

DOI: 10.19650/j.cnki.cjsi.J2107749

改进 Mask R-CNN 的交通场景多目标快速检测与分割^{*}

伍锡如¹, 邱涛涛¹, 王耀南²

(1. 桂林电子科技大学电子工程与自动化学院 桂林 541004; 2. 湖南大学电气与信息工程学院 长沙 410082)

摘要:针对智能驾驶中出现的交通场景多目标检测与分割效率低、鲁棒性差等问题,提出一种改进的 Mask R-CNN 交通场景多目标快速检测与分割方法。首先采用轻量级 MobileNet 作为骨干网络,有效减少网络参数并压缩模型体积,提升后续嵌入式端的算法移植能力,其次通过优化 FPN 与骨干网络卷积结构,保证高底层之间特征信息的完整传递,通过调整超参数得到交通场景多目标检测与分割改进网络模型。设计不同交通场景下的对比实验,改进网络能够准确实现多目标的检测与分割,平均检测精度可达 85.2%。在 ApolloScape 和 NuScence 数据集上进行迁移实验,改进网络展示出良好的泛化能力。本文所提出的改进骨干网络与网络结构优化,能够适应多种复杂交通场景,完成交通场景多目标的快速检测与分割,为智能驾驶提供了理论依据与技术方案。

关键词: MobileNet; 目标检测; 实例分割; 深度学习

中图分类号: TP183 TP391.41 TH744 **文献标识码:** A **国家标准学科分类代码:** 520.6040

Multi-object detection and segmentation for traffic scene based on improved Mask R-CNN

Wu Xiru¹, Qiu Taotao¹, Wang Yaonan²

(1. School of Electronic Engineering and Automation, Guilin University of Electronic Technology, Guilin 541004, China;

2. College of Electrical and Information Engineering, Hunan University, Changsha 410082, China)

Abstract: Aiming at the problems of low efficiency and poor robustness of multi-object detection and segmentation for traffic scenes in intelligent driving, a fast Multi-object detection and segmentation for traffic scene based on improved Mask R-CNN is proposed. Firstly, in order to effectively reduce network parameters and compress model volume, the lightweight MobileNet is used as the backbone network to improve the ability of transplant algorithms on the subsequent embedded side. Secondly, by optimizing the convolution structure of FPN and backbone network to ensure the complete transfer of feature information between high-level layers, the improved network model for multi-object detection and segmentation in traffic scenes is obtained by adjusting hyperparameters. Comparative experiments are conducted under different traffic scenarios, the improved network can accurately realize the detection and segmentation of multiple objects and the average detection accuracy can reach 85.2%. Migration experiments are carried out on the ApolloScape and NuScence dataset to improve the network, which show good generalization capabilities. The improved backbone structure and network structure optimization proposed in this paper can adapt to a variety of complex traffic scenarios and complete the fast detection and segmentation of multiple object in traffic scenarios. It provides theoretical basis and technical solutions for intelligent driving.

Keywords: MobileNet; object detection; instance segmentation; deep learning

收稿日期: 2021-04-12 Received Date: 2021-04-12

^{*} 基金项目: 国家自然科学基金(61863007)、广西自然科学基金(2020GXNSFDA238029)、广西研究生教育创新计划(YCSW2020159, 2021YCX122)、广西研究生教育创新计划(C20YJM00BX0M)项目资助

0 引言

交通场景多目标检测是交通管理、城市安全和无人驾驶的核心研究领域,也是智能交通系统中数据前端的关键研究课题^[1-3]。交通场景中随时出现的行人、车辆、交通标志以及交通灯是影响智能自动驾驶的主要因素,受到环境以及人为因素干扰,比如遮挡、光照不均匀以及恶劣天气等情况,导致检测难度加大。因此,交通场景多目标快速检测是热点研究的问题。

现有的交通场景多目标检测主要基于先验知识,如背景差分法、光流法等。其中,光流法^[4-6]通过将图像的灰度值与二维速度场相结合,获得交通目标的视觉检测结果。背景差分法实现运动场景下目标检测^[7-8],建立图像序列的背景模板,通过融合高斯、码本模型、Vibe 算法来匹配灰度值的参数,检测当前图像帧与背景模板之间的差分得到当前场景下的目标检测。基于图像处理的方向梯度直方图、局部图像纹理特征等算法^[9-10],将场景下多目标提取特征并传输至支持向量机、迭代等分类器进行目标分割。通过特征提取的方法,特征质量直接影响到分割的性能,需要不断更新参数以及模板匹配来获取最优特征,每个特征模板限定特定任务,无法满足多目标的快速检测。

