



Traffic Flow Prediction Model Based on Variant Hybrid Multi-Hop Graph Convolution

Xuejun Li¹, Xiaoqi Yin^{1,*}, Chenbin Gu¹, Yudong Yang¹ and Kharudin Bin Ali²

¹School of Electronic Information Engineering, Huaiyin Institute of Technology, Huaian 223003, China

²Faculty of Electrical and Automation Engineering Technology, University College TATI, Kemaman 24000, Malaysia

Abstract

Accurate traffic flow prediction is a crucial step in building an intelligent transportation system, and it is of great significance for alleviating urban traffic congestion and optimizing travel routes. Due to the complex spatial topology of the transportation network and the highly nonlinear temporal dynamic characteristics of the flow data, traditional prediction methods are difficult to fully capture the inherent spatio-temporal dependencies. Therefore, this paper proposes a traffic flow prediction model based on variant hybrid multi-hop graph convolution. Firstly, by introducing a multi-hop graph convolution operator, the model explicitly aggregates the spatial information of multiple-order neighborhoods in the traffic sensor network to capture long-distance spatial dependencies. Secondly, a variant hybrid graph convolution module is designed, which combines Chebyshev polynomial approximation and adaptive adjacency matrix, while balancing computational efficiency and enhancing the model's ability to

mine implicit spatial correlations. In the temporal dimension, we adopt a method combining gated temporal convolution and attention mechanisms to dynamically capture the temporal change patterns of traffic flow. Evaluations across public datasets reveal that our method demonstrates consistent advantages when benchmarked against baseline approaches. The model achieves better performance in short-term and medium-term traffic flow prediction tasks, significantly reducing prediction errors and verifying the effectiveness and advancement of the proposed model.

Keywords: traffic flow prediction, deep learning, spatio-temporal convolutional layer, variant hybrid multi-hop graph convolution.

1 Introduction

With the continuous acceleration of global urbanization, traffic congestion has become a key issue that restricts urban economic development, reduces residents' quality of life, and causes environmental pollution. According to statistics, the average commuting time lost each year by residents in major cities worldwide due to congestion amounts to several tens of hours, and it has also led to significant fuel waste and carbon emissions. In this context, developing intelligent transportation systems is regarded as the core path to achieve refined urban



Academic Editor:

Jianlei Kong

Submitted: 03 October 2025

Accepted: 13 November 2025

Published: 09 April 2026

Vol. 3, No. 2, 2026.

10.62762/TIS.2025.954751

*Corresponding author:

✉ Xiaoqi Yin

hy_xuebao2009@126.com

Citation

Li, X., Yin, X., Gu, C., Yang, Y., & Ali, K. B. (2026). Traffic Flow Prediction Model Based on Variant Hybrid Multi-Hop Graph Convolution. *ICCK Transactions on Intelligent Systematics*, 3(2), 81–93.

© 2026 ICCK (Institute of Central Computation and Knowledge)

traffic governance and improve the operational efficiency of road networks. The effectiveness of many functions of intelligent transportation systems (such as dynamic signal control, congestion warning, and route guidance) highly depends on the precise and forward-looking perception ability of future traffic conditions. However, the transportation system is a typical complex spatiotemporal system, and its state is dynamically influenced by multiple factors such as road network structure, traffic flow interaction, weather events, and even sudden accidents, making accurate prediction extremely challenging. Therefore, developing prediction models that can deeply integrate the spatiotemporal characteristics of road networks and accurately capture the long-range dependency of traffic flow is not only an academic challenge but also a key technical support for achieving the transformation from 'passive response' to 'active control' in traffic governance models, ultimately alleviating urban congestion and improving public travel experience. This research is carried out under the impetus of this practical need.

Based on the actual needs of intelligent transportation system construction, this research focuses on solving the prediction problems in two core scenarios: regional coordinated signal control and intelligent control of expressway ramps. In the regional coordinated signal control scenario, addressing the issue that traditional fixed timing schemes cannot adapt to the dynamic changes of tidal traffic flow, this research aims to develop a model that can provide precise traffic flow predictions for the next 30-60 minutes. This will provide forward-looking decision support for dynamic green wave zones and regional coordinated signal control systems, thereby improving the overall traffic efficiency of the main roads. In the expressway ramp control scenario, to solve the congestion problem on the main road caused by disorderly merging of ramps, this research is dedicated to building a model that can predict the traffic flow state for the next 15-30 minutes, enabling the early identification of potential congestion bottlenecks, and providing technical support for ramp intelligent speed-limiting measures and congestion warning systems. To achieve these application goals, this research will focus on addressing key technical challenges such as long-term spatial dependency modeling and dynamic spatiotemporal correlation capture, ensuring the accuracy and practicality of the model in short-term and medium-to-long-term predictions.

In recent years, the spatio-temporal graph neural

network (STGNN) based on deep learning has become the mainstream paradigm for traffic flow prediction, with its core being the simultaneous modeling of the spatial topology of the road network and the temporal dynamics of traffic flow [1, 2]. STGCN [3], as a pioneering work in this field, constructs a classic spatio-temporal block stacking architecture by combining Chebyshev polynomial approximation spectral graph convolution with gated temporal convolution. However, subsequent research gradually exposed three core technical bottlenecks of this framework: Firstly, the limited receptive field of traditional graph convolution is unable to capture the long-range non-local spatial correlations in the road network [7]. Secondly, the static graph structure defined by a fixed distance threshold cannot accurately depict the dynamic spatial dependencies caused by real-time factors [6]. Finally, the approach of simply stacking convolution layers to expand the receptive field ignores the heterogeneity of information from different hop neighbors, resulting in insufficient modeling ability for mixed multi-hop spatial dependencies [4].

To address these challenges, researchers have proposed improvement schemes from different perspectives. In terms of expanding the receptive field, Graph WaveNet [4] combines multiple graph convolution layers to implicitly learn multi-hop dependencies and introduces an adaptive adjacency matrix to explore hidden spatial relationships. In terms of graph structure learning, STADGCN specifically designs an adaptive dynamic graph convolution module, while Bansal et al. [8] uses an attention mechanism to dynamically learn the strength of node interconnections. In terms of introducing new architectures, models based on Transformer show great potential, such as studies that use graph Transformer to dynamically encode spatio-temporal interactions [9], or develop heterogeneous spatio-temporal graph Transformer to capture long-range dependencies [10]. Additionally, multi-dimensional feature fusion has become a trend, such as MDAS-GNN [11] that integrates multi-dimensional features such as traffic risks. To improve the model's generalization on spatial time series with complex structural dependencies, some works propose structure-preserving contrastive learning frameworks for spatio-temporal sequences [12]. Meanwhile, some studies explore the application of spatio-temporal graph networks in related fields, such as traffic accident prediction [5]

and electric vehicle charging demand prediction [13], which provide cross-domain references for traffic flow prediction.

Based on this research, this paper proposes a traffic flow prediction model based on variant hybrid multi-hop graph convolution (VH-MGCN). The innovative work mainly lies in the following three aspects:

- Firstly, in terms of spatial modeling, we propose the variant hybrid multi-hop graph convolution (VH-MGCN) core module. Unlike Graph WaveNet and other models that simply stack the same convolution layers, this module parallelly aggregates multi-order neighborhood information within a specified hop range, explicitly capturing the mixed spatial dependencies from local to remote, and naturally retains the heterogeneity of features from different hop neighbors.
- Secondly, in terms of graph structure learning, we design an adaptive mechanism that works in conjunction with multi-hop convolution. Unlike methods such as ASTGCN that only adjust weights on a fixed graph structure, we introduce a learnable adaptive adjacency matrix that interacts with the physical graph structure, enabling the model to autonomously learn implicit spatial correlations not given in the data, thereby more accurately depicting dynamic traffic patterns.
- Thirdly, in terms of the model architecture, the integration of these two innovative points constitutes a unified framework that can adaptively capture complex spatio-temporal dependencies.

