Short-term Traffic Flow Prediction With Residual Graph Attention Network

Xijun Zhang, Guangjie Yu, Jiyang Shang, Baoqi Zhang

Abstract—Traffic flow prediction has been essential for traffic management and road network planning. However, the complex urban road network and the strong spatial-temporal correlation of traffic flow data make this problem difficult. Existing prediction methods cannot fully utilize the spatial-temporal correlations in traffic flow data. Therefore, we propose a deep learning model called ResGAT-ABiGRU which combines Residual Network (ResNet), Graph Attention Network (GAT), Attention Mechanism, and the Bidirectional Gated Recurrent Unit (BiGRU). Firstly, GAT is used to capture the spatial correlations of traffic flow data, and then the time characteristics are extracted by Bidirectional GRU. Secondly, the ResNet module stacks multiple GAT layers and designs the attention mechanism to assign weights for different flow sequences to further capture spatial relations. Finally, we obtain the output through the fully connected layers. Validation traffic data from California, USA, is used for verification. The results show that the ResGAT-ABiGRU model proposed in this paper has higher prediction accuracy. Compared the model's performance with the Gated Recurrent Unit (GRU) baseline model, and the root means square error (RMSE) is reduced by 22.75%, and compared to the T-gcn model, the root mean square error is reduced by 3.29%.

Index Terms—Intelligent transportation, Residual neural network, Graph attention network, Bidirectional Gated Recurrent Unit

I. Introduction

With the rapid development of the economy and the improvement of people's living standards, the number of motor vehicles continues to increase. The large number of motor vehicles has brought serious traffic congestion to the city and posed great safety hazards to the traveling public. In order to alleviate traffic congestion and develop accurate travel routes for travelers, intelligent transportation systems have been increasingly emphasized and studied by scholars. The core role of intelligent transportation system is road network planning and traffic control, but the core of these two tasks is to achieve accurate traffic flow prediction. Traffic flow data is mainly collected from sensors at different times and sections, so the traffic flow data itself has temporal and spatial characteristics, and the prediction of traffic flow is essentially a process of temporal and spatial modeling. There are two common approaches to the urban traffic flow prediction problem:

1) The traditional statistical prediction models[3]. According to the periodicity of traffic state evolution, some nonparametric models, such as K-nearest neighbors[4] (KNNs), is used to predict traffic speed and traffic flow. Support vector machines[5] (SVM) improves the accuracy of predictions by capturing the complexity within the traffic system. Some researchers combine multiple models to improve performance and a single road segment's traffic flow prediction accuracy. Some researchers introduce time-series forecasting models into the traffic forecasting problem by correlating observed traffic state evolutions. Autoregressive integrated moving average (ARIMA) is a typical method that includes several essential traffic flow characteristics, such as internal correlation and its impact on the short-term future[6]. Wang et al.[7] propose a hybrid EMDW-LSSVM model to predict short-term traffic, decomposing and reconstructing original traffic flow into stable sub-sequences with enhanced predictability and reduced noise.

Although statistical methods are used in many studies, which effectively extract dynamic spatial-temporal correlations in traffic flow data. These methods are still limited to single road segment scenarios and simple data processing methods. It limites the further application of statistical methods in complex urban traffic flow prediction.

2) The prediction model based on neural networks. Zhang et al.[8] propose a ST-ResNet model based on a residual convolution unit to predict the urban pedestrian flow. Although the method extracts the spatial-temporal characteristics of traffic data, the input format is limited to standard gridded data. Therefore, it cannot apply to the complex urban road network with graph structure. Cao et al.[9] propose a short-term traffic flow prediction model based on LSTM, which considers the time series characteristics of traffic flow data. However, the actual traffic flow not only contains temporal factors but also spatial factors. Pan et al.[10] propose a model that includes GRU and SVR. Considering the time series characteristics of traffic flow, compared with LSTM, which has simpler structure. It improves the operation speed, and SVR enhances the prediction accuracy of the whole model. However, the latent spatial features are not considerd in the traffic flow data. Zhao et al.[11] propose a hybrid model combining GCN and GRU for traffic flow prediction, which uses GCN and GRU to extract spatial and temporal features respectively.

Manuscript received November 22, 2021; revised September 02, 2022. This work was supported in part by National Science Fund subsidized project(61461027,61762059), Gansu Provincial Science and Technology Program Funding Project(21ZD4GA028) and Gansu Province Higher Education Innovation Fund-funded Project(2021A-028)

X. J. Zhang is an associate professor of Lanzhou University of Technology, Lanzhou, China, 730050. (Corresponding author to provide phone: 15593106088; e-mail: zhangxijun198079@sina.com)

G. J. Yu is a postgraduate student of Lanzhou University of Technology, Lanzhou, China, 730050. (e-mail: 2607927419@qq.com).

