

控制与决策

Control and Decision

基于卷积长短时记忆神经网络的城市轨道交通短时客流预测

王秋雯, 陈彦如, 刘媛春

引用本文:

王秋雯, 陈彦如, 刘媛春. 基于卷积长短时记忆神经网络的城市轨道交通短时客流预测[J]. *控制与决策*, 2021, 36(11): 2760–2770.

在线阅读 View online: <https://doi.org/10.13195/j.kzyjc.2020.0501>

您可能感兴趣的其他文章

Articles you may be interested in

基于FWADE-ELM的短时交通流预测方法

Short-term traffic flow forecasting based on hybrid FWADE-ELM

控制与决策. 2021, 36(4): 925–932 <https://doi.org/10.13195/j.kzyjc.2019.1103>

基于FWADE-ELM的短时交通流预测方法

Short-term traffic flow forecasting based on hybrid FWADE-ELM

控制与决策. 2021, 36(4): 925–932 <https://doi.org/10.13195/j.kzyjc.2019.1103>

迁移学习引导的变源域长短时记忆网络建筑负荷预测

Load forecasting of buildings using LSTM based on transfer learning with variable source domain

控制与决策. 2021, 36(10): 2328–2338 <https://doi.org/10.13195/j.kzyjc.2020.0215>

结合注意力机制的循环神经网络复述识别模型

Recurrent neural networks based paraphrase identification model combined with attention mechanism

控制与决策. 2021, 36(1): 152–158 <https://doi.org/10.13195/j.kzyjc.2019.0638>

融合长短时记忆机制的机械臂多场景快速运动规划

Multi-scene rapid motion planning combining with long and short time memory mechanisms for manipulators

控制与决策. 2020, 35(12): 2968–2976 <https://doi.org/10.13195/j.kzyjc.2018.1387>

基于卷积长短时记忆神经网络的城市轨道交通 短时客流预测

王秋雯¹, 陈彦如^{1†}, 刘媛春²

(1. 西南交通大学 经济管理学院, 成都 610031; 2. 江西师范大学 软件学院, 南昌 330022)

摘要: 我国城市轨道交通正处在快速发展阶段, 城轨交通短时客流预测对保障运营安全、优化线网结构, 进而构建智慧城市具有重要意义. 城轨短时客流除了具有周期性、随机性等时间特征之外, 跨时段的断面客流具有相似性, 并且相邻站点客流之间存在空间联系. 对此, 充分考虑以上城轨短时客流的时空特征, 基于卷积长短时记忆神经网络(ConvLSTM)与自适应 k -means 聚类算法, 提出城轨短时客流预测的深度学习模型 k -ConvLSTM, 并通过实验对模型关键参数进行寻优; 同时, 基于深圳市地铁 IC 卡的真实客流数据对模型的有效性进行检验. 结果表明, k -ConvLSTM 在均方根误差、绝对误差均值、绝对误差百分比方面, 均优于仅考虑时空特征的深度学习模型——卷积网络(CNN)与长短时记忆网络(LSTM)的并行混合模型和 ConvLSTM 内嵌式网络模型, 仅考虑时间特征的深度学习模型——LSTM 网络和双向长短时记忆网络(Bi-LSTM), 以及浅层机器学习模型——BP 神经网络和支持向量回归模型(SVR).

关键词: 城轨交通短时客流; 时空特征; 跨时段客流聚类; 卷积长短时记忆神经网络

中图分类号: TP181

文献标志码: A

DOI: 10.13195/j.kzyjc.2020.0501

开放科学(资源服务)标识码(OSID):



引用格式: 王秋雯, 陈彦如, 刘媛春. 基于卷积长短时记忆神经网络的城市轨道交通短时客流预测[J]. 控制与决策, 2021, 36(11): 2760-2770.

Metro short-term traffic flow prediction with ConvLSTM

WANG Qiu-wen¹, CHEN Yan-ru^{1†}, LIU Yuan-chun²

(1. School of Economics and Management, Southwest Jiaotong University, Chengdu 610031, China; 2. School of Software, Jiangxi Normal University, Nanchang 330022, China)

Abstract: China's urban rail transit is developing rapidly. Short-term passenger flow prediction is of great significance for operational safety, network optimization, and then smart city building. While the urban rail passenger flow is cyclical and random in the aspect of temporal characteristics, passenger flows in certain time slots are similar and passenger flows at adjacent stations are spatially correlated. Considering the above spatiotemporal characteristics, this research proposes a deep learning model, named k -means-ConvLSTM (k -ConvLSTM) for urban rail short-term passenger flow prediction based on the ConvLSTM and the adaptive k -means clustering algorithm. Experiments are designed to optimize the key parameters of the model. Also, in order to examine the performance of the proposed model, abundant experiments are conducted based on the real passenger flow data of the Shenzhen Metro IC card. The results show that proposed k -ConvLSTM model performs better than deep learning models that only consider spatiotemporal characteristics — parallel architecture comprising the convolutional neural networks (CNN) and the long short-term memory network (LSTM), the ConvLSTM, and deep learning models that only consider temporal characteristics — the LSTM and the bi-directional long short-term memory network (Bi-LSTM), and shallow learning models — the back propagation neural network (BPNN) and the support vector regression model (SVR), in terms of root mean square errors, mean absolute errors and mean absolute percentage errors.

Keywords: urban rail short-term traffic flow; spatiotemporal characteristics; passenger flows in different time slots clustering; ConvLSTM

收稿日期: 2020-04-30; 修回日期: 2020-08-01.

基金项目: 国家重点研发计划项目(2018YFB1601402).

责任编辑: 陈虹.

†通讯作者. E-mail: chenyanru@swjtu.cn.

0 引言

随着我国城市化进程的加快,交通拥堵和环境污染等城市交通问题日益加剧.建设环保、准时、大运量的城市轨道交通已成为解决城市交通问题的重要措施^[1].然而,由于客流时空分布的失衡,城轨交通断面客流过饱和的情况时有发生.如上海城轨在工作日部分断面客流饱和度高达130%^[2],由此带来巨大的安全隐患.同时,由于城轨的网络化运营,一旦某个站点运营出现异常,将会迅速波及整个网络,产生严重的安全威胁和巨大的经济损失.因此,准确预测城轨交通客流量,并据此制定科学的运营方案,对于保证轨道交通的运营安全、降低运营成本都具有重要的现实意义.

目前,城轨交通客流短时预测方法主要有以下几种:1)数理统计模型,如卡尔曼滤波(Kalman)模型^[3]、灰色模型(grey model, GM)^[4]等;2)人工智能模型,如BP(back propagation)神经网络^[5]、支持向量机(support vector machine, SVM)^[6]、长短时记忆网络(long short term memory network, LSTM)等^[7-8];3)基于以上两类的组合模型,如深度信任网络(deep belief network, DBN)与SVM的组合^[9]、GM与差分整合移动平均自回归模型的组合^[10]、GM与马尔科夫模型的组合^[11]、经验模态分解与BP神经网络的组合^[12]等.以上研究对城轨客流短时预测具有重要的参考价值,但也存在一些不足:数理统计模型能较好地刻画中长期客流的变化趋势,但对于数据的层次分布和随机变化不够敏感,针对随机性较强的短时客流,其预测性能较为有限;人工智能算法的输入数据类型更为灵活,能凭借自身学习能力和自适应能力对非线性规律进行有效捕捉,尤其深度学习因具有更深层次的结构,且强调对特征的学习,能较为准确地描述输入与输出的复杂关联;但单一的数理统计模型或人工智能模型均存在自身的局限性;混合模型则综合各种单一预测模型的优势,在一定程度上能够提高预测精度.此外,以上研究在城轨短时预测中仅考虑了目标站点客流的时间分布,而城轨交通网络化的运营意味着客流除了受到自身历史时段的影响,还会受到相邻站点客流分布的影响,因此,同时考虑目标站点的客流时间特征以及关联站点客流的的空间分布特征能够更加全面地描述城轨短时客流的变化规律.基于此,文献[13]使用主成分分析法(PCA)提取地铁客流的时间特征和空间特征,赋予特征相应权重并构成向量输入随机森林模型中进行预测.该研究全面考虑了客流的时空特征,但特征一旦构成向量就将

