Building Urban Traffic Flow Prediction Model Using Spatio-Temporal Graph Convolutional Networks

Jing Sun[1], Yajuan Liu[1], Yanlei Dou[2], Minggang Wang[3]
[1]Urban Construction Institutes, Chongqing Energy College, Chongqing 402260, China
[2]Road and Bridge Design and Research Institute, Shanghai Municipal Engineering Design Institute (Group) Co., Ltd., Shanghai 200092, China
[3]Road and Bridge Design and Research Institute, Chongqing Expressway Group Co., Ltd., Chongqing 401121, China
Corresponding author: Jing Sun (e-mail: 18584667236abc@sina.com)

ABSTRACT To address the inadequacy of existing models in capturing spatial position information and multi-scale periodic patterns in traffic flow, this paper proposes a Position-Aware Spatio-Temporal Graph Convolutional Network (PASTGCN). The model incorporates a spatio-temporal position embedding module that encodes geographic coordinates and topological attributes into spatial embeddings, combined with sinusoidal encoding for periodic temporal representations. A dual graph learning mechanism fuses a static distance graph based on a Gaussian kernel with an attention-based dynamic graph, while spatial-aware dilated causal convolution enables temporally modulated spatial feature injection. Multigranularity periodic branches further enhance the modeling of daily and weekly patterns. Extensive experiments on the METR-LA dataset show that PASTGCN achieves an RMSE of 2.53 (veh/5 min) and an R2 of 0.982, outperforming several contemporary baselines including DCRNN, STGCN, Graph WaveNet, ASTGCN, AGCRN, and STAEformer. Ablation studies confirm the contribution of each module, with the dynamic graph learning alone reducing RMSE by 0.34. The model also maintains high efficiency with an inference latency of 18.7 ms and a throughput of 53.5 samples/second, demonstrating its suitability for real-time urban traffic prediction. Beyond intelligent transportation applications, the proposed position-aware spatio-temporal modeling strategy provides a practical reference for information fusion and dynamic state estimation in large-scale wireless sensing networks and electromagnetic signal-driven monitoring systems, supporting future intelligent perception and communication infrastructures. This high accuracy and low-latency capability also makes PASTGCN particularly promising for supporting intelligent logistics scheduling in time-sensitive industries.

INDEX TERMS urban traffic prediction, spatio-temporal graph convolutional network, position-aware learning, wireless sensing networks, spatiotemporal information fusion

I. INTRODUCTION

URBAN traffic flow prediction is a core component of intelligent transportation systems and is of great significance for traffic congestion warning, route planning, and urban governance [1,2]. Specifically, for industries reliant on efficient urban logistics, such as the apparel sector, precise traffic prediction is crucial for optimizing supply chains, reducing delivery delays, and managing the transport of materials and finished goods. With the continuous expansion of urban road network scale [3,4] and the increasing complexity of traffic dynamics, traffic flow shows significant nonlinear, spatio-temporal coupling, and multi-scale periodic characteristics. However, existing methods generally have problems such as insufficient utilization of sensor node geographic position information, coarse-grained modeling of periodic time patterns, and insufficient interaction of spatio-temporal features, resulting in limited prediction accuracy and generalization ability in complex urban environments. Especially in non-stationary scenarios such as peak hours [5] and holidays [6], the model's insufficient response to the dynamic evolution of spatial dependencies and position sensitivity restricts its practical application value. Therefore, it is urgent to build a new prediction framework that can deeply integrate position-aware and multi-scale spatio-temporal dependencies to improve the modeling accuracy and robustness under complex traffic systems.

This paper applies a traffic flow prediction method based on PASTGCN, aiming to more precisely model the complex dynamic spatio-temporal dependencies in urban road networks. This paper designs an explicit spatio-temporal position embedding module as the initial input of the model. A spatial-temporal alternating modeling strategy is adopted. A parallel periodic feature extraction branch is designed, and a dynamic graph regularization term is applied to improve model stability. t-SNE (t-Distributed Stochastic Neighbor Embedding) visualization results demonstrate that PASTGCN can distinguish traffic p atterns d uring different time periods (morning rush hour, evening rush hour, off-peak period, and nighttime off-peak period), with clear cluster boundaries. Ablation experiments show that removing dynamic graph learning (w/o DynGraph) increases the RMSE to 2.87 (vehicles/5 minutes), while removing the periodic branch (w/o PeriodicBranch) reduces the R^2 by 0.039, validating the necessity of each module. PASTGCN achieves 3.2G FLOPs (Floating Point Operations) and 18.7ms latency, while maintaining a high throughput of 53.5 samples/second. In multi-dataset validation, the model performs robustly on PeMS-BAY (Performance Measurement System - Bay Area) (325 nodes), PeMSD4 (Performance Measurement System District 4) (307 nodes), and PeMSD8 (Performance Measurement System District 8) (170 nodes), with predicted curves closely matching the true values and demonstrating strong generalization capabilities. PASTGCN exhibits significant a dvantages i n spatio-temporal feature decoupling, computational efficiency, and cross-domain adaptability, demonstrating its effectiveness and robustness in complex urban traffic environments.

II. RELATED WORK

In the early stage, traffic flow prediction research mainly relied on traditional time series analysis methods, such as the autoregressive integrated moving average model [7,8] and its variants, which achieved short-term prediction by modeling linear trends and periodic components. However, these methods assume that the data follows a stationary distribution, making it difficult to capture the nonlinear dynamics and external disturbances in traffic flow, and unable to model spatial correlation [9]. Models such as support vector regression [10,11] and Kalman filtering are applied. Although they have improved in local nonlinear fitting, they are still limited to single-point prediction and lack the ability to perceive the overall structure of the road network. With the development of deep learning, recurrent neural networks and their gated variants have been widely used in traffic prediction tasks due to their ability to model sequence dependencies [12,13], which can effectively capture long-term dependencies in the time dimension. Researchers have tried to apply multiple LSTM (Long Short-Term Memory) units in parallel to each sensor node, and applied spatial interactions through a fully connected layer to initially achieve spatio-temporal joint modeling [14]. Due to the fixed and fully connected dependencies between all nodes, they fail to reflect the sparse and dynamic spatial associations in the real road network. In addition, convolutional neural networks have also been used to extract local spatio-temporal patterns, but they are only applicable to regular grid structures and cannot directly handle the irregular, non-Euclidean spatial topology of urban road networks [15,16]. Therefore, a new neural network paradigm that can naturally model graph-structured data and deeply integrate spatio-temporal dynamics is urgently needed.

In recent years, the rapid development of graph neural networks (GNNs) has provided a powerful tool for modeling complex traffic networks. By abstracting urban road networks into a graph structure, where nodes represent monitored road sections and edges represent spatial connections or functional associations, GNN can effectively aggregate neighbor information and capture non-local spatial dependencies [17]. On this basis, a series of spatio-temporal graph convolutional networks (STGCNs) have been proposed, marking the entry of traffic prediction into a new stage of deep spatio-temporal modeling [18,19]. STGCN combines graph convolution with temporal convolution, achieving end-to-end prediction by stacking spatio-temporal convolution blocks, significantly improving the model's expressive power [20]. DCRNN (Diffusion Convolutional Recurrent Neural Network) applies diffusion convolution to simulate the bidirectional propagation mechanism of traffic flow, and combines the encoder-decoder architecture with the attention mechanism to further enhance the modeling capability of long-term dynamics [21]. Graph WaveNet, through an adaptive adjacency matrix learning mechanism, can discover implicit spatial dependencies without relying on prior graph structure, achieving leading performance on multiple datasets [22,23]. ASTGCN (Attention Based Spatial-Temporal Graph Convolutional Network) applies a spatio-temporal attention module to model the weight distribution in the temporal and spatial dimensions, respectively, thereby enhancing the recognition of periodic and trend patterns [24,25]. The GMAN (Graph Multi-Attention Network) further integrates a gating mechanism with multi-head attention to achieve dynamic perception of the global spatio-temporal context [26]. Although the above methods have made significant progress, most models do not make sufficient use of the geographic position information of nodes. The position is usually only used to construct a static graph and is not integrated into the deep model as a learnable semantic feature.

To enhance the model's ability to perceive spatial structure, recent research has begun to explore the application of position encoding mechanisms in GNNs. Inspired by position encoding in natural language processing, some work has attempted to embed node indexes or coordinate information into the feature space [27], but most of them use fixed encoding methods and lack fine-grained modeling of geographic semantics. In addition, although dynamic graph learning mechanisms can partially reflect the functional relationship between positions, they often ignore the prior constraints of absolute and relative positions, resulting in unreasonable attention distribution. In terms of temporal modeling, although TCN (Temporal Convolutional Network) and Transformer have been widely adopted, their temporal convolution process is usually independent of spatial representation updates, resulting in insufficient fusion of spatio-temporal features [28,29]. At the same time, multi-task learning has gradually become an effective strategy to improve the generalization ability of models. Existing studies have attempted to jointly predict speed and flow, or combine anomaly detection tasks to enhance feature representation capabilities [30]. However, given the strong periodicity [31,32] characteristics of traffic flow, how to design auxiliary tasks to explicitly guide the model to learn periodic residual patterns still lacks systematic exploration. In summary, existing methods still have significant shortcomings in position-aware, deep spatio-temporal fusion mechanisms, and explicit modeling of periodic patterns. The PASTGCN model applied in this paper systematically improves these key issues. By constructing position-aware embeddings, designing spatio-temporal synchronized convolutional blocks, and applying multi-granularity periodic branches, it achieves a more precise depiction of complex traffic dynamics, driving the development of smarter and more robust urban traffic prediction.

