Proceedings of the 2023 International Conference on Mechatronics and Smart Systems
DOI: 10.54254/2755-2721/9/20230076

Traffic flow prediction based on T-GCN in extreme weather: A case study of Beijing

Xiaotao Wu[1, †], Zhuoheng Ying[2, †] and Haoyue Zhang[3, 4, †]

[1] Lanzhou University College of Information Science and Engineering, Lanzhou University, Lanzhou, 730000, China
[2] Chang'an Dublin International College of Transportation, Chang'an University, Xi'an, 710064, China
[3] Tianjin University College of Management & Economicsian, Tianjin University, Tianjin, 300072, China

[4] 3020209067@tju.edu.cn
[†] These authors contributed equally

Abstract. In recent years, frequent extreme weather events have significantly impacted traffic management and urban planning. Consequently, accurate and real-time traffic prediction has become crucial for effective urban traffic planning, management, and control. This study aims to predict road traffic flow in Beijing under extreme weather conditions. To achieve this goal, the study analyses historical weather data of Beijing and identifies four typical extreme weather conditions. Using traffic data collected between 2014 and 2016, the study then employs the time series graph convolutional network (T-GCN) model to predict traffic flow under extreme weather conditions.

Keywords: deep learning, T-GCN, extreme weather, traffic flow prediction.

1. Introduction

In recent years, China's rapid urbanization has resulted in an exponential increase in traffic demand, leading to increased traffic pressure within cities. Extreme weather conditions have emerged as a significant factor contributing to the destruction of traffic conditions. Heavy rainfall, as a form of extreme weather, refers to accumulated precipitation of 50mm or above within 24 hours and 30mm or above within 12 hours, which can cause floods, landslides, debris flows, and other disasters [1]. Every year, there are numerous incidents of road disruptions, vehicle and pedestrian strandings due to extreme weather conditions. Storm weather has altered the traffic environment of city roads, reduced visibility and disrupting drivers' view. Water accumulation on the road surface reduces the adhesion coefficient, further decreasing vehicle braking performance. Rainstorms, in particular, exacerbate these issues due to higher and stronger rainfall leading to increased road water accumulation. Consequently, road traffic during rainstorms is often in a degraded state [2]. It is critical to note that road traffic plays an indispensable role in meeting citizens' daily travel needs and maintaining the normal operation of the urban system. Rainstorms can cause waterlogging and flooding, leading to traffic interruptions or paralysis, which can significantly impact citizens' travel and urban life [3]. Due to frequent and unpredictable extreme weather and the complexity of modern road traffic in China, large-scale traffic delays and accidents caused by bad weather are commonplace, resulting in significant economic losses and casualties. Therefore, accurate and real-time prediction of extreme weather and its impact on traffic flow is of significant research importance.

There are two distinct categories of methods that are employed for the purpose of predicting short-term traffic flow on urban highways. The first is to use traditional methods based on time series analysis such as multiple linear regression models to predict short-term traffic flow [4]. As an illustration, in one study, Kumar and Vanajakshi used a serial ARIMA model to forecast traffic flow with limited data [5]. In another study, Kumar et al. dished a data-efficient method. This approach is based on the Kalman filter model, which solves the problem of relying on a great deal of data for development [6].

In recent years, machine learning methods have gained popularity in short-term traffic flow prediction of metropolitan highways. Some models, which include SVR and KNN, continuously modify their parameters through adaptive learning to capture complicated nonlinear interactions (KNN). While these models have the potential to outperform traditional methods, the pursuit of high accuracy can result in extended training times and greater data sample requirements [7]. To address these issues, Li et al. [8] introduced a prediction approach that utilizes kernel function switching. This technique selects the most suitable kernel function for predicting short-term traffic flow, taking into account the variations in SVR models that use different kernel functions for processing traffic flow at various times. Similarly, Cai et al. [9] enhanced the prediction accuracy of KNN model by considering the spatio-temporal correlation of road network using Gaussian weighting method. Overall, machine learning methods provide a promising approach for this task. However, the selection of appropriate models and feature extraction methods should be carefully considered to balance accuracy and efficiency, especially for large-scale road networks.

