2018 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovations

Deep Convolutional Mesh RNN for Urban Traffic Passenger Flows Prediction

Zhene Zou[1], Hao Peng[1,2], Lin Liu[1], Guixi Xiong[1], Bowen Du[1,2], Md Zakirul Alam Bhuiyan[3], Yuntao Long[4], Da Li[5]

[1] State Key Laboratory of Software Development Environment, Beihang University, Beijing, China

[2] Beijing Advanced Innovation Center for Big Data and Brain Computing, Beihang University, Beijing, China

[3] Department of Computer and Information Sciences, Fordham University, New York, USA

[4] Institutes of Science and Development, Chinese Academy of Sciences, Beijing, China

[5] Electrical and Computer Engineering Department, University of Missouri-Columbia, Columbia, USA

Email: zouzhene,penghao,liulin,xgx,dubowen@buaa.edu.cn, mbhuiyan3@fordham.edu, yuntaol@casipm.ac.cn, dlx7f@mail.missouri.edu,

***Abstract*—Urban traffic passenger flows prediction is practically important to facilitate many real applications including transportation management and public safety. Sustained and rapid economic growth requires an orderly organization, and planning is an indispensable part of an orderly organization process. The reduction in travel efficiency due to traffic congestion, as well as energy and various pollution issues from the transportation sector, have become the bottleneck for the further development of the city and are the most troublesome topic for governments in all countries. Recently, deep learning performs the excellent ability to extract high dimensional spatial-temporal characters in regression and classification tasks. In this paper, we propose a deep learning model based on CNN and RNN, which takes matrixed traffic as input, uses CNN to extract traffic characteristics, and uses RNN to predict the evolution of features to achieve traffic flow prediction. Instead of traditional rnn models, we design a new type of RNN structure unit that can process time data in multiple time dimensions at the same time. Using a network-like RNN model, the evolution of traffic flow in different time dimensions is fully considered, and the interaction between different time dimensions is taken into account to predict the traffic flow of the target time series. The prediction of each data in the sequence has real data as input instead of merely taking the output of the previous moment as the input of the next moment. Experiments show that our model can significantly improve the prediction accuracy for real traffic passenger flow datasets.**

***Index Terms*—Traffic Passenger Flows Prediction; CNN; RNN; Urban Computing**

I. Introduction

Predicting traffic flows in a city is of great importance to public security and traffic management. As an important link, urban transportation is linked to urban economic activities at one end and urban social activities at the other end. Urban transport is conducive to the sustainable and healthy development of urban economy [1], [2], [3]. In the field of transportation, the importance of planning is reflected by the increasing traffic problems. In both developed and developing countries, the reduction in travel efficiency due to traffic congestion, as well as energy and various pollution issues from the transportation sector, have become the bottleneck for the further development of the city and are the most troublesome topic for governments in all countries. On the one hand, traffic problems is becoming increasingly serious. How to make full use of existing road resources and improve traffic efficiency has become one of the focuses of future traffic research. On the other hand, major cities in large and medium-sized cities in China are synchronizing with the world in the collection of information on city roads and the perfection of transmission facilities. This allows us to obtain a lot of traffic data and make traffic forecasts more effective. As urbanization progresses and the urban population grows, it becomes increasingly challenging to accurately predict the inflow and outflow of each transportation site. This is mainly due to the fact that traffic flow may be affected by various dynamic and complex factors such as dynamic traffic routes, upgrading of transport facilities, complex traffic flow, peak hour effects, and other extreme factors such as bad weather. Actually the change of passenger flow is affected by multiply and complex factors including urban planning, population growth, gregarious distribution effect and stochastic events like accidents or bad weathers [4].

Most traditional methods represent urban traffic passenger flow as two matrices, each of which represents the city's regional inflow or outflow data. Although the regional traffic passenger flow model is a direct and effective way to characterize the traffic flow in a city, it ignores the independence and interaction between different traffic lines. Therefore, they may not be suitable for traffic-level or site-free traffic forecast.

