Short-term Traffic Flow Prediction with LSTM Recurrent Neural Network

Danqing Kang, Yisheng Lv, Yuan-yuan Chen

***Abstract*—Accurate and timely short-term traffic flow prediction plays an important role in intelligent transportation management and control. Traffic flow prediction has a long history and is still a difficult problem due to intrinsically highly nonlinear and stochastic characteristics of complex transportation systems. In this paper, we employ the long short-term memory (LSTM) recurrent neural network to analyze the effects of various input settings on the LSTM prediction performances. Flow, speed, and occupancy at the same detector station are used as inputs to predict traffic flow. The results show that the inclusion of occupancy/speed information may help to enhance the performance of the model overall. Further, we include as inputs traffic variables from the upstream and/or downstream detector stations for traffic flow prediction. The evaluation of such spatial-temporal input interactions show that the inclusion of both downstream and upstream traffic information is useful in improving prediction accuracy.**

Keywords—traffic flow prediction, long short-term memory model (LSTM), deep Learning, upstream traffic information, downstream traffic information

I. Introduction

As a key technology in intelligent transportation systems (ITS), traffic flow prediction has gained more and more attention with the rapid development and deployment of ITS. However, transportation systems are essential large-scale time-variant complex systems consisting of various nonlinear and stochastic dynamical processes at multiple levels. Short-term traffic flow prediction has received a lot of research interest over the past four decades, but continues to be a challenge until today [1]-[6].

Short-term traffic flow prediction aims at estimating traffic flow in the near future based on historical traffic data. A great deal of works have been done on the subject of traffic flow prediction and an extensive variety of models based on different theories have been developed [7]-[9]. In general, traffic flow prediction models can be divided into two major categories: parametric approach and non-parametric approach.

Parametric models include time-series models, Kalman filtering models, etc. Nonparametric models include k-nearest neighbor (KNN) methods, support vector regression (SVR), artificial neural networks (ANNs), etc. Due to the stochastic and nonlinear nature of traffic flow, parametric models cannot describe it accurately with analysis formulas. Hence, researchers attach great attention to nonparametric approaches in the field of traffic flow prediction [10].

A widely-used technique for traffic flow prediction is based on time-series methods, and Auto Regressive Integrated Moving Average (ARIMA) is a typical parametric regression model, which assumes that the traffic condition is a stationary process where the mean, variance and auto-correlation are unchanged. ARIMA was proposed in the 1970s to predict short-term freeway traffic data and then a series of variant models based on ARIMA have been proposed, such as Kohonen-ARIMA [11], subset ARIMA [12], ARMA [13], seasonal ARIMA [14]. In addition to the ARIMA-like time series model, other types of time series models are also used for traffic flow prediction [15].

KNN is a type of instance-based learning, or lazy learning, where the function is only approximated locally and all computation is deferred until classification. Chang et al. presented a dynamic multi-interval traffic flow prediction model based on the KNN nonparametric regression [16]. Davis and Nihan used the KNN method for short-term freeway traffic forecasting and argued that the KNN method performed comparably with but not better than the linear time-series approach [17].

SVR assumes that there is some definite mapping between the historical observations of the traffic and the future values. The essence of SVR is to map data into a high-dimensional feature space via a nonlinear relationship and then performs linear regression within this space [18]. Castro-Neto et al. used online support vector regression (OL-SVR) to predict traffic flow under typical and atypical traffic conditions [19]. Jeong et al. further developed an online learning weighted support-vector regression (OLWSVR) based on OLSVR [20].

ANN is another popular prediction strategy due to its flexible model structure, learning ability and adaptability. Different from the statistical methods, ANN does not require underlying assumptions regarding data [21]. A Bayesian combined neural network approach for short-term freeway

This work was partially supported by National Natural Science Foundation of China (61533019, 61233001, 61603381, 71472174, 71232006) and the Early Career Development Award of SKLMCCS.

Danqing Kang is with Harbin University of Science and Technology Harbin 150080, China, and he is also affiliated with State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing, 100190, China. (e-mail: kangdanqing2015@ia.ac.cn)

Yisheng Lv is with State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing, 100190, China, and he is also affiliated with Qingdao Academy of Intelligent Industries, Qingdao, Shandong, 266109, China. (e-mail: yisheng.lv@ia.ac.cn). Corresponding author of this paper.

