Urban Traffic State Prediction Based on SA-LSTM

Yu Jingyu¹⁺, a*, Wei Haiping¹,b, Guo Hongwei¹,c, Cai Yafeng¹,d

¹Geospatial Information Department, Information Engineering University, Zhengzhou, Henan, 450001, China

¹Corresponding author’s email: yjy16638057247@163.com

Abstract: Traffic prediction is the key to intelligent traffic guidance, management and control. Aiming at the spatiotemporal characteristics of urban traffic and the complexity of nonlinear variation of traffic parameters, a long short-term memory network (LSTM) urban traffic condition prediction model (SA-LSTM) based on self-attention mechanism is proposed. SA-LSTM uses the self-attention mechanism to assign different weights to traffic state information in different space and time, reflecting the spatio-temporal correlation of traffic data prediction. It can also avoid the vanishing gradient problem faced by traditional RNN and improve the defect that LSTM cannot accurately express the different importance and spatio-temporal characteristic of traffic information. Based on SA-LSTM, experiments were conducted on the Shenzhen road network data and floating car data. LSTM and SA-LSTM were selected for Comparative verification experiments, and the results confirmed that SA-LSTM is superior to LSTM in multiple evaluation indicators. Moreover, the road spatio-temporal correlations obtained by traffic data analysis and obtained by model learning are highly consistent, which proves that SA-LSTM can precisely learn and express the spatio-temporal characteristics and changing trend of the traffic.

1. Introduction

Accurate prediction of traffic conditions (such as traffic flow, running speed and passing time) is the critical technology of intelligent transportation systems (ITS), which can help traffic authorities to take reasonable preventive measures to deal with possible traffic congestion and provide smarter choices for ordinary citizens who have travel needs. Therefore, traffic condition prediction has always been a hot topic in research on intelligent transportation system, and it is still an urgent problem to be solved for geographic information and traffic researchers.

Due to the randomness of urban traffic flow, the accuracy of traffic condition prediction model has not been improved. To solve this problem, researchers have proposed many methods and technologies to model the evolution of traffic conditions. Existing schemes for prediction can be roughly divided into three categories: parametric methods, nonparametric methods and hybrid methods. Parametric methods include ARIMA model [1], Periodic ARIMA model [2, 3] and Kalman filter [4, 5]. Parametric methods are widely used in traffic condition prediction, but these methods are sensitive to traffic data under different conditions. Nonparametric methods include artificial neural networks (ANN) [6-9], k-Nearest Neighbors [10-14], support vector regression model (SVR) [15, 16] and Bayesian model [17, 18]. Compared with parametric methods, nonparametric methods have higher prediction performance and accuracy, but these methods need a lot of historical data and training. The hybrid method is mainly a combination of parametric method and nonparametric method [19-29]. Although the prediction accuracy of nonparametric
methods and hybrid methods is superior to that of parametric methods, these two methods mainly consider the data close to the prediction time point in time series, and cannot fully reveal the spatial-temporal characteristics of traffic condition data. Reference [30] has summarized the existing traffic condition prediction algorithms from 2004 to 2013. Reference [31] has studied the traffic condition prediction system, summarized the previous work on traffic condition prediction, and highlighted the change trend of traffic conditions, providing direction of research for future work. Reference [32] has summarized the latest technical achievements in the field of traffic condition prediction, and conducted in-depth analysis of unresolved technical challenges.

With the rapid progress of the Internet of Things, 5G communication technology and Internet, the Global Positioning System (GPS) is widely applied and popularized. Floating cars generally refer to taxis, buses, police cars, etc., which are equipped with vehicle GPS navigation systems. Floating car is the major means for mobile monitors to get traffic data, which has the characteristics of all-day working, wide urban coverage and low cost. It is difficult to deal with massive floating car data by conventional parametric methods. However, non-parametric methods, most of which are shallow in architecture, cannot express underlying correlation and implicit traffic information. As a newly emerging machine learning method, deep learning has attracted extensive attention from academic and industrial community. Therefore, traffic condition prediction based on the deep learning method has become a new trend.