随着深度学习技术的快速发展,神经网络在目标检测与分割领域中取得了显著的优势^[11-13]。例如

Wu 等^[14]提出了利用 Faster R-CNN 实现不同尺度人脸检测,提升了人脸检测效率。郑少武等^[15]提出一种基于激光点云与图像信息融合的交通环境车辆检测方法,实现了复杂交通环境下的车辆检测。范宜标等^[16]结合卷积神经网络与多尺度特征图设计了一种新型车辆检测算法,获得车辆更准确的检测。文献^[17]中采用融合 CNN 不同尺度的特征图的方法实现行人图像的检测。徐谦等^[18]利用 RT-SegNet 分割网络完成了多种道路环境的语义分割。

本文针对交通场景下多目标快速检测与分割问题。设计一种改进的 Mask R-CNN 网络,采用 MobileNet 取代原始 ResNet-50/101 作为骨干网络,降低模型参数同时减少检测时间。通过优化特征提取网络卷积结构,提升目标特征在高底层之间的传递能力。进行多组实验对比并对检测结果进行评估,单张图像检测时间为 0.402 s,平均检测精度可达 85.2%。在迁移数据集上的验证实验表明,改进后的网络具有良好的泛化性能。

1 Mask R-CNN 检测框架

Mask R-CNN 是两阶段网络,它在 Faster R-CNN 基础上添加一个并行分支以预测目标实例分割,可以实现端到端的多目标检测与分割。改进网络整体架构如图 1 所示。改进网络主要由 4 部分组成:1) MobileNet 特征提取网络;2) 区域建议网络(region proposal network, RPN);3) RoI Align;4) 像素级实例分割输出。

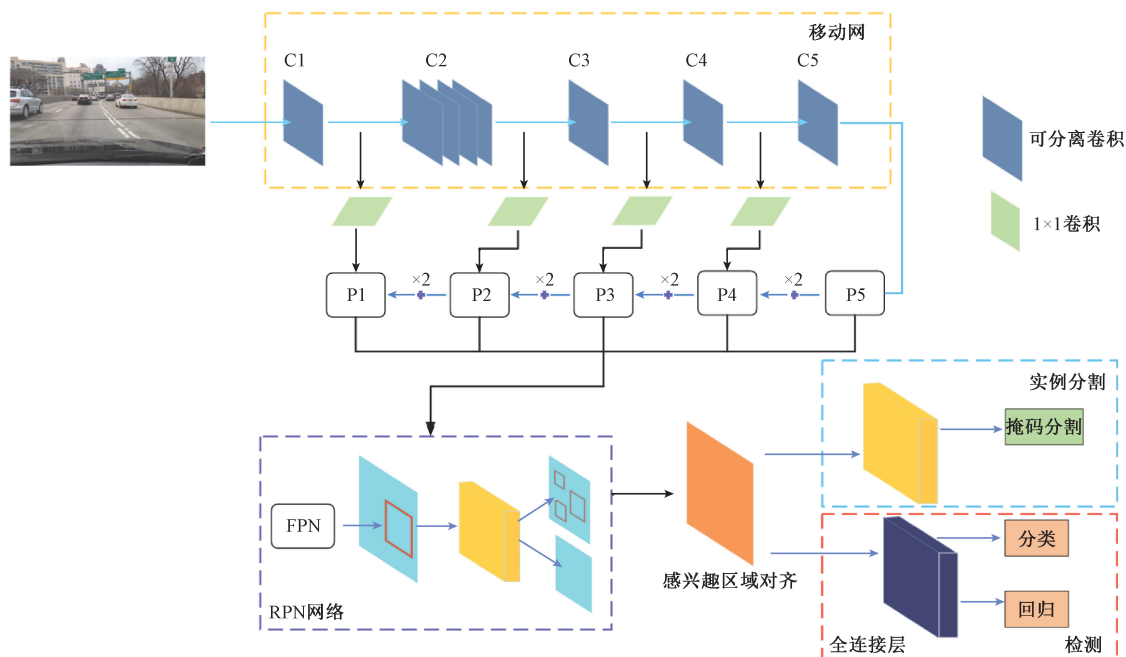


图 1 改进的 Mask R-CNN 架构

Fig. 1 Improved Mask R-CNN architecture

1.1 特征金字塔网络

Mask R-CNN 中特征金字塔网络 (feature pyramid network, FPN) 结合骨干网络, 实现对输入图像的特征提取及融合, 由 1×1 卷积层把图像特征横向连接至 FPN, FPN 通过上采样操作实现从低层至高层特征图的融合, 增强高层语义特征。对于不同尺度的行人、车辆、交通标志等目标, FPN 可以输出不同尺度的特征图, 来匹配交通目标检测中出现的多尺度目标。

1.2 RPN 网络

RPN 结构是全卷积网络, 如图 2 所示, 接收 FPN 最后一层卷积层输出的特征图 P1~P5, 输出一组对应的矩形区域建议。RPN 使用一个滑动窗口在特征图上滑动来提取区域建议, 对应的滑动窗口被映射成一个低维度的特征。该网络同时提供多个预测区域建议, 矩形区域建议将作为后续 RoI Align 层网络的输入。

