

Received March 13, 2021, accepted March 30, 2021, date of publication April 2, 2021, date of current version April 14, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3070625

Expressway Exit Traffic Flow Prediction for ETC and MTC Charging System Based on Entry Traffic Flows and LSTM Model

ZHU CHEN¹, BANGYU WU¹ , (Member, IEEE), BIN LI² , AND HOUSONG RUAN¹

¹School of Mathematics and Statistics, Xi'an Jiaotong University, Xi'an 710049, China

²School of Automobile, Chang'an University, Xi'an 710046, China

Corresponding author: Bangyu Wu (bangyuwu@xjtu.edu.cn)


This work was supported in part by the Natural Science Basic Research Program of Shaanxi under Grant 2020JM-018, in part by the National Research Program for Key Issues in Air Pollution Control of China under Grant DQGG0207-03, and in part by the National Key Research Project of China under Grant 2017YFC0212103.

ABSTRACT The Expressway (controlled-access highways) of China is the longest in the world and plays an important role in people's daily life. Accurate short-term traffic prediction is essential for travel schedule and active traffic management. There are two coexisting charging systems for expressway in China, Electronic Toll Collection (ETC) and Manual Toll Collection (MTC), which have different passing capacity and variation pattern. In this work, we demonstrate that the exit traffic flow prediction at Shanghai Xinqiao toll station using entry traffic flows from multiple close-related stations with Long Short-Term Memory (LSTM) model. Based on the origin-destination (OD) traffic data of a month, we present a new method to predict the exit station's traffic flow in the future 5 minutes. After deleting abnormal data, we select 12 of the 109 entry toll stations for the experiment. The traffic flow of these 12 entry stations account for 86% of the total exit traffic flow. This method uses the spatial-temporal matrix to deal with different three scenes that are ETC and MTC charging systems individually, the mix of ETC and MTC. We use the LSTM model with various lengths of flow sequence and amounts of hidden layer neurons for three different scenes. Lastly, we validate our model and carefully select the hyperparameters for better prediction accuracy by three evaluation metrics. The experimental results demonstrate that predicting the ETC is the best in the three scenes.

INDEX TERMS Short-term traffic flow prediction, spatial-temporal matrix, LSTM, OD traffic data, ETC and MTC.

I. INTRODUCTION

With the development of China's economy, the number of motor vehicles is increasing rapidly, which leads to a series of traffic problems: traffic congestions, traffic accidents, environmental pollution, and so on [1]. To alleviate these traffic problems, researchers pay more and more attention to the Intelligent Transportation Systems (ITS) [2]–[5], which are the set of applications and technological systems created with the aim of improving safety and efficiency in road transport. Accurate prediction of traffic information of the ITS (i.e., travel time, traffic speed, traffic flow, etc.) provides reliable real-time road-traffic information for travelers and managers, effectively reducing environmental pollution, and alleviating traffic congestions and traffic accidents [6]–[8]. Traffic flow prediction is an essential indicator of the ITS,

The associate editor coordinating the review of this manuscript and approving it for publication was Shaohua Wan .

which can be divided into three categories by the length of the forecast period: short-term forecast, medium-term forecast, and long-term forecast, respectively [9]. The prediction of the traffic flow of less than one hour is often called short-term traffic flow forecasting [10].

In this paper, based on origin-destination (OD) data of the Xinqiao toll station in Shanghai, China, we use the Long Short-Term Memory (LSTM) model for short-term traffic flow prediction. Main contributions of the paper are presented as follows:

- 1) We utilize the real traffic data for the model. Significantly, we find that the traffic flow of ETC is different with MTC, where the MTC has more peak values than the ETC. Therefore, separating the ETC and MTC datasets will improve the prediction accuracy naturally.
- 2) We use multi entry stations' traffic data to predict the corresponding exit station's traffic flow in the future 5 minutes.

- 3) We use the LSTM model with various lengths of flow sequence and amounts of hidden layer neurons for three different scenes that are ETC and MTC charging systems individually, the mix of ETC and MTC. The experimental results demonstrate that the ETC is the best in the three scenes. The RMSE value of ETC, of mixed ETC and MTC, and MTC are 13.15, 14.37, and 16.23.

The rest of this paper is organized as follows. Section 2 summarizes the existing literature on short-term traffic flow prediction. Section 3 presents the LSTM architecture and builds the traffic flow prediction network. Section 4 is the data description and data preprocessing. Section 5 gives the experiment results. Finally, section 6 concludes the paper.

II. LITERATURE REVIEW

Over the past few decades, many scholars have been proposed many methods to predict the short-term traffic flow. The existing methods mainly can be classified into three categories: parametric approach, non-parametric approach and hybrid approach [11].

The parametric approach includes time-series methods and Kalman filtering [12]–[14]. The most used time series models are the autoregressive integrated moving average (ARIMA) model [15] and its many variants: the Time-Oriented ARIMA [16], seasonal autoregressive integrated moving average (SARIMA) [17], Kohonen-ARIMA [18], subset ARIMA [19].

However, due to the stochastic and nonlinear feature for traffic flow [20], more researchers pay attention to non-parametric approaches in the traffic flow prediction field, such as k-nearest neighbor (k-NN) [21], Bayesian network [22], support vector regression (SVR) [23], support vector machine (SVM) [24]. Hybrid models combine the advantages of different methods. For example, Tan *et al.* [25] proposed an aggregation approach for short-term traffic flow prediction which was based on the moving average (MA), exponential smoothing (ES), ARIMA, and neural network (NN) models. Hong *et al.* [26] combined SVR with the algorithm-simulated annealing algorithm (GA-SA). Tang *et al.* [27] integrated the Fuzzy C-Means (FCM) with the Genetic Algorithm (GA) for missing traffic volume data estimation. Chen *et al.* [28] examined the impact of periodic component on three statistical models (i.e., space time (ST) model, vector autoregressive (VAR) model, ARIMA model) and three machine learning approaches (i.e., SVM model, multi-layer perceptron (MLP) model, recurrent neural network (RNN) model), results showed that the proposed hybrid prediction approach is effective for both statistical and machine learning models in short-term speed prediction.

However, compared to the hybrid models, a single model is simpler in application and faster in calculation, and the prediction accuracy can satisfy the requirements in actual engineering applications. With the development of artificial

intelligence, short-term traffic forecast based on deep learning approach has become a new trend [29]. Deep learning methods can present traffic features without prior knowledge due to the extraordinary ability to capture the uncertainty and complex nonlinearity [30], which achieve better performance for traffic parameters prediction [2]. Common deep learning algorithms include deep residual networks, cyclic neural networks, and convolutional neural networks, which self-learning capability highlights the features of short-term traffic flow prediction [31]. A variety of deep learning methods are used in traffic forecasting. Such as feed forward neural network (FFNN) [32], backpropagation (BP) neural network [33], stacked autoencoder (SAE) model [2], radial basis function (RBF) network [34]. Zhang *et al.* [35] proposed Deep Multi-Scale Learning Model for trajectory classification and understand the mobility of moving objects based on classification results. The recurrent neural network (RNN) is considered a suitable method to capture the temporal and spatial evolution of traffic flow among deep learning models on traffic flow prediction [36]. However, traditional RNNs suffer a major drawback of gradient vanishing or exploding [37]. To solve the problem, Hochreiter and Schmidhuber [38] proposed long short-term memory (LSTM) architecture. The latest research found that the LSTM has a high prediction effect in the short-term traffic flow prediction. Yu *et al.* [39] focus on forecasting the short-term traffic based on the data of private cars and minibuses operating on the Chang Tai highway with the deep learning method – LSTM. The prediction results are outstanding not only under the workday and the weekend but also under the unusual traffic status include the festival and the rainy day. Poonia *et al.* [40] applied the LSTM model on four months of data collected by MNIT College, and the data aggregated into a 5-min interval. Ma *et al.* [41] compared LSTM with three different typologies of the RNN, other nonparametric and parametric approaches, which are SVM, Time Series, and Kalman Filter, for urban travel time prediction. The numerical experiments demonstrate that the LSTM outperforms other algorithms in terms of accuracy and stability. Tian *et al.* [42] applied LSTM to short-term traffic flow prediction comparing with other models, such as random walk (RW), SVM, FFNN, and SAE. The result showed that the LSTM prediction model achieved the highest accuracy and generalized best among these models. Wang *et al.* [43] used the LSTM and the gated recurrent unit (GRU) models on trucks' GPS data. As a result, the average prediction accuracy throughout both peak and off-peak periods, LSTM is better than GRU with improved accuracy of 4.1%. Li *et al.* [44] compared different models which are constructed by different RNN layers. Experiment results indicate the performances of LSTM is better than GRU model.

III. METHODOLOGY

The LSTM architecture consists of one input layer, one recurrent hidden layer, and one output layer [49]. Additionally, the basic unit of the recurrent hidden layer is a memory block. Each block includes one or more self-connected memory

cells and three multiplicative units: the input gate, output gate, and forget gate—that provide continuous analogs of write, read and reset operations for the cells [45].

Short-term prediction of expressway exiting flow can help to recognize the future traffic state of the traffic network [29]. In this paper, based on the previous and current entry stations' traffic data, we predict the next time period's traffic flow of the exit station. First, we get a spatial-temporal matrix for entry stations' traffic data. Second, based on this matrix, we use the LSTM network to traffic flow prediction under different scenes: the mixed ETC and MTC, ETC, and MTC. The detail of the methodology is described in this section.

A. TRAFFIC FLOW FORECAST RELATIONSHIP BETWEEN ENTRY STATIONS AND THE EXIT STATION

The exit station's traffic flow not only relies on the sequential patterns in the temporal dimension but also depends on corresponding entry stations in the spatial dimension. Temporal correlation occurs between the traffic flow of the next time interval and the historical traffic data. Spatial correlation refers to the exit station's traffic flow relies on its corresponding entry stations' traffic flow.

In this paper, we select the Q entry stations and set the length of flow sequence as T . We show the prediction relationship for the ETC and MTC datasets. The prediction correlation of ETC and MTC datasets is shown in Figure 1.

