

Received 11 October 2022, accepted 19 November 2022, date of publication 23 November 2022, date of current version 30 November 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3224184

RESEARCH ARTICLE

A Novel Hybrid Deep Neural Network Prediction Model for Shield Tunneling Machine Thrust

CHENG CHEN¹, BEN WU¹, PENGJIAO JIA¹, AND ZHANSHENG WANG²

¹School of Rail Transportation, Soochow University, Suzhou, Jiangsu 215000, China

²Suzhou Rail Transit Group Company Ltd., Suzhou, Jiangsu 215004, China

Corresponding author: Pengjiao Jia (pjia@suda.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 52108380, in part by the Natural Science Foundation of Jiangsu Province under Grant BK20210721, in part by the Postdoctoral Research Foundation of China under Grant 2021M702400, and in part by the Natural Science foundation for Colleges and Universities in Jiangsu Province under Grant 21HJB560018.

ABSTRACT Shield thrust is a critical operational parameter during shield driving, which is of vital significance for adjusting operational parameters and ensuring efficient and safe propulsion of shield tunneling machine. In this paper, a novel hybrid prediction model (CLM) combining attention mechanism, convolutional neural networks (CNN) and Bi-directional long short-term memory (BiLSTM) network is proposed for shield thrust prediction. Correlation analysis based on Maximal Information Coefficient (MIC) between the thrust and other parameters is first conducted to select optimal parameters and reduce input dimension. An attention mechanism is introduced into CNN to distinguish the importance of different features, with the convolution layer and pooling layer further extracting dimension features of the data. Then, a BiLSTM neural network integrating first attention layer is employed to extract time-varying characteristics of the data, with a second attention layer added to capture important time information. Field data during shield cutting bridge piles are investigated to support and validate the effectiveness and superiority of the proposed method. Results show that the proposed CLM model are general enough to avoid overfitting problems and have good performance at prediction. The predicted value match reasonably well the monitoring data, with coefficient of determination (R^2) equaling to 0.85, root mean square error (RMSE) equaling to 0.05, mean absolute error (MAE) equaling to 0.02. The CLM model in this paper can accurately predict the thrust even under complicated construction conditions, which provides reference for similar industrial application.

INDEX TERMS Shield machine, deep learning, thrust prediction, high-dimension, time series.

I. INTRODUCTION

With advancing tunnelling technique and booming demands for underground transportation, the subway mileage has increased dramatically worldwide in recent decades, and is expected to keep increasing in the future. Dense subway network and limited construction sites frequently have had subway tunnels cross existing underground structures, such as steel reinforced concrete piles supporting bridges or buildings [1], [2]. The concrete piles, if encountered, used to be pulled out from ground surface or cut manually through vertical shaft or from machine face. The manual pile removal is costly, risky and time consuming [3]. More and more projects

The associate editor coordinating the review of this manuscript and approving it for publication was Sunil Karamchandani¹.

start to cut through piles directly using shield machines equipped with strong disc or shell cutters.

During the cutting process, the change of the shield thrust is a result of complicated machine-ground-structure interaction which is a function of a variety of parameters, including machine parameters (e.g., torque, thrust, chamber pressure, cutter wear, and cutterhead temperature), ground properties (e.g., soil type, strength, and stiffness), structure properties (e.g., diameter, concrete strength, and steel arrangement) and so on [4], [5]. Shield thrust is important since its value directly affects the safety of superstructure and shield cutting efficiency. Therefore, precisely predicting the shield load can help engineers adjust operational parameters before shield cutting, which ensures the safety of superstructure and shield machine. The prediction method of the shield thrust

mainly includes theoretical analysis, numerical simulation, and machine learning. Previous prediction methods for shield load are composed of two types: theoretical analysis methods and numerical simulation methods. For theoretical analysis methods, the Colorado School of Mines (CSM) was first proposed and widely used in the calculation of shield loads [6]. In order to study the impact of geological condition, Wang et al. [7] proposed a new thrust force model under the single geological condition based on an assumption that the shield excavation is very close to equilibrium. Considering the limitations of the single geological conditions, Zhang et al. [8] established a theoretical prediction model for shield loads based on the impact of soil-rock interbedded ground on shield loads, and the proposed model had been proved to be effective. Zhou et al. [9] established a theoretical torque prediction model under mixed geological conditions. In addition, the operational and structural parameters are also important factors that affect shield loads, Yagiz et al. [10] established theoretical prediction model to predict shield loads using polynomial regression under geological conditions and operational parameters. Zhang et al. [11] proposed a prediction model for shield loads through combined analysis of geological, operational and structural parameters. For numerical simulation methods, Han et al. [12] simulated the shield driving using three-dimensional finite element method, which can obtain cutterhead torque variation curve under different geological conditions. Wu and co-workers [13] established a three-dimensional model estimating the average value of cutterhead torque in a certain distance or time. Faramarzi et al. [14] used the discrete element method to estimate the TBM torque and thrust, and achieved higher prediction accuracy than theoretical analysis. These theoretical analysis and numerical simulation methods provide certain guidance for the shield driving. However, the prediction accuracy is affected by various factors, such as complex geological conditions and cutter wear, which are not considered in theoretical and numerical methods. Besides, the change of the thrust is a complex dynamic process, thus it can hardly be clearly figure out simply through theoretical and numerical analysis. Over the years, since the monitoring method of shield construction is becoming more and more diversified, such as automation equipment, fiber, and so on, there are more available data with large scales and high dimensions for shield load analysis. Hence, one burning issue is to choose appropriate method to analyze shield loads and extract data features.

In order to effectively explore the variation of shield load and precisely predict the value of shield thrust, this paper proposes a CNN-BiLSTM-Multiattention (CLM) method for shield thrust prediction. Due to convolutional neural networks (CNNs) perform well at capturing the dimension features of the data, bi-directional long short-term memory network (BiLSTM) can extract the time-varying characteristics of time series and attention mechanism can focus on important information to increase fitting abilities of the prediction models, CLM model in this paper can not only captures the

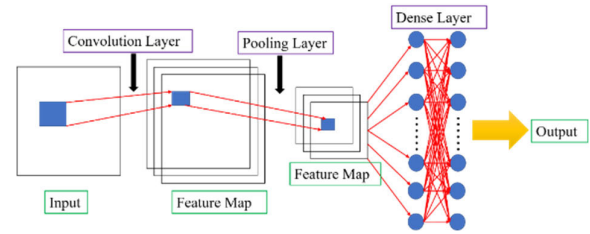


FIGURE 1. CNN structure diagram.

time-varying characteristics of data, but also extracts the data dimensions features, and highlights the important dimension and key time information to improve the prediction accuracy.

The study is organized as follows: section 2 presents the related work about shield thrust prediction. Section 3 introduces the materials. Section 4 explains the proposed algorithm. Section 5 presents the preparing work, including data preprocessing, experimental environment, model establishment and metrics. Section 6 discusses the results. The conclusion is drawn in Section 7.

II. RELATED WORK

A. DEEP LEARNING

1) CONVOLUTIONAL NEURAL NETWORKS (CNN)

CNNs are similar to ordinary neural networks in that neurons are composed of learnable weights and bias constants. As shown in Figure 1, CNN is mainly composed of convolutional layers, pooling layers and fully connected layers. The input can be regarded as a grayscale image. Each convolutional layer is consisted of several convolutional units, and the parameters of each convolutional unit are optimized by the back-propagation algorithm. The purpose of the convolutional operation is to extract different features of input parameters. The first convolutional layer containing several filters is to obtain feature maps. Every filter is a weight matrix with local connections and shared weights, which can convolve an original image to a corresponding feature map that can be considered an image representation extracted by the filter. The first convolutional layer may only extract some low-level features such as edges, lines, and corners. More network layers can iteratively extract more complex features from the low-level features. A nonlinear layer (or activation layer) is usually applied immediately after each convolutional layer. The purpose is to introduce nonlinear features. Moreover, a pooling layer is inserted periodically between successive convolutional layers. It aims to gradually reduce the spatial size of the data volume, thus reducing the number of parameters in the network and avoiding over-fitting problem. The fully connected (dense) layer can be used to map the final output to linearly separable space. The input can be regarded as a grayscale image. The convolution formula is expressed as [15]:

$$a_{i,j} = f \left(\sum_{m=0}^k \sum_{n=0}^k w_{m,n} x_{i+m,j+n} + w_b \right) \quad (1)$$

where $x_{i,j}$ represents the i row and j column of the input image, $w_{m,n}$ is the m row and n column of $k \times k$ weights matrix, w_b is the filter bias, f is an activation function, and $a_{i,j}$ is the value of the i row and j column of the feature map.

2) LONG SHORT-TERM MEMORY NETWORK (LSTM)

The LSTM algorithm [16], [17], [18] is a variant structure of RNN, which can avoid the problem of gradient disappearance and explosion caused by long sequence training of RNN. LSTM is mainly composed of three parts: forget gate, input gate and output gate. The forget gate is mainly used to determine the retention and discarding of information in the cell, the input gate mainly determines the partial input and shielding of the information, and the output gate mainly determines the output information. The cell type is shown in Fig. 2.