J. Y. Shang is a postgraduate student of Lanzhou University of Technology, Lanzhou, China, 730050. (e-mail: 1546983280@qq.com).

B. Q. Zhang is a postgraduate student of Lanzhou University of Technology, Lanzhou, China, 730050. (e-mail: 2217437265@qq.com).

However, the lack of consideration of a deep GCN network results in incomplete spatial components extracted. Pan et al.[12] propose a new hybrid model (GNN-PSO) of grey system theory and neural networks with particle swarm optimization to salivate traffic congestion and improve the mobility of transportation, which can exploit sufficiently the characteristics of the grey system model requiring fewer data. Guo et al.[13] propose the attention-based spatial-temporal graph convolutional neural network (STSGCN) to predict traffic flow. The model considers the temporal correlations of traffic flow data from multiple perspectives and uses GCN to capture the spatial correlations but very shallow GCN layers. Zhang et al.[14] propose a model that combines GAT and TCN to predict traffic flow. GAT is used to obtain spatial dependencies, and TCN is used to get temporal dependencies. Moreover, a multi-head attention mechanism is introduced to obtain spatial-temporal correlation coefficients to improve the prediction accuracy. However, the model only uses a single-layer GAT network and cannot extract the deep spatial features of traffic data. Liu et al.[15] propose a novel end-to-end deep learning architecture that extracts spatial-temporal information and analyzes historical traffic flow data. Ruan et al.[16] propose a pre-trained nodal model with multiple fusion layers architecture which can utilize the connections and the time component of nodes to predict the value of the next node and capture spatial-temporal characteristics of data.

To solve the problems mentioned above, we propose a novel ResGAT-ABiGRU model. On one side, GAT with residual structure is used to extract spatial features of traffic flow data. Besides, attention-based GRU model is used to extract temporal features of traffic flow data. The main contributions of this paper are as follows:

1) The graphs are used to denote the complex road network topology to solve the problem that the input matrix of traditional grid traffic flow data is too sparse.

2) The residual operations are integrated into the GAT network to build a deeper network model to achieve the purpose of mining the latent space characteristics of traffic flow data. Meanwhile, the residual operation can avoid the degradation of the deep network model and the problem of vanishing gradients.

3) The bidirectional GRU model is used to extract the temporal features of traffic flow data and incorporate an attention mechanism to dynamically assign different weights to traffic flow data in different periods.

II. Problem description

A. Road network topology

Undirected graph data structure $G = (V, E)$ to model the road network and map the roads to the nodes in the graph. V denotes the set of nodes, where v_i denotes the i_{th} node and N is the total number of nodes; E denotes the set of edges. The original graph connection is shown in Fig. 1.

Because each node influences itself and the adjacent nodes, the new graph connection can be obtained after adding the self-loop, which is shown in Fig. 2.

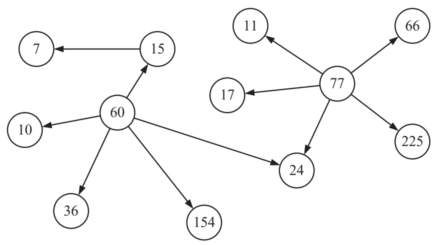

Fig. 1. The graph structures mapped by road network connections

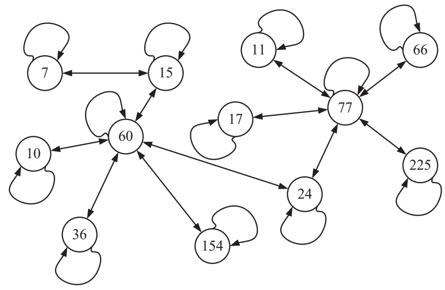

Fig. 2. Adding self-loop for each node and converting the directed graph into the undirected graph

B. Feature matrix

Feature matrix is denoted by $H \in \mathbb{R}^{N \times F}$. We consider the traffic flow of a road as the characteristic value of a node, where N denotes the number of nodes and F denotes the number of features.

C. Short time traffic flow prediction problem description

Traffic flow prediction is mainly to predict the traffic flow data at a specific time in the future by modeling and analyzing the historical traffic flow data of the road network. For example, we can predict the traffic flow data $f_s(t)$ of sensor s in the future t period using k historical traffic flow data $f_s(t-n)$ $(n = 1, 2, 3..., k)$.