Deep Belief Networks (DBN) consists of several layers of neurons, and its component is Restricted Boltzmann Machine (RBM). Reference [33] has proposed a framework for traffic condition prediction using DBN and multi-task learning. Reference [34] has proposed a DBN-based method to predict traffic flow, including the historical traffic flow, weather data and event-based data. Reference [35] has used DBN to predict traffic flow and weather data respectively, and merged each prediction result with DADA fusion technology. Reference [36] has found the optimal hyper-parameters of DBN model by genetic algorithms. Autoencoder is another basic structure of neural networks apart from RBM. The stacked autoencoder (SAE) model is a deep neural network model composed of multi-layer sparse autoencoders. Reference [37] has used SAE model to learn general traffic characteristics. Reference [38] has evaluated the different performances of SAE model in traffic condition prediction during the day and at night respectively. In order to improve the prediction accuracy, a SAE-Levenberg-Marquardt model has been proposed, and Taguchi method has been developed to optimize the model structure in reference [39]. Reference [40] has introduced an adaptive enhancement scheme for SAE networks.

In recent years, compared with other deep learning structures used to process sequence data, Recurrent Neural Network (RNN) has shown the most excellent performance. Reference [41] has used RBM and RNN architectures to model and predict traffic jams. However, the traditional RNN faces the problems of gradient vanishing and gradient exploding. In order to solve these problems, researchers have proposed the LSTM. Because LSTM can automatically calculate the optimal time span and capture the characteristics of time series with longer time span, LSTM model can be used in traffic condition prediction to achieve better performance. Reference [42] has developed the long-term dependencies to capture the traffic flow series through the use of LSTM. Reference [43] has used LSTM to learn more abstract representations in nonlinear traffic condition data. LSTM is very successful in traffic condition prediction, but it seldom considers the spatial-temporal characteristics of traffic conditions, and cannot distinguish the importance of traffic condition information under different time and space in traffic condition prediction tasks.

Therefore, in order to solve the gradient vanishing problem faced by RNN and the inability of LSTM to accurately express the spatial-temporal characteristics and different importance of traffic condition information, this paper proposes a LSTM traffic condition prediction model based on self-attention mechanism.
2. Data Presentation

2.1. Floating Car Data
In this paper, an experiment on the traffic condition prediction is carried out based on the floating car data of local-regional roads in Shenzhen. This data set mainly includes part of the main roads and the surrounding associated roads of Fulong Road, Xinzhou Road and Beihuan Blvd in Shenzhen. The data comes from Mathematics Modeling & Cross Technology (m2ct.org), which includes GPS data of all floating cars such as taxis, buses, coaches, dump trucks, semi-trailers and training vehicles passing through the area from March 25 to 31, 2018. Data fields include positioning time, license plate, longitude, latitude, driving speed and satellite speed.

2.2. Data Preprocessing

Data preprocessing includes floating car data cleaning and road section selection. There is signal drift and distortion in the raw floating car data, which can easily lead to the discrepancy between the calculation results and actual results if being used directly without being processed, and a large amount of abnormal data will increase the computational complexity and task difficulty. Therefore, the floating car data needs to be preprocessed, including deleting redundant and invalid data or eliminating them by the rules.

1) Filter redundant data: filtering and deleting redundant data with exactly the same recorded contents, such as duplicate records, and data whose position and speed have not changed for a long time.

2) Delete invalid data: invalid data include data distributed outside the survey region and whose speed far exceeds the maximum speed limit of urban roads.

3) Road section selection: in order to accurately express the traffic conditions of the city, road sections are selected based on the length of urban roads, road intersections and main road entrances and exits in the survey region, and 10 sub-sections of 5 sections of roads as shown in Figure 1 are selected for research. Among them, \( l_0 \) and \( l_7 \) are predicted road sections, and \( l_0 - l_9 \) are survey sections related to the predicted sections in the input models.

4) Set a time sequence of speed prediction as 5 minutes, count floating car data. Calculating the average driving speed of each sub-section within 5 minutes, then the speed sequence length of each sub-section is \( 12 \times 24 \times 7 = 2016 \).