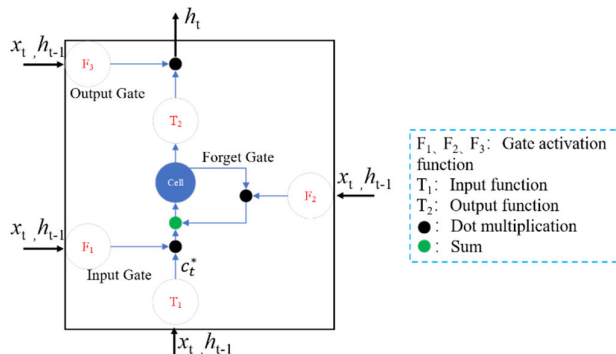


FIGURE 2. LSTM structure diagram.

The candidate state is expressed as

$$c_t^* = \tanh(W_{xc}x_t + W_{hc}h_{t-1} + b_c) \quad (2)$$

where x_t is the input at the current moment; h_{t-1} is the output of the previous neural unit; W_{xc} is the weighting of the input parameter x_t and the memory unit; W_{hc} is the weighting of the h_{t-1} and the memory unit; b_c is bias vector.

The input gate, forget gate and output gate are calculated as

$$I_t = \sigma(W_{xi}x_t + W_{hi}h_{t-1} + b_i) \quad (3)$$

$$F_t = \sigma(W_{xf}x_t + W_{hf}h_{t-1} + b_f) \quad (4)$$

$$O_t = \sigma(W_{xo}x_t + W_{ho}h_{t-1} + b_o) \quad (5)$$

where W_{xi} , W_{xf} and W_{xo} are the weighting of the input parameter x_t and gate units, respectively; W_{hi} , W_{hf} and W_{ho} is the weighting of the input parameter h_{t-1} and gate units; b_i , b_f and b_o are bias vectors of three gate units; and σ is sigmoid function with value ranging from 0 to 1.

The update status in the neural unit cell is:

$$c_t = F_t c_{t-1} + c_t^* I_t \quad (6)$$

where c_{t-1} is the stored value at the previous moment.

The output value of the LSTM unit is:

$$h_t = O_t \tanh(c_t) \quad (7)$$

where h_t is the output value of the neural unit at the current moment.

3) ATTENTION MECHANISM

Attention mechanism is firstly applied in the human visual system [19], it can capture the important and ignore the unimportant from enormous information. Currently, attention mechanism is an important component of the neural network, which is widely used in areas of natural language processing, statistical learning and so on. Moreover, it has been widely used in the RNN and LSTM algorithms to solve time series tasks. Attention mechanism can extract important time information, then assign different weights to information at different moments. The calculation process of the Attention mechanism can be summarized into three stages: (1) calculating the similarity or correlation between *Query* and *Key*; (2) normalizing the original scores in the first stage; (3) weighting the summation of *Value* according to the weight coefficients. As shown in Figure 3.

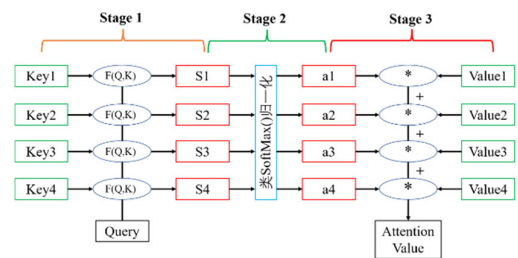


FIGURE 3. Attention structure diagram.

The calculation formula is expressed as:

$$\begin{aligned} Attention(Query, Source) &= \sum_{i=1}^{L_x} Similarity(Query, Key_i) \cdot Value_i \end{aligned} \quad (8)$$

where *Source* is input parameters, $\{x_1, x_2, x_3 \dots x_n\}$; $Value_i$ is the value of each element in *Source*; L_x is the length of the input parameter; *Query* is an element of the output parameter; Key_i can be regarded as the address of each element in *Source*.

B. INTELLIGENT PREDICTION FOR SHIELD LOAD

In addition to the theoretical analysis methods and numerical simulation methods, with the rapid development of machine learning techniques, the internal characteristics of a vast amount of monitoring data can be explored and fed back to engineering constructions. Deep learning as an important branch of machine learning, deep learning has been widely used in engineering to predict the shield load over the past years due to the ability to extract the law of data. Gao et al. [20] adopted three different recurrent neural network (RNN) models to predict TBM thrust and thrust in real-time based on in-situ operating data. Zhang et al. [21] found that LSTM model is better suitable for predicting the shield load than random forest (RF) model. Chen et al. [22]

predicted torque and thrust based on an improved LSTM algorithm, and making it possible to adjust the TBM tunneling parameters in real time. Qin et al. [23] combined convolutional neural network (CNN) and long short-term memory (LSTM) to extract implicit features and sequential features for cutterhead torque prediction. Zhou et al. [24] proposed a multi-step shield load and attitude prediction method of shield tunneling machine based on WCNN-LSTM neural network. Shi et al. [25] proposed a novel hybrid multi-step prediction model for shield machine cutterhead torque. The model, combining variational mode decomposition (VMD), empirical wavelet transform (EWT) and long short-term memory (LSTM) network, can accurately predict cutterhead torque of shield tunneling machine in multiple time steps. Xu et al. [26] successfully predicted the shield thrust using five different statistical and ensemble machine learning methods.

It can be seen that the deep learning methods have had many achievements on load prediction. However, their accuracy and practicability still need improvement. On the one hand, prediction methods above mainly study the time-varying characteristics of the data, and ignore the dimension features of the data. On the other hand, these existing predicting methods don't take it into account the influence of the different features and time information on the predicting results.

C. CONTRIBUTIONS

The contributions and innovations of this paper are concluded as follows:

- (1) In this work, we propose a novel hybrid model for precise shield thrust prediction. The proposed CLM model combines CNN algorithm, BiLSTM algorithm and attention mechanism. The operational, geological, structure and tunnel parameters are selected as input, and the output is the thrust at the next time.
- (2) The CLM model integrates advantages of various algorithms. It not only captures dimensions features and the time-varying characteristics of time series, but also highlights important dimensions and key time information. The CLM model can avoid the overfitting problems on the training data set and has stronger generalization ability.
- (3) Compared with existing prediction models, the proposed CLM prediction model has higher prediction accuracy and overcomes the shortcomings of traditional methods that cannot effectively learn the important dimension feature and key time information of shield thrust data.

III. MATERIALS

Using the Suzhou metro line No. 2 project as a testbed, this paper investigates shield machine cutting steel reinforced concrete piles. The diameter of the piles ranges from 1 m to 1.2 m, as shown in Fig.4 and 5. The subsurface strata consist of a layer of Miscellaneous Fill (Stratum ①) underlain by natural soils which consist of Clay (Stratum ②₁), Silty

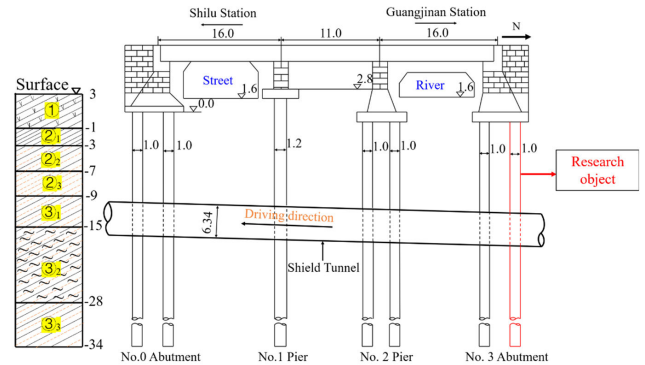


FIGURE 4. Profile view (units: m).



FIGURE 5. Plane view.

Clay (Stratum ②₂), Silt (Stratum ②₃), Silty Clay (Stratum ③₁), Sandy Silt (Stratum ③₂) and Silty Clay (Stratum ③₃). The natural soils are generally in the state of medium stiff or medium dense. The shield machine is expected to pass through the piles in Silty Clay (Stratum ③₁) and Sandy Silt (Stratum ③₂).

According to the previous studies [23], [25], the data types usually include geological parameters, shield operating parameters and tunnel parameters during the shield driving. The structure parameters should also be taken into account because the structure parameters have directly impact on the shield load when shield machine needs to cut the structure. Therefore, the data types include geological parameters, shield operating parameters, structure parameters and tunnel parameters in Suzhou metro line No. 2 project. Operational parameters can be obtained from the monitoring center of the shield machine. Geological parameters, structure parameters and tunnel parameters can be obtained from engineering investigation report. The monitoring period was 10 s. 19000 datasets in total was sampled during shield machine cutting the first pile on the right line.

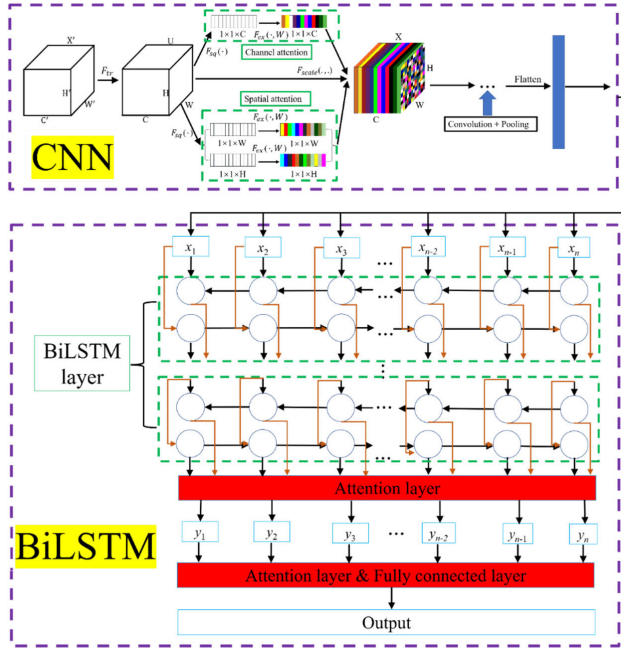


FIGURE 6. CLM model architecture.