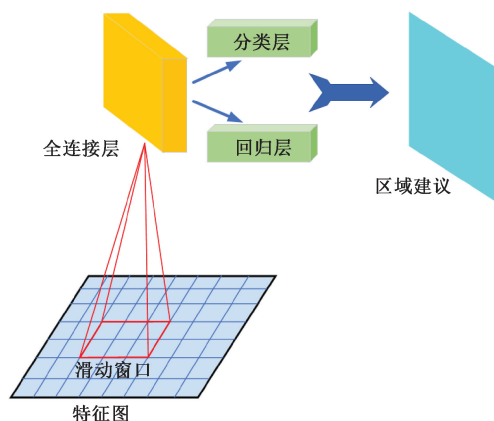


图 2 RPN 网络

Fig. 2 RPN network

1.3 RoI Align 层

RoI Align 层通过遍历每一个候选区域, 能以浮点数的形式计算回归框的位置, 保留像素坐标的小数部分而不作量化处理, 有效保留图像远景中小目标的像素信息。对于 RPN 输入的矩形区域建议, RoI Align 层以双线性内插和池化方法, 产生 7×7 的特征图作为分类与回归修正的输入, 14×14 的特征图是作为掩码分割的输入, 整个特征聚集过程是连续进行的。

2 改进 Mask R-CNN

2.1 骨干网络改进

Mask R-CNN 常用的骨干网络为 ResNet-50/101 残差网络。该网络对每层输入做一个 Reference (X), 通过加深层数并学习得到可优化的残差函数, 随着网络层数的不断加深, 在网络模型训练中, 大量的卷积计算使得权

重参数更新速度慢, 影响目标的快速检测。

为了实现交通场景多目标快速检测与分割, 本文利用 MobileNet 可分离卷积结构以及降低卷积计算成本的特点, 取代 Mask R-CNN 的 ResNet-50/101 网络。MobileNet 作为轻量级高效模型, 取消了网络层的叠加, 使用 3×3 逐通道卷积与 1×1 逐点卷积的结构组合, 实现目标特征的提取以及特征信息的融合。表 1 所示为 MobileNet 网络结构以及网络参数分布, 对于可分离卷积结构, 3×3 的逐通道卷积层参数占比少, 网络层的计算量以及参数占比主要集中在的 1×1 逐点卷积中, 逐通道卷积参数占比的减少使得 MobileNet 能够加快矩阵运算的速度, 提高了网络训练效率, 实现了快速检测。

表 1 MobileNet 网络结构与参数分布

Table 1 MobileNet network structure and parameter distribution

类型	计算量/%	参数占比/%
Conv 1×1	94.86	74.59
Conv DW 3×3	3.06	1.06
Conv 3×3	1.19	0.02
全连接层	0.18	24.33

改进后的骨干网络结合 FPN 可划分为自底向上和自顶向下两个路径。自底向上的 MobileNet 结构会生成 C1-C5 高饱和度的特征图。自顶向下路径是从高层特征图开始进行最邻近上采样。两个路径对图像点对点特征求和来得到 P1-P5 特征层, 经过整合之后的 P1-P5 会送入 RPN 网络进行后续计算。

RPN 网络中锚的作用是获取类标签。对获得的 9 个锚点用来预测 9 个候选区域, 每个候选区域用一个四维向量 (x, y, w, h) 来表示, x, y, w, h 分别代表边界框的中心点坐标及其高和宽。通过式 (1) 实现检测框与真实框的重叠校准

$$F[A_x, A_y, A_w, A_h] = [G'_x, G'_y, G'_w, G'_h] \quad (1)$$

式中: A_x, A_y, A_w, A_h 表示检测框的 4 个坐标向量; G'_x, G'_y, G'_w, G'_h 表示真实框的 4 个坐标向量。

计算边界框的回归:

$$\begin{aligned} t_x &= (x - x_a) / w_a \\ t_y &= (y - y_a) / h_a \\ t_w &= \log(x / x_a) \\ t_h &= \log(h / h_a) \end{aligned} \quad (2)$$

式中: x 和 y 表示边界框的中心坐标; t_x, t_y 表示平移因子; t_w, t_h 表示缩放因子。

网络中多任务的损失函数定义为:

$$L = L_{cls} + L_{box} + L_{mask} \quad (3)$$

式中: L_{cls} 为分类损失; L_{box} 为回归损失; 掩码分支对于每个

RoI 会有一个 Km^2 维度的输出; L_{mask} 为平均交叉熵损失。

实例分割损失即平均交叉熵损失,当标定框检测为目标类别,则使用该类别交叉熵作为误差来计算,其余损失不计算入内,定义为:

$$L_{mask} = -\frac{1}{n^2} \sum_{1 \leq i, j \leq n} [y_{ij} \log \hat{y}_{ij}^k + (1 - y_{ij}) \log (1 - \hat{y}_{ij}^k)] \quad (4)$$

