

文章编号: 1671-1637(2026)01-0046-29

交通预测中的时空图神经网络研究综述: 从模型解构到发展路径

贾兴利^{*1}, 曲远海¹, 朱浩然¹, 杨宏志¹, 姚慧², 李孟晖³

(1. 长安大学 公路学院, 陕西 西安 710064; 2. 陕西省交通规划设计研究院有限公司, 陕西 西安 710065; 3. 中交路桥建设有限公司, 北京 101113)

摘要:为厘清交通预测模型的发展路径,探索未来交通预测发展方向,通过系统化的文献分析方法确立了以时空图神经网络为主导的技术发展方向;基于时空图神经网络框架特征,构建了从数据预处理、动静态图构建、时空特征提取到特征融合的全流程分析体系;系统梳理了典型交通预测任务及其对应的开源数据集;归纳了基于拓扑关系、距离特性、相似度计算等静态图构建方法,以及动态图直接优化与特征优化等前沿图构建技术;从时间与空间两大维度总结分析了当前较新的时空特征建模与空间特征建模方法,并通过 Graph WaveNet 与 DCRNN 两大典型案例阐释了时空特征融合机制;针对深层网络训练中的梯度异常和性能衰退问题,总结了通用信息传递的解决方案;探索了对比学习、预训练机制、因果推理、混合专家模型等新兴技术与交通预测的结合路径。研究结果表明:无论是中国还是国外,时空图神经网络在交通预测的应用研究已经逐渐白热化,其中中国在统计中以 1 671 篇的发文量占据榜首;现有研究主要聚焦于提高模型对于时空特征的记忆能力与构建最优图结构,而这类优化方案已进入模型性能提升与效率提高的平衡期;在时间建模方面,现有研究仍在寻找计算效率与运行性能的平衡点,而空间建模则成为现有模型效率提升的主要阻碍;根据对过去研究的总结梳理,未来的突破方向或将集中于新型预测场景的开拓、模型可解释能力的提升、现实物理约束的添加、新型学习策略的创新以及工业级部署方案的探索,为智慧交通系统建设提供更坚实的技术支撑。

关键词:智能交通系统;交通时空预测;综述;时空图神经网络;时空建模;图结构

中图分类号:U491.14 **文献标志码:**A **DOI:**10.19818/j.cnki.1671-1637.2026.01.003

Research review on STGNN in traffic prediction: From model deconstruction to development path

JIA Xing-li^{*1}, QU Yuan-hai¹, ZHU Hao-ran¹, YANG Hong-zhi¹, YAO Hui², LI Meng-hui³

(1. School of Highway, Chang'an University, Xi'an 710064, Shaanxi, China; 2. Shaanxi Province Transport Planning Design and Research Institute Co., Ltd., Xi'an 710065, Shaanxi, China; 3. Road & Bridge International Co., Ltd., Beijing 101113, China)

Abstract: To clarify the development path of traffic prediction models and to explore future development directions of traffic prediction, a systematic literature analysis approach was

出版历程: 2025-02-10 收稿, 2025-04-19 修回, 2025-06-06 录用

基金项目: 国家重点研发计划(2020YFC1512003); 陕西交通科研项目(24-41X); 长安大学中央高校基本科研业务费专项资金项目(300102212203); 陕西省技术创新引导计划基金项目(2025QCY-KXJ-139); 陕西省重点研发计划(2025SF-YBXM-283)

* 作者简介: 贾兴利(1986-), 男, 山东济宁人, 教授, 博士生导师, 博士, E-mail: jiaxingli@chd.edu.cn.

引用格式: 贾兴利, 曲远海, 朱浩然, 等. 交通预测中的时空图神经网络研究综述: 从模型解构到发展路径[J]. 交通运输工程学报, 2026, 26(1): 46-74.

Citation: JIA Xing-li, QU Yuan-hai, ZHU Hao-ran, et al. Research review on STGNN in traffic prediction: From model deconstruction to development path[J]. Journal of Traffic and Transportation Engineering, 2026, 26(1): 46-74.

adopted, and a technology development direction dominated by spatiotemporal graph neural networks (STGNN) was established. Based on the framework characteristics of STGNNs, a full-process analysis system was constructed, including data preprocessing, static and dynamic graph construction, spatiotemporal feature extraction, and feature fusion. Typical traffic prediction tasks and their corresponding open-source datasets were systematically reviewed. Static graph construction methods based on topological relationships, distance properties, and similarity calculations were summarized, and frontier graph construction technologies were summarized, including direct optimization of dynamic graphs and feature optimization. Current temporal feature modeling methods and spatial feature modeling methods were analyzed from the two dimensions of time and space, and the spatiotemporal feature fusion mechanism was illustrated through two typical cases of Graph WaveNet and DCRNN. For the problems of gradient anomalies and performance degradation in deep network training, general solutions based on information propagation were summarized. The integration paths of emerging technologies were explored, including contrastive learning, pre-training mechanisms, causal reasoning, and mixture-of-experts models with traffic prediction. Analysis results show that applications of STGNNs in traffic prediction have gradually intensified both domestically and internationally, and China ranks first with 1 671 publications in the statistics. Existing studies are mainly focused on improving model memory ability for spatio-temporal features and constructing optimal graph structures, and such optimization schemes have reached a balance between model performance and efficiency. In temporal modeling, a balance between computational efficiency and operational performance is still being explored, while spatial modeling has become the main obstacle to efficiency improvement of existing models. Based on the summary and review of previous studies, future breakthrough directions are expected to be concentrated on the exploration of novel prediction scenarios, improvement of model interpretability, incorporation of real-world physical constraints, innovation of learning strategies, and exploration of industrial-level deployment solutions, so that stronger technical support is provided for intelligent transportation systems.

Keywords: intelligent transportation system; traffic spatio-temporal prediction; review; spatio-temporal graph neural network; spatio-temporal modeling; graph structure

Publication history: Received 2025-02-10; Received in revised form 2025-04-19; Accepted 2025-06-06

Funding: National Key R&D Program of China (2020YFC1512003); Science and Technology Project of Shaanxi Department of Transportation (24-41X); Fundamental Research Funds for the Central Universities (300102212203); Technology Innovation Guidance Program of Shaanxi Province (Fund)(2025QCY-KXJ-139); Key R&D Program of Shaanxi Province (2025SF-YBXM-283)

* **Corresponding author:** JIA Xing-li, professor, PhD, E-mail: jiaxingli@chd.edu.cn.

0 引 言

随着中国经济的高速发展与城市化进程的持续推进,有限的道路承载力与机动车保有量激增之间的结构性矛盾日益凸显,已成为制约现代交通发展的核心瓶颈。在此背景下,智能交通系统(Intelligent Transportation System, ITS)的创新应用为破解这一难题提供了有效路径^[1]。ITS通过深

度整合物联网(Internet of Things, IoT)、边缘计算、移动通信等新一代信息通信技术,构建起“人-车-路-环境”多维协同的智能管控体系,借助实时信息交互与智能决策算法实现交通系统各子系统的动态协同优化。作为 ITS 的关键技术支撑,交通预测通过前瞻性研判路网运行态势,赋能管理部门实施主动式交通管控,对预防交通拥堵蔓延、阻断级联失效风险具有重要意义。因此,构建精准高效的交通

预测模型既是实现智慧交通治理的核心技术支撑,也是当前 ITS 领域最具前沿性的研究方向之一。

交通预测模型的核心架构由底层数据与预测方法共同构成,呈现出显著的双向促进特征。从技术演进视角分析,路域测控系统、浮动车、ETC 门架、无人机遥感以及北斗动态定位等多元感知技术的迭代升级,实现了交通数据从单模态同构型向多源异构复合体的范式转变^[2],为高精度时空预测奠定了数据基础。而在方法论层面则呈现出从先验模型驱动到深度数据驱动的方法迁移,传统基于模型驱动的预测体系主要依托交通波动理论^[3]和到达流分布^[4]假设进行微观车辆行为建模,或通过跟驰模型^[5]构建车辆交互动力学仿真系统,其中元胞传输模型^[6]在宏观路网建模中展现出独特优势。此类机理驱动方法在小样本封闭场景具有良好适用性,但其参数化建模方式受限于固有理论假设,在应对多源异构大数据时存在特征表征能力不足的固有缺陷。相较而言,数据驱动模型通过构建非参数化特征建模框架,利用历史数据的时空关联性建立自适应学习机制,早期研究聚焦单一路段的时间序列建模,典型如历史均值模型^[7]、多元线性回归、移动平均模型^[8]以及差分自回归模型^[9]等方法。尽管这类方法能够有效挖掘时序数据的平稳性特征,但在处理交通流数据中普遍存在的非线性、非平稳性以及突变响应等复杂特性时,其动态建模能力与适应性仍面临挑战。

与统计学方法的建模范式不同,机器学习模型通过数据驱动的参数空间映射实现特征表征,其架构具有自适应性特征。典型模型如支持向量机(Support Vector Machine, SVM)^[10]、随机森林(Random Forest, RF)^[11]以及 K 近邻模型(K Nearest Neighbor, KNN)^[12],通过构建非线性特征提取机制突破传统统计学的线性假设约束,实证研究表明浅层机器学习模型相较于统计学方法具有显著优势^[13],其高维空间映射能力有效提升了小样本数据的特征解耦效率。数据规模的指数级增长及误差反向传播算法的出现^[14]推动了深度学习网络的发展,基于特征提取范式的差异可将其划分为深度信念网络(Deep Belief Network, DBN)^[15]、卷积神经网络(Convolutional Neural Network, CNN)^[16]、循环神经网络(Recurrent Neural Network, RNN)^[17],及其变式长短期记忆网络(Long Short-term Memory, LSTM)^[18]、门控循环单元(Gated Recurrent Unit, GRU)^[19]等以及 Transformer^[20]

等类别。不少研究将深度学习模型应用于交通预测任务并取得了优秀的效果:马飞等^[21]通过结合格莱姆角场、CNN 与 LSTM 将外部气象条件融入交通预测模型中;胡立伟等^[22]将双向 LSTM 与注意力机制融合提出一种行车风险预测方法;Ma 等^[23]基于 CNN-GRU 提出一种适应于交通预测的特征选择算法。然而,RNN 架构通过门控机制调控时序信息流,但其序列建模存在无法并行计算的固有约束,CNN 模型则基于离散卷积核的局部感知特性实现空间特征聚合,而 Transformer 凭借自注意力机制在图像处理^[24]与序列建模^[25]领域引发架构革新,其交通预测应用^[26]已验证了其跨任务建模有效性。需指出的是,尽管 CNN 与 Transformer 在并行计算方面具有优势,但二者在非欧氏交通网络拓扑结构的表征学习方面仍存在理论局限性。

近年来,图卷积网络(Graph Convolutional Network, GCN)在非欧氏空间节点表征学习中的突破性进展,推动了时空图神经网络(Spatio-temporal Graph Neural Network, STGNN)^[27]的发展与进步。研究者通过融合传统深度学习架构与图拓扑建模技术,构建出具有时空联合特征提取能力的预测框架,这一架构创新在交通流预测领域取得显著成效并成为主流方法。Dong 等^[28]通过超网络提出了一种时空元参数学习范式用于时空异质性学习的 STGNN 模型;Li 等^[29]提出一种时空图融合神经网络,从多个语义环境中提取多尺度时间依赖关系并构建了动态自适应图实现精确时空预测。伴随 STGNN 的应用拓展,相关综述研究呈现体系化发展趋势。Wang 等^[30]构建了跨领域应用图谱,涵盖交通流态、需求响应及气象关联等多维度应用场景;崔建勋等^[31]系统阐释了涵盖 STGNN 模型在内的短时交通深度学习预测技术;邹慧琪等^[32]以 STGNN 算法在交通预测中的应用为主线,整理分析近年来关于时空图神经网络的最新研究进展。聚焦交通预测的专项研究^[33-34]虽深化了模型优化方向,但现有研究多聚焦于模型架构创新,而在数据治理体系构建与模块化架构解耦分析方面存在显著方法论缺位。

相较于以往的交通预测综述,本文的贡献主要如下。

(1)针对交通预测任务中的时空耦合特性,从数据模态、图结构构建、时空特征解耦建模以及特征耦合机制 4 个维度,对 STGNN 模型框架展开系统性调研。揭示了不同方法在道路网络动态演化表征与

时空依赖协同建模方面的技术演进路径。

(2)系统梳理了时空图神经网络在交通预测领域的理论基础与应用实践,重点剖析了先进时空图神经网络的创新模式,概括性梳理总结了现有先进方法的优势与代表性模型。

(3)结合现有 AI 前沿技术及 STGNN 模型发展脉络,从多维方向对时空图神经网络方法的发展趋势进行试验性探讨,为后续方法论创新提供参考依据。

1 文献统计与框架搭建

1.1 文献来源与趋势

本研究采用系统性文献计量方法,基于中国知网(China National Knowledge Infrastructure, CNKI)和科学网(Web of Science, WoS)核心数据库进行检索。考虑交通预测本身带有时间属性,中文文献以“主题=(时空 OR 空间)AND(交通预测)”为逻辑表达式,时间跨度设定为 2017~2025 年(涵盖 DCRNN 模型^[27]提出后的关键演进阶段),经筛选后获得有效文献 1 724 篇,其中核心期刊论文 524 篇、学位论文1 174 篇、会议文献 26 篇。英文文献采用“Topic=(Spatiotemporal OR Spatial OR Spatio) AND (Traffic prediction OR Traffic forecast)”检索关键词,共获取相关文献 20 010 篇。图 1 展示了 2017~2024 年(排除 2025 年数据)国内外学术期刊发文量演变趋势(WoS 数据限定文献类型为研究论文 Article,CNKI 数据限定为学术期刊论文)。由图 1 可知:时空交通预测研究呈现显著增长规律且全球范围内年度发文量保持近似线性增长态势,这印证了智慧交通发展进程中,交通供需矛盾催生精准时空预测需求,同时也表明了时空交通预测仍具有较好的发展前景。

1.2 关键词分析

对 CNKI 与 WoS 的导出文献利用 Cite Space 进行整理,规定 CNKI 文献为 2017~2025 年所有文献,考虑 WoS 文献较多,对 2022~2025 年核心集前 500 期刊进行统计分析,绘制关键词聚类图如图 2 所示。规定同一文章关键词出现次数为 1 次,诸如“Neural network”与“Neural networks”类似重复词取频次较高关键词,去除“Model”“Flow”“Study”“Paper”“Data prediction”类似研究方向无意义词汇,获得关键词共现频次如表 1 所示。

结合图 2 与表 1 可知:国内外现有时空交通预测方法主要为深度学习模型;国内外所用共同方法

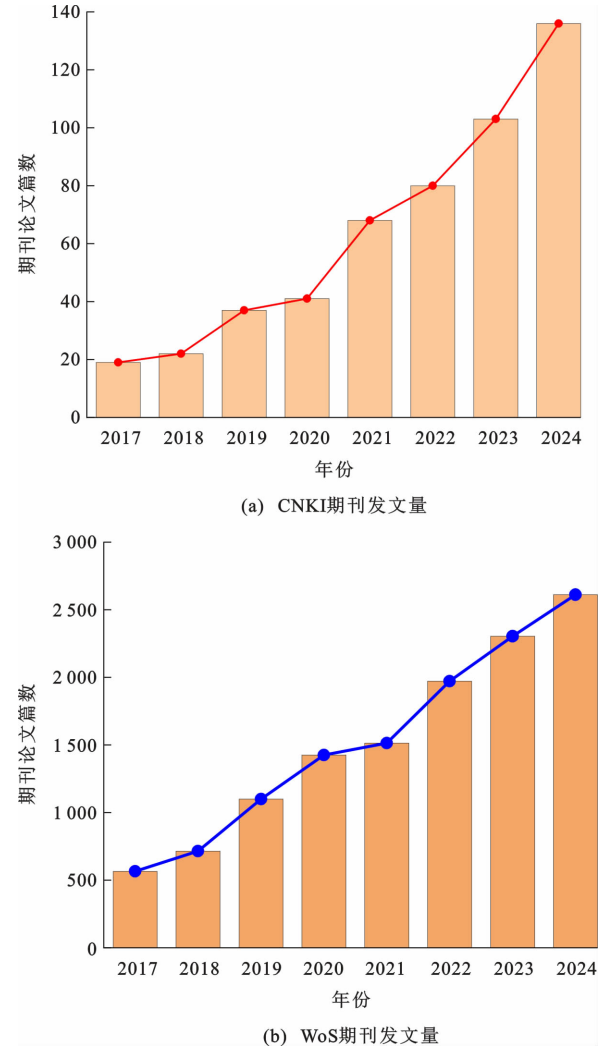


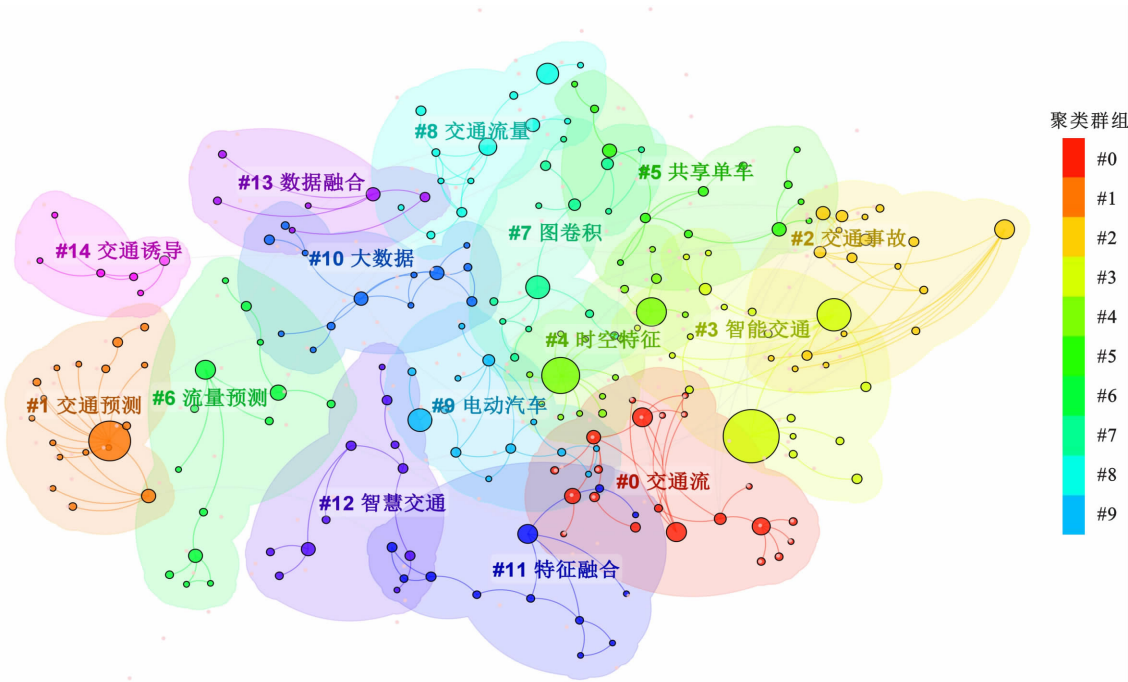
图 1 国内外发文趋势统计
Fig. 1 Trend statistics of domestic and international publications

表 1 关键词共现频次排序

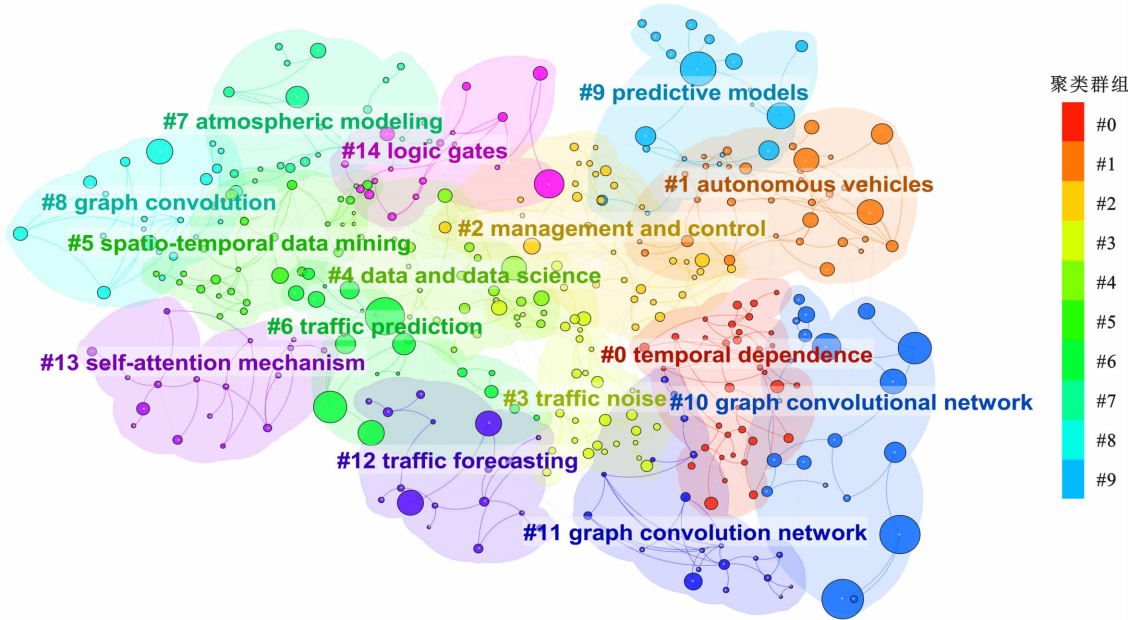
Table 1 Keyword co-occurrence frequency ranking

