WILEY | Hindawi

Research Article

The Prediction of Multistep Traffic Flow Based on AST-GCN-LSTM

Fan Hou,[1] Yue Zhang,[1] Xinli Fu,[3] Lele Jiao,[1] and Wen Zheng ⓘ[1,2]

[1]*Institute of Public-safety and Big Data, College of Data Science, Taiyuan University of Technology, Taiyuan 030060, China*
[2]*Center for Big Data Research in Health, Changzhi Medical College, Changzhi 046000, China*
[3]*College of Foreign Languages, Taiyuan University of Technology, Taiyuan 030060, China*

Correspondence should be addressed to Wen Zheng; zhengwen@tyut.edu.cn

Aiming at the traffic flow prediction problem of the traffic network, this paper proposes a multistep traffic flow prediction model based on attention-based spatial-temporal-graph neural network-long short-term memory neural network (AST-GCN-LSTM). The model can capture the complex spatial dependence of road nodes on the road network and use LSGC (local spectrogram convolution) to capture spatial correlation features from the K-order local neighbors of the road segment nodes in the road network. It is more accurate to extract the information of neighbor nodes by replacing the single-hop neighborhood matrix with K-order local neighborhoods to expand the receptive field of graph convolution. The high-order neighborhood of road nodes is also fully considered instead of only extracting features from first-order neighbor nodes. In addition, an external attribute enhancement unit is designed to extract external factors (weather, point of interest, time, etc.) that affect traffic flow in order to improve the accuracy of the model's traffic flow prediction. The experimental results show that when considering the static, dynamic, and static and dynamic combination, the model has excellent performance: RMSE (4.0406, 4.0362, 4.0234), MAE (2.7184, 2.7044, 2.7030), accuracy (0.7132, 0.7190, 0.7223).

1. Introduction

Traffic forecasting is an important field in the research of intelligent transportation [1], and effective traffic flow forecasting can alleviate traffic congestion, travel planning, and traffic management for individual drivers and decision-makers [2, 3]. The complex temporal and spatial correlations between traffic flows will show huge differences affected by external emergencies [4], dynamic factors, and static factors. Ahmed [5] and others proposed an autoregressive integrated moving average model (ARIMA) model that can only deal with nonstationary time series data. It is difficult to explore connections between dynamic data and is no longer suitable for current application scenarios. In addition, though traditional linear methods such as a series of Kalman filtering methods proposed and improved by Stephanedes [6], Xie (2007) [7], Ojeda (2013) [8], Guo [9] have improved the accuracy of traffic prediction in some aspects, its ability to fit nonlinear traffic flow data is still poor, and it increases the prediction time [10–12].

With the development of computer capabilities, typical machine learning methods, such as support vector regression (SVR) [13, 14], k-nearest neighbor algorithm [15, 16] K-NN (K-NearestNeighbor), and decision tree models [17–19], can dig out the essential laws and rich information hidden in traffic flow from massive data [20], and better promote the development process of traffic flow forecasting.

The emergence of deep neural network models has enabled the development of the potential of artificial intelligence in traffic prediction. Although some simple network structures can improve the accuracy of model traffic prediction [21], there are problems such as slow convergence, prone to over-fitting, and prone to error values [22]. Compared with the traditional neural network model, recurrent neural network (RNN) [23], long short-term memory network (LSTM) [24], and gate recurrent unit

(GRU) [25] can effectively use the self-loop system and learn time series features to improve the effectiveness of prediction. Therefore, it is used as a component of each model to predict traffic speed, travel time, traffic flow, etc.

In order to capture the spatial dependencies in the traffic road network, researchers [26] extract spatial features combined with convolutional neural networks (CNN) from two-dimensional spatiotemporal traffic data. The description of the traffic structure using two-dimensional spatio-temporal data is not accurate and does not conform to the complex road network conditions in real life so some scholars [27] have begun to try to convert the structure of the traffic network into images and use CNN to learn the traffic images in order to capture the spatial characteristics. However, there is more or less noise in the images converted by the traffic network structure, and the existence of noise will inevitably cause CNN to capture false spatial relationships. Traditional methods based on CNN cannot essentially deal with the topological structure and physical properties of the traffic network. Recent studies [28, 29] also tried to convert the traffic state data into a three-dimensional (3D) matrix and use 3D convolutional networks to extract characteristics in deeper levels. Researchers [30] learned the traffic network as a graph and extract features from the graph structure of the traffic network using convolution operators based on graphs, which effectively learns the changes in traffic flow under the temporal and spatial attributes and achieves great forecast results.

After considering the dynamic change characteristics of the traffic network, this paper proposes the AST-GCN-LSTM model, which can predict the future traffic state according to historical traffic flow information on roads and external auxiliary information. Since traffic flow is affected by a variety of external factors (actual factors) such as weather, holiday, and time, in this article, we predict the traffic speed in the future based on the traffic speed in the past period of time and the external factors that affect the traffic flow. This is of great significance for realizing dynamic traffic signal optimization, dynamic traffic management planning, and traffic management decision [31].

2. Method

2.1. The Introduction of Basic Algorithms

2.1.1. Graph Neural Network (GCN). The transportation network can be regarded as a graph composed of nodes and edges, so the transportation network as a graph structure has been used for dynamic shortest path routing [32], traffic congestion analysis [33], and dynamic traffic allocation [34].

The most commonly used method for our research on graph networks is to introduce a spectrum frame in the spectrum domain [35] and obtain the spectrogram convolution model by designing the spectrum convolution based on the graph Laplacian matrix. In order to reduce the number of parameters and save the amount of calculation, we use the local spectrogram convolution with polynomial filter, but the Laplacian matrix power operation still requires a lot of calculation and high complexity, and to reduce the

complexity, the Chebyshev polynomial is introduced to calculate the K-order local convolution, which can reduce the computational time complexity from the square level to the linear level.

As shown in Figure 1, the spectrogram convolution model using Chebyshev polynomial approximation can capture features from the K-order local neighbors of the vertices in the graph, fully taking into account the high-order neighborhood of the node instead of extracting features from the single-hop neighborhood only. This chapter expands the receptive field of graph convolution by replacing the single-hop neighborhood matrix with the K-order local neighborhood, which can extract the information of neighbor nodes more accurately.

