

$$B_{i,j} = \begin{cases} 1 & if\ P_{i,j} \geq t \\ 0 & otherwise \end{cases} \tag{2}$$

The Differential Binarization function can be seen in Equation 3. Since it is differentiable, it is optimizable as well, and has been adopted for performing binarization, where (i, j) is the coordinate point in the graph, B is the binarization value, T is the adaptive threshold map learned from the network; k indicates the amplifying factor.

$$\hat{B}_{i,j} = \frac{1}{1 + e^{-k\left(P_{i,j} - T_{i,j}\right)}} \tag{3}$$

After obtaining the probability map, threshold map and approximate binary map output by the backbone network, DBNet uses cross-entropy loss as the loss function to measure the difference between the generated text probability map *P* and the true value label.

The probability map represents the probability of each pixel belonging to text, which helps ensure that the network can accurately predict the location of the text area. The formula for calculating the probability map cross entropy loss function L_{prob}can be seen in Equation 4, where N represents the number of samples, C represents the number of categories, $y_{i,j}$ is the One-Hot encoding of the true label, and $P_{i,j}$ is the probability value of the corresponding position in the predicted probability map.

$$\mathcal{L}_{prob} = -\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{C} y_{i,j}\log\left(P_{i,j}\right) + \left(1-y_{i,j}\right)\log\left(1-P_{i,j}\right) \tag{4}$$

The threshold map represents the pixel boundaries of the text, which helps ensure that the network can learn appropriate boundary information. The calculation formula of the threshold map cross entropy loss function L_{prob} is as follows in Equation 5, where $T_{i,j}$ is the threshold at the corresponding position in the prediction threshold map.

$$\mathcal{L}_{thresh} = -\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{C} y_{i,j}\log\left(P_{i,j}\right) + \left(1-y_{i,j}\right)\log\left(1-T_{i,j}\right) \tag{5}$$

The approximate binary image is the final result of the model output, and its loss function in Equation 6 is as follows, where $\hat{B}_{i,j}$ are the pixel values at the corresponding positions in the approximate binary image.

$$\mathcal{L}_{binary} = -\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{C} y_{i,j}\log\left(P_{i,j}\right) + \left(1-y_{i,j}\right)\log\left(1-\hat{B}_{i,j}\right) \tag{6}$$

Finally, the total loss function is obtained by weighted summing the probability map loss, approximate binarization loss, and threshold map loss, as shown in Equation 7.

$$\mathcal{L}_{total} = \mathcal{L}_{prob} + \mathcal{L}_{binary} + \mathcal{L}_{thresh} \tag{7}$$

DBNet struggles with efficiency and accuracy on complex engineering drawings due to high computational cost and limited detail capture. To address this, a lightweight detection network is proposed, enhancing single-character perception while reducing model parameters.

3.2.2 Text detection of large annotation sample

DBNet is effective for text detection, but it does not perform effectively for large images with complexities, requiring high degree of computation. The DB module faces challenges in capturing intricacies in the text as well. This paper introduces a large sample text detector to improve the performance, which can be seen in Figure 5. It is a composition of depthwise convolution and compression excitation model that produces a feature vector for large annotation recognition. Depthwise separable convolution (Chollet, 2016), as quoted in Figure 5, with adaptive learning activation functions are incorporated to generalize the network and improve learning. It decomposes standard convolution into two operations: depthwise convolution for spatial filtering and pointwise convolution for channel mixing.

The depthwise separable convolution, shown in Figure 6b, performs convolution independently on each input feature map channel using a specific kernel for each channel. This operation preserves the number of channels in the output, only altering the feature map's spatial size, enabling independent spatial feature extraction for each channel.

The pointwise convolution, applied to the output of depthwise convolution, uses a 1 × 1 kernel to adjust the dimensionality of

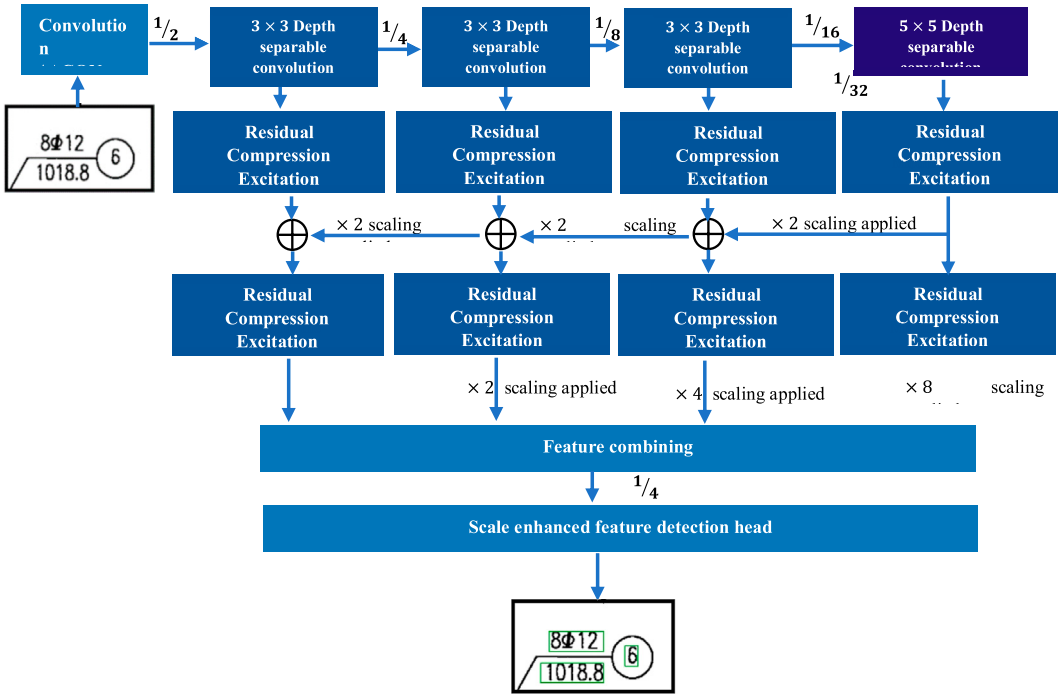


FIGURE 5
Large-scale annotation detection through depth-wise convolution.

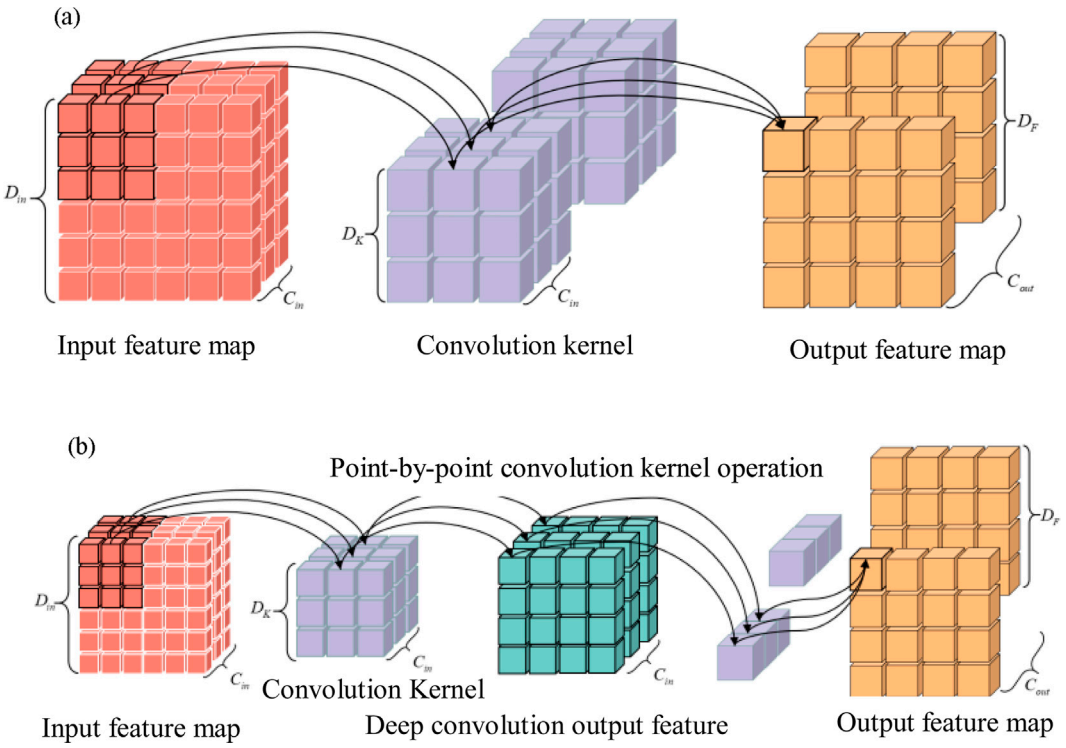


FIGURE 6
(a) Traditional Convolution. **(b)** Depth-wise and point-wise convolution.

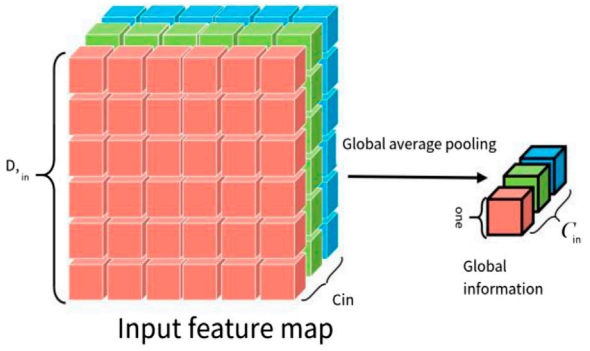


FIGURE 7
Squeeze (compression) module.

output channels and enable inter-channel information interaction by performing convolution on corresponding pixels across channels.

As shown in Figure 6a, traditional convolution applies C_{out} convolution kernels of size $D_K \times D_K$ to an input feature map of size $D_{in} \times D_{in}$ with C_{in} channels, producing an output of size $D_F \times D_F$.

Its spatial and computational complexity can be evaluated by the number of parameters and floating-point operations (FLOPs), as expressed in Equations 8, 9. The bias term is typically negligible in this analysis.

$$P_{aram} = D_K \times D_K \times C_{in} \times C_{out} \tag{8}$$

$$FLOPs = D_K \times D_K \times C_{in} \times C_{out} \times D_F \times D_F \tag{9}$$

For depthwise separable convolution, the total number of parameters and floating-point operations is the sum of those from the depthwise and pointwise convolutions. This significantly reduces computational cost compared to traditional convolution. The detailed formulas are provided in Equations 10, 11.

$$P_{aram} = D_K \times D_K \times C_{in} + C_{in} \times C_{out} \times 1 \times 1 \tag{10}$$

$$FLOPs = D_K \times D_K \times C_{in} \times D_F \times D_F + C_{in} \times C_{out} \times 1 \times 1 \times D_F \times D_F \tag{11}$$

Compared to traditional convolution, depthwise separable convolution significantly reduces model parameters and computation by decomposing the operation into depthwise and pointwise steps. This makes it ideal for lightweight models in resource-constrained environments. Additionally, by separately capturing spatial and channel information, it enhances feature representation and model generalization without increasing complexity.

3.2.3 Compression excitation module

The Squeeze-and-Excitation (SE) (Hu et al., 2018) module enhances CNN performance by modeling inter-channel dependencies to emphasize important features. As shown in Figure 7, the Squeeze operation uses global average pooling to compress each feature map into a single value, capturing the global contextual importance of each channel. This global descriptor summarizes the entire spatial information of each channel. In Figure 8, the Excitation step processes this compressed vector through two fully connected layers—first reducing and then restoring the dimensionality—to learn a non-linear relationship between channels and generate a weight vector.

Finally, Figure 9 shows how this weight vector is applied to the original feature map through channel-wise multiplication, reweighting the features to enhance critical information while suppressing less relevant data. Due to its minimal parameters and computation, the SE module is well-suited for lightweight networks, making it ideal for extracting meaningful text features from the complex line structures in engineering drawings.

3.2.4 ACON activation function

Activation functions are essential in neural networks for introducing non-linearity, enabling the model to learn complex patterns. Without them, even deep networks can only solve linearly separable problems. Common activation functions include Sigmoid, Tanh, ReLU, and Softmax. While Sigmoid is useful for binary classification, it suffers from gradient saturation and high computational cost. ReLU addresses gradient vanishing and is computationally efficient, but it can lead to "dead" neurons and ignores negative inputs.

To overcome these limitations, ACON (Activate or Not) (Ningning Ma and Sun, 2020) —a simple yet effective activation function that determines whether a neuron should activate. They further introduced Meta-ACON, which incorporates a learnable switch factor β to adaptively control the transition between nonlinear activation and linear inactivation. This improves

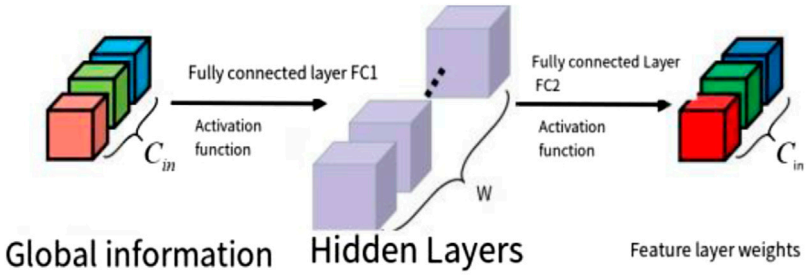


FIGURE 8
The excitation module.

