

doi: 10.3969/j.issn.1672-6073.2026.04.019

融合多时间粒度与图注意力的地铁短时进站客流预测方法

沈鹏举^{1,2}, 宋丽英^{1,2}

(1. 北京交通大学交通运输学院, 北京 100044;

2. 北京交通大学综合交通运输大数据应用技术交通运输行业重点实验室, 北京 100044)

摘要:为解决地铁短时进站客流预测中多粒度时间特征融合不足、站点空间关联异质性刻画不充分等问题,提出一种融合多粒度感知与图注意力机制的地铁短时客流预测模型(multi-granularity and attention-enhanced passenger flow, MGAPF)。模型在时间维度构建4种不同感受野的一维卷积分支,并通过粒度注意力机制自适应融合多尺度时间特征;在空间维度引入图注意力机制,动态学习邻接站点间的关联强度;同时融合节假日、周末及天气等外部特征,并结合门控融合结构、残差连接和Transformer编码器完成时空特征交互与预测输出。基于杭州市地铁真实运营数据的实验结果表明,在10 min与30 min客流预测任务中,MGAPF的MAPE(平均绝对百分比误差)分别达到16.93%与9.52%,相较次优模型分别下降约4.5%与7.8%,且在MAE(平均绝对误差)、RMSE(均方根误差)及预测稳定性方面表现更优,表明融合多时间粒度特征、站点空间关联及外部影响因素能够有效提升地铁短时进站客流预测的精度与稳定性。

关键词:城市轨道交通; 客流预测; 多时间粒度; 图注意力网络; 深度学习

中图分类号: U231

文献标志码: A

文章编号: 1672-6073(2026)04-0161-09

Short-Term Metro Inbound Passenger Flow Prediction Using Multi-Granularity Temporal Features and Graph Attention

Shen Pengju^{1,2}, Song Liying^{1,2}

(1. School of Traffic and Transportation, Beijing Jiaotong University, Beijing 100044;

2. Key Laboratory of Transport Industry of Big Data Application Technologies for Comprehensive Transport of Ministry of Transport, Beijing Jiaotong University, Beijing 100044)

Abstract: To address the insufficient fusion of multi-granularity temporal features and inadequate modeling of spatial heterogeneity among stations in short-term metro inbound passenger flow forecasting, this paper proposes the Multi-Granularity and Attention-enhanced Passenger Flow Prediction model (MGAPF). In the temporal dimension, four one-dimensional convolutional branches with different receptive fields are constructed, and a granularity attention mechanism is used to adaptively fuse multi-scale temporal features. In the spatial dimension, a graph attention mechanism is introduced to dynamically learn the correlation strengths among adjacent stations. External temporal contextual features, including holidays, weekends, and weather conditions, are also incorporated. A gated fusion structure, residual connections, and a Transformer encoder are further combined to perform spatiotemporal feature interaction and generate prediction outputs. Experimental results based on real operational data from Hangzhou Metro show that MGAPF achieves MAPE values of 16.93% and 9.52% in the 10-minute and 30-minute forecasting tasks, respectively, representing error reductions of approximately 4.5% and 7.8% compared with the second-best baseline model. MGAPF also performs better in terms of MAE, RMSE, and prediction stability, indicating that the fusion of multi-

收稿日期: 2025-07-29 修回日期: 2026-06-30

第一作者: 沈鹏举, 男, 博士研究生, 研究方向为交通大数据挖掘, 21114022@bjtu.edu.cn

通信作者: 宋丽英, 女, 教授, 研究方向为智能交通建模方法, lysong@bjtu.edu.cn

基金项目: 广西科技重大专项(桂科 AA23062021-2); 国铁集团科研开发计划课题(P2025X002)

引用格式: 沈鹏举, 宋丽英. 融合多时间粒度与图注意力的地铁短时进站客流预测方法[J]. 都市轨道交通, 2026, 39(4): 161-169.

granularity temporal features, spatial correlations among stations, and external factors can effectively improve the accuracy and stability of short-term metro inbound passenger flow prediction.

Keywords: urban rail transit; passenger flow prediction; multi-granularity; graph attention network; deep learning

城市轨道交通因其安全、高效、低碳等特性,已成为城市居民通勤出行的主要方式之一。在快速城市化背景下,精准的短时客流预测对于保障轨道交通系统运行效率与服务水平具有重要意义,特别是在列车运行组织、客流引导与运营调度等应用场景中发挥着关键作用。随着刷卡记录与移动支付等技术的广泛应用,轨道交通客流数据在采样频率与覆盖范围方面持续提升,为高精度建模提供了坚实的数据基础^[1]。与此同时,客流变化除受历史时序规律影响外,还会受到周末、节假日及天气变化等外部因素的共同作用,使得地铁客流预测呈现更加复杂的时空动态特征。

现有研究广泛采用序列建模方法以刻画客流的时间演化特性,其中,基于支持向量回归(support vector regression, SVR)的预测模型能够有效处理客流数据中的非线性变化特征^[2],门控循环单元(gated recurrent unit, GRU)适用于提取短期动态信息^[3],长短期记忆网络(long short-term memory, LSTM)增强了对长期依赖关系的建模能力^[4]。在此基础上,部分学者结合 LSTM 与 Transformer 构建编码-解码结构,以拓展模型对复杂时间模式的表达能力^[5]。此外,卷积与循环结构的融合被用于提升模型对局部波动的建模能力,而双向循环结构则有助于增强序列表达的稳定性与深度^[6]。尽管上述方法在时间建模方面取得了一定进展,但大多数方法仍采用单一时间粒度进行建模,难以同时兼顾短时突变、周期波动与长期趋势等多尺度特征。为缓解粒度建模的局限性,有研究引入多时间尺度输入以增强模型对不同时间模式的表达能力^[7],并结合多粒度特征与深层结构设计,提升对跨尺度动态的感知效果^[8],但这类方法大多基于简单拼接策略整合不同粒度特征,缺乏对各时间粒度相对贡献程度的动态调节能力,难以充分挖掘不同时间尺度之间的互补信息。

此外,地铁客流还受到站点空间拓扑结构的显著影响,图结构建模逐渐成为表达空间依赖关系的主流方式^[9]。部分方法通过将图模型与序列模型串联构建联合结构,以刻画时空耦合特征^[10],也有研究将图结构直接嵌入时间建模流程中,以增强空间信息在序列演化过程中的表达能力^[11]。然而,现有方法普遍采用固定邻接矩阵或普通图卷积结构,默认所有邻接节点具有相同影响权重,难以刻画不同邻接站点之间客流