2.1.2. GCN-LSTM Structure. To capture the complex spatial correlation and dynamic time correlation of traffic data in the real world, we have added a long- and short-term memory neural network LSTM. LSTM is an improved recurrent neural network (RNN), and LSTM has better performance than ARIMA when the training time series is long enough [36, 37]. The basic unit of the hidden layer of LSTM is a special cell unit, not a traditional neuron node. It is this special memory unit that enables LSTM to successfully solve the defect of RNN gradient explosion and also capture the temporal correlation of traffic flow. The overall structure of the GCN-LSTM structure is shown in Figure 2. In order to capture the complex spatial correlation and dynamic time correlation of traffic data in the real world, we combine the GCN with LSTM models. The function of the GCN model is to generate a graph of the traffic information of the road segment based on a given graph structure. It learns the representation of the road segment by integrating the characteristics of the local neighbors of the node and captures the spatial dependence of these road segments in the road network at each timestamp. Then, these time-varying feature representations are input into the LSTM model to capture the time dependence [38, 39].

2.2. AST-GCN-LSTM Spatiotemporal Graph Convolution Model

2.2.1. Attribute Augmentation Unit. On the basis of the GCN-LSTM traffic flow model introduced in Section 2.1.2, we have added an attribute augmentation unit. As shown in Figure 3, static external attribute features and dynamic external attribute features expand the dimensions of the original traffic feature matrix through attribute augmentation units.

At time t, traffic information matrix X_t is extracted from the historical feature matrix $\mathcal{X} = (X_1, \ldots, X_t)$, and $\{L^{t-w}, L^t\}$ is the set of dynamic attribute features of $w + 1$ time windows. In different timestamps, the static attribute feature set of H is always unchanged. L and H are merged to generate an augmented matrix P^t. The problem of multistep traffic flow prediction can be expressed as

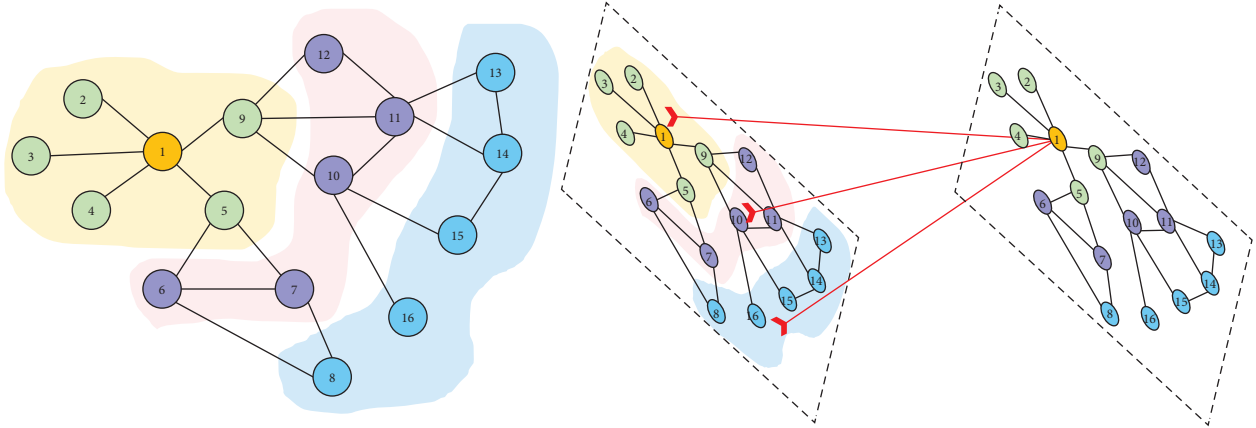$$\widehat{y} = f(A, X, P). \tag{1}$$

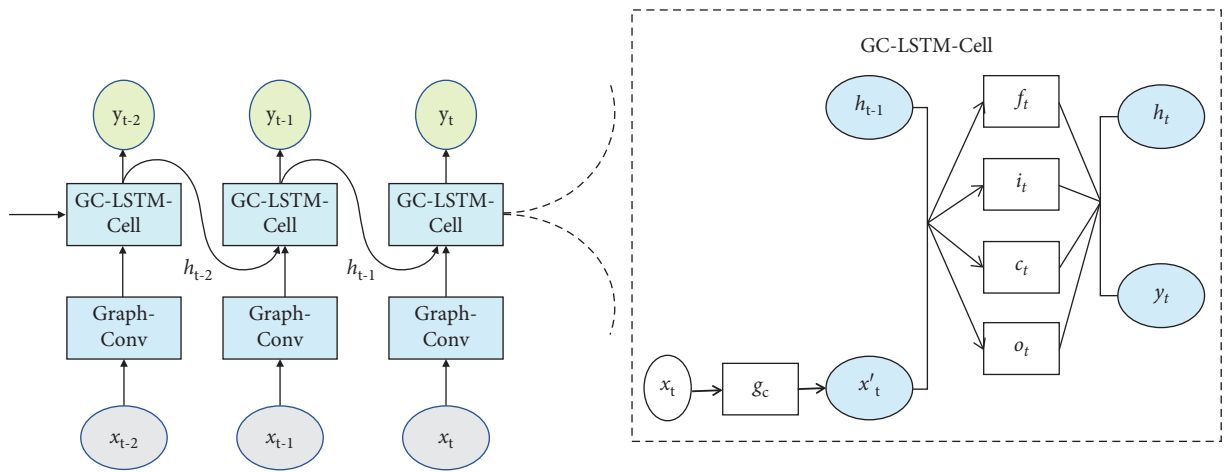FIGURE 1: K-hop neighbors of graph convolution.

FIGURE 2: The architecture of the GCN-LSTM model. i_t is the input gate, o_t is the output gate, f_t is the forget gate, and c_t is the input unit state.
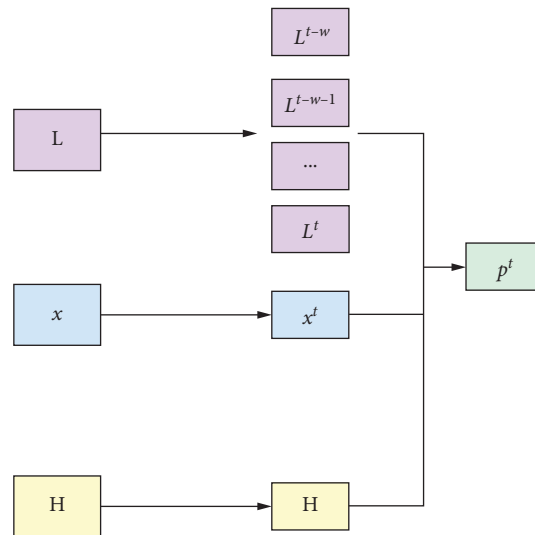

FIGURE 3: The Attribute augmentation unit.

