

# Traffic Flow Prediction Model Based on Spatio-Temporal Dilated Graph Convolution

**Xiufang Sun, Jianbo Li\*, Zhiqiang Lv and Chuanhao Dong**

College of Computer Science and Technology, Qingdao University  
Qingdao, CN 266071- China

[e-mail:sunflowerxf@163.com; lijianbo@188.com; lvzq7614@163.com; 1551231201@qq.com]

\*Corresponding author: Jianbo Li

*Received March 20, 2020; revised June 18, 2020; accepted June 25, 2020;  
published September 30, 2020*

---

## **Abstract**

With the increase of motor vehicles and tourism demand, some traffic problems gradually appear, such as traffic congestion, safety accidents and insufficient allocation of traffic resources. Facing these challenges, a model of Spatio-Temporal Dilated Convolutional Network (STDGCN) is proposed for assistance of extracting highly nonlinear and complex characteristics to accurately predict the future traffic flow. In particular, we model the traffic as undirected graphs, on which graph convolutions are built to extract spatial feature informations. Furthermore, a dilated convolution is deployed into graph convolution for capturing multi-scale contextual messages. The proposed STDGCN integrates the dilated convolution into the graph convolution, which realizes the extraction of the spatial and temporal characteristics of traffic flow data, as well as features of road occupancy. To observe the performance of the proposed model, we compare with it with four rivals. We also employ four indicators for evaluation. The experimental results show STDGCN's effectiveness. The prediction accuracy is improved by 17% in comparison with the traditional prediction methods on various real-world traffic datasets.

---

**Keywords:** Traffic Flow Prediction, Deep Learning, Graph Convolution, Dilated Convolution

---

This research was supported in part by National Key Research and Development Plan Key Special Projects under Grant No. 2018YFB2100303, Key Research and Development Plan Project of Shandong Province under Grant No. 2016GGX101032, Program for Innovative Postdoctoral Talents in Shandong Province under Grant No. 40618030001, National Natural Science Foundation of China under Grant No. 61802216, and Postdoctoral Science Foundation of China under Grant No.2018M642613.

## 1. Introduction

With the rapid increase of car ownership, the motor vehicle ownership in China reached 340 million until June 2019, of which 250 million were motor vehicles. Besides, there were 420 million motor vehicle drivers, of which 380 million were motorists. Along with the emergence and rapid development of the sharing economy, a series of shared transportation modes such as shared bicycles and Internet-linked taxis have sprung up like mushrooms. In 2018, about 20 billion passengers took online taxis in China, accounting for 36.3% of the total number of taxis. In other words, at least one in every three taxi riders uses an online taxi. Increasing demand for transportation and these convenient shared transportations have also increased motor vehicle travel. The number of urban vehicle ownership and transportation needs are increasing, but urban transportation infrastructure resources remain stable. Improper traffic signal preemption in an emergency will also have an impact on traffic [1]. Therefore, urban traffic loads are becoming increasingly serious. Then there are problems such as traffic congestion, safety accidents and insufficient allocation of traffic resources [2].

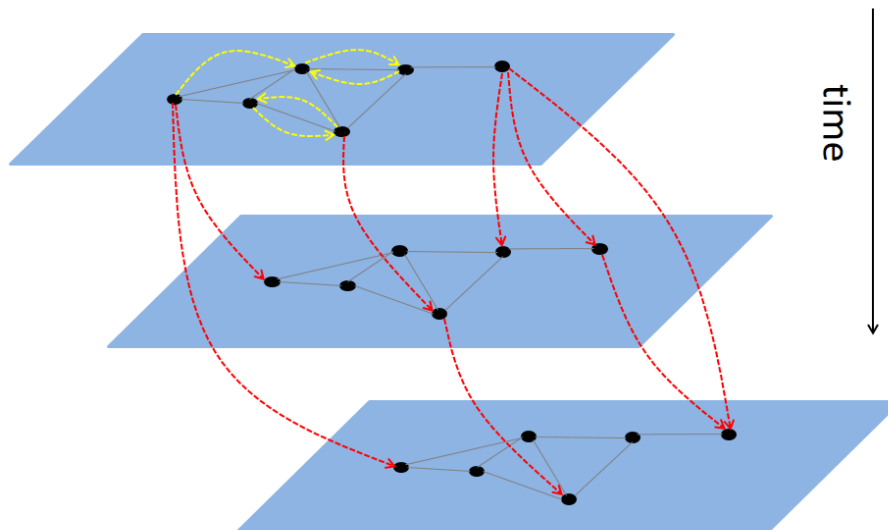
While solving the problem of urban traffic congestion, the usual approach is to increase or widen the scale of urban roads, and build basic transportation facilities such as subways or viaducts in transportation hubs, or to increase traffic signs to guide the road conditions [3]. But in reality, for any city, the traffic network cannot be changed arbitrarily and restrictively, and the traffic signs are greatly affected by the complex background and shooting angle [4]. So, the traditional solution shows its insurmountable limitations. At the same time, the growth rate of motor vehicles is far greater than the speed of the construction of transportation facilities. The main contradiction of urban traffic congestion is manifested between the rapid growth of motor vehicle ownership and the caution and slowness necessary for urban renewal and construction, and it is getting worse. Therefore, with the in-depth study of traffic problems, the design of traffic countermeasures must gradually shift from the hard countermeasures which focus on facility supply to the soft and hard coordination method which combine facility supply and demand management.

Nowadays, information collection technology is constantly upgrading, which brings great convenience to information collection. For example, the Internet of things has been successfully applied in environmental detection [5], and it has become an important part of infrastructure [6]. At present, with the development of information technology, most cities in China have begun to build an integrated platform for traffic information, including traffic information collection equipment, information transmission equipment and traffic information release platform. Internet of vehicles can also be used to exchange road information between vehicles [7]. The management, processing and releasing of these dynamic data need a comprehensive platform to control the real-time update. Intelligent traffic management system is the center of traffic data management. It can provide drivers with traffic information, recommend reference strategies for the selection of traffic routes, and offer information such as driving time and driving services. The current equipment used in the city also provides such information, but only the historical traffic information is released. As the transportation system is a dynamic and complex system, it is easy to be interfered by the external conditions. The emergencies or other factors can cause great differences to the traffic data. Therefore, the current system equipment does not achieve the dynamic and real-time information release. Traffic managers and the drivers should master various trend of traffic state in the road network, find out the essential law of traffic flow from a large number of data, and make

decisions flexibly to improve the efficiency of road traffic. Therefore, traffic prediction plays an important role for traffic management and controls departments to take traffic guidance measures. The research on traffic prediction model [8] has the exploding popularity in recent years.