序号	CNKI		WoS	
	关键词	频次	关键词	频次
1	深度学习	195	Deep learning	290
2	交通预测	100	Prediction	280
3	时空特征	84	Predictive models	258
4	智能交通	65	Traffic prediction	219
5	时空数据	54	Traffic flow prediction	187
6	时空特性	34	Neural network	184
7	图卷积	34	Graph convolutional network	150
8	时空预测	25	Attention mechanism	134
9	神经网络	23	Feature extraction	121
10	城市交通	23	Trajectory prediction	120

为图卷积神经网络,即引言中所述 STGNN 类模型。交通预测模型构建与 STGNN 模型结构 2 大框架确定文章思路,而针对关键词及聚类中出现的特征融



(a) CNKI关键词聚类



(b) WoS关键词聚类

图 2 国内外关键词聚类

Fig. 2 Domestic and international keyword clustering

合、Feature extraction、Attention mechanism 则在后续方法论整理中进行统一概述。

1.3 主要研究国家分析

利用 Cite Space 对 WoS 2022~2025 年核心集前 500 期刊约 2 000 篇文献进行统计分析,主要研究国家及平均被引量如表 2 所示。由表 2 可知:中国发文量位居第 1 且远超位居第 2 的美国,可见中国对于智慧交通的重视;中国平均每篇被引量约为

8 篇,位居第 7,远低于新加坡的 28 篇,同时低于澳大利亚、加拿大,与韩国、英格兰、美国相近;虽然中国已取得较好成果,但后续更应注重研究质量,争取保质保量的高质量发展。

1.4 行文结构搭建

STGNN 框架如图 3 所示,从图 3 可知:STGNN 模型通过空间建模模块与时间建模模块将输入数据与图结构转换成隐藏状态,并通过全连接

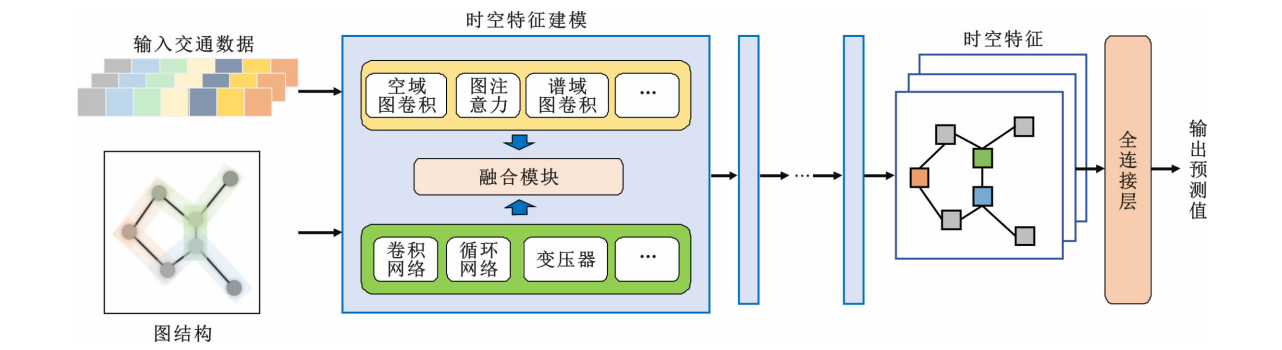


图 3 STGNN 模型框架

Fig. 3 Framework of STGNN model

表 2 主要研究国家分布

Table 2 Distribution of main research countries

国家	发文篇数	发文量排序	平均被引篇数	平均被引量排序
中国	1 671	1	8. 01	7
美国	300	2	8. 89	4
英格兰	101	3	8. 38	6
澳大利亚	98	4	12. 31	3
加拿大	73	5	12. 51	2
印度	65	6	4. 12	10
韩国	63	7	8. 40	5
新加坡	59	8	27. 59	1
日本	45	9	5. 53	9
德国	38	10	7. 44	8

层输出预测值。本综述行文组织以图 3 框架为依托,建立图 4 所示组织结构。从图 4(b)可以看出:论文以 STGNN 模型搭建过程为主线,以各模块发展为支线整理交通预测中的 STGNN 网络;第 1 部分介绍了交通预测任务的定义以及常用的公开数据集,这些数据被广泛应用于交通预测任务;第 2 部分从静态与动态角度介绍了 STGNN 中时空图的构建方法;第 3 部分介绍了 STGNN 模型的时空建模部分,从空间特征建模、时间特征建模以及融合模块介绍了主流建模方法;最终以整体 STGNN 模型为研究对象,整理了先进 STGNN 预测方法并给出可能的未来研究方向。

2 交通预测问题分类及数据整理

2.1 交通预测问题分类

交通预测问题设计包括交通状态预测、交通需求预测、交通事故预测、行程时间预测、车辆轨迹预测等多种形式预测,本节以基于 STGNN 模型的多变量交通预测为研究对象,将交通预测任务根据输入数据形式,简要分为状态预测类任务与轨迹预测类任务,本文以状态预测类任务为主、轨迹预测类任

务为辅进行调研分析。

2.1.1 状态预测类任务

状态预测类任务以交通状态预测(交通流量^[35]、车辆运行速度^[36]等)为主,其数据来自于交通传感器,每个传感器生成一定时间间隔的时间序列数据,这些时间序列数据被称为关联时间数据(Correlated Time Series, CTS)^[37],CTS 不仅仅包含时间序列内的时间依赖性关系,还包含时间序列之间的空间相关性关系。这种空间关系可以时空图 G 的形式进行表示。交通预测任务根据预测序列长度可被定义为单步预测任务与多步预测任务,单步预测任务通过学习输入序列预测未来 1 个时间步数据

$$\tilde{v}_{t+P+1} = S_1(v_{t+1}, v_{t+2}, \dots, v_{t+P}; \mathbf{A}; \mathbf{W}) \quad (1)$$

式中: $S_1(\cdot)$ 为单步预测函数,即执行单步预测的交通预测模型; $\mathbf{W}$ 为可训练矩阵; $\tilde{v}_{t+P+1}$ 为未来一个时间步的交通状态预测值; $\mathbf{A}$ 为邻接矩阵; $v_{t+1} \sim v_{t+P}$ 为 $t+1 \sim t+P$ 时刻输入到模型的交通状态数据(流量、速度、密度等), P 为输入时间步长。

多步预测则是一次性预测 Q 步交通数据,其过程为

$$\tilde{v}_{t+P+1:t+P+Q} = S_2(v_{t+1}, v_{t+2}, \dots, v_{t+P}; \mathbf{A}; \mathbf{W}) \quad (2)$$

式中: $S_2(\cdot)$ 为多步预测函数,即执行多步预测的交通预测模型; $\tilde{v}_{t+P+1:t+P+Q}$ 为第 1 步到第 Q 步的输出。

除交通状态预测外,诸如行程时间预测^[38]、交通事故预测(交通事故发生率^[39]、交通事故发生时间^[40]、交通拥堵持续时间^[41])以及交通需求预测(订单需求^[42]、电动车充电需求^[43]、共享单车需求^[44]等)均以相似形式执行,其区别主要在于外部事件的嵌入以及 CTS 数据类型。

2.1.2 轨迹预测类任务

车辆轨迹及行人轨迹预测是交通预测的主要任

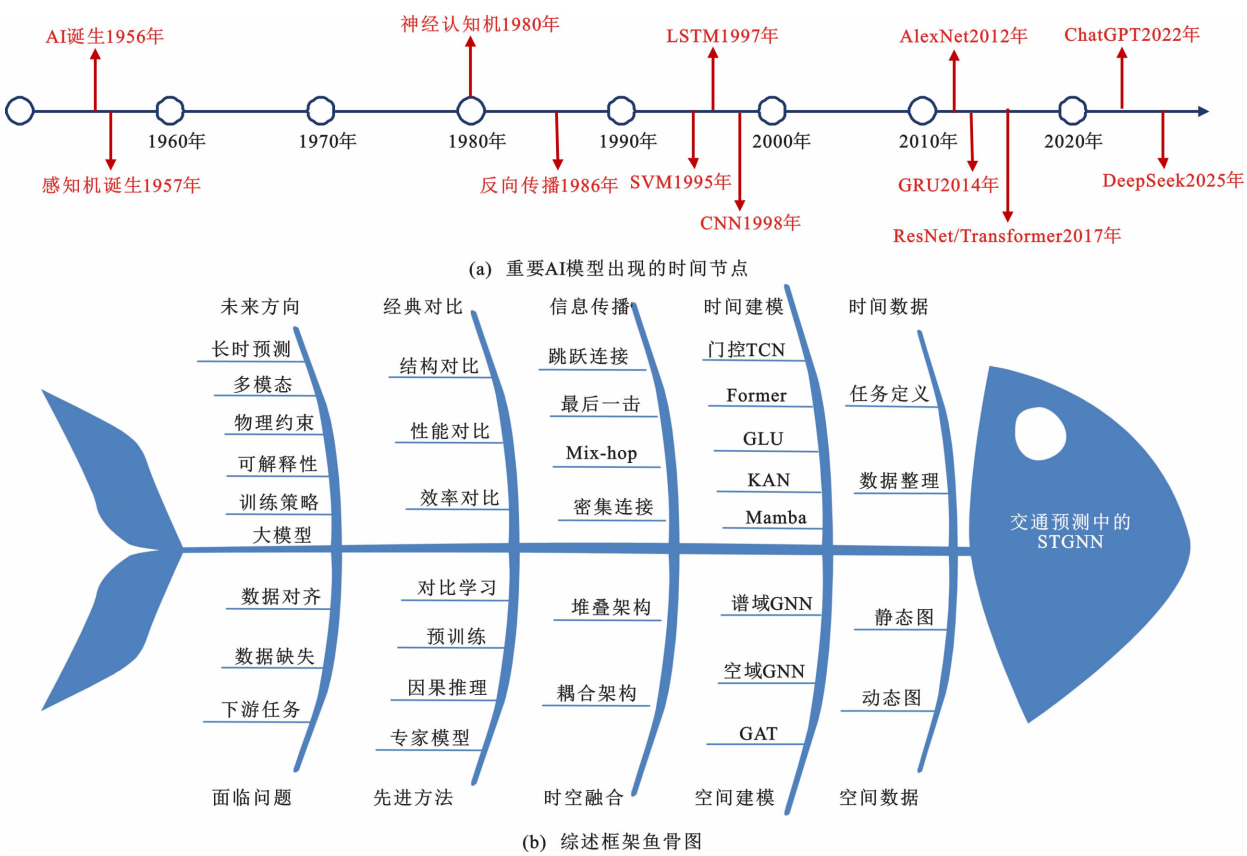


图 4 综述框架与人工智能发展路径

Fig. 4 Framework of review and development of AI

务之一,也是实现高等级自动驾驶的重要环节,与交通数据预测类任务相似,其预测可基于物理模型、深度学习、强化学习、传统机器学习多种方式^[45],由于不同目标轨迹存在相互关系,轨迹预测也可利用STGNN类模型解决^[46],通过对定义的时空图和历史数据的学习以预测未来坐标,从而实现车辆轨迹预测,其过程可被表示为

$$(\tilde{x}_{t+P+1:t+P+Q}, \tilde{y}_{t+P+1:t+P+Q}) = S_3[(x_{t:t+P}, y_{t:t+P}); e_{t:t+P}; \mathbf{A}; \mathbf{W}]$$

(3)

式中: $S_3(\cdot)$ 为轨迹预测函数; $(x_{t:t+P}, y_{t:t+P})$ 为 t 时刻到 $t+P$ 时刻的历史坐标位置; $(\tilde{x}_{t+P+1:t+P+Q}, \tilde{y}_{t+P+1:t+P+Q})$ 为 $t+P+1$ 时刻到 $t+P+Q$ 时刻的未来坐标数据; $e_{t:t+P}$ 为速度以及加速度等历史车辆状态数据。

2.2 交通预测公开数据整理

如表 3 所示,本节论文对各种任务部分可用公开数据集进行分类及溯源整理,并且将数据介绍(常用数据格式及内容)、所用模型及参考文献列入表 3,以供后续研究参考,其中所列预测任务的含义并非数据仅用于该任务,而是根据主要应用场景分类,例如 TaxiNYC 数据集、TaxiBJ 数据集也可用于

行程时间预测任务。

3 时空图构建方法综述

不管是单步预测还是多步预测,都需要在训练过程中不断更新时空图,根据时空图结构是否动态变化可分为图 5(a)所示静态图与图 5(b)所示动态图 2 种,图中 $\mathbf{W}$ 为可训练权重,图 5(a)脚标为不同边的编号,图 5(b)中 $\mathbf{W}_{t,1}$ 为第 1 条边在时刻 t 的权重。由图 5 可知:静态图连接关系仅与预定义矩阵有关,输入节点属性并不影响边的权重;动态图连接关系随着节点属性的变化而变化,同时其连接关系及边的权重值随着时间的变化而发生动态演变。为更好地区分静态图和动态图,本节分别以 2 种图的构建方法为研究对象进行总结分析。

3.1 静态时空图

静态时空图即预定义先验知识图,通过在训练前定义节点空间关系或挖掘交通数据相关性预定义邻接矩阵,并输入到模型当中进行特征学习以挖掘交通空间特征,交通预测中的静态图常用的静态图包括拓扑图、距离图、相似性图以及交互图等,本小节以不同静态图中的邻接矩阵为对象进行调研分析。

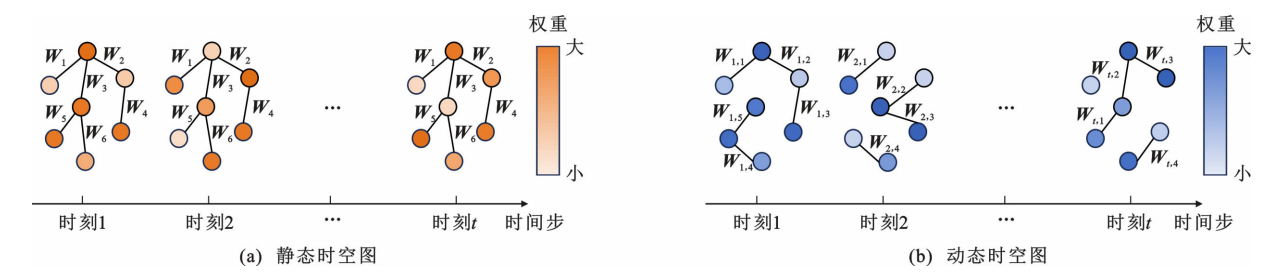


图 5 时空图变化过程

Fig. 5 Changing process in spatio-temporal graphs

表 3 不同交通预测任务可用公开数据集统计

预测任务		数据名称	数据介绍	部分应用模型
交通状态	PeMS	PeMS-Bay	PeMS-Bay 为加州高速公路 2017-01-01 到 2017-05-01 交通速度数据集, 包含 325 个节点,共 52 116 条数据,采样间隔为 5 min	DCRNN ^[27] AutoSTS ^[47] MTGNN ^[48]
			已处理数据: https://github.com/liyaguang/DCRNN (已征求作者同意)	
		PeMS03	PeMS03 为加州高速公路 2018-09-01 到 2018-11-30 交通数据,包括 358 个节点,共 26 208 条数据,常用数据为流量数据,采样间隔为 5 min	STGODE ^[49] STSGCN ^[50]
		PeMS04	PeMS04 为加州高速公路 2018-01-01 到 2018-02-28 交通数据, 包括 307 个节点,共 16 992 条数据,常用数据为交通流量、交通速度以及占有率,采样间隔为 5 min	Z-GCNETs ^[51] ASTGCN ^[52] STFGCN ^[29]
		PeMS07	PeMS07 为加州高速公路 2017-05-01 到 2017-08-31 交通数据,包括 883 个节点,共 28 224 条数据,常用数据为流量数据,采样间隔为 5 min	STGCN ^[53] MVSTGCN ^[54]
		PeMS08	PeMS08 为加州高速公路 2016-07-01 到 2016-08-31 交通数据, 包括 170 个节点,共 17 856 条数据,常用数据为交通流量、交通速度以及占有率,采样间隔为 5 min	ST-AE ^[35] StemGNN ^[55]
		数据源链接: http://pems.dot.ca.gov/		
	Metr-LA		Metr-LA 为洛杉矶地区 2012-03-01 到 2012-06-30 交通速度数据集,包含 207 个节点,共 24 272 条数据,采样间隔为 5 min	GWN ^[56] LightCTS ^[57]
			数据源链接: https://www.metro.net/ 已处理数据: https://github.com/liyaguang/DCRNN (已征求作者同意)	
	EXPY-TKY		EXPY-TKY 为东京地区 2021-12-01 到 2021-12-31 交通速度数据, 包含 1 843 条链接,共 13 248 条数据,采样间隔为 10 min	MegaCRN ^[58] TESTAM ^[59]
			数据源链接: https://en.wikipedia.org/wiki/Shuto_Expressway 已处理数据链接: https://github.com/deepkashiwa20/MegaCRN (MIT License)	
交通需求 行人流量	TaxiNYC		TaxiNYC 为纽约出租车 2011-01-01 到 2016-06-30 期间 3 500 万条行程数据,每条行程数据包括上车时间、下车时间、上车经纬度、下车经纬度、行程距离	CCRNN ^[42] MDSR ^[60]
			数据源链接: https://www1.nyc.gov/site/tlc/about/tle-trip-record-data.page	
	TaxiBJ		TaxiBJ 包含北京出租车 2013-07-01 到 2017-10-30、2014-03-01 到 2014-06-30、2015-03-01 到 2015-06-30、2015-11-01 到 2016-6-30 四个时间段超过 34 000 辆出租车行程数据	ST-ResNet ^[61]
			处理数据链接: https://github.com/topazape/ST-ResNet (MIT License)	
	BikeNYC		BikeNYC 包含纽约超过 6 800 辆自行车在 2014-04-01 到 2014-09-30 期间的行程数据,包括出行时长、起止站点 ID、起止时间	ST-ResNet ^[61] STMN ^[62]
			数据源链接: https://www.citibikenyc.com/system-data 处理数据链接: https://github.com/topazape/ST-ResNet (MIT License)	

