

基于 YOLOv3-tiny 的二轮车头盔检测

杨国亮 李世聪 邹俊峰 龚家仁

(江西理工大学电气工程与自动化学院 江西 赣州 341000)

摘要 针对二轮车驾乘人员头盔佩戴问题,提出一种基于 YOLOv3-tiny 的轻量化头盔检测模型。将原始模型主干网络进行轻量化处理,减少检测模型的参数量,在网络中添加 U 型特征二次融合模块,引入关于边框距离的 DIoU 损失函数,用于提高检测模型的特征提取能力和识别精度。在测试集上的实验表明,改进后的模型相比原 YOLOv3-tiny 模型表现出更高的查全率和 mAP 及 F1 指标,且在保持较小参数量的同时,具有优于深度网络 YOLOv3 的检测性能。

关键词 头盔检测 轻量化网络 特征融合 边框损失函数

中图分类号 TP391.41

文献标志码 A

DOI:10.3969/j.issn.1000-386x.2024.05.023

HELMET DETECTION OF TWO WHEELED VEHICLE BASED ON YOLOV3-TINY

Yang Guoliang Li Shicong Zou Junfeng Gong Jiaren

(School of Electrical Engineering and Automation, Jiangxi University of Science and Technology, Ganzhou 341000, Jiangxi, China)

Abstract Aimed at the helmet wearing problem of two-wheeled vehicle drivers and passengers, a lightweighted helmet detection model based on YOLOv3-tiny is proposed. The original model backbone network was light-weighted to reduce the amount of parameters of the detection model, and a U-shaped feature secondary fusion module was added to the network. The DIoU loss function about the distance of bounding boxes was introduced to improve the feature extraction ability of the detection model and recognition accuracy. Experiments on the test set show that the improved model exhibits higher recall rate and mAP and F1-score than the original YOLOv3-tiny model, and it has better detection performance than the deep network YOLOv3 while maintaining a small amount of parameters.

Keywords Helmet detection Lightweighted network Feature fusion Bounding box loss function

0 引言

二轮车具有价格便宜、出行便捷等一系列的优点,因此一直以来受我国广大消费者的青睐,是我国人民群众的主要出行工具。由于二轮车的安全保障相对较差,在二轮车交通事故中,发生人员伤亡的可能性相对更高。《2015 年全球道路安全现状报告》^[1]表明,在我国道路交通事故中,二轮车相比其他出行方式更容易产生伤亡。同时,因头部损伤致死人数占我国道路交通事故死亡人数的 66.4%,在该类损伤致死人数中,二轮车驾乘人员占比较高,而二轮车头盔可以保护驾乘人员头部,降低驾乘人员伤亡风险^[2]。因此,利用先

进的机器学习方法,如卷积神经网络^[3],监督道路上的二轮车驾乘人员,督促驾乘人员佩戴安全头盔,对减少事故伤亡率、降低事故伤亡风险和道路交通安全有着重大的意义。

目前国内对于二轮车头盔的佩戴检测研究尚少,国外对于二轮车头盔佩戴检测有较多的研究。在传统图像处理方法上,文献[4-5]提取候选区域的方向梯度直方图(HOG)特征,用提取到的 HOG 特征训练分类器如支持向量机(SVM)分类驾乘人员是否佩戴头盔。在基于深度学习的图像处理方法上,文献[6-8]利用目标检测算法如 YOLOv2,在第一阶段先检测出所有二轮车驾乘人员,后在第二阶段检测驾乘人员是否佩戴头盔。文献[9]使用 Caffe 框架下预训练好的

目标检测模型进行二轮车检测,检测出的目标预测框接着送入 InceptionV3 分类网络,进行头盔与其他物体的二分类。文献[10]在固定分类网络为 GoogLeNet 的前提下,对比了三类二轮车检测方法,结果表明采用基于卷积神经网络的检测方法(YOLO),相比其他两类传统图像处理方法,可显著提升二轮车头盔检测方法的精确程度。

本文对文献[7]中二轮车头盔检测的第二阶段进行改进,对所使用的复杂检测模型进行轻量化和特征增强,旨在使改进后的检测模型具有可迁移到移动设备端的能力,同时保持优良的检测能力。本文选择 YOLO 系列^[11-13]目标检测算法的轻量化版本 YOLOv3-tiny 检测模型进行改进。主要工作有:(1)将原始检测模型所用主干网络替换为轻量化网络。(2)对更换主干网络后的检测模型添加 U 型特征二次融合模块进行特征增强。(3)在检测算法损失函数中添加 DIoU 损失。