Traffic prediction in the prediction model refers to the prediction of future traffic changes based on historical traffic data. Traffic prediction refers to the real-time prediction of the traffic flow at the next decision moment  $t + \Delta t$  and even several times at time  $t$  [9].

Changes in transportation are affected by many factors, such as time, surrounding environment, and complex factors such as weather. The influence of traffic on time is mainly reflected in historical time. The traffic conditions in historical time will affect the traffic conditions in subsequent times; the traffic conditions in the spatial dimension will be affected by surrounding conditions. As shown in Fig. 1. The yellow line represents the influence in space, the red line represents the influence in time. Traffic is affected by time and space. In previous studies, it often focused on one aspect of research, and ignored the impact of the other. For traffic flow prediction, if only considering the impact of time or space, it is not able to get high accuracy prediction results. The existing prediction methods consider the influence of time and space on the traffic flow to predict. Although those methods improve the accuracy of prediction to a certain extent, but it is still incomplete. For example, for the same number of vehicles, the four lanes road will not be congested, but it is likely to cause congestion on the single lane road, resulting in slow vehicle speed, thus affecting the prediction of traffic flow and reducing the accuracy of prediction. Therefore, lane occupancy rate is also an important factor affecting traffic flow prediction.



**Fig. 1.** Spatiotemporal map of traffic flow

The main contribution of this paper as follows:

- (1) We introduce strategies to model the temporal and spatial dependence of traffic flow and the impact of surrounding road occupancy.
- (2) The generalized graph is used to model the traffic network to avoid discrete traffic flows and disrupting space dependence.
- (3) We use dilated convolution to extract features, fuse temporal and spatial features, and then fuse traffic occupancy.

- (4) To the best of our knowledge, for the first time, we add the influence of road occupancy into the prediction of traffic flow.

Using these features and adopting the convolutional neural network structure to predict traffic flow, the prediction accuracy is improved, the training speed is improved, and the required parameters are reduced.

The organizational structure of this paper is as follows. Section 2 briefly introduces the related work. Section 3 describes the definition of the problem. In Section 4, the system model is presented and the mathematical description is given. In Section 5, the model is simulated and the experimental results are analyzed. Section 6 summarizes the work of this paper.

## 2. Related Work

### 2.1 Classification of Traffic Prediction

Traffic prediction can be divided into two categories: the prediction of vehicle index and the prediction of vehicle derived behavior. The traffic research based on vehicle index usually focuses on three basic variables such as flow, speed, and density. These three variables are used as standards for measuring current traffic conditions and predicting future traffic conditions. The derivative behavior prediction of vehicles is generally represented by trajectory prediction [10]. According to the length of forecast time, traffic forecast can be divided into three categories: short-term forecast, medium-term forecast and long-term forecast. The short-term forecast refers to the situation with short time series interval and forecast period, such as 5-30 minutes. The medium-term forecast and long-term forecast refer to the situation with long time series interval and forecast period, such as one hour, half day, one day or even longer.

### 2.2 Prediction Methods Based on Deep Learning

The current mainstream methods are: classic statistical and deep learning models [11,12,13]. The linear theoretical model based on statistics contains historical average(HA) [14], time series method and Kalman filter [15,16]. In the analysis of time series, the autoregressive integrated moving average model (ARIMA) and its variants are built based on traditional statistical methods [17,18]. These are a type of linear model, which have the advantages of simple structure and fast calculation. However, the traffic data owns the nonlinear characteristics with strong randomness and uncertainty. The linear model is subject to the stable distribution of the time series, and therefore, it does not consider the influence of space-time factors on the traffic prediction [19]. The prediction accuracy is low, and the ability to resist interference is poor. These methods lack the ability to represent highly non-linear traffic flows. In order to meet the non-linear characteristics of the data, people also proposed non-linear prediction models such as wavelet-based theoretical models, chaotic theoretical models, and non-parametric regression models. The non-parametric regression model has high prediction accuracy and good error distribution. It is applicable to the short-term traffic forecast with emergencies, but this method is still inadequate and needs to meet the complexity of "proximity" matching and neighbor search with a huge amount. Recently, traditional statistical methods have been impacted by deep learning methods in traffic prediction. These models have the ability to obtain higher accuracy and model more complex data, such as K nearest neighbors (KNN), support vector machines (SVM) [20,21] and neural networks.

Because traffic flow has complex features such as non-linearity and randomness, traditional traffic prediction cannot extract more accurate specific features from complex feature expressions, and it cannot take full advantage of multi-attribute features in traffic data. The method is deficient in capturing higher-dimensional features and performing fusion prediction, while deep learning makes up for the deficiencies in traditional methods. Hinton et al. [22] proposed a fast learning algorithm based on Deep Belief Networks (DBN) [23,24]. This algorithm uses unsupervised greedy pre-training methods to obtain the weight parameters of the model. Through multi-layer representation learning, we can obtain a representation that can better cover the data features. And the layer-by-layer training method reduces the difficulty of deep neural network training and promotes the application of deep learning in various aspects. However, these methods are difficult to extract temporal and spatial features in a fully connected manner, and due to strict restrictions on spatial attributes, the representation capabilities of these methods are severely restricted and cannot be fully expressed.

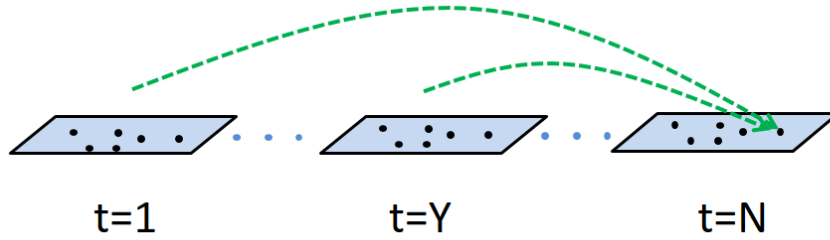
In order to show that traffic is affected by spatial characteristics, Shi et al. [25,26] proposed a convolutional LSTM, which is an extension of a fully connected LSTM with embedded convolutional layers. Although the features of time and space are extracted, conventional convolution is used, which can only be applied to conventional network structures, but not road networks with graph structure characteristics. In addition, the model based on recurrent network has a large amount of calculation, which makes it easy to increase the error and difficulty to train in the calculation process. Yu et al. [27] proposed a model STGCN that combines temporal and spatial features. The paper used graph convolution to capture the temporal and spatial characteristics of traffic flow but did not consider the impact of other factors on traffic flow.