续表 3

预测任务	数据名称	数据介绍	部分应用模型
交通事故	NYC Accident	NYC Accident 为多个时间段交通事故数据,如 GSNet 使用为 2013-01-01 到 2013-12-31 数据,RiskSeq 使用为 2017-01-01 到 2017-05-31 数据,包括交通事故记录,出租车订单、兴趣点以及天气等多类数据	GSNet ^[39] Riskoracle ^[63]
		数据源链接:https://opendata.cityofnewyork.us/ 已处理数据:https://github.com/Echohhhhhh/GSNet(已征求作者同意)	
	Chicago Accident	Chicago Accident 为 2016-02-01 到 2016-09-30 交通事故数据,包含 44 000+交通事故记录以及 1 744 000+出租车订单等多类数据	GSNet ^[39]
		数据源链接:https://data.cityofchicago.org 已处理数据:https://github.com/Echohhhhhh/GSNet(已征求作者同意)	
车辆轨迹 行程时间	Porto dataset	包含葡萄牙波尔图 2013-01-07 到 2014-06-30 的 442 辆出租车超过 300 000 条轨迹数据,其内容包括司机标号、行程开始时间、行程总时间等 9 种类型数据	STGNN-TTE ^[64] START ^[65]
		数据源链接:https://www.kaggle.com/datasets/crailitap/taxi-trajectory	
	T-Drive	T-Drive ^[66-67] 包含北京地区 2008-02-02 到 2008-02-08 期间 10 357 辆出租车约 1 500 条 GPS 轨迹数据	PreCLN ^[68]
		链接:https://www.microsoft.com/en-us/research/publication/t-drive-trajectory-data-sample/	
车辆轨迹	NGSIM	NGSIM 包含 I-80 和 US-1012 个高速公路段以及兰克希姆大道和桃树街 2 个干道段 9 206 条车辆轨迹数据,内容含车辆行驶状态、车辆坐标、车辆长度等多种数据,采样频率为 10 Hz	CDSTraj ^[69] STDAN ^[70]
		数据源链接:https://data.transportation.gov/Automobiles/Next-Generation-Simulation-NGSIM-Vehicle-Trajector/8ect-6jqj/about_data	
	HighD ^[71]	HighD 为德国科隆 6 条约 420 m 高速公路通过无人机采集的包含 110 000+辆汽车,总里程 45 000 km 的行驶数据,其中有约 5 600 条车辆变道记录,采样频率为 25 Hz	CDSTraj ^[69] STDAN ^[70]
		数据源链接:https://levelxdata.com/highd-dataset/	
	Argoverse ^[72]	Argoverse 包括 3D 追踪与运动轨迹 2 部分,其中 3D 追踪包含 LiDAR 和视觉传感器从 113 个场景采集的数据,轨迹数据包含 1 000+小时驾驶时间提取的 324 557 条轨迹数据	LaneGCN ^[73] STGACN ^[74]
		数据源链接:https://www.argoverse.org/	

https://outreach.didichuxing.com/en 可对 DiDi Chengdu 与 DiDi Xi'an 出租车 GPS 轨迹数据进行申请^[75]

3.1.1 拓扑矩阵

早期研究通过最直接的空间关系表达节点空间特征,通过对偶映射^[76]、原始映射^[77]、Space L 模型以及 Space P 模型^[78]等方式构建道路网拓扑模型,考虑同一条道路上下游^[79]或车站客流转移^[80]往往存在微观交通的关联性,这种拓扑关联性可以利用连通性关系表达^[52,81]。简单的拓扑邻接矩阵被定义为一个基于(0,1)的二值矩阵,当 2 个节点存在拓扑上的连接关系则表示为 1,否则表示为 0

$$a_{i,j}=\begin{cases}1 & \text{节点 } i \text{ 与节点 } j \text{ 相连} \\ 0 & \text{其他情况}\end{cases}\quad (4)$$

式中: $a_{i,j}$ 为邻接矩阵第 i 行第 j 列元素,即表达第 i 个节点与第 j 个节点关系的元素。

3.1.2 距离矩阵

拓扑矩阵仅能反映节点连接性关系,而忽视了不在同一道路的 2 个节点仍可能具有级联相关性,因此部分研究者根据托布勒的地理学第一定律“任何事物都是与其他事物相关的,只不过相近的事物关联更紧密”^[82]定义距离矩阵。相较于拓扑矩阵,距离矩阵不依赖拓扑关系,仅靠节点之间的距离信息表达空间特征,其元素通过与距离相关的核函数进行计算^[27,53],现有研究多利用带阈值的高斯核函数矩阵^[83-84]表示节点距离关系以构建距离邻接矩阵

$$a_{i,j}=\begin{cases}\exp\left(-\frac{d_{i,j}^2}{\gamma^2}\right) & a_{i,j}\geqslant\varepsilon \\ 0 & a_{i,j}<\varepsilon\end{cases}\quad (5)$$

式中: $d_{i,j}$ 为第 i 个与第 j 个节点的空间距离, $i\neq$

j ; γ 为控制邻接矩阵分布的超参数; ϵ 为预定义阈值。

3.1.3 相似性矩阵

事实上,仅依靠拓扑矩阵或距离矩阵很难完全反映现实世界交通节点关联性,由图 6(a)可知:学校与医院距离较近,若以距离矩阵计算则会赋予 2 个节点较高权重关系,然而由于反向问题,实际上 2 个节点相关性并不强;而医院与银行虽然距离较远,但处于同一道路上下游,反而具有较强相关性。图 6(b)表明:图中存在 2 个兴趣点(Point of Interest, POI),分别为游乐场与企业,2 个兴趣点的存在使得不在同一路段且距离较远的停车场、居民区 1、居民区 2 与居民区 3 在工作日与休息日均具有较强关联性,从而形成 POI 特性,这种特性并不能通过拓扑矩阵或距离矩阵进行表达。

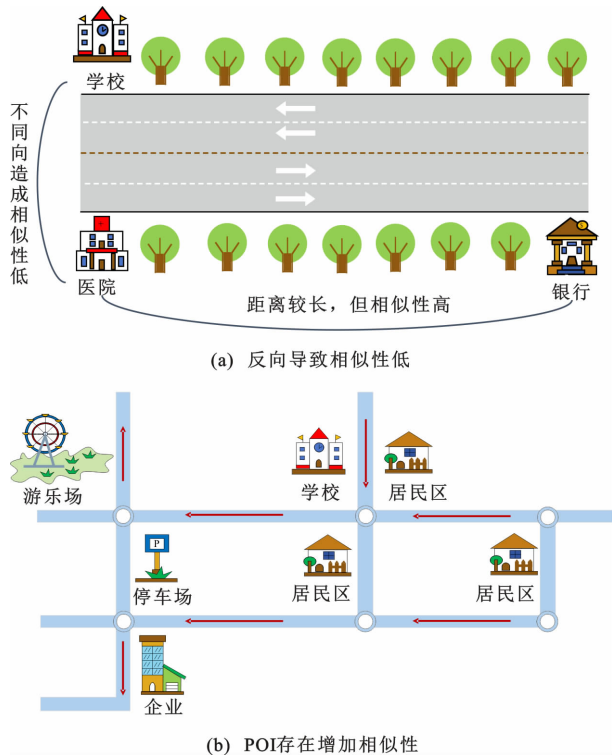


图 6 交通节点之间的特殊空间关系

Fig. 6 Special spatial relationships between transportation nodes

无论是拓扑矩阵还是距离矩阵,均期望通过定义一定的空间约束从而描述节点之间的时间相似性,因此可以认为交通预测任务中的空间特征描述事实是为了更好地捕捉节点之间的时间依赖性关系,后续研究者省略了这个步骤,直接从时间维度预定义相似性矩阵^[85-86],一定程度上克服了连通性与距离性这样的直接空间关系无法完全表达节点相似性的问题。

相似性矩阵通过在时间维度计算语义相近程度

以反映节点之间的关联性,简单来说,相似性矩阵通过度量函数直接计算节点属性相似程度,这种属性不仅可以是输入时间序列^[87],还可以是 POI 属性^[88],其中较为简单的方法为 Pearson 相关性系数^[89-90],如下所示

$$a_{i,j} = \begin{cases} \frac{\sum_{t=1}^P (v_{t,i} - \bar{v}_i)(v_{t,j} - \bar{v}_j)}{\sqrt{\sum_{t=1}^P (v_{t,i} - \bar{v}_i)^2} \sqrt{\sum_{t=1}^P (v_{t,j} - \bar{v}_j)^2}} & a_{i,j} \geq \epsilon \\ 0 & a_{i,j} < \epsilon \end{cases} \quad (6)$$

式中: $v_{t,i}$ 为第 i 个节点的第 t 个交通状态数据; $v_{t,j}$ 为第 j 个节点的第 t 个交通状态数据; $\bar{v}_i$ 、 $\bar{v}_j$ 分别为第 i 、 j ($i \neq j$) 个节点在输入时间步长 P 范围内的均值。

Pearson 相关性系数由于其核心理论限制仅能反映节点的线性相关性而忽视非线性关系,且 Pearson 相关性系数要求序列严格对齐,具有较高的噪声敏感性,并不适应于数据缺失或含噪场景的预定义图构建,因此部分研究人员利用动态时间调整(Dynamic Time Warping, DTW)^[29,50]构建预定义相似矩阵,相较于 Pearson 相关性系数,DTW 通过动态规划对齐时间轴,相位与噪声敏感性较低,其相似性计算方法可被简化的表示为

$$a_{i,j} = \begin{cases} 1 & \text{DTW}(v_i, v_j) > \epsilon \\ 0 & \text{其他情况} \end{cases} \quad (7)$$

式中:DTW($\cdot$)为 DTW 函数; v_i 、 v_j 分别为第 i 、 j 个节点的交通数据组成的向量。

为解决 DTW 高计算复杂度问题,其效率改进版本 Fast-DTW^[54]方法也常常被用于序列相似性计算,通过粗化原始时间序列,利用尽可能少的数据表达同一曲线,Fast-DTW 将计算复杂度降至线性。

3.1.4 交互矩阵

相似性矩阵仅能反映连续交通流情况下的节点关联性,而无法反映诸如乘车需求或是自行车流量等具有迁移关系的节点交互特征^[91],不同位置的交互可从信息流角度表达二者的关联性^[92],对于移动性特征来说,2 个节点流量交互占比越高,其连接越为亲密,这种交互邻接矩阵构建方式可被表示为

$$a_{i,j} = \begin{cases} \frac{v_{P,i \rightarrow j}}{\sum_{m \in N_i} v_{P,i \rightarrow m}} & v_{P,i \rightarrow j} > 0 \\ 0 & \text{其他情况} \end{cases} \quad (8)$$

式中: $v_{P,i \rightarrow j}$ 为输入时间步长 P 内节点 i 到节点 j ($i \neq j$) 的转移状态; $v_{P,i \rightarrow m}$ 为时间跨度 P 内节点 i

到其他节点 m 的转移状态; N_i 为节点集。

3.1.5 其他预定义矩阵

除了上述常用预定义矩阵外,为适应特殊场景的预测任务,也有学者定义特殊的预定义矩阵表达节点空间关系,如 Shao 等^[93]通过定义时空训练数据近似直方图并结合欧氏距离构建如式(9)所示的启发式图

$$a_{i,j} = \begin{cases} \exp\left(-\frac{\|\omega_{i,j}\|^2}{\gamma^2}\right) & a_{i,j} \geq \epsilon \\ 0 & a_{i,j} < \epsilon \end{cases} \quad (9)$$

$$\omega_{i,j} = \sqrt{(\alpha_i - \alpha_j)^2 + (\beta_i - \beta_j)^2}$$

式中: $\omega_{i,j}$ 为节点 i 与节点 j ($i \neq j$) 拟合参数的欧氏距离; α_i, α_j 与 β_i, β_j 均为拟合参数,在计算中首先创建直方图表示时空数据概况,并利用函数 $f(x) = \alpha e^{-\beta x}$ 逼近直方图,以计算 α, β, x 为输入数据。

Liu 等^[94]在进行海运轨迹预测时定义了式(10)所示的尺寸相似性矩阵

$$a_{i,j} = \begin{cases} \frac{1}{2} & 0 < L_{i,j} < L_a \\ \frac{3}{4} & L_a \leq L_{i,j} \leq L_b \\ 1 & L_b < L_{i,j} < L_{\max} \end{cases} \quad (10)$$

式中: $L_{i,j}$ 为第 i 与 j 个节点(这里的节点为船舶)的长度差; $L_{\max}$ 为最大差; L_a, L_b 均为预定义的超参数。

Xu 等^[88]则在一个 3×3 网格中将目标区域与相邻区域连接以捕捉空间邻域特征。

3.2 动态时空图

基于时间的预定义矩阵一定程度上克服了基于空间关系的预定义矩阵无法完全表达节点关联性的缺陷,然而由于其计算特性,不管是相关性矩阵还是交互矩阵,均可能涉及到测试集信息泄露的问题,而且预定义矩阵并不具有泛化能力。此外,交通网络在宏观视角表现为长期静态性,在微观视角表现为短期动态性,单一静态时空图可能无法完全反映交通节点的相关关系,因此部分学者通过同时定义多个预定义邻接矩阵以期最大程度表达现实交通网络^[92-94]。然而已有研究表明交通预测的主要计算复杂度来自于空间算子^[57],即表达时间序列的时间张量与表达空间关系的邻接矩阵执行点积计算的过程,增加输入预定义矩阵的数量会大大提高模型训练对于算力的要求。为克服上述问题,部分研究人员通过模型的训练过程自适应生成最适合模型的动态时空图,根据是否依赖交通数据,本文将交通动态

时空图分为直接优化类动态图与特征优化类动态图。

3.2.1 直接优化类动态图

直接优化类动态图不依赖于交通数据,而是直接将可训练参数矩阵作为输入,这种方法起源于 Graph WaveNet^[56],并由于其简单有效的特性应用于多个交通预测模型^[95-96],该方法首先定义可训练嵌入矩阵 E_1, E_2 ,随后通过点积形式计算直接优化矩阵,该过程可被表示为

$$\tilde{A}_{\text{adp}} = \text{Softmax}[\text{ReLU}(E_1 \cdot E_2^T)] \quad (11)$$

式中: $\text{ReLU}(\cdot)$ 为负值归零函数; $\text{Softmax}(\cdot)$ 为归一化函数; $\tilde{A}_{\text{adp}}$ 为直接优化矩阵; $\cdot$ 为点积运算。

然而上述直接优化图虽然具有一定的自适应能力,实际测试却表明这种形式的动态图并不能摆脱对于预定义矩阵的依赖性,为摆脱这种依赖关系, Wu 等^[48]提出一种有向图学习方法,表示为

$$M_1 = \tanh(\mathcal{J}E_3 \cdot W_1) \quad (12)$$

$$M_2 = \tanh(\mathcal{J}E_4 \cdot W_2) \quad (13)$$

$$\tilde{A}_{\text{adp}} = \text{ReLU}\{\tanh[\mathcal{J}(M_1 \cdot M_2^T - M_2 \cdot M_1^T)]\} \quad (14)$$

式中: $\mathcal{J}$ 为控制激活函数饱和率的超参数; M_1, M_2 分别为上下游状态的可训练矩阵; $\tanh(\cdot)$ 为激活函数; W_1, W_2 分别为上下游可训练嵌入矩阵的可训练权重; E_3, E_4 分别为上下游空间嵌入矩阵。

Wu 等^[48]的方法假设交通预测中的变量变化是单向的,一个节点的变化会对下游节点产生影响,而下游节点则很难反向影响上游节点,并且实例研究表明这种方法并不依赖于预定义矩阵,单一自适应矩阵依然可以取得杰出的效果。

后续也有模型^[97-98]参考上述形式对式(11)进行改写,通过简单的源节点与目标节点嵌入生成稀疏的可训练动态邻接矩阵,如下所示

$$\tilde{A}_{\text{adp}} = \text{Softmax}[\text{ReLU}(E_1 \cdot E_2^T - E_2 \cdot E_1^T)] \quad (15)$$

部分研究者将记忆增强^[99]的概念引入自适应图构建方法中构建增强优化图^[58-59],提高了直接嵌入图对于输入特征的记忆能力,这一定程度克服了直接优化图无法捕捉空间异质性的问题,通过定义可训练嵌入矩阵 E 作为源节点库并结合可训练矩阵 W ,利用门控网络计算隐藏状态与查询记忆项的相似性进行构图,如下所示

$$\tilde{A}_{\text{adp}} = \text{Softmax}\{\text{ReLU}[(E \cdot W)(E \cdot W)^T]\} \quad (16)$$

3.2.2 特征优化类动态图

直接优化图直接嵌入可训练矩阵构建空间特征图是非常简便的,然而这种完全依赖模型训练的图

构建方法具有极强的不可解释性,且在部分情况可能学习出不符合物理世界规律的错误空间关系,因此部分研究通过对交通数据的特征学习构建最优节点属性动态图。基于相关性的特征优化图 $\hat{\mathbf{A}}_{\text{adp}}$ 通过对窗口长度的输入交通序列进行学习并以一定规则构建动态时空图,Cao 等^[55] 最先利用自注意力机制反映节点属性的相似关系,如式(17)所示,这是由于自注意力所计算的自注意力分数一定程度反映了节点的属性相似性,其数学意义可被简单理解为输入向量的点积,由于向量点积值越大,向量越相似,因此自注意力构图的过程可被理解为针对向量夹角余弦值 $\cos(\tau)$ 的计算过程,表示为

$$\hat{\mathbf{A}}_{\text{adp}} = \text{Softmax} \left[\frac{(\mathbf{W}_3 \cdot \mathbf{v}) \cdot (\mathbf{W}_4 \cdot \mathbf{v})^T}{\sqrt{n}} \right] \quad (17)$$

$(\mathbf{W}_3 \cdot \mathbf{v}) \cdot (\mathbf{W}_4 \cdot \mathbf{v})^T = |\mathbf{W}_3 \cdot \mathbf{v}| |\mathbf{W}_4 \cdot \mathbf{v}| \cos(\tau)$ (18)
式中: $\mathbf{W}_3$ 为可训练查询矩阵; $\mathbf{W}_4$ 为可训练键矩阵; τ 为向量夹角; n 为输入向量 $\mathbf{v}$ 的长度。

自注意力作为一种特征优化类动态图被应用于多项研究^[59,98,100],然而由于自注意力本身的缺陷,基于注意力的动态图存在噪声敏感、极化反应以及趋势忽略等多种问题,因此后续学者针对注意力图缺陷做出多项改进措施:Shin 等^[101] 提出一种基于注意力的余弦相似性图,在计算过程中将查询向量与键向量转化为单位向量构图以捕捉交通流量的趋势变化特征;Jin 等^[102] 通过在构图及图卷积过程添加ReLU函数消除负连接以保持自注意力图的稀疏性;吴永庆等^[103] 将自注意力与直接优化策略相结合提出一种新的特征优化图构建方法;Ta 等^[104-105] 则利用多层卷积代替单层线性投影以捕捉交通数据更深层的特征从而构建最优属性相似图。

事实上,除开自注意力与卷积 2 种构图方法,也有模型通过定义新的度量函数^[106] 衡量节点的相近关系,常见方法有余项相似度^[107-108]、径向基函数^[109] 或者组合深度学习网络^[110],此外,基于分布采用概率建模生成方法^[111-112] 也常被用于动态时空图构建,部分方法已被作者用于交通预测测试,部分方法尚未与下游交通预测任务相结合,这种融合测试可作为未来应用研究方向之一。

上述方法通过引入度量函数衡量节点属性之间的相似性关系,而除相关性外,时空数据在空间和时间分布模式上表现出不均匀性和复杂性^[113],这种特性被称为时空异质性,如图 7 所示。给定城市三大区域,假设图 7 右侧为三大区域某节点交通流量变化,则由图 7 可知:尽管区域间可能存在相似的峰值变化,但其交通流量变化可能存在不同,且同一节点的交通流量在不同时间也可能呈现出不同的交通模式;基于相似性的度量函数可以逐个节点衡量节点关联性,却无法衡量这种包含区域特性的异质性关系。

过去几年,不少工作以人工划分网格^[61,63] 或聚类算法^[114] 捕捉空间异质性,然而基于数据预处理的异质性构图方法虽然可以捕捉不同的区域模式,却也使得模型不得不以 2 阶段形式进行交通预测,同时预处理方法并不适用大场景或跨城市异质性检测,因此一些研究者将神经网络引入异质性处理中生成异质性特征优化图。Agarwal 等^[115] 利用双曲时间卷积提取时间特征并通过双曲超图聚合节点特征;侯越等^[116] 考虑异质性的时间特征与空间特征,提出时滞影响下的动态图反映当前时间步的异质性关系。

