HAG-MTF: Higher-Order Adaptive Generative Graph for Massive Traffic Forecasting in Industry 5.0

LEI WANG*, Tianjin University, China
HUAMING WU†, Tianjin University, China
FAN ZHANG*, Sino-Singapore Tianjin Eco-City Data Bureau, China
KEQIU LI, Tianjin University, China
WEI YU†, Zhejiang Yuexiu University of Foreign Languages, China
SHUO CHEN, Tianjin Shenglian Intelligent Technology Development Co., Ltd., China

With the evolution of urban smart transportation, the complexity of urban traffic networks escalates, emphasizing the importance of large-scale traffic data prediction in traffic management and urban planning. Traditional spatiotemporal graph models, such as Graph-WaveNet and MTGCN, face exponentially increasing computational complexity as the spatial dimensions expand. To address this challenge, we propose a novel Higher-order Adaptive Generative graph for Massive Traffic Forecasting (HAG-MTF) approach, which utilizes generative AI and high-order graph structures to model the intricate spatial dependencies in large-scale traffic data. The HAG-MTF incorporates a high-order dimensionality reduction module to optimize traffic node processing, utilizing prior graph relationships to generate a fusion graph that dynamically incorporates neighborhood information for efficient, localized graph convolution. The model further incorporates the high-order spatiotemporal relationship extraction module (H-net), enhancing the capacity and speed of traffic data processing while boosting prediction accuracy for complex spatial structures. Furthermore, HAG-MTF introduces a fusion loss function that hierarchically balances multiple objectives, ensuring both precision and computational efficiency. HAG-MTF adaptively handles large-scale real-world traffic data, meeting the needs of traffic controllers and urban planners for predicting massive datasets in practical settings. It supports efficient, flexible interactions via parameter tuning and model outputs, ultimately integrating human insights into traffic analysis and decision-making. This dynamic human-machine collaboration differs from non-Industry 5.0 approaches, which rely on purely automated systems without human input. Those lead to inflexible, brittle conclusions and recommendations, neglecting shifts in traffic patterns driven by human behavior. Extensive experiments on real-world traffic datasets demonstrate that HAG-MTF significantly improves processing efficiency for high-complexity spatial data while delivering precise, human-informed predictions through generative AI-driven operations.

CCS Concepts: • **Networks** → **Traffic engineering algorithms**; • **Theory of computation** → **Dynamic graph algorithms**.

Additional Key Words and Phrases: Spatiotemporal big data, Traffic prediction, High-order generative graph convolution

*Both authors contribute equally to this research.
†Corresponding author.

Authors' Contact Information: Lei Wang, Tianjin University, Tianjin, China, wanglei2019@tju.edu.cn; Huaming Wu, Tianjin University, Tianjin, China, whming@tju.edu.cn; Fan Zhang, Sino-Singapore Tianjin Eco-City Data Bureau, Tianjin, China, zhangf@eco-city.gov.cn; Keqiu Li, Tianjin University, Tianjin, China, keqiu@tju.edu.cn; Wei Yu, Zhejiang Yuexiu University of Foreign Languages, Shaoxing, China, weiyu@zyufl.edu.cn; Shuo Chen, Tianjin Shenglian Intelligent Technology Development Co., Ltd., Tianjin, China, 18630631289@163.com.

ACM 1556-4703/2025/10-ART
https://doi.org/10.1145/3772723

1 INTRODUCTION

The prediction of traffic flow is crucial for various applications such as congestion forecasting, vehicle flow measurement, and autonomous driving [15]. This field has evolved from traditional methods involving feature engineering and time-series modeling to more advanced approaches based on Graph Neural Networks (GNNs). However, the practical implementation of these models necessitates the integration of high-frequency data from extensive sensor networks. For instance, a city with 100 intersections may generate over 10 million daily data points for second-level traffic control, presenting significant computational challenges. Meanwhile, complex real-world traffic conditions, such as traffic congestion, traffic accidents, and rain weather disturbances [7], introduce increasingly high-dimensional features into traffic data, leading to a substantial rise in data dimensionality. To address this, current techniques often rely on downsampling or data segmentation to manage the complexity. However, this comes at the expense of data integrity and the challenge of balancing efficiency with real-world demands.

The rapid growth of urban traffic networks also presents increasing challenges to traffic data analysis and prediction. For example, London has over 5719 lane sensors as indicated in Table 1. This trend especially affects models that handle large temporal scales, such as involving over 50,000 time steps and 1,000 nodes spatial scales.

Table 1. **Dataset Descriptions**

Overview of traffic and spatiotemporal datasets by city, sensors, time span, interval, and data size.

Dataset	City	Sensors/Lanes	Time Span	Interval (min)	Data Size
METR-LA	Los Angeles	207	2012-2017	5	3,011,904,000
PEMS-BAY	Bay Area	325	2017	5	1,617,408,000
Solar Power	Phoenix	5 (e.g., irradiance, temp)	2017-2020	15	6,727,680
Traffic	Los Angeles	4	2015-2017	60	5,045,760
PeMS04	California	384	2018	5	3,822,059,520
PeMS08	California	170	2016	5	846,028,800
UTD19	London	5719	2015-2016	5	41,803,776,000

Challenges in accessing traffic trends and cyclical data stem from incomplete sensor coverage leading to data loss, real-time data delays due to the city network's size, and interference from weather or congestion. Furthermore, limited computing resources impede accurate trend extraction and cycle detection, necessitating advanced data dimension compression and efficient feature extraction algorithms.

Existing spatiotemporal models face computational challenges with large-scale traffic data. Graph WaveNet employs diffusion convolution and multi-layer adaptive GCNs, but becomes impractical as node counts exceed 1,500 due to rising complexity. DCRNN [16] uses a diffusion process and recursive neural networks to capture dependencies, yet struggles with scalability. ST-GCN[32] integrates graph and temporal convolutions for multi-node modeling but lacks depth for accurate large-scale predictions. GMAN [36] leverages multi-graph attention for precision, but its high complexity limits real-time processing. These shortcomings highlight the need for innovative, efficient solutions that maintain accuracy in large-scale traffic forecasting.

Higher-order dependencies in traffic data, such as shared patterns across lanes at intersections, enable generative AI to map similarities and compress road network scales. This enhances the adaptability and efficiency of autonomous traffic systems while improving prediction accuracy within human-centered Industry 5.0 frameworks. It signals a shift in traffic intelligence toward human-focused practices, where AI and automation augment rather than replace human skills.

To address this issue, we introduce HAG-MTF, a novel spatio-temporal prediction model. This model incorporates high-order feature compression and expansion modules to efficiently predict large-scale spatial nodes within computational constraints. By enhancing the viability of traffic data prediction technology, HAG-MTF offers crucial predictive techniques for urban traffic management and planning.

- This study introduces a distributed higher-order graph convolution algorithm powered by generative AI, which captures high-order dependencies in spatiotemporal graphs by modeling mappings, grouping, and convolving neighboring traffic data nodes. This approach efficiently reduces computational complexity in large-scale graph structures, enhancing adaptability and precision in traffic forecasting.
- A novel module integrates spatial neighborhood information in traffic spatiotemporal data by coupling neighboring node weights, embedding global spatial details, and hierarchically fusing spatiotemporal information, enhancing prediction model performance without increasing computational load.
- Extensive experiments on real-world traffic datasets show that HAG-MTF outperforms existing methods in managing high-frequency, long-sequence, multi-node spatiotemporal data, offering superior prediction accuracy, enhanced node-processing capacity, faster speeds, and robust generalization across diverse traffic scenarios.

In traffic forecasting scenarios, the HAG-MTF model integrates human intent through adaptive analysis and prediction of large-scale real-world traffic data. It supports judgment and decision adjustments for real-time events and urban traffic planning. Industry 5.0 applications in smart traffic facilitate hybrid systems, where human experts collaborate with AI to ensure safe operations and effective oversight. This approach contrasts sharply with non-Industry 5.0 methods, which rely on full automation and often overlook expert awareness, human factors, or real-world experiences and objective changes. As a result, they produce rigid and potentially biased outcomes in urban traffic.

2 RELATED WORK

Existing traffic flow forecasting methods can be broadly categorized into four groups: classical statistical approaches, time series-based methods, spatiotemporal neural network-based methods, and approaches that leverage higher-order structural information. The details of each category are outlined below:

2.1 Classical Traffic Flow Prediction Methods

Classical traffic flow prediction methods, rooted in traditional machine learning, target specific traffic data traits. The Historical Average (HA) model [14] leverages historical time series averages for periodic sequences, while fully connected neural networks capture basic linear patterns in simpler time series. HMMHMPI [22] employs Hidden Markov Models to integrate spatial attributes, enhancing regional traffic evolution modeling. The ALTO model [1] uses low-rank tensor decomposition with controlled randomness to avoid local optima, and ARIMA-SVM [3] merges SVM with ARIMA for accurate short-term urban traffic forecasts [19]. Though effective, these methods, developed earlier, struggle with the coarse data granularity and fail to meet the complexity and precision demands of modern traffic prediction tasks [25].

2.2 Time Series-based Traffic Flow Prediction Methods

Advancing insights into traffic flow data have led researchers to reframe prediction as time series forecasting [21], leveraging deep sequence models under the hidden Markov hypothesis [6]. The GRU model employs gating to mitigate vanishing gradient issues, adeptly handling long-term dependencies. CNN-Traffic [37] reframes prediction as a classification task using convolutional neural networks for fixed-interval traffic analysis. 3D-ConvLSTMNet [9] integrates 3D CNN with LSTM to capture both short- and long-term traffic features, enhancing performance. MF-CNN [26] incorporates external factors like periodicity, weather, and holidays into traditional

CNNs for improved efficiency. MAPredRNN [10] adds a multi-head attention mechanism to RNNs, better capturing complex temporal dependencies in traffic data.

The above methods deeply consider the long-term and short-term dependence of traffic flow data in the time dimension, as well as the influencing factors [27] in the task of traffic flow data prediction, such as weather, holidays, etc. However, such methods ignore the spatial characteristics of traffic flow data, such as the connectivity relationship of each intersection.

2.3 Spatiotemporal Neural Network-based Traffic Flow Prediction Methods

Advances in graph theory and GNNs have led researchers to model traffic flow data as a fusion of temporal and spatial features, developing spatiotemporal neural network-based prediction methods. Graph convolutional neural networks [4] utilize adjacency matrices to capture spatial correlations, excelling in traffic network modeling. STGCN [32] integrates graph and temporal convolutions for effective spatiotemporal prediction. STFGCN [2] introduces a spatial-temporal multi-factor fusion graph convolutional network for traffic flow prediction. TCGCN [30] employs a multi-channel cross-attention model to address cross-dimensional dependencies. Crossformer [35] adapts the Transformer framework with cross-dimensional attention for multi-dimensional sequence analysis. These spatiotemporal data prediction methods have large spatiotemporal information monitoring modules, which are difficult to apply to the analysis of large-scale spatial node traffic data.

