applied sciences

MDPI

Article

Accurate Prediction of Tunnel Face Deformations in the Rock Tunnel Construction Process via High-Granularity Monitoring Data and Attention-Based Deep Learning Model

Mingliang Zhou [1], Zhenhua Xing [2], Cong Nie [1,*], Zhunguang Shi [2], Bo Hou [2] and Kang Fu [3,*]

[1] Department of Geotechnical Engineering, Tongji University, Shanghai 200092, China
[2] China Construction Third Engineering Bureau Group Co., Ltd., Wuhan 430070, China
[3] China State Construction Railway Investment & Engineering Group Co., Ltd., Beijing 100053, China
* Correspondence: 2132586@tongji.edu.cn (C.N.); fukang@cscec.com (K.F.)

Abstract: Monitoring and predicting the deformation of surrounding rocks in the rock tunnel construction process is of great significance. This study implemented a wireless sensor network (WSN), including gateway transmission, relay point, and sensor nodes, to obtain high granularity deformation data during construction. A transformer model is proposed, which considers the construction sequence into the positional embedding and has an attention module to deeply learn the high dimensionality correlation between the nearby deformation data and the tunnel face deformation. The attention-enhanced LSTM model and the LSTM model are also constructed to compare them with the performance of the transformer model. A site study conducted on a shallow buried tunnel section suggested an excellent performance of the proposed WSN system. The transformer model shows the best performance in terms of the model prediction results, which can extract more information from the time sequence data than the attention-enhanced LSTM and LSTM models. The proposed system has great value as guidance and reference for the construction of rock tunnel projects in complex and unfavourable geological conditions.

Keywords: rock tunnel construction; tunnel face deformation; tunnel monitoring; deep learning model; attention module

check for updates

Citation: Zhou, M.; Xing, Z.; Nie, C.; Shi, Z.; Hou, B.; Fu, K. Accurate Prediction of Tunnel Face Deformations in the Rock Tunnel Construction Process via High-Granularity Monitoring Data and Attention-Based Deep Learning Model. *Appl. Sci.* **2022**, *12*, 9523. https://doi.org/10.3390/app12199523

Academic Editor: Daniel Dias

Received: 30 August 2022
Accepted: 19 September 2022
Published: 22 September 2022

1. Introduction

The rapid growth of traffic demand has led to the increasing construction of rock tunnels, especially in China [1,2]. Although tunnel boring machines have gained popularity in the last two decades due to the complex and uncertain geological conditions in rock tunnel engineering, the drilling and blasting method is still the most widely used construction process in rock tunnel excavation [3,4]. The problem of large squeezing deformation of the excavated tunnel face is often encountered during the working progress in a rock tunnel [5] under high in situ stress or unfavourable geological conditions consisting of a weak surrounding rock mass. Therefore, grasping the tunnel face deformation during the rock tunnel construction process is an important research topic for civil engineering scholars [6–8].

During rock tunnel excavation, an unstable surrounding rock mass or potential collapse failure can delay the construction process and substantially increase the construction cost [9–11]. Monitoring and predicting the deformation of surrounding rocks in the rock tunnel construction process are of great significance. Site engineers often adopt a comprehensive monitoring scheme to obtain the surrounding rock deformation data of the tunnel and make engineering judgments based on the monitoring data to ensure the construction safety of the rock tunnel [12–14]. Accurate and in-time prediction of the tunnel face deformation can help the site engineers to conduct adequate treatment to maintain the stability of the excavated surrounding rock.

The total station is a widely adopted monitoring tool in tunnel construction used to obtain the surrounding rock deformations. Reflecting targets are implemented in the tunnel linings, and then the site engineers perform regular measurements of the movements of those targets and then determine the deformation values. The frequency of monitoring tunnel deformation via the total station is limited to a few times a day due to the high labour cost and the heavy equipment [15–17]. However, a higher granularity of the deformation data is required to make in-time decisions during the rock tunnel construction process in unfavourable geological conditions. Therefore, an advanced monitoring method should be implemented to obtain higher granularity deformation data so that field engineers can better predict the stability of the exposed tunnel face [18,19]. In practice, the robotic total station (RTS) is often adopted in the tunnel deformation monitoring scheme. Nevertheless, due to the high cost of the RTS equipment and lack of protection measures, the RTS monitoring locations are always at a certain distance away from the excavating tunnel face. Therefore, the ability of the RTS to measure the deformation data and predict the stability near the tunnel face is limited in rock tunnelling construction.

Wireless sensor networks (WSNs) have received much research interest in recent years for their promising potential applications in health monitoring of civil infrastructure, such as in the railway industry [20,21], pipeline inspection [22], and long-span bridges [23]. Sensors in the WSNs, often equipped with micro-electromechanical systems (MEMS) technology, can acquire and stably transmit high granularity monitoring data to the cloud server, which significantly assists the experienced engineers in the office to respond quickly. In recent years, WSNs have been used to monitor tunnel health status in the metro systems of many cities [24–26]. The deformations of the tunnel linings are often monitored in the operation stage to aid in the maintenance of the tunnel. Nevertheless, few studies have shown its application in rock tunnel construction. The underlying reason lies in the harsh construction environment and the dynamic construction process in the rock tunnel site. In addition, monitoring data obtained from the rock tunnel construction practice are often noisy and challenging to interpret. There is a lack of adequate data analytical methods to utilize the noisy monitoring data to predict the possible tunnel face deformation in the coming construction period.

With recent advances in deep learning and data mining techniques, opportunities have opened up for developing a data-driven prediction model for the construction phase of the rock tunnel practice with dynamic construction processes. The main goal of the data-driven model is to provide a practical tool to predict the time sequence deformation responses in the construction process with the site monitoring input. In this context, recurrent neural networks (RNN) and long short-term memory (LSTM) networks are the go-to solutions for sequence data processing in geotechnical practice [27,28], which originate from the problem of natural language processing (NLP). When provided with a data sequence, an RNN processes the first input data in the sequence and feeds the result into the layer that processes the following input. This enables the RNN to keep track of the entire data sequence instead of processing each piece of data separately. LSTM, the successor to RNN, can solve the vanishing gradients problem to some degree and handle more extensive data sequences. However, RNN and LSTM need to process data sequentially, which cannot take advantage of parallel computing hardware and graphics processing units (GPU) in training and prediction. In addition, RNN and LSTM cannot handle long sequences of data and can only capture the relations between inputs at nearby positions/time steps in the data sequence.

Accurate prediction of tunnel face deformation requires consideration of the whole data sequence depending on the rock tunnel construction process. In this regard, the transformer model was introduced in 2017 by a team at Google Brain, a deep learning model that adopts the self-attention mechanism [29]. Compared to RNN or LSTM, the transformer model follows an encoder–decoder structure, which does not rely on recurrence and convolutions to generate an output. The attention mechanism provides context for any position in the input sequence and differentially weights the significance of each part of the input data. Hence, the transformer model makes it possible to process entire sequences in parallel and scale the speed and capacity of sequential deep learning models to unprecedented rates. Although the attention mechanism and the transformer model have been widely adopted in NLP tasks and other sequence data features, to the best of the authors' knowledge, there is limited study on adopting an attention-based deep learning model to perform tunnel deformation predictions in rock tunnelling practice.

