

基于跨尺度交互注意力与动态图注意力卷积的交通流预测^①



许志强¹, 平 鹏², 徐 扬¹

¹(南京信息工程大学 计算机学院, 南京 210044)

²(南通大学 交通与土木工程学院, 南通 226019)

通信作者: 徐扬, E-mail: yangxu@nuist.edu.cn

摘 要: 准确的交通流预测是保障智能交通系统安全运行与高效调度的关键. 尽管相关研究已取得显著进展, 但由于交通网络复杂的时空关联特性以及交通数据本身的非线性分布, 当前预测模型仍面临以下挑战: 时空动态建模精度不足、时间依赖与空间异质性难以协同表征、时空特征融合有待深化. 针对上述问题, 本文提出一种基于跨尺度交互注意力与动态图注意力卷积的交通流预测模型——DSAGACN. 在时间维度上, 设计跨尺度交互注意力机制, 通过多尺度局部上下文感知与动态交互, 构建“点-趋势-全局”的多层次时间感受野, 从而有效捕捉交通流动态变化; 在空间维度上, 引入动态图注意力卷积模块并结合空间位置嵌入, 利用注意力机制自适应捕捉空间相关性并表征异质性; 在特征融合层面, 提出一种高效融合机制, 以增强时空特征的联合表达并降低计算复杂度. 基于两个真实交通流数据集的实验结果表明, DSAGACN 能充分提取并融合交通数据的时空特征, 在预测精度上优于现有的主流模型.

关键词: 交通流预测; 时空数据; 注意力机制; 图神经网络

引用格式: 许志强, 平鹏, 徐扬. 基于跨尺度交互注意力与动态图注意力卷积的交通流预测. 计算机系统应用. <http://www.c-s-a.org.cn/1003-3254/10266.html>

Traffic Flow Prediction Based on Cross-scale Interactive Attention and Dynamic Graph Attention Convolution

XU Zhi-Qiang¹, PING Peng², XU Yang¹

¹(School of Computer Science, Nanjing University of Information Science & Technology, Nanjing 210044, China)

²(School of Transportation and Civil Engineering, Nantong University, Nantong 226019, China)

Abstract: Accurate traffic flow prediction is crucial for ensuring the safe operation and efficient scheduling of intelligent transportation systems. Despite significant progress, current prediction models still face challenges due to the complex spatio-temporal correlations in traffic networks and the nonlinear nature of traffic data. These challenges include insufficient accuracy in spatio-temporal dynamic modeling, difficulties in jointly representing temporal dependencies and spatial heterogeneity, and the need for deeper integration of spatio-temporal features. To address these issues, this study proposes a traffic flow prediction model, termed as DSAGACN, based on cross-scale interactive attention and dynamic graph attention convolutions. In the temporal dimension, a cross-scale interactive attention mechanism is designed. Through multi-scale local context awareness and dynamic interactions, a multi-level temporal receptive field spanning “point, trend, and global” scales is developed, thus effectively capturing dynamic traffic flow variations. In the spatial dimension, a dynamic graph attention convolution module is introduced and combined with spatial positional embedding.

① 基金项目: 国家自然科学基金 (52442218)

收稿时间: 2026-02-04; 修改时间: 2026-02-27, 2026-03-16; 采用时间: 2026-03-27; csa 在线出版时间: 2026-08-21

Attention mechanisms are leveraged to adaptively capture spatial correlations and represent heterogeneity. At the feature fusion level, an efficient fusion mechanism is proposed to enhance the joint representation of spatio-temporal features while reducing computational complexity. Experimental results on two real-world traffic datasets show that DSAGACN effectively extracts and integrates spatiotemporal features from traffic data, achieving higher prediction accuracy than existing mainstream models.

Key words: traffic flow prediction; spatio-temporal data; attention mechanism; graph neural network (GNN)

随着城市化进程加速与机动车保有量激增, 交通拥堵与安全事故等问题日益凸显, 严重制约城市运行效率与居民出行体验^[1]. 在此背景下, 智能交通系统 (intelligent transportation system, ITS) 成为缓解交通压力、提升交通管理水平核心手段, 而交通流预测作为 ITS 的关键技术支撑, 其预测精度直接决定了交通信号优化、路径规划、应急调度等功能的有效性, 对保障交通系统安全性、稳定性与运行效率具有重要现实意义^[2].

近年来, 学界围绕交通流预测展开了大量研究, 方法从早期的自回归综合移动平均 (autoregressive integrated moving average, ARIMA) 等统计模型, 逐步发展到支持向量机 (support vector machine, SVM)、随机森林等机器学习算法, 再到当前以卷积神经网络 (convolutional neural network, CNN)、长短期记忆 (long short-term memory, LSTM) 网络、图神经网络 (graph neural network, GNN)^[3]、Transformer^[4]等为代表的深度学习模型, 预测性能不断提升. 然而, 交通流数据具有显著的时空复杂性^[5], 现有模型在应用中仍存在多方面技术瓶颈: 一是时空动态建模精度有限, 难以同时刻画多尺度时间规律与复杂空间依赖^[6]; 二是在时间与空间特征的协同表达上仍显不足, 未能充分适配交通网络的空间异质性, 即不同路段或节点在交通流的生成机制、运行特征及演化规律上表现出的显著差异^[7]; 三是时空特征的融合机制设计偏浅, 多局限于特征层面的简单拼接或加权叠加, 尚未建立时空维度间的动态关联与互补约束, 导致信息交互深度不足, 预测结果在精度与稳定性上仍难以满足实际应用需求^[8].

针对上述问题, 本文提出一种基于跨尺度交互注意力与动态图注意力卷积的交通流预测模型 DSAGACN (dynamic scale-aware graph attention convolutional network).

本文的主要工作与贡献如下.

(1) 提出一种跨尺度交互注意力机制, 针对交通流兼具瞬时波动、局部趋势与全局周期的多尺度时间特性, 引入多尺度因果卷积强化对时间片瞬时波动与趋势的局部建模能力, 通过跨尺度动态交互实现不同时间片交通流特征的互补融合, 从而统一捕捉交通流“点-趋势-全局”的多层次动态特征, 适配交通流复杂的时间演化模式.

(2) 设计结合空间位置嵌入与动态图注意力卷积的空间建模模块. 针对路网节点间固有的空间异质性, 在图卷积结构中引入空间位置嵌入以刻画节点的特征属性与区域差异, 同时利用动态图注意力机制自适应调整节点间的关联强度, 以适配交通流拥堵传导、路段关联随时间动态变化的特性, 显著提升模型对路网动态交通状态的适配能力, 增强模型稳定性.

(3) 构建基于门控机制的时空特征融合模块与轻量化解码器, 针对交通流时空强耦合特性, 通过门控单元自适应控制不同尺度时空特征的信息流通比例, 强化深层时空特征的融合表达; 同时采用轻量化解码器设计降低模型复杂度, 在保证交通流量预测精度的前提下提升计算效率, 以满足实际路网预测场景中的实时性要求.