3. SA-LSTM for urban traffic prediction

3.1. Problem Description
Define the size \( T \times N \times S \) of tensor \( X \), which represents GPS data of floating cars on the selected...
survey road sections. Among them, \( T \) represents the length of the data time series, \( N \) represents the number of road sections, and \( S \) represents traffic condition data. Then \( X \) can be expressed as:

\[
X = [X_1, X_2, \cdots, X_{n}, X_T]
\]

\[
X_t = [x_1^t, x_2^t, \cdots, x_N^t, x_N^T]
\]

Among them, \( x_t^i \) is a vector with the dimension of \( S \), representing in the \( t \) time series, the traffic condition data of the \( i \) road section, including but not limited to average speed, traffic flow, occupancy rate or any combination thereof.

The goal of short-term road traffic speed prediction task is to predict the road traffic condition in future time series lengths based on the current and historical average road traffic speeds. The road traffic conditions are usually represented by the average road traffic speed. The prediction problem is formalized as a learning function \( h(\cdot) \) which maps the \( T \) time series of traffic conditions to the \( T + 1 \) time series:

\[
X \xrightarrow{h(\cdot)} X_{T+1}
\]

After obtaining \( X_{T+1} \), we can attach \( X_{T+1} \) to \( X \) to get \( X' \). Repeatedly using the above formula, the traffic condition prediction of multiple time series can be carried out.

3.2. Long Short-term Memory Networks

RNN is a neural network specially used to process data containing time series. Unlike convolutional neural network, RNN allows previously input data to remain in the internal state of the neural network, and previously input data will continuously affect the output of the neural network. RNN has shown extraordinary ability in modeling nonlinear time series problems. However, with the increase of time span, the gradient of RNN may vanish due to excessively deep feedforward neural network. Therefore, LSTM is proposed to overcome the shortcomings of classic RNN.
In order to solve the problem of vanishing gradient, a LSTM structure with forget cells is proposed. The memory cells of LSTM can forget remote or useless information after a certain period of time. The architecture of LSTM consists of several LSTM cells. Each cell contains one or more self-join memory cells and three gates, namely input gate, forget gate and output gate. The typical structure of LSTM cells containing a memory unit is shown in Figure 2. The input gate acquires newly input data from the outside and processes the new data. The forget gate decides when to forget the previous data, thus selecting the optimal time span for the input sequence. The output gate obtains all the calculated results and generates output for LSTM cells. In the traffic condition prediction model, a linear regression layer is usually applied to the output of LSTM cells to determine the final output.

As shown in Figure 2, \( N \) is set as a dimension of data input for a time series, \( H \) is the number of cells in the hidden layer, and \( C \) is the number of memory cells. Subscripts \( i \), \( f \) and \( o \) refer to input gate, forget gate and output gate respectively. \( w_{jk} \) is the connection weights between component \( j \) and component \( k \). \( a^t_j \) is the network input of component \( j \) at the time of \( t \), and \( b^t_j \) is the value of the activation function in the same component. \( s^t_c \) is the state of the memory cell at the time of \( t \). \( \sigma \) is the activation function of the gate, usually the sigmoid function. \( g \) and \( h \) are the activation functions of the input and output of the memory unit respectively, usually tanh functions. The analytical expression for each component of the LSTM model is as follows:

1. The input gate needs to obtain newly input data from the outside, process the new data and update the model state. Its expression is:
   \[
   a^t_i = \sum_{n=1}^{N} w_{ni} x^t_n + \sum_{h=1}^{H} w_{hi} b^t_{h-1} + \sum_{c=1}^{C} w_{ci} s^t_{c-1} \\
   b^t_i = \sigma(a^t_i)
   \]

2. The forget gate takes the output of the previous sequence and the input of the current sequence as the input together, and obtains the parts of the data sequence that need to be forgotten and retained respectively through the activation function. The formula of forget gate is as follows:
   \[
   a^t_f = \sum_{n=1}^{N} w_{nf} x^t_n + \sum_{h=1}^{H} w_{hf} b^t_{h-1} + \sum_{c=1}^{C} w_{cf} s^t_{c-1} \\
   b^t_f = \sigma(a^t_f)
   \]

3. The output gate confirms the output content and processes the content memory cells. Its formula is:
   \[
   a^t_o = \sum_{n=1}^{N} w_{no} x^t_n + \sum_{h=1}^{H} w_{ho} b^t_{h-1} + \sum_{c=1}^{C} w_{co} s^t_{c-1} \\
   b^t_o = \sigma(a^t_o)
   \]

4. The state of the memory cell represents the memory information of the model at this time. The expression of the state of the memory cell is:
   \[
   s^t_c = b^t_f s^t_{c-1} + b^t_i g(a^t_c) \\
   s^t_c = b^t_i h(a^t_c)
   \]

With the functions of different gates, LSTM model has the ability to process time series data with arbitrary time span.

