Computer Modeling in
Engineering & Sciences

Tech Science Press

https://doi.org/10.32604/cmes.2026.082647

ARTICLE

A Hybrid SSA-CNN-LSTM-Transformer Model for Predicting Settlement of Buildings Adjacent to Shield Tunnels

Jiawang Zou, Annan Jiang*, Xinzhi Wang and Hao Huang

College of Transportation Engineering, Dalian Maritime University, Dalian, China

*Corresponding Author: Annan Jiang. Email: jiangannan@163.com

Received: 19 March 2026; Accepted: 05 June 2026; Published: 27 July 2026

ABSTRACT: Accurate forecasting of settlement in buildings adjacent to shield tunnels remains a critical challenge in underground engineering due to complex spatiotemporal interactions and nonlinear relationships among multi-source monitoring data and construction parameters. To address this issue, a Sparrow Search Algorithm (SSA)-optimized Convolutional Neural Network–Long Short-Term Memory–Transformer (CNN-LSTM-Transformer) hybrid framework is proposed, explicitly incorporating the relative spatial relationship between the shield excavation face and adjacent structures. In this framework, the Convolutional Neural Network (CNN) module extracts spatial features from monitoring data and tunneling parameters, capturing interdependencies among different construction indicators and reflecting local spatial heterogeneity of building deformation, while the Long Short-Term Memory (LSTM) network is employed to model temporal dependencies in multi-factor time-series data, enabling the network to learn long-term settlement evolution patterns. Notably, a Transformer-based self-attention mechanism is incorporated to improve the model's ability to capture global dependencies, highlight critical construction stages—such as the shield undercrossing period—and identify key influencing features in the multivariate data. Key hyperparameters are optimized using the Sparrow Search Algorithm (SSA), improving prediction performance, model robustness, and stability while identifying an optimal configuration. The framework is validated through a case study using real monitoring data from the east extension of Section 02 of the Nanchang Metro. Results demonstrate that the SSA-CNN-LSTM-Transformer model achieves significantly higher precision, reducing the maximum prediction error by over 80% compared to traditional LSTM and empirical Peck methods. This approach provides a reliable computational tool for proactive risk assessment, safety management, and informed decision-making in complex urban underground engineering applications.

KEYWORDS: Shield tunnelling; adjacent building settlement; convolutional neural network; long short-term memory network; transformer-based attention; sparrow search algorithm

1 Introduction

With the rapid advancement of urban modernization, subway tunnel projects have become increasingly prevalent in urban construction, reflecting continuous improvements in urban transportation systems as well as growing demands for transport efficiency, environmental sustainability, and urban development [1–3]. Subway tunnels are typically constructed beneath densely built urban areas, where shield tunnelling is one of the most commonly adopted construction methods. However, tunnel excavation inevitably disturbs surrounding soil and rock strata, which may induce settlement of adjacent buildings and underground structures, potentially leading to structural damage and safety risks [4–6].

As an important component of the external environment affected by tunnel construction, the settlement and deformation of buildings adjacent to shield tunnels have a direct impact on construction safety and risk control. To systematically analyze these complex structural and geotechnical responses, Soga et al. [7] investigated the long-term tunnel behavior and ground movements during tunnelling in clayey soils. Despite such fundamental insights, it remains difficult for traditional statistical and empirical prediction methods to capture the complex nonlinear relationships among multiple influencing factors and their temporal dependencies [8–11].

To address these limitations, neural network-based deep learning models have been increasingly adopted to analyze and process real-time monitoring data [12–16]. For instance, Elmousalami [16] highlighted the significant potential of artificial intelligence and parametric estimation in construction engineering. Deep learning models such as CNNs can extract spatial features from multivariate monitoring data, while LSTM and Bi-LSTM networks capture temporal dependencies [17–19]. Hybrid models combining CNN and LSTM, often augmented with attention mechanisms, have demonstrated improved prediction performance in settlement monitoring, landslide displacement prediction, dam deformation, and railway track geometry analysis [20–23].

Recent advances include small-sample prediction using graph convolutional networks [20], CNN-BiLSTM with self-attention mechanisms [21], and surrogate modeling frameworks for spatiotemporal prediction [19]. In these studies, hybrid architectures and attention-based mechanisms have demonstrated strong capability in capturing nonlinear spatiotemporal dependencies in underground engineering applications [22–24]. In addition, recent studies have investigated tunnel construction risks, excavation-induced geological hazards, underground construction management, and artificial intelligence-assisted decision-making in tunnelling engineering [25–29]. Deep learning is a subfield of machine learning that utilizes deep neural networks to model complex nonlinear relationships [30].

However, prediction models based on a single algorithm often suffer from limitations in robustness and stability. To address these issues, hybrid and attention-augmented models have been proposed. Yu et al. [31] developed an Attention–ResNet–LSTM model for tunnel boring machine advance rate prediction, while Yu et al. [12] applied an LSTM-based model for settlement prediction. CNN-LSTM-based frameworks have also been widely adopted in geotechnical engineering, demonstrating improved prediction accuracy in excavation and deformation monitoring problems [13–15,17,21]. Similar hybrid or improved models have been successfully applied in landslide displacement prediction, dam deformation monitoring, railway track geometry analysis, and TBM performance evaluation [14,15,17,21,32].

Despite these advances, several challenges remain in the prediction of building settlement induced by shield tunnelling. Existing studies have mainly focused on improving prediction accuracy through model architecture design, while the integration of construction-related physical information into the modeling process is often insufficient. In particular, the spatial relationship between shield tunnelling progression and building deformation is rarely explicitly incorporated into the input representation, which limits the interpretability and adaptability of models in practical engineering scenarios. Furthermore, under complex construction conditions, conventional models may have difficulty effectively identifying critical influencing variables and key time steps in multivariate time-series data. In addition, the selection of model hyperparameters is commonly based on empirical experience, reducing model stability and limiting generalization performance.

While deep learning has been extensively applied to shield tunnelling settlement prediction, existing studies predominantly focus on greenfield conditions. In dense urban environments, the settlement behavior of buildings adjacent to shield tunnels is significantly influenced by factors such as structural stiffness, foundation type, and the dynamic interaction between tunnelling processes and surrounding strata. Traditional empirical models, such as the Peck formula, may have limitations in capturing these interactions under complex urban conditions [33]. Therefore, further research is needed to improve the applicability and adaptability of intelligent prediction models in practical engineering environments.

To address these issues, in this study, an SSA-CNN-LSTM-Transformer hybrid prediction framework is developed for settlement prediction of buildings adjacent to shield tunnels. By integrating convolutional neural networks, long short-term memory networks, and a Transformer-based attention mechanism, the proposed model improves the capability of capturing nonlinear spatiotemporal deformation characteristics during shield tunnelling. In the proposed framework, the relative spatial relationship between the shield excavation face and the building location, together with multi-point settlement and inclination monitoring data, is incorporated into the model input variables to characterize the temporal evolution of deformation under tunnelling disturbance. The CNN module extracts spatial correlations among multiple monitoring parameters, while the LSTM module captures long-term temporal dependencies in settlement sequences. Furthermore, the Transformer-based attention mechanism improves the model's ability to identify critical influencing features and key time steps in multivariate monitoring data.

To improve model adaptability and reduce dependence on empirical parameter selection, the Sparrow Search Algorithm (SSA) is introduced to optimize key hyperparameters, including convolution kernel size, number of filters, learning rate, and number of hidden neurons. Based on monitoring data from the east extension project of Nanchang Metro Line 2, comparative experiments and ablation analyses are conducted to verify the effectiveness and engineering applicability of the proposed framework.

2 Deformation Prediction Model Influenced by Multiple Factors

2.1 Long Short-Term Memory Neural Network

Long short-term memory (LSTM) networks are an advanced architecture derived from recurrent neural networks (RNNs). While traditional RNNs, as illustrated in Fig. 1, can model sequential data, they are prone to gradient vanishing or explosion issues when dealing with long-range dependencies, often caused by sensitive weight initialization (e.g., overly large or small weights). These limitations significantly reduce prediction accuracy.

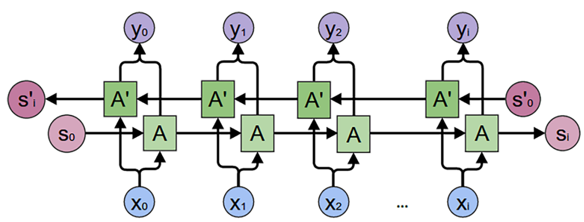

Figure 1: RNN network.

To address these issues, LSTM introduces a gated memory structure consisting of forget gates, input gates, and output gates, as shown in Fig. 2. Through selective information retention and updating, LSTM networks can effectively preserve long-term temporal dependencies and improve training stability.

The number of hidden neurons is an important hyperparameter affecting the learning capability and prediction performance of the LSTM network [34]. An insufficient number of neurons may limit the ability of the model to capture complex temporal dependencies, whereas excessive neurons may increase computational cost and overfitting risk. Therefore, the number of hidden neurons should be reasonably selected according to the complexity of the engineering dataset and prediction task.

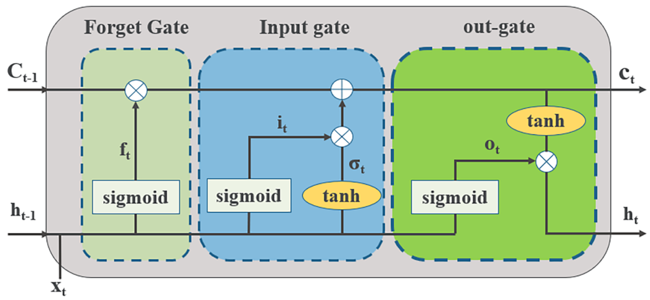

Figure 2: LSTM network.

Generally, the number of neurons is estimated by the following formula:

$$A = C \cdot (m + n), \tag{1}$$

where C is a constant, usually depending on the specific task and network structure; m is the dimension of the input feature; n is the number of categories or the dimension of the target value of the output.

In practical applications, the number of neurons can also be adjusted based on experience and experiment, and different numbers can be tried, and the best configuration can be selected through methods such as cross-validation.

