陈艳, 徐坤财, 张硕, 等. 基于强化学习和曼巴相关图卷积网络的交通流量预测[J]. 智能计算机与应用,2026,16(5):73-81. DOI:10.20169/j.issn.2095-2163.25121801

基于强化学习和曼巴相关图卷积网络的交通流量预测

陈 艳, 徐坤财, 张 硕, 张金贵, 张 宁
(贵阳信息科技学院 智能工程学院, 贵阳 550025)

摘 要: 准确的交通流预测是智能交通系统的核心任务之一,能为交通管理部门提供有效的决策支持,缓解城市交通拥堵、优化路网运行效率。然而,传统预测方法在处理复杂的时空非线性关系和动态交通模式时存在局限性。为此,提出了一种结合强化学习的 Mamba-相关性图卷积网络(Mamba-HiLoGCN)方法。首先,将 Mamba 与 CorGCN 相结合,以灵活探索时空依赖性并准确捕捉不同路段交通流的相互作用。其次,通过将相关注意力机制与 HiLo 注意力机制相结合,自适应地调整对高频和低频特征的关注,提高了捕捉短期波动和长期趋势的能力。最后,DDPG 通过与环境的交互来优化预测策略,并自适应地调整邻接矩阵,从而确保该方法在复杂多变的交通场景中保持高精度。在多个真实交通数据集上的实验表明,Mamba-HiLoGCN 在各项评价指标上均显著优于 18 种基线模型,包括 HA、ARIMA、LSTM、GRU、CNN、DCRNN、GMAN、T-GCN、STGCN、ASTGCN、STSGCN、AGCRN、STGODE、STCGCN、Bi-STAT、ASTMGODE、ASTTN 和 DDGformer,验证了其在提升预测精度与鲁棒性方面的有效性。

关键词: 智能交通系统; 交通流预测; Mamba-CorGCN 模型; HiLo 关联注意力机制; DDPG 算法
中图分类号: TP391.4 **文献标志码**: A **文章编号**: 2095-2163(2026)05-0073-09

Traffic flow prediction based on reinforcement learning and Mamba-Correlation Graph Convolutional Networks

CHEN Yan, XU Kuncai, ZHANG Shuo, ZHANG Jingui, ZHANG Ning
(College of Intelligent Engineering, Guiyang Institute of Information Science and Technology, Guiyang 550025, China)

Abstract: Accurate traffic flow forecasting is one of the core tasks of intelligent transportation systems, providing effective decision support for traffic management authorities to alleviate urban congestion and optimize road network efficiency. However, traditional forecasting methods have limitations when handling complex spatio-temporal nonlinear relationships and dynamic traffic patterns. To this end, we propose a Mamba-Correlation Graph Convolutional Network (Mamba-HiLoGCN) method incorporating reinforcement learning. First, Mamba is combined with CorGCN to flexibly explore spatio-temporal dependencies and accurately capture interactions between traffic flows on different road segments. Second, by integrating a correlation attention mechanism with a HiLo attention mechanism, it adaptively adjusts focus on high-frequency and low-frequency features, enhancing the ability to capture both short-term fluctuations and long-term trends. Finally, DDPG optimizes prediction strategies through interaction with the environment and adaptively adjusts the adjacency matrix, ensuring high accuracy in complex and dynamic traffic scenarios. Experiments conducted on multiple real-world traffic datasets demonstrate that Mamba-HiLoGCN significantly outperforms 18 baseline models across all evaluation metrics. These include HA, ARIMA, LSTM, GRU, CNN, DCRNN, GMAN, T-GCN, STGCN, ASTGCN, STSGCN, AGCRN, STGODE, STCGCN, Bi-STAT, ASTMGODE, ASTTN and DDGformer, validating its effectiveness in enhancing prediction accuracy and robustness.

Key words: intelligent transportation systems; traffic flow prediction; Mamba-CorGCN model; HiLo Correlation Attention; DDPG algorithm

0 引 言

交通流预测是智能交通系统的关键技术[1],其通过分析历史数据和当前交通状况来预测未来的交通趋势[2]。传统的交通流预测方法(回归模型[3-4]、时间序列分析[5-6]和物理模型[7])在一定程度上能

基金项目: 贵州省自然科学研究项目(青年项目)(黔教技[2024]279 号);贵阳信息科技学院校级课题(GKY2025-XJ-46)。
作者简介: 徐坤财(1997—),男,硕士,讲师。主要研究方向:类脑计算,医学图像处理; 张 硕(2003—),男,本科生。主要研究方向:智慧城市,人工智能,大数据; 张金贵(2003—),男,本科生。主要研究方向:智慧城市,人工智能,大数据; 张 宁(1993—),男, 硕士,讲师。主要研究方向:机器视觉,灾害监测。
通信作者: 陈 艳(1997—),女,硕士,助教。主要研究方向:智慧城市,人工智能,大数据。Email:chenyan0909.edu@gmail.com。
收稿日期: 2025-12-18

够捕捉交通流的变化规律。但往往难以处理非线性关系,对复杂交通模式的适应性有限,并且无法有效整合来自多源的数据。随着深度学习和深度强化学习的广泛应用[8],特别是在时空特征建模[9-10]、多源数据融合[11]和传感器相关性分析[12]方面的进步,为交通流量预测提供了新的研究思路。此外,Mamba[13]能够优化预测模型参数、筛选关键特征,并与深度学习模型结合,从而进一步提高预测的准确性和效率,适用于广泛的交通场景。

传统交通流量预测方法主要包括回归模型[14]、时间序列分析方法[15]、物理模型[16]、机器学习方法[17]和统计学习方法[18]。Rajeh 等[19]采用多元线性回归单元(MLRU)方法,将时间序列预测转化为多元线性回归问题,并通过将注意力层与 MLRU 输出相结合,直接基于历史流量预测未来交通流量。Salamanis 等[20]提出"稀疏特征回归"方法,通过从交通数据中提取精心筛选的特征并施加稀疏性,实现快速准确的大规模交通预测。Miao 等[21]采用误差季节性自回归积分滑动平均(ESARIMA)方法,该方法结合自回归积分滑动平均(ARIMA)模型、经验模态分解(EMD)和奇异值分解(SVD)技术进行交通流量预测。