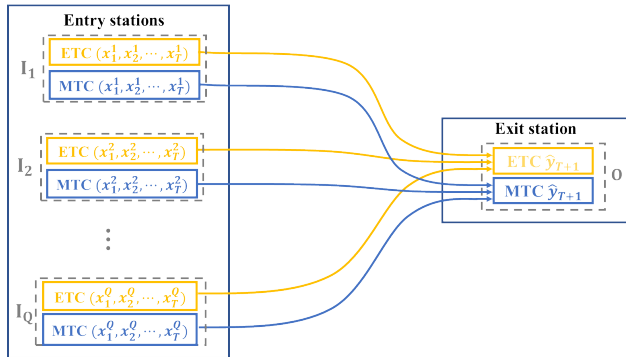


FIGURE 1. The spatial-temporal of ETC and MTC datasets. $I_i (i = 1, \dots, Q)$ represents the i_{th} entry station, O is the exit station. $(x_1^i, x_2^i, \dots, x_T^i, \dots, x_T^i)'$ ($t = 1, \dots, T$) is the traffic flow of the i_{th} entry station at the previous time step t . Every yellow arrowed line represents the ETC toll station, and each blue arrowed line refers to the MTC toll station.

B. INPUT MATRIX OF THE MODEL

Through the analysis, the traffic flow of the previous time t at entry stations as input. We write the input matrix as:

$$\begin{pmatrix} x_1^1 & x_1^2 & \dots & x_1^Q \\ x_2^1 & x_2^2 & \dots & x_2^Q \\ \vdots & \vdots & \ddots & \vdots \\ x_T^1 & x_T^2 & \dots & x_T^Q \end{pmatrix} \quad (1)$$

The matrix $\in \mathbb{R}^{T \times Q}$, T is the temporal length, and Q is the spatial length. $x_t^i (t \in [1, T], i \in [1, Q])$ is the

i_{th} entry station's traffic data at the time period t . The matrix is determined by the time period and works as input data in the LSTM network.

It is necessary to normalize the feature data because it is hard to convergent when the value range of some features are varying widely [44]. In this paper, we utilize linear normalization to traffic data to a range from 0 to 1. The normalization function defined as below.

$$norm(x_t^i) = \frac{x_t^i - \min(x_t^i)}{\max(x_t^i) - \min(x_t^i)}. \quad (2)$$

C. SHORT-TERM TRAFFIC FLOW PREDICTION MODEL BASE ON LSTM

In this paper, we use LSTM as our experimental model. The model's structure is composed with one input layer, one LSTM layer, one fully connected layer, and one output layer. The structure of the model is shown in Figure 2. The data flow at the first moment is shown in Figure 3.

The following equations show that how the LSTM network works.

$$f_t = \sigma([h_{t-1}, X_t] \cdot [W_{fh}, W_{fx}]' + [b_{fh}, b_{fx}]) \quad (3)$$

Equation (3) represents the Forget Gate. h_{t-1} is the hidden state of this layer at time $t-1$. X_t is normalized as the input at time t . W_{fh} and W_{fx} are weight matrices of the Forget Gate. b_{fh} and b_{fx} are bias vectors of the Forget Gate. The symbol \cdot represents multiplication of two matrices. σ is the sigmoid activation function.

$$i_t = \sigma([h_{t-1}, X_t] \cdot [W_{ih}, W_{ix}]' + [b_{ih}, b_{ix}]) \quad (4)$$

The activation of the Input Gate in the LSTM is described in Equation (4). W_{ih} and W_{ix} are weight matrices of the Input Gate. b_{ih} and b_{ix} are bias vector of the Input Gate.

$$\tilde{C}_t = \tanh([h_{t-1}, X_t] \cdot [W_{ch}, W_{cx}]' + [b_{ch}, b_{cx}]) \quad (5)$$

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \quad (6)$$

The output of the Cell State - C_t works out by both the Forget Gate and the Input Gate, where C_{t-1} is the output of the Cell State at time $t-1$, \tanh is the activation function.

$$o_t = \sigma([h_{t-1}, X_t] \cdot [W_{oh}, W_{ox}]' + [b_{oh}, b_{ox}]) \quad (7)$$

$$h_t = o_t * \tanh(C_t) \quad (8)$$

Finally, the output of the LSTM layer is h_t , which is achieved by the results of the Cell State and the Output Gate. The symbol $*$ represents the element-wise multiplication. W_{oh} and W_{ox} are weight matrices of the Output Gate.

Therefore, f_t, i_t, o_t, C_t , and $h_t \in \mathbb{R}^{1 \times u}$, W_{fh}, W_{ih}, W_{ch} , and $W_{oh} \in \mathbb{R}^{u \times u}$, W_{fx}, W_{ix}, W_{cx} , and $W_{ox} \in \mathbb{R}^{Q \times u}$.

In order to convert h_t into the predicted value, one fully connected layer is added after LSTM layer. The calculation formula is shown below:

$$\hat{y}_{T+1} = \sigma(h_T W_T + b_T) \quad (9)$$

where $W_T \in \mathbb{R}^{u \times 1}$, $b_T \in \mathbb{R}$.

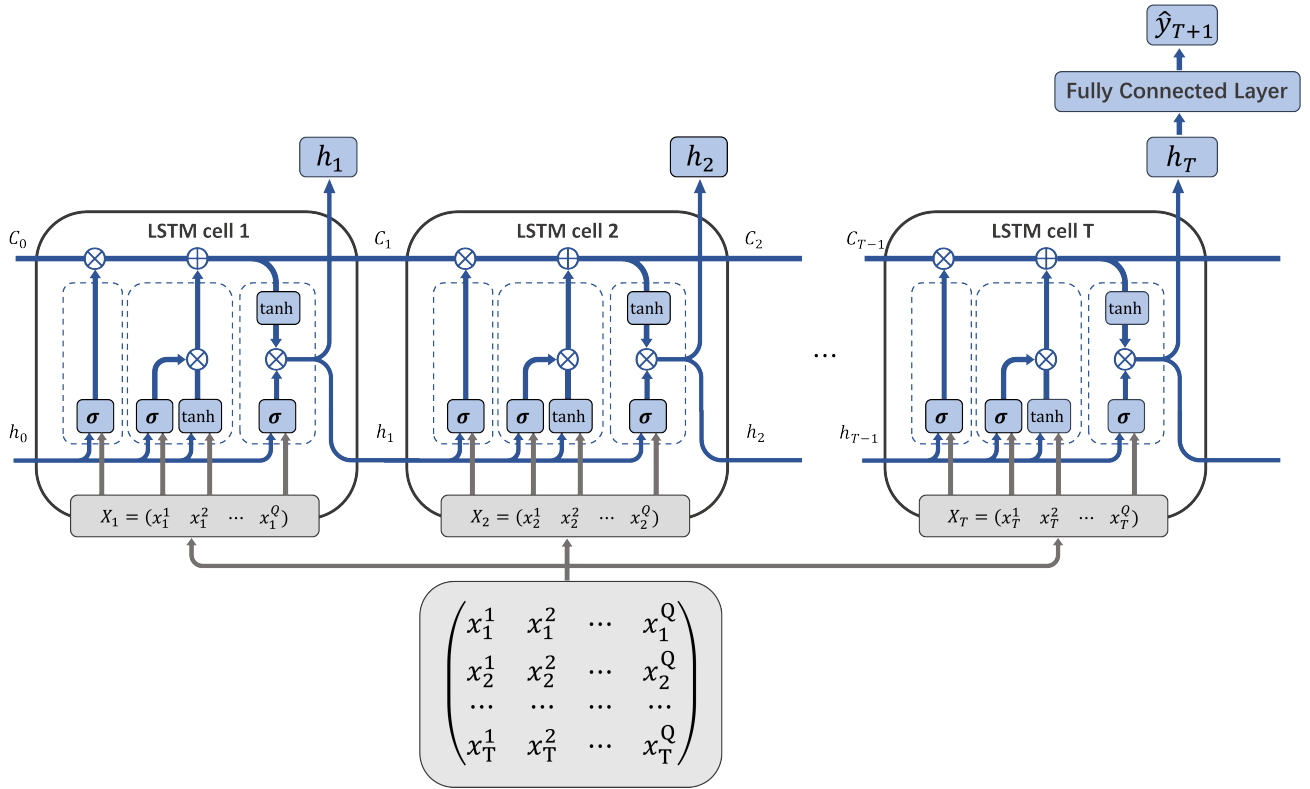


FIGURE 2. The experimental model—LSTM.

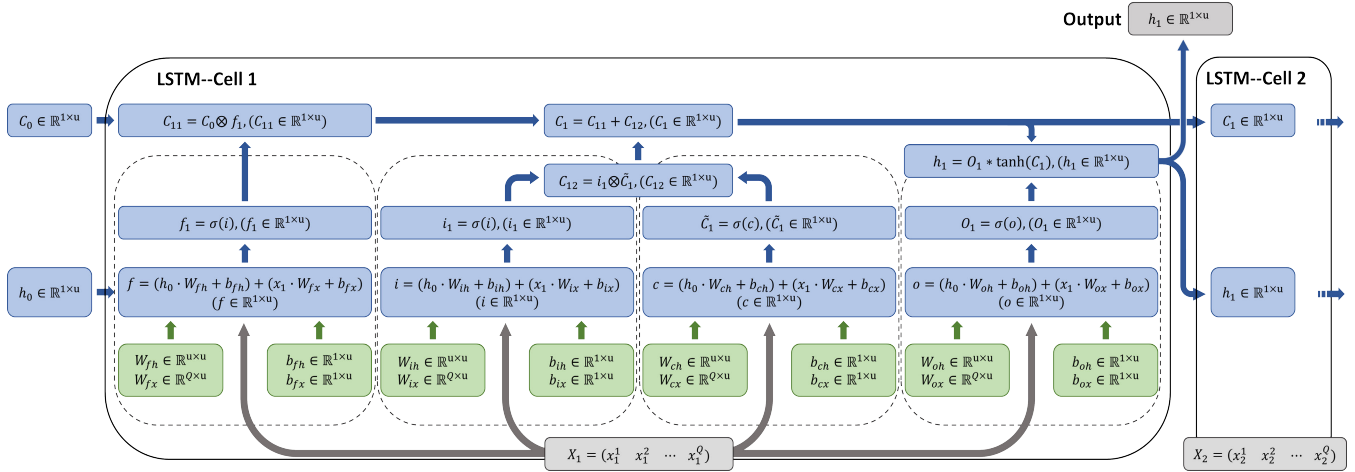


FIGURE 3. The details of data flow at the first moment, u represents the neuron numbers of the hidden layer.

