

YOLOv5-M3:基于轻量级网络的车辆检测模型

刘超阳^{1,2},范菁^{1,2},曲金帅^{1,2},左金花^{1,2},唐玉敏^{1,2}

(1. 云南民族大学 电气信息工程学院,云南 昆明 650500;

2. 云南民族大学 云南省高校信息与通信安全灾备重点实验室,云南 昆明 650500)

摘要:针对车辆检测模型参数量大,多尺度车辆目标检测困难、重叠目标、小目标和遮挡目标容易漏检的问题,提出了一种轻量化的YOLOv5-MobileNetv3检测算法。首先,利用 k -means算法来提高锚框维数聚类效果和检测速度。其次,将YOLOv5的原始骨干网络CSPDarknet53替换为MobileNetv3进行特征提取。再次,在检测网络中,用CIoU损失函数代替GIOU损失函数,快速准确地定位图像目标区域,生成边界框,预测目标类别。使用交叉熵损失函数(cross entropy loss)作为分类损失函数。最后利用 k -means算法来提高锚框维数聚类效果和检测速度。结果表明,与YOLOv5网络相比,改进后的YOLOv5-M3检测准确率提高了5.0%,模型参数数量减少了46%,训练时间减少44.9%。改进后的YOLOv5-M3网络更小、训练时间更短、更准确地识别出目标,提高了自动驾驶系统中车辆和行人的目标检测的准确性,也为实现智能交通系统提供了一种选择。

关键词:目标检测;YOLOv5;深度学习;MobileNetv3;自动驾驶

中图分类号:TP391 **文献标志码:**A **文章编号:**1672-8513(2024)06-0760-07

近年来,许多基于深度学习的算法已经展示了良好的目标检测性能。然而,在真实交通场景中检测物体仍然是一个挑战。一些研究人员使用复杂的模型来获得较高的交通目标检测性能。然而,由于车辆的车载计算资源有限,这些复杂的模型无法部署在嵌入式设备中,或者无法实现自动驾驶过程中的实时检测。当部署在车载计算单元上时,提高这种模型的检测精度仍然具有挑战性。

国内对自动驾驶目标检测已有一定的研究。Guo等^[1]提出了一种单阶段目标检测网络(YOLOv3-promote),用于检测城市复杂环境中的车辆和行人,首先,空间金字塔池化用于融合图像中的局部和全局特征,以更有效地检测图像中尺寸差异较大的目标;其次,在特征图中增加了注意力机制,从而增强特征网络区分目标对象和背景的能力;Zhang等^[2]提出了一种改进的YOLOv3目标检测算法YOLO-R。将YOLOv3原有的3个特征尺度增加到4个,以降低小目标的漏检率;并通过稀疏训练和稀疏训练后对模型中不重要的通道进行修剪,可以减小模型

尺寸,加快检测速度,防止过拟合;Liu等^[3]提出了在YOLOv4的第三个yolo检测头模块前面添加了空间金字塔池,以优化深层网络特征的提取。然后,在优化网络的基础上,采用剪枝策略简化目标检测算法Tidy YOLOv4。

基于以上目标检测算法的检测模型,模型复杂度高、计算量大、推理速度慢,难以在终端设备部署。针对以上问题,提出了基于轻量化网络的YOLOv5-M3算法,以减少模型参数量,提高模型推理速度,提高特征提取能力。经过模型优化和损失函数的调整后的YOLOv5-M3,能够有效地提高检测精度。从而成功地完成了自动驾驶系统中的车辆的检测。

1 网络结构设计

Yolov5-M3目标检测网络使用MobileNetv3^[4]改进了Yolov5的特征提取网络。与Yolov5s相比,Yolov5-M3的参数和计算量更少,适用于嵌入式设备。并且在使用数据增强后,提高了网络的检测精度。

