ARTICLE

The Analysis of the Correlation between SPT and CPT Based on CNN-GA and Liquefaction Discrimination Research

Ruihan Bai[1], Feng Shen[2,*], Zihao Zhao[3], Zhiping Zhang[4] and Qisi Yu[4]

[1]Key Laboratory of the Ministry of Education for Geomechanics and Embankment Engineering, College of Civil and Transportation Engineering, Hohai University, Nanjing, 210098, China

[2]School of Civil Engineering, Suzhou University of Science and Technology, Suzhou, 215011, China

[3]School of Civil Engineering, Shenyang Jianzhu University, Shenyang, 110168, China

[4]College of Civil Engineering, Tongji University, Shanghai, 200092, China

*Corresponding Author: Feng Shen. Email: shenfeng1023@163.com

ABSTRACT

The objective of this study is to investigate the methods for soil liquefaction discrimination. Typically, predicting soil liquefaction potential involves conducting the standard penetration test (SPT), which requires field testing and can be time-consuming and labor-intensive. In contrast, the cone penetration test (CPT) provides a more convenient method and offers detailed and continuous information about soil layers. In this study, the feature matrix based on CPT data is proposed to predict the standard penetration test blow count N. The feature matrix comprises the CPT characteristic parameters at specific depths, such as tip resistance q_c, sleeve resistance f_s, and depth H. To fuse the features on the matrix, the convolutional neural network (CNN) is employed for feature extraction. Additionally, Genetic Algorithm (GA) is utilized to obtain the best combination of convolutional kernels and the number of neurons. The study evaluated the robustness of the proposed model using multiple engineering field data sets. Results demonstrated that the proposed model outperformed conventional methods in predicting N values for various soil categories, including sandy silt, silty sand, and clayey silt. Finally, the proposed model was employed for liquefaction discrimination. The liquefaction discrimination based on the predicted N values was compared with the measured N values, and the results showed that the discrimination results were in 75% agreement. The study has important practical application value for foundation liquefaction engineering. Also, the novel method adopted in this research provides new ideas and methods for research in related fields, which is of great academic significance.

KEYWORDS

CNN; liquefaction discrimination; SPT; CPT

1 Introduction

Soil liquefaction refers to the compaction of saturated soil caused by vibrations that lead to a rapid increase in pore water pressure, correspondingly resulting in a gradual decrease in effective stress until it disappears ultimately [1–5]. Eventually, the soil is unable to resist any external load. As a geological

hazard with severe consequences, soil liquefaction usually leads to uneven settlement of foundations, building collapse, and slope instability, seriously threatening people's lives [6–10]. Many studies have been conducted on the mechanism of soil liquefaction and related discrimination procedures, such as the Seed simplified method [11–15], laboratory studies [16,17], and liquefaction discrimination methods based on artificial intelligence [18–21].

As an earthquake-prone country, China has accumulated valuable experience in liquefaction discrimination [22–24]. The discrimination method currently used in China is the empirical analysis method that mainly includes the SPT-based and CPT-based approaches, which is summarized by investigating and studying many liquefaction and non-liquefaction data collected over the years. Among them, the liquefaction discrimination theory of the SPT-based approach is relatively mature due to its technical maturity, and liquefaction discrimination parameters derived from SPT data can be directly applied to existing design specifications and standards. However, the test result of SPT is significantly influenced by the operation of the geotechnical personnel, such as the drilling method, penetration speed, and perpendicularity of the drill pipe, which leads to difficulty in obtaining test data. In contrast, the CPT test is a simpler and faster process that causes less damage to the soil layer and allows for continuous recording of soil information. However, the quality of the data referenced by the CPT-based method is not high enough, especially in complex sites where the lack of data is more prominent, which limits its prediction accuracy and reliability. Additionally, compared with the SPT-based liquefaction discrimination method, the CPT method is based on relatively less research on liquefaction phenomena, so the understanding of liquefaction phenomena is limited, and its liquefaction discrimination mechanism needs further in-depth study. In summary, developing an economical and simple method to combine the two liquefaction discrimination methods can better meet the requirements of practical engineering applications, such as obtaining the standard penetration test blow count N by CPT parameters for liquefaction discrimination.

Table 1 lists studies about empirical equations for predicting N based on one or more CPT parameters, and these methods mainly use regression analysis to establish the relationship. Acka [25] investigated the correlation between SPT and CPT data in the UAE and proposed an empirical equation to predict CPT values from SPT data. Dos Santoz et al. [26] developed correlation models between SPT-CPT and DPL-CPT data for sandy soils in Brazil. Shahri et al. [27] performed statistical analysis to develop a high-accuracy correlation model between SPT and CPT data. Zhao et al. [28] established correlations between SPT and CPT parameters for liquefaction evaluation in China, proposing two correlation equations for N value and q_c. Li et al. [29] used the least squares method to establish a regression model between N with q_c and investigated the impact of three factors on liquefaction discrimination. Liu et al. [30] determined the selection principle of SPT-CPT comparison holes and the acquisition method of parameters in several actual projects, and further used statistical methods to fit the scatter plot formed by N and q_c values. The emergence of artificial intelligence has given rise to the widespread use of intelligent techniques for problem-solving in science and engineering. Among these techniques, machine learning has proven to be a powerful tool for extracting information from data. Recently, it has been increasingly applied to predict the N value using CPT parameters. Tarawneh [31] analyzed a dataset consisting of both SPT and CPT data and used the Artificial Neural Networks (ANN) model to predict the N value from the CPT data. The study found that the ANN model provided accurate predictions of the N value, demonstrating the potential for using this approach as an alternative to traditional SPT testing. Fernando et al. [32] explored the feasibility of using ANN to predict SPT values and compared the prediction accuracy with and without data normalization. The study found that an ANN model using CPT data and soil properties as input variables could accurately predict SPT values, with data normalization further improving the

accuracy and reliability of the predictions. Gupta et al. [33] investigated the correlations between SPT and CPT data for liquefaction assessment using the statistical programming language R. The study compared different correlation methods and found that the developed correlations using the random forest regression method provided the best prediction accuracy.

Table 1: The study of correlation analysis between SPT and CPT

Studies	Soil category	Relationship
Acka [25]	Sand	$(n) = q_c/N = 0.77$
	Silty sand	$(n) = q_c/N = 0.7$
	Sandy silt	$(n) = q_c/N = 0.58$
Dos Santoz et al. [26]	Sandy soil	$(n) = q_c/N = 0.44$
Shahri et al. [27]	Silty sand	$q_c = 0.282N^{1.212}$
	Clay	$q_c = 0.409N^{0.779}$
	Sandy silt	$q_c = 0.563N - 0.366$
	Sand	$q_c = 0.605N - 0.842$
	Gravelly sand to sand	$q_c = 0.3975N^{1.13}$
Zhao et al. [28]	Silt	$N = 2.089q_c + 1.485$
	Silty sand	$N = 1.384\, q_c + 9.063$
Li et al. [29]	All categories of soils	$N = [2.834 - 0.106\ln(q_c) - 0.073(q_c)0.54]q_c$
Liu et al. [30]	Silt	$N = 1.970\, q_c + 1.485$
	Silty sand	$N = 1.408\, q_c + 9.008$

To match CPT and SPT data, many of the above studies have employed the method of averaging CPT parameters or selecting representative values within a given soil layer to correspond to N values. This method has the advantage of being simple and easy to implement. However, it is important to note that this approach has some limitations. One limitation is that this method may overlook critical local or inhomogeneous features. Averaging or selecting representative values may smooth out small-scale variations within the soil layer, and these variations may be important for understanding the geotechnical behavior of the soil. Another potential limitation of averaging or selecting representative values for CPT data is that it may obscure the variability between different categories of soil layers. Soil layers can exhibit different geotechnical properties, and ignoring this variability can lead to inaccurate correlations between the CPT and SPT data. When performing CPT and SPT data matching, factors such as soil properties, structure, and variability characteristics usually need to be considered to obtain accurate and reliable matching results.

To address a common problem faced by current SPT-CPT correlation studies, this study proposed to correspond to N values by establishing the CPT feature matrix. The CPT feature matrix consists of the CPT parameters of the two CPT exploration holes closest to the SPT exploration holes within a given soil layer. The numerical matrix divides the CPT data of different exploration holes into different column vectors. By doing so, the CPT feature matrix approach can make full use of the information

from multiple exploration holes, allowing for the extraction of richer information regarding soil characteristics. Additionally, the numerical matrix approach divides the CPT data into different row vectors based on the different depths of CPT parameters. This method enables the extraction of feature information at different depths, which accurately reflects the changing characteristics of soil properties. In contrast, the averaging method simply averages the CPT data, which may not fully capture the local variations in depth and can lead to inaccuracies in characteristic soil estimations. Drawing inspiration from the characteristics of Convolutional Neural Network (CNN) in image recognition [34,35], the paperproposed that a sliding window function performing the convolution principle can be effectively applied to the constructed CPT feature matrix. As the convolution kernel slides over the matrix, the features are integrated by multiplying the convolution kernel with the input matrix. The CNN network performs higher-level feature extraction on the input CPT feature matrix to predict the N value.