Yuan-yuan Chen is with State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing, 100190, China. (e-mail: yy-chen5133@ia.ac.cn)

978-1-5386-1526-3/17/$31.00 ©2017 IEEE

traffic flow prediction was presented in [22]. A genetically designed regression and time delay neural network models was proposed traffic prediction on different types of roads in [23]. Vlahogianni et al. optimized the neural network with genetic approach and applied the model to short-term traffic flow prediction [24].

Recently, as the deep learning takes great success in dealing with numerous engineering and technical problems [25]-[28], such as natural language processing, objects detection, etc., more and more researchers have been trying to apply the deep neural network to the traffic prediction. Lv et al. firstly applied a deep architecture model using autoencoders as building blocks to represent traffic flow features for prediction [29]. Duan et al. evaluated the performance of the SAE model for traffic flow prediction at different times [30]. Xiong et al. proposed a deep learning based model to predict bus travel time [31].

The long short-term memory (LSTM) model, a specially designed recurrent neural network (RNN) architecture, can learn information with long time spans and determine the optimal time lags in an automatic manner. It is particularly useful for time series prediction with long temporal dependency. Ma, et al. applied LSTM for traffic speed prediction with remote microware sensor data [32]. Tian et al. proposed LSTM RNN for traffic flow prediction and showed that LSTM NN have better performance than most of the non-parametric models [33]. Li et al. evaluated the performance of the LSTM and GRU model for traffic flow prediction [34]. Previous LSTM based traffic flow prediction model mainly consider flow alone as inputs and we use the LSTM model to adopt various combinations of different types of information as inputs, such as occupancy, speed and the neighboring traffic flow.

It is believed that traffic flow prediction performances can be enhanced if incorporating flow, speed, and occupancy from the current and/or neighoring stations as inputs. Actually, traffic state variations of neighboring detector stations are usually correlative. In this paper, we employ the LSTM model to analyze the effects of different input settings on the traffic flow prediction performances. Flow, speed, and occupancy at the same detector station are used as inputs to predict traffic flow. Further, we include traffic variables from the upstream and downstream detector stations as inputs to the LSTM model for traffic flow prediction.

The rest of this paper is organized as follows. Section II presents the LSTM model, and its counterparts for traffic flow prediction with various input settings. Section III analyzes experimental results. Section IV makes the conclusions of this paper.

II. METHODOLOGY

In this section, we mainly present the LSTM model and its counterparts for traffic flow prediction with various input settings. Also, we introduce the data sources used in the subsequent experiments..

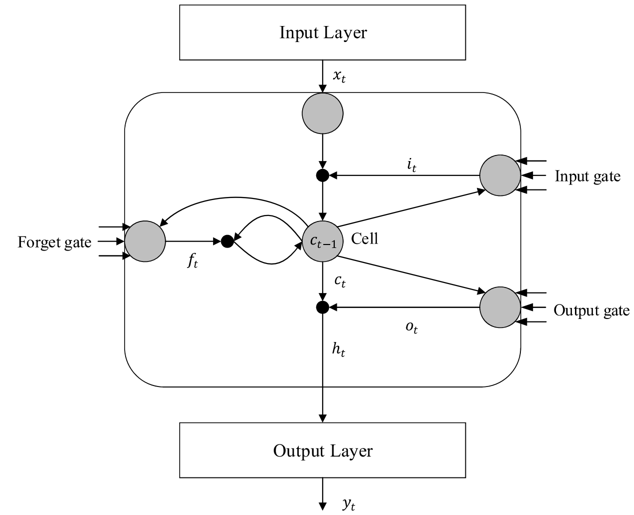

Fig. 1. Structure of LSTM Memory Block

A. LSTM RNN

Traditional RNN architecture has the so called vanishing gradient problem. To overcome such disadvantage, certain structure of RNNs such as LSTM were proposed, which was designed to give the memory cells ability to determine when to forget certain information, thus determining the optimal time lags for time series problems. These features are particularly desirable for short-term traffic flow prediction in the transportation domain because of its long-standing memory ability.