基于真实交通流量数据集的实验结果表明, 本文的 DSAGACN 模型在预测精度上显著优于现有主流模型, 可为智能交通系统的实际应用提供更可靠的技术支持.

1 相关工作

1.1 传统方法和统计模型预测

早期交通流预测方法主要基于数理统计体系, 代表性方法包括历史平均 (historical average, HA) 法、自回归综合移动平均 (ARIMA) 法及向量自回归 (vector autoregressive, VAR) 法. 此类方法均以交通流数据分布已知且保持稳定为核心前提假设, 通过挖掘与拟合

历史数据的统计规律进行预测,但预测精度有限.随着研究的深入推进,学者们在统计方法基础上引入非线性改进策略,形成支持向量机(SVM)、随机森林等为代表的机器学习模型.但此类模型仍局限于“仅聚焦时间维度或单一节点”的分析框架^[9],当交通流数据规模扩大、时空关联性增强时,模型的特征提取能力与计算效率难以满足实际应用需求,进而逐渐被更适用于时空数据建模的深度学习方法取代^[10].

1.2 基于图神经网络的交通流预测

鉴于交通网络天然的图结构特性,可将交叉路口抽象为图节点、路段抽象为边,而路段的连通关系、长度与通行能力等属性可作为边的权重.因此图神经网络(GNN)成为交通流空间相关性建模的主流技术,其核心在于通过节点间的信息传递与聚合^[11],有效捕捉路网中非线性空间关联.早期基于图卷积网络(graph convolutional network, GCN)的交通流预测模型,如STGCN^[12],通过将交通网络构建为静态图,并结合一维卷积捕捉时间动态,初步实现时空特征的融合.该模型系统地将图卷积引入交通流预测,验证了图结构对空间相关性建模的有效性.然而,STGCN采用固定的邻接矩阵,无法动态适应交通流随时间变化的空间关联强度,且对路网空间异质性的适配能力不足.DSTAGNN^[13]通过数据驱动的方式构建动态时空感知图,根据交通流实时数据调整节点邻接关系,提升模型对时空动态变化的适配性;虽然图卷积模型在空间建模上取得显著进展,但多数模型仍存在“时间建模依赖传统卷积或循环单元,难以捕捉多尺度时间规律”的问题,且时空特征融合多为浅层拼接,未能实现深层次的信息交互^[14].

1.3 基于Transformer的交通流预测

Transformer模型凭借自注意力机制的全局依赖捕捉能力,为交通流量预测的多尺度时空相关性建模提供了新范式.其无需依赖循环结构,可并行处理长时序交通数据,能有效捕捉日内高峰、周内周期等多尺度时间规律,契合预测任务对长时序全局特征刻画的需求.但将其直接迁移至交通流量预测时,受数据特性与路网结构约束,场景适配难题尚未完全解决:一是标准自注意力计算复杂度随序列长度平方增长,面对大规模路网与长时序数据时,计算开销大、推理效率低,难以满足实时预测需求^[15];二是交通流具有显著局部关联性与周期性,而标准自注意力过度侧重全局关联,未

能嵌入该类先验信息,易形成无效全局感受野,导致交通流时间相关性建模精度不足,难以适配复杂路网场景^[16].为解决上述问题,学者们从模型轻量化与注意力机制改进两方面展开研究.如Informer^[17],通过优化网络架构设计与改进自注意力机制,同时兼顾降低长序列计算效率与全局时间依赖精准捕捉;CASTNet^[18]通过改进自注意力机制,同时考虑路网局部连接性与全局时间上下文,增强时空特征的协同建模能力;STWave^[19]则设计解纠缠-融合改进框架优化计算效率,同时结合高效谱图注意力,实现多变量交通流数据的时空特征解耦与动态融合;但现有研究仍存在短板,多数方案难以在模型复杂度、计算效率与预测性能三者间实现理想平衡,且对交通流时空耦合规律的适配性仍有提升空间.

综上所述,当前技术演进方向明确聚焦于“更精准的时空动态建模、更高效的时空特征融合、更优的模型复杂度与性能平衡”^[20].基于现有方法在多尺度时间规律捕捉、动态空间相关性及异质性适配、深层次时空融合这3方面关键问题上的不足,本文构建DSAGACN模型,旨在通过针对性的模块改进设计突破技术瓶颈,从而提升交通流预测的性能.

2 模型设计与实现

2.1 问题定义

交通流预测是典型的时间序列预测任务,其核心目标是基于交通网络中各监测点的历史交通流数据(各段时间内通过各节点的车辆数量),对未来特定节点的交通流状态进行精准预测.本文将交通网络定义为一个无向图 $G=(V,E)$,其中 V 是 N 个节点的集合,每个节点表示一个交通监控,例如交通探测器或观测站. E 是路网中边的集合,表示节点之间的连接关系.

交通网络 G 在时间片 t 上包含 N 个节点的观测结果表示为交通信号矩阵 $X_t=(x_{t,1},x_{t,2},\dots,x_{t,N})^T \in R^{N \times C}$,其中 $x_{t,v} \in R^C$ 表示节点 v 在时间片 t 内的交通信号特征, C 是特征维度.交通流预测为基于过去 T_h 个时间片交通信号矩阵序列 $X=(X_{t-T_h+1},X_{t-T_h+2},\dots,X_t) \in R^{N \times C \times T_h}$,预测未来在 T_p 个时间片上的交通信号矩阵序列 $Y=(X_{t+1},X_{t+2},\dots,X_{t+T_p}) \in R^{N \times C \times T_p}$.

交通数据源于人类日常活动,呈现出显著的趋势性、周期性与空间异质性.图1展示了交通流的变化趋势常随时间呈现规律性重复,表明趋势与周期性是交通预测模型设计中需重点考虑的因素.

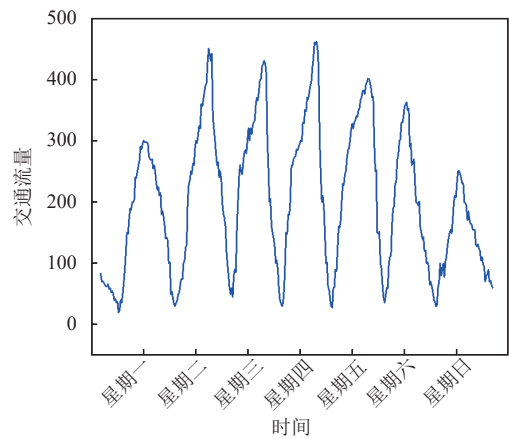


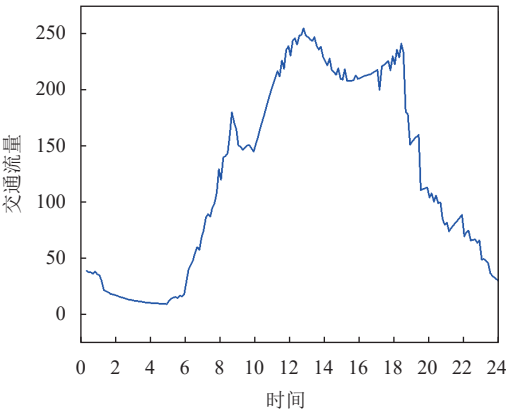
图1 节点D在1星期内的流量变化

如图2展示了选取 PEMS04 数据集中具有典型差异的两个监测节点 A 和 D 的日内流量变化特征,节点 A (图 2(a)) 位于生活街道上,流量主要在 (0, 200) 区间波动,节点 D (图 2(b)) 位于高速公路上,流量集中在 (200, 600) 区间,不同路网节点呈现差异化的流量分布模式,体现了交通流的空间异质性,这一特性同样对精准交通预测至关重要。

