

REVIEW

A systematic review of computer vision-based assessment of concrete properties

Mati Ullah  · Vagelis Plevris 

Received: 6 February 2025 / Accepted: 3 February 2026

© The Author(s) 2026

Abstract

Recent advancements in image processing and convolutional neural networks have revolutionized the assessment of concrete properties. These techniques provide non-destructive and automated solutions to analyze aggregates, air voids, cement phases, and surface defects in concrete and offer significant improvements over traditional manual methods. CNNs have been developed for tasks like segmentation, classification, and quantification of concrete properties under varied conditions. This paper provides a comprehensive review of the latest developments in computer vision applications for evaluating concrete properties. It explores state-of-the-art image processing and CNN architectures being utilized in concrete properties assessment. Then, the application of these techniques to enhance the precision and reliability of concrete analysis is discussed. The review is organized into various categories based on the specific applications and advancements of CNNs in analyzing concrete properties. Challenges such as data variability, feature overlap, and imaging inconsistencies are critically discussed. The potential directions for future research are proposed, integrating advanced multimodal imaging and deep learning frameworks. This review aims to guide researchers and practitioners in advancing the field of concrete property assessment and structural health monitoring of concrete structures.

Keywords Concrete properties · Image processing · Convolutional neural networks · Non-destructive testing · Deep learning · Structural health monitoring

Abbreviations

AI	Artificial intelligence
ANN	Artificial neural networks
AS	Adaptive sampling
ASPP	Atrous spatial pyramid pooling
BSE	Backscattered Electron
CNN	Convolutional neural network
CV	Computer vision
DCNN	Deep convolutional neural network
DIA	Digital image analysis
DL	Deep learning



DSLR	Digital single-lens reflex (camera)
DT	Destructive testing
FE	Flat and elongated
FE-SEM	Field emission SEM
SEM	Scanning electron microscopy
FPN	Feature pyramid networks
GAI	Global aggregate index
GPUs	Graphics processing units
HCN	Holistically nested convolutional network
HWC	High-workability concrete
IoU	Intersection over union
ISRN	Improved structure relation network
LAI	Local aggregate index
LWAC	Lightweight aggregate concrete
mAP	Mean average precision
MAPE	Mean absolute percentage error
MVTS	Multivariate time series analysis
NDT	Non-destructive testing
RA	Recycled aggregates
RAC-NET	Recycled aggregates characterization network
ReLU	Rectified linear unit
RMSE	Root mean square error
RNN	Recurrent neural network
RoI	Region of interest
RPN	Region proposal network
SAM	Segment anything model
SCC	Self-consolidating concrete
SHM	Structural health monitoring
SI	Segregation index
SVM	Support vector machine
SVR	Support vector regression
TWS	Trainable weka segmentation

1 Introduction

Concrete is one of the most widely used building materials in modern construction because of its versatility, durability, and cost-effectiveness. The durability, serviceability, and structural performance of concrete depend mainly on the proportioning and mixing of the materials during production and exposure to the environment. Hence, an accurate prediction of concrete properties i.e., compressive strength, tensile strength, elasticity, durability, thermal conductivity, etc., is essential for structural design and Structural Health Monitoring (SHM) to have sustainable and resilient structures capable of surviving different environmental conditions, loads, and stresses [1].

Engineers and construction professionals rely on accurate predictions to ensure the safety, reliability, and efficiency of buildings, bridges, dams, and other infrastructures. Traditional methods for assessing concrete properties have often relied on manual testing and physical measurements. These methods can be time-consuming and labor-intensive [2]. This highlights the need for advanced methods to predict concrete properties, enabling

non-destructive testing. Such methods allow engineers to evaluate concrete properties without damaging the structure, supporting real-time monitoring of existing infrastructure.

Over the past few years, deep learning (DL) and image processing have found broad applications in solving classification as well as prediction problems in various scientific and engineering areas. Among many other applications, these methods can be applied to the problems related to SHM of civil infrastructures. Typically, SHM involves monitoring processes where data are collected by measuring physical parameters to build a measurable understanding of the current condition of a structure. Physical parameters may be continuously measured to monitor structural integrity in real time [3, 4]. Computer vision (CV) techniques are introduced as low-cost and reliable approaches to estimating physical parameters and monitoring the health condition of concrete structures. Image processing and Convolutional Neural Networks (CNNs) are employed in CV to make predictions about material characteristics based on information obtained from images [5–7]. The incorporation of these state-of-the-art techniques has revolutionized the field of concrete technology. High-resolution images obtained by Scanning Electron Microscopy (SEM) or digital microscopy capture intricate surface and internal structural features of concrete. Image processing techniques such as background subtraction, template matching, Canny filter, Morphological detectors, Haar transform, Sobel edge detector and texture recognition methods can be used to extract useful information from images [8–10]. This information can then be utilized to predict the properties and composition of concrete and numerous factors influencing its performance. Alternatively, CNN models can also be trained using pre-processed images. These neural networks are able to extract information that is essential for estimating physical characteristics because they can recognize complicated patterns and correlations in images [11]. The CNN models can generalize and accurately predict various properties based on new input images when trained on large datasets of labeled concrete images.

The application of CV in civil engineering has been explored in several review papers, with a focus on construction site monitoring and inspection [12–16]. Jiang et al. [17] carried out a comprehensive review on detecting and retrofitting tunnel linings through image processing. The article emphasized the importance of identifying deteriorations like cracks, leakage, and voids to ensure safety and extend the service life of tunnels. Alsakka et al. [18] conducted a scoping assessment on CV applications in offsite construction, highlighting areas that require more research and reviewing the tasks, algorithms and limitations in this field. Meanwhile, Ma et al. [19] reviewed automated quality inspection methods for precast concrete components. The authors focused on dimensional and surface quality inspections as well as improvements in data collection. Deng et al. [20] reviewed CV-based solutions for fracture analysis in civil infrastructures. Baduge et al. [21] examined the uses of artificial intelligence (AI), and machine learning in the construction sector 4.0, exploring building management, offsite manufacturing, and structural design. The paper also looked at how sensors and smart vision can be used to collect data. Jayaram [22] reviewed a wide range of applications for CV in civil engineering, encompassing domains such as infrastructure inspection and SHM. Mostafa and Hegazy [23] examined image-based analytic methods in the construction industry, reviewing more than 100 scientific articles to find important applications in damage assessment, progress monitoring and safety.

This review paper aims to provide a comprehensive overview of the state-of-the-art in concrete properties prediction using image processing and CNNs. It will encompass a wide range of properties including characteristics of aggregates, cement paste, surface bugholes, and internal air voids. Additionally, the paper discusses the prediction of mechanical properties of concrete in both laboratory settings and real-world construction scenarios, highlighting the potential application in SHM of concrete. The scope of this review paper encompasses a comprehensive analysis of existing literature, case studies, and technological advancements related to the use of image processing and CNNs in predicting various concrete properties. By synthesizing this information, the paper aims to provide a valuable resource for researchers, practitioners, and professionals in the construction and engineering domains.

The work is organized as follows: Sect. 2 explains the review process, procedures involved in data collection and processing and the bibliometric analysis. Sections 3 and 4 provide a comprehensive analysis and discussion with a focus on describing datasets, image processing methods, CNN models, and their applications in concrete

technology. Section 5 discusses the key observations, challenges in using CNNs for the evaluation of concrete properties and future research methods to address these shortcomings. Section 6 provides the final conclusions of the review paper.

2 Review methodology and bibliometric analysis

We used Scopus to identify relevant studies and conduct a bibliometric assessment of the literature on the application of image processing and CNNs in predicting concrete properties. Scopus was selected due to its comprehensive coverage of high-quality and respected journals in the fields of DL, image processing, and concrete technology. Figure 1 provides an overview of the review methodology. First, the data were gathered by retrieving papers from the Scopus database, using a thorough search with keywords related to image processing, CNN, CV and concrete properties.

The search query used for the review, covering the period from January 2015 to October 2024, was as follows: (*“concrete”*) AND ((*“image processing”* OR *“computer vision”* OR *“digital image analysis”* OR *“machine vision”*) OR (*“deep learning”* OR *“CNN”* OR *“convolutional neural network”* OR *“neural network”*)) AND (*“segmentation”* OR *“classification”* OR *“detection”* OR *“aggregate characteristics”* OR *“cement”* OR *“air voids”* OR *“bughole”* OR *“segregation”* OR *“mechanical properties”* OR *“porosity”* OR *“permeability”* OR *“concrete properties”*). Papers with these terms in the title, abstract, or keywords were included. To ensure the reliability and quality of the sources, only English-language papers from peer-reviewed academic journals were considered. The search results were further refined by excluding book chapters, editorials, conference proceedings, and other unrelated sources, and irrelevant articles were filtered out. A manual review of the search results was conducted to eliminate any remaining irrelevant papers. This process led to the identification of 82 journal articles deemed suitable for analysis.

Additionally, a bibliometric analysis was performed to examine trends, influential publications, and collaborations in the field. The results and discussion sections provide a detailed exploration of the image processing and DL techniques used in the reviewed studies, their application in assessing concrete properties, challenges in implementing CNNs, and recommendations for future research directions based on identified gaps in the literature.

2.1 An overview of the sampled literature

The quantitative analysis of CV for concrete properties assessment highlights significant trends and advancements in this research domain. Figure 2 shows the annual number of publications related to CV-based concrete properties assessment over the period from January 2015 to October 2024, revealing a consistent increase in interest and the number of related studies. This upward trend reflects the growing integration of image processing and CNNs in material characterization and quality evaluation of concrete.

The geographical distribution of studies demonstrates that most of the DL-based research originates from a few key countries/regions. Figure 3 represents the five high-frequency contributors, accounting for more than 90% of relevant research. Notably, China, USA, South Korea, Turkey and England are recognized as leading contributors, with frequent collaborations between China and other countries. This collaboration highlights the global interest in using image processing and CNN-based methods for concrete analysis.

2.2 Keywords analysis

Keywords serve as concise indicators of the core content of an article, enabling an understanding of the research scope and focus. In this review, we conducted a keyword-based analysis to map the knowledge domains and explore their interrelationships. Using co-occurrence network visualization, we identified research hotspots in the

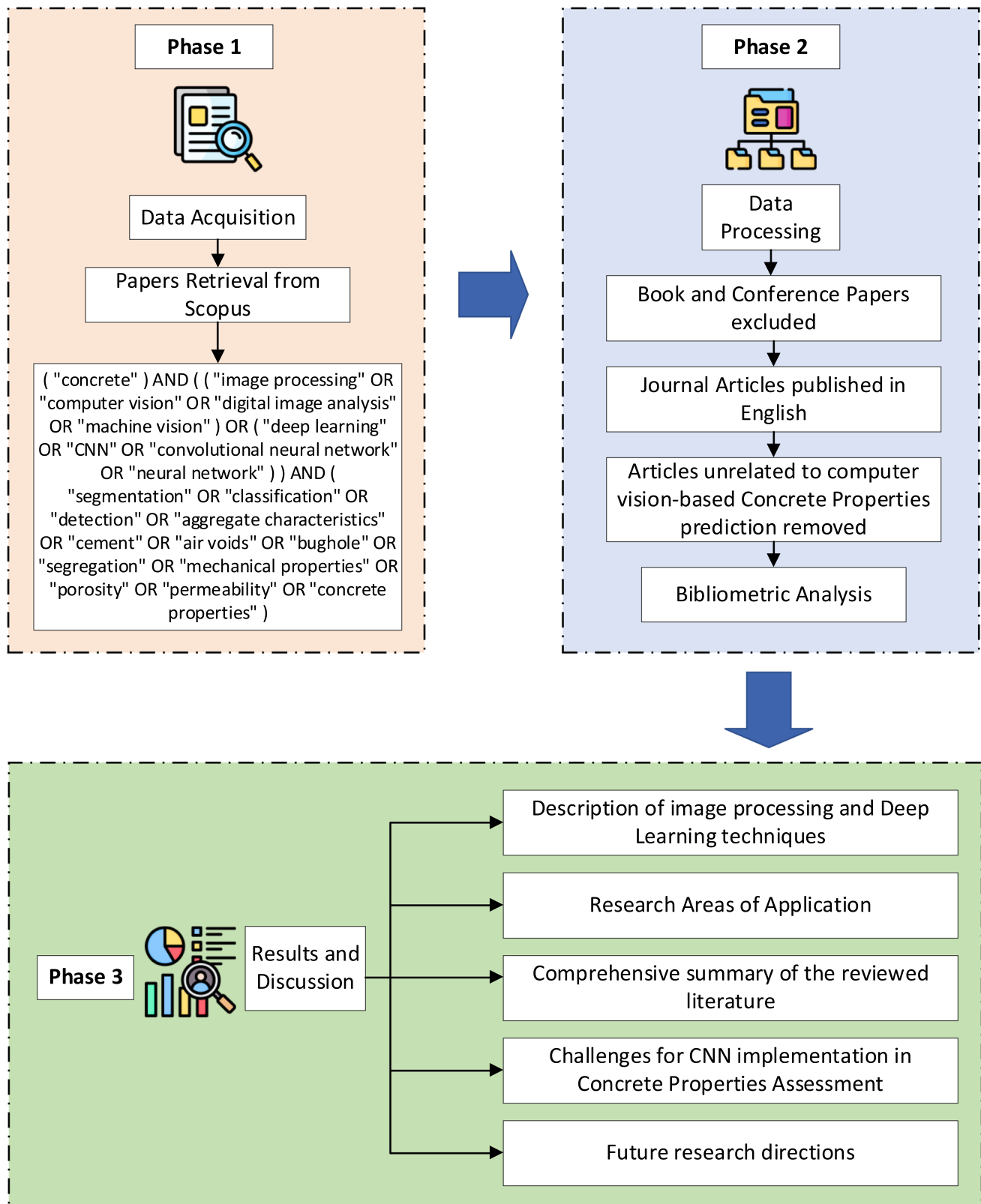


Fig. 1 Review methodology

Fig. 2 Number of publications during the 10-year period 2015–2024

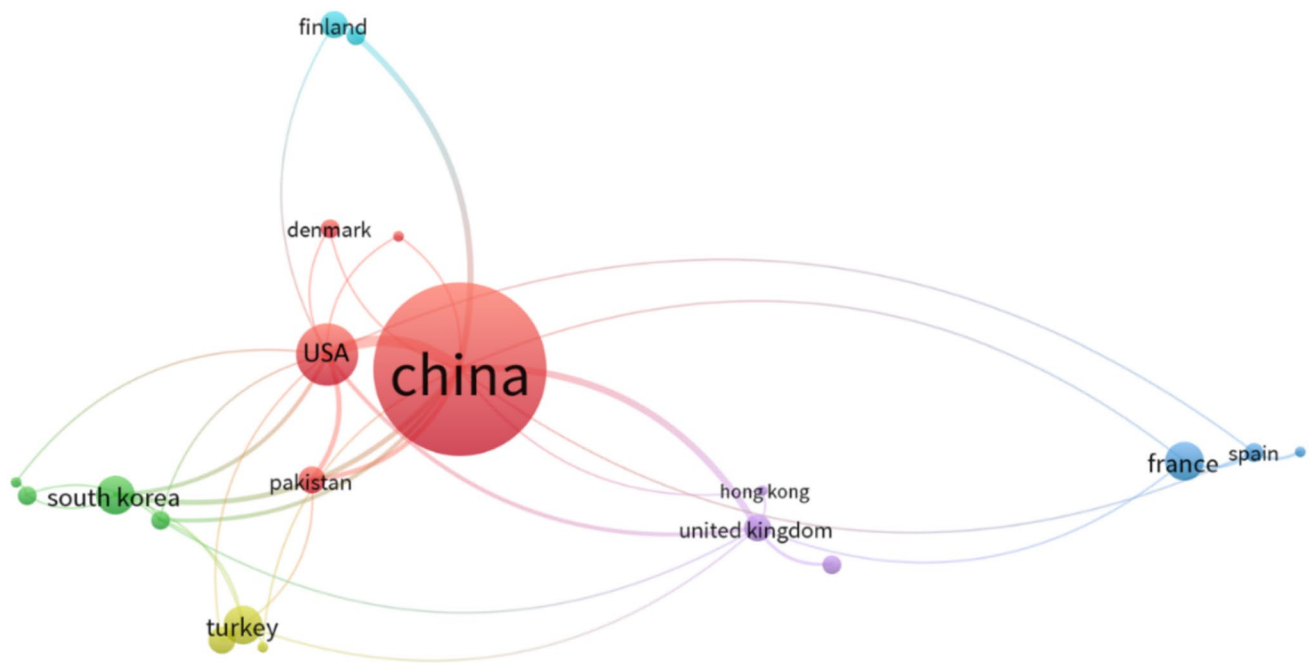
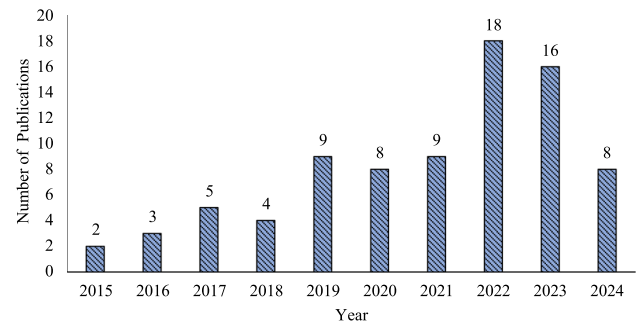


Fig. 3 Co-occurrence network mapping of countries with high publication frequency

application of CV for concrete properties assessment. Keywords with a frequency of three or more occurrences were included in the analysis. As shown in Fig. 4, the font size represents the frequency of keywords, indicating their importance, while the thickness of connecting lines signifies the degree of correlation between them. Larger font sizes highlight frequently studied topics, and thicker lines indicate stronger relationships among research themes.

The keyword mapping analysis highlights the growing interdisciplinary nature of research in concrete properties assessment by integrating image processing and CV. By examining trends in frequently co-occurring keywords, this review identifies a strong focus on segmentation tasks, air void quantification, and mechanical property prediction, with CNNs serving as the dominant algorithmic framework. These trends emphasize the significant role of CV techniques in enhancing the precision, efficiency, and automation of concrete analysis processes. We analyze the keyword network from several dimensions: image analysis targets, concrete properties, DL methods, and applications.

A detailed examination of the sample literature reveals seven major applications of image processing and CNNs in the field of concrete technology, summarized in Fig. 5. The primary focus areas in the literature include the segmentation and characterization of aggregates, air voids, surface bugholes, cement particles and hydration analysis, segregation, and predicting mechanical properties of concrete.

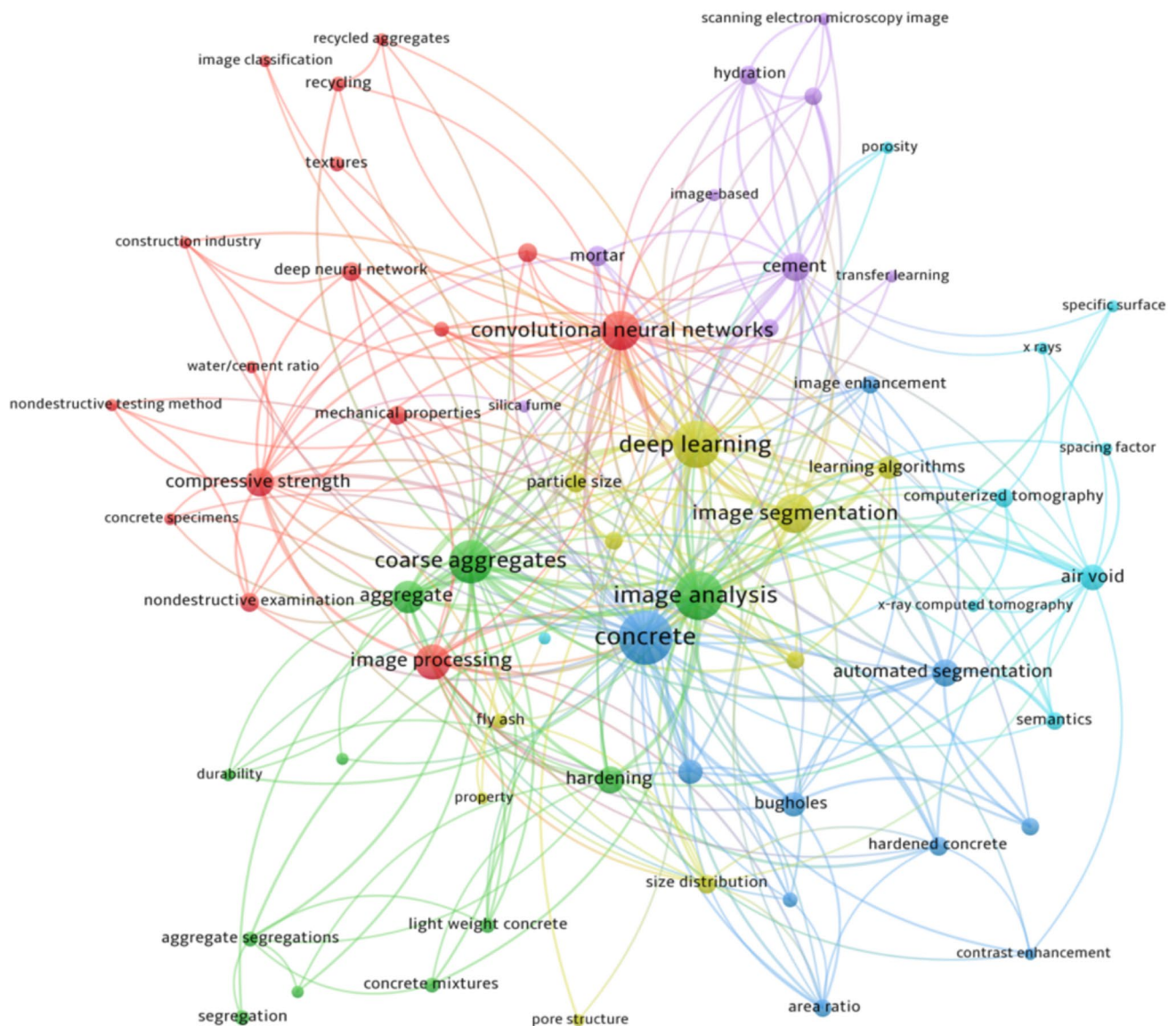


Fig. 4 Co-occurrence network mapping of keywords

3 Background on image processing and CNN techniques

Gaining insight into image processing and CNN techniques is essential for providing readers with a foundational knowledge of these technologies and their significance in analyzing concrete properties. Image processing lays the groundwork by preparing and enhancing raw images to extract meaningful features. CNNs leverage these processed images to perform complex tasks like segmentation and prediction with high accuracy. Integration of these methods effectively addresses challenges like noise, uneven lighting, and intricate feature extraction in concrete analysis.

3.1 Image processing

Analysis of concrete using image processing follows a technical approach to determine and measure different features. This procedure consists of several steps, each customized to address certain challenges associated with

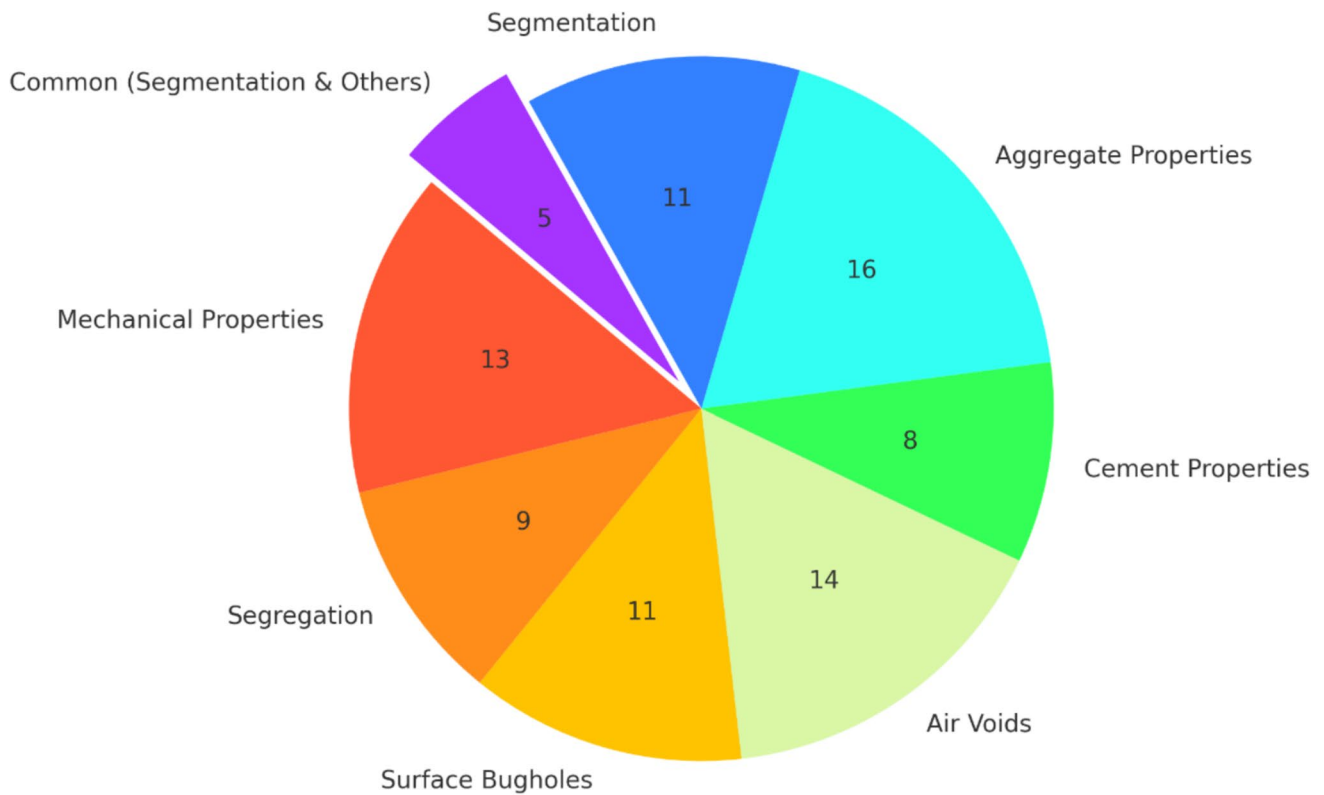


Fig. 5 Categories-wise details of the selected sample literature

concrete imaging such as variations in lighting conditions, surface texture and artifacts of noise. The success of subsequent analyses depends on the careful execution of these steps starting with the preparation of raw images to ensure accuracy and clarity.

