Virtual mix design： prediction of compressive strength of concrete with industrial wastes using deep data augmentation

Ning Chen

Lecturer (Ph.D.), Beijing Key Laboratory of Traffic Engineering, Beijing University of Technology, No.100 Pingleyuan, Chaoyang District, Beijing, China; Toyota Transportation Research Institute, 3-17 Motoshiro-cho, Toyota City, Aichi, Japan; email: chenningbjut@bjut.edu.cn

Shibo Zhao

Beijing Key Laboratory of Traffic Engineering, Beijing University of Technology, No.100 Pingleyuan, Chaoyang District, Beijing, China; email: zhaoshibo@emails.bjut.edu

Zhiwei Gao

Lecturer (Ph.D.), James Watt School of Engineering, University of Glasgow, Glasgow G12 8QQ, UK; email: Zhiwei.gao@glasgow.ac.uk

Dawei Wang

Professor (Ph.D.), School of Transportation Science and Engineering, Harbin Institute of Technology, China; Institute of Highway Engineering, RWTH Aachen University, D52074 Aachen, Germany; email: wang@isac.rwth-aachen.de

Pengfei Liu

Senior researcher (Ph.D.), Institute of Highway Engineering, RWTH Aachen University, D52074 Aachen, Germany; email: liu@isac.rwth-aachen.de

Markus Oeser

Research Director (PH.D.), Institute of Highway Engineering, RWTH Aachen University, D52074 Aachen, Germany; email: oeser@isac.rwth-aachen.de

Yue Hou

Associate professor (Ph.D.), Beijing Key Laboratory of Traffic Engineering, Beijing University of Technology, No.100 Pingleyuan, Chaoyang District, Beijing, China; email: yuehou@bjut.edu.cn

Linbing Wang

Professor (Ph.D.), Department of Civil and Environmental Engineering, Virginia Tech, Blacksburg, VA24061, USA; email: wangl@vt.edu

ABSTRACT

The adding of industrial wastes, including blast furnace slag and fly ash, to concrete materials will not only improve the working performance, but also significantly reduce the carbon emissions and promote the green development in civil engineering area. The traditional material designs are mainly indoor laboratory-based, which is complex and time-consuming. In this study, a virtual material design method, including deep data augmentation methods and deep learning methods, was employed to predict the compressive strength of concrete with industrial wastes. Three types of Generative Adversarial Networks (GANs) were employed to augment the original data and the results were evaluated. The test was conducted based on a small experiment dataset from previous literature, comparing with traditional machine learning methods. Test results show that the deep learning methods have the highest accuracy in compressive strength prediction, increasing from 0.90 to 0.98 (Visual Geometry Group, VGG) and from 0.83 to 0.96 (One-Dimensional Convolutional Neural Network, 1D CNN) after deep data augmentation, where the prediction accuracy of Random Forest (RF) and Support Vector Regressive (SVR) in traditional machine learning algorithms increase from 0.91 to 0.96 and from 0.78 to 0.86, respectively. In addition, a lightweight deep convolutional neural network was designed based on the augmented dataset. The results show that the lightweight model can improve the computation efficiency, reduce the complexity of the model compared with the original model, and reach a great prediction accuracy. The proposed study can facilitate the concrete material design with industrial wastes with less labor and time cost compared with traditional ones, thus can provide a cleaner solution for the whole industry.

KEYWORDS: virtual material design; compressive strength prediction, data augmentation, deep learning, lightweight model.

1. BACKGROUND AND INTRODUCTION

Concrete material plays an important role in the construction of rigid pavement. High strength, good stability and good durability have always been the advantages of cement pavement, thus the corresponding material properties are very necessary for the service life and service quality [1]. Meanwhile, civil engineers have always considered the green development in concrete materials area as a top priority [2]. Nowadays, industrial wastes have become a significant problem. The adding of industrial wastes, including blast furnace slag and fly ash, to concrete materials will not only improve the working performance, but also significantly reduce the carbon emissions and promote the green development in civil engineering area. For instance, the High Performance Concrete (HPC) will add some auxiliary cementitious materials such as fly ash, slag and chemical admixtures [3].

According to ASTM C39 [4], concrete mixtures are defined as high performance cement concrete if the early compressive strength is 20 to 28 MPa at 3 to 12 hours or 1 to 3 days. After the completion of concrete mix design, the compressive strength will be tested to verify its mechanical performance. Generally, this complex test process has huge time-cost and is labor-consuming, which may limit the development of using some other industrial wastes as additives in concrete materials. Besides, even slightly change of the additive contents may significantly change the properties of concrete materials [5–7]. Therefore, it is necessary to find a fast, reliable, and time-saving method for the green material design of concrete with industrial wastes.

In recent years, with the fast development of Machine Learning (ML) method, it has become more and more widely used for structural and material design in civil engineering, especially in the prediction of Concrete Compressive Strength (CCS) [8–10]. Chithra et al. (2016) established the prediction model of compressive strength of high-performance concrete with nano silica and copper slag replacing part of cement and fine aggregate respectively based on Multiple Regression Analysis (MRA) and Artificial Neural Network (ANN) [11]. Khashman and Akpinar (2017) found that the ANN was efficient in predicting and classifying the CCS [12]. Behnood et al. (2017) used the M5P model tree algorithm to predict the compressive strength of normal concrete and HPC [13]. Kaloop et al. (2020) studied Multiple Adaptive Regression Spline model (MARS) as a feature extraction method to design the optimal input of HPC [14]. Feng et al. (2020) used the adaptive boosting algorithm to integrate several weak learning machines to predict the CCS [15]. In addition, a lot of deep learning-based researches have been conducted in properties evaluation of concrete structures, providing references for further studies in this area [16–20].

In sum, the main advantages of these studies are: 1) Some simple machine learning algorithms can well learn the potential physical relationships based on small data sets; 2) The ensemble algorithms can combine several weak learning machine methods to make better judgment [21]. Meanwhile, they also have some shortcomings: 1) At present, many of the tests are based on small test dataset. 2) few studies use the deep learning methods to investigate the deeper relationship between different material design factors. Nevertheless, the investigation of deeper relationship caters for mass data support. However, it is difficult to collect adequate data considering the huge labor and time cost during the concrete material test process. Thus, how to conduct the data augmentation based on small test dataset has become a problem for civil engineers, where the Generative Adversarial Network (GAN) can be used as a powerful tool. For example, Frid-Adar et al. (2018) proposed a method of generating synthetic medical images based on GAN, and improved the performance of CNN in medical image classification [22]. Liu et al. (2021) proposed an image generation method based on Variational Autoencoders (VAE) and GANs fusion network to solve the problem of insufficient data in leukocyte classification [23]. Zhang et al. (2020) modified the original GAN structure system and trained it on the face image data set [24]. Pei et al. (2020) proposed a 3D augmented convolution network (3DACN) to extract time series information and solve the serious data imbalance problem [25]. Li et al. (2018) combined the reinforcement learning and generative confrontation network to expand the original data set, and help to improve its generalization ability in the process of supervised training [26].

Considering the achievement of deep learning methods, this study conducted the following virtual material design studies: first, two kinds of machine learning models were used to predict the compressive strength of concrete with industrial wastes, including traditional machine learning methods (SVR and RF) and deep learning methods (VGG and 1D CNN). Based on the limited experimental data, the original data was augmented using GAN, Deep Convolutional GAN (DCGAN) and Wasserstein GAN with Gradient Penalty (WGAN-GP) methods. Later, two kinds of machine learning models were trained and tested for comparisons on datasets before and after augmentation, to test the effects of deep data augmentation. Finally, the prediction accuracies of two kinds of machine learning algorithms were compared. In order to improve the computation efficiency, the two lightweight deep convolutional neural networks were further employed and analyzed. The technical route of this study is shown in Fig. 1.

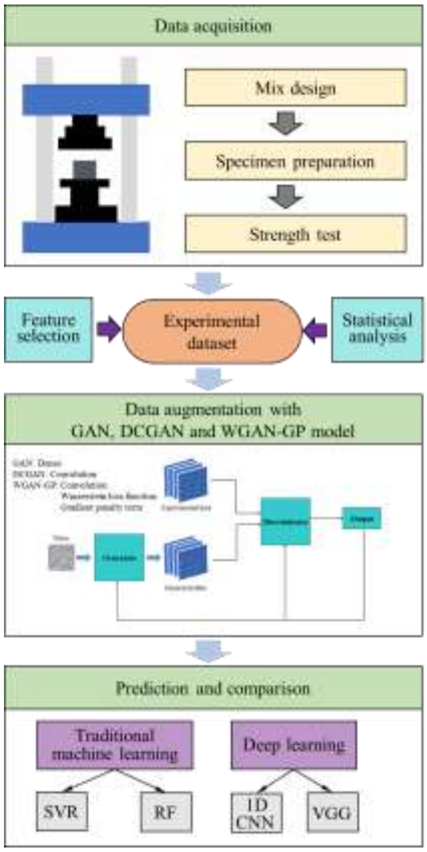


Fig. 1. Flowchart of this study.

2. DATASET

In this study, 1030 pieces of experimental data of concrete with industrial wastes originally collected by Yeh from GitHub [27] was used as the original dataset. After the concrete solidified for a period of time under normal conditions, the compressive strength was obtained through the typical compressive test on 150-mm-high cylindrical specimens. There are currently nine parameters in the experimental data set, and each input parameter has a certain effect on the ultimate compressive strength. Table 1 lists the name, unit, maximum/minimum value, average value and standard deviation of test material parameters. Note that X_2 blast furnace slag and X_3 fly ash are generally considered as industrial wastes.

