

文章编号:1006-2467(2021)02-0124-07

DOI: 10.16183/j.cnki.jsjtu.2020.99.009

基于空洞-稠密网络的交通拥堵预测模型

石 敏, 蔡少委, 易清明
(暨南大学 信息科学技术学院, 广州 510632)

摘 要: 在利用卷积神经网络模型对短时交通拥堵情况等预测场景进行预测时,由于模型的卷积池化操作过程会丢失部分数据,使得目标位置的信息出现丢失及特征的分辨率持续下降,导致模型的预测能力降低. 针对此,本文提出一种空洞-稠密神经网络模型. 首先,利用空洞卷积用较少的网络参数获取更大感受野的特点,充分提取出复杂多变的数据时空特征. 其次,通过下采样及稠密网络的等值映射,解决参数在神经网络层数增加过程出现退化的问题. 最后,取实际的城市道路平均车速数据块对网络结构的有效性进行验证. 结果表明:同卷积神经网络模型相比,该网络结构预测平均绝对误差降低 3%~23%.

关键词: 空洞-稠密网络;时空特征;卷积神经网络;短时交通拥堵预测

中图分类号: TP 391; U 495 **文献标志码:** A

A Traffic Congestion Prediction Model Based on Dilated-Dense Network

SHI Min, CAI Shaowei, YI Qingming
(School of Information Science and Technology, Jinan University, Guangzhou 510632, China)

Abstract: When using the convolutional neural network (CNN) model to predict short-term traffic congestion, due to the convolution pooling operation of the model, part of the data for the information of the target position will be lost, resulting in the decline of the resolution of the output features and the decrease in the predictive ability of the model. To solve this problem, this paper proposes a dilated-dense neural network model. First, it uses dilated convolution to obtain the characteristics of a larger receptive field with fewer network parameters, and fully extracts complex and variable data spatio-temporal characteristics. Then, through down-sampling and equivalent mapping of dense network, it solves the problem of parameter degradation in the process of increasing layers of neural network. Finally, it uses the actual urban road average speed data blocks to verify the validity of the model. The results show that compared with the convolutional neural network model, the average absolute error of the network structure prediction is reduced by 3% to 23%.

Key words: dilated-dense network; spatio-temporal characteristics; convolutional neural networks; short-term traffic congestion prediction

收稿日期:2019-10-29
基金项目:国家青年科学基金(61603153),广州市产业技术重大攻关技术项目(201802010028),广州市“羊城创新创业领军人才支持计划”(2019019)
作者简介:石 敏(1984-),女,湖北省襄阳市人,副教授,主要研究方向为信号处理、图像及图像处理研究.
通信作者:蔡少委,男,硕士,电话(Tel.):020-85220689,E-mail: caishaowei@stu2017.jnu.edu.cn.

在具有不确定性、非线性或时变性的数据中,提取数据的时空特征是获取数据信息的一种重要方式,研究数据的时空特征信息提取方法对于提升模型的预测能力具有重要意义^[1-4]. 在一些复杂的预测场景如在短时交通拥堵预测中,由于车辆速度容易受到包括邻近道路交通拥堵情况、节假日等干扰因素的影响,使得车速数据波动较大且异常值较多. 而常态和异常时的数据特征同样重要,因此需要有效的方法来提取这些数据的时空特征信息. 随着近年来深度学习模型的广泛应用,越来越多的深度学习预测模型通过不同的改进方式来获取并利用数据的时空特征. 罗文慧等^[5]利用卷积神经网络(CNN)能够提取数据时空特征的特点,将 CNN 和 SVR(Support Vector Regression)相结合应用于短时交通流预测. Deng 等^[6]提出基于 CNN 的随机子空间学习方法,将数据转换成图像的形式并进行时空特征信息的提取,从而提高了模型的预测能力. Lin 等^[7]利用稀疏自动编码器提取数据的时空特征信息并作为 LSTM(Long Short-Term Memory)的输入,通过实际交通数据验证了模型的有效性. Kang 等^[8]将车流量、车辆速度和检测器占有率等数据同时作为 LSTM 的输入方式,提高了输入数据特征的多样性,并达到不同数据特征共同影响预测结果的目的. An 等^[9]通过利用残差网络来增加卷积的层数并联合不同时间间隔输入建立起一个模糊卷积神经网络深度模型,使得数据在时间上的特征信息能被充分利用. 上述研究主要采用了 LSTM 和 CNN 两种模型, LSTM 更适合时间序列等具有波动性较小、异常值较少、变化规律较明显的场景预测,在复杂的预测场景中,其对时空特征信息的提取不够充分. 而 CNN 模型由于感受野的变化能够提取出不相邻数据之间的结构信息,可以通过这个特点来获取数据的时空特征信息,用于数据波动性较大、异常值较多的预测场景中^[10-12].