Lastly, our model uses an effective gradient descent Adam model for fitting, Mean Square Error (MSE) loss function to optimize our model. Adam can be defined by the following equations [46].

$$m_n = \beta_1 m_{n-1} + (1 - \beta_1) g_n \quad (10)$$

$$v_n = \beta_2 v_{n-1} + (1 - \beta_2) g_n^2 \quad (11)$$

$$\hat{m}_n = \frac{m_n}{(1 - \beta_1^n)} \quad (12)$$

$$\hat{v}_n = \frac{v_n}{(1 - \beta_2^n)} \quad (13)$$

$$\theta_n = \theta_{n-1} - \alpha \cdot \frac{\hat{m}_n}{\sqrt{\hat{v}_n + \epsilon}} \quad (14)$$

where n represents the iterative time step. m_n represents the biased first-moment estimation, v_n is biased second raw moment estimation, \hat{m}_n is bias-corrected first-moment estimation, \hat{v}_n is the bias-corrected second raw moment estimation. θ_n is the parameter [49]. In this paper, the

TABLE 1. percentage of traffic flow from 12 main entry stations to the exit station.

| | | | |
|-------------|----------|----------|----------|
| ID | I_1 | I_2 | I_3 |
| Proportions | 1.8964 | 2.0967 | 1.5380 |
| ID | I_4 | I_5 | I_6 |
| Proportions | 3.9349 | 3.1285 | 2.8153 |
| ID | I_7 | I_8 | I_9 |
| Proportions | 13.3973 | 11.7363 | 14.0401 |
| ID | I_{10} | I_{11} | I_{12} |
| Proportions | 3.2091 | 23.9797 | 4.3032 |

parameters that need to be optimized with Adam are: $\theta_n = \{W_{fh}, W_{ih}, W_{ch}, W_{oh}, W_{fx}, W_{ix}, W_{cx}\}$ and W_{ox} . Furthermore, the settings for the parameters are $\alpha = 0.001$, $\beta_1 = 0.9$, $\beta_2 = 0.999$, $\epsilon = 10^{-8}$.

IV. DATA DESCRIPTION AND PREPROCESSING

A. DATA DESCRIPTION

In Shanghai, the most traffic volume appears in Xinqiao toll station, where congests heavily, especially in rush hours of morning and afternoon. This traffic network includes 109 entry stations and 1 corresponding exit station.

The expressway data contains some information for the vehicle entering and exiting expressway toll station including location, time, vehicle type, mileage, and whether the vehicle passes through the ETC toll lane.

In this study, the traffic data are aggregated into 5-min interval, 24 hours a day, from August 1st 2019 to August 31st 2019. Specifically, we choose 12 main stations from the 109 entry stations to predict the corresponding exit station's traffic flow in the next 5-min. 86% of this exit station's traffic flow comes from these 12 main entry stations. We calculate the percentage of traffic flow from each entry station to the exit station and show it in Table 1. The location of selected 12 entry toll stations and the exit station are marked on the map of Shanghai in Figure 4.

B. DATA PREPROCESSING

After the analysis, it is found that the original dataset has some abnormalities. Specifically, due to the damage to some detectors or traffic accidents, traffic data contains the negative, zero records of driving time, and zero mileage. These abnormal data might have a negative influence on the forecasting models and should be filtered out before forecasting [47]. We simply delete the abnormal data in the datasets. Finally, the deleted entry stations' traffic data records account for 3.35%. The deleted exit station's traffic data records account for 3.63%.

For directly visualization, we draw traffic flow curves of 12 entry stations and the exit station for three scenes: the mixed ETC and MTC, ETC, MTC from Figure 5 to Figure 8, respectively. Moreover, we also do some statistical analysis

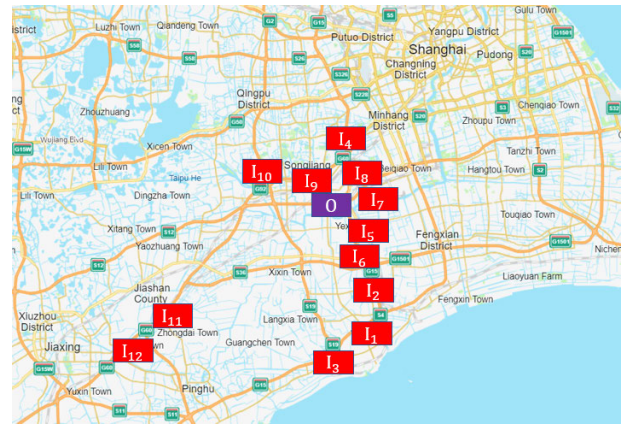


FIGURE 4. Locations of entry and the exit stations in the map. Red boxes represent the 12 entry stations. The purple box represents the exit station. (Map is from <https://cn.bing.com/maps/#>).

in Figure 9, which presents the box plots of the data. We can see that the variance of ETC is relative greater but the outlier of MTC is significantly more, especially the entry station I_8 , I_9 and I_{11} . In a word, the distributions of the ETC and MTC are different. Motivated by this, we are trying to predict the traffic flow by separating the ETC and MTC for better prediction accuracy.

In this study, the traffic flow of entry stations and the exit station are divided into two datasets — the front 75% data for training and the remaining 25% for testing.

V. EXPERIMENTS

To compare the performance among three different scenes, we use Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and Root Mean Square Error (RMSE) to evaluate our forecasting results. MAE can better reflect the actual situation of forecast error. MAPE represents the prediction error in a percentage, and it is the case of identifying the significant error when MAE is very small [18]. The equations are the following:

$$MAE = \frac{1}{m} \sum_{i=1}^m |\hat{y}_i - y_i| \quad (15)$$

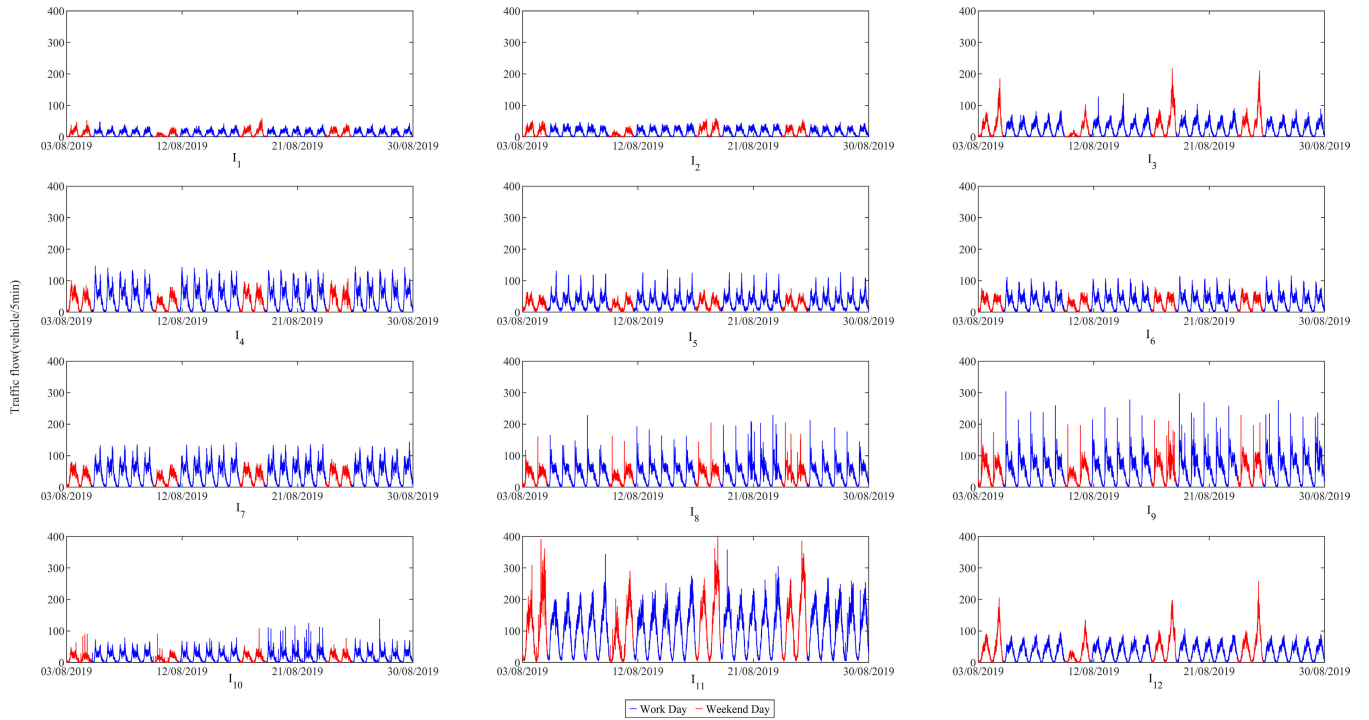


FIGURE 5. Traffic flow curves of 12 entry stations for mix ETC and MTC toll lanes.

