



研究与开发

基于优化卷积神经网络的车辆特征识别算法研究

陈暄¹, 吴吉义^{2,3}

(1. 浙江工业职业技术学院, 浙江 绍兴 312000;

2. 浙江省人工智能学会, 浙江 杭州 310027;

3. 浙江大学智能教育研究中心, 浙江 杭州 310027)

摘要: 针对道路场景图像中不同距离目标车辆特征存在识别效果弱、精度低的问题, 提出一种基于优化卷积神经网络的车辆特征识别算法。首先, 采用基于 PAN 模型的多尺度输入获取不同距离的目标车辆特征; 其次, 在卷积神经网络结构中加入多池化、BN 层和 Leaky ReLU 激活函数改进网络模型的性能, 通过引入混合注意力机制, 集中关注车辆图像中的重要特征和区域, 从而增强了网络模型的泛化能力; 最后, 通过构建多层次卷积神经网络结构完成对车辆的特征效果识别。仿真实验结果表明, 在单一场景的 BIT-Vehicle 数据库中, 本文算法相比 CNN、R-CNN、ABC-CNN、Faster R-CNN、AlexNet、VGG16 和 YOLOV8 在单一目标和多目标识别率方面分别提高了 16.75%、10.9%、4%、3.7%、2.46%、1.3%、1% 和 17.8%、10.5%、2.5%、3.8%、2.7%、1.1%、1.3%, 在复杂场景的 UA-DETRAC 数据库中, 本文算法相比其他算法在不同距离目标车辆识别中获得了更加精确的效果。

关键词: 车辆识别; 卷积神经网络; 多尺度输入

中图分类号: TP391.14

文献标志码: A

doi: 10.11959/j.issn.1000-0801.2023188

Research on vehicle feature recognition algorithm based on optimized convolutional neural network

CHEN Xuan¹, WU Jiyi^{2,3}

1. Zhejiang Industry Polytechnic College, Shaoxing 312000, China

2. Zhejiang Federation of Artificial Intelligence, Hangzhou 310027, China

3. Intelligent Education Research Center, Zhejiang University, Hangzhou 310027, China

Abstract: To address the issue of weak identification and low accuracy in recognizing features of target vehicles at different distances in road scene images, a vehicle feature recognition algorithm based on optimized convolutional neural network (CNN) was proposed. Firstly, a multi-scale input based on the PAN model was employed to capture

收稿日期: 2023-06-25; 修回日期: 2023-10-10

通信作者: 吴吉义, cloudLab@139.com

基金项目: 国家自然科学基金资助项目 (No.61702151, No.61702320, No.61772334); 国家重点研发计划项目 (No.2018YFB1003800); 浙江省哲学社会科学规划课题 (No.23NDJC369YB)

Foundation Items: The National Natural Science Foundation of China (No.61702151, No.61702320, No.61772334), The National Key Research and Development Program of China (No.2018YFB1003800), Zhejiang Provincial Philosophy and Social Sciences Planning Project (No.23NDJC369YB)



target vehicle features at varying distances. Subsequently, improvements were made to the network model by incorporating multi-pool, batch normalization (BN) layers, and Leaky ReLU activation functions within the CNN architecture. Furthermore, the generalization ability of the network model was enhanced by introducing a hybrid attention mechanism that focuses on important features and regions in the vehicle image. Lastly, a multi-level CNN structure was constructed to achieve feature recognition for vehicles. Simulation experiment results conducted on the BIT-Vehicle database within a single scene show the proposed algorithm's significant enhancements in single-object and multi-object recognition rates compared to CNN, R-CNN, ABC-CNN, Faster R-CNN, AlexNet, VGG16, and YOLOV8. Specifically, improvements of 16.75%, 10.9%, 4%, 3.7%, 2.46%, 1.3%, and 1% in single-object recognition, as well as 17.8%, 10.5%, 2.5%, 3.8%, 2.7%, 1.1%, and 1.3% in multi-object recognition, have been demonstrated by the proposed algorithm, respectively. Over the more complex UA-DETRAC datasets, more precise results have been also achieved by the proposed algorithm in recognizing target vehicles at various distances compared to other algorithms.

Key words: vehicle recognition, convolutional neural network, multi-scale input

0 引言

随着信息技术和车联网技术的快速发展,车辆自动驾驶系统和交通管理系统都需要高效准确全天候的车辆识别系统进行辅助和管理。而卷积神经网络(convolutional neural network, CNN)凭借出色的性能在车辆特征图像识别领域获得巨大的成功^[1-2]。学者以 CNN 为主要技术手段在车辆识别方面展开深入的研究。有的学者关注 CNN 应用结果方面的研究,如文献[3]提出基于背景分割法卷积神经网络的车辆识别方法,它将输入视频分割成与前景或背景相对应的区域降低 CNN 的计算量并提高识别精度;文献[4]提出基于 CNN 的车牌识别系统,该系统在多行、倾斜和多字体车牌方面识别率达到了 98.13%;文献[5]提出基于 CNN 的自动驾驶车辆的多模态融合识别系统,在高精度和适应性方面具有不错的效果;文献[6]提出无须标注车辆样本的 CNN 车辆检测和分类方法,该类方法的检测准确率约为 98.5%。以上方法获得了较好的识别率,但是如何降低识别模型的复杂度,优化识别时间依然是当前应用领域的主要问题。有的学者针对 CNN 自身性能提升开展研究,如采用粒子群算法^[7-8]、人工蜂群算法^[9]、鲸鱼优化算法^[10]等对 CNN 的