基于深度学习的交通流预测模型能够通过自动学习历史数据中的时空特征,有效捕捉复杂的非线性关系和动态变化,从而显著提升预测精度和实时性能。这类模型包含卷积神经网络(CNN)[22]、循环神经网络(RNN)、LSTM[23]、GCN[24]、混合模型(CNN+ LSTM 、GCN + LSTM)[25]以及生成对抗网络(GAN)。这些模型在多源数据融合和长期预测方面展现出显著优势。Cini 等[26]设计了一种深度集成模型(DEM),通过自适应组合各模型的预测性能,将 CNN、LSTM 和 GRU 作为基础学习器。Liu 等[27]选择了一种结合 CNN、BiGRU 和加法注意力机制(AAM)的 CNN-BiGRU-AAM 模型来提升预测性能。Lin 等[28]研究了一种基于异构数据融合的交通流预测模型,通过结合 GCN 和 GRU 来捕捉节点的空间特征和时间依赖性。

为此,本文提出了 Mamba-HiLoGCN 模块,该模块有效提高了交通流预测的准确性和实时响应能力。Mamba - HiLoGCN 由 Mamba - CorGCN、HiLo CorAttention 和 DDPG 组成。Mamba-CorGCN 模块能够捕捉传感器节点之间的时空相关性,并优化交通流的空间依赖性建模。HiLo CorAttention 模块结合了相关注意力机制和 HiLo 注意力机制,动态调整对高频和低频特征的关注,更好地适应交通流的波动和长期趋势。DDPG 模块自动调整 CorGCN 的邻接矩阵,以提高交通流预测的准确性和实时性。

基于上述模块设计与实验验证。本工作的贡献总结如下:

(1) Mamba-HiLoGCN 由 Mamba-CorGCN、HiLo CorAttention 和 DDPG 组成,每个模块在处理交通流数据中都发挥着独特的作用,共同提高了预测的准确性和实时响应能力。

(2) Mamba-CorGCN 将 Mamba 与 CorGCN 相结合,以有效建模传感器节点之间的时空依赖性。Mamba 结构通过灵活的结构调整增强了复杂时空特征的建模能力,而 CorGCN 则捕捉传感器节点之间的动态相关性,并准确反映不同路段间交通流的相互作用,从而提高了预测精度。

(3) HiLo CorAttention 模块包括相关注意力机制和 HiLo 注意力机制,通过动态调整注意力权重来自适应地增强对高频和低频特征的关注。高频特征可以捕捉短期波动,而低频特征可以学习长期趋势。

(4) DDPG 根据交通流的变化自动调整 CorGCN 的邻接矩阵,使该方法能够灵活应对复杂的交通场景,并在不同时间和条件下始终保持较高的预测精度。

1 Mamba-HiLoGCN 算法

Mamba - HiLoGCN 由 Mamba - CorGCN、HiLo CorAttention 和 DDPG 组成,如图 1 所示。Mamba-CorGCN 通过动态捕捉节点间的相关性来提高时空预测精度。HiLo CorAttention 通过调整注意力权重,自适应地增强对高频和低频特征的关注,高频特征捕捉短期波动,而低频特征学习长期趋势。DDPG 根据交通流的动态变化自适应地调整 CorGCN 的邻接矩阵,并确保在不同时间步和不同条件下保持高精度的预测结果。总之,Mamba-HiLoGCN 在处理复杂动态交通流场景时具 有出色的适应性和高效率,并能提供高精度的预测结果。

1.1 Spatiotemporal Correlation Information (时空相联信息)

最大信息系数(MIC)方法可以有效地衡量传感器序列之间的相关性。因此,提出了两种基于 MIC 的时空表征,包括空间相关信息(SC)和时间相关信息(TC)。SC 有效地从空间数据中捕捉复杂的依赖结构。TC 捕捉时间序列数据中的非线性和时变关系。通过这两种时空表征,有效地捕捉了交通流预测中可能被忽略的细微时空关系,并提高了理解和预测复杂交通模式的能力。MIC 方法的计算如下:

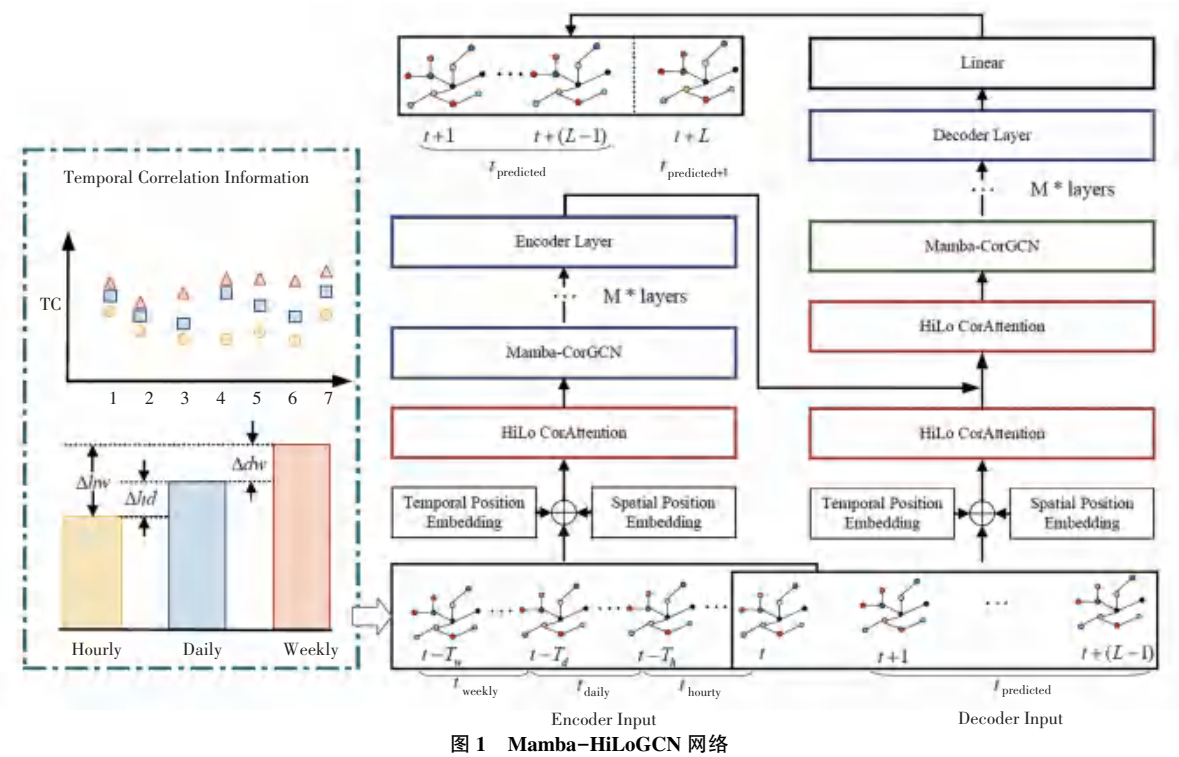


图 1　Mamba-HiLoGCN 网络

Fig. 1　Mamba-HiLoGCN Network

$$\mathrm{MIC}(X,Y) = \max_{|G|\leqslant B(n)} \frac{I(X,Y)}{\log\min(|G_X|,|G_Y|)} \quad (1)$$

其中，X、Y 为随机变量；$I(X,Y)$ 为 X 和 Y 的互信息；G 为网格划分；$|G_X|$、$|G_Y|$ 为 X 和 Y 方向的网格数；$B(n)$ 为网格数上限。

空间相关信息(SC)通过最大信息系数(MIC)量化不同传感器序列间的空间依赖性。传统邻接矩阵仅能表示节点间是否存在连接(0 或 1)，而 SC 能够捕捉传感器之间复杂动态的非线性空间关系，其值域为[0,1]。具体而言，给定两个传感器序列 X 和 Y，其空间相关性定义为 $\mathrm{SC}(X,Y) = \mathrm{MIC}(X,Y)$。较高的 SC 值表明传感器间具有强空间关联性和相似交通模式，较低值则反映空间特征存在显著差异。

交通流数据表现出显著的时间依赖性，如，交通流量的高峰时段变化往往遵循更一致的模式，而节假日和工作日的交通特征可能差异显著。因此，引入 TC 来探索和捕捉时间序列数据中不同时间步之间的相似模式。TC 定义如下：

$$\mathrm{TC}(X)_i^c = \frac{1}{T}\sum_{t=1}^{T} \mathrm{MIC}(X_{i,t:t+\tau}^c, \tilde{X}_{i,t:t+\tau}^c) \quad (2)$$

其中，$\mathrm{TC}(X)_i^c$ 表示传感器 i 在属性 c 下的时间相关性，描述了传感器 i 的数据序列在时间维度上相关性的强度；T 是时间序列的长度；t 是当前时间步，表示时间序列中的第 i 个时刻，i 为传感器的编号。

1.2　Mamba-CorGCN

为提高时空特征提取能力，如图 2 所示，提出了 Mamba-CorGCN 模块。该模块通过多尺度设计，深入探索交通网络中的局部和全局相关性，并充分利用交通传感器数据的时空信息，为交通预测提供更准确和鲁棒的解决方案。Mamba-CorGCN 模块通过融合 Mamba 的高效特征提取能力与 CorGCN 的空间依赖性建模优势，实现了对复杂交通流模式的多尺度捕捉。CorGCN 突破传统图卷积仅考虑邻接关系的限制，引入 $\mathrm{SC}(X,Y)$ 作为动态权重因子，并结合强化学习动态优化邻接矩阵，有效表征节点间非线性的时空交互关系。Mamba 组件采用选择性状态空间模型，通过非线性变换和自适应特征选择机制，显著增强了模型对高维时空特征的提取能力。该双模块架构既能深入挖掘交通网络的局部细节特征，又能捕捉全局时空依赖关系，为交通流预测提供了更精确的时空表征。

1.3　HiLo CorAttention

传统的交通流预测模型通常对高频和低频特征的处理较为简单，要么使用固定权重进行特征融合，要么忽略不同频率特征对预测准确性的不同贡献。HiLo CorAttention 模块通过动态调整注意力权重，自适应地关注不同时间尺度上的高频和低频特征，有效提高了对短期波动和长期趋势的关注。下面是对高频注意力(Hi-Fi)和低频注意力(Lo-Fi)的具体描述。

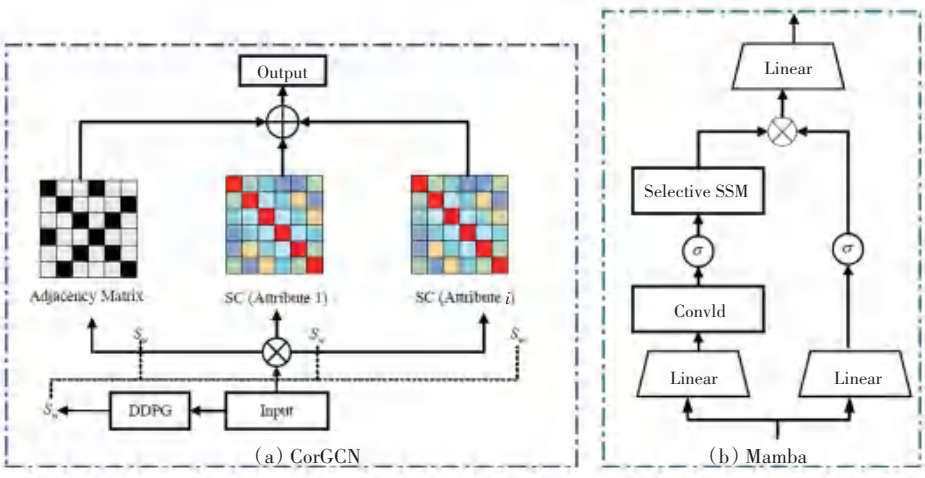


图 2 网络图 CorGCN 和 Mamba

Fig. 2 Network diagrams of CorGCN and Mamba

(1)高频注意力 (Hi-Fi)。Hi-Fi 通过多头注意力机制捕捉输入特征中的高频信息。具体来说,首先计算查询 (q)、键 (k) 和值 (v)。因为传统的注意力机制可能受限于单一特征表示,难以充分表达复杂的语义信息。因此,将 k 和 q 分别建模为内容 (k_i^C,q_i^C) 和位置 (k_i^R,q_i^R) 信息,通过拼接进行组合。接着,引入 SC 来建模和优化 k 的相关性,从而更准确地描述高频特征的相关性。然后,基于优化后的 k,计算 k 和 v 之间的相似度,并通过 Softmax 函数获得注意力权重。最后,将这些注意力权重应用于 v 的加权求和,以生成加权输出,如公式(3)~(10)所示。