IV. THE PROPOSED MULTIATTENTION-CNN-BILSTM PREDICTION MODEL

A. FRAMEWORK OVERVIEW

This paper proposed a hybrid model (CLM), combining CNN algorithm, BiLSTM algorithm and multi-attention mechanism to predict the thrust. Figure 6 shows that the CLM model consists of two modules, with the mixed domain attention introduced into the CNN algorithm and the attention mechanism integrated into BiLSTM algorithm.

B. MIXED DOMAIN ATTENTION-CNN COMPONENT

CNN algorithm can extract the dimension features of the data but fails to highlight important features while attention mechanism can capture important information. Therefore, CNN algorithm combined with attention mechanism can not only extract the dimension feature of the data, but also highlight important features.

As shown in the CNN module of the Fig. 6, mixed domain attention network is introduced into the CNN structure. Considering that the attention mechanism can distinguish the importance of data in different directions, it can assign weights to the channel domain and the spatial domain, then further extract data features to improve prediction accuracy. Detailed steps of the channel domain attention are as follows: first, matrix U is obtained after a convolution (F_{tr}) of the original matrix X' . then subjected to the maximum pooling operation ($F_{sq}(\cdot)$) to obtain compressed unit-length matrix ($1 \times 1 \times C$), as shown in Eqn. 9.

$$Z_c = F_{sq}(U_c) = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W u_c(i, j) \quad (9)$$

where u_c is the feature map of the c th channel, $u_c(i, j)$ refers to i th row unit of j th column of the feature map of the c th channel; H and W is the feature map height and width respectively. Z_c is the scalar that is the output of the c th element.

Second, the matrix is followed by two fully connected layers ($F_{ex}(\cdot, W)$) to obtain an updated matrix, as shown in Eqn. 10.

$$S_c = F_{ex}(z, W) = \sigma(g(z, W)) = \sigma(W_2 \sigma(W_1 z)) \quad (10)$$

where σ is the tanh activation function, $W_1 \in R^{C/r \times C}$, $W_2 \in R^{C \times C/r}$ represents weights in two fully connected layers. Z is the scalar that is obtained by the squeeze operator. $\sigma(W_1 z)$ represents a fully connected layer is activated by an activation function. S_c is the weight matrix. Similarly, S_w and S_h of the spatial attention domain also can be obtained. Then, two updated matrices in the spatial domain are fused by the tanh activation function, finally the above obtained matrix are fused with an updated matrix in the channel domain, as shown in Eqn. 11.

$$M(U) = \sigma(S_c + \sigma(S_h + S_w)) \quad (11)$$

where S_c , S_w and S_h are the weight matrices.

Then updated matrices in the mixed domain are combined with the original matrix to obtain a new matrix X , as shown in Eqn. 12. Followed by convolution and pooling layers to extract the data feature. The fully connected layer converts data from graphs into vectors, serving as input of BiLSTM algorithm.

$$X = F_{scale}(U, M(U)) = U \cdot M(U) \quad (12)$$

C. ATTENTION-BILSTM COMPONENT

The Bi-directional LSTM (BiLSTM) is a variant structure of LSTM, which can better capture the forward and backward changes in the time dimension by combining forward LSTM and backward LSTM. The output value of the BiLSTM unit is:

$$y_t = \vec{h}_t + \overleftarrow{h}_t \quad (13)$$

where \vec{h}_t is the value of the forward LSTM, \overleftarrow{h}_t is the value of the backward LSTM.

Some studies have shown that the prediction accuracy of the multi-layer LSTM structures is better than the single-layer structure [27], so the multi-layer BiLSTM structures are used in this algorithm to extract the time series characteristics of the data. Since \vec{h}_t and \overleftarrow{h}_t contains different information of previous moments in the BiLSTM network, they have different effects on the y_t . Weighting assignment in attention mechanism allows those important information presenting larger influence. Therefore, they are assigned with different weights. The y_t is expressed as:

$$y_t = \alpha \vec{h}_t + \beta \overleftarrow{h}_t \quad (14)$$

where α , β are the weights of the forward and backward, respectively.

TABLE 1. Experiment configuration.

Parameters	Notes
Operation system	windows 10
CPU	Intel(R) Core (TM) i7-10710U CPU @ 1.10GHz
GPU	GeForce MX350
Python	3.9
PyTorch	1.9.1

Considering the influence of output y_i at different times on final output, the attention layer is introduced to learn the weight of y_i at the i moment automatically. Outputs of the algorithm are finally obtained. The Y_i is expressed as:

$$Y_i = \sum_{i=1}^n \alpha_i y_i \quad (15)$$

where α_i is the weight at the i moment.

D. FLOWCHART OF THE PROPOSED MODEL

To study the prediction performance of the proposed model, this study mainly analyzes shield machine cutting the first reinforced concrete pile on the upper tunnel. The flow chart of the proposed model is depicted in Fig.7. First, different parameters are collected, and then input parameters dimensions will be selected using correlation analysis. Next, the optimal hyperparameters of the five prediction models can be obtained by training. Finally, the proposed model can be evaluated and verified through comparison with other four models and testing thrusts can be obtained.

V. EXPERIMENTS

A. DATASET AND DATA PREPROCESSING

In order to facilitate the training of neural networks, we use the Max-min method to scale the data into the range $[-1, 1]$, as shown in Eqn. 16

$$x' = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \quad (16)$$

where x is the raw data; x_{\max} and x_{\min} are the maximum and minimum of the raw data, respectively; and x' is the value after normalization.

Through trial-and-error [28], the mean square error (MSE) of the model is set as evaluation criteria, 80% (15200 sets) of the data are used as training subsets, the 15% (2850 sets) are used as validation subsets and the rest 5% (950sets) are used as testing data set, as shown in Fig. 8.

B. EXPERIMENTAL ENVIRONMENT

The hardware and software on the computer used in the experiment are listed on the Table 1.

C. MODEL ESTABLISHMENT AND HYPERPARAMETER

Shield thrust is affected by time and space [29], so the spatial-temporal matrix can be set as the input of the CNN neural network, as shown in Fig. 9. t_n is the time of shield tunneling and S_m is the shield driving distance, represented as H and W in Fig. 6, building up the spatial domain of the CNN structure. C_k is the feature dimension of shield driving, which constitutes the channel domain of the CNN structure. The spatial-temporal matrix of the shield thrust can be expressed by Eqn. 17. The spatial-temporal matrix is the input layer of the CNN structure, with the dimensional features of the data extracted by convolution layer and pooling layer and the importance of the dimensional features extracted by the mixed domain attention mechanism. After a series of processing, such as convolution, pooling, and flattening, a one-dimensional array is obtained. It is the input layer of LSTM model. The time series characteristics of the data can be extracted by BiLSTM model, and key time node information are extracted by the attention mechanism. Finally, the prediction results are obtained through the fully connected layer and the output layer. Table 2 presents the framework of the CLM model for thrust prediction:

$$A = \begin{bmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \dots & \alpha_{2n} \\ \vdots & \vdots & \dots & \vdots \\ \alpha_{m1} & \alpha_{m2} & \dots & \alpha_{mn} \end{bmatrix} \quad (17)$$

where a_{ij} is the element of the matrix in i row and j column, it represents the shield thrust value when the shield driving at the i s and the distance is j mm.

The prediction performance of the neural network is affected by many factors and hyperparameters of the neural network is an important factor. However, there is no mature theory for effectively selection of these hyperparameters [30]. The proposed algorithm is established by repeated experiments to determine the optimal hyperparameters, the parameters as shown in Table 3.

D. BASELINE MODEL

In order to verify the prediction accuracy of the proposed model, four widely used prediction models, namely, Random Forest (RF), Support Vector Machine (SVR), Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM) algorithms are introduced for comparison with the CLM algorithm. In each experiment, the model is trained on training subsets. The model performance is analyzed on the validation subsets to adjust the hyperparameters, and the predictive accuracy are finally verified on the test dataset.