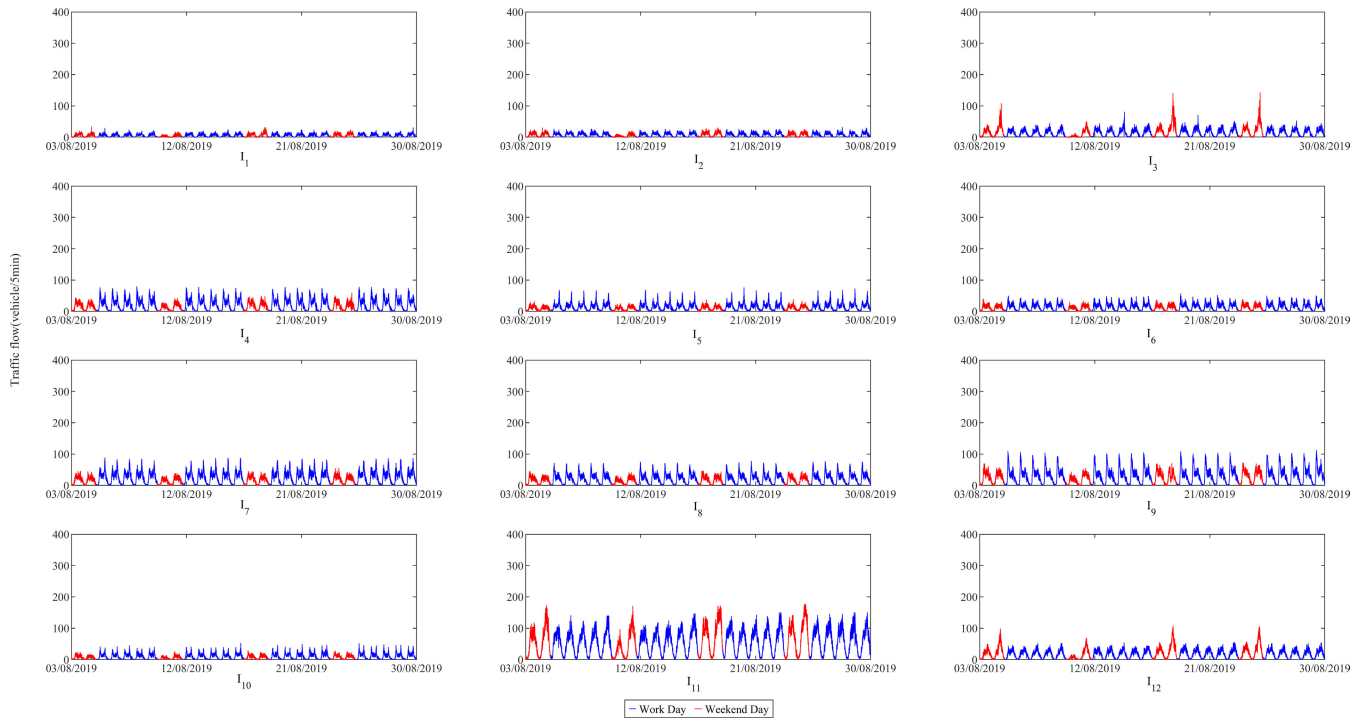


FIGURE 6. Traffic flow curves of 12 entry stations for ETC toll lane.

$$MAPE = \frac{100\%}{m} \sum_{i=1}^m \left| \frac{\hat{y}_i - y_i}{y_i} \right| \quad (16)$$

$$RMSE = \sqrt{\frac{1}{m} \sum_{i=1}^m (y_i - \hat{y}_i)^2} \quad (17)$$

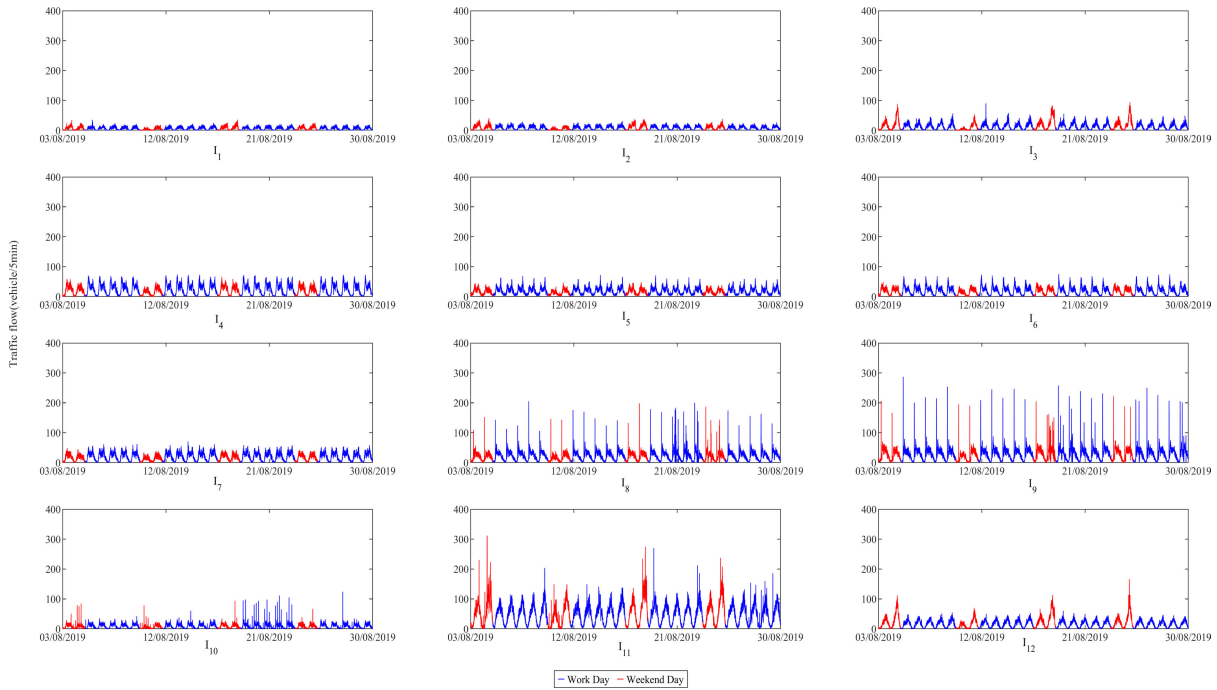


FIGURE 7. Traffic flow curves of 12 entry stations for MTC toll lane.

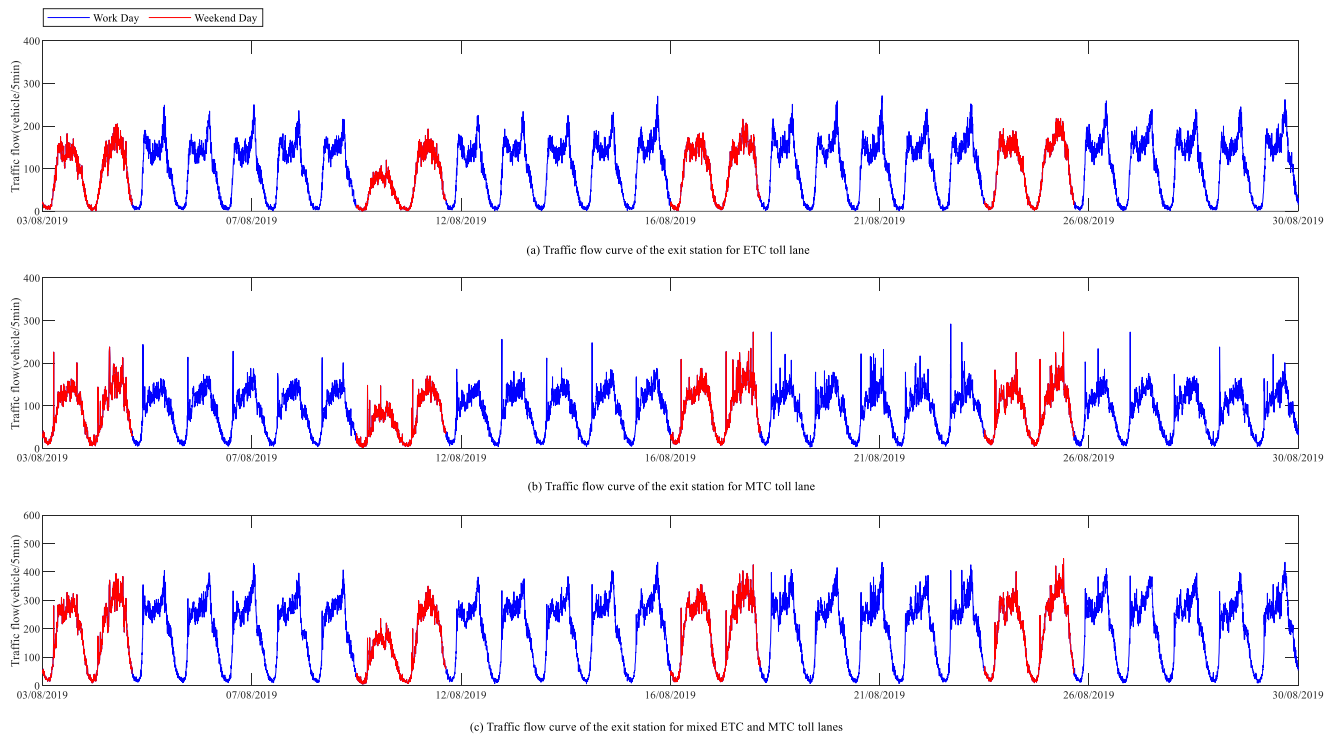


FIGURE 8. Traffic flow curves of the exit station for three scenes.

where y_i represents the ground truth, and \hat{y}_i is the prediction value. The lower the value of the evaluation indicator, the better the model performance.

We compare the LSTM network performance under different hyper parameters to indicate that our network architecture

is effective. Two hyper parameters are the length of flow sequence T and the amount of the hidden layer neuron u . The value range of T is set to range of $\{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$. The value range of u is $u = \{10, 30, 50, 70, 90, 110\}$.

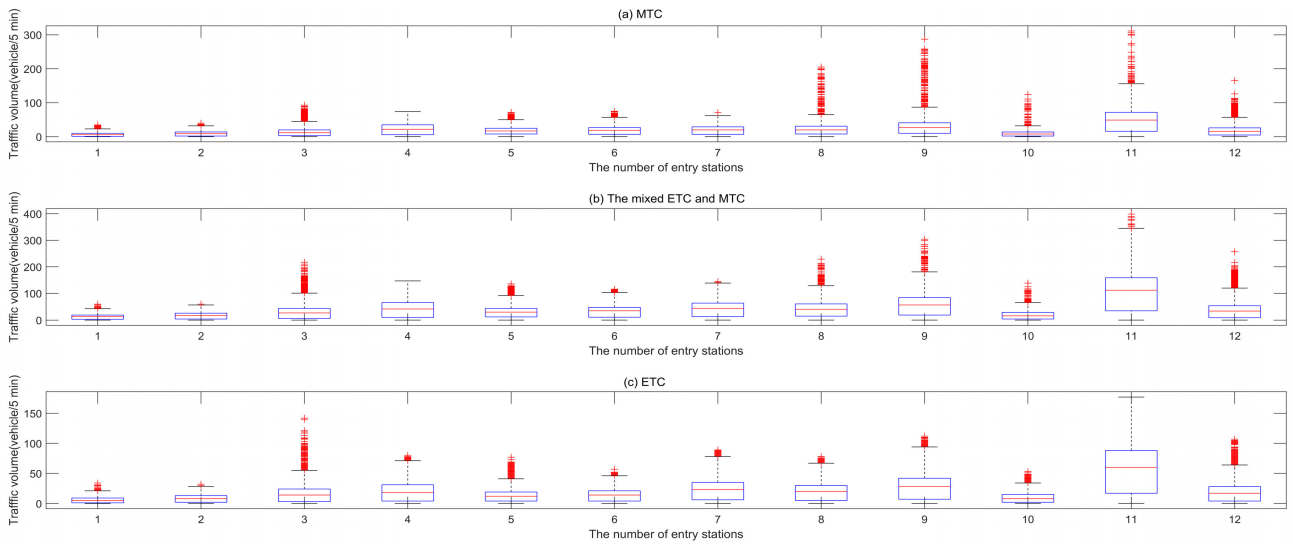


FIGURE 9. The boxplots of the 12 entry stations.

