

# 基于改进 Mask R-CNN 的交通监控视频 车辆检测算法

陆 飞<sup>1,2</sup>, 沈世斌<sup>1,3</sup>, 苏晓云<sup>1,2</sup>, 谢 非<sup>1,2,3</sup>, 章 悦<sup>1</sup>, 刘益剑<sup>1,2,3</sup>

(1.南京师范大学电气与自动化工程学院,江苏 南京 210023)

(2.江苏省三维打印装备与制造重点实验室,江苏 南京 210023)

(3.南京智能高端装备产业研究院有限公司,江苏 南京 210023)

**[摘要]** 针对交通监控视频车辆检测常易受到遮挡导致目标车辆出现漏检或误检的问题,提出一种基于改进 Mask R-CNN 的交通监控视频车辆检测算法。采用基于 bottleneck 结构的主干网络,提高主干网络提取特征的能力;通过基于预测 mask 分数的掩码分支,融合目标的类别分数和掩码质量分数,提高车辆的掩码质量;通过基于 Arcface Loss 的目标检测损失函数设计,提高不同特征之间的可判别性,提高目标的检测精度。实验结果表明,改进的 Mask R-CNN 模型可更好地检测到被遮挡的车辆,目标车辆的检测精度超过 Faster R-CNN、YOLO v3 和 Mask R-CNN 模型,可解决目标车辆漏检或误检问题。

**[关键词]** 目标检测,交通监控,Mask R-CNN,掩码预测

**[中图分类号]** TP391 **[文献标志码]** A **[文章编号]** 1672-1292(2020)04-0044-07

## Vehicle Detection Algorithm Based on Improved Mask R-CNN in Traffic Surveillance Video

Lu Fei<sup>1,2</sup>, Shen Shibin<sup>1,3</sup>, Su Xiaoyun<sup>1,2</sup>, Xie Fei<sup>1,2,3</sup>, Zhang Yue<sup>1</sup>, Liu Yijian<sup>1,2,3</sup>

(1.School of Electrical and Automation Engineering, Nanjing Normal University, Nanjing 210023, China)

(2.Jiangsu Key Laboratory of 3D Printing Equipment and Manufacturing, Nanjing 210023, China)

(3.Nanjing Industry Institute for Advanced Intelligent Equipment Co., Ltd., Nanjing 210023, China)

**Abstract:** Aiming at the problem of missing detection or wrong detection of target vehicles caused by occlusion in traffic surveillance video, an improved vehicle detection algorithm based on Mask R-CNN traffic surveillance video is proposed. Firstly, the backbone network based on the bottleneck structure is used to improve the ability of extracting features from the backbone network. Then, the mask branch based on the predicted mask score is used to fuse the target's category score and mask quality score to improve the vehicle's mask quality. Finally, the target detection loss function based on Arcface Loss can improve the discriminability between different features and improve the detection accuracy of the target. The experimental results show that the improved Mask R-CNN model can better detect the shielded vehicle, and that the detection accuracy of the target vehicle is higher than those of the Faster R-CNN, YOLO v3 and Mask R-CNN model, thus solving the problem of missing or wrong detection of the target vehicle.

**Key words:** target detection, traffic surveillance, Mask R-CNN, mask prediction

基于深度学习的目标检测技术在无人驾驶、智能监控、智能交通等领域有着广泛的应用。随着城市道路车辆违反交通法规的事件不断增加,对监控视频中的车辆快速精准地检测已成为城市交通管理的一项重要研究课题。快速且准确地检测复杂场景中不同尺度和类型的车辆已成为目前最具挑战性的课题之一<sup>[1-2]</sup>。

收稿日期:2020-06-11.