1 算法模型

头盔检测是一项定位并分类图片中二轮车驾乘人员是否佩戴头盔的机器视觉任务,该任务的主流解决方案是利用成熟的目标检测算法如 YOLOv3、SSD300 等,对待检测图片进行多阶段检测得到定位与分类结果。前述用于头盔检测的目标检测算法如 YOLOv3 可以得到较好的检测结果,但在易用性和可迁移能力两个方面上,由于所用的目标检测算法依赖的检测模型参数量过于庞大,如 YOLOv3 的参数量为 6.15×10^7 ,使检测模型迁移到如记录仪等移动设备上有着一一定的困难。

为轻量化上述目标检测模型,具体做法有减小网络深度、收缩网络宽度和卷积替换等,但随之出现的问题是较少参数的神经网络性能表现不如深度神经网络。为使较少参数的头盔检测模型达到较高的精确程度,本文认为应满足以下基础条件:(1)简洁高效的轻量化主干网络。(2)合理有效的特征复用方法。(3)算法层面的改进。满足上述基础条件的较少参数量头盔检测模型,能在获取到与深度神经网络同样丰富的特征信息的同时,增强深层特征的语义性和检测算法的精确性,以提升头盔检测结果的正确性。

本文首先考虑使用参数量较小的 YOLOv3-tiny 检测模型完成头盔检测任务,但该检测模型的检测结果并不理想,故将 YOLOv3-tiny 检测模型作为基准模型进行改进。为了进一步减少基准模型的参数量,使检测模型在可迁移到的移动设备上有更大的选择空间,

本文采用深度可分离卷积替换标准卷积搭建检测模型的主干卷积神经网络,通过合理设计网络结构,避免更换小参数量主干网络后的检测模型出现性能退化。

卷积神经网络的参数量侧面反映了网络的学习能力,故减小头盔检测模型参数量必然会带来卷积神经网络学习能力的降低。为在轻量化头盔检测模型的同时保持优良的检测能力,本文设计一种 U 型二次特征融合模块和引入一类基于边框距离的 DIoU 损失函数,其中 U 型二次特征融合模块用于增强轻量化头盔检测模型的特征提取能力,DIoU 损失函数用于提高头盔检测算法所生成预测框的精确程度。综上,本文所改进的基于 YOLOv3-tiny 的头盔检测模型的整体网络结构如图 1 所示。

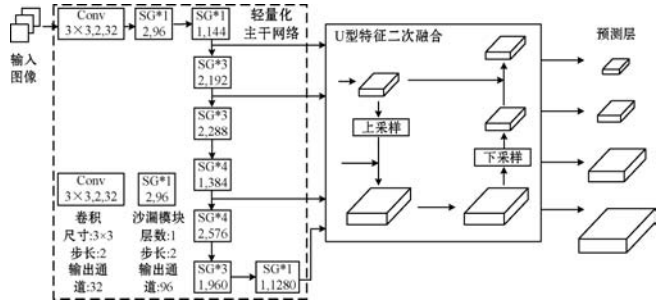


图 1 本文整体网络结构

1.1 轻量化主干网络

本文参考文献[14-16]的轻量化方法,采用基于深度可分离卷积的卷积神经网络作为检测模型主干网络,该网络基本组件由图 2 中包含深度可分离卷积的多个网络层构成,将该基本组件记作沙漏层。沙漏层中首尾的两组 3×3 逐深度卷积可以缓解采用典型瓶颈层中先减少通道数量再提取特征的方法所带来的特征信息丢失,获取更丰富的特征图空间信息。沙漏层中间的两组减少和扩充特征图通道数量的 1×1 卷积构成了沙漏层中的瓶颈模块,该两组卷积用于实现内部降维和固定最终输出特征图通道数,减少沙漏层的参数量和计算资源消耗。综合上述四组卷积的特点,在经过沙漏层前向通路的处理后,所得到的前向输出特征图仍保持着丰富的高维空间特征信息。跳层连接将沙漏层输入和前向输出特征图以逐元素相加的方式进行融合生成最终输出特征图,两者的高维特征信息互相补充,增强了沙漏层的特征描述能力。