This study focuses on tunnel deformations in the construction process of rock tunnel excavation. Wireless tilt and distance sensor (WTDS) nodes were adopted in the WSNs to measure the deformations of excavated surrounding rocks. The sensor locations were carefully selected according to the engineering safety risks. A triangular truss protection cover was designed and deployed on site to ensure the durability of the WTDS in the harsh environment of rock tunnel excavation using the drill and blasting method. This study adopts three deep learning models for predicting the tunnel face deformation based on the monitoring data, namely LSTM, attention-enhanced LSTM, and the transformer model. These three models correlate the deformation of the nearby transverse tunnel section with the tunnel face deformation prediction to provide early warning of the risk of surrounding rock deformation. A tunnel site case study was conducted over three weeks to obtain high-granularity tunnel deformation data from the WTDS nodes. The obtained monitoring data were processed to create a dataset to train and assess the performance of the three deep learning models. Overall, the transformer model provides sound performance to predict the tunnel face deformation. The proposed system has great value as guidance and reference for the construction of rock tunnel projects in complex and unfavourable geological conditions.

2. Rock Tunnel Deformation Monitoring Scheme

2.1. Sensor for Tunnel Deformation Monitoring

In rock tunnel construction, the lining and face deformations of the surrounding rocks are important indicators for site engineers. A typical monitoring scheme includes the horizontal convergence and the crown settlement of a transverse section in front of the rock tunnel excavation. Based on successful monitoring applications in the metro tunnel maintenance, the Omni WTDSs produced by the WISENMESHNET® were adopted in this study to achieve high-granularity monitoring data.

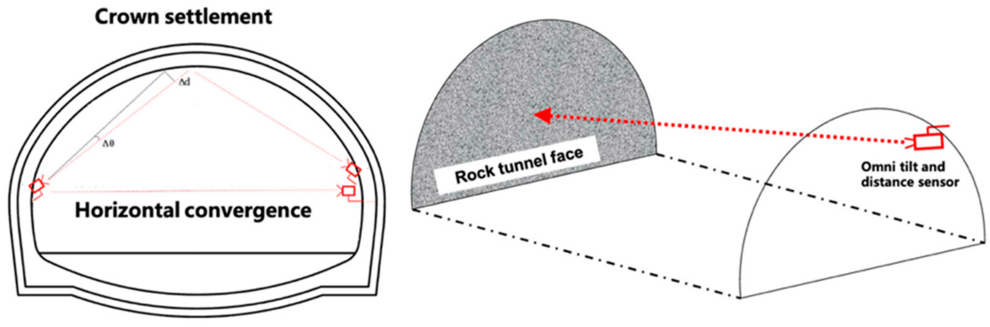
The WTDS is designed to work in a harsh outdoor environment with an IP66-rated enclosure, which has the highest dust-tight protection and the third-highest level of water protection. The WTDS also has high precision during sampling and solid immunity to radio interference. The accuracy and resolution of the sensor are summarized in Table 1. The high-precision laser distance module can accurately measure the distance change between two points of the tunnel structure. The high-precision MEMS Omni tilt sensing module can accurately measure the tilt attached to tunnel linings.

Table 1. The accuracy and resolution of the Omni tilt and distance sensor.

Monitoring item	Accuracy	Resolution
Inclination	0.01°	0.0025°
Distance	0.5 mm	0.1 mm

This sensor can monitor tunnel deformation with high-granularity data and low power consumption. It can operate for more than 12 months with a monitoring frequency of 5 min. Therefore, high-granularity data can be obtained using this type of sensor, and the monitoring and data transmission frequency can be remotely controlled in the cloud-based online platform.

Two sensors are located on one side of the tunnel to monitor the rock tunnel deformation of a typical tunnel transverse section. The other is located on the opposite side (Figure 1). The horizontal convergence and the tunnel crown settlement are measured by interpreting the tilt of the tunnel surface and the distance change between the sensor and the targeting points in the space of the tunnel structure. To monitor the tunnel face deformation, a sensor is placed slightly further from the excavation face, and the laser points at the tunnel face's middle section. The tunnel face deformation can be derived based on the changing distance and the initial shorting angle between the tunnel face plane and the laser beam line.



(**a**) (**b**)

Figure 1. Sensor locations for monitoring deformations in: (**a**) typical tunnel transverse section; (**b**) tunnel face.

In this study, a reading was taken every five minutes to achieve high granularity of the data whilst ensuring a reasonable life span and power consumption. During the installation, all the WTDS nodes should face the same direction, and clear notes must be taken so that a monitored structure's laser distance sensor direction can be correctly interpreted.

It should be noted that movement of the sensor relative to the attached tunnel lining can affect the accuracy of the measurements. Therefore, fixing the sensor to the primary tunnel lining should be relatively rigid to avoid significant movement during the rock tunnel construction. The monitoring data should be regularly analysed to check if the sensor node has considerable movement. During the blasting stage of the tunnel excavation, flying rock masses are likely to hit and destroy the installed sensors. The heavy dust generated during this construction process can also block the laser beam and affect the distance measurement. Therefore, protection measures are required for the sensors, and notable signs should be implemented at the sensor location.

2.2. Protection Cover of the Sensor

A specific cover was designed and implemented to protect the sensors from high-impact collisions during the drill and blasting construction process. The cover is designed with a triangular shape using a 304 stainless steel plate with 2 mm thickness (Figure 2a). The cover size is designed according to the size of the sensor. The protection cover has rubber sheets on the outside surfaces, designed to buffer the impact of flying rock slag after the on-site excavation in the rock tunnel (Figure 2b).

The cover protection's length, width, and height are designed to have enough safety space based on the sensor sizes (100 × 100 × 60 mm). As a result, the installed cover can cover each sensor from the top and the front side towards the excavating tunnel face. In addition to the design, the laser beam can shoot through the gap of the cover truss bar to the target point on the tunnel linings (Figure 3). To have the cover firmly fixed on the primary tunnel lining structure, three bolt holes are distributed along the arc-welded stainless-steel plate. During the deployment of the protection cover, three steel expansion bolts are used to connect the tunnel lining and the cover. Overall, this protective cover can ensure that the monitoring sensor is not damaged during the rock tunnel construction process.

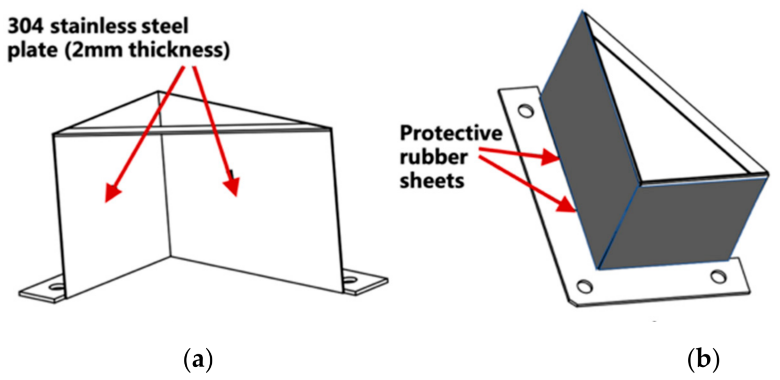


Figure 2. Protective cover for the sensors: (**a**) stainless steel plate of the cover, (**b**) protective rubber sheets on the outer cover surface.

Figure 3. Site photos on the deployment of the sensors and the protective cover: (**a**) cover installation; (**b**) cover and sensors after installation.

2.3. Data Transmission via the WSN

Before any WTDS nodes in the tunnel are switched on, a gateway must be deployed in the tunnel to transmit the monitoring data properly. The gateway should be installed at the tunnel entrance for a typical rock tunnel site to ensure a stable open-air mobile 4G signal source. The stable data transmission coverage area of the gateway can reach up to 300 m. In most cases, there are more than 300 m between the gateway location (next to the tunnel entrance) and the sensor location (nearby the tunnel face). Therefore, it is necessary to install relay sensor nodes to transmit the data.

