

DOI: 10.13382/j.jemi.B2104495

结合注意力机制和密集连接网络的车辆检测方法^{*}

梁继然^{1,2} 陈 壮¹ 董国军³ 陈 琦¹ 许延雷¹

(1. 天津大学微电子学院 天津 300072; 2. 天津市成像与感知微电子技术重点实验室 天津 300072;
3. 天津七一二通信广播有限公司 天津 300462)

摘 要:为提高算法对车辆检测的准确性,解决原有算法在复杂交通场景下对车辆检测效果不佳的问题,提出一种基于注意力机制和改进密集连接网络结构的车辆检测方法。首先在过渡层中使用 SoftPool 整合密集块之间的特征信息;其次通过轻量化通道注意力机制加强有效通道特征的表达,将其作为 Darknet-53 的深层特征提取层;引入 CIOU 损失作为模型的边界框位置预测损失项,使用深度可分离卷积缩减模型体积;与原算法相比 mAP 值提高 2.6%,模型体积缩减为原来的 42%,实验证明本算法在复杂交通场景下具有良好的检测性能。

关键词: 车辆检测;密集连接网络;注意力机制;SoftPool

中图分类号: TP391.4 **文献标识码:** A **国家标准学科分类代码:** 520.6040

Vehicle detection method combining attention mechanism and dense connection network

Liang Jiran^{1,2} Chen Zhuang¹ Dong Guojun³ Chen Qi¹ Xu Yanlei¹

(1. School of Microelectronics Tianjin University, Tianjin 300072, China;
2. Tianjin Key Laboratory of Imaging and Sensing Microelectronics Technology, Tianjin 300072, China;
3. Tianjin 712 Communication and Broadcasting Corporation, Tianjin 300462, China)

Abstract: To improve the accuracy of the algorithm for vehicle detection and solve the problem that the original algorithm is not effective in the complex traffic scene, a vehicle detection method based on attention mechanism and improved densely connection network structure was proposed. Firstly, SoftPool was used in the transition layer to consolidate the characteristic information between the dense blocks. Secondly, the expression of effective channel features was enhanced by the lightweight channel attention mechanism, it was used as the deep feature extraction layer of Darknet-53. The CIOU loss was used as the prediction loss term of the bounding box position of the model, and reduce the model volume using deep separable convolution. Compared with the original algorithm, the mAP value is increased by 2.6%, and the model volume is reduced to 42%. Experimental results show that the algorithm has good detection performance in complex traffic scene.

Keywords: vehicle detection; densely connected network; attention mechanism; SoftPool

0 引 言

近年来,随着智慧城市的不断发展,智能化交通系统逐渐成为未来城市交通系统新的发展方向,车辆检测是智能交通系统的重要组成部分,在车流量统计、违章停车、自动驾驶等方面具有很大的应用潜力^[1]。但是车辆

检测易受障碍物遮挡、尺度变化、光照强度等影响,造成错检和漏检现象,车辆检测算法的鲁棒性亟待提高^[2]。随着深度学习理论的发展,极大地提高了卷积神经网络(convolutional neural network, CNN)的特征提取能力,与传统方法相比,卷积神经网络自动学习、抽取蕴含大量信息的图像特征,在视觉图像检测方面取得了突破性进展^[3]。其主要分为基于候选区域的双阶段(two-stage)目

标检测算法和基于回归的单阶段(one-stage)目标检测算法^[4-5]。传统目标检测算法需要人为设计特征算子,难以适应不同的检测场景,泛化能力较弱,基于卷积神经网络的目标检测方法具有强大的特征提取能力和学习能力,克服了人为设计的特征算子泛化能力弱和鲁棒性差的问题^[6]。近年来,随着众多研究者们对深度学习的深入研究,基于深度学习的图像处理方法已经在文本检测、农业植被检测、车辆检测等领域当中都取得了令人满意的结果^[7]。