基金项目:国家自然科学基金项目(61601228、41974033、61803208)、江苏省自然科学基金项目(BK20161021、BK20180726)、江苏省高校自然科学基金项目(17KJB510031)。

通讯作者:沈世斌,高级实验员,研究方向:嵌入式系统、目标检测与跟踪、机器视觉与图像处理。E-mail:63018@njnu.edu.cn

近年来,国内外学者针对基于深度学习的目标检测方法展开了广泛而深入的研究. Girshick 等<sup>[3]</sup>首次提出 R-CNN 模型,通过 Selective Search 算法<sup>[4]</sup>提取目标对象的候选框,将候选框的大小调整为一个固定尺寸,然后输入到卷积神经网络中提取特征,采用线性回归策略调整目标的边界框,使用 SVM 分类器预测目标的类别,但产生候选框的过程耗时长,且候选框之间存在大量重叠. He 等<sup>[5]</sup>提出了空间金字塔池化网络(SPPNet),克服了 R-CNN 冗余的特征计算,在目标检测时对整张图像只进行一次特征计算,然后引入空间金字塔池化层通过 CNN 生成任意感兴趣区域. Girshick 等<sup>[6]</sup>提出了 Fast R-CNN 算法,采用 Softmax 函数代替原来的 SVM 分类器,将目标的分类与回归加入到神经网络中,减少了运算时间. Ren 等<sup>[7]</sup>在 Fast R-CNN 的基础上作了改进,提出了 Faster R-CNN 模型, Faster R-CNN 主要贡献在于引入区域建议网络(RPN)<sup>[8]</sup>代替了原来的选择性搜索策略,并将对象候选框的提取、区域归一化以及分类与回归等操作都集中在端到端框架中. Redmon 等<sup>[9]</sup>提出 YOLO 算法,首次将目标检测当作回归问题来处理,无需产生候选框,直接通过卷积神经网络预测目标的类别和位置,提高了检测的速度,但精度却较低. He 等<sup>[10]</sup>提出了 Mask R-CNN 模型,结合了经典的目标检测算法 Faster R-CNN 和经典的语义分割算法 FCN<sup>[11]</sup>, Faster R-CNN 可以既快又准地实现目标检测功能, FCN 可以精准实现语义分割功能,由于发现了 ROI Pooling<sup>[12]</sup>中所存在的像素偏差问题,提出了对应的 ROI Align 策略,加上 FCN 精准的像素 mask,可获得高准确率. 江昆鹏等<sup>[13]</sup>调整了区域建议网络,在 anchor 的基础上增加了两个尺度,提高了目标的识别精度. 朱有产等<sup>[14]</sup>优化 Mask R-CNN 主干网络的深度,并在卷积操作中加入开运算层,减少图片的边缘噪声. 石杰等<sup>[15]</sup>在 Mask R-CNN 的基础上加入 LH R-CNN,提高了运算速度.

由于交通监控视频中的目标车辆存在目标尺寸和尺度的巨大差异,车辆之间存在相互遮挡问题,导致目标车辆检测存在误检或漏检. 针对 Mask R-CNN 在车辆检测上出现误检或漏检的问题,本文基于改进的 Mask R-CNN 的交通监控视频车辆检测算法,在 Mask R-CNN 通用结构框架基础上,设计了基于 bottleneck 结构的主干网络、基于预测 mask 分数的掩码分支和基于 Arcface Loss 的目标检测损失函数,解决视频监控中部分目标由于存在遮挡而导致漏检或误检的问题,提高了目标的检测精度和目标特征的判别能力.

## 1 基于改进 Mask R-CNN 的交通监控视频车辆检测算法

### 1.1 改进的 Mask R-CNN 算法的结构框架

Mask R-CNN 模型将目标的分类分数作为最终目标分割的分数,即评判分类、边界框修正以及掩码分割效果取决于该分类分数. 在背景杂乱、目标相互遮挡等复杂场景下,可能会出现目标被漏检或误检,以及目标掩码质量低的问题. 针对上述问题,本文在 Mask R-CNN 结构框架基础上,设计了基于 bottleneck 结构的主干网络、基于预测 mask 分数的掩码分支和基于 Arcface Loss 的目标检测损失函数,改进后的 Mask R-CNN 结构框架图如图 1 所示.