The model learns the complex spatial dependence, dynamic time dependence, and external dependence in traffic data. In this model, the gate structure and hidden state in LSTM are unchanged, but the input is replaced by the graph convolution feature. At time t, the input gate, output gate, forget gate, and input unit are defined as formulas (2) to (7):

$$f_t = \sigma\left(W_f \cdot [g_c(p_t, A), h_{t-1}] + b_f\right), \qquad (2)$$

$$i_t = \sigma\left(W_i \cdot [gc(P_t, A), h_{t-1}] + b_i\right), \qquad (3)$$

$$\tilde{c}_t = \tanh\left(W_c \cdot [gc(P_t, A), h_{t-1}] + b_c\right), \qquad (4)$$

$$c_t = f_t \circ c_{t-1} + i_t \circ \tilde{c}, \qquad (5)$$

$$o_t = \sigma\left(W_\circ \cdot [gc(P_t, A), h_{t-1}] + b_\circ\right), \qquad (6)$$

$$h_t = o_t \circ \tanh(c_t). \qquad (7)$$

The sign "•" is a matrix multiplication operator, W_f, W_i, W_c, and W_o are weight matrices that map the input to the states of three gates and input units, and b_f, b_i, b_c, and b_o are four deviation vectors. s is the activation function of the gate, which is usually the sigmoid function. Tanh is the hyperbolic tangent function, and $gc(\cdot)$ represents the graph convolution operation (Chebyshev polynomial approximation).

2.2.2. Loss Function.

In the process of model training, loss is chosen as the training target to optimize the error of multistep prediction and make the prediction result close to the real traffic state. Therefore, the loss function used in multistep traffic prediction AST-GCN-LSTM can be expressed as follows:

$$\text{Loss} = L(y_T, \hat{y}_T) = \|y_T - \hat{y}_T\| + \rho L_{\text{reg}}. \qquad (8)$$

$L(\bullet)$ is a function to calculate the error between the predicted $\hat{y}_T$ value and the true value y_T. Here, L_{reg} represents a regular term avoiding over-fitting of the model, and ρ is a hyperparameter that is learnable in the network.

2.2.3. AST-GCN-LSTM Spatiotemporal Graph Convolution Model.

The GCN-LSTM traffic flow model introduced in Section 2.1.2 is combined with the attribute expansion unit in Section 2.2.1. A multistep traffic flow prediction model (AST-GCN-LSTM) that considers external factors is also proposed. This model fully takes the external attribute characteristics that affect the traffic flow into account. Figure 4 shows the overall framework of the model, which is mainly composed of data preprocessing, attribute expansion, and spatiotemporal graph convolutional layers. In the model, we set the number of neural units in all hidden layers to 64, the batch size to 64, the learning rate to 0.001, the order of the Chebyshev polynomial to 3, and the maximum number of training iterations to 3000. The Adam optimizer is used to train the model. The data set is divided into two parts, 80% of the data are used for training, and 20% of the data are used for testing. After dividing the data set into two parts, we generate sequence samples through a time window whose width is $T + T'$.

2.3. Data Set.

The traffic speed data set of the real-world road network is used in this article to evaluate the model performance. This public data set contains the taxi trajectory data of every 15 minutes setting on 156 roads from January 1 to January 31, 2015. The data sampling location is Luohu District, Shenzhen, Guangdong Province. The data mainly include the following 4 parts:

(1) Adjacency matrix: the data set selects 156 roads, so the size of the adjacency matrix A is 156 ∗ 156. The adjacency matrix represents the connectivity between segments. Each row of the matrix represents a road. If there are links connecting nodes i and j, then the element in the adjacency matrix $A_{i,j} = 1$. If there are no links connecting nodes i and j, the element in the adjacency matrix $A_{i,j} = 0$.

(2) Feature matrix: the feature matrix size of the data set is 2976 ∗ 156. The feature matrix is the speed value of 156 roads in 31 days. Each column represents a road, and each row represents the traffic speed value of 156 roads at a certain time τ. Speed information is collected every 15 minutes.

(3) Static attribute characteristic matrix: the point-of-interest (POI) information on 156 roads is provided in the data set. POI categories include the following nine types: catering, business, shopping, transportation, education, life, medical care, accommodation, and others. When determining the POI category of each road, the POI distribution on each road is calculated firstly, and then, proportions of the various categories of POIs are calculated. After comparing the proportions of the various categories of POIs, the POI with the largest proportion is used as the static feature of the road. The size of the static attribute feature matrix is 156 ∗ 1.

(4) Dynamic attribute feature matrix: weather conditions of every 15 minutes in January are provided by the data set, which can be divided into five categories: light rain, heavy rain, cloudy, foggy, and sunny. Time information includes the time of day, weekdays, and weekends. Because they will have a significant impact on the traffic state, this section also takes it into consideration. The size of the three types of external attribute feature matrices is 2976 ∗ 156.