Table 1 Different variables in concrete material design

Parameter	Unit	Minimum value	Maximum value	Mean	Standard deviation	Type
X_1: Cement	kg/m^3	540.00	102.00	281.17	104.46	Input
X_2: Blast furnace slag	kg/m^3	359.40	0.00	73.90	86.24	Input
X_3: Fly ash	kg/m^3	200.10	0.00	54.19	63.97	Input
X_4: Water	kg/m^3	247.00	121.75	181.57	21.35	Input
X_5: Superplasticizer	kg/m^3	32.20	0.00	6.20	5.97	Input
X_6: Coarse aggregate	kg/m^3	1145.00	801.00	972.92	77.72	Input
X_7: Fine aggregate	kg/m^3	992.60	594.00	773.58	80.14	Input
X_8: Age	days	365.00	1.00	45.66	63.14	Input
Y: Strength	MPa	82.60	2.33	35.82	16.70	Output

3. METHODOLIGIES

3.1 Traditional machine learning methods

Support Vector Machine (SVM), belonging to the Generalised Linear Classifier family and based on Vapnik-Chervonenkis Dimension theory, was first developed by Vladimir N. Vapnik for linear models in 1963 and later extended for non-linear data training in 1995 by Cortes and Vapnik [28–32]. The architecture of the SVR model is presented in Fig. 2 (a).

Random Forest (RF) is the advancement of decision-tree based ensemble algorithm, which is widely used in classification, regression and other works [33]. It is established, by constructing a group of randomly created decision trees and forecasting the class that is mode of the classification or the regression of the individual trees [34,35]. The RF structure is shown in Fig. 2 (b) and it can be described by the following Eq. (3).

$$y = \frac{1}{N}\sum_{i=1}^{N} t_i(x) \tag{3}$$

where N is number of trees, $t_i(x)$ is prediction value of each individual tree, and y is final prediction of random forest.

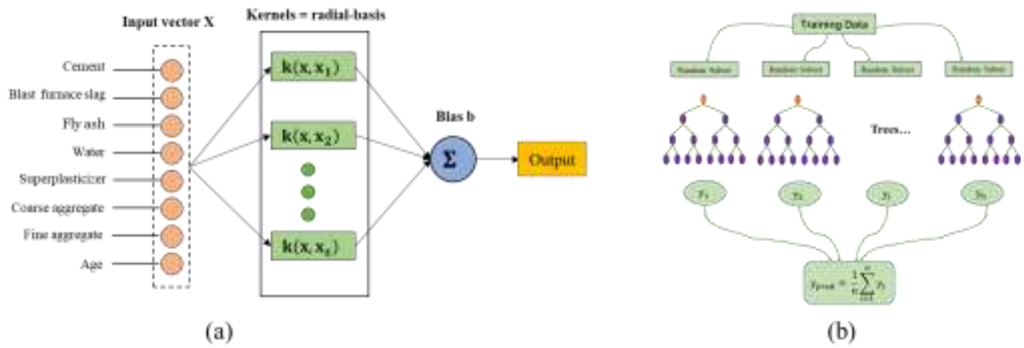


Fig. 2. The structures of various machine learning methods: (a) The architecture of the SVR model [36]. (b) The architecture of RF model.

3.2 Deep learning methods

In this study, we predicted the compressive strength of concrete based on deep learning methods, including 1D CNN and VGG.

With the rapid development of computer technology, deep learning models such as CNN were widely used in the field of computer vision [37–39]. Since Kiranyaz et al. (2015) proposed the first compact and adaptive 1D CNN [40], this structure has become popular with the advanced performance in various signal processing applications such as structural health monitoring, structural damage detection, speech recognition, etc. [41–43].

Multiple sets of data measured through the experiment can also be considered as a type of signal. Therefore, this study analyzed and predicted the experimental data based on 1D CNN [44]. The architecture of 1D CNN used in this study is shown in Fig. 3 (a).

In the CNN-layers, the 1D forward propagation (1D-FP) is defined by Eq. (5).

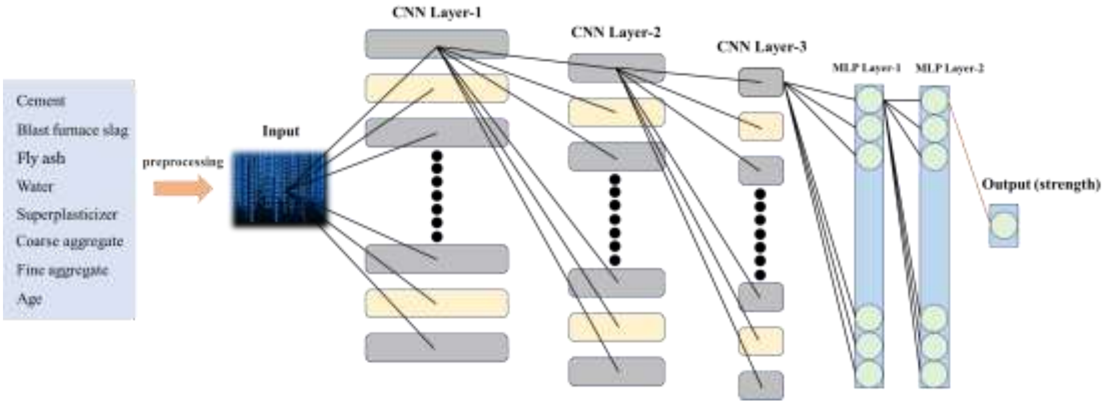
$$x_k^l = b_k^l + \sum_{i=1}^{N_{l-1}} conv1D\left(w_{ik}^{l-1}, s_i^{l-1}\right) \tag{5}$$

where x_k^l is defined as the input, b_k^l is defined as the bias of the k^{th} neuron at layer l, s_i^{l-1} is defined as the output of the i^{th} neuron at layer $l-1$, and w_{ik}^{l-1} is defined as the 1D filter kernel from the i^{th} neuron at layer $l-1$ to the k^{th} neuron at layer l.

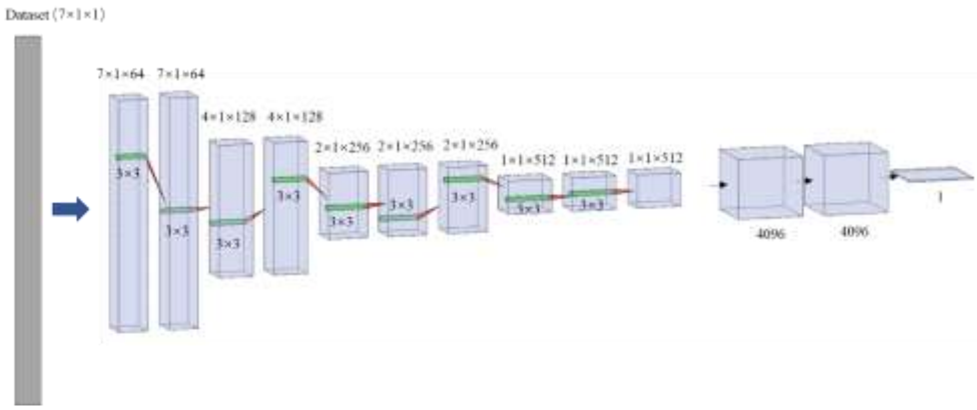
A deep architecture with a very small (3×3) convolutional filter is proposed by Simonyan et al. (2015), named VGG-16, which has achieved good results in image classification and localization tasks [45].

The two-dimensional convolutional structure was generally used to process image data, but this study applied it to a simple table form to validate the responses of model on regression problems.

The number of network layers and parameters were modified accordingly. The architecture of VGG is shown in Fig. 3 (b).



(a) 1D CNN



(b) VGG

Fig. 3. 1D CNN and VGG structure used in this study

3.3 Deep data augmentation methods

The GAN was first proposed by Goodfellow et al. (2014), which consists of two key networks: a generator G network and a discriminator D that contest with each other [46]. The goal of generator is to produce data that are as distributed as real data, but in fact the data of output could confuse the discriminator, while the goal of discriminator is to distinguish whether the data of input are from the original dataset or the generation of the generator. Thus, the competition between the generator and discriminator that can be formalized as a minimax function, as shown in Eq. (6) [47].

$$\min_{G}\max_{D} V(D,G) = E_{x\sim p(x)}[\log \mathrm{D}(x)] + E_{z\sim p_z(z)}\left[\log\left(1-\mathrm{D}\big(G(z)\big)\right)\right] \quad (6)$$

where $p(x)$ is the training data distribution, $p_z(z)$ is the prior distribution of the generative network, and z is a noise vector sampled from the model distribution $p_z(z)$.

DCGAN was used in this study, as shown in Fig. 4, is obtained by modifying the architecture of the original GANs, which replace fully-connected layer and pooling layer with convolutional layers. WGAN-GP was also used for data augmentation in this study for the reason that it can solve the problem of mode collapse occurring during training and generate higher quality images. The highlights is that model measures the distance between the distribution of the true samples and the generated samples using the Earth-Mover distance W (P_g , P_r) and consider directly constraining the gradient norm of the critic's output with respect to its input. The new objective function is as shown in Eq. (7) [48].

$$L = \underset{\tilde{x}\sim P_g}{\mathrm{E}}[D(\tilde{x})] - \underset{x\sim P_r}{\mathrm{E}}[D(x)] + \lambda \underset{\hat{x}\sim P_{\hat{x}}}{\mathrm{E}}[(\|\nabla_{\hat{x}} D(\hat{x})\|_2 - 1)^2] \quad (7)$$

where $x \sim P_r$ is the true distribution, $\hat{x} \sim P_g$ is the generated distribution, and $\hat{x} \sim P_{\hat{x}}$ is the sample in the real distribution and the generated distribution.

