

文章编号:1671-6833(2025)05-0068-09

一种面向交通流量预测的自适应时空图卷积网络

张震^{1,2}, 刘博¹, 李卓², 张学忠³

(1. 郑州大学 河南先进技术研究院, 河南 郑州 450001; 2. 郑州大学 电气与信息工程学院, 河南 郑州 450001; 3. 国网周口供电公司, 河南 周口 466000)

摘要:针对现有交通流量预测方法未能充分利用节点属性指导图结构学习,以及在捕获复杂时空相关性方面存在局限性等问题,提出了一种结合自适应图结构学习和时空卷积架构的自适应时空图卷积网络(AdpSTGCN)。首先,设计一种基于节点属性的自适应图结构学习方法,从全局和局部两个视角动态学习道路网络的空间关系;其次,提出一种专用的时空卷积网络架构,有效地捕获交通流量中的时空相关性,进一步提升模型对复杂时空关系的建模能力,同时引入递进式训练策略来解决模型训练中可学习参数过多和数据稀疏性问题;最后,在高速公路交通数据集 METR-La、PEMS-Bay 中分别进行了 15、30、60 min 的交通流量预测实验。实验结果表明:AdpSTGCN 模型相较于多个基线模型,在 MAE、RMSE、MAPE 3 个预测误差指标上均表现最优。这说明该模型在未来短期和长期交通流量预测任务上均具有更优的建模能力,为城市交通疏导提供了理论依据。

关键词:交通流量预测;自适应图结构;节点属性;图卷积网络;时空相关性

中图分类号: TP391 **文献标志码:** A **doi:**10.13705/j.issn.1671-6833.2025.05.011

随着城市化进程的快速推进,城市道路网络日益复杂,交通拥堵问题愈发突出。相关统计数据表明,中国城市交通拥堵造成的年均经济损失已达 1 600 亿元,且这一数值仍呈上升态势^[1]。为缓解城市交通拥堵及其带来的经济损失,提升城市道路交通状况与通行能力成为迫切需求。实时监测城市道路的交通流量并预测未来相关数据,可为交通管理部门提供决策依据。基于此,管理部门能够及时采取疏导措施,有效缓解交通拥堵、提升通行效率,最终实现基础设施优化和降低经济损失。

交通流量数据具有高维度、非线性和异构性等特征,涵盖速度、密度等多维度变量^[2]。虽然已有研究利用支持向量机(support vector machine, SVM)^[3]和 K-最近邻算法(K-nearest neighbor, KNN)^[4]等机器学习方法实现了交通流量预测,但是预测的精度受限。深度学习技术的兴起为该领域带来新突破,其中图卷积网络(graph convolution network, GCN)^[5]的出现使得卷积操作得以适配路网拓扑结构。Li 等^[6]提出的 DCRNN 将交通网络数据的输入输出关系建模为一种扩散过程,利用双向随机

游走算法来捕捉时空依赖分布;Yu 等^[7]提出的 STGCN 模型,结合了时序卷积网络和图卷积网络,实现了路网时空相关性的联合建模。然而,这类方法在动态相关性捕捉和多周期模式识别方面仍存在局限。因此,Guo 等^[8]提出了一种基于注意力的时空图卷积网络(attention based spatial-temporal graph convolutional network, ASTGCN),通过设计近期交通流、日周期性、周周期性 3 个独立组件,结合时空注意力机制实现了动态特征的有效捕获;Zheng 等^[9]进一步提出 GMAN 模型,采用编码器-解码器架构用于建模时空因素的影响,在模拟动态相关性的同时缓解了误差传播问题,提升了长期预测性能。

上述方法捕获空间相关性主要依赖预定义的静态图结构,但此类结构不足以表征复杂的交通网络。针对这一问题,Zhang 等^[10]提出了 AdapGL 算法,通过多种空间学习视角生成适应性的图结构,该算法通过使用四组可学习参数,在图卷积操作中获取最优依赖图的表示形式。Wu 等^[11]和 Bai 等^[12]则通过可学习二维邻接矩阵对复杂道路网络进行建模。

收稿日期:2026-05-17;修订日期:2026-07-24

基金项目:河南省重点研发专项(231111211600)

作者简介:张震(1966—),男,河南郑州人,郑州大学教授,博士,博士生导师,主要从事计算机视觉、交通流量预测的研究,E-mail:zhangzhen66@126.com。

引用本文:张震,刘博,李卓,等.一种面向交通流量预测的自适应时空图卷积网络[J].郑州大学学报(工学版),2026,47(5):68-76.[Zhang Zhen, Liu Bo, Li Zhuo, et al. An adaptive spatial-temporal graph convolutional network for traffic flow forecasting [J]. Journal of Zhengzhou University (Engineering Science), 2026, 47(5): 68-76.]

然而,这些方法主要依赖可学习参数生成道路网络结构,而未充分考虑节点特征的影响,这在数据规模较小或稀疏时限制了模型性能的进一步优化。Lin 等^[13]提出的 DCGCN 虽然引入了动态因果图结构以捕捉交通流节点中的因果关系,但未显式融合节点特征信息,对节点自身特征的利用相对较弱。在交通流量预测中,节点特征能够反映道路交通状况的变化,在生成邻域图结构时具有更高的有效性。基于上述分析,本文提出了一种面向交通流量预测的自适应时空图卷积网络模型(adaptive spatial-temporal graph convolutional network for traffic forecasting, AdpSTGCN),该模型基于节点属性特征推导出最优图结构,能够动态捕捉时空耦合的复杂空间关系,实现对道路网络拓扑结构的动态表征,进而对未来交通状况进行精准预测。