3.1.1 Image pre-processing

Concrete surface images often exhibit artifacts like shadows, noise, and uneven lighting, reducing clarity and analysis accuracy. Image pre-processing techniques enhance image quality while retaining critical features such as aggregates and air voids. Among these techniques, spatial domain filtering [24, 25], including mean, median, and morphology filtering techniques are particularly effective for processing large datasets. Mean Filtering [26] smooths Gaussian noise efficiently but may blur edges, potentially losing features like cracks and air voids. Median filtering [26] is a nonlinear approach that removes impulse noise while preserving edges, ideal for detecting low grayscale features. Morphology filtering [27] detects features and removes "salt and pepper" noise through operations like opening (erosion-dilation) and closing (dilation-erosion) to enhance image connectivity and fill gaps. These methods collectively improve image quality for accurate analysis of concrete features.

3.1.2 Image enhancement

Image enhancement techniques are applied after filtering to address the blurring and indistinct edges caused by grayscale blending. These methods improve contrast and edge details, enabling clearer identification of features such as aggregate boundaries and pores. Grayscale transformation [28] adjusts the intensity range to highlight target features like aggregates and air voids, enhancing visibility for better segmentation and analysis. Similarly, histogram equalization [29] revises grayscale intensities, increasing detail visibility and improving feature

extraction by enhancing the contrast of concrete images. These techniques collectively ensure that critical features are distinguishable from the surrounding matrix, facilitating accurate image analysis.

3.1.3 Segmentation

Image segmentation is a crucial step following noise reduction and contrast enhancement to identify key features in concrete images. Various methods are used for segmentation. Edge detection [30] highlights boundaries by detecting intensity variations using mathematical operators. The Roberts Operator calculates gradients diagonally, the Prewitt Operator detects gradients horizontally and vertically, while the Laplacian Operator efficiently identifies edges by determining intensity contrasts [31]. The morphological techniques [32] like dilation, erosion, opening, and closing enhance image structures by removing noise, improving object shapes, and filling gaps. They are often combined with edge detection or thresholding for better segmentation accuracy. The threshold segmentation method [33] distinguishes features by setting an intensity threshold. Pixels below the threshold are classified as background, while those above represent aggregates or voids. Global thresholding works for evenly lit images, while adaptive thresholding adjusts for uneven lighting to improve precision. The Watershed algorithm [34] separates overlapping objects by analyzing the image as a topographic map. It identifies ridges (high-intensity edges) and basins (low-intensity regions) using a flooding process, effectively defining boundaries between adjacent objects.

3.1.4 Features extraction techniques

The Feret rectangle method [35] is the most effective image processing technique for analyzing the shape and size of objects employing smaller area bounding rectangles. It is based on Feret diameter, which is defined as the greatest distance between any two points at the boundary of a particle in order to measure rectangle sizes. Especially for calculating the aggregate sizes, this method gives both the maximum and minimum Feret diameters. Both these diameters are required for the accurate measurement of elongated and irregular shapes. The Euclidean distance calculation [36] is another spatial analytical approach that features in nearest-neighbor computations and it is applied to examine the image clustering patterns and spatial organization. It computes straight line distance between the centroids of neighboring aggregates and infers valuable quantitative information concerning their spatial relationships. Without these measurements, improving the understanding of aggregate distribution, which significantly influences concrete properties like strength, workability and susceptibility to segregation would be impossible. After segmentation, the centroids for all aggregate particles are measured and the distance between each pair of particles is calculated. Since each particle has a unique nearest-neighbor distance, this measure is equivalent to computing the shortest distance to any other centroid. This gives crucial information about the spatial consistency and clustering patterns of the aggregate distribution.

3.2 Convolutional neural networks

Convolutional Neural Networks (CNNs) are a specialized DL architecture designed for image-based tasks, known for their ability to automatically learn hierarchical spatial features [37]. CNNs consist of multiple layers, including convolutional layers for feature extraction, pooling layers to reduce dimensionality, activation functions to introduce non-linearity, and fully connected layers for classification. Figure 6 represents the architecture of a typical CNN with convolutional, pooling and fully connected layers. The convolutional layer, a core component of CNNs, applies filters (kernels) that slide over the input image, performing element-wise multiplication and summation to produce feature maps. These maps emphasize key image characteristics, such as edges, textures, or complex shapes. While earlier layers capture low-level features like edges, deeper layers learn high-level patterns crucial for classification. Convolutional operations also minimize the number of learnable parameters by applying filters to local patches, preserving spatial information. Key hyperparameters such as kernel size, stride, and

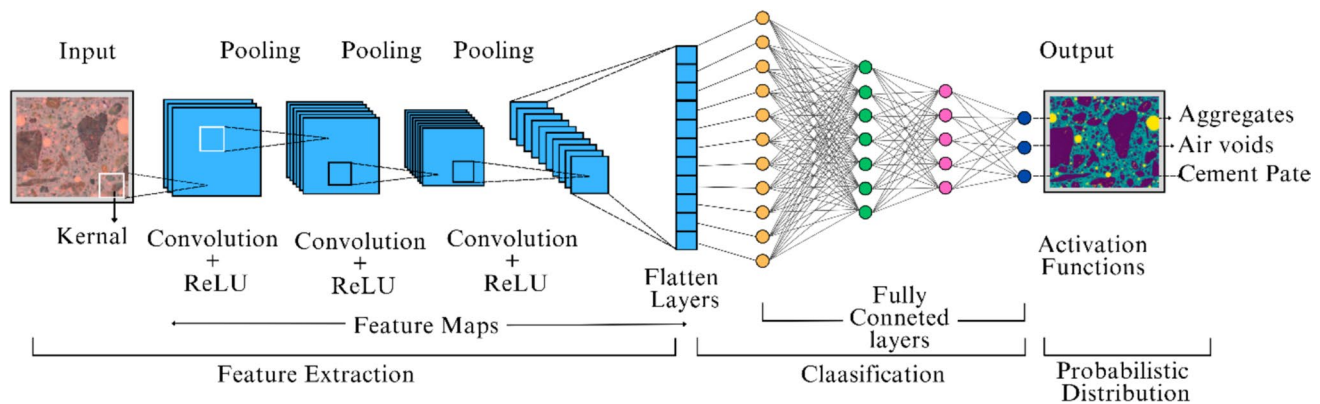


Fig. 6 A typical CNN architecture with convolutional and pooling layers

padding significantly affect the resolution of the resulting feature maps. The pooling layers follow convolutional layers to downsample feature maps, retaining essential features while reducing spatial dimensions and computational complexity. Common pooling techniques include max-pooling, which selects the maximum value within a region, and average-pooling, which computes the average value. These methods enhance translational invariance, enabling CNNs to recognize patterns regardless of their location within the input image.

To introduce non-linearity, activation functions such as the Rectified Linear Unit (ReLU) are applied to the outputs of convolutional layers. ReLU replaces negative values with zero, accelerating training and mitigating vanishing gradient issues. After convolutional and pooling layers, the output is flattened into a vector and passed through one or more fully connected layers. These layers aggregate the extracted features and map them to class probabilities using functions like Softmax, which normalizes scores for multi-class classification. The training of CNNs involves forward propagation to compute predictions and backpropagation to update the network parameters by minimizing the prediction error. During forward propagation, the input image passes through successive layers, and the network generates an output. A loss function, such as cross-entropy loss, measures the discrepancy between the predicted output and the actual label. Backpropagation then computes the gradient of the loss with respect to each weight using the chain rule and propagates these gradients backward through the network. Optimization algorithms like Stochastic Gradient Descent and Adam use these gradients to update the weights efficiently. Due to their ability to share weights across spatial regions and reduce the number of parameters, CNNs are highly efficient at learning rich and hierarchical feature representations. This makes them especially well-suited for image-based tasks, often outperforming traditional fully connected neural networks.

The computationally intensive training process has been accelerated by Graphics Processing Units (GPUs) and frameworks like TensorFlow, PyTorch, and Keras, enabling parallel processing. These capabilities make CNNs particularly well-suited for analyzing concrete images. They are extensively used for tasks such as aggregate segmentation, air void detection, and surface defect analysis. CNNs extract hierarchical spatial patterns from high-resolution images acquired via X-ray Computed Tomography (XCT), Scanning Electron Microscopy (SEM), or digital cameras, facilitating automated defect detection, material characterization, and accurate predictions of concrete mechanical properties [38, 39].

3.2.1 Object detection and segmentation CNN models

Object detection and segmentation models are essential for identifying, localizing, and analyzing objects in images. These models not only detect objects but also segment them at the pixel level, providing detailed insights into their structure and features. By leveraging advancements in DL, these architectures achieve remarkable accuracy and efficiency, making them highly applicable in fields such as medical imaging, autonomous systems, and material analysis. Mask R-CNN [40] is a popular DL model for object detection, localization, and pixel-level

segmentation. It combines a Region Proposal Network (RPN) for region generation, a feature extraction backbone like ResNet with Feature Pyramid Networks (FPN) for multi-scale features, and region of interest (RoI) Align to maintain spatial accuracy. Its precision makes it ideal for applications in material analysis, autonomous vehicles, and medical imaging.

Faster R-CNN [41] extends Mask R-CNN by introducing a learnable RPN for efficient region proposal generation, replacing traditional methods. It uses a two-stage approach: RPN identifies regions, and the second stage classifies and refines bounding boxes. Its end-to-end training and compatibility with backbones like ResNet [42] and VGGNet [43] make it a standard in object detection. U-Net [44] is an encoder-decoder CNN for segmentation, widely used in biomedical imaging. The encoder captures context through convolution and pooling layers, while the decoder restores spatial details using up-sampling. Skip connections merge local and contextual features, enabling high-resolution segmentation even with limited datasets. DeepLabv3+ [45] is a semantic segmentation model that enhances boundary accuracy with Atrous Spatial Pyramid Pooling (ASPP) and a lightweight decoder. ASPP captures multi-scale spatial details, while depthwise separable convolutions reduce computational complexity, making it efficient for detailed segmentation in real-world applications.

The YOLO family of object detection models [46] has become one of the most prominent and widely adopted frameworks for real-time object detection and segmentation due to its exceptional speed and efficiency. Unlike region-based methods such as Mask R-CNN or Faster R-CNN, YOLO processes the entire image in a single pass, predicting bounding boxes and class probabilities simultaneously, making it highly suitable for time-sensitive applications [47]. Recent versions, particularly YOLOv9, YOLOv10, and YOLOv11 [47], have demonstrated significant improvements in detection accuracy, inference speed, and scalability. These advancements are driven by the integration of neural architecture search, quantization-aware training, and transformer-based components, enhancing their performance for edge deployment and resource-constrained environments. In concrete-related applications, these models have been increasingly employed for detecting surface defects like bugholes, classifying aggregates in real time on conveyor belts, analyzing air voids, and monitoring concrete structures for deterioration. Given their evolving architecture and balance of accuracy, speed, and resource efficiency, the YOLO family offers significant promise for advancing automated inspection and real-time structural health monitoring in the construction industry.

3.2.2 Backbone architectures for feature extraction

Segmentation models often rely on powerful backbone architectures for feature extraction, which serve as the foundation for various CV tasks. These backbones process raw input data to extract meaningful features, which are then utilized by more advanced layers for tasks like object detection, segmentation, and classification. Backbones are designed to balance computational efficiency with accuracy, enabling robust performance across a wide range of applications. AlexNet [48] is a breakthrough architecture that demonstrated the power of DL for large-scale image recognition. It uses stacked convolutional layers for feature extraction, along with ReLU activation, max pooling, and dropout to improve performance and prevent overfitting. AlexNet laid the foundation for modern DL architectures and remains influential in feature extraction tasks. VGGNet [43] simplifies DL architectures with sequential 3×3 convolutional layers, enabling deeper networks without excessive complexity. Variants like VGG-16 and VGG-19 are widely used for tasks such as ImageNet classification. VGGNet excels in extracting hierarchical features and recognizing fine details, making it a common choice for foundational tasks in feature extraction. The Inception [49] network employs parallel convolutions at different scales, including 1×1 , 3×3 , and 5×5 , to capture both local and global features. Dimensionality reduction via 1×1 convolutions enhances computational efficiency. Later versions like Inception-v3 and Inception-v4 include additional optimizations like factorized convolutions and auxiliary classifiers, making them effective for diverse tasks in feature extraction and image classification.

Building on these foundational backbones, other architectures have introduced innovative approaches to feature extraction and efficiency optimization, further expanding their applicability in CV. Xception [50] builds upon

the Inception architecture by replacing standard convolutions with depthwise separable convolutions, enhancing computational efficiency while retaining high accuracy. It is used in tasks like semantic segmentation and is often combined with models like DeepLabv3+. ResNet [42] is an important backbone that introduces residual connections to enable effective training of very deep neural networks by mitigating the vanishing gradient problem. Residual mappings simplify optimization and improve accuracy. Variants like ResNet-50 and ResNet-101 are widely used backbones in advanced models such as Faster R-CNN and Mask R-CNN, supporting applications in object detection, segmentation, and classification. DenseNet [51] connects each layer directly to every other layer, promoting maximum feature reuse and minimizing redundant computations. Its dense connectivity structure enhances gradient flow and reduces the number of parameters, making it highly efficient for deep networks. EfficientNet [52] is a family of models that balances accuracy and efficiency using compound scaling, which simultaneously adjusts the network's depth, width, and input resolution. The original EfficientNet versions (B0 to B7) use mobile inverted bottleneck (MBConv) layers and depthwise separable convolutions to achieve high accuracy with fewer parameters. These models have been employed in concrete image analysis tasks such as aggregate and air void segmentation due to their lightweight design and strong generalization capabilities. Building upon this, EfficientNetV2 [53] was introduced with further enhancements in training efficiency, parameter utilization, and accuracy. EfficientNetV2 utilizes a combination of fused MBConv and regular convolutions for faster training and better performance on smaller datasets, which is particularly relevant for concrete imaging where annotated datasets are often limited.

3.3 Evaluation metrics for CNN-based model assessment

To effectively assess and compare the performance of CNNs and image processing models used in concrete property evaluation, a variety of performance metrics are employed. These metrics provide quantitative insights into a model's ability to accurately segment, classify, or detect features such as aggregates, voids, cracks, or surface defects. Among the most commonly used metrics in the reviewed literature are Accuracy, Precision, Recall, F1-score, mIoU, mAP, and RMSE [54, 55], which are especially relevant in classification and segmentation tasks. Below is a detailed explanation of each.

3.3.1 Accuracy

Accuracy is the most intuitive and widely used evaluation metric. It measures the proportion of correctly predicted outcomes (both positive and negative) among the total number of predictions made by the model. Accuracy provides an overall measure of correctness. However, it can be misleading when applied to imbalanced datasets, common in concrete images where background pixels vastly outnumber feature pixels. In such cases, a model may achieve high accuracy simply by predicting the majority class. It can be mathematically expressed as

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

where TP (True Positive) refers to correctly predicted positive instances (e.g., correctly identified aggregate pixels), TN (True Negative) denotes correctly predicted negative instances (e.g., correctly identified background pixels), FP (False Positive) represents incorrectly predicted positive instances (e.g., misclassified background or noise as aggregate), and FN (False Negative) indicates missed positive instances (e.g., actual aggregate pixels incorrectly classified as background).

3.3.2 Precision

Precision is the ratio of correctly predicted positive observations to the total predicted positive observations. It reflects the model's ability to avoid false positives. Precision is important when false positives can lead to costly

or incorrect decisions, such as overestimating air void content or misidentifying bugholes on a surface. A high precision value means that the model produces fewer incorrect positive predictions, indicating better reliability in feature identification. The mathematical formulation is as follows:

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

3.3.3 Recall

Recall, also known as sensitivity or true positive rate, measures the model's ability to correctly identify all relevant instances of the positive class. High recall ensures that the most relevant features (e.g., air voids, aggregates) are detected. This is particularly important in safety-critical applications like structural health monitoring, where missing a defect could have severe consequences. A low recall means the model is failing to identify significant portions of the target feature. It is defined by the following equation:

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

3.3.4 F1-score

F1-score is the harmonic mean of precision and recall. It provides a balanced metric that considers both false positives and false negatives. The F1-score is particularly useful when there is an uneven class distribution or when a balance between precision and recall is desired. In concrete analysis tasks, such as segmentation of aggregates or air voids, the F1-score provides a single metric that summarizes both the model's accuracy in identifying features and its tendency to generate false alarms. The metric is calculated using the equation:

$$F1 \text{ Score} = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall} \quad (4)$$

3.3.5 Mean intersection over union

Mean Intersection over Union (mIoU) is a critical metric for evaluating semantic segmentation models. It measures the average overlap between the predicted segmentation mask and the ground truth mask across all classes. mIoU is one of the most commonly reported metrics in concrete image segmentation tasks, such as detecting aggregates, air voids, or cement paste phases. It provides a balanced measurement of how well the model segments each class. A higher mIoU indicates that the predicted areas closely match the ground truth, making it ideal for evaluating pixel-level accuracy in multi-class or binary segmentation problems. Mathematically, it is given as

$$IOU = \frac{TP}{TP + FP + FN} \quad (5)$$

$$mIOU = \frac{1}{N} \sum_{i=1}^N IOU_i \quad (6)$$

where N is the number of classes.

3.3.6 Mean average precision

Mean Average Precision (mAP) is a standard metric for evaluating object detection models. It combines precision and recall across different intersection over union (IoU) thresholds to assess how well a model can localize

and classify multiple objects. mAP is highly useful in concrete applications involving object detection, such as identifying aggregates, voids, or surface bugholes. It is especially critical for instance segmentation tasks where a model needs to not only identify the presence of an object but also accurately locate and delineate its boundaries. mAP@0.5 (IoU threshold of 0.5) and mAP@0.75 are commonly reported variants in reviewed literature. This metric can be quantified using the following expression:

$$AP = \int_0^1 Precision(r) dr \quad (7)$$

$$mAP = \frac{1}{N} \sum_{i=1}^N AP_i \quad (8)$$

where r is recall, N is the number of classes.

3.3.7 Root mean square error (RMSE)

Root Mean Square Error (RMSE) is a regression metric that measures the average magnitude of prediction errors. It calculates the square root of the average squared differences between predicted and actual values. RMSE is often used in concrete research when predicting physical properties such as aggregate size, gradation, cement content, or compressive strength. It penalizes larger errors more than smaller ones, making it useful when outliers significantly impact model performance. It is mathematically represented as:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2} \quad (9)$$

where $\hat{y}_i$ is the predicted value and y_i is the true value.

3.3.8 Mean absolute percentage error

Mean Absolute Percentage Error (MAPE) expresses the prediction accuracy as a percentage. It calculates the average absolute percentage difference between predicted and true values. MAPE is commonly used in applications where interpretability in percentage terms is desirable, such as comparing estimated and actual material proportions or performance indicators. It is also employed in models predicting mechanical properties and micro-structural distributions. Mathematically, it is given as

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{\hat{y}_i - y_i}{y_i} \right| \times 100 \quad (10)$$

where $\hat{y}_i$ is the predicted value and y_i is the true value.

4 Application of image processing and CNN for assessment of concrete properties

Image processing and CNNs have revolutionized the assessment and maintenance of concrete structures. These techniques have improved accuracy and efficiency by automatically detecting and quantifying concrete features such as aggregates, air voids, surface bug holes, and surface defects like spalling and delamination. Figure 7 represents an overview of CV-based methods for various applications in concrete properties assessment. Image processing and CNNs analyze images to identify patterns and features, enabling the assessment of different concrete properties. In the following sections, this article comprehensively analyzes the segmentation of concrete

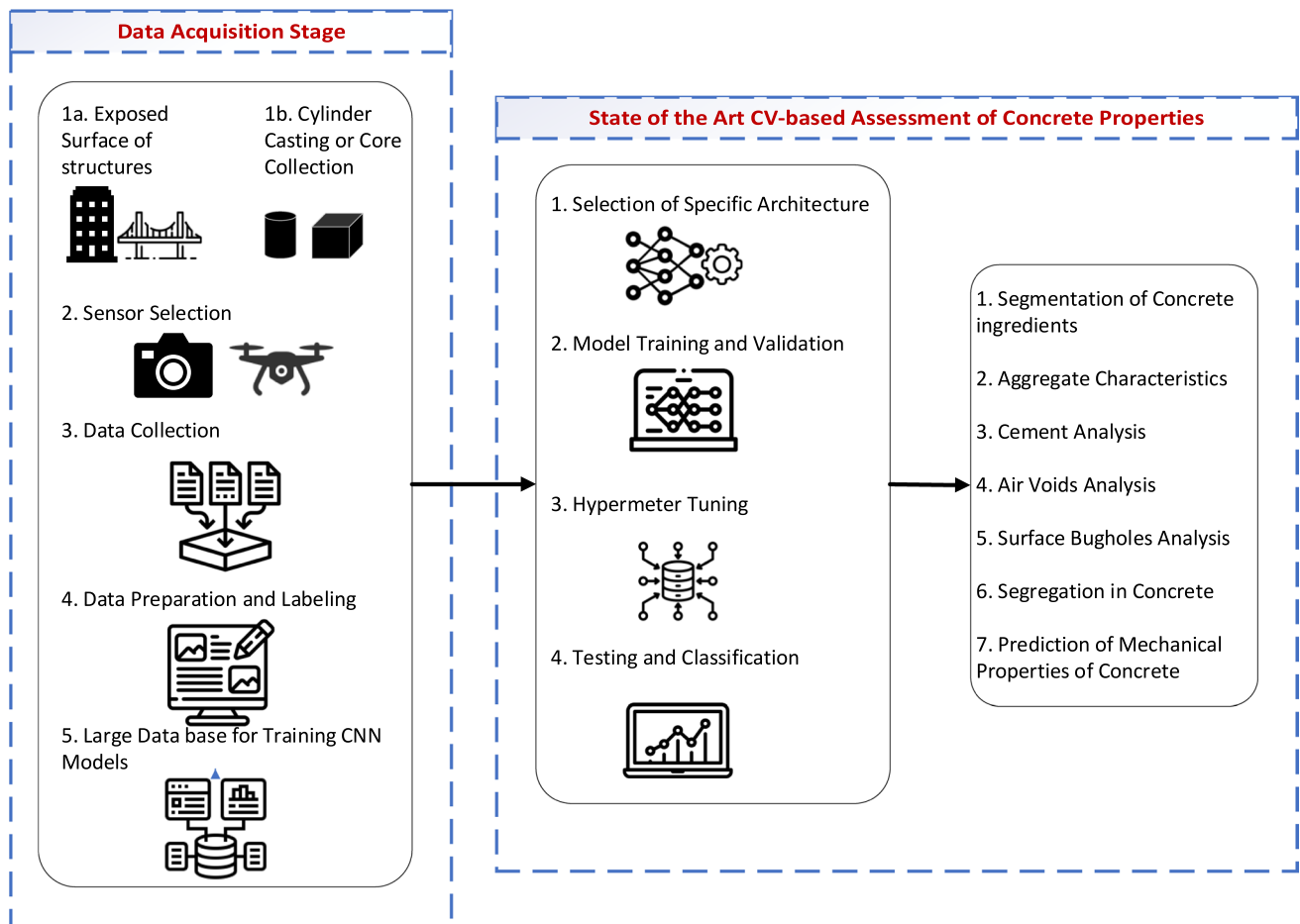


Fig. 7 A schematic view of the state-of-the-art CV-based operations in concrete properties assessment

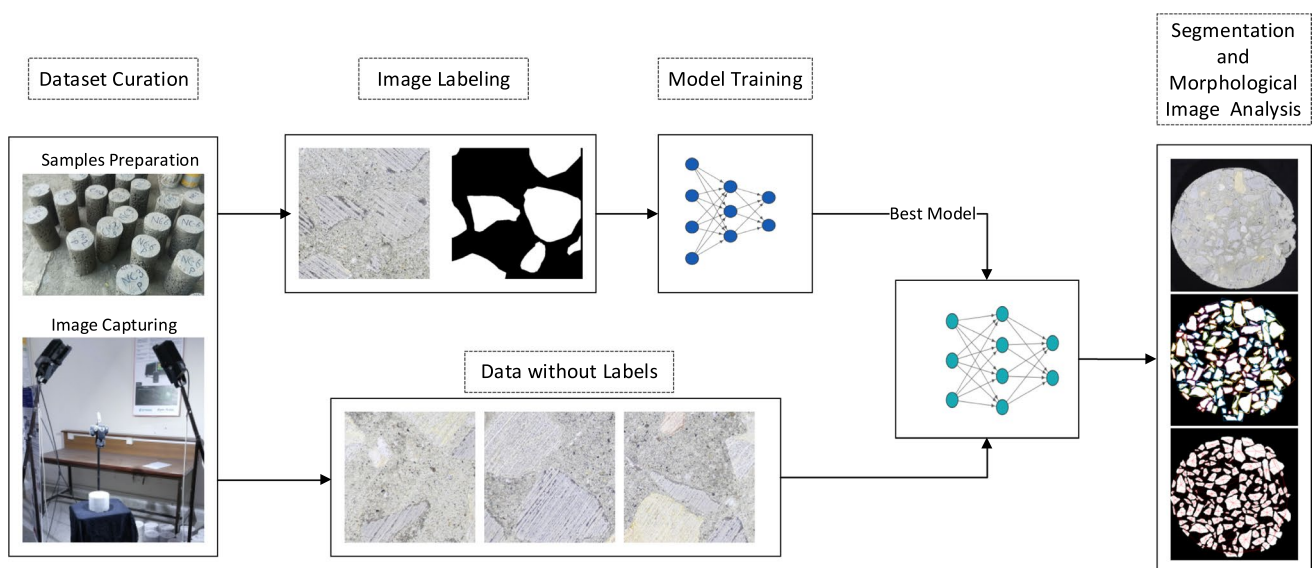


Fig. 8 CNN-based segmentation of concrete images for prediction and analysis

ingredients, aggregate characteristics, cement particles and hydration analysis, air voids, bugholes, segregation and mechanical properties of concrete.