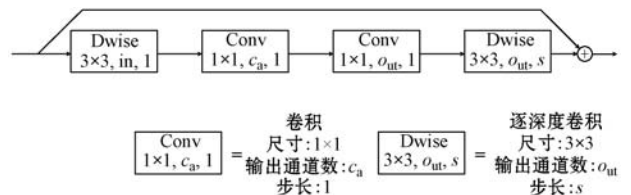


图 2 沙漏层

在图 1 中,SG 模块表示内含多个连续堆叠的沙漏层,轻量化主干网络主要由 SG 模块连续堆叠构成。每个 SG 模块内含的沙漏层的输出通道数相等,即图 2 中的 o_m 值均相等。为减少参数数量和计算资源消耗,调整主干网络中所有沙漏层的第一个降维卷积输出通道数即图 2 中的 c_a 值,设为该沙漏层输出通道数的六分之一,即 $c_a = o_m/6$ 。在步长为 2 的 SG 模块中,仅有首层沙漏层的步长值即图 2 中的 s ,设置为 2,用于下采样特征图减小尺寸,而首层后续堆叠的沙漏层步长 s 均设置为 1。在所有沙漏层中,仅有在输入与输出特征图尺寸相同与通道数相等时,跳层连接才存在。

1.2 U 型二次特征融合模块

本文在轻量化主干网络后增加一个 U 型二次特征融合模块,以增强检测模型在检测过程中所提取到的特征,图 3 为 U 型二次特征融合模块的整体结构。考虑到角度距离等因素,导致图片中检测对象头盔的尺度变化范围较大,特征融合模块采用轻量化主干网络中最后四次下采样前的四种不同尺寸的特征图作为输入,输入的四种特征图在尺寸上具有从小到大依次倍增的关系,例如 13×13 、 26×26 、 52×52 、 104×104 。

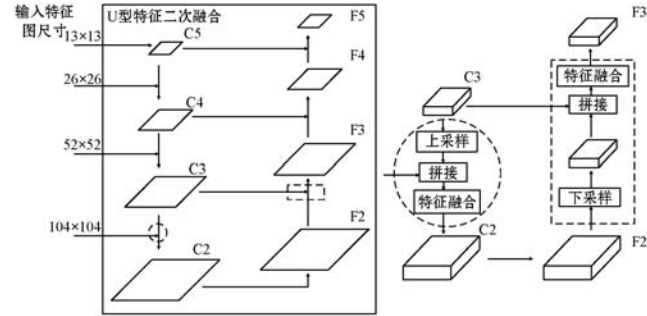


图 3 U 型二次特征融合模块

如图 3 所示,从模块内中间特征图的尺寸变化角度来看,特征融合模块可以分为自顶向下和自底向上两条路径,分别记为 P1 和 P2。输入的四个尺寸的特征图首先要在自顶向下路径 P1 内前向传播,尺寸最小的特征图位于自顶向下路径 P1 上的顶部,记为 C5。除开位于底部的最大尺寸特征图外,其余三种尺寸特征图各自经过上采样层扩充一倍,扩充后与输入中尺寸相符的原特征图进行特征融合,分别得到中间特征图 C4、C3、C2。考虑到逐元素相加的特征融合方式会损失一部分特征信息,故将两个相同尺寸的特征图的所有通道进行拼接,拼接所得的特征图经过两组连续的 3×3 卷积,使两种特征图上的特征信息进行通道级的融合。在自底向上路径 P2 中,自顶向下路径 P1 所生成的底部中间特征图 C2 即是自底向上路径 P2 的起点特征图 F2。特征图 F2 在自底向上路径 P2 中共经

过三次下采样,每次下采样后通过旁侧连接分别与自顶向下路径 P1 中尺寸匹配的中间特征图 C5、C4、C3 拼接所有通道,拼接后进行二次特征融合,得到 U 型二次特征融合模块的输出特征图 F5、F4、F3、F2。

在 U 型二次特征融合模块的两条路径 P1 和 P2 中,路径 P1 自顶向下地将主干网络所提取到的深层特征与浅层特征互相融合,将深层特征连续补充进较高分辨率的特征图中,路径 P2 将融合了丰富特征信息的较高分辨率特征图连续下采样并与 P1 进行特征融合,在较低分辨率特征图上得到了更为精细的特征信息。