1 AdpSTGCN 模型设计

1.1 交通流量预测问题定义

交通流量预测是指根据历史交通流量数据,采用某种预测模型或算法对其进行训练或分析,从而预测未来一段时间的交通流量信息。本文将道路网络表示为加权有向图 $G=(V,E,A)$ 。其中, V 表示道路网络中传感器的节点集合,节点数量 $|V|=N$; E 表示节点之间的边集合,反映了节点间连接的权重; $A \in \mathbf{R}^{N \times N}$ 是一个加权邻接矩阵,表示任意一对节点之间的关联程度。在 t 时刻,图 G 上的交通历史记录表示为一个图信号 $X_t \in \mathbf{R}^{N \times F}$,其中, F 表示每个节点属性的维度。交通预测问题旨在学习一个函数 f ,该函数能够根据 S 个历史图信号 $X=[X_{(t-S+1)}, \dots, X_t] \in \mathbf{R}^{S \times N \times F}$ 和交通网络的 $A \in \mathbf{R}^{N \times N}$ 的情况下,预测未来 T 个时间步的图信号 $\hat{Y}=[\hat{X}_{(t-S+1)}, \dots, \hat{X}_t] \in \mathbf{R}^{T \times N \times F}$,即 $\hat{Y}=f(X,A)$ 。

1.2 AdpSTGCN 模型框架

AdpSTGCN 模型框架如图 1 所示。该框架由一个自适应图结构学习组件和一个多步交通状况预测组件构成。自适应图结构学习组件旨在自动学习图结构,通过综合考虑图中所有节点的历史信息,以数据驱动的方式生成最优图邻接矩阵。多步交通状况预测组件旨在预测多步交通状况,它堆叠了多个时空层(ST-Layers),并基于学习到的图邻接矩阵进行预测以提高性能。在每个时空层中,首先通过门控时间卷积(gate temporal convolution, GTCN)模块来提取全局时间特征,以增强模型的时序建模能力。随后,通过图扩散卷积循环单元(graph diffusion convolu-

tion recurrent unit, GDCRU)模块进一步提取全局空间特征,并捕捉局部时间依赖性。通过跳跃连接(skip connection)融合多个堆叠时空块(ST-Blocks)中不同时空层的特征,AdpSTGCN 能够从不同维度综合考虑时空特征,从而精准地预测交通流量。AdpSTGCN 的学习过程为

$$A^* = g(X,A); \tag{1}$$

$$\hat{Y} = h(X,A^*). \tag{2}$$

式中: $g(X,A)$ 用于自适应学习图结构,该函数以节点属性 X 和邻接矩阵 A 作为输入,生成学习到的最优图结构 A^* ; $h(X,A^*)$ 基于学习到的图结构 A^* 进行多步交通状况预测,并输出最终的预测结果 $\hat{Y} \in \mathbf{R}^{T \times N \times F}$ 。

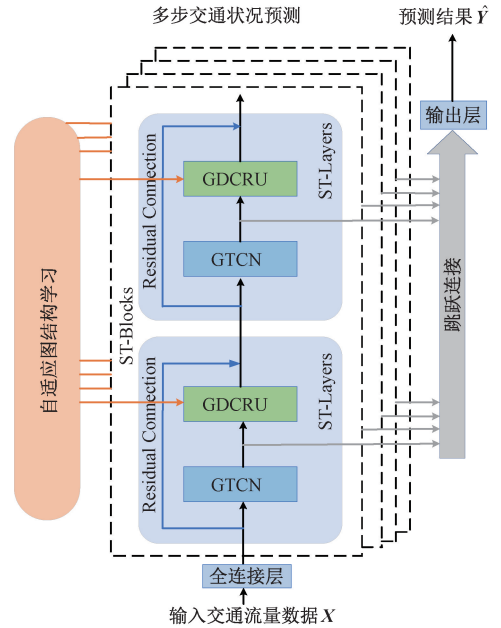


图 1 AdpSTGCN 模型框架

Figure 1 AdpSTGCN model framework

1.3 自适应图结构学习

图 2 展示了自适应图结构学习框架。传统方法依赖单一可学习矩阵拟合路网结构,但忽略了节点属性变化在道路结构学习中的作用,导致静态图在小样本或稀疏数据场景下难以优化的问题。为此,本文构建了一种自适应图结构学习组件,以数据驱动的方式获得最优图邻接矩阵。具体而言,图结构学习分为全局图结构学习和局部图结构学习两部分。全局图结构变化较为缓慢,因为从长期视角看,交通网络整体上相对稳定^[14]。然而,交通网络可能因外部因素(如局部地区的极端天气或短期的交通事故)而发生局部变化。因此,构建的模型需要从全局和局部视角自适应地推断图结构。

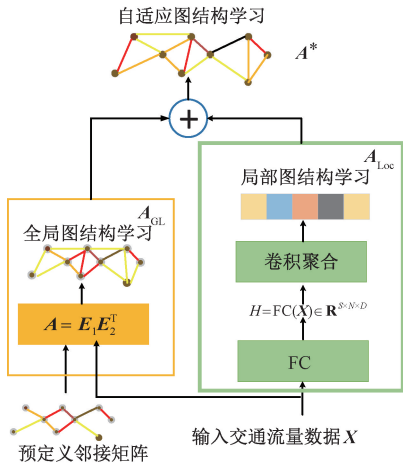


图2 自适应图结构学习框架

Figure 2 Adaptive graph structure learning framework

1.3.1 全局图结构学习

从全局视角来看,节点之间的依赖关系相对稳定。传统的预定义邻接矩阵通常通过欧氏距离来衡量节点之间的关系。然而,欧氏距离难以准确反映实际的道路拓扑结构^[15]。因此,本文采用全局学习的视角来学习交通网络。为了更精准地捕捉交通网络中节点间复杂的空间依赖关系,本文使用两个可训练向量 $E_1, E_2 \in \mathbf{R}^{N \times k}$ 来表示空间依赖关系,其中 k 为隐藏维度,用于减少参数数量。两个可训练向量能够从不同维度对交通流量节点特征进行编码,一个向量可以侧重于节点的地理位置特征,另一个向量可以关注节点周边的交通流量模式、区域功能等特征,从而更全面地捕捉节点之间的空间依赖关系。在训练过程中,考虑到交通网络原始预定义邻接矩阵直接优化难度大^[16],同时为保留原始图中的有效信息,原始预定义邻接矩阵借助残差连接并通过更新 ψ 来进行调整优化。

$$\psi = \sigma(\text{ReLU}(E_1 E_2^T)); \quad (3)$$

$$A_{GL} = \psi + A. \quad (4)$$

式中: E_1 为源节点的嵌入向量; E_2 为目标节点的嵌入向量,通过点积操作计算源节点和目标节点之间的空间依赖权重; ReLU 用于消除节点之间的弱连接;函数 $\sigma(\cdot)$ 为 softmax 函数,用于对全局信息邻接矩阵进行归一化处理。

