

基于混合 Transformer 的对外客运枢纽 抵站客流预测模型

于江波¹, 翁剑成¹, 林鹏飞², 孙宇星³, 柴娇龙³

(1. 北京工业大学 交通工程北京市重点实验室, 北京 100124; 2. 北京工业大学 信息学部, 北京 100124; 3. 北京市交通委员会 北京市运输事业发展中心, 北京 100088)

摘 要: 本文综合考虑节假日、恶劣天气、疫情防控政策等因素影响, 以 2021 年 3 月~2023 年 5 月北京首都国际机场及 2021 年 12 月~2023 年 7 月的 7 个高铁站点的抵站客流数据为基础, 提出了一种融合卷积神经网络、长短期记忆网络的混合 Transformer 枢纽抵站客流预测模型, 分析多因素影响下抵站客流规模的时变规律。结果表明, 混合 Transformer 模型相比于传统机器学习和深度学习模型, 具有更小的误差, 对所有场站的客流预测准确率超过了 80%, 其中对 6 个场站客流的预测准确率在 90% 左右, 证明该模型具有良好的适用性。

关键词: 交通运输系统工程; 对外客运枢纽; 城际出行; 客流预测; 混合 Transformer 模型

中图分类号: U125 **文献标志码:** A **文章编号:** 1671-5497(2025)07-2251-09

DOI: 10.13229/j.cnki.jdxbgxb.20231091

Passenger flow prediction model of external transportation hub based on hybrid Transformer

YU Jiang-bo¹, WENG Jian-cheng¹, LIN Peng-fei², SUN Yu-xing³, CHAI Jiao-long³

(1. Beijing Key Laboratory of Traffic Engineering, Beijing University of Technology, Beijing 100124, China; 2. Faculty of Information Technology, Beijing University of Technology, Beijing 100124, China; 3. Beijing Development Center of Transport, Beijing Municipal Commission of Transport, Beijing 100088, China)

Abstract: The impact of various external factors such as holidays, weather conditions, and epidemic prevention policies was considered in this paper, based on the passenger arrivals date at Beijing Capital International Airport between March 2021 and May 2023, and seven major high-speed railway stations from December 2021 to July 2023, a hybrid Transformer model integrating convolutional neural networks, long short-term memory networks is proposed for hub-to-station passenger flow prediction, analyzing the time-varying patterns of passenger flow volume at arrival stations under the influence of multiple factors. The empirical results reveal that the hybrid Transformer model outperforms conventional machine learning and deep learning models in prediction accuracy, achieving over 80% across all stations, and nearing 90% for six of them, this evidences the model's robust applicability.

收稿日期: 2023-10-11.

基金项目: 国家自然科学基金项目 (52072011).

作者简介: 于江波 (1996-), 男, 博士研究生. 研究方向: 智能交通. E-mail: y1728467259@gmail.com

通信作者: 翁剑成 (1981-), 男, 教授, 博士. 研究方向: 智能交通与交通大数据. E-mail: youthweng@bjut.edu.cn

Key words: engineering of communications and transportation system; external transportation hubs; intercity mobility; passenger flow prediction; hybrid Transformer model

0 引言

对外客运枢纽是包含民航、铁路或省际客运等至少一种城市对外客运交通方式和若干市内客运交通方式,通过完善的换乘服务体系所连接的固定设备和活动设备所形成的从事旅客运输的一类交通枢纽^[1]。考虑特殊条件对客运枢纽的抵站客流规模进行精准预测,将有效帮助管理者及时掌握抵站客流变化趋势,对提高对外交通枢纽的交通运输接续保障能力具有重要意义。

对外客运枢纽的客流规模受到多种时空因素的影响,有学者重点考虑天气^[2]、星期^[3]或者假期^[4]等因素开展城市内交通站点客流预测研究。Zhang 等^[5]考虑天气和空气质量因素构建深度学习集成模型进行短时客流预测,提升了模型的预测精度;Sha 等^[6]分析每周不同天的客流特征,发现工作日和非工作日的客流具有一定差异。疫情条件对客流预测^[7]和乘坐飞机的旅客的出行行为^[8]均产生了重要影响,而这些因素对枢纽客流规模的影响程度仍有待探究。

在充分考虑多因素影响的情况下,如何准确提取影响因素并进行量化,建立能够快速计算、准确率高的预测模型仍值得探究。梯度提升算法^[9]、长短期记忆(Long short-term memory, LSTM)^[10,11]模型及其改进形式^[12-14]在客流预测相关研究中得到广泛应用。基于 Transformer 的新型网络结构也被证明能够提升时间序列预测的准确性^[15-17]。在恶劣天气、节假日以及前期疫情防控政策变化背景下,对外客运枢纽抵站客流规模存在较大的波动与不确定性,鲜有研究将 Transformer 模型应用于城市的对外客运枢纽抵站客流预测。综上所述,本文拟通过引入 Transformer 模型结构处理数据长期依赖关系,融合卷积神经网络(Convolutional neural networks, CNN)和 LSTM 神经网络构建 CNN-LSTM-Transformer 混合模型(CL-Transformer),以提高预测模型的客流趋势表征能力,并以北京首都国际机场及北京市 7 个高铁站点的抵站客流为例,检验本文构建的模型对于特殊事件影响下的客流预测精度以及模型结构对于不

同场站的适用性,以期为城市对外客运枢纽客流规模预判、接续运输保障和调度组织提供精准的客流数据支撑。

1 研究数据基础及影响因素

研究所用数据主要分为客流数据、天气数据、疫情相关数据、时间数据等 4 类。客流数据来源于北京市接续运输与春运节假日保障系统,包含 2021 年 3 月 2 日至 2023 年 5 月 10 日北京首都国际机场实际抵站客流数据统计和预售票量数据,以及 2021 年 12 月 1 日至 2023 年 7 月 8 日北京各铁路枢纽站点的实际抵站客流数据统计和预售票量数据,数据统计粒度为小时,预售票数据为到站前一日 15:00 前获得小时粒度统计数据。天气数据统计粒度为小时,包含最高与最低气温、风速。北京首都国际机场抵站客流数据集因素样例如表 1 所示。高铁站点抵站客流数据集中不考虑准点率及班次信息,其他影响因素相同。在各类影响因素中,星期、天气、节假日特征、进返京限制政策等为类别变量,其余为连续变量。

从国家卫健委网站及本地宝获取每日新增本土确诊病例数与北京疫情防控措施情况,并将疫情防控分为 3 级:第一级管控最为严格,在新冠肺炎病例数增长较快的情况下,对中高风险地区人员严格限制进返京;第二级管控较为宽松,在本土新冠肺炎病例增长较少的情况下,按照人员旅居史情况限制进返京;第三级管控最为宽松,在国内无新冠肺炎病例增长的情况下,低风险地区仅需健康宝绿码即可进返京,具体分级标准见表 2。2022 年 12 月 24 日后,疫情形势逐渐好转,不再进行病例统计与管控分级。

2 客流特征及影响因素分析

城市对外客运枢纽抵站客流在不同日期、不同时间的分布不同。首先对客流数据的整体特征进行分析,图 1 展示了 2021 年 3 月~2023 年 3 月北京首都国际机场客流状况,该时间段包含了疫情期间及政策放开初期的客流状况。从日期维度上看,2022 年 8 月份与 11 月份客流明显减少;从小时维度上看,凌晨 0:00~3:00 时段抵站客流急

表 1 初始数据集因素样例				
Table 1 Factor samples of original dataset				
因素	2021/3/2	2021/3/2	2021/3/2	...
时间	8:00~9:00	9:00~10:00	10:00~11:00	...
实际抵站客流/人次	584	1 032	2 239	...
预售票/张	1 775	1 980	2 759	...
星期	2	2	2	...
节假日	4	4	4	...
月份	3	3	3	...
季节	1	1	1	...
风速/(m·s ⁻¹)	1	2	2	...
天气	雾霾	雾霾	雾霾	...
最低温度/℃	-1.2	-1.2	-1.2	...
最高温度/℃	5	5	5	...
航班准点率/%	100	100	100	...
航班数量/架次	6	13	19	...
新增病例数	0	0	0	...
进返京限制政策	1	1	1	...

表 2 北京市城际出行(进返京)的疫情防控政策示例	
Table 2 Exaple of epidemic prevention and control measures for itercity centering and returning to Beijing in Beijing	
疫情防控等级	政策
第一级管控	中高风险地区所在市全域人员限制进京,中高风险地区人员不得进京。全面升级北京首都国际机场、大兴国际机场防控措施。所有人员进(返)京须持 48 h 内核酸检测阴性证明和“北京健康宝”绿码
第二级管控	有 1 例及以上本土新冠病毒感染者所在地级市(直辖市、副省级城市的县、市、区)其他县人员来京的须持“北京健康宝”绿码和登机登车前 48 h 内核酸检测阴性证明;高风险地区人员暂不许返京
第三级管控	国内低风险地区进返京人员仅需持健康通行码绿码即可进京;国内中高风险地区及全域实行封闭管理的人员,原则上不能进京

剧减少,6:00~12:00 客流呈逐步增长的趋势,12:00 至第二天 0:00 抵站客流均保持当日较高的水平。

图 2 表明在 2022 年 7 月底和 10 月底每日新增本土病例数量开始增长,受进返京管控措施的影响,抵站客流显著减少。

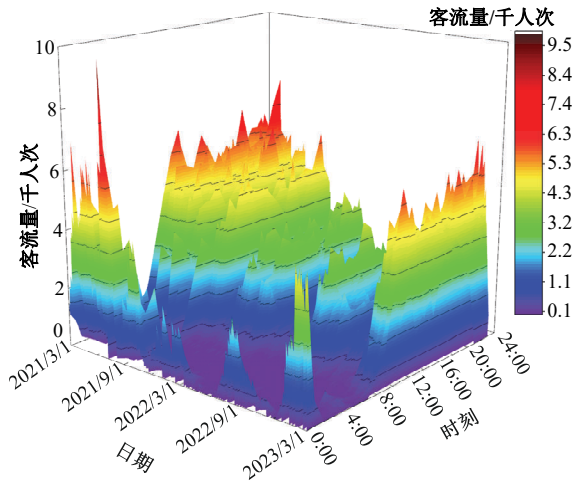


图 1 北京首都国际机场抵站客流整体特征
Fig. 1 Overall tendency of passenger flow arriving at Beijing Capital International Airport

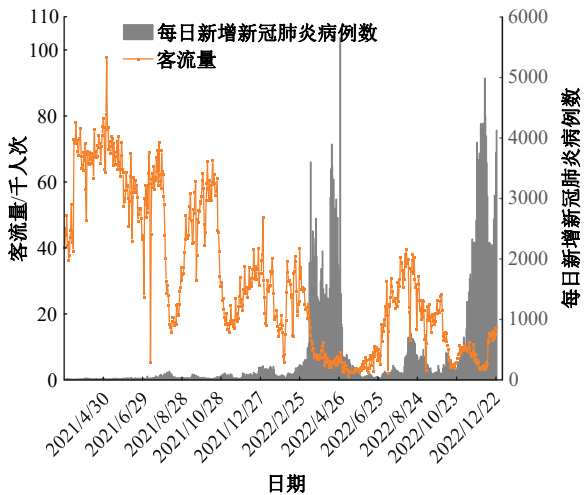


图 2 客流与每日新增病例日变特征对比
Fig. 2 Comparison of daily variation characteristics between passenger flow and number of daily new COVID-19 cases

3 CL-Transformer 模型构建

3.1 模型构建

在恶劣天气、节假日叠加疫情防控策略变化的条件下,对外客运枢纽旅客抵站特征存在较大的不确定性。CNN 具有较强的特征提取能力,有利于从复杂的影响因素中提取关键特征,但并不擅长处理时间序列数据中的长期依赖关系;LSTM 是一种特殊的循环神经网络,能够弥补 CNN 的缺点,但 LSTM 在实际应用中可能遇到梯度消失或者爆炸的问题,影响预测的准确性。Transformer 模型的注意力机制能够实现并行处理数据,相比于 LSTM,能够更有效地处理数据

的长期依赖关系。通过融合 3 种类型的神经网络结构建立混合模型,能够实现数据的重要特征提取,并有效降低局部冗余,捕获数据中存在的长期依赖关系,并且混合模型在处理不同类型的时间序列数据时,具有更好的灵活性和鲁棒性,有利于实现多个场站的精准预测。

本文构建基于混合 CL-Transformer 模型的枢纽抵站客流预测模型结构如图 3 所示。首先利用 Savitzky-Golay(SG)滤波器对抵站客流数据进行降噪,并与其他因素关联匹配,构造训练与验证数据集,之后对模型进行训练和参数调优,最终输出模型的预测结果。

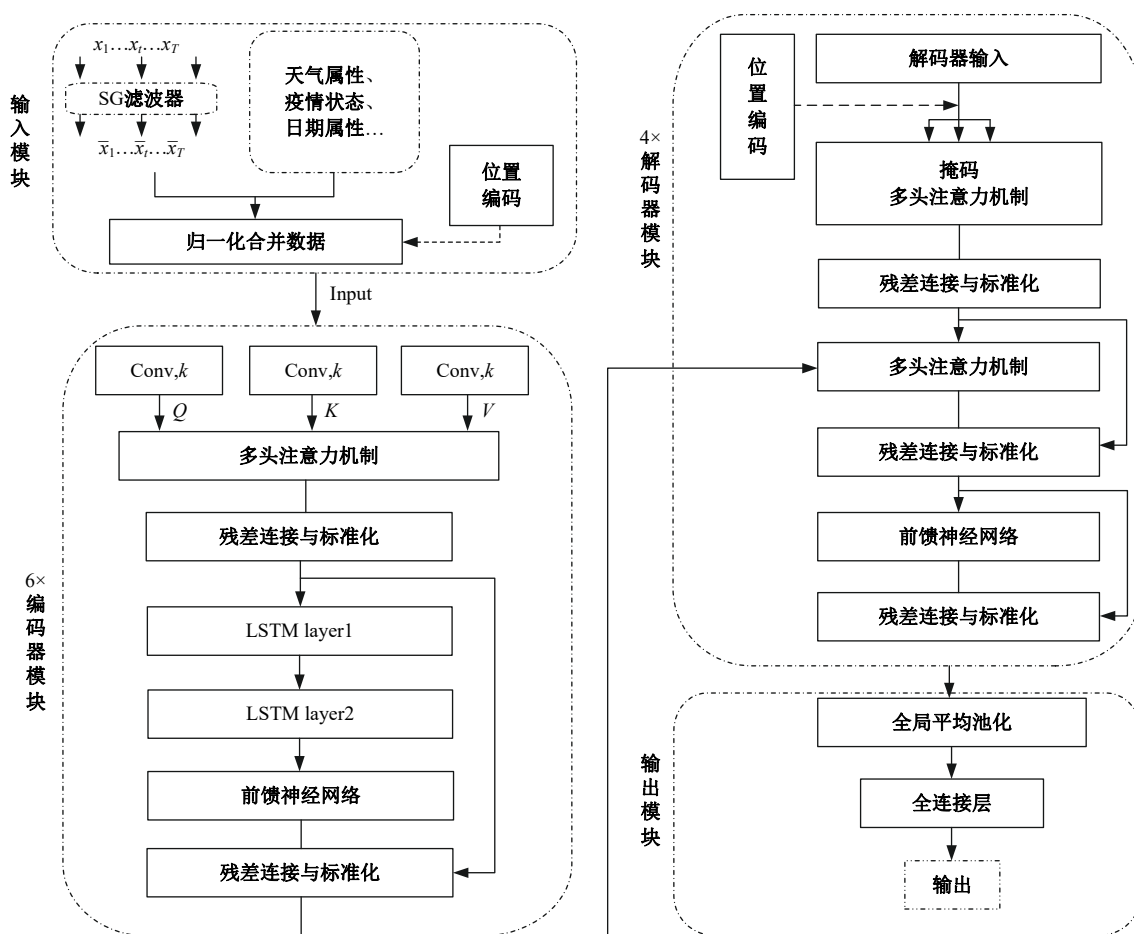


图 3 混合 CL-Transformer 模型结构

Fig.3 The structures of hybrid CL-Transformer model

为了消除客流时间序列中存在的随机波动和噪声,本文利用 SG 滤波器对数据进行预处理。处理过程如下:

$$X_j^* = \frac{\sum_{i=-m}^{i=m} C_i X_{j+1}}{N} \quad (1)$$

式中: X_j^* 为时间序列的拟合值; X_{j+1} 为窗口中的时间序列数据点; C_i 为第 i 个时间序列数据的系数; N 为卷积数目; m 为滤波窗口大小,本文中设置滤波窗口为 25; j 为原始时间序列数据集的系数。

预测模型利用卷积层进行数据特征的提取和映射, k 为卷积核的数量,输入数据与不同的卷积

核进行卷积运算,可以得到包含输入数据不同特征的映射矩阵 X^{m+1} ,可表示为:

$$X^{m+1} = \text{ReLU}(\sum W^m \cdot X^m + B^m) \quad (2)$$

式中: W^m 为该层的权重; X^m 、 X^{m+1} 分别为 m 层及其下一层的特征输入; B^m 为该层的偏置。

传统 Transformer 模型中没有循环和卷积结构,本文将 LSTM 网络嵌入 Transformer 的编码器中,增强对数据长期依赖信息的学习。LSTM 的遗忘门将前一层隐藏状态的信息 h_{t-1} 和当前输入信息 x_t 传递到 sigmoid 函数中,对信息进行筛选,保留输出值 f_t 接近于 1 的值,删除接近于 0 的值,如式(3)所示,其中 W_f 为权重, b_f 为偏置。在

输入门中通过线性运算得到输入矩阵 G_t , 如式(4)所示。利用 tan 函数创造新的候选值向量 $\widetilde{C}_t$, 存储 x_t 和 h_{t-1} 中的信息, 如式(5)所示。将 sigmoid 函数输出与 tan 输出相乘, 利用 sigmoid 的输出值决定需要保留的 tan 函数输出值, 之后将隐藏状态作为当前层的输出, 将新的单元状态和隐藏单元状态输入到下一步长中, 生成单元状态 C_t , 如式(6)所示。

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (3)$$

$$G_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (4)$$

$$\widetilde{C}_t = \tan(W_c \cdot [h_{t-1}, x_t] + b_c) \quad (5)$$

$$C_t = f_t * C_{t-1} + G_t * \widetilde{C}_t \quad (6)$$

根据 Query-Key-Value 模型, Transformer 中的注意力机制可以表示为:

$$\text{Attention}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = \text{softmax}\left(\frac{\mathbf{Q}\mathbf{K}^T}{\sqrt{D_K}}\right)\mathbf{V} \quad (7)$$

式中: $\mathbf{Q}, \mathbf{K}, \mathbf{V}$ 为 CNN 输出的特征映射; $\mathbf{Q}$ 为查询矩阵; $\mathbf{K}$ 为键矩阵; $\mathbf{V}$ 为值矩阵; D_K 为矩阵 $\mathbf{K}$ 的维度。Transformer 使用注意力机制学习不同类别数据间的相互依赖关系, 通过拼接方式将单头自注意力变为多头自注意力, 可表示为:

$$\begin{aligned} \text{MultiHeadAttn}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) &= \\ \text{Concat}(\text{head}_1, \dots, \text{head}_H) \mathbf{W}^o & \\ \text{where head}_i &= \text{Attention}(\mathbf{Q}\mathbf{W}_i^Q, \mathbf{K}\mathbf{W}_i^K, \mathbf{V}\mathbf{W}_i^V) \end{aligned} \quad (8)$$

前馈神经网络为全连接神经网络的形式, 连接多头自注意力输出, 可表示为:

$$\text{FFN}(H') = \text{ReLU}(H'W^1 + b^1)W^2 + b^2 \quad (9)$$

式中: H' 为前一层的输出; W^1, W^2, b^1, b^2 为训练参数, 通过残差连接方式连接各层。

$$H' = \text{LayerNorm}(\text{SelfAttn}(X) + X) \quad (10)$$

$$H = \text{LayerNorm}(\text{FFN}(H') + H') \quad (11)$$

式中: $\text{SelfAttn}(\cdot)$ 为自注意力模块; $\text{LayerNorm}(\cdot)$ 为归一化层。

3.2 模型训练和影响因素重要度分析

为了进一步降低模型的复杂度, 筛选反映客流特征及变化规律的典型指标, 引入基于置换特征重要性算法计算特征间的相对重要性 FI。平均绝对误差 (Mean absolute error, MAE) 值越大表明该特征的重要程度越高^[18], 特征 p 的重要性分数 FI_p 可以表示为

$$\text{FI}_p = \frac{E_p - \text{Min}(E_p)}{\text{Max}(E_p) - \text{Min}(E_p)} \quad (12)$$

式中: E 为误差。

不同因素的相对重要性如表 3 所示, 结果表明, 进返京限制政策与本土新增新冠肺炎病例数对于抵站客流预测均具有相对重要的影响。

表 3 影响因素间的相对重要度

Table 3 Relative importance of influencing factors

影响因素	相对重要度	排序
时段属性	1	1
航班数	0.92	2
预售票量	0.84	3
天气状况	0.89	4
进返京限制政策	0.85	5
本土新增新冠肺炎病例数	0.73	6
准点率	0.72	7
日期属性	0.71	8
节假日	0.65	9
风速	0.59	10
月份	0.47	11
季节	0.32	12
最低气温	0.27	13
最高气温	0	14

3.3 模型精度验证及结果分析

本文基于 TensorFlow 框架, 首先选取 2021 年 3 月 2 日至 2022 年 12 月 22 日共 15 840 h 数据样本为基础, 以前 80% 的数据为训练样本, 后 20% 的数据为测试样本, 构建 CL-Transformer 模型, 并与 GBDT、LSTM、Transformer、CNN-Transformer 和 LSTM-Transformer 模型进行比较。利用 MAE、均方根误差 (Root mean square error, RMSE) 和准确率 (Accuracy) 对模型的预测结果进行评价, 各指标计算公式如下:

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i| \quad (13)$$

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (|y_i - \hat{y}_i|)^2} \quad (14)$$

$$\text{Accuracy} = 1 - \frac{\sum_{i=1}^N |y_i - \hat{y}_i|}{\sum_{i=1}^N y_i} \quad (15)$$

式中: y_i 为 i 时刻客流真实值; $\hat{y}_i$ 为模型的客流预测值; N 为总样本数。

本文利用 Grid Search 方法确定模型训练的批大小 (batch size)、学习率等最优参数, 采用 3 折交叉验证方法确定具有良好泛化能力的模型。各模型的最优参数如表 4 所示。

表 4 不同类型模型的最优参数结果

Table 4 Optimal parameter results for different models

模型类型	模型名称	最优参数
深度学习模型	LSTM、Transformer	学习率:0.001
		Epoch:150
		批大小:48
		优化器:Adam
	CNN-Transformer、LSTM-Transformer、CL-Transformer	学习率:0.001
		Epoch:150
		批大小:128
		优化器:Adam
机器学习模型	GBDT	评估器数量:1 000
		最大深度:7
		最小样本数量:50
		学习率:0.002

模型验证结果如表 5 所示。结果表明,CL-Transformer 算法具有更加优秀的预测性能,对各场站预测结果的 MAE、RMSE 结果最小,准确率最高,对北京首都国际机场、北京西站、北京南站、北京北站、清河站的预测准确率在 90% 左右,对于北京丰台站的预测准确率较为理想,为 83%,对于北京朝阳站的预测准确率较低,为 81.3%。图 4 展示了客流量较大的北京首都国际机场、北京站、北京西站、北京南站和准确率相对较低的北京朝阳站、北京丰台站连续 4 天(共 96 h)的各场站客流预测结果。可以发现,CL-Transformer 模型对于客流变化的节点更为敏感,更为接近实际客流状态,能够更好地表现随时间的变化趋势。

表 5 不同客流预测模型性能比较

Table 5 Performance comparison of different passenger flow predicting models

场站	验证指标	GBDT	LSTM	Transformer	CNN-Transformer	LSTM-Transformer	CL-Transformer
北京首都国际机场	MAE	375.98	269.8	230.6	134.6	144.8	88.6
	RMSE	432.01	359.1	462.3	196.1	191.9	134.4
	Accuracy/%	71.8	80.1	72.5	87.5	86.6	91.8
北京站	MAE	723.6	875.9	676.8	538.9	746.2	440.3
	RMSE	1 824.3	4 868.3	2 470.2	2 388.9	2 486.9	2 222.3
	Accuracy/%	43.2	66.8	74.3	79.5	71.7	83.3
北京西站	MAE	1 379.5	943.2	634.3	784.3	1 020	590
	RMSE	2 657.6	4 155.6	2 360.4	2 997.5	2 954.9	1 418.3
	Accuracy/%	51.2	82.4	88.2	85.3	81	89
北京南站	MAE	1 877.6	1 336.7	1 242.7	958.6	1 515.4	697.4
	RMSE	3 728.9	4 034.2	3 375.5	3 951.3	3 881.7	3 088.5
	Accuracy/%	39.5	79.7	81.1	85.4	77.1	89.4
北京北站	MAE	139.1	202.2	62.4	43.4	106.7	29.13
	RMSE	187.8	423.2	83.7	70.7	175.4	51.8
	Accuracy/%	54.8	64.9	89.2	92.5	81	95.9
清河站	MAE	244.6	56.4	134.3	45.1	164.3	44.7
	RMSE	420	129.8	195.7	115.2	322.5	117.6
	Accuracy/%	43.1	93.4	84.3	94.8	80.9	94.8
北京朝阳站	MAE	501.5	734.8	623.6	533.9	762	492.7
	RMSE	2 108.7	4 648.3	3 811.9	3 523.7	3 675.8	3 337.1
	Accuracy/%	37.5	61.1	65.3	68.9	55.5	81.3
北京丰台站	MAE	648.1	505.2	634.37	460.9	813.5	350.6
	RMSE	1 269.8	1 308.9	1 769.9	1 500.8	1 861.3	1 289.6
	Accuracy/%	46.6	75.7	69.6	77.9	61	83

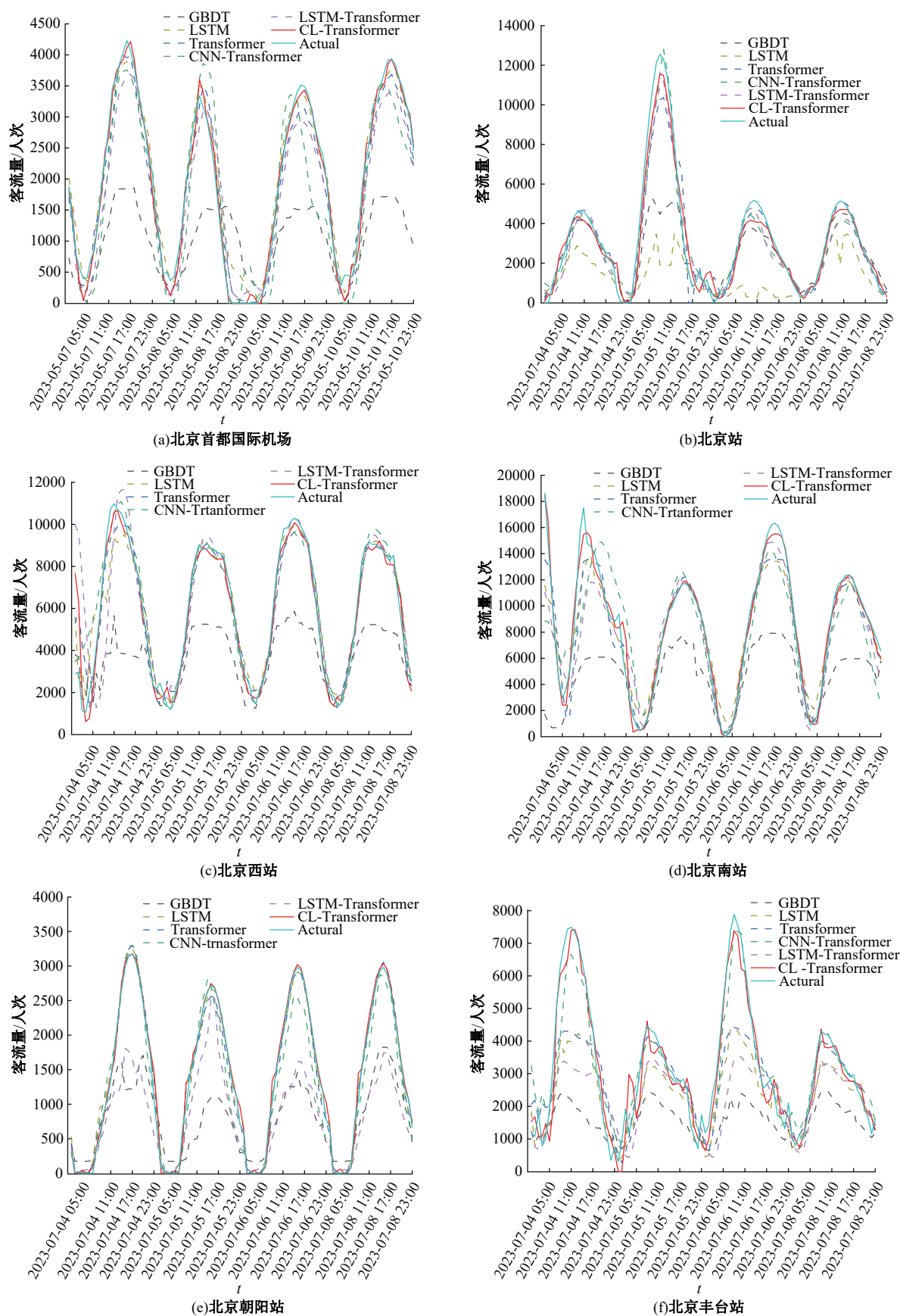


图 4 不同模型对各个枢纽抵站客流预测结果比较

Fig. 4 Comparison of prediction results from different models for arrival passenger flow at various hubs

4 结 论

(1)本文考虑了本土新增新冠肺炎病例数及防控政策等因素的影响,以北京首都国际机场及7个高铁站点的抵站客流数据为基础,通过将CNN、LSTM与Transformer模型有效结合,构建的CL-Transformer模型提升了客流预测精度。相比于GBDT、LSTM、Transformer,以及CNN-Transformer和LSTM-Transformer等混合模型,CL-Transformer模型对北京市8个对外客运枢纽的抵站客流规模预测准确率超过80%,对其中6个场站的预测准确率在90%左右,表现出良好的预测准确率与稳定性。

(2)本文研究结果表明:置换特征重要性方法可以应用于深度学习算法特征重要性的识别。不仅通常考虑的时段属性、预售票量、天气状况、日期属性、节假日属性均对抵站客流预测产生重要影响,而且疫情相关的因素同样对抵站客流预测具有重要影响。

(3)本文提出的CL-Transformer模型能较好地适应和满足政策管控等特殊场景中的客流预测需求。未来可以考虑航班出发地的天气等时空因素,以及出发地与目的地之间的联系紧密程度,构建客流预测模型,为多个场站提供更准确的预测方案。

参考文献:

- [1] 刘咏. 城市对外综合客运枢纽功能定位及相关问题研究[D]. 四川:西南交通大学交通运输学院, 2008.
Liu Yong. Research on functional orientation and related issues of urban comprehensive passenger transport hubs[D]. Sichuan: School of Transportation, Southwest Jiaotong University, 2008.
- [2] Liu L, Chen R C, Zhu S. Passenger flow prediction using weather data for metro systems[C]//Proceedings of the 2018 Conference on Technologies and Applications of Artificial Intelligence. Taichung: IEEE, 2018: 70-73.
- [3] 巫威眺, 曾坤, 周伟, 等. 基于多源数据和响应面优化的公交客流预测深度学习[J]. 吉林大学学报:工学版, 2023, 53(7): 2001-2015.
Wu Wei-tiao, Zeng Kun, Zhou Wei, et al. Deep learning method for bus passenger flow prediction based on multi-source data and surrogate-based optimization[J]. Journal of Jilin University(Engineering and Technology Edition), 2023, 53(7): 2001-2015.
- [4] Guo Z, Zhao X, Chen Y, et al. Short-term passenger flow forecast of urban rail transit based on GPR and KRR[J]. IET Intelligent Transport Systems, 2019, 13(9): 1374-1382.
- [5] Zhang J, Chen F, Cui Z, et al. Deep learning architecture for short-term passenger flow forecasting in urban rail transit[J]. IEEE Transactions on Intelligent Transportation Systems, 2021, 22(11): 7004-7014.
- [6] Sha S, Li J, Zhang K, et al. RNN-based subway passenger flow rolling prediction[J]. IEEE Access, 2020, 8: 15232-15240.
- [7] Jiao F, Huang L, Song R, et al. An improved STL-LSTM model for daily bus passenger flow prediction during the COVID-19 pandemic[J]. Sensors, 2021, 21(17): No. 5950.
- [8] Zhang L, Yang H, Wang K, et al. The impact of COVID-19 on airline passenger travel behavior: An exploratory analysis on the chinese aviation market[J]. Journal of Air Transport Management, 2021, 95: No. 102084.
- [9] Qian H, Yang Z, Weng J, et al. Short-term passenger flow prediction of metro stations around sports events based on afc data[C]//The 20th COTA International Conference of Transportation Professionals, Online, 2020:70-73.
- [10] Hochreiter S, Schmidhuber J. Long short-term memory[J]. Neural Computation, 1997, 9(8): 1735-1780.
- [11] 张惠臻, 高正凯, 李建强, 等. 基于循环神经网络的城市轨道交通短时客流预测[J]. 吉林大学学报:工学版, 2023, 53(2): 430-438.
Zhang Hui-zhen, Gao Zheng-kai, Li Jian-qiang, et al. Short-term passenger flow forecasting of urban rail transit based on recurrent neural network[J]. Journal of Jilin University(Engineering and Technology Edition), 2023, 53(2): 430-438.
- [12] Zhang X, Zhong S, Mahadevan S. Airport surface movement prediction and safety assessment with spatial-temporal graph convolutional neural network[J]. Transportation Research Part C: Emerging Technologies, 2022, 144: No. 103873.
- [13] Zhang J, Che H, Chen F, et al. Short-term origin-destination demand prediction in urban rail transit systems: a channel-wise attentive split-convolutional neural network method[J]. Transportation Research Part C: Emerging Technologies, 2021, 124: No.

- 102928.
- [14] Cheng Z, Lu J, Zhou H, et al. Short-term traffic flow prediction: an integrated method of econometrics and hybrid deep learning[J]. IEEE Transactions on Intelligent Transportation Systems, 2022, 23(6): 5231-5244.
- [15] Hu S, Xiong C. High-dimensional population inflow time series forecasting via an interpretable hierarchical transformer[J]. Transportation Research Part C: Emerging Technologies, 2023, 146: No. 103962.
- [16] Zhou H, Li J, Zhang S, et al. Expanding the prediction capacity in long sequence time-series forecasting [J]. Artificial Intelligence, 2023, 318: No. 103886.
- [17] Ramana K, Srivastava G, Kumar M R, et al. A vision transformer approach for traffic congestion prediction in urban areas[J]. IEEE Transactions on Intelligent Transportation Systems, 2023, 24(4): 3922-3934.
- [18] Fisher A, Rudin C, Dominici F. All models are wrong, but many are useful: Learning a variable's importance by studying an entire class of prediction models simultaneously[J]. Journal of Machine Learning Research, 2019, 20: 1-8.