In the proposed SSA-CNN-LSTM-Transformer framework, the LSTM module serves as the core temporal modeling component. It captures long-term dependencies in building deformation sequences, allowing the model to learn cumulative settlement evolution patterns induced by continuous shield tunneling. By preserving historical information over extended time horizons, the LSTM module provides a stable temporal foundation for the subsequent attention-based feature refinement.

2.2 Convolutional Neural Networks

Convolutional neural networks (CNNs) represent a deep learning architecture comparable to multilayer perceptrons, with notable advantages in modeling and analyzing data arranged in grid-like patterns. The network operates by extracting and abstracting features from the input through convolutional layers, activation functions, and pooling layers. Convolutional layers produce feature maps using a set of learnable filters (kernels) that scan the input data to capture local patterns. Pooling layers downsample the feature maps, retaining essential information while improving the model's translation invariance. The resulting features are then passed to fully connected layers for classification or regression, as illustrated in Fig. 3.

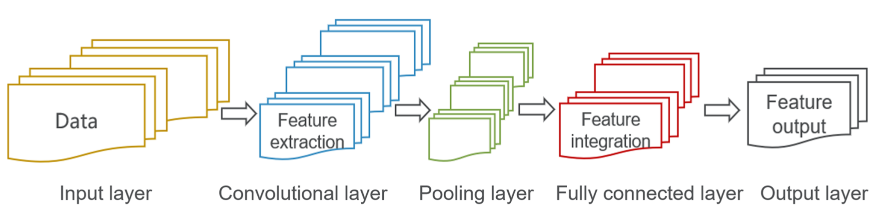

Figure 3: CNN network.

The convolutional layer is a fundamental component of the CNN architecture, utilizing a set of learnable kernels to extract spatial features from the input data. By sliding these kernels across the input, the network effectively encodes local patterns and hierarchical representations. The selection of kernel size is crucial, as it directly impacts the model's ability to represent features at various scales. The output dimension of the feature map generated by a convolutional layer is calculated as follows:

$$w_{out} = \left\lfloor \frac{w_{in} + 2p - k}{s} \right\rfloor + 1, \tag{2}$$

where w_{out} denotes the width of the output feature map, w_{in} represents the width of the input, k is the kernel size, s is the stride, and p is the zero-padding size. The $\lfloor \cdot \rfloor$ symbol denotes the floor function.

Within the proposed hybrid architecture, the CNN module is employed as a spatial feature extractor. By applying convolution operations to multi-point settlement and inclination monitoring data, CNN effectively captures local spatial correlations among different monitoring locations, as well as short-term temporal variations. This enables the model to represent the spatial heterogeneity of building deformation induced by shield tunneling before temporal dependency modeling is performed by the LSTM.

2.3 Sparrow Search Algorithm

The Sparrow Search Algorithm (SSA) is adopted in this study to optimize the hyperparameters of the proposed CNN–LSTM–Transformer model. As a population-based metaheuristic algorithm, SSA simulates the foraging and anti-predation behaviors of sparrows to iteratively search for optimal solutions within the parameter space. Each individual in the population represents a candidate set of hyperparameters, and the quality of the solution is evaluated using a predefined fitness function. Through the collaborative interaction among explorers, followers, and vigilant individuals, SSA is able to effectively balance global exploration and local exploitation, thereby reducing the risk of premature convergence to local optima [35].

With each iteration, the explorer's position is revised in accordance with the following equation:

$$x_{i,j}^{t+1} = \begin{cases} x_{i,j}^{t} \cdot \exp\left(\dfrac{-i}{\alpha \cdot iter_{max}}\right) & R_2 > ST \\ \chi_{i,j}^{t} + Q \cdot L & R_2 \le ST \end{cases}, \tag{3}$$

here, t, $\chi_{i,j}^{t}$ denotes the position of the i-th sparrow in the j-th dimension. $iter_{max}$ is the total number of iterations, and α is a random value in [0, 1]. The parameters $R_2 \in [0,1]$ and $ST \in [0.5,1]$ correspond to the warning and safety thresholds, respectively. Q follows a normal distribution, and L is a $1 \times d$ vector with all entries equal to 1. When $R_2 < ST$, no predators are detected, enabling explorers to perform a global search. Conversely, $R_2 \ge ST$ signals nearby predators, prompting sparrows to adopt protective behavior. During foraging, certain follower sparrows continuously track explorers, moving immediately toward better food sources when an explorer discovers them. The position update rule for followers is defined as follows:

$$x_{i,j}^{t+1} = \begin{cases} Q \cdot \exp\left(\dfrac{\chi_{worst}^{t} - x_{i,j}^{t}}{i^2}\right) i > n/2 \\ \chi_{p}^{t+1} + \left| x_{i,j}^{t} - x_{p}^{t+1} \right| \cdot A^{+} \cdot L \end{cases}, \tag{4}$$

here, x_{p}^{t} indicates the position of the best-performing explorer, while x_{worst}^{t} corresponds to the globally worst position. The population size is denoted by n. $A$is a $1 \times d$ vector whose entries are independently set to either 1 or −1, and A^{+} is defined as follows:

$$A^{+} = A^{T}\left(AA^{T}\right)^{-1}, \tag{5}$$

when $i > n/2$, the i-th follower with lower fitness is considered to be in poor condition and must move elsewhere to search for food. In the algorithm, it is assumed that some individuals are alert to potential threats, and their initial positions (the watchers) are randomly assigned within the population:

$$x_{i,j}^{t+1} = \begin{cases} x_{best}^{t} + \beta \cdot \left| x_{i,j}^{t} - x_{best}^{t} \right| & f_i > f_g \\ x_{i,j}^{t} + k \cdot \left(\dfrac{\left| x_{i,j}^{t} - x_{worst}^{t} \right|}{(f_i - f_w) + \varepsilon} \right) & f_i = f_g \end{cases}, \tag{6}$$

here, x_{best}^{t} denotes the global best position at iteration t. The parameter β, drawn from a standard normal distribution, controls the step size, while $K \in [-1, 1]$ governs the search direction. f_i represents the fitness of the i-th sparrow, and f_g and f_w denote the best and worst fitness values in the current population, respectively. A small constant ε is introduced to avoid division by zero. When $f_i > f_g$, the sparrow is located at the boundary of the population; when $f_i = f_g$, it indicates that a central individual perceives potential risk and remains close to its neighbors.

In this study, SSA is applied to optimize critical hyperparameters of the model, including the number of LSTM neurons, convolutional kernel size, number of convolutional filters, and attention-related parameters. The optimization process is driven by model performance on the validation dataset, enabling adaptive selection of parameter configurations. Compared with manual tuning, this approach improves the stability and robustness of the model while enhancing prediction accuracy and generalization capability. Consequently, the integration of SSA provides an effective strategy for hyperparameter optimization in complex engineering prediction tasks.

2.4 Attention Mechanism (Transformer)

The Transformer architecture, built on the self-attention mechanism, was first introduced to address sequence modeling tasks. Its core concept lies in modeling global dependencies by computing pairwise interactions among elements within an input sequence through the Query (Q), Key (K), and Value (V) representations. The most common form is the scaled dot-product self-attention, expressed as:

$$Attention\,(Q, K, V) = softmax\left(\frac{QK^{T}}{\sqrt{d_k}}\right)V, \tag{7}$$

the matrices Q, K, and V encode the query, key, and value, respectively, and d_k indicates the size of each key vector.

Through this mechanism, the model can directly measure the relative importance of different positions in a sequence, enabling effective capture of long-range dependencies without relying on recurrent structures.

To further enhance representation capability, multi-head attention (MHA) is employed, which projects the input into multiple subspaces and performs self-attention in parallel:

$$MHA\,(Q, K, V) = Concat\,(head_1, \ldots, head_h)\,W^{O}, \tag{8}$$

$$head_i = Attention(QW_i^{Q}, KW_i^{K}, VW_i^{V}), \tag{9}$$

where h denotes the count of attention heads, and W_i^{Q}, W_i^{K}, and W_i^{V} are projection matrices that can be optimized during training.

In the multi-head attention mechanism, the input feature sequence is first transformed into three distinct functional representations: Query (Q), Key (K), and Value (V). The projection matrices W_i^Q, W_i^K, and W_i^V play a crucial role in this process by mapping the original high-dimensional input into different representation subspaces. Specifically, for each attention head i, the query matrix Q represents the current state for which the model is seeking information, the key matrix K contains the contextual information for all time steps, and the value matrix V holds the actual feature content to be propagated.

In the dot-product interaction between Q and K, the attention weights are calculated, which effectively measure the relevance between the current construction stage and historical monitoring states. By employing parallel heads, the model can simultaneously focus on diverse feature dimensions—for instance, one head might track the spatial influence of shield machine proximity (S), while another emphasizes abrupt changes in building inclination (Q). Finally, the outputs from all heads are concatenated and transformed using an output projection matrix W^O, so that the network can integrate these multifaceted temporal dependency patterns into a unified representation of the structural settlement evolution.

Different attention heads learn diverse temporal dependency patterns and feature interactions from different representation subspaces.

In the context of adjacent building settlement, these projection matrices serve a critical physical function: The mechanism automatically identifies and assigns higher weights to specific time steps—such as when the shield machine is directly undercrossing the structure or during periods of significant construction parameter fluctuations. In this way, the model can prioritize key physical disturbances over long-term monitoring sequences, effectively learning the non-linear mapping between tunneling progression and structural response while reducing redundancy in time series.

In this study, a lightweight Transformer encoder containing a single multi-head self-attention layer with four attention heads was integrated into the CNN-LSTM architecture to avoid excessive model complexity under limited data conditions. Different attention heads were designed to capture diverse temporal patterns and feature interactions, thereby improving the model's ability to characterize complex nonlinear deformation evolution.

2.5 SSA-CNN-LSTM-Transformer

During shield tunnel construction, settlement responses of adjacent buildings exhibit pronounced spatiotemporal correlation characteristics, manifested by spatial heterogeneity among monitoring points and temporal phase differences induced by construction progression. In addition, building deformation at the same location may occur in multiple forms, such as vertical settlement and inclination, which further increases the complexity of settlement prediction and places higher demands on modeling spatiotemporal dependencies. Collectively, the CNN, LSTM, attention mechanism, and SSA optimization form a unified spatiotemporal modeling framework, in which spatial feature extraction, temporal dependency learning, feature reweighting, and parameter optimization are jointly realized for accurate settlement prediction under shield tunneling conditions.