目前基于卷积神经网络的车辆检测算法主要研究方向有多尺度特征融合、聚合通道特征和数据增强等方法。Hao等^[8]提出多目标角池化层的神经网络结构,提取图像中车辆的局部信息和全局信息,可以有效检测出被阻塞车辆,但是检测速度有待提升。Hua等^[9]提出一种聚合通道特征的检测算法,提高了车辆检测算法的检测精度。Jin等^[10]在模型中加入深度通道来补充车辆细节,添加样本挖掘策略对模型进行训练,提高了识别准确率。上述工作通过不同角度的改进方式提高了车辆检测算法的检测性能。但在实际车辆检测任务中,车辆易受建筑物、树木等障碍物遮挡或由于距离较远导致车辆在图像中所占像素点较少,存在对小目标和被遮挡目标检测精度不高的问题。此外,在工程应用中车辆检测一般以监控设备为载体,大多数应用场景对于实时性要求较高且算法模型一般部署在计算能力有限的嵌入式设备之中,模型的运算量过大会制约算法的泛化部署能力^[11]。因此在增强网络检测能力的同时应减小网络复杂度,缩小模型体积,使性能优越的CNN算法模型能够部署在计算资源有限的移动终端和嵌入式设备上。

本文将一种轻量化注意力机制引入到密集连接网络^[12](DenseNet)作为模型的深层信息特征提取层,使有效信息得到更好的表达,使用SoftPool^[13]对密集连接网络中的特征图进行缩减,保留更多细粒度信息,将改进密集连接网络输出的特征图与浅层特征图融合对不同尺寸的目标进行预测输出,改进模型的损失函数加快模型的收敛,采用深度可分离卷积减少算法参数,缩小模型体积,最后通过实验对比验证本文所提方法的合理性和可行性。

1 网络结构

YOLO v3^[14]算法通过多尺度特征融合分别对大中小尺寸的目标进行预测输出。模型中浅层特征图包含待检测目标的位置信息,深层特征图包含待检测目标丰富的语义信息,通过深浅层特征融合,YOLO v3算法对大尺寸和中等尺寸的目标检测取得了良好的效果。但是在背景复杂,待检测对象较远,目标较小的情况下,对目标的检

测还存在一定的缺陷^[15]。YOLO v3算法通过8倍下采样的特征图对小目标进行预测输出,当待检测目标的像素小于 8×8 时,对小尺寸目标的检测准确率会有明显的下降。针对小目标检测任务,综合考虑在嵌入式设备中的泛化部署能力,提出了一种多尺度特征融合的具有注意力机制的密集连接网络模型,网络结构如图1所示,其中CA_Den表示改进后的具有注意力机制的密集连接模块。在神经网络训练过程中,生成的众多特征图中并不都是待检测目标的特征信息,其中也包含一些复杂的背景信息,对最终的目标预测结果几乎没有贡献甚至产生干扰,尤其小目标极易受到背景信息干扰,改进后的密集连接网络将特征信息输入到注意力机制层,重新分配不同通道的特征权重,将输入特征与生成的通道权重进行相乘,抑制无效的背景信息,增强待检测目标特征信息的表达;为在下采样过程中保留更多的细粒度信息,在密集连接网络的过渡层中采用SoftPool来缩小特征图的尺寸,由于在目标检测任务中深层次信息占据着更重要的位置,因此将原YOLO v3算法中的第4和第5个残差块改为具有通道注意力机制的密集连接方式,使其作为预测支路的高维度信息的输入。使用CIOU^[16]损失作为模型的边界框位置预测损失项,使预测框与目标真实框更加接近,加快收敛速度。最后,考虑到嵌入式设备的内存和计算资源有限以及实时性目标检测任务的需求,使用深度可分离卷积来实现网络中的标准卷积操作,以获得体积小、检测精度高的目标检测模型,更有利于在资源有限的环境中进行部署。