Traffic flows affect each other and do not exist independently. Traffic flows are also affected by surrounding road conditions. Therefore, we can model the traffic flow, make the individual traffic flows connected to each other on the generalized graph, and retain their interconnectedness instead of discretely, add the impact of road occupancy on the traffic flow, and predict the traffic flow.

### 3. Problem Definition

#### 3.1 Definition of Traffic Flow Forecast

Traffic flow forecasting is a type of time series forecasting, which uses traffic flow as a predictive indicator. We predict the future traffic depending on the historical traffic flow, that is, the number of vehicles passing through the history is used to predict the number of vehicles in the future. Traffic flow prediction is a time series prediction problem. The prediction process is to give a specified number of nodes. In detail, the prediction process is conducted based on the possible traffic flow of the given first  $Y$  observation samples after the next  $N$  time stamps, each node includes information such as the number of vehicles, time and space information that affects traffic changes, as shown in Fig. 2. We select sample information from the past hour of 64 monitoring stations to predict the traffic flow in the next 45 minutes.



**Fig. 2.** Traffic flow forecast

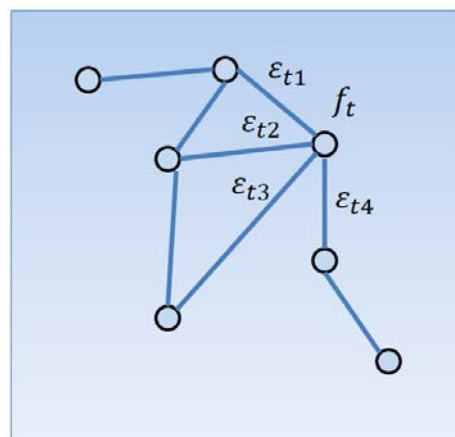
We describe the traffic flow prediction in a mathematical form as:

$$\hat{f}_{t+1, \dots, t+N} = \arg \max_{t+1, \dots, t+N} \log P(f_{t+1, \dots, t+N} | f_{t-Y+1, \dots, t}) \quad (1)$$

where  $f_t \in \mathbb{R}^n$  is a traffic vector with  $n$  monitoring stations that we selected at time stamp  $t$ . Each vector records the observed flow over a distance.

### 3.2 Introduction of Road Network

As shown in **Fig. 3**, recording the traffic flow information, each node  $f_t$  depends on each other, these nodes are connected in pairs. The vectors are connected to each other to form a network structure called road network. We therefore define  $f_t$  in an undirected graph, undirected graphs can be defined as  $G = (f_t, \varepsilon, M)$ . Node  $f_t$  can be defined as a signal with weight in graph  $G$ . In the graph,  $f_t$  represents a finite set of vertices, the number of vertices is the number of monitoring stations we choose,  $\varepsilon$  represents the set of connected edges between vertices, and  $M$  represents an adjacency matrix.



**Fig. 3.** Graph structure of traffic flow

### 3.3 Presentation of Graph Convolution

Kipf et al. [28] and others first proposed graph convolution, and applied the convolutional neural network commonly used for images in deep learning to graph data. Graph convolution

is mostly used in computer vision, and the processing of graph data is more complete. From the previous introduction to the road network, we know that the monitoring stations are connected to each other, so the data of a single monitoring station cannot represent all the information of this monitoring station. Connections are likely to cause feature deviations, so considering the information of neighboring nodes will give more complete information than considering the single feature of a single node, so graph convolution is a good choice. We model the traffic vectors of containing traffic flow as undirected graphs, on which convolutions are employed to extract features. However, convolutions on traditional grids cannot be applied to generalized graphs. We therefore utilize graph convolution to extract spatiotemporal features. Graph convolution is a method of using convolution on a graph. We first use the Fourier transform on the graph and then use the convolution theorem, so that the product of two Fourier transforms can be used to represent the convolution operation. The Fourier transform is operated in the spectral domain by introducing the spectral framework into the model and using the convolution in the spectral domain. We express the graph through the Laplace matrix in the following form:

$$L = I_n - D^{-1/2}MD^{-1/2} \in R^{n \times n} \quad (2)$$

In the above formula,  $I_n$  is the identity matrix,  $M$  is the adjacency matrix, and  $D \in R^{n \times n}$  is the diagonal degree matrix. Eigendecomposing the Laplacian matrix  $L$ , we further obtains:

$$L = I_n - D^{-1/2}MD^{-1/2} = U\Lambda U^T \in R^{n \times n} \quad (3)$$

where  $\Lambda \in R^{n \times n}$  is a matrix of eigenvalues of  $L$ .  $U \in R^{n \times n}$  is the fourier basis of the graph. Graph convolution can be expressed as:

$$K * gx = K(L)x = K(U\Lambda U^T)x = UK(\Lambda)U^T x \quad (4)$$

where  $x$  denotes the traffic flow on the graph, which is transformed based on the properties of the Laplacian matrix. By the definition of equation (4), a graph signal  $x$  is processed by a convolution kernel  $K$  by performing a Fourier transform on the graph.

### 3.4 Dilated Convolution

Yu F. [29] and others initially proposed the application of Dilated Convolution in semantic segmentation. Dilated Convolution is an operation of adding holes in the convolution kernel, which results in expanding the receptive field. Accordingly, the required information can be extracted in a new way, and the calculation complexity is reduced. Dilated convolution is able to capture multi-scale contextual messages. We therefore use dilated convolution in the information extraction part, which more effectively captures the time and space information of traffic flow. Then, the thinking of dilated convolution is applied to fuse the information we have captured.

## 4. Spatio-Temporal Dilated Graph Convolution Model

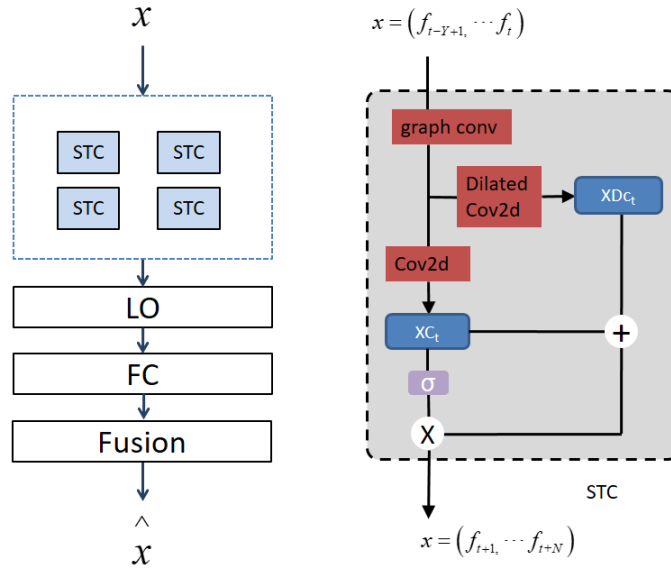
### 4.1 Model Architecture

**Fig. 4** shows the basic architecture of the spatiotemporal dilated graph convolution (STDGCN). The main part of the model consists of the spatio-temporal convolution block STC and subsequent fusion components. The STC consists of a convolution kernel in the time dimension and a convolution in the space dimension. STC can extract the characteristics of time and space, and predict the traffic flow in the future by merging lane occupancy. Capable of capturing the impact of time and space on traffic flow and integrating the impact of lane