被固化,特征有效性得不到后续预测结果的反馈,从而影响预测精度.卷积神经网络(convolutional neural networks, CNN)无需人工对空间分布进行预处理和额外的特征抽取等复杂操作,而是以其特有的细粒度特征提取方式自动处理空间数据;在处理时间特征时,LSTM能够有效避免因数据间隔跨度较长而导致的有效信息消失问题.文献[14]采用并行的CNN和LSTM分别对城轨客流的的空间特征与时间特征进行提取,并对未来20 min的客流进行预测.该研究较精细地提取了城轨客流的时间特征与空间特征,但仍存在一定的局限性,如在CNN与LSTM组成的并行结构中,二者的输入与输出相对独立,且忽略了对不同特征之间关系的提取,因而影响最终预测精度.此外在时间特征方面,现有研究仅考虑客流历史数据,而对客流隐式时间特征鲜有关注,如跨时段的断面客流相似性等^[15].

由于细粒度时空特征的完整提取对高随机性的城轨短时客流的预测精度有重要影响,本文构建一种自适应 k -means聚类法与卷积长短时记忆网络(convolutional long short-term memory, ConvLSTM)相结合的 k -ConvLSTM深度学习模型对城轨短时客流进行预测.其中:ConvLSTM融合了CNN强大的空间特征提取能力与LSTM高效的时序关系专注力,采用在LSTM中内嵌CNN的方式来提取城轨短时客流的时空特征,以避免LSTM与CNN并行组合的缺陷;而自适应 k -means聚类法将通过与主网络层并行的聚类过程,对跨时段的断面客流特征进行提取.三者的结合能够以更细的粒度充分提取城轨短时客流的时空特征,有效降低个体出行行为的不确定性对整体客流分布的随机影响,提升预测精度.此外,传统 k -means聚类法作为无监督式聚类,其聚类数 k 通常需要事先选定;而本文所提出的自适应 k -means聚类法,通过与监督式深度网络连接,使 k 可以实现自适应调整,以准确反映跨时段的断面客流特征.

1 k -ConvLSTM模型构建

ConvLSTM模型^[16]于2015年提出,目前已被应用于地面交通流预测、天气预测和文本分类等方面^[17-18].区别于CNN与LSTM并行的混合模型,ConvLSTM模型是嵌入式的一体化结构,文献[18]指出,该结构的特征提取能力和融合能力优于前者.本文提出的 k -ConvLSTM模型主要通过以下3个方面实现城轨短时客流的预测:1)考虑客流时间分布的纵向随机性、横向周期性和客流空间分布,构建基于前 d 天的 q 个历史时段客流数据的“时空图像”,

作为有监督的主模型的输入数据,并采用 ConvLSTM 网络层对预测目标预测站点及其关联站点的时空关系逐层提取;2)探究跨时段的断面客流相似性,以各个历史时段的客流量作为聚类样本点,通过 k -means 算法生成聚类时段标签,输入网络层以细化时间层

次;3)融合城轨客流的时空特征与跨时段的聚类信息,向前输出城轨短时客流预测值并计算误差,通过误差的逆向反馈实现网络权值和 k -means 聚类数 k 的自动调整.用于城市轨道交通短时客流预测的 k -ConvLSTM 模型框架如图1所示.

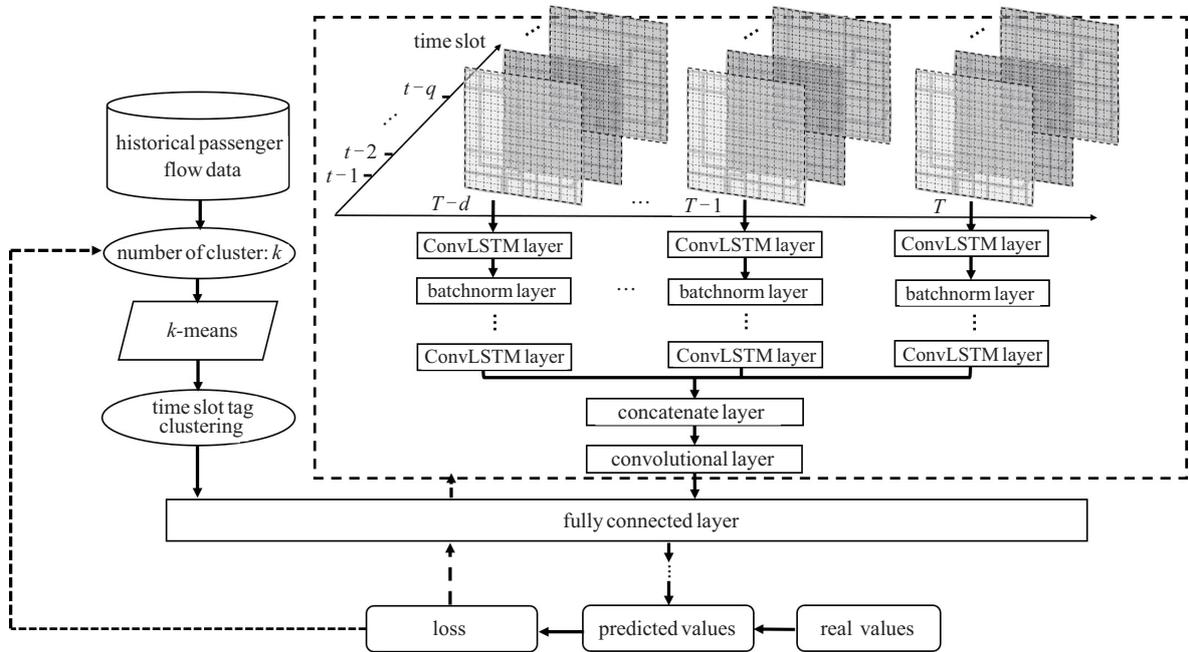


图1 k -ConvLSTM预测模型框架

1.1 输入数据集构造

城轨交通的运营线路整体呈网络形式,各站点的客流分布存在一定的空间联系,并且这种联系在地理位置紧邻、连接居民点与市区商业中心的站点之间尤为明显.因此,城轨客流短时预测除了考虑目标站点客流量的时间分布,还需考虑其关联站点客流的时空分布,本文采用 ConvLSTM 网络层对城轨客流历史的时间和空间特征进行提取.

ConvLSTM 网络对空间特征的学习,主要通过扫描格点化数据的“扫描”来实现.通常,地面交通流的预测可通过广布城市“感知器”以及车辆自带的定位系统等获得实时移动轨迹的图像数据.而城轨列车由于其运行范围包括地面及地下,其客流轨迹图像不易获取,因此,本文考虑客流的出行偏好与站点的相关性,通过类比真实图像构造城轨客流的时空分布“图像”.

