Typhoon Track Prediction Based on CNN-ATTN-LSTMs Hybrid Model

Yuchen Zhang
Alibaba Cloud College of Big Data Application
Zhuhai College of Science and Technology
Zhuhai,China
Yorksen2001@outlook.com

Zijun Liang
Faculty of Science and Technology
University of Macau
Macau,China
3520731426@qq.com

Jiadong Lu
Alibaba Cloud College of Big Data Application
Zhuhai College of Science and Technology
Zhuhai,China
1870151544@qq.com

Bangting Yu
college of mathematics and information
South China Agricultural University
Guangzhou,China
chichun_leung@163.com

Rouying Wu
Alibaba Cloud College of Big Data Application
Zhuhai College of Science and Technology
Zhuhai,China
wurouying@stu.zcst.edu.cn

Meixuan Jiang
Alibaba Cloud College of Big Data Application
Zhuhai College of Science and Technology
Zhuhai,China
1302350091@qq.com

Wei Lv
Alibaba Cloud College of Big Data Application
Zhuhai College of Science and Technology
Zhuhai,China
luwei@jluzh.edu.cn

***Abstract*—The accurate prediction of the specific path and trajectory of typhoons can greatly reduce the disasters caused by typhoons, and it is also of great significance for relevant institutions to conduct risk assessment. With the development of artificial intelligence computer field and deep learning technology, more and more related technologies of deep neural network are used in the field of meteorology. This paper proposes a typhoon trajectory prediction model with a hybrid mechanism of CNN-ATTN-LSTMs incorporating the Attention mechanism. The model mixed mode is used to fully mine the characteristics of typhoon trajectory data, aiming to improve the accuracy of typhoon trajectory prediction. This paper uses the best typhoon track dataset from the Tropical Cyclone Data Center of the China Meteorological Administration (Shanghai Typhoon Research Institute), and compares the CNN-ATTN-LSTMs model with the separate LSTM and TCN models to predict the typhoon track. The results show that the error distance indicators of the CNN-ATTN-LSTMs model are better than the single LSTM model and the TCN model. That is, on the basis of the LSTM typhoon trajectory prediction model, the typhoon trajectory prediction accuracy is further improved.**

Keywords—lstm, cnn, ybrid neural, typhoon track prediction.

I INTRODUCTION

Typhoon is a typical tropical weather system and one of the important forms of ocean-atmosphere interaction.[1] my country is located in the Northwest Pacific region, and the Northwest Pacific region is one of the places with the highest incidence of typhoons in the world. Every year in the typhoon season, some areas of my country will cause different degrees of economic losses and casualties due to typhoon disasters. Targeted, timely and effective prediction of typhoon tracks can provide effective data support for disaster prevention and mitigation departments and related institutions, and then To achieve the purpose of reducing casualties and economic losses. However, as we all know, there are many factors that affect the typhoon trajectory, and after the typhoon actually makes landfall, the typhoon trajectory will also be affected by land, topography, and coastline water depth [2]. Therefore, typhoon trajectory prediction is a very important and challenging research direction and research topic.

Early typhoon trajectory prediction mainly relied on thermodynamics and aerodynamics [3] to analyze the typhoon in detail, combined with the analysis of the typhoon landing point on the complex coastline and the factors affecting the land topography in the coastal area, and established a unique rule of thumb in the field of typhoon trajectory prediction. However, **under the influence of subjective consciousness, this empirical method presents a situation of low efficiency,** and also requires a lot of manpower and material resources, and the accuracy and timeliness of prediction are difficult to meet the demand. With the wide application of neural networks, and the typhoon has obvious nonlinear structure and the big data characteristics of related image data [4-5], the multi-layer deep neural network (Deep Neural Networks, DNN) from the convolutional neural network

[6](Convolutional Neural Network, CNN) to recurrent neural network [7] (recurrent/recursive neural network, RNN) and other models have gradually begun to be used in typhoon prediction models, hoping to obtain more convenient and accurate prediction effects.