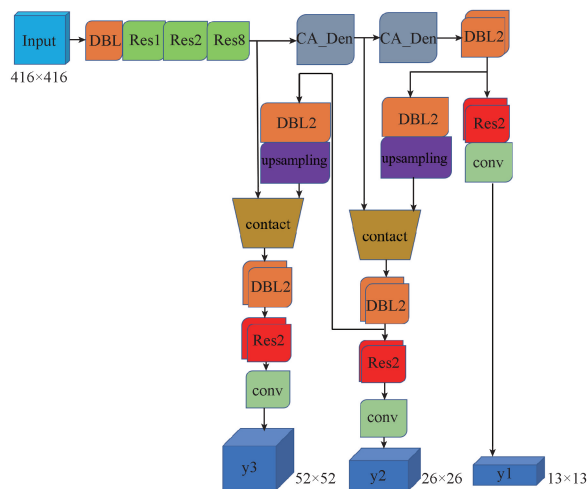


图1 网络结构

Fig. 1 Structure of model network

2 改进的密集连接网络结构

密集连接网络结构将提取到的特征图进行跨网络层的连接,在信息传递过程中实现特征复用,其结构包括密

集连接单元(DenseBlock)和转换层(transition layer)两部分,密集连接单元中每个特征图的大小相同,在通道维度上进行连接,相邻密集连接单元之间通过转换层进行连接,转换层使用步长为 2 的平均值池化缩小特征图尺寸。改进后的密集连接网络在转换层采用 SoftPool 来实现特征图尺寸的缩小,目的是为减少平均值池化过程中特征信息的损失,保留待检测目标的关键信息;其次在密集连接网络结构中引入通道注意力机制,为尽量不增加模型的复杂度,采用轻量化通道注意力机制 ECANet^[17],具体实现方式为利用全局平均值池化获得每个通道的特征信息,再使用一维卷积执行卷积操作,最后通过 Sigmoid 激活函数得到不同通道的权重,筛选出对最终目标检测贡献程度较大的关键信息,进而得到高质量的特征图。图 2 展示了改进的具有注意力机制的密集连接网络结构。

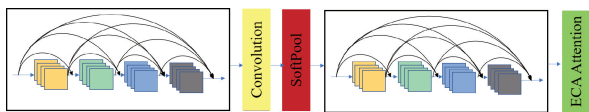


图 2 改进的密集连接网络结构

Fig. 2 Structure of improved dense connection

2.1 轻量化通道注意力机制

通道注意力机制是目标检测过程中抑制干扰信息的重要手段,通过对通道域赋予不同的权重进而输出高质量的特征响应图,例如 SENet 经过全局平均池化后使用两个全连接层和 Sigmoid 函数生成通道权重,但是使用全连接层使模型的参数量和计算量增加,且通道之间的关联性不高,为了不增加网络的运算负担,本研究引入 ECA 通道注意力机制,在不降低维度的情况下通过全局平均池化进行特征整合得到 $1 \times 1 \times C$ 的特征图信息,使用一维卷积实现跨通道交互,卷积核的大小 k 决定了通道交互的范围,通过 Sigmoid 函数生成新的权重参数,最后将生成的权重参数与原通道特征图相乘,进行不同权重的分配,相比于 SENet 等注意力机制,ECA 具有更加轻量化的结构,仅增加极少量模型的参数量和计算量,图 3 为 ECA 注意力机制的示意图。

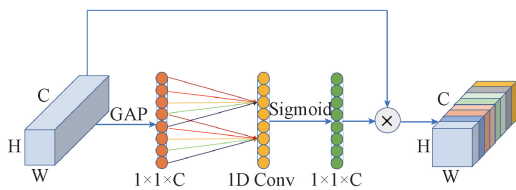


图 3 ECA 结构

Fig. 3 Structure of ECA

2.2 SoftPool

池化层在卷积神经网络中起到缩小特征图尺寸的作用,

在池化过程中应尽量减少特征信息的丢失,平均值池化能够反映该通道的整体情况,对于待检测大目标来说,目标占据了特征图的大部分,使用区域平均值池化可以在特征图中较好的体现通道的响应情况,但对于小目标来说,小目标只占据了特征图上很小的一部分,因此区域平均值池化并不能很好的保留小目标的特征信息。在真实交通场景下,由于尺度变化或由于障碍物干扰被部分遮挡,车辆在图像中所占像素较少,平均值池化不能很好的保留其有效信息,因此本文采用 SoftPool 对密集连接网络中的特征图进行尺寸的缩减,SoftPool 是可微的,以自然指数为基础结合 Softmax 加权方法保留输入的特征属性,在激活区域内,每个激活因子被分配不同的权重,通过非线性变换实现池化内核中激活因子的加权求和,如式(1)所示, a_i 表示激活因子, W_i 为权重因子,计算方式为激活因子的自然指数与区域内的所有激活因子的自然指数之和的比值, $\tilde{a}$ 表示 SoftPool 的输出值,由内核邻域内所有激活因子进行加权求和得到。

$$\tilde{a} = \sum_{i \in R} W_i * a_i = \sum_{i \in R} \frac{e^{a_i} * a_i}{\sum_{j \in R} e^{a_j}} \quad (1)$$