通常,一幅完整的图像包含多个图像通道,而每个通道又由数个不可分割的小方格组成,并且小方格有明确的位置和被分配的像素值,从而构成图像的色彩和线条特征.基于此,为合理表达目标站点与关联

站点之间的空间分布关系,本文根据以下步骤构造相应的客流“时空图像”:1)依据站点在平面图上的相对位置,对平面图进行网格化分割,使得目标站点和所有关联站点均被划入对应的小方格内,并且每个小方格至多包含一个站点;2)依据地铁 IC 卡记录,获取相同统计时段下各站客流量作为“时空图像”的“像素值”填入对应站点的小格内,无站点的方格则填入0;3)城轨网络平面图将转换为长为 M 个方格、宽为 N 个方格的客流“时空图像”.其中:目标站点 f 的坐标定义为 (a, b) , $1 < a < M, 1 < b < N$;关联站点的坐标定义为 (m, n) , $m = 1, 2, \dots, M, n = 1, 2, \dots, N, m \neq a, n \neq b$.用 $x_{T,t}^f$ 表示目标站点 f 第 T 日时刻 t 的客流时空“图像”, $x_{T,t}^{a,b} \geq 0, x_{T,t}^{m,n} \geq 0$ 分别对应目标站点 f 及其关联站点在第 T 日时刻 t 的“像素值”,即此刻各站点客流量,则时空信息构造如图2所示.

城轨客流受到多个历史时段客流的影响,针对站点 f 第 T 日时刻 t 的客流预测,应充分考虑当前时刻紧邻的前 q 个时段客流以及前 d 天的客流整体分布的相关性,则 ConvLSTM 模型输入数据集可以表示为

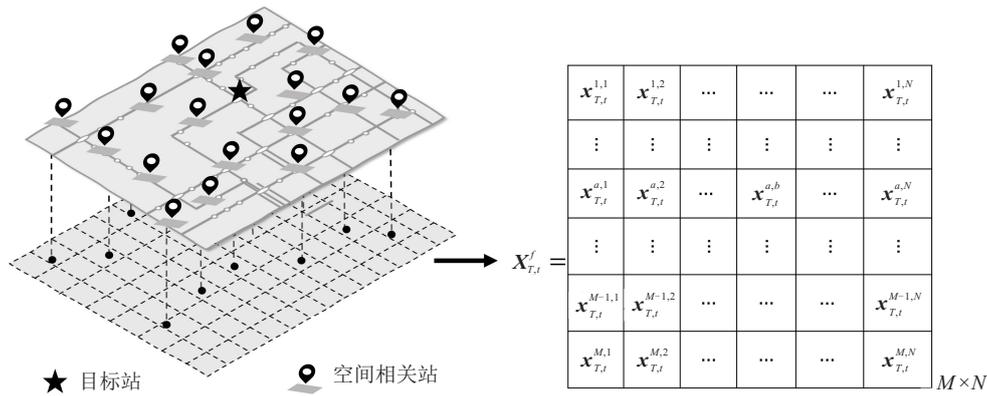


图2 城轨客流时空信息构造

dataset $_{T,t}^f =$

$$\begin{bmatrix} \mathbf{X}_{T,t-1}^f & \cdots & \mathbf{X}_{T-d+1,t-1}^f & \mathbf{X}_{T-d,t-1}^f \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{X}_{T,t-q+1}^f & \cdots & \mathbf{X}_{T-d+1,t-q+1}^f & \mathbf{X}_{T-d,t-q+1}^f \\ \mathbf{X}_{T,t-q}^f & \cdots & \mathbf{X}_{T-d+1,t-q}^f & \mathbf{X}_{T-d,t-q}^f \end{bmatrix}$$

其中

$$\mathbf{X}_{T-i,t-j}^f = \begin{bmatrix} \mathbf{x}_{T-i,t-j}^{1,1} & \mathbf{x}_{T-i,t-j}^{1,2} & \cdots \\ \vdots & \vdots & \vdots \\ \mathbf{x}_{T-i,t-j}^{a,1} & \mathbf{x}_{T-i,t-j}^{a,2} & \cdots \\ \vdots & \vdots & \vdots \\ \mathbf{x}_{T-i,t-j}^{M-1,1} & \mathbf{x}_{T-i,t-j}^{M-1,2} & \cdots \\ \vdots & \vdots & \vdots \\ \mathbf{x}_{T-i,t-j}^{M,1} & \mathbf{x}_{T-i,t-j}^{M,2} & \cdots \end{bmatrix} \rightarrow \begin{bmatrix} \cdots & \mathbf{x}_{T-i,t-j}^{1,b} & \cdots & \mathbf{x}_{T-i,t-j}^{1,N} \\ \vdots & \vdots & \vdots & \vdots \\ \cdots & \mathbf{x}_{T-i,t-j}^{a,b} & \cdots & \mathbf{x}_{T-i,t-j}^{a,N} \\ \vdots & \vdots & \vdots & \vdots \\ \cdots & \mathbf{x}_{T-i,t-j}^{M-1,b} & \cdots & \mathbf{x}_{T-i,t-j}^{M-1,N} \\ \cdots & \mathbf{x}_{T-i,t-j}^{M,b} & \cdots & \mathbf{x}_{T-i,t-j}^{M,N} \end{bmatrix}$$

$$i = 0, 1, \dots, d, j = 1, 2, \dots, q.$$

1.2 时段聚类标签产生

城市轨道交通作为公共出行方式之一,其客流量变化与社会群体的日常出行规律相关.当前研究主要以“日”为单位考虑各日之间城轨客流时间分布的相似性与差异性,忽略了当日非相邻时段之间的客流特征^[19].而对于城市轨道交通短时客流,一日内跨时段的客流分布也存在相似性,如图3所示,时段07:00~09:30与时段18:00~20:30的客流、时段10:00~11:30与时

段14:00~15:30的客流以及12:00~15:30内多个时段的客流都分别呈现相似的变化趋势.

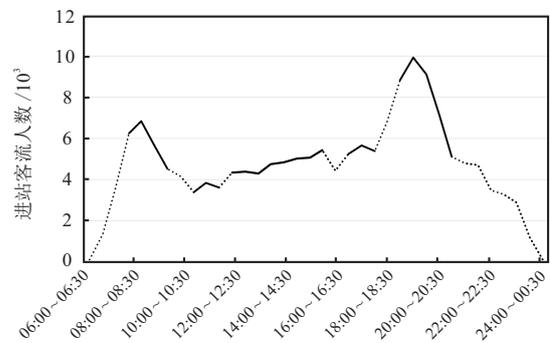


图3 跨时段的客流分布

为更细粒度地刻画城轨交通跨时段的短时客流时间分布,本文设计了自适应k-means聚类法对具有相似客流变化规律的时段进行聚类,属于同一类别的短时客流具有相同的时段聚类标签.该标签将被输入至主模型,并与其他客流特征一起作为回归任务的历史客流依据.传统k-means因收敛速度快、聚类结果优且需调参数少而被广泛应用于数据分析,但该方法需事先指定聚类数k,并且聚类结果的优劣难以获得有效反馈.在本文所提出的自适应k-means聚类法中,主模型开始训练前,时段聚类标签将依据随机选取的聚类数k生成,并在每次主模型训练结束后,依据模型训练误差的反馈进行自适应调整,直至模型性能达到要求.

自适应k-means聚类法分为两步:1)计算每个样本点与k个聚类中心的距离,将样本点逐个归至与其距离最小的聚类簇下,并在所有样本点都归类后更新聚类中心,不断重复以上过程直至两次聚类中心的差值符合终止要求^[20];2)将得到的聚类结果输入主模型中,并依据模型训练结果对聚类结果进行适应性调整.假设预测第T天时刻t的城轨客流,每日总时段为

$$\left\lceil \frac{(\text{每日记录起始时间} - \text{每日记录终止时间})(h)}{\text{单位统计时长}(h)} \right\rceil,$$

其中 $\lceil \cdot \rceil$ 表示向上取整,可选聚类数 $K = (k_1, k_2, \dots, k_l)$.

