

面向无人驾驶场景下的道路多目标检测算法

牛文杰 伊力哈木·亚尔买买提

(新疆大学电气工程学院 新疆 乌鲁木齐 830046)

摘要 针对无人驾驶场景下目标检测算法误检率高的问题,设计一种改进 YOLOv3 的多目标检测算法。该文在原始特征提取网络 Darknet53 中引入分组卷积核替换标准卷积核,降低了卷积操作的计算量;改进原始 YOLOv3 的特征融合方法,使不同尺度的特征层融合更加充分,对遮挡目标和小目标的检测效果有明显提升;构建 CIoU 位置损失函数,提示网络收敛效果。实验结果表明,改进的 YOLOv3 算法平均精确度提高了 1.71%,误检率降低了 12%,明显优于原始算法。

关键词 无人驾驶 多目标检测 分组卷积 YOLOv3 CIoU 损失函数

中图分类号 TP391.4

文献标志码 A

DOI:10.3969/j.issn.1000-386x.2024.08.041

A MULTI-TARGET DETECTION ALGORITHM OF ROAD FOR UNMANNED DRIVING SCENE

Niu Wenjie Yilihamu · Yaermaimaiti

(School of Electrical Engineering, Xinjiang University, Urumqi 830046, Xinjiang, China)

Abstract Aimed at the problem of high false detection rates of object detection in unmanned driving scene, a multi-target detection algorithm with improved YOLOv3 is proposed. The groups convolution kernel was introduced into the original feature network Darknet53 to replace the original convolution kernel, which reduced the complexity of convolution operation. The original feature fusion was improved to make the fusion of different scales more fully, and it improved the detection effect of occluded and small targets. The CIoU loss function was constructed to make the network convergence better. Experimental results show that the average accuracy of the improved YOLOv3 algorithm is increased by 1.71%, and the false detection rate is reduced by 12%, which is significantly better than the YOLOv3 algorithm.

Keywords Driverless Multi-target detection Group convolution YOLOv3 CIoU loss function

0 引言

目标检测任务是无人驾驶系统中必要的一环,对于无人驾驶系统来说,目标检测相当于无人驾驶车辆的眼睛,通过感知前方车辆、行人情况得到道路信息,从而决定车辆运行状态,是无人驾驶领域不可缺少的部分。

传统的目标检测算法首先使用滑动窗口对图像进行特征提取,然后使用分类器对图像进行分类^[1]。比如 Hog 与 Svm 算法的结合^[2],以及 Haar 特征提取与

Adaboost 分类器进行目标检测^[3]。这些传统算法在检测精度能达到很好的效果,但在实际应用场景下却达不到检测速度的要求。对于无人驾驶方面的多目标检测来说,实时性是非常重要的指标要求,只有在运行速度上达到要求,才是保证安全驾驶的前提。

基于深度学习的目标检测算法可以同时检测路面状况的多种信息,无论从快速性还是精确度上都相对于传统检测算法更胜一筹。基于深度学习的目标检测算法可以分成两类:one-stage 和 two-stage 算法。two-stage 算法中最典型的是 R-CNN^[4] 系列算法,该系列算法的特点是网络结构比较复杂,精度高,但是运行时

间较慢。另一类 one-stage 目标检测算法有 SSD^[5-6]、YOLO^[7-9]等,其特点是将整个检测过程简化为回归问题,检测速度比较快。

YOLOv3 算法是基于深度学习的 one-stage 目标检测算法。该算法使用 Darknet-53 作为特征提取网络,采用多尺度融合方法,提升了对不同尺度目标的检测效果,但是在无人驾驶任务下,由于背景比较复杂,在远处车辆或者遮挡目标等情况下,存在将背景信息错误检测为车辆和人的问题,导致检测效果没有那么理想,错误检测率比较高。鉴于以上因素,本文提出在原始 YOLOv3 算法的基础上训练出一种改进型的算法结构。首先利用分组卷积设计出一种新的网络并行结构,减少网络的参数量并且提升特征提取效果;采用改进的特征融合方法,改善对不同尺度目标的检测效果;然后对于边界框的回归损失函数,改变原始 MSE 损失函数为 CIoU 损失函数,使网络收敛效果更加明显。最后通过实验证明在不改变网络原始深度的情况下可以有效减少计算量,降低错误检测率,提高准确率。特别对于遮挡目标和小目标有更好的检测效果。并且与 SSD、Faster-RCNN、YOLOv3 等主流算法进行比较,可以看出效果更好,鲁棒性更强。