In recent years, the deep learning method has achieved great success in computer vision and natural language processing [5], [6], [7], and has also achieved good results in other areas [8]. The achievements of deep learning in the field of transportation also deserve our attention.Many researchers have used deep learning methods for traffic data processing and have achieved good results [9], [10]. For example, considering the traffic flow matrix as a snapshot of a fixed time interval, a convolutional neural network (CNN) uses a multi-layer convolution structure to extract the spatial correlation of passenger flows. Some works also use deep learning models to extract

978-1-5386-9380-3/18/$31.00 ©2018 IEEE
DOI 10.1109/SmartWorld.2018.00227

the spatio-temporal traffic passenger flows features [11], [12]. For example, considering the traffic flow matrix as a snapshot of a fixed time interval, a convolutional neural network (CNN) uses a multi-layer convolution structure to extract the spatial dependence of the passenger flow [13]. The Recurrent Neural Network (RNN) continuously updates the internal state of the network to study the temporal characteristics of traffic evolution [14]. As a combination of CNN and RNN, the hybrid model is designed to simultaneously extract spatio-temporal features to improve prediction accuracy [15].

In this paper, we propose a deep learning model based on CNN and RNN, which takes matrixed traffic as input, uses CNN to extract traffic characteristics, and uses RNN to predict the evolution of features to achieve traffic flow prediction. In summary, the main contributions of this article are as follows.

- We divide the traffic flow into input traffic and output traffic, and construct two traffic flow matrices. At the same time, using the CNN-based model to extract the spatial characteristics of traffic flow.
- We propose a new type of rnn structure unit that can process time data in multiple time dimensions at the same time. Using a network-like RNN model, the evolution of traffic flow in different time dimensions is fully considered, and the interaction between different time dimensions is taken into account to predict the traffic flow of the target time series. The result of the prediction is a time series. The prediction of each data in the sequence has real data as input instead of merely taking the output of the previous moment as the input of the next moment.
- Considering the influence of external factors on traffic flow prediction, traffic flow can be adjusted in real time according to external factors. Unlike other models, we fuse the effects of external factors on our results into our main model rather than adding them as independent factors and results.
- We evaluate the performance of the proposed model for real traffic passenger flow data sets and use baseline experiments to prove the effectiveness of our model.

The rest of our paper is organized as follow. In section 2, we review the related works for traffic flow prediction. In section 3 , we introduce the architecture and the details of prediction algorithms. In section 4, we verify the validity of our model with the real traffic flow dataset. In section 5, we conclude the paper.

II. Related Work

In this section, we will briefly review the work related with the paper. Traffic passenger flow prediction plays a very important role in Intelligent Transportation System (ITS) [16]. In recent decades, various methods based on various theories have been proposed and verified through actual data sets of different scenarios, thereby improving the stability and safety of the entire transportation system. Traffic flow forecasting model can be roughly divided into traditional shallow forecasting model and deep learning based forecasting model.

Traffic prediction problems mainly pay more attention on parametric methods include methods based on ARIMA [17], [18] and SVM-based methods [19]; nonparametric methods include K-nearest neighbor (KNN) nonparametric regression methods [20] and Gaussian Process methods [21]. Neural networks (NNs) have also been used on a large scale in the transportation sector [22]. Neural network-based prediction methods can fit complex nonlinear relationships, so it is suitable for traffic prediction [23], [24]. In order to make the prediction model more effective, some researchers have tried to consider various factors, some of which combine multiple traffic data to predict, some consider traffic-related events together with traffic flow, including traffic events, weather, festivals, etc [25], [26]. Traditional traffic forecasting methods are usually used to forecast traffic at a specific location, and it is not possible to predict multiple locations within a single area. Our model fully considers the spatial correlation of multiple locations in the area and predicts their traffic flow at the same time.

Recently, some researchers have also tried to apply the deep learning model to traffic prediction and achieved good results [1], [2]. The existing deep learning model for traffic flow prediction inspired by the computer vision algorithm. The space-time deep learning model has been applied to urban traffic speed prediction tasks such as ST-ResNet [1]. It considers the evolution law of traffic flow in different time dimensions, considers the influence of external factors on traffic flow, and uses residual neural network method to predict spatio-temporal flow, so as to achieve very good performance. In order to model the evolution of traffic flow over time, the RNN model continuously updates its cellular state by accepting time-series passenger flow inputs [27]. In particular, the improved RNN, LSTM, uses input gates, ignores gates and output gates to control changes in their own state, solves the problem of gradient disappearance to a certain extent. Researchers have used the LSTM model to predict traffic flow and achieved good results [14]. [28] proposes the popular DCRNN model, which treats traffic flow by graph convolution and predicts future traffic flow sequences. Although [28] can predict the traffic flow of a time series, its model uses the real data to predict the value except for the first time. The prediction of the back time only uses the results of the previous moments, but there is no real data as input. [1] proposed the popular ST-ResNet model which samples at a regular internals for closeness, period, trend and external influence on Residual Networks and gains a hopeful performance. Although [1] considers different evolutionary rules of traffic flow in multiple time dimensions, it predicts multiple time dimensions, and finally fuses multiple predictive values, ignoring the influence of multiple time dimensions on each other. Most models do not consider the impact of external factors on traffic flow, such as weather, holidays, workdays, etc. A few models take into account the influence of external factors, but it predicts traffic flow as an independent part of external factors and fuses the results with the results of the main model.