自适应 k -means时段聚类法步骤如下.

步骤1):

输入:随机选取聚类数 $k \in K$,第 $T - i$ 天的时段客流 $A_{T-i} = (a_{T-i}^1, a_{T-i}^2, \dots, a_{T-i}^r)$,任意 $i \in [1, d]$,迭代终止,阈值 σ_1 ,迭代次数 E_1 ;

输出:时段聚类标签 $L = (L^1, L^2, \dots, L^k)$.

聚类簇 $\delta = (\delta_1, \delta_2, \dots, \delta_k)$,初始化聚类中心

$C = (\text{center}_{\delta_1}, \text{center}_{\delta_2}, \dots, \text{center}_{\delta_k}), \Delta = +\infty$

for $e_1 = 1, 2, \dots, E_1$

for $a_{T-i}^l \in A_{T-i}$

计算 $\delta_n = \underset{\delta_j \in \delta}{\operatorname{argmin}} \operatorname{dist}(a_{T-i}^l, \text{center}_{\delta_j})$

将 a_{T-i}^l 划分至聚类簇 δ_n 中

end for

更新 $\text{center}_{\delta_j} = \frac{1}{|\delta_j|} \sum_{a_{T-i}^l \in \delta_j} a_{T-i}^l$ ($|\cdot|$ 表示 \cdot 中

元素个数), $j = 1, 2, \dots, k$

$$\Delta = \sum_{j=1}^k \sum_{a_{T-i}^l \in \delta_j} \|a_{T-i}^l - \text{center}'_{\delta_j}\|^2$$

if $\Delta < \sigma_1$

then 输出各时段对应的聚类标签 L^j 至网络

break

end if

end for

步骤2):

输入:时段聚类标签 $L^j (j = 1, 2, \dots, k)$,迭代终止,阈值 σ_2 ,训练次数 E_2 ;

输出:城轨短时客流预测模型.

for $e_2 = 1, 2, \dots, E_2$

时段聚类标签加入主模型训练

计算模型训练误差 ε

end for

if $\varepsilon < \sigma_2$

then 输出短时客流预测模型

break

else

更新聚类数 $k' = \{x | \forall x \in K \text{ 且 } x \neq k\}$

重复步骤1)、步骤2)

end if

1.3 主要网络层

1.3.1 卷积长短时记忆网络层(ConvLSTM layer)

ConvLSTM网络层是一种复合的网络结构,其关键组成部分是分布在时间轴上的卷积结构和贯穿记忆体的信息流 $C_{T,t-j} (j = 1, 2, \dots, q)$.卷积结构提取客流的空间分布特征;而 $C_{T,t-j}$ 作为综合性的信息流,在不断更新和累积的同时向下传递,能在跨度较长的信息处理过程中保持客流时间特征的有效性.ConvLSTM对客流时空信息的处理与传递过程如图4所示.

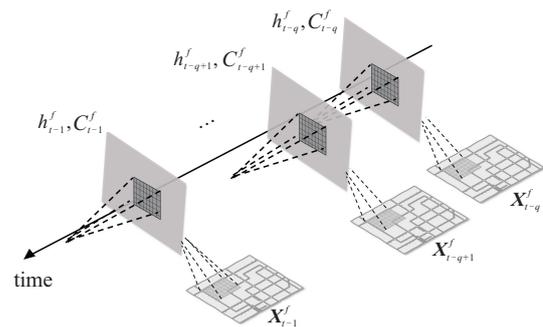


图4 ConvLSTM对客流时空信息的处理与传递

假设站点 f 第 T 日时刻 t 的客流 $p_{T,t}^f \in R$,考虑与时刻 t 紧邻的前 q 个统计时段的客流影响,将前 q 个统计时段的客流时空数据 $\mathbf{X}_{T,t-j}^f \in R^{M \times N} (j = 1, 2, \dots, q)$ 作为主网络模型的输入.为有效提取各时段的时空分布特征,指定大小为 \bar{h} 的卷积核 $\mathbf{W} \in R^{\bar{h} \times \bar{h}} (\bar{h} \ll M, \bar{h} \ll N)$,卷积核移动步长 α ,以及“时空图像”边缘的填充大小 β ,对 $\mathbf{X}_{T,t}^f$ 进行卷积运算.以二维卷积为例,计算过程^[21]如下:

$$y_{lz}^f = b_{T,t-j}^f + \sum_{u=1}^m \sum_{v=1}^n w_{uv} \cdot x_{l+u-1, z+v-1},$$

$$y_{lz} \in (\mathbf{X}_{T,t-j}^f)', w \in W, l \in (0, M'), z \in (0, N'); \quad (1)$$

$$M' = \left\lfloor \frac{M - \bar{h} + 2 \times \alpha}{\beta} + 1 \right\rfloor, \quad (2)$$

$$N' = \left\lfloor \frac{M - \bar{h} + 2 \times \alpha}{\beta} + 1 \right\rfloor. \quad (3)$$

其中: $\lfloor \cdot \rfloor$ 表示计算结果向下取整, $b_{T,t-j}^f$ 与 $(\mathbf{X}_{T,t-j}^f)'$ 分别对应偏置量和卷积后的客流时空特征图.同时,为降低网络复杂度,并提取主要特征,引入池化操作对卷积后的时空特征图进行压缩.

考虑各时段之间的时序依赖关系,客流时空信息将由3个交互的功能性细胞结构——遗忘门、输入门、输出门^[22]进一步筛选和更新.其中,依据上一时段的状态信息 $\mathbf{h}_{T,t-j-1}^f$ 、细胞综合状态 $\mathbf{C}_{T,t-j-1}^f$ 和当前输入 $(\mathbf{X}_{T,t-j}^f)'$ ，“遗忘门”和“输入门”将分

别产生 $C_{T,t-j-1}^f$ 被保留的比例 $\theta_{T,t-j}^f$ 以及候选状态 $\hat{C}_{T,t-j}^f$ 被选择的比例 $S_{T,t-j}^f$, 由此实现无用信息的删除和有用信息的更新。“输出门”则将根据 $h_{T,t-j-1}^f$ 和 $(X_{T,t-j}^f)'$ 来判断当前客流的状态特征 $O_{T,t-j}^f$, 并以此产生当下的状态信息 $h_{T,t-j}^f$. 依据客流相关时段数, 这个过程将重复多次, 并最终输出提取到的客流时空规律 h_T^f .

以此类推, 前 d 天同时段客流的时空规律 h_{T-i}^f ($i = 1, 2, \dots, d$) 将分别由结构相同的 ConvLSTM 模块处理得到. ConvLSTM 网络层的具体计算过程^[16] 如下:

遗忘门

$$\theta_{T,t-j}^f = \sigma(W_\theta * [h_{T,t-j-1}^f, (X_{T,t-j}^f)'] + W_\theta \circ C_{T,t-j-1}^f + b_\theta); \quad (4)$$

输入门

$$S_{T,t-j}^f = \sigma(W_s * [h_{T,t-j-1}^f, (X_{T,t-j}^f)'] + W_s \circ C_{T,t-j-1}^f + b_s); \quad (5)$$

过渡状态

$$\hat{C}_{T,t-j}^f = g(W_c * [h_{T,t-j-1}^f, (X_{T,t-j}^f)'] + b_c); \quad (6)$$

综合状态

$$C_{T,t-j}^f = \theta_{T,t-j}^f \circ C_{T,t-j-1}^f + S_{T,t-j}^f \circ \hat{C}_{T,t-j}^f; \quad (7)$$

输出门

$$O_{T,t-j}^f = \sigma(W_o * [h_{T,t-j-1}^f, (X_{T,t-j}^f)'] + W_o \circ C_{T,t-j}^f + b_o); \quad (8)$$

当前状态

$$h_{T,t-j}^f = O_{T,t-j}^f \circ g(C_{T,t-j}^f), \quad j = 1, 2, \dots, q. \quad (9)$$

其中: * 表示卷积操作; \circ 表示 Hadamard 乘积; W 为卷积核的权重系数矩阵 (如 W_θ 表示“遗忘门”的权重系数矩阵); b 为偏置向量 (如 b_θ 表示“遗忘门”的偏置向量); $\sigma(\cdot)$ 、 $g(\cdot)$ 表示神经网络的激活函数 sigmoid 和 tanh; C 、 θ 、 S 则分别对应客流信息在神经细胞内的存储状态、遗忘门输出的遗忘因子和输入门输出的状态更新因子.

