

Research Article

Prediction on tensile mechanical properties of engineered cementitious composites using a fully automated deep learning method

Liuyi Liu, Beixin Xie, Jinghao Huang, Jiahao Wen, Ge Zhang[✉], Jiaxiang Lin

School of Civil and Transportation Engineering, Guangdong University of Technology, Guangzhou 510006, China

✉ Address correspondence to Ge Zhang, gezhang@gdut.edu.cn

Received: 24 May 2026; Revised: 6 September 2026; Accepted: 30 September 2026

© The Author(s) 2026. This is an open access article under the terms of the Creative Commons Attribution 4.0 International License (CC BY 4.0, <http://creativecommons.org/licenses/by/4.0/>).

ABSTRACT

Accurately predicting tensile mechanical properties of Engineered Cementitious Composites (ECC), including tensile strength (TS) and ultimate tensile strain (UTS), remains challenging, particularly for the ultimate tensile strain, which exhibits high data dispersion that hinders reliable predictions. While machine learning (ML) has shown significant promise in predicting the mechanical properties of ECC, it faces several challenges: (1) limited predictive accuracy; (2) complexity and time-consuming nature of hyperparameter tuning; (3) insufficient research into predicting the UTS of ECC. To overcome these challenges, this study proposes a fully automated deep learning model enhanced by Bayesian optimization for predicting the tensile mechanical properties of polyethylene fiber-reinforced engineered cementitious composites (PE-ECC). First, convolutional neural network (CNN) model was employed to enhance feature extraction and predictive accuracy. Second, Bayesian optimization was utilized to automatically construct the network and optimize various hyperparameters, resulting in an optimized network. This research demonstrates the effectiveness of

the proposed Bayesian-optimized convolutional neural network (BO-CNN) method in predicting PE-ECC's TS and UTS, utilizing a dataset of 154 experimental samples from published literature, achieving coefficients of determination (R^2) of 0.911 and 0.775, respectively. The BO-CNN method not only outperforms CNNs and artificial neural networks (ANNs) optimized through grid search but also surpasses other eleven traditional ML models. Furthermore, SHapley Additive exPlanations (SHAP) was employed to interpret input features' contributions in the prediction of tensile mechanical properties. Overall, the proposed method provides an efficient and reliable solution for predicting the tensile mechanical properties of ECC.

Keywords: Engineered cementitious composites; polyethylene fiber; deep learning model; tensile mechanical property; Bayesian optimization

1. Introduction

Engineered Cementitious Composites (ECC) is a high-performance fiber-reinforced concrete (FRC) that has garnered significant attention due to its unique pseudo strain-hardening behavior and exceptional multiple-cracking tensile performance [1–3]. Unlike conventional concrete, ECC does not exhibit brittle fracture under tension. Instead, it demonstrates ductility similar to that of metallic materials, forming fine cracks throughout the tensile process without complete failure. The composition of ECC primarily consists of a matrix (cement, sand, water, etc.) and fiber. Recent studies indicate that the tensile behavior of ECC depends not only on the volume fraction of fiber [4–8] but also on several other factors: fiber properties (mechanical properties, aspect ratio, etc.) [9, 10], matrix properties (initial flaw size distribution, inherent mechanical properties, etc.) [11–13], fiber-matrix interface properties (chemical bonding, frictional interaction, etc.) [14], and concrete admixture properties (type, volume, etc.) [15]. However, fully understanding the tensile mechanical properties of ECC under the influence of these various factors relies on extensive experimentation, which is often uneconomical and time-consuming. Therefore, there is an urgent need to develop methods that can swiftly, efficiently, and accurately predict the mechanical properties of ECC using available experimental data.

Due to its advantages in addressing complex nonlinear problems, machine learning (ML) has been widely adopted in civil engineering [16], with applications such as structural damage detection and the prediction of mechanical properties [17-19]. The most commonly used ML methods for predicting concrete mechanical properties can be categorized into four main groups: empirical risk minimization (e.g., artificial neural network, ANN), structural risk minimization (e.g., support vector machine, SVM), ensemble algorithms (e.g., decision tree, DT), and evolutionary algorithms [20]. For instance, Fan et al. [21] utilized SVM to predict the compressive strength of concrete based on its constituent materials. Behnood and Golashfani [22] employed the M5P algorithm to model various properties of concrete containing waste foundry sand, including compressive strength, modulus of elasticity, flexural strength, and splitting tensile strength. Nguyen et al. [23] applied ensemble algorithms such as DT and extreme gradient boosting (XGBoost) to predict the strain at peak tensile stress and energy absorption capacity of ultrahigh-performance fiber-reinforced concrete (UHPFRC). Alyami et al. [24] implemented seven ML algorithms, including support vector regression (SVR), DT, SVR-bagging, SVR-boosting, random forests (RF), gradient boosting (GB), and gene expression programming (GEP) to predict the compressive strength of 3D-printed fiber-reinforced concrete (3DP-FRC).

In addition to the aforementioned ML algorithms, neural networks, as a significant class of ML algorithms, have also begun to be applied in predicting the mechanical properties of concrete. Uddin et al. [25] employed ANNs to predict the compressive strength and tensile strength of polyvinyl alcohol fiber-reinforced engineered cementitious composite (PVA-ECC). Malik et al. [26] used an ANN model that incorporates ten input parameters, including all the ingredients of the mix design and the properties of the fibers utilized, to predict the complete tensile stress-strain curve of ECC in terms of a bilinear idealized curve. Furthermore, some researchers have begun to compare the predictive capabilities of neural networks with other ML models in evaluating the mechanical properties of concrete. Feng et al. [27] employed backpropagation neural networks (BPNN) alongside three other ML models, including RF, linear regression (LR), and multiple linear regression (MLR) to predict the uniaxial compressive strength and corresponding peak strain of rubber-modified recycled aggregate

concrete (RRAC) [28, 29]. Validation based on a dataset comprising 120 samples revealed that BPNN exhibited superior predictive performance compared to the other ML algorithms. Lyngdoh et al. predicted the strength of concrete using various ML approaches, including polynomial regression (PR), least absolute shrinkage and selection operator (LASSO), SVM, RF, XGBoost, and ANN. They found that ANN and XGBoost demonstrated excellent predictive efficiency when dealing with datasets containing a higher degree of missing data [30]. These results underscore the strong capabilities of neural network models in accurately predicting the mechanical properties of ECC and various other types of concrete.

Unlike traditional neural networks and other ML algorithms mentioned above, convolutional neural networks (CNNs) represent an emerging class of deep learning (DL) techniques [31,32]. With multiple built-in hidden layers, CNN can extract deeper features from input samples and utilize them effectively, thereby enhancing the network's predictive capabilities in complex nonlinear problems. Researchers have applied CNN in various fields, including structural engineering, earthquake prevention, and disaster mitigation. Notably, image-based detection and evaluation of concrete surface cracks are widely employed [33–36]. Shaswat et al. [37] optimized CNN to predict concrete slump values by hybridizing sea lion prediction results. Deng et al. [38] developed a CNN model to estimate the compressive strength of recycled aggregate concrete, demonstrating that CNN exhibited greater adaptability and efficacy in learning deeper features from input variables compared to traditional ANN model. Additionally, Li et al. [39] validated the effectiveness of CNN in predicting concrete compressive strength while accounting for the influence of coarse aggregate type and gradation. Their findings indicated that the prediction performance of the CNN model, which considers aggregate types and gradations, outperformed that of the ANN model. These studies indicate that CNNs can provide effective feature extraction and improved predictive performance in concrete-property regression. However, their suitability for predicting the tensile properties of ECC requires further investigation.

Current research identifies several shortcomings in models that predict the mechanical properties of ECC and other concrete materials: (1) traditional ML models, such as RF and SVM, face

limitations in feature extraction and in addressing complex nonlinear relationships, which negatively impact predictive accuracy; (2) the optimization processes for these ML models often require manual adjustments or extensive hyperparameter tuning, constraining their optimization potential, reducing predictive accuracy and efficiency, and prolonging development timelines [40–42]; (3) Furthermore, there is a scarcity of studies utilizing ML to predict the mechanical properties of ECC, particularly ultimate tensile strain, due to its higher data dispersion, which complicates predictions [26, 43–46]. Further exploration in this area is warranted.

This study presents an advanced predictive methodology for assessing the tensile properties of polyethylene fiber-reinforced ECC (PE-ECC), addressing the previously identified limitations. The approach employed a CNN optimized through Bayesian optimization techniques, referred to as BO-CNN, to predict the tensile properties of ECC. A dataset consisting of 154 samples, compiled from existing literature [7, 46–89], was utilized for this purpose. The model incorporated ten input features, including matrix and fiber properties, and was specifically designed to predict two primary output targets: tensile strength (TS) and ultimate tensile strain (UTS). The efficacy of the BO-CNN model was rigorously evaluated using various performance metrics, including the correlation coefficient (R), standard deviation (SD) and so on. Its superiority was demonstrated through comparative analyses with other thirteen representative models. Additionally, SHapley Additive exPlanations (SHAP) was conducted to elucidate the contribution of each input feature to the model's outputs. The findings of this study are significant, as they present a fully automated prediction framework for the tensile strength and ultimate tensile strain of ECC, achieving optimal predictive performance while functioning independently of human intervention.

2. Database

2.1. Overview

The first step in developing a predictive model is data collection. This study utilized a dataset compiled from published literature on ECC materials, consisting of 154 samples of PE-ECC. The raw materials included ordinary Portland cement 52.5R (C), fly ash (FA), ground granulated blast-furnace slag (GGBS), lime powder (LP), silica fume (SF), sand (S), water (W), and polyethylene (PE) fiber.

Specimens were cured under standard conditions for 28 days and subjected to uniaxial tensile tests to determine tensile strength and ultimate tensile strain. The dataset encompassed 12 features: 10 input features and 2 output targets. The input features pertained to matrix and fiber properties, specifically: the mass ratios of cement to binder materials (C/B), fly ash to binder materials (FA/B), ground granulated blast-furnace slag to binder materials (GGBS/B), lime powder to binder materials (LP/B), silica fume to binder materials (SF/B), sand to binder materials (S/B), water to binder materials (W/B), fiber volume fraction (V_f), fiber length (L_f), and fiber diameter (D_f). Here, the mass of the binder materials (B) is defined as the sum of C, FA, GGBS, LP and SF. The output targets were tensile strength and ultimate tensile strain. Table 1 presents the statistical information and range of features in the dataset, highlighting their variation.

Table 1. Statistical details of the features in the dataset.

	Features	Unit	Min	Max	Mean	SD
Inputs	C/B	/	0.200	0.830	0.509	0.163
	FA/B	/	0.000	0.800	0.182	0.244
	GGBS/B	/	0.000	0.590	0.177	0.215
	LP/B	/	0.000	0.490	0.038	0.083
	SF/B	/	0.000	0.400	0.095	0.090
	S/B	/	0.059	0.912	0.337	0.114
	W/B	/	0.131	0.420	0.216	0.065
	V_f	%	0.500	3.00	1.700	0.500
	L_f	mm	6.000	18.000	15.279	3.069
	D_f	μm	12.000	26.900	23.323	2.118
Outputs	TS	MPa	2.780	18.810	8.447	3.722
	UTS	%	0.990	13.430	5.847	2.497

2.2. Data description

Fig. 1 illustrates the probability density of the mass ratios of various matrix components to binder materials. The density distributions of these mass ratios range from 0 to 1, with distinct peaks, indicating a wide variation in the dataset. This diversity enhances the reliability and generalization ability of the implemented ML models, suggesting that the samples in this study are both diverse and representative. Fig. 2 displays the density distributions of fiber properties, including V_f , L_f and D_f . The fiber volume fraction predominantly clusters around 1% and 2%, while the fiber length mainly concentrates at 12 mm and 18 mm, and the fiber diameter is primarily observed at 18 μm and 24 μm . The statistics presented in Fig. 2 further support the representativeness of the samples, aligning with practical engineering applications.

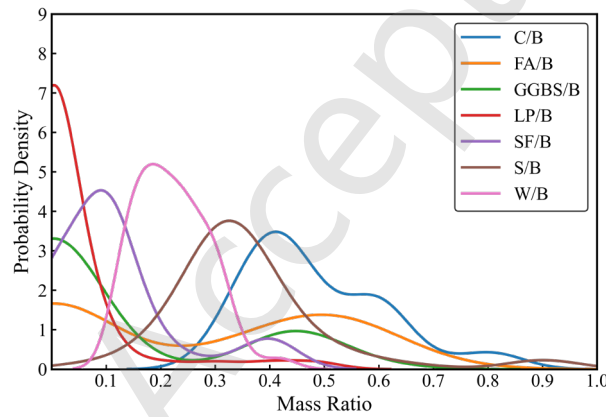


Fig. 1. Density distributions of the mass ratios of various matrix components to binders.

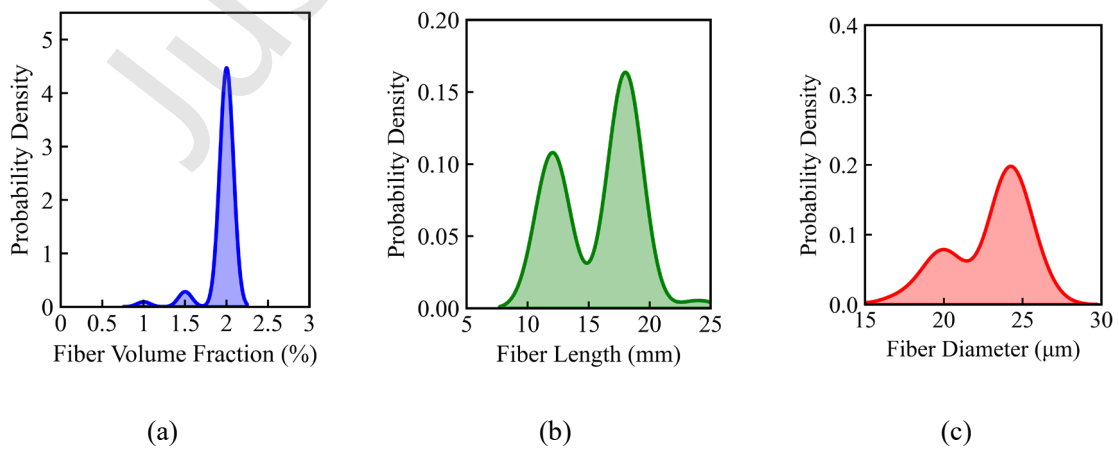


Fig. 2. Density distributions of fiber properties: (a) V_f ; (b) L_f ; (c) D_f .