Deep learning has gained significant traction in the machine learning field in recent years. While it is true that these models (such as RNN and LSTM) exhibit adeptness in learning long-term correlations in traffic flow sequences and capturing intricate temporal characteristics, they often neglect the spatial features of traffic networks. This inadequacy poses a challenge in the prediction of traffic flow between road networks. To address this issue, Bruna et al. [10] used the Graph Convolutional Networks (GCN) to solve the spatial correlation in road networks problem. By using graph structure information to extract the road network topology and processing the irregular data in the graph structure, traffic prediction tasks can be completed.

Subsequently, Zhu et al. [11] used a novel prediction model that combines RNN-GCN and Belief Rule Base (BRB). The model begins by using BRB for data fusion, RNN and GCN models are then employed to obtain temporal correlations of traffic data. Similarly, Wu et al. [12] employed GCN to extract topological structure features from traffic data and LSTM structure to extract time features. The integration of these features with ResNet was carried out with the objective of optimizing the overall model. Consequently, the model succeeded in predicting traffic flow. Zhang et al. [13] used LSTM neurons to extract temporal correlations and convolutional transmission to extract spatial correlations to predict traffic flow.

While deep learning presents enormous opportunities, the existing models still have limitations. They rarely consider factors such as road networks and extreme weather and only focus on predicting traffic on a single road. Thus, the prediction accuracy needs to be improved. This study built on previous prediction models to extract the road network topology and consider the traffic state of the road network sections and the spatio-temporal correlation of traffic flow. The T-GCN model was chosen as the preferred model to address the challenge of predicting short-term traffic flow under extreme weather conditions. This approach aims to improve the accuracy of predicting traffic flow on urban roads during extreme weather, which is currently an unmet need in the field. By leveraging the T-GCN model, the researchers hope to enhance the accuracy of result, particularly under conditions of severe weather, such as heavy rain, snow, and high winds. This would be a significant improvement over existing approaches, which often struggle to accurately predict traffic flow during such conditions.

Researchers are able to account for the temporal and spatial dependencies in traffic flow data, which can be particularly important during extreme weather events, when traffic patterns can change rapidly and unpredictably. Overall, this approach has the potential to significantly enhance our ability to predict traffic flow under extreme weather conditions, which could have important implications for transportation planning and management.

2. Method

2.1. Overview of the problem

The traffic prediction problem pertains to forecasting future traffic conditions in different segments of a road network by utilizing historical traffic data. In light of the prevailing occurrence of extreme weather conditions across the globe, this study takes into account the processing of weather and traffic data to obtain a traffic dataset applicable to extreme weather conditions. Subsequently, the T-GCN model is used to forecast traffic flow data in Beijing under such extreme weather circumstances.

The T-GCN model consists of two primary components, as illustrated in Figure 1: GCN, which enables the processing of graph-structured data, and GRU, which enhances the model's ability to capture temporal dependencies.

