

基于非局部操作的深度卷积神经网络车位占用检测算法

申铨京^① 沈哲^① 黄永平^① 王玉^{*②}

^①(吉林大学计算机科学与技术学院 长春 130012)

^②(吉林大学应用技术学院 长春 130012)

摘要: 随着城市交通智能化发展,准确高效地获取可用车位对于解决日益严峻的停车难问题至关重要。该文提出一种基于非局部操作的深度卷积神经网络车位占用检测算法。针对停车位图像特性,引入非局部操作,度量远距离像素间的相似性,直接获取边缘高频特征;使用小卷积核获取局部细节特征;以端到端的方式训练网络。实验中,通过设置不同卷积核尺寸和非局部模块层数,优化网络结构。实验结果表明,该文所提算法与传统的基于纹理特征的车位占用检测算法相比,无论在预测精度还是模型的泛化性能,均具有显著的优势。与当前广泛应用的基于局部特征提取的卷积神经网络相比,该算法具有较大的优势。在真实场景中,该算法同样具有较高精度,具备实际应用价值。

关键词: 车位占用检测; 纹理特征; 卷积神经网络; 非局部操作

中图分类号: TN911.73; TP391

文献标识码: A

文章编号: 1009-5896(2020)09-2269-08

DOI: 10.11999/JEIT190349

Deep Convolutional Neural Network for Parking Space Occupancy Detection Based on Non-local Operation

SHEN Xuanjing^① SHEN Zhe^① HUANG Yongping^① WANG Yu^{*②}

^①(College of Computer Science and Technology, Jilin University, Changchun 130012, China)

^②(College of Applied Technology, Jilin University, Changchun 130012, China)

Abstract: With the intelligent development of urban traffic, accurate and efficient access to available parking spaces is essential to solve the increasingly difficult problem of parking difficulties. Therefore, this paper proposes a deep convolutional neural network parking occupancy detection algorithm based on non-local operation. For the image characteristics of parking spaces, non-local operations are introduced, the similarity between distant pixels is measured, and the high-frequency features of the edges are directly obtained. The local details are obtained by using small convolution kernels, and the network is trained in an end-to-end manner. In the experiment, the network structure is optimized by setting different convolution kernel sizes and non-local module layers. The experimental results show that compared with the traditional texture feature-based parking space occupancy detection algorithm, the proposed algorithm has significant advantages in both prediction accuracy and generalization performance of the model. At the same time, compared with the currently widely used convolutional neural network based on local feature extraction, the algorithm also has great advantages. In real scenes, the algorithm also has high precision and has practical application value.

Key words: Parking space occupancy detection; Texture feature; Convolutional neural network; Non-local operation

收稿日期: 2019-05-17; 改回日期: 2020-01-04; 网络出版: 2020-07-01

*通信作者: 王玉 wangyu001@jlu.edu.cn

基金项目: 智慧法院智能化服务技术研究及支撑平台开发(2018YFC0830100), 国家自然科学基金(61672259, 61876070), 国家自然科学基金青年科学基金(61602203), 吉林省科技发展计划重点科技研发项目(20180201064SF), 吉林省优秀青年人才基金(20180520020JH)

Foundation Items: The Intelligent Court Intelligent Service Technology Research and Support Platform Development (2018YFC0830100), The National Natural Science Foundation of China (61672259, 61876070), The National Natural Science Foundation of China Youth Science Foundation (61602203), The Key Scientific and Technological R & D Projects of Jilin Province Science and Technology Development Plan(20180201064SF), Jilin Province Outstanding Young Talent Fund Project (20180520020JH)

1 引言

近年来, 私家车人均占有量显著增长, 停车难问题越来越严重。随着智慧城市的推广, 基于云平台的停车场应运而生。准确高效地获取车位状态, 并及时反馈给用户, 对于解决停车难问题具有很高的价值。传统停车场通常利用地磁传感器检测车位状态, 安装和维护较复杂, 而且一个地磁传感器只能检测一个车位状态, 极大增加了成本^[1]。相反, 通过视觉算法判断车位状态, 不仅提高检测精度和效率、降低成本, 同时可以监控停车场内物理环境, 获取更多其他信息。多年来, 研究者们已经提出多种基于摄像头的视觉算法^[2]。2002年Dan^[3]提取颜色特征向量并通过SVM分类。2007年Tsai等人^[4]从克服车位间遮挡的角度出发, 他们将3个相邻车位作为一个单元, 将颜色特征直方图送入SVM中分类。2013年Huang等人^[5]提出一种贝叶斯分层框架, 对停车位空间进行3D建模, 类似地, Delibaltov等人^[6]建立车位空间的体积块, 提取车辆的LBP和TB纹理特征, 经过SVM分类, 并进行融合, 计算车辆在车位体积块中的概率, 从而检测停车位状态。上述算法大多通过人工设计车位图像特征, 利用支持向量机等分类算法实现车位占用的检测, 人工设计特征的好坏决定算法的精度和泛化性能, 然而, 现实场景具有多样性, 人工设计的特征过于单一, 针对某个特定场景是可行的, 当应用到其他场景时, 效果变差。随着深度学习^[7]兴起, 通过大数据端到端地训练模型, 使模型自动学习特征, 很好地解决了人工设计特征算法泛化性较差的难题。因此, 本文提出一种端到端的基于非局部操作的深度卷积神经网络算法, 使用非局部操作直接获取图像边缘高频特征, 同时使用小卷积核获取图像局部细节特征。实验中, 探索网络的最优架构, 比较不同非局部模块层数和不同尺寸卷积核对模型的影响。在PKLot^[8]和CNRPark^[9]两个停车位数据集上, 与传统的基于纹理特征方法和基于卷积神经网络方法相比, 算法具有较高精度, 在真实场景中具备实际应用价值。

