

面向短时地铁客流量预测的混合深度学习模型

彭桐歆^{1,2}, 韩 勇^{1,2}, 王 程³, 张志浩^{1,2}

(1.中国海洋大学 信息科学与工程学院, 山东 青岛 266100; 2.青岛海洋科学与技术国家实验室 区域海洋动力学与数值模拟功能实验室, 山东 青岛 266237; 3.青岛市市北区党建引领基层治理推进中心, 山东 青岛 266000)

摘 要: 城市交通客流量精准预测是智能交通系统的重要环节,是有效管控交通、规划最佳出行线路的关键。目前城市交通客流量短时预测研究主要集中在利用深度学习模型进行时空特征的提取,忽略了对模型优化的研究。针对短时地铁客流量预测存在的问题,提出一种混合深度学习模型 ResGRU_{Metro},将卷积神经网络、残差单元和门控循环单元相结合,捕获流量数据的时空特征。针对深度学习模型常用的损失函数难以对交通客流量峰值进行精准预测的问题,引入面向短时交通流量预测的加权平方误差,根据交通客流量的大小为预测误差赋予不同权重,并加大对交通客流量峰值处误差的惩罚,使神经网络在反向传播时更加关注峰值处的预测和误差,从而提升交通客流量峰值的预测精度。此外,通过耦合天气、空气质量等外部因子,改善模型的整体预测性能,增强模型的稳定性。实验结果表明,相比 LR、PSVR、CNN 等典型的预测模型,ResGRU_{Metro} 模型有更高的预测精度,能够准确预测交通客流量的峰值。

关键词: 智能交通;短时客流量预测;时空特征;残差神经网络;门控循环单元;加权平方误差

开放科学(资源服务)标志码(OSID):



中文引用格式: 彭桐歆, 韩勇, 王程, 等. 面向短时地铁客流量预测的混合深度学习模型[J]. 计算机工程, 2022, 48(5): 297-305.

英文引用格式: PENG T X, HAN Y, WANG C, et al. Hybrid deep-learning model for short-term metro passenger flow prediction[J]. Computer Engineering, 2022, 48(5): 297-305.

Hybrid Deep-learning Model for Short-term Metro Passenger Flow Prediction

PENG Tongxin^{1,2}, HAN Yong^{1,2}, WANG Cheng³, ZHANG Zhihao^{1,2}

(1.College of Information Science and Engineering, Ocean University of China, Qingdao, Shandong 266100, China; 2.Laboratory for Regional Oceanography and Numerical Modeling, Qingdao National Laboratory for Marine Science and Technology, Qingdao, Shandong 266237, China; 3.Center of Grassroots Governance Led by the Chinese Communist Party in Shibei District, Qingdao, Shandong 266000, China)

[Abstract] Accurate prediction of traffic passenger flow is an essential part of an Intelligent Transportation System (ITS), which helps manage traffic and plan the best travel routes. Presently, the short-term prediction of urban traffic passenger flow mainly focuses on using deep-learning models to extract spatiotemporal features, which neglects the model optimization. A hybrid deep learning model, ResGRU_{Metro}, is proposed to predict short-term metro passenger flow. It combines Convolutional Neural Networks (CNN), Residual Units (ResUnits), and Gated Recurrent Units (GRU) to capture the spatiotemporal dependency of flow data. A Weighted Square Error (WSE) with a bias towards peak traffic passenger flow is proposed to address the problem that the classic loss function used in deep learning models cannot capture the peak hours features. The WSE applies different weights to the prediction errors according to the traffic passenger flow, which may increase the penalty for the peak traffic passenger flow error. Additionally, it helps the neural network pay more attention to the prediction and error of the peak traffic passenger flow during the backpropagation period, which may improve the prediction accuracy of the peak traffic passenger flow. In addition, the model incorporates external factors, such as weather and air quality, which improves the overall prediction performance and enhances the model stability. The results show that the hybrid ResGRU_{Metro} model has a more accurate prediction ability than some typical prediction models such as LR, PSVR, and CNN. In addition, hybrid ResGRU_{Metro} can accurately predict the peak value of passenger traffic flow.

基金项目: 山东省自然科学基金面上项目(ZR2020MD020)。

作者简介: 彭桐歆(1996—),女,硕士研究生,主研方向为时空大数据挖掘;韩 勇(通信作者),教授;王 程、张志浩,硕士研究生。

收稿日期: 2021-03-29 **修回日期:** 2021-07-07 **E-mail:** yonghan@ouc.edu.cn

[Key words] intelligent transportation; short-term passenger flow prediction; spatio-temporal feature; Recurrent Neural Networks(RNN); Gated Recurrent Units (GRU); Weighted Square Error (WSE)

DOI:10.19678/j.issn.1000-3428.0061309

0 概述

随着传感器技术的飞速发展,记录人们日常行为的城市时空大数据如公交卡数据、出租车轨迹等越来越丰富。这些数据可以模拟和预测人类的活动模式,揭示人类社会活动及其运动规律,为构建智能交通系统(Intelligent Transportation System, ITS)提供可靠支持。通过分析地铁站的历史刷卡记录,对不同站点的短时客流量进行预测,不仅能方便出行者规划线路,而且能为管理部门调度车次提供依据,对促进智慧城市的建设具有重要的现实意义。

城市交通客流量预测的研究历程可归结为统计学方法、传统机器学习方法和深度学习方法3个阶段。统计学方法包括差分整合移动平均自回归模型(Autoregressive Integrated Moving Average Model, ARIMA)^[1]及其变体模型^[2-3]、逻辑回归模型(Logistic Regression, LR)^[4]、卡尔曼滤波^[5]等。统计学方法能捕获时空序列中的线性特征,但对数据中的非线性特征却无法提取。因此,研究人员引入了传统机器学习方法,传统机器学习方法包括决策树、聚类、贝叶斯方法、支持向量机、最大期望算法、Adaboost、人工神经网络等。ZHANG等^[6]采用支持向量机(Support Vector Machine, SVM)对交通流量进行预测,LI等^[7]融合季节性ARIMA和SVR对地铁客流量进行预测,取得较好的结果。此外,还有不少研究人员在交通客流量预测上采用贝叶斯方法^[8]、BP神经网络^[9]、随机森林^[10]等方法,均取得了良好的结果。

然而在大数据时代,复杂且庞大的时空数据使建立特征工程的难度增加。相比统计学和传统机器学习方法,深度学习属于机器学习的一个分支,它利用深度神经网络来解决特征表达的过程。深度神经网络是一种相对特殊的神经网络,它包含较多的隐含人工神经网络层,可以学习复杂的非线性关系,以减小特征工程的复杂度。ZHAO等^[11]利用长短期记忆(Long Short-Term Memory, LSTM)网络进行交通流量预测。LI等^[12]提出基于残差LSTM的交通流量预测方法,显式建模特征维度之间的依赖关系。ZHANG等^[13]利用结构相对简单的门控循环单元(Gated Recurrent Units, GRU)预测交通流。ZHANG等^[14]将城市范围的时空流量序列表达为规则格网数据,以小时、天和周的周期提取历史观察值,并送入多层卷积神经网络(Conventional Neural Networks, CNN)进行建模。ZHAO等^[15]提出混合深度学习模型ResNet-CNN1D,对青岛市地铁客流量进行建模预测。

在现有的研究中,单纯使用循环神经网络(Recurrent Neural Networks, RNN)^[16-17]只能捕获数据的时间特征,使用单一的CNN^[18-19]只能捕获数据的空间特征,部分模型缺乏对流量数据时空特征的捕获,且忽视了历史周期片段对目标时刻的影响。此外,现有研究主要集

中在如何将深度学习模型应用于短时交通流量预测,忽略了对模型的优化研究。交通流量典型的特征之一是交通峰值,现有模型对峰值的捕捉能力不够,极少关注天气、空气状况等多源外部信息。

为精准预测短时地铁客流量,本文构建一种混合深度学习模型ResGRU_{Metro},通过融合残差神经网络(Residual Neural Network, ResNet)提取交通流量的空间特征,运用GRU提取流量的时间特征。该模型将近邻、日、周周期模式的时间片流量数据作为输入,利用混合深度学习模型挖掘数据的时空特征,使用加权平方误差(Weighted Square Error, WSE)对模型损失函数进行改进,加大对交通流量序列峰值处误差的惩罚,提高模型对交通流量中峰值的预测能力。此外,模型耦合天气、空气质量、节假日等外部因子,改进模型整体预测精度,得到目标时刻所有站点的客流量预测结果。

1 问题定义

本文拟解决的问题为利用混合深度学习模型对特定时间片的地铁客流量作精准预测。地铁客流量受历史时段客流量的影响,也受天气等外部因素的影响,该预测问题如图1所示,可用公式表示为:

$$F_{\text{prediction}} = f(X, E, W) \quad (1)$$

其中: $F_{\text{prediction}}$ 为预测问题; f 为解决预测问题的混合深度学习模型; X 为历史客流量输入; E 为外部因子输入; W 为可学习权重。

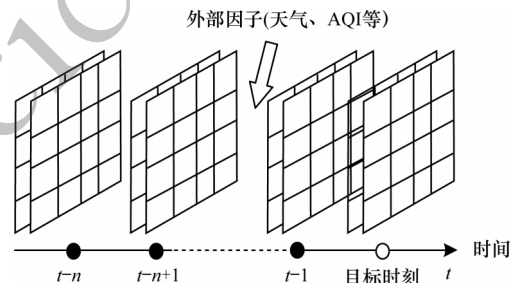


图1 预测问题定义

Fig.1 Definition of the prediction problem

历史客流输入包括近邻、日周期和周周期模式的3部分历史客流量数据,应用3种不同周期的模式能够更好地研究地铁客流的周期性及其周期时间片对目标时刻客流的影响,以达到更精准的预测。近邻模式表示最接近预测时刻的若干时间片数据,日周期和周周期模式与目标时刻处于相同时间片,但在日或周的周期循环中。如预测目标是周六上午7:00~7:10的地铁网络中各站点的进出站客流量,那么近邻模式的输入就是最接近周六上午7:00的若干时间片的客流量,日周期模式的输入是前N天每天上午7:00~7:10的客流量,周周期模式的输入是前N周每周周六上午7:00~7:10的客流量。3种模式的输入可用式(2)表示:

$$\begin{aligned} I_c &= (X_{t-1}, X_{t-2}, \dots, X_{t-c}) \\ I_d &= (X_{t-1 \times n}, X_{t-2 \times n}, \dots, X_{t-d \times n}) \\ I_w &= (X_{t-1 \times 7 \times n}, X_{t-2 \times 7 \times n}, \dots, X_{t-w \times 7 \times n}) \end{aligned} \quad (2)$$

其中: X_t 表示第 t 个时间片的客流量数据, 即目标时间片客流量; I_c 、 I_d 和 I_w 分别表示近邻、日周期和周周期模式的历史观测数据; I_c 、 I_d 和 I_w 中的时间片的数量分别是 c 、 d 和 w ; 一天的时间片总数是 n 。

2 短时地铁客流预测模型

2.1 模型架构

地铁客流量是典型的时空数据, 其时间特征表现在历史客流量对预测时间点客流量的影响, 空间特征表现为前一站或前几站的客流量对当前站点的影响。因此, 深度学习预测模型需考虑地铁客流量的时空属性。本文提出的混合深度学习模型简称 ResGRU_{Metro}, 其结构如图2所示。模型包括3个部分, 第1部分用于捕获空间特征, 将 CNN 与残差单元 (Residual Units, ResUnits) 融合, 构建 ResNet, 第2部分主要利用 GRU 挖掘时间特征。此外, 利用 LSTM 耦合温度、空气质量等外部因子。作为第3部分改善模型预测的整体能力。最后利用参数矩阵融合的方式, 将时空特征和外部特征融合输出, 通过激活函数得到最终预测值。为更好地捕捉客流量的峰值, 设计 WSE 作为损失函数反向传播训练, 加强对流量峰值的捕获。

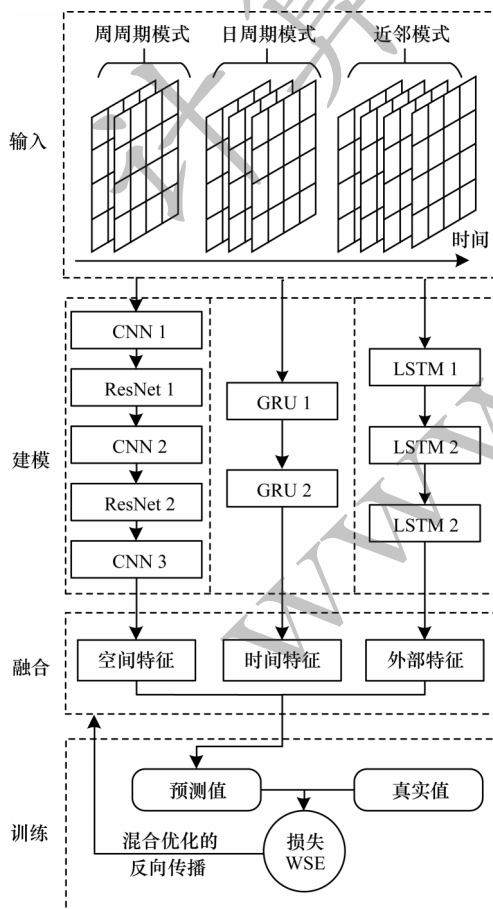


图2 ResGRU_{Metro} 模型结构

Fig.2 Structure of ResGRU_{Metro} model

2.2 时空特征捕捉

CNN 是一类包含卷积计算且具有深度结构的神经网络, 擅长处理具有类似于网格结构数据的神经网络, 例如时间序列数据和图像数据, CNN 将其分别看作是在时间轴上经过有规律地采样形成的一维网络和二维像素网格。每个 CNN 单层网络包括卷积、非线性变换、下采样 3 个部分, 每层的输入输出为一组向量构成的特征图, CNN 通过 Kernel 系统逐层对输入信号完成映射, 并对提取的特征进行非线性映射。最后, 通过下采样阶段对前两步得到的特征信息进行采样。图3是一个在二维向量上进行卷积运算的例子, 图中仅处理全位于图像中的输入, 由图3可以说明输出张量是由输入张量通过卷积得到的。就地铁客流量而言, 每个地铁站有由历史客流量组成的特征向量, 为捕捉地铁站间的空间关联特征, 将地铁网络数据转换为二维图像, 水平轴表示时间, 垂直轴代表地铁站, 每个位置的值为历史地铁客流量, 通过对图像执行卷积来进行预测。

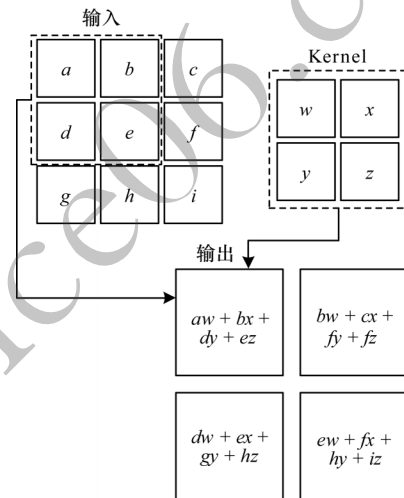


图3 卷积示例图

Fig.3 Convolution example diagram

传统 CNN 无法达到较深结构, 随着隐藏层的增加, 会出现梯度消失、网络退化、梯度爆炸等问题, 损害模型的学习能力。本文将残差结构融入到 CNN 中, 构建 ResCNN 单元。一个典型的 ResCNN 单元的结构如式(3)所示:

$$X^{L+1} = F_r(X^L) + X^L \quad (3)$$

其中: X^L 为第 L 个残差单元的输入; X^{L+1} 为第 $(L+1)$ 个残差单元的输出; F_r 是残差函数。

残差结构的原理是学习关于 X^L 的残差函数 F_r , 文中 1 个 ResCNN 单元由 2 个堆叠的 CNN 层和 1 个捷径连接组成, 如图4所示。本文采用 ReLU 作为残差结构中的激活函数, 通过堆叠若干个 ResCNN 单元, 形成一个残差网络 ResNet, 以达到更深层的特征获取。

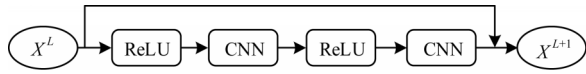


图4 ResCNN单元结构

Fig.4 Structure of ResCNN unit

为建模交通客流量在时间维度上的依赖,考虑利用RNN的循环机制,但传统RNN很难保持长期的依赖性,且存在梯度消失、梯度爆炸的问题^[20]。LSTM和GRU是具有特殊结构的RNN,通过门控机制有针对性地保留信息,能够学习时间跨度相对较长的依赖关系,解决传统RNN存在的问题。相较于LSTM,GRU在保持LSTM原有效果的前提下有更简单的结构和更快的学习速度,因此本文采用GRU捕获交通客流量的时间特征。GRU的结构如图5所示,1个GRU单元包含2个门,即重置门和更新门,重置门用来丢弃与预测无关的历史信息,更新门用于控制历史状态信息对当前时刻的影响程度,更新门的值越大,状态信息保留越多。

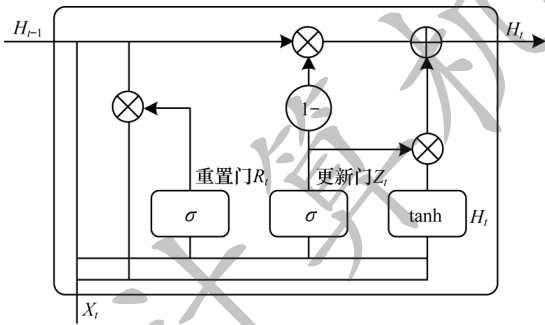


图5 GRU单元结构

Fig.5 Structure of the GRU unit

具体传播公式如式(4)所示:

$$R_t = \sigma(W_{XR}X_t + W_{HR}H_{t-1} + b_R)$$

$$Z_t = \sigma(W_{XZ}X_t + W_{HZ}H_{t-1} + b_Z)$$

$$\tilde{H}_t = \tanh(W_{XH}X_t + W_{HH}(R_t \odot H_{t-1}) + b_H)$$

$$H_t = Z_t \odot H_{t-1} + (1 - Z_t) \odot \tilde{H}_t \quad (4)$$

其中: R_t 、 Z_t 、 \tilde{H}_t 和 H_t 分别为重置门、更新门、当前时刻的候选隐藏状态和隐藏状态; W_{XR} 、 W_{XZ} 和 W_{XH} 分别为输入层到重置门、更新门和隐藏状态的权重矩阵; W_{HR} 、 W_{HZ} 和 W_{HH} 分别为隐藏状态到重置门、更新门和另一隐藏状态的权重矩阵; b_R 、 b_Z 、 b_H 分别为更新门、重置门和候选隐藏状态的偏置向量; H_{t-1} 为历史时刻的隐藏状态; X_t 为当前时刻的原始输入; σ 表示激活函数; \odot 表示哈达玛积。

2.3 外部特征耦合

外部因素包括天气、重大事件、空气质量等影响人们的出行计划,进而对地铁客流量产生影响。例如,暴风雨天气会降低人们的出行欲望,节假日等重

大事件会使城市交通迎来新的高峰,空气质量较差的日子人们大概率选择居家活动,这些外部因素潜移默化地影响人们的出行决策。目前,只有部分客流量预测模型引入了外部因素,且对空气质量的关注相对较少。本文选取的11个外部指标可以分为3类,包括天气状况(最高气温、最低气温、是否阴雨)、特殊事件(是否节假日)和空气质量(AQI、PM2.5、PM10、NO₂、CO、O₃、SO₂浓度),利用LSTM网络对这些因子的时间序列进行分析,捕捉外部因子影响特征,优化模型预测精度。

由于3个组成部分对预测结果的影响程度不同,本文采用参数矩阵方法对3个部分进行融合。参数矩阵方法由ZHANG等^[21]提出,具体步骤为:

- 1)初始化参数矩阵,矩阵元素值范围为[0,1];
- 2)各部分原始输出与矩阵做哈达玛积运算后相加融合;
- 3)通过迭代训练获得最小损失值的方式并确定最优权重。具体如式(5)所示:

$$O = \sigma(W_1 \odot O_1 + W_2 \odot O_2 + W_3 \odot O_3) \quad (5)$$

其中: O 表示模型的最终输出; \odot 表示哈达玛积; W_1 、 W_2 和 W_3 表示可学习权重; O_1 、 O_2 和 O_3 分别是模型前3个部分的输出; σ 是相应的激活函数。

2.4 训练损失优化

短时交通流量预测是一种典型的回归问题,研究人员普遍采用均方误差(Mean Square Error, MSE)作为训练的损失函数。MSE是对所有样本预测值和真实值的方差取平均值,所有样本在计算误差时均会被附以相同的权重,它更侧重于误差的整体计量。然而在真实生活中,无论是对于出行者或是交通的管理者而言,都更加关注交通时序的峰值,以方便对出行时间进行合理规划及线路车次管理。MSE采用算数平均的方式,难以有针对性地满足对于交通时序峰值的精准预测。为解决此问题,本文设计了一种面向短时交通流量时间序列预测的损失函数加权平方误差,使其更加关注对于交通流量序列峰值的预测误差,从而起到对损失函数进行优化的效果。WSE的计算过程如式(6)所示:

$$W_{WSE} = \frac{\sum_{i=1}^n y_i \times (\hat{y}_i - y_i)^2}{\sum_{i=1}^n y_i} \quad (6)$$

其中: \hat{y}_i 为预测值; y_i 为真实值; n 为总时间切片量; i 为某一特定的时间切片。

WSE在计算损失时根据交通流量的大小为预测误差赋予不同权重, 加大了对交通流量时间序列峰值处误差的惩罚, 这使得神经网络在反向传播时, 会更加关注峰值处的预测和误差, 最终在交通流量峰值处取得更为精确的预测。

3 实验结果与分析

3.1 数据集与评价指标

数据选自2015年4月上海市地铁进出站刷卡数据, 共有289个站点, 覆盖全市14条地铁线, 平均每天有900万条刷卡记录。据市民活动情况, 本文选取6:40~23:00为目标时间段进行客流量研究, 以10 min为时间切片, 即每10 min统计一次客流量, 则30天内共划分出2 940个时间片。外部因子中的天气数据来自天气网数据, 以h为单位的空气质量数据来自中国环境监测总站提供的城市空气质量实时监测数据。训练数据来自4月1日~4月23日, 验证数据来自4月24日和4月25日(含1个工作日和1个非工作日), 测试数据来自4月26日~4月30日(含4个工作日和1个非工作日)。

为评估模型预测效果, 选用平均绝对误差(Mean Absolute Error, MAE)、均方根误差(Root Mean Square Error, RMSE)、加权平均绝对百分比误差(Weighed Mean Absolute Percentage Error, WMAPE)进行量化分析, 误差计算公式为:

$$M_{MAE} = \frac{1}{n} \sum_{i=1}^n |\hat{y}_i - y_i|$$

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

$$W_{WMAPE} = \sum_{i=1}^n \left(\frac{y_i}{\sum_{j=1}^n y_j} \left| \frac{\hat{y}_i - y_i}{y_i} \right| \right) \quad (7)$$

3.2 实验环境与模型参数

本文实验在GPU平台运行, 运行环境为python3.6, 利用python的第三方库scikit-learn、Keras和TensorFlow搭建交通流量预测模型。以10 min为时间间隔统计流量数据, 通过多次实验测试和调参, 对模型的卷积核尺寸、CNN神经元个数、残差单元个数、CRU神经元个数进行讨论, 结果如图6所示。需说明的是由于MAE、WMAPE指标波动较小, 这里仅用RMSE进行比较判断。最终模型的超参数设置: 卷积核大小为3×3; 3层CNN神经元个数均为128; 残差单元个数分别均为16, 2层GRU神经元的个数均为256, 根据经验设置3层LSTM神经元, 个数分别为256、128、64; 输入数据周周期模式、日周期模式和近邻模式时间切片分别为7、1、1; 在模型训练过程中, 学习率为 1×10^{-3} , 迭代次数为600次, 使用验证集早停策略以避免过拟合, 采用Adam优化器, 损失函数为WSE。

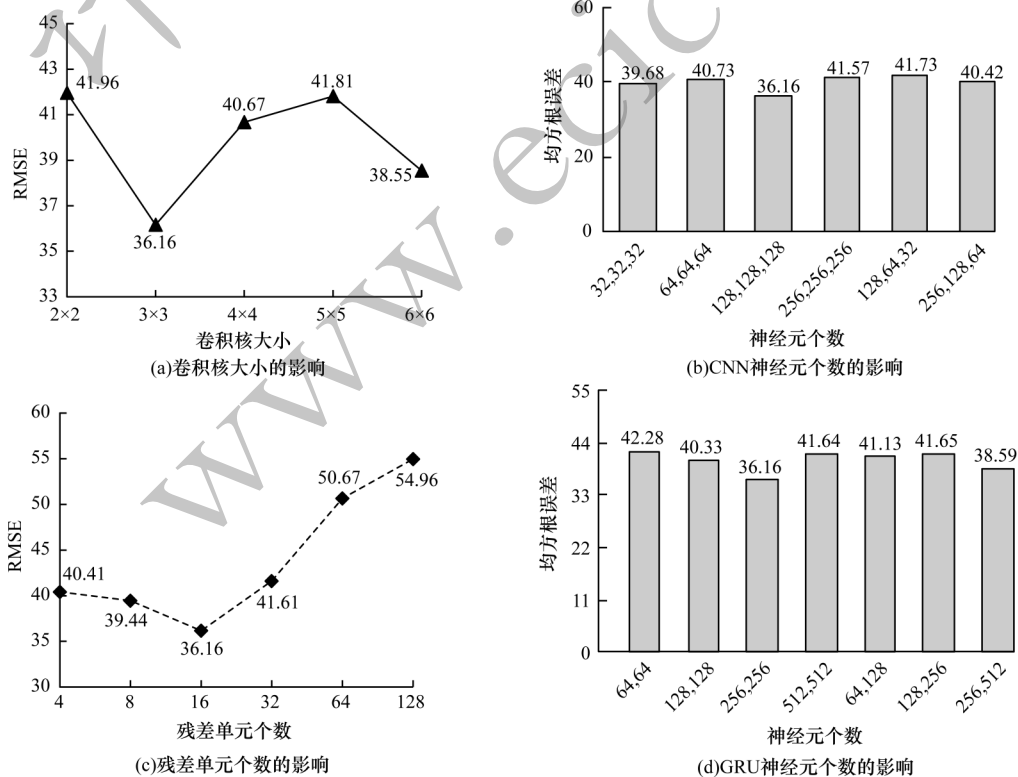


图6 本文模型的超参数调参

Fig.6 Optimal hyperparameters of model in this paper

3.3 不同基准模型结果

分别以 10 min、20 min 和 30 min 为时间间隔进行流量统计,将本文提出的 ResGRU_{Metro} 模型与已有经典模型进行比较,验证模型的预测性能。基准模型包括 1 个统计学方法、1 个机器学习方法和 4 个深度学习方法。

1)线性回归分析 LR^[22]:一种统计学方法,用于确定两种及以上变量间的依赖关系。

2)支持向量机 PSVR^[23]:一种机器学习方法,scikit-learn 核函数类型为“poly”。

3)卷积神经网络 CNN^[24]:一种深度学习方法。

4)ResNet:加入残差单元的卷积神经网络^[21]。

5)门控循环单元 GRU^[25],深度学习方法。

6)ResNet+GRU:ResGRU_{Metro} 模型除去天气因子和 WSE 损失函数。

对比实验结果如表 1 所示。

表 1 地铁站客流量预测实验结果

Table 1 Experiment results of passenger flow prediction for metro stations

模型	平均绝对误差/人次			均方根误差/人次			加权平均绝对百分比误差/%		
	10 min	20 min	30 min	10 min	20 min	30 min	10 min	20 min	30 min
LR 模型	22.03	56.28	79.50	52.18	85.54	127.09	12.35	10.99	9.44
PSVR 模型	21.92	48.60	44.70	47.85	84.70	117.17	12.38	10.75	9.13
CNN 模型	21.85	32.73	45.31	47.64	82.44	124.51	12.67	9.51	9.04
ResNet 模型	21.68	30.15	44.28	45.92	72.34	109.00	11.97	9.11	8.74
GRU 模型	21.11	28.05	42.56	38.86	67.32	94.67	10.31	9.53	8.66
ResNet+GRU 模型	20.74	28.32	41.71	37.61	65.56	92.30	10.74	9.18	8.51
ResGRU _{Metro} 模型	20.13	27.41	41.56	36.16	63.45	88.01	10.18	9.08	8.25

由表 1 可知,本文提出的 ResGRU_{Metro} 模型在 MAE、RMSE 和 WMAPE 这 3 种指标上较其他方法均取得了较优的结果。深度学习方法整体优于统计学方法和机器学习方法,考虑时空特征的深度学习方法结果优于仅考虑空间特征的深度学习方法。

以 10 min 切片具体分析,相比 LR 模型,ResGRU_{Metro} 模型的 MAE 相对减少了 8.62%,RMSE 相对减少了 30.70%,WMAPE 相对减少了 2.17 个百分点,这可能是因为 LR 模型只能捕捉时间序列中的线性关系,对于非线性特征难以捕获。将 ResGRU_{Metro} 模型与 PSVR 对比,ResGRU_{Metro} 的 MAE 相对减少了 8.17%,RMSE 相对减少了 24.43%,WMAPE 相对减少了 2.20 个百分点,这说明 ResGRU_{Metro} 在交通流量预测上的效果优于传统机器学习模型。与仅捕捉空间关联特征的 CNN 和 ResNet 对比,与仅捕捉时间关联特征的 GRU 对比,ResGRU_{Metro} 均取得了较好的效果,这体现了时空混合模型的优越性。相比 ResNet+GRU 模型,ResGRU_{Metro} 的 MAE 相对减少了 2.94%,RMSE 相对减少了 3.86%,WMAPE 相对减少了 0.56 个百分点,表明引入天气因子以及改进的损失函数能提高模型预测能力。

3.4 模型峰值捕获能力

为直观展示 ResGRU_{Metro} 模型捕获峰值的能力,本文讨论了在特定站点下不同时间片的进出流量预测

情况。徐家汇是上海中央活动区之一,也是上海十大商业中心之一,日进出客流量较大,具有一定研究意义,因此选取徐家汇地铁站作为研究对象。

图 7 是徐家汇站 4 月 26 日~4 月 28 日(含 1 个非工作日和 2 个工作日)不同时间切片预测值与真实值的比较结果。由图 7 可知,各时间切片 ResGRU_{Metro} 模型预测结果与真实值拟合较好,且在峰值的捕获上有优越性。为进一步验证使用 WSE 作为损失函数对于 ResGRU_{Metro} 捕获峰值特征的改进,选取以 MSE 作为损失函数的 ResGRU_{MSE} 模型作为基准模型,与以 WSE 作为损失函数的模型 ResGRU_{WSE} 进行对比实验。计算了所有站点的模型误差,结果如表 2 所示。由表 2 可知,ResGRU_{WSE} 模型的 MAE 和 RMSE 均优于 ResGRU_{MSE} 模型,但 WMAPE 稍有不足。以 10 min 切片为例,与 ResGRU_{MSE} 模型相比,ResGRU_{WSE} 模型的 MAE 相对减少了 2.89%,RMSE 相对减少了 3.13%,WMAPE 相对增加了 0.28 个百分点。MAE 是衡量模型性能的最基本的方法,实验结果表明,在所有时间切片下 ResGRU_{WSE} 模型均具有较低的 MAE 值,表明其具备良好的预测性能。RMSE 可以放大模型中的较大偏差,RMSE 越小则证明 ResGRU_{WSE} 模型的稳定性越好。但 ResGRU_{WSE} 模型具有低 MAE、低 RMSE 与高 WMAPE,表明预测误差可能主要来自流量数据中的低值,而不是来源于峰值。

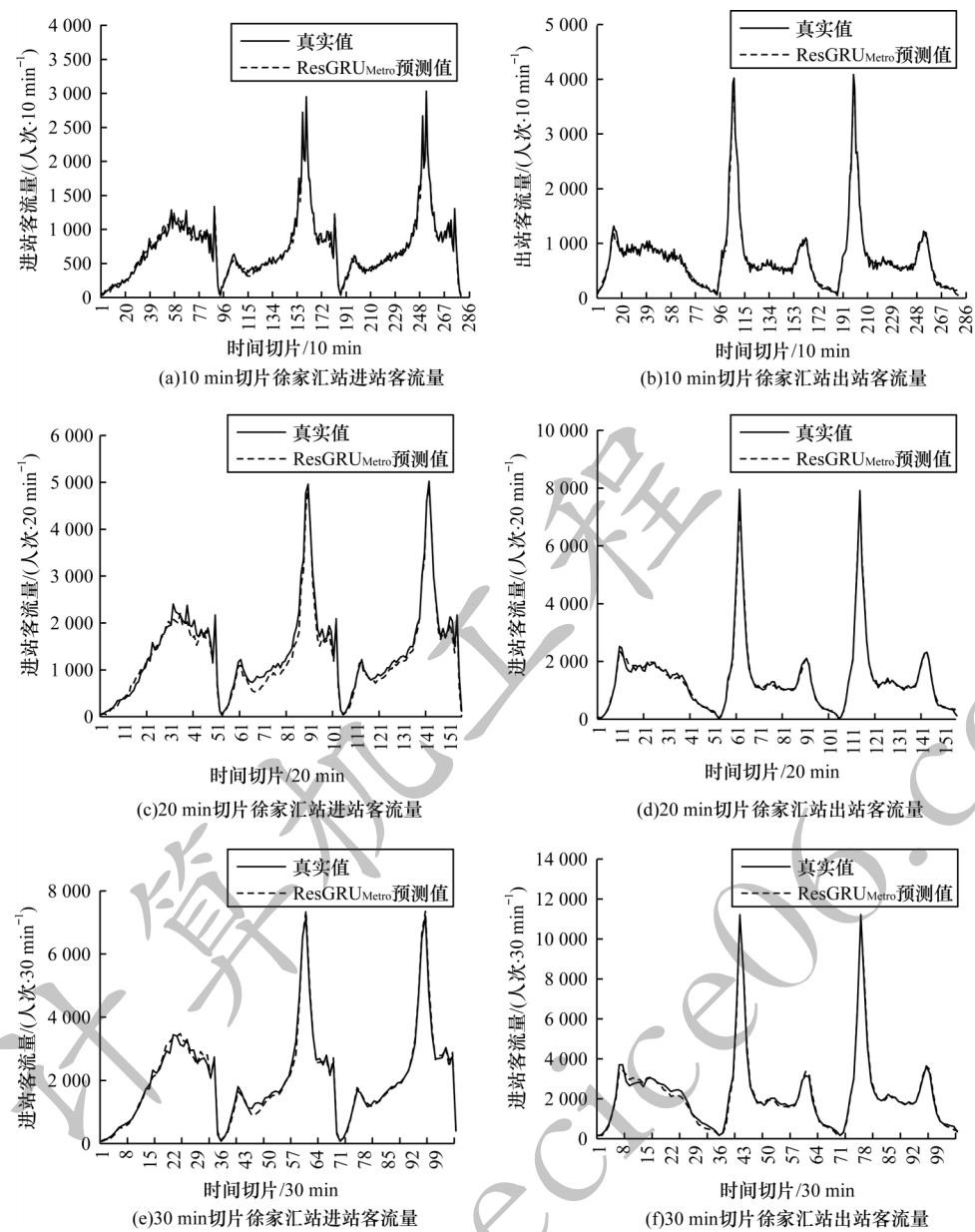


图7 在不同时间切片下徐家汇站进出客流预测结果

Fig.7 Prediction results of passenger inflow and outflow for Xujiahui metro station in different time intervals

模型	平均绝对误差/人次			均方根误差/人次			加权平均绝对百分比误差/%		
	10 min	20 min	30 min	10 min	20 min	30 min	10 min	20 min	30 min
ResGRU _{MSE} 模型	20.73	28.75	41.83	37.33	65.32	91.32	9.90	9.17	8.07
ResGRU _{WSE} 模型	20.13	27.41	41.56	36.16	63.45	88.01	10.18	9.08	8.25

为进一步分析ResGRU_{MSE}和ResGRU_{WSE}模型在不同大小流量处的预测误差,将不同时间切片的流量数据分别划分为4个流量阶段,分别对这4个阶段绘制误差柱状图,如图8所示。可以看出在流量高峰阶段,ResGRU_{WSE}的表现明显优于ResGRU_{MSE},而流量低峰阶段ResGRU_{WSE}的表现略差于ResGRU_{MSE},这可能是由于损失函数WSE使模型更加关注流量序列中的峰

值,从而在一定程度上弱化了对低值处的关注程度。在真实生活中,无论是对出行者亦或是交通的管理者而言,均更加关注交通流量时间序列的峰值,以方便出行时间的合理规划及车辆调度。虽然从总体上看,ResGRU_{WSE}模型的WMAPE并非最低,但是它具有较低的MAE与RMSE,能够更精准地预测交通流量中的峰值,整体预测能力也较强。

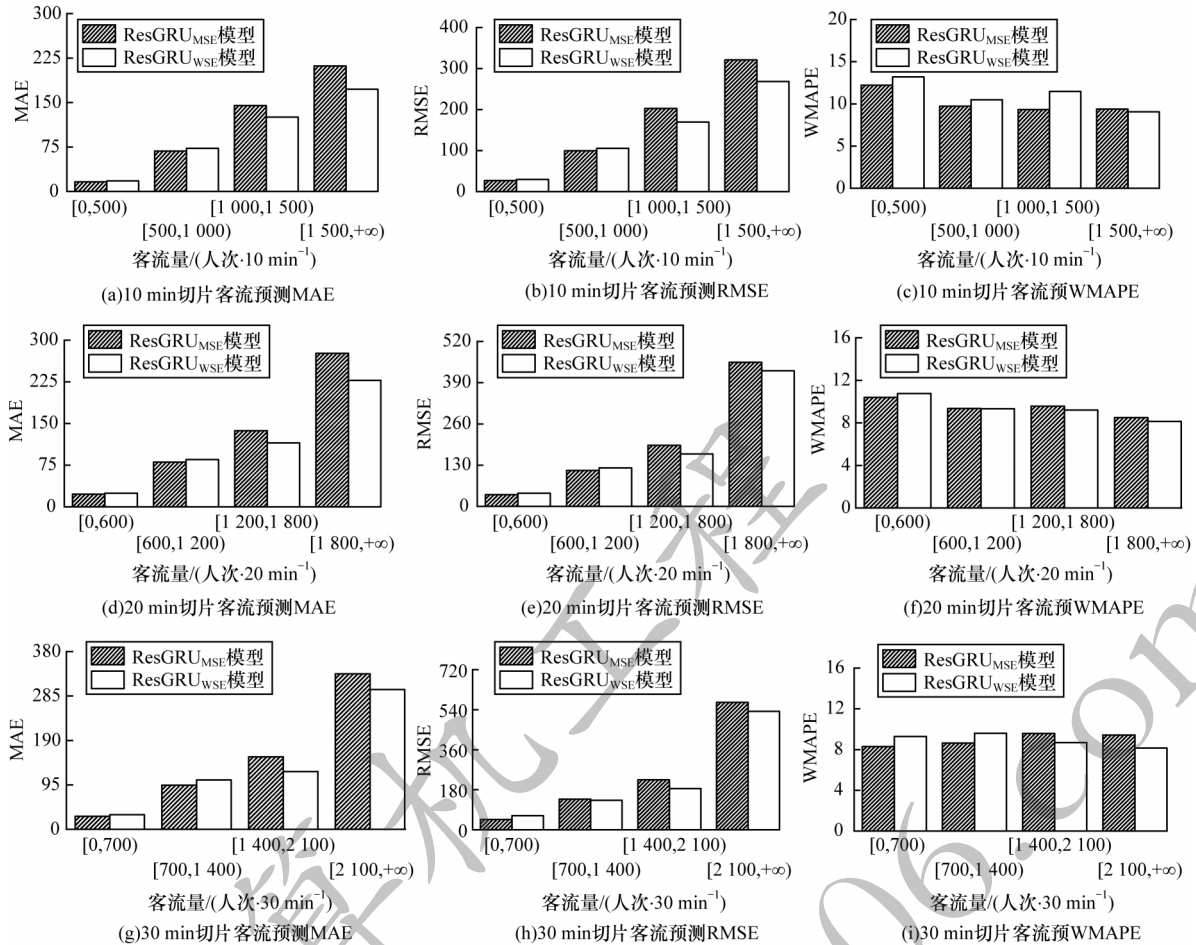


图8 不同流量阶段客流预测误差

Fig.8 Prediction error of passenger flow in different flow stage

4 结束语

针对短时地铁客流量预测问题,本文从模型优化的角度出发,提出一种混合深度学习模型 $\text{ResGRU}_{\text{Metro}}$ 。将卷积神经网络与残差单元相结合,用于捕捉流量的空间特征,并利用门控循环单元捕捉流量的时间特征。为改进模型的预测能力,耦合天气、空气质量等外部因子,在模型训练上使用加权平方误差作为损失函数,提升模型流量峰值的预测能力。此外,探究 $\text{ResGRU}_{\text{Metro}}$ 模型在特定站点不同时间切片下的预测能力,对比以 MSE 和 WSE 作为损失函数的模型预测效果,并量化不同流量阶段的预测误差。实验结果表明,与 LR、PSVR、CNN 等经典模型对比, $\text{ResGRU}_{\text{Metro}}$ 具有较低的 MAE、RMSE 和 WMAPE 与较高的预测精度。下一步将寻找更长时间的交通流量序列对模型进行验证,并将混合模型应用于交通流量的多步预测问题中,以提高模型的鲁棒性。

参考文献

- [1] LIU S Y, LIU S, TIAN Y, et al. Research on forecast of rail traffic flow based on ARIMA model[J]. Journal of Physics: Conference Series, 2021, 1792(1): 12065-12066.
- [2] CHENG T, WANG J Q, HAWORTH J, et al. A dynamic