1.3.2 局部图结构学习

从局部视角来看,当遇到交通事件时(如高峰时段、天气状况或交通事故),静态路网中的空间关系可能发生显著变化。由于节点与记录交通状况的属性相关联,本研究通过挖掘节点属性中的相关信息来描述这些突发波动。对于给定的节点属性 $X \in \mathbf{R}^{S \times N \times F}$,首先通过全连接网络将属性维度从 F 扩展

到 D ,公式如下:

$$H = \text{FC}(X) \in \mathbf{R}^{S \times N \times D}. \quad (5)$$

式中: H 表示转换后的节点属性。随后,为捕捉短期时间段 S 内交通流序列的临时空间关系,采用时间维度卷积聚合转换后输入 H 的节点特征,其公式如下:

$$M_i = K \sum_{p=1}^p X_{p,n,d} \circ \quad (6)$$

式中:卷积核 $K \in \mathbf{R}^P$ 用于在时间维度上聚合特征,进而完成交通流序列的卷积处理; n 表示节点索引; d 表示特征维度索引; p 表示时间维度索引; M_i 表示通过对节点 i 在所有时间维度上的属性进行聚合后得到的值。

完成时间维度上的聚合后,本文设计了一种度量学习方法,通过学习一对节点表示的度量函数 $\phi(\cdot, \cdot)$ 来推导节点之间的学习关系,其公式如下:

$$A_{i,j}^{\text{Loc}} = \phi(M_i, M_j) = M_i \cdot M_j^T, 1 \leq i, j \leq N. \quad (7)$$

式中: $A_{i,j}^{\text{Loc}}$ 表示节点 i 和节点 j 之间的学习关系。由于采用节点操作来表示节点间的接近性,该度量学习函数可以简化为矩阵乘法形式:

$$A^{\text{Loc}} = M \cdot M^T. \quad (8)$$

1.3.3 多层级图结构融合

在完成全局和局部图结构学习后,需要将这两部分的图结构信息进行融合,以构建一个能够全面反映交通网络空间依赖关系的最优图邻接矩阵,用于后续的多步交通流量预测任务。本文通过使用 ReLU 激活函数对两个邻接矩阵求和,然后归一化可得到最优的图结构,公式如下:

$$A^* = \text{Norm}(\text{ReLU}(A_{GL} + A_{Loc})). \quad (9)$$

1.4 多步交通状况预测

针对交通状况中复杂的时空关性,本文设计了一种时空图卷积网络来预测未来的多步交通状况。具体而言,设计了图扩散卷积循环单元(GDCRU)模块,旨在精准捕捉全局空间相关性与局部时间依赖性。其核心思路是将双向图扩散卷积嵌入门控循环单元架构中^[17]。在每个循环单元里,图扩散卷积能够处理来自上一循环单元输出的过往空间特征以及当前交通状况中的空间相关信息,以此强化空间信息的挖掘与整合能力。与此同时,借助门控循环单元自身的特性,该模块亦能够敏锐地捕捉局部时间维度上的依赖关系,进而实现对交通数据时空特性更为细致与准确的表征与建模。为了进一步捕获全局时间相关性,本文设计了 GTCN 模块。在门控时间卷积模块中,采用了扩张时域因果卷积^[18](dilated temporal causal convolution, DTCC),这种卷积方

式通过引入扩张因子,巧妙地扩大了卷积的感受野,使得模型能够在不增加过多计算量的情况下,跨越更广泛的时间区间去捕捉数据中的时间依赖特征。

1.4.1 图扩散卷积循环单元

DCRNN^[6]是一种探究节点表示的图卷积方法,其提出的双向扩散卷积能有效捕捉上下游交通的影响。本文将交通状态的变化建模为有向图上的信号扩散过程,并采用图扩散卷积来提取空间特征。其公式如下:

$$\mathbf{Z} = \sum_{k=0}^K \mathbf{P}_f^k \mathbf{X} \mathbf{W}_{k1} + \mathbf{P}_b^k \mathbf{X} \mathbf{W}_{k2}; \quad (10)$$

$$\mathbf{P}_f = \mathbf{A} / \text{rowsum}(\mathbf{A}); \quad (11)$$

$$\mathbf{P}_b = \mathbf{A}^T / \text{rowsum}(\mathbf{A}^T). \quad (12)$$

式中: $\mathbf{P}_f$ 和 $\mathbf{P}_b$ 分别为前向与后向传递矩阵; K 为扩散步数; $\mathbf{W}_{k1}$ 和 $\mathbf{W}_{k2}$ 为可训练参数。通过自适应图结构学习组件可以获得最优图结构 $\mathbf{A}^*$ 后,图扩散卷积层将在最优图结构上进行学习。具体方法是在式(10)中将 $\mathbf{P}$ 替换为 $\mathbf{A}^*$,本文定义基于 $\mathbf{A}^*$ 的图扩散卷积层表示为 $\tilde{\mathbf{Z}}$ 。