III. Model

In this section, we will detail the traffic flow prediction deep learning model based on CNN and RNN. The architecture of this model is shown in Figure 1. The model consists of three main parts, residual convolutional layer, mesh RNN layer and fusion layer. As shown, we can see that the input data is first processed by an convolution filter.

The parameter sharing convolution module is introduced to extract the spatial-temporal features among multiple channels in traffic passenger flow. Next, using the features extracted by CNN as input, the mesh RNN model is used to predict the evolution of features in different time dimensions. Finally, using predicted traffic flow features, combined with external event factors, predict traffic flow over a period of time in the future.

A. Spatial Feature Extraction

Deep CNN has higher accuracy than traditional algorithms in the field of image recognition and classification. The convolution module extracts high-dimensional features from the color matrix to interpret the information contained in the image. In the scene of traffic flow prediction, we also consider the multidimensional traffic flow matrix as a multi-dimensional color image. For each moment of traffic flow input, the size of the graph matrix is $(M, N, 2 \times D)$, M, N represent the spatial coordinates of the traffic flow, and D denotes the traffic flow is distributed on D traffic lines, $2 \times D$ indicates that each site's traffic flow is divided into incoming traffic and outgoing traffic. We implement a multi-layer convolutional layer, which transforms the input matrix into a feature vector containing a high-dimensional representation of the flow dimension model. We need a deep network to capture very large urban dependencies. To solve this problem, we used residual unit in our model. Each Residual unit includes two batch normalizations layer, two activation layer with a "Relu" function and two convolutional layer with a kernel of size $(3, 3)$, as Figure2 shows. When using batch data to train the model, the input and output distribution of each intermediate layer changes partially. The batch normalization layer maintains the normal distribution of batch data and improves training efficiency.

B. RNN Cell

Since our RNN model needs to consider the evolution of traffic flow in different time dimensions at the same time, the traditional RNN cell obviously cannot solve this problem. Therefore, we have designed a new RNN cell that can receive traffic flow data from different time dimensions, taking into account the mutual influence between them, memorizing the evolution of traffic flow in different time dimensions, and outputting Traffic flow characteristics are predicted in different time dimensions, as Figure 3 shows. The data received by this RNN cell contains three kinds of data, spatial characteristics of the traffic flow at the current moment (obtained by the feature extraction module), the output of the RNN cell at the last moment of the time dimension 1, C1, H1, and the RNN cell at the last moment of the time dimension 2, C2, H2. The three kinds of data are merged. As shown in the figure, two results are obtained. The characteristics of traffic flow at the next moment in time dimension 1,C1',H1', and the characteristics of traffic flow at the next moment in time dimension 2, C2', H2'.

C. Mesh RNN

For the previous models, they only historical data of the traffic flow in the last few hours are considered, or, although the evolution of traffic flow in different time dimensions is considered, the predictions in different time dimensions are predicted separately. These models lack interaction between traffic flows in different time dimensions. Because traffic flow data shows different rules at different time spans, considering the evolution of traffic flow in different time dimensions can improve the accuracy of traffic flow prediction. Our experimental results also verify this point. In our model, we consider two temporal dimensions of traffic flow data as input data, namely traffic flow data for the last few hours and traffic flow data at the same time in recent days. That is, we not only considered the evolution of traffic flow data in the last few hours, but also considered the evolution of traffic flow data at the same moment in recent days.

D. Transfer

Through the above-mentioned mesh RNN model, we can learn the traffic flow characteristics of the target time. This traffic flow feature consists of two parts, the characteristics of traffic flow in the hourly dimension and the characteristics of traffic flow in the daily dimension, denoted as X_1, X_2 . Through the two traffic flow characteristics and historical traffic flow data, the traffic flow at the target time can be predicted. First, X_1 generates a matrix of the same size as the traffic stream matrix through the fully connected network, denoted as X_{11}, and thenX_1 generates a matrix with the same size as the traffic stream matrix and activates it with "sigmod" function, which is denoted as X_{12}. In the same way, we can get X_{21}, X_{22}. By calculating $Y = X_{11} + X_{12} \times M + X_{21} + X_{22} \times N$, the traffic flow at the target time can be calculated, where M is the traffic flow of an hour ago and N is the traffic flow at the same time a day ago.