(2)低频注意力 (Lo-Fi)。Lo-Fi 与 Hi-Fi 类似,都使用多头注意力机制来捕捉输入特征中的重要信息,并通过计算 q 和 k 之间的相似度来生成注意力权重。两者不同之处在于,k 和 v 的低频部分通过池化操作从输入特征中进行下采样得到,而高频部分则直接使用原始输入特征计算。Lo-Fi 的计算公式如下:

$$\begin{cases}\text{Hi-Fi}: c_i^{KV} = \boldsymbol{W}^{DKV} x_i \\ \text{Lo-Fi}: c_i^{KV} = \boldsymbol{W}^{DKV}\text{Pool}(x_i)\end{cases} \tag{3}$$

$$\begin{cases}\text{Hi-Fi}: c_i^{Q} = \boldsymbol{W}^{DQ} x_i \\ \text{Lo-Fi}: c_i^{Q} = \boldsymbol{W}^{DQ}\text{Pool}(x_i)\end{cases} \tag{4}$$

$$\begin{cases} k_i^C = \boldsymbol{W}^{UK} c_i^{KV} \\ k_i^R = \text{RoPE}(\boldsymbol{W}^{KR} x_i) \\ k_{i,i} = [k_{i,i}^C; k_i^R] \\ k_{i,i} = \mathcal{F}(\text{SC}, K) = \mathcal{F}\dfrac{1}{C}\displaystyle\sum_{c=1}^{C}\frac{\exp(\text{SC})}{\displaystyle\sum_{u=1}^{n}\exp(\text{SC})} k_{i,i} \end{cases} \tag{5}$$

$$\begin{cases} q_t^C = \boldsymbol{W}^{UQ} c_t^Q, \\ q_t^R = \text{RoPE}(\boldsymbol{W}^{QR} c_t^Q) \\ q_{t,i} = [q_{t,i}^C; q_{t,i}^R] \end{cases} \tag{6}$$

$$v_t^C = \boldsymbol{W}^{UV} c_t^{KV} \tag{7}$$

$$o_{t,i} = \sum_{j=1}^{t} \text{Softmax}_j\left(\frac{q_{t,i}^{\mathrm{T}} k_{j,i}}{\sqrt{d_h + d_{R\,h}}}\right) v_{j,i}^C \tag{8}$$

$$\begin{cases}\text{Hi-Fi}: y_h = \boldsymbol{W}_h^O[o_{t,1}; o_{t,2}; \cdots; o_{t,n_h}] \\ \text{Lo-Fi}: y_l = \boldsymbol{W}_l^O[o_{t,1}; o_{t,2}; \cdots; o_{t,n_h}]\end{cases} \tag{9}$$

$$Y = \text{Concat}(y_h, y_l) \tag{10}$$

其中,$\boldsymbol{W}^{DKV}$、$\boldsymbol{W}^{DQ}$、$\boldsymbol{W}^{UK}$、$\boldsymbol{W}^{KR}$、$\boldsymbol{W}^{UQ}$、$\boldsymbol{W}^{QR}$ 和 $\boldsymbol{W}^{UV}$ 分别是 Hi-Fi 和 Lo-Fi 的查询、键和值的权重矩阵;$\boldsymbol{W}_h^O$ 和 $\boldsymbol{W}_l^O$ 分别表示 Hi-Fi 和 Lo-Fi 输出映射的权重矩阵;RoPE 是一种在注意力机制中引入位置信息的位置编码方法,通过矩阵旋转将位置信息嵌入到输入向量中。

Hi-Fi 和 Lo-Fi 分别针对高频和低频信息设计,通过不同的特征处理方法(包括直接使用原始特征和池化降维)实现多尺度特征提取。两种模型共享相同的计算框架,但在输入特征处理上有所不同。通过结合 Hi-Fi 和 Lo-Fi,模型能够同时捕捉细节和全局依赖关系,从而在复杂任务中表现出更强的性能。

1.4 DDPG

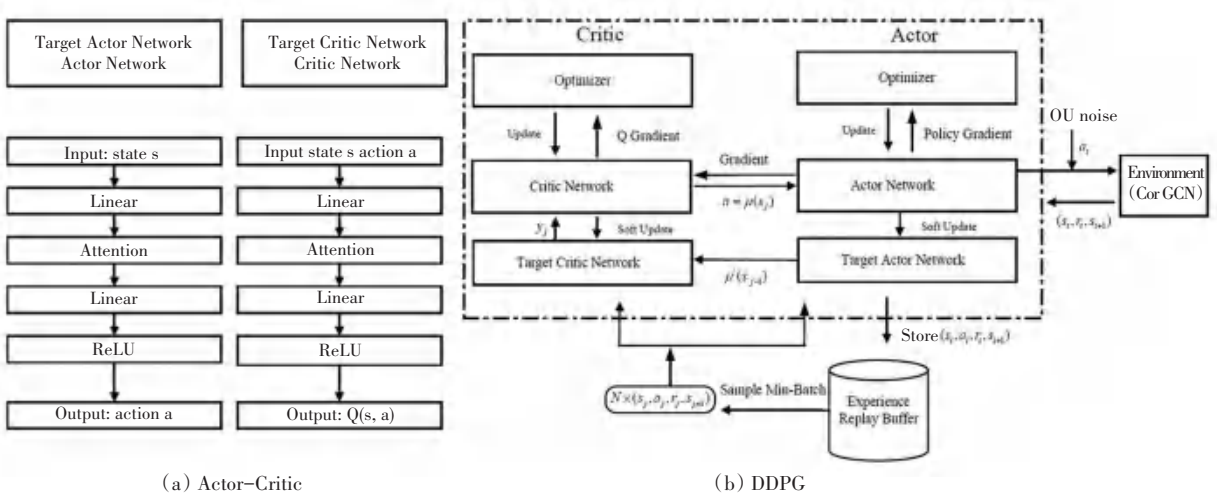
DDPG 算法通过强化学习框架实现交通流预测策略的自主优化,如图 3 所示。该算法采用 Actor-Critic 架构,其中 Actor 网络负责生成动态调整的邻接矩阵作为动作,Critic 网络则评估这些动作的长期价值。通过与环境交互收集经验数据,利用经验回放机制打破数据相关性,并通过添加探索噪声平衡