E. METRICS

The prediction accuracy is measured by the error between predicted value \hat{y} and instrumented value y_i . The RMSE (root mean square error), MAE (mean absolute error) and R^2 (determination coefficient) are used to evaluate the accuracy

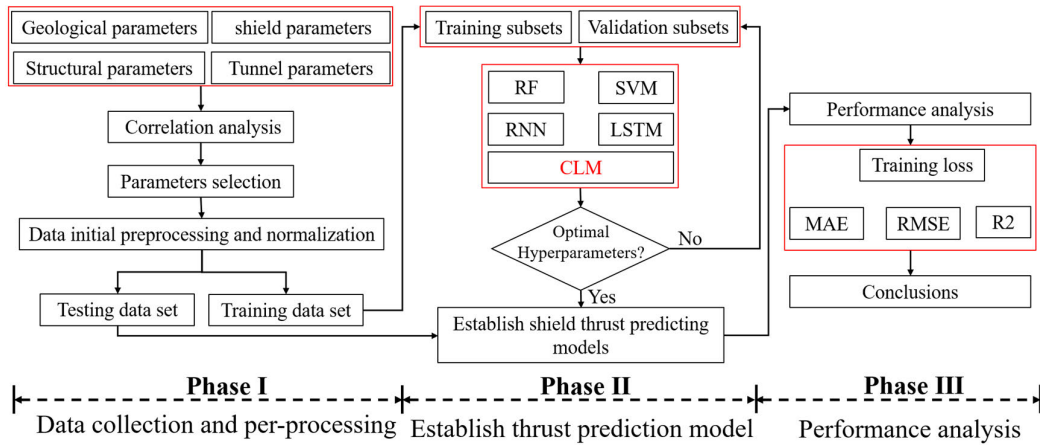


FIGURE 7. Flowchart of the proposed model.

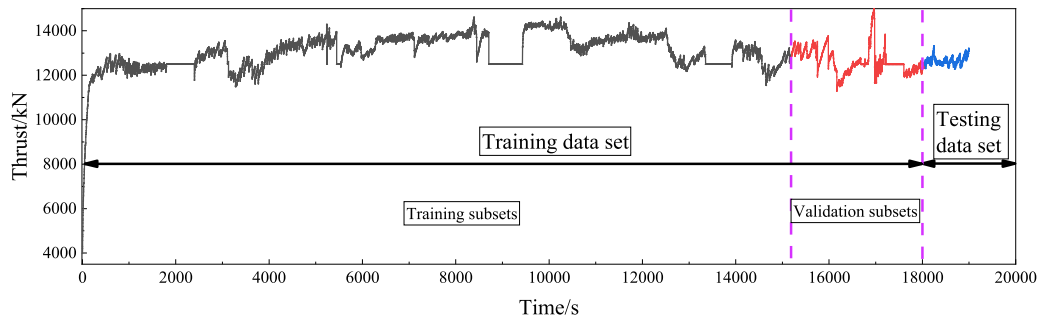


FIGURE 8. Thrust data set.

TABLE 2. Framework of the CLM model for thrust prediction.

Algorithm 1

Input: Given dataset: $X_{train}, X_{test}, y_{train}, y_{test}$. Each input sample is the data of the previous 50 moments.

Output: Predicted shield thrust. Each output sample is the corresponding predicted thrust value at the next moment.

1. Start
- 2: Set random hyperparameter (filter, kernel_size, BiLSTM layer, batch_size, iteration, dropout rate, optimizer, Loss function, Activation function)
- 3: Normalize data in between (0, 1)
- 4: First, mixed domain attention is applied to update the input matrix.
- 5: Next, CNN is applied to extract dimension features of data.
- 6: Reshape the output of CNN to one-dimensional vector.
- 7: Then, LSTM network combined with the first attention layer is applied to capture the temporal features of data.
- 8: LSTM structure is followed by the second attention layer, which is applied to assign different weight to the different moments.
9. Predicted output
- 10: Stop

TABLE 3. Optimal hyperparameters.

Model	Hyperparameters
CLM	Filter = 64, Kernel_size = 3, BiLSTM layer = 2, Number of BiLSTM = 32, Each batch = 64, Iteration = 200, Max pooling = 2, Dense = 1, Dropout rate = 0.3, Optimizer = Adam, Loss function= MSE, Activation function = ReLu

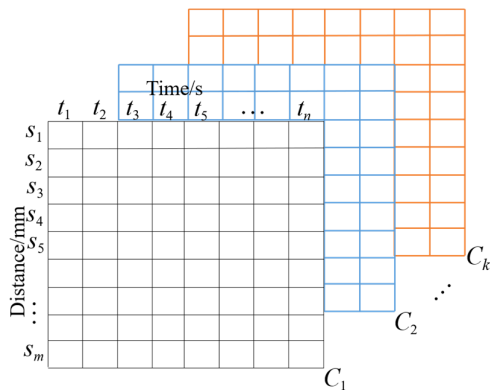


FIGURE 9. Space-time matrix of shield thrust.

of regressions [31]. The evaluation indicators are as follows:

$$\alpha_{MAE} = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (18)$$

$$\alpha_{RMSE} = \left[\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \right]^{1/2} \quad (19)$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (20)$$

VI. RESULTS AND DISCUSSIONS

A. PARAMETER SELECTION AND CORRELATION ANALYSIS

Many parameters can affect the shield thrust during the shield machine cutting through bridge piles. They are mainly divided into four kinds: geological parameters, shield operational parameters, tunnel parameters and structural parameters. Each kind of parameter also includes many sub-parameters. Therefore, the number of parameters is very large and it is unrealistic to choose all parameters as the input. This section will choose key parameters that can reflect thrust. Previous studies [32] indicated that the shield thrust is mainly affected by some operational parameters during the shield driving, such as cutter temperature, driving speed, shield attitude, penetration depth, grouting amount, grouting pressure and chamber earth pressure. In addition, some geological and tunnel parameters also can affect the shield thrust, such as buried depth ratio, friction angle and cohesion [33]. According to project characteristics, pile diameter, tensile strength of reinforcing bar and concrete strength selected as influencing factor. Based on the above analysis, 13 parameters

TABLE 4. Statistical description of 13 key parameters related to shield thrust.

Parameters	Features	Mean	Min	Max
geological parameters	friction angle (°)	18.5	12.8	24.1
	Cohesion (kPa)	23.2	6.2	40.2
	grouting pressure (MPa)	0.39	0	0.77
shield operational parameters	grouting amount (L)	0.085	0	0.17
	cutter temperature (°C)	52.8	48.3	57.2
	chamber earth pressure (MPa)	0.2	0.16	0.24
tunnel parameters	driving speed (mm/min)	12.5	1	24
	shield attitude (°)	1.24	1.21	1.27
	Penetration depth (mm/min)	6.5	0.7	12.3
structural parameters	buried depth ratio	1.7	1.9	1.4
	pile diameter (m)	1.1	1.0	1.2
structural parameters	tensile strength of reinforcing bar (MPa)	492.5	455	530
	concrete strength (MPa)	37	30	44

are selected to predict thrust, Table 4 shows the statistical analyses on 13 parameters.

Previous studies [34] have shown that the high dimension of the input parameters can cause overfitting in models and reduce computational efficiency. Therefore, this section will study the correlation between the input and output parameters using Maximal Information Coefficient (MIC) [35] to remove redundant input parameters. Figure 10 shows that the correlation coefficient (C) between cutter temperature, driving speed, shield attitude, penetration depth, grouting amount, grouting pressure, chamber earth pressure, buried depth ratio, friction angle, cohesion, pile diameter, tensile strength of reinforcing bar, concrete strength and thrust is

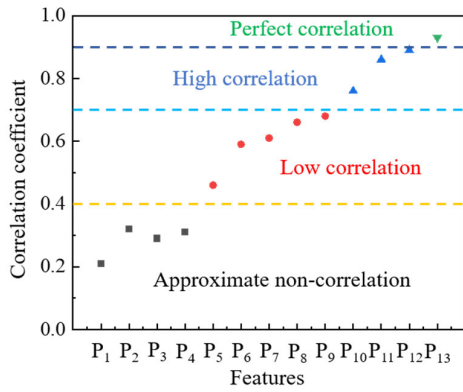


FIGURE 10. Correlation coefficient between the input features and thrust.

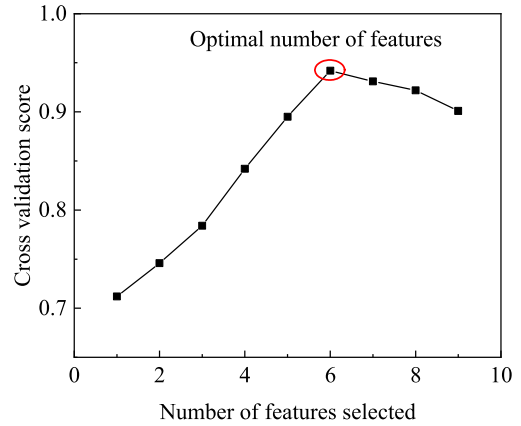


FIGURE 12. Optimal number of features selected by RF-RFE.

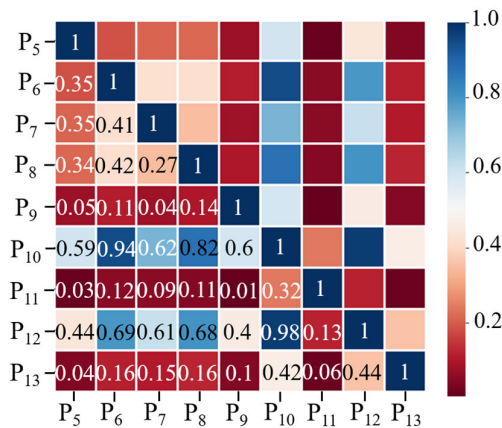


FIGURE 11. Correlation coefficient between the input features.