3.3. SA-LSTM for Urban Traffic Prediction

Attention mechanism enables the model to focus on and fully learn critical information. It is not a complete model, but it is realized on the basis of encoder-decoder model. It is a general mechanism that can apply to any time series model. When processing time series data, LSTM is usually used to encode the sequence. Through convolution or pooling, the sequence will be encoded into a vector with fixed
length as output. However, the conventional encoding method pays equal attention to different elements in the sequence, which cannot reflect the varied importance of different elements in a sequence. Therefore, an attention mechanism is needed to help the encoder adjust the degree of attention it pays to different elements. In this paper, the importance of different traffic information is adjusted at any moment based on the self-attention mechanism. The time series vector which can correctly express the global trends of traffic condition is calculated for traffic prediction task.

This paper proposes a LSTM traffic condition prediction model based on self-attention mechanism (SA-LSTM), which solves the problem of vanishing gradient of neural network with long time span, and reflects the varied importance of different elements in the same time series.

SA-LSTM model is a result of improvement on the basis of LSTM model. Self-attention mechanism is used to distinguish the importance of input information of the model. The model uses LSTM and self-attention mechanism to model traffic condition prediction. The model structure is shown in Figure 3. SA-LSTM model is roughly divided into three parts: LSTM layer, attention layer and fully connected layer.

![Figure 3. Structure diagram of SA-LSTM cell](image)

(1) LSTM layer: enter the historical time series data $X = [X_1, X_2, \ldots, X_t]$ before the $t+1$ time series to be predicted. After encoding of LSTM layer, memory and forgetting, the output vector $q_t = LSTM(X_t)$ of the hidden layer is obtained.

(2) Attention layer: the input $q$ of the attention layer is often called a Query. The query of self-attention mechanism is usually the transformation of input data, and the query of traditional attention mechanisms comes from the outside. After obtaining Query $q$, the self-attention mechanism will get the attention distribution $\alpha_t$ through different scoring mechanisms $s$ (including additive model, dot product model and bilinear model). In this paper, the scaled dot product model is used as the scoring mechanism of self-attention and softmax calculation. Attention distribution $\alpha_t$ represents the degree of importance of the $t$ data when querying $q_t$ from time series data. Finally, the self-attention
mechanism will use attention distribution $\alpha_t$ to re-encode input data $X$ and obtain Attention Value as the output of this layer. In SA-LSTM model, the calculation process of attention layer is as follows:

$$\alpha_t = \text{softmax}(s(X_t, q_t)) = \text{softmax} \left( \frac{X_t^T q_t}{\sqrt{N \times S}} \right)$$

$$V_t = \text{att}(q_t, X) = \sum_{i=1}^{T} \alpha_t X_t$$

(3) Fully connected layer: the output $V$ obtained in the attention layer is input into the fully connected layer, and the predicted value $Y^*_t = [y_1, y_2, ..., y_t]$ is calculated and output through the fully connected method. Among them $y_t = f(uV_t + c)$, $w$ is the weight matrix of the fully connected layer, and $c$ is the bias vector.

In the training process, the parameters of SA-LSTM model are continuously trained and optimized, and finally the training is stopped after the model meets the times of training or the parameters stop to update. In order to minimize the training error and avoid local minimum at the same time, the Adam optimizer improved based on Stochastic Gradient Descent (SGD) optimizer with adaptive learning rate is applied for back propagation. Meanwhile, neural network is famous for its superior expression ability, but has a tendency of overfitting. For a long time, a large number of regularization methods have been proposed to reduce over-fitting. Dropout method proposed in 2012 is a very effective method for training neural networks. SA-LSTM uses dropout method to prevent over-fitting of training data in the neural network.