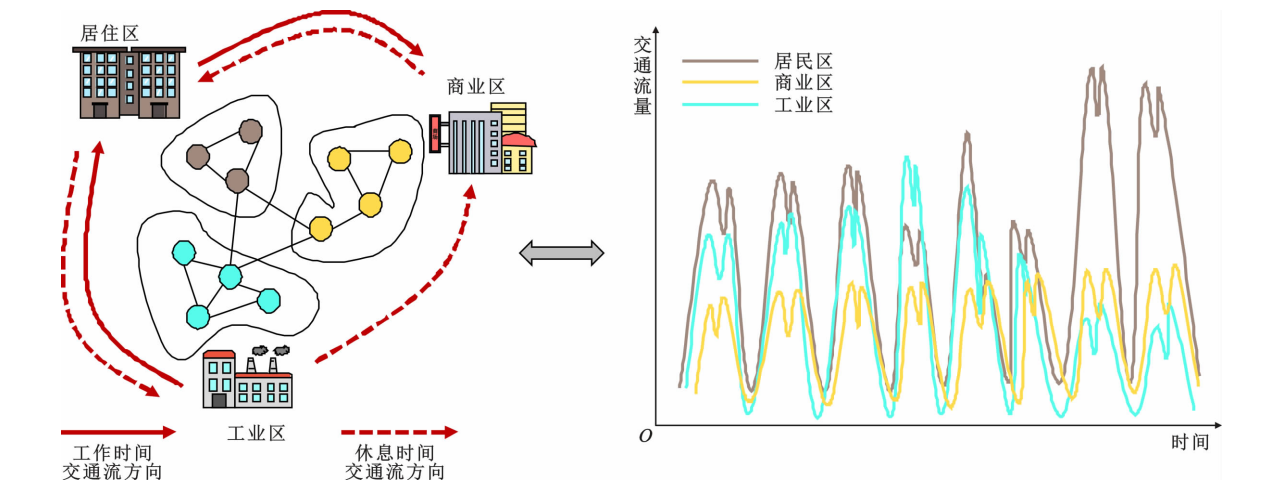


图 7 交通时空数据异质性
Fig. 7 Heterogeneity of traffic spatio-temporal data

从上述图构建方法的描述中不难看出,节点之间的连接关系并非单一的,2个节点可能同时存在相似性、异质性或更高阶的空间关系,通过建立多种空间图不但极大提高了计算的时间与空间复杂度,还难以全面地捕捉节点的高阶连接模式,因此多项研究利用超图捕捉更深层节点空间关联性,超图与普通图结构对比如图8所示。由图8(a)可知:普通图结构通过一条边连接2个节点以表示节点关系,而超图则可通过超边连接多个节点;在图8(b)案例中超图仅利用3条边就可以连接7个节点。因此,通过超图可更好地表示诸如地铁乘客不同出行模式及路径^[117]、交通节点多种模式关联^[118]等高阶时空关系,更好地反映现实世界节点之间的多样性空间关联性的同时可降低模型训练复杂程度。

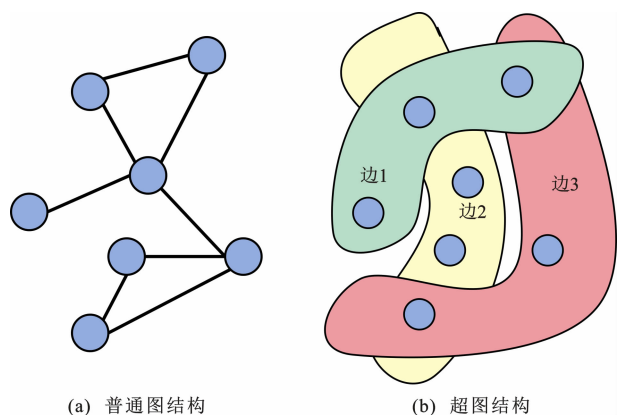


图8 超图与普通图结构对比

Fig. 8 Comparison between hypergraph and regular graph structures

3.3 小 结

现有研究从静态与动态入手开展时空图构建相关工作,取得了一定的成果。对于时空图神经网络而言,时空图往往即代表研究区域路网的空间特征。本章梳理分析了静态时空图到动态时空图的演进过程,确定了从包含直接空间关系的拓扑图与距离图,到依托时间特征的相似性图,进而到直接优化图、特征优化图等自适应动态图的发展过程。为最大程度地表达真实世界的路网拓扑关系,现有研究越来越倾向于更为复杂的图结构,这往往代表了更高的时间复杂度与空间复杂度,然而这种提升对于实际预测效果的影响是十分微弱的,如何衡量计算效率与预测性能将是未来图构建领域一大挑战。

4 时空特征建模方法综述

本节从图3所示的STGNN模型结构入手,以

时间特征建模方法、空间特征建模方法进行统计,同时对多个模型时空特征融合方法进行统计,梳理论文增加模型深度所用策略。

4.1 时间特征建模方法

针对传统时间特征建模方法(CNN、RNN、双向类模型^[119]等),近年来已有多个综述进行详细调研^[34,113],本文已在引言中进行简要介绍,此处不再过多介绍,仅对相对较新的方法进行分类梳理。

4.1.1 时间卷积网络(Temporal Convolutional Network, TCN)

传统CNN强调窗口空间关系,因此常被用于空间特征提取,其设计并未考虑时间维度上下文依赖性,因此在交通预测时可能导致信息丢失问题,而RNN结构的每个时间步输入依赖于上个时间步隐藏状态的输出,因此难以并行运算,这大大限制了模型在大规模交通数据上的预测能力。而TCN方法严格来说仍属于CNN方法,其基于CNN的特性提高了模型的并行计算能力,同时引入被广泛用于LSTM、GRU等模型的门控机制,通过门控控制信息流通过TCN^[120],该过程可被表示为

$$\mathbf{v}_{\text{out}} = \tanh[g_l(\mathbf{v}_{\text{in}})] \odot \text{Sigmoid}[f_l(\mathbf{v}_{\text{in}})] \quad (19)$$

式中: $g_l(\cdot)$ 为第 l 层TCN的门控分支; $f_l(\cdot)$ 为第 l 层TCN的滤波分支; $\mathbf{v}_{\text{in}}$ 、 $\mathbf{v}_{\text{out}}$ 分别为输入与输出的交通数据; $\odot$ 为哈达玛积符号。

考虑神经网络在传递过程中倾向于赋予重要信息较高的可训练参数权重,即重要信息输出值更大,其Sigmoid归一化更趋近1,因此通过门控,TCN控制了信息流通过率。为增加感受野,TCN往往引入因果膨胀策略,通过在标准因果卷积中加入膨胀因子扩大模型感受野,以 $g_l(\mathbf{v}_{t-1})$ 为例,其计算过程可表示为

$$g_l(\mathbf{v}_{t-1}) = \mathbf{v}_{t-1} * \boldsymbol{\delta}_l \quad (20)$$

式中: $*$ 为以 $\boldsymbol{\delta}_l$ 为卷积核的因果膨胀卷积; $\boldsymbol{\delta}_l$ 为第 l 层网络的卷积核; $\mathbf{v}_{t-1}$ 为第 $l-1$ 层网络输入交通数据组成的张量。

实际预测中通常交通预测设置膨胀因子 d 从1开始以2的指数次幂膨胀^[48]。然而扩散卷积不仅可以逐步扩张TCN的感受野,同时也可能带来TCN的信息丢失问题,这限制了TCN对于长期依赖性关系的记忆能力,迫使相关模型不得不增加TCN层数来获取全局信息,而TCN本身存在的冗余连接^[57]同样增加了模型的复杂程度。

4.1.2 Former类模型

事实上传统Transformer并不适合交通预测,

以 $\sin(x)$ 为例,利用 Python 绘制 $\sin(x)$, 交替 $\sin(x)$ [$\sin(x)$ 与 $\sin(3x)$ 交替], 尖峰 $\sin(x)$ [$\sin(x)$ 随机点添加高斯噪声]的时序图及注意力分数图,如图 9 所示。图 9(b)实线为添加噪声变化,虚线为源 $\sin(x)$ 。其中 $\sin(x)$ 与尖峰 $\sin(x)$ 由 $-10\pi \sim 10\pi$ 范围内随机 150 个相同间隔点生成,交替 $\sin(x)$ 以 $-5\pi \sim 5\pi$ 范围内随机 150 个相同间隔点生成,注意力分数计算方式为 Transformer 模型

去除可训练参数,选取第 50~100 个点,记为节点 0~50 以可视化注意分数。从图 9 可以看出:注意力更倾向于将数值聚焦于大值。这种现象为自注意力极化^[121],这种极化会导致 Former 类模型忽略季节性特征;极化效应甚至会导致模型聚焦于高斯噪声。而除极化效应导致的季节性特征忽视外^[122],Former 类模型还存在自注意力计算复杂度较高^[123]、token 输入方式存在缺陷^[124]等多种挑战。

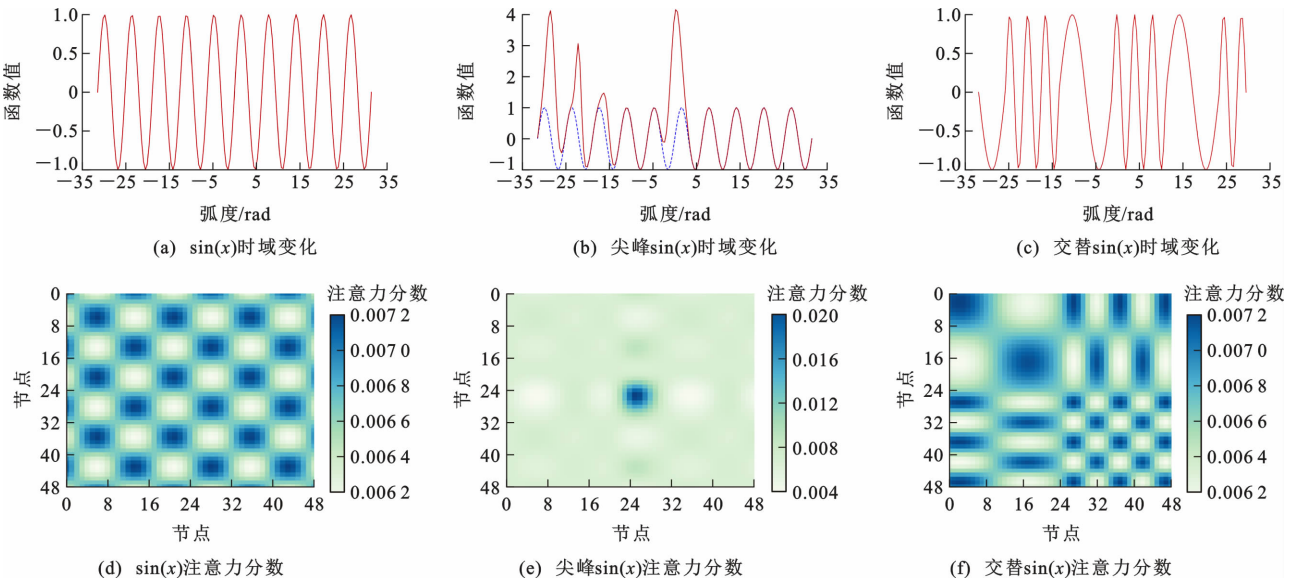


图 9 $\sin(x)$ 、尖峰 $\sin(x)$ 、交替 $\sin(x)$ 时域图像及注意力分数

Fig. 9 $\sin(x)$, peak $\sin(x)$, alternating $\sin(x)$ time-domain images and attention scores

为克服 Transformer 模型本身计算的问题,研究人员对模型本身做出了多种改进。Liu 等^[124]认为将多变量作为一个 Token 计算注意力抹除了变量相关性,进而提出一种新的 Token 输入方法;Liu 等^[125]通过滑动均值与噪声滤波器将序列分解为趋势项-季节项-余项,克服了自注意力计算中存在的噪声问题;Zhou 等^[126]提出了新型的稀疏注意力计算方法;Liu 等^[127]聚焦嵌入方式,提出了适合时空预测而非序列建模的新型序列嵌入方法。

4.1.3 门控线性单元(Gated Linear Unit, GLU)

RNN 类模型不断记忆过去信息而大大增加模型空间复杂度,同时每个状态需要输入上一个时间步的隐藏状态的特性限制了传统 RNN 结构的并行运算能力,因此 Dauphin 等^[128]认为在序列建模任务中无限的上下文记忆是非必要的,从而提出简单的线性门控单元并证明了这种方法已足够实现序列建模,这种线性单元输入严格按照时序位置保留信息从而提升了预测性能,并且通过并行预算加快了运算速度,其计算式为

$$\mathbf{v}_{\text{out}} = (\mathbf{W}_5 \cdot \mathbf{v}_{\text{in}} + \mathbf{b}_1) \odot \text{Sigmoid}(\mathbf{W}_6 \cdot \mathbf{v}_{\text{in}} + \mathbf{b}_2) \quad (21)$$

式中: $\mathbf{W}_5$ 为门控分支的可训练权重参数; $\mathbf{W}_6$ 为滤波分支的可训练权重参数; $\mathbf{b}_1$ 和 $\mathbf{b}_2$ 均为相应偏置项。

事实上交通预测下一个时间步的预测其实主要依赖于上个时间步信息,更远的时间特征对于交通预测并非如此重要^[57],因此,多项研究将 GLU 应用于交通预测领域^[53,55]。然而与 GLU 高效运算能力相对应的,是其序列记忆能力的减弱,这导致 GLU 难以承担长期预测任务。

4.1.4 Kolmogorov-Arnold 网络(Kolmogorov-Arnold Networks, KAN)

传统多层感知机 (Multi-layer Perceptron, MLP) 在处理高维数据时,由于数据分布变得稀疏而难以衡量数据的相似性,且计算成本随着数据维度的升高而迅速增大,而 KAN^[129]为 2024 年最新提出的上下文建模方法。针对 MLP 模型难以处理高维数据的问题,该模型将复杂高维函数分解为简单的一维函数组合,其与 MLP 模型主要架构区别如图 10 所示。从图 10 可以看出:基于通用近似定理的全连接 MLP 模型主要依赖于线性学习与非线性函数激活,大大提高了模型计算复杂性并且降低了

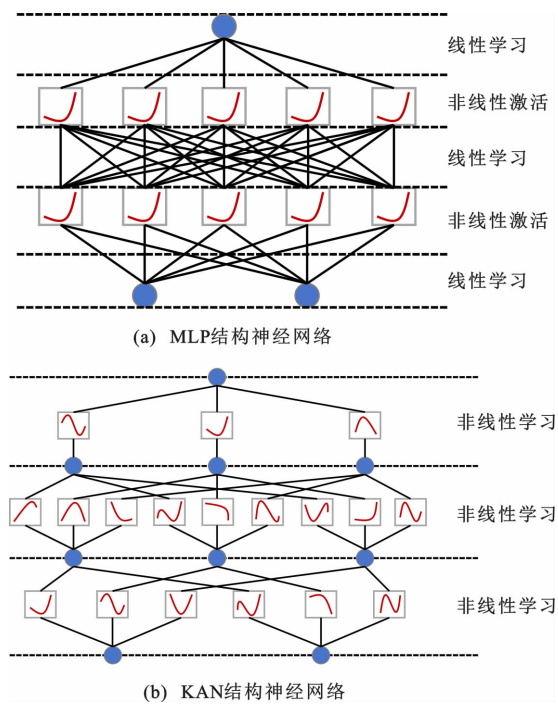


图 10 MLP 与 KAN 结构对比

Fig. 10 Comparison between MLP and KAN structures

模型可解释性;KAN 基于 Kolmogorov-Arnold 定理,每个多元连续函数都可以表示为单变量连续函数的 2 层嵌套叠加,提高了模型的可解释能力。KAN 的计算公式可表达为

$$f(\mathbf{v}) = \sum_{s=1}^{2n+1} \phi_s \left[\sum_{t=1}^n \varphi_{s,t}(v_t) \right] \tag{22}$$

式中: $\phi_s(\cdot)$ 为内部一元函数, s 为外部函数索引; $\varphi_{s,t}(v_t)$ 为外部的一元函数, v_t 为输入第 t 个数据。

通过利用可学习一维函数代替权重参数,并且将其参数化为 B 样条函数,KAN 降低了全连接 MLP 的计算复杂度,且作者在文中通过多个案例说明了 KAN 对于 MLP 的优势,然而小样本数据情况下,KAN 可能过拟合从而降低模型的泛化能力,且由于其专门化的性质,KAN 可能并不适用于现有的基于 GPU 的架构体系,虽然已有多项研究将其应用于时间特征建模^[130-132],KAN 在时空交通预测领域的应用还有待开发。

4.1.5 Mamba

Mamba 是 Gu 等^[133]提出的基于 S4 架构的新型序列建模模型,传统 RNN 架构虽然每个时间步均会输入之前所有时间步的隐藏信息,然而随着时间的推移,RNN 会逐渐遗忘较远时间步的信息,而在每个时间步输入之前隐藏状态的特性限制了 RNN 的并行计算能力,而 Mamba 则将所有信号输入压缩为系数向量,并通过状态空间模型(State

Space Model, SSM)与参数选择相互集成使得模型专注于相关输入,这使得模型拥有高效数据处理能力的同时能记忆更长的序列,提高了模型的长时建模能力。如图 11 所示,图中 SiLU 为激活函数。由图 11 可知:Mamba 首先利用线性投影与卷积生成特征,通过可选择变量重要性的 SSM 与生成特征输入状态表示输出。SSM 的离散化版本可被表示为

$$\tilde{\mathbf{h}}_t = \overline{\mathbf{E}}_5 \cdot \mathbf{h}_t + \overline{\mathbf{E}}_6 \cdot \mathbf{v} \tag{23}$$

$$\mathbf{O}_t = \mathbf{E}_7 \mathbf{h}_t \tag{24}$$

$$\overline{\mathbf{E}}_5 = \exp(\Delta \mathbf{E}_5) \tag{25}$$

$$\overline{\mathbf{E}}_6 = (\Delta \mathbf{E}_5)^{-1} [\exp(\Delta \mathbf{E}_5) - \mathbf{I}] \Delta \mathbf{E}_6 \tag{26}$$

$$\Delta = \text{softplus}\{\text{Broadcast}[\text{Linear}(\mathbf{I})] + \boldsymbol{\theta}\} \tag{27}$$

式中: $\mathbf{E}_5$ 、 $\mathbf{E}_6$ 为嵌入可训练矩阵; $\overline{\mathbf{E}}_5$ 、 $\overline{\mathbf{E}}_6$ 为嵌入矩阵的离散化形式;Linear($\cdot$) 为线性投影函数;Broadcast($\cdot$) 为通过广播机制将一维数据广播至高维空间的广播函数; $\boldsymbol{\theta}$ 为恒定权重矩阵; $\mathbf{I}$ 为单位矩阵; $\mathbf{h}_t$ 为 t 时刻的隐藏状态; $\mathbf{O}_t$ 为 t 时刻 SSM 的输出; $\tilde{\mathbf{h}}_t$ 为包含当前时刻与之前时刻合并的隐藏状态。

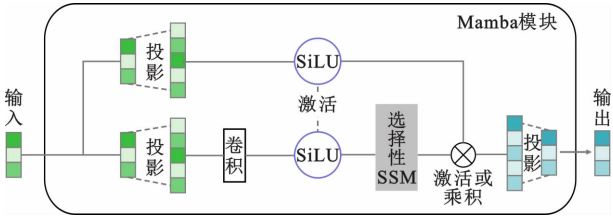


图 11 Mamba 模块

Fig. 11 Mamba module

作为具有近线性复杂度的序列模型,Mamba 效率与性能均优于 Transformer 模型,近年来,其在交通预测领域的应用逐渐加深。Lin 等^[134]提出双向图卷积 Mamba 模型,将 Mamba 与图卷积网络集成在一起,并使用双向 Mamba 架构捕获历史交通数据中的长期时间依赖关系;Mehravian 等^[135]提出了一个用于流量预测的自适应图 Mamba 模型;Shao 等^[136]从理论上证明了 Mamba 与 Transformer 的互补性并通过二者的融合提高了交通预测效率。

4.2 空间特征建模方法

图神经网络方法在处理非欧空间特征建模方面展现了卓越的效能。鉴于交通节点间存在的空间依赖性本质上具有非欧几里得特性,使得图神经网络(Graph Neural Network, GNN)非常适合被应用于交通网络的空间特征建模中。通过其独特的架构,GNN 能够有效建模复杂交通系统内部节点间的多层次关联与依赖关系,从而为交通流预测提供了强

有力的支持,根据其处理视角不同,GNN 通常被分为谱域图卷积、空域图卷积与图注意力网络。

4.2.1 谱域图卷积

谱域图卷积通常基于图傅里叶变换将信号转换到谱域中进行特征提取,该过程将空域空间的遍历过程转化为谱域空间的图滤波过程^[137],与离散傅里叶变换利用拉普拉斯算子作为特征函数相似,图傅里叶变换利用图特征向量矩阵 $\mathbf{U}$ 及其转置 $\mathbf{U}^T$ 实现空域与谱域傅里叶变换,表示为

$$F(\mathbf{v}) = \mathbf{U}^T \mathbf{v} \quad (28)$$

$$F^{-1}(\mathbf{v}) = \mathbf{U} \cdot \mathbf{v} \quad (29)$$

式中: $F(\mathbf{v})$ 为图傅里叶变换; $F^{-1}(\mathbf{v})$ 为图逆傅里叶变换。

卷积定理规定卷积的傅里叶变换等于函数傅里叶变换的乘积^[138],因此图卷积可被写为

$$\mathbf{U}\{\text{diag}[c(\lambda)]\}\mathbf{U}^T \mathbf{v} = \mathbf{U}\boldsymbol{\delta}\mathbf{U}^T \mathbf{v} \quad (30)$$