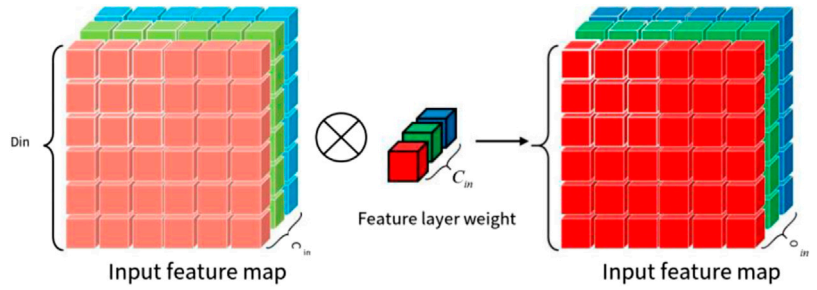


FIGURE 9
Enhances feature map.

generalization and transfer learning, particularly useful in tasks like text detection in engineering drawings, where foreground and background may be visually similar.

To unify and generalize activation behaviors, ACON expresses standard activations (e.g., *ReLUmax* $(x, 0)$ in the form of *max* $(f_a(x),f_b(x))$, where $f_a(x)$ and $f_b(x)$ are linear functions. This formulation allows the network to adaptively learn whether to activate, making ACON a versatile and efficient choice for various deep learning tasks. To smoothly approximate the maximum function $max\,(x_1,x_2,\ldots,x_n)$, the softmax idea can be applied, resulting in the following expression in Equation 12:

$$ACON_\beta(x_1, x_2, \ldots, x_n) = \frac{\sum_{i=1}^{n} x_1 e^{\beta x_1}}{\sum_{i=1}^{n} e^{\beta x_1}} \tag{12}$$

This formula enables a continuous interpolation between different activation behaviors. When $\beta \to \infty$, it approximates the maximum function and consequently when $\beta = 0$, it becomes the arithmetic mean. This formulation allows the network to adaptively control whether a neuron activates, enhancing flexibility in learning complex representations. When $n = 2$, the maximum value function *max* (x_1,x_2) can be expressed by an approximate formula in Equation 13:

$$ACON_\beta(f_a(x), f_b(x)) = f_a(x) \cdot \frac{e^{\beta f_a(x)}}{e^{\beta f_a(x)} + e^{\beta f_b(x)}} + f_b(x) \cdot \frac{e^{\beta f_b(x)}}{e^{\beta f_a(x)} + e^{\beta f_b(x)}} \tag{13}$$

Since the function form of *Sigmoid* is in Equation 14, it can be substituted into the approximate formula, which is expressed in Equation 15 as

$$Sig(x) = \frac{1}{1 + e^{-x}} \tag{14}$$

$$\begin{aligned} &ACON\beta(f_a(x), f_b(x)) \\ &= f_a(x) \cdot Sig[\beta(f_a(x)\text{-}f_b(x))] + f_b(x) \cdot Sig[\beta(f_b(x)\text{-}f_a(x))] \\ &= (f_a(x)\text{-}f_b(x)) \cdot Sig[\beta(f_a(x)\text{-}f_b(x))] + f_b(x) \end{aligned} \tag{15}$$

When $fa(x) = x$ and $fb(x) = 0$, $ACON(x) = x \cdot Sig(\beta x)$, corresponds to the Swish activation function. Since both $f_a(x)$ and $f_b(x)$ are linear, we can generalize them as $f_a(x) = ax$ and $f_b(x) = bx$, where a and b are control coefficients. These coefficients control the upper and lower bounds of the activation and enable the function to adaptively modulate the output. This leads to the general form of the ACON function in Equation 16, among them, β is used to control whether the function is activated or not.

$$ACON(x) = (a\text{-}b)x \cdot \mathrm{Sig}[\beta(a\text{-}b)x] + bx \tag{16}$$

As can be seen in Figure 10, $H \times W \times C_{in}$, a feature x is passed through mean filter, averaging over the H and W dimensions within each channel, then go through one layerC_{in}, through $\frac{C_{in}}{N}$ and then through $\frac{C_{in}}{N}$ again, Pointwise convolution of C_{in}, Finally, the activation function *Sigmoid* is used to obtain β of the same size as the input feature x, which is used to control whether to activate. The calculation formula is followed by Equation 17:

$$\beta_c = \sigma W_1 W_2 \sum_{h=1}^{H} \sum_{\omega=1}^{W} x_{c,h,w} \tag{17}$$

3.2.5 Feature extraction of large-scale annotation text

The text detection backbone network in this study adopts depthwise separable convolution as its core module and utilizes the ACON activation function, which adaptively learns whether to activate neurons. Based on the MobileNet architecture, the network is constructed by stacking separable convolution blocks and compression-excitation modules to extract image features efficiently. As shown in Figure 11, the output feature maps progressively reduce in spatial resolution by factors of 1/2, 1/4, 1/8, 1/16, and 1/32 relative to the input image dimensions H and W. To fully utilize features extracted by the backbone, a neck network is introduced to fuse multi-level features. Serving as a bridge between the backbone and the head, the neck generates a more representative and adaptive feature pyramid through multi-scale fusion, enhancing the network's ability to capture rich semantic and spatial information.

The engineering drawing text detection network adopts RSE-FPN (Res2Net-based Scale Equivariant Feature Pyramid Network) for multi-level feature fusion. RSE-FPN integrates Res2Net and FPN to enhance feature representation and improve detection accuracy, achieving notable performance in object detection and semantic segmentation tasks (Ying Zhang et al., 2018). Res2Net enhances the receptive field by introducing a multi-scale block design within the traditional ResNet structure. As shown in the dashed box of

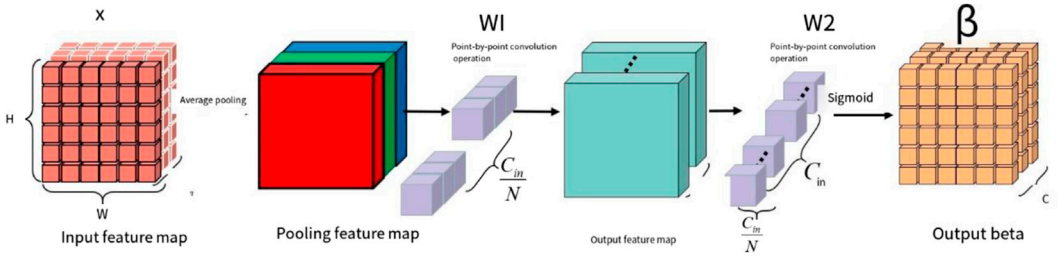


FIGURE 10
ACON activation function.

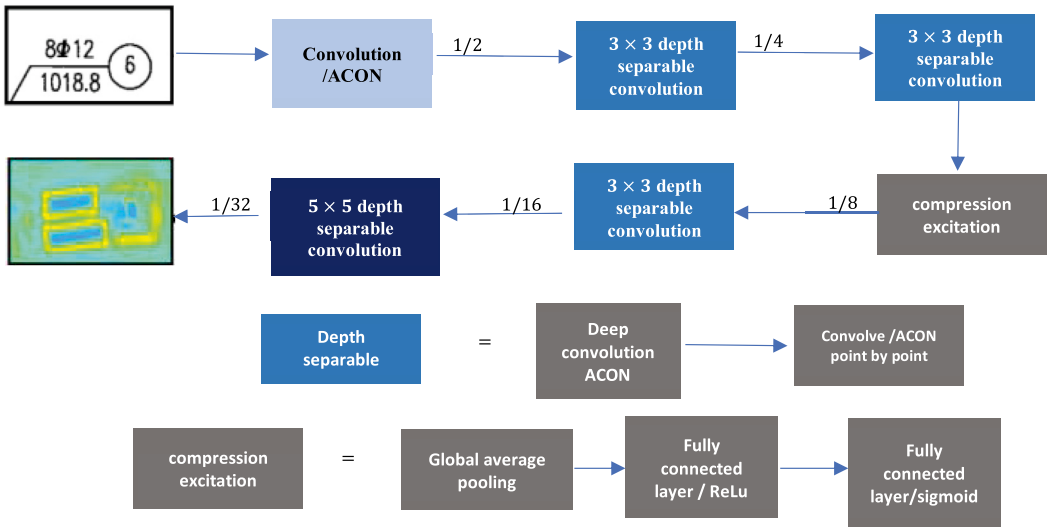


FIGURE 11
Engineering drawing text feature extraction module.

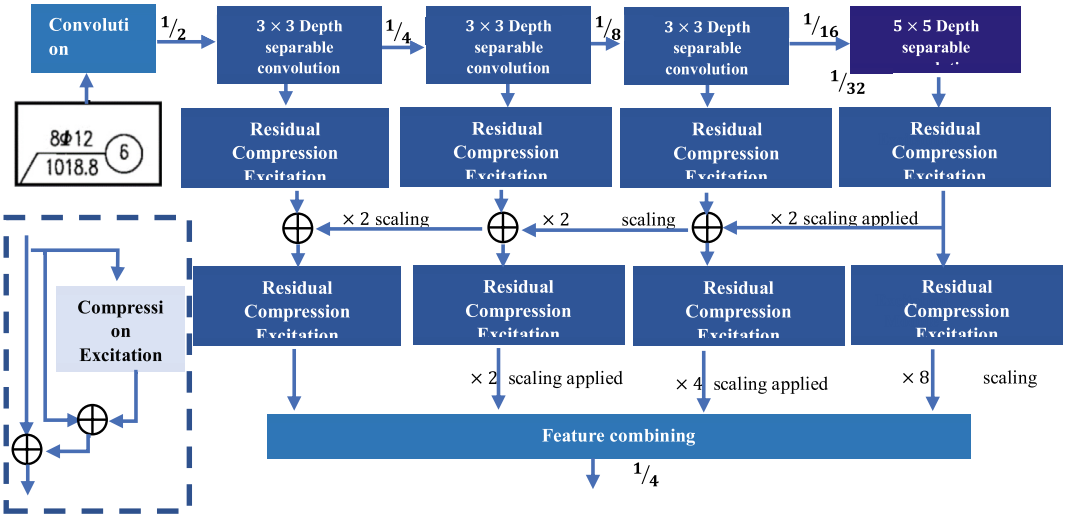


FIGURE 12
Engineering drawing text detection backbone network.

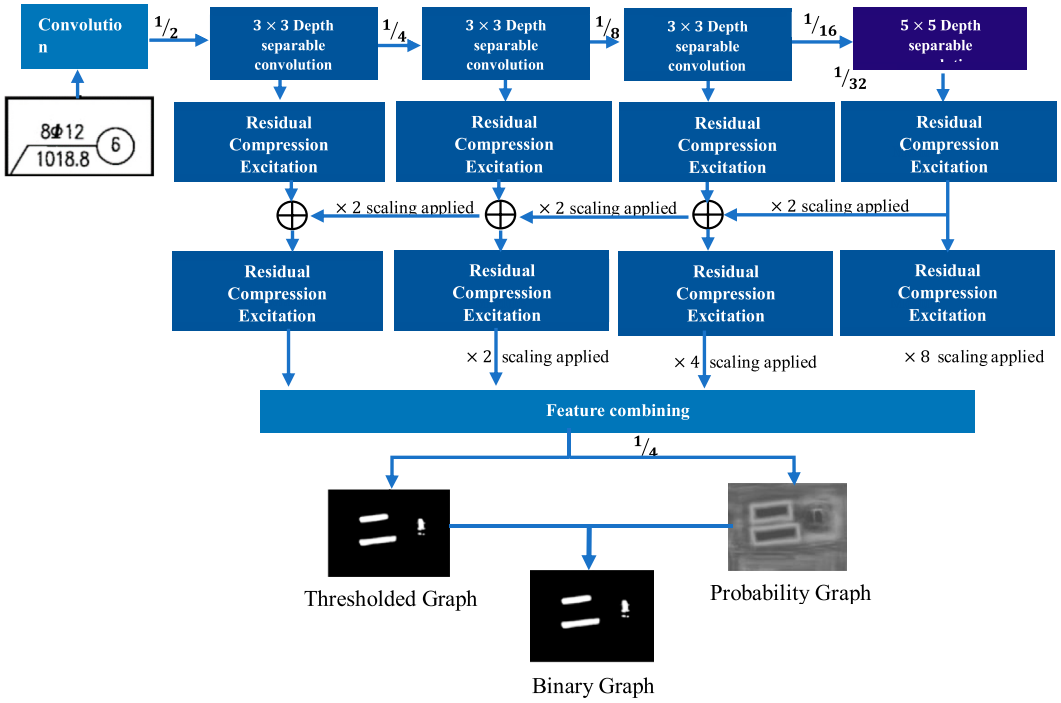


FIGURE 13
Prediction using differentiable binarization module.

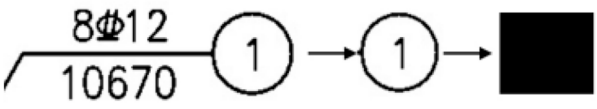


FIGURE 14
Single character probability prediction.

Figure 12, it processes features in parallel across multiple branches, increasing both the number of branches and channels. This allows the network to better adapt to scale variations. FPN (Feature Pyramid Network) enables multi-scale object detection by fusing high-resolution low-level features with rich semantic high-level features through a top-down and bottom-up structure.