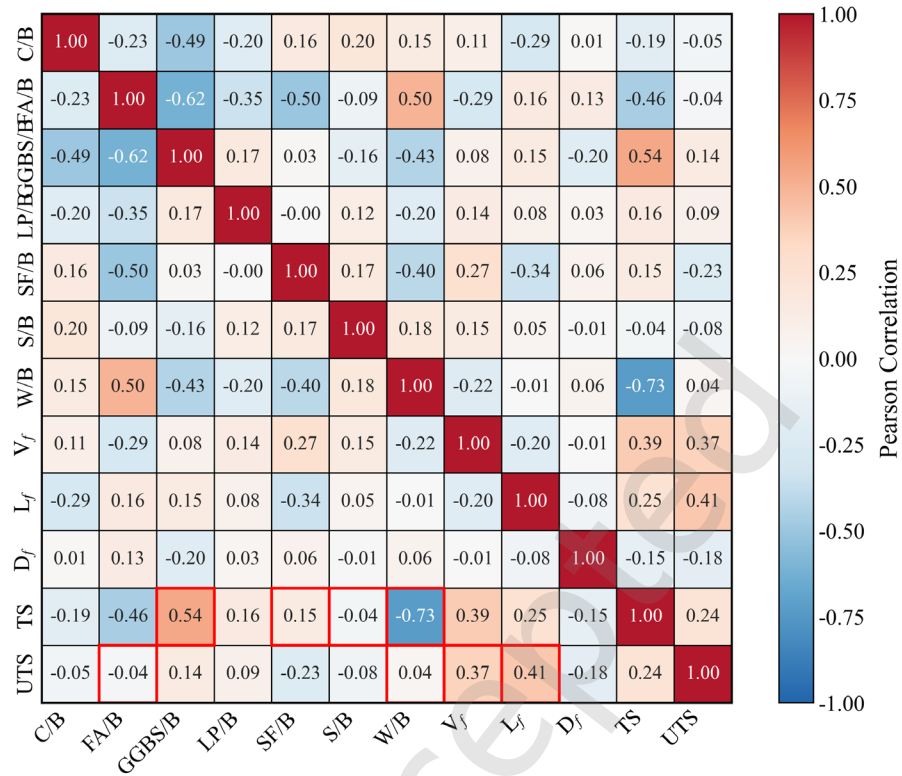


Fig. 3. Correlation heatmap of the features in the dataset.

Fig. 3 presents a heatmap of the dataset features, where the correlation coefficients indicate the strength of linear relationships between each feature. Notably, the W/B significantly influences the tensile strength of PE-ECC, exhibiting a strong negative correlation with a coefficient of -0.73. This finding is consistent with experimental results, which show that a higher W/B ratio leads to lower tensile strength in specimens. Conversely, the GGBS/B positively impacts tensile strength, with a correlation coefficient value of 0.54. For ultimate tensile strain, both V_f and L_f show positive correlations, with R values of 0.37 and 0.41, respectively. This suggests that increasing fiber volume fraction and using longer fibers can enhance ultimate tensile strain. In contrast, the SF/B and D_f negatively affect ultimate tensile strain, with R values of -0.23 and -0.18, respectively, indicating that higher SF content and larger fiber diameter may slightly reduce ultimate tensile strain. Overall, Fig. 3 illustrates that W/B and GGBS/B ratios have significant effects on the tensile strength of PE-ECC, while fiber properties exert a more pronounced influence on ultimate tensile strain performance.

2.3. Data preprocessing

Before training the models, the input data were normalized using min-max normalization, as described in Eq. (1). This process transformed all features into the range [0, 1], which can accelerate and stabilize the training process. Furthermore, min-max normalization reduces the impact of dimensional diversity, ensuring that all input features contribute equally to the model output.

The normalized value X_{nor} is calculated using the following formula:

$$X_{nor} = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (1)$$

where X_{min} and X_{max} denote the minimum and maximum values of the input feature, respectively.

2.4. Cross validation

Cross-validation [90] is an effective technique for mitigating overfitting issues that may arise from a limited dataset by making use of all available data for both training and testing. In this study, a 9-fold cross-validation (CV) method was employed, as illustrated in Fig. 4. The dataset was initially divided into a 9-fold CV set (90%) and a test set (10%). The 9-fold CV set was then further split into 9 folds, with 8 folds utilized for training the ML model, while the remaining fold was reserved for performance validation. During the training phase, the model's hyperparameters were optimized using a Bayesian optimization technique. After each training iteration, evaluation metrics, such as mean squared error (MSE), were computed. This process was repeated nine times to complete the 9-fold CV. Ultimately, the model with the optimal hyperparameters was validated on the test set to assess its predictive performance. This strategic division of the data ensures a fair evaluation on unseen data and enhances the model's generalization capability.

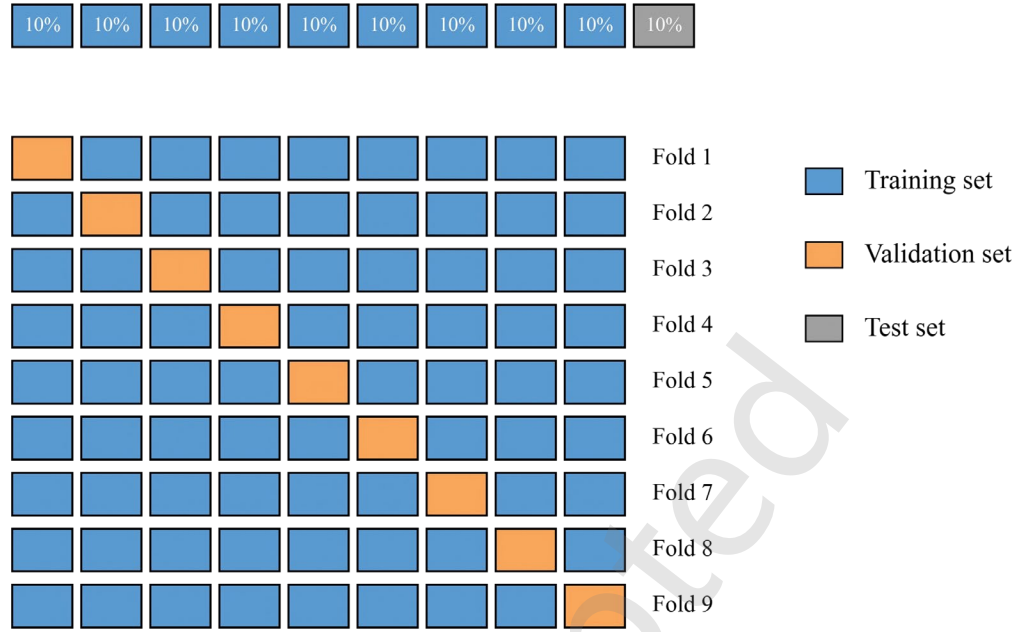


Fig. 4. 9-fold cross validation.

Additionally, early stopping was implemented on the validation dataset to prevent overfitting of the current ML model [91]. This technique halts training if the model's performance does not improve over a specified number of epochs. The predefined threshold for this criterion is referred to as patience. In this study, the maximum number of epochs was set to 2000, with a performance improvement threshold of 0.001 and a patience value of 50. Specifically, if the MSE of the validation set does not decrease by at least 0.001 for 50 consecutive epochs, training will cease, and the weights from the epoch with the best validation performance will be considered the optimal weights for the model.

3. Methodology

This paper proposes a novel ML model for predicting the tensile mechanical properties of PE-ECC using CNN and Bayesian optimization. The proposed model incorporated ten input features and two output targets. The input features included matrix component properties (C/B, FA/B, GGBS/B, LP/B, SF/B, S/B, W/B) and fiber properties (V_f , L_f , D_f). The output targets were tensile strength and ultimate tensile strain. The predictive performance of the model was evaluated using seven evaluation metrics: R , coefficient of determination (R^2), MSE, root mean square error (RMSE), mean absolute error (MAE), mean absolute percentage error (MAPE) and SD. Consequently, the proposed BO-CNN

model was compared with a grid search-optimized CNN (GS-CNN), a grid search-optimized ANN (GS-ANN), and other eleven conventional ML models, such as RF, SVM, and XGBoost. Fig. 5 illustrates the flowchart of the proposed predictive model for the tensile mechanical properties of PE-ECC. This section also focuses on the fundamental principles of the two techniques, CNN and Bayesian optimization, and briefly introduces the principles of other comparative ML methods and the grid search method.

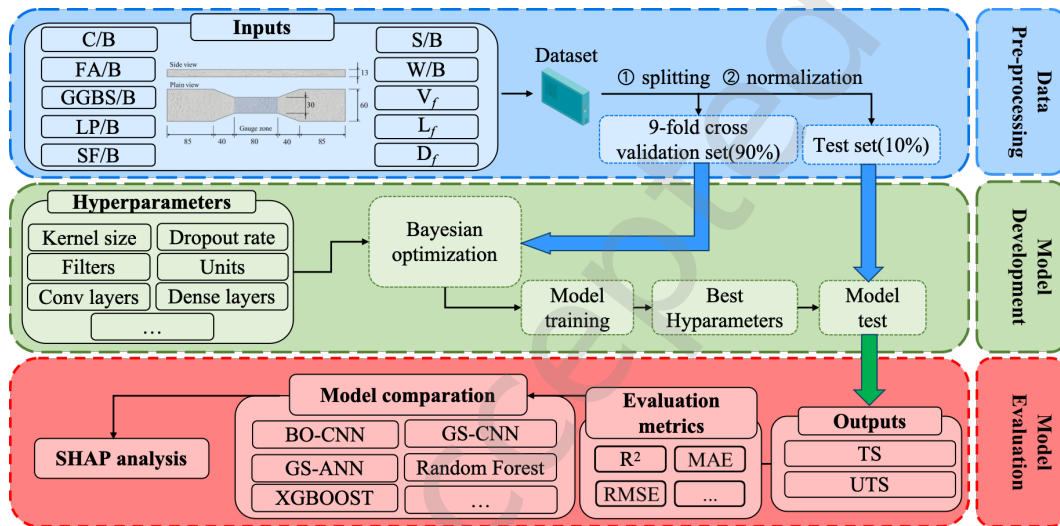


Fig. 5. Flowchart of the proposed predictive model for tensile mechanical properties of PE-ECC.

3.1. Convolutional neural network

As a type of feed-forward neural network, CNN excels at extracting features from data with convolutional structures. Unlike traditional feature extraction methods, CNN eliminates the need for manual feature selection. The success of CNN is largely attributed to its local connectivity and weight sharing, which not only reduce the number of weights but also simplify network optimization while mitigating the risk of overfitting [92].

The most commonly used layers in a CNN include: (1) the input layer; (2) the convolutional layer; (3) rectified linear units (ReLU); (4) the pooling layer and (5) the fully connected layer, as illustrated in Fig. 6. The input layer typically receives input data formatted as matrices (e.g., images), which may have equal dimensions in width and height. Features are extracted by applying convolutional filters to the input matrix.

The core structure of a CNN is the convolutional layer, which is characterized by local receptive fields. In this configuration, each neuron in the convolutional layer connects only to a small region of the input domain. This design allows the CNN to effectively capture spatial correlations between units, thereby enhancing the accuracy of classification or regression tasks. The mathematical representation of the convolution operation is given by:

$$S(i,j)=(I*K)(i,j)=\sum_m\sum_n I(i+m,j+n)\cdot K(m,n) \quad (2)$$

where $S(i,j)$ is the input to the next layer, I represents the input data, K denotes the 2D kernel function or filter matrix, i and j are variables, and m and n the indices within the convolution kernel.

The activation function endows the CNN model with nonlinear capabilities, enabling it to establish a nonlinear relationship between input features and output values. In this study, various activation functions were utilized to enhance the model's ability to fit nonlinear relationships. For example, the ReLU activation function was employed due to its advantages in conserving computational resources and reducing issues like gradient vanishing and overfitting because of its sparse characteristics.

The pooling layer, also known as the subsampling layer, typically follows the convolutional layer and does not involve weights or activation functions. Its primary purpose is to reduce the spatial dimensions of the feature maps through downsampling, thereby decreasing computational complexity. The most prevalent pooling algorithm is max pooling, which selects regions from the input feature map, retaining the maximum value within each region while discarding the other values.

Finally, the output from the last convolutional or pooling layer is transformed into a one-dimensional vector. This vector is then connected to fully connected layers containing learnable weights, linking the input and output vectors. The feature vectors generated by the convolutional or pooling layers are mapped to the final output of the model through a subset of fully connected layers, which yield either classification probabilities or regression values for different learning tasks. In regression tasks, the final fully connected layer typically has a single output node with an activation

function, allowing the model to predict continuous values. The activation function is applied after each hidden layer to enhance the model's capacity to capture nonlinearities.

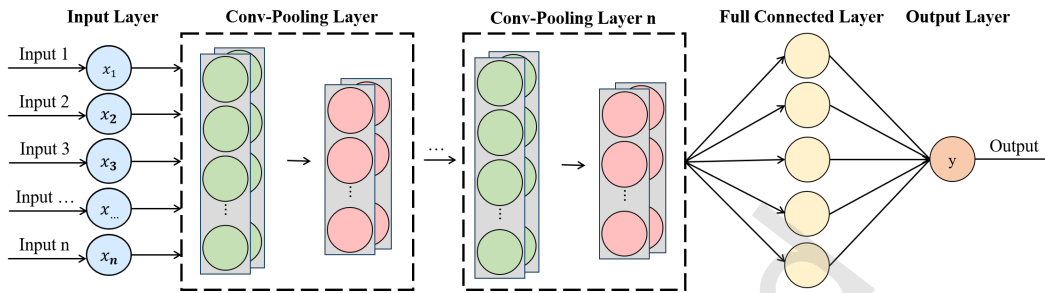


Fig. 6. Architecture of CNN model.

3.2. Bayesian Optimization

When employing CNN for regression tasks, configuring hyperparameters such as the number of hidden layers, the number of neurons, and the learning rate is critical before training. Typically, researchers establish these hyperparameters based on empirical evidence, which, while cost-effective, can often result in suboptimal model performance. Bayesian optimization is a powerful global optimization algorithm designed to identify the optimal solution within the parameter space [93, 94]. Mathematically, this process involves seeking the global optimum of an unknown objective function [95]:

$$z^* = \arg \max_{z \in Z} f(z) \quad (3)$$

where z^* denotes the parameters being optimized, Z represents the set of all parameters under consideration, $f(z)$ is the objective function to be optimized, and z is a parameter vector in the parameter space Z . Bayesian optimization typically requires two key components to function effectively.

First, the prior function's form is established based on the assumed distribution of the objective function. A Gaussian process is commonly selected for this purpose due to its flexibility and tractability. Second, the construction of an acquisition function is crucial, as it determines the next point to explore in the parameter space, utilizing the updated posterior distribution to guide the decision-making process.

In practice, the selection of subsequent sampling points in Bayesian optimization considers both exploitation and exploration. Exploitation focuses on searching around the currently identified optimal solution to refine it, while exploration targets regions of the parameter space that have not yet been assessed, thus avoiding entrapment in local optima. Common acquisition functions include the Expected Improvement (EI), Upper Confidence Bound (UCB), and Probability of Improvement (PI) functions. For this research, the UCB function was chosen as the primary sampling strategy due to its effectiveness in balancing exploitation and exploration. The mathematical formulation of the UCB function is given by:

$$UCB(z)=\mu(z)+\varepsilon(z) \quad (4)$$

where $\mu(z)$ and $\varepsilon(z)$ are the mean and standard deviation of the joint posterior distribution of the objective function obtained using a Gaussian process, respectively.

From Eq. 4, it is evident that by adjusting the magnitude of the parameter ε , the trade-off between exploration and exploitation can be controlled effectively, thus guiding the selection of sampling points in the parameter space.

