ANNALS OF APPLIED SCIENCES

MIDITERRANEAN JOURNALS

Original Research

DOI: 10.55085/aas.2023.715

Traffic Flow Prediction: A Spatiotemporal Convolution Model Considering Traffic Accident Influence Coefficient Matrix

Xue Liu, Zhiqiang Lv, Jianbo Li, Zhaobin Ma, Benjia Chu, Fengqian Xia, Yang Liu
College of Computer Science & Technology
Qingdao University, Qingdao, Shandong 266001 China

ABSTRACT

Accurate and timely traffic flow prediction plays a vital role in traffic planning. But the work is more complicated for the road where the traffic accident occurred and the roads around it. This is because the influence of traffic accidents not only acts on the traffic flow of the current road but also quickly spreads to the surrounding road. Traffic accidents differ from other external factors, and their influence on surrounding traffic flow is distinct. Therefore, this work proposes an influence coefficient matrix to express the degree of influence between any two roads to quickly capture the impact of traffic accidents on different road traffic flows. Moreover, this work proposes a hybrid network model based on graph and temporal convolution. To address the spatial dependence between traffic flow data and roads, we selected a graph convolutional network that can be used to analyze the complicated non-Euclidean spatial data in order to extract the spatial dependence. Taking into consideration the temporal dependence of traffic flow data, the temporal convolution model is chosen in this work to model the temporal dependence of the data. Compared to traditional statistical models, single deep learning models, and complex spatiotemporal convolutional models, our model's performance has been improved by 30% to 50%.

KEYWORDS: Traffic Planning, Traffic Flow Prediction, Graph Convolution Network, Time Convolution Network, Traffic Accident Influence Coefficient Matrix, The Spatial-Temporal Correlation.

Received: 8 Aug 2023
Revised: 11 Sept 2023
Accepted: 23 Sept 2023
Published: 22 Nov 2023

Correspondence: Jianbo Li
College of Computer Science & Technology. Qingdao University, Qingdao, Shandong 266001 China
Email: lijianbo@ubinet.cn

Cite this article as: Liu X, Lv Z, Li J, Ma Z, Chu B, Xia F, et al. Traffic Flow Prediction: A Spatiotemporal Convolution Model Considering Traffic Accident Influence Coefficient Matrix. Ann Appl Sci. 2023:715.
https://doi.org/10.55085/aas.2023.715

Authors' contributions
The participation of each author corresponds to the criteria of authorship and contributorship emphasized in the Recommendations for the Conduct, Reporting, Editing, and Publication of Scholarly work in Medical Journals of the International Committee of Medical Journal Editors. Indeed, all the authors have actively participated in the redaction, the revision of the manuscript, and provided approval for this final revised version.

Acknowledgments
None

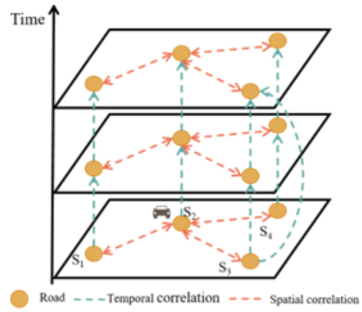
Funding
No funding was received from any organization to conduct the present study.

Conflict of interest
The author declares that there is no conflict of interest regarding the publication of this article.

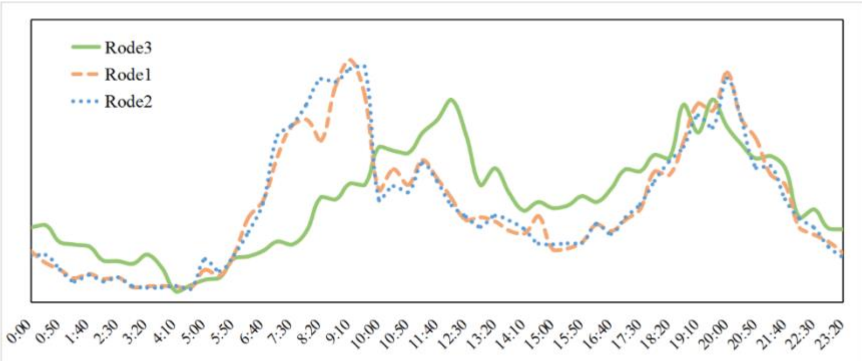
1. INTRODUCTION

With the widespread application of Intelligent Transportation Systems (ITS), processing traffic prediction through intelligent computing has received more and more attention. Among them, accurate and timely traffic flow prediction has become one of the most vital challenges in ITS. Based on the predicted traffic flow, relevant departments can better develop relevant traffic planning strategies to reduce urban congestion, improve traffic efficiency, and decrease the number of public accidents. Meanwhile, travelers can choose the appropriate travel route based on current traffic conditions, thereby reducing time costs and financial losses [1].

However, since traffic flow prediction is a typical spatiotemporal process, the key problem to achieving traffic flow prediction is how to extract the time dependence and spatial correlation [2]. Most current methods for predicting traffic flow are based on graph architecture and sequence learning models, which have been able to extract spatiotemporal information better. However, when predicting traffic flow, we found two problems: first, when a traffic accident occurs in a certain area, the influence generated by that accident will quickly spread to its surrounding areas, causing congestion phenomenon and thus affecting the traffic flow in the surrounding areas [3]. Similarly, the state of the area where a traffic accident occurs at that moment also affects the state of traffic at subsequent moments. That is, traffic accidents can have a spatial and temporal influence on traffic flow. This work uses a graph structure to describe the spatial structure of traffic flow, where nodes denote roads in the region and edges denote the connectivity between two roads. As shown in Fig. 1(a), the yellow line indicates the influence on the surrounding area and the green line indicates the influence on future moments when a traffic accident occurs. That is, a traffic accident affects the traffic flow in the adjacent area and the following period. Therefore, it is important to fully analyze the influence relationship between different regions to increase traffic flow prediction accuracy [4]. Second, traffic accidents differ from other external factors (e.g., weather and holidays). Other external factors affect each road in a region to the same extent; i.e., they are consistent. However, traffic accidents affect each road differently. The degree of influence between the two areas depends not only on the distance and connectivity but also on the traffic patterns of the two areas. Regions with similar traffic patterns have similar spatiotemporal correlations. As shown in Fig. 1(b), the traffic modes of both Road 1 and Road 2 are car lanes, while that of Road 3 is bicycle lanes. Therefore, the trend of traffic flow changes on Road 1 and Road 2 is similar, but there is a difference between the trend of traffic flow changes on Road 3. Therefore, modeling the relationship between traffic patterns and the degree of interaction between the two areas is crucial in traffic flow forecasting [5].



(a) Composition of spatiotemporal correlation



(b) The influence of traffic patterns on traffic flow

Fig. 1. The complexity of traffic accident influence on traffic flow.

Based on the above research, it can be concluded that when both roads belong to car lanes, the trend of traffic flow changes is highly similar. If a traffic accident occurs on one road during peak commuting hours, based on the principle that drivers usually choose the same traffic mode, the traffic flow on the other road will be significantly affected, and it is prone to a sudden increase or even blockage of traffic flow. However, for Road 3, with a bicycle lane and a similar distance, the impact of traffic flow is relatively small. Therefore, this work constructs an influence coefficient matrix according to the changing trend of traffic flow between different regions and the distance between regions, which is used to express the degree of influence of traffic accidents between any two regions. Meanwhile, this work proposes a hybrid deep learning framework, namely, a hybrid network model based on graph and temporal convolution (GTCN). This model could thoroughly analyze and predict the spatiotemporal feature of the traffic flow. A convolutional map network is employed to extract the spatial correlation among different regions, and a temporal convolutional network is employed to obtain the temporal correlation of the traffic flow. This work was tested using real datasets, and the results of the experiments show that this model outperforms other current models in predicting traffic flow.

This work proposes a model GDCN for predicting traffic flow, which is used to capture the spatiotemporal correlation of traffic flow data. To address the varying degrees of impact of traffic accidents among different regions, this work proposes a coefficient matrix to represent the degree of impact of traffic accidents in different regions. This work contributed to the following:

(1) This work considers the influence of traffic accidents in one area on the traffic flow in other areas, that is, the capture of the diffusion effect between nodes. This work proposes constructing the influence coefficient matrix according to the changing traffic flow and distance trend in different regions. The coefficient matrix represents the degree of influence of traffic accidents in each region.

(2) This work introduces a hybrid model that fuses convolutional graph and temporal convolutional networks (GTCN). The graphical convolution layer realizes the extraction of spatial correlation for the traffic flow data, and the temporal convolution layer realizes the extraction of temporal dependence for the traffic flow data. Thus, the data's spatial and temporal correlation can be fully analyzed and extracted, and prediction accuracy is improved.