1.3.2 卷积层(convolutional layer)

一定时间范围内, 考虑前 d 天客流分布的周期性, 客流时空规律 ($h_T^f, h_{T-1}^f, \dots, h_{T-d}^f$) 将分别由对应的 ConvLSTM 网络块提取得到. 为保证模型能学习到普适性的规律并在实际预测中展现相当的泛化能力, 使用卷积层对时空规律 ($h_T^f, h_{T-1}^f, \dots, h_{T-d}^f$) 中相对独立的特征进行过滤和降维.

1.3.3 全连接层(fully connected layer)

考虑细粒度时间因素——同日跨时段的客流分布相似性, 以降低客流随机性对预测结果的影响, 使用全连接层对 k -means 聚类结果 L^j ($j = 1, 2, \dots, k$) 和过滤后的时空特征信息进行完整表达并充分融合. 为实现融合信息与客流值的映射关系构建, 通过堆叠多层全连接层对信息的空间维度进行转换和整合信息, 提升模型的学习能力. 最终, 输出时段 t 时的客流预测值 $P_{T,t}^f$.

k -ConvLSTM 模型的训练过程包含信息的前向传递与误差的反向反馈. 在输出层输出 $P_{T,t}^f$ 后, 通过对比 $P_{T,t}^f$ 与样本实际客流值 $x_{T,t}^f$, 误差 ε 将被逐层回传至各网络层及 k -means 聚类模块, 指导网络层的权重调整, 并调动 k -means 聚类模块重新调整聚类簇数.

1.3.4 其他网络层

为避免在反向传播过程中, 网络整体训练速度因每批次数据的差异或低层网络梯度的消失而变慢, 在网络层间穿插批规范化层 (batchnorm layer) 对数据进行规范化. 此外, 使用融合层 (concatenate layer) 对第 $T-i$ ($i = 0, 1, \dots, d$) 各天的时空特征进行聚合.

2 实验与分析

2.1 实验数据

本文以深圳北站地铁站为预测目标站点, 其位于深圳铁路“四主四辅”客运格局的核心位置, 是深圳市接驳功能最为齐全的特大型综合交通枢纽, 同时也是地铁4号线和地铁5号线换乘站, 日常客流极大且非线性特征明显, 准确预测深圳北站短时客流具有较强的现实意义. 选取深圳地铁真实IC卡历史记录作为 k -ConvLSTM 城轨短时客流预测模型的数据来源. 因各站点运营时间范围存在差异, 为保证数据的完整, 指定每日记录时段为6:30~23:30, 并进行异常值处理. 过短或过长的客流统计时长缺乏实际意义, 不利于站点的实际运营安排和紧急预警布置, 因此, 本文选择15 min的时间间隔作为城轨短时客流统计时长, 每日统计时段为68.

根据相关研究, 63%以上的深圳地铁IC卡用户搭乘地铁的时长在30 min以内^[23], 故可认为路程在30 min以内的站点之间客流流动频繁, 具有较强的相关性. 因此, 选择从深圳北站出发, 30 min内可达的59个站点作为深圳北站的时空关联站点. 根据站点在地铁网络图中的相对位置构建 15×17 的地理网格, 作为输入主模型的客流时空分布“图像”.

为更直观地分析深圳北站的客流特征, 随机选取调查时间内一周的进站客流, 并与中间站——民治站进行对比, 对比结果如图5、图6所示.

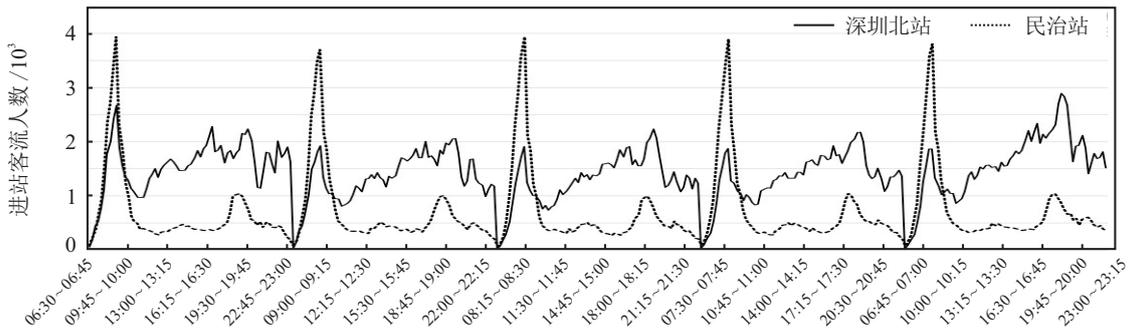


图5 工作日进站客流对比

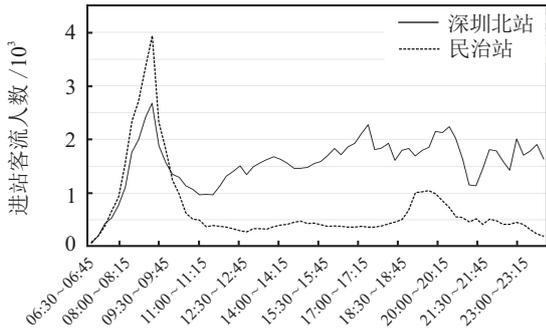


图6 周末进站客流对比

由图5、图6可以看到,深圳北站工作日各天进站客流曲线的整体分布相似,具有周期性特征,但因客流量巨大,随机性强于一般站点.其中:工作日客流主要集中在中午11:00后的时段,早高峰出现在8:00~10:00之间,晚高峰为17:00~21:00之间;周末各时段客流量分布相对均匀,但单日客流总量高于工作日,并且没有明显早、晚高峰.另外,深圳北站因同时承担枢纽站和换乘站的角色,即便是在工作日的非通勤时段,进站客流也十分可观,也正因如此,深圳北站工作日潮汐现象较民治站并不明显.相比之下,民治站工作日的进站客流峰值为3706人/15min,非工作日平均进站客流远小于深圳北站,且客流曲线随时间的波动频率也相对较低.

为进一步确定工作日各日及非工作各日之间相关性,依据特征指标——最大值、最小值、均值、标准

表1 工作日客流相关性统计

时间	指标	周一	周二	周三	周四	周五
周一	相关系数	1.00	0.90	0.91	0.91	0.86
	显著性(双尾)	-	0	0	0	0
周二	相关系数	0.90	1.00	0.96	0.96	0.92
	显著性(双尾)	0	-	0	0	0
周三	相关系数	0.91	0.96	1.00	0.98	0.95
	显著性(双尾)	0	0	-	0	0
周四	相关系数	0.91	0.96	0.98	1.00	0.96
	显著性(双尾)	0	0	0	-	0
周五	相关系数	0.86	0.92	0.95	0.96	1.00
	显著性(双尾)	0	0	0	0	-

表2 周末客流相关性统计

时间	指标	周六	周日
周六	相关系数	1.00	0.89
	显著性(双尾)	-	0
周日	相关系数	0.89	1.00
	显著性(双尾)	0	-

差、偏度等对工作日之间与非工作日之间的进站客流进行皮尔森相关性分析,结果如表1、表2所示.

为保证模型的预测精度,视皮尔森相关系数在0.85以上为相关性显著,本文将数据集分为工作日与非工作日两个部分,分别进行预测.

2.2 关键参数选择

2.2.1 相邻时段数选择

当前时段的城轨短时客流量会受到之前 q 个相邻时段客流量的影响,为确定合理的 q 值,综合考虑每日统计时段总数与日常出行规律,分别选择 $q = 2, 3, 4, 5$ 进行测试,此时ConvLSTM卷积核大小为 5×5 .以深圳北站历史城轨短时客流量作为训练数据,设定最大训练次数为1000,训练损失连续30次无法再改进则训练停止.模型输入为预测时间前一周同天的前 q 个时段的时空矩阵、前一天前 q 个时段的时空矩阵以及当天前 q 个时段的时空矩阵,输出为当前时刻客流预测值, q 取不同值时的模型平均相对误差如表3所示.

表3 q 取不同值的平均相对误差

时间	时段数 q			
	2	3	4	5
周一	97.83	83.70	118.28	109.51
周二	87.04	76.51	106.73	96.95
周三	86.69	81.06	100.12	98.37
周四	87.97	77.14	105.83	95.44
周五	96.89	87.74	112.16	115.57
平均	91.28	81.23	108.62	103.17
周末	49.18	67.38	89.25	81.76

由上可知:当 $q = 3$ 时的平均训练损失和相对误差最小,说明紧邻的前3时段与当前时段的客流量联

系紧密;而在周末,出行人流随时间呈近似正态分布的趋势,波动相对平稳,可见,不同 q 值所对应结果的差别并不大,其中 $q = 2$ 时相对误差最小.因此,在实例研究中,分别设定工作日 $q = 3$,周末 $q = 2$.

2.2.2 卷积核大小选择

ConvLSTM网络层通过卷积核对客流数据的子区域进行特征局部提取和信息抽象化,其中“局部”的范围大小取决于卷积核.一般来说,卷积核越大,所对应区域内提取到的信息越多,最终获得的全局特征越好,但另一方面,卷积核过大会导致计算量的暴增,不利于模型深度的增加,模型计算性能也会降低.因此,基于先前确定的相邻时段数,并综合考虑本文所构造的时空“图像”尺寸及相关研究经验,分别选取大小为 2×2 、 3×3 、 5×5 、 7×7 的卷积核进行测试,模型预测误差如表4所示.

表4 不同大小卷积核的预测误差

时间	卷积核大小			
	2×2	3×3	5×5	7×7
周一	113.34	82.05	88.18	94.68
周二	104.26	83.38	75.32	87.78
周三	106.63	88.31	72.06	95.49
周四	110.26	89.17	75.14	90.05
周五	97.66	92.35	84.74	118.68
平均	106.43	87.05	79.09	97.34
周末	85.31	76.48	52.88	81.55

由表4可以看到,卷积核大小对预测精度有明显的影响.由于城轨短时客流数据存在稀疏值,卷积核大小为 2×2 时,提取局部特征的范围过小,特征得不到有效的表征,不论是工作日还是周末,实际模型的预测效果都不够理想.相比之下,卷积核大小为 5×5 时,城轨短时客流在工作日与周末的时空特征都得到了较为丰富地表达,模型实际预测精度较高,因此,本文实例选取的卷积核大小为 5×5 .

此外,结合相关研究经验以及预测时段与前一天、上周同时段的周期性关系^[24],本文选取 $d = 1, 7$.

2.3 模型性能分析

本文将数据集分为工作日和周末两个部分.当预测第 T 日时刻 t 的城轨短时客流时,网络层输入数据为第 T 日、第 $T - 1$ 日以及第 $T - 7$ 日的前 q 个时段客流量,其中工作日与周末模型中 q 分别为3和2.进行 k -means聚类时,考虑到数据的多样性及周期性,选择第 $T - 14$ 日的历史客流 $\mathbf{X}_k = (\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_r)$ 作为聚类样本(r 为客流聚类样本点数).为避免聚类结果过于稀疏,以每小时作为一个聚类样本点,该小时内相邻4个时段的短时客流量为描述该样本点的

4个属性,同一样本点中4个时段的聚类标签相同.例如当自适应 k -means聚类簇数为 k 时,样本点 $\mathbf{X}_1 = (x_1^1, x_1^2, x_1^3, x_1^4)$ 与 $\mathbf{X}_2 = (x_2^1, x_2^2, x_2^3, x_2^4)$ 分别被划分至聚类簇 δ_a, δ_b 中,则 \mathbf{X}_1 与 \mathbf{X}_2 中各时段对应的聚类标签分别为 $L^{\delta_a}, L^{\delta_b}$.因此,预测模型输入为 $\mathbf{X} = [\text{dataset}_{T,t}^f, \mathbf{X}_k]$.

本文使用数据为深圳市地铁2019年4月~9月各站点的短时客流数据,数据总量为 $(30 \times 3 + 31 \times 3) \times 68 = 12444$ 组,为避免特殊节假日客流剧烈波动对模型训练结果的影响,剔除其中4天特殊节假日的客流数据,则最终使用数据为12172组.其中:每组数据共包含4个输入数据与1个输出数据,主网络模型输入数据为同天当前预测时段的前3个时段(工作日)或前2个时段(非工作日)的客流量以及前第1天和前第7天的同时段客流量,聚类模型输入为前第14天的客流量,输出则为当前时段的预测客流量.由此,在输入数据与输出数据之间建立起一一对应的关系.为提高模型的稳定性,避免模型在训练过程中出现过学现象,随机打乱各输入-输出数据对的顺序,以保证训练好的模型在实际应用中具有相当的泛化能力,同时,为提升预测精度和模型训练速度,将输入数据进行归一化处理.在数据集中,划分2019年7月1日~2019年7月6日的数据为测试集,其余为训练集.

综合考虑数据输入维度、问题类型以及先前的参数选择试验,设定第 $T - i$ ($i = 0, 1, 7$)日的时空特征提取模块由3层ConvLSTM网络和1层批规范化层构成,每层ConvLSTM网络卷积核大小为 5×5 ,个数为30;聚合后的时空信息与聚类信息将由3层全连接层进一步处理以输出预测值,全连接层节点个数为(256,192,144);网络激活函数为“relu”函数,并采用adam算法对网络进行优化.使用python语言对模型进行编程建模,借助Google Colaboratory平台执行代码以训练模型.

同时,为验证本文所提出模型的有效性,分别选取另外6种客流预测模型进行对比:1)深度学习模型,包括考虑时间依赖关系的长短时记忆网络(LSTM)和双向长短时记忆网络(Bi-LSTM)、考虑时空特征的并行CNN与LSTM混合模型、内嵌式ConvLSTM混合模型;2)浅层机器学习模型,包括支持向量回归模型(support vector regression, SVR)和BP神经网络.预测结果如图7所示.