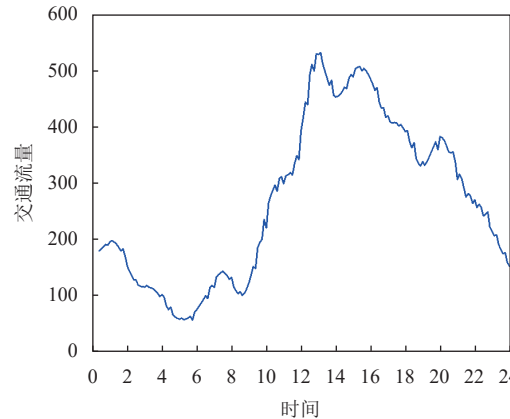
2.2 模型框架

本文提出的 DSAGACN 模型总体框架如图3所示.模型由嵌入层、编码器和解码器构成:嵌入层进行时空位置嵌入;编码器部分由多个级联的时空特征提取层构成,每层包含跨尺度交互注意力模块和动态图注意力卷积模块,前者用于建模交通数据在时间维度上的多尺度动态变化,后者旨在捕捉交通数据在空间维度上的动态关联.解码器包含时空特征融合模块和

全连接输出层.时空特征融合层实现高效的时空特征融合,然后通过全连接层进行映射降维输出预测结果。



(a) 节点A



(b) 节点D

图2 节点 A 和 D 在 1 天内的流量变化

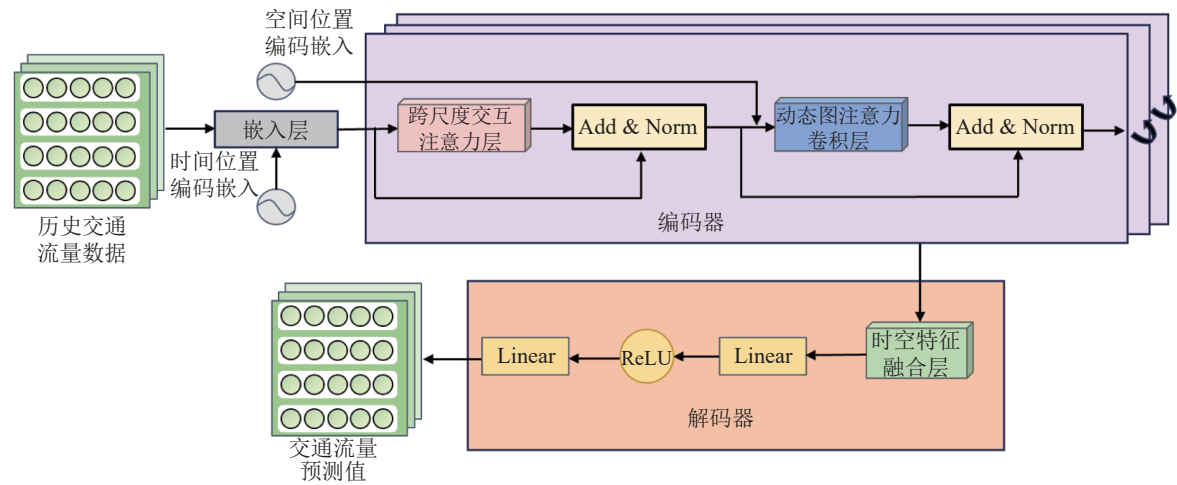


图3 DSAGACN 模型框架

2.3 时空位置嵌入层

在时空预测模型中, 时空维度上邻近的路网节点更具有特征相关性, 故而将时间与空间的位置顺序信息引入模型, 对交通流时序建模任务具有重要意义.

嵌入层首先将原始序列输入 $X \in \mathbb{R}^{N \times C \times T_h}$ 经高维映射得到特征序列 $X \in \mathbb{R}^{N \times d_{\text{model}} \times T_h}$, 随后对其执行时空位置编码操作, 最终将时间位置嵌入矩阵 E_{TP} 与空间位置嵌入矩阵 E_{SP} 叠加至映射后的原始特征序列中, 完成位置信息与原始信息的融合.

时间位置嵌入使用正弦和余弦函数结合嵌入维度进行编码 E_{TP} , 为每个位置 t 的输入元素固定位置嵌入, 每个向量维度 $1 \leq d \leq d_{\text{model}}$, 如式 (1) 所示:

$$\begin{cases} E_{TP}(t, 2d) = \sin\left(\frac{t}{k(d)^{2d/d_{\text{model}}}}\right) \\ E_{TP}(t, 2d+1) = \cos\left(\frac{t}{k(d)^{2d/d_{\text{model}}}}\right) \end{cases} \quad (1)$$

其中, t 表示输入中每个元素的相对索引, d_{model} 表示模型嵌入维度, $k(d) = 2^{(\log_2(d_{\text{model}}))/d_{\text{model}}}$ 表示动态底数.

空间位置嵌入是针对空间异质性进行设计. 基于传统 GCN 实现路网节点的位置嵌入: 首先为每个节点初始化一个独立的嵌入向量, 其维度设为 d_{model} ; 随后借助 GCN 的邻域特征聚合机制, 对各节点的初始嵌入向量进行更新优化, 使更新后的节点嵌入矩阵 E_{SP} 能够有效编码交通路网的固有空间结构信息.

2.4 跨尺度交互注意力模块

针对自注意力机制在交通流预测中难以兼顾局部细节与宏观趋势的问题, 本文提出一种跨尺度交互注意力机制, 结构示意图如图 4.

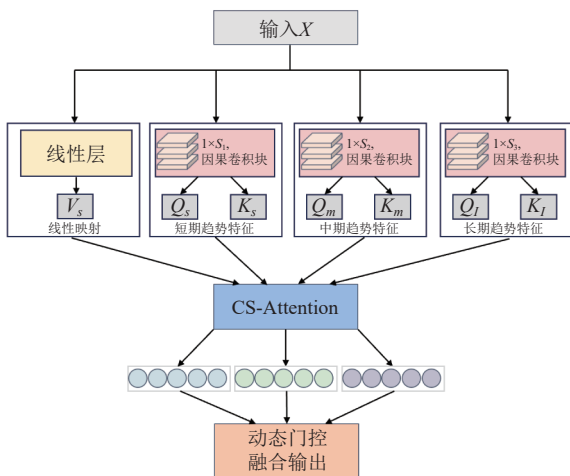


图 4 跨尺度交互注意力

该机制将输入交通时序特征 $X \in \mathbb{R}^{N \times d_{\text{model}} \times T_h}$ 按时间尺度划分为“点-趋势-全局”这 3 个层次, 分别表征单步状态、局部趋势与长程演化特征.