3.3. Grid search and other comparative ML models

To evaluate the effectiveness of CNNs and Bayesian optimization techniques within the BO-CNN framework for predicting the tensile mechanical properties of PE-ECC, we also employed the classic grid search (GS) technique and compared the results with those of eleven other ML models. GS is a widely utilized global optimization technique designed to search for optimal solutions within the hyperparameter space [96]. A key advantage of GS lies in its simplicity and intuitiveness, which facilitates straightforward implementation. It ensures that each parameter combination is evaluated fairly, thereby guaranteeing the identification of optimal solutions. However, this approach can be particularly time-consuming when dealing with a large hyperparameter space or when training each model is costly. Despite its limitations in efficiently managing high-dimensional hyperparameter spaces, GS remains a prevalent parameter optimization technique in the field of ML, especially suitable for scenarios with a relatively small number of hyperparameters.

For comparative analysis, the comparison set comprised eleven conventional ML models, including LR, ridge regression (RR), LASSO, DT, RF, SVR, k-nearest neighbors (kNN), GB, XGBoost, Light Gradient Boosting Machine (LightGBM), and Categorical Boosting (CatBoost), together with GS-ANN and GS-CNN. All tunable comparison models were optimized using grid search, whereas the proposed BO-CNN was optimized using Bayesian optimization. Thus, the benchmark covers a progressive range of models from conventional machine learning to ANN and CNN architectures. In particular, GS-ANN provides a grid-search-optimized MLP benchmark, while XGBoost, LightGBM, and CatBoost provide strong modern gradient-boosting benchmarks. The theoretical foundations and algorithmic details of these models can be found in the relevant literature [97-103].

3.4. Model performance evaluation metrics

Seven evaluation metrics were selected to assess the performance of the ML model. These metrics are R , R^2 , MSE, RMSE, SD, MAE and MAPE. The values of R and R^2 indicate how well the predicted values align with the observed outputs, with higher values signifying a better goodness of fit. Conversely, MSE, RMSE, MAE, and MAPE quantify the discrepancies between predicted and experimental values, with lower values indicating higher prediction accuracy.

$$R = \frac{\sum_{i=1}^n (Y_{pre} - \overline{Y_{pre}}) \cdot (Y_{exp} - \overline{Y_{exp}})}{\sqrt{\sum_{i=1}^n (Y_{pre} - \overline{Y_{pre}})^2} \cdot \sqrt{\sum_{i=1}^n (Y_{exp} - \overline{Y_{exp}})^2}} \quad (5)$$

$$R^2 = \frac{\sum_{i=1}^n (Y_{pre} - \overline{Y_{exp}})^2}{\sum_{i=1}^n (Y_{exp} - \overline{Y_{exp}})^2} \quad (6)$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (Y_{pre} - Y_{exp})^2 \quad (7)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (Y_{pre} - Y_{exp})^2} \quad (8)$$

$$SD = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (Y_{exp} - \overline{Y_{exp}})^2} \quad (9)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |Y_{pre} - Y_{exp}| \quad (10)$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n \frac{|Y_{pre} - Y_{exp}|}{Y_{exp}} \cdot 100 \quad (11)$$

where n represents the total number of samples, Y_{exp} denotes the experimental values, and Y_{pre} refers to the predicted values.

3.5. Hyperparameter tuning

To fully leverage the advantages of Bayesian optimization for network construction and automatic hyperparameter tuning, this study meticulously defined the model's search space to enhance the predictive performance of the BO-CNN. Specifically, Bayesian optimization was employed to explore key architectural and training hyperparameters, including the number of convolutional layers, the number of kernels per layer, and the kernel sizes, thereby facilitating more effective feature extraction. Furthermore, hyperparameters such as the dropout rate were adjusted to mitigate overfitting and enhance the model's generalization capabilities. The optimized hyperparameters for this model are summarized in Table 2, while the optimal parameters identified through Bayesian optimization are detailed in Appendix A and B.

Table 2. The Bayesian optimized hyperparameters in the BO-CNN model.

Hyperparameter	Range	Hyperparameter	Range
Convolutional layer number	1 to 5	Number of neurons of each fully connected layer	1 to 256
		Activation function of each fully connected layer	ReLU; Leaky ReLU
Number of filters of each convolutional layer	1 to 256	Regularization type for each fully connected layer	L1; L2; L1 and L2; None

Activation function of		Activation function of	
each convolutional layer	ReLU; Leaky ReLU	each fully connected layer	ReLU; Leaky ReLU
Batch normalization		Batch normalization	
for each convolutional layer	Yes; No	for each fully connected layer	Yes; No
Spatial down sampling		Dropout rate of each fully connected layer	
type of each convolutional layer	Pooling layer; Transposed convolution layer; None		0 to 0.5 (step size 0.05)
Transformation layer type	Flatten layer; Global pooling layer	Optimizer type	Adam; RMSprop; SGD
Number of fully connected layers	1 to 5 layers	Learning rate for optimizer	0.0001 to 1

Using the BO-CNN model to predict tensile strength as an example, the optimization process is illustrated in Fig. 7. The model's initial performance was unstable. However, as the optimization progressed, the performance gradually improved and eventually stabilized. As shown in Fig. 7, the best performance was achieved at the 137th trial, where the test set attained a R^2 value of 0.911 and an MSE value of 0.500. Furthermore, the cumulative mean of R^2 converged and remained stable after approximately 800 to 1000 trials, indicating that Bayesian optimization had successfully identified the optimal hyperparameter configuration and that further exploration of the parameter space ceased to yield additional improvements.

Using the BO-CNN model for tensile strength as an example, Fig. 7 illustrates the hyperparameter optimization process. For each trial, the validation R^2 values were calculated for the nine folds and then averaged. The hyperparameter configuration obtained at the 137th trial achieved the highest mean validation R^2 and was therefore selected as the optimal configuration. Using this

configuration, the resulting BO-CNN models achieved an average test-set R^2 of 0.911 and an average MSE of 0.500. Fig. 7 presents the cumulative average of the nine-fold mean validation R^2 over the optimization trials. This cumulative average gradually stabilized after approximately 800 to 1000 trials, indicating that the overall performance of the optimization process had become stable.

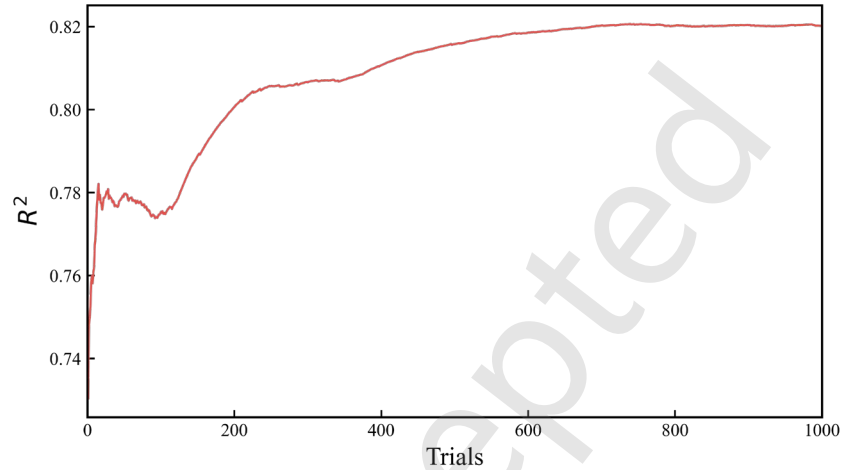


Fig. 7. Bayesian optimization process of cumulative average of nine-fold mean validation R^2 .

The optimization process for selected hyperparameters is depicted in Fig. 8. As shown in Figs. 8(a) and 8(c), the number of convolutional layers converged towards 2, while the number of fully connected layers approaches 5, indicating these configurations yield higher R^2 values compared to other settings. For the transformation layer type, Fig. 8(b) illustrates that the model ultimately selected the flatten layer, which demonstrated stable performance within the CNN architecture and retains the features processed by the convolutional layers more effectively than global pooling layers. Regarding the optimizer, Fig. 8(d) shows that the model favors RMSProp, as it consistently provides excellent and stable performance while minimizing computational resources in most scenarios. Overall, it is evident from Fig. 8 that the hyperparameter values were relatively dispersed during the initial trials. However, in subsequent trials, these values gradually converged in the region that maximized the objective function. As an informed search method, Bayesian optimization continuously learns from prior optimizations, facilitating the rapid identification of an optimal base ML model with enhanced predictive accuracy while reducing human intervention and computational costs.

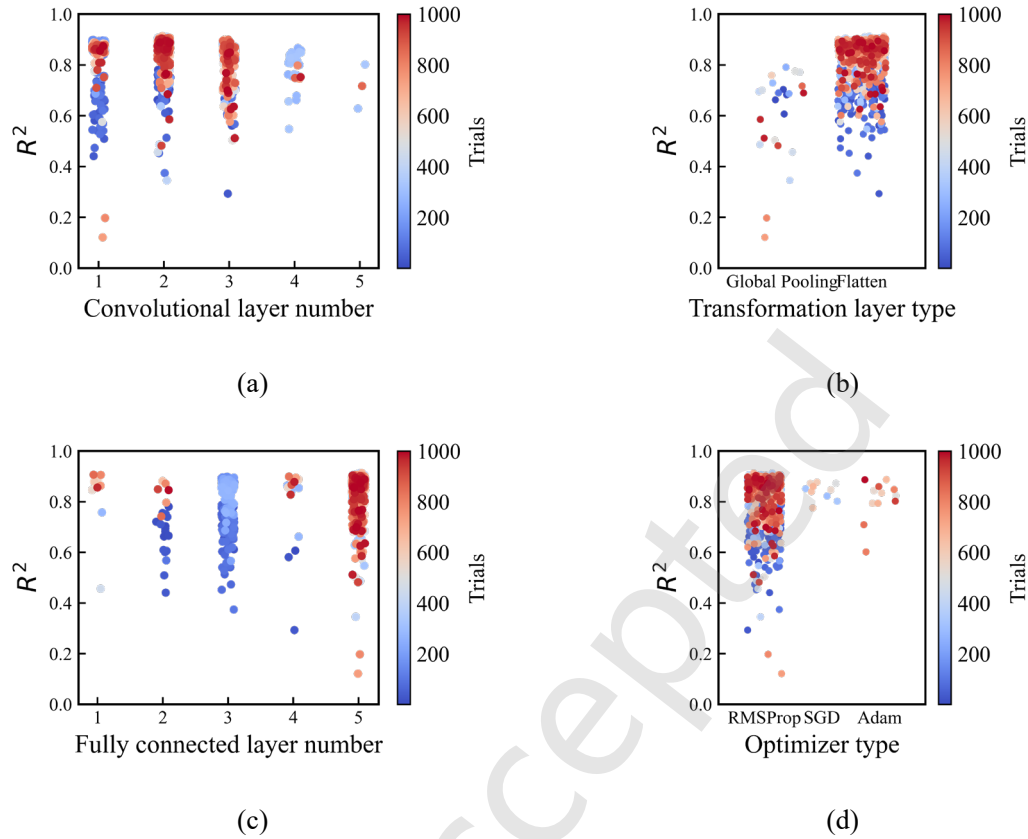


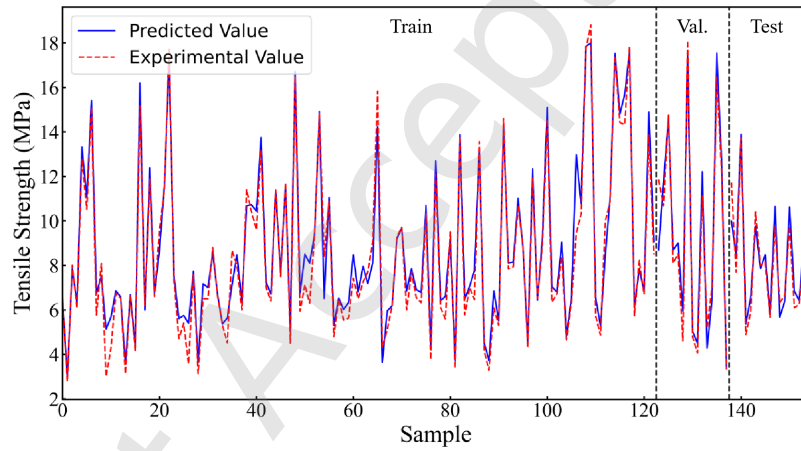
Fig. 8. Bayesian optimization process of each hyperparameter: (a) convolutional layer number; (b) transformation layer type; (c) fully connected layer number; (d) optimizer type.

4. Results and discussions

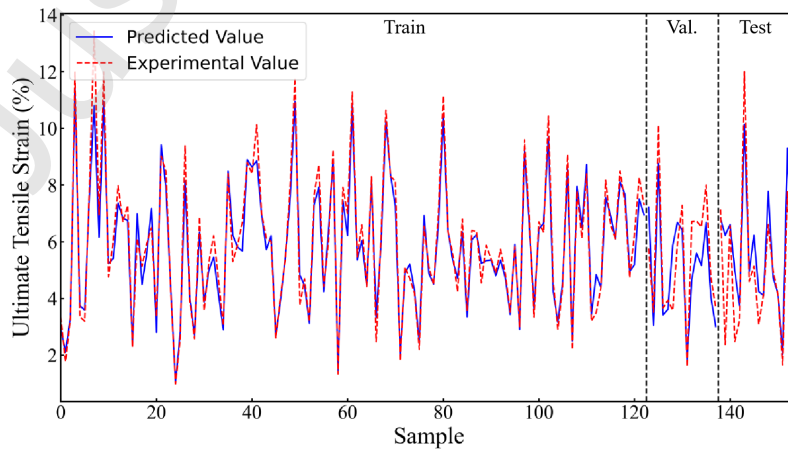
4.1. BO-CNN model performance evaluation

This section provides a comprehensive analysis of the prediction results from the optimized BO-CNN model on each set of the ninth fold, aimed at evaluating its predictive accuracy and generalization capabilities. During training, both the training loss and validation loss steadily decreased and converged toward zero, indicating that the model was free from underfitting. Meanwhile, the validation loss remained consistently above the training loss without divergence, confirming the absence of overfitting. Early stopping was adopted to terminate training once the validation loss failed to improve for 50 consecutive epochs, which effectively mitigated overfitting risk and conserved computational resources.

Figs. 9(a) and 9(b) compare the experimental values with the predicted values obtained from the BO-CNN models. Table 3 demonstrates that the model predicts PE-ECC tensile strength with a high degree of accuracy ($R^2 = 0.914$, MAPE = 0.074). However, the model's accuracy for ultimate tensile strain is lower, with an R^2 value of 0.759 and a MAPE value of 0.279. Ultimate tensile strain is more sensitive to factors such as fiber distribution and orientation, fiber – matrix interfacial behavior, microcrack evolution, and variations in casting and loading processes, and consequently exhibits greater data dispersion. Compared with the model developed by Altayeb et al. [104], which achieved an R^2 value of 0.61 for ultimate tensile strain, the BO-CNN model can therefore be regarded as reliable in predicting the ultimate tensile strain of ECC.



(a)



(b)

Fig. 9. Performance of the BO-CNN model in predicting PE-ECC: (a) tensile strength; (b) ultimate tensile strain.