本文采用基于 RNN 的模型来捕获时间相关性。门控循环单元(gated recurrent unit, GRU)通过较少的参数有效解决梯度消失问题,并能够捕获长时间序列的依赖关系。为了同时捕获全局空间相关性和局部时间的依赖关系,本文将 GRU 中的矩阵乘法替换为图扩散卷积,并将其命名为图扩散卷积循环单元(graph diffusion convolution recurrent unit, GDCRU)。其公式如下所示:

$$u_t = \sigma(\tilde{\mathbf{Z}}_u(\mathbf{X}_t \parallel \mathbf{H}_{t-1})); \quad (13)$$

$$r_t = \sigma(\tilde{\mathbf{Z}}_r(\mathbf{X}_t \parallel \mathbf{H}_{t-1})); \quad (14)$$

$$\mathbf{C}_t = \sigma(\tilde{\mathbf{Z}}_c(\mathbf{X}_t \parallel (r_t \odot \mathbf{H}_{t-1}))); \quad (15)$$

$$\mathbf{H}_t = u_t \odot \mathbf{H}_{t-1} + (1 - u_t) \odot \mathbf{C}_t. \quad (16)$$

式中: $\mathbf{X}_t$ 为 t 时刻的输入特征; $\mathbf{H}_{t-1}$ 表示 $(t-1)$ 时刻从 GDCRU 输出的隐藏状态; r_t 和 u_t 分别为 t 时刻的重置门和更新门; $\tilde{\mathbf{Z}}_r$ 、 $\tilde{\mathbf{Z}}_u$ 、 $\tilde{\mathbf{Z}}_c$ 为扩散卷积层的可学习参数。图扩散卷积循环单元(GDCRU)将图扩散卷积捕获的空间相关性与门控循环单元(GRU)建模的时间相关性相结合,实现了交通数据中空间与时间相关性的同时学习,从而增强模型捕捉节点间依赖关系的能力。

1.4.2 门控时间卷积

扩张时域因果卷积(DTCC)是一种改进的卷积操作,通过引入扩张因子(dilation factor),显著扩大

了网络的时间感受野,使输出仅依赖于当前及历史时间步的数据,而不包含未来信息。扩张时域因果卷积的公式为

$$x * f(t) = \sum_{s=0}^{K-1} f(s) \cdot x(t - g \cdot s). \quad (17)$$

式中: $x \in \mathbf{R}^T$ 表示时间序列; g 表示扩张因子; $*$ 表示卷积运算符。通过逐步增加扩张因子 g ,模型在时间域上的感受野呈指数增长,使得模型能够堆叠较少的因果卷积层来捕获全局时间序列的信息。

为了更有效地探索全局时间相关性,本文借鉴了 Kong 等^[19]之前的工作,设计了一个门控时间卷积模块(GTCN)。该模块结合门控机制和扩张时域因果卷积,旨在高效地捕获复杂的全局时间相关性,并能处理长期依赖问题,其结构如图3所示。GTCN的每一层包含两个参数设置相同的扩张时域因果卷积。为了确保足够的时间窗口,本文在第 i 层中设置扩张因子为 2^{i-1} 。这种设计使感受野随网络层数的增加呈指数级增长,从而有效捕获时间序列中的全局时间相关性。GTCN 的公式如下:

$$\text{gated} = \sigma(f * \mathbf{X} + b); \quad (18)$$

$$\text{GTCN}(\mathbf{X}) = \text{gated} \otimes (g * \mathbf{X} + c) + (1 - \text{gated}) \otimes (h * \mathbf{X} + g). \quad (19)$$

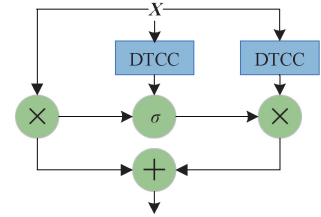


图3 GTCN 模块结构

Figure 3 Structure of GTCN module

1.4.3 多步预测的输出层

不同于以往的编码器-解码器模型,其下一步的预测依赖于前一步的预测结果,本研究通过输出层直接获取所有节点在未来 T 个时间步的交通状态预测结果,具体如下:

$$\hat{\mathbf{Y}} = \text{MLP}(\text{ReLU}(\text{MLP}(\text{ReLU}(\mathbf{Z}_s))))). \quad (20)$$

式中: $\mathbf{Z}_s$ 表示通过跳跃连接沿特征维度堆叠各 ST-Layer 融合后的特征; $\text{ReLU}(\cdot)$ 表示非线性激活层; $\text{MLP}(\cdot)$ 表示多层感知机。

在模型训练的过程中,与现有的研究一致,采用平均绝对误差作为训练目标,并对多步预测任务的损失进行优化。AdpSTGCN 的多步交通预测损失函数定义如下:

$$L(\mathbf{Y}, \hat{\mathbf{Y}}) = \frac{1}{T \cdot N \cdot D} \sum_{i=1}^T \sum_{j=1}^N \sum_{k=1}^D |\mathbf{Y}_{i,j,k} - \hat{\mathbf{Y}}_{i,j,k}|. \quad (21)$$

式中: $Y_{i,j,k}$ 为真实值; $\hat{Y}_{i,j,k}$ 为第 i 个时间步所有节点的预测值。

1.5 训练算法

在训练过程中, AdpSTGCN 模型的所有参数均可以通过梯度下降进行优化, 然而, 模型包含大量可学习参数且训练数据较为稀疏, 采用传统的端到端的训练方式可能导致模型训练不充分, 进而影响性能表现^[20]。

为了解决上述问题, 本文采用了递进式训练方法来优化模型训练过程。首先, 将全局层级图结构学习模块与交通状态预测组件相结合, 进行预训练 A_{GL} , 公式如下所示:

$$\hat{Y} = h(X, A_{GL})。 \quad (22)$$

此预训练策略为 A_{GL} 提供了良好的初始化, 从而使 AdpSTGCN 更易于收敛。接着, 通过式(9)将预训练的 A_{Loc} 与 A_{GL} 相结合, 并通过式(1)和式(2)以及损失函数式(21)对 AdpSTGCN 进行训练。递进式训练如算法 1 所示。