式中: y_{ij} 为 $n \times n$ 区域中坐标 (i, j) 标签值; $\hat{y}_{ij}^k$ 为该点第 k

类预测值。

2.2 MobileNet 改进

MobileNet 的基本单元是深度级可分离卷积,如图 3 所示,由逐通道卷积和逐点卷积组成。卷积中的卷积核作用于所有的通道上,而逐通道卷积针对每个通道采用不同的卷积核,一个卷积核对应一个输入通道。逐点卷积的结构为 1×1 的卷积核,作用是将逐通道卷积的输出再结合,降低计算量和模型参数量。

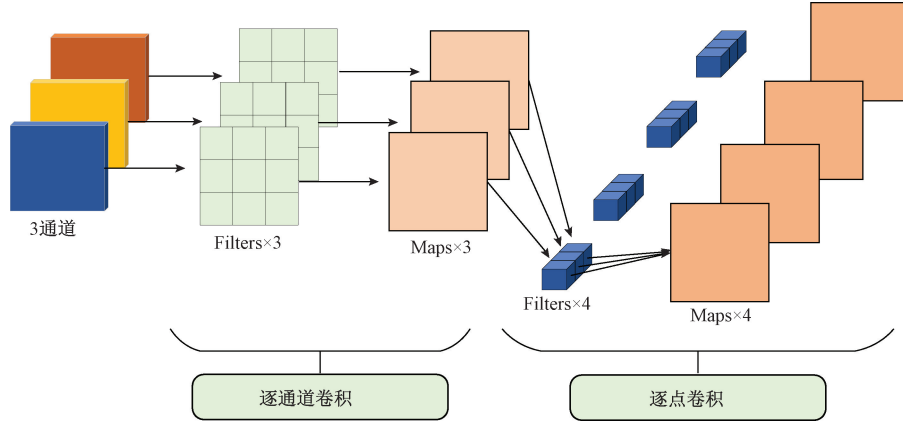


图 3 深度可分离卷积结构

Fig. 3 Depth separable convolution structure

针对交通场景出现的小目标快速检测与分割,对逐通道卷积与逐点卷积的卷积特征尺寸进行优化,扩张为 $[32, 64, 128, 256, 512, 1024]$,并优化步长等参数。改进后的 3×3 逐通道卷积计算如算法 1 所示,对于输入特征图 I ,加载卷积核权重 K ,改进后的特征尺寸用于计算每个输入通道特征图,遍历每个通道后,输出相应的特征图 O 。

算法 1: 逐通道卷积计算 (步长为 2)

输入: 输入特征图 I , 卷积核权重 K

输出: 输出特征图 O

设定: 输出通道 $outc$, 输出高度 $outh$, 输出宽度 $outw$

1. For $outc \in [1, 2, 3]$ Do
2. For $outh$ Do
3. For $outw$ Do
4. 加载特征图数据 I
5. 加载卷积核权重 K
6. For $\sum_{x=1}^3 \sum_{y=1}^3 [x][y] \cdot \ker[x][y]$ (x, y 为通道数与卷积核数, $\ker[x][y]$ 为卷积核权重系数)
7. 循环结束
8. 输出特征图 O

逐通道卷积与逐点卷积层之后的特征需要进行归一化操作,采用过滤器响应归一化 (filter response normalization, FRN) 取代批量归一化 (batch normalization, BN), 实现可分离卷积归一化的改进,加速逐通道与逐点层之间输入数据的归一化处理。FRN 可以加速网络的训练,能够降低网络特征处理误差,提升网络小批量训练的收敛速度。FRN 的操作建立在每个通道上的样本归一化,以二次范数的平均值推动模型训练加速,提升检测精度。

FRN 对通道的归一化计算方式:

$$v^2 = \sum_i x_i^2 / N \quad (5)$$

$$\hat{x} = \frac{x}{\sqrt{v^2 + \omega}} \quad (6)$$

式中: N 为维度; x 为一个维度的向量; v^2 为向量 x 的均方差数; $\hat{x}$ 为均值, ω 为常量。

归一化之后进行缩放和平移:

$$y = \gamma \hat{x} + \beta \quad (7)$$

式中: y 为仿射变换; γ 和 β 为学习参数。

给定输入特征图大小为 $L_F \times L_F \times M$, 输出特征图大小 $L_F \times L_F \times N$ 。其中, L_F 表示特征图的宽、高, M, N 表示指的是输入、输出通道数, 对与卷积 $D_K \times D_K$, 计算如下:

$$L_F \times L_F \times M \times N \times D_K \times D_K \quad (8)$$

对于逐通道卷积和逐点卷积的计算分别为:

$$L_F \times L_F \times M \times D_K \times D_K \quad (9)$$

$$L_F \times L_F \times M \times N \quad (10)$$

综合式(8)~(10),得出可分离卷积计算为:

$$L_F \times L_F \times M \times D_K \times D_K + L_F \times L_F \times M \times N \quad (11)$$

