

基于改进卷积神经网络的短时公交客流预测

陈深进 薛 洋

(华南理工大学电子与信息学院 广州 510641)

摘 要 针对城市公交客流存在随机性、时变性和不确定性的问题,文中提出了一种基于无监督特征学习理论和改进卷积神经网络的短时公交站点客流预测模型,以为市民提供实时、准确、有效的公交出行服务。运用无监督学习的方法对公交客流出行特征表达进行提取,利用大量已有数据集描述不同日期、不同时间段的短时客流的特征表达。为了防止和减少过拟合现象,运用改进卷积神经网络 DropSample 训练方法构造一个高效且高可信度的模型预测系统。在训练过程中,使用 Adam 算法的优化器对模型进行优化,更新网络模型参数,为自适应性学习率设置不同的参数。利用公交客流算法模型对广州实际公交站点的客流进行预测,实验结果表明:改进 CNN 网络模型的均方根误差为 229.539,平均绝对百分比误差为 0.117,相比于 CNN 网络模型、多元线性回归模型、卡尔曼滤波模型和 BP 神经网络模型,该模型的预测精度和可靠性更高。实例证明所提方法的预测误差更小,改进模型和算法具有实用性和可靠性。

关键词 公交客流,卷积神经网络,无监督学习,DropSample 训练方法,模型预测系统

中图分类号 TP181 文献标识码 A DOI 10.11896/j.issn.1002-137X.2019.05.027

Short-term Bus Passenger Flow Prediction Based on Improved Convolutional Neural Network

CHEN Shen-jin XUE Yang

(School of Electronic and Information Engineering, South China University of Technology, Guangzhou 510641, China)

Abstract Aiming at the random, time-varying and uncertain problems of urban public transport passenger flow, this paper proposed an unsupervised feature learning theory and an improved convolutional neural network based short-term bus station passenger flow prediction, which provides real-time, accurate and effective bus travel services for citizens. In order to prevent and reduce the occurrence of over-fitting, an efficient and reliable model prediction system based on DropSample training method of improved convolutional neural network is constructed. During the training process, the model can be used to describe the short-term passenger flow in different dates and different time periods. The optimizer of Adam algorithm is used to optimize the model, the network model parameters are updated, and different parameters for the adaptive learning rate are set. The results show that the root mean square error of the improved CNN network model is 229.539 and the average absolute percentage error is 0.117. Compared with CNN network model, multiple linear regression model, Kalman filter model and BP neural network model, this model is more accurate and reliable. The prediction error of the proposed method is smaller, and an example proves that the improved model and algorithm are practical and reliable.

Keywords Bus passenger flow, Convolutional neural network, Unsupervised learning, DropSample training method, Model prediction system

1 引言

随着城市经济的迅速发展,交通需求不断增大,大城市陷入了交通拥堵状态。为了提高道路的通行能力,减少交通事故,改善城市公共交通环境,须逐步建立和完善先进的公共交通系统(Advanced Public Transport System, APTS)。智能交通系统(Intelligent Traffic System, ITS)不仅需要合理的公交

线网规划,还要对公交客流进行精准的、实时的预测,适时对公交运营调度进行调整和优化,为市民提供实时、准确、有效的公交出行服务,实现公交资源的动态配置优化,节约市民的出行时间,有效缓解交通拥堵,从而达到绿色环保出行的目的。短时公交站点客流预测是调整公交发车频率、优化公交调度资源的重要依据。因此,研究公交站点短时客流预测具有重要的现实意义。

收稿日期:2019-01-18 返修日期:2019-03-17 本文受广东省应用型科技研发重大专项资金项目(2015B010131004)资助。

陈深进(1970-),男,博士生,主要研究领域为智能交通、机器学习、云计算;薛洋(1977-),女,博士,副教授,主要研究领域为人体行为模式识别、机器学习、人机交互信息处理,E-mail:yxue@scut.edu.cn(通信作者)。

短时公交站点客流预测主要用于城市公共交通的组织和控制,对实时性和精度要求较高。影响公交客流的因素很多,如站点位置、时间、交通路况、天气、国家政策等,部分客流处于动态不确定环境中,公交客流随着环境条件的变化而变化,精准的公交客流预测对城市智能公交调度起着决定性的作用。公交客流预测是指在当前时刻对下一决策时刻乃至以后若干时刻的客流做出实时预测。短时公交客流预测^[1]的时间小于 15 min(本文为 5 min)。目前,国内外公交客流预测的算法分为两种:第一种为线性估计模型,如自回归移动平均(Autoregressive Integrated Moving Average, ARIMA)模型^[2]、多元线性回归(Multivariable Linear Regression, MLR)模型^[3]、卡尔曼滤波(Kalman Filtering)模型^[4-5]等;第二种为机器学习的方法,如支持向量机^[6]、神经网络^[7]、深度学习^[8-9]等。线性估计模型存在如下缺点:1)目前,该模型解决不了非线性问题和客流存在的随机性、时变性和不确定性等问题;2)该模型对短时公交客流的预测不准确。深度学习中,基于无监督学习^[10-11]的卷积神经网络算法对具有长期依赖性的大规模时序数据具有良好的学习效果,在短时公交客流预测领域具有较好的应用效果^[12]。

本文提出了一种适用于广州的基于改进卷积神经网络^[13-14]的短时公交客流预测模型和算法,以分析站点、线路、交通小区的客流分布,挖掘公交客流的出行特征,为掌握城市公交客流状况提供技术支持;利用无监督学习方法获得不同日期、不同时间段的短时客流的有效高层语义表达特征,借助卷积神经网络的局部连接方式、特征提取不变性和可靠且高效的预测性能来预测短时公交客流^[15]的发展变化趋势。

2 研究方案及技术实现

2.1 基于无监督学习的公交客流出行特征的表达与提取

公交客流出行特征是对短时客流对象的描述和表征,有效的特征会在整个模型预测过程中发挥积极的作用。公交客流的出行特征表现为公交客流的时间分布特性、动态特性和短时特性。时间分布特性反映了工作日客流、周末客流和节假日客流的不同分布规律。动态特性反映了公交客流时间分布的不均匀性、方向的不均衡性以及客流分布的周期变动特性。短时特性反映了相邻时段、相邻天、相邻周和相邻月的客流变化的相关性。工作日客流与非工作日相比,人群出行规律、出行人群的特征均不同,客流高峰时段的客流密集程度也不同,这使得客流量数据在工作日和非工作日呈现出不同的特征。

鉴于无监督特征学习的强大表述能力和对大量数据样本的充分利用^[16],借助大量源数据集,包括公交 IC 卡刷卡数据、公交进出站数据、公交车视频检测数据样本、公交站点视频检测数据样本,利用无监督学习方法训练多个网络,形成具有深度网络结构的层叠自动编码器,以获得不同日期、不同时间段的短时客流的有效高层语义表达特征。该特征既具有不同时间分布特性,又具有周期变动性。

基于无监督特征学习方法学习到的特征表达的数据量非

常庞大,无法通过神经网络全连接方式来计算,须借助卷积神经网络的局部连接方式、特征提取不变性以及可靠且高效的预测性能来预测短时公交客流^[18]。

受深度置信网络^[17](Deep Belief Network, DBN)思想的启发, Bengio 等 2007 年提出了层叠自动编码器(Stacked Auto-Encoder, SAE)的深度学习结构。SAE 与 DBN 的主要区别在于组成深度结构的基本单元由受限玻尔兹曼机(Restricted Boltzmann Machine, RBM)变为自动编码器^[18-19](Auto-Encoder, AE),它是一种典型的无监督学习的双层神经网络结构,通过输出对输入进行重建使得隐藏层为输入层的压缩或稀疏表示,可以看成是对输入数据的特征学习。SAE 的构造过程与 DBN 的构造过程相似,如图 1 所示。SAE 每次都无监督地预训练一个 Auto-Encoder,然后将隐藏层作为下一个 Auto-Encoder 的输入层,最后将所有训练好的 Auto-Encoders 叠加,并外加一个输出层组成一个深度结构,利用反向传播算法(Back Propagation, BP)对深度网络结构进行训练。采用无监督学习方法对多个自动编码器进行学习,层叠各种自动编码器的编码部分用于短时客流的特征提取。