After the trained model is obtained, the test data set is input into the model for prediction and the prediction error between the predicted result and the actual value is calculated. The loss function of SA-LSTM model is:

$$\text{loss} = \| Y_t - Y^*_t \| + L$$

Among them, $Y_t$ represents the actual value, and $Y^*_t$ represents the predicted result. $L$ represents L2 regularizers, which can avoid over-fitting problems together with dropout mechanism.

3.4. Performance Indicators

In order to extensively verify the predictive accuracy of the model, this paper selects Mean Absolute Percentage Error (MAPE) and Root Mean Square Error (RMSE) as evaluation functions. The calculation formulas are as follows:

$$\text{MAPE} = \frac{1}{n} \sum_{i=1}^{n} \left| \frac{y_i - y_i^*}{y_i} \right|$$

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{n} (y_i - y_i^*)^2}{n}}$$

Among them, $y_i$ is the actual value, $y_i^*$ is the predicted result, and $n$ is the quantity of samples. MAPE indicates the percentage of deviation between the predicted data and the actual data. The smaller the value, the higher the prediction accuracy. RMSE indicates the sample standard deviation of the residual error between the predicted data and the actual data. The smaller the value, the better the fitting of the model.

4. Experiment Results and Analysis

4.1. Parameter Setting

The prediction task of this experiment is to use the road average speed data of 12 known time series (5 minutes for each time series, and 1 hour in total) to predict the road average speed of the next 3-time
series. The data set is divided into training dataset and test dataset in sequence according to the ratio of 7: 3. Adam method is used as the optimizer of the model in the experiment. The initial learning rate is set to 0.001, and each time 25% of the training is completed, it is attenuated to 10% of the original. The Batch size is set to 64, and the learning times epochs are set to 100. The probability of dropout is 20%.

4.2. Experiment Results
According to the selected data and parameter setting, the LSTM model and SA-LSTM model are trained and tested respectively. The RMSE and MAPE obtained from the experiment are shown in Table 1.

| Predicted road sections | Prediction time | MAPE LSTM | MAPE SA-LSTM | RMSE LSTM | RMSE SA-LSTM |
|------------------------|----------------|-----------|--------------|-----------|--------------|
| Road section l₀        | 5min           | 0.125     | 0.091        | 1.608     | 1.559        |
|                        | 10min          | 0.184     | 0.124        | 1.987     | 1.966        |
|                        | 15min          | 0.257     | 0.167        | 2.398     | 2.248        |
| Road section l₀        | 5min           | 0.152     | 0.138        | 1.836     | 1.687        |
|                        | 10min          | 0.197     | 0.164        | 2.058     | 1.952        |
|                        | 15min          | 0.273     | 0.215        | 2.419     | 2.251        |

4.3. Experiment Analysis
From the road speed prediction results in Table 1, it can be seen that the prediction accuracy of the same model for different road sections is somewhat different, indicating that there are spatial differences in traffic condition change trends of different road sections, resulting in large differences in learning and prediction effects of the model for different road sections. However, considering all predicted road sections and prediction time, the MAPE and RMSE of the prediction results of SA-LSTM model are better than those of traditional LSTM. The average MAPE and RMSE of SA-LSTM model were 0.150 and 1.944, which were 24.24% and 5.22% smaller than those of LSTM model respectively. It shows that the self-attention mechanism effectively improves the accuracy of road speed prediction and reflects the right trend of traffic speed change between road sections.

By analyzing the prediction accuracy of different prediction time, it can be seen that the longer the time span of road speed prediction tasks, the more SA-LSTM model is ahead of LSTM model. It shows that self-attention mechanism can help SA-LSTM model to provide long-term traffic speed prediction more accurately, and provide more accurate and long-term suggestions for traffic authorities and ordinary travelers.