4.1 Segmentation of concrete ingredients

Segmentation of concrete ingredients is an essential step in concrete forensic analysis to gain insight into the distribution and properties of aggregates, air voids, and surface defects that may influence concrete strength and durability. The overall workflow of the CNN-based segmentation process is illustrated in Fig. 8, showing steps from imaging and data labeling to model training and morphological image analysis. Image processing and CNNs have proven to be highly effective for image segmentation tasks. Various authors have developed CNN models that can precisely detect and segment aggregates and air voids in concrete images. Table 1 represents a summary of CV-based segmentation of different features in concrete images. The Mask R-CNN model has been widely used to enhance segmentation accuracy for various concrete components. Zhang et al. [56] developed an improved Mask R-CNN algorithm to segment and analyze the pore structure of pervious concrete. The modified algorithm showed improvements in both convergence speed and edge segmentation precision and resulted in a more accurate structure analysis for smaller pores. Performance metrics like mean average precision (mAP) and intersection over union (IoU) verified that the enhanced Mask R-CNN outperformed traditional methods in segmenting and detecting pores properly.

In another study, Zhang et al. [57] explored Mask R-CNN for aggregate particle segmentation, achieving over 92% accuracy for particles sized 2.36 mm to 16 mm. However, accuracy decreased slightly for smaller particles. The study also found that increased mud content had no significant effect on identification accuracy or precision. Tian et al. [58] proposed a method for segmenting meso-damage in concrete computed tomography images using Mask R-CNN, enabling quick segmentation without parameter adjustments. The model improved recognition accuracy with pixel-level instance segmentation of air voids and was validated on CT images under uniaxial compression. However, the method proposed in this study showed significant errors when extracting aggregates due to their complex shape characteristics. With the increasing complexity of concrete analysis, researchers have explored U-Net variations to enhance segmentation precision. Bangaru et al. [59] developed an automatic method for microstructure analysis of concrete using SEM images and U-Net-based semantic segmentation. The approach attained an accuracy of 94.43% for SEM images taken of seven different specimens under various magnifications. The U-Net model achieved a successful segmentation of the images into four different categories: aggregates, hydrated and anhydrous cement as well as pores. The approach gave a cost-effective solution for petrographic analysis concerning the investigation of failures in concrete. However, it struggled to distinguish pores from voids in some cases.

Wang et al. [60] presented a Mamba-based dual network enhanced by a triple-attention (TA-DCMamUNet), designed to segment concrete aggregates and evaluate their distribution uniformity. The model integrates Vision Mamba with a U-Net architecture, leveraging the state space module and a triple hybrid attention mechanism to enhance local and global feature extraction. TA-DCMamUNet was evaluated using both a publicly available dataset featuring colored aggregates and a custom-built dataset with gray aggregates. It demonstrated higher performance compared to other CNN and transformer-based approaches, achieving better F1 scores and mean IoU values. Additionally, it introduces a uniformity coefficient for assessing aggregate distribution uniformity, which demonstrates the model's capability to handle complex concrete images effectively. Zhang and Wang [61] presented an innovative approach for analyzing reinforced concrete corrosion using XCT and a U-Net model.

By automating the segmentation of XCT images into phases like rebar, cement mortar, corrosion products, pores, and cracks, the study achieves high segmentation accuracy (99.18%) and IoU of 82.1%. The framework facilitates 3D visualization and quantification of non-uniform corrosion, and provides insights for durability assessment. Lux et al. [62] presented a novel lightweight instance segmentation network based on ResUNet-a architecture for segmenting individual aggregates. This method facilitated the isolation of each aggregate from images for subsequent analysis and characterization. The segmentation masks generated by the network were

Table 1 Summary of CV-based segmentation of concrete ingredients

References	Dataset and image modality	Image processing technique/CNN Model	Application focus	Performance metrics			
				mIoU	Precision	Recall	F1
Zhang et al. [56]	180 images with 512×512 pixel resolution of pervious concrete samples	Improved Mask R-CNN with FPN	Pore structure of pervious concrete	0.93	—	—	—
Zhang et al. [57]	1000 images 2048×2048 pixels resolution of aggregates, specifically limestone, basalt, and pebble	Mask R-CNN, ResNet-101, FPN	Aggregate particles	0.93	—	—	—
Tian et al. [58]	2400 CT scanned images with 512×512 pixel resolution of concrete cylindrical samples	Mask R-CNN ResNet-101, FCN	Meso-damage in concrete, including voids and cracks	0.95	—	—	—
Bangaru et al. [59]	SEM Images with 256×256 pixels resolution	U-Net	Porosity and degree of hydration in concrete	—	0.95	0.94	0.94
Wang et al. [60]	2,208 images of hardened concrete specimens with a resolution of 256×256 pixels	TA-DCMamUNet	Concrete aggregate segmentation	0.85	92.31	0.91	0.91
Zhang and Wang [61]	2032 XCT images with resolution of 256×256 pixels of reinforced concrete specimens subjected to galvanostatic accelerated corrosion tests	Data augmentation techniques U-Net, SegNet, FCN	Automated segmentation of XCT images for steel corrosion analysis in concrete	0.82	0.917	0.827	0.87
Lux et al. [62]	Images with 256×256 pixels resolution of 6 classes of recycled aggregates	ResUNet	Recycled aggregates	—	—	0.847 (MAPE)	—
Wang et al. [63]	1144 subimages, each with a resolution of 480×480 pixels of concrete cylinders cut lengthwise	Otsu thresholding, U-Net, PSPNet, Modified ResNeXt50	Aggregates in concrete suspension	0.93	0.97	0.95	0.96
Song et al. [64]	Images with 800×800 pixels resolution of eight types of concrete mix cylinders	DeepLabv3 framework with ResNet-101 as the backbone	Air voids and aggregates segmentation for petrographic analysis	0.92	—	—	—
Ullah et al. [65]	9572 images with resolution of 512×512 pixels of concrete cylindrical samples with gray-colored aggregates and varying SCM content	DeepLabv3+ with encoders; ResNet101 MobileNet, Inception-ResNetV2, Xception, EfficientNetV2-S SegFormer SwinTV2-S, Feret diameter method, Nearest neighbor analysis	Coarse aggregate segmentation for forensic analysis of concrete	0.93	0.95	0.96	0.96
Saha et al. [66]	800 X-ray tomographic images of 7 days cured concrete specimens	Thresholding by Otsu's method	Air voids, aggregates, and mortar	0.91	0.92	0.89	0.91
Zeng et al. [67]	1800 images with 256×256 resolution	PANet (Path Aggregation Network), ResNet50	Fly ash microspheres in images,	0.82 to 0.92	—	—	—
Qian et al. [68]	180 patched SEM images with 800×800 pixels resolution of cement Sample	Fully Convolutional Networks (FCN), PointRend and MidPointRend	Unhydrated regions in cement samples	0.81	—	—	—
Hilloul et al. [69]	2024 images with 608×608 pixel resolution of various concrete, mortar, and cement paste samples	Mask R-CNN with ResNet101	Air voids and aggregates in concrete microscopic images	0.910	—	—	—
Huang et al. [70]	164 stockpile aggregate images collected from various quarries	Mask R-CNN, FCN	Stockpile aggregate images	0.93	0.95	—	—
Arafin et al. [71]	5,187 images of concrete surfaces having resolution of 224×224 pixels	U-Net, PSP-Net with VGG19, ResNet50 InceptionV3 and EfficientNetB3 as backbones	Concrete surface defects	0.91	0.94	0.97	0.95

qualitatively accurate as indicated by low area mean absolute percentage error (MAPE) loss. This made it possible to measure morphological properties such as particle size distribution and angularity, which provide significant insights into the nature of the aggregates.

DeepLab-based models have been employed to improve segmentation through encoder-decoder architectures and attention mechanisms. Wang et al. [63] addressed the drawbacks associated with traditional methods in evaluating concrete aggregates. In this work, an encoder-decoder architecture was employed by using a modified ResNeXt50 network as the encoder and the DeepLabv3+ decoder. A squeeze-and-excitation block was further added to perform dynamic fine-tuning of channel-wise feature responses for accurate feature extraction. Further improvement in model performance was achieved by applying a combined loss function that yielded better precision, recall, F1 and IoU scores on the concrete aggregate dataset. Song et al. [64] proposed an image-based concrete petrography analysis using a CNN model for the segmentation of concrete samples without any color treatment. The ResNet-101 served as the backbone in the DeepLabv3+ model and was tested on several sets of paired concrete images and labels. The segmentation accuracy was determined through manual identification and achieved an improvement in both accuracy and processing time over traditional color-based methods. Ullah et al. [65] carried out research work for segmenting gray-colored coarse aggregates from standard concrete images without color treatment or microscopy. Several encoders were utilized in the DeepLabv3+ architecture and trained on a self-curated dataset, and segmentation outputs were evaluated using metrics like recall, precision, and F1-score. Furthermore, the SegFormer model with a SwinTV2+ backbone was also trained and achieved the highest segmentation accuracy with an mIoU of 94.7%, surpassing the DeepLabv3+ architecture with the EfficientNetV2+ encoder. This indicates that Transformer-based architecture holds significant promise and merits consideration in future research directions, particularly for complex segmentation tasks in concrete imaging, where CNNs may have limitations.

Further advancements in segmentation techniques have been achieved through hybrid and attention-based models. A two-stage image processing method was reported by Saha et al. [66] for segmenting voids, aggregates, and mortar in concrete using images obtained from X-ray microtomography. In the first stage, voids were detected using gray value thresholding. In the second stage, a machine learning technique was applied for segmenting aggregates and mortar. Logistic regression was employed to model a decision boundary between the aggregate and mortar phases based on radial distance, gray value, and filtered embossed image features. The method achieved 95% accuracy in detecting aggregates. However, the proposed method has limitations in the precise detection of phases in concrete due to beam hardening and overlapping gray values. Zeng et al. [67] explored the PANet model for segmenting and analyzing fly ash microspheres using a simple optical microscope. This approach addressed limitations in traditional particle analysis methods by effectively minimizing errors caused by irregularly shaped particles, which can distort the overall particle size distribution in fly ash. Qian et al. [68] developed a segmentation framework for SEM images of cement paste samples. The key innovation was optimizing the PointRend algorithm to capture region boundaries. This resulted in more accurate and continuous segmentation. The framework was tailored to SEM images and outperformed methods like Fully Convolutional Networks (FCN) and Trainable Weka Segmentation (TWS).

It was also used to estimate the hydration degree of cement paste which aligned well with experimental data. However, on large images, it failed to include fine details and often gave a zig-zag boundary. This limitation shows that the framework might not be suitable for all image sizes and might need further fine-tuning for larger images. Hilloulin et al. [69] presented a modular DL segmentation model developed for the analysis of microscopic concrete images. The model consisted of two CNN models dedicated explicitly to air-void and aggregate detection. The model was trained and tested on distinct images at various magnifications and different samples, including concrete, mortar, and cement paste and demonstrated perfect capability in the segmentation of air voids and aggregates in concrete images. Among the important contributions made by this study is a simplified approach for the calculation of protected paste volume, distance-to-air-void, and characterization of air-void systems related to freeze-thaw cycles. However, the algorithm performed poorly when tested for larger air voids due to the overlapping between images. Huang et al. [70] utilized Mask R-CNN for automated segmentation

and morphological analysis of densely stacked stockpile aggregate images. The model outperformed traditional methods like watershed segmentation in accuracy and robustness. It enabled the efficient extraction of particle size and shape metrics, achieving a high precision of 87.4% and a completeness of 80.2% for practical field applications in aggregate quality control.

Arafin et al. [71] presented a DL approach for detecting and classifying concrete defects using CNN-based classifiers and semantic segmentation. Three CNN classifiers (VGG19, ResNet50, InceptionV3) and two segmentation models (U-Net and PSP-Net) with four backbone architectures including EfficientNetB3 and InceptionV3 were evaluated. InceptionV3 achieved the best classification accuracy of 91.98%, while EfficientNetB3-based U-Net yielded the highest F1-score of 95.66%. The study highlights the effectiveness of encoder-decoder architectures and emphasizes the impact of optimizers and learning rates on performance.

4.2 Aggregate characteristics

Analyzing the characteristics of aggregates from digital images of concrete offers a highly efficient and automated method for quality assessment of concrete. Integration of image processing and CNNs into concrete forensic analysis significantly reduces the need for manual inspection and ensures precise identification and measurement of aggregate properties such as size, shape, distribution, roundness, and texture.

Figure 9 represents the image analysis of aggregate size and roundness and the nearest-neighbor distance distribution in a cut concrete section of a cylinder. Table 2 represents a summary of the CV-based assessment of aggregate properties. The analysis of aggregate characteristics in hardened concrete has been significantly advanced through image processing and CNNs. Pourebrahimi et al. [72] introduced a comprehensive method for analyzing the size distribution of aggregates from micro and macrostructure images of hardened concrete. A simulation using images has been created to evaluate how aggregates pass through sieve openings by analyzing their relative translational and rotational movements within the images. A 2D image analysis method is introduced by Han et al. [73] for the assessment of coarse aggregate characteristics in concrete, including particle size, grading, shape, orientation, distribution, and spacing, as well as mortar-to-aggregate ratios. The study validated 2D results compared to 3D data and checked the agreement with aggregate sieving curves and rheology-based segregation predictions. The impact of stochastic errors on homogeneity was evaluated, and measures were proposed for mitigating segregation in self-consolidating concrete. Baek et al. [74] investigated the Segment Anything Model (SAM) for analyzing coarse aggregate characteristics in hardened concrete. The study introduced a novel lighting-based method to separate pores misclassified as aggregates. The SAM achieved aggregate area ratio estimates with less than 7% error compared to actual volumes, though errors in size distribution were noted due to the 2D

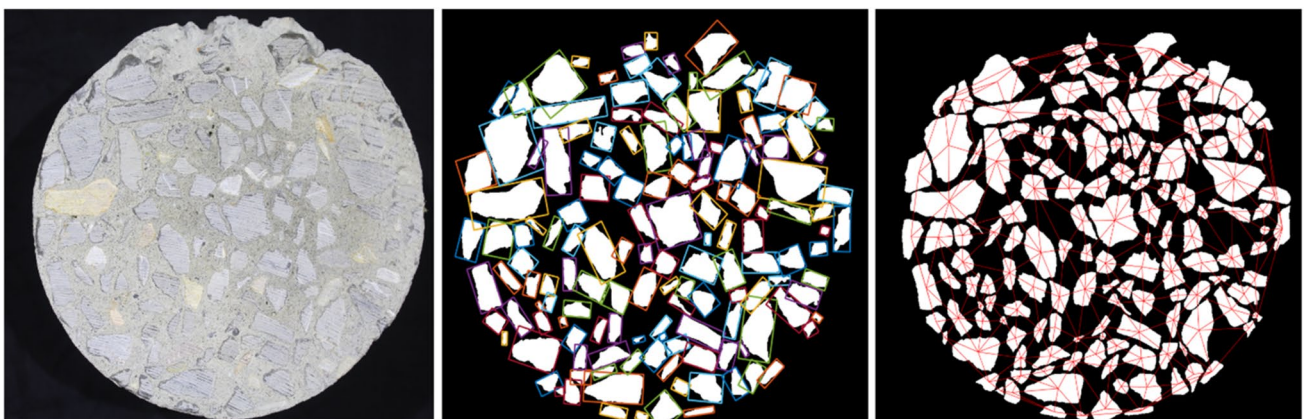


Fig. 9 Cut concrete cylinder (left), Aggregates size and roundness analysis (middle) and nearest-neighbor distance analysis (right)

Table 2 Summary of CV-based assessment of aggregates properties

References	Dataset source	Image processing/CNN model	Application focus
Zhang et al. [57]	1000 images with 2048 × 2048 pixels resolution of different types of aggregates, including diabase, limestone, basalt, and pebbles	Color enhancement, Gaussian noise addition, Mask R-CNN	Aggregate particle identification and gradation analysis
Lux et al. [62]	540 images with a resolution of 8192 × 4096 pixels of 16 classes of recycled aggregates	Classical image analysis, RACNET	Class, mass, size distribution and angularity of RA
Huang et al. [70]	164 Images of Stockpile aggregate images	VGG Image Annotating, Mask R-CNN	Equivalent sizes and flat and elongated (FE) ratios
Pourebrahimi et al. [72]	160 flatbed scanned images with a resolution 600 dp of hardened concrete cross-section samples	Otsu thresholding, Sieve simulation and optimizing rules (SSOR),	Aggregate size distribution
Han et al. [73]	Images with 3456 × 2304 pixels resolution of concrete cylinder specimens	Thresholding methods, Otsu's method	Aggregate size, grading, roundness, orientation, distributing homogeneity and mortar to coarse aggregate area ratio
Baek et al. [74]	Cross-sectional images of hardened concrete specimens with a resolution of 16,320 × 12,240 pixels	Segment Anything Model (SAM)	Coarse aggregates and mortar in concrete sections
Wang et al. [75]	Borehole wall images	Thresholding and binarization	Gradation, shape characteristic parameters, and distribution uniformity of coarse aggregates in foamed concrete components
Hoong et al. [76]	36,000 images of individual grains classified according to their nature	Binarization, marker-controlled watershed, ResNet34, ResNet50	Composition of RA
Hu et al. [77]	Images collected directly above the conveyor belt	Threshold segmentation, Watershed algorithm, Mask R-CNN	Angularity, flatness, elongation and gradation of aggregates
Yang et al. [78]	3D point cloud data of the aggregate on a conveyor belt	Multi-threshold segmentation, PoinNet, ISRNet	Particle size classification and gradation detection of aggregates
Qin et al. [79]	550 images of 6720 × 4480 pixels, collected directly above the conveyor belt	Color enhancement, Gaussian noise addition, AS Mask R-CNN	Detection and classification of particle size of mixed aggregates
Ma et al. [80]	Images of 800 Brick-concrete recycled coarse aggregate	X-ray tomography imaging and Laser scanning technology method	Morphological characteristics of RA
Ye et al. [81]	1784 images having a resolution of 640 × 640 pixels of steel slag concrete specimens	Canny Edge Detection, Segment Anything Model (SAM), improved YOLOv8	Evaluation of aggregate uniformity Detection of segregation issues in steel slag concrete
Zang et al. [82]	1000 images of 416 × 416 pixels resolution of manufactured limestone aggregates	Local threshold segmentation, YOLOv4	Particle size distribution measurement
Liu et al. [83]	235 images with a resolution of 608 × 608 pixels of aggregate in batching plant storage bins	YOLOv5x variant	Non-contact measurement of aggregate level in concrete batching plant storage bins
Yuan et al. [84]	525 images with a 1920 × 1080 pixels resolution of aggregates	Grayscale analysis, YOLOv5-Ytiny	Detection and classification of aggregate types (stones, small stones, machine-made sand, surface sand)

projection of 3D shapes. Despite challenges such as pore misclassification and resolution limitations, SAM offers a cost-effective, efficient alternative for concrete analysis, with potential for refinement through complementary algorithms to enhance 3D accuracy from 2D data. Wang et al. [75] presented a novel method for evaluating the size, shape, and distribution of coarse aggregates within concrete structures by using images captured with a borehole camera. The study also evaluated the distribution uniformity, gradation, shape characteristics, aspect ratio, inclination angle, and surface texture of the coarse aggregates.

Recycled aggregates present unique challenges and opportunities for characterization due to their variability in shape and composition. Zhang et al. [57] presented an innovative approach for aggregate particle identification and gradation analysis using the Mask R-CNN framework. The proposed framework demonstrated high accuracy and robustness, even in the presence of interference factors such as mud content. A gradation analysis method was presented based on equating aggregates to ellipses or circles, which showed close agreement with sieving test results. Hoong et al. [76] proposed an automated approach to assess the aggregate composition in the construction industry. Marker-controlled watershed, ResNet34, and ResNet50 models were used to learn from a labeled database of 36,000 images of individual grains, which achieved a validation accuracy of 97%. Images were compared with manual sorting and showed less than 2% difference in mass for most recycled aggregates (RA). The authors reported that grains of a given nature have a constant form and density for determination of the mass of a grain may not always hold true due to variations in grain shape and density. Recycled Aggregates Characterization Network (RACNET) was developed by Lux et al. [62] to determine binary mask, class and mass of individual aggregates based on 2D images. The proposed method could support quick characterization and monitoring of recycled aggregates in recycling platforms. However, the method relies on a simplified geometrical approach to estimate mass, which may not be fully accurate in cases where the aggregate's shape and density vary significantly. Online monitoring of aggregate characteristics in industrial applications has been enhanced through CV techniques. Mask R-CNN was utilized for stockpile aggregate image morphological analysis by Huang et al. [70]. Equivalent sizes and flat and elongated (FE) ratios of segmented particles were calculated through morphological analysis. The technique also gave an insight into particle size distribution and the distribution of the FE ratio, both of which are required to understand aggregate assemblies well regarding their composition and packing. Hu et al. [77] developed an online measurement method for coarse aggregate gradation in limestone mixtures with different proportions, water content, and mud content using the Watershed and Mask R-CNN model.

The particle shape parameters including maximum projected angularity, sphericity, elongation and flatness were calculated based on the boundary contour of each aggregate particle extracted. The Mask R-CNN proved more effective in achieving satisfactory segmentation results for different particle size ranges. The complexity of the background and stacking of objects resulted in traditional segmentation methods being less effective than the proposed methods. Yang et al. [78] introduced 3D point cloud and DL method for particle size classification and gradation of aggregate. The study proposed a multi-threshold segmentation that could reduce the amount of data that needed to be processed and improve prediction speed. The authors applied improved structure relation to improve the feature extraction capability of Improved Structure Relation Network (ISRNet) for aggregate point cloud and incorporated it with an adaptive sampling (AS) module that replaced the old Farthest Point Sampling to choose more important sampling points. Qin et al. [79] presented a noncontact, intelligent method for detecting aggregate size grading based on a background of big data. The authors presented an AS Mask RCNN model for hybrid stacked aggregate particle size detection through a fine-tuned migration learning training strategy. Three different experiments were performed and in all three experimental scenarios, the AS Mask RCNN network model achieved an accuracy of more than 89.13%. Morphological characteristics and fractal analysis of coarse aggregates obtained from construction and demolition waste were examined by Ma et al. [80] to understand how their type and size influence aggregate morphology and fractal dimensions. It was found that aggregate morphology largely fits into a normal distribution, and the fractal dimension serves as an effective measure of the irregularity and complexity of coarse aggregates.