参数进行优化,虽然能够提高网络性能,但这些元启发式算法自身性能有限,增加了模型计算量。还有的学者针对快速区域卷积神经网络(Faster R-CNN)开展研究,如文献[11]对 Faster R-CNN 模型在车辆特征识别方面的应用进行了综述,肯定了该模型的识别效果;文献[12]提出在不同的天气环境中使用该模型进行车辆目标识别检测,实验结果表明该算法在真实环境中也具有较为不错的效果;文献[13]采用该模型进行车辆特征识别,通过 VTTI 数据库展示较好的识别效果,但消耗了大量的计算时间;文献[14]采用 Faster R-CNN 模型解决部分外形相似的车辆之间的误检问题,该算法具有较高的精确度、较小的误检率和更好的鲁棒性;文献[15]使用 Faster R-CNN 模型对车辆目标图像进行多尺度的特征提取,实验结果表明能够获得较为精确的识别效果但需要较多的时间。以上方法的识别率均获得提升,但依然存在训练复杂、无法适应不同尺寸目标的问题。

基于此,提出一种基于优化卷积神经网络的车辆特征识别算法,它使用多尺度图像处理技术获取不同尺度的车辆图像特征,利用多池化(multi-pool)、批归一化(batch normalization, BN)层和 Leaky ReLU 激活函数优化网络结构,引入混

合注意力机制提高模型的泛化效果,通过仿真实验验证不同场景的车辆特征识别性能。

1 卷积神经网络

传统神经网络在图像处理过程中数据量较大,图像数字化处理无法保留原来的图像特征,使得图像识别率降低。CNN 凭借特有的权重共享、权重数量较少等特点能够有效地提升图像处理的识别率而受到广泛应用。它主要包括输入层、卷积层、池化层、全连接层和输出层。本文仅对中间 3 层进行简单介绍。

(1) 卷积层

卷积层的作用将原始图像输入 CNN 中进行图像特征提取。在卷积层中设置多种卷积核,按照设定的卷积核大小对原始图像进行窗口滑动提取图像特征,再利用激活函数对输出结果进行非线性变化,从而获得非线性特征图。

(2) 池化层

池化层的作用是对卷积层输出的特征图像进行降维,减小特征图的空间尺寸,降低信息的数据维度,从而减少计算量。

(3) 全连接层

全连接层的作用是将一个特征空间线性变换到另一个特征空间中。它将经过卷积层、池化层处理获得的特征整合在一起,然后根据非线性激活函数自由调节学习过程,从而完成对目标的识别分类。

2 优化卷积神经网络

针对现有 CNN 在车辆特征识别中存在不同距离目标车辆识别精度低、图像鲁棒性弱的问题,从 4 个方面进行优化,首先通过多尺度图像处理获得精确的车辆图像特征,其次利用改进的卷积神经网络结构提升模型识别能力,接着引入混合注意力机制加强模型泛化能力,最后构建多层次卷积神经网络结构用于车辆特征识别。

2.1 多尺度图像处理

传统的 CNN 主要通过逐层抽象方式提取目标特征,其中感受视野范围是获取目标特征的关键。但是在复杂道路场景下的车辆识别中,如果视野太小,则只能观察到局部特征,存在丢失车辆关键信息的可能;反之,可能会观察到周围场景很多无效的信息,增大图像特征提取的难度。采用多尺度图像处理能够使得网络对不同尺度的物体的体积、纹理、结构等进行感知,极端天气、街道建筑物、光线等复杂场景会增大车辆图像提取难度。采用多尺度图像处理能够对不同距离的车辆目标实现较好的识别效果。本文使用路径聚合网络(path aggregation network, PAN)模型^[16]获取车辆图像特征信息,利用该模型自下而上的路径技术增加整个特征层次结构,从而获得不同尺度的图像大小特征,提高模型的鲁棒性。PAN 模型如图 1 所示。

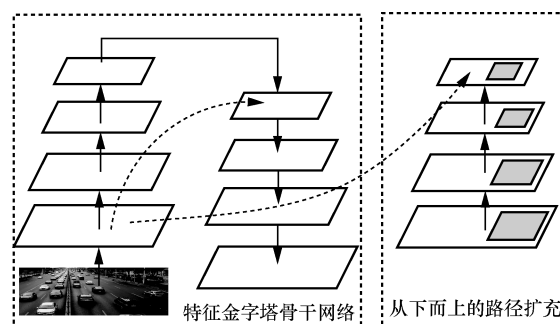


图 1 PAN 模型

2.2 改进的卷积神经网络

为了更好地发挥 CNN 在车辆识别中的效果。本文在卷积层依次加入多池化、BN 层和 Leaky ReLU 激活函数,通过这些措施提升 CNN 的性能和识别准确率。改进的卷积神经网络结构如图 2 所示。

(1) 采用多池化结构

在车辆特征识别中,为了更好地获得目标特征及感知能力,本文借助了 YOLOV3 中空间金字塔池化(spatial pyramid pooling, SPP)架

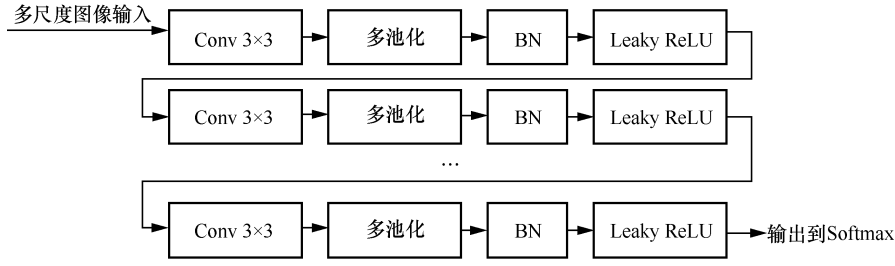


图2 改进的卷积神经网络结构

构的设计思想。在 SPP 架构的 4 个分支结构中，第一个分支保留原始输入的信息，即不进行池化或降采样，第二、第三、第四分支池化核分别设定为 5×5 、 9×9 及 13×13 的最大池化，经过池化处理后的特征图像尺寸和深度保持不变。通过池化层的设置方式能够避免 CNN 中需要固定尺寸的滤波器，对不同任意尺度的特征进行采样和池化以获得不同尺度目标的信息。结合复杂场景中的车辆特征信息，本文利用 SPP 架构对多尺度的输入图像进行处理，为当前的网络模型提供更高的灵活性和鲁棒性。

