



Research
Civil Engineering Materials—Review

Advancements in Machine Learning for the Development of Cementitious Composites Toward an Intelligent and Green Lifecycle: A State-of-the-Art Review

Jinyang Jiang^{*}, Junlin Lin, Lin Jin, Fengjuan Wang, Zhiyong Liu, Yingze Li, Zeyu Lu

State Key Laboratory of Materials for Major Infrastructure Engineering, School of Materials Science and Engineering, Southeast University, Nanjing 211189, China

ARTICLE INFO

Article history:

Received 25 August 2024

Revised 22 October 2025

Accepted 13 February 2026

Available online xxxx

Keywords:

Machine learning

Cementitious composites

Intelligent

Green

Lifecycle

ABSTRACT

As fundamental construction materials, cementitious composites face significant challenges under conventional development approaches, including carbon-intensive production, resource-intensive experimentation, and inefficient design processes. With the emergence of machine learning (ML) as a transformative solution to these limitations, this study presents a state-of-the-art review of existing research to highlight its potential in advancing the development of cementitious composites with intelligent and green lifecycles. The review first provides a foundational introduction to ML concepts and then proposes a novel four-quadrant classification framework to systematically organize current ML applications in the field. The ML-driven innovations integrate the component–structure–process–performance relationships of cementitious composites through sustainable material selection, effective characterization, accurate performance prediction, and optimized inverse design, collectively promoting a paradigm shift toward intelligent and green lifecycles. Furthermore, critical implementation challenges are examined across technical, methodological, and operational dimensions, together with corresponding solution strategies. This review ultimately offers both a conceptual framework and practical implementation guidelines for the development of next-generation sustainable construction materials.

© 2026 THE AUTHORS. Published by Elsevier LTD on behalf of Chinese Academy of Engineering and Higher Education Press Limited Company. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

Cementitious composites are fundamental to civil engineering development, with consumption levels second only to water [1–3]. However, traditional research and practical approaches for these materials are often labor-intensive, time-consuming, and environmentally detrimental. For example, the carbon footprint of cement, the core component of cementitious composites, accounts for approximately 7% of global carbon emissions, primarily due to the calcination process involved in clinker production [4–6]. Moreover, current research and design practices largely rely on experimental trial-and-error methods, resulting in inefficiencies and increased material waste. These limitations highlight the urgent need for intelligent strategies that can improve research efficiency while reducing environmental impacts [7].

Current efforts to promote the development of low-carbon cementitious composites primarily focus on reducing dependence

on traditional cement clinker. One widely adopted strategy involves the use of supplementary cementitious materials (SCMs), such as fly ash, ground granulated blast furnace slag, and silica fume, to partially replace cement clinker in composite mixtures [8–14]. Building upon this approach, limestone calcined clay cement (LC³), which combines both limestone and calcined clay, enables further clinker reduction and represents a significant innovation in low-carbon composite production [15]. In parallel, numerical modeling techniques have been increasingly investigated to optimize material composition and performance [16–18]. Nevertheless, the inherent complexity of cementitious composites, arising from their diverse constituents and multiscale structural features, poses challenges for accurate data acquisition and leads to strong interdependencies among composition, processing, structure, and performance [19].

The rapid development of artificial intelligence (AI), particularly machine learning (ML) [20–22], offers sustainable and efficient solutions for advancing cementitious composite research [23–25]. The ML employs data-driven algorithms that allow computational models to learn patterns from data, make decisions, and pre-

^{*} Corresponding author.

E-mail address: jiangjinyang16@163.com (J. Jiang).

<https://doi.org/10.1016/j.eng.2026.02.012>

2095–8099/© 2026 THE AUTHORS. Published by Elsevier LTD on behalf of Chinese Academy of Engineering and Higher Education Press Limited Company. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

dict outcomes without being explicitly programmed [26,27]. In cementitious composite research, ML has emerged as a powerful tool for enabling intelligent and sustainable transformations across the material lifecycle. During raw material selection, ML optimizes mixture formulations by maximizing the use of SCMs and recycled constituents, thereby reducing clinker consumption and associated carbon emissions [28]. In the design stage, ML facilitates inverse design by leveraging well-trained predictive models to identify optimal compositions, directly reducing material consumption associated with trial-and-error experimentation [29–33]. During production, ML enhances manufacturing processes and construction techniques to minimize energy consumption and waste generation while improving resource efficiency [34,35]. In performance evaluation, ML enables efficient and nondestructive assessment of key properties, such as mechanical behavior and durability, which are traditionally time- and labor-intensive to measure [36–38]. In addition, ML automates microstructural image analysis and feature extraction, which are essential for phase identification (e.g., solid particles and pores) and for advancing fundamental material understanding [39–42]. For in-service monitoring, ML supports the prediction of long-term performance and service life, enabling proactive maintenance strategies that enhance durability and reduce environmental impacts [43–45]. At the end-of-life stage, ML contributes to recyclability assessment and waste valorization, thereby promoting circular economy principles for cementitious materials [46–48]. Collectively, these ML-driven innovations establish new research paradigms and foster eco-efficient practices for cementitious composites [49]. Given these significant advances, a comprehensive review is necessary to synthesize current knowledge, identify critical research gaps, and guide the future development of sustainable and intelligent cementitious composites.

Several existing reviews have documented progress in ML applications for cementitious composite research. Most studies have focused primarily on two aspects: ① performance prediction, such as mechanical properties and durability, based on material composition and environmental parameters [50–52], and ② application-oriented categorizations, including material characterization, property prediction, and crack detection [53]. Although ML's capability for parameter-driven performance prediction is widely recognized in the materials community, underlying mechanisms and implementation workflows are often obscured by specialized terminology and perceived “black-box” characteristics. At the same time, rapid advances in AI-based image applications, such as microstructural analysis and damage identification, have gained increasing attention [54], yet systematic dissemination, particularly among practitioners in civil engineering, remains limited. Moreover, existing studies lack a holistic framework that clearly articulates the interconnections among diverse ML applications and their collective role in advancing green and intelligent cementitious materials. These gaps underscore the need for a state-of-the-art review that synthesizes cross-disciplinary insights, clarifies methodological pipelines, and establishes actionable guidelines for scalable implementation.

Accordingly, this study presents a comprehensive and systematic review of ML applications in cementitious composites, with the aim of guiding the field toward an intelligent and green lifecycle. This objective is achieved through four primary tasks: ① providing foundational knowledge to bridge the gap for researchers without AI expertise; ② developing a novel classification framework to systematically organize existing research; ③ examining the contributions of ML in intelligent and green design across the material lifecycle; and ④ presenting a practical challenge–solution matrix to support real-world deployment. This review is expected to bridge the knowledge gap between the AI and cement research

communities while offering a roadmap for the development of next-generation sustainable construction materials.

2. Background

This section provides a concise introduction to ML, including its fundamental concepts, general workflows, and key terminology, with the aim of enabling researchers from nonspecialized backgrounds to better follow with this study. Additional details are provided in the Appendix A.

2.1. ML overview and general workflow

The ML is a core branch of AI that enables computers to learn from data and autonomously improve their performance through algorithms and statistical models, without explicit programming [23,55]. Deep learning (DL), a specialized subset of ML, employs multi-layer neural networks (NNs) to perform advanced pattern recognition and data analysis (Fig. S1 in Appendix A) [56,57].

The fundamental principle of ML is its data-driven approach, which differs markedly from conventional problem-solving approaches (Fig. 1). Traditional methods typically aim to derive an exact solution, denoted as $F(x)$, but often suffer from inefficiency, high cost, and limited adaptability. In contrast, data-driven approach provides a more automatic and efficient alternative by generating an approximate solution $H(x)$ which iteratively approaches the true solution by leveraging available data [58]. The discrepancy between $H(x)$ and $F(x)$ is quantified using a loss function, and the core objective of ML is to minimize this function through an optimization process.

A typical data-driven workflow consists of several key steps, including data acquisition and preprocessing, model construction, training, evaluation, and application. These steps are described in detail in the following subsections.

2.2. Data acquisition and preprocessing

Depending on the specific objective, ML tasks can be categorized into four primary types: classification, regression, clustering, and dimensionality reduction (Section S2 in Appendix A). Accordingly, ML approaches can also be classified based on learning strategies, including supervised, unsupervised, semi-supervised, and reinforcement learnings. Regardless of the task type, data acquisition and preprocessing are essential steps, as both data quantity and quality directly influence ML model performance [59]. Data acquisition involves collecting information from diverse sources, such as experimental measurements, numerical simulations, and open-access databases. Although data may appear in various forms to humans, including images, text, and audio, all data are ultimately converted into numerical representations and preprocessed before being fed into an ML model. Based on their format and semantic meaning, data can be broadly classified as image and non-image data. Data preprocessing aims to prepare raw data in a form that maximizes the model's learning efficiency. Common preprocessing procedures include data cleaning, integration, transformation, augmentation, and feature engineering [60–62]. A summary of data types and corresponding preprocessing methods is provided in Table S1 in Appendix A.

2.3. ML models

The ML model constitutes the core component of the data-driven framework, and its performance depends strongly on the task requirements and the characteristics of the available data [63]. A wide range of ML models based on different algorithmic

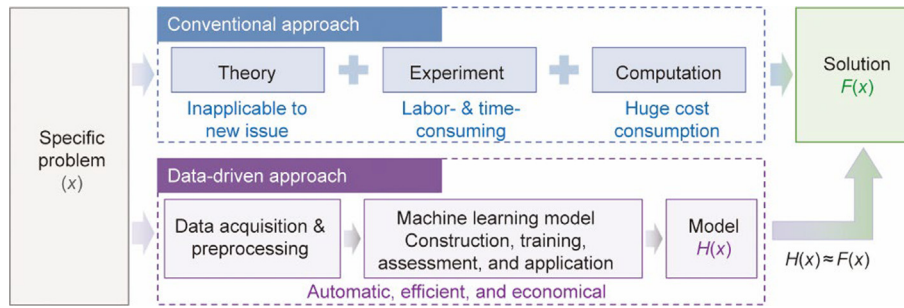


Fig. 1. Comparison between conventional and data-driven approaches for problem solving.

principles has been developed and can be broadly categorized according to their underlying architectures. For non-image data, artificial neural networks (ANN), support vector machines (SVM), and decision trees (DT) are among the most widely recognized and commonly used algorithms. For image-based data, convolutional neural networks (CNN) and generative adversarial networks (GAN) represent two fundamental model architectures. Schematic illustrations of these models are presented in Fig. 2. Additional mainstream ML models, along with their principles, advantages, limitations, and application scenarios, are summarized in Table S2 in Appendix A.

2.4. Model training, assessment, and analysis

Dataset partitioning is a critical step in ensuring the reliability and generalizability of ML models. Typically, datasets are divided into three subsets: training, validation, and testing sets, each serving a unique purpose in model development [64,65]. The training set usually comprises the largest portion of the data, often at least 60%, and is used to learn the relationships between inputs and outputs. The validation set is employed during training to guide model architecture selection and hyperparameter tuning, whereas the testing set is reserved for the final evaluation of model performance after the training is complete.

Model evaluation is essential for assessing effectiveness and reliability, and the choice of evaluation metrics depends on the specific task objectives [66]. For classification tasks, evaluation focuses on the correctness of predicted class labels, assessing the model's ability to distinguish among categories. Common metrics include confusion matrices, accuracy, precision, recall, and the F1-score. For regression tasks, evaluation emphasizes the accuracy of continuous predictions by measuring the agreement between predicted and observed values. Typical metrics include the coefficient of determination (R^2) and mean error (ME). Detailed definitions and calculation formulas are provided in Table S3 in Appendix A.

Commonly used loss functions for regression tasks include zero-one loss, L1 loss, mean squared error (MSE) loss, squared loss, logarithmic loss, and exponential loss, whereas cross-entropy loss and hinge loss are typically employed for classification tasks. It should be noted that loss functions are primarily used during the training process to guide parameter optimization, whereas evaluation metrics are applied to assess the final model's performance and generalization ability on unseen data.

To enhance transparency and reliability, advanced interpretability techniques are increasingly integrated into ML models to elucidate the data-driven factors underlying predictions [67]. Common interpretability methods include Shapley additive explanations (SHAP) [68,69], partial dependence plots (PDP) [70], and the cosine amplitude method (CAM) [71,72], as summarized in

Table S4 in Appendix A. These techniques provide insights into model decision logic and enhance the credibility and applicability of ML models in materials science.

3. ML in cementitious composites research

This section provides a comprehensive review of ML applications in cementitious composite research. It further highlights the critical role of ML in enabling intelligent and green innovations in cementitious composites, underscoring its significant potential to promote sustainability in both academic research and engineering practices.

3.1. Overview

The literature collection and analysis methodology consisted of three sequential stages, as summarized in Table 1. Stage 1 involved a comprehensive search of the Web of Science database to ensure broad coverage of relevant studies. Stage 2 focused on filtering the retrieved literature based on publication quality and year to maintain both rigor and diversity. Stage 3 entailed a detailed screening of the remaining studies to ensure that the selected literature was representative of current research trends. As illustrated in Fig. 3, the number of publications addressing ML applications in cementitious composites has increased steadily from 2001 to 2025. The accompanying keyword co-occurrence network further demonstrates the broad thematic scope of this research area, highlighting the diversity and interdisciplinary nature of existing ML applications.

Current ML applications in cementitious composite research can be categorized into seven specific tasks. As shown in Fig. 4, these tasks are traditionally classified according to ML perspectives such as learning paradigms and model types. However, with the rapid evolution of ML methodologies, multiple learning strategies are often integrated within a single task, resulting in blurred boundaries among categories. Furthermore, the increasing complexity of technical terminology and algorithmic diversity can pose challenges for materials researchers, potentially reducing readability and accessibility.

To address these challenges, we propose an alternative categorization (Fig. 5). This new approach classifies ML applications into four distinct quadrants based on the nature of the input and output data—specifically, whether they are image-based or non-image-based. This refined classification provides a clearer perspective on the diversity of ML applications in cementitious composite research, focusing on data characteristics rather than the complexity of the algorithms. Additionally, the quadrant-based approach offers a strategic advantage in selecting the most appropriate ML techniques for various research scenarios, thereby contributing to the synergistic development of intelligent and green lifecycle practices in the field.

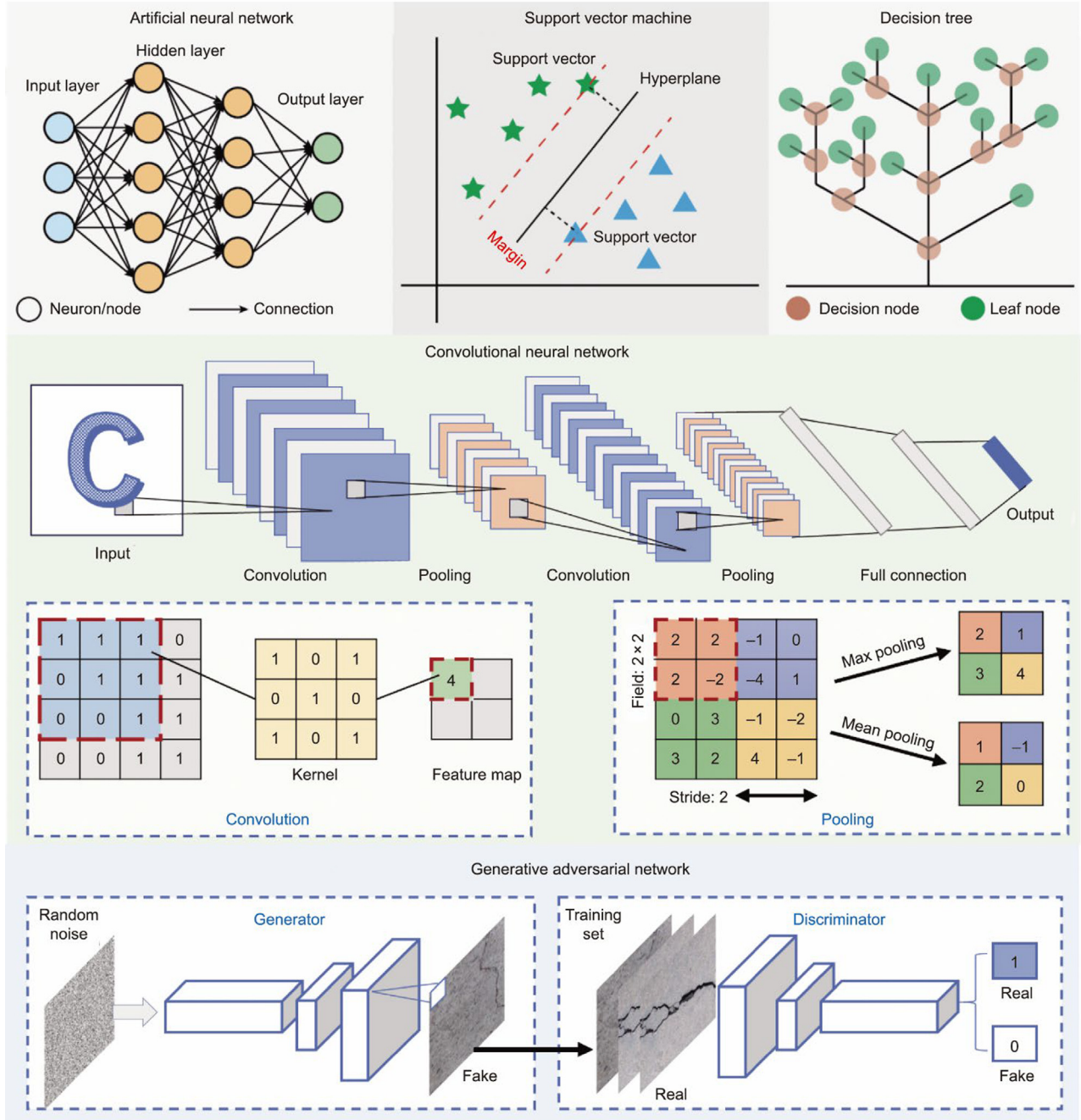


Fig. 2. Schematic of ANN, SVM, DT, CNN, and GAN, along with their key components.

3.2. Quadrant I: from non-image to non-image

Quadrant I encompasses applications in which both input and output data are non-image-based, with performance prediction and inverse design representing the predominant tasks. Fig. 6 illustrates the classical workflow for these applications, which consists of three main phases: data acquisition, model construction and evaluation, and practical application. In the initial phase, a comprehensive database of matched input–output pairs is established. Input variables may include material composition, environmental conditions, or processing parameters, whereas outputs typically correspond to performance indicators such as mechanical proper-

ties and durability. After data preprocessing and dataset partitioning, the data are used to train and iteratively refine ML models through evaluation and optimization, resulting in a well-trained predictive model.

The trained model is theoretically applicable to any combination of input variables within the defined data range and serves two primary functions. First, it enables rapid and nondestructive performance evaluation by predicting material properties based on known inputs, thereby reducing the need for extensive and costly experimental or numerical efforts [73]. Second, the model supports inverse design by iteratively adjusting input parameters to achieve target performance outputs, allowing the identification

Table 1
Literature collection steps and constraints.