传播强度的异质性。相较之下,图注意力机制(graph attention network, GAT)能够动态学习不同邻接站点对目标站点的影响程度,从而更加符合实际地铁网络中复杂的空间关联特征。

近两年,相关研究进一步从物理约束、多源特征筛选与深层时序建模等角度改进城市轨道交通客流预测模型,说明融合外部因素、空间关联与时间依赖仍是提升预测精度的重要方向^[12]。然而,仅依赖单一时序模型或固定图结构,仍难以同时刻画短时局部波动、多时间尺度变化和站点间异质空间影响。因此,如何联合多粒度时间建模、空间异质关联学习和长时间依赖建模,构建适用于地铁短时客流预测的统一时空预测框架,仍是当前研究的关键。

针对上述问题,本文提出一种融合多时间粒度感知机制与图注意力网络的地铁短时客流预测模型(multi-granularity and attention-enhanced passenger flow, MGAPF)。模型通过多粒度卷积分支提取不同时间尺度特征,并利用粒度注意力机制动态调节各尺度的重要性;通过图注意力机制学习邻接站点间的关联强度,增强空间依赖建模能力;进一步融合节假日、周末和天气等外部信息,并结合门控融合结构、残差连接和 Transformer 编码器完成预测。本文预测对象为各站点未来时间步的进站客流量,不包括出站量与换乘量。

1 问题描述

1.1 数据说明

本研究使用2019年1月1日至26日的杭州地铁刷卡记录数据,涵盖1号线、2号线和4号线共81个站点,累计记录量超过7000万条。该数据具有连续时间跨度、完整刷卡记录和统一字段口径,能够支持不同模型在相同数据条件下开展公平对比。受运营数据开放性和脱敏获取条件限制,本文未进一步获取更新年度的同口径客流数据,后续将结合更多年份和更新线网数据进一步检验模型的跨时期适应性。每条记录包含刷卡时间、站点编号与进出站标识,统计单位为“人次”。以各站点未来时间步的进站客流量作为预测对象。表1列出了建模所用的主要字段。

为统一网络结构表示,对各线路站点编号进行重新编码,确保编号在整体网络中具备连续性与唯一性。

由于编号为 54 的站点存在严重数据缺失, 予以剔除, 最终保留 80 个有效站点。图 1 展示了编号处理后的站点拓扑结构。

表 1 地铁刷卡数据字段说明

Table 1 Description of metro smart-card data fields

名称	类型	说明	示例
time	String	刷卡时间	2019-01-11 12:30
LineID	String	地铁线路编号	2
stationID	int	地铁站编号	10
deviceID	int	刷卡设备编号	2688
status	int	进出站状态	1

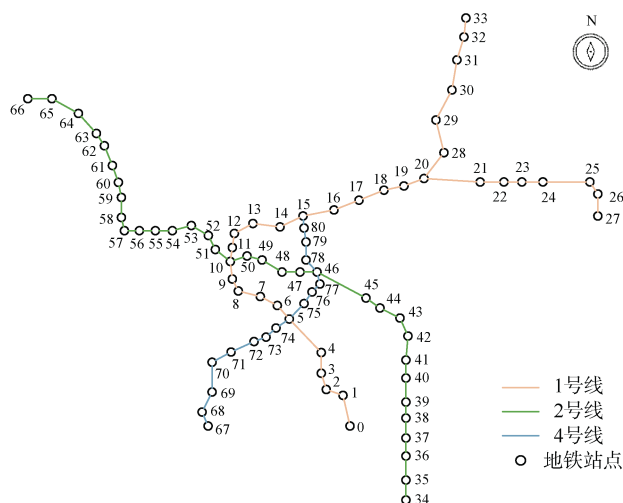


图 1 2019 年杭州市地铁网络站点

Figure 1 Hangzhou metro network (2019)

按自然日对刷卡记录进行时间聚合, 并分别构建 10 min 和 30 min 目标粒度下的预测数据集; 在案例展示中按全日 10 min 时间步呈现客流变化过程, 以便观察完整日内波动特征。模型训练过程中, 在给定目标预测粒度的历史进站客流序列基础上, 通过不同感受野的一维卷积分支提取短时局部波动与较长时间趋势特征。除历史客流序列外, 为刻画客流周期性变化规律并反映外部环境变化对居民出行需求的影响, 进一步融合节假日、周末及天气等外部影响因素。其中, 节假日与周末用于刻画客流周期性变化规律, 天气因素用于反映外部环境变化对居民出行需求的影响。上述外部时间上下文特征在数据预处理阶段与对应时间步客流序列对齐, 作为辅助输入特征共同输入模型, 以增强模型对复杂客流场景的适应能力与鲁棒性。

此外, 为降低异常数据对模型训练的干扰, 在数据预处理阶段对缺失记录、重复记录及异常时间段数据进行统一筛选与清洗, 并采用统一时间粒度聚合方

式构建多站点时序客流数据集, 以保证模型输入数据的一致性与稳定性。

1.2 研究问题

本研究面向地铁站点短时进站客流预测, 旨在基于历史客流序列、外部影响因素与地铁网络结构, 预测未来时间步各站点的进站客流规模。主要以 10 min 与 30 min 尺度下的短时客流预测为主实验任务, 为地铁实时运营组织与客流调度提供参考; 同时补充不同预测时间粒度下的敏感性分析, 以考察预测目标粒度变化对模型误差的影响。

首先, 地铁客流数据呈现出多尺度动态特征: 不同时间粒度下的序列反映出各自侧重的变化模式, 不同序列存在显著的互补性。其中, 较小时间粒度更有助于捕捉短时波动与局部突变, 较大时间粒度则更有利于反映整体趋势与周期性变化。然而, 现有方法大多采用简单拼接方式融合多粒度信息, 忽略了粒度间的结构差异与权重关系, 难以有效提取多尺度时序特征。因此, 如何根据预测任务自适应刻画不同时间粒度的相对贡献, 是提升短时客流预测精度的关键。