We evaluate the performance of the proposed model by using the real-world dataset PEMS04. The sampling interval of the detector is $5min$. We divide the original data into several groups, each group contains six historical data and one label data. Data preprocessing is shown in Fig. 3.

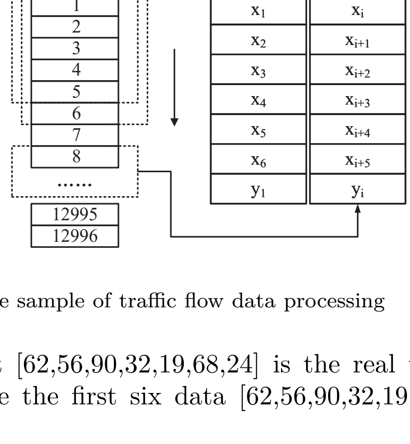

Fig. 3. The sample of traffic flow data processing

The list [62,56,90,32,19,68,24] is the real traffic flow data. Take the first six data [62,56,90,32,19,68] as the historical data and the seventh data as the label data. And the sliding window moves one unit at a time. Finally, the preprocessed data is divided into a training set and a testing set according to the proportion, as shown in Fig. 4.

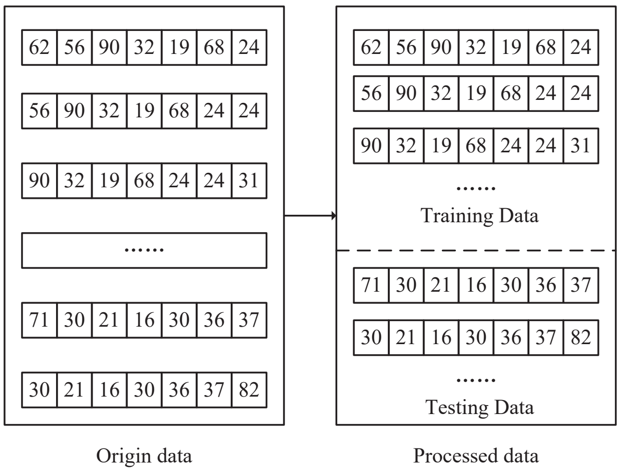

Fig. 4. The example of data processing and train-test dataset dividing

III. combined model

We propose the ResGAT-ABiGRU combined model to extract the spatial-temporal correlations in the traffic flow data. A graph neural network can effectively model the complex road network topology, capture the spatial correlation among traffic flow data, and assign different weights to neighbor nodes dynamically through the attention mechanism. In addition, the deep network model can mine the potential spatial correlations between the data through a residual connection, which can avoid the problems of the deep network model, such as model degradation and gradient disappearance. Otherwise, the traffic flow data is also typical time-series data, so we propose an attention-based GRU to capture the temporal features in traffic flow data. The overall structure of the model is shown in Fig. 5.

1) Firstly, we construct the model's input through the original data file, including data input matrix $H = \left\{h_1, h_2, h_3, ..., h_n\right\}$, $h_i \in \mathbb{R}^F$, and graph connection matrix A. where h_i denotes the feature vector of node i, and F denotes the dimension of the vector.

2) A new feature vector $H^\circ = \left\{h_1^\circ, h_2^\circ, h_3^\circ, ..., h_n^\circ\right\}$ can be obtained by linearly changing the input feature matrix H. The linear transformation operates as follows:

$$h_i^\circ = Wh_i \tag{1}$$

where $h_i^\circ \in \mathbb{R}^{F'}$, $W \in \mathbb{R}^{F \times F'}$ is the linear transformation matrix, F' denotes the dimension of the feature vector after linear transformation.

3) The correlation coefficient between nodes is calculated through the attention-based GAT Layer. The attention layer in GAT is shown in Fig. 6. The GAT Layer operation formula is as follows:

$$e_{ij} = LaekyReLU\left(a^T\left[Wh_i \,||\, Wh_j\right]\right) \tag{2}$$

where e_{ij} denotes the importance of node j to node i; a is the initialized learnable matrix; $LaekyReLu$ is used as the nonlinear activation function.

$$\alpha_{ij} = softmax_j\left(e_{ij}\right) = \frac{exp\left(e_{ij}\right)}{\sum_{k \in \mathbb{N}i} exp\left(e_{ik}\right)} \tag{3}$$

where $\mathbb{N}i$ is the first-order neighbor node of node i. In order to make the correlation coefficient easy to compare between different nodes, we use the $softmax$ function to normalize data.

$$h_i' = \sigma\left(\sum_{j \in Ni} \alpha_{ij} h_j\right) \tag{4}$$

4) We adopt a stacked multi-layer neural network to extract the deep spatial correlation between traffic flow data.

$$H'' = Batchnorm\left(GAT\left(H'\right)\right) \tag{5}$$

where $Batchnorm$ layer speeds up neural network training, accelerates convergence speed and increases model stability.