(3) This work uses real datasets for experimental evaluation. The results show that the hybrid model composed of the graph convolution and the temporal convolution models has a relatively high accuracy in predicting traffic flow. In the meantime, the fusion of the influence coefficient matrix and traffic flow data as inputs to the model is more conducive to capturing traffic accident influence on traffic flow. Compared with existing models for the prediction of traffic flow, this work proposes a model

with higher prediction accuracy; RMSE, MAE, and MAPE decreased to 18.72, 12.23, and 0.2, respectively.

The rest of this article is arranged as follows: In Section 2, relevant work is reviewed. In Section 3, the relevant data design and model design were mainly introduced. In Section 4, we mainly introduce experiments and result analysis based on real datasets. The experiment is summarized in Section 5.

2. RELATED WORK

The traffic flow prediction problem has been a widespread concern. Current traffic flow prediction faces two main challenges: extracting the spatiotemporal correlation of traffic flow data and capturing the influence of traffic accidents on traffic flow. Many works presented in the literature have explored how to improve prediction accuracy.

2.1. Traffic flow prediction

Examples of traditional statistical methods include the historical average (HA) [6] and autoregressive integrated moving average (ARIMA). The HA model is the result of taking the average of historical data as the prediction. The ARIMA model predicts future traffic flows by analyzing the correlation between historical and current data. These methods have the characteristics of simple and fast calculation and strong interpretability. However, statistical methods cannot process complex and nonlinear traffic flow data [7].

Machine learning methods solve such problems very well. Traditional machine learning models can handle more complex data. Common models are the K-nearest neighbor algorithm (KNN) [8], support vector machine (SVM) [9], and decision tree (DT) [10]. However, traditional machine learning models rely on feature engineering and expert experience, so their prediction accuracy could be higher for highly nonlinear data.

With the maturity of data acquisition technology and the continuous improvement of computing power, deep learning methods have been gradually applied to traffic forecasting. Huang et al. proposed using a deep belief learning network (DBN) [11]. The network is a special deep neural network formed by stacking Boltzmann machines to learn the features of traffic flow data fully. However, such a dense network makes extracting spatial and temporal features from the input data difficult. The performance of this class of models is greatly degraded under conditions that strictly limit or even completely ignore spatial correlations. To be able to solve this problem, recurrent neural network (RNN) [12], long short-term memory network (LSTM) [13], and gated recurrent unit (GRU) [14] models began to be applied to traffic prediction. Among them, RNN captures the temporal dependency of traffic flow data by setting a self-circulation mechanism. LSTM adds a gating unit based on RNN so that the storage unit can continuously store the updated data, solving the long-term dependence problem. GRU is a variant of the LSTM model. Compared with the LSTM model, the structure is simpler and the parameters are fewer. Convolutional neural network (CNN) is a typical feedforward propagation deep learning network, which has achieved significant achievements in the field of image analysis [15]. Therefore, Zhang et al. [16] proposed treating urban traffic flow as pixel values and historical traffic flow data as a set of images and using CNN to predict traffic flow images at the next timestamp. In this way, the extraction of spatial correlation of data is realized. However, using traditional CNNs to process topological graph data usually requires traversing all possible sequences of node appearances in the graph as input to the model, which leads to problems such as extensive computation and slow training speed [17]. The birth of a graph convolution network (GCN) is a good solution to such problems. GCN can aggregate the features of nodes near a node and learn the features of nodes through weighted aggregation to perform a series of prediction tasks [18]. GCNs have shown better performance in various traffic applications. Zhu et al. [19] used GCN to extract positional features through positional attributes; Peng et al. [20] proposed to construct a new correlation dynamic graph based on historical traffic flow and use a model integrated with GCN and LSTM to mine the spatiotemporal correlation of the data. Li et al. [21] captured the spatiotemporal correlation of data by incorporating an adaptive learning matrix into GCN.

As traffic data's spatial and temporal correlations can be better captured, researchers have begun to fuse the two for traffic safety prediction. Yu et al. [22] proposed a spatiotemporal graph convolutional network (STGCN) that effectively captures comprehensive spatiotemporal correlations by modeling multiscale traffic networks. Similarly, the spatiotemporal attention-based graph convolutional network (ASTGCN) was proposed by Guo et al. [23]. The model captures the spatial correlation among different locations through spatial attention and the temporal correlation among different times through temporal attention. The model performs better in traffic prediction and improves the accuracy of traffic safety assessment. Wang et al. [37] proposed a traffic gate graph neural network (traffic-GGCN) for the real-time fusion of spatiotemporal representation modeling and applied GGRU-based modules to explore and aggregate spatial interactions to extract temporal correlations through real-time fusion.

A more detailed discussion about these actuators, their mathematical modeling, and incorporation into the governing equations of motion of the rotating flexible structure is not part of the scope of this work. For this work, the idea that adding piezoelectric actuators acting along the flexible structure adds external forces on the right side of Eq. (6) is sufficient.

Equation (5) can also be written as follows [7, 22]:

$$\ddot{\nu} + \left(r + x \right) \sin\alpha\ddot{\theta} + \cos\alpha\ddot{s} - \nu\sin^2\alpha\dot{\theta}^2 + \rho Ag\cos\alpha + \frac{EI}{\rho A}\nu^{iv} = q_{piezo}(x,t), \tag{6}$$

where $q_{piezo}(x, t)$in Eq. (6) is the force applied by the piezoelectric actuator to the beam. The external force $q_{piezo}(x, t)$ is also the control force to be applied along the flexible structure.

2.2. Capturing the influence of traffic accidents

None of the models above consider the influence on local traffic congestion from traffic accidents, i.e., capturing the diffuse influence between regions. In the last few years, several disciplines have studied the influence of traffic accidents extensively. These studies are mostly based on theoretical modeling and simulation and can be divided into three categories: deterministic queuing theory or shock wave theory [24], heuristics and simulations [25], and microscopic modeling of driver behavior [26]. However, the results of such models rely on theoretical simulations of road network traffic rather than on actual collected traffic data. Therefore, to solve this type of problem, most current literature treats traffic accidents as external factors. They are fused with traffic flow data through learnable parameters and used as inputs to the model for prediction. For example, Yu et al. [13] proposed using traffic incidents as disruption signals. They used a stacked autoencoder to extract potential features. Historical traffic flows are input to a deep neural model composed of LSTM-based units for time-dependent modeling. The final combination of the two completes predicting traffic flow in abnormal situations (traffic accidents). However, traffic accidents have varying degrees of influence on the road where the accident occurred and other roads; i.e., the influence of traffic accidents is inconsistent. Therefore, the combination of traffic accidents with spatial information is beginning to emerge to further capture the influence generated by traffic accidents. For example, Liu et al. [3] proposed to detect anomalies (traffic accidents) in complex traffic environments by using location coding. In that article, it was pointed out that traffic accidents greatly influence traffic congestion in their area, so centralized processing can improve prediction performance. The quantitative features of traffic accidents can be extracted, and the potential of traffic congestion can be represented using location coding. Fukuda et al. [27] proposed adding traffic accident features to the input data and constructing features of whether a traffic accident occurred as features for locating sensors while using a model constructed based on graphical convolutional networks to predict the traffic flow under the influence of accidents. Liu et al. [28] redefined the information on traffic accidents and combined it with traffic flow data to capture the impact of traffic accidents on traffic flow while solving the problem of traffic accidents having a smaller scale than traffic flow data. Moreover, Pan et al. [29] proposed quantifying the spread of traffic accident influences, i.e., providing the spatial (affected area) and temporal (traffic flow reduction) aspects as the final prediction results. Inspired by this study, this study proposes constructing a matrix of influence coefficients to describe the degree of mutual influence between any two regions to quantify the spread of traffic accident influence. Meanwhile, this work proposes a hybrid model that integrates a graphical convolutional network and a temporal convolutional network. This also achieves the extraction of the spatiotemporal correlation of traffic flow data. GCNs have shown better performance in various traffic applications.

3. DATA DESIGN

3.1 Problem definition

As typical spatiotemporal data, the traffic flow data are correlated in temporal and spatial dimensions. Therefore, when predicting traffic flow, traffic flow data of historical temporal and traffic flow data of its adjacent road nodes should be considered. In short, the traffic flow prediction problem is a time series prediction problem considering spatial characteristics. Traffic flow is used as a predictor to predict the future traffic flow using the traffic flow at historical time. That is, the traffic flow in the first d periods of *n* nodes $[X_{t-d}, \dots, X_{t-2}, X_{t-1}]$ is used to predict the traffic flow in the future *q* periods, as shown in equation (1):

$(Y_{t-q-1}, \dots, Y_{t-1}, Y_t) = M((\dots\dots, X_{t-2}, X_{t-1}), G)$, (1)

where $X_t = (x_t^1, \dots, \dots, x_t^n) \in R^n$ and x_t^n represent the traffic flow of the *n*-th node at time *t*. Likewise, $Y_t = (y_t^1, \dots, \dots, y_t^n) \in R^n$, $\mathrm{y_t^n}$ represents the traffic flow at the *n*-th node for time *t*. *M* represents the modeling method.