STGNFs [33] uses conditional normalized flows and a spatiotemporal fusion network to learn probabilistic relationships between historical and future traffic data. DCGCN [28] incorporates causality via dynamic Bayesian networks, generating temporal causal graphs to depict nonlinear traffic flow changes. Additionally, preprocessing techniques like EfficientDeRain+ [8] enhance sensor data quality in rain using RainMix augmentation and uncertainty filtering, enhancing spatiotemporal traffic prediction robustness. These probabilistic models of spatiotemporal data are studied by probabilistic methods, and the training results are uncertain. In practice, the stability of the training results should be improved.

The aforementioned methods design a series of spatiotemporal neural network-based approaches [11] aimed at enhancing traffic flow prediction accuracy, effectively improving both the model's accuracy and receptive field. However, they also introduce significant challenges [20], particularly in terms of increased model complexity, which complicates their application to real-world tasks.

2.4 High-order Structure-based Traffic Flow Prediction Methods

Recent research highlights the significance of incorporating higher-order dependencies in isotemporal traffic flow data [24]. Methods such as STHGCN [29] and [17] have been instrumental in improving prediction accuracy by separately modeling time, space, and dimensional attributes. While STHGCN excels in capturing intricate traffic characteristics, its computational efficiency is compromised by its multi-layered structure [34], particularly with increasing data volume and node counts (Zhang et al., 2024). Similarly, HSPGNN effectively addresses missing spatiotemporal data through a spatial attention mechanism and a dynamic Laplace matrix, utilizing a general non-homogeneous partial differential equation to reconstruct time series values [13]. Nonetheless, both approaches encounter challenges in adapting to large-scale datasets due to their complexity.

To address the challenges posed by high time and space complexity, suboptimal performance with large-scale data, and obstacles in deploying models for practical traffic flow applications, this chapter presents an effective spatiotemporal data compression module utilizing high-order graphs, in conjunction with a robust prediction approach named HAG-MTF. This methodology incorporates a modular framework with bidirectional diffusion convolution and a Global Node Information Embedding Module, providing enhanced computational efficiency and flexibility. Through dimensionality reduction while retaining crucial spatiotemporal patterns, HAG-MTF

surpasses existing techniques in large-scale scenarios, ensuring resilient performance and scalability without significant computational overhead.

3 METHODOLOGY

3.1 Problem Definition

This section focuses on the formal definition related to high-order spatial networks in traffic flow prediction. This section introduces the commonly used concepts and symbols in this chapter.

Road Network G: A road network is typically defined as a weighted directed graph, denoted by $G = (V, E, A, X)$, used to describe the inherent topological configuration of traffic road lanes. Here, the node set $V = \{v_0, \cdots, v_S\}$ represents a collection of S intersections containing lanes, while the edge set E signifies the relationships between intersection nodes, where $(v_i, v_j) \in E$ indicates a direct connection between intersection v_i and v_j by road or lane. The adjacency matrix $A \in \mathbb{R}^{S\times S}$ serves as the weight matrix to express the strength of the connection between these nodes quantitatively.

The intrinsic traffic data on the road network G is considered as attribute features of the nodes in set V, represented by matrix $X \in \mathbb{R}^{S\times F\times T}$. Here, F represents the number of node attribute features, including traffic speed, traffic flow, and traffic density, while T represents the duration of the historical time series, and S denotes the number of nodes.

High-order Road Network $\hat{G}$: A high-order road network is typically represented as a weighted directed graph $\hat{G} = (\hat{V}, \hat{E}, \hat{A}, \hat{X})$, aggregated from elements in G. This is achieved by defining a mapping $\pi : \{V, E, \delta\} \rightarrow \{\hat{V}, \hat{E}\}$, where nodes from the original network are mapped such that for any node $\pi(\hat{v}) \in \hat{V}$, $\pi^{-1}(\hat{v}) \in V$. Here, π can be any similarity-based node merging algorithm, which combines nodes $v_1, v_2, \ldots, v_\delta$ with similarity measures into a super node $\hat{v}$, with the input parameter δ denoting the compression ratio.

Traffic Flow Prediction: Let $x_t \in \mathbb{R}^{S\times F}$ represent the traffic data at each intersection or lane at time t. The historical traffic data with a length of T time steps is denoted as $X = \{x_0, \cdots, x_T\} \in \mathbb{R}^{S\times F\times T}$. The traffic data for the next P time steps can be predicted as follows:

$$\hat{Y} = f_\theta\left(A, X\right),\ X \in \mathbb{R}^{S\times F\times T}, \tag{1}$$

where $\hat{Y} = \{\hat{x}_{T+1}, \cdots, \hat{x}_{T+P}\}$ represents the predicted outputs generated by the model f_θ, where θ denotes the model parameters. The training objective of the model f_θ is to progressively minimize the discrepancy between $\hat{Y}$ and the corresponding ground truth $Y = \{x_{T+1}, \cdots, x_{T+P}\}$.

3.2 Framework

To effectively capture the spatiotemporal dynamics of complex systems, the HAG-MTF spatiotemporal data prediction framework, as illustrated in Fig. 1, is composed of four core components, namely, the spatiotemporal neighborhood information coupling module, the high-order spatial information extraction module, the spatiotemporal evolution pattern modeling module (denoted as H-net), and the high-order spatial data decoding and mapping module. The HAG-MTF framework systematically addresses the high-order features of spatiotemporal data by efficiently extracting spatial high-order information, dynamically learning the high-order graph structure, and incorporating temporal dependencies to predict future states. The core modules of the model are as follows:

(1) **Spatial-Temporal Neighborhood Information Coupling Module**: By integrating spatial weights with data features and incorporating the spatiotemporal information of neighboring data nodes [12], the module enhances the spatial information representation of sequential data without increasing the computational load.

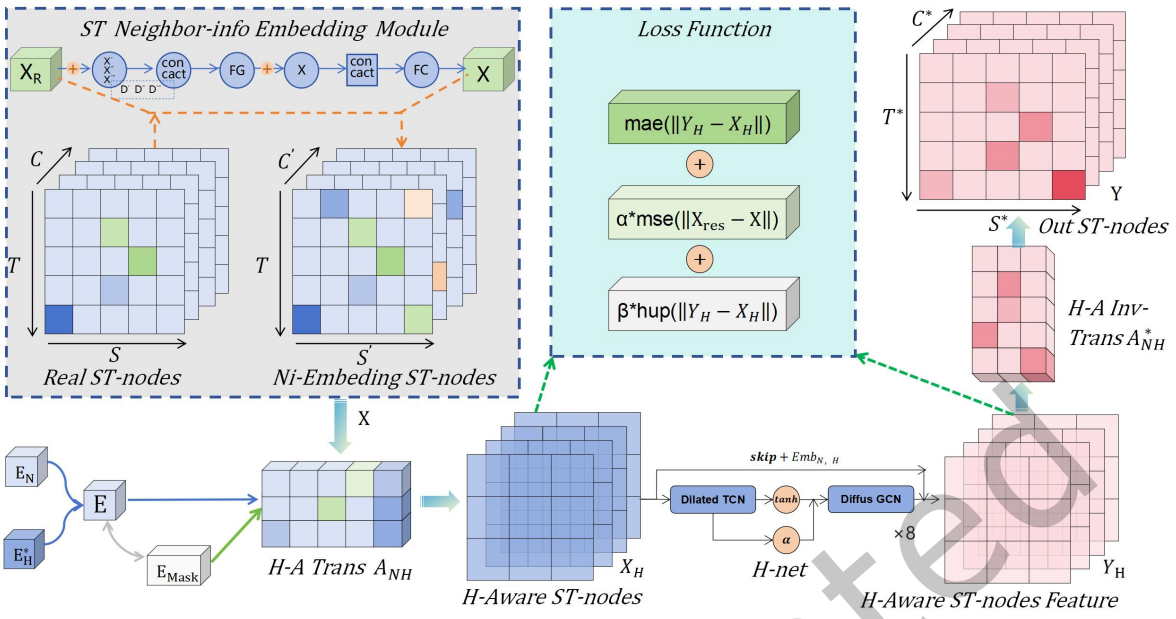

Fig. 1. The architecture of HAG-MTF. This framework integrates four core components that capture the spatiotemporal dynamics information: a. Spatiotemporal neighborhood coupling, b. Spatial information extraction, c. High-order spatiotemporal analysis network, d. High-order ST-data decoding.

(2) **High-order Spatial Information Extraction Module**: Through high-order spatial transformations [23] and feature simplification, it generates high-order, low-dimensional spatiotemporal information representations, simplifies and optimizes data dimensions, and further extracts high-order features.
(3) **Spatiotemporal Evolution Pattern Modeling**: Using a chain-like, time-causal convolutional main architecture [35], it captures high-order, low-dimensional spatiotemporal data features. The module includes submodules such as diffusion graph convolution, causal convolution, gate structures, and attention mechanisms.
(4) **Reverse High-order Decoding**: It uses high-order features to reconstruct predicted values of the original space, ensuring prediction accuracy.

Through the collaborative interaction of the above modules, the HAG-MTF algorithm framework effectively addresses issues such as excessively large spatial dimensions, complexity, nonlinearity, and dynamics in high-order, low-dimensional data for spatiotemporal prediction. It provides a scientifically sound solution for modeling and predicting large-scale, complex system data [5] in traffic information.

3.3 Neighborhood Spatial Information Fusion Module. (NSIF)

In spatiotemporal modeling, accurately representing the relationship between nodes and their neighbors is essential for precise prediction. The original data often contains spatial features with redundancies or incompleteness, complicating feature aggregation and increasing computational complexity. To address this issue, we introduce a spatio-temporal neighborhood information coupling module. This module combines spatial features of target nodes through weighted aggregation and convolution optimization. By prioritizing important neighbor nodes (e.g., high traffic and short distance links) over less significant ones (e.g., low traffic and long distance links), the module incorporates key neighborhood information into the target node. It achieves this by utilizing

trainable adaptive weights for the fusion of neighborhood node information. This approach enhances the model's spatial information representation capability without adding to the computational burden. The formulation of the modular approach is as follows:

For the raw traffic data input X_{raw}, the method for coupling spatial information from neighboring nodes is formulated as:

$$X = \text{Conv2d}\left(X_{\text{raw}} + g_i\left(\sum_{i=1}^{n} \alpha_j X_{\text{raw}}^{(j)}\right)\right), \tag{2}$$

where X_{raw} represents the raw data of the current traffic node, and each $X_{\text{raw}}^{(i)}$ denotes the data of the i-th neighboring traffic node. The weight parameter α_i is a learnable coefficient that adjusts the contribution of each neighboring node's data in the fusion process.

The module integrates spatial feature information by considering neighboring nodes that are close to the current node. Finally, a two-dimensional convolution (Conv2d) is applied to the weighted aggregated data to extract features for subsequent analysis. The module for spatial feature fusion and extraction of hidden information integrates weighted spatial data from neighboring nodes with two-dimensional convolution operations. By incorporating spatial information from original data nodes, it enriches the representation of neighborhood node information to deepen the understanding of spatial relationships within the model. This process extracts crucial features for decision-making and prediction, while disregarding distant spatial nodes' mutual relationships. This approach uniquely contributes to the analysis and prediction of non-uniformly distributed actual traffic flow patterns.