Table 3. R^2 and MAPE values of the BO-CNN model for the test set predictions.

Output feature	R^2	MAPE
Tensile strength	0.914	0.074
Ultimate tensile strain	0.759	0.279

To further evaluate the reliability of the BO-CNN predictions, Fig. 10 presents the error distributions for the training, validation, and test subsets. For tensile strength, most prediction errors are concentrated near zero, and the training and test subsets exhibit similar error standard deviations of 0.685 and 0.695 MPa, respectively. Although the validation subset has a relatively larger standard deviation of 1.085 MPa, the overall standard deviation remains 0.737 MPa. For ultimate tensile strain, the prediction errors are also predominantly distributed around zero, with an overall standard deviation of 0.753 percentage points. Despite some variation among the three subsets, the concentrated error distributions and limited overall standard deviations demonstrate that the BO-CNN model maintains stable error characteristics and reliable predictive performance for both tensile strength and ultimate tensile strain, supporting its application in predicting the tensile mechanical properties of PE-ECC.

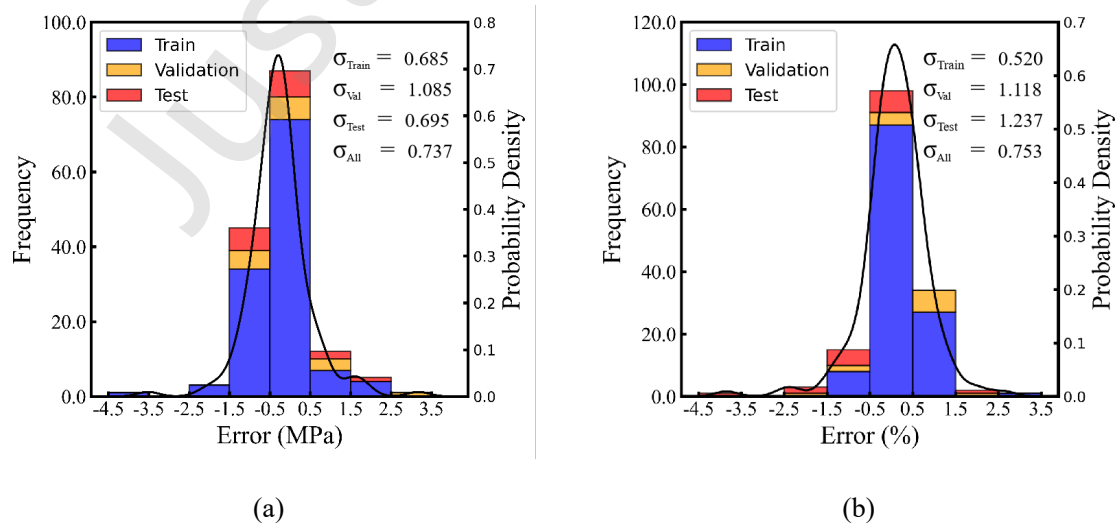


Fig. 10. Error distribution histograms for BO-CNN model predictions: (a) tensile strength; (b) ultimate

tensile strain.

4.2. Comparative assessment of the BO-CNN model and other neural network models optimized by GS techniques

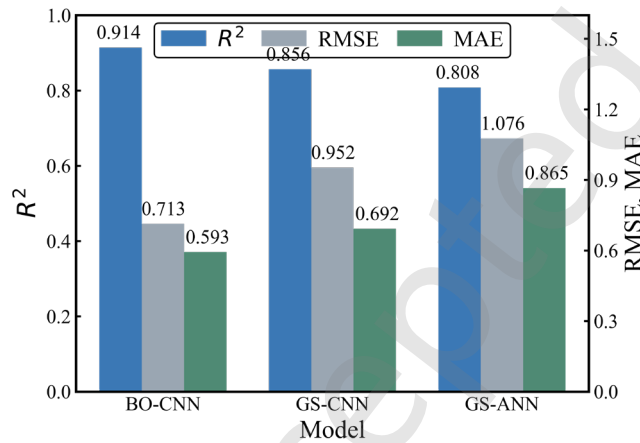
This section evaluates the effectiveness of CNN and Bayesian optimization technique within the BO-CNN framework for predicting the tensile mechanical properties of PE-ECC. For comparison, the classic GS method and ANN were introduced. The GS method is a widely used hyperparameter optimization technique, while ANN is a model commonly employed for predicting the mechanical properties of concrete. In this study, the GS-CNN and GS-ANN were used as comparative models to assess the prediction accuracy of three models, including BO-CNN, for both tensile strength and ultimate tensile strain.

The predicted values generated by the BO-CNN model closely align with the experimental values, indicating superior predictive performance. Fig. 11 presents the evaluation metrics (R^2 , RMSE, MAE) for predicting tensile strength and ultimate tensile strain across the models. For tensile strength, as shown in Fig. 11(a), the R^2 values for BO-CNN, GS-CNN, and GS-ANN are 0.91, 0.86 and 0.81, respectively, demonstrating that BO-CNN achieves the highest accuracy, followed by GS-CNN and GS-ANN. The RMSE values for these models are 0.71, 0.95 and 1.08, while the MAE values are 0.59, 0.69 and 0.86, indicating that BO-CNN exhibits the smallest prediction bias and error, with GS-CNN following and GS-ANN lagging behind.

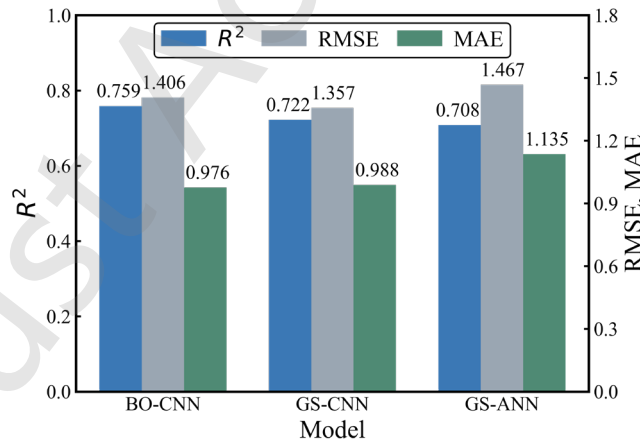
These results indicate that, in predicting the tensile properties of PE-ECC, the CNN architecture was more effective than the ANN architecture in processing the ten input features and capturing local relationships among them, thereby achieving better predictive performance.

Furthermore, the findings indicate that Bayesian optimization techniques outperform traditional grid search methods, which is particularly significant for neural network models, where hyperparameters greatly influence performance. Exhaustive searches within limited hyperparameter spaces may fail to fully exploit the model's potential. In contrast, Bayesian optimization allows for the identification of optimal hyperparameters within a vast search space, thereby enhancing predictive performance.

Similar conclusions can be drawn for ultimate tensile strain, as illustrated in Fig. 11(b). Notably, predicting ultimate tensile strain involves a higher degree of data dispersion, further emphasizing the advantages of the proposed BO-CNN model. The BO-CNN model improves R^2 by 5.12% and 7.20% compared to GS-CNN and GS-ANN, respectively.



(a)



(b)

Fig. 11. Evaluated metrics of BO-CNN, GS-CNN, and GS-ANN models on the test set for PE-ECC:

(a) tensile strength; (b) ultimate tensile strain.

4.3. Comparative assessment of the BO-CNN model and other ML models

This section explores the advantages of the BO-CNN model compared to other ML models. The benchmark follows a progressive hierarchy from grid-search-optimized conventional ML models to

GS-ANN and GS-CNN, allowing BO-CNN to be evaluated against representative traditional machine-learning, ANN, and CNN configurations. Table 4 summarizes the predictive performance of various models for predicting tensile strength on the test set, evaluated using three metrics: R^2 , RMSE, and MAPE. As shown in Table 4, among all evaluated ML models, the BO-CNN achieves the highest accuracy, with an R^2 value of 0.911. Additionally, it exhibits the lowest RMSE and MAPE values, recorded at 0.701 and 0.066, respectively, indicating that this model has the smallest prediction bias and error. The CNN architecture employed in the BO-CNN framework serves as a representative DL model, adept at extracting intricate features related to target outputs and effectively capturing complex nonlinear relationships between multiple input features and tensile strength in PE-ECC. The Bayesian optimization technique applied in the BO-CNN model automates hyperparameter tuning, enhancing the network's performance and achieving optimal predictive accuracy.

Other ML models, such as GS-CNN and RF, also demonstrate commendable predictive accuracies, with R^2 values exceeding 0.8. Conversely, models such as kNN, SVR and LR exhibit poorer performance, with R^2 values falling below 0.7. Specifically, the kNN determines output values based on distances to the k nearest samples, which may not be appropriate for this particular problem. Moreover, the selected SVR configuration may not fully capture the complex nonlinear relationships in the present dataset, and training SVR can be computationally intensive. LR, by assuming a linear relationship between variables, limits its ability to model nonlinear problems.

As indicated in Table 4, the prediction performance for ultimate tensile strain is generally inferior to that for tensile strength across all models. This discrepancy can primarily be attributed to the sensitivity associated with the casting and loading processes of ECC, resulting in higher data variability for ultimate tensile strain and consequently increasing the difficulty of model predictions. Nonetheless, in the context of predicting ultimate tensile strain, which is characterized by high data dispersion and more complex nonlinear relationships, the BO-CNN model achieves the highest predictive accuracy ($R^2 = 0.775$) and the lowest prediction errors (RMSE = 1.174, MAPE = 0.238) compared to other ML models.

These figures also clearly demonstrate that BO-CNN consistently delivers superior predictive performance, whether predicting tensile strength or ultimate tensile strain. From a modeling perspective, BO-CNN exhibits more pronounced advantages over traditional models (such as kNN, LR and SVR) than over neural networks (like GS-ANN and GS-CNN), as evidenced by the greater increase in the R^2 metric and the smaller reductions in RMSE and MAPE. Notably, the advantages of BO-CNN are particularly evident in predicting ultimate tensile strain, which is characterized by higher data variability. This implies that as problem complexity increases, the suitability of the BO-CNN model becomes increasingly apparent.

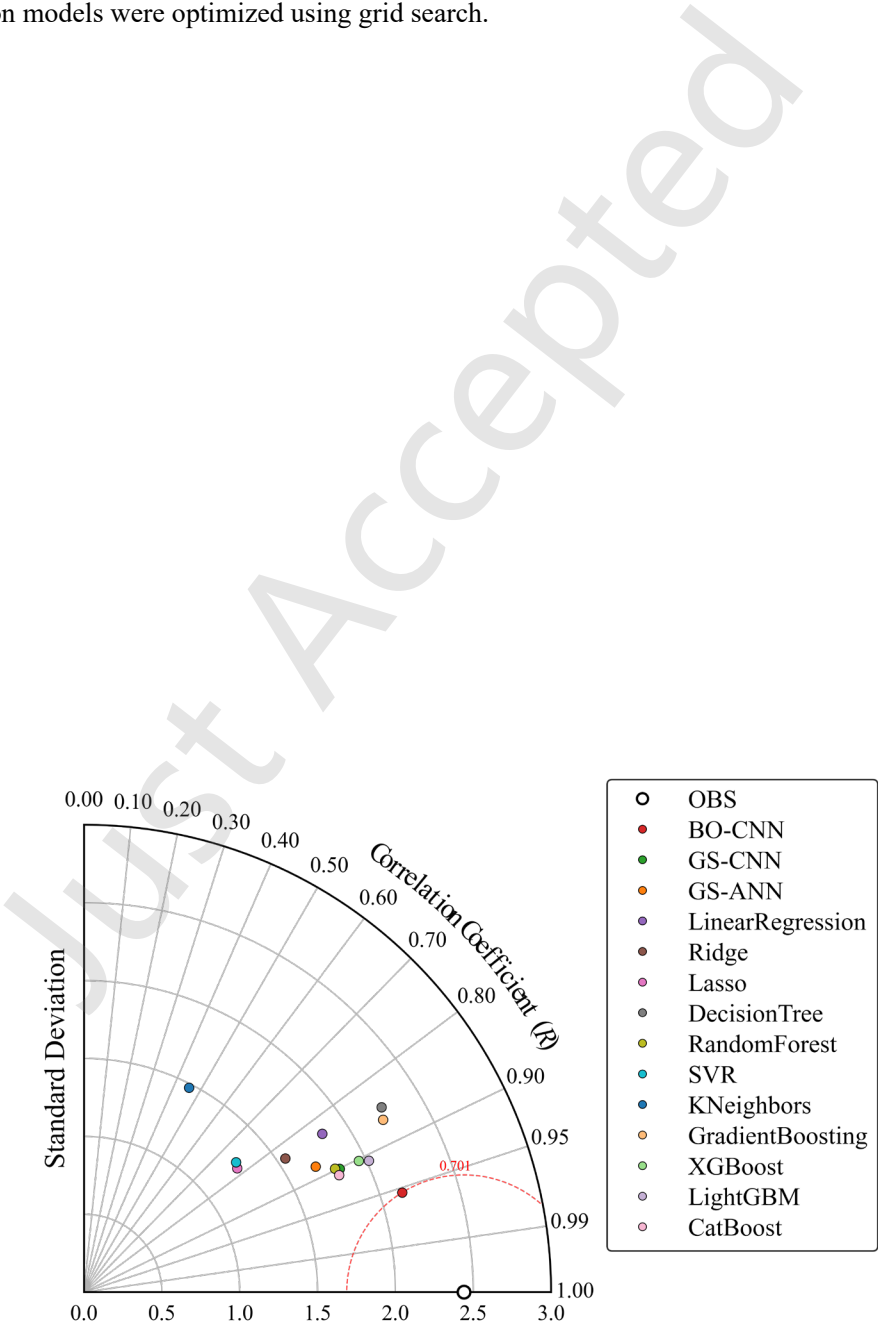
Table 4. Performance evaluation metrics of models for tensile properties on the test set.

Models	Tensile strength	Tensile strength	Tensile strength	Ultimate tensile strain	Ultimate tensile strain	Ultimate tensile strain
Models	R^2	RMSE	MAPE	R^2	RMSE	MAPE
BO-CNN	0.911	0.701	0.066	0.775	1.174	0.238
GS-CNN	0.811	1.024	0.110	0.715	1.307	0.257
GS-ANN	0.773	1.120	0.124	0.698	1.364	0.248
Linear Regression	0.694	1.309	0.147	0.105	2.366	0.511
Ridge	0.694	1.309	0.160	0.133	2.330	0.518
Lasso	0.605	1.487	0.186	0.158	2.297	0.526
Decision Tree	0.722	1.229	0.120	0.250	2.153	0.417
Random Forest	0.805	1.039	0.102	0.534	1.705	0.371
SVR	0.578	1.536	0.156	0.302	2.088	0.472
KNeighbors	0.209	2.103	0.238	0.223	2.206	0.518
Gradient	0.751	1.166	0.110	0.268	2.127	0.416

Boosting

XGBoost	0.814	1.015	0.098	0.588	1.603	0.338
LightGBM	0.825	0.988	0.087	0.578	1.623	0.386
CatBoost	0.826	0.985	0.103	0.504	1.763	0.397

Note: The proposed BO-CNN was optimized using Bayesian optimization, whereas all tunable comparison models were optimized using grid search.