2 VH-MGCN Traffic Flow Prediction Algorithm

2.1 Graph Convolutional Network (GCN)

GCN aims to extend the convolution operation from regular Euclidean data (such as images) to graph data with non-Euclidean structures. The core idea is to generate new representations of nodes by aggregating the features of the node itself and those of its neighboring nodes. The first-order approximation of spectral graph convolution is one of the most influential works among them. Its layered propagation rule is:

$$H^{(l+1)} = \sigma \left(D^{-\frac{1}{2}} A D^{-\frac{1}{2}} H^{(l)} W^{(l)} \right) \quad (1)$$

where A is the adjacency matrix with self-loops, $D^{-\frac{1}{2}}$ is the degree matrix of A , $H^{(l)}$ is the node feature matrix at the l -th layer, $W^{(l)}$ is the trainable weight matrix, and σ denotes the nonlinear activation function. This formulation clearly illustrates the two key steps of feature transformation and neighborhood aggregation, laying the foundation for numerous subsequent graph neural network models.

2.2 Multi-Head Attention

The attention mechanism enables the model to dynamically assign different importance weights to different parts of the input data, thereby achieving a focus on key information. As illustrated in Figure 1, the scaled dot-product attention is one of its classic forms:

$$\text{Attention}(Q, K, V) = \text{softmax} \left(\frac{QK^T}{\sqrt{d_k}} \right) V \quad (2)$$

where Q , K and V are all obtained by applying linear transformations to the input. The core of the mechanism lies in computing a weight distribution through the interaction between Q and K , followed by a weighted summation over V . In multi-head attention, Q , K and V undergo separate linear mappings, and attention is computed independently for each head; the outputs are then concatenated to form multi-head attention. This allows the model to jointly attend to information from different representation subspaces. Subsequently, this mechanism has been widely adopted in graph neural networks to capture dynamic and non-local dependencies between nodes.

2.3 VH-MGCN Convolutional Layer

The effectiveness of graph convolution technology in the field of spatial feature extraction has been verified. An approximate calculation method for Chebyshev spectrum filters was first proposed. This method simplifies the definition of convolution operation, transforming the originally complex calculation process into a simple neighborhood averaging operator. The specific expression of the graph convolutional layer constructed in [3] can be described as:

$$Z = AXW \quad (3)$$

where $X \in \mathbb{R}^{N \times D}$ and $Z \in \mathbb{R}^{N \times D'}$ denote the input and output node feature matrices, respectively; $W \in \mathbb{R}^{D \times D'}$ is a trainable weight matrix; and $A \in \mathbb{R}^{N \times N}$ is the normalized adjacency matrix with self-loops. As illustrated in Figure 2, this model first

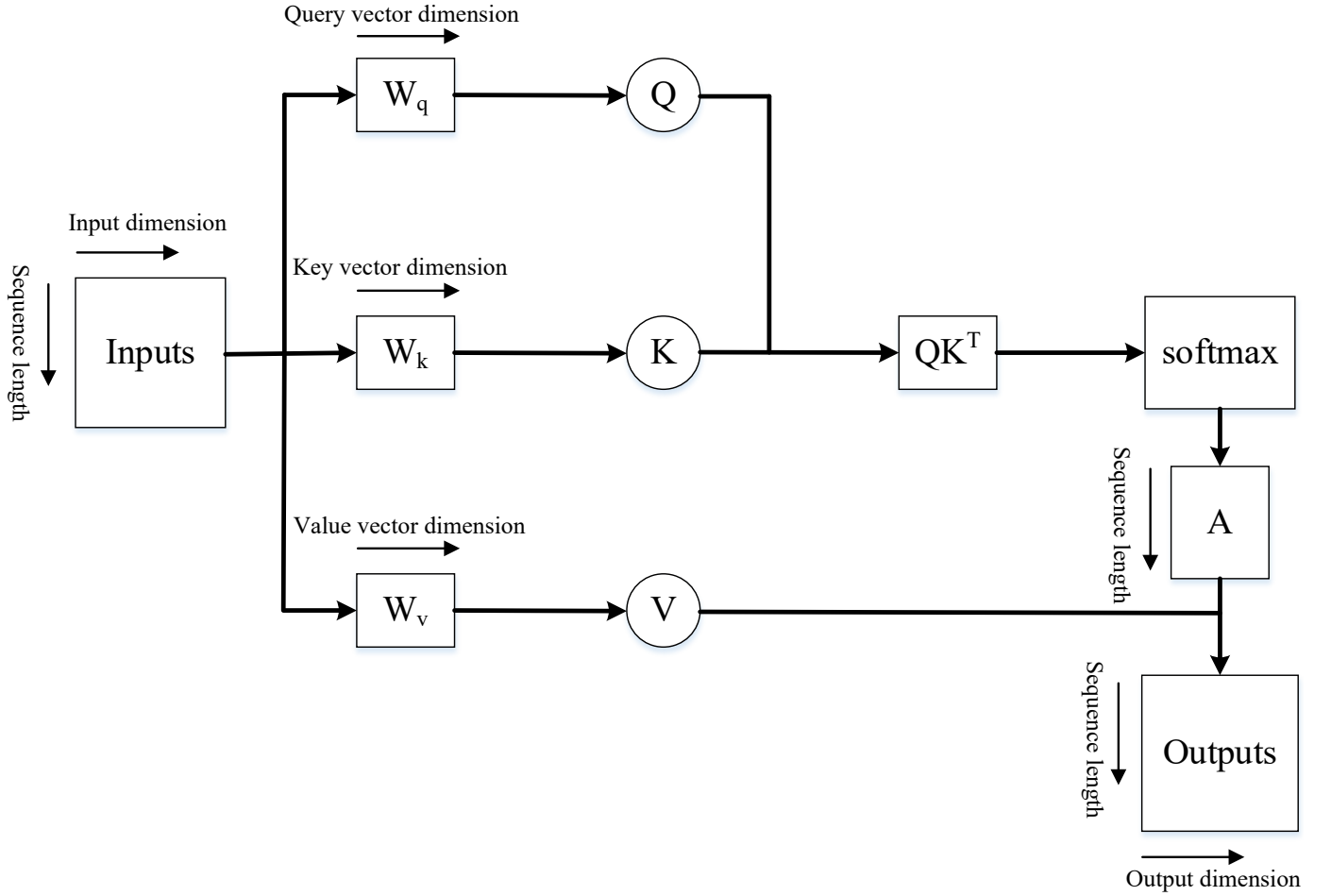


Figure 1. Schematic diagram of the attention mechanism structure.

aggregates features from the first-order neighborhood of each node according to the graph structure, then applies a linear transformation and dimensional mapping to the aggregated features. Although this simplified convolution mechanism significantly reduces computational cost, it also limits the model's expressive power and introduces several associated drawbacks.

(1) **Limited representational capacity.** The simplified graph convolution architecture exhibits substantial representational limitations, as it cannot effectively handle mixed information from multi-hop neighborhoods. Even for the simplest two-hop Delta operator $f(\sigma(AX) - \sigma(A^2X))$, this model lacks a corresponding modeling mechanism.