In order to further verify whether the self-attention mechanism can correctly learn the nonlinear spatial-temporal correlation of traffic condition information, this paper has analyzed the road speed correlation of the input road sections l₀ - l₀. Spearman correlation coefficient pair is selected to quantify l₀ - l₀ road speed correlation. The calculation formula of the Spearman correlation coefficient is as follows:

\[
\rho = \frac{\sum_i (x_i - \bar{x}) (y_i - \bar{y})}{\sqrt{\sum_i (x_i - \bar{x})^2 (y_i - \bar{y})^2}}
\]

(18)

Among them, \(x\) and \(y\) represent the road speed data of the two road sections to be calculated respectively. According to the above formula, the thermodynamic diagram of the Spearman correlation coefficient between road sections of l₀ - l₀ is obtained.
As shown in Figure 4, the road section $l_0$ is highly correlated with road sections $l_2$ and $l_8$, and the road section $l_7$ is highly correlated with $l_3$ and $l_9$. Among them, the correlation coefficient of $l_6$ and $l_8$, $l_7$ and $l_9$ are as high as 0.94 and 0.9 respectively, which is consistent with the road network structure as shown in Figure 1. The adjacent interlinked road sections have similar trends of road speed change. At the same time, the SA-LSTM model is analyzed to obtain the mean of attention distribution parameters $\alpha$ of the self-attention layer:

![Correlation between roads.](image)

| $l_0$ | $l_1$ | $l_2$ | $l_3$ | $l_4$ | $l_5$ | $l_6$ | $l_7$ | $l_8$ | $l_9$ |
|------|------|------|------|------|------|------|------|------|------|
| 0.009 | 0.037 | 0.085 | 0.088 | 0.012 | 0.042 | 0.231 | 0.232 | 0.135 | 0.129 |

As can be seen from Table 2, in the learning process of SA-LSTM, $l_6$ and $l_7$ as the predicted road sections have obtained the maximum attention distribution. Then, adjacent road sections $l_2$, $l_3$, $l_8$, and $l_9$ have obtained relatively wide attention distribution. However, $l_0$ and $l_4$, which are not connected with the predicted road section, have obtained little attention distribution. This learning result is highly similar to road network structure and its correlation analysis, which proves that SA-LSTM model can effectively learn and express the spatial-temporal correlation and weight difference of traffic condition information, and shows that SA-LSTM model has improved the accuracy in traffic condition prediction.

5. Conclusion and Prospect
Traffic condition prediction is a key component of urban intelligent transportation system. The changing trend of traffic conditions varies greatly with time and road sections, so its short-term and long-term prediction is difficult to a certain degree. In the light of the problem of traffic condition prediction, this paper proposed a LSTM model based on self-attention mechanism. SA-LSTM model used self-attention mechanism to mine the temporal and spatial relationships of road traffic conditions in different road
sections, and quantified the importance of different spatial-temporal traffic conditions through attention distribution, which solved the problem that traditional LSTM model ignored the spatial-temporal correlation in road traffic condition prediction. The experimental results show that the accuracy of SA-LSTM was significantly improved compared with the traditional LSTM model, and the effectiveness of self-attention mechanism in learning and expressing the spatial-temporal correlation of traffic conditions in different road sections was verified.

The SA-LSTM model proposed in this paper still needs to be improved: first, due to the short time span and small coverage of the data itself, it is impossible to predict the long-term and large-scale traffic conditions. Second, the prediction includes less road traffic condition information, and the traffic condition prediction including multi-information fusion such as traffic flow and passing time needs further exploration in the future.

References

[1] Voort, M., Dougherty, M., Watson, S. (1996) Combining kohonen maps with arima time series models to forecast traffic flow. Transportation Research Part C Emerging Technologies, 4(5): 307-318.

[2] Williams, B. M., L. A. Hoel. (2003) Modeling and Forecasting Vehicular Traffic Flow as a Seasonal ARIMA Process: Theoretical Basis and Empirical Results. Journal of Transportation Engineering 129(6): 664-672.

[3] Shi, G., Guo, J., Huang, W., Williams, B. M. (2014) Modeling seasonal heteroscedasticity in vehicular traffic condition series using a seasonal adjustment approach. Journal of Transportation Engineering, 140(5): 1053-1058.

[4] Ojeda, L. L., Kibangou, A. Y., CC D Wit. (2013) Adaptive Kalman Filtering for Multi-Step ahead Traffic Flow Prediction. American Control Conference (ACC), 2013. IEEE.

[5] Guo, J., Huang, W., Williams, B M. (2014) Adaptive kalman filter approach for stochastic short-term traffic flow rate prediction and uncertainty quantification. Transportation Research Part C, 43, pp. 50-64.