其次, 地铁站点间的客流影响具有结构性与异质性。为此, 城市地铁系统可抽象为图结构 $G=(V, E, A)$, 其中, V 为地铁站点集合, E 为站点间的物理联通关系, A 为站点的邻接矩阵。目前多数研究采用图卷积网络(graph convolutional network, GCN)构建地铁网络空间结构, 默认相邻节点影响等权, 但它难以反映站点间关联强度的差异, 限制了空间特征的表达能力。而在实际地铁网络中, 不同邻接站点因区位功能、换乘关系和客流规模差异, 对目标站点客流变化的影响并不相同。因此, 相较于普通 GCN 的固定权重传播方式, 图注意力机制能够动态学习不同邻接站点对目标站点的影响程度, 更适合刻画地铁站点间异质性的空间关联关系。

此外, 短时客流变化不仅受到历史时序规律与空间结构关系影响, 还会受到节假日、周末及天气等外部因素的调节。这类外部因素通常不改变地铁网络的物理拓扑结构, 但会影响不同日期和天气条件下的整体出行需求、峰值强度与时段分布特征。因此, 本研究将外部影响因素作为辅助输入变量, 与历史客流序列共同参与预测建模, 以增强模型对复杂客流场景的适应能力。与此同时, Transformer 结构在时间依赖建模方面具有较强表达能力, 能够进一步捕捉历史序列中的全局依赖关系。但仅依赖 Transformer 难以充分

利用地铁客流中的局部波动特征与站点空间拓扑信息, 因此需要将多粒度时间特征提取、空间异质关联学习和长时间依赖建模进行协同融合。

综合上述背景, 本文的研究问题可表述为: 在给定多粒度历史客流序列 $X_t^{(h)}$ 和地铁结构图 G 的条件下, 学习映射函数 f :

$$\left(\left\{X_{t-P+1:t}^{(h)}\right\}, C_t, G\right) \rightarrow \hat{X}_{t+1}^{(H)} \quad (1)$$

式中, h 为历史客流序列的时间粒度索引; P 为历史序列长度; H 为目标预测时间粒度; C_t 为节假日、周末、天气等外部特征; $\hat{X}_{t+1}^{(H)}$ 为预测结果。

通过多粒度注意力机制刻画不同时间尺度的贡献差异, 通过图注意力机制学习站点间异质关联强度, 并结合 Transformer 编码器建模时间依赖关系, 从而构建面向地铁短时进站客流预测的统一时空建模框架。

2 模型构建

MGAPF 模型整体结构由 4 个功能子模块构成: 多粒度卷积特征提取、空间图注意力编码、时空门控融合以及基于 Transformer 的时序建模。各模块功能相互协作, 共同完成对多站点、多时间尺度的地铁短时进站客流预测。在输入层, 将历史进站客流序列与节假日、周末、天气等外部特征按时间步对齐, 并共同构成输入特征张量, 以增强模型对不同日期属性和外部环境变化下客流波动的表达能力。模型整体结构如图 2 所示, 具体建模流程如下。

2.1 多粒度卷积模块

为捕捉不同时间尺度上的动态变化, 本模型在时间维度上设计了多分支的一维卷积结构。地铁短时客流同时具有局部突变、短周期波动和整体趋势变化等特征, 单一时间尺度往往难以兼顾不同类型的变化规律。因此, 通过多个卷积核尺度构造不同感受野的时间特征提取分支, 并进一步利用粒度注意力机制自适应刻画不同时间粒度对预测任务的相对贡献, 从而避免简单拼接方式难以区分不同粒度重要性的问题。设输入序列为张量 $\mathbf{X} \in \mathbb{R}^{B \times N \times T \times C}$, 其中, B 为样本数, N 为地铁站数量, T 为时间步数, C 为输入特征维度, 包括历史进站客流特征以及节假日、周末、天气等外部

时间上下文特征。第 g 个卷积分支采用感受野为 k_g 的一维卷积对完整输入特征张量进行特征提取, 并经过标准化与 ReLU 非线性变换, 得到对应的时间特征。

$$\mathbf{H}_t^{(g)} = \text{ReLU}\left(\text{BN}\left(\text{Conv1D}_{k_g}(\mathbf{X})\right)\right) \quad (2)$$

式中, $\mathbf{H}_t^{(g)}$ 表示第 g 个卷积分支在时刻 t 提取的时间特征; $\text{Conv1D}_{k_g}(\cdot)$ 、 $\text{BN}(\cdot)$ 、 $\text{ReLU}(\cdot)$ 分别表示卷积核大小为 k_g 的一维卷积、批归一化和 ReLU 激活函数。

卷积结果经线性映射投影至统一维度 d :

$$\hat{\mathbf{H}}_t^{(g)} = \mathbf{H}_t^{(g)} \cdot \mathbf{W}^{(g)} \in \mathbb{R}^{B \times N \times T \times d} \quad (3)$$

式中, $\mathbf{W}^{(g)} \in \mathbb{R}^{C \times d}$ 为第 g 个分支的线性投影矩阵, 维度 d 为预设超参数。多个分支输出通过可学习的权重 α_g 进行融合, 形成最终的时间特征表示:

$$\mathbf{H}_t = \sum_{g=1}^G \alpha_g \cdot \hat{\mathbf{H}}_t^{(g)}, \alpha_g = \frac{e^{\beta_g}}{\sum_{g'=1}^G e^{\beta_{g'}}} \quad (4)$$