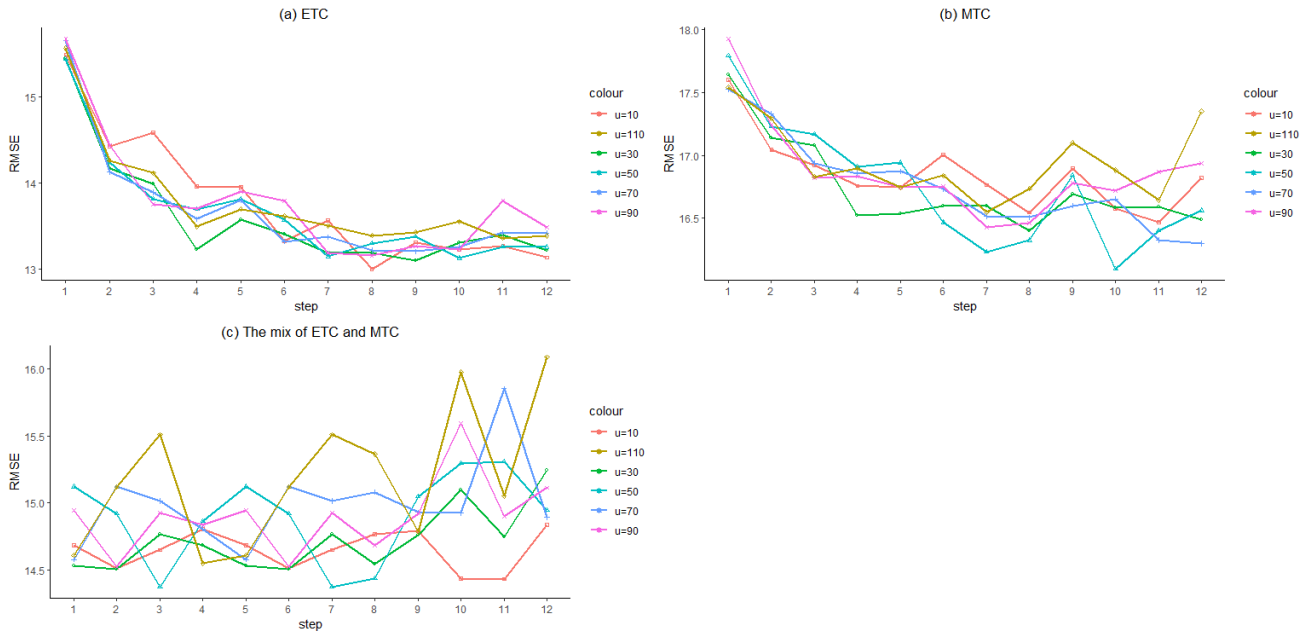


FIGURE 10. RMSE values of different T and u.

All our experiments are implemented with the Keras framework. We build the LSTM network, where the activation function of the LSTM layer is tanh and the output layer is a fully connected layer with the sigmoid activation function. We set the training epochs is to 100. For each T and u, we calculate the average of all the evaluation metrics for five times' experiments. The average results of RMSE, MAE and MAPE are shown in Table 2, Table 3 and Table 4, respectively. The best results are represented in bold font. Here we mainly focus on RMSE because that Chai [48] indicated that the RMSE usually is better at revealing model

performance. We also show the trend of RMSE results in Figure 10. All of these results show that: From Figure 10, dividing the traffic data into ETC and MTC can improve the prediction accuracy. Because the RMSE curves of the mixed ETC and MTC have no obvious downward trend. For ETC and MTC, the change trend of these RMSE curves is obvious.

- 1) From Figure 10, dividing the traffic data into ETC and MTC can improve the prediction accuracy. Because the RMSE curves of the mixed ETC and MTC have no obvious downward trend. For ETC and MTC, the change trend of these RMSE curves is obvious.

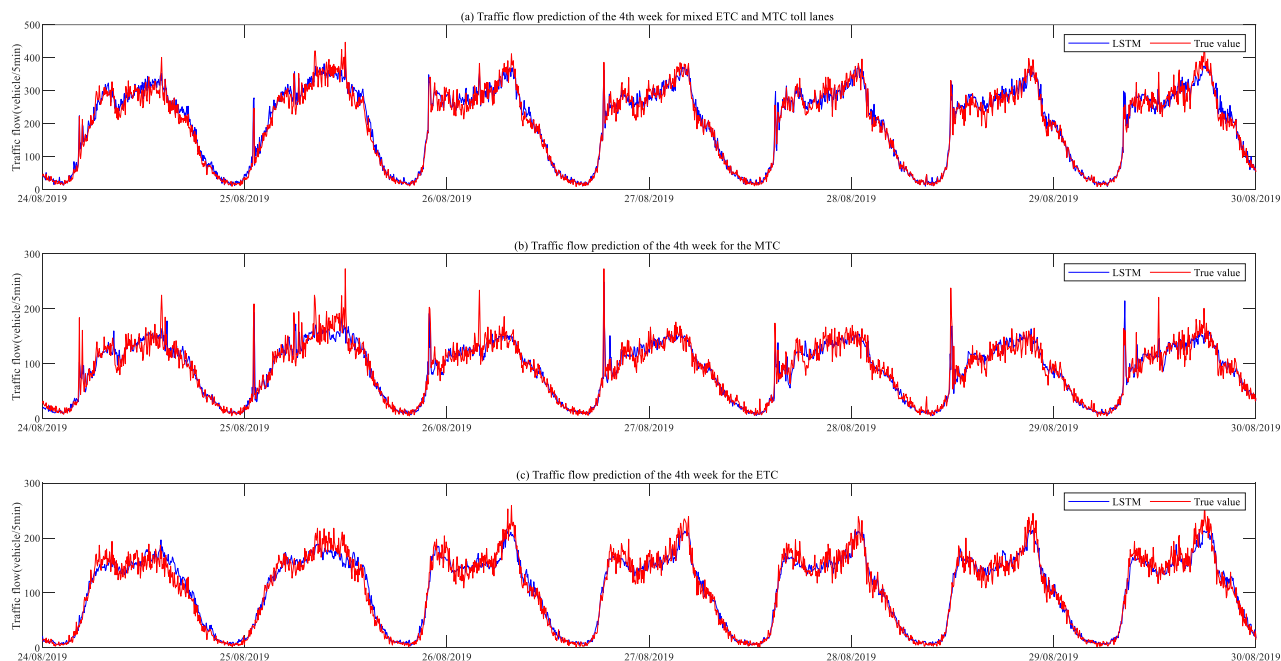


FIGURE 11. Prediction results for three scenes.

TABLE 2. Short-term traffic flow prediction result errors of RMSE.

| | The mix of ETC and MTC | | | | | | ETC | | | | | | MTC | | | | | |
|------|------------------------|-------|--------------|-------|-------|-------|-------|-------|--------------|-------|-------|-------|-------|-------|--------------|-------|-------|-------|
| | u=10 | u=30 | u=50 | u=70 | u=90 | u=110 | u=10 | u=30 | u=50 | u=70 | u=90 | u=110 | u=10 | u=30 | u=50 | u=70 | u=90 | u=110 |
| T=1 | 14.68 | 14.53 | 15.12 | 14.58 | 14.95 | 14.61 | 15.49 | 15.45 | 15.44 | 15.66 | 15.67 | 15.57 | 17.60 | 17.65 | 17.79 | 17.53 | 17.93 | 17.54 |
| T=2 | 14.51 | 14.51 | 14.92 | 15.12 | 14.53 | 15.12 | 14.43 | 14.17 | 14.24 | 14.13 | 14.44 | 14.26 | 17.04 | 17.14 | 17.23 | 17.33 | 17.24 | 17.29 |
| T=3 | 14.65 | 14.76 | 14.37 | 15.02 | 14.93 | 15.51 | 14.58 | 13.99 | 13.81 | 13.89 | 13.75 | 14.12 | 16.92 | 17.08 | 17.17 | 16.94 | 16.82 | 16.83 |
| T=4 | 14.81 | 14.68 | 14.86 | 14.81 | 14.84 | 14.55 | 13.96 | 13.23 | 13.70 | 13.58 | 13.70 | 13.49 | 16.76 | 16.52 | 16.91 | 16.86 | 16.83 | 16.90 |
| T=5 | 14.68 | 14.53 | 15.12 | 14.58 | 14.95 | 14.61 | 13.95 | 13.58 | 13.82 | 13.80 | 13.90 | 13.70 | 16.75 | 16.54 | 16.94 | 16.87 | 16.75 | 16.75 |
| T=6 | 14.51 | 14.51 | 14.92 | 15.12 | 14.53 | 15.12 | 13.33 | 13.40 | 13.57 | 13.32 | 13.79 | 13.61 | 17.01 | 16.60 | 16.47 | 16.74 | 16.75 | 16.84 |
| T=7 | 14.65 | 14.76 | 14.37 | 15.02 | 14.93 | 15.51 | 13.56 | 13.19 | 13.15 | 13.37 | 13.19 | 13.50 | 16.76 | 16.60 | 16.23 | 16.52 | 16.43 | 16.55 |
| T=8 | 14.77 | 14.54 | 14.44 | 15.08 | 14.68 | 15.37 | 13.00 | 13.19 | 13.30 | 13.22 | 13.16 | 13.39 | 16.55 | 16.40 | 16.33 | 16.51 | 16.46 | 16.74 |
| T=9 | 14.79 | 14.76 | 15.05 | 14.93 | 14.92 | 14.78 | 13.31 | 13.10 | 13.38 | 13.20 | 13.27 | 13.43 | 16.90 | 16.69 | 16.84 | 16.60 | 16.78 | 17.10 |
| T=10 | 14.43 | 15.10 | 15.30 | 14.93 | 15.59 | 15.98 | 13.22 | 13.31 | 13.13 | 13.26 | 13.23 | 13.55 | 16.58 | 16.58 | 16.10 | 16.66 | 16.72 | 16.88 |
| T=11 | 14.43 | 14.75 | 15.31 | 15.85 | 14.90 | 15.05 | 13.26 | 13.39 | 13.26 | 13.42 | 13.79 | 13.36 | 16.47 | 16.59 | 16.40 | 16.33 | 16.87 | 16.65 |
| T=12 | 14.84 | 15.24 | 14.95 | 14.89 | 15.11 | 16.09 | 13.13 | 13.22 | 13.25 | 13.42 | 13.48 | 13.39 | 16.82 | 16.49 | 16.56 | 16.30 | 16.94 | 17.35 |