本文组织结构如下: 第2节介绍本文所提出的车位占用检测算法; 第3节是本文的实验结果和分析; 第4节中测试了算法在真实场景中的检测效果; 第5节总结全文并展望未来。

2 基于非局部操作的深度卷积神经网络

2.1 非局部操作

近年来, 卷积神经网络在计算机视觉领域取得了突破性进展, 卷积操作其实质是一种局部操作,

卷积层无法覆盖全局信息, 卷积核以一定的步长提取局部区域内的特征, 通过增加卷积层数获取全局特征。然而, 这种方式对特征的传递并不灵活, 而且, 盲目增加卷积层数, 往往会增加模型训练的难度, 使网络难于优化, 甚至造成过拟合。

非局部操作的思想来源于图像去噪领域的NL-mean^[10]和处理时间序列信息的Self-attention^[11]方法。非局部操作由Wang等人^[12]提出, 主要解决视频中人体动作的识别问题。作者认为传统的卷积层只能覆盖局部的信息, 无法涵盖全局信息, 而非局部网络可以融合长程的时间序列和空间序列信息。式(1)给出了非局部操作的一般性定义。

$$y_i = \frac{1}{C(x)} \sum_{\forall j} f(x_i, x_j) g(x_j) \quad (1)$$

式(1)中 x 表示输入信号(图片、序列、视频特征), y 表示输出信号, y 的尺寸与 x 相同。 i 和 j 是输入输出信号的某个位置(空间、时间、时空)索引, 使用一个二元函数 f 来度量远距离特征间(位置 i 和所有位置 j)的相似性, 一元函数 g 是输入信号在位置 j 的响应函数, $C(x)$ 表示归一化因子。

非局部操作考虑了所有位置($\forall j$)信息, 相反, 卷积操作是针对输入的局部邻域信息的加权求和(例如, 对于1D尺寸为3的卷积核, $i-1 < j < i+1$)。非局部操作也不同于全连接操作, 非局部操作利用两点的相似性对每个位置的特征做加权, 而全连接操作则是学习位置相关的权重参数, 直接对每个位置加权, 非局部操作输出与输入保持一致, 而全连接需要固定输入输出尺寸, 丢失了位置信息。

为了更好地接入当前主流的网络架构中, 作者将非局部操作包装成非局部模块。公式为

$$\mathbf{W}_Z \mathbf{Y} + \mathbf{X} \quad (2)$$

式(2)中 \mathbf{W}_Z 为可学习的权值, "+ \mathbf{X} "表示一个残差连接^[13], \mathbf{Z} 为模块输出。残差连接在不改变网络结构的情况下, 可将非局部模块接入到任何预训练的网络中进行迁移学习。以空间信息(图片)为输入, 图1详细阐述了一个实例化的非局部模块内部的细节, 使用嵌入高斯函数实例化函数 f , 则

$$f(x_i, x_j) = e^{\theta^T(x_i) \phi(x_j)} \quad (3)$$

式(3)中嵌入项 $\theta(x_i) = \mathbf{W}_\theta x_i$, $\phi(x_j) = \mathbf{W}_\phi x_j$, \mathbf{W}_θ 和 \mathbf{W}_ϕ 为网络自动学习的权值。响应函数 $g(x_i) = \mathbf{W}_g x_i$, \mathbf{W}_g 同样为自动学习的权值。归一化因子 $C(x) = \sum_{\forall j} f(x_i, x_j)$, 不难发现, 对于式(1), 给定 i , $\frac{1}{C(x)} \sum_{\forall j} f(x_i, x_j)$ 即为沿着维度 j 的softmax计算式, 则式(1)等价于

$$y = \text{softmax}(\mathbf{x}^T \mathbf{W}_\theta^T \mathbf{W}_\phi^T \mathbf{x})g(\mathbf{x}) \quad (4)$$