考虑 CNN 提取数据时空特征信息的方法还存在一定存在缺陷,即模型在完成卷积过程后,需要通过池化过程来对特征进行压缩,并对二维数据进行填充,该过程会使图像特征的空间信息出现丢失. 针对此问题,本文提出了空洞-稠密网络结构模型. 该结构利用空洞卷积能灵活控制卷积采样间隔的特点^[13],在降低模型复杂度的同时可以减小池化层的作用. 此外,本文开辟了第 2 条下采样通道,用于提取数据的显著特征,然后将空洞卷积通道与下采样通道在输出部分进行稠密连接^[14-16],使得更多的时空特征信息被传递,从而保证了模型结构的深度和

预测精度,在道路拥堵预测等方面取得较好的预测效果.

1 空洞-稠密网络结构构造

1.1 模型网络结构

本文沿用 CNN 的特征提取和分类器的方式,将输入通过 1 层卷积后传递到以多个稠密块为核心的稠密层中充分提取特征,稠密块之间以卷积核大小为 1×1 的卷积层进行连接,最后通过两层全连接层进行输出,算法结构流程如图 1 所示. 其中稠密块的数量取决于输入矩阵的大小,全连接层为普通的前馈网络,激活函数为 Sigmoid.

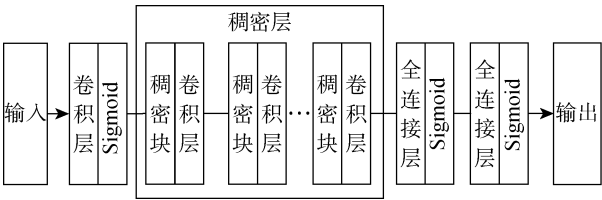


图 1 算法结构流程图
Fig. 1 Flowchart of algorithm structure

1.2 输入矩阵构造

考虑到同一类型的数据在相邻的时间间隔上具有一定的延续性且不同采集数据区域具有关联性的特点,本文将 i 个测量周期内 j 个不同区域采集数据以二维矩阵 \mathbf{I} 的形式进行输入:

$$\mathbf{I} = \begin{bmatrix} x_{1,0} & x_{1,1} & \cdots & x_{1,j} \\ x_{2,0} & x_{2,1} & \cdots & x_{2,j} \\ \vdots & \vdots & & \vdots \\ x_{i,0} & x_{i,1} & \cdots & x_{i,j} \end{bmatrix} \quad (1)$$

式中: $x_{i,j}$ 表示第 i 区域在 j 时刻的采集数据, $i \geq 1$, $j \geq 0$. 每 1 行元素表示 1 个采集点从采集时间 0 到采集时间 j 的 j 个测量数据,每 1 列元素表示同一时刻不同采集点的 i 个测量数据,从而构成基于时空特征的输入矩阵,并作为卷积层的输入.

1.3 稠密块构造

稠密块主要由两条通道构成,第 1 条空洞卷积通道由 3 个卷积块组成,其中每个卷积块由 1 层空洞卷积层、批量归一化层和 Sigmoid 激活函数层构成,卷积块的输入为 \mathbf{X} ,输出为 \mathbf{Y} ,通过改变空洞卷积的 k 值可以改变卷积核大小,并通过采用间隔 r 设置卷积核空洞的大小. 第 2 条下采样通道由两层最大值池化层和 1 层 1×1 的卷积层组成. 稠密块的输入为 \mathbf{I}' ,空洞卷积通道和下采样通道在通道输出部分进行稠密连接,并输出 \mathbf{O} ,稠密块的结构如图 2 所示.