1 YOLOv3 基本原理

2016 年 YOLO (You Only Look Once) 系列算法的提出把目标检测问题从候选框生成和分类问题简化成回归问题。YOLOv3 是 YOLO 系列的第三代算法,其预测过程是将图像信息输入卷积神经网络后,使用边界框回归策略预测出物体的边界框信息和类别,使用这种预测方法也加快了检测的速度,从而达到实时性的要求。

1.1 Darknet53 网络

YOLOv3^[10-11] 算法采用 Darknet53 作为主干特征提取网络,Darknet53 是一个由多个残差结构^[12]组合而成的深层卷积网络,引入残差结构能够防止网络过深时造成的梯度爆炸和消失现象。Darknet53 网络是由卷积层(Conv2d)连接批次归一化层(BN)和非线性激活函数(Leaky ReLU)组成的残差结构堆叠而成。主干网络组成的模块如图 1 所示,卷积层+批次归一化层+激活函数构成一个 CBL 模块,两个 CBL 模块组成一个 Res 模块,CBL 模块和 Res 模块堆叠成为 Darknet53 主干特征提取网络。

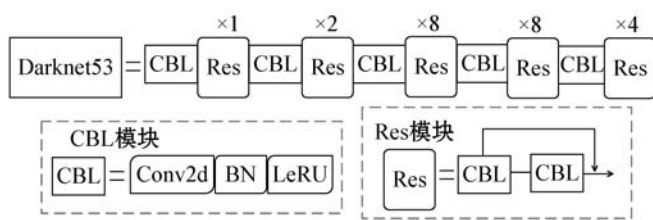


图 1 Darknet53 网络结构

1.2 检测过程

从 Darknet53 不同深度得到三个大小不同的特征层后,将三个特征层输入特征金字塔网络 (Feature Pyramid Networks, FPN), FPN 能使不同深度的特征层的语义信息进行融合,经过 FPN 网络后,输出三种不同尺度特征图,将不同大小的特征图分割成 $S \times S$ ($13 \times 13, 26 \times 26, 52 \times 52$) 的网格,每个网格对应 N 个预测框,其中预测框包含预测目标的坐标,目标得分和类别置信度信息。其中目标的坐标信息并不是直接得到预测框的坐标信息,而是通过式(1) - 式(5)转换。

$$b_x = \sigma(t_x) + C_x \quad (1)$$

$$b_y = \sigma(t_y) + C_y \quad (2)$$

$$b_h = p_h e^{t_h} \quad (3)$$

$$b_w = p_w e^{t_w} \quad (4)$$

$$\sigma(t) = \frac{1}{1 + e^{-t}} \quad (5)$$

式中: t_x, t_y 为网络预测坐标的偏移量; t_w, t_h 是预测尺度的偏移量; b_x, b_y, b_w, b_h 表示目标框的中心坐标和宽高; C_x 和 C_y 对应特征图每个网格左上角坐标; p_w 和 p_h 为预设在先验框映射到特征图的宽高。根据式(1) - 式(5)得到预测边界框的坐标 b_x, b_y, b_w, b_h , 此时预测框中含有大量重复和错误的预测目标,根据边界框预测目标得分和类别置信度,过滤掉得分低的边界框信息,最后通过非极大值抑制 (NMS) 得到最终预测目标。检测过程如图 2 所示。