Stage	Stage 1: comprehensiveness	Stage 2: diversity	Stage 3: representation
Principle	Internet search	Literature screening	Literature analysis
Constraint	Keyword 1: "cement" or "concrete" or "cementitious" Keyword 2: "artificial intelligence" or "machine learning" or "deep learning" Language: English Document types: Article	Q1 journals in the field: Cement and Concrete Research; Cement and Concrete Composites; Construction and Building Materials; and Journal of Building Engineering High-quality general interest journals (impact factor > 10): Engineering Composites Part B: Engineering Citation: > 10 Year: 2019–2025	Performance prediction with non-image inputs and outputs: 120 random samples in equal proportions according to the performance distribution Related to images: further breakdown of research and selection of representative papers

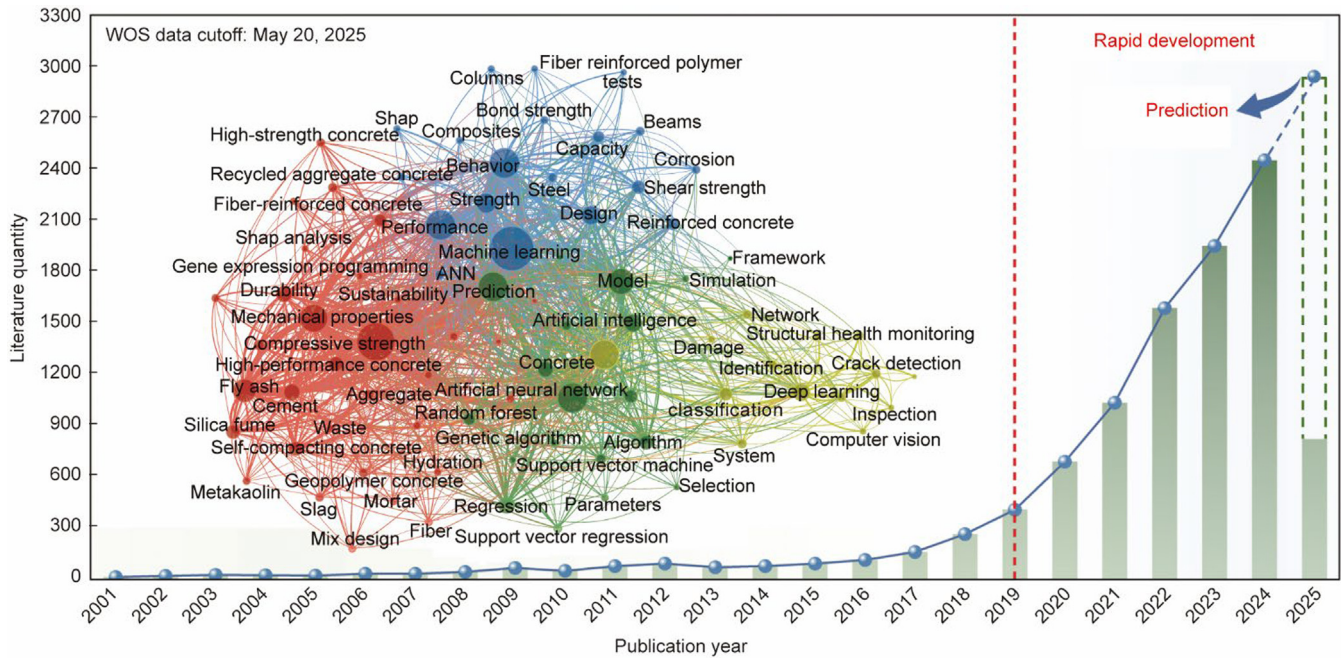


Fig. 3. Growth in publication volume and keyword co-occurrence network for ML applications in cementitious composites research (2001–2025). The literature volume for 2025 represents a forecast based on current growth trends).

of optimal design strategies. Finally, the predicted results can be validated against experimental or field data to ensure accuracy and reliability [74].

3.2.1. Statistical results of ML applications to performance prediction

A statistical analysis was conducted based on 120 research articles published over the past five years [42,73–191] to identify prevailing practices and emerging trends in ML-based performance prediction for cementitious composites. The analysis encompasses a broad range of aspects, including input and output parameters, dataset characteristics, modeling approaches, and evaluation methods, which are further supported by detailed visualization and discussion.

Fig. 7(a) [42,73–191] presents a chord diagram illustrating the relationships between input parameters, such as material composition and exposure conditions, and output parameters, which include key performance indicators of cementitious composites, such as fresh properties, mechanical properties, and durability. Cement, water, and aggregates emerge as the most frequently used input variables, reflecting their fundamental roles in cementitious composite systems. Despite the diversity of constituent materials, ML-based prediction approaches have been successfully applied to a wide range of cementitious composites, including systems incorporating SCMs [76,77], chemical additives such as superplasticizers [78,79], nanomaterials [80], and various fiber reinforce-

ments [81], provided that their contents can be expressed as numerical data. Prediction accuracy and robustness can be further enhanced by incorporating additional information, such as physical properties of the components [74,82], curing regimes and environmental conditions [83,84], and processing parameters [42,85], even when these variables are represented in categorical form. Fig. 7(b) [42,73–191] further classifies the output parameters into detailed subcategories, demonstrating the wide range of performance criteria addressed in the literature. Among these, compressive strength is the most extensively studied property, not only within the mechanical category but also overall, accounting for nearly half of all output parameters. This is likely due to its fundamental importance and ease of measurement. Studies related to structure and durability constitute 16.3% of the total, with permeability (6.8%) receiving the greatest attention within this category. Other durability-related indicators, such as creep and carbonation depth, are also represented. In contrast, research focusing on fresh properties (8.2%) and functional performance (4.1%) remains relatively limited. Table 2 summarizes the specific output parameters corresponding to these categories. Overall, any quantifiable and describable characteristic of cementitious composites can serve as an input or output parameter in ML-based approaches.

Fig. 8 [42,73–191] summarizes the key characteristics of ML features reported in the reviewed cementitious composite studies. Dataset sizes vary considerably, ranging from 50 to 30 000 samples

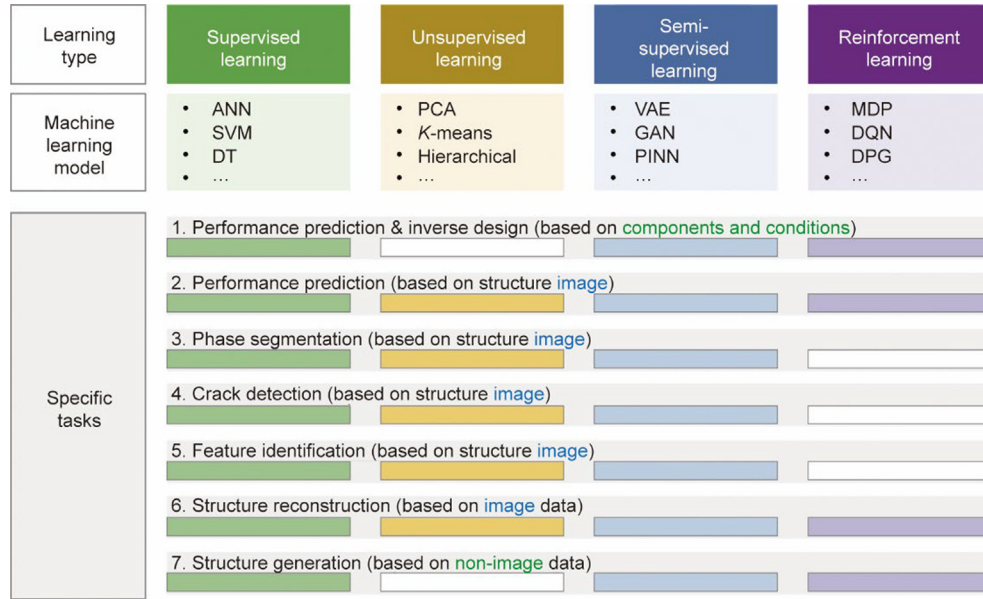


Fig. 4. Categorization of seven specific tasks in cementitious composite research based on ML learning types. PCA: principal component analysis; VAE: variational autoencoder; PINN: physics-informed neural network; MDP: Markov decision process; DQN: deep Q-network; DPG: deterministic policy gradient.

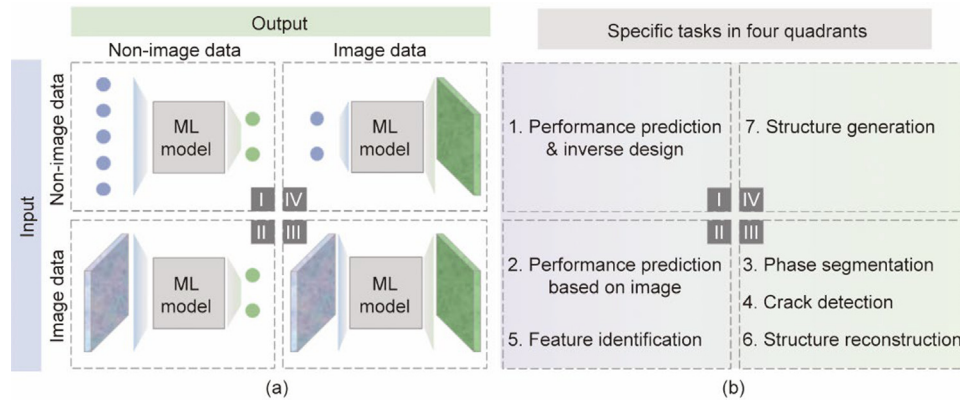


Fig. 5. (a) Division into four quadrants (I–IV) based on input and output data types (image vs non-image); (b) distribution of seven specific tasks across the four quadrants.

(Fig. 8(a)), although approximately half of the datasets contain between 100 and 400 samples. This distribution highlights several significant challenges, including limited availability of large-scale data, the consequent need for effective data generation strategies, and the development of ML methods suitable for small-sample scenarios. As shown in Fig. 8(b), most studies employ model configurations with 4–12 input variables and a single output, indicating a tendency toward focused prediction of specific performance properties while balancing model complexity and data availability. Regarding modeling approaches, Fig. 8(c) shows that NN-, DT-, and SVM-based algorithms are the most frequently adopted, owing to their robustness in capturing the nonlinear behavior of complex material systems. Ensemble learning (EL) techniques, which integrate multiple models to enhance predictive performance, have also attracted increasing attention [91]. For a given task, comparisons among different models are commonly conducted to identify the most suitable approach based on evaluation outcomes. In terms of model evaluation (Fig. 8(d)), MSE/root mean square error (RMSE), R^2 , mean absolute error (MAE), and mean relative error (MRE)/mean absolute percentage error (MAPE) are the most widely used metrics, while other specialized indicators are occasionally applied in specific contexts. With respect to data partition-

ing, the combined training and validation set typically accounts for 70% or 80% of the dataset (Fig. 8(e)). For validation strategies (Fig. 8(f)), k -fold cross-validation (CV) is the most commonly adopted approach, with ten-fold CV being the most prevalent [99], followed by five-fold CV [100] and other methods such as leave-one-out cross-validation (LOOCV) [102,192]. Fig. 8(g) indicates that sensitivity analysis remains underutilized, appearing in only about half of the reviewed studies. When employed, SHAP (39%), tree-based feature importance (26%), and PDP (14%) are the most commonly used techniques [193,194]. Greater emphasis on model interpretability is therefore necessary to enhance practical applicability and to extract physically meaningful insights from ML models.

Representative examples of the three sensitivity analysis techniques are shown in Fig. 9 [195–197]. The SHAP approach (Fig. 9 (a) [195]) visualizes the influence of material composition and curing conditions on the drying shrinkage of magnesium silicate hydrate cement, ranking features by global importance using mean absolute SHAP values, and depicting individual sample effects with colored markers. In contrast, tree-based feature importance (Fig. 9 (b) [196]) directly ranks variable contributions to early cracking risk in foundation slabs, identifying slab thickness (h), hardening rate, and upper surface temperature (T_{up}) as dominant factors.

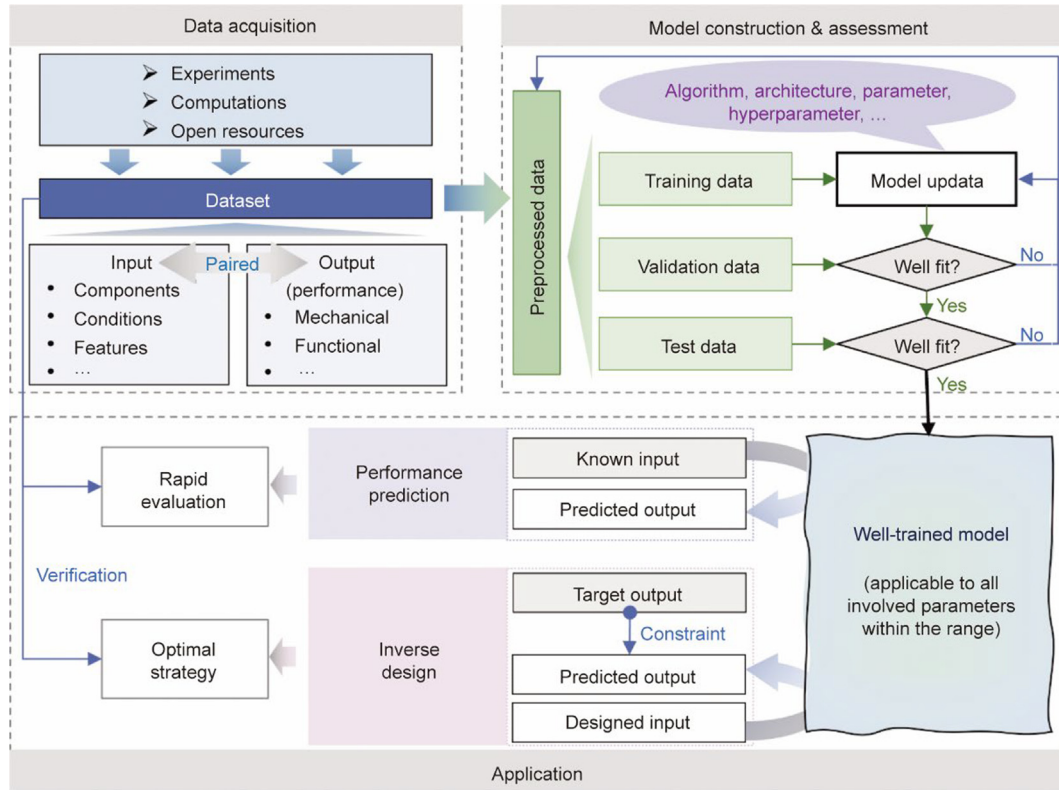


Fig. 6. Workflow overview for ML-based performance prediction and inverse design in cementitious composites.

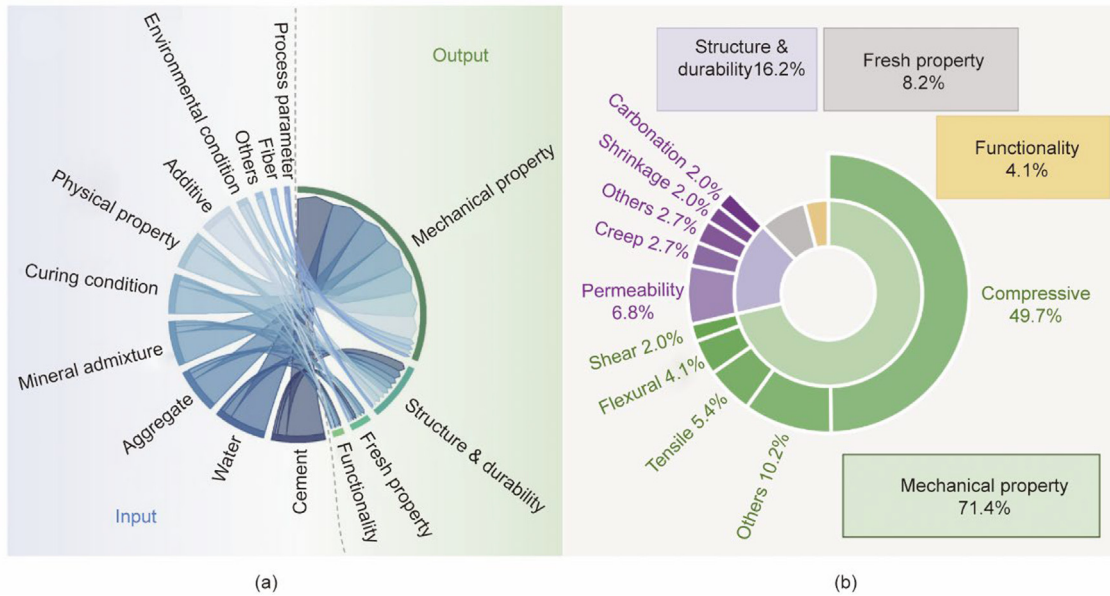


Fig. 7. (a) Relationships between input and output parameters; (b) detailed statistical categorization of output parameters [42,73–191].

The PDP method (Fig. 9(c) [197]) illustrates the global effects of input variables on the compressive strength of stabilized soil, revealing a nonlinear interaction between solid waste content (P_{sw}) and SiO_2 content (C_{si}) via a color-mapped representation.

3.2.2. Physics-informed ML prediction

As noted previously, the core principle of ML involves iteratively refining nonlinear parameters through data-driven approaches to

develop models that closely approximate real-world solutions. However, the effectiveness of these models largely depends on data quality and availability. Although ML models can achieve accurate performance predictions, they often lack direct physical interpretability, requiring additional analyses to elucidate the underlying physical mechanisms. While sensitivity analysis can highlight the relative importance of input variables, it does not inherently explain the physical processes governing material behavior.

Table 2

Specific output parameters corresponding to the categories shown in Fig. 7(b) [42,73–191].

Category	Specific parameters
Fresh property	Slump, apparent viscosity, cohesion, plastic viscosity, and setting time
Functionality	Thermal conductivity, Gamma rays' linear attenuation coefficient, heat flow rate, and electrical resistivity
Mechanical property (others)	Yield stress, relative dynamic elastic modulus, peak strain, and bond strength
Structure & durability (others)	Mass loss rate, porosity, and alkali-silica reaction expansion

Recent advancements, particularly physics-informed neural networks (PINNs), have emerged as promising approaches for cementitious composite research. The PINNs integrate physical laws, typically expressed as partial differential equations (PDEs), directly into the ML model architecture. This integration enables more accurate and physically consistent predictions, even when training data are limited [198]. For example, Zhang et al. [199] developed a PINN-based model, RheologyNet, to predict the thixotropic behavior of cementitious materials by embedding complex rheological PDEs involving the structural parameter, shear stress, and apparent viscosity. This framework enables efficient rheological predictions under data-scarce conditions. Similarly, Wan et al.

[200] applied PINNs to identify the chloride ion diffusion coefficient during unsteady-state immersion and accelerated chloride migration tests.

By integrating governing physical principles, PINNs offer deeper insights into material behavior under diverse conditions, which is essential for advancing the design of intelligent and green cementitious materials. As illustrated in Fig. 10 [201], Rahman and Lu [201] proposed a PINN-based model, EcoBlendNet, that embeds PDEs describing heat generation rates and energy consumption during cement hydration. This model predicts hydration temperature, early-age strength, and carbon emissions for cementitious composites with varying mixture proportions of cement, fly ash, and slag. Quantitative analysis showed that replacing 45%–80% of cement with fly ash and slag could reduce carbon dioxide emissions during hydration by 60%–80%, demonstrating a practical and environmentally friendly solution for the design of green cementitious composites.

3.2.3. Inverse design based on target performance

Inverse design represents an advanced application of predictive modeling, enabling the optimization of material properties by iteratively adjusting input parameters to achieve predefined target outcomes [202]. By leveraging the predictive capabilities of ML models, this approach enables researchers and engineers to tailor material compositions and processing conditions to meet desired performance metrics in an intelligent and sustainable manner,

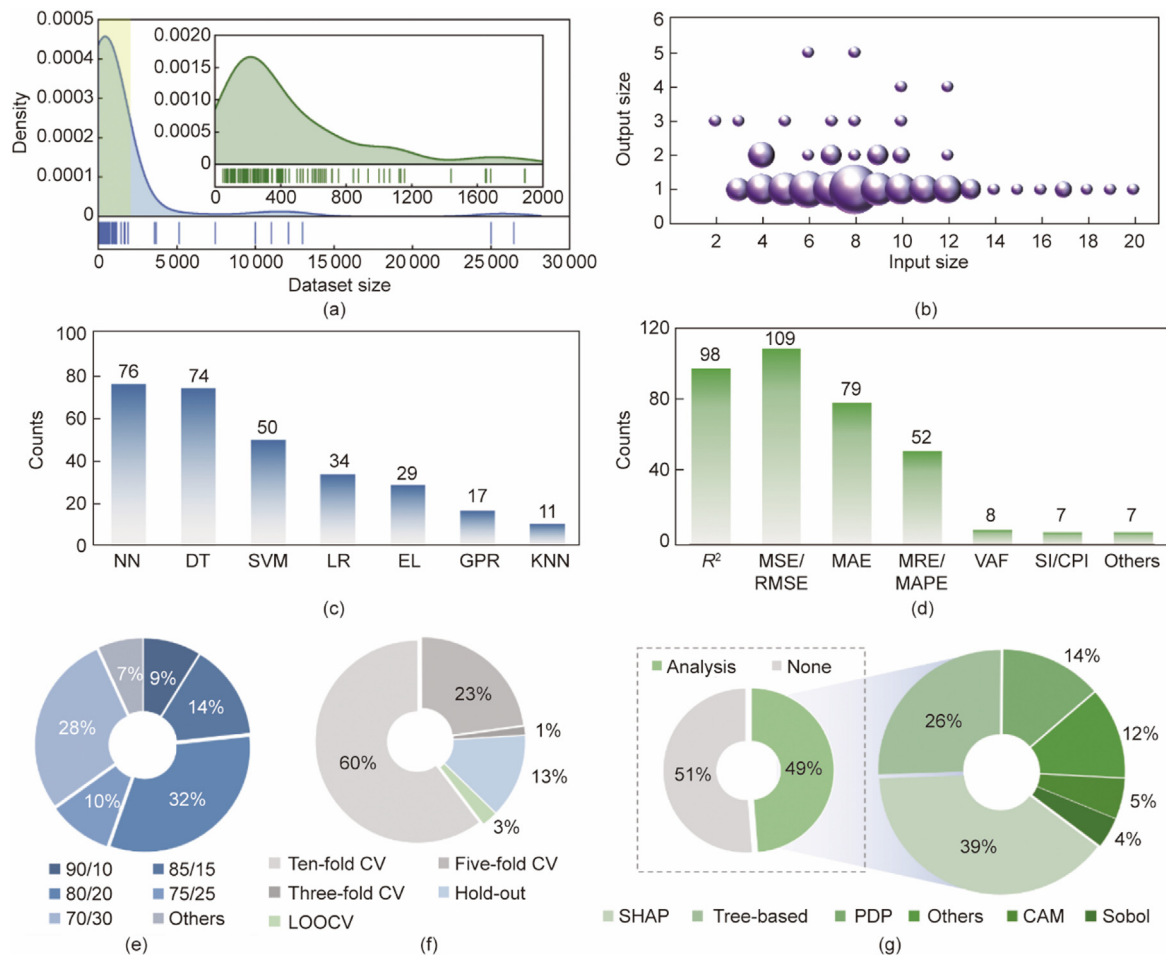


Fig. 8. Statistical summary of key ML features in reviewed studies [42,73–191]. (a) dataset sizes; (b) numbers of input and output variables; (c) primary ML algorithms; (d) evaluation metrics; (e) proportions of generalized training (training and validation) and test sets; (f) validation strategies; (g) sensitivity analysis methods. LR: logistic regression; GPR: gaussian process regression; KNN: k -nearest neighbor; VAF: variance accounted for; SI: scatter index; CPI: composite performance index.

