

文章编号: 1003-1421(2026)10-0132-12

引用格式: 朱沁萱, 陈韬, 张培华, 等. 基于多特征改进 CNN-BiLSTM 的地铁车站级短时客流预测[J]. 铁道运输与经济, 2026, 48(10): 132-143, 171.

ZHU Qinxuan, CHEN Tao, ZHANG Peihua, et al. Short-Term Passenger Flow Prediction at Metro Station Level Based on Multi-Feature Improved CNN-BiLSTM[J]. Railway Transport and Economy, 2026, 48(10): 132-143, 171.

基于多特征改进 CNN-BiLSTM 的地铁车站级 短时客流预测

朱沁萱^{1, 2}, 陈 韬^{1, 3, 4}, 张培华⁵, 谢欣怡⁶, 朱逸夫⁵

(1. 西南交通大学 交通运输与物流学院, 四川 成都 611756; 2. 成都地铁运营有限公司 地保监护中心, 四川 成都 610094; 3. 西南交通大学 综合交通运输智能化国家地方联合工程实验室, 四川 成都 611756; 4. 西南交通大学 综合交通大数据应用技术国家工程实验室, 四川 成都 611756; 5. 成都地铁运营有限公司 运营三公司, 四川 成都 610094; 6. 成都地铁运营有限公司 工电维保公司, 四川 成都 610094)

摘 要: 随着城市轨道交通线网的不断扩张, 地铁线路日益纵横交错, 车站数量也在持续增加。提高客流短时预测精度, 对制定客流引导策略尤为重要, 因此, 研究将通过客流数据分析、损失函数优化, 构建结合集成学习策略的改进 CNN-BiLSTM 模型实现车站级的短时客流预测。首先通过区分普通站和换乘站的数据分类分析, 设计基于自相关函数的滑动窗口数据划分方法, 并构建有限数据集下的多特征增强数据集; 其次通过正余弦函数将时间特征转化为周期性循环特征, 并设计 MSE 与 R^2 的加权组合损失函数作为优化目标; 同时模型融合多尺度卷积、时空注意力机制及 BiLSTM/GRU 模块以增强预测能力; 进一步引入集成学习策略, 通过 Adaboost 对 5 个改进模型的预测结果加权计算, 获得更稳定的客流预测值; 最后实证分析表明, 构建的多特征改进 CNN-BiLSTM 模型具有优异的预测效果。

关键词: 短时客流预测; 多特征; 多尺度卷积; 改进 CNN-BiLSTM; 注意力机制; 集成学习
中图分类号: U-9 **文献标识码:** A **DOI:** 10.16668/j.cnki.issn.1003-1421.20250310002

Short-Term Passenger Flow Prediction at Metro Station Level Based on Multi-Feature Improved CNN-BiLSTM

ZHU Qinxuan^{1, 2}, CHEN Tao^{1, 3, 4}, ZHANG Peihua⁵, XIE Xinyi⁶, ZHU Yifu⁵

(1. School of Transportation and Logistics, Southwest Jiaotong University, Chengdu 611756, Sichuan, China; 2. Subway Control Protection Area Monitoring Center, Chengdu Metro Operation Co., Ltd., Chengdu 610094, Sichuan, China; 3. National United Engineering Laboratory of Integrated and Intelligent Transportation, Southwest Jiaotong University, Chengdu 611756, Sichuan, China; 4. National Engineering Laboratory for Big Data Application in Integrated Transportation, Southwest Jiaotong University, Chengdu 611756, Sichuan, China; 5. The Third Operation Company, Chengdu Metro Operation Co., Ltd., Chengdu 610094, Sichuan, China; 6. Industrial and Electrical Maintenance Company, Chengdu Metro Operation Co., Ltd., Chengdu 610094, Sichuan, China)

Abstract: With the continuous expansion of the urban rail transit network, metro lines become increasingly interconnected, and the number of stations continues to increase. Improving the accuracy of short-term

收稿日期: 2025-03-10

通信作者: 陈韬(1981—), 女, 广西柳州人, 西南交通大学交通运输与物流学院副教授。

基金项目: 国家重点研发计划项目(2022YFB4300502-1)

passenger flow prediction is crucial for formulating passenger flow guidance strategies; therefore, this paper constructed an improved CNN-BiLSTM model combining ensemble learning strategies through passenger flow data analysis and loss function optimization to achieve station-level short-term passenger flow prediction. First, through the classification analysis of data distinguishing regular stations and transfer stations, a sliding window data partitioning method based on the autocorrelation function was designed, and a multi-feature enhanced dataset under a limited dataset was constructed. Second, the time features were transformed into periodic cyclic features through sine and cosine functions, and a weighted combination loss function of MSE and R^2 was designed as the optimization objective. Meanwhile, the model integrated multi-scale convolution, spatiotemporal attention mechanism, and BiLSTM/GRU module to enhance the prediction capability. Furthermore, the ensemble learning strategy was introduced, and Adaboost was used to weight and calculate the prediction results of five improved models to obtain more stable passenger flow prediction values. Finally, the empirical analysis shows that the constructed multi-feature improved CNN-BiLSTM model has excellent prediction effects.

Keywords: Short-Term Passenger Flow Prediction; Multi-Feature; Multi-Scale Convolution; Improved CNN-BiLSTM; Attention Mechanism; Ensemble Learning

0 引言

随着城市规模的持续扩张,城市轨道交通线网变得愈发复杂,普通站和换乘站的数量也随之增多。这使得地铁车站的客流特征更难以预测,给车站管理带来了新的挑战。因此,建立合理的客流引导策略,对于地铁车站的管理工作具有重要的指导性意义。而客流的时间序列数据为低维数据,如何从有限维度的数据中深度挖掘数据的全部特征,从而提高预测精度和地铁场景的泛化性是值得研究的。

地铁客流受多种时空、环境、社会及运营因素的综合影响,在数据预处理和特征分析阶段,相关研究多从时序特征、空间特征或是外部特征等入手,以数据特征工程或数据增强的方式从多方面深入挖掘客流数据特征。为了利用轨道交通数据进行客流特征分析以及城市功能划分匹配,曹敬洽等^[1]基于不平衡系数指标,对北京市地铁各站点的乘客出行时空分布特征开展聚类分析,探讨居民出行行为规律与城市功能空间结构之间的不平衡关联性,旨在为城市轨道交通运营组织优化策略提供理论依据。蔡昌俊等^[2]基于周期性分析,结合周期与短期相关性的乘积特性,构建乘积 ARIMA 模型预测地铁客流量。对于节假日一般会引起地铁客运显著的客流量增长,Yao 等^[3]构建了节假日客流分布预测模型以应对节假日乘客出行行为的特殊性;Zhang 等^[4]提出一种基于深度学习的时空注意力融合网络来深入研究元旦期间城市轨道交通系统的短期客流预测问题。

短时客流的预测算法早期以基于数理统计的时间序列模型为主,例如 Smith 等^[5]针对高速公路交通

流量预估未来 15 min 的交通流量的问题,研究开发并测试了 4 种模型:历史平均模型、时间序列模型、神经网络模型和非参数回归模型。张丽莉等^[6]将用地类型与随机森林模型(Random Forest, RF)结合的方式对地铁出站客流量进行预测。Yao 等^[7]使用随机森林回归、线性回归、贝叶斯岭回归(Bayesian Ridge Regression)、弹性网络(Elastic Network)和梯度提升回归(Gradient Boosting Regression) 5 种模型进行对比实验,结果表明多源数据和随机森林回归模型的预测精度最为理想。在机器学习领域,自深度学习再次爆发以来,各类深度学习模型特别是组合模型凭借其良好的性能在各个领域大放异彩,交通预测方向同样迎来革新。Ma 等^[8]提出图卷积(GCN)和双向长短期记忆网络(BiLSTM)结合的短时客流预测模型,划分近期、日周期、周周期时间模式,分别建模其时空依赖,最终通过参数矩阵融合输出结果。Zhang 等^[9]提出一种结合 Transformer 网络和长短期记忆网络(LSTM),同时融合 GCN 的混合模型。Zeng 等^[10]提出一种基于完全自适应噪声集合经验模态分解(CEEMDAN)与 LSTM 的组合预测模型,以提高城市轨道交通短时客流预测精度。其他短时客流预测方法还有 CNN-LSTM-Attention^[11]组合模型或 Random Forest-LSTM^[12]结合模型等。