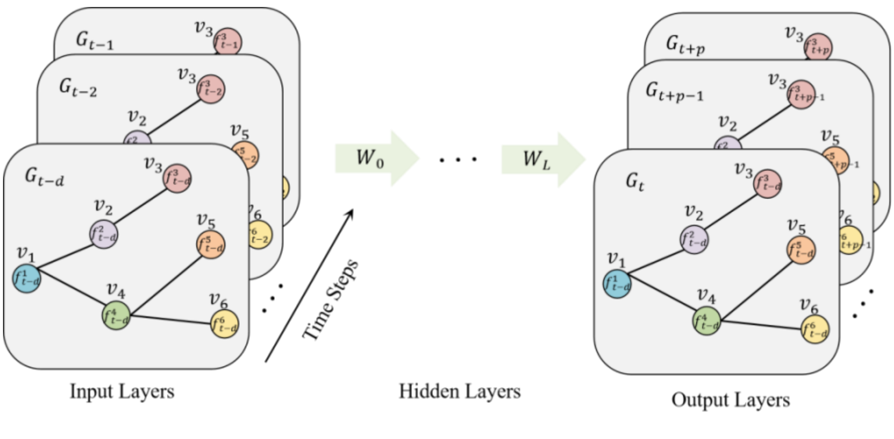


Fig. 2. *d* input channels obtain the traffic flow at time *p* through the mapping function *M*.

As shown in **Fig. 2**, in this work, the graph structure G is used to represent the traffic graph. $G = (V, E)$ is a nonweighted matrix representing the spatial dependencies between traffic roads. Among them, node V of the graph represents the collection of roads. V_i represents the i-th road and $V = \{\dots, v_2, \dots, v_n\}$. Edge E represents the set of intersections, reflecting the connections between roads. This work uses an adjacency matrix A to store the connections between traffic roads. The definition of A is shown in (2):

$$A_{ij} = \begin{cases} 1, & e_{ij} \in E \\ 0, & e_{ij} \notin E' \end{cases} \quad (2)$$

where e_{ij} represents the connection between v_i and v_j. If e_{ij}=1, there is a connection between v_i and v_j. If e_{ij}=0, there is no connection between v_i and v_j.

This work constructs a matrix $B \in R^{T \times n}$ based on the traffic accident dataset to represent the occurrence of traffic accidents on roads, where T denotes the historical time duration and n is the number of road nodes. b_i^t represents whether a traffic accident occurred on road i at time t. b_i^t= 0 means no traffic accident occurred on road i at time t; b_i^t= 1 means a traffic accident occurred on road i at time t. Based on this, a dimensional 01 matrix is constituted.

3.2 Model design

The model structure proposed in this article is shown in Fig. 3, which mainly consists of the temporal convolution layer and the spatial convolution layer. Among them, the temporal convolution layer adopts a temporal convolution network to extract the temporal correlation of data. In contrast, the spatial convolution layer uses a graph convolution network to extract the spatial correlation of data. The proposed traffic accident impact coefficient matrix is embedded in the input layer and traffic flow is the final input of the model. This section mainly provides a detailed introduction to the impact coefficient matrix and model.

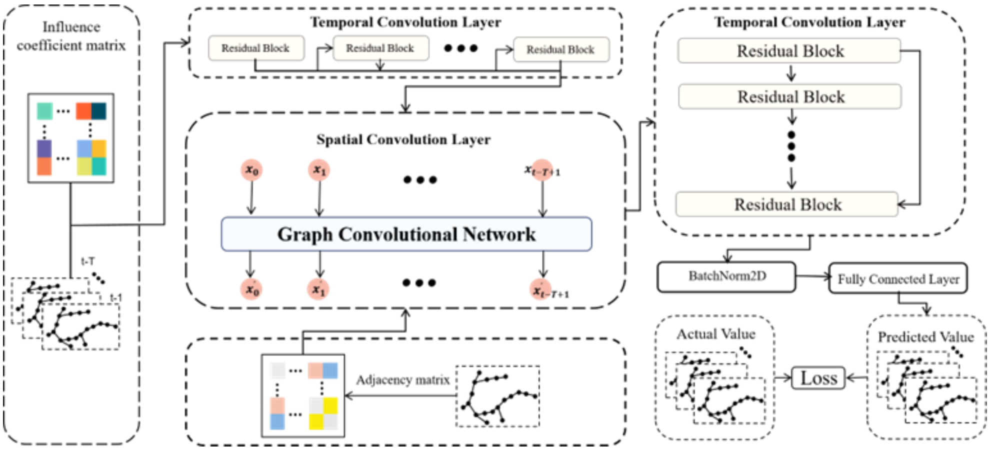


Fig. 3. The structure of the model.

3.2.1 Traffic accident influence coefficient matrix

For traffic flow prediction, not only the spatiotemporal correlation of data but also the influence of traffic accidents on traffic flow should be considered. Traffic accidents are different from other external factors, such as weather, which have the same impact on any road. However, the impact of traffic accidents on different roads varies depending on their time and space. In addition, the degree of influence between roads is related not only to distance and connectivity but also to the relationship between traffic modes. The impact between roads with the same traffic mode will be greater. This is mainly because roads in the same traffic mode often have similar behavioral characteristics and patterns, so their interaction may be more prominent. As shown in **Fig. 4**, when a traffic accident occurs on a road, traffic participants are more likely to choose a road with the same traffic mode as an alternative path, resulting in an increase in traffic flow on the road and a smaller change in traffic flow on the road. Based on the above reasons, this work proposes constructing a traffic accident influence coefficient matrix to capture the influence diffusion between nodes.

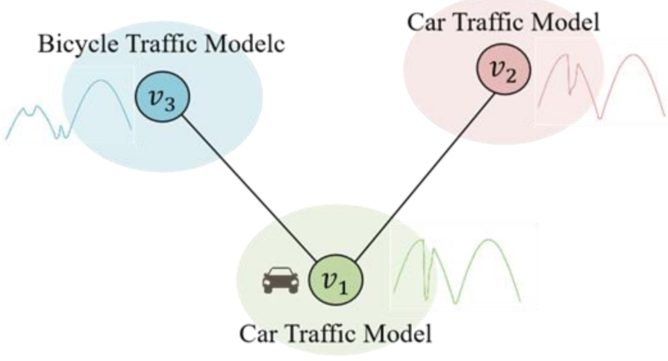


Fig. 4. The influence of traffic patterns on traffic flow.

The calculation process of the influence coefficient matrix will be described in detail below. Since the changing trend of traffic flow can well reflect the traffic pattern in the area, this work uses the traffic flow of the road as a measure of the similarity of the traffic patterns of any two roads. Likewise, the degree of influence between roads is also influenced by distance. When the distance between the two roads is shorter, the degree of influence is more significant. Conversely, the influence is reduced when the two roads are farther apart.

Therefore, this work uses the Pearson correlation coefficient to calculate the similarity of traffic patterns between two roads and then divides it by the distance between the two, making it positively correlated with the similarity of traffic patterns and negatively correlated with distance. Based on this, the impact coefficient of traffic accidents, i.e., the degree of impact between two points, is obtained. The specific calculation formula is as follows:

$$cor = \frac{p_{ij}}{d_{ij}}, \quad (3)$$

where p_{ij} is the similarity degree of traffic patterns between node *i* and node *j* and the specific calculation is (4); d_{ij} is the distance between node *i* and node *j*; x_i represents the traffic flow data of node *i*.

$$p_{ij} = \frac{z\sum x_i x_j - \sum x_i \sum x_j}{\sqrt{z\sum (x_i)^2 - (\sum x_i)^2}\sqrt{z\sum (x_j)^2 - (\sum x_j)^2}}, \quad (4)$$

The *B* matrix (01 matrix) constructed in Section 3.1 is multiplied by the traffic accident influence coefficient matrix to obtain the degree of influence of each road node by the accident at each moment. This matrix is denoted as $C \in R^{T\times n}$. Specifically, it is shown in **Fig. 5**, where $c_{v_2}^{t_2}$ =1 indicates that the road has a traffic accident at the moment and $c_{v_2}^{t_2}$ =0.7 is denoted that the road is affected by a road traffic accident to the extent of 0.7 at the moment.