(2) 引入 Leaky ReLU 激活函数

传统 CNN 一般使用 Sigmoid 或者 ReLU 激活函数，但是这些函数在训练过程中容易出现神经元“死亡现象”，导致在反向传播中出现梯度消失的情况。因此使用 Leaky ReLU 激活函数 $f(x)$ ，它是一种修正线性单元激活函数的变体，它解决了传统激活函数的零梯度问题。

$$f(x) = \max\{ax, x\} \quad (1)$$

其中， a 为固定斜率，当 $x \leq 0$ 时，Leaky ReLU 激活函数可以进行反向传播，有效解决 ReLU 激活函数出现的神经元“死亡现象”和梯度消失问题。但由于 Leaky ReLU 激活函数的值域缺乏固定区间，因此需要对 Leaky ReLU 激活函数所得结果进行归一化处理。

$$b_{x,y}^i = \frac{a_{x,y}^i}{\left(k + \alpha \sum_{j=\max\{0, i-n/2\}}^{\min\{N-1, i+n/2\}} (\alpha_{x,y}^j)^2 \right)^\beta} \quad (2)$$

其中， $b_{x,y}^i$ 为归一化后的激活函数值， $a_{x,y}^i$ 表示第 i 个核位置 (x, y) 经过 Leaky ReLU 激活函数处理后神经元的输出， n 为特征图像的数目， N 为特征图像输出通道的总数， α 为缩放因子， β 为指数项。

(3) 加入 BN 层

为了更好地在模型后期获得分类结果，在 CNN 中加入 BN 层用来处理训练过程中特征数据分布发生改变的问题。它通过对卷积层的输入进行归一化，使输入数据的均值接近 0、标准差接近 1，从而减少内部协变量偏移的影响，使得网络更容易训练，有助于梯度的传播和提高反向传播过程中的数值稳定性，从而加快了模型的收敛速度。

2.3 混合注意力机制

传统的 CNN 主要以大量的训练样本为基础，但训练数据较少导致无法有效地提取样本特征，造成网络性能下降。为了提高输入特征数据的感知和理解能力，增强网络对重要特征和区域的关注，将通道注意力机制和空间注意力机制组成混合注意力机制引入车辆识别中。

(1) 通道注意力机制

通道注意力机制主要利用特征通道间的关系生成通道注意图，并为显著目标分配更大的权重，从而提高模型的性能。其过程是对输入图像特征进行最大池化和平均池化操作，并通过聚合获得特征映射的空间信息。

$$M_c(F) = \sigma(M_{Lp}(P_{avg}(F)) \oplus \otimes M_{Lp}(P_{max}(F))) \quad (3)$$

其中, F 为输入特征; σ 为 sigmoid 函数; M_{LP} 为共享全连接层, 由卷积层和 Leaky ReLU 激活函数组成; $\oplus \otimes$ 为拼接操作; P_{avg} 和 P_{max} 分别表示平均池化和最大池化操作; $M_c(F)$ 为该模块的输出。

(2) 空间注意力机制

空间注意力机制是一种用于图像或特征图处理的注意力机制, 它的主要目标是加权关注图像或特征图中的重要区域, 以便更有效地处理视觉信息, 它通过计算每一个空间位置的注意力权重, 使得网络聚焦在图像特定区域重要部分。

$$M_s(F) = \sigma(f^{7 \times 7}(P_{avg}(F)) \oplus \otimes M_{LP}(P_{max}(F))) \quad (4)$$

其中, $f^{7 \times 7}$ 为 7×7 大小的卷积核, $\oplus \otimes$ 为拼接操作, $M_s(F)$ 为该模块的输出。

2.4 多层次卷积神经网络结构

结合以上对卷积神经网络的改进措施, 本文提出一种多层次卷积神经网络结构, 如图 3 所示。

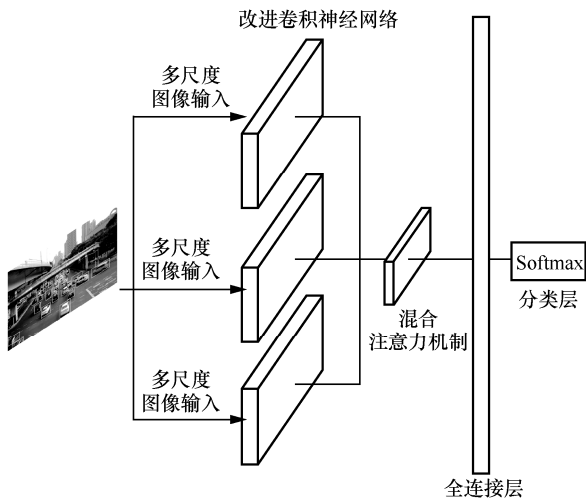


图3 多层次卷积神经网络结构

图3所示的多层次卷积神经网络由3个相对独立的改进卷积神经网络、1个混合注意力机制、1个全连接层和1个分类层组成。第一个改进卷积

神经网络感受野的大小为 $67 \text{ pixel} \times 67 \text{ pixel}$, 采用 3×3 、 5×5 、 7×7 的池化核; 第二个改进卷积神经网络感受野的大小为 $131 \text{ pixel} \times 131 \text{ pixel}$, 采用 5×5 、 7×7 、 9×9 池化核; 第三个改进卷积神经网络感受野的大小为 $195 \text{ pixel} \times 195 \text{ pixel}$, 采用 7×7 、 9×9 、 11×11 池化核。3个改进卷积神经网络结构与混合注意力机制的融合关系采用文献[17]的设计思想进行构建, 表达式如下:

$$O_{att} = \sum_{i=1}^3 \frac{\exp[\tanh(W_{att,i} O_{CNN,i}) + b_{att,i}]}{\sum_j \exp[\tanh(W_{att,j} O_{CNN,j}) + b_{att,j}]} \times O_{CNN,i} \quad (5)$$