综合来看,地铁短时客流预测算法历经持续的优化迭代,当前已迈入以深度学习技术为核心的新阶段。研究中将系统考量地铁网络客流数据的特征属性、模型算法的精确度及相关预测性能指标等关键要素,通过构建分类车站的深度学习预测模型开展进出站客流预测研究,并最终通过实验结果进行验证分析。

1 数据特征增强及划分策略分析

1.1 数据分析及特征增强

客流数据为成都地铁系统2024年1—8月客流记录数据,统计时段为每天6:00—24:00,以1 h为采样颗粒,每个车站的原始数据为4 392条记录。划分数据按照滑动窗口的方式,时间步长根据自相关系数ACF^[13]计算得出,将70%的数据集归为训练集,将15%的数据集归为验证集,余下的15%作为测试集。并且在划分数据集时,设置偏移量等于时间步长,确保滑动窗口能够正确地覆盖到所有相关的历史数据。根据地铁车站的一般区分,将其分为普通站和换乘站2种类型。另外,考虑到客流不均衡性和周期性特点,用0,1,2,3,4,5,6来代表周一到周日,用0和1分别表示工作日和节假日,同时将每天的时间用6—23的整数划分为间隔1 h的时间节点。普通站06:00—24:00数据样例如表1所示,换乘站06:00—24:00数据样例如表2所示。

表1 普通站06:00—24:00数据样例
Tab.1 Data samples of regular station during 06:00—24:00

序号	转化时段	星期	车站	进站客流/人次	出站客流/人次	换乘客流/人次	工作日/节假日
1	6	0	三岔	187	26	0	1
2	7	0	三岔	327	235	0	1
⋮	⋮	⋮	三岔	⋮	⋮	⋮	⋮
18	23	0	三岔	49	253	0	1

将上述的地铁站点特征进行规范化整理,如公式(1)所示。

$$flow_i = \begin{bmatrix} in_{F_1}^i & out_{F_1}^i & t_{F_1}^i & v_{F_1}^i & t_{\max F_1}^i & h_1^i & w_1^i & H_1^i \\ in_{F_2}^i & out_{F_2}^i & t_{F_2}^i & v_{F_2}^i & t_{\max F_2}^i & h_2^i & w_2^i & H_2^i \\ in_{F_3}^i & out_{F_3}^i & t_{F_3}^i & v_{F_3}^i & t_{\max F_3}^i & h_3^i & w_3^i & H_3^i \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ in_{F_n}^i & out_{F_n}^i & t_{F_n}^i & v_{F_n}^i & t_{\max F_n}^i & h_n^i & w_n^i & H_n^i \\ in_{F_{n+1}}^i & out_{F_{n+1}}^i & & & & & & \end{bmatrix} = \quad (1)$$

式中: $flow_i$ ($i = 1, 2, \dots, n$) 为第 i 个站点的客流时

间序列,即为模型输入; $in_{F_n}^i$ 为车站 i 在时间点 n 的小时进站客流量,人次; $out_{F_n}^i$ 为车站 i 在时间点 n 的小时出站客流量,人次; $t_{F_n}^i$ 为车站 i 在时间点 n 的小时换乘客流量,人次; $v_{F_n}^i$ 为车站 i 在时间点 n 的小时进出站总客流量,人次; $t_{\max F_n}^i$ 为车站 i 在一天中的最大换乘客流量,人次; h_n^i 为车站 i 在时间点 n 的客流量对应的小时时刻; w_n^i 为车站 i 在时间点 n 对应星期几; H_n^i 为车站 i 在时间点 n 是否为节假日; $in_{F_{n+1}}^i$ 为预测 $n+1$ 时刻的进站量,人次; $out_{F_{n+1}}^i$ 为预测 $n+1$ 时刻的出站量,人次。

研究中,设计的特征工程通过挖掘时间维度的周期性、趋势性和动态关联性增强模型对复杂规律的捕捉能力。具体而言,周期性的时间特征通过三角函数转化为周期性坐标。时段转化算法为 $h_sin = \sin(2\pi \times \text{hour}/24)$, $h_cos = \cos(2\pi \times \text{hour}/24)$ 。星期数转换算法为 $w_sin = \sin(2\pi \times (\text{week} + 1)/7)$, $w_cos = \cos(2\pi \times (\text{week} + 1)/7)$ 。如此,既保留了时间的周期性邻近关系,又避免了独热编码导致的高维稀疏问题,这里 $\text{week} + 1$ 是因为在前述 week 进行编码时是从0开始的,此处转化为周期性坐标进行修正。此外,进出站客流的交互特征,进出站比率、和值、差值进一步扩充了特征维度的信息密度,通过计算进站与出站客流的比率 $in_out_r = \text{进站客流量}/\text{出站客流量}$,和值 $in_out_r = \text{进站客流量} + \text{出站客流量}$,差值 $in_out_d = \text{进站客流量} - \text{出站客流量}$,可量化车站瞬时客流的动态平衡关系。这些特征使得模型输入增加7个维度,如公式(2)所示,揭示了原始数据中时间按周期和数据交互的逻辑,同时避免了冗余特征对模型鲁棒性的干扰。通过上述分析,结合前述对输入向量各个维度的规范化整理,将新增的7维输入特征与公式(1)中的8维输入特征水平并置成为15维特征,建立新的客流时间序列矩阵。最后需要说明的是,建立的车站级客流时间序列矩阵中的数据都已经进行离差标准化。

表2 换乘站06:00—24:00数据样例
Tab.2 Data samples of transfer station during 06:00—24:00

序号	转化时段	星期	车站	进站客流/人次	出站客流/人次	换乘客流/人次	最大换乘客流/人次	工作日/节假日
1	6	0	火车南	798	493	1 183	11 471	1
2	7	0	火车南	1 086	1 529	3 507	11 471	1
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
18	23	0	火车南	861	757	2 023	11 471	1

$$\begin{bmatrix} h_sin_1^i & h_cos_1^i & w_sin_1^i & w_cos_1^i & in_out_r_1^i & in_out_s_1^i & in_out_d_1^i \\ h_sin_2^i & h_cos_2^i & w_sin_2^i & w_cos_2^i & in_out_r_2^i & in_out_s_2^i & in_out_d_2^i \\ h_sin_3^i & h_cos_3^i & w_sin_3^i & w_cos_3^i & in_out_r_3^i & in_out_s_3^i & in_out_d_3^i \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ h_sin_n^i & h_cos_n^i & w_sin_n^i & w_cos_n^i & in_out_r_n^i & in_out_s_n^i & in_out_d_n^i \end{bmatrix} \quad (2)$$

式中： $h_sin_n^i$ 为车站 i 在时间点 n 的小时转化正弦坐标值； $h_cos_n^i$ 为车站 i 在时间点 n 的小时转化余弦坐标值； $w_sin_n^i$ 为车站 i 在时间点 n 的星期转化正弦坐标值； $w_cos_n^i$ 为车站 i 在时间点 n 的星期转化余弦坐标值； $in_out_r_n^i$ 为车站 i 在时间点 n 的进站与出站客流的比率； $in_out_s_n^i$ 为车站 i 在时间点 n 的进站与出站客流的和值，人次； $in_out_d_n^i$ 为车站 i 在时间点 n 的进站与出站客流的差值，人次。

1.2 数据划分策略分析

在上述理论的基础上考虑到客流时间序列的特殊性，在处理数据集时除了按比例划分训练集、验证集、测试集外，还需重点考虑选择合适的时间步长。用自相关系数(ACF, Auto-Correlation Function)

来确定最优模型预测时间步长，其计算如公式(3)所示。

$$\rho_{x_t, x_{t+h}} = \frac{\sum_{i=1}^{n-h} (x_i - \bar{X})(x_{i+h} - \bar{X})}{\sum_{i=1}^n (x_i - \bar{X})^2} \quad (3)$$