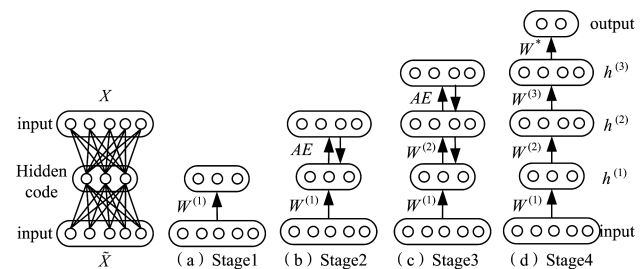


图 1 层叠自动编码器的形成过程

Fig. 1 Formation process of cascade automatic coding machine

2.2 卷积神经网络模型

卷积神经网络(Convolutional Neural Networks, CNN)是多层感知机(Multi-Layer Perceptron, MLP)的变种模型^[20],是近年迅速发展并引起全球高度重视的一种高效识别方法。CNN 是从生物学概念中演化过来的。从 2012 年开始,科学家们不断提出层次更深、性能更强的卷积神经网络模型: AlexNet^[21], VGGNet^[22] 和 ResNet^[23] 等。

CNN 结构由卷积层(Convolutional Layer)和聚合层(Pooling Layer)交替组成,每一层的输出为一组特征图(Feature Map)。第一层为特征提取层,前一层的局部接收域与每个神经元的输入相连,提取该局部的特征后,将确定与其他特征间的位置关系。第二层为特征映射层, CNN 网络计算层由 N 个特征映射组成,每个特征映射是一个平面,平面上所有神经元的权值相同。特征映射结构将 ReLu 函数作为 CNN 激活函数,使得特征映射具有位移不变性。神经网络的卷积运算是通过稀疏连接、参数共享、等变表示的重要特性来改进学习训练系统。

2.3 网络模型的改进

2.3.1 网络模型结构的改进与 dropout 技术

研究发现,运用多隐层的卷积神经网络^[24-27]可以提取到数据的本质特征,有利于对公交客流数据特征进行分类。

如果不断地增加网络卷积层的深度,将导致神经网络的性能低下。因此,对 CNN 网络结构的改进方法有两种:一种方法为适当增加卷积层,另一种方法为适当调整卷积核的大小^[28]。本文对 9 种不同 CNN 网络结构的性能进行了测试,结果如表 1 所列。

表 1 9 种不同 CNN 网络结构的性能对比

Table 1 Performance comparison of 9 CNNs with different structures

序号	卷积层数	卷积核	训练集准确率	测试集准确率
1	2	2×2	1.00	0.89
2	2	3×3	1.00	0.91
3	2	4×4	1.00	0.93
4	3	2×2	1.00	0.85
5	3	3×3	1.00	0.95
6	3	4×4	1.00	0.92
7	4	2×2	0.94	0.63
8	4	3×3	0.98	0.84
9	4	4×4	1.00	0.87

由表 1 可知,当卷积核大小相同时,3 层卷积层的 CNN 在测试集上的准确率最高;当卷积层数相同时,3×3 的卷积核 CNN 的性能最好。在确保 CNN 网络结构性能相差较小且 CNN 规模不大的前提下,将 CNN 网络结构改进为 3 层卷积层,且每层卷积核为 3×3。

通过比较网络在训练集和测试集上的准确率可知,网络结构的改进虽然可以提高 CNN 网络的性能,但同时也会产生过拟合的问题。研究发现,训练中出现过拟合问题是因为 CNN 全连接层的参数完全按照训练数据的特征识别结果更新,如果不对成熟的训练数据进行分类,则无法对测试数据进行准确的提取和分类。因此,本文引入 dropout 技术^[29],其中的参数设为 0.75,在深度学习网络的训练过程中,使参数以 0.75 的概率从网络中丢弃(不进行训练更新),提高了网络的泛化能力。

第二层采样层进行 2×2 最大池化时,加强了特征的稀疏表达。输出层改进为由 Softmax 函数计算结果,其表达式如式(1)所示:

$$p_j = \frac{e^{f_j}}{\sum_j e^{f_j}} \quad (1)$$

其中, y_i 为第 i 个输入特征的标签, f_j 表示输出层输出向量 f 的第 j 个元素, p_j 表示输入特征属于第 j 类的概率。

2.3.2 dropout 优化方案

为防止和减少过拟合,在全连接层后加入 dropout^[25] 的优化方案。dropout 是在每层神经元单元训练开始时随机删除一定比例的隐藏层神经元,以保证输入层和输出层的个数不变。根据 BP 神经网络学习算法对卷积神经网络中的参数进行更新学习,完成第一次迭代更新,第二次迭代中随机删除部分神经元,直到完成训练。

没有 dropout 时的神经网络的训练过程为:

$$z_i^{(t+1)} = w_i^{(t+1)} y^t + b_i^{(t+1)} \quad (2)$$

$$y_i^{(t+1)} = f(z_i^{(t+1)}) \quad (3)$$

为了达到 ensemble 的特性,使用 dropout 后神经网络的训练过程为:

$$r_j^t \sim \text{Bernoulli}(p) \quad (4)$$

$$\tilde{y}^{(t)} = r^{(t)} * y^t \quad (5)$$

$$z_i^{(t+1)} = w_i^{(t+1)} \tilde{y}^t + b_i^{(t+1)} \quad (6)$$

$$y_i^{(t+1)} = f(z_i^{(t+1)}) \quad (7)$$

其中, l 是隐藏层索引; z 表示 l 层的向量输入; y 是每一层的输出; y_0 是输入层; w 是权值参数; f 表示激活函数。预测时,每一个单元的参数要预乘以 p 。

改进 1 在模型训练过程中,不仅要求模型结构、权值收敛情况、分类的识别率等要合理,还要考虑训练过程所需的硬件计算能力和存储能力等。通过压缩减少卷积层的学习参数总量,将卷积核修改为 3×3,在全连接层后加入 dropout 层。

改进 2 卷积的作用是提取特征,在识别的过程中,输入特征提取得越精细,识别的精度就越高。CNN 网络通过不同的卷积层和池化层级交替堆叠生成通道,通过设计卷积核大小及不同池化实现所需要的效果。然而,在一系列操作过程中,通过一条通道、一种网络结构对特征图进行卷积,可能所提取的特征图不够充分。针对公交客流具有结构化数据和非结构化数据的特点,提出了结合两条不同网络通道对不同的特征进行卷积操作的方法,通过级联(Concat)操作将两条网络通道的输出结果进行结合。具体的操作为将降维后的两类特征数据并行分为两条支路运行:一条支路采用卷积层 3×3 的卷积核进行卷积操作;另一条支路在前一条支路的基础上进行残差操作。在下次特征图进行降维前,通过级联操作结合两条支路上的特征图,从而形成一个新的特征图。