图1中输入特征图 \mathbf{X} 尺寸 $N \times H \times W \times C$ ， N 为batchsize， H 为高度， W 为宽度， C 为通道数量。首先，本文以 1×1 随机初始化的卷积核来实例化 \mathbf{W}_θ 和 \mathbf{W}_ϕ 两个权值，并设置它们的通道数量为 $C/2$ ，这种瓶颈结构^[13]降低了1/2的计算代价。然而，对 $\forall j$ 位置的计算代价仍然是很大的，通过在 ϕ 和 g 之后增加 2×2 的最大值池化层，这并不影响非局部操作的性能，反而只需要1/4的计算代价。最后，本文设置 \mathbf{W}_z 为 1×1 随机初始化的卷积核，其数量为 C ，从而保证输出 \mathbf{Z} 与输入 \mathbf{X} 具有相同的尺寸($N \times H \times W \times C$)。

2.2 模型结构

停车位占用检测，可理解为车位图像的二分类问题。Amato等人^[14]提出mAlexnet卷积神经网络用于解决停车位占用检测问题，mAlexnet只有3层卷积层和2层全连接层，相比Alexnet^[15]，其具有较浅的网络层数和更少的参数。事实上，对于图2中的停车位图像，车辆几乎占据整张停车位图像的全

部面积，需要更关注远距离像素间的联系，从而捕获车辆轮廓等边缘高频信息。然而，通过提取局部特征，堆叠卷积操作的卷积神经网络对车位图像远距离边缘信息的传递并不灵活。

图3为模型的结构图，其中小矩形块表示卷积操作，大矩形块表示非局部操作，圆点表示全连接层的神经元。模型使用3个卷积层，每层均由 5×5 的小卷积核来提取局部细节特征，使用3个非局部模块提取全局特征，每个卷积层之后接入非线性激活函数ReLU，每个非局部模块之后接入最大值池化层，然后经过2个全连接层和softmax函数输出预测结果。第1层全连接层包含96个神经元，第2层包含2个神经元。在第1层全连接层中使用50%的Dropout和L2正则化操作加快模型的收敛速度并抑制过拟合。第1个卷积层中使用32个尺寸为 5×5 的卷积核，由于第1个卷积层的输入是原始图片，信息冗余较大，本文设置卷积核的滑动步长为4，粗略提取局部信息。在第2，第3层卷积层中设置滑动步长为1，同时增加卷积核数量至64个，提取更精细的局部信息。将非局部模块接入到每层卷积层之后，卷积层所提取的局部信息特征图经过非局部模块的作用，远距离的信息之间得到融合，使提取的特征图更有益预测分类结果。

对于输入，PKLot数据集中图片间的高度和宽度并不统一，最小高度和宽度介于58和31像素之间，最大高度和宽度介于176和85像素之间^[16]。本文使用插值法将所有输入图片统一成 224×224 的尺

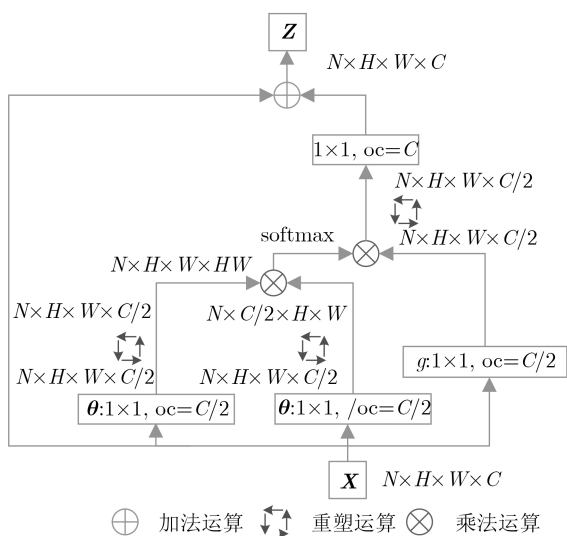


图1 非局部模块



图2 停车位图像

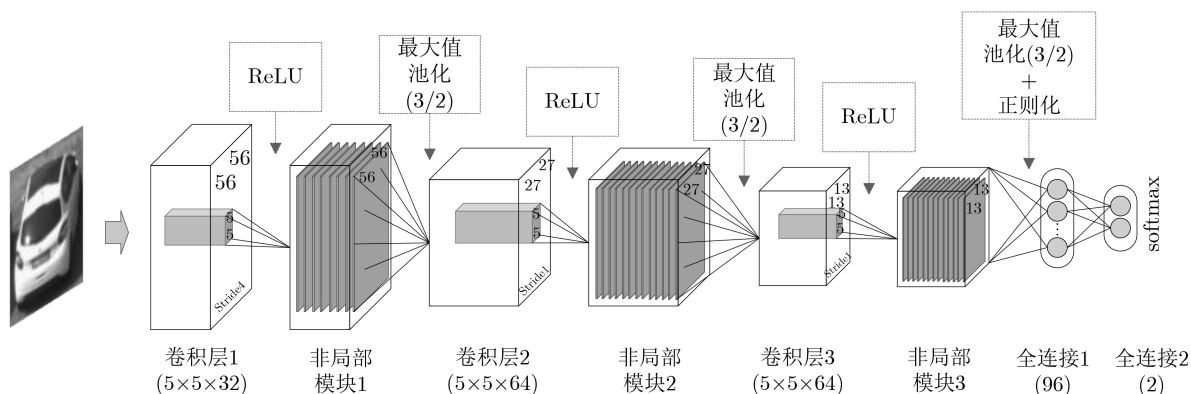


图3 模型结构图

寸, 并进行随机翻转, 归一化处理。采用梯度下降法, 以64张随机打乱顺序的图片为一个批量训练模型。因此, 针对上述问题, 本文提出基于非局部操作的深度卷积神经网络。主要贡献概括如下3点:

(1) 针对停车位问题, 引入非局部操作, 通过非局部模块直接构建特征图中远程特征的联系, 降低堆叠卷积操作获取远程信息的不灵活性。

(2) 通过小卷积核捕获车辆的细节信息, 使用3层卷积层、2层全连接层, 拥有较少的参数量。

(3) 本文算法在PKLot数据集上取得较高的准确率和更强的泛化能力。在真实场景中同样具有较高准确率。

3 实验结果和分析

为验证算法性能, 本文在PKLot数据集和CNRPark数据集上进行实验。PKLot包含695899张停车位图片, 分为UFPR04, UFPR05和PUCPR 3个子集, 每个子集的训练集和测试集划分比例为1:1。所有图片都是在晴天、雨天和多云等不同天气条件下, 在30 d内, 以5 min的时间间隔, 分别在UFPR和PUCPR户外停车场拍摄获取。CNRPark数据集包含12584张停车位图片。其遮挡条件更严格, 包括局部甚至全部遮挡, 如路灯、树叶等; CNRPark包含亮度更低的图片, 有益于模型应对现实中天色较暗时摄像头捕捉图像不清晰的情况。CNRPark拍摄角度更低, 更接近现实中路侧停车位的使用场景。

本文使用Tensorflow框架训练模型, 所有的实验都是基于Intel (R) Core (TM) i7-7800X CPU @ 4 GHz, 64 GB RAM和两张NVIDIA GeForce GTX 1080 Ti GPU完成。使用Adam优化器, 使用随机高斯函数初始化卷积核权值, 设置卷积核初始权值均值为0, 标准差为0.01, 设置偏置项初始值为0。随机从PKLot测试集选出一半作为验证集, 实验中, 完整训练12轮数据集, 为使优化器更好地更新卷积核权值, 使用可衰减的学习率, 当验证集的精度停止增长时, 减小学习率, 学习率初始值设置为0.01, 每完整训练4轮数据, 使学习率衰减10%。

3.1 优化模型结构

Amato等人已经探讨使用3层卷积层和2层全连接层mAlexnet的预测准确率较高, 然而, 在本文模型中, 为使卷积操作更好地结合非局部操作, 在使用3层卷积层和2层全连接层基础之上, 探讨不同卷积核尺寸对预测准确率的影响, 同时增加非局部模块层数, 以优化模型结构。我们发现UFPR04和UFPR05的拍摄视角低于PUCPR, 更接近真实场

景。UFPR04数量较少, 利于训练, 因此, 本文在UFPR04训练集上训练模型, 然后在UFPR04, UFPR05和PUCPR测试集上检验模型性能。

首先, 只在第1层卷积层之后接入非局部模块, 并分别设置卷积核尺寸为 3×3 , 5×5 , 7×7 , 9×9 和 11×11 。在UFPR04训练集上训练12轮, 图4(a)中给出5种卷积核的训练准确率曲线。可以看出, 在经过1轮训练, 5种尺寸的训练准确率均在90%以上, 其中 5×5 的卷积核准确率较高; 随后5轮, 除 9×9 的卷积核在第5轮准确率有所下降, 其他卷积核准确率均稳步上升; 在7~12轮, 准确率曲线趋于平稳, 在第12轮, 3×3 , 5×5 的准确率为99.97%, 99.90%, 而 9×9 , 11×11 的准确率只有99.44%和99.40%, 说明较小卷积核训练准确率要高于较大卷积核。

图4(b), 4(c), 4(d)中分别给出在UFPR04, UFPR05和PUCPR测试集上每轮次的测试准确率曲线。UFPR04测试集上, 3×3 , 5×5 , 7×7 的准确率相对接近, 9×9 , 11×11 明显落后, 而在UFPR05和PUCPR测试集上, 小卷积核优于大卷积核的现象更加明显, 表1给出了第12轮详细的准确率对比, 尽管 3×3 的训练准确率较高, 然而在3个子测试集上 5×5 的准确率都是最高的。因此, 本文设置卷积核的尺寸为 5×5 。对于接入不同层数的非局部模块, 同样采用上述的方法实验, 在使用 5×5 的卷积核情况下, 比较分别接入1, 2, 3层非局部模块的准确率。图5为接入1, 2, 3层非局部模块的准确率曲线图, 在UFPR04训练集和测试集上(图5(a), 图5(b))3种层数几乎无准确率差, 而在不同的子测试集UFPR05, PUCPR上, 使用3层非局部模块效果较好。表2详细给出第12轮不同层数非局部模块的准确率对比。最终, 本文确定使用3层非局部模块, 即每层卷积层之后均接入非局部模块, 使用 5×5 的小卷积核。