(2) **Narrow receptive field.** Conventional graph convolution aggregates only first-order neighborhood information within a single layer, resulting in a limited perceptual scope. Obtaining higher-order neighborhood information requires stacking multiple network layers, which inevitably increases model complexity and training difficulty. To

overcome this limitation, Li et al. [1] introduced diffusion-convolution layers in spatio-temporal modeling. This design innovatively integrates the information flow from consecutive K -hop neighborhoods. It can be viewed as an extension of Equation (3), with its mathematical formulation given by:

$$Z = \sum_{k=0}^K P^k X W_k \quad (4)$$

This technical solution first performs operations on the node feature matrix and the k -th power of the adjacency matrix separately, thereby completing the feature aggregation of the neighborhood node information within the k -hop range of each node. Subsequently, by introducing a trainable parameter matrix, a linear transformation operation is performed on the aggregated internal node features. Finally, the sum operation is executed on the representation vectors generated by different hop counts.

Here, the power operation of the adjacency matrix needs to distinguish between two network topologies:

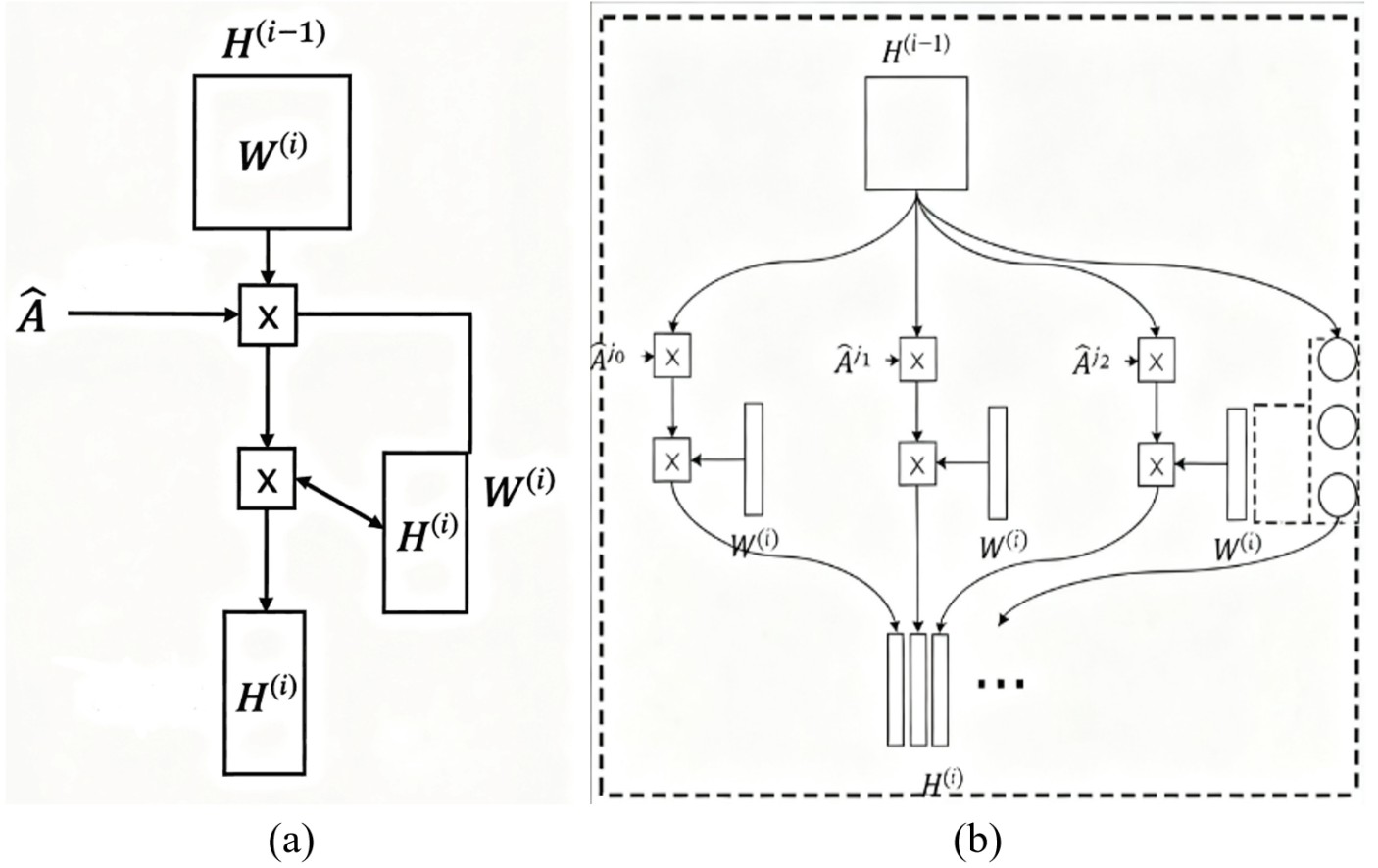


Figure 2. (a) Traditional GC layer. (b) Our GC layer.

in the undirected graph scenario, the calculation method is the standard adjacency matrix power operation; while in the directed graph environment, information diffusion needs to consider two directions: forward propagation

$$P_j^f = A^j / \text{ROWSUM}(A^j)$$

and backward propagation

$$P_j^b = (A^\top)^j / \text{ROWSUM}((A^\top)^j).$$

At this time, the forward transition matrix and the backward transition matrix need to be defined separately to complete the calculation.

Li et al. [1] diffusion convolutional architecture imposes a predefined constraint on the extraction of spatial dependencies, limiting its learning capacity to K-step neighborhood interactions. This design has certain rationality in conventional semi-supervised learning tasks because such tasks usually rely more on local low-order neighborhood information.

Capturing long-range spatial-temporal dependencies becomes crucial in traffic forecasting, as the rapid displacement of vehicles means that high-order neighborhood connections often provide essential contextual cues for medium and long-term prediction horizons.

To break through this limitation, the MixHop graph convolution architecture removed the order constraint of the diffusion convolution, enabling the model to simultaneously aggregate neighborhood information of any order, and only requiring the use of two layers of network stacking to achieve collaborative learning of mixed neighborhood features:

$$H^{(i+1)} = \bigcup_{j \in Q} \sigma \left(p_j^f H^{(i)} W_{j_1}^{(i)} + p_j^b H^{(i)} W_{j_2}^{(i)} \right) \quad (5)$$

In this architecture, the hyperparameter Q serves as the configuration set of the receptive field, enumerating the orders of neighborhoods over which feature aggregation is performed. Let $H^{(i)}$ denote the output of the i -th layer, σ be an optional activation function, and $\bigcup$ represent the concatenation of features along the channel dimension. The contrast with conventional

graph convolution is illustrated in Figure 2: when Q is the singleton $\{1\}$, the model reduces to conventional graph convolution; when Q is the consecutive integer set $\{0, 1, \dots, K\}$, it becomes functionally equivalent to dilated convolution. To capture potential spatial dependencies not explicitly represented by the original graph structure, this paper introduces an adaptive adjacency matrix, defined tensorially as:

$$\hat{A}_{\text{adp}} = \text{SoftMax}(\text{ReLU}(E_1 E_2^T)) \quad (6)$$

where $E_1 \in \mathbb{R}^{N \times C}$ and $E_2 \in \mathbb{R}^{N \times C}$ are source-node and target-node embeddings learned from the topology learner, respectively. The node embeddings E_1 and E_2 are randomly initialized and jointly optimized with the full model parameters during end-to-end training. Their product yields a matrix of potential connection strengths between nodes. A non-linear filter via ReLU removes insignificant connections, followed by stochastic normalization through SoftMax to form an adaptive adjacency matrix represented as a probability distribution.