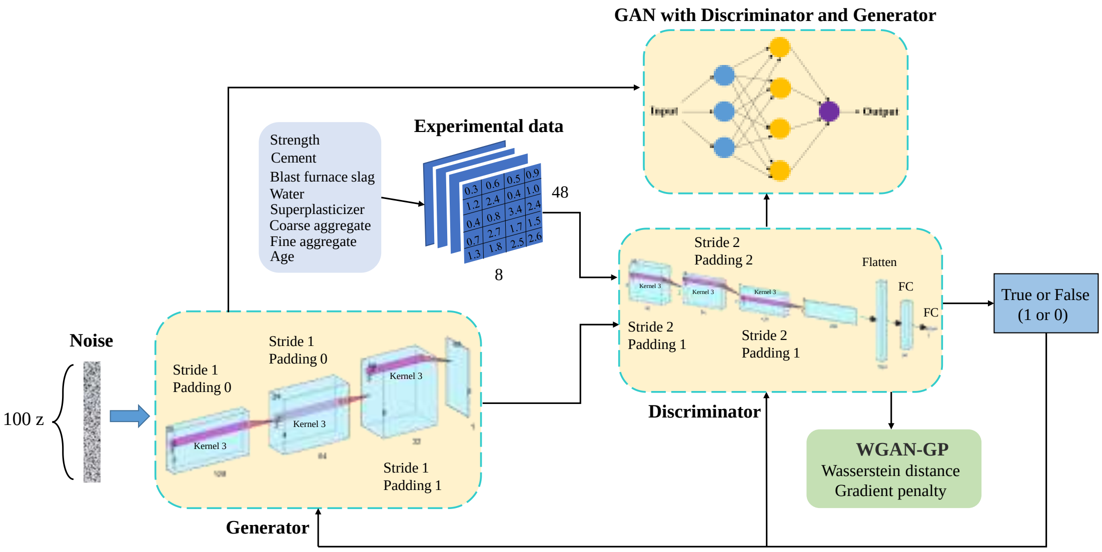


Fig. 4. GAN [48], WGAN-GP [50] and DCGAN [24] structures used in this study

4. PERFORMANCE OF MODELS BASED ON THE ORIGINAL SMALL DATASET

This section describes predicting the compressive strength of concrete with industrial wastes and comparing their performances based on the original small dataset. The advantages and disadvantages of traditional machine learning methods and deep learning methods are compared. And the collected experimental data [27] were randomly divided into training set (90%) and test set (10%) by sk-learn in python, in which the training set was used to train the model and the test set was used to evaluate the model.

4.1 Tuning model hyperparameters

In this study, the traditional machine learning models, including the support vector regression model and random forest model, were used. The SVR model selected Radial Basis Function (RBF) as the kernel function, whose parameter γ is 0.8 and the penalty factor C is 10; The RF model parameter settings refer to Feng et al. [15], where the Maximum iteration number is 200 and Learning rate is 0.2. The parameters of decision trees as weak learner are: Maximum depth blow root is 50, Minimum samples for split is 5, Minimum samples of leaf node is 2 and Minimum impurity is 10^{-4}.

The structure of one-dimensional convolutional neural network model based on deep learning is four convolution layers and two fully connected layers, in which the numbers of convolution kernels are 16, 64, 32 and 32 respectively, the sizes of convolution kernels are 3, 2, 1 and 1, the numbers of units in the fully connected layer are 256 and 64, and the activation function is Rectified Liner Unit (ReLU). The other used deep learning model VGG removes three convolutional layers with 512 convolutional kernels on the basis of the original VGG-16. Besides, the number of unit nodes in the output layer is changed to 1 and the activation function is Linear.

4.2 Statistical measures for model evaluation

To measure the performance of the predicting model of machine learning and deep learning, four types of indicators are presented, as shown in (7), (8), (9) and (10).

$$\mathrm{MAE} = \frac{\sum_{i=1}^{N}|Y_i - \hat{Y}_i|}{N} \tag{7}$$

$$\mathrm{MSE} = \frac{\sum_{i=1}^{N}(Y_i - \hat{Y}_i)^2}{N} \tag{8}$$

$$\mathrm{MAPE} = \frac{100\%}{N}\sum_{i=1}^{N}\left|\frac{Y_i - \hat{Y}_i}{Y_i}\right| \tag{9}$$

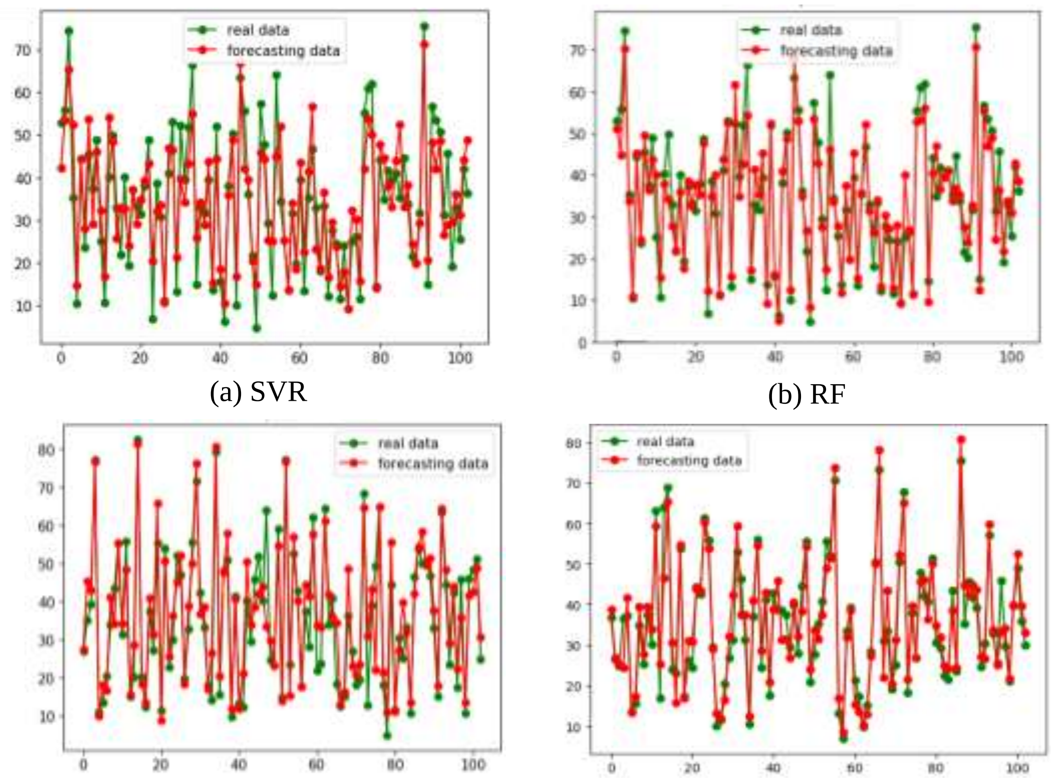
$$\mathrm{R}^2 = 1 - \frac{\sum_{i=1}^{N}(Y_i - \hat{Y}_i)^2}{\sum_{i=1}^{N}(Y_i - \bar{Y})^2} \tag{10}$$

where Y_i and $\hat{Y}_i$ are the tested and predicted values, respectively; $\bar{Y}$ is the mean value of all the tested values; N is the total number of samples in the test. MAE is used to measure the absolute error associated with the predicted value, MAPE is used to measure the percentage of the prediction error of model, and MSE is used to measure the relative error associated with the predicted value.

The measure R^2 shows the extent of the linear correlation between the predicted and real values, where the closer R^2 is to 1, the better the performance of the model is.

4.3 Model results

Fig. 5 shows the results of two kinds of machine learning methods before data augmentation. The line plots show the comparisons of concrete compressive strength real values and forecasting values on the test set, and the regression chart shows the comparisons of real values and forecasting values of concrete compressive strength on the training and test sets.



(a) SVR (b) RF

(c) 1D CNN (d) VGG

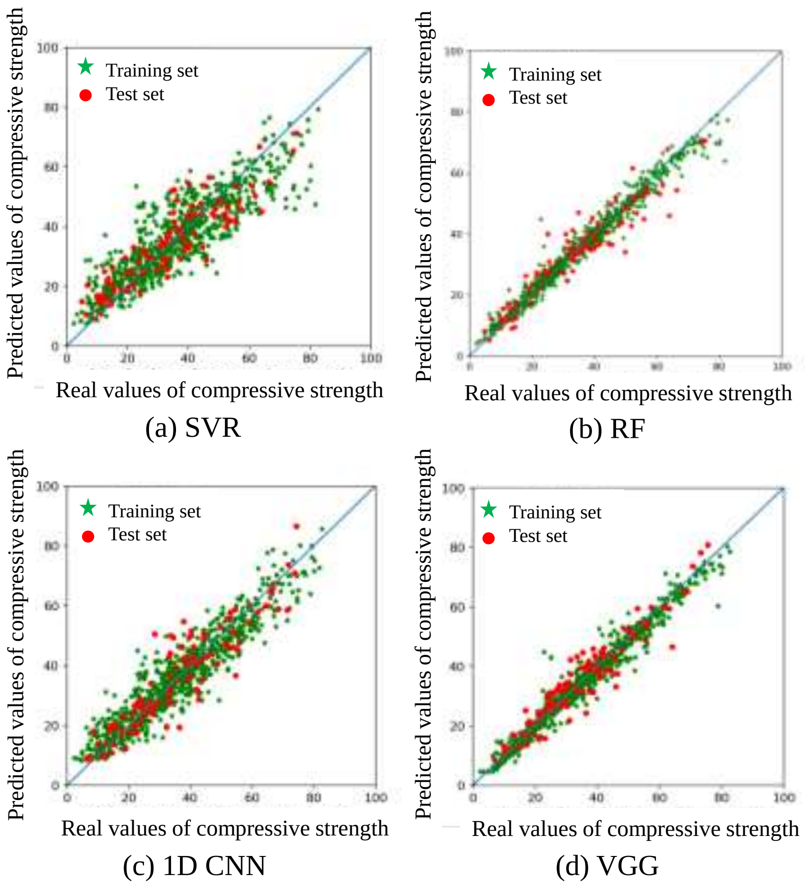


Fig. 5. Results before data augmentation

Overall, the prediction results of concrete compressive strength based on both machine learning and deep learning approaches perform well, as shown in Fig. 5. Obviously, RF based on decision trees and deep convolutional network structure perform higher prediction accuracy. Scatter plots of RF and VGG model show that the relationship between the test and predicted values is very close to the linear function $y = x$, especially on the training set where the green scatter points is closely coincided with the diagonal and the test set is slightly more disperse. Compared to the RF and VGG model, the scatter plot of the 1D CNN model does not fit well and its prediction is slightly worse. The worst prediction performance of these prediction models is the SVR model, which has large dispersity of scatter points and large deviations in coincidence can also be seen in the line plots generated from real and predicted data.