算法 1 AdpSTGCN 的递进式训练算法。

输入: 训练集 D_{train} 、邻接矩阵 $A \in \mathbf{R}^{N \times N}$;

输出: 训练后的 AdpSTGCN 模型。

- ① 初始化模型的所有参数;
- ② 重复执行以下步骤(步骤③~⑧);
- ③ 从训练集 D_{train} 中随机选取一个批次 $\{X, Y\}$;
- ④ 使用式(4)更新 $A_{GL} \leftarrow \psi + A$;
- ⑤ 使用式(22)计算预测值 $\hat{Y} = h(X, A_{GL})$;
- ⑥ 使用式(21)计算损失函数 $L(Y, \hat{Y})$;
- ⑦ 更新参数 $\theta_1 \leftarrow \theta_1 - \alpha_1 \nabla_{\theta_1} L$;
- ⑧ 更新邻接矩阵 $A_{GL} \leftarrow A_{GL} - \alpha_1 \nabla_{A_{GL}} L$;
- ⑨ 直到满足停止准则;
- ⑩ 重复执行以下步骤(步骤⑪~⑯);
- ⑪ 从训练集 D_{train} 中随机选取一个批次 $\{X, Y\}$;
- ⑫ 使用式(9)生成 $A^* \leftarrow \text{FUSE}(A_{GL}, A_{Loc})$;
- ⑬ 使用式(2)计算预测值 $\hat{Y} = h(X, A^*)$;
- ⑭ 使用式(21)计算损失函数 $L(Y, \hat{Y})$;
- ⑮ 更新参数 $\theta_1 \leftarrow \theta_1 - \alpha_1 \nabla_{\theta_1} L$;
- ⑯ 更新参数 $\theta_2 \leftarrow \theta_2 - \alpha_2 \nabla_{\theta_2} L$;
- ⑰ 直到满足停止准则;
- ⑱ 输出训练后的 AdpSTGCN 模型。

设 θ_1 表示多步交通状态预测组件中的可训练参数, θ_2 表示自适应图结构学习组件中的可训练参数, 它们的梯度分别为 $\nabla_{\theta_1} L$ 和 $\nabla_{\theta_2} L$ 。设 α_1 为初始阶段的学习率, β_1 和 β_2 为后续阶段的学习率。如算

法 1 所示, 训练集和预定义的邻接矩阵作为输入, 然后通过迭代全局层级图结构, 直至收敛。最后通过梯度下降更新整个模型, 直至满足停止准则。

2 实验部分

2.1 数据集

本文在两个真实交通流量数据集上对 AdpSTGCN 模型的性能进行了评估。

(1) METR-La: 该数据集涵盖 2012 年 3 月 1 日至 6 月 30 日洛杉矶高速公路的交通流量数据, 包含 207 个传感器节点, 记录时长达 4 个月。

(2) PEMS-Bay: 该数据集收集了 2017 年 1 月 1 日至 6 月 30 日旧金山湾区的交通速度数据, 涉及 325 个传感器节点, 覆盖时长为 6 个月。

本文对两个数据集的交通流量时间序列进行处理: 将采样间隔统一调整每条数据为 5 min, 按 7:1:2 的比例划分为训练集、验证集和测试集, 并对划分后的数据实施归一化处理。数据集的详细信息如表 1 所示。

表 1 数据集详细信息

Table 1 Dataset details

数据集	节点数量	步长	间隔/min	目标特征
METR-La	207	34 272	5	交通流量
PEMS-Bay	325	52 116	5	交通流量

2.2 对比模型和评价指标

实验使用了误差指标来评估模型, 分别为平均绝对误差(MAE)、均方根误差(RMSE)和平均绝对百分比误差(MAPE)作为实验的评价指标。评估指标的定义如下:

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|; \quad (23)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}; \quad (24)$$

$$MAPE = \frac{100\%}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right|。 \quad (25)$$

式中: y_i 为真实值; $\hat{y}_i$ 为预测值; n 为测试样本的数量。MAE、RMSE、MAPE 反映了预测值与真实值之间的差异, 当 MAE、RMSE、MAPE 值越小, 表明预测模型的性能越好。

本文选择 HA^[21]、ARIMA^[22]、FC-LSTM^[23]、DCRNN^[6]、STGCN^[7]、Graph WaveNet^[11]、GMAN^[9]、DMSTGCN^[24]、PGCN^[25]等 9 种同样可以用于交通流量预测领域的模型作为基准方法进行对比实验。

2.3 实验设置

本实验采用 CUDA 11.0 深度学习框架和 CuD-

NN 网络加速库,并基于 PyTorch 1.12.0 框架实现。所有实验均在配备 GRID T4-8Q (8 GB 显存) GPU, QEMU Virtual CPU (内存 32 GB) 的计算平台上运行。实验使用过去 1 h 的交通速度来预测未来 1 h 的交通速度,即 $S = T = 12$ 。为了进行时空特征学习,本文堆叠了 4 个 ST-Blocks,每个 ST-Block 包含两个 ST-Layer。为了将整个输入序列都考虑在内,本文将第一个时空层中 GTCN 模块的扩张因子设定为 1,第二个时空层中 GTCN 模块的扩张因子设定为 2。训练中,实验数据的批处理大小为 32,轮次 *epoch* 设置为 100,使用学习率为 0.001 的 Adam 优化器进行实验优化。

2.4 实验结果及分析

表 2 展示了 AdpSTGCN 模型与基准模型在两个数据集上对未来 15、30 和 60 min 的预测性能。实验结果表明,AdpSTGCN 在两个数据集上的评估指标均优于现有基线模型。

传统统计模型 (HA、ARIMA) 与非图结构深度学习模型 (FC-LSTM) 预测性能有限。其原因在于仅捕捉时间相关性,忽视交通数据复杂空间关联性,而实际预测中空间特征至关重要。以 DCRNN、STGCN 为代表的时空图卷积模型,通过图卷积建模空间关系提升了预测精度。STGCN 作为单步预测