A convolutional neural network (CNN) model is usually a neural network used to process image features, and has powerful feature extraction and mining capabilities. There are many researchers using typhoon satellite cloud image data to extract and analyze the features in the spiral part, so as to improve the prediction accuracy of typhoon trajectory. However, the way to extract the image is to directly read the satellite cloud image data. If the resolution of the data image sent is low, the prediction effect will be greatly affected. In the study of Gao Shan et al., it has occurred that the accuracy of CNN image prediction is greatly reduced due to the complex atmospheric factors in the formation of typhoons and the insignificant spiral radius of cloud images [8]. **Therefore, CNN has certain defects in the feature extraction of typhoon cloud images, and cannot complete high-precision typhoon trajectory prediction.**

Recurrent Neural Networks (RNN) [9-11] **is a kind of neural network suitable for processing time series data**, which has been widely used in many fields, but there will be the problem of "vanishing gradient" in the later iteration. Hochreiter et al. (1997) first proposed a long short-term memory network [12-13] (Long Short-Term Memory, LSTM) by adding input gates, forgetting gates and output gates, so that the weight of the network self-circulation can be changed, thereby avoiding " Gradient vanishing" problem, suitable for processing and forecasting long-delayed events in time series. Therefore, a large number of studies have shown that, considering typhoon data as time series data, the LSTM algorithm can be of great help to this research topic. Xu Gaoyang et al. [14] applied **the LSTM network in typhoon trajectory prediction, and combined with the dynamic time warping algorithm, the effect of predicting typhoon trajectory was better.**

With the wide application of deep learning technology in the computer field, the data and information contained in it grow exponentially, the degree of connection between data increases, and the influencing factors are different. The neural network variable feature set constructed by traditional methods cannot be fully mined. For the connection of non-continuous features in high-dimensional space, traditional LSTM has been unable to perfectly solve the classification or prediction problems with high degree of connection between data or many data influencing factors [15]. In view of the significant advantages of convolutional neural networks in data feature extraction and dimensionality reduction [16-17], this paper is based on previous studies [18-20], based on convolutional neural networks (CNN) and long short-term memory networks (LSTM), **a dual-model hybrid network mechanism with attention mechanism (CNN-ATTN-LSTMs) is proposed.**

The optimal path data of typhoon from 2010 to 2019 is matrixed, and the feature correlation is extracted through the CNN network, which is input into the long-term and short-term memory neural network, and the different inputs are scored and assigned weights. The features are connected to get the final prediction result.

II THE HYBRID MODELS OF CNN-ATTN-LSTMS

In order to fully consider some characteristics of typhoon track prediction in detail focusing and feature-related, combined with the influence of historical typhoon track data on typhoon track prediction, this paper combines the advantages of CNN network and LSTM network, and integrates the module of attention mechanism, and proposes CNN- The basic structure of ATTN-LSTMs hybrid neural network has a good effect on the prediction of typhoon trajectory.

Figure 1 shows the flow chart of the entire model, the specific steps are as follows:

- Step1: Convert the selected representative typhoon data into a trajectory matrix, which is input into the CNN network as input
- Step2: Extract features from the input through the CNN network, that is, extract the correlation between the time series of typhoon tracks.
- Step3: Input the extracted data to the LSTM layer according to the divided time step, and then use the scoring function to assign different weights to different inputs, that is, to expand the difference in influencing factors.
- Step4: Finally, perform feature integration to obtain prediction results.