2.4. Evaluation Index.

Three commonly used traffic forecasting indicators are as follows: mean absolute error (MAE), accuracy, and root mean square error (RMSE) [31] are used to evaluate the performance of the proposed model and the comparison model. The formula is from (9) to (11):
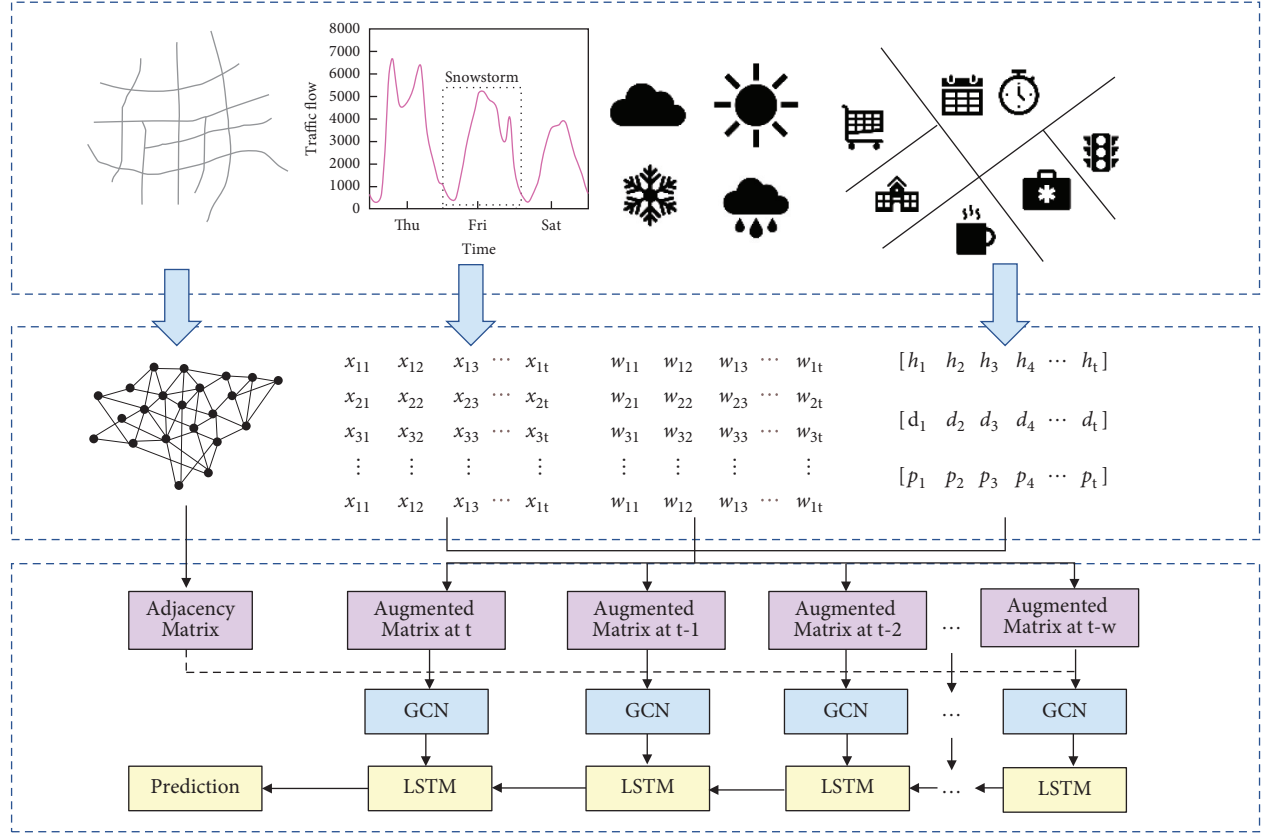
FIGURE 4: The architecture of the AST-GCN-LSTM model.

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^{n} \left| y_T - \hat{y}_T \right|, \qquad (9)$$

$$\text{accuracy} = 1 - \frac{\left\| y_T - \hat{y}_T \right\|_F}{\left\| y_T \right\|_F}, \qquad (10)$$

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} \left(y_T - \hat{y}_T \right)^2}. \qquad (11)$$

where n is the total number of test sets and y_T and $\hat{y}_T$ represent the true and predicted values of the flow.

3. Analysis of Experimental Results

3.1. Analysis of Static Attribute. In order to evaluate the overall performance of our proposed AST-GCN-LSTM model, we compare it with other traditional and common models. These models are as follows:

(1) Historical average model (HA): HA models the traffic flow as a seasonal cyclical process and uses the average value of the previous seasons (for example, the flow value of the same time period in the previous days) as the predictive value

(2) Autoregressive integrated moving average model (ARIMA): the autoregressive integrated moving

average model (ARIMA) with Kalman filter is widely used in time series forecasting. It predicts the series by fitting time series data.

(3) Support vector regression (SVR): linear support vector machine is used to predict the regression task of traffic flow sequence

(4) Diffusion convolution recurrent neural network (DCRNN): diffusion convolution recurrent neural network formulates the diffusion process in graph convolution and uses a two-way random walk to capture the spatial correlation of the traffic flow in graphs. An encoder-decoder is used to capture the temporal correlation of the traffic flow, and the diffusion convolution GCN is combined with the recursive model in prediction.

(5) GCN-LSTM: the combination of LSGC and LSTM model using Chebyshev polynomial approximation is introduced in Section 2.1.1.

Among them, HA, ARIMA, and SVR are traditional nonneural network models, DCRNN is a deep learning model that can capture spatial features, and GCN-LSTM is a deep learning model that comprehensively considers the spatial features and dynamic correlation of traffic data.

Table 1 shows the overall prediction performance of the AST-GCN-LSTM model and five representative methods. Three indicators, root mean square error (RMSE), mean

TABLE 1: Performance comparison of different methods.

T (min)	Metrics	HA	ARIMA	SVR	DCRNN	GCN-LSTM	AST-GCN-LSTM
	RMSE	4.2951	7.2406	4.1455	4.5000	4.1193	4.0234
15	MAE	2.7815	4.9824	2.6233	3.1700	2.7701	2.7030
	Accuracy	0.7008	0.4463	0.7112	0.2913	0.7129	0.7223
	RMSE	4.2951	6.7899	4.1628	4.5600	4.1207	4.0508
30	MAE	2.7815	4.6765	2.6875	3.2300	2.7739	2.7244
	Accuracy	0.7008	0.3845	0.7100	0.2970	0.7126	0.7196
	RMSE	4.2951	6.7852	4.1885	4.6000	4.1252	4.0587
45	MAE	2.7815	4.6734	2.7359	3.2700	2.7753	2.7346
	Accuracy	0.7008	0.3847	0.7082	0.3021	0.7123	0.7172
	RMSE	4.2951	6.7708	4.2156	4.6400	4.1262	4.0689
60	MAE	2.7815	4.6655	2.7751	3.3100	2.7811	2.7403
	Accuracy	0.7008	0.3851	0.7063	0.3069	0.7119	0.7165

absolute error (MAE), and accuracy (accuracy) evaluation, are used in the comparison of performances.