To effectively capture these characteristics, this study proposes an SSA-CNN-LSTM-Transformer hybrid model for multivariate time series settlement prediction, as illustrated in Fig. 4. The proposed model combines CNN and LSTM networks with a Transformer-based attention mechanism, while the Sparrow Search Algorithm (SSA) is used to optimize its hyperparameters. The CNN module is used to extract local spatial correlations and short-term temporal features from multi-source monitoring data through temporal convolution, transforming high-dimensional inputs into compact feature representations. Subsequently, the features are processed by the LSTM module, enabling the network to learn long-term temporal dependencies via its gated recurrence mechanism. To address the limited ability of conventional CNN-LSTM architectures to selectively emphasize critical information, a Transformer-based attention module is introduced to dynamically reweight the extracted features, highlighting key variables and critical time steps associated with shield construction. Moreover, SSA is employed to fine-tune critical model hyperparameters, such as convolutional kernel size, LSTM neuron count, and attention-related settings, which enhances both prediction accuracy and model robustness. The resulting model is then tested on an independent dataset, with its performance assessed using multiple evaluation metrics.

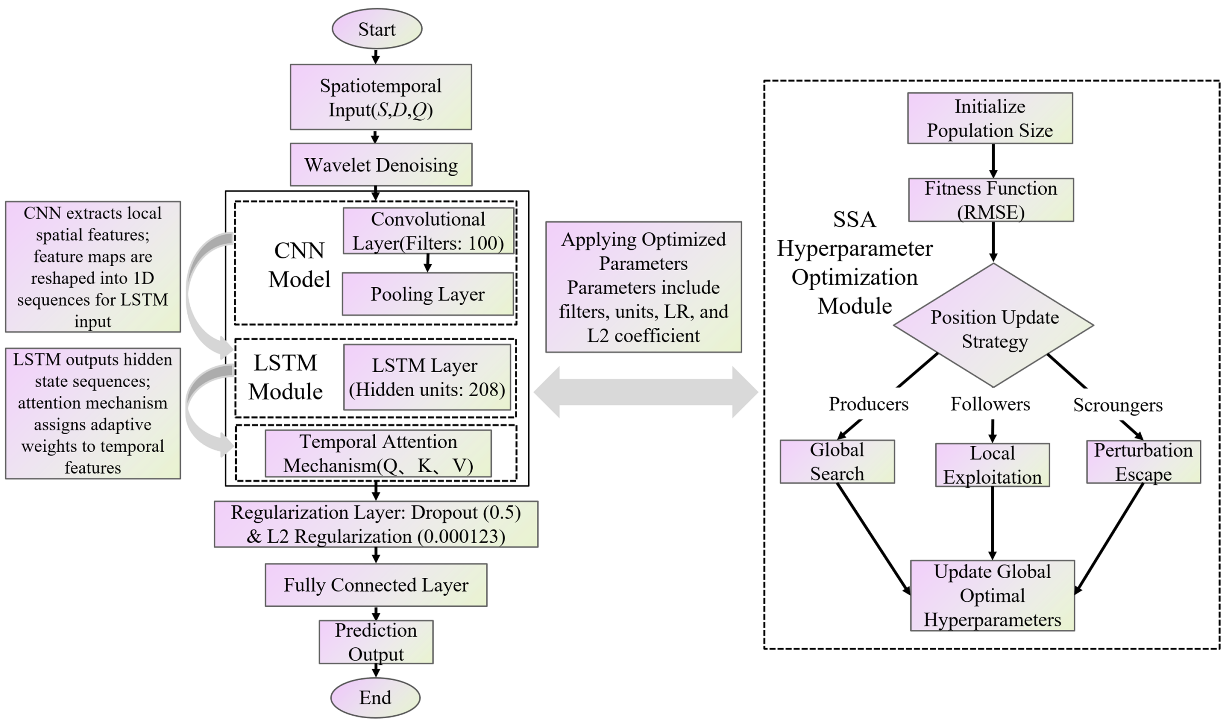

Figure 4: Schematic diagram of the proposed SSA-CNN-LSTM-Transformer framework.

Considering the specificity of engineering data, to prevent the deep network from overfitting on small samples, a Dropout layer with a rate of 0.5 was introduced after the self-attention layer, and an L2 regularization factor was configured. These regularization strategies, along with the optimized parameters, effectively penalize model complexity, ensuring smoothness when handling nonlinear settlement fluctuations.

2.6 Shield Tunneling-Oriented Modeling Strategy Considering Spatial Effects and Input–Output Construction

During shield tunneling, the deformation of buildings adjacent to the tunnel is mainly governed by the interaction between construction-induced ground disturbance and the spatial relationship between the advancing shield machine and the surrounding structures. As the shield machine progresses along the tunnel alignment, the magnitude and distribution of soil disturbance vary continuously in space, resulting in time-dependent and spatially non-uniform settlement responses of buildings. Therefore, in an effective deformation prediction model, the spatial evolution of shield tunneling effects and their coupling with building deformation monitoring data should be explicitly considered.

In this study, the spatial influence of shield tunneling is incorporated into the model through the construction of physically meaningful input variables derived from engineering monitoring data. The relative distance S between the shield excavation face and the reference point of the building is introduced as a key spatial indicator. This variable reflects the positional relationship between the shield machine and the building during construction and implicitly characterizes different tunneling stages, including shield approaching, passing beneath, and moving away from the building. By incorporating S into the input sequence, the spatial evolution of construction disturbance is transformed into a learnable temporal feature in the prediction framework.

To capture the spatial heterogeneity of building deformation induced by shield tunneling, settlement monitoring data from multiple locations on the building are selected as input variables. Settlement measurements at different monitoring points represent the non-uniform deformation responses of various parts of the structure and exhibit strong spatial correlation due to the integrity of the building–soil system. Multi-point settlement data is simultaneously used, so that the model can learn spatial interaction patterns among monitoring points, thus improving its ability to represent spatial effects in a data-driven framework.

In addition to settlement data, inclination monitoring data are incorporated to further describe the deformation characteristics of the building. Inclination reflects differential settlement behavior and provides complementary information on the overall deformation tendency of the structure under shield tunneling conditions. Through the combined use of settlement and inclination data, the model can capture multiple deformation modes of buildings influenced by shield construction.

Based on the above considerations, the model input is constructed as a multivariate time series consisting of the relative shield–building distance S, multi-point settlement value D, and corresponding inclination measurement Q. The output of the model is defined as the future settlement at representative building monitoring points, ensuring that the prediction results are directly interpretable and consistent with conventional engineering monitoring indices. This input–output construction establishes a clear mapping between shield tunneling progression, observed deformation responses, and future settlement evolution.

Through this shield tunneling-oriented modeling strategy, spatial effects and deformation characteristics are effectively embedded into the prediction framework using readily available monitoring data. The proposed input–output design provides a rational and engineering-consistent basis for the application of the SSA-CNN-LSTM-Transformer model to settlement prediction of buildings adjacent to shield tunnels, thus supporting practical deformation assessment and construction risk management.

To reduce random environmental noise and measurement interference in the field monitoring data, a Wavelet Denoising method was adopted in the data preprocessing stage. Compared with conventional smoothing techniques, wavelet denoising can effectively separate low-frequency deformation trends from high-frequency noise components while preserving the local characteristics and temporal evolution features of settlement signals. Considering the non-stationary characteristics of shield tunnelling settlement data, the Daubechies wavelet basis (db4) was selected due to its favorable time-frequency localization capability and suitability for engineering monitoring signal processing. Specifically, a 3-level wavelet decomposition was performed on the original monitoring sequence, as this level of decomposition provides an optimal balance by isolating environmental noise without filtering out critical local deformation features.

During the decomposition process, the low-frequency components representing the principal deformation trend were retained, while the high-frequency components associated with environmental disturbance, sensor fluctuation, and random measurement noise were processed using a soft-thresholding strategy. The soft-threshold function was adopted to avoid abrupt signal mutation caused by hard-threshold processing and to ensure smoother reconstruction of the monitoring sequence. In this process, the universal threshold λ is typically calculated as follows:

$$\lambda = \sigma\sqrt{2\ln(N)} \tag{10}$$

where λ denotes the threshold value, σ represents the standard deviation of the noise, and N indicates the length of the monitoring signal sequence. This parameterization ensures that only coefficients exceeding the noise floor contribute to the reconstruction, effectively filtering out sensor fluctuations and site disturbances. After threshold filtering, the denoised signal was reconstructed through inverse wavelet transformation to obtain the final preprocessed dataset. This process effectively suppressed high-frequency interference while preserving the overall cumulative settlement evolution characteristics and local deformation variation trends, thereby improving the stability and reliability of subsequent deep learning model training and prediction.

To ensure reproducibility and maximize the predictive capability of the proposed framework, the Sparrow Search Algorithm (SSA) was incorporated to automatically optimize critical hyperparameters of the SSA-CNN-LSTM-Transformer model. As a population-based metaheuristic optimization algorithm, SSA simulates the foraging and anti-predation behaviors of sparrows to iteratively explore the parameter space and identify the optimal network configuration. In this study, the Root Mean Square Error (RMSE) on the validation dataset was adopted as the fitness function for SSA optimization. The search space included several key hyperparameters, namely the convolution kernel size, the number of convolutional filters, the learning rate, and the number of LSTM hidden neurons. The SSA population size was set as 30, and the maximum iteration number was set to 50. Through the collaborative mechanism of producers, followers, and vigilant individuals, SSA effectively balanced global exploration and local exploitation, thereby reducing the dependence on empirical manual parameter tuning and improving the adaptability of the prediction framework.

Furthermore, to reduce the influence of randomness during network initialization and training, all experiments were conducted under fixed random seeds and repeated independently multiple times. The reported results correspond to the average prediction performance obtained from repeated runs to ensure statistical reliability and robustness.

To rigorously evaluate the generalization capability of the proposed SSA-CNN-LSTM-Transformer model under a constrained engineering dataset (consisting of 150 continuous temporal samples per monitoring point), a 5-fold cross-validation strategy was adopted. Using a sliding-window technique with a lag window size of 24 time steps, the original multivariate monitoring sequence of each point was transformed into 127 overlapping spatiotemporal feature samples. A many-to-one prediction structure was employed, in which the settlement value at time step $t + 1$ was predicted using monitoring information from the previous 24 time steps. The model input consisted of multivariate monitoring variables, including settlement measurements, inclination responses, and tunnelling construction parameters, while the output corresponded to the predicted settlement value at the target monitoring point.