残差网络容易被改进和优化,可通过适当增加网络深度来提高准确率。其核心是解决了增加 CNN 卷积网络深度产生的退化问题,通过增加网络深度来提高网络性能。残差学习结构主要是通过前向神经网络和 shortcut 连接实现,shortcut 连接相当于执行了同等的映射,如式(8)和式(9)所示:

$$y = F(x, W_i) + x \quad (8)$$

$$F = W_2 \sigma(W, x) \quad (9)$$

$F(x)$ 与 x 是逐元素相加的,其中 $F(x)$ 表示拟合残差函数。若两者的维度不同,则须针对 x 执行一个线性映射匹配维度,如式(10)所示:

$$y = F(x, W_i) + W_i x \quad (10)$$

通过实验发现,用来学习残差的网络层数应大于 1,否则残差网络退化为线性,卷积层 $layers=3$ 时网络的性能较稳定。

整个网络通过端到端的反向传播进行训练,用卷积层进行深度残差学习,主要原理是在初始卷积层结构的外部使用短接(shortcut)操作,以组成基本的残差模块。通过逐级累加残差模块,可以成功地缓解网络随深度增加而产生退化的问题,从而提升了网络的整体性能。

本文运用了双网络级联融合:在上次降维前,对支路一的 CNN 卷积神经网络和支路二含有残差模块的卷积神经网络进行级联操作,通过这种方式将两个不同的特征数据进行了整合,最后整体输出。

模块间的参数如式(11)所示:

$$P = I \times K \times O \quad (11)$$

其中, I 表示输入特征个数, O 表示输出特征个数, K 表示卷积核的面积大小。在单个模块中,双网络级联结构的参数相当于两条支路参数的总和。

2.3.3 模型设置

双网络级联结构是由多个模块组成的,支路一采用了CNN结构,在平均池化降维前采用 3×3 的卷积核进行卷积;支路二在CNN结构的基础上加入了残差操作。双网络级联结构层级的具体设置如表2所列,其中Conv1表示支路一,Conv2表示支路二。

表2 双网络级联层级参数的配置

Table 2 Parameter configuration of dual network cascade level

序号	层结构	卷积核大小/ 步长/pad	输出维度
1	input	—	$454 \times 454 \times 6$
2	Conv0	$11 \times 11 / 4 / 0$	$110 \times 110 \times 256$
3	Conv1(1-3层)	$3 \times 3 / 1 / 1$	$56 \times 56 \times 256$
4	Conv2(1-3层)残差	$3 \times 3 / 1 / 1$	$56 \times 56 \times 256$
5	级联网络	—	$56 \times 56 \times 256$
6	Conv1(4-6层)	$3 \times 3 / 1 / 1$	$30 \times 30 \times 512$
7	Conv2(4-6层)残差	$3 \times 3 / 1 / 1$	$30 \times 30 \times 512$
8	级联网络	—	$30 \times 30 \times 1024$
9	全连接	—	202×512

2.4 模型预测训练

本文将DropSample训练方法^[30]用于深度网络的学习中,这种新的训练技术加快了深度网络的收敛速度,实现了大规模的神经网络训练。在训练过程中,每个训练样本有一个配额,DropSample衡量了该样本被抽取用于训练的概率。配额每一次的更新都取决于该样本被CNN所掌握的程度,如该样本通过CNN后的输出置信度较大,则说明样本被熟练掌握,配额相应减少,并且该样本下次被抽取作为训练样本的概率也会降低;相反,配额提高,该样本下次被抽取作为训练样本的概率会增大。DropSample不仅能将掌握熟练的样本从训练集中逐渐减少,而且还能通过配额更新规则将训练集中噪音较大的样本剔除,给还没被熟练掌握的困难样本更多的机会,以提高深度学习训练的效率。

假设有 m 个训练样本 (x_i, y_i) ,输入向量为 $(n+1)$ 维,CNN的输出有 k 类,则CNN的输出为:

$$h_{\theta_j}(x_i) = \frac{\exp(\theta_j^T x_i)}{\sum_{t=1}^k \exp(\theta_t^T x_i)}, j=1, 2, \dots, k \quad (12)$$

其中, θ_j 是对应的第 j 个CNN输出的参数,包含权值和偏移量。带正则化项的代价函数 $J(\theta)$ 用交叉熵的形式表示,如式(13)所示:

$$J(\theta) = -\frac{1}{m} \left[\sum_{i=1}^m \sum_{j=1}^k 1\{y_i=j\} \log \frac{\exp(\theta_j^T x_i)}{\sum_{t=1}^k \exp(\theta_t^T x_i)} \right] + \frac{\lambda}{2} \sum_{i=1}^m \sum_{j=0}^n \theta_{ij}^2 \quad (13)$$

代价函数 $J(\theta)$ 对 θ_j 误差项的偏导数如式(14)所示:

$$\nabla_{\theta_j} J(\theta) = -\frac{1}{m} \sum_{i=1}^m [x_i (1\{y_i=j\} - p(y_i=j|x_i;\theta))] + \lambda \theta_j \quad (14)$$

其中, p 为softmax输出的概率分布, $1\{\cdot\}$ 是指示函数, λ 为正则化因子。每次迭代后,参数更新的表达式如式(15)所示:

$$\theta_j := \theta_j - \alpha \nabla_{\theta_j} J(\theta), j=1, 2, \dots, k \quad (15)$$

其中, α 为参数学习率。由式(14)和式(15)可知,误差项可以依据softmax输出分成3个子误差项。令 p_i 作为第 i 个样本在CNN中的预测类别的概率,通过阈值 T_1 和 T_2 将 p_i 分隔

成3个子集 M_1, M_2, M_3 ,这3个子集分别对应熟练掌握样本项($T_2 < p_i \leq 1$)、未熟练掌握样本项($T_1 \leq p_i \leq T_2$)、噪声样本项($0 \leq p_i \leq T_1$),具体如式(16)所示:

$$\begin{aligned} \nabla_{\theta_j} J(\theta) &= E_1 + E_2 + E_3 + \lambda \theta_j \\ &= -\frac{1}{m} \left\{ \sum_{i_1 \in M_1} [x_{i_1} (1\{y_{i_1}=j\} - p(y_{i_1}=j|x_{i_1}; \theta))] + \sum_{i_2 \in M_2} [x_{i_2} (1\{y_{i_2}=j\} - p(y_{i_2}=j|x_{i_2}; \theta))] + \sum_{i_3 \in M_3} [x_{i_3} (1\{y_{i_3}=j\} - p(y_{i_3}=j|x_{i_3}; \theta))] \right\} + \lambda \theta_j \quad (16) \end{aligned}$$

其中, E_1, E_2 和 E_3 分别代表以上3个子误差项。针对不同的子误差项,对不同的样本配额更新策略。

基本的配额更新公式为:

$$q_i^t = q_i^{t-1} f(p_i^t) \quad (17)$$

其中, $f(\cdot)$ 是配额更新方程, p_i^t 是第 i 个样本在第 t 次训练中softmax输出的预测概率值, q_i^t 是第 i 个样本在第 t 次训练中的配额值。受上一次配额的影响,每次配额的更新变化不大,这有利于配额更新的稳定性。将 $f(\cdot)$ 设置为两种函数:指数分层函数和多级分层函数。其中,指数分层函数 $f_1(\cdot)$ 如式(18)所示:

$$f_1(p_i^t) = \begin{cases} -\exp(-\gamma p_i^t) + 1, & 0 \leq p_i^t < T_1 \\ -\exp(\beta(p_i^t - 1)) + 1, & T_2 < p_i^t \leq 1 \\ 1/q_i^{t-1}, & \text{otherwise} \end{cases} \quad (18)$$

其中, β 和 γ 是倾斜因子,可调整更新的速率。

多级分层函数 $f_2(\cdot)$ 是一个多级的三层函数,如式(19)所示:

$$f_2(p_i^t) = \begin{cases} a_{1h}, & \{L_{1h} \leq p_i^t \leq U_{1h} | 0 \leq L_{1h} < U_{1h} < T_1, \\ & h=1, 2, \dots, l_1, i \in M_1\} \\ a_{2h}, & \{L_{2h} \leq p_i^t \leq U_{2h} | T_1 \leq L_{2h} < U_{2h} \leq T_2, \\ & h=1, 2, \dots, l_2, i \in M_2\} \\ a_{3h}, & \{L_{3h} \leq p_i^t \leq U_{3h} T_2 < L_{3h} < U_{3h} \leq 1, \\ & h=1, 2, \dots, l_3, i \in M_3\} \end{cases} \quad (19)$$