为更直观地说明接入非局部模块的影响, 本文可视化了特征图。图6展示输入一张停车位图片, 网络前向推导时, 第1层特征图的可视化结果。其中图6(a)为经过第1层卷积层之后的32张子特征图的可视化结果, 图6(b)为32张子特征图相加融合的特征图, 图6(c)为经过第1层非局部模块之后的32张子特征图可视化结果, 图6(d)为相加融合之后的特征图。显然, 特征图中车辆轮廓信息更清晰, 非局部操作有效融合远距离高频特征, 使特征更丰富。

3.2 数据集内部实验

首先, 在UFPR04, UFPR05和PUCPR 3个子数据集各自的训练集、验证集和测试集上实验, 模

型分别取得了99.85%, 99.60%, 99.92%的测试准确率。表3给出了本文方法与mAlexnet方法^[14]和基于纹理特征的LPQ^[17]以及均值融合方法的测试准确率详细对比。

尽管在表3中, 本文方法相较于其他方法只有不足1%的准确率优势。然而, 上述实验仅仅在各自子数据集, 即同一种拍摄角度下实验的, 这不能完全突出本文方法的性能, 因此, 我们在子数据集

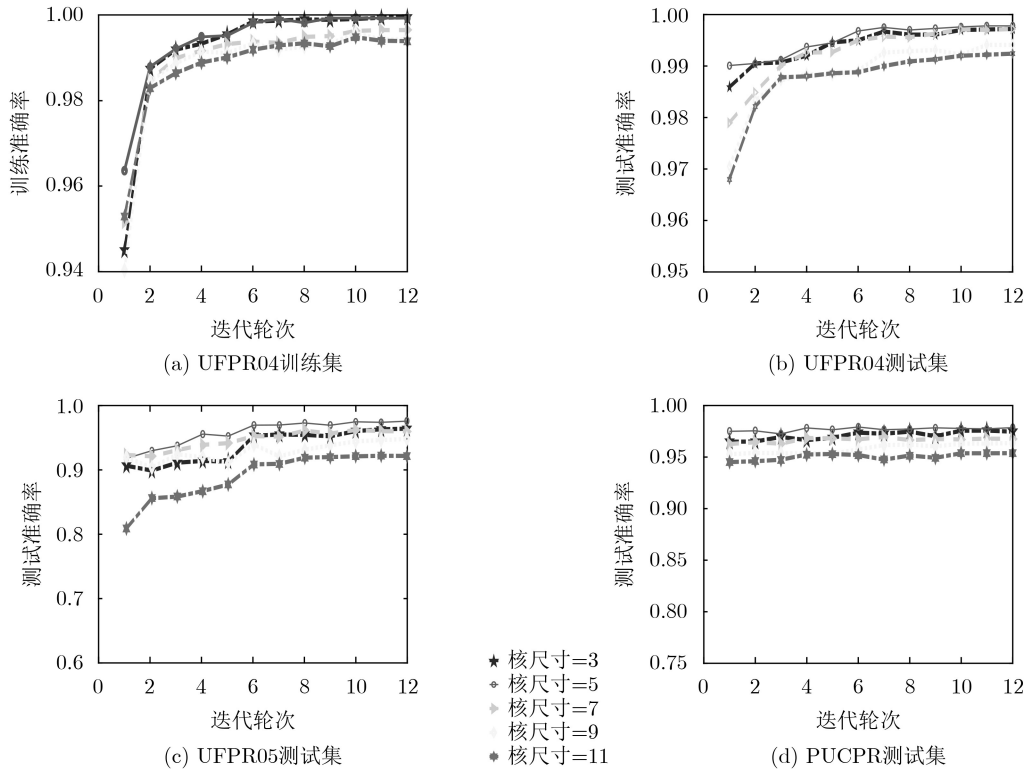


图4 不同卷积核尺寸的准确率曲线图

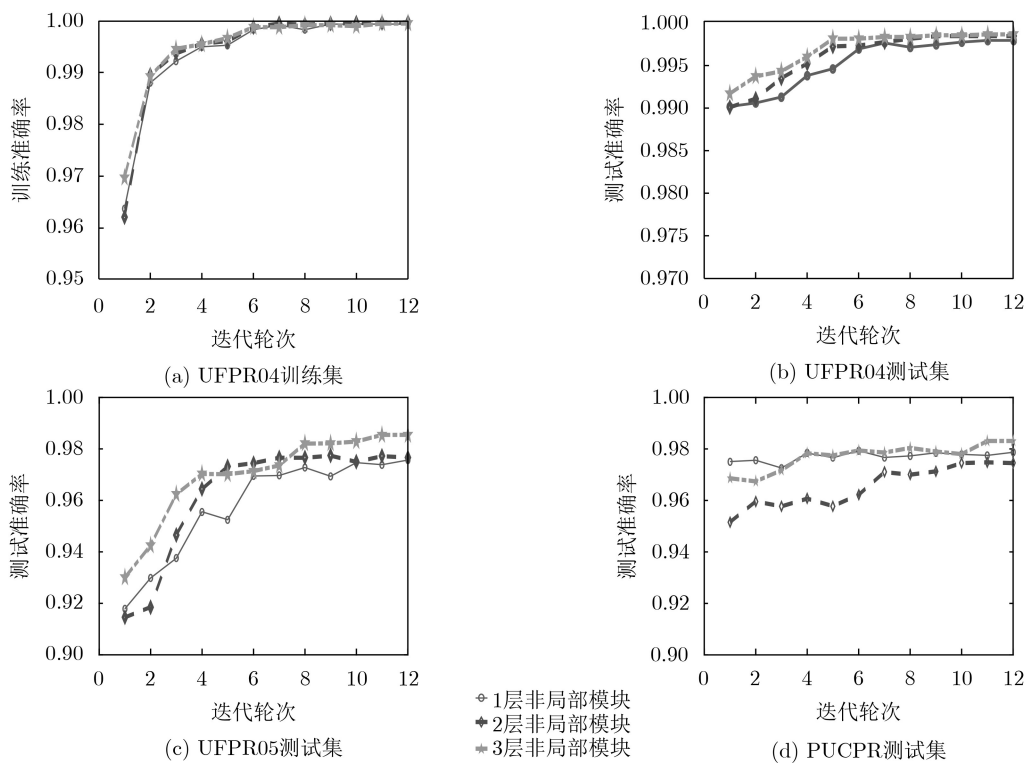


图5 不同层数非局部模块的准确率曲线图

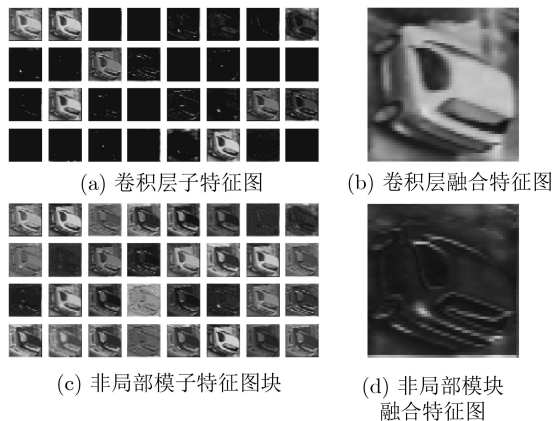


图6 可视化的特征图

表1 不同卷积核尺寸的准确率详细对比(%)

卷积核尺寸	训练精度	测试精度		
		UFPR04	UFPR05	PUCPR
3	99.97	99.74	96.40	97.48
5	99.90	99.78	97.67	97.85
7	99.56	99.72	96.00	96.78
9	99.44	99.41	94.81	96.38
11	99.41	99.25	92.18	95.39

表2 不同层数非局部模块的准确率详细对比(%)

非局部模块层数	训练精度	测试精度		
		UFPR04	UFPR05	PUCPR
1	99.90	99.78	97.67	97.85