Typical LSTM is composed of one input layer, one recurrent hidden layer whose basic unit is memory block, and one output layer. The memory block contains memory cells with self-connections memorizing the temporal state, and three adaptive, multiplicative gating units: the input, the output and the forget gates to control information flow in the block. The three additional gates provide continuous analogues of write, read and reset operations on the block. Multiplicative gates can learn to open and close, and thus LSTM memory cells can store and access information over long periods of time, thereby mitigating the vanishing gradient problem. Fig. 1 gives an illustration of the LSTM memory block.

Suppose that the input of historical traffic flow sequence is denoted as $\mathbf{x} = (x_1, x_2, ..., x_T)$, where T is the prediction period, the hidden state of memory block $\mathbf{h} = (h_1, h_2, ..., h_T)$, and then the real output sequence $\mathbf{y} = (y_1, y_2, ..., y_T)$ can be iteratively calculated by following the equations:

$$y_t = W_{hy} h_t + b_y \tag{1}$$

$$h_t = H(W_{xh} x_t + W_{hh} h_{t-1} + b_h) \tag{2}$$

Where W denotes weight matrices (e.g. W_{xh} is the input-hidden weight matrix), b denotes bias vectors, and H is the hidden layer function, which can be computed in the following formulas:

$$i_t = \sigma(W_{xi} x_t + W_{hi} h_{t-1} + W_{ci} c_{t-1} + b_i) \tag{3}$$

$$f_t = \sigma(W_{xf} x_t + W_{hf} h_{t-1} + W_{cf} c_{t-1} + b_f) \tag{4}$$

$$c_t = f_t c_{t-1} + i_t g(W_{xc} x_t + w_{hc} h_{t-1} + b_c) \quad (5)$$

$$o_t = \sigma(W_{xo} x_t + W_{ho} h_{t-1} + W_{co} c_t + b_o) \quad (6)$$

$$h_t = o_t h(c_t) \quad (7)$$

The i, f, o, and c are the input gate, forget gate, output gate and activation vectors respectively. Where $\sigma(\cdot)$ is the standard logistic sigmoid function defined in Eq. (8):

$$\sigma(x) = \frac{1}{1+e^{-x}} \quad (8)$$

$g(\cdot)$ is a centered logistic sigmoid function with range [-2, 2]:

$$g(x) = \frac{4}{1+e^{-x}} - 2 \quad (9)$$

$h(\cdot)$ is a centered logistic sigmoid function with range [-1,1]:

$$h(x) = \frac{2}{1+e^{-x}} - 1 \quad (10)$$

The objective function is to minimize the sum of square errors, which is given by the following equation:

$$e_t = \sum_{t=1}^{n} (y_t - p_t)^2 \quad (11)$$

p_t represents the predicted traffic flow value at time step t.

To minimize training error and meanwhile avoid local minimal points, Adam optimizer, a modification of stochastic gradient descent (SGD) optimizer with adaptive learning rates, is applied for back propagation through time (BPTT) in this paper. At the same time, we use the normalization method to preprocess the traffic flow data and apply weighted L1 and L2 regularization methods and dropout methods for LSTM to reduce overfitting.

In general, the more layers the LSTM has, the stronger the learning ability of the model, but it is also easier to be overfitting. Here, we use two LSTM layers and a dense layer to capture the characteristics of traffic flow dynamics and achieve satisfactory results.

B. Data Descripition and Experiment Design

The data used in this study is downloaded from the open-access traffic flow database named California Performance Measurement System (PeMS). We use the traffic data of five loop detector stations located at north bound freeway I5-S, Ripon city, San Joaquin County, California. The freeway has three lanes under surveillance. The data were collected from October 1, 2009 to November 30, 2009 with the updating frequency of 30s and then the data are aggregated into 5-min interval for each detector station. The data include a total of 17556 sample points, of which 2/3 was used for training while the remaining 1/3 was used for testing, where the key information include flow, occupancy and speed. To ensure a more reliable result, missing and erroneous records were properly remedied using temporally adjacent records.