3.4 Spatiotemporal High-Order Feature Aggregation Module

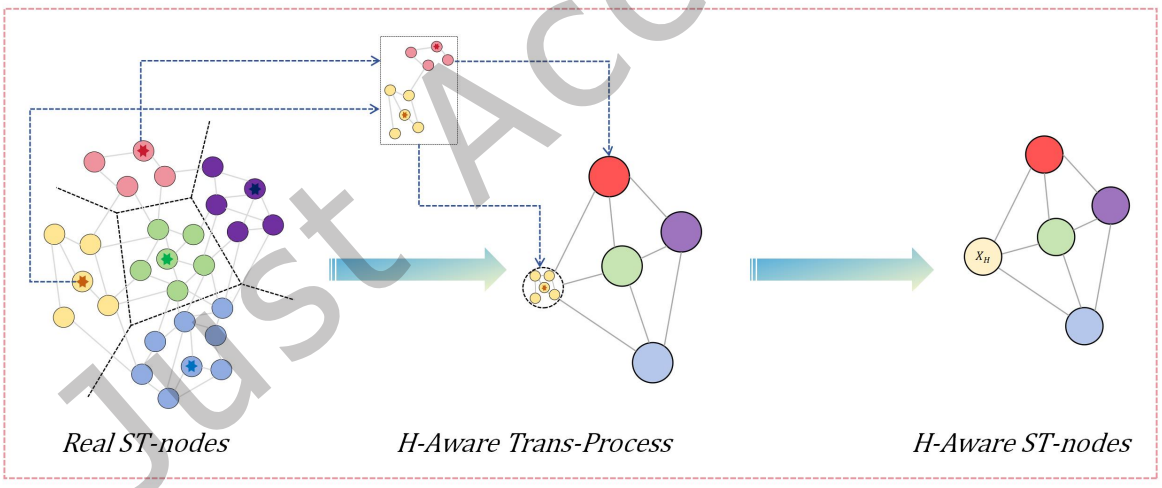

Fig. 2. High-order transformation graph. The module employs high-order node temporal feature averaging and spatial information compression/reconstruction transformations to process the high-order features of complex spatiotemporal data.

The HAG-MTF framework employs a spatiotemporal high-order feature aggregation module to generate a compact, information-rich feature representation by reducing input data dimensionality, simplifying spatial information, and compressing redundant features into a lower-dimensional form. After spatial domain information

coupling, the low-order spatial feature tensor of the original data is denoted as $E_N \in \mathbb{R}^{N\times d}$, where N represents the number of low-order node spatial features and d is the dimension of each feature. After dimensionality reduction, this tensor is transformed into high-order spatial features, represented as $E_H \in \mathbb{R}^{H\times d}$, where H denotes the number of high-order node spatial features and $H < N$, This transformation not only reduces the number of spatial dimensions but also aggregates and preserves essential information.

In the high-order feature transformation, we utilize feature alignment and multiple activation functions to extract and integrate both conventional and high-order features, enhancing their representational capacity through a salient feature iterative enhancement method. The fusion process is expressed as follows:

$$A'_{NH} = \tanh\left(\tanh(g_\theta(E_N)) \odot \tanh(g_\phi(E^*_H))\right), \tag{3}$$

where g_θ and g_ϕ denote parameterized projection networks, and $\odot$ is the Hadamard product.

To mitigate information degradation post-high-order transformation and enhance feature expressiveness, we introduce a mask generation mechanism using a parameter-enhanced iterative strategy. This approach selects the top$K$ contributing elements from fused features, zero-padding non-salient ones to preserve tensor shape consistency. The mask matrix is formulated as follows:

$$A_{\text{mask}} = \text{TopK}\left(A^{'}_{NH}\right)^{A_i=1}_{k=10} + \text{TopK}^{'}\left(A^{'}_{NH}\right)^{A'_i=0}_{k'>10}. \tag{4}$$

To continuously reinforce the salient features within the data, we select the top $K = 10$ most significant features from the transformation module A'_{NH} after introducing random noise. The corresponding positions in the index matrix are set to 1, while all other positions are set to 0. Iteratively repeating this process yields the reinforced high-order transformation feature mask matrix A_{mask}.

A small amount of random noise is then added to the original high-order feature matrix A'_{NH} to enhance feature diversity and improve the model's robustness. The noise magnitude is controlled by a constant of 0.01, ensuring that it does not significantly alter the features, thereby maintaining stability when processing uncertain or noisy data.

Finally, the high-order feature matrix is element-wise multiplied with the generated mask A_{mask}. This operation further filters the important feature information by selecting key features while ignoring unnecessary ones, reducing computational complexity and enhancing efficiency. Consequently, this leads to improved performance in subsequent tasks. The formula is as follows:

$$A_{NH} = \left(A^{'}_{NH} + \text{Rand}(A^{'}_{NH}) \times 0.01\right) \times A_{\text{mask}}. \tag{5}$$

The dimension reduction transformation matrix $\mathbf{A_{NH}} \in \mathbb{R}^{N\times H}$ compresses spatial information in traffic data into a compact feature representation via dimensionality reduction. Processed through the high-order spatiotemporal module, upscaling and downscaling operations map these low-dimensional features back to a high-dimensional space, restoring data post-degenerate feature analysis for a concise, effective representation. Embedded into the high-order spatiotemporal network, this refined representation captures long-range dependencies and global distribution characteristics, enhancing the model's perception of the overall spatiotemporal structure.

Within the HAG-MTF framework, the dimensionality ascending and descending modules address redundancy issues in high-dimensional traffic data by compressing and simplifying spatial node dimensions. This process effectively reduces computational complexity, significantly enhancing the scalability of real-time data processing models, which is of practical importance. Moreover, it has demonstrated superior efficiency in handling large-scale data, thereby improving the scalability of real-time traffic prediction and resource optimization within urban systems. Through global embedding, the modules minimize noise while preserving essential information, ultimately ensuring efficient model calculations without sacrificing accuracy. Furthermore, the framework exhibits

adaptability to diverse datasets, thereby bolstering the effectiveness, scalability, and generalization capabilities of the model.

3.5 High-order Spatiotemporal Evolution Module

Traffic data's spatiotemporal variability necessitates concurrent spatial and temporal modeling, yet traditional approaches falter with high-order data and escalating spatial feature dimensions. To address this, the High-order spatiotemporal Evolution Module integrates two advanced modules: the Global Node Information Embedding Module, which aggregates temporal features from high-order spatial data to reduce dimensionality while simulating spatial dynamics, and the High-Order Spatiotemporal Analysis Network (H-net), which enhances modeling via diffusion convolution for long-range spatial interactions and gated convolutions for multi-scale temporal variations.

Reverse high-order transformation then restores features to their original dimensions, followed by convolution to produce the final output. Within this, HAG-MTF excels in initialization performance, leveraging bidirectional diffusion convolution to capture complex forward and reverse spatial relationships [18] and a modular H-net design for adaptive feature and dependency learning. This ensures high predictive accuracy in large-scale graph scenarios, computational efficiency, and robust initialization over prolonged operations [31].

3.5.1 Global Node Information Embedding. In higher-order tensor representations of traffic data, dimensionality reduction inevitably leads to feature loss during spatial compression. To mitigate this challenge, we propose a global-aware embedding mechanism that enhances residual learning through three coordinated operations: (1) temporal feature aggregation, (2) adaptive spatial reconstruction, and (3) structural dimension calibration. This approach effectively preserves essential spatiotemporal patterns during the dimensionality reduction process.

Given input tensor $X \in \mathbb{R}^{b\times d\times h\times t}$ where b, d, h, t denote batch size, node dimension, hidden size, and temporal length respectively, the module first extracts *global temporal context* through dimension-wise averaging:

$$Q' = \mathrm{Tanh}(g_\theta(\frac{1}{t}\sum_{k=1}^{t} X_{:,:,:,k})), \tag{6}$$

where $g_\theta(\cdot)$ denotes learnable linear projection.

Subsequently, we construct *dynamic embedding matrices* $\mathrm{E_N} \in \mathbb{R}^{d\times r}$ and $\mathrm{E_H} \in \mathbb{R}^{h\times r}$ through hierarchical transformation:

$$\mathrm{E_N}, \mathrm{E_H} = \mathrm{Tanh}(g_\phi(\mathrm{Tanh}(g_\psi(\mathcal{N}(0,1))))), \tag{7}$$

$$Q = Q' \odot (\mathrm{E}_N \otimes \widehat{\mathrm{E}_H}), \tag{8}$$

where $\otimes$ denotes tensor product and $\odot$ represents element-wise multiplication. As illustrated in Fig. 2, this *dual-embedding mechanism* establishes learnable mappings between original and compressed spaces through: (1) Stochastic initialization with normal distribution $\mathcal{N}(0,1)$, (2) Nonlinear projection via g_ψ and g_ϕ, (3) Dimensional alignment using hyperbolic tangent activations.

3.5.2 Temporal Order-Preserving Causal Convolution. H-net employs Gated convolution to process temporal data, Causal convolution is used to handle temporal information, and Dilated convolution is employed to compress feature dimensions into the time dimension, enabling the model to capture time features of high-order data at different scales. The formulas are as follows:

$$\mathcal{T} = \tanh\big(W_f * X + b_f\big) \times \sigma\big(W_g * X + b_g\big), \tag{9}$$

$$X^{(t)} = \tanh\big(W_t * X^{(t-1)}\big), \tag{10}$$

$$RF_{n+1} = RF_n + (k + (k-1)(d-1) - 1)\, S, \tag{11}$$

where $\mathcal{T}$ is the output of gated convolution, $X^{(t)}$ is the output at time step t, W and b are convolution and bias parameters, RF is the receptive field size, k is the kernel size, d is the dilation rate, and S is the stride.

Gated Convolution. The **Multi-layer Gated Recognition Module (MGSM)**, defined by Eq. (9), serves as a specialized multi-layer gated convolution mechanism within the high-order spatiotemporal feature extraction module H-net. It adaptively regulates information flow through a gating mechanism that combines tanh and sigmoid activations for selective propagation. The module captures long-range dependencies, boosts model expressiveness, and focuses on the identification and processing of temporal features in complex traffic sequences.

Causal Convolution. Eq. (10) ensures temporal outputs depend only on current and past inputs, capturing high-order spatiotemporal dependencies. **Dilated Convolution.** Eq. (11) expands the receptive field in the temporal dimension by systematically adjusting the dilation rate d, allowing rapid expansion of the receptive field in a multi-layer causal network. This enhances the model's ability to capture global temporal information across all frequencies without increasing parameters or computational costs.

