Prediction of dam horizontal displacement based on CNN-LSTM and attention mechanism

Yuhao Liu*, Xiao Feng

School of Civil Engineering, Chongqing Jiaotong University, Chongqing, China
**Corresponding author: 1457251985@qq.com*

Abstract: *Aiming at the characteristics of dam safety monitoring data sequence with few samples, short sequence and nonlinearity, a dam horizontal displacement prediction method based on attention mechanism, convolutional neural network (CNN) and long short-term memory (LSTM) is proposed. This method can reduce the loss of historical information and improve the prediction accuracy. First, the missing values are supplemented by linear interpolation to improve the integrity of the data. Then the abstract feature data extracted by CNN is mapped to the predicted value of LSTM, and then optimized through attention mechanism. Finally, the model is trained and verified with the monitoring data of a concrete gravity dam in Chongqing as a sample. Experimental results show that the root mean square error (RMSE), mean absolute percentage error (MAPE) and fit (R^2) of the CNN-LSTM hybrid model based on attention mechanism are 0.3882, 0.7121% and 0.9543, respectively. The prediction accuracy of the new model is better than the CNN-LSTM model and the LSTM neural network model.*

Keywords: *Horizontal displacement prediction of dam, CNN-LSTM, Attention mechanism, Linear interpolation method.*

1. Introduction

Establishing a mathematical monitoring model based on measured data is an important task of dam safety monitoring and analysis. The dam safety monitoring model is used as a mathematical expression describing the nonlinear mapping relationship between dam deformation and causes, and can be used to evaluate and predict the time-varying state of the dam structure during use. In dam safety monitoring, multiple regression models and time series models are commonly used to describe and predict the magnitude of dam effects [1]. The classic time series forecasting model requires the data to have a certain degree of stationarity and linear correlation, so it cannot deal with nonlinear problems. With the rise of artificial intelligence, many traditional machine learning models have been applied to dam deformation prediction. In recent years, some signal processing and artificial intelligence methods have been widely used in dam safety prediction modeling. Such as neural network, wavelet network, empirical mode decomposition, fractal, support vector machine (SVM) [2-4], etc., but traditional machine learning models are difficult to adapt to the big data environment. Deep learning is a frontier field of artificial intelligence and has been widely used in time series data forecasting in recent years. LSTM mode is a special RNN prediction mode with powerful information capture and storage functions [5]. The prediction model of a single algorithm usually only contains part of the information in the data, and the prediction model based on the hybrid algorithm can more fully extract effective information from the data, reduce the parameters caused by the prediction error, improve the prediction efficiency, and the accuracy of the prediction [6-7]. Because the water level data has a certain periodicity, and the current dam deformation data has a certain correlation with the past data. Therefore, in order to make full use of the time series of dam detection data and improve the accuracy of dam horizontal displacement prediction, a CNN-LSTM hybrid model based on attention mechanism is proposed. Taking a concrete gravity dam in Chongqing as an example, the prediction accuracy of the model is improved and the validity of the model is verified.

2. Build experimental models

2.1 CNN model

CNN model is generally divided into five parts, which are input layer, convolution layer, activation

ISSN 2663-1563 Vol. 3, Issue 1: 14-19, DOI: 10.25236/AJAGE.2021.030103

function layer, pooling layer and full connection layer. The convolution layer and the pooling layer are specially designed data processing layers, which are used to filter the input data and extract useful information. The activation layer makes the output feature non-linear mapping. The pooling layer sieves the features, extracts the most representative features, and reduces the dimension of the features [8].The full connection layer will summarize the learned features and map them into two-dimensional feature output. The convolution formula is:

$$T(i, j) = \sum_{k=1}^{n} (X_k W_k)(i, j) \tag{1}$$

In Formula 1, T is the characteristic sequence after the volume; N is the length of input data X_k, W_k is the kernel sequence of convolution [9].

2.2 LSTM model

Long Short Term Memory networks (LSTM), a kind of neural network with memory function, is a variant of RNN. LSTM has excellent performance in temporal data processing, and has been widely used in natural language processing and other fields. LSTM uses input gate, output gate and forgetting gate to control the information [10]. A single LSTM neuron is shown in Figure 1, in which the activation function Sigmoid is represented, and the tanh function resizes the value. Output range is [-1,1].