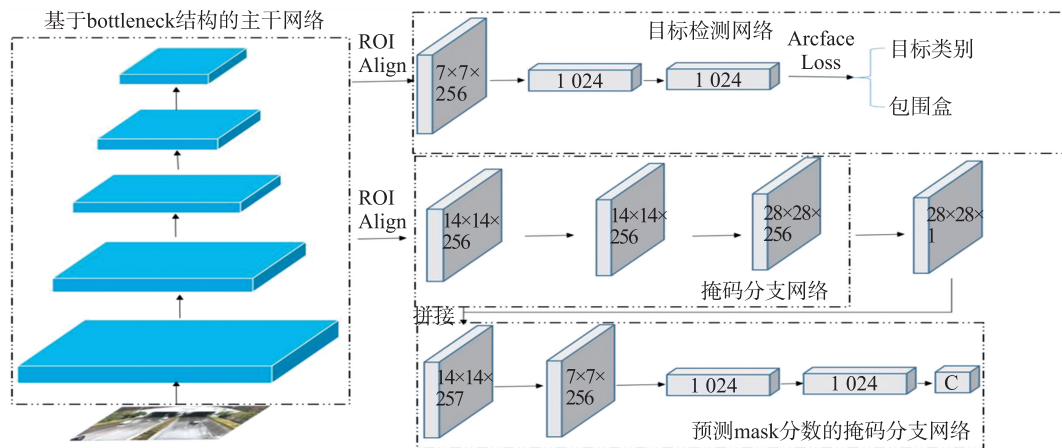


图 1 改进的 Mask R-CNN 算法的结构框架图

Fig. 1 The structure of improved Mask R-CNN

## 1.2 基于 bottleneck 结构的主干网络设计

Mask R-CNN 采用残差网络 ResNet<sup>[16]</sup> 作为主干网络来完成图像特征的提取. ResNet 网络分为 ResNet50、ResNet101、ResNet152,这 3 种网络的主要区别在于层数不同,层数越深,提取图像的特征就越丰富,越能反映图像抽象的语义信息,但层数的加深往往会带来梯度消失或爆炸、不易训练等问题. 由于 bottleneck 结构能有效减少卷积的参数个数和计算量,深度提取目标特征,为了在性能和运行速度之间保持平衡,本文选用 bottleneck 作为残差网络的主干结构.

基于 bottleneck 结构的主干网络结构如图 2 所示. ResNet 使用 bottleneck 残差结构, Batch Normalization 层用于加快收敛速度. ResNet 共有 5 组卷积,第一组卷积的输入大小为  $224 \times 224$ ,第五组卷积的输出大小为  $7 \times 7$ ,总共缩小 5 次,每次缩小 2 倍,最终输入图像缩小了 32 倍,减少了参数的数量,从而减少了计算量.

## 1.3 基于预测 mask 分数的掩码分支设计

Mask R-CNN 通过掩码分支对目标车辆进行掩码分割,会得到一个沿着目标车辆轮廓包围的闭合区域,闭合区域为目标分割的掩码. 由于 Mask R-CNN 预测掩码质量的分数和类别分类的分数相同,因此 Mask R-CNN 不能完全评估目标车辆的掩码分割质量. 针对该问题,本文在掩码分支网络的基础上增加了基于预测 mask 分数的掩码分支,用于单独评估目标车辆的掩码质量.

基于预测 mask 分数的掩码分支网络是将目标的分割掩码与 ROI 特征图拼接在一起作为输入,在拼接之前,需要对目标的分割掩码进行最大值池化,保持与 ROI 特征图的尺寸大小一致. 基于预测 mask 分数的掩码分支网络结构如图 3 所示,图中使用 3 个卷积层分别为  $14 \times 14 \times 257$ ,  $14 \times 14 \times 256$ ,  $7 \times 7 \times 256$ ; 3 个全连接层分别为 1 024, 1 024, 1. 对于 3 个全连接层,遵循 Mask R-CNN 目标检测网络设计原则,将前两个 FC 层的输出设置为 1024,最后一个全连接层的输出设置为类别数目,掩码分支只对目标车辆的分割掩码进行评分,即类别数目为 1.