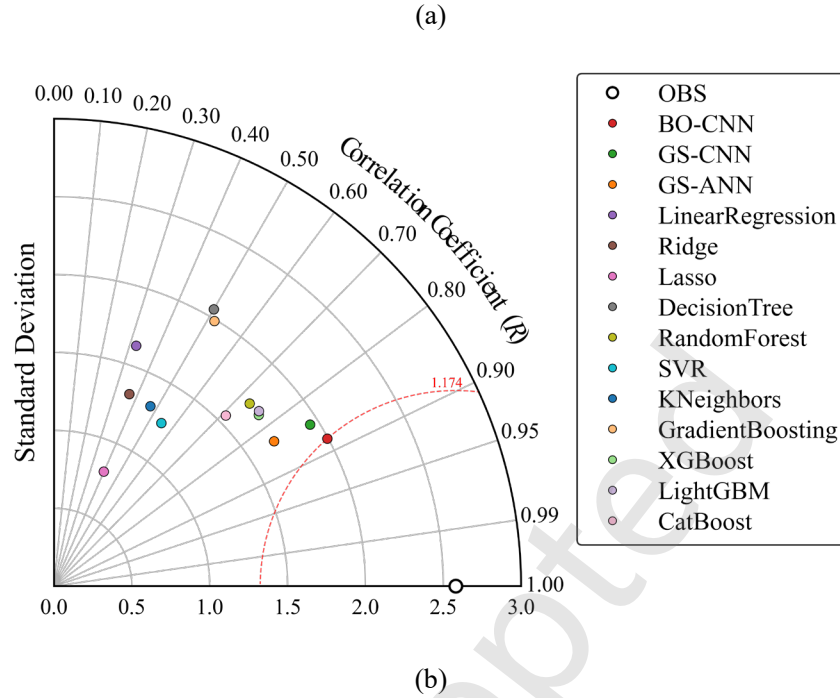


Fig. 12. Taylor diagrams of BO-CNN and other ML models for PE-ECC property prediction: (a) tensile strength; (b) ultimate tensile strain.

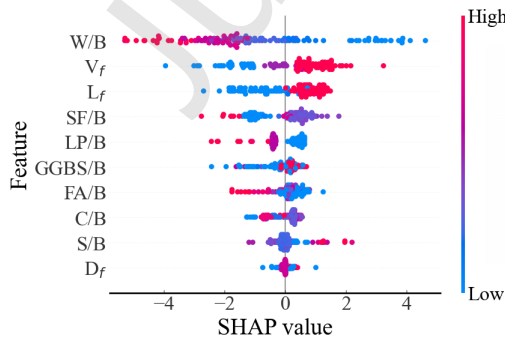
4.4. Interpretability of the BO-CNN models

As a neural network model, the BO-CNN often features more complex and non-linear architectures, resulting in its classification as a black-box model. The hierarchical structure of models makes visual interpretation challenging. To address this, SHAP [105] was employed to elucidate the workings of models, especially DL architectures with numerous parameters. SHAP uses Shapley values to quantify the contribution of each feature across all possible feature combinations, thereby identifying the features that influence specific model predictions and assisting users in understanding the decision-making process. Given that the BO-CNN demonstrates superior performance in predicting both the tensile strength and ultimate tensile strain of PE-ECC, we focused our interpretation on these metrics using SHAP.

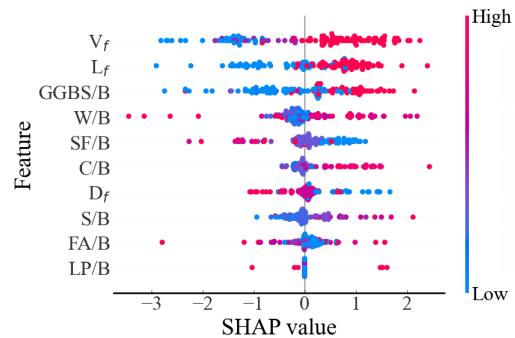
Fig. 13 presents the SHAP values associated with various features for the predictions of tensile strength and ultimate tensile strain as derived from the BO-CNN model. In the SHAP summary plots, features are ordered by their mean absolute SHAP magnitude. The horizontal position indicates the direction and magnitude of the SHAP value, while the color gradient represents the normalized feature

value, with red and blue denoting high and low feature values, respectively. As shown in Fig. 13(a), W/B has the largest global SHAP magnitude in tensile strength predictions. Specifically, an increase in W/B is associated with a decrease in predicted tensile strength, while lower W/B values correspond to higher tensile strength, indicating a negative correlation between W/B and tensile strength. This finding is consistent with previous experimental studies [106]. Physically, this trend is consistent with the denser and stronger matrix generally formed at a lower water-to-binder ratio. These results affirm that the BO-CNN model effectively captures the critical impact of W/B on tensile strength, yielding predictions that closely match empirical data. Similarly, Fig. 13(b) shows that V_f and fiber length L_f have the largest mean absolute SHAP magnitudes for ultimate tensile strain, followed by GGBS/B and W/B. Their relatively broad SHAP distributions indicate that the predicted ultimate tensile strain is particularly sensitive to variations in fiber content, fiber geometry, and matrix composition. This trend is physically consistent with the roles of fiber content and fiber length in fiber bridging and multiple cracking. This observation aligns with prior research [107], highlighting the significance of fiber properties in determining the tensile mechanical properties of fiber-reinforced ECC.

Overall, SHAP visualizations provide valuable insights into the features driving model predictions. By comparing the magnitude and direction of the feature contributions, the analysis clarifies how individual input variables are associated with the predicted tensile properties of PE-ECC. These findings enhance the interpretability of the model and may provide useful guidance for future material-design and optimization studies.



(a)



(b)

Fig. 13. SHAP values for the BO-CNN model's predictions of PE-ECC: (a) tensile strength; (b) ultimate tensile strain.

5. Conclusions

This study aimed to develop a high-precision, fully automated prediction model for the tensile strength and ultimate tensile strain of PE-ECC, eliminating the need for manual intervention. The proposed model was based on a CNN architecture and combining with Bayesian optimization to automate the modeling process and optimize hyperparameters, thereby enhancing both modeling efficiency and prediction accuracy. The model's performance was evaluated by comparing its accuracy and generalization capabilities against GS-CNN, GS-ANN, and other eleven traditional ML models across multiple evaluation metrics. Additionally, SHAP was employed to enhance the interpretability of the BO-CNN model's predictions. The main conclusions of this study are as follows:

The proposed BO-CNN model outperforms other thirteen models, achieving the highest prediction accuracy and the lowest prediction error. For tensile strength prediction, the R^2 , RMSE, and MAPE values are 0.911, 0.701, and 0.066, respectively. For ultimate tensile strain prediction, these values are 0.775, 1.174, and 0.238. Notably, the BO-CNN model shows significant advantages in predicting ultimate tensile strain, which is characterized by higher data dispersion, compared to other models.

The BO-CNN model also surpasses traditional GS-optimized neural network models, including GS-CNN and GS-ANN. For tensile strength prediction, the accuracy of the BO-CNN model improves by 12.28% over GS-CNN and by 17.82% over GS-ANN. For ultimate tensile strain prediction, it shows an accuracy improvement of 8.32% compared to GS-CNN and 10.99% compared to GS-ANN. Since hyperparameters significantly influence neural network performance, exhaustive searches within limited hyperparameter spaces may not fully exploit the model's potential. Utilizing Bayesian optimization allows for the identification of optimal hyperparameters within a virtually infinite search space, ensuring optimal model performance.

SHAP analysis of the BO-CNN model revealed the contribution of each input feature to the model's predictions, illustrating how output values vary with changes in each feature. The analysis indicates that the W/B , V_f and L_f have the most significant impacts on PE-ECC tensile mechanical

properties, aligning with existing experimental findings. This interpretability enhances the model's practical applicability and provides valuable insights for material design and optimization.

Overall, the proposed BO-CNN framework provides an accurate, automated, and interpretable approach for predicting the tensile properties of PE-ECC. Future work will extend the framework to mixture proportion optimization to facilitate the performance-based design of PE-ECC with desired tensile properties.

Acknowledgments

This work was supported in part by Postdoctoral Research Foundation of China (2021M700886), Guangzhou Science and Technology Project (SL2023A04J01633), Guangdong Provincial Key Laboratory of Intelligent Disaster Prevention and Emergency Technologies for Urban Lifeline Engineering (2022B1212010016).

Author contribution statement

“Author 1 is responsible for Work 1. ...All the authors have approved the final manuscript.” CRediT author statement (<https://credit.niso.org/>) is recommended to use, and all the listed authors should meet the definition of Author in ICMJE recommendations (<https://www.icmje.org/>).

Data availability

“Not applicable.”

OR “The data that support the findings of this study could be downloaded from xxxxxxxx / is available from the corresponding author upon reasonable request.”

Use of AI statement

“None.”

OR “During the preparation of this work, the author(s) used [NAME TOOL / SERVICE] in order to [REASON]. After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the publication.”

Declaration of competing interest

The author(s) have no competing interests to declare that are relevant to the content of this article.

References

- [1] V.C. Li, On Engineered Cementitious Composites (ECC), *J. Adv. Concr. Technol.* 1 (2003) 215–230.
- [2] C. Wu, V.C. Li, CFRP-ECC hybrid for strengthening of the concrete structures, *Compos. Struct.* 178 (2017) 372–382.
- [3] C. Wu, V.C. Li, Thermal-mechanical behaviors of CFRP-ECC hybrid under elevated temperatures, *Compos. Part B Eng.* 110 (2017) 255–266.
- [4] Z. Lin, V.C. Li, Crack bridging in fiber reinforced cementitious composites with slip-hardening interfaces, *J. Mech. Phys. Solids* 45 (1997) 763–787.
- [5] V.C. Li, C.K.Y. Leung, Steady - State and Multiple Cracking of Short Random Fiber Composites, *J. Eng. Mech.* 118 (1992) 2246 – 2264.
- [6] K. Yu, L. Li, J. Yu, Y. Wang, J. Ye, Q. Xu, Direct tensile properties of engineered cementitious composites: A review, *Constr. Build. Mater.* 165 (2018) 346–362.
- [7] Y. Huo, T. Liu, D. Lu, X. Han, H. Sun, J. Huang, X. Ye, C. Zhang, Z. Chen, Y. Yang, Dynamic tensile properties of steel fiber reinforced polyethylene fiber-engineered/strain-hardening cementitious composites (PE-ECC/SHCC) at high strain rate, *Cem. Concr. Compos.* 143 (2023) 105234.
- [8] J. Zhou, J. Pan, C.K.Y. Leung, Mechanical Behavior of Fiber-Reinforced Engineered Cementitious Composites in Uniaxial Compression, *J. Mater. Civ. Eng.* 27 (2015) 04014111.
- [9] N. Shanmugasundaram, S. Praveenkumar, Mechanical properties of engineered cementitious composites (ECC) incorporating different mineral admixtures and fibre: a review, *J. Build. Pathol. Rehabil.* 7 (2022) 40.
- [10] J.-X. Lin, R.-H. Luo, J.-Y. Su, Y.-C. Guo, W.-S. Chen, Coarse synthetic fibers (PP and POM) as a replacement to steel fibers in UHPC: Tensile behavior, environmental and economic assessment, *Constr. Build. Mater.* 412 (2024) 134654.
- [11] J.-X. Lin, R.-A. Liu, L.-Y. Liu, K.-Y. Zhuo, Z.-B. Chen, Y.-C. Guo, High-strength and high-toughness alkali-activated composite materials: Optimizing mechanical properties through synergistic

utilization of steel slag, ground granulated blast furnace slag, and fly ash, *Constr. Build. Mater.* 422 (2024) 135811.

[12] M. Zhu, B. Chen, M. Wu, J. Han, Effects of different mixing ratio parameters on mechanical properties of cost-effective green engineered cementitious composites (ECC), *Constr. Build. Mater.* 328 (2022) 127093.

[13] S.-J. Chen, R.-A. Liu, X.-Y. Huang, Z.-M. Yang, J.-X. Lin, PET particles modify strain hardening cementitious composites: An approach to introduce defects to enhance deformation capacity, *Constr. Build. Mater.* 447 (2024) 138064.

[14] Y. Ding, K. Yu, M. Li, A review on high-strength engineered cementitious composites (HS-ECC): Design, mechanical property and structural application, *Structures* 35 (2022) 903–921.

[15] K.-X. Zhuo, G. Chen, R.-H. Luo, Y.-W. Chen, D.-H. Li, J.-X. Lin, Effect of Na_2CO_3 Replacement Quantity and Activator Modulus on Static Mechanical and Environmental Behaviours of Alkali-Activated-Strain-Hardening-Ultra-High-Performance Concrete, *Buildings* 14 (2024) 681.

[16] M.I. Jordan, T.M. Mitchell, Machine learning: Trends, perspectives, and prospects, *Science* 349 (2015) 255–260.

[17] O. Avci, O. Abdeljaber, S. Kiranyaz, M. Hussein, M. Gabbouj, D.J. Inman, A review of vibration-based damage detection in civil structures: From traditional methods to Machine Learning and Deep Learning applications, *Mech. Syst. Signal Process.* 147 (2021) 107077.

[18] Z. Chen, Q. Liu, Z. Ding, F. Liu, Automated structural resilience evaluation based on a multi-scale Transformer network using field monitoring data, *Mech. Syst. Signal Process.* 222 (2025) 111813.

[19] Y. An, E. Chatzi, S.-H. Sim, S. Laflamme, B. Blachowski, J. Ou, Recent progress and future trends on damage identification methods for bridge structures, *Struct. Control Health Monit.* 26 (2019) e2416.

[20] W. Ben Chaabene, M. Flah, M.L. Nehdi, Machine learning prediction of mechanical properties of concrete: Critical review, *Constr. Build. Mater.* 260 (2020) 119889.