Forget gate is used to control whether the state of the previous moment is retained to the current state of the neuron to realize the screening of memory. Input the state value of the previous moment and the current input value into the activation function Sigmoid to get an importance value to determine the information update, and then through the tanh function to process the state value of the previous moment and the input information to get the candidate cell state. The output gate controls the final output of the unit state [11]. The unit state is filtered by the output gate and compressed by the tanh function to obtain the final output of the unit.The calculation details of LSTM are as follows:

Forgetting gate can be described as shown in Equation (2):

$$f_t = sigmoid(w_f[h_t, x_t] + b_f) \tag{2}$$

The input gate can be described as Equation (3-4):

$$i_t = sigmoid(w_i[h_{t-1}, x_t] + b_i) \tag{3}$$

$$C_t^* = \tanh(w_c[h_{t-1}, x_t] + b_c) \tag{4}$$

The information conduction C_t is shown in Equation (5):

$$C_t = C_{t-1} f_t + i_t C_t \tag{5}$$

Output gate O_t is shown in Equations (6-7):

$$O_t = sigmoid(w_o[h_{t-1}, x_t] + b_o) \tag{6}$$

$$h_t = O_t^* \tanh(C_t) \tag{7}$$

In Formula 1-6, w_f, w_i, w_c, and w_o are weight indexes, b_f, b_i, b_c, and b_o are offset. σ is the Logistic Sigmoid function; f_t, i_t and O_t respectively represent the state of the forget gate at time t, the output state of the input gate and the output gate [12]; C_t represents the state of the memory unit at time t. C_t^* represents the candidate memory unit state at time t.

2.3 Attention mechanism

Attention mechanism is a simulation of the human brain's attention resource allocation mechanism, the human brain at a particular moment will focus on the need to focus on the area, reduce or even ignore the attention to other areas, in order to get more the need to pay attention to details, inhibit other useless information, ignore irrelevant information and enlarge the required information [13]. By means of probability distribution, attention mechanism gives enough attention to key information and highlights the influence of important information, so as to improve the accuracy of the model. Attention mechanism can effectively improve the situation that LSTM loses information due to too long sequence, and at the same time, it replaces the original random weight distribution method with probability distribution

ISSN 2663-1563 Vol. 3, Issue 1: 14-19, DOI: 10.25236/AJAGE.2021.030103

method.

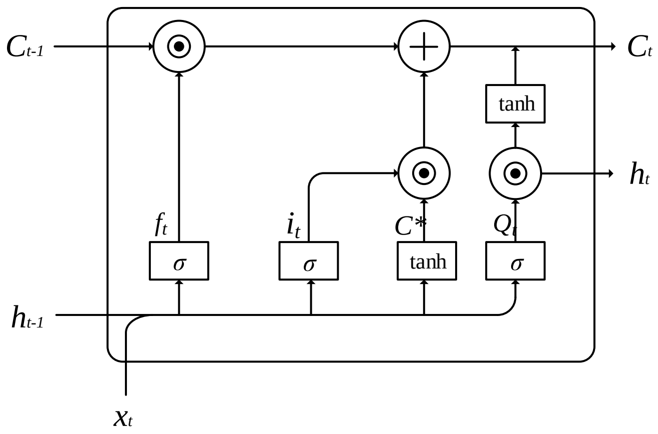

Figure 1: Single neuron structure of LSTM

2.4 Attention mechanism and CNN-LSTM hybrid model

In this paper, the CNN-LSTM hybrid model is applied to dam deformation prediction. As the upper layer of the model structure, CNN is used to extract and screen the main features of the data through the convolutional layer and the pooling layer, and then it is flattened through the full connection layer. Then, the weights were assigned through the attention layer [14], finally, the predicted horizontal displacement of the dam is obtained by input LSTM. The flowchart of CNN-LSTM model based on attention mechanism is shown in Figure 2.