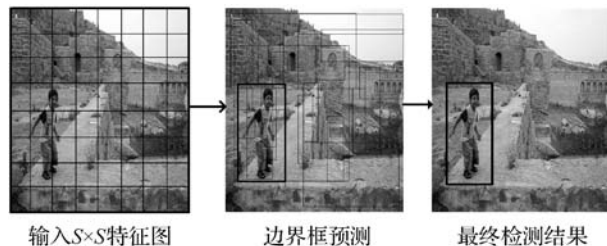


图 2 检测过程

2 改进 YOLOv3 的相关工作

2.1 G-darknet 主干网络

DarkNet53 网络由大量残差块堆叠而成,并且网络后端需要进行 FPN 网络的特征融合,虽然提取能力有

了提高,但导致有大量参数需要进行处理,从而会使检测速度降低。2012 年在 ImageNet 国际大赛上, Alex-Net^[13] 中引入了分组卷积^[14] 的思想,并且以远超第二名的成绩获得了冠军。分组卷积是介于普通卷积和深度可分离卷积的一个折中方案,引入分组卷积(Group Convolution)代替原始标准卷积,可以达到减少计算量的效果。利用组卷积核的思想,使每一组卷积负责输出层的一部分,这样将输出特征层通过多个卷积核得到。这也能够增强模型的泛化性能,并且降低了训练参数,使网络可以提取出更具有泛化性的特征。分组卷积降低参数的原理为,当输入特征图的尺寸为 $N_1 \times N_2 \times C_1$,卷积核尺寸为 $F_1 \times F_2 \times C_1$ 时,若输出特征图尺寸为 $N_1 \times N_2 \times C_2$,则每次标准卷积操作需要的参数量为 $F_1 \times F_2 \times C_1 \times C_2$ 。当使用分组卷积去替代标准卷积核之后,将输入的特征图按照通道数分成 M 组,可以得到每一组输入特征图的尺寸为 $N_1 \times N_2 \times C_1/M$,则对应每一组的卷积核尺寸为 $F_1 \times F_2 \times C_1 \times C_2/M$,经过不同组分别对特征图卷积之后,可以得到每组的输出特征图尺寸为 $N_1 \times N_2 \times C_2/M$ 然后将每组得到的 M 组进行拼接操作,最终得到标准卷积一样大小的特征图。标准卷积核的计算的参数为 $F_1 \times F_2 \times C_1 \times C_2$ 分组卷积的参数量为 $F_1 \times F_2 \times C_1 \times C_2/M$,通过上面的表述可以看出利用分组卷积可以减少卷积操作的计算量和参数量,达到减少计算复杂度的效果。分组卷积原理如图 3 所示。

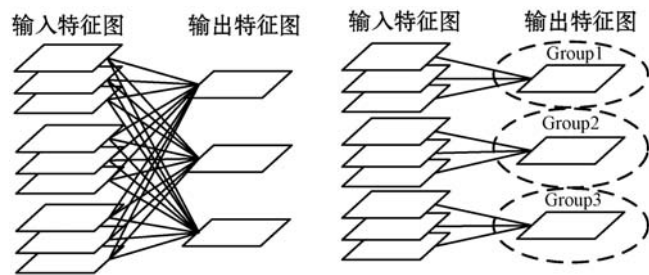


图3 分组卷积原理

为了使模型训练效果更加平滑,激活函数的改进是十分重要的。激活函数通过神经元的输入映射到输出端,激活函数能增加模型的非线性因素,从而达到线性模型达不到的泛化性能。我们使用 Mish 激活函数替换原始 YOLOv3 模型的 Leaky ReLU 激活函数,Mish 相比 Leaky ReLU 具有更好的平滑性,激活函数的平滑性可以更大程度地让网络模型接收信息,从而加强模型的泛化性能。我们将上述对于主干特征提取网络的改进,重新设计了一种新的残差结构 G-Res 模块,将残差结构中 3×3 卷积改进成组数为 2 的分组卷积,替换掉原始 Darknet53 主干提取网络的最后一组残差结构的 3×3 的卷积核。并且将网络结构中的原始 Leaky

ReLU 激活函数改为 Mish 激活函数,将分组卷积,归一化层和新的激活函数堆叠,组成一个新的主干网络结构 G-Darknet53,模块具体结构如图 4 所示。

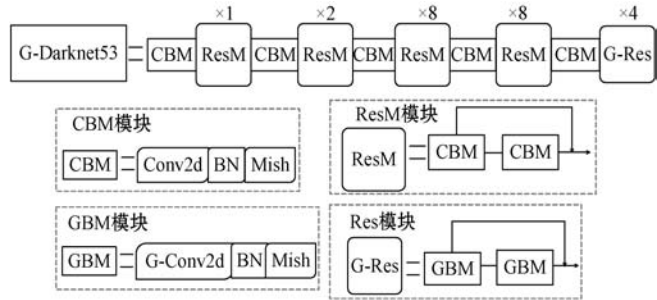


图4 G-Darknet 主干网络结构

2.2 改进的特征融合网络

经过特征提取后的图像数据在网络的不同深度能表达不同的语义信息。浅层网络能够提取到纹理和绝大部分共有的特征,而对于深层网络提取到轮廓等具有代表性的特征,实现不同尺度的特征层融合,能够更大程度地保留特征图的语义信息。FPN 网络是从顶部特征层向下部进行信息传递。在无人驾驶场景实验中,为了提高对于远处目标和被遮挡目标的检测能力,需要利用更多不同特征层的细节信息和语义信息,所以需要改进 FPN 网络。PANet^[15] (Path Aggregation Network) 是一种应用于实例分割的特征融合方法。FPN 为自顶向下,把顶层的特征信息传递下来,增强了语义信息,但是对于定位信息的传递效果不大。PANet 在自顶向下的特征融合后增加一个自底向上的金字塔结构,这样的结构利用底层特征图的定位信息,提升了网络的定位能力,改善了模型对于小目标物体和被遮挡物体的检测能力,PANet 网络结构如图 5 所示。