式中： $\rho_{x_t, x_{t+h}}$ 为自相关系数； $\bar{X}$ 为时间序列的均值； h 为时间步长，即滞后数。

根据地铁车站的真实数据，分别计算普通站与换乘站的自相关系数，普通站进出站客流量自相关图如图1所示，换乘站进出站客流量自相关图如图2所示。

根据计算结果和自相关图，普通站的进站客流和出站客流在滞后阶数为18时的自相关系数约为

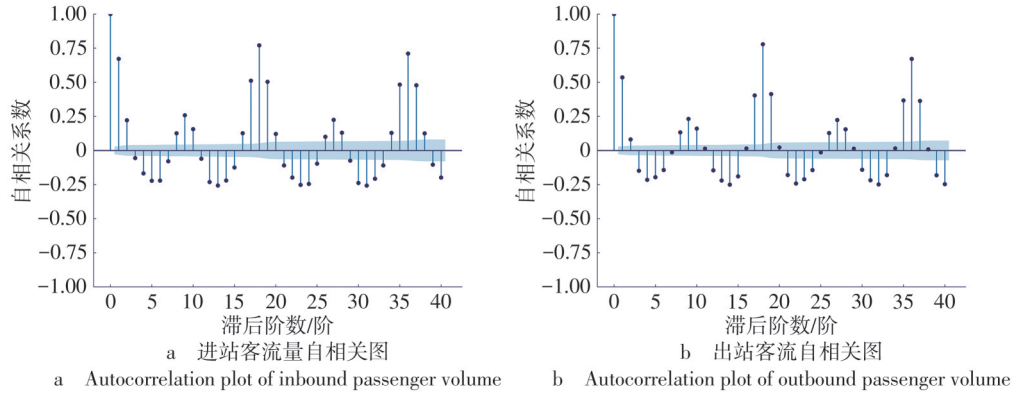


图1 普通站客流量自相关图

Fig.1 Autocorrelation plot of passenger volume at regular station

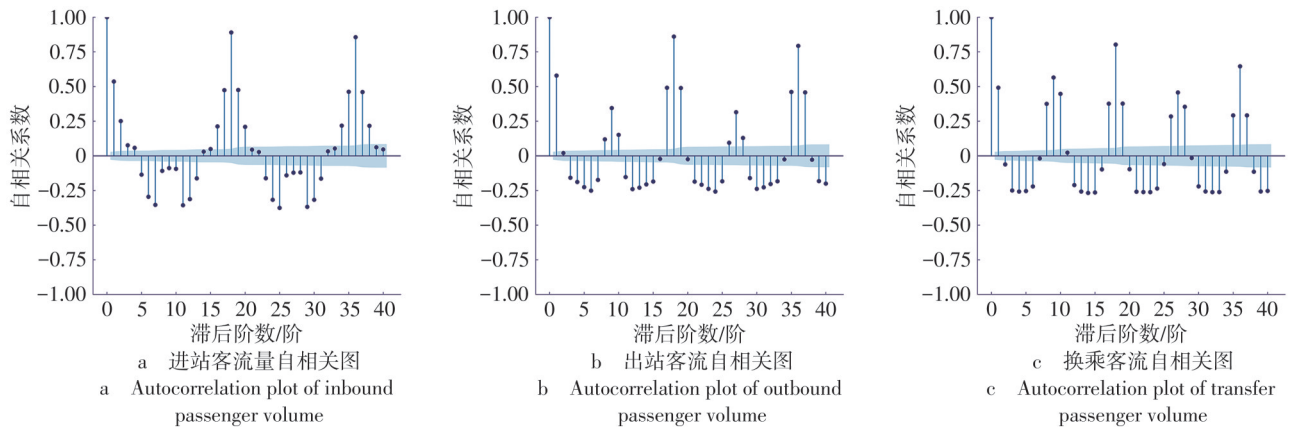


图2 换乘站客流量自相关图

Fig.2 Autocorrelation plot of passenger volume at transfer station

0.770 6和0.778 4,可以得到进出站客流量自相关系数最大滞后阶数都是18,这也跟实际情况一天的18个时段是相吻合的,因此预测时间步长采用18;同样,换乘站的最大滞后阶数都是18。则2种站点都用18为步长,采用滑动窗^[14]的方式进行模型预测,滑动窗口示意如图3所示。一个时间窗口中存在18个时间步,这18个时间步的数据用于预测下一个时间步的进站客流和出站客流数值。

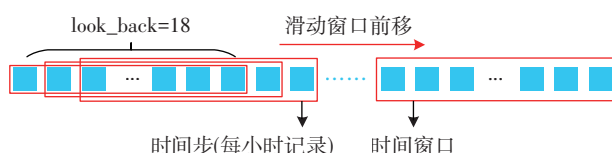


图3 滑动窗口示意
Fig.3 Schematic diagram of sliding window

2 改进CNN-BiLSTM模型构建

城市轨道交通客流特性会受到多种因素的影响,具有不均衡性、周期性^[15]和显著的潮汐现象,其中换乘站客流量相对较多,波动性较大,为做好地铁站点客流的动态管理和提前预警,提高地铁短时客流预测精度,构建合适的客流预测模型尤为重要。为了突破传统单一分层架构的局限性并增强一般卷积神经网络-双向长短期记忆网络(CNN-BiLSTM)模型的预测能力,研究的核心在于通过混合架构与注意力机制的协同设计,增强复合模型对多尺度时空依赖的捕捉能力,同时应用模块化设计平衡复杂度与计算资源。

2.1 多尺度卷积

CNN以局部连接的特点,只对输入数据的特定区域进行操作,在轨道交通运输中,有助于识别不同时间段内的乘客流动规律,并且可以自动提取数据特征,无需人工干预,通过堆叠多个卷积层,CNN可以在不同尺度上捕获信息,例如,在轨道交通系统中,既存在短期波动(如高峰时段的变化),也存在长期趋势(如周期性变化)。首先采用类似Inception^[16]的多尺度卷积模块,设计了包含1×1,3×3,5×5的卷积分支结构,可并行提取类似短期波动的高频局部特征、半天或一天的中等尺度周期性以及周、月模式的长周期趋势。

研究设计的多尺度卷积模块如图4所示,采用了3层并行分支结构,每个分支由2次连续的同尺

寸卷积操作与批标准化(Batch Normalization, BN)构成,最终通过通道拼接融合多尺度特征。输入层同时进入3个分支,首层卷积提取基础特征后,第2层卷积深化特征抽象,形成“输入→Conv1→Conv2→BN”的双层非线性变换路径,较经典Inception的单卷积结构增强了特征表达能力。同时,归因于其双卷积结构对客流时序中非线性关系的精细刻画,以及多尺度融合对站点客流差异的显式建模,较传统单分支CNN神经网络或浅层Inception神经网络更具时空特征解耦优势。

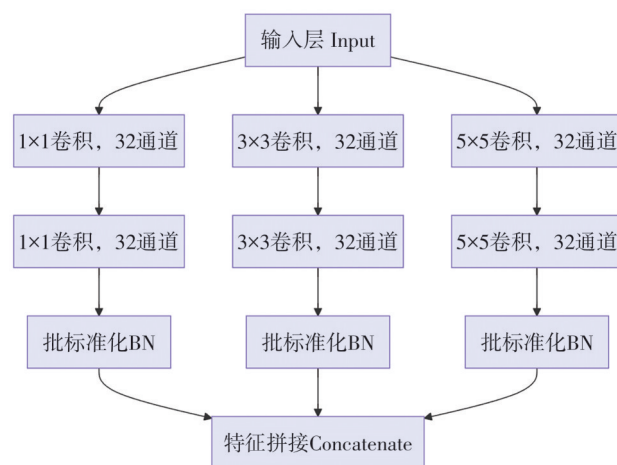


图4 多尺度卷积模块
Fig.4 Multi-scale convolution module

2.2 时空注意力机制

在模型中引入时间和空间注意力机制,这是一种通过动态调整特征权重来增强深度学习模型对关键时空信息捕获能力的技术,其核心思想是模拟人类认知系统对重要信息的聚焦过程。在地铁客流预测场景中,时间注意力机制通过可学习的权重矩阵自动识别客流时序中的关键时间点,而空间注意力机制则着重捕捉路网中关键节点的客流变化特征。