A typical WSN in the rock tunnel consists of the WTDS nodes, the relay devices, the gateway, and the cloud server. The monitoring data obtained from the sensors are transmitted to the gateway at the entrance through wireless relay devices. Then, the monitoring data are sent to the client through mobile operators and cloud servers. The monitoring frequency can be adjusted according to the site requirements for up to 1 min per reading.

3. Dataset Creation of the Rock Tunnel Construction

3.1. Monitoring Data Structure

The WTDS takes six data dimensions for each measurement: X-axis angle, Y-axis angle, Z-axis angle, laser distance, temperature, and time stamp. The tilt values are based on the orientation of the three-axis angles (Figure 4). In the rock tunnel site, the sensor should be oriented with at least one axis marked on the label parallel to the horizontal plane so that the tilting angle data can be recognized and interpreted. For interpolation purposes, the Z-axis lies along the tunnel lining surface, and the X-axis and Y-axis lie in the plane of the transverse section. Generally, the Z-axis and X-axis tilt values remain stable without notable change during the monitoring time. The tunnel crown settlement and lining convergence values are thus computed using the laser distance data and the Y-axis tilt values. Although the temperature fluctuates at different times of the day, it does not influence the measurements significantly. The timestamp is the recorded data transmission time during the whole monitoring period. Therefore, the deformation data of the rock tunnel can be obtained with a time sequence format.

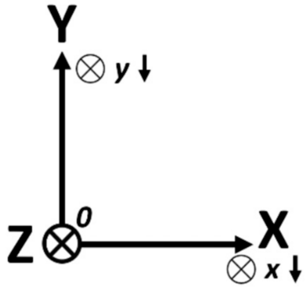


Figure 4. Three axes of the WTDS.

The deformation values monitored at the transverse section slightly away from the excavated tunnel face are processed as the input, which includes the crown settlement and the horizontal convergence. The deformation values at the tunnel face are used as the output. The time series of the monitoring values is computed according to the monitoring frequency. For training of the deep learning model, the amount of data in the training, validation, and testing stages depends on the monitoring period and the granularity of the monitoring data.

3.2. Deformation Computing Method

Three sensor nodes are deployed on the primary tunnel lining in a typical transverse section near to the tunnel face (as shown in Figure 1a). The Y-axis of the sensor's built-in coordinate system is the same as the laser irradiation direction. The tilt angle, β, between the Y-axis and the transverse sectional plane is used to compute the crown settlement and horizontal convergence. The measured laser distance is the real-time distance between the targeting point and the sensor. For the measurements at time i, d_i is the laser distance and β_i is the Y-axis tilt angle. The computed values are S_i (crown settlement) and C_i (horizontal convergence), which can be derived as follows:

$$S_i = d_i \sin \beta_i - d_0 \sin \beta_0 \quad (1)$$

$$C_i = d_i \cos \beta_i - d_0 \cos \beta_0 \quad (2)$$

where the subscript 0 denotes the initial moment for time i, a negative S value indicates the tunnel crown displacing in the direction of gravity, and a negative C value indicates the tunnel lining converging horizontally.

The above computation assumes that the sensors implemented for monitoring are rigidly attached to the tunnel lining to reflect the structural deformation. However, in an actual rock tunnel site during the construction process, the harsh environment near the rock tunnel face is complex. The sensor position and the laser beam are heavily influenced by the impact of the excavation blasting and other construction activities. Those unfavourable conditions can affect the tilt angle and laser distance measurements.

From the site experience, the following scenarios are often encountered while monitoring the rock tunnel construction. Suppose the distance measurement suddenly decreases but returns to an expected value after a short period. In that case, construction personnel or equipment can consider a temporary laser beam blockage. Thus, those monitoring data can be eliminated if the distance data continue to be abnormal. Checking the sensor in person and verifying if dust is blocking the laser sensor or some obstacles in front of the sensor is advised. If the tilt angle of any axis has a significant change at a particular time point, it can be considered as a movement of the sensor, which leads to a significant difference in the measured value. Other scenarios can also significantly affect the measured values. The laser lens can also be blocked by heavy dust and present invalid sampling values or abnormally short-distance values due to the drill and blasting construction activities.

Although the adopted protective cover can mitigate some of those problems, the monitoring data quality is highly influenced by the construction activities. Therefore, it is advised that the site engineers should regularly check the sensor positions and clean the laser lens. In addition, the computed tunnel deformations should be periodically compared with manual inspection and another monitoring method to ensure the reliability of the measurements.

3.3. Data Filtering Method

The computed crown settlement and convergence are reflections of the actual structure deformation, the surrounding construction activities, and the influence of the sensor movement. A filtering method is proposed to quantitatively distinguish and decouple the direction of the measuring point caused by the sensor and the actual deformation of the structure.

When the laser distance values over a short period are "NaN" (not a number), the distance values in this period are replaced as the average before and after these time distance values. A threshold value of 5 mm for the laser distance is set to eliminate sudden changes due to the movement of the sensor or blockage by the construction activities. In addition to laser ranging, the sensor node has a built-in three-axis gravitational accelerometer. Sensor movements can be reflected in the change of tilt in the three axes with a changing tile angle concerning the horizontal plane. Regarding the tilt filtering, a threshold value of 0.01° is selected to eliminate rotational movement of the sensor.

After the first data filtering step, those monitoring data above the set threshold of angle and distance are thus corrected to the values before the sudden change or are eliminated for the disturbance period to reflect the actual structure movement.

To eliminate the slight fluctuations of the tilt angle and the laser ranging, a data process algorithm is further applied after the first filtering step to deal with missing values and outliers in the data. The "3-sigma limits" method includes the data within three standard deviations from a mean and excludes the outlier values. The mean and variance of the settlement and convergence rates for a certain monitoring period are computed to set the upper and lower control limits. As a result, the remaining data are considered the actual tunnel deformation, which can be used to train the deep learning model.

The input and output should be normalized for training and testing the deep learning models with their respective maximum and minimum values. All the normalized values range between zero and one, so the trained deep learning model limits the search in a fixed domain to achieve higher prediction accuracy. The sigmoid activation function is often adopted in deep learning models, which can restrict the output between zero and one. After the model training is completed, the predicted value in the prediction stage is rescaled to obtain the deformation predicted value corresponding to the accurate monitoring interval.

4. Dataset Creation of the Rock Tunnel Construction

This study proposes two attention-based deep learning models to perform time-series prediction on the monitoring data. One adopts the LSTM model, and the other assumes the transformer model. The attention module in the deep learning model generally requires more iterations than the CNN (convolution neural network) model or MLP (multilayer perceptron) model to achieve convergence.

4.1. The Attention Module

The attention mechanism was originally a signal-processing mechanism discovered in the study of human vision and has been applied to time-series modelling with great success in recent years [30]. Recurrent neural networks such as LSTM can effectively establish a time-series dependency on data. Still, this dependency is explicitly transmitted from adjacent units and cannot be used for global time-series dependencies of non-adjacent data. For example, LSTM shows good performance on the continuous deformation monitoring value of the tunnel structure since the deformation of two adjacent moments can establish a strong time-series dependence. Nevertheless, for the monitored deformation values in the rock tunnel construction, the data are not uniform over the period, and two monitoring data points in different construction cycles may have a strong correlation.

