Hindawi Complexity Volume 2021, Article ID 5548988, 18 pages https://doi.org/10.1155/2021/5548988



Research Article

On the Training Algorithms for Artificial Neural Network in Predicting the Shear Strength of Deep Beams

Thuy-Anh Nguyen D, Hai-Bang Ly D, Hai-Van Thi Mai D, and Van Quan Tran D

University of Transport Technology, Hanoi 100000, Vietnam

Correspondence should be addressed to Thuy-Anh Nguyen; anhnt@utt.edu.vn

Received 17 February 2021; Revised 21 April 2021; Accepted 13 May 2021; Published 24 May 2021

Academic Editor: Zhen Zhang

Copyright © 2021 Thuy-Anh Nguyen et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This study aims to predict the shear strength of reinforced concrete (RC) deep beams based on artificial neural network (ANN) using four training algorithms, namely, Levenberg–Marquardt (ANN-LM), quasi-Newton method (ANN-QN), conjugate gradient (ANN-CG), and gradient descent (ANN-GD). A database containing 106 results of RC deep beam shear strength tests is collected and used to investigate the performance of the four proposed algorithms. The ANN training phase uses 70% of data, randomly taken from the collected dataset, whereas the remaining 30% of data are used for the algorithms' evaluation process. The ANN structure consists of an input layer with 9 neurons corresponding to 9 input parameters, a hidden layer of 10 neurons, and an output layer with 1 neuron representing the shear strength of RC deep beams. The performance evaluation of the models is performed using statistical criteria, including the correlation coefficient (R), root mean square error (RMSE), mean absolute error (MAE), and mean absolute percentage error (MAPE). The results show that the ANN-CG model has the best prediction performance with R = 0.992, RMSE = 14.02, MAE = 14.24, and MAPE = 6.84. The results of this study show that the ANN-CG model can accurately predict the shear strength of RC deep beams, representing a promising and useful alternative design solution for structural engineers.

1. Introduction

Deep beams are defined as load-bearing structural elements in the form of simple beams, in which a considerable amount of load is transferred to the supports by a combined compression force of load and jet. Deep beams are characterized by a larger beam depth compared to conventional beams, classified by the ratio of the length of the cut span to the beam depth (a/h) or on the ratio between calculated span length and beam height (l/h). Several design codes have given the conditions for defining deep beams. For instance, according to IS Code 456-2000, the deep beam is defined by a ratio of effective span-to-overall depth (l/h), which does not exceed 2.0 for the simple beam and 2.5 for the continuous beam [1]. Besides, the ACI 318-14 [2] classifies a beam as a deep beam if it satisfies the following: (a) the spacing does not exceed four times of overall structural depth, or (b) the cutting span does not exceed twice the overall part depth.

According to Eurocode 2 (EC2) [3], when the ratio l/h is less than three, the beam is considered a deep beam.

Currently, RC deep beams are widely used in structural works, such as transfer beam, wall foundation, foundation pile cap, floor partition wall, and shear wall [4]. In particular, deep beams play a crucial role in the design of large structures as well as small structures. In several specific cases for architectural purposes, the buildings are designed without using any columns for a very large span. In this case, if normal beams are used, failures such as bending failures might occur. So using deep beams is an effective solution that could increase the durability of structures [5–7]. Due to the large height of deep beams, the primary type of damage is shear damage [8-10]. For deep beams, cracks often appear quite early, in the direction of primary compressive stress, perpendicular to the direction of tensile stress. In many cases, the crack appears vertical or inclined when the beam is damaged by shear force. This leads to a sudden malfunction

of the beam when the beam's height increases [11]. In deep beams, the shear capacity can be 2 to 3 times greater than that determined by the calculation method obtained with conventional beams. Therefore, the shear stress in the high beam cannot be ignored compared with the conventional bending beam. The stress distribution is not linear even in the elastic phase. At given ultimate stress, the stress field is not the same parabolic shape as the conventional beams anymore, which is also a significant reason for slippage problems in deep beams [5].

In the past several decades, many methods have been proposed to analyze the shear strength of deep beams, including the strut-and-tie method (STM) [12, 13] and the upper limit theorem of plasticity theory [14, 15]. Based on the STM, theoretical methods to calculate the shear strength are proposed, such as compression field theory (CFT) and modified compression field theory (MCFT) [16, 17], the theory of softened strut-and-tie model (SSTM) considering the compression softening of concrete [18, 19], and strutand-tie model based on the crack band theory [20, 21]. Besides, the current design codes, such as ACI 318–14 [2], EN 1992-1-1:2004 [3], and CSA A23.3-04 [22], have recommended the STM approach as a deep beam design tool. In addition, some in-depth studies have been carried out to analyze the shear behavior of deep beams as well as determine the most critical parameters affecting the shear strength. According to studies [4, 23-25], several important parameters have been identified, including compressive strength of concrete, yield strength of longitudinal and transverse reinforcement, the ratio of effective depth to breadth, as well as the main reinforcement ratio. In fact, the relationship between the parameters and the shear capacity of deep beams is nonlinear [9, 12, 23]. Consequently, building an accurate model that can accurately estimate shear strength based on mathematical equations is challenging [26]. Meanwhile, the deep beam shear strength obtained by experimental tests or numerical analysis is more or less limited because of the complexity of such kind of material and beam structure [12, 23]. To overcome these difficulties and to improve the ability to estimate the shear strength of deep beams, artificial intelligence (AI) approaches have been used in several investigations [27, 28].

Indeed, the construction field has effectively applied AI models to solve many problems such as geotechnical [29, 30], building materials [31, 32], structure analysis, and design [33–35]. The application of AI models for problems related to the shear strength of deep beams has been studied by many scientists. Goh's first study in 1995 [36] applied the artificial neural network (ANN) model to predict beam shear resistance with 6 input parameters. Later, Sanad and Saka [37] also checked the effectiveness of the ANN model in predicting the shear strength of deep beams using 10 input parameters related to the geometry and material properties. The results showed that ANN provides an effective alternative solution in predicting the shear artificial neural network of reinforced concrete (RC) deep beams. It is obvious that the ANN algorithm is a widely used machine learning (ML) prediction tool, but the selection of an appropriate ANN algorithm is still being questioned. In fact, it

is challenging to find the best ANN model that could accurately predict the target and optimize many factors, such as the processing speed, numerical precision, and memory requirements. Such an optimization problem lies in the learning process in a neural network and could be solved by using an appropriate training algorithm. In fact, the ANN algorithm contains four principal training algorithms, including Levenberg-Marquardt (ANN-LM), quasi-Newton method (ANN-QN), conjugate gradient (ANN-CG), and gradient descent (ANN-GD). A given training algorithm might be suitable for a given problem but might fail in another case [38]. Gradient descent is the slowest training algorithm but requires less memory than the other three algorithms. The fastest algorithm is Levenberg-Marquardt, which requires the most memory. Therefore, an in-depth investigation is crucial to determine the best training algorithm in general and in predicting the shear strength of deep beams in particular. Besides, the basis of selecting the best ANN black-box raises a number of fundamental questions, especially the criterion to define the best one. In the field of ML, the performance evaluation of the models is assessed by different metrics [39-41], namely, the correlation coefficient (R) or the coefficient of determination (R^2) , mean absolute error (MAE), root mean square error (RMSE), and mean absolute percentage error (MAPE). A concise evaluation and comparison of different criteria need to be conducted to confirm ML models' effectiveness.

Therefore, in this study, the procedure to determine the best ANN algorithm is conducted through different ANN training algorithms and evaluation metrics, with the highest aim is to accurately and reliably predict the shear strength of the deep beam. To achieve this goal, in the first step, the construction of the deep beam database is conducted by gathering different experimental results published in the literature. The general theory of ANN models is then presented, including four previously mentioned training algorithms. An architecture of ANN models is proposed, along with an extensive investigation on the ANN epoch numbers. The best ANN algorithm is deduced by comparing different performance metrics and the corresponding probability density functions, taking into account the random sampling effect, while constructing the two datasets. Finally, the representative results in predicting the shear strength of deep beams are presented and compared with several existing prediction results in the available literature.