Table 2 shows the measurement of four indicators on the test set by different algorithms. Except SVR, the R^2 of other models are all above 0.8. Among them, the R^2 of VGG is 0.90, MSE and MAE have the lowest values, which are 26.41 and 3.38 respectively. Fig. 6 presents the effects of maximum iteration number and maximum depth blow root on the performances of the RF model. Moreover, RF also show good performance in predicting compressive strength, where R^2 = 0.89, MSE = 29.03 and MAE = 3.94.

Table 2 Performances of various models

ML Model	Performance			
	R^2	MSE	MAE	MAPE
SVR	0.78	58.08	6.11	23.89%
RF	0.89	29.03	3.94	13.82%
1D CNN	0.83	42.53	5.04	17.97%
VGG	0.90	26.41	3.38	13.85%

According to the analysis, the RF model and modified VGG model can output the predicted concrete compressive strength with very high precision. However, the deep learning methods including 1D CNN does not show its advantages for the original small dataset. The reason might be that one-dimensional deep convolutional structures require large amounts of dataset.

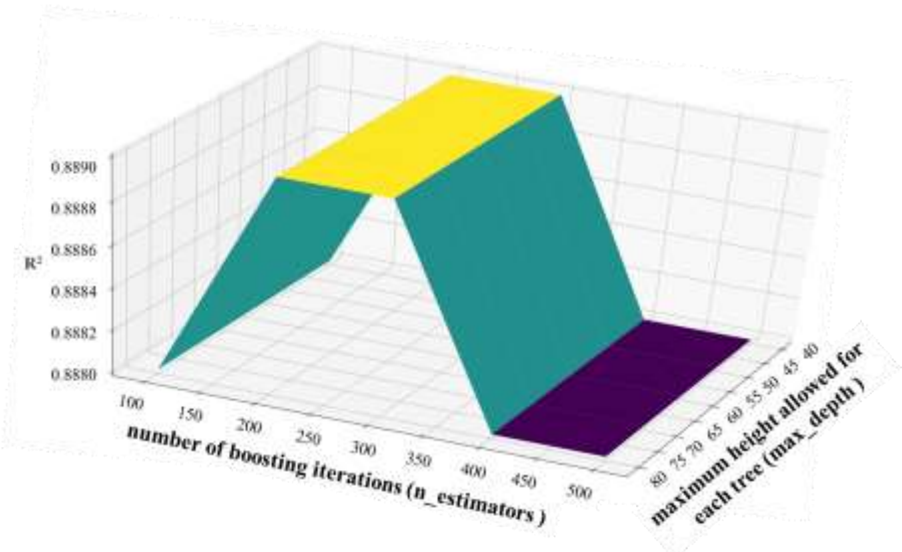


Fig. 6. Effects of maximum iteration number and maximum depth blow root on the performance (R^2) of RF in prediction of compression strength

To further demonstrate the prediction accuracy of the four machine learning models on the original small dataset, a 10-fold cross validation approach was utilized. The method involves equally dividing the experimental data into ten subsets, nine of which are used as the training set and the remaining one is used to validate the predictive ability of the models. After complete training cycles of the model for ten times, we can obtain a test result of the ten times and take the average of them.

It is clear from the Fig. 7 that although the results of 10 folds for RF, 1D CNN and VGG have some fluctuations, they all show a high average accuracy of 0.894, 0.885 and 0.923 respectively. Unfortunately, the mean accuracy of SVR is only 0.754. Therefore, the results of the 10-fold cross validation and the automatic dividing of the dataset are almost the same.

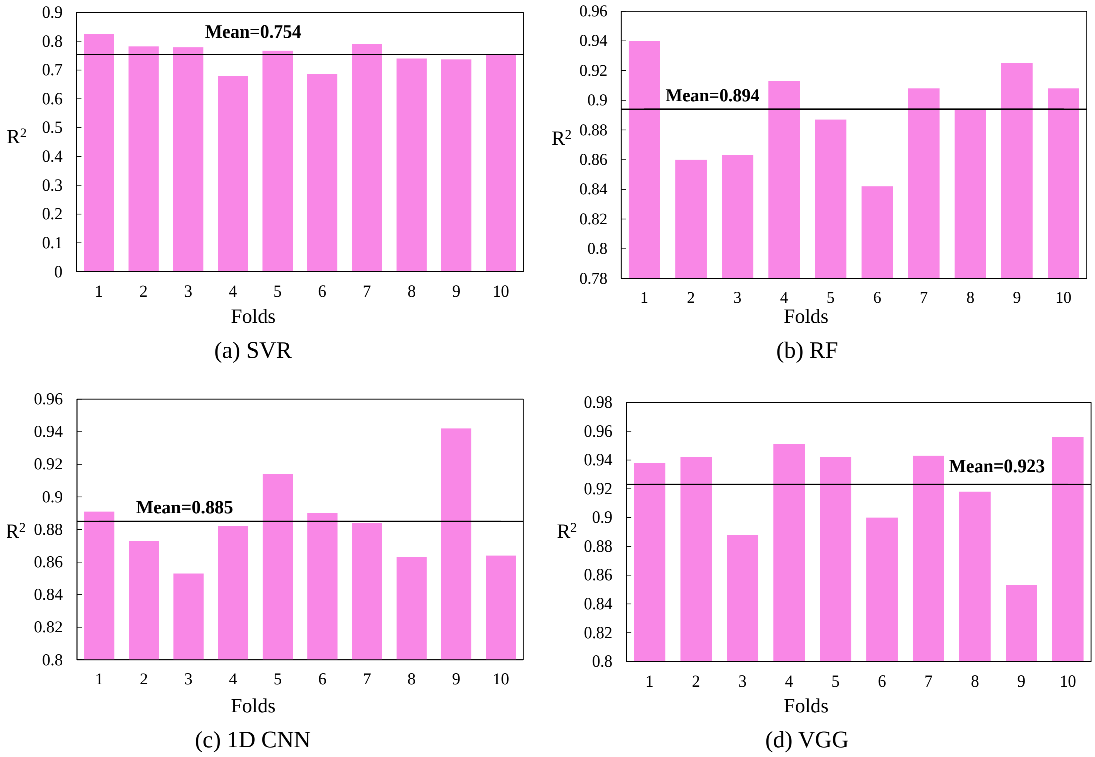


Fig. 7. Results of 10-fold cross validation

5. PERFORMANCE OF MODELS AFTER DATA AUGMENTATION

5.1 Sensitivity analysis of input variables

There are different effects of the concrete compressive strength according to different components. Among these components, two industrial wastes components, the blast furnace slag and fly ash, are considered in this study. Firstly, a sensitivity analysis of different input variables is conducted. And the results are shown in Fig. 8. As it can be seen, the cement and age are the two most important factors that affect the compressive strength. Admixtures like superplasticizer, blast furnace slag and fly ash effect less on the concrete compressive strength. It is concluded that, for the current test dataset, only blast furnace slag makes a great influence on the compressive strength of concrete, among the two kinds of industrial wastes. Therefore, we finally get seven inputs including cement, water, age, superplasticizer, blast furnace slag, fine aggregate and coarse aggregate.

The rankings of feature importance are shown in Fig. 9. Shapley Additive Explanation (SHAP) values indicate the impact of parameters on the model output, which is used to describe the specific impact of each feature on the predicted value of the model to achieve the purpose of explaining the model. It can be seen that the higher the SHAP values of age and cement, the more positive the impact on the final model output. The larger the water consumption is, the lower the predicted value of concrete compressive strength will be. The result that, the blast furnace slag from the two industrial wastes has made a more positive influence on the compressive strength of concrete, has been found.