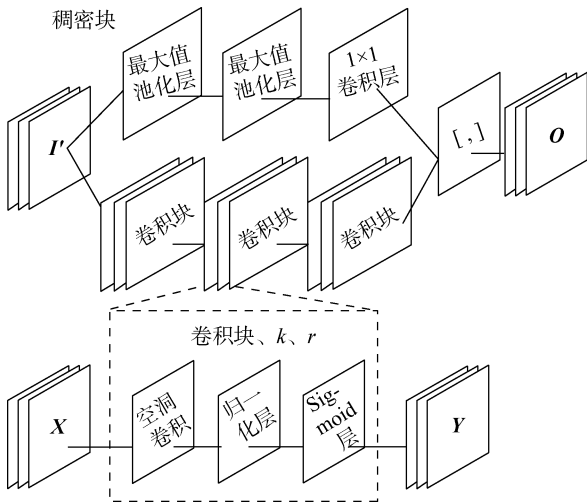


图 2 稠密块结构图

Fig. 2 Structure diagram of dense block

1.4 空洞卷积和下采样通道构造

空洞卷积是在普通的卷积核内部加入空洞,主要是通过控制 r 的大小来实现,如图 3 所示.假设 r 为 1、2 和 3 时, 3×3 大小的卷积核下的感受野随 r 的增大而增大.此外,随着 r 的增大,空洞卷积的卷积核参数数量与传统卷积核的参数数量相同,而卷积过程的输出能够包含数据更大范围内的特征信息.

假设空洞卷积的卷积核大小为 (k_1, k_2) ,输入二维数据大小为 (n, m) ,则随着 r 的变化,卷积核大小变为 (k'_1, k'_2) ,特征图大小变为 (n_1, m_2) ,卷积核大小随 r 的变化满足:

$$k'_1 = k_1 + r(k_1 - 1) \quad (2)$$

$$k'_2 = k_2 + r(k_2 - 1) \quad (3)$$

卷积核提取的特征大小满足:

$$n_1 = [n + 2p - k - r(k - 1)] / (s + 1) \quad (4)$$

$$m_1 = [m + 2p - k - r(k - 1)] / (s + 1) \quad (5)$$

式中: p 为卷积核移动的步长; s 为特征图填充的像素数.假设空洞卷积输入 X' 的大小为 $M \times N$,卷积核权重 W 的大小为 $m \times n$,并且 $M \geq m, N \geq n$,则卷积操作为

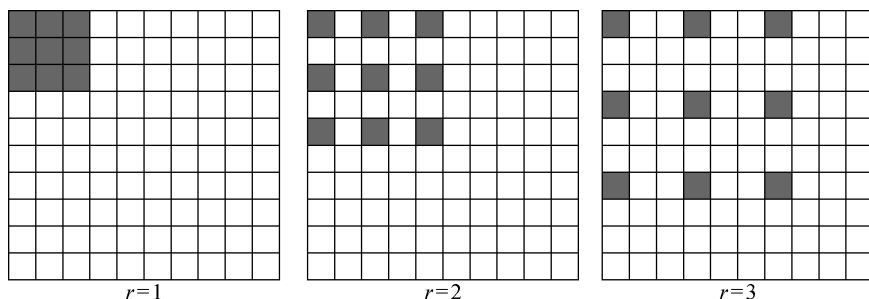


图 3 不同采样间隔下的感受野变化图

Fig. 3 Change in receptive field at different sampling intervals

$$C_{ij} = \sum_{s=1}^m \sum_{u=1}^n X'_{i+m-s, j+n-1} \times W_{su} \quad (6)$$

式中: C_{ij} 为卷积得到的元素, i 和 j 分别为卷积输出矩阵的行和列,满足 $1 \leq i \leq M - m + 1, 1 \leq j \leq N - n + 1$; $X'_{i+m-s, j+n-1}$ 为卷积输入的元素; W_{su} 为卷积核的元素.

为了避免空洞卷积块中的参数更新导致空洞卷积层的输出发生剧烈变化,本文在空洞卷积层后增加 1 层批量归一化层.通过引入小批量数据的均值和标准差来调整输出,从而使整体模型输出更加稳定.用 $f_1(x)$ 表示空洞卷积层的输出, $f_2(x)$ 表示批量归一化过程的输出, $f_3(x)$ 表示批量归一化层输出,则数据归一化过程为