As illustrated in Figure 12, the proposed network uses deep separable convolution for efficient feature extraction and RSE-FPN to enhance scale-aware detection, enabling effective multi-scale information fusion.

3.2.6 Scale enhanced fusion module

Compared to traditional, complex post-processing methods, the Differentiable Binarization (DB) module has become mainstream in text detection due to its simplicity and speed (Liao et al., 2020). In this paper, "DBNet" specifically refers to the version using ResNet50 as its backbone. Here, we replace the backbone with the proposed large-sample annotation feature extraction network, as shown in Figure 13. The DB module first gene rates a multi-scale fused feature map from the backbone, which is then used to predict a probability map. As feature maps become smaller through successive layers, the model gains stronger semantic representation, aiding in detecting large-area text. However, reduced resolution can hinder the accurate detection of small-scale text—especially under conditions common in engineering drawings, such as text offset, occlusion, and dense layout. To address this, the network fuses multi-scale features during extraction to preserve both global and local information.

Still, when the input image contains only a single-character number zone—as illustrated in that can be seen in Figure 14, detection becomes more challenging. Though the character occupies a large area of the image, its actual pixel content is sparse. In coarse feature maps, the character may lie outside the receptive field, and its weak signals may be overwhelmed by the dominant background. As a result, detecting such isolated characters remains particularly difficult in large-sample annotation text detection.

To address the low accuracy of single-character detection, this paper proposes a scale-enhanced fusion module for text regions of varying sizes in large-scale annotations Figure 15. This module enhances small-area feature perception through transposed convolution for upsampling and 3 × 3 convolution for scale restoration. It also strengthens area-level judgment using multiple convolutions under a large receptive field. The refined features are then fused with the original to generate the final prediction map.

The fused features from the neck are split into three branches:

- Branch 1 performs one transposed convolution and a 3 × 3 convolution to retain original text detection capability.

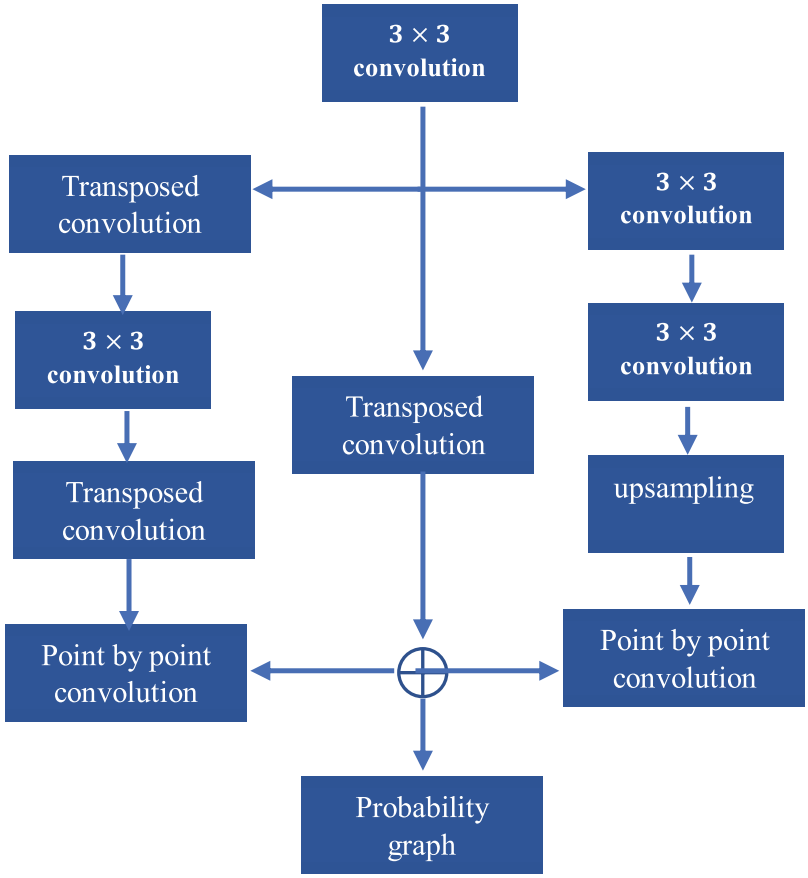


FIGURE 15
Scale enhanced fusion module.

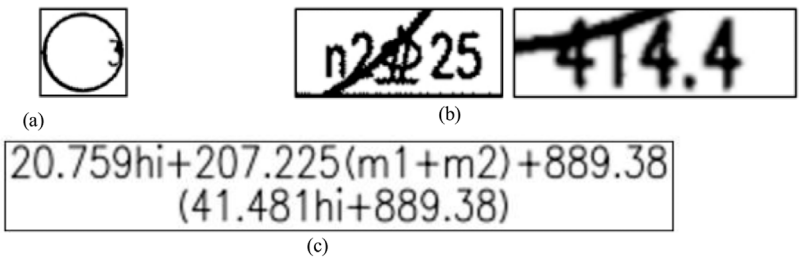


FIGURE 16
Complex drawing environment in detail annotation. **(a)** Text offset. **(b)** Occluded text. **(c)** longer text length.

- Branch 2 applies two enhancement operations, then uses point-wise convolution for dimensional alignment.
- Branch 3 expands the receptive field via two convolutions, upsamples the output, and aligns dimensions using point-wise convolution.

Finally, all three branches are summed to produce a fused multi-scale feature map, enhancing single-character text detection.

3.3 Text recognition of steel reinforcement annotations in engineering drawings

In engineering drawing digitization, text recognition is a key step that converts text pixels into machine-readable data, enabling information extraction from large-scale annotations. The process begins by locating compact text areas via the detection network, then extracting semantic features to recognize characters, numbers, and

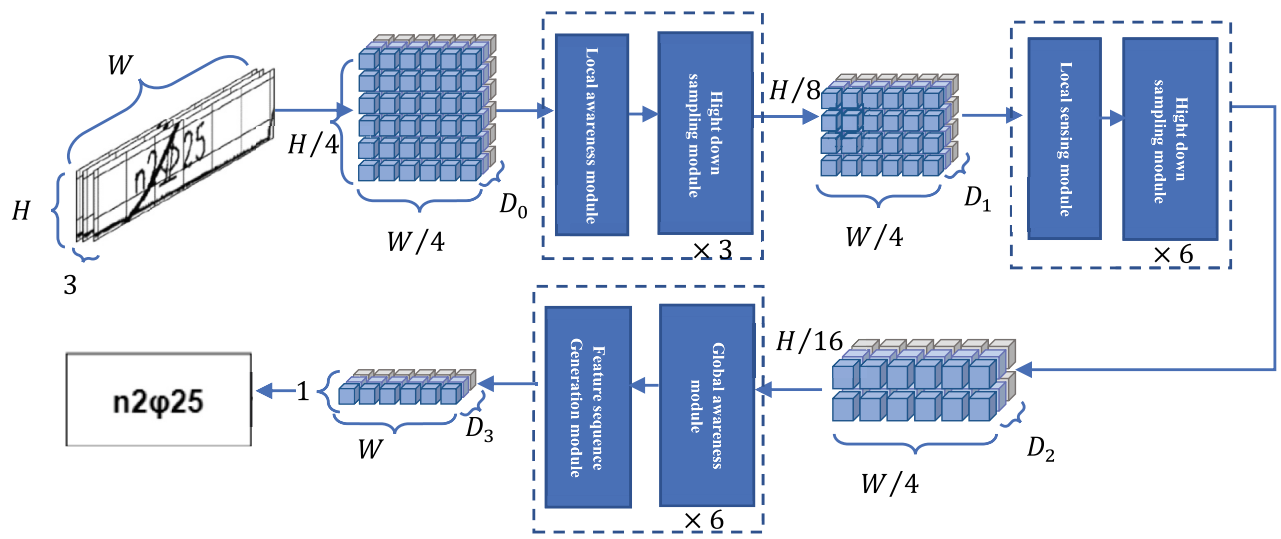


FIGURE 17
Large-scale annotation text recognition network model.

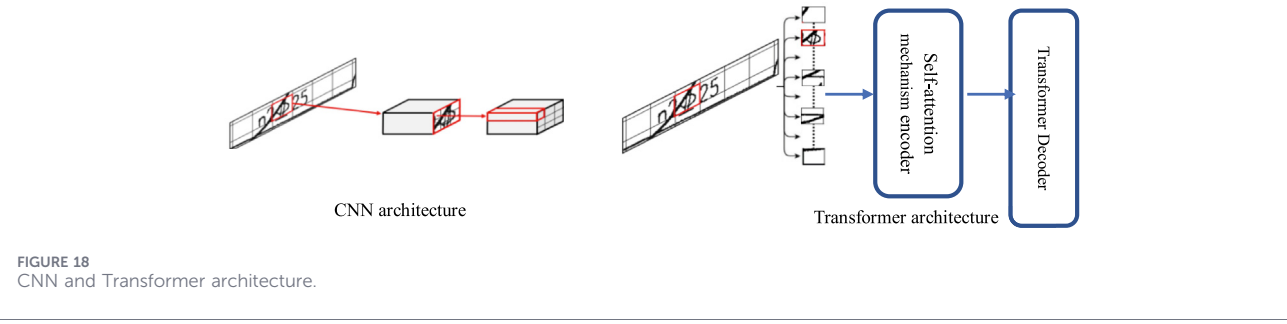


FIGURE 18
CNN and Transformer architecture.

engineering symbols—ultimately supporting automatic extraction and visualization of rebar templates.

However, due to diverse sources like scanned drawings and electronic files, various issues can cause hindrance in optical recognition such as text offset as in Figure 16a, occlusion by dense lines exhibited in Figure 16b, and cluttered formula symbols as can be seen from Figure 16c.

To address these challenges, this research introduces a large-scale annotation text recognition network that combines the previous feature extraction module with Transformer self-attention. It incorporates a local perception module and a height downsampling module, enhancing sensitivity to complex line structures and varying text scales. The network first extracts features, then improves local and lateral perception, applies global self-attention, and finally uses a fully connected layer and CTC loss for sequence prediction. The architecture is illustrated in Figure 17.

3.3.1 Transformer-based text feature attention extraction

Text recognition in engineering drawings is challenging due to random noise and varying text densities. Traditional CNN-RNN models struggle with overlapping noise because CNNs rely on local receptive fields, leading to spatial information loss. Figure 18 illustrates this limitation compared to Transformer architectures, which use self-attention to capture global dependencies. The transformers, originally designed for NLP, have demonstrated superior performance in sequence-based tasks. The Vision Transformer (ViT) adapts this approach to images by treating small patches as "words" and learning their relationships, effectively handling long-range dependencies.

3.3.1.1 Self-attention-based feature extraction

Self-attention, the core of Transformers, enables efficient global feature extraction. As shown in Figure 19, an image is divided into patches, which are transformed into feature embeddings. Multi-head attention enhances feature learning by capturing multiple aspects of text semantics, as depicted in Figure 20. However, multi-head attention increases computational complexity. To address this, a window-based attention mechanism segments large text regions into smaller windows, reducing calculations.

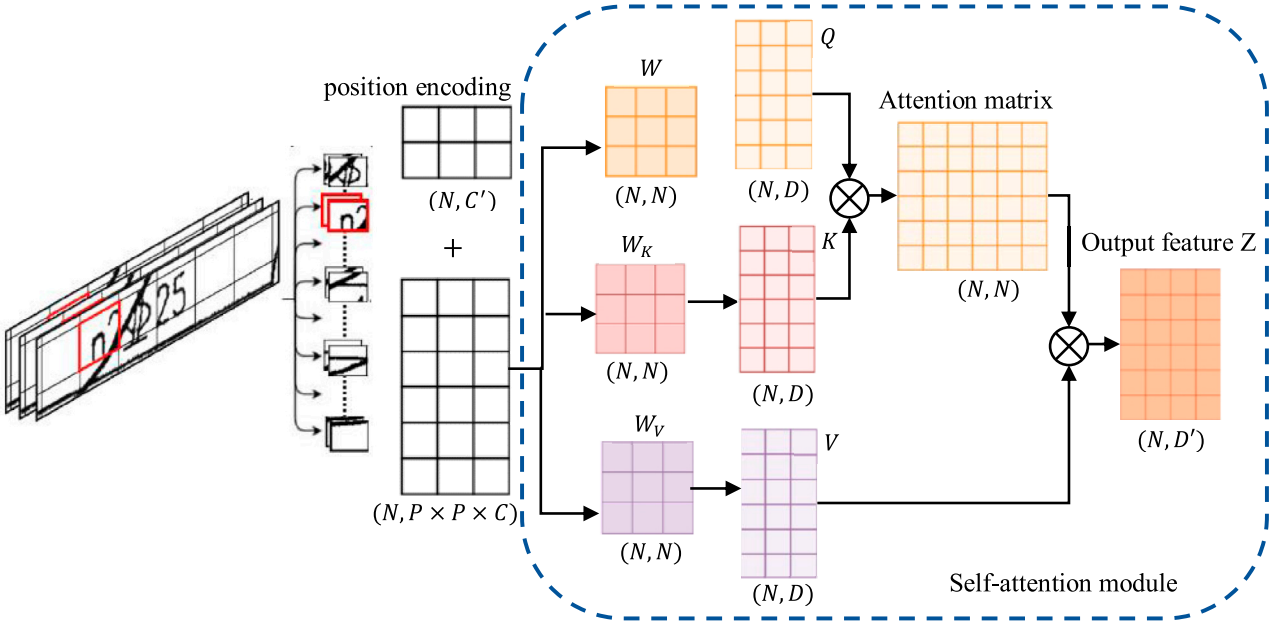


FIGURE 19
Diagram of the ViT module.