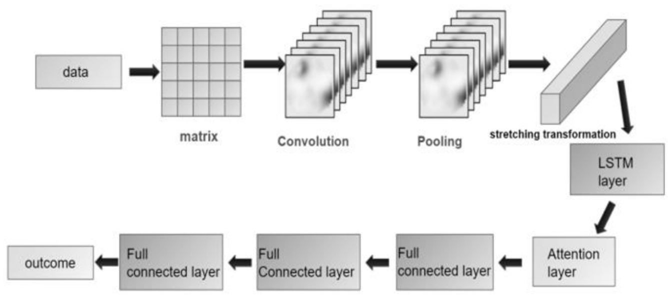

Fig.1. CNN-ATTN-LSTMs Hybrid Network Structure

A. The Models of CNN

Convolutional Neural Network (CNN) is a widely used neural network in the field of deep learning, and has achieved remarkable application results in image recognition, speech recognition, text classification and other fields. Since CNNs are mostly used to complete image processing tasks, studies have shown that CNNs also have strong applicability in prediction tasks. The principle is to use the ability of the convolution kernel to feel the situation of the data for a period of time, and make predictions based on the situation of the data for a period of time. CNN is composed of input layer, convolution layer, pooling layer, fully connected layer and output layer. The convolution and pooling layers are used for feature engineering, and the fully connected layer is used for feature weighting. The network itself has "local links" and "" The "weight sharing" feature simplifies the complexity of network links, improves the model's ability to extract abstract features, and alleviates the problems of slow training and overfitting of fully connected networks to a certain extent. The structure of the CNN network is shown in Figure 2.

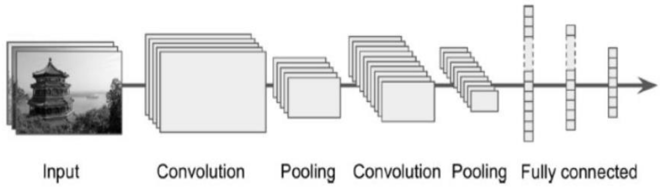

Fig.2. CNN network structure

This paper continues the previous research direction of our research group, **using one-dimensional convolution Conv1d to extract features through 64 convolution kernels, then activate through the Relu function, then perform pooling through max-pooling, and add dropout to prevent Over-fitting, according to the above convolution and pooling principles, calculate and obtain the feature matrix. Finally, the variable feature is stretched into a vector, which is used as the input variable of the next module**.

B. Modules that Introduce Attention Mechanisms

Attention (attention) mechanism literally means the attention to the same transaction, and its core logic is "from paying attention to all to focus". The structure of the attention mechanism is shown in Figure 2. At present Attention is a mechanism (Mechanism) used to improve the effect of the RNN-based Encoder + Decoder model, but the attention model can be regarded as a general idea, which itself does not depend on a specific framework, generally called Attention Mechanism . In this paper, it is applied based on the LSTM model, and weights are assigned to different inputs through the scoring function as the input of the next step. The basic steps are divided into three steps:

- Step 1: Calculate the similarity between query and key to get the weight value
- Step 2: Normalize the weight values to obtain directly usable weights
- Step 3: Weighted summation of weights and values

C. The Models of LSTM

The LSTM model is a variation of the RNN model proposed by Hochreiter etc. in 1997, which learns long-term dependent information of time series, and keeps the useful key information as much as possible. Compared with RNN, it solves the problem of long-term dependence by adding a new cell state and three gates mechanism. LSTM consists of several repetitive modules, one module contains at least temporal state, cell state and three gates. The gate of LSTM is a method to let information pass selectively, and contains a sigmoid neural network layer and a pointwise multiplication operation.

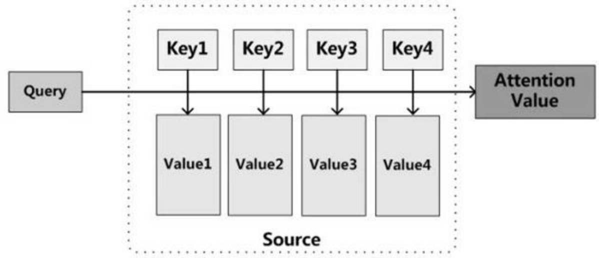

Fig.3. Attention Mechanism Model

The three LSTM gate contain the input gate, the forgetting gate and the output gate, are used to determine the discarding of information, the updating of information and the output of information, respectively.

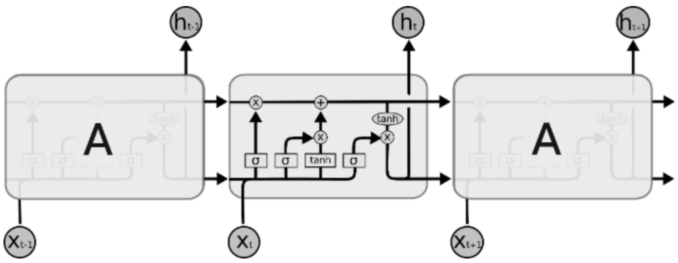

Fig.4. Neural Cell Structure

The first step in an LSTM network is to get h_{t-1} and x_t and than output a value between 0 and 1 to C_{t-1}.The formula for the calculation is as follows.