式中: $\boldsymbol{\delta} = \text{diag}[c(\lambda)]$ 为需要设计的图卷积核; λ 为特征值。

式(30)也可写为下述形式

$$F^{-1}[F(\boldsymbol{\delta}) \odot F(\mathbf{v})] = \mathbf{U}(\mathbf{U}^T \cdot \boldsymbol{\delta}) \odot (\mathbf{U}^T \mathbf{v}) \quad (31)$$

近年来,基于 Chebshev 多项式的截断扩展方法^[139]逐渐成为主流,以 ChebNet 为例,这类图卷积方法可被总结为

$$\mathbf{W}\star \mathbf{v} = \sum_{z=0}^Z \mathbf{W}_z \Psi_z(\tilde{\mathbf{L}}) \mathbf{v} \quad (32)$$

式中: $\tilde{\mathbf{L}}$ 为拉普拉斯矩阵 $\mathbf{L}$ 归一化表示,其计算为 $\tilde{\mathbf{L}} = 2\mathbf{L}/\lambda_{\max} - \mathbf{I}$; $\lambda_{\max}$ 为最大特征值; $\Psi_z(\tilde{\mathbf{L}})$ 为 z 阶截断 Chebshev 多项式; Z 为截断数; $\star$ 为图卷积操作。

截断 Chebshev 多项式方法在多个模型^[52-53,55]中得到了应用。此外,近几年也有学者对卷积核拟合多项式进行拓展,设计了以 Bernstein 多项式^[140]为代表的多个新型傅里叶图滤波器^[141]。

4.2.2 空域图卷积

谱域图卷积模型由于过于依赖拉普拉斯矩阵而不适合动态图场景,为克服这种依赖性,Kipf 等^[142]通过空间域图卷积实现节点消息传递^[143],其计算式如下

$$\mathbf{W}\star \mathbf{v} = \mathbf{W}(\mathbf{I} + \mathbf{D}^{-\frac{1}{2}}\mathbf{A}\mathbf{D}^{-\frac{1}{2}}) \mathbf{v} \quad (33)$$

式中: $\mathbf{D}$ 为邻接矩阵 $\mathbf{A}$ 的度矩阵。

与 Kipf 等^[142]相类似,Li 等^[27]将交通流动态建模为扩散过程,通过扩散卷积实现节点消息传递,同时考虑交通流的双向性,将图卷积的过程定义为

$$\mathbf{v}_{\text{out}} = \sum_{k=0}^K \mathbf{D}^{-1} \mathbf{A} \mathbf{v} \mathbf{W}_{7,k} + \sum_{k=0}^K \mathbf{D}^{-1} \mathbf{A}^T \mathbf{v} \mathbf{W}_{8,k} \quad (34)$$

式中: K 为转移步数; $\mathbf{W}_{7,k}$ 、 $\mathbf{W}_{8,k}$ 分别为第 k 步转移步数前向交通流与后向交通流的训练参数。

然而上述方法需要完整图输入的特性使其并不适合大场景空间特征建模,且无法泛化到训练过程中没有出现的节点,由此 Hamilton 等^[144]提出了一种大图采样学习模型,其过程可被表达为

$$\mathbf{h}_{N_i,t} \leftarrow \text{Aggregator}(\{\mathbf{h}_{q,t}, \forall q \in N_i\}) \quad (35)$$

$$\mathbf{h}_{i,t} \leftarrow \sigma[\mathbf{W}\text{Cat}(\mathbf{h}_{i,t-1} | \mathbf{h}_{N_i,t})] \quad (36)$$

式中:Aggregator($\cdot$)为定义的聚合函数; $\mathbf{h}_{N_i,t}$ 为节点 i 经过聚合操作后的嵌入形式; $\mathbf{h}_{i,t}$ 为第 i 个节点在 t 时刻的嵌入形式; σ 为可选择的激活函数;Cat($\cdot$)为叠加算子; $\mathbf{h}_{q,t}$ 为 t 时刻 i 的相邻节点 q 的隐藏状态。

在交通时空预测领域,也有学者利用常微分方程(Ordinary Differential Equation, ODE)方式^[145]实现时间与空间的连续快速建模或是利用变分自编码器^[146]的想法实现动态图不确定性特征建模。

4.2.3 图注意力网络(Graph Attention Network, GAT)

与自注意力想法类似,Velickovic 等^[147]提出了一种图上的注意力以加权聚合邻居节点信息并将其命名为图注意力网络,这种邻居信息加权聚合的操作可有效增强模型对噪声的鲁棒性,相较于传统图卷积,图注意力网络具有更强的表达能力,能够学习到节点直接复杂度交互关系,且通过注意力可以抑制无关邻居权重,区分关键节点,一定程度上缓解了过平滑的问题,其过程可被定义为

$$a_{i,u} = \frac{\exp\{\text{LeakyReLU}[\rho([\mathbf{W}\mathbf{h}_{i,t-1} \| \mathbf{W}\mathbf{h}_{u,t-1}])]\}}{\sum_{m \in N_i} \exp\{\text{LeakyReLU}[\rho([\mathbf{W}\mathbf{h}_{i,t-1} \| \mathbf{W}\mathbf{h}_{m,t-1}])]\}} \quad (37)$$

式中: $a_{i,u}$ 为节点 i 与相邻节点 u 的注意力关系; $[\cdot \| \cdot]$ 为拼接操作; $\rho(\cdot)$ 为将拼接后的高维特征映射为实数的函数; m 为针对邻居节点集 N_i 的索引;LeakyReLU($\cdot$)为 LeakyReLU 激活函数。

以注意力分数为权重聚合目标节点与邻接节点,其过程可被表示为

$$\mathbf{h}_{i,t+1} = \sigma\left(\sum_{u \in N_i} a_{i,u} \mathbf{W}\mathbf{h}_{u,t}\right) \quad (38)$$

图注意力网络通过计算所有邻居对注意力权重确定邻居重要性的特点大大增加了模型计算的复杂程度,因此 Zhang 等^[148]在传统 GAT 注意力的基础

上提出利用卷积子网络自适应分配多头注意力,利用键-值注意力机制与点积计算方式而非全连接计算注意力分数。

4.3 信息传播与时空融合策略

4.3.1 信息传播策略

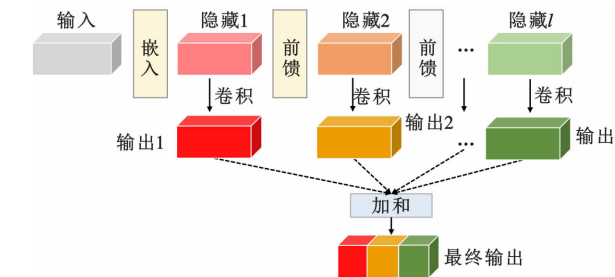
通常来说神经网络的层数越深,往往性能越好^[149],然而过深的神经网络可能存在过拟合、模型退化以及梯度爆炸的风险,因此需要引入层与层、模块与模块之间的信息传播策略,这种策略以残差连接^[150]最为常见,考虑残差连接应用的普遍程度,本文仅对几种典型并且有效的交通时空预测信息传播策略进行梳理分析。

(1) 跳跃连接

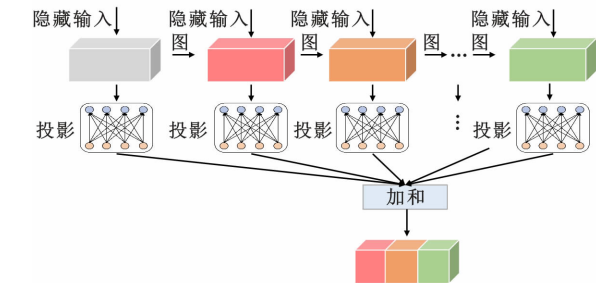
跳跃连接是一种常用的增加模型深度,减少模型退化的信息传播策略^[56,98,101],其框架如图 12(a)所示。由图 12(a)可知:跳跃连接在交通时空预测中的表现形式为通过卷积层控制不同时空层的输出,并通过加和形式集成每个时空层捕捉到的时空特征。

(2) 最后一击连接

最后一击连接为 Lai 等^[57]提出的跳跃连接压缩改进版本,其结构如图 12(b)所示。由图 12(b)可知:最后一击连接考虑下一步预测仅与最近时间步信息有关,将跳跃连接的卷积层改变为最后一个时间步的保留,从而减少与邻接矩阵点积过程的计算



(a) 跳跃连接



(c) Mix-hop传播

图 12 不同层数增加策略

Fig. 12 Strategies for increasing different layers

复杂度。这种策略在不丢失信息的同时,大大降低了模型训练复杂程度,并在实际测试中取得了优秀的结果。

(3) 混合跳跃(Mix-hop)传播

Mix-hop 传播策略为一种针对图卷积的信息保留以缓解过平滑问题的传播策略,最早由 Sami 等^[151-152]提出,这种思想在交通时空预测领域的应用主要体现在 Wu 等^[48]的模型中,其结构如图 12(c)所示,由图 12(c)可知:Mix-hop 连接将输入到图卷积模块前的信息与图卷积后的信息通过加权求和的方法输出结果,增强了模型对于误差的鲁棒性。关于 Mix-hop 传播可以表示为

$$h_{out} = \mu h_{in} + (1 - \mu) \tilde{A} h_{in} \tag{39}$$

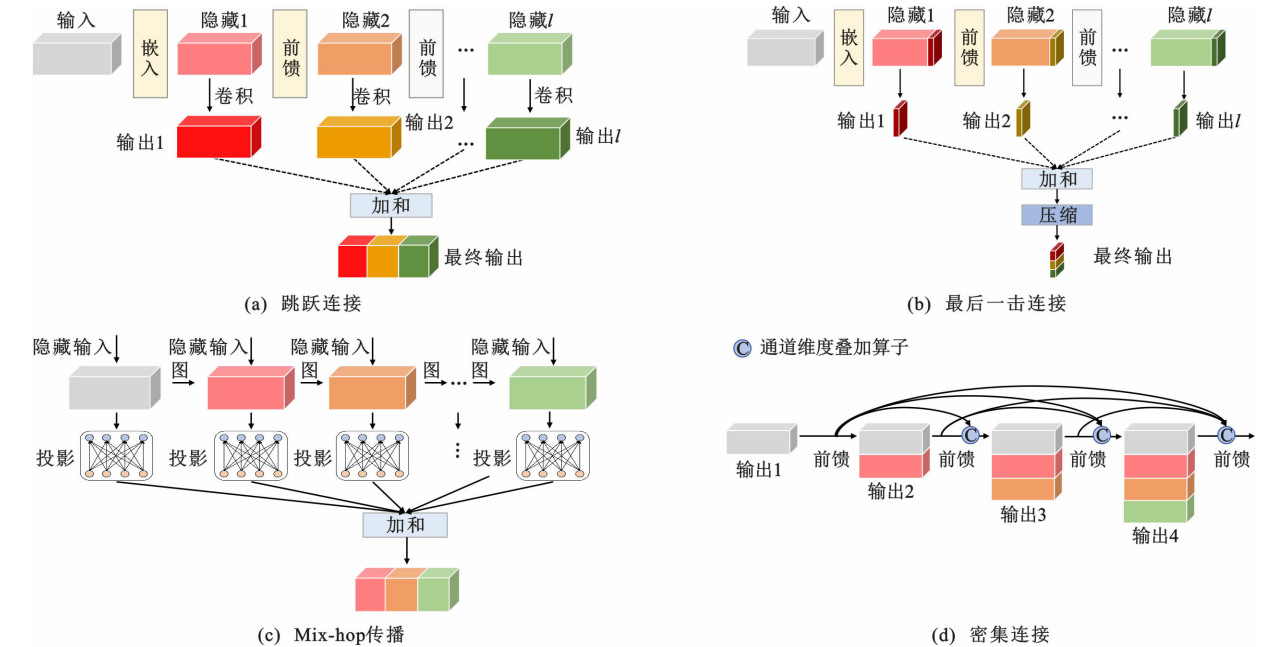
式中: μ 为控制源节点保留比例的超参数; h_{in} 为输入的隐藏状态张量; $\tilde{A}$ 为归一化后的邻接矩阵。

(4) 密集连接

密集连接是一种针对残差连接的改进连接方案^[153],这种连接由于其卓越的性能同时也被应用于交通时空预测领域^[154-156],其结构如图 12(d)所示,由图 12(d)可知:与残差连接不同,密集连接期望在第 l 层前馈时保留前 $l-1$ 层前馈的时空图以降低模型训练过程的误差传播与误差累积问题,并减轻反向传播过程中出现的梯度消失。

4.3.2 时空融合策略

除时间特征建模和空间特征建模外,如何有效



(d) 密集连接

将时空特征结合并获得最优解也是搭建交通时空预测模型的关键步骤之一。现有时空融合策略可分为堆叠类^[39,53,98]与耦合类^[27,157]。堆叠式以图13(a)的 Graph WaveNet 为例,耦合式由图 13(b)所示的 DCRNN^[27] 为例,其计算过程如下

$$r_t = \sigma(W_r \star [v_t, h_{t-1}] + b_r) \tag{40}$$

$$u_t = \sigma(W_u \star [v_t, h_{t-1}] + b_u) \tag{41}$$

$$C_t = \tanh\{W_C \star [v_t, (r_t \odot h_{t-1})] + b_C\} \tag{42}$$

$$h_t = u_t \odot h_{t-1} + (1 - u_t) \odot C_t \tag{43}$$

式中: r_t 为第 t 个时间步的重置门输出; u_t 为第 t 个时间步更新门输出; C_t 为第 t 个时间步候选隐藏状态输出; b_r 、 b_u 、 b_C 为 3 个状态的偏置项。

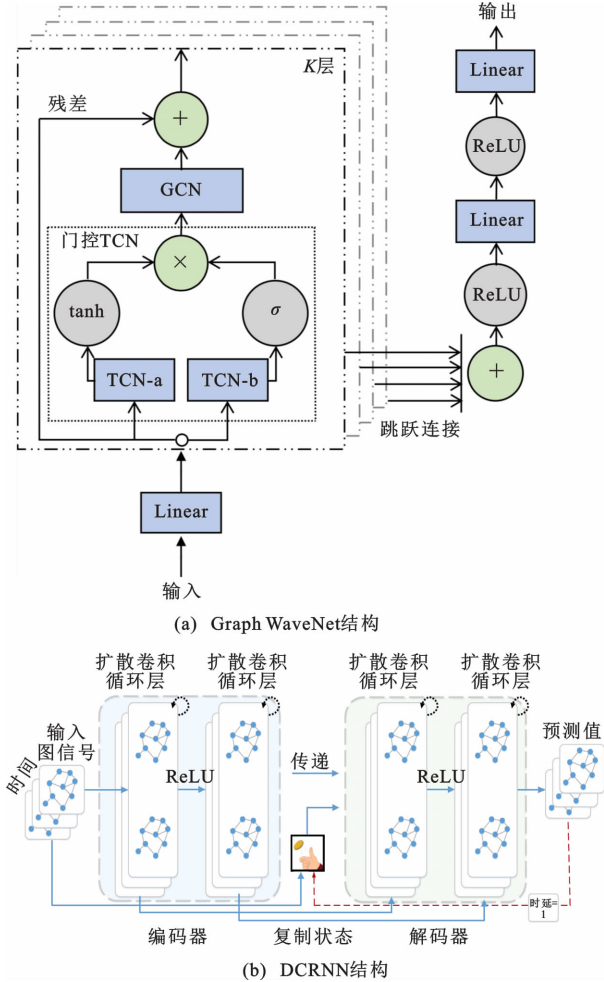


图 13 典型堆叠架构与耦合架构

Fig. 13 Typical stacked and coupled architectures

图 13(a)中 GCN 为图卷积,TCN 为门控时间卷积,ReLU 为激活函数,Linear 为线性前馈层。由图 13 可知:堆叠类通过将时间特征建模模块与空间特征建模模块通过并联或串联的形式堆叠在一起,这种结构以层堆叠的形式提取时间特征与空间特征,这种变化在 Graph WaveNet 中体现为门控

TCN 与扩散 GCN 的交替堆叠,信息流在整个模型中呈现线性传递的姿态;耦合式架构通常将空间特征建模模块嵌入时间特征建模模块,DCRNN 对 GRU 改动主要在于内部运算符,而 GRU 整体运算过程并没有发生改变。

4.4 经典交通时空预测模型对比

为更好对比经典时空预测模型的性能表现,本文以交通状态预测中的几个经典数据集为依托,对比了各年经典模型的架构设计以及其预测均方根误差(Root Mean Square Error, RMSE)、平均绝对误差(Mean Absolute Error, MAE)以及平均绝对百分比误差(Mean Absolute Percentage Error, MAPE),如表 4 所示,表中结果为 Metr-LA 与 PeMS-Bay 第 12 步预测及 PeMS04、PeMS08 的预测均值,最优效果加粗表示。

从表 4 可以看出:现有时间建模方法与空间建模方法在预测性能上并无严格优劣之分,通过合理的架构,任何方法都能得到最优预测效果;相较于 DCRNN 与 ASTGCN^[52],Graph WaveNet 首次引入了自适应图构建方法,因此该模型在所列举数据集上取得了近乎跨越式的提升;HimNet^[28]通过自适应图元参数学习提高了模型对于时空异质性的捕捉能力并取得了最优效果;GTS^[155]通过优化图分布学习概率图,并利用可微分采样的方式得到离散图结构;DDSTGCN^[154]引入了超图及超图卷积捕捉更高阶的空间关系;统计表中共列举了 3 种 Former 模型,相较于 ASTGCN,GMAN^[156]的改动主要在于模型结构的调整与图构建方法的改进,而 STAEformer^[127]则完全舍弃了图卷积,通过更改嵌入方法的手段取得了极为优秀的结果,这从侧面反映了 Former 模型在交通预测任务上的优势。

近年模型更倾向于 Former 模型而非扩散 TCN 或是 RNN 架构,其中一部分原因可能是由于 Transformer 在自然语言处理取得了巨大成功,而其序列建模的思想与交通预测极为相近,吸引了部分研究者的注意。而针对空间维度做注意力的 Transformer 天然能够建模空间特征(类似于特征优化图),这使得诸如 STAEformer 等 Former 模型仅通过更改嵌入方式即取得非常优秀的结果。相比之下,虽然基于扩散 TCN 的模型如 DDSTGCN 取得了相近于最优的效果,且 2019 年的 Graph WaveNet 在近 3 年依然具有竞争力,然而由于其扩散特性,卷积核大小与层数息息相关,造成在改变预

表 4 近年最先进模型预测效果统计

Table 4 Recent statistics on the prediction performance of state-of-the-art models

对比领域		DCRNN (2018 年)	ASTGCN (2019 年)	Graph Wave Net(2019 年)	GMAN (2020 年)	GTS (2021 年)	DDSTGCN (2022 年)	STAEformer (2023 年)	HimNet (2024 年)
时间建模		GRU	Former	TCN	Former	CNN	TCN	Former	GRU
空间建模		空域 GCN	谱域 GCN	空域 GCN	GAT	空域 GCN	空域 GCN	无	空域 GCN
融合策略		耦合式	堆叠式	堆叠式	堆叠式	堆叠式	堆叠式	堆叠式	耦合式
Metr-LA	RMSE	7.59	7.75	7.37	7.35	7.44	7.13	7.02	7.22
	MAE	3.60	3.83	3.53	3.44	3.59	3.44	3.34	3.37
	MAPE%	10.5	10.8	10.0	10.1	10.3	9.7	9.7	9.8
PeMS-Bay	RMSE	4.74	4.81	4.52	1.92	4.60	4.37	4.34	4.32
	MAE	2.07	2.15	1.95	4.49	2.06	1.89	1.88	1.84
	MAPE/%	4.9	5.1	4.6	4.5	4.9	4.5	4.4	4.3
PeMS04	RMSE	31.26	35.22	29.92	31.60	32.95	31.05	30.18	29.88
	MAE	19.63	22.93	18.53	19.14	20.96	19.61	18.22	18.14
	MAPE/%	13.6	16.6	12.9	13.2	14.7	13.7	12.0	12.0
PeMS08	RMSE	24.17	28.16	23.39	24.92	26.08	24.16	23.25	23.22
	MAE	15.22	18.61	14.40	15.31	16.49	15.30	13.46	13.57
	MAPE/%	10.2	13.1	9.2	10.1	10.5	10.5	8.9	9.0