III. METHODS

A. PROBLEM DEFINITION AND MODEL ARCHITECTURE

This study aims to address the short-term prediction of urban-level traffic flow. By constructing a PASTGCN, the dynamic spatio-temporal dependencies in complex road networks can be precisely modeled. Given an urban traffic network [33,34], its topology can be formalized as an undirected graph G=(V,E), where the node set V={v1,v2,...,vN} represents N traffic sensors or road monitoring points distributed throughout the network, and the edge set E⊆V×V represents the spatial adjacency between nodes.

In a discrete time series, it is assumed that Xt∈RN×F is the input traffic state matrix at time t, where each row Xi,t∈RF represents the F-dimensional observation features of node vi at time t. The model takes a historical observation sequence Xt-L+1:t=[Xt-L+1,Xt-L+2,...,Xt] of length L as input, combines the graph structure G, node position information P∈RN×Dp, and time context Tt, and predicts the traffic state in the next Tpred time steps:

$$\hat{X}_{t+1:t+T_{pred}} = f_\theta(X_{t-L+1:t}, G, P, T_t) \tag{1}$$

In Formula 1, fθ is the proposed PASTGCN model, and the parameter is represented by θ. The goal is to make the predicted output as close to the true value as possible.

All modules of this model are jointly trained in an end-to-end manner, fully integrating spatial topology, position information, temporal dynamics, and periodic patterns. The overall architecture of this model is shown in Figure 1.

The core architecture of the PASTGCN model applied in this paper consists of a spatio-temporal position embedding module, a multi-granularity periodic feature branch, and stacked spatio-temporal synchronized convolutional blocks. The model first maps node geographic coordinates into a d-dimensional spatial embedding using a spatial encoder. Simultaneously, timestamp features are converted into temporal embeddings using an embedding lookup table and sinusoidal position encoding. These two are jointly mapped to generate a unified spatio-temporal position encoding as a priori input. To explicitly capture the periodicity of traffic flow, the model designs parallel dilated convolutional branches, employing daily and weekly TCN structures to extract multi-scale temporal features.

The backbone network consists of four layers of spatio-temporally synchronized convolutional blocks, each employing an alternating spatial-temporal modeling strategy. In the spatial modeling stage, a dual graph learning mechanism is used to fuse a static road network distance graph with an attention-based dynamic graph, generating a spatial representation through graph convolution. In the temporal modeling stage, dilated causal convolutions are used to extract multi-scale temporal patterns. Spatial representations are then injected into each layer to achieve spatially-aware temporal modeling. Residual connections ensure gradient flow. Multi-granularity periodic features are dynamically fused with backbone features through a gating mechanism, and a 1×1 convolution outputs the prediction for the next step. The model jointly optimizes the Mean Absolute Error (MAE) loss, the periodic residual reconstruction loss, and a dynamic graph regularization term, and its performance is validated on a real-world traffic dataset. The entire architecture, through end-to-end training, precisely models dynamic spatio-temporal dependencies in complex road networks.

B. SPATIO-TEMPORAL POSITION EMBEDDING MODULE

C. SPATIAL POSITION EMBEDDING

To enhance the model's ability to perceive the geographical distribution characteristics and topological role differences of sensor nodes in urban road networks, this paper designs an explicit spatial position embedding module. Traditional GNNs rely only on adjacency relationships for message transmission, ignoring the absolute position information of nodes and their structural importance in the network, resulting in limited modeling capabilities for long-distance dependencies and non-local spatial patterns.

This module jointly encodes the geographic coordinates and topological attribute features of each node into a learnable spatial position vector as the prior spatial knowledge input of the model. The original coordinates are normalized and then concatenated with the topological attributes to form a joint spatial feature:

$$f_i^{space} = [p_i, a_i] \in \mathbb{R}^{2+D_a} \tag{2}$$

A nonlinear mapping is performed through a parameterized multi-layer perceptron:

$$e_i^{space} = MLP_{space}(f_i^{space}) \in \mathbb{R}^d \tag{3}$$

The Euclidean distance between any two nodes is defined as:

$$\delta_{ij} = \|p_i - p_j\|_2 \tag{4}$$

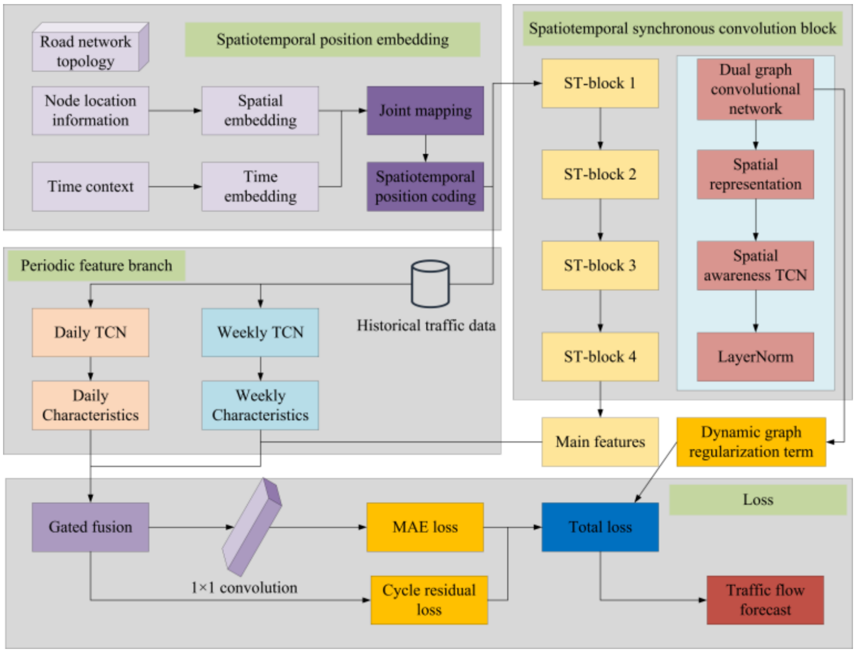

FIGURE 1. Overall architecture of this model

A radial basis function kernel is used to map it to a relative position offset:

$$\phi(\delta_{ij}) = \exp\left(-\frac{\delta_{ij}^2}{\sigma^2}\right) \tag{5}$$

In Formula 5, σ is the scale parameter that controls the decay rate.

In this paper, geographic coordinates (such as latitude and longitude) reflect the absolute physical location of nodes, used to calculate Euclidean distance and construct distance-based static spatial relationships.

Topological attributes describe the role and connection patterns of nodes within the network structure, respectively characterizing the spatial geometric information and structural functional information of the location. By jointly encoding these two into a unified node representation, the model can simultaneously perceive the absolute location of a node and its structural importance within the network, thus more precisely distinguishing between functionally similar but geographically dispersed nodes, or nodes geographically close but with different topological roles.

Compared to standard fixed node embeddings, the proposed joint location-aware embedding offers the following advantages: First, it injects geographic and topological priors into the model in a learnable manner, enhancing its ability to capture long-distance dependencies and non-local spatial patterns, enabling the model to identify traffic associations beyond local neighborhoods. Second, this embedding, in conjunction with temporal embeddings, forms a spatiotemporally consistent node representation, allowing the model to dynamically adjust the spatial semantics of nodes across different time periods, thereby more accurately modeling the nonlinear evolution of traffic states, such as functional changes in the same road segment during morning and evening rush hours. This explicit spatiotemporal location awareness mechanism enables PASTGCN to better distinguish between periodic patterns and dynamic fluctuations caused by sudden events, thereby improving the robustness and accuracy of predictions.

D. PERIODIC TIME EMBEDDING

This paper designs a structured periodic time embedding module that explicitly converts the original timestamp into a low-dimensional vector representation rich in periodic semantics. Unlike simple time step indexing, this module fully exploits the hierarchical structural characteristics of time, distinguishes time semantics of different granularities, and fuses them into a unified time position encoding through a learnable mechanism.

Categorical time variables: including binary flags such as hour (0–23), week (0–6), whether it is a weekday, whether it is a holiday, whether it is in the morning and evening rush hour, etc. For each categorical variable ck∈C, it is mapped to a dk-dimensional vector through a learnable embedding lookup table:

$$e_k = E_k[c_k], \quad E_k \in \mathbb{R}^{|V_k| \times d_k} \tag{6}$$

Continuous periodic variable: "time of week", using sine/cosine position encoding to preserve its periodic continuity:

$$PE(x)_i = \begin{cases} \sin\left(2\pi x \cdot 10000^{-2i/d}\right), & \text{if } i \text{ even,} \\ \cos\left(2\pi x \cdot 10000^{-2i/d}\right), & \text{if } i \text{ odd.} \end{cases} \tag{7}$$

After concatenating all categorical embeddings with the continuous encoding vector, they are input into a shared multi-layer perceptron for nonlinear fusion:

$$e_t^{time} = MLP_{time}([e_1, e_2, \ldots, e_K, PE(x_1), \ldots]) \in \mathbb{R}^d \tag{8}$$

E. JOINT SPATIO-TEMPORAL POSITION ENCODING

To achieve collaborative perception of spatial layout and temporal regularity in traffic systems, this paper proposes a learnable joint spatio-temporal position encoding mechanism that deeply fuses independently modeled spatial position embeddings with periodic time embeddings. The spatial and temporal embedding vectors are concatenated in the feature dimension:

$$z_i^{joint} = [e_i^{space}, e_t^{time}] \in \mathbb{R}^{2d} \tag{9}$$

Nonlinear fusion and dimensionality compression are performed through a parameterized multi-layer perceptron:

$$E_i^{pos} = MLP_{fusion}(z_i^{joint}) \in \mathbb{R}^d \tag{10}$$

This joint encoding mechanism breaks through the limitation of traditional models that process spatio-temporal information separately and constructs a deep spatio-temporal modeling paradigm for position perception.