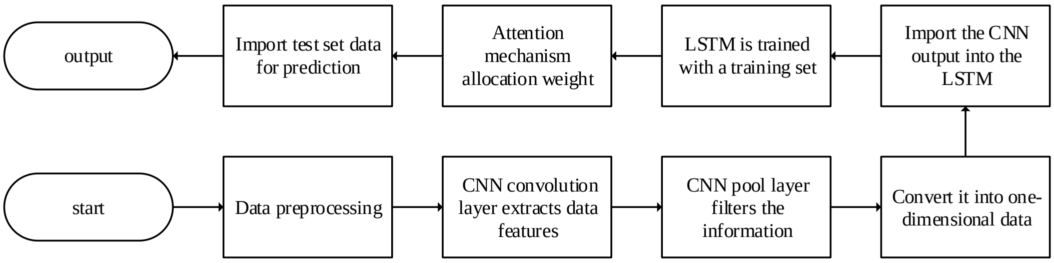

Figure 2: Flow chart of the proposed model

3. Experimental analysis

3.1 Data selection

The experimental data came from the horizontal displacement measurement data along the river flow of a reinforced concrete gravity dam No. 6 observation point from January 2002 to October 2016. The observation period was one month, and there were 178 sample data in total.

(1) Data difference processing

In order to improve the accuracy of prediction and provide accurate and concise data for the prediction model, it is necessary to process the test data in advance. In this paper, linear interpolation method is used to complement the missing values [15]. The calculation formula of linear interpolation method is as follows:

$$x_k = x_i + (x_j - x_i) \times \frac{k-i}{j-i} \tag{8}$$

In Formula 8, x_k is the value to be completed; x_i is the known data bit before x_k; x_j is one bit known after x_k.

(2) Data set partitioning and normalization

Firstly, the data of 178 period were divided into training set, verification set and test set according to

ISSN 2663-1563 Vol. 3, Issue 1: 14-19, DOI: 10.25236/AJAGE.2021.030103

7:1.5:1.5, the training set is used to train the model to find the best weight and bias, the verification set is used to filter the training model to find the best hyperparameters, and the test set is only used to evaluate the performance of the trained model. In order to accelerate the gradient descent speed, the temperature, water level and other parameters affecting the dam displacement are normalized according to the formula [16], and all input parameters are scaled to the range [0,1] using the maximum and minimum standardization. The normalization formula is as follows:

$$\tilde{x} = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \tag{9}$$

In Equation 9, $\tilde{x}$ is the normalized value; x_{max} and x_{min} are the maximum and minimum values of input parameters respectively.

3.2 Model evaluation criteria

In this paper, the commonly used regression model evaluation indexes RMSE, MAPE and R^2 were selected to prove the accuracy of the model. RMSE and MAPE are closer to 0. The closer R^2 is to 1, the higher the accuracy of the model.

3.3 Parameter Settings

In order to verify the effectiveness of CNN-LSTM by introducing attention mechanism, the model is implemented on the basis of the TensorFlow deep learning framework and compared with the LSTM and CNN-LSTM models. The structure of the CNN-LSTM model based on attention mechanism It consists of an input layer, a convolutional layer, a pooling layer, an LSTM layer with integrated attention mechanism, and an output layer. In this model, the size of the time window, the number of convolution kernels, the size of batch processing, the number of training times and the number of hidden layer units are set as super parameters. In order to reduce the influence of human factors on the model, the value of the super parameter is obtained through the following series of experiments: the values of the sliding window are designated as 4, 5 and 6, respectively, and the three experiments are different in the execution time of Windows. The number of cells in the LSTM hidden layer is 68. The training time is directly determined by the model error loss. Figure 3 shows the loss curve of attention mechanism and CNN-LSTM hybrid model. As the number of iterations increases, the training loss curve and the verification loss curve tend to zero. When the number of iterations reaches 1080, the verification loss curve no longer drops. Choose trial and error to determine the number of sliding Windows. Table 2 shows the performance of the hybrid model under different sliding windows. It can be seen from Table 2 that when the sliding window is 4, the performance is the best, RMSE is 0.4872, MAPE is 2.7431%, and R^2 is 0.9543. As the sliding window increases, RMSE and MAPE also increase, while R^2 decreases.