5) However, when the model depth increases, there will be a problem of model degradation. As the number of model layers deepens, we get bad results. Therefore, we use the residual connection in model building, and every two GAT operations constitute a residual module, as shown in Fig. 7. The formula for the residual operation is as follows:

$$H_{(l+1)} = Relu\left(H_l + \vartheta\left(H_l\right)\right) \tag{6}$$

where $H_{(l+1)}$ is the output of the residual structure; H_l is the direct mapping part; H_l is the residual operation part; $Relu$ is the nonlinear activation function. We use residual structure to mine the potential spatial relation and avoid the problem of model degradation. Since it is difficult to train the identity mapping function $H_{(l+1)} = H_l$, the function is converted into a residual function $H_{(l+1)} = H_l + \vartheta\left(H_l\right)$. When $\vartheta\left(H_l\right) = 0$, which can be converted into an identity mapping function, but it is simpler to train the residual function.

6) We extract the temporal features of the traffic flow data by GRU modul, which can be defined as follows:

$$z_t = \sigma\left(W_z \cdot \left[h_{t-1}, H_t\right]\right) \tag{7}$$

$$r_t = \sigma\left(W_r \cdot \left[h_{t-1}, H_t\right]\right) \tag{8}$$

$$\tilde{h_t} = tanh\left(W_{\tilde{h}} \cdot \left[r_t * h_{t-1}, H_t\right]\right) \tag{9}$$

$$h_t = \left(1 - z_t\right) * h_{t-1} + z_t * \tilde{h_t} \tag{10}$$

$$H_t' = \sigma\left(W_o \cdot h_t\right) \tag{11}$$

where H_t denotes the input of the current moment; h_{t-1} denotes the output of the hidden layer at the previous moment; $W_z, W_r, W_{\tilde{h}}$ and W_o denote the corresponding weight matrix; $*$ denotes the multiplication of matrix elements; H_t' denotes the output of the GRU.

Due to the forgetfulness of the recurrent neural network, the information contained in the last state of the model is incomplete. Therefore, two GRU models are trained simultaneously, one is forward training, and the