During model training, the Mean Squared Error (MSE) was adopted as the loss function, and the Adam optimizer was employed for network parameter updating. The batch size was set to 16, and the maximum number of training epochs was set to 200. The initial learning rate was not manually assigned but was adaptively determined through the SSA optimization process, with the optimal learning rate identified as 0.005813 for the final model configuration.

3 Engineering Application

For an extended assessment of the proposed SSA-CNN-LSTM-Transformer hybrid model, a practical engineering case from the eastern extension of Nanchang Metro Line 2 was selected. The study area, located in the Xleng section, features a typical plain terrain and complex surrounding conditions, including roads, street-side commercial buildings, residential communities, and dense above-ground and underground pipelines. In this section, shield tunneling is employed, with the tunnel alignment passing beneath a high-density building cluster.

The selected case building adopts an independent foundation with a burial depth of approximately 2.25 m. The interval tunnel passes beneath the building with a minimum vertical clearance of about 13.24 m. According to engineering risk assessment criteria, the building is classified as a high-risk structure. During shield tunneling, continuous deformation monitoring was conducted along the tunnel alignment to ensure construction safety and to prevent potential engineering and environmental hazards. By comparing the measured deformation responses with model-predicted results, the effectiveness and rationality of construction control can be evaluated, providing technical support for timely adjustment and optimization of subsequent construction measures. The spatial relationship between the shield tunnel and the adjacent building, as well as the layout of building deformation monitoring points, is illustrated in Fig. 5.

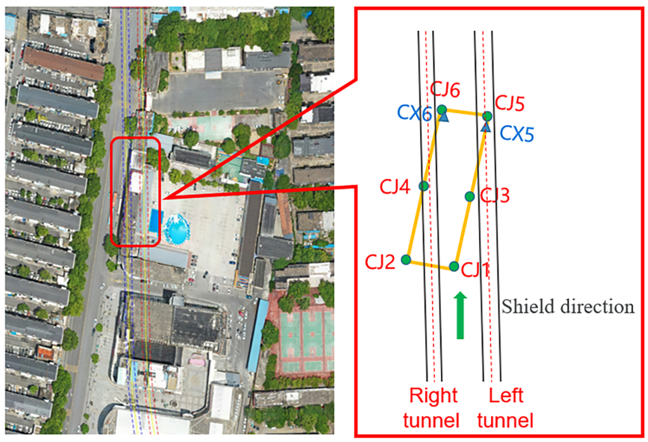

Figure 5: Adjacent building and Monitoring point distribution.

To support quantitative analysis and model validation, the geometric center of the building and the horizontal projection of the shield machine position were adopted as reference points for spatial characterization. Settlement monitoring data from points CJ1–CJ6, illustrated in Fig. 6, together with inclination monitoring data from points CX5 and CX6, shown in Fig. 7, were collected synchronously during shield advancement, forming the primary monitoring dataset for engineering application. The tilt values are expressed in per mile (‰), representing the structural inclination ratio. These monitoring points were selected to comprehensively reflect the deformation behavior of the building along and across the tunnel alignment, ensuring that both overall settlement trends and differential deformation characteristics were effectively captured.

The integration of these multi-source data is grounded in the specific physical constraints of engineering monitoring. In shield tunneling, the settlement of adjacent buildings is primarily governed by strong local spatiotemporal dependencies. During the critical undercrossing period, the structural response is mainly influenced by the dynamic interaction between the advancing excavation face and the building's foundation in the immediate vicinity. Consequently, historical data far removed from the current construction stage contribute less to the immediate settlement trend, which underscores that locally correlated monitoring sequences and spatial features are more critical for effective prediction than a vast volume of independent samples.

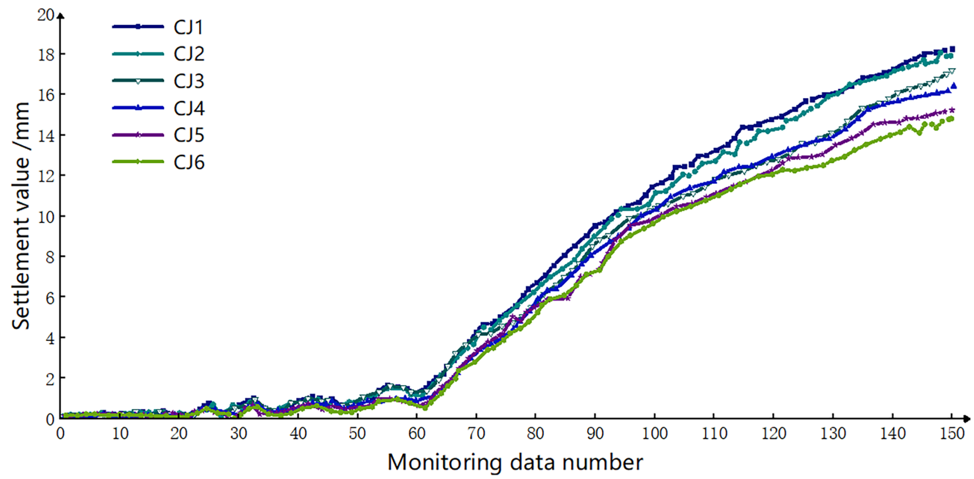

Figure 6: Building settlement monitoring data.

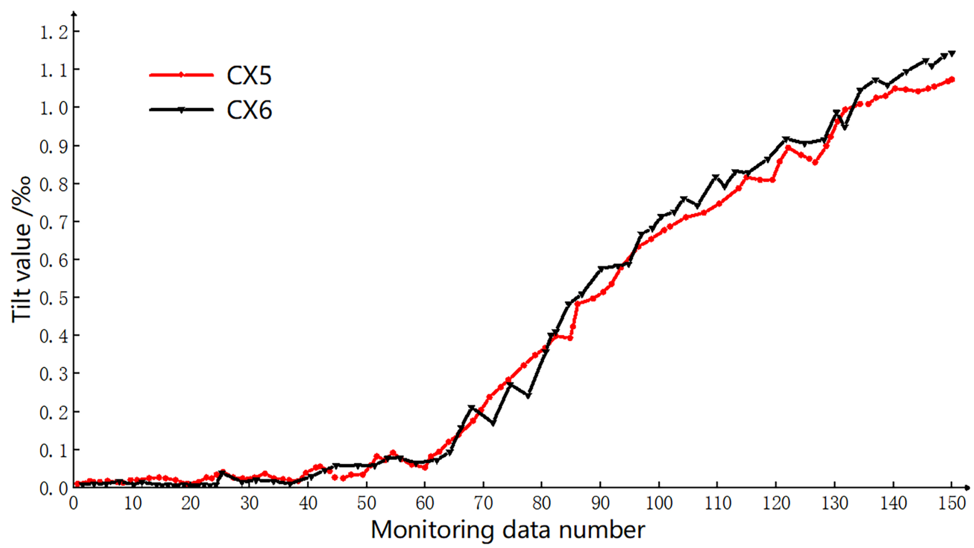

Figure 7: Building tilt monitoring data.

Based on this rationale, building deformation monitoring data comprising 150 synchronized, continuous temporal samples for each monitoring point, collected from 5 November to 10 November 2023, were utilized as the input dataset. During this period, the shield machine advanced at an average rate of eight rings per day, with each ring having a width of 1.2 m. Based on the advancement rate, the temporal evolution of shield tunneling progression was synchronized with the deformation monitoring schedule. Due to the influence of external factors such as temperature variation, weather conditions, and site disturbances, the raw monitoring data contained a certain level of noise. Therefore, a denoising procedure was applied prior to model training (as described in Section 2.6) to improve data quality and prediction reliability.

For model evaluation, settlement measurements at monitoring points CJ5 and CJ6 were selected as prediction targets, as these points are located near the region most sensitive to shield-induced disturbance and exhibit representative deformation characteristics. Based on the processed dataset, 50 consecutive samples (from the 101st to the 150th data points) were used for prediction and compared with corresponding measured values to assess model performance.

The time-series partitioning follows a uniform sampling interval (every 1 h). Although the shield advancement rate (approximately 8 rings per day) may exhibit minor fluctuations, such variations in construction intensity are explicitly encoded in the input variable S (relative distance). By utilizing a sliding window technique, the model maps the non-uniform physical advancement onto a uniform temporal axis, ensuring the synchronous learning of temporal dependencies and spatial disturbances.

In the proposed model, the inclination data from Points CX5 and CX6 are incorporated as auxiliary input variables (Q) rather than prediction targets. These points characterize the structural inclination ratio and differential settlement behavior, providing crucial complementary information on the overall deformation tendency and spatial heterogeneity of the building under shield tunneling disturbance. In contrast, settlement measurements at Monitoring Points CJ5 and CJ6 were selected as the primary prediction targets because they are situated in the region most sensitive to construction-induced disturbance and exhibit the most representative deformation characteristics for risk assessment. By using this input-output strategy, the model can leverage tilting responses to enhance the spatiotemporal representation of the structural system, while focusing its predictive output on the most critical settlement indicators for proactive safety management.

For assessing the performance of the proposed model in actual shield tunneling operations, several neural network architectures were implemented and trained in MATLAB, including a standard RNN, LSTM, bidirectional LSTM (BiLSTM), an attention-enhanced LSTM-Transformer, and the SSA-CNN-LSTM-Transformer model. All models were trained and tested using identical monitoring datasets and consistent data preprocessing and partition strategies to ensure a fair and objective comparison.

To ensure a fair comparison, all baseline models were trained using identical data preprocessing procedures, dataset partition strategies, and optimization settings. Key hyperparameters of the baseline recurrent neural network models were adjusted through consistent parameter tuning procedures to avoid bias caused by default parameter configurations.

The selected benchmark models cover both recurrent-based time-series prediction approaches and Transformer-based sequence modeling paradigms commonly used in deformation monitoring studies. By comparing the prediction performance of these models under identical shield tunneling conditions, the applicability and effectiveness of different time-series modeling strategies in capturing shield-induced building deformation evolution can be systematically evaluated. The comparison focuses on overall prediction accuracy and stability at the engineering application level, while the detailed contributions of individual model components are further investigated through ablation experiments presented in a subsequent section.