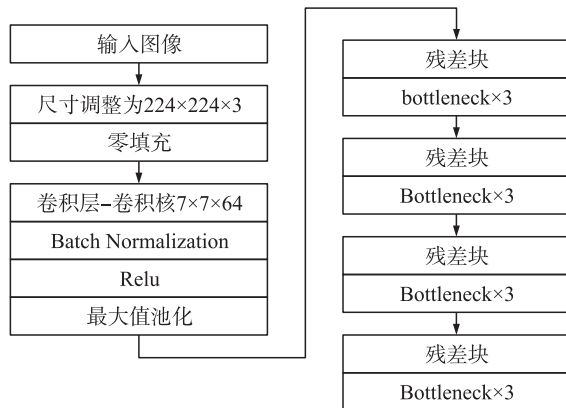


图 2 基于 bottleneck 结构的主干网络结构图

Fig. 2 The structure of backbone network based on bottleneck

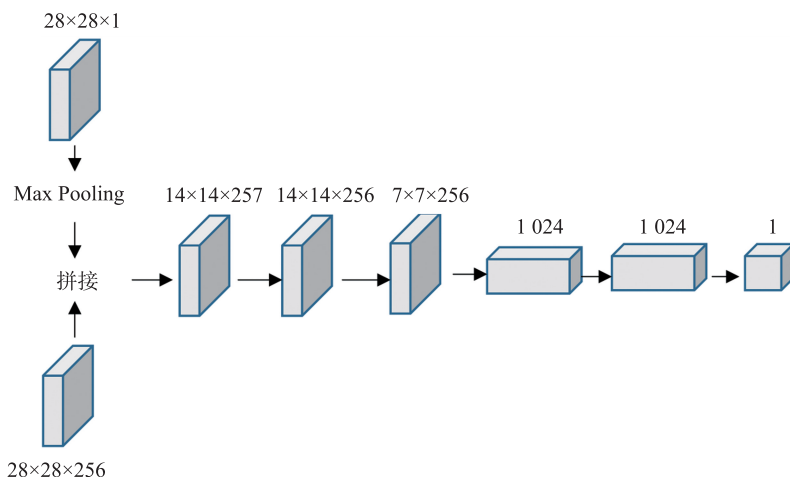


图 3 基于预测 mask 分数的掩码分支网络结构图

Fig. 3 The structure of mask branch network based on predicting mask score

基于预测 mask 分数的掩码分支通过对掩码分支网络输出的目标掩码与真实标签值掩码重叠面积 (IoU) 的预测,来精确地评估掩码的质量. 最终的评分细则定义如下:

$$S_{\text{mask}} = S_{\text{cls}} \times S_{\text{IoU}}, \quad (1)$$

式中,  $S_{\text{mask}}$  为最终的掩码质量分数;  $S_{\text{cls}}$  为 Mask R-CNN 中所预测的类别分类分数;  $S_{\text{IoU}}$  为预测的 IoU. 掩码分

割分支与基于预测 mask 分数的掩码分支相互独立. 在训练阶段,分类正确的目标车辆掩码才会输入到基于预测 mask 分数的掩码分支,即  $S_{cls}$  必须取正值. 然后,基于预测 mask 分数的掩码分支对分类正确的目标车辆掩码预测对应的 IoU,不断减小  $S_{mask}$  和  $S_{IoU}$  之间的差距,使得  $S_{mask}$  接近于  $S_{IoU}$ ,保证目标类别的掩码质量与掩码分数保持一致. 改进后的 Mask R-CNN 既能确保分类的准确,又能评判掩码质量的好坏.

#### 1.4 基于 Arcface Loss 的目标检测损失函数设计

Softmax Loss 作为目标检测模型中常用的损失函数,用来输出目标类别的概率,以保证目标类别的可分性,但若两个类别的特征相似度高时,使用 Softmax Loss 进行相似度学习,会无法区分出这两个类别. 在 Softmax Loss 基础上提出的 A-softmax Loss<sup>[17]</sup>,加入了角度距离,使得特征具有可判别性:

$$L_{ang} = \frac{1}{N} \sum_i -\log \left( \frac{e^{\|x_i\| \psi(\theta_{y_i}, i)}}{e^{\|x_i\| \psi(\theta_{y_i}, i)} + \sum_{j \neq y_i} e^{\|x_i\| \cos(\theta_j, i)}} \right), \quad (2)$$

式中,  $\psi(\theta_{y_i}, i) = (-1)^k \cos(m\theta_{y_i}, i) - 2k$ ,  $\theta_{y_i, i} \in \left[ \frac{k\pi}{m}, \frac{(k+1)\pi}{m} \right]$ ,  $k \in [0, m-1]$ .

以这种方式拉近网络模型类内距离,放大类间距离,使得类别间的特征更具可分性,但由于夹角增加到原来的  $m$  倍,网络模型训练难度变大,导致损失函数的损失值很难收敛.