The detector station P is the target for traffic flow prediction, and detector stations U3, U2 and U1 are the upstream detector stations of P, while D1, D2 and D3 are the downstream detector stations. The traffic flow variation of the given site P is closely related with the upstream and downstream traffic conditions.

TABLE I. Various Combinations of Flow, Speed, and Occupancy

Variables Combinations	Flow	Occupancy	Speed
S1	*		
S2	*	*	
S3	*		*
S4	*	*	*

Using only flow data does not warrant the best prediction results. Speed data or occupancy data need to be used simultaneously with flow data to improve prediction performance. We firstly investigate various combinations of flow, speed, and occupancy as the inputs of LSTM to predict traffic flow, which is depicted in Table I. From Table I we can see a total of four variable combinations S1, S2, S3, S4 and the variables contained in each combination are marked with an asterisk in the table. We test all these four methods on 7 detector stations, and the results are shown in Table II.

The evolution of traffic flow can be considered as a temporal and spatial process. Therefore, we explore further how they interact and what it can improve flow prediction performances. We investigate whether the inclusion of upstream and/or downstream data would improve traffic flow prediction. Table III shows the configurations of such spatial information for the LSTM model. Herein, we use only traffic flow data of the current and neighboring stations.

As can be seen from Table III that 16 different combinations of detector stations can be used as inputs to the prediction model, and the variables contained in each combination are marked with an asterisk. That means we should build 16 models based on the LSTM model to uncover the effects of the upstream and/or downstream traffic condition on the prediction accuracy. The input variables are the different combinations of historical traffic flow of detector stations U3, U2, U1, P, D1, D2, D3 and the response is traffic flow of detector station P at the next time steps. Through comparing the prediction accuracy of different input settings, the optimal variable combination can be acquired for the freeway short-term traffic prediction model.

C. Index of Performance

To evaluate the effectiveness of the model for prediction, we use three performance indexes, which are the root mean square error (RMSE), mean absolute error (MAE), and the mean absolute percentage error (MAPE). They are defined as:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left|f_i - \hat{f}_i\right|^2} \quad (12)$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n} | f_i - \hat{f}_i | \quad (13)$$

$$MAPE = \frac{1}{n}\sum_{i=1}^{n} \frac{| f_i - \hat{f}_i |}{f_i} \quad (14)$$

where f_i and $\hat{f}_i$ are the real and predicted traffic flow of the given detector station respectively.

TABLE II. Performance of Different Input Variable Combinations on 7 Detector Stations

Detector station	Performance index	Input variable combinations			
		S1	S2	S3	S4
U3	RMSE	14.09	13.86	13.99	13.98
	MAE	10.37	10.17	10.33	10.33
	MAPE(%)	13.16	12.92	13.06	13.15
U2	RMSE	21.09	20.84	20.92	20.68
	MAE	14.04	13.93	13.94	13.73
	MAPE(%)	11.24	10.92	10.84	10.75
U1	RMSE	19.64	19.37	19.65	19.66
	MAE	14.28	14.00	14.25	14.22
	MAPE(%)	10.66	10.42	10.59	10.53
P	RMSE	23.61	23.39	23.41	23.45
	MAE	16.16	16.09	16.05	15.97
	MAPE(%)	13.22	13.04	12.88	12.84
D1	RMSE	18.55	18.50	18.48	18.42
	MAE	13.63	13.65	13.55	13.51
	MAPE(%)	10.31	10.35	10.27	10.23
D2	RMSE	22.27	21.91	22.58	22.32
	MAE	16.05	15.82	16.41	16.18
	MAPE(%)	11.53	11.35	11.79	11.59
D3	RMSE	19.07	19.01	18.76	18.70
	MAE	14.27	14.23	14.09	14.07
	MAPE(%)	9.92	9.93	9.84	9.81

III. EXPERIMENTS AND DISCUSSION

A. Combined with Speed and/or Occupancy

We compared the effects of different input variable combinations listed in Table I on the performance of traffic flow prediction. There are four possible input combinations: only flow (S1), flow and occupancy (S2), flow and speed (S3), flow occupancy and speed (S4). We test all these choices on the 7 detector stations, and present the results in Table II.