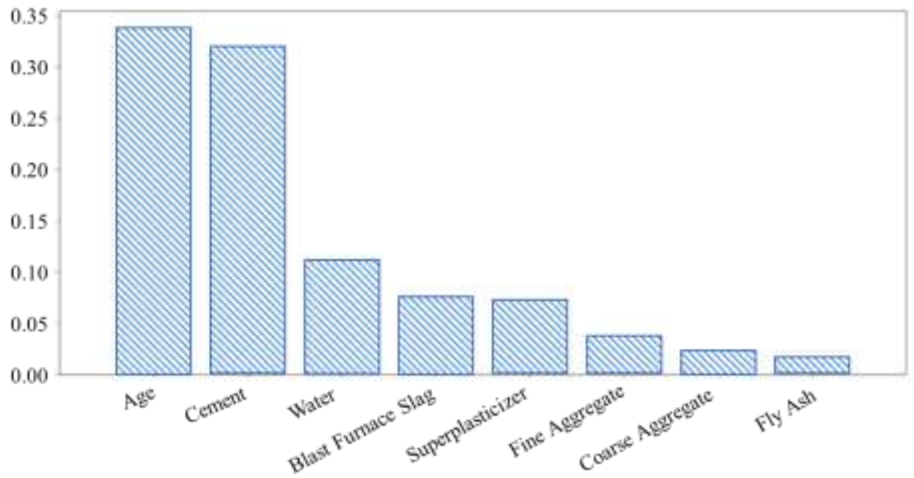


Fig. 8. Feature importance

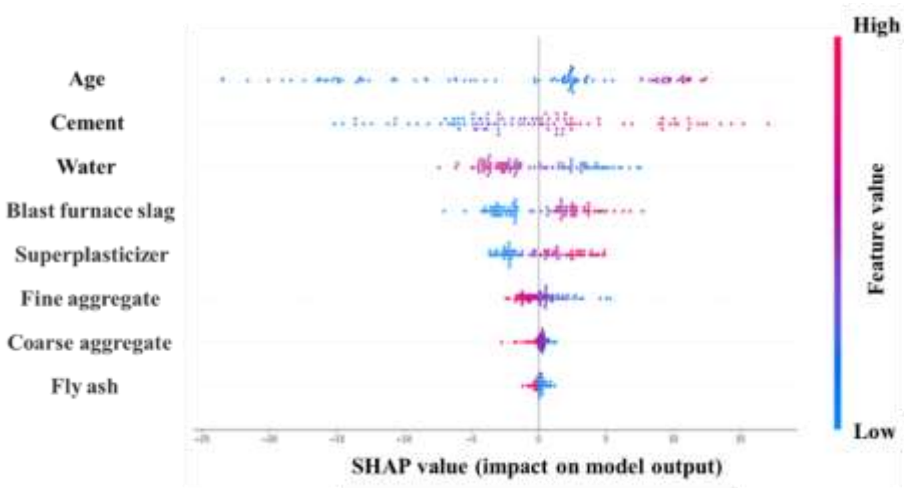


Fig. 9. SHAP beeswarm plot of the feature importance

5.2 Results after data augmentation

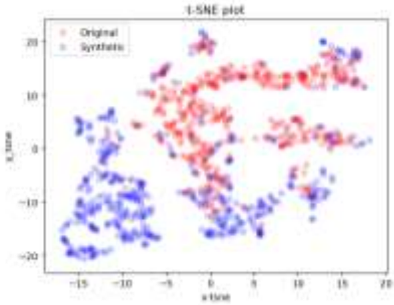
In the study, the existing data set was organized into several parts as 48×8 matrices. Then, DCGAN was used to gradually optimize the generator and discriminator through adversarial training, so as to generated high-quality new training data. In addition, the effects of data augmentation were compared with those of GAN and WGAN-GP. Because of the inevitable random noise in the training process, these new data as virtual test data may deviate from the real test data. After data augmentation, the size of dataset is augmented to 50208, which is large enough to predict the concrete compressive strength with satisfactory accuracy.

The original small dataset is augmented by three adversarial generative networks, GAN, DCGAN and WGAN-GP. As it is difficult to visualize the difference between the distribution of the augmented dataset and the original small dataset, we applied t-SNE (t-distributed stochastic neighbor embedding), a machine learning algorithm for data dimensionality reduction, for data visualization [49]. By comparing the variability of the data generated by the three GAN networks with the original data distribution, it can be seen that DCGAN is more suitable for data augmentation. The visualization results are shown in the Fig. 10. Red color represents the original data, and blue color represents the man-made data.

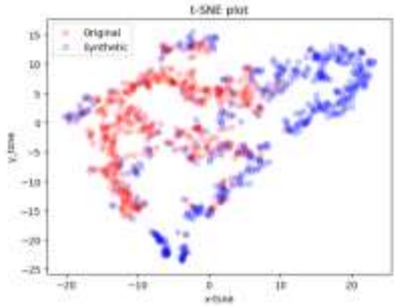
From Fig. 10, it can be found that the red and blue points are best fused by DCGAN, while only a small part of the GAN and WGAN-GP synthesized data overlap with the original data, and most of them are completely separated. The loss function (JS divergence) in DCGAN is not suitable for measuring the distance between the generated data distribution and the real data distribution, where WGAN-GP is an improvement on DCGAN that solves the problem of mode collapse occurring during training and can generate high quality images. However, in this study, WGAN-GP did not perform well in generating test data. Furthermore, DCGAN performed better than WGAN-GP. The possible reason may be the Wasserstein distance and Gradient Penalty item. Though Wasserstein distance and Gradient Penalty item are ways to balance the generator and discriminator training process, they also decrease the convergence rate relatively. In this study, as mode collapse had not occurred, Wasserstein distance and Gradient Penalty item may lead the WGAN network's convergence rate too slow to get a better result compared with DCGAN.

The DCGAN network is mainly composed of generator and discriminator. For the generator network, it consists of four convolution layers. The depths of these four layers are 128, 64, 32, 1 respectively. In addition, the first three convolution layers are added with up-sampling layer. To get the 48×8 dimension feature map to fit the discriminator's input, kernel dimension, stride and padding are designed. The tuple of layer, (Kernel Dimension, Stride, Padding), are (3×3, 1, 0), (3×3, 1, 0), and (3×3, 1, 1) respectively. And the changing process of feature map size is from 100 to 32×2 to 24×4 to 48×8 to 48×8. Finally, a 48×8 dimension of output vector has been gotten by generator network. For the discriminator network, it contains four convolution layers and three fully connected layers. The depths of these four convolution layers are 32, 64, 128, and 256 respectively. In addition, the structure of the convolution layers are (3×3, 2, 1), (3×3, 2, 2), and (3×3, 2, 1) as Kernel Dimension, Stride, and Padding. In the generator network, except the last layer uses *tanh* function, the other layers use ReLU activation function; in the discriminator network, all layers use the LeakyReLU activation function. Batch normalization is used to stabilize the learning and training process. To get a bool type output value for judging the input data with 48×8 dimension is true or false, the changing process of feature map size is from 48×8 to 24×4 to 12×2 to 7×2 to 4×1. Then the fully layers convert the input dimension from $4\times1\times256$ to 1024 to 64 to 1, as presented in Fig. 4. The network architecture of WGAN-GP is same to DCGAN.

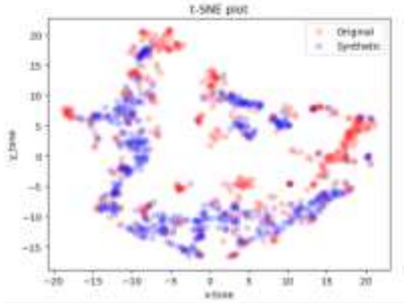
The augmented seven input variables and one output variable will be input into the prediction model for training. Fig. 11 shows the prediction results of each model after data augmentation.



(a) GAN



(b) WGAN-GP



(c) DCGAN

Fig. 10. t-SNE visualization

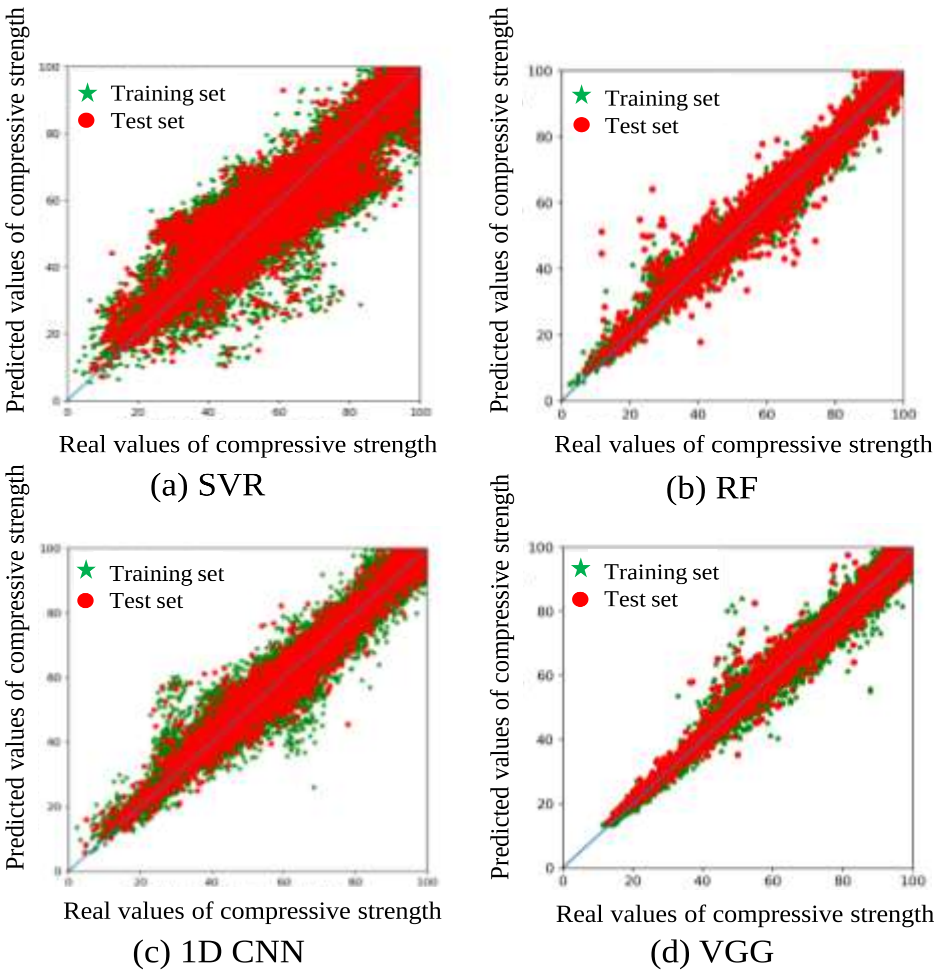


Fig. 11. Model result after data augmentation

The test results show that the prediction accuracy of the model is obviously improved after data augmentation. The red data points that closely fit on the diagonal represent the test set, which shows that both of the deep convolutional structures and ensemble algorithm have good prediction results on the test set. As shown in the Fig. 12, the accuracy of 1D CNN increases significantly, where R^2 is from 0.83 to 0.96; the accuracy of SVR does not increase significantly, where R^2 is only improved by 0.08. The reasons might be that the intrinsic mechanism of deep learning model require a significant large amount of data while the simple machine learning methods do not have such requirement. However, the improvement in accuracy of the RF model and a deep structure like VGG are not very significant may because the result is already excellent in the original dataset.