可分离卷积比卷积计算量减少比为:

$$\frac{L_F \times L_F \times M \times D_K \times D_K + L_F \times L_F \times M \times N}{L_F \times L_F \times M \times N \times D_K \times D_K} = \frac{1}{D_K^2} + \frac{1}{N} \quad (12)$$

3 实验分析

3.1 实验数据集

实验使用伯克利大学提供的无人驾驶开源数据集 BDD-100 K 和自己采集制作的实景图像,共同作为训练网络数据集,图像分辨率 1 280×720。样本类别涵盖交通场景中的行人、车辆、交通标志、交通灯等目标。使用 Labelme 标注软件将数据集打标获得 JSON 格式文件。数据集样本总数为 10 万张,数据集按比例进行划分,其中的 70% 作为训练样本集,20% 作为验证样本集,10% 作为测试样本集。数据集格式如表 2 所示。

表 2 数据集标注格式

Table 2 Dataset annotation format

数据文件	功能作用
Pic_img	网络训练原图
Info. yaml	标注位置文件
JSON File	标注后的转译文件
Label_viz	标注后的分割图像

3.2 实验结果

本实验在 TensorFlow 平台下进行,操作系统为 Windows 7,计算机处理器为 Intel © Core™ i7-6700CPU,内存 32 G,同时使用 NVIDIA GeForce RTX 2080Ti 来实现 GPU 加速训练。

训练改进网络的损失变化曲线如图 4 所示,设定改进网络训练参数,初始学习率为 1×10^{-3} ,训练周期为 160,动量系数为 0.9,一个训练周期为 1 000 步,分别在训练周期数为 80 和 120 时改变学习率为 2×10^{-4} 和 1×10^{-4} 。图 4(a) 表示改进网络整体损失变化曲线,图 4(b)~(d) 分别表示网络中 Bbox、Class 以及 Mask 的损失变化曲线。由图可知,4 种损失整体下降趋于平稳, Loss、Mask 以及 Bbox 3 种损失在前期下降较快,后期整体下降趋于平稳且收敛,Class 中期下降要快于前期,后期逐渐稳定收敛。

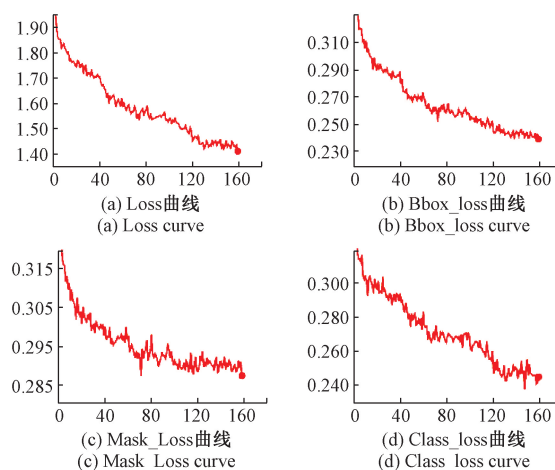


图 4 损失函数变化曲线

Fig. 4 Loss function curves

表 3 所示为网络训练所得到的模型体积与单张图像检测时间。ResNet-50/101 两种骨干网络所训练出的模型体积大小对应为 171.0 M 与 244.0 M, MobileNet 体积为 91.3 M。由体积比值公式(13)计算可得,相比较于 ResNet-50/101, MobileNet 模型体积分别降低了 46.6% 和 62.6%,表明改进网络的模型体积得到了极大的压缩,有效节省存储内存。在单张图像检测时间上, ResNet-101 的检测时间为 0.831 s,而 MobileNet 的检测时间为 0.402 s,提升检测速度效果显著,实现快速性需求。

$$r = 1 - \frac{M_1}{M_1 - M_2} \quad (13)$$

式中: M 表示网络模型体积; M_1, M_2 分别表示选取两个对比模型; r 为体积比。

表 3 训练网络模型大小

Table 3 Network model size

网络模型	模型体积/M	检测时间/s
ResNet-50 + Mask R-CNN	171.0	0.792
ResNet-101 + Mask R-CNN	244.0	0.831
MobileNet + Mask R-CNN	91.3	0.402

为了验证改进网络模型在不同交通场景下的多目标检测能力,选择训练周期数为 160 的网络模型进行测试,在 10% 测试样本集中分类选取样本开展多组对比实验,实验检测结果如图 5~8 所示。

图 5 为自然交通场景下多目标检测与分割结果,其中,图 5(a)~(d) 的交通场景分别为复杂公路、高速公路、城市公路和郊区公路。从检测结果可知,改进的 Mask R-CNN 网络模型能够有效实现车辆、交通标志、交通灯等目标的检测与分割,检测结果与实际判定结果相符合,且目标的标定框范围较精准。