Table 1: Performance of the model under different sliding Windows

Number of sliding Windows	RMSE/mm	MAPE/%	R^2
4	0.4872	2.7431	0.9543
5	0.4983	2.8533	0.9433
6	0.5005	2.9432	0.9235

3.4 Model comparison and analysis

Set the sliding window of the three models of CNN-LSTM model, LSTM model, and CNN-LSTM model based on attention mechanism to 4, then determine the model training set and validation set, and input the test set and validation set to the model for training. Then use the trained model for simulation and prediction experiments. Using RMSE, MAPE and R^2 are used to evaluate and compare different models. The simulation results are shown in Table 2. According to Table 2, RMSE and MAPE of CNN-LSTM model based on attention mechanism are 0.1756mm and 1.4561% lower than those of CNN-LSTM model, and 0.3807mm and 4.3514% lower than those of LSTM model, R^2 is 0.0187 and 0.0638 higher than CNN-LSTM model and LSTM model, respectively. Experimental results show that the CNN-LSTM hybrid model based on attention mechanism is better than LSTM and CNN-LSTM in predicting the horizontal displacement of the dam. Figure 4 shows the comparison between the predicted and actual horizontal displacement of the dam. The abscissa is the number of months, and the time is 24 months from October 2014 to October 2016. Compared with the LSTM model and the CNN-LSTM model, the prediction curve of the CNN-LSTM hybrid model based on attention mechanism is closer to the true curve.

ISSN 2663-1563 Vol. 3, Issue 1: 14-19, DOI: 10.25236/AJAGE.2021.030103

Table 2: Prediction performance of different models

Model	RMSE/mm	MAPE/%	R^2
LSTM	0.7689	5.0635	0.8905
CNN-LSTM	0.5638	2.1683	0.9356
Attention mechanism and CNN-LSTM	0.3882	0.7121	0.9543

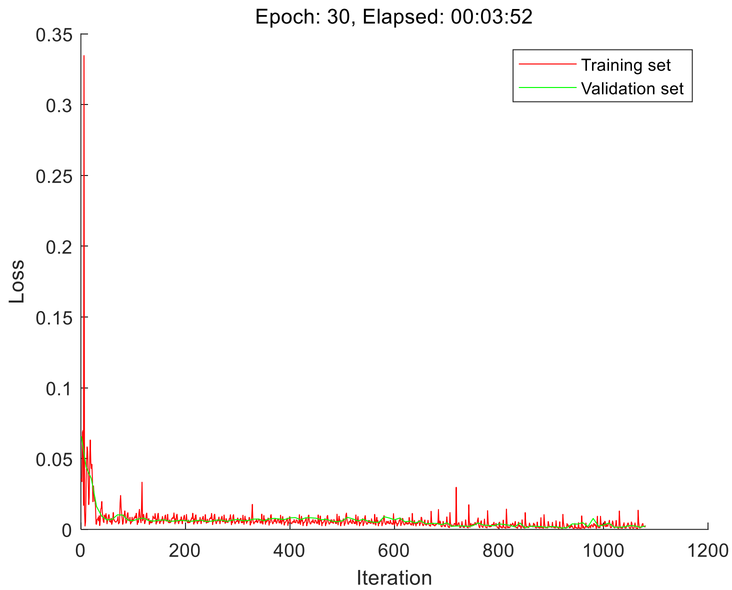

Figure 3: Loss curve of CNN-LSTM model based on attention mechanism

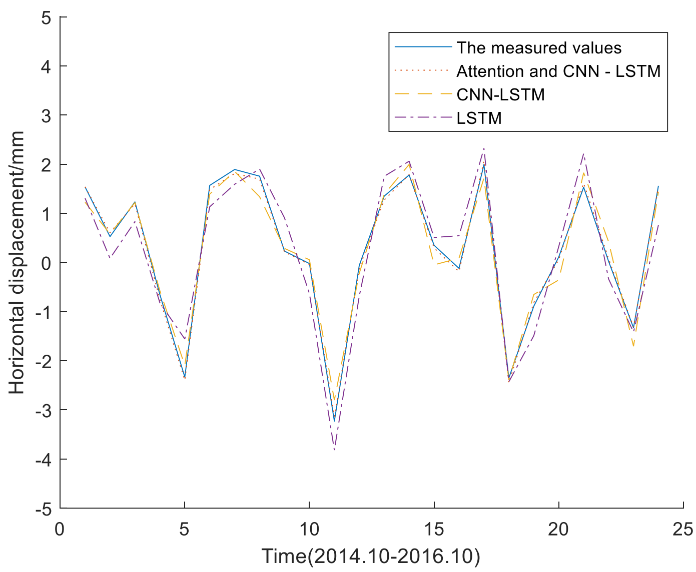

Figure 4: Comparison chart of the predicted horizontal displacement of the dam with the actual result