Several YOLO-based models have been used to improve the evaluation, detection, and classification of aggregates in concrete. Ye et al. [81] proposed an improved YOLOv8-based method for evaluating the uniformity of

steel slag aggregates in concrete. The model incorporates ConvNeXt V2-R and an efficient multiscale attention mechanism, achieving high segmentation accuracy with mAP50 of 98.3%. A comprehensive uniformity coefficient was introduced to quantify aggregate distribution. The method proved effective for detecting segregation and assessing uniformity, especially at a 50% steel slag replacement rate. Zang et al. [82] developed a DL-based method using YOLOv4, distance transformation, and local thresholding to evaluate the size and shape of manufactured aggregates. The method achieved over 90% accuracy in segmenting overlapping particles and assessing roundness and axial coefficients, outperforming traditional image processing and U-Net models. Liu et al. [83] detected concrete aggregate levels in storage bins using monocular imaging and YOLOv5 DL. A single camera, angled at 45°, captures images to identify aggregate peak and valley positions. The method maps pixel coordinates to actual heights through a projection model. Compared to traditional sensor-based or corner detection techniques, this approach achieved high accuracy (95.43%) and real-time monitoring, offering a low-cost, reliable solution for aggregate level detection in concrete batching plants. Yuan et al. [84] developed YOLOv5-Ytiny, a lightweight and efficient DL model for real-time detection and classification of construction aggregates. The authors significantly reduce model size and computation time by modifying the original YOLOv5 architecture and replacing the C3 module with a simpler CI module, compressing the Neck structure, and switching to a CIoU loss function. Despite its compact design, YOLOv5-Ytiny achieves a high mAP of 99.6% and outperforms other models in speed, making it ideal for real-time applications in resource-limited environments.

4.3 Cement analysis

Image processing and CNNs have revolutionized material characterization in cement and concrete technology. By enabling precise segmentation and feature extraction, image processing identifies critical parameters like

Table 3 Summary of CV-based analysis of cement in concrete

References	Dataset source	Image processing technique/CNN Model	Application focus
Qian et al. [68]	115 SEM images with resolution of 3840×2400 pixels of cement paste samples	FCN, PointRend algorithm	Degree of hydration in paste samples
Sheiati et al. [86]	222 Images with resolution of 1024×640 pixels of cement paste samples of various hydration ages	Histogram thresholding, median filter algorithm, VGG-Unet	Evolution of solid phases, porosity and hydration in hydrated cement paste samples
Wong and Buenfeld [87]	BSE images with 2560×2048 resolution of OPC pastes with varying w/c ratios and curing ages	Histogram thresholding, Overflow method	Estimation of the original cement content, and degree of hydration
Gaël et al. [88]	SEM images with a resolution of 640×480 pixels of cementitious by-products	Gray scale equalization segmentation, binarization, hole filling filter, open morphological filter, and minimal particle size filter	Quantifying the anhydrous cement proportions in cementitious by-product
Zhao et al. [89]	BSE images with 3 different magnifications of cement paste samples	Grey-level histograms, Image quantitative analysis	Volume fraction of unhydrated cement
Qian et al. [90]	3600 SEM images with magnification 500 of cement paste mixtures and mortar samples	VGG16	End-to-end classification of cement paste and mortar with different w/c and silica fume contents
Käthler et al. [91]	Flatbed scanned images with 300 dpi resolution of polished concrete specimens	Trainable WEKA segmentation	Quantifying the cement content in concrete specimens, Chloride analyses in concrete structures
Chen et al. [92]	150 images with various dimensions of cement fragments of various sizes	Watershed, OTSU, HCN developed from a trimmed VGG net	Estimating cement fragment distributions

particle size, shape, and distribution, which are essential for understanding cement reactivity and its impact on concrete properties. These methods facilitate accurate analysis, reduce manual effort, and provide insights into microstructural characteristics that influence hydration kinetics and performance prediction. The integration of image processing and DL offers transformative potential for real-time quality control, predictive modeling, and sustainable advancements in cementitious materials. Several studies provide insights and in-depth analysis of the application of image processing and CNN in analyzing cement particles in concrete. Table 3 shows a summary of the CV-based analysis of cement in concrete. Qian et al. [68] developed a DL-based segmentation algorithm, optimized from PointRend [85], for precise identification of cement particles in SEM images, using 2267 labeled particles from samples with 0.2 and 0.4 w/c ratios cured over various periods. The segmentation results enabled accurate hydration degree calculations, aligned well with experimental data, and facilitated the analysis of particle irregularity and roundness across different curing conditions. The VGG-Unet model was applied by Sheiati et al. [86] for solid phase and porosity recognition from hydrated cement paste BSE images. The model attained 93% accuracy using 222 images across all hydration stages from 1 day to 2 years. A study was conducted by Wong and Buenfeld [87] using high-resolution field emission scanning electron microscopy (FE-SEM) in back-scattered electron (BSE) mode to analyze hardened cement paste. Images at 500 magnifications were segmented to measure area and volume fractions of unreacted cement, hydration products, and capillary pores. A strong correlation was observed between estimated and actual values for water content, cement content, and water-cement ratio across different samples and curing ages. Gaël et al. [88] quantified anhydrous cement in materials using SEM image analysis and particle swarm optimization. SEM images were analyzed for the anhydrous grain distribution of two cementitious by-products (CB1 and CB2). The method achieved noise elimination, prevented over-segmentation and allowed precise measurement of cement fragment sizes.

In addition to characterization methods, CNNs models have been applied for process monitoring and predictive modeling in cement production. Zhao et al. [89] developed a soft sensor model utilizing multivariate time series analysis combined with a convolutional neural network (MVTSCNN) to monitor the free calcium oxide (f-CaO) content in cement clinker production in real time. The model was developed using a dataset of historical data from a cement plant. The methodology involved analyzing the coupling relationship in the production system, selecting time series that may have direct impacts on the f-CaO content, and using a multivariate time series feature extraction method and a new timing matching method to deal with time-varying delay. Results indicated that, the MVTSCNN model has high-quality control, high production process optimization, and intelligent control in the production of cement clinker.

The proposed model enabled the automation of phase quantification in cementitious materials with accuracy comparable to expert analysis. Qian et al. [90] developed a modified VGGNet model to classify cement pastes and mortars in SEM images. The dataset comprised 1300 SEM images of cement paste and mortar samples with varying water-to-binder ratios and silica fume contents. Three transfer learning strategies were applied toward getting better classification performance on a small dataset. The results indicated high accuracies, 99.86% for the training set and 91.87% for the test set.

Käthler et al. [91] proposed a method to quantify cement content in concrete specimens using image analysis. Concrete slices were colored, polished, and scanned and then the images were processed with Fiji software and the TWS plugin [93], classifying pixels into aggregates, cement paste, and air voids. Area fractions for aggregates and cement paste were computed, converting cement paste volume to mass percentage. Results indicated considerable variability in cement content across specimens, impacting chloride analysis accuracy. Chen et al. [92] introduced a hybrid approach for cement fragment segmentation and distribution estimation, tackling issues like fragment overlap and image stains. The model was trained on a dataset of 150 images, with data augmentation to boost accuracy and multi-scale adaptability. The authors used a holistically nested convolutional network (HCN) for segmentation, applying morphological operations in post-processing to improve results. This method minimized noise, prevented over-segmentation, and accurately measured cement fragment sizes, proving effective for industrial cement image segmentation.

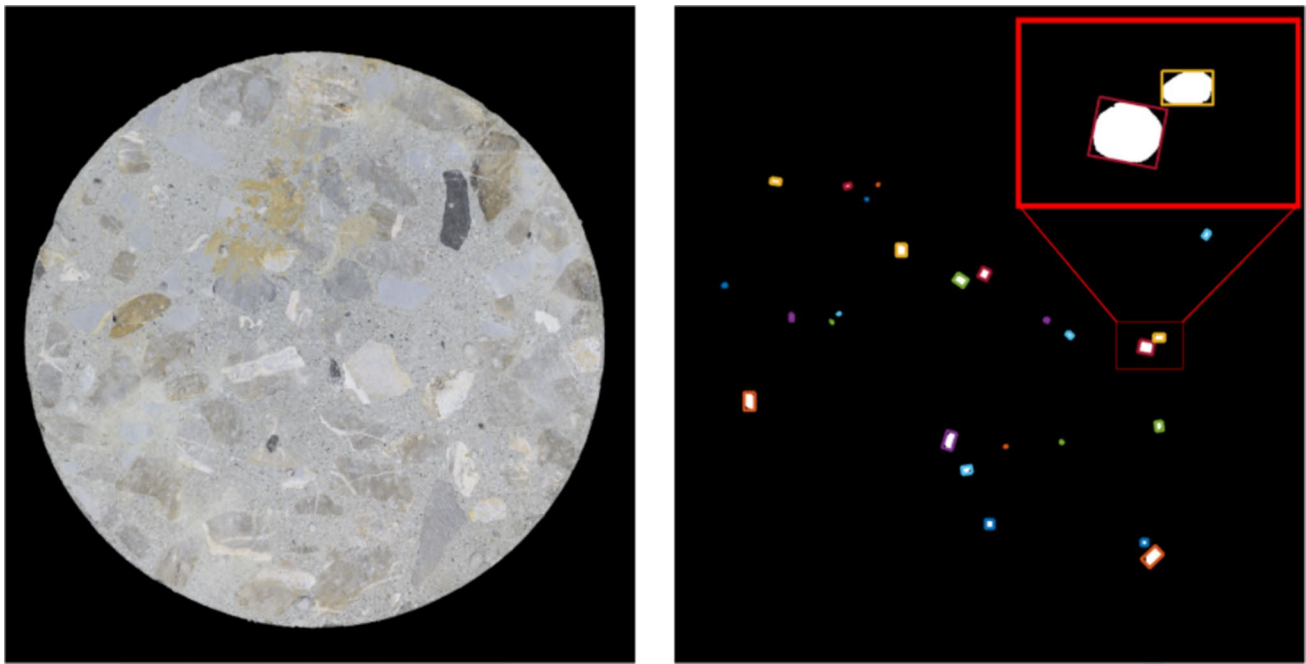


Fig. 10 Original concrete surface image (left) and corresponding air voids characterization (right)

4.4 Air voids analysis

Air voids in concrete are an important feature that can affect the durability and strength of the material. Air voids are produced due to the presence of air bubbles or pockets within the concrete mixture, and their distribution and size can impact the overall performance of the material. Figure 10 shows the original concrete surface and corresponding air voids characterization. Specific amounts of entrapped air or air voids in concrete are essential for enhancing freeze–thaw resistance, durability and workability [94]. However, excessive air void proportion can lead to a decrease in mechanical properties [95]. Assessing the air void ratio in concrete is crucial for achieving the desired balance between durability and mechanical properties. Traditionally, the analysis of air voids in concrete has been largely based on visual or physical inspection. Recent innovations in image processing and CNNs have led to the development of automated systems for accurate analysis of air void characteristics in concrete. Table 4 represents a summary of CV-based analysis of air voids in concrete. Tao et al. [96] employed photometric stereo and DL for air-void detection in hardened concrete without contrast enhancement. The authors captured multiple images under varied lighting and used the photometric stereo technique to reconstruct a 3D surface normal map. This 3D map was input into a U-Net CNN model to segment air voids. Accuracy was assessed with the modified point count method showing that the CNN model detected air voids accurately, achieving over 0.9 accuracy. Specific accuracy rates for air content, specific surface, and spacing factors were 0.92, 0.91, and 0.89, respectively.

A segmentation-based method was developed by Wei et al. [97] to detect air voids in concrete. It used a unique zero-angle light source and a modified PANet model. This light source highlighted air void features in concrete cross-sections, which are hard to capture with regular lighting. The modified PANet model effectively learned these features for improved detection and analysis. IoU values ranged from 0.599 to 0.926 with conventional lighting, and from 0.857 to 0.950 with the special light source. Tian et al. [58] proposed a Mask R-CNN model for segmentation of air voids and internal cracks in concrete CT images. Trained on 5000 high-resolution CT scans, the model achieved accuracies of 0.957 for voids and 0.931 for cracks. CT scanning and Mask R-CNN were used by Zhang et al. [56], for pore analysis in the pervious concrete. The improved Mask R-CNN model was applied

Table 4 Summary of CV-based analysis of air voids in concrete

References	Dataset source and image modality	Image processing technique/CNN architecture	Application focus
Zhang et al. [56]	180 Images with 512×512 pixel resolution of pervious concrete samples	Improved Mask R-CNN with FPN	Pore structure of pervious concrete
Tian et al. [58]	5000 CT scanning images with 512×512 pixels of concrete under uniaxial compression load	ResNet101 and FPN	Characterization and quantification of concrete voids and cracks in concrete meso-damage
Tao et al. [96]	10,200 Images with 256×256 pixels of concrete samples	U-Net	Identification of air voids
Wei et al. [97]	1160 images with 2048×1536 pixels resolution of cylindrical concrete samples cut horizontally	FPN used as the backbone in the PANet	Quantitative assessment of the air void structures in concrete
Lu et al. [98]	X-ray CT images with 1018×1024 pixels pixel resolution of Concrete Cylinders	Active contour line algorithm, grayscale thresholding, 3D spherical solidity thresholding	Size of entrapped air void and distribution in LWC
Yuan et al. [99]	500 high-resolution with 1024×768 pixels resolution of concrete specimens cut lengthwise	Modified version of ResNet-50,	Characterization of air voids and frost resistance
Koenig [100]	Incident light microscopy (ILM) and micro X-ray computed tomography images of mortars and concretes	Threshold segmentation, circle-similar voxel clusters	Distribution of micro air voids, specific surface and spacing factor
Chen et al. [101]	X-ray CT scanned images with $27 \mu\text{m}/\text{pixel}$ resolution of limestone and calcined clay-based 3D printable cementitious materials	Trainable Weka Segmentation	Characterizing air voids in 3D printed cementitious materials in terms of content, shape, size, and distribution
Wolter et al. [102]	Images with pixel size of $5.97 \mu\text{m}$ of five concrete mixes cylinders	Photogrammetry techniques for Point clouds generation, Delaunay triangulation	Detect and analyze air voids in concrete specimens
Hong et al. [103]	Images of Specimens inclosing Columnar limestone aggregates in mortar	Infrared Thermography with digital image correlation methods	Detecting and assessing void defects in early-age concrete
Fonseca and Scherer [104]	Images with a resolution of 3200 dpi of Mortar Samples with different proportions	Otsu thresholding algorithms for segmentation, and pixel connectivity analysis	Quantifying the air content, spacing, size distribution and specific surface of the bubbles in mortar samples
Murotani et al. [105]	Concrete slices images with $30 \mu\text{m}/\text{pixel}$ obtained from prism specimens	Image analysis with ImageJ software, segmentation of phases, 2D spatial point process	Quantifying the air voids distribution, estimate the spacing factor based on the characteristic distance
Ojala et al. [106]	Images with 1200 dpi resolution of concrete drilled core samples	Digital image analysis coupled with a stereology mathematical algorithm	Characterizing the total air content, air-entrained voids, and specific surface of voids in hardened concrete
Kim et al. [107]	BSE images with $0.5376 \mu\text{m}$ per pixel resolution of concrete samples	Edge detection, brightness thresholding, size and shape filters	Assessing the porosity of bonded mortar in recycled aggregate concrete

for pore identification, achieving mAP scores of 0.9750 and 0.9410 for mAP50 and mAP75, respectively. The study found that larger dataset sizes improved model accuracy, with segmented pore edges aligning more closely with real pore structures. XCT has been widely used to analyze air voids in concrete.

Lu et al. [98] used XCT to determine air void size distribution in pipe insulation lightweight concrete. The dataset included 3D XCT scans of industrial concrete specimens at varying resolutions. The active contour approach

helped set thresholds and a spherical solidity parameter to distinguish voids in paste and aggregates. A stepwise method simplified the process into two-phase grayscale thresholding, isolating lightweight aggregate particles on 2D slices. Thresholds were then applied between solid aggregate portions and paste, followed by a 3D shape-based criterion to isolate air voids in the paste. Yuan et al. [99] developed a correlation between microscopic air voids and concrete's frost resistance is explored using industrial CT. Concrete specimens were scanned with a microfocus X-ray source, achieving micro-level pixel resolution. Air content, average void diameter, air void distribution, and spacing factor were analyzed with VG Studio Max 2.2 software. The study examined air void characteristics under freeze–thaw cycles in salt solution. Results showed that industrial CT accurately measures microscopic air voids, offering insights into frost resistance. Air content measured by industrial CT was slightly higher than traditional pore structure analyzers. Koenig [100] proposed two methods, incident light microscopy (ILM) and micro-X-ray computed tomography to analyze air voids in cementitious materials. ILM detected air void parameters, while micro-X-ray computed tomography analyzed the size, surface, and location of individual air voids. The study examined the effects of sample preparation, resolution, and sharpness on image analysis. It concluded that micro X-ray computed tomography is suitable for analyzing single air void characteristics in cementitious materials. Chen et al. [101] studied air voids in 3D-printed cementitious materials, including limestone and calcined clay. Optical image scanning and X-ray CT were used to quantify air void content, distribution, pore size, and shape. Results showed that smaller air voids (10–1000 μm) were uniformly distributed in the layer regions, while larger voids (1000–6000 μm) were mainly between printed filaments, resulting in higher local porosity. Alternative imaging techniques such as photogrammetry and digital image analysis (DIA) have been applied for air void characterization. Wolter et al. [102] explored automated air void analysis in hardened concrete using close-range photogrammetry without contrast enhancement. Polished concrete sections with different air contents were analyzed. Point cloud data were processed to identify air voids through steps like pre-processing, surface identification, air void detection, and binary image generation. This digital method estimated air void parameters effectively. However, it struggled to capture smaller voids, suggesting a need for higher camera resolution and improved noise reduction in point clouds. Ojala et al. [106] used digital image analysis (DIA) and Schwartz-Saltykov conversion to examine size, shape, and distribution of air voids in hardened concrete samples. A stereology-based algorithm analyzed volume and area distributions of air voids. Key parameters, including total air content, specific surface, and Powers spacing factor, were calculated. The method effectively measured air void characteristics, proving useful for quality control of air-entrained concrete. Fonseca and Scherer [104] proposed an image analysis method using Otsu's thresholding to assess air voids in mortar and concrete. Samples were prepared with black acrylic ink and white barium sulfate powder. A flatbed scanner captured the images and binary images were created using Otsu's algorithm. Key air void properties air content, specific surface, size distribution, spacing were calculated using stereological analysis. The study also compared different darkening agents and white powders, recommending an optimal preparation method for improved accuracy.

Several studies have focused on the statistical and spatial analysis of air voids in concrete. Murotani et al. [105] examined air void distribution in concrete using spatial point process analysis. 2D images of air voids are transformed into spatial points. Statistical methods, such as the K-function and L-function, analyze these points. The authors also propose a method to estimate air void spacing. Hong et al. [103] examined void and crack detection in early-age concrete using digital image correlation and passive infrared thermography. The study assessed how defect size affects temperature and analyzed the void impact on concrete's mechanical properties. Both experimental and numerical approaches were used to understand early-age concrete behavior and potential defects. This research provides valuable insights into concrete performance and durability. Kim et al. [107] explored the impact of bonded mortar content on the properties of recycled aggregate concrete using BSE imaging. The study quantified porosity and highlighted its inverse relationship with mechanical strength by employing a pore segmentation technique and brightness histogram analysis.

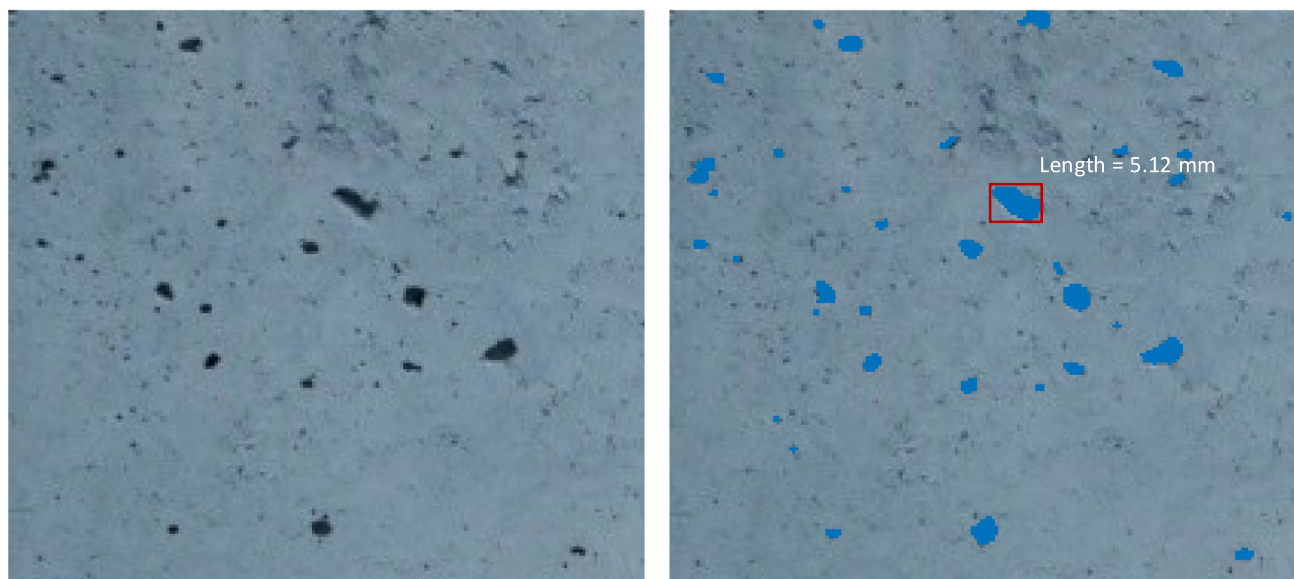


Fig. 11 Original concrete surface image (left) and corresponding Bughole quantification (right)

Table 5 Summary of CV-based analysis of surface bugholes

References	Dataset source and image modality	Image processing technique/ CNN architecture	Application focus
Liu and Yang [109]	Surface Images of Specimens having size of 150 mm × 150 mm × 150 mm	Image contrast enhancement, OTSU thresholding	Detection of surface bugholes in concrete
Zhao et al. [110]	1000 images of concrete surfaces	Extraction of a 108-dimensional texture feature using Gabor filters and gray-level run lengths	Classification of surface bugholes in concrete surface images
Yoshitake et al. [112]	Images with 5184 × 3456 pixels resolution of Tunnel-lining concrete surfaces	Thresholding, RGB color analysis	Bughole distribution in tunnel lining concrete
Liu et al. [113]	Surface Images of Specimens having size of 150 mm × 150 mm × 150 mm	Image contrast enhancement, OTSU thresholding	Factors influencing bugholes
Wei et al. [114]	428 raw images with 256 × 256 pixels resolution of concrete surface bugholes	Mask R-CNN, instance segmentation	Bughole detection and quantification
Wei et al. [115]	3,215 images with 256 × 256 pixels resolution of concrete surface	Data augmentation techniques, Mask R-CNN	Concrete quality assessment based on bughole rating
Sun et al. [116]	16,662 images with 256 × 256 pixels resolution of concrete surface of buildings	Deeplabv3 + with Xception-65 as encoder	Bughole detection and quantification
Zhou et al. [117]	4000 concrete surface images with 640 × 480 pixel resolution	DeepLabv3 + with Xception as encoder backbone	Evaluating pore size distribution
Wei et al. [118]	10,000 sub-images with 512 × 512 pixels resolution of concrete specimens	Otsu Thresholding, DCNN inception modules	Automated detection and evaluation of bugholes
Yao et al. [119]	800 images with resolution of 256 × 256 pixels of concrete surface	Otsu Thresholding, DCNN	Bughole detection
Tian et al. [120]	980 images with resolution of 640 × 640 pixels of bridge pier surfaces	OTSU thresholding, YOLOv5s	Bughole detection

4.5 Surface bughole analysis

Bugholes are undesirable voids or holes that appear on the surface of concrete structures, affecting their aesthetics and durability. The presence of bugholes can lead to issues such as reduced adhesion properties of materials applied to the surface and premature degradation of reinforced concrete structures [108]. Bugholes also pose challenges for subsequent processes like painting, as they need to be filled before coating application, adding labor-intensive and costly steps. Figure 11 shows the analyzed image with detected and labeled air voids along with their measured parameters. Various research studies have focused on methods for quantitatively evaluating concrete surfaces using image processing and CNNs to detect and quantify bugholes on concrete surfaces. Table 5 represents a summary of CV-based analysis of surface bugholes. Traditional image processing techniques have also been used for bughole analysis. Liu and Yang [109] proposed image processing techniques such as gray conversion, enhancement, threshold segmentation, and filtering for bugholes detection on the concrete surface. The evaluation parameters were the area ratio and maximum diameter of the bugholes. A MATLAB-based software for bughole detection was developed evaluating the quality of concrete surface quality. A variable selection method was used by Zhao et al. [110], to classify concrete surface voids with Gabor filter and gray-level texture features. The variable selection strategy involves seven steps, including base classifier selection, sample division, scoring, variable ordering, variable selection, evaluation metrics, and ensemble classification. This method, tested on 1,000 concrete texture images, outperformed the L-SHADE-SVM-SVD method Hoang and Nguyen [111], with accuracy, precision, and recall of 0.919, 0.904, and 0.938, respectively.