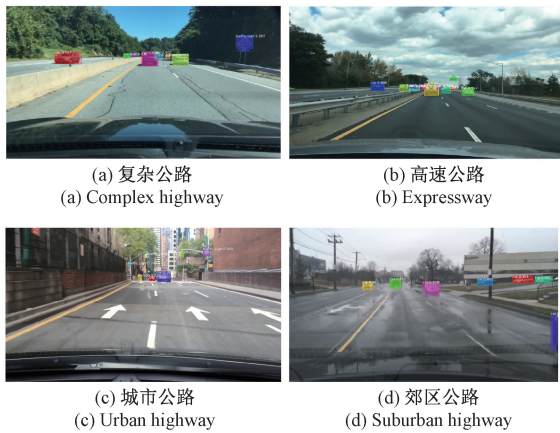


图 5 自然场景下检测结果

Fig. 5 Detection results in natural scenes

图 6 所示为光照交通场景下实验结果,图 6(a)为阴影遮挡情况、图 6(b)为光照遮挡情况。从图中可以看出,即使图像中车辆、交通标志等目标差异性受光照影响变化较大,颜色以及纹理受到干扰,本文提出的方法可以很好的完成检测以及分割任务。

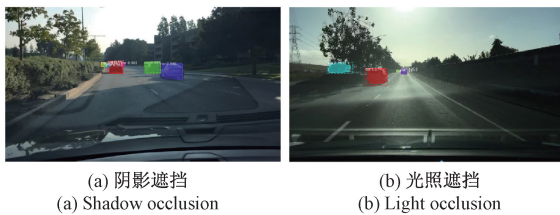


图 6 光照场景下检测结果

Fig. 6 Detection results under lighting scenes

图 7 为黑暗交通场景下的实验结果,针对图 7(a)双行道场景、图 7(b)单行道场景,本文方法对于黑暗场景下多目标也能够进行快速检测与分割,在应对低辨识度状况下,具有较好的可行性与鲁棒性。

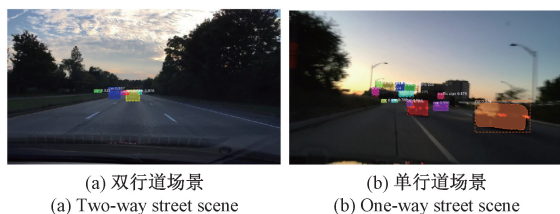


图 7 黑暗场景下检测结果

Fig. 7 Detection results in dark scenes

图 8 给出了天气、车辆运动导致模糊以及远景等不同场景检测结果,图 8(a)为雪景天气、图 8(b)为雨景天气,受到背景干扰,导致图像中的车辆目标特征不明显。

图 8(c)为模糊场景、图 8(d)为远景场景下,车辆目标会受到形变以及像素模糊影响。由实验结果可知,在天气、模糊与远景场景下,改进网络能够保证多目标快速检测与分割。

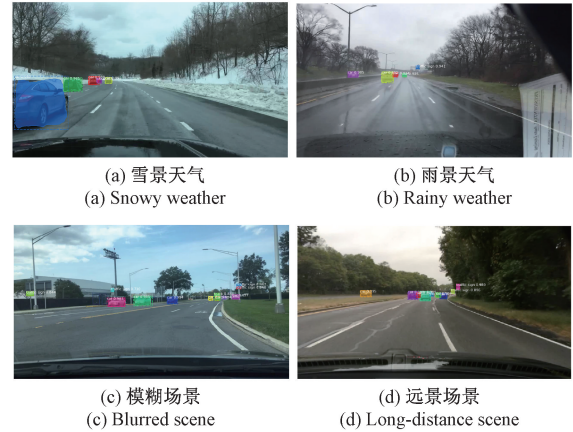


图 8 模糊及远景场景下检测结果

Fig. 8 Detection results in blurred and distant scenes

采用检测精度 P 、召回率 R 指标来衡量改进网络模型性能,计算公式如式(14)、(15)所示。

$$P = \frac{TP}{TP + FP} \quad (14)$$

$$R = \frac{TP}{TP + FN} \quad (15)$$

式中: TP 为正确识别并标定目标的数目; FP 为错误识别但标定目标的数目; FN 为错误识别且未能标定目标的数目。

网络训练时,分别选取迭代周期为 40、80、120 以及 160 的模型进行评估,网络训练损失与检测精度如表 4 所示,随着网络训练周期的增加以及学习率的改变,网络损失能够有效下降且趋于收敛,表明改进网络模型检测精度是不断提升的。

表 4 网络训练损失与检测精度

Table 4 Network training loss and detection accuracy

训练周期	网络损失	检测精度/%
40	1.652	73.3
80	1.537	79.4
120	1.493	83.8
160	1.364	84.7

选取周期数为 160 所训练的网络模型,从测试样本集中随机选取图像,对其进行测试实验以及计算检测精度,所得到的目标类别的检测精度如柱状图 9 表示。测试实验中平均检测精度达到 85.2%。说明改进网络能够