4 Result Analysis

4.1 Analysis of Prediction Results

After training on the multivariate time-series dataset, the last 50 settlement values at monitoring points CJ5 and CJ6 were predicted by each neural network model. The corresponding prediction results and error distributions are illustrated in Figs. 8–11. As shown in Figs. 8 and 10, during the initial prediction interval, both the conventional LSTM model and other baseline recurrent models exhibit relatively good agreement with the measured settlement values. However, as the prediction horizon extends, noticeable deviations gradually emerge, indicating the limited ability of these models to capture long-term dependencies and complex spatiotemporal relationships.

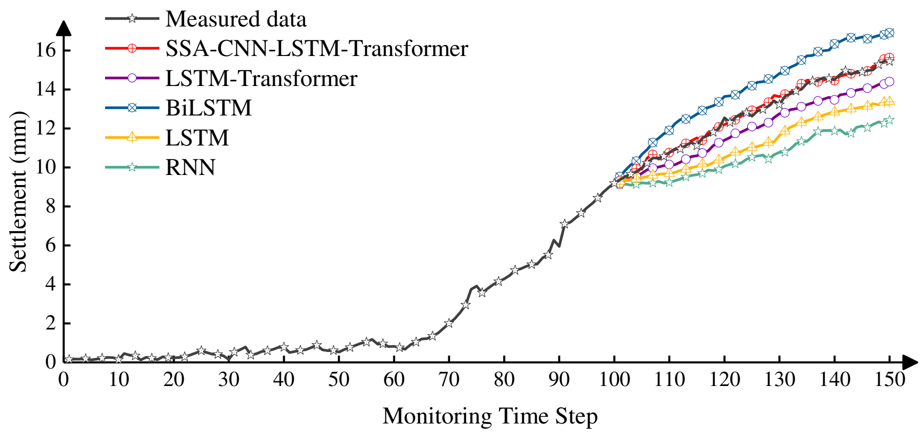

Figure 8: Predicted value and measured value of the CJ5 measurement point.

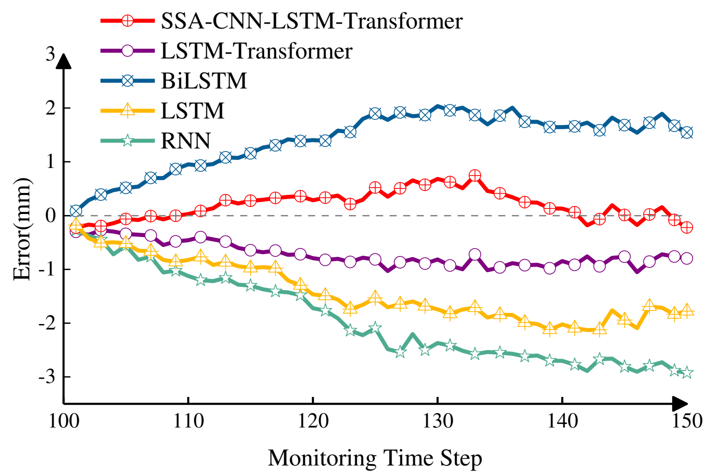

Figure 9: Error of the CJ5 measurement point prediction value.

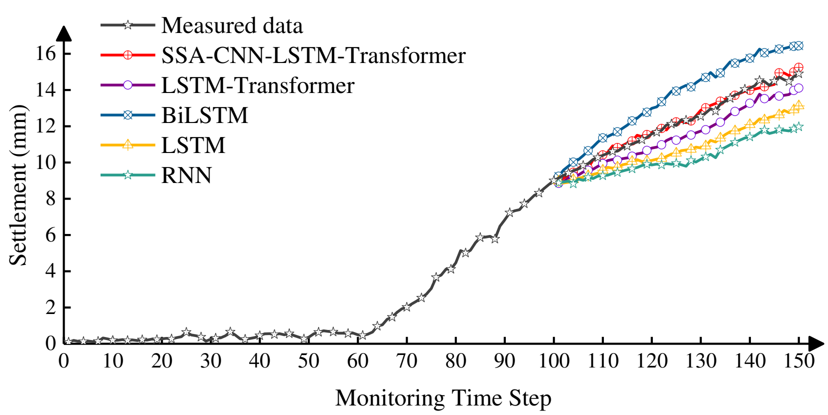

Figure 10: Comparison of the CJ6 predicted and measured values.

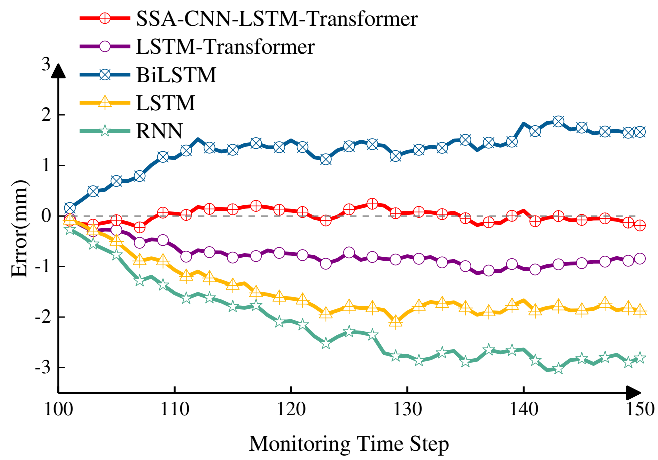

Figure 11: Error of the CJ6 predicted value of the measuring point.

In contrast, the proposed SSA-CNN-LSTM-Transformer hybrid model demonstrates significantly improved prediction performance throughout the entire prediction period. The incorporation of the CNN module enables effective extraction of spatial correlations among multiple monitoring points, while the LSTM component enhances the modeling of long-term temporal evolution. Furthermore, the Transformer-based self-attention mechanism adaptively reweights key features and critical time steps, allowing the model to emphasize influential construction stages and deformation responses. As a result, the proposed model exhibits enhanced robustness against error accumulation and effectively mitigates the degradation in prediction accuracy commonly observed in conventional recurrent models.

Further error analysis based on Figs. 9 and 11 shows that the forecasting deviations of the LSTM and other baseline models increase and become more unstable as the prediction horizon extends, indicating a gradual reduction in predictive reliability. In contrast, the SSA-CNN-LSTM-Transformer model maintains relatively stable deviations, fluctuating within a narrow range of approximately −1.0 to 1.0 mm. This behavior demonstrates enhanced prediction stability and accuracy, emphasizing the effectiveness of the attention-augmented hybrid architecture for long-term settlement forecasting.

Considering that the Peck empirical method is used to produce smooth theoretical settlement profiles rather than dynamic time-series fluctuations, its results are quantitatively compared in tabular form rather than directly superimposed in the time-series prediction figures to maintain figure readability.

Quantitative comparison results are summarized in Tables 1 and 2, which present the maximum error, mean error, and error variance of all models at monitoring points CJ5 and CJ6. As shown, the SSA-CNN-LSTM-Transformer model achieves a maximum error of only 0.419 mm at CJ5, compared with 2.129 mm for the traditional LSTM model, representing a reduction of approximately 80%. The mean error decreases from 1.416 to 0.103 mm, while the error variance drops from 0.303 to 0.032 mm^2. At CJ6, the model's maximum error is 0.383 mm, which is nearly 81% lower than the 2.104 mm observed for the LSTM model, with corresponding substantial reductions in mean error and error variance.

Compared with the LSTM-Transformer and BiLSTM models, the SSA-CNN-LSTM-Transformer also demonstrates more stable performance across all metrics. For instance, at CJ5, its mean error is approximately 85% lower and its error variance is about 34% lower than those of the LSTM-Transformer, indicating a stronger capability in capturing complex time-series correlations and nonlinear deformation behavior. Overall, the proposed model significantly outperforms traditional RNN, LSTM, BiLSTM, and LSTM-Transformer models in both prediction accuracy and stability.

Table 1: Comparison of model error metrics at CJ5.

Prediction Model	Maximum Error/mm	Mean Error/mm	Error Variance/mm^2
SSA-CNN-LSTM-Transformer	0.419	0.103	0.032
LSTM-Transformer	1.051	0.719	0.049
BiLSTM	2.035	1.403	0.267
LSTM	2.129	1.416	0.303
RNN	2.924	1.935	0.663
Peck	5.009	2.128	3.067

Table 2: Comparison of model error metrics at CJ6.

Prediction Model	Maximum Error/mm	Mean Error/mm	Error Variance/mm^2
SSA-CNN-LSTM-Transformer	0.383	0.093	0.029
LSTM-Transformer	1.133	0.767	0.059
BiLSTM	1.871	1.309	0.145
LSTM	2.104	1.491	0.274
RNN	3.051	2.160	0.595
Peck	4.787	1.821	3.354

Furthermore, the classic Peck formula was implemented as a physics-based baseline to evaluate the model's performance against traditional empirical methods. The Peck results exhibit a Mean Absolute Error (MAE) of 2.128 and 1.821 mm at CJ5 and CJ6, respectively, An error variance exceeds 3.0. This significant discrepancy underscores the limitations of empirical models in adjacent building scenarios: In the Peck formula a static "greenfield" profile is assumed, and the constraints imposed by structural stiffness are ignored. Consequently, it provides a rigid estimate that fails to capture the dynamic feedback during tunneling. In contrast, the proposed SSA-CNN-LSTM-Transformer model achieves a precision improvement of over 90% compared to this physical baseline, demonstrating substantial potential for practical engineering risk assessment and safety management.

Considering the relatively limited size of the engineering monitoring dataset and the strong temporal dependence characteristics of shield tunnel settlement data, a 5-fold cross-validation strategy was introduced to comprehensively evaluate the robustness and generalization capability of the proposed SSA-CNN-LSTM-Transformer model. Compared with a single train–test split, 5-fold cross-validation can effectively reduce the uncertainty caused by sample partitioning and improve the reliability of model evaluation under small-sample conditions. In addition, this strategy allows all available samples to participate in both training and validation processes, thereby enhancing the utilization efficiency of the monitoring data and providing a more objective assessment of the predictive performance of the proposed model. The corresponding cross-validation results are presented in Table 3.