Yoshitake et al. [112] used color image analysis to improve bughole evaluation on tunnel-lining concrete. Small bugholes (less than 1 mm) were detected, which visual inspection often misses. This method enhanced area calculations of randomly distributed bugholes on concrete surfaces. Liu et al. [113] used image processing to analyze the effects of concrete properties on bughole area ratio, size distribution, and maximum diameter. The authors found that an increased W/C ratio reduced the area ratio of large bugholes and increased the number of small ones. For W/C ratios between 0.3 and 0.4, bugholes larger than 4 mm decreased their number and the area ratio as well. Correspondingly, with W/C ratio increments, smaller bugholes (0–2 mm) increase their number and also the area ratio. The author reported that superplasticizer added to sand and fly ash gave a successful reduction in bughole formation. Several studies have employed CNN models for automatic detection and quantification of bugholes. Wei et al. [114] modified Mask R-CNN to detect and quantify bughole defects in concrete surfaces using a dataset of 428 images. The model achieved an average precision of 0.89, outperforming traditional methods. In their quantification test on 70 bugholes, over 68% had an area error rate below 10%, and more than 74% had a diameter error rate below 10%.

A Mask R-CNN model was developed by Wei et al. [115], for detecting bugholes on concrete surfaces. A dataset of 3,215 images was collected using mobile cameras. The model was tested against the CIB scale for bughole rating. It achieved a bughole mask precision of 90.8%, allowing automatic detection and area calculation of each bughole for surface quality assessment. Sun et al. [116] developed a modified DeepLabv3+ model for detecting cracks and bugholes simultaneously. Transfer learning optimized the model with pretrained Xception backbones. It achieved 97.63% AP for cracks, 93.53% for bugholes, a mAP of 95.58%, and an MIoU of 81.87%. This model outperformed standard DeepLabv3+ and FCN in detecting small defects. Zhou et al. [117] analyzed bugholes larger than 10 μm in concrete using the DeepLabv3+, deep learning model. The model demonstrated high accuracy, achieving an IoU of 0.83, and effectively analyzed structural parameters such as bughole size, porosity, and shape factor. Compared to other models (e.g., FCN and SegNet) and traditional techniques, the proposed approach was found to be faster, more robust to lighting and spatial transformations, and required minimal preprocessing.

Other studies have focused on the application of CNNs for enhanced bughole recognition. Wei et al. [118] applied an image-based approach using DCNN for automated bughole detection on concrete surfaces. The proposed method effectively detected both large and small bugholes with high pixel accuracy. Above 99% pixel accuracy was achieved with a precision and recall of 89.46% and 83.87%, respectively. The compressive strength

of concrete was also assessed based on relating attributes of bughole numbers to the compressive strength. Yao et al. [119], utilized a DCNN using inception modules to detect bughole regions. The network was trained on 800 images (256×256 pixels) and outperformed traditional methods like Otsu and Laplacian of Gaussian algorithms. The DCNN achieved 96.55% accuracy, while Otsu and Laplacian of Gaussian achieved 87.93% and 82.76%, respectively. Tian et al. [120] proposed an improved YOLOv5 model for detecting surface bubbles in fair-faced concrete. The model introduces area-based K-means for better anchor box generation and integrates the efficient channel attention mechanism and feature fusion to enhance feature extraction. Tested on a custom dataset of 980 images, the model achieved superior accuracy compared to YOLOv5s, YOLOv6s with a 7.1% increase in mAP50. This approach offered fast and accurate detection, making it suitable for real-time concrete quality assessment.

4.6 Segregation analysis

The separation of coarse and fine aggregates during the mixing and placing process affects the quality and durability of concrete. Aggregate segregation is visually analyzed through aggregate distribution, as shown in Fig. 12. The uneven distribution of aggregates and cement paste results in areas with different strengths and durability. Researchers have been actively exploring innovative ways to monitor and detect segregation in concrete. The application of image processing and CNNs. utilize the power of CV and DL to automatically identify and measure segregation within concrete samples as shown in Table 6. Several studies have focused on using image-based methods to quantify segregation in lightweight aggregate concrete (LWAC). Solak et al. [121] introduced a segregation index to quantify segregation in LWAC and evaluated its performance against methods like ultrasonic pulse velocity, density measurements, and other segregation indices [122–125]. Using binarized images of LWAC specimens, the segregation index (SI) was calculated through a point-counting approach. This involved computing a local aggregate index (LAI) for each section (ratio of black to total pixels) and a global aggregate

Fig. 12 Concrete segregation analysis showing original and segmented images and aggregate distribution regions [126]

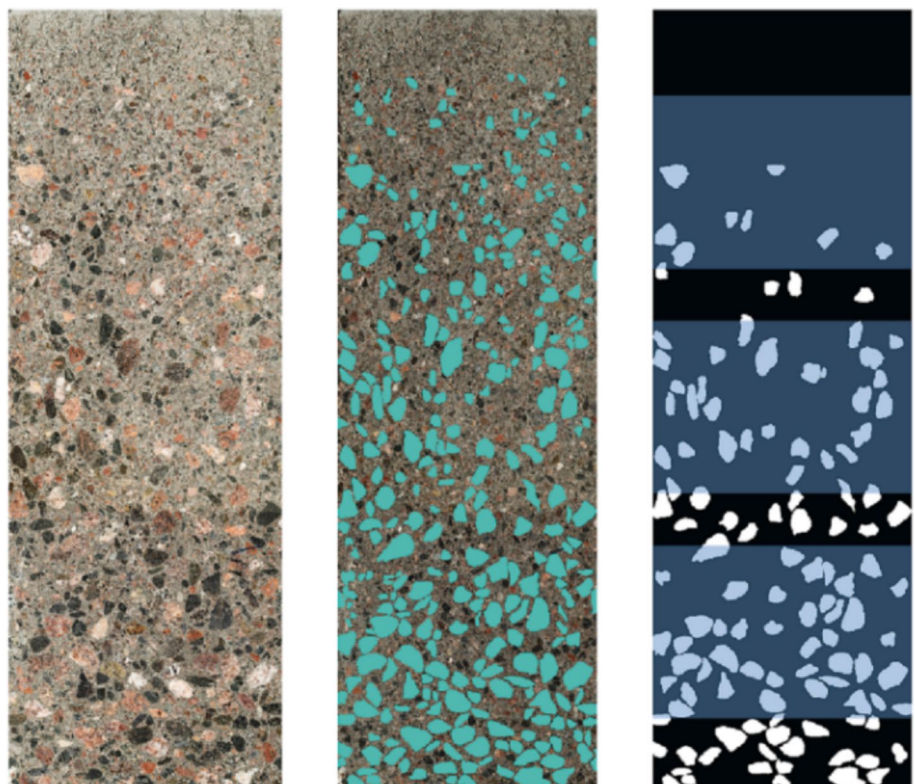


Table 6 Summary of CV-based assessment of concrete segregation

References	Dataset source and image modality	Image processing technique/CNN architecture	Application focus
Solak et al. [121]	202 images with resolution of 4752×3168 pixels, longitudinally saw-cut LWAC specimens	Contrast and threshold adjustments, Binarization, Matrix data extraction for aggregate and mortar identification	Estimation of segregation index in LWAC
Solak et al. [127]	Images obtained from cylindrical samples of LWAC	Generation of a binary matrix and Calculation of aggregate volumetric fraction and local aggregate volumetric fraction	Analyzing the influence of segregation caused by excessive vibration on the compressive strength of LWAC
Tenza-Abril et al. [128]	Images obtained from cylindrical samples of LWAC	Binarization, cropping, and filtering methods from the Catalano framework	Segregation indices in LWAC
Cui et al. [129]	60 images with resolution of 2000×2000 pixels of concrete slices	Calculation of geometric parameters (centroid, area, roundness) using convex hulls and Delaunay triangulation	Evaluating the static segregation of concrete aggregates and understanding the distribution of aggregates within concrete
Yoon et al. [130]	Cross-sections of hardened self-consolidating concrete (SCC) specimens	Global thresholding, Image binarization	Evaluating the degree of segregation in self-consolidating concrete (SCC)
Nili et al. [131]	Cylindrical concrete specimens 600 pixels per inch	Binary thresholding and dynamic thresholding methods (Otsu's method)	Evaluating the segregation resistance of self-consolidating concrete (SCC)
Ozen and Guler [132]	Pixel size measuring $0.169 \text{ mm} \times 0.169 \text{ mm}$	Grayscale segmentation, thresholding for object extraction, and morphological processing (erosion-dilation cycles) to clean the binary images and separate overlapping particles	Quantifying segregation in Portland cement concrete based on the spatial distribution of aggregate size fraction
Ojala et al. [126]	Digitally treated images of concrete, having a resolution of 9000×3000 pixels	Manual Segmentation and automatic calculation of the ratio of coarse aggregates	Evaluating the segregation of self-consolidating concrete (SCC)
Li et al. [133]	51,950 patched images of resolution 300×300 pixels of actual construction sites	Improved EfficientNetV2	Real-time monitoring of concrete vibration quality

index (GAI) for the entire image. The SI was derived by comparing LAI and GAI, offering a reliable image-based alternative to traditional experimental methods.

An image analysis method proposed by Solak et al. [127], assessed the impact of segregation on the compressive strength of LWAC. The study calculated a local aggregate volumetric index and local absolute difference coefficients from binary images to derive the Solak Index, quantifying segregation levels. Results indicated that single-layer versus two-layer molding significantly influenced LWAC's mechanical properties. Excessive segregation led to well-defined compression cones in 90% of cases, with cone size increasing alongside segregation intensity. Tenza-Abril et al. [128] developed a smartphone app for analyzing segregation indices in LWAC using image analysis. The app captures and processes images of LWAC samples, counting aggregate (black pixels) and cement mortar (white pixels) to calculate local and overall segregation indices. However, the simplified binarization process used in the algorithm may lead to overestimating or underestimating the black-and-white matrices, which could impact the accuracy of the segregation analysis results.

Other studies have focused on the analysis of static segregation in high-workability concrete (HWC) and self-consolidating concrete (SCC). Cui et al. [129] evaluated static segregation in HWC by analyzing 2000×2000 pixel cross-sectional images to identify aggregates in hardened concrete. Convex hulls and Delaunay triangulation were applied to calculate aggregate distribution characteristics, examining factors like aggregate gradation,

maximum diameter, and mortar rheology. Results indicated that higher large aggregate content reduced stability, increasing the risk of uneven distribution and static segregation in HWC. Similarly, Yoon et al. [130] used 3D point cloud analysis and digital image processing, including grayscale conversion, segmentation, and binarization, to assess segregation in SCC. This involved comparing the volume of segregation regions from 3D point cloud data with static segregation from image processing, achieving strong agreement and a 2.4% average error, affirming method reliability according to ASTM and AASHTO standards. Nili et al. [131] used image processing and Otsu thresholding to evaluate SCC segregation resistance by measuring aggregate distribution and mortar layer thickness in fly ash and silica fume mixes, finding enhanced resistance with these additions. Ozen and Guler [132] developed a digital imaging segregation index based on aggregate size distribution in concrete cross-sections, using maximum Feret diameters for size distribution analysis across binary images.

Another innovative approach for segregation analysis involved AC impedance spectroscopy combined with image processing. Ojala et al. [126] used ACIS alongside image processing to examine how coarse aggregate ratios varied at different depths in concrete specimens subjected to varying vibration times. The image analysis revealed significant segregation differences, even when specimens were vibrated under the same frequency and duration, despite using similar concrete mixtures. Li et al. [133] proposed a novel machine vision-based method for real-time monitoring of concrete vibration quality. The approach integrates an improved EfficientNetV2 model with image patching and a temporal fusion strategy to classify vibration quality into four levels: unqualified, middle, qualified, and over-vibration. The method achieved a high classification accuracy of 96.47%, outperforming other mainstream models. Temporal fusion significantly improved accuracy and robustness under varying lighting conditions. The approach can be successfully deployed in real-world construction settings, demonstrating its practical viability and potential to replace subjective human judgment in vibration quality assessment.

4.7 Mechanical properties prediction

An accurate prediction of mechanical concrete properties, i.e., compressive strength, tensile strength, elasticity, durability, thermal conductivity, etc., is vital for designing SHM sustainable and resilient structures that can withstand various environmental conditions, loads, and stresses. Compressive strength f_c and tensile strength f_t are the main indicators used in structural health assessment [134]. Traditionally, the mechanical properties of



Fig. 13 Testing setup for assessing mechanical properties of concrete; Compressive strength test using a universal testing machine (left), Ultrasonic pulse velocity (middle), Rebound hammer test (right)

Table 7 Summary of CV-based assessment of concrete mechanical properties

References	Dataset source and image modality	Image processing technique/CNN architecture	Application focus
Lee et al. [138]	332 DSLR camera images with a resolution of $4,096 \times 2,160$ pixels and 300 microscopic images with a resolution of $1,920 \times 1,080$ pixels	AlexNet	Assessing concrete compressive strength using surface image characteristics
Fan et al. [135]	Geopolymer Concrete samples having size of 1000×1000 pixels	Mesostructure model based on threshold segmentation of images	Assessment of the mechanical characteristics (peak stress, strain at peak stress and elastic modulus,) of geopolymer concrete and the impact of aggregate composition on its properties
Li et al. [136]	Fractured surface images of split cubic concrete samples	binarization, and opening operation processing	Influence of heterogeneity on the splitting process of concrete
Woldeamanuel et al. [137]	Infrared and true color images 1000 images from construction sites	FCN, SegNet, DeepLabv3+, ResNet-18 and ResNet-50	Early age strength of concrete and optimize formwork removal times in construction projects
Dogan et al. [139]	96 images with 3888×2592 pixels resolution of cylindrical samples	Image digitization, Segmentation, ANN; Levenberg–Marquardt (LM) back-propagation algorithm	Compressive and tensile strength of concrete samples
Waris et al. [140]	324 images with 6000×4000 pixel resolution of 54 different concrete cylinders	ANN; Levenberg–Marquardt (LM) back-propagation algorithm and Adaptive Neuro-Fuzzy Inference System	Compressive and tensile strength of concrete samples
Celik and Ozdemir [141]	1453 surface images with resolution of 256×256 pixels, captured from 203 core samples	MPaConvNet, SVR	Prediction of compressive strength of concrete
Jang et al. [142]	5145 images with resolution of $1,920 \times 1,080$ captured with portable digital microscope, of 27 Cylindrical specimens	AlexNet, GoogLeNet, and ResNet	Compressive strength of concrete samples
Kozubal et al. [143]	1000 standardized images of concrete interface surfaces with resolution of 224×224 pixels	Deep classifier, Alexnet and Resnet-50	Concrete interface strength based on surface morphology
Shiuly et al. [144]	120 images of mortar samples and 100 images of concrete samples of size 1600×1200 pixels	AlexNet, GoogleNet, VGG19, ResNet, and Inception-ResNet-V2	Compressive strengths of mortar and concrete samples
Fang et al. [145]	1080 images with resolution of 3264×2448 pixels of split concrete cylinders	Morphological operations Canny edge detection U-Net, SegNet, PSPNet	Impact of mesoscopic pore features on the splitting-tensile Strength of cellular concrete
Güçlüer [146]	Images of hardened concrete test samples produced with different types of aggregates	Image enhancement, Threshold Segmentation	Relationship between the shape characteristics of aggregates and the mechanical behavior of concrete
Mahankali and Valikala [147]	Surface images of concrete cube samples	Histogram analysis	Assessing the compressive strength of concrete using NDT and image processing

concrete are predicted through destructive and non-destructive testing as shown in Fig. 13. Civil engineers and construction professionals rely on accurate predictions of these properties to ensure the safety, reliability, and efficiency of buildings, bridges, dams, and other infrastructure. However, traditional testing methods, while accurate, are often time-consuming, destructive, and require extensive sample preparation. These limitations make it difficult to perform continuous monitoring or large-scale assessments.

In contrast, the integration of CNNs and image processing techniques offers a non-destructive, efficient, and scalable alternative for predicting concrete mechanical properties. Numerous recent studies have utilized CNNs and image processing to assess the concrete mechanical properties as summarized in Table 7. Fan et al. [135] presented a rapid method for predicting geopolymers concrete properties using image processing. Processed gray images and threshold segmentation generated mesoscopic models for numerical simulation, achieving low prediction errors of 1.50% (peak stress) and 9.70% (elastic modulus).

Li et al. [136] investigated the mechanical behavior of concrete under uniaxial compression and tensile splitting tests using image processing techniques. Threshold segmentation converts fracture surface images into binary form, enabling interface reconstruction in a numerical model. The heterogeneous model, assigning material properties to different regions, better replicates fracture initiation and splitting sequences, achieving a tensile strength of 3.770 MPa with 1.5% error, compared to 6.6% error in the homogeneous model. Woldeamanuel et al. [137] proposed a method to estimate concrete strength by integrating thermography and image segmentation. A thermal imaging camera and a trained segmentation model identify the concrete area, which is mapped onto a thermal image to analyze temperatures. Using a maturity-based equation, early-age concrete strength (1–7 days) was predicted with ± 3 MPa accuracy, achieving over 80% prediction accuracy and a 30% cost reduction.

Several studies have combined image processing techniques with ANN and other machine learning methods to predict concrete's mechanical properties. Dogan et al. [139] proposed a method combining image processing and ANN to predict concrete's mechanical properties. Statistical features extracted from images of cylindrical samples were used as ANN inputs, achieving a correlation of 97.18% to 99.87% with experimental results. This highlights the approach as an effective alternative to traditional testing methods. Waris et al. [140] proposed an image processing and neural networks method, integrating image processing and ANN to predict the compressive and split-tensile strength of hybrid concrete. High-resolution images (6000×4000 pixels) of 162 concrete samples with cement replacement ratios (0%, 15%, 25%) were analyzed for statistical features like mean, skewness, and entropy. Trained with destructive test (DT) results and calibrated with non-destructive tests (NDT) like ultrasonic pulse velocity and smith hammer, the ANN and Adaptive Neuro-Fuzzy Inference Systems models achieved R^2 values of 99.7% and 97.8%, respectively. Celik and Ozdemir [141] presented a novel non-destructive technique for predicting concrete compressive strength by integrating CNN and SVR methods. The proposed CNN architecture, MPaConvNet, analyzes 1453 surface images of concrete core samples, captured under diverse real-world conditions, to extract critical features, which are then utilized by the SVR model for accurate strength estimation. The method demonstrates outstanding performance with an R^2 value of 0.9959 and a low error rate (MAPE = 0.0209), surpassing traditional and existing AI-based approaches. Using Gradient-weighted Class Activation Mapping, the model identifies significant image regions influencing compressive strength, such as aggregate size and void ratio. This efficient and cost-effective model shows great promise for SHM and post-earthquake damage assessments.

Lee et al. [138] explored the application of DCNNs, AlexNet model, to predict the compressive strength of concrete from surface images taken with a DSLR camera and a microscope. They created 4,709 and 3,846 image datasets for the DSLR and microscope categories, respectively. The performance of the DCNN models was assessed by comparing their results to those obtained NDT methods. The DCNN model using DSLR images yielded lower RMSE (3.779) and MAPE (12.25%), indicating better prediction accuracy. In contrast, the DCNN model using microscope images showed higher RMSE (4.875) and MAPE (18.31%). These findings suggest that the DSLR images, which have a broader range, are more effective for feature extraction and achieving greater accuracy in predicting concrete compressive strength. Jang et al. [142] introduced a model using CNNs with portable digital microscope images to estimate compressive strength. Compared to ANN-based models, DCNNs,

especially ResNet, showed superior accuracy ($R^2=0.96$), lower RMSE (2.08 MPa), and MAPE (5.98). Kozubal et al. [143] proposed a framework to evaluate concrete-to-concrete interface strength using surface morphology analysis. Standardized concrete surface images and experimental data were used to develop a deep classifier for quick assessment. Deep features from AlexNet and ResNet-50 with classifiers demonstrated high precision and accuracy, with ResNet-50 showing robust performance. A strength model and assessment flowchart were also presented, highlighting the method's potential for estimating shear strength via surface texture properties. Shiuly et al. [144] investigated the use of SVM alongside five DCNN models, AlexNet, GoogleNet, VGG19, ResNet, and Inception-ResNet-V2 to assess their effectiveness in predicting the compressive strength of cement mortar and concrete based on digital microscopic images. Among the models tested, Inception-ResNet-V2 yielded the most accurate strength predictions, achieving R^2 values near 0.9 for mortar and approximately 0.85 for concrete. The RMSE was also calculated for each model, and Inception-ResNet-V2 produced the smallest RMSE values.

Surface morphology and aggregate characteristics extracted by image processing have been considered by some studies for strength estimation. Fang et al. [145] examined how pores affect the splitting-tensile strength of cellular concrete using specimens with controlled porosity. A deep-learning-based framework quantified meso-characteristics of pores on splitting surfaces, revealing that higher porosity and pore size reduce tensile strength. Key factors influencing strength were identified as pore size and surface porosity, leading to an empirical model for strength prediction. Güçlüer [146] examined the relationship between aggregate shape properties and the mechanical properties of concrete using image analysis to transform 2D aggregate behaviors into 3D. It evaluates geomechanical properties, surface roughness, and their effect on bond strength, compressive strength, and ultrasonic pulse velocity. Findings show that increased aggregate aspect ratio, roundness, and surface roughness enhance compressive strength and ultrasonic pulse velocity. Mahankali and Valikala [147] utilized high-resolution image processing to quantify air voids, cement, and aggregate proportions, aiding in strength prediction. Digital image analysis effectively predicts compressive strength, with results closely matching destructive and non-destructive methods like the rebound hammer and ultrasonic pulse velocity tests, showing minimal discrepancies with the SonReb method (-0.0028%) and image analysis (-9.382%) at 28 days.

5 Discussion

The studies reviewed in this paper demonstrate the application of various image processing techniques and CNN models, for the detection and characterization of various properties of concrete. Table 8 shows the merits and limitations of image processing and CNN in predicting various properties of concrete. CNN Models such as Mask R-CNN, U-Net, and DeepLabv3+ have demonstrated remarkable accuracy in segmentation, over traditional thresholding and morphological techniques, with mIoU values consistently exceeding 90% [56, 64]. The comparative analysis of different segmentation models reveals that incorporating modules like attention mechanisms and hybrid approaches can improve segmentation results [60, 63]. However, CNN struggled to distinguish overlapping phases or materials due to similar gray scales in images, especially in SEM or X-ray tomography. Furthermore, the performance of these models degrades when dealing with small and overlapping air voids [59, 69].

Traditional image processing techniques such as thresholding, watershed segmentation, and binarization work well for aggregate characterization in concrete due to their simplicity and computational efficiency [72]. However, these techniques primarily classified aggregates based on size and area but failed to capture more detailed shape and texture characteristics. CNN-based segmentation achieved high precision over traditional image processing techniques, even in complex aggregate distributions, as demonstrated by Zhang et al. [57] who used Mask R-CNN for gradation analysis and achieved robust performance despite the presence of mud content. Hybrid approaches that integrate image processing with CNN offer a more comprehensive understanding of aggregate distribution and uniformity [75, 76]. However, hybrid models involve complex architecture, increasing computational requirements and processing time. Three-dimensional imaging techniques such as XCT and point cloud

Table 8 Merits and limitations of image processing and CNN in concrete properties assessment

Concrete property	Merits	Limitations
Segmentation	1. Reduces manual intervention, speeding up processes 2. Attention mechanisms improve feature extraction 3. Provides detailed measurements (pore structure, aggregate size, hydration)	1. Requires large, labeled datasets, which are costly and time-consuming 2. Difficult to distinguish overlapping phases or materials due to similar gray scales in images, (e.g., SEM, X-ray tomography) 3. Image quality, resolution, and noise affect accuracy
Aggregate characteristics	1. Robust characterization of recycled aggregates, even with high mud content 2. Automated methods provide near real-time results, enhancing field-scale evaluations and reducing reliance on labor-intensive processes 3. 3D data and multi-feature fusion improve gradation detection	1. Overlapping aggregates and low contrast affect detection accuracy 2. Variations in aggregate shapes, sizes, and compositions across different regions reduce the generalizability of certain models 3. Fine-grained aggregates are challenging to detect in 2D
Cement analysis	1. Enables rapid, automated phase quantification (hydration, porosity in cement paste) 2. Adaptable to various imaging methods i.e., SEM, BSE and hyperspectral analysis	1. Distinguishing similar phases (overlapping fragments, stains) is difficult 2. Results from complex analyses like hyperspectral imaging can be non-trivial, and a lack of standardized protocols can limit reproducibility
Air voids	1. Reduces manual labor compared to traditional methods like ASTM C457 2. Combining infrared thermography and DIC enable detection of air voids and cracks in early-age concrete	1. Equipment like industrial CT scanners and zero-angle light sources can be expensive, limiting accessibility for routine testing 2. The method struggles with capturing the smallest air voids
Surface bugholes	1. Supports grading against standard scales (e.g., CIB reference) 2. Multi-scale analysis improves detection in poor conditions	1. Despite improvements, these methods still struggle with artifacts like dust, surface discoloration, or mixed imperfections 2. Techniques for small samples may not scale to large surfaces
Concrete segregation	1. Provides detailed measurements of aggregate segregation 2. Correlates segregation with material properties (size, gradation)	1. Limited to hardened concrete, not usable for fresh concrete 2. Relies on uniformity assumptions, which may not always hold
Mechanical properties	1. Avoids physical damage to structures, suitable for ongoing projects 2. Techniques like infrared thermography and CNN-based segmentation provide faster results compared to traditional destructive methods	1. Sensitive to variations in concrete mix design, requiring re-training 2. Gray color aggregates can lead to poor segmentation and CNNs may lack the important feature needed for precise mechanical property prediction

analysis provide the most detailed morphological insights into aggregate properties. Unlike 2D methods, 3D imaging predicted aggregate shape, size, texture, and spatial distribution with higher accuracy [78].