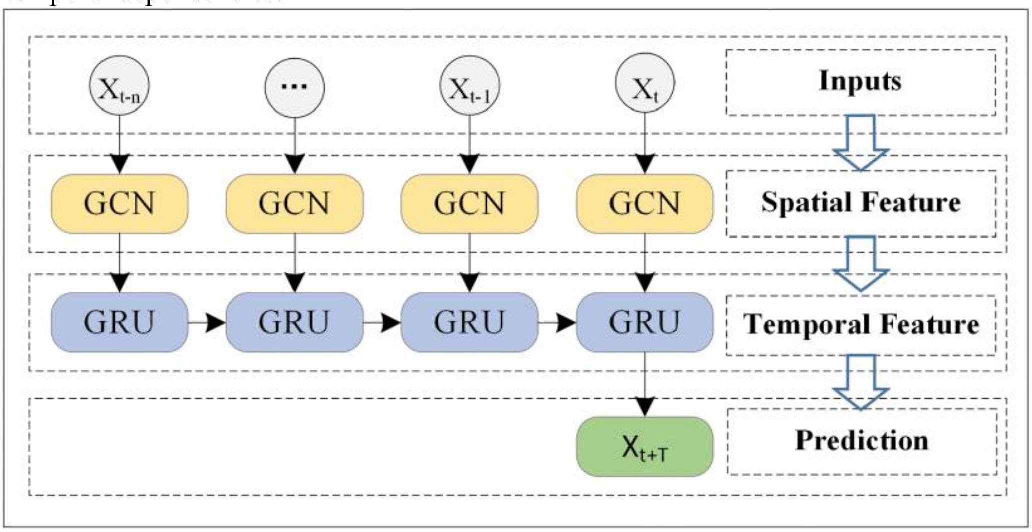


Figure 1. Overview. We take historical traffic data of the road network as input and prediction of traffic conditions in Beijing under extreme weather conditions.

2.2. Extreme weather analysis

2.2.1. Definition of extreme weather. Extreme weather is defined as local weather conditions that can be destructive and are not conducive to human activities. Before predicting the traffic flow of Beijing's road network under extreme weather conditions, it is necessary to establish criteria for extreme weather. The criteria is shown in Table 1.

Table 1. Definition of extreme weather.

Extreme Weather Types	Weather Conditions
Heavy Foggy Weather	The weather with visibility less than 1 km
Strong Wind Weather	The wind speed is higher than 45 km / h, the branches of the trees swing, the umbrella is difficult and the sea surface appears large waves
High-Temperature Weather	Temperature greater than or equal to 34 degrees Celsius
Low-Temperature Weather	Temperature is less than or equal to minus 8 degrees Celsius

DOI: 10.54254/2755-2721/9/20230076

2.2.2. Weather screening. Through the above definition of extreme weather types (Table 1), the weather without weather conditions is set as normal weather, and the weather data is screened based on this. During the screening process, the weather data meets any of the above indicators, and the selection can be completed directly. Repeat the data in the meteorological database to obtain all extreme weather data.

2.3. Problem definition

This paper considers traffic flow prediction problem with the mapping function f on the basis of constructing the feature matrix X and the road network topology G, and further calculating the traffic flow at T moments in the future. The graph $G = (V, E, Y)$ is defined to represent the topology of the road network. Each road section is regarded as a node, V is the set of road nodes, N is the number of nodes, E is the set of edges, and $V = \{v_1, v_2, \dots, v_M\}$. The characteristic matrix $Y^{N\times P}$ is defined, and $Y_t \in R^{N\times i}$ represents the characteristic of each road nodes in the road network at time i.

2.4. Spatial dependence model

GCN clarifies the relationship between different nodes by using adjacency matrix A, $A \in R^{N\times N}$. Specifically, we use a 2-layer GCN model as shown in Equation (1):

$$f(X, A) = \sigma\big(\widehat{A} Relu\big(\widehat{A} X W_0\big) W_1\big) \tag{1}$$

where A and X represent the adjacency matrix and the characteristic matrix respectively, $\sigma(\cdot)$ and $Relu()$ are the activation functions, and W_0 and W_1 represent the weight matrices of the first and second layers respectively.