This topology-learning mechanism is then integrated into the MixHop architecture, resulting in the following improved graph-convolutional layer:

$$H^{(i+1)} = \bigcup_{j \in Q} \sigma \left(p_j^f H^{(i)} W_{j1}^{(i)} + p_j^b H^{(i)} W_{j2}^{(i)} + A_{\text{adp}} H^{(i)} W_{j3}^{(i)} \right) \quad (7)$$

In the module design, two MixHop graph-convolutional layers are cascaded to model mixed-neighborhood information. Finally, a fully-connected layer is appended to transform the output, thereby defining the complete graph-convolutional computation unit.

2.4 VH-MGCN Model Framework

As shown in Figure 3, The overall design of the VH-MGCN model follows the encoder-decoder (or stack of temporal-spatial modules) paradigm. The most important core innovation of this model lies in its spatial dependency modeling unit - the variant hybrid multi-hop graph convolution layer. Compared with the single graph convolution used by STGCN, the diffusion convolution used by DCRNN, and the implicit multi-hop stacking of Graph WaveNet, the VH-MGCN layer designed in this paper achieves two major breakthroughs: 1) Explicit and parallel multi-hop feature fusion, which directly controls the receptive field range and distinguishes the importance of different neighborhood orders; 2) Collaborative modeling of static topology and dynamic implicit

association, by compensating for the shortcomings of predefined graphs through an adaptive adjacency matrix.

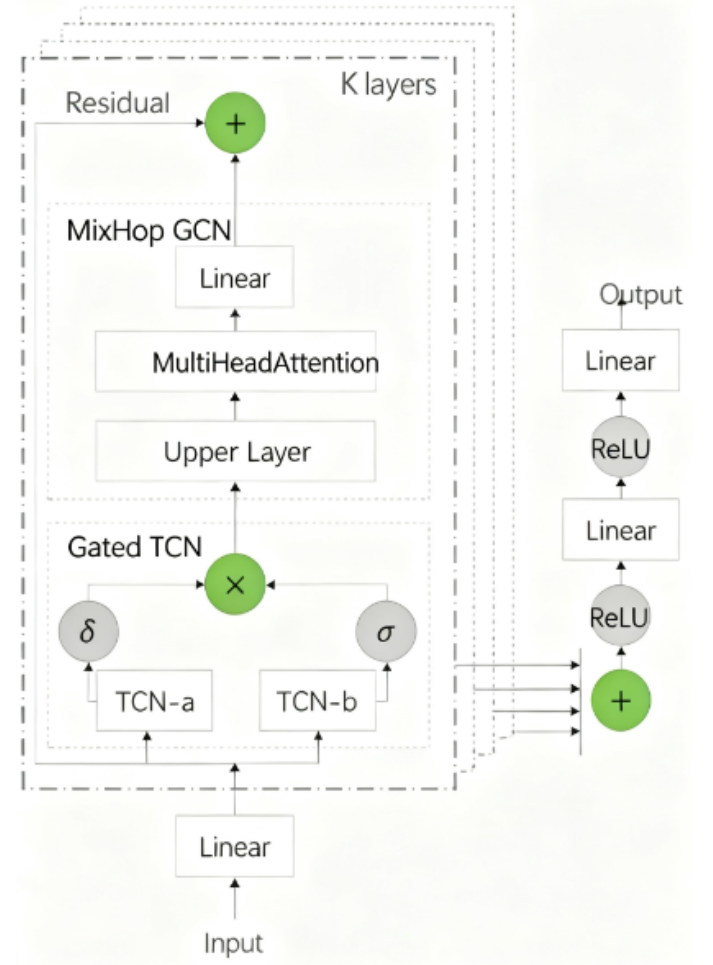


Figure 3. The framework of VH-MGCN.

The calculation formula of VH-MGCN consists of a spatio-temporal module composed of MixHop graph convolution and gated temporal convolution, which are responsible for spatial dependency modeling and time pattern mining respectively. The temporal component adopts a parallel design of dual-path expansion causal convolution (TCN-a/TCN-b), while the spatial component includes two iterations of MixHop graph convolution and dimension projection operations. Through the stacking of multiple spatio-temporal layers, multi-scale feature extraction is formed: the shallow layer extracts local temporal patterns, and the deep layer captures global temporal dependencies. The outputs of all spatio-temporal layers are fed back to the prediction layer after being summed by the scalar operation. The model uses minimizing the mean absolute error (MAE) as the training criterion, and its mathematical formalization

is as follows:

$$\text{Loss} = \frac{1}{TND} \sum_{i=1}^T \sum_{j=1}^N \sum_{k=1}^D \left| Y_{jk}^{(i)} - \hat{Y}_{jk}^{(i)} \right| \quad (8)$$

where $Y \in \mathbb{R}^{T \times N \times D}$ denotes the ground truth, and $\hat{Y} \in \mathbb{R}^{T \times N \times D}$ corresponds to the predicted values.

3 Experiment Results

3.1 Experimental dataset

This experiment collected data with a time granularity of 5 minutes. The topological connectivity between nodes was based on road network distances, and the quantification was carried out using a threshold-based Gaussian kernel function. The data set was divided into three parts according to the time sequence: the training set (70%), the validation set (10%), and the test set (20%). During the model training stage, the training data was randomly rearranged, and the Z-score normalization method was applied to eliminate the dimensional differences among features. The data set is shown in Table 1.

Table 1. METR-LA and PEMS8.

Dataset	Nodes	Edges	Time Steps
METR-LA	207	1515	34272
PEMS8	170	295	17856

3.2 Implementation Details

All experiments were implemented in PyTorch 2.1 and trained on a single NVIDIA GeForce RTX 3090 GPU with 24 GB memory, running Ubuntu 20.04 with Python 3.9 and CUDA 11.8. The Adam optimizer was used with an initial learning rate of 1×10^{-3} , which decays by a factor of 0.5 every 10 epochs. The model was trained for 100 epochs with a batch size of 64. Early stopping with a patience of 15 epochs was applied based on validation MAE. The receptive field configuration set is $Q = \{1, 2, 3\}$, the number of spatio-temporal layers is 4, the hidden dimension is 64, and the node embedding dimension for the adaptive adjacency matrix is 10. Dropout with a rate of 0.3 was applied after each graph convolution operation. The input sequence length is 12 time steps (60 minutes), and the prediction horizon covers 3, 6, and 12 steps, corresponding to 15, 30, and 60 minutes respectively. All baseline models were reproduced using their official implementations with the hyperparameters reported in their respective papers.

3.3 Introduction to the Baseline Model

To test the performance of the model, this section will compare the VH-MGCN model with the following models:

- STGCN [3]: The spatio-temporal graph convolutional network combines gated temporal convolution and graph convolution.
- DCRNN [1]: Utilizing the encoder-decoder method in combination with recurrent neural networks and graph convolutional networks based on random walks.
- ARIMA [14]: The autoregressive integrated moving average model belongs to the time series prediction model.
- WaveNet [15]: A dilated causal convolutional architecture originally proposed for audio generation, whose temporal sequence modeling design has been adapted as a baseline for evaluating time-series prediction performance in traffic forecasting.
- Graph WaveNet [4]: Combining graph convolution with dilated causal convolution and proposing an adaptive adjacency matrix.
- MRA-BGCN [7]: A multi-range attention dual-component graph convolutional network was proposed.
- ASTGCN [16]: Spatial-temporal graph convolutional network based on attention mechanism.
- LSGCN [17]: Using graph convolution to handle spatial dependencies, and proposing Gated Linear Unit (GLU) to handle temporal dependencies.