The High-Order Spatiotemporal Analysis Network (H-net) employs Diffusion Graph Convolution to model the characteristics of high-order spatial traffic data due to its ability to effectively capture complex spatial dependencies. This method leverages a normalized bidirectional power propagation mechanism, aggregating node features across both forward and reverse information flows. This bidirectional approach excels in handling the dynamic, non-uniform spatial relationships in traffic networks, enhancing the expressive capacity for long-range dependencies. The adjustable diffusion step length K allows flexibility in controlling the propagation range, making it ideal for large-scale, high-order data.

3.5.3 Spatial-Dependency Diffusion Graph Convolution. To capture the spatial information contained in high-order spatial data, we first apply Diffusion Graph Convolution, which employs a diffusion propagation mechanism. This process, also known as the normalized bidirectional power propagation of the adjacency matrix, aggregates node features for extraction. The module is designed to enhance the expressive capacity of high-order spatial features by utilizing bidirectional diffusion convolution that accounts for both forward and reverse information flows. The specific method is as follows:

$$\mathcal{X}' = \sum_{k=0}^{K-1} \left(\mathbf{D}^{-1}\mathbf{A}\right)^k \mathbf{X}\Theta_k + \sum_{k=0}^{K-1} \left(\mathbf{D}\mathbf{A}^\top\right)^k \mathbf{X}\Theta_k, \tag{12}$$

where the adjacency matrix for high-order spatial data is embedded after the initial matrix undergoes a high-order transformation. $\mathbf{D}$ is the degree matrix, with its diagonal elements given by $\mathbf{D}_{ii} = \sum_j \mathbf{A}_{ij}$. The degree matrix helps to aggregate the neighborhood information of nodes, thereby better capturing the spatial dependencies among them. And $\mathbf{A}_{ij}$ is the **Spatial Prior Information Embeding (SPIE)** adjacency matrix. The SPIE module embeds prior spatial knowledge from raw input data into the H-net following dimensionality reduction. The H-net module employs bidirectional diffusion convolution and Gated units to model high-order features by the $X' \in \mathbb{R}^{T\times d\times N_H}$, compressing spatial data to E_H while retaining global and neighborhood details, widening receptive fields, and enhancing prediction accuracy and efficiency in large-scale traffic forecasting without extra computational burden. Θ_k denotes the trainable parameters of the kth convolution layer, and K is the diffusion step length, which determines the range of information propagation.

3.6 Loss Function and Optimization Objective

3.6.1 Composite Loss Optimization Module. A key challenge in the model is the discrepancy that can arise when transforming between high-dimensional and low-dimensional data. Relying on a single error metric is insufficient for a comprehensive evaluation of prediction performance. To address this, we have designed a composite loss function that simultaneously considers high-order error, global error, and dynamic adjustments, thereby imposing constraints on the model from multiple perspectives.

The selection of this composite loss function is specifically designed for traffic prediction applications. It incorporates high-order errors to capture intricate spatial relationships, global errors to maintain consistency throughout traffic networks, and dynamic adjustments to accommodate real-time fluctuations such as congestion or weather conditions. This comprehensive methodology improves resilience and is in line with the demand for precise and scalable predictions in evolving urban settings, as evidenced by various traffic datasets. The formulation of the loss function is outlined below:

$$\mathcal{L} = \frac{1}{n}\sum_{i=1}^{n}\left((Y_H)_i - (X_H)_i\right) + \alpha\frac{1}{n}\sum_{i=1}^{n}\left((X_{\text{rec}})_i - X_i\right)^2 + \beta\,\text{HUP}\left(Y_H, X_H\right). \tag{13}$$

Where Y_H and X_H denote predicted and input high-order data, respectively, focusing on fitting high-order features, while X_{rec} represents reconstructed data from upscaling and downscaling via inverse normalization, assessing transformation fidelity against original data X. Hyperparameters α and β balance global error and dynamic penalty weights, ensuring comprehensive optimization of high-order feature fitting, data consistency, and refined constraints. The loss comprises three parts: the overall error between Y_H and X_H to evaluate high-order feature fitting, the difference between X_{rec} and X to maintain consistency across transformations, and the High Uncertainty Penalty (HUP) to heighten focus on regions with significant high-order errors, boosting sensitivity to critical data in complex scenarios.

3.6.2 Data upscaling and downscaling Reconstruction. To accurately assess data fidelity during upscaling and downscaling, we introduce a reconstruction formula:

$$\mathcal{X}_{\text{rec}} = f_{\text{MLP}}\left(g_i\left(\text{Tanh}\left(X \times \mathbf{A}_{\text{NH}}\right) \times \mathbf{A}^*_{\text{NH}}\right)\right). \tag{14}$$

The matrices A_{NH} and A^*_{NH} serve as transformation matrices for upscaling and downscaling, respectively. The activation function tanh introduces nonlinearity. The functions g_i and f_{MLP} denote a linear transformation and a multi-layer perceptron (MLP) module, respectively, for adaptive fine-tuning following dimensionality changes. Comparing X_{rec} with the original data ensures consistency and integrity during the upscaling and downscaling processes.

3.6.3 Dynamic Penalty Term (HUP). This term imposes additional penalties when the prediction error exceeds a threshold, focusing the model's optimization on high-error regions. It is formulated as a piecewise function that applies no penalty for small errors, while larger errors incur quadratic penalties, thereby increasing sensitivity to critical data. The HUP is defined as:

$$\text{hup}(e) = \begin{cases} \gamma \cdot e^2, & \text{if } e > \delta, \\ 0, & \text{otherwise.} \end{cases} \tag{15}$$

The model employs a loss function where $e = ||\hat{y}_i - x_i||$ denotes the absolute prediction error, δ (set to 5) is the error threshold triggering additional penalties, and γ (set to 0.1) modulates penalty strength. For example, an error $e = 8$ incurs a penalty of $0.1 \cdot 8^2 = 6.4$, while $e = 3$ incurs none, prioritizing correction of larger discrepancies. The first term loss function assesses absolute errors in high-order spatial data post-degeneration, underscoring its impact on accuracy, while the 2nd portion evaluates average loss across original data, minimizing errors from dimensionality transformations. The 3rd component, the dynamic High-Order Uncertainty Penalty (HUP), enhances sensitivity to extreme variations by weighting larger errors, and optimizing prediction. Experiments confirm this formulation significantly boosts performance.

The HAG-MTF framework effectively captures the spatiotemporal dynamics of complex systems through its four core components. It adeptly extracts high-order spatial information, learns graph structures dynamically, and integrates temporal dependencies to ensure precise and resilient predictions of future states. The algorithmic process for HAG-MTF is as shown in Algorithm 1.

We offer a simple step-by-step overview to interpret Algorithm 1. For each time step, combine spatial data from neighbors using weighted fusion and convolution. And then, create high-order embeddings, boosting key features with a TopK mask and slight random noise for robustness. Use bidirectional diffusion convolution to capture forward and backward spatial links. After that, apply gated causal convolution to manage time-series flow, and restore features to original dimensions via upsampling and a multi-layer perceptron. At last, compute a combined loss covering high-order errors, reconstruction, and penalties for big uncertainties, minimize the loss with the Adam optimizer, and output predictions and reconstructed data. The detailed algorithmic process is as described in Algorithm 1.

Algorithm 1 HAG-MTF Spatiotemporal Data Prediction

Input: Traffic tensor $\mathcal{X}_{\text{raw}} \in \mathbb{R}^{B\times S\times F\times T}$, adjacency matrix $\mathbf{A}$, hyperparameters S, K, α, β
Output: Predictions $\hat{\mathcal{Y}}$, reconstructions $\tilde{\mathcal{X}}$

Initialize parameters $\Theta, \mathbf{W}, \{\theta_k\}_{k=0}^{K-1}$
for time step $t = 1$ **to** T **do** ▷ Spatial aggregation
  for node $i = 1$ **to** N **do**
    $\mathbf{X}^{(t,i)} = \text{Conv2D}\Big(\mathcal{X}_{\text{raw}}^{(t,i)} + g_i\big(\sum_{j\in\mathcal{N}_i} \alpha_j \mathcal{X}_{\text{raw}}^{(t,j)}\big)\Big)$
  end for
end for
$\mathbf{E}_N, \mathbf{E}_H = \text{N}(\mathcal{X}), \text{H}^*(\mathcal{X})$ ▷ High-order features
$\mathbf{A}'_{NH} = \tanh(g(\mathbf{E}_N) \odot g(\mathbf{E}_H^*))$
$\mathbf{A}_{NH} = \big(\mathbf{A}'_{NH} + \eta \cdot \text{Rand}(\mathbf{A}'_{NH}) * 0.01\big) \otimes \text{TopK}_{10}(\mathbf{A}'_{NH})^{A_i=1}$
$\mathbf{X}' = \sum_{k=0}^{K-1} \Big[(\mathbf{D}^{-1}\mathbf{A})^k \mathbf{X}\theta_k + (\mathbf{D}^{-1}\mathbf{A}^\top)^k \mathbf{X}\theta_k\Big]$ ▷ Bidiffusion
$\mathcal{T} = \tanh(\mathbf{W}_f * \mathbf{X} + \mathbf{b}_f) \odot \sigma(\mathbf{W}_g * \mathbf{X} + \mathbf{b}_g)$ ▷ Gated tempoconv
$\mathbf{X}^t = \tanh(\mathbf{W} * \mathbf{X}^{t-1})$ **for** $t = 2 : T$ ▷ Causal convolution
$\tilde{\mathcal{X}} = \text{MLP}\big(\tanh(\mathbf{X}\mathbf{A}_{NH}) \cdot \mathbf{A}_{NH}^\top\big)$ ▷ Reconstruction
$\mathcal{L} = \|\hat{\mathcal{Y}} - \mathcal{Y}\|_2^2 + \alpha\|\tilde{\mathcal{X}} - \mathcal{X}\|_2^2 + \beta\mathcal{L}_{\text{hup}}$ ▷ Composite loss
Minimize $\mathcal{L}$ via Adam
return $\hat{\mathcal{Y}}, \tilde{\mathcal{X}}$

3.7 Analysis of Module Complexity

Building on the problem definition for traffic data and the methodological details outlined earlier, we examine the complexity of the method's core modules to lay the groundwork for discussing subsequent experiments and applications.

- **Neighborhood Spatial Information Fusion (NSIF):** NSIF acts as the preprocessing module for spatial information coupling in the HAG-MTF model. It fuses spatial details from adjacent nodes into the raw input $X_{\text{raw}} \in \mathbb{R}^{N\times T\times d}$ using a graph convolution-style aggregation, as shown in Eq. 2. The computation of $A^k X_{\text{raw}}$ uses sparse matrix multiplication, where $k = 3$ is the number of neighbor nodes, E is the number of edges in the adjacency matrix, and d is the node feature dimension. The module's complexity is $O(E \cdot d \cdot T) \times 3$.
- **Spatial Prior Information Embedding (SPIE):** SPIE embeds prior spatial information from the raw data into the high-order spatiotemporal feature processing module (H-net) after dimensionality reduction. It models the high-order features $X' \in \mathbb{R}^{T\times d\times N}$ using bidirectional diffusion convolution and GRU-like units. The SPIE module compresses the spatial information to a reduced dimension E_H, resulting in a per-operation cost of $O(E_H \cdot d \cdot T)$ for the bidirectional diffusion convolution. The gated temporal

convolution operates along the time steps T per layer, with a complexity of $O(T \cdot E_H \cdot d^2)$. The full H-net module stacks $L = 8$ layers, yielding a total complexity of $O((E_H \cdot d \cdot T + T \cdot E_H \cdot d^2)) \times 8$.