其中, $W_{att,j}$ 为混合注意力层的权值, $b_{att,j}$ 为混合注意力层的偏移值, $O_{CNN,i}$ 为 CNN 的输出结果 ($i \in [1, 2, 3]$), $\frac{\exp[]}{\sum_j \exp[]}$ 为每一个 CNN 层的权重。

tanh 激活函数作用是将每一个网络输出值 $W \times O + b$ 控制在 $[-1, 1]$ 。

3 仿真实验与分析

为了更好地验证本文算法的识别效果, 搭建了基于 TensorFlow 深度学习框架, 实验硬件方面采用酷睿 i5 的 CPU, 内存为 16 GB DDR, 硬盘为 1 TB, 软件环境为 Windows10 操作系统, 利用 Spyder 编译工具进行 Python 编程。对比算法为 CNN、区域卷积神经网络 (region-CNN, R-CNN)、人工蜂群优化卷积神经网络 (artificial bee colony-CNN, ABC-CNN) [18]、Faster R-CNN [19]、AlexNet、VGG16 和 YOLOV8。在单一和复杂两个场景下验证本文算法的性能。

3.1 本文算法性能

为了验证本文算法的性能, 将 8 种算法的精度进行了对比, 8 种算法的训练精度对比如图 4 所示。图 4 显示随着训练次数逐渐增加, 8 种算法的训练精度都呈现不同程度的上升趋势。当训



训练次数达到 400 时,本文算法率先趋近稳定并始终保持平缓状态。而 CNN 的训练精度在整个训练过程中一直处于上升趋势且具有较为明显的震荡;R-CNN 和 ABC-CNN 的精度上升趋势较为平缓,且出现了微弱的震荡;Faster R-CNN 的精度在训练次数达到 500 后才出现较为平缓的现象;AlexNet 和 VGG16 的精度在训练次数达到 450 后出现了较为平缓的现象;YOLOV8 大致与本文算法的精度结果相当,但稍微低于本文算法。因此,纵观整个训练过程,本文算法的精度优于其他 7 种算法,证明了本文算法具有较为明显的识别效果。

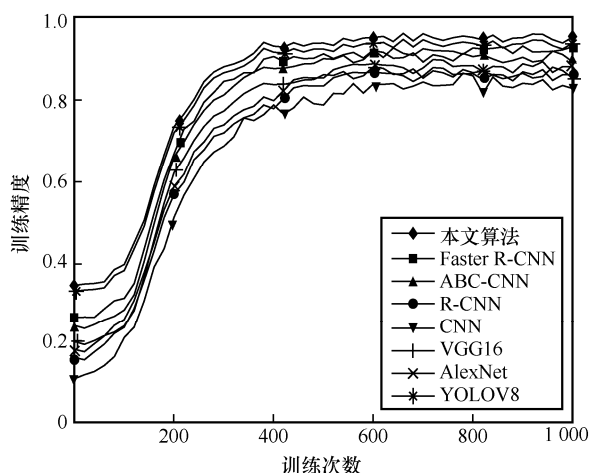


图 4 8 种算法的训练精度对比

3.2 单一场景车辆特征数据库对比

为了说明本文算法在单一车辆场景中的识别效果,使用文献[20]提供的 BIT-Vehicle 数据集中的车辆图片作为验证本文算法的图片库,该数据库一共包含轿车(sedan)、运动型多功能汽车(SUV)、卡车(truck)、公交车(bus)、小型公交车(microbus)和小型货车(minivan) 6 种不同类型的车辆。将该数据库中的 9 850 张图片分为两个部分,按照 6:4 分为训练数据集(5 910 张车辆样本图片)和测试数据集(3 490 张车辆图片)。部分 BIT-Vehicle 数据集样本如图 5 所示。