2. Significance of the Research Study

Accurate prediction of the deep beam shear strength is crucial in the construction design. Although some machine learning models have been proposed to predict the shear strength of deep beams in the available literature, namely, genetic-simulated annealing [4], backpropagation neural network [42], artificial neural network [43], gene expression programming [43], support vector machine [42], multivariate adaptive regression splines [42], smart artificial firefly colony algorithm and least squares support vector regression [24], and adaptive neural fuzzy inference system [44], the prediction accuracy and reliability could be further

improved. Therefore, different contributions of the present investigation could be pointed out by the following ideas:

- (1) Four representative training algorithms for the ANN model are investigated to predict the shear strength of deep beams, in which the training epoch of each model is fine-tuned.
- (2) The reliability of ANN models is carefully evaluated by Monte Carlo simulations with random sampling strategy to construct the database.
- (3) The model using the conjugate gradient algorithm (ANN-CG) containing 10 neurons in the hidden layer is deduced as the best predictor.
- (4) The performance of the best ANN-CG architecture is compared with 10 previously published works in the literature and achieved the highest value of the correlation coefficient (*R*) and lowest values of mean absolute error (MAE). Thus, the simplicity and effectiveness of the proposed approach using ANN-CG are confirmed.

3. Database Construction

In this study, the database used to develop the ML models is collected from published research. The dataset includes 106 test results of the shear strength of deep beams. Specifically, 19 test results of high-strength RC deep beams are collected in the study by Tan et al. [45], 52 test results from the work of Smith et al. [46], and 35 test results from the work of Kong et al. [47]. This database includes various parameters affecting the shear strength of RC deep beams (denoted as V), including the ratio of effective span to effective depth (L/d), ratio of effective depth to breadth (d/b_w) , ratio of shear span to effective depth (a/d), concrete cylinder strength (f_c) , yield strength of horizontal reinforcement (f_{yh}), yield strength of vertical web reinforcement (f_{yy}), ratio of horizontal web reinforcement (ρ_h), ratio of longitudinal reinforcement to concrete area (ρ_s), and ratio of vertical web reinforcement $(\rho_{\rm v})$. Representative information on these parameters is detailed in Table 1. Besides, the histograms of each input and output parameter are shown in Figure 1. The beam test diagram and a schematic illustration of RC deep beams are illustrated in Figure 2. Prior to the training process of ANN models, all input and output values are normalized in the range of [0, 1] and then converted back to the initial range of values for the sake of clarity and postprocessing processes.

The database is randomly divided into two parts, representing the generation of the random sampling effect. The first part (containing 70% of the total data, 74 samples) is used to train the ANN network, called the training part. The second part (using the remaining 30% of data, 32 samples) is used to verify the ANN models, referred to as the testing part. The random sampling effect generates variability in the input space of the training part, which considerably affects the accuracy of the ML models. Besides, the meaning of separating the training and testing parts in machine learning problems is to fully assess the accuracy of the models, as the testing data are entirely unknown to the model during the

training phase. In general, the prediction capacity of the model is the most important factor. Therefore, the results in the next sections only focus on the evaluation criteria of the testing parts.

4. General Presentation of ANN Models

An artificial neural network (ANN) is a computational model that is built based on the human brain with many biological neurons. It consists of many artificial neurons, interconnected in a network, including input and output data. From input data to come up with a complete result or output, a set of learning rules is used. It is called the backpropagation or backward propagation of error. The structure of a backpropagation network is a combination of different layers, including the input layer, the output layer, and the hidden layer. The input layer is the first layer, the output layer is the last one, and the connection between the two layers is the hidden layer, which might contain one or many hidden layers (Figure 3).

During the training phase of the algorithm, ANN learns to recognize patterns from the input data. Then, it compares the produced result with the desired output. The difference between the two results is adjusted through a backward working process until such a difference is lower than a predefined criterion. Therefore, to train a neural network, the selection of an appropriate training algorithm is very important. The training algorithms are the underlying engines for building neural network models with the goal of training features or patterns from the input data so that a set of internal model parameters can be found to optimize the model's accuracy. There are many types of training algorithms, but frequently used ones can be listed as gradient descent, conjugate gradient, quasi-Newton method, and Levenberg–Marquardt algorithms.

4.1. Training Algorithms of ANN Model

4.1.1. Gradient Descent Algorithm (ANN-GD). Gradient descent is an iterative optimization algorithm used in ML and deep learning problems with the goal of finding a set of internal variables for model optimization. Inside, the "gradient" is the rate of inclination or declination of a slope, and the "descent" means descending. Gradient descent often performs in 3 steps, namely, (1) internal variable initialization, (2) evaluating the model based on the internal variable and loss function, and (3) updating internal variables in the direction of finding optimal points. The gradient descent method possesses the iteration step by

$$\mathbf{w}^{(i+1)} = \mathbf{w}^{(i)} - \nabla f(\mathbf{w}^{(i)}) \eta^{(i)}, \tag{1}$$

where $\mathbf{w}^{(i)}$ is the set of variables to be updated, $\nabla f(\mathbf{w}^{(i)})$ is the gradient of the loss function f according to set $\mathbf{w}^{(i)}$, η is the training rate, and $i = 0, 1, ..., \eta$ can be a fixed value or determined by one-dimensional optimization along the training direction per step. The nature of the optimization process of the loss function is finding the suitable points to minimize or maximize the loss function. The goal of the

Table 1: Summary of the input and output variables of deep beams used in this study.

	Notation	Unit	Min	Median	Average	Max	Std ^a	SK ^b
Ratio of effective span to effective depth	L/d	_	1.05	3.08	2.88	5.38	1.12	0.08
Ratio of effective depth to breadth	$d/b_{ m w}$	_	2.84	2.99	4.44	10.03	2.24	1.49
Ratio of shear span to effective depth	a/d	_	0.27	1.00	1.01	2.70	0.049	0.40
Concrete cylinder strength	$f_{\rm c}$	$10^{-1} \text{ x kN/mm}^2$	0.16	0.21	0.26	0.59	0.11	16.91
Yield strength of horizontal reinforcement	$f_{ m yh}$	kN/mm ²	0.00	0.48	0.40	0.50	0.15	-1.60
Yield strength of vertical web reinforcement	f_{yy}	kN/mm ²	0.00	0.38	0.35	0.48	0.18	-1.11
Ratio of horizontal web reinforcement	$ ho_{ m h}$	10^{-2}	0.00	0.45	0.49	2.45	0.56	193.12
Ratio of longitudinal reinforcement to concrete area	$ ho_{ m s}$	_	0.00	0.01	0.01	0.02	0.00	-0.48
Ratio of vertical web reinforcement	$ ho_{ m v}$	10^{-2}	0.00	0.48	0.51	2.45	0.54	213.04
Shear strength	V	kN	74.00	161.00	203.58	675.00	121.93	2.20

^aStandard deviation. ^bSkewness.

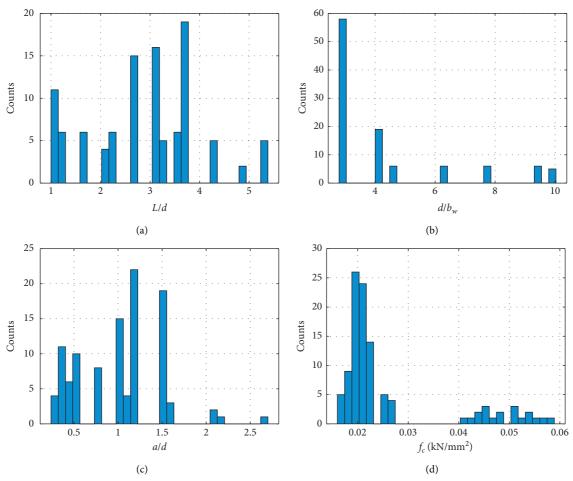


FIGURE 1: Continued.

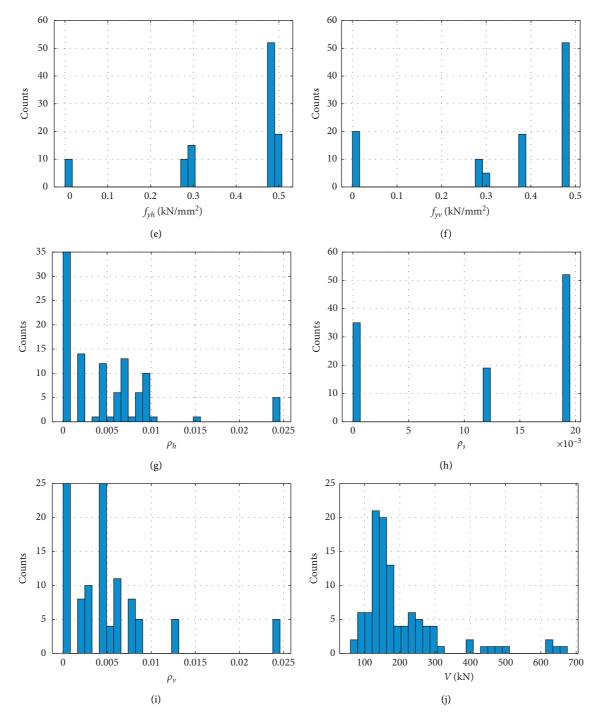


FIGURE 1: Histograms of input and output variables used in this study: (a) ratio of effective span to effective depth; (b) ratio of effective depth to breadth; (c) ratio of shear span to effective depth; (d) concrete cylinder strength; (e) yield strength of horizontal reinforcement; (f) yield strength of vertical web reinforcement; (g) ratio of horizontal web reinforcement; (h) ratio of longitudinal reinforcement to concrete area; (i) ratio of vertical web reinforcement; (j) deep beams' shear strength.

gradient descent method is to find such global minimum points. The stopping criterion of the gradient descent method can be (i) the maximum number of epochs reached, (ii) the value of the loss function is small enough, and the accuracy of the model is large enough, and (iii) the value of the loss function remains stable after a finite number of epochs. The gradient descent algorithm is often used with the big neural networks. The advantage of this method lies in

the storage of the gradient vector, instead of the Hessian matrix. The diagram for the training process with the gradient descent is shown in Figure 4.