2.2.1 空间注意力机制

空间注意力机制如图5所示,空间注意力模块是使用最大池化和平均池化来获取每个空间位置的最大值和平均值。由于卷积之后会产生多个通道,

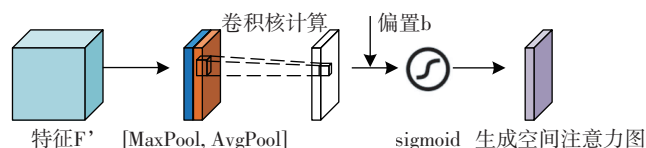


图5 空间注意力机制
Fig.5 Spatial attention mechanism

空间注意力会在每一个特征点的通道上进行最大池化和平均池化操作,得到2个池化结果矩阵后,拼接2个矩阵,再通过一个卷积层和 Sigmoid 函数来学习每个空间位置的权重。最后,将权重应用于特征图上的每个空间位置,以产生具有增强空间重要性的特征。

2.2.2 时间注意力机制

时间注意力机制旨在通过自适应学习不同时间步的权重,突出任务重要的时间点或时间段,抑制无关或冗余的时间信息。其核心是为时间序列中的每个时间步分配一个权重,通过加权融合特征,提升模型对关键时间依赖关系的捕捉能力,时间注意力机制计算过程如图6所示。输入特征张量 $X \in R^{T \times d}$, T 为时间步长,这里为18, d 为特征维度。在具体的计算过程中, Q, K, V 在标准实现中通常通过线性变换得到,但研究中使用的是基本的缩放点积注意力,没有额外的线性投影,即 $Q = K = V$ 等于输入 X ,那么直接计算: Query $Q = X \in R^{T \times d}$, Key $K = X \in R^{T \times d}$, Value $V = X \in R^{T \times d}$ 。其他为相似度矩阵 S 、注意力权重 A 、注意力输出 Z 、残差连接 R 、层归一化 $\hat{R}$ 的计算。

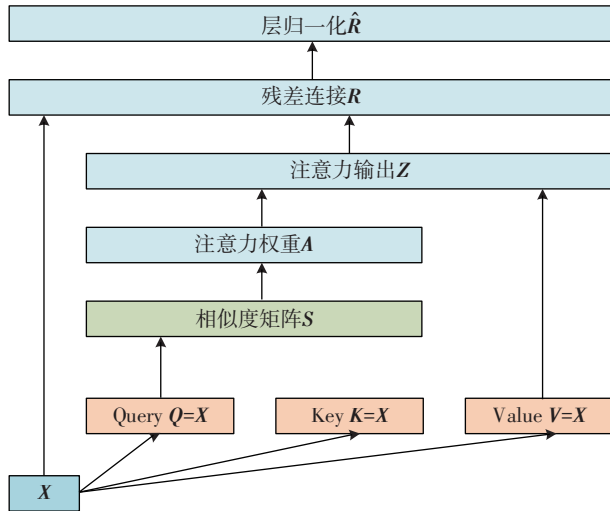


图6 时间注意力机制计算过程

Fig.6 Computation process of temporal attention mechanism

2.3 BiLSTM 及总体结构设计

长短期记忆网络(LSTM)^[17]是一种特殊类型的循环神经网络(RNN), LSTM 能够有效地学习和记住长时间间隔内的依赖关系,非常适合用于处理具有长时间关联的地铁客流预测数据。而 BiLSTM 是 LSTM

的一种变体,它通过同时处理时间序列数据的正向和反向信息来捕捉更加全面的时间依赖性, BiLSTM 包含2个 LSTM 层:一个按照正常时间顺序处理输入(前向 LSTM),另一个则逆序处理输入(后向 LSTM)^[18]。

门控循环单元结构(GRU)是 LSTM 的简化版本,将 LSTM 的3个门(输入门、遗忘门、输出门)合并为2个门(更新门、重置门),减少了参数数量和计算复杂度。GRU 与 LSTM 在应用场景上各有侧重, LSTM 在复杂序列任务中更具优势,而 GRU 因结构简洁在轻量级任务中更高效,考虑到模型已构建的 BiLSTM 模块,为了在预测的场景中的有效性,选择了 GRU 来构建全连接层之前的数据处理模块。

根据上述模型结构的改进分析,构建改进 CNN-BiLSTM 模型。模型以 CNN-BiLSTM 架构设计为基础,首先设计多尺度卷积模块,通过并行布置 $1 \times 1, 3 \times 3, 5 \times 5$ 等3种卷积分支,分别提取局部细粒度特征、中程空间关联和广域上下文信息,每个分支采用2次同尺寸卷积层与批标准化操作,经通道维度拼接形成融合多尺度感知能力的特征图。随后引入空间注意力机制,将全局平均池化与最大池化特征拼接后,通过 7×7 卷积核生成空间权重热力图,并与多尺度特征进行逐位置通道缩放,使轨道交通枢纽等重点区域的特征响应显著增强,卷积中填充采用保持特征图尺寸的“same”设置。时序处理阶段, BiLSTM 层设置了128个隐藏单元捕获前后向时序依赖,再采用时间注意力机制计算时间维度相似度矩阵,通过 Softmax 归一化生成时段权重分布,结合残差连接与层归一化强化关键时间片段的特征表达。最终结合门控循环单元 GRU 的门控机制,利用更新门与重置门自适应融合历史状态与当前输入,并与 BiLSTM 输出的双向时序特征拼接,形成兼具局部细节感知与长程时空关联的融合表征,由全连接层输出预测结果,其中设置了64个隐藏单元。该结构通过多尺度空间特征提取、注意力引导的时空特征选择,以及门控循环单元的动态时序建模三重机制,系统性地增强了模型对地铁客流时空异质性的建模能力。改进 CNN-BiLSTM 模型结构如图7所示,整体结构通过各模块的协同,实现对多维特征数据的高效编码,有效应对地铁场景中大客流、线路换乘等复杂时空变化的预测需求。