收稿日期:2022-09-14。

作者简介:刘超阳(1997-),男,硕士研究生。主要从事深度学习研究。

通信作者:范菁(1976-),女,傣族,教授,硕士生导师。主要从事深度学习、计算机网络、智能控制研究。

总体架构从 3 个方面改进了原有网络的设计。一是使用轻量级骨干网络 MobileNetv3 代替 YOLOv5 的原始 CSPDarknet53 骨干网络进行特征提取。二是引入 CIoU 损失函数, 增加边界框的宽高比信息, 增强鲁棒性。DIOU_nms 预测框筛选机制用于提高重叠目标的筛选性能。最后将检测类别分类的 loss 由 soft max 替换为 sigmoid。三是通过 k -means 重新得到的锚框尺度, 能有效地提升了 YOLOv5 算法的性能和检测精度。这种组合可以提高网络的学习速率, 然后进行特征融合以减少数据丢失。

1.1 YOLOv5 - M3 网络

YOLOv5 网络的总体框架由 Anchors、Backbone、Neck 和 Prediction 组成。在输入模块中, 采用 Mosaic 数据增强、自适应锚框计算和自适应图像缩放。主干模块采用焦点结构和 BCSP 结构。颈部模块采用 FPN + PAN 结构。最后, 在预测模块中采用 GIoU^[5] 的损耗函数, 利用 NMS 非最大抑制函数^[6] 对目标进行检测和分类。结合 MobileNetv3 对 YOLOv5 骨干网进行了改进, 在保证检测速度和轻量级的前提下提高了检测精度。

MobileNetv3 是一个轻量级的神经网络。本文用 MobileNetV3 代替 CSPDarkNet - 53 作为 YOLOv5 的骨干网络。替换可以有效减少图像处理所需的内存, 加快车辆和行人等小目标的检测。MobileNetV3 - Small 融合了 MobileNetV1^[7] 的深度可分离卷积和 MobileNetV2^[8] 的带线性瓶颈的倒置残差结构, 并引

入了轻量级注意力机制, 从而降低了特征图的计算负荷, 加快了特征图的传播速度。MobileNet 作为基础网络, 可以提取图像特征, 在尽可能保证性能的同时减少参数量。因此, 网络非常适合移动设备。通过平衡时间和精度要求, 根据宽度和分辨率因子构建具有适当精度和速度的 MobileNet。网络结构的基本思想是将通道之间的相关性完全分离, 从而大大减少计算量和参数量。

YOLOv5 - M3 算法包括 2 部分。一是位于前端的 MobileNetv3 网络, 通过可以提取目标的初始特征; 二是后端多尺度特征检测网络, 获取前端网络骨干网输出 $19 \times 19 \times 255$ 、 $38 \times 38 \times 255$ 和 $76 \times 76 \times 255$, 3 个尺度的特征图, 然后送到本算法的 PAN + FPN 特征融合层。

在特征融合网络中, 本文使用了结合跨阶段局部连接结构的卷积模块 C3。经过 2 次上采样和 2 次下采样后, 对特征图大小相同的特征层进行通道融合, 发送到本文算法的检测层网络。最终输出 $19 \times 19 \times 255$ 、 $38 \times 38 \times 255$ 和 $76 \times 76 \times 255$ 的 3 个尺度的特征图用于计算 3 个尺度的预测框在不同条件下的典型特征。最后, 通过非极大值抑制 (NMS) 模块过滤掉重复的预测目标。该模型能有效提取特征图信息, 准确识别车辆和行人和位置; 具有位移不变性和检测速度快特点, 对变化的目标具有较好的鲁棒性。同时, 改进后的 YOLOv5 - M3 网络参数的数量比 YOLOv5 减少了 46%。