4.1.2. Conjugate Gradient Algorithm (ANN-CG). The conjugate gradient algorithm could be considered as one of the algorithms to improve the convergence rate of the artificial

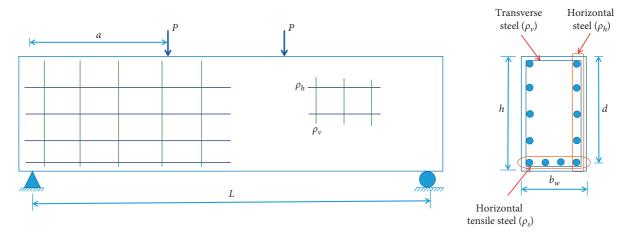


FIGURE 2: Schematic illustration of reinforced concrete deep beams.

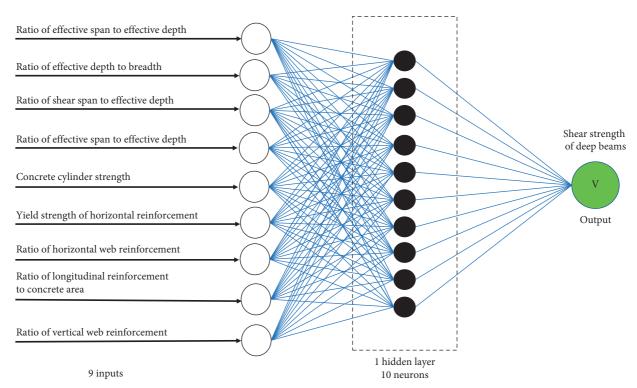


FIGURE 3: ANN structure used in this study.

neural network, being the intermediate between gradient descent and Newton's method. The advantage of this approach lies in the fact that there is no need to evaluate, store, and reverse the Hessian matrix. In this algorithm, the search is performed along with conjugate directions, which produce generally faster convergence than gradient descent directions. These training directions are conjugated concerning the Hessian matrix. In this algorithm, the sequence of training directions is built using the following formula:

$$y^{(i+1)} = v^{(i+1)} + y^{(i)}c^{(i)},$$
 (2)

with the initial training direction vector

$$y^{(0)} = -v^{(0)},\tag{3}$$

where y is the training direction vector, c is the conjugate parameter, and i = 0, 1, ...

The training direction, in all the cases, is reset to the gradient's negative [48]. The parameters' improvement process with the conjugate gradient algorithm is defined by

$$w^{(i+1)} = w^{(i)} + y^{(i)} \eta^{(i)}, (4)$$

where $i = 0, 1, ..., \eta$ is the training rate, usually found by line minimization. The diagram for the training process with the conjugate gradient is shown in Figure 4.

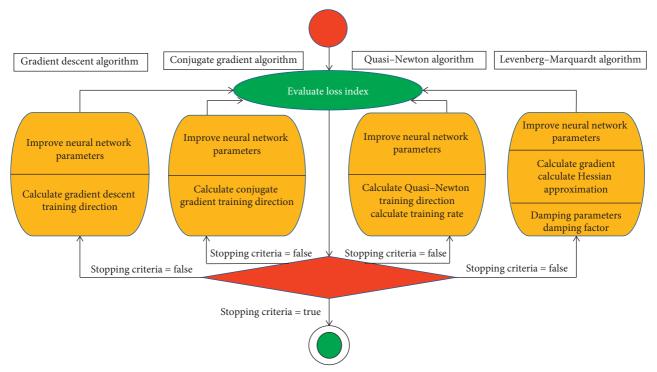


FIGURE 4: Diagrams of the training algorithms of the ANN model.

4.1.3. Quasi-Newton Algorithm (ANN-QN). The advantage of the quasi-Newton method is that it is computationally inexpensive because it does not need many operations to evaluate the Hessian matrix and calculate the corresponding inverse. An approximation value to the inverse Hessian matrix is built at each iteration. It is computed using only information on the first derivatives of the loss function. The Hessian matrix is composed of the second partial derivatives of the loss function. The quasi-Newton formula is presented by

$$w^{(i+1)} = w^{(i)} - (G^{(i)}g^{(i)})\eta^{(i)}, \tag{5}$$

where $G^{(i)}$ is the inverse Hessian approximation. The quasi-Newton method is commonly used because it is faster than gradient descent and conjugate gradient. The diagram of the quasi-Newton method is shown in Figure 4.

4.1.4. Levenberg-Marquardt Algorithm (ANN-LM). The Levenberg-Marquardt (LM) algorithm, also called the damped least squares method, is used to solve nonlinear least squares problems. Instead of computing the exact Hessian matrix, this algorithm calculates with the gradient vector and the Jacobian matrix. The loss function is expressed as a sum of squared errors as

$$f = \sum_{i=1}^{a} u_i^2,$$
 (6)

with a is the number of instances in the dataset and u is the vector of all error terms. The Jacobian matrix of the loss function is defined as follows:

$$A_{i,j} = \frac{\partial u_i}{\partial w_i},\tag{7}$$

for i = 1, ..., a and j = 1, ..., b and a is the number of instances in the dataset, b is the number of parameters in the neural network, and A is the Jacobian matrix. The size of the Jacobian matrix is [a, b]. The gradient vector of the loss function is calculated as

$$\nabla f = 2A^T . u. \tag{8}$$

The Hessian matrix is approximately computed by

$$Bf \approx 2A^T.A + \beta.I,\tag{9}$$

where B is the Hessian matrix, β is a damping factor that ensures the positive of the Hessian, and I is the identity matrix. The large parameter β is chosen in the first step. Next, if there is an error in any iteration, β will be increased by some factor. On the contrary, if the loss decreases, β will be decreased so that the Levenberg–Marquardt algorithm approaches the Newton method. Finally, the parameters' improvement process using the Levenberg–Marquardt algorithm is defined as

$$w^{(i+1)} = w^{(i)} - (A^{(i)T}.A^{(i)} + \beta^{(i)}I)^{-1} (2A^{(i)T}.u^{(i)}).$$
 (10)

for i = 0, 1, ...

The diagram of the ANN-LM training algorithms is shown in Figure 4.

4.2. Validation of Models. To evaluate the performance of the machine learning models, in this investigation, four indexes,

namely, root mean square error (RMSE), mean absolute error (MAE), mean absolute percentage error (MAPE), and correlation coefficient (R) are used. The RMSE is used to evaluate the difference between the actual and predicted values. MAE shows the average error of the actual and predicted values. The MAPE is defined as the difference between the actual and predicted values and then divided by the actual value. Specifically, the lower the RMSE, MAE, and MAPE values, the higher the accuracy of the models and the better the performance of the models. On the contrary, the higher R values mean higher model performance. The R value varies in the range from -1 to 1. The R values close to 0 show the poor performance of the model and close to 1 means good accuracy. The values of RMSE, MAE, MAPE, and R are defined by the following formulas:

$$RMSE = \sqrt{\frac{1}{N} \sum_{j=1}^{N} (Q_{j,AV} - Q_{j,PV})^{2}},$$

$$MAE = \frac{1}{N} \sum_{j=1}^{N} |Q_{j,AV} - Q_{j,PV}|,$$

$$MAPE = \frac{1}{N} \sum_{j=1}^{N} \frac{|Q_{j,AV} - Q_{j,PV}|}{Q_{j,AV}} \times 100\%,$$

$$R = \frac{\sum_{j=1}^{N} (Q_{j,AV} - \overline{Q}_{AV})(Q_{j,PV} - \overline{Q}_{PV})}{\sum_{j=1}^{N} (Q_{j,AV} - \overline{Q}_{AV})^{2} \sum_{j=1}^{N} (Q_{j,PV} - \overline{Q}_{PV})^{2}},$$
(1

where Q_{AV} and \overline{Q}_{AV} are the actual and the average values and Q_{PV} and \overline{Q}_{PV} are the predicted and the average predicted values.