探索与利用。这种端到端的训练方式使模型能够自适应不同交通场景，实时调整预测策略，显著提升了复杂动态环境下的预测准确性和鲁棒性。Action、State 和 Reward function 的定义描述如下：

（1）Action：动作定义为 CorGCN 中的图邻接矩阵，表示交通节点之间的连接关系。动作由演员网络生成，用于调整交通流预测策略，优化图邻接矩阵，提高了模型适应不同交通情况的能力。

（2）State：状态是交通流预测模型的输出，通常包含交通状况信息，如交通流量、速度和拥堵水平。状态信息对于调整预测策略至关重要，确保智能体能够实时响应交通流的变化。

（3）Reward function：奖励通过计算预测交通流与实际交通流之间的误差来定义。预测误差越小，奖励越高，表明模型的预测性能越好。



（a）Actor-Critic　　（b）DDPG

图 3　网络图 Actor-Critic 和 DDPG

Fig. 3　Network diagrams of Actor-Critic and DDPG

2　实验

2.1　数据集

为全面验证 Mamba-HiLoGCN 的性能，如表 1 所示，选取高速公路和地铁两类典型交通数据集进行广泛实验。高速公路交通数据集源自加州交通流监测系统（PEMS），选取了 4 个区域（PEMS03、PEMS04、PEMS07 和 PEMS08），这些区域包含的核心交通流量、速度及占有率数据充分体现了高速公路交通复杂的时空特征。地铁客流数据集来自杭州地铁系统（HZME），包含各站点的乘客进出站流量，反映了不同时段地铁交通的动态特性。本文使用 Z-Score 对交通数据进行归一化。每个数据集按 6∶2∶2 的比例分为训练集、验证集和测试集。

表 1　数据集统计信息

Table 1　Statistics of datasets

数据类型集	数据集名称	节点数量	时间范围	时间粒度/min
高速公路交通数据集	PeMS03	358	2018/9/1-2018/11/30	5
	PeMS04	307	2018/1/1-2018/2/28	5
	PeMS07	883	2017/5/1-2017/8/31	5
	PeMS08	170	2016/7/1-2016/8/31	5
地铁客流数据集	HZME（in）	80	2019/1/1-2019/1/26	5
	HZME（out）	80	2019/1/1-2019/1/26	5

2.2　实验设置

Mamba-HiLoGCN 方法基于 PyTorch GPU 框架开发。在训练过程中，选用均方误差（MSE）作为损失函数来衡量预测值与真实值之间的差异，同时采用 Adam 优化器进行模型参数优化。经实验验证，将 Adam 的学习率设置为 0.001，既能保证训练效率又能有效提升模型性能。如表 2 所示，为高效训练模型并实现交通流预测，针对不同数据集（PEMS03、PEMS04、PEMS07、PEMS08、HZME(in)和 HZME(out)）设置了相应的超参数。

表 2 不同数据集的超参数

Table 2 Hyperparameters for different datasets

数据集	编码器层数	解码器层数	卷积核大小	注意力头数	批次大小	Top-U
PEMS03	4	4	3	8	8	5
PEMS04	4	4	3	8	8	5
PEMS07	3	3	3	8	2	5
PEMS08	4	4	3	8	8	5
HZME(in)	4	4	3	8	16	5
HZME(out)	4	4	3	4	16	5

采用平均绝对误差(MAE)、均方根误差(RMSE)和平均绝对百分比误差(MAPE)三项指标来评估 Mamba-HiLoGCN 的性能。MAE、RMSE 和 MAPE 的计算公式如下:

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n} | F_i - F_i^* | \tag{11}$$

$$\mathrm{MAPE} = \frac{1}{n}\sum_{i=1}^{n} \left| \frac{F_i^* - F_i}{F_i^*} \right| \times 100\% \tag{12}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n} (F_i - F_i^*)^2} \tag{13}$$

其中,F_i 表示预测的交通流量,F_i^* 表示真实的交通流量。

2.3 实验结果与分析

2.3.1 实验结果

如表 3 所示,与 8 种基准方法相比,Mamba-HiLoGCN 在不同场景下具有较高的预测精度。不同数据集上的性能差异主要源于数据特性和模型架构的匹配程度,PEMS 数据具有规律性强、路网结构清晰的特点,因此擅长捕捉时空依赖的方法表现良好。而 HZME 则呈现显著的潮汐效应和进出站流量不对称性,其中 HZME(out)因突发性更强导致所有模型的预测误差显著更高,例如,Mamba-HiLoGCN 在 HZME(out) 上的 RMSE 相对于其他方法显著下降。最终,Mamba-HiLoGCN 凭借其状态空间架构对长序列和非平稳时序的卓越建模能力,在所有数据集上均取得最佳或次优性能,显示出其强大的泛化性和对复杂时空动态的捕捉能力。例如,在 PeMS03 数据集上,Mamba-HiLoGCN 的 MAE 值为 14.20,Mamba-HiLoGCN 与 Bi-STAT、ASTM-GODE、ASTTN 和 DDGformer 相比,误差分别降低了约 7.73%、11.19%、5.27%和 2.47%。这些结果进一步验证了 Mamba-HiLoGCN 在大规模交通流预测任务中的优异性能。