According to Table 1, it can be concluded that, from the results of the 15-minute prediction window, compared with the traditional models, HA, ARIMA, and SVR, the RMSE value of the AST-GCN-LSTM model decreases by 3.07%, 44.43%, and 2.95%. Compared with the HA model and the SVR model, the accuracy value is increased by 14.69% and 1.56%, respectively. This shows that HA, ARIMA, and SVR cannot compete with other methods because the data have complex spatiotemporal correlation and high-dimensional features, and nonneural network methods are not suitable for such network-wide prediction tasks Accounting for external attribute features, the RMSE value of the AST-GCN-LSTM model that takes all external attributes into account is 10.59% and 2.33% lower than that of the DCRNN model and GCN-LSTM model. The value of MAE is lower than that of the DCRNN model and GCN-LSTM model and reduces by 14.73% and 2.42%. According to Table 1, compared with traditional methods and other methods based on deep learning, the model proposed by this article has achieved significant improvements proving the effectiveness of the model.

3.2. Analysis of the External Attribute.
In order to verify the influence of external attribute characteristics in traffic flow prediction, corresponding comparative experiments are done. The experimental settings are divided into four kinds as follows: adding static attribute characteristics only, adding dynamic attribute characteristics only, adding dynamic and static external attribute characteristics at the same time, and not adding external attributes characteristics. The results are shown in Figure 5. Yellow is the result of adding static attribute characteristics. Gray is the result of adding dynamic attribute features. Blue is the result of adding dynamic and static external attributes at the same time.

It can be seen from Figure 5 that when only dynamic attribute features are considered, the value of AST-GCN-LSTM (dynamic) RMSE is 10.31% and 2.02% lower than that of DCRNN and GCN-LSTM models. The value of MAE is lower than that of DCRNN and GCN-LSTM models and reduced by 14.69% and 2.37%. When only static attributes are considered,
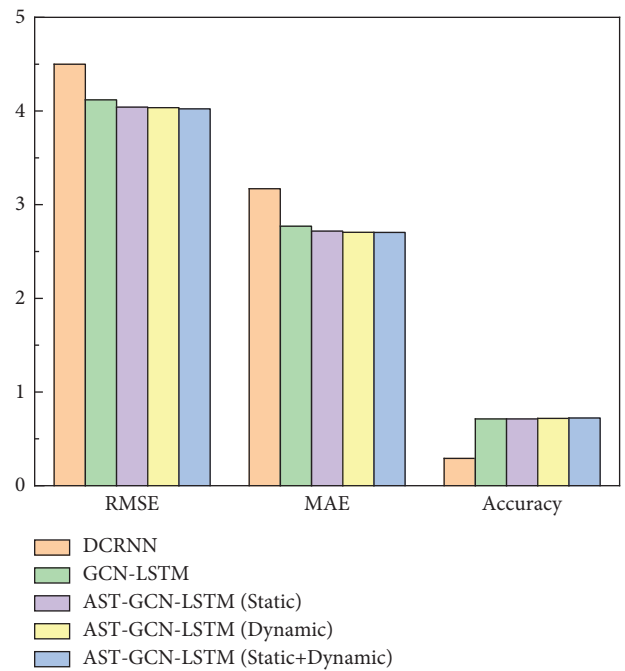

FIGURE 5: Experiments under different conditions.

the value of AST-GCN-LSTM (static) RMSE is reduced by 10.21% and 1.91% compared with DCRNN and GCN-LSTM models, and the value of MAE is reduced by 14.25% and 1.87% compared with DCRNN and GCN-LSTM models. When static factors and dynamic factors are considered at the same time, the RMSE value of the AST-GCN-LSTM model is reduced by 10.59% and 2.33% compared with the DCRNN model and the GCN-LSTM model, and the value of MAE is reduced by 14.73% and 2.42% compared with the DCRNN model and the GCN-LSTM model.

It can be seen from Figure 5 that the model performance when only dynamic attribute features are considered is better than the model performance when only static attribute features are considered. This also indirectly illustrates the importance of considering dynamic external attribute features, and we also observed that when static and dynamic factors are considered at the same time, the performance of the model is optimal.
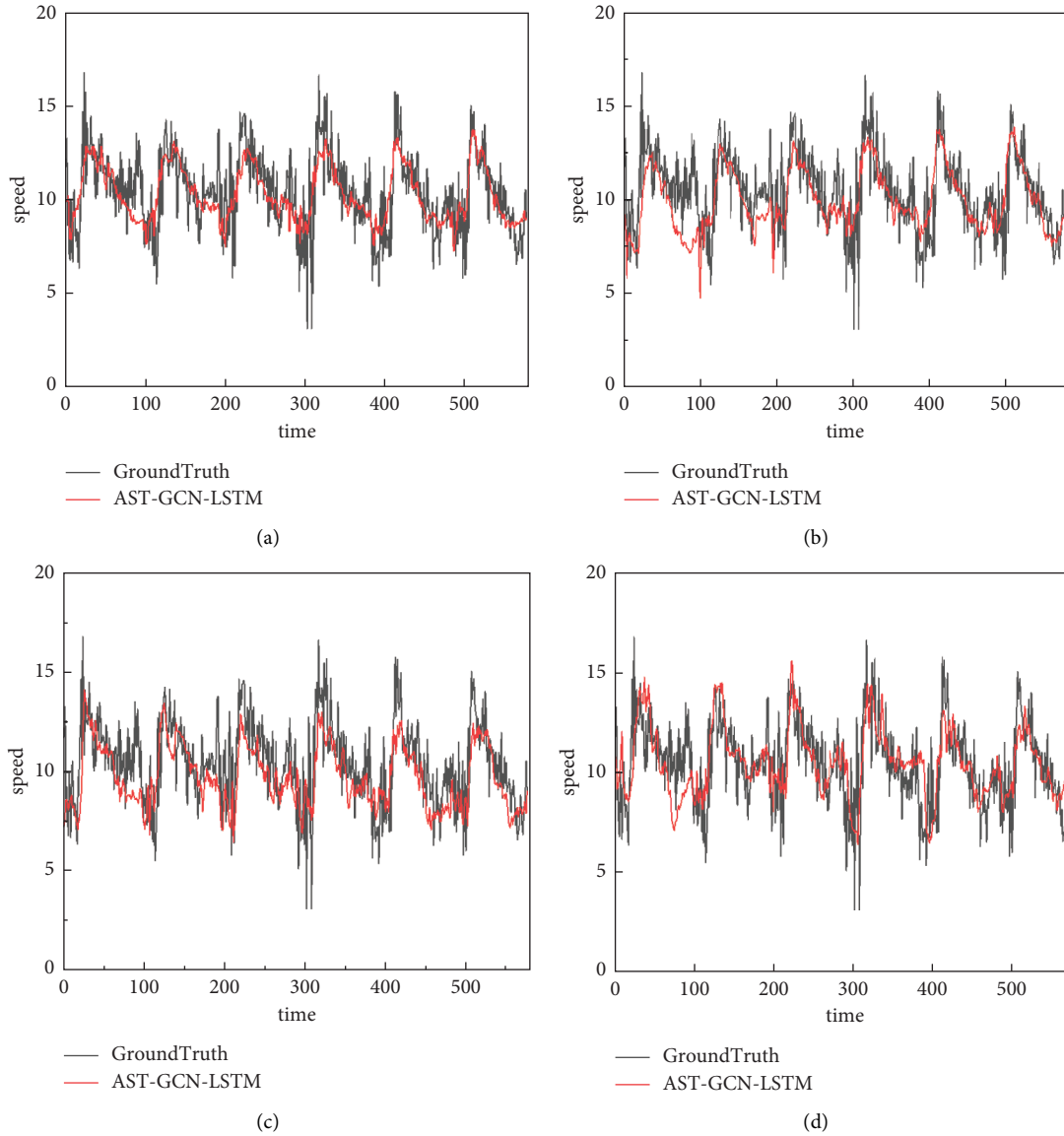
(a)