5. Methodology Flowchart

In this study, the flowchart of the proposed methodology includes the following steps:

- (a) Data collection: this is the first step, and the dataset is built by gathering data from the available literature. All data are randomly divided into 2 parts: training data and testing data, in which the training part accounts for 70% of the dataset and the testing part accounts for 30% of the dataset.
- (b) Building models: in this step, the data of the training part was used for training the models based on training algorithms such as gradient descent, conjugate gradient, quasi-Newton, and Levenberg-Marquardt.
- (c) Model validation: in this final step, the data of the testing part is applied to validate the proposed models. Statistical indicators including RMSE, MAE, MAPE, and *R* are utilized to evaluate the models.

A schematic diagram of the methodology is illustrated in Figure 5.

6. Results and Discussion

The definition of the ANN structure is critical in solving problems [49, 50]. In the case that the number of input and output is fixed, the performance of the ANN model depends on the hidden layer number and the neuron number in each hidden layer. Cybenko [51] and Bound [52] have succeeded in using a single hidden layer model in classifying the input variables for model processing. Besides, some studies [53-55] have shown that an ANN model with only one hidden layer could be enough to successfully explore a complex nonlinear relationship between input(s) and output. Therefore, one hidden layer is proposed for the structure of the ANN model in this investigation. Moreover, semiempirical relationships proposed by Nagendra [56], Tamura [57], and some investigations [58-60] have recommended that the neuron number of the hidden layer is equal to the total number of inputs and outputs. In the current database, the number of input and output representing deep beams' shear strength is equal to 9 and 1, respectively. Therefore, 10 neurons in the hidden layer ANN is proposed. The sigmoid activation function for the hidden layer is selected, while the activation function for the output layer is a linear function. The cost function has been chosen as the mean square error one. Due to the random sampling effect, the number of simulations is proposed 300 times to obtain reliable results.

The main purpose of this work is to investigate the performance of four ANN models to predict the shear strength of deep beams, trained by the four algorithms, namely, Levenberg–Marquardt (ANN-LM), quasi-Newton method (ANN-QN), conjugate gradient (ANN-CG), and gradient descent (ANN-GD). The training process is repeated until the network output error reaches an acceptable value (less than the initial specified error threshold). In this study, the network training is performed with various epoch numbers, ranging from 100 to 1000 with a step of 100. Finally, Table 2 summarizes the characteristics of the ANN models proposed in this study.

6.1. Comparison of ANN Models' Prediction Capability. The results of the network training by different algorithms are evaluated by the values of criteria R, RMSE, MAE, and MAPE. Figures 6(a)-6(d) show the mean and std values of R, RMSE, MAE, and MAPE in function of different epoch numbers for the testing part of the ANN-LM algorithm. Similarly, Figures 7(a)–7(d) show the mean and std values of R, RMSE, MAE, and MAPE in function of different epoch numbers for the testing parts obtained by using ANN-QN, ANN-CG, and ANN-GD algorithms. For the ANN-LM model, the mean and std of R values decrease with a higher number of epochs, and the mean value and std of RMSE, MAE, and MAPE increase. This behavior shows that the accuracy of the ANN-LM model is highly affected by the number of epochs. It means that, with a higher number of epochs, the accuracy of the ANN-LM model decreases. It could be confirmed that the ANN-LM model can produce

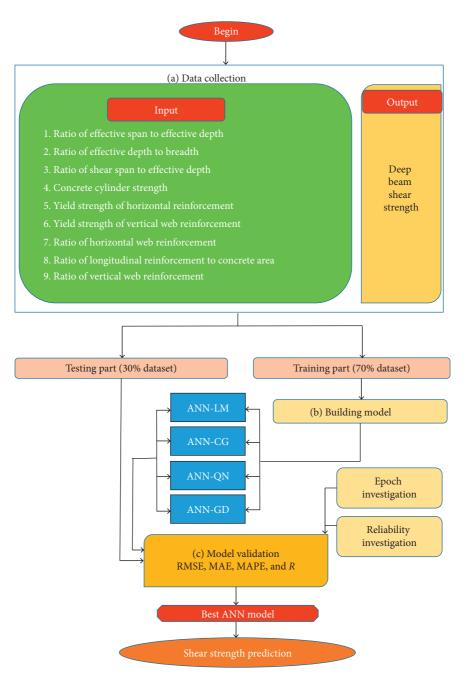


FIGURE 5: Methodology flowchart of the present study.

Table 2: Summary of ANN training algorithms and structure used in this study.

Parameter	Parameter	Description			
	Neurons in input layer	9			
	Neurons in output layer	1			
	Hidden layer activation function	Sigmoid			
Fix	Output layer activation function	Linear			
FIX	Cost function	Mean square error (MSE)			
	Number of hidden layer	1			
	Neurons in hidden layer	10			
	Number of simulations	300			
		Levenberg-Marquardt (ANN-LM)			
	Tuoining algorithms	Conjugate gradient (ANN-CG)			
Investigation	Training algorithms	Quasi-Newton method (ANN-QN)			
_		Gradient descent (ANN-GD)			
	Number of epochs	Varying from 100 to 1000 with a step of 10			

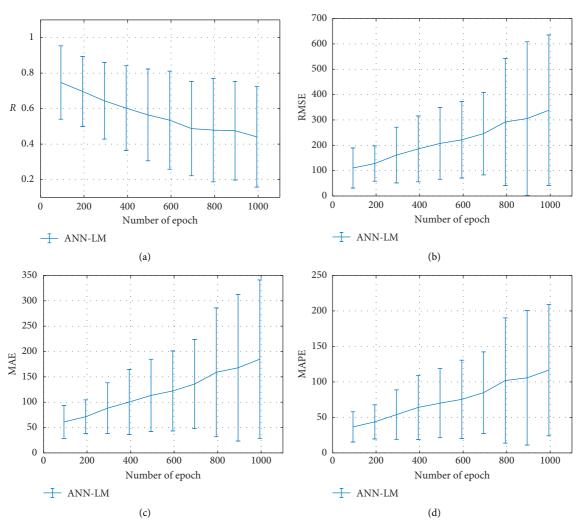
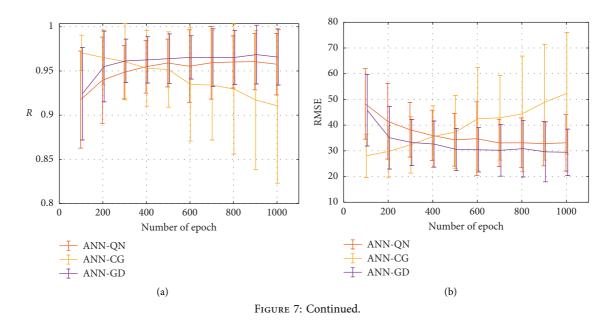


FIGURE 6: Mean value and std value of (a) R, (b) RMSE, (c) MAE, and (d) MAPE in function of the epoch number for the testing part of ANN-LM.



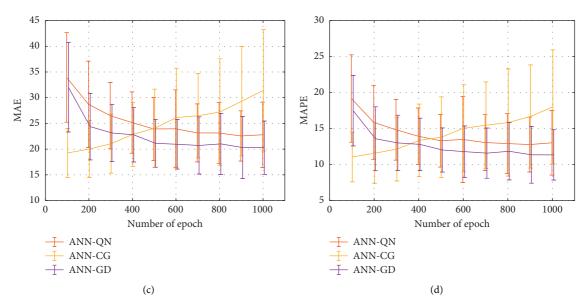


FIGURE 7: Mean value and std value of (a) R, (b) RMSE, (c) MAE, and (d) MAPE in function of the epoch number for the testing part of ANN-QN, ANN-CG, and ANN-GD.

high accuracy results with high speed, and the same conclusion could be drawn for the case of the ANN-CG model. However, an opposite conclusion is found for the case of ANN-QN and ANN-GD models, in which the mean and std values of *R* increase and those of RMSE, MAE, and MAPE decrease with a higher number of epochs. Thus, the accuracy of ANN-QN and ANN-GD models increases with a higher number of epochs.

Moreover, Table 3 details the values and std of R, RMSE, MAE, and MAPE of four models with different epoch numbers, varying from 100 to 1000 with a step of 100. It is found that the accuracy of the ANN-LM model is very low, where the maximum value of R is only 0.747 at 100 epochs. Similarly, for the ANN-CG model, at 100 epochs, the highest R value is 0.971. Therefore, the optimal ANN-LM and ANN-CG model is at 100 epochs. In contrast, with ANN-QN and ANN-GD algorithms, the highest value of R is R = 0.961 and R = 0.969, respectively. Besides, the std values of the three criteria RMSE, MAE, and MAPE of the ANN-LM model are the highest compared to the other models. This shows that the ANN-LM model has the lowest accuracy among the 4 models.