其中, L_h 和 U_h 分别表示各层中第 h 级的上界和下界,参数 a_{1h}, a_{2h} 和 a_{3h} 为第 h 级的更新因子。由于 M_2 是未熟练掌握的样本,因此设定 $1/a_i^{t-1}$ 为不减少配额。式(19)简化为:

$$f_2(p_i^t) = \begin{cases} a_{1h}, & \{L_{1h} \leq p_i^t \leq U_{1h} | 0 \leq L_{1h} < U_{1h} < T_1, \\ & h=1, 2, \dots, l_1, i \in M_1\} \\ a_{3h}, & \{L_{3h} \leq p_i^t \leq U_{3h} | T_2 < L_{3h} < U_{3h} \leq 1, \\ & h=1, 2, \dots, l_3, i \in M_3\} \\ 1/q_i^{t-1}, & \text{otherwise} \end{cases} \quad (20)$$

在CNN训练后期,大量的噪声过大的样本和被熟练掌握的样本会导致CNN模型过拟合,所以须降低这些样本的配额,让未被熟练掌握的样本集中在一起继续训练,以防止过拟合,达到高效率的目的。DropX训练方法是相兼容,可在短时公交客流预测中灵活运用,能构建高效率、高性能的深度学习网络框架。

通过对广州B3线路6个站点、BRT棠下站的客流进行调查,选取6个分类的样本数据和优选参数建立预测训练模

型,并训练样本。对 B3 线路 6 个站点未来时段(2017-07-10 17:30-18:30)公交站点上车客流进行预测与分析的结果如表 3 所列。对 BRT 棠下站未来一周内(2017-07-17-2017-07-23)早上 7:00-8:00 的公交客流数据进行预测分析的结果如表 4 所列。对未来一月(2017-07)内 B3 线路 6 个站点早/晚高峰的公交客流数据进行预测分析的结果如表 5 和表 6 所列。利用 CNN 算法优化数据集,可以得到高效率、高性能的卷积神经网络训练集。训练样本的分类会直接影响短时公交客流的预测,利用卷积神经网络进行预测得到的预测结果的精确度更高。

表 3 未来时段公交站点上车客流预测(2017-07-10 17:30-18:30)
Table 3 Forecast of passenger flow at bus stations in future period (2017-07-10 17:30 to 18:30)

站名	实际上车客流	预测上车客流	预测绝对误差	预测相对误差/%
松南路口	574	586	12	2.09
罗冲围客运站	346	379	33	9.54
富力半岛花园	415	378	-37	-8.92
西场电器城	682	614	-68	-9.97
和平新村	414	411	-3	-0.72
东风西路	391	408	17	4.35

表 4 未来一周内 BRT 棠下站早上 7:00-8:00 的客流预测(2017-07-17-2017-07-23)

Table 4 Forecast of passenger flow from 7:00 to 8:00 a. m. in the next week(2017-07-17-2017-07-23)

星期	7:00-7:05	7:05-7:10	7:10-7:15	7:15-7:20	7:20-7:25	7:30-7:35	7:35-7:40	7:40-7:45	7:45-7:50	7:50-7:55	7:50-8:00
一	2426	2711	2928	3738	3835	3918	3815	3739	3603	3427	3075
二	2325	2675	2833	3680	3693	3680	3624	3602	3539	3311	2986
三	2317	2698	2817	3674	3730	3766	3701	3630	3588	3233	2816
四	2379	2576	2844	3755	3796	3799	3733	3659	3522	3276	2911
五	2530	2839	3027	3652	3702	3745	3708	3619	3520	3215	2895
六	2213	2378	2738	3472	3529	3579	3538	3507	3435	3053	2717
日	2014	2363	2702	3359	3494	3517	3376	3205	3104	2983	2694

表 5 未来一个月 B3 线路 6 个站点早高峰客流预测(2017-07)

Table 5 Prediction of early peak passenger flow for 6 stations on B3 line in the next month(2017-07)

站名	站点早高峰客流	对向站点早高峰客流	站点早高峰/对向站点早高峰
松南路口	24258	1560	15.55
罗冲围客运站	24540	1612	15.22
富力半岛花园	22480	1690	13.30
西场电器城	19471	1924	10.12
和平新村	21601	2462	8.77
东风西路	11202	1294	8.65

表 6 未来一个月 B3 线路 6 个站点晚高峰客流预测(2017-07)

Table 6 Prediction of late peak passenger flow for six stations on B3 line in the next month(2017-07)

站名	站点晚高峰客流	对向站点晚高峰客流	站点晚高峰/对向站点晚高峰
松南路口	11524	1633	7.05
罗冲围客运站	25029	3779	6.62
富力半岛花园	6021	1029	5.85
西场电器城	6712	1164	5.76
和平新村	11319	2021	5.60
东风西路	12801	2455	5.21

2.5 基于改进 CNN 的深度学习模型

本文设计的基于改进卷积神经网络的短时公交客流预测模型如图 2 所示,该模型主要由输入层、隐含层、输出层、模型训练和模型预测组成。输入层为全连接层,须对样本数据进行处理,以满足 CNN 输入的要求。隐含层由多个 CNN 神经元组成。输出层通过一个全连接网络将隐含层的多个结果重新映射成模型期望的结果。在训练过程中,使用 Adam^[31](Adaptive Moment Estimation)算法的优化器对模型进行优化,使模型的预测结果可视化。羊城通 IC 卡刷卡数据、公交进出站数据为结构化数据,从公交站点、公交车上采集的客流视频图像为非结构化数据。从公交客流数据序列中提取特征数据,对公交客流数据进行分类,本文采用多分支卷积神经网络模型进行客流预测。

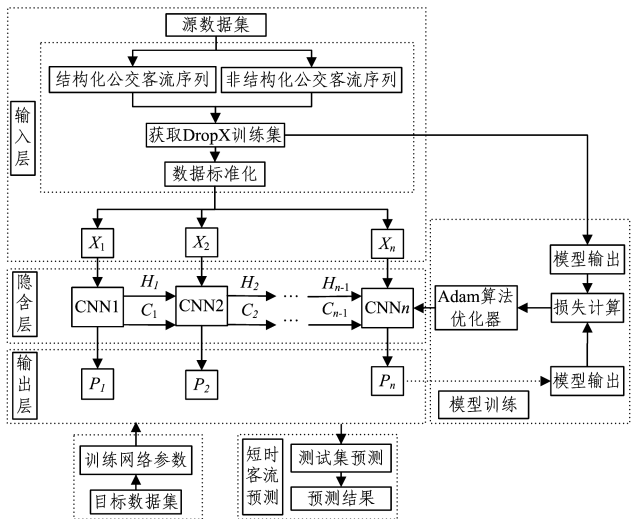


图 2 基于改进卷积神经网络的短时公交客流预测模型

Fig. 2 Short-term bus passenger flow prediction model based on improved convolution neural network

(1)BP 卷积神经网络中算法的实现

BP 卷积神经网络是一种反向传播神经网络,对非线性系统具备很强的模拟能力。通过对训练数据的不断学习,以及对网络权值和阈值的不断更新,网络误差逐渐减小,接近期望输出。