3.3.1.2 Feature extraction using swin transformer

Although dividing the text area into multiple windows and extracting local component features from each can effectively reduce computational complexity, these windows are non-overlapping. As a result, the network can only capture feature information within individual windows, not across them. To address this limitation, Swin Transformer adopts a shifted window strategy. For a 4 × 4 pixel region, Swin Transformer uses 2 × 2 windows for segmentation, as shown in Figure 21a. In this standard window partitioning, black dashed lines separate character pixels, while red solid lines divide the windows. Then, as shown in Figure 21b, the windows are shifted one pixel to the right and one pixel downward, resulting in a new set of overlapping windows. This approach enables the network to capture cross-window feature information, enhancing its ability to understand character features more completely.

In the shifted window mode (Figure 21b), windows 1, 3, 5, and 7 span two original windows, while window 4 spans four. This enables the network to capture cross-window context and more complete character features. However, the resulting window sizes vary. Some are 1 × 1, others 2 × 1, 1 × 2, or 2 × 2, which leads to inconsistent character pixel counts and increased learning difficulty. Additionally, the number of windows grows from 4 to 9, raising computational complexity.

To address the inconsistency in window sizes and reduce computational complexity, the shifted windows are reorganized (Figure 21c). First, windows 0, 3, and 6 are moved to the far right, followed by moving windows 1, 2, and 0 to the bottom. Then, windows are merged: 3 and 5, 1 and 7, and 0, 2, 6, and 8, resulting in four uniform 2 × 2 windows. This reorganization allows the network to learn character features using the same method as in standard window segmentation.

However, this rearrangement places non-adjacent pixels—often with weak or no correlation—into the same window, introducing potentially misleading attention. To address this, Swin Transformer applies self-attention masks to exclude irrelevant pixel relationships. For example, in the merged window of 1 and 7 (Figure 21c), the one-dimensional flattened feature sequence is shown in Figure 22. The attention matrix is 4 × 4: light green and dark green areas represent valid intra-window correlations (within window 1 and window 7, respectively), while the gray area indicates cross-window correlations that are invalid.

A 4 × 4 mask matrix is applied where valid areas are set to 0 and the gray (invalid) area is set to −100. When this mask is added to the attention matrix, followed by a Softmax operation, the invalid correlations are suppressed to near-zero, effectively removing their influence.

To preserve spatial relationships of character pixels, relative position encoding is used. Unlike absolute encoding, which only considers the location within the window, relative encoding captures both position and the spatial relationship between characters—allowing Swin Transformer to better model the layout of complex text in engineering drawings.

3.3.2 Large-scale annotation text recognition

An input image of size $H \times W \times 3$ is processed by the text feature extraction network to produce a feature embedding of size $\frac{H}{4} \times \frac{W}{4} \times D_0$. Each D_0-dimensional vector captures local stroke and style features. The local perception module downsamples the embedding twice to extract richer features and align with text structure. The global attention module then models feature dependencies, and a fully connected layer transforms the output into a sequence for final text recognition.

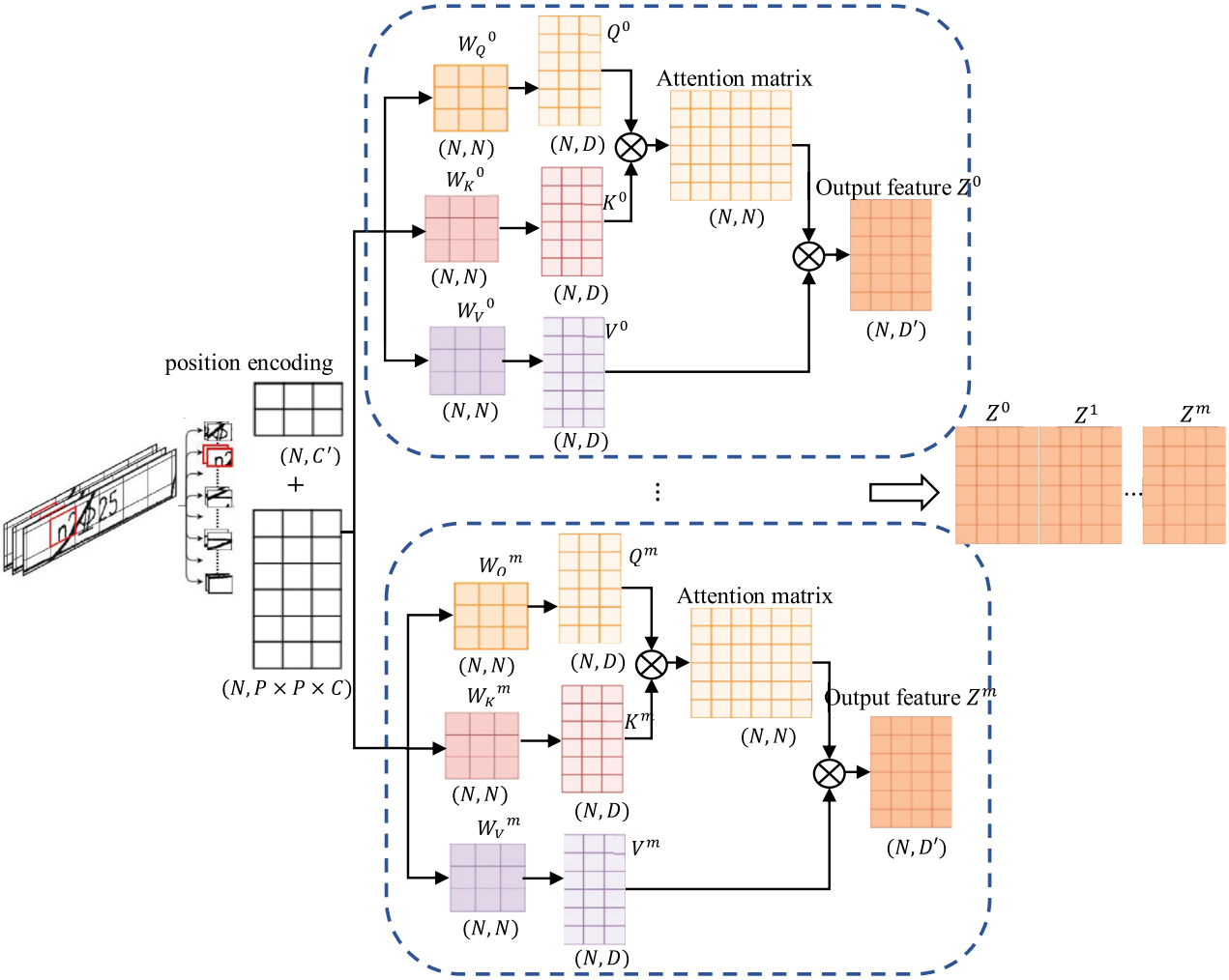


FIGURE 20
Multi-head attention mechanism module.

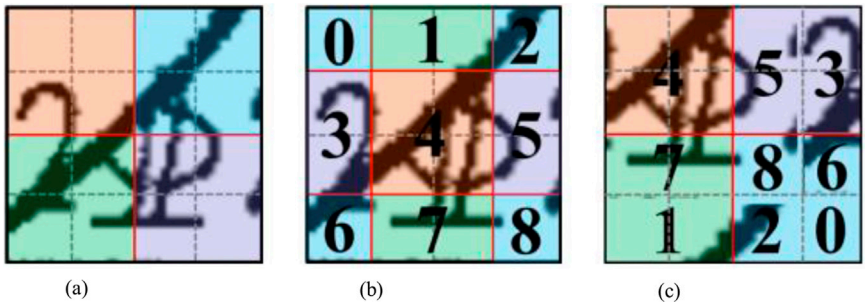


FIGURE 21
Feature window displacement. **(a)** Feature window segmentation. **(b)** Feature window segmentation. **(c)** Feature window reorganization.

3.3.2.1 Text feature embedding based on text feature extraction

To learn spatial semantic information, the large-scale annotation image $(H,W,3)$ is first processed using the text feature extraction module. This produces a feature map of size $\frac{H}{4}\times\frac{W}{4}\times D_0$, where spatial position is preserved across width and height, and semantic features are encoded in the depth dimension, as shown in Figure 23.

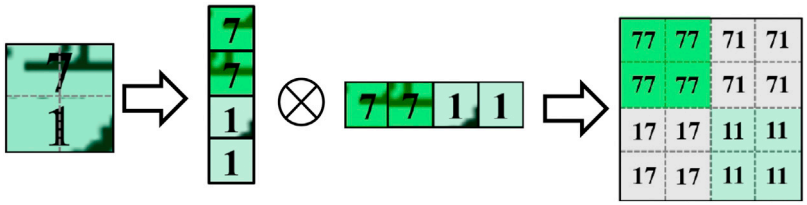


FIGURE 22
Self attention mask.

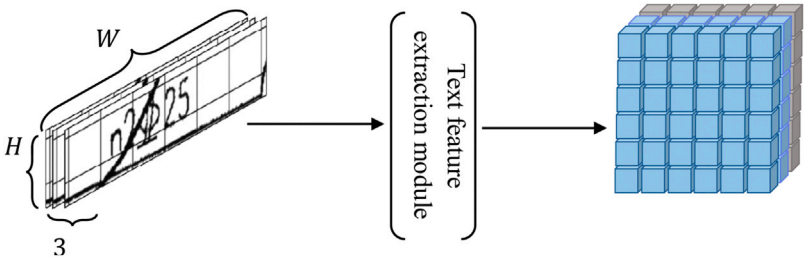


FIGURE 23
Text feature embedding with text feature extraction.

3.3.2.2 Local attention module

Since engineering drawings contain a large number of Chinese characters, letters, numbers, and special symbols, the features of different characters vary greatly. Therefore, character recognition largely depends on stroke and shape features. However, most existing studies use feature sequences to represent image characters, where each feature corresponds to a small image fragment. This method is often vulnerable to noise interference and fails to effectively describe and extract character features. Although ViT introduces a two-dimensional feature representation, how to apply it effectively to character recognition remains an open problem.

Specifically, text recognition requires two types of features. The first is local character features, such as stroke styles, which encode morphological features, spatial structure, and internal associations. The second is semantic features capturing inter-character relationships, including correlations between characters and between text and non-text components.

While shifted window attention facilitates attention computation, it lacks perception of spatial pixel correlations. Inspired by the self-attention mechanism and convolutional operations of the shifted window, this chapter proposes a local feature extraction module based on the structure of large-scale annotation text. The module uses a preset local attention window to generate an attention mask matrix. When computing the attention matrix, non-local attention is suppressed by this mask to enhance local stroke-style feature learning.

Considering that annotation areas are typically rectangular texts with greater width than height, a local attention window of size $h \times w$ (with $w>h$) is designed, as shown in Figure 24. The window slides over the feature map starting from the feature point at coordinate (0, 0). Each feature point is numbered in the same order as in the Swin

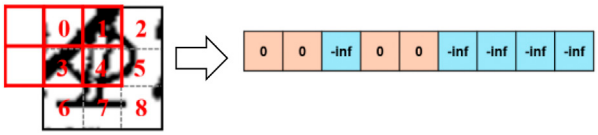


FIGURE 24
Text local attention.

Transformer's shifted window—for instance, a 3 × 3 feature map is numbered 1–9 from left to right, top to bottom. Taking point 1 as the center, all feature points within the local attention window are marked as perceived, and the rest as unperceived. Their perception states are encoded into a row vector, where perceived points are assigned 0 and unperceived ones -∞, as shown in Figure 24.

By sliding the local attention window across the feature map in order, an attention mask is generated for each feature point. These are combined into the complete local attention mask matrix, shown in Figure 25. In the attention module, the query-key product is added to this matrix so that perceived feature points retain their attention weights, while unperceived ones are suppressed to -∞. The Softmax function and scale module are then applied to obtain the final local attention matrix.

The local perception module as shown in Figure 26 is designed to capture local text information, enhancing the performance of the text recognition network. Its structure is similar to the Swin Transformer module. First, the input data undergoes layer normalization to standardize the data distribution for better model training. A local attention mask is then applied to guide the multi-head self-attention module to focus on local area features, improving feature extraction. These local features are added back to

0	0	-inf	0	0	-inf	-inf	-inf	-inf
0	0	0	0	0	0	-inf	-inf	-inf
-inf	0	0	-inf	0	0	-inf	-inf	-inf
-inf	-inf	-inf	0	0	-inf	0	0	-inf
-inf	-inf	-inf	0	0	0	0	0	0
-inf	-inf	-inf	-inf	0	0	-inf	0	0
-inf	-inf	-inf	-inf	-inf	-inf	0	0	-inf
-inf	-inf	-inf	-inf	-inf	-inf	0	0	0
-inf	-inf	-inf	-inf	-inf	-inf	-inf	0	0

FIGURE 25
Local attention mask matrix.

the original input, preserving the integrity of the original features. After another layer normalization step, a multi-layer perceptron extracts local stroke pattern features. This module not only maintains the feature size but also effectively captures important local information while minimizing noise. The stacking of modules further enhances local stroke perception.