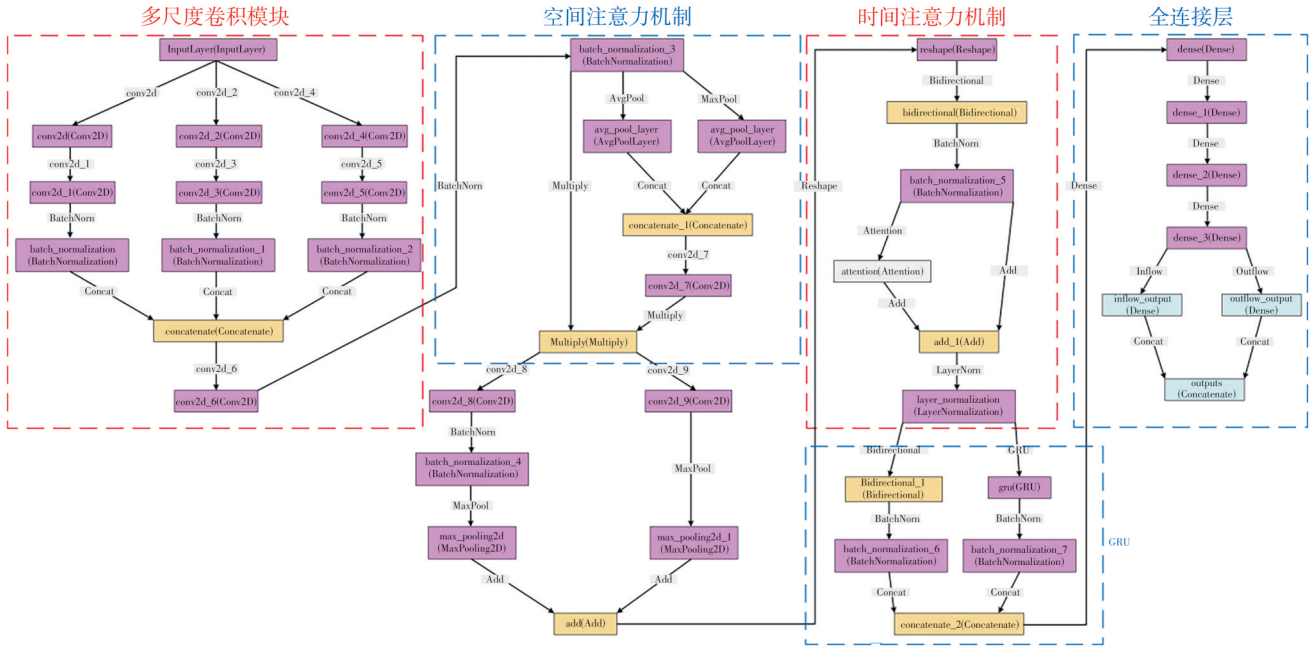


图7 改进CNN-BiLSTM模型结构
Fig.7 Structure of improved CNN-BiLSTM model

2.4 集成学习策略

通过集成学习整合多个改进CNN-BiLSTM训练的回归模型,可构建兼具高精度与强稳定性的强预测系统。集成学习的核心思想在于,即使部分个体模型存在预测偏差,通过群体决策机制仍能有效降低整体预测误差。集成学习的构建需解决基模型的生成策略与模型的组合方式选择这2个关键问题。改进CNN-BiLSTM模型属于基模型,因此只需要设计模型的组合方式即可。

基于上述分析,研究提出了一种基于改进CNN-BiLSTM模型的集成学习客流预测策略,集成学习组合方式如图8所示,该集成系统由多个回归基模型构成,其最终输出通过个体模型预测值的加权组合生成。

首先分别计算第 $i \in \{1, 2, 3, 4, 5\}$ 个基模型在验证集上的预测结果与真实值的决定系数(R^2),进

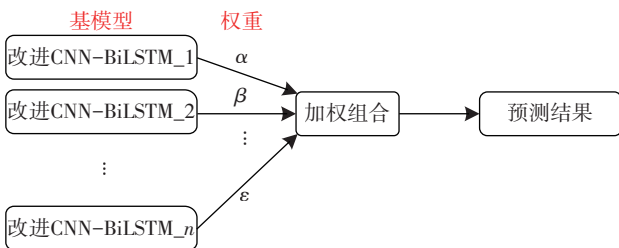


图8 集成学习组合方式
Fig.8 Combination method of ensemble learning

出站客流的 R^2 分别记为 $inflow_r2_i$ 和 $outflow_r2_i$,使用平均 R^2 值作为当前第 i 个基模型的权重。具体加权计算如公式(4)所示。

$$avg_r2_i = inflow_r2_i \times 0.5 + outflow_r2_i \times 0.5 \quad (4)$$

式中: avg_r2_i 为第 i 个模型的权重; $inflow_r2_i$ 为第 i 个模型进站客流预测的决定系数; $outflow_r2_i$ 为第 i 个模型出站客流预测的决定系数。

基于综合均值 avg_r2_i ,通过指数函数放大优质基模型的权重并归一化,计算方法如公式(5)、公式(6)所示。

$$weigh_i = e^{avg_r2_i \times 5} \quad (5)$$

$$norm_weight_i = \frac{weight_i}{\sum_{j=1}^5 weight_j} \quad (6)$$

式中: $weigh_i$ 为第 i 个模型放大后的权重; $norm_weight_i$ 为第 i 个模型归一化后的权重。

最终,预测结果通过公式(7)、公式(8)计算得到。

$$\hat{y}_{final, in} = \sum_{i=1}^5 norm_weight_i \times \hat{y}_{i, in} \quad (7)$$

$$\hat{y}_{final, out} = \sum_{i=1}^5 norm_weight_i \times \hat{y}_{i, out} \quad (8)$$

式中: $\hat{y}_{i, in}$ 和 $\hat{y}_{i, out}$ 分别为第 i 个基模型对进站和出站客流的预测值,人次; $\hat{y}_{final, in}$ 和 $\hat{y}_{final, out}$ 分别为最终进站和出站客流预测值,人次。

2.5 模型结构整体框架

模型结构整体框架如图9所示, 研究中针对地铁客流时间序列预测, 提出一种基础深度学习模型改进、结合集成学习的时间序列预测方法。基础深度学习模型改进是指改进CNN-BiLSTM模型在预测阶段进行时间序列预测并获取预测结果, 而结合集成学习是指将5个基模型的预测结果加权计算后作为最终预测结果。总的来说, 利用预测模型的监督学习满足了地铁客流不平衡、周期变化、潮汐性等因素扰动下的有效预测, 具有很好的预测正确率, 而集成学习进一步增强了预测结果的准确性并提升预测结果的稳定性。

3 结果分析

3.1 评价指标及损失函数

采用均方根误差(RMSE)、平均绝对误差(MAE)、决定系数(R^2)和平均绝对百分比误差(MAPE)等4个指标来评价预测结果的优劣, 分别表示为 r_{RMSE} , r_{MAE} ,

r_{R^2} , r_{MAPE}

$$r_{RMSE} = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (9)$$

$$r_{MAE} = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (10)$$

$$r_{R^2} = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{Y})^2} \quad (11)$$

$$r_{MAPE} = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \times 100\% \quad (12)$$

式中: y_i 为客流实际值, 人次; $\hat{y}_i$ 为客流预测值, 人次; $\bar{Y}$ 为客流实际值平均数, 人次。

将 r_{R^2} 融合进损失函数进行分析, 研究通过设计多目标损失函数将两者结合, 构建 r_{MSE} 与 r_{R^2} 的加权组合损失函数作为损失函数(loss)。

$$r_{loss} = \alpha \cdot r_{MSE} + \beta \cdot (1 - r_{R^2}) \quad (13)$$

$$r_{MSE} = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (14)$$

式中: α 和 β 为各自的权重值, 且 $\alpha + \beta = 1$; y_i 为进出站客流实际值; $\hat{y}_i$ 为进出站客流预测值; n 为样本数。为了一定程度抑制数据噪声的影响, 在均分的基础上增大 r_{R^2} 相关项权重, 直接采用了 $\alpha = 0.4$, $\beta = 0.6$ 的权重分布, 以优先保证全局趋势的稳定性。

3.2 对比实验和改进CNN-BiLSTM模型验证

以相同的训练集数据, 构建CNN模型、LSTM模型、BiLSTM模型、CNN-LSTM模型、CNN-BiLSTM和改进CNN-BiLSTM模型, 进行对比实验分析。为了对比分析特征的数量对预测模型的拟合能力, 以CNN-BiLSTM模型建立仅进出站客流为输入的2维特征模型, 和以公式(1)为输入的8维特征模型进行验证实验。三岔站CNN-BiLSTM模型训练过程如图10所示, 三岔站CNN-BiLSTM模型预测结果如图11所示, 火车南站CNN-BiLSTM模型训练过程如图12所示, 火车南站CNN-BiLSTM模型预测结果如图13所示, 三岔站模型预测评价指标如表3所示, 火车南站模型预测评价指标如表4所示。