$$f_2(x) = [f_1(x) - \mu_0] / \sqrt{\sigma_0^2 + \xi} \quad (7)$$

$$f_3(x) = \gamma f_2(x) + \beta \quad (8)$$

式中: μ_0 为小批量数据的均值; σ_0^2 为小批量数据的方差; ξ 为 1 个比较小的常数; γ 为拉升系数; β 为偏移系数. γ 和 β 参数通过模型训练学习得到,训练过程中归一化层的重要性得到调整,当 $\gamma = \sqrt{\sigma_0^2 + \xi}$ 且 $\beta = \mu_0$ 时归一化层不起作用.归一化层输出经过激活函数 Sigmoid 后得到卷积块的输出为

$$Y = \text{Sigmoid}(f_3(x)) \quad (9)$$

由于最大值池化层具有保留数据的显著特征且能够进行尺度放缩的特点^[2],本文选择最大值池化层对输入稠密块的数据进行下采样,并将下采样得到的特征通过卷积核大小为 1×1 的卷积层后作为该通道的输出,使得显著的特征能够传递到输出部分,并且特征块能够与空洞卷积层输出特征块大小相适应.

考虑目前特征块连接的方式主要包括残差和稠密连接两种方式,相对于残差网络,稠密连接结合了残差连接能够使网络更好地收敛的优点,同时仅需更少的参数数量,能够降低模型整体参数数量.因此本文的稠密块中两条通道在输出部分采用稠密进行

连接,稠密连接为

$$\mathbf{O} = [\mathbf{f} \ \mathbf{y}] \tag{10}$$

式中: \mathbf{O} 为输出,将空洞卷积通道的输出 \mathbf{f} 和下采样通道的输出 \mathbf{y} 进行稠密连接。

1.5 网络训练

空洞-稠密网络结构训练的过程由两部分构成,输入矩阵经过网络模型进行正向传播得到下一检测周期的预测数据.对这些预测数据进行均方差分析,损失函数为

$$L = \frac{1}{2N} \sum_{i=1}^m \| \mathbf{H}_i - \mathbf{Z}_i \|^2_2 \tag{11}$$

式中: N 为输入样本的批量; \mathbf{H}_i 为空洞-稠密网络模型的输出值; \mathbf{Z}_i 为数据的真实值.考虑本文选取的实测数据具有复杂多变且存在噪声的特点,因此本文选择 Adam(Adaptive Moment Estimation)^[17] 梯度下降算法来更新模型参数,假设 a 时刻,损失函数 L 对于参数的一阶导数为 g_a ,则梯度的一阶矩估计和二阶矩估计式为

$$m_a = \beta_1 m_{a-1} + (1 - \beta_1) g_a \tag{12}$$

$$\delta_a = \beta_2 \delta_{a-1} + (1 - \beta_2) g_a^2 \tag{13}$$

$$m'_a = m_a / (1 - \beta_1^a) \tag{14}$$

$$\delta'_a = \delta_a / (1 - \beta_2^a) \tag{15}$$

式中: m_a 为梯度的指数移动均值; β_1 为一阶矩估计的指数衰减率; δ_a 为平方梯度; β_2 为二阶矩估计的指数衰减率; m'_a 、 δ'_a 为修正矩估计后的值,从而得到参数的权值 θ_a 更新为 θ_{a+1} :

$$\theta_{a+1} = \theta_a - \frac{\eta m'_a}{\sqrt{\delta'_a} + \epsilon} \tag{16}$$

式中: η 为学习率; ϵ 为维持数值稳定性的常熟,设为 10^{-8} .Adam 算法能够通过梯度的一阶矩估计和二阶矩估计对不同模型参数自适应不同学习率,使得模型参数变化比较平稳有利于数据特征的提取。

2 实验及分析

2.1 数据预处理

为了验证空洞-稠密网络结构的有效性,本文选择对城市交通拥堵情况进行预测.本文数据来源于 OpenITS 联盟提供的 2016 年 8 月 1 日至 2016 年 9 月 30 日的广州 214 条匿名路段,主要是主干路和快速路的实测车速数据,数据监测周期为 10 min.选取 00:00~24:00 时间段的车速数据进行训练和预测,以数据的前 55 d 共 221 760 个数据作为训练数据,后 6 d 共 24 192 个数据作为预测数据.由于车辆在发生拥堵时数据的波动较大,模型会出现受局部数据影响较大的情况,导致预测结果与真实数据