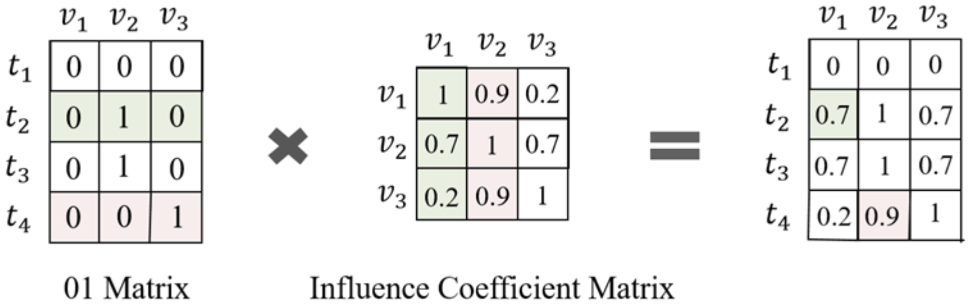


Fig. 5. Schematic diagram of the 01 matrix multiplied by the influence coefficient matrix.

3.2.2 Spatial dependency modeling

Traffic flow changes as the traffic road topology shifts. Commonly used CNN modeling methods can extract the spatiotemporal features of data. However, they are only applicable to Euclidean data structures and cannot be used to capture the spatiotemporal correlation of complex data. Graph convolutional networks make up for this shortcoming. The model can handle non-Euclidean data well and has achieved good results in image classification and document analysis [30-31]. Because of the complex spatiotemporal correlation of traffic flow data, this work uses a graph convolution model to capture the spatial correlation of traffic flows. The GCN model structure is shown in Fig. 6.

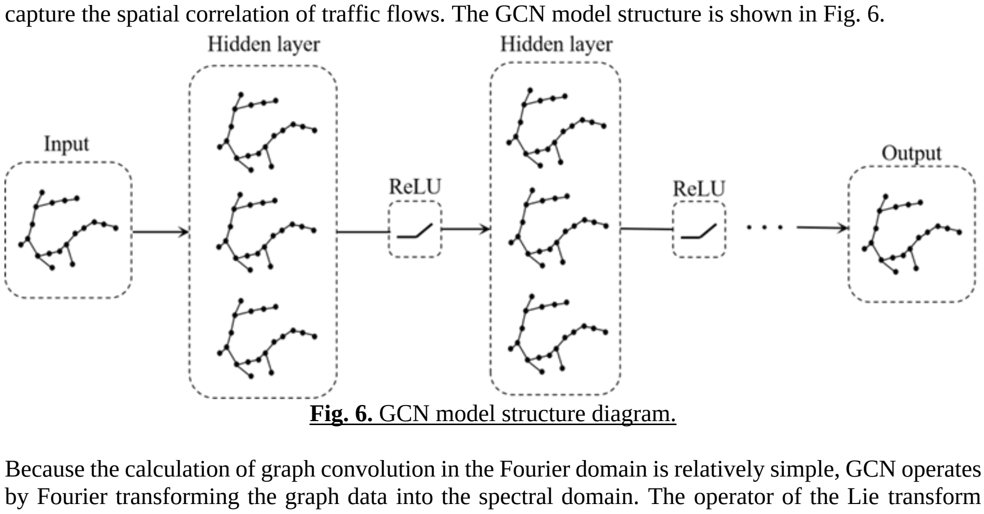


Fig. 6. GCN model structure diagram.

Because the calculation of graph convolution in the Fourier domain is relatively simple, GCN operates by Fourier transforming the graph data into the spectral domain. The operator of the Lie transform becomes the feature vector corresponding to the graph, which is also the core idea of applying the Fourier transform to the graph [32-33]. The definition of the Laplace matrix is shown in (5). To prevent the functions of the model from degrading during the training process due to different scales, the eigenvalues of the Laplacian matrix are normalized in this work to obtain the normalized Laplacian matrix. Its calculation formula is shown in (6):

$$L = D - A, \quad (5)$$

$$L^{sys} = D^{-\frac{1}{2}} L D^{-\frac{1}{2}}, \quad (6)$$

where L represents the Laplacian matrix, A is the adjacency matrix of the graph, and D is the degree matrix of the node.

Some follow-up work has been done by reducing the computational complexity from O(n^2) to linear, including using Chebyshev polynomials to approximate the convolution kernel and by stacking multiple local graph convolution layers with first-order approximations of Laplacian graphs. Among them, the Chebyshev polynomial greatly reduces the computational cost by avoiding the eigenvalue decomposition required to find the Fourier basis. The relevant calculation is shown in (7), where β_k represents the Chebyshev polynomial coefficient, the specific calculation process of $\widetilde{L^{sys}}$ is (7), I represents the identity matrix, and λ_{max} represents the largest eigenvalue of the Laplace matrix. The specific process of $T_k(\widetilde{L^{sys}})$ is (8). Based on (6), the parameters of the convolution kernel are reduced to k, with k being a constant. Equation (9) shows the recursive definition of the Chebyshev polynomial [34]:

$$g_\theta(L^{sys}) = \sum_{k=1}^{k} \beta_k T_k(\widetilde{L^{sys}}), \quad (7)$$

$$\widetilde{L^{sys}} = \frac{2}{\lambda_{max}} L^{sys} - I, \quad (8)$$

$$\begin{cases} T_0(\widetilde{L^{sys}}) = 1 \\ T_1(\widetilde{L^{sys}}) = \widetilde{L^{sys}} \\ T_k(\widetilde{L^{sys}}) = 2\widetilde{L^{sys}} \cdot T_{k-1}(\widetilde{L^{sys}}) - T_{k-2}(\widetilde{L^{sys}}) \end{cases}. \quad (9)$$

Although traffic accidents can affect traffic flow significantly, due to the small number of traffic accident samples, it is very easy to treat the phenomenon of a sudden drop in traffic flow due to traffic accidents as abnormal data during the training process. Therefore, this work adds the matrix C mentioned in Section 3.2.1 to the traffic flow data, which can be applied to increase the importance of such data and to better catch the influence of traffic accidents on traffic flow. The influence of traffic accidents on traffic flow is better captured. After this operation, the input data are expanded to four dimensions. The data are then dimensionally transformed and taken as new inputs into the model.

Over time, traffic flow at a historical time can have different effects on future traffic flow. In short, traffic flow data has an obvious time correlation: time series data. One of the most desirable properties of a predictive model is that the output value depends on all the input data. This can be achieved by continuously deepening the network as the input sequence length grows. But this also causes the problem that the model has many parameters and takes longer to train. Moreover, the problem of the gradient disappearing as the number of model layers increases is easy to occur. For the above reasons, this work adopts a temporal convolutional network (TCN) to extract temporal features [35].

TCN can extract temporal features of traffic flow across time steps. Therefore, this work uses TCN as the temporal data correlation capture model. TCN has three main modules: causal convolution, dilated convolution, and residual modules.

Causal convolution refers to the convolution operation that considers only the first half; i.e., only the inputs before the prediction time step are convolved. Since all the data in causal convolution strictly obey temporal causality, the value of time t for each layer of the model in this experiment depends only on the information before time t (i.e., [0, t]) and is therefore adapted to handle time series data. The expression of the formula is shown in (10):

$$Y_t = \sum_{k=1}^{k} f_i \cdot X_{t-k+1}. \quad (10)$$

Dilated convolution permits the input data to be sampled in intervals according to the sample rate during the convolution process, and the sampling rate is controlled by a dilation factor d, as shown in **Fig. 7**.

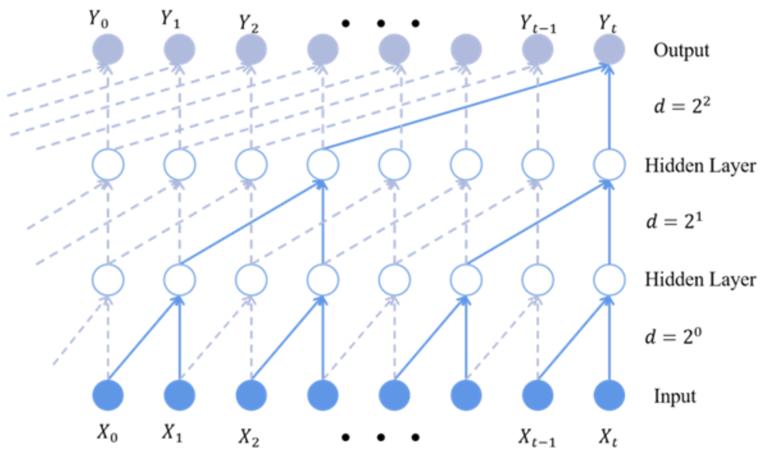


Fig. 7. Diagram of dilated convolution.