Instead of traditional time-series deep learning, the attention module captures the global temporal correlation of input data, which provides a new perspective on temporal modelling. The input of the attention mechanism includes the weight matrices for query (Q), key (K), and value (V), respectively. A softmax function is then applied to normalize the weights. Adopting the widely used dot product attention module, the output can be computed as:

$$\text{Attention}(Q,K,V) = \text{softmax}(\frac{QK^T}{\sqrt{d_k}})V \tag{3}$$

where the weight represents the correlation between the query (Q) and the key (K), and the attention weights are divided by the feature dimension of the model (d_k), which is the square root of the dimension of the key vectors. This operation is used to scale the result of the inner product of Q and K and stabilize gradients during training. The general multi-head attention method is adopted in the attention module, which can map Q, K, and V into a higher dimensional space and apply an attention module in its subspace. As a result, the attention module can combine the results of each subspace and then map the output.

4.2. Attention-Enhanced LSTM Model

The LSTM model has a unique input gate, forget gate, output gate structure, and serial hidden units, which can capture the correlation of long-distance time-series data and predict the future. Nevertheless, when a long time series contains many time steps, the data correlation between distant time steps is difficult for the model to learn. In this study, an attention module is added to the original structure of the LSTM to improve the performance of learning long-distance data correlation. As shown in Figure 5, a multi-layer LSTM model structure is constructed to extract the multiple hidden states of the output with respect to the input deformation sequence data. A dot product self-attention module is applied to consider all the outputs of the hidden states from the last LSTM layers. The outputs of the attention module are then processed through the linear layer and sigmoid activation function to obtain the final regression output prediction with the same sequence length as the input.

4.3. Transformer Model Structure

The constructed transformer model has two encoder layers and two decoder layers (Figure 6). The sequence length of the input encoder is a time series composed of measurement points with two dimensions (the crown settlement and the horizontal convergence of a transverse section). The predicted value output by the decoder has the same measurement points as the inputs with one dimension, which is a time sequence consisting of the monitoring deformation values at the tunnel face.

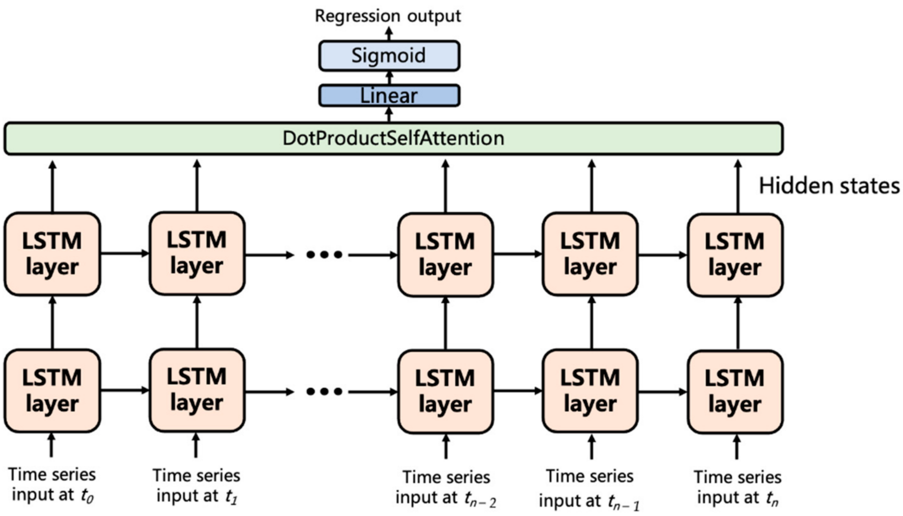


Figure 5. Model structure of the attention-enhanced LSTM.

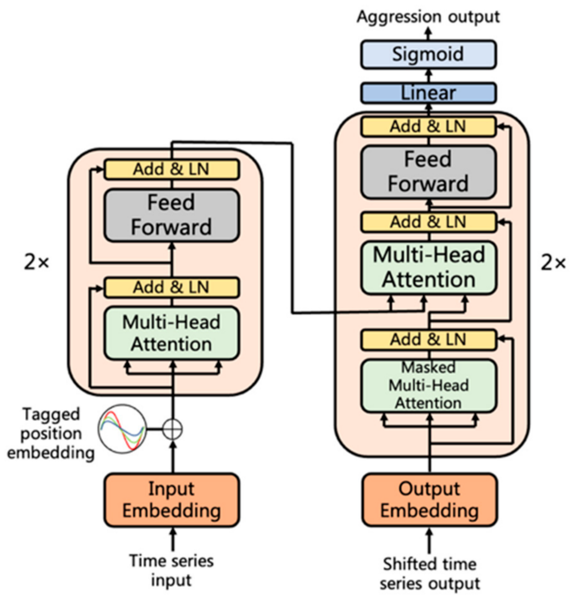


Figure 6. Model structure of the transformer.

To let the encoder layer learn the rich data information in the input sequence, the transformer model upgrades the features of the input sequence from a low dimension to a high dimension in the input embedding. In this study, the parameter dimension of the input time-series multi-head attention block is upgraded from 2 to 18, and the sequence length is selected as 30. The number of hidden layer units in the feed-forward network module (FFN) is 10. Since the feature dimension is not large, the number of heads in the multi-head attention mechanism is set to three.

The decoder layer has a similar structure to the encoder layer, which has an output dimension of one. Using the same approach as shown in the attention-enhanced LSTM, the outputs of the decoder layer are then processed through the linear layer and sigmoid activation function. The output prediction should have the same sequence length as the input.

An essential feature in the constructed transformer model is the positional embedding, widely used in the BERT (Bidirectional Encoder Representations from Transformers) model [31,32] for pre-training bidirectional representations for language understanding. The positional embedding is designed and placed before the encoder but not included in the decoder stage. The positional embedding assigns tags denoting the different process stages of rock tunnel construction, which can adjust the positional encoding amplitude at various construction stages. As a result, the proposed transformer model can incorporate the process information into the model's input and hide the process information from the output.

This study's proposed positional tags represent different process stages in rock tunnel construction. For ease of recognition, the excavation–slagging–stand stage is represented by the tag value of one, the advanced support stage is represented by the tag value of two, and the shotcrete lining stage is represented by the tag value of three. The site engineer manually records each construction stage period so that the monitoring data can align with the reordered time to assign the positional tags.

The computed inputs (crown settlements and the horizontal convergence values) corresponding to each construction stage for a specific monitoring period are illustrated in Figure 7 with different tag values. It can be seen that the assigned tags can differentiate the deformation values with similar magnitude, which informs the transformer model with additional site information. Hence, the transformer model can know the construction process, whereas the attention-enhanced LSTM model only learns the data features from the monitoring sensors.

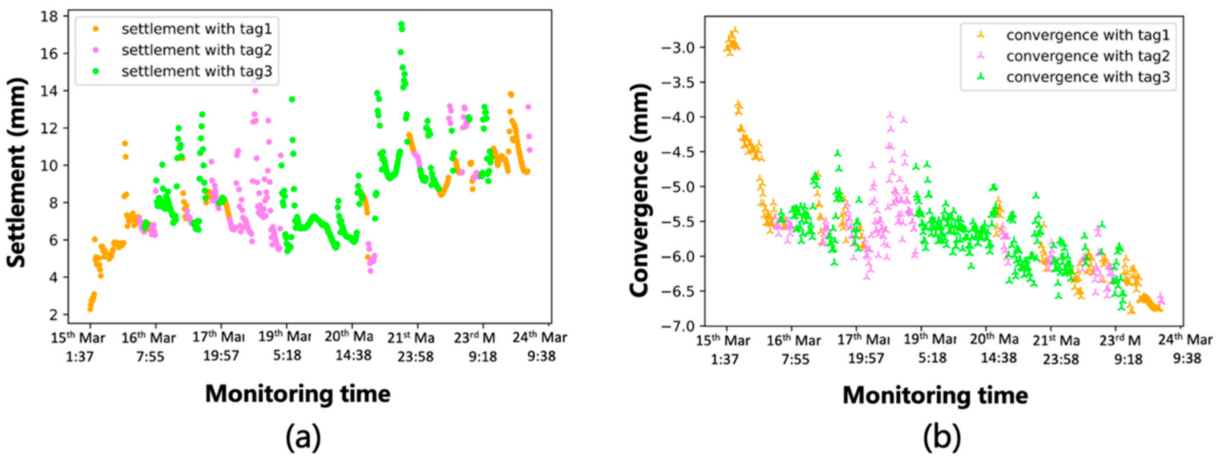


Figure 7. The positional tags of (**a**) the crown settlements and (**b**) the horizontal convergence values corresponding to different construction stages.