拟合程度较差,预测精度较低.所以本文通过 Z-score 标准化方法对输入数据进行预处理,预处理后数据符合标准的正态分布,减小了数值差异的影响,预处理公式为

$$\mu' = (x - \mu) / \sigma \tag{17}$$

式中: x 为输入矩阵 \mathbf{I} 中的数据; μ 为输入矩阵样本的均值; σ 为输入矩阵样本的标准差.通过标准化后使得输出数据 μ' 值在 $[0,1]$ 之间。

2.2 网络结构参数

考虑输入矩阵大小和空洞卷积的采样间隔对模型预测精度的影响,本文将模型的卷积核的统一设置为 3×3 大小,模型的学习率为 0.01,批量大小为 32,稠密块个数为 3,在 Python 的 IDE PyCharm 中进行实验.通过对 3 种不同采样间隔条件下输入不同的道路数进行测试,道路数量每次增加 4 条,模型的训练迭代次数为 8 000,得到损失函数值,如表 1 所示,其中 Loss 为损失函数值。

可以看出, Loss 越小,真实值与预测值的误差越小.由于输入矩阵中道路数决定输入矩阵的行数,因此当道路数为 28,采样间隔为 1、2、4 或 1、2、5 时不能满足本文结构的要求,因此损 Loss 为空,适合采样间隔为 1、2、3,其他情况同理.此外,输入矩阵增大,需要增大采样间隔来减少对预测精度的影响。

表 1 道路数和采样间隔配置表

Tab. 1 Road number and sampling interval configuration

道路数	Loss		
	采样间隔		
	1、2、3	1、2、4	1、2、5
28	0.002 3	—	—
32	0.002 6	0.002 6	—
36	0.002 9	0.003 1	0.002 7
38	0.003 4	0.003 2	0.003 0
40	0.003 7	0.003 3	0.003 3
42	0.003 9	0.003 6	0.003 5

2.3 直观效果对比分析

选择 9 月 26 日和 9 月 27 日相邻两天的数据作为假日和正常日的测试数据集并在同一框图内进行表示.根据表 1 的数据选取道路数为 28,采样间隔分别为 1、2、3 进行观察,结果如图 4 所示.图中: t 为时间, v 为车速. $t=0 \sim 24$ h 所对应数值为 9 月 26 日监测周期内所对应的车辆速度变化情况, $t=24 \sim 48$ h 所对应的数值为 9 月 27 日监测周期内所对应的车辆速度变化情况.本文随机列举其中 4 条道路实际速度测量值和预测值,如图 5 所示,4 个曲线图