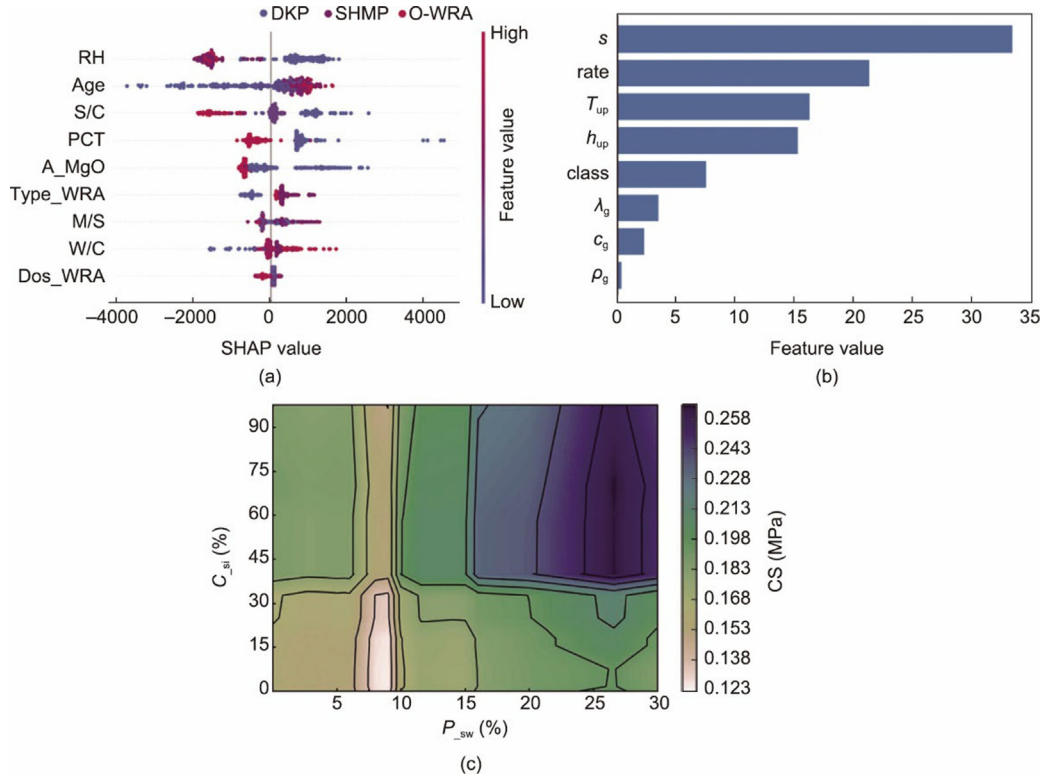


Fig. 9. Examples of sensitivity analysis methods. (a) SHAP (Shapley additive explanations). Reproduced from Ref. [195] with permission. (b) Feature importance [196]. (c) PDP. Reproduced from Ref. [197] with permission. RH: relative humidity; S/C: sand-to-binder ratio; PCT: pre-curing duration; A_MgO: MgO reactivity; Type_WRA: type of water-reducing agent; M/S: magnesium-to-silicate ratio; W/C: water-to-cement ratio; Dos_WRA: dosage of water-reducing agent; DKP: dipotassium hydrogen phosphate; SHMP: sodium hexametaphosphate; O-WRA: other water-reducing agents; s: slab thickness; rate: hardening rate; T_{up} : temperature on the upper surface; h_{up} : heat transfer coefficient on the upper surface; class: concrete compressive strength class; λ_g : soil thermal conductivity coefficient; c_g : soil specific heat capacity; ρ_g : soil density; P_{sw} : solid waste content; C_{si} : SiO_2 content; CS: compressive strength.

thereby reducing reliance on extensive experimental testing [203]. Inverse design is particularly valuable for developing cementitious composites that simultaneously satisfy multiple objectives, including mechanical performance, economic constraints, and environmental considerations [204].

For instance, Li et al. [205] proposed an intelligent design framework for high-strength, low-carbon rice husk ash concrete (RHC). A multi-objective optimization was performed considering compressive strength, material cost, embodied carbon, and embodied energy, resulting in reductions of 6.30% in the carbon footprint and 4.71% in embodied energy for the optimized RHC mixture. Cheng et al. [206] introduced a multi-objective optimization strategy for low-carbon self-compacting concrete (SCC) by integrating a random forest (RF) model with the non-dominated sorting genetic algorithm II (NSGA-II). In this framework, carbon emissions, 28-day rapid chloride penetration, and porosity were defined as primary constraints. The NSGA-II was employed to generate a Pareto-optimal solution set (Fig. 11(a) [206]), followed by a multi-factorial comparison of SCC mixtures designed using the AI-based approach (Fig. 11(b) [206]). The optimized C50 SCC mixture not only met strength requirements and improved durability but also reduced carbon emissions by 10.1% (Fig. 11(c) [206]). This methodology provides an intelligent multi-objective optimization framework that lays a solid foundation for the development and application of low-carbon cementitious composites.

In summary, advancements within Quadrant I—encompassing performance prediction and inverse design—have significantly contributed to the intelligent and green development of cementitious composites. The ML-based performance prediction models provide efficient, data-driven insights into material behavior,

enabling rapid evaluation of key properties without extensive experimental testing. Physics-informed approaches, such as PINNs, further enhance predictive reliability by embedding governing physical laws, even under data-limited conditions. Building upon these predictive models, inverse design facilitates the systematic optimization of material compositions and processing parameters to meet targeted performance objectives. Collectively, these approaches streamline material development, reduce resource consumption, and promote the creation of eco-friendly, high-performance cementitious composites.

3.3. Quadrant II: from image to non-image

In Quadrant II, image data serve as the input, while the outputs are non-image data. Research in this quadrant reflects a shift from traditionally describable non-image inputs toward image-based data, such as microstructural images of cementitious composites. Applications can be broadly classified into two primary approaches, depending on whether the output is categorical or numerical: feature identification through image classification and performance prediction using image-based regression models.

3.3.1. Feature identification based on image classification

Microstructural features play a critical role in cementitious composite research, as they not only reflect variations in material composition and processing methods but also directly influence overall material performance [207]. Conventional analyses of these features rely heavily on expert knowledge and extensive manual effort. Moreover, traditional analytical techniques based on predefined physical descriptors and statistical functions often require

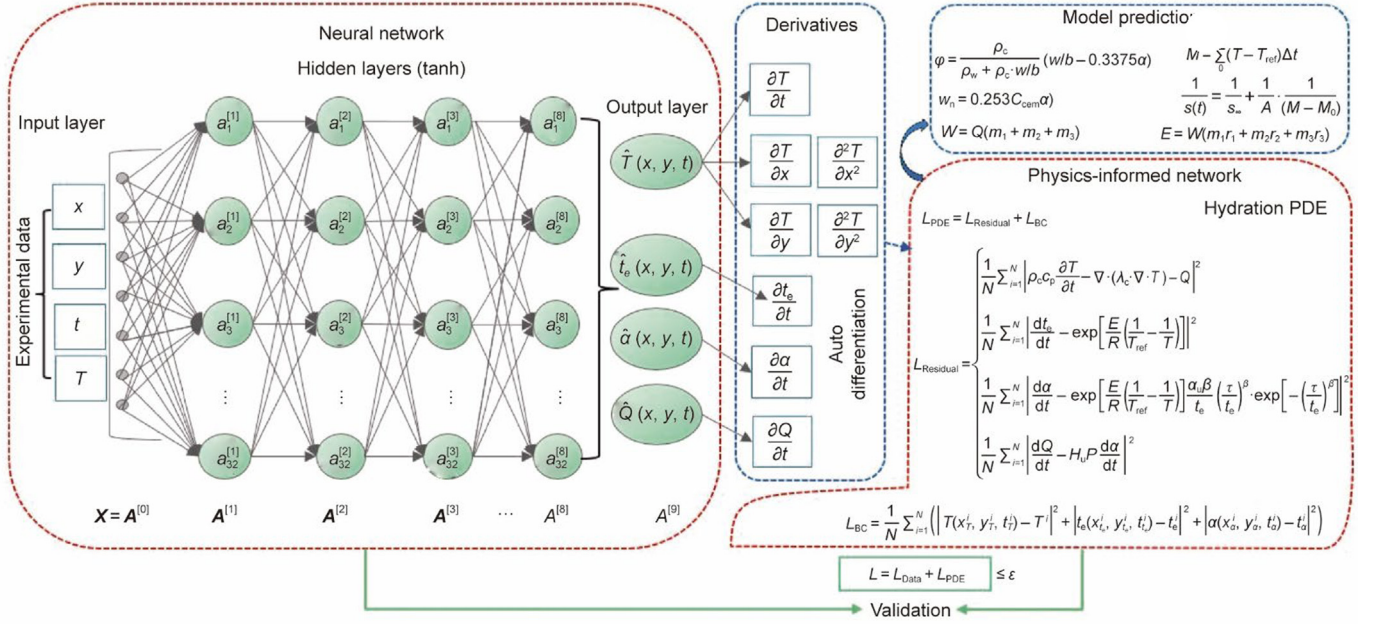


Fig. 10. Architecture of EcoBlendNet. x and y : spatial locations; t : time; T : initial temperature; $\hat{T}(x, y, t)$: predicted temperature; $\hat{t}_e(x, y, t)$: predicted equivalent age; $\hat{\alpha}(x, y, t)$: predicted degree of hydration; $\hat{Q}(x, y, t)$: predicted heat generation rate; tanh: hyperbolic tangent activation function; $\mathbf{X}$: input feature vector; $\mathbf{A}^{(l)}$: activation vector of the l th hidden layer; $a_n^{(l)}$: activation output of the n th neuron in the l th hidden layer; φ : porosity; ρ_c : concrete density; ρ_w : density of water; w/b : water-to-binder ratio; M : concrete maturity at age t ; T_{ref} : reference temperature; Δt : time interval; w_n : chemically bound water content; C_{cem} : cement content; α : degree of hydration; $S(t)$: compressive strength at time t ; S_{∞} : compressive strength after 28 d of curing; M_0 : offset maturity; J : regression coefficient; W : total energy produced; m_1, m_2, m_3 : mass of cement, fly ash, and slag, respectively; E : CO₂ emissions; r_1, r_2, r_3 : emissions factor of cement, fly ash, and slag, respectively; L_{PDE} : partial differential equation residual loss; $L_{Residual}$: residuals of constitutive equations; L_{BC} : residuals occurring at the boundary; L_{Data} : data loss; L : total loss; t_e : theoretical equivalent age; N : number of training data points for training; c_p : heat capacity at constant pressure; λ_c : thermal conductivity of concrete; H_u : total heat of hydration; P : total amount of cement, fly ash, and slag; α_u : ultimate degree of hydration; β : hydration shape parameter; τ : age parameter; R : universal gas constant; x^i, y^i, t^i : the i th training data point at space x, y , and time t , respectively; T^i : actual temperature of the i th training data point; t_e^i : actual equivalent age of the i th training data point; α^i : actual degree of hydration of the i th training data point; ε : error tolerance. Reproduced from Ref. [201] with permission.

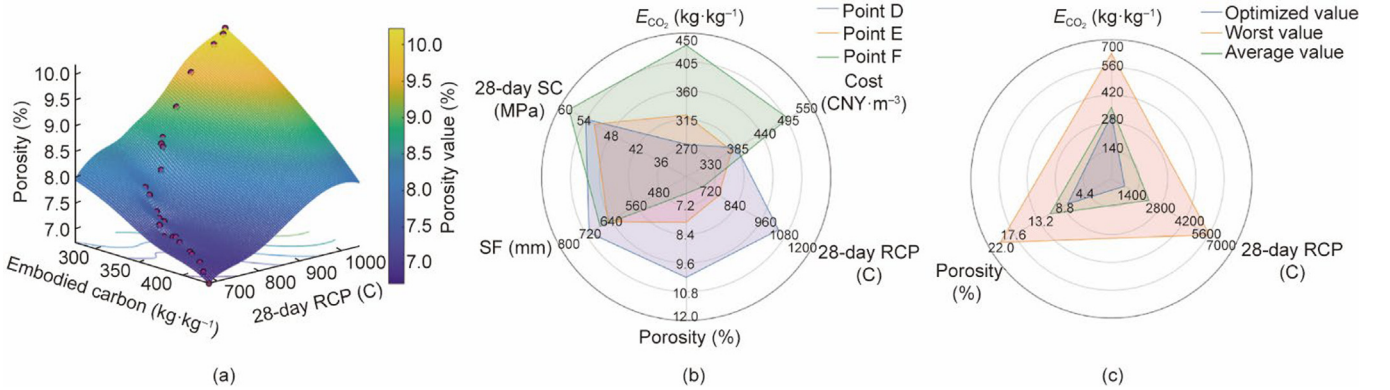


Fig. 11. (a) Pareto-optimal solutions in three-dimensional space for C50; (b) comparison of SCC mixtures designed using the proposed AI method; (c) comparison between optimized and reference SCC mixtures. RCP: rapid chloride permeability; E_{CO_2} : embodied carbon; SC: compressive strength; SF: slump flow; point D/E/F: concrete with different mix proportions. Reproduced from Ref. [206] with permission.

the prior selection of attributes, such as porosity, solidity, and spatial distribution, which may overlook subtle but influential characteristics. In contrast, ML provides a transformative approach by enabling the automated extraction and classification of microstructural features directly from image data. Unlike traditional methods, ML can accommodate the diversity and complexity of microstructures across different cementitious composites and autonomously identify the most informative features for achieving optimal classification performance [208].

When analyzing microstructural features, it is essential to ensure that the studied image size adequately represents the material, a concept referred to as the representative volume element

(RVE). Within an ML framework, Lin et al. [209] proposed a method for determining the RVE of cementitious composites using a straightforward strategy (Fig. 12 [209]). In their study, binarized backscattered electron (BSE) images of three types of cementitious composites were used as inputs, and a CNN model was trained to classify the materials. By preserving image geometry while systematically varying pixel resolution, the authors observed changes in classification accuracy, with 100% accuracy indicating that the image size was sufficiently representative. This approach established a foundation for identifying the minimum image size required for reliable ML-based analysis of cementitious composites. For studies in which image size is fixed, efforts have focused

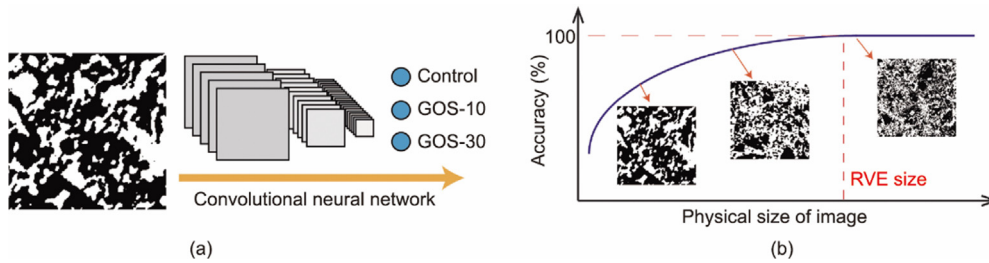


Fig. 12. Schematic illustration of (a) CNN-based classification of different cementitious composites and (b) ML-based estimation of the RVE. Control: ordinary Portland cement (OPC); GOS-10, GOS-30: OPC reinforced by graphene oxide-silica (GOS) with thickness of 10 and 30 nm, respectively. Reproduced from Ref. [209] with permission.

on improving ML model accuracy through architectural and input optimization. For example, Zhang et al. [210] employed modified time-frequency spectrograms as inputs to classify damage-induced acoustic emission signals in ultra-high-performance concrete (UHPC). To balance computational efficiency and classification accuracy, a lightweight network based on ResNet18 was developed and trained from scratch. This model achieved a higher overall classification accuracy (93.94%) and a shorter computational response time compared with traditional approaches (84.1%). As classification accuracy improves, ML models are expected to capture increasingly representative and discriminative microstructural features.

Beyond classification, visualization of features extracted by ML models can reveal patterns and correlations that may be overlooked by human experts, thereby enabling deeper insights into material behavior and supporting the development of advanced, high-performance, and environmentally sustainable cementitious materials. For instance, Lin et al. [209] applied occlusion sensitivity analysis to identify the most influential regions within microstructural images for classification. Their results showed that nano-reinforced cementitious composites exhibit more complex spatial correlations than that of the control samples and that pore spatial arrangement is more critical than porosity alone. Similarly, Sui et al. [211] performed microstructural analyses of LC³ and ordinary Portland cement (OPC) (Fig. 13 [211]), identifying distinct differences in spatial correlation patterns. By visualizing CNN decision-making using deep Taylor decomposition, the study revealed that

LC³ exhibits lower pore connectivity and higher solid particle connectivity, providing insights into its microstructural characteristics and potential performance advantages.

Overall, automated feature identification enhances both the accuracy and efficiency of microstructural analysis and facilitates the exploration of correlations between microstructure and material performance. These capabilities, in turn, inform the optimization of material composition and processing parameters, providing critical technical support for the development of green and low-carbon cementitious systems.

3.3.2. Performance prediction based on image regression

This subsection focuses on the application of image-based regression models to predict the performance of cementitious composites using microstructural images. Unlike the approaches discussed in Quadrant I, which rely on non-image data, image regression leverages the rich and detailed information embedded in microstructural images, which are more directly correlated with material properties. By analyzing these images, researchers can obtain deeper insights into the relationships between microstructural characteristics and macroscopic performance, thereby providing a scientific basis for material design and optimization [212].

For instance, Liang et al. [213] successfully predicted the short-term creep modulus of cement paste using a deep convolutional neural network (DCNN) trained on grayscale X-ray computed tomography images (Fig. 14 [213]). By combining a microscopic lattice model with data-augmentation techniques, the authors

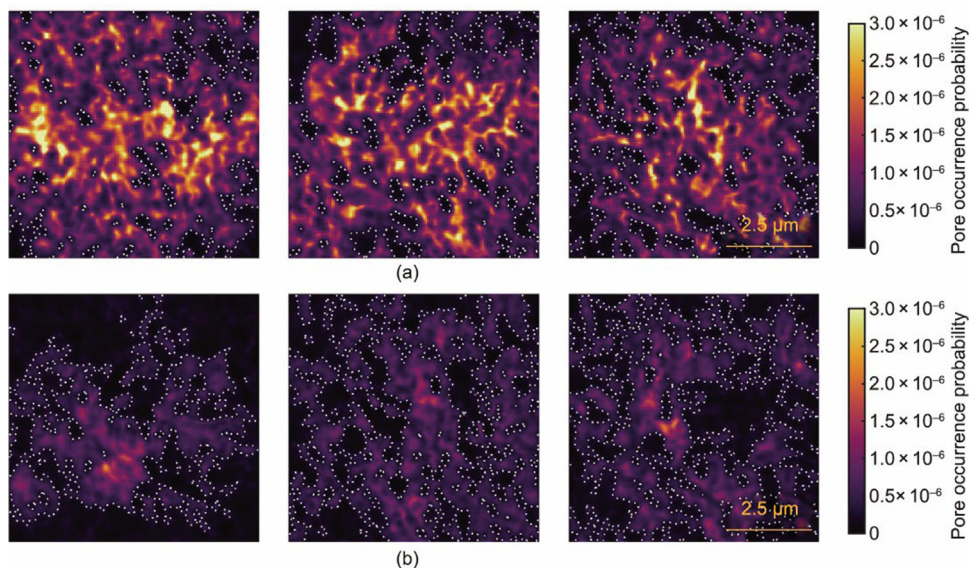


Fig. 13. Representative heatmaps illustrating shared characteristic pattern changes for (a) OPC and (b) LC³. Heatmaps were generated by iteratively analyzing the most similar shared characteristic patterns using deep Taylor decomposition. Reproduced from Ref. [211] with permission.