由图7可知,整体上,各模型都能在不同程度上对未来客流进行有效预测.其中,因周末客流实际变动规律相对平稳,各时段的客流差值相对工作日较

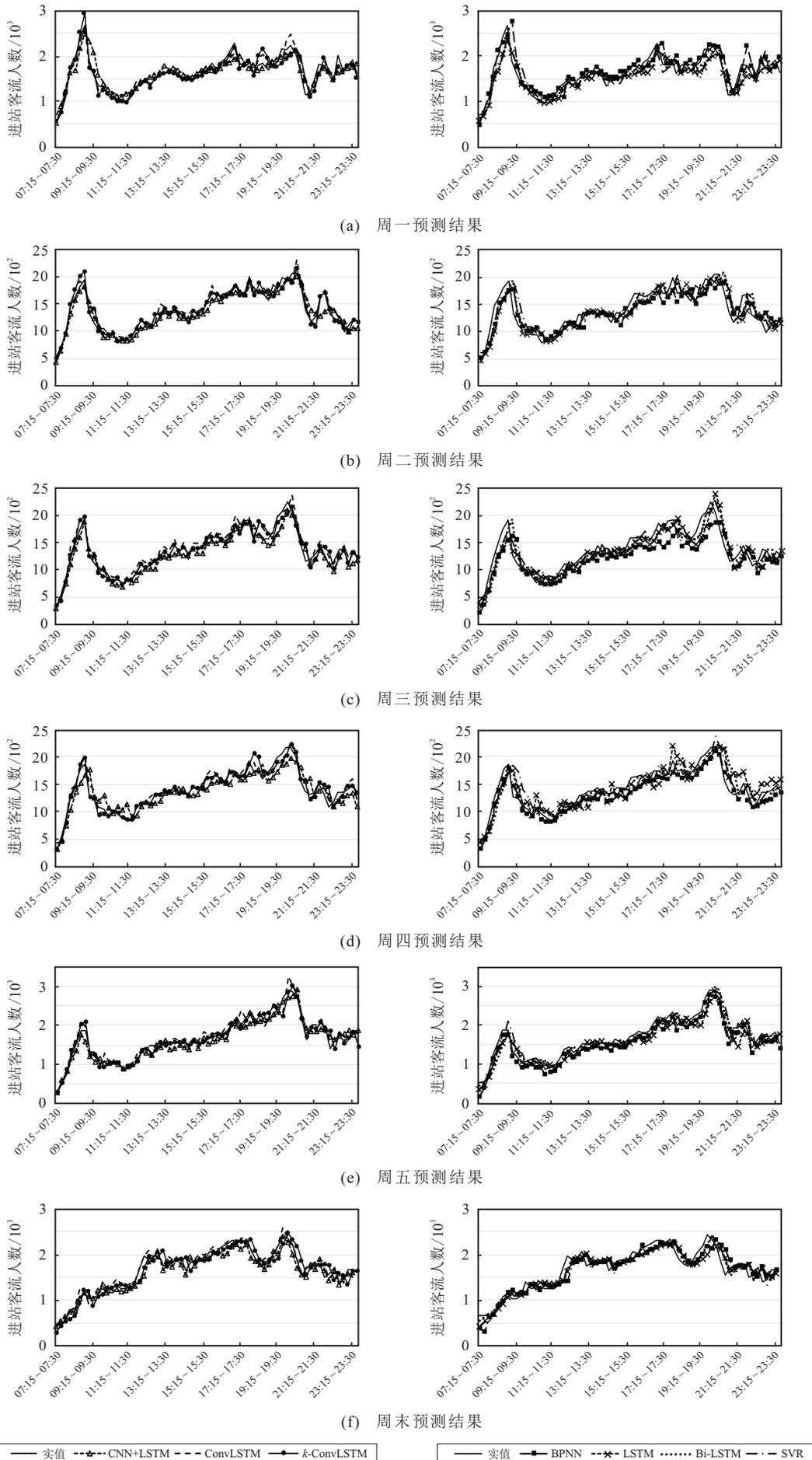


图7 深圳北站一周客流预测结果

小,直观上看,各模型周末预测结果均优于工作日预测结果.此外,仅考虑时间因素的4种模型(LSTM、Bi-LSTM、BPNN和SVR)客流预测结果相对较差;而同时考虑时空特征的并行CNN与LSTM混合模型、内嵌式ConvLSTM混合模型以及本文所提出的 k -ConvLSTM模型因对客流内在规律的把握更为全面,预测客流量曲线更贴近实际客流曲线,表现较为优秀,尤其是本文所提出的 k -ConvLSTM模型预测结果最为理想,预测客流量曲线与实际客流量曲线趋势一致,拟合效果最好.

为了更精确地给出预测精度,本文分别选取均方根误差(RMSE)、平均绝对误差(MAE)、平均绝对百分比误差(MAPE)作为评价指标,同时,鉴于深圳北站是典型的枢纽站和换乘站,工作日早晚潮汐现象不明显,为进一步验证所提出模型的有效性,选取潮汐现象较为明显的中间站——民治站进行实验.对所有模型的预测性能进行测度,结果如表5所示.

表5 模型评价指标

预测模型	时间	深圳北站			民治站		
		RMSE	MAE	MAPE/%	RMSE	MAE	MAPE/%
k -ConvLstm	工作日	136.66	86.29	5.30	59.67	45.77	8.61
	周末	81.51	52.88	3.26	58.39	47.52	7.48
ConvLstm	工作日	144.49	107.88	7.27	67.35	51.42	8.94
	周末	95.26	65.68	4.91	61.66	50.71	7.73
CNN+LSTM	工作日	168.90	137.76	8.47	73.13	61.62	9.82
	周末	124.88	95.46	5.96	63.48	54.32	8.98
SVR	工作日	187.63	150.39	9.34	135.89	109.63	14.34
	周末	154.13	126.48	8.01	104.61	90.57	11.89
BPNN	工作日	186.13	151.43	8.83	135.06	105.77	13.19
	周末	176.07	136.49	8.82	99.34	84.12	11.02
LSTM	工作日	182.53	144.94	8.56	109.13	76.85	10.67
	周末	159.22	123.51	7.50	94.17	69.03	9.70
Bi-LSTM	工作日	178.99	140.73	8.33	105.96	71.58	10.20
	周末	150.97	120.29	7.03	90.01	57.87	8.98

由表5可知,本文所设计的 k -ConvLSTM城轨短时客流预测模型的3个评价指标明显优于其他6种模型,说明对进站客流时间特征的细化、以及对时空特征较为全面的提取及融合,能有效改善客流预测精度.

在对比模型中:ConvLSTM模型由于采取了内嵌结构对时空特征进行提取,预测结果优于并行的CNN+LSTM模型,但由于这两个模型均未考虑跨时段的断面客流特征,对特征的细化及融合有所欠缺,较本文提出的 k -ConvLSTM,其客流预测值的

平均绝对误差较大,对单个时段客流的预测精度较低;LSTM与Bi-LSTM分别对数据时间依赖关系进行单向和双向的提取,Bi-LSTM结果优于LSTM,但由于两者均未考虑客流空间特征,导致预测结果的误差都较大;BPNN模型逐层传递客流历史信息以训练网络,对客流特征的识别无针对性且存在梯度消失的情况,因而对客流的拟合度并不高;SVR数据输入类型和特征的提取在一定程度上依赖于手工,对复杂数据内部特征的自动识别能力较弱,且对参数的选择极为敏感,因此,针对随机性较强的短时客流数据,这两类浅层预测模型的结果均不理想.