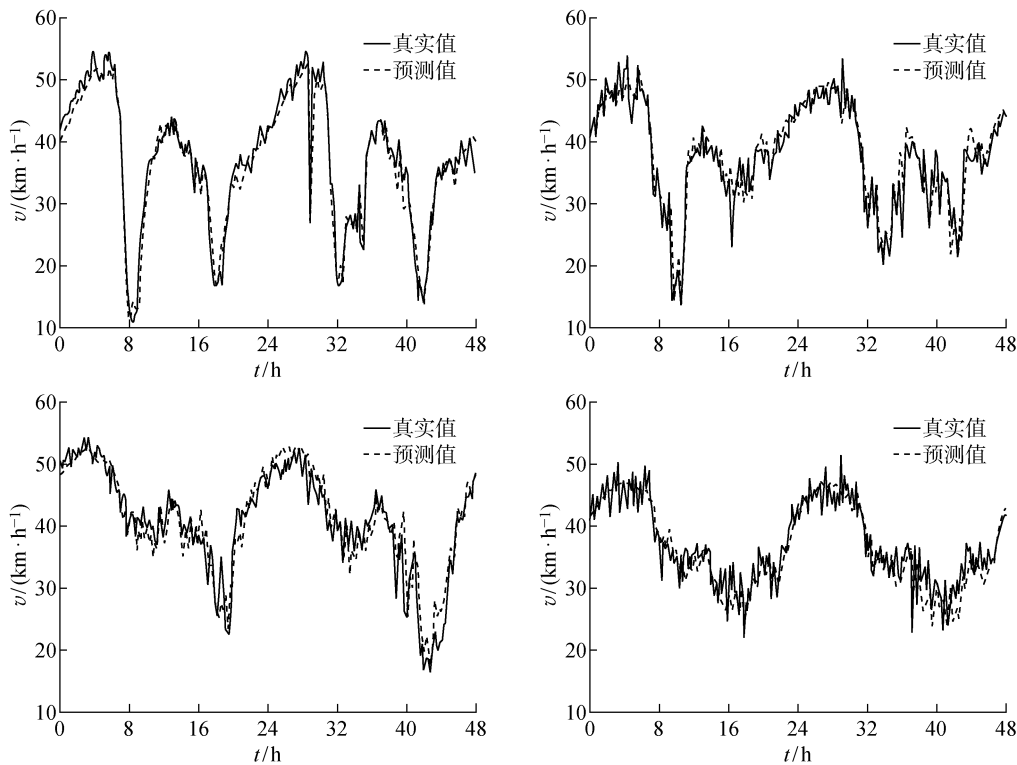


图 4 多条道路真实值与预测值对比

Fig. 4 Comparison of true and predicted values of multiple roads

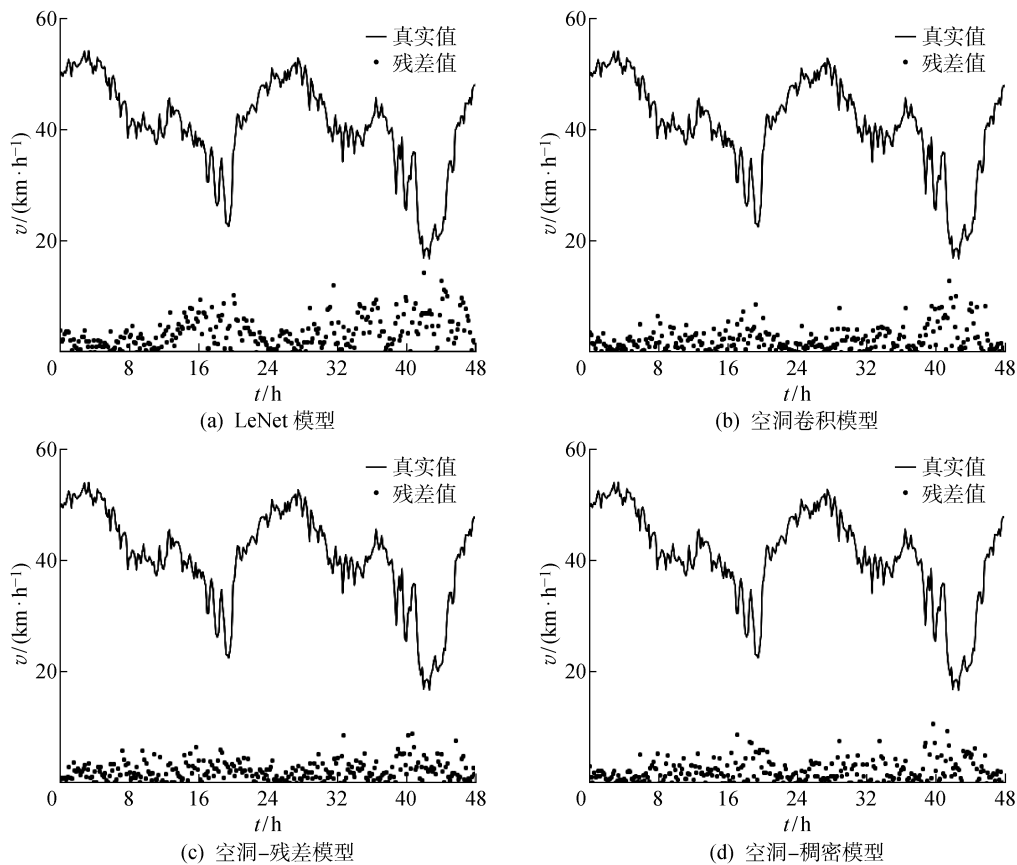


图 5 4 种模型在正常工作日和假日的残差对比

Fig. 5 Comparison of residuals between four models on normal working days and holiday

分别为 4 条道路的车辆速度真实值与预测值变化情况.可以看出,真实值与预测值有相似的变化趋势,拟合程度较高,在第 1 天 8 h 和第 2 天 12 h 左右道路发生拥堵时,车速预测值与真实值之间的误差较小,在其他拥堵情况较平缓的时间段,预测值曲线与真实值的重合度较高.

空洞-稠密结构的预测曲线能够较贴切地反映正常工作日及假日的真实车流量的变化情况,说明该结构能够有效提取出输入矩阵中波动较大点或异常值的特征,且能够正确预测正常值部分的变化情况.此外模型的预测数据的是由 28 条道路的预测数据构成,因此本文模型可以在不增加计算量的同时对 28 道路进行预测,大大降低整体车辆拥堵预测系统的复杂度.