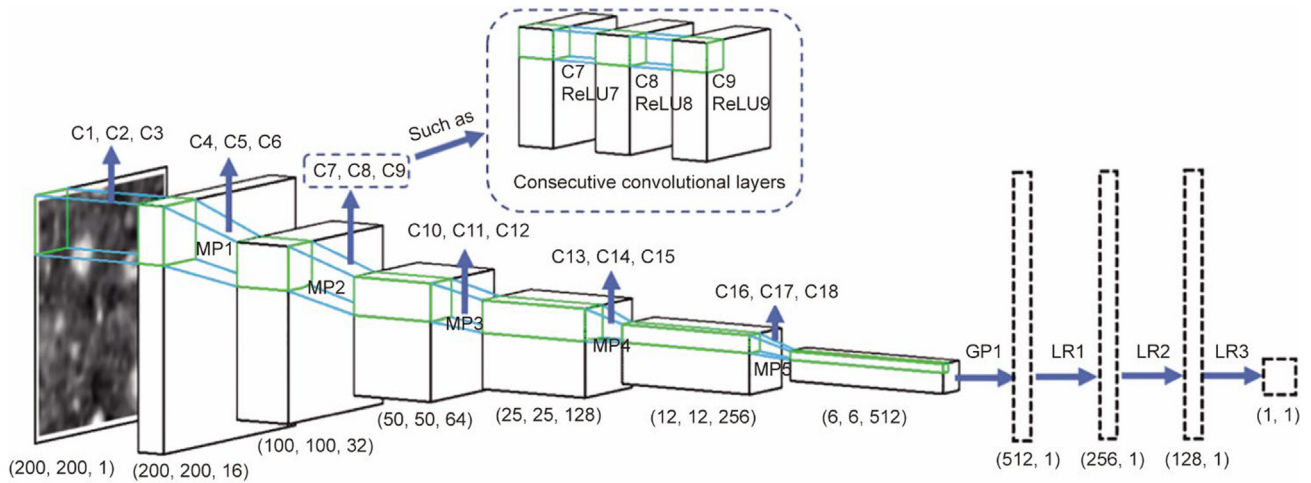


Fig. 14. Architecture of DCNN-1 [213]. C: convolutional layer; MP: max pooling layer; ReLU: rectified linear unit activation; GP: global average pooling layer; LR: linear activation layer.

constructed a dataset comprising 18 920 samples, each containing a microstructural image paired with its corresponding creep modulus. The results showed that the trained DCNN effectively captured the local importance of different microstructural phases, particularly the contributions of unhydrated cement particles and high-density calcium-silicate-hydrate (C-S-H) to creep behavior, achieving R^2 values consistently exceeding 0.96. This study addressed the computational inefficiency of conventional finite element methods and demonstrated the ability of DCNNs to extract and integrate local features from highly heterogeneous microstructures, offering an efficient alternative for performance prediction.

The limited ability of two-dimensional (2D) microstructural images to capture the full spatial structure of cementitious composites often leads to increased prediction errors. In contrast, three-dimensional (3D) microstructural images provide a more accurate representation of hydration states and phase spatial distributions, thereby offering a more comprehensive basis for performance prediction. Guo et al. [214] proposed a method that utilizes 3D microstructural images of cement samples to predict compressive strength. A deep belief network was employed for unsupervised pretraining followed by supervised fine-tuning to learn the nonlinear relationship between microstructural images and compressive strength, resulting in a reduction of MAE by at least 8.49% compared with traditional 2D image-based methods. In addition to improving predictive accuracy, recent developments have focused on enhancing the interpretability and reliability of DL models, which have been successfully applied to reveal intrinsic relationships between image features and material properties. Komninos et al. [215] introduced a transformer-based multihead self-attention mechanism to model spatiotemporal features in crack image sequences. This approach explicitly visualized the model's attention regions across both spatial and temporal domains, thereby improving predictions of remaining useful life for structural health monitoring applications.

In summary, ML techniques in this quadrant automate the extraction of critical microstructural features, uncover patterns and relationships that may be overlooked by conventional methods, and provide valuable insights for advanced material design. Image-based regression models, in particular, enable a more accurate understanding of how microstructural characteristics influence macroscopic performance. When combined with numerical simulation images, these approaches have strong potential to enhance material design applications. By integrating experimental

and simulated data, researchers can more accurately predict and optimize the performance of cementitious composites, thereby supporting their development along intelligent and green pathways.

3.4. Quadrant III: From image to image

In Quadrant III, both the inputs and outputs consist of image data, with an emphasis on methods that transform images into other forms of image-based representations. This quadrant primarily encompasses two approaches: image segmentation and image reconstruction.

3.4.1. Image segmentation

The microstructure of cementitious composites is highly complex, comprising multiple solid phases, such as unhydrated particles and various hydration products, as well as intricate pore networks. Accurate image segmentation is therefore essential for advancing research, as it provides a detailed understanding of material composition and its relationship to performance. Traditional image segmentation methods typically rely on manual operations or predefined algorithms that require substantial expertise and often struggle with the inherent variability and complexity of cementitious materials [216]. These conventional approaches commonly involve thresholding, edge detection, and region-growing techniques, which are limited in their ability to distinguish between similar phases or resolve fine structural details [217]. In contrast, ML techniques, particularly DL models, are capable of learning complex patterns and features from large datasets, enabling more accurate and consistent segmentation outcomes [218–220]. As a result, these methods substantially reduce reliance on expert knowledge while improving the efficiency and scalability of the segmentation process [221,222].

In cementitious composite research, image segmentation applications can be broadly classified into two categories: phase segmentation and crack identification. The key distinction between these applications lies in their objectives: Phase segmentation focuses on identifying multiple microstructural components within an image, whereas crack identification is dedicated exclusively to detecting cracks.

Phase segmentation typically involves distinguishing among hydrates, unhydrated particles, and voids, which is critical for understanding hydration mechanisms and mechanical properties. For instance, Sheiati et al. [223] applied DL techniques to segment

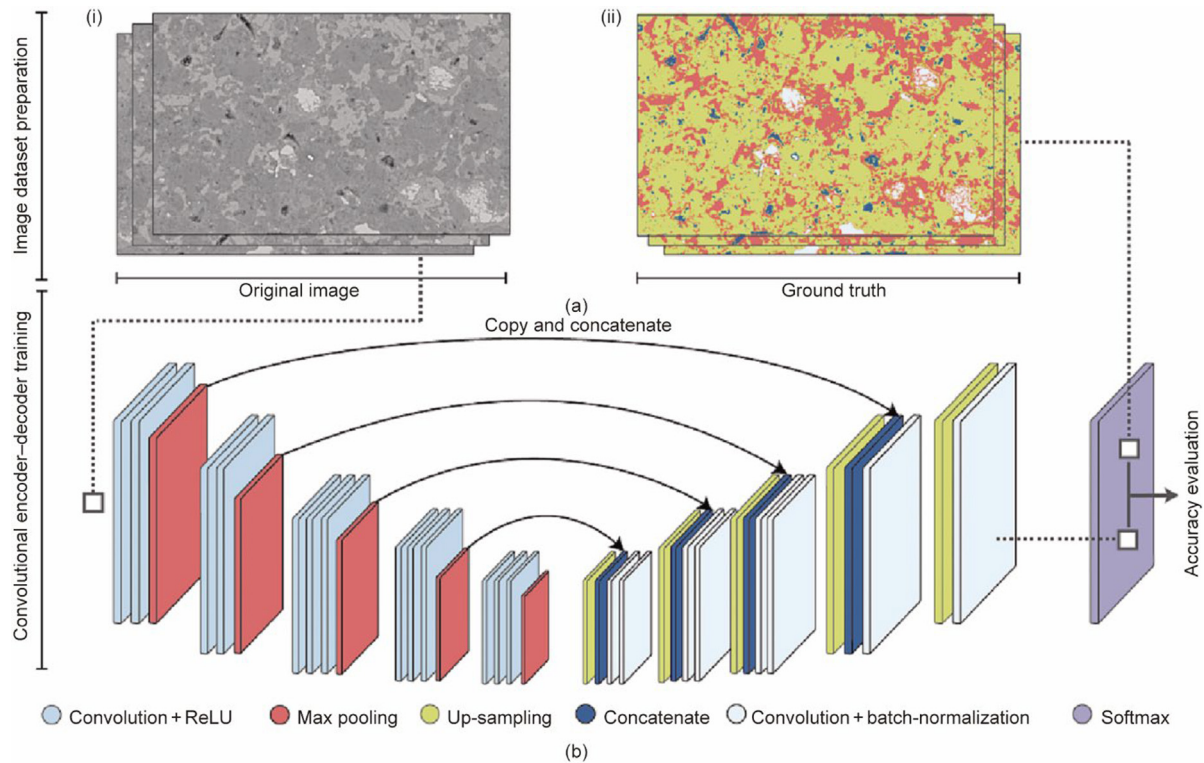


Fig. 15. (a) Image datasets: (a-i) original BSE image; (a-ii) ground truth. (b) VGG-U-Net architecture [223].

pores, portlandite, other hydrates, and unreacted components in cement paste (Fig. 15 [223]). In their study, a high-quality dataset of BSE images was compiled from cement paste samples at different hydration times. Based on prior knowledge, ground truth labels were generated using histogram thresholding, and salt-and-pepper noise was removed using a median filter. A total of 140 images were randomly selected for training and validation, while 82 images were reserved for testing. A visual geometry group (VGG) network was employed for feature extraction and downsampling, and a U-Net architecture was used for upsampling to restore spatial resolution and preserve detailed information. The model was trained using the Adam optimizer with early stopping to mitigate overfitting. The resulting accuracies for the training, validation, and testing datasets were 91%, 90%, and 93%, respectively. Yu et al. [224] performed a comparative evaluation of U-Net, DeepLab V3+, YOLACT, and Mask R-CNN for segmenting pores, aggregates, and paste phases in cementitious composites. Among these models, Mask R-CNN outperformed the others, achieving the highest intersection over union (IoU) and F1-score, demonstrating superior precision and robustness in phase segmentation tasks.

A variety of ML-based strategies have also been developed for crack identification. Hao et al. [225] proposed a high-precision automated semantic segmentation approach using a dual pre-modification strategy to detect cracks in engineered cementitious composites (ECCs) under different conditions. Their network, inspired by U-Net, consists of five convolutional blocks and replaces traditional fully connected layers with deconvolutional layers (Fig. 16 [225]). This design integrates low-level geometric features with high-level semantic information by cross-stacking predictions from multiple layers, leading to a significant improvement in mean IoU (mIoU) to 98.82% and an increase in crack classification accuracy from 92.28% to 99.87%. However, this approach requires large sample sizes and high-quality images. To address limitations related to insufficient data and substandard image quality, Han et al. [226] introduced an enhanced YOLOv5 model

combined with a Wasserstein generative adversarial network (WGAN) for concrete surface defect detection. In addition, the CSNet strategy proposed by Liang et al. [227] demonstrated strong competitiveness in crack identification tasks across multiple public datasets characterized by class imbalance.

3.4.2. Image reconstruction

Image reconstruction in cementitious composite research focuses on enhancing, refining, or fully reconstructing microstructural images to more accurately represent internal material structures. This process is essential for precise characterization and simulation, as it enables the correction of imaging artifacts, enhancement of image resolution, and generation of high-fidelity representations from incomplete or low-quality data. Owing to the multiphase and multiscale nature of cementitious composites, traditional reconstruction methods often struggle to accurately reproduce original microstructural characteristics [228–230]. In contrast, advanced ML techniques, particularly DL, leverage large volumes of image data to effectively extract features and generate virtual images that preserve statistical properties similar to those of the originals, thereby enabling more reliable and detailed reconstructions [231]. With the rapid development of ML models, image reconstruction has been successfully applied to cementitious composites for multiple purposes, including reconstruction across different dimensions, prediction of dynamic microstructural evolution, and data augmentation of structural images.

Microstructural reconstruction plays a critical role in accurately representing the internal structure of cementitious composites and is fundamental to precise material characterization. Hong et al. [232] employed a CycleGAN framework combined with a CNN-based autoencoder connection strategy to reconstruct cement paste microstructures (Fig. 17(a) [232]). Their approach followed a two-step process, involving the generation of 2D microstructures from random noise and subsequent 3D reconstruction. The reconstructed structures were validated using metrics such as phase

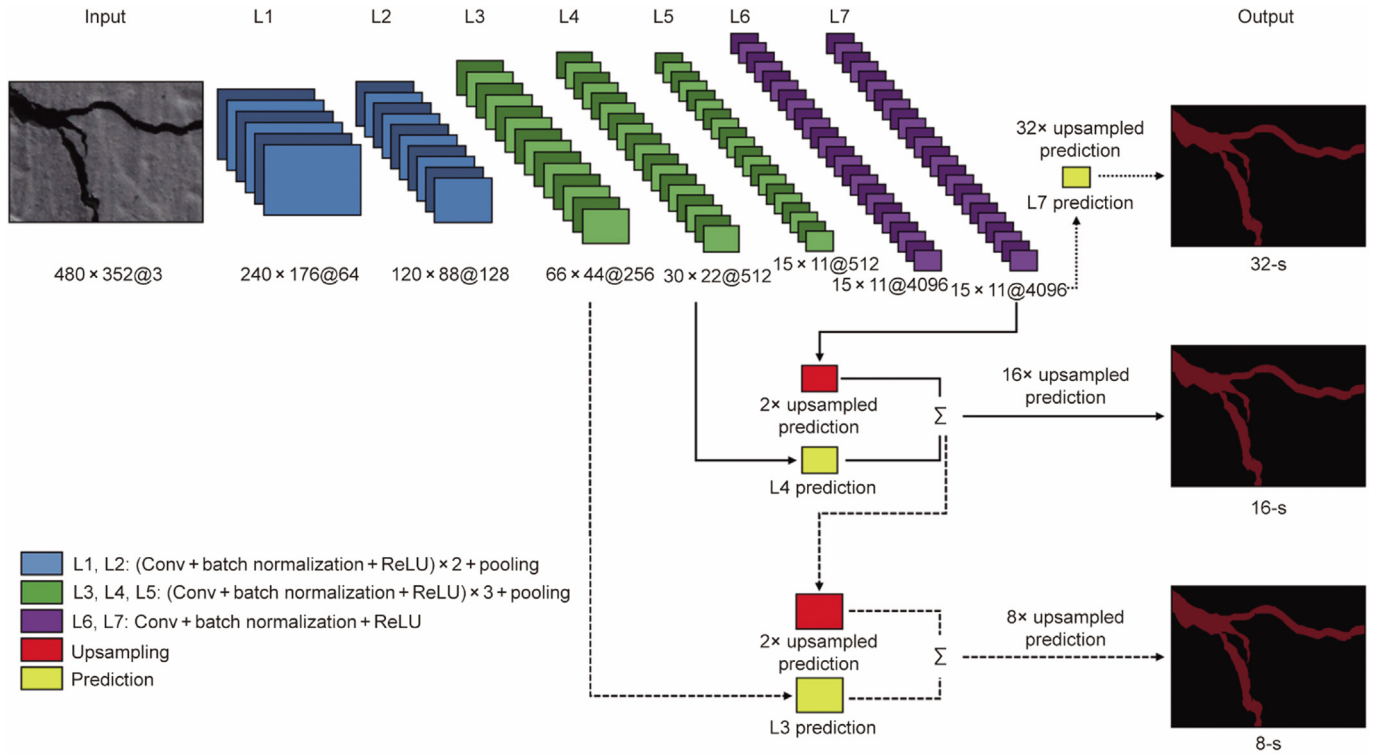


Fig. 16. Framework of the proposed semantic segmentation network for microcracks in ECC. 32-s: the 32-fold upsampled prediction of L7 prediction; 16-s: the 16-fold upsampled prediction of the cross-superimposition of two-fold expansion of L7 prediction and L4 prediction; 8-s: the 8-fold upsampled prediction of the cross-superimposition of two-fold expansion of the cross-superimposition above and L3 prediction; Conv: convolution. Reproduced from Ref. [225] with permission.

volume fraction and continuity. Building on this work, Zhao et al. [233] proposed a low-cost framework, CEM3DMG, to generate realistic 3D microstructures from a single 2D BSE image. By integrating multiscale learning with blockwise synthesis, this method enabled large-scale 3D reconstruction with high structural fidelity. However, such approaches may omit information at different structural orders, potentially limiting comprehensive insight into material microstructures. To address this issue, Lin et al. [234] introduced a novel framework that successfully reconstructed the microstructure of a central region based on surrounding neighborhood information while preserving both low- and high-order features. Quantitative analysis revealed that spatial correlations within the microstructure significantly influenced reconstruction accuracy.

Furthermore, image reconstruction techniques have been extended to support the investigations of microstructural evolution. Yao et al. [235] proposed a DL-based image translation approach to predict the *in situ* evolution of microporous structures during cement hydration (Fig. 17(b) [235]). Using a CycleGAN framework, this method achieved high-resolution bidirectional translation between images representing low and high degrees of hydration without requiring paired datasets or complex physical formulations. Validation against experimental data demonstrated high predictive accuracy ($R^2 \geq 0.91$), particularly for porosity, pore size distribution, and shape characteristics. This study collectively highlights the effectiveness of DL, especially GAN-based models, in realistically and efficiently predicting complex microstructural evolution in cementitious materials.

The manual acquisition of image datasets, particularly those depicting damage in cementitious composites, is often labor-intensive and time-consuming, posing significant challenges for the development of efficient data-driven models. To address this limitation, Huang et al. [236] proposed an unsupervised image-

to-image translation method using CycleGAN to generate underwater dam crack images from above-water images (Fig. 17(c) [236]). Building on this concept, Li and Zhao [237] employed a conditional GAN to generate high-resolution images of concrete damage, focusing on cracks and spalling. The synthesized images exhibited a high degree of realism, indicating strong potential for substantially supporting DL-based damage detection networks through data augmentation.

In summary, Quadrant III emphasizes both the extraction and representation of image features, underscoring the transformative potential of image-to-image ML techniques in cementitious composite research. Existing studies demonstrate that DL excels in accurately segmenting complex microstructures and reconstructing high-fidelity images, even when source data are incomplete or of low quality. These approaches not only advance the understanding of microstructural characteristics but also enhance visualization and modeling capabilities beyond those achievable with traditional methods. The application of DL in this quadrant provides robust and scalable solutions for analyzing and simulating internal material structures, thereby promoting the intelligent development of cementitious composites through improved accuracy, efficiency, and reduced reliance on physical experimentation.

3.5. Quadrant IV: from non-image to image

In Quadrant IV, the input is non-image data and the output is image data. Applications in this quadrant focus on transforming non-image data into image data, with a typical application being image generation, a process that opens new possibilities for understanding and visualizing the properties of cementitious composites.