- [21] Z. Fan, R. Chiong, Z. Hu, Y. Lin, A fuzzy weighted relative error support vector machine for reverse prediction of concrete components, *Comput. Struct.* 230 (2020) 106171.
- [22] A. Behnood, E.M. Golafshani, Machine learning study of the mechanical properties of concretes containing waste foundry sand, *Constr. Build. Mater.* 243 (2020) 118152.
- [23] N.-H. Nguyen, J. Abellán-García, S. Lee, T.-K. Nguyen, T.P. Vo, Simultaneous prediction the strain and energy absorption capacity of ultra-high performance fiber reinforced concretes by using multi-output regression model, *Constr. Build. Mater.* 384 (2023) 131418.
- [24] M. Alyami, M. Khan, M. Fawad, R. Nawaz, A.W.A. Hammad, T. Najeh, Y. Gamil, Predictive modeling for compressive strength of 3D printed fiber-reinforced concrete using machine learning algorithms, *Case Stud. Constr. Mater.* 20 (2024) e02728.
- [25] M. Nasir Uddin, L. Li, A. Ahmed, K. Yahya Mohammed Almajhali, Prediction of PVA fiber effect in Engineered Composite cement (ECC) by Artificial neural Network (ANN), *Mater. Today Proc.* 65 (2022) 537–542.
- [26] U.J. Malik, S.A. Khokhar, M. Hammad, R.A. Khushnood, F.A. Najam, F. Ali, M. Shahid, ANN-based predictive mimicker for the constitutive model of engineered cementitious composites (ECC), *Constr. Build. Mater.* 420 (2024) 135530.
- [27] W. Feng, Y. Wang, J. Sun, Y. Tang, D. Wu, Z. Jiang, J. Wang, X. Wang, Prediction of thermo-mechanical properties of rubber-modified recycled aggregate concrete, *Constr. Build. Mater.* 318 (2022) 125970.
- [28] Z.-M. Yang, H.-S. Pan, Z.-X. Jiang, J.-H. Lv, G.-W. Ruan, H.-M. Lai, J.-X. Lin, Pseudo strain-hardening alkali-activated composites with up to 100 % rubber aggregate: Static mechanical properties analysis and constitutive model development, *Constr. Build. Mater.* 439 (2024) 137338.
- [29] Z.-X. Jiang, J.-Y. Su, G.-W. Ruan, J.-H. Lv, Z.-M. Yang, H.-M. Lai, J.-X. Lin, Tensile behavior of rubberized high strength-high ductile concrete under elevated temperature, *Constr. Build. Mater.* 446 (2024) 138036.
- [30] G.A. Lyngdoh, M. Zaki, N.M.A. Krishnan, S. Das, Prediction of concrete strengths enabled by missing data imputation and interpretable machine learning, *Cem. Concr. Compos.* 128 (2022) 104414.

- [31] N.K. Benamara, M. Val-Calvo, J.R. Álvarez-Sánchez, A. Díaz-Morcillo, J.M. Ferrández-Vicente, E. Fernández-Jover, T.B. Stambouli, Real-time facial expression recognition using smoothed deep neural network ensemble, *Integr. Comput.-Aided Eng.* 28 (2021) 97–111.
- [32] X. Liang, Image-based post-disaster inspection of reinforced concrete bridge systems using deep learning with Bayesian optimization, *Comput.-Aided Civ. Infrastruct. Eng.* 34 (2019) 415–430.
- [33] S. Li, X. Zhao, G. Zhou, Automatic pixel-level multiple damage detection of concrete structure using fully convolutional network, *Comput.-Aided Civ. Infrastruct. Eng.* 34 (2019) 616–634.
- [34] S. Xiao, Y. Cai, Y. Guo, J. Lin, G. Liu, X. Lan, Y. Song, Experimental Study on Axial Compressive Performance of Polyvinyl Alcohol Fibers Reinforced Fly Ash—Slag Geopolymer Composites, *Polymers* 14 (2022) 142.
- [35] Y. Okazaki, S. Okazaki, S. Asamoto, P. Chun, Applicability of machine learning to a crack model in concrete bridges, *Comput.-Aided Civ. Infrastruct. Eng.* 35 (2020) 775–792.
- [36] M.H. Rafiei, H. Adeli, NEEWS: A novel earthquake early warning model using neural dynamic classification and neural dynamic optimization, *Soil Dyn. Earthq. Eng.* 100 (2017) 417–427.
- [37] K. Shaswat, Concrete slump prediction modeling with a fine-tuned convolutional neural network: hybridizing sea lion and dragonfly algorithms, *Environ. Sci. Pollut. Res.* 29 (2022) 43758–43769.
- [38] F. Deng, Y. He, S. Zhou, Y. Yu, H. Cheng, X. Wu, Compressive strength prediction of recycled concrete based on deep learning, *Constr. Build. Mater.* 175 (2018) 562–569.
- [39] K. Li, L. Pan, X. Guo, Y.F. Wang, Hybrid random aggregation model and Bayesian optimization-based convolutional neural network for estimating the concrete compressive strength, *Comput.-Aided Civ. Infrastruct. Eng.* 39 (2024) 559–574.
- [40] D. Zhang, Y. Shen, Z. Huang, X. Xie, Auto machine learning-based modelling and prediction of excavation-induced tunnel displacement, *J. Rock Mech. Geotech. Eng.* 14 (2022) 1100–1114.
- [41] Q. Zhang, W. Hu, Z. Liu, J. Tan, TBM performance prediction with Bayesian optimization and automated machine learning, *Tunn. Undergr. Space Technol.* 103 (2020) 103493.

- [42] S. Dong, X. Wu, X. Qi, C. Affolter, G.P. Terrasi, G. Xian, Prediction model of long-term tensile strength of glass fiber reinforced polymer bars exposed to alkaline solution based on Bayesian optimized artificial neural network, *Constr. Build. Mater.* 400 (2023) 132885.
- [43] J. Yu, Y. Weng, J. Yu, W. Chen, S. Lu, K. Yu, Generative AI for performance-based design of engineered cementitious composite, *Compos. Part B Eng.* 266 (2023) 110993.
- [44] L. Shi, S.T.K. Lin, Y. Lu, L. Ye, Y.X. Zhang, Artificial neural network based mechanical and electrical property prediction of engineered cementitious composites, *Constr. Build. Mater.* 174 (2018) 667–674.
- [45] K.M.A. Hossain, M.S. Anwar, S.G. Samani, Regression and artificial neural network models for strength properties of engineered cementitious composites, *Neural Comput. Appl.* 29 (2018) 631–645.
- [46] W. Chen, R. Fediuk, J. Yu, K. Nikolayevich, N. Vatin, D. Bazarov, K. Yu, Performance prediction and analysis of engineered cementitious composites based on machine learning, *Dev. Built Environ.* 18 (2024) 100459.
- [47] I. Curosu, M. Liebscher, V. Mechtcherine, C. Bellmann, S. Michel, Tensile behavior of high-strength strain-hardening cement-based composites (HS-SHCC) made with high-performance polyethylene, aramid and PBO fibers, *Cem. Concr. Res.* 98 (2017) 71–81.
- [48] B.-Y. Deng, L.-Z. Li, D. Tan, M.N. Uddin, Z.-W. Cai, K.-Q. Yu, Sustainable and cost-effective ultra-lightweight engineered cementitious composite: Design and material characterization, *Cem. Concr. Compos.* 136 (2023) 104895.
- [49] Z. Zhang, Z. Li, J. He, X. Shi, High-strength engineered cementitious composites with nanosilica incorporated: Mechanical performance and autogenous self-healing behavior, *Cem. Concr. Compos.* 135 (2023) 104837.
- [50] A. Sheta, X. Ma, Y. Zhuge, M.A. ElGawady, J.E. Mills, A. Singh, E.-S. Abd-Elal, Structural performance of novel thin-walled composite cold-formed steel/PE-ECC beams, *Thin-Walled Struct.* 162 (2021) 107586.
- [51] W. Zhou, Y. Zhang, L. Ma, V.C. Li, Influence of printing parameters on 3D printing engineered cementitious composites (3DP-ECC), *Cem. Concr. Compos.* 130 (2022) 104562.

- [52] Z. Feng, Y. Zhou, L. Sui, Z. Zhu, Optimal design of a low-cost high-performance hybrid fiber engineered cementitious composites, *Constr. Build. Mater.* 345 (2022) 128372.
- [53] Q. Qiu, J. Zhu, J.-G. Dai, In-situ X-ray microcomputed tomography monitoring of steel corrosion in engineered cementitious composite (ECC), *Constr. Build. Mater.* 262 (2020) 120844.
- [54] Z. Zhang, A. Yuvaraj, J. Di, S. Qian, Matrix design of light weight, high strength, high ductility ECC, *Constr. Build. Mater.* 210 (2019) 188–197.
- [55] K.-Q. Yu, J.-G. Dai, Z.-D. Lu, C.-S. Poon, Rate-dependent tensile properties of ultra-high performance engineered cementitious composites (UHP-ECC), *Cem. Concr. Compos.* 93 (2018) 218–234.
- [56] Y. Ding, K.-Q. Yu, J. Yu, S. Xu, Structural behaviors of ultra-high performance engineered cementitious composites (UHP-ECC) beams subjected to bending-experimental study, *Constr. Build. Mater.* 177 (2018) 102–115.
- [57] J. Luo, Z. Cai, K. Yu, W. Zhu, Z. Lu, Temperature impact on the micro-structures and mechanical properties of high-strength engineered cementitious composites, *Constr. Build. Mater.* 226 (2019) 686–698.
- [58] H. Zhu, W.-H. Hu, M. Mehthel, T. Villette, O.S. Vidal, W.N. Nasser, V.C. Li, Engineered cementitious composites (ECC) with a high volume of volcanic ash: Rheological, mechanical, and micro performance, *Cem. Concr. Compos.* 139 (2023) 105051.
- [59] K. Yu, Y. Ding, Y.X. Zhang, Size effects on tensile properties and compressive strength of engineered cementitious composites, *Cem. Concr. Compos.* 113 (2020) 103691.
- [60] H. Zhou, J. Wu, X. Wang, Y. Chen, X. Du, S. Yu, Performance of engineered cementitious composite (ECC) monolithic and composite slabs subjected to near-field blast, *Eng. Struct.* 279 (2023) 115561.
- [61] Y. Zhou, Q. Zhong, F. Xing, L. Sui, Z. Huang, M. Guo, Influence of cyclic loading on the tensile fracture characteristics of ultra-high performance engineered cementitious composites, *Constr. Build. Mater.* 240 (2020) 117937.

- [62] K. Yu, Y. Wang, J. Yu, S. Xu, A strain-hardening cementitious composites with the tensile capacity up to 8%, *Constr. Build. Mater.* 137 (2017) 410–419.
- [63] H. Gou, M. Sofi, Z. Zhang, M. Zhu, P. Mendis, Matrix design and performance development of eco-friendly high-strength, lightweight engineered cementitious composites, *J. Build. Eng.* 77 (2023) 107254.
- [64] Z. Xu, H. Sun, Y. Huo, M. Jia, Z. Chen, Y. Yang, Ti₃C₂T_x MXene modified polyethylene fibers for enhancing interface properties of strain-hardening cementitious composites, *Cem. Concr. Compos.* 145 (2024) 105358.
- [65] Y. Li, J. Li, E.-H. Yang, X. Guan, Mechanism study of crack propagation in river sand Engineered Cementitious Composites (ECC), *Cem. Concr. Compos.* 128 (2022) 104434.
- [66] D.-Y. Lei, L.-P. Guo, Y. Li, Z. Zheng, J.-P. Liu, S.-C. Li, P.-G. Wang, C.-C. Li, V. Mechtcherine, Z.-H. Li, D.-Z. Zeng, B.-M. Zhong, The investigating on mechanical properties of ultra-high strength and ultra-high ductility cementitious composites (UHS-UHDCC), *J. Build. Eng.* 43 (2021) 102486.
- [67] Q. Zhang, J. Xiao, K. Zhang, Z. Duan, Mechanical behavior of seawater sea-sand recycled concrete columns confined by engineered cementitious composite under eccentric compression, *J. Build. Eng.* 45 (2022) 103497.
- [68] B.-T. Huang, J. Yu, J.-Q. Wu, J.-G. Dai, C.K.Y. Leung, Seawater sea-sand Engineered Cementitious Composites (SS-ECC) for marine and coastal applications, *Compos. Commun.* 20 (2020) 100353.
- [69] B.-T. Huang, J.-Q. Wu, J. Yu, J.-G. Dai, C.K.Y. Leung, V.C. Li, Seawater sea-sand engineered/strain-hardening cementitious composites (ECC/SHCC): Assessment and modeling of crack characteristics, *Cem. Concr. Res.* 140 (2021) 106292.
- [70] L. Sui, Q. Zhong, K. Yu, F. Xing, P. Li, Y. Zhou, Flexural Fatigue Properties of Ultra-High Performance Engineered Cementitious Composites (UHP-ECC) Reinforced by Polymer Fibers, *Polymers* 10 (2018) 892.
- [71] W.-H. Mao, J.-P. Liu, Y. Ding, High-modulus and low-shrinkage hybrid-fiber reinforced engineered cementitious composites (ECC), *Mater. Struct.* 55 (2022) 87.

- [72] Y. Wang, F. Liu, J. Yu, F. Dong, J. Ye, Effect of polyethylene fiber content on physical and mechanical properties of engineered cementitious composites, *Constr. Build. Mater.* 251 (2020) 118917.
- [73] J. Ye, C. Cui, J. Yu, K. Yu, J. Xiao, Fresh and anisotropic-mechanical properties of 3D printable ultra-high ductile concrete with crumb rubber, *Compos. Part B Eng.* 211 (2021) 108639.
- [74] K.-Q. Yu, Z.-D. Lu, J.-G. Dai, S.P. Shah, Direct Tensile Properties and Stress–Strain Model of UHP-ECC, *J. Mater. Civ. Eng.* 32 (2020) 04019334.
- [75] Y. Guo, Z. Chen, X. Qin, A. Shen, S. Zhou, Z. Lyu, Evolution mechanism of microscopic pores in pavement concrete under multi-field coupling, *Constr. Build. Mater.* 173 (2018) 381–393.
- [76] Q. Liao, Y. Su, J. Yu, K. Yu, Compression-shear performance and failure criteria of seawater sea-sand engineered cementitious composites with polyethylene fibers, *Constr. Build. Mater.* 345 (2022) 128386.
- [77] J. Wei, C. Wu, Y. Chen, C.K.Y. Leung, Shear strengthening of reinforced concrete beams with high strength strain-hardening cementitious composites (HS-SHCC), *Mater. Struct.* 53 (2020) 102.
- [78] L. Liang, J. Yang, G. Lv, Z. Lei, X. Li, Q. Liu, Surface-Functionalized Nanocelluloses as Viscosity-Modifying Agents in Engineered Cementitious Composites, *Front. Mater.* 8 (2021).
- [79] Z. Cai, F. Liu, J. Yu, K. Yu, L. Tian, Development of ultra-high ductility engineered cementitious composites as a novel and resilient fireproof coating, *Constr. Build. Mater.* 288 (2021) 123090.
- [80] J. Yu, F. Jiang, K. Yu, F. Dong, X. Duan, Deformability enhancement of fiber-reinforced cementitious composite by incorporating recycled powder, *J. Reinf. Plast. Compos.* 39 (2020) 119–131.
- [81] J. Jiang, J. Luo, J. Yu, Z. Wang, Performance Improvement of a Fiber-Reinforced Polymer Bar for a Reinforced Sea Sand and Seawater Concrete Beam in the Serviceability Limit State, *Sensors* 19 (2019) 654.
- [82] Z.-W. Cai, J.-T. Yu, X.-Z. Duan, B.-Y. Deng, Z.-D. Lu, K.-Q. Yu, Enhancing the strain-hardening performance of ultra-high performance concrete by tailoring matrix toughness and fiber parameters, *Constr. Build. Mater.* 395 (2023) 132335.