2.5. Time-dependent model

The RNN is a widely used model for sequence data analysis, but it suffers from issues such as gradient vanishing and exploding. To overcome these limitations, its variants LSTM and GRU have been introduced. Both of these models can capture the spatio-temporal dependencies of the sequence data effectively. Although the LSTM has a more complex structure and longer training time, it can learn more long-term dependencies than the GRU. However, to reduce the computational cost, we choose the GRU for our study to capture the temporal dependencies of the traffic flow. The specific formula for the GRU model is as follows:

$$r_t = \sigma(W_r[h_{t-1}, x_t]) \tag{2}$$

$$\tilde{h}_t = tanh(W * [r_t * h_{t-1}, x_t]) \tag{3}$$

$$z_t = \sigma(W_z * [h_{t-1}, x_t]) \tag{4}$$

$$h_t = z_t * h_{t-1} + (1 - z_t) * \tilde{h}_t \tag{5}$$

where r_t is a reset gate utilized to regulate the degree of disregard for the previous state information; h_{t-1} refers the hidden state at time t-1; x_t represents the traffic information at time t ; z_t is an update gate used to regulate the degree to which the previous moment's state information enters the current state; $\tilde{h}_t$ signifies the memory content stored at time t; and h_t represents the output state at time t, Please refer to Figure 2 for the specific calculation process.

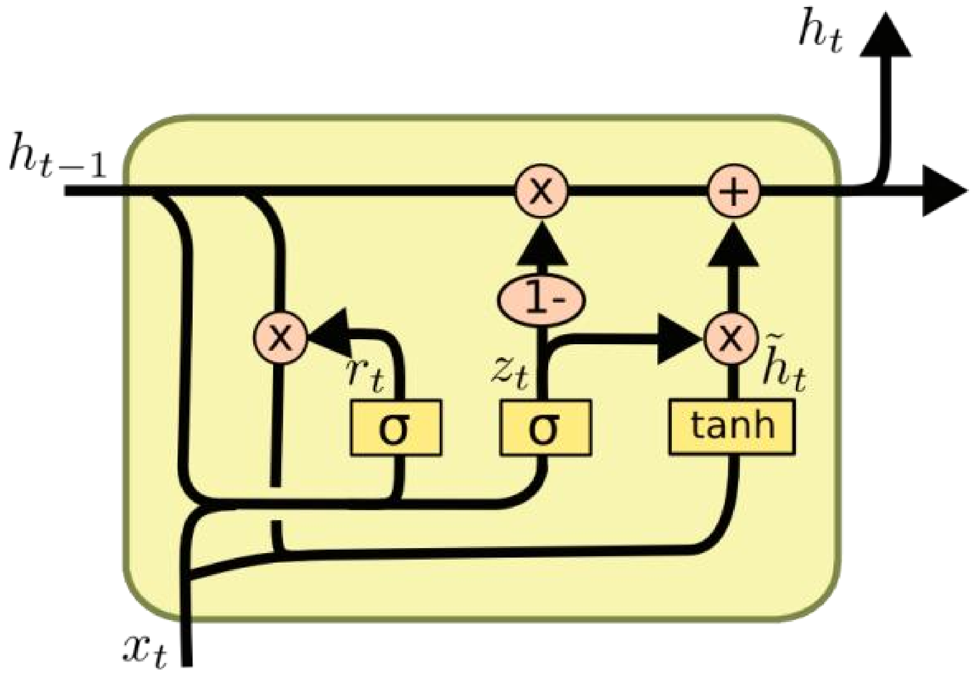


Figure 2. The gated recurrent unit model.

2.6. Temporal graph convolution network

Combining the spatial dependence model GCN with the time dependence model GRU, the T-GCN is established to achieve the task. The specific calculation process can be expressed as:

$$z_t = \sigma(W_z[f(X,A), h_{t-1}] + b_z) \tag{6}$$

$$r_t = \sigma(W_r[f(X,A), h_{t-1}] + b_r) \tag{7}$$

$$\tilde{h}_t = \tanh(W_{\tilde{h}}[f(X,A), r_t * h_{t-1}] + b_c) \tag{8}$$

$$h_t = z_t * h_{t-1} + (1 - z_t) * \tilde{h}_t \tag{9}$$

where $f(X, A)$ represents the graph convolution process, b represents the deviation in the training process, and W represents the weight in the training process.