occupancy. Details are shown on the left side of **Fig. 4**. STC is composed of a graph convolution block that extracts spatial features and a time convolution block that extracts temporal features. The right side of **Fig. 4** shows the composition of the time convolution block in detail. It is composed of 2D conventional convolution and 2D dilated convolution, which can effectively capture the impact of time series on traffic flow.

After the spatio-temporal feature extraction, the influence of the lane occupancy is added, and after the full connection, the output results are fused through the parameter matrix to finally output the prediction results. Assume that the monitoring data of  $m$  monitoring stations are selected, the time sampling frequency is  $s$  times a day, and there are  $m*s$  nodes each day, and the traffic flow at the future time  $T_f$  is predicted based on the node information of the historical time period  $T_h$  set by the experiment. The model uses the features captured by each component and the influence of lane occupancy to obtain the final prediction result.



**Fig. 4.** model structure

## 4.2 Spatial Characteristics

In order to effectively extract the spatial features, we use graph convolution to calculate the spatial relationship of traffic flow, as presented in Section 3.2. In section 3.2 we introduced the use of graph convolution in detail. Graph convolution can effectively extract the features of the data mapped to the graph. But the time complexity of the kernel calculation reaches  $O(n^2)$  due to the Fourier basis multiplication in formula (4). While dealing with a large amount of traffic data, we generate large graphs showing the connections between the data. Such situation leads to an excessively high calculation complexity. In order to solve this problem, we take an approximate strategy to simplify it.

We first consider Chebyshev polynomial approximation as:

$$K * gx = K(L)x \approx \sum_{h=0}^{H-1} K_h T_h \left( \frac{\tilde{L}}{L} \right) x \quad (5)$$



To reduce the parameters, the convolution kernel  $K$  is restricted by a polynomial, where  $K_h \in R^{n \times n}$  is the Chebyshev polynomial coefficient. The parameter  $\theta \in R^K$  is a vector of polynomial coefficients and  $H$  is the size of the graph convolution kernel. The Chebyshev polynomial approximation is used to approximate the kernel as  $H-1$  expansion, that is

$$K(\Lambda) \approx \sum_{h=0}^{H-1} K_h T_h(\tilde{\Lambda}) \quad . \quad \tilde{\Lambda} = 2\Lambda / \lambda_{\max} - I_n \quad , \quad \lambda_{\max} \text{ is the largest eigenvalue of } L \quad .$$

$T_h(\tilde{L}) \in R^{n \times n}$  is the scaled Laplacian matrix. Recursive approximation using Chebyshev polynomials can be normalized. Suppose  $\lambda_{\max} = 2$ . Equation (5) can be abbreviated as:

$$K * gx = K_0 x + K_1 \left( \frac{2}{\lambda_{\max}} L - I_n \right) x = K_0 x - K_1 \left( D^{-\frac{1}{2}} M D^{\frac{1}{2}} \right) x \quad (6)$$

where  $K_0$  and  $K_1$  are two shared parameters, in order to simplify the use of the same parameter instead of the representation,  $K = K_0 = -K_1$ ,  $\tilde{M} = M + I_n$ ,  $\tilde{D}_{ii} = \sum_i \tilde{M}_{ij}$ . After renormalization, the graph convolution can be expressed as:

$$K * gx = K \left( I_n + D^{-\frac{1}{2}} M D^{\frac{1}{2}} \right) x = K \left( \tilde{D}^{-\frac{1}{2}} \tilde{M} \tilde{D}^{\frac{1}{2}} \right) x \quad (7)$$

### 4.3 Convolution in Time Dimension

After extracting the spatial features through graph convolution, the temporal features are captured by temporal convolution blocks. The temporal convolution block is composed of a combination of dilated convolution and two-dimensional convolution to capture temporal features, as shown in the right side of Fig. 4. This component contains two parts of convolution. The corresponding items of the matrix processed after 2D conventional convolution and 2D dilated convolution are multiplied. We introduce the sigmoid gate  $\sigma$  to control the processing of the time series of the current state, mining time features by stacking time layers. Model the temporal characteristics of traffic flow and extract temporal characteristics. The use of dilated convolution can increase the field of view and better extract the time characteristics that affect the traffic flow.

### 4.4 Lane Occupancy

As we all know, in order to ensure the safety of vehicles, a certain distance is maintained between the vehicles, including the distance traveled side by side in the lane and the distance maintained by the front and rear vehicles. Therefore, the lanes can only be fully occupied for a short period of time, and the occupancy rate of the lanes is not very high most of the time. Especially on highways with fast speeds, it is almost impossible to maintain a high occupancy rate for a long time. However, the relationship between lane occupancy and traffic flow is very complicated. Although the high occupancy rate can reflect to a certain extent that the current traffic flow through the lane is relatively high, it also has a negative impact on the traffic volume. We collated the data and found that in the next stage when the lane occupancy rate becomes higher, the traffic flow does not increase linearly, but shows a certain downward trend, as shown in Fig. 5. The blue line represents the change in traffic flow, and the orange

curve represents the occupancy rate of the lane. The lane occupancy rate here is the average value between the lanes, and the lane occupancy rate of the four lanes is the average value obtained by the four lanes.