We use light gray shading to mark the best performance of the four different input combinations, and we can easily find that using additional occupancy and/or speed information to assist the prediction of traffic flow is useful. In general, the best performance is achieved by the S2 and S4 combinations, which both includes occupancy information as inputs to the model. It may imply that occupancy information is helpful for improving the performance of the traffic flow prediction model.

But the speed information is uncertain for boosting the accuracy of the model. It can help the model to make an improvement on most detector station but can not exclude the occurrence of exceptions on some detector stations. Using traffic flow concurrently with speed information can not always get a better performance.

B. Use The Neighboring Traffic Information

LSTM provides a flexible framework to adopt various combinations of different types of information as inputs for the prediction models. In this part, totally 16 short-term prediction models with different input settings are built to uncover the effects of the upstream and downstream traffic condition on the prediction performance. The detail information of these models can be found in Table III.

The input of each model are different combinations of historical traffic flow of the 7 detector stations U3, U2, U1, P, D1, D2, D3 and the response is the traffic flow of the detector station P at the next time steps. Through comparing the prediction accuracy of different models, the optimal variable combination can be acquired for the short-term traffic flow prediction.

The prediction accuracy of all models is ranked as shown in Table III. The top three high-accuracy models marked with light gray shading in Table III are Model 16, Model 15 and Model 11, which all take historical traffic flow of multiple detector stations as inputs. And Model 16 which achieves the best performance using historical traffic flow data of all the 7 detector stations. These three models have an remarkable performance improvement by 3.96%, 3.28%, 3.25% in MAPE

TABLE III. Performance of 16 LSTM Models with Neighbor Flows as Inputs

Model	Number of input variables	Detector station for prediction							RMSE	MAE	MAPE (%)	Sorting
		U3	U2	U1	P	D1	D2	D3				
1	1				*				23.61	16.16	13.22	16
2	2			*	*				23.45	16.37	13.15	15
3					*	*			18.82	13.73	10.50	8
4	3		*	*	*				23.34	16.28	13.05	14
5				*	*	*			18.54	13.51	10.41	6
6					*	*	*		21.99	15.87	11.42	12
7	4	*	*	*	*				23.17	16.17	13.04	13
8			*	*	*	*			18.49	13.43	10.29	5
9				*	*	*	*		21.96	15.85	11.22	11
10					*	*	*	*	18.78	14.02	10.17	7
11	5	*	*	*	*	*			17.77	12.89	9.97	3
12			*	*	*	*	*		21.92	15.76	11.20	10
13				*	*	*	*	*	18.32	13.56	9.96	4
14	6	*	*	*	*	*	*		21.27	15.26	10.91	9
15			*	*	*	*	*	*	18.26	13.55	9.94	2
16	7	*	*	*	*	*	*	*	17.18	12.79	9.26	1

compared to the basic Model 1 which only uses traffic flow as the input at detector station P, respectively.

From Table III, overall using more historical data of neighboring detector stations can further improve the performance of the model for traffic flow prediction. The three worst-accuarcy models, i.e., Model 1, Model 2, and Model 4, all use less neighboring flows compared to Model 16 and other superior moels.

In consequence, according to the performance of different models demonstrated in Table III, we can see that the prediction accuracy of short-term traffic flow for a given detector station is influenced obviously by the upstream and/or downstream traffic condition. And historical traffic flow data of the neighboring stations can be used to enhance the accuarcy of the prediction model.

IV. Conclusion

In this paper, we have studied the effects of various inputs, including traffic flow, occupancy and speed, and the neighboring traffic information on the performance of short-term traffic flow prediction. Long short-term memory (LSTM) recurrent neural network, which can learn time series with long time dependency is applied to provide a flexible framework to adopt different combinations of variables. Several empirical analyses are conducted using traffic data collected from California Performance Measurement System (PeMS). The results have suggested that:

1. Using traffic flow alone as input variables can yield the acceptable performance of traffic flow prediction, while using traffic flow together with occupancy/speed as inputs may yield better results than using traffic flow alone.

2. The inclusion of the neighboring traffic flow information can improve the performance of traffic flow prediction.

Short-term traffic flow prediction is vital for intelligent transportation systems, and more spatial and temporal traffic information could be taken into account for the accurate traffic prediction in a larger scale road network in the future work.