3.4 Experimental results and analysis

3.4.1 VH-MGCN Ablation Experiments

We set up the following baseline models to compare with the VH-MGCN model:

Ablation Model A: Without Multi-Hop (w/o Multi-Hop): Replace the hybrid multi-hop graph convolutional layer in the VH-MGCN model with a regular single-hop graph convolutional layer. The purpose is to verify the necessity of the innovative point of explicitly parallel aggregation of multi-order neighborhood information.

Ablation Model B: Without Adaptive Graph (w/o Adaptive Graph): Remove the adaptive adjacency

Table 2. Results of the ablation experiments on the METR-LA dataset.

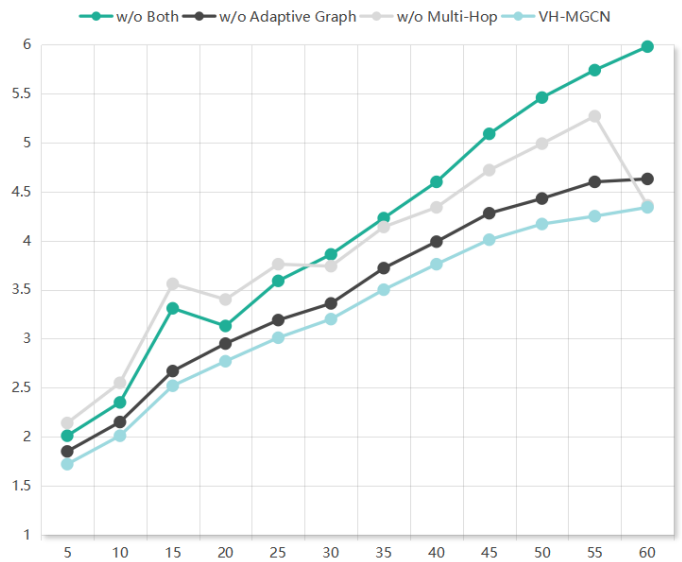
Model	15 min			30 min			60 min		
	MAE	RMSE	MAPE (%)	MAE	RMSE	MAPE (%)	MAE	RMSE	MAPE (%)
VH-MGCN (Ours)	3.01	5.93	6.99	3.34	7.09	8.50	3.86	7.59	9.89
w/o Multi-Hop	3.05	6.01	7.12	3.52	7.35	8.95	4.25	8.12	10.85
w/o Adaptive Graph	3.08	6.08	7.20	3.45	7.28	8.78	4.01	7.82	10.12
w/o Both	3.15	6.20	7.35	3.65	7.52	9.20	4.52	8.45	11.40

matrix A_{adp} and only retain the static adjacency matrix based on road network distance. The purpose is to verify the effectiveness of the data-driven adaptive adjacency matrix in capturing implicit spatial dependencies.

Ablation Model C: Without Multi-Hop and Without Adaptive Graph (w/o Both): Make both modifications simultaneously to obtain a baseline model that only uses single-hop static graph convolution. The purpose is to verify the synergy of the two innovative points.

The results of the ablation experiments (see Table 2) clearly validate the effectiveness of each innovative component in the VH-MGCN model.

(1) The core role of the multi-hop mechanism: Compared with the complete model, the model without the multi-hop mechanism (w/o Multi-Hop) has a performance decline on all prediction time scales, especially in the 60-minute long-term prediction, where the MAE index has significantly increased by approximately 10.1%. This fully demonstrates that explicitly and concurrently aggregating multi-order neighborhood information is crucial for capturing the long-range propagation and global trend of traffic flow, and it compensates for the fundamental defect of the traditional single-hop graph convolution with a limited receptive field.

**Figure 4.** Comparison of the prediction results of the ablation experiment.

(2) The stabilizing contribution of the adaptive graph: The model without the adaptive adjacency matrix (w/o Adaptive Graph) also shows significant performance degradation, indicating that learning implicit spatial dependencies from the data can stably improve the model performance.

(3) The synergy effect of components: The model with the worst performance is the one that simultaneously

Table 3. Performance comparison of VH-MGCN and other baseline models.

Data	Model	15 min			30 min			60 min		
		MAE	RMSE	MAPE (%)	MAE	RMSE	MAPE (%)	MAE	RMSE	MAPE (%)
METR-LA	ARIMA	2.85	7.62	8.23	5.10	9.73	10.73	7.30	10.98	12.40
	WaveNet	2.81	6.72	8.12	3.83	7.80	9.53	4.73	8.21	13.40
	DCRNN	2.89	5.27	7.95	3.52	7.03	9.22	4.70	8.39	11.20
	LSGCN	2.82	5.33	7.77	3.86	7.01	8.99	4.33	8.17	10.87
	GWN	3.09	5.99	7.62	3.95	6.35	8.72	4.20	8.29	9.92
	MRA-BGCN	3.22	5.88	6.97	3.58	6.84	8.65	3.67	7.89	10.00
	Ours	2.79	5.25	6.93	3.34	6.35	8.50	3.51	7.69	9.89
PEMSD8	DCRNN	1.39	2.87	2.65	1.78	3.89	4.23	2.30	5.63	4.20
	STGCN	1.23	2.79	2.38	1.95	3.72	4.11	2.69	5.23	3.95
	ASTGCN	1.41	3.09	2.97	1.52	3.75	4.06	1.92	5.03	4.02
	Ours	1.20	2.65	2.29	1.49	3.48	3.21	1.88	4.82	3.55

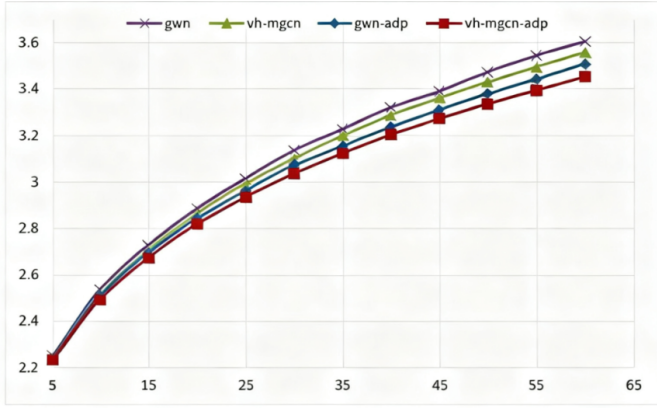


Figure 5. Performance comparison between VH-MGCN and GWN.

removes two innovative components (w/o Both), which indicates that multi-hop perception and dynamic graph learning are mutually reinforcing and working together. They respectively enhance the spatial modeling ability from the two dimensions of 'receptive field breadth' and 'association relationship accuracy', and their combination produces an effect of $1+1>2$, jointly constituting the core of the superior performance of the VH-MGCN model. The comparative performance of these ablation models across different prediction horizons is visually summarized in Figure 4.

3.4.2 Performance comparison of VH-MGCN with other baseline models

A quantitative assessment of prediction errors is presented in Table 3, which offers a comparative analysis of the proposed framework against established baselines over a range of time intervals, 15-minute, 30-minute and 60-minute predictions (or performances). It is evident that the model proposed in this paper has achieved better performance results on both datasets.