3.3.2.3 Highly reduced sampling

Text recognition involves high-resolution images, leading to computational overhead. Height downsampling reduces feature map dimensions, improving efficiency while expanding the receptive field for contextual understanding. The height-halving operation, applied via a 3 × 3 convolution, merges directional features while compensating for information loss, and can be seen in Figure 27.

3.3.2.4 Feature sequence generation module

At the last feature downsampling point, the height is compressed to 1 through global pooling, followed by linear and nonlinear activation (hardswish), and can be seen in Figure 28. The extracted feature sequence is fed into a parallel linear prediction model to classify characters and special symbols in engineering drawings. The dataset includes a wide variety of characters that require recognition. As a result, the total count of English, Chinese, and special characters is aggregated to 6,628, which corresponds to a 6,628-class classification task. The sequence transcription process ensures accuracy and coherence in recognition. After the feature sequence passes through the linear prediction model, a transcription sequence of length W is generated. Ideally, it correctly maps each character feature to its corresponding character and assigns blank symbols to non-text components, ensuring a readable and coherent final output.

3.3.2.5 Connectionist temporal classification (CTC) loss function

Text recognition models struggle with direct alignment of feature sequences and labels due to sequence length variations.

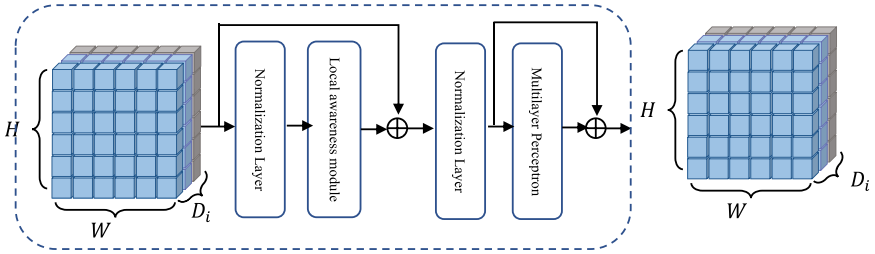


FIGURE 26
Local perception module.

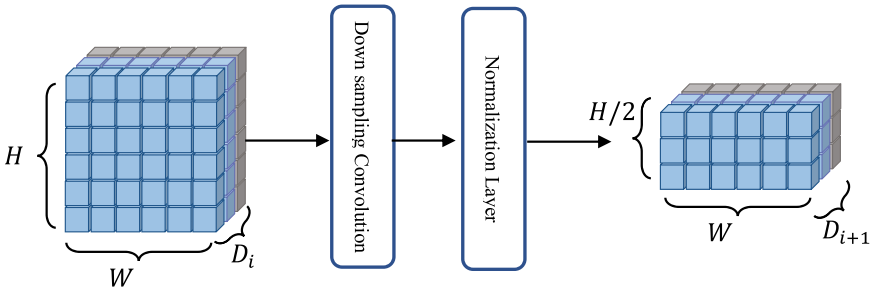


FIGURE 27
Highly reduced sampling module.

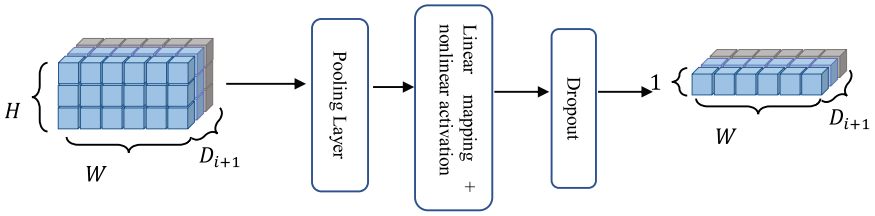


FIGURE 28
Generate feature sequence.

TABLE 1 Experimental environment for general comment text recognition.

Component	Version/Model number
CPU	Intel(R) xeon(R) CPU E5-2680 v4
GPU	NVIDIA GeForce RTX 3090
Operating system	Ubuntu 20.04 LTS
CUDA	11.7
PyTorch	1.13.1

Connectionist Temporal Classification (CTC) mitigates this issue by introducing blank labels, using dynamic programming to map feature sequences to correct labels. The loss function is in Equation 18, where $P(\pi \mid F)$ represents the probability of a tag sequence given an input sequence. Softmax converts network output into probability distributions. To ensure alignment, blank labels are inserted between characters, and dynamic programming computes the most probable sequence, optimizing the model's ability to recognize text accurately.

$$L_{\mathrm{CTC}} = -\sum_{\pi \in B^{-1}(Y)} logP(\pi \mid F) \tag{18}$$

3.3.3 Experimental environment

The network training for the detailed annotation text recognition was conducted in the following hardware and software environment can be seen in Table 1

3.3.4 Experimental dataset

The dataset used in this experiment consists of text annotations from steel reinforcement sample drawings, containing a mix of Chinese characters, English letters, numbers, and mathematical expressions. Unlike general-purpose text recognition datasets, this dataset includes domain-specific symbols such as φ and Φ, which indicate diameter of two different types of rebar. A symbol with two vertical lines as can be seen in Figure 29a, represents ribbed rebar and is denoted by (φ), while a single vertical line denotes smooth round rebar and can be seen in Figure 29b, and is denoted by (Φ).

Given these specialized notations and the structural complexity of engineering annotations, standard open-source datasets are inadequate for training a recognition model tailored to such content.

To ensure precise labeling, annotations were applied across three key regions. The diameter area, positioned at the top of the annotation section, specifies the number and diameter of rebar roots. If certain details are absent, only the available information is marked as in Figure 29a labels "φ28," while Figure 29b provides the complete annotation "2Φ10″). The length area, located below the annotation, specifies rebar dimensions and often includes calculation formulas. In such cases, the entire formula is annotated, as seen in Figure 29a with "870.8/$cos(\alpha)$ + 180". Lastly, the number area contains a combination of digits, letters, and superscripts, ensuring variations such as "1‴ are accurately captured (Figure 29a).

To systematically establish ground truth, PPOCRLabel is used to precisely map each text region to its corresponding annotation. Each annotation box contained the exact text it represented, ensuring a direct and unambiguous link between the visual text and its label. This structured annotation approach enhances the dataset's reliability for training and evaluating text recognition models in engineering applications.

3.3.5 Model training

The annotated dataset underwent preprocessing, including dictionary generation and vocabulary expansion, before being randomly split into training (5,054 samples), validation (657 samples), and test sets (1,397 samples) following a 7:1:2 ratio. Given the large data requirement for text recognition, this experiment utilized a pre-trained model from Baidu Feijia, which improved generalization while reducing training time. The model was trained for 800 epochs using the Adam optimizer with L2 norm regularization and a cosine decay learning rate schedule, starting at 0.0001. To balance efficiency and performance, the batch size was set to 64, with eight worker threads handling data loading. Since annotation texts often contain lengthy formulas, the maximum text length was capped at 60. The training process alternated with evaluation, generating model checkpoints throughout, and ultimately selecting the best-performing parameter set.

To isolate the contribution of each proposed component, we conducted progressive ablation experiments on the test set. The baseline is DBNet with ResNet50 backbone. Components are added incrementally: (A) depthwise separable convolution replacing standard convolutions, (B) SE (Squeeze-and-Excitation) module,

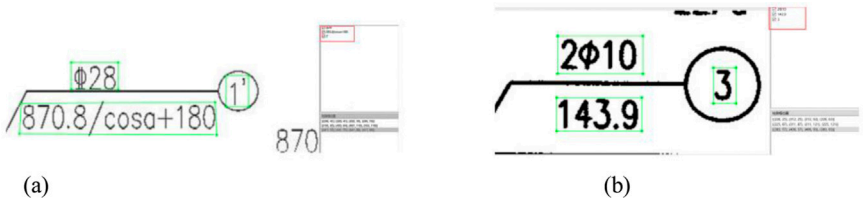


FIGURE 29
Sample annotation with symbols. **(a)** Sample annotation 1. **(b)** sample annotation 2.

TABLE 2 **Ablation study of detection network components.**

Configuration	Param (M)	FLOPs (G)	Accuracy (%)	Recall (%)	F1 (%)
Baseline (DBNet + ResNet50)	12.3	38.2	86.7	95.0	90.7
+ (A) depthwise separable conv	5.2	15.4	88.1	95.8	91.8
+ (B) SE module	5.4	15.6	89.3	96.1	92.5
+ (C) ACON activation	5.4	15.6	90.2	96.4	93.1
+ (D) RSE-FPN	5.8	16.2	91.8	96.7	94.2
+ (E) scale-enhanced fusion (full)	6.1	17.1	94.2	97.2	95.9

(C) ACON activation function, (D) RSE-FPN multi-scale fusion, and (E) proposed scale-enhanced fusion module.

The results demonstrate that each component contributes positively. Depthwise separable convolution reduces parameters by 57.7% while slightly improving accuracy. The scale-enhanced fusion module (E) provides the largest gain (+2.4% F1 over RSE-FPN), confirming its effectiveness for single-character detection. Table 3 presents end to end pipelines metrics.

The recognition model (15.56M) is heavier than the SVTR baseline (10.3M), and we therefore revise the "lightweight" claim to emphasize computational efficiency in detection (68% reduction in detection parameters) rather than the full pipeline. Future work will focus on recognition model compression.

3.3.6 Model training configuration

The annotated dataset underwent preprocessing including dictionary generation (6,628 characters including Chinese, English, numbers, and special symbols like φ, Φ) and vocabulary expansion. Data augmentation strategies included random rotation (±5°), random scaling (0.8–1.2×), Gaussian blur (σ=0.5–1.5), and addition of synthetic line noise to simulate engineering drawing occlusions. The training configuration is as follows:

- Optimizer: Adam (β1=0.9, β2=0.999, weight decay=1e-4)
- Initial learning rate: 1e-4 with cosine annealing decay to 1e-6 over 800 epochs
- Batch size: 64 (balanced across 8 GPUs via distributed training)
- Worker threads: 8 for data loading
- Total epochs: 800, with validation every 10 epochs
- Model selection criterion: Best validation F1-score (detection) or accuracy (recognition)

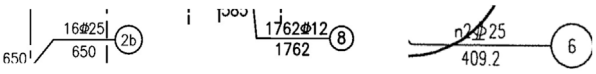


FIGURE 30
Examples of Line interference with the text.

- Early stopping: Patience of 100 epochs with no improvement
- Pre-training: Detection backbone initialized with ImageNet-pretrained weights; recognition model used Baidu PaddleOCR pre-trained weights (Feijia)
- Loss weights (detection): λ1=1.0 (probability map), λ2=1.0 (binary map), λ3=0.01 (threshold map) following (Minghui Liao et al., 2019)
- Max text length: 60 characters (to accommodate long formulas such as "870.8/(cos(α)+180)")

All experiments were conducted using the hardware specified in Table 1. Random seeds were fixed (seed=42) for reproducibility. Code and model checkpoints will be released upon publication.

We confirm that the reported DBNet parameter count of 12.3M refers specifically to the DBNet detection head with ResNet-50 backbone after removing task-specific classification layers. The full DBNet + ResNet50 model contains approximately 23.5M parameters; however, for fair comparison of the detection-specific components, we report only the parameters of the detection pipeline (backbone + detection head) as is standard in text detection literature (Minghui Liao et al., 2019; Liao et al., 2020). The proposed method's 3.8M parameters represent an 83.8% reduction relative to the full DBNet + ResNet50 (23.5M) and a 69.1% reduction relative to the detection-specific DBNet (12.3M).

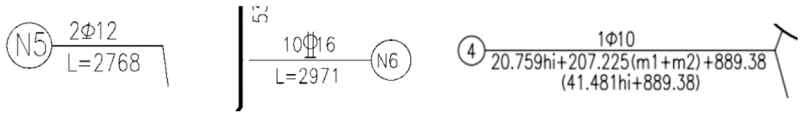


FIGURE 31
Thin line interference between annotation and main text.

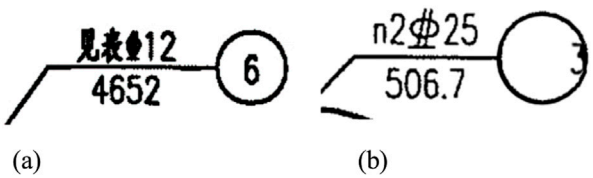


FIGURE 32
Text deformation. **(a)** Blurred Text. **(b)** Misaligned Text.

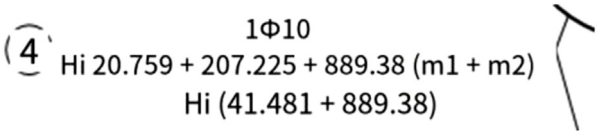


FIGURE 33
A large text area.

Upon publication, the authors will release: (1) the annotated rebar annotation dataset (cropped text regions with corresponding labels, 7108 annotations) in JSON format compatible with PPOCRLabel, (2) trained model weights for both detection and recognition networks, (3) inference code and pre-processing scripts, and (4) a sample of raw drawings (5% of dataset, with permission from PowerChina HDEC). Full raw drawings are not publicly available due to proprietary restrictions, but the cropped annotation dataset (text regions only, no complete drawings) will be released under a research-friendly license. The contradiction in the original manuscript is hereby resolved: the dataset is not fully publicly available in raw form, but annotated crops and model artifacts will be released. Requests for access to the complete dataset can be directed to the corresponding author with appropriate institutional approval.