Based on the assigned positional tags, the amplitude of the positional encoding should be modified to fit the application before the encoder model. By adjusting the amplitude, the position tags of the input data are integrated to preserve the process information in a long time series. In this study, the positional encoding value can be computed as follows:

$$\mathrm{PE}(pos, 2i) = tag \times \sin\left(\frac{pos}{1000^{2i/d_k}}\right) \quad (4)$$

$$\mathrm{PE}(pos, 2i+1) = tag \times \cos\left(\frac{pos}{1000^{2i/d_k}}\right) \quad (5)$$

where the *tag* value corresponds to the construction stages mentioned above, *pos* is the position of the data in the entire input sequence at a specific moment, i represents the dimension, and d_k is the feature dimension of the model. Position embedding applies sine and cosine functions at odd and even positions to identify the feature difference at different dimensions.

5. Site Experiment Results

5.1. Project Background

The site study is based on the construction of the Tongluoshan tunnel of the Chongqing Rail Transit Line 15 project, Chongqing, China. The total length of the tunnel is 4031 m.

It is designed with a tunnel span of 17.4–12.9 m and a tunnel height of 11.5–12.8 m. The designed longitudinal slope of this tunnel is a “herringbone” slope with a slope gradient of 1.0% to 2.3% and to −0.6%.

The Tongluoshan tunnel has a variety of unfavourable geological conditions, including methane gas, coal seams, dissolved depressions, karst caves, dissolved troughs, water-rich fault fracture zones, and shallowly buried silt (backfill areas). Among those challenging conditions, one of the most significant concerns is the tunnel face stability in the shallow buried section with a minimum burial depth of about 5 m and a grade V surrounding rock condition (Figure 8). The green lines in Figure 8 represent the top and bottom outlines of the tunnel excavation, and the red line is the tunnel floor surface for laying the rail track. In the section YK76 + 410~YK76 + 490, the tunnel crown is in the strongly weathered rock layer and backfill. The main tunnel section is in moderately weathered rock formations.

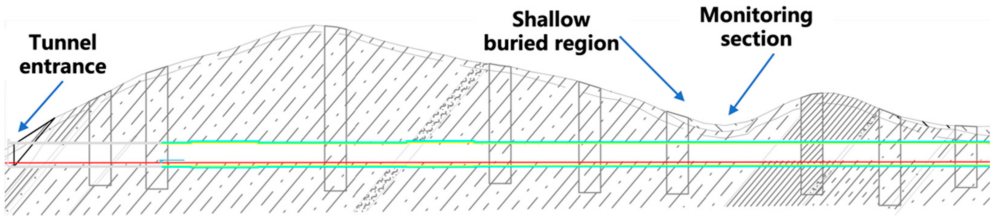


Figure 8. The geological layout of the Tongluoshan mountain at the shallowly buried tunnel region.

This study implemented the monitoring systems in this shallow buried region to mitigate the construction risks. It should be noted that this section was excavated by the CD method, and the upper pilot hole was excavated strictly by non-blasting excavation. The tunnel was constructed from the shallowly buried region using the drill and blasting method.

5.2. Tunnel Monitoring Scheme

The longitudinal arrangement of sensors of the WSN monitoring system is shown in Figure 9. To monitor the tunnel deformation during construction in a shallow buried region, four WTDS nodes were deployed, three of which monitored the deformation of a transverse section near the tunnel face, and one sensor monitored the face deformation. The sensor locations in the transverse section were based on the arrangement in Figure 1a. The sensor for monitoring the tunnel face deformation was based on the configuration in Figure 1b.

Based on the layout of the three sensors in monitoring Section 1, the horizontal convergence is measured by one of the sensors. The other two sensors were pointed towards a reflection target at the tunnel crown of this transverse section to obtain the settlement deformation. To facilitate installation and maintenance of the sensors, they were installed at about 2 m above the excavated invert so that the site engineers could notice the sensors without accidental contact. When the CD method was used for construction, the three sensors were installed on the same side of the tunnel lining according to the on-site construction conditions.

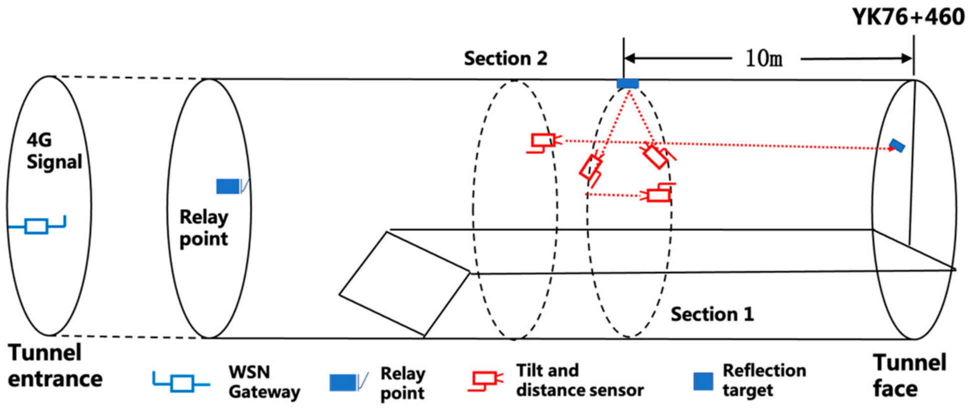


Figure 9. Longitudinal schematic diagram of the sensor layout in each section.

Monitoring Section 2 was implemented to monitor the deformation of the tunnel face. The installation and layout position of the sensor in Section 2 needed to be adjusted according to the site conditions. In principle, it should be as close to monitoring Section 1 as possible and should not be more than 30 m away from the tunnel face. If conditions permit, it should be arranged at the upper half of the transverse section to ensure that the front interference does not affect the laser sensor. When the distance between the excavation tunnel face and monitoring Section 1 exceeds 20 m, it is necessary to arrange on-site personnel for dynamic installation and layout. The basic installation and layout method is to move the monitoring sensors of Section 1 forward to about 10 m from the face and move the sensor in Section 2 accordingly.

Since the tunnel face of the monitoring region is about 500 m away from the tunnel entrance, a relay device was implemented at the midpoint between the tunnel face and the entrance (about 250 m away from the tunnel entrance), which could transmit the monitoring data to the gateway. A gateway was installed at the entrance to send the data to the cloud server. The monitoring time started on 2 March 2022 and lasted for three weeks until the tunnel excavation passed beyond the shallow buried region and stopped on 24 March 2022. The monitoring data were set to record once every 5 min.

Based on the monitoring data in Section 1, the tilt and distance of the two sensors pointing towards the tunnel crown can be obtained. The settlement over the three weeks can be obtained using the data filtering method introduced, as shown in Figure 10a. The horizontal convergence of the transverse section can be similarly obtained (Figure 10b). It should be noted that the crown settlement presented in Figure 10 has eliminated the significant movement of the sensors. The horizontal convergence has eliminated the outlier values due to the blockage of the laser beam caused by construction equipment, slagging, and site personnel.