图 4 为 SoftPool 的双向传播过程,保留了小目标和部分被遮挡目标等所占像素较少的目标的特征信息。

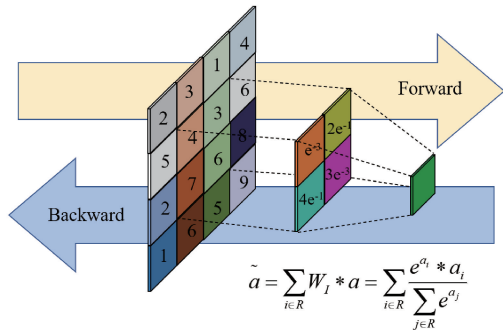


图 4 SoftPool 原理

Fig. 4 Principle of SoftPool

2.3 深度可分离卷积

由于特征提取网络中的卷积层在算法的参数量和计算量中占据了很大的比重,本研究利用深度可分离卷积降低参数冗余,缩小模型体积,使其更易于在嵌入式设备中部署。深度可分离卷积可以分为逐通道卷积和逐点卷积两个步骤,在逐通道卷积阶段使用和图像通道数相等的卷积核对相应通道进行卷积操作,逐点卷积阶段采用 1×1 卷积将相同空间位置的特征信息进行组合,如图 5 所示,假设有 $D_c \times D_c$ 尺寸的输入特征图,通道数为 M ,卷积核的尺寸为 $k \times k$,输出的特征图尺寸为 $D_f \times D_f$,通道数为 N ,则标准卷积的运算量可表示为 $k \times k \times M \times D_f \times D_f \times N$,深度可分离卷积的总计算量可由下式表示:

$k \times k \times M \times D_F \times D_F + M \times D_F \times D_F \times N$ (2)

与标准卷积相比,其压缩量可用下式表示:

$$P = \frac{k \times k \times M \times D_F \times D_F + M \times D_F \times D_F \times N}{k \times k \times M \times D_F \times D_F \times N} = \frac{1}{N} + \frac{1}{k^2}$$

(3)

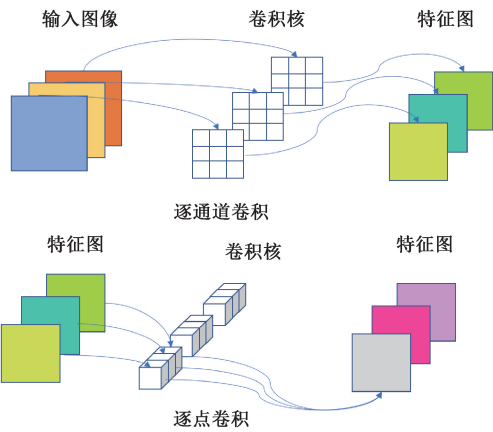


图5 深度可分离卷积

Fig.5 Depth separable convolution

2.4 损失函数

YOLO v3 采用交并比损失 (IOU) 函数作为目标框位置回归损失函数的重要参考依据,但是 IOU 损失函数有可能出现梯度为 0 和无法表达目标框与真实框之间的对齐情况的问题,本研究以 CIOU 损失作为模型的边界框预测损失项,CIOU 函数直接最小化预测框和目标框的归一化距离从而加速模型的收敛,通过增加检测框与目标框的长宽比一致性的衡量参数使损失函数朝着重叠区域增加的方向进行优化。CIOU 函数可用式(4)表示:

$$L_{CIOU} = 1 - IOU + \frac{\rho(b, b^{gt})}{c^2} + \alpha v$$

(4)

其中, $\rho(b, b^{gt})$ 表示检测框和真实框的中心点间的欧氏距离, c 表示最小闭合凹面的对角线距离, α 表示平衡比例的参数, v 表示衡量长宽比一致性的参数。 α 和 v 可分别用式(5)和(6)表示。

$$v = \frac{4}{\pi} \left(\arctan \frac{\omega^{gt}}{h^{gt}} - \arctan \frac{w}{h} \right)^2$$