首先通过多尺度因果卷积提取不同感受野下的特征表示, 如式 (2) 所示:

$$H_s = \text{Conv}_{k_s}^{\text{causal}}(X), s \in \{1, 2, \dots, S\} \quad (2)$$

其中, k_s 为第 s 个尺度的卷积核大小. 较小的卷积核聚焦局部波动, 较大的卷积核捕获长期依赖, 从而完成从局部到全局的时间特征重构. 为后续注意力计算提供多尺度的上下文特征输入.

各尺度上, 对卷积特征 H_s 进行查询 Q_s 与键 K_s 的映射, 以表征趋势信息; 对原始输入 X 进行值 V_s 映射, 保留单一时间原始点级信息, 三者采用独立的可学习投影矩阵并映射至相同维度空间, 有效缓解特征来源不同可能产生的尺度偏移和分布差异, 如式 (3) 所示:

$$Q_s = H_s W_Q^{(s)}, K_s = H_s W_K^{(s)}, V_s = X W_V^{(s)} \quad (3)$$

注意力权重计算引入时间因果约束与局部上下文保留机制, 通过“点-趋势”关联建模, 如式 (4), 在局部约束下分配权重, 增强注意力与交通流实际演化逻辑的一致性.

$$A_{cs} = \text{Softmax} \left(\frac{\exp(Q_s K_s^T)}{\sum_{i'=1}^T \exp(Q_s K_{s:,i'}^T)} \right) V_s \quad (4)$$

其中, $\exp(Q_s K_s^T)$ 表示对每个时间步的内积进行指数化, $\sum_{i'=1}^T \exp(Q_s K_{s:,i'}^T)$ 表示同一查询位置的所有键内积进行归一化求和, 进一步削弱尺度影响, 确保不同来源的特征能够在归一化求和中有有效对齐. $\text{Softmax}(\cdot)$ 实现时间维度的归一化.

为融合多尺度信息, 本文设计动态门控融合策略. 首先计算各尺度时间平均特征 H_{sv} , 与全局特征 H_{glob} 拼接后生成尺度自适应权重, 如式 (5)–(7).

$$H_{sv} = \frac{1}{T} \sum_{t=1}^T H_s(t, :) \quad (5)$$

$$H_{\text{glob}} = \text{Pooling} \left(\sum_{s=1}^S \alpha_s H_{sv} \right) \quad (6)$$

$$g_s = \sigma(W_g [H_{sv}; H_{\text{glob}}]) \quad (7)$$

其中, α_s 为可学习的尺度权重系数, $\text{Pooling}(\cdot)$ 表示全

局聚合操作. $\sigma(\cdot)$ 为 Sigmoid 函数, 输出 $g_s \in (0, 1)$ 表示尺度 s 的动态权重系数.

最终通过加权聚合实现“点-趋势-全局”多层次特征的统一建模, 如式 (8) 所示:

$$Z = \sum_{s=1}^S g_s \odot A_{cs} \quad (8)$$

该策略能依据交通状态 (如拥堵、畅通、波动) 动态调整各尺度贡献, 增强模型对不同交通态势的适应性与特征表达能力, 实现更稳健的特征表达.

本机制的贡献在于从单尺度静态建模转向多尺度动态交互建模, 通过因果卷积与门控融合补全了常规注意力在局部依赖建模上的不足, 实现了多层次特征的协同感知, 提升了模型对复杂时序结构的刻画能力与预测性能.

2.5 动态图注意力卷积模块

针对交通路网复杂拓扑结构, 传统 GCN 虽能通过节点特征聚合学习空间表征, 但其固有的时间静态假设限制了建模效果, 即固定的邻接矩阵无法反映交通网络随时间的动态关联变化, 也难以刻画不同路段在功能、通行能力等方面的空间异质性.

DSTAGNN 等现有动态图模型虽突破了传统 GCN 的静态邻接限制, 但在构建动态邻接矩阵时存在动态图无拓扑约束 (路网原始连接结构)、空间异质性感知不足等问题. 为此, 本文提出一种融合空间位置嵌入与动态图注意力卷积的空间建模模块 DGACN, 该方法通过静态结构约束保证路网拓扑合理性; 动态相关矩阵捕捉交通流时变关联; 空间位置嵌入显式刻画异质性, 实现动态空间建模的三重优化. 构建了“静态结构约束下的动态演化”机制, 实现了动态性与异质性的协同建模.

首先, 通过空间位置嵌入增强对路网异质性的感知. 在动态图卷积前融合节点位置编码, 使模型显式考虑地理结构与功能差异, 提升对空间异质特征的刻画能力.

DGACN 保留传统 GCN 静态结构约束, A 为静态结构约束矩阵, 编码了路网的基础拓扑连接关系, 定义如式 (9) 所示, 其中 $\tilde{A}$ 为在 A 的基础上添加自环后的矩阵, $\tilde{D}$ 为 $\tilde{A}$ 对应的度矩阵.

$$A = \tilde{D}^{-\frac{1}{2}} \tilde{A} \tilde{D}^{-\frac{1}{2}} \quad (9)$$

最后, 引入动态相关性建模, 通过时变空间相关性矩阵 S_t 实时调整节点间连接强度, 实现图结构的动态

重构. 时变空间相关性矩阵由给定节点表示 $Z_t^{(l-1)}$ (跨尺度交互注意力层的输出) 作为输入计算生成, 如式 (10) 所示, 其元素 S_{ij} 表示节点 i 与 j 在当前时刻的相关性强度, 从而捕捉随时间演化的空间依赖.

$$S_t = \text{Softmax} \left(\frac{Z_t^{(l-1)} Z_t^{(l-1)T}}{\sqrt{d_{\text{model}}}} \right) \in R^{N \times N} \quad (10)$$

基于上述内容, 结合式 (9) 和式 (10), DGACN 的整体计算过程如式 (11) 所示:

$$X_t^{(l)} = \text{DGACN}(Z_t^{(l-1)}) = \sigma((A \odot S_t) Z_t^{(l-1)} W^{(l)}) \quad (11)$$

其中, A 为静态结构约束矩阵, S_t 为动态相关性矩阵, 二者先通过 Hadamard 积实现结构与动态权重的融合; $W^{(l)}$ 为投影矩阵, $\sigma(\cdot)$ 为非线性激活函数. 最终的输出结果为 $X^{(l)} = (X_{t-T_h+1}^{(l)}, X_{t-T_h+2}^{(l)}, \dots, X_t^{(l)})$, 表征了路网在不同时间步的空间相关特征, 为后续交通流预测提供了自适应、动态且位置敏感的空间表征.