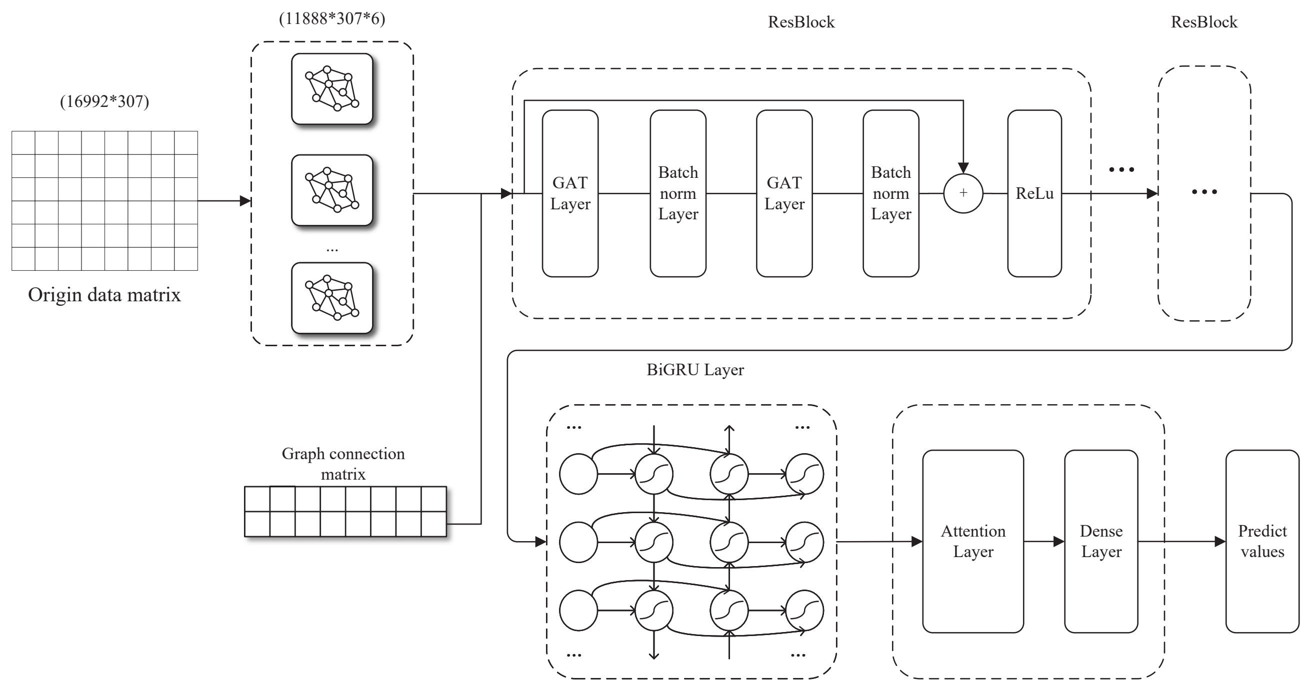

Fig. 5. The proposed deep learning based model for short-term traffic flow prediction

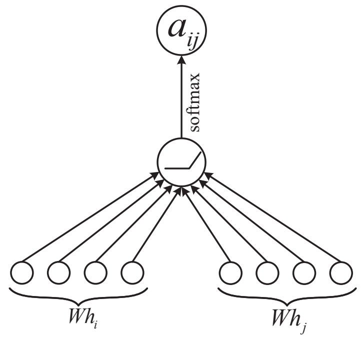

Fig. 6. The correlation coefficient of two groups of node features in GAT

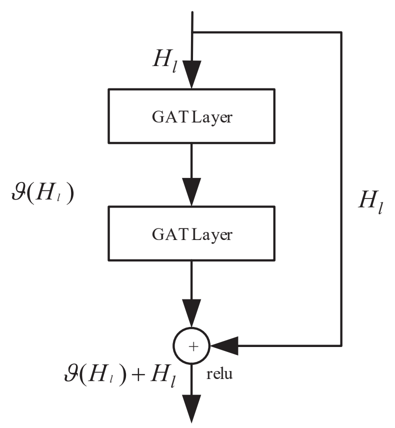

Fig. 7. The residual connection based on GAT layer

other is reverse training. The model's output depends on the final state quantity of the GRU, so that the model uses the past and future information simultaneously. The bidirectional GRU formula is as follows:

$$H_t' = \sigma\left(W_o \cdot \left[h_t^{\rightarrow}, h_t^{\leftarrow}\right]\right) \tag{12}$$

$$output = Dense(Attention(H_t')) \tag{13}$$

where $h^{\rightarrow}$ and $h^{\leftarrow}$ denote the state quantities of the final output of the forward and reverse models, respectively, and *output* is the final result of the model.

IV. Experimental data and parameter setting

A. Dataset

We select the California open dataset PEMS04 as the experimental data. The data contains a total of 307 roads, 59 days from January to March 2013. Each road contains three characteristics (traffic flow, occupancy, and average speed). We analyze the traffic flow as the unique feature the road. The sampling interval of the data is $5min$, and each road contains 16992 traffic flow data. We choose 70% of the data as training data and the rest as testing data. To use the data more conveniently, speed up the convergence of the model and reduce the impact of the origin data on the prediction accuracy, the collected data needs to be normalized. The normalization operation formula is shown in formula(14):

$$X' = \frac{x - x_{min}}{x_{max} - x_{min}} \tag{14}$$

where X' denotes the normalized data; X denotes the original data; x_{min} denotes the minimum value, and x_{max} denotes the maximum value.

B. Performance evaluation indicators

To effectively evaluate the traffic flow prediction model, it is necessary to introduce the evaluation index of the model. We use mean absolute error (MAE) and root means square error (RMSE) as evaluation criteria to measure the prediction accuracy. The formula is defined as follows:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|x_i - \hat{x_i}| \tag{15}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}|x_i = \hat{x_i}|} \tag{16}$$

where n denotes the number of training samples, x_i denotes the real traffic flow, and $\hat{x_i}$ denotes the predicted traffic flow.

C. Model parameter setting

The appropriate setting of the hyperparameters of the prediction model directly affects the accuracy of the prediction results. The structure of the ResGAT-ABiGRU short-term traffic flow prediction model proposed in this paper mainly includes GAT, residual module, and bidirectional GRU module. Through experimental comparison, the overall parameters of the model are shown in TABLE I.