In summary, this paper presents a novel approach for establishing the correlation between N and CPT parameters using the CNN model. The Genetic Algorithm (GA) is utilized to optimize the CNN network hyperparameters in a data-driven manner. The model can predict the values of N based on the CPT parameters. Finally, the predicted N values are expected to be used for assessing the liquefaction potential based on the discriminatory liquefaction method outlined in the Chinese Specification for Seismic Design of Buildings.

The remainder of this paper is as follows: Section 2 introduces the experimental data used in this paper and data pre-processing work; Section 3 introduces the principle of CNN and GA; Section 4 presents the establishment of the optimized CNN model; Section 5 shows a comparative study with the empirical methods; Section 6 demonstrates the feasibility of utilizing the predicted N values obtained from proposed model for liquefaction discrimination.

2 Experimental Data Introduction

2.1 Data Source

The standard penetration test blow count N is a crucial geotechnical parameter that reflects soil resistance properties. It is determined by counting the number of blows required to drive a split-spoon sampler 30 cm into the soil using standard energy of 60 joules per blow during the SPT (Fig. 1a). The tip resistance q_c and sleeve resistance f_s, obtained through the CPT, are equally important geotechnical parameters for evaluating soil mechanical properties. The q_c value measures the resistance of a cone-shaped penetrometer to penetration into the soil at a constant rate, while the f_s value measures the resistance of the penetrometer's side friction against the soil (Fig. 1b).

The experimental data were collected from five investigation engineering sites in Shanghai, China, as shown in Fig. 1c. The whole data included 138 SPT exploration holes and 198 CPT exploration holes, with the number of exploration holes varying across each site. Specifically, the number of exploration holes for each site is statistically illustrated as follows:

Engineering site 1: 25 SPT exploration holes and 41 CPT exploration holes

Engineering site 2: 42 SPT exploration holes and 63 CPT exploration holes

Engineering site 3: 7 SPT exploration holes and 14 CPT exploration holes

Engineering site 4: 20 SPT exploration holes and 27 CPT exploration holes

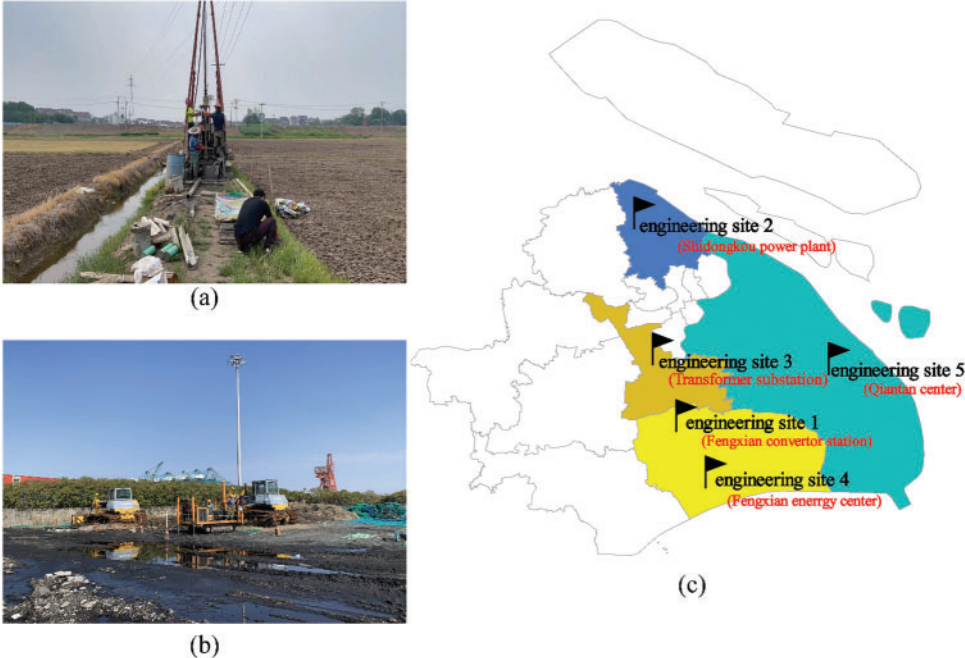Engineering site 5: 44 SPT exploration holes and 53 CPT exploration holes

Figure 1: Data source: (a) Standard penetration test; (b) Cone penetration test (c) The location of the engineering site

2.2 Abnormal Data Filtering

This research investigates three categories of soil layers: sandy silt, silty sand, and clayey silt. To ensure that the data used for correlation analysis is of good quality and reduces the risk of overfitting during model training, the box plot method [36] is used to identify and remove outliers from the experimental dataset. It works by first arranging the data in ascending order and then calculating statistical metrics such as the median value, quartiles Q1 and Q3, and the interquartile range (IQR). The upper limit (UL) and lower limit (LL) are then calculated based on the quartiles and IQR. Data points outside the UL and LL are considered outliers and removed from the dataset. Figs. 2–4 illustrate the distribution of both raw and filtered data for different soil categories, including sandy silt, silty sand, and clayey silt. The removal of data outliers leads to a more concentrated data distribution, significantly reducing the impact of outliers on data analysis and improving the reliability of data analysis.
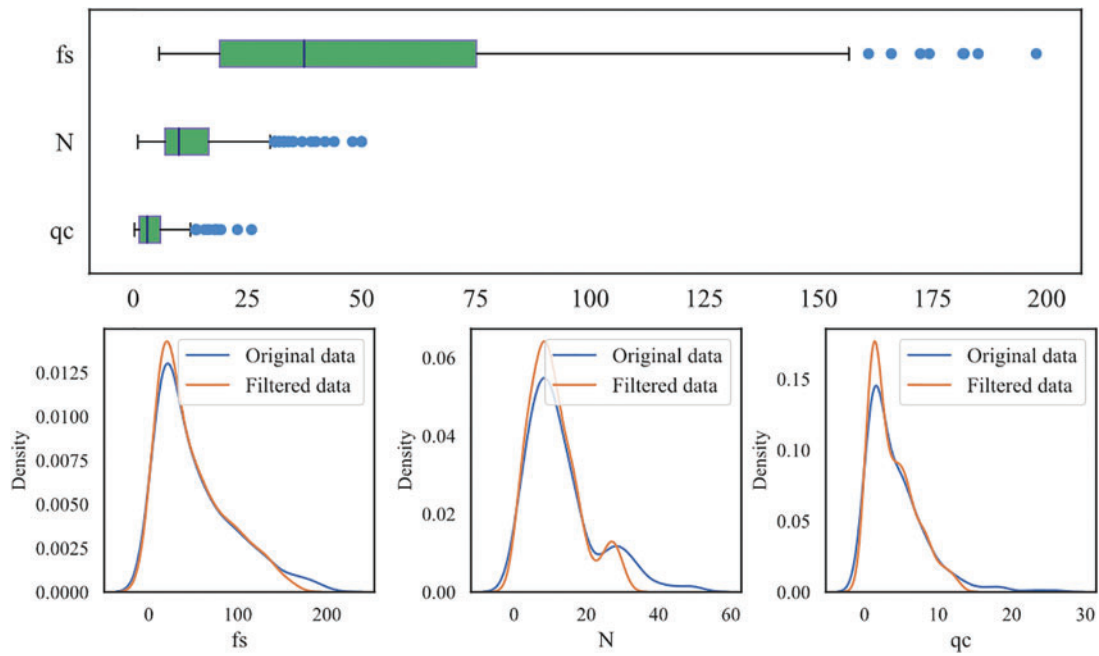
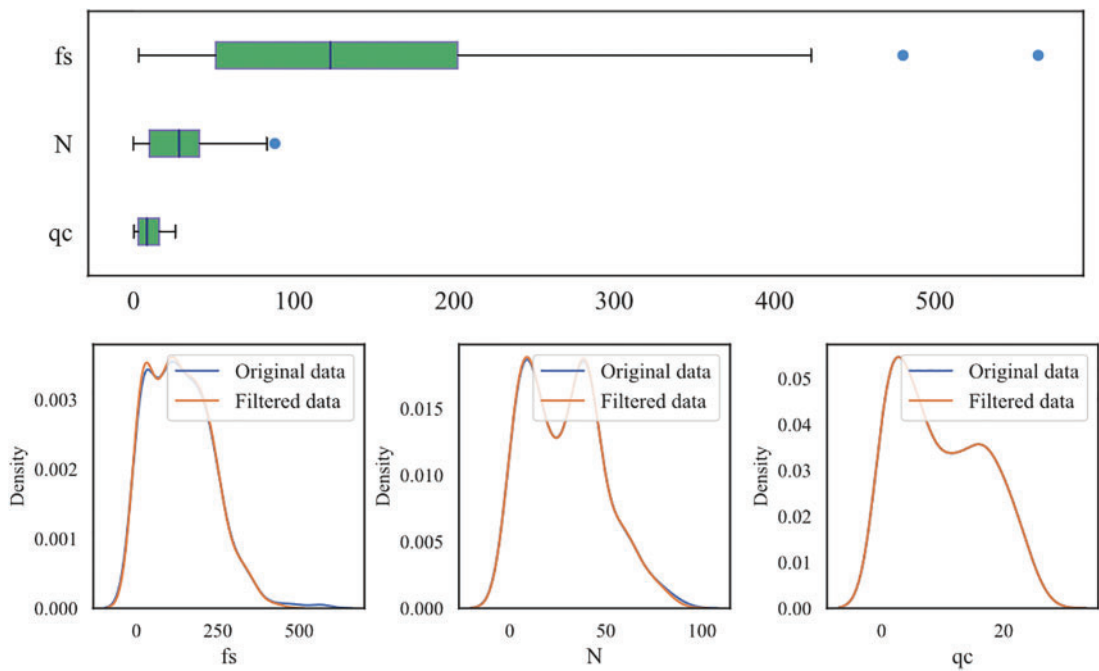**Figure 2:** Abnormal data filtering (sandy silt)