E. Sequence Prediction

The proposed model not only can be used to predict traffic flow data at a certain moment, but also can be used to predict traffic flow at multiple consecutive moments. The prediction of traffic flow at a single time is basically the same as previously stated. For traffic flow prediction at multiple consecutive moments, the traffic flow prediction at the first moment is also consistent with the previous one. For the prediction of the following moments, the input of the hour dimension corresponding to the RNN cell at the time of prediction is actually not the real data but the output of the model at the previous moment. This will cause some errors in the prediction of the current time. Fortunately, our model considers multiple time dimensions. The input of the day dimension corresponding to the RNN cell is real data, that is, the traffic flow at the same time yesterday is real data, which greatly Improve the reliability of model prediction. For the problem that the left part of input just mentioned is not real data, we also proposed a solution. In training, we can use the real data of the hourly dimension as the input data of the

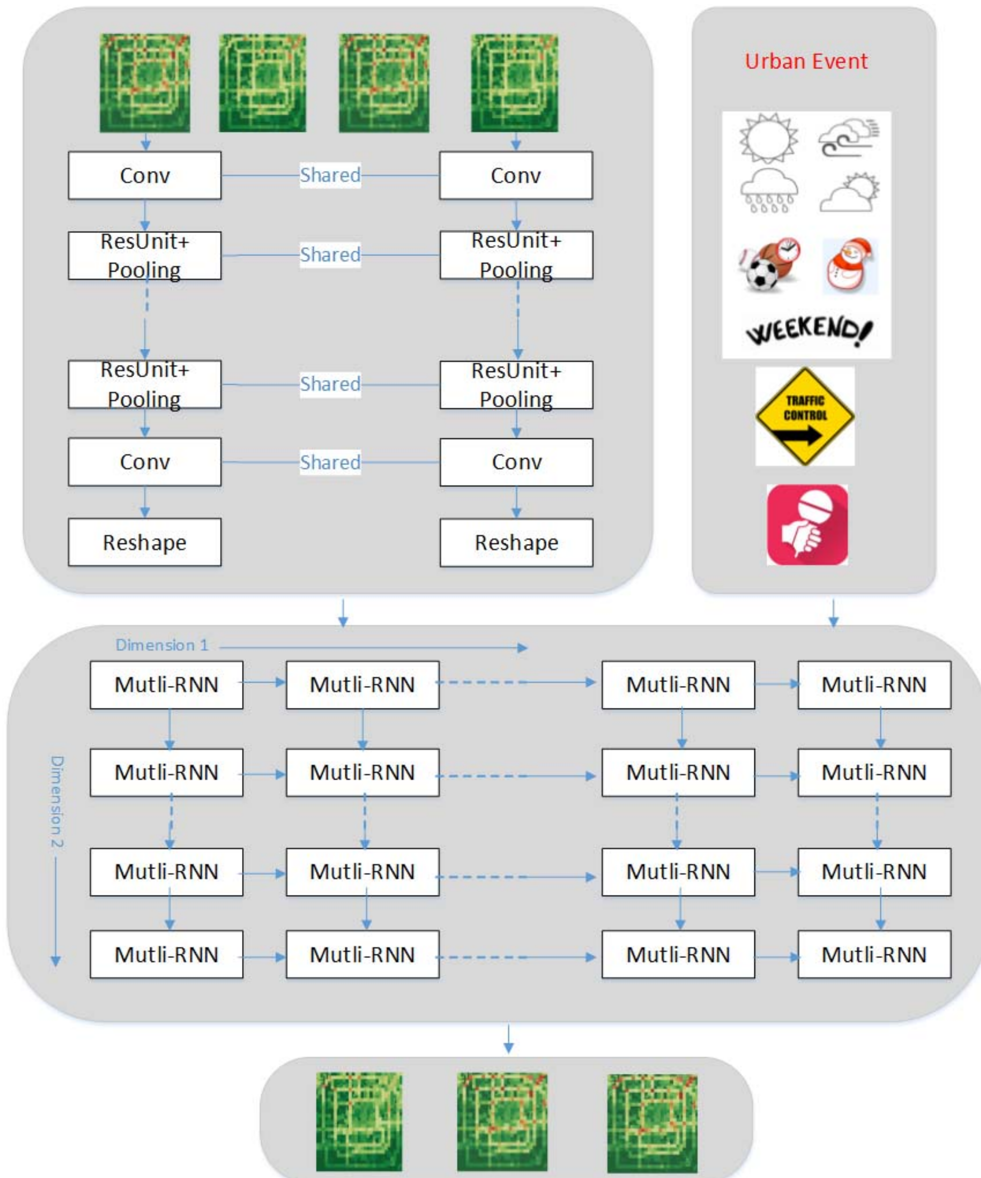

Fig. 1. The framework of the proposed DCMRNN. Conv: Convolution; ResUnit+Pooling: Residual Unit and Max Pooling; cell: the cell of mesh RNN.

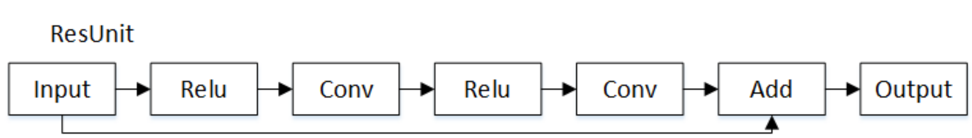

Fig. 2. Residual unit

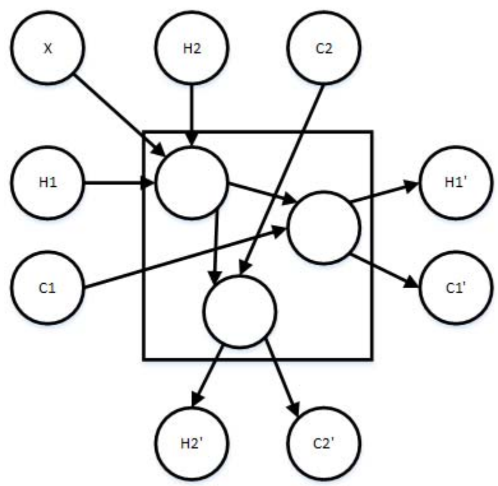

Fig. 3. RNN Cell

current moment with a variable probability. In the prediction, the predicted value of the previous moment is used as the input of the current moment. Need to maintain a variable to control the probability of using real data.

F. External Component