模型,递归多步时误差随步长累积,DCRNN 采用 Seq2Seq 架构优化多步预测,但二者均依赖空间距离确定的固定邻接矩阵,难以准确刻画真实交通网络关联性。

Graph WaveNet 结合图卷积与因果卷积实现多步预测,利用自适应参数生成图结构,但未引入节点属性且生成固定拉普拉斯矩阵,缺乏动态表征能力。GMAN 通过多头注意力机制缓解多步误差累积,在 60 min 预测中接近 AdpSTGCN,但其依赖全局注意力建模空间关系,未融合道路拓扑结构且计算复杂度高。DMSTGCN 借助动态图生成算法捕获时空动态特性,中短期预测性能优于 GMAN,但其依赖外部特征生成图结构,跨数据集稳定性不足。PGCN 构建渐进图结合扩张卷积,在 PEMS-Bay 数据集 15 min 预测中 *MAE* 接近 AdpSTGCN,但缺乏对全局稳定特征与局部动态波动的协同建模。相比之下,AdpSTGCN 通过全局-局部双视角自适应推断动态图结构,兼顾时空相关性的稳定性与波动性建模,在多类模型对比中展现更强综合性能。

为了验证模型的鲁棒性,本文将 METR-La 数据集按月份分为 4 个子集,并与实验效果较好的 Graph WaveNet 模型进行比较,结果如表 3 所示。

表 2 不同模型在 METR-La 和 PEMS-Bay 数据集上的预测性能
Table 2 Prediction performance of different models on the METR-La and PEMS-Bay datasets

数据集	模型	15 min			30 min			60 min		
		<i>MAE</i>	<i>RMSE</i>	<i>MAPE</i> /%	<i>MAE</i>	<i>RMSE</i>	<i>MAPE</i> /%	<i>MAE</i>	<i>RMSE</i>	<i>MAPE</i> /%
METR-La	HA	4.16	7.80	13.00	4.16	7.80	13.00	4.16	7.80	13.00
	ARIMA	3.99	8.21	9.60	5.15	10.45	12.70	6.90	13.23	17.40
	FC-LSTM	3.44	6.30	9.60	3.77	7.23	10.90	4.37	8.69	14.00
	DCRNN	2.77	5.38	7.30	3.15	6.45	8.80	3.60	7.60	10.50
	STGCN	2.88	5.74	7.62	3.47	7.24	9.57	4.59	9.40	12.70
	Graph WaveNet	2.69	5.15	6.90	3.07	6.22	8.37	3.53	7.37	10.01
	GMAN	2.77	5.48	7.25	3.08	6.34	8.43	3.44	7.35	9.87
	DMSTGCN	2.85	5.54	7.54	3.26	6.56	9.19	3.72	7.55	10.96
	PGCN	2.70	5.16	6.98	3.08	6.22	8.38	3.54	7.36	9.94
	AdpSTGCN	2.64	5.04	6.76	3.01	6.01	8.17	3.42	7.08	9.83
PEMS-Bay	HA	2.88	5.59	6.80	2.88	5.59	6.80	2.88	5.59	6.80
	ARIMA	1.62	3.30	3.50	2.33	4.76	5.40	3.38	6.50	8.30
	FC-LSTM	2.05	4.19	4.80	2.20	4.55	5.20	2.37	4.96	5.70
	DCRNN	1.38	2.95	2.90	1.74	3.97	3.90	2.07	4.47	4.90
	STGCN	1.36	2.96	2.90	1.81	4.27	4.17	2.49	5.69	5.79
	Graph WaveNet	1.30	2.74	2.73	1.63	3.70	3.67	1.95	4.52	4.63
	GMAN	1.34	2.91	2.86	1.63	3.76	3.68	1.86	4.46	4.37
	DMSTGCN	1.33	2.83	2.80	1.67	3.79	3.81	1.99	4.54	4.78
	PGCN	1.30	2.73	2.72	1.62	3.67	3.63	1.92	4.45	4.55
	AdpSTGCN	1.29	2.71	2.69	1.60	3.60	3.59	1.85	4.32	4.48

表 3 模型在不同时间段和不同子数据集上 RMSE 对比
Table 3 RMSE comparison of the model on different time period and sub-datasets

模型	时间段	RMSE	
		Graph WaveNet	AdpSTGCN
1	2012-03-01—2012-03-31	9.47	7.34
2	2012-04-01—2012-04-30	7.53	6.51
3	2012-05-01—2012-05-31	8.72	6.92
4	2012-06-01—2012-06-30	6.32	5.34

通过实验数据发现,AdpSTGCN 的表现优于 Graph WaveNet,从而表明 AdpSTGCN 模型具有较高的鲁棒性。

2.5 消融实验

为了验证 AdpSTGCN 不同子模块的有效性,本文在 METR-La 和 PEMS-Bay 数据集上进行了消融实验。AdpSTGCN 的变体包括以下几个方面。

表 4 AdpSTGCN 在不同数据集上消融实验结果

Table 4 Experimental results of AdpSTGCN ablation on different datasets

数据集	模型	15 min			30 min			60 min		
		MAE	RMSE	MAPE/%	MAE	RMSE	MAPE/%	MAE	RMSE	MAPE/%
METR-La	w/o AL	2.79	5.31	7.30	3.25	6.43	8.72	3.72	7.48	10.33
	w/o GDCRU	2.68	5.08	6.82	3.05	6.12	8.27	3.50	7.20	9.87
	w/o GTCN	2.70	5.10	6.86	3.08	6.15	8.30	3.52	7.18	9.84
	AdpSTGCN	2.64	5.04	6.76	3.01	6.01	8.17	3.42	7.08	9.83
PEMS-Bay	w/o AL	1.34	2.80	2.85	1.70	3.83	3.89	1.97	4.56	4.96
	w/o GDCRU	1.30	2.72	2.71	1.62	3.65	3.62	1.88	4.34	4.49
	w/o GTCN	1.31	2.74	2.73	1.63	3.66	3.68	1.90	4.38	4.48
	AdpSTGCN	1.29	2.71	2.69	1.60	3.60	3.59	1.85	4.32	4.48