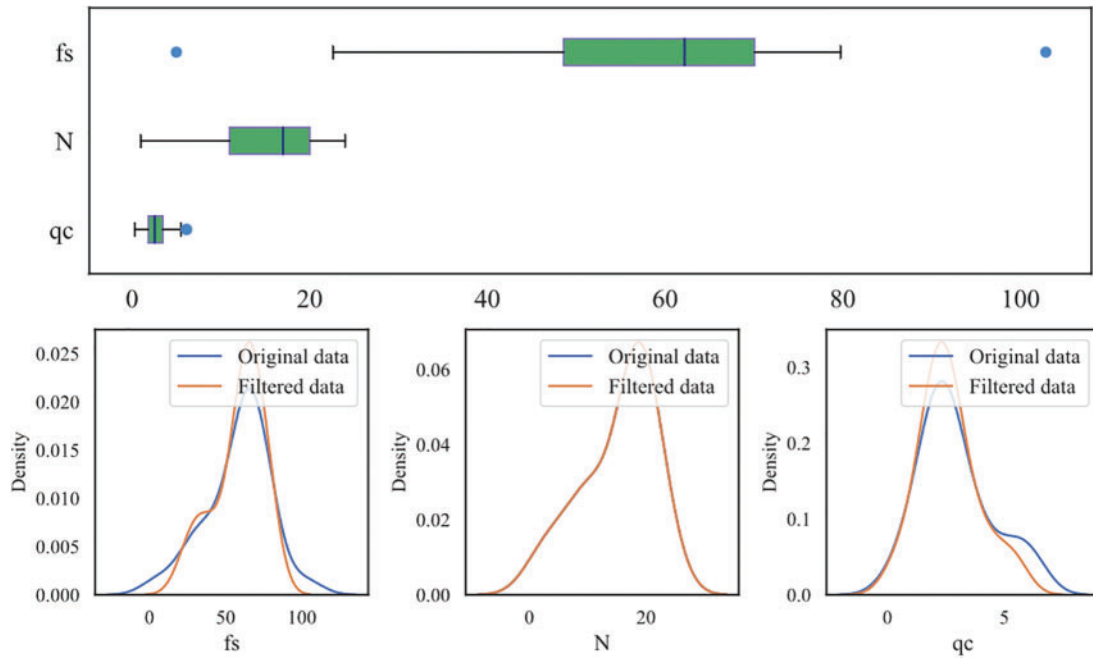**Figure 3:** Abnormal data filtering (silty sand)

Figure 4: Abnormal data filtering (clayey silt)

2.3 The Matching Principle between CPT and SPT Data

The data of SPT is discrete, while the data of CPT is continuous, and a principle needs to be established to match the SPT and CPT data. This study proposes a solution using the CPT feature matrix to match the N value. First, the maximum distance between CPT and SPT was set at 30 m, and only data within this range was considered for correlation analysis. In the vertical direction, the upper and lower 0.2 m of each depth of the N value were taken as the soil's upper and lower characteristic boundary to capture the detailed variance characteristics of CPT data. For the horizontal direction, the CPT hole with the two closest distances to the SPT exploration hole was selected. The feature matrix corresponding to the N was established based on the q_c, f_s, and H of the two selected CPT holes (Fig. 5). Note that SPT exploration holes without a nearby CPT hole within 30 m were eliminated. If only one CPT hole was available, its feature information was duplicated to establish the feature matrix.

As described in Eq. (1), the CPT feature matrix consists of the CPT parameters of the two CPT exploration holes closest to the SPT exploration holes within a given soil layer.

$$X = \left(x^{(1)}, H, x^{(2)}\right) = \begin{vmatrix} q_c^{(1)} & f_s^{(1)} & H_{i+0.2} & f_s^{(2)} & q_c^{(2)} \\ q_c^{(1)} & f_s^{(1)} & H_{i+0.1} & f_s^{(2)} & q_c^{(2)} \\ q_c^{(1)} & f_s^{(1)} & H_i & f_s^{(2)} & q_c^{(2)} \\ q_c^{(1)} & f_s^{(1)} & H_{i-0.1} & f_s^{(2)} & q_c^{(2)} \\ q_c^{(1)} & f_s^{(1)} & H_{i-0.2} & f_s^{(2)} & q_c^{(2)} \end{vmatrix}_{5 \times 5} \tag{1}$$

where X refers to the CPT characteristic matrix corresponding to N, x refers to the CPT hole with the two closest distances to the SPT hole, i refers to the penetration depth of SPT, and the subscripts (1) and (2) at the top right corner of x refer to the number of CPT exploration hole that is closest to the SPT exploration hole, respectively.
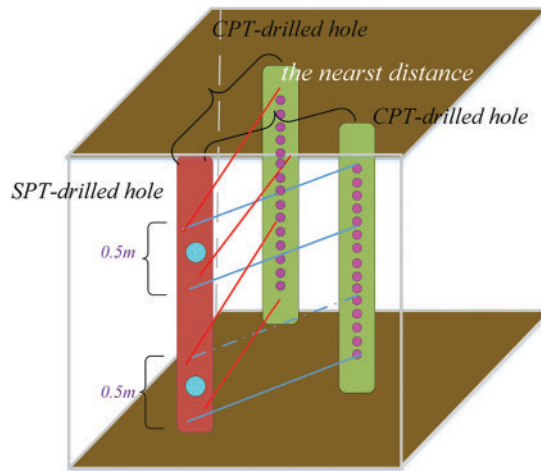
Figure 5: The matching principle between SPT and CPT

The proposed method utilizes a CNN model to establish the relationship between the CPT feature matrix and N. Through the sliding window function, the convolution kernel integrates the q_c and f_s information from different depths of the first CPT exploration hole (Fig. 6a), then integrates the fs information and depth H (Fig. 6b). Subsequently, the window integrates CPT information from the second exploration hole (Figs. 6c and 6d). When the window slides down to the next level, it can integrate CPT information from different depths (Fig. 6e).
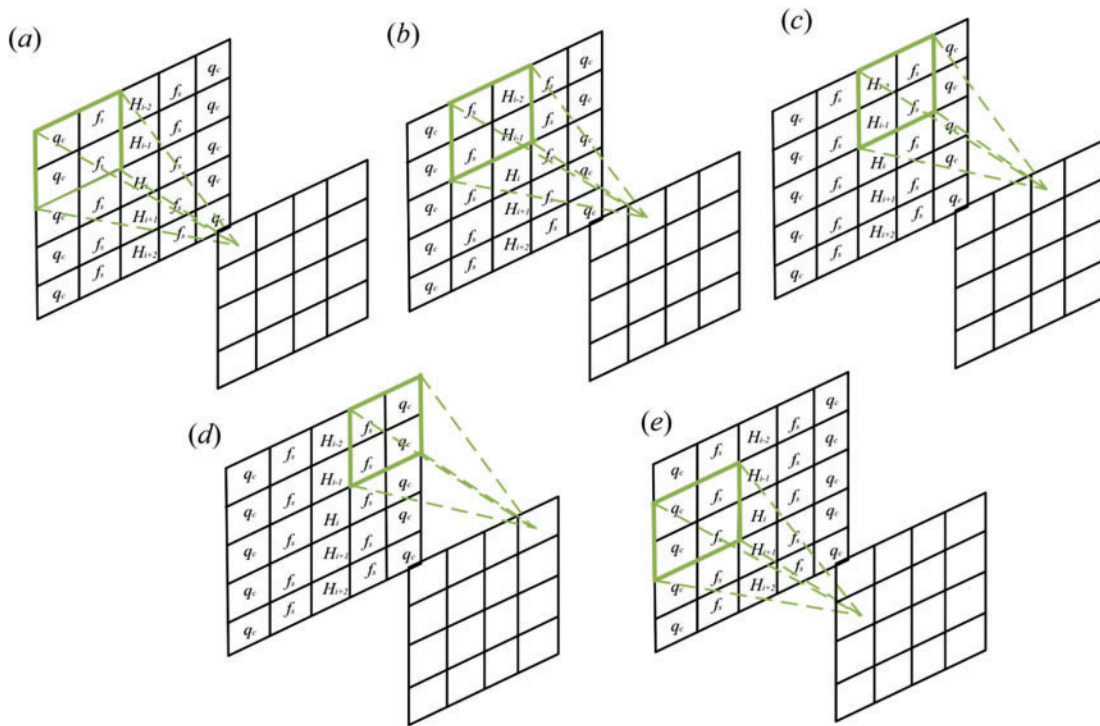

Figure 6: The principle of feature fusion based on the feature matrix

3 Analytical Method

3.1 Machine Learning Algorithm CNN

CNN is a powerful machine learning algorithm that uses convolutional operations for feature extraction and a layered combination of neural networks to learn complex data efficiently. The convolution operation refers to sliding a data window to different locations of the input data, then performing dot product and accumulation operations on the data in the window with a convolution kernel (also called a filter) to obtain the output feature map, as shown in Eq. (2):

$$a_{ij} = f\left(\sum_{m=0}^{2}\sum_{n=0}^{2}\omega_{m,n}x_{i+m,j+n} + \omega_b\right) \tag{2}$$

where x_{ij} refers to the value of the i_{th} row and j_{th} column of the input matrix, a_{ij} refers to the value of the i_{th} row and j_{th} column of the output matrix, f refers to the activation function, $w_{m,n}$ refers to the weight assigned by the filter to the m_{th} row and n_{th} column of the input matrix, w_b refers to the bias assigned by the filter.