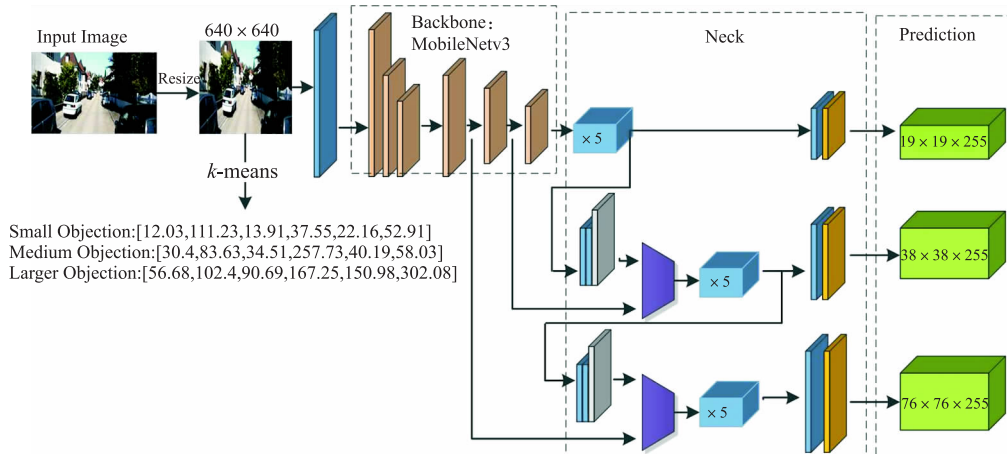


图 1 YOLOv5 - M3 网络结构

1.2 基于 K -means 聚类的先验框重选取

YOLOv5 的作者 COCO 数据集上使用了 k -means 聚类方法, 选择了 9 个聚类中心并将其分为 3 组, 以预测 3 种类型的目标: 大, 中, 小。但是 COCO 数据集总共有 80 个类别, 涵盖了从小的物体, 例如:

牙刷和老鼠, 到大的物体: 例如沙发, 大象和火车。但是, 本文仅检测车辆, 分为两类: 车辆和非车辆。因此, 根据 COCO 数据集聚类的锚框并不完全适用于本文。

K -means 算法^[9] 是一种典型的聚类算法, 它通

过最小化样本与聚类中心^[10]之间的距离,得到最优的 K 个聚类中心.本文使用 K -means算法对本文中使用的数据集进行重新聚类,以获得一个新的锚框.所得到的锚框是:(12.03, 111.23)、(13.91, 37.55)、(22.16, 52.91)、(30.4, 83.63)、(34.51, 257.73)、(40.19, 58.03)、(56.68, 102.4)、(90.69, 167.25)、(150.98, 302.08).

1.3 损失函数优化

1.3.1 边界框损失函数优化

损失函数是衡量网络预测结果质量的工具.本节阐述了YOLOv5网络的损失函数.此外,对原网络的损失函数进行了优化,检测任务的损失函数一般由分类损失函数和回归损失函数2部分组成.YOLOv5使用GIoU损失函数作为bbox(bounding box)的损失.与IoU^[11]损失函数相比,GIoU增加了错误检测的惩罚.检测误差越大,惩罚越重.在训练过程中,对于不同大小的预测框,可以得到相对较好的检测结果.但是当预测框和目标框重叠时,GIoU的效果会和IoU一样.

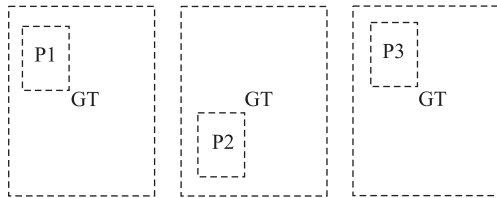


图2 GIoU退化为IoU示例

本文针对上述问题,采用CIoU对YOLOv5网络的损耗函数进行了修正.与GIoU计算预测框和目标框的交集和并集不同,CIoU计算预测框和目标框的中心点之间的欧氏距离,因此CIoU可以解决使用GIoU时出现的问题.预测框中有3个重要因素,即预测框和目标框中心点之间的距离、重叠区域和纵横比.当预测框和目标框重叠时,损失值大于GIoU的CIoU可以更好地描述当前位置信息.CIoU同时考虑预测框和目标框的重叠区域以及它们的中心点之间的距离.边界框纵横比的一致性是另一个重要的几何因素,CIoU可以使2个中心点之间的距离归一化,从而加速网络的收敛.

$$L_{CIoU} = \frac{\rho^2(X, Y)}{b^2} + \alpha v. \quad (1)$$

$$v = \frac{4}{\pi^2} \left(\arctan \frac{Y_w}{Y_h} - \arctan \frac{X_w}{X_h} \right)^2. \quad (2)$$

$$CIoU = 1 - IoU + L_{CIoU}. \quad (3)$$

其中, L 为交叉熵损失函数, X 表示预测框, Y 表示目

标框. ρ 表示 X 和 Y 的2个预测框的中心点之间的欧式距离, b 表示可以同时包含 X 和 Y 的最小闭合区域的对角线距离, α 是平衡比参数, v 表示相似度参数测量预测框和目标框的长度和宽度.