- **Multi-layer Gated Recognition Module (MGSM)**: MGSM is a specialized multi-layer gated convolution mechanism within the high-order spatiotemporal feature extraction module H-net, dedicated to identifying and processing temporal features. By stacking multiple layers of gated convolutions (with L = 8), it adaptively manages information flow, preserves temporal sequencing, and broadens the receptive field. Its complexity is two main operations: bidirectional diffusion convolution for spatial handling and gated temporal convolution across layers. As a result, the overall complexity of MGSM (H-net) is $O((E_H \cdot d \cdot T + T \cdot E_H \cdot d^2) \times 8)$. With high-order compression yielding $E_H << N$(where N is the original number of spatial nodes), this complexity remains quadratic in the original spatial dimension N, specifically, an $O(N^2) \cdot d$ scaling that leverages the efficiency gains from dimensionality reduction.

In summary, the total complexity of HAG-MTF is $O(8 \times (3 \times E \cdot d \cdot T + E_H^2 \cdot d + T \cdot E_H \cdot d^2))$. As shown in Table 4, HAG-MTF can process data with over 1,500 spatial nodes from the UTD19 dataset, while baselines like GraphWaveNet fail due to GPU memory overflow. This design strikes a balance between accuracy and efficiency for real-time Industry 5.0 applications, delivering strong prediction metrics unlike other deep learning frameworks built on complete graph structures, which often cannot finish running because of their heavy computational load.

4 EXPERIMENTS

4.1 Experimental Settings

4.1.1 Dataset and Baseline. The proposed HAG-MTF model has been evaluated on multiple publicly available datasets spanning various domains, including traffic flow, energy generation, and meteorology. The study utilizes METR-LA, PEMS-BAY, Solar, Traffic, Weather, PeMS04, PeMS08, and UTD-19 datasets. And the proposed method is compared against the following baseline methods: HAG-MTF, HA, FNN, GRU, GCN, STGCN, STHSL, Crossformer, AGCRN, SVR, DGCRN, and AM-GCN.

The HAG-MTF model was evaluated using publicly available datasets with diverse characteristics. Specifically, the model was tested on traffic flow data from METR-LA and PEMS-BAY, both of which are characterized by dense road network structures. METR-LA was used to assess the model's sensitivity to traffic bursts and to verify its robustness in handling missing data. In contrast, PEMS-BAY was employed to evaluate the model's ability to manage complex topologies and capture long-term dependencies. Additionally, the model was tested on the high-speed traffic datasets PEMS04 and PEMS08, which feature sparse spatial nodes, to assess its performance in modeling sparse nodes and implicit spatial relationships. Finally, a medium-sized traffic dataset was used to confirm the model's stability in predicting traffic patterns for medium-sized road networks, particularly in managing seasonal variations and missing data.

The UTD-19 dataset comprises a vast road network with diverse inter-city traffic patterns and intricate spatio-temporal dependencies, serving as a rigorous benchmark for evaluating prediction models. Specifically, it assesses the algorithm's efficacy in long-range spatial association mining and computational efficiency. Additionally, the dataset includes solar energy generation and meteorological data, each varying in scale and complexity. These diverse datasets offer a robust foundation for evaluating the spatiotemporal predictive capabilities of the model under test.

4.1.2 Evaluation Metrics. To evaluate the performance of HAG-MTF, two widely used metrics are employed to quantify the relative prediction error.

- *MSE* (Mean Squared Error) emphasizes larger errors through squaring and computes the average of squared differences, which can be computed as:

$$MSE = \frac{1}{QN} \sum_{t=1}^{Q} ||Y_t - \hat{Y}_t||_2^2. \tag{16}$$

- *MAE* (Mean Absolute Error) provides the average magnitude of errors using absolute differences, which can be computed as:

$$MAE = \frac{1}{QN} \sum_{t=1}^{Q} |Y_t - \hat{Y}_t|. \tag{17}$$

Other Settings. This study utilized an RTX 4090 GPU cluster for efficient parallel training of deep learning models, with Python 3.8 ensuring portable and maintainable code and PyTorch 2.1.1, with its flexible API and optimizations, enabling rapid model iteration and enhanced performance; for a fair comparison, all baseline models employed Huber loss for its demonstrated effectiveness, with a consistent batch size of 64 and a dataset partitioned into training, validation, and test sets at an 8:1:1 ratio, selecting the best-performing model after 50 epochs of training on the validation set for testing on the test set.

4.2 Experimental Results

4.2.1 Baseline Comparisons. We assessed the HAG-MTF model using a range of publicly available datasets. The datasets were selected based on their varying scales, temporal resolutions, and connections to practical traffic forecasting problems, such as urban congestion, sparse network configurations, long-term trends, data incompleteness, and weather impacts. The METR-LA dataset examined the model's response to traffic surges and its robustness against missing values in dense networks, while PEMS-BAY evaluated its handling of intricate topologies and prolonged dependencies. These two medium-sized datasets also verified the model's stability in the face of seasonal fluctuations. PEMS04 and PEMS08 gauged their ability to capture implicit spatial connections. UTD-19 served as a benchmark for long-distance associations and computational efficiency in expansive systems. Solar and weather data tested the model's performance on sparse datasets and its adaptability to environmental variations. This selection of datasets validates the model's effectiveness and supports reliable outcomes in extensive, fine-grained spatiotemporal traffic prediction tasks.

Experimental results in Table 2 and Table 3 show that HAG-MTF consistently achieves top or near-top performance across datasets and prediction intervals. HAG-MTF markedly surpasses traditional models like GCN and STGCN, sustains superior performance across all intervals on the Solar dataset, and excels on the Weather and PeMS datasets, demonstrating robust capability in both short- and long-term predictions.

The model is trained using the Adam optimizer with an initial learning rate of 0.001, decayed by a factor of 0.1 every 50 epochs, and a batch size of 32. Convergence is achieved when the validation loss stabilizes for 10 consecutive epochs or falls below 0.01. To mitigate overfitting, a dropout rate of 0.1 is applied during training.

The superior performance of HAG-MTF can be attributed to its advanced spatiotemporal modeling, which combines higher-order graph convolution with a spatial information fusion module. This approach efficiently captures both temporal and spatial features, along with non-linear spatiotemporal dependencies, enabling accurate predictions across datasets characterized by significant temporal variations and spatial complexity.

4.2.2 Efficiency Experiment. The efficiency experiment aims to assess the algorithm's performance in processing a significant volume of node spatiotemporal data. Table 4 illustrates that HAG-MTF employs adaptive neighborhood embedding, higher-order transformation, and spatial graphics compression to reduce the dimensionality of such data. It demonstrates notable efficiency and accuracy in handling large-scale node data, effectively processing Los Angeles traffic data in the UTD19 dataset using an RTX-2080Ti graphics processor with limited computational

Table 2. **Baseline Comparisons (MSE)**

Method		HAG-MTF	HA	FNN	GRU	GCN	STGCN	STHSL	Crossformer
METR-LA	15m.	**25.983**	75.075	28.425	32.238	56.336	95.019	32.120	46.352
	30m.	**32.380**	118.091	36.120	43.302	66.272	103.960	43.686	72.592
	1h.	**42.362**	189.279	46.612	59.338	83.939	119.961	64.772	96.196
PEMS-BAY	15m.	8.871	9.501	11.647	6.741	44.936	131.055	**6.337**	8.443
	30m.	12.945	18.798	15.354	12.536	46.573	115.063	**11.972**	16.656
	1h.	**16.856**	36.438	19.639	19.294	49.810	162.193	20.581	-
Solar	30m.	5.843	**4.891**	12.642	11.741	58.600	166.111	14.736	16.405
	1h.	**8.016**	12.537	12.884	17.766	103.157	371.089	14.878	38.394
	2h.	**13.206**	33.528	16.765	30.924	98.217	211.283	129.028	44.714
Traffic	3h.	**0.001**	0.003	0.001	0.002	0.002	0.006	0.002	-
	6h.	**0.001**	0.004	0.001	0.002	0.003	0.005	0.003	-
	12h.	**0.002**	0.005	0.002	0.002	0.003	0.004	0.003	-
Weather	3h.	**661.181**	791.558	985.900	832.029	103.5K	679.301	1.3K	683.787
	6h.	**856.330**	1.2K	1.7K	1.2K	103.7K	895.496	5.8K	921.287
	12h.	**1.1K**	2.2K	2.4K	1.8K	104.2K	1.4K	27.6K	1.4K
PSM04	15m.	961.880	1.1K	1.5K	1.1K	4.5K	7.7K	**951.134**	1.2K
	30m.	**943.937**	1.5K	1.7K	1.2K	4.7K	7.2K	947.983	2.0K
	1h.	**1.0K**	2.7K	1.7K	1.5K	5.0K	7.9K	1.0K	2.4K
PSM08	15m.	**560.165**	658.759	1.1K	631.344	3.3K	5.7K	582.550	774.455
	30m.	**597.541**	959.105	1.3K	780.775	3.4K	8.7K	604.886	1.8K
	1h.	**617.755**	1.9K	1.5K	932.173	3.6K	10.2K	644.150	2.0K

resources. This dataset comprises over 1500 spatial nodes. The HAG-MTF's prediction performance consistently surpasses that of baseline methods, while approaches like 3DFormer, Crossformer, DGCRN, and AM-GCN struggle to manage replication tasks involving 1500 spatial nodes due to their extensive parameter spatiotemporal modules. Deep learning models with fixed topologies, such as AGCRN, STGCN, and FNN, exhibit inadequate prediction performance in intricate and variable graph structures and are incapable of executing long-term prediction tasks spanning one hour. Apart from the models introduced in this study, only HA, SVR, and GRU can accomplish the task, albeit with prediction accuracy below the desired standard.

The HAG-MTF method utilizes high-order spatial node compression to decrease data dimensionality, enhance computational efficiency, preserve spatial information richness, demonstrate its effectiveness in handling intricate spatio-temporal dynamics, and is well-suited for analyzing and predicting large-scale spatio-temporal data in real-world traffic scenarios.

4.2.3 Ablation Study. The ablation study confirms the critical significance of key modules within HAG-MTF for spatiotemporal prediction. These include the incorporation of higher-order compression loss in the loss function, spatial diffusion convolution in the higher-order spatiotemporal analysis module, time-gated convolution, and training initial convolution. Empirical findings across various datasets unequivocally underscore the pivotal role of these modules in enhancing the predictive efficacy of the model.