2.6 时空特征融合模块

交通流量预测中, 不同路网节点的流量变化往往呈现出差异化的时间模式. 为更有效地融合并校正交通流量数据中的多层次时空特征, 本文提出了一种时空特征融合门控 tanh 单元 (spatio-temporal fusion gated tanh unit, STF-GTU). 该模块通过引入多尺度门控卷积结构, 对不同时间感受野下的特征进行融合与筛选, 从而提升模型对复杂动态交通模式的预测精度.

传统递归单元 (如 GRU^[21]) 虽能有效捕捉时间依赖性, 但因其序列计算特性, 存在计算开销大、并行性受限的不足. GTU 使用门控卷积替代递归计算, 提升了运行效率, 但其单尺度门控机制难以在多层次时空特征融合过程中保持信息的一致性.

针对上述问题, 本文将门控机制扩展至多尺度融合框架, 结构示意图如图 5.

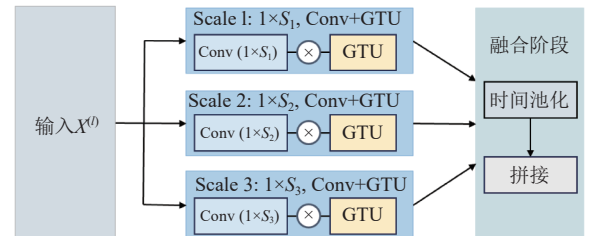


图5 时空特征融合层

STF-GTU 不再将多尺度卷积仅视为时间特征的逐层提取器, 而是将其作为时空特征融合的核心载体,

通过多个尺度的门控路径对特征进行自适应加权与调节,使模型能够在短期与长期时间依赖之间实现动态平衡,在维持较低计算负担的同时增强时空信息的整合能力.

具体而言,时空特征融合模块输入为编码器输出的张量 $X^{(l)} \in R^{N \times T_h \times d_{\text{model}}}$,其中 N 为路网节点数, T_h 为历史时间步长, d_{model} 为特征嵌入维度.

传统 GTU 通过卷积核 $\Gamma \in R^{1 \times S \times C \times 2C}$ 在时间维度执

$$\begin{aligned} X_{\text{out}}^{(l)} &= \text{STF-GTU}(X^{(l)}) \\ &= \text{ReLU}\left(\text{Concat}\left(\text{Pooling}\left(\Gamma_1 *_{\tau} X^{(l)}\right), \text{Pooling}\left(\Gamma_2 *_{\tau} X^{(l)}\right), \text{Pooling}\left(\Gamma_3 *_{\tau} X^{(l)}\right)\right) + X^{(l)}\right) \end{aligned} \quad (13)$$

其中, Γ_1 、 Γ_2 、 Γ_3 分别为感受野大小为 $1 \times S_1$ 、 $1 \times S_2$ 、 $1 \times S_3$ 的卷积核. $\text{Concat}(\cdot)$ 表示沿时间维度拼接不同尺度的门控输出. 为保持输出时间长度与输入 T_h 一致,对拼接后的特征进行窗口大小为 W 的池化下采样,如式 (14) 所示:

$$(3T_h - (S_1 + S_2 + S_3 - 3)) / W = T_h \quad (14)$$

融合特征经 ReLU 激活后,通过残差连接得到 $X_{\text{out}}^{(l)} \in R^{N \times T_h \times d_{\text{model}}}$,再经全连接层降维映射,最终输出未来 T_p 个时间片的交通流量预测序列 $Y \in R^{N \times C \times T_p}$.

STF-GTU 借助多尺度门控融合的设计,在时空特征融合过程中自适应地保留不同尺度特征间的关联信息,实现特征层面上的时空动态对齐与信息重构,从而为后续预测任务提供结构化、层次化的时空特征表示.

3 实验分析

3.1 数据集和实验环境

为了验证本文提出的 DSAGACN 模型的性能,选取真实高速公路场景下的 PEMS04 和 PEMS07 数据集作为实验数据来源.数据集是由加利福尼亚州交通部性能测量系统 (PEMS) 每隔 30 s 实时收集,每 5 min 为间隔进行数据汇总处理,并视为一个时间步.不同数据集的采集地点与时间不同,涵盖多个区域与时间段的高频交通流量数据,被广泛应用于交通流预测任务.数据集具体信息如表 1 所示.

表 1 数据集特点描述

数据集	节点	边	时间步数	时间范围
PEMS04	307	340	16 992	2018/9/1–2018/11/30
PEMS07	883	866	28 224	2017/5/1–2017/8/31

本文实验使用 Python 语言与 PyTorch 深度学习框架完成模型搭建;实验将数据集按照时间以 6:2:2 的比

行门控运算,输出表示如式 (12) 所示:

$$\Gamma *_{\tau} X^{(l)} = \phi(E) \odot \sigma(F) \quad (12)$$

其中, $*_{\tau}$ 为门控卷积算子, $\phi(\cdot)$ 为 tanh 函数, $\sigma(\cdot)$ 为 Sigmoid 函数, E 和 F 分别为卷积结果的前、后半部分.该结构通过门控机制调节信息流,实现对时序动态的非线性调制.

本文在此基础上,依据预测所需的时间跨度将 GTU 扩展为多尺度门控融合架构,定义 STF-GTU 如式 (13) 所示:

例划分为训练集、验证集和测试集.模型输入为过去 1 h (12 个时间步) 的交通流量序列,用于推断未来 1 h (12 个时间步) 的交通状态.在训练设置方面,PEMS04 使用批次大小为 16,PEMS07 较大使用批次大小为 4,将学习率固定为 0.001,训练轮数分别设定为 100 次与 200 次,并选用 Adam 优化算法进行参数更新.

3.2 基线模型

本次实验选择与以下 8 个基线方法进行对比.

(1) STSGCN^[22]: 通过精心设计的时空同步建模机制有效捕捉复杂的局部时空相关性.

(2) STGODE^[23]: 提出了一个新的图卷积的连续表示增加图卷积的深度,扩大空间感受野,捕捉更深层次的时空依赖.

(3) DSTAGNN^[13]: 采用数据驱动的动态时空感知图代替传统图卷积中预定义的静态图.

(4) STG-NCDE^[24]: 提出了时空图神经控制微分方程 (STG-NCDE) 方法.设计了两个控制微分方程,一个用于时间处理,另一个用于空间处理.组合成一个单独的框架.

(5) STWave^[19]: 设计了一个新的时空网络来适配解纠缠-融合框架.同时设计了一个全新的高效谱图注意力捕捉多变量间动态变化的空间信息.

(6) SSTBAN^[25]: 设计了包含空间瓶颈注意力 (SBA) 和时间瓶颈注意力 (TBA) 的 STBA 模块.

(7) CASTNet^[18]: 使用新型注意力机制与图卷积,使网络能同时考虑局部连接性和全局上下文的时间特征,又增强了空间特征提取能力.

(8) IEDSFAN^[26]: 设计了多图融合门控机制 (MGFG),通过控制静态图和动态图的信息权重实现两者融合,动态捕捉隐藏的时空相关性.