Based on the monitoring data in Section 2, the tunnel face deformations are computed and plotted in Figure 10c. Due to the dynamic construction process, the tunnel face distance from monitoring Section 2 changes over different construction cycles. The obtained monitoring data are filtered to reflect the actual movement of the tunnel face in each construction cycle and eliminate the extended distances caused by the excavation. As shown in Figure 10, the filtered monitoring data of crown settlement and horizontal convergence agree well with the total station measurements. They can reflect the dynamic changing of deformations during tunnel construction. Nevertheless, there is no measured data on the tunnel face deformation from the total station, and it is hard to discover the correlation between the measured deformations between the selected transverse section (Figure 10a,b) and the tunnel face (Figure 10c).

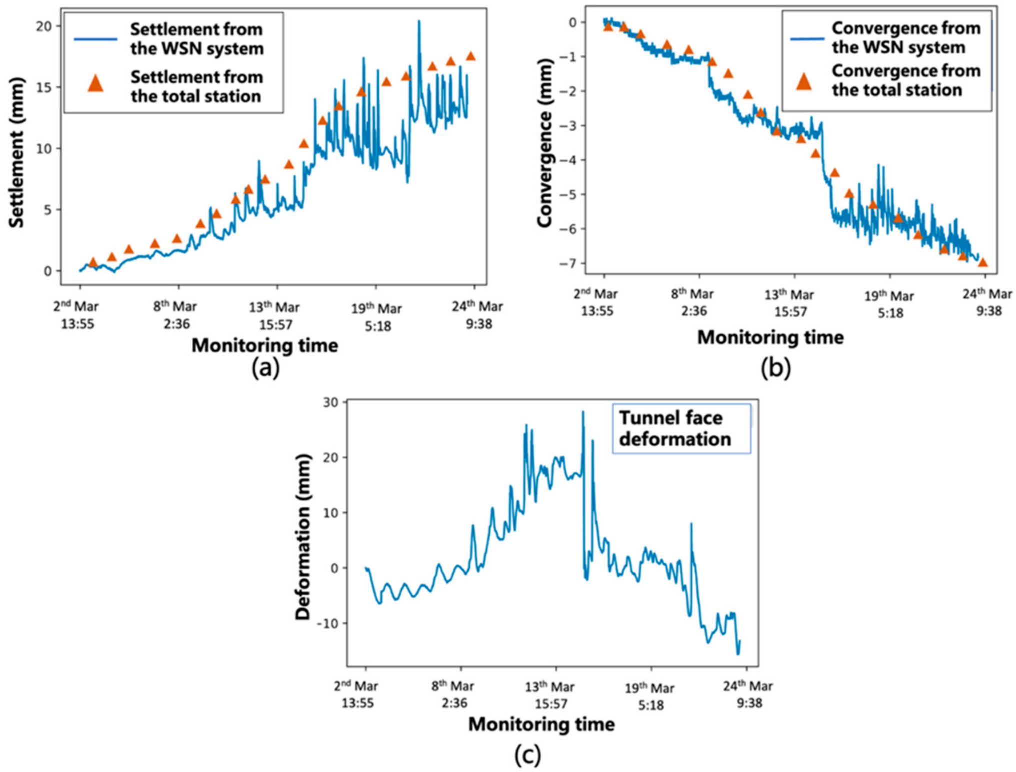


Figure 10. Monitored deformation data of the transverse section (Section 1) in front of the tunnel face: (**a**) crown settlement; (**b**) horizontal convergence; (**c**) tunnel face deformation.

5.3. Training and Testing Process

The input and output data are divided into 80% training data and 20% testing data. At the same time, the training part is randomly split into a training set and validation set with a ratio of 9:1. The loss function adopts the mean square error (MSE) function, and the teacher-forcing method is used to speed up the training of the transformer model. The Adam optimizer is adopted in the training process. The initial learning rate is 1×10^{-4}, and the learning rate decay strategy adopted reduces to a plateau (patience = 10). The three deep learning models are trained on the same desktop with an Intel Core i7-9700k CPU and an Nvidia RTX3050 GPU with 4 GB of memory. The training software environment is configured with PyTorch framework on the Ubuntu 22.04 system LTS.

Three deep learning models were adopted in this site experiment to compare the performance of different models for the prediction of tunnel face deformation. Because of the optimum model structures, a grid search method was applied to determine the optimal transformer parameters and model architecture with the MSE on the validation set as the criterion. All three models had the same input dimension of two (input_size) and output dimension of one (output_size). The model structure parameters that were strongly related to the performance of the model were the number of layers (num_layers), the hidden layer feature dimension (hidden_size) of the feed-forward network (FFN), the time steps (max_len) of the input time series, and the data feature dimension (model_dim) after embedding in the transformer.

Nine different model structures of the transformer model and the corresponding RMSE values of the testing data are shown in Table 2. Based on this grid search approach, the best model structure among the nine cases was determined based on the RMSE value for the testing data. It can be seen that the increased layers, model dimension, hidden size, and length of the input time steps can enhance the model performance. However, as shown in Case 9, too complex a model structure can cause overfitting and less optimal prediction performance. Therefore, the model structure in Case 8 was selected.

Table 2. Transformer parameter tuning and final model selection based on validation data.

Case No.	Num_Layers	Model_Dim	Hidden_Size	Max_Len	Validation Data RMSE/mm
1	1	6	6	6	2.786
2	1	6	6	8	0.909
3	1	12	10	9	1.053
4	1	12	12	9	0.868
5	2	6	6	6	1.248
6	2	6	6	9	0.943
7	2	6	6	24	0.954
8	2	18	10	30	0.784
9	3	24	18	48	0.832

Using this adapted model structure, the loss curves of the transformer model training and validation are shown in Figure 11. After 20,000 epochs, the transformer model tends to converge, and the loss of the validation set drops to 0.0008. Based on the optimized structure of the transformer model, both LSTM models adopted similar model structures for performance comparison (Table 3).

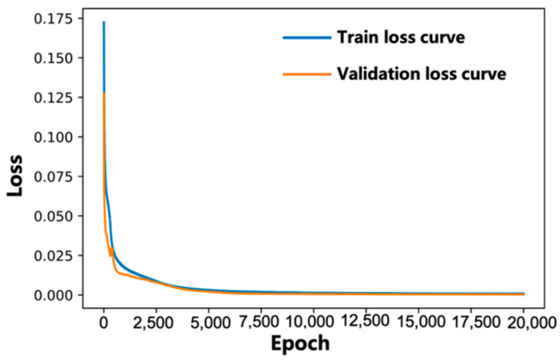


Figure 11. The training and validation loss curves of the transformer model.

Table 3. The model structures of different deep learning models.

Model	Num_Layers	Model_Dim	Input_Size	Hidden_Size	Output_Size	Max_Len
LSTM	2	30	2	10	1	30
Attention-enhanced LSTM	2	\	2	10	1	30
Transformer	2	\	2	10	1	30

5.4. Deformation Prediction Comparison

Three deep learning models were trained and tested based on the adopted model structure settings in Table 3. The quantitative measured RMSE values in the validation and testing stages of the three deep learning models are shown in Table 4. The transformer model had the lowest RMSE value and outperformed the attention-enhanced LSTM model and the LSTM model. The two LSTM models have similar RMSE values, whereas the attention-enhanced LSTM gives a slightly smaller RMSE value. The attention module included in the attention-enhanced LSTM scored the hidden state better than the standard LSTM. The attention module can automatically increase the weights and significance of more relevant data inputs in the time series of the monitoring data, and the model prediction accuracy was slightly improved from the LSTM model (Table 4).