Generally speaking, the higher the sampling rate of the layer, the larger the value. That is, the dilation convolution enables the valid window size to grow exponentially as the number of layers increases, thus obtaining a larger field of sensation with fewer layers. The calculation formula is shown in (11):

$$Y_t = \sum_{i=0}^{k} f_i \cdot X_{t-i \cdot d}. \quad (11)$$

The deepening of the network makes the training process very complicated. This problem can be well solved by using residual structures instead of traditional structures, as shown in **Fig. 8.**

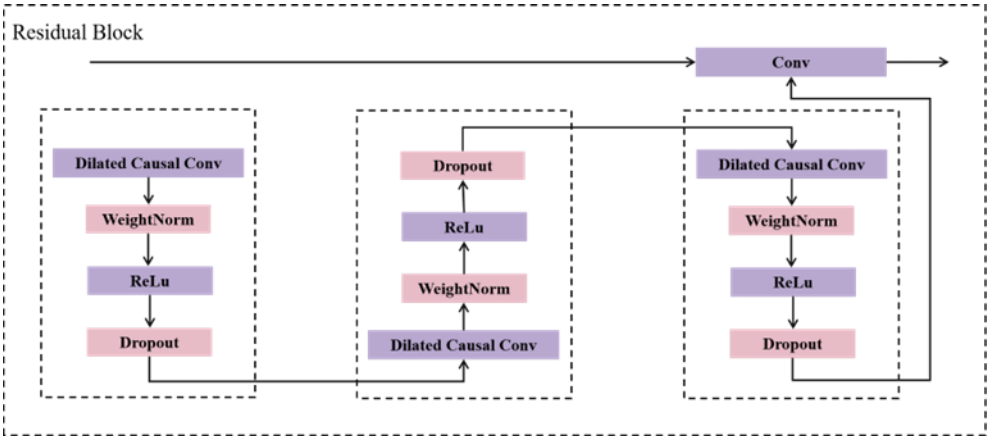


Fig. 8. Residual module structure diagram.

Let X be the input value of the residual module, the identity mapping function across layers is $F(\cdot)$, and the result will add the input value X. Hence, the calculation formula of the output value of the residual module is as follows:

$$o = \mathrm{Tanh}(X + \mathrm{F}(X)). \qquad (12)$$

4. EXPERIMENT

4.1 Dataset introduction

This work uses a real dataset from California, USA, with a longitude range of [37.928, 38.0124] and a dimension range of [-121.941, -121.750]. The work selects 19 roads as the most trained nodes—55% of the dataset as the training set and 45% as the testing set are used. Data on traffic flow from May 31 until June 30, 2021 (30 days in total), were chosen as the training set, and the sampled raw data were collected at five-minute intervals. That is, a sensor collects 288 pieces of data daily, totaling 8640 five-minute periods. To verify the performance of the model, the test set uses 25 days of traffic flow data from June 1 to June 25, 2021. The testing set data consisted of 7200 five-minute time slots. This work mainly uses linear interpolation to fill in missing values and standardize the data. To verify the stability and robustness of the GTCN model, this work selects all 29 roads in the dataset and abstracts them as nodes, constructing a topological structure diagram based on longitude and latitude. GTCN achieved good results on both data, verifying the effectiveness and stability of the model.

4.2 Experimental setup

All experiments were conducted and tested on Windows 10(CPU: Intel(R) W-2133 CPU@3.60GHz; GPU: NVIDIA GeForce RTX 2080 Ti). The input data used in this work is 12 five-minute periods, and the output data is 3 five-minute periods. That is, the traffic flow in the next 15 minutes is predicted based on the data of the past 1 hour. In the spatial convolution layer, the convolution kernel is approximated using Chebyshev’s first-order polynomial to reduce the difficulty of model training. In the temporal convolutional layer, the initial dilation rate of TCN is set to 2, and the convolution kernel is set the same for 3 and 5 layers of TCN. In the training part of the model, the experiment batch size is set to 32 and the learning rate is set to 0.001. The optimizer uses Adam to train the model and the number of iterations is set to 1000. The training algorithm employed in this experiment is depicted in Algorithm 1.

Algorithm 1: Model Training

N=19（The number of node）,*S*=12（The number of input）

Input: Initialize Adjacency Matrix A $\in R^{N\times N}$, Influence Coefficient Matrix $cor \in R^{N\times N}$, 01 Matrix $B \in R^{T\times N}$

1: for v<N do
2: $p_v \leftarrow [p_{v1}, p_{v2}, \dots p_{vN}]$
3: $D_v \leftarrow [d_{v1}, d_{v2}, \dots d_{vN}]$
4: $Cor_v \leftarrow [\frac{p_{v1}}{d_{v1}}, \frac{p_{v1}}{d_{v1}}, \dots, \frac{p_{v1}}{d_{v1}}]$
5: $\mathrm{v} \leftarrow \mathrm{v} + 1$
6: end for
7: C← B · Cor
8: Construct the model as shown in Fig. 2
9: repeat
10: Input a batch of data into the network
11: $X_t' \leftarrow TCN(X_t, C_t)$
12: $X_t'' \leftarrow GCN(X_t', A)$
13: $Y_{pred} \leftarrow FCN(Batch\ Norm2D(TCN(X_t'')))$
14: Update network parameters and use Adam algorithm to minimize L2 loss between y_{pred} and y
15: Until Reached the designated number of rounds
16: Return to training model

4.3 Baselines

To prove the performance of the model, this work selects the following seven baselines for comparison. The specific introduction of the model is as follows:

1) HA: This model is a traditional statistical model, which is an averaging of flows over the same period in the historical data, and uses the calculated result as the predicted value.

2) ARIMA: The model predicts traffic flow by analyzing the relationship between historical traffic flow data and current flow data.

3) GCN: The model is a graph convolution network (GNN) using convolution operations, which can catch the spatial correlation of traffic flow data well. The main goal of GCN is to extract the spatial features of topological graphs. This work sets a layer of GCN for prediction, with input and output channels set to 1 and 16, and convolutional kernel size set to 1×3. GCN updates the learnable parameter matrix W.

4) LSTM: The model handles the transfer state of the data by a gating mechanism that comprises an input gate, an output gate, and a forget gate.

5) STGCN: The model consists of GCN and TCN, which can not only extract the spatial features of traffic flow data but also capture the most basic temporal features consistently. This work uses two spatiotemporal convolution modules, with 64 channels in each module. The kernel size of graph convolution and time convolution is set to 1×3.

6) ASTGCN: The model uses an attention mechanism for extracting temporal features and a graph convolution model for extracting spatial features. Compared to STGCN, the model can capture dynamic spatiotemporal correlations in traffic flow data. Among them, the number of channels for GCN and TCN in each submodule is set to 1 and 64, GCN is approximated using first-order Chebyshev polynomials, and the convolution kernel size for time convolution layers is 1×3.

7) GCNNM: Graph Convolutional Neural Network. The model consists of a graph convolution model and an attention encoder model. Among them, the graph convolution model is taken to capture the spatial correlation of traffic flow data, the convolution kernel size is 1×3, and the first-order Chebyshev approximation is used. The attention encoder model is taken to capture the temporal correlation of the data; the convolutional kernel size is also 1×3. The model also considers the impact of traffic accidents and embeds representations in the input conversion layer [36].

4.4 Evaluation indicators

In terms of evaluating model performance, this work uses three metrics used to evaluate the traffic flow prediction function. They are mean absolute error (MAE), mean absolute percentage error (MAPE), root mean square error (RMSE), and symmetric mean absolute percentage error (SMAPE), and the specific calculation formulas are as follows:

$$\mathrm{MAE} = \frac{1}{NT_p}\sum_{j=1}^{N}\sum_{i=1}^{T_p}|\hat{Y}_i^j - Y_i^j|, \quad (13)$$

$$\mathrm{MAPE} = \frac{1}{NT_p}\sum_{j=1}^{N}\sum_{i=1}^{T_p}\frac{|\hat{Y}_i^j - Y_i^j|}{Y_i^j}, \quad (14)$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{NT_p}\sum_{j=1}^{N}\sum_{i=1}^{T_p}(\hat{Y}_i^j - Y_i^j)}, \quad (15)$$

$$\mathrm{MAPE} = \frac{1}{NT_p}\sum_{j=1}^{N}\sum_{i=1}^{T_p}\frac{|\hat{Y}_i^j - Y_i^j|}{\left(|\hat{Y}_i^j| - |Y_i^j|\right)/2}, \quad (16)$$

where $\hat{Y}_i^j$ and Y_i^j represent the predicted flow and actual flow of the *j*-th node at the time *i*, respectively. *N* stands for the number of nodes. T_p stands for the length of the time window for prediction. SMAPE is an improvement of the MAPE indicator, which avoids the problem of a small percentage difference in predicted values due to a small actual value.

4.5 Quantitative analysis

Table 1 shows how our model and the baseline model function to predict traffic flow within 15 minutes. Comparing the two, the deep learning model proves more effective than employing traditional statistical models for time series data. Deep learning models can catch the temporal features inherent in traffic flow data very well. In the deep learning models, models that take temporal correlation and traffic topology into account (e.g., STGCN, ASTGCN, GCNNM, and our model) perform obviously better than the traditional deep learning models (LSTM, GCN). Our model takes longer to run than the traditional model, mainly due to the fused and complex nature of the model. The increased complexity of time also improves the accuracy of predictions. Also, since this work catches the influence of traffic accidents on other roads and constructs a traffic influence coefficient matrix, the influence of traffic accidents on traffic flow can be better caught. Our model performs optimally for predictions dealing with traffic flow data that includes the influence of traffic accidents.

The evaluation of the model should not only consider the performance of the model but also consider the computational cost, as shown in **Table 2**. Taking the experimental dataset as an example, the running time of GTCN per round is 0.519 seconds, second only to GCN and STGCN. Although GCN, STGCN, and ASTGCN are all models composed of convolutional networks, ASTGCN is a multichannel composite network composed of three periods, and the calculations of the three periods are not parallel, resulting in a longer computational time. Similarly, LSTM time series models and traditional statistical models, such as HA and ARIMA, do not have the ability to perform parallel computing, resulting in longer runtime. In summary, GTCN can achieve parallel computing and effectively reduce running time.

Table 1. Quantitative analysis results.

Model	PeMS19				PeMS29			
	RMSE	MAE	MAPE	SMAPE	RMSE	MAE	MAPE	SMAPE
HA	**44.93**	**30.14**	**1.34**	**1.455**	**40.25**	**29.63**	**1.20**	**1.34**
ARIMA	**27.19**	**43.27**	**0.86**	**0.981**	**22.85**	**41.33**	**0.78**	**0.82**
GCN	**18.83**	**12.72**	**0.59**	**0.670**	**16.68**	**11.69**	**0.54**	**0.61**
LSTM	**38.151**	**26.84**	**0.251**	**2.661**	**33.34**	**24.68**	**0.22**	**0.24**
STGCN	**27.91**	**17.44**	**1.122**	**1.198**	**22.49**	**16.88**	**1.02**	**1.21**
ASTGCN	**29.196**	**19.57**	**1.662**	**1.775**	**27.64**	**17.89**	**1.44**	**1.43**
GCNNM	**38.85**	**28.32**	**1.875**	**1.977**	**32.58**	**27.65**	**1.64**	**1.71**
Withou Cor-GTCN	**28.0079**	**18.933**	**0.268**	**0.271**	**22.35**	**17.84**	**0.22**	**0.23**
Without TCN-GTCN	**19.72**	**13.54**	**0.28**	**0.3001**	**18.54**	**12.21**	**0.24**	**0.247**
Without GCN-GTCN	**19.13**	**12.88**	**0.24**	**0.2574**	**18.23**	**11.69**	**0.20**	**0.214**
Our model	**18.724**	**12.233**	**0.2009**	**0.2249**	**16.45**	**11.02**	**0.198**	**0.207**

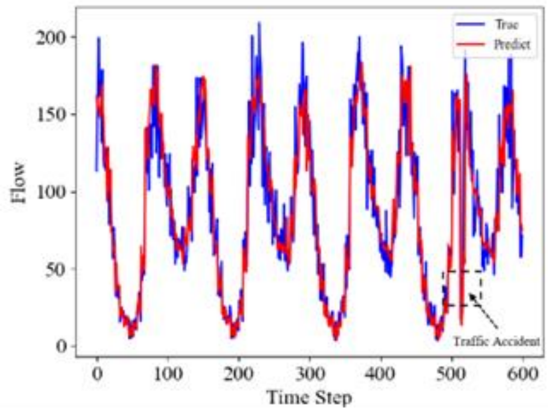
Table 2. Model runtime comparison.

Model	HA	ARIMA	GCN	LSTM	STGCN	ASTGCN	GTCN
Time	**1.02**	**1.24**	**0.418**	**4.103**	**0.346**	**2.532**	**0.519**

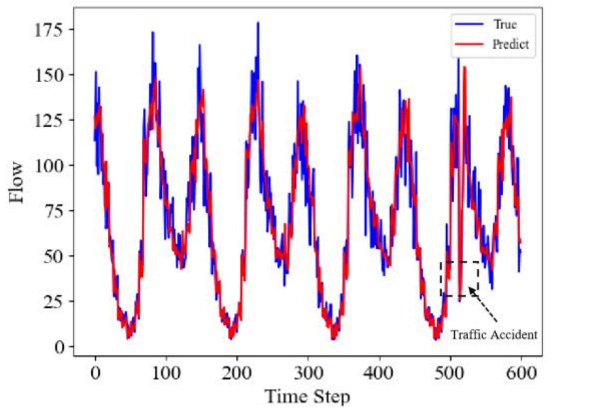
4.6 Qualitative analysis

To prove the superiority of our model, several types of baselines are set up in this experiment to compare with it, namely, traditional statistical models and deep learning models. The experimental results are plotted in the form of traffic flow data graphs, which reflect whether the trend of the model is consistent with the actual traffic flow data. Traditional statistical models mainly include HA and ARIMA, and deep learning models include GCN, LSTM, STGCN, ASTGCN, and GCNNM.

Fig. 9 shows a comparison chart of the traffic predicted by the model proposed in this work. Fig. 9(a) shows the influence of traffic accidents on the road on traffic flow. Fig. 9(b) shows the influence on the traffic flow of the road adjacent to the road where the traffic accident occurred. It can be seen that using the traffic accident influence coefficient matrix, the diffusion of the influence of the traffic accident is well captured to accurately predict the Changes in road traffic flow and how traffic flows on adjacent roads are affected.



(a) Traffic flow data on roads where traffic accidents occurred



(b) Traffic flow on nearby roads affected by traffic accidents

Fig. 9. Comparison chart of the predicted traffic flow by GTCN model.

Fig. 10 shows a plot of traffic predicted by other models versus the true value. Fig. 10(a) and Fig. 10(b) are the comparison charts of the traditional statistical model HA model and ARIMA model. It can be found that the value ratio predicted by such a model is smaller than the actual value. This is mainly because the structure of such models is relatively simple, cannot handle complex and nonlinear traffic flow data well, and cannot capture such phenomena as a sudden decrease in traffic flow caused by traffic accidents.

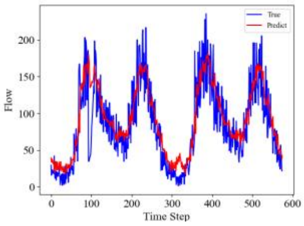
Fig. 10(c) and Fig. 10(d) are the comparison diagrams of the simple deep learning LSTM and GCN models. Observing the pictures, it can be found that the performance of this type of model has been improved to some extent compared with the traditional statistical model. However, there is still a deviation between the actual and predicted values, mainly because such models do not fully consider the spatiotemporal correlation of traffic flow data. LSTM is primarily used to process time series data and extract the temporal correlation of traffic flow but ignores its spatial correlation. Therefore, the sudden decrease in traffic flow on the road where the traffic accident occurred can be predicted, but the change in the traffic flow on the affected road cannot be well captured. Similarly, the forecast data of GCN is still small, its volatility is large, and some fluctuation trends are quite different from the fluctuation trend of the actual value. This is mainly because GCN is a graph convolutional network that can extract the spatial correlation of analytical data. Traffic flow data is greatly affected by spatial correlation. Since the data’s temporal correlation is not considered, there is a partial bias in the temporal correlation reflected in the fluctuations.

Fig. 10(e) and Fig. 10(f) show the prediction results of STGCN and ASTGCN. The predictions from the two models are about the same. The predicted value of this type of model is consistent with the real value, and there is a slight deviation locally, which is slightly smaller than the real value. Likewise, the capture of the influence of traffic incidents on traffic flow by such models has improved compared to previous models, but there is still a bias. There is a lag when the predicted value drops sharply compared to the actual value. This is mainly due to the insufficient capture of the influence of traffic accidents in this model. Fig. 10(g) shows the comparison between the prediction results of the GCNNM model and the real values. The model is more suitable for the predicted value of traffic flow and the real value, but the prediction performance of the influence of traffic accidents on traffic flow is poor; that is, the influence cannot be captured. The coefficient influence matrix proposed in this work can well represent the degree of influence between any two places. According to the matrix, the degree of influence between two areas at any time can be quickly captured. In a traffic accident, its influence can also be captured soon. As shown in Fig. 10(h), the predicted value of the model proposed in this work is very close to the actual value.