Softmax 权重中的参数 β_g 为可学习标量, 用于衡量第 g 个卷积分支在最终时间特征融合中的重要性。通过该机制, 模型能够根据不同预测场景动态调整各

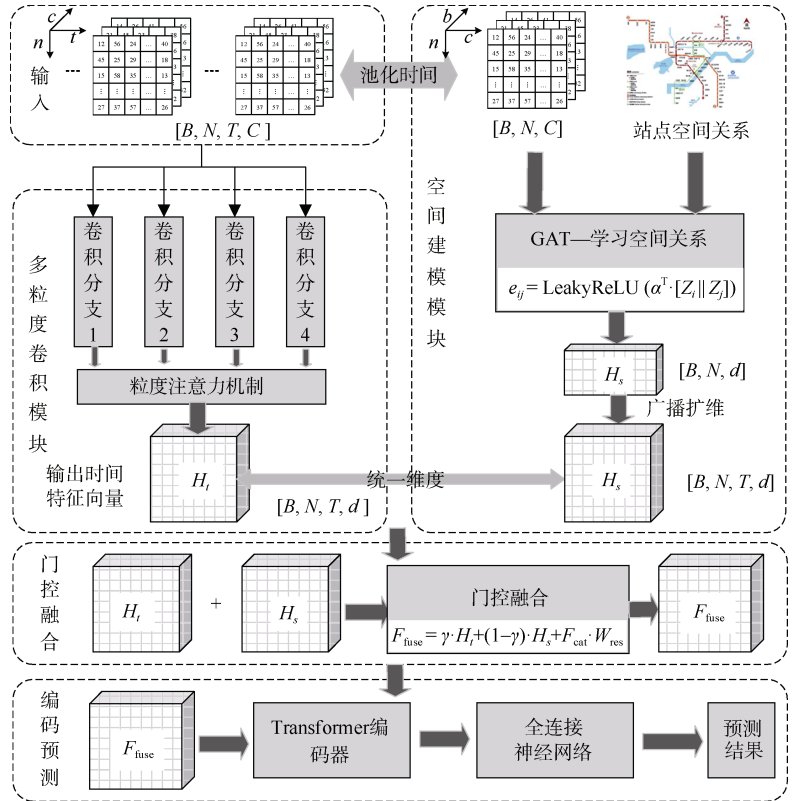


图 2 MGAPF 模型架构

Figure 2 Overall architecture of the MGAPF model

时间尺度特征的贡献,而不是将不同粒度特征进行固定拼接,从而提升多尺度时序特征表达能力。

2.2 空间建模模块

考虑到空间结构中各站点间的复杂依赖关系,模型引入图注意力机制对空间信息进行建模。地铁站点之间虽然具有明确的物理连接关系,但不同相邻站点对目标站点客流变化的影响强度并不完全相同。普通GCN通常基于固定邻接矩阵进行特征传播,默认相邻节点影响权重相同,难以表达站点间客流关联强度的差异。相比之下,GAT能够在邻接约束下动态学习不同邻接站点的影响权重,更适合刻画地铁网络中的空间异质关联。输入为特征张量 $\mathbf{X}_s \in \mathbb{R}^{B \times N \times T \times C}$,表示时间-空间的原始特征。为便于空间关系建模,先在时间维度进行池化处理,获得每个站点的聚合表示为 $\mathbf{X}_s^{\text{pool}} \in \mathbb{R}^{B \times N \times C}$,随后对输入特征进行整合,以构造节点级表示:

$$\mathbf{h}_i = \text{Aggregate}(\mathbf{X}_s^{\text{pool}}[:, i, :]) \in \mathbb{R}^{B \times 1} \quad (5)$$

式中, $\mathbf{h}_i$ 为第 i 个站点的节点级表示, $\text{Aggregate}(\cdot)$ 表示对该站点池化特征的聚合操作; $\mathbb{R}^{B \times 1}$ 表示维度为 $B \times 1$ 的实数空间。

得到所有位置的嵌入表示 $\mathbf{H}_s \in \mathbb{R}^{B \times N \times d_{in}}$, $\mathbf{H}_s$ 为所有站点的节点表示矩阵, N 为站点数, d_{in} 为节点输入特征维度,再经线性映射得到第 i 个站点的嵌入表示 $\mathbf{Z}_i$:

$$\mathbf{Z}_i = \mathbf{H}_s[:, i, :] \cdot \mathbf{W} \in \mathbb{R}^{B \times d} \quad (6)$$

式中, $\mathbf{W} \in \mathbb{R}^{d_{in} \times d}$ 为线性投影矩阵, d 为节点特征的目标嵌入维度, d_{in} 为池化后输入特征的维度。基于节点对 (i, j) 的注意力打分计算如下:

$$e_{ij} = \text{LeakyReLU}(\alpha^T \cdot [\mathbf{Z}_i \parallel \mathbf{Z}_j]) \quad (7)$$

式中, e_{ij} 为节点 i 与邻接节点 j 之间的未归一化注意力得分, $\text{LeakyReLU}(\cdot)$ 为激活函数, $\mathbf{Z}_i$ 和 $\mathbf{Z}_j$ 分别为节点 i 与节点 j 的嵌入表示。 $\alpha \in \mathbb{R}^{2d}$ 为用于注意力打分的可学习参数向量。结合邻接矩阵对注意力进行筛选归一化:

$$\alpha_{ij} = \frac{\exp(e_{ij}) \cdot A_{ij}}{\sum_k \exp(e_{ik}) \cdot A_{ik}} \quad (8)$$

式中, α_{ij} 为节点 i 对邻接节点 j 的归一化注意力权重。最终空间表示为邻接节点特征的加权和:

$$\mathbf{H}_{GAT}[:, i, :] = \sum_{j=1}^N \alpha_{ij} \cdot \mathbf{Z}_j \quad (9)$$

将式(9)得到的空间特征 $\mathbf{H}_{GAT}$ 记为 $\mathbf{H}_s^{\text{spatial}} \in \mathbb{R}^{B \times N \times d}$, 为匹配时间特征维度, 将其在时间维度上进行广播扩展, 得到最终的空间特征向量为 $\mathbf{H}_s \in \mathbb{R}^{B \times N \times T \times d}$ 。通过上述建模过程, 空间模块不仅利用地铁站点的物理邻接关系约束信息传播范围, 还能够自适应区分不同邻接站点目标站点客流变化的影响强度, 从而增强空间特征表达能力。

2.3 门控融合模块

为了有效融合时间与空间信息, 模型设计了门控融合结构。考虑到时间特征与空间特征在表达重点上存在差异, 通过门控机制动态调节两类特征的融合比例, 从而增强模型对复杂时空关联关系的自适应表达能力。设时间特征、空间特征分别为 $\mathbf{H}_t \in \mathbb{R}^{B \times N \times T \times d}$ 、 $\mathbf{H}_s \in \mathbb{R}^{B \times N \times T \times d}$, 两者已在维度上对齐, 将其拼接得到时空联合特征 $\mathbf{F}_{\text{cat}}$, 其中“ $\parallel$ ”表示特征拼接:

$$\mathbf{F}_{\text{cat}} = [\mathbf{H}_t \parallel \mathbf{H}_s] \in \mathbb{R}^{B \times N \times T \times 2d} \quad (10)$$

通过 MLP 和 Sigmoid 激活生成融合系数 γ :

$$\gamma = \sigma(\text{ReLU}(\mathbf{F}_{\text{cat}} \cdot \mathbf{W}_1) \cdot \mathbf{W}_2) \quad (11)$$