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \tag{1}$$

Where W_f is the forgetting gate weight. b_f is the forgetting gate bias term.

The next step consists of two parts, one is the sigmoid layer called the "input gate layer" that determines what values we will update. Second, the tanh layer creates a new vector of candidate values $\tilde{C}_t$ that is added to the state. The formula for the calculation is as follows.

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \tag{2}$$

$$\tilde{C}_t = tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \tag{3}$$

Where W_i is the input gate weight.h_{t-1} is the implied state.x_t is the input data.b_i is the input gate bias term.W_c is the implied state weight; and b_c is the implied state bias term; and $\tilde{C}_t$ is the current memory cell.

Eventually, we determine the values to be output. We will run a sigmoid layer to determine which part of the cell state will be output. Next, we process the cell state through the tanh layer to get a value between -1 and 1, and multiply it with the output of the sigmoid gate, and finally we output the fraction to be output. The formula for the calculation is as follows.

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \tag{4}$$

$$h_t = o_t * \tanh(C_t) \tag{5}$$

Where W_o is the output gate weight.b_o is the output gate bias term.

III EXPERIMENT

D. Experimental Data

This paper uses the best data from the Tropical Cyclone Data Center of the China Meteorological Administration (Shanghai Typhoon Research Institute) to obtain the best track data of about 200 typhoons from 2000 to 2019, and process the data in detail.

Accuracy evaluation index

When using a model for forecasting, the error is calculated based on the difference between the predicted coordinates and the actual coordinates. Using three accuracy evaluation indicators, the root mean square error (RMSE) can well reflect the offset of coordinate points, and the mean absolute error (MAE) can well reflect the average value of the distance between the model predicted value samples and the real values. The absolute percentage error (MAPE) measures the deviation. Therefore, using the above three accuracy evaluation indicators to verify the error of the CNN-ATTN-LSTMs model for typhoon prediction can more intuitively reflect the reliability of the model. The formulas for RMSE , MAE and MAPE are expressed as follows:

$$RMSE = \sqrt{\frac{1}{N}\sum_{t=1}^{N}((x_t - \widehat{x_t})^2) + (y_t - \widehat{y_t})^2)} \tag{1}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|\widehat{y_i} - y_i| \tag{2}$$

$$MAPE = \frac{100\%}{n}\sum_{i=1}^{n}\left|\frac{\widehat{y_i} - y_i}{y_i}\right| \tag{3}$$

L1 Paradigm:Loss=sum {|x-x'|+|y-y'|} (4)

Where, x_t , y_t are the true value of longitude and latitude, $\widehat{y_t}$, $\widehat{x_t}$ are the corresponding predicted value, and T is the TTH predicted value. A represents the difference between the latitude of the real data and the longitude of the forecast data, and B represents the difference between the longitude of the real data and the latitude of the forecast data. R represents the radius of the earth, taking 6378.137 km.

For the above four precision evaluation indexes, the smaller the value of their values, the better the robustness of the model and the stronger the prediction ability of new data.

During training, the appropriate model structure and model parameters are determined by observing their changes.