The experimental data confirm that the VH-MGCN model proposed in this study has achieved multiple breakthroughs in prediction performance. Compared with the classic ARIMA model, in the 15/30/60-minute prediction tasks on the METR-LA dataset, the MAE index has decreased. In the horizontal comparison with the family of spatial-temporal graph convolutional networks, this model has achieved stable improvements on top of the benchmark models such as Graph WaveNet, STGCN and ASTGCN. Compared with the optimal benchmark model MRA-BGCN, VH-MGCN shows better performance at 15 minutes, while its MAE values are 0.24 and 0.16 ahead in the 30-minute and 60-minute predictions

respectively. Although MRA-BGCN also explores multi-neighborhood information fusion, the attention mechanism it adopts is inferior to the variant hybrid convolution architecture designed in this study in terms of accuracy and efficiency. These comparisons fully highlight the unique value of VH-MGCN in long-term prediction scenarios.

In the architecture design of the graph convolution layer, the core contribution of this paper lies in proposing a variant hybrid multi-hop graph convolution (VH-MGCN), aiming to significantly enhance the model's ability to perceive high-order neighborhoods and mixed multi-hop information. This design is an extension and deepening of the Graph WaveNet framework, and innovatively introduces an adaptive adjacency matrix (VH-MGCN-Adp) as its supplementary mechanism. As shown in Figure 5, in the 5-minute short-term prediction, VH-MGCN did not bring significant performance improvement compared to the baseline model. However, as the prediction window was extended to 60 minutes, its advantages gradually became prominent. This phenomenon can be attributed to the differences in spatial dependency at different time scales: in short-term prediction, local low-order neighborhood information is mainly relied upon, and introducing complex multi-hop operations at this time may actually interfere with the extraction of key features, causing VH-MGCN to degrade to ordinary graph convolution; while in long-term prediction, local information is insufficient to provide sufficient support for inference, and at this time, VH-MGCN can adaptively enhance the contribution weights of high-order neighborhood information, thereby more effectively modeling long-range spatial dependencies and obtaining more accurate prediction results.

3.5 Comparative Experiment

To verify the predictive ability of the VH-MGCN model in real-world scenarios, a representative time series regression prediction model based on Transformer was selected for comparative experiments. Transformer uses the self-attention mechanism to globally model the sequence dependencies, aiming to verify the performance and computational efficiency of the proposed model in complex dependency patterns. This study conducted a visual analysis of the output results of the model on two real datasets. Figures 6, 7, 8 and 9 respectively show the comparison of the predicted values and the actual values of a randomly selected single node at different time scales. It can be

seen that the predicted curves and the actual curves are highly consistent in their changing trends, despite the different lengths of the time series. The model has a stable ability to track the fluctuation patterns of traffic speed, which fully validates the accuracy and reliability of VH-MGCN in real-time prediction applications.

To more intuitively evaluate the predictive performance of the model, Figures 6, 7, 8 and 9 show the comparison of the prediction effects of VH-MGCN and the Transformer model in typical time periods of the test set. It can be observed that in the morning rush hour when the traffic conditions undergo drastic changes, the VH-MGCN model proposed in this paper can more accurately track the dynamic changes of the true values. In particular, it significantly outperforms the Transformer model in capturing the peak of traffic flow and responding to sudden changes. In contrast, the prediction curve of the Transformer model shows a clear smoothing effect, lagging behind in responding to rapid changes in traffic flow, and underestimating the peaks in many places.

This visualization result qualitatively confirms the advantages of the model in this paper: VH-MGCN, through its innovative variant of the hybrid multi-hop graph convolution structure, can effectively capture the long-range spatial propagation characteristics of traffic flow in the road network, thereby responding more sensitively to regional traffic state changes. Although the Transformer model has advantages in modeling the time dimension, it has limitations in handling complex spatial dependencies, which makes it difficult to accurately predict the sudden changes in traffic flow.

4 Conclusion

This paper proposes a traffic flow prediction model based on Variant Hybrid Multi-hop Graph Convolution (VH-MGCN). The main theoretical and application breakthroughs of this work lie in the following: Firstly, from the model architecture, an explicit parallel multi-hop graph convolution operator is defined and implemented, providing a new solution for how to efficiently and differently fuse multi-order neighborhood information in graph neural networks; Secondly, from the methodology, the physical-guided static graph and data-driven dynamic



Figure 6. Prediction results of Transformer on PEMS8 at 5, 15, 30 and 60 minute horizons.

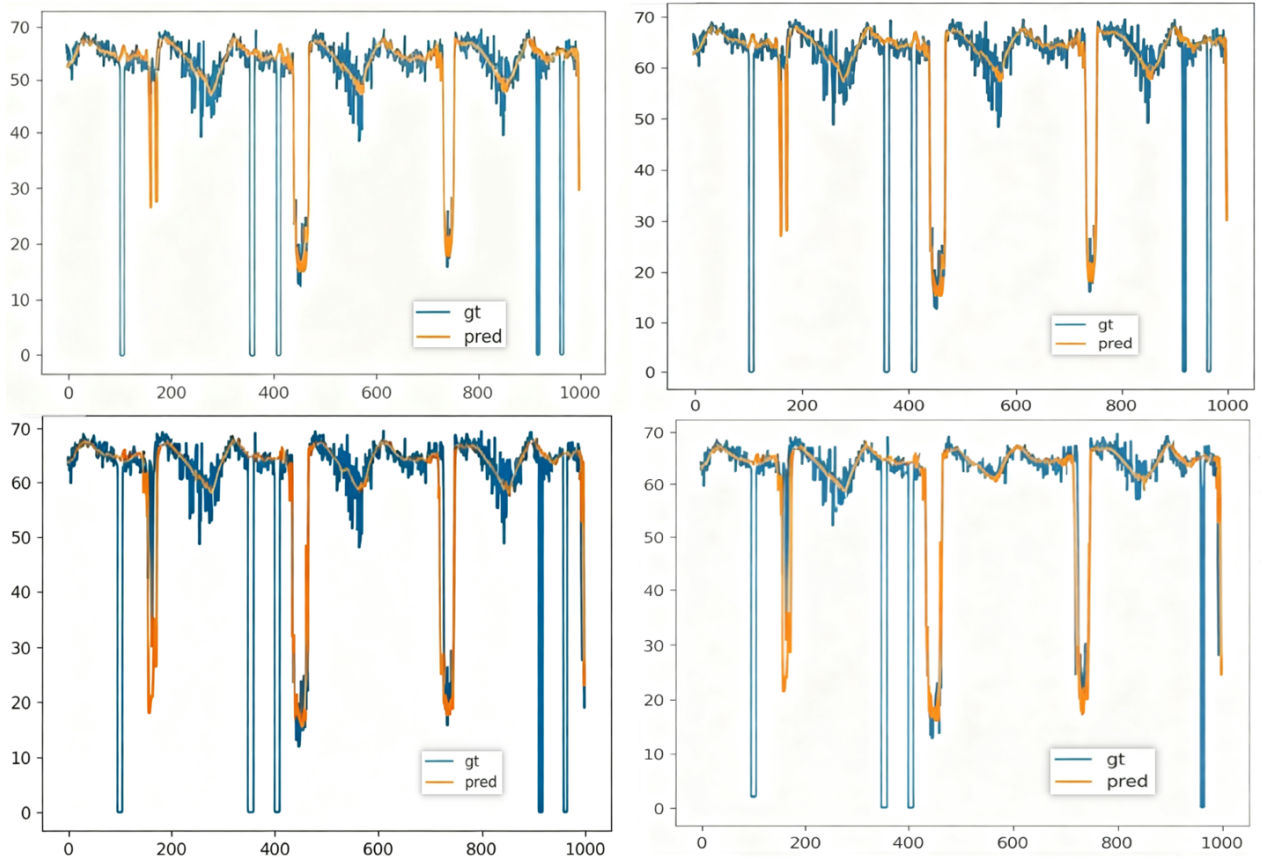


Figure 7. Prediction results of Transformer on METR-LA at 5, 15, 30 and 60 minute horizons.