BP 卷积神经网络分全连接层、卷积核的反向传播、池化层的反向通路(Backward Pass)。全连接层的 BP 计算与传统的前馈神经网络相同,卷积层的反向传播是一个与前向传播类似的交叉相关计算:

$$\left(\frac{\partial E}{\partial A}\right)_{i,j}^i = \sum_{k=1}^k \sum_{x=1}^f \sum_{y=1}^f w_k^{i+1}(x,y) \left(\frac{\partial E}{\partial A}\right)_{s_0 i+x, s_0 j+y, k}^{i+1} f'(A_{i,j}) \quad (21)$$

$$w^i = w^{i+1} - \alpha \left(\frac{\partial E}{\partial A}\right)_k = w^{i+1} - \alpha [A^{i+1} \left(\frac{\partial E}{\partial A}\right)_k^{i+1}] \quad (22)$$

其中, E 为误差函数计算的误差, f' 为激励函数的导数, α 为

学习速率。若卷积核的前向传播使用卷积计算,则反向传播对卷积核翻转页也使用卷积计算。CNN 使用了 BP 框架内的随机梯度下降(Stochastic Gradient Descent, SGD)和其变体,如 Adam 算法。SGD 在每次迭代中随机选择样本来计算梯度,在大量学习样本的情形下,有利于信息的筛选。SGD 在迭代初期能快速收敛,计算复杂度更小。

改进 CNN 网络是对 CNN 卷积神经网络进行优化,更新网络参数 θ ,采用 Adam 优化算法替换 SGD 算法以最小化损失函数。SGD 算法保持单一的学习率(α)以更新所有的权重,在网络学习训练过程中学习率未发生变化。Adam 算法优化器计算梯度的一阶矩估计和二阶矩估计,对自适应性学习率设置不同的参数。Adam 算法优化器的内存需求低、计算效率高,其梯度对角缩放(Diagonal Rescaling)具有不变性。网络参数的更新过程如下:

$$g_t = \frac{1}{n} \nabla_{\theta} \sum_i L(F(Y_i, \theta), X_i) \quad (23)$$

$$m_t = u \times m_{t-1} + (1-u) \times g_t \quad (24)$$

$$n_t = v \times n_{t-1} + (1-v) \times g_t^2 \quad (25)$$

$$\hat{m}_t = \frac{m_t}{1-u^t} \quad (26)$$

$$\hat{n}_t = \frac{n_t}{1-v^t} \quad (27)$$

$$\Delta\theta_t = -\eta \times \frac{\hat{m}_t}{\sqrt{\hat{n}_t} + \epsilon} \quad (28)$$

$$\theta_{t+1} = \theta_t + \Delta\theta_t \quad (29)$$

其中,均方误差函数 $L(\theta)$ 对 θ 的梯度表示为 g_t ,梯度一阶矩估计表示为 m_t ,梯度二阶矩估计表示为 n_t , m_t 偏差修正表示为 \hat{m}_t , n_t 偏差修正表示为 \hat{n}_t 。 u 和 v 表示矩估计的指数衰减率, η 表示步长, ϵ 表示数值稳定的小常数, $\Delta\theta_t$ 表示计算的 θ_t 更新值, θ_{t+1} 为 $t+1$ 时刻的 θ 值,对 θ_t 和 $\Delta\theta_t$ 的值求和,并应用到 θ_{t+1} 。

Adam 算法优化器首先初始化参数向量、一阶矩向量和二阶矩向量;然后对每个部分进行循环迭代更新,使参数 θ 收敛。在时间步 $t+1$ 上对一阶矩估计和二阶矩估计进行偏差更新,计算一阶矩估计的偏差修正和二阶矩估计的偏差修正,在时间步 $t+1$ 上对参数 θ 所求的梯度进行目标函数的更新;最后用上述计算的值更新模型参数 θ 。

(2) 卷积神经网络的运算加速

我们将重点研究卷积神经网络的参数压缩方法和运算加速方法,以提升深度学习的实际应用价值,以便于构建相关的公共交通移动互联应用。压缩模型参数的方案如下:1)借鉴 Hinton 等提出的蒸馏方法把大网络或者多网络输出的结果作为监督来重新训练一个参数更少的网络,或者在此方法的基础上用大网络的每一个隐含层输出指导训练小网络对应隐含层的参数;2)通过对卷积核加低秩约束,采用把神经网络参数张量分解为多个低秩张量的方法,并重新训练网络,达到同时降低存储量和运算量的目的。

运算加速的研究思路如下:1)可以采用硬件上的并行化技术(如集群技术、GPU 并行化、CPU 的 SIMD 指令和 ARM 的 NEON 指令^[32]);2)针对卷积运算的特点,可以对数据和

参数进行快速傅里叶变换并进行乘积运算再反变换回来,已证明这种加速算法在卷积模板较大的情况具有明显优势。该加速算法适用于突发异常事件的高性能、低存储深度学习模型的构建及训练。本文研究了深度神经网络的参数压缩方法和运算加速方法,探讨了兼顾效率和性能的深度学习框架。

实验环境配置:实验所用的计算机配置为服务器平台,E5-2600 v4 处理器,内存 2TB,2 块 GPU 卡,Ubuntu 16.04 版本的 Linux 操作系统,深度学习框架为 caffe,软件编程环境为 Python。

3 模型应用验证

3.1 公交客流算法模型

根据客流量的时间分布,在不同时段采用不同的发车间隔,使客流量较大时段的车流密度较大。按照客流分布特征对公交公司的人员、车辆、物资等资源进行合理调配,使企业在运营投入和运营效益之间达到利益的最优化,实现人、车、线的最佳匹配。假设公交公司按设定方案发车运行,乘客在不同时段以不同的客流密度随机到站乘车,从而形成了公交车运营的全过程。乘客上下车模型的设定如下。

(1) 站间行车时间(min)

$$\tau_1 = 0, \tau_i = \frac{L_i}{V_i}, i = 2, \dots, n$$

其中, τ 表示第 $i-1$ 站到第 i 站间的行车时间(含在第 i 站的停车时间)。

(2) 首站发车时间表

首站发车的时间表为 $T = (T_0, T_1, \dots, T_k, \dots, T_m)$,其中 T_0 表示第一辆车到达起点站 S_1 的时间; T_k 表示第 k 辆车驶离起点站 S_1 的时间, $k=1, 2, \dots, m$; T_m 表示最后一辆车驶离起点站 S_1 的时间。 T_{ki} 表示第 k 辆车离开第 i 站的时间, $T_{ki} = T_k + \sum_{j=1}^i \tau_j$,全天内第 i 站的有效候车时间为 $[T_0 + \tau_i, T_m + \tau_i]$, $i = 1, \dots, n$,公交车线路运营的工作时间为 $[T_0, T_m + \tau_n]$,全天共发行 m 辆公交车。

(3) 上下车人数分布

假设乘客每一站、每一时段上车、下车的人数服从等概率分布,则 $U_i(t)$ 表示 t 时刻之前到第 i 站上车的总人数,如式(30)所示:

$$U_i(t) = \begin{cases} \frac{U_{i1}}{\tau - \tau_i} \times (t - \tau_i) \\ U_{i1} + \frac{U_{i2}}{\tau} \times (t - \tau_i) \\ \vdots \\ U_{i1} + U_{i2} + \dots + U_{i(Q-2)} + \frac{U_{i(Q-1)}}{\tau} \times (t - (Q-2)\tau) \\ U_{i1} + \dots + U_{i(Q-1)} + \frac{U_{iQ}}{\tau - \sum_{j=1}^n \tau_j + \tau_i} \times (t - (Q-1)\tau) \end{cases} \quad (30)$$

$$\tau_i \leq t \leq \tau$$

$$\tau < t \leq 2\tau$$

$$(Q-2)\tau < t \leq (Q-1)2\tau$$

$$(Q-1)\tau < t \leq Q\tau - \sum_{j=1}^n \tau_j + \tau_i$$

$D_i(t)$ 表示 t 时刻之前在第 i 站下车的总人数,如式(31)所示:

$$D_i(t) = \begin{cases} \frac{D_{i1}}{\tau - \tau_i} \times (t - \tau_i) \\ D_{i1} + \frac{D_{i2}}{\tau} \times (t - \tau) \\ \vdots \\ D_{i1} + D_{i2} + \dots + D_{i(Q-2)} + \frac{D_{i(Q-1)}}{\tau} \times (t - (Q-2)\tau) \\ D_{i1} + \dots + D_{i(Q-1)} + \frac{D_{iQ}}{\tau - \sum_{j=1}^n \tau_j + \tau_i} \times (t - (Q-1)\tau) \end{cases} \quad (31)$$

$$\tau_i \leq t \leq \tau$$

$$\tau < t \leq 2\tau$$

$$(Q-2)\tau < t \leq (Q-1)2\tau$$

$$(Q-1)\tau < t \leq Q\tau - \sum_{j=1}^n \tau_j + \tau_i$$

$PU_i(T_k, t)$ 表示第 k 辆车离开第 i 站后,间隔 t min 在第 i 站上车的人数,如式(32)所示:

$$PU_i(T_k, t) = U_i(T_k + \tau_i + t) - U_i(T_k + \tau_i) \quad (32)$$

$PD_i(T_k, t)$ 表示第 k 辆车离开第 i 站后,间隔 t min 在第 i 站下车的人数,如式(33)所示:

$$PD_i(T_k, t) = D_i(T_k + \tau_i + t) - D_i(T_k + \tau_i) \quad (33)$$

$B_i(T_k, t)$ 表示第 k 辆车发出后,间隔 t min 从始发站所发的第 $i+1$ 辆车离开第 i 站时,车上所载客的人数,如式(34)所示:

$$B_i(T_k, t) = \sum_{j=1}^i (PU_j(T_k, t) - PD_j(T_k, t)) \quad (34)$$

根据公交客流模型算法计算松南路口、罗冲围客运站、富力半岛花园、西场电器城、和平新村、东风西路 6 个站点在 2017-07-10 17:30-18:30 上车的客流量,预测结果如表 7 所列。

表 7 6 个站点的上车客流量情况(2017-07-10 17:30-18:30)

Table 7 Passenger flow on six stations(2017-07-10 17:30 to 18:30)

站点	5:30-5:35	5:35-5:40	5:40-5:45	5:45-5:50	5:50-5:55	5:55-6:00	6:00-6:05	6:05-6:10	6:10-6:15	6:15-6:20	6:20-6:25	6:25-6:30
松南路口	37	36	38	39	41	44	45	49	57	56	65	67
罗冲围客运站	10	15	16	20	23	27	30	43	51	44	34	33
富力半岛花园	13	17	19	25	32	33	42	47	50	57	43	37
西场电器城	37	39	44	45	63	68	71	79	86	76	43	31
和平新村	10	18	21	23	25	38	39	45	48	57	48	42
东风西路	8	11	19	24	27	31	43	49	46	53	47	33

公交线路客流呈现出空间分布和时间分布的不均匀性,客流特征不仅会影响城市公交的行车组织与计划,还会影响长远的线网规划,因此研究公交客流的规律十分重要。对广州 B3 线路 6 个站点的客流进行预测的结果表明:影响客流的因素主要有天气(最低温度、最高温度、当前降雨量、当天降雨量、是否有雷阵雨)、日期类型(非节假日、节假日)和重大事件活动(演唱会、博览会等)。

3.2 基于改进 CNN 的短时公交客流预测

首先设定获取训练集与获取测试集的函数,经分析将训练集与测试集的个数比设为 8:1,将训练集数据与测试集数据进行数据标准化;然后为 CNN 网络输入层和输出层的权重、偏置设定变量;最后根据基于改进 CNN 的短时公交客流量预测模型确定输入数据维度和输出数据维度。模型采用 AdamOptimizer 优化器进行优化,经过反复的实验分析可以得到,当模型参数 $learning\ rate = 0.02$, $batch_size = 4600$, $timesteps = 64$,训练次数为 15000 时,基于改进 CNN 的短时公交客流量预测模型的预测效果最好。

3.3 模型分析比较

3.3.1 基于多元线性回归的公交客流预测

多元线性回归是一种适用于预测任务多因素分析的预测方法。MLR 用于多因素预测分析的模型表示为 $y = a_0 + a_1x_1 + a_2x_2 + \dots + a_{n-1}x_{n-1} + a_nx_n + \epsilon$ 。其中, y 为预测值; x_1, x_2, \dots, x_n 代表不同的影响因素; a_1, a_2, \dots, a_n 代表每个 x_i ($i=0, \dots, n$) 对应的回归系数; a_0 和 ϵ 分别代表偏置项和误差项。多元线性回归模型特征提取的复杂度不高,特征值构建

符合公交客流特性,适用于公交客流预测的模型。

3.3.2 基于卡尔曼滤波的公交客流预测

卡尔曼滤波采用由状态方程和观测方程组成的线性随机系统的状态空间模型来描述滤波器,利用状态方程的递推性,按线性无偏最小均方误差估计准则,用递推算法对该滤波器的状态变量做最佳估计,以求得滤掉噪声的有用信号的最佳估计。卡尔曼滤波不仅可以用于信号的滤波和估计,而且可以用于模型参数的估计,适用于公交客流状况的预测。

3.3.3 基于 BP 神经网络的公交客流预测

BP 神经网络是一种前反馈性质的网络,称为反向传播神经网络,其学习过程由误差反向传播和模式正向传播组成。BP 神经网络应用广泛,对非线性系统有很强的模拟能力。不断学习训练数据,不断更新网络的权值和阈值,可以使网络误差逐渐减小而接近期望输出。以广州公交客流 OD B3 线路的有效路径作为样本,将乘客选择的路径作为输入层的输入神经元,将路径的选择权值作为输出层的输出神经元,对同一公交站点的客流 OD 得到的有效路径选择权值进行均一化,因此,BP 神经网络适用于公交客流的预测。

3.4 模型误差的分析与比较

为了比较改进 CNN 网络、CNN 网络、MLR、卡尔曼滤波、BP 神经网络模型的预测性能和效果,计算两种误差评价指标,即均方根误差(RMSE)和平均绝对百分比误差(MAPE),如式(35)和式(36)所示:

$$E_{RMS} = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2} \quad (35)$$

$$E_{MAP} = \frac{1}{N} \sum_{i=1}^N \left| \frac{y_i - \hat{y}_i}{y_i} \right| \times 100\% \quad (36)$$

RMSE 表示预测总体的可靠性:该值越大,预测误差越大,说明预测值的平均误差越不可靠;该值越小,预测误差越小,说明平均误差越可靠。MAPE 表示相对的平均偏离程度和可靠性程度。

CNN 网络模型、改进 CNN 网络模型的预测结果如表 8、表 9 所列,CNN 模型和改进 CNN 模型的参数设置相同: $learning\ rate = 0.02, batch_size = 4600, timesteps = 64$ 。当训练次数为 15000 时,改进 CNN 网络模型比 CNN 网络模型的精度更高,其中 $learning\ rate$ 表示学习率, $batch_size$ 表示样本总数, $timesteps$ 表示可利用的时间序列长度。

表 8 不同参数设置的 CNN 模型预测结果的比较

Table 8 Comparison of predicted results of CNN models with different parameters

<i>learning rate</i>	<i>batch_size</i>	<i>timesteps</i>	训练参数	评价指标	
				E_{RMS}	$E_{MAP}/\%$
0.200	4600	64	15000	357.897	0.349
0.020	4600	128	15000	585.325	0.459
0.020	4600	64	15000	312.276	0.193
0.020	4600	256	15000	448.914	0.449
0.020	4600	64	10000	423.765	0.354
0.020	4600	64	20000	370.446	0.334
0.040	4600	64	15000	436.589	0.417
0.002	4600	64	15000	485.791	0.435