CNN models such as VGG-Unet and FCN, outperformed traditional image processing techniques in cement analysis, especially in identifying hydrated and unhydrated cement particles [68]. Integration of Otsu thresholding with CNN significantly reduced over-segmentation errors in cement fragment analysis [92]. MVTS-CNN provided a base for real-time monitoring and process optimization of cement clinker production [89]. These approaches enhanced quality control and eliminated manual testing errors, making them a valuable tool in industrial applications. CNN models, while offering superior performance in cement analysis are dependent on high-quality labeled SEM/BSE images [87].

In air void detection and classification, CNN models showed superior performance compared to image processing techniques such as Otsu thresholding and histogram-based segmentation [103]. Combining photometric stereo techniques with the U-Net model significantly improved air void detection accuracy [96]. XCT images with grayscale thresholding and active contour methods achieved precise quantification of entrapped air voids

[98]. Mask R-CNN and DeepLabv3+ have demonstrated superior performance in bughole detection and quantification compared to traditional image processing methods [117, 118]. Hybrid approach that integrated DCNN with inception modules, reduced false positives and enhanced model generalization across different surface textures [119].

Mask R-CNN models have successfully detected aggregate distribution patterns in concrete, reducing the reliance on manual inspections and improving the objectivity of segregation analysis. Image-based segregation indices, such as the local and global aggregate indices, have provided more reliable and consistent evaluations compared to conventional experimental methods [121]. Three-dimensional point cloud analysis has proven effective in detecting static segregation in SCC, achieving low error rates and confirming the reliability of image-based segregation monitoring [130]. The use of convex hulls and Delaunay triangulation techniques has provided significant insights into aggregate distribution [129].

Image processing and CNNs enable non-destructive and real-time prediction of mechanical properties, which is crucial for SHM and ensuring the safety and reliability of infrastructure. The analysis of image-based features, including pore size, aggregate distribution and surface roughness is crucial in determining mechanical properties [139]. The hybrid CNN models such as MPaConvNet and CNN-SVR utilizing these features have demonstrated exceptionally high accuracy in compressive strength of concrete, with correlation values exceeding 0.99 [141]. Similarly, ANN models integrated with image processing techniques have also achieved predictive accuracy above 97%, indicating the effectiveness of these techniques in estimating concrete strength [140]. Despite these advancements, the CV-based methods for the prediction of mechanical properties of concrete are sensitive to variations in concrete mix design and may require re-training or calibration for different datasets. Furthermore, concrete with gray-colored aggregates may cause CNNs to fail in extracting important aggregate features necessary for precise mechanical properties prediction.

5.1 Datasets

While the advancement of image processing and CNN-based models is central to the non-destructive evaluation of concrete properties, the foundation of these models lies in the quality, size, and diversity of the datasets used for training and evaluation. The performance and generalization capability of these models are directly influenced by the datasets they are trained on. Datasets enable models to learn critical spatial and textural features associated with concrete aggregates, air voids, cement phases, and surface defects. Therefore, the development of accurate and reliable CV models requires rich, well-annotated, and representative datasets.

In concrete analysis, dataset characteristics vary widely depending on the imaging modality (e.g., digital camera, SEM, X-ray computed tomography), the type of features being analyzed (aggregates, pores, cement, etc.), and the complexity of the segmentation or classification task. Datasets used in segmentation tasks, particularly at the pixel level, are often much smaller than classification datasets due to the labor-intensive nature of manual annotation. Annotating features like pores or overlapping aggregates demands a high degree of precision, which increases both time and cost. To facilitate comparison across studies and to guide researchers in dataset selection, Table 9 summarizes several key datasets reported in the literature. The table includes information such as dataset name (or source reference), image resolution, number of samples, type of concrete properties analyzed, annotation level (pixel-level, object-level, etc.), and key limitations. Manual labeling remains a significant challenge in concrete image datasets. For instance, segmenting air voids or differentiating between hydrated and anhydrous cement in SEM images often requires expert knowledge and meticulous effort. Similarly, capturing variations in aggregate shape, size, and texture under different lighting or imaging conditions adds to dataset complexity. The size and diversity of datasets greatly affect the robustness of CNN models. Larger and more heterogeneous datasets help models learn more generalized features, minimizing overfitting and enhancing performance on unseen data. Conversely, models trained on small, narrow datasets may perform well in controlled environments but fail under real-world variations. For example, models trained only on colored aggregate images may struggle when applied to grayscale or X-ray computed tomography images unless adequately generalized. Researchers

Table 9 Summary of image datasets for concrete properties predictions

Dataset referenceS	Characteristics	Annotation type	Strengths	Weaknesses
Zhang et al. [56]	180 images with 512×512 pixel resolution of pervious concrete samples	Pixel-level manual annotations of pores	1. Accurate representation of internal pore structures 2. Useful for pore segmentation tasks 3. Supports quantitative analysis	1. Requires specialized CT imaging equipment 2. Limited variability in aggregate materials 3. Few external conditions captured
Wang et al. [60]	2,208 images of hardened concrete specimens with a resolution of 256×256 pixels	Pixel-level manual annotations of aggregates	1. Reflects realistic construction scenarios 2. Suitable for evaluating robustness 3. Enables uniformity distribution analysis 4. Varying aggregate shapes/sizes	1. Image heterogeneity complicates annotation 2. Annotation quality varies
Zhang and Wang [61]	2032 XCT images with a resolution of 256×256 pixels of reinforced concrete specimens	Pixel-level manual annotations of steel rebar, corrosion products, and pores	1. Detailed, non-destructive imaging 2. Good for corrosion visualization 3. Enables 3D reconstruction	1. Limited scenarios 2. Quality is dependent on XCT settings 3. Requires advanced imaging equipment
Ullah et al. [65]	9572 images with a resolution of 512×512 pixels of concrete cylindrical samples with gray-colored aggregates and varying SCM content	Pixel-level manual annotations of coarse aggregates	1. Easy capture, suitable for the field 2. No special treatments 3. Realistic conditions 4. Variety in mix designs	1. Difficult differentiation of similar colors 2. Lower detail than microscopy
Hilloulin et al. [69]	2024 images with 608×608 pixel resolution of various concrete, mortar, and cement paste samples	Pixel-level manual annotations of aggregates and air voids	1. High segmentation accuracy 2. Microstructural insights 3. Enable freeze–thaw resistance analysis	1. Requires skilled sample preparation 2. Less scalable 3. Limited field usability
Song et al. [64]	Images with 800×800 pixels resolution of eight types of concrete mix cylinders	Pixel-level manual annotations of aggregates, paste, and air voids	1. Reduces segmentation time 2. No color treatments required	1. Requires polished surfaces 2. Quality varies with prep techniques 3. Limited external variability
Wang et al. [63]	1144 subimages, each with a resolution of 480×480 pixels of concrete cylinders cut lengthwise	Manual, pixel-level annotations for aggregates	1. Robust to complex backgrounds	1. Limited to sedimentation test images 2. Generalization to other aggregates is untested
Qian et al. [68]	180 patched SEM images with 800×800 pixels resolution of cement Sample	Manual, pixel-level annotations for cement particles	1. Effective boundary segmentation 2. No preprocessing needed 3. Detailed particle shape analysis	1. Performance is dependent on high-quality SEM images 2. Limited to cement particles
Bangaru et al. [59]	SEM Images with 256×256 pixels resolution	Manual, pixel-level annotations for aggregates, and pores	1. Comprehensive microstructure segmentation (aggregates, hydrated cement, anhydrous cement, pores)	1. Magnification sensitive 2. Limited dataset diversity 3. SEM-dependent
Saha et al. [66]	800 X-ray tomographic images of 7 days cured concrete specimens	Manual, pixel-level annotations for aggregates, mortar, and voids	1. Clear differentiation of voids, aggregates, and mortar phases	1. Limited scalability to diverse concrete mixtures, 2. Requires significant manual setup
Zhang et al. [57]	1000 images with 2048×2048 pixels resolution of aggregates	Manual, pixel-level annotations for aggregate particles	1. Accurate particle identification and gradation analysis 2. Robustness improved via transfer learning	1. Transfer learning needed for robustness across aggregate types 2. sensitive to image quality
Tian et al. [58]	2400 CT scanned images with 512×512 pixel resolution of concrete cylindrical samples	Manual, pixel-level annotations (masks) for cracks and voids	1. Good repeatability 2. allows detailed 3D reconstruction	1. Limited to CT scan images 2. Performance is sensitive to crack and void shape complexity

Table 9 (continued)

Dataset referenceS	Characteristics	Annotation type	Strengths	Weaknesses
Huang et al. [70]	164 stockpile aggregate images collected from various quarries	Manual, pixel-level annotations for aggregate particles	1. Efficient segmentation of overlapping aggregates 2. Good correlation with manual methods 3. Practical field applications	1. Depending on image angle and lighting conditions, 2. May miss deeply stacked particles
Zeng et al. [67]	1800 images with 256×256 resolution	Manual, pixel-level annotations for fly ash spherical particles	1. Rapid, on-site application 2. High accuracy for spherical particle segmentation 3. Directly correlates to quality parameters	1. Potential issues with varying fly ash particle density 2. Limited applicability beyond spherical particles
Qin et al. [79]	550 images of $6720 \text{ pixels} \times 4480 \text{ pixels}$, collected directly above the conveyor belt	Manual, instance annotation (pixel-level) for aggregates	1. High resolution 2. Various moisture and sand conditions 3. Real-time applications	1. Limited diversity (specific pebbles), 2. Time-consuming annotation
Yang et al. [78]	3D point cloud data of the aggregate on a conveyor belt	Manual, point cloud segmentation for coarse aggregates	1. Rich 3D geometric data 2. Adaptable to real-time grading	1. Requires specialized equipment 2. Data processing complexity
Hoong et al. [76]	36,000 images of individual grains classified according to their nature	Manual, bounding-box annotation, grain classification	1. An extensive labeled dataset 2. Suitable for real-time classification	1. Requires preprocessing (washing, drying) 2. Assumptions on grain shapes and densities
Hu et al. [77]	Images collected directly above the conveyor belt	Manual, pixel-level annotations for coarse aggregates	1. Effective in real-time environments 2. Robust against overlapping aggregates	1. Illumination sensitivity 2. Limited generalization to different conditions
Baek et al. [74]	Cross-sectional images of hardened concrete specimens with a resolution of $16,320 \times 12,240$ pixels	Manual, pixel-level annotations for coarse aggregates and pores	1. Minimal preprocessing required 2. Effective separation of aggregates and pores	1. Limitations in distinguishing complex aggregates
Chen et al. [92]	150 images with various dimensions of cement fragments of various sizes	Manual, pixel-level annotation for cement fragments	1. High segmentation precision 2. Improved accuracy in fragment size calculation 3. Robust against noise	1. Overlapping fragment issue 2. Manual selection of thresholds required 3. Sensitive to staining
Qian et al. [90]	3600 SEM images with a magnification 500 of cement paste mixtures and mortar samples	Manual, pixel-level annotations for cement paste microstructure	1. Good generality across mixtures	1. Dependent on SEM quality and preparation
Sheiati et al. [86]	222 Images with a resolution of 1024×640 pixels of cement paste samples of various hydration ages	Manual, pixel-level annotations for cementitious phases and porosity	1. Reliable quantification of hydration phases 2. Variation in imaging conditions	1. Grey-level overlap 2. Initial manual annotation intensive
Wei et al. [114]	428 raw images with 256×256 pixels resolution of concrete surface bugholes	Manual, pixel-level bugholes	1. Enables quantification of area and diameter of defects	1. Requires intensive initial manual labeling 2. Limited dataset size
Sun et al. [116]	16,662 images with 256×256 pixels resolution of concrete surface of buildings	Manual, pixel-level. Cracks, bugholes	1. Large-scale 2. Diverse defects 3. Supports transfer learning strategies	1. Limited generalization to different conditions 2. Requires preprocessing like contrast enhancement, thresholding complexity
Yuan et al. [99]	500 high-resolution with 1024×768 pixels resolution of concrete specimens cut lengthwise	Manual, pixel-level Air voids	1. Enables precise measurement of air void morphology and distribution	1. Complex handling and setup 2. Requires preprocessing like contrast enhancement, thresholding complexity

Table 9 (continued)

Dataset referenceS	Characteristics	Annotation type	Strengths	Weaknesses
Lee et al. [138]	332 DSLR camera images with a resolution of $4,096 \times 2,160$ pixels and 300 microscopic images with a resolution of $1,920 \times 1,080$ pixels	Pixel-level surface features	1. Limited dataset size 2. Controlled laboratory conditions 3. Reduced quality in microscopic images	1. High-quality, realistic images (DSLR) 2. Dual modality (DSLR & microscope) 3. Wide range of image features
Celik and Ozdemir [141]	1453 surface images with a resolution of 256×256 pixels, captured from 203 core samples	Surface features (voids, aggregate distribution)	1. Variability in image quality due to uncontrolled conditions 2. Potential annotation inconsistencies 3. Complexity due to heterogeneity of real-world structures	1. Large and diverse real-building samples 2. Variety in lighting and angles 3. Represents practical construction scenarios
Dogan et al. [139]	96 images with 3888×2592 pixels resolution of cylindrical samples	Surface features, voids	1. Limited diversity (controlled lab setup) 2. Smaller dataset size 3. Lack of real-world variability	1. Comprehensive parameter variation 2. Consistent imaging conditions 3. High-resolution surface imaging
Woldeamanuel et al. [137]	Infrared and true color images 1000 images from construction sites	concrete regions, thermal areas	1. Difficulty in precise alignment of dual image types 2. Interference from external factors (shadows, weather) 3. Dependence on thermal camera specifications	1. Real construction site images 2. Integration of thermal and visual data 3. Wide area coverage per image
Güçlüer [146]	Images of hardened concrete test samples produced with different types of aggregates	Surface features	1. Diverse aggregate textures 2. Rich variety from different crushers; clearly distinguishable aggregate properties	1. Limited to aggregate textural data 2. Does not cover concrete internal microstructure or defects
Fang et al. [145]	1080 images with a resolution of 3264×2448 pixels of split concrete cylinders	Surface features, pores	1. Precisely controlled pore characteristics 2. Ideal for studying pore structure effects	1. Limited variability due to controlled lab conditions 2. Not representative of diverse field conditions
Shiuly et al. [144]	120 images of mortar samples and 100 images of concrete samples of size 1600×1200 pixels	Mortar and concrete microstructures' surface features	1. High-resolution images with multiple scales 2. Good diversity of mortar and concrete mixes	1. Sensitive to microscope calibration 2. Labor-intensive sample preparation 3. Limited dataset size
Kozubal et al. [143]	1000 standardized images of concrete interface surfaces with a resolution of 224×224 pixels	Surface features and interface surface roughness	1. Unique imaging of surface roughness 2. Standardized illumination enhances data consistency	1. Highly controlled, lacks real-world variability 2. Heavily dependent on precise lighting conditions

must carefully evaluate the applicability, resolution, and annotation quality of datasets before employing them in model development. Choosing datasets aligned with the specific imaging conditions and structural characteristics of the concrete under study is essential for reliable predictions.

5.2 Practical implications and pathways for implementation

Recent advances in CV have demonstrated remarkable potential for assessing concrete properties. However, translating these advancements into everyday engineering practice requires making the technology accessible to practitioners. Many CV-based models proposed in research studies rely on high-end computational resources or require specialized expertise, hindering their widespread adoption in the field. To bridge this gap, the focus must shift toward user-friendly solutions that integrate into existing workflows. In practice, civil engineers, site managers, and quality control engineers should be able to leverage vision-based tools without requiring expert knowledge of machine learning or programming. The goal is to make these technologies as accessible and intuitive as traditional testing methods, enabling practitioners to visualize, quantify, and monitor material characteristics with minimal setup. For professionals new to CV, the following step-by-step pathway offers a practical roadmap for gradually implementing these methods in real-world applications.

5.2.1 Identify the use case and imaging needs:

Concrete properties to be evaluated should be identified first, as this will dictate the type of images required. For visible surface features (such as bugholes or areas of honeycombing/spalling), a standard digital camera or smart-phone is usually sufficient. Images should be captured under typical field conditions to ensure that subsequent models or analyses remain representative of real construction scenarios. For evaluating mechanical properties or internal features like air voids and aggregate dispersion in a polished concrete sample, higher-resolution imaging (e.g., a flatbed scanner or optical microscope image of a concrete cross-section) may be needed.

5.2.2 Collect and prepare example images:

A diverse image dataset should be compiled to capture the full range of conditions relevant to the intended assessment scenario. The dataset should include concrete surfaces exhibiting varying textures, lighting conditions, and defect severities, along with defect-free samples for reference. A scale or reference object should be included in some images for accurate calibration. For aggregate or void analysis, images should display a clear contrast between different phases. This can be achieved through consistent lighting or by using a plain background when photographing loose aggregates or polished sections.

5.2.3 Choose an appropriate model or technique

With the image dataset prepared, an informed choice between a pre-trained learning model and a classical image-processing pipeline should be made. For more straightforward or specialized tasks, a classical approach might suffice, for example, thresholding a grayscale image of a concrete slice to measure air-void content or using edge detection to find specific features. For complex and variable tasks, pretrained DL models are recommended because they can handle variations automatically. For instance, when segmenting aggregates or identifying surface defects, a pre-trained segmentation model such as Deeplabv3, U-Net, Mask R-CNN variant, or an object detector such as YOLO, may be used. However, when pre-trained models or readily available tools do not meet specific requirements, programmers or machine learning engineers can be engaged to develop customized CNN architectures or fine-tuned models tailored to the particular images at hand. This collaboration ensures optimal performance and enables field practitioners to benefit from robust AI solutions without requiring advanced coding expertise.

5.2.4 Set up software and hardware:

A practical setup for CV-based concrete assessment can be achieved with readily available tools and moderate computing resources. A standard workstation or laptop is generally sufficient, especially when using efficient models like the smallest YOLOv8 [148] segmentation model. This model is optimized for fast inference and can run on regular CPUs or modest GPUs commonly available in engineering offices. The software environment should include essential libraries when using Python-based implementations, such as Ultralytics YOLO [149], TensorFlow [150], PyTorch [151], and OpenCV [152]. These frameworks provide detailed documentation and ready-to-use examples, making configuration and initial experimentation straightforward. For practitioners opting for a no-code solution, open-source platforms like Fiji or ImageJ [153] offer a convenient alternative. These tools operate on standard computers and support plugins or macros for measurement and feature extraction without requiring programming expertise. For large-scale computation such as training DL networks or processing extensive image datasets, access to cloud-based GPUs or online computing services offers a cost-effective option.

5.2.5 Pilot test and calibrate

Before full implementation, the selected vision-based system should be tested on representative samples and its outputs compared with conventional measurements or expert evaluations. This preliminary validation builds confidence in the system's accuracy and ensures its results align with established engineering practices. For instance, when developing an image-based approach to evaluate aggregate gradation, a sample batch can be photographed and analyzed using the vision model and then subjected to a standard sieve analysis on the same material. The resulting gradation curves should show close agreement between both methods. Similarly, when analyzing air void parameters from concrete section images, automated results such as air content, void size distribution, and spacing factor should be compared with those obtained through ASTM C457 [154] microscopic analysis or benchmarked against typical acceptable ranges.

In cases where discrepancies arise, system calibration is essential. A standard calibration involves establishing a pixel-to-millimeter scale to ensure that image-derived dimensions accurately represent real-world units. This can be achieved by including an object of known size within the image to compute the necessary scale factor. Adjustment of model parameters may also be required, for example, refining the detection confidence threshold that classifies a feature as a void, aggregate, or cement matrix. These pilot trials also offer an opportunity to fine-tune lighting conditions, camera angles, and algorithmic settings to enhance consistency and robustness. Once this calibration phase is completed, the system should produce reliable outputs that correspond closely with conventional measurements, within an acceptable margin of error, ensuring readiness for broader application in practical engineering contexts.

5.2.6 Integrate into workflow and upscale

Once validated and calibrated, the CV system can be integrated into routine inspection and quality control workflows. Incorporation into daily operations may be as simple as adding a step to existing inspection routines, such as capturing images of specific structural elements after a site visit and processing them through the model to generate an automated condition report. At the early stage of implementation, the automated workflow should operate in parallel with manual inspections to confirm that key observations and defects are consistently detected. This parallel phase helps establish trust in the model's outputs and ensures alignment with established quality control practices.

5.3 Challenges for computer vision implementation in concrete properties assessment

5.3.1 Data imbalance and limited training datasets

One of the major issues that need to be addressed in implementing CV for the assessment of concrete properties is the imbalance and scarcity of diversified high-quality datasets. Factors like pore shape, aggregate type, and surface defects have very wide variations because of factors such as material mixtures, exposure to weather conditions, and method of curing [155]. All this calls for extensive fieldwork, data collection under varying conditions, and detailed annotation, which is highly time-consuming and resource-intensive [63]. Further, most datasets pertaining to concrete quality are highly imbalanced; for example, some types of defects (e.g., fine air pores) are vastly lower in number than others (e.g., cracks on the surface) [156]. This could train the model in a biased manner such that the algorithm is optimized to detect the more common types while not looking for subtle or infrequent features that are also important for structural evaluation [65]. Moreover, the existing publicly available datasets on concrete property evaluation are relatively small and narrow, primarily concentrated on particular applications [157, 158]. Such limitations have held back the formation of general models that can take up diversified tasks from cement analysis to segmenting aggregates and air voids.

5.3.2 Variability in concrete properties and environmental conditions

Variability in concrete properties and environmental conditions is a huge challenge to the deployment of CV techniques for the assessment of concrete properties. Surface textures can range widely in concrete structures because of variations in composition, mix design, modes of curing, and characteristics of aggregates [159]. Morphological variations due to the above-mentioned factors are further diversified, based on the influencing environmental parameters like weather conditions, lighting, and surface contamination which make imaging data noisy and inconsistent [80]. Contrast variation with changes in lighting conditions and shadows on the surface of the concrete produces an often striking visual impression. The size of the aggregate, its type, and distribution in a sample of concrete can also cause differences that impair the robustness of the model in CV. Another major reason which further complicates the accurate detection and quantification of air voids and bugholes is the effect of the season and environmental factors [118]. Rain, fog, and dust are such additional factors under which defect detection becomes more challenging. Surface wetness may mask the defects and voids as their shadows or glaring artifacts from imbalanced illumination may appear as defects. Before these preconditions, the assessment may be such that it gives a false positive or a false negative thereby reducing the reliability.

5.3.3 Defect overlap and complexity

The overlapping and complex relationship between features and defects in concrete structures presents major challenges for CV systems to do proper identification, classification and quantification. For instance, concrete surfaces exhibit sometimes overlapping features such as crack intersections with voids and discoloration blending with spalling or rough textures that hide microcracks [96]. Such overlapping can lead to a wrong result of misclassification or not being detected fully.

Different features often look alike. For example, under some conditions, bugholes can look like surface discolorations or fine gray color aggregates can look like air voids. Such morphological resemblance adds complexity to defect classification so the models need to be based on subtle features for differentiation [66].