4. Conclusion

In this paper, a hybrid CNN-LSTM model based on attention mechanism is proposed for dam horizontal displacement prediction.This method not only solves the problems of time series fluctuation and random disturbance term in horizontal displacement data of DAMS, but also combines attention mechanism, LCNN model and LSTM to have their own advantages.Compared with CNN-LSTM model and LSTM model, the CNN-LSTM hybrid model based on attention mechanism can effectively improve the prediction accuracy of horizontal displacement of DAMS.The results show that the proposed model improves the overall accuracy of dam horizontal displacement prediction, ensures a better local prediction

ISSN 2663-1563 Vol. 3, Issue 1: 14-19, DOI: 10.25236/AJAGE.2021.030103

value, and has a higher adaptability.It provides a new choice for dam horizontal displacement monitoring, and has a certain reference value for landslide and bridge horizontal displacement monitoring.

References

[1] M.L. Du, J.Y. Gao and L.B. Zhang (2020). Time-series prediction and completion of dam monitoring data. Hydroelectric Power, vol.46, no.11, p.111-115.
[2] F.L. Li (2020). Analysis and prediction of dam safety monitoring data of Baishi reservoir. Heilongjiang Water Conservancy Technology, vol.48, no.8, p.52-56.
[3] Y.H. Wang and H.Z. Su (2020). Dam deformation prediction based on PCA-GWO-SVM. The People of The Yellow River, vol.42, no.11, p.111-115.
[4] M.J. Li, J.X. Wang and J.Y. Pan (2021). Dam deformation prediction based on improved GWO and SVM. Hydroelectric Power, vol.47, no.3, p.89-93.
[5] X.K. Zhou, Y.Y. Hu and W Liang (2021). Variational LSTM Enhanced Anomaly Detection for Industrial Big Data. IEEE Transactions on Industrial Informatics, vol.17, no.5, p.3469-3477.
[6] L. Ren, J.B, Dong and X.K. Wang (2021). A Data-Driven Auto-CNN-LSTM Prediction Model for Lithium-Ion Battery Remaining Useful Life. IEEE Transactions on Industrial Informatics, vol.17, no.5, p.3478-3487.
[7] B.C. Li, H.J. Yu and J.Y. Liu (2021). Short-term PV power generation prediction based on Kmeans and CEEMD-PE-LSTM. Hydroelectric Energy Science, vol.39, no.4, p.204-208.
[8] H.C. Shin, H.R. Roth and M.C. Gao (2016). Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics and Transfer Learning. IEEE Transactions on Medical Imaging, vol.35, no.5, p. 1285-1298.
[9] L. Zheng, Y. Yang and Q. Tian (2018). SIFT Meets CNN: A Decade Survey of Instance Retrieval. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.40, no.5, p.1224-1244.
[10] Y.B. Shen, X.L. Zhang and Y. Xia (2021). BI-LSTM neural network for residual service life prediction of bearings. Journal of Vibration Engineering, vol.34, no.2, p. 411-420.
[11] X. Liang and Z.H. Zhang (2021). Dynamic fluid level depth of oil well based on short-term energy and LSTM. Computers and Modernization, vol.32, no.4, p.15-19+26.
[12] Z. Zhao, W.H. Chen and X.M. Wu (2017). LSTM network: a deep learning approach for short-term traffic forecast. IET Intelligent Transport Systems, vol.11, no.2, p.68-75.
[13] X.Y. Guo, J.H. Lian and H.M. Li (2021). Prediction algorithm of ammonia concentration in chicken house based on two-stage attention mechanism and LSTM. Journal of China Agricultural University, vol.26, no.6, p.187-195.
[14] G.J. Liu and X.B. Wang (2020). Text classification of marketing news based on CNN+LSTM attention. Computer Technology and Development, vol.30, no.11, p.59-63.
[15] X.M. Liu, D.B. Zhao and R.Q. Xiong (2011). Image Interpolation Via Regularized Local Linear Regression. IEEE Transactions on Image Processing, vol.20, no.12, p.3455-3469.
[16] M.D. Zhen, B.L. Zhang and V.D. Radulescu (2021). Normalized solutions for nonlinear coupled fractional systems: low and high perturbations in the attractive case. Discrete and Continuous Dynamical Systems, vol.41, no.6, p.2653-2676.