2.4 模型残差对比分析

为了验证模型的鲁棒性,本文引入 LeNet 模型^[11]、空洞卷积模型^[13]、空洞-残差结构模型及空洞-稠密结构模型,并对这些模型在正常工作日和假日两天的真实值与预测值之间的残差值进行分析.根据表 1 数据选取道路数为 28,采样间隔分别为 1、2、3 进行分析,结果如图 5 所示.可以看出,LeNet 模型在不同时刻的残差值相对其他采用空洞卷积的模型的残差值较大,而在速度变化较明显的时间段 4 种模型的残差值变化都较明显.空洞-残差结构模型和空洞-稠密结构模型相对空洞卷积模型的残差值更小,在整体上空洞-稠密卷积结构模型比空洞-残差模型的残差值分布在更小的区域.从测试结果

来看,空洞-稠密结构模型的残差值相比较其他 3 种模型在车速正常状态和异常状态的情况下能保持较小值.

2.5 客观参数对比分析

分析 LeNet 模型、空洞卷积模型及空洞-残差结构模型的平均绝对误差(MAE)和均方根误差(RMSE)来验证模型结构的有效性,并且保证网络模型参数的一致性.MAE 反映了预测值误差的实际情况,而 RMSE 反映了预测的误差分布情况,MAE 和 RMSE 值越小,预测模型拟合程度更好,预测精度更高,其表达式分别为

$$MAE = \frac{1}{2M} \sum_{i=1}^M |F_i - Z_i| \tag{18}$$

$$RMSE = \frac{1}{2M} \sqrt{\sum_{i=1}^M (F_i - Z_i)^2} \tag{19}$$

式中: F_i 为预测值; Z_i 为实际测量值; M 为测量值的数量.

4 种模型通过训练集进行训练后,分别对 9 月 26 日和 9 月 27 日两天车辆速度进行预测.选择图 4 中第 1 条道路数据,分析其 MAE 和 RMSE.

4 种模型的 MAE 和 RMSE 如表 2 所示,相比 LeNet 模型和空洞卷积模型,空洞-残差网络结构模型和空洞-稠密结构模型的 MAE 和 RMSE 都有一定的降低,并且空洞-稠密结构相对空洞残差结构模型 MAE 和 RMSE 更小.在测试集中,相比其他 3 种模型,空洞-稠密模型的 MAE 降低了约 3% 至 23%,而 RMSE 降低了 2% 至 26%.

表 2 空洞-稠密结构和其他 3 种模型的 MAE 及 RMSE 比较
Tab. 2 Comparison of MAE and RMSE between dilated convolution-dense network and the other three models

模型	MAE			RMSE		
	假日	正常日	测试集	假日	正常日	测试集
LeNet	3.16	3.29	3.17	4.93	5.01	4.67
空洞卷积	2.61	2.65	2.72	3.73	3.79	3.86
空洞-残差结构	2.44	2.46	2.51	3.41	3.52	3.50
空洞-稠密结构	2.35	2.43	2.44	3.32	3.49	3.43

3 结语

设计了用于车辆拥堵预测的空洞-稠密神经网络模型.该模型能更好地提取复杂数据的时空特征以及不同道路之间的关联性信息,为预测过程提供更多的判断依据,从而提高了车辆拥堵的预测性能.应用模型对广州市的道路拥堵情况进行预测,并与其他模型进行了比较.结果表明空洞-稠密网络结构

的预测值与真实值的误差相对较小,能够较好地预测车辆速度.而且模型能够实现同时对多条道路进行预测,在错综复杂、数量庞大的城市道路的拥堵预测中可以作为一种有效的方法.由于城市路段与高速路段受干扰因素的不同,表现出对模型结构要求的差异性.因此结合高速路段的相关因素,提高模型的泛化能力并研究模型在其他预测场景中的适用性是未来研究的重要内容.