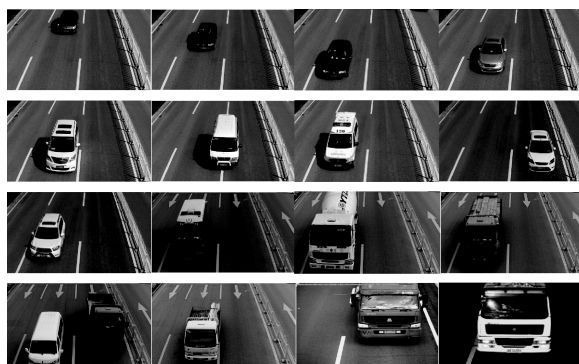


图 5 部分 BIT-Vehicle 数据集样本

8 种算法的单一车辆识别率对比见表 1。从表 1 发现这些算法对 6 种类型车辆的识别率都不相同,但是本文算法的优势较为明显。从这 6 种类型的车辆图片来看,周围场景的颜色对车辆具有一定的影响,而本文算法的多尺度图像处理降低了这些无效元素的影响,从而使得车辆的特征提取更加准确,它相比 CNN、R-CNN、ABC-RNN、Faster R-CNN、AlexNet、VGG16 和 YOLOV8 识别率分别提升了 16.75%、10.9%、4%、3.7%、2.46%、1.3%和 1%。8 种算法的多目标车辆识别率对比见表 2,本文算法相比单一车辆的识别率有所降低,主要是由于不同车辆处于同一个场景中使得特征提取过程受到了影响,同时外界光线、摄像角度、多目标车辆的重叠降低了识别率,但本文算法保持一定的优势,它相比 CNN、R-CNN、ABC-RNN、Faster R-CNN、AlexNet、VGG16 和 YOLOV8 识别率提升了 17.8%、10.5%、2.5%、3.8%、2.7%、1.5%和 1.1%。8 种算法对随机的 1 000 张数据库图片的识别时间对比见表 3。从表 3 可以发现,本文算法相比 CNN、R-CNN 具有明显优势,相比 ABC-CNN、Faster R-CNN、AlexNet 和 VGG16 算法优势较为一般,而和 YOLOV8 算法时间几乎相当,这说明了优化卷积神经网络由于结构优化导致识别时间有所增加,但总体效果影响不大。

召回率 R 和准确率 P 是衡量模型识别的重要方

表 1 8 种算法的单一车辆识别率对比

算法	sedan	SUV	truck	bus	minivan	microbus
本文算法	96.32%	95.61%	92.14%	94.32%	93.89%	94.78%
CNN	78.38%	82.41%	79.23%	81.92%	82.14%	81.62%
R-CNN	84.28%	87.29%	81.82%	85.23%	86.24%	86.19%
ABC-CNN	90.73%	89.31%	88.29%	93.02%	91.21%	92.18%
Faster R-CNN	91.21%	88.16%	89.27%	93.32%	91.82%	92.91%
AlexNet	92.78%	91.40%	90.75%	93.46%	92.04%	93.01%
VGG16	94.19%	93.39%	91.74%	93.52%	92.76%	93.85%
YOLOV8	95.12%	94.02%	91.84%	93.75%	92.79%	93.92%

表 2 8 种算法的多目标车辆识别率对比

算法	sedan	SUV	truck	bus	minivan	microbus
本文算法	92.13%	90.13%	90.23%	91.32%	90.41%	92.38%
CNN	76.18%	79.31%	75.73%	78.93%	78.34%	75.32%
R-CNN	81.72%	85.17%	78.92%	82.43%	83.14%	83.27%
ABC-CNN	89.34%	88.48%	87.91%	89.03%	88.71%	89.31%
Faster R-CNN	88.31%	89.27%	83.71%	88.92%	87.31%	88.93%
AlexNet	90.78%	89.40%	84.75%	89.47%	88.34%	89.17%
VGG16	91.24%	89.42%	87.32%	90.43%	89.04%	91.05%
YOLOV8	91.39%	90.01%	88.19%	90.47%	89.13%	91.38%

表 3 8 种算法对随机的 1 000 张数据库图片的识别时间对比 (单位: s)

算法	sedan	SUV	truck	bus	minivan	microbus
本文算法	12.183	13.238	11.147	11.611	12.523	12.416
CNN	15.982	15.298	14.327	14.317	15.925	15.291
R-CNN	14.925	14.198	13.782	12.893	14.197	14.981
ABC-CNN	13.891	13.278	12.982	12.892	13.183	13.261
Faster R-CNN	13.782	14.372	12.931	12.927	13.894	13.936
AlexNet	13.153	14.176	12.369	12.725	13.126	13.619
VGG16	13.013	13.739	12.319	12.074	13.015	12.946
YOLOV8	12.186	13.341	11.159	11.657	12.576	12.519

法, 本文使用 4 种条件下的本文算法计算 BIT-Vehicle 数据集中 6 类不同车辆的准确-召回率 PR 进行对比。

其中召回率 R 和准确率 P 的计算式如下:

$$R = \frac{T_p}{T_p + F_n} \quad (6)$$

$$P = \frac{T_p}{T_p + F_p} \quad (7)$$

其中, T_p 和 F_p 分别表示正样本和负样本被正确识别的数目, F_n 表示正样本被误判为负样本的数目。不同条件下本文算法的 6 种车型 PR 如图 6 所示。

