

一种基于 YOLOv5s 的红外图像目标检测改进算法

李晓佩, 张寅宝, 李严培, 姚芸星

(郑州大学南校地球科学与技术学院, 河南 郑州 450000)

摘要: 受热红外成像方式限制, 交通场景下红外图像存在对比度低、目标尺度和姿态的多样性以及目标之间的相互遮挡问题, 从而造成检测精度下降, 部分目标出现漏检、误检的情况。本文在 YOLOv5s 的基础上提出一种改进算法: 在数据处理方面, 使用 AHE 算法对训练集图像进行部分数据增强; 在模型改进方面, 通过引入跨域迁移学习策略、插入通道注意力机制 SENet、改进损失函数 GIoU 为 α -CIoU 对 YOLOv5s 进行改进。并通过消融实验的方式, 在自制数据集上对夜间道路环境下的电动自行车驾驶行为进行检测。实验结果表明, 改进后的算法对单人驾驶电动自行车行为检测的平均精度达到了 95.9%, 比 YOLOv5s 的检测精度提高了 3.1%; 对载人驾驶电动自行车行为检测的平均精度达到了 88.4%, 比 YOLOv5s 的检测精度提高了 9.5%; 总类别检测的平均精度达到了 92.2%, 比 YOLOv5s 的检测精度提高了 6.4%, 有效降低了红外目标漏检、误检的概率。

关键词: YOLOv5s; 红外目标检测; 跨域迁移学习; SENet; α -CIoU Loss

中图分类号: TP391; TN219 **文献标识码:** A **DOI:** 10.3969/j.issn.1001-5078.2023.07.010

An improved method of infrared image target detection based on YOLOv5s

LI Xiao-pei, ZHANG Yin-bao, LI Yan-pei, YAO Yun-xing

(The School of Geo-Science & Technology, Zhengzhou University, Zhengzhou 450000, China)

Abstract: Due to the limitations of heated infrared imaging mode, infrared images in traffic scenes suffer from low contrast, diversity of target dimensions and postures, and mutual occlusion between targets, which leads to the decrease of detection accuracy and the false detection of some targets. In this paper, an improved algorithm based on YOLOv5s is proposed in terms of data processing, the AHE algorithm is used for partial data enhancement of the training set images; in terms of model improvement, YOLOv5s is improved by introducing a cross-domain transfer learning strategy, inserting channel attention mechanism SENet and improving the loss function GIoU to α -CIoU. By means of ablation experiments, the driving behavior of electric bicycle in night road environment is detected on self-made data set. The experimental results show that the improved algorithm achieves an average accuracy of 95.9% for single-person driving electric bicycle behavior detection, which is 3.1% higher than that of YOLOv5s; an average accuracy of 88.4% for manned electric bicycles behavior detection, which is 9.5% higher than the detection accuracy of YOLOv5s; and an average accuracy of 92.2% for total category detection, which is 6.4% higher than the detection accuracy of YOLOv5s, effectively reducing the probability of missing and false detection of infrared targets.

Keywords: YOLOv5s; infrared target detection; cross-domain transfer learning; SENet; α -CIoU Loss

作者简介: 李晓佩(1997-), 女, 硕士研究生, 研究方向为深度学习, 图像处理。

通讯作者: 张寅宝(1980-), 男, 副教授, 硕士生导师, 主要研究方向为目标检测, 图像处理。E-mail: zhangyinbao@zzu.edu.cn

收稿日期: 2022-09-17

1 引言

根据中国统计年鉴,2020 年全国非机动车交通事故共 29969 起,造成 4373 人死亡,33636 人受伤,直接财产损失达到了 6271.4 万元。电动自行车作为非机动车的典型代表,数量庞大、不易制定规范、事故发生率高,朱新宇等^[1]根据相关数据分析,2015~2020 年驾驶电动自行车肇事死亡人数增长了 86.6%,其中乘坐者死亡人数达到了电动车总体死亡人数的 10% 以上。夜间光照条件不良的情况下,电动自行车驾驶人员对周围环境的感知能力降低,更易引发交通事故。因此,采用红外图像对夜间电动自行车驾驶行为检测具有现实意义。

