Expressway Traffic Flow Prediction Model Based on Bi-LSTM Neural Networks

Xiaojuan XUE\textsuperscript{1a}, Yuanhua JIA\textsuperscript{2b}, Shengyou WANG\textsuperscript{3c}\textsuperscript{*}

\textsuperscript{1}School of Traffic and Transportation Beijing jiaotong university Beijing, China, CCCC First Highway Consultants Co., Ltd Xian, China
\textsuperscript{2}School of Traffic and Transportation Beijing jiaotong university Beijing, China
\textsuperscript{3}School of Traffic and Transportation Beijing jiaotong university Beijing, China
\textsuperscript{a}59022038@qq.com, \textsuperscript{b}yhjia@bjtu.edu.cn
\textsuperscript{c}\textsuperscript{*}Corresponding author: 18114043@bjtu.edu.cn

Abstract—In this paper, the expressway traffic flow prediction model based on the Bi-LSTM is designed, and four sections of expressway are applied to the model for training and evaluation. Through training and verification, the results show that the average prediction accuracy, MAPE, RMSE, and MAE of the proposed model is 89.54\%, 10.46\%, 31.55, and 24.58, respectively. In addition, in order to evaluate the effect of the proposed model, this paper introduces the ARIMA model for comparison. It is found that the prediction accuracy of the Bi-LSTM is 18.30\% higher than ARIMA, the RMSE is reduced by 31.85, and the MAE is reduced by 26.32. The results show that the proposed Bi-LSTM model exhibits higher prediction performance. Afterwards, this paper makes a comparative analysis of the predicted value and the original value of each road section. The results show that the proposed model has a certain lag, and the forecast value of traffic flow is low for rush hours, however, the forecast value of traffic flow is higher in the low peak hours.

1. INTRODUCTION

With the development of economy and the advance of urbanization, the continuous increase of the number of vehicles has brought more and more serious urban traffic congestion. Therefore, accurate traffic flow prediction and strategic decision-making for the management to control and manage the traffic situation play an important role in alleviating the road pressure. The traditional traffic flow model prediction is the parameter model, such as ARIMA\textsuperscript{[1]}, Kalman filter \textsuperscript{[2]}, et al. However, the simple structure of the parameter model will produce large errors for the non-linear prediction problem of traffic flow. At present, non-parametric models, including Support Vector Machine (SVM), BP neural network and deep learning model\textsuperscript{[3-4]}, have made great progress in traffic flow prediction. In recent years, the deep learning approach has proved to be a great improvement in terms of accuracy in predicting traffic flow\textsuperscript{[5]}. In the prediction of traffic flow time series, the Recurrent Neural Network (RNN) has the advantages of strong memory, parameter sharing, complete turing and nonlinear feature extraction\textsuperscript{[7]}. Although many scholars have verified the validity of RNN model, the traffic flow has complex nonlinear relationship, and the prediction effect depends on the collected historical data to a great extent. In addition, the computational cost of complex models affects the application of models in practice. Therefore, scholars are constantly developing novel traffic flow prediction methods. There are many variants on the basis of the RNN, among which the better methods include the Long Short Term
Memory Neural Network (LSTM) and Bi-directional Long Short Term Memory Neural Network (Bi-LSTM)[8-12]. Ma adopts the LSTM method to verify the accuracy of traffic data prediction [13]. However, LSTM only considers unidirectional traffic flow data, and few scholars have studied the bi-directional traffic flow time sequence learning. This paper establishes the expressway traffic flow prediction model based on Bi-LSTM method to predict expressway traffic flow, and the model result was evaluated and compared with the traditional ARIMA model. It is hoped that the research in this paper can provide reference for the application of models in the future intelligent transportation system and the method research of scholars.

2. Bi-LSTM

Bi-directional Long Short-Term Memory (Bi-LSTM) is a time-recurrent neural network, which is composed of forward LSTM and backward LSTM. LSTM is a variant of the traditional recurrent neural network. Since the traditional recurrent neural network has only a relatively simple neural network module chain structure, there are problems that the gradient disappears and explodes that cannot be solved. Therefore, LSTM uses a special memory module to overcome the above problems. The memory model has designed three gates, namely the input gate, output gate and forget gate. The three gates can control the opening and closing of the gate under the influence of the activation function. The input gate controls the transfer of the input layer; the forget gate controls whether the historical data of the current memory module is abandoned, and outputs the output of the gate control information. The design of LSTM has obvious promotion effect for solving the long-term dependence problem of traditional recurrent neural network. In the traffic flow time series, considering the Change Law of forward and backward data is helpful to improve the fitting precision of the model. The Bi-LSTM structure is shown in Figure 1.