0.59, 0.86, 0.89, 0.93, 0.46, 0.31, 0.76, 0.21, 0.32, 0.29, 0.61, 0.68, 0.66, respectively. Previous studies [36] have shown that if $C > 0.9$, it indicates that input and output parameters are perfect correlation; if $0.7 < C < 0.9$, it indicates that input and output parameters are high correlation; if $0.4 < C < 0.7$, it indicates that input and output parameters are low correlation; if $C < 0.4$, it indicates that input and output parameters are approximate non-correlation. Therefore, grouting pressure, buried depth ratio, friction angle and cohesion are excluded, the remainder parameters were selected for thrust prediction.

The remainder parameters can be further reduced due to the redundancy between the selected input parameters that also can cause the overfitting in the machine learning. Fig.11 presents the correlation coefficient between the input features, it can be seen that the correlation coefficient (C) between cutter temperature and chamber earth pressure is 0.94, between shield attitude and chamber earth pressure is 0.98, between tensile strength of reinforcing bar and chamber earth pressure is 0.82, it indicating that there is perfect correlation between cutter temperature, shield attitude and chamber earth pressure, respectively. There is high correlation between tensile strength of reinforcing bar and chamber earth pressure. In order to ensure the independence of the

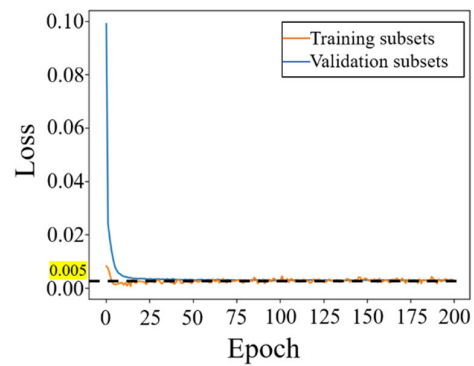


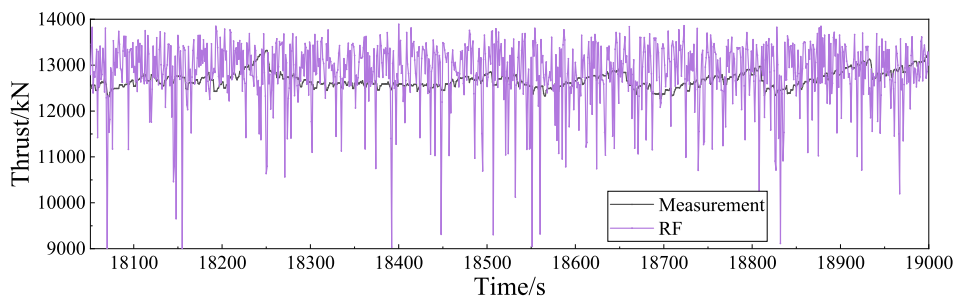
FIGURE 13. Training and validation loss during training process.

input parameters, we can abandon some parameters using the random forest-recursive feature elimination (RF-RFE) [37]. The raw data are randomly divided into a training set (80%) and a test set (20%), and then the RF-RFE was used for feature optimization. The RF-RFE selects different feature sets according to the importance of each feature, then calculates the accuracy of each feature set. Finally, the feature set with less features and high classification accuracy is regarded as the optimal feature set. As shown in Fig.12, cutter temperature, shield attitude and tensile strength of reinforcing bar can be excluded, the remainder 6 parameters were selected for thrust prediction.

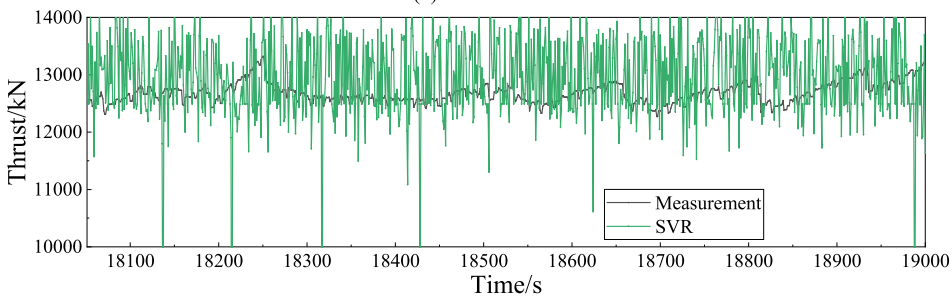
where P_1 to P_{13} represent buried depth ratio, friction angle, cohesion, grouting pressure, grouting amount, cutter temperature, pile diameter, tensile strength of reinforcing bar, concrete strength, chamber earth pressure, driving speed, shield attitude, penetration depth, respectively.

B. PERFORMANCE OF THE CLM MODEL

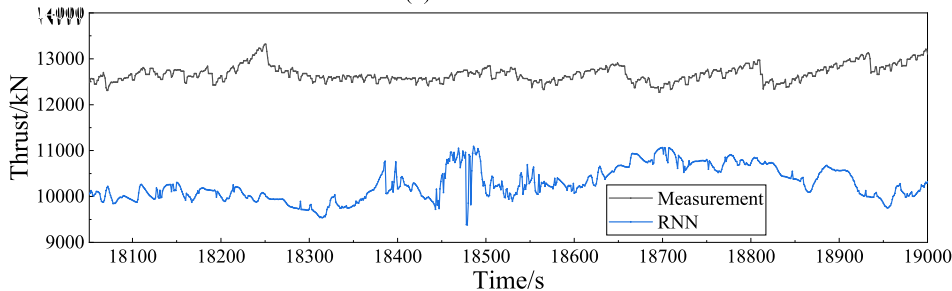
Figure 13 presents loss values of the thrust on training subsets and validation subsets respectively. The results show that with the increase of epochs, the loss values rapidly decreased and tend to be stable, with the last loss values on training subsets and validation subsets almost the same (0.005). This indicates that the CLM model has a good prediction performance on



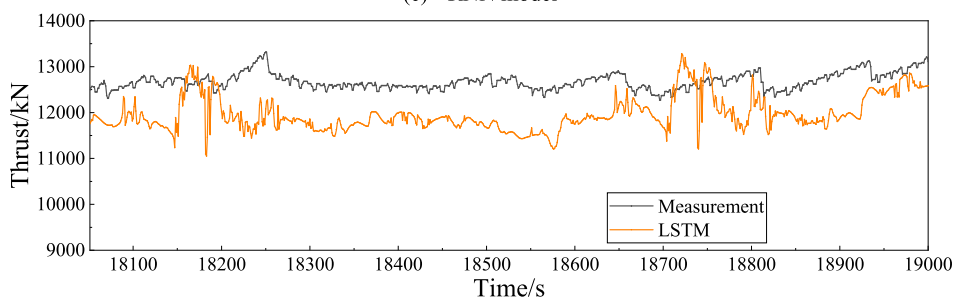
(a) RF model



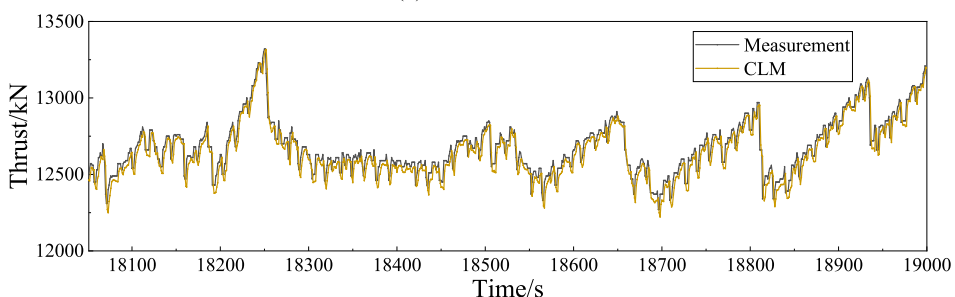
(b) SVR model



(c) RNN model



(d) LSTM model



(e) CLM model

FIGURE 14. Prediction curve of the different models.

TABLE 5. Optimal hyperparameters of different algorithms.

Algorithms	Hyperparameters
RF	Max_Features = 400, Max_Depth = 8, N_Estimators = 60, Min_Sample_Leaf = 2, Min_Sample_Split = 2, Random_State = 10
SVR	Kernel = RBF, Gamma = 0.006, C = 2
RNN	RNN layer = 4, Number of the RNN units = 32, Each batch = 64, Iteration = 100, Dense = 1, Dropout rate = 0.2, Optimizer = Adam, Loss function = MSE, Activation function = ReLu
LSTM	LSTM layer = 3, Number of the LSTM units = 32, Each batch = 64, Iteration = 150, Dense = 1, Dropout rate = 0.3, Optimizer = Adam, Loss function = MSE, Activation function = ReLu

training dataset and can avoid the overfitting problems on the training data set. In addition, there is a few fluctuations on the loss curve on the validation subsets, which indicates that the CLM model has good generalization ability. Some fluctuations were observed in the loss curves on the training subsets due to the existence of noise in the raw data. The loss curves rapidly appeared to be stable after certain number of iterations, which demonstrates the reliability of the model.

C. COMPARISON WITH EXISTING METHODS

In this section, we compare the CLM algorithm with Random Forest (RF), Support Vector Machine (SVR), Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM) algorithms using the thrust data to demonstrate effectiveness of the CLM model. In order to ensure the rationality of the results, the optimal hyperparameters of the other algorithms are selected by training, and then the prediction results are compared with the LSTM algorithm. The optimal parameters of the other algorithms are shown in the Table 5.