3.3 评估指标

为了评估 DSAGACN 模型性能, 使用平均绝对误差 (*MAE*)、均方根误差 (*RMSE*) 和平均绝对百分比误差 (*MAPE*) 这 3 个作为评估指标, 来计算现实流量值与预测值之间的误差. 它们从不同维度衡量预测误差, 联合运用能更系统地反映模型的预测精度与稳健性能. *MAE* 直观呈现了预测误差的算术平均水平, *RMSE* 对较大幅度的预测偏差赋予了更高权重, *MAPE* 则以百分比形式刻画了误差的相对幅度. 这 3 类指标的数值越小, 表明模型的预测效果越优, 其定义如式 (15)–(17):

$$MAE = \frac{1}{N} \sum_{i=1}^N |Y_i - \hat{Y}_i| \quad (15)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (Y_i - \hat{Y}_i)^2} \quad (16)$$

$$MAPE = \frac{1}{N} \sum_{i=1}^N \frac{|Y_i - \hat{Y}_i|}{Y_i} \times 100\% \quad (17)$$

其中, N 是测试数据集的大小, Y_i 和 $\hat{Y}_i$ 分别是实际值和预测值. 对于这 3 个指标, 值越小表示交通流预测结果越好.

3.4 对比实验

为系统评估本文提出的 DSAGACN 模型在交通流预测任务中的性能, 将其与 8 种主流基线模型在 PEMS04 和 PEMS07 两个公开数据集上进行了量化对比, 综合评价指标包括 *MAE*、*RMSE* 与 *MAPE*. 详细结果见表 2, 后续表格中各指标最优值已用粗体标出. 总体而言, DSAGACN 在所有指标上均显著优于基线模型, 体现了其在时空数据分析中的有效性与稳健性.

表 2 对比实验结果

模型	PEMS04			PEMS07		
	<i>MAE</i>	<i>RMSE</i>	<i>MAPE</i> (%)	<i>MAE</i>	<i>RMSE</i>	<i>MAPE</i> (%)
STSGCN	21.19	33.65	13.90	24.26	39.03	10.21
STGODE	20.84	32.82	13.77	22.99	37.54	10.14
DSTAGNN	19.30	31.46	12.70	21.42	34.51	9.01
STG-NCDE	19.21	31.09	12.76	20.53	33.84	8.80
STWave	21.20	34.47	14.32	20.33	34.03	8.58
SSTBAN	18.89	31.21	12.62	20.45	37.55	9.31
CAST-Net	20.53	32.62	14.63	23.28	38.81	10.17
IEDSFAN	19.24	31.37	12.68	20.63	33.86	8.89
DSAGACN	18.73	30.80	12.47	20.17	33.89	8.45

具体来看, 在 PEMS04 数据集上, DSAGACN 各项评价指标全面领先, 尤其在反映极端误差抑制能力的

RMSE 指标上, 较表现最优的基线模型有明显降低. 在路网结构更复杂、监测点密度更高的 PEMS07 数据集上, 模型在 *MAE* 与 *MAPE* 上依然保持领先, *RMSE* 虽较其在 PEMS04 上的结果略有波动, 但仍居于所有模型前列. 这表明 DSAGACN 所采用的特征提取与融合机制能够适应不同数据分布, 具备较强的泛化能力.

值得注意的是, PEMS07 数据集具有更高的特征冗余与更不规则的拓扑结构, 增加了模型在关键信息筛选与时空依赖建模方面的难度. 这可能是 DSAGACN 在该数据集上 *RMSE* 出现小幅波动的主要原因. 尽管如此, 模型整体仍表现出良好的性能.

从建模机制角度分析, DSAGACN 的有效性得益于其针对现有方法的不足提出 3 方面改进: 在时间维度上, 通过分层时序特征提取结构, 同时捕捉瞬时特征、短时依赖与长期趋势; 在空间维度上, 引入异质性感知机制, 动态建模节点间差异化的关联强度; 在时空融合层面, 设计跨模态交互模块, 实现深层特征整合, 避免单维度表征的局限. 实验结果表明, 该设计能够更精准地刻画交通流的动态演化规律, 从而显著提升预测精度.

图 6 展示了 DSAGACN 在 PEMS04 与 PEMS07 数据集上的多步预测结果, DSAGACN 对未来 12 个时间步的交通流量展现出稳定且可解释的预测性能, 具体表现分析如下.

在 PEMS04 数据集上, 随着预测步长从第 1 个时间步延伸至第 12 个时间步, 误差虽随步长增加呈累积趋势, 但整体上升较为平缓, 表明模型在较长预测范围内仍能保持结构稳定. 在 PEMS07 数据集上, 模型各步误差同样呈现类似的渐进上升特征. 但误差增长趋势仍保持一致, 未出现急剧发散或波动, 反映出模型在捕捉长期时序依赖方面的有效性与鲁棒性.

3.5 消融实验

为了进一步评估 DSAGACN 中不同组件的影响, 本文设计了 3 个 DSAGACN 变体版本进行了消融实验, 并在数据集 PEMS04 上分析了实验结果. 虽然更改了组件, 所有变体模型都具有与 DSAGACN 相同的设置, 除了下述差异.

(1) w/o SE&GACN: 它移除了空间位置嵌入和图注意力卷积, 改为传统图卷积代替, 以研究动态调整空间相关强度而不是仅基于静态拓扑关系的作用.

(2) w/o CSIA: 它移除了跨尺度交互注意力模块,

改为基础注意力代替,以研究在预测时同时考虑“点-趋势-全局”对序列信息建模的重要性.

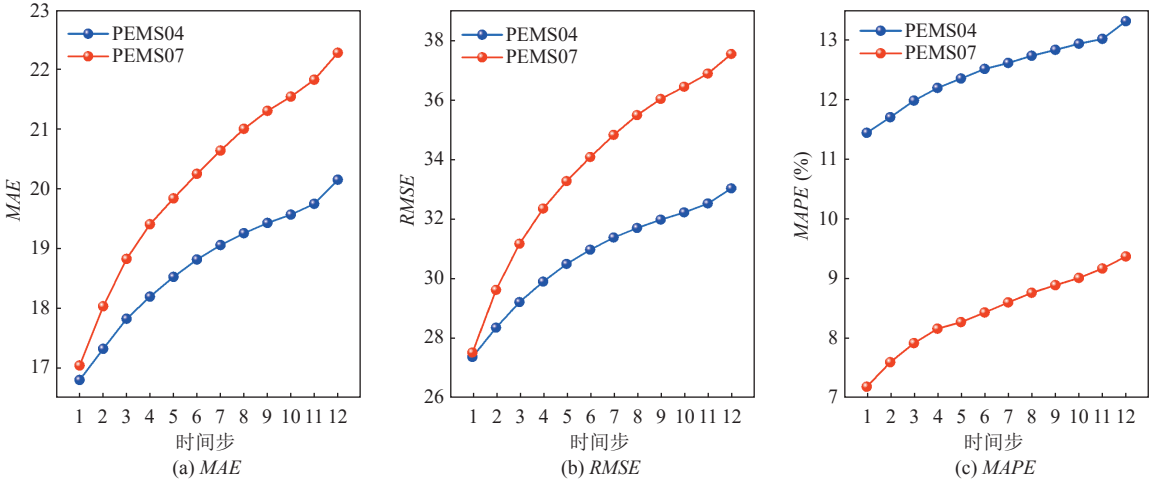


图6 两个数据集上的预测误差

(3) w/o TSF: 它移除了时空特征融合模块, 改为仅用全连接层融合预测, 以研究强化时空特征表达和深层次融合对预测的重要性.