CNN consists of multiple convolutional, pooling, activation functions, and fully connected layers (Fig. 7). The input values are passed through the layers to the final output layer. Eq. (3) is the formula for calculating the difference between the output and the true values.

$$J = \frac{1}{2}\left(y_{fianl} - y_{true}\right)^2 \tag{3}$$

During the training process, CNN improves the accuracy and generalization of the model by continuously adjusting the parameters of the convolutional kernel and the weights between neurons, as described in Eqs. (4) and (5):

$$(W, b) = \arg\min\left(J\left(w, b\right)\right) \tag{4}$$

$$(W, b)^{n+1} = (W, b)^n - r \cdot grad\left(J\left(W, b\right)^n\right) \tag{5}$$

where r refers to the learning rate, and $grad$ refers to the derivative of the loss function J to the w or b.
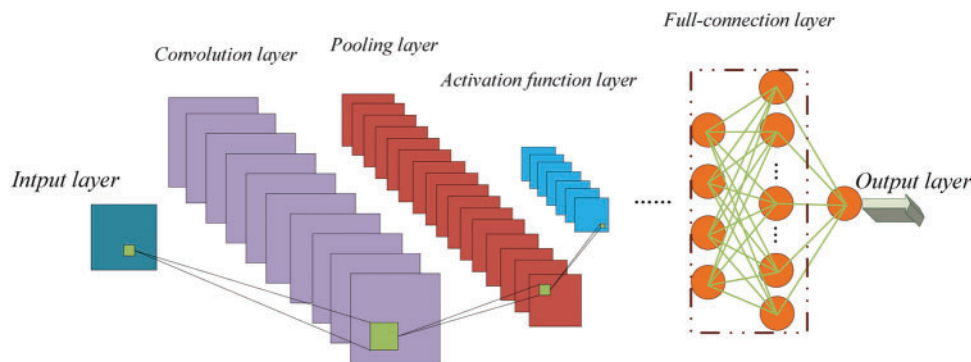
Figure 7: CNN algorithm

3.2 Genetic Algorithm

The Genetic Algorithm (GA) is a powerful optimization algorithm that simulates the process of natural selection and biological evolution in nature. As shown in Fig. 8, the GA converts multiple problem solutions into corresponding chromosomes, which are then evaluated based on their fitness to the problem. The chromosomes with higher fitness are selected and undergo genetic operations, including gene exchange and mutation. The process of selection, elimination, mutation, and exchange

is repeated multiple times until the remaining chromosomal genes represent the most optimal solution to the problem.

Figure 8: Genetic algorithm
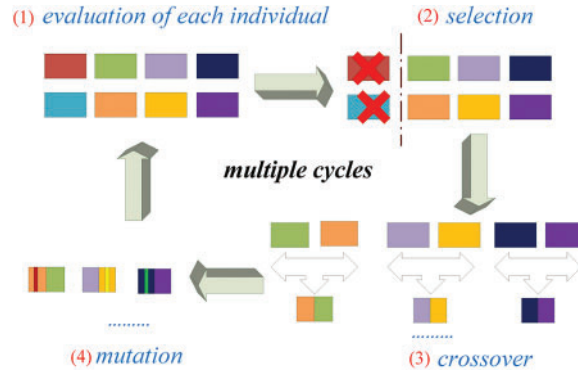
3.3 Evaluation Metric

Mean absolute error (MAE), root mean square error (RMSE), mean squared error (MSE), and the coefficient of determination R^2 are the metrics used to evaluate the model performance in this paper. MAE represents the average of the absolute error, which can better reflect the actual situation of the prediction value error. RMSE is a measure of the dispersion of the sample distribution, indicating the deviation between the observed value and the true value, and is often used as a standard for measuring the prediction results of machine learning models. MSE is a more common error calculation in statistics, which is the expectation of the error square between the true value and the predicted value. R^2 is a metric used to evaluate regression models, which indicates the proportion of the variance in the data that the model can explain. Generally, the closer to 1, the smaller the error between the predicted and true values. The above metrics are calculated by:

$$MAE = \sqrt{\frac{1}{N} \sum_{i=1}^{N} \left| \hat{y}_i - y_i \right|} \tag{6}$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^{N} \left(\hat{y}_i - y_i \right)^2} \tag{7}$$

$$MSE = \frac{1}{N} \sum_{i=1}^{N} \left(\hat{y}_i - y_i \right)^2 \tag{8}$$

$$R^2 = \frac{1 - \sum_{i=1}^{N} \left(\hat{y}_i - y_i \right)^2}{\sum_{i=1}^{N} \left(y_i - \bar{y}_i \right)^2} \tag{9}$$

where y_i is the true value, $\hat{y}_i$ is the predicted value, $\bar{y}_i$ is the average value of the true values, and N is the number of data.

To certify the robustness of the model, the cross-validation method was introduced in this study. The method evenly divides the whole data into several groups, selecting one as the testing dataset,

while the corresponding remaining part is the training dataset. Each data group was selected as the testing dataset and avoided overlapping with the data from the training dataset. Take the average of all testing dataset scores as the final model performance score, which is calculated in Eq. (10):

$$CV_{(k)} = \frac{1}{k} \sum_{i=1}^{k} R_i^2 \tag{10}$$

where k refers to the number of groups of all datasets that were divided, and R^2_i refers to the R^2 value when the i_{th} group dataset is the testing dataset.

4 The Optimization of the CNN Model

The structure of the CNN model, especially the number of filters per convolutional layer and neurons per full-connection layer, significantly impacts the model performance. Therefore, optimizing the model structure is crucial to achieving high accuracy and efficiency. In this study, the genetic algorithm was utilized to optimize the structure of the model. The designed CNN network comprises five layers, including three convolutional layers and two fully connected layers. The structural information of the model was encoded as chromosomes, as presented in Fig. 9.
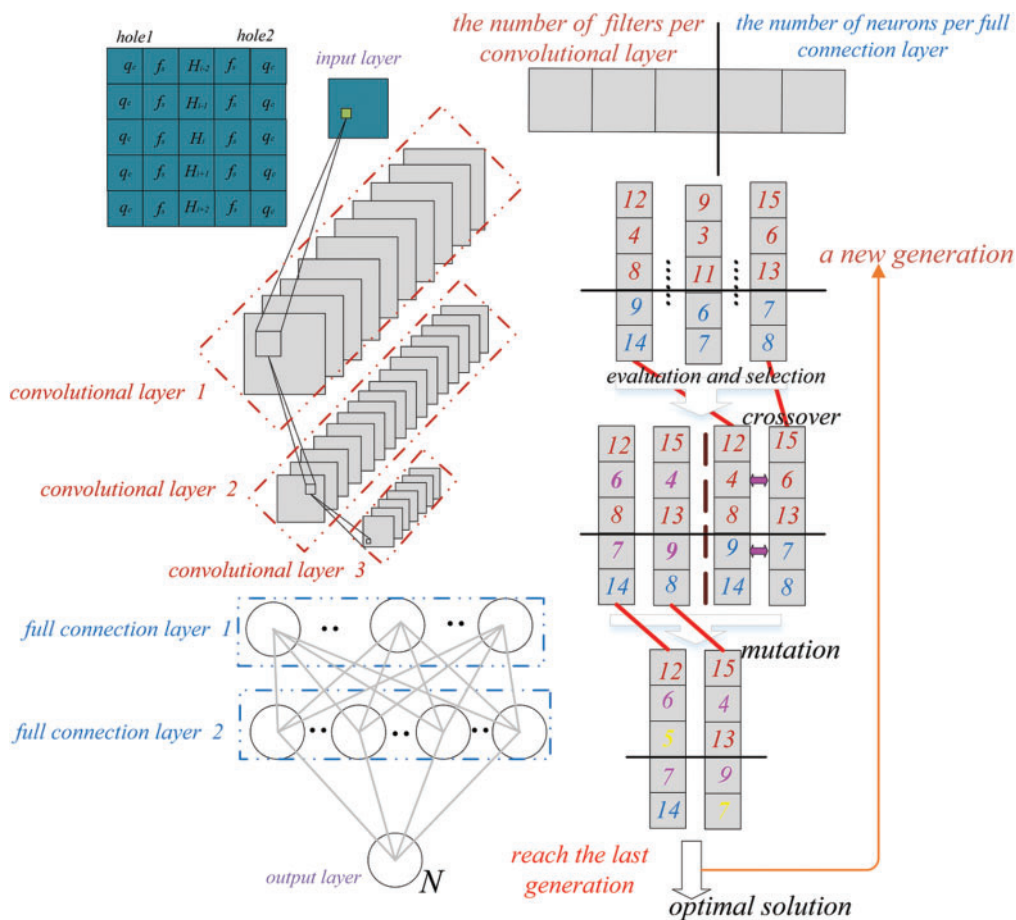