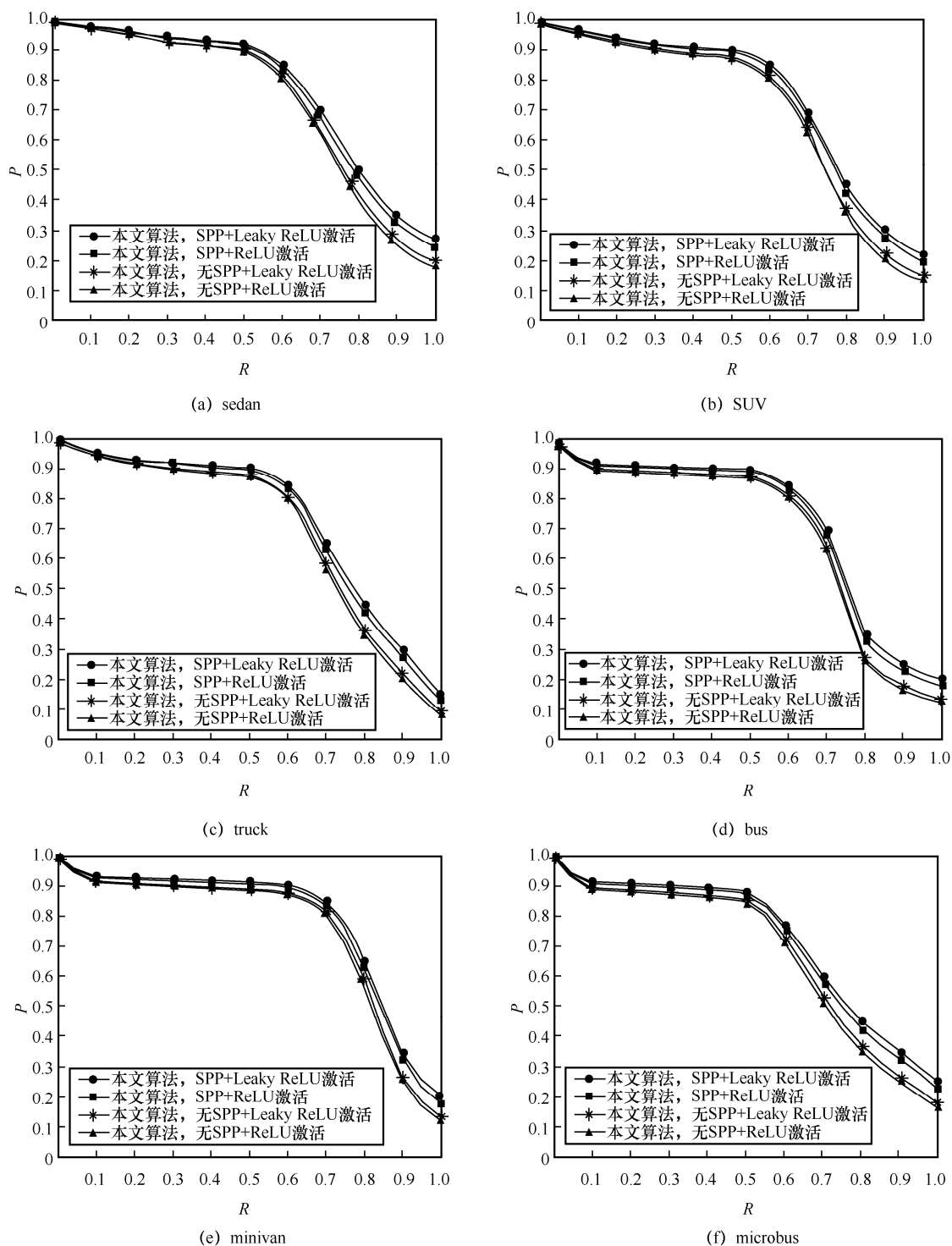


图6 不同条件下本文算法的6种车型PR

从6种车型来看,虽然4种条件下的PR都呈现下降趋势,但是SPP+Leaky ReLU激活函数的PR具有较为明显的优势,这说明增加SPP和Leaky ReLU函数的效果明显。8种算法的6种车型PR

如图7所示。从图7发现随着 R 值逐渐增大,8种算法的 P 值总体呈现逐渐下降的趋势,而本文算法的 R 值在 $[0,0.6]$ 内较为平缓,相比其他算法 R 值平缓范围为 $[0,0.5]$ 具有更加稳定的效果。