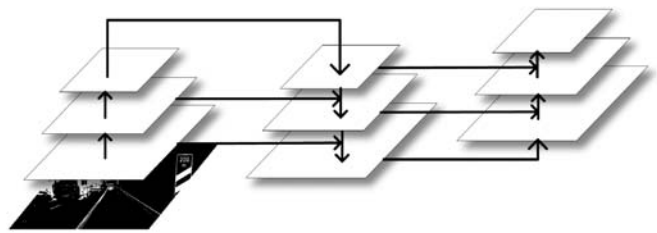


图5 PANet 网络结构

2.3 损失函数的改进

One-stages 目标检测最主要的任务是使先验框不断拟合目标框的位置和大小从而得到最好的检测效果,所以损失函数的优越性决定了网络训练的精度。用来衡量我们先验框与目标框拟合程度性能指标是交并比 IoU,原始 YOLOv3 算法是通过目标物体中心点的位置坐标和宽高进行 MSE 均方误差损失,但是这类损失函数把边界框位置信息的各个参数当成独立的变量

对待,并且与 IoU 指标并没有很强的关系。如果将 IoU 作为损失函数进行优化,会存在两个主要问题,首先当预测框与目标没有交集时,导致没有重叠部分,使损失函数没有梯度无法进行反向传播算法。其次由于位置信息不同,IoU 大小相同,这时的位置信息无法通过 IoU 表示出来,从而会导致不同的优化效果。所以我们引入 CIoU (Complete-IoU) 作为损失函数的一部分,CIoU 将目标框与回归框的距离、重叠的程度、尺度、宽高比都考虑进去,通过 CIoU 损失函数能够有效地使损失函数进行收敛。这样可以在目标框和预测框无重叠的情况下,通过其他尺度进行边界框的拟合任务,从而进行更加精准有效的回归。CIoU 的具体计算公式及其损失函数如式(6) - 式(8)所示。

$$v = \frac{4}{\pi^2} \left(\arctan \frac{\omega^2}{h^{gt}} - \arctan \frac{\omega}{h} \right)^2 \quad (6)$$

$$R_{CIoU} = \frac{\rho^2(b, b^{gt})}{c^2} + \alpha v \quad (7)$$

$$L_{CIoU} = 1 - IoU + \frac{\beta^2(b, b^{gt})}{c^2} + \alpha v \quad (8)$$

式中: R_{CIoU} 为 CIoU 计算的定义; L_{CIoU} 表示 CIoU 损失函数; b, b^{gt} 分别表示预测框与目标框的中心点的位置, $\rho^2(b, b^{gt})$ 代表两个坐标点的欧氏距离; c 表示能够同时包围目标框和预测框的对角线距离; v 用来衡量长宽比的相似性; $\omega^{gt}, h^{gt}, \omega, h$ 分别表示目标框和预测框的宽高; α 表示权重参数。CIoU 原理如图 6 所示。

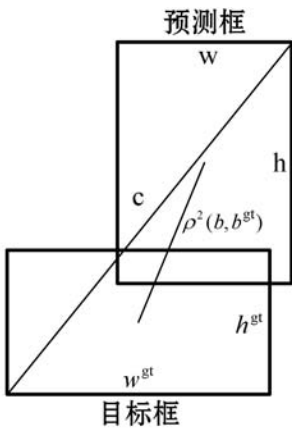


图 6 CIoU 原理

3 改进 YOLOv3 目标检测算法

本文采用 G-Darknet53 作为改进算法的主干特征提取网络。G-Darknet53 的具体网络结构是由 3×3 的卷积层为首层和 4 个分别包含 1、2、8、8 个 ResM 模块组成的残差结构,后经过由 4 个 G-Res 模块组成的残差模块进行信息流动,残差模块前使用 3×3 的卷积层 + BN 层 + Mish 激活函数的组合进行下采样操作。数据