3 结 论

针对城轨交通短时客流预测,现有研究多关注于客流数据的历史浅层时间特征,忽略了深层时间信息和空间特征.因此,为进一步提高客流的预测精度,本文提出了能同时考虑多层次时间特征和空间特征的方法——基于自适应 k -means和ConvLSTM网络的深度学习方法.该方法使用嵌入式ConvLSTM网络对时空特征进行提取,并通过自适应 k -means时段聚类实现对跨时段客流特征的细化,保证了客流特征的充分融合以及内在规律的全面表达.为使客流特征得到有效提取,本文通过实验对工作日和周末的客流历史关键参数——紧邻时段数 q 进行选择.结果显示,工作日紧邻前3个时段的客流对当前预测时段的客流量影响最大,而周末客流则主要受紧邻前2个时段客流的影响.此外,为保证模型基本性能的稳定,考虑数据规模及模型深度,对卷积核大小进行了确定.实验表明,大小为 5×5 的卷积核对目标站点及关联站点的客流空间分布特征的提取较为全面,提取到的空间信息能帮助后续回归任务的学习及最终模型精度的提升.

为验证本文所提出模型的有效性,同时设计了CNN+LSTM、ConvLSTM两种考虑时空因素的深层模型,以及仅考虑时间因素的LSTM、Bi-LSTM两种深层模型及SVR、BPNN两种浅层模型.研究结果表明,相比于同时考虑时空因素的模型,仅考虑历史客流时间因素的模型受随机性影响较强,预测结果不太理想.而在考虑时空特征的模型中,本文所设计的 k -ConvLSTM模型预测精度最佳,评价指标RMSE、MAE、MAPE均优于其他模型,从而验证了该模型在城轨短时客流预测上的有效性.

下一步,可以深入探究不同类型地铁站点(如中间站、换乘站、终点站等)的具体客流特征,并建立相适应的模型进行预测,提升模型的泛化能力.

参考文献(References)

- [1] 徐成永,叶轩,宣晶.轨道交通运行效果评估、客流特征分析和对策研究[J].都市轨道交通,2019,32(6):44-50.
(Xu C Y, Ye X, Xuan J. Operational evaluation, characteristic analysis and development strategy of urban rail transit in Beijing[J]. Urban Rapid Rail Transit, 2019, 32(6): 44-50.)
- [2] 周慧娟,李蓓,刘小明.过饱和轨道交通车站客流动态协调控制研究现状与展望[J].自动化博览,2019(12):48-53.
(Zhou H J, Li B, Liu X M. Passenger flow dynamic and coordinated control on oversaturation urban rail station: Current status and perspective[J]. Automation Panorama, 2019(12): 48-53.)
- [3] Jiao P P, Li R M, Ibrahim A, et al. Three revised Kalman filtering models for short-term rail transit passenger flow prediction[J]. Mathematical Problems in Engineering: Theory, Methods and Applications, 2016(3): 1-10.
- [4] Wang Y G, Ma J F, Zhang J. Metro passenger flow forecast with a novel Markov-Grey model[J]. Periodica Polytechnica Transportation Engineering, 2019, 48(1): 70-75.
- [5] Özuysal M, Tayfur G, Tanyel S. Passenger flows estimation of light rail transit (LRT) system in Izmir, Turkey using multiple regression and ANN methods[J]. Promet-Traffic & Transportation, 2012, 24(1): 1-14.
- [6] Liu S, Yao E. Holiday passenger flow forecasting based on the modified least-square support vector machine for the metro system[J]. Journal of Transportation Engineering, 2017, 143(2): 04016005.1-04016005.8.
- [7] Chen Q, Wen D, Li X, et al. Empirical mode decomposition based long short-term memory neural network forecasting model for the short-term metro passenger flow[J]. PloS One, 2019, 14(9): e0222365.
- [8] 崔洪涛,陈晓旭,杨超,等.基于深度长短期记忆网络的地铁进站客流预测[J].城市轨道交通研究,2019,22(9):41-45.
(Cui H T, Chen X X, Yang C, et al. Forecast of subway inbound passenger flow based on DLSTM recurrent network[J]. Urban Mass Transit, 2019, 22(9): 41-45.)
- [9] 龙小强,李捷,陈彦如.基于深度学习的城市轨道交通短时客流量预测[J].控制与决策,2019,34(8):1589-1600.
(Long X Q, Li J, Chen Y R. Metro short-term traffic flow prediction with deep learning[J]. Control and Decision, 2019, 34(8): 1589-1600.)
- [10] Jia Y, He P, Liu S, et al. A combined forecasting model for passenger flow based on GM and ARMA[J]. International Journal of Hybrid Information Technology, 2016, 9(2): 215-226.
- [11] Pan J, Ma C Q. Passenger flow forecast based on improved grey Markov model[J]. Technology & Economy in Areas of Communications, 2018, 20(6): 52-56.
- [12] Fu C L, Huang M, Sha Z R. Short-term forecast of passenger flow into an urban rail transit station based on EEMD-BP[J]. Railway Transport and Economy, 2020, 42(3): 105-111.
- [13] Ma Y L, Du X H, Li M Z. Forecast of subway passenger flow in different periods based on PCA-RF method[J]. Chinese Railways, 2019(5): 61-64.
- [14] Ma X L, Zhang J Y, Du B W, et al. Parallel architecture of convolutional bi-directional LSTM neural networks for network-wide metro ridership prediction[J]. IEEE Transactions on Intelligent Transportation Systems, 2019, 20(6): 2278-2288.
- [15] Wei Y, Chen M C. Forecasting the short-term metro passenger flow with empirical mode decomposition and neural networks[J]. Transportation Research Part C: Emerging Technologies, 2012, 21(1): 148-162.
- [16] Xing J S, Chen Z, Wang H, et al. Convolutional LSTM network: A machine learning approach for precipitation nowcasting[C]. Advances in Neural Information Processing Systems. Montreal, 2015: 802-810.
- [17] Lin Y, Zhang J, Liu H. Deep learning based short-term air traffic flow prediction considering temporal-spatial correlation[J]. Aerospace Science and Technology, 2019, 93: 105113.
- [18] Liu Y P, Zheng H F, Feng X X, et al. Short-term traffic flow prediction with Conv-LSTM[C]. The 9th International Conference on Wireless Communications and Signal Processing (WCSP). Nanjing: IEEE, 2017: 1-6.
- [19] Liu X J. Analysis of urban rail transit passenger flow based on big data[J]. Technology Wind, 2019(21): 215.
- [20] Arora P, Varshney S. Analysis of k -means and k -medoids algorithm for big data[J]. Procedia Computer Science, 2016, 78: 507-512.
- [21] Krizhevsky A, Sutskever I, Hinton G E. Imagenet classification with deep convolutional neural networks[C]. Advances in Neural Information Processing Systems. Lake Tahoe, 2012: 1097-1105.
- [22] Sundermeyer M, Schlüter R, Ney H. LSTM neural networks for language modeling[C]. The 13th Annual Conference of the International Speech Communication Association. Portland, 2012: 194-197.
- [23] Yu Z J. Passenger flow characteristics analysis based on IC card data of Shenzhen metro system[J]. China Transportation Review, 2019, 41(4): 115-119.
- [24] Zhang W D, Chen F, Wang Z J, et al. Similarity measurement of metro travel rules based on multi-time granularities[J]. Journal of the China Railway Society, 2018, 40(4): 9-17.

作者简介

王秋雯(1995—),女,硕士生,从事机器学习的研究, E-mail: 173638358@qq.com;

陈彦如(1974—),女,教授,博士生导师,从事机器学习、物流与交通规划管理等研究, E-mail: chenyanru@swjtu.cn;

刘媛春(1975—),女,讲师,硕士,从事信息系统的研究, E-mail: 52997035@qq.com.

(责任编辑:李君玲)