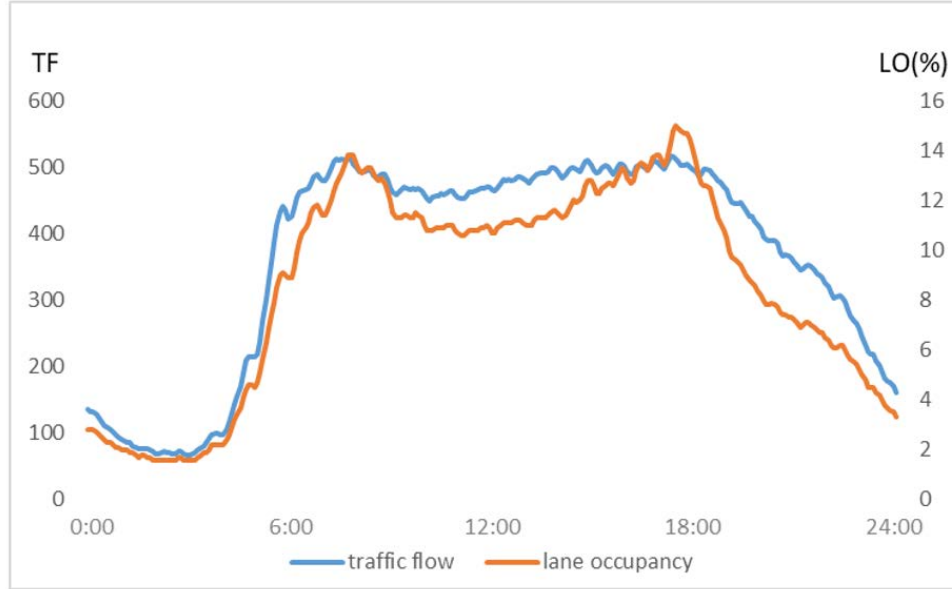


Fig. 5. Traffic Flow and Lane Occupancy

In order to add the influence of the lane occupancy in the prediction process, we increase the lane occupancy plate LO, and in this section we process the matrix after extracting the time features. Lane occupancy and traffic flow affect each other, and their influence is non-linear. It is clear that a lower lane occupancy rate will not affect the traffic flow in the future, so we have set a threshold of 10% based on the lane occupancy rate. In our model, the lane occupancy rate exceeds 10%, and the lane occupancy rate will affect the lane occupancy rate in the future.  $O^b$  is the parameter we set for this.

#### 4.5 Feature Fusion

After graph convolution, the node's information contains the information of neighboring nodes. After dilated convolution, the node's information is covered by the information including neighboring time slices. Therefore, after STC, the temporal and spatial characteristics of node data are extracted. After multiple layers of extraction, more long-term information in space and time can be obtained. And on this basis, the lane occupancy rate is added to consider the influence of the entrance lane occupancy rate on the node traffic flow.

Modeling input  $f^b$  and  $f^{b+1}$  outputs, we obtain the following formula,

$$f^{b+1} = \text{ReLU}(T_1^b * \text{ReLU}(K^b(T_0^b * f^b))) * O^b \quad (8)$$

where  $T_0^b$  and  $T_1^b$  are the upper and lower time layers,  $K^b$  is the spectral convolution kernel, and  $O^b$  is the lane occupancy rate.  $\text{ReLU}()$  represents the activation function.

The loss function of STDGCN model is defined as:

$$L(\hat{f}, P) = \sum_t \left\| \hat{f}(f_{t-Y+1}, \dots, f_t, P) - f_{t+1} \right\|^2 \quad (9)$$

where  $P$  is the set of all parameters in the model,  $f_{t+1}$  is the fact data, and  $\hat{f}$  is the prediction result of the model.

## 5. Experiments

### 5.1 Data Set and Experimental Configurations

The data used is the 50 days of data in 64 roads of Caltrans performance measurement system (PeMS) zone 7 (hereinafter referred to as PeMS7) excluding holidays, 25 days as training set, 10 days as evaluation set, and 15 days as test set. These monitoring stations are located in major urban parts of the California highway system. The data of each monitoring station in these data sets includes traffic, speed and lane occupancy with time stamp, which are collected with 30s as the sampling interval, and contain the time and location information of the detector. Part of the data was selected for aggregation processing, and the training set and test set were divided.

The interval of the data set is set to 5mins. Therefore, each vertex in the road network has 288 data points per day. After removing nodes that are too large or too small, the data is cleaned up using linear interpolation to fill in missing values. And zero-mean normalization is performed so that the processed data set can have an average value of zero.

From the introduction of the road network, we know that each node is connected to each other, and the adjacency matrix of the road network is represented by the distance between the monitoring stations we selected. The adjacency matrix  $M$  in Section 3.2 is calculated by the following formula:

$$M_{ij} = \begin{cases} \exp\left(-\frac{d_{ij}^2}{\sigma^2}\right), & i \neq j \text{ and } \exp\left(-\frac{d_{ij}^2}{\sigma^2}\right) \geq \delta \\ 0, & \text{otherwise} \end{cases} \quad (10)$$

$d_{ij}$  represents the distance between nodes, by calculating, we get  $M_{ij}$ ,  $M_{ij}$  represents the weight of the edges between nodes,  $\sigma^2$  and  $\delta$  are the threshold to adjust the distribution and sparsity of matrix  $M_{ij}$ .

All the experiments in this paper are trained and tested on a Linux cluster. The hardware configuration information are as follows, CPU: Intel(R) Xeon(R) CPU E5-2620 v4 @ 2.10GHz, GPU : NVIDIA Corporation GV100GL [Tesla V100 DGXS 32GB] (rev a1).

### 5.2 Experimental Results

We tested the STDGCN model on the dataset PeMS7. The traffic data for the next 9 time periods (45 mins) is predicted based on the flow from the previous 12 time periods (60 mins). We compare the proposed model with its four rivals, including the historical average (HA) [7], the autoregressive integral moving average method (ARIMA) [10-11], long short-term memory network (LSTM) [18-19] and Spatio-Temporal Graph Convolutional Networks (STGCN) [20]:

HA: historical average method. The prediction of future time is obtained by averaging the traffic flow of historical timestamps. In order to maintain consistency, we select 12 historical timestamps to predict the traffic of the next timestamp;

ARIMA: Autoregressive Integral Moving Average Method, this method is a classic method based on time series prediction in traffic research;

LSTM: Long Short-Term Memory Network, which is a more commonly used RNN model;

STGCN: Spatio-Temporal Graph Convolutional Networks defines matrices on a spatio-temporal graph convolution model on the graph, enabling parameter sharing of convolution kernels.

To estimate the performance of different methods, we employ the mean absolute error (MAE), root mean square error (RMSE), and mean absolute percentage error (MAPE) as evaluation indicators. Generally speaking, the smaller the values of these three indicators are, the smaller the error is, and the higher the prediction accuracy is. The calculation formula is as follows:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \hat{x}_i)^2} \quad (11)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |x_i - \hat{x}_i| \quad (12)$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{x_i - \hat{x}_i}{x_i} \right| \quad (13)$$

The experimental results of this experiment are predictions made with a 5-minute cycle, which predicts the traffic changes in the next 45 minutes, which is a short-term forecast and a medium-term forecast. During the experiment, the parameters we used are as follows: set the initial learning rate to 1e-3, for every 5 training groups, the learning rate will be 70% of the current learning rate. To simplify the calculation, we set  $O^b$  to 0.93.