经过 G-Darknet53 网络后输出三种大小不同的特征层,将特征层传递给改进的特征融合网络 PANet 进行特征融合。G-Res block4 模块输出的特征层经过上采样等操作和 ResM block3 输出的特征层进行融合,在融合之后的特征图继续进行卷积和上采样的操作,在于 ResM block2 的特征图进行融合。最后融合后的特征图一方面作为 YOLO Head 的输出,通过 YOLO Head1 进行尺度为 52×52 的预测。另一方面进行下采样的操作经过与 ResM block3 模块卷积操作后的特征图进行融合,通过 YOLO Head2 预测出 26×26 的结果。同理可得到 YOLO Head3 的输入,之后进行 13×13 预测结果。训练时 CIoU 损失函数对预测框的信息和目标框进行不断地拟合,将距离、重叠的程度、尺度、宽高比等都考虑进去,使拟合过程更加充分,拟合程度高。预测框不断接近目标框的位置等信息,最后得到实际预测框。网络结构如图 7 所示。

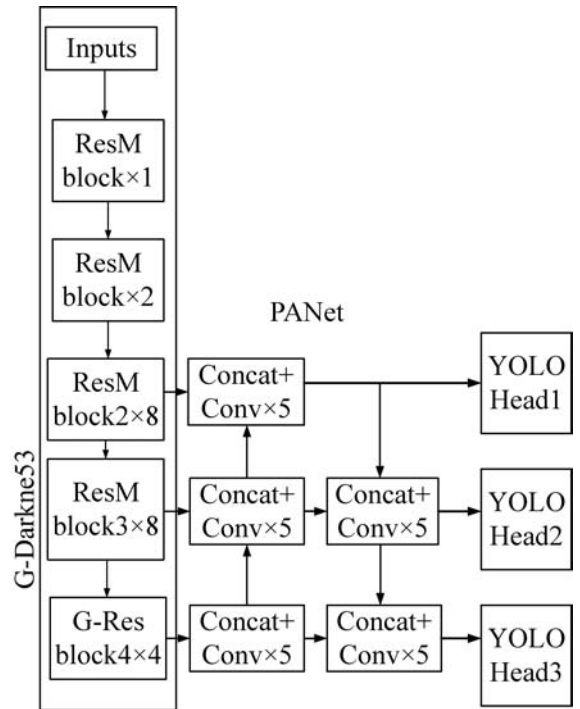


图 7 改进 YOLOv3 算法网络结构

与原始的 YOLOv3 算法比较,我们的改进网络在主干网络深层处使用分组卷积替换原始的标准卷积,从而使每层特征提取是由不同组的卷积核得到,这也增加了提取特征的泛化性,通过能够充分利用不同尺度的特征图信息,进行特征融合,改善了对于不同大小目标的提取能力提升算法的泛化性和精确度。算法基本流程如下:

- (1) 数据预处理。将图像数据尺寸转化为 416×416 大小后输入检测网络。
- (2) 特征提取。使用 G-Darknet53 网络对处理的图像数据进行特征提取。

(3) 特征融合。使用 PANet 网络对提取特征进行融合。

(4) 输出预测结果。将预测的目标框信息通过 YOLO Head 输出。

(5) 检测结果。通过非极大值抑制(NMS)优化目标框的检测结果,并将最终结果标注在图像中,输出目标检测结果。

4 数据集与实验设计

4.1 数据集处理

实验环境平台配置:Ubuntu 16.04.7 操作系统; GTX 1080 Ti 显卡;32 GB 内存空间;CPU 为 Intel E5-2678 处理器;采用 Python 为编程语言;深度学习框架为 PyTorch 1.5。为了验证改进 YOLOv3 算法的实用性,我们选用有高速公路,城市街区,乡村小道等场景的 KITTI^[16] 公开数据集,KITTI 数据集包含 7 000 幅道路场景图片。首先对原始数据进行处理,保留标签“VAN”“CAR”“Cyclist”和“Pedestrian”类别,合并成“Car”“Person”和“Truck”三个类别。同时将实验数据格式转换成 VOC2007 格式的数据集,其中按 8:1:1 的比例把数据分为训练集,验证集和测试集。在训练时将原始数据的大小转化为 416×416 的格式,并且对转换后的图像加入随机颜色填补缺失的像素以防止图像失真,数据处理后如图 8 所示。