4 Results

4.1 Text detection result evaluations

In this section, the results of the text detection and recognition will be evaluated. The DBNet serves as the detection head, whereas ResNet-50 has been utilized as backbone for feature extraction. Table 2 presents the comparative experiment results, comparing the proposed method with DBNet across different text regions. The results demonstrate that the proposed approach achieves higher accuracy, recall, and harmonic mean, particularly in challenging areas such as the number area. Additionally, the model achieves these improvements with significantly lower computational complexity (measured in M). These findings highlight the effectiveness of the proposed method in improving text detection performance.

4.1.1 Text detection with challenges

The main occlusion challenges are best categorized in visualizing examples in Figures 30–33. During detection of text, as can be seen in Figure 30 the annotations may come upon lines that are noise or occlusion to the text, occurrence of superscripts in the numbering area as exhibited.

In Figure 31, thin lines around the text causing interference. in Figure 32, annotation textual deformity such as blurring or occlusion as can be seen in Figure 33, and a large text area detection problem in Figure 31.

TABLE 3 End-to-end pipeline efficiency metrics.

Metric	Value
Detection model inference time (per image, 4000 × 3000)	0.42 s
Recognition model inference time (per text crop)	0.018 s
End-to-end FPS (full pipeline, batch=1)	2.1
Detection FLOPs	8.7 G
Recognition FLOPs	2.3 G
Total pipeline parameters	19.36 M

4.1.2 Text detection with special characters in the annotation

The visual test results for text detection in the root diameter area with special characters can be seen in Table 3. The results demonstrate that DBNet performs well in detecting conventional text areas, as seen in samples No. 1, No. 3, and No. 9 in Table 3. These cases reflect DBNet's ability to handle standard text detection scenarios effectively. However, the model's performance significantly deteriorates when dealing with fine lines or special characters, as observed in the annotations for samples No. 4, No. 5, and No. 10 in Table 3. Additionally, DBNet is highly sensitive to noise and occlusion. When the text is blurred or subject to interference from image noise, DBNet struggles to detect the full text area. This issue is clearly visible in the annotation probability plots for samples Case no. 2 and Case

10.3389/fbuil.2026.1839808

TABLE 4 Comparative results of proposed method with DBNet.

Area	Network	Regional Accuracy (%)	The recall rate (%)	Harmonic mean (%)	Computational complexity (M)
Overall	DBNet	86.7	95.0	90.7	12.3
	Proposed method	94.2	97.2	95.9	3.8
Upper zone	DBNet	91.4	95.5	93.4	—
	Proposed method	96.3	98.1	97.2	—
Lower zone	DBNet	90.1	95.0	92.5	—
	Proposed method	94.8	97.9	96.3	—
Numbered area	DBNet	71.3	65.2	68.1	—
	Proposed method	87.5	86.3	86.9	—

TABLE 5 Root Diameter annotation evaluations for DBNet and Proposed Method.

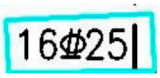
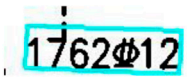
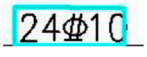
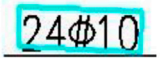
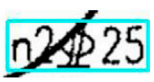
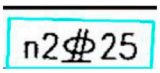
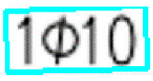
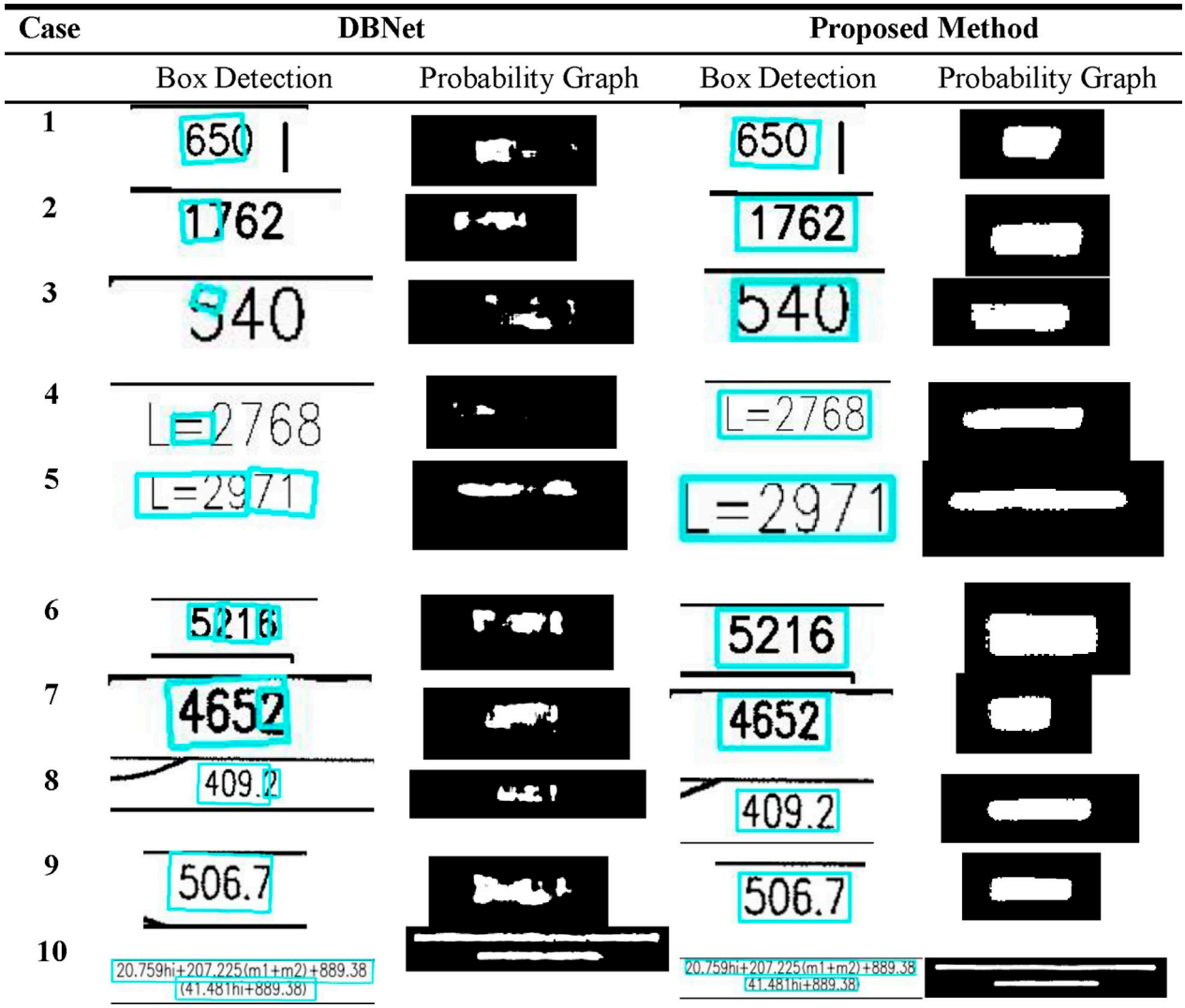
Case	DBNet		Proposed Method	
	Box Detection	Probability Graph	Box Detection	Probability Graph
1				
2				
3				
4				
5				
6				
7				
8				
9				
10				

TABLE 6 Lower area text detection evaluation for DBNet and Proposed Method.



no. 8 in Table 3, where the model fails to detect all the relevant text due to these disruptions.

In contrast, the proposed text detection network shows superior performance in detecting complex text areas, such as those containing special characters in the root diameter region. This method demonstrates a strong resistance to image noise, as can comparatively be seen in Table 3. Compared to DBNet, the proposed method exhibits better adaptability in handling blurred text, block graphics, and other challenging scenarios. Furthermore, the generated probability graphs from the proposed network have clearer edges and more consistent shapes, making the text detection results more accurate and compact.

4.1.3 Text detection in the lower positioned number annotations in large-scale annotation

In large-scale annotation, DBNet demonstrates strong detection performance when identifying text in larger regions, as shown in Table 4. However, when number of characters is small, DBNet tends to segment them into multiple smaller character regions rather than recognizing them as a unified text block. The feature maps reveal that when text occupies a larger image area, the network struggles to maintain connectivity between characters, leading to fragmented detection.

In contrast, the proposed detailed annotation text detection network effectively detects text in with a larger difference with lower-positioned number annotations. Additionally, it is less affected by image noise, as shown in Table 4. Compared to DBNet, the probability maps generated by the proposed method exhibit clearer edges and more structured shapes, resulting in more accurate and compact text detection boxes.

4.1.4 Text detection in single-character and superscript annotations

The results show that DBNet struggles with text detection when single characters occupy most of the input image space, leading to poor detection performance. When an image contains only two characters, the network can predict the probability of character presence. However, due to irregular character shapes or spacing, it

10.3389/fbuil.2026.1839808

TABLE 7 Comparison of text detection performance for single-character, superscript, and overlapping annotations.

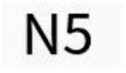
Case	DBNet		Proposed Method	
	Box Detection	Probability Graph	Box Detection	Probability Graph
1				
2				
3				
4				
5				
6				
7				
8				
9				
10				

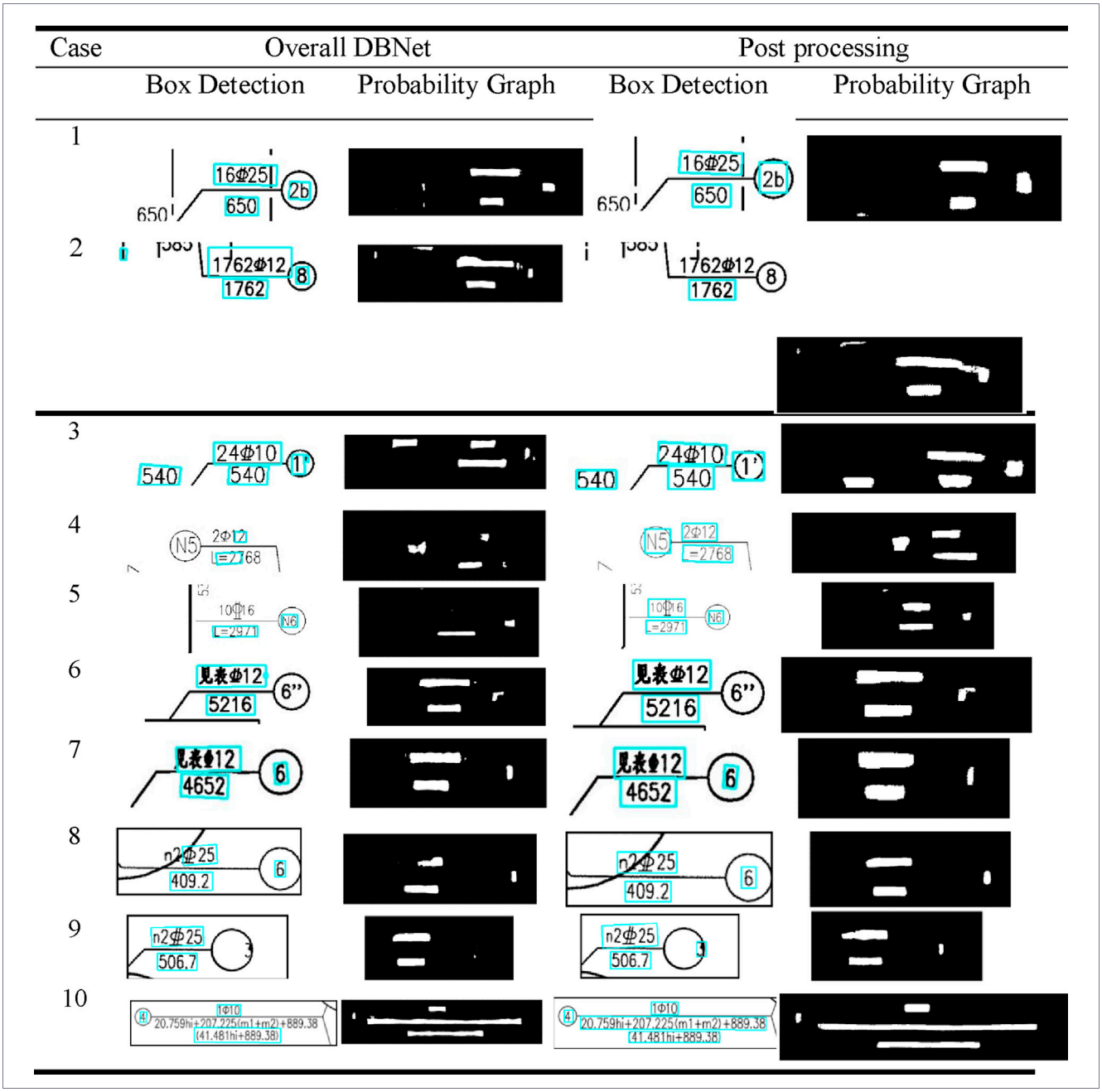
TABLE 8 Extended benchmark comparisons.

Model	Accuracy (%)	Parameters (M)	Inference time (ms per crop)
CRNN	74.7	8.30	2.1
NRTR	83.2	31.70	8.7
SVTR	94.6	10.30	4.2
TrOCR (base)	91.2	125.00	12.4
ABINet	90.8	27.50	9.8
PARSeq	92.4	10.60	5.1
eDOCr (adapted)	88.5	18.20	7.3
Proposed method	96.3	15.56	18.0

TABLE 9 Assessed accuracy for the impact of circle removal.