(5)

$$\alpha = \frac{v}{(1 - IOU) + v}$$

(6)

其中, ω 和 h 分别表示检测框的宽和高, ω^{gt} 和 h^{gt} 分别表示真实框的宽和高。

损失函数可以反映模型的预测值和真实值之间的一致程度,其值越小,则模型的鲁棒性越好,良好的损失函数有助于损失值的降低,同时可以加快网络的收敛速度。本研究采用采用 CIOU 损失、分类误差和置信度误

差作为模型的损失函数组成。表 1 为改进损失函数前后模型在训练过程中的训练次数和损失值比较。从表 1 可以看出使用 CIOU 作为损失项,模型的收敛速度更快,损失收敛值更小,对于提高目标的检测精度和促进模型收敛起到积极作用。

表 1 不同损失函数训练参数

Table 1 Training parameters of different loss functions		
损失函数	损失函数值	训练周期
Loss	1.21	45
L_{CIOU}	0.98	45

3 实验结果

3.1 实验平台及数据集

本实验选取 UA-DETRAC^[18] 作为实验数据集,该数据集由北京和天津的道路监控摄像头拍摄而成,整个数据集包括公共汽车、厢式货车等多种类型汽车图像,并且包括多云、夜晚、晴天及雨天多种天气情况,同时包含被其他车辆或障碍物遮挡的车辆,保证了数据的多样性,符合真实场景下的交通状况。

3.2 模型的训练

本算法在 PyTorch 框架下进行,训练及测试的 CPU 为 i7 处理器,显卡为 Tesla V100,操作系统为 Ubuntu16.04。在训练过程中学习率初始值设为 0.001,动量参数为 0.9,批量大小设为 64,共训练 45 个训练周期 (epochs),为方便描述,将具有注意力机制的密集连接网络命名为 AD-YOLO v3,将使用深度可分离卷积进行轻量化操作的网络模型命名为 DW-AD-YOLO v3,对训练过程中的损失函数值进行记录,随着训练周期次数的增加,3 种模型在训练集上的损失函数变化趋势如图 6 所示。

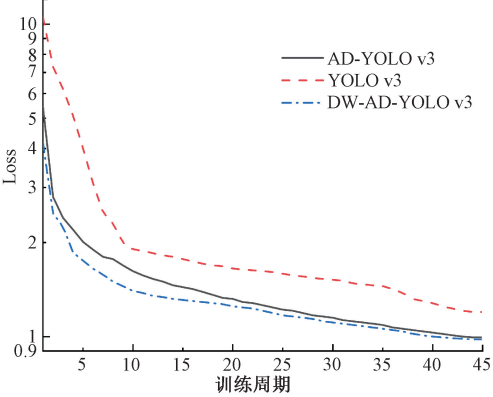


图 6 不同模型的损失函数曲线

Fig.6 The loss function curves of different models

由图 6 可知,DW-AD-YOLO v3 在训练集上收敛速度最快,经过 45 个 epochs 后模型达到拟合状态。

3.3 模型评价指标及实验对比

本研究采用精确率 P (precision)、召回率 R (recall)、平均类别精度 (mean average precision, mAP)、帧率 (FPS)、模型参数量 (parameters) 这 5 项评价指标来评估模型的性能。其中精确率和召回率的定义可用式 (7) 和 (8) 来表示。

$$P = \frac{TP}{TP + FP} \tag{7}$$

$$R = \frac{TP}{TP + FN} \tag{8}$$

其中, TP 表示模型正确检测出来的目标个数, FP 表示模型错误检测出来的目标个数, FN 表示模型漏检的正确的目标个数。通过对 YOLO v3、AD-YOLO v3、DW-AD-YOLO v3 3 种算法进行对比,表 2 展示了以上 3 种算法的对比结果。

表 2 不同模型检测效果对比
Table 2 Comparison of detection effects of different models

模型	mAP/%	Precision/%	Recall/%	Parameters	FPS/(frame · s ⁻¹)
YOLO v3	86.1	71.2	87.4	61 523 734	42
AD-YOLO v3	89.2	72.6	92.5	47 861 146	36
DW-AD-YOLO v3	88.7	72.3	91.3	26 107 152	45