1.3 DIoU 损失函数

在基于锚框的目标检测算法中,交并比也称 IoU,是评价预测框和对应真实框重叠程度的指标,如图 4 (a) 所示,其定义为:

$$\frac{|P \cap G_T|}{|P \cup G_T|} \quad (1)$$

式中: P 和 G_T 为预测框和真实框的面积。交并比只关注两个边框的重叠区域对两个边框并集面积的占比,无法精确反映两个边框的重合程度,故对检测算法中所使用的 IoU 均替换为如式(2)所示的新定义。

$$I_{oU} = \frac{|P \cap G_T|}{|P \cup G_T|} - \frac{|C - (P \cup G_T)|}{|C|} \quad (2)$$

式中: C 是预测框和真实框的最小包围框的面积,如图 4(b) 所示。式(2)中第二项比值表示了两个边框的非重叠区域(不包括边框自身区域)占最小包围框的占比,对于重合程度越好的两个边框,该比值越小,即惩罚项越小,IoU 值则越大。

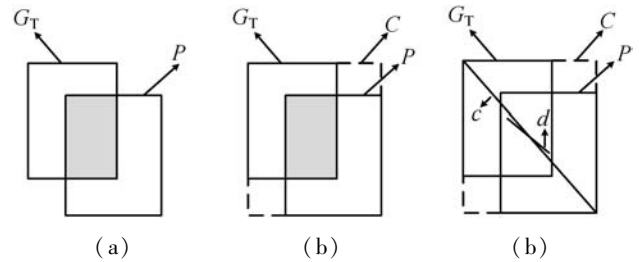


图 4 本文所使用 DIoU 定义

预测框和真实框的几何距离也是衡量重合程度的重要指标之一,故在设计 IoU 损失函数时,添加边框距离项 $DIoU^{[17]}$ 作为惩罚项,最终使用的 $DIoU$ 损失函数如下:

$$R_{DIoU} = \frac{\rho^2(d^p, d^{G_T})}{c^2} \quad (3)$$

$$L_{DIoU} = 1 - I_{oU} + R_{DIoU} \quad (4)$$

式中: ρ 表示两个边框中心点的欧氏距离即图 4(c) 中的 d ; c 表示包围两个边框的最小包围框的对角线长度,如图 4(c) 所示。在式(4)的 $DIoU$ 损失函数中,IoU

项关注预测框与真实框的非重叠区域和重叠区域,边框距离项 DIoU 则关注预测框和真实框的距离,将 DIoU 损失函数放入检测算法的整体损失函数中共同优化,可以得到与真实框精确重合的预测框。

2 实验与结果分析

实验环境:64 位 Ubuntu 16.04 操作系统,CPU 为 Intel i7-4790 @ 3.60 GHz 八核处理器,GPU 为 NVIDIA GeForce GTX 1070,并选择 PyTorch 深度学习框架来测试和评价模型。

2.1 实验数据

本文使用 BikeHelmets 数据集来训练和测试模型,该数据集共包含 761 幅图片,每幅图片中包含单个或多个二轮车驾驶员及乘坐人员,这两类人员被划分为佩戴头盔和未佩戴头盔两个类别,图片中的类别与定位信息被标注并保存在与图片对应的标签文件中,按 8:2 的比例将数据集划分为训练集和测试集。由于原数据集存在较多的错误标注和遗漏标注,故对原数据集重新进行标注,以丰富数据集内训练样本,平衡样本类别分布。为缓解数据集内较少的样本量带来的神经网络过拟合,在数据集封装程序中,使用水平翻转、比例缩放等数据扩充方法将待训练的图片进行扩充,再前向传播给网络模型。

2.2 评价指标

本文选取查全率 (R)、类别平均精度均值 (mAP) 和 F1 指标 (F1-score) 作为评价模型性能的主要指标,查全率和辅助指标精确率 (P) 定义如下:

$$R = \frac{T_p}{T_p + F_N} = \frac{T_p}{\text{len}(G_{Ts})} \quad (5)$$

$$P = \frac{T_p}{T_p + N_p} = \frac{T_p}{\text{len}(D_{ets})} \quad (6)$$

式中: T_p 表示检测模型生成的检测框里与真实框的 IoU 值大于指定阈值的边框数量,同一个真实框只计算一次,即正确检测出的目标数量; N_p 表示检测框里与真实框的 IoU 值小于指定阈值的边框数量,即错误检测出的目标数量; F_N 表示没有检测出来的真实框的数量; G_{Ts} 和 D_{ets} 分别表示真实框的集合和检测框的集合。