**Table 1.** Performance comparison of different approaches on the dataset PeMS7.

Model	RMSE	MAE	MAPE
HA	40.35	27.02	12.5
ARIMA	36.80	20.53	11.82
LSTM	34.46	20.69	10.94
STGCN	34.01	23.87	5.17
STDGCN (ours)	27.60	17.15	7.97

Due to the high demand of traffic flow on time real-time, the closer time is to the current time node, the greater the reference value. We put the three experimental indexes of the first prediction period (the next five minutes) in Table 1. As can be seen from **Table 1**, with the improvement of research methods, the error rate shows a downward trend, compared with other models, our model achieves the best in the two evaluation indexes of RMSE and MAE, but lacks in MAPE. The traditional method cannot consider multiple factors, although we did not achieve the best in every index, we did not lag far behind in the value of MAPE. It is undeniable that our accuracy rate is higher than the comparison method in **Table 1**. Compared with other research methods, our method adds the influence of lane occupancy rate on traffic flow prediction, and our two indicators of error rate are lower than other methods, so our model considering multiple factors can be superior to the traditional model.

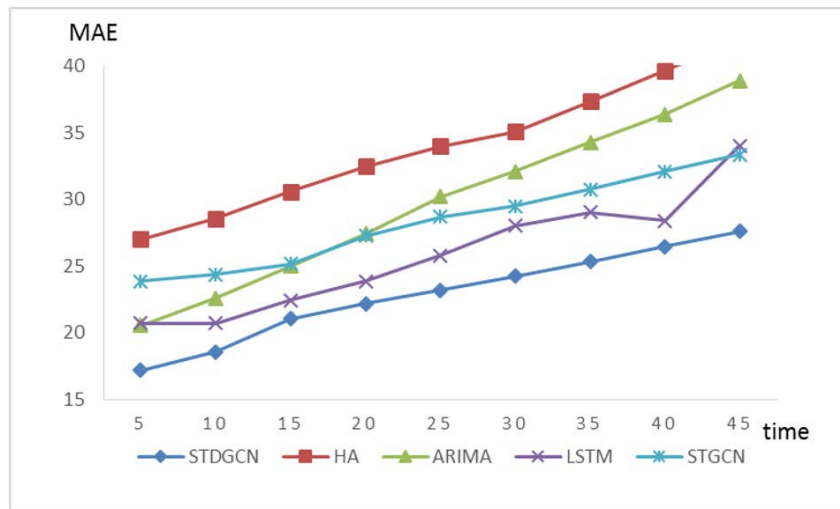


Fig. 6. MAE Comparison

From the experimental results in Fig. 6, the errors generated by all methods have an increasing trend with the extension of the prediction time. In terms of MAE, the historical average method performs worst, the performance of HA method is the worst, the error is the highest of several methods, so the accuracy is the lowest; although ARIMA has a low error in a short time, even in the first 20 minutes, it is lower than STGCN's error, but its error growth rate is very fast, with the increase of prediction time, the error increases gradually; LSTM performed well in the early stage, but its error increased rapidly after 40 minutes, and it was suitable for short-term prediction, but in the medium-term and long-term prediction, it was insufficient; although the error rate of STGCN is not the lowest, the growth rate of stgcn is relatively low, which may achieve good results in long-term prediction. STDGCN outperforms the four alternatives, which achieves the lowest error and has an accuracy rate of about 17% higher than the existing optimal method. In addition, the error growth rate of STDGCN is also relatively slow in comparison with other rivals.

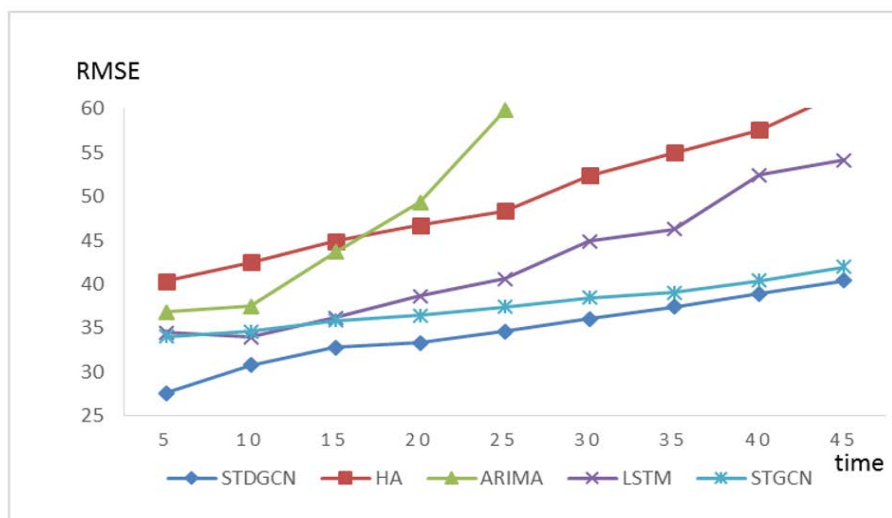


Fig. 7. RMSE Comparison

The RMSE of the experimental results is shown in Fig. 7. The HA method still has relatively high errors; although the error within 5 minutes is not very high, the ARIMA method has a relatively high growth rate and becomes the method with the largest error among all methods at the 17th minute; the LSTM method has a good prediction result in a short time, but the error growth rate gradually increases with time, so LSTM is available for short-term prediction; compared with the traditional method, STGCN has good performance, the accuracy of this method is relatively high, and the error growth rate is slow; STDGCN has a low initial error and a slow growth rate in a short time, which is therefore more suitable for short-term and medium-term forecasts.