在神经网络结构中,每一次卷积操作相当于对输入图像数据的一次非线性变换,非线性变换的复杂度也随着神经网络层数的加深而增加,密集连接网络允许在网络中复用特征,相比于单一的连接方式更容易获得性能良好的分类决策函数,引入注意力机制后增强了有效特征信息的表达,SoftPool 减少了池化过程中特征信息的损失,保留了较多的目标特征细粒度信息,由表 2 可以看出在引入注意力机制和 SoftPool 后模型的召回率得到了明显的提升,对于漏检问题得到了较大的改善;使用 CIOW 作为边界框预测损失项,其收敛速度更快,收敛值也更小,对于模型的训练起到一定的促进作用;目前的车辆检测算法大多部署在嵌入式设备中,但这些嵌入式设备的计算能力有限难以满足大规模神经网络的运行,相比于 AD-YOLO v3、DW-AD-YOLO v3 的参数量和计算量更少,更适合在这些计算资源有限的嵌入式设备上进行部署,此外 DW-AD-YOLO v3 的检测速度达到 45 FPS,能够完成一些实时性要求较高的车辆检测任务;从精确率、召回率和平均类别精度来看,DW-AD-YOLO v3 与 AD-YOLO v3 没有过大差距,也能够满足车辆检测的要求。综上所述,DW-AD-YOLO v3 的性能要更加均衡,更能满足车辆检测任务的要求。

3.4 SoftPool 对密集连接结构的影响

为了进一步验证 SoftPool 在密集连接网络中的作用,本文同时设计了基于平均值池化 (average pooling) 和最大值池化 (max pooling) 的密集连接网络结构,将其嵌入到本文所提模型的相同位置,在数据集上进行实验并比较,以精确率和召回率作为评价指标,其结果如表 3 所示。可以看出 SoftPool 的检测效果要优于平均值池化和最大值池化的检测效果,在密集连接网络中,由于卷积层之间实现特征复用,平均值池化相比于最大值池化能够

保留更多的全局信息,所以检测效果要比最大值池化更好,但是平均值池化和最大值池化都会在特征图缩小的过程中丢失掉部分关键信息,难以保留小目标的信息,而 SoftPool 通过权重激活使池化区域内的所有激活都对最终输出做出贡献,在保证特征图缩减的同时尽量减少了池化操作过程带来的损失,在性能上要优于平均值池化和最大值池化。

表 3 不同池化方式的检测结果
Table 3 Detection results of different pooling

方法	Precision/%	Recall/%
SoftPool	72.3	91.3
Average Pooling	71.5	89.8
Max Pooling	71.3	88.3

3.5 与主流算法对比

为进一步验证本研究模型的有效性,与其他主流目标算法进行对比实验。实验结果如表 4 所示。可以看出,与二阶段的典型算法 Faster-RCNN^[19] 相比虽然检测能力有所不及,但在速度上却远超 Faster-RCNN,可以满足实时性较高的任务需求,与一阶段目标检测算法 SSD^[20] 和无锚框的目标检测算法 CenterNet^[21] 相比,本研究模型在检测精度和检测速度上都具有一定的优势,而且相较于其他两种算法,模型的规模更小,更适合在终端部署。

表 4 不同检测算法的检测效果对比
Table 4 Comparison of detection results of different detection algorithms

算法	mAP/%	FPS/(frame · s ⁻¹)	Size/MB
Faster-RCNN	96.4	8	488
SSD	83.1	15	136
CenterNet	82.2	33	285
Ours	88.7	45	103

UA-DETRAC 数据集包括晴天、多云、夜晚、雨天 4 种天气情况,同时根据车辆目标的尺度大小、光照强度、不同遮挡程度分为简单、中等、困难 3 个部分,表 5 展示了在不同天气条件下和不同困难程度下本文所提算法的评

估结果。可以看出本文所研究算法在不同条件下的检测效果良好,在中小尺度、困难、夜晚条件下相较于其他检测算法具有一定的优势。

表 5 UA-DETRAC 数据集在不同算法下的 mAP 值
Table 5 mAP values of UA-DETRAC datasets in different algorithm (%)