F. SPATIO-TEMPORAL SYNCHRONOUS CONVOLUTION BLOCK

G. SPATIAL MODELING: DUAL GRAPH LEARNING MECHANISM

This paper applies a dual graph learning mechanism in the spatio-temporal synchronous convolution block to collaboratively model static topological constraints and time-varying functional associations. The static graph is a structural graph driven by geographical priors. Based on the geographical distribution of sensor nodes, a static adjacency matrix determined by physical distance is constructed to encode the topological prior of the road network.

For any node pair (vi,vj), its connection weight is defined as the Gaussian kernel function:

$$A_{ij}^{static} \&= \begin{cases} \exp\left(-\frac{\|p_i - p_j\|^2}{\sigma^2}\right), & \text{if} \|p_i - p_j\| < \tau \\ 0, & \text{otherwise} \end{cases} \tag{11}$$

In Formula 11, pi and pj are longitude and latitude coordinates; σ controls the decay rate; τ is the cutoff threshold (3 kilometers), ensuring sparse connectivity.

The Gaussian weighted adjacency matrix of the road network distance is shown in Figure 2.

The Gaussian weighted adjacency matrix of road network distance is a key tool for modeling spatial dependencies in traffic prediction. The Euclidean distance between nodes is converted into continuous weights through the Gaussian kernel function, otherwise it is 0. The truncation threshold is set to 3 kilometers to maintain the sparsity of the matrix. The matrix has strict symmetry (the other half of the symmetric data is not shown), reflecting the bidirectional characteristics of traffic impact, while capturing local spatial dependencies through continuous weights. The closer the distance, the higher the interaction intensity of nodes, which conforms to the physical law that the traffic flow of adjacent sections in the real road network is more correlated.

This paper uses Euclidean distance between nodes, rather than path length or travel time. Physical distance was chosen because it offers stability and computational simplicity in static graph construction and effectively reflects the direct impact of geographical proximity on traffic conditions.

To address the sensitivity of the Gaussian kernel variance parameter σ^2 to the distance threshold τ, this experiment conducted parameter ablation analysis on the METR-LA dataset. When σ^2 varies within the range of 0.1 to 1.0, the model performance (RMSE) fluctuation is less than 0.05 (veh/5min), indicating that the model has a certain robustness to σ^2. Ultimately, $\sigma = 0.5$ (i.e., $\sigma^2 = 0.25$) was chosen to balance the strength and decay rate of local correlations. The distance threshold τ was set to 3 kilometers, a value based on statistical analysis of the traffic impact range in actual road networks.

The dynamic graph is a feature-driven time-varying dependency graph, and the soft attention mechanism is used to calculate the dependency intensity between nodes:

$$\alpha_{ij} = \frac{\exp\left(\text{LeakyReLU}\left(a^T[Wh_i^{(k)} \| Wh_j^{(k)}]\right)\right)}{\sum_{l=1}^{N} \exp\left(\text{LeakyReLU}\left(a^T[Wh_i^{(k)} \| Wh_l^{(k)}]\right)\right)} \tag{12}$$

In Formula 12, a is the learnable attention vector, and the obtained attention weights constitute the dynamic adjacency matrix.

The static and dynamic graphs are convexly combined to generate a fused adjacency matrix:

$$\tilde{A}^{(k)} = \lambda \cdot \hat{A}^{static} + (1 - \lambda) \cdot A_{dynamic}^{(k)}, \quad \lambda \in (0, 1) \tag{13}$$

The first-order graph convolution is used to perform message passing on the fused graph:

$$H' = \sigma\left(\tilde{D}^{-1/2} \tilde{A} \tilde{D}^{-1/2} H^{(k)} W^{(k)}\right) \tag{14}$$

In Formula 14, H' is the updated node representation, which incorporates non-local spatial context information.

The proposed dual-graph learning mechanism does not simply combine static and dynamic graphs linearly.

Instead, it employs a collaborative modeling mechanism to adaptively capture spatiotemporal functional dependencies while maintaining geographic topological rationality. Specifically, the static Gaussian kernel graph ensures the physical interpretability and local continuity of spatial relationships, while the dynamic attention graph identifies

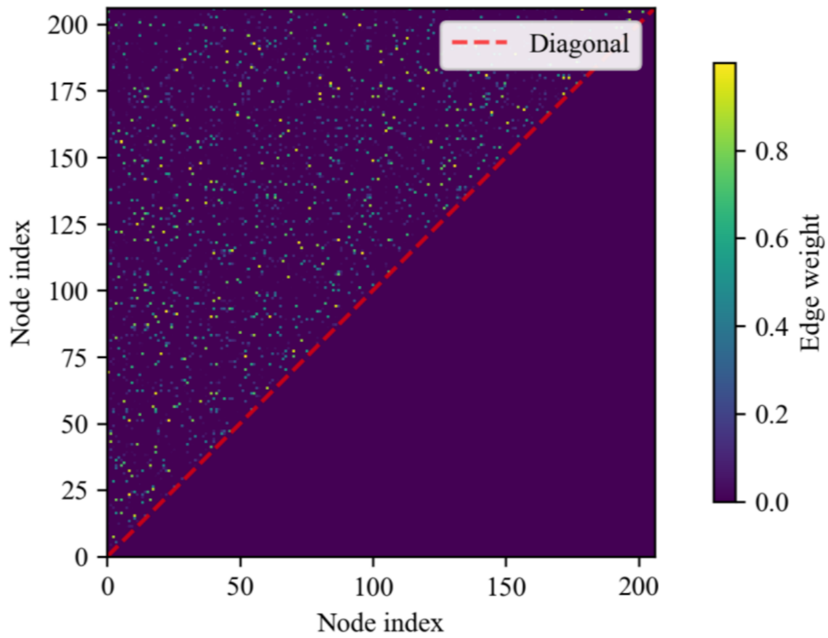

FIGURE 2. Gaussian weighted adjacency matrix of road network distance

nonlocal, time-varying correlations. The two are weighted by a learnable fusion coefficient λ and constrained by a structure regularization term to prevent the dynamic graph from deviating excessively from geographic priors, thereby enhancing the model's stability and generalization ability.

Compared to existing methods that rely solely on dynamic attention or adaptive adjacency matrices, this mechanism explicitly injects geographic semantics and topological constraints while maintaining the flexibility of dynamic modeling, thus achieving more accurate and stable spatial dependency modeling in complex traffic scenarios.

H. TEMPORAL MODELING: SPATIAL-AWARE CAUSAL CONVOLUTION

This paper adopts dilated causal convolution as the core component of temporal modeling in the spatio-temporal synchronous convolution block. Causal convolution ensures that the output at any moment depends only on the current and historical inputs, satisfying the causal constraints of time series prediction; the dilation mechanism captures long-term dependencies through exponentially growing receptive fields, and can cover a long historical window without stacking a large number of layers.

In the k-th layer of the spatio-temporal synchronous convolution block, a one-dimensional causal convolution with K=3 is first performed independently on each node, with a dilation rate of dk=2k-1, to achieve multi-scale temporal pattern extraction:

$$Y = \mathrm{Conv1D}_{causal}(X; \text{kernel size} = 3, \text{dilation} = d_k) \tag{15}$$

This paper proposes a spatial-aware temporal modeling paradigm, which explicitly injects the spatial representation from the graph convolution after each layer of temporal convolution. H' is expanded into a time series form through a learnable projection matrix:

$$\mathrm{Proj}(H') = H'W_p \in \mathbb{R}^{N\times d} \tag{16}$$

It is replicated T times along the time dimension to form $\tilde{H}' \in \mathbb{R}^{N\times T\times d}$, and is added element-wise to the temporal convolution output:

$$Z = Z_{temp} + \tilde{H}' \tag{17}$$

A spatial gating bias is applied in temporal modeling so that the feature update at each time step is modulated by the current spatial dependency structure.

I. RESIDUAL CONNECTION AND INTERLAYER STACKING

To construct a deep and stable spatio-temporal feature extraction backbone network, this paper designs a standardized spatio-temporal block (ST-Block) as the basic building block and achieves efficient training of multi-layer networks through residual connection and stacking strategy.

H' is projected and superimposed on the temporal feature through the spatial injection mechanism to obtain the intermediate output Z. Then, layer normalization and dropout are applied to improve training stability and generalization ability:

$$Z_{norm} = \mathrm{LayerNorm}(Z) \tag{18}$$

$$Z_{drop} = \mathrm{Dropout}(Z_{norm}) \tag{19}$$

The input and output are fused through the identity residual connection:

$$X_{out} = X_{in} + F(X_{in}) \tag{20}$$

Multiple ST-Blocks are stacked layer by layer to form the backbone encoder. Assuming there are K layers in total, the overall forward propagation processes are:

$$X^{(1)} = \text{Embedding}(X) \tag{21}$$

$$X^{(k+1)} = X^{(k)} + \text{ST-Block}^{(k)}(X^{(k)}), \quad k = 1, 2, \ldots, K \tag{22}$$

J. MULTI-GRANULARITY PERIODIC FEATURE EXTRACTION BRANCH

This paper designs a multi-granularity periodic feature extraction branch to explicitly model two main inherent periodic patterns in traffic flow: daily cycle (24 hours) and weekly cycle (7 days). These two granularities are chosen based on the observation that traffic flow exhibits distinct morning and evening peaks, off-peak hours, and nighttime troughs within a single day (daily cycle), while systematic differences exist between weekdays and weekends (weekly cycle).