本文通过融合 A-Softmax Loss 和 Softmax Loss 的特性,设计基于 Arcface Loss 的目标检测损失函数, Arcface Loss<sup>[18]</sup> 计算公式为:

$$L = -\frac{1}{N} \sum_{i=1}^N \log \left( \frac{e^{s(\cos(\theta_{y_i} + m))}}{e^{s(\cos(\theta_{y_i} + m))} + \sum_{j \neq y_i} e^{s \cos(\theta_j)}} \right), \quad (3)$$

式中,  $s$  为固定尺度因子;  $m$  为正整数.

图 4 所示为 Arcface Loss 原理示意图. 首先,对目标特征向量  $X_i$  和权重  $W_{y_i}$  进行归一化操作,并使用反余弦函数来计算目标特征向量和权重之间的夹角  $\theta_{y_i}$ . 将附加的角余量  $m$  添加到目标角度上. 然后,使用余弦函数再次获得余弦距离,并乘上固定尺度因子  $s$ ,将以上结果输入到 Softmax 层中,以完成目标识别工作.

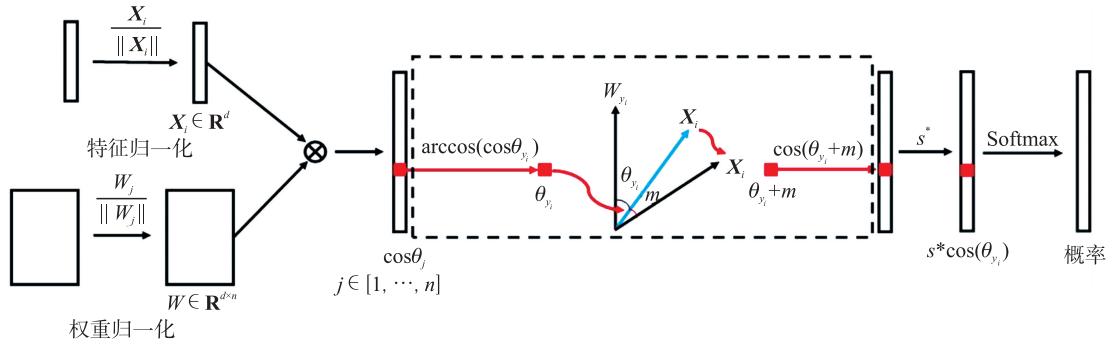


图 4 Arcface Loss 原理示意图

Fig. 4 Schematic diagram of Arcface Loss

由于 A-Softmax Loss 计算特征相似度时采用余弦距离,而 Arcface Loss 计算特征相似度时采用角度距离,计算更加方便. A-Softmax Loss 让对应类别的  $W$  和  $X$  的夹角增加到原来的  $m$  倍,而 Arcface Loss 只在原来夹角的基础上增加一个角余量  $m$ ,不需要借助退火方法,网络的训练难度降低.

改进后的 Mask R-CNN 网络的损失函数表达式为:

$$L = L_{cls} + L_{box} + L_{mask} + L_{score}. \quad (4)$$

其中,

$$L_{cls} = -\frac{1}{N} \sum_{i=1}^N \log \left( \frac{e^{s(\cos(\theta_{y_i} + m))}}{e^{s(\cos(\theta_{y_i} + m))} + \sum_{j \neq y_i} e^{s \cos(\theta_j)}} \right), \quad (5)$$

式中,  $L_{cls}$  为目标检测网络中类别分类的损失值;  $N$  为区域候选网络获得的候选框个数;  $\cos(\theta_{y_i})$  为对应的每个候选框分类正确的概率.

$$L_{box} = \sum_{i \in (x, y, w, h)} \text{smooth}_{L1}(t_i - t_i^*), \quad (6)$$

式中,  $L_{\text{box}}$  为车辆检测框回归的损失值;  $t_i$  为检测框的 4 个坐标数据组成的向量;  $t_i^*$  为标准框的 4 个坐标数据组成的向量.

$$L_{\text{mask}} = - \sum_i p_i \log(y_i) + (1 - p_i) \log(1 - y_i),$$

(7)

式中,  $L_{\text{mask}}$  为掩码分支网络的损失值;  $i$  为每个 ROI 的索引数;  $p_i$  为 ROI 二分类的概率值;  $y_i$  为第  $i$  个输出.  $L_{\text{mask}}$  仅在 ROI 的正样本上定义.

$$L_{\text{score}} = \sum_{i=0}^m (y - y^*)^2,$$

(8)

式中,  $L_{\text{score}}$  为基于预测 mask 分数的掩码分支的损失值;  $m$  为目标掩码的总数;  $y$  为掩码分支网络输出目标掩码的类别分数;  $y^*$  为基于预测 mask 分数的掩码分支对相应的目标掩码预测的 IoU.