\[ i_t = \sigma(W_{ix} x_t + W_{ih} h_{t-1} + W_{ic} c_{t-1} + b_i) \]  
\[ f_t = \sigma(W_{fx} x_t + W_{fh} h_{t-1} + W_{fc} c_{t-1} + b_f) \]  
\[ o_t = \sigma(W_{ox} x_t + W_{oh} h_{t-1} + W_{oc} c_t + b_o) \]

\( i, f, o \) represents input gate, forget gate, and output gate, respectively. \( W \) represents the weight matrix; \( b \) represents the bias vector, and \( \sigma \) represents the sigmoid function. In order to verify the model results, this paper uses MAE, MAPE and RMSE evaluation parameters for evaluation, where the formula of RMSE is shown in Eq. (4), (5), (6).

\[ RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (\hat{y}_i - y_i)^2} \]  
\[ MAPE = \frac{100\%}{n} \sum_{i=1}^{n} \frac{|\hat{y}_i - y_i|}{y_i} \]  
\[ MAE = \frac{1}{n} \sum_{i=1}^{n} |\hat{y}_i - y_i| \]
In the Eq. (4) - (6), $\hat{y}_t$ represents original traffic flow, $y_t$ represents predicted traffic flow, and $n$ represents the number of samples.

3. DATA
In this paper, the traffic flow data of 4 sections of a certain expressway in Zhengzhou City is used. The time is from 2nd to 6th December 2019, a total of 5 day, and it is a working day from Monday to Friday. The time interval is 1 hour, and each section has 120 sets of data. The first three days of data are used for model training, and the next two days of data are used for model evaluation. The four road sections are numbered S1, S2, S3, S4, and the average flow of the four road sections is 240.61veh/h, 261.95veh/h, 204.98veh/h, and 200.39veh/h, respectively. Before the model training, ADF stability analysis must be performed on the data. The ADF values of the four road sections after difference and log stability processing are shown in Table 1.

| Section | S1          | S2          | S3          | S4          |
|---------|-------------|-------------|-------------|-------------|
| Inspection value | -3.798      | -2.701      | -3.798      | -3.798      |
| P-value  | 0.003       | 0.012       | 0.004       | 0.001       |
| Standard P-value | 1% -3.809  | 1% -3.321   | 1% -3.132   | 1% -2.721   |
|         | 5% -3.021   | 5% -3.031   | 5% -2.621   | 5% -2.301   |
|         | 10% -2.65   | 10% -2.312  | 10% -2.025  | 10% -1.719  |

As can be seen from Table 1, after log and differential processing, the P value is less than 0.05, and the four road sections have met the stability requirements. The following will predict the flow of the 4 road sections after the stability treatment. The model in this paper uses Python language and is trained and evaluated based on Keras.

4. RESULT
The four road sections are input into proposed method for training and evaluation. The number of epoch is set to 200, the time window step size is set to 4, the Dropout is set to 0.01, and the number of batches is 1. The comparison between the predicted value and the original value of each road segment is shown in Figure 2.
As can be seen from figure 2, the predicted value of Bi-LSTM has a certain delay compared with the original value, and the predicted value of Bi-LSTM is low for high peak flow and high for low peak flow. In addition, in order to evaluate the prediction effect of the proposed method, the ARIMA model was introduced for comparative analysis. The model evaluation results are shown in Table 2.

| Model  | S1    | S2    | S3    | S4    | Mean   |
|--------|-------|-------|-------|-------|--------|
| Bi-LSTM|       |       |       |       |        |
| RMSE   | 33.45 | 38.05 | 28.77 | 25.96 | 31.55  |
| MAPE   | 10.86 | 11.32 | 11.48 | 9.08  | 10.46  |
| MAE    | 26.89 | 28.06 | 24.13 | 19.23 | 24.58  |
| ARIMA  |       |       |       |       |        |
| RMSE   | 54.92 | 82.89 | 48.15 | 67.67 | 63.41  |
| MAPE   | 22.15 | 37.46 | 22.94 | 32.48 | 28.76  |
| MAE    | 44.60 | 67.96 | 41.13 | 47.69 | 50.89  |

It can be seen from Table 2 that the RMSE, MAPE and MAE of the S1 section are 33.45, 10.86 and 26.89, respectively, which are reduced by 44.85, 27.14 and 39.90 for the ARIMA model. The RMSE, MAPE, and MAE of the S2 section were 38.05, 10.32, and 28.06, respectively, which were reduced by 19.38, 11.45, and 17.00 for the ARIMA model, respectively. The RMSE, MAPE and MAE of the S3 section were 28.77, 11.48 and 24.13 respectively, which decreased by 41.72, 23.31 and 28.46 for the ARIMA model, respectively. The RMSE, MAPE and MAE of the S4 section are 31.55, 10.46 and 24.58 respectively, which are reduced by 31.85, 18.30 and 26.32 for the ARIMA model, respectively. The above analysis shows that the prediction accuracy of Bi-LSTM is significantly improved compared to ARIMA, and the average prediction accuracy is improved by 18.30%. The average prediction parameters of Bi-LSTM and ARIMA are shown in Figure 3.
It can be seen from Figure 4 that average MAPE obtained by the Bi-LSTM method after predicting four road sections is 10.46, which is 18.30 lower than ARIMA’s 41.26. The average RMSE of Bi-LSTM method is 31.55, which is 31.85 lower than ARIMA. The average MAE of Bi-LSTM method is 24.58, which is 26.32 lower than ARIMA. The proposed Bi-LSTM method has higher prediction accuracy, and the prediction accuracy is 18.30% higher than that of the ARIMA model.