The daily cycle branch uses a 5-minute time step and constructs a receptive field of approximately 288 steps (24 hours) by stacking multiple dilated causal convolutional layers, focusing on capturing recurring patterns within the day, such as commuting tides and midday off-peak hours. The weekly cycle branch, using the day as the basic unit, expands to a 7-day range with a larger dilation rate (d=288) to model traffic differences between weekdays and weekends, as well as cross-day patterns such as holidays.

Urban traffic flow has multi-scale periodicity [35-36], with daily cycle (24 hours) and weekly cycle (7 days).

This paper designs a parallel multi-granularity periodic feature extraction branch to explicitly model periodic components at different time scales, enhancing the model's ability to decouple and reconstruct regular traffic patterns.

Two parallel lightweight dilated causal convolutional subnetworks are used to extract diurnal and weekly periodic features, respectively.

The diurnal branch uses a dilation rate of d=12 and stacks multiple layers to construct a receptive field of approximately 288 steps, focusing on capturing daily recurring patterns such as morning and evening rush hours and commuting patterns.

The weekly branch uses a larger dilation rate of d=288 to extract cross-daily periodic trends, such as weekday and weekend traffic differences and holiday effects.

Each branch consists of several layers of TCNs equipped with gated activation functions to enhance nonlinear representation, ultimately outputting a low-dimensional periodic feature tensor.

To effectively integrate the backbone spatio-temporal features with the main network, this paper proposes a gated fusion mechanism that adaptively balances the contributions of the backbone dynamic prediction and the periodic prior. The three are concatenated in the feature dimension:

$$C = [F_{min}, F_{daily}, F_{weekly}] \in \mathbb{R}^{N \times T_{pred} \times 3d} \tag{23}$$

Gate weights are generated through projection on the learnable weight matrix and Sigmoid activation:

$$G = \sigma(CW_g) \in \mathbb{R}^{N \times T_{pred} \times d} \tag{24}$$

Fused features are combined through gated weighting:

$$F_{fused} = G \odot F_{main} + (1 - G) \odot (F_{daily} + F_{weekly}) \tag{25}$$

K. MULTI-TASK OUTPUT LAYER AND LOSS FUNCTION

L. OUTPUT PREDICTION

After obtaining the final spatio-temporal-periodic joint representation F_{fused} in the feature fusion stage, it is mapped to the specific traffic state prediction value $\hat{Y} \in \mathbb{R}^{N \times T_{pred} \times F_{out}}$. This paper uses a lightweight 1×1 convolutional layer as the output decoder.

The high-dimensional fusion feature Ffused is projected into the original data space to generate the prediction results for the next Tpred time steps:

$$\hat{Y} = [\hat{X}_{t+1}, \hat{X}_{t+2}, \ldots, \hat{X}_{t+T_{pred}}], \quad \hat{X}_{t+\tau} \in \mathbb{R}^{N \times F_{out}} \tag{26}$$

M. MULTI-TASK LEARNING OBJECTIVES

This paper adopts a multi-task learning framework to jointly optimize the main prediction task and the auxiliary reconstruction task, and applies a structural regularization term to constrain the dynamic graph learning behavior. The mean absolute error that is robust to outliers is used as the main loss function:

$$L_{pred} = \frac{1}{NT_{pred}} \sum_{i=1}^{N} \sum_{t=1}^{T_{pred}} |\hat{y}_{i,t} - y_{i,t}| \tag{27}$$

In Formula 27, $\hat{y}$i,t and yi,t are the predicted value and the true value of node i at time t, respectively. The periodic residual reconstruction is designed as an auxiliary task, and the periodic baseline is constructed.

The sliding window average is used:

$$\bar{X}_t = \text{PeriodicAverage}(X_{t-k \cdot T_{daily}}, k = 1, 2, \ldots, 7) \tag{28}$$

The residual term is calculated:

$$\Delta X_t = X_t - \bar{X}_t \tag{29}$$

This residual reflects the instantaneous fluctuation component that deviates from the periodic law. The model adds an additional branch to the backbone network to predict the residual and applies L2 reconstruction loss:

$$L_{res}=\frac{1}{NT_{pred}}\sum_{i=1}^{N}\sum_{t=1}^{T_{pred}}\left\|\hat{\Delta}x_{i,t}-\Delta x_{i,t}\right\|_2^2 \tag{30}$$

Combining the main task, auxiliary tasks, and graph structure constraints, the total loss function is defined as:

$$L = L_{pred} + \beta L_{res} + \gamma R_{graph} \tag{31}$$

In Formula 31, β>0 and γ>0 are task weight hyperparameters that balance the contribution of each objective.

The third term is the dynamic graph regularization term:

$$R_{graph}=\|A_{dynamic}\|_F^2=\sum_{i=1}^{N}\sum_{j=1}^{N}\left(\alpha_{ij}\right)^2 \tag{32}$$

N. DYNAMIC GRAPH STRUCTURE REGULARIZATION

To constrain the structural rationality of the dynamic attention graph and prevent overfitting or the formation of non-physical dependencies, this paper introduces a Frobenius norm-based structure regularization term to explicitly constrain the dynamic adjacency matrix $A_{dynamic}$.

The formula for this regularization term is $R_{graph} = \|A_{dynamic}\|_F^2 = \sum_{i=1}^{N}\sum_{j=1}^{N}\alpha_{ij}^2$, where α_{ij} is the dynamic attention weight between node i and node j. This regularization penalizes the sum of squares of the attention weights, ensuring that the dynamic graph maintains a certain degree of sparsity and avoiding globally dense connections or over-attention to irrelevant nodes. Simultaneously, combined with the constraints of the static geographical prior graph, this mechanism ensures that the dynamic graph does not significantly deviate from the spatial logic of the actual road network during the learning process, thereby improving the model's generalization ability and interpretability in complex traffic scenarios.

The structural regularization term $R_{graph} = \|A_{dynamic}\|_F^2$ penalizes the Frobenius norm of the dynamic attention weight matrix, aiming to constrain the sparsity and smoothness of the dynamic graph. Specifically:

Sparseness constraint: This term penalizes the sum of squares of all attention weights, prompting the model to compress insignificant or redundant weights to near zero while maintaining necessary connections, thus avoiding globally dense connections and enhancing the interpretability and computational efficiency of the dynamic graph.

Smoothness constraint: Combined with the static geographical prior graph, this regularization encourages the dynamic graph to maintain a certain degree of spatial continuity, i.e., geographically proximate or functionally similar nodes should have similar attention distributions, avoiding non-physical, abrupt dependencies.

During training, this regularization is optimized together with the main loss, ensuring that the dynamic graph learns time-varying functional associations without deviating from the actual topological logic of the road network, thereby improving the model's generalization ability and stability in complex traffic scenarios.

IV. EXPERIMENTAL SETUP

A. DATASETS

To comprehensively evaluate the model's predictive performance and generalization capabilities, the experiment uses four public, real-world urban traffic datasets: METR-LA (primarily), supplemented by PeMS-BAY, PeMSD4, and PeMSD8. These datasets cover diverse urban areas, road network sizes, and traffic characteristics.

METR-LA: collected from the Los Angeles County freeway system, it contains 207 fixed sensors that record vehicle speed data every five minutes.

PeMS-BAY: a monitoring network deployed by the California Department of Transportation (Caltrans) in the San Francisco Bay Area, encompassing 325 monitoring stations. It provides multi-dimensional information such as traffic flow and occupancy at a five-minute granularity, and exhibits strong periodicity and regional linkage.

PeMSD4 and PeMSD8: subsets of the California Traffic Performance Monitoring System in District 4 (North Bay) and District 8 (Southern California), respectively, containing 307 and 170 sensors. They focus on modeling short-term fluctuations in local road networks and are suitable for high-resolution prediction validation.

Dataset statistics are shown in Table 1.

TABLE 1. Dataset statistics

Dataset	Number of nodes	Number of edges	Total time steps
METR-LA	207	1515	34272
PeMS-BAY	325	2369	52116
PeMSD4	307	1876	16992
PeMSD8	170	1231	17856

All datasets are processed with missing value interpolation, Z-score normalization, and timestamp alignment, and a static adjacency graph is constructed based on geographic distance. Cross-dataset comparison experiments are conducted to verify the robustness and transferability of the model. The dataset is divided into training set, validation set, and test set according to a 3:1:1 ratio.

To ensure the rigor of the evaluation and eliminate the possibility of data leakage, this paper adopts a strict time-order partitioning strategy: the dataset is divided into a training set (first 60%), a validation set (middle 20%), and a test set (last 20%) in chronological order, ensuring that the test set data is completely invisible during training. The reported high R^2 value is mainly due to the explicit spatiotemporal location encoding and multi-granularity periodic branch design in the PASTGCN model. These modules can more accurately capture the inherent periodicity and spatiotemporal dependence in traffic flow, thus achieving a high goodness of fit on the test set. At the same time, the model effectively alleviates overfitting through dynamic graph regularization, multi-task learning, and early stopping strategies, ensuring its generalization performance.