2	99.96	99.81	97.65	97.55
3	99.95	99.85	98.55	98.35

表3 不同方法的PKLot子数据集内测试准确率(%)

训练集	UFPR04	UFPR05	PUCPR
测试集	UFPR04	UFPR05	PUCPR
本文方法	99.85	99.62	99.92
mAlexnet	99.54	99.49	99.90
LPQu	99.50	98.90	99.58
Mean	99.64	99.30	99.61

间进行交叉训练测试, 分别使用1个子数据集训练模型, 然后通过其他2个子数据集测试性能。在表4中, 本文方法相较mAlexnet方法有着1%~3%准确率优势, 然而, 传统的基于纹理等方法只有不足90%的准确率。

3.3 数据集间实验

CNRPark数据集比PKLot数据集更接近真实拍摄场景, 本文在CNRPark数据集上检验模型的泛化能力。为减少训练时间, 选择在2 d时间内拍

表4 不同方法的PKLot子数据集间测试准确率(%)

训练集	测试集	方法	精度
UFPR05		本文方法	98.55
		mAlexnet ^[14]	93.29
		LPQg ^[18]	84.92
UFPR04		Max	88.33
		本文方法	98.31
PUCPR		mAlexnet ^[14]	98.27
		LPQg ^[18]	84.25
		Mean	88.40
UFPR04		本文方法	94.45
		mAlexnet ^[14]	93.69
		LPQg ^[18]	85.76
		Mean	85.53
UFPR05		本文方法	95.87
		mAlexnet ^[14]	92.72
		LPQu ^[17]	87.74
PUCPR		Mean	89.83
		本文方法	99.24
		mAlexnet ^[14]	98.03
		LPQg ^[18]	87.15
UFPR04		Mean	88.88
		本文方法	98.89
		mAlexnet ^[14]	96.00
PUCPR		LBPr ^[19]	82.78
		Mean	84.20
		本文方法	98.89