TABLE 3. Short-term traffic flow prediction result errors of MAE.

| | The mix of ETC and MTC | | | | | | ETC | | | | | | MTC | | | | | |
|------|------------------------|------|-------------|------|------|-------|------|------|-------------|------|------|-------|------|------|-------------|------|------|-------|
| | u=10 | u=30 | u=50 | u=70 | u=90 | u=110 | u=10 | u=30 | u=50 | u=70 | u=90 | u=110 | u=10 | u=30 | u=50 | u=70 | u=90 | u=110 |
| T=1 | 3.42 | 3.40 | 3.44 | 3.40 | 3.39 | 3.40 | 3.36 | 3.38 | 3.33 | 3.37 | 3.38 | 3.36 | 3.38 | 3.37 | 3.38 | 3.37 | 3.40 | 3.37 |
| T=2 | 3.38 | 3.35 | 3.39 | 3.39 | 3.39 | 3.36 | 3.26 | 3.22 | 3.23 | 3.23 | 3.26 | 3.24 | 3.31 | 3.31 | 3.31 | 3.31 | 3.31 | 3.31 |
| T=3 | 3.34 | 3.34 | 3.31 | 3.34 | 3.31 | 3.28 | 3.30 | 3.20 | 3.21 | 3.22 | 3.20 | 3.18 | 3.30 | 3.30 | 3.31 | 3.29 | 3.27 | 3.27 |
| T=4 | 3.27 | 3.27 | 3.30 | 3.28 | 3.29 | 3.27 | 3.21 | 3.13 | 3.18 | 3.18 | 3.15 | 3.16 | 3.29 | 3.27 | 3.28 | 3.29 | 3.30 | 3.29 |
| T=5 | 3.28 | 3.27 | 3.32 | 3.25 | 3.30 | 3.26 | 3.21 | 3.17 | 3.19 | 3.20 | 3.16 | 3.15 | 3.27 | 3.24 | 3.29 | 3.27 | 3.34 | 3.28 |
| T=6 | 3.27 | 3.26 | 3.27 | 3.31 | 3.25 | 3.32 | 3.16 | 3.15 | 3.16 | 3.14 | 3.16 | 3.16 | 3.28 | 3.27 | 3.26 | 3.27 | 3.30 | 3.29 |
| T=7 | 3.28 | 3.28 | 3.23 | 3.30 | 3.28 | 3.36 | 3.16 | 3.12 | 3.10 | 3.13 | 3.14 | 3.14 | 3.27 | 3.25 | 3.22 | 3.25 | 3.24 | 3.29 |
| T=8 | 3.30 | 3.25 | 3.24 | 3.30 | 3.27 | 3.33 | 3.11 | 3.10 | 3.12 | 3.12 | 3.13 | 3.12 | 3.24 | 3.22 | 3.23 | 3.24 | 3.24 | 3.26 |
| T=9 | 3.30 | 3.28 | 3.30 | 3.28 | 3.28 | 3.26 | 3.14 | 3.14 | 3.12 | 3.12 | 3.11 | 3.13 | 3.31 | 3.26 | 3.28 | 3.25 | 3.27 | 3.27 |
| T=10 | 3.26 | 3.30 | 3.32 | 3.28 | 3.33 | 3.37 | 3.13 | 3.12 | 3.11 | 3.13 | 3.12 | 3.17 | 3.25 | 3.26 | 3.24 | 3.25 | 3.24 | 3.28 |
| T=11 | 3.25 | 3.27 | 3.31 | 3.36 | 3.27 | 3.30 | 3.13 | 3.15 | 3.10 | 3.13 | 3.18 | 3.13 | 3.23 | 3.24 | 3.23 | 3.28 | 3.27 | 3.26 |
| T=12 | 3.29 | 3.32 | 3.28 | 3.30 | 3.29 | 3.39 | 3.12 | 3.14 | 3.13 | 3.15 | 3.16 | 3.13 | 3.24 | 3.24 | 3.23 | 3.23 | 3.28 | 3.28 |

- From Figure 10 (a) and (b), the RMSE decreases as the length of the flow sequence change from 1 to 7 and increases after 8. Obviously, 7 is the best length of the flow sequence for our model. When the value of u is from 70 to 110, the network has overfitting.
- Increasing the number of hidden layer neurons can typically help the network extract more features that

can improve the training accuracy. But too many neurons may cause the overfitting problems, which is verified in Figure 10. In our experiment, we choose 50 neurons in hidden layer for better prediction accuracy.

As a result, the length of flow sequence is set to 7 and the value of u is 50, which is best for the network.

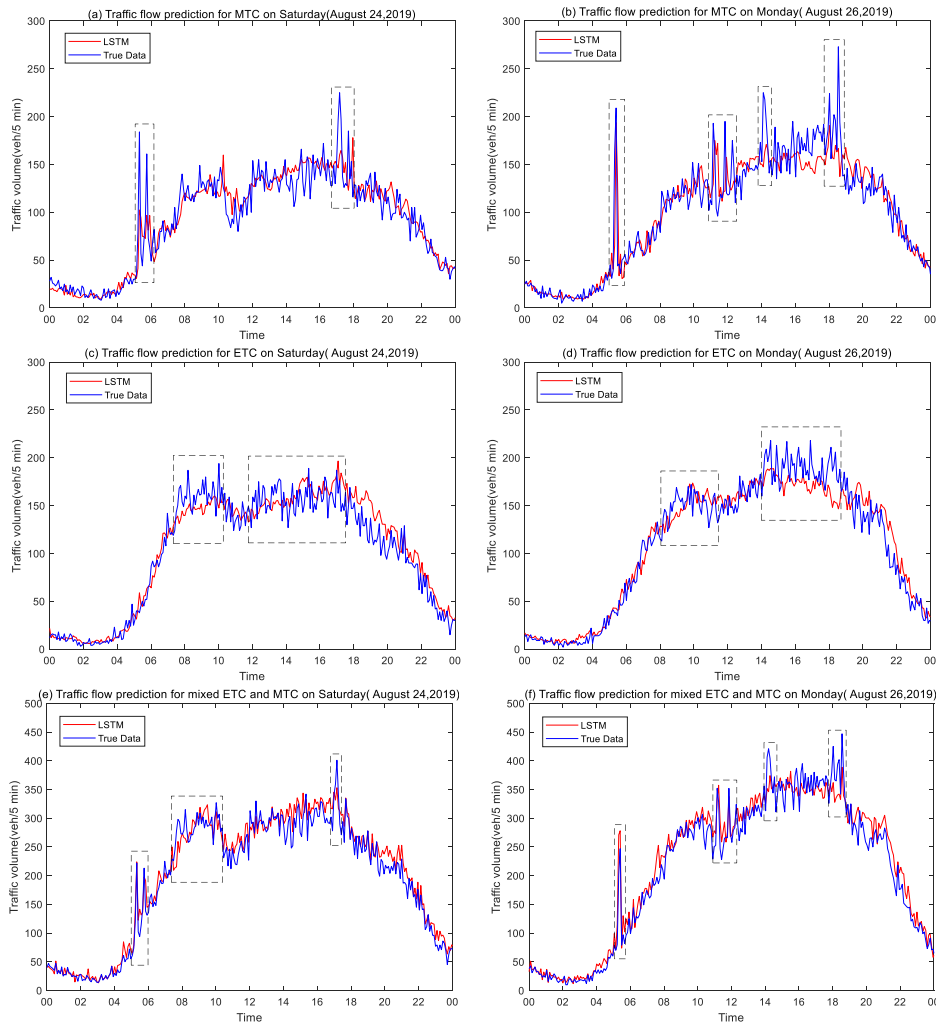


FIGURE 12. The highlighted peak hour of prediction results.

TABLE 4. Short-term traffic flow prediction result errors of MAPE.