(b)

(c)

(d)

FIGURE 6: Visualization of the prediction window. (a–d) The prediction effect of 15, 30, 45, and 60 minutes.

In summary, considering the external information has a good effect on the prediction of the model under actual conditions.

3.3. Performance in Different Forecast Periods. All tests of this model use 60 minutes as the historical time window, which means four observation data points are used to predict the traffic conditions in the future 15, 30, 45, and 60 minutes ($H = 1, 2, 3, 4$). Figure 6 shows the visualization results of 15-, 30-, 45-, and 60-minute forecast windows. Each graph is the prediction result from January 26, 2015, to January 31, 2015.

It can be seen from Table 2 that when the traffic flow prediction window is 15 minutes, the RMSE value of the AST-GCN-LSTM model is reduced by 10.59% and 2.33% compared with the DCRNN model and the GCN-LSTM

TABLE 2: Performance comparison of different prediction durations.

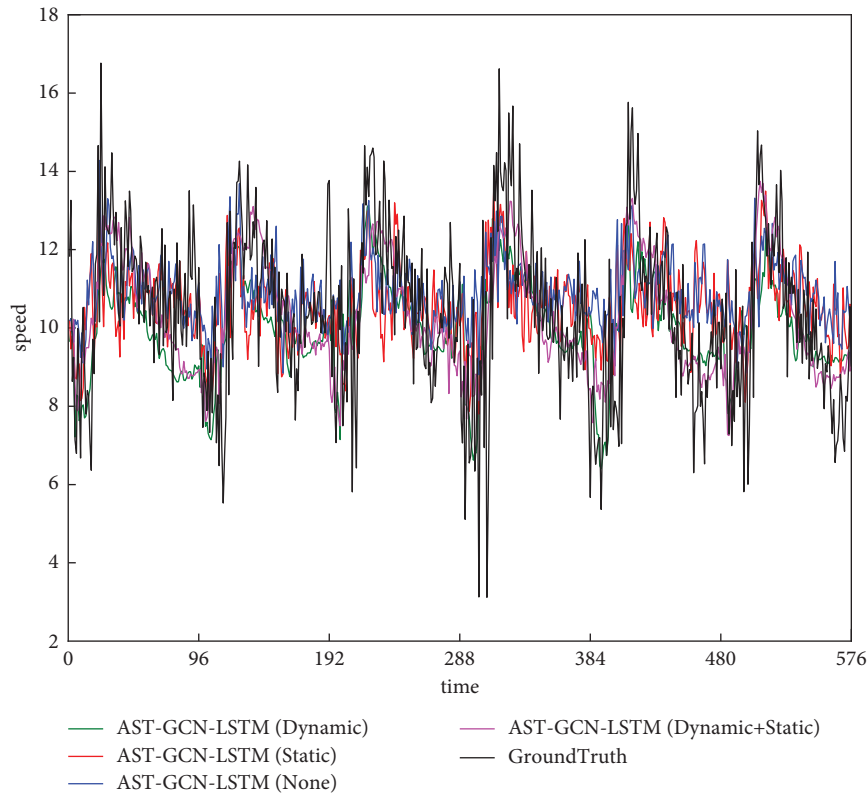| T (min)	Metric	DCRNN	GCN-LSTM	AST-GCN-LSTM
	RMSE	4.5000	4.1193	4.0234
15	MAE	3.1700	2.7701	2.7030
	Accuracy	0.2913	0.7129	0.7223
	RMSE	4.5600	4.1207	4.0508
30	MAE	3.2300	2.7739	2.7244
	Accuracy	0.2970	0.7126	0.7196
	RMSE	4.6000	4.1252	4.0587
45	MAE	3.2700	2.7753	2.7346
	Accuracy	0.3021	0.7123	0.7172
	RMSE	4.6400	4.1262	4.0689
60	MAE	3.3100	2.7811	2.7403
	Accuracy	0.3069	0.7119	0.7165

FIGURE 7: Comparison among forecasting enhanced by different external information. The green line is the dynamic attribute prediction of the model, the red line is the static attribute prediction of the model, the blue line is the model without external attribute prediction, the pink line is the model dynamic and static attribute prediction, and the black line is the real traffic flow.

model. Compared with the DCRNN model and the GCN-LSTM model, the value of MAE is reduced by 14.73% and 2.42%. When the traffic flow prediction window is 30 minutes, the RMSE value of the AST-GCN-LSTM model is decreased by 11.17% and 1.70% compared with the DCRNN model and the GCN-LSTM model. The MAE value is decreased by 15.65% and 1.78% compared with the DCRNN model and the GCN-LSTM model. When the traffic flow prediction window is 45 minutes, the RMSE value of the AST-GCN-LSTM model is 11.77% and 1.61% lower than the DCRNN model and the GCN-LSTM model, and the MAE value is 16.37% and 1.47% lower than the DCRNN model and the GCN-LSTM model.