5.3.4 Lack of standardized imaging protocols

Simple photography to advanced and sophisticated techniques such as scanning electron microscopy, X-ray computed tomography, and thermal imaging are some examples of concrete imaging methods [90, 160]. The resolution, depth, and contrast characteristics of each method can be so distinct that it may affect the detection and quantification of concrete defects or properties. Data inconsistencies can arise due to a lack of standardization, which hinders the interoperability of CV systems across different datasets and use cases. Furthermore, images are greatly influenced by the imaging protocols in the field conditions, such as variations in lighting, angle, and distance of the camera [161]. Such variations could distort the appearance of the features, such as aggregates, bugholes or air voids, thus leading to errors in detection and classification. Minor details might be concealed by inconsistent lighting, and oblique angles may alter perspective distortion. The lack of standardized scales for capturing and analyzing images complicates feature quantification and model training.

5.3.5 Influence of additives on computer vision-based assessment

The incorporation of supplementary cementitious materials (SCMs) and chemical admixtures introduces complexities in CV-based assessment of concrete properties. SCMs such as fly ash, silica fume, and slag alter key attributes like color, surface roughness, and void structures [65]. Additionally, SCMs and chemical admixtures affect workability and aggregate distribution [162]. Hydration and curing-induced temporal changes can influence defect characteristics such as pore size. A limited dataset representing only standard concrete means that

the model is not as robust and would require retraining for different applications. These variations challenge the accuracy that can be achieved with CV-based assessment, particularly in air void and cement analysis in concrete.

5.4 Future research directions

5.4.1 Benchmarking with the latest vision architectures

Future research should focus on benchmarking existing concrete datasets using the latest DL architectures to establish performance baselines and enhance reproducibility. Advanced one-stage detectors such as YOLOv10 [163], YOLOv11 [164], and YOLOv12 [165] offer significant improvements in inference speed and quantization awareness, supported by transformer-assisted attention mechanisms. Similarly, Vision Transformer (ViT)-based models, including Swin Transformer V2 [166], SegFormer [167], and ConvNeXt V2 [168], provide global context modeling that overcomes the local-feature bias of CNNs. Hybrid approaches that integrate traditional IP with transformer-based models and hierarchical classification can further enhance multi-class feature detection and segmentation. These methods can enable more precise recognition and categorization of complex concrete features such as aggregates, air voids, and surface defects across multiple scales and environmental conditions. Applying and systematically comparing these architectures on diverse concrete datasets identified in this review will help establish a unified benchmarking framework and guide future model development in this domain.

5.4.2 Integration of multi-modal imaging

Ultrasonic testing, thermal imaging, computed tomography, and ground penetrating radar are some of the advanced NDT techniques. All these NDT techniques can be integrated with CV algorithms to give a more holistic view of the subsurface anomalies. Such volumetric data, once acquired, can include void boundaries and irregularities within them, providing volumetric information that can significantly enhance void detection and characterization for CV algorithm processing. The performance of these systems in real-world applications needs to be validated for conducting a thorough and accurate assessment of a structure.

5.4.3 Synthetic data generation

Synthetic data generation has emerged as a pivotal solution to the challenges posed by limited and imbalanced datasets in CV. Techniques such as generative adversarial networks enable the creation of realistic representations of defects like cracks, voids, and spalling under diverse conditions. Synthetic data offers advantages such as addressing data imbalance by generating rare defect scenarios, reducing costs and time associated with manual data collection, and tailoring datasets to simulate specific conditions.

5.4.4 Integration with smart infrastructure

Integrating CV with smart infrastructure modernizes concrete assessment and health monitoring by combining IoT sensors, real-time analytics, and CV-based detection. CV systems process data from cameras, drones, and thermal or LiDAR sensors, enabling autonomous analysis and reporting through IoT platforms. This integration supports automated alerts, predictive maintenance, and lifecycle assessments while enhancing decision-making through combined visual and performance data. Future research should focus on developing lightweight CV algorithms, data fusion, and standardization which will ensure efficient, reliable, and durable infrastructure health monitoring.

5.4.5 Potential use of video data

Most existing approaches in concrete properties assessment focus solely on static images, neglecting the potential of video data. High-definition videos can provide richer temporal and spatial information for detecting surface bugholes in concrete structures. Future research should explore processing video frames as sequential datasets using RNNs. This approach can enhance the accuracy and robustness of SHM systems in real-world applications.

5.5 Economic and sustainability considerations

Implementing advanced CV-based inspection and monitoring systems does entail a significant upfront investment. High-resolution cameras and computing hardware need to be deployed, and substantial effort is required to collect and label training data. However, these initial costs are offset by the following long-term benefits.

5.5.1 Reduced labor and inspection costs

Once operational, automated visual inspection drastically lowers the need for labor-intensive manual checks. AI-driven systems can continuously monitor structures without large inspection crews, cutting down on labor expenses and time. Besides cost savings, it also minimizes the safety risks to personnel by reducing the frequency of dangerous on-site inspections.

5.5.2 Savings from non-destructive testing

The CV approach is inherently non-destructive as it evaluates concrete quality and detects defects without harming the structure. This avoids the costs of extracting samples or drilling cores required by traditional testing methods. By preventing damage during testing, engineers save on repair expenses and keep the structure fully serviceable during inspections. Furthermore, frequent monitoring can be performed with minimal disruption, providing a more complete picture of structural health at a lower overall cost.

5.5.3 Preventative maintenance and extended lifespan

Automated inspection enables a shift from reactive fixes to proactive maintenance. Early detection of defects allows engineers to intervene with minor repairs before problems escalate. Such preventive upkeep is far less expensive than emergency repairs after a major failure. Over time, this approach extends the service life of critical infrastructure by preserving its condition. In fact, continuous SHM and timely maintenance can increase intervals between major overhauls without compromising safety, yielding longer effective lifespan for bridges and buildings.

5.5.4 Sustainability and global priorities

Enhancing infrastructure durability through such technologies contributes to major sustainability goals in engineering. Global initiatives in civil engineering emphasize maximizing the service life of existing structures rather than resource-intensive new construction. By ensuring high construction quality and enabling timely repairs, CV-based methods help structures reach their full design life with safe performance. This reduces the need for premature replacements and the associated consumption of materials like cement and steel, whose production carries heavy environmental costs. The CV-based approach is a key enabler of a more sustainable construction industry as it promotes efficient use of resources, minimizes waste from demolition or over-maintenance, and curbs the carbon footprint of infrastructure through extended longevity.

6 Conclusion

This review examined the application of image processing and CNNs in assessing various concrete properties. The integration of image processing and CNNs has revolutionized these assessments, offering automated, non-destructive, and precise methodologies. A structured review approach was followed, with a bibliometric analysis using Scopus to identify relevant studies. A total of 82 peer-reviewed journal articles were selected based on a predefined set of exclusion criteria. Seven major applications of image processing and CNNs are identified in this review, including segmentation, aggregate characteristics, cement, air voids, surface defects (such as bugholes), segregation, and mechanical properties like compressive and tensile strength. The key insights from this review are summarized as follows.

Advanced CNN-based models, including Mask R-CNN, U-Net and DeepLabv3+, have demonstrated improvements in segmenting concrete components such as aggregates, air voids, cement particles, and bugholes. The Encoder-decoder architectures, such as DeepLabv3+ with EfficientNetV2, have achieved superior segmentation accuracy, particularly in challenging tasks involving gray-colored aggregates. Attention mechanisms and multi-scale feature extraction techniques have further enhanced model robustness in distinguishing complex concrete microstructures. Studies utilizing CT scanning, industrial X-ray tomography, and CV-based segmentation have provided more detailed assessments of air void characteristics, including size, shape, and distribution. The use of photometric stereo imaging combined with U-Net models has enabled more precise air void detection without contrast enhancement. CNN models like PANet, Mask R-CNN, and ResUNet have enabled automated quantification of aggregate size distribution, gradation, and roundness, reducing reliance on labor-intensive sieving methods. Multivariate time series analysis combined with CNNs has provided a novel approach for real-time monitoring of cement clinker quality. Image-based methods have successfully introduced segregation indices to quantify aggregate distribution uniformity, particularly in SCC and LWAC. AC impedance spectroscopy combined with image processing techniques has also proven effective in detecting aggregate segregation in fresh concrete mixtures. The prediction of concrete mechanical properties has been significantly enhanced through CV techniques. The hybrid CNN models, such as MPaConvNet and CNN-SVR, utilizing the aggregate distribution features have demonstrated exceptionally high accuracy in the compressive strength of concrete.

The review emphasizes practical implementation pathways that enable engineers to adopt vision-based inspection systems. Integrating these approaches into quality control workflows allows real-time, non-destructive evaluation of concrete structures. Moreover, the economic and sustainability discussion highlights that although initial investments in imaging and computation are required, automated inspection significantly lowers labor and testing costs, minimizes material waste, and enhances long-term sustainability by extending structural lifespan and reducing carbon-intensive maintenance. Furthermore, the review identifies critical challenges and future research directions in CV-based concrete properties assessment. The application of CV in concrete properties analysis faces challenges such as data variability, limited standardized datasets, and the complexity of overlapping features. These issues can be addressed by the integration of multi-modal imaging, synthetic data generation, and enhanced feature detection frameworks. The incorporation of real-time analytics through IoT and smart infrastructure can present a promising direction for autonomous and comprehensive SHM. Future research should prioritize the development of robust, scalable models capable of handling diverse environmental conditions and mix designs. The applicability and reliability of these techniques in real-world scenarios can be enhanced by leveraging video data, exploring hybrid DL architectures and integrating CV with traditional non-destructive testing methods. This review underscores the pivotal role of CV in modernizing concrete forensic analysis and provides valuable insights for researchers and practitioners to advance the field of concrete properties prediction.

Acknowledgments The first author would like to thank Qatar University for the sponsorship provided through the Graduate Assistantship program. Open Access funding provided by the Qatar National Library.

Authors' contributions The first author (MU) was responsible for conducting the literature review, analyzing and synthesizing the collected information, and drafting the manuscript. The second author (VP) provided guidance throughout the

research process, including defining the scope of the review, establishing the review methodology, and ensuring adherence to academic standards. The second author also contributed by reviewing and refining the manuscript, suggesting critical additions and improvements, and ensuring the scientific rigor and clarity of the work. Both authors reviewed and approved the final version of the manuscript.

Funding Open Access funding provided by the Qatar National Library. This research received no external funding.

Data availability The data that support the findings of this study are available from the corresponding author, Vagelis Plevris, upon reasonable request.

Declarations

Competing interests 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.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Prakash, G., R. Dugalam, M. Barbosh, and A. Sadhu (2022) Recent advancement of concrete dam health monitoring technology: a systematic literature review. In *Proceedings of the Structures*. Elsevier. p. 766–784., <https://doi.org/10.1016/j.istruc.2022.08.021>.
2. Tuan NM, Van Hau Q, Chin S, Park S (2021) In-situ concrete slump test incorporating deep learning and stereo vision. *Autom Constr* 121:103432. <https://doi.org/10.1016/j.autcon.2020.103432>
3. Flah M, Nunez I, Ben Chaabene W, Nehdi ML (2021) Machine learning algorithms in civil structural health monitoring: a systematic review. *Arch Comput Methods Eng* 28(4):2621–2643. <https://doi.org/10.1007/s11831-020-09471-9>
4. Jia J, Li Y (2023) Deep learning for structural health monitoring: data, algorithms, applications, challenges, and trends. *Sensors* 23(21):8824. <https://doi.org/10.3390/s23218824>
5. Başıyigit C, Comak B, Kılınçarslan Ş, Üncü İS (2012) Assessment of concrete compressive strength by image processing technique. *Constr Build Mater* 37:526–532. <https://doi.org/10.1016/j.conbuildmat.2012.07.055>
6. Qayyum W et al (2023) Predicting characteristics of cracks in concrete structure using convolutional neural network and image processing. *Front Mater* 10:1210543. <https://doi.org/10.3389/fmats.2023.1210543>
7. Karathanasopoulos N, Hadjidoukas P (2024) Deep learning based automated fracture identification in material characterization experiments. *Adv Eng Inform* 60:102402. <https://doi.org/10.1016/j.aei.2024.102402>
8. Kumar D, Pandey RC, Mishra AK (2024) A review of image features extraction techniques and their applications in image forensic. *Multimedia Tools Appl* 83(40):87801–87902. <https://doi.org/10.1007/s11042-023-17950-x>
9. Uchida S (2013) Image processing and recognition for biological images. *Dev Growth Differ* 55(4):523–549. <https://doi.org/10.1111/dgd.12054>
10. Kumar C, Sinha AK (2023) Automated crack detection and a web tool using image processing techniques in concrete structures. *Russ J Nondestr Test* 59(11):1119–1135. <https://doi.org/10.1134/S1061830923600569>
11. Moein MM et al (2023) Predictive models for concrete properties using machine learning and deep learning approaches: a review. *J Build Eng* 63:105444. <https://doi.org/10.1016/j.jobbe.2022.105444>
12. Seo J, Han S, Lee S, Kim H (2015) Computer vision techniques for construction safety and health monitoring. *Adv Eng Inform* 29(2):239–251. <https://doi.org/10.1016/j.aei.2015.02.001>
13. Koch C, Georgieva K, Kasireddy V, Akinci B, Fieguth P (2015) A review on computer vision based defect detection and condition assessment of concrete and asphalt civil infrastructure. *Adv Eng Inform* 29(2):196–210. <https://doi.org/10.1016/j.aei.2015.01.008>
14. Bai R, Wang M, Zhang Z, Lu J, Shen F (2023) Automated construction site monitoring based on improved YOLOv8-seg instance segmentation algorithm. *IEEE Access*. <https://doi.org/10.1109/ACCESS.2023.3340895>

15. Xu S et al (2021) Computer vision techniques in construction: a critical review. *Arch Comput Methods Eng*. <https://doi.org/10.1007/s11831-020-09504-3>
16. Kapadia H, Patel PV, Patel JB, Kanani N (2023) Implementation of computer vision technique for crack monitoring in concrete structure. *J Inst Eng (India): Series A* 104(1):111–123. <https://doi.org/10.1007/s40030-022-00695-5>
17. Jiang Y et al (2023) Tunnel lining detection and retrofitting. *Autom Constr* 152:104881. <https://doi.org/10.1016/j.autcon.2023.104881>
18. Alsakka F, Assaf S, El-Chami I, Al-Hussein M (2023) Computer vision applications in offsite construction. *Autom Constr* 154:104980. <https://doi.org/10.1016/j.autcon.2023.104980>
19. Ma Z, Liu Y, Li J (2023) Review on automated quality inspection of precast concrete components. *Autom Constr* 150:104828. <https://doi.org/10.1016/j.autcon.2023.104828>
20. Deng J, Singh A, Zhou Y, Lu Y, Lee VC-S (2022) Review on computer vision-based crack detection and quantification methodologies for civil structures. *Constr Build Mater* 356:129238. <https://doi.org/10.1016/j.conbuildmat.2022.129238>
21. Baduge SK et al (2022) Artificial intelligence and smart vision for building and construction 4.0: machine and deep learning methods and applications. *Autom Constr* 141:104440. <https://doi.org/10.1016/j.autcon.2022.104440>
22. Jayaram M (2023) Computer vision applications in construction material and structural health monitoring: a scoping review. *Mater Today: Proc*. <https://doi.org/10.1016/j.matpr.2023.06.031>
23. Mostafa K, Hegazy T (2021) Review of image-based analysis and applications in construction. *Autom Constr* 122:103516. <https://doi.org/10.1016/j.autcon.2020.103516>
24. Hoang N-D, Nguyen Q-L (2018) Fast local Laplacian-based steerable and Sobel filters integrated with adaptive boosting classification tree for automatic recognition of asphalt pavement cracks. *Adv Civ Eng* 2018(1):5989246. <https://doi.org/10.1155/2018/5989246>
25. Roy V, Shukla S (2013) Spatial and transform domain filtering method for image de-noising: a review. *Int J Modern Education Comput Sci* 5(7):41. <https://doi.org/10.5815/ijmecs.2013.07.05>
26. Fonseca, L.M.G., L.M. Namikawa, and E.F. Castejon (2009) Digital image processing in remote sensing. In *Proceedings of the 2009 tutorials of the XXII Brazilian symposium on computer graphics and image processing*. IEEE. p. 59–71, <https://doi.org/10.1109/SIBGRAPI-Tutorials.2009.13>.
27. De Natale FG, Boato G (2017) Detecting morphological filtering of binary images. *IEEE Trans Inf Forensics Secur* 12(5):1207–1217. <https://doi.org/10.1109/TIFS.2017.2656472>
28. Saravanan, C. Color image to grayscale image conversion. In *Proceedings of the 2010 Second International Conference on Computer Engineering and Applications*. 2010. IEEE. p. 196–199., <https://doi.org/10.1109/ICCEA.2010.192>.
29. Mustafa, W.A. and M.M.M. Abdul Kader. A review of histogram equalization techniques in image enhancement application. In *Proceedings of the Journal of Physics: Conference Series*. 2018. IOP Publishing. p. 012026., <https://doi.org/10.1088/1742-6596/1019/1/012026>.
30. Kang C-C, Wang W-J (2007) A novel edge detection method based on the maximizing objective function. *Pattern Recognit* 40(2):609–618. <https://doi.org/10.1016/j.patcog.2006.03.016>
31. Wang X (2007) Laplacian operator-based edge detectors. *IEEE Trans Pattern Anal Mach Intell* 29(5):886–890. <https://doi.org/10.1109/TPAMI.2007.1027>
32. Lotufo RA, Audigier R, Saúde AV, Machado RC (2023) *Morphological image processing in microscope image processing*. Elsevier, Amsterdam
33. Al-Amri, S.S. and N.V. Kalyankar (2010) Image segmentation by using threshold techniques. *arXiv preprint arXiv:1005.4020*, <https://doi.org/10.48550/arXiv.1005.4020>.
34. Bieniek A, Moga A (2000) An efficient watershed algorithm based on connected components. *Pattern Recognit* 33(6):907–916. [https://doi.org/10.1016/S0031-3203\(99\)00154-5](https://doi.org/10.1016/S0031-3203(99)00154-5)
35. Walton W (1948) Feret's statistical diameter as a measure of particle size. *Nature* 162(4113):329–330. <https://doi.org/10.1038/162329b0>
36. Gower JC (1985) Properties of Euclidean and non-Euclidean distance matrices. *Linear Algebra Appl* 67:81–97. [https://doi.org/10.1016/0024-3795\(85\)90187-9](https://doi.org/10.1016/0024-3795(85)90187-9)
37. Kattenborn T, Leitloff J, Schiefer F, Hinz S (2021) Review on Convolutional Neural Networks (CNN) in vegetation remote sensing. *ISPRS J Photogramm Remote Sens* 173:24–49. <https://doi.org/10.1016/j.isprsjprs.2020.12.010>
38. Chen L et al (2021) Review of image classification algorithms based on convolutional neural networks. *Remote Sens* 13(22):4712. <https://doi.org/10.3390/rs13224712>
39. Chen X et al (2022) Recent advances and clinical applications of deep learning in medical image analysis. *Med Image Anal* 79:102444. <https://doi.org/10.1016/j.media.2022.102444>
40. Bharati P, Pramanik A (2019) Deep learning techniques—R-CNN to mask R-CNN: a survey. *Comput Intell in Pattern Recognition: Proc of CIPR 2019*:657–668. https://doi.org/10.1007/978-981-13-9042-5_56
41. Ren S, He K, Girshick R, Sun J (2016) Faster R-CNN: towards real-time object detection with region proposal networks. *IEEE Trans Pattern Anal Mach Intell* 39(6):1137–1149. <https://doi.org/10.1109/TPAMI.2016.2577031>

42. Targ, S., D. Almeida, and K. Lyman (2016), Resnet in resnet: generalizing residual architectures. arXiv preprint [arXiv:1603.08029](https://arxiv.org/abs/1603.08029). <https://doi.org/10.48550/arXiv.1603.08029>.
43. Muhammad, U., W. Wang, S.P. Chattha, and S. Ali. (2018) Pre-trained VGGNet architecture for remote-sensing image scene classification. In Proceedings of the 2018 24th International Conference on Pattern Recognition (ICPR). IEEE. p. 1622–1627., <https://doi.org/10.1109/ICPR.2018.8545591>.
44. Siddique N, Paheding S, Elkin CP, Devabhaktuni V (2021) U-net and its variants for medical image segmentation: a review of theory and applications. *IEEE Access* 9:82031–82057. <https://doi.org/10.1109/ACCESS.2021.3086020>
45. Peng H et al (2020) Semantic segmentation of litchi branches using DeepLabV3+ model. *IEEE Access* 8:164546–164555. <https://doi.org/10.1109/ACCESS.2020.3021739>
46. Du, J. (2018) Understanding of object detection based on CNN family and YOLO. In Proceedings of the Journal of Physics: Conference Series. IOP Publishing. p. 012029, <https://doi.org/10.1088/1742-6596/1004/1/012029>.
47. Sharma A, Kumar V, Longchamps L (2024) Comparative performance of YOLOv8, YOLOv9, YOLOv10, YOLOv11 and faster R-CNN models for detection of multiple weed species. *Smart Agri Technol* 9:100648. <https://doi.org/10.1016/j.atech.2024.100648>
48. Shanthi T, Sabeenian R (2019) Modified Alexnet architecture for classification of diabetic retinopathy images. *Comput Electr Eng* 76:56–64. <https://doi.org/10.1016/j.compeleceng.2019.03.004>
49. Szegedy C, Ioffe S, Vanhoucke V, Alemi A (2017) Inception-v4, inception-resnet and the impact of residual connections on learning. *Proc AAAI Conf Artif Intell*. <https://doi.org/10.1609/aaai.v31i1.11231>
50. Chollet, F. Xception: Deep learning with depthwise separable convolutions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2017. p. 1251–1258., <https://doi.org/10.48550/arXiv.1706.03059>.
51. Zhang K, Guo Y, Wang X, Yuan J, Ding Q (2019) Multiple feature reweight densenet for image classification. *IEEE Access* 7:9872–9880. <https://doi.org/10.1109/ACCESS.2018.2890127>
52. Marques G, Agarwal D, la De Torre Díez I (2020) Automated medical diagnosis of COVID-19 through EfficientNet convolutional neural network. *Appl Soft Comput* 96:106691. <https://doi.org/10.1016/j.asoc.2020.106691>
53. Tan, M. and Q. Le (2021) Efficientnetv2: Smaller models and faster training. In Proceedings of the International conference on machine learning. PMLR. p. 10096–10106.
54. Plevris, V., G. Solorzano, N.P. Bakas, and M.E.A. Ben Seghier (2022) Investigation of performance metrics in regression analysis and machine learning-based prediction models. <https://doi.org/10.23967/eccomas.2022.155>.
55. Plevris V (2025) Assessing uncertainty in image-based monitoring: addressing false positives, false negatives, and base rate bias in structural health evaluation. *Stoch Environ Res Risk Assess* 39(3):959–972. <https://doi.org/10.1007/s00477-024-02898-7>
56. Zhang H et al (2022) Analyzing the pore structure of pervious concrete based on the deep learning framework of Mask R-CNN. *Constr Build Mater* 318:125987. <https://doi.org/10.1016/j.conbuildmat.2021.125987>
57. Zhang R et al (2023) Aggregate particle identification and gradation analysis method based on the deep learning network of Mask R-CNN. *Mater Today Commun* 35:106269. <https://doi.org/10.1016/j.mtcomm.2023.106269>
58. Tian W et al (2021) Meso-structure segmentation of concrete CT image based on mask and regional convolution neural network. *Mater Des* 208:109919. <https://doi.org/10.1016/j.matdes.2021.109919>
59. Bangaru SS, Wang C, Zhou X, Hassan M (2022) Scanning electron microscopy (SEM) image segmentation for microstructure analysis of concrete using U-net convolutional neural network. *Autom Constr* 144:104602. <https://doi.org/10.1016/j.autcon.2022.104602>
60. Wang D et al (2024) Evaluation of aggregate distribution uniformity using Vision Mamba-based dual networks for concrete aggregate segmentation. *Expert Syst with Appl*. <https://doi.org/10.1016/j.eswa.2024.126076>
61. Zhang M, Wang W (2024) Deep learning-based extraction and quantification of features in XCT images of steel corrosion in concrete. *Case Stud Constr Mater* 20:e02717. <https://doi.org/10.1016/j.cscm.2023.e02717>
62. Lux J, Hoong JDLH, Mahieux P-Y, Turcry P (2023) Classification and estimation of the mass composition of recycled aggregates by deep neural networks. *Comput Ind* 148:103889. <https://doi.org/10.1016/j.compind.2023.103889>
63. Wang W, Su C, Zhang H (2022) Automatic segmentation of concrete aggregate using convolutional neural network. *Autom Constr* 134:104106. <https://doi.org/10.1016/j.autcon.2021.104106>
64. Song Y, Huang Z, Shen C, Shi H, Lange DA (2020) Deep learning-based automated image segmentation for concrete petrographic analysis. *Cem Concr Res* 135:106118. <https://doi.org/10.1016/j.cemconres.2020.106118>
65. Ullah M, Mir J, Husain SS, Shahid MLUR, Ahmad A (2024) Concrete forensic analysis using deep learning-based coarse aggregate segmentation. *Autom Constr* 162:105372. <https://doi.org/10.1016/j.autcon.2024.105372>
66. Saha SK, Pradhan S, Barai SV (2020) Use of machine learning based technique to X-ray microtomographic images of concrete for phase segmentation at meso-scale. *Constr Build Mater* 249:118744. <https://doi.org/10.1016/j.conbuildmat.2020.118744>
67. Zeng Z et al (2022) Deep learning enabled particle analysis for quality assurance of construction materials. *Autom Constr* 140:104374. <https://doi.org/10.1016/j.autcon.2022.104374>
68. Qian H, Li Y, Yang J, Xie L, Tan KH (2023) Segmentation and analysis of cement particles in cement paste with deep learning. *Cem Concr Compos* 136:104819. <https://doi.org/10.1016/j.cemconcomp.2022.104819>