Figure 9: The optimization of the CNN model structure

Each chromosome consists of five parameters, with the first three determining the number of filters per convolutional layer, ranging from 1 to 20, and the last two determining the number of neurons per fully connected layer, ranging from 1 to 15. The fitness of each chromosome was calculated using Eq. (11), where chromosomes with higher fitness correspond to superior model structures and are thus more likely to be inherited to the next generation. The probability of inheritance for each chromosome was determined using Eq. (12), which calculates the ratio of the fitness of that chromosome to the sum of the fitness values of the entire population. The chromosomes with greater fitness have a higher probability of being selected and could potentially be selected multiple times, further improving the average performance of the next-generation model. The hyperparameters of the genetic algorithm are summarised in Table 2. For more information, refer to Appendix A for the pseudocode of optimizing the CNN model using the genetic algorithm.

$$fitness_i = CV_{(10,R^2)} (population_i) = CV_{(10,R^2)} (Model_i (a, b)) \tag{11}$$

$$probability_i = \frac{fitness_i}{\sum\limits_{k=1}^{K} fitness_k} \tag{12}$$

where i refers to each population, K refers to the number of populations, and a, b refers to the number of filters per convolutional layer and the number of neurons per full connection layer.

Table 2: GA parameters used for hyper-parameter tuning

GA parameter	Description
Fitness function	Eq. (11)
Selection method	Eq. (12)
Genetic possibility	Crossover (80%) Mutation (5%)
Number of chromosomes	50
Number of generations	20

The optimal combination of filter and neuron numbers for the model could be identified by using the GA algorithm for tuning the hyperparameters of the CNN model. Fig. 10 illustrates the convergence of the average population fitness after several generations of selection, exchange, and mutation, indicating the effectiveness of the Genetic Algorithm in hyperparameter tuning. In our study, the optimized model had three convolutional layers with 9, 7, and 13 filters, respectively, and two full-connection layers with 12 and 8 neurons. The use of the Genetic Algorithm greatly reduced the workload of the optimization process compared to permutation or variable analysis, which is time-consuming and inefficient.

Table 3 compares the performance of the optimized CNN model with other machine-learning algorithms. The results indicate that the optimized CNN model outperforms the other models with a minimum error, demonstrating that it is the most suitable model for this particular scenario. The R2 values obtained for the optimized CNN model (0.797) are higher than those obtained for the other models, including Linear Regression (0.743), SVR (0.746), XGBRegressor (0.757), and AdaBoostRegressor (0.682). Furthermore, the feature matrix computation based on the convolution principle is interpretable in fusing multifaceted features, which enhances the researchers' comprehension of the

model. Therefore, the optimized CNN model is a promising approach for predicting the outcome of this specific task.
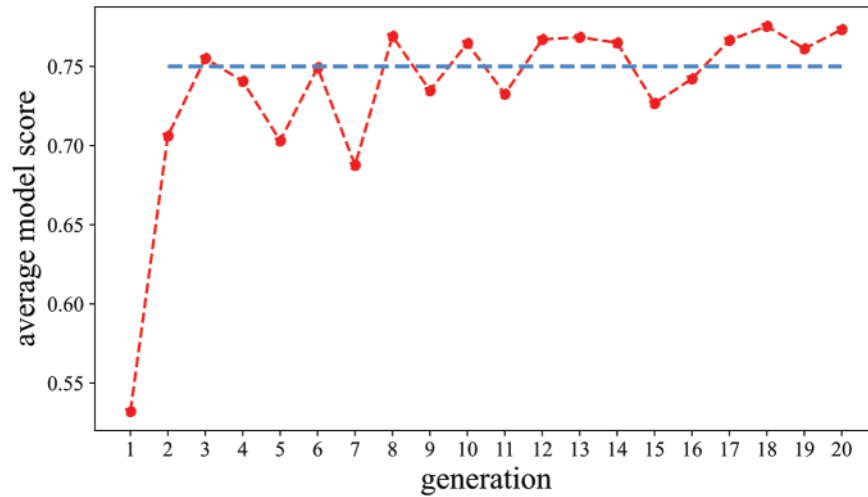

Figure 10: Average value of model score *vs*. generation

Table 3: Comparative results with other machine learning models

Algorithm	R^2
LinearRegression	0.743
SVR	0.746
XGBRegressor	0.757
AdaBoostRegressor	0.682
CNN-GA	0.797

5 Analysis of the Built Model

5.1 Comparison with the Equation Methods

This section compares the prediction of the CNN-GA model with equation methods in terms of N values for various soil categories. The characteristics of datasets from the same engineering site are likely to be similar. Therefore, a model trained on data from one engineering site may perform well when tested on data from the same site, but it may not perform as well when tested on data from a different site. To ensure the robustness and generalizability of the models across different sites, the training and testing datasets were divided based on engineering sites. A cross-validation approach was utilized, where the test set consisted of data from one engineering site while the training set comprised data from the other engineering sites. This approach ensured that the models were trained and tested on a diverse set of data, enabling them to be more robust and generalizable. This method provides a stringent test of the generalizability of our model, as it ensures that there is no overlap or similarity between the data in the training and testing sets.

This section only analyzes the performance of the models when data from engineering site 1 is used as the testing set and the data from the remaining engineering sites are used as the training set,

and other cases are detailed in Appendix B. The soil at engineering site 1 contains two main categories, including sandy silt and silty sand, and this section analyzes them separately. The comparison results are presented through scatter distribution and frequency distribution graphs. The scatter distribution graph shows the true and predicted values on the horizontal and vertical axes, respectively. Points closer to the diagonal indicate a smaller discrepancy between the predicted and true values. In a frequency distribution plot, the ratio of predicted to true values for a given sample is represented by the variable n. The frequency distribution curve of n is more centralized at $x = 1$ means the prediction result is accurate.

For sandy silt (Fig. 11), Acka [25] and Shahri et al. [27] consistently underestimated true N values, and Zhao et al. [28], Li et al. [29], and Liu et al. [30] performed poorly for data with true values above 30. In contrast, the CNN-GA model provided more accurate predictions, especially for data with smaller true values.
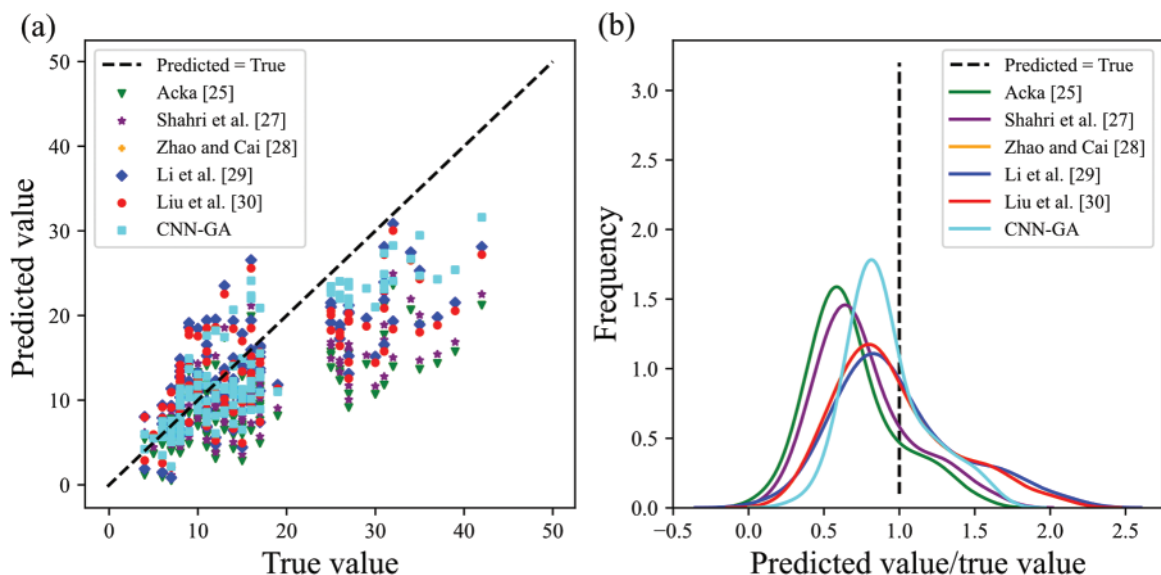

Figure 11: The comparative prediction results of different methods for N (sandy silt)

Compared to the CNN-GA model, the traditional methods of Acka [25], Dos Santoz et al. [26], Shahri et al. [27], Zhao et al. [28], Li et al. [29], and Liu et al. [30] for predicting N values in silty sand yielded poor results (Fig. 12). Specifically, Acka [25], Dos Santoz et al. [26], and Shahri et al. [27] produced predicted values that were significantly far from the true N values; Zhao et al. [28], Li et al. [29], and Liu et al. [30] consistently underestimated the true N values. In contrast, the CNN-GA model demonstrated superior predictive performance for silty sand, outperforming traditional methods in terms of accuracy.