通过进站量和出站量在2维与8维特征下的模

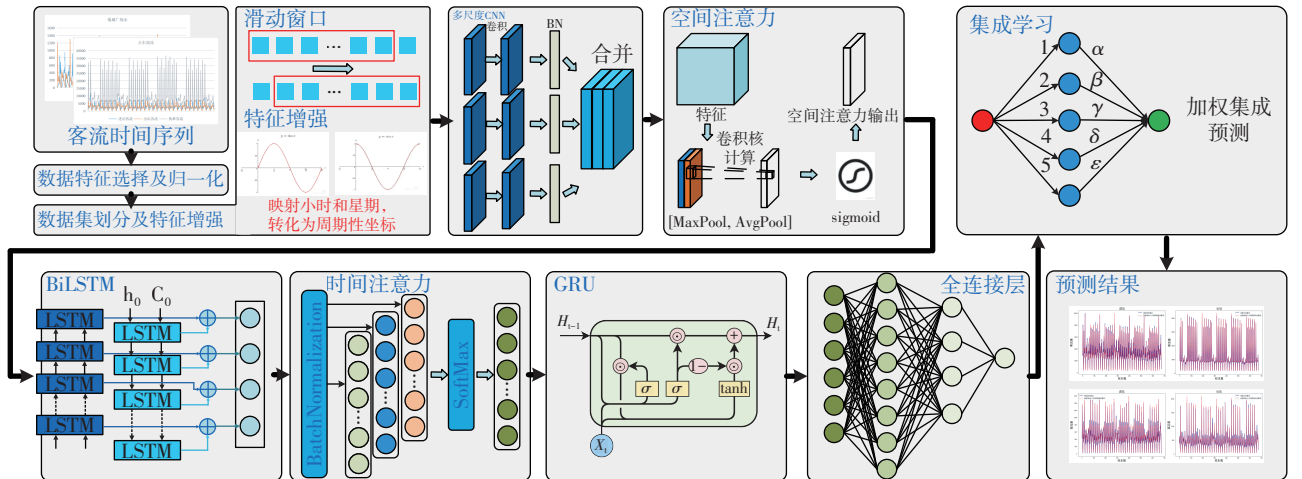
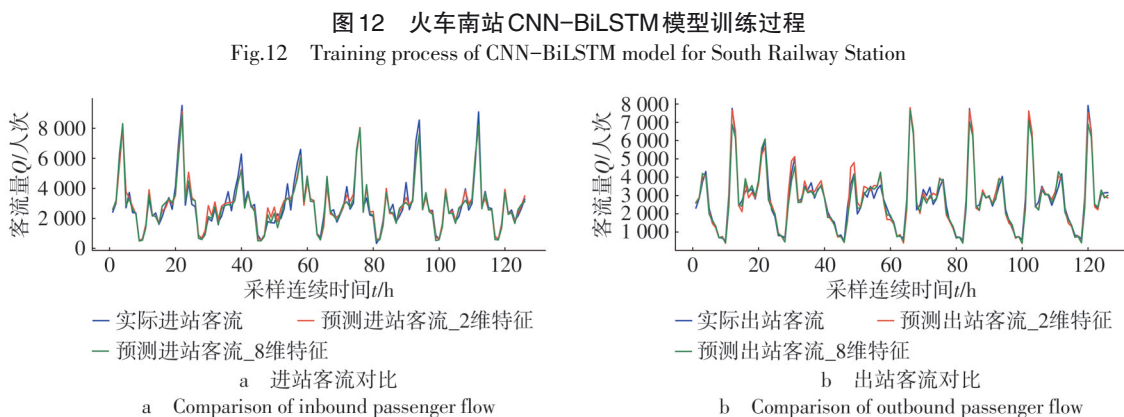
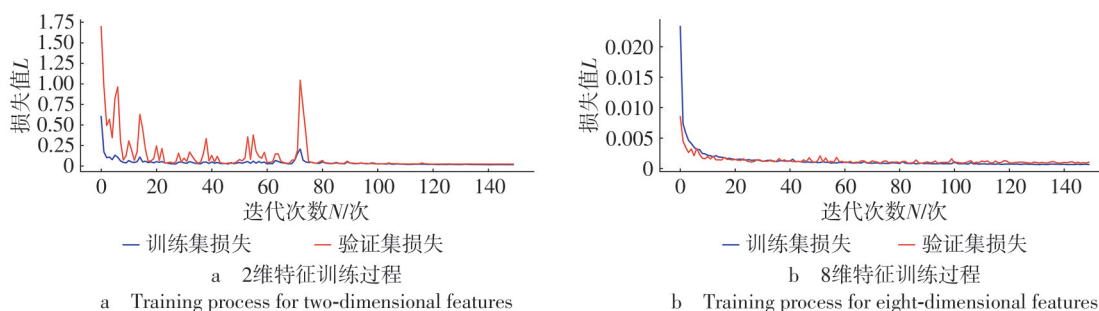
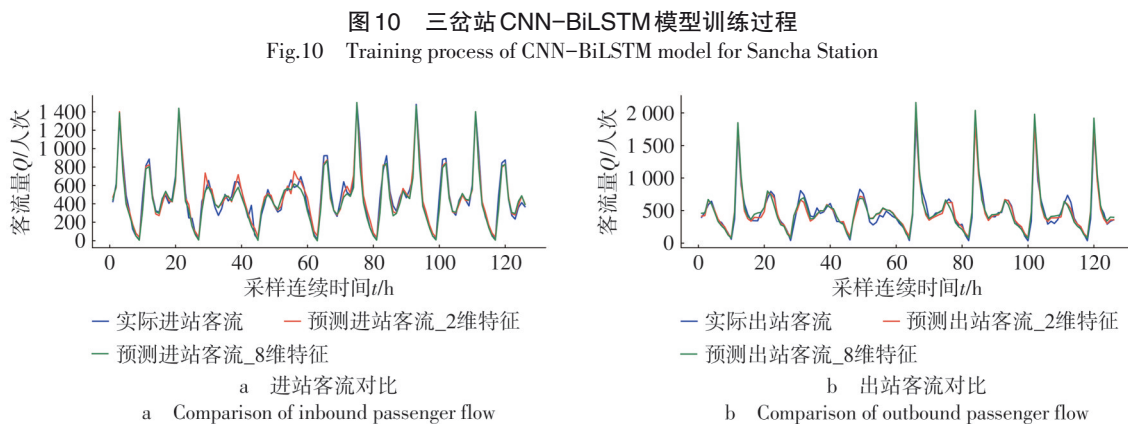
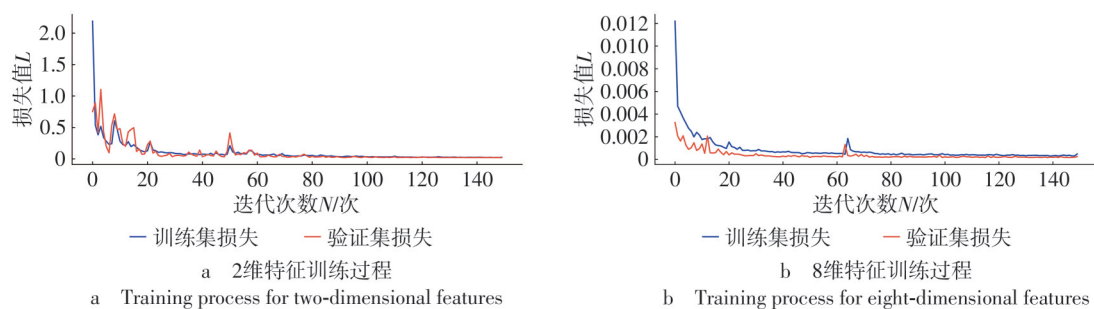


图9 模型结构整体框架

Fig.9 Overall framework of model structure



型性能指标 RMSE, MAE, R^2 和 MAPE 的对比分析发现, 8 维特征在整体上显著优于 2 维特征, 但在不同指标和场景中表现存在细微差异。在普通站进站

量预测中, 8 维特征的 RMSE 从 0.019 2 降至 0.017 6, 降幅达 8.3%, MAE 从 0.014 3 降至 0.013 6, 降幅 4.9%, R^2 从 0.934 4 提升至 0.945 0, 同时 MAPE 从

表3 三岔站模型预测评价指标

Tab.3 Prediction evaluation metrics of model for Sancha Station

模型名称	特征维度	进出站	RMSE	MAE	R^2	MAPE/%
CNN-BiLSTM	2维	进站量	0.019 2	0.014 3	0.934 4	27.32
		出站量	0.015 6	0.010 8	0.949 9	21.39
	8维	进站量	0.017 6	0.013 6	0.945 0	16.15
		出站量	0.015 5	0.011 7	0.950 7	19.14

表4 火车南站模型预测评价指标

Tab.4 Prediction evaluation metrics of model for South Railway Station

模型名称	特征维度	进出站	RMSE	MAE	R^2	MAPE/%
CNN-BiLSTM	2维	进站量	0.035 5	0.025 6	0.940 1	29.44
		出站量	0.032 2	0.024 4	0.970 0	11.76
	8维	进站量	0.029 5	0.021 4	0.963 7	26.47
		出站量	0.019 0	0.014 7	0.991 2	10.40

27.32% 提升到 16.15%，表明模型在进站量预测中通过增加特征维度有效降低了预测误差并增强了对数据变异的解释能力。而出站量预测中，8 维特征的 RMSE 仅小幅下降 0.6%，MAE 反而略微上升 0.000 9，但 R^2 实现了 0.000 8 的提升，MAPE 减小 2.25 个百分点，说明高维特征对出站量的误差控制效果有限，但模型解释力仍有所增强。值得注意的是，尽管 8 维特征在出站量的 MAE 指标上表现欠佳，但 R^2 的微升可能反映了其捕捉复杂模式的能力，而 RMSE 的微小改善则表明极端误差的控制相对稳定，且

MAPE 表现出模型较高的预测准确度。综合来看，8 维特征在进出站量客流预测中展现出显著优势。

在换乘站客流预测中，8 维特征相较于 2 维特征显著提升了模型性能，尤其在出站量预测中表现更为突出。对于进站量，8 维特征的 RMSE 从 0.035 5 降至 0.029 5，降幅达 16.9%，MAE 从 0.025 6 降至 0.021 4，降幅达 16.4%， R^2 从 0.940 1 提升至 0.963 7，提升了 2.5%，MAPE 提升约 3 个百分点，表明多维特征增强了模型对进站模式的捕捉能力。而出站量的优化更为显著，RMSE 从 0.032 2 锐减至 0.019 0，降幅达 41.0%，MAE 从 0.024 4 降至 0.014 7，降幅达 39.8%， R^2 从 0.970 0 跃升至 0.991 2，提升了 2.2%，并且 MAPE 还有 1.36 个百分点的模型精确度提高，说明 8 维特征能更精准地解析出站行为的复杂特征。整体而言，特征维度的扩展使模型误差显著降低，拟合度接近完美，且出站量预测的提升幅度远超进站量，可能源于出站客流受更多外部因素影响，而多维特征能更全面地建模这些动态关系。

改进 CNN-BiLSTM 模型训练过程如图 14 所示，三岔站改进 CNN-BiLSTM 模型预测结果如图 15 所示，火车南站改进 CNN-BiLSTM 模型预测结果如图 16 所示，三岔站模型预测评价指标如表 5 所示，火车南站模型预测评价指标如表 6 所示。

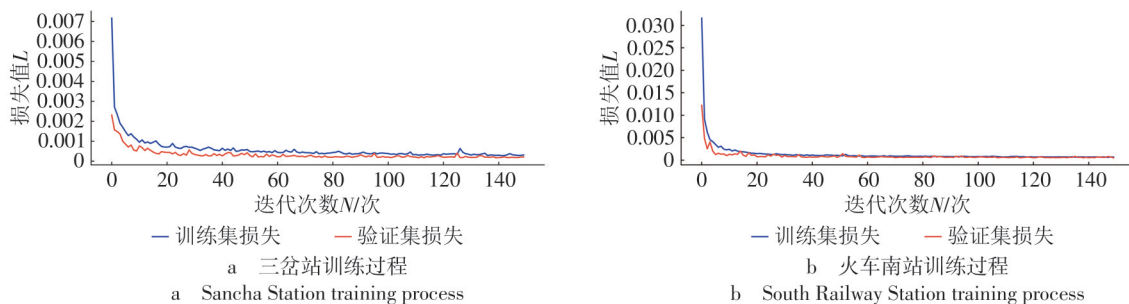


图14 改进 CNN-BiLSTM 模型训练过程
Fig.14 Training process of improved CNN-BiLSTM model

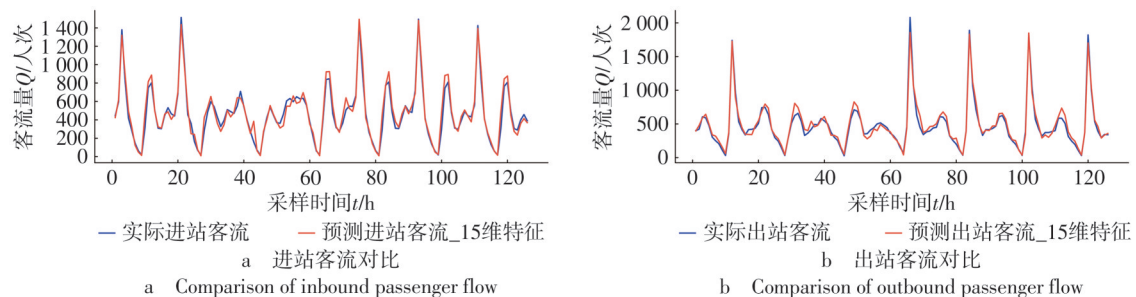


图15 三岔站改进 CNN-BiLSTM 模型预测结果
Fig.15 Prediction results of improved CNN-BiLSTM model for Sancha Station

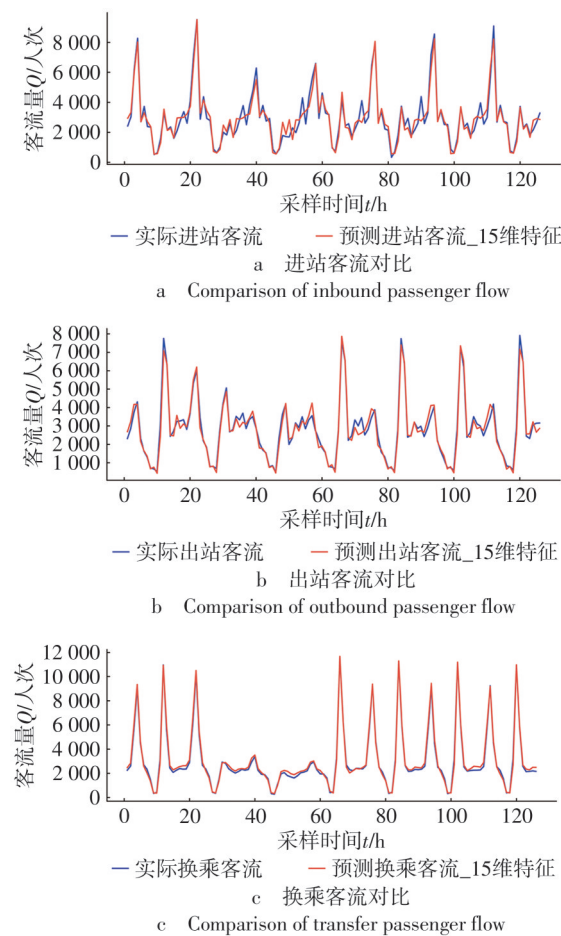


图 16 火车南站改进 CNN-BiLSTM 模型预测结果

Fig.16 Prediction results of improved CNN-BiLSTM model for South Railway Station

表 5 三岔站模型预测评价指标

Tab.5 Prediction evaluation metrics of model for Sancha Station						
模型名称	特征维度	进出站	RMSE	MAE	R^2	MAPE/%
CNN	8 维	进站量	0.021 7	0.016 3	0.916 6	26.21
		出站量	0.015 9	0.011 9	0.948 4	17.81
LSTM	8 维	进站量	0.023 1	0.016 9	0.905 3	24.57
		出站量	0.020 5	0.014 4	0.913 7	23.94
BiLSTM	8 维	进站量	0.021 1	0.015 7	0.920 6	36.42
		出站量	0.016 9	0.011 7	0.941 6	18.19
CNN-LSTM	8 维	进站量	0.017 9	0.013 5	0.943 1	21.54
		出站量	0.017 6	0.011 9	0.936 2	15.75
CNN-BiLSTM	8 维	进站量	0.017 6	0.013 6	0.945 0	16.15
		出站量	0.015 5	0.011 7	0.950 7	19.14
改进 CNN-BiLSTM	15 维	进站量	0.016 4	0.012 2	0.952 2	14.05
		出站量	0.014 9	0.010 3	0.954 7	12.18

由图 14、图 15 可以看出，改进 CNN-BiLSTM 在 10 步左右收敛，收敛速度较快，与 CNN-BiLSTM 模型基本一样。根据表 5、表 6 中评价指标的对比，改进 CNN-BiLSTM 模型在预测精度上均有显著提升，