有效学习目标特征并以高精度完成交通场景下多目标的快速检测与分割。

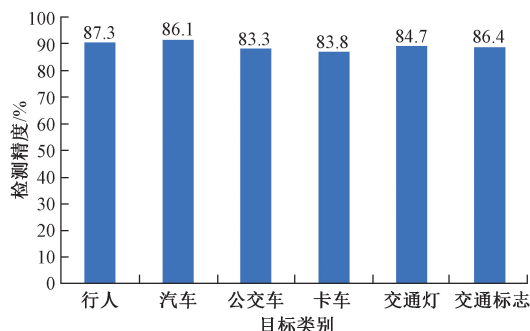


图9 交通场景多目标检测精度

Fig. 9 Multi-object detection accuracy in traffic scenes

本文以 Apollo Scape 和 NuScence 数据集进行迁移实验,两个数据集均为真实交通场景驾驶状态下采集的实时图像。实验从数据集中随机选取不同场景下的样本图像进行测试。迁移实验检测结果如图 10 所示。

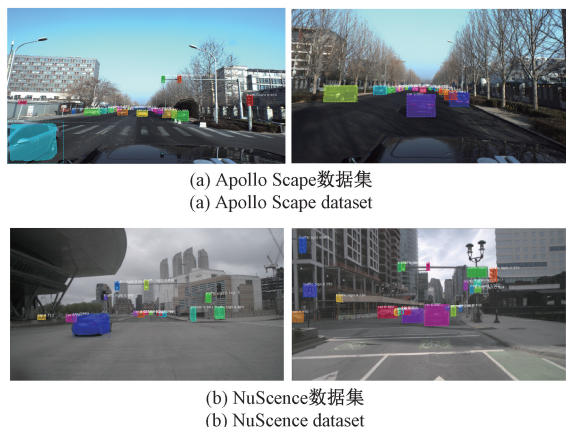


图10 迁移实验检测结果

Fig. 10 Migration test results

实验结果表明,对于未经训练的新样本数据,改进网络模型依然能够完成各种复杂交通场景下行人、车辆、交通标志等目标的快速检测与分割,证明所提方法是有效可行的。

4 结 论

交通场景多目标的检测与分割是智能驾驶研究中的关键技术。本文以 Mask R-CNN 为基础,以 MobileNet 作为骨干网络并结合 FPN 来实现目标特征的提取与融合,降低模型参数与体积,单张图像检测速度可达 0.402 s。改进可分离卷积网络层,在高低层之间充分利用网络特征数据,使得特征信息更加精确丰富,实现了交通场景多

目标快速检测与分割,平均检测精度可达 85.2%,证明改进网络的有效性。通过迁移实验表明改进算法具有良好的泛化能力,为智能驾驶以及无人驾驶提供了理论与技术支持。

参考文献

- [1] JIN M Z, ZHANG Q, WANG H, et al. Research on intelligent transportation system based on internet of things [J]. International Journal of Heavy Vehicle Systems, 2020, 27(3): 247-257.
- [2] LIAN Y Q, ZHANG G Q, LEE J Y, et al. Review on big data applications in safety research of intelligent transportation systems and connected/automated vehicles [J]. Accident; Analysis and Prevention, 2020, 146, DOI: 10.1016/j.aap.2020.105711.
- [3] 夏元清, 闫策, 王笑京, 等. 智能交通信息物理融合云控制系统 [J]. 自动化学报, 2019, 45(1): 132-142.
- XIA Y Q, YAN C, WANG X J, et al. Intelligent transportation cyber-physical cloud control systems [J]. Acta Automatica Sinica, 2019, 45(1): 132-142.
- [4] 李成美, 白宏阳, 郭宏伟, 等. 一种改进光流法的运动目标检测及跟踪算法 [J]. 仪器仪表学报, 2018, 39(5): 249-256.
- LI CH M, BAI H Y, GUO H W, et al. Moving object detection and tracking based on improved optical flow method [J]. Chinese Journal of Scientific Instrument, 2018, 39(5): 249-256.
- [5] CHENG X, ZHOU J M, ZHAO X M, et al. Vehicle ego-localization based on the fusion of optical flow and feature points matching [J]. IEEE Access, 2019, 7: 167310-167319.
- [6] ZHANG Y S, XU H, WU J Q, et al. An automatic background filtering method for detection of road users in heavy traffics using roadside 3-D LiDAR sensors with noises [J]. IEEE Sensors Journal, 2020, 20(12): 6596-6604.
- [7] ZHANG K, YANG K, LI S Y, et al. A difference-based local contrast method for infrared small target detection under complex background [J]. IEEE Access, 2019, 7: 105503-105513.
- [8] 刘晓, 崔光照, 李正周, 等. 基于视觉系统分层的小目标运动检测 [J]. 光学精密工程, 2019, 27(10): 2251-2262.
- LIU X, CUI G ZH, LI ZH ZH, et al. Small target motion detection based on layering of vision system [J]. Optics and Precision Engineering, 2019, 27(10): 2251-2262.
- [9] ZHANG Y S, ZHAO C H, SHI W, et al. Vehicles detection for illumination changes urban traffic scenes