- spatial weight matrix and localized space-time autoregressive integrated moving average for network modeling[J]. Geographical Analysis, 2014, 46(1): 75-97.
- [3] MILENKOVIĆ M, ŠVADLENKA L, MELICHAR V, et al. SARIMA modelling approach for railway passenger flow forecasting [EB/OL]. [2021-02-29]. https://www.researchgate.net/publication/292088237_SARIMA_modelling_approach_for_railway_passenger_flow_forecasting.
- [4] SMITH B L, WILLIAMS B M, OSWALD R K. Comparison of parametric and nonparametric models for traffic flow forecasting[J]. Transportation Research Part C: Emerging Technologies, 2002, 10(4): 303-321.
- [5] LIANG S D, MA M H, HE S X, et al. Short-term passenger flow prediction in urban public transport: Kalman filtering combined K-nearest neighbor approach[J]. IEEE Access, 2019, 7: 120937-120949.
- [6] ZHANG Y, LIU Y C. Traffic forecasting using least squares support vector machines[J]. Transportmetrica, 2009, 5(3): 193-213.
- [7] LI W, SUI L Y, ZHOU M, et al. Short-term passenger flow forecast for urban rail transit based on multi-source data [EB/OL]. [2021-02-29]. <https://link.springer.com/article/10.1186/s13638-020-01881-4>.
- [8] ROOS J, BONNEVAY S, GAVIN G. Short-term urban rail passenger flow forecasting: a dynamic Bayesian network approach[C]//Proceedings of the 15th IEEE International Conference on Machine Learning and Applications. Washington D. C., USA: IEEE Press, 2016: 1034-1039.

- [9] ZHANG S Y, LIU Z K, SHEN F T, et al. A prediction model of buses passenger flow based on neural networks [J]. *Journal of Physics: Conference Series*, 2020, 1656(1): 12-20.
- [10] LIU L J, CHEN R C, ZHAO Q F, et al. Applying a multistage of input feature combination to random forest for improving MRT passenger flow prediction[J]. *Journal of Ambient Intelligence and Humanized Computing*, 2019, 10(11): 4515-4532.
- [11] ZHAO Z, CHEN W H, WU X M, et al. LSTM network: a deep learning approach for short-term traffic forecast[J]. *IET Intelligent Transport Systems*, 2017, 11(2): 68-75.
- [12] 李月龙, 唐德华, 姜桂圆, 等. 基于维度加权的残差LSTM短期交通流量预测[J]. *计算机工程*, 2019, 45(6): 1-5.
LI Y L, TANG D H, JIANG G Y, et al. Short term traffic flow forecasting based on dimension weighted residual LSTM[J]. *Computer Engineering*, 2019, 45(6): 1-5. (in Chinese)
- [13] ZHANG D, KABUKA M R. Combining weather condition data to predict traffic flow: a GRU-based deep learning approach [J]. *IET Intelligent Transport Systems*, 2018, 12(7): 578-585.
- [14] ZHANG J B, ZHENG Y, QI D K, et al. DNN-based prediction model for spatio-temporal data[C]//*Proceedings of the 24th ACM Sigspatial International Conference on Advances in Geographic Information Systems*. New York, USA: ACM Press, 2016: 92-95.
- [15] 赵建立, 石敬诗, 孙秋霞, 等. 基于混合深度学习的地铁站进出客流量短时预测[J]. *交通运输系统工程与信息*, 2020, 20(5): 128-134.
ZHAO J L, SHI J S, SUN Q X, et al. Short-time inflow and outflow prediction of metro stations based on hybrid deep learning[J]. *Journal of Transportation Systems Engineering and Information Technology*, 2020, 20(5): 128-134. (in Chinese)
- [16] ZHANG J L, CHEN F, SHEN Q. Cluster-based LSTM network for short-term passenger flow forecasting in urban rail transit[J]. *IEEE Access*, 2019, 7: 147653-147671.
- [17] 王祥雪, 许伦辉. 基于深度学习的短时交通流预测研究[J]. *交通运输系统工程与信息*, 2018, 18(1): 81-88.
WANG X X, XU L H. Short-term traffic flow prediction based on deep learning [J]. *Journal of Transportation Systems Engineering and Information Technology*, 2018, 18(1): 81-88. (in Chinese)
- [18] YU D, LIU Y, YU X. A data grouping CNN algorithm for short-term traffic flow forecasting [C]//*Proceedings of International Asia-Pacific Web Conference*. Berlin, Germany: Springer, 2016: 92-103.
- [19] YU F, WEI D, ZHANG S T, et al. 3D CNN-based accurate prediction for large-scale traffic flow[C]//*Proceedings of the 4th International Conference on Intelligent Transportation Engineering*. Washington D. C., USA: IEEE Press, 2019: 99-103.
- [20] LI Y. The combination of CNN, RNN, and DNN for relation extraction [C]//*Proceedings of the 2nd International Conference on Computing and Data Science*. Washington D. C., USA: IEEE Press, 2021: 105-114.
- [21] ZHANG J B, ZHENG Y Z, QI D K. Deep spatio-temporal residual networks for citywide crowd flows prediction[C]//*Proceedings of the 31th AAAI Conference on Artificial Intelligence*. San Francisco, USA: AAAI Press, 2016: 1655-1661.
- [22] OBER P B. Introduction to linear regression analysis[J]. *Journal of Applied Statistics*, 2013, 40(11/12): 2775-2776.
- [23] SAUNDERS C, STITSON M O, WESTON J, et al. Support vector machine[J]. *Computer Science*, 2002, 1(4): 1-28.
- [24] MA X L, DAI Z, HE Z B, et al. Learning traffic as images: a deep convolutional neural network for large-scale transportation network speed prediction[J]. *Sensors*, 2017, 17(4): 818.
- [25] CHO K, VAN MERRIENBOER B, GULCEHRE C, et al. Learning phrase representations using RNN encoder-decoder for statistical machine translation[C]//*Proceedings of 2014 Conference on Empirical Methods in Natural Language Processing*. Stroudsburg, USA: Association for Computational Linguistics, 2014: 1724-1734.

编辑 赖玉玲