传统非机动车检测主要是人工提取图像中的颜色、形状等特征,然后通过支持向量机、Adaboost 等分类器识别^[2]。近几年随着深度学习的发展,现代目标检测算法成为非机动车检测识别的主流方法,有效克服了传统非机动车检测方法中人工构建特征不全、非机动车外观多变、鲁棒性低、窗口冗余等问题,一般分为 one-stage 目标检测算法和 two-stage 目标检测算法^[3]两类。YOLO 作为典型的 one-stage 目标检测算法近些年来被广泛应用于红外目标检测^[4-7]中,本文将 YOLOv5s 应用于夜间交通场景下电动自行车驾驶行为检测,分别在数据预处理阶段和 YOLOv5s 模型本身两个方面进行改进,最后通过消融实验的方式验证了改进算法的有效性。

2 改进算法

使用 YOLOv5s 实现夜间道路环境下电动自行车驾驶行为检测的方法主要包括三个步骤:数据预处理、网络训练和检测结果输出。在数据预处理阶段包括数据采集、筛选图像、数据划分和数据集制作四个步骤,本文在数据预处理中增加部分数据增强步骤,目的在于改善由红外成像方式引起的图像质量不高、目标边界模糊造成的检测精度下降问题,针对图像中含有目标类别选择增强算法,起到丰富目标特征、提升检测精度的效果。在网络训练阶段选用 YOLOv5s 模型,该算法主要包括四个部分:Input、Backbone、Neck、Prediction。在 Input 中引入跨域迁移学习策略,用于解决使用小样本数据集对网络进行训练造成的模型不稳定问题;在网络的主干部分

Backbone 中添加基于通道注意力机制的 SENet 模块,用于解决由于目标尺度和姿态多样性造成的部分目标漏检、误检问题。在 Prediction 中使用 α -CIoU_Loss 替换原来的 GIoU_Loss,提升模型的检测速度、检测精度。本文流程如图 1 所示。

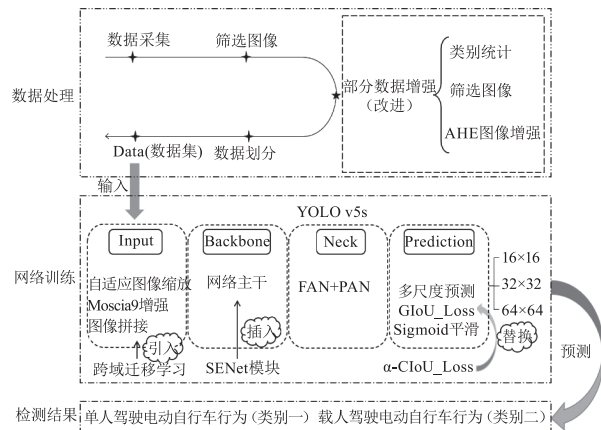


图 1 本文流程图

Fig. 1 Flow chart of proposed

最后输出两个类别的检测结果:单人驾驶电动自行车行为,以下统称为类别一;载人驾驶电动自行车行为,以下统称为类别二。

2.1 数据处理改进

在夜间环境下,热红外成像方式具有不受光照条件制约的优势,但受设备性能的影响,红外图像相较于可见光图像存在分辨率低、目标边缘模糊的缺点,从而造成使用 YOLOv5s 在红外图像上进行检测时精度不高。为解决这一问题,本文选取了两种算法:自适应直方图均衡(Adaptive Histogram Equalization, AHE)、限制对比度自适应直方图均衡(Contrast Limited Adaptive Histogram Equalization, CLAHE)方法对图像进行增强。两种方法的增强效果如图 2 所示。



图 2 增强效果对比图

Fig. 2 Contrast image of enhanced effect

在交通场景下的红外图像中,成像范围内环境温度变化较小,图像整体表现为对比度低。驾驶电动自行车行为目标在图像中表现为亮部少、暗部多

的一种轮廓形态,电动自行车的车体部分温度接近环境温度,在图像中表现为亮度低,其中车轮部分在运动过程中温度较高,在图像中表现为一个亮度较高的环形轮廓,驾驶人和乘坐人与环境温度相差较大,在图像中表现为亮度高,两类目标中类别二的亮部区域多于类别一。普通的直方图均衡化算法对图像中全部像素使用相同的直方图进行变换,图像中较亮和较暗区域对比度得不到有效提升。AHE 算法对每个像素的周边矩形范围内的像素做直方图均衡化处理,增强后的图像表现为亮部更强暗部更暗,通过提升局部特征的对比度来提升整幅图像的对比度,同时放大图像中的噪声;CLAHE 算法能够有效地抑制 AHE 算法产生的噪声,但同时也降低了图像的对比度,不利于网络训练。