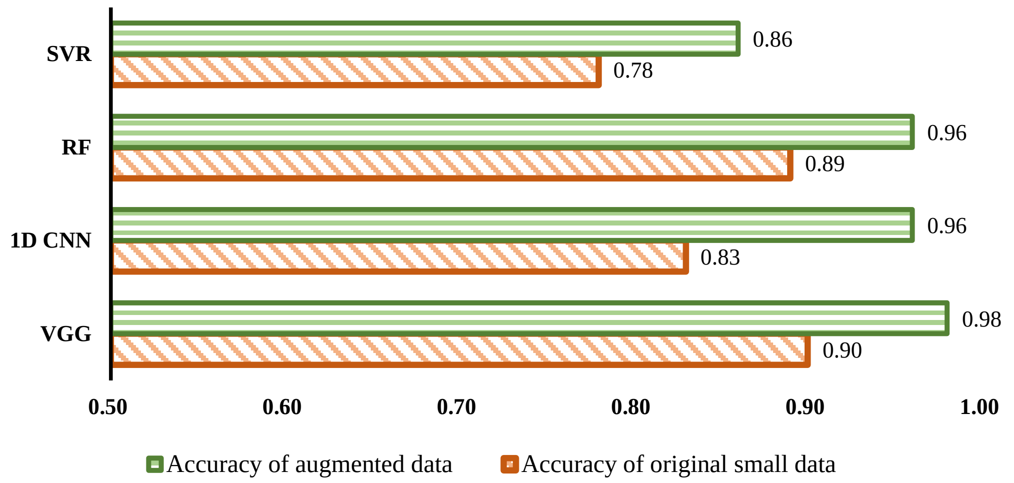


Fig. 12. Comparison of model accuracy before and after data augmentation

A new dataset was made by mixing 80 samples of new data provided by Yeh [50] and 90 samples of original data for deep learning model validation. The pre-trained model was obtained by training the 1D CNN and VGG on the original small dataset and the augmented dataset separately and then validated on the new hybrid dataset. This new dataset has 170 samples and each sample contained 8 variables i.e. Cement, Blast furnace slag, Water, Superplasticizer, Coarse aggregate, Fine aggregate, Age and Compressive strength.

The prediction results of 1D CNN and VGG on the new dataset are presented in Fig. 13. As can be seen, the proposed deep learning pre-trained model on augmented data improves the prediction accuracy by 0.16 for 1D CNN and 0.20 for VGG, respectively, compared to the pre-trained model on a smaller dataset. It can also be observed from the line Fig. 14 that the gap between the real data and the predicted data is significantly reduced for the two pre-trained models with different data scales. These results therefore suggest that data augmentation techniques can better improve prediction performance and robustness on deep learning models.

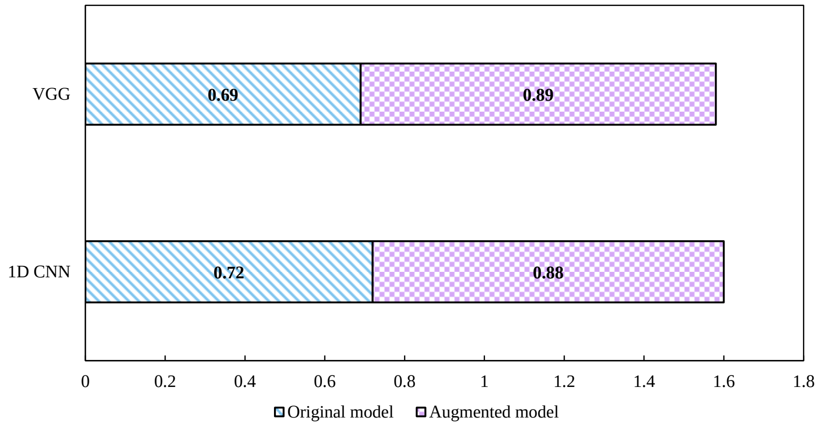
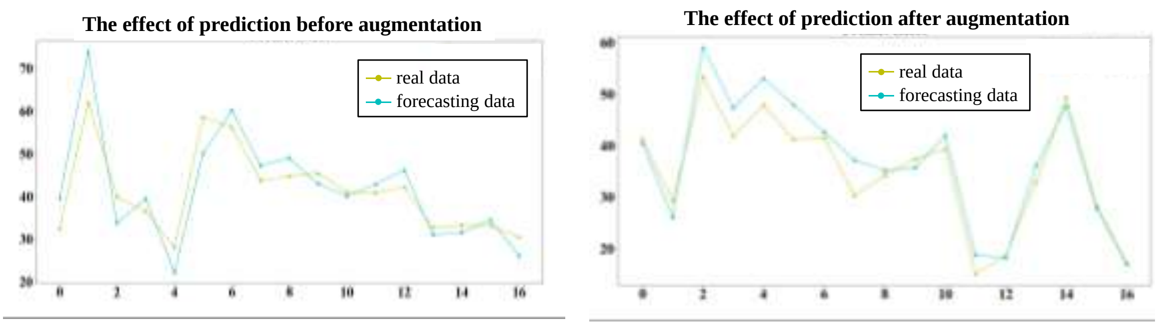
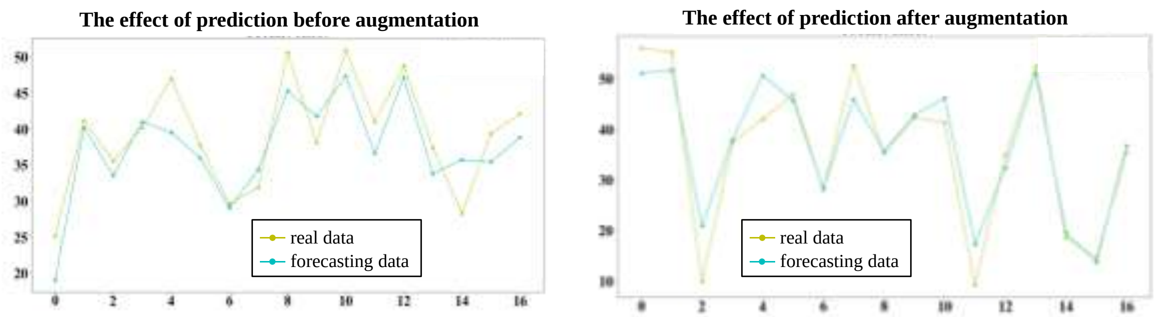


Fig. 13. Comparison of model accuracy before and after data augmentation in new dataset



(a) 1D CNN



(b) VGG

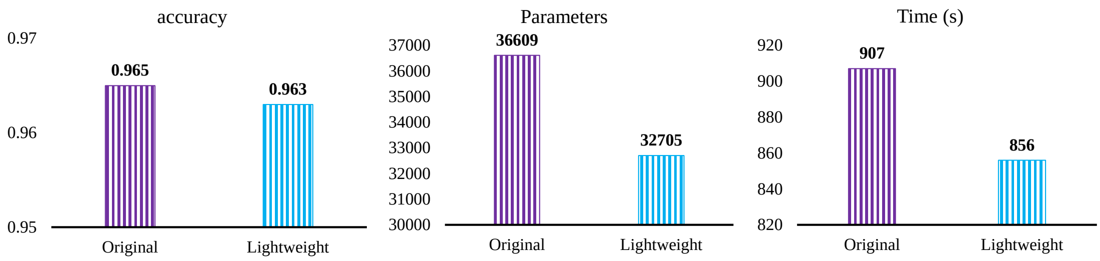
Fig. 14. Comparison of the effect of prediction for deep learning model before and after data augmentation

5.3 Performance based on lightweight CNNs

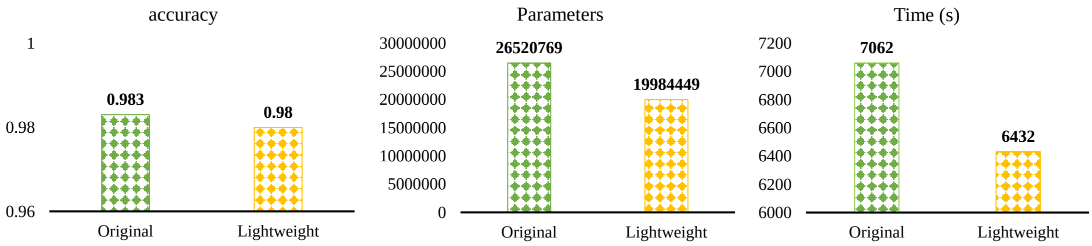
Convolutional neural network performs well on dealing with large amount of data, but the model hyperparameters are usually huge and the computation time is long. To improve the computation efficiency, the original four layer 1D CNN and VGG models were partially lightweight designed based on depthwise separable convolution in this study[51–53]. Specifically, each channel of the input image was convoluted by a filter, and then the convolution result of the first step was further convoluted by the point-to-point convolution.

In order to ensure that the model can reduce the amount of computation without excessively affecting its prediction accuracy, the convolution neural network did not need to carry out all lightweight design, only the standard 1D-convolution of the second and third layer were replaced by the depth separable 1D-convolution for 1D CNN and the standard 2D-convolution which was 256 and 512 kernels number were lightweight designed for VGG.

The augmented dataset was used to train the lightweight-designed convolutional neural networks, and compared with its original model in training time, model parameters number and prediction accuracy. The structure of network with 16 convolution kernels in the first layer has been used previously, but the learning rate reduction strategy which was also applied to VGG was improved. For every ten generations, the training of the model was checked. When the performance indicators of the model were no longer improved, the original learning rate was reduced to 75% of its value. By adjusting the number of convolution kernels of the initial model (with 16 convolution kernels in the first layer), it was found that the performance of the model with 64 convolution kernels in the first layer improved. Therefore, a model with 64 convolution kernels of first layer was chosen. Fig. 15 shows the differences between the lightweight model and the original model. It can be seen that the accuracy of the model decreases after lightweight design, but the decreases of computational time and parameters are more obvious. The computational time of 1D CNN dropped by approximately 5% from 907 to 856, while VGG (2D CNN) dropped by approximately 8% from 7062 to 6432. Thus, lightweight design with 2D CNN is more efficient. In addition, Fig. 16 shows the downward trend of MSE training in the initial stage of the lightweight model. Compared with the original model, the initial loss of the lightweight model is higher, but then the decreasing rate is higher. After a few iterations of training, the loss values (MSE) tends to coincide and have no significant differences.