When the traffic flow prediction window is 60 minutes, the RMSE value of the AST-GCN-LSTM model is decreased by 12.31% and 1.39% compared with the DCRNN model and the GCN-LSTM model. The MAE value is decreased by 17.21% and 1.47% compared with the DCRNN model and the GCN-LSTM model. The above conclusions show the robustness and stability of our proposed model in long-term prediction.

For different prediction times, the AST-GCN-LSTM model proposed in this paper can predict traffic speed well. The results show that this model can capture the changing trend of traffic speed very well, which also verifies the effectiveness of our model in multistep traffic flow prediction.

Comparing the prediction values of the 15-minute and 60-minute prediction windows, we can see that the prediction effect of the 15-minute window in the short-term

prediction is closer to the true value, which also shows that the model can better capture short-term dependence.

In order to test the effectiveness of adding static and dynamic external attribute features, we visualized the model prediction results. Figure 7 shows a comparison of prediction results between models with static external attributes, dynamic external attributes, and models without external attributes.

From the visualization results in Figure 7, it can be found that the deviation between the predicted result of AST-GCN-LSTM and the real speed value is smaller than that of AST-GCN-LSTM (static attribute) and AST-GCN-LSTM (dynamic attribute), which indicates that the diversity of external information can better promote prediction.

4. Conclusions

This paper uses the proposed AST-GCN-LSTM model to obtain dynamic attribute features by adding the attribute augmentation unit structure of external factors. After the feature matrix is augmented, the Chebyshev polynomial approximation spectrogram convolution model is used for feature extraction. This model can capture the spatial characteristics of traffic flow from the K-order local neighbors of the vertices in the graph. The K-order local neighborhood can replace the single-hop neighborhood matrix to expand the receptive field of the graph convolution, which can more accurately extract the information of

neighbor nodes. After the information is extracted, the characteristic representation of the information that changes over time is input into the LSTM model to capture the time dependence. By analyzing the performance of the proposed model, including the performance analysis of external attribute characteristics and the performance analysis of different prediction windows, and comparing with different baseline models to verify the effectiveness of the proposed model, it solves the inability of the previous traffic prediction models. The external factors affecting traffic flow are fully considered.

Results show that the AST-GCN-LSTM model can not only fully consider the spatial relationship of road nodes but also capture the time dependence of traffic flow and effectively improve the accuracy of traffic prediction. In addition, the AST-GCN-LSTM model is suitable for both road network traffic flow prediction and midterm and long-term traffic flow prediction and multistep prediction.

Data Availability

The data set in this article can be obtained by contacting the corresponding author.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

Acknowledgments

References

[1] J. Li, F. Guo, A. Sivakumar, Y. Dong, and R. Krishnan, "Transferability improvement in short-term traffic prediction using stacked LSTM network," *Transportation Research Part C: Emerging Technologies*, vol. 124, Article ID 102977, 2021.

[2] M. Cao, V. Li, and V. Chan, "A CNN-LSTM Model for Traffic Speed Prediction," in *Proceedings of the 2020 IEEE 91st Vehicular Technology Conference (VTC2020-Spring)*, Antwerp, Belgium, May 2020.

[3] L. Li, Y. Yang, and Z. Yuan, "A spatial-temporal approach for traffic status analysis and prediction based on bi-lstm structure," *Modern Physics Letters B*, vol. 35, no. 31, 2021.

[4] S. Wang, Q. Zhuo, H. Yan, Q. M. Li, and Y. Qi, "A network traffic prediction method based on LSTM," *ZTE Communications*, vol. 17, no. 2, 2019.

[5] M. S. Ahmed and A. R. Cook, "Analysis of freeway traffic time-series data by using box-jenkins techniques," *Traffic Time-Series Data by Using Box-Jenkins Techniques*, pp. 1–9, 1979.

[6] I. Okutani and Y. J. Stephanedes, "Dynamic prediction of traffic volume through Kalman filtering theory," *Transportation Research Part B: Methodological*, vol. 18, no. 1, pp. 1–11, 1984.

[7] Y. Zhang, Z. Ye, and Y. Xie, "Short-term traffic volume forecasting using kalman filter with discrete wavelet decomposition," *Computer-Aided Civil and Infrastructure Engineering*, vol. 22, no. 5, pp. 326–334, 2007.

[8] L. L. Ojeda, A. Y. Kibangou, and C. C. D. Wit, "Adaptive Kalman Filtering for Multi-step Ahead Traffic Flow Prediction," in *Proceedings of the 2013 American Control Conference*, pp. 4724–4729, Washington, DC, USA, June 2013.

[9] J. Guo, W. Huang, and B. M. Williams, "Adaptive Kalman filter approach for stochastic short-term traffic flow rate prediction and uncertainty quantification," *Transportation Research Part C: Emerging Technologies*, vol. 43, no. 2, pp. 50–64, 2014.

[10] M. Y. Aghdam, S. K. Tabbakh, S. M. Chabok, and M. khairabadi, "Optimization of air traffic management efficiency based on deep learning enriched by the long short-term memory (LSTM) and extreme learning machine (ELM)," *Journal of Big Data*, vol. 8, no. 1, pp. 1–26, 2021.

[11] Y. Li, S. Chai, Z. Ma, and G. Wang, "A hybrid deep learning framework for long-term traffic flow prediction," *IEEE Access*, vol. 9, p. 1, 2021.

[12] T. Li, A. Ni, C. Zhang, and G. Xiao, "Short-term traffic congestion prediction with Conv-BiLSTM considering spatio-temporal features," *IET Intelligent Transport Systems*, vol. 14, no. 1, 2021.

[13] M. Zhang, Z. Yaobao, H. Ganglong, and C. Gang, "Accurate multisteps traffic flow prediction based on SVM," *Mathematical Problems in Engineering*, vol. 12, no. 6, pp. 1–8, 2013.