表 3 六个数据集上不同模型的性能比较

Table 3 Performance comparison of different models across six datasets

数据集	指标	STSGCN[28]	AGCRN[29]	STGDDE[30]	STCGCN[31]	Bi-STAT[32]	ASTMGODE[33]	ASTTN[34]	DDGformer[35]	Mamba-HiLoGCN
PeMS03	MAE	17.49	15.97	16.50	16.67	15.39	15.99	14.99	14.56	14.20
	MAPE/%	29.21	28.11	27.85	27.72	25.87	24.37	23.97	23.33	23.01
	RMSE	16.78	15.23	16.69	16.31	15.48	14.58	13.48	13.68	13.14
PeMS04	MAE	21.20	19.83	20.84	20.79	18.31	19.80	18.89	18.09	17.22
	MAPE/%	33.61	32.28	32.82	33.63	30.83	30.53	30.22	30.02	29.61
	RMSE	13.84	12.98	13.78	13.73	12.33	13.09	12.23	11.43	10.34
PeMS07	MAE	24.22	21.22	22.59	23.50	20.25	21.33	20.35	18.97	18.13
	MAPE/%	39.04	35.19	37.58	37.29	33.24	32.03	30.33	30.13	30.10
	RMSE	10.22	9.16	10.15	9.72	8.56	8.78	7.98	7.78	6.92
PeMS08	MAE	17.16	16.05	16.81	17.85	13.46	13.06	12.16	12.04	11.31
	MAPE/%	26.86	25.23	25.98	26.48	23.13	23.13	22.03	21.93	21.65
	RMSE	10.92	10.19	10.62	10.64	9.55	9.76	8.66	8.16	7.92
HZME(in)	MAE	12.86	11.84	11.34	11.02	11.31	10.51	10.11	9.91	9.44
	MAPE/%	23.40	24.40	22.01	18.91	17.21	16.23	16.01	15.71	15.01
	RMSE	28.10	30.10	40.44	22.40	21.30	20.52	20.20	19.98	19.32
HZME(out)	MAE	18.72	19.45	19.43	17.67	17.11	16.78	16.08	15.89	15.02
	MAPE/%	33.19	33.81	33.57	30.88	30.27	28.87	28.17	27.27	26.01
	RMSE	76.72	88.79	88.34	70.65	70.34	59.94	59.14	57.24	50.12

2.3.2　消融实验

Mamba－HiLoGCN 包含 Mamba－CorGCN、HiLo CorAttention 和 DDPG 三个核心模块，如图 4、图 5 所示，为验证上述三个模块有效性，本文利用 HZME（in）和 HZME（out）两个数据集，对各组件性能进行评估。结果表明，移除 Mamba－CorGCN 会导致预测曲线显著偏离真实值，说明 Mamba－CorGCN 作为时空特征提取的核心，在捕捉路段间动态关系中起主导作用。移除 DDPG 后，模型在交通波动剧烈时段（如高峰时段）性能明显下降，表明其通过动态调整图结构增强了模型对交通状态变化的适应能力。移除 HiLo CorAttention 模块 后则使预测曲线在平稳段仍接近真实值，但在快速变化段出现平滑过度现象，说明该机制通过多尺度特征加权有效增强了对短期波动和长期趋势的区分能力。综上所述，Mamba－HiLoGCN 通过整合三个模块的优势，在不同数据集和场景中展现出更高的预测精度。

如图 6、图 7 所示，Mamba－HiLoGCN 最接近真实交通流，在整体趋势和局部特征方面均具有出色的预测精度。特别是在复杂波动和高峰时段，Mamba－HiLoGCN 通过 HiLo CorAttention、Mamba－CorGCN 和 DDPG 的动态优化机制，实现了对交通流快速变化的精准捕捉。其预测曲线与真实曲线高度吻合。