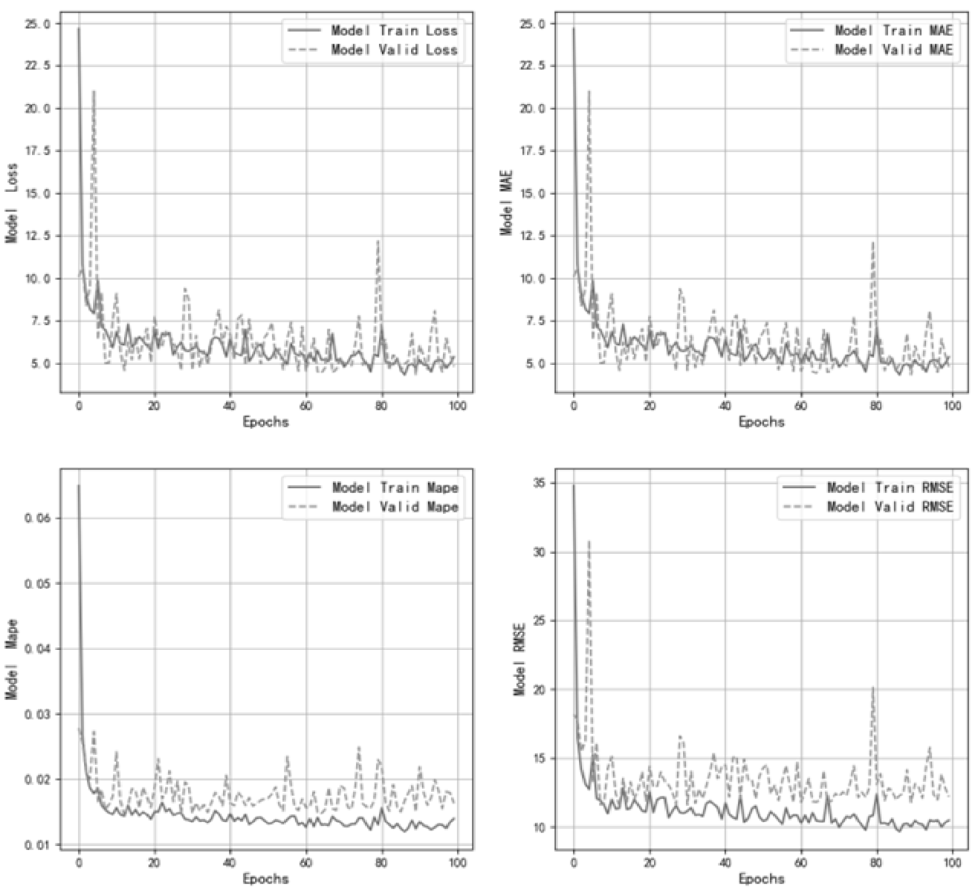

Fig.5. Prediction from the TCN Model

E. Analysis of Results

Since typhoon prediction requires timeliness, the experiment selects data with a minimum time interval of 6 hours for training time step. Then all the typhoon track data is divided, 75% of the data is used as the training set, which is used to train the parameters of the CNN-LSTM prediction model of different models, and the remaining 25% of the data is used as the validation set, which is used to verify the learning of the model Effect.

In order to verify the improvement of the prediction performance by adding a CNN mechanism for extracting the time-column dimension correlation, a TCN Model and a separate LSTM model and a CNN-LSTM hybrid mechanism model without an attention mechanism module and a Predictions from the CNN-ATTN-LSTMs model module are used. The models predict 6 hours trajectories respectively.According to the prediction results in Figure 5,6,7,8.