- [83] Y. Jiangtao, J. Fangming, X. Qiong, Y. Qi, L. Mi, Mechanical and Self-Healing Performances of Crumb Rubber Modified High-Strength Engineered Cementitious Composites, *Front. Mater.* 8 (2021).
- [84] P. Men, X. Wang, D. Liu, Z. Zhang, Q. Zhang, Y. Lu, On use of polyvinylpyrrolidone to modify polyethylene fibers for improving tensile properties of high strength ECC, *Constr. Build. Mater.* 417 (2024) 135354.
- [85] L. Li, Z. Cai, K. Yu, Y.X. Zhang, Y. Ding, Performance-based design of all-grade strain hardening cementitious composites with compressive strengths from 40 MPa to 120 MPa, *Cem. Concr. Compos.* 97 (2019) 202–217.
- [86] L.-S. Wu, Z.-H. Yu, C. Zhang, Z. Yuan, T. Bangi, RETRACTED: Shrinkage and tensile properties of ultra-high-performance engineered cementitious composites (UHP-ECC) containing superabsorbent polymers (SAP) and united expansion agent (UEA), *Constr. Build. Mater.* 339 (2022) 127697.
- [87] Z. Lu, J. Yao, C.K.Y. Leung, Using graphene oxide to strengthen the bond between PE fiber and matrix to improve the strain hardening behavior of SHCC, *Cem. Concr. Res.* 126 (2019) 105899.
- [88] K.-Q. Yu, J.-T. Yu, J.-G. Dai, Z.-D. Lu, S.P. Shah, Development of ultra-high performance engineered cementitious composites using polyethylene (PE) fibers, *Constr. Build. Mater.* 158 (2018) 217–227.
- [89] B.-T. Huang, Y.-T. Wang, J.-Q. Wu, J. Yu, J.-G. Dai, C.K. Leung, Effect of fiber content on mechanical performance and cracking characteristics of ultra-high-performance seawater sea-sand concrete (UHP-SSC), *Adv. Struct. Eng.* 24 (2021) 1182–1195.
- [90] R. Kohavi, A study of cross-validation and bootstrap for accuracy estimation and model selection, *Morgan Kaufm. Publ., Inc.* 2 (1995) 1137–1143.
- [91] R. Caruana, S. Lawrence, C. Giles, Overfitting in Neural Nets: Backpropagation, Conjugate Gradient, and Early Stopping, *MIT Press*, 13 (2000).
- [92] T. Lindeberg, Scale Invariant Feature Transform, *Scholarp.* 7 (2012) 10491.
- [93] K. Kawaguchi, L.P. Kaelbling, T. Lozano-Pérez, Bayesian Optimization with Exponential Convergence, *Curr. Assoc. Inc.* 28 (2015).

- [94] J. Snoek, H. Larochelle, R.P. Adams, Practical Bayesian Optimization of Machine Learning Algorithms, *Curr. Assoc. Inc.* 25 (2012).
- [95] B. Shahriari, K. Swersky, Z. Wang, R.P. Adams, N. de Freitas, Taking the Human Out of the Loop: A Review of Bayesian Optimization, *Proc. IEEE* 104 (2016) 148–175.
- [96] J. Bergstra, R. Bardenet, Y. Bengio, B. Kégl, Algorithms for Hyper-Parameter Optimization, *Curr. Assoc. Inc.* 24 (2011).
- [97] X. Wu, V. Kumar, J. Ross Quinlan, J. Ghosh, Q. Yang, H. Motoda, G.J. McLachlan, A. Ng, B. Liu, P.S. Yu, Z.-H. Zhou, M. Steinbach, D.J. Hand, D. Steinberg, Top 10 algorithms in data mining, *Knowl. Inf. Syst.* 14 (2008) 1–37.
- [98] J.H. Friedman, Greedy Function Approximation: A Gradient Boosting Machine, *Ann. Stat.* 29 (2001) 1189–1232.
- [99] G. Ke, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, Q. Ye, T.-Y. Liu, LightGBM: A Highly Efficient Gradient Boosting Decision Tree, *Curr. Assoc. Inc.* 30 (2017).
- [100] L. Prokhorenkova, G. Gusev, A. Vorobev, A.V. Dorogush, A. Gulin, CatBoost: unbiased boosting with categorical features, *Process. Syst.* 31(2018).
- [101] M. Krzywinski, N. Altman, Classification and regression trees, *Nat. Methods* 14 (2017) 757–758.
- [102] T. Chen, C. Guestrin, XGBoost: A Scalable Tree Boosting System, *Assoc. Comput. Mach.* (2016) 785–794.
- [103] L. Breiman, Random Forests, *Mach. Learn.* 45 (2001) 5–32.
- [104] M. Altayeb, X. Wang, T.H. Musa, An ensemble method for predicting the mechanical properties of strain hardening cementitious composites, *Constr. Build. Mater.* 286 (2021) 122807.
- [105] I.U. Ekanayake, D.P.P. Meddage, U. Rathnayake, A novel approach to explain the black-box nature of machine learning in compressive strength predictions of concrete using Shapley additive explanations (SHAP), *Case Stud. Constr. Mater.* 16 (2022) e01059.
- [106] Ranade R., Li V.C., Stults M.D., Heard W.F., Rushing T.S., Composite Properties of High-Strength, High-Ductility Concrete., *Mater. J.* 110 (2013) 413-422.

[107] S.H. Said, H.A. Razak, The effect of synthetic polyethylene fiber on the strain hardening behavior of engineered cementitious composite (ECC), Mater. Des. 86 (2015) 447–457.

Appendix

Appendix 1. Optimal hyperparameters of the BO-CNN model for tensile strength.

Hyperparameter	Optimal value	Hyperparameter	Optimal value
Spatial transformation type of the first convolutional layer	Transposed convolution layer	Batch normalization for the second fully connected layer	No
Activation function of the first convolutional layer	Leaky ReLU	Regularization type of the second fully connected layer	L2
Batch normalization for the first convolutional layer	Yes	Number of neurons of the second fully connected layer	49
Number of filters of the first convolutional layer	17	Activation function of the third fully connected layer	ReLU
Filter size of the first convolutional layer	2	Batch normalization for the third fully connected layer	No
Filter stride of the first convolutional layer	1	Regularization type of the third fully connected layer	None
Spatial transformation type of the second convolutional layer	None	Number of neurons of the third fully connected layer	128
Activation function of the second convolutional layer	ReLU	Activation function of the fourth fully connected layer	ReLU
Batch normalization for the second convolutional layer	Yes	Batch normalization for the fourth fully connected layer	No
Number of filters of the second	91	Regularization type of the	L2

convolutional layer		fourth fully connected layer	
Filter size of the second convolutional layer	1	Number of neurons of the fourth fully connected layer	43
Filter stride of the second convolutional layer	2	Activation function of the fifth fully connected layer	ReLU
Activation function of the first fully connected layer	ReLU	Batch normalization for the fifth fully connected layer	No
Batch normalization for the first fully connected layer	No	Regularization type of the fifth fully connected layer	L2
Regularization type of the first fully connected layer	L1	Number of neurons of the fifth fully connected layer	24
Number of neurons of the first fully connected layer	55	Optimizer type	RMSProp
Dropout rate of the first fully connected layer	0	Optimizer learning rate	0.000187
Activation function of the second fully connected layer	ReLU		

Appendix 2. Optimal hyperparameters of the BO-CNN model for ultimate tensile strain.

Hyperparameter	Optimal value	Hyperparameter	Optimal value
----------------	---------------	----------------	---------------

Activation function of		Batch normalization	
the first convolutional layer	ReLU	for the first fully connected layer	No
Batch normalization for the first convolutional layer	No	Regularization type of the first fully connected layer	L1 and L2
Number of filters of the first convolutional layer	125	Number of neurons of the first fully connected layer	226
Filter size of the first convolutional layer	1	Dropout rate of the first fully connected layer	No
Filter stride of the first convolutional layer	1	Activation function for the second fully connected layer	ReLU
Spatial transformation type of the first convolutional layer	Pooling layer	Batch normalization for the second fully connected layer	No
Type of the first pooling layer	Max pooling	Regularization type of the second fully connected layer	None
Kernel size of the first pooling layer	2	Number of neurons of the second fully connected layer	163
Kernel stride of the first pooling layer	1	Activation function of the third fully	Leaky ReLU

		connected layer	
Activation function of		Batch normalization	
the second	ReLU	for the third fully	No
convolutional layer		connected layer	
Batch normalization		Regularization type of	
for the second	Yes	the third fully	L2
convolutional layer		connected layer	
Number of filters of		Number of neurons of	
the second	205	the third fully	114
convolutional layer		connected layer	
Filter size of the		Activation function of	
second convolutional	3	the fourth fully	ReLU
layer		connected layer	
Filter stride of the		Batch normalization	
second convolutional	1	for the fourth fully	Yes
layer		connected layer	
Spatial transformation		Regularization type of	
type of the second	Transposed	the fourth fully	L1
convolutional layer	convolution layer	connected layer	
Number of filters of		Number of neurons of	
the second transposed	123	the fourth fully	65
convolutional layer		connected layer	
Kernel size of the			
second transposed	3	Optimizer type	RMSProp
convolution layer			
Kernel stride of the	2	Optimizer learning	0.000187

second transposed rate
convolution layer
Activation function of
the first fully Leaky ReLU
connected layer

Appendix C. Database used in this study.

	C	FA	GGBS	LP	SF	S	W	V_f	L_f	D_f	TS	UTS
								%	mm	μm	MPa	%
Y. Huo [7]	0.450	0.450	0.000	0.000	0.100	0.890	0.180	1.50	18	20	10.89	4.42
W. Chen [46]	0.700	0.300	0.000	0.000	0.000	0.300	0.270	2.00	18	21	5.74	7.40
W. Chen [46]	0.700	0.300	0.000	0.000	0.000	0.300	0.270	2.00	12	25	6.50	6.85
W. Chen [46]	0.700	0.300	0.000	0.000	0.000	0.300	0.270	2.00	18	25	8.41	8.71
W. Chen [46]	0.700	0.300	0.000	0.000	0.000	0.300	0.270	2.00	18	27	6.35	6.33
W.	0.700	0.300	0.000	0.000	0.000	0.300	0.270	1.00	12	21	6.19	5.14

	C	FA	GGBS	LP	SF	S	W	V_f	Lf	D_f	TS	UTS
								%	mm	μm	MPa	%
Chen												
[46]												
W.												
Chen	0.700	0.300	0.000	0.000	0.000	0.300	0.270	1.50	18	24	5.77	13.43
[46]												
W.												
Chen	0.700	0.300	0.000	0.000	0.000	0.300	0.270	1.50	12	25	6.48	9.25
[46]												
W.												
Chen	0.700	0.300	0.000	0.000	0.000	0.300	0.270	1.50	18	25	8.66	8.40
[46]												
W.												
Chen	0.200	0.800	0.000	0.000	0.000	0.190	0.270	1.00	18	27	3.77	6.55
[46]												
W.												
Chen	0.200	0.800	0.000	0.000	0.000	0.190	0.270	1.00	12	27	3.31	3.68
[46]												
W.												
Chen	0.200	0.800	0.000	0.000	0.000	0.190	0.270	0.75	18	24	4.33	2.48
[46]												
W.												
Chen	0.200	0.800	0.000	0.000	0.000	0.190	0.270	0.75	18	21	3.13	6.70
[46]												
W.	0.200	0.800	0.000	0.000	0.000	0.190	0.270	1.00	18	25	5.17	6.73

	C	FA	GGBS	LP	SF	S	W	V_f	L_f	D_f	TS	UTS
								%	mm	μm	MPa	%
Chen												
[46]												
W.												
Chen	0.200	0.800	0.000	0.000	0.000	0.190	0.270	1.00	12	24	3.14	2.58
[46]												
W.												
Chen	0.200	0.800	0.000	0.000	0.000	0.190	0.270	0.75	18	27	3.29	4.55
[46]												
W.												
Chen	0.200	0.800	0.000	0.000	0.000	0.190	0.270	0.75	12	27	2.78	1.78
[46]												
W.												
Chen	0.700	0.300	0.000	0.000	0.000	0.300	0.270	1.00	18	21	6.66	7.28
[46]												
W.												
Chen	0.700	0.300	0.000	0.000	0.000	0.300	0.270	1.00	18	24	6.62	6.21
[46]												
W.												
Chen	0.700	0.300	0.000	0.000	0.000	0.300	0.270	1.00	18	25	6.70	6.24
[46]												
W.												
Chen	0.700	0.300	0.000	0.000	0.000	0.300	0.270	1.00	18	27	6.09	5.87
[46]												
W.	0.700	0.300	0.000	0.000	0.000	0.300	0.270	1.50	12	21	5.94	11.78

	C	FA	GGBS	LP	SF	S	W	V_f	L_f	D_f	TS	UTS
								%	mm	μm	MPa	%
Chen												
[46]												
W.												
Chen	0.700	0.300	0.000	0.000	0.000	0.300	0.270	1.50	18	21	6.71	10.43
[46]												
W.												
Chen	0.700	0.300	0.000	0.000	0.000	0.300	0.270	2.00	12	21	5.59	6.95
[46]												
W.												
Chen	0.700	0.300	0.000	0.000	0.000	0.300	0.270	2.00	18	25	6.60	6.50
[46]												
I.												
Curosu	0.830	0.000	0.000	0.000	0.170	0.080	0.180	2.00	6	20	7.60	3.90
[47]												
B.-Y.												
Deng	0.461	0.433	0.000	0.000	0.106	0.912	0.320	1.00	18	24	5.54	1.34
[48]												
Z.												
Zhang	0.506	0.404	0.000	0.000	0.090	0.303	0.204	2.00	12	26	8.01	3.27
[49]												
Z.												
Zhang	0.327	0.586	0.000	0.000	0.087	0.293	0.197	2.00	12	26	6.11	4.48
[49]												
A. Sheta	0.453	0.000	0.460	0.000	0.087	0.355	0.302	2.00	12	12	7.00	7.00

	C	FA	GGBS	LP	SF	S	W	V_f	Lf	D_f	TS	UTS
								%	mm	μm	MPa	%
[50]												
W.												
Zhou	0.470	0.440	0.000	0.000	0.090	0.290	0.240	1.50	12	24	7.84	5.73
[51]												
Z. Feng												
[52]												
Q. Qiu	0.412	0.000	0.443	0.058	0.087	0.364	0.303	2.00	18	25	5.74	8.39
[53]												
Z.												
Zhang	0.455	0.545	0.000	0.000	0.000	0.164	0.260	2.00	13	17	5.63	7.92
[54]												
Z.												
Zhang	0.505	0.404	0.000	0.000	0.091	0.303	0.202	2.00	18	26	9.76	4.16
[54]												
K.-Q.												
Yu [55]	0.400	0.000	0.429	0.057	0.114	0.286	0.131	2.00	12	24	12.79	3.35
K.-Q.												
Yu [55]	0.400	0.000	0.429	0.057	0.114	0.286	0.131	2.00	18	24	18.10	5.92
K.-Q.												
Yu [55]	0.400	0.000	0.429	0.057	0.114	0.286	0.131	2.00	18	20	18.81	6.13
Y. Ding												
[56]	0.400	0.000	0.429	0.057	0.114	0.286	0.131	2.00	18	24	16.46	8.00
J. Luo												
[57]	0.357	0.000	0.446	0.107	0.090	0.298	0.149	2.00	18	24	9.46	9.08