TABLE I
The model hyperparameter setting

name	value
Number of residual block	2
Number of GRU hidden layer neurons	100
Back propagation algorithm	Adam
Learning rate	0.005
Batch size	64
Loss function	L1Loss

V. Experimental results and analysis

1) We compared the prediction results of the GAT model with and without residual connections. The experimental results are shown in TABLE II.

TABLE II
The comparison of results with and without reidual connection in multi-GAT layer

Number of GAT layers	w/ residual		w/o residual	
	RMSE	MAE	RMSE	MAE
2	30.25	19.77	56.48	35.34
4	30.06	19.69	46.87	29.99
6	30.42	19.89	56.53	34.83
8	29.98	19.42	56.10	34.44
10	30.01	19.45	52.61	33.41

As shown in TABLE II, the RMSE and MAE of the prediction model with residual connections are lower than those without residual connections. Therefore, adding residual connections when building a deep network model can effectively solve the problem that the gradient disappears during backpropagation and improve the prediction accuracy. However, building a deep model will reduce the computing speed and increase memory consumption. Finally, we select 4-layer GAT to construct the model through a large number of experiments.

2) To verify the effectiveness of the attention mechanism and residual connection, we test the accuracy of GRU-Attention and ResGAT. To show the advantages of the proposed model, we compare the ResGAT-ABiGRU model with the T-gcn model presented in the literature[11] and two baseline models, SVR and GRU. Finally, we carry out the accuracy test of the above model for 5 minutes and show the experimental results in TABLE III, the prediction performance in Fig. 8.

TABLE III
The results of comparison models on dataset

Methods	Metrics	
	RMSE	MAE
SVR	44.56	28.70
GRU	38.36	23.06
GRU-Attention	37.33	22.52
T-gcn(*literature*[11])	30.64	19.26
ResGAT2	30.06	19.69
ResGAT-ABiGRU	29.63	18.56

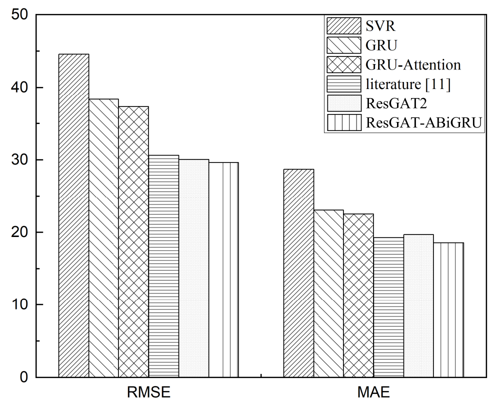

Fig. 8. The comparison diagram of different models' prediction performance

Compared with SVR, the RMSE and MAE of the combined model proposed in this paper are reduced by 33.51% and 35.33%, respectively. Compared with GRU, RMSE decreased by 22.75%, and MAE decreased by 19.51%. Compared with GRU-attention, RMSE decreased by 20.63%, and MAE decreased by 17.58%. Compared with ResGAT2, RMSE and MAE decreased by 1.43% and 5.74%, respectively. Compared with the literature[11], RMSE decreased by 3.29%, and MAE decreased by 3.63%.

In order to be more intuitive to show the model fitting of the predicted values and the actual value, select the model predicted results in the 4 hours of data, a total of 64 data, model and contrast model for proposed respectively mapped the effect graph, as shown in Fig. 9.

As we can see from Fig. 9, the prediction results of the ResGAT-ABiGRU combination model reflect the actual data trend with few abnormal fluctuations in the predicted data. The prediction results of the baseline model GRU model can only reflect the general trend of the accurate data, with a significant fluctuation. The prediction results of the model in literature [11] can be