式中, $\text{ReLU}(\cdot)$ 为激活函数, $\sigma(\cdot)$ 为 Sigmoid 函数, $\mathbf{W}_1$ 和 $\mathbf{W}_2$ 分别为多层感知机第一层和第二层的权重矩阵。最终融合输出结合残差路径表达为:

$$\mathbf{F}_{\text{fuse}} = \gamma \cdot \mathbf{H}_t + (1 - \gamma) \cdot \mathbf{H}_s + \mathbf{F}_{\text{cat}} \cdot \mathbf{W}_{\text{res}} \quad (12)$$

式中, $\mathbf{W}_{\text{res}} \in \mathbb{R}^{2d \times d}$ 为残差映射权重, 用于匹配维度并增强信息保留能力。通过门控融合机制, 模型能够根据不同时间步与不同站点的特征变化, 自适应平衡时间依赖特征与空间关联特征的重要性, 从而提升复杂客流场景下的特征表达能力与稳定性。

2.4 编码预测模块

为了捕捉融合后特征中的时间依赖性, 模型采用标准 Transformer 编码器对其序列建模。设融合后的特征为 $\mathbf{F}_{\text{fuse}} \in \mathbb{R}^{B \times N \times T \times d}$, 为了适配 Transformer 结构, 按时间维度将其组织为序列形式, 并记为编码器输入序列 $\mathbf{F}_{\text{token}}$:

$$\mathbf{F}_{\text{token}} = \mathbf{F}_{\text{fuse}} \in \mathbb{R}^{B \times N \times T \times d} \quad (13)$$

经 Transformer 编码器 $\text{TransEncoder}(\cdot)$ 处理后, 得到编码特征 $\mathbf{F}_{\text{enc}}$:

$$\mathbf{F}_{\text{enc}} = \text{TransEncoder}(\mathbf{F}_{\text{token}}) \in \mathbb{R}^{B \times N \times T \times d} \quad (14)$$

最终模型取输出序列的最后一时刻的编码结果作为预测表示, 并通过全连接层获得下一时刻的目标预测值:

$$\hat{Y} = F_{enc}[:, :, -1, :] \cdot W_{out} \in \mathbb{R}^{B \times N} \tag{15}$$

式中， $\hat{Y}$ 为下一时刻的客流预测结果， W_{out} 为全连接层权重，Transformer 编码结果的最后时刻输出经全连接层映射为预测值。由于前序模块已融合多粒度时间特征、空间异质关联特征以及外部特征信息，因此，Transformer 编码器能在统一时空特征空间中进一步建模复杂时间依赖关系，从而提升短时客流预测精度与模型鲁棒性。

3 实验验证

3.1 超参数设置

为提升模型在多粒度时间特征和空间结构建模下的泛化能力，本文采用 Optuna 框架对关键超参数进行自动搜索与优化，最终选取在验证集上表现最优的参数组合作为模型配置。在 Transformer 模块中，编码维度设为 64，注意力头数为 4，层数为 3；图注意力模块输出维度设为 64。多粒度卷积模块采用感受野分别为 3、5、7 和 11 的卷积核，用于提取不同尺度的时序特征。其中，较小卷积核主要用于捕捉短时局部波动，较大卷积核则用于提取较长时间范围内的趋势变化特征。

训练过程中使用 Adam 优化器，初始学习率设为 0.001，批次大小为 16，最多训练 200 轮，并引入早停策略，当验证性能在 100 个连续轮次内无提升时终止训练。所有实验均基于 80 个地铁站点的进站客流数据展开，输入序列长度为 18，预测目标为未来一个时间步内的进站客流量，损失函数为均方误差(MSE)，并采用验证集最优策略保存模型。模型输入除历史客流序列外，还包含节假日、周末及天气等外部特征。

3.2 模型比较

模型性能评估采用 3 项指标：平均绝对误差 MAE、均方根误差 RMSE 和平均绝对百分比误差 MAPE(仅统计真实值大于 10 的样本)。模型比较实验均在统一数据集与实验设定下独立运行 10 次，结果以预测误差的均值与标准差表示(记作 $x \pm y$ ，其中 x 为平均值， y 为标准差)，以全面评估模型的精度、稳定性与泛化能力。

为全面验证 MGAPF 模型的有效性，选取 6 种具有代表性的基线方法作为对比，涵盖时间建模、空间建模以及时空联合建模 3 类典型策略。其中，LSTM 与 Transformer 分别代表基于循环神经网络与自注意力机制的时间序列预测方法；GCN 用于建模地铁网络中

的空间拓扑关联；GCN-LSTM 与 GCN-Transformer 分别结合图卷积与时序模型，实现基础的时空联合建模；ResGCN-Transformer 则在 GCN-Transformer 基础上引入多粒度残差图卷积与并行时间分支，具备一定的多尺度建模能力。所有模型均在 10 min 与 30 min 两个目标时间粒度下分别进行训练与测试，具体结果分别如表 2、表 3 所示。

表 2 各模型在 10 min 预测任务下的结果比较

Table 2 Performance comparison of different models in the 10-minute prediction granularity

模型	MAE	RMSE	MAPE/%
MGAPF	17.11 ± 0.32	26.09 ± 0.51	16.93 ± 0.49
ResGCN-Transformer	17.97 ± 0.35	27.14 ± 0.47	17.73 ± 0.54
GCN-Transformer	18.31 ± 0.37	27.61 ± 0.48	17.80 ± 0.35
GCN-LSTM	18.75 ± 0.55	28.11 ± 0.97	17.91 ± 0.41
LSTM	20.93 ± 0.42	29.52 ± 1.21	19.82 ± 0.74
GCN	55.52 ± 1.52	96.34 ± 3.07	57.83 ± 1.56
Transformer	18.33 ± 0.23	27.85 ± 0.36	17.86 ± 0.49

表 3 各模型在 30 min 预测任务下的结果比较

Table 3 Performance comparison of different models in the 30-minute prediction granularity