2 实验与分析

2.1 车辆标记及训练参数设置

为了更好地检测出交通监控视频图像中的目标车辆,实验首先使用构建的数据集对改进后的 Mask R-CNN 网络进行多次迭代训练,再将迭代训练后的网络模型应用到实际的交通监控视频图像中,以检测出目标车辆.实验从构建的数据集中选取了 50 000 张目标车辆图片,使用 Labelme 软件对 50 000 张车辆图片进行人工标注,标注出每辆车的类别、边界框及标签值掩码.标注的数据集一共有 3 类,分别为 car、bus、truck.图 5 所示为人工标注结果示意图.实验按 4:1 比例划分训练集和测试集,其中 40 000 张为训练图像,10 000 张为测试图像.



图 5 人工标注结果示意图  
Fig. 5 Schematic diagram of manual annotation results

实验中网络训练参数的设置如表 1 所示.

表 1 改进的 Mask R-CNN 网络训练参数设置  
Table 1 The training parameter setting of improved Mask R-CNN network

网络训练参数	参数值	网络训练参数	参数值
区域候选网络中 NMS 阈值	0.8	权重衰减量	0.000 000 5
ROI 正负样本比例	0.33	动量	0.9
基础学习率	0.001	优化方法	SGD

2.2 实验结果

实验在构建的数据集上对 Mask R-CNN 模型和改进后的 Mask R-CNN 模型分别进行训练,每次输入 16 张图片,训练 50 轮,共迭代 36 000 次,在迭代 20 000 次时学习率设置为 0.000 1.改进后的 Mask R-CNN 网络模型的损失函数曲线如图 6 所示.

图 6 显示,在训练和测试阶段前期,损失值下降速率较快,后期曲线趋向平稳,改进后的 Mask R-CNN 网络模型收敛.将待检测的交通视频图像转换成 800×800 尺寸大小,分别输入到迭代训练好的 Mask R-CNN 和改进后的 Mask R-CNN 中,输出所有目标车辆的边界框位置、掩

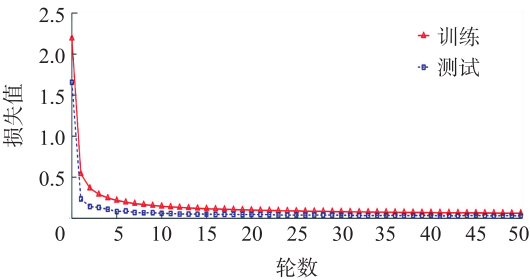


图 6 改进的 Mask R-CNN 网络在构建的数据集上的损失函数曲线图  
Fig. 6 The loss function curve of improved Mask R-CNN network on the constructed dataset

码及最终的掩码质量分数.

图 7 和图 8 分别为交通监控场景 1、2 的实验结果图,其中(a)为 Mask R-CNN 的处理结果,(b)为改进后的 Mask R-CNN 的处理结果. 在图 7 的交通监控场景 1 中,车辆数量较少且互相无遮挡,两种算法均能检测和分割出所有的运动车辆,最终的掩码质量分数相差不大,均能反映出真实的掩码质量. 在图 8 的交通监控场景 2 中,车辆数量较多且互相之间遮挡严重,Mask R-CNN 不仅漏检车辆,且部分车辆因相互遮挡导致掩码分割效果不理想,但最终的掩码质量分数却很高,即掩码质量分数与实际的掩码质量不一致;改进后的 Mask R-CNN 不仅能检测出所有的目标车辆,且最终的掩码质量分数能真实地反映真实的掩码质量.