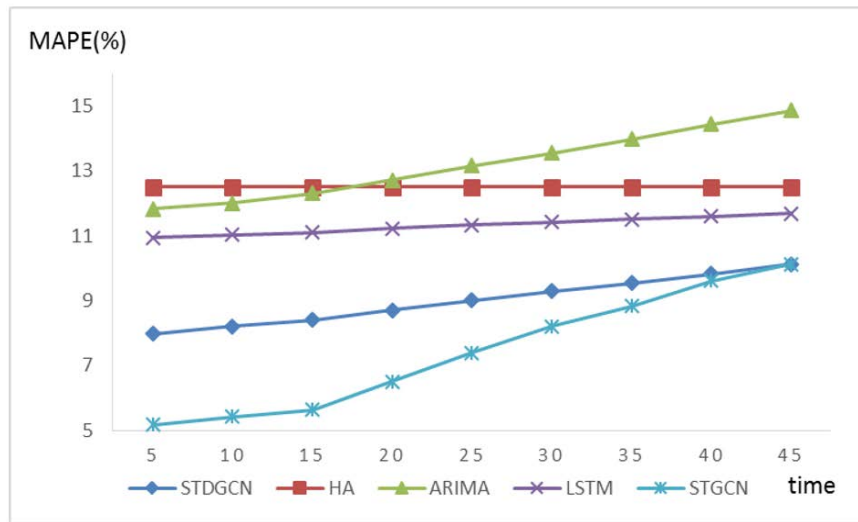


Fig. 8. MAPE Comparison

Fig. 8 shows that these traditional methods have relatively stable performance. The growth rate of error rate is basically unchanged, but the base number of errors is relatively high. Under this standard, STGCN performance is the best, the error value is small, but the error growth rate of our method is lower than its growth rate. With the increase of prediction time, at 45 minutes, the error rates of the two methods are basically the same, but we have a lower growth rate. In the long-term prediction, our method prediction accuracy is better than STGCN.

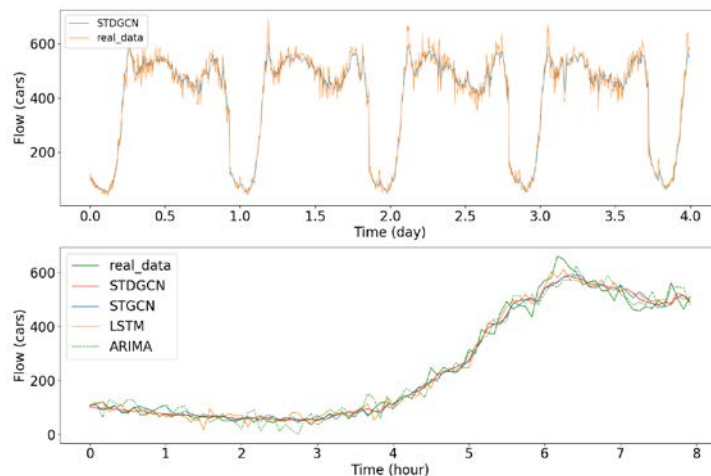


Fig. 9. Comparison of Traffic Flow Predictions

The comparison of traffic flow is shown in **Fig. 9**. The first half of the **Fig. 9** is a comparison between the predicted value of our method and the real traffic volume in four days. We can see intuitively that our method can predict the change of traffic flow in the future. The second half of the **Fig. 9** is the comparison between the predicted traffic flow of four methods in our paper and the real traffic flow in eight hours (From the introduction of the previous part, we know that HA method has the worst performance. In order to make the figure clear and intuitive, we did not reflect the predicted traffic flow of HA method on the figure). From **Fig. 9**, we can see that in these methods, our method can better reflect the change trend of traffic flow in the whole prediction time, and it is closer to the real data, with smaller error.

It can be seen that deploying a dilated convolution structure and a graph convolution structure in the model effectively capture time and space information, and incorporate the influence of lane occupancy to consider the influence of peripheral factors, which can reduce parameters and improve training efficiency.

## 6. Conclusions and Future Work

This paper proposes a model based on dilated convolution and graph convolution, which simultaneously integrate temporal and spatial features to add the impact of road occupancy. To test the performance of the proposed model, experiments are performed on real highway traffic data. The results demonstrate that STDGCN reduced the required parameters and time complexity, as well as improve the prediction accuracy. This model is also applicable to the traffic flow prediction problem of other traffic roads. In the next work, we hope to introduce other strategies to optimize the network structure by considering the impact of complex factors such as weather, simplify parameters, and apply it to a wider field.

## References

- [1] Jiao Yao, Kaimin Zhang, Yaxuan Dai, Jin Wang, "Power Function-based Signal Recovery Transition Optimization Model of Emergency Traffic," *Journal of Supercomputing*, vol. 74, pp.7003-7023, 2018. [Article \(CrossRef Link\)](#).
- [2] Ryder Benjamin, Dahlinger Andre, Gahr Bernhard, Zundritsch Peter, Wortmann Felix.and Fleisch Elgar, "Spatial prediction of traffic accidents with critical driving events-Insights from a nationwide field study," *Transportation research part A: policy and practice*, vol. 124, pp. 611-626, 2019. [Article \(CrossRef Link\)](#).
- [3] Shuren Zhou, Wenlong Liang, Junguo Li, Jeong-Uk Kim, "Improved VGG Model for Road Traffic Sign Recognition," *CMC-Computers, Materials & Continua*, vol. 57, no. 1, pp.11-24, 2018. [Article \(CrossRef Link\)](#).
- [4] Jianming Zhang, Wei Wang, Chaoquan Lu, Jin Wang, Arun Kumar Sangaiah, "Lightweight deep network for traffic sign classification," *Annals of Telecommunications*, vol. 75, pp. 369-379, 2020. [Article \(CrossRef Link\)](#).
- [5] Baowei Wang, Weiwen Kong, Hui Guan, Neal N. Xiong, "Air Quality Forecasting Based on Gated Recurrent Long Short Term Memory Model in Internet of Things," *IEEE Access*, 7, 69524-69534, 2019. [Article \(CrossRef Link\)](#).
- [6] Baowei Wang, Weiwen Kong, Naixue Xiong, "A dual-chaining watermark scheme for data integrity protection in Internet of Things," *Cmc-computers Materials & Continua*, 58(3), 679-695, 2019. [Article \(CrossRef Link\)](#).
- [7] Belghachi Mohammed and Debab Naouel, "An Efficient Greedy Traffic Aware Routing Scheme for Internet of Vehicles," *CMC-COMPUTERS MATERIALS & CONTINUA*, 60(3), 959-972, 2019. [Article \(CrossRef Link\)](#).