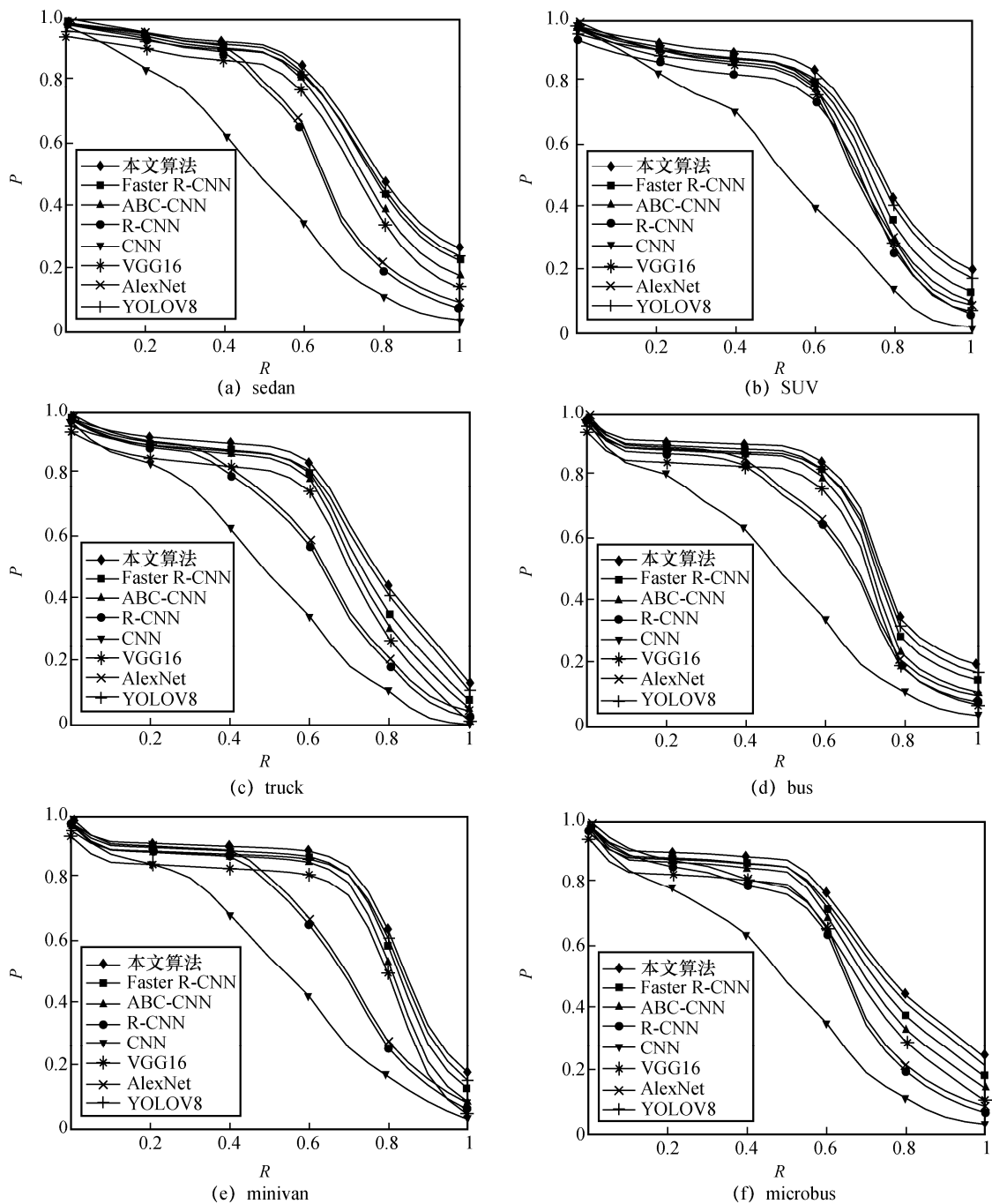


图7 8种算法的6种车型PR

3.3 复杂场景下的车辆目标验证

为了验证复杂场景下的识别效果,选取交通数据集 UA-DETRAC^[21]作为本文模型的训练测试样本,该数据集中包含了基于不同路侧视角、不同时段和不同车辆拥堵度的车辆图像。它满足了场景复杂、数据规模大的要求。本文选取白天正

面和夜晚侧面作为识别场景进行研究。8种算法的白天和夜晚识别效果分别如图8、图9所示。

图8和图9分别展示了白天和晚间场景下的8种算法对车辆目标的识别效果,方框代表不同算法的识别效果。在白天道路正面图像中,所有算法基本上都能够检测出车辆特征,但是本文算



法相比其他算法能够明显获取不同位置的车辆特征,降低了漏检率和误检率,特别是道路图像中的密集车辆,本文算法的识别优势比较明显,如图8(h)中远处的目标依然能够被识别,验证了本文模型性能的优越性。在晚间道路侧向图像中,8种算法的识别效果都呈现了不同程度的下降。从总体上看,本文算法在识别方面具有明显的优势,当目标与周围的场景存在较小的差异,如图9(h)中公共汽车左上角出现的5辆小轿车时,采用本文算法依然能够获得较好的识别结果。

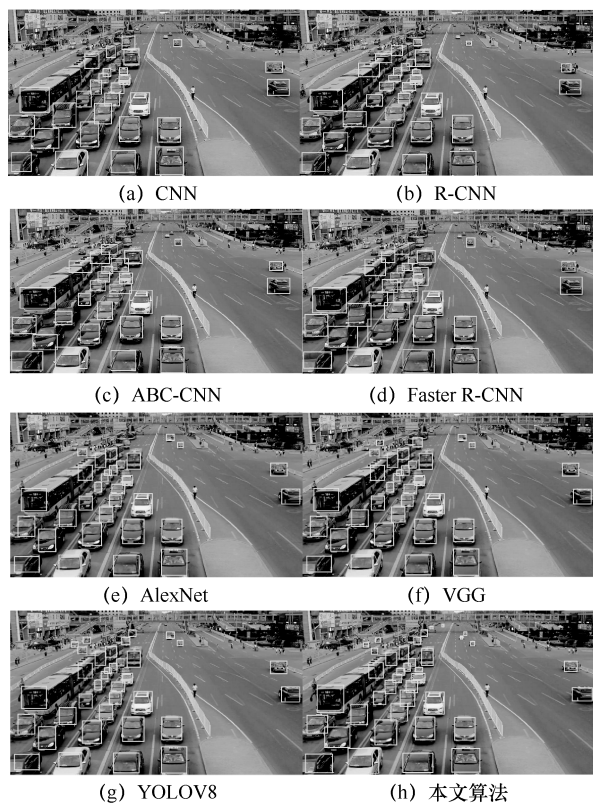


图8 8种算法的白天识别效果

综上所述,通过理论阐述和仿真实验说明了本文算法能有效提升车辆特征识别的效果。该模型能够捕获图像中不同距离目标车辆的特征,降低交通场景对图像特征的干扰,同时实现了对不同尺度的图像特征进行采样和池化,有效避免了神经元的消失和梯度为0的缺点,提高了模型泛化能力,该网络相比普通的卷积神经网络,无论

在应用范围还是自身性能结构上都具有较好的优势。

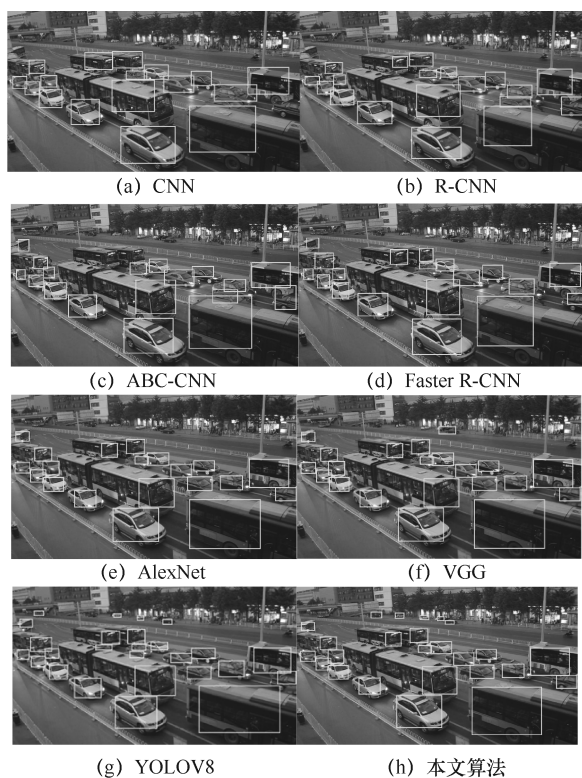


图9 8种算法的夜晚识别效果

4 结束语

本文提出了一种优化的卷积神经网络算法。该算法结合多尺度图像处理、多池化结构、BN层、Leaky ReLU 激活函数、混合注意力机制等多种方法提高道路场景图像中不同距离目标车辆特征的识别精度。目前,有关CNN在下一步研究主要体现在:模型结构设计、跨域和跨模态应用、弱监督和无监督学习、可解释性和鲁棒性、轻量级和移动端应用、联合学习和多任务学习等方面。笔者将在模型结构设计和多任务学习方面继续开展研究。

参考文献:

- [1] ZHANG J P, WANG F Y, WANG K F, et al. Data-driven intelligent transportation systems: a survey[J]. IEEE Transactions on Intelligent Transportation Systems, 2011, 12(4): 1624-1639.