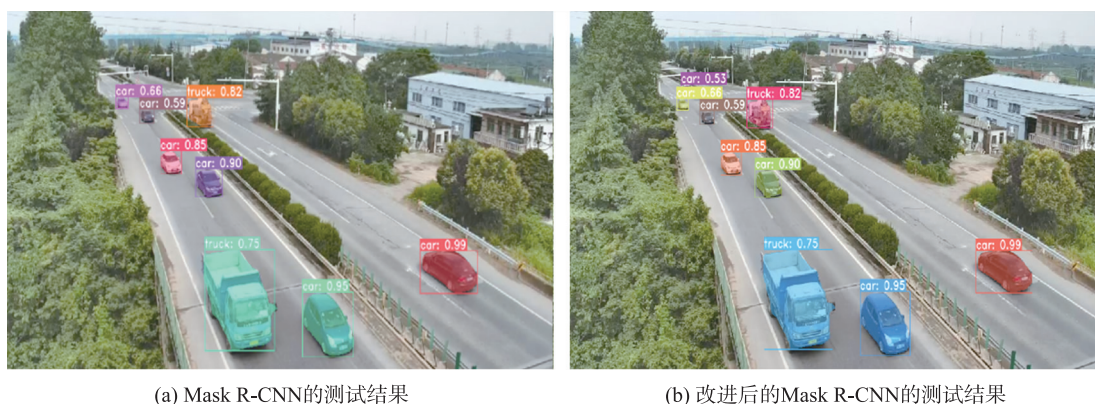


图7 交通监控场景1的实验结果图

**Fig. 7** Experimental results of traffic monitoring scenario 1

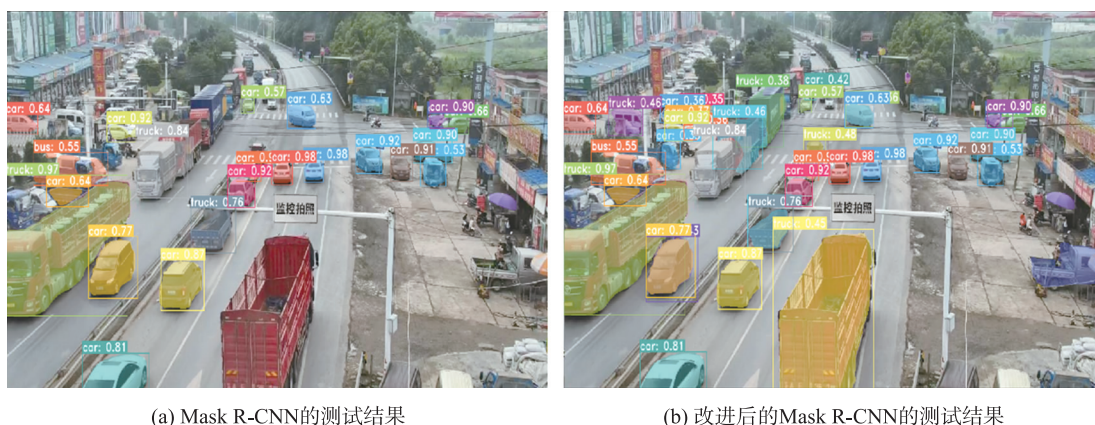


图 8 交通监控场景 2 的实验结果图

**Fig. 8** Experimental results of traffic monitoring scenario 2

可通过计算准确率和召回率来评价网络模型的检测性能:

$$P_{\text{precision}} = \frac{\text{TP}}{\text{TP} + \text{FP}}, \quad (9)$$

$$R_{\text{recall}} = \frac{\text{TP}}{\text{TP} + \text{FN}}, \quad (10)$$

式中, TP 为正确划分为正样本的个数; FP 为错误划分为正样本的个数; FN 为错误划分为负样本的个数。

$P-R$  曲线即以召回率和准确率作为横、纵坐标的曲线,  $P-R$  曲线所围面积即为类别精度 (average precision, AP), mAP 就是所有类的 AP 平均值, AP 和 mAP 可用于评判目标检测性能. FPS 为每一秒模型检测的图片张数, 可用于评价网络模型的检测速度.

本实验与 Faster R-CNN、YOLO v3 及 Mask R-CNN 模型在测试集上的 FPS、mAP 测试结果对比如表 2 所示. 结果表明,改进的 Mask R-CNN 模型在满足速度的

表 2 4 种模型在构建数据集上的测试结果

**Table 2** Test results of four models on constructed datasets

方法	FPS	mAP/%
Faster R-CNN	7.6	85.6
YOLO v3	28.2	77.3
Mask R-CNN	12.3	88.5
改进的 Mask R-CNN	15.4	94.8

同时,准确率要高于其余3种模型.