算法	整体	简单	中等	困难	多云	夜晚	雨天	晴天
SpotNet ^[22]	86.80	97.58	92.57	76.80	89.38	89.53	80.93	91.42
MVSD-SPP ^[23]	85.29	96.04	89.42	76.55	88.00	88.67	78.90	88.91
Ours	88.70	98.53	94.78	78.96	90.46	91.25	81.87	92.13

3.6 检测效果展示

图 7 展示了本研究所提出的算法在测试集中对车辆的检测效果。其中图 7(a)为车辆目标密集时的检测效果,图 7(b)为车辆目标受到遮挡时的检测效果,图 7(c)为夜晚场景下的车辆检测效果,图 7(d)为光线不足时车辆的检测效果。可以看出本研究所提出的模型在城市交通场景中能较好的完成车辆检测的任务,具有较好的鲁棒性和一定的应用价值。

参考文献

[1] ERHAN D, SZEGEDY C, TOSHEV A, et al. Scalable object detection using deep neural networks [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2014: 2147-2154.

[2] HU X W, XU X M, XIAO Y J, et al. SINet: A scale-insensitive convolutional neural network for fast vehicle detection [J]. IEEE Transactions on Intelligent Transportation Systems, 2019, 20(3): 1010-1019.

[3] 蓝金辉,王迪,申小盼. 卷积神经网络在视觉图像检测的研究进展 [J]. 仪器仪表学报, 2020, 41(4): 167-182.

LAN J H, WANG D, SHEN X P. Research progress on visual image detection based on convolutional neural network [J]. Chinese Journal of Scientific Instrument, 2020, 41(4): 167-182.

[4] REDMON J, FARHADI A. YOLO9000: Better, faster, stronger [C]. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26 2017, Honolulu, HI, USA. New York: IEEE, 2017: 6517-6525.

[5] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: Unified, real-time object detection [C]. 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, NV, USA, 2016: 779-788.

[6] 化嫣然,张卓,龙赛,等. 基于改进 YOLO 算法的遥感图像目标检测 [J]. 电子测量技术, 2020, 43(24): 87-92.

HUA Y R, ZHANG ZH, LONG S, et al. Remote sensing image target detection based on improved YOLO algorithm [J]. Electronic Measurement Technology, 2020, 43(24): 87-92.

[7] 张晓丽,董显. 面向铁路货车车号定位的 Faster R-CNN 卷积神经网络 [J]. 电子测量与仪器学报, 2020, 34(10): 65-73.

ZHANG X L, DONG Y. Faster R-CNN convolution neural network for the location of freight train number [J]. Journal of Electronic Measurement and Instrumentation,

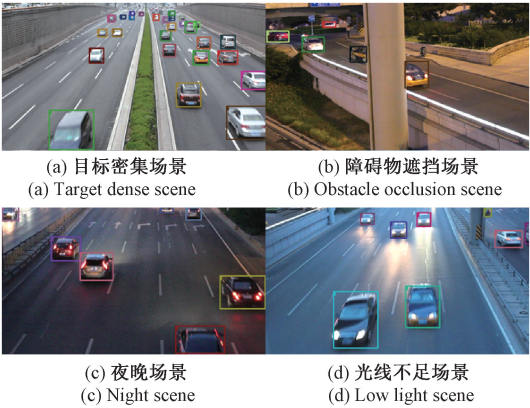


图 7 不同场景下模型对车辆的检测效果
Fig. 7 The detection effect of the model on vehicles in different scenes

4 结 论

为提高复杂交通场景下的车辆检测精度,提出一种具有注意力机制的密集连接方式,并在转换层中使用 SoftPool 整合密集块中包含的特征信息,保留更多的车辆特征信息,将改进后的密集连接结构作为模型的深层特征提取层,同时引入 CIOU 损失作为模型的边界框位置损失项;提高了对小目标车辆、被遮挡目标车辆等背景相对复杂情况下的检测能力;能够满足复杂交通场景下的车辆检测任务需求,为城市车辆检测提供参考依据。为保证模型的实用性,如何在不增加模型复杂度的基础上继续提升检测精度是后续工作的研究重点。