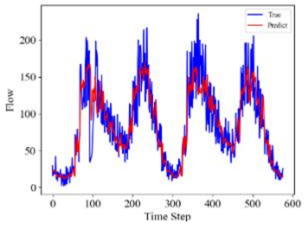
4.7 Ablation studies

The model presented in this work is formed by combining components with different functions. The effect of different components on the total function of the model and the function of each component can be determined through ablation experiments. In this experiment, an ablation experiment is set up for three parts in the model: the temporal convolution module, the spatial convolution module, and the influence coefficient matrix, and the experimental results are given. Fig. 11(a) shows the performance of these three ablation experiments compared to the original model.

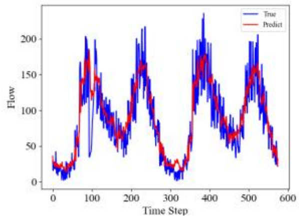
The experimental results without considering the influence coefficient matrix of traffic accidents are shown in Fig. 11(b). This work captures this influence by fusing the accident influence coefficient matrix with raw traffic flow data before inputting it into the model. The main purpose of removing this coefficient matrix is to demonstrate whether there is an influence on the performance of the overall model without capturing this effect. Observing Fig. 11(b), it is observed that the prediction results of the GTCN without coefficient matrix are relatively consistent with the actual values. Still, it cannot accurately predict the great changes in traffic flow caused by traffic accidents. This is mainly because the frequency of traffic accidents is small, and the number of road nodes affected by a traffic accident is also small compared with the number of road nodes under normal conditions. Therefore, in modeling, the data of sudden drops in traffic caused by traffic accidents may be regarded as abnormal data and not considered. Therefore, before inputting the original traffic flow data, the influence coefficient matrix is embedded, which can increase the model’s emphasis on this type of data to accurately capture this type of influence and improve the accuracy of traffic flow prediction. Experiments show that the influence coefficient matrix greatly influences the function of the whole model.



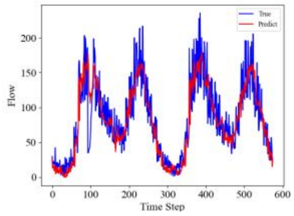
(a) The result predicted by HA



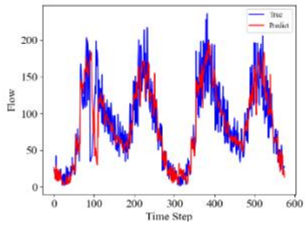
(b) The result predicted by ARIMA



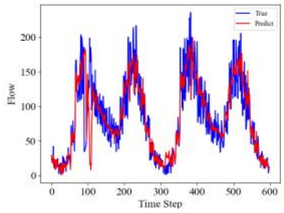
(c) The result predicted by LSTM



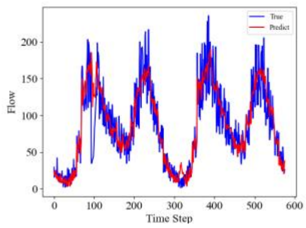
(d) The result predicted by GCN



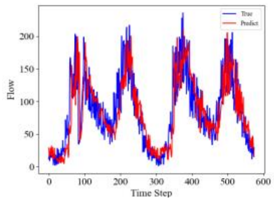
(e) The result predicted by STGCN



(f) The result predicted by ASTGCN



(g) The result predicted by GCNNM

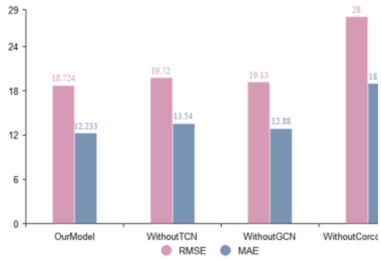


(h) The result predicted by the model proposed in this work

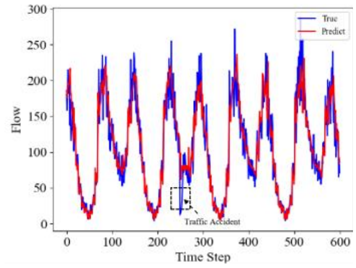
Fig. 10. Comparison between the predicted value of the baseline models and the actual.

The results after removing the temporal convolution module are shown in Fig. 11(c). The main purpose of removing the TCN module is to explore the influence of extracting temporal correlation on the whole functionality of the model. As can be shown in the figure, the GTCN model without the TCN module can still catch he influence of traffic accidents on traffic flow. However, due to the removal of the temporal convolution module, there is a phenomenon that the lag of the predicted impact generation time and the fluctuation trend of the traffic flow are not consistent enough. Experimental results show the importance of the temporal convolution module to the overall model.

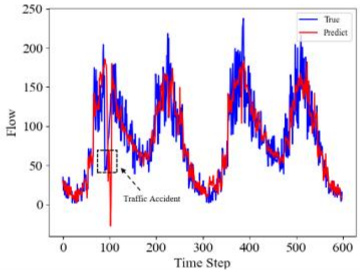
The results of removing the spatial convolution module GCN are shown in Fig. 11(d). The main purpose of removing the GCN module is to explore the influence of extracting spatial correlation on the whole functionality of the model. The figure shows that although the GTCN model without the GCN module can capture the influence on traffic accidents according to the influence coefficient matrix, it cannot predict the degree of influence well. This is mainly because after removing the graph convolution module, the model cannot aggregate the relevant information of neighboring nodes to the central node on time. That is, the spatial correlation cannot be sufficiently extracted. Therefore, the overall traffic flow has a situation where the predicted value is smaller than the real value. The experimental results show that the spatial convolution module is important to the overall model performance.



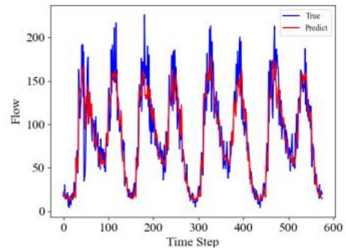
（a） Ablation experiment results



（b） Ablation experiments to remove the influence of the coefficient matrix



（c） Ablation experiment results with removal of temporal convolution module



（d） Ablation experiment results with removal of spatial convolution module

Fig. 11. Comparison of ablation experiment results.

5. CONCLUSION AND FUTURE WORK

To better catch the influence of traffic accidents on traffic flow and accurately predict the traffic flow on the road where the accident occurred and the nearby roads, this work proposes to construct an influence coefficient matrix to represent the degree of influence between any two road nodes and embed this matrix into the traffic flow data to form a new model input. After fully analyzing the influencing factors of traffic accidents, this work proposes a hybrid model combining graph convolutional network and spatiotemporal convolutional network (GTCN) for mining the spatiotemporal correlation of traffic flow data. The results show that the prediction results of this model are closer to the real values and outperform other models.

Future research focuses on the following aspects: introducing other external factors affecting traffic flow into the model, such as weather factors and holiday factors, to further improve the prediction accuracy of the model.

REFERENCES

[1] Jabbarpour MR, Zarrabi H, Khokhar RH, Shamshirband S, Choo KKR. Applications of Computational Intelligence in Vehicle Traffic Congestion problem: a Survey. Soft Computing. 2018 Apr 1;22(7):2299–320. https://doi.org/10.1007/s00500-017-2492-z

[2] Jiang W, Xiao Y, Liu Y, Liu Q, Li Z. Bi-GRCN: A Spatio-Temporal Traffic Flow Prediction Model Based on Graph Neural Network. Dong J, editor. Journal of Advanced Transportation. 2022 Feb 1;2022:1–12. https://doi.org/10.1155/2022/5221362

[3] Liu Z, Zhang R, Wang C, Jiang H. Spatial-Temporal Conv-Sequence Learning With Accident Encoding for Traffic Flow Prediction. IEEE Transactions on Network Science and Engineering. 2022 May 1;9(3):1765–75. https://doi.org/10.1109/tnse.2022.3152983

[4] Pan Z, Zhang W, Liang Y, Zhang W, Yu Y, Zhang J, et al. Spatio-Temporal Meta Learning for Urban Traffic Prediction. IEEE Transactions on Knowledge and Data Engineering. 2022 Mar 1;34(3):1462–76. https://doi.org/10.1109/tkde.2020.2995855

[5] Li W, Wang X, Zhang Y, Wu Q. Traffic Flow Prediction over Muti-Sensor Data Correlation with Graph Convolution Network. Neurocomputing. 2021 Feb;427:50–63. https://doi.org/10.1016/j.neucom.2020.11.032