The experimental results across the five folds for Monitoring Point CJ5 yielded a mean Root Mean Square Error (RMSE) of 0.1753 mm and a mean Absolute Error (MAE) of 0.1326 mm. Notably, the low standard deviation (±0.0201 mm for RMSE) demonstrates that the model maintains consistent predictive performance across different data partitions. Similarly, for CJ6, the mean RMSE and MAE were 0.1717 and 0.1305 mm, respectively. The standard deviation was ±0.0223 mm. This statistical evidence confirms that the hybrid architecture, reinforced by Dropout (0.5) and L2 regularization (0.000123), effectively mitigates the risk of overfitting and provides reliable forecasts despite the limited engineering sample size.

Table 3: 5-fold cross-validation results for settlement prediction at CJ5 and CJ6.

Fold	Metric	CJ5	CJ6
Fold1	RMSE (mm)/MAE (mm)	0.1725/0.1085	0.1520/0.1147
Fold2	RMSE (mm)/MAE (mm)	0.2091/0.1628	0.1893/0.1450
Fold3	RMSE (mm)/MAE (mm)	0.1735/0.1249	0.1616/0.1234
Fold4	RMSE (mm)/MAE (mm)	0.1588/0.1494	0.2009/0.1551
Fold5	RMSE (mm)/MAE (mm)	0.1625/0.1174	0.1546/0.1144

By explicitly integrating spatial feature extraction, long-term temporal modeling, and a Transformer-based attention mechanism, the proposed model effectively captures the coupled spatiotemporal effects of building deformation and shield tunneling processes. Consequently, it not only achieves higher prediction accuracy but also maintains stable performance during long-term forecasting, thus improving the reliability of settlement prediction under complex engineering conditions.

Ablation analyses were made to examine the role of each component in the model, The quantitative results are summarized in Tables 4 and 5, and qualitative distributions are shown in Figs. 12 and 13. The scatter distributions of predicted vs. measured settlement values at Monitoring Points CJ5 and CJ6 exhibit pronounced differences among the tested models. The predictions of the single LSTM model deviate considerably from the ideal prediction line, indicating limited accuracy.

Table 4: Quantitative ablation analysis results of different model components at Monitoring Point CJ5.

Prediction Model	Maximum Error/mm	Mean Error/mm	Error Variance/mm^2
SSA-CNN-LSTM-Transformer	0.419	0.103	0.032
CNN-LSTM-Transformer	1.035	0.694	0.033
LSTM-Transformer	1.051	0.719	0.049
LSTM	2.129	1.416	0.303

Table 5: Quantitative ablation analysis results of different model components at Monitoring Point CJ6.

Prediction Model	Maximum Error/mm	Mean Error/mm	Error Variance/mm^2
SSA-CNN-LSTM-Transformer	0.383	0.093	0.029
CNN-LSTM-Transformer	0.918	0.692	0.031
LSTM-Transformer	1.133	0.767	0.059
LSTM	2.104	1.491	0.274

By introducing the Transformer attention mechanism, the LSTM–Transformer model achieves higher prediction accuracy; however, its scatter distribution reveals a consistent underestimation tendency, which may pose safety risks in practical engineering by failing to warn of maximum settlement. With the further integration of the CNN module, although the improvement in numerical error metrics is relatively incremental as shown in Tables 4 and 5, the CNN–LSTM–Transformer model demonstrates superior engineering significance. Specifically, its predicted values tend to be slightly higher than the measured data, shifting the distribution toward the upper side of the ideal line. This "conservative prediction" characteristic is crucial for construction safety, as it provides a safety margin for risk assessment. Finally, the SSA–CNN–LSTM–Transformer model, optimized via the Sparrow Search Algorithm, exhibits the tightest scatter distribution.

The predicted points are most closely aligned with the ideal line. This highlights the superior accuracy and stability achieved by the final hybrid architecture. These findings indicate that the integration of CNN spatial extraction specifically improves the model's reliability for proactive risk control in shield tunneling projects while each component enhances numerical performance.

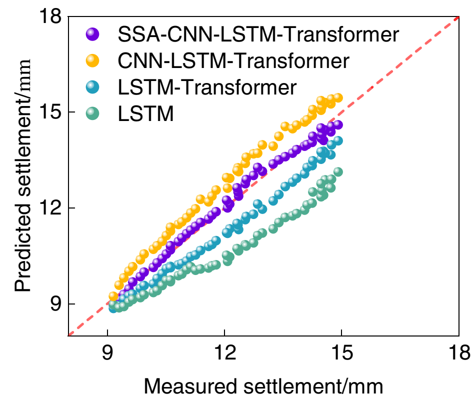

Figure 12: Stepwise ablation comparison chart of CJ5 monitoring point settlement prediction results.

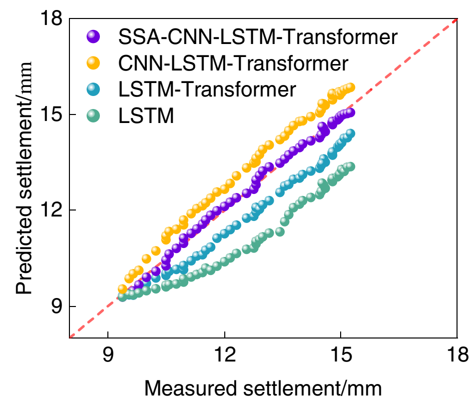

Figure 13: Stepwise ablation comparison chart of CJ6 monitoring point settlement prediction results.

4.2 Parameter Analysis

In applied forecasting, the predictive performance of deep learning models depends heavily on hyperparameter selection. To further improve prediction accuracy and reliability, key parameters of the SSA-CNN-LSTM-Transformer hybrid model were systematically examined and tuned. The Sparrow Search Algorithm (SSA) was employed to optimize critical hyperparameters of the convolutional and recurrent layers, enhancing feature extraction efficiency while reducing information loss and training instability.

The effects of convolution kernel size and stride on model performance are illustrated in Fig. 14. The results show that these parameters play a crucial role in determining the model's feature extraction effectiveness and overall stability. Among the tested configurations, a convolution kernel size of 6 × 5 with a stride of 1 achieves the lowest RMSE, indicating the best overall prediction performance. This configuration enables sufficient extraction of spatial features while avoiding excessive information loss during convolution operations. In contrast, inappropriate kernel sizes or stride settings either reduce feature resolution or increase model complexity, which leads to degraded prediction accuracy and reduced robustness.

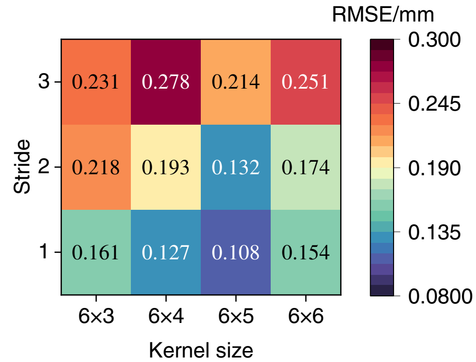

Figure 14: Mean square deviation of the calculated results of different convolutional layer parameters.

Fig. 15 illustrates the relationship between the number of convolutional kernels and the RMSE. It can be observed that the number of kernels has a nonlinear impact on model performance, with distinct differences in RMSE for different kernel quantities. When the number of kernels is relatively small, the RMSE remains high—for example, at 40 kernels, the RMSE reaches a maximum of 0.241—indicating insufficient spatial feature extraction and limited capability to capture the correlations among multivariate inputs. As the kernel count increases, the RMSE gradually decreases, reaching a minimum of 0.108 at 100 kernels, which demonstrates that the model can effectively extract spatial features and achieves optimal predictive performance.

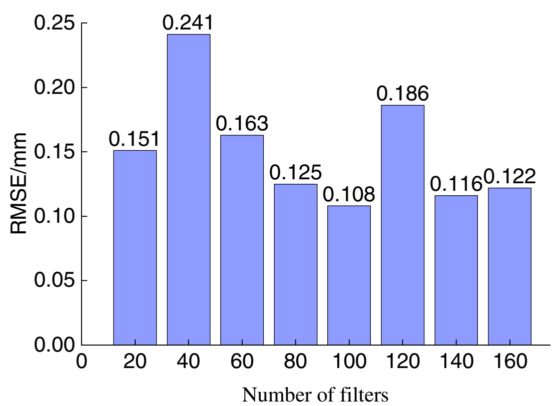

Figure 15: Root mean square error of prediction results for different convolution kernel numbers.

However, further increasing the number of kernels leads to fluctuations and a rise in RMSE. At 120 kernels, the RMSE increases to 0.186, suggesting that an excessive number of kernels may introduce redundant features, increase model complexity, and adversely affect generalization. Although the RMSE slightly decreases at 140 and 160 kernels, it remains higher than the optimal configuration. These observations indicate that too few kernels limit feature representation, while too many kernels can cause overfitting and reduced generalization ability. Therefore, in this study, 100 convolutional kernels are selected as the optimal parameter setting, ensuring both high prediction accuracy and model stability.

Fig. 16 depicts how varying the number of neurons in the LSTM hidden layer affects model performance. The selection of neuron number is crucial for achieving a balance between fitting accuracy and generalization ability. When the number of neurons is small, the model exhibits clear underfitting behavior, failing to adequately capture the intrinsic temporal patterns of the settlement data, which leads to large prediction errors despite relatively short training time. Conversely, an excessively large number of neurons causes overfitting, characterized by increased prediction errors and significantly prolonged training time due to overlearning of noise and local details.

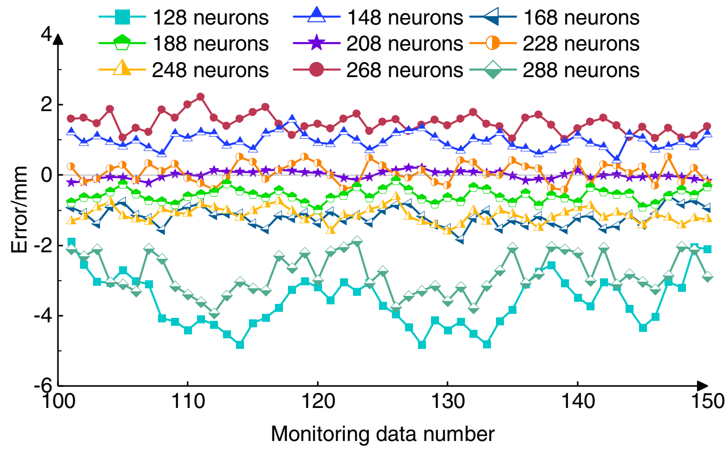

Figure 16: Absolute error of different neuronal parameter models.