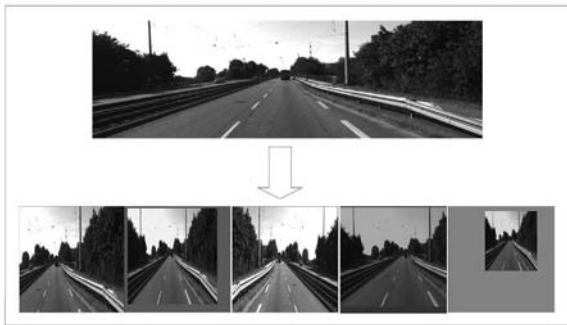


图 8 处理后的数据

4.2 训练过程

训练时对图像数据进行随机翻转、裁剪和色域调整等数据处理操作,从而增强模型的泛化能力与鲁棒性。并且为了更好地训练我们的改进网络,使用 COCO 公共数据集,训练我们的原始网络和改进网络,训练得到一定效果后,保留训练权重为预训练权重。我们采用迁移学习的思想载入 COCO 数据集上训练的预权重给我们的改进模型,使用迁移学习能够缩短模型的训练时间并且得到更好的学习效果。载入与训练权重后开始训练 KITTI 数据集,为了增加训练速度,我

们先冻结主干特征提取网络参数,在保持主干特征提取网络参数不变的情况下,训练后端网络的模型。训练时采用 Adam 优化器,初始学习率设定为 0.001,批量设定为 64。使用固定步长衰减(StepLR)策略,每次间隔一个 epoch,衰减率 gamma 设为 0.95,进行学习率的调整,进行 50 个 epoch 的训练。冻结训练后,在解冻主干特征提取网络,训练整个网络的参数以批量为 8,学习率降低至原来的 0.1 倍后继续训练 100 个 epoch,直至 loss 值收敛,并保留权重模型为最终的模型。

5 实验结果分析

5.1 参数量对比

模型训练完以后,载入最终训练权重作为最终模型进行检测。通过对比发现引入分组卷积后,网络模型参数量有明显降低,证明分组卷积的引入能够在不改变网络模型深度和宽度的情况下减少参数量,使模型变得更加精简。参数对比表如表 1 所示。

表 1 原始网络和改进网络的参数对比

算法名称	参数量	参数量大小/MB
YOLOv3	61 949 149	236.32
改进 YOLOv3	56 660 577	216.14

5.2 loss 曲线对比

改进后 YOLOv3 的损失函数使用 Ciou 损失函数,通过实验效果可以证明 Ciou 相比原始 MSE 损失函数具有更好的性能,能够使 loss 曲线较原始算法下降更快,而且最后 loss 值逐渐稳定并且比原始算法收敛效果更好,loss 下降曲线如图 9 所示。

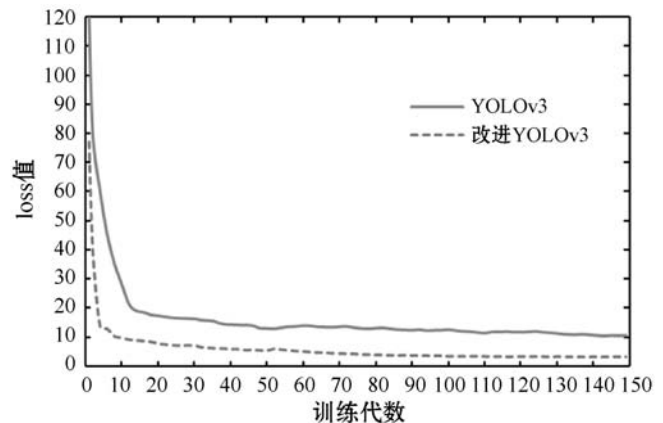


图 9 loss 曲线对比

5.3 误检率对比

为了验证改进 YOLOv3 算法能有效降低算法的误检率,我们使用假阳性率曲线作为对比,假阳性率即误

检率,通过误检率以判断算法是否过多将背景等信息错误检测成我们的目标信息。在无人驾驶检测任务中,根据误检率评价算法的检测性能。误检率(FPR)定义如式(9)所示。

$$R_{FP} = \frac{F_p}{F_p + T_N} \quad (9)$$

式中: T_p 是正样本正确预测为正类, F_p 为负样本被预测为正类, F_p 表示正样本错误被预测为负类, T_N 表示负样本正确预测为负类。通过误检率曲线可以得出,改进后的 YOLOv3 算法较原始算法的误检率很大改善。误检率曲线如图 10 所示。其中,MFPR 表示平均误检率,Car、Person、Truck 表示这几类单独的误检率。