模型	MAE	RMSE	MAPE/%
MGAPF	29.02 ± 0.74	44.31 ± 0.73	9.52 ± 0.38
ResGCN-Transformer	31.99 ± 0.93	48.42 ± 1.46	10.33 ± 0.42
GCN-Transformer	32.63 ± 1.20	48.82 ± 1.29	10.83 ± 0.69
GCN-LSTM	33.43 ± 0.98	49.11 ± 1.28	11.01 ± 0.32
LSTM	35.93 ± 0.69	52.43 ± 1.92	13.52 ± 0.78
GCN	105.67 ± 4.28	154.73 ± 5.12	58.37 ± 2.45
Transformer	33.95 ± 0.59	50.42 ± 1.01	10.94 ± 0.47

从表 2 和表 3 的结果可以看出，MGAPF 模型在 10 min 和 30 min 两个预测粒度下均取得最优性能。在 10 min 预测任务中，相较于次优模型 ResGCN-Transformer，MGAPF 的 MAE、RMSE 和 MAPE 分别下降约 4.8%、3.9%和 4.5%；在 30 min 预测任务中，MAPE 的相对降幅进一步扩大至约 7.8%。同时，MGAPF 在不同实验轮次下的误差标准差整体较小，表明模型具有较好的稳定性与泛化能力。为进一步分析预测目标时间粒度变化对模型性能的影响，补充 5、10、15 min 和 30 min 粒度下的敏感性分析，结果如表 4 所示。

由表 4 可知，预测时间粒度变化会同时改变客流聚合规模和序列波动特征。随着粒度由 5 min 增大至 30 min，平均单步进站客流明显增大，MAE 与 RMSE 整体随之上升；而 MAPE 总体下降，说明较大时间粒度

表 4 不同预测时间粒度下 MGAPF 模型性能敏感性分析

Table 4 Sensitivity analysis of MGAPF at different prediction time granularities

预测粒度	平均单步进站客流/人次	MAE	RMSE	MAPE/%
5 min	69	14.06	21.43	21.60
10 min	137	17.11	26.09	16.93
15 min	205	20.13	30.71	14.80
30 min	404	29.02	44.31	9.52

下随机波动被一定程度平滑, 相对误差降低, 但时间分辨率也随之下降。5 min 粒度能够反映更细的短时变化, 但低客流和局部波动影响更明显; 15 min 粒度在误差水平和时间分辨率上总体介于 10 min 与 30 min 之间。由此可见, 预测时间粒度会影响模型误差表现, 本文选取 10 min 与 30 min 作为主实验粒度, 能够分别代表短时运营响应和较稳定趋势预测两类型典型场景。

从不同类型模型的结果表现来看, GCN 由于仅建模空间结构, 缺乏时间动态特征表达能力, 在两种预测粒度下均表现出较大的预测误差, 且随着预测时间粒度增大, 误差进一步放大。LSTM 与 Transformer 虽然能够有效建模时间序列动态变化, 但未考虑地铁站点之间的空间关联关系, 因此预测能力受到限制。GCN-LSTM 与 GCN-Transformer 通过结合图卷积与时间建模模块, 实现了基础的时空联合建模, 预测性能相比单一时间模型有所提升, 但在多尺度时间特征表达方面仍存在不足。ResGCN-Transformer 进一步引入多尺度图卷积与并行时间分支, 实现了初步的多粒度建模能力, 但其不同时间粒度特征仍采用固定拼接方式进行融合, 难以根据不同预测场景动态调整各时间尺度的重要性。

相比之下, MGAPF 通过可学习的粒度加权机制自适应融合多尺度时序特征, 并利用图注意力机制刻画邻接站点间空间影响强度的差异。同时, 节假日、周末及天气等外部特征信息的引入, 有助于模型识别不同日期属性和外部环境下的客流波动规律。配合时空门控融合与残差连接设计, MGAPF 在时空联合建模中表现出更好的预测精度与稳定性。

在实验环境配置下, MGAPF 模型参数规模为 0.224 M, 单批次推理时间 218 ms; ResGCN-Transformer 与 GCN-Transformer 的参数规模分别为 0.201 M 和 0.152 M, 对应推理时间分别为 183 ms 和 146 ms。结果表明, MGAPF 在提升预测精度的同时, 仅带来有限的复杂度增加, 能够满足地铁短时客流预测的应用需求。

3.3 结果分析

表 2 和表 3 已从 80 个有效站点整体层面对模型预测性能进行了验证。为进一步展示模型在不同站点功能与日期属性下的拟合效果, 选取火车东站(站点 15)与钱江路站(站点 46)作为典型案例, 并展示其在 2019 年 1 月 11 日(周五, 工作日)与 1 月 12 日(周六, 休息日) 2 天的 10 min 粒度预测结果, 分别如图 3、图 4 所示。图中横轴为时间步序号, 单位为步; 每步对应 10 min, 全日共 144 个时间步, 其中 0~143 步分别对应当日 0:00—23:50; 纵轴为对应时间步内的进站客流量。火车东站为交通枢纽型高客流站点, 客流规模较大且高峰特征明显; 钱江路站为商住混合型换乘站, 客流变化受通勤、商业活动及周末出行等因素共同影响。上述站点和日期分别覆盖枢纽型与复合功能型站点、工作日与周末两类型典型场景, 可用于辅助分析模型在不同客流节律下的适应能力。

从客流形成机制看, 火车东站受铁路到发、城市通勤和换乘接驳等因素影响, 进站客流在工作日早晚通勤时段、铁路集中到发时段及周末出行集中时段均出现快速上升; MGAPF 中的多粒度卷积分支有助于