- [2] COLLINS R T, LIPTON A J, KANADE T. Introduction to the special section on video surveillance[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000, 22(8): 745-746.
- [3] CHAROUH Z, EZZOUHRI A, GHOGHO M, et al. A resource-efficient CNN-based method for moving vehicle detection[J]. Sensors, 2022, 22(3): 1193.
- [4] RANJITHKUMAR S, CHENTHUR PANDIAN S. Automatic license plate recognition system for vehicles using a CNN[J]. Computers, Materials & Continua, 2022, 71(1): 35-50.
- [5] SHI R W, YANG S C, CHEN Y Y, et al. CNN-Transformer for visual-tactile fusion applied in road recognition of autonomous vehicles[J]. Pattern Recognition Letters, 2023, 166: 200-208.
- [6] SATYANARAYANA G S R, DESHMUKH P, DAS S K. Vehicle detection and classification with spatio-temporal information obtained from CNN[J]. Displays, 2022, 75: 102294.
- [7] SOON F C, KHAW H Y, CHUAH J H, et al. Hyper-parameters optimisation of deep CNN architecture for vehicle logo recognition[J]. IET Intelligent Transport Systems, 2018, 12(8): 939-946.
- [8] 王上, 唐欢容. 一种基于混合粒子群优化算法的深度卷积神经网络架构搜索方法[J]. 计算机应用研究, 2023, 40(7): 2019-2024.
- WANG S, TANG H R. Deep convolutional neural architecture search method based on hybrid particle swarm optimization algorithm[J]. Application Research of Computers, 2023, 40(7): 2019-2024.
- [9] KIYMAÇ E, KAYA Y. A novel automated CNN arrhythmia classifier with memory-enhanced artificial hummingbird algorithm[J]. Expert Systems With Applications, 2023(213): 119162.
- [10] GHASEMI DAREHNAEI Z, SHOKOUHIFAR M, YAZDAN-JOUEI H, et al. SI-EDTL: swarm intelligence ensemble deep transfer learning for multiple vehicle detection in UAV images[J]. Concurrency and Computation: Practice and Experience, 2022, 34(5): e6726.
- [11] MAITY M, BANERJEE S, SINHA CHAUDHURI S. Faster R-CNN and YOLO based vehicle detection: a survey[C]//Proceedings of 2021 5th International Conference on Computing Methodologies and Communication (ICCMC). Piscataway: IEEE Press, 2021: 1442-1447.
- [12] GHOSH R. On-road vehicle detection in varying weather conditions using Faster R-CNN with several region proposal networks[J]. Multimedia Tools and Applications, 2021, 80(17): 25985-25999.
- [13] HSU S C, HUANG C L, CHUANG C H. Vehicle detection using simplified Fast R-CNN[C]//Proceedings of 2018 International Workshop on Advanced Image Technology (IWAIT). Piscataway: IEEE Press, 2018: 1-3.
- [14] 宁俊, 王年, 朱明. 基于改进 Faster R-CNN 的车辆类型识别算法[J]. 安徽大学学报(自然科学版), 2021, 45(3): 26-33.
- NING J, WANG N, ZHU M. Vehicle type recognition algorithm based on the improved Faster R-CNN[J]. Journal of Anhui University (Natural Sciences), 2021, 45(3): 26-33.
- [15] LUO J Q, FANG H S, SHAO F M, et al. Multi-scale traffic vehicle detection based on Faster R-CNN with NAS optimization and feature enrichment[J]. Defence Technology, 2021, 17(4): 1542-1554.
- [16] LIU S, QI L, QIN H F, et al. Path aggregation network for instance segmentation[C]//Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2018: 8759-8768.
- [17] GAO X, WANG J F, ZHOU M Z. The research of resource allocation method based on GCN-LSTM in 5G network[J]. IEEE Communications Letters, 2023, 27(3): 926-930.
- [18] ERKAN U, TOKTAS A, USTUN D. Hyperparameter optimization of deep CNN classifier for plant species identification using artificial bee colony algorithm[J]. Journal of Ambient Intelligence and Humanized Computing, 2023, 14(7): 8827-8838.
- [19] ZHANG X L, CUI J, LIU H J, et al. Weed identification in soybean seedling stage based on optimized Faster R-CNN algorithm[J]. Agriculture, 2023, 13(1): 175.
- [20] DONG Z, WU Y W, PEI M T, et al. Vehicle type classification using a semisupervised convolutional neural network[J]. IEEE Transactions on Intelligent Transportation Systems, 2015, 16(4): 2247-2256.
- [21] WEN L Y, DU D W, CAI Z W, et al. UA-DETRAC: a new benchmark and protocol for multi-object detection and tracking[J]. Computer Vision and Image Understanding, 2020(193): 102907.

[作者简介]



陈暄 (1979—), 男, 浙江工业职业技术学院副教授, 主要研究方向为云计算、人工智能。



吴吉义 (1980—), 男, 博士, 浙江大学高级工程师, 主要研究方向为服务计算、人工智能。