Compared with the baseline model, PASTGCN is competitive in RMSE, while the significant improvement in R2 reflects its stronger interpretability and structural adaptability.

B. PREPROCESSING

This paper uses linear interpolation to fill in consecutive missing segments in the time dimension:

$$x(t) = x(t_1) + \frac{t - t_1}{t_2 - t_1}\left(x(t_2) - x(t_1)\right), \quad t \in (t_1, t_2) \quad (33)$$

In Formula 33, x(t1) and x(t2) are the most recent valid observations. For residual missing values after interpolation, the previous and next neighboring values are used to fill in the missing values.

Traffic data may contain abnormal readings caused by equipment errors or emergencies. The Z-score method is used to identify and eliminate statistically significant outliers:

$$z_{i,t} = \frac{x_{i,t} - \mu_i}{\sigma_i} \quad (34)$$

In Formula 34, μ_{i} and σ_{i} are the mean and standard deviation of the node in the historical data, respectively. If |zi,t|>3, xi,t is determined to be an outlier and replaced by the linear interpolation result.

C. BASELINE MODEL AND EVALUATION INDICATORS

To fully analyze the performance of this model, it is compared with the existing model. The models used for comparison are: DCRNN, STGCN, Graph WaveNet, ASTGCN, GMAN, MTGNN (Multivariate Time Series Forecasting with Graph Neural Networks), AGCRN (Adaptive Graph Convolutional Recurrent Network), and STAEformer (Spatio-Temporal Adaptive Embedding Transformer). To comprehensively and objectively evaluate the prediction performance of the models, this paper adopts four evaluation indicators that are widely used in traffic fl ow prediction tasks. MAE measures the mean absolute deviation between the predicted value and the true value, and the formula is:

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n} |\hat{y}_i - y_i| \quad (35)$$

RMSE reflects the square root of the mean square of the prediction error, and the formula is:

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n} (\hat{y}_i - y_i)^2} \quad (36)$$

MAPE is:

$$\mathrm{MAPE} = \frac{1}{n}\sum_{i=1}^{n} \left|\frac{\hat{y}_i - y_i}{y_i}\right| \times 100\% \quad (37)$$

The formula for the coefficient of determination R2 is:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \quad (38)$$

D. EXPERIMENTAL DETAILS

The PASTGCN model applied in this paper is implemented based on the PyTorch deep learning framework and is combined with PyTorch Geometric to efficiently construct graph convolution operations. The optimizer used is Adam, with an initial learning rate of 1×10-3 and momentum parameters of β1=0.9 and β2=0.999.

The learning rate strategy used is ReduceLROnPlateau; the training batch size is set to 32; the maximum training epochs are 100; early stopping is used to prevent overfitting. The model input history window length is L=12 (corresponding to 1 hour of 5-minute granularity data), and the prediction time is 12 steps ahead (that is, 60 minutes). The hidden layer dimension is uniformly set to d=64, and four layers of spatio-temporal synchronized convolutional blocks (ST-Blocks) are stacked.

The model parameter settings for this paper are shown in Table 2.

TABLE 2. Model parameter settings

Parameter category	Hyperparameter	Value
Model structure	History window length	12
	Prediction step	12
	Hidden dimension	64
	Number of ST-Block layers	4
	Temporal convolution kernel size	3
	Dilation rate	[1, 2, 4, 1] (loop per layer)
	Optimizer	Adam
	Initial Learning Rate	1×10^{-3}
Optimize configuration	β_1, β_2	0.9, 0.999
	Batch size	32
	Maximum training rounds	100
	Early stopping	patience = 10

V. RESULTS

A. TRAFFIC FLOW PREDICTION PERFORMANCE

The results of the traffic flow prediction loss on the METR-LA dataset are shown in Figure 3.

In Figure 3, the horizontal axis represents epochs, and the vertical axis represents loss. PASTGCN achieves faster loss reduction and convergence to a lower level, primarily due to its innovative position-aware spatio-temporal modeling architecture. The explicit spatio-temporal position embedding module, by fusing geographic coordinates with temporal semantic encoding, provides the model with strong prior knowledge, accelerating feature learning during the initial

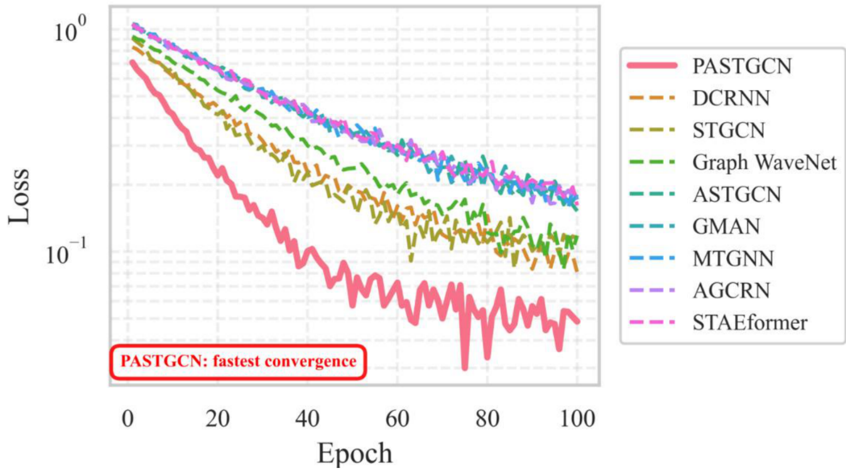

FIGURE 3. Traffic flow prediction loss change

training phase. The dual graph learning mechanism, which coordinates static road network constraints with dynamic attention dependencies, enables spatial modeling to maintain geographic plausibility while capturing functional relationships. This results in more precisely spatial relationship representation than baseline models with a single graph structure (STGCN uses only a static graph, while DCRNN relies on a diffusion process). The spatial injection mechanism in the spatio-temporal synchronous convolutional block enables spatial-aware modeling in the temporal dimension, avoiding the feature interaction loss associated with traditional serial architectures. Furthermore, multi-granularity periodic branches explicitly decouple daily/weekly cyclical patterns, enhancing modeling of long-term traffic flow patterns through a gated fusion mechanism. A multi-task learning framework jointly optimizes the main prediction task and periodic residual reconstruction, further improving model generalization. In contrast, pure attention models such as GMAN are prone to illogical spatial correlations due to their lack of geographic constraints, and general-purpose graph networks such as MTGNN are not specifically designed for the cyclical nature of traffic data, requiring more training rounds to achieve suboptimal convergence. PASTGCN achieves efficient learning and precise prediction of complex traffic dynamics by systematically integrating position priors, dynamic graph learning, and periodic modeling.

This paper uses t-SNE to visualize the distribution and clustering of samples from different time periods in low-dimensional space. Information on traffic flow characteristics for each time period is shown in Table 3.

TABLE 3. Information on traffic flow characteristics for each time period

Type	Time period	Features
Morning peak	7:00–10:00	Concentrated travel, heavy commuter traffic, and severe congestion
Evening peak	17:00–20:00	High road loads during return trips
Off-peak	10:00–17:00	Stable traffic flow with minimal fluctuations
Nighttime low	20:00–7:00	Low traffic flow, with some roads near empty

The t-SNE visualization results are shown in Figure 4.

PASTGCN exhibits clearer time-of-day cluster separation in the t-SNE visualization. This is fundamentally due to the model architecture's ability to precisely model spatio-temporal heterogeneity. The position-aware embedding module integrates geographic coordinates and topological attributes to reflect differences in road network status at different times of day in the spatial distribution. For example, during the morning rush hour, nodes on main roads form dense clusters due to commuting pressure, while during the evening off-peak hours, the embedding vectors of these nodes shift due to a sudden drop in traffic. The dynamic attention in the dual graph learning mechanism captures time-of-day dependencies, automatically clustering nodes with similar functions in a low-dimensional space. The static geographic distance constraint prevents physical position distortion caused by excessive clustering. In contrast, DCRNN relies solely on diffusion processes, making it difficult to distinguish the directional differences between morning and evening peaks.

STGCN's fixed graph structure cannot adapt to the dynamic correlations between different time periods.

While ASTGCN incorporates attention, its lack of a position prior leads to blurred cluster boundaries.

PASTGCN's periodic branch explicitly encodes daily/weekly cycles, clearly distinguishing the stable patterns during off-peak hours from the nighttime trough in the latent space. Traditional models, however, lack sufficient coupling of periodic and spatio-temporal features, resulting in overlapping samples from these periods.

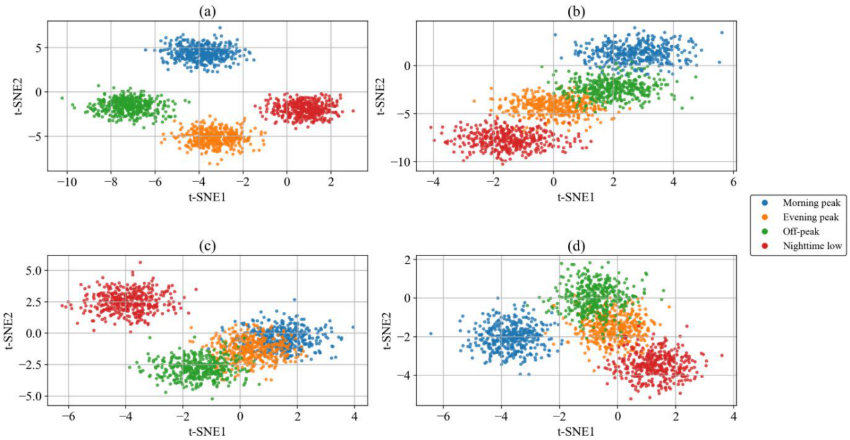

FIGURE 4. t-SNE visualization results; Figure 4(a). PASTGCN; Figure 4(b). DCRNN; Figure 4(c). STGCN; Figure 4(d). ASTGCN

B. COMPARISON BETWEEN PREDICTION AND ACTUAL RESULTS

Figure 5 shows a comparison of predictions and results for various models in the METR-LA dataset.