- 2020,34(10):65-73.
- [8] HAO L Y, LI J, GUO G. A multi-target corner pooling-based neural network for vehicle detection[J]. Neural Computing and Applications, 2020, 32 (18): 14497-14506.
- [9] HUA J, SHI Y, XIE C J, et al. Pedestrian-and vehicle-detection algorithm based on improved aggregated channel features[J]. IEEE Access, 2021, (9): 25885-25897.
- [10] JIN T, ZHANG D, DING F, et al. A vehicle detection algorithm in complex traffic scenes [C]. Twelfth International Conference on Digital Image Processing, 2020.
- [11] 崔家华,张云洲,王争,等.面向嵌入式平台的轻量级目标检测网络[J].光学学报, 2019, 39(4):307-313.
CUI J H, ZHANG Y ZH, WANG ZH, et al. Light-weight object detection networks for embedded platform [J]. Acta Optica Sinica, 2019, 39(4): 307-313.
- [12] HUANG G, LIU Z, MAATEN L, et al. Densely connected convolutional networks [C]. 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, HI, USA, 2017: 2261-2269.
- [13] ALEXANDROS S A, RONALD P, GRIGORIOS K. Refining activation downsampling with SoftPool [EB/OL]. [2021-02-02]. <https://arxiv.org/abs/2101.00440>.
- [14] REDMON J, FARHADI A. YOLOv3: An incremental improvement [C]. Conference Computer Vision and Pattern Recognition, 2018, arXiv: 1804.02767.
- [15] 曹红燕,沈小林,刘长明,等.改进 YOLOv3 的红外目标检测算法[J].电子测量与仪器学报,2020,34(8): 188-194.
CAO H Y, SHEN X L, LIU CH M et al. Improved infrared target detection algorithm of YOLOv3 [J]. Journal of Electronic Measurement and Instrumentation, 2020, 34(8): 188-194.
- [16] ZHENG Z, WANG P, LIU W, et al. Distance-IoU loss: Faster and better learning for bounding box regression[C]. Proceedings of the AAAI Conference on Artificial Intelligence, 2020, 34(7): 12993-13000.
- [17] WANG Q, WU B, ZHU P, et al. ECA-Net: Efficient channel attention for deep convolutional neural networks [C]. Proceedings of the Conference on Computer Vision and Pattern Recognition, Piscataway, NJ: IEEE Press, 2019: 1-12.
- [18] WEN L, DU D, CAI Z, et al. UA-DETRAC: A new benchmark and protocol for multi-object detection and tracking[J]. Computer Vision and Image Understanding, 2020, 193: 102907.
- [19] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2015, 39(6): 1137-1149.
- [20] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single shot multibox detector [C]. Proceedings of the 2016 European Conference on Computer Vision. Amsterdam, The Netherlands: Springer, 2016: 21-37.
- [21] ZHOU X Y, WANG D Q, KRÄHENBÜHL P. Objects as points [EB/OL]. [2020-04-02] <https://arxiv.org/abs/1904.07850>.
- [22] PERRRAULT H, BILODEAU G A, SAUNIERN, et al. SpotNet: Self-attention multi-task network for object detection [C]. Proceedings of the 17th Conference on Computer and Robot Vision (CRV), ELECTRONETWORK, 2020: 230-237.
- [23] KIM K J, KIM P K, CHUNG Y S, et al. Multi-scale detector for accurate vehicle detection in traffic surveillance data [J]. IEEE Access, 2019, (7): 78311-78319.

作者简介



梁继然 (通信作者), 2008 年于天津大学获得博士学位, 现为天津大学副教授, 主要研究方向为功能敏感器件集成与传感器智能应用。

E-mail: liang_jiran@tju.edu.cn

Liang Jiran (Corresponding author)

received his Ph. D. degree from Tianjin University in 2008. Now he is an associate professor in Tianjin University. His main research interests include functional sensitive device integration and sensor intelligent application.



陈壮, 2015 年于山东师范大学获得学士学位, 现为天津大学研究生, 主要研究方向为图像处理和深度学习。

E-mail: chenzhuang6677@163.com

Chen Zhuang received his B. Sc.

degree from Shandong Normal University in 2015. He is now a M. Sc. candidate at Tianjin University. His main research interests include image processing and deep learning.