5. CONCLUSION
In this paper, the Bi-LSTM expressway traffic flow prediction model is designed, and four expressway sections are applied to the model respectively, and the model is trained and evaluated. Prior to model training, this paper conducted unit root ADF tests on four road sections and performed logarithmic and differential processing to make the data more stable. Through training and verification, the results show that the average prediction accuracy of the model is 89.54%, MAPE is 10.46, RMSE is 31.55, and MAE is 24.58. In addition, in order to evaluate the effect of the proposed model, this paper introduces the ARIMA model for comparison. It is found that the prediction accuracy of Bi-LSTM is 18.30% higher than ARIMA, the RMSE is reduced by 31.85, and the MAE is reduced by 26.32. The results show that the proposed Bi-LSTM model exhibits higher prediction performance. Furthermore, this paper makes a comparative analysis of the predicted value and the original value of each road section. The results show that the model has a certain hysteresis, and the predicted value of traffic flow during peak hours is low, and the predicted value of traffic flow during low peak hours is high. In the future, methods to improve the prediction accuracy of the model are worth further mining, especially in terms of spatiotemporal characteristics.

ACKNOWLEDGMENT
This research was supported by the Fundamental Research Funds for the Central Universities (NO. 2019YJS107), China.

REFERENCES
[1] S. Kumar, L. Vanajakshi. “Short-term traffic flow prediction using seasonal ARIMA model with limited input data,” European Transport Research Review, 2015, 7(3), 1-9.
[2] J. Guo, W. Huang, William, M. Billy. “Adaptive Kalman filter approach for stochastic short-term traffic flow rate prediction and uncertainty quantification,” Transportation Research Part C: Emerging Technologies, 2014, 43: 50-64.
[3] H. Yang, T. Dillon, Y. Chen. “Optimized Structure of the Traffic Flow Forecasting Model With a Deep Learning Approach,” IEEE Transactions on Neural Networks and Learning Systems, 2017, 28(10) : 2371-2381.
[4] L. Duan, W. Kang, et al. “Traffic Flow Prediction With Big Data: A Deep Learning Approach”, IEEE Transactions on Intelligent Transportation Systems, 2015, 16(2): 865-873.
[5] K. Kumar, M. Parida. “Short term traffic flow prediction in heterogeneous condition using artificial neural network”, Transport, 2015, 30(4): 397-405.
[6] Z. Zhao, W. Chen, X. Wu, P. Chen, J. Liu. “LSTM network: a deep learning approach for short-term traffic forecast”, IET Intelligent Transport Systems, 2017, 11(2): 68-75.

[7] K. Cho, B. Merrienboer, C. Gulcehre, D. Bahdanau, F. Bougares, et al. “Learning phrase representations using RNN Encoder-Decoder for Statistical Machine Translation”, Conference on Empirical Methods in Natural Language Processing (EMNLP), 2014, 1724-1734.

[8] J. Guo, Z. Xie, Y. Qin, L. Jia, Y. Wang. “Short-term abnormal passenger flow prediction based on the fusion of SVR and LSTM”, IEEE Access, 2019, 7: 42946-42955.

[9] X. Luo, D. Li, Y. Yang, S. Zhang. “Spatiotemporal traffic flow prediction with KNN and LSTM”, Journal of Advanced Transportation, 2019, 1-10.

[10] N. Petersen, F. Rodrigues, and F. Pereira. “Multi-output bus travel time prediction with Convolutional LSTM neural networks source code”, Expert Systems with Applications, 2017, 120(15): 426-435.

[11] X. Wang, L. Xu, K. Chen. “Data-driven short-term forecasting for urban road network traffic based on data processing and LSTM-RNN”, Arabian Journal for Science and Engineering, 2017, 4(44): 3043-3060.

[12] D. Yang, K. Chen, M. Yang, X. Zhao. “Urban rail transit passenger flow forecast based on LSTM with enhanced long-term features”, IET Intelligent Transport Systems. 2019, 13(10): 1475-1482.

[13] X. Ma, Z. Tao, Y. Wang, et al. “Long short term memory neural network for traffic speed prediction using remote microwave sensor data”, Transportation Research Part C, 2015, 54: 187-197.