1.3.2 分类损失函数优化

YOLOv5的置信度损失函数使用二元交叉熵损失函数(BCE with logits loss)来计算样本的类概率损失的,可以用于多标签分类的,一个目标可以属于一个或者多个类别且类别存在一种包括关系.但是通常对于分类任务来说,输出的标签是互斥的.

本文针对上述问题,采用交叉熵损失函数(cross entropy loss)代替原始网络中的二元交叉熵损失函数来解决数据集中的类别具有互斥关系的问题.交叉熵损失函数中将会用到softmax函数将所有类别的预测值转换为总和为1的概率值,并分类为概率最高的那一类.进而解决类别互斥的问题,加快网络的收敛.

$$\text{Loss}(p, q) = -X_x(p(x) \log q(x)). \quad (4)$$

其中 p 为真实标签值, q 为预测值.

2 实验结果和分析

2.1 实验配置

本实验的模型训练的实验环境为: Intel(R) Xeon(R) Gold 5218R CPU, 软件平台 Ubuntu, 深度学习模型采用 Pytorch 搭建, CUDA10.1A, CUDNN10.1 以支持 GPU 使用. 对改进前后的 YOLOv5 算法分别进行训练, 训练模型的参数设置: Epochs 设为 300, Batch size 是指一次训练的样本数目, 与显卡的显存大小有关, 将其设为 15, 输入的分辨率值为 640 × 640, 初始学习率为 0.01, 动量为 0.937, 预设衰减系数 0.0005, 训练时间为 3.215 h.

2.2 数据集

KITTI 数据集^[12-13]由德国卡尔斯鲁厄技术研究所和丰田美国技术研究所共同创建. KITTI 包含城市、农村和高速公路等多种真实场景的图像数据, 每幅图像包含车辆和行人以及各种阴影、不同的光照、遮挡和截断, 为算法的鲁棒性提供了参考. KITTI 原始数据集的标签分为 8 类: 汽车、货车、卡车、行人、行人(坐着)、自行车、有轨电车和杂项. 但是, 由于车联网应用中自动驾驶的主要目标是检测车辆和行人的目标, 本文将原来的八类标签改为两类, 将汽车、货车、卡车、自行车和有轨电车划分为车辆; 行人和行人(坐着)划分为行人, 并删除了杂项类别. 最后两个类别是汽车和行人. 本文选取了数据集中的

7 481 幅图像作为实验数据, 并分配了数据集的十分之一作为验证集.

将数据集里的图片模仿 VOC2007 数据集格式, 利用 Labeling 标注软件依次对这些图片中的车辆进行标注外围框, 转化为训练所需要的 xml 格式.

2.3 评价标准

IoU、召回率、精确度和平均精确度 (mAP) 用于评估改进后的 YOLOv5.

IoU 是模型生成的 ground truth 与 predicted bounding 的交集, 如式 (5) 所示, 其中 B_{pb} 为预测边界框; B_{gt} 为目标边界框; $area()$ 是目标边界框和预测边界的联合形状或交集形状的面积. 当满足 IoU 大于阈值 TIOU 的条件时, 认为边界框是正确的结果. 基于此原理测量目标边界框与预测值之间的相关性; 相关性和值服从比例关系.

$$IoU = \frac{area(B_{pb} \cap B_{gt})}{area(B_{pb} \cup B_{gt})} \geq T_{IoU}. \quad (5)$$

召回率 (Recall): 召回率仍然是针对某一个具体类别而言的, 即预测边界框和所有目标边界框的比值, 公式为:

$$Recall = \frac{TP}{TP + FN} \times 100\%. \quad (6)$$

精准率 (Precision): 精准率是针对某一个具体类别而言的, 用于描述预测出来的正例占所有正例的比率, 公式为:

$$Precision = \frac{TP}{TP + FP} \times 100\%. \quad (7)$$

但一般情况下召回率和精准率很难都维持在高水平, 由此就需要一个参数来综合这 2 个参数, 使用 mAP 值来衡量检测网络的算法性能, 其适用于多标签图像分类, 计算公式如式 (8).

$$mAP = \frac{\sum_{k=1}^N P(k) \Delta R(k)}{c}. \quad (8)$$

式中的 N 表示验证集中的样本个数, $P(k)$ 是精准率 Precision 在同时识别 k 个样本时的大小, $\Delta R(k)$ 表示召回率 Recall 在检测样本个数从 $k - 1$ 个变为 k 个时的变化情况, c 则是多分类检测任务重类别的个数.