The high-order dimensionality reduction module optimizes computational efficiency for large-scale data by compressing redundant features without major information loss. The generative fusion graph enhances spatial dependency modeling through dynamic neighborhood integration, improving relationship capture in dense

Table 3. **Baseline Comparisons (MAE)**

Method		HAG-MTF	HA	FNN	GRU	GCN	STGCN	STHSL	Crossformer
METR-LA	15m.	2.742	8.665	3.051	2.952	4.756	3.849	2.974	**2.577**
	30m.	**2.968**	10.867	3.307	3.290	5.071	5.211	3.281	3.926
	1h.	**3.294**	13.758	3.680	3.860	5.613	5.359	4.251	4.437
PEMS-BAY	15m.	1.540	3.082	1.833	1.226	2.933	5.344	**1.213**	1.422
	30m.	1.748	4.336	2.019	1.536	3.056	5.036	**1.513**	1.811
	1h.	1.965	6.036	2.202	**1.892**	3.250	5.776	1.989	-
Solar	30m.	**1.450**	2.212	2.268	2.312	4.982	10.051	2.577	2.500
	1h.	**1.664**	3.541	2.310	2.855	6.822	13.841	2.523	3.923
	2h.	**2.220**	5.790	2.644	3.700	6.730	10.065	8.194	4.035
Traffic	3h.	0.017	0.052	**0.015**	0.019	0.029	0.048	0.026	-
	6h.	**0.017**	0.065	0.017	0.024	0.031	0.042	0.027	-
	12h.	**0.017**	0.070	0.019	0.021	0.030	0.041	0.027	-
Weather	3h.	6.516	28.135	14.775	**5.487**	174.780	6.884	15.493	7.229
	6h.	7.644	34.887	22.917	**7.062**	175.480	8.127	28.902	7.746
	12h.	**8.134**	47.185	21.633	9.575	177.075	10.158	77.558	10.145
PSM04	15m.	**19.773**	33.255	25.167	20.870	40.942	72.171	19.894	22.615
	30m.	19.849	38.255	27.176	22.538	42.580	70.217	**19.721**	30.386
	1h.	**20.814**	52.347	27.550	24.818	44.788	72.972	20.938	33.705
PSM08	15m.	**15.781**	25.666	23.073	16.374	37.382	65.827	16.016	19.768
	30m.	**15.897**	30.969	24.639	17.956	38.847	81.033	16.024	28.380
	1h.	**15.972**	43.539	25.887	19.422	40.522	83.297	16.489	30.260

Table 4. **Efficiency Experiment**

		Model							
Metric	Step	AGCRN	GRU	HA	FNN	SVR	STGCN	3DFormer Crossformer DGCRN AM-GCN	HAG-MTF
MSE	15min.	150.1K	257.2K	10.4K	**8.7K**	164.4K	110.2K	-	9.2K
	30min.	152.5K	263.8K	22.9K	13.2K	166.8K	123.7K	-	**12.5K**
	1h.	-	214.9K	46.2K	-	170.1K	-	-	**23.8K**
MAE	15min.	269.798	393.614	62.937	59.327	283.394	273.520	-	**49.289**
	30min.	269.305	395.649	83.156	**64.790**	286.445	301.980	-	82.756
	1h.	-	344.024	127.529	-	290.514	-	-	**102.185**

networks. The H-net module boosts accuracy via hierarchical fusion of temporal and spatial features, aiding long-sequence handling. The fusion loss function aids balanced optimization, enhancing stability and convergence while preventing overfitting. Notably, the NSIF module has a pivotal impact, reducing errors by embedding global details and coupling node weights, essential for Information extraction of data compression.

As shown in Table 5, experimental evidence underscores the importance of the time-gated convolution and spatial diffusion convolution components in effectively capturing complex spatiotemporal features and

Table 5. **Ablation Study**

Method		HAG-MTF		Wo. Second Loss		Wo. GCN in H-net		Wo. Head Conv		Wo. Gate in H-net	
Metric		MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE
METR-LA	15m.	**25.983**	**2.742**	26.107	2.769	26.867	2.798	27.569	2.850	41.718	3.481
	30m.	**32.380**	**2.968**	34.191	3.057	33.297	3.021	33.440	3.057	43.647	3.447
	1h.	**42.362**	**3.294**	43.093	3.336	43.440	3.340	42.509	3.331	63.033	4.091
PEMS-BAY	15m.	**8.871**	**1.540**	9.033	1.556	8.989	1.564	10.375	1.707	25.523	2.422
	30m.	**12.045**	**1.718**	12.136	1.745	12.091	1.721	12.902	1.830	25.917	2.465
	1h.	16.496	1.965	**16.482**	1.974	16.507	**1.965**	17.810	2.044	26.304	2.486
Solar	30m.	**5.843**	1.450	6.113	1.535	5.876	**1.440**	8.702	1.857	6.399	1.532
	1h.	**8.016**	**1.664**	8.071	1.707	8.237	1.735	11.761	2.195	8.968	1.825
	2h.	**13.206**	**2.220**	13.398	2.236	13.899	2.264	18.644	2.775	15.131	2.480
Traffic	3h.	0.001	0.017	0.001	0.017	0.001	0.017	**0.001**	**0.016**	0.002	0.020
	6h.	**0.001**	0.017	0.001	0.017	0.001	**0.017**	0.001	0.018	0.003	0.037
	12h.	**0.001**	0.017	0.001	0.017	0.001	**0.017**	0.001	0.018	0.001	0.017
Weather	3h.	**661.181**	**6.516**	749.860	8.695	723.939	8.233	671.397	7.026	680.445	6.623
	6h.	**856.330**	**7.644**	955.735	9.827	1.1K	10.137	905.162	8.580	1.1K	11.528
	12h.	**1.1K**	**8.134**	1.3K	11.281	1.6K	12.844	1.6K	11.001	1.5K	11.879

establishing dynamic dependencies within spatiotemporal data. These results emphasize the critical interplay between temporal and spatial feature analysis modules, which is fundamental to higher-order spatiotemporal analysis frameworks.

Furthermore, ablation experiments underscore the essential role of high-order compression loss in counterbalancing spatial compression loss. The discrepancy between compressed high-order features and original data is integrated into the loss function to effectively manage information loss during high-order transformation. These results affirm that the modular architecture of HAG-MTF effectively harmonizes feature extraction, dynamic dependency modeling, and regularization processes.

4.2.4 Parametric Analysis. Analysis of the ST Block underscores HAG-MTF's exceptional spatiotemporal feature modeling and its sensitivity to depth optimization. Its modular design unifies high-dimensional spatial representation, feature extraction, and dynamic adjacency matrix learning, delivering a robust framework for capturing intricate spatiotemporal dependencies.

A. ST Block Depth Analysis. Experiments on St-Block layer variations, as detailed in Table 6, demonstrate that the 8-layer configuration of the HAG-MTF framework consistently outperforms alternatives across datasets, achieving an MSE of 25.983 on the METR-LA 15-minute prediction task, and an MSE of 13.206 for the Solar 2-hour prediction. The 8-layer's performance surpasses 6 and 10-layer setups by effectively capturing nonlinear, dynamic spatiotemporal features and long-term dependencies. This robust performance arises from synergistic modules' graph convolution for spatial dependencies, causal convolution for temporal dynamics, and adaptive graph learning for latent node relationships, though extending to 10 layers reveals diminishing returns, suggesting a trade-off where excessive depth risks inefficiency, noise, or overfitting. For complex datasets like Weather, the 8-layer setup remains optimal despite higher errors, indicating room for enhanced multi-scale modeling, yet it

Table 6. **ST Block Depth Analysis Experiment**

Method		Layers 4		Layers 6		Layers 8		Layers 10	
Metric		MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE
METR-LA	15m.	28.116	2.861	30.010	2.924	**25.983**	**2.742**	28.393	2.910
	30m.	35.032	3.135	46.767	3.467	**32.380**	**2.968**	34.685	3.135
	1h.	-	-	50.589	3.580	**42.362**	**3.294**	44.531	3.433
PEMS-BAY	15m.	10.600	1.724	13.143	1.856	**8.871**	**1.540**	10.825	1.740
	30m.	14.245	1.847	23.316	2.587	12.945	**1.748**	**12.216**	1.750
	1h.	-	-	26.313	2.491	**16.856**	**1.965**	17.665	2.038
Solar	30m.	6.449	1.551	14.497	2.363	**5.843**	**1.450**	8.240	1.810
	1h.	8.840	1.796	15.718	2.553	**8.016**	**1.664**	11.891	2.214
	2h.	-	-	18.866	2.789	**13.206**	**2.220**	17.206	2.657
Traffic	3h.	0.002	0.020	0.001	0.018	**0.001**	0.017	0.001	**0.016**
	6h.	0.001	0.017	0.002	0.022	**0.001**	**0.017**	0.002	0.021
	12h.	-	-	0.002	0.026	**0.002**	**0.017**	0.002	0.022
Weather	3h.	694.242	7.523	895.496	8.127	**661.181**	**6.516**	666.095	7.357
	6h.	1.0K	9.075	1.2K	14.512	**856.330**	**7.644**	1.1K	10.114
	12h.	-	-	1.4K	10.158	**1.1K**	**8.134**	1.5K	11.737

strikes a balance between depth and efficiency, reinforcing HAG-MTF's adaptability and rigor in spatiotemporal prediction tasks.

B. High-Order Mapping Module Compression Ratio Analysis. The analysis of the high-order mapping module's compression ratio, as shown in Fig. 3, underscores HAG-MTF's efficiency in optimizing feature representation and dynamic modeling. A 1/3 compression ratio consistently yields superior performance across datasets, striking an optimal balance between redundancy elimination and information retention. This ratio enhances dimensionality reduction, spatial dynamics extraction, and spatiotemporal feature fusion.

To further illustrate the impact of spatial compression rates on model performance, we examine the trends in Fig. 3, which depicts MAE and MSE variations across ratios from 1/1 to 1/5 on the PeMS04 dataset. As the compression rate rises from 1/1 (no compression) to 1/3, both metrics exhibit a steady decline, indicating improved prediction accuracy. This gain arises from the high-order dimensionality reduction module's ability to distill essential spatiotemporal dependencies, thereby alleviating noise and redundancy in raw high-dimensional data. For example, on PeMS04, MAE decreases from 28.5 at 1/1 to 22.3 at 1/3, as compressed features preserve key higher-order patterns, such as multi-hop neighbor correlations, while cutting computational overhead by about 65%.

However, beyond 1/3 (e.g., at 1/4 or 1/5), errors increase sharply due to excessive information loss, where critical local dependencies are discarded, resulting in underfitting for complex scenarios. These trends affirm the 1/3 ratio as optimal, minimizing errors (with an 18% MSE reduction) while fostering scalability and efficiency. This choice aligns with HAG-MTF's design to circumvent the quadratic complexity issues in baselines like GraphWaveNet.

Overall, this optimal ratio addresses the pitfalls of low compression (suboptimal performance) and high compression (feature degradation). The module's integration of high-dimensional spatial representations with a reverse decoding strategy further ensures efficient feature extraction, achieving a robust equilibrium between predictive accuracy and computational demands.