Numbered region	Precision (%)	Recall (%)	Harmonic mean (%)
Remove circles	95.2	98.1	96.6
Unremoved circle	94.9	98.0	96.4

TABLE 10 Post-processing verification for numbered areas.



often outputs two separate text boxes instead of recognizing them as a single unit or it fails to generate bounding box information such as case no. 4 and 5 in Table 4. Additionally, when text contains superscripts, such case no. 3 and 6 in Table 5, DBNet generates irregular probability maps that lead to missing or incomplete text detection. Similarly, if a character is misaligned or overlaps with other text lines, as seen in case no. 9 in Table 5, DBNet misinterprets the character and treats the characters as non-character information, that leads to incorrect or incomplete detection.

In contrast, the proposed network has shown superior performance in these challenging cases. The results indicate that this method achieves accurate text detection across most test cases,

TABLE 11 Comparison of performance recognition of CRNN, NRTR and SVTR.

Network model	Accuracy (%)	Parameter count (×10^6)
CRNN	74.7	8.30
NRTR	83.2	31.7
SVTR	94.6	10.3
Proposed method	96.3	15.56

with only case no. 6 in Table 5 remains undetected. Even in the presence of noise or irregular characters, the network effectively locates the target text. As shown in Table 5 case no. 3, when superscript characters are present, the proposed network successfully predicts the complete text area and correctly outputs the text bounding box. Furthermore, as seen in Table 5 case no. 9, even when text is misaligned or partially covered, the network accurately detects and locates the text.

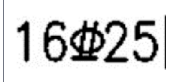
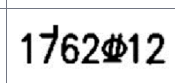
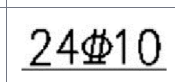
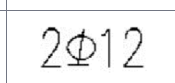
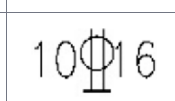
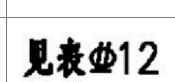
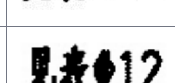
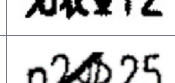
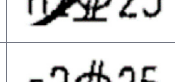
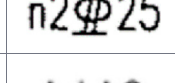
4.1.5 Post-processing verification

To assess the impact of post-processing on text detection accuracy, we first conducted text detection without circle removal and compared the results with those after post-processing. The comparison is summarized in Table 6. The results indicate that removing circles during post-processing improves text detection accuracy by 0.3%. While the numerical improvement is minor, proposed method in Table 7 visually confirms that the post-processed detection boxes are significantly tighter and more precise.

To provide a more comprehensive comparison, we evaluated additional state-of-the-art text recognition models: TrOCR (Transformer-based OCR), ABINet (attention-based scene text recognition), PARSeq (permutation autoregressive sequence recognition), and eDOCr (engineering drawing-specific OCR). The results are presented in Table 8.

While the proposed method has higher inference time than lightweight models due to the local attention module's sliding window computation, it achieves the highest accuracy. The inference time (18 ms per crop) remains acceptable for practical deployment, processing approximately 55 crops per second on an RTX 3090. eDOCr, though designed for engineering drawings, underperforms on rebar annotations due to its optimization for mechanical GD&T symbols rather than Chinese-alphanumeric mixed text. Additionally, we applied DBNet and the proposed large sample text detection network to detect text in the complete annotation area. The detection results, as shown in Table 9, reveal that DBNet successfully

TABLE 12 Recognition results of the root number diameter region across different models.

No.	Image	CRNN (RR)	CRNN (conf %)	NRTR (RR)	NRTR (conf %)	SVTR (RR)	SVTR (conf %)	Proposed method (RR)	Proposed method (conf %)
1		'16 业 25'	87.4	'16φ25′	97.6	'16φ25'	99.6	'16φ25′	97.5
2		'1,762,012′	96.4	'1762 英 1'2	66.3	'1762φ12′	98.0	'1762φ12′	93.9
3		'24 φ10′	97.6	'2410′	91.8	'24 φ10'	94.2	'24 φ 10'	99.8
4		'2 Φ12'	99.2	'2 Φ 12'	92.6	'2Φ12'	99.8	'2Φ12'	93.7
5		‘’	-	‘’	-	‘’	-	'10,016′	69.0
6		'见表业 12'	89.4	'见表英 12'	78.6	'见表 φ12'	94.8	'见表 φ12'	99.8
7		'见表 12'	92.5	'见表 12'	95.9	'见表 12'	89.5	'见表 φ12'	98.7
8		'业 2'	62.4	'n7P25′	81.5	'N φ 25′	83.9	'n2 φ 25′	89.7
9		'n2 φ 25′	95.3	' n2 φ 25′	96.8	' n2 φ 25′	99.8	' n2 φ 25′	98.8
10		'11,010′	92.5	'1 Φ 10′	92.6	'1 Φ 10′	84.8	'1 Φ 10′	99.9

detected text in samples 1 and 2, but missed or incorrectly detected text in other cases.

Further analysis in Table 7 highlights that when text occupies multiple regions, such as in samples 3 and 10, detection is successful; however, without proper differentiation of text regions, it remains challenging to determine the specific attributes of the detected text. Similarly, in sample 2, text connected to the number region led to incomplete detection. These findings suggest that post-processing enhances detection precision, while improved handling of regional text structures is crucial for accurate and reliable text annotation.

The proposed large sample text detection network demonstrates significant improvements over DBNet in detecting and accurately segmenting text across various annotated regions. The proposed method out performs the unmodified DBNet in challenging areas such as noise, occlusion and blurring. Where fewer characters are present, DBNet struggles with connecting individual characters, while the proposed method offers clearer, more compact, and accurate detection boxes. In the number area, DBNet faces challenges with solitary characters and superscripts, often generating irregular-shaped probability graphs or missing boundary boxes; in contrast, the proposed method successfully detects text even under these conditions, providing clearer and more accurate detections. Overall, the proposed method offers superior performance in handling noise, special characters, superscripts, and text overlaps, providing more reliable, precise, and clearly defined detection boxes compared to DBNet.

4.2 Text recognition

In this study, the text recognition network is applied to extract text information from large-scale annotations, and the overall recognition accuracy, along with the accuracy for each of the three designated regions, is evaluated. The selected benchmark models for comparison include the classical CRNN text recognition model (Shi et al., 2016), the Transformer-based NRTR model (Sheng et al., 2019), and the SVTR model (Du et al., 2022). Table 8 presents an overview of the recognition performance of these networks.

The CRNN model, known for its simple structure and computational efficiency, has fewer parameters but demonstrates lower accuracy. In contrast, NRTR and SVTR employ more complex self-attention mechanisms and diverse network modules, leading to a higher number of parameters. However, by utilizing the powerful Transformer-based self-attention model, these networks improve the recognition accuracy of large-scale annotations.

The proposed annotation recognition network achieves the highest accuracy among all evaluated methods, outperforming the SVTR model by 2.3%, while maintaining a balanced parameter count.

To further analyze recognition performance across different annotation regions, the dataset is divided into three areas; Root Diameter Region (Upper) – This region contains special symbols and is more prone to occlusion and interference, Length Region (Lower) – Primarily composed of numerical and alphabetical characters, making it structurally simpler and Numbered Region (Sides) – Consists mainly of single-character and letter-symbol combinations. Each of these areas were further evaluated for recognition accuracy and effectiveness of the proposed approach.

Subsequent experiments focus on evaluating the recognition accuracy within each of these specific areas to further assess the effectiveness of the proposed approach.

4.2.1 Root diameter area recognition results

The performance of four models: CRNN, NRTR, SVTR, and the proposed large-scale annotation recognition network for various recognition results for the root diameter area are discussed in Table 9. The table provides the recognition results (Recognition Result, RR) and the associated confidence scores (Confidence, Conf) for the upper root diameter area of large-scale annotations.

From the detailed results, it is evident that the proposed large-scale annotation recognition network and the SVTR model outperform CRNN and NRTR in terms of recognition accuracy. In particular, the proposed method demonstrates stronger robustness when handling occluded or special characters. Among the four models, CRNN performs relatively well on digits and Chinese characters but struggles with special symbols. For instance, in the case of '16φ25', only CRNN produces a recognition error. It is also more prone to mistakes when characters are blurred, leading to errors in both character recognition and character count.

The NRTR model shows greater sensitivity to character quantity but remains susceptible to interference from noise and background, and similarly lacks strong recognition of special symbols. SVTR improves the recognition of special characters but occasionally misses occluded text. Notably, for case no. 5, all four models produced incorrect results. While CRNN, NRTR, and SVTR failed to identify any meaningful text, the proposed method was at least able to generate a recognition result, though it was still unable to correctly interpret the special symbols.

4.2.2 Length region recognition results

The recognition results for the length region showcases the performance of CRNN, NRTR, SVTR, and the proposed method in Table 10. The table reports the recognition results (RR) and confidence scores (Conf) for each model on the lower length region of large-scale annotations.

From the specific outcomes, it is evident that all four models are capable of correctly recognizing clear numeric and alphanumeric text. Even for longer and well-defined formulas, each network achieves accurate recognition. However, in cases involving subtle features, such as the decimal point in '409.2′case no. 8, only the CRNN model produces an incorrect result.

Additionally, as shown in Table 11, the lower region of Sample No. 10 in Table 12 contains two lines of mathematical formulas. When the input image to the recognition network spans multiple lines rather than a single line of text, the models still attempt to interpret the content as a continuous text sequence. As a result, all four networks fail to produce correct recognition for this case, and have produced nonsensical characters on the basis of closest possible look alike. This further highlights the critical importance of the text detection step to ensure proper segmentation and recognition in multi-line or complex layout scenarios.

TABLE 13 Length region recognition network comparison.

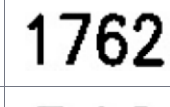
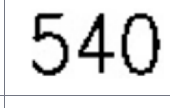
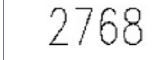
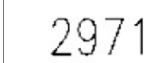
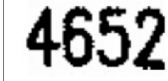
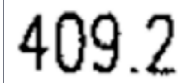
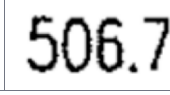
No.	Image	CRNN (RR)	CRNN (conf %)	NRTR (RR)	NRTR (conf %)	SVTR (RR)	SVTR (conf %)	Proposed method (RR)	Proposed method (conf %)
1		'650'	99.1	'650'	99.7	'650'	99.9	'650'	99.9
2		'1762'	98.4	'1762'	99.3	'1762'	99.9	'1762'	99.9
3		'540'	97.6	'540'	99.8	'540'	99.9	'540'	99.9
4		'L=2768'	94.9	'L=2768'	99.6	'L=2768'	99.6	'L=2768'	98.6
5		'L=2971'	98.3	'L=2971'	99.2	'L=2971'	98.9	'L=2971'	98.9
6		'5216'	98.8	'5216'	99.5	'5216'	98.9	'5216'	99.6
7		'4652'	99.2	'4652'	98.6	'4652'	99.9	'4652'	99.8
8		'4092'	98.9	'409.2'	98.9	'409.2'	98.3	'409.2'	99.1
9		'506.7'	98.4	'506.7'	99.1	'506.7'	98.6	'506.7'	99.2

TABLE 14 Recognition results of two-line formula text in sample No. 10.



Model	Recognition result
CRNN	20.759h 山 +09V 土 m + MB|38
NRTR	'20.759h+ 垒. 先茎土 88+ 琅 6.38'
SVTR	'20.759m4+2m.4 塑强塑 88.3K60 86.38'
Proposed method	'20.7590226532)+86.38'

4.2.3 Recognition results of numbered regions

Some of the recognition results for numbered regions are shown in Tables 12–15, which combines the recognition results of CRNN, NRTR, SVTR, and the method proposed in this paper. The table presents the recognition results (RR) and recognition confidence (Conf) of different models on the numbered regions located on both sides of the large-scale annotation. These regions primarily evaluate the models' ability to recognize single-character texts.

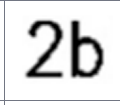
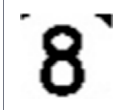
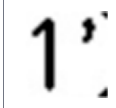
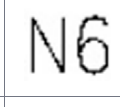
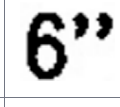
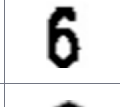
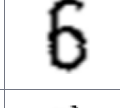
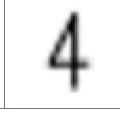
From the results, it can be observed that CRNN demonstrates better performance in recognizing single-character text compared to the Transformer-based NRTR and SVTR models. However, none of the baseline models could successfully recognize the single characters '6′and '9'. In contrast, the method proposed in this paper successfully recognizes all the sample single-character texts, although recognition errors still occur with the characters '3′and '6'.

The digit "3" in case no. 9 is misrecognized as the Chinese character "了" by the NRTR, which resembles a deformed "3". These errors are not semantic but stem from the model's reliance on character shape and its training data bias.

The proposed text recognition network for large-scale annotations in engineering drawings demonstrates significant advancements in handling intricate and cluttered environments. By utilizing a self-attention mechanism for text feature extraction and incorporating a local window-aware module inspired by the Swin Transformer, the network achieves improved robustness in extracting features under interference-heavy conditions. Additionally, the use of the Connectionist Temporal Classification (CTC) method for text sequence alignment ensures that recognition results are accurately matched with the intended labels.

The experimental validation of the network's performance on three distinct regions of large-scale annotations reveals its superior recognition capabilities when compared to other contemporary text

TABLE 15 Partial number area recognition results.