The previous model rarely considered the impact of external factors on the outcome of the forecast. Some models take into account the influence of external factors, but the method used is far-fetched. External factors are used as an independent module to directly calculate the influence of external factors on traffic flow through external factors.We convert external factors into vector representations, denoted as E. E generates a matrix of the same size as the traffic stream matrix through the fully connected network, denoted as E_1, and then$E$ generates a matrix with the same size as the traffic stream matrix and activates it with "sigmod" function, which is denoted as E_2. By calculating $Z = E_1 + E_2 \times Y$, the traffic flow at the target time can be calculated, where Y is the traffic flow without the influences of external factors.

IV. EXPERIMENT

In this section, we will conduct extensive experiments to evaluate the proposed model.

A. Datasets

We selected the following two datasets for evaluation, the New York City Bicycle Dataset and the Beijing Subway Dataset. We can see that these two data sets represent different types of traffic patterns. Table I lists the statistics for these two dataset statistics.

- New York Bikes Dataset: This data was published by the New York City Bike System from April 1st to September 30th, 2014. Each trip data includes trip duration, start and end station IDs, and start and end times. We adopt the last 10 days of data as test data and the rest of the data as training data.

Dataset	SubwayBJ	BikeNYC
Data type	e-card	Bike rent
Location	Beijing	New York
Time Span	7/1/2016-10/1/2016	4/1/2014-9/30/2014
Data frequency	30 minutes	1 hour
Grid region size	(128,128)	(16,8)

TABLE I
DATASET STATISTICS

- Beijing Subway Dataset: Beijing subway data comes from the people's departure records of their Beijing metro-card. The data spans from July 1 to October 30, 2016. There are 18 lines in the Beijing Subway. The recording interval is 30 minutes. For this work, we obtained two types of crowd flow and a matrix representation of 2×18 channels. We built a 128×128 grid area in Beijing for this data set. We use the first three and a half months of data for training and the remaining half of the data is used for testing.