[6] Hamed MM, Al-Masaeid HR, Said ZMB. Short-Term Prediction of Traffic Volume in Urban Arterials. Journal of Transportation Engineering. 1995 May;121(3):249–54. https://doi.org/ 10.1061/(asce)0733-947x(1995)121:3(249)

[7] Yang X, Zou Y, Tang J, Liang J, Ijaz M. Evaluation of Short-Term Freeway Speed Prediction Based on Periodic Analysis Using Statistical Models and Machine Learning Models. Journal of Advanced Transportation. 2020 Jan 20;2020:1–16. https://doi.org/10.1155/2020/9628957

[8] Sun B, Cheng W, Goswami P, Bai G. Short-term traffic forecasting using self-adjusting k-nearest neighbours. IET Intelligent Transport Systems. 2018 Feb 1;12(1):41–8. https://doi.org/10.1049/iet-its.2016.0263

[9] Wang X, Zhang N, Zhang Y, Shi Z. Forecasting of Short-Term Metro Ridership with Support Vector Machine Online Model. Journal of Advanced Transportation. 2018 Jun 27;2018:1–13. https://doi.org/10.1155/2018/3189238

[10] Crosby H, Davis P, Jarvis SA. Spatially-Intensive Decision Tree Prediction of Traffic Flow across the Entire UK Road Network [Internet]. IEEE Xplore. 2016. p. 116–9. https://doi.org/10.1109/DS-RT.2016.19

[11] Huang W, Song G, Hong H, Xie K. Deep Architecture for Traffic Flow Prediction: Deep Belief Networks With Multitask Learning. IEEE Transactions on Intelligent Transportation Systems [Internet]. 2014 Oct 1;15(5):2191–201. https://doi.org/10.1109/TITS.2014.2311123

[12] Duives D, Wang G, Kim J. Forecasting Pedestrian Movements Using Recurrent Neural Networks: An Application of Crowd Monitoring Data. Sensors [Internet]. 2019 Jan 18;19(2):382. https://doi.org/10.3390/s19020382

[13] Yu R, Li Y, Shahabi C, Demiryurek U, Liu Y. Deep Learning: A Generic Approach for Extreme Condition Traffic Forecasting. Proceedings of the 2017 SIAM International Conference on Data Mining. 2017 Jun 9;777–85. https://doi.org/10.1137/1.9781611974973.87

[14] Zhang D, Kabuka MR. Combining weather condition data to predict traffic flow: a GRU-based deep learning approach. IET Intelligent Transport Systems. 2018 Sep 1;12(7):578–85. https://doi.org/ 10.1049/iet-its.2017.0313

[15] Yamashita R, Nishio M, Do RKG, Togashi K. Convolutional neural networks: an overview and application in radiology. Insights into Imaging. 2018 Jun 22;9(4):611–29. https://doi.org/ 10.1007/s13244-018-0639-9

[16] Zhang J, Zheng Y, Qi D. Deep Spatio-Temporal Residual Networks for Citywide Crowd Flows Prediction. Proceedings of the AAAI Conference on Artificial Intelligence. 2017 Feb 12;31(1). https://doi.org/10.1609/aaai.v31i1.10735

[17] Defferrard M, Bresson X, Vandergheynst P. Convolutional Neural Networks on Graphs with Fast Localized Spectral Filtering. Advances in Neural Information Processing Systems. 2017 Feb. https://doi.org/10.48550/arXiv.1606.09375

[18] Cao S, Wu L, Wu J, Wu D, Li Q. A spatio-temporal sequence-to-sequence network for traffic flow prediction. Information Sciences. 2022 Sep;610:185–203. https://doi.org/10.1016/j.ins.2022.07.125

[19] Bai J, Zhu J, Song Y, Zhao L, Hou Z, Du R, et al. A3T-GCN: Attention Temporal Graph Convolutional Network for Traffic Forecasting. ISPRS international journal of geo-information. 2021 Jul 15;10(7):485–5. https://doi.org/10.3390/ijgi10070485

[20] Peng H, Wang H, Du B, Bhuiyan MZA, Ma H, Liu J, et al. Spatial Temporal Incidence Dynamic Graph Neural Networks for Traffic Flow Forecasting. Information Sciences. 2020 Jan. https://doi.org/10.1016/j.ins.2020.01.043

[21] Yu B, Yin H, Zhu Z. Spatio-Temporal Graph Convolutional Networks: A Deep Learning Framework for Traffic Forecasting. Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence. 2018;3634–40. https://doi.org/10.24963/ijcai.2018/505

[22] Wang S, Cao J, S. Yu P. Deep Learning for Spatio-Temporal Data Mining: A Survey. IEEE Transactions on Knowledge and Data Engineering. 2022 Aug;34(8):3681–700. https://doi.org/10.1109/TKDE.2020.3025580

[23] Lawson TW, Lovell DJ, Daganzo CF. Using Input-Output Diagram To Determine Spatial and Temporal Extents of a Queue Upstream of a Bottleneck. Transportation Research Record: Journal of the Transportation Research Board. 1997 Jan;1572(1):140–7. https://doi.org/10.3141/1572-17

[24] Pal R, Sinha KC. Simulation Model for Evaluating and Improving Effectiveness of Freeway Service Patrol Programs. Journal of transportation engineering. 2002 Jul 1;128(4):355–65. https://doi.org/10.1061/(asce)0733-947x(2002)128:4(355)

[25] Wang Z, Murray-Tuite P. A Cellular Automata Approach to Estimate Incident-Related Travel Time on Interstate 66 in Near Real Time. 2010 Mar 1.

[26] Fukuda S, Uchida H, Fujii H, Yamada T. Short-term Prediction of Traffic Flow under Incident Conditions using Graph Convolutional RNN and Traffic Simulation. IET Intelligent Transport Systems. 2020 Apr 27. https://doi.org/10.1049/iet-its.2019.0778

[27] Liu Y, Wu C, Wen J, Xiao X, Chen Z. A Grey Convolutional Neural Network Model for Traffic Flow Prediction under Traffic Accidents. Neurocomputing. 2022 May. https://doi.org/10.1016/j.neucom.2022.05.072

[28] Pan B, Ugur Demiryurek, Shahabi C, Gupta C. Forecasting Spatiotemporal Impact of Traffic Incidents on Road Networks. 2013 Dec 1. https://doi.org/10.1109/icdm.2013.44

[29] Lei B, Zhang P, Suo Y, Li N. SAX-STGCN: Dynamic Spatio-Temporal Graph Convolutional Networks for Traffic Flow Prediction. IEEE Access. 2022;10:107022–31. https://doi.org/ 10.1109/access.2022.3211518

[30] Lv Z, Li J, Dong C, Li H, Xu Z. Deep learning in the COVID-19 epidemic: A deep model for urban traffic revitalization index. Data & Knowledge Engineering. 2021 Jul;101912. https://doi.org/10.1016/j.datak.2021.101912

[31] Zheng Y, Wang S, Dong C, Li W, Zheng W, Yu J. Urban road traffic flow prediction: A graph convolutional network embedded with wavelet decomposition and attention mechanism. Physica A: Statistical Mechanics and its Applications. 2022 Dec 1;608:128274–4. https://doi.org/ 10.1016/j.physa.2022.128274

[32] Wang Y, Lv Z, Sheng Z, Sun H, Zhao A. A deep spatio-temporal meta-learning model for urban traffic revitalization index prediction in the COVID-19 pandemic. Advanced Engineering Informatics. 2022 Jun;101678. https://doi.org/10.1016/j.aei.2022.101678

[33] Zhiqiang Lv, Wang X, Chen Z, Li J, Li H, Xu Z. A new approach to COVID-19 data mining: A deep spatial–temporal prediction model based on tree structure for traffic revitalization index. Data and Knowledge Engineering. 2023 Jul 1;146:102193–3. https://doi.org/10.1016/j.datak.2023.102193

[34] Gao H, Jia H, Yang L. An Improved CEEMDAN-FE-TCN Model for Highway Traffic Flow Prediction. Journal of Advanced Transportation. 2022 May 10;2022:1–20. https://doi.org/10.1155/2022/2265000

[35] Zhiqiang Lv, Li J, Dong C, Xu Z. DeepSTF: A Deep Spatial–Temporal Forecast Model of Taxi Flow. The Computer Journal. 2021 Nov 16. https://doi.org/10.1093/comjnl/bxab178

[36] Wang Y, Zheng J, Du Y, Huang C, Li P. Traffic-GGNN: Predicting Traffic Flow via Attentional Spatial-Temporal Gated Graph Neural Networks. 2022 Oct 1;23(10):18423–32. https://doi.org/10.1109/tits.2022.3168590