No.	Image	CRNN (RR)	CRNN (conf %)	NRTR (RR)	NRTR (conf %)	SVTR (RR)	SVTR (conf %)	Proposed method (RR)	Proposed method (conf %)
1		'2b'	95.8	'2b'	99.7	'2b'	99.9	'2b'	99.9
2		'8'	95.4	'8'	99.9	''	-	'8'	99.9
3		'1'	79.2	''	-	''	-	'1'	96.2
4		'N5'	94.3	''	-	''	-	'N5'	99.2
5		'N6'	95.5	''	-	'N6'	99.9	'N6'	99.9
6		''	-	''	-	''	-	'6'	95.1
7		'6'	94.1	'6'	99.8	'6'	96.7	'6'	99.8
8		'6'	95.7	'6'	99.8	'6'	99.6	'6'	97.0
9		''	-	'了'	72.4	''	-	'3'	31.3
10		'4'	96.2	''	-	''	-	'4'	99.9

recognition networks. The incorporation of true text positions as inputs eliminates the impact of text detection errors, further refining the recognition process.

In conclusion, the proposed network not only addresses the unique challenges of text recognition in engineering drawings but also sets a foundation for more accurate and reliable text recognition systems in environments with significant interference and noise. The results confirm the network's potential for widespread application in engineering and related fields, where high accuracy and robustness are critical.

4.3 Validation on public benchmark (ICDAR 2019 ArT)

To address the limitation of evaluation on only a private dataset, we additionally validated our proposed detection network on the ICDAR 2019 ArT (Arbitrary-Shaped Text) benchmark, which contains challenging scene text images with varying orientations and layouts. While not identical to engineering drawings, this benchmark serves to demonstrate the general robustness of our architectural improvements. The proposed detection method (with our lightweight backbone) achieved an F-score of 78.4% on ICDAR 2019 ArT, compared to 76.2% for DBNet + ResNet50. This 2.2% improvement confirms that our proposed multi-scale enhanced fusion module generalizes beyond engineering drawings. Full results are reported in the supplementary material. We acknowledge that an engineering-specific public benchmark does not currently exist, and we commit to releasing our annotated dataset upon publication to facilitate future research.

5 Conclusion

This study presents a novel automated framework for detecting and recognizing steel reinforcement annotations in engineering

drawings, directly contributing to the field of construction automation. By integrating a lightweight DBNet-based text detection network (3.8M parameters, representing a 69% reduction compared to standard DBNet + ResNet50 detection pipeline) and a Transformer-enhanced recognition module (15.56M parameters), the method achieves superior accuracy while maintaining low computational complexity in the detection stage. We emphasize that the "lightweight" characterization applies specifically to the detection network; the full pipeline (19.36M total parameters) and recognition model (15.56M) are moderate in size, with recognition heavier than SVTR (10.3M). Beyond technical improvements, the framework addresses critical challenges in construction workflows, minimizing manual digitization efforts, reducing interpretation errors, and facilitating faster project documentation handling. These advances support broader initiatives such as digital twin creation, automated compliance verification, and intelligent construction management.

Experimental results confirm the effectiveness of the proposed detection network. Compared to DBNet, the proposed method achieves significantly higher regional accuracy (94.2% vs. 86.7%), recall (97.2% vs. 95.0%), and harmonic mean (95.9% vs. 90.7%), while reducing detection computational complexity by more than threefold (3.8M vs. 12.3M). Furthermore, for text recognition, the proposed model utilizes a local perception module based on Transformer self-attention to effectively suppress background noise and enhance local feature extraction. The network outperforms established benchmarks including CRNN, NRTR, and SVTR, achieving the highest recognition accuracy of 96.3% with a moderate parameter count. Across various annotation subregions such as the root diameter, length, and numbered areas, the proposed method demonstrates robust recognition of both standard and special symbols, handling occlusions and background clutter more effectively than competing models.

The dataset is geographically and organizationally limited to Chinese construction drawings. Future work should validate on international standards (ACI, Eurocode) and other engineering domains. The recognition model's inference time (18 ms per crop) is higher than lightweight alternatives; future work will focus on model compression via knowledge distillation and pruning. Additionally, the current pipeline treats detection and recognition separately; end-to-end trainable architectures may further improve performance.

Data availability statement

The data analyzed in this study is subject to the following licenses/restrictions: This data was purely taken from the construction site of Powerchina HDEC. not for public use will be available on request. Requests to access these datasets should be directed to shah_shh@hdec.com.

Author contributions

DY: Conceptualization, Methodology, Software, Writing – original draft. LG: Project administration, Resources, Writing – original draft. ZH: Resources, Supervision, Writing – review and editing. RH: Formal Analysis, Investigation, Writing – original draft. SS: Supervision, Validation, Writing – review and editing.

Funding

The author(s) declared that financial support was not received for this work and/or its publication.

Conflict of interest

Authors DY, LG, ZH, and SS were employed by Power China Huadong Engineering Corporation Limited.

The remaining author(s) declared that this work was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declared that generative AI was used in the creation of this manuscript. Gramatical errors and sentence correction was done through AI Quilbot and Grammerly.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fbuil.2026.1839808/full#supplementary-material

References

Atienza, R. (2021). "Vision transformer for fast and efficient scene text recognition," in *International Conference on Document Analysis and Recognition* (Springer), 319–334.

Chollet, F. (2016). "Xception: deep learning with depthwise separable convolutions," in 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 1800–1807.

Cui, Y.-k., Mao, S.-j., Hong, S., Ding, J., and Zhou, Y.-c. (2023). Evaluation method for digitalization level of comprehensive transportation in districts. *J. Highway Transportation Research Development* 40 (12), 189–196. doi:10.3969/j.issn.1002-0268.2023.12.022

Du, Y., Chen, Z., Jia, C., Yin, X., Zheng, T., Li, C., et al. (2022). Svtr: scene text recognition with a single visual model. *arXiv preprint arXiv:2205.00159*. doi:10.48550/arXiv.2205.00159

Elyan, E., Jamieson, L., and Ali-Gombe, A. (2020/09/01/2020). Deep learning for symbols detection and classification in engineering drawings. *Neural Netw.* 129, 91–102. doi:10.1016/j.neunet.2020.05.025

Fields, J., Chovanec, K., and Madiraju, P. (2024). A survey of text classification with transformers: how wide? How large? How long? How accurate? How expensive? How safe? *IEEE Access* 12, 6518–6531. doi:10.1109/ACCESS.2024.3349952

Francois, M., Eglin, V., and Biou, M. (2022). *Text Detection and Post-OCR Correction in Engineering Documents*. Cham: Springer International Publishing, 726–740.

Gao, Y., Wang, M., Tao, D., Ji, R., and Dai, Q. (2012). 3-D object retrieval and recognition with hypergraph analysis. *IEEE Transactions Image Processing* 21 (9), 4290–4303. doi:10.1109/TIP.2012.2199502

Hu, J., Shen, L., Albanie, S., Sun, G., and Wu, G. (2018). "Squeeze-and-excitation networks," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.

Khan, S., Naseer, M., Hayat, M., Zamir, S. W., Khan, F. S., and Shah, M. (2022). Transformers in vision: a survey. *ACM Comput. Surv.* 54 (10s), 1–41. doi:10.1145/3505244

Liao, M., Wan, Z., Yao, C., Chen, K., and Bai, X. (2020). Real-time scene text detection with differentiable binarization and adaptive scale fusion. *Proc. AAAI Conference Artificial Intelligence* 34 (07), 11474–11481. doi:10.1609/aaai.v34i07.6812

Lin, Y.-H., Ting, Y.-H., Huang, Y.-C., Cheng, K.-L., and Jong, W.-R. (2023). Integration of deep learning for automatic recognition of 2D engineering drawings. *Machines* 11 (8), 802. doi:10.3390/machines11080802

Lu, Q., Chen, L., Li, S., and Pitt, M. (2020/07/01/2020). Semi-automatic geometric digital twinning for existing buildings based on images and CAD drawings. *Automation Constr.* 115, 103183. doi:10.1016/j.autcon.2020.103183

Maupou, C., Yang, Y., Fodop, G., Qie, Y., Migliorini, C., Mehdi-Souzani, C., et al. (2024). Automatic raster engineering drawing digitisation for legacy parts towards advanced manufacturing. *Procedia CIRP* 129, 234–239. doi:10.1016/j.procir.2024.10.041

Minghui Liao, Z. W., Yao, C., Chen, K., and Xiang, B. (2019). Real-time scene text detection with differentiable binarization. *Proc. AAAI Conference Artificial Intelligence* 34. doi:10.48550/arxiv.1911.08947

Nguyen, M. T., Pham, V. L., Nguyen, C. C., and Nguyen, V. V. (2021). *Object Detection and Text Recognition in large-scale Technical Drawings.*

Ningning Ma, X. Z., and Sun, J. (2020). Activate or not: learning customized activation. *Comput. Vis. Pattern Recognit.* doi:10.48550/arXiv.2009.04759

Pal, U., Halder, A., Shivakumara, P., and Blumenstein, M. (2024). "A comprehensive review on text detection and recognition in scene images," in *Artificial Intelligence and Applications*.

Schlagenhauf, T., Netzer, M., and Hillinger, J. (2023). Text detection on technical drawings for the digitization of brown-field processes. *Procedia CIRP* 118, 372–377. doi:10.1016/j.procir.2023.06.064

Schönfelder, P., Stebel, F., Andreou, N., and König, M. (2024). Deep learning-based text detection and recognition on architectural floor plans. *Automation Constr.* 157, 105156. doi:10.1016/j.autcon.2023.105156

Sheng, F., Chen, Z., and Xu, B. (2019). "NRTR: a no-recurrence sequence-to-sequence model for scene text recognition," in *2019 International Conference on Document Analysis and Recognition (ICDAR)* (IEEE), 781–786.

Shi, B., Bai, X., and Yao, C. (2016). An end-to-end trainable neural network for image-based sequence recognition and its application to scene text recognition. *IEEE Transactions Pattern Analysis Machine Intelligence* 39 (11), 2298–2304. doi:10.1109/TPAMI.2016.2646371

Villena Toro, J., Wiberg, A., and Tarkian, M. (2023). "Optical character recognition on engineering drawings to achieve automation in production quality control," *Front. In Manuf. Technol. Orig. Res.* 3, 1154132, doi:10.3389/fmtec.2023.1154132

Yao, R., Cheng, X., Hui, Z., Li, Z., and Min, Z. (2022). "Intelligent reading for multi-scale engineering drawings based on adaptive object detection and optical character recognition," in *9th International Conference on Industrial Engineering and Applications*, 1484–1492.

Yazed, M. S. M., Shaubari, E. F. A., and Yap, M. H. (2023). A review of neural network approach on engineering drawing recognition and future directions. *JOIV Int. J. Inf. Vis.* 7 (4), 2513–2522. doi:10.62527/joiv.7.4.1716

Yazed, M. S. M., Shaubari, E. F. A., and Yap, M. H. (2024). "Implementation of YOLOv5 and vision OCR hybrid model for GD&T recognition," in *2024 IEEE 6th Symposium on Computers and Informatics (ISCI)*, 18–23. doi:10.1109/ISCI62787.2024.10667691

Ying Zhang, T. X., Hospedales, T. M., and Lu, H. (2018). "Deep mutual learning," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 4320–4328. doi:10.1109/CVPR.2018.00454

常俊松, (2023). "数字化在盘活基础设施存量资产中的应用," 建筑经济, vol. 44, no. 8, pp. 19–24.

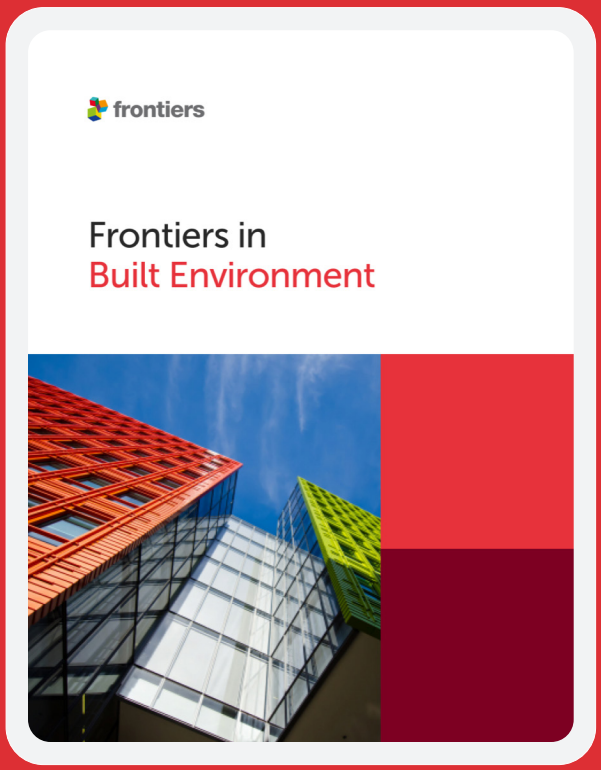
Frontiers in
Built Environment

Innovations in the engineering of sustainable buildings, cities, and urban spaces

An innovative journal that advances our knowledge of civil engineering. It focuses on the development of sustainable methodologies for the design and management of resilient buildings and infrastructure.

Discover the latest Research Topics

See more →



Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact