A Hybrid Deep Learning Approach for Traffic Flow Prediction in Highway Domain

Zhe Wang[1,2](✉), Weilong Ding[1,2], and Hui Wang[3]

[1] School of Information Science and Technology, North China University of Technology, Beijing 100144, China
wwwzzz7116@163.com
[2] Beijing Key Laboratory on Integration and Analysis of Large-Scale Stream Data, Beijing 100144, China
[3] Beijing China-Power Information Technology Company Limited, Beijing 100192, China

Abstract. With the development of cities, intercity highway plays a vital role in people's daily travel. The traffic flow on the highway network is also increasingly concerned by road managers and participants. However, due to the influence of highway network topology and extra feature such as weather, accurate traffic flow prediction becomes hard to achieve. It is difficult to construct a multidimensional feature matrix and predict the traffic flow of the network at one time. A novel prediction method based on a hybrid deep learning model is proposed, which can learn multidimensional feature and predict network-wise traffic flow efficiently. The experiment shows that the prediction accuracy of this method is significantly better than existing methods, and it has a good performance during the prediction.

Keywords: Traffic flow prediction · Deep learning · Spatio-temporal data

1 Introduction

With the development of industrialization and infrastructure, the connection between cities have become more convenient. People's work and life are inseparable from the intercity highway. During the upgrades of road infrastructure and the increase in the number of vehicles, the traffic situation of highway has also been paid attention by city officer. In recent years, the rapid development of intelligent transportation systems has brought new solutions to traffic management in the domain of highway transportation [1]. By integrating and analyzing data generated by Internet of Things (IoT) devices, the intelligent transportation system can better provide suggestions and predictions for traffic management. One of the most important function in the intelligent transportation system is traffic flow evaluation. It can help managers better detect road conditions and make reasonable plans. So, many researchers are focusing on traffic flow prediction with different time and space scope.

With the development of IoT technology, types of data become diverse. These data can have more or less impact on the traffic flow [2]. For example, the impact of extreme

H. Gao et al. (Eds.): CollaborateCom 2020, LNICST 350, pp. 253–267, 2021.
https://doi.org/10.1007/978-3-030-67540-0_15

weather such as heavy rain and fog is obvious. As analogously for the weather data, the willing of travel on holiday and weekends is higher than that of weekdays. For weekdays, in morning and evening peak, traffic flow is significantly higher than the noon. Therefore, how to effectively employ the appropriate data type and fully explore the relationship between the data for traffic flow prediction has to be considered. At the same time, for the highway, it is a loop topology. The toll stations are closely related to distance. When a vehicle enters the highway from a certain station, it is necessary to pass the adjacent segments. Therefore, in highway domain, the benefits for the traffic flow prediction in the whole network are more valuable than the single one. So how to collect the data and predict the network traffic flow at the same time is the focus of the research.

In this paper, we propose a novel network traffic flow prediction method based on relevant data. We present a hybrid deep learning model to predict traffic flow. By collecting weather, date and other type data, we structure a multidimensional feature model. We predict the network traffic flow through fully convolution network (FCN) and the long short-term memory (LSTM) network. We solved the problems of multidimensional feature learning and network traffic flow prediction.

The rest of this paper is organized as follows: Sect. 2 explains the background of motivation and the related work; Sect. 3 introduces the feature modeling and the network traffic flow prediction method in detail; Sect. 4 evaluates the method with related experiments and results to show the performance. The fifth section summarizes the research and conclusion.

2 Background

2.1 Motivation

With the rapid growth of vehicle ownership, road congestion has become more serious. The intelligent transportation system allows road managers to manage traffic more intuitively and comprehensively. As an important function of the intelligent transportation system, traffic flow prediction has always been the focus of research and development. For example, Highway Big Data Analysis System in Chinese most populated province, Henan, can conduct business analysis by integrating historical toll station data [3]. Toll station data reveal the traffic flow around it. When the road becomes congested, the toll station data in the highway network will increase. Therefore, toll station data is the basis for us to predict traffic flow. Through the intelligent equipment and other information collection system installed in the toll station, we can get a record when vehicle passing toll station. As shown in Table 1 below.

As we can see, the toll record has three parts of dimension. In entity dimension, we use the vehicle details to identity the vehicle. In time dimension we can know the spatiotemporal attribute of vehicle. Like the time vehicle enter and exit the highway. In space dimension, we know the location information that the vehicle enters and exits the highway network. We can aggregate each record to determine the traffic flow at each toll station. It is very important for network traffic flow prediction. Through the data cleaning process, we can get the complete trajectory of the vehicles in highway network, including the static information of the vehicle and the dynamic information that vehicle passing through toll station.

Table 1. Highway toll station record

Attribute	Notation	Type
collector_id	Toll collector identity	Entity
car_id	Vehicle identity	
vehicle_type	Vehicle type	
etc_id	Vehicle ETC card identity	
entry_time	Vehicle entry timestamp	Time
exit_time	Vehicle exit timestamp	
entry_station	Identity of entry station	Space
entry_lane	Lane number of entry station	
exit_station	Identity of exit station	
exit_lane	Lane number of exit station	

Using this data, Highway Big Data Analysis System can not only build individual profile for vehicles, but also makes overall planning and analysis for highway network. As shown in the Fig. 1 below, the system can calculate the traffic flow information of each station on the highway day by day.

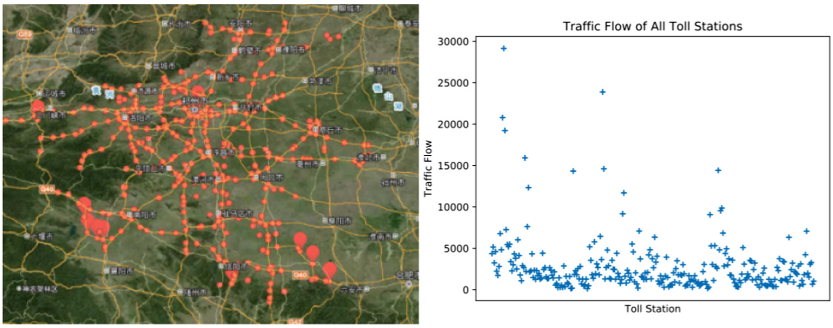

Fig. 1. Traffic flow of all toll stations in one day

As shown in the Fig. 1, the traffic flow is quite different at toll stations. Due to the agglomeration and planning of urbanization, the traffic flow around *zhengzhou*, the provincial capital of Henan Province, is significantly larger than that of its distant cities. Therefore, traffic flow prediction for key regional station can't represent the network situation. If we can predict traffic flow of all the toll stations in network at a time, it will provide global strategy support for highway managers. For network traffic flow prediction, due to the loop topology of the highway network, the traffic flow at the associated stations will also inevitably affect the traffic flow at a given station [3]. From entering the toll station to ending the journey, it will definitely pass adjacent segments which contain other toll stations. Those toll stations will also have impact on road traffic. Therefore, the traffic flow of associated station should be fully considered when predict. At the same time, there are other factors that will affect the traffic flow too.

One is the weather. When people are faced with heavy rain and fog, travel plans will be affected. It will lead to the fluctuation of traffic flow. Secondly, date type will also become one of influencing factors. From the daily analysis, the arrival of holidays will directly lead to a peak travel, and the traffic flow of that day will increase than usual. Therefore, how to fully consider weather, date, and other influencing factors, and find the correlation between data is the key issue of flow prediction. At the same time, choosing the appropriate prediction model will also let the traffic flow prediction of network achieved better results. In the era of rapid development of deep learning, shallow machine learning and deep machine learning have their own characteristics. How to choose a model that suitable for data and goals is also the focus of our research.

2.2 Related Work

As traffic flow prediction has gradually become a hot issue, more and more researchers have begun to use big data to predict traffic flow [4]. However, with the development of technology, lots of data in different types can be used. So, there are still challenges in traffic flow prediction domain. The first is the dimension of the features. Different data types will have different effects on traffic flow prediction. The second is prediction accuracy. With the development of deep learning, many researchers have begun to use deep learning methods to improve prediction accuracy. We divide the related work into two perspectives: feature dimension selection and model selection.

In feature dimension selection perspective, the traffic flow seems the traditional selection. Smith and Demetsky [5] use statistical and machine learning for prediction at first. It just used traffic flow data for analyze. However, with the development of science and technology, more and more dimensions of feature can be selected to improve the effect of traffic flow prediction. Now we often treat traffic flow prediction as a time series problem. Kumar and Vanajakshi [6] use seasonal Autoregressive Integrated Moving Average mode (ARIMA) model with limited input data for traffic flow prediction. Due to the structure of ARIMA model, only temporal dimension can be use. Now, more researchers begin to consider the impact of multiple features. Compared with a prediction model based on time series, these studies on the impact of multiple features are closer to the facts. For example, for prediction of air quality, Shengdong and Du [7] added meteorological data and population to compare and study the effect of these features on the experimental results. Meteorological data also can affect traffic flow. It gives us a inspire to ponder problem. Similar as highway domain, the prediction of pedestrian flow at railway stations is also based on loop topology. Niu [8] considers the impact of meteorological data and neighboring stations. In our research, associated stations also affect the traffic flow. In traffic flow study of Rong [9], he also referred to the driving speed and average speed of the vehicle. Cheng [10] regards the occupancy of roads as one of the features that affect the traffic flow. They all use multidimensional features for traffic flow analyzation and prediction, but their problem is that they can only predict one station one time. They lack the ability to learn multidimensional features on the entire network. Therefore, we chose meteorological data and date type for multidimensional feature. Moreover, in the domain of traffic flow prediction, the network traffic flow prediction has not been fully studied by researchers. Most researchers focus on the single station traffic flow, and then choose a better prediction method [8–10]. We focus on predicting the traffic flow of network at once. Because we want to learn the multidimensional feature of network precisely.

The second perspective is model selection. With the development of machine learning, model has become diverse. ARIMA is a typical time series learning model that can handle continuous data in the time dimension. Smith et al. introduced the construction and learning process of time series in detail [11–14]. But for the multidimensional feature matrix, the method of constructing time series is not suitable. Other machine learning methods such as Support Vector Regression (SVR), K-Nearest Neighbor (KNN) [15], etc. are also very popular. Luo [16–19] and others used KNN to learn the relationship between adjacent data points to predict traffic flow. Castro-Neto et al. [20, 21] used the kernel method in SVR to map the feature to high dimensions, and then made traffic flow prediction, which also achieved good results. However, with the development of deep learning, more researchers have begun to use deep learning methods to predict traffic flow, and achieved better results than shallow learning [22–25]. Zhang [26] et al. use deep belief network for feature learning and prediction. For traffic flow with time attributes, the time-step-oriented deep learning method LSTM is widely used, and for learning the feature matrix, the convolution method can also learn the relationship between features well. Therefore, a fully convolutional network and cooperates with the LSTM network is proposed as a hybrid deep learning model. It can fully learn the multidimensional feature, and predict the traffic flow of the whole network.

3 Method and Models

3.1 Feature Modeling

In highway traffic flow prediction domain, we find that there are many data types can affect it. Therefore, we propose a feature modeling method based on multiple types of data, which will uniformly model the various feature. We use variety of data to construct a multidimensional feature model suitable for network traffic flow prediction. Then we use a hybrid deep learning model to fit the feature model. Finally, we predict the traffic flow of network toll stations one time. For the prediction time range, we define daily traffic flow as the research background. The definition as follow:

Definition 1 (Daily traffic flow). For one toll station s in highway network, daily traffic flow is described as TF_s^t, represent the summation of vehicles exiting the toll station s in day t. Here, s is in the set of toll stations in highway network S, and t is the current day.

We choose the exiting record as the basis data for the summation of traffic flow. Because it is a completely record for the vehicle. In Chinese highway domain, only when the vehicle leaves toll station means tolls would be charged. After a complete record, we can still analyze the incoming information of the vehicle. Daily traffic flow is the main dimension for the prediction. But from elaborative observation, we found that other dimensions also affect the traffic flow of highway toll stations. Like weather, date type and related stations traffic flow, etc. those data types must be fully considered when modeling feature, so there a definition of feature modeling:

Definition 2 (multidimensional feature vector). In the day t, the multidimensional feature vector of daily traffic flow is $V_s^t = \left(W_s^t, D_t, Vol_s^t\right)$, toll station s $\in$ S in highway network. $W_s^t \in \{0, 1\}$ is weather category in date t, $D_t \in \{0, 1, 2\}$ is date type category in date t, $Vol_s^t = \left(TF_{s_1}^t, TF_{s_2}^t, TF_{s_3}^t, \ldots, TF_{s_n}^t\right)$, n $\in \mathbb{Z}^+$ represent *Daily traffic flow* on n nearest station's *Daily traffic flow* around *s* computed by the Euclidean distance in day t.

The following is further illustrations for the three features.

(1) In meteorological feature, the occurrence of extreme weather will affect the daily traffic flow of highway network. The raw weather data we have has the following properties: $Weather_s^t$ = *(weather condition, temperature, wind speed and direction)* in toll station s in highway network and day t. When weather condition is heavy rain, fog, or the degree of wind bigger than six-level, daily traffic flow of toll station s will drop. So, we define extreme weather $= (weathercondition = rain \vee fog \vee snow) \vee (windspeed > 6)$. And in matrix of daily traffic flow, we define

$$W_s^t = \begin{cases} 1, & extreme\ weather \\ 0, & otherwise \end{cases}$$

(2) In date feature, due to the holiday period, the highway network will adjust the charging mode, which leads to a significant influence in traffic flow. Therefore, the date feature is also one of the important factors that affect the traffic flow. Based on existing statistics, we define D_t in date t as follow:

$$D_t = \begin{cases} 2, & if\ t\ is\ a\ holiday, \\ 1, & if\ t\ is\ a\ weekend, \\ 0, & otherwise. \end{cases}$$

(3) In the last feature, Vol_s^t represent the spatial related traffic flow with toll station *s*. As we all know, highway network is a loop topology. The network traffic flow is influenced by all toll stations. For the toll station *s*, its adjacent stations' daily traffic flow is considered. The adjacent station means the toll station nearby *s* in spatial dimension. The traffic flow of adjacent station will influence the traffic flow of *s* apparently. We calculate the Euclidean distance between each toll station, take *n* nearest toll stations as the spatial relationship for feature modeling in spatial dimension.

With multidimensional feature vector, we can build a feature matrix for network toll stations. It is construct by all toll station so that we can use it to reflect network traffic flow. The definition as follow:

Definition 3 (Network toll station feature matrix). In day t, the network toll station feature matrix $M^t = \left(V_1^t, V_2^t, V_3^t, \cdots, V_k^t\right)$. k is the number of all toll stations in highway network.

Network toll station feature matrix represent the highway network's multidimensional feature that can reflect the network traffic flow in day t. It also is the feature modeling for the traffic flow prediction method. After feature modeling, due to the difference in the value range between each dimension, features with small range will be ignored, so the normalization method is used to reduce the prediction error.

3.2 Traffic Flow Prediction

In the following, we describe a traffic flow prediction method, based on hybrid deep learning model. The structure is shown in Fig. 2.

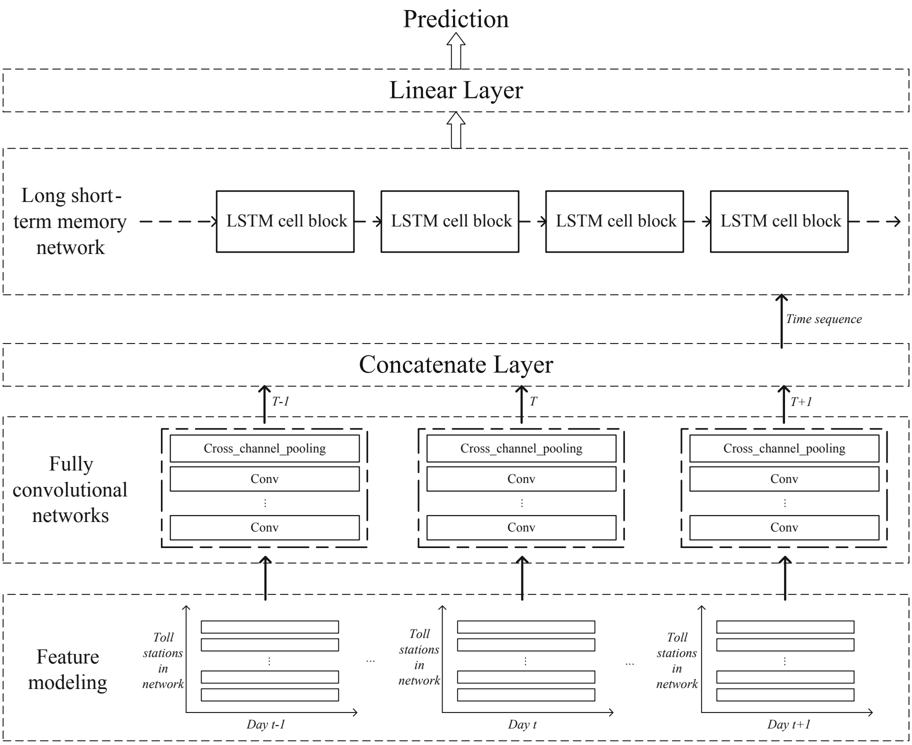

Fig. 2. Network traffic flow prediction method

The method is composed of fully convolution network (FCN) and the long short-term memory (LSTM). As we can see in Fig. 2, we propose to use FCN to fit multidimensional feature. Then we use concatenation layer to restrict the data into time series, then let LSTM to learn it. Finally, we use linear layer to restore network traffic flow. This combination considers both spatial and temporal of multidimensional feature. First of all, in feature modeling, the spatial relationship of toll stations in the highway network is fully considered. Based on one related work [27], three stations with min Euclidean distances from the toll station s have the most obvious impact, so n = 3 in *Definition* 2.

Three adjacent stations *daily traffic flow* was selected for feature modeling, plus weather and date type factors, which constituted the five-dimensional feature of each toll station. In order to learn the spatiality feature, we chose to use a special convolutional network. After that, we can see from Table 1 that the toll station data of the highway network has both spatial and temporal dimension, and it is not enough to only consider the spatiality. Therefore, after the convolution network, we add a concatenation layer to construct a time series and input to the LSTM network to learning temporal feature. Finally, a network traffic flow prediction method is formed through hybrid deep learning architecture to make prediction. The following is a detailed description of each component.

For feature modeling, we build *Network toll station feature matrix* M^t for all toll stations in network. This matrix is composed of all toll stations and *multidimensional feature vector* of each station. It contains spatial correlation feature and other dimension features that affect the traffic flow, like weather and date type. First, we use convolutional neural network for feature learning. Convolutional neural network (CNN) was first used in the field of image recognition. Through the calculation of pixels and convolution kernels, it learns the relationship of each pixel to achieve the role. The CNN involves many individual processing steps [28]. The general process includes convolution operation, pooling operation and the final fully connected layer with activation function. In our case, we use Cross-channel pooling instead of generate pooling operation. We complete CNN as follow functions.

$$\mathrm{o}\left(M^t\right) = \mathrm{Relu}\left(\mathrm{Conv}\left(M^t, \mathrm{kernel}\right)\right) \tag{1}$$

$$\mathrm{c}\left(M^t\right) = \mathrm{Cross\_channel\_pooling}\left(\mathrm{o}\left(M^t\right)\right) \tag{2}$$

where M^t is input feature matrix of day t. In Eq. 1, we choose hetero convolution kernel to calculate convolution. Usually, the pixels of the image are irregularly distributed. One feature in image often component by lots of pixels in different directions. The convolution kernel usually uses a square, so they can learn the feature form every direction. In recent years, with the development of convolutional neural networks, new technologies such as dilated convolution and deformable convolutional have emerged [29]. However, the design of the convolution kernel is essentially to learn the correlation feature in the matrix. We already considered the spatial dimension in *multidimensional feature vector*. There is no need to compute each station correlation on it. Therefore, in this paper we designed a hetero convolution kernel form, only compute the column direction of the feature matrix. In this way, the connection between each feature vector is cut, and only the respective multidimensional feature needs to be learned. First, we designed a set of 1×2 kernels to calculate with matrix. This is designed for the weather and date type dimension in the multidimensional feature vector. In feature modeling, the label encoding method is used for the feature of these two dimensions, which is different from the daily traffic flow dimension. After that, we use ReLU function as the activation function. Second, we also execute a convolution operation with a set of 1×2 convolution kernels, and continue to learn the features of each vector. Finally, we use a set of 1×3 convolution kernels to end the convolution operation, which reduces the dimensionality of the multidimensional feature vector of each station in the matrix in to a traffic flow value. After that, we uniformly perform pooling operations on the channel, as shown in

Eq. 2. Generally, pooling operations have maximum pooling or average pooling, which play a role in input to reduce the dimension. Cross-channel pooling is a method for reducing the size of state and number of parameters needed to have a given number of filters in the model [30]. After three times convolution operations, the multidimensional feature vector of each station in feature matrix is reduced to a one-dimensional daily traffic flow value. But on the channel, the number of convolution kernels causes channel increment. So, we use cross-channel pooling to reduce channel dimension. Our pooling method only averages channels, which we call channel subsampling, to achieve the purpose of channel dimensional reduction. So that, we can retain the feature matrix structure in one channel. At the same time, we discarded the fully connection layer of CNN, and keep the current feature matrix structure. After learning, feature matrix has turn to one-dimensional feature from multidimensional vector, as $\mathrm{c}\left(M^{t}\right)$. So, the output $\mathrm{c}\left(M^{t}\right)$ is the predicted value of traffic flow for each station in highway network in day t. We use real daily traffic flow as training label to training the FCN we designed. After training, we get daily traffic flow matrix on each day.

After getting the daily traffic flow matrix of all toll stations in the network, in order to use temporal of data, we will process the matrix in concatenation layer into time series. We set the learning time range as 15, so after concatenation layer we can get a time series data as input. Then we choose LSTM network for feature learning in time dimension. LSTM is an improved model of recurrent neural network (RNN). The classic LSTM cell block contains four components, namely input gate, forget gate, output gate and cell state. These four components together make LSTM cell block have memory and function of learning time dimension. The corresponding calculation steps are as follows:

$$i_t = \sigma\left(\mathrm{W_x^i c(M^t) + W_h^i h_{t-1} + b_i}\right) \tag{3}$$

$$o_t = \sigma\left(\mathrm{W_x^o c(M^t) + W_h^o h_{t-1} + b_o}\right) \tag{4}$$

$$f_t = \sigma\left(\mathrm{W_x^f c(M^t) + W_h^f h_{t-1} + b_f}\right) \tag{5}$$

$$c_t = f_t \odot c_{t-1} + i_t \odot \tanh\left(W_h^c \mathrm{c}\left(M^t\right) + W_h^c h_{t-1} + b_c\right) \tag{6}$$

$$h_t = \mathrm{o}_t \odot \tanh(\mathrm{c}_t) \tag{7}$$

where $\mathrm{c}\left(M^{t}\right)$, c_t, and h_t are input, cell state, and cell output respectively at time t. W_* and b_* are weight and bias vectors connecting different gates. $\odot$ denotes an element-wise product. As shown in the above equation, the input gate can control how much new information enters the cell state, the forget gate can choose how much information is discarded, and the output gate can control how much information is passed into the next time step or output. Cell state records the state information of the cell with the change of time to complete the memory function of LSTM. By serializing the feature matrix of toll stations in whole network, LSTM can also learn the time dimension features of the data and improve the prediction accuracy. We put the time series data generated by the concatenation layer as input into LSTM. We use fifteen days feature matrix for training, and use the daily traffic flow of the next day as the label for learning. Finally, we get the predicted value as traffic flow of all toll stations in highway network. The combination of the above two deep learning model has fully studied the spatio-temporal and multidimensional feature of the data. At the same time, through the improvement of the convolution network, the method can predict the highway network traffic flow at one time.