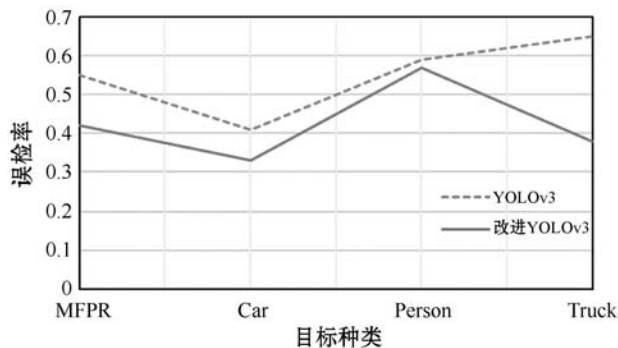


图 10 误检率曲线对比

5.4 mAP 进行对比

mAP 值是检测算法常用的衡量算法性能的指标,mAP 为精确率和召回率围成的曲线的面积。在改进网络模型彻底收敛后,使用测试数据集在原始网络模型进行检测,从实验结果可以看出模型最佳的平均 mAP 达到 70.32,将本文使用的数据集分别在 SSD、原始 YOLOv3、Faster-RCNN 训练进行对比,测试结果如表 2 所示,可以看出,本文改进的 YOLOv3 算法相比其他算法每一类的 mAP 均有提高,并且比原始 YOLOv3 在 Truck 类提高了将近 4%,这也证明 PANet 特征融合方法比原始特征融合方法更加有效。

表 2 不同算法 mAP 对比

检测算法	主干网络	mAP	Car	Person	Truck
SSD	VGG-16	38.29	56.12	32.50	26.26
Faster-Rcnn	Resnet-50	52.39	70.26	46.29	40.63
YOLOv3	Darknet53	68.61	75.88	58.33	71.63
改进 YOLOv3	GDarknet53	70.32	77.00	58.83	75.14

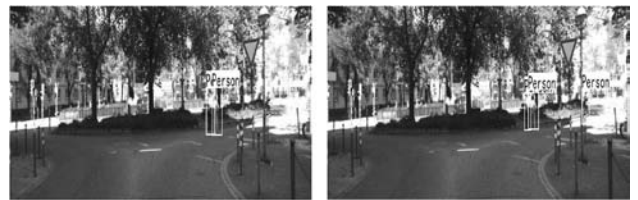
5.5 效果对比

原始和改进算法对比效果如图 11 所示。左边为原始 YOLOv3 算法检测结果,右边为改进 YOLOv3 算法检测结果,对于图 11(a)复杂场景下远方车辆,原始

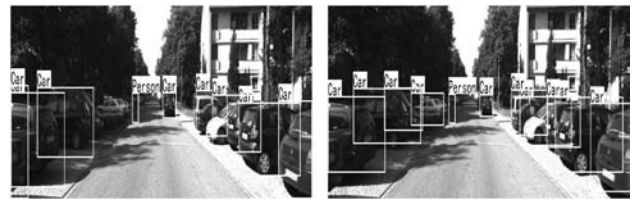
YOLOv3 算法并没有检测出来,但我们的改进的检测算法将其检测出来。对于图 11(b)行人的检测,对于图片右方的行人,原始算法并没有将其检测出来。对于图 11(c)、(d)、(e)来说在有遮挡的情况下,改进算法能够更有效地将遮挡物体检测出来。对于图 11(f)原始算法并没有将 Truck 正确分类,而我们的改进算法得到了正确的结果。综上所述,对于多数有遮挡的物体和复杂场景下的我们可以得到更好的效果,证明了算法的有效性。



(a) test 1



(b) test 2



(c) test 3



(d) test 4



(e) test 5



(f) test 6

图 11 原算法(左)和改进算法(右)在 KITTI 数据上检测结果

利用改进的 YOLOv3 算法对现实场景下的道路信息进行检测,如图 12 所示。可以看出,4 种不同的道路环境车辆都可以被检测到,且精确度较高,因此采用改进 YOLOv3 网络模型在现实场景下检测效果较好,具有一定的稳定性。