2.7. Loss function

The objective is to reduce the disparity between the projected and observed traffic flow. To accomplish this, Q_t and $\widehat{Q}_t$ signify the actual and predicted traffic flow, respectively, in various ways. The T-GCN model's loss function, as displayed in formula (10). Moreover, to prevent overfitting, a second term L_{reg} is incorporated, where λ represents a hyperparameter.

$$loss = \|Q_t - \widehat{Q}_t\| + \lambda L_{reg} \tag{10}$$

3. Experiments

3.1. Data description

It is clear that the experiment based on the real transportation data in Chinese. The data are mainly collected from the official website for use. To begin with, we collected taxi data from Beijing for the years 2014, 2015, and 2016.The three years are all occupied with abundant rainy days. We stored the data in a MYSQL database and analysed the extreme weather data in each year using Navicat. Based on this analysis, we selected the 2014 data for our study. The time interval for the data was set to 15 minutes, and we calculated the traffic flow matrix for each 15-minute interval.

For the TaxiBJ dataset, we normalized the data using minimum-maximum normalization and scaled it to the range of [0,1]. The dataset was bifurcated into a training set comprising 80% of the total data, with the aim of utilizing them for model training and evaluation.

3.2. Metrics

To evaluate the T-GCN model, we utilize four metrics:

RMSE:

$$RMSE = \sqrt{\frac{1}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N}\left(q_n^m - \widehat{q_n^m}\right)} \tag{11}$$

MAE:

$$MAE = \frac{1}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N}\left|q_n^m - \widehat{q_n^m}\right| \tag{12}$$

Accuracy:

$$Accuracy = 1 - \frac{\|Q-\widehat{Q}\|_F}{\|Q\|_F} \tag{13}$$

Coefficient of Determination (R^2):

$$Accuracy = \frac{\sum_{m=1}^{M}\sum_{n=1}^{N}\left(q_n^m - \widehat{q_n^m}\right)^2}{\sum_{m=1}^{M}\sum_{n=1}^{N}\left(q_n^m - \overline{Q}\right)^2} \tag{14}$$

3.3. Experimental results

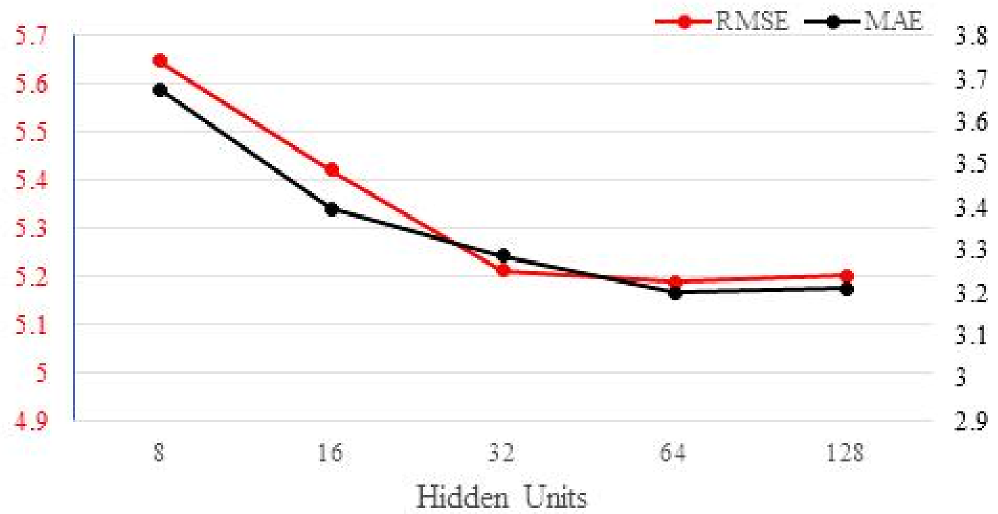
Parameter settings are crucial in the T-GCN model. We mainly focus on four key parameters: learning rate, batch size, and hidden unit. In our experiment, the learning rate is set to 0.001, the batch size is set to 32, and the training period is set to 3000 through manual adjustment. We also set the number of hidden units to optimize the performance of the model.