B. Baselines and Benchmark

We compared the proposed depth prediction model with 6 traditional benchmark algorithms or models.

- Auto-Regressive Integrated Moving Average (ARIMA): The theory of auto-regressive integrated sliding average algorithm was proposed and applied to traffic flow prediction first in the early 1970's.
- SARIMA: It is a seasonal ARIMA model.
- Vector Auto-Regressive (VAR): It captures the pairwise relationship between all traffic, which is an advanced space-time model, and is computationally expensive.
- ST-ANN: This is a very deep model. It takes as input the relevant traffic passenger data of space (nearly 8 regional values) and time (8 previous time intervals) and then feeds them to the artificial neural network.
- DeepST: It uses a deep neural network (DNN)-based predictive model and models spatio-temporal data into temporal compactness, cycle, and seasonal trends.
- ST-ResNet : It is the most advanced deep convolution-based residual network currently used to predict future urban passenger flow.

In order to evaluate the accuracy of different models, we propose to use the root mean squared error (RMSE) as a measure

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

Where $\hat{y}_i$ and y_i are the corresponding predicted traffic flow matrix and future real traffic flow matrix, m is the number of samples of the verification set or test set.

C. Experimental Settings

All experiments of this paper are conducted on 4 core Intel Core I5-6500 @3.20GHz cpu, with 12GB RAM and Nvidia Geforce GTX-1060 GPU. The operating system and software

Model	RMSE in SubwayBJ	RMSE in BikeNYC
ARIMA	0.0193	10.07
SARIMA	0.0217	10.56
VAR	0.0174	9.92
DeepST	0.0113	7.43
ST-ResNet	0.0079	6.33
DCMRNN	0.0068	5.82

TABLE II
COMPARISON AMONG DIFFERENT METHODS ON SUBWAYBJ AND BIKENYC

platforms are Windows 10, Keras 2.0.5, Tensorflow 1.2.0 and Python 3.5. The size of convolutional kernels is 3×3. The Batch size is set to 32 and we use Adam optimizer to train the models.

D. Results of different models

Table II shows the results of our DCMRNN model and other baseline for SubwayBJ and BikeNYC datasets. In SubwayBJ dataset, the DCMRNN model can reduce the RMSE to 0.0068, which is the best performance of all the methods. Similarly, the DCMRNN model can reduce the RMSE to 5.82 in BikeNYC dataset.

V. CONCLUSIONS

In this paper, we propose a new model based on CNN and mesh RNN for predicting crowd flow in urban traffic routes. Our DCMRNN model fuses residual network-based CNN networks to extract features of complex traffic flows. Considering that the traffic flow shows different evolution rules in different time dimensions, and the mutual influence of traffic flow between different time dimensions, a comprehensive RNN model is used to learn the evolution of traffic flow over time. At the same time, we consider some factors that will affect traffic flow, and consider the impact of these factors on traffic flow in the model. The experimental results show that our modeling method is effective and greatly improves the accuracy of the traffic prediction tasks.

VI. ACKNOWLEDGMENTS

The corresponding authors are Bowen Du and Hao Peng. This work is supported by NSFC program (No.61772151,61702552,51778033). We also thank the anonymous reviewers for their valuable comments and suggestions that help improve the quality of this manuscript.