Traditional methods for obtaining images, such as direct imaging techniques (e.g., microscopy and X-ray tomography), involve

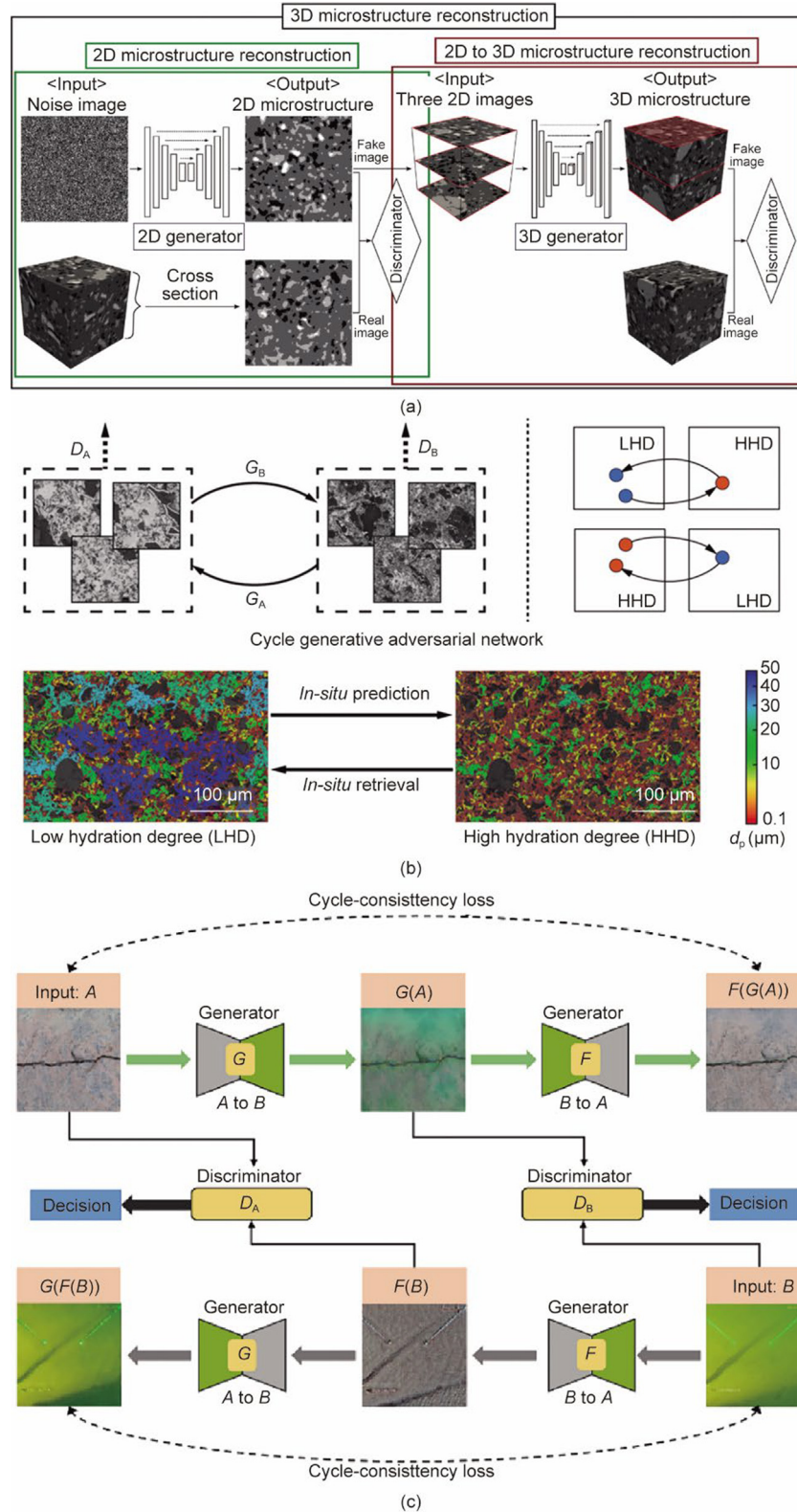


Fig. 17. (a) Modified CycleGAN training network for reconstructing 3D cement paste microstructures. Reproduced from Ref. [232] with permission. (b) 3D microstructural image generation framework (CEM3DMG). Reproduced from Ref. [235] with permission. (c) DL-based image translation of microporous structure evolution during cement hydration. Reproduced from Ref. [236] with permission. D_A and D_B : associated adversarial discriminators; G_B : mapping microstructures of low hydration degree (LHD) to high hydration degree (HHD); G_A : mapping HHD to LHD; d_p : equivalent pore diameter; G and F : two generators. Images of A and B domains are used as inputs to generators G and F to generate images $G(A)$ and $F(B)$; $F(G(A))$ and $G(F(B))$ are images generated by inputting $G(A)$ and $F(B)$ into F and G , respectively.

capturing visual data directly from physical samples, which are often labor-intensive and time-consuming [238]. In comparison, ML-based image generation leverages structured data inputs

(e.g., numerical parameters or categorical labels) to create visual representations, which significantly improves efficiency at minimal cost.

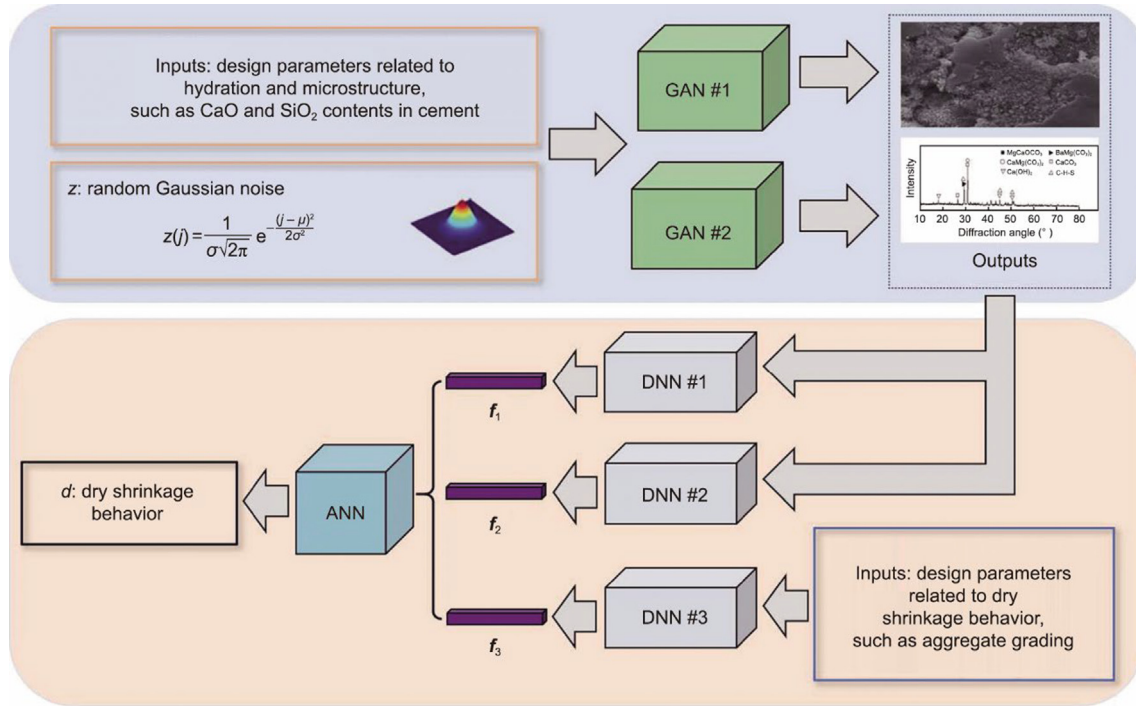


Fig. 18. Illustration of the DL framework for image generation and subsequent analysis. C-H-S: calcium silicate hydrate; DNN: deep neural networks; j : independent random variable representing the value domain in the Gaussian distribution equation; σ : standard deviation of the distribution; μ : mean of the distribution; f_1, f_2, f_3 : feature vectors output by DNN #1, DNN #2, and DNN #3, respectively. Reproduced from Ref. [239] with permission.

Tong et al. [239] proposed a new method based on a GAN to characterize the hydration process and microstructural features of cement-emulsified asphalt composites (Fig. 18 [239]). The generator received the design parameters and random Gaussian noise to generate synthetic X-ray diffraction (XRD) spectra and scanning electron microscope (SEM) images, whereas the discriminator was trained to optimize the quality of the generator output by distinguishing between the real and synthetic data. An average Euclidean distance of 0.001 was achieved between the synthetic and real XRD spectra, and the synthetic SEM images could clearly characterize microstructures such as cement hydration products, asphalt, and pores. This method of generating characterization images from non-image data offers a new avenue for data augmentation, potentially supporting other studies with additional data resources.

The generation of material microstructures based on text prompts has opened new avenues for intelligent designs. Zheng et al. [240] proposed a generative DL framework called Txt2Microstruct-Net that enables the direct generation of 3D material microstructures from natural language descriptions (Fig. 19 [240]). This method established a unified multimodal embedding space. Thus, it overcomes the limitations of traditional inverse design approaches that are based solely on structural parameters or performance indicators. Furthermore, the generated microstructures can be manufactured, providing direct support for fabrication technologies such as 3D printing. Although this approach has not yet been applied to cementitious composites, its underlying concept of “language-driven morphology generation with direct manufacturing linkage” offers novel insights for the intelligent design and digital fabrication of cementitious composite materials.

In summary, research within Quadrant IV is still in its early stages of development. While current advancements in other fields [241–243] demonstrate the potential of ML-based techniques to

efficiently create visual representations from abstract data, much work remains to be done on cementitious composites. Future efforts should focus on refining these methods to enhance their accuracy and applicability, expand the range of non-image inputs that can be effectively transformed, and integrate these generated images into broader workflows such as material design and performance prediction. Advancing this area of research will not only provide new insights into the properties and behaviors of cementitious composites but also foster innovative approaches to material design and optimization.

3.6. Contributions of the four quadrants in green and intelligent design

Material design involves a deep understanding and strategic manipulation of the fundamental attributes of a material to achieve the desired characteristics, which encapsulate four interconnected pillars: components, structure, process, and performance. In the era of data-driven science, the intricate relationship among these four cores lays a solid foundation for the dynamic and precise design of materials, as illustrated in Fig. 20. Specifically, the components, structures, and processes serve as the three cornerstones, jointly forming a logical starting point for performance prediction. Multi-objective performance, as the ultimate goal, constitutes a feedback mechanism that ensures continuous improvement and iterative upgrades in material design. Furthermore, internal relationships exist among the components, structures, and processes. Particular pairs of components and processes can induce unique structures. In contrast, an in-depth analysis and understanding of the structural features can guide the selection of components and optimization of preparation processes. This virtuous and interactive cycle propels the progress and development of materials.

Accordingly, the entire cycle can be categorized into four quadrants, where ML is used to provide efficient and innovative alterna-

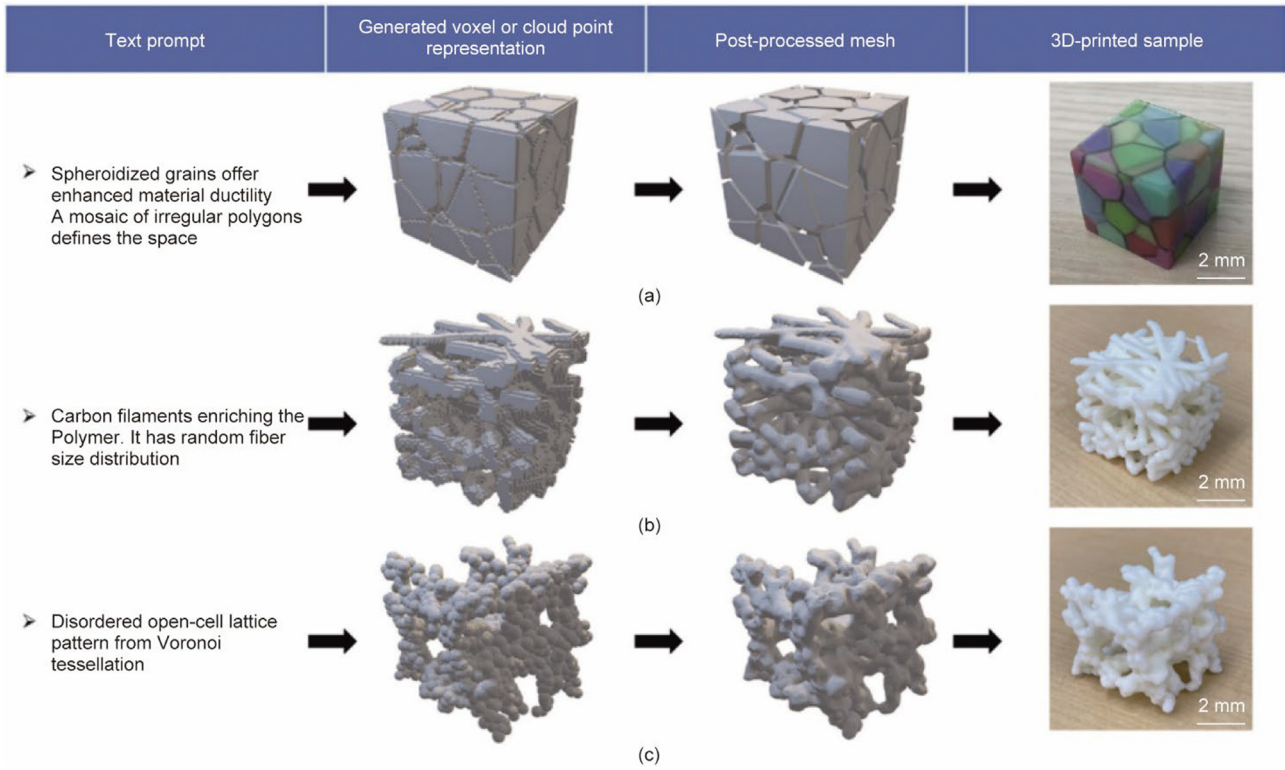


Fig. 19. Microstructures generated by the Txt2Microstruct-Net model using text prompts [240]. (a) Generating and fabricating a voxel-based microstructure with multimaterials. (b) Generating and fabricating a voxel-based microstructure with a single material. (c) Generating and fabricating a point-cloud-based microstructure.

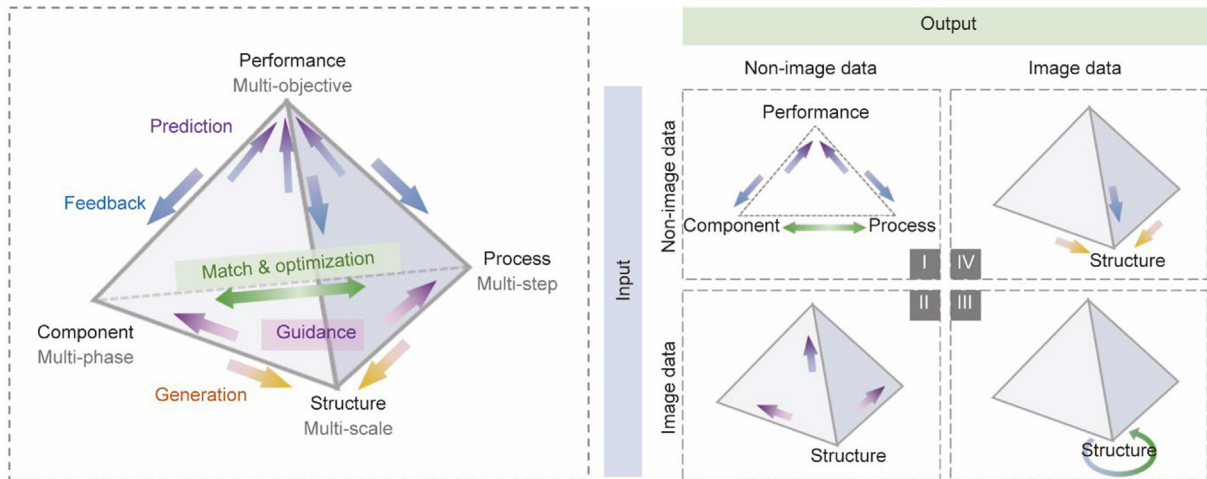


Fig. 20. Intricate relationship between component, structure, process, and performance.

tives that complement conventional approaches. The synergy of the four quadrants contributes significantly to the green and smart design of cementitious composites, as summarized below.

(1) The non-image data almost comprehensively encompass aspects of the components, processes, and performance of cementitious composites, as discussed in Section 3.1. By harnessing the power of data analytics on vast datasets, ML can discern patterns and correlations that would otherwise remain elusive when compared with conventional methods. This enables the optimization of the mix design for cementitious composites that minimizes environmental impact while maximizing performance outcomes. Similarly, fine-tuning process parameters leads to more efficient

and ecofriendly manufacturing practices. The integration of ML in Quadrant I aligns perfectly with the principles of green design, but the potential of this approach to drive innovation and sustainability in the engineering sector remains unresolved.

(2) For any material, the structure is the fundamental determinant of the performance, acting as a bridge between the components, processes, and performance. In contrast to the simplistic correlations between component and process performances, the image data of structural information offer a more authentic and evidence-based perspective. In Quadrant II, deconstruction analysis of the characteristic features detected by ML enables researchers to visualize the structure–performance correlation, revealing

an unprecedented level of detail. These structural insights can be leveraged to guide the components and process parameters instead of relying solely on trial-and-error or empirical knowledge.

(3) The structure of cementitious composites is well recognized as exceedingly complex and is characterized by multiphase and multiscale properties with an intricate spatiotemporal evolution mechanism. The exploration in Quadrant III focused on the structural information of cementitious composites, aiming to enhance the efficiency and precision of image processing. This endeavor strives to automate and intellectualize the extraction and transformation of image information, fostering a more streamlined and intelligent approach to understanding composite structures.

(4) The majority of image data utilized in other quadrants were sourced from real-world photographs captured at varying scales, a process heavily reliant on photography equipment and often time-consuming and labor-intensive. By contrast, exploration within Quadrant IV endeavors to overcome this limitation by leveraging the understanding of how components, processes, and performance factors influence structures. The objective was to generate structural images directly based on these descriptive parameters, thereby bypassing the need for physical photography. Although this approach is still in its nascent stages, development in this direction would contribute to a crucial breakthrough in the green design of cementitious composites, ushering in an era of experiment-free design. By eliminating reliance on resource-intensive laboratory experiments and promoting data-driven simulations, this paradigm shift holds immense significance for realizing green and intelligent design methodologies.

3.7. Case studies in industrial applications

Despite the significant progress, ML applications in cementitious composites remain largely confined to simulations. Transitioning ML from theory to practice necessitates the integration of diverse technologies and addresses real-world engineering complexities. Currently, ML is pivotal in two key stages: material design and structural health monitoring. This section reviews the case studies that highlight the practical industrial value of ML.

In materials design, ML has significantly optimized the mix design and production processes of cementitious composites. Ge et al. [244] employed conditional variational autoencoders to formulate low-carbon concrete and balance environmental impact data with component properties to meet strength requirements. This low-carbon concrete formulation has been applied in the construction of Meta's data center, reducing carbon emissions by up to 50%. Converge (London, UK) developed a strength prediction engine that integrated local weather, curing histories, and real-time data [245]. This tool forecasts the strength development days in advance and provides crucial decision support. Its implementation in the London City Airport (UK) expansion project substantially improved productivity, while reducing time and costs.

In the field of structural health monitoring, ML has significantly advanced the intelligent monitoring of infrastructure. Mirzazade et al. [246] combined computer vision and DL techniques to develop a drone system for automated bridge damage detection. In the health monitoring of the Pahtajokk Bridge (Sweden), the system not only generated an accurate 3D point-cloud model but also successfully quantified bridge damage, providing valuable maintenance data support. Nguyen et al. [247] proposed an automatic system that uses transfer learning and CNNs to detect concrete cracks in real time. The model, trained with 3000 images and deployed on NVIDIA Jetson Nano devices (NVIDIA Corporation, USA), demonstrated robust field performance in complex environments.

In summary, these cases demonstrate the successful translation of research and active industrial exploration, confirming the engi-

neering value of ML in cementitious composites. Future applications are expected to span the entire lifecycle of these materials, enhance their performance, optimize design processes, and advance sustainable development.

4. Challenges

Despite the significant advantages of ML for enabling a green and intelligent future for cementitious composites, several challenges remain. Four major issues are outlined below.

4.1. Data availability

The quality, quantity, and representativeness of data are critical to the success of ML applications [248–250]. However, research on cementitious composites is hindered by nonuniform data formats, dimensions, and levels of precision arising from disparate experimental practices and data-recording methods across institutions. In addition, the limited availability of data for innovative research restricts the development of high-dimensional datasets. These inconsistencies and data shortages significantly impede the efficient application of ML. Addressing these issues requires collaborative efforts to establish standardized data protocols and shared data platforms that can overcome data fragmentation and privacy barriers. Such platforms should comprehensively incorporate diverse data types, which can be grouped into three primary categories: data sources, material data, and ecological-economic attributes. Data sources should include information about data origin and recording conditions to ensure traceability and authenticity. Material data should be organized around four core dimensions: components, structure, process, and performance. Ecological-economic attributes should capture both economic indicators and environmental impact metrics, such as material cost, embodied carbon, and lifecycle energy consumption.

4.2. Algorithm selection

Selecting an appropriate ML algorithm is challenging due to the diversity of research objectives, data characteristics, and performance requirements in cementitious composite studies [251]. Algorithm choice is often based on researcher experience, which may lead to suboptimal outcomes, particularly when data quality is inconsistent and certain algorithms exhibit unstable performance. This challenge highlights the need for systematic pre-analysis of data characteristics and careful evaluation of algorithmic strengths and limitations, including those of emerging methods. In addition, the adoption of hybrid or multi-model fusion strategies should be considered to improve adaptability and predictive performance across different application scenarios.

4.3. Model interpretability

The ML models, especially DL models, are frequently criticized as “black boxes” due to their limited transparency in decision-making processes [252–254]. For critical engineering materials such as cementitious composites, this lack of interpretability undermines trust and hinders broader adoption. Without sufficient interpretability, diagnosing model errors, identifying biases, and understanding the mechanisms governing material systems become difficult. Therefore, enhancing model interpretability is essential. Existing interpretation techniques for non-image data should be further refined to provide clearer physical meaning, while interpretability methods for image-based models should be improved to strengthen model credibility and facilitate practical implementation.