4 Experiment

4.1 Settings

The experiment used real data from the toll stations of the highway network, which was derived from the Henan Highway Management System supported by the project. At the same time, it also collects the weather information of all toll stations. The data from May to September of Henan highway was used for model training and analysis. These large amounts of data are stored in a distributed cluster. The cluster has one master node and two slave nodes. The configuration is Intel (R) Xeon (R) CPU E5–4607 2.20 GHz, 32 GB RAM and 80 TB storage. Distributed computing frameworks such as Hadoop, Spark and HBASE are installed on it for data cleaning and feature modeling. The framework is built and trained on a computer with Intel (R) Core (R) CPU i7–7700 2.80 GHz, 32 GB RAM, 4 TB storage and two NVIDIA 1080 GPUs.

The open source deep learning framework TensorFlow 1.12.0 supports the construction of this method, and Scikit-learn 0.20.3 is used to build models of counterparts. We named our method as hybrid deep learning method (HDM) to compare with four machine learning models. They are: KNN is a classic supervised statistical machine learning model that classifies and regressions by analyzing the distance relationship of data; ARIMA is the most widely used prediction model for time series data; SVR is a machine learning model based on kernel method. it can map data to high dimensions through the kernel to make regression predictions; Gradient boost regression tree (GBRT) is an integrated machine learning method that learns by reducing the loss between different models. It has been widely used in recent years. Then we chose three predictive metrics to evaluate the method. First is root mean square error (RMSE) as Eq. 8; the second is mean absolute percentage error (MAPE) as Eq. 9 and R-square as Eq. 10 is the last metric.

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\hat{y}_i - y_i\right)^2} \tag{8}$$

$$\mathrm{MAPE} = \frac{100\%}{n}\sum_{i=1}^{n}\left|\frac{\hat{y}_i - y_i}{y_i}\right| \tag{9}$$

$$\mathrm{R-square} = 1 - \frac{\sum_{i=1}^{n}\left(\hat{y}_i - y_i\right)^2}{\sum_{i=1}^{n}\left(y_i - \bar{y}\right)^2} \tag{10}$$

where $\hat{y}_i$ represents the predicted value, y_i represents ground truth value and n is the number of test dataset.

4.2 Evaluation

We designed two experiments for method evaluation. The first is a comparative experiment to verify the effectiveness and accuracy of HDM by comparing with other methods. Secondly, we use HDM and predicted all toll stations traffic flow one time, and compare with traditional LSTM prediction method.

Experiment 1: Prediction Effects in a Station. The data from September 2017 was used as the test set. In the experiment, all 274 toll stations in Henan Province constituted the whole network. Among them, we randomly selected *zhengzhounan* toll station for analysis. The prediction results of HDM and the comparison model are shown in Fig. 3:

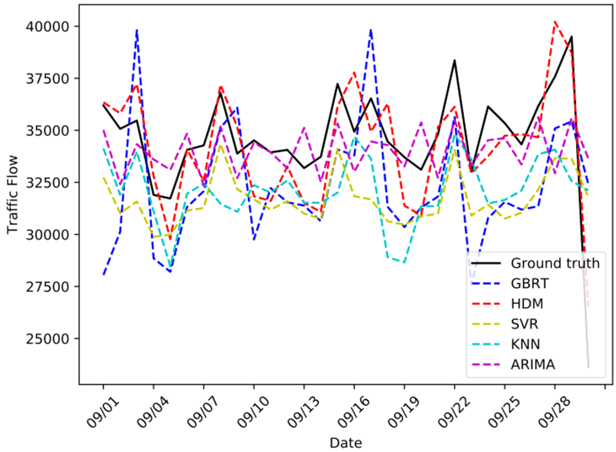

Fig. 3. The real and prediction traffic flow in *zhengzhounan*

Through Fig. 3, we can intuitively see that the HDM fits the real value best and has good performance in the time period of wave peak and wave trough. At the same time, the fluctuation of GBRT is worse than HDM. Although the overall trend is close to the true value, the prediction of the turning point is lacking. Compared with the SVR and ARIAM models, our method's prediction is more accurate. And like the KNN model, it can fit the real curve well. So, by analyzing the prediction results, we get Table 2:

Table 2. Prediction performances of different models

	HDM	GBRT	SVR	KNN	ARIMA
RMSE	1983.95	3933.77	3692.97	2583.35	3597.36
MAPE (%)	41.69	54.68	59.72	43.69	53.64
R-square	0.894	0.863	0.704	0.883	0.873

From Table 2, we can see that, comparing all models, the HDM has the best fitting effect on the traffic flow. The RMSE is kept within two thousand, indicating that the error between the predicted value and the true value is kept within an acceptable range. At the same time, the R-square value is closest to 1, better than the KNN model, which represents the best fitting effect on real traffic flow. Through this experiment, we can know that HDM has good prediction accuracy at the key station *zhengzhounan*, and better than other comparison models.