2.4 实验结果分析

图 3 是模型训练过程中的损失降低曲线前后对比, 改进后的损失函数收敛的速度更快, 损失值降得更低.

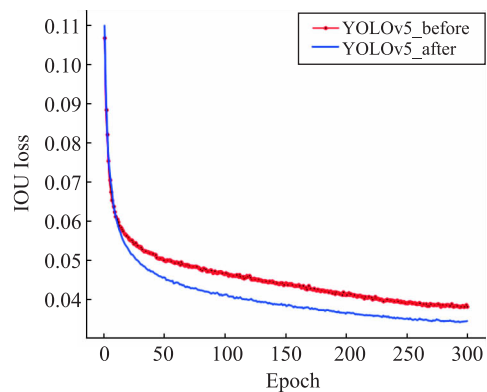


图 3 改进前后的 YOLOv5 算法损失函数对比图

如图 4 所示, 可以看出来改进后的 YOLOv5 算法的 mAP 曲线比原来的网络有所提高, 并且与原来的算法相比精度也有所提升. 改进前后的 YOLOv5 算法的各种指标如表 1 所示. 从表 1 中可以看出, 文章提出的改进方法将模型的 mAP 提升了 5.0%, 达到了 88.8%. 在精度和召回率方面, 相比较原来的方法分别提升了 7.1% 和 2.6%, 证明了文章的方法能够有效地提高对车辆的目标检测能力.

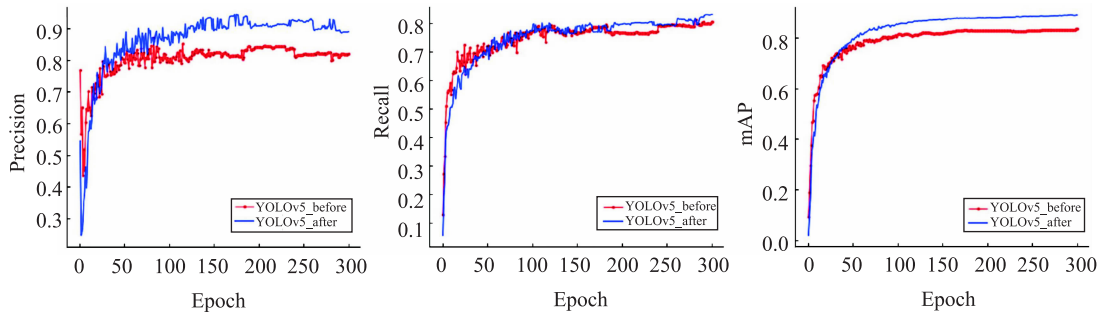


图 4 改进前后的 YOLOv5 算法实验对比图

表 1 综合指标测试结果

模型	P/%	R/%	mAP/%	时间/h	Weight/MB
YOLOv5	81.8	80.3	83.4	5.832	13.6
YOLOv5 - M3	88.9	82.9	88.8	3.215	7.3

从表 1 中可以看出改进后的 YOLOv5 - M3 不光准确率有所提升,而且训练的时间更短,模型更小.

KITTI 测试集用于车辆检测测试,各种复杂环境下的测试结果如图 5~9 所示.其中,(a)和(b)分别基于原始 YOLOV5 和改进 YOLOV5 进行对比.图 4 显示,在阴影环境中,原始 YOLOV5 检测不阴影环境下的车辆物体,存在漏检的问题.但改进后的 YOLOV5 已经实现了对所有车辆物体的检测,并在一定程度上提高了置信度得分.从图 5 可以看出,原始 YOLOV5 存在对小型车辆物体检测缺失和定位不准确的情况,改进后的 YOLOV5 实现了对多尺度车辆物体的有效检测和准确定位.从图 6 可以看出,

多个车辆物体不同程度地被阻塞,原始 YOLOV5 对严重阻塞的车辆物体造成了大量的检测缺失,可以观察到车辆物体的远距离定位不准确.相比之下,改进后的 YOLOV5 系统已经实现了对所有汽车和货车的有效检测.图 7 显示车辆物体位于道路交叉口,是典型的交通事故易发区域.原 YOLOV5 仅检测具有明显特征信息的车辆对象,改进后的 YOLOV5 正确检测所有车辆物体,有效提高相应的置信度分数.图 8 显示,在交通堵塞环境中,物体密度较高,且大多是回顾性的.原 YOLOV5 导致远视场对车辆物体的检测缺失,改进后的 YOLOV5 仍然实现了对所有车辆物体的有效检测.