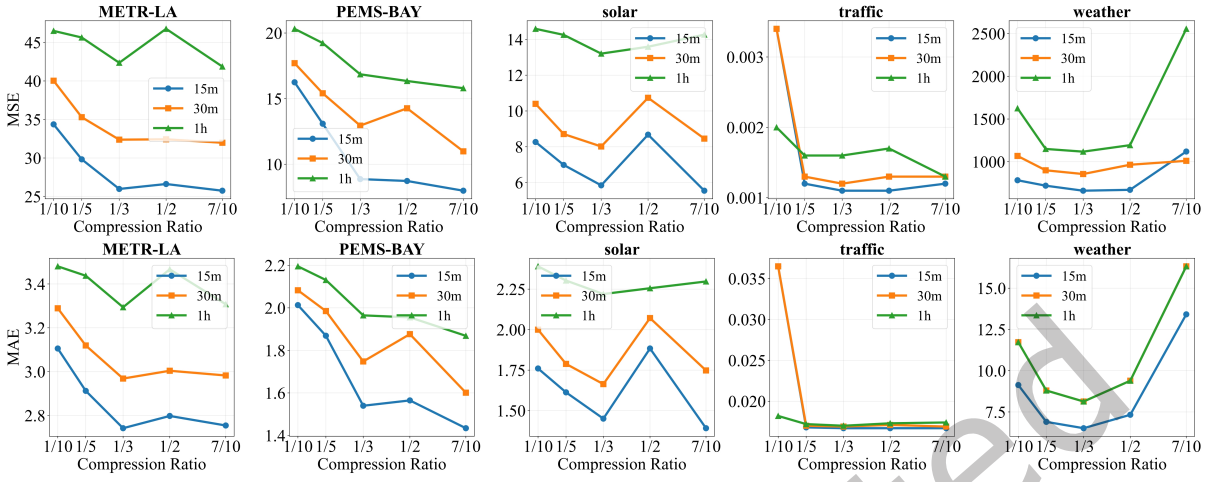

Fig. 3. The figure analyzes the high-order compression-decompression module's performance with varying compression ratios. The experimental diagram shows that the MAE and MSE values are relatively low when the compression ratio of multiple datasets is 1/3, showing a parabolic relationship between compression and prediction accuracy.

C. The Dimension in High-order Analysis Modules Study. As shown in Table 7, the experiment results on embedding dimensions within the HAG-MTF framework reveal that a 128-dimensional embedding optimally balances feature representation and computational efficiency across diverse datasets and prediction tasks. Outperforming the 192-dimensional variant, it delivers sufficient expressive capacity with lower computational cost, minimizing noise and inefficiencies tied to higher dimensions. HAG-MTF?s adaptive architecture?integrating graph convolution, causal convolution, and attention mechanisms?synergistically captures dynamic spatiotemporal relationships, ensuring robustness and efficiency. This equilibrium underscores the model?s emphasis on accuracy and resource efficiency, laying a foundation for future enhancements via dynamic embedding and advanced dimensionality reduction techniques.

4.2.5 Components Study. The enhanced spatiotemporal feature modeling within the HAG-MTF framework is achieved through the synergistic integration of NSIF, MGSM, and SPIE modules. This collaborative approach facilitates dynamic information exchange and mutual reinforcement among the components, enabling more effective analysis of high-order data features and adept capture of nonlinear dynamics within intricate datasets. As illustrated in the experimental Fig. 4, the combined NSIF, MGSM, and SPIE modules demonstrate superior performance in predictive tasks across various datasets. Notably, the holistic configuration outperforms individual modules in isolation, underscoring their pivotal roles in spatial interdependence and multi-layer feature extraction.

The collaboration between NSIF and SPIE enhances the spatial information receptive field of spatiotemporal data and integrates prior spatial information into high-order spatiotemporal data analysis. This collaboration closely interacts with the MGSM module, which analyzes long-sequence spatiotemporal data characteristics, ensuring the incorporation of original spatial information into the multi-gate structure's high-order spatiotemporal analysis module while preserving the original characteristics. Additionally, compressing spatiotemporal information enhances model efficiency significantly. Thus, these three modules are essential and work closely together.

In the context of the 30-minute solar mission, the integrated model demonstrates superior performance compared to a standalone MGSM configuration. This suggests that although MGSM can compress time and predict over medium to long durations, additional elements within the model enhance its capacity for representing

Table 7. **Experiment of The Dimension in High-order Analysis Modules**

Method		Dimension 256		Dimension 192		Dimension 128		Dimension 64	
Metric		MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE
METR-LA	15m.	27.846	2.824	26.107	2.757	**25.983**	**2.742**	28.904	2.900
	30m.	33.299	3.042	**32.014**	**2.648**	32.380	2.968	35.011	3.123
	1h.	42.226	3.341	42.362	3.294	**41.819**	**3.208**	43.327	3.327
PEMS-BAY	15m.	10.324	1.719	**8.653**	**1.540**	8.871	1.471	10.692	1.762
	30m.	14.358	1.874	**12.207**	1.799	12.945	**1.748**	13.921	1.871
	1h.	16.914	2.012	**16.633**	**1.704**	16.856	1.965	17.125	2.031
solar	30m.	9.547	1.994	**5.415**	**1.309**	5.843	1.450	9.723	2.001
	1h.	12.875	2.254	8.096	1.704	**8.016**	**1.664**	11.746	2.174
	2h.	19.585	2.857	13.406	2.238	**13.292**	**2.220**	18.319	2.761
traffic	3h.	0.001	**0.016**	0.002	0.020	**0.001**	0.017	0.001	0.016
	6h.	0.001	0.018	0.001	0.017	**0.001**	**0.017**	0.002	0.020
	12h.	**0.001**	0.018	0.001	0.018	0.002	**0.017**	0.002	0.020
weather	3h.	740.103	8.205	695.880	8.146	**661.181**	**6.516**	691.394	7.705
	6h.	929.805	9.436	948.674	8.081	**856.330**	**7.644**	897.144	8.102
	12h.	1.5K	10.623	1.6K	10.354	**1.1K**	**8.134**	1.2K	8.990

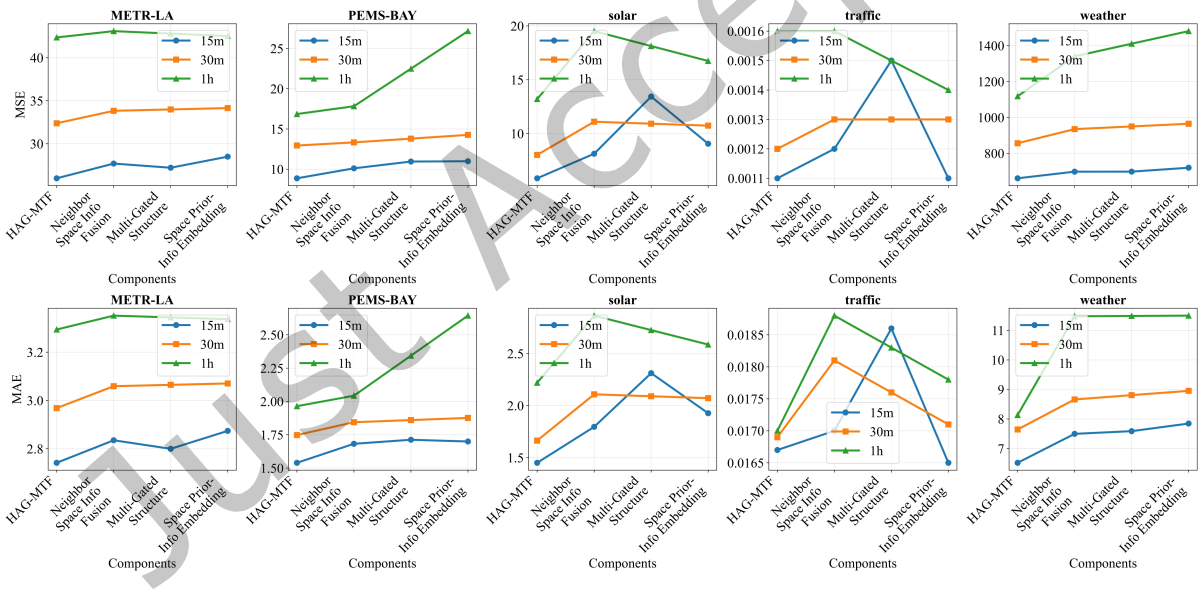

Fig. 4. The figure compares the prediction performance of the model composed of different components (Original, NSIF, NSIF + SPIE, MGSM) under multiple data sets, illustrating the contribution of each optimized module to the prediction accuracy.

high-dimensional spatial features and encoding extensive spatiotemporal data. Despite certain modules like higher-order transformations leading to decreased accuracy and efficiency, potential enhancements in module

synchronization and task-specific optimization may enhance the adaptability of HAG-MTF for spatiotemporal prediction tasks.

5 CONCLUSION AND FUTURE WORK

This paper introduces a distributed higher-order graph convolution algorithm using generative AI. It captures high-order dependencies in spatiotemporal traffic graphs by mapping, grouping, and convolving neighboring nodes. This reduces computational complexity in large-scale graphs, improving adaptability and precision. We also present a module that fuses spatial neighborhood information by weighting nodes, embedding global features, and combining spatiotemporal data, boosting performance without extra computational cost. Experiments on real-world traffic datasets show HAG-MTF excels in handling high-frequency, long-sequence, multi-node data, with better accuracy, speed, node capacity, and generalization.

While our experiments demonstrate robust performance on urban datasets like METR-LA and UTD-19, which feature dense and inter-city networks, the model's generalizability to other scenarios warrants further discussion. For instance, in rural networks with sparser topologies and fewer sensors, HAG-MTF's high-order dimensionality reduction and adaptive graph fusion could maintain efficiency by compressing sparse nodes without losing key dependencies, as qualitatively suggested by its handling of implicit spatial relationships in PEMS04/PEMS08. In extreme weather, where data shows sudden disruptions like reduced visibility or altered flows, the model's component may boost resilience by dynamically integrating environmental factors, akin to its performance on weather-impacted solar and meteorological data.

Furthermore, HAG-MTF captures the essence of Industry 5.0's focus on people by using AI for flexible traffic predictions in everyday settings, promoting closer teamwork between humans and machines while giving traffic managers the tools for in-depth research and judgment for smarter oversight. Looking ahead, to push beyond what Industry 5.0 currently requires, we should concentrate on building tougher hybrid setups by blending in real-time data from multiple sources, reinforcement learning, big models, and agent-based tools. This approach affords individuals supplementary channels for intervening in and steering traffic flows. For instance, through the configuration of adjustable intelligent emergency alerts or the execution of rapid safety risk assessments, thereby augmenting the scalability, robustness, and adaptability of intelligent transportation systems.

ACKNOWLEDGMENTS

This work was supported by the National Key R&D Program of China under Grant No. 2023YFC3321501 and the Emerging Frontiers Cultivation Program of Tianjin University Interdisciplinary Center.