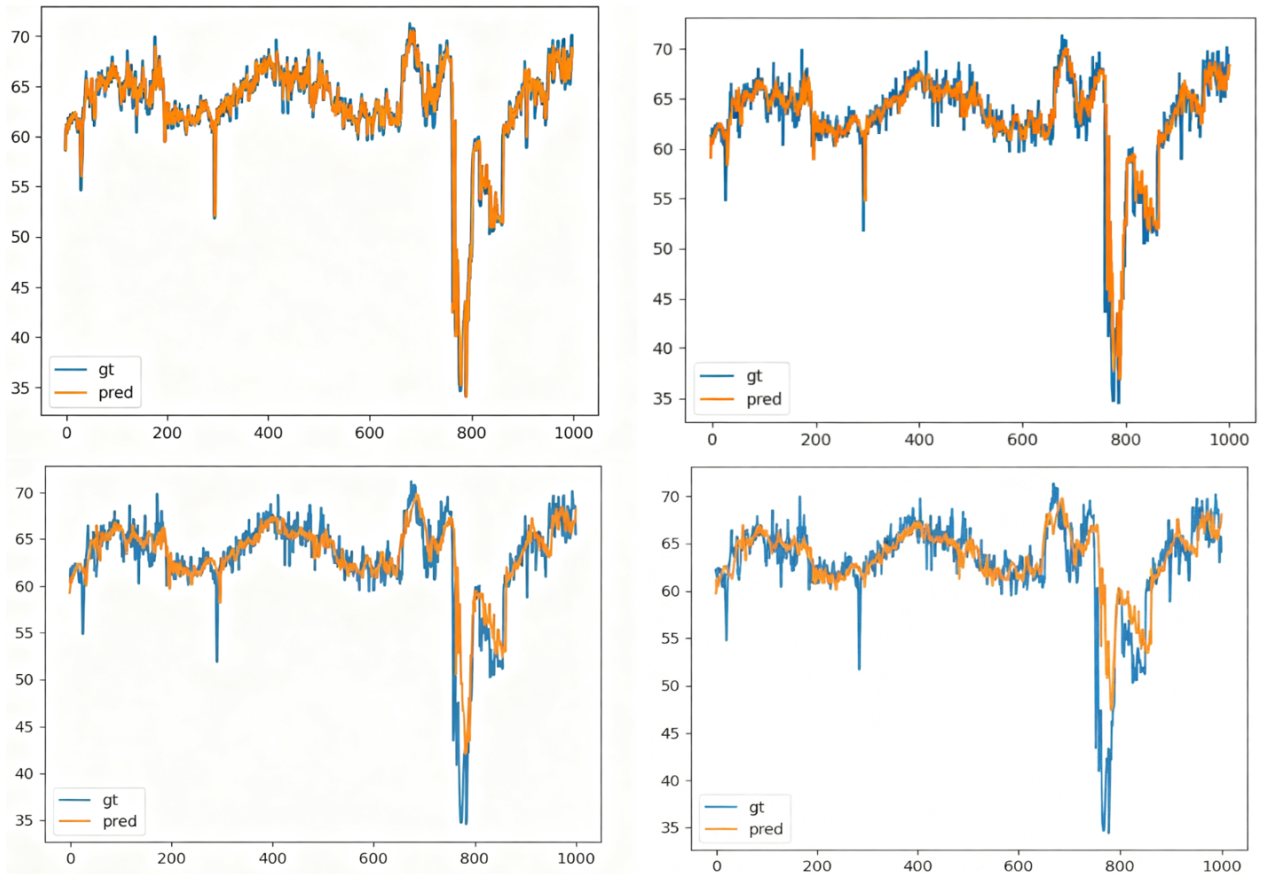


Figure 8. Prediction results of VH-MGCN on PEMS8 at 5, 15, 30 and 60 minute horizons.

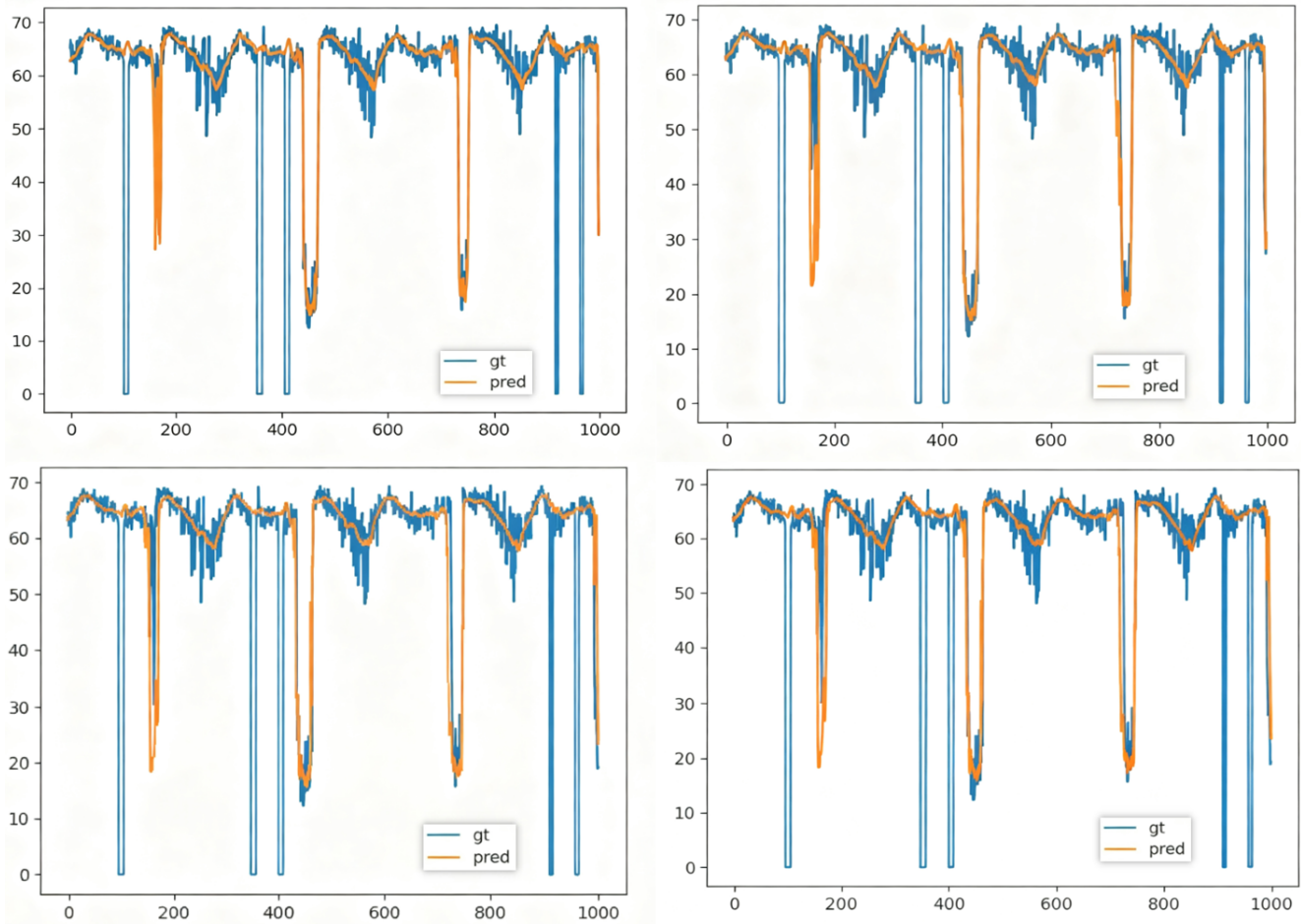


Figure 9. Prediction results of VH-MGCN on METR-LA at 5, 15, 30 and 60 minute horizons.

graph have been successfully trained collaboratively, enhancing the model's generalization ability to the complex spatiotemporal dynamics of the real world. Experiments on multiple real datasets not only prove the superior performance of the model, but also confirm the effectiveness of each innovative module through ablation studies. The comparison with multiple spatiotemporal graph convolution baselines further confirms that VH-MGCN not only stably improves the consistency of predictions, but also has significant advantages in accuracy and robustness.