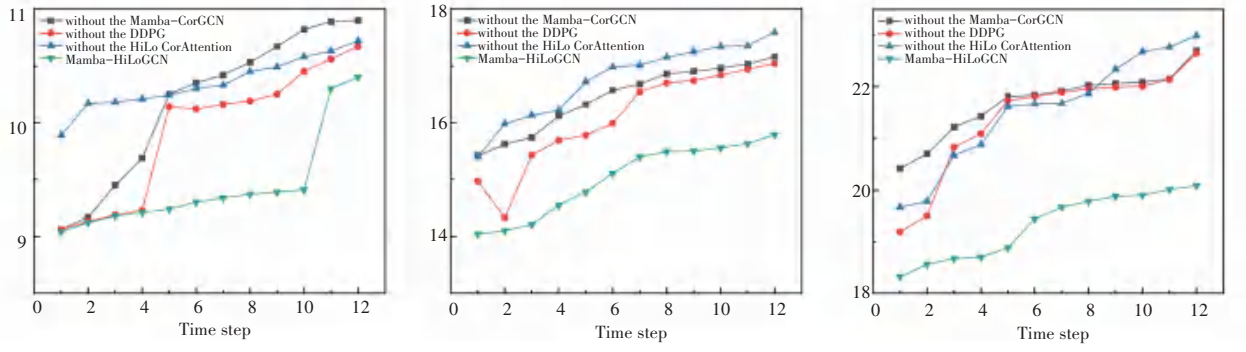


图 4　HZME（in）评价指标结果

Fig. 4　HZME（in）evaluation metric results

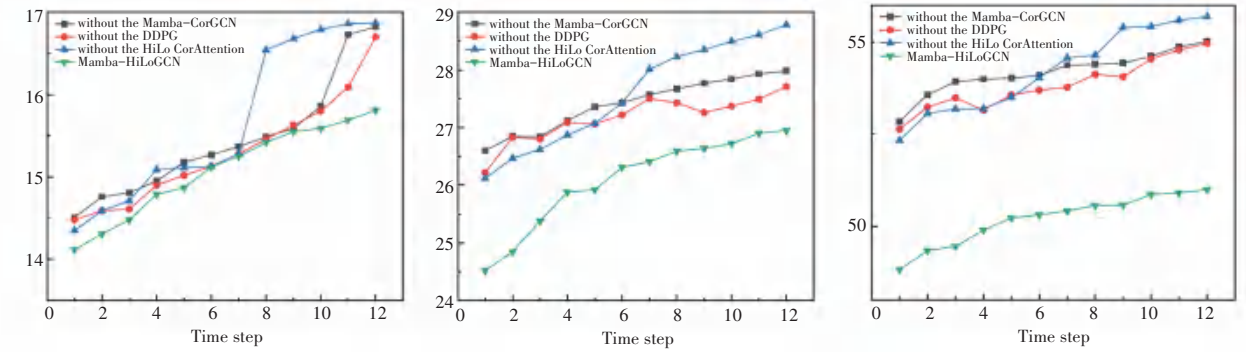


图 5　HZME（out）评价指标结果

Fig. 5　HZME（out）evaluation metric results

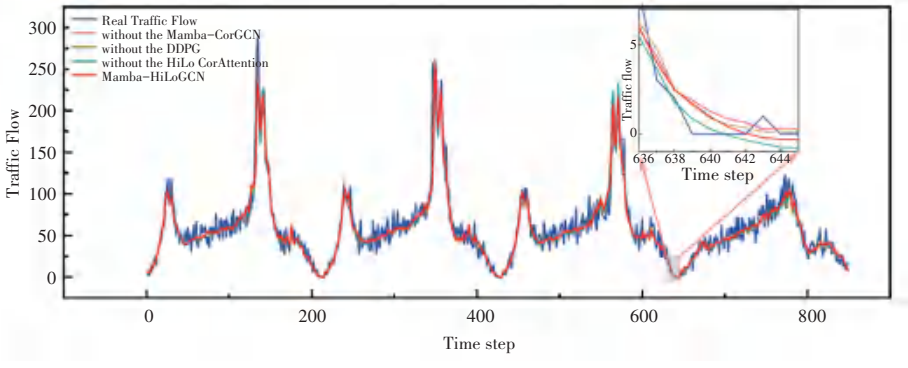


图 6　HZME（in）预测结果

Fig. 6　HZME（in）prediction results

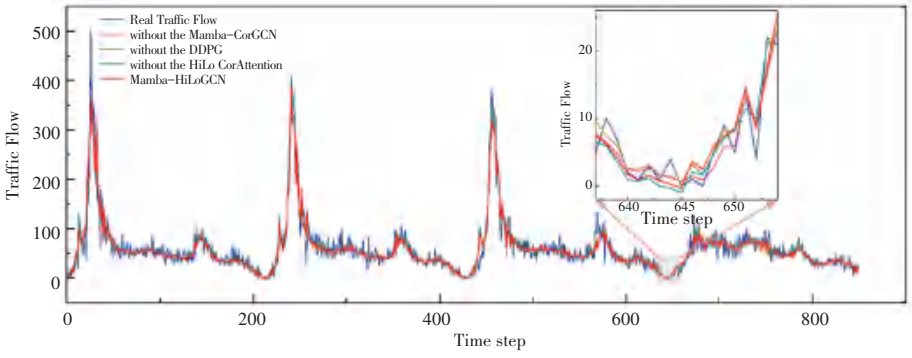


图 7 HZME (out) 预测结果

Fig. 7 HZME (out) prediction results

2.3.3 时间成本研究

本文在 PyTorch 2.0.0、Python 3.8 和 CUDA 11.8 环境中实现了 Mamba-HiLoGCN 方法。表 4 为 Mamba-HiLoGCN 各核心模块在不同规模数据集上的时间成本。其中,SC 因需要处理复杂的空间依赖关系而耗时较长,而 TC 凭借其轻量级设计展现出更高的计算效率。Mamba-CorGCN 在时空特征提取方面实现了较好的平衡,时间消耗处于中等水平。HiLo CorAttention 因需要动态调整高低频特征的注意力权重而稍慢。DDPG 需要通过大量试错来优化策略,这种迭代学习方法虽然有效,但也带来了较多时间消耗。表 4 为 Mamba-CorGCN、HiLo CorAttention 和 DDPG 在计算复杂度和功能实现方面的关系,为后续性能优化提供了重要参考。

表 4 计算效率

Table 4 Computational efficiency

数据集	SC	TC	Mamba-CorGCN	HiLo CorAttention	DDPG
PeMS03	0:55:00	0:01:29	0:32:00	0:40:00	1:10:00
PeMS04	0:48:00	0:01:07	0:25:00	0:31:00	1:00:00
PeMS07	1:20:00	0:01:58	0:40:00	0:46:00	1:50:00
PeMS08	0:20:00	0:00:43	0:18:00	0:25:00	0:50:00
HZME(in)	0:00:17	0:00:12	0:00:50	0:01:00	0:15:00
HZME(out)	0:00:17	0:00:12	0:00:50	0:01:00	0:15:00

3 结束语

本文提出了一种基于 Mamba-CorGCN、HiLo CorAttention 和 DDPG 的交通流预测方法 Mamba-HiLoGCN。Mamba-CorGCN 用于提取时空特征。DDPG 自动适应不同时间步和场景的邻接矩阵以提高预测精度。HiLo CorAttention 使该方法能够自适应地增加对高频和低频特征的关注,提高了方法对交通流变化的适应性和鲁棒性。本文使用 HZME(in)、HZME(out)、PeMS03、PeMS04、PeMS07 和 PeMS08 数据集对 Mamba-CorGCN 的性能进行了实验评估。实验结果表明,Mamba-CorGCN 的性能优于 HA、ARIMA、LSTM、GRU、CNN、DCRNN、GMAN、T-GCN、STGCN、ASTGCN、STSGCN、AGCRN、STGODE、STCGCN、Bi-STAT、ASTMGODE、ASTTN 和 DDGformer。

未来,Mamba-HiLoGCN 将扩展到空气质量预测领域。通过整合气象数据、污染源分布等多模态信息,进一步验证 Mamba-HiLoGCN 框架在动态图结构建模和长短期特征捕捉方面的泛化能力,最终实现交通流预测与空气质量预测的协同优化,为智慧城市建设提供统一的时空决策支持平台。

参考文献

[1] 叶明壕. 基于改进 LSTM 与分层动态图卷积的交通流预测[J/OL]. 智能计算机与应用 (2025-12-16)[2026-04-10]. DOI: 10.20169/j.issn.2095-2163.25112603.

[2] 陈以, 齐兴宇, 胡水源, 等. 基于 VMD 多阶段优化的短时交通流预测研究[J]. 计算机仿真,2025,42(1):126-132.

[3] SUN Y, LIAO K. A hybrid model for metro passengers flow prediction[J]. Systems Science & Control Engineering, 2023, 11(1): 2191632.

[4] 张天逸, 孙毅然, 刘凡琪, 等. 基于 K 均值聚类算法与 RBF 神经网络的交通流预测方法[J]. 智能计算机与应用,2020,10(8):148-151.

[5] MÉNDEZ M, MERAYO M G, NÚÑEZ M. Long-term traffic flow forecasting using a hybrid CNN-BiLSTM model [J].