Experiment 2: Network Prediction. In this experiment, we use HDM to predict network traffic flow at once. And use normal LSTM model to predict traffic flow for comparison. We chose four key toll station for experiment, they are *puyangnan, zhengzhouxinqu, xinxiang and xuanyuanguli.* There daily traffic flow is different. We use LSTM four times. The single toll station prediction uses the same training set and uses LSTM neural network model in [16] for time series data learning. The experiment results are shown in Fig. 4:

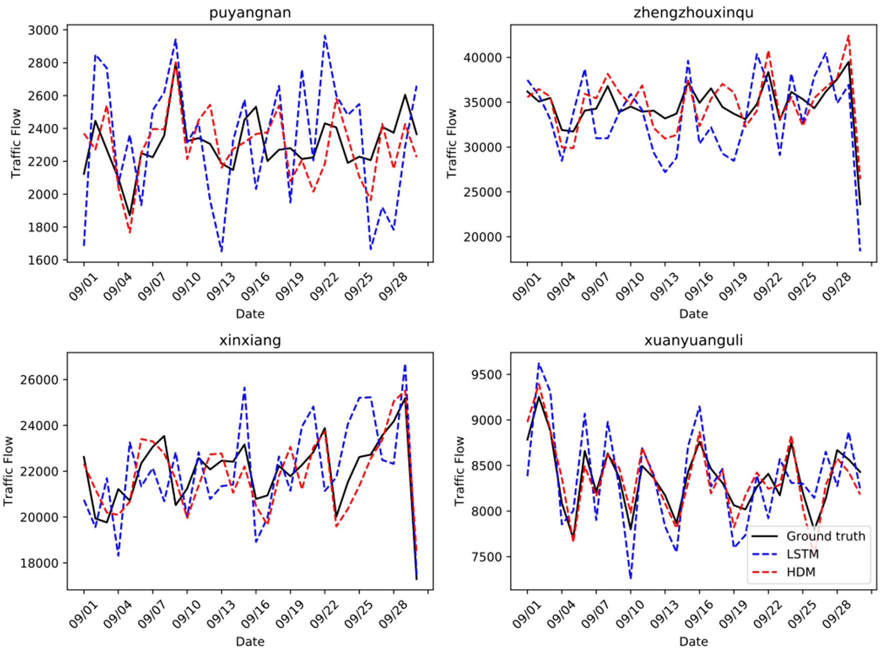

Fig. 4. The prediction traffic flow with HDM and LSTM

We can see from the Fig. 4 from the selected toll stations, the prediction performance of HDM is better than the single station traffic flow prediction. We calculate the average of the metrics of the four station to Table 3. According to the results in Table 3, the R-square is better than the LSTM model and is closer to the true value. It can be seen from the results that HDM method is not affected by the increase of the parameter of the feature matrix, but the learning ability is better than the single station prediction method. Therefore, the network traffic flow prediction method is fully verified, and the expected prediction effect can be achieved.

Table 3. Prediction performances

	HDM	LSTM
RMSE	1085.72	1489.56
MAPE (%)	32.71	38.96
R-square	0.843	0.794

5 Conclusion

In this paper, we proposed a new multidimensional feature-based network traffic flow prediction method. A hybrid deep learning model consists of a fully convolutional neural network and long-short term memory neural network is proposed. It can fully study the spatio-temporal of highway toll station data. We considered the spatial feature of data and some other feature can affect traffic flow like weather and date type. We use this for feature modeling. After learning this toll station feature matrix, the method can well predict the traffic flow of highway network. Experiments show that the method has a good performance compared with the shallow machine learning method. At the same time, it has a good effect on the learning of multidimensional feature. In the future research, we would further subdivide the dimensions of features, and consider more relationship between with traffic flow. We can also improve the performance of neural network in depth to predict more precise.

Acknowledgment. This work was supported by National Key R&D Program of China (No. 2018YFB1402500), National Natural Science Foundation of China (No. 61702014) and Beijing Municipal Natural Science Foundation (No. 4192020).

References

1. Avineri, E.: Soft computing applications in traffic and transport systems: a review. In: Hoffmann, F., Köppen, M., Klawonn, F., Roy, R. (eds.) Soft Computing: Methodologies and Applications, vol. 32, pp. 17–25. Springer, Heidelberg (2005). https://doi.org/10.1007/3-540-32400-3_2
2. Ding, W., Zhang, S., Zhao, Z.: A collaborative calculation on real-time stream in smart cities. Simul. Model. Pract. Theory **73**(4), 72–82 (2017)
3. Wang, X., Ding, W.: A short-term traffic prediction method on big data in highway domain. J. Comput. Appl. **39**(01), 93–98 (2019)
4. Zheng, Y., Capra, L., Wolfson, O., et al.: Urban computing: concepts, methodologies, and applications. ACM Trans. Intell. Syst. Technol. (TIST) **5**(3), 1–55 (2014)
5. Smith, B.L., Demetsky, M.J.: Traffic flow forecasting: comparison of modeling approaches. J. Transp. Eng. **123**(4), 261–266 (1997)
6. Kumar, S.V., Vanajakshi, L.: Short-term traffic flow prediction using seasonal ARIMA model with limited input data. Eur. Transp. Res. Rev. **7**(3), 21 (2015)
7. Du, S., Li, T., Yang, Y., Horng, S.: Deep air quality forecasting using hybrid deep learning framework. IEEE Trans. Knowl. Data Eng. (2019). https://doi.org/10.1109/TKDE.2019.2954510