表3和图7展示了DSAGACN及其3种变体在PEMS04数据集上的性能对比.

表3 PEMS04数据集上的消融结果			
模型	MAE	RMSE	MAPE (%)
w/o SE&GACN	19.23	31.28	12.86
w/o CSIA	19.03	31.02	12.63
w/o TSF	18.91	30.91	12.62
DSAGACN	18.73	30.80	12.47

结果显示, 各变体模型的预测性能均显著低于完

整DSAGACN, 且不同组件的移除对模型影响程度存在明显差异: 时间建模能力的缺失对性能影响最大, 表明跨尺度时序特征提取是提升预测精度的关键; 动态空间建模次之, 说明适应节点间异质关联的结构优于静态图卷积; 时空特征融合模块的影响相对较小, 但仍对最终性能具有稳定贡献.

上述消融实验结果表明, DSAGACN中跨尺度时序建模、动态空间感知与时空特征融合这3大组件具有协同增强作用, 共同支撑了模型在交通流预测任务中的优越性能, 验证了本文所提完整架构设计的合理性与有效性.

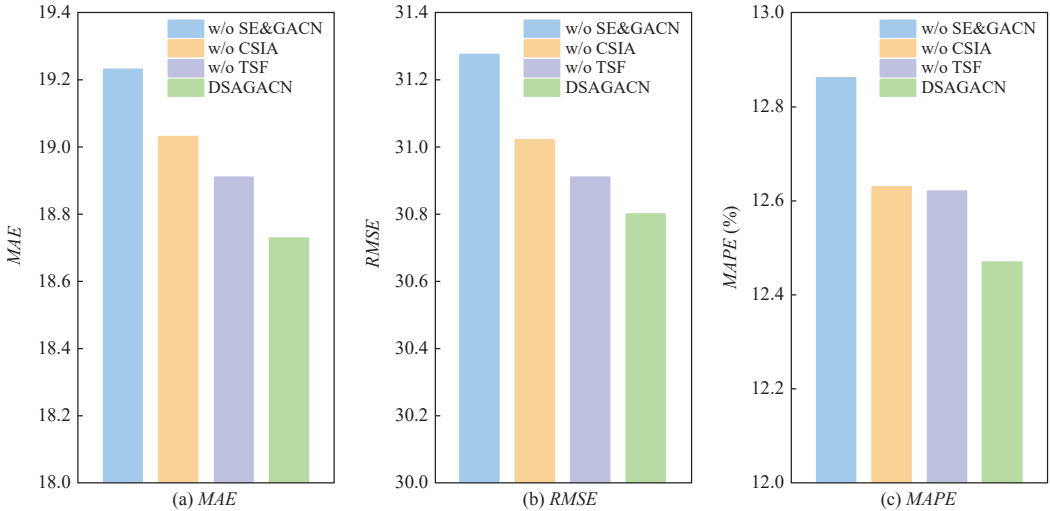


图7 PEMS04数据集上的消融结果

3.6 计算时间

为系统评估模型的计算效率, 本文在 PEMS04 数据集上测试了 DSAGACN 与多个基线模型的训练与推理时间, 测试均在本文实验环境下进行, 结果如表 4 所示。指标包括每轮训练时间 (s) 和单次推理时间 (s)。

表 4 PEMS04 数据集上不同模型的计算时间 (s)

模型	每轮训练时间	推理时间
STSGCN	57.55	14.39
DSTAGNN	150.37	10.21
STWave	70.02	6.74
SSTBAN	137.23	7.53
DSAGACN	49.69	5.82

从表 4 可见, 不同模型的计算开销差异显著。其中, DSTAGNN 因结构复杂、并行度低, 训练时间最高; STSGCN 在推理阶段产生较多额外开销; SSTBAN 体现出时空瓶颈注意力在平衡计算开销方面的优势, STWave 体现出设计的解纠缠-融合框架在计算效率方面表现较好。

相比之下, DSAGACN 在训练与推理阶段均展现出最优的计算效率: 训练时间为 49.69 s, 推理时间为 5.82 s, 均为对比模型中最低。具体而言, 其训练时间较 DSTAGNN 降低约 66.9%, 推理时间降低 43.0%; 较 STSGCN 训练时间降低 13.7%, 推理时间降低 59.5%; 即便与效率较高的 STWave 相比, 训练时间仍降低 29.0%, 推理时间降低 13.6%。这一结果得益于 DSAGACN 在解码器结构上的优化设计, 使其在维持高预测精度的同时, 显著降低了计算负担。DSAGACN 不仅在预测精度上优于现有基准模型, 在计算效率方面也具备明显优势, 体现了其在实际部署中的可行性与实用性。

4 结论与展望

本文提出一种融合跨尺度交互注意力与动态图注意力卷积的交通流预测模型 DSAGACN, 以应对时空动态建模精度有限、时空协同表征不足及特征融合深度欠缺等挑战。该模型构建了多维度协同的时空特征提取与融合框架。在 PEMS04、PEMS07 等真实交通数据集上的实验结果表明, DSAGACN 在 MAE、RMSE 和 MAPE 这 3 项关键评价指标上均显著优于 STSGCN、DSTAGNN 等 8 种主流基线模型, 验证了其在时空特征建模方面的优越性。消融实验进一步表明, 跨尺度时间建模、动态空间建模以及时空特征融合这 3 大核心模块的协同作用是模型性能提升的关键, 印证了模型

架构设计的合理性。效率测试显示, 该模型在保持预测精度的同时, 具备更优的训练与推理速度, 有利于实际系统部署。然而, 在面对节点规模急剧增长的大规模城市路网时, 模型仍面临计算复杂度升高、关键时空特征捕获能力下降等挑战。未来工作将着重优化空间建模模块的并行计算效率, 并探索适用于超大规模路网的轻量化动态图卷积结构。