[6] Smith, B. L., Demetsky, M. J. (1994). MJ, Short-term traffic flow prediction: neural network approach. Transportation Research Record, 1453, 98-104.

[7] Zhang, Y., Song, Z., Weng, X., Xie, Y. (2019) A new soil-water characteristic curve model for unsaturated loess based on wetting-induced pore deformation. Geofluids, (3): 1-14.

[8] Wang, Q., Sun, H. (2019) Traffic structure optimization in historic districts based on green transportation and sustainable development concept. Advances in Civil Engineering, 1-18.

[9] Chen, D. (2017) Research on traffic flow prediction in the big data environment based on the improved RBF neural network. IEEE Transactions on Industrial Informatics, pp. (4), 1-1.

[10] Bernaś, M., Placzek, B., Porwik, P., Pamula, T. (2015) Segmentation of vehicle detector data for improved k-nearest neighbours-based traffic flow prediction. Intelligent Transport Systems Iet, 9(3), 264-274.

[11] Wu, S., Yang, Z., Zhu, X., Yu, B. (2014) Improved k-nn for short-term traffic forecasting using temporal and spatial information. Journal of Transportation Engineering, 140(7), 04014026.

[12] Dell’Acqua, P., Bellotti, F., Berta, R., De Gloria, A. (2015) Time-Aware Multivariate Nearest Neighbor Regression Methods for Traffic Flow Prediction. IEEE Transactions on Intelligent Transportation Systems. 16(6): 3393-3402.

[13] Cai, P., Wang, Y., Lu, G., Chen, P., Ding, C., Sun, J. (2016) A spatiotemporal correlative k-nearest neighbor model for short-term traffic multistep forecasting. Transportation Research Part C Emerging Technologies., 62, pp. 21–34.

[14] Sun, B., Cheng, W., Goswami, P., Bai, G. (2018) Short-term traffic forecasting using self-adjusting k-nearest neighbours. IET Intelligent Transport Systems., 12. pp. 41–48.
[15] Castro-Neto, M., Jeong, Y. S., Jeong, M. K., Han, L. D. (2009). Online-svr for short-term traffic flow prediction under typical and atypical traffic conditions. Expert Systems with Applications, 36(3-part-P2), 6164-6173.

[16] Sun, Y., Leng, B., Guan, W. (2015) A novel wavelet-SVM short-time passenger flow prediction in Beijing subway system. Neurocomputing, 166(oct.20), 109-121.

[17] Jian, W., Deng, W., Guo, Y. (2014) New Bayesian combination method for short-term traffic flow forecasting. Transportation Research Part C, 43(pt.2), 79-94.

[18] Xu, Y., Kong, Q. J., Klette, R., Liu, Y. (2014) Accurate and interpretable Bayesian MARS for traffic flow prediction. IEEE Transactions on Intelligent Transportation Systems, 15(6), 2457-2469.

[19] Wang, K., Qiu, J., Niu, F., Chen, J. (2016) Vibration Response Characteristics of the Cross Tunnel Structure. Shock & Vibration, 16.

[20] Xia, D., Wang, B., Li, Y., Zhang, Z. (2016) A distributed spatial–temporal weighted model on MapReduce for short-term traffic flow forecasting. Neurocomputing, 179, pp. 246–263.

[21] Moretti, F., Pizzuti, S., Panzieri, S., & Annunziato, M. (2015) Urban traffic flow forecasting through statistical and neural network bagging ensemble hybrid modeling. Neurocomputing, 167, 3-7.

[22] Zhou, Z., Ren, C., Xu, G., Zhan, H., Liu, T. (2019) Dynamic failure mode and dynamic response of high slope using shaking table test. Shock and vibration, 2019.

[23] Zhang, Y., Weng, X., Song, Z., Sun, Y. (2019) Modeling of loess soaking induced impacts on a metro tunnel using a water soaking system in centrifuge. Geofluids, 2019.

[24] Kim, Y. J., Hong, J. S. (2015) Urban traffic flow prediction system using a multifactor pattern recognition model. IEEE Transactions on Intelligent Transportation Systems, 16(5), 2744-2755.