表 9 不同参数设置的改进 CNN 模型预测结果的比较

Table 9 Comparison of predicted results of improved CNN model with different parameters

<i>learning rate</i>	<i>batch_size</i>	<i>timesteps</i>	训练参数	评价指标	
				E_{RMS}	$E_{MAP}/\%$
0.200	4600	64	15000	246.236	0.237
0.020	4600	128	15000	453.178	0.318
0.020	4600	64	15000	229.539	0.117
0.020	4600	256	15000	326.6412	0.335
0.020	4600	64	10000	309.320	0.246
0.020	4600	64	20000	265.437	0.215
0.040	4600	64	15000	348.361	0.342
0.002	4600	64	15000	381.653	0.319

通过对 CNN 网络模型、改进 CNN 网络模型的预测结果进行比较,发现 $batch_size$ 增大到一定程度时,其确定的下降方向已基本不再变化,当样本总数 4600 时,模型效果较好。样本总体由全数据集所确定的方向表示,确定的极值所在方向更准确。 $timesteps$ 表示可利用的时间序列长度。从表 8、表 9 中可以发现, $timesteps = 128$ 时 CNN 模型和改进 CNN 模型的精度比 $timesteps = 64$ 时的精度更低,耗时也 longer。这说明过高的 $timesteps$ 会增加模型的收敛时间,提高模型的训练难度,有可能降低模型的精度。 $learning\ rate$ 过高或过低导致的模型过拟合或欠拟合的问题,同样会影响模型的精度。不同参数设置下,改进 CNN 网络、CNN 网络、MLR、卡尔曼滤波、BP 神经网络模型的预测结果比较如表 10 所列。表 10 中的结果说明,相比于 CNN 网络、MLR、卡尔曼滤波、BP 神经网络,改进 CNN 网络的预测精度和可靠性更高。

表 10 改进 CNN 网络、CNN 网络、MLR、卡尔曼滤波、BP 神经网络预测结果的比较

Table 10 Comparison of forecasting results of improved CNN network, CNN network, MLR, Kalman filter and BP neural network

预测模型	评价指标	
	E_{RMS}	$E_{MAP}/\%$
改进 CNN 网络	229.539	0.117
CNN 网络	312.276	0.193
MLR	586.439	0.485
卡尔曼滤波	413.945	0.378
BP 神经网络	376.470	0.362

针对广州 B3 线路 6 个公交站高峰客流、中峰客流和低峰客流数据,分别利用改进 CNN 网络、CNN 网络、MLR、卡尔曼滤波、BP 神经网络模型进行预测,这些模型的耗时情况如表 11 所列。

表 11 改进 CNN 网络、CNN 网络、MLR、卡尔曼滤波、BP 神经网络的耗时比较

Table 11 Time-consuming comparison of improved CNN network, CNN network, MLR, Kalman filter and BP neural network

类型	程序运行时间/s	时间间隔		
		高峰/min	中峰/min	低峰/min
改进 CNN 网络	50	5	8	10
CNN 网络	46	5	8	10
MLR	16	5	8	10
卡尔曼滤波	30	5	8	10
BP 神经网络	42	5	8	10

通过实验分析可知,线性估计模型(MLR、卡尔曼滤波)的程序运行耗时比非线性估计模型(CNN 网络、BP 神经网络)少,虽然 MLR、卡尔曼滤波模型运行的效率较高,但是其无法解决非线性问题和客流存在的随机性、时变性和不确定性等问题,以及对短时公交客流预测不准确的问题。改进 CNN 网络模型虽然运行耗时较长,但预测精度和可靠性更高。由广州公交客流的实际情况可知,改进 CNN 网络在解决公交客流预测方面是有效且可靠的,其程序运行耗时为 50s,完全能满足短时公交客流预测的时效性要求。

4 客流应用分析

短时公交站点客流数据预测为智能公交系统的规划管理、调度运营提供了科学的依据,提高了公交公司的效益,改善了公交的服务质量,提升了乘客的满意度。利用采集到的实时、动态的客流信息,动态调整线路发班计划,可以最大程度地提高公交的运营效率。在各公交车站,乘客不断地随机到达,等待上车。当公交车到站时,乘客上下车。根据公交客流量的时间分布特点,在不同时段可采用不同的发车间隔,由此可使客流量较大的时段的车流密度也较大。公交车按设定的客流预测方案发车运行,乘客在不同时段以不同的客流密度随机到站乘车,便形成了实时的动态公交车辆运营的全过程。通过对广州公交 3 条试点线路进行验证,可知所提模型的成效显著,主要表现在以下两方面^[33]。

(1)通过对公交线路的调查及短时公交站点客流预测模型的数据指标分析,针对广州公交 81 线路、132 线路、581 线路的线路特征采取相关措施来优化预测模型:1)增加线路运

力,高峰期抽调车辆支援,确保线路高峰期运力配置到位,有效疏运高峰客流;2)针对客流高断面班次,采取常规公交与大站快车组合、常规公交与区间车组合的调度优化模型,尤其对早晚高峰客流运营进行重点监控。在 3 条线路公交车上安装视频客流检测设备,结合智能公交调度系统,对 3 条试点线

路实施相关措施前后的重要特征数据进行比较,结果表明,采取相关措施后客量、载程、车日班次、日均班次、车日行程、早高峰出车率、晚高峰出车率等特征指标方面均有较大的提升。公交 81 线路、132 线路、581 线路试点前后重要特征数据的对比如表 12 所列。

表 12 试点线路实施前后重要特征数据的对比

Table 12 Comparison of important characteristic data before and after the implementation of pilot lines

线路	81			132			581		
	实施前	实施后	同比/%	实施前	实施后	同比/%	实施前	实施后	同比/%
客量	667129	673327	0.93	628985	632075	0.49	728644	733577	0.68
载程	227490	228134	0.28	189573	190572	0.53	248139	253998	2.36
车日班次	14.8	14.9	0.68	21.9	22.4	2.28	8.5	8.6	1.18
日均班次	276	281	1.81	332	333	0.30	168	171	1.79
车日行程	216	218	0.93	213	223	4.69	233	233	0.00
早高峰出车率	94.92%	95.01%	0.09	99.67%	101.62%	1.96	98.18%	98.86%	0.69
晚高峰出车率	94.74%	95.01%	0.28	101.32%	103.96%	2.55	97.73%	98.64%	0.93

通过优化公交短时客流预测模型系统,有助于增加公交客流量,提高公交运营收入,提升公交整体服务水平和市民出行满意度。据统计分析,2018 年 1 月至 2018 年 8 月期间,公交 81 线路车日班次同比上升 4.43%,载客里程同比增加 19.58%。

短时公交站点客流预测的准确性、有效性、实时性是制定公交调度运营和公交线网规划的重要依据,应用公交客流监测与预测结果,为公交线网优化、公交调度优化等工作提供了科学的依据。

结束语 本文在对广州市现有公交客流实际调查分析的基础上,提出了一种基于无监督学习的改进卷积神经网络的短时公交站点客流预测方法,解决了卷积神经网络模型、多元线性回归模型、卡尔曼滤波模型、BP 神经网络等传统模型不能有效从源数据集中提取有效特征而导致模型有效性和可靠性低的问题,验证了改进卷积神经网络模型在公交站点短时客流预测中的准确性以及在智能交通系统中的适用性,拓展了深度学习方法的应用范围,并通过实例证明了模型和算法的实用性和可靠性。在现有研究条件下,将继续增加深度 CNN 网络隐藏层的层数,建立更有效、更可靠、更准确的短时公交站点客流预测模型,并将这些模型应用于不同区域、不同公交站点的客流预测,为城市公交线网规划提供科学的依据,为公交调度组织和运营管理提供科学的决策,以应对突发客流,有效缓解交通拥堵,实现市民绿色出行的目的。

参 考 文 献

[1] BAI Y, SUN Z, ZENG B, et al. A multi-pattern deep fusion model for short-term bus passenger flow forecasting[J]. Applied Soft Computing, 2017, 58(1): 669-680.