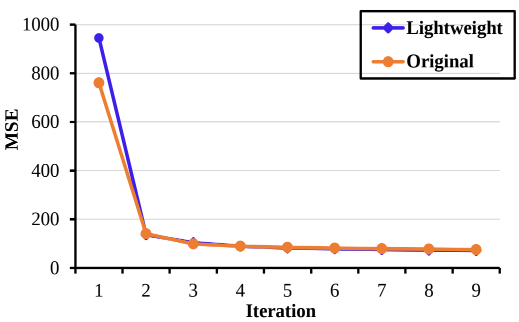


(a) 1D CNN

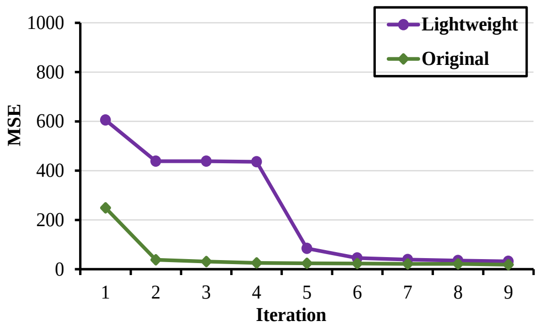


(a) VGG

Fig. 15. Comparison of accuracy, parameters and time of lightweight model and corresponding original model for 1D CNN and VGG



(a) 1D CNN



(b) VGG

Fig. 16. Comparison of loss (MSE) descent curves of lightweight model and corresponding original model for 1D CNN and VGG

In order to improve the prediction accuracy of lightweight 1D CNN, the model is further optimized. At the end of the training process, it is found that there is a great difference between the verification set and the training set, indicating the possible existing of over fitting problem. The reason may be that the data distribution changes greatly with the multi-layer operation of the network. Therefore, Batch Normalization (convolutional layers) and Dropout (fully connected layers) were added to solve the problem, which ensures the nonlinear expression ability of the model and speeds up the training speed.

Pooling layers can also alleviate the risk of over fitting to a certain extent, but excessive pooling layers will lead to some key information loss. For the prediction of concrete compressive strength, the nonlinear relationship between the input variables including cement content and water consumption, etc. is complex and changeable, which is closely related to the value of compressive strength. This will affect the final result of prediction, so it is reasonable to delete parts of the maximum pooling layer.

In the full connection layer, more units were added, where there exist 512 units in the first layer and 256 units in the second layer. It is helpful to aggregate the multiple feature information extracted from the convolution layer and output the predicted value. The number of iterations of model training is 400. As the performance of the model was no longer improved, after 366 generations, the training stopped. In the training set, the MSE is 11.52, MAE is 2.49 and MAPE is 4.56. Finally, the accuracy of the test set $R^2 = 0.97$ is improved by about 0.01.

6. CONCLUSIONS

In this study, a virtual material design of concrete with industrial wastes using deep data augmentation based on limited experiment data was proposed. The traditional machine learning methods and deep learning methods were used to predict the compressive strength of concrete materials. A total of 1030 pieces of concrete test data were used, including 8 input variables (Portland cement, water, coarse aggregate, fine aggregate, superplasticizer, blast furnace slag, fly ash and age) and 1 output variable (compressive strength). The DCGAN was used to perform deep augmentation on the original small dataset comparing to GAN and WGAN-GP. In addition, lightweight one-dimensional convolutional neural network and VGG were designed to improve the computation efficiency of the model while retaining the original prediction accuracy as far as possible. According to the experiment results, the following conclusions can be concluded:

(1) For the original small dataset, the RF model and VGG model both show good prediction performance on the test set. The R^2 is 0.89 for RF method and 0.90 for VGG method, which are higher than SVM ($R^2 = 0.78$) and 1D CNN ($R^2 = 0.83$).

(2) DCGAN was used to augment the original 1030 pieces of original data to 50208 pieces due to its good generation effect compare to GAN and WGAN-GP. For the augmented large dataset,

the deep learning method has better performance. R^2 of 1D CNN is dramatically increased from 0.83 to 0.96, while SVR is only increased by 0.08. Nevertheless, it should be noted that although it is possible to verify the effectiveness of data augmentation using new data, the current deep learning model developed in this study has limited generalization capabilities. In the future, we will continue to improve this model.

(3) The variable importance score of random forest showed that age and cement content are the two most important characteristics, where the influence of fly ash content is the lowest. Therefore, this variable was not introduced in the data augmentation process. It was also discovered through SHAP library that age, cement content and water consumption are considered to be highly important characteristics, in which age and cement content are positively correlated with the predicted value, while water consumption is negatively correlated with the predicted value. In addition, the effects of two industrial wastes on the mechanical properties of concrete materials were discussed. Although this study achieves a positive/negative effect of input parameters of the model on the concrete compressive strength based on the SHAP method, more detailed parametric studies are still required in future research to build a simplified intelligent analytical model.

(4) The lightweight convolution structure was applied to the original 1D CNN and VGG. After training on the augmented large dataset, it was found that the computational efficiency of the model is improved. The results also indicate a significant reduction in its parameters without a significant reduction in prediction accuracy.

ACKNOWLEDGEMENT

This work was supported by the International Research Cooperation Seed Fund of Beijing University of Technology (No. 2021A05), Opening project fund of Materials Service Safety Assessment Facilities (MSAF-2021-109), Talent Promotion Program by Beijing Association for Science and Technology, and the Construction of Service Capability of Scientific and Technological Innovation-Municipal Level of Fundamental Research Funds (Scientific Research Categories) of Beijing City.

REFERENCES

[1] Y. Wei, W. Kong, Y. Wang, Strengthening mechanism of fracture properties by nano materials for cementitious materials subject to early-age frost attack, Cement and Concrete Composites. 119 (2021) 104025. https://doi.org/10.1016/j.cemconcomp.2021.104025.

[2] S. Hansen, P. Sadeghian, Recycled gypsum powder from waste drywalls combined with fly ash for partial cement replacement in concrete, Journal of Cleaner Production. 274 (2020) 122785. https://doi.org/10.1016/j.jclepro.2020.122785.

[3] G.L. Vieira, J.Z. Schiavon, P.M. Borges, S.R. da Silva, J.J. de Oliveira Andrade, influence of recycled aggregate replacement and fly ash content in performance of pervious concrete mixtures, Journal of Cleaner Production. 271 (2020) 122665. https://doi.org/10.1016/j.jclepro.2020.122665.

[4] ASTM C39 / C39M–21, Standard Test Method for Compressive Strength of Cylindrical Concrete Specimens, ASTM International 2021 West Conshohocken, PA.

[5] D. Feng, X. Ren, J. Li, Stochastic damage hysteretic model for concrete based on micromechanical approach, International Journal of Non-Linear Mechanics. 83 (2016) 15–25. https://doi.org/10.1016/j.ijnonlinmec.2016.03.012.

[6] S. Liang, Y. Wei, Effects of water-to-cement ratio and curing age on microscopic creep and creep recovery of hardened cement pastes by microindentation, Cement and Concrete Composites. 113 (2020) 103619. https://doi.org/10.1016/j.cemconcomp.2020.103619.

[7] X. Shi, M. Mirsayar, A. Mukhopadhyay, D. Zollinger, Characterization of two-parameter fracture properties of portland cement concrete containing reclaimed asphalt pavement aggregates by semicircular bending specimens, Cement and Concrete Composites. 95 (2019) 56–69. https://doi.org/10.1016/j.cemconcomp.2018.10.013.

[8] Q.X. Lieu, K.T. Nguyen, K.D. Dang, S. Lee, J. Kang, J. Lee, An adaptive surrogate model to structural reliability analysis using deep neural network, Expert Systems with Applications. 189 (2022) 116104. https://doi.org/10.1016/j.eswa.2021.116104.

[9] S. Lee, S. Park, T. Kim, Q.X. Lieu, J. Lee, Damage quantification in truss structures by limited sensor-based surrogate model, Applied Acoustics. 172 (2021) 107547. https://doi.org/10.1016/j.apacoust.2020.107547.

[10] A.T. Huynh, Q.D. Nguyen, Q.L. Xuan, B. Magee, T. Chung, K.T. Tran, K.T. Nguyen, A Machine Learning-Assisted Numerical Predictor for Compressive Strength of Geopolymer Concrete Based on Experimental Data and Sensitivity Analysis, Appl. Sci.-Basel. 10 (2020) 7726. https://doi.org/10.3390/app10217726.

[11] S. Chithra, S.R.R.S. Kumar, K. Chinnaraju, F. Alfin Ashmita, A comparative study on the compressive strength prediction models for High Performance Concrete containing nano silica and copper slag using regression analysis and Artificial Neural Networks, Construction and Building Materials. 114 (2016) 528–535. https://doi.org/10.1016/j.conbuildmat.2016.03.214.

[12] A. Khashman, P. Akpinar, Non-Destructive Prediction of Concrete Compressive Strength Using Neural Networks, Procedia Computer Science. 108 (2017) 2358–2362. https://doi.org/10.1016/j.procs.2017.05.039.