8. Niu, Z., Sun, Q.: Study of railway passenger volume forecast based on grey forecasting model. In: 2016 International Conference on Logistics, Informatics and Service Sciences (LISS), pp. 1–4. IEEE (2016)
9. Rong, Y., Zhang, X., Feng, X., et al.: Comparative analysis for traffic flow forecasting models with real-life data in Beijing. Adv. Mech. Eng. **7**(12), 1687814015620324 (2015)
10. Cheng, A., Jiang, X., Li, Y., et al.: Multiple sources and multiple measures based traffic flow prediction using the chaos theory and support vector regression method. Phys. A: Stat. Mech. Appl. **466**, 422–434 (2017)
11. Smith, B.L., Williams, B.M., Oswald, R.K.: Comparison of parametric and nonparametric models for traffic flow forecasting. Transp. Res. Part C: Emerg. Technol. **10**(4), 303–321 (2002)
12. Sun, S., Zhang, C., Yu, G.: A Bayesian network approach to traffic flow forecasting. IEEE Trans. Intell. Transp. Syst. **7**(1), 124–132 (2006)
13. Li, L., He, S., Zhang, J., et al.: Short-term highway traffic flow prediction based on a hybrid strategy considering temporal–spatial information. J. Adv. Transp. **50**(8), 2029–2040 (2016)
14. Ghosh, B., Basu, B., O'Mahony, M.: multivariate short-term traffic flow forecasting using time-series analysis. IEEE Trans. Intell. Transp. Syst. **10**(2), 246–254 (2009)
15. Hong, W.C., Dong, Y., Zheng, F., et al.: Forecasting urban traffic flow by SVR with continuous ACO. Appl. Math. Model. **35**(3), 1282–1291 (2011)
16. Luo, X., Li, D., Yang, Y., Zhang, S.: Spatiotemporal traffic flow prediction with KNN and LSTM. J. Adv. Transp. **2019**, article ID 4145353, 10 p. (2019). https://doi.org/10.1155/2019/4145353
17. Cai, P., Wang, Y., Lu, G., Chen, P., Ding, C., Sun, J.: A spatiotemporal correlative k-nearest neighbor model for short term traffic multistep forecasting. Transp. Res. Part C: Emerg. Technol. **62**, 21–34 (2016)
18. Dell'acqua, P., Bellotti, F., Berta, R., De Gloria, A.: Time aware multivariate nearest neighbor regression methods for traffic flow prediction. IEEE Trans. Intell. Transp. Syst. **16**(6), 3393–3402 (2015)
19. Sun, B., Cheng, W., Goswami, P., Bai, G.: Short-term traffic forecasting using self-adjusting k-nearest neighbors. IET Intel. Transport Syst. **12**(1), 41–48 (2018)
20. Castro-Neto, M., Jeong, Y.-S., Jeong, M.-K., Han, L.D.: Online-SVR for short-term traffic flow prediction under typical and atypical traffic conditions. Expert Syst. Appl. **36**(3), 6164–6173 (2009)
21. Sun, Y., Leng, B., Guan, W.: A novel wavelet-SVM short time passenger flow prediction in Beijing subway system. Neurocomputing **166**, 109–121 (2015)
22. Ding, W., Xia, Y., Wang, Z., et al.: An ensemble-learning method for potential traffic hotspots detection on heterogeneous spatio-temporal data in highway domain. J. Cloud Comput. **9**(1), 1–11 (2020)
23. Moretti, F., Pizzuti, S., Panzieri, S., et al.: Urban traffic flow forecasting through statistical and neural network bagging ensemble hybrid modeling. Neurocomputing **167**, 3–7 (2015)
24. Priambodo, B.: Traffic flow prediction model based on neighboring roads using neural network and multiple regression. J. Inf. Commun. Technol. **17**(4), 513–535 (2020)
25. Jia, Y., Wu, J., Benakiva, M., et al.: Rainfall-integrated traffic speed prediction using deep learning method. IET Intel. Transp. Syst. **11**(9), 531–536 (2017)
26. Zhang, Y., Huang, G.: Traffic flow prediction model based on deep belief network and genetic algorithm. IET Intell. Transp. Syst. **12**(6), 533–541 (2018)
27. Ding, W., Wang, X., Zhao, Z.: CO-STAR: a collaborative prediction service or short-term trends on continuous spatio-temporal data. Future Gener. Comput. Syst. **102**, 481–493 (2020)
28. McLaughlin, N., del Rincon, J.M., Miller, P.: Recurrent convolutional network for video-based person re-identification. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 1325–1334 (2016)

29. Yu, F., Koltun, V., Funkhouser, T.: Dilated residual networks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 472–480 (2017)
30. Goodfellow, I., Warde-Farley, D., Mirza, M., et al.: Maxout networks. In: International Conference on Machine Learning, pp. 1319–1327 (2013)