摄的PKLot数据集(包含不同天气的UFPR04, UFPR05和PUCPR子数据集)训练模型。根据遮挡程度, 称CNRPark数据集中路灯、少部分树叶遮挡为轻度遮挡, 称较多树叶遮挡为重度遮挡。从CNRPark数据集中选出轻度、重度遮挡以及亮度较暗的车位图片分为3组, 每组各1500张, 并随机从每组中选取1000张图片用于测试模型。本文方法在CNRPark全部数据上准确率达91.2%, mAlexnet只有82.90%, 而LPQ和LBP只有63%~65%的准确率。对于轻度遮挡和亮度较暗情况, 基于深度学习的方法效果明显好于传统人工设计特征的方法。本文方法更关注车位图像的整体信息, 对于轻度遮挡鲁棒性高于mAlexnet方法。然而, 对于重度遮挡, 树叶占据了大部分车位面积, 几种方法准确率都只有50%左右, 模型已经无法准确区分占用和空缺。图7给出本文方法与其他5种方法分别在CNR-Park全部数据、轻度遮挡、重度遮挡和亮度较暗4组测试结果的准确率对比。

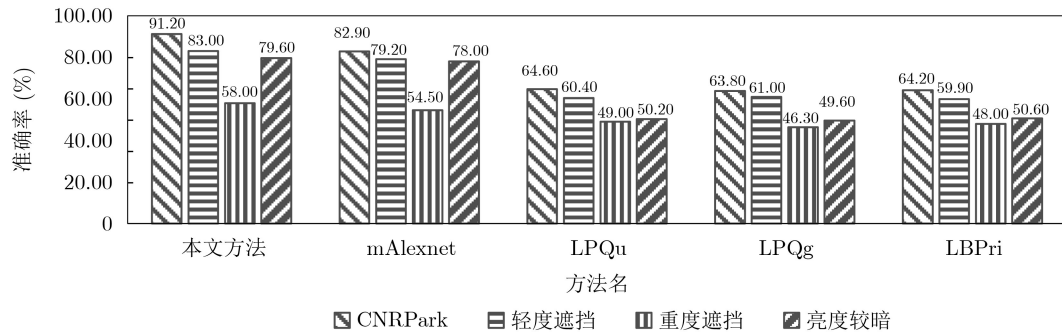


图7 PKLot, CNRPark数据集间实验准确率对比柱状图

4 真实场景测试

本文测试了模型在真实场景中的效果。利用TensorFlow Serving框架将模型部署至云端，通过海康威视网络摄像头(型号DS-2CD1021FD-IW16 mm焦距)捕捉停车区域图片，针对停车位分割，本文方法是在摄像头捕获的第1帧图片上通过鼠标人工框定车位，并为车位分配ID，程序会将车位的位置ID信息保存至json文件中。在之后的检测过程中读取json文件，根据位置信息自动分割出停车

位，使用REST协议将车位图片发送至云端，由模型自动返回检测结果。图8(a)，图8(b)为吉林大学前卫南区匡亚明停车场部分检测结果。其中图8(a)为白天拍摄，图8(b)图为夜间拍摄，检测框内为车位id和状态标识，“Occupied”表示占用状态，“Vacant”表示空缺状态。针对遮挡和光照较暗情况，白天28个车位中正确检测出27个车位状态，夜间28个车位中正确检测出26个车位的状态，模型具有较高的检测精度。



图8 匡亚明楼停车场车位占用情况检测结果

5 结束语

本文针对在停车位占用检测问题中，传统的基于纹理特征检测算法，存在泛化能力差准确率问题，而基于卷积操作的神经网络算法，高频特征的传递不够灵活。本文提出基于非局部操作的深度卷积神经网络算法，使用小卷积核提取局部细节特征，融合非局部操作，直接提取车位图像的全局高频特征。实验结果表明，本算法具有较高的准确率和更强的泛化性能。对于未来停车位占用检测的研究工作，本文认为现有停车位数据集仍然无法涵盖真实场景诸多因素，如春季冬季等季节变化、更严苛的遮挡情况。因此，数据集的扩增对未来停车位检测的研究至关重要。模型的轻量化，边缘终端设备的部署等研究同样具有重要意义。

参考文献

- [1] CAICEDO F, BLAZQUEZ C, and MIRANDA P. Prediction of parking space availability in real time[J]. *Expert Systems with Applications*, 2012, 39(8): 7281–7290. doi: [10.1016/j.eswa.2012.01.091](https://doi.org/10.1016/j.eswa.2012.01.091).
- [2] DEL POSTIGO C G, TORRES J, and MENÉNDEZ J M. Vacant parking area estimation through background subtraction and transience map analysis[J]. *IET Intelligent Transport Systems*, 2015, 9(9): 835–841. doi: [10.1049/iet-its.2014.0090](https://doi.org/10.1049/iet-its.2014.0090).
- [3] DAN N. Parking management system and method[P]. US, 20030144890, 2003.
- [4] TSAI L W, HSIEH J W, and FAN K C. Vehicle detection using normalized color and edge map[J]. *IEEE Transactions on Image Processing*, 2007, 16(3): 850–864. doi: [10.1109/tip.2007.891147](https://doi.org/10.1109/tip.2007.891147).
- [5] HUANG C C, TAI Yushu, and WANG S J. Vacant parking space detection based on plane-based Bayesian hierarchical framework[J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2013, 23(9): 1598–1610. doi: [10.1109/](https://doi.org/10.1109/)