测步长时需改变相关模型的 TCN 层数,这极大制约了模型的自适应性。此外,根据表 4 的统计,HimNet 取得了最优结果,与 DCRNN 类似,该模型利用 GCN 代替了 GRU 中的线性单元,这种架构具有极高的空间复杂度(HimNet 在 Metr-LA 数据集的内存占用约为 6 056 MB,GWN 则为 1 994 MB,几乎产生了 3 倍差距),而普通的 GRU 受限于线性投影而不具有并行运算的能力。

为更好地衡量模型预测的实时性,本文对上述模型的运行效率进行了整理,所有模型均在 NVIDIA GeForce RTX 5070 Ti Laptop GPU 上运行,相关模型均为 Pytorch 版本,具体运行时间如图 14 所示,图中训练时长代表一个 Epoch 的训练时长,从图 14 可以看出:现有高性能模型往往代表了低效率,这说明现有模型在一定程度上依靠参数的堆叠记忆数据并通过增加模型复杂程度提高模型泛化能力,这一定程度上制约了模型的工业级部署,考虑实际模型应用需求,未来相关模型研究的一大方向可能是在保持模型整体性能同时实现模型本身的轻量化。

4.5 先进交通时空预测方法综述

近年来,随着交通时空预测模型的不断发展,传统模型或模型组合已难以取得更优秀的效果,由此诞生了许多其他的交通预测方法,这些方法通过最先进的深度学习技术使得模型更好地捕捉输入序列

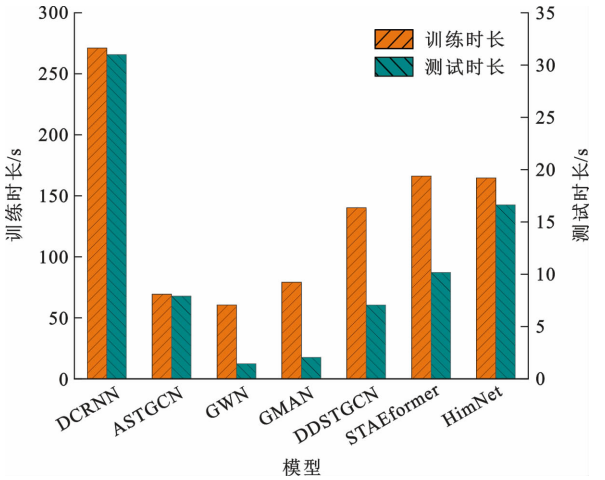


图 14 模型运行效率对比分析

Fig. 14 Comparative analysis of model running efficiency

的时空特征从而表达出较其他 STGNN 模型更好的效果,本节以对比学习、预训练、因果推理、混合专家模型以及多模态网络为研究对象,对几种典型交通时空预测模型进行总结梳理。

(1) 基于对比表征学习的交通时空预测

对比学习通过在嵌入空间构造正负样本,将锚点推离负样本的同时推进正样本,这种策略往往依赖于前期数据的特征增强,具体为添加噪声、翻转、平移、放缩等多种方式^[158]。将同一序列的不同增强视角视为正样本,而同一批次中的其他样本被视作负样本形成正负样本对比学习。这种方法被应用于多个模型中:例如 Zheng 等^[159]通过将高斯噪声

注入训练样本生成增强序列,同时引入 InfoNCE^[160] 损失实现原始序列解耦;Lan 等^[161] 提出一种分层语义波对比学习方法,通过全局与局部 2 种交互形式设计语义波表示并以无监督学习解决轨迹预测长尾挑战。

除序列对比预测外,图对比也是一种常见的交通时空预测对比学习方法^[162],通过对比学习策略从输入序列提取节点特征以构建最优动态图^[163]。图对比学习策略由于其优秀的性能吸引了众多学者关注:例如刘伟等^[164] 对动态图的边、节点以及属性进行遮挡与插值,并筛选负样本构建图对比学习;Jia 等^[165] 通过在节点视角和图视角 2 种视角执行随机图增强来生成 2 个不同的自适应增强图,以确保 2 种图编码嵌入的一致性为目标生成异质性结构感知图以捕捉节点的异质性关系。

(2) 基于预训练的交通时空预测

预训练是一种基于相似数据的预先训练策略,将训练好的模型作用于自己的数据集并对模型进行微调,这种训练策略可有效解决跨城市交通预测^[166]、小样本预测^[167] 以及缺失数据插补预测^[168] 等问题,此外,也有模型将长序列切片并进行掩码自动编码而非 2 阶段训练实现预训练交通预测^[169],通过预训练进行交通预测可有效减少模型训练时间,增加数据利用效率并降低模型训练成本从而提高 STGNN 类模型在交通时空预测领域的可伸缩性与可部署性。

(3) 基于因果推理的交通时空预测

因果关系指“因”事件与“果”事件的相互关系,在交通预测中,因果关系可以表示为外界事件(天气、事故以及人类活动等)与交通状态变化或车辆轨迹变化的关系,通过因果推理可有效将外部事件作用嵌入交通预测模型并提高模型对于事件攻击的鲁棒性^[170]。从另一方面,基于因果推理的模型往往具有更高的可解释性^[171]。此外,因果推理模型也常被应用于交通事故预测任务^[172]。

(4) 基于专家模型的交通时空预测

近期,DeepSeek R1 模型的爆火为公众对国产大模型的信心打了一针强心针,同时也引起了学术界对于 MoE 架构的广泛关注。MoE 模型不仅在生成式取得了巨大突破,基于 MoE 模型的交通时空预测方法也不断涌现^[173-174],这种方法克服了传统模型的数据特化,极大增加了模型的适应能力,吸引了众多研究人员的兴趣:Lee 等^[59] 从构图方法出发构建 3 种交通预测专家,通过选择最优专家提高模

型对不同数据集的泛化能力;Jiang 等^[175] 设计了周期模式与趋势模式 2 类专家并将专家与稀疏门控自适应图学习器级联以实现更为精确及可解释的交通拥堵预测。

4.6 小结

本节针对 STGNN 模型常用的时间建模方法、空间建模方法、时空融合策略以及信息传递策略进行了梳理分析,同时聚焦时空一体化预测,从经典模型对比及先进建模方法 2 方面总结了现有模型的预测性能及近年研究人员的研究方向。在时间建模方法方面,现有方法仍在长时记忆能力与并行运算能力之间抉择,尽管 Mamba 看起来为这种决策提供了一种可能的解决方案,然而 Mamba 虽然提高了模型长时记忆能力的同时具有并行运算能力,其计算精度仍需进一步提升。而在空间特征建模方面,目前无论是 GCN 还是 GAT,均并未克服过平滑问题,并且空间建模方法所带来的计算复杂度是极大的,例如 Graph WaveNet 有 75.23% 的 FLOPS 来自于其空间特征建模部分,而近年来的模型如 STAEformer、LightCTS 均说明空间特征建模在交通预测并非必要的,如何平衡 GCN 的性能与效率,是值得思考的问题。对于信息传播策略与时空融合策略方法,现有的传递与架构方法仍是非常初级的,如何在性能、效率与通用性之间找到平衡,是未来相关方法的一大挑战。

5 结语

5.1 未来发展方向总结

从前几节综述梳理中不难发现,未来交通时空预测仍以 STGNN 类模型为主,而近几年针对 STGNN 模型的研究仍集中于新模型结构与新构图方法,而最新的论文报告结果显示,改进模型结构对于模型预测效果的影响已愈发微弱,未来研究必将趋近于新数据、新任务、新指标,因此对未来几种可能的研究方向总结如下。

(1) 如何实现长时交通预测

近年来的预测模型仍以短期预测为主,绝大多数模型通过输入 12 步预测 12 步的形式验证模型的有效性,然而对于 Metr-LA、PeMS-Bay 等数据来说,12 步预测仅代表未来一个小时的交通数据,事实上这对于交通组织来说是远远不够的。近几年已有部分研究探索尝试超越 12 步的长时预测模型搭建技巧,然而交通预测由于复杂的节点分布而产生了远超其他类型预测任务的计算复杂度,如 Guo

等^[176]的研究表明,在 Seattle 数据集 48 步预测时,诸如 DCRNN 等复杂度较高的模型在 A4000 显卡的验证时间高达 123 s,且占用了 1 331 MB 的内存,而 MAPE 为 14.91%,传统意义上较为轻量化的模型如 Graph WaveNet 也耗时 32 s,占用 2 597 MB 内存,其 MAPE 为 15.62%,二者性能相较于 12 步预测 12 步时均出现了约 30% 的性能下降。如何在保持长时预测性能的同时降低模型的相应时间,是未来研究人员应当思考的问题。

(2) 如何利用多模态数据

随着包括无人机(Unmanned Aerial Vehicle, UAV)、ETC 门架、浮动车等多源数据不断应用于交通预测任务,如何处理同目标多源数据成为了交通研究的又一大研究方向。多模态预测模型能够更精确地捕捉不同交通方式间的相互作用及其对整体交通流的影响,现有研究方法多利用单一交通数据作为输入,而少数关于多模态的研究主要针对于多种交通模式^[177](自行车、出租车、私家车等)、多种数据形式^[178](流量、速度、占有率、事件等),然而将多数据格式(图像、文本、视频等)的模态融合应用于交通预测仍有待研究。

(3) 如何添加现实物理约束

近年来交通领域涌现大量融合物理约束的预测模型(如物理信息神经网络 PINN^[179]或是一些 ODE 方法^[49]),然而现有方法多采用优化损失函数或是在图卷积引入常微分方程考虑物理约束,关于物理方程与数据驱动的耦合多停留在浅层,而缺乏对于高阶物理约束的探索,如何将诸如流体模型或是动态边界条件作为物理约束嵌入预测模型也是未来探索方向之一。

(4) 如何增加模型可解释性

无论是传统优化图的预测模型,还是基于对比学习的预测模型,抑或是基于 MoE 的预测模型,均未克服神经网络存在的黑盒问题,这种黑盒特性大大增加了模型勘误的难度,因此可解释性一直是深度学习主要研究方向之一,通过提高交通预测模型的可解释性不仅可以确定哪些时空特征对预测结果产生正面效益,也可以降低交通预测大模型的修正成本,根据目前的研究成果,基于因果推理^[180]或基于 KAN 的交通预测模型可能是潜在的可解释性研究方向,未来可通过构建可解释的时空权重分配机制,揭示外部因素或特殊时段对整体预测的贡献程度。

(5) 有没有新的模型训练方式

诸如对比学习^[159]、课程学习^[181]、联邦学习^[182]、迁移学习^[183]等多种学习形式已被应用于交通时空预测模型,通过正负样本形成对比、由简到繁逐步学习、多方参与联合学习、相似任务知识迁移等多种形式降低模型训练难度,提高模型预测结果并增加模型可部署性,未来新的学习策略必将成为新型交通时空预测模型的研究方向之一。

(6) 如何架构交通预测大模型

大语言模型不仅是交通预测领域未来发展目标,也是实现 ITS 的关键一步。相较于小模型,大语言模型具有更多的可训练参数,更优秀的数据记忆能力以及更好的预测效果。主流交通预测大模型诸如 TF-LLM^[184]、UniST^[185]等模型在多种交通预测场景取得了优秀的效果并展现演出零样本泛化能力,然而现有大模型仍存在计算成本较高、时空数据多样性有限、大模型“幻觉”以及可解释性较低等问题^[186],因此未来关于交通预测大模型的研究可能需要注重引入新的模型架构,同时提高模型的可解释性与多任务通用能力。

根据中国《交通强国建设纲要》不难看出,未来 ITS 将越来越趋近于智能化、安全化,同时推动交通大数据一体化的应用,而美国的《智能交通系统(ITS)战略规划 2020~2025》则包括新兴和使能技术、网络安全、数据访问和交换、自动驾驶、完整出行-ITS4US、加速 ITS 部署,以推动 ITS 技术的全生命周期发展等 ITS 六大规划领域。对于基于时空图神经网络的交通预测而言,从 2017 年 GCN 诞生开始,最早的交通预测模型诸如 DCRNN^[27]或是 STGCN^[53]通过简单组合时空特征建模方法取得了超越交通预测的效果,而后研究者们将目标聚焦于图表达能力以求最大程度地反映现实节点的空间关系^[55-56],抑或是引入物理约束方法约束空间建模过程^[49],而随着相关模型进一步发展,相关领域学者开始意识到单一增加模型复杂度以追求更优秀性能的发展路线可能并不利于模型的工业化部署,由此衍生出了轻量化模型^[57]或是改进学习策略^[159,181]的预测模型,而随着 DeepSeek 的爆火,越来越多企业和高校开始相关架构大模型的本地部署,而图 4(a)所示的重要 AI 模型或算法出现时间节点表明:从 2017 年 Transformer 出现至今,AI 领域主要研究成果集中在大模型的架构、微调与部署。基于时空图神经网络的交通预测作为 AI 领域的分支,其发展与 AI 大方向发展基本一致,同时考虑诸如无人机、浮动车、ETC 等数据采集的手段的不断发展与应用

可以推断,未来交通预测领域或许也将集中于多模态、低成本的大模型发展,通过集成架构模型,实现实时的交通预测。

5.2 面临的问题与挑战

随着交通预测的发展,各种新型的交通预测模型层出不穷,尽管近年来的 STGNN 模型性能提升越来越高,在未来发展过程中仍可能存在以下问题需要进行系统性探索。

(1)多模态数据对齐问题

事实上阻碍多模态数据发展的因素不仅仅源于模型本身,更多的来自于外部数据的限制,由于数据采集传感器的限制,所获取视频、图像、交通状态、外部事件文本未必来源于同一节点的同一段时间,而同时架设部署多种传感器的代价是十分昂贵的,在获取数据后的处理又是十分繁琐的。

(2)数据大范围缺失问题

据统计,国内外交通局高速公路数据丢失比例可达到 16%~93%,加拿大亚伯罕部分时间数据缺失高达 90%,美国明尼苏达交通数据也存在 40% 以上的数据缺失问题^[187]。尽管有部分研究开始探索数据补全算法^[188],其数据可用性以及极端情况数据的补全精度仍有待探讨。

(3)交通预测下游衔接问题

现有交通预测算法更倾向于预测本身的精度与效率问题,而忽视了预测与下游应用问题的衔接,交通预测服务于下游任务而同时受到下游任务的反作用影响,将交通预测任务与下游任务(交通管控、自动驾驶)相互衔接是未来预测模型需要思考的一大挑战。

尽管未来的交通预测任务仍面临的巨大的挑战,其前景仍是非常宽广的,如何探索新的模型发展、实现模型工业部署、克服发展问题挑战值得每一个相关领域研究人员思考。

参考文献:

References:

[1] 冯 晓,陈思龙.改善城市道路机动车排放污染的智能交通手段[J]. 交通运输工程学报,2002,2(2):73-77.
FENG Xiao, CHEN Si-long. An ITS method to decrease motor-vehicle pollution in urban area[J]. Journal of Traffic and Transportation Engineering, 2002, 2(2): 73-77.

[2] FENG K R, LIN N. Reconstructing and analyzing the traffic flow during evacuation in Hurricane Irma (2017) [J]. Transportation Research Part D: Transport and Environment, 2021, 94: 102788.

[3] 臧 华,彭国雄.高速道路异常状况下车辆排队长度的预测模

型[J]. 交通与计算机,2003,21(3):10-12.

ZANG Hua, PENG Guo-xiong. A forecast model of queuing length in expressway emergency[J]. Computer and Communications, 2003, 21(3): 10-12.

[4] LARTEY J D. Predicting traffic congestion: A queuing perspective[J]. Open Journal of Modelling and Simulation, 2014, 2(2): 57-66.

[5] 王永泉,陈花玲,毛文雄.基于跟驰理论的道路交通噪声预测模型研究[J]. 系统仿真学报,2004,16(11):2413-2416.
WANG Yong-quan, CHEN Hua-ling, MAO Wen-xiong. Research on road traffic noise prediction model based on car-following theory [J]. Acta Simulata Systematica Sinica, 2004, 16(11): 2413-2416.

[6] DAGANZO C F. The cell transmission model: A dynamic representation of highway traffic consistent with the hydrodynamic theory[J]. Transportation Research Part B: Methodological, 1994, 28(4): 269-287.

[7] XU X, ZHANG L L, KONG Q, et al. Enhanced-historical average for long-term prediction [C] // IEEE. 2022 2nd International Conference on Computer, Control and Robotics (ICCCR). New York: IEEE, 2022: 115-119.

[8] HAMED M M, AL-MASAEID H R, SAID Z M B. Short-term prediction of traffic volume in urban arterials [J]. Journal of Transportation Engineering, 1995, 121(3): 249-254.

[9] SHAHRIARI S, GHASRI M, SISSON S A, et al. Ensemble of ARIMA: Combining parametric and bootstrapping technique for traffic flow prediction[J]. Transportmetrica A: Transport Science, 2020, 16(3): 1552-1573.

[10] CORTES C, VAPNIK V. Support-vector networks[J]. Machine Learning, 1995, 20(3): 273-297.

[11] BREIMAN L. Random forests[J]. Machine Learning, 2001, 45(1): 5-32.

[12] COVER T, HART P. Nearest neighbor pattern classification [J]. IEEE Transactions on Information Theory, 1967, 13 (1): 21-27.

[13] WU C H, HO J M, LEE D T. Travel-time prediction with support vector regression [J]. IEEE Transactions on Intelligent Transportation Systems, 2004, 5(4): 276-281.

[14] RUMELHART D E, HINTON G E, WILLIAMS R J. Learning representations by back-propagating errors [J]. Nature, 1986, 323: 533-536.

[15] BAO X X, JIANG D, YANG X F, et al. An improved deep belief network for traffic prediction considering weather factors[J]. Alexandria Engineering Journal, 2021, 60(1): 413-420.

[16] ZHANG W B, YU Y H, QI Y, et al. Short-term traffic flow prediction based on spatio-temporal analysis and CNN deep learning[J]. Transportmetrica A: Transport Science, 2019, 15(2): 1688-1711.

[17] WILLIAMS R J, ZIPSER D. A learning algorithm for continually running fully recurrent neural networks [J]. Neural Computation, 1989, 1(2): 270-280.

[18] HOCHREITER S, SCHMIDHUBER J. Long short-term memory[J]. Neural Computation, 1997, 9(8): 1735-1780.