为了进一步验证自适应图结构组件及其他方法的优越性,本文在 METR-La 数据集上开展了一组消融实验。消融实验中,保持模型其他模块结构不变,将自适应图结构组件分别替换为预定义邻接矩阵及自适应图邻接矩阵^[11]。结果如表 5 所示。

表 5 不同图结构模块的性能比较

Table 5 Performance comparison of different graph structure modules

模块	MAE	RMSE	MAPE/%
预定义邻接矩阵	3.21	6.22	8.58
自适应图邻接矩阵	3.06	6.01	8.29
自适应图结构	2.96	5.89	8.05

由表 5 可知,使用预定义邻接矩阵和自适应图邻接矩阵的模型相较于自适应图结构的模型误差显著更高,这是因为前两者难以有效捕捉节点间的动态时空关联。

2.6 不同训练策略的对比分析

为了证明本文所设计的递进式训练方法的有效性,本文在 METR-La 数据集上将两种训练方法得到

(1)w/o AL:去除自适应图结构学习组件,使用预定义的静态图来训练模型。探究自适应图学习组件对捕获动态交通模式的作用。

(2)w/o GDCRU:去除多步交通预测组件中的图扩散卷积循环单元模块,改用传统的 GCN 模块来训练模型,以分析 GDCRU 模块在捕获时空关系中的优势。

(3)w/o GTCN:去除多步交通预测组件中的 GTCN 模块,探讨该模块对建模全局时间相关性的影响。

表 4 展示了消融实验中各模型变体在 15、30、60 min 的对比预测结果,所有变体的具体设置除上述差异外均与 AdpSTGCN 设置保持一致。消融实验结果显示,AdpSTGCN 的性能显著优于其他变体模型,充分验证了模型各组件的有效性。

的结果进行了对比,其结果如表 6 所示。

表 6 不同训练策略的性能比较

Table 6 Performance comparison of different training strategies

训练方法	MAE	RMSE	MAPE/%
端到端训练	3.15	6.04	8.42
递进式训练	2.96	5.89	8.05

由对比结果可以清晰看出,递进式训练方法在各项评估指标上均优于端到端训练。这是因为传统端到端训练方法在处理 AdpSTGCN 模型时,模型包含大量可学习参数且训练数据较为稀疏,难以对所有参数进行充分优化,容易陷入局部最优解,导致模型训练不充分,进而影响预测性能。而递进式训练方法采用分阶段训练策略,首先将全局层级图结构学习模块与交通状态预测组件相结合,对全局图结构进行预训练。这一过程为模型提供了良好的初始化条件,使得后续训练能够更高效地进行。在预训练阶段,模型学习到了交通网络中节点间的大致空间关系和长期趋势,为后续捕捉局部动态变化奠定

了坚实基础。之后,将预训练的结果与局部图结构相结合,对整个模型进行微调。这种逐步优化的方式能够有效降低模型训练难度,使模型更好地收敛,充分学习到数据中的时空特征,从而在交通流量预测任务中表现更优。

3 结论

本文针对交通流量预测中的关键挑战,提出了一种结合自适应图结构学习和时空卷积架构的自适应时空图卷积网络(AdpSTGCN)。通过设计基于节点属性的自适应图结构学习组件,从全局和局部两个视角动态建模复杂交通网络的空间依赖关系,同时构建了多层次的时空卷积网络架构,以更高效地捕捉交通流量中的时空相关性。此外,本文引入了递进式训练策略,有效缓解了模型参数过多和数据稀疏性带来的训练难度。实验结果表明,所提出的AdpSTGCN模型在METR-La和PEMS-Bay等公共交通数据集上,相较于当前主流模型,在短期和长期交通流量预测任务中均取得了显著的性能提升。通过自适应图结构学习组件,模型能够动态捕捉道路网络中的空间变化,而基于时空卷积的深度架构则有效挖掘了交通流量数据的时空相关性,从而提高了预测精度。

参考文献:

[1] Wang Jiahao, Li Wenbin, Guo Shiyao, et al. Urban traffic flow prediction based on global spatiotemporal graph convolutional neural network [J]. Computer Science, 2024, 51(S2): 534-542. [王佳昊, 黎文斌, 郭仕尧, 等. 基于全局时空图卷积神经网络的城市交通流量预测[J]. 计算机科学, 2024, 51(增刊2): 534-542.]

[2] Yang Gaofei, Xu Rui, Qin Ming, et al. Short-term traffic volume forecasting based on ARMA and Kalman filter [J]. Journal of Zhengzhou University (Engineering Science), 2017, 38(2): 36-40. [杨高飞, 徐睿, 秦鸣, 等. 基于 ARMA 和卡尔曼滤波的短时交通预测[J]. 郑州大学学报(工学版), 2017, 38(2): 36-40.]

[3] Yan He, Zhang Tianan, Qi Yong, et al. Short-term traffic flow prediction based on a hybrid optimization algorithm[J]. Applied Mathematical Modelling, 2022, 102: 385-404.

[4] Lin Guancen, Lin Aijing, Gu Danlei. Using support vector regression and K-nearest neighbors for short-term traffic flow prediction based on maximal information coefficient[J]. Information Sciences, 2022, 608: 517-531.

[5] Zhang Si, Tong Hanghang, Xu Jiejun, et al. Graph convolutional networks: a comprehensive review[J]. Com-

putational Social Networks, 2019, 6: 11.

[6] Li Yaguang, Yu R, Shahabi C, et al. Diffusion convolutional recurrent neural network: data-driven traffic forecasting[PP/OL]. V1. arXiv, (2017-07-06) [2025-06-19]. <https://doi.org/10.48550/arXiv.1707.01926>.