[13] A. Behnood, V. Behnood, M. Modiri Gharehveran, K.E. Alyamac, Prediction of the compressive strength of normal and high-performance concretes using M5P model tree algorithm, Construction and Building Materials. 142 (2017) 199–207.

https://doi.org/10.1016/j.conbuildmat.2017.03.061.
[14] M.R. Kaloop, D. Kumar, P. Samui, J.W. Hu, D. Kim, Compressive strength prediction of high-performance concrete using gradient tree boosting machine, Construction and Building Materials. 264 (2020) 120198. https://doi.org/10.1016/j.conbuildmat.2020.120198.
[15] D.-C. Feng, Z.-T. Liu, X.-D. Wang, Y. Chen, J.-Q. Chang, D.-F. Wei, Z.-M. Jiang, Machine learning-based compressive strength prediction for concrete: An adaptive boosting approach, Construction and Building Materials. 230 (2020) 117000. https://doi.org/10.1016/j.conbuildmat.2019.117000.
[16] O.R. Abuodeh, J.A. Abdalla, R.A. Hawileh, Assessment of compressive strength of Ultra-high Performance Concrete using deep machine learning techniques, Applied Soft Computing. 95 (2020) 106552. https://doi.org/10.1016/j.asoc.2020.106552.
[17] J.A. Abdalla, A. Elsanosi, A. Abdelwahab, Modeling and simulation of shear resistance of R/C beams using artificial neural network, Journal of the Franklin Institute. 344 (2007) 741–756. https://doi.org/10.1016/j.jfranklin.2005.12.005.
[18] M.Z. Naser, V. Kodur, H.-T. Thai, R. Hawileh, J. Abdalla, V.V. Degtyarev, StructuresNet and FireNet: Benchmarking databases and machine learning algorithms in structural and fire engineering domains, Journal of Building Engineering. 44 (2021) 102977. https://doi.org/10.1016/j.jobe.2021.102977.
[19] M.Z. Naser, AI-based cognitive framework for evaluating response of concrete structures in extreme conditions, Engineering Applications of Artificial Intelligence. 81 (2019) 437–449. https://doi.org/10.1016/j.engappai.2019.03.004.
[20] M.Z. Naser, V.K. Kodur, Explainable machine learning using real, synthetic and augmented fire tests to predict fire resistance and spalling of RC columns, Engineering Structures. 253 (2022) 113824. https://doi.org/10.1016/j.engstruct.2021.113824.
[21] S.-Z. Chen, S.-Y. Zhang, W.-S. Han, G. Wu, Ensemble learning based approach for FRP-concrete bond strength prediction, Construction and Building Materials. 302 (2021) 124230. https://doi.org/10.1016/j.conbuildmat.2021.124230.
[22] M. Frid-Adar, I. Diamant, E. Klang, M. Amitai, J. Goldberger, H. Greenspan, GAN-based synthetic medical image augmentation for increased CNN performance in liver lesion classification, Neurocomputing. 321 (2018) 321–331. https://doi.org/10.1016/j.neucom.2018.09.013.
[23] K. Liu, R. Shuai, L. Ma, ZeXu, Cells image generation method based on VAE-SGAN, Procedia Computer Science. 183 (2021) 589–595. https://doi.org/10.1016/j.procs.2021.02.101.
[24] Z. Zhang, X. Pan, S. Jiang, P. Zhao, High-quality face image generation based on generative adversarial networks, Journal of Visual Communication and Image Representation. 71 (2020) 102719. https://doi.org/10.1016/j.jvcir.2019.102719.
[25] S. Pei, T. Shen, X. Wang, C. Gu, Z. Ning, X. Ye, N. Xiong, 3DACN: 3D Augmented convolutional network for time series data, Information Sciences. 513 (2020) 17–29. https://doi.org/10.1016/j.ins.2019.11.040.
[26] Y. Li, Q. Pan, S. Wang, T. Yang, E. Cambria, A Generative Model for category text generation, Information Sciences. 450 (2018) 301–315. https://doi.org/10.1016/j.ins.2018.03.050.
[27] I.-C. Yeh, Modeling of strength of high-performance concrete using artificial neural networks, Cement and Concrete Research. 28 (1998) 1797–1808. https://doi.org/10.1016/S0008-

8846(98)00165-3.

[28] G.C. Batista, D.L. Oliveira, O. Saotome, W.L.S. Silva, A low-power asynchronous hardware implementation of a novel SVM classifier, with an application in a speech recognition system, Microelectronics Journal. 105 (2020) 104907. https://doi.org/10.1016/j.mejo.2020.104907.

[29] C. Cortes, V. Vapnik, Support-Vector Networks, Mach. Learn. 20 (1995) 273–297. https://doi.org/10.1023/A:1022627411411.

[30] C. Leslie, E. Eskin, W.S. Noble, The spectrum kernel: a string kernel for SVM protein classification., Pacific Symposium on Biocomputing. Pacific Symposium on Biocomputing. (2002) 564–75.

[31] D.A. Otchere, T.O. Arbi Ganat, R. Gholami, S. Ridha, Application of supervised machine learning paradigms in the prediction of petroleum reservoir properties: Comparative analysis of ANN and SVM models, Journal of Petroleum Science and Engineering. 200 (2021) 108182. https://doi.org/10.1016/j.petrol.2020.108182.

[32] V.N. Vapnik, A.Y. Lerner, Recognition of patterns with help of generalized portraits, Avtomat. i Telemekh. 24 (1963) 774–780.

[33] L. Breiman, Random forests, Mach. Learn. 45 (2001) 5–32. https://doi.org/10.1023/A:1010933404324.

[34] H. Gong, Y. Sun, X. Shu, B. Huang, Use of random forests regression for predicting IRI of asphalt pavements, Constr. Build. Mater. 189 (2018) 890–897. https://doi.org/10.1016/j.conbuildmat.2018.09.017.

[35] T.K. Ho, Random decision forests, in: Proceedings of 3rd International Conference on Document Analysis and Recognition, 1995: pp. 278–282 vol.1. https://doi.org/10.1109/ICDAR.1995.598994.

[36] S. Abdollahpour, A. Kosari-Moghaddam, M. Bannayan, Prediction of wheat moisture content at harvest time through ANN and SVR modeling techniques, Information Processing in Agriculture. 7 (2020) 500–510. https://doi.org/10.1016/j.inpa.2020.01.003.

[37] K. Gopalakrishnan, S.K. Khaitan, A. Choudhary, A. Agrawal, Deep Convolutional Neural Networks with transfer learning for computer vision-based data-driven pavement distress detection, Constr. Build. Mater. 157 (2017) 322–330. https://doi.org/10.1016/j.conbuildmat.2017.09.110.

[38] Z. Tong, J. Gao, H. Zhang, Recognition, location, measurement, and 3D reconstruction of concealed cracks using convolutional neural networks, Constr. Build. Mater. 146 (2017) 775–787. https://doi.org/10.1016/j.conbuildmat.2017.04.097.

[39] A. Zhang, K.C.P. Wang, B. Li, E. Yang, X. Dai, Y. Peng, Y. Fei, Y. Liu, J.Q. Li, C. Chen, Automated Pixel-Level Pavement Crack Detection on 3D Asphalt Surfaces Using a Deep-Learning Network, Comput.-Aided Civil Infrastruct. Eng. 32 (2017) 805–819. https://doi.org/10.1111/mice.12297.

[40] S. Kiranyaz, T. Ince, R. Hamila, M. Gabbouj, Convolutional Neural Networks for patient-specific ECG classification, in: 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2015: pp. 2608–2611. https://doi.org/10.1109/EMBC.2015.7318926.

[41] O. Abdeljaber, O. Avci, M.S. Kiranyaz, B. Boashash, H. Sodano, D.J. Inman, 1-D CNNs for structural damage detection: Verification on a structural health monitoring benchmark data, Neurocomputing. 275 (2018) 1308–1317. https://doi.org/10.1016/j.neucom.2017.09.069.

[42] O. Abdeljaber, O. Avci, S. Kiranyaz, M. Gabbouj, D.J. Inman, Real-time vibration-based structural damage detection using one-dimensional convolutional neural networks, Journal of Sound and Vibration. 388 (2017) 154–170. https://doi.org/10.1016/j.jsv.2016.10.043.

[43] T. Sercu, C. Puhrsch, B. Kingsbury, Y. LeCun, Very deep multilingual convolutional neural networks for LVCSR, in: 2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2016: pp. 4955–4959. https://doi.org/10.1109/ICASSP.2016.7472620.

[44] S. Kiranyaz, O. Avci, O. Abdeljaber, T. Ince, M. Gabbouj, D.J. Inman, 1D convolutional neural networks and applications: A survey, Mechanical Systems and Signal Processing. 151 (2021) 107398. https://doi.org/10.1016/j.ymssp.2020.107398.

[45] K. Simonyan, A. Zisserman, Very Deep Convolutional Networks for Large-Scale Image Recognition, ArXiv:1409.1556 [Cs]. (2015). http://arxiv.org/abs/1409.1556 (accessed September 26, 2021).

[46] I.J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, Y. Bengio, Generative Adversarial Nets, in: Z. Ghahramani, M. Welling, C. Cortes, N.D. Lawrence, K.Q. Weinberger (Eds.), Advances in Neural Information Processing Systems 27 (Nips 2014), Neural Information Processing Systems (nips), La Jolla, 2014: pp. 2672–2680.

[47] Q. Cai, M. Abdel-Aty, J. Yuan, J. Lee, Y. Wu, Real-time crash prediction on expressways using deep generative models, Transportation Research Part C: Emerging Technologies. 117 (2020) 102697. https://doi.org/10.1016/j.trc.2020.102697.

[48] I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin, A. Courville, Improved Training of Wasserstein GANs, ArXiv:1704.00028 [Cs, Stat]. (2017). http://arxiv.org/abs/1704.00028 (accessed September 27, 2021).

[49] L. van der Maaten, G. Hinton, Visualizing Data using t-SNE, J. Mach. Learn. Res. 9 (2008) 2579–2605.

[50] I.-C. Yeh, Modeling slump flow of concrete using second-order regressions and artificial neural networks, Cement and Concrete Composites. 29 (2007) 474–480. https://doi.org/10.1016/j.cemconcomp.2007.02.001.

[51] Y. Hou, Q. Li, Q. Han, B. Peng, L. Wang, X. Gu, D. Wang, MobileCrack: Object Classification in Asphalt Pavements Using an Adaptive Lightweight Deep Learning, J. Transp. Eng. Pt. B-Pavements. 147 (2021) 04020092. https://doi.org/10.1061/JPEODX.0000245.

[52] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, L.-C. Chen, MobileNetV2: Inverted Residuals and Linear Bottlenecks, in: 31st Meeting of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, CVPR 2018, June 18, 2018 - June 22, 2018, IEEE Computer Society, Salt Lake City, UT, United states, 2018: pp. 4510–4520. https://doi.org/10.1109/CVPR.2018.00474.

[53] Y. Hou, Q. Li, C. Zhang, G. Lu, Z. Ye, Y. Chen, L. Wang, D. Cao, The State-of-the-Art Review on Applications of Intrusive Sensing, Image Processing Techniques, and Machine Learning Methods in Pavement Monitoring and Analysis, Engineering. 7 (2021) 845–856. https://doi.org/10.1016/j.eng.2020.07.030.