图 5 基于原始和改进的 YOLOv5 在阴影环境下的车辆检测测试结果



图 6 基于原 YOLOv5 和改进 YOLOv5 的多尺度目标车辆检测测试结果



图 7 基于原始 YOLOv5 和改进 YOLOv5 的车辆遮挡检测结果



图 8 基于原 YOLOv5 和改进 YOLOv5 的道路交叉口车辆检测测试结果



图 9 基于原始和改进 YOLOv5 的交通堵塞环境下车辆检测测试结果

车辆检测测试结果表明,改进后的 YOLOV5 网络的检测性能具有很大的优势,这主要归功于

YOLOv5 基础结构和损失函数的改进.所提出的车辆检测算法对于复杂的交通环境和道路场景具有优

异的鲁棒性和环境适应性,检测精度进一步提高。

将该算法与其他方法进行了比较,以检验车辆检测算法的技术水平.表3显示了基于KITTI数据集的算法性能比较统计数据。

表3表明,所有算法都在KITTI数据集的基础上进行了车辆检测实验.mAP值是模型性能比较的主要评价指标.与现有的基于深度学习的检测方法相比,改进的YOLOV5网络模型使算法能够同时获得精度和实时性,有利于实现快速准确的车辆自动检测.同时,也减少了复杂环境中漏检的概率.这对于保证智能车在真实交通场景中的准确、全面运行具有重要意义,有助于大大减少交通事故的发生,充分保护人们的生命财产安全。

从图9可以看到,改进模型后的YOLOv5可以检测到更多的车辆物体,减少漏检率并且对于遥远的微小物体车辆也具有良好的检测结果。



图10 改进前后YOLOv5在测试集上的对比结果(左边改进后YOLOv5,右边原始YOLOv5)

表3 基于KITTI数据集上与其它算法的性能比较统计

模型	Car	Person	mAP/% 0.5:0.95
SSD ^[14]	85.1	48.1	61.3
YOLOv3 ^[15]	79.5	83.1	80.5
GaussianYOLOv3 ^[16]	87.3	79.9	83.6
SqueezeDet ^[17]	82.9	70.4	76.7
SqueezeDet + ^[17]	85.5	73.7	80.4
Conf - MobileDet ^[18]	90.4	69.5	79.8
YOLOv5	91.0	80.6	83.4
YOLOv5 - M3	96.7	81.0	88.8

图10是改进前后的YOLOv5模型在测试集上的比较结果.在不降低检测速度的情况下,所提模型提高了模型在目标检测中的精度和召回率.与其他模型相比,改进后的YOLOv5模型可以更有效地用于现实场景中的微小车辆目标检测。

3 结语

在准确率和工作效率方面,YOLOv5算法具有突出的性能优势,对于确保智能汽车在真实交通场景中的准确高效运行具有重要意义,有利于大幅度降低交通事故发生率,充分保护人民生命财产安全.未来,可以继续关注极端条件下的车辆检测算法和算法的实现,进一步提升算法的综合性能和实用意义。

参考文献:

- [1] GUO M, XUE D, LI P, et al. Vehicle pedestrian detection method based on spatial pyramid pooling and attention mechanism[J]. Information, 2020, 11(12): 583.
- [2] ZHANG X, DING S. Research on traffic target detection method based on improved YOLOv3[C]//2021 4th International Conference on Intelligent Autonomous Systems (ICoIAS). IEEE, 2021: 39-43.
- [3] LIU X, YAN Y, GAN H. Research on pedestrian detection algorithm in driverless urban traffic environment[C]//

MATEC Web of Conferences. EDP Sciences, 2021, 336:06002.