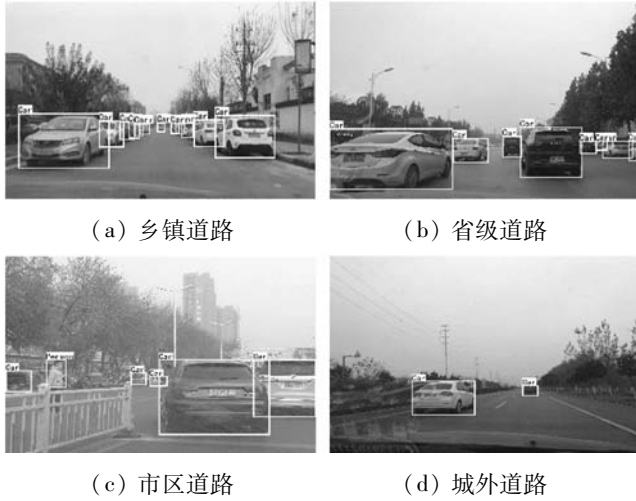


图 12 改进算法在现实场景检测结果

6 结 语

本文提出了改进 YOLOv3 的无人驾驶场景下多目标检测算法,用于解决误检率高和难检目标检测效果差的问题。首先,在原算法主干特征网络的基础上进行改进,引进了分组卷积和 Mish 激活函数设计新的主干特征提取网络结构 G-darknet53,在不改变原始网络深度和宽度的基础上减少了卷积核计算参数量,通过分组卷积核的思想得到更好的泛化性能。然后改进了原始特征融合方法(PANet),加强了特征提取的能力,改进了对于小目标和遮挡目标的检测能力。最后改进了损失函数,减少了预测框的回归误差。实验结果证明,改进算法能够降低原始算法的误检率,使 loss 值收敛效果更好,提高了对于车辆以及行人的准确率和召回率,改善了对于遮挡目标和远处目标的检测效果,因此对于自动驾驶目标检测算法有一定的应用价值,对于具有一定遮挡的物体也有不错的效果。但对于距离过远的车辆和行人检测效果还存在欠缺,下一步解决这个难题是重要的探索方向。

参 考 文 献

- [1] 李学宾. 基于树莓派的实时目标检测[D]. 武汉:华中科技大学,2019.
- [2] Radman A, Zainal N, Suandi S A, et al. Automated segmentation of iris images acquired in an unconstrained environment using HOG-SVM and GrowCut[J]. Digital Signal Processing,2017,64:60-70.
- [3] Bong K, Choi S, Kim C, et al. A low-power convolutional neural network face recognition processor and a CIS integrated with always-on face detector[J]. IEEE Journal of Solid-State Circuits,2018,53(1):115-123.
- [4] Girshick R, Donahue J, Darrell T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]//IEEE Conference on Computer Vision and Pattern Recognition,2014:580-587.
- [5] Liu W, Anguelov D, Erhan D, et al. SSD: Single shot multibox detector[C]//European Conference on Computer Vision Springer International Publishing,2016:21-37.
- [6] 王鹏,陆振宇,詹天明,等. 基于 PDSSD 改进型神经网络的小目标检测算法[J]. 计算机应用与软件,2021,38(1):149-156.
- [7] Redmon J, Divvala S, Girshick R, et al. You only look once: Unified, real-time object detection[C]//IEEE Conference on Computer Vision and Pattern Recognition,2016:779-788.
- [8] Redmon J, Farhadi A. YOLO 9000: Better, faster, stronger[C]//IEEE Conference on Computer Vision and Pattern Recognition,2017:6517-6525.
- [9] Redmon J, Farhadi A. YOLOv3: An incremental improvement[EB]. arXiv:1804.02767,2018.
- [10] 鞠默然,罗海波,王仲博,等. 一种改进的 YOLOv3 算法及其在小目标检测中的应用[J]. 光学学报,2019,39(7):245-252.
- [11] 卢官有,顾正弘. 改进的 YOLOv3 案件包裹中危险品检测算法[J]. 计算机应用与软件,2021,38(1):197-203.
- [12] Khan R U, Zhang X S, Kumar R. Analysis of ResNet and GoogleNet models for malware detection[J]. Journal of Computer Virology and Hacking Techniques, 2019, 15(4):29-37.
- [13] Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks[C]//Neural Information Processing Systems,2012:1097-1105.
- [14] 杨志贤,周宁宁. 基于分组卷积和特征图级联的轻量级目标检测[J]. 计算机应用研究,2020,38(5):1590-1594.
- [15] Liu S, Qi L, Qin H F, et al. Path aggregation network for instance segmentation[C]//IEEE Conference on Computer Vision and Pattern Recognition,2018:8759-8768.
- [16] Geiger A, Lenz P, Urtasun R. Are we ready for autonomous driving? The KITTI vision benchmark suite[C]//IEEE Conference on Computer Vision and Pattern Recognition,2010:3354-3361.