Within a moderate range of neuron numbers, the model demonstrates stable training behavior and lower RMSE values, indicating a favorable trade-off between model complexity and generalization capability. Notably, the optimal performance is achieved when the number of neurons is set to 208, at which point the model attains the highest prediction accuracy and strong generalization ability. The optimal neuron number identified by the Sparrow Search Algorithm is consistent with this result, further confirming the effectiveness and rationality of SSA in optimizing LSTM hyperparameters within the CNN-LSTM-Transformer framework.

Through the iterative search of the Sparrow Search Algorithm (SSA), the optimal hyperparameter configuration was identified: the number of convolutional filters was set to 100, with a kernel size of 6×5 and a stride of 1. Furthermore, the optimal number of LSTM hidden neurons was determined to be 208. This configuration enables sufficient extraction of spatial features while avoiding excessive information loss or overfitting.

Overall, the parameter analysis results demonstrate that SSA can accurately and efficiently identify optimal hyperparameter configurations for the CNN-LSTM-Transformer model. By jointly optimizing convolution kernel size, stride, kernel number, and LSTM neuron count, the proposed model achieves improved prediction accuracy, stability, and robustness. Moreover, in practical engineering applications, as construction progresses and spatial effects evolve, reference data for post-construction prediction continuously accumulate, leading to dynamic changes in the input matrix. Under such conditions, manual parameter tuning becomes time-consuming and labor-intensive. The introduction of SSA enables adaptive and efficient parameter optimization, providing a practical and reliable solution for real-world engineering prediction tasks.

To further validate the effectiveness of the Sparrow Search Algorithm (SSA) in hyperparameter optimization, a comparative study was conducted against two commonly used methods: Particle Swarm Optimization (PSO) and Bayesian Optimization (BO). The optimization targets included critical parameters affecting both the spatial and temporal modeling layers, such as the number of convolutional kernels (Filters), kernel size ($Kernel_H \times Kernel_W$), convolutional stride (s), the number of LSTM hidden neurons (d_{model}), the initial learning rate (LR), and the L2 regularization factor (L_2). All algorithms were evaluated on the same multivariate engineering dataset using the Root Mean Square Error (RMSE) as the fitness function, with identical evaluation budgets to ensure a fair and rigorous comparison.

The hyperparameter search space was formally defined with convolutional filters ranging from 16 to 128, kernel dimensions from 3 × 3 to 9 × 9, the convolutional stride spanning across discrete integers from 1 to 3, LSTM hidden units from 32 to 256, the learning rate from 1e−4 to 1e−2, and the L2 factor from 1e−4 to 0.1. Regarding algorithm-specific settings, SSA was configured with a population of 30 and 50 iterations, utilizing a role-based structure consisting of 20% producers for global exploration and 10%–20% vigilant individuals to trigger random position adjustments, preventing the model from falling into local optima. PSO was similarly configured with 30 particles and 50 iterations, maintaining an inertia weight of $w = 0.8$ and acceleration coefficients of $c_1 = c_2 = 1.5$. Bayesian Optimization utilized a Gaussian Process (GP) surrogate model with an Expected Improvement (EI) acquisition function over 50 evaluation cycles to balance the exploration of unknown parameter regions and the exploitation of known optimal configurations.

Analysis of the convergence characteristics reveals distinct behaviors among the three methods. While PSO demonstrated a rapid initial error reduction during the early exploration phase, it tended to stagnate at local optima around the 20th iteration due to its simpler update mechanism. Bayesian Optimization provided a steady but slower refinement process, facing challenges in efficiently navigating the high-dimensional and non-linear parameter space inherent in tunneling monitoring data. In contrast, SSA leveraged its unique "producer-follower-vigilant" mechanism to successfully navigate the complex landscape. As shown in Table 6, SSA identified a global optimal configuration—consisting of 100 convolutional filters, a 6 × 5 kernel size, 208 LSTM neurons, and a learning rate of 0.00581—achieving a minimum RMSE of 0.108 mm. This represents a precision improvement of approximately 50% over the RMSE values of 0.218 mm and 0.217 mm achieved by PSO and BO, respectively, confirming SSA's superior search efficiency, stability, and robustness for complex nonlinear engineering datasets.

Table 6: Comparison of hyperparameter optimization results among SSA, PSO, and Bayesian optimization (BO).

Optimization Algorithm	Number of Convolution Kernels	Kernel Size	Stride	Number of Hidden Neurons	Learning Rate	RMSE (mm)
SSA	100	6 × 5	1	208	0.00581	0.108
PSO	16	7 × 3	1	87	0.00108	0.218
BO	123	9 × 4	3	164	0.00013	0.217

As consolidated in the empirical results, Table 6 reports the global optimal fitness value achieved during the hyperparameter optimization phase, whereas Table 3 delineates the averaged cross-validation performance under diversified data partitions. Mechanistically, since the 5-fold cross-validation scheme rigorously evaluates the structural generalization capability across multiple distinct data splits rather than merely reflecting the peak performance on a single validation subset, the averaged cross-validation RMSE values (0.1753 mm for CJ5 and 0.1717 mm for CJ6) are theoretically and expectedly higher than the deterministic optimal fitness value (0.108 mm) convergence tracked during the SSA exploration. This numerical discrepancy further substantiates that while the optimal configuration marks the upper bound of the model's single-run precision, the framework maintains an exceptionally stable and low error variance under rigorous data perturbations. Consequently, these synergistic findings robustly demonstrate that the Sparrow Search Algorithm (SSA) provides a highly efficient, responsive, and statistically stable hyperparameter mapping mechanism, showing distinct algorithmic superiority when navigating small-sample, highly non-linear engineering monitoring datasets.

5 Conclusions

Building upon earlier work on the use of neural networks for engineering deformation forecasting, this research extends the conventional single LSTM framework by integrating a Convolutional Neural Network (CNN) layer and a Transformer module. Shield tunneling construction progress is incorporated as a key input variable, while the spatial correlation of deformation among different building locations is explicitly considered. On this basis, an SSA-CNN-LSTM-Transformer hybrid model is constructed to make multivariate time series prediction of building settlement induced by shield tunneling.

(1) To address the limitations of traditional models in representing multi-point spatial correlations and construction dynamics, in this research, an engineering-oriented spatiotemporal prediction mechanism that explicitly incorporates spatial influences and multi-source data coupling is established. The relative distance S between the shield excavation face and the building is introduced as a key spatial indicator, transforming distinct tunneling stages—including approaching, passing beneath, and moving away—into learnable temporal features in the model input. Simultaneously, the CNN module is employed to extract spatial correlations between multi-point settlement values D and inclination data Q, effectively capturing the non-uniform deformation and spatial heterogeneity inherent in the integrated building–soil system. This framework is further enhanced by a Transformer-based self-attention mechanism that identifies critical construction stages, such as the undercrossing period, and highlights key influencing features in the multivariate data, thus strengthening the model's representation of global dynamic relationships and long-term dependencies. To ensure robustness during implementation, the raw data undergoes db4 Wavelet Denoising to filter environmental noise while preserving physical trends, and a sliding-window strategy with a lag size of 24 is used to synchronize physical advancement with a uniform temporal axis. Furthermore, to mitigate overfitting risks for the constrained dataset consisting of 150 continuous temporal samples per monitoring point, the architecture incorporates a Dropout rate of 0.5 and L2 regularization of 0.000123.

(2) Experimental results from thorough comparisons demonstrate the clear advantages of the SSA-CNN-LSTM-Transformer model in the prediction of settlement. At Monitoring Points CJ5 and CJ6, the prediction errors of the model remain within a narrow range of approximately −1.0 to 1.0 mm throughout the prediction horizon, showing stable trends without noticeable error accumulation. Quantitative evaluation indicates that, the maximum prediction error is reduced by over 80%, while the mean error is reduced by approximately 90% (from 1.416 to 0.103 mm at CJ5) compared with the traditional LSTM model. Significant improvements are also observed compared with RNN, BiLSTM, and LSTM-Transformer models, as well as the empirical Peck formula, confirming the model's clear superiority in both accuracy and stability. These findings indicate that the SSA-CNN-LSTM-Transformer model captures nonlinear deformation patterns and long-term temporal dependencies more effectively, resulting in greater robustness and reliability compared with traditional recurrent neural networks. Consequently, the model demonstrates strong potential for practical engineering applications, offering a dependable tool for forecasting complex structural deformations and supporting proactive risk management.

(3) The introduction of the Sparrow Search Algorithm (SSA) significantly improves hyperparameter configuration efficiency, where the optimization process is driven by model performance on the validation dataset to enable adaptive parameter selection. Comparative studies demonstrate that the SSA-optimized model achieves an RMSE of 0.108 mm, which is substantially lower than the errors obtained via Particle Swarm Optimization (PSO, 0.218 mm) or Bayesian Optimization (BO, 0.217 mm). Furthermore, SSA demonstrates superior global search effectiveness in identifying the optimal configuration—consisting of 208 LSTM hidden neurons, 100 convolutional filters, and a

6 × 5 kernel size—thus preventing the model from falling into local optima and ensuring higher stability and robustness for nonlinear engineering datasets.

Furthermore, the proposed framework offers significant practical value for engineering safety management. It can be integrated into real-time monitoring platforms to provide 24-h advance warnings of differential settlement risks, so that field engineers can dynamically adjust key tunneling parameters—such as thrust, torque, and grouting volume—facilitating a transition from passive monitoring to proactive risk control for high-risk buildings in urban tunneling projects.

Although the proposed framework achieved stable prediction performance under the current engineering dataset, the monitoring data were collected from a single shield tunneling project with a limited sample size. Therefore, the generalization capability of the model under different geological conditions, tunnel configurations, and building types requires further validation.

Future studies will incorporate larger multi-project datasets and more diverse engineering conditions to further improve the robustness and generalization performance of the proposed framework.