- [4] ZHAO L, WANG L. A new lightweight network based on MobileNetV3[J]. KSII Transactions on Internet and Information Systems (TIIS), 2022, 16(1): 1-15.
- [5] REZATOFI H, TSOI N, GWAK J Y, et al. Generalized intersection over union: a metric and a loss for bounding box regression[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2019: 658-666.
- [6] ORO D, FERNÁNDEZ C, MARTORELL X, et al. Work-efficient parallel non-maximum suppression for embedded GPU architectures[C]//2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2016: 1026-1030.
- [7] HOWARD A G, ZHU M, CHEN B, et al. Mobilenets: efficient convolutional neural networks for mobile vision applications[J]. arXiv preprint arXiv:1704.04861, 2017.
- [8] SANDLER M, HOWARD A, ZHU M, et al. Mobilenetv2: inverted residuals and linear bottlenecks [C]//Proceedings

- of the IEEE Conference on Computer Vision and Pattern Recognition. 2018:4510 – 4520.
- [9] 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.
- [10] ARTHUR D, VASSILVITSKII S. K – means + + : the advantages of careful seeding[R]. Stanford, 2006.
- [11] GEIGER A, LENZ P, URTASUN R. Are we ready for autonomous driving? the kitti vision benchmark suite [C]//2012 IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2012:3354 – 3361.
- [12] GEIGER A, LENZ P, STILLER C, et al. Vision meets robotics: the kitti dataset [J]. The International Journal of Robotics Research, 2013, 32(11):1231 – 1237.
- [13] WU B, IANDOLA F, JIN P H, et al. Squeezedet: unified, small, low power fully convolutional neural networks for real – time object detection for autonomous driving [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops. 2017:129 – 137.
- [14] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. Yolov4: optimal speed and accuracy of object detection [J]. arXiv preprint arXiv:2004.10934, 2020.
- [15] REDMON J, FARHADI A. Yolov3: an incremental improvement [J]. arXiv preprint, 2018:1804.02767.
- [16] CHOI J, CHUN D, KIM H, et al. Gaussian yolov3: an accurate and fast object detector using localization uncertainty for autonomous driving [C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. 2019:502 – 511.
- [17] 刘超阳, 曲金帅, 范菁, 等. 基于改进 YOLOv5 算法的车辆目标检测 [J]. 云南民族大学学报(自然科学版), 2022, 31(6):749 – 754.
- [18] LI W, LIU K. Confidence – aware object detection based on mobileNetv2 for autonomous driving [J]. Sensors, 2021, 21(7):2380 – 2380.

YOLOv5 – M3 : a lightweight network-based object detection model for vehicles and pedestrians

LIU Chao-yang^{1,2}, FAN Jing^{1,2}, QU Jin-shuai^{1,2}, ZUO Jin-hua^{1,2}, TANG Yu-min^{1,2}

(1. School of Electrical and Information Engineering, Yunnan Minzu University, Kunming 650500, China; 2. Yunnan Provincial Key Laboratory of Information and Communication Security Disaster Recovery, Kunming 650500, China)

Abstract: In view of the large amount of parameters of the vehicle detection model, the detection of multi – scale vehicle objects is difficult, and the overlapping objects, small objects and occluded objects are easily missed. A lightweight YOLOv5 – MobileNetv3 detection algorithm is proposed. First, the K – means algorithm is used to improve the effect and detection speed of anchor box dimension clustering. Second, replace the original backbone network CSP-Darknet53 of YOLOv5 with MobileNetv3 for feature extraction. Thirdly, in the detection network, the CIOU loss function is used instead of the GIOU loss function to quickly and accurately locate the target area of the image, generate bounding boxes, and predict the target category. Use Cross Entropy Loss as the classification loss function. Finally, the K – means algorithm is used to improve the effect and detection speed of anchor box dimension clustering. The results show that compared with the YOLOv5 network, the improved YOLOv5 – M3 improves the detection accuracy by 5.0%, reduces the number of model parameters by 46%, and reduces the training time by 44.9%. The improved YOLOv5 – M3 network is smaller, has shorter training time, and recognizes objects more accurately, which improves the accuracy of object detection for vehicles and pedestrians in autonomous driving systems, and also provides an option for implementing intelligent transportation systems.

Key words: object detection; YOLOv5; deep learning; MobileNetv3; autonomous driving

(责任编辑 梁志茂)