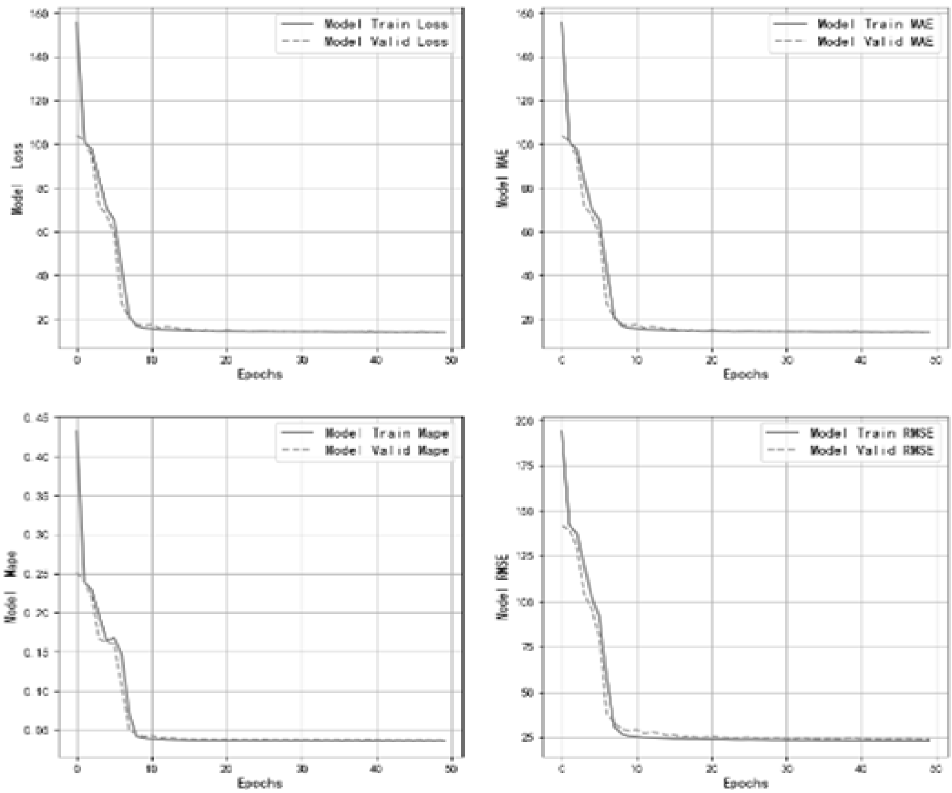

Fig.6. Predictions from the Separate LSTM Model

As can be seen from the figure, **the RMSE and MAEP, MAE data obtained from the experiment by the model CNN-ATTN-LSTM introducing the attention module are lower than those of other comparison models.** In particular, the reduction in predicting longer time series is higher big. Because although LSTM also has the ability to strengthen time series and long memory, under the CNN-LSTM structure, long sequences will greatly reduce the prediction accuracy of the model. The attention enhancement module can solve this problem and can re-allocate the weight parameters of the time series information.