- [8] Huey-Kuo Chen, and Che-Jung Wu, "Travel time prediction using empirical mode decomposition and gray theory: Example of National Central University bus in Taiwan," *Transportation research record*, vol. 2324(1), pp.11-19, 2012. [Article \(CrossRef Link\)](#).
- [9] Billy M Williams, and A.Hoel Lester, "Modeling and forecasting vehicular traffic flow as a seasonal ARIMA process: Theoretical basis and empirical results," *Journal of transportation engineering*, vol. 129, no. 6, pp. 664-672, 2003. [Article \(CrossRef Link\)](#).
- [10] Xia, Zhuoqun, Zhenzhen Hu, and Junpeng Luo, "UPTP Vehicle Trajectory Prediction Based on User Preference Under Complexity Environment," *Wireless Personal Communications*, vol. 97, pp. 4651-4665, 2017. [Article \(CrossRef Link\)](#).
- [11] Hongyu Sun, Henry Liu, Heng Xiao, Rachel He, Ran Bin, "Use of local linear regression model for short-term traffic forecasting," *Transportation Research Record*, vol. 1836, pp. 143-150, 2003. [Article \(CrossRef Link\)](#).
- [12] Min Wanli, and Laura Wynter, "Real-time road traffic prediction with spatio-temporal correlations," *Transportation Research Part C: Emerging Technologies*, vol. 19, no. 4, pp. 606-616, 2011. [Article \(CrossRef Link\)](#).
- [13] Vlahogianni, Eleni I., Matthew G. Karlaftis, and John C. Golias, "Optimized and meta-optimized neural networks for short-term traffic flow prediction: A genetic approach," *Transportation Research Part C: Emerging Technologies*, vol. 13, no. 3, pp.211-234, 2005. [Article \(CrossRef Link\)](#).
- [14] Vlahogianni, Eleni I, "Computational intelligence and optimization for transportation big data: challenges and opportunities," *Engineering and Applied Sciences Optimization*, Springer, Cham, pp. 107-128, 2015. [Article \(CrossRef Link\)](#).
- [15] EDES, YORGOS J. STEPHAN, Panos G. Michalopoulos, and Roger A. Plum, "Improved estimation of traffic flow for real-time control," *Transportation Research Record*, vol. 95, pp. 28-39, 1980. [Article \(CrossRef Link\)](#).
- [16] Ahmed, Mohammed S., and Allen R. Cook, "Analysis of freeway traffic time-series data by using Box-Jenkins techniques," No. 722, pp. 1-9, 1979. [Article \(CrossRef Link\)](#).
- [17] Okutani, Iwao, and Yorgos J. Stephanedes, "Dynamic prediction of traffic volume through Kalman filtering theory," *Transportation Research Part B: Methodological*, vol. 18, no. 1, pp. 1-11, 1984. [Article \(CrossRef Link\)](#).
- [18] Hui Feng Ji, Aigong Xu, Xin Sui, Lanyong Li, "The applied research of Kalman in the dynamic travel time prediction," in *Proc. of 2010 18th International Conference on Geoinformatics. IEEE*, pp.1-5, 2010. [Article \(CrossRef Link\)](#).
- [19] Defferrard, Michaël, Xavier Bresson, and Pierre Vandergheynst, "Convolutional neural networks on graphs with fast localized spectral filtering," in *Proc. of NIPS'16: Proceedings of the 30th International Conference on Neural Information Processing Systems*, pp. 3844-3852, 2016. [Article \(CrossRef Link\)](#).
- [20] Yisheng Lv, Yanjie Duan, Wenwen Kang, Zhengxi Li, "Traffic flow prediction with big data: a deep learning approach," *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 2, pp. 865-873, 2015. [Article \(CrossRef Link\)](#).
- [21] Quanjun Chen, Xuan Song, Harutoshi Yamada, Ryosuke Shibasaki, "Learning deep representation from big and heterogeneous data for traffic accident inference," in *Proc. of Thirtieth AAAI Conference on Artificial Intelligence*, pp. 338-344, 2016. [Article \(CrossRef Link\)](#).
- [22] Hinton, Geoffrey E., Simon Osindero, and Yee-Whye Teh, "A fast learning algorithm for deep belief nets," *Neural computation*, vol. 18, no. 7, pp. 1527-1554, 2006. [Article \(CrossRef Link\)](#).
- [23] Jia, Yuhuan, Jianping Wu and Yiman Du, "Traffic speed prediction using deep learning method," in *Proc. of 2016 IEEE 19th International Conference on Intelligent Transportation Systems (ITSC). IEEE*, pp. 1217-1222, 2016. [Article \(CrossRef Link\)](#).
- [24] Huang, Wenhao, Wenhao Huang, Guojie Song, Haikun Hong, Kunqing Xie, "Deep architecture for traffic flow prediction: deep belief networks with multitask learning," *IEEE Transactions on Intelligent Transportation Systems*, vol. 15, no. 5, pp. 2191-2201, 2014. [Article \(CrossRef Link\)](#).

- [25] Xingjian Shi, Zhourong Chen, Hao Wang, Dit-Yan Yeung, Wai Kin Wong, Wang-chun WOO, “Convolutional LSTM network: A machine learning approach for precipitation nowcasting,” in *Proc of NIPS'15: Proceedings of the 28th International Conference on Neural Information Processing Systems - Volume 1*, pp. 802-810, 2015. [Article \(CrossRef Link\)](#).
- [26] Sepp Hochreiter, and Jürgen Schmidhuber, “Long short-term memory,” *Neural computation*, vol. 9, no. 8, pp. 1735-1780, 1997. [Article \(CrossRef Link\)](#).
- [27] Yu, Bing, Haoteng Yin, and Zhanxing Zhu, “Spatio-temporal graph convolutional networks: A deep learning framework for traffic forecasting,” in *Proc. of the Twenty-Seventh International Joint Conference on Artificial Intelligence (IJCAI-18)*, 3634-3640, 2017. [Article \(CrossRef Link\)](#).
- [28] Kipf, Thomas N., and Max Welling, “Semi-supervised classification with graph convolutional networks,” *arXiv preprint arXiv:1609.02907*, 2016. [Article \(CrossRef Link\)](#).
- [29] Yu, Fisher, and Vladlen Koltun, “Multi-scale context aggregation by dilated convolutions,” *arXiv preprint arXiv:1511.07122*, 2015. [Article \(CrossRef Link\)](#).



**Xiufang Sun** is currently pursuing the master degree at the School of Computer Science and Technology, Qingdao University. She received the bachelor degree in Digital media technology from the Qingdao University, China in 2017. Her research interest is urban traffic forecast and mobile social networks.



**Jianbo Li** is currently a professor at the Computer Science and Technology College of Qingdao University. He received the bachelor and master degrees in computer science from Qingdao University, China in 2002 and 2005, respectively, and the PhD degree in computer science from the University of Science and Technology of China in 2009. His research interests include urban computing, mobile social networks, and data offloading.



**Zhiqiang Lv** is currently pursuing the master degree at the College of Computer Science & Technology, Qingdao University. He received the bachelor degree in the College of Information & Electrical Engineering, Ludong University in 2019. His research field is preprocessing of big data analysis for urban computing and intelligent transportation.



**Chuanhao Dong** is currently earning his master's degree in computer science and technology from Qingdao University. He received his bachelor's degree from Taiyuan University of Science and Technology in 2019. His research direction is urban computing.