As shown in Table 6, the MAE, RMSE and R2 in the datasets are 0.02, 0.05 and 0.85 for the CLM algorithm, 7.53, 6.62 and 0.61 for LSTM, 9.98, 11.63 and 0.50 for RNN, 9.21, 10.68 and 0.52 for SVR, 10.21, 13.44 and 0.49 for RF, respectively. Hence, the MAE and RMSE of the CLM algorithm are lower than other algorithms, the R2 of the CLM algorithm are higher than other algorithms. The results show that the prediction accuracy of the CLM, RNN and LSTM algorithms are higher than that of the other two machine learning algorithms. The reason is that the CLM, RNN and LSTM algorithms can effectively extract the time series features from the data, and can effectively learn the changing law of the time-varying bridge structure sequence. Prediction accuracy of the CLM, LSTM algorithm is higher than that of the RNN algorithm, this is because the RNN algorithm cannot learn the long-term dependencies for time series data due to the gradient disappearance and gradient explosion. Prediction accuracy of the CLM algorithm is higher than that of the LSTM algorithm, this is because the LSTM algorithm cannot learn the important time node information for time series data.

In addition, the prediction accuracy of SVR is higher than RF, the reason is that SVR is suitable for processing continuous data, while RF are suitable for processing discrete data. The same conclusions also can be drawn from Fig.14, it presents the prediction and measurement curves of the five prediction models in the testing data set. Compared with the RF, SVR, RNN and LSTM, the coincidence degree of the proposed CLM model is higher than other models.

As shown in Table 7, the time of the training subsets, validation subsets and testing dataset are 365.62, 276.31 and 89.64s for the CLM algorithm; 225.54, 106.43 and 33.37s for LSTM; 107.64, 49.76 and 28.21s for RNN; 88.76, 40.25 and 17.12s for SVR; 67.32, 25.33 and 9.43s for RF, respectively. We all known that deep learning algorithms require more time for training than non-deep learning algorithms. Hence, it can be seen that the time of the CLM, LSTM and RNN algorithms are longer than RF and SVR algorithms. The running time of the CLM algorithm are longer than LSTM and RNN algorithms, the reason is that the proposed algorithm needs the addition of CNN structure to extract the data characteristic. Despite the longest running time, the CLM algorithm gives the best prediction results. In fact, the testing time is more important than training time, because the model can be trained in advance and the testing time directly reflect how long the prediction results can be applied to the actual prediction. The testing time of the CLM algorithm is 89.64s, which is acceptable. Moreover, the running time of the proposed model will be further reduced with the development of the computer hardware performance in the future.

D. ABLATION EXPERIMENT

The CLM model is mainly composed of two components: CNN algorithm fused with attention mechanism and BLSTM algorithm fused with attention mechanism. In order to further analyze the validation of each module, ablation experiments are carried out. The algorithms are CNN, BiLSTM, CNN + attention, BiLSTM + attention and CLM.

The testing results are shown in Table 8, the MAE and RMSE of the CLM algorithm are lower than other algorithms,

TABLE 6. Model evaluation of thrust.

Algorithm	MAE	RMSE	R ²
RF	10.21	13.44	0.49
SVR	9.21	10.68	0.52
RNN	9.98	11.63	0.50
LSTM	7.53	6.62	0.61
CLM	0.02	0.05	0.85

TABLE 7. Training and testing time for each model in thrust prediction.

Models	Time cost		
	Training subsets/s	Validation subsets/s	Testing data set/s
RF	67.32	25.33	9.43
SVR	88.76	40.25	17.12
RNN	107.64	49.76	28.21
LSTM	225.54	106.43	33.37
CLM	365.62	276.31	89.64

TABLE 8. Model evaluation of thrust.

Algorithm	MAE	RMSE	R ²
CNN	12.12	22.57	0.44
BiLSTM	5.65	10.68	0.63
CNN- Attention	3.05	8.11	0.67
BiLSTM-Attention	0.03	0.12	0.77
CLM	0.02	0.05	0.85

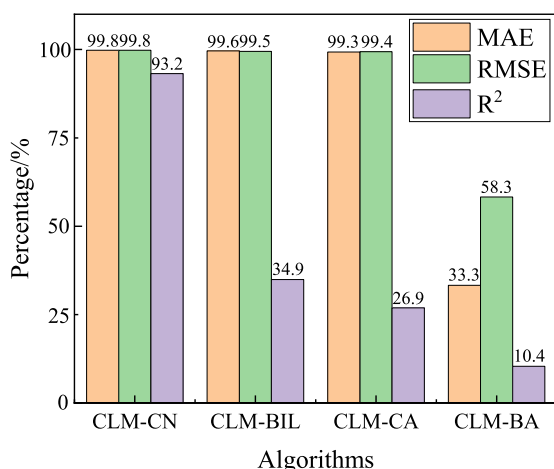
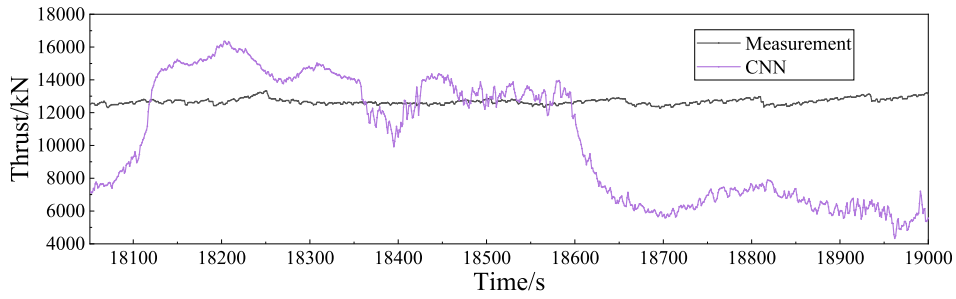


FIGURE 15. Comparison of the different models.

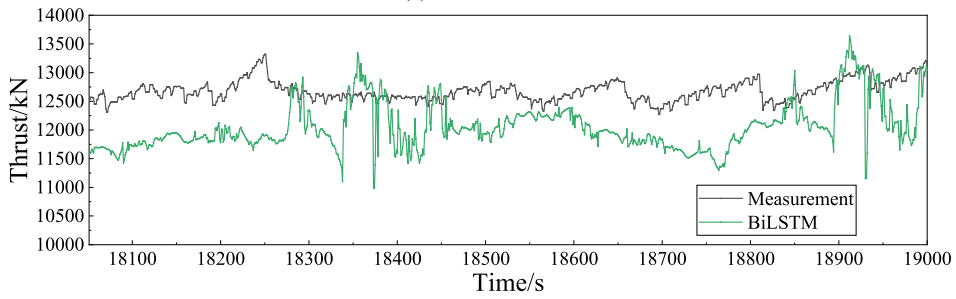
the R² of the CLM algorithm is higher than other algorithms. The results show that the prediction accuracy of the CLM algorithm is higher than that of the other four machine learning algorithms. In order to intuitively analyze the prediction accuracy of the proposed model, a comparison chart is drawn,

as shown in Figure 15. MAE and RMSE through CLM model are smaller than that through CNN-Attention model by 99.3% and 99.4%, respectively. R² through CLM model is larger than that through CNN-Attention model by 26.9%. The reason is that the BiLSTM model can recognize the long-short term dependency for time series data and attention mechanism in BiLSTM model can catch the key time node information. MAE and RMSE through CLM model are smaller than that through BiLSTM-Attention model by 33.3% and 58.3%, respectively. R² through CLM model is larger than that through BiLSTM-Attention model by 10.4%. The reason is that the CNN model can deal with multi-dimensional problem with better accuracy and efficiency, and the attention mechanism in CNN model can catch important features. MAE and RMSE of the BiLSTM-Attention model are smaller than BiLSTM model. R² of the BiLSTM-Attention model is larger than BiLSTM model. The same conclusion can be drawn from CNN-Attention and CNN models. It indicates that attention mechanism is beneficial to improve the prediction accuracy.

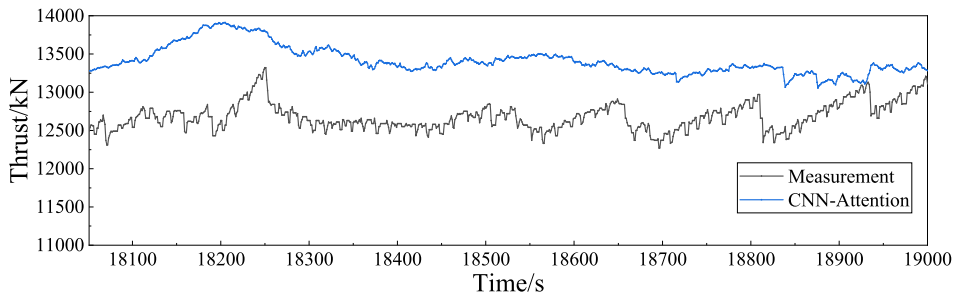
The prediction effect of the model is intuitively evaluated by comparing prediction data and monitoring data, as shown