Next, the values of criteria RMSE, MAE, and MAPE of the three remaining models are compared. With the ANN-CG model, the values of these criteria are the lowest at 100 epochs, compared with the lowest value of the ANN-QN model at 900 epochs and the lowest value of the ANN-GD model at 1000 epochs. Through evaluation and analysis, it is found that the ANN-CG model is the model with the best accuracy with the least number of epochs. Considering the case of large numbers of epochs, it can be seen that the ANN-GD model is superior to the ANN-QN model. Therefore, a reliability evaluation of the three models is performed in the following sections. The lowest accuracy ANN-LM model for shear beam prediction is not proposed for the next investigation.

6.2. Reliability Evaluation of the Best ANN Training Algorithms. The main purpose of this section is to evaluate the reliability of the three models, including the optimal ANN-CG at 100 epochs, the optimal ANN-GD at 900 epochs, and the optimal ANN-QN at 1000 epochs. Figure 8 shows the distribution of the probability density function (PDF) of the four statistical criteria for the training part, namely, R (Figure 8(a)), RMSE (Figure 8(b)), MAE (Figure 8(c)), and MAPE (Figure 8(d)), over 300 simulations performed using the mentioned ANN structures. Meanwhile, Figures 9(a)-9(d) show the corresponding distribution of the probability density function for the testing part. According to observation, the PDF curves for the four statistical criteria of the training and testing parts of the ANN-CG model are the narrowest, and the best values of the four criteria are better than the two other algorithms.

Simultaneously, Table 4 presents in detail the values of the four statistical criteria (maximum, minimum, average, and standard deviation) according to the three proposed ANN models. Considering the testing part, the values of average and standard deviation for the case of R are 0.961 and 0.032 for ANN-QN, 0.971 and 0.02 for ANN-CG, and 0.966 and 0.032 for ANN-GD, respectively. For RMSE, these are 32.8 and 8.62 for ANN-QN, 28.07 and 8.44 for ANN-CG, and 29.44 and 9.01 for ANN-GD. In the case of MAE, the average and standard deviation are 22.56 and 4.89 for ANN-QN, 19.26 and 4.75 for ANN-CG, and 20.28 and 5.18 for ANN-GD. Finally, these values of MAPE are 12.78 and 3.81, respectively, for ANN-QN, 11.05 and 3.45 for ANN-CG, and 11.35 and 3.5 for ANN-GD. Thus, in terms of the average value, the ANN-CG model outperforms the other two models, with the average value of R being the highest. Meanwhile, the mean values of RMSE, MAE, and MAPE are the lowest. More importantly, the standard deviation values obtained from the ANN-CG model are also the lowest, which shows that ANN-CG is the most stable and reliable

TABLE 3: Summary of R, RMSE, MAE, and MAPE mean values and std values of 4 ANN models for the testing part.

	Epochs										
	100	200	300	400	500	600	700	800	900	1000	
Mean (R)										
LM	0.747	0.697	0.644	0.603	0.565	0.535	0.488	0.479	0.476	0.441	
QN	0.918	0.940	0.949	0.955	0.959	0.955	0.959	0.960	0.961	0.958	
CG	0.971	0.965	0.961	0.953	0.952	0.935	0.934	0.930	0.917	0.911	
GD	0.924	0.955	0.961	0.963	0.964	0.965	0.965	0.965	0.969	0.966	
Std (R)											
LM	0.207	0.197	0.216	0.239	0.259	0.276	0.266	0.292	0.278	0.283	
QN	0.055	0.049	0.030	0.030	0.027	0.041	0.041	0.030	0.032	0.035	
CG	0.020	0.031	0.043	0.043	0.043	0.064	0.062	0.074	0.079	0.088	
GD	0.052	0.040	0.025	0.026	0.028	0.025	0.033	0.031	0.033	0.032	
Mean (
LM	110.43	128.02	161.42	185.82	207.44	221.53	246.05	292.03	305.51	338.32	
QN	48.24	41.48	38.20	35.99	34.29	34.71	33.08	33.13	32.80	33.12	
CG	28.07	29.76	32.38	35.54	37.34	42.46	42.85	44.37	48.90	52.36	
GD	45.81	35.08	33.24	32.63	30.54	30.41	30.22	30.86	29.60	29.44	
Std (RN											
LM	79.10	69.85	109.87	129.88	141.54	150.88	162.70	250.70	302.82	296.92	
QN	13.72	14.69	10.64	9.81	10.26	14.38	9.19	9.70	8.62	10.98	
CG	8.44	9.98	11.03	11.93	14.19	19.98	16.45	22.47	22.53	23.66	
GD	13.94	12.16	8.93	8.97	8.15	8.61	10.08	11.02	11.65	9.01	
Mean (
LM	61.02	71.41	88.19	100.41	113.37	122.23	135.83	158.97	167.94	184.97	
QN	33.94	28.72	26.51	25.16	23.92	23.95	23.15	23.14	22.56	22.79	
CG	19.26	20.02	21.06	22.86	24.07	26.13	26.49	27.22	29.32	31.44	
GD	32.02	24.37	23.14	22.78	21.14	20.94	20.72	21.02	20.30	20.28	
Std (M.											
LM	32.37	33.30	50.19	64.17	71.34	78.82	87.98	127.07	144.57	156.18	
QN	8.70	8.39	6.45	5.97	6.13	7.58	5.64	5.90	4.89	6.35	
CG	4.75	5.56	5.74	6.23	7.59	9.59	8.21	10.36	10.62	11.80	
GD	8.69	6.47	5.57	5.28	4.66	4.82	5.56	5.94	6.04	5.18	
Mean (MAPE)										
LM	36.62	43.71	54.01	64.17	70.05	75.47	84.91	102.03	105.79	116.68	
QN	19.17	15.87	14.83	13.94	13.30	13.49	13.06	12.91	12.78	13.02	
CG	11.05	11.57	12.15	13.35	13.81	15.04	15.46	15.84	16.67	18.02	
GD	17.48	13.58	13.01	12.80	12.02	11.79	11.58	11.85	11.35	11.35	
Std (M.											
LM	21.34	24.06	34.86	45.16	48.65	55.14	57.45	88.24	94.75	92.62	
QN	6.07	5.14	4.22	3.91	3.65	5.99	3.87	4.16	3.81	4.49	
CG	3.45	4.23	4.43	5.04	5.60	6.05	6.03	7.44	7.17	7.92	
GD	4.89	4.44	3.85	3.64	3.07	3.57	3.49	3.98	3.95	3.50	

method. The results show that the ANN-CG is the most reliable training algorithm for predicting the shear strength of the deep beam. Therefore, the ANN-CG model is chosen to predict the shear strength of the deep beam in the next section.

6.3. Prediction of Beam Shear Strength Using the Best ANN Model. In this section, the capability of the ANN-CG model to predict the shear strength of deep beams is investigated. The selection criterion of the best model depends on the values of the four statistical criteria used in this study. Therefore, 4 cases of performance evaluation are considered, namely, (1) maximum value of R, (2) minimum value of RMSE, (3) minimum value of MAE, and (4) minimum value of MAPE. With respect to each case, the values of the

criteria, as well as the standard deviation and mean error, are detailed in Table 5.

In the first case, the maximum value of R is 0.993 for both training and testing parts. For the second case, the minimum RMSE value is 14.73 for the training part and 14.02 for the testing part. The minimum value of MAE is considered in Case 3, where MAE = 9.88 for the training part, and MAE = 10.06 for the testing part. The last case finds the minimum MAPE of 6 and 5.79 with the training and testing parts, respectively.

In analyzing the results presented in Table 5, the prediction performance is evaluated through 4 criteria for the testing part. The maximal value of *R* is slightly different, considering cases 1, 3, and 4. The difference between the Case 4 and min MAPE values of cases 1, 2, and 3 is relatively small, especially when comparing with those of RMSE and

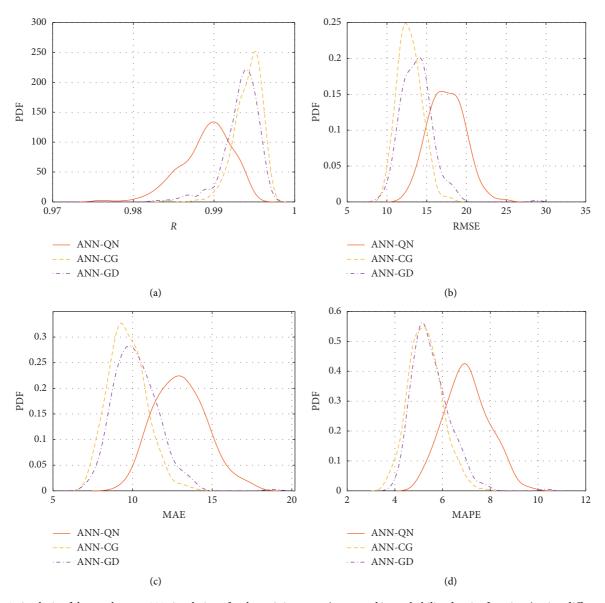


FIGURE 8: Analysis of the results over 300 simulations for the training parts (presented in probability density functions) using different ANN structures. (a) *R*. (b) RMSE. (c) MAE. (d) MAPE.