对比使用 AHE 算法和 CLAHE 算法增强后数据集与未使用算法增强的数据集的训练结果发现,使用 AHE 算法增强后的数据集中类别二的训练精度有明显上升,类别一的训练精度略有下降;使用 CLAHE 算法增强后的数据集中类别一、类别二的训练精度均略有下降。基于此,提出一种数据集数据处理改进方法——部分数据增强方法,首先对数据集中每幅图像所含的目标类别进行统计,然后根据图像中所含目标类别选择增强算法,对含有类别一的图像做自适应直方图均衡化(AHE)处理,对含有类别二的图像不做处理,同时含有类别一、类别二的图像不做处理。

2.2 YOLOv5s 模型改进

YOLOv5s 模型 (YOLOv5s-5.0 版本) 主要包括 4 个部分: ①Input 部分, 主要包括图像的自适应缩放, Moscia9 数据增强。②网络的主干部分 Backbone, 通过特征提取网络对特征信息进行提取。③Neck 部分, 主要包括 FPN + PAN 结构, 用以连接网络的 Backbone 和 Prediction 部分。④Prediction 部分, 包括多尺度预测, 在三个尺度 (16×16 , 32×32 , 64×64) 上对输出目标框、标签类别进行预测, 边界框损失函数采用 GIoU_Loss 提升预测框的回归速度和精度, 使用 Sigmoid 激活函数进行平滑以便求导。

本文在 Input 部分通过预加载权重的方式引入跨域迁移学习策略,降低目标检测任务对硬件设备的依赖性,同时迁移学习可见光图像训练所得网络

的边缘概率和网络参数,提高使用小样本红外数据集对网络训练的精度;在 Backbone 部分插入基于通道注意力机制的 SENet 模块,改善目标间互相遮挡造成的漏检、误检的问题;在 Prediction 部分使用 α -CIoU_Loss 替换 GIoU_Loss,提升网络梯度更新速度、增加高 IoU 锚框的输出数量。改进后的网络结构如图 3 所示。

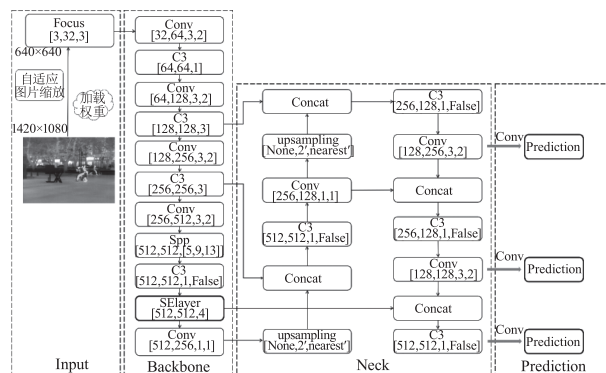


图3 改进后 YOLOv5s 网络结构

Fig. 3 Improved YOLOv5s network structure

2.2.1 跨域迁移学习

在目标检测任务中,往往需要大量的图像数据对模型进行训练以提高精度。受硬件设备的约束以及人工标注成本制约,用于模型训练的数据集较小时,会导致训练得到的模型检测精度低、泛化能力不强。针对这一问题,冯毅雄等^[8]提出一种集成迁移学习的轴件表面缺陷实时检测方法,解决了小样本和实时检测效率低的问题;蒋文萍等^[9]采取多重迁移学习的方法通过对 400 张图片的火灾数据集训练得到火灾探测模型,得到了较好的目标精度;王卓^[10]等使用跨域迁移与域内迁移学习相结合的学习策略解决了苹果数据集样本量小的问题。

基于前人的研究,为了降低目标检测任务对硬件设备的依赖和人工标注成本,针对红外目标的检测,本文在 YOLOv5s 模型的基础上提出一种基于小样本数据集的跨域迁移学习策略,即在 YOLOv5s 的 Input 部分预加载 COCO 数据集的训练权重进行训练。跨域迁移学习策略如图 4 所示。

跨域迁移学习就是从大数据集样本、多种类(80类)目标训练得到的用于可见光图像目标检测的网络权重,应用于小样本数据集、较少种类(2类)目标训练的红外目标检测任务中进行再训练。从目