The PASTGCN model's prediction curves show high consistency with real traffic flow in both overall trends and local fluctuations, outperforming leading baseline models such as DCRNN, STGCN, and ASTGCN.

Specifically, PASTGCN precisely captures the rate of change and peak intensity during peak traffic periods and maintains stable predictions during nighttime troughs and off-peak periods, without significant overshoot or lag. In contrast, DCRNN overreacts to local fluctuations, exhibiting some oscillation. STGCN and ASTGCN exhibit phase lag during peak periods. The superior performance of PASTGCN is primarily attributed to its position-aware mechanism and dual graph learning architecture. By explicitly encoding node geographic coordinates and temporal semantics, the model enhances prior knowledge of spatial layout and periodic patterns. The weighted fusion of dynamic and static graphs effectively identifies time-varying spatial dependencies, improving the accuracy of congestion propagation modeling. Furthermore, the spatial-aware temporal convolution and multi-task learning framework further enhance the deep coupling of spatio-temporal features and the ability to reconstruct residual dynamics. As a result, PASTGCN not only accurately fits periodic baselines but also sensitively responds to short-term non-stationary changes, reducing prediction bias and achieving more robust and precise traffic state deduction.

C. ABLATION EXPERIMENT RESULTS

Ablation experiments are conducted on the METR-LA dataset to analyze the role of each module in this paper's model. The results are shown in Table 4.

TABLE 4. Ablation experiment results

Model	PosEmb	DynGraph	Spatial Injection	Periodic Branch	Multi Task	RMSE (veh/5min)	R^2
PASTGCN	✓	✓	✓	✓	✓	2.53	0.982
w/o PosEmb	×	✓	✓	✓	✓	2.79	0.961
w/o DynGraph	✓	×	✓	✓	✓	2.87	0.968
w/o Spatial-Injection	✓	✓	×	✓	✓	2.81	0.954
w/o Period-icBranch	✓	✓	✓	×	✓	2.72	0.943
w/o MultiTask	✓	✓	✓	✓	×	2.68	0.963
w/o AllCompo-nents	×	×	×	×	×	3.15	0.892

Notes: PASTGCN is the complete model, including all proposed modules; w/o PosEmb removes the spatio-temporal position embedding module; w/o DynGraph uses only the static graph, removing dynamic graph learning; w/o SpatialInjection removes the temporal convolution process for spatial feature injection; w/o PeriodicBranch removes the multi-granularity periodic feature extraction branch; w/o MultiTask retains only the main task loss and removes periodic residual reconstruction; w/o AllComponents retains only the backbone GCN+TCN without any improved modules.

Ablation experiments show that PASTGCN outperforms all variants on the METR-LA dataset, validating the effectiveness and synergy of each module. The complete model achieves the best performance, achieving an RMSE of 2.53 (veh/5min) and an R2 of 0.982, surpassing the ablation version across the board, demonstrating that the proposed components together form an efficient spatio-temporal modeling system. Removing dynamic graph learning (w/o DynGraph) results in an increase in RMSE to 2.87 (veh/5min) and a decrease in R2 by 0.014, demonstrating that the adaptive attention mechanism plays a crucial role in capturing time-varying spatial dependencies such as congestion propagation and regional linkages. In contrast, removing spatial-aware temporal convolution (w/o SpatialInjection) also results in a significant performance loss, suggesting that traditional serial spatio-temporal modeling suffers from information

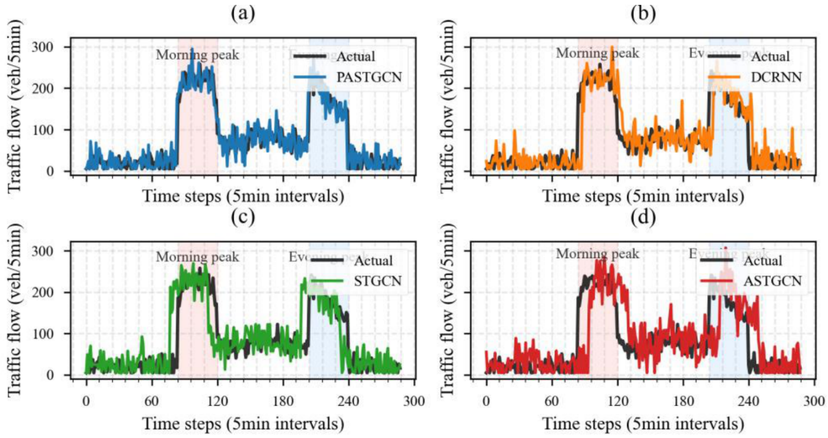

FIGURE 5. Comparison of predictions and results for the METR-LA dataset; Figure 5 (a). Predictions and results of the PASTGCN model; Figure 5 (b). Predictions and results of the DCRNN model; Figure 5 (c). Predictions and results of the STGCN model; Figure 5 (d). Predictions and results of the ASTGCN model

fragmentation. The proposed feature injection mechanism effectively enables continuous modulation of temporal dynamics by spatial context, improving the model's responsiveness to non-stationary processes.

Further analysis reveals that omitting spatio-temporal position embedding (w/o PosEmb) increases the RMSE to 2.79 (veh/5min) and decreases the R^2 to 0.961, demonstrating that explicitly encoding geographic coordinates and temporal semantics significantly enhances the model's prior understanding of node functional roles and cyclical patterns, avoiding the instability associated with relying solely on data-driven learning. While the impact of multi-task learning and the removal of periodic branches is relatively minor, it remains significant. The w/o PeriodicBranch causes an R^2 decrease of 0.039, demonstrating that explicitly modeling multi-scale periodicity improves long-term trend fitting. The performance decrease of w/o MultiTask reflects the role of auxiliary tasks in promoting feature decoupling and training stability. The simplest model of w/o AllComponents performs the worst, highlighting the importance of the integrated architecture in this paper. Modules are collaboratively optimized to achieve high-accuracy traffic flow prediction.

D. EFFICIENCY

Efficiency analysis evaluates the computational overhead and response speed of a model in real-world deployment, measuring its feasibility on edge devices and ensuring high-accuracy predictions while meeting engineering requirements for low latency and high throughput. Figure 6 shows an efficiency comparison on the METR-LA dataset.

The PASTGCN applied in this paper achieves an excellent balance between efficiency and performance. While its FLOPs (3.2G) are slightly higher than STGCN (2.9G), it is lower than complex graph attention models such as DCRNN (5.8G), Graph WaveNet (7.1G), and STAEformer (9.6G). This demonstrates that PASTGCN avoids expensive recursive or self-attention mechanisms in its architectural design and effectively controls computational complexity by employing a lightweight dual graph fusion and spatial injection strategy. In terms of inference latency, PASTGCN achieves 18.7ms, outperforming most baseline models and only slightly higher than STGCN (15.3ms), but significantly lower than GMAN (51.2ms) and STAEformer (62.8ms), demonstrating its excellent real-time performance due to its feedforward architecture. Its throughput reaches 53.5 samples/second, exceeding both GMAN and STAEformer, demonstrating its high efficiency in batch processing scenarios. The feasibility of PASTGCN is reflected in its engineering practicality and deployment flexibility. While slightly delayed compared to STGCN, this yields superior prediction accuracy with limited FLOPs increase, making it suitable for deployment on edge computing units. Compared to moderately complex models such as MTGNN and AGCRN, PASTGCN achieves higher throughput at lower latency, demonstrating its efficient module design and memory-friendly access. In summary, PASTGCN maintains advanced prediction performance while meeting the real-time response (<100ms) and high-concurrency processing requirements of urban-scale traffic systems, demonstrating promising practical application prospects.

E. IMPACT OF AGGREGATION WINDOW

Analyzing the impact of the aggregation window aims to assess the robustness and applicability of the model at different time granularities and to reveal its sensitivity to input history length. Aggregation windows are set to 5 minutes, 20 minutes, 35 minutes, and 50 minutes. The results on the METR-LA dataset are shown in Figure 7.

5-minute aggregation window: PASTGCN outperforms other models with a MAE of 2.53 (veh/5 min) and a MAPE of 11.2%. At this time, the raw data is noisy, requiring high model robustness. PASTGCN, leveraging explicit position encoding and dynamic graph learning, effectively distinguishes local fluctuations from real traffic conditions, avoids overfitting to noise, and demonstrates strong anti-interference capabilities.