[2] XUE R, JIAN, CHEN S K. Short-term bus passenger demand prediction based on time series model and interactive multiple model approach[J]. Discrete Dynamics in Nature and Society, 2015, 66(1): 61-78.

[3] LI X S, WANG S L. Estimation of parameters of linear regression model with linear constraints [J]. Statistical Research,

2016, 33(11): 86-92. (in Chinese)

李小胜, 王申令. 带线性约束的多元线性回归模型参数估计[J]. 统计研究, 2016, 33(11): 86-92.

[4] RASYIDI M A, KIM J, RYU K R. Short-Term Prediction of Vehicle Speed on Main City Roads using the k-Nearest Neighbor Algorithm [J]. Journal of Intelligence and Information Systems (S0925-9902), 2014, 20(1): 121-131.

[5] ZHANG C H, SONG R, SUN Y. Kalman Filter-Based Short Term Passenger Flow Forecasting on Bus Stop [J]. Journal of Transportation Systems Engineering and Information Technology, 2011, 11(4): 154-159. (in Chinese)

张春辉, 宋瑞, 孙杨. 基于卡尔曼滤波的公交站点短时客流预测 [J]. 交通运输系统工程与信息, 2011, 11(4): 154-159.

[6] YANG X F, LIU L F. Short-term passenger flow prediction at bus stops based on Support Vector Machine Based on AP clustering [J]. Journal of Wuhan University of Technology (Transportation Science and Engineering Edition), 2016, 40(1): 36-40. (in Chinese)

杨信丰, 刘兰芬. 基于 AP 聚类的支持向量机公交站点短时客流预测 [J]. 武汉理工大学学报(交通科学与工程版), 2016, 40(1): 36-40.

[7] COSTARELLI D, VINTI G. Pointwise and uniform approximation by multivariate neural network operators of the max-product type [J]. Neural Networks, 2016, 81(9): 81-90.

[8] VIEIRA S, PINAYA W H L, MECHELLI A. Using deep learning to investigate the neuroimaging correlates of psychiatric and neurological disorders: methods and applications [J]. Neuroscience and Biobehavioral Reviews, 2017, 74: 58-75.

[9] 焦李成, 赵进, 杨淑媛, 等. 深度学习、优化与识别 [M]. 北京: 清华大学出版社, 2017: 100-120.

[10] LEE H, GROSSE R, RANGANATH R, et al. Unsupervised learning of hierarchical representations with convolutional deep belief networks [J]. Communications of the ACM, 2011, 54(10): 95-103.

[11] HE Y Q, LI B Q. A Combinatorial Learning Model Learning Rate Strategy [J]. Journal of Automation, 2016, 42(6): 954-958. (in Chinese)

贺昱曜, 李宝奇. 一种组合型的深度学习模型学习率策略 [J]. 自

- 动化学报, 2016, 42(6): 954-958.
- [12] WANG X X, XU L H. Research on Short-term Traffic Flow Forecasting Based on Deep Learning [J]. *Transportation System Engineering and Information*, 2018, 18(1): 82-87. (in Chinese)
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): 82-87.
- [13] DENG L, YU D. Deep learning: methods and applications [J]. *Foundations and Trends in Signal Processing*, 2014, 7(4): 197-387.
- [14] AREL I, ROSE D C, KARNOWSKI T P. Deep machine learning a new frontier in artificial intelligence research [J]. *IEEE Computational Intelligence Magazine*, 2010, 5(4): 13-18.
- [15] DONG H Z, LIU Q, FU F J. A string invariant prediction method for short-term bus passenger flow [J]. *Pattern Recognition and Artificial Intelligence*, 2018, 31(9): 846-855. (in Chinese)
董红召, 刘倩, 付凤杰. 短时公交客流的弦不变量预测方法[J]. *模式识别与人工智能*, 2018, 31(9): 846-855.
- [16] FENG W, XIE L, ZENG J, et al. Audio-visual human recognition using semi-supervised spectral learning and hidden Markov models [J]. *Journal of Visual Languages and Computing*, 2009, 20(3): 188-195.
- [17] HINTON G E, OSINDERO S, TEH Y W. A fast learning algorithm for deep belief nets [J]. *Neural Computation*, 2006, 18(7): 1527-1554.
- [18] SABOKROU M, FATHY M, HOSEINI M. Video anomaly detection and localisation based on the sparsity and reconstruction error of auto-encoder [J]. *Electronics Letters*, 2016, 52(13): 1122-1124.
- [19] DAHL G E, YU D, DENG L, et al. Context-dependent pre-trained deep neural networks for large vocabulary speech recognition [J]. *IEEE Transactions on Audio, Speech, and Language Processing*, 2012, 20(1): 30-42.
- [20] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks [C] // *Proceedings of the 26th Annual Conference on Neural Information Processing Systems 2012*. South Lake Tahoe, NV, USA, 2012: 1097-1105.
- [21] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition [J]. *arXiv: 1409. 1556*, 2014.
- [22] HE K M, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition [C] // *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition*. Las Vegas, NV, USA, 2016: 770-778.
- [23] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition [J]. *arXiv: 1409. 1556v6*, 2014.
- [24] CHEN L C, PAPANDREOU G, KOKKINOS I, et al. Semantic image segmentation with deep convolutional nets and fully connected CRFs [J]. *Computer Science*, 2016, 26(4): 357-361.
- [25] HINTON G E, SALAKHUTDINOV R R. Reducing the dimensionality of data with neural networks [J]. *Science*, 2006, 313(5786): 504-507.
- [26] HINTON G E, SALAKHUTDINOV R R. Reducing the dimensionality of data with neural networks [J]. *Science*, 2006, 313(5786): 504-507.
- [27] DENG L, ABDEL-HAMID O, YU D. A deep convolutional neural network using heterogeneous pooling for trading acoustic invariance with phonetic confusion [C] // *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. 2013: 6669-6673.
- [28] HINTON G E, SRIVASTAVA N, KRIZHEVSKY A, et al. Improving neural networks by preventing co-adaptating of feature detectors [J]. *Computer Science*, 2012, 3(4): 212-223.
- [29] BA J, FREY B. Adaptive dropout for training deep neural networks [C] // *Advances in Neural Information Processing Systems (NIPS)*. 2013: 3084-3092.
- [30] YANG W, JIN L, TAO D, et al. DropSample: A new training method to enhance deep convolutional neural networks for large-scale unconstrained handwritten Chinese character recognition [J]. *Pattern Recognition*, 2016, 4(58): 190-203.
- [31] KINGMA D P, BA J. Adam: A method for stochastic optimization [J]. *arXiv: 1412. 6980v8*, 2014.
- [32] PASZKE A, GROSS S, CHINTALA S, et al. PyTorch: Tensors and dynamic neural networks in Python with strong GPU acceleration [OL]. <http://pytorch.org>.
- [33] CHEN S J, XUE Y. Research on Real-time Dynamic Bus Scheduling Based on Unsupervised Learning [J]. *Journal of Chongqing University of Posts and Telecommunications (Natural Science Edition)*, 2019, 31(2): 192-199. (in Chinese)
陈深进, 薛洋. 基于无监督学习的实时公交动态调度的研究[J]. *重庆邮电大学学报(自然科学版)*, 2019, 31(2): 192-199.