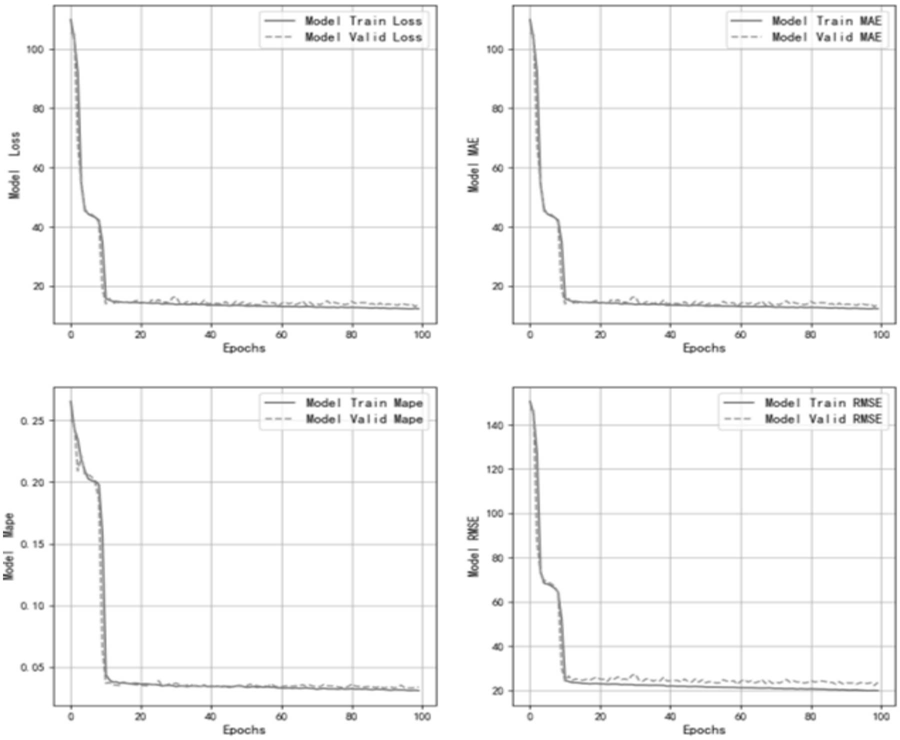

Fig.7. Predictions from the Separate CNN-LSTM Model

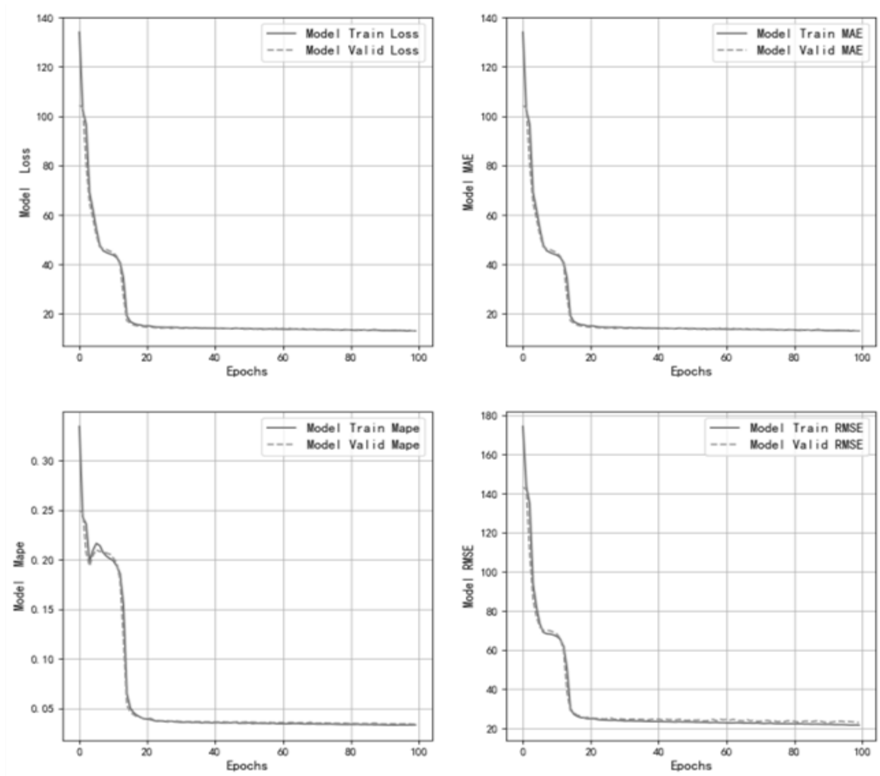

Fig.8. Predictions from the CNN-ATTN-LSTMs Model

From the above table 1, it can be seen that the model CNN-Attention-LSTM with the introduction of attention mechanism obtained from the experiments Valid_loss and RMSE and MAE and MAPE data are lower than those of the CNN-LSTM model alone, so it can be recognized that the CNN-Attention-LSTM model is better trained compared to the above two models.（The

prediction effect of the TCN model is poor, and this table does not compare）

TABLE.1 Comparison of Accuracy Indicators for Multiple Models

	Valid_ loss	Test_ MAE	Test_ MAPE	Test_ RSME
LSTM	14.0347	14.0163	0.0338	23.6839
CNN-LSTM	13.5674	13.5419	0.0364	23.3553
CNN-ATTN-LSTMs	13.077	13.0621	0.0342	22.954

F. Case Verification

This paper selects the most influential typhoon case "Korovan"(Tropical Storm Krovanh, International Number: 2023) in 2020 for verification. The comparison between the model prediction of typhoon and the actual observation value is shown in Figure 9. The experiment time is from 00:00:00 on December 18, 2020 to 6:00:0 on December 25, 2020, and the typhoon rules represented by the red track in the legend Observed values, the green trace represents the predicted value obtained by the CNN-ATTN-LSTMs hybrid model, and the blue trace represents the predicted value obtained by using the LSTM model alone.

Fig.9. Comparison of actual observed and predicted values

IV CONCLUSION

The practical research significance of the CNN-ATTN-LSTM hybrid model proposed in this paper for the application of typhoon trajectory prediction is that it fully exploits the features and temporal information of typhoon trajectory data, adds CNN to the input data to extract the temporal feature correlation, and weights the hidden states in it to be able to obtain the time-dependent enhanced sequences, and finally constructs the corresponding deep neural network for typhoon The corresponding deep neural network is finally constructed to analyze and predict the typhoon trajectory. **Compared with the pure LSTM typhoon track prediction model and the CNN+LSTM model with feature extraction, it is obvious that the CNN-ATTN-LSTM model can improve the prediction accuracy.**

Since there are many factors influencing the typhoon trajectory, and the influence on the typhoon trajectory may be large or small, the influence of other related elements on the typhoon trajectory should also be fully considered in the subsequent research, and the factors influencing the typhoon trajectory will be added as feature variables in the next research, and a complex neural network model containing many elements will be designed to improve the prediction warp.