Table 4 presents the error metrics of various models for different soil categories, including sandy silt, silty sand, and clayey silt. The results demonstrate that the CNN-GA model studied in this paper outperformed other studies, achieving lower MAE, RMSE, and MSE values for all three soil categories. In addition to the improved accuracy, the CNN-GA model is a comprehensive model that predicts soil parameters without prior knowledge of soil category, which makes it an attractive alternative to other methods.
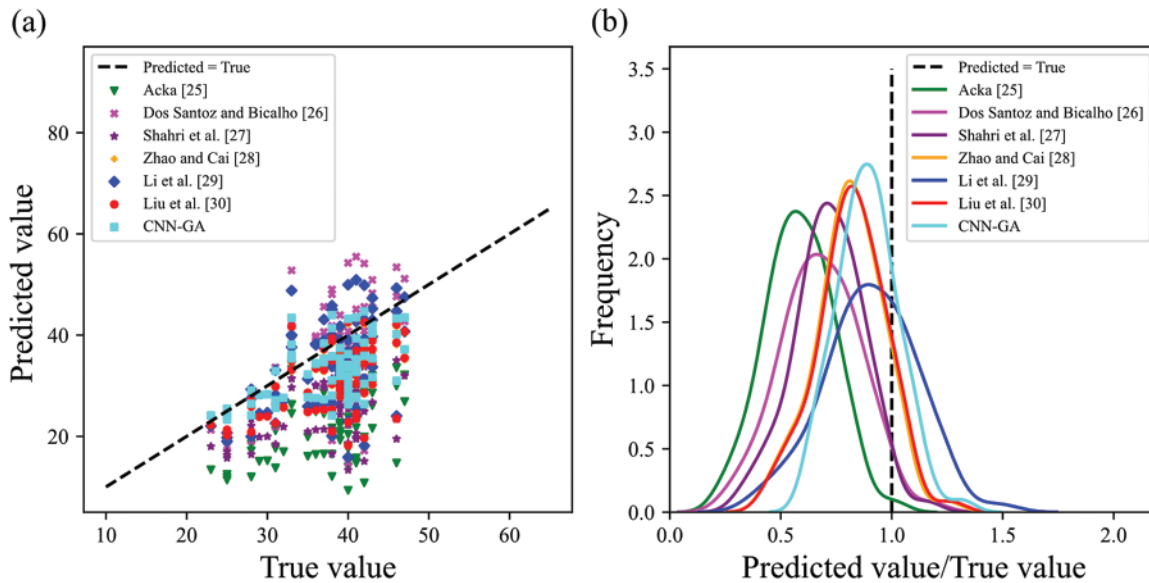
Figure 12: The comparative prediction results of different methods for N (silty sand)

Table 4: The performance comparison of the different method

Different methods	MAE	RMSE	MSE	Soil category
Acka [25]	6.479	8.476	71.840	Sandt silt
Shahri et al. [27]	5.890	7.829	61.296	
Zhao et al. [28]	4.796	6.381	40.715	
Li et al. [29]	4.721	6.221	38.751	
Liu et al. [30]	5.018	6.695	44.827	
CNN-GA	3.746	4.708	22.164	
Acka [25]	15.675	16.866	284.465	Silty sand
Dos Santoz et al. [26]	7.656	9.620	92.541	
Shahri et al. [27]	10.868	12.454	155.108	
Zhao et al. [28]	7.751	9.513	90.489	
Li et al. [29]	6.787	8.798	77.426	
Liu et al. [30]	7.553	9.320	86.861	
CNN-GA	6.129	7.384	54.528	

5.2 The Analysis of CNN-GA Model Prediction Results

This section further analyzed the prediction accuracy of our proposed model for N values at different soil depths. Fig. 13 illustrates the discrepancy between the predicted and true values of N, and the prediction results for relatively deep soil layers are not well. However, the liquefaction discrimination problem focuses primarily on soils within 20 m below ground, so the poor predicted results for deep soils have less impact on the latter study of liquefaction discrimination. Table 5 summarizes the error metrics of different categories of soils in different depth ranges. The result showed that the prediction for the soil with a depth of less than 20 m has a minor error, which is more beneficial

to the subsequent liquefaction discrimination study. The results demonstrate that the model has minor prediction errors for soils with a depth of less than 20 m, which is particularly beneficial for subsequent liquefaction discrimination studies.
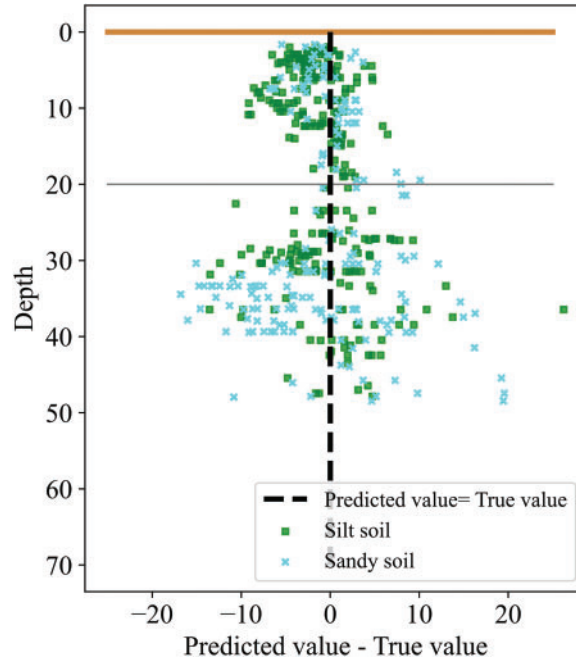

Figure 13: Analysis of CNN-GA model prediction results

Table 5: The comparative prediction error metric of different penetration depth H

	$0 < h < 20$ (m)			$h > 20$ (m)		
	MAE	RMSE	MSE	MAE	RMSE	MSE
Sandy silt	2.379	3.074	9.451	4.437	5.816	33.830
Silty sand	2.409	3.289	10.820	6.441	8.285	68.634

6 The Application of Liquefaction Discrimination Based on the Built Model

This section demonstrates the feasibility of utilizing the predicted N values for liquefaction discrimination. The approach for liquefaction discrimination utilized in this study was based on the code for the seismic design of buildings (GB50011-2010). According to this code, liquefaction is considered to occur when the standard penetration test blow count N of the saturated soil is less than or equal to the critical value N_{cr}. The formula for N_{cr} is given in Eq. (13):

$$N_{cr} = N_0 \beta \left[\ln\left(0.6 d_s + 1.5\right) - 0.1 d_w \right] \sqrt{3/\rho_c} \tag{13}$$

where N_0 is the reference value of the SPT for liquidation discrimination, d_s is the penetration depth (m), d_w is the depth of the groundwater level (m), ρ_c is the percentage of clay, and β is the regulation factor.

After determining the N_{cr} value based on soil factors, the predicted and measured N values were compared with N_{cr} for each exploration hole to assess the liquefaction potential. Take engineering site 1 as an illustration, and the liquefaction discrimination results are summarized in Table 6. The number 0 in the table indicates that the sample is classified as liquefied, while the number 1 indicates that the sample is classified as non-liquefied. $\sqrt{}$ represents consistent discriminatory results, while $\times$ represents inconsistent discriminatory results. The results showed that the discriminatory results based on predicted N values had a consistency of 75% with the results obtained from measured N values, indicating the reliability of the proposed model in evaluating the liquefaction potential of an engineering site. Moreover, using predicted N values can significantly reduce the workload and time required for geotechnical personnel to perform liquefaction assessments, particularly when many SPT exploration holes need to be tested. At the initial stage of liquefaction discrimination, a preliminary assessment can be conducted using the predicted N values, and the SPT can be selectively performed based on the preliminary results for further evaluation. This simplified and efficient approach can enhance the effectiveness of the geotechnical investigation, making our proposed model a valuable tool for geotechnical engineers.