[7] Yu Bing, Yin Haoteng, Zhu Zhanxing. Spatio-temporal graph convolutional networks: a deep learning framework for traffic forecasting[PP/OL]. V4. arXiv (2018-07-12) [2026-01-19]. <https://doi.org/10.48550/arXiv.1709.04875>.

[8] Guo Shengnan, Lin Youfang, Feng Ning, et al. Attention based spatial-temporal graph convolutional networks for traffic flow forecasting[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2019, 33(1): 922-929.

[9] Zheng Chuanpan, Fan Xiaoliang, Wang Cheng, et al. G-MAN: a graph multi-attention network for traffic prediction[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2020, 34(1): 1234-1241.

[10] Zhang Wei, Zhu Fenghua, Lv Yisheng, 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.

[11] Wu Zonghan, Pan Shirui, Long Guodong, et al. Graph WaveNet for deep spatial-temporal graph modeling [PP/OL]. V1. arXiv (2019-05-31) [2026-01-19]. <https://doi.org/10.48550/arXiv.1906.00121>.

[12] Bai Lei, Bai Lei, Yao Lina, et al. Adaptive graph convolutional recurrent network for traffic forecasting[C]//Proceedings of the 34th International Conference on Neural Information Processing Systems. New York: ACM, 2020: 17804-17815.

[13] Lin Junpeng, Li Ziyue, Li Zhishuai, et al. Dynamic causal graph convolutional network for traffic prediction [C]//Proceedings of the 2023 IEEE 19th International Conference on Automation Science and Engineering (CASE). Piscataway: IEEE, 2023: 1-8.

[14] Diao Zulong, Wang Xin, Zhang Dafang, et al. Dynamic spatial-temporal graph convolutional neural networks for traffic forecasting[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2019, 33(1): 890-897.

[15] Lyu Mingqi, Hong Zhaoxiong, Chen Ling, et al. Temporal multi-graph convolutional network for traffic flow prediction[J]. IEEE Transactions on Intelligent Transportation Systems, 2021, 22(6): 3337-3348.

[16] Zhang Hong, Wang Hongyan, Chen Linlong, et al. Traffic flow forecasting based on transformer with diffusion graph attention network[J]. International Journal of Au-

- tomotive Technology, 2024, 25(3): 455–468.
- [17] Liu Haiyang, Zhu Chunjiang, Zhang Detian, et al. Attention-based spatial-temporal graph convolutional recurrent networks for traffic forecasting[C]//Advanced data mining and applications. Cham: Springer Nature Switzerland, 2023: 630–645.
- [18] Weng Wenchao, Fan Jin, Wu Huifeng, et al. A decomposition dynamic graph convolutional recurrent network for traffic forecasting[J]. Pattern Recognition, 2023, 142: 109670.
- [19] Kong Xiangyuan, Xing Weiwei, Wei Xiang, et al. STGAT: spatial-temporal graph attention networks for traffic flow forecasting[J]. IEEE Access, 2020, 8: 134363–134372.
- [20] Sha Ning, Wu Xiaochun, Wen Jinpeng, et al. Adaptive spatio-temporal graph convolutional network with attention mechanism for mobile edge network traffic prediction[J]. Cluster Computing, 2024, 27(9): 13257–13272.
- [21] Li Yaguang, Shahabi C. A brief overview of machine learning methods for short-term traffic forecasting and future directions[J]. SIGSPATIAL Special, 2018, 10(1): 3–9.
- [22] Pan Bei, Demiryurek U, Shahabi C. Utilizing real-world transportation data for accurate traffic prediction[C]//Proceedings of the 2012 IEEE 12th International Conference on Data Mining. Piscataway: IEEE, 2012: 595–604.
- [23] Sutskever I, Sutskever I, Vinyals O, et al. Sequence to sequence learning with neural networks[C]//Proceedings of the 28th International Conference on Neural Information Processing Systems. New York: ACM, 2014: 3104–3112.
- [24] Han Liangzhe, Du Bowen, Sun Leilei, et al. Dynamic and multi-faceted spatio-temporal deep learning for traffic speed forecasting[C]//Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining. New York: ACM, 2021: 547–555.
- [25] 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.

An Adaptive Spatial-Temporal Graph Convolutional Network for Traffic Flow Forecasting

ZHANG Zhen^{1,2}, LIU Bo¹, LI Zhuo², ZHANG Xuezhong³

(1. School of Henan Institute of Advanced Technology, Zhengzhou University, Zhengzhou 450001, China; 2. School of Electrical and Information Engineering, Zhengzhou University, Zhengzhou 450001, China; 3. State Grid Zhoukou Power Supply Company, Zhoukou 466000, China)

Abstract: To address the limitations of existing traffic flow prediction methods in fully utilizing node attributes to guide graph structure learning and capturing complex spatio-temporal dependencies, in this study an Adaptive Spatio-Temporal Graph Convolutional Network (AdpSTGCN) integrating adaptive graph structure learning with spatio-temporal convolutional architecture was proposed. Firstly, an adaptive graph structure learning method based on node attributes was designed to dynamically capture spatial relationships in road networks from both global and local perspectives. Secondly, a dedicated spatio-temporal convolutional architecture was developed to effectively model spatio-temporal correlations in traffic flow patterns, further enhancing the model's capability to handle complex spatio-temporal relationships. A progressive training strategy was introduced to address challenges of excessive learnable parameters and data sparsity during model training. Finally, experimental evaluations on highway traffic datasets (METR-La and PEMS-Bay) demonstrated the model's performance in 15, 30, and 60 minutes traffic flow prediction tasks. Experimental results showed that the AdpSTGCN model achieved the best performance among multiple baseline models in terms of three prediction error metrics: *MAE*, *RMSE*, and *MAPE*. These findings indicate the model's superior modeling capabilities for both short-term and long-term traffic flow prediction tasks, providing a theoretical foundation for urban traffic management strategies.

Keywords: traffic flow prediction; adaptive graph structure; node attributes; graph convolutional network; spatio-temporal correlation