- [19] CHO K, VAN MERRIENBOER B, GULCEHRE C, et al. Learning phrase representations using RNN encoder-decoder for statistical machine translation[C]//ACL. Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP). Doha: ACL, 2014: 1724-1734.
- [20] ASHISH V, NOAM S, NIKI P, et al. Attention is all you need[C]//NeurIPS Foundation. Conference on Neural Information Processing Systems. Cambridge: MIT Press, 2017: 6000-6010.
- [21] 马 飞, 杨治杰, 王江博, 等. 基于气象-交通多通道数据融合的短时交通流速度预测模型[J]. 交通运输工程学报, 2024, 24(6): 183-196.
- MA Fei, YANG Zhi-jie, WANG Jiang-bo, et al. Short-term traffic flow speed prediction model based on meteorological-traffic multi-channel data fusion[J]. Journal of Traffic and Transportation Engineering, 2024, 24(6): 183-196.
- [22] 胡立伟, 侯 智, 赵雪亭, 等. 基于交通事故文本挖掘的高速公路行车风险预测模型改进研究[J/OL]. 西南交通大学学报, 2025, <http://kns.cnki.net/kcms/detail/51.1277.U.20240509.1523.010.html>.
- HU Li-wei, HOU Zhi, ZHAO Xue-ting, et al. Research on improvement of highway traffic risk prediction model based on traffic accident text mining[J/OL]. Journal of Southwest Jiaotong University, 2025, <http://kns.cnki.net/kcms/detail/51.1277.U.20240509.1523.010.html>.
- [23] MA C X, ZHAO Y P, DAI G W, et al. A novel STFSA-CNN-GRU hybrid model for short-term traffic speed prediction[J]. IEEE Transactions on Intelligent Transportation Systems, 2023, 24(4): 3728-3737.
- [24] BEECHE C, SINGH J P, LEADER J K, et al. Super U-Net: A modularized generalizable architecture[J]. Pattern Recognition, 2022, 128: 108669.
- [25] ZHANG Z B, WU S, JIANG D W, et al. BERT-JAM: Maximizing the utilization of BERT for neural machine translation[J]. Neurocomputing, 2021, 460: 84-94.
- [26] 杨国亮, 习 浩, 龚家仁, 等. 基于 Transformer 的短时交通流时空预测[J]. 计算机应用与软件, 2024, 41(3): 169-173, 225.
- YANG Guo-liang, XI Hao, GONG Jia-ren, et al. Short term traffic flow forecasting based on transformer[J]. Computer Applications and Software, 2024, 41(3): 169-173, 225.
- [27] LI Y G, YU R, CYRUS S, et al. Diffusion convolutional recurrent neural network: Data-driven traffic forecasting [C]//YOSHUA B, YANN L. International Conference on Learning Representations. Portland: OpenReview. net, 2018: 1-16.
- [28] DONG Z, JIANG R H, GAO H T, et al. Heterogeneity-informed meta-parameter learning for spatiotemporal time series forecasting[C]//ACM. Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining. New York: ACM, 2024: 631-641.
- [29] LI H, LIU J, HAN S Y, et al. STFGCN: Spatial-temporal fusion graph convolutional network for traffic prediction[J]. Expert Systems with Applications, 2024, 255: 124648.
- [30] WANG S Z, CAO J N, YU P S. Deep learning for spatio-temporal data mining: A survey[J]. IEEE Transactions on Knowledge and Data Engineering, 2022, 34(8): 3681-3700.
- [31] 崔建勋, 要 甲, 赵泊媛. 基于深度学习的短期交通流预测方法综述[J]. 交通运输工程学报, 2024, 24(2): 50-64.
- CUI Jian-xun, YAO Jia, ZHAO Bo-yuan. Review on short-term traffic flow prediction methods based on deep learning [J]. Journal of Traffic and Transportation Engineering, 2024, 24(2): 50-64.
- [32] 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[J]. 软件学报, 2025, 36(4): 1811-1843.
- ZOU Hui-qi, SHI Bin-ze, SONG Ling-yun, et al. Survey on complex spatio-temporal data mining methods based on graph neural networks[J]. Journal of Software, 2025, 36(4): 1811-1843.
- [33] 卢佐安, 邓锦程, 韩金丽, 等. 图神经网络在交通预测中的应用综述[J]. 交通运输工程学报, 2023, 23(5): 39-61.
- HU Zuo-an, DENG Jin-cheng, HAN Jin-li, et al. Review on application of graph neural network in traffic prediction[J]. Journal of Traffic and Transportation Engineering, 2023, 23(5): 39-61.
- [34] 熊章友, 李卫军, 朱晓娟, 等. 基于深度学习的短时交通流预测研究综述[J]. 计算机工程与应用, 2025, 61(11): 67-82.
- XIONG Zhang-you, LI Wei-jun, ZHU Xiao-juan, et al. Short-term traffic flow prediction based on deep learning[J]. Computer Engineering and Applications, 2025, 61(11): 67-82.
- [35] LIU M Z, ZHU T Y, YE J C, et al. Spatio-temporal autoencoder for traffic flow prediction[J]. IEEE Transactions on Intelligent Transportation Systems, 2023, 24(5): 5516-5526.
- [36] ABDELRAOUF A, ABDEL-ATY M, YUAN J H. Utilizing attention-based multi-encoder-decoder neural networks for freeway traffic speed prediction[J]. IEEE Transactions on Intelligent Transportation Systems, 2022, 23(8): 11960-11969.
- [37] WU X L, ZHANG D L, GUO C J, et al. AutoCTS: Automated correlated time series forecasting [J]. Proceedings of the VLDB Endowment, 2022, 15(4): 971-983.
- [38] ZOU G J, LAI Z L, MA C X, et al. When will we arrive a novel multi-task spatio-temporal attention network based on individual preference for estimating travel time[J]. IEEE Transactions on Intelligent Transportation Systems, 2023, 24(10): 11438-11452.
- [39] WANG B B, LIN Y F, GUO S N, et al. GSNet: Learning spatial-temporal correlations from geographical and semantic aspects for traffic accident risk forecasting [C]//AAAI. Proceedings of the AAAI Conference on Artificial Intelligence. Washington DC: AAAI Press, 2021: 4402-4409.
- [40] WANG Z N, JIANG R H, XUE H, et al. Event-aware multimodal mobility nowcasting[C]//AAAI. Proceedings of the AAAI Conference on Artificial Intelligence. Washington DC: AAAI Press, 2022: 4228-4236.
- [41] 贾兴利, 李双庆, 杨宏志, 等. 基于 ATT-LSTM 模型的高速公路交通事故持续时长预测[J]. 交通信息与安全, 2022, 40(5): 61-69.

- JIA Xing-li, LI Shuang-qing, YANG Hong-zhi, et al. Prediction of the duration of freeway traffic incidents based on an ATT-LSTM model[J]. *Journal of Transport Information and Safe*, 2022, 40(5): 61-69.
- [42] YE J C, SUN L L, DU B W, et al. Coupled layer-wise graph convolution for transportation demand prediction [C] // AAAI. Proceedings of the AAAI Conference on Artificial Intelligence. Washington DC: AAAI Press, 2021: 4616-4625.
- [43] WANG S Y, ZHUGE C X, SHAO C F, et al. Short-term electric vehicle charging demand prediction: A deep learning approach[J]. *Applied Energy*, 2023, 340: 121032.
- [44] LIANG Y B, DING F Y, HUANG G, et al. Deep trip generation with graph neural networks for bike sharing system expansion [J]. *Transportation Research Part C: Emerging Technologies*, 2023, 154: 104241.
- [45] HUANG Y J, DU J T, YANG Z R, et al. A survey on trajectory-prediction methods for autonomous driving [J]. *IEEE Transactions on Intelligent Vehicles*, 2022, 7(3): 652-674.
- [46] CHEN G X, HU L, ZHANG Q S, et al. ST-LSTM: Spatio-temporal graph based long short-term memory network for vehicle trajectory prediction[C] // IEEE. 2020 IEEE International Conference on Image Processing (ICIP). New York: IEEE, 2020: 608-612.
- [47] LI F X, YAN H, JIN G Y, et al. Automated spatio-temporal synchronous modeling with multiple graphs for traffic prediction[C] // ACM. Proceedings of the 31st ACM International Conference on Information & Knowledge Management. New York: ACM, 2022: 1084-1093.
- [48] WU Z H, PAN S R, LONG G D, et al. Connecting the dots: Multivariate time series forecasting with graph neural networks[C] // ACM. Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. New York: ACM, 2020: 753-763.
- [49] FANG Z, LONG Q Q, SONG G J, et al. Spatial-temporal graph ODE networks for traffic flow forecasting[C] // ACM. Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining. New York: ACM, 2021: 364-373.
- [50] SONG C, LIN Y F, GUO S N, et al. Spatial-temporal synchronous graph convolutional networks: A new framework for spatial-temporal network data forecasting [C] // AAAI. Proceedings of the AAAI Conference on Artificial Intelligence. Washington DC: AAAI Press, 2020: 914-921.
- [51] CHEN Y Z, DOMINGUEZ I S, GEL Y R. Z-GCNets: Time zigzags at graph convolutional networks for time series forecasting[C] // ACM. International Conference on Machine Learning. New York: ACM, 2021: 1684-1694.
- [52] GUO S N, LIN Y F, FENG N, et al. Attention based spatial-temporal graph convolutional networks for traffic flow forecasting[C] // AAAI. Proceedings of the AAAI Conference on Artificial Intelligence. Washington DC: AAAI Press, 2019: 922-929.
- [53] YU B, YIN H T, ZHU Z X. Spatio-temporal graph convolutional networks: A deep learning framework for traffic forecasting [C] // AAAI. Proceedings of the 27th International Joint Conference on Artificial Intelligence. Washington DC: AAAI Press, 2018: 3634-3640.
- [54] 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[J]. *软件学报*, 2024, 35(4): 1751-1773.
- ZHAO Wen-zhu, YUAN Guan, ZHANG Yan-mei, et al. Multi-view fused spatial-temporal dynamic GCN for urban traffic flow prediction[J]. *Journal of Software*, 2024, 35(4): 1751-1773.
- [55] CAO D F, WANG Y J, DUAN J Y, et al. Spectral temporal graph neural network for multivariate time-series forecasting [C] // NeurIPS Foundation. Conference on Neural Information Processing Systems. Cambridge: MIT Press, 2020: 17766-17778.
- [56] WU Z H, PAN S R, LONG G D, et al. Graph WaveNet for deep spatial-temporal graph modeling [C] // AAAI. Proceedings of the 28th International Joint Conference on Artificial Intelligence. Washington DC: AAAI Press, 2019: 1907-1913.
- [57] LAI Z C, ZHANG D L, LI H, et al. LightCTS*: Light-weight correlated time series forecasting enhanced with model distillation[J]. *IEEE Transactions on Knowledge and Data Engineering*, 2024, 36(12): 8695-8710.
- [58] JIANG R H, WANG Z N, YONG J W, et al. Spatio-temporal meta-graph learning for traffic forecasting [C] // AAAI. Proceedings of the AAAI Conference on Artificial Intelligence. Washington DC: AAAI Press, 2023: 8078-8086.
- [59] LEE H, KO S. TESTAM: A time-enhanced spatio-temporal attention model with mixture of experts[C] // YOSHUA B, YANN L. International Conference on Learning Representations. Portland: OpenReview.net, 2024: 1-19.
- [60] LIU D C, WANG J, SHANG S, et al. MSDR: Multi-step dependency relation networks for spatial temporal forecasting [C] // ACM. Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining. New York: ACM, 2022: 1042-1050.
- [61] ZHANG J B, ZHENG Y, QI D K. Deep spatio-temporal residual networks for citywide crowd flows prediction[C] // AAAI. AAAI Conference on Artificial Intelligence. Washington DC: AAAI Press, 2017: 1-7.
- [62] LI X Y, XU Y, CHEN Q, et al. Short-term forecast of bicycle usage in bike sharing systems: A spatial-temporal memory network [J]. *IEEE Transactions on Intelligent Transportation Systems*, 2022, 23(8): 10923-10934.
- [63] ZHOU Z Y, WANG Y, XIE X K, et al. RiskOracle: A minute-level citywide traffic accident forecasting framework [C] // AAAI. Proceedings of the AAAI Conference on Artificial Intelligence. Washington DC: AAAI Press, 2020: 1258-1265.
- [64] JIN G Y, WANG M, ZHANG J L, et al. STGNN-TTE: Travel time estimation via spatial-temporal graph neural network[J]. *Future Generation Computer Systems*, 2022, 126: 70-81.

- [65] JIANG J W, PAN D Y, REN H X, et al. Self-supervised trajectory representation learning with temporal regularities and travel semantics[C]// IEEE. 2023 IEEE 39th International Conference on Data Engineering (ICDE). New York; IEEE, 2023: 843-855.
- [66] YUAN J, ZHENG Y, XIE X, et al. Driving with knowledge from the physical world[C]// ACM. Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. New York; ACM, 2011: 316-324.
- [67] YUAN J, ZHENG Y, ZHANG C Y, et al. T-drive: Driving directions based on taxi trajectories[C]// ACM. Proceedings of the 18th SIGSPATIAL International Conference on Advances in Geographic Information Systems. New York; ACM, 2010: 99-108.
- [68] YAN B Q, ZHAO G, SONG L X, et al. PreCLN: Pretrained-based contrastive learning network for vehicle trajectory prediction[J]. World Wide Web, 2023, 26(4): 1853-1875.
- [69] LIAO H C, LI X L, LI Y K, et al. Characterized diffusion and spatial-temporal interaction network for trajectory prediction in autonomous driving[C]// AAAI. International Joint Conference on Artificial Intelligence. Washington DC: AAAI Press, 2024: 1-9.
- [70] CHEN X B, ZHANG H J, ZHAO F, et al. Intention-aware vehicle trajectory prediction based on spatial-temporal dynamic attention network for Internet of vehicles[J]. IEEE Transactions on Intelligent Transportation Systems, 2022, 23(10): 19471-19483.
- [71] KRAJEWSKI R, BOCK J, KLOEKER L, et al. The highD dataset: A drone dataset of naturalistic vehicle trajectories on German highways for validation of highly automated driving systems[C]// IEEE. 2018 21st International Conference on Intelligent Transportation Systems (ITSC). New York; IEEE, 2018: 2118-2125.
- [72] CHANG M-F, LAMBERT J, SANGKLOY P, et al. Argoverse: 3D tracking and forecasting with rich maps[C]// IEEE. 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). New York; IEEE, 2019: 8740-8749.
- [73] LIANG M, YANG B, HU R, et al. Learning lane graph representations for motion forecasting [C] // Springer. Proceedings of 16th European Conference on Computer Vision (ECCV). Munich; Springer, 2020: 541-556.
- [74] 袁 静, 夏 英. 基于时空图注意力卷积神经网络的车辆轨迹预测[J]. 计算机科学, 2024, 51(12): 157-165.
YUAN Jing, XIA Ying. Vehicle trajectory prediction based on spatial-temporal graph attention convolutional network [J]. Computer Science, 2024, 51(12): 157-165.
- [75] CHANG Y C, QI J Z, LIANG Y X, et al. Contrastive trajectory similarity learning with dual-feature attention[C]// IEEE. 2023 IEEE 39th International Conference on Data Engineering (ICDE). New York; IEEE, 2023: 2933-2945.
- [76] 付相君. 基于路网拓扑结构的相关路段流量预测研究[D]. 重庆: 重庆交通大学, 2020.
FU Xiang-jun. Research on traffic forecasting of related sections based on topological structure of road network[D]. Chongqing: Chongqing Jiaotong University, 2020.
- [77] 温振国. 基于复杂网络理论的陕西省高速公路网结构特性及鲁棒性研究[D]. 西安: 长安大学, 2019.
WEN Zhen-guo. Study on structural characteristics and robustness of Shaanxi expressway network based on complex network theory[D]. Xi'an: Chang'an University, 2019.
- [78] 陈 希, 钱江海, 韩定定. L 空间和 P 空间模型下的树状结构网络[J]. 计算机应用研究, 2015, 32(1): 45-47.
CHEN Xi, QIAN Jiang-hai, HAN Ding-ding. Tree network under space L and space P model[J]. Application Research of Computers, 2015, 32(1): 45-47.
- [79] WANG X Y, MA Y, WANG Y Q, et al. Traffic flow prediction via spatial temporal graph neural network [C] // ACM. Proceedings of The Web Conference 2020. New York; ACM, 2020: 1082-1092.
- [80] GAO A, ZHENG L J, WANG Z X, et al. Attention based short-term metro passenger flow prediction [M]. Munich: Springer International Publishing, 2021.
- [81] GUO K, HU Y L, QIAN Z, et al. Optimized graph convolution recurrent neural network for traffic prediction [J]. IEEE Transactions on Intelligent Transportation Systems, 2021, 22(2): 1138-1149.
- [82] 孙 俊, 潘玉君, 和瑞芳, 等. 地理学第一定律之争及其对地理学理论建设的启示[J]. 地理研究, 2012, 31(10): 1749-1763.
SUN Jun, PAN Yu-jun, HE Rui-fang, et al. The enlightenment of geographical theories construction from the first law of geography and its debates[J]. Geographical Research, 2012, 31(10): 1749-1763.
- [83] 谭会生, 杨 威, 严舒琪. 交通速度预测时空图卷积网络及其 FPGA 实现研究[J]. 电子测量技术, 2024, 47(18): 108-119.
TAN Hui-sheng, YANG Wei, YAN Shu-qi. Research on spatio-temporal graph convolutional network for traffic speed prediction and their FPGA implementation [J]. Electronic Measurement Technology, 2024, 47(18): 108-119.
- [84] ZHAO L, SONG Y J, ZHANG C, et al. T-GCN: A temporal graph convolutional network for traffic prediction[J]. IEEE Transactions on Intelligent Transportation Systems, 2020, 21(9): 3848-3858.
- [85] 祁 朵, 毛政元. 基于自适应时序剖分与 KNN 的短时交通流量预测[J]. 地球信息科学学报, 2022, 24(2): 339-351.
QI Duo, MAO Zheng-yuan. Short-term traffic flow prediction based on adaptive time slice and KNN[J]. Journal of Geo-information Science, 2022, 24(2): 339-351.
- [86] LV M Q, HONG Z X, CHEN L, et al. Temporal multi-graph convolutional network for traffic flow prediction[J]. IEEE Transactions on Intelligent Transportation Systems, 2021, 22(6): 3337-3348.
- [87] LI P, WANG S, ZHAO H T, et al. IG-net: An interaction graph network model for metro passenger flow forecasting [J]. IEEE Transactions on Intelligent Transportation Systems, 2023, 24(4): 4147-4157.
- [88] XU G, LI Y G, WANG L Y, et al. Spatiotemporal multi-graph convolution network for ride-hailing demand forecasting

- [C] // AAAI. Proceedings of the AAAI Conference on Artificial Intelligence. Washington DC: AAAI Press, 2019; 3656-3663.
- [89] LIU L B, CHEN J W, WU H F, et al. Physical-virtual collaboration modeling for intra- and inter-station metro ridership prediction[J]. IEEE Transactions on Intelligent Transportation Systems, 2022, 23(4): 3377-3391.
- [90] 包银鑫,曹 阳,施 隼. 基于改进时空残差卷积神经网络的都市路网短时交通流预测[J]. 计算机应用, 2022, 42(1): 258-264.
- BAO Yin-xin, CAO Yang, SHI Quan. Improved spatio-temporal residual convolutional neural network for urban road network short-term traffic flow prediction[J]. Journal of Computer Applications, 2022, 42(1): 258-264.
- [91] CHAI D, WANG L Y, YANG Q. Bike flow prediction with multi-graph convolutional networks[C] // ACM. Proceedings of the 26th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems. New York: ACM, 2018; 397-400.
- [92] 濮稳稳. 多类别交通参与者轨迹预测算法研究[D]. 长沙: 湖南大学, 2021.
- PU Wen-wen. Research on trajectory prediction algorithm of multi-class traffic participants [D]. Changsha: Hunan University, 2021.
- [93] SHAO W, JIN Z L, WANG S, et al. Long-term spatio-temporal forecasting via dynamic multiple-graph attention[C] // AAAI. Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence. Washington DC: AAAI Press, 2022; 1-8.
- [94] LIU R W, LIANG M H, NIE J T, et al. STMGCN: Mobile edge computing-empowered vessel trajectory prediction using spatio-temporal multigraph convolutional network[J]. IEEE Transactions on Industrial Informatics, 2022, 18(11): 7977-7987.
- [95] LUO X L, ZHU C J, ZHANG D T, et al. Dynamic graph convolutional network with attention fusion for traffic flow prediction[C] // Springer. Proceedings of 26th European Conference on Artificial Intelligence (ECAI). Munich: Springer, 2023; 1-8.
- [96] MA Y, LOU H J, YAN M, et al. Spatio-temporal fusion graph convolutional network for traffic flow forecasting[J]. Information Fusion, 2024, 104: 102196.
- [97] ZHANG W, ZHU F H, LV Y S, et al. AdapGL: An adaptive graph learning algorithm for traffic prediction based on spatiotemporal neural networks[J]. Transportation Research Part C: Emerging Technologies, 2022, 139: 103659.
- [98] QU Y H, JIA X L, GUO J H, et al. MSSTGNN: Multi-scaled spatio-temporal graph neural networks for short- and long-term traffic prediction[J]. Knowledge-based Systems, 2024, 306: 112716.
- [99] CHAUHAN V K, ZHOU J D, LU P, et al. A brief review of hypernetworks in deep learning[J]. Artificial Intelligence Review, 2024, 57(9): 250.
- [100] PENG Y F, GUO Y Y, HAO R, et al. Network traffic prediction with attention-based spatial-temporal graph network[J]. Computer Networks, 2024, 243: 110296.
- [101] SHIN Y, YOON Y. PGCN: Progressive graph convolutional networks for spatial-temporal traffic forecasting[J]. IEEE Transactions on Intelligent Transportation Systems, 2024, 25(7): 7633-7644.
- [102] JIN G Y, LIU C X, XI Z X, et al. Adaptive dual-view WaveNet for urban spatial-temporal event prediction[J]. Information Sciences, 2022, 588: 315-330.
- [103] 吴永庆,姜正宇. 基于解耦动态时空卷积循环网络的交通流预测[J/OL]. 计算机工程, 2025, DOI:10.19678/j.issn.1000-3428.0070319.
- WU Yong-qing, JIANG Zheng-yu. Traffic flow prediction based on decoupled dynamic spatial-temporal convolutional recurrent network [J/OL]. Computer Engineering, 2025, DOI: 10.19678/j.issn.1000-3428.0070319.
- [104] TA X X, LIU Z H, HU X, et al. Adaptive spatio-temporal graph neural network for traffic forecasting[J]. Knowledge-Based Systems, 2022, 242: 108199.
- [105] LV Z Q, WANG X T, CHENG Z S, et al. ST-TDCN: A two-channel tree-structure spatial-temporal convolutional network model for traffic velocity prediction[J]. Expert Systems with Applications, 2024, 257: 125053.
- [106] WANG F, SUN J M. Survey on distance metric learning and dimensionality reduction in data mining[J]. Data Mining and Knowledge Discovery, 2015, 29(2): 534-564.
- [107] WOJKE N, BEWLEY A. Deep cosine metric learning for person re-identification [C] // IEEE. 2018 IEEE Winter Conference on Applications of Computer Vision (WACV). New York: IEEE, 2018; 748-756.
- [108] CHEN Y, WU L F, ZAKI M J. Iterative deep graph learning for graph neural networks: Better and robust node embeddings[C] // NeurIPS Foundation. Proceedings of 34th Conference on Neural Information Processing Systems. Cambridge: MIT Press, 2020; 19314-19326.
- [109] LI R Y, WANG S, ZHU F Y, et al. Adaptive graph convolutional neural networks[C] // AAAI. Proceedings of the AAAI Conference on Artificial Intelligence. Washington DC: AAAI Press, 2018; 3546-3553.
- [110] LIU H Y, YANG D H, LIU X Z, et al. TodyNet: Temporal dynamic graph neural network for multivariate time series classification[J]. Information Sciences, 2024, 677: 120914.
- [111] LUCA F, MATHIAS N, MASSIMILIANO P, et al. Learning discrete structures for graph neural networks[C] // ACM. Proceedings of 36th International Conference on Machine Learning. New York: ACM, 2019; 1972-1982.
- [112] ZHENG C, ZONG B, CHENG W, et al. Robust graph representation learning via neural sparsification[C] // ACM. Proceedings of 37th International Conference on Machine Learning. New York: ACM, 2020; 11458-11468.
- [113] JIN G Y, LIANG Y X, FANG Y C, et al. Spatio-temporal graph neural networks for predictive learning in urban computing: A survey[J]. IEEE Transactions on Knowledge and Data Engineering, 2024, 36(10): 5388-5408.
- [114] GUO K, HU Y L, SUN Y F, et al. Hierarchical graph convolution network for traffic forecasting [C] // AAAI.