4.4. Model reliability

Model reliability is essential to ensure that ML predictions are accurate, robust, and generalizable to unseen data. Research on cementitious composites encompasses a wide range of experimental conditions and data distributions, depending on the specific application context [255]. Consequently, an algorithm that performs well in one scenario may fail in another due to overfitting or limited generalizability. To mitigate these risks, rigorous hyperparameter optimization prior to training is required. Moreover, ML models must undergo continuous monitoring, evaluation, and updating through long-term data supplementation to maintain their reliability and adaptability in practical engineering applications.

5. Future innovations

In the future, the comprehensive integration of AI—driven by advances in high-throughput databases, computational capabilities, and algorithmic development—is expected to significantly influence the development of cementitious composites throughout their entire lifecycle (Fig. 21). This integrated framework enables intelligent and green transformations at every stage.

(1) **Raw materials phase.** At this stage, raw material selection can be optimized using large-scale databases that incorporate carbon footprint information. Ensemble models capable of predicting engineering performance can be used to balance performance requirements with carbon emissions, enabling the identification of optimal low-carbon materials and promoting sustainability from the earliest stage of the material lifecycle.

(2) **New composite design phase.** This phase benefits from a fully digitalized design workflow that integrates predictive and generative models. For example, by combining multiscale modeling (MD) with GNNs, structural images can be generated directly from detailed compositional information and subsequently used for performance prediction. Iterative improvements achieved

through this digital loop support comprehensive data-driven research and innovation in composite material design.

(3) **Preparation and production phase.** The ML models that link preparation processes to performance metrics can guide the optimization of manufacturing procedures. By quantifying the influence of individual processing steps, high-impact and scalable processes can be strategically refined. For instance, ML-assisted building information modeling (BIM) and 3D printing can be used to optimize printing paths and process parameters, thereby improving production efficiency and construction quality through intelligent methods.

(4) **Performance evaluation phase.** In this phase, nondestructive testing techniques enable efficient assessment of material performance. Models based on CNNs and transformer architectures with feature fusion can be trained using image data from multiple sources. By rapidly correlating surface imagery with performance indicators, evaluation becomes faster and more reliable, reducing the need for destructive testing and improving overall assessment efficiency.

(5) **Service monitoring phase.** Long-term performance monitoring can be enhanced through intelligent sensing devices that detect signals related to material health, such as cracks or deformations. When combined with models that link signal features to performance characteristics, these systems enable continuous and automated structural health monitoring. The integration of long short-term memory (LSTM), recurrent neural network (RNN), and temporal convolutional networks (TCNs) supports real-time data analysis from Internet of things (IoT) sensors, facilitating proactive maintenance and extending service life.

(6) **Waste disposal phase.** In waste management, advanced models capable of recognizing and evaluating particle morphology can be developed to assess recyclability. By integrating these insights into comprehensive material property databases, intelligent systems can support waste classification and processing. For example, deep reinforcement learning (DRL) and deep Bayesian optimization (DBO) can be applied to automate waste sorting and recycling, ensuring high-quality recycled materials and maximizing reuse value. This approach promotes efficient waste management, supports a circular economy, and minimizes environmental impact.

Overall, this integrated AI-driven framework has the potential to enhance the performance and sustainability of cementitious composites across their entire lifecycle, while driving a comprehensive transition toward intelligent and green practices in the construction materials industry.

6. Conclusions

This review systematically examines the potential of ML to advance cementitious composites toward intelligent and green lifecycles. The main contributions of this review are summarized as follows:

(1) A foundational knowledge bridge is established by clarifying core ML concepts, workflows, and terminology in an accessible manner for researchers without specialized AI expertise.

(2) A novel data-centric taxonomy is proposed, organizing existing research into a four-quadrant classification framework based on input and output data types. This framework is supported by a quantitative analysis of 120 representative studies, and complemented by illustrative case examples.

(3) The contributions of the four quadrants in green and intelligent design are examined by mapping ML advancements onto the composition–structure–process–property paradigm.

(4) A practical challenge–solution matrix is presented, which identifies key barriers to real-world deployment and proposes targeted solutions at the technical, methodological, and practical levels to facilitate broader implementation.

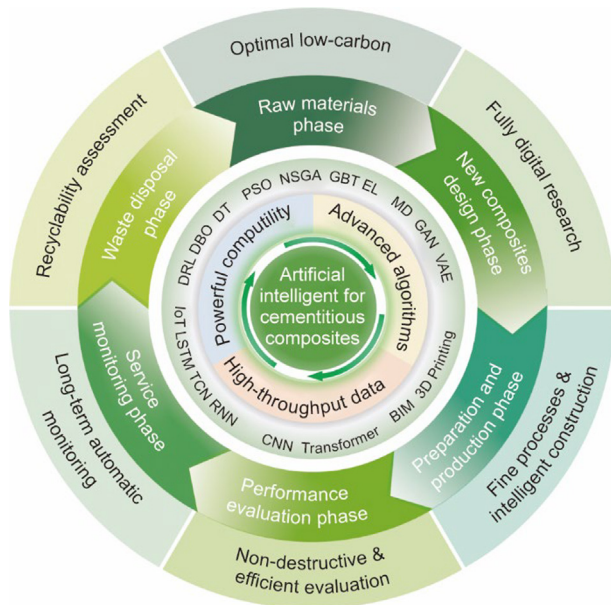


Fig. 21. Prospects for an intelligent and green lifecycle of cementitious composites. DRL: deep reinforcement learning; DBO: deep bayesian optimization; PSO: particle swarm optimization; NSGA: non-dominated sorting genetic algorithm; GBT: gradient boosting tree; MD: multiscale modeling; BIM: building information modeling; TCN: temporal convolutional network; RNN: recurrent neural network; LSTM: long short-term memory; IoT: Internet of things.

Collectively, this review helps reduce the knowledge gap between AI and materials science and expands the application scope of ML in cementitious composite research.

From the perspective of the four quadrants, the main contributions of ML for cementitious composites can be summarized as follows:

(1) Non-image to non-image: ML enables material performance prediction and inverse design, streamlining material development, reducing resource consumption, and supporting the creation of sustainable, high-performance cementitious composites.

(2) Image to non-image: ML excels in automated feature extraction from microstructural images, facilitating accurate material classification and establishing correlations between microstructure and properties. These capabilities not only guide production processes but also enable innovative nondestructive testing methods.

(3) Image to image: DL demonstrates accurate performance in the segmentation of complex microstructures and image reconstruction, enhancing the understanding of microstructural features and supporting structural damage detection.

(4) Non-image to image: Although still at an early stage, this quadrant shows promise in generating visual representations from nonvisual data, enabling data augmentation and reducing reliance on experimental testing. Across all four quadrants, ML-driven innovations advance cementitious composites by fundamentally integrating the component–structure–process–performance relationships. This integration, manifested as intelligent material selection, characterization, prediction, and inverse design, collectively promotes the paradigm shift toward intelligent and green lifecycles that is central to this review.

The comprehensive integration of AI with other emerging technologies is expected to further enhance the performance and sustainability of cementitious composites across their entire lifecycles. Looking forward, the proposed frameworks, including the four-quadrant classification, can serve as foundational references for developing standardized protocols in data collection, model validation, and performance benchmarking. In addition, insights into ML-driven carbon reduction may inform future policymaking and regulatory guidelines for sustainable construction, thereby supporting the intelligent and green transformation of the construction industry.

CRediT authorship contribution statement

Jinyang Jiang: Writing – original draft, Supervision, Funding acquisition, Conceptualization. **Junlin Lin:** Writing – review & editing, Visualization, Software, Methodology, Investigation, Conceptualization. **Lin Jin:** Writing – review & editing, Visualization, Investigation, Data curation. **Fengjuan Wang:** Writing – review & editing. **Zhiyong Liu:** Writing – review & editing. **Yingze Li:** Writing – review & editing. **Zeyu Lu:** Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors are grateful for the financial support from the Major Research Plan of the National Natural Science Foundation of China (52293431) and the Special Fund for Carbon Peak and Carbon Neutrality Science and Technology Innovation of Jiangsu Province (BE2023851).

Appendix A. Supplementary material

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.eng.2026.02.012>.

References

- [1] Schmidt W, Alexander M, John V. Education for sustainable use of cement based materials. *Cem Concr Res* 2018;114:103–14.
- [2] Shi C, Qu B, Provis JL. Recent progress in low-carbon binders. *Cem Concr Res* 2019;122:227–50.
- [3] Felipe Arbeláez Pérez O, Senior Arrieta V, Hernán Gómez Ospina J, Herrera Herrera S, Ferney Rodríguez Rojas C, María Santis Navarro A. Carbon dioxide emissions from traditional and modified concrete. A review. *Environ Dev* 2024;52:101036.
- [4] Cheng D, Reiner DM, Yang F, Cui C, Meng J, Shan Y, et al. Projecting future carbon emissions from cement production in developing countries. *Nat Commun* 2023;14(1):8213.
- [5] Monteiro PJM, Miller SA, Horvath A. Towards sustainable concrete. *Nat Mater* 2017;16(7):698–9.
- [6] Hanifa M, Agarwal R, Sharma U, Thapliyal PC, Singh LP. A review on CO₂ capture and sequestration in the construction industry: emerging approaches and commercialised technologies. *J CO₂ Util* 2023;67:102292.
- [7] Leslie M. Construction industry innovation takes aim at reducing carbon emissions. *Engineering* 2022;19:7–10.
- [8] Provis JL, Bernal SA, Zhang Z. The decarbonization of construction—how can alkali-activated materials contribute? *Engineering* 2024;37:18–21.
- [9] Barbhuiya S, Kanavaris F, Das BB, Idrees M. Decarbonising cement and concrete production: strategies, challenges and pathways for sustainable development. *J Build Eng* 2024;86:108861.
- [10] Coffetti D, Crotti E, Gazzaniga G, Carrara M, Pastore T, Coppola L. Pathways towards sustainable concrete. *Cem Concr Res* 2022;154:106718.
- [11] He P, Drissi S, Hu X, Liu J, Shi C. Properties of CO₂-cured cement incorporating fly ash and slag subjected to further water curing. *Cem Concr Compos* 2024;152:105633.
- [12] Beskopylny AN, Stel'Makh SA, Shcherban' EM, Saidumov MS, Abumuslimov A, Mezhdidov D, et al. Eco-friendly foam concrete with improved physical and mechanical properties, modified with fly ash and reinforced with coconut fibers. *Constr Mater. Prod* 2025;8(1):1.
- [13] Bespolitov DV, Kononova NA, Dabizha ON, Pankov PP, Rush EA. Influence of the mechanical activation of fly ash on strength of ground concrete based on waste production. *Ecol Ind Russ* 2021;25(11):36–41.
- [14] Kharun M, Klyuev S, Koroteev D, Chiadighikaobi PC, Fediuk R, Olisov A, et al. Heat treatment of basalt fiber reinforced expanded clay concrete with increased strength for cast-in-situ construction. *Fibers* 2020;8(11):67.
- [15] Scrivener K, Martirena F, Bishnoi S, Maity S. Calcined clay limestone cements (LC³). *Cem Concr Res* 2018;114:49–56.
- [16] Tanhadoust A, Emadi SAA, Nasrollahpour S, Dabbaghi F, Nehdi ML. Optimal design of sustainable recycled rubber-filled concrete using life cycle assessment and multi-objective optimization. *Constr Build Mater* 2023;402:132878.
- [17] Yu Y, Fang G, Kurda R, Sabuj AR, Zhao XY. An agile, intelligent and scalable framework for mix design optimization of green concrete incorporating recycled aggregates from precast rejects. *Case Stud Constr Mater* 2024;20:e03156.
- [18] Pfeiffer OP, Gong K, Severson KA, Chen J, Gregory JR, Ghosh S, et al. Bayesian design of concrete with amortized Gaussian processes and multi-objective optimization. *Cem Concr Res* 2024;177:107406.
- [19] Van Damme H. Concrete material science: past, present, and future innovations. *Cem Concr Res* 2018;112:5–24.
- [20] Raabe D, Mianroodi JR, Neugebauer J. Accelerating the design of compositionally complex materials via physics-informed artificial intelligence. *Nat Comput Sci* 2023;3(3):198–209.
- [21] Malik S, Muhammad K, Waheed Y. Artificial intelligence and industrial applications—a revolution in modern industries. *Ain Shams Eng J* 2024;15(9):102886.
- [22] Gallegos M, Vassilev-Galindo V, Poltavsky I, Martín Pendas Á, Tkatchenko A. Explainable chemical artificial intelligence from accurate machine learning of real-space chemical descriptors. *Nat Commun* 2024;15(1):4345.
- [23] Jordan MI, Mitchell TM. Machine learning: trends, perspectives, and prospects. *Science* 2015;349(6245):255–60.
- [24] Vityaev EE, Goncharov SS, Sviridenko DI. Task-driven approach to artificial intelligence. *Cogn Syst Res* 2023;81:50–6.
- [25] Wu H, Zhang J, Bao Z, Wang G, Wang W, Yang Y, et al. Runoff modeling in ungauged catchments using machine learning algorithm-based model parameters regionalization methodology. *Engineering* 2023;28:93–104.
- [26] Zheng X, Li P, Wu X. Data stream classification based on extreme learning machine: a review. *Big Data Res* 2022;30:100356.
- [27] Chandwani V, Agrawal V, Nagar R. Modeling slump of ready mix concrete using genetic algorithms assisted training of artificial neural networks. *Expert Syst Appl* 2015;42(2):885–93.
- [28] Jiang JY, Liu Y, Lin JL, Yang T, Wang F. Machine learning assisted design of low-carbon aluminosilicate cementitious composites with diverse raw

- materials and target mechanical strength. *Case Stud Constr Mater* 2025;22:e04664.
- [29] Li Y, Liu Y, Lin H, Jin C. Study of flexural strength of concrete containing mineral admixtures based on machine learning. *Sci Rep* 2023;13(1):18061.
 - [30] Dai J, Yang X, He J, Wang Q, Zhang Z. Machine learning prediction of electric flux in concrete and mix proportion optimization design. *Mater Today Commun* 2024;38:107778.
 - [31] Li WH, Xiong CC, Zhou Y, Chen W, Zheng Y, Lin W, et al. Insights on the mechanical properties and failure mechanisms of calcium silicate hydrates based on deep-learning potential molecular dynamics. *Cem Concr Res* 2024;186:107690.
 - [32] Wu ZY, Pan H, Huang P, Tang J, She W. Biomimetic mechanical robust cement-resin composites with machine learning-assisted gradient hierarchical structures. *Adv Mater* 2024;36(35):2405183.
 - [33] Golafshani EM, Behnood A, Kim T, Ngo T, Kashani A. A framework for low-carbon mix design of recycled aggregate concrete with supplementary cementitious materials using machine learning and optimization algorithms. *Structures* 2024;61:106143.
 - [34] Butakov EB, Abdurakipov SS, Neznamov VY, Alekseenko SV. On Increasing the efficiency of a cement clinker kiln using machine learning. *J Eng Thermophys* 2024;33(4):675–82.
 - [35] Zheng W, Liu H, Zhou XY, Xue XJ, Li H, Liu JX. A soft sensor model based on an improved semi-supervised stacked autoencoder for just-in-time updating of cement clinker production process data f-CaO. *Meas Sci Technol* 2024;35(5):056203.
 - [36] Li H, Lyngdoh G, Krishnan NMA, Das S. Machine learning guided design of microencapsulated phase change materials-incorporated concretes for enhanced freeze-thaw durability. *Cem Concr Compos* 2023;140:105090.
 - [37] Elshaarawy MK, Alsaadawi MM, Hamed AK. Machine learning and interactive GUI for concrete compressive strength prediction. *Sci Rep* 2024;14(1):16694.
 - [38] Jin L, Duan J, Jin Y, Xue P, Zhou P. Prediction of HPC compressive strength based on machine learning. *Sci Rep* 2024;14(1):16776.
 - [39] Das AK, Suthar D, Leung CKY. Machine learning based crack mode classification from unlabeled acoustic emission waveform features. *Cem Concr Res* 2019;121:42–57.
 - [40] Aravind N, Nagajothi S, Elavenil S. Machine learning model for predicting the crack detection and pattern recognition of geopolymer concrete beams. *Constr Build Mater* 2021;297:123785.
 - [41] Althoei F, Amin MN, Khan K, Usman MM, Khan MA, Javed MF, et al. Machine learning based computational approach for crack width detection of self-healing concrete. *Case Stud Constr Mater* 2022;17:e01610.
 - [42] Albert C, Isgor OB, Angst U. Exploring machine learning to predict the pore solution composition of hardened cementitious systems. *Cem Concr Res* 2022;162:107001.
 - [43] Vedrtam A, Kumar S, Barluenga G, Chaturvedi S. Early crack detection using modified spectral clustering method assisted with FE analysis for distress anticipation in cement-based composites. *Sci Rep* 2021;11(1):19685.
 - [44] Jia H, Qiao G, Han P. Machine learning algorithms in the environmental corrosion evaluation of reinforced concrete structures-a review. *Cem Concr Compos* 2022;133:104725.
 - [45] Taffese WZ, Wally GB, Magalhães FC, Espinosa-Leal L. Concrete aging factor prediction using machine learning. *Mater Today Commun* 2024;40:109527.
 - [46] Qayyum A, Javed MF, Asghar R, Iqtidar A, Alabduljabbar H, Khan MA, et al. Promoting the sustainable construction: a scientometric review on the utilization of waste glass in concrete. *Rev Adv Mater Sci* 2024;63(1):20240036.
 - [47] Bodade V, Kadrolli V. Machine learning based energy efficiency analysis with concrete waste reduction techniques and carbon footprint modelling. *Adv Concr Constr* 2024;18(2):135–46.
 - [48] Rambabu GV, Swetha RR, Parwekar P, Gangir P, Amutha S, Sivaramaraju VV. Energy efficiency and concrete waste management based on machine learning in sustainable construction. *Adv Concr Constr* 2024;18(2):125–33.
 - [49] Huang Z, Zhou H, Miao Z, Tang H, Lin B, Zhuang W. Life-Cycle Carbon Emissions (LCCE) of buildings: implications, calculations, and reductions. *Engineering* 2024;35:115–39.
 - [50] Ni BW, Rahman MZ, Guo SC, Zhu D. A review on properties and multi-objective performance predictions of concrete based on machine learning models. *Mater Today Commun* 2025;44:112017.
 - [51] Li FY, Rana MS, Qurashi MA. Advanced machine learning techniques for predicting concrete mechanical properties: a comprehensive review of models and methodologies. *Multiscale Multidiscip Model Exp Des* 2025;8:110.
 - [52] Moein MM, Saradar A, Rahmati K, Mousavinejad SHG, Bristow J, Aramali V, et al. Predictive models for concrete properties using machine learning and deep learning approaches: a review. *J Build Eng* 2023;63(Pt A):105444.
 - [53] Dinesh A, Prasad R. Predictive models in machine learning for strength and life cycle assessment of concrete structures. *Autom Constr* 2024;162:105412.
 - [54] Gamil Y. Machine learning in concrete technology: a review of current researches, trends, and applications. *Front Built Environ* 2023;9:1145591.
 - [55] Nilsson NJ. Principles of artificial intelligence. Berlin: Springer Science & Business Media; 1982.
 - [56] LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature* 2015;521(7553):436–44.
 - [57] Lei N, An D, Guo Y, Su K, Liu S, Luo Z, et al. A geometric understanding of deep learning. *Engineering* 2020;6(3):361–74.
 - [58] Patil DJ, Mason H. Data driven. Sebastopol: O'Reilly Media, Inc.; 2015.
 - [59] Zhou L, Pan S, Wang J, Vasilakos AV. Machine learning on big data: opportunities and challenges. *Neurocomputing* 2017;237:350–61.
 - [60] Wang H, Wang Y, Wang X, Yin WX, Yu TC, Xue CH, et al. Multimodal machine learning guides low carbon aeration strategies in urban wastewater treatment. *Engineering* 2024;36:51–62.
 - [61] Zhou Y, Meng S, Lou Y, Kong Q. Physics-informed deep learning-based real-time structural response prediction method. *Engineering* 2024;35:140–57.
 - [62] Zhou T, Song Z, Sundmacher K. Big data creates new opportunities for materials research: a review on methods and applications of machine learning for materials design. *Engineering* 2019;5(6):1017–26.
 - [63] Mahesh B. Machine learning algorithms-a review. *Int J Sci Res* 2020;9(1):381–6.
 - [64] Tan J, Yang J, Wu S, Chen G, Zhao J. A critical look at the current train/test split in machine learning. *arXiv:2106.04525*.
 - [65] Nti IK, Nyarko-Boateng O, Aning J. Performance of machine learning algorithms with different K values in K-fold CrossValidation. *Int J Inf Technol Comput Sci* 2021;13(6):61–71.
 - [66] Zhai X, Yin Y, Pellegrino JW, Haudek KC, Shi L. Applying machine learning in science assessment: a systematic review. *Stud Sci Educ* 2020;56(1):111–51.
 - [67] Linardatos P, Papastefanopoulos V, Kotsiantis S. Explainable ai: a review of machine learning interpretability methods. *Entropy* 2020;23(1):18.
 - [68] Antwarg L, Miller RM, Shapira B, Rokach L. Explaining anomalies detected by autoencoders using shapley additive explanations. *Expert Syst Appl* 2021;186:115736.
 - [69] Lundberg SM, Lee S. A unified approach to interpreting model predictions. In: von Luxburg U, Guyon I, Bengio S, Wallach H, Fergus R, editors. *Proceedings of the 31st International Conference on Neural Information Processing Systems (NIPS)*; 2017 Dec 4–9; Long Beach, CA, USA. Red Hook: Curran Associates Inc.; 2017. p. 4768–77.
 - [70] Greenwell BM. pdp: an R package for constructing partial dependence plots. *R J* 2017;9(1):421.
 - [71] Liu K, Zou C, Zhang X, Yan J. Innovative prediction models for the frost durability of recycled aggregate concrete using soft computing methods. *J Build Eng* 2021;34:101822.
 - [72] Amlashi AT, Abdollahi SM, Goodarzi S, Ghanizadeh AR. Soft computing based formulations for slump, compressive strength, and elastic modulus of bentonite plastic concrete. *J Clean Prod* 2019;230:1197–216.
 - [73] Yang LF, Lai BL, Xu R, Hu X, Su H, Cusatis G, et al. Prediction of alkali-silica reaction expansion of concrete using artificial neural networks. *Cem Concr Compos* 2023;140:105073.
 - [74] Lyngdoh GA, Zaki M, Krishnan N, Das S. Prediction of concrete strengths enabled by missing data imputation and interpretable machine learning. *Cem Concr Compos* 2022;128:104414.
 - [75] Machello C, Aghabalaie Baghaei K, Bazli M, Hadigheh A, Rajabipour A, Arashpour M, et al. Tree-based machine learning approach to modelling tensile strength retention of fibre reinforced polymer composites exposed to elevated temperatures. *Compos, Part B Eng* 2024;270:111132.
 - [76] Han TH, Bhat R, Ponduru SA, Sarkar A, Huang J, Sant G, et al. Deep learning to predict the hydration and performance of fly ash-containing cementitious binders. *Cem Concr Res* 2023;165:107093.
 - [77] Jaf DKI, Abdulrahman PI, Mohammed AS, Kurda R, Qaidi SMA, Asteris PG. Machine learning techniques and multi-scale models to evaluate the impact of silicon dioxide (SiO₂) and calcium oxide (CaO) in fly ash on the compressive strength of green concrete. *Constr Build Mater* 2023;400:132604.
 - [78] Al Martini S, Sabouni R, Khartabil A, Wakjira TG, Shahria AM. Development and strength prediction of sustainable concrete having binary and ternary cementitious blends and incorporating recycled aggregates from demolished UAE buildings: experimental and machine learning-based studies. *Constr Build Mater* 2023;380:131278.
 - [79] Khan MI, Abbas YM, Fares G, Alqahtani FK. Strength prediction and optimization for ultrahigh-performance concrete with low-carbon cementitious materials-XG boost model and experimental validation. *Constr Build Mater* 2023;387:131606.
 - [80] Adel H, Palizban S, Sharifi SS, Ilchi Ghazaan M, Habibnejad KA. Predicting mechanical properties of carbon nanotube-reinforced cementitious nanocomposites using interpretable ensemble learning models. *Constr Build Mater* 2022;354:129209.
 - [81] Almustaafa MK, Nehdi ML. Machine learning prediction of structural response of steel fiber-reinforced concrete beams subjected to far-field blast loading. *Cem Concr Compos* 2022;126:104378.
 - [82] Li H, Lin JJ, Zhao DW, Shi G, Wu H, Wei T, et al. A BFRC compressive strength prediction method via kernel extreme learning machine-genetic algorithm. *Constr Build Mater* 2022;344:128076.
 - [83] Yang JL, Zeng BW, Ni Z, Fan Y, Hang Z, Wang Y, et al. Comparison of traditional and automated machine learning approaches in predicting the compressive strength of graphene oxide/cement composites. *Constr Build Mater* 2023;394:132179.
 - [84] Liu KH, Dai ZH, Zhang RB, Zheng J, Zhu J, Yang X. Prediction of the sulfate resistance for recycled aggregate concrete based on ensemble learning algorithms. *Constr Build Mater* 2022;317:125917.
 - [85] Zhang JR, Niu WJ, Yang YZ, Hou D, Dong B. Machine learning prediction models for compressive strength of calcined sludge-cement composites. *Constr Build Mater* 2022;346:128442.