REFERENCES

[1] J. Zhang, Y. Zheng, and D. Qi, "Deep spatio-temporal residual networks for citywide crowd flows prediction," 2016.
[2] Y. Lv, Y. Duan, W. Kang, Z. Li, and F. Y. Wang, "Traffic flow prediction with big data: A deep learning approach," *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 2, pp. 865–873, 2015.
[3] T. J. Kim, L. L. Wiggins, and J. R. Wright, "Expert systems: Applications to urban planning," vol. 2, p. 173182, 1990.
[4] M. X, D. Z, H. Z, M. J, W. Y, and W. Y, "Learning traffic as images: A deep convolutional neural network for large-scale transportation network speed prediction," *Sensors (Basel, Switzerland)*, vol. 17, no. 4, 2017.
[5] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," pp. 770–778, 2015.
[6] Y. S. Y. L. Hao Peng, Jianxin Li, "Incrementally learning the hierarchical softmax function for neural language models," 2017.
[7] H. Peng, J. Li, Y. He, Y. Liu, M. Bap, L. Wang, Y. Song, and Q. Yang, "Large-scale hierarchical text classification with recursively regularized deep graph-cnn," in *WWW*, 2018.
[8] S. Manoharan, "Short term traffic flow prediction using deep learning approach," December 2016.
[9] N. G. Polson and V. O. Sokolov, "Deep learning for short-term traffic flow prediction," *Transportation Research Part C Emerging Technologies*, vol. 79, pp. 1–17, 2017.
[10] A. Abadi, T. Rajabioun, and P. A. Ioannou, "Traffic flow prediction for road transportation networks with limited traffic data," *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 2, pp. 653–662, 2015.
[11] S. Zhang, G. Wu, J. P. Costeira, and J. M. F. Moura, "Fcn-rlstm: Deep spatio-temporal neural networks for vehicle counting in city cameras," in *IEEE International Conference on Computer Vision*, 2017, pp. 3687–3696.
[12] X. Cheng, R. Zhang, J. Zhou, and W. Xu, "Deeptransport: Learning spatial-temporal dependency for traffic condition forecasting," 2017.
[13] X. Ma, H. Yu, Y. Wang, and Y. Wang, "Large-scale transportation network congestion evolution prediction using deep learning theory," *Plos One*, vol. 10, no. 3, p. e0119044, 2015.
[14] Y. Tian and L. Pan, "Predicting short-term traffic flow by long short-term memory recurrent neural network," in *IEEE International Conference on Smart City/socialcom/sustaincom*, 2015, pp. 153–158.
[15] H. Yu, Z. Wu, S. Wang, Y. Wang, and X. Ma, "Spatiotemporal recurrent convolutional networks for traffic prediction in transportation networks:," *Sensors*, 2017.
[16] N. Zhang, F. Y. Wang, F. Zhu, D. Zhao, and S. Tang, "Dynacas: Computational experiments and decision support for its," *IEEE Intelligent Systems*, vol. 23, no. 6, pp. 19–23, 2008.
[17] B. Williams, P. Durvasula, and D. Brown, "Urban freeway traffic flow prediction: Application of seasonal autoregressive integrated moving average and exponential smoothing models," *Transportation Research Record*, vol. 1644, no. 1, pp. 132–141, 1998.
[18] B. M. Williams and L. A. Hoel, "Modeling and forecasting vehicular traffic flow as a seasonal stochastic time series process," *Dissertation Abstracts International, Volume: 60-01, Section: B, page: 0292.;Adviser: Lester A. Hoel*, 1999.
[19] M. Tan, L. I. Yingjun, and X. U. Jianmin, "A hybrid arima and svm model for traffic flow prediction based on wavelet denoising," *Journal of Highway & Transportation Research & Development*, vol. 26, no. 7, pp. 127–131, 2009.
[20] B. Sun, W. Cheng, P. Goswami, and G. Bai, "Flow-aware wpt k-nearest neighbours regression for short-term traffic prediction," in *Computers and Communications*, 2017, pp. 48–53.
[21] S. Thajchayapong, J. A. Barria, and E. Garcia-Trevino, "Lane-level traffic estimations using microscopic traffic variables," in *International IEEE Conference on Intelligent Transportation Systems*, 2010, pp. 1189–1194.
[22] C. Qiu, C. Wang, X. Zuo, and B. Fang, "A bayesian regularized neural network approach to short-term traffic speed prediction," pp. 2215–2220, 2011.
[23] S. Y. Yun, S. Namkoong, R. J.-H, S. S.-W, and C. J.-U, "A performance evaluation of neural network models in traffic volume forecasting," *Mathematical & Computer Modelling*, vol. 27, no. 9-11, pp. 293–310, 1998.
[24] V. Alarcon-Aquino and J. Barria, "Multiresolution fir neural-network-based learning algorithm applied to network traffic prediction," *IEEE Transactions on Systems Man & Cybernetics Part C*, vol. 36, no. 2, pp. 208–220, 2006.
[25] M. Cools, E. Moons, and G. Wets, "Assessing the impact of weather on traffic intensity," *Wea Climate Soc*, vol. 2, no. 1, pp. 60–68, 2008.
[26] E. Mai and R. Hranac, "Twitter interactions as a data source for transportation incidents," *Transportation Research Board Annual Meeting*, 2013.
[27] 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," *Transportation Research Part C*, vol. 54, pp. 187–197, 2015.
[28] Y. Li, R. Yu, C. Shahabi, and Y. Liu, "Diffusion convolutional recurrent neural network: Data-driven traffic forecasting," 2017.