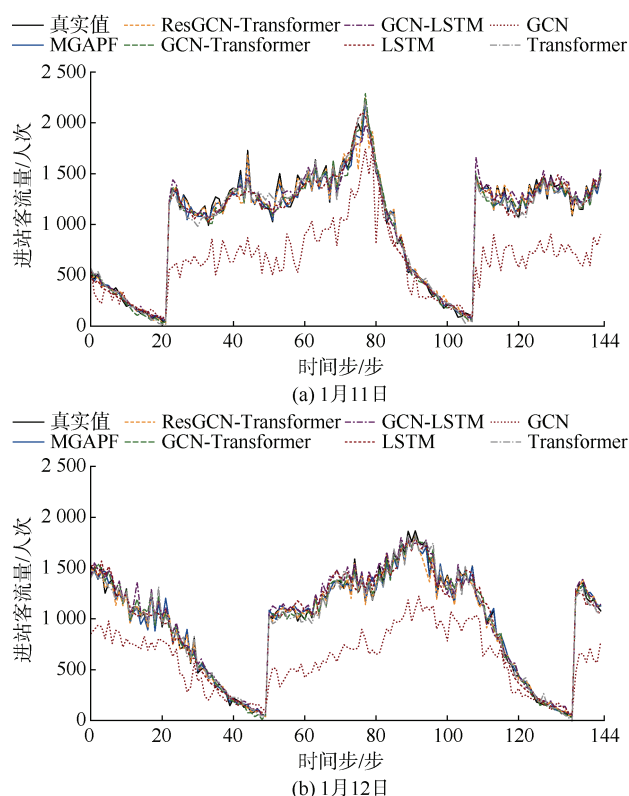


图 3 火车东站进站客流预测效果

Figure 3 Inbound passenger flow prediction for East Railway Station

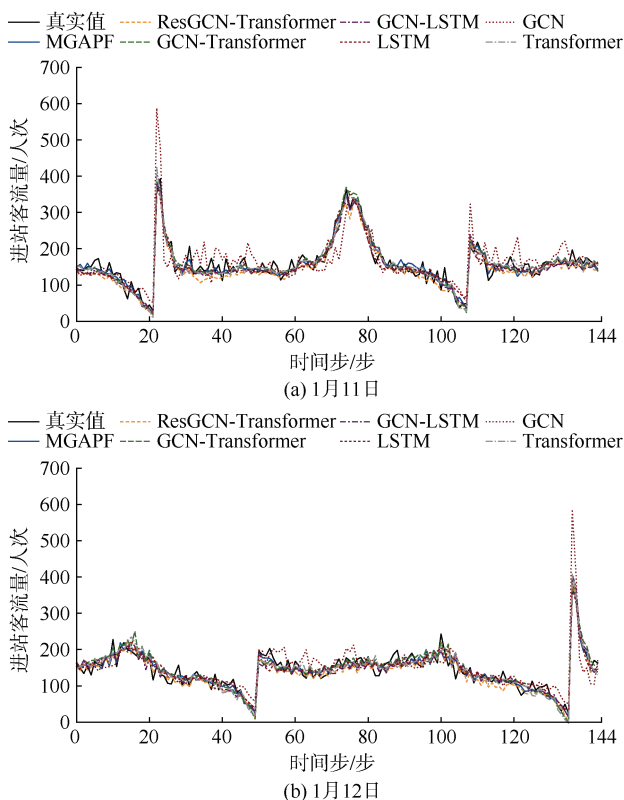


图4 钱江路站进站客流预测效果
Figure 4 Inbound passenger flow prediction for Qianjiang Road Station

捕捉这种短时突增与回落过程,图注意力模块则能够在物理邻接关系约束下刻画相邻站点对目标站点客流变化的影响。钱江路站兼具居住、商业和换乘功能,工作日客流更受通勤节律影响,周末则更容易体现商业出行和休闲出行特征;模型引入的周末、节假日及天气等外部特征,有助于区分不同日期属性下的客流分布差异。结果表明,MGAPF在高峰突增、局部转折和周末非周期波动区间均保持较好的跟踪能力。

尽管在部分低客流时段存在轻微高估现象,但整体拟合趋势与真实客流数据保持较高一致性,特别是在客流快速变化区间,MGAPF表现出更快的响应能力与更稳定的预测精度。

综合来看,MGAPF在典型枢纽型和复合功能型站点、工作日和周末等场景下均保持较好的预测效果,说明多粒度时间建模、图注意力机制和外部信息融合能够有效提升复杂场景下的短时进站客流预测能力。本文原始数据包含进出站状态字段,出站客流可按类似方式构建预测序列,但实验对象限定为进站客流;换乘客流则需进一步结合OD识别、路径推断或站内换乘监测数据,相关问题将在后续研究中展开。

4 结论

本文面向地铁短时进站客流预测任务,提出融合多粒度感知与图注意力机制的MGAPF模型。模型在时间维度通过多尺度卷积与粒度注意力机制实现不同时间尺度特征的自适应融合,在空间维度采用图注意力机制动态学习站点间的空间关联强度,进一步引入节假日、周末及天气等外部时间上下文信息以增强模型对复杂客流变化的表达能力。研究主要结论如下:

1) 基于杭州市地铁运营数据的实验结果表明,MGAPF在10 min与30 min预测任务下均取得较优性能,MAPE分别为16.93%与9.52%,相较于对比模型具有更高的预测精度和稳定性。

2) 典型站点分析进一步表明,MGAPF能够较好地捕捉客流高峰、局部转折及复杂波动趋势,在不同站点类型和不同日期属性下均具有较好的适应性。同时,模型参数规模与推理时间保持在合理范围内,具备一定的工程应用基础。

3) 未来工作将进一步结合更新年份、更长时间跨度和更多城市轨道交通网络数据,验证模型在不同线网规模、客流结构和运营时期下的泛化能力,并探索动态图结构和更丰富外部因素在短时客流预测中的作用。

参考文献

- [1] 凌晨,许心越,叶子扬,等.城市轨道交通短时客流预测研究综述及展望[J].铁道运输与经济,2025,47(7):44-55.
Ling Chen, Xu Xinyue, Ye Ziyang, et al. Review and prospects of short-term passenger flow prediction in urban rail transit[J]. Railway Transport and Economy, 2025, 47(7): 44-55.
- [2] 胡松,常征,杨贝,等.基于PSO-SVR模型的轨道交通客流预测[J].都市轨道交通,2025,38(5):50-56.
Hu Song, Chang Zheng, Yang Bei, et al. Urban rail passenger flow prediction using particle swarm optimization-tuned SVR model[J]. Urban Rapid Rail Transit, 2025, 38(5): 50-56.
- [3] Zhan Shuguang, Cai Yi, Xiu Cong, et al. Parallel framework of a multi-graph convolutional network and gated recurrent unit for spatial-temporal metro passenger flow prediction[J]. Expert Systems with Applications, 2024, 251: 123982.
- [4] Wang Tao, Song Jianjun, Zhang Jing, et al. Short-term metro passenger flow prediction based on hybrid spatiotemporal extraction and multi-feature fusion[J]. Tunnelling and Underground Space Technology, 2025, 159: 106491.
- [5] 张思楠,李树彬,曹永军.基于LSTM-Transformer的城市轨道交通短时客流预测[J].物流科技,2024,47(14):103-106.
Zhang Sinan, Li Shubin, Cao Yongjun. Short-term passenger flow prediction of urban rail transit based on LSTM-transformer[J]. Logistics Sci-tech, 2024, 47(14): 103-106.
- [6] 帅春燕,谢亚威,单君,等.基于SSA-SVR模型的城市