References

[1] Y. J. Duan, Y. S. Lv, J. Zhang, X. L. Zhao, F. Y. Wang, Deep Learning for Control: The State of the Art and Prospects, Acta Automatica Sinica, Vol. 42 Issue (5): 643-654, 2106

[2] Vlahogianni, Matthew G. Karlaftis, John C. Golias, Short-term traffic forecasting Where we are and where we're going, Transportation Research Part C: Emerging Technologies, Volume 43:3-19, June 2016

[3] X. Zheng, W. Chen, P. Wang, D. Y. Shen, Big Data for Social Transportation, Transactions on Intelligent Transportation System, 17(3): 620-630, 2016

[4] X. Wang, X. Zheng, Q. P. Zhang, T. Wang, Crowdsourcing in ITS: The State of the Work and the Networking, IEEE Transactions on Intelligent Transportation System, 17(6): 1596-1605, 2016

[5] L. Li, Y. S. Lv, F. Y. Wang, Traffic Signal Timing via Deep Reinforcement Learning. IEEE/CAA Journal of Automatica Sinica, 3(3): 247-254, 2016

[6] Y. S. Lv, Y. Chen, X. Q. Zhang, Y. J. Duan, Social Media Based Transportation Research: the State of the Work and the Networking, IEEE/CAA Journal of Automatica Sinica,4(1): 19-26, 2017

[7] E. I. Vlahogianni, J. C. Golias, and M. G. Karlaftis, Short-term traffic forecasting: Overview of objectives and methods, Transportation Research Part C., vol. 24, no. 5, pp. 533–557, Sep. 2004

[8] C. P. Van Hinsbergen, J. W. Van Lint, and F. M. Sanders, Short term traffic prediction models, presented at the ITS World Congress, Beijing, China, 2007

[9] L. Li, X. N. Su, Y. Zhang, Y. Lin, Trend Modeling for Traffic Time Series Analysis: An Integrated Study, IEEE Transactions on Intelligent Transportation System, VOL. 16, NO. 6, 12, 2015

[10] J. Zhang, F. Y. Wang, K. Wang, W. H. Lin, X. Xu, Data-Driven Intelligent Transportation Systems : A Survey, IEEE Transactions on Intelligent Transportation Systems, 12(4):1624-1639, 2001

[11] M. vanderVoort, M. Dougherty, and S. Watson, Combining Kohonen maps with ARIMA time series models to forecast traffic flow, Transportation Research Part C: Emerging Technologies, vol. 4, no. 5, pp. 307–318, Oct. 1996

[12] S. Lee and D. Fambro, Application of subset autoregressive integrated moving average model for short-term freeway traffic volume forecasting, Transportation Research Part C: Emerging Technologies, vol. 1678, pp. 179–188, 1999

[13] Y. Kamarianakis and P. Prastacos, Forecasting traffic flow conditions in an urban network—Comparison of multivariate and univariate approaches, Transportation Research Part C, no. 1857, Transporation Network Modeling 2003: Planning and Administration. pp. 74–84, 2003

[14] B. M. Williams and L. A. Hoel, Modeling and forecasting vehicular traffic flow as a seasonal ARIMA process: Theoretical basis and empirical results, Transportation Research Part C: Emerging Technologies, vol. 129, no. 6, pp. 664–672, Nov./Dec. 2003

[15] B. Ghosh, B. Basu, and M. O'Mahony, Multivariate short-term traffic flow forecasting using time-series analysis, IEEE Transactions on Intelligent Transportation System, vol. 10, no. 2, pp. 246–254, Jun. 2009

[16] H. Chang, Y. Lee, B. Yoon, and S. Baek, Dynamic near-term traffic flow prediction: System oriented approach based on past experiences, IEEE Transactions on Intelligent Transportation System, vol. 6, no. 3, pp. 292–305, Sep. 2012

[17] G. A. Davis and N. L. Nihan, Nonparametric regression and short-term freeway traffic forecasting, Transportation Research Part C: Emerging Technologies, vol. 117, no. 2, pp. 178–188, Mar./Apr. 1991