- Proceedings of the AAAI Conference on Artificial Intelligence. Washington DC: AAAI Press, 2021: 151-159.
- [115] AGARWAL S, SAWHNEY R, THAKKAR M, et al. THINK: Temporal hypergraph hyperbolic network[C] // IEEE. 2022 IEEE International Conference on Data Mining (ICDM). New York: IEEE, 2022: 849-854.
- [116] 侯 越, 张 鑫, 裘著涛, 等. 特殊路网拓扑解构下的时空异质化交通流预测[J/OL]. 铁道科学与工程学报, 2025, 22(7): 2932-2945.
- HOU Yue, ZHANG Xin, XI Zhu-tao, et al. Spatio-temporal heterogeneous traffic flow prediction under special road network topology deconstruction[J/OL]. Journal of Railway Science and Engineering, 2025, 22(7): 2932-2945.
- [117] WANG J C, ZHANG Y, WEI Y, et al. Metro passenger flow prediction via dynamic hypergraph convolution networks [J]. IEEE Transactions on Intelligent Transportation Systems, 2021, 22(12): 7891-7903.
- [118] ZHAO Y S, LUO X, JU W, et al. Dynamic hypergraph structure learning for traffic flow forecasting [C] // IEEE. 2023 IEEE 39th International Conference on Data Engineering (ICDE). New York: IEEE, 2023: 2303-2316.
- [119] SCHUSTER M, PALIWAL K K. Bidirectional recurrent neural networks [J]. IEEE Transactions on Signal Processing, 1997, 45(11): 2673-2681.
- [120] LEA C, VIDAL R, REITER A, et al. Temporal convolutional networks: A unified approach to action segmentation[C] // Springer. Proceedings of 12th European Conference on Computer Vision (ECCV). Munich: Springer, 2016: 47-54.
- [121] ZHANG X Y, JIN X Y, GOPALSWAMY K, et al. First detrend then attend: Rethinking attention for time-series forecasting[C] // NeurIPS Foundation. Conference on Neural Information Processing Systems. Cambridge: MIT Press, 2022: 1-11.
- [122] ZHOU T, MA Z Q, WEN Q S, et al. FEDformer: Frequency enhanced decomposed transformer for long-term series forecasting[C] // ACM. International Conference on Machine Learning. New York: ACM, 2022: 27268-27286.
- [123] ZHOU H Y, ZHANG S H, PENG J Q, et al. Informer: Beyond efficient transformer for long sequence time-series forecasting[C] // AAAI. Proceedings of the AAAI Conference on Artificial Intelligence. Washington DC: AAAI Press, 2021: 11106-11115.
- [124] LIU Y, TU T G, ZHANG H R, et al. Transformer: Inverted transformers are effective for time series forecasting[C] // YOSHUA B, YANN L. International Conference on Learning Representations. Portland: OpenReview.net, 2024: 1-25.
- [125] LIU Y M, WU X, TANG Y, et al. Decomposition with feature attention and graph convolution network for traffic forecasting [J]. Knowledge-based Systems, 2024, 300: 112193.
- [126] ZHOU H Y, ZHANG S H, PENG J Q, et al. Informer: Beyond efficient transformer for long sequence time-series forecasting[C] // AAAI. Proceedings of the AAAI Conference on Artificial Intelligence. Washington DC: AAAI, 2021: 11106-11115.
- [127] LIU H C, DONG Z, JIANG R H, et al. Spatio-temporal adaptive embedding makes vanilla transformer SOTA for traffic forecasting[C] // ACM. Proceedings of the 32nd ACM International Conference on Information and Knowledge Management. New York: ACM, 2023: 4125-4129.
- [128] DAUPHIN Y N, FAN A, AULI M, et al. Language modeling with gated convolution networks[C] // ACM. International Conference on Machine Learning. New York: ACM, 2017: 933-941.
- [129] LIU Z M, WANG Y X, VAIDYA S, et al. KAN: Kolmogorov-arnold network[C] // YUE Y. International Conference on Learning Representations. Amherst: OpenReview.net, 2025: 1-50.
- [130] GAO Y, HU Z H, CHEN W A, et al. A revolutionary neural network architecture with interpretability and flexibility based on Kolmogorov-Arnold for solar radiation and temperature forecasting[J]. Applied Energy, 2025, 378: 124844.
- [131] LIVIERIS I E. C-KAN: A new approach for integrating convolutional layers with Kolmogorov-Arnold networks for time-series forecasting [J]. Mathematics, 2024, 12(19): 3022.
- [132] LIU X F, YANG Z L, GUO Y J, et al. A novel correlation feature self-assigned Kolmogorov-Arnold networks for multi-energy load forecasting in integrated energy systems [J]. Energy Conversion and Management, 2025, 325: 119388.
- [133] GU A, DAO T. Mamba: Linear-time sequence modeling with selective state spaces[C] // YEJIN C, DANNY Z. Proceedings of 1st Conference on Language Modeling. Amherst: OpenReview.net, 2024: 1-36.
- [134] LIN W X, 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.
- [135] MEHRABIAN A, WONG V W S. A-gamba: An adaptive graph-mamba model for traffic prediction in wireless cellular networks[J]. IEEE Wireless Communications Letters, 2025, 14(6): 1801-1805.
- [136] SHAO Z Q, WANG Z, YAO X S, et al. ST-MambaSync: Complement the power of Mamba and Transformer fusion for less computational cost in spatial-temporal traffic forecasting [J]. Information Fusion, 2025, 117: 102872.
- [137] ZHOU J, CUI G Q, HU S D, et al. Graph neural networks: A review of methods and applications[J]. AI Open, 2020, 1: 57-81.
- [138] 向 仪, 冯 强. 傅里叶正、余弦变换的加权卷积及其应用 [J]. 浙江大学学报(理学版), 2023, 50(3): 266-272.
- XIANG Yi, FENG Qiang. Weighted convolution of the Fourier sine-cosine transform and its application[J]. Journal of Zhejiang University (Science Edition), 2023, 50(3): 266-272.
- [139] HAMMOND D K, VANDERGHEYNST P, GRIBONVAL R. Wavelets on graphs via spectral graph theory[J]. Applied and Computational Harmonic Analysis, 2011, 30(2): 129-150.
- [140] HE M G, WEI Z W, HUANG Z F, et al. BernNet: Learning

- arbitrary graph spectral filters via Bernstein approximation [C] // NeurIPS Foundation. Conference on Neural Information Processing Systems. Cambridge; MIT Press, 2021; 1-16.
- [141] CHIEN E L, PENG J H, LI P, et al. Adaptive universal generalized PageRank graph neural network[C] // YOSHUA B, YANN L. International Conference on Learning Representations. Amherst; OpenReview.net, 2021; 1-24.
- [142] KIPF T N, WELING M. Semi-supervised classification with graph convolutional networks[C] // YOSHUA B, YANN L. International Conference on Learning Representations. Amherst; OpenReview.net, 2017; 1-14.
- [143] GILMER J, SCHOENHOLZ S S, RILEY P F, et al. Neural message passing for quantum chemistry [C] // ACM. International Conference on Machine Learning. New York; ACM, 2017; 1263-1272.
- [144] HAMILTON W L, YING R, LESKOVEC J. Inductive representation learning on large graphs [C] // NeurIPS Foundation. Conference on Neural Information Processing Systems. Cambridge; MIT Press, 2017; 1025-1035.
- [145] LOUIS-PASCAL A C X, QU M, TANG J. Continuous graph neural networks [C] // ACM. International Conference on Machine Learning. New York; ACM, 2020; 1-15.
- [146] KIPF T N, WELING M. Variational graph auto-encoders [C] // NeurIPS Foundation. Conference on Neural Information Processing Systems. Cambridge; MIT Press, 2016; 1-3.
- [147] VELICKOVIC P, CUCURULL G, CASANOVA A, et al. Graph attention networks [C] // YOSHUA B, YANN L. International Conference on Learning Representations. Amherst; OpenReview.net, 2018; 1-12.
- [148] ZHANG J N, SHI X J, XIE J Y, et al. GaAN: Gated attention networks for learning on large and spatio-temporal graphs [C] // AUAI. Conference on Uncertainty in Artificial Intelligence. Berkeley; AUAI, 2018; 339-349.
- [149] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks [J]. Communications of the ACM, 2017, 60(6): 84-90.
- [150] HE K M, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition [C] // IEEE. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). New York; IEEE, 2016; 770-778.
- [151] SAMI A E H, BRYAN P, AMOL K, et al. MixHop: Higher-order graph convolutional architectures via sparsified neighborhood mixing [C] // ACM. International Conference on Machine Learning. New York; ACM, 2019; 21-29.
- [152] CHEN F W, PAN S R, JIANG J, et al. DAGCN: Dual attention graph convolutional networks [C] // IEEE. 2019 International Joint Conference on Neural Networks (IJCNN). New York; IEEE, 2019; 1-8.
- [153] HUANG G, LIU Z, VAN DER MAATEN L, et al. Densely connected convolutional networks [C] // IEEE. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). New York; IEEE, 2017; 2261-2269.
- [154] SUN Y F, JIANG X H, HU Y L, et al. Dual dynamic spatial-temporal graph convolution network for traffic prediction [J]. IEEE Transactions on Intelligent Transportation Systems, 2022, 23(12): 23680-23693.
- [155] SHANG C, CHEN J, BI J B. Discrete graph structure learning for forecasting multiple time series [C] // YOSHUA B, YANN L. International Conference on Learning Representations. Amherst; OpenReview.net, 2021; 1-14.
- [156] ZHENG C P, FAN X L, WANG C, et al. GMAN: A graph multi-attention network for traffic prediction [C] // AAAI. Proceedings of the AAAI Conference on Artificial Intelligence. Washington DC; AAAI Press, 2020; 1234-1241.
- [157] ZHANG Z C, LIN X, LI M, et al. A customized deep learning approach to integrate network-scale online traffic data imputation and prediction [J]. Transportation Research Part C: Emerging Technologies, 2021, 132: 103372.
- [158] WEN Q S, SUN L, YANG F, et al. Time series data augmentation for deep learning: A survey [C] // AAAI. Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence. Washington DC; AAAI Press, 2021; 1-8.
- [159] ZHENG X, BAGLOEE S A, SARVI M. TRECK: Long-term traffic forecasting with contrastive representation learning [J]. IEEE Transactions on Intelligent Transportation Systems, 2024, 25(11): 16964-16977.
- [160] WANG X, AL-BASHABSHEH A, ZHAO C, et al. Smoothed noise contrastive mutual information neural estimation [J]. Journal of the Franklin Institute, 2023, 360(16): 12415-12435.
- [161] LAN Z X, REN Y L, YU H Y, et al. Hi-SCL: Fighting long-tailed challenges in trajectory prediction with hierarchical wave-semantic contrastive learning [J]. Transportation Research Part C: Emerging Technologies, 2024, 165: 104735.
- [162] QU Y S, RONG J, LI Z L, et al. ST-A-PGCL: Spatio-temporal adaptive periodical graph contrastive learning for traffic prediction under real scenarios [J]. Knowledge-based Systems, 2023, 272: 110591.
- [163] YOU Y N, CHEN T L, SUI Y D, et al. Graph contrastive learning with augmentations [C] // NeurIPS Foundation. Conference on Neural Information Processing Systems. Cambridge; MIT Press, 2020; 1-12.
- [164] 刘伟,贾素玲.基于图对比学习的稳健交通流量预测 [J]. 北京航空航天大学学报, 2024, 50(1): 122-133.
- LIU Wei, JIA Su-ling. Robust traffic flow prediction based on graph contrastive learning [J]. Journal of Beijing University of Aeronautics and Astronautics, 2024, 50(1): 122-133.
- [165] JIA R, GAO K, LIU Y, et al. I-CLTP: Integrated contrastive learning with transformer framework for traffic state prediction and network-wide analysis [J]. Transportation Research Part C: Emerging Technologies, 2025, 171: 104979.
- [166] JIN Y L, CHEN K, YANG Q. Selective cross-city transfer learning for traffic prediction via source city region re-weighting [C] // ACM. Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data

- Mining. New York; ACM, 2022; 734-741.
- [167] 赖培源,李 程,王增辉,等.基于图提示微调的交通流量预测[J].计算机研究与发展,2024,61(8):2020-2029.
LAI Pei-yuan, LI Cheng, WANG Zeng-hui, et al. Traffic flow prediction based on graph prompt-finetuning[J]. Journal of Computer Research and Development, 2024, 61(8): 2020-2029.
- [168] QU Y S, LI Z L, ZHAO X H, et al. Towards real-world traffic prediction and data imputation: A multi-task pretraining and fine-tuning approach [J]. Information Sciences, 2024, 657: 119972.
- [169] SHAO Z Z, ZHANG Z, WANG F, et al. Pre-training enhanced spatial-temporal graph neural network for multivariate time series forecasting[C]//ACM. Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining. New York; ACM, 2022; 1567-1577.
- [170] LI C, LIU W, YANG H. Deep causal inference for understanding the impact of meteorological variations on traffic [J]. Transportation Research Part C: Emerging Technologies, 2024, 165: 104744.
- [171] LIU J M, LIN H, WANG X D, et al. Reliable trajectory prediction in scene fusion based on spatio-temporal structure causal model[J]. Information Fusion, 2024, 107: 102309.
- [172] HU J J, BAI J, YANG J Y, et al. Crash risk prediction using sparse collision data: Granger causal inference and graph convolutional network approaches[J]. Expert Systems with Applications, 2025, 259: 125315.
- [173] YIN J, LI B. Long-short-term expert attention neural networks for traffic flow prediction[C]//Springer. Advanced Intelligent Computing Technology and Applications. Munich; Springer, 2024; 3-14.
- [174] 李舒豪. ST-MoE:用于交通预测去偏的时空混合专家框架[D].广州:广州大学,2023.
LI Shu-hao. ST-MoE: A spatio-temporal mixed expert framework for traffic forecasting depolarization [D]. Guangzhou: Guangzhou University, 2023.
- [175] JIANG W Z, HAN J D, LIU H, et al. Interpretable cascading mixture-of-experts for urban traffic congestion prediction [C]//ACM. Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining. New York; ACM, 2024; 5206-5217.
- [176] GUO S N, LIN Y F, GONG L T, et al. Self-supervised spatial-temporal bottleneck attentive network for efficient long-term traffic forecasting[C]//IEEE. 2023 IEEE 39th International Conference on Data Engineering (ICDE). New York; IEEE, 2023; 1585-1596.
- [177] ZHANG D R, YAN J N, POLAT K, et al. Multimodal joint prediction of traffic spatial-temporal data with graph sparse attention mechanism and bidirectional temporal convolutional network[J]. Advanced Engineering Informatics, 2024, 62: 102533.
- [178] FENG S F, FENG X X, XU L X, et al. BTG-GTAN: Federated traffic flow prediction with multimodal feature fusion considering anomalies[C]//IEEE. 2024 9th International Conference on Computer and Communication Systems (ICCCS). New York; IEEE, 2024; 462-467.
- [179] WANG B, QIN A K, SHAFIEI S, et al. Training physics-informed neural networks via multi-task optimization for traffic density prediction[C]//IEEE. 2023 International Joint Conference on Neural Networks (IJCNN). New York; IEEE, 2023; 1-7.
- [180] DENG P, ZHAO Y, LIU J T, et al. Spatio-temporal neural structural causal models for bike flow prediction[J]//AAAI. Proceedings of the AAAI Conference on Artificial Intelligence. Washington DC: AAAI Press, 2023; 4242-4249.
- [181] WANG H J, CHEN J Y, PAN T, et al. Easy begun is half done: Spatial-temporal graph modeling with ST-curriculum dropout[C]//AAAI. Proceedings of the AAAI Conference on Artificial Intelligence. Washington DC: AAAI Press, 2023; 4668-4675.
- [182] GUO H, MEESE C, LI W X, et al. B2SFL: A bi-level blockchained architecture for secure federated learning-based traffic prediction [J]. IEEE Transactions on Services Computing, 2023, 16(6): 4360-4374.
- [183] LU B, GAN X Y, ZHANG W N, et al. Spatio-temporal graph few-shot learning with cross-city knowledge transfer [C]//ACM. Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining. New York; ACM, 2022; 1162-1172.
- [184] GUO X S, ZHANG Q M, JIANG J Y, et al. Towards explainable traffic flow prediction with large language models [J]. Communications in Transportation Research, 2024, 4: 100150.
- [185] YUAN Y, DING J T, FENG J, et al. UniST: A prompt-empowered universal model for urban spatio-temporal prediction [C]//ACM. Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining. New York; ACM, 2024; 4095-4106.
- [186] 肖建力,邱 雪,张 扬,等.交通大模型综述[J].交通运输工程学报,2025,25(1):8-28.
XIAO Jian-li, QIU Xue, ZHANG Yang, et al. Review on large language models in transportation[J]. Journal of Traffic and Transportation Engineering, 2025, 25(1): 8-28.
- [187] 董瀚萱.数据丢失情况下高速路网交通状态预测方法研究[D].南京:东南大学,2022.
DONG Han-xuan. Research on traffic state prediction method of expressway network under data loss[D]. Nanjing: Southeast University, 2022.
- [188] 赵永梅,董云卫.低秩张量补全的时空交通数据预测[J].交通运输工程学报,2024,24(4):243-258.
ZHAO Yong-mei, DONG Yun-wei. Spatio-temporal traffic data prediction based on low-rank tensor completion [J]. Journal of Traffic and Transportation Engineering, 2024, 24(4): 243-258.