REFERENCES

[1] Zhang Zhiwei, 2019. A case study on the response of the upper ocean to typhoons in the northwestern Pacific Ocean [J]. Ocean Bulletin, 38(5): 562-568

[2] Jin Hua, Tang Jiaxiang, Dai Yuhan, et al., 2012. Analysis of operational forecast errors and causes of typhoon tracks in my country [J]. Meteorology, 38(6): 695-700.

[3] Huang X Y, Jin L. An artificial intelligence prediction model based on principal component analysis for typhoon tracks[J]. Chinese J Atmospheric Sci, 2013, 37(5): 1154-1164.

[4] Liang Peng. Research on media early warning of typhoon disasters (2001-2010) under the path of big data [D]. Central China Normal University, 2014.

[5] Wang Xuyang. Review of traditional models and neural networks for predicting typhoon tracks [J]. Scientific Consulting (Science and Technology Management), 2020(01):62-65.

[6] Zeiler M D, Fergus R. Visualizing and Understanding Convolutional Networks[J]. 2013

[7] Jeffrey L. Elman. Finding structure in time[J]. Cognitive Science, 14(2):179-211.

[8] Gao Shan. Research on typhoon intensity prediction based on deep learning [D]. Guangxi University, 2021. DOI: 10.27034/d.cnki.ggxiu.2021.000721.

[9] TOKGZA, NALG, 2018. A RNN based time series approach for fore－casting turkish electricity load[C]. 2018 26th Signal Processingand Communications Applications Conference (SIU): 1-4.

[10] Yang Li, Wu Yuqian, Wang Junli, Liu Yili. Review of Recurrent Neural Network Research [J]. Computer Applications, 2018, 38(S2): 1-6+26.

[11] Xia Yulu. A review of the development of recurrent neural networks [J]. Computer Knowledge and Technology, 2019, 15(21): 182-184. DOI: 10.14004/j.cnki.ckt.2019.2379.

[12] Staudemeyer R C, Morris E R. Understanding LSTM--a tutorial into Long Short-Term Memory Recurrent Neural Networks[J]. arXiv preprint arXiv:1909.09586, 2019.

[13] Gao Chengliang. Research on the representation method of text context-dependent features based on LSTM [D]. Hebei University of Science and Technology, 2019.

[14] Xu Gaoyang, Liu Yao, 2019. Application of LSTM network in typhoon track prediction [J]. Computer and Modernization, 285(5): 68-72.

[15] Qi, Liu Dongxu, Song Wei, Huang Dongmei, Du Yanling. Typhoon track prediction model based on dual attention mechanism [J]. Ocean Bulletin, 2021,40(04):387-395.

[16] Li Zhishuai, Lv Yisheng, Xiong Gang. Short-term traffic flow prediction based on graph convolutional neural network and attention mechanism [J]. Traffic Engineering, 2019,19(04):15-19+28.DOI:10.13986 /j.cnki.jote.2019.04.003.

[17] Zhang Xue. Movie box office prediction based on deep learning convolutional neural network [D]. Capital University of Economics and Business, 2017.

[18] Huang Jie, Zhang Feng, Du Zhenhong, Liu Renyi, Cao Xiaopei. PM_(2.5) Hourly Concentration Prediction Based on RNN-CNN Integrated Deep Learning Model [J]. Journal of Zhejiang University (Science Edition), 2019,46(03):370-379.

[19] Zhao Hongrui, Xue Lei. Research on stock prediction based on LSTM-CNN-CBAM model [J]. Computer Engineering and Applications, 2021, 57(03): 203-207.

[20] Dou Min. Design and implementation of video semantic analysis system based on CNN and LSTM [D]. Nanjing University of Posts and Telecommunications, 2018.