参考文献

- 1 Kashyap AA, Raviraj S, Devarakonda A, *et al.* Traffic flow prediction models—A review of deep learning techniques. *Cogent Engineering*, 2022, 9(1): 2010510. [doi: [10.1080/23311916.2021.2010510](https://doi.org/10.1080/23311916.2021.2010510)]
- 2 Gan L, Chu WB, Li GF, *et al.* Large models for intelligent transportation systems and autonomous vehicles: A survey. *Advanced Engineering Informatics*, 2024, 62: 102786. [doi: [10.1016/j.aei.2024.102786](https://doi.org/10.1016/j.aei.2024.102786)]
- 3 Liu YJ, Liu QX, Zhang JW, *et al.* Multivariate time-series forecasting with temporal polynomial graph neural networks. *Proceedings of the 36th International Conference on Neural Information Processing Systems*. New Orleans: Curran Associates Inc., 2022. 1411.
- 4 Vaswani A, Shazeer N, Parmar N, *et al.* Attention is all you need. *Proceedings of the 31st International Conference on Neural Information Processing Systems*. Long Beach: Curran Associates Inc., 2017. 6000–6010.
- 5 王雨松, 吴向东, 尤晨欣, 等. 基于 DWT-GCN 的短时交通流预测. *计算机系统应用*, 2022, 31(9): 306–312. [doi: [10.15888/j.cnki.csa.008682](https://doi.org/10.15888/j.cnki.csa.008682)]
- 6 赵文竹, 袁冠, 张艳梅, 等. 多视角融合的时空动态 GCN 城市交通流量预测. *软件学报*, 2024, 35(4): 1751–1773. [doi: [10.13328/j.cnki.jos.007018](https://doi.org/10.13328/j.cnki.jos.007018)]
- 7 Han K, Wang YH, Chen HT, *et al.* A survey on vision Transformer. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2023, 45(1): 87–110. [doi: [10.1109/TPAMI.2022.3152247](https://doi.org/10.1109/TPAMI.2022.3152247)]
- 8 邹慧琪, 史彬泽, 宋凌云, 等. 基于图神经网络的复杂时空数据挖掘方法综述. *软件学报*, 2025, 36(4): 1811–1843. [doi: [10.13328/j.cnki.jos.007275](https://doi.org/10.13328/j.cnki.jos.007275)]
- 9 Ji JH, Wang JY, Huang C, *et al.* Spatio-temporal self-supervised learning for traffic flow prediction. *Proceedings of the 2023 AAAI Conference on Artificial Intelligence*. Washington: AAAI, 2023. 4356–4364.
- 10 Ye BL, Zhang MJ, Li LX, *et al.* A survey of traffic flow prediction methods based on long short-term memory

- networks. IEEE Intelligent Transportation Systems Magazine, 2024, 16(5): 87–112. [doi: [10.1109/MITS.2024.3400679](https://doi.org/10.1109/MITS.2024.3400679)]
- 11 曹威, 王兴, 邹复民, 金彪, 等. 融合多种时空自注意力机制的 Transformer 交通流预测模型. 计算机系统应用, 2024, 33(4): 82–92. [doi: [10.15888/j.cnki.csa.009451](https://doi.org/10.15888/j.cnki.csa.009451)]
- 12 Yu B, Yin HT, Zhu ZX. Spatio-temporal graph convolutional networks: A deep learning framework for traffic forecasting. Proceedings of the 27th International Joint Conference on Artificial Intelligence. Stockholm: IJCAI.org, 2018. 3634–3640. [doi: [10.24963/ijcai.2018/505](https://doi.org/10.24963/ijcai.2018/505)]
- 13 Lan SY, Ma YT, Huang WK, *et al.* DSTAGNN: Dynamic spatial-temporal aware graph neural network for traffic flow forecasting. Proceedings of the 39th International Conference on Machine Learning. Baltimore: PMLR, 2022. 11906–11917.
- 14 Liu AY, Zhang YY. Spatial-temporal dynamic graph convolutional network with interactive learning for traffic forecasting. IEEE Transactions on Intelligent Transportation Systems, 2024, 25(7): 7645–7660. [doi: [10.1109/TITS.2024.3362145](https://doi.org/10.1109/TITS.2024.3362145)]
- 15 Reza S, Ferreira MC, Machado JJM, *et al.* A multi-head attention-based Transformer model for traffic flow forecasting with a comparative analysis to recurrent neural networks. Expert Systems with Applications, 2022, 202: 117275. [doi: [10.1016/j.eswa.2022.117275](https://doi.org/10.1016/j.eswa.2022.117275)]
- 16 Ata KIM, Hassan MK, Ismaeel AG, *et al.* A multi-layer CNN-GRUSKIP model based on Transformer for spatial-temporal traffic flow prediction. Ain Shams Engineering Journal, 2024, 15(12): 103045. [doi: [10.1016/j.asej.2024.103045](https://doi.org/10.1016/j.asej.2024.103045)]
- 17 Zhou HY, Zhang SH, Peng JQ, *et al.* Informer: Beyond efficient Transformer for long sequence time-series forecasting. Proceedings of the 35th AAAI Conference on Artificial Intelligence. AAAI, 2021. 11106–11115.
- 18 Chen ZX, Zhao SJ, Zeng J, *et al.* CASTNet: Convolution augmented graph sampling Transformer network for traffic flow forecasting. Proceedings of the 27th International Conference on Computer Supported Cooperative Work in Design (CSCWD). Tianjin: IEEE, 2024. 1292–1297.
- 19 Fang YC, Qin YJ, Luo HY, *et al.* When spatio-temporal meet wavelets: Disentangled traffic forecasting via efficient spectral graph attention networks. Proceedings of the 39th International Conference on Data Engineering. Anaheim: IEEE, 2023. 517–529.
- 20 Liu YF, Wu CZ, Wen JH, *et al.* A grey convolutional neural network model for traffic flow prediction under traffic accidents. Neurocomputing, 2022, 500: 761–775. [doi: [10.1016/j.neucom.2022.05.072](https://doi.org/10.1016/j.neucom.2022.05.072)]
- 21 Hu N, Zhang DF, Xie K, *et al.* Multi-range bidirectional mask graph convolution based GRU networks for traffic prediction. Journal of Systems Architecture, 2022, 133: 102775. [doi: [10.1016/j.sysarc.2022.102775](https://doi.org/10.1016/j.sysarc.2022.102775)]
- 22 Song C, Lin YF, Guo SN, *et al.* Spatial-temporal synchronous graph convolutional networks: A new framework for spatial-temporal network data forecasting. Proceedings of the 34th AAAI Conference on Artificial Intelligence. New York: AAAI, 2020. 914–921.
- 23 Fang Z, Long QQ, Song GJ, *et al.* Spatial-temporal graph ODE networks for traffic flow forecasting. Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining. Singapore: ACM, 2021. 364–373.
- 24 Choi J, Choi H, Hwang J, *et al.* Graph neural controlled differential equations for traffic forecasting. Proceedings of the 36th AAAI Conference on Artificial Intelligence. AAAI, 2022. 6367–6374.
- 25 Guo SN, Lin YF, Gong LT, *et al.* Self-supervised spatial-temporal bottleneck attentive network for efficient long-term traffic forecasting. Proceedings of the 39th International Conference on Data Engineering. Anaheim: IEEE, 2023. 1585–1596.
- 26 Yu LF, Wang ZL, Yang WX, *et al.* IEDSFAN: Information enhancement and dynamic-static fusion attention network for traffic flow forecasting. Complex & Intelligent Systems, 2025, 11(1): 28. [doi: [10.1007/s40747-024-01663-1](https://doi.org/10.1007/s40747-024-01663-1)]
- (校对责编: 张重毅)