[25] Zhou, Z., Zhu, S., Kong, X., Lei, J., Liu, T. (2019) Optimization analysis of settlement parameters for post-grouting piles in loess area of Shaanxi, China. Advances in Civil Engineering, 2019.

[26] Huang, M. L. (2015) Intersection traffic flow forecasting based on v-GSVR with a new hybrid evolutionary algorithm. Neurocomputing, 147, 343-349.

[27] Luo, X., Niu, L., Zhang, S. (2018) An algorithm for traffic flow prediction based on improved SARIMA and GA. KSCE Journal of Civil Engineering, 22(10), 4107-4115.

[28] Luo, X., Li, D., Zhang, S. (2019) Traffic flow prediction during the holidays based on DFT and SVR. Journal of Sensors, 2019.

[29] Lai, J., Qiu, J., Fan, H., Zhang, Q., Hu, Z., Wang, J., Chen, J. (2016) Fiber Bragg grating sensors-based in situ monitoring and safety assessment of loess tunnel. Journal of Sensors, 2016.

[30] Vlahogianni, E. I., Karlaftis, M. G., Golas, J. C. (2014) Short-term traffic forecasting: Where we are and where we’re going. Transportation Research Part C: Emerging Technologies, 43, 3-19.

[31] Suhas, S., Kalyan, V. V., Katti, M., Prakash, B. A., Naveena, C. (2017) A comprehensive review on traffic prediction for intelligent transport system. In 2017 International Conference on Recent Advances in Electronics and Communication Technology (ICRAECT), pp. 138-143.

[32] Lana, I., Del Ser, J., Velez, M., Vlahogianni, E. I. (2018) Road traffic forecasting: Recent advances and new challenges. IEEE Intelligent Transportation Systems Magazine, 10(2), 93-109.

[33] Huang, W., Song, G., Hong, H., Xie, K. (2014) Deep architecture for traffic flow prediction: deep belief networks with multitask learning. IEEE Transactions on Intelligent Transportation Systems, 15(5), 2191-2201.
[34] Soua, R., Koesdwiady, A., Karray, F. (2016) Big-data-generated traffic flow prediction using deep learning and dempster-shafer theory. In 2016 International joint conference on neural networks (IJCNN), (pp. 3195-3202).

[35] Koesdwiady, A., Soua, R., Karray, F. (2016) Improving traffic flow prediction with weather information in connected cars: A deep learning approach. IEEE Transactions on Vehicular Technology, 65(12), 9508-9517.

[36] Zhang, Y., Huang, G. (2018) Traffic flow prediction model based on deep belief network and genetic algorithm. IET Intelligent Transport Systems, 12(6), 533-541.

[37] Lv, Y., Duan, Y., Kang, W., Li, Z., Wang, F. Y. (2014) Traffic flow prediction with big data: a deep learning approach. IEEE Transactions on Intelligent Transportation Systems, 16(2), 865-873.

[38] Duan, Y., Lv, Y., Wang, F. Y. (2016) Performance evaluation of the deep learning approach for traffic flow prediction at different times. In 2016 IEEE International Conference on Service Operations and Logistics, and Informatics (SOLI), pp. 223-227.

[39] Yang, H. F., Dillon, T. S., Chen, Y. P. P. (2016) Optimized structure of the traffic flow forecasting model with a deep learning approach. IEEE transactions on neural networks and learning systems, 28(10), 2371-2381.

[40] Zhou, T., Han, G., Xu, X., Lin, Z., Han, C., Huang, Y., Qin, J. (2017) δ-agree AdaBoost stacked autoencoder for short-term traffic flow forecasting. Neurocomputing, 247, 31-38.

[41] Ma, X., Yu, H., Wang, Y., Wang, Y. (2015) Large-scale transportation network congestion evolution prediction using deep learning theory. PloS one, 10(3), e0119044.

[42] Ma, X., Tao, Z., Wang, Y., Yu, H., Wang, Y. (2015) Long short-term memory neural network for traffic speed prediction using remote microwave sensor data. Transportation Research Part C: Emerging Technologies, 54, 187-197.

[43] Shao, H., Soong, B. H. (2016) Traffic flow prediction with long short-term memory networks (LSTMs). In 2016 IEEE Region 10 Conference (TENCON), pp. 2986-2989.