表 6 火车南站模型预测评价指标

Tab.6 Prediction evaluation metrics of model for South Railway Station						
模型名称	特征维度	进出站	RMSE	MAE	R^2	MAPE/%
CNN	8 维	进站量	0.038 2	0.028 2	0.939 0	27.31
		出站量	0.019 6	0.014 3	0.990 7	12.36
		换乘量	0.017 4	0.013 6	0.992 6	12.73
LSTM	8 维	进站量	0.036 6	0.026 6	0.944 0	38.78
		出站量	0.019 4	0.015 3	0.990 9	12.56
		换乘量	0.023 1	0.017 2	0.987 0	13.23
BiLSTM	8 维	进站量	0.044 7	0.032 3	0.916 5	36.33
		出站量	0.025 2	0.019 7	0.984 6	13.66
		换乘量	0.023 6	0.018 7	0.986 5	17.69
CNN-LSTM	8 维	进站量	0.030 7	0.022 7	0.960 6	31.58
		出站量	0.020 2	0.015 2	0.990 1	10.56
		换乘量	0.017 2	0.012 9	0.992 8	8.00
CNN-BiLSTM	8 维	进站量	0.029 5	0.021 4	0.963 7	26.47
		出站量	0.019 0	0.014 7	0.991 2	10.40
		换乘量	0.016 4	0.011 8	0.993 5	8.69
改进 CNN-BiLSTM	15 维	进站量	0.021 2	0.022 3	0.969 4	19.36
		出站量	0.014 0	0.010 5	0.995 2	13.02
		换乘量	0.015 8	0.012 7	0.993 9	8.32

在普通站点上，与其他 5 种模型相比，进站客流预测中的 RMSE，MAE， R^2 和 MAPE 分别提升 29.0%，7.8%，5.2% 和 61.4 个百分点，出站客流同样提升了 27.3%，28.5%，4.5% 和 49.1 个百分点；在换乘站点上，改进 CNN-BiLSTM 模型以其对数据的深度挖掘能力，展现出更明显的优势，进站客流预测中的 RMSE，MAE 和 R^2 分别提升了 52.6%，32.0% 和 5.8%，出站客流提升 44.4%，46.7% 和 1.1%，观察 MAPE 指标可得，改进模型在进站客流预测上有大幅提升，相比仅次于该模型的 CNN-BiLSTM 模型也有 7.11 个百分点的提升，使其低于 20%，则进站、出站和换乘量的预测都具有较高的准确度，可为决策者提供决策依据。在换乘站中，同时对换乘量进行了预测分析，3 种组合模型均优于单一模型，4 种评价指标综合来看，改进 CNN-BiLSTM 模型与其他 2 种组合模型相比，表现出更好的均衡性，展现了模型的稳定性和普适性。

4 结束语

以成都地铁三岔站(普通站)和火车南站(换乘站)实际客流数据为基础，进行多维度特征输入模型在 2 类车站的分类讨论，以探究模型的适用性。

首先,根据地铁车站客流周期性和不均衡性,在 CNN-BiLSTM 的基础模型结构上,设计多尺度卷积,学习客流短期(小时)、中长期(天)和长期(周)的时间和空间特征;其次,引入空间注意力机制和时间注意力机制,增强模型对关键时空信息的捕获能力,着重捕捉路网中关键节点的客流变化特征;同时,保留 BiLSTM 层处理客流数据的时序性,结合门控循环单元 GRU,降低模型过拟合风险,形成兼具局部细节感知与长程时空关联的融合模型,系统性地增强了模型对地铁客流时空异质性的建模能力;最后以集成学习思想增强模型的稳定性,排除个体模型偏差的影响。通过与其他模型的对比验证,所提出的改进 CNN-BiLSTM 模型不论在普通站还是换乘站都有较高的预测精度,具有良好的适用性和泛化性。

参考文献:

- [1] 曹敬涵, 许 琰, 孙立山, 等. 基于轨道交通数据的客流特征与城市功能结构分析[J]. 都市快轨交通, 2021, 34(2): 71-78, 85.
CAO Jinghan, XU Yan, SUN Lishan, et al. Passenger Flow Characteristics and Analysis of Urban Functional Structure Based on Rail Transit Data[J]. Urban Rapid Rail Transit, 2021, 34(2): 71-78, 85.
- [2] 蔡昌俊, 姚恩建, 王梅英, 等. 基于乘积 ARIMA 模型的城市轨道交通进出站客流量预测[J]. 北京交通大学学报, 2014, 38(2): 135-140.
CAI Changjun, YAO Enjian, WANG Meiyong, et al. Prediction of Urban Railway Station's Entrance and Exit Passenger Flow Based on Multiply ARIMA Model[J]. Journal of Beijing Jiaotong University, 2014, 38(2): 135-140.
- [3] YAO E J, HONG J Y, PAN L, et al. Forecasting Passenger Flow Distribution on Holidays for Urban Rail Transit Based on Destination Choice Behavior Analysis[J]. Journal of Advanced Transportation, 2021, 2021: 9922660.
- [4] ZHANG S X, ZHANG J L, YANG L X, et al. Spatiotemporal Attention Fusion Network for Short-Term Passenger Flow Prediction on New Year's Day Holiday in Urban Rail Transit System[J]. IEEE Intelligent Transportation Systems Magazine, 2023, 15(5): 59-77.
- [5] SMITH B L, DEMETSKY M J. Traffic Flow Forecasting: Comparison of Modeling Approaches[J]. Journal of Transportation Engineering, 1997, 123(4): 261-266.
- [6] 张丽莉, 宁亚琴. 基于改进随机森林模型的地铁客流量预测[J]. 公路工程, 2025, 50(2): 226-232.
ZHANG Lili, NING Yaqin. Subway Passenger Flow Prediction Based on Improved Random Forest Model[J]. Highway Engineering, 2025, 50(2): 226-232.
- [7] YAO K, GAO G, LIU Y, et al. A Stable Passenger Flow Forecast Approach for Newly Opened Metro Stations Based on Multi-Source Data and Random Forest Regression Model[C]// 2022 3rd International Conference on Intelligent Design (ICID). Xi'an: IEEE, 2022: 249-254.
- [8] MA D L, GUO Y T, MA S Z. Short-Term Subway Passenger Flow Prediction Based on GCN-BiLSTM[J]. IOP Conference Series: Earth and Environmental Science, 2021, 693(1): 012005.
- [9] ZHANG Z H, HAN Y, PENG T X, et al. A Comprehensive Spatio-Temporal Model for Subway Passenger Flow Prediction[J]. ISPRS International Journal of Geo-Information, 2022, 11(6): 341.
- [10] ZENG L, LI Z N, YANG J, et al. CEEMDAN-IPSO-LSTM: A Novel Model for Short-Term Passenger Flow Prediction in Urban Rail Transit Systems[J]. International Journal of Environmental Research and Public Health, 2022, 19(24): 16433.
- [11] 黄昕宇. 基于注意力机制的 LSTM 区域人流量预测研究[D]. 南昌: 南昌大学, 2023.
- [12] LIN S F, TIAN H Y. Short-Term Metro Passenger Flow Prediction Based on Random Forest and LSTM[C]//2020 IEEE 4th Information Technology, Networking, Electronic and Automation Control Conference (ITNEC). Chongqing: IEEE, 2020: 2520-2526.
- [13] GENG X R, MA Y, CAI W N, et al. Evaluation of Models for Multi-Step Forecasting of Hand, Foot and Mouth Disease Using Multi-Input Multi-Output: A Case Study of Chengdu, China[J]. PLoS Neglected Tropical Diseases, 2023, 17(9): e0011587.
- [14] 李 洁, 彭其渊, 文 超. 基于 LSTM 深度神经网络的高速铁路短期客流预测研究[J]. 系统工程理论与实践, 2021, 41(10): 2669-2682.
LI Jie, PENG Qiyuan, WEN Chao. Short Term Passenger Flow Prediction of High Speed Railway Based on LSTM Deep

(下转第 171 页)