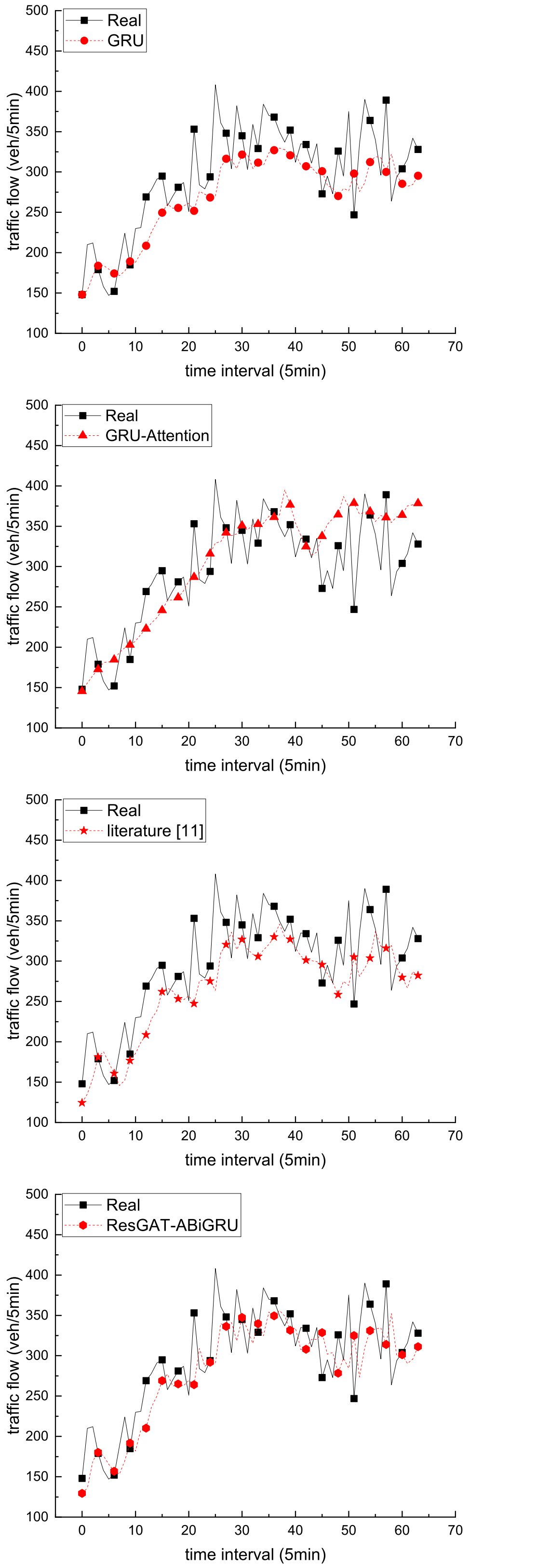

Fig. 9. The comparison of different models' fitting results of traffic flow prediction on real value

compared with real data to a large extent. However, compared with our model, the degree of the fitting is not high. The ResGAT-ABiGRU combined prediction model presented in this paper can extract the spatiotemporal features in the traffic flow data with higher prediction accuracy.

VI. Conclusion

We propose a model to deal with the short-term traffic flow prediction problem in complex road networks. Firstly, we use GAT to extract the spatial features of complex road networks and stack multiple layers of GAT through residual connections to extract the deep spatial correlations between road networks. Secondly, we extract the temporal characteristics of traffic flow data through bidirectional GRU fused with an attention mechanism. Furthermore, the experimental results on the accurate data set show that the model in this paper is superior to other baseline models and the model of the literature [11]. However, this paper does not consider the periodic characteristics of the data. Subsequent research can model workday and non-workday data separately, which can better reflect the periodicity of the data and can also consider the impact of multi-dimensional factors (such as temperature, PM2.5, etc.) on traffic flow.

References

[1] N. Zhao, J. Yuan& X. Han, "Survey on Intelligent Transportation System," Computer Ence.

[2] Y. Jin, E. Tan, L. Li, G. Wang, J. Wang, & P. Wang, "Hybrid Traffic Forecasting Model with Fusion of Multiple Spatial Toll Collection Data and Remote Microwave Sensor Data," IEEE Access. 6 pp. 79211-79221 (2018)

[3] H. Roh, "Assessing the Effect of Snowfall and Cold Temperature on a Commuter Highway Traffic Volume Using Several Layers of Statistical Methods," Transportation Engineering. 2 pp. 100022 (2020)

[4] Z. Tao, C. Xian, & X. Meiping, "Short-Term Traffic Flow Prediction Method Based on K Nearest Neighbor None-Parametric Regression," Systems Engineering Theory & Practice. 2 pp. 376-384 (2010)

[5] X. Feng, X. Ling, H. Zheng, Z. Chen, & Y. Xu, "Adaptive Multi-Kernel SVM with Spatial–Temporal Correlation for Short-Term Traffic Flow Prediction," IEEE Transactions on Intelligent Transportation Systems. 20, 2001-2013 (2018)

[6] S. Liu, S. Liu, Y. Tian, Q. Sun, & Y. Tang, "Research on Forecast of Rail Traffic Flow Based on ARIMA Model," Journal of Physics: Conference Series. 1792, 012065 (2021)