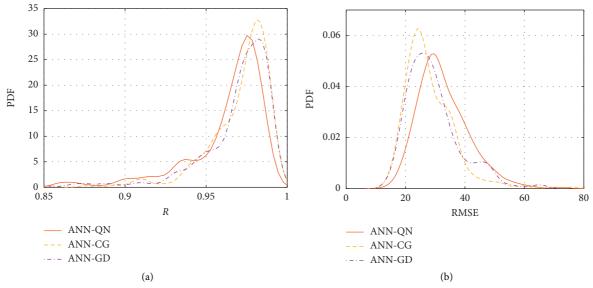


FIGURE 9: Continued.

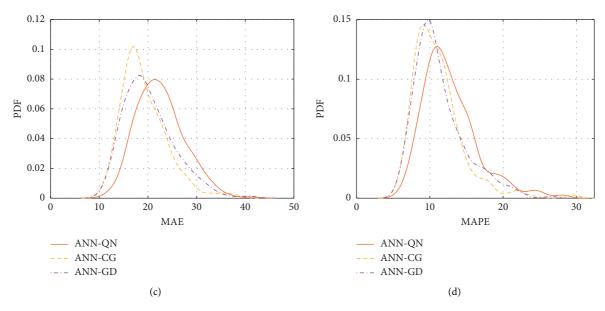


FIGURE 9: Analysis of the results over 300 simulations for the testing parts (presented in probability density functions) using different ANN structures. (a) *R.* (b) RMSE. (c) MAE. (d) MAPE.

Table 4: Summary of different quality assessment criteria over 300 simulations with different ANN structures.

0:4:	ANN	I-QN	ANN	1-CG	ANN-GD		
Criteria	Train	Test	Train	Test	Train	Test	
R							
Min	0.976	0.705	0.984	0.854	0.954	0.764	
Average	0.989	0.961	0.994	0.971	0.993	0.966	
Max	0.995	0.993	0.997	0.993	0.997	0.994	
Std	0.003	0.032	0.002	0.020	0.003	0.032	
RMSE							
Min	12.28	14.51	9.37	14.02	9.20	14.15	
Average	17.49	32.80	12.86	28.07	13.76	29.44	
Max	24.73	71.99	18.06	76.57	28.84	66.07	
Std	2.23	8.62	1.51	8.44	2.04	9.01	
MAE							
Min	9.18	11.07	7.28	10.60	7.21	11.25	
Average	13.01	22.56	9.68	19.26	10.23	20.28	
Max	17.52	41.24	13.84	38.72	18.80	41.54	
Std	1.61	4.89	1.17	4.75	1.40	5.18	
MAPE							
Min	5.03	6.00	3.72	5.79	4.02	5.97	
Average	7.02	12.78	5.29	11.05	5.52	11.35	
Max	9.61	28.26	7.70	29.70	10.33	26.58	
Std	0.92	3.81	0.69	3.45	0.79	3.50	

MAE values. It means that the Cases 2 and 3 have higher performance compared with the two other cases. However, Case 3 possesses an RMSE value higher than that of Case 2. Therefore, Case 2 has the best performance of shear strength prediction.

The error diagrams of the training and the testing parts of the ANN-CG model are presented in Figures 10(a) and 10(b). According to the results, the number of samples with errors out of the range from -30 to 40 kN is small (only 2 samples) for the training part. The error of the testing part is

Table 5: Summary of different quality assessment criteria for the best ANN-CG model.

	RMSE	MAE	Err. Mean	Err. Std	R	MAPE	
Case 1: Max	(R)					<u>.</u>	
Training set	12.53	9.82	0.16	12.61	0.993	5.79	
Testing set	17.92	13.30	0.77	18.19	0.993	6.47	
Case 2: Min	(RMSE)						
Training set	14.73	10.90	-0.45	14.82	0.993	5.89	
Testing set	14.02	11.24	0.12	14.24	0.992	6.84	
Case 3: Min	(MAE)						
Training set	13.02	9.88	0.17	13.11	0.995	5.20	
Testing set	15.49	10.60	-2.47	15.54	0.981	6.92	
Case 4: Min (MAPE)							
Training set	13.48	10.20	0.10	13.57	0.992	6.00	
Testing set	26.49	15.41	-3.66	26.65	0.990	5.79	

lower than that of the training phase. Besides, the cumulative red lines show that 80% of error is in the range from -20 to $20 \, kN$ for the training part, whereas about 90% of error is within -15 to $20 \, kN$ for the testing part.

Finally, Figures 11(a) and 11(b) show a regression model representing the correlation between the actual and predicted shear strength values for the training and testing parts, respectively. A linear fit is also applied and plotted in each case. The values of R calculated for the training part are R = 0.993 and R = 0.992 for the testing part, respectively. The values of RMSE, MAE, and MAPE for the training and testing parts are shown in Table 5.

Finally, the results of this investigation are compared with the results previously published with some other predictive methods, summarized in Table 6. Using the artificial neural network-conjugate gradient (ANN-CG) in this study, the performance of shear strength prediction of the deep beam seems to be the best with the highest value of *R*,

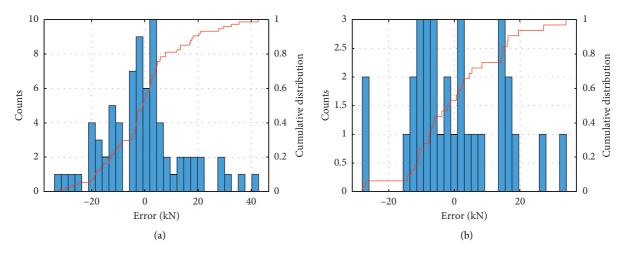


Figure 10: Error histogram of target and output values and cumulative distribution of error for the best ANN-CG model. (a) Training part. (b) Testing part.

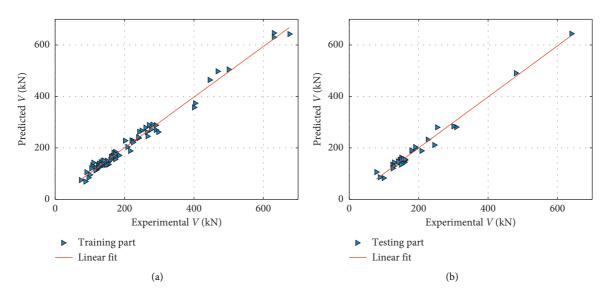


FIGURE 11: Regression graphs for the best ANN-CG model. (a) Training part. (b) Testing part.

Table 6: Comparison of prediction performance between this investigation and currently popular AI techniques.

Model		Statistical criteria			
Model	R	RMSE	MAE	MAPE	
Genetic-simulated annealing (GSA) [4]	0.929	_	_		
Backpropagation neural network (BPNN) [42]	0.916	34.032	_	11.273	
Radial basis function neural network (RBFNN) [42]	0.9767	20.29	_	7.63	
Artificial neural network (ANN) [43]	0.9711	42.27	30.28	_	
Gene expression programming (GEP) [43]	0.9654	51.57	40.99	_	
Support vector machine (SVM) [42]	0.9465	30.134	_	14.435	
Multivariate adaptive regression splines (EMARS) [42]	0.986	13.011	_	5.887	
Genetic-simulated annealing (GSA) [4]	0.929	_	_	12.3	
Smart artificial firefly colony algorithm and least squares support vector regression (SFA LS-SVR) [24]	0.941	_	_	8.87	
Adaptive neural fuzzy inference system (ANFIS) [44]	0.984	34.76	25.24	_	
Artificial neural network-conjugate gradient (ANN-CG) of this study	0.992	14.02	11.24	6.84	

IW (10 ×	(9)								LW (1×10)	BL (1×10)	BO (1×1)
-0.576	-0.075	-0.679	0.078	-0.113	-0.950	1.181	-0.257	0.343	0.410	1.589	-0.390
0.367	0.052	0.518	-0.424	-0.739	0.350	0.784	0.819	-0.877	-0.285	-1.485	_
1.281	-0.106	-0.286	0.648	-0.339	0.771	-0.897	0.585	-0.540	0.645	-0.663	_
0.278	0.060	-1.788	0.298	-0.343	-0.402	0.014	0.061	0.887	0.976	-0.924	_
-0.469	-0.099	0.137	-0.538	1.265	-0.285	0.417	-0.103	-1.010	0.307	0.384	_
-0.036	-0.891	0.371	-0.054	0.743	0.517	-0.866	-0.274	0.671	-0.083	-0.067	_
-0.440	0.312	-1.232	-0.598	-0.621	-1.295	0.010	-0.475	-0.229	0.019	-0.601	_
-0.290	-0.348	-0.482	0.505	1.006	-0.248	-0.507	-0.133	0.930	0.234	-1.274	_
0.395	0.257	-0.233	-0.362	0.718	-0.166	0.734	-0.552	-1.248	0.471	1.490	_
-0.311	0.658	0.839	-0.248	0.276	0.947	0.875	0.519	-0.070	0.398	-1.798	_

TABLE 7: Final values of weights and biases of the typical ANN-CG model (Case 2).