仅用精确率与查全率中的单个指标无法评价检测模型之间的优劣,故另选取 F1 指标 (F1-score) 来评价检测模型,该指标是查全率和精确率的调和平均,可作为评价检测模型优劣的指标,该指标定义如下:

$$F_{1\text{-score}} = 2 \times \frac{P \times R}{P + R} \quad (7)$$

在评价目标检测模型时,常用类别平均精度均值 (mAP) 指标衡量检测模型的预测框的分类和定位是否准确,AP 表示某一类别的精确率与查全率所画出的 P-R 曲线下的面积,定义如下:

$$A_p = \int_0^1 P dR \quad (8)$$

通常来说,某一类别的 P-R 曲线越往右上方凸,即 AP 值越大,则表明检测模型对于该类别的检测能力越优秀。AP 是检测模型对某一类别的准确率和查全率的综合评价,mAP 表示所有类别的 AP 指标的平均值,可作为评价检测模型分类与定位性能的综合指标。

2.3 实验结果分析

在原数据集上划分出的测试集共包含 271 个佩戴头盔类别与未佩戴头盔类别的真实目标,为验证改进模型的性能,将 YOLOv3 检测模型、原始 YOLOv3-tiny 检测模型和改进后的检测模型在测试集上的检测性能进行对比,评价指标中的查全率 (R) 与 F1 指标 (F1-score) 均取测试集所含头盔佩戴与否即两种类别的对应指标的平均值作为最终对比结果,对比结果如表 1 和表 2 所示。

表 1 对比实验结果

模型	R/%	mAP/%	参数量/($\times 10^6$)
YOLOv3	88.0	85.6	61.50
YOLOv3-tiny	83.4	80.9	8.67
本文改进模型	87.8	88.1	9.23

表 2 参数量对比实验结果

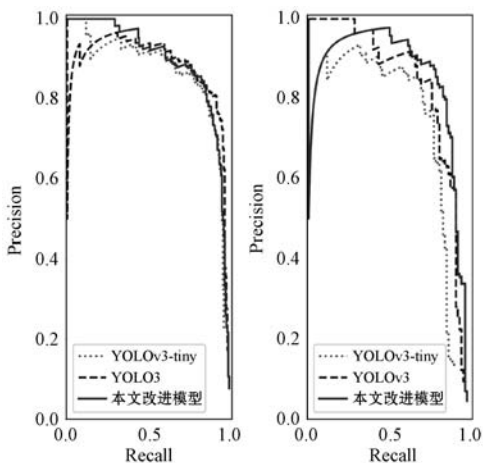
模型	参数量 /($\times 10^6$)	点乘操作数 /($\times 10^9$)	F1-score /%
YOLOv3	61.50	37.76	79.9
YOLOv3-tiny	8.67	2.75	76.0
本文改进模型	9.23	13.20	82.2

从表 1 可以看出,与原 YOLOv3-tiny 检测模型相比,由于特征融合模块的引入使得参数量略微增加,但改进后的检测模型相比原检测模型具有更高的查全率和 mAP,分别在查全率和 mAP 上提升了 4.4 个百分点和 7.2 个百分点。与深度卷积神经网络模型 YOLOv3 相比,改进模型所需的参数量减小到其参数量的五分之一,但在 mAP 评价指标上提升了 2.5 百分点,在查全率 R 上相比于 YOLOv3 具有相接近的表现。

表 2 是三类实验模型在模型参数量、检测一幅 416×416 图片所需操作数和 F1 指标三个方面的对比

结果。可以看出, YOLOv3-tiny 模型的参数量最小, 检测同一幅图片所需的操作数也最小, 但同时其检测模型的 F1 指标也最低。YOLOv3 模型的 F1 指标相比 YOLOv3-tiny 模型提高了 3.9 百分点, 但在参数量和点乘操作数需求上相比 YOLOv3-tiny 模型呈指数级增长。在改进原 YOLOv3-tiny 模型时, 将增强原检测模型检测能力所必需的点乘操作数以尽量小的参数量实现, 可以看出, 改进模型参数量略高于原 YOLOv3-tiny 检测模型, 点乘操作数为 YOLOv3 的三分之一, 但 F1 指标要显著高于 YOLOv3-tiny 和 YOLOv3。