- [86] Aghabalaee Baghaei K, Hadigheh SA. Durability assessment of FRP-to-concrete bonded connections under moisture condition using data-driven machine learning-based approaches. *Compos Struct* 2021;114576.
- [87] Xue J, Shao JF, Burlon N. Estimation of constituent properties of concrete materials with an artificial neural network based method. *Cem Concr Res* 2021;150:106614.
- [88] Charrier M, Ouellet-Plamondon CM. Artificial neural network for the prediction of the fresh properties of cementitious materials. *Cem Concr Res* 2022;156:106761.
- [89] Han FL, Lv Y, Liu Y, Zhang X, Yu W, Cheng C, et al. Exploring interpretable ensemble learning to predict mechanical strength and thermal conductivity of aerogel-incorporated concrete. *Constr Build Mater* 2023;392:131781.
- [90] Shahmansouri AA, Yazdani M, Hosseini M, Akbarzadeh Bengar H, Farrokh GH. The prediction analysis of compressive strength and electrical resistivity of environmentally friendly concrete incorporating natural zeolite using artificial neural network. *Constr Build Mater* 2022;317:125876.
- [91] Cai R, Han TH, Liao WY, Huang J, Li D, Kumar A, et al. Prediction of surface chloride concentration of marine concrete using ensemble machine learning. *Cem Concr Res* 2020;136:106164.
- [92] Asteris PG, Skentou AD, Bardhan A, Samui P, Pilakoutas K. Predicting concrete compressive strength using hybrid ensembling of surrogate machine learning models. *Cem Concr Res* 2021;145:106449.
- [93] Liu KH, Alam MS, Zhu J, Zheng J, Chi L. Prediction of carbonation depth for recycled aggregate concrete using ANN hybridized with swarm intelligence algorithms. *Constr Build Mater* 2021;301:124382.
- [94] Zhang TR, Zhang YZ, Wang QH, Aganyira AK, Fang Y. Experimental study and machine learning prediction on compressive strength of spontaneous-combustion coal gangue aggregate concrete. *J Build Eng* 2023;71:106518.
- [95] Peng YM, Unluer C. Modeling the mechanical properties of recycled aggregate concrete using hybrid machine learning algorithms. *Resour Conserv Recycl* 2023;190:106812.
- [96] Alabdullah AA, Iqbal M, Zahid M, et al. Prediction of rapid chloride penetration resistance of metakaolin based high strength concrete using light GBM and XGBoost models by incorporating SHAP analysis. *Constr Build Mater* 2022;345:128296.
- [97] Munir MJ, Kazmi S, Wu YF, Lin X, Ahmad MR. Development of novel design strength model for sustainable concrete columns: a new machine learning-based approach. *J Clean Prod* 2022;357:131988.
- [98] Zhang LV, Marani A, Nehdi ML. Chemistry-informed machine learning prediction of compressive strength for alkali-activated materials. *Constr Build Mater* 2022;316:126103.
- [99] Zou D, Wu L, Hao Y, Xu L, Chen JJ. Composition-strength relationship study of ultrahigh performance fiber reinforced concrete (UHPFRC) using an interpretable data-driven approach. *Constr Build Mater* 2023;392:131973.
- [100] Feng WH, Wang YF, Sun JB, Tang Y, Wu D, Jiang Z, et al. Prediction of thermo-mechanical properties of rubber-modified recycled aggregate concrete. *Constr Build Mater* 2022;318:125970.
- [101] Ford E, Maneparambil K, Kumar A, Sant G, Neithalath N. Transfer (machine) learning approaches coupled with target data augmentation to predict the mechanical properties of concrete. *Mach Learn Appl* 2022;8:100271.
- [102] Ke XY, Duan Y. A Bayesian machine learning approach for inverse prediction of high-performance concrete ingredients with targeted performance. *Constr Build Mater* 2021;270:121424.
- [103] Shen JL, Li Y, Lin H, Li H, Lv J, Feng S, et al. Prediction of compressive strength of alkali-activated construction demolition waste geopolymers using ensemble machine learning. *Constr Build Mater* 2022;360:129600.
- [104] Long WJ, Cheng BY, Luo SY, Li L, Mei L. Interpretable auto-tune machine learning prediction of strength and flow properties for self-compacting concrete. *Constr Build Mater* 2023;393:132101.
- [105] Güçlüer K, Özbeyaz A, Göymen S, Günaydin O. A comparative investigation using machine learning methods for concrete compressive strength estimation. *Mater Today Commun* 2021;27:102278.
- [106] Chen HG, Yang JJ, Chen XH. A convolution-based deep learning approach for estimating compressive strength of fiber reinforced concrete at elevated temperatures. *Constr Build Mater* 2021;313:125437.
- [107] Feng JP, Zhang HW, Gao K, Liao Y, Yang J, Wu G. A machine learning and game theory-based approach for predicting creep behavior of recycled aggregate concrete. *Case Stud Constr Mater* 2022;17:e01653.
- [108] Taffese WZ, Espinosa-Leal L. A machine learning method for predicting the chloride migration coefficient of concrete. *Constr Build Mater* 2022;348:128566.
- [109] Jong SC, Ong D, Oh E. A novel Bayesian inference method for predicting optimum strength gain in sustainable geomaterials for greener construction. *Constr Build Mater* 2022;344:128255.
- [110] Zeng ZY, Zhu ZY, Yao W, Wang Z, Wang C, Wei Y, et al. Accurate prediction of concrete compressive strength based on explainable features using deep learning. *Constr Build Mater* 2022;329:127082.
- [111] Yao XF, Liu X, Sun JB, Wang B, Wang Y, Yang M, et al. AI-based performance prediction for 3D-printed concrete considering anisotropy and steam curing condition. *Constr Build Mater* 2023;375:130898.
- [112] Ghosh A, Ransinchung G. Application of machine learning algorithm to assess the efficacy of varying industrial wastes and curing methods on strength development of geopolymer concrete. *Constr Build Mater* 2022;341:127828.
- [113] Nyakilla EE, Jun G, Kasimu NA, Robert EF, Innocent N, Mohamedy T, et al. Application of machine learning in the prediction of compressive, and shear bond strengths from the experimental data in oil well cement at 80 °C. Ensemble trees boosting approach. *Constr Build Mater* 2022;317:125778.
- [114] Young BA, Hall A, Pilon L, Gupta P, Sant G. Can the compressive strength of concrete be estimated from knowledge of the mixture proportions?: new insights from statistical analysis and machine learning methods. *Cem Concr Res* 2019;115:379–88.
- [115] Rodrigues AP, Marathe S, Fernandes R, Shikha A, Shree N. Comparative analysis of machine learning techniques in the prediction of the strength of structural concrete. *Mater Today Proc* 2023;88:6–13.
- [116] Çaliskan A, Demirhan S, Tekin R. Comparison of different machine learning methods for estimating compressive strength of mortars. *Constr Build Mater* 2022;335:127490.
- [117] Sevim UK, Bilgic HH, Cansiz OF, Ozturk M, Atis CD. Compressive strength prediction models for cementitious composites with fly ash using machine learning techniques. *Constr Build Mater* 2021;271:121584.
- [118] Rajender A, Samanta AK. Compressive strength prediction of metakaolin based high-performance concrete with machine learning. *Mater Today Proc* 2023. In press.
- [119] Shah S, Chen B, Zahid M, Ahmad MR. Compressive strength prediction of one-part alkali activated material enabled by interpretable machine learning. *Constr Build Mater* 2022;360:129534.
- [120] Shamim Ansari S, Muhammad Ibrahim S, Danish Hasan S. Conventional and ensemble machine learning models to predict the compressive strength of fly ash based geopolymer concrete. *Mater Today Proc*. In press.
- [121] Jayasinghe T, Chen BW, Zhang ZR, et al. Data-driven shear strength predictions of recycled aggregate concrete beams with/without shear reinforcement by applying machine learning approaches. *Constr Build Mater* 2023;387:131604.
- [122] Lv ZQ, Jiang AN, Liang B. Development of eco-efficiency concrete containing diatomite and iron ore tailings: mechanical properties and strength prediction using deep learning. *Constr Build Mater* 2022;327:126930.
- [123] Biswas R, Li EM, Zhang N, Kumar S, Rai B, Zhou J. Development of hybrid models using metaheuristic optimization techniques to predict the carbonation depth of fly ash concrete. *Constr Build Mater* 2022;346:128483.
- [124] Mai H, Nguyen MH, Ly HB. Development of machine learning methods to predict the compressive strength of fiber-reinforced self-compacting concrete and sensitivity analysis. *Constr Build Mater* 2023;367:130339.
- [125] Gunes B, Karatosun S, Gunes O. Drilling resistance testing combined with SonReb methods for nondestructive estimation of concrete strength. *Constr Build Mater* 2023;362:129700.
- [126] Zhang M, Zhang C, Zhang JF, Wang L, Wang F. Effect of composition and curing on alkali activated fly ash-slag binders: machine learning prediction with a random forest-genetic algorithm hybrid model. *Constr Build Mater* 2023;366:129940.
- [127] Feng JP, Zhang HW, Gao K, Liao Y, Gao W, Wu G. Efficient creep prediction of recycled aggregate concrete via machine learning algorithms. *Constr Build Mater* 2022;360:129497.
- [128] Zhang C, Zhu ZD, Liu F, Yang Y, Wan Y, Huo W, et al. Efficient machine learning method for evaluating compressive strength of cement stabilized soft soil. *Constr Build Mater* 2023;392:131887.
- [129] Nguyen H, Vu T, Vo TP, Thai HT. Efficient machine learning models for prediction of concrete strengths. *Constr Build Mater* 2021;266:120950.
- [130] Sathiparan N, Jeyanthan P. Predicting compressive strength of cement-stabilized earth blocks using machine learning models incorporating cement content, ultrasonic pulse velocity, and electrical resistivity. *Nondestruct Test Eval* 2024;39(5):1045–69.
- [131] Mai H, Trinh SH, Ly HB. Enhancing compressive strength prediction of roller compacted concrete using machine learning techniques. *Measurement* 2023;218:113196.
- [132] Salami BA, Iqbal M, Abdulraheem A, Jalal FE, Alimi W, Jamal A, et al. Estimating compressive strength of lightweight foamed concrete using neural, genetic and ensemble machine learning approaches. *Cem Concr Compos* 2022;133:104721.
- [133] Gao XF, Yang J, Zhu H, Xu J. Estimation of rubberized concrete frost resistance using machine learning techniques. *Constr Build Mater* 2023;371:130778.
- [134] Basaran B, Kalkan I, Bergil E, Erdal E. Estimation of the FRP-concrete bond strength with code formulations and machine learning algorithms. *Compos Struct* 2021;268:113972.
- [135] Sami B, Sami B, Kumar P, Ahmed AN, Amieghemen GE, Sherif MM, et al. Feasibility analysis for predicting the compressive and tensile strength of concrete using machine learning algorithms. *Case Stud Constr Mater* 2023;18:e01893.
- [136] Li QF, Song ZM. High-performance concrete strength prediction based on ensemble learning. *Constr Build Mater* 2022;324:126694.
- [137] Tanyildizi H, Marani A, Türk K, Nehdi ML. Hybrid deep learning model for concrete incorporating microencapsulated phase change materials. *Constr Build Mater* 2022;319:126146.
- [138] Liang MF, Chang Z, Wan Z, Gan Y, Schlangen E, Šavija B. Interpretable ensemble-machine-learning models for predicting creep behavior of concrete. *Cem Concr Compos* 2022;125:104295.
- [139] Hilloulin B, Tran V. Interpretable machine learning model for autogenous shrinkage prediction of low-carbon cementitious materials. *Constr Build Mater* 2023;396:132343.
- [140] Imran H, Ibrahim M, Al-Shoukry S, Rustam F, Ashraf I. Latest concrete materials dataset and ensemble prediction model for concrete compressive