- employing adaptive local texture feature background model [J]. IET Intelligent Transport Systems, 2018, 12(10): 1283-1290.
- [10] 孙伟, 施顺顺, 张小瑞, 等. 基于联合特征和压缩字典学习的车辆类型识别算法[J]. 汽车工程, 2018, 40(3): 259-264, 276.
- SUN W, SHI SH SH, ZHANG X R, et al. Vehicle type recognition algorithm based on joint features and compressed dictionary learning [J]. Automotive Engineering, 2018, 40(3): 259-264, 276.
- [11] 王中宇, 倪显扬, 尚振东, 等. 利用卷积神经网络的自动驾驶场景语义分割[J]. 光学精密工程, 2019, 27(11): 2429-2438.
- WANG ZH Y, NI X Y, SHANG ZH D, et al. Autonomous driving semantic segmentation with convolution neural networks [J]. Optics and Precision Engineering, 2019, 27(11): 2429-2438.
- [12] FANG F, LI L Y, ZHU H Y, et al. Combining Faster R-CNN and model-driven clustering for elongated object detection[J]. IEEE Transactions on Image Processing, 2020, 29: 2052-2065.
- [13] MOU L C, ZHU X X. Vehicle instance segmentation from aerial image and video using a multitask learning residual fully convolutional network [J]. IEEE Transactions on Geoscience and Remote Sensing, 2018, 56(11): 6699-6711.
- [14] WU W Q, YIN Y J, WANG X G, et al. Face detection with different scales based on Faster R-CNN[J]. IEEE Transactions on Cybernetics, 2019, 49(11): 4017-4028.
- [15] 郑少武, 李巍华, 胡坚耀. 基于激光点云与图像信息融合的交通环境车辆检测[J]. 仪器仪表学报, 2019, 40(12): 143-151.
- ZHENG SH W, LI W H, HU J Y. Vehicle detection in the traffic environment based on the fusion of laser point cloud and image information [J]. Chinese Journal of Scientific Instrument, 2019, 40(12): 143-151.
- [16] 范宜标, 卢玮, 傅智河. 自动驾驶场景下的鲁棒车辆检测[J]. 电子测量与仪器学报, 2018, 32(12): 60-65.
- FAN Y B, LIU W, FU ZH H. Robust vehicle detection for automatic driving [J]. Journal of Electronic Measurement and Instrumentation, 2018, 32(12): 60-65.
- [17] 刘丹, 马同伟. 结合语义信息的行人检测方法[J]. 电子测量与仪器学报, 2019, 33(1): 54-60.
- LIU D, MA T W. Pedestrian detection method based on semantic information [J]. Journal of Electronic Measurement and Instrumentation, 2019, 33(1): 54-60.
- [18] 徐谦, 李颖, 王刚, 等. 基于深度学习图像语义分割的机器人环境感知[J]. 吉林大学学报(工学版), 2019, 49(1): 248-260.
- XU Q, LI Y, WANG G, et al. Robotic environment sensing based on semantic segmentation by deep learning [J]. Journal of Jilin University (Engineering and Technology Edition), 2019, 49(1): 248-260.

作者简介



伍锡如, 2012 年于湖南大学获得博士学位, 现为桂林电子科技大学教授, 主要研究方向为机器人控制、复杂网络、深度学习。

E-mail: xiruwu520@163.com

Wu Xiru received his Ph. D. degree from Hunan University in 2012. He is currently a professor in the Guilin University of Electronic Technology. His main research interests include robot control, complex network and deep learning.



邱涛涛 (通信作者), 2018 年于河南理工大学获得学士学位, 现为桂林电子科技大学硕士研究生, 主要研究方向为深度学习、神经网络、计算机视觉。

E-mail: 18339171275@163.com

Qiu Taotao (Corresponding author) received his Bachelor's degree from Henan Polytechnic University in 2018. He is currently a graduate student in the Guilin University of Electronic Technology. His main research interests include deep learning, neural networks and computer vision.



王耀南, 博士生导师, 中国工程院院士, 现为湖南大学教授, 主要研究方向为智能机器人控制、模式识别、智能信息处理和图像识别。

E-mail: yaonan@hnu.edu.cn

Wang Yaonan Ph. D. supervisor, Academician of Chinese Academy of Engineering. He is currently a professor of Hunan University. His main research interests include intelligent robot control, pattern recognition, intelligent information processing and image recognition.