图 5 为三个对比实验检测模型在测试集内两个类别上的 P-R 曲线图。可以看出, 对于佩戴头盔类别, 改进模型的 P-R 曲线仅在部分小区间低于 YOLOv3 模型, 这表明改进模型关于检测佩戴头盔类别的性能表现要优于 YOLOv3-tiny 原始模型与 YOLOv3 模型。关于未佩戴头盔类别, 改进模型的 P-R 曲线在查全率较低的区间内低于 YOLOv3 模型, 同时明显高于 YOLOv3-tiny 原始模型曲线, 这表明在检测未佩戴头盔类别方面, 改进模型的检测能力要显著强于原始模型 YOLOv3-tiny, 而与 YOLOv3 模型具有相接近的检测性能。



(a) 佩戴头盔 (b) 未佩戴头盔

图 5 P-R 曲线

为测试各个模块的改进效果, 设置四组消融实验, 分别为: (1) 基准实验, 测试模型为基准模型 YOLOv3-tiny, 记为 baseline; (2) 轻量化网络实验, 对基准模型的主干网络进行轻量化处理后进行测试, 记为 M1; (3) U 型特征二次融合实验, 在 M1 实验的测试模型上加入 U 型特征二次融合模块进行测试, 记为 U1; (4) DIoU 实验, 将 U1 实验模型中的检测算法中加入 DIoU 损失函数进行测试, 记为 D1。

从表 3 可以看出, 用于改进检测模型的三个模块不同程度地提升了查全率和 mAP 指标。其中轻量化主干网络主要作用为将原检测模型参数量减少至接近一半, 同时保持与原检测模型相接近的查全率和略高

于原检测模型的 mAP 指标。在轻量化网络减少了原检测模型近一半参数量的基础上, U 型特征二次融合模块虽然带来了较多的参数量, 但增强了检测模型的特征提取能力, 有效提升了评价指标。在加入上述两个模块的基础上, 在检测算法中添加 DIoU 损失, 在不增加参数量的情况下, 使检测算法更加精确, 显著提升了检测模型的检测能力。

表 3 消融实验结果

实验模型	轻量化网络	特征融合	DIoU	R/%	mAP/%	参数量 / ($\times 10^6$)
baseline				83.4	80.9	8.67
M1	√			84.0	83.8	4.99
U1	√	√		85.0	86.1	9.23
D1	√	√	√	87.8	88.1	9.23

2.4 测试图片结果对比

从测试集中随机挑选三幅图片, 用于对比检测模型在测试图片上的检测结果。从图 6 检测结果可以看出, 原 YOLOv3-tiny 模型在测试图片上的漏检和错检现象比较严重, 原因在于其主干网络简浅, 无法提取到丰富的深层特征。YOLOv3 检测模型在测试图片上没有出现错检现象, 但对于较小的目标仍存在漏检现象, 原因在于其检测尺度不够多, 小目标在大检测尺度上不容易被检测处理。改进模型克服了原 YOLOv3-tiny 检测模型的缺点, 相比上述两个检测模型, 很好地抑制了错检漏检现象, 正确检测到较多数量的真实框, 具有良好的检测效果。



(a) YOLOv3-tiny (b) YOLOv3 (c) 改进模型

图 6 测试集部分图片检测结果

2.5 现实路口检测效果

为验证改进模型在实际应用场景下的效果, 本文根据赣州市某交通路口摄取的通行视频, 以抽帧的方法制作了一个现实场景数据集。该现实场景数据集共包含 238 个真实目标, 每个真实目标均标注其所属是否佩戴头盔的类别。将本文改进模型在现实场景数据

集上进行二轮车头盔检测,检测结果如表 4 所示。

表 4 现实场景数据集检测结果 (%)