参考文献:

- [1] CHEN P, DING C, LU G, *et al.* Short-term traffic states forecasting considering spatial-temporal impact on an urban expressway[J]. **Transportation Research Record**, 2016, 2594(1): 61-72.
- [2] ZHANG W, YU Y, QI Y. Short-term traffic flow prediction based on spatio-temporal analysis and CNN deep learning[J]. **Transportation Research Record**, 2019, 15 (2): 1688-1711.
- [3] 陈琛, 吴青, 高嵩. 基于时空因素的水路短期货运量预测[J]. **上海交通大学学报**, 2019, 53 (5): 556-562.
- CHEN Chen, WU Qing, GAO Song. Short-term shipping freight volume prediction based on temporal-spatial features [J]. **Journal of Shanghai Jiao Tong University**, 2019, 53(5): 556-562.
- [4] FENG X, LING X, ZHENG H, *et al.* Adaptive multi-kernel SVM with spatial-temporal correlation for short-term traffic flow prediction [J]. **IEEE Transactions on Intelligent Transportation Systems**, 2018, 20(6): 2001-2013.
- [5] 罗文慧, 董宝田, 王泽胜. 基于 CNN-SVR 混合深度学习模型的短时交通流预测[J]. **交通运输系统工程与信息**, 2017, 17(5): 68-74.
- LUO Wenhui, DONG Baotian, WANG Zesheng. Short-term traffic flow prediction based on CNN-SVR hybrid deep learning model [J]. **Journal of Transportation Systems Engineering and Information Technology**, 2017, 17(5): 68-74.
- [6] DENG S, JIA S, CHEN J. Exploring spatial-temporal relations via deep convolutional neural networks for traffic flow prediction with incomplete data[J]. **Applied Soft Computing**, 2019, 78(3): 712-721.
- [7] LIN F, XU Y, YANG Y, *et al.* A Spatial-temporal hybrid model for short-term traffic prediction[J]. **Mathematical Problems in Engineering**, 2019, 29(1): 1-12.
- [8] KANG D, LV Y, CHEN Y. Short-term traffic flow prediction with LSTM recurrent neural network [C]//**2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC)**. Piscataway, NJ, USA: IEEE, 2017: 1-6.
- [9] AN J, FU L, HU M, *et al.* A novel fuzzy-based convolutional neural network method to traffic flow prediction with uncertain traffic accident information [J]. **IEEE Access**, 2019, 21(1): 20708-20722.
- [10] WANG P, LI L, JIN Y, *et al.* Detection of unwanted traffic congestion based on existing surveillance system using in freeway via a CNN-architecture trafficNet [C]//**2018 13th IEEE Conference on Industrial Electronics and Applications (ICIEA)**. Piscataway, NJ, USA: IEEE, 2018: 1134-1139.
- [11] 张伟斌, 余英豪, 戚湧, 等. 基于时空分析和 CNN 的交通流量短时预测方法[C]//**第十三届中国智能交通年会大会**. 北京: 电子工业出版社, 2018: 349-361.
- ZHANG Weibin, YU Yinghao, QI Yong, *et al.* Short-term traffic flow prediction based on spatio-temporal analysis and CNN deep learning [C]//**The 13th China Intelligent Transportation Annual Meeting**. Beijing: Publishing House of Electronics Industry, 2018: 349-361.
- [12] LIU Q, WANG B, ZHU Y, *et al.* Short-term traffic speed forecasting based on attention convolutional neural network for arterials[J]. **Computer-Aided Civil and Infrastructure Engineering**, 2018, 33(11): 999-1016.
- [13] YU F, KOLTUN V. Multi-scale context aggregation by dilated convolutions[C]//**International Conference on Learning Representations (ICLR)**. Montevideo, Uruguay: IEEE, 2016: 1-13.
- [14] HUANG G, LIU Z, VAN DER, *et al.* Densely connected convolutional networks[C]//**Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)**. Piscataway, NJ, USA: IEEE, 2017: 2261-2269.
- [15] DOLZ J, GOPINATH K, YUAN J, *et al.* Hyperdense-Net: A hyper-densely connected CNN for multi-modal image segmentation[J]. **IEEE Transactions on Medical Imaging**, 2018, 38(5): 1-13.
- [16] ZHANG C, ZHANG H, YUAN D, *et al.* Citywide cellular traffic prediction based on densely connected convolutional neural networks[J]. **IEEE Communications Letters**, 2018, 22(8): 1656-1659.
- [17] KINGMA D K, BA J L. Adam: A method for stochastic optimization[C]//**International Conference for Learning Representations (ICLR)**. Montevideo, Uruguay: IEEE, 2018: 6980-6995.

(本文编辑: 陈晓燕)