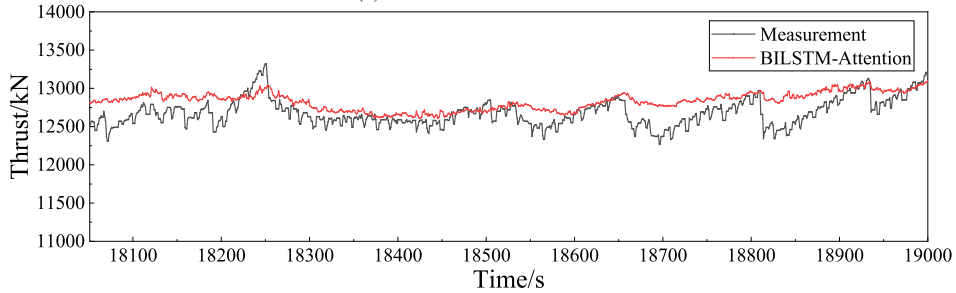
(a) CNN model



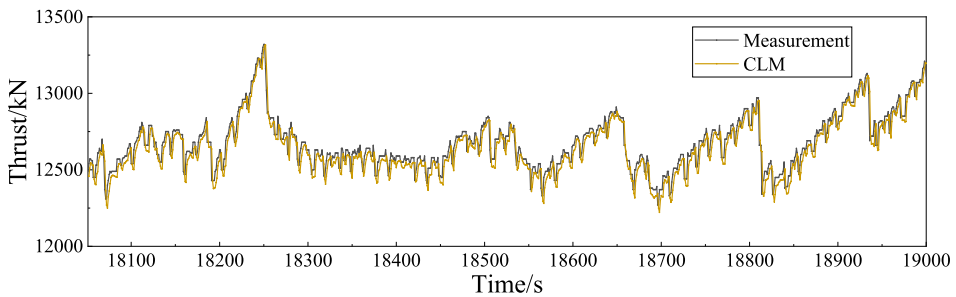
(b) BiLSTM model



(c) CNN-Attention model

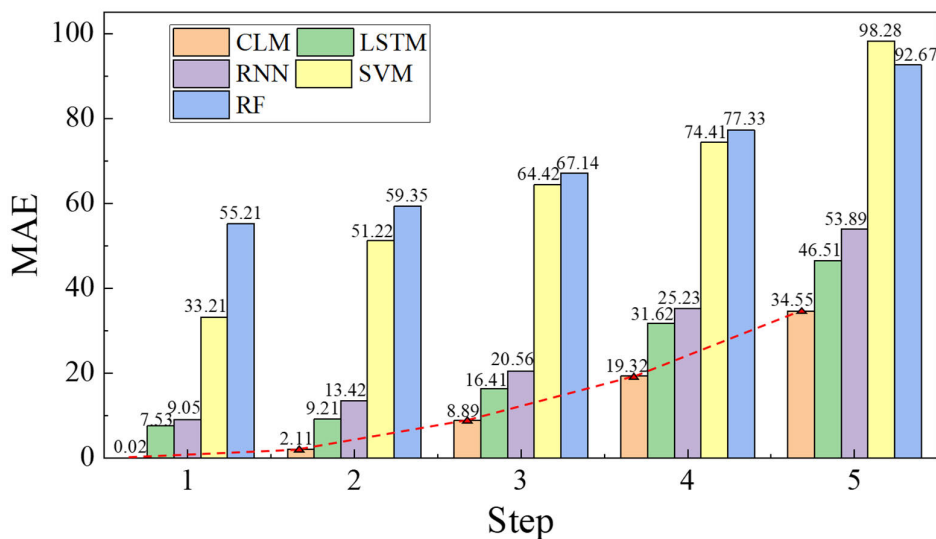


(d) BiLSTM-Attention model

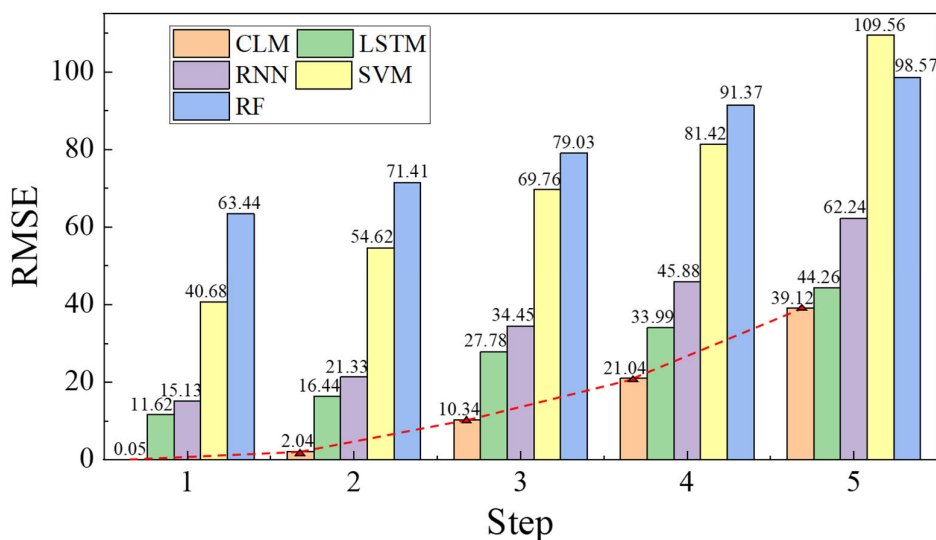


(e) CNN-Attention model

FIGURE 16. Thrust prediction results of the different models.



(a) MAE of five steps prediction



(b) RMSE of five steps prediction

FIGURE 17. Evaluation on multi-step prediction.

in Fig. 16. It can be found that the predicted value of proposed model is more consistent with the monitoring data, while the prediction values of CNN model is quite different from the monitoring data. The prediction performance of the BiLSTM model is better than CNN model, which indicates that BiLSTM model can better deal with a large volume of multi-dimensional and multivariate time series data.

Where CN: CNN; BIL: BiLSTM; CA: CNN-Attention; BA: BiLSTM-Attention; CLM: CNN-BiLSTM-Multiattention.

E. MULTI-STEP PREDICTION

Currently, the multi-step prediction of time series is more valuable than the single-step prediction in some engineering fields. Engineering construction is urgent and

single-step prediction cannot meet warning requirements. Therefore, multi-step prediction is more valuable for better construction measures. In this section, the multi-step prediction performance of the proposed model is evaluated. The thrust of next five time-steps were predicted and the CLM model compared with RF, SVR, RNN and LSTM models. MAE and RMSE are used as evaluation index.

Figure 17 shows that the MAE and RMSE of all models are larger with bigger time steps, which indicates that the prediction performance of the models is getting worse and worse with bigger time steps. The reason is that the prediction results depend on the prediction of the last step and the prediction errors will be accumulated. However, compared with the SVR, RF, RNN and LSTM, the MAE and RMSE of CLM model are minimal. The results show that the prediction

performance of CLM model is the best among all models. Moreover, it can be observed that the curve of the MAE and RMSE in CLM model begin to increase after the two time-steps, it indicates that the prediction performance of the CLM model rapidly drops after two time-steps. Therefore, how to improve the prediction performance of the CLM model after two time-steps is what to study in following study.

VII. CONCLUSION

In this paper, we propose a novel hybrid deep neural network for precise thrust prediction of shield tunneling machine. Correlation analysis based on Maximal Information Coefficient (MIC) between parameters and the thrust are conducted for parameter selection and input dimension reduction, then the dimensions features of shield thrust are extracted by CNN layer and important features are captured by mix domain attention mechanism. The time-varying characteristics are extracted by BiLSTM layer, and the attention mechanism can capture the important time node information. In order to validate the effectiveness, generalization and superiority of the proposed method, some experiments have been conducted based on real project data, and comparisons are made with existing machine learning models.

The results show that parameter selection is important in thrust prediction with construction data since each feature contributes differently and the dimensions of the input parameters can cause overfitting in models and reduce computational efficiency. The proposed model has a good performance, which can avoid overfitting problems and has good generalization ability. Compared with four existing prediction models, the proposed method shows higher prediction performance in terms of determination coefficient (R^2), root mean square error (RMSE) and mean absolute error (MAE). Moreover, the validation of each component of the proposed model were analyzed using ablation experiment. The results demonstrate that the proposed CLM model integrates the advantages of CNN, BiLSTM and attention mechanism, it indicates that the proposed CLM model can accurately predict shield thrust during shield driving.

It can be found that although the multi-step prediction performance of the proposed CLM model is better than existing four models, prediction performance of the CLM model rapidly drops after two time-steps. Therefore, in the future, efforts will be made to on improving the multi-step prediction performance of the CLM model. Meanwhile, the proposed model also can be applied in other fields.