### 3 结论

针对交通监控视频中车辆检测的目标车辆存在遮挡,导致目标车辆存在漏检或误检的问题,提出了一种基于改进的Mask R-CNN的交通监控视频车辆检测算法.实验结果表明,相比较于Faster R-CNN、YoLo v3和Mask R-CNN网络模型,本文所提出的改进模型可以更好地提高目标车辆的检测精度,即使存在车辆目标互相遮挡的情况,依然能精准获得车辆目标的类别、位置及掩码,为交通监控视频中的车辆检测提供了有效的途径.

### [参考文献](References)

- [1] 沈连丰,张瑞,朱亚萍,等.面向自动驾驶的车辆精确实时定位算法[J].电子与信息学报,2020,42(1):28-35.
- [2] ANALA M R,MAKKER M,ASHOK A. Anomaly detection in surveillance videos[C]//2019 26th International Conference on High Performance Computing,Data and Analytics Workshop. Hyderabad,India;IEEE Computer Society,2019:93-98.
- [3] GIRSHICK R,DONAHUE J,DARRELL T,et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]//The IEEE Conference on Computer Vision and Pattern Recognition(CVPR). Columbus,USA;IEEE Computer Society,2014:580-587.
- [4] UIJLINGS J R R,SANDE K E A,GEVERS T,et al. Selective search for object recognition[J]. International Journal of Computer Vision,2013,104(2):154-171.
- [5] HE K,YU X. Spatial pyramid pooling in deep convolutional networks for visual recognition[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence,2015,37(9):1904-1916.
- [6] GIRSHICK R B. Fast R-CNN[C]//The IEEE International Conference on Computer Vision(ICCV). Santiago,Chile;IEEE Computer Society,2015:1440-1448.
- [7] 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.
- [8] ZHANG X,LI B,HU H. Scale-aware hierarchical loss:a multipath RPN for multi-scale pedestrian detection[C]//IEEE Visual Communications and Image Processing(VICIP). Petersburg,USA;IEEE Computer Society,2017:1-4.
- [9] REDMON J,DIVVALA S K,GIRSHICK R B,et al. You only look once:Unified,real-time object detection[C]//IEEE International Conference on Computer Vision and Pattern Recognition. Las Vegas,USA;IEEE Computer Society,2016:779-788.
- [10] HE K,GKIOXARI G,DOLLAR P,et al. Mask R-CNN[C]//IEEE International Conference on Computer Vision and Pattern Recognition. Hawaii,USA;IEEE computer Society,2017:2961-2969.
- [11] LI Y,QI H,DAI J. Fully convolutional instance-aware semantic segmentation[C]//IEEE Conference on Computer Vision and Pattern Recognition. Hawaii,USA;IEEE Computer Society,2017:4438-4446.
- [12] 陈泽,叶学义,钱丁炜,等. 基于改进的Faster R-CNN小尺度行人检测[J]. 计算机工程,2020,46(9):226-232.
- [13] 江昆鹏,闫洪涛,杨红卫,等. 改进Mask R-CNN的细粒度车型识别算法[J]. 软件,2020,41(3):1-5.
- [14] 朱有产,王雯瑶. 基于改进Mask R-CNN的绝缘子目标识别方法[J]. 微电子学与计算机,2020,37(2):69-74.
- [15] 石杰,周亚丽,张奇志. 基于改进Mask R-CNN和Kinect的服务机器人物品识别系统[J]. 仪器仪表学报,2019,40(4):216-228.
- [16] 马素刚,赵祥模,侯志强,等. 一种基于ResNet网络特征的视觉目标跟踪算法[J]. 北京邮电大学学报,2020,43(2):129-134.
- [17] LIU W,WEN Y,YU Z,et al. SphereFace:deep hypersphere embedding for face recognition[C]//IEEE Conference on Computer Vision and Pattern Recognition(CVPR). Hawaii,USA;IEEE Computer Society,2017:6738-6746.
- [18] DENG J,GUO J,ZAFEIRIOU S. ArcFace:additive angular margin loss for deep face recognition[C]//IEEE Conference on Computer Vision and Pattern Recognition(CVPR). Los Angeles,USA;IEEE Computer Society,2019:4685-4694.

[责任编辑:严海琳]