Table 6: The comparison of liquefaction discrimination results

Borehole No.	Liquefaction discrimination		
	Measured values	Predicted values	Consistency
C01	0	0	$\sqrt{}$
C07	0	0	$\sqrt{}$
C13	0	0	$\sqrt{}$
C17	0	0	$\sqrt{}$
C19	1	1	$\sqrt{}$
C21	0	0	$\sqrt{}$
C25	0	1	$\times$
C27	0	0	$\sqrt{}$
C34	0	0	$\sqrt{}$
C36	0	0	$\sqrt{}$
C41	1	0	$\times$
C45	0	0	$\sqrt{}$
C46	0	0	$\sqrt{}$
C48	0	0	$\sqrt{}$
C50	1	0	$\times$
C52	1	0	$\times$
C55	1	1	$\sqrt{}$
C57	0	0	$\sqrt{}$
C59	0	0	$\sqrt{}$
C61	1	0	$\times$

7 Conclusion

The CPT feature matrix is a novel feature extraction method proposed in this study, which converts the traditional CPT test data into a matrix form for processing. Each element in the feature matrix represents the CPT data at a specific depth, such as tip resistance q_c and sleeve resistance f_s. By convolving the feature matrix, the multidimensional features were fused, effectively improving the prediction accuracy of the N value. Additionally, the Genetic Algorithm (GA) was employed to optimize the model and improve its accuracy and generalization. The study verified the prediction performance of the model for different soil categories, and the results showed that the method has higher prediction accuracy than the conventional method. Then the study validates the feasibility of liquefaction discrimination based on the developed model. The study utilized the predicted N values obtained from the model for liquefaction discrimination and compared the results with the measured N values, using the liquefaction discrimination criteria outlined in the seismic design code for buildings (GB50011-2010). The experimental results show that the agreement between the liquefaction discrimination results based on the predicted N values and the measured N values is 75%, which indicates that the proposed model has reliability in assessing the liquefaction potential of the engineering site.

Given the high level of uncertainty associated with geotechnical engineering, the feature matrix proposed in this study has the potential for broad applications in other geoengineering fields. The use of multidimensional features can improve the accuracy and reliability of predictions in contexts where data is often limited and uncertainty is high. Although this study yielded some meaningful results, some shortcomings need to be acknowledged. The primary deficiencies of this study include:

1. While the proposed method has shown promising results for liquefaction discrimination, its reliability needs to be further verified in multiple regions. There is also scope for refinement and enhancement of the method in the future. To this end, we intend to conduct further studies and collect data on liquefaction cases in China to verify the effectiveness of our proposed method.

2. The current study did not include a comparison with other state-of-the-art AI methods, such as related deep learning methods, because it would involve the authors' source code and experimental data. In future studies, we plan to include more comprehensive and systematic comparisons with other methods to provide a more thorough understanding of the effectiveness of our proposed method.

Author Contributions: Ruihan Bai wrote the main manuscript text and performed the experiments. Feng Shen gave guidance and suggestions for method of this article, as well as reviewed and edited the manuscript. Zihao Zhao summarized the methods of the existing related research. Zhiping Zhang drew the diagrams (model structure in the article and provided suggestions for improving the article's layout. Qisi Yu contributed to the check, review, and editing of the article. All authors reviewed and revised the manuscript.

Availability of Data and Materials: The data supporting the conclusions of this article are included within the article. Any queries regarding these data may be directed to the author Ruihan Bai (bairuihanup@163.com).

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

1. Seed, H. B., Idriss, I. M. (1967). Analysis of soil liquefaction: Niigata earthquake. *Journal of the Soil Mechanics and Foundations Division, 93(3),* 83–108.
2. Huang, Y., Yu, M. (2013). Review of soil liquefaction characteristics during major earthquakes of the twenty-first century. *Natural Hazards, 65,* 2375–2384.
3. Singh, S. (1996). Liquefaction characteristics of silts. *Geotechnical & Geological Engineering, 14,* 1–19.
4. Shu, S., Ge, B., Wu, Y., Zhang, F. (2023). Probabilistic assessment on 3D stability and failure mechanism of undrained slopes based on the kinematic approach of limit analysis. *International Journal of Geomechanics, 23(1),* 06022037.
5. Shibata, T., Oka, F., Ozawa, Y. (1996). Characteristics of ground deformation due to liquefaction. *Soils and Foundations, 36,* 65–79.
6. Goda, K., Atkinson, G. M., Hunter, J. A., Crow, H., Motazedian, D. (2011). Probabilistic liquefaction hazard analysis for four Canadian cities. *Bulletin of the Seismological Society of America, 101(1),* 190–201.
7. Subedi, M., Acharya, I. P. (2022). Liquefaction hazard assessment and ground failure probability analysis in the Kathmandu Valley of Nepal. *Geoenvironmental Disasters, 9(1),* 1–17.
8. Piya, B., van Westen, C., Woldai, T. (2004). Geological database for liquefaction hazard analysis in the Kathmandu Valley, Nepal. *Journal of Nepal Geological Society, 30,* 141–152.
9. Putti, S. P., Satyam, N. (2018). Ground response analysis and liquefaction hazard assessment for Vishakhapatnam City. *Innovative Infrastructure Solutions, 3,* 1–14.
10. Ansari, A., Zahoor, F., Rao, K. S., Jain, A. K. (2022). Liquefaction hazard assessment in a seismically active region of Himalayas using geotechnical and geophysical investigations: A case study of the Jammu Region. *Bulletin of Engineering Geology and the Environment, 81(9),* 349.
11. Seed, H. B., Idriss, I. M. (1971). Simplified procedure for evaluating soil liquefaction potential. *Journal of the Soil Mechanics and Foundations Division, 97(9),* 1249–1273.
12. Cetin, K. O., Seed, R. B., Kayen, R. E., Moss, R. E., Bilge, H. T. et al. (2018). Examination of differences between three SPT-based seismic soil liquefaction triggering relationships. *Soil Dynamics and Earthquake Engineering, 113,* 75–86.
13. Robertson, P. K., Campanella, R. G. (1985). Liquefaction potential of sands using the CPT. *Journal of Geotechnical Engineering, 111(3),* 384–403.
14. Robertson, P. K., Wride, C. E. (1998). Evaluating cyclic liquefaction potential using the cone penetration test. *Canadian Geotechnical Journal, 35(3),* 442–459.
15. Juang, C. H., Chen, C. J., Jiang, T., Andrus, R. D. (2000). Risk-based liquefaction potential evaluation using standard penetration tests. *Canadian Geotechnical Journal, 37(6),* 1195–1208.
16. Missoum, H., Belkhatir, M., Bendani, K., Maliki, M. (2013). Laboratory investigation into the effects of silty fines on liquefaction susceptibility of Chlef (Algeria) sandy soils. *Geotechnical and Geological Engineering, 31,* 279–296.
17. Liu, J. (2020). Influence of fines contents on soil liquefaction resistance in cyclic triaxial test. *Geotechnical and Geological Engineering, 38(5),* 4735–4751.
18. Zhao, Z., Duan, W., Cai, G., Wu, M., Liu, S. (2022). CPT-based fully probabilistic seismic liquefaction potential assessment to reduce uncertainty: Integrating XGBoost algorithm with Bayesian theorem. *Computers and Geotechnics, 149,* 104868.
19. Zhao, Z., Duan, W., Cai, G. (2021). A novel PSO-KELM based soil liquefaction potential evaluation system using CPT and Vs measurements. *Soil Dynamics and Earthquake Engineering, 150,* 106930.

20. Fang, Y., Jairi, I., Pirhadi, N. (2023). Neural transfer learning for soil liquefaction tests. *Computers & Geosciences, 171,* 105282.

21. Zhang, Y., Xie, Y., Zhang, Y., Qiu, J., Wu, S. (2021). The adoption of deep neural network (DNN) to the prediction of soil liquefaction based on shear wave velocity. *Bulletin of Engineering Geology and the Environment, 80,* 5053–5060.

22. Bray, J. D., Sancio, R. B. (2006). Assessment of the liquefaction susceptibility of fine-grained soils. *Journal of Geotechnical and Geoenvironmental Engineering, 132(9),* 1165–1177.

23. Shengcong, F., Tatsuoka, F. (1984). Soil liquefaction during Haicheng and Tangshan earthquake in China; A review. *Soils and Foundations, 24(4),* 11–29.

24. Huang, Y., Jiang, X. (2010). Field-observed phenomena of seismic liquefaction and subsidence during the 2008 Wenchuan earthquake in China. *Natural Hazards, 54,* 839–850.

25. Akca, N. (2003). Correlation of SPT-CPT data from the United Arab Emirates. *Engineering Geology, 67(3–4),* 219–231.

26. Dos Santos, M. D., Bicalho, K. V. (2017). Proposals of SPT-CPT and DPL-CPT correlations for sandy soils in Brazil. *Journal of Rock Mechanics and Geotechnical Engineering, 9(6),* 1152–1158.

27. Shahri, A., Juhlin, C., Malemir, A. (2014). A reliable correlation of SPT-CPT data for Southwest of Sweden. *Electronic Journal of Geotechnical Engineering, 19,* 1013–1032.

28. Zhao, X., Cai, G. (2015). SPT-CPT correlation and its application for liquefaction evaluation in China. *Marine Georesources & Geotechnology, 33(3),* 272–281.

29. Li, G., Guo, W., Qiao, L. (2018). Study on the method of site liquefaction discrimination based on CPT in Jiangsu area. *Site Investigation Science and Technology, 221(1),* 10–13.

30. Liu, X., Huang, Z., Ren, Y., Chen, N. (2012). The analysis of SPT-CPT correlation and influencing factors. *Science Technology and Engineering, 12(28),* 7443–7448.

31. Tarawneh, B. (2017). Predicting standard penetration test N-value from cone penetration test data using artificial neural networks. *Geoscience Frontiers, 8(1),* 199–204.