- strength containing RCA and GGBFS materials. *Constr Build Mater* 2022;325:126525.
- [141] Mehta V. Machine learning approach for predicting concrete compressive, splitting tensile, and flexural strength with waste foundry sand. *J Build Eng* 2023;70:106363.
- [142] Cosgun C. Machine learning for the prediction of evaluation of existing reinforced concrete structures performance against earthquakes. *Structures* 2023;50:1994–2003.
- [143] Dias PP, Jayasinghe LB, Waldmann D. Machine learning in mix design of *Miscanthus* lightweight concrete. *Constr Build Mater* 2021;302:124191.
- [144] Shah MI, Javed MF, Aslam F, Alabduljabbar H. Machine learning modeling integrating experimental analysis for predicting the properties of sugarcane bagasse ash concrete. *Constr Build Mater* 2022;314:125634.
- [145] Nunez I, Nehdi ML. Machine learning prediction of carbonation depth in recycled aggregate concrete incorporating SCMs. *Constr Build Mater* 2021;287:123027.
- [146] Chi L, Wang M, Liu KH, Lu S, Kan L, Xia X, et al. Machine learning prediction of compressive strength of concrete with resistivity modification. *Mater Today Commun* 2023;36:106470.
- [147] El Mir A, El-Zahab S, Sbartaï ZM, Homsî F, Saliba J, El-Hassan H. Machine learning prediction of concrete compressive strength using rebound hammer test. *J Build Eng* 2023;64:105538.
- [148] Eyo EU, Abbey SJ. Machine learning regression and classification algorithms utilised for strength prediction of OPC/by-product materials improved soils. *Constr Build Mater* 2021;284:122817.
- [149] Gomaa E, Han TH, Elgawady M, Huang J, Kumar A. Machine learning to predict properties of fresh and hardened alkali-activated concrete. *Cem Concr Compos* 2021;115:103863.
- [150] Shamsabadi EA, Roshan N, Hadigheh SA, Nehdi ML, Khodabakhshian A, Ghalehnovi M. Machine learning-based compressive strength modelling of concrete incorporating waste marble powder. *Constr Build Mater* 2022;324:126592.
- [151] Almasabha G, Al-Shboul KF, Shehadeh A, Alshboul O. Machine learning-based models for predicting the shear strength of synthetic fiber reinforced concrete beams without stirrups. *Structures* 2023;52:299–311.
- [152] Moaf FO, Kazemi F, Abdelgader HS, Kurpińska M. Machine learning-based prediction of preplaced aggregate concrete characteristics. *Eng Appl Artif Intell* 2023;123:106387.
- [153] Kazemi F, Asgarkhani N, Jankowski R. Machine learning-based seismic fragility and seismic vulnerability assessment of reinforced concrete structures. *Soil Dyn Earthq Eng* 2023;166:107761.
- [154] Kim HK, Lim Y, Tafesse M, Kim GM, Yang B. Micromechanics-integrated machine learning approaches to predict the mechanical behaviors of concrete containing crushed clay brick aggregates. *Constr Build Mater* 2022;317:125840.
- [155] Antunes R. Multivariate regression model for peak temperatures in massive elements statistically verified by artificial neural networks. *Constr Build Mater* 2022;316:126072.
- [156] Go C, Kwak YJ, Kwag S, Eem S, Lee S, Ju B. On developing accurate prediction models for residual tensile strength of GFRP bars under alkaline-concrete environment using a combined ensemble machine learning methods. *Case Stud Constr Mater* 2023;18:e02157.
- [157] Hafez H, Teirelbar A, Kurda R, Tošić N, de la Fuente A. Pre-bcc: a novel integrated machine learning framework for predicting mechanical and durability properties of blended cement concrete. *Constr Build Mater* 2022;352:129019.
- [158] Fan DQ, Yu R, Fu SY, Yue L, Wu C, Shui Z, et al. Precise design and characteristics prediction of Ultra-High Performance Concrete (UHPC) based on artificial intelligence techniques. *Cem Concr Compos* 2021;122:104171.
- [159] Jeon D, Jung J, Park J, Min J, Oh JE, Moon J, et al. Predicting airborne chloride deposition in marine bridge structures using an artificial neural network model. *Constr Build Mater* 2022;337:127623.
- [160] Bhat R, Han TH, Ponduru SA, Reka A, Huang J, Sant G, et al. Predicting compressive strength of alkali-activated systems based on the network topology and phase assemblages using tree-structure computing algorithms. *Constr Build Mater* 2022;336:127557.
- [161] Adesanya E, Aladejare A, Adediran A, Lawal A, Illikainen M. Predicting shrinkage of alkali-activated blast furnace-fly ash mortars using artificial neural network (ANN). *Cem Concr Compos* 2021;124:104265.
- [162] Chao ZM, Wang MY, Sun YN, Xu X, Yue W, Yang C, et al. Predicting stress-dependent gas permeability of cement mortar with different relative moisture contents based on hybrid ensemble artificial intelligence algorithms. *Constr Build Mater* 2022;348:128660.
- [163] Song HW, Ahmad A, Farooq F, Ostrowski KA, Maślak M, Czarnecki S, et al. Predicting the compressive strength of concrete with fly ash admixture using machine learning algorithms. *Constr Build Mater* 2021;308:125021.
- [164] Penido RE, de Paixao RCF, Costa LCB, Peixoto RAF, Cury AA, Mendes JC. Predicting the compressive strength of steelmaking slag concrete with machine learning-considerations on developing a mix design tool. *Constr Build Mater* 2022;341:127896.
- [165] Wang Z, Liu TX, Long ZL, Wang J, Zhang J. Predicting the drift capacity of precast concrete columns using explainable machine learning approach. *Eng Struct* 2023;282:115771.
- [166] Navarrete I, La Fé-Perdomo I, Ramos-Grez JA, et al. Predicting the evolution of static yield stress with time of blended cement paste through a machine learning approach. *Constr Build Mater* 2023;371:130632.
- [167] Tanyildizi H. Predicting the geopolymerization process of fly ash-based geopolymer using deep long short-term memory and machine learning. *Cem Concr Compos* 2021;123:104177.
- [168] Canbek O, Xu QZ, Mei YJ, Washburn NR, Kurtis KE. Predicting the rheology of limestone calcined clay cements (LC³): linking composition and hydration kinetics to yield stress through Machine Learning. *Cem Concr Res* 2022;160:106925.
- [169] Xi B, Huang ZW, Al-Obaidi S, Ferrara L. Predicting ultra high-performance concrete self-healing performance using hybrid models based on metaheuristic optimization techniques. *Constr Build Mater* 2023;381:131261.
- [170] Sun YB, Cheng H, Zhang SZ, Mohan MK, Ye G, De Schutter G. Prediction & optimization of alkali-activated concrete based on the random forest machine learning algorithm. *Constr Build Mater* 2023;385:131519.
- [171] Fan CC, Zheng YX, Wang SQ, Ma J. Prediction of bond strength of reinforced concrete structures based on feature selection and GWO-SVR model. *Constr Build Mater* 2023;400:132602.
- [172] Liu QF, Iqbal MF, Yang J, Lu X, Zhang P, Rauf M. Prediction of chloride diffusivity in concrete using artificial neural network: modelling and performance evaluation. *Constr Build Mater* 2021;268:121082.
- [173] Sun Z, Li YL, Li YQ, Su L, He W. Prediction of chloride ion concentration distribution in basalt-polypropylene fiber reinforced concrete based on optimized machine learning algorithm. *Mater Today Commun* 2023;36:106565.
- [174] Parhi SK, Patro SK. Prediction of compressive strength of geopolymer concrete using a hybrid ensemble of grey wolf optimized machine learning estimators. *J Build Eng* 2023;71:106521.
- [175] Kocak B, Pınarç B, Güvenç U, Kocak Y. Prediction of compressive strengths of pumice-and diatomite-containing cement mortars with artificial intelligence-based applications. *Constr Build Mater* 2023;385:131516.
- [176] Cao C. Prediction of concrete porosity using machine learning. *Results Eng* 2023;17:100794.
- [177] Dehestani A, Kazemi F, Abdi R, Nitka M. Prediction of fracture toughness in fibre-reinforced concrete, mortar, and rocks using various machine learning techniques. *Eng Fract Mech* 2022;276:108914.
- [178] Nagaraju TV, Mantena S, Azab M, Alisha SS, El Hachem C, Adamu M, et al. Prediction of high strength ternary blended concrete containing different silica proportions using machine learning approaches. *Results Eng* 2023;17:100973.
- [179] Zhang H, Guo QQ, Xu LY. Prediction of long-term prestress loss for prestressed concrete cylinder structures using machine learning. *Eng Struct* 2023;279:115577.
- [180] Amirî M, Hatami F. Prediction of mechanical and durability characteristics of concrete including slag and recycled aggregate concrete with artificial neural networks (ANNs). *Constr Build Mater* 2022;325:126839.
- [181] Hosseinzadeh M, Dehestani M, Hosseinzadeh A. Prediction of mechanical properties of recycled aggregate fly ash concrete employing machine learning algorithms. *J Build Eng* 2023;76:107006.
- [182] Zhang XQ, Akbar MZ, Zheng W. Prediction of seven-day compressive strength of field concrete. *Constr Build Mater* 2021;305:124604.
- [183] Chen RY, Song JJ, Xu MB, Wang X, Yin Z, Liu T, et al. Prediction of the corrosion depth of oil well cement corroded by carbon dioxide using GA-BP neural network. *Constr Build Mater* 2023;394:132127.
- [184] Congro M, Monteiro V, Brandao A, Santos BF, Roehl D, Silva FA. Prediction of the residual flexural strength of fiber reinforced concrete using artificial neural networks. *Constr Build Mater* 2021;303:124502.
- [185] Shanmugasundaram N, Praveenkumar S, Gayathiri K, Divya S. Prediction on compressive strength of engineered cementitious composites using Machine learning approach. *Constr Build Mater* 2022;342:127933.
- [186] Mei Y, Sun YH, Li F, Xu X, Zhang A, Shen J. Probabilistic prediction model of steel to concrete bond failure under high temperature by machine learning. *Eng Fail Anal* 2022;142:106786.
- [187] Wang NL, Xia ZJ, Amin MN, Ahmad W, Khan K, Althoey F, et al. Sustainable strategy of eggshell waste usage in cementitious composites: an integral testing and computational study for compressive behavior in aggressive environment. *Constr Build Mater* 2023;386:131536.
- [188] Mai H, Nguyen MH, Trinh SH, Ly HB. Toward improved prediction of recycled brick aggregate concrete compressive strength by designing ensemble machine learning models. *Constr Build Mater* 2023;369:130613.
- [189] El-Sayed AA, Fathy IN, Tayeh BA, Almeshal I. Using artificial neural networks for predicting mechanical and radiation shielding properties of different nano-concretes exposed to elevated temperature. *Constr Build Mater* 2022;324:126663.
- [190] Chen N, Zhao SB, Gao ZW, Wang D, Liu P, Oeser M, et al. Virtual mix design: prediction of compressive strength of concrete with industrial wastes using deep data augmentation. *Constr Build Mater* 2022;323:126580.
- [191] Chen Z, Lin J, Sagoe-Crentsil K, Duan W. Development of hybrid machine learning-based carbonation models with weighting function. *Constr Build Mater* 2022;321:126359.
- [192] Shi Y, Yang Z, Ma S, Kang PL, Shang C, Hu P, et al. Machine learning for chemistry: basics and applications. *Engineering* 2023;27:70–83.
- [193] Li Y, Liu YJ, Miao YC, Liu Z, Jiang J, Lin J. Development of heat-resistant tunnel muck-based shotcrete for geothermal environments: dual drive of combining explainable machine learning and microstructure characterization. *Constr Build Mater* 2025;473:140994.

- [194] Sobuz MHR, Rahman MM, Aayaz R, Al-Rashed WS, Datta SD, Safayet MA, et al. Combined influence of modified recycled concrete aggregate and metakaolin on high-strength concrete production: experimental assessment and machine learning quantifications with advanced SHAP and PDP analyses. *Constr Build Mater* 2025;461:139897.
- [195] Luo X, Li Y, Lin H, Hao M, Wang H. Machine learning-driven modeling and importance analysis of drying shrinkage behavior in magnesium silicate hydrate cement. *J Build Eng* 2025;107:112721.
- [196] Kondratieva TN, Tyurina VS, Chepurmenko AS. Predicting the risk of early cracking in massive monolithic foundation slabs using artificial intelligence algorithms. *Constr Mater Prod* 2025;8(1):6.
- [197] Yao QL, Tu YL, Yang JH. Predicting the compressive strength of solid waste-cement stabilized compacted soil using machine learning model. *Mater Today Commun* 2025;44:11882.
- [198] Song C, Xiao R, Zhang C, Zhao X, Sun B. Simulation-free reliability analysis with importance sampling-based adaptive training physics-informed neural networks: method and application to chloride penetration. *Reliab Eng Syst Saf* 2024;246:110083.
- [199] Zhang TJ, Wang DL, Lu Y. RheologyNet: a physics-informed neural network solution to evaluate the thixotropic properties of cementitious materials. *Cem Concr Res* 2023;168:107157.
- [200] Wan Y, Zheng W, Wang Y. Identification of chloride diffusion coefficient in concrete using physics-informed neural networks. *Constr Build Mater* 2023;393:132049.
- [201] Rahman MA, Lu Y. EcoBlendNet: a physics-informed neural network for optimizing supplementary material replacement to reduce the carbon footprint during cement hydration. *J Clean Prod* 2024;464:142777.
- [202] Liu Q, Andersen LV, Wu M. Prediction of concrete abrasion depth and computational design optimization of concrete mixtures. *Cem Concr Compos* 2024;148:105431.
- [203] Ali Hadi K, Ridha AS. Deep learning techniques in concrete powder mix designing. *Open Eng* 2024;14(1):20220588.
- [204] Shahrokhishahraki M, Malekpour M, Mirvalad S, Faraone G. Machine learning predictions for optimal cement content in sustainable concrete constructions. *J Build Eng* 2024;82:108160.
- [205] Li Z, Gao Y, Zhu Z, Tian W. Data-guided for discovering high-strength, cost-effective, and low-carbon rice husk ash concrete. *J CO₂ Util* 2024;83:102786.
- [206] Cheng B, Mei L, Long W, Luo Q, Zhang J, Xiong C, et al. Data driven multi-objective design for low-carbon self-compacting concrete considering durability. *J Clean Prod* 2024;450:141947.
- [207] Wang W, Chen SJ, Duan WH, Sagoe-Crentsil K, Pathirage CSN, Li L, et al. Revealing microstructural modifications of graphene oxide-modified cement via deep learning and nanoporosity mapping: implications for structural materials' performance. *ACS Appl Nano Mater* 2022;5(5):7092–102.
- [208] Szafraniec M, Omiotek Z, Barnat-Hunek D. Water absorption prediction of nanopolymer hydrophobized concrete surface using texture analysis and machine learning algorithms. *Constr Build Mater* 2023;375:130969.
- [209] Lin JL, Liu YM, Sui H, Sagoe-Crentsil K, Duan W. Microstructure of graphene oxide-silica-reinforced OPC composites: image-based characterization and nano-identification through deep learning. *Cem Concr Res* 2022;154:106737.
- [210] Zhang R, Yan XF, Guo L. Deep learning-based classification of damage-induced acoustic emission signals in UHPC. *Constr Build Mater* 2022;356:129885.
- [211] Sui H, Wang W, Lin JL, Tang ZQ, Yang DS, Duan W. Spatial correlation and pore morphology analysis of limestone calcined clay cement (LC³) via machine learning and image-based characterisation. *Constr Build Mater* 2023;401:132721.
- [212] Nguyen HD, Wang W, Yao XP, Sagoe-Crentsil K, Duan W. Porosity-modulus mapping enhanced nanomechanical analysis of heterogeneous materials. *J Mater Sci* 2023;58(24):10058–72.
- [213] Liang MF, Gan YD, Chang Z, Zhi W, Erik S, Branko Š. Microstructure-informed deep convolutional neural network for predicting short-term creep modulus of cement paste. *Cem Concr Res* 2022;152:106681.
- [214] Guo JF, Li MH, Wang L, Yang B, Zhang L, Chen Z, et al. Estimating cement compressive strength using three-dimensional microstructure images and deep belief network. *Eng Appl Artif Intell* 2020;88:103378.
- [215] Komninos P, Verraest A, Eleftheroglou N, Zarouchas D. Intelligent fatigue damage tracking and prognostics of composite structures utilizing raw images via interpretable deep learning. *Compos B Eng* 2024;287:111863.
- [216] Hou Y, Li Q, Zhang C, Lu G, Ye Z, Chen Y, et al. The state-of-the-art review on applications of intrusive sensing, image processing techniques, and machine learning methods in pavement monitoring and analysis. *Engineering* 2021;7(6):845–56.
- [217] Zhang YH, Zheng JQ, Sun W, Shan L. Image recognition method of building wall cracks based on feature distribution. *Soft Comput* 2020;24(11):8285–94.
- [218] Zhang Q, Barri K, Babanajad SK, Alavi AH. Real-time detection of cracks on concrete bridge decks using deep learning in the frequency domain. *Engineering* 2021;7(12):1786–96.
- [219] Bao Y, Chen Z, Wei S, Xu Y, Tang Z, Li H. The state of the art of data science and engineering in structural health monitoring. *Engineering* 2019;5(2):234–42.
- [220] Du B, Ye J, Zhu H, Sun L, Du Y. Intelligent monitoring system based on spatio-temporal data for underground space infrastructure. *Engineering* 2023;25:194–203.
- [221] Shi ZQ, Jin N, Chen DB, Ai D. A comparison study of semantic segmentation networks for crack detection in construction materials. *Constr Build Mater* 2024;414:134850.
- [222] Shi MN, Hua TB, Yang ZH, Tan C, Wen Y. Weakly supervised deep learning-based concrete aggregates automatic segmentation for assessing separation degree. *J Build Eng* 2024;82:108342.
- [223] Sheiati S, Nguyen H, Kinnunen P, Ranjbar N. Cementitious phase quantification using deep learning. *Cem Concr Res* 2023;172:107231.
- [224] Yu F, Cai H, Zhang H, Hu M, Zhang R, Gao Z. Automatic identification method for three-phase structure of pervious concrete based on deep learning network of Mask R-CNN. *Constr Build Mater* 2024;420:135534.
- [225] Hao ZX, Lu C, Li ZH. Highly accurate and automatic semantic segmentation of multiple cracks in engineered cementitious composites (ECC) under dual pre-modification deep-learning strategy. *Cem Concr Res* 2023;165:107066.
- [226] Han YM, Wang L, Wang YQ, Geng Z. Intelligent small sample defect detection of concrete surface using novel deep learning integrating improved YOLOv5. *IEEE-CAA J Automatica Sin* 2024;11(2):545–7.
- [227] Liang J, Gu XY, Jiang D, Zhang Q. CNN-based network with multi-scale context feature and attention mechanism for automatic pavement crack segmentation. *Autom Constr* 2024;164:105482.
- [228] Jiang YQ, Pang DD, Li CD, Wang J. A method of concrete damage detection and localization based on weakly supervised learning. *Comput Aided Civ Infrastruct Eng* 2024;39(7):1042–60.
- [229] Song Y, Huang ZL, Shen CY, Shi H, Lange DA. Deep learning-based automated image segmentation for concrete petrographic analysis. *Cem Concr Res* 2020;135:106118.
- [230] Saha SK, Pradhan S, Barai SV. Use of machine learning based technique to X-ray microtomographic images of concrete for phase segmentation at meso-scale. *Constr Build Mater* 2020;249:118874.
- [231] Chen W, Iyer A, Bostanabad R. Data centric design: a new approach to design of microstructural material systems. *Engineering* 2022;10:89–98.
- [232] Hong SW, Kim SY, Park K, Terada K, Lee H, Han TS. Mechanical property evaluation of 3D multi-phase cement paste microstructures reconstructed using generative adversarial networks. *Cem Concr Compos* 2024;152:105646.
- [233] Zhao X, Wang L, Li QF, Chen H, Liu S, Hou P, et al. 3D microstructural generation from 2D images of cement paste using generative adversarial networks. *Cem Concr Res* 2025;187:107726.
- [234] Lin JL, Chen SJ, Wang W, Pathirage CSN, Li L, Sagoe-Crentsil K, et al. Transregional spatial correlation revealed by deep learning and implications for material characterisation and reconstruction. *Mater Charact* 2021;178:111268.
- [235] Yao XP, Fang HY, Du MR, Feng H, Zhai K, Lin J, et al. Evolution of the microporous structure in cement hydration: a deep learning-based image translation method. *J Build Eng* 2024;94:110065.
- [236] Huang B, Kang F, Li XY, Zhu S. Underwater dam crack image generation based on unsupervised image-to-image translation. *Autom Constr* 2024;163:105430.
- [237] Li SY, Zhao XF. High-resolution concrete damage image synthesis using conditional generative adversarial network. *Autom Constr* 2023;147:104739.
- [238] Qoku E, Xu K, Li J, Monteiro PJM, Kurtis KE. Advances in imaging, scattering, spectroscopy, and machine learning-aided approaches for multiscale characterization of cementitious systems. *Cem Concr Res* 2023;174:107335.
- [239] Tong Z, Wang Z, Wang X, Ma Y, Guo H, Liu C. Characterization of hydration and dry shrinkage behavior of cement emulsified asphalt composites using deep learning. *Constr Build Mater* 2021;274:121898.
- [240] Zheng XY, Watanabe I, Paik J, Li J, Guo X, Naito M. Text-to-microstructure generation using generative deep learning. *Small* 2024;20(37):2402685.
- [241] Zhang H, Xu T, Li H, Zhang S, Wang X, Huang X, et al. StackGAN++: realistic image synthesis with stacked generative adversarial networks. *IEEE Trans Pattern Anal Mach Intell* 2019;41(8):1947–62.
- [242] Dong P, Wu L, Li R, Meng X, Meng L. Text to image synthesis with multi-granularity feature aware enhancement generative adversarial networks. *Comput Vis Image Underst* 2024;245:104042.
- [243] Tan YX, Lee CP, Neo M, Lim KM, Lim JY. Text-to-image synthesis with self-supervised bi-stage generative adversarial network. *Pattern Recognit Lett* 2023;169:43–9.
- [244] Ge X, Goodwin RT, Yu H, Romero P, Abdelrahman O, Sudhakar A, et al. Accelerated design and deployment of low-carbon concrete for data centers. In: Bainomugisha E, Brunette W, Dell N, Sharma A, editors. *Proceedings of the ACM International Conference Proceeding Series: 2022 Jun 29–Jul 1; Seattle, WA, USA*. New York City: Association for Computing Machinery (ACM); 2022. p. 340–52.
- [245] Converge launches AI machine learning program for concrete strength prediction [Internet]. Greater Manchester: Highways Today; 2019 Nov 27 [2025 May 21]. Available from: <https://highways.today/2019/11/27/converge-ai-concrete-strength/>.
- [246] Mirzazade A, Popescu C, Blanksvärd T, Täljsten B. Workflow for off-site bridge inspection using automatic damage detection-case study of the Pahtajokk Bridge. *Remote Sens* 2021;13(14):2665.
- [247] Nguyen CL, Nguyen A, Brown J, Byrne T, Ngo BT, Luong CX. Optimising concrete crack detection: a study of transfer learning with application on NVIDIA Jetson Nano. *Sensors* 2024;24(23):7818.
- [248] Greener JG, Kandathil SM, Moffat L, Jones DT. A guide to machine learning for biologists. *Nat Rev Mol Cell Biol* 2022;23(1):40–55.
- [249] Qi XB, Chen GF, Li Y, Cheng X, Li C. Applying neural-network-based machine learning to additive manufacturing: current applications, challenges, and future perspectives. *Engineering* 2019;5(4):721–9.
- [250] Fiona Zhao Y, Xie J, Sun L. On the data quality and imbalance in machine learning-based design and manufacturing—a systematic review. *Engineering* 2025;45:105–31.

- [251] Tezsezen E, Yigci D, Ahmadpour A, Tasoglu S. AI-based metamaterial design. *ACS Appl Mater Interfaces* 2024;16(23):29547–69.
- [252] Zhang Z, Li S, Feng H, Zhou X, Xu N, Li H, et al. Machine learning for bridge wind engineering. *Adv Wind Eng* 2024;1(1):100002.
- [253] Dobbelaere MR, Plehiers PP, Van de Vijver R, Stevens CV, Van Geem KM. Machine learning in chemical engineering: strengths, weaknesses, opportunities, and threats. *Engineering* 2021;7(9):1201–11.
- [254] Gao Y, Wang JY, Xu XX. Machine learning in construction and demolition waste management: progress, challenges, and future directions. *Autom Constr* 2024;162:105380.
- [255] Dev B, Rahman MA, Islam MJ, Rahman MZ, Zhu D. Properties prediction of composites based on machine learning models: a focus on statistical index approaches. *Mater Today Commun* 2024;38:107659.