DECLARATION OF COMPETING INTEREST

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

REFERENCES

- [1] X. Zhang, C. P. Zhang, and J. C. Wang, "Effect of closely spaced twin tunnel construction beneath an existing subway station: A case study," *J. Test. Eval.*, vol. 46, no. 4, pp. 1559–1573, 2018.
- [2] M. L. Yin, H. Jiang, Y. S. Jiang, Z. Y. Sun, and Q. L. Wu, "Effect of the excavation clearance of an under-crossing shield tunnel on existing shield tunnels," *Tunnelling Underground Space Technol.*, vol. 78, pp. 245–258, Aug. 2018.
- [3] M. Nematollahi and D. Dias, "Interaction between an underground parking and twin tunnels—Case of the Shiraz subway line," *Tunnelling Underground Space Technol.*, vol. 95, Jan. 2020, Art. no. 103150.
- [4] H. Xiao, Z. Chen, R. Cao, Y. Cao, L. Zhao, and Y. Zhao, "Prediction of shield machine posture using the GRU algorithm with adaptive boosting: A case study of Chengdu subway project," *Transp. Geotechnics*, vol. 37, Nov. 2022, Art. no. 100837.
- [5] J. Su, Y. Wang, X. Niu, S. Sha, and J. Yu, "Prediction of ground surface settlement by shield tunneling using XGBoost and Bayesian optimization," *Eng. Appl. Artif. Intell.*, vol. 114, Sep. 2022, Art. no. 105020.
- [6] J. Rostami, "Development of a force estimation model for rock fragmentation with disc cutters through theoretical modeling and physical measurement of crushed zone pressure," M.S. thesis, Colorado School Mines, Golden, CO, USA, May 1997.
- [7] L. T. Wang, G. Gong, H. Shi, and H. Yang, "Modeling and analysis of thrust force for EPB shield tunneling machine," *Autom. Construct.*, vol. 27, pp. 138–146, Nov. 2012.
- [8] Q. Zhang, T. Huang, G. Huang, Z. Cai, and Y. Kang, "Theoretical model for loads prediction on shield tunneling machine with consideration of soil-rock interbedded ground," *Sci. China Technol. Sci.*, vol. 56, no. 9, pp. 2259–2267, 2013.
- [9] X.-P. Zhou and S.-F. Zhai, "Estimation of the cutterhead torque for Earth pressure balance TBM under mixed-face conditions," *Tunnelling Underground Space Technol.*, vol. 74, pp. 217–229, Apr. 2018.
- [10] S. Yagiz, "New equations for predicting the field penetration index of tunnel boring machines in fractured rock mass," *Arabian J. Geosci.*, vol. 10, no. 2, p. 33, Jan. 2017.
- [11] Q. Zhang, C. Qu, Z. Cai, Y. Kang, and T. Huang, "Modeling of the thrust and torque acting on shield machines during tunneling," *Automat. Construct.*, vol. 40, pp. 60–67, Apr. 2014.
- [12] M. D. Han, Z. X. Cai, C. Y. Qu, and L. S. Jin, "Dynamic numerical simulation of cutterhead loads in TBM tunnelling," *Tunnelling Underground Space Technol.*, vol. 70, pp. 286–298, Nov. 2017.
- [13] L. Wu, T. Guan, and L. Lei, "Discrete element model for performance analysis of cutterhead excavation system of EPB machine," *Tunnelling Underground Space Technol.*, vol. 37, pp. 37–44, Aug. 2013.
- [14] L. Faramarzi, A. Kheradmandian, and A. Azhari, "Evaluation and optimization of the effective parameters on the shield TBM performance: Torque and thrust-using discrete element method (DEM)," *Geotech. Geol. Eng.*, vol. 38, no. 3, pp. 2745–2759, 2020.
- [15] C. Zhou, H. Xu, L. Ding, L. Wei, and Y. Zhou, "Dynamic prediction for attitude and position in shield tunneling: A deep learning method," *Autom. Construct.*, vol. 105, Sep. 2019, Art. no. 102840.
- [16] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997.
- [17] L. Yu, J. Qu, F. Gao, and Y. Tian, "A novel hierarchical algorithm for bearing fault diagnosis based on stacked LSTM," *Shock Vib.*, vol. 12, pp. 1–10, Jan. 2019.
- [18] J. Zhang, P. Wang, R. Yan, and R. X. Gao, "Long short-term memory for machine remaining life prediction," *J. Manuf. Syst.*, vol. 48, pp. 78–86, Jul. 2018.
- [19] M. Corbetta and G. L. Shulman, "Control of goal-directed and stimulus-driven attention in the brain," *Nature Rev. Neurosci.*, vol. 3, no. 3, pp. 201–215, 2002.
- [20] X. Gao, M. Shi, X. Song, C. Zhang, and H. Zhang, "Recurrent neural networks for real-time prediction of TBM operating parameters," *Automat. Construct.*, vol. 98, pp. 225–235, Feb. 2019.
- [21] P. Zhang, H.-N. Wu, R.-P. Chen, T. Dai, F.-Y. Meng, and H.-B. Wang, "A critical evaluation of machine learning and deep learning in shield-ground interaction prediction," *Tunnelling Underground Space Technol.*, vol. 106, Dec. 2020, Art. no. 103593.
- [22] H. Chen, C. Xiao, Z. Yao, H. Jiang, T. Zhang, and Y. Guan, "Prediction of TBM tunneling parameters through an LSTM neural network," in *Proc. IEEE Int. Conf. Robot. Biomimetics (ROBIO)*, Dec. 2019, pp. 702–707.
- [23] C. Qin, G. Shi, J. Tao, H. Yu, and C. Liu, "Precise cutterhead torque prediction for shield tunneling machines using a novel hybrid deep neural network," *Mech. Syst. Signal Processing.*, vol. 151, no. 6, 2021, Art. no. 107386.

- [24] C. Zhou, H. Xu, L. Ding, L. Wei, and Y. Zhou, "Dynamic prediction for attitude and position in shield tunneling: A deep learning method," *Autom. Construct.*, vol. 105, Sep. 2019, Art. no. 102840.
- [25] G. Shi, C. Qin, J. Tao, and C. Liu, "A VMD-EWT-LSTM-based multi-step prediction approach for shield tunneling machine cutterhead torque," *Knowl.-Based Syst.*, vol. 228, Sep. 2021, Art. no. 107213.
- [26] C. Xu, X. Liu, E. Wang, and S. Wang, "Prediction of tunnel boring machine operating parameters using various machine learning algorithms," *Tunnelling Underground Space Technol.*, vol. 109, Mar. 2021, Art. no. 103699.
- [27] A. J. Lu, "From single layer networks to multilayer networks—structure, dynamics and function," *Modern Phys.*, vol. 27, no. 4, pp. 3–8, 2015.
- [28] D.-M. Zhang, J.-Z. Zhang, H.-W. Huang, C.-C. Qi, and C.-Y. Chang, "Machine learning-based prediction of soil compression modulus with application of 1D settlement," *J. Zhejiang Univ. Sci. A*, vol. 21, no. 6, pp. 430–444, Jun. 2020.
- [29] G. M. Yu, J. Lei, P. H. Zhang, Z. Chen, and G. W. Liang, "Exploration and prospect of 'five-dimensional' space effect of ground surface settlement caused by shield tunnel construction," *J. Qingdao Univ. Technol.*, vol. 43, no. 2, pp. 1–10, 2022.
- [30] H. Shao, H. Jiang, F. Wang, and H. Zhao, "An enhancement deep feature fusion method for rotating machinery fault diagnosis," *Knowl.-Based Syst.*, vol. 119, pp. 200–220, Mar. 2017.
- [31] W. Zhang, R. Zhang, C. Wu, A. T. C. Goh, S. Lacasse, Z. Liu, and H. Liu, "State-of-the-art review of soft computing applications in underground excavations," *Geosci. Frontiers*, vol. 11, no. 4, pp. 1095–1106, Jul. 2020.
- [32] Q. Zhai, "Study on safety management of TBM construction of Qingdao metro line1," Lanzhou Jiaotong Univ., Lanzhou, China, Tech. Rep., 2021.
- [33] Z. L. Li, "Dynamic prediction of surface deformation induced by shield tunneling based on maximal information coefficient-long short-term memory," *Tunnel Construct.*, vol. 41, no. 2, pp. 199–205, 2021.
- [34] L. Yu and H. Liu, "Efficient feature selection via analysis of relevance and redundancy," *J. Mach. Learn. Res.*, vol. 5, no. 10, pp. 1205–1224, 2004.
- [35] D. N. Reshef, Y. A. Reshef, H. K. Finucane, S. R. Grossman, G. McVean, P. J. Turnbaugh, E. S. Lander, M. Mitzenmacher, and P. C. Sabeti, "Detecting novel associations in large data sets," *Science*, vol. 334, no. 6062, pp. 1518–1524, 2011.
- [36] T. K. Koo and M. Y. Li, "A guideline of selecting and reporting intra-class correlation coefficients for reliability research," *J. Chiropractic Med.*, vol. 15, no. 2, pp. 155–163, 2016.
- [37] Z. Liu, L. Li, X. Fang, W. Qi, J. Shen, H. Zhou, and Y. Zhang, "Hard-rock tunnel lithology prediction with TBM construction big data using a global-attention-mechanism-based LSTM network," *Autom. Construct.*, vol. 125, May 2021, Art. no. 103647.

CHENG CHEN received the M.Sc. degree in civil engineering from the Beijing University of Civil Engineering and Architecture, China, in 2019. He is currently pursuing the Ph.D. degree with the School of Rail Transportation, Soochow University, China. His research interests include underground engineering and machine learning.

BEN WU received the Ph.D. degree in civil engineering from Soochow University, China, in 2021. He is currently a Lecturer with the School of Rail Transportation, Soochow University. His research interests include underground engineering and machine learning.

PENGJIAO JIA received the M.Sc. and Ph.D. degrees in geotechnical engineering from Northeastern University, China, in 2020. He is currently an Associate Professor with the School of Rail Transportation, Soochow University, China. His research interests include underground engineering and machine learning.

ZHANSHENG WANG received the M.Sc. degree in geotechnical engineering from Beijing Jiaotong University, China. He is currently a Professorate Senior Engineer with Suzhou Rail Transit Group Company Ltd., China. His research interests include geotechnical engineering and shield tunnel.

• • •