Table 4. Accuracy comparison of LSTM, attention-enhanced LSTM, and transformer with a similar model structure.

Model	Validation Data RMSE/mm	Testing Data RMSE/mm
LSTM	2.035	8.120
Attention-enhanced LSTM	1.942	7.135
Transformer	0.784	3.810

The prediction performance of the three models is shown in Figure 12. Both LSTM models showed poor predictions at both the training and testing stages. The LSTM model does not capture the trend of the tunnel face deformation. The results suggest that the LSTM model has a problem with gradient disappearance for the time-series prediction of long sequences. Compared to the LSTM model, the attention-enhanced LSTM model automatically evaluates the scores in the hidden layer units, highlighting the modelling effect of the most appropriate time steps and thus improving the prediction of the prediction stage. Although the correlation of long-distance time sequence data can be trained with the additional attentional module in the attention-enhanced LSTM model, it still suffers from the gradient vanishing problem.

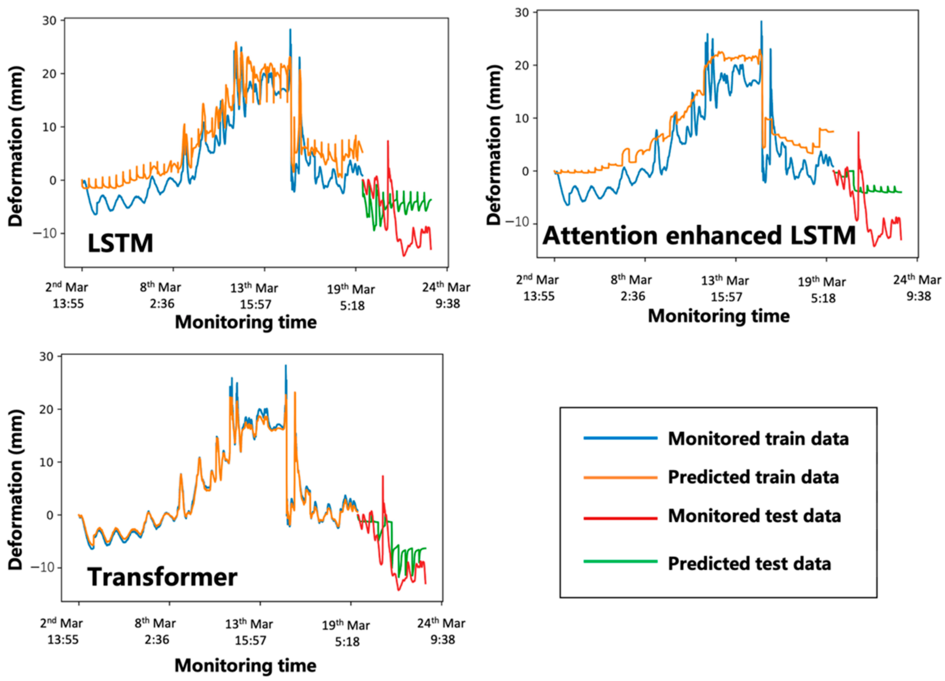


Figure 12. The training and testing results of the LSTM model, the attention-enhanced LSTM model, and the transformer model.

Compared to the two LSTM models, the transformer model showed good prediction performance at both the training and testing stages (Figure 12). The transformer model expands the feature dimension through embedding and extracts the correlation of time-series data in the higher-dimensional feature space through the multi-head attention mechanism. In addition, the transformer model fully considers the potential similarity of the deformation data sequence under different processes through the positional embedding. The encoder structure in the transformer fully extracts time sequence information in high-dimensional space, and the decoder structure has a better decoding and mapping function than the sliding window mechanism in the LSTM. The transformer can extract more information from the time sequence data than the attention-enhanced LSTM and LSTM.

It should be noted that the teacher forcing used in the transformer model training phase improves the speed of model training. However, it is not suitable for use in the testing phase because the positional embedding value needs to be input into the decoder. In practice, the positional coding value in the testing stage remains unknown. Therefore, the auto-aggression method was adopted to generate the transformer model's prediction output. Due to the influence of the construction activities in the tunnel, the monitoring value has errors, but it can still reflect the trend of tunnel face deformation. The predicted value by the transformer model can accurately capture the same trend as the monitored deformation measurements.

6. Conclusions

It is vital to grasp the tunnel face deformation information for rock tunnels using the drilling and blasting construction method. This study implemented a WSN, including gateway transmission, relay point, and sensor nodes, to obtain high-granularity deformation data during construction. A transformer model was proposed, which considers the construction sequence into the positional bedding and has an attention module to deep learn the high dimensionality correlation between the nearby deformation data and the tunnel face deformation. An attention-enhanced LSTM model and the LSTM model were also constructed to predict the tunnel face deformation and compare it to the performance of the transformer model.

A site study conducted on a shallow buried tunnel section suggested good performance of the proposed WSN system. The transformer model showed the best performance in terms of the model prediction results, whereas the attention-enhanced LSTM and LSTM models showed less accurate prediction results. The prediction results suggest that the transformer can extract more information from the time sequence data than the attention-enhanced LSTM and LSTM. Instead of accurately predicting the deformation value, the proposed approach can capture the dynamic changing of deformations during tunnel construction, allowing the site engineers to determine the possible risks involved in the rock tunnel construction process.

It should be noted that the deformation monitoring data during the construction phase only constitute part of the site conditions. The tunnel design specifications, the embedding depth, and the surrounding rock conditions should also be considered to understand the tunnel risks. Therefore, computer vision-based techniques can be adopted in future studies to obtain rich features of the surrounding rock mass and identify the discontinuities, weak interlayers, and water leakages in the excavated sections. With multisource data input, predicting the tunnel face deformation and the potential risks associated with the construction process can become more accurate and reliable.

Author Contributions: Conceptualization, M.Z., Z.X., C.N. and K.F.; data curation, M.Z., C.N. and Z.S.; formal analysis, M.Z. and C.N.; funding acquisition, M.Z., Z.X. and K.F.; investigation, M.Z. and C.N.; methodology, M.Z., Z.X., C.N., B.H. and K.F.; project administration, M.Z., Z.X. and K.F.; resources, M.Z., Z.X., B.H. and K.F.; software, M.Z. and C.N.; supervision, M.Z., Z.X. and K.F.; validation, M.Z., C.N., Z.S. and K.F.; visualization, M.Z., C.N. and Z.S.; writing—original draft, M.Z., Z.X., C.N., Z.S., B.H. and K.F.; writing—review and editing, M.Z., Z.X., C.N., Z.S., B.H. and K.F. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the China State Construction Railway Investment & Engineering Group, grant number CSCECZJTT-2021-07.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: The financial support is gratefully acknowledged. In addition, we also thank Jinghan Chang, Jianhong Man, Qihao Jiang, and Linghan Ouyang from Tongji University, China, for their valuable works in the monitoring data collection of this paper.