We evaluated the performance of the T-GCN model against the SVR method on the TaxiBJ dataset, conducting prediction tasks at 15, 30, 45, and 60-minute intervals. Table 2 presents the models' prediction outcomes. Remarkably, the T-GCN model outperforms the SVR method in terms of all evaluation metrics for almost all prediction intervals, affirming the effectiveness of the T-GCN model for spatio-temporal traffic prediction tasks.

Table 2. Prediction results of T-GCN model and SVR method on TaxiBJ dataset.

Time	Model	Metric			
		RMSE	MAE	Accuracy	R^2
15min	SVR	5.4251	3.1514	0.9012	0.8459
	T-GCN	5.1876	3.1989	0.9117	0.8602
30min	SVR	5.3241	3.0147	0.8948	0.8498
	T-GCN	5.0517	3.0948	0.9084	0.8543
45min	SVR	5.5109	3.0946	0.90648	0.8395
	T-GCN	5.1142	3.1129	0.9101	0.8519
60min	SVR	5.3264	3.1148	0.9011	0.8465
	T-GCN	5.0945	3.1475	0.9093	0.8504

We conducted experiments with different numbers of hidden units in the T-GCN model to choose the optimal value, as it is a crucial parameter that can significantly affect the prediction accuracy. The range of hidden units [8, 16, 32, 64, 128] was tested and the prediction results were compared to determine the optimal value, based on the analysis of the variation in prediction accuracy.



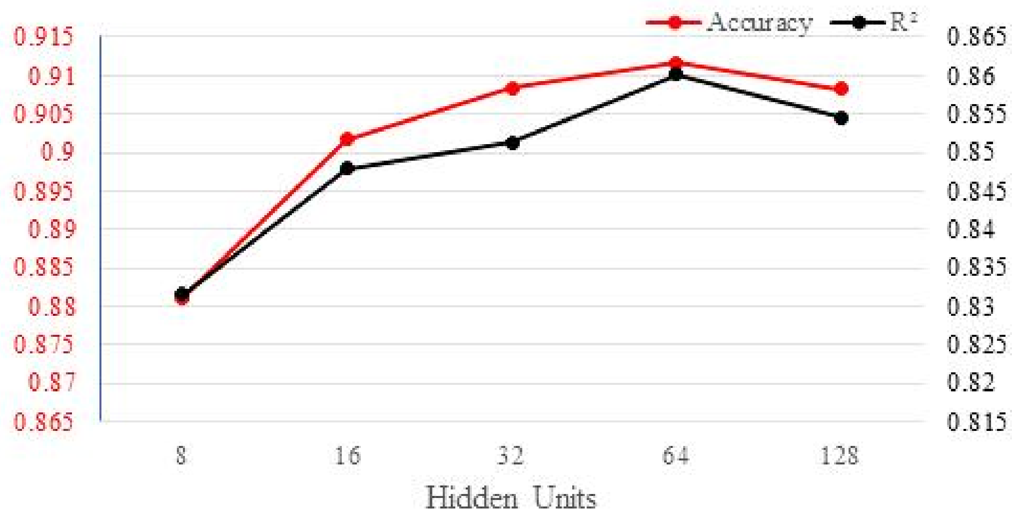


Figure 3. Comparison of predictive performance under different hidden units.

Based on the information presented in Figure 3, the horizontal axis displays the quantity of hidden units, while the vertical axis depicts changes in various evaluation indicators. Notably, the minimum values for both RMSE and MAE are achieved when the number of hidden units is set to 64, accompanied by corresponding maximum values for Accuracy and R^2. Thus, we can conclude that the optimal number of hidden units for superior prediction performance is 64.

3.4. Model interpretation

Figure 4 presents the visualization outcomes of the T-GCN model for distinct forecasting time intervals: (a) 15 minutes, (b) 30 minutes, (c) 45 minutes, and (d) 60 minutes. The color proximity between the projected and actual lines indicates superior prediction performance. From (a) to (d), it can be concluded that over time, the proximity between the projected line and the actual line gradually increases, but the two lines remain basically close. The results from Figure 4 demonstrate the T-GCN model's ability to precisely capture spatio-temporal correlations and dynamic variations in traffic flow, consequently forecasting the future traffic situation with remarkable accuracy.

DOI: 10.54254/2755-2721/9/20230076

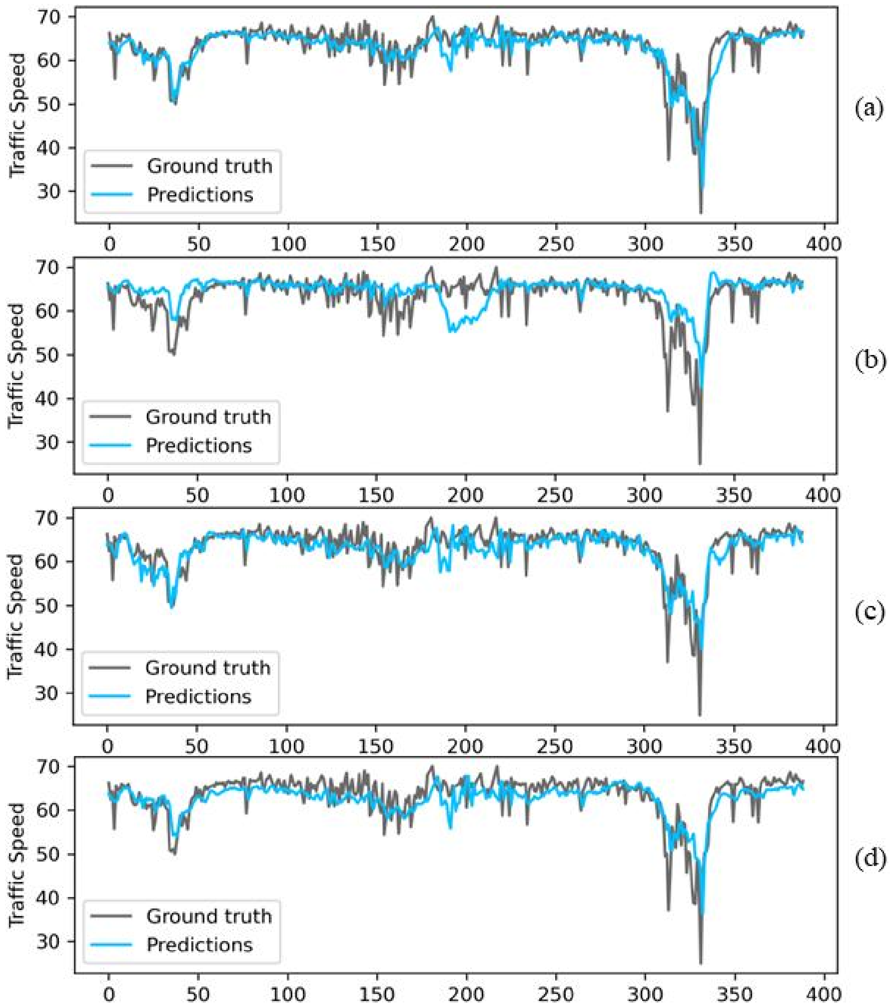


Figure 4. The visualization results for prediction horizons of 15, 30, 45, and 60 minutes.

4. Future work

In future research, there is a need to investigate the impact of external features on the accuracy of the T-GCN model in predicting urban traffic. Specifically, the incorporation of external factors such as road infrastructure, points of interest, and social events can provide valuable insights into traffic patterns and enable the T-GCN model to make more accurate predictions.

To achieve this objective, a transformer combination model can be employed, which has demonstrated its ability to optimize longer-term time-series predictions and can be fine-tuned for specific tasks, enabling the model to adapt to various traffic flow prediction scenarios. Furthermore, the transformer combination model can be extended to time series problems in other domains, providing a more comprehensive understanding of spatio-temporal modeling beyond urban traffic forecasting.

Moreover, the proposed transformer combination model can be improved by exploring the best combination of T-GCN and transformer models, optimizing the hyperparameters, and incorporating additional features. Additionally, an investigation of the potential impact of different types of external features and their optimal integration with the T-GCN and transformer models is needed to improve the model's accuracy further.

5. Conclusion

The investigation highlights a time series T-GCN model capable of predicting traffic flow during extreme weather conditions. The urban traffic network, as a new communication system is modeled as a graph, with roads portrayed as panel points and edges linking them together. The node attributes in the graph represent the traffic at different time intervals. By employing graph convolutional networks, the T-GCN model captures spatial dependencies by analyzing the graph's topology, and temporal dependencies by utilizing a GRU model to capture dynamic changes in node attributes. The T-GCN model's effectiveness in addressing spatiotemporal traffic prediction is evidenced in its superior performance when compared to the support vector regression (SVR) model. The results demonstrate the T-GCN model's potential to effectively capture spatiotemporal features from traffic data and its potential utilization in other spatiotemporal prediction tasks. T-GCN model works as a new convolutional neural network model will get more prominent applications

References

[1] Xiao G, Wei X. 2001. J. Urban disaster analysis and disaster reduction construction in Shanghai. Catastrophology, (01).
[2] Ji X. 2021. Study on dynamic Traffic characteristics and Traffic flow allocation Method of urban road under Rainstorm weather. Beijing transportation University, PhD dissertation.
[3] Li M Y. 2019. Study on urban Commuting exposure and traffic characteristics under rainstorm and waterlogging. Shanghai: East China Normal University. PhD dissertation.
[4] Yin X., Wu G., Wei J., et al. 2020. J. A comprehensive survey on traffic prediction. arXiv preprint arXiv:2004.08555:1-16.
[5] Kumar S V, Vanajakshi L. 2015. J. Short-term traffic flow prediction using seasonal ARIMA model with limited input data. European Transport Research Review, 7(3): 1-9.
[6] Kumar S V. 2017. J. Traffic flow prediction using Kalman filtering technique. Procedia Engineering, 187: 582- 587.
[7] Xiang X W, Lun H X. 2018. J. Short-term Traffic Flow Prediction Based on Deep Learning. Transportation system engineering and information, 18(1):81-88.
[8] Li L W, Zhang J, Yang F, Ran B. 2017. J. Short-term traffic volume forecasting model based on kernel function switching and Support vector Regression. Journal of Southeast University (Natural Science Edition),47(5):1032-1036.
[9] Cai P, Wang Y, Lu G, et al. 2016. J. A spatiotemporal correlative k-nearest neighbor model for short-term traffic multistep forecasting. Transportation Research Part C: Emerging Technologies, 62: 21-34.
[10] Bruna J, Zaremba W, Szlam A, et al. 2013. J. Spectral networks and locally connected networks on graphs. arXiv preprint arXiv:1312.6203.
[11] Xie Y X, Zhou G Z, He W H, et al. 2020. J. A Novel Traffic Flow Forecasting Method Based on RNN-GCN and BRB. Journal of Advanced Transportation:1-11.
[12] Wu Z, Huang M, Zhao A. 2021. J. Traffic prediction based on GCN-LSTM model. Journal of Physics: Conference Series. IOP Publishing, 1972(1): 012107.
[13] Zhang Y, Hu Y, Xin D R. 2021. J. Short-term traffic flow forecasting method based on deep learning considering spatio-temporal Correlation. Journal of Intelligent Science and Technology, 3(2): 172-178.