20-minute aggregation window: PASTGCN achieves optimal performance with a MAE of 2.38 (veh/5 min) and a

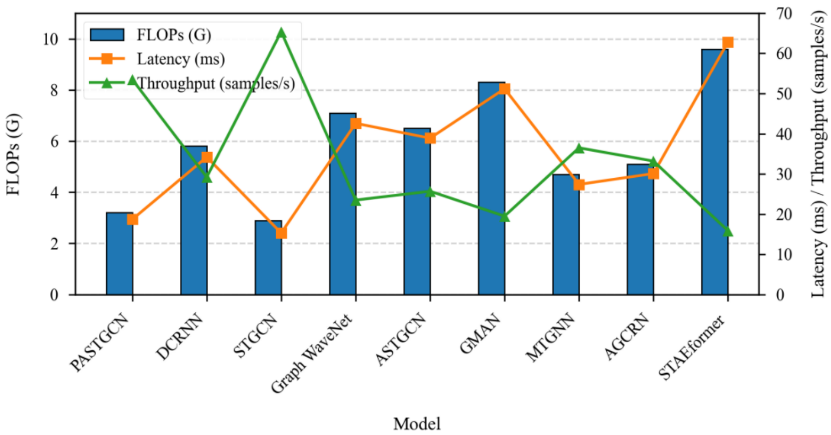

FIGURE 6. Efficiency comparison results

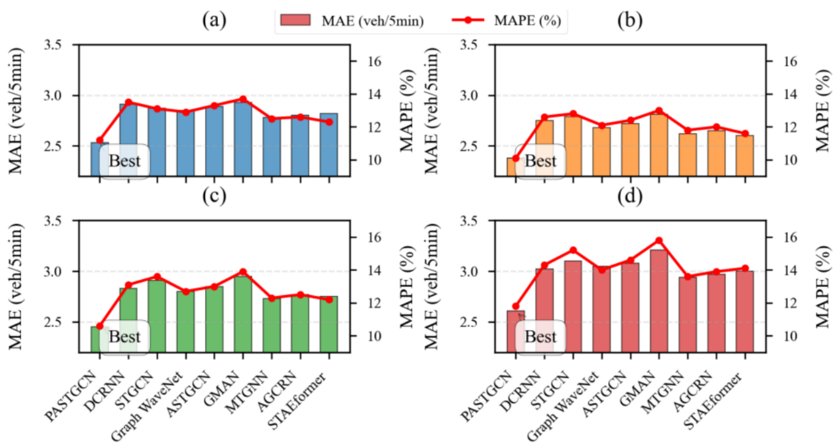

FIGURE 7. Impact of aggregation window; Figure 7(a). 5 min; Figure 7(b). 20 min; Figure 7(c). 35 min; Figure 7(d). 50 min

MAPE of 10.1%, outperforming all baselines. This granularity balances noise suppression and information preservation. Its multi-granularity periodic branches and spatial-aware temporal convolutions fully capture smoothed periodic trends. Simultaneously, the dual graph mechanism precisely models mesoscale spatial dependencies, achieving optimal spatio-temporal synergy.

35-minute aggregation window: as the aggregation granularity increases, traffic dynamic details are further smoothed, but PASTGCN maintains its lead, achieving a MAE of 2.45 (veh/5min) and a MAPE of 10.6%.

Compared to other models, its multi-task framework more accurately models the periodic baseline, and its spatial injection mechanism maintains its effectiveness in modeling medium- and long-term dependencies, ensuring prediction stability.

50-minute aggregation window: over-aggregation of data distorts short-term variations, causing performance degradation for all models. PASTGCN, however, achieves a MAE of 2.61 (veh/5min) and a MAPE of 11.8%. Despite this, its relative advantage remains significant, thanks to its structural regularization and residual reconstruction tasks, which make it more robust to information loss.

F. MULTI-DATASET VALIDATION

Compared with multiple datasets, including PeMS-BAY, PeMSD4, and PeMSD8, with a 5-minute aggregation window, the comparison of traffic flow predictions with actual values is shown in Figure 8.

On three independent datasets, PeMS-BAY, PeMSD4, and PeMSD8, PASTGCN's predictions closely match real-world traffic flows, outperforming mainstream models such as DCRNN and STGCN, demonstrating its superior generalization and cross-domain adaptability. In PeMS-BAY, PASTGCN precisely captures the coordinated congestion evolution of multiple highways, particularly during the morning and evening peak transitions, exhibiting minimal phase delay and amplitude deviation. In PeMSD4 and PeMSD8, despite lower sensor density and more localized traffic dynamics, PASTGCN precisely captures traffic flow fluctuations at complex intersections and ramps, demonstrating its robustness to diverse road network topologies and spatial scales. The model's performance advantage stems primarily from its explicit modeling of position-aware and dynamic spatial dependencies. Through geocoding and dual graph learning, the model can adaptively identify functional associations even without a complete prior graph

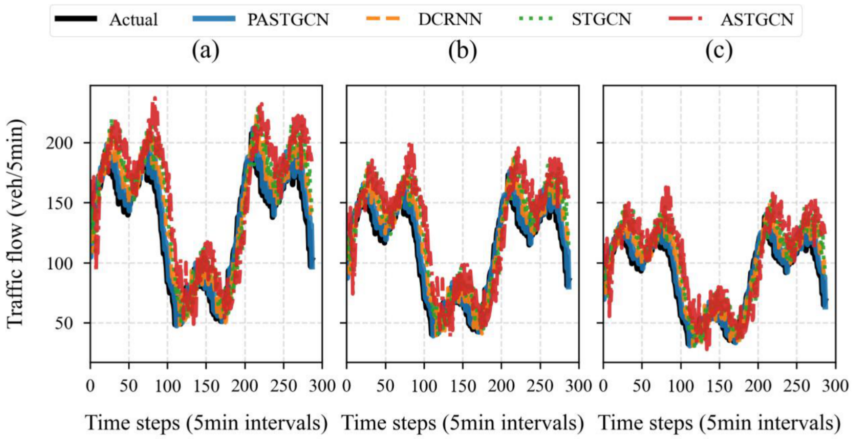

FIGURE 8. Multi-dataset validation; Figure 8(a). PeMS-BAY; Figure 8(b). PeMSD4; Figure 8(c). PeMSD8

structure. Spatial-aware temporal convolution and a multi-task periodic reconstruction mechanism further enhance its ability to extract region-specific periodic patterns. Furthermore, the unified model architecture eliminates the need for structural redesign for different datasets, demonstrating PASTGCN's excellent transferability and deployment potential. Comprehensive results demonstrate that the model not only adapts to diverse urban areas, road network densities, and traffic characteristics, but also consistently maintains optimal prediction performance in complex, heterogeneous environments, demonstrating its broad application prospects.

VI. CONCLUSION

The proposed PASTGCN effectively addresses the shortcomings of traditional models in spatial position-aware, periodic pattern modeling, and the lack of spatial-temporal feature interaction. Its core contributions are: 1) designing explicit spatio-temporal position encoding to deeply integrate geographic coordinates with temporal semantics and enhance the model's prior knowledge; 2) proposing dual graph learning and spatial-aware temporal convolution to achieve collaborative modeling of dynamic spatial dependencies and multi-scale temporal features; 3) applying multi-granularity periodic branches and a multi-task learning framework to explicitly capture the cyclical patterns of traffic flow. Experiments show that PASTGCN outperforms existing models in multiple metrics, such as an RMSE of only 2.53 (veh/5min). This research can provide more precise congestion warnings and route planning support for urban intelligent transportation systems. For the industry, this translates to more reliable logistics, enabling better management of just-in-time manufacturing supply chains and optimizing the distribution of apparel to retail centers. Its efficient computing capabilities are also suitable for edge device deployment, driving the development of real-time and intelligent traffic management. This real-time capability is particularly valuable for dynamic logistics systems,ensuring that transportation assets can be rerouted efficiently in response to sudden congestion.

AUTHOR CONTRIBUTIONS

Conceptualization – Jing Sun; methodology – Jing Sun, Yajuan Liu, Yanlei Dou and Minggang Wang; investigation – Jing Sun; writing-original draft preparation – Jing Sun, Yajuan Liu, Yanlei Dou and Minggang Wang. All authors have read and agreed to the published version of the manuscript.

CONFLICTS OF INTEREST

The authors declare no conflict of interest.

FUNDING

This research received no external funding.

ACKNOWLEDGEMENTS

Not applicable.

REFERENCES

[1] Y. Y. Pu, W. H. Wang, Q. Zhu, and P. P. Chen, "Urban short-term traffic flow prediction algorithm based on CNN-ResNet-LSTM model," Journal of Beijing University of Posts and Telecommunications, vol. 43, no. 5, pp. 9-14, 2020, doi: 10.13190/j.jbupt.2019-243.
[2] H. Li, S. Zhang, B. Cao, and J. Fan, "Traffic flow prediction based on urban road checkpoint data," Journal of Chongqing University, vol. 43, no. 11, pp. 29-40, 2020, doi: 10.11835/j.issn.1000-582X.2020.11.004.
[3] H. Yang, Z. Li, and Y. Qi, "Predicting traffic propagation flow in urban road network with multi-graph convolutional network," Complex & Intelligent Systems, vol. 10, no. 1, pp. 23-35, 2024, doi: 10.1007/s40747-023-01099-z.
[4] S. Wu, "Spatiotemporal dynamic forecasting and analysis of regional traffic flow in urban road networks using deep learning convolutional neural network," IEEE transactions on intelligent transportation systems, vol. 23, no. 2, pp. 1607-1615, 2021, doi: 10.1109/TITS.2021.3098461.
[5] Liu Yongle and Gu Yuanli, "Prediction of spatio-temporal characteristics of highway traffic flow based on CNN-BiLSTM," Technology & Economy in Areas of Communications, vol. 24, no. 1, pp. 9-18, 2022, [Online]. Available: http://xsqk.hljit.edu.cn/Jweb_jtkj/CN/Y2022/V24/I1/9.
[6] X. F. Ji and Y. C. Ge, "Holiday highway traffic flow prediction method based on deep learning," Journal of System Simulation, vol. 32, no. 6, pp. 1164-1171, 2020, doi: 10.16182/j.issn1004731x.joss.19-0565.
[7] N. Sadeghi Gargari, R. Panahi, H. Akbari, and A. K. Ng, "Long-term traffic forecast using neural network and seasonal autoregressive integrated moving average: case of a container port," Transportation Research Record, vol. 2676, no. 8, pp. 236-252, 2022, doi: 10.1177/03611981221083311.
[8] X. Xu, X. Jin, D. Xiao, C. Ma, and S. C. Wong, "A hybrid autoregressive fractionally integrated moving average and nonlinear autoregressive neural network model for short-term traffic flow prediction," Journal of