检测类别	R	mAP	F1-score
佩戴头盔	83.8	82.6	76.3
未佩戴头盔	75.4	70.1	68.4

根据表 4 可知,改进模型在实际使用中检测未佩戴头盔驾乘人员时,检测效果差于检测佩戴头盔驾乘人员。造成该现象的原因有以下两点:(1) 在检测离摄像点较远处目标时,未佩戴头盔人员的头部所占像素数较少,容易造成遗漏;(2) 受新冠疫情影响,现实交通中二轮车驾乘人员佩戴口罩者居多,而本文所用训练数据集中不包含佩戴口罩的二轮车驾乘人员实例,导致改进模型在识别佩戴口罩目标时识别难度加大。针对图 7 中改进模型在实际使用时突显出的问题,可通过设置取景框和在训练集中增加样例等方法解决,这时也是本文未来研究的方向和目标。



图 7 现实场景检测效果

3 结 语

本文提出一种用于检测二轮车驾乘人员头盔佩戴情况的轻量化检测模型,以 YOLOv3-tiny 作为基准模型,对该模型主干网络进行轻量化处理,并添加特征融合模块和 DIoU 损失函数增强模型检测能力。测试结果表明,相比于基准模型和深度卷积神经网络检测模型,本文提出的轻量化检测模型以较少的参数量达到了较高的查全率和 mAP 指标及 F1 指标,同时在实际使用中保持了良好的检测能力。

参 考 文 献

- [1] World Health Organization. Global status report on road safety 2015[R]. World Health Organization, 2015.
- [2] 李金刚,戴帅. 国外摩托车和自行车安全头盔佩戴经验以及对我国的启示[J]. 汽车与安全, 2017(2): 21-25.
- [3] 罗会兰,易慧. 基于三流卷积神经网络模型的图像分类方法[J]. 江西理工大学学报, 2019, 40(5): 86-92.
- [4] Mukhtar A, Tang T. Vision based motorcycle detection using HOG features[C]//2015 IEEE International Conference on Signal and Image Processing Applications, 2015: 452-456.
- [5] Silva R, Aires K, Veras R. Detection of helmets on motorcyclists[J]. Multimedia Tools and Applications, 2017, 77: 5659-5683.
- [6] Devadiga K, Gujarathi Y, Khanapurkar P, et al. Real time automatic helmet detection of bike riders[J]. International Journal for Innovative Research in Science and Technology, 2018, 4(11): 146-148.
- [7] Mistry J, Misraa A, Agarwal M, et al. An automatic detection of helmeted and non-helmeted motorcyclist with license plate extraction using convolutional neural network [C]//2017 7th International Conference on Image Processing Theory, Tools and Applications, 2017: 1-6.
- [8] Raj K, Chairat A, Timtong V, et al. Helmet violation processing using deep learning [C]//2018 International Workshop on Advanced Image Technology, 2018: 1-4.
- [9] Rohith C, Nair S, Nair P, et al. An efficient helmet detection for MVD using deep learning [C]//2019 3rd International Conference on Trends in Electronics and Informatics, 2019: 282-286.
- [10] Chairat A, Dailey M, Limsoonthrakul S, et al. Low cost, high performance automatic motorcycle helmet violation detection [C]//2020 IEEE Winter Conference on Applications of Computer Vision, 2020: 3549-3557.
- [11] 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, 2016: 779-788.
- [12] Redmon J, Farhadi A. YOLO9000: Better, faster, stronger [C]//2017 IEEE Conference on Computer Vision and Pattern Recognition, 2017: 6517-6525.
- [13] Redmon J, Farhadi A. Yolov3: An incremental improvement [EB]. arXiv:1804.02767, 2018.
- [14] Howard A, Zhu M, Chen B, et al. Mobilenets: Efficient convolutional neural networks for mobile vision applications [EB]. arXiv:1704.04861, 2017.
- [15] Sandler M, Howard A, Zhu M, et al. Mobilenetv2: Inverted residuals and linear bottlenecks [C]//2018 IEEE Conference on Computer Vision and Pattern Recognition, 2018: 4510-4520.
- [16] Zhou D, Hou Q, Chen Y, et al. Rethinking bottleneck structure for efficient mobile network design [EB]. arXiv: 2007.02269, 2020.
- [17] Zheng Z, Wang P, Liu W, et al. Distance-IoU loss: Faster and better learning for bounding box regression [C]//AAAI Conference on Artificial Intelligence, 2020: 12993-13000.