标种类丰富、样本量大的数据集训练得到的用于可见光图像目标检测的模型中学习其特征空间、边缘概率和网络参数,进而提升 YOLOv5s 应用于小样本红外数据集训练的泛化能力和检测精度。

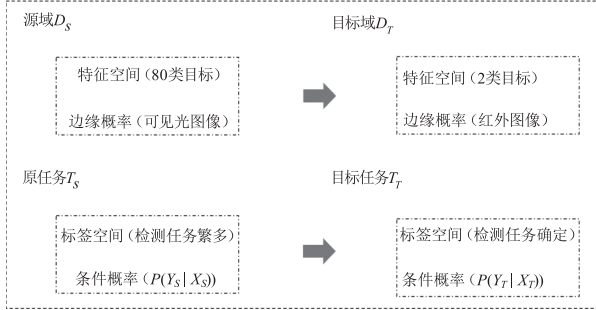


图 4 跨域迁移学习策略图

Fig. 4 Cross-domain transfer learning strategy diagram

2.2.2 SENet 模块

在交通场景下,道路环境复杂,目标物体之间、目标物体与其他物体之间不可避免的会出现互相遮挡的情况,造成可用检测特征减少,从而导致漏检、误检,检测精度下降问题。为解决检测特征减少造成的检测精度下降问题,本文在 Backbone 部分特征提取网络中的 C3 层和全连接层之间插入基于通道注意力机制的 SENet^[11] 模块,对提取到的共 512 个通道赋予不同权值,对通道中的无效特征进行抑制、有效特征进行增强,起到降低信息冗余的作用,从而降低目标的漏检、误检概率。SENet 模块如图 5 所示。

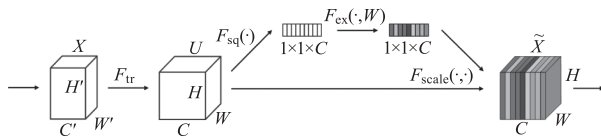


图 5 SENet 模块图

Fig. 5 SENet module diagram

SENet 模块采取重定标策略实现对通道的加权。首先,对于给定输入的 X 经过卷积操作后通道数由 C' 变换成 C ,然后通过 Squeeze 操作($F_{sq}(\cdot)$)将二维通道特征变换成实数实现通道加权,接着 Excitation 操作($F_{ex}(\cdot, W)$)通过 W 参数建立起通道间的相关性关系,最后 Scale 操作($F_{scale}(\cdot)$)将 Excitation 操作得到的相关关系与 C 个通道的特征图进行点乘,实现在通道维度上对原始通道中特征图层的重新加权。

2.2.3 α -CIoU_Loss

YOLOv5s 的 Prediction 部分对边界框的预测采

用 GIoU^[12] 作为损失函数,解决了交并比 (Intersection over Union, IoU) 在预测框与真实框在没有重叠情况下的梯度消失问题,但是并未解决预测框与真实框的宽和高相同时,GIoU 退化为 IoU 的问题。有学者在 GIoU 的基础上提出 CIoU^[13] 损失函数,进一步考虑 IoU 中的重叠面积、中心点距离和长宽比以及正则化项,解决了 GIoU 中出现的损失函数退化问题,有效促进了网络梯度的更新速率,但是并没有考虑到高 IoU 目标与低 IoU 目标之间的权重分配问题。本文在 CIoU 的基础上进一步考虑,使用 α -CIoU^[14] 代替原网络中的 GIoU 作为边界框预测的损失函数,通过控制权重系数 α ,使检测器根据所有预测对象的 IoU 自适应的降低和增加其相对损失。 α -CIoU 的定义为:

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

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

$$\beta = \frac{v}{(1 - IoU) + v} \quad (3)$$

式中, w^{gt} 为预测框宽; h^{gt} 为预测框高; w 为真实框宽; h 为真实框高; $\rho^2(b, b^{gt})$ 为预测框中心点 b 与真实框中心点 b^{gt} 的欧氏距离; c 为可以包含预测框和真实框的最小闭合区域的对角线长度; α 可取大于零的数。

3 实验

3.1 实验环境

实验参数设置: epochs 设置为 300, batch-size 设置为 8, 初始学习率、学习率动量、学习率周期权重衰减系数分别设置为: 0.01、0.937、0.2、0.0005, 图像输入大小为 640×640 。具体实验环境如表 1 所示。