Intelligent Transportation Systems, vol. 27, no. 1, pp. 1-18, 2023, doi: 10.1080/15472450.2021.1977639.
[9] V. Rajalakshmi and S. Ganesh Vaidyanathan, “Hybrid time-series forecasting models for traffic flow prediction,” Promet-Traffic&Transportation, vol. 34, no. 4, pp. 537-549, 2022, doi: 10.7307/ptt.v34i4.3998.
[10] N. Nidhi and D. K. Lobiyal, “Traffic flow prediction using support vector regression,” International Journal of Information Technology, vol. 14, no. 2, pp. 619-626, 2022, doi: 10.1007/s41870-021-00852-2.
[11] H. Xu and C. Jiang, “Deep belief network-based support vector regression method for traffic flow forecasting,” Neural Computing and Applications, vol. 32, no. 7, pp. 2027-2036, 2020, doi: 10.1007/s00521-019-04339-x.
[12] Y. He, P. Huang, W. Hong, Q. Luo, L. Li, L. Tsui K, et al., “In-depth insights into the application of Recurrent Neural Networks (RNNs) in traffic prediction: A comprehensive review,” Algorithms, vol. 17, no. 9, pp. 398, 2024, doi: 10.3390/a17090398.
[13] F. M. Awan, R. Minerva, and N. Crespi, “Improving road traffic forecasting using air pollution and atmospheric data: Experiments based on LSTM recurrent neural networks,” Sensors, vol. 20, no. 13, pp. 3749, 2020, doi: 10.3390/s20133749.
[14] C. Zhang, Z. Ni, and C. Berger, “Spatial-temporal-spectral lstm: A transferable model for pedestrian trajectory prediction,” IEEE Transactions on Intelligent Vehicles, vol. 9, no. 1, pp. 2836-2849, 2023, doi: 10.1109/TIV.2023.3285804.
[15] H. Qiu, Q. Zheng, M. Msahli, G. Memmi, M. Qiu, J. Lu, et al., “Topological graph convolutional network-based urban traffic flow and density prediction,” IEEE transactions on intelligent transportation systems, vol. 22, no. 7, pp. 4560-4569, 2020, doi: 10.1109/TITS.2020.3032882.
[16] N. Hu, D. Zhang, K. Xie, W. Liang, and M. Y. Hsieh, “Graph learning-based spatial-temporal graph convolutional neural networks for traffic forecasting,” Connection Science, vol. 34, no. 1, pp. 429-448, 2022, doi: 10.1080/09540091.2021.2006607.
[17] Z. Chen, J. H. Liu, B. Zhao, S. X. Yuan, and J. Kang, “Taxi demand prediction for multi-factor urban road network based on GCN and TCN,” Control and Decision, vol. 38, no. 4, pp. 1031-1038, 2023, doi: 10.13195/j.kzyjc.2021.0814.
[18] Z. Y. Gu, C. Chen, J. J. Zheng, and D. H. Sun, “Traffic flow prediction based on spatio-temporal graph convolutional recurrent neural network,” Control and Decision, vol. 37, no. 3, pp. 645-653, 2022, doi: 10.13195/j.kzyjc.2020.1445.
[19] S. Liu, H. Chen, X. Y. Chen, and J. J. He, “Dual-branch spatio-temporal graph convolutional neural network for traffic flow prediction,” Information and Control, vol. 52, no. 3, pp. 391-404, 2023, doi: 10.13976/j.cnki.xk.2023.2092.
[20] Z. Ma, Z. Lv, X. Xin, Z. Cheng, F. Xia, J. Li, et al., “Spatio-Temporal Heterogeneous Graph-Based Convolutional Networks for Traffic Flow Forecasting,” Transportation Research Record, vol. 2678, no. 8, pp. 120-133, 2024, doi: 10.1177/03611981231213878.
[21] S. Fukuda, H. Uchida, H. Fujii, and T. Yamada, “Short-term prediction of traffic flow under incident conditions using graph convolutional recurrent neural network and traffic simulation,” IET Intelligent Transport Systems, vol. 14, no. 8, pp. 936-946, 2020, doi: 10.1049/iet-its.2019.0778.
[22] S. Karim, M. Mehmud, Z. Alamgir, and S. Shahid, “Dynamic spatial correlation in graph wavenet for road traffic prediction,” Transportation research record, vol. 2677, no. 7, pp. 90-100, 2023, doi: 10.1177/03611981221151024.
[23] C. Tian and W. K. Chan, “Spatial-temporal attention wavenet: A deep learning framework for traffic prediction considering spatial-temporal dependencies,” IET Intelligent Transport Systems, vol. 15, no. 4, pp. 549-561, 2021, doi: 10.1049/itr2.12044.
[24] F. Hou, Y. Zhang, X. Fu, L. Jiao, and W. Zheng, “The Prediction of Multistep Traffic Flow Based on AST-GCN-LSTM,” Journal of Advanced Transportation, vol. 2021, no. 1, Art. no. 9513170, 2021, doi: 10.1155/2021/9513170.
[25] Y. Wang and C. Jing, “Spatiotemporal graph convolutional network for multi-scale traffic forecasting,” ISPRS International Journal of Geo-Information, vol. 11, no. 2, pp. 102, 2022, doi: 10.3390/ijgi11020102.
[26] L. Peng, X. Liao, T. Li, X. Guo, and X. Wang, “An overview based on the overall architecture of traffic forecasting,” Data Science and Engineering, vol. 9, no. 3, pp. 341-359, 2024, doi: 10.1007/s41019-024-00246-x.
[27] W. T. Wang, X. Q. Wang, L. X. Li, Y. H. Tao, and H. Lin, “A review of the construction and application of spatio-temporal graph neural networks in traffic flow prediction research,” Journal of Computer Engineering & Applications, vol. 60, no. 8, pp. 31, 2024, doi: 10.3778/j.issn.1002-8331.2307-0133.
[28] C. X. Dong, X. Wei, K. P. Zhang, and Y. C. Wang, “Large-scale road network traffic flow prediction based on Graph Transformer,” Industrial Engineering Journal, vol. 26, no. 5, pp. 159-167, 2023, doi: 10.3969/j.issn.1007-7375.2023.05.018.
[29] X. Li, X. Zhang, L. A. Yu, and S. Y. Wang, “Research on short-term passenger flow prediction of scenic spots based on improved Transformer model,” Chinese Journal of Management Science, vol. 33, no. 2, pp. 105-117, 2025, doi: 10.16381/j.cnki.issn1003-207x.2023.1927.
[30] Y. Zhang, Y. Yang, W. Zhou, H. Wang, and X. Ouyang, “Multi-city traffic flow forecasting via multi-task learning,” Applied Intelligence, vol. 51, no. 10, pp. 6895-6913, 2021, doi: 10.1007/s10489-020-02074-8.
[31] Y. Zhang, S. Liu, P. Zhang, and B. Li, “GRU-and Transformer-based periodicity fusion network for traffic forecasting,” Electronics, vol. 12, no. 24, pp. 4988, 2023, doi: 10.3390/electronics12244988.
[32] L. Cai, K. Janowicz, G. Mai, B. Yan, and R. Zhu, “Traffic transformer: Capturing the continuity and periodicity of time series for traffic forecasting,” Transactions in GIS, vol. 24, no. 3, pp. 736-755, 2020, doi: 10.1111/tgis.12644.
[33] Y. A. Pan, F. Li, A. Li, Z. Niu, and Z. Liu, “Urban intersection traffic flow prediction: A physics-guided stepwise framework utilizing spatio-temporal graph neural network algorithms,” Multimodal Transportation, vol. 4, no. 2, Art. no. 100207, 2025, doi: 10.1016/j.multra.2025.100207.
[34] Y. Zhang, M. Lu, and H. Li, “Urban traffic flow forecast based on FastGCRNN,” Journal of Advanced Transportation, vol. 2020, no. 1, Art. no. 8859538, 2020, doi: 10.1155/2020/8859538.
[35] S. Li, L. Yang, and X. H. Zhao, “Urban area short-term traffic congestion prediction algorithm based on deep learning,” Science Technology and Engineering, vol. 23, no. 25, pp. 10866-10878, 2023, doi: 10.3969/j.issn.1671-1815.2023.25.028.
[36] Y. K. Zhang, J. P. Lu, Zhang Sc, X. P. Ji, Y. T. Luo, W. Chen, et al., “Urban traffic accident short-term risk prediction method based on road network structure,” CAAI Transactions on Intelligent Systems, vol. 15, no. 4, pp. 663-671, 2020, doi: 10.11992/tis.201910002.