Engineering Applications of Artificial Intelligence, 2023, 121: 106041.
[6] 戴洪波, 曾献辉. 基于 BP 神经网络的高速公路短时交通流预测[J]. 智能计算机与应用,2015,5(4):36-38.
[7] AL-HUTHAIFI R, LI T, AL-HUDA Z, et al. FedGODE: Secure traffic flow prediction based on federated learning and graph ordinary differential equation networks[J]. Knowledge-based Systems, 2024, 299: 112029.
[8] 朱凯利, 朱海龙, 刘靖宇, 等. 基于图卷积神经网络的交通流量预测[J]. 智能计算机与应用,2019,9(6):168-170.
[9] WEI S, HU D, WEI F, et al. Traffic flow prediction with multi-feature spatio-temporal coupling based on peak time embedding[J]. The Journal of Supercomputing, 2024, 80(16): 23442-23470.
[10] 陈艳, 徐坤财, 李云松, 等. 基于多模态时空变分图自编码器的交通流预测[J]. 武汉理工大学学报,2025,47(9):101-106.
[11] FANG S, PAN X, XIANG S, et al. Meta-MSNet: Meta-learning based multi-source data fusion for traffic flow prediction[J]. IEEE Signal Processing Letters, 2020, 28: 6-10.
[12] KUMAR R, MOREIRA J M, CHANDRA J. DyGCN-LSTM: A dynamic GCN-LSTM based encoder-decoder framework for multistep traffic prediction[J]. Applied Intelligence, 2023, 53(21): 25388-25411.
[13] YUAN D, XUE J, SU J, et al. ST-Mamba: Spatial-temporal mamba for traffic flow estimation recovery using limited data[C]//Proceeding of 2024 IEEE/CIC International Conference on Communications in China. Piscataway, NJ: IEEE, 2024: 1928-1933.
[14] LI Z, WANG L, WANG D, et al. Short-term traffic-flow forecasting based on an integrated model combining bagging and stacking considering weight coefficient[J]. Electronics, 2022, 11(9): 1467.
[15] LI W, SUI L, ZHOU M, et al. Short-term passenger flow forecast for urban rail transit based on multi-source data[J]. EURASIP Journal on Wireless Communications and Networking, 2021, 2021(1): 9.
[16] CHEN J, CAO J, HU M. A rapid semi-analytical approach for modeling traffic flow on changing road conditions and its application[J]. Applied Mathematics and Computation, 2024, 481: 128932.
[17] ZENG W, WANG K, ZHOU J, et al. Traffic flow prediction based on hybrid deep learning models considering missing data and multiple factors[J]. Sustainability, 2023, 15(14): 11092.
[18] QU L, ZHOU Y, LI J, et al. HMM-based map matching and spatiotemporal analysis for matching errors with taxi trajectories[J]. ISPRS International Journal of Geo-Information, 2023, 12(8): 330.
[19] RAJEH T M, LI T, LI C, et al. Modeling multi-regional temporal correlation with gated recurrent unit and multiple linear regression for urban traffic flow prediction[J]. Knowledge-based Systems, 2023, 262:110237.
[20] SALAMANIS A I, GRAVVANIS G A, KOTSIANTIS S B, et al. Novel sparse feature regression method for traffic forecasting[J]. International Journal on Artificial Intelligence Step Tools, 2023, 32(1): 2350008.
[21] MIAO Y, BAI X, CAO Y, et al. A novel short-term traffic prediction model based on svd and arima with blockchain in industrial internet of things[J]. IEEE Internet of Things Journal, 2023, 10(24): 21217-21226.
[22] JIA C, ZHANG W, HE Y, et al. Context based spatial-temporal graph convolutional networks for traffic prediction[J]. Knowledge-Based Systems, 2025, 310: 112933.
[23] KUO R J, KUNARSITO D A. Residual stacked gated recurrent unit with encoder-decoder architecture and an attention mechanism for temporal traffic prediction[J]. Soft Computing-A Fusion of Foundations, Methodologies & Applications, 2022, 26(17): 14259-14276.
[24] LIN W, ZHANG Z, REN G, et al. MGCN: Mamba-integrated spatiotemporal graph convolutional network for long-term traffic forecasting[J]. Knowledge-based Systems, 2025, 309: 112875.
[25] LI Y, CHAI S, MA Z, et al. A hybrid deep learning framework for long-term traffic flow prediction[J]. IEEE Access, 2021, 9: 11264-11271.
[26] CINI N, AYDIN Z. A deep ensemble approach for long-term traffic flow prediction[J]. Arabian Journal for Science and Engineering, 2024, 49(9): 12377-12392.
[27] LIU S, LIN W, WANG Y, et al. Convolutional neural network-based bidirectional gated recurrent unit-additive attention mechanism hybrid deep neural networks for short-term traffic flow prediction[J]. Sustainability, 2024, 16(5): 1986.
[28] LIN Y, SONG C, GUO S, et al. Spatial-temporal synchronous graph convolutional networks: A new framework for spatial-temporal network data forecasting[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2020, 34(1): 914-921.
[29] BAI L, YAO L, LI C, et al. Adaptive graph convolutional recurrent network for traffic forecasting[C]//Proceedings of Advances in Neural Information Processing Systems. NeurIPS, 2020: 17804-17815.
[30] FANG Z, LONG Q, SONG G, et al. Spatial-temporal graph ode networks for traffic flow forecasting[C]//Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining. New York: ACM, 2021: 364-373.
[31] BAO Y, HUANG J, SHEN Q, et al. Spatial-temporal complex graph convolution network for traffic flow prediction[J]. Engineering Applications of Artificial Intelligence, 2023, 121: 106044.
[32] CHEN C, LIU Y, CHEN L, et al. Bidirectional spatial-temporal adaptive transformer for urban traffic flow forecasting[J]. IEEE Transactions on Neural Networks and Learning Systems, 2022, 34(10): 6913-6925.
[33] CHEN Y T, LI C, LI S. Attention based spatial-temporal multi-graph ordinary differential equation network for traffic flow prediction[J]. Engineering Applications of Artificial Intelligence, 2025, 149: 110526.
[34] XUE Z, HUANG L, NING Q. ASTTN: An adaptive spatial-temporal transformer network for traffic flow prediction[J]. Engineering Applications of Artificial Intelligence, 2025, 148: 110263.
[35] LI Y, XU H, ZHANG T, et al. DDGformer: Direction-and distance-aware graph transformer for traffic flow prediction[J]. Knowledge-based Systems, 2024, 302: 112381.