REFERENCES

[1] Xiaobo Chen, Kaiyuan Wang, Zuoyong Li, Yu Zhang, and Qiaolin Ye. 2023. A novel nonconvex low-rank tensor completion approach for traffic sensor data recovery from incomplete measurements. *IEEE Transactions on Instrumentation and Measurement* 72 (2023), 1–15.

[2] Ying-Ting Chen, An Liu, Cheng Li, Shuang Li, and Xiao Yang. 2025. STFGCN:Traffic flow prediction based on spatial-temporal multi factor fusion graph convolutional networks. *Scientific Reports* 15, 1 (2025), 12612.

[3] Z. Chi and L. Shi. 2018. Short-Term Traffic Flow Forecasting Using ARIMA-SVM Algorithm and R. In *2018 5th International Conference on Information Science and Control Engineering (ICISCE)*. IEEE, Zhengzhou, China, 517–522.

[4] Wenhao Chu, Chunxiao Zhang, Heng Li, Laifu Zhang, Dingtao Shen, and Rongrong Li. 2024. SHAP-powered insights into spatiotemporal effects: Unlocking explainable Bayesian-neural-network urban flood forecasting. *International Journal of Applied Earth Observation and Geoinformation* 131 (2024), 103972.

[5] Jingtao Ding, Yu Zheng, Huandong Wang, Carlo Vittorio Cannistraci, Jianxi Gao, Yong Li, and Chuan Shi. 2025. Artificial Intelligence for Complex Network: Potential, Methodology and Application. In *Companion Proceedings of the ACM on Web Conference 2025 (WWW '25)*. ACM, Sydney, NSW, Australia, 5–8.

[6] Shai Fine, Yoram Singer, and Naftali Tishby. 1998. The hierarchical hidden Markov model: Analysis and applications. *Machine learning* 32 (1998), 41–62.

[7] Qing Guo, Hua Qi, Jingyang Sun, Felix Juefei-Xu, Lei Ma, Di Lin, Wei Feng, and Song Wang. 2024. EfficientDeRain+: Learning Uncertainty-Aware Filtering via RainMix Augmentation for High-Efficiency Deraining. *Int. J. Comput. Vision* 133, 4 (Nov. 2024), 2111–2135.

[8] Qing Guo, Hua Qi, Jingyang Sun, Felix Juefei-Xu, Lei Ma, Di Lin, Wei Feng, and Song Wang. 2025. EfficientDeRain+: Learning Uncertainty-Aware Filtering via RainMix Augmentation for High-Efficiency Deraining. *International Journal of Computer Vision* 133 (2025), 2111–2135.

[9] Lihua He and Wuman Luo. 2022. 3D-ConvLSTMNet: A deep spatio-temporal model for traffic flow prediction. In *2022 23rd IEEE International Conference on Mobile Data Management (MDM)*. IEEE, Paphos, Cyprus, 147–152.

[10] Xiaohui Huang, Yuan Jiang, and Jie Tang. 2023. MAPredRNN: multi-attention predictive RNN for traffic flow prediction by dynamic spatio-temporal data fusion. *Applied Intelligence* 53, 16 (2023), 19372–19383.

[11] Mouna Jiber, Imad Lamouik, Yahyaouy Ali, and My Abdelouahed Sabri. 2018. Traffic flow prediction using neural network. In *2018 International Conference on Intelligent Systems and Computer Vision (ISCV)*. IEEE, Fez, Morocco, 1–4.

[12] Guangyin Jin, Hengyu Sha, Zhexu Xi, and Jincai Huang. 2023. Urban hotspot forecasting via automated spatio-temporal information fusion. *Applied Soft Computing* 136 (2023), 110087.

[13] Ming Jin, Qingsong Wen, Yuxuan Liang, Chaoli Zhang, Siqiao Xue, Xue Wang, James Zhang, Yi Wang, Haifeng Chen, Xiaoli Li, et al. 2023. Large models for time series and spatio-temporal data: A survey and outlook. *arXiv preprint arXiv:2310.10196* (2023).

[14] Tae-Hwy Lee, Yundong Tu, and Aman Ullah. 2015. Forecasting equity premium: Global historical average versus local historical average and constraints. *Journal of Business & Economic Statistics* 33, 3 (2015), 393–402.

[15] Jinlong Li, Baolu Li, Zhengzhong Tu, Xinyu Liu, Qing Guo, Felix Juefei-Xu, Runsheng Xu, and Hongkai Yu. 2024. Light the Night: A Multi-Condition Diffusion Framework for Unpaired Low-Light Enhancement in Autonomous Driving . In *2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE Computer Society, Los Alamitos, CA, USA, 15205–15215.

[16] Yaguang Li, Rose Yu, Cyrus Shahabi, and Yan Liu. 2018. Diffusion Convolutional Recurrent Neural Network: Data-Driven Traffic Forecasting. In *International Conference on Learning Representations (ICLR)*. ICLR, Vancouver, Canada.

[17] Guojun Liang, Prayag Tiwari, Sławomir Nowaczyk, and Stefan Byttner. 2024. Higher-order Spatio-temporal Physics-incorporated Graph Neural Network for Multivariate Time Series Imputation. In *Proceedings of the 33rd ACM International Conference on Information and Knowledge Management (CIKM)*. ACM, Birmingham, United Kingdom, 1356–1366. Published: 21 October 2024.

[18] Dongya Liu, Xinqi Zheng, Chunxiao Zhang, and Hongbin Wang. 2017. A new temporal–spatial dynamics method of simulating land-use change. *Ecological modelling* 350 (2017), 1–10.

[19] Iraj Lohrasbinasab, Amin Shahraki, Amir Taherkordi, and Anca Delia Jurcut. 2022. From statistical-to machine learning-based network traffic prediction. *Transactions on Emerging Telecommunications Technologies* 33, 4 (2022), e4394.

[20] Maryam M Najafabadi, Flavio Villanustre, Taghi M Khoshgoftaar, Naeem Seliya, Randall Wald, and Edin Muharemagic. 2015. Deep learning applications and challenges in big data analytics. *Journal of big data* 2 (2015), 1–21.

[21] Ignacio-Iker Prado-Rujas, Antonio García-Dopico, Emilio Serrano, M Luisa Córdoba, and María S Pérez. 2024. A multivariable sensor-agnostic framework for spatio-temporal air quality forecasting based on Deep Learning. *Engineering Applications of Artificial Intelligence* 127 (2024), 107271.

[22] Yan Qi and Sherif Ishak. 2014. A Hidden Markov Model for short term prediction of traffic conditions on freeways. *Transportation Research Part C: Emerging Technologies* 43 (2014), 95–111.

[23] Yongming Rao, Wenliang Zhao, Yansong Tang, Jie Zhou, Ser Nam Lim, and Jiwen Lu. 2022. Hornet: Efficient high-order spatial interactions with recursive gated convolutions. *Advances in Neural Information Processing Systems* 35 (2022), 10353–10366.

[24] Mandana Saebi, Jian Xu, Lance M Kaplan, Bruno Ribeiro, and Nitesh V Chawla. 2020. Efficient modeling of higher-order dependencies in networks: from algorithm to application for anomaly detection. *EPJ Data Science* 9, 1 (2020), 15.

[25] Joakim Skarding, Bogdan Gabrys, and Katarzyna Musial. 2021. Foundations and modeling of dynamic networks using dynamic graph neural networks: A survey. *iEEE Access* 9 (2021), 79143–79168.

[26] Dan Tang, Liu Tang, Wei Shi, Sijia Zhan, and Qiuwei Yang. 2021. MF-CNN: a new approach for LDoS attack detection based on multi-feature fusion and CNN. *Mobile Networks and Applications* 26, 4 (2021), 1705–1722.

[27] Eleni I Vlahogianni, Matthew G Karlaftis, and John C Golias. 2014. Short-term traffic forecasting: Where we are and where we're going. *Transportation Research Part C: Emerging Technologies* 43 (2014), 3–19.

[28] Chenxi Wang, Huizhen Zhang, Shuilin Yao, and Minglei Liu. 2022. DCGCN: Double-channel graph convolutional network for passenger flow prediction in urban rail transit. In *2022 8th International Conference on Big Data Computing and Communications (BigCom)*. IEEE, Beijing, China, 304–313. Published: 06-07 August 2022.

[29] Jun Wang, Wenjun Wang, Wei Yu, Xueli Liu, Keyong Jia, Xiaoming Li, Min Zhong, Yueheng Sun, and Yuqing Xu. 2022. STHGCN: A spatiotemporal prediction framework based on higher-order graph convolution networks. *Knowledge-Based Systems* 258 (2022), 109985.

[30] Lei Wang, Deke Guo, Huaming Wu, Keqiu Li, and Wei Yu. 2024. TC-GCN: Triple cross-attention and graph convolutional network for traffic forecasting. *Information Fusion* 105 (2024), 102229.

[31] Sipei Wu, Haiou Wang, and Kai Hong Luo. 2024. A robust autoregressive long-term spatiotemporal forecasting framework for surrogate-based turbulent combustion modeling via deep learning. *Energy and AI* 15 (2024), 100333.

[32] Bing Yu, Haoteng Yin, and Zhanxing Zhu. 2018. Spatio-Temporal Graph Convolutional Networks: A Deep Learning Framework for Traffic Forecasting. In *Proceedings of the 27th International Joint Conference on Artificial Intelligence (IJCAI)*. International Joint Conferences on Artificial Intelligence Organization, Stockholm, Sweden, 3634–3640.

[33] Le Zhang, Wei Cheng, Shuo Zhang, Ji Xing, Xuefeng Chen, Lin Gao, Zhao Xu, Ruzhen Yang, Junying Hong, and Yingfei Ma. 2024. Spatial-Temporal Graph Conditionalized Normalizing Flows for Nuclear Power Plant Multivariate Anomaly Detection. *IEEE Transactions on Industrial Informatics* 20, 11 (2024), 12945–12957.

[34] Yudong Zhang, Pengkun Wang, Binwu Wang, Xu Wang, Zhe Zhao, Zhengyang Zhou, Lei Bai, and Yang Wang. 2024. Adaptive and Interactive Multi-Level Spatio-Temporal Network for Traffic Forecasting. *IEEE Transactions on Intelligent Transportation Systems* 25, 10 (2024), 14070–14086.

[35] Yunhao Zhang and Junchi Yan. 2023. Crossformer: Transformer utilizing cross-dimension dependency for multivariate time series forecasting. In *Proceedings of the Eleventh International Conference on Learning Representations (ICLR)*. International Conference on Learning Representations, Kigali, Rwanda.

[36] Chuanpan Zheng, Xiaoliang Fan, Cheng Wang, and Jianzhong Qi. 2020. GMAN: A Graph Multi-Attention Network for Traffic Prediction. *Proceedings of the AAAI Conference on Artificial Intelligence* 34, 1 (2020), 1234–1241.

[37] Huiyi Zhou, Yong Wang, Xiaochun Lei, and Yuming Liu. 2017. A method of improved CNN traffic classification. In *2017 13th International Conference on Computational Intelligence and Security (CIS)*. IEEE, Guangzhou, China, 177–181.