| | The mix of ETC and MTC | | | | | | ETC | | | | | | MTC | | | | | |
|------|------------------------|------|-------------|------|------|-------|-------|-------|--------------|-------|-------|-------|-------|-------|--------------|-------|-------|-------|
| | u=10 | u=30 | u=50 | u=70 | u=90 | u=110 | u=10 | u=30 | u=50 | u=70 | u=90 | u=110 | u=10 | u=30 | u=50 | u=70 | u=90 | u=110 |
| T=1 | 7.81 | 7.72 | 7.79 | 7.85 | 7.73 | 7.70 | 17.57 | 17.66 | 17.51 | 17.82 | 17.51 | 17.71 | 16.59 | 16.60 | 16.52 | 16.48 | 16.38 | 16.55 |
| T=2 | 8.18 | 8.01 | 7.79 | 8.05 | 7.81 | 8.29 | 16.75 | 16.48 | 16.83 | 17.11 | 16.31 | 17.20 | 15.68 | 15.47 | 16.14 | 15.70 | 15.37 | 15.86 |
| T=3 | 8.46 | 8.18 | 7.53 | 7.66 | 7.84 | 7.52 | 16.90 | 16.00 | 16.46 | 17.94 | 16.58 | 16.03 | 15.27 | 15.31 | 15.25 | 15.34 | 14.81 | 15.96 |
| T=4 | 8.14 | 7.97 | 7.52 | 7.61 | 7.61 | 7.83 | 15.72 | 15.57 | 16.08 | 16.03 | 15.74 | 15.59 | 15.49 | 15.02 | 14.93 | 14.69 | 15.15 | 15.90 |
| T=5 | 8.03 | 7.93 | 7.47 | 7.56 | 7.74 | 7.47 | 15.87 | 16.38 | 15.94 | 15.86 | 16.04 | 15.69 | 15.43 | 15.16 | 15.07 | 14.60 | 16.07 | 15.35 |
| T=6 | 7.86 | 7.91 | 7.47 | 7.61 | 7.59 | 7.83 | 16.35 | 15.71 | 15.52 | 17.29 | 16.23 | 16.15 | 15.11 | 14.82 | 15.35 | 15.13 | 15.21 | 15.02 |
| T=7 | 8.08 | 7.95 | 7.42 | 7.68 | 7.48 | 8.37 | 16.35 | 15.00 | 15.56 | 15.33 | 16.34 | 15.49 | 14.79 | 14.96 | 14.57 | 14.92 | 14.56 | 15.50 |
| T=8 | 8.50 | 7.87 | 7.46 | 7.68 | 7.48 | 8.12 | 15.54 | 15.33 | 15.82 | 15.35 | 15.70 | 15.76 | 15.10 | 14.53 | 14.37 | 14.56 | 15.27 | 15.84 |
| T=9 | 8.17 | 8.45 | 7.53 | 7.90 | 7.72 | 8.09 | 15.70 | 15.79 | 15.01 | 15.65 | 15.24 | 17.64 | 15.05 | 14.85 | 14.71 | 15.03 | 15.37 | 15.07 |
| T=10 | 8.33 | 7.83 | 7.63 | 7.42 | 7.72 | 7.91 | 15.68 | 15.09 | 14.90 | 15.14 | 15.05 | 15.81 | 15.34 | 15.38 | 16.30 | 15.24 | 14.71 | 14.90 |
| T=11 | 8.12 | 7.91 | 8.09 | 7.58 | 7.76 | 8.10 | 15.92 | 15.84 | 15.03 | 15.34 | 14.92 | 15.80 | 14.65 | 14.77 | 14.51 | 15.83 | 14.91 | 14.75 |
| T=12 | 7.99 | 7.93 | 7.48 | 8.49 | 7.50 | 7.80 | 15.21 | 16.29 | 16.55 | 15.25 | 15.03 | 15.03 | 15.45 | 14.51 | 14.54 | 15.66 | 15.12 | 14.62 |

To show the prediction results in more detail, we highlight the peak hours for prediction results on Monday and Saturday. From the Figure 12, MTC dataset has more peak hours in workdays. As a result, MTC may cause the traffic congestion. Therefore, it's necessary for the government to develop more ETC lanes that traffic management department can alleviate the traffic congestion.

VI. CONCLUSION

In this paper, we focus on the short-term traffic flow prediction based on the traffic data of Xinqiao toll station in Shanghai of China and deep learning method — LSTM. Our application of the LSTM model to the actual traffic data has largely proved that the purpose of creating the model is to adapt to the actual data.

First, we divided the actual traffic data into ETC and MTC sequences. Second, we try different time periods in the three scenes. Finally, we get the best time period is 7, and the experimental results demonstrate that the ETC is the best in the three scenes.

From our results, dividing the traffic data into ETC and MTC condition can improve the prediction accuracy. With the increasing portion of develops the ETC toll lanes, the traffic efficiency of the station is enhanced and we can improve the accuracy for traffic flow prediction. In a word, our study is of great significance to the industry.

REFERENCES

- [1] J. Zhang, F.-Y. Wang, K. Wang, W.-H. Lin, X. Xu, and C. Chen, "Data-driven intelligent transportation systems: A survey," *IEEE Trans. Intell. Transp. Syst.*, vol. 12, no. 4, pp. 1624–1639, Dec. 2011, doi: [10.1109/TITS.2011.2158001](https://doi.org/10.1109/TITS.2011.2158001).
- [2] Y. Lv, Y. Duan, W. Kang, Z. Li, and F.-Y. Wang, "Traffic flow prediction with big data: A deep learning approach," *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 2, pp. 865–873, Apr. 2015, doi: [10.1109/TITS.2014.2345663](https://doi.org/10.1109/TITS.2014.2345663).
- [3] S. Fernandez and T. Ito, "Driver classification for intelligent transportation systems using fuzzy logic," in *Proc. IEEE 19th Int. Conf. Intell. Transp. Syst. (ITSC)*, Nov. 2016, pp. 1212–1216.
- [4] C. Chen, B. Liu, S. Wan, P. Qiao, and Q. Pei, "An edge traffic flow detection scheme based on deep learning in an intelligent transportation system," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 3, pp. 1840–1852, Mar. 2021, doi: [10.1109/TITS.2020.3025687](https://doi.org/10.1109/TITS.2020.3025687).
- [5] S. Wan, X. Xu, T. Wang, and Z. Gu, "An intelligent video analysis method for abnormal event detection in intelligent transportation systems," *IEEE Trans. Intell. Transp. Syst.*, early access, Sep. 9, 2020, doi: [10.1109/TITS.2020.3017505](https://doi.org/10.1109/TITS.2020.3017505).
- [6] X. Yang, Y. Zou, J. Tang, J. Liang, and M. Ijaz, "Evaluation of short-term freeway speed prediction based on periodic analysis using statistical models and machine learning models," *J. Adv. Transp.*, vol. 2020, pp. 1–16, Jan. 2020, doi: [10.1155/2020/9628957](https://doi.org/10.1155/2020/9628957).
- [7] X. Miao, B. Wu, Y. Zou, and L. Wu, "Examining the impact of different periodic functions on short-term freeway travel time prediction approaches," *J. Adv. Transp.*, vol. 2020, pp. 1–15, Aug. 2020.
- [8] A. E. Essien, G. Chukwkelu, and C. Giannetti, "A scalable deep convolutional LSTM neural network for large-scale urban traffic flow prediction using recurrence plots," in *Proc. IEEE AFRICON*, Accra, Ghana, Sep. 2019, pp. 1–7.
- [9] Z. Hou and X. Li, "Repeatability and similarity of freeway traffic flow and long-term prediction under big data," *IEEE Trans. Intell. Transp. Syst.*, vol. 17, no. 6, pp. 1786–1796, Jun. 2016.
- [10] M. Bennet Wondimagegnehu and G. Alemu, "Short time traffic flow identification and prediction at road junctions using recurrent multilayer perceptron," in *Proc. IEEE AFRICON*, Sep. 2017, pp. 15–20.
- [11] A. Koesdwiary, R. Soua, and F. Karray, "Improving traffic flow prediction with weather information in connected cars: A deep learning approach," *IEEE Trans. Veh. Technol.*, vol. 65, no. 12, pp. 9508–9517, Dec. 2016.
- [12] O. Iwao and S. J. Yorgos, "Dynamic prediction of traffic flow through Kalman filtering theory," *Transp. Res. B, Methodol.*, vol. 18, no. 1, pp. 1–11, Feb. 1984, doi: [10.1016/0191-2615\(84\)90002-X](https://doi.org/10.1016/0191-2615(84)90002-X).
- [13] Y.-S. Gong and Y. Zhang, "Research of short-term traffic volume prediction based on Kalman filtering," in *Proc. 6th Int. Conf. Intell. Netw. Intell. Syst.*, Shenyang, China, Nov. 2013, pp. 99–102, doi: [10.1109/ICINIS.2013.32](https://doi.org/10.1109/ICINIS.2013.32).
- [14] Y. Kamarianakis and P. Prastacos, "Forecasting traffic flow conditions in an urban network: Comparison of multivariate and univariate approaches," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 1857, no. 1, pp. 74–84, Jan. 2003, doi: [10.3141/1857-09](https://doi.org/10.3141/1857-09).
- [15] B. M. Williams, "Multivariate vehicular traffic flow prediction: Evaluation of ARIMAX modeling," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 1776, no. 1, pp. 194–200, Jan. 2001, doi: [10.3141/1776-25](https://doi.org/10.3141/1776-25).
- [16] H. Dong, L. Jia, X. Sun, C. Li, and Y. Qin, "Road traffic flow prediction with a time-oriented ARIMA model," in *Proc. 5th Int. Joint Conf. INC, IMS IDC*, Seoul, South Korea, Aug. 2009, pp. 1649–1652, doi: [10.1109/NCM.2009.224](https://doi.org/10.1109/NCM.2009.224).
- [17] B. M. Williams and L. A. Hoel, "Modeling and forecasting vehicular traffic flow as a seasonal ARIMA process: Theoretical basis and empirical results," *J. Transp. Eng.*, vol. 129, no. 6, pp. 664–672, Nov. 2003, doi: [10.1061/\(ASCE\)0733-947X\(2003\)129:6\(664\)](https://doi.org/10.1061/(ASCE)0733-947X(2003)129:6(664)).
- [18] M. Van Der Voort, M. Dougherty, and S. Watson, "Combining Kohonen maps with ARIMA time series models to forecast traffic flow," *Transp. Res. C, Emerg. Technol.*, vol. 4, no. 5, pp. 307–318, 1996, doi: [10.1016/S0968-090X\(97\)82903-8](https://doi.org/10.1016/S0968-090X(97)82903-8).
- [19] S. Lee and D. B. Fambro, "Application of subset autoregressive integrated moving average model for short-term freeway traffic volume forecasting," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 1678, no. 1, pp. 179–188, Jan. 1999, doi: [10.3141/1678-22](https://doi.org/10.3141/1678-22).
- [20] H. Zheng, F. Lin, X. Feng, and Y. Chen, "A hybrid deep learning model with attention-based conv-LSTM networks for short-term traffic flow prediction," *IEEE Trans. Intell. Transp. Syst.*, early access, Jun. 9, 2020, doi: [10.1109/TITS.2020.2997352](https://doi.org/10.1109/TITS.2020.2997352).
- [21] T. Zhang, L. Hu, Z. Liu, and Y. Zhang, "Nonparametric regression for the short-term traffic flow forecasting," in *Proc. Int. Conf. Mechanic Autom. Control Eng.*, Wuhan, China, Jun. 2010, pp. 2850–2853, doi: [10.1109/MACE.2010.5536292](https://doi.org/10.1109/MACE.2010.5536292).
- [22] S. Sun, C. Zhang, and G. Yu, "A Bayesian network approach to traffic flow forecasting," *IEEE Trans. Intell. Transp. Syst.*, vol. 7, no. 1, pp. 124–132, Mar. 2006, doi: [10.1109/TITS.2006.869623](https://doi.org/10.1109/TITS.2006.869623).
- [23] Y. S. Jeong, Y. J. Byon, M. M. Castro-Neto, and S. M. Easa, "Supervised weighting-online learning algorithm for short-term traffic flow prediction," *IEEE Trans. Intell. Transp. Syst.*, vol. 14, no. 4, pp. 1700–1707, Dec. 2013, doi: [10.1109/TITS.2013.2267735](https://doi.org/10.1109/TITS.2013.2267735).
- [24] Y. Zhang and Y. Liu, "Traffic forecasting using least squares support vector machines," *Transportmetrica*, vol. 5, no. 3, pp. 193–213, Sep. 2009.
- [25] M.-C. Tan, S. C. Wong, J.-M. Xu, Z.-R. Guan, and P. Zhang, "An aggregation approach to short-term traffic flow prediction," *IEEE Trans. Intell. Transp. Syst.*, vol. 10, no. 1, pp. 60–69, Mar. 2009, doi: [10.1109/TITS.2008.2011693](https://doi.org/10.1109/TITS.2008.2011693).
- [26] W.-C. Hong, Y. Dong, F. Zheng, and S. Y. Wei, "Hybrid evolutionary algorithms in a SVR traffic flow forecasting model," *Appl. Math. Comput.*, vol. 217, no. 15, pp. 6733–6747, Apr. 2011.
- [27] J. Tang, G. Zhang, Y. Wang, H. Wang, and F. Liu, "A hybrid approach to integrate fuzzy C-means based imputation method with genetic algorithm for missing traffic volume data estimation," *Transp. Res. C, Emerg. Technol.*, vol. 51, pp. 29–40, Feb. 2015.
- [28] C. Chen, Z. Liu, S. Wan, J. Luan, and Q. Pei, "Traffic flow prediction based on deep learning in Internet of vehicles," *IEEE Trans. Intell. Transp. Syst.*, early access, Oct. 6, 2020, doi: [10.1109/TITS.2020.3025856](https://doi.org/10.1109/TITS.2020.3025856).
- [29] X. Wang, K. An, L. Tang, and X. Chen, "Short term prediction of freeway exiting volume based on SVM and KNN," *Int. J. Transp. Sci. Technol.*, vol. 4, no. 3, pp. 337–352, Sep. 2015, doi: [10.1260/2046-0430.4.3.337](https://doi.org/10.1260/2046-0430.4.3.337).
- [30] W. Huang, G. Song, H. Hong, and K. Xie, "Deep architecture for traffic flow prediction: Deep belief networks with multitask learning," *IEEE Trans. Intell. Transp. Syst.*, vol. 15, no. 5, pp. 2191–2201, Oct. 2014, doi: [10.1109/TITS.2014.2311123](https://doi.org/10.1109/TITS.2014.2311123).
- [31] G. Dai, C. Ma, and X. Xu, "Short-term traffic flow prediction method for urban road sections based on space-time analysis and GRU," *IEEE Access*, vol. 7, pp. 143025–143035, 2019, doi: [10.1109/ACCESS.2019.2941280](https://doi.org/10.1109/ACCESS.2019.2941280).
- [32] D. Park and L. R. Rilett, "Forecasting freeway link travel times with a multilayer feedforward neural network," *Comput.-Aided Civil Infrastruct. Eng.*, vol. 14, no. 5, pp. 357–367, Sep. 1999, doi: [10.1111/0885-9507.00154](https://doi.org/10.1111/0885-9507.00154).
- [33] X. Guo and Q. Zhu, "A traffic flow forecasting model based on BP neural network," in *Proc. 2nd Int. Conf. Power Electron. Intell. Transp. Syst. (PEITS)*, Dec. 2009, vol. 3, no. 3, pp. 311–314, doi: [10.1109/PEITS.2009.5406865](https://doi.org/10.1109/PEITS.2009.5406865).
- [34] B. Park, C. Messer, and T. Urbanik, *Short-Term Traffic Flow Forecasting Using Radial Basis Function Neural Network*. Newbury Park, CA, USA: Sage, 1998, pp. 39–47, doi: [10.3141/1651-06](https://doi.org/10.3141/1651-06).
- [35] R. Zhang, P. Xie, C. Wang, G. Liu, and S. Wan, "Classifying transportation mode and speed from trajectory data via deep multi-scale learning," *Comput. Netw.*, vol. 162, Oct. 2019, Art. no. 106861, doi: [10.1016/j.comnet.2019.106861](https://doi.org/10.1016/j.comnet.2019.106861).
- [36] Z. Zhao, W. Chen, X. Wu, P. C. Y. Chen, and J. Liu, "LSTM network: A deep learning approach for short-term traffic forecast," *IET Intell. Transp. Syst.*, vol. 11, no. 2, pp. 68–75, Mar. 2017, doi: [10.1049/iet-its.2016.0208](https://doi.org/10.1049/iet-its.2016.0208).
- [37] A. Graves, *Supervised Sequence Labelling With Recurrent Neural Networks*, vol. 385. London, U.K.: Springer, 2012.