- 轨道交通短时进站客流预测[J]. 都市快轨交通, 2022, 35(5): 76-83.
- Shuai Chunyan, Xie Yawei, Shan Jun, et al. Prediction of short-term inbound passenger flow of urban rail transit based on the singular spectrum analysis and support vector regression model[J]. Urban Rapid Rail Transit, 2022, 35(5): 76-83.
- [7] 杨越迪, 潘保霏, 刘军, 等. 多分辨率多粒度时空特征提取的轨道交通短时 OD 客流预测方法[J]. 铁道学报, 2024, 46(11): 1-11.
- Yang Yuedi, Pan Baofei, Liu Jun, et al. A short-term OD passenger flow prediction method for rail transit based on multi-resolution and multi-granularity spatio-temporal feature extraction[J]. Journal of the China Railway Society, 2024, 46(11): 1-11.
- [8] 杜姿晨, 郑长江, 郑树康, 等. 基于 ResNet-GCN-Transformer 的多时间粒度地铁短时客流预测[J]. 交通运输工程与信息学报, 2026, 24(1): 64-79.
- Du Zichen, Zheng Changjiang, Zheng Shukang, et al. ResNet-GCN-Transformer-based prediction of short-time passenger flow for multi-time granularity subway[J]. Journal of Transportation Engineering and Information, 2026, 24(1): 64-79.
- [9] Lu Wenbo, Zhang Yong, Vu H L, et al. A novel integrative prediction framework for metro passenger flow[J]. Journal of Intelligent Transportation Systems, 2026, 30(3): 490-515.
- [10] 刘晓磊, 段征宇, 余庆, 等. 基于图卷积循环神经网络的城市轨道交通客流预测[J]. 华南理工大学学报(自然科学版), 2022, 50(3): 21-27.
- Liu Xiaolei, Duan Zhengyu, Yu Qing, et al. Passenger flow forecast of urban rail transit based on graph convolution and recurrent neural network[J]. Journal of South China University of Technology (Natural Science Edition), 2022, 50(3): 21-27.
- [11] Liu Yang, Liu Zhiyuan, Jia Ruo. DeepPF: a deep learning based architecture for metro passenger flow prediction[J]. Transportation Research Part C: Emerging Technologies, 2019, 101: 18-34.
- [12] 李卓, 李怡, 巨玉祥, 等. 结合时序直觉模糊粗糙集和 TCN-BiGRU 的城市轨道交通客流预测[J]. 兰州交通大学学报, 2025, 44(5): 107-116.
- Li Zhuo, Li Yi, Ju Yuxiang, et al. Integration of temporal intuitionistic fuzzy rough sets and TCN-BiGRU for urban rail transit passenger flow forecasting[J]. Journal of Lanzhou Jiaotong University, 2025, 44(5): 107-116.

(编辑: 傅依萱)

(上接第 152 页)

- [5] 孙波. 克诺尔空气制动系统原理分析[J]. 铁道技术监督, 2014, 42(6): 46-48.
- Sun Bo. Analyzing the principles of knorr air braking system[J]. Railway Quality Control, 2014, 42(6): 46-48.
- [6] 赵文文, 魏晓东, 陈特放. 基于 FTA 和模糊 Petri 网的城轨列车制动系统可靠性分析[J]. 铁道科学与工程学报, 2021, 18(5): 1158-1166.
- Zhao Wenwen, Wei Xiaodong, Chen Tefang. Reliability analysis of urban rail train braking system based on FTA and fuzzy Petri net[J]. Journal of Railway Science and Engineering, 2021, 18(5): 1158-1166.
- [7] 张济民, 刘志刚, 沈霞. 基于深度学习的轨道交通制动系统故障预测与健康研究综述[J]. 交通运输工程学报, 2022, 22(3): 1-15.
- Zhang Jimin, Liu Zhigang, Shen Xia. Review of fault prediction and health management for rail transit braking systems based on deep learning[J]. Journal of Traffic and Transportation Engineering, 2022, 22(3): 1-15.
- [8] 胡辉, 王俊, 蔡煜. 基于数据驱动的城轨车辆制动系统状态评估与剩余寿命预测[J]. 中国铁道科学, 2023, 44(1): 145-154.
- Hu Hui, Wang Jun, Cai Xuan. Data-driven condition assessment and remaining useful life prediction of urban rail vehicle braking systems[J]. China Railway Science, 2023, 44(1): 145-154.
- [9] 李国栋, 陈小平, 杨帆. 基于数字孪生的高速列车制动系统故障诊断方法[J]. 机械工程学报, 2021, 57(10): 43-52.
- Li Guodong, Chen Xiaoping, Yang Fan. Fault diagnosis method for high-speed train braking systems based on digital twin[J]. Journal of Mechanical Engineering, 2021, 57(10): 43-52.
- [10] 王寿峰, 徐广增. 城轨车辆车控制动系统与架控制动系统对比分析[J]. 铁道机车车辆, 2019, 39(1): 114-117.
- Wang Shoufeng, Xu Guangzeng. Contrastive analysis between vehicle control brake system and bogie control brake system of urban rail vehicle[J]. Railway Locomotive & Car, 2019, 39(1): 114-117.
- [11] 杜守忠, 吴志明. EP2002 制动系统的控制过程研究[J]. 电力机车与城轨车辆, 2013, 36(4): 35-38.
- Du Shouzhong, Wu Zhiming. Control process study of brake system EP2002[J]. Electric Locomotives & Mass Transit Vehicles, 2013, 36(4): 35-38.
- [12] 万宇. 广州地铁 4 号线列车 EP2002 制动系统介绍及故障分析[J]. 铁道机车车辆, 2010, 30(6): 65-69.
- Wan Yu. EP2002 brake control system introduction and failure analysis for Guangzhou line No.4[J]. Railway Locomotive & Car, 2010, 30(6): 65-69.
- [13] 史富强. 西安地铁 1 号线车辆 EP2002 制动系统[J]. 电子设计工程, 2013, 21(24): 187-190.
- Shi Fuqiang. Xi'an Metro Line 1 vehicle braking system of EP2002[J]. Electronic Design Engineering, 2013, 21(24): 187-190.

(编辑: 王艳菊)