[18] K. Y. Chan, T. S. Dillon, J. Singh, and E. Chang, Neural-network-based models for short-term traffic flow forecasting using a hybrid exponential smoothing and Levenberg–Marquardt algorithm, IEEE Transactions on Intelligent Transportation System, vol. 13, no. 2, pp. 644–654, Jun. 2012

[19] M. Castro-Neto, Y. S. Jeong, M.-K. Jeong, and L. D. Han, Onlinesvr for short-term traffic flow prediction under typical and atypical traffic conditions, Expert systems with applications, vol. 36, no. 3, pp. 6164–6173, 2009

[20] 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 Transactions on Intelligent Transportation System, vol. 14, no. 4, pp. 1700–1707, Dec. 2013

[21] Karlaftis, M.G., Vlahogianni, E.I, Statistical methods versus neural networks in transportation research: differences, similarities and some insights. Transportation Research Part C: Emerging Technologies, 19, 387–399, 2011

[22] W. Z. Zheng, D. H. Lee, and Q. X. Shi, Short-term freeway traffic flow prediction: Bayesian combined neural network approach, Transportation Research Part C: Emerging Technologies, vol. 132, no. 2, pp. 114–121, Feb. 2006

[23] M. Zhong, S. Sharma, and P. Lingras, Short-term traffic prediction on different types of roads with genetically designed regression and time delay neural network models, Journal of Computing in Civil Engineering, vol. 19, no. 1, pp. 94–103, Jan. 2005

[24] E. I. Vlahogianni,M. G. Karlaftis, and J. C. Golias, Optimized and metaoptimized neural networks for short-term traffic flow prediction: A genetic approach, Transportation Research Part C: Emerging Technologies, vol. 13, no. 3, pp. 211–234, Jun. 2005

[25] K. F. Wang, C. Gou, N. N. Zheng, Parallel vision for perception and understanding of complex scenes: methods, framework, and perspectives, Artificial Intelligence Review, vol. 48, no. 3, pp. 298-328, Oct. 2017

[26] K. F. Wang, C. Gou,Y. J. Duan, Y. L. Lin, Generative adversary networks: introduction and outlook, IEEE/CAA Journal of Automatica Sinica, 4(3):389-39, 2017

[27] Y. Yuan, F. Y. Wang, Towards blockchain-based intelligent transportation systems, In Proc. of the 2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC2016), Rio de Janeiro, Brazil, 2016

[28] Y. Yuan, F. Y. Wang, Blockchain: the state of the art and future trends, Acta Automatica Sinica, 42(4):481-494, 2016

[29] Y. S. Lv, Y. J. Duan, W. W. Kang, Z. X. Li, F. Y. Wang, Traffic Flow Prediction With Big Data: A Deep Learning Approach, IEEE Transactions on Intelligent Transportation Systems, Volume: 16, Issue: 2 Pages: 865 – 873, 2015

[30] Y. J. Duan, Y. S. Lv, W. W. Kang, Performance Evaluation of the Deep Learning Approach for Traffic Flow Prediction at Different Times, IEEE 17th International Conference on Intelligent Transportation Systems (ITSC). OCT 08-11, 2016

[31] G. Xiong, W. Kang, F. Wang, F. Zhu, Y. Lv, X. Dong, J. Riekki, and S. Pirttikangas, Continuous Travel Time Prediction for Transit Signal Priority Based on a Deep Network, 2015 IEEE 18th International Conference on Intelligent Transportation Systems (ITSC), pp. 523-528, 2015

[32] X. Ma, Z. Tao, Y. Wang, Long short-term memory neural network for traffic speed prediction using remote microwave sensor data, transportation Research Part C: Emerging Technologies, Volume: 54 Pages-197

[33] Y. X. Tian, P. Li, Predicting Short-term Traffic Flow by Long Short Term Memory Recurrent Neural Network, 2015 IEEE International Conference on Smart City/SocialCom/SustainCom (SmartCity), Pages: 153 – 158, 2015

[34] R. Fu, Z. Zhang, L. Li, Using LSTM and GRU neural network methods for traffic flow prediction, Youth Academic Annual Conference of Chinese Association of Automation (YAC), Pages: 324 – 328, 2016