- [38] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997.
- [39] L. Yu, J. Zhao, Y. Gao, and W. Lin, "Short-term traffic flow prediction based on deep learning network," in *Proc. Int. Conf. Robots Intell. Syst. (ICRIS)*, Haikou, China, Jun. 2019, pp. 466–469, doi: [10.1109/ICRIS.2019.00122](https://doi.org/10.1109/ICRIS.2019.00122).
- [40] P. Poonia and V. K. Jain, "Short-term traffic flow prediction: Using LSTM," in *Proc. Int. Conf. Emerg. Trends Commun., Control Comput. (ICONC3)*, Lakshmanagarh, India, Feb. 2020, pp. 1–4, doi: [10.1109/ICONC345789.2020.9117329](https://doi.org/10.1109/ICONC345789.2020.9117329).
- [41] X. Ma, Z. Tao, Y. Wang, H. Yu, and Y. Wang, "Long short-term memory neural network for traffic speed prediction using remote microwave sensor data," *Transp. Res. C, Emerg. Technol.*, vol. 54, pp. 187–197, May 2015.
- [42] Y. Tian and L. Pan, "Predicting short-term traffic flow by long short-term memory recurrent neural network," in *Proc. IEEE Int. Conf. Smart City/SocialCom/SustainCom (SmartCity)*, Chengdu, China, Dec. 2015, pp. 153–158.
- [43] S. Wang, J. Zhao, C. Shao, C. Dong, and C. Yin, "Truck traffic flow prediction based on LSTM and GRU methods with sampled GPS data," *IEEE Access*, vol. 8, pp. 208158–208169, 2020, doi: [10.1109/ACCESS.2020.3038788](https://doi.org/10.1109/ACCESS.2020.3038788).
- [44] J. Li, L. Gao, W. Song, L. Wei, and Y. Shi, "Short term traffic flow prediction based on LSTM," in *Proc. 9th Int. Conf. Intell. Control Inf. Process. (ICICIP)*, Wanzhou, China, Nov. 2018, pp. 251–255, doi: [10.1109/ICICIP.2018.8606716](https://doi.org/10.1109/ICICIP.2018.8606716).
- [45] E. Coutinho, G. Trigeorgis, S. Zafeiriou, and B. Schuller, "Automatically estimating emotion in music with deep long-short term memory recurrent neural networks," in *Proc. Work. Notes MediaEval Workshop*, Sep. 2015.
- [46] K. Diederik and B. Jimmy, "Adam: A method for stochastic optimization," in *Proc. Int. Conf. Learn. Represent.*, 2015, pp. 1–15.
- [47] D. Ma, X. Song, and P. Li, "Daily traffic flow forecasting through a contextual convolutional recurrent neural network modeling inter- and intra-day traffic patterns," *IEEE Trans. Intell. Transp. Syst.*, early access, Feb. 24, 2020, doi: [10.1109/TITS.2020.2973279](https://doi.org/10.1109/TITS.2020.2973279).
- [48] T. Chai and R. R. Draxler, "Root mean square error (RMSE) or mean absolute error (MAE)?—Arguments against avoiding RMSE in the literature," *Geoscientific Model Develop.*, vol. 7, no. 3, pp. 1247–1250, 2014, doi: [10.5194/gmd-7-1247-2014](https://doi.org/10.5194/gmd-7-1247-2014).
- [49] H. James, "Named entity recognition with long short-term memory," in *Association for Computational Linguistics*, vol. 4. New York, NY, USA, pp. 172–175, doi: [10.3115/1119176.1119202](https://doi.org/10.3115/1119176.1119202).



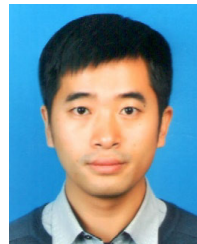
ZHU CHEN received the B.S. degree from the School of Mathematical Sciences, Guizhou Normal University, Guiyang, China, in 2019. She is currently pursuing the master's degree with the School of Mathematics and Statistics, Xi'an Jiaotong University, Xi'an, China.

Her research interest includes time series forecast.



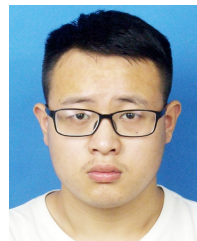
BANGYU WU (Member, IEEE) received the B.S. degree in information engineering and the Ph.D. degree in information and communication engineering from Xi'an Jiaotong University, Xi'an, China, in 2005 and 2012, respectively.

From 2007 to 2011, he was a Visiting Scholar with the University of California at Santa Cruz. He also worked as a (Senior) Geophysicist with Statoil (Beijing) Technology Service Company Ltd., from 2012 to 2015. He is currently an Associate Professor with the School of Mathematics and Statistics, Xi'an Jiaotong University. His research interests include seismic wave modeling and migration/inversion, signal processing, and machine learning.



BIN LI was born in Huanglong, Yan'an, Shaanxi, China, in 1982. He received the B.S., M.S., and Ph.D. degrees in automobile engineering from Chang'an University, in 2005, 2008, and 2013.

From 2014 to 2019, he was a Lecturer with the School of Automobile, Chang'an University. Since 2020, he has been a Vice Chairman with the Department of Automobile Engineering, Chang'an University. He is the coauthor of six books, more than 15 articles. His research interests include cooperative vehicle infrastructure systems, optimization design technology of electric vehicle, and traffic big data analysis. He is a member of working committee of automated driving of China.



HOUSONG RUAN received the B.S. degree from the Anhui University of Finance and Economics, China, in 2019. He is currently pursuing the master's degree with Xi'an Jiaotong University.

His research interests include traffic flow forecast and seismic wave inversion.

...