69. Hilloulin B, Bekrine I, Schmitt E, Loukili A (2022) Modular deep learning segmentation algorithm for concrete microscopic images. *Constr Build Mater* 349:128736. <https://doi.org/10.1016/j.conbuildmat.2022.128736>
70. Huang H, Luo J, Tutumluer E, Hart JM, Stolba AJ (2020) Automated segmentation and morphological analyses of stockpile aggregate images using deep convolutional neural networks. *Transp Res Rec* 2674(10):285–298. <https://doi.org/10.1177/036119812094388>
71. Arafin P, Billah AM, Issa A (2024) Deep learning-based concrete defects classification and detection using semantic segmentation. *Struct Health Monit* 23(1):383–409. <https://doi.org/10.1177/14759217231168212>
72. Pourebrahimi M, Shahhosseini V, Ramezaniapour AA (2023) Innovative sieve simulation and microstructure image analysis techniques for estimation of aggregate size distribution in hardened concrete. *Constr Build Mater* 384:131456. <https://doi.org/10.1016/j.conbuildmat.2023.131456>
73. Han J, Wang K, Wang X, Monteiro PJ (2016) 2D image analysis method for evaluating coarse aggregate characteristic and distribution in concrete. *Constr Build Mater* 127:30–42. <https://doi.org/10.1016/j.conbuildmat.2016.09.120>
74. Baek S, Koh S, Kim W, Kim T (2024) Evaluation of coarse aggregate properties in hardened concrete based on segment anything model (SAM). *Constr Build Mater* 453:139013. <https://doi.org/10.1016/j.conbuildmat.2024.139013>
75. Wang C, Han Z, Wang Y, Wang C, Chen S (2023) A novel method for measuring and analyzing the characteristics of coarse aggregates in concrete. *J Build Eng* 78:107730. <https://doi.org/10.1016/j.jobbe.2023.107730>
76. Hoong JDLH, Lux J, Mahieux P-Y, Turcry P, Ait-Mokhtar A (2020) Determination of the composition of recycled aggregates using a deep learning-based image analysis. *Autom Constr* 116:103204. <https://doi.org/10.1016/j.autcon.2020.103204>
77. Hu X et al (2022) Online measurement and segmentation algorithm of coarse aggregate based on deep learning and experimental comparison. *Constr Build Mater* 327:127033. <https://doi.org/10.1016/j.conbuildmat.2022.127033>
78. Yang M et al (2023) A coarse aggregate gradation detection method based on 3D point cloud. *Constr Build Mater* 377:131111. <https://doi.org/10.1016/j.conbuildmat.2023.131111>
79. Qin J et al (2023) Deep learning-based software and hardware framework for a noncontact inspection platform for aggregate grading. *Measurement* 211:112634. <https://doi.org/10.1016/j.measurement.2023.112634>
80. Ma K et al (2021) The morphological characteristics of brick-concrete recycled coarse aggregate based on the digital image processing technique. *J Build Eng* 44:103292. <https://doi.org/10.1016/j.jobbe.2021.103292>
81. Ye G et al (2024) Evaluation method for uniformity of steel slag concrete aggregate based on improved YOLOv8. *J Build Eng* 98:111046. <https://doi.org/10.1016/j.jobbe.2024.111046>
82. Zang B, Peng X, Zhong X, Zhao C, Zhou K (2023) The size distribution measurement and shape quality evaluation method of manufactured aggregate material based on deep learning. *J Test Eval*. <https://doi.org/10.1520/JTE20220529>
83. Liu M, Yue S, Li S, Du Y, Li B (2022) Research on intelligent detection of concrete aggregate level based on monocular imaging. *Measurement* 194:111036. <https://doi.org/10.1016/j.measurement.2022.111036>
84. Yuan S et al (2022) YOLOv5-Ytiny: a miniature aggregate detection and classification model. *Electronics* 11(11):1743. <https://doi.org/10.3390/electronics11111743>
85. Kirillov, A., Y. Wu, K. He, and R. Girshick (2020). PointRend: image segmentation as rendering. In *Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. p. 9796–9805. <https://doi.org/10.1109/CVPR42600.2020.00982>.
86. Sheiati S, Nguyen H, Kinnunen P, Ranjbar N (2023) Cementitious phase quantification using deep learning. *Cem Concr Res* 172:107231. <https://doi.org/10.1016/j.cemconres.2023.107231>
87. Wong H, Buenfeld N (2009) Determining the water–cement ratio, cement content, water content and degree of hydration of hardened cement paste: method development and validation on paste samples. *Cem Concr Res* 39(10):957–965. <https://doi.org/10.1016/j.cemconres.2009.06.013>
88. Gaël B, Christelle T, Gilles E, Sandrine G, Tristan S-F (2016) Determination of the proportion of anhydrous cement using SEM image analysis. *Constr Build Mater* 126:157–164. <https://doi.org/10.1016/j.conbuildmat.2016.09.037>
89. Zhao Y, Ding B, Zhang Y, Yang L, Hao X (2021) Online cement clinker quality monitoring: a soft sensor model based on multivariate time series analysis and CNN. *ISA Trans* 117:180–195. <https://doi.org/10.1016/j.isatra.2021.01.058>
90. Qian H, Li Y, Yang J, Xie L, Tan KH (2022) Image-based microstructure classification of mortar and paste using convolutional neural networks and transfer learning. *Cem Concr Compos* 129:104496. <https://doi.org/10.1016/j.cemconcomp.2022.104496>
91. Käthler CB, Angst UM, Wagner M, Elsener B (2017) Image analysis for determination of cement content in concrete to improve accuracy of chloride analyses. *Cem Concr Res* 99:1–7. <https://doi.org/10.1016/j.cemconres.2017.04.007>
92. Chen H, Jin Y, Li G, Chu B (2018) Automated cement fragment image segmentation and distribution estimation via a holistically-nested convolutional network and morphological analysis. *Powder Technol* 339:306–313. <https://doi.org/10.1016/j.powtec.2018.08.015>
93. Arganda-Carreras I et al (2017) Trainable weka segmentation: a machine learning tool for microscopy pixel classification. *Bioinformatics* 33(15):2424–2426. <https://doi.org/10.1093/bioinformatics/btx180>

94. Nguyen TT, Bui HH, Ngo TD, Nguyen GD (2017) Experimental and numerical investigation of influence of air-voids on the compressive behaviour of foamed concrete. *Mater Des* 130:103–119. <https://doi.org/10.1016/j.matdes.2017.05.054>
95. Al-Maharma AY, Patil SP, Markert B (2020) Effects of porosity on the mechanical properties of additively manufactured components: a critical review. *Mater Res Express* 7(12):122001. <https://doi.org/10.1088/2053-1591/abce5d>
96. Tao J et al (2022) Deep learning based automated segmentation of air-void system in hardened concrete surface using three dimensional reconstructed images. *Constr Build Mater* 324:126717. <https://doi.org/10.1016/j.conbuildmat.2022.126717>
97. Wei Y et al (2021) Automated detection and segmentation of concrete air voids using zero-angle light source and deep learning. *Autom Constr* 130:103877. <https://doi.org/10.1016/j.autcon.2021.103877>
98. Lu H, Alymov E, Shah S, Peterson K (2017) Measurement of air void system in lightweight concrete by X-ray computed tomography. *Constr Build Mater* 152:467–483. <https://doi.org/10.1016/j.conbuildmat.2017.06.180>
99. Yuan J, Wu Y, Zhang J (2018) Characterization of air voids and frost resistance of concrete based on industrial computerized tomographical technology. *Constr Build Mater* 168:975–983. <https://doi.org/10.1016/j.conbuildmat.2018.01.117>
100. Koenig A (2020) Analysis of air voids in cementitious materials using micro X-ray computed tomography (μ XCT). *Constr Build Mater* 244:118313. <https://doi.org/10.1016/j.conbuildmat.2020.118313>
101. Chen Y, Çopuroğlu O, Rodriguez CR, de Mendonca Filho FF, Schlangen E (2021) Characterization of air-void systems in 3D printed cementitious materials using optical image scanning and X-ray computed tomography. *Mater Charact* 173:110948. <https://doi.org/10.1016/j.matchar.2021.110948>
102. Wolter S, Uhre FAH, Hasholt MT, Dahl VA, Anton F (2019) Air void analysis of hardened concrete by means of photogrammetry. *Constr Build Mater* 226:953–964. <https://doi.org/10.1016/j.conbuildmat.2019.07.203>
103. Hong M, Lei D, Hu F, Chen Z (2023) Assessment of void and crack defects in early-age concrete. *J Build Eng* 70:106372. <https://doi.org/10.1016/j.jobe.2023.106372>
104. Fonseca PC, Scherer GW (2015) An image analysis procedure to quantify the air void system of mortar and concrete. *Mater Struct* 48:3087–3098. <https://doi.org/10.1617/s11527-014-0381-9>
105. Murotani T, Igarashi S, Koto H (2019) Distribution analysis and modeling of air voids in concrete as spatial point processes. *Cem Concr Res* 115:124–132. <https://doi.org/10.1016/j.cemconres.2018.10.008>
106. Ojala T, Chen Y, Punkki J, Al-Neshawy F (2021) Characteristics of entrained air voids in hardened concrete with the method of digital image analysis coupled with schwartz-saltykov conversion. *Materials* 14(9):2439. <https://doi.org/10.3390/ma14092439>
107. Kim Y, Hanif A, Usman M, Park W (2019) Influence of bonded mortar of recycled concrete aggregates on interfacial characteristics—porosity assessment based on pore segmentation from backscattered electron image analysis. *Constr Build Mater* 212:149–163. <https://doi.org/10.1016/j.conbuildmat.2019.03.265>
108. Ozkul T, Kucuk I (2011) Design and optimization of an instrument for measuring bughole rating of concrete surfaces. *J Franklin Inst* 348(7):1377–1392. <https://doi.org/10.1016/j.jfranklin.2010.04.004>
109. Liu B, Yang T (2017) Image analysis for detection of bugholes on concrete surface. *Constr Build Mater* 137:432–440. <https://doi.org/10.1016/j.conbuildmat.2017.01.098>
110. Zhao Z, Liu T, Zhao X (2021) Variable selection from image texture feature for automatic classification of concrete surface voids. *Comput Intell Neurosci* 2021(1):5538573. <https://doi.org/10.1155/2021/5538573>
111. Hoang N-D, Nguyen Q-L (2020) A novel approach for automatic detection of concrete surface voids using image texture analysis and history-based adaptive differential evolution optimized support vector machine. *Adv Civil Eng* 2020(1):4190682. <https://doi.org/10.1155/2018/5989246>
112. Yoshitake I, Maeda T, Hieda M (2018) Image analysis for the detection and quantification of concrete bugholes in a tunnel lining. *Case Stud Constr Mater* 8:116–130. <https://doi.org/10.1016/j.cscm.2018.01.002>
113. Liu B, Yang T, Xie Y (2017) Factors influencing bugholes on concrete surface analyzed by image processing technology. *Constr Build Mater* 153:897–907. <https://doi.org/10.1016/j.conbuildmat.2017.07.156>
114. Wei F, Yao G, Yang Y, Sun Y (2019) Instance-level recognition and quantification for concrete surface bughole based on deep learning. *Autom Constr* 107:102920. <https://doi.org/10.1016/j.autcon.2019.102920>
115. Wei F et al (2022) Deep learning-based automatic detection and evaluation on concrete surface bugholes. *Comput Model Eng Sci* 131(2):619–637. <https://doi.org/10.32604/cmescs.2022.019082>
116. Sun Y, Yang Y, Yao G, Wei F, Wong M (2021) Autonomous crack and bughole detection for concrete surface image based on deep learning. *IEEE Access* 9:85709–85720. <https://doi.org/10.1109/ACCESS.2021.3088292>
117. Zhou S et al (2019) Quick image analysis of concrete pore structure based on deep learning. *Constr Build Mater* 208:144–157. <https://doi.org/10.1016/j.conbuildmat.2019.03.006>
118. Wei W, Ding L, Luo H, Li C, Li G (2021) Automated bughole detection and quality performance assessment of concrete using image processing and deep convolutional neural networks. *Constr Build Mater* 281:122576. <https://doi.org/10.1016/j.conbuildmat.2021.122576>

119. Yao G, Wei F, Yang Y, Sun Y (2019) Deep-learning-based bughole detection for concrete surface image. *Adv Civil Eng* 2019(1):8582963. <https://doi.org/10.1155/2019/8582963>
120. Tian W et al (2024) An improved YOLOv5 model for concrete bubble detection based on area K-means and ECANet. *Mathematics* 12(17):2777. <https://doi.org/10.3390/math12172777>
121. Solak AM, Tenza-Abril AJ, Baeza-Brotons F, Benavente D (2019) Proposing a new method based on image analysis to estimate the segregation index of lightweight aggregate concretes. *Materials* 12(21):3642. <https://doi.org/10.3390/ma12213642>
122. Ke, Y., Caractérisation du comportement mécanique des bétons de granulats légers: expérience. Cergy: University of Cergy-Pontoise, 2008 DOI.
123. Ke Y, Beaucour A, Ortolá S, Dumontet H, Cabrillac R (2009) Influence of volume fraction and characteristics of lightweight aggregates on the mechanical properties of concrete. *Constr Build Mater* 23(8):2821–2828. <https://doi.org/10.1016/j.conbuildmat.2009.02.038>
124. Esmaeilkhani B, Khayat K, Yahia A, Feys D (2014) Effects of mix design parameters and rheological properties on dynamic stability of self-consolidating concrete. *Cem Concr Compos* 54:21–28. <https://doi.org/10.1016/j.cemconcomp.2014.03.001>
125. Navarrete I, Lopez M (2016) Estimating the segregation of concrete based on mixture design and vibratory energy. *Constr Build Mater* 122:384–390. <https://doi.org/10.1016/j.conbuildmat.2016.06.066>
126. Ojala T, Ahmed H, Kuusela P, Seppänen A, Punkki J (2023) Monitoring of concrete segregation using AC impedance spectroscopy. *Constr Build Mater* 384:131453. <https://doi.org/10.1016/j.conbuildmat.2023.131453>
127. Solak AM, Tenza-Abril AJ, García-Vera VE (2022) Adopting an image analysis method to study the influence of segregation on the compressive strength of lightweight aggregate concretes. *Constr Build Mater* 323:126594. <https://doi.org/10.1016/j.conbuildmat.2022.126594>
128. Tenza-Abril AJ, Compañ-Rosique P, Satorre-Cuerda R, Solak AM, Gavotti Freschi D (2021) Smartphone application for determining the segregation index of lightweight aggregate concrete. *Sustainability* 13(22):12564. <https://doi.org/10.3390/su132212564>
129. Cui W, Miao R-C, Yan W-S, Song H-F, Jiang Z-A (2022) Static segregation of fresh high workable concrete based on an image processing method. *Constr Build Mater* 361:129708. <https://doi.org/10.1016/j.conbuildmat.2022.129708>
130. Yoon J, Li Z, Kim H (2024) Evaluation of aggregate segregation in self-consolidating concrete using 3D point cloud analysis. *J Build Eng* 82:108199. <https://doi.org/10.1016/j.jobbe.2023.108199>
131. Nili M, Razmara M, Sadeghi M, Razmara M (2018) Automatic image analysis process to appraise segregation resistance of self-consolidating concrete. *Mag Concr Res* 70(8):390–399. <https://doi.org/10.1680/jmacr.16.00417>
132. Ozen M, Guler M (2020) Quantification of segregation in portland cement concrete based on spatial distribution of aggregate size fractions. *Image Anal Stereol* 39(3):147–159. <https://doi.org/10.5566/ias.2318>
133. Li T et al (2024) A machine vision approach with temporal fusion strategy for concrete vibration quality monitoring. *Appl Soft Comput* 160:111684. <https://doi.org/10.1016/j.asoc.2024.111684>
134. Badarloo B, Kari A, Jafari F (2018) Experimental and numerical study to determine the relationship between tensile strength and compressive strength of concrete. *Civ Eng J* 4(11):2787–2800. <https://doi.org/10.28991/cej-03091198>
135. Fan L, Ding H, Zhong W, Xu C (2023) A rapid approach for determining the mechanical properties of geopolymer concrete based on image processing technology (IPT). *Constr Build Mater* 378:131165. <https://doi.org/10.1016/j.conbuildmat.2023.131165>
136. Li X, Lin H, Chen W, Liang S, Huang L (2023) A numerical study on the tensile splitting of concrete with digital image processing. *J Mater Res Technol* 25:1626–1641. <https://doi.org/10.1016/j.jmrt.2023.06.026>
137. Woldeamanuel MM, Kim T, Cho S, Kim H-K (2023) Estimation of concrete strength using thermography integrated with deep-learning-based image segmentation: case studies and economic analysis. *Expert Syst Appl* 213:119249. <https://doi.org/10.1016/j.eswa.2022.119249>
138. Lee S, Ahn Y, Kim HY (2020) Predicting concrete compressive strength using deep convolutional neural network based on image characteristics. *Comput Mater Contin*. <https://doi.org/10.32604/cmc.2020.011104>
139. Dogan G, Arslan MH, Ceylan M (2017) Concrete compressive strength detection using image processing based new test method. *Measurement* 109:137–148. <https://doi.org/10.1016/j.measurement.2017.05.051>
140. Waris MI, Plevris V, Mir J, Chairman N, Ahmad A (2022) An alternative approach for measuring the mechanical properties of hybrid concrete through image processing and machine learning. *Constr Build Mater* 328:126899
141. Celik G, Ozdemir M (2024) Determination of concrete compressive strength from surface images with the integration of CNN and SVR methods. *Measurement* 238:115331. <https://doi.org/10.1016/j.measurement.2024.115331>
142. Jang Y, Ahn Y, Kim HY (2019) Estimating compressive strength of concrete using deep convolutional neural networks with digital microscope images. *J Comput Civil Eng* 33(3):04019018. [https://doi.org/10.1061/\(ASCE\)CP.1943-5487.0000837](https://doi.org/10.1061/(ASCE)CP.1943-5487.0000837)
143. Kozubal JV et al (2022) Automatic strength assessment of the virtually modelled concrete interfaces based on shadow-light images. *Constr Build Mater* 359:129296. <https://doi.org/10.1016/j.conbuildmat.2022.129296>

144. Shiuly A, Dutta D, Mondal A (2022) Assessing compressive strengths of mortar and concrete from digital images by machine learning techniques. *Front Struct Civ Eng* 16(3):347–358. <https://doi.org/10.1007/s11709-022-0819-z>
145. Fang X et al (2022) Influence of mesoscopic pore characteristics on the splitting-tensile strength of cellular concrete through deep-learning based image segmentation. *Constr Build Mater* 315:125335. <https://doi.org/10.1016/j.conbuildmat.2021.125335>
146. Güçlüer K (2020) Investigation of the effects of aggregate textural properties on compressive strength (CS) and ultrasonic pulse velocity (UPV) of concrete. *J Build Eng* 27:100949. <https://doi.org/10.1016/j.jobbe.2019.100949>
147. Mahankali S, Valikala G (2022) Comparison of compressive strength of M30 grade concrete with destructive and nondestructive procedures using digital image processing as a technique. *Adv Civil Eng* 2022(1):4649660. <https://doi.org/10.1155/2022/4649660>
148. Wang G et al (2023) UAV-YOLOv8: a small-object-detection model based on improved YOLOv8 for UAV aerial photography scenarios. *Sensors* 23(16):7190. <https://doi.org/10.3390/s23167190>
149. Baldovino RG et al (2024) Comprehensive analysis on ultralytics-supported YOLO models for detection and recognition of large office objects for indoor navigation. *Procedia Comput Sci* 246:3851–3858. <https://doi.org/10.1016/j.procs.2024.09.158>
150. Pang B, Nijkamp E, Wu YN (2020) Deep learning with tensorflow: a review. *J Educ Behav Stat* 45(2):227–248. <https://doi.org/10.3102/1076998619872761>
151. Ketkar N, Moolayil J (2021) Introduction to pytorch. Deep learning with python: learn best practices of deep learning models with PyTorch. Springer, pp 27–91. https://doi.org/10.1007/978-1-4842-5364-9_2
152. Bradski G (2000) The opencv library. *Dr Dobb's J: Software Tools for the Professional Programmer* 25(11):120–123
153. Collins TJ (2007) ImageJ for microscopy. *Biotechniques* 43(sup1):S25–S30. <https://doi.org/10.2144/000112517>
154. Lindquist W, Montney R (2019) Comparison of spacing factors as measured by the air-void analyzer and ASTM C457. *Int J Pavement Eng* 20(2):183–190. <https://doi.org/10.1080/10298436.2017.1279485>
155. Ulusoy U (2023) A review of particle shape effects on material properties for various engineering applications: from macro to nanoscale. *Minerals* 13(1):91. <https://doi.org/10.3390/min13010091>
156. Fan C-L, Chung Y-J (2022) Supervised machine learning-based detection of concrete efflorescence. *Symmetry* 14(11):2384. <https://doi.org/10.3390/sym14112384>
157. Flotzinger, J., P.J. Rösch, and T. Braml (2024) dacl10k: benchmark for semantic bridge damage segmentation. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*. p. 8626–8635.
158. Coenen M, Schack T, Beyer D, Heipke C, Haist M (2021) Semi-supervised segmentation of concrete aggregate using consensus regularisation and prior guidance. *ISPRS Ann Photogrammetry, Remote Sens Spatial Inf Sci* 2:83–91
159. Pan X, Shi Z, Shi C, Ling T-C, Li N (2017) A review on concrete surface treatment Part I: types and mechanisms. *Constr Build Mater* 132:578–590. <https://doi.org/10.1016/j.conbuildmat.2016.12.025>
160. Qoku E, Xu K, Li J, Monteiro PJ, Kurtis KE (2023) Advances in imaging, scattering, spectroscopy, and machine learning-aided approaches for multiscale characterization of cementitious systems. *Cem Concr Res* 174:107335. <https://doi.org/10.1016/j.cemconres.2023.107335>
161. Nguyen HN, Lee C (2018) Effects of hyper-parameters and dataset on CNN training. *J IEEE* 22(1):14–20. <https://doi.org/10.7471/ikjee.2018.22.1.14>
162. Plank J, Sakai E, Miao C, Yu C, Hong J (2015) Chemical admixtures—chemistry, applications and their impact on concrete microstructure and durability. *Cem Concr Res* 78:81–99. <https://doi.org/10.1016/j.cemconres.2015.05.016>
163. Wang A et al (2024) Yolov10: real-time end-to-end object detection. *Adv Neural Inf Process Syst* 37:107984–108011. <https://doi.org/10.52202/079017-3429>
164. Khanam, R. and M. Hussain (2024) Yolov11: an overview of the key architectural enhancements. *arXiv preprint arXiv:2410.17725*. <https://doi.org/10.48550/arXiv.2410.17725>
165. Tian, Y., Q. Ye, and D. Doermann (2025) Yolov12: attention-centric real-time object detectors. *arXiv preprint arXiv:2502.12524*. <https://doi.org/10.48550/arXiv.2502.12524>
166. Wei J, Chen J, Wang Y, Luo H, Li W (2023) Improved deep learning image classification algorithm based on Swin Transformer V2. *PeerJ Comput Sci* 9:e1665. <https://doi.org/10.7717/peerj-cs.1665>
167. Mahboob Z, Khan MA, Lodhi E, Nawaz T, Khan US (2024) Using segformer for effective semantic cell segmentation for fault detection in photovoltaic arrays. *IEEE J Photovolt*. <https://doi.org/10.1109/JPHOTOV.2024.3450009>
168. Sajol, M.S.I., A.J. Hasan, M.S. Islam, and M.S. Rahman (2024) A convnext v2 approach to document image analysis: enhancing high-accuracy classification. In *Proceedings of the 2024 IEEE 3rd Conference on Information Technology and Data Science (CITDS)*. IEEE. p. 1–6. <https://doi.org/10.1109/CITDS62610.2024.10791343>.

Authors and Affiliations

Mati Ullah¹  · **Vagelis Plevris¹** 

✉ Vagelis Plevris
vplevris@qu.edu.qa
Mati Ullah
matiullah@qu.edu.qa

¹ Department of Civil and Environmental Engineering, College of Engineering, Qatar University, Doha, Qatar