Note. IW = the input to hidden layer weights; LW = the hidden layer to output layer weights; BL = the input to hidden layer biases; BO = the hidden layer to output layer bias.

the lowest value of MAE, and almost the lowest values of RMSE and MAPE. More importantly, while comparing the four algorithms proposed in this study, the ANN-CG appears as the best predictor with respect to the accuracy in estimating the shear strength of the deep beam as well as less computation time is required (i.e., best performance at 100 iterations). Furthermore, the computation memory and cost are less demanded in comparison with other algorithms. It implies that the prediction of deep beam shear strength would not require a high-performance computer with the use of the ANN-CG algorithm. Usually, hybrid ML algorithms take a longer computation time than standalone ones. However, given the prediction accuracy achieved in this study, the development of a hybrid approach would not be necessary. Overall, this confirms the effectiveness of ANN-CG proposed in this study, suggesting a promising and useful alternative design solution for structural engineers. For practical applications, the final weight and bias values of the best ANN-CG model are given in Table 7 and could be used to develop a supporting numerical tool for estimation of shear strength of deep beams.

7. Conclusion

In this study, the neural network (ANN) model is proposed to predict the shear strength of deep beams. For this purpose, a database of 106 results from shear tests of RC deep beams is built from the available literature. The ANN model is built with 9 input parameters divided into two groups, namely, the geometric size parameter group and the parameter group representing the material properties. Four training algorithms of ANN are explored, namely, the Levenberg-Marquardt (ANN-LM), quasi-Newton method (ANN-QN), conjugate gradient (ANN-CG), and gradient descent (ANN-GD). The prediction performance of different ANN training algorithms is compared. Four different statistical criteria, namely, the correlation coefficient (R), root mean square error (RMSE), mean absolute error (MAE), and mean absolute percentage error (MAPE), are introduced to validate and evaluate the performance of the ANN model. The conjugate gradient (CG) algorithm is chosen as the best training ANN algorithm for predicting the shear strength of deep beams. With the ANN-CG model chosen as best, four cases corresponding to four different black-boxes are

studied. The results show that the crucial information to choose an accurate machine learning model might lie on the criterion that the smallest value of RMSE is obtained. Besides, the analysis of error between predicted and actual shear strength shows that the ANN model can be a promising numerical tool that could considerably avoid time-consuming and costly experimental procedures. Despite an extensive investigation on different potential training algorithms and epochs, this study is only conducted on one ANN architecture. Therefore, regardless of the highest and outstanding prediction accuracy achieved, it is interesting to perform another investigation related to the neuron number and the hidden layer number to, possibly, enhance the performance of the ANN-CG model, or to further decrease the computation time by decreasing the neuron in the hidden layer.

Data Availability

The data used to support the findings of the study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- Bureau of Indian Standards, Concrete, Plain and Reinforced, pp. 1–114, Bureau of Indian Standards, New Dehli, India, 2000.
- [2] ACI 318-14, ACI 318-14 Building Code Requirements for Structural Concrete. 2014.
- [3] British Standards Institution, Eurocode 2: Design of Concrete Structures: Part 1-1: General Rules and Rules for Buildings, British Standards Institution, London, UK, 2004.
- [4] A. H. Gandomi, A. H. Alavi, D. M. Shadmehri, and M. G. Sahab, "An empirical model for shear capacity of RC deep beams using genetic-simulated annealing," *Archives of Civil and Mechanical Engineering*, vol. 13, no. 3, pp. 354–369, 2013.
- [5] G. Harsha and p raju, "Shear Strength of Deep Beams: A State of Art," Apr. 2019.
- [6] K. Ismail, "Shear Behaviour of Reinforced Concrete Deep Beams," 2016.

[7] D. B. Birrcher, Design of Reinforced Concrete Deep Beams for Strength and Serviceability, Ph. D Thesis, The University of Texas, Austin, USA, 2009.

- [8] A. Arabzadeh, R. Aghayari, and A. R. Rahai, "Investigation of experimental and analytical shear strength of reinforced concrete deep beams," *International Journal of Civil Engi*neering, vol. 9, no. 3, pp. 207–214, 2011.
- [9] M. Shahnewaz, A. Rteil, and M. S. Alam, "Shear strength of reinforced concrete deep beams - a review with improved model by genetic algorithm and reliability analysis," *Structures*, vol. 23, pp. 494–508, 2020.
- [10] M. Tasleema, M. A. Kumar, and J. L. Raj, "Evaluation of shear strength of deep beams using artificial neural networks," *International Journal of Recent Technology and Engineering*, vol. 7, no. 6C2, pp. 341–345, 2019.
- [11] K.-H. Yang, H.-S. Chung, E.-T. Lee, and H.-C. Eun, "Shear characteristics of high-strength concrete deep beams without shear reinforcements," *Engineering Structures*, vol. 25, no. 10, pp. 1343–1352, 2003.
- [12] T. S. Tan and L. W. Weng, "A strut-and-tie model for deep beams subjected to combined top-and-bottom loading," *Structural Engineer*, vol. 75, no. 13, 1997.
- [13] J. Park and D. Kuchma, "Strut-and-tie model analysis for strength prediction of deep beams," ACI Structural Journal, vol. 104, pp. 657–666, 2007.
- [14] C. Y. Tang and K. H. Tan, "Interactive mechanical model for shear strength of deep beams," *Journal of Structural Engineering*, vol. 130, no. 10, pp. 1534–1544, 2004.
- [15] M. Lezgy-Nazargah, "A four-variable global-local shear deformation theory for the analysis of deep curved laminated composite beams," *Acta Mechanica*, vol. 231, no. 4, pp. 1403–1434, 2020.
- [16] M. P. Collins, "The modified compression-field theory for reinforced concrete elements subjected to shear," ACI Journal Proceedings, vol. 832 pages, 1986.
- [17] M. P. Vecchio and F. J. Collins, "Compression response of cracked reinforced concrete," *Journal of Structural Engineering*, vol. 119, pp. 3590–3610, 1993.
- [18] S.-J. Hwang and H.-J. Lee, "Strength prediction for discontinuity regions by softened strut-and-tie model," *Journal of Structural Engineering*, vol. 128, 2002.
- [19] S.-J. Hwang, W.-Y. Lu, and H.-J. Lee, "Shear strength prediction for deep beams," ACI Structural Journal, vol. 97, pp. 367–376, 2000.
- [20] K.-H. Yang and A. F. Ashour, "Strut-and-Tie model based on crack band theory for deep beams," *Journal of Structural Engineering*, vol. 137, no. 10, pp. 1030–1038, 2011.
- [21] Z. Bazant and J. Planas, Fracture and Size Effect in Concrete and Other Quasibrittle Materials, CRC Press, Washington, DC, USA, 2019.
- [22] Canadian Standards Association, Design of Concrete Structures; CSA A23.3-04, Canadian Standards Association, Mississauga, ON, Canada, 2004.
- [23] M. Pal and S. Deswal, "Support vector regression based shear strength modelling of deep beams," *Computers & Structures*, vol. 89, no. 13-14, pp. 1430–1439, 2011.
- [24] J.-S. Chou, N.-T. Ngo, and A.-D. Pham, "Shear strength prediction in reinforced concrete deep beams using natureinspired metaheuristic support vector regression," *Journal of Computing in Civil Engineering*, vol. 30, no. 1, Article ID 04015002, 2016.
- [25] M. Y. Mansour, M. Dicleli, J. Y. Lee, and J. Zhang, "Predicting the shear strength of reinforced concrete beams using artificial