	C	FA	GGBS	LP	SF	S	W	V_f	Lf	D_f	TS	UTS
								%	mm	μm	MPa	%
H. Zhu [58]	0.625	0.375	0.000	0.000	0.000	0.300	0.250	2.00	12	20	7.00	6.80
H. Zhu [58]	0.455	0.545	0.000	0.000	0.000	0.300	0.250	2.00	12	20	7.70	6.60
H. Zhu [58]	0.313	0.688	0.000	0.000	0.000	0.300	0.250	2.00	12	20	5.30	5.40
K. Yu [59]	0.357	0.536	0.000	0.000	0.107	0.429	0.200	2.00	18	24	8.49	7.29
K. Yu [59]	0.357	0.536	0.000	0.000	0.107	0.428	0.250	2.00	18	24	6.32	6.61
K. Yu [59]	0.357	0.536	0.000	0.000	0.108	0.429	0.300	2.00	18	24	4.80	7.28
K. Yu [59]	0.357	0.536	0.000	0.000	0.108	0.429	0.350	2.00	18	24	4.15	6.36
K. Yu [59]	0.357	0.536	0.000	0.000	0.108	0.429	0.420	2.00	18	24	3.38	6.35
H. Zhou [60]	0.438	0.562	0.000	0.000	0.000	0.059	0.280	2.00	12	24	8.80	5.30
Y. Zhou [61]	0.412	0.000	0.441	0.088	0.059	0.294	0.135	2.00	18	25	15.10	7.80
K. Yu [62]	0.455	0.545	0.000	0.000	0.000	0.364	0.240	2.00	18	25	6.52	7.98
K. Yu [62]	0.823	0.177	0.000	0.000	0.000	0.528	0.278	2.00	18	25	6.07	11.98

	C	FA	GGBS	LP	SF	S	W	V_f	Lf	D_f	TS	UTS
								%	mm	μm	MPa	%
K. Yu [62]	0.385	0.000	0.500	0.000	0.115	0.538	0.177	2.00	18	25	17.84	8.49
H. Gou [63]	0.545	0.436	0.000	0.000	0.019	0.327	0.158	2.00	18	22	11.40	6.01
Z. Xu [64]	0.550	0.350	0.000	0.000	0.100	0.400	0.180	1.50	18	20	6.30	4.69
Y. Li [65]	0.600	0.300	0.000	0.000	0.100	0.650	0.160	2.00	18	19	11.88	5.94
Y. Li [65]	0.800	0.000	0.000	0.000	0.200	0.300	0.180	2.00	18	24	11.70	4.10
Y. Li [65]	0.600	0.000	0.000	0.400	0.000	0.360	0.230	2.00	12	24	4.89	2.48
Y. Li [65]	0.600	0.000	0.000	0.300	0.100	0.360	0.230	2.00	12	24	5.66	3.09
Y. Li [65]	0.600	0.000	0.000	0.200	0.200	0.360	0.230	2.00	12	24	5.89	3.23
Y. Li [65]	0.600	0.000	0.000	0.000	0.400	0.360	0.230	2.00	12	24	6.55	4.71
Y. Li [65]	0.600	0.000	0.000	0.000	0.400	0.360	0.210	2.00	12	24	6.74	6.06
Y. Li [65]	0.600	0.000	0.000	0.000	0.400	0.360	0.190	2.00	12	24	7.00	6.53
Y. Li [65]	0.600	0.000	0.000	0.000	0.400	0.360	0.170	2.00	12	24	6.65	4.98

	C	FA	GGBS	LP	SF	S	W	V_f	Lf	D_f	TS	UTS
								%	mm	μm	MPa	%
Y. Li [65]	0.600	0.000	0.000	0.000	0.400	0.360	0.150	2.00	12	24	6.32	4.18
D.-Y. Lei [66]	0.600	0.000	0.000	0.100	0.300	0.360	0.230	2.00	12	24	6.35	4.42
Q. Zhang [67]	0.320	0.000	0.590	0.000	0.090	0.300	0.200	2.00	18	20	10.52	3.20
B.-T. Huang [68]	0.800	0.000	0.000	0.000	0.200	0.300	0.180	2.00	12	24	8.00	5.30
B.-T. Huang [69]	0.800	0.000	0.000	0.000	0.200	0.300	0.180	2.00	12	24	8.14	5.14
B.-T. Huang [69]	0.800	0.000	0.000	0.000	0.200	0.300	0.180	1.00	12	24	6.09	1.65
B.-T. Huang [69]	0.800	0.000	0.000	0.000	0.200	0.300	0.180	1.50	12	24	8.23	4.73
B.-T. Huang [69]	0.800	0.000	0.000	0.000	0.200	0.300	0.180	2.00	12	24	7.69	2.38
B.-T. Huang	0.800	0.000	0.000	0.000	0.200	0.300	0.180	2.00	12	24	7.42	7.00

	C	FA	GGBS	LP	SF	S	W	V_f	Lf	D_f	TS	UTS
								%	mm	μm	MPa	%
Y.												
Wang [72]	0.700	0.300	0.000	0.000	0.000	0.300	0.250	0.75	18	24	4.06	1.59
Y.												
Wang [72]	0.700	0.300	0.000	0.000	0.000	0.300	0.250	1.00	18	24	4.51	3.04
Y.												
Wang [72]	0.700	0.300	0.000	0.000	0.000	0.300	0.250	1.50	18	24	3.57	9.38
Y.												
Wang [72]	0.700	0.300	0.000	0.000	0.000	0.300	0.250	2.00	18	24	3.00	11.99
J. Ye [73]	0.560	0.340	0.000	0.000	0.100	0.360	0.310	2.00	12	24	8.08	6.95
K.-Q. Yu [74]	0.550	0.350	0.000	0.000	0.100	0.400	0.180	1.50	18	20	9.08	4.42
K.-Q. Yu [74]	0.412	0.000	0.441	0.059	0.088	0.290	0.135	1.50	12	24	10.29	2.20
K.-Q. Yu [74]	0.412	0.000	0.441	0.059	0.088	0.290	0.135	1.75	12	24	10.68	3.52
K.-Q. Yu [74]	0.412	0.000	0.441	0.059	0.088	0.290	0.135	2.25	12	24	13.56	6.40
K.-Q. Yu [74]	0.412	0.000	0.441	0.059	0.088	0.290	0.135	2.50	12	24	13.73	6.56

	C	FA	GGBS	LP	SF	S	W	V_f	Lf	D_f	TS	UTS
								%	mm	μm	MPa	%
K.-Q. Yu [74]	0.412	0.000	0.441	0.059	0.088	0.290	0.135	3.00	12	24	14.42	6.60
K.-Q. Yu [74]	0.412	0.000	0.441	0.059	0.088	0.290	0.135	1.50	18	24	13.79	5.57
K.-Q. Yu [74]	0.412	0.000	0.441	0.059	0.088	0.290	0.135	1.75	18	24	14.55	6.70
K.-Q. Yu [74]	0.412	0.000	0.441	0.059	0.088	0.290	0.135	2.25	18	24	17.75	7.78
K.-Q. Yu [74]	0.412	0.000	0.441	0.059	0.088	0.290	0.135	1.25	18	20	12.23	4.74
K.-Q. Yu [74]	0.412	0.000	0.441	0.059	0.088	0.290	0.135	1.50	18	20	14.61	4.90
K.-Q. Yu [74]	0.412	0.000	0.441	0.059	0.088	0.290	0.135	1.75	18	20	14.33	6.11
Y. Guo [75]	0.730	0.000	0.000	0.000	0.270	0.450	0.160	2.00	12	24	8.17	3.96
Y. Guo [75]	0.660	0.000	0.060	0.000	0.270	0.460	0.160	2.00	12	24	8.86	3.34
Y. Guo [75]	0.600	0.000	0.130	0.000	0.280	0.460	0.160	2.00	12	24	9.34	3.18
Y. Guo [75]	0.530	0.000	0.190	0.000	0.280	0.470	0.160	2.00	12	24	8.52	3.94
Y. Guo [75]	0.460	0.000	0.260	0.000	0.280	0.470	0.160	2.00	12	24	7.94	4.85

	C	FA	GGBS	LP	SF	S	W	V_f	Lf	D_f	TS	UTS
								%	mm	μm	MPa	%
Y. Guo [75]	0.380	0.000	0.330	0.000	0.290	0.480	0.160	2.00	12	24	7.26	5.47
Q. Liao [76]	0.480	0.440	0.000	0.000	0.080	0.300	0.260	0.50	18	24	4.71	0.99
Q. Liao [76]	0.480	0.440	0.000	0.000	0.080	0.300	0.260	1.00	18	24	9.56	3.36
Q. Liao [76]	0.480	0.440	0.000	0.000	0.080	0.300	0.260	1.50	18	24	11.87	5.76
J. Wei [77]	0.800	0.000	0.000	0.000	0.200	0.300	0.200	2.00	12	24	10.00	3.50
L. Liang [78]	0.500	0.400	0.000	0.000	0.100	0.300	0.220	2.00	18	24	9.71	8.15
Z. Cai [79]	0.470	0.000	0.440	0.000	0.090	0.290	0.140	2.00	18	24	15.16	6.11
Z. Cai [79]	0.470	0.000	0.440	0.000	0.090	0.290	0.150	2.00	18	24	13.86	8.27
Z. Cai [79]	0.360	0.000	0.440	0.110	0.090	0.290	0.140	2.00	18	24	15.83	8.29
Z. Cai [79]	0.240	0.000	0.440	0.220	0.090	0.290	0.140	2.00	18	24	14.83	7.64
J. Yu [80]	0.450	0.550	0.000	0.000	0.000	0.360	0.280	2.00	18	25	4.27	4.76
J. Jiang [81]	0.450	0.550	0.000	0.000	0.000	0.360	0.230	2.00	18	24	6.25	7.79

	C	FA	GGBS	LP	SF	S	W	V_f	Lf	D_f	TS	UTS
								%	mm	μm	MPa	%
Z.-W. Cai [82]	0.440	0.000	0.370	0.100	0.090	0.310	0.180	2.00	12	24	6.50	3.58
Z.-W. Cai [82]	0.410	0.000	0.440	0.060	0.090	0.300	0.250	2.00	12	24	5.78	4.24
Z.-W. Cai [82]	0.410	0.000	0.440	0.060	0.090	0.350	0.320	2.00	12	24	4.61	3.58
Y. Jiangtao [83]	0.440	0.480	0.000	0.000	0.090	0.400	0.190	2.00	18	24	11.73	7.14
P. Men [84]	0.510	0.400	0.000	0.000	0.090	0.300	0.200	2.00	12	26	6.47	3.55
L. Li [85]	0.412	0.000	0.441	0.059	0.088	0.294	0.135	1.00	18	24	10.30	2.20
L. Li [85]	0.412	0.000	0.441	0.059	0.088	0.294	0.135	1.00	18	20	10.70	3.52
L. Li [85]	0.412	0.000	0.441	0.059	0.088	0.294	0.135	2.00	12	24	12.10	5.11
L. Li [85]	0.412	0.000	0.441	0.059	0.088	0.294	0.135	2.00	18	24	16.50	8.00
L. Li [85]	0.412	0.000	0.441	0.059	0.088	0.294	0.135	2.00	18	20	17.80	8.49
L. Li [85]	0.413	0.000	0.443	0.058	0.087	0.349	0.320	1.00	12	24	4.30	3.01
L. Li	0.413	0.000	0.443	0.058	0.087	0.349	0.320	1.00	18	24	4.50	5.40

	C	FA	GGBS	LP	SF	S	W	V_f	Lf	D_f	TS	UTS
								%	mm	μm	MPa	%
[85]												
L. Li	0.413	0.000	0.443	0.058	0.087	0.349	0.320	1.00	18	20	4.80	6.03
[85]												
L. Li	0.413	0.000	0.443	0.058	0.087	0.349	0.320	2.00	12	24	5.10	6.18
[85]												
L. Li	0.413	0.000	0.443	0.058	0.087	0.349	0.320	2.00	18	24	6.20	10.64
[85]												
L. Li	0.413	0.000	0.443	0.058	0.087	0.349	0.320	2.00	18	20	6.50	11.28
[85]												
L. Li	0.413	0.000	0.442	0.058	0.087	0.304	0.220	1.00	12	24	5.40	2.59
[85]												
L. Li	0.413	0.000	0.442	0.058	0.087	0.304	0.220	1.00	18	24	6.00	5.91
[85]												
L. Li	0.413	0.000	0.442	0.058	0.087	0.304	0.220	1.00	18	20	6.40	6.13
[85]												
L. Li	0.413	0.000	0.442	0.058	0.087	0.304	0.220	2.00	12	24	6.50	6.88
[85]												
L. Li	0.413	0.000	0.442	0.058	0.087	0.304	0.220	2.00	18	24	9.50	11.14
[85]												
L. Li	0.413	0.000	0.442	0.058	0.087	0.304	0.220	2.00	18	20	10.40	12.00
[85]												
L. Li	0.413	0.000	0.442	0.058	0.087	0.292	0.180	1.00	18	24	7.30	4.31
[85]												
L. Li	0.413	0.000	0.442	0.058	0.087	0.292	0.180	1.00	18	20	7.70	5.11

	C	FA	GGBS	LP	SF	S	W	V_f	Lf	D_f	TS	UTS
								%	mm	μm	MPa	%
[85]												
L. Li	0.413	0.000	0.442	0.058	0.087	0.292	0.180	2.00	12	24	8.90	5.82
[85]												
L. Li	0.413	0.000	0.442	0.058	0.087	0.292	0.180	2.00	18	24	11.40	9.02
[85]												
L. Li	0.413	0.000	0.442	0.058	0.087	0.292	0.180	2.00	18	20	11.90	9.60
[85]												
L. Li	0.413	0.000	0.442	0.058	0.087	0.292	0.180	1.00	12	24	6.60	3.36
[85]												
L. Li	0.413	0.000	0.443	0.058	0.090	0.292	0.160	1.00	12	24	7.50	2.61
[85]												
L. Li	0.413	0.000	0.443	0.058	0.090	0.292	0.160	1.00	18	24	8.10	3.66
[85]												
L. Li	0.413	0.000	0.443	0.058	0.090	0.292	0.160	1.00	18	20	8.40	4.59
[85]												
L. Li	0.413	0.000	0.443	0.058	0.090	0.292	0.160	2.00	12	24	11.40	6.69
[85]												
L. Li	0.413	0.000	0.443	0.058	0.090	0.292	0.160	2.00	18	24	13.20	10.13
[85]												
L. Li	0.413	0.000	0.443	0.058	0.090	0.292	0.160	2.00	18	20	14.80	10.11
[85]												
L.-S.	0.580	0.090	0.150	0.000	0.180	0.530	0.170	2.00	12	24	10.88	4.39
Wu [86]												
Z. Lu	0.800	0.000	0.000	0.000	0.200	0.300	0.200	2.00	12	24	7.13	3.75

	C	FA	GGBS	LP	SF	S	W	V_f	Lf	D_f	TS	UTS
								%	mm	μm	MPa	%
[87]												
K.-Q.												
Yu [88]	0.470	0.000	0.440	0.000	0.090	0.294	0.135	2.00	18	20	17.42	8.17
B.-T.												
Huang	0.800	0.000	0.000	0.000	0.200	0.300	0.180	1.00	12	24	6.00	1.80
[89]												
B.-T.												
Huang	0.800	0.000	0.000	0.000	0.200	0.300	0.180	1.50	12	24	8.00	4.60
[89]												

Data availability

Data will be made available on request.