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

References

1. Zhao, Z.; Cao, Q. The development of urban underground space from the perspective of urban economy. *Procedia Eng.* **2011**, *21*, 767–770. [CrossRef]
2. Vähäaho, I. Underground space planning in Helsinki. *J. Rock Mech. Geotech. Eng.* **2014**, *6*, 387–398. [CrossRef]
3. Rostami, J. Performance prediction of hard rock Tunnel Boring Machines (TBMs) in difficult ground. *Tunn. Undergr. Space Technol.* **2016**, *57*, 173–182. [CrossRef]
4. Ebrahim, F. Uniformly distributed lace design for hard rock TBMs. *Tunn. Undergr. Space Technol.* **2021**, *110*, 103829.
5. Liu, J.K.; Jiang, Y.J. Optimized ANN model for predicting rock mass quality ahead of tunnel face using measure-while-drilling data. *Bull. Eng. Geol. Environ.* **2021**, *80*, 320–336. [CrossRef]
6. Bo, Y.; Liu, Q.; Huang, X.; Pan, Y. Real-time hard-rock tunnel prediction model for rock mass classification using CatBoost integrated with Sequential Model-Based Optimization. *Tunn. Undergr. Space Technol.* **2022**, *124*, 104448. [CrossRef]
7. Xue, Y.; Liu, H.; Bai, C.; Su, M.; Qiu, D.; Zhou, B.; Yu, Y.; Jiang, X. Extension prediction model of soft rock tunnel deformation grade based on entropy weight method and rough set. *Environ. Earth Sci.* **2022**, *81*, 24. [CrossRef]
8. Lai, J.; Wang, X.; Qiu, J.; Chen, J.; Hu, Z.; Wang, H. Extreme deformation characteristics and countermeasures for a tunnel in difficult grounds in southern Shaanxi, China. *Environ. Earth Sci.* **2018**, *77*, 706. [CrossRef]
9. Wang, Y.Q.; Chang, H.T.; Wang, J.Y.; Shi, X.L.; Qiu, J.L. Countermeasures to treat collapse during the construction of road tunnel in fault zone: A case study from the Yezhuping Tunnel in south Qinling, China. *Environ. Earth Sci.* **2019**, *78*, 16. [CrossRef]
10. Hencher, S.R. The Glendoe Tunnel Collapse in Scotland. *Rock Mech. Rock Eng.* **2019**, *52*, 4033–4055. [CrossRef]
11. Tanii, K.; Hirota, M.; Kikuchi, Y.; Kamatani, N.; Takahashi, H. Design and construction of countermeasure for record-breaking land collapse. *Proc. Tunn. Eng. JSCE* **2002**, *12*, 297–302. [CrossRef]
12. Gao, S.M.; Chen, J.P.; Zuo, C.Q.; Wang, W.; Sun, Y. Structure optimization for the support system in soft rock tunnel based on numerical analysis and field monitoring. *Geotech. Geol. Eng.* **2016**, *34*, 1089–1099. [CrossRef]
13. Qiao, S.; Cai, Z.; Tan, J.; Xu, P.; Zhang, Y. Analysis of collapse mechanism and treatment evaluation of a deeply buried hard rock tunnel. *Appl. Sci.* **2020**, *10*, 4294. [CrossRef]
14. Zhu, G.Q.; Feng, X.T.; Pan, P.Z.; Zhou, Y.Y.; Yang, C.X.; Li, Z.W.; Taiwakuli, Y. Real-time monitoring of the development of brittle fracture in hard rock tunnels based on physical model test. *Tunn. Undergr. Space Technol.* **2022**, *119*, 104240. [CrossRef]
15. Wang, F.; Yin, S.; Guo, A.; Wang, Z.; Mi, M.; Qi, G.; Ma, J.; Zhang, H. Frame structure and engineering applications of the multisource system cloud service platform of monitoring of the soft rock tunnel. *Geofluids* **2021**, *2021*, 6672732. [CrossRef]
16. Luo, Y.; Chen, J.; Xi, W.; Zhao, P.; Qiao, X.; Deng, X.; Liu, Q. Analysis of tunnel displacement accuracy with total station. *Measurement* **2016**, *83*, 29–37. [CrossRef]
17. Zhou, M.; Shadabfar, M.; Xue, Y.; Zhang, Y.; Huang, H. Probabilistic analysis of tunnel roof deflection under sequential excavation using ANN-based Monte Carlo simulation and simplified reliability approach. *ASCE-ASME J. Risk Uncertain. Eng. Syst. Part A Civ. Eng.* **2021**, *7*, 04021043. [CrossRef]
18. Luo, Y.; Chen, J.; Xi, W.; Zhao, P.; Li, J.; Qiao, X.; Liu, Q. Application of a total station with RDM to monitor tunnel displacement. *J. Perform. Constr. Facil.* **2017**, *31*, 04017030. [CrossRef]
19. Zhou, Z.; Zhang, J.; Gong, C. Automatic detection method of tunnel lining multi-defects via an enhanced You Only Look Once v4 network, *Comput-Aided. Civ. Inf.* **2022**, *37*, 762–780.
20. Hodge, V.J.; O'Keefe, S.; Weeks, M.; Moulds, A. Wireless sensor networks for condition monitoring in the railway industry: A survey. *IEEE Trans. Intell. Transp. Syst.* **2014**, *16*, 1088–1106. [CrossRef]
21. Huang, Z.K.; Zhang, D.M.; Pitilakis, K.; Tsinidis, G.; Huang, H.W.; Zhang, D.M.; Argyroudis, S. Resilience assessment of tunnels: Framework and application for tunnels in alluvial deposits exposed to seismic hazard. *Soil Dyn. Earthq. Eng.* **2022**, *162*, 107456. [CrossRef]
22. Meribout, M. A wireless sensor network-based infrastructure for real-time and online pipeline inspection. *IEEE Sens. J.* **2011**, *11*, 2966–2972. [CrossRef]
23. Kim, S.; Pakzad, S.; Culler, D.; Demmel, J.; Fenves, G.; Glaser, S.; Turon, M. Health monitoring of civil infrastructures using wireless sensor networks. In Proceedings of the 6th International Conference on Information Processing in Sensor Networks, Cambridge, MA, USA, 25–27 April 2007; pp. 254–263.
24. Hoult, N.; Bennett, P.J.; Stoianov, I.; Fidler, P.; Maksimović, Č.; Middleton, C.; Graham, N.; Soga, K. Wireless sensor networks: Creating 'smart infrastructure'. *Proc. Inst. Civ. Eng. Civ. Eng.* **2009**, *162*, 136–143. [CrossRef]
25. Hang, Y.; Luo, Y.; Lai, H.; Huang, Y. A new localisation strategy with wireless sensor networks for tunnel space model. *Int. J. High Perform. Comput. Netw.* **2019**, *14*, 249–256. [CrossRef]
26. Zhang, J.Z.; Huang, H.W.; Zhang, D.M.; Zhou, M.L.; Tang, C.; Liu, D.J. Effect of ground surface surcharge on deformational performance of tunnel in spatially variable soil. *Comput. Geotech.* **2021**, *136*, 104229. [CrossRef]
27. Zhang, W.; Li, H.; Li, Y.; Liu, H.; Chen, Y.; Ding, X. Application of deep learning algorithms in geotechnical engineering: A short critical review. *Artif. Intell. Rev.* **2021**, *54*, 5633–5673. [CrossRef]
28. Xie, J.; Huang, J.; Zeng, C.; Jiang, S.H.; Podlich, N. Systematic literature review on data-driven models for predictive maintenance of railway track: Implications in geotechnical engineering. *Geosciences* **2020**, *10*, 425. [CrossRef]

29. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, Ł.; Polosukhin, I. Attention is all you need. In Proceedings of the 31st Conference on Neural Information Processing Systems (NIPS 2017), Long Beach, CA, USA, 4–9 December 2017.
30. Niu, Z.; Zhong, G.; Yu, H. A review on the attention mechanism of deep learning. *Neurocomputing* **2021**, *452*, 48–62. [CrossRef]
31. Devlin, J.; Chang, M.W.; Lee, K.; Toutanova, K. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv* **2018**, arXiv:1810.04805.
32. Acheampong, F.A.; Nunoo-Mensah, H.; Chen, W. Transformer models for text-based emotion detection: A review of BERT-based approaches. *Artif. Intell. Rev.* **2021**, *54*, 5789–5829. [CrossRef]