- neural networks," Engineering Structures, vol. 26, no. 6, pp. 781-799, 2004.
- [26] J. Amani and R. Moeini, "Prediction of shear strength of reinforced concrete beams using adaptive neuro-fuzzy inference system and artificial neural network," *Scientia Iranica*, vol. 19, no. 2, pp. 242–248, 2012.
- [27] X.-H. Tan, W.-H. Bi, X.-L. Hou, and W. Wang, "Reliability analysis using radial basis function networks and support vector machines," *Computers and Geotechnics*, vol. 38, no. 2, pp. 178–186, 2011.
- [28] J.-S. Chou, M.-Y. Cheng, and Y.-W. Wu, "Improving classification accuracy of project dispute resolution using hybrid artificial intelligence and support vector machine models," *Expert Systems with Applications*, vol. 40, no. 6, pp. 2263–2274, 2013.
- [29] M. D. Nguyen, B. T. Pham, T. T. Tuyen et al., "Development of an artificial intelligence approach for prediction of consolidation coefficient of soft soil: a sensitivity analysis," *The Open Construction and Building Technology Journal*, vol. 13, no. 1, pp. 178–188, 2019.
- [30] H.-B. Ly and B. T. Pham, "Prediction of shear strength of soil using direct shear test and support vector machine model," *The Open Construction and Building Technology Journal*, vol. 14, no. 1, pp. 41–50, 2020.
- [31] D. Dao, H.-B. Ly, S. Trinh, T.-T. Le, and B. Pham, "Artificial intelligence approaches for prediction of compressive strength of geopolymer concrete," *Materials*, vol. 12, no. 6, p. 983, 2019.
- [32] H.-B. Ly, B. Thai Pham, D. Van Dao, V. Minh Le, L. Minh Le, and T.-T. Le, "Improvement of ANFIS model for prediction of compressive strength of manufactured sand concrete," *Applied Sciences*, vol. 9, no. 18, p. 3841, 2019.
- [33] H.-B. Ly, L. M. Le, H. T. Duong et al., "Hybrid artificial intelligence approaches for predicting critical buckling load of structural members under compression considering the influence of initial geometric imperfections," *Applied Sciences*, vol. 9, no. 11, p. 2258, 2019.
- [34] Q. H. Nguyen, H.-B. Ly, V. Q. Tran et al., "A novel hybrid model based on a feedforward neural network and one step secant algorithm for prediction of load-bearing capacity of rectangular concrete-filled steel tube columns," *Molecules*, vol. 25, no. 15, p. 3486, 2020.
- [35] H. Q. Nguyen, H.-B. Ly, V. Q. Tran, T.-A. Nguyen, T.-T. Le, and B. T. Pham, "Optimization of artificial intelligence system by evolutionary algorithm for prediction of axial capacity of rectangular concrete filled steel tubes under compression," *Materials*, vol. 13, no. 5, p. 1205, 2020.
- [36] A. T. C. Goh, "Prediction of ultimate shear strength of deep beams using neural networks," ACI Structural Journal, vol. 921 page, 1995.
- [37] A. Sanad and M. Saka, "Prediction of ultimate shear strength of reinforced-concrete deep beams using neural networks," *Journal of Structural Engineering*, vol. 127, 2001.
- [38] M. T. K. Niazi, Arshad, J. Ahmad, F. Alqahtani, F. A. Baotham, and F. Abu-Amara, "Prediction of critical flashover voltage of high voltage insulators leveraging bootstrap neural network," *Electronics*, vol. 9, no. 10, pp. 1620–10, 2020.
- [39] A. J. Jakeman, R. A. Letcher, and J. P. Norton, "Ten iterative steps in development and evaluation of environmental models," *Environmental Modelling & Software*, vol. 21, no. 5, pp. 602–614, 2006.
- [40] G. Vazquez Amabile, B. A. Engel, and D. Flanagan, "Modeling and risk analysis of nonpoint-source pollution caused by

atrazine using SWAT," *Transactions of the ASABE (American Society of Agricultural and Biological Engineers)*, vol. 49, pp. 667–678, 2006.

- [41] J. C. Bathurst, J. Ewen, G. Parkin, P. E. 'O'Connell, and J. D. Cooper, "Validation of catchment models for predicting land-use and climate change impacts. 3. Blind validation for internal and outlet responses," *Journal of Hydrology*, vol. 287, no. 1–4, pp. 74–94, 2004.
- [42] M.-Y. Cheng and M.-T. Cao, "Evolutionary multivariate adaptive regression splines for estimating shear strength in reinforced-concrete deep beams," *Engineering Applications of Artificial Intelligence*, vol. 28, pp. 86–96, 2014.
- [43] A. H. Gandomi, G. J. Yun, and A. H. Alavi, "An evolutionary approach for modeling of shear strength of RC deep beams," *Materials and Structures*, vol. 46, no. 12, pp. 2109–2119, 2013.
- [44] A. Khajeh, S. R. Mousavi, and M. Rakhshani Mehr, "Adaptive neural fuzzy inference system models for predicting the shear strength of reinforced concrete deep beams," *Journal Of Rehabilitation In Civil Engineering*, vol. 3, no. 1, pp. 14–23, 2015.
- [45] F.-K. K. Teng and L.I. Guan, "High-strength concrete deep beams with effective span and shear span variations," ACI Structural Journal, vol. 92, no. 4, 1995.
- [46] A. S. Vantsiotis, "Shear strength of deep beams," ACI Journal Proceedings, vol. 79, no. 3, 1982.
- [47] P. J. R. David and F. Cole, "Web reinforcement effects on deep beams," *ACI Journal Proceedings*, vol. 67, no. 12, 1992.
- [48] P. S. Sandhu and S. Chhabra, "A comparative analysis of conjugate gradient algorithms & PSO based neural network approaches for reusability evaluation of procedure based software systems," *Chiang Mai Journal of Science*, vol. 38, no. 2, pp. 123–135, 2011.
- [49] P. Sarir, J. Chen, P. G. Asteris, D. J. Armaghani, and M. M. Tahir, "Developing GEP tree-based, neuro-swarm, and whale optimization models for evaluation of bearing capacity of concrete-filled steel tube columns," *Engineering with Computers*, vol. 37, no. 1, pp. 1–19, 2021.
- [50] D. V. Dao, H. Adeli, H.-B. Ly et al., "A sensitivity and robustness analysis of GPR and ANN for high-performance concrete compressive strength prediction using a Monte Carlo simulation," *Sustainability*, vol. 12, no. 3, p. 830, 2020.
- [51] G. Cybenko, "Approximation by superpositions of a sigmoidal function," *Mathematics of Control, Signals, and Sys*tems, vol. 2, no. 4, pp. 303–314, 1989.
- [52] L. M. Bounds and W. Mathew, "A multilayer perceptron network for the diagnosis of low back pain," in *Proceedings of* the IEEE 1988 International Conference On Neural Networks, no. 2, pp. 481–489, San Diego, CA, USA, 1988.
- [53] E. T. Mohamad, R. S. Faradonbeh, D. J. Armaghani, M. Monjezi, and M. Z. A. Majid, "An optimized ANN model based on genetic algorithm for predicting ripping production," *Neural Computing and Applications*, vol. 28, no. S1, pp. 393–406, 2016.
- [54] T. N. Singh, V. K. Singh, and S. Sinha, "Prediction of cadmium removal using an artificial neural network and a neuro-fuzzy technique," *Mine Water and the Environment*, vol. 25, no. 4, pp. 214–219, 2006.
- [55] B. Gordan, D. Jahed Armaghani, M. Hajihassani, and M. Monjezi, "Prediction of seismic slope stability through combination of particle swarm optimization and neural network," *Engineering With Computers*, vol. 32, no. 1, pp. 85–97, 2015.
- [56] S. Nagendra, Practical Aspects of Using Neural Networks: Necessary Preliminary Specifications, GE Research and

- Development Center, New York, NY, USA, 1998, http://citeseerx.ist.psu.edu/viewdoc/citations; jsessionid=0DE4DB7A72E3FD87F294E49E95FEAA2F? doi=10.1.1.
- [57] S. Tamura, "Method of determining an optimal number of neurons contained in hidden layers of a neural network," Jan. 21, 1997.
- [58] K. G. Sheela and S. N. Deepa, "Review on methods to fix number of hidden neurons in neural networks," *Mathe-matical Problems in Engineering*, vol. 2013, Article ID 425740, 11 pages, 2013.
- [59] K. Hornik, "Approximation capabilities of multilayer feedforward networks," *Neural Networks*, vol. 4, no. 2, pp. 251–257, 1991.
- [60] A. Kuri, "Closed determination of the number of neurons in the hidden layer of a multi-layered perceptron network," *Soft Computing*, vol. 21, pp. 1–13, 2016.