[14] X. Feng, X. Ling, H. Zheng, Z. Chen, and Y. Xu, "Adaptive multi-kernel SVM with spatial-temporal correlation for short-term traffic flow prediction," *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 99, pp. 1–13, 2018.

[15] D. Fan and X. Zhang, "Short-term traffic flow prediction method based on balanced binary tree and K-nearest neighbor nonparametric regression," in *Proceedings of the International Conference on Modelling, Simulation and Applied Mathematics*, Bangkok, Thailand, March 2017.

[16] S. Wu, Z. Yang, X. Zhu, and B. Yu, "Improved k-nn for short-term traffic forecasting using temporal and spatial information," *Journal of Transportation Engineering*, vol. 140, no. 7, Article ID 04014026, 2014.

[17] D. Xu, Y. Wang, P. Peng, S. Beilun, Z. Deng, and H. Guo, "Real-time road traffic state prediction based on kernel-KNN," *Transportmetrica: Transportation Science*, vol. 16, no. 1, pp. 104–118, 2020.

[18] R. V. Leshem, "Traffic flow prediction using adaboost algorithm with random forests as a weak learner," *Proceedings of World Academy of Science, Engineering and Technology*, vol. 19, no. -, pp. 193–198, 2007.

[19] Y. Zhang and A. Haghani, "A gradient boosting method to improve travel time prediction," *Transportation Research Part C: Emerging Technologies*, vol. 58, no. SEP.B, pp. 308–324, 2015.

[20] S. Yang, J. Wu, Y. Du, Y. He, and X. Chen, "Ensemble learning for short-term traffic prediction based on gradient boosting machine," *Journal of Sensors*, vol. 2017, Article ID 7074143, 15 pages, 2017.

[21] W. Huang, G. Song, H. Hong, and K. Xie, "Deep architecture for traffic flow prediction: deep belief networks with multitask learning," *IEEE Transactions on Intelligent Transportation Systems*, vol. 15, no. 5, pp. 2191–2201, 2014.

[22] Y. Lv, Y. Duan, W. Kang, and Z. Li, "Traffic flow prediction with big data: a deep learning approach," *Intelligent transportation systems*, vol. 16, no. 2, pp. 865–873, 2015.

[23] A. Graves, *Long Short-Term Memory*Springer, Heidelberg, Germany, 2012.

[24] R. Fu, Z. Zhang, and L. Li, "Using LSTM and GRU Neural Network Methods for Traffic Flow prediction," in *Proceedings of the 2016 31st Youth Academic Annual Conference of Chinese Association of Automation (YAC)*, pp. 324–328, Wuhan, Hubei Province China, November 2016.

[25] S. Hart, E. J. Hultink, N. Tzokas, and H. R. Commandeur, "Industrial companies' evaluation criteria in new product development gates," *Journal of Product Innovation Management*, vol. 20, no. 1, pp. 22–36, 2003.

[26] Y. Li, R. Yu, C. Shahabi, and L. Yan, "Diffusion convolutional recurrent neural network: data-driven traffic forecasting," in *Proceedings of the Sixth International Conference on Learning Representations*, Vancouver, Canada, May 2018.

[27] H. Yu, Z. Wu, S. Wang, Y. Wang, and X. Ma, "Spatiotemporal recurrent convolutional networks for traffic prediction in transportation networks," *Sensors*, vol. 17, no. 7, 2017.

[28] N. Hoang, L. K. Minh, W. Tao, and C. Chen, "Deep learning methods in transportation domain: a review," *Iet Intelligent Transport Systems*, vol. 12, no. 9, 2018.

[29] X. Ma, Z. Dai, Z. He, J. Ma, Y. Wang, and Y. Wang, "Learning traffic as images: a deep convolutional neural network for large-scale transportation network speed prediction," *Sensors*, vol. 17, no. 4, 2017.

[30] C. Zhang and P. Patras, "Long-term mobile traffic forecasting using deep spatio-temporal neural networks," in *Proceedings of the Eighteenth ACM International Symposium on Mobile Ad Hoc Networking and Computing*, Chennai, India, July 2017.

[31] E. I. Vlahogianni, M. G. Karlaftis, and J. C. Golias, "Short-term traffic forecasting: where we are and where we're going," *Transportation Research Part C: Emerging Technologies*, vol. 43, no. 1, 2014.

[32] et al., "Exploiting spatio-temporal correlations with multiple 3D convolutional neural networks for citywide vehicle flow prediction," in *Proceedings of the 2018 IEEE International Conference on Data Mining (ICDM)*, vol. 60, pp. 893–898, Singapore, November 2018.

[33] S. Guo, Y. Lin, S. Li, Z. Chen, and H. Wan, "Deep spatial-temporal 3D convolutional neural networks for traffic data forecasting," *IEEE Transactions on Intelligent Transportation Systems*, vol. 99, pp. 1–14, 2019.

[34] Y. Sun, X. Yu, R. Bie, and H. Song, "Discovering time-dependent shortest path on traffic graph for drivers towards green driving," *Journal of Network and Computer Applications*, vol. 83, pp. 204–212, 2016.

[35] H. Sun, J. Wu, D. Ma, and J. Long, "Spatial distribution complexities of traffic congestion and bottlenecks in different network topologies," *Applied Mathematical Modelling*, vol. 38, no. 2, pp. 496–505, 2014.

[36] A. Azari, P. Papapetrou, S. Denic, and G. Peters, "Cellular traffic prediction and classification: a comparative evaluation of lstm and arima," in *Proceedings of the International Conference on Discovery Science*, Split, Croatia, October 2019.

[37] T. Nguyen, "Spatiotemporal tile-based attention-guided lstms for traffic video prediction," 2019, https://arxiv.org/abs/1910.11030.

[38] G. Kalafatas and S. Peeta, *An Exact Graph Structure for Dynamic Traffic Assignment: Formulation, Properties, and Computational Experience*, Transportation Research Board Meeting, Washington, DC USA, 2007.

[39] Z. Ding, *Diversified Ensemble Classifiers for Highly Imbalanced Data Learning and its Application in bioinformatics*, Georgia State University, Atlanta, GA USA, 2011.