Data Availability Statement

Data will be made available on request.

Funding

This work was supported without any funding.

Conflicts of Interest

The authors declare no conflicts of interest.

AI Use Statement

The authors declare that no generative AI was used in the preparation of this manuscript.

Ethical Approval and Consent to Participate

Not applicable.

References

- [1] Li, Y., Yu, R., Shahabi, C., & Liu, Y. (2017). Diffusion convolutional recurrent neural network: Data-driven traffic forecasting. *arXiv preprint arXiv:1707.01926*.
- [2] Jia, T., & Yan, P. (2020). Predicting citywide road traffic flow using deep spatiotemporal neural networks. *IEEE Transactions on Intelligent Transportation Systems*, 22(5), 3101-3111. [[CrossRef](#)]
- [3] Yu, B., Yin, H., & Zhu, Z. (2018, July). Spatio-temporal graph convolutional networks: a deep learning framework for traffic forecasting. In *Proceedings of the 27th International Joint Conference on Artificial Intelligence* (pp. 3634-3640).
- [4] Wu, Z., Pan, S., Long, G., Jiang, J., & Zhang, C. (2019, August). Graph wavenet for deep spatial-temporal

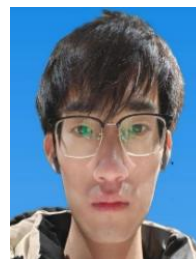
- graph modeling. In *Proceedings of the 28th International Joint Conference on Artificial Intelligence* (pp. 1907-1913).
- [5] Yu, L., Du, B., Hu, X., Sun, L., Han, L., & Lv, W. (2021). Deep spatio-temporal graph convolutional network for traffic accident prediction. *Neurocomputing*, 423, 135-147. [CrossRef]
- [6] Shi, Y., Cui, W., Wang, R., Lou, J., & Shen, Q. (2025). STADGCN: spatial-temporal adaptive dynamic graph convolutional network for traffic flow prediction. *Neural Computing and Applications*, 1-14. [CrossRef]
- [7] Chen, W., Chen, L., Xie, Y., Cao, W., Gao, Y., & Feng, X. (2020, April). Multi-range attentive bicomponent graph convolutional network for traffic forecasting. In *Proceedings of the AAAI conference on artificial intelligence* (Vol. 34, No. 04, pp. 3529-3536). [CrossRef]
- [8] Bansal, N., & Bali, R. S. (2026). Graph Convolutional Network and Attention for Traffic Prediction-A Deep Learning Approach. *Recent Advances in Computer Science and Communications*, 19(2), E26662558351567. [CrossRef]
- [9] Ouyang, N., Ao, L., Cai, Q., Wan, W., Ren, X., He, X., & Sheng, K. (2024). Graph transformer-based dynamic edge interaction encoding for traffic prediction. *IEEE Transactions on Intelligent Transportation Systems*. [CrossRef]
- [10] Shi, G., Luo, L., Song, Y., Wang, H., & Zhang, Y. (2024). Deep transformer-based heterogeneous spatiotemporal graph learning for geographical traffic forecasting. *iScience*, 27(1), 108-122. [CrossRef]
- [11] Gao, Z. (2025). MDAS-GNN: Multi-Dimensional Spatiotemporal GNN with Spatial Diffusion for Urban Traffic Risk Forecasting. *arXiv preprint arXiv:2510.27197*.
- [12] Jiao, Y., van Cranenburgh, S., Calvert, S., & van Lint, H. (2025). Structure-preserving contrastive learning for spatial time series. *arXiv preprint arXiv:2502.06380*.
- [13] Tupayachi, J., Camur, M. C., Heaslip, K., & Li, X. (2025). Spatio-Temporal Graph Convolutional Networks for EV Charging Demand Forecasting Using Real-World Multi-Modal Data Integration. *arXiv preprint arXiv:2510.09048*.
- [14] Katambire, V. N., Musabe, R., Uwitonze, A., & Mukanyiligira, D. (2023). Forecasting the traffic flow by using ARIMA and LSTM models: case of Muhima junction. *Forecasting*, 5(4), 616-628. [CrossRef]
- [15] Van Den Oord, A., Dieleman, S., Zen, H., Simonyan, K., Vinyals, O., Graves, A., ... & Kavukcuoglu, K. (2016). Wavenet: A generative model for raw audio. *arXiv preprint arXiv:1609.03499*.
- [16] Guo, S., Lin, Y., Feng, N., Song, C., & Wan, H. (2019, July). Attention based spatial-temporal graph convolutional networks for traffic flow forecasting. In *Proceedings of the AAAI conference on artificial intelligence* (Vol. 33, No. 01, pp. 922-929). [CrossRef]
- [17] Huang, R., Huang, C., Liu, Y., Dai, G., & Kong, W. (2020, July). LSGCN: Long short-term traffic prediction with graph convolutional networks. In *Proceedings of the 29th International Joint Conference on Artificial Intelligence* (Vol. 7, pp. 2355-2361). [CrossRef]



Xuejun Li was born in Zhoukou, Henan Province, China. He received the B.S. degree in Electronic Science and Technology from Huaiyin Institute of Technology, China. He is currently pursuing the M.S. degree at Huaiyin Institute of Technology. His research interests include traffic information engineering and control. (Email: lixuejun1013@163.com)



Xiaoqi Yin is a Professor and Master's Supervisor at Huaiyin Institute of Technology, China. She has long been engaged in research on signal processing and measurement technology. She is the corresponding author of this paper. (Email: hy_xuebao2009@126.com)



Chenbin Gu was born in Huai'an, Jiangsu Province, China. He received the B.S. degree in Electronic Science and Technology from Huaiyin Institute of Technology, China. He is currently pursuing the M.S. degree at Huaiyin Institute of Technology. His research interests include traffic information engineering and control. (Email: gcb31830127@126.com)



Kharudin Bin Ali received the Diploma degree in Industrial Automation (Mechatronics) and the bachelor's degree (Hons.) in Engineering Technology (Mechatronics) from TATI University College in 2005 and 2012, respectively. He received the M.Sc. degree in Mechatronic Engineering from University Malaysia Pahang in 2017, and the Ph.D. degree from University Tenaga Nasional (UNITEN), Malaysia, in 2020. He became a Professional Technologist in 2018. He has published a total of 33 papers in ISI- and Scopus-indexed journals and international conferences and has supervised around 85 graduate students and two postgraduate (master's) students. He has also published one book on eddy current inspection.

His current research interests include non-destructive testing, sensor system design, embedded systems, the Internet of Things, and artificial intelligence. He has extensive experience in industrial work, teaching, and training since 2004. His industrial experience includes ESCATEC Mechatronics and TATI University College. He has also received several awards, including the Best Paper Award at the ICON 2013 Conference and the IGRAD 2018 Conference, as well as the Excellent Worker Award from TATI University College in 2012 and 2018. (Email: kharudin@uctati.edu.my)