[7] S. Wang, Y. Gu, H. Uchida, H. Fujii, & S. Yoshimura, "A hybrid short-term traffic flow forecasting method based on EMDW-LSSVM," 2020 IEEE 23rd International Conference On Intelligent Transportation Systems (ITSC). pp. 1-6 (2020)

[8] J. Zhang, Y. Zheng, & D. Qi, "Deep Spatial-temporal Residual Networks for Citywide Crowd Flows Prediction," Thirty-first AAAI Conference on Artificial Intelligence. (2017)

[9] B. Cao, & M. Gao, "Research on the Short-Term Traffic Flow Prediction Based on LSTM," Modern Computer. (2018)

[10] W. Pan, & D. Chen, "Research on Short-Term Traffic Flow Prediction Based on GRU-SVR," Computer Technology and Development., 1-7 (2019)

[11] L. Zhao, Y. Song, C. Zhang, Y. Liu, P. Wang, T. Lin, M. Deng, & H. Li, "T-gcn: A Temporal Graph Convolutional Network for Traffic Prediction," IEEE Transactions on Intelligent Transportation Systems. 21, 3848-3858 (2019)

[12] Y. Pan, & Y. Shi, "Short-Term Traffic Forecasting Based on Grey Neural Network with Particle Swarm Optimization," Proceedings of the World Congress on Engineering and Computer Science. 2 (2016)

[13] S. Guo, Y. Lin, N. Feng, C. Song, & H. Wan, "Attention Based Spatial-Temporal Graph Convolutional Networks for Traffic Flow Forecasting," Proceedings of the AAAI Conference on Artificial Intelligence. 33, 922-929 (2019)

[14] K. Zhang, F. He, Z. Zhang, X. Lin, & M. Li, "Graph Attention Temporal Convolutional Network for Traffic Speed Forecasting on Road Networks," Transportmetrica B: Transport Dynamics. 9, 153-171 (2021)

[15] Y. Liu, H. Zheng, X. Feng, & Z. Chen, "Short-term traffic flow prediction with Conv-LSTM," 2017 9th International Conference on Wireless Communications and Signal Processing (WCSP). pp. 1-6 (2017)

[16] T. Ruan, C. H. Jin, L. Xu, J. C. Ding, S. Y. Ying, M.H. Wu, and H. Q. Li, "A Novel Fusion Framework for Sequential Data using Pre-trained Model," IAENG International Journal of Computer Science, vol. 47, no.3, pp593-598, (2020)

Xijun Zhang is an associate professor and master's supervisor. Currently, he is the vice president of The School of Computer and Communication, Lanzhou University of Technology, and a famous teacher of innovation and entrepreneurship education in Gansu Province. Outstanding achievements in teaching, scientific research, innovation, and entrepreneurship. Published more than 20 technical papers and three teaching research papers; Participated in compiling two textbooks and edited one book. Presided over one innovation and entrepreneurship teaching reform project of the Education Department of Gansu Province and participated in one local creation and entrepreneurship specialty construction reform project. Instruct students to compete in more than 50 events. Personal successively won the honorary title: Lanzhou university of science and technology outstanding union activists, Lanzhou university of technology, exceptional communist party member, Lanzhou university of science and technology advanced individual of employment, Lanzhou polytechnic university top ten headteacher, graduated from Lanzhou university of technology design excellent teachers, Lanzhou university of technology course teaching third competition prize, won three consecutive years of school teaching quality excellent.

Guangjie Yu is now a graduate student of Software Engineering in the School of Computer and Communication, Lanzhou University of Technology. He learned Java, C, C++, PHP, and other development languages during his undergraduate period. He has developed and designed website projects independently and has rich development experience. He learned artificial intelligence, machine learning, and other courses during his postgraduate study. He is proficient in building machine learning models with Python and won the provincial first prize and national second prize in the artificial Intelligence Application Competition in his first year of postgraduate study. Besides, he had developed in an all-around way in morality, intelligence, and sports.

Jiyang Shang is a graduate student of Software Engineering in the School of Computer and Communication, Lanzhou University of Technology. Moreover, he has a positive learning attitude, a wide range of hobbies, works sedate, has mature thinking, has good communication skills, can bear hardships and stand hard work, and emphasizes team cooperation.

Baoqi Zhang is a graduate student of Software Engineering in the School of Computer and Communication, Lanzhou University of Technology. He studied the path selection algorithm based on big data of traffic tracks during his postgraduate study. He mastered the establishment and implementation of machine learning algorithms, Python programming, and some heuristic intelligent algorithm models.