32. Fernando, H., Nugroho, S. A., Suryanita, R., Kikumoto, M. (2021). Prediction of SPT value based on CPT data and soil properties using ANN with and without normalization. *International Journal of Artificial Intelligence Research, 5(2),* 123–131.

33. Gupta, A. K., Alla, V., Suneel Kumar, G., Behera, R. N. (2022). Development of correlations between SPT-CPT data for liquefaction assessment using R. http://dspace.nitrkl.ac.in/dspace/handle/2080/3820

34. Jin, K. H., McCann, M. T., Froustey, E., Unser, M. (2017). Deep convolutional neural network for inverse problems in imaging. *IEEE Transactions on Image Processing, 26(9),* 4509–4522.

35. Rawat, W., Wang, Z. (2017). Deep convolutional neural networks for image classification: A comprehensive review. *Neural Computation, 29(9),* 2352–2449.

36. Williamson, D. F., Parker, R. A., Kendrick, J. S. (1989). The box plot: A simple visual method to interpret data. *Annals of Internal Medicine, 110(11),* 916–921.

Appendix A.

Algorithm 1: Genetic Algorithm for Optimizing CNN Hyperparameters
Input: Fitness evaluation function F (Eq. (11)), Training data D
Output: Optimized hyperparameters for the CNN model

Algorithm 1 (Continued)

1: *DNA size*←5 // The first three values are the number of convolutional filters corresponding to the three convolutional layers, and the last two values are the number of neurons corresponding to the two fully connected layers

2: *Population size*←*m*

3: *Crossover rate*←p_c

4: *Mutation rate*←p_m

5: *Generations*←*n*

6: Initialize an array of size (*Population size*, *DNA size*) with zeros and assign it to *P*(*t*)

7: **for** *i* in the range of *Population size* **do**

8:   Generate three random integers from 1 to 20 and assign them to *Pop filters*

9:   Generate three random integers from 1 to 15 and assign them to *Pop neutrons*

10:   Concatenate *Pop filters* and *Pop neutrons* into a single array and assign it to *Pop stack*

11:   Assign *Pop stack* to the i_{th} row of *P*(*t*)

12: **end for**

13: **for** *i* in the range of *Generations* **do**

14:   Initialize an array of size *Population size* with zeros and assign it to *Fitness*

15:   **for** *i* in the range of *Population size* **do**

16:     Convert the i_{th} row of *P*(*t*) to a list and assign it to *Pop list*

17:     Convert each element in the *Pop list* to build the *CNN model*

18:     Evaluate the fitness of the *CNN model* using the fitness evaluation function *F* and training dataset *D*

19:     Assign the fitness to the i_{th} element of *Fitness*

20:   **end for**

21:   Select a new population *P′*(t) from the current population *P*(t) using tournament selection based on *Fitness*

22:   **for** each *individual* in *P′*(*t*) **do** // Crossover(*P′*(*t*)) with probability p_c

23:     **if** rand() < p_c **then**

24:       Select another individual *mate* from *P′*(*t*)

25:       Choose a random crossover point *cross_point*

26:       *individual* [*cross_point*]←*mate* [*cross_point*]

27:       *mate* [*cross_point*]←*individual* [*cross_point*]

28:     **end if**

29:     **for** each *point* in *DNA size* **do** // Mutation(*P′*(*t*)) with probability p_m

30:       **if** rand() < p_m **then**

31:         **if** *point* >3 **then** // change the number of neurons corresponding to the two fully connected layers

32:           Generate a random integer from 1 to 15 and assign it to *individual* [*point*]

33:         **else** // change the number of convolutional filters corresponding to the three convolutional layers

34:           Generate a random integer from 1 to 20 and assign it to *individual* [*point*]

35:         **end if**

36:       **end if**

37:     **end for**

38: **end for**

Algorithm 1 (Continued)

39:   Evaluate the fitness of $P'(t)$ using the fitness evaluation function F and training dataset D
40:   Merge $P(t)$ and $P'(t)$ into a combined population Q
41:   Sort Q in descending order of fitness
42:   Select the top *Population size* individuals from Q to form the new population $P(t+1)$
43:   Set the best individual in $P(t+1)$ as the new optimal hyperparameters
44: **end for**
45: Return the highest-fitness individual as the result

Note:

1. The first three values of each DNA are the number of convolutional filters corresponding to the three convolutional layers, and the last two values of each DNA are the number of neurons corresponding to the two fully connected layers.

2. At the crossover point, the initial three genes encoding the number of filters in each convolutional layer can be interchanged between two individuals. Moreover, at an additional randomly-selected crossover point, the final two genes representing the number of neurons in the two fully connected layers can be exchanged between the same two individuals. This process generates two novel individuals that amalgamate the genetic information of the original two individuals.

3. Mutation provides a mechanism to explore new regions of the search space and can prevent premature convergence to sub-optimal solutions. The genes representing the number of filters in the convolutional layer or the number of neurons in the fully connected layer can randomly generate new values during mutation operation.

Appendix B.

Table: The performance comparison of the different method

Testing dataset	Different methods	MAE	RMSE	MSE	Soil category
	Acka [25]	6.479	8.476	71.840	
	Shahri et al. [27]	5.890	7.829	61.296	
	Zhao et al. [28]	4.796	6.381	40.715	Sandy silt
	Li et al. [29]	4.721	6.221	38.751	
	Liu et al. [30]	5.018	6.695	44.827	
Engineering site 1 (including sandy silt and silty sand)	CNN-GA	3.746	4.708	22.164	
	Acka [25]	15.675	16.866	284.465	
	Dos Santoz et al. [26]	7.656	9.620	92.541	
	Shahri et al. [27]	10.868	12.454	155.108	
	Zhao et al. [28]	7.751	9.513	90.489	Silty sand
	Li et al. [29]	6.787	8.798	77.426	
	Liu et al. [30]	7.553	9.320	86.861	
	CNN-GA	6.129	7.384	54.528	

(Continued)

Table (continued)

Testing dataset	Different methods	MAE	RMSE	MSE	Soil category
Engineering site 2 (including sandy silt and silty sand)	Acka [25]	6.298	8.894	79.106	Sandy silt
	Shahri et al. [27]	5.689	8.392	70.431	
	Zhao et al. [28]	4.981	7.517	56.501	
	Li et al. [29]	5.260	7.755	60.137	
	Liu et al. [30]	5.062	7.652	58.557	
	CNN-GA	2.911	4.446	19.769	
	Acka [25]	6.187	10.960	120.125	Silty sand
	Dos Santoz et al. [26]	4.693	8.098	65.583	
	Shahri et al. [27]	4.859	9.2554	85.637	
	Zhao et al. [28]	7.226	9.634	92.807	
	Li et al. [29]	4.464	7.978	63.655	
	Liu et al. [30]	7.204	9.579	91.763	
	CNN-GA	2.369	3.597	12.935	
Engineering site 3 (including sandy silt, silty sand, and clayey silt)	Acka [25]	8.311	10.593	112.201	Sandy silt
	Shahri et al. [27]	7.793	9.941	98.820	
	Zhao et al. [28]	7.067	8.628	74.435	
	Li et al. [29]	6.889	8.537	72.886	
	Liu et al. [30]	7.218	8.864	78.569	
	CNN-GA	4.022	5.207	27.114	
	Acka [25]	18.306	18.645	347.647	Silty sand
	Dos Santoz et al. [26]	12.814	12.949	167.685	
	Shahri et al. [27]	14.523	14.822	219.706	
	Zhao et al. [28]	9.533	10.218	104.400	
	Li et al. [29]	12.003	12.155	147.745	
	Liu et al. [30]	9.432	10.097	101.948	
	CNN-GA	3.691	4.390	19.276	
	Li et al. [29]	12.003	12.155	147.745	Clayey silt
	CNN-GA	3.691	4.390	19.276	
Engineering site 4 (including sandy silt and silty sand)	Acka [25]	5.054	6.590	43.426	Sandy silt
	Shahri et al. [27]	4.322	6.030	36.367	
	Zhao et al. [28]	3.500	5.557	30.879	
	Li et al. [29]	3.820	5.786	33.478	
	Liu et al. [30]	3.966	5.887	34.662	
	CNN-GA	3.463	5.422	29.402	
	Acka [25]	6.573	9.010	82.807	
	Dos Santoz et al. [26]	4.039	5.107	20.086	
	Shahri et al. [27]	4.731	6.804	46.288	

Table (continued)

Testing dataset	Different methods	MAE	RMSE	MSE	Soil category
	Zhao et al. [28]	6.179	7.224	52.189	Silty sand
	Li et al. [29]	3.767	4.931	24.318	
	Liu et al. [30]	6.113	7.123	50.736	
	CNN-GA	3.656	5.117	26.186	
Engineering site 5 (silty sand)	Acka [25]	22.162	27.757	770.452	Silty sand
	Dos Santoz et al. [26]	14.611	20.291	411.749	
	Shahri et al. [27]	18.175	24.226	586.899	
	Zhao et al. [28]	15.332	21.651	461.784	
	Li et al. [29]	14.626	20.365	414.745	
	Liu et al. [30]	15..194	21.477	461.276	
	CNN-GA	11.052	13.834	191.383	