- tcsvt.2013.2254961.
- [6] DELIBALTOV D, WU Wencheng, LOCE R P, *et al.* Parking lot occupancy determination from lamp-post camera images[C]. The 16th International IEEE Conference on Intelligent Transportation Systems, The Hague, Netherlands, 2013: 2387–2392. doi: [10.1109/itsc.2013.6728584](https://doi.org/10.1109/itsc.2013.6728584).
- [7] LECUN Y, BENGIO Y, and HINTON G E. Deep learning[J]. *Nature*, 2015, 521(7553): 436–444. doi: [10.1038/nature14539](https://doi.org/10.1038/nature14539).
- [8] DE ALMEIDA P R L, OLIVEIRA L S, BRITTO JR A S, *et al.* PKLot—a robust dataset for parking lot classification[J]. *Expert Systems with Applications*, 2015, 42(11): 4937–4949. doi: [10.1016/j.eswa.2015.02.009](https://doi.org/10.1016/j.eswa.2015.02.009).
- [9] AMATO G, CARRARA F, FALCHI F, *et al.* Car parking occupancy detection using smart camera networks and deep learning[C]. 2016 IEEE Symposium on Computers and Communication, Messina, Italy, 2016: 1212–1217. doi: [10.1109/iscc.2016.7543901](https://doi.org/10.1109/iscc.2016.7543901).
- [10] BUADES A, COLL B, and MOREL J M. A non-local algorithm for image denoising[C]. 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, San Diego, USA, 2005: 60–65. doi: [10.1109/cvpr.2005.38](https://doi.org/10.1109/cvpr.2005.38).
- [11] VASWANI A, SHAZEER N, PARMAR N, *et al.* Attention is all you need[C]. The 31st Conference on Neural Information Processing Systems, Long Beach, USA, 2017: 5998–6008.
- [12] WANG Xiaolong, GIRSHICK R, GUPTA A, *et al.* Non-local neural networks[C]. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, USA, 2018: 7794–7803. doi: [10.1109/cvpr.2018.00813](https://doi.org/10.1109/cvpr.2018.00813).
- [13] HE Kaiming, ZHANG Xiangyu, REN Shaoqing, *et al.* Deep residual learning for image recognition[C]. 2016 IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, USA, 2016: 770–778. doi: [10.1109/cvpr.2016.90](https://doi.org/10.1109/cvpr.2016.90).
- [14] AMATO G, CARRARA F, FALCHI F, *et al.* Deep learning for decentralized parking lot occupancy detection[J]. *Expert Systems with Applications*, 2017, 72: 327–334. doi: [10.1016/j.eswa.2016.10.055](https://doi.org/10.1016/j.eswa.2016.10.055).
- [15] KRIZHEVSKY A, SUTSKEVER I, and HINTON G E. ImageNet classification with deep convolutional neural networks[C]. The 25th International Conference on Neural Information Processing Systems, Red Hook, USA, 2012: 1097–1105.
- [16] NURULLAYEV S and LEE S W. Generalized parking occupancy analysis based on dilated convolutional neural network[J]. *Sensors*, 2019, 19(2): 277. doi: [10.3390/s19020277](https://doi.org/10.3390/s19020277).
- [17] OJANSIVU V and HEIKKILÄ J. Blur insensitive texture classification using local phase quantization[C]. The 3rd International Conference on Image and Signal Processing, Cherbourg-Octeville, France, 2008: 236–243. doi: [10.1007/978-3-540-69905-7_27](https://doi.org/10.1007/978-3-540-69905-7_27).
- [18] RAHTU E, HEIKKILÄ J, OJANSIVU V, *et al.* Local phase quantization for blur-insensitive image analysis[J]. *Image and Vision Computing*, 2012, 30(8): 501–512. doi: [10.1016/j.imavis.2012.04.001](https://doi.org/10.1016/j.imavis.2012.04.001).
- [19] OJALA T, PIETIKAINEN M, and MAENPAA T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2002, 24(7): 971–987. doi: [10.1109/tpami.2002.1017623](https://doi.org/10.1109/tpami.2002.1017623).
- 申铨京: 男, 1958年生, 博士, 教授, 研究方向为图像处理与模式识别、多媒体信息安全、智能控制技术。
- 沈 哲: 男, 1995年生, 硕士生, 研究方向为图像处理与模式识别。
- 黄永平: 男, 1964年生, 博士, 副教授, 研究方向为图像处理与模式识别、智能控制与嵌入式系统。
- 王 玉: 男, 1983年生, 博士, 副教授, 研究方向为图像处理与模式识别、多媒体信息技术。

责任编辑: 陈 倩