Acknowledgement: Not applicable.

Funding Statement: This work was supported by the National Natural Science Foundation of China (No. 52578465) and the Liaoning Provincial Science and Technology Plan Joint Project: Key Research and Development (R&D) Plan Project (Grant No. 2025JH2/101800303).

Author Contributions: Conceptualization, Jiawang Zou and Annan Jiang; methodology, Jiawang Zou; model development and software implementation, Jiawang Zou; data preprocessing and curation, Jiawang Zou; validation and formal analysis, Jiawang Zou, Xinzhi Wang and Hao Huang; investigation and engineering case analysis, Jiawang Zou and Xinzhi Wang; visualization, Jiawang Zou and Hao Huang; writing—original draft preparation, Jiawang Zou; writing—review and editing, Jiawang Zou and Annan Jiang; supervision, Annan Jiang; project administration, Annan Jiang; funding acquisition, Annan Jiang. All authors reviewed and approved the final version of the manuscript.

Availability of Data and Materials: The data that support the findings of this study are available from the corresponding author upon reasonable request.

Ethics Approval: Not applicable.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript

LSTM	Long Short-Term Memory
CNN	Convolutional Neural Network
SSA	Sparrow Search Algorithm
RNN	Recurrent Neural Network
BiLSTM	Bidirectional Long Short-Term Memory
LSTM-Transformer	Long Short-Term Memory–Transformer
CNN-LSTM-Transformer	Convolutional Neural Network–Long Short-Term Memory–Transformer
SSA-CNN-LSTM-Transformer	Sparrow Search Algorithm–Convolutional Neural Network–Long Short-Term Memory–Transformer

References

1. Sousa RL, Einstein HH. Lessons from accidents during tunnel construction. Tunn Undergr Space Technol. 2021;113:103916. doi:10.1016/j.tust.2021.103916.
2. Vlasov SN, Makovsky LV, Merkin VE. Accidents in transportation and subway tunnels: construction to operation. Moscow, Russia: Elex-KM Publishers; 2001.
3. Xu T, Song Z, Guo D, Song Y. A cloud model-based risk assessment methodology for tunneling-induced damage to existing tunnel. Adv Civ Eng. 2020;2020(1):8898362. doi:10.1155/2020/8898362.
4. Moosazadeh S, Namazi E, Aghababaei H, Marto A, Mohamad H, Hajihassani M. Prediction of building damage induced by tunnelling through an optimized artificial neural network. Eng Comput. 2019;35(2):579–91. doi:10.1007/s00366-018-0615-5.
5. Shen SL, Wu HN, Cui YJ, Yin ZY. Long-term settlement behaviour of metro tunnels in the soft deposits of Shanghai. Tunn Undergr Space Technol. 2014;40(11):309–23. doi:10.1016/j.tust.2013.10.013.
6. Shen Y, Zhang D, Wang R, Li J, Huang Z. SBD-K-medoids-based long-term settlement analysis of shield tunnel. Transp Geotech. 2023;42(344):101053. doi:10.1016/j.trgeo.2023.101053.
7. Soga K, Laver RG, Li Z. Long-term tunnel behaviour and ground movements after tunnelling in clayey soils. Undergr Space. 2017;2(3):149–67. doi:10.1016/j.undsp.2017.08.001.
8. Song ZP, Jiang AN, Jiang ZB. Back analysis of geomechanical parameters using hybrid algorithm based on difference evolution and extreme learning machine. Math Probl Eng. 2015;2015(6):821534. doi:10.1155/2015/821534.
9. Song Z, Yang Z, Huo R, Zhang Y. Inversion analysis method for tunnel and underground space engineering: a short review. Appl Sci. 2023;13(9):5454. doi:10.3390/app13095454.
10. Wu Q, Yan B, Zhang C, Wang L, Ning G, Yu B. Displacement prediction of tunnel surrounding rock: a comparison of support vector machine and artificial neural network. Math Probl Eng. 2014;2014(1):351496. doi:10.1155/2014/351496.
11. Jearsiripongkul T, Keawsawasvong S, Thongchom C, Ngamkhanong C. Prediction of the stability of various tunnel shapes based on hoek-brown failure criterion using artificial neural network (ANN). Sustainability. 2022;14(8):4533. doi:10.3390/su14084533.
12. Yu K, Wu L, Zhang Q, Liu Y, Chen J, Li D. Settlement prediction of buildings surrounding excavations based on ALSTM model. Eng Investig. 2023;51(3):63–72. (In Chinese).
13. Hong YC, Qian JG, Ye YX, Cheng L. Application of CNN-LSTM model based on spatiotemporal correlation characteristics in deformation prediction of excavation engineering. Chin J Geotech Eng. 2021;43(S2):108–11. (In Chinese). doi:10.11779/CJGE2021S2026.
14. Wang CY, Li LM, Wen ZZ, Zhang MY, Wei XW. Dynamic prediction of landslide displacement based on time series and CNN-LSTM. Foreign Electron Meas Technol. 2022;41(3):1–8. (In Chinese). doi:10.19652/j.cnki.femt.2103321.
15. Chen ZY, Fan LT, Zhang YP, Xiao HH, Wang L. Knowledge-based and data-based machine learning in intelligent TBM construction. China Civ Eng J. 2024;57(6):1–12. (In Chinese). doi:10.15951/j.tmgcxb.23090744.
16. Elmousalami HH. Artificial intelligence and parametric construction cost estimate modeling: state-of-the-art review. J Constr Eng Manag. 2020;146(1):03119008. doi:10.1061/(asce)co.1943-7862.0001678.
17. Wang X, Bai Y, Liu X. Prediction of railroad track geometry change using a hybrid CNN-LSTM spatial-temporal model. Adv Eng Inform. 2023;58:102235. doi:10.1016/j.aei.2023.102235.
18. Xu X, Zheng Y, Tong Z, Zhu P. Research on structural health monitoring database construction and intelligent diagnosis method based on CNN-LSTM integrated model. Struct Health Monit. 2025;4:619. doi:10.1177/14759217251373102.
19. Liu W, Tong L, Zhang Y, Li H, Xiao Y, Wu H. Spatiotemporal-information-driven surrogate modeling for predictive analysis of shield tunneling-induced settlement. Transp Geotech. 2026;57:101850. doi:10.1016/j.trgeo.2025.101850.
20. Li J, Huang H, Wang G. A small-sample prediction model for ground surface settlement in shield tunneling based on adjacent-ring graph convolutional networks (GCN-SSPM). Buildings. 2025;15(19):3519. doi:10.3390/buildings15193519.

21. Yin Q, Zhou Y, Rao J. An intelligent prediction method for surface settlement of shield tunnel construction based on CNN-BiLSTM-SA. KSCE J Civ Eng. 2025;29(4):100061. doi:10.1016/j.kscej.2024.100061.
22. Liu L, Zhou W, Gutierrez M. Effectiveness of predicting tunneling-induced ground settlements using machine learning methods with small datasets. J Rock Mech Geotech Eng. 2022;14(4):1028–41. doi:10.1016/j.jrmge.2021.08.018.
23. Zhang Y, Zhong W, Li Y, Wen L. A deep learning prediction model of DenseNet-LSTM for concrete gravity dam deformation based on feature selection. Eng Struct. 2023;295(1):116827. doi:10.1016/j.engstruct.2023.116827.
24. Zou N, Cheng HZ, Dai T, Yang ZH. Real-time surface settlement prediction in shield tunneling using a time-series framework and interpretable machine learning. Results Eng. 2026;29:109478. doi:10.1016/j.rineng.2026.109478.
25. Guo D, Song Z, Xu T, Zhang Y, Ding L. Coupling analysis of tunnel construction risk in complex geology and construction factors. J Constr Eng Manag. 2022;148(9):04022097. doi:10.1061/(asce)co.1943-7862.0002360.
26. Tian X, Song Z, Zhang Y. Monitoring and reinforcement of landslide induced by tunnel excavation: a case study from Xiamaixi tunnel. Tunn Undergr Space Technol. 2021;110(7):103796. doi:10.1016/j.tust.2020.103796.
27. Tian X, Song Z, Wang H, Zhang Y, Wang J. Evolution characteristics of the surrounding rock pressure and construction techniques: a case study from Taoshuping tunnel. Tunn Undergr Space Technol. 2022;125(1–4):104522. doi:10.1016/j.tust.2022.104522.
28. Antoniou F, Aretoulis G, Giannoulakis D, Konstantinidis D. Cost and material quantities prediction models for the construction of underground metro stations. Buildings. 2023;13(2):382. doi:10.3390/buildings13020382.
29. Mahmoodzadeh A, Mohammadi M, Daraei A, Faraj RH, Mohammed Dler Omer R, Sherwani AFH. Decision-making in tunneling using artificial intelligence tools. Tunn Undergr Space Technol. 2020;103:103514. doi:10.1016/j.tust.2020.103514.
30. Domingos P. A few useful things to know about machine learning. Commun ACM. 2012;55(10):78–87. doi:10.1145/2347736.2347755.
31. Yu S, Zhang Z, Wang S, Huang X, Lei Q. A performance-based hybrid deep learning model for predicting TBM advance rate using Attention-ResNet-LSTM. J Rock Mech Geotech Eng. 2024;16(1):65–80. doi:10.1016/j.jrmge.2023.06.010.
32. Liu J, Jiang Y, Ishizu S, Sakaguchi O. Estimation of tunnel support pattern selection using artificial neural network. Arab J Geosci. 2020;13(9):321. doi:10.1007/s12517-020-05311-z.
33. Peck RB. Deep excavations and tunneling in soft ground. In: Proceedings of the 7th International Conference on Soil Mechanics and Foundation Engineering; 1969 Aug 25–29; Mexico City, Mexico. p. 225–90.
34. Qin Y, Zhou J, Xiao D, Qin C, Qian Q. High-precision cutterhead torque prediction for tunnel boring machines using an attention-based embedded LSTM neural network. Measurement. 2024;224:113888. doi:10.1016/j.measurement.2023.113888.
35. Liu X, Li K, Jiang A, Fang Q, Zhang R. Prediction interaction responses between railway subgrade and shield tunnelling using machine learning with sparrow search algorithm. Transp Geotech. 2024;44(9):101169. doi:10.1016/j.trgeo.2023.101169.