表 1 实验环境

Tab. 1 Experimental environment

名称	环境参数
操作系统	Win10 64bit
CPU	Inter(R) UHD Graphics 630
GPU	NVIDIA GeForce RTX 2080 SUPRER
语言	Python3. 8. 3
神经网络框架	Pytorch, CUDA10. 2

3.2 精度指标

精确率 (Precision) 是指目标检测正确的概率,

计算公式如下:

$$Precision = \frac{TP}{TP + FP} \quad (4)$$

召回率(Recall)是指所有正样本中正确识别的概率,计算公式如下:

$$Recall = \frac{TP}{TP + FN} \quad (5)$$

平均准确率(Average Precision, AP)为 Precision-Recall 曲线与 x 轴所围成的面积,计算公式如

下所示:

$$AP = \int_0^1 p(r) dr \quad (6)$$

实验选取平均准确率精度(Average Precision, AP)作为评价指标,对模型改进效果进行评定。

3.3 数据集

3.3.1 数据采集

数据采集工具使用红外摄像机,其参数设置如表 2 所示。

表 2 数据采集参数设置

Tab. 2 Data collection parameter settings

色板模式	分辨率	量程/℃	校正参数					
			温度/℃	反射温度/℃	环境温度/℃	湿度	发射率	距离/m
黑热	1420 × 1080	-20 ~ 120	0	20	20	0.67	0.98	10

针对两种电动自行车驾驶行为,为了更好的适应实际环境,分别采集多段含有一定样本数量的视频序列。数据采集时间选择晚上的 8 点到 10 点,采集地点选择道路交叉口,现场采集环境如图 6 所示。



图 6 现场采集环境

Fig. 6 On-site collection environment

3.3.2 数据集制作

对拍摄到的 30 段红外视频序列按照 1 张/秒的速度进行截取,筛选出含有电动自行车驾驶行为的红外图像共 1681 张,按照 8 : 1 : 1 的比例将其划分为训练集、验证集和测试集,称为数据集一。遵循 YOLOv3^[15] 的方法, YOLOv5 沿用由 k-means 在 COCO 数据集上聚类得到的锚框来预测边界框。边界框的标注工具使用 LabelImg,生成 yolo 格式标注文本。得到两种类别的标签数量共 1917 个,类别一共 1514 个,类别二共 403 个。表 3 为数据集信息。

为验证数据预处理方法中数据集部分数据增强方法的有效性,制作数据集二,即对数据集一中训练集使用部分数据增强方法进行处理,不改变标签的

类别和数量。部分数据集二的图像如图 7 所示。其中, A 组为训练集中部分只包含类别一的图像, B 组为训练集中只包含类别二的图像, C 组为训练集中同时包含两个类别的图像。

表 3 数据集信息

Tab. 3 Data set information

类别	数量
红外图像/张	1681
类别/个	2
标签/个	1917
类别一(单人驾驶电动自行车行为)/个	1514
类别二(载人驾驶电动自行车行为)/个	403

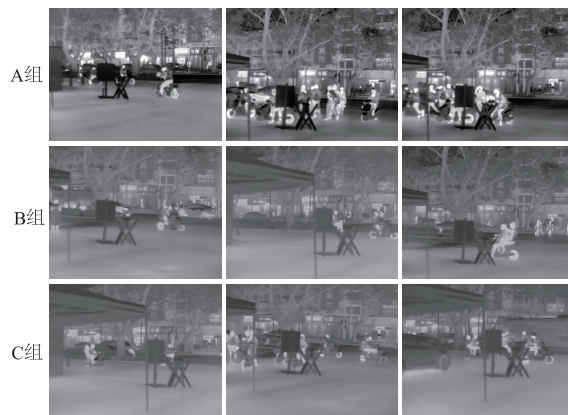


图 7 部分数据集二图像

Fig. 7 Partial data from dataset 2

3.4 消融实验

为了验证本文提出的数据处理改进方法中的部分数据增强方法和算法改进方法中的三种改进方法

分别对类别一和类别二的检测效果影响,采取消融实验的方式分别验证四种改进方式的效果,具体实验设置如表 4 所示。

以实验一为基础实验,实验二用来验证数据处理改进方法的有效性;实验三至实验五分别用来验

证 YOLOv5s 模型改进方法的有效性,实验三用来验证跨域迁移学习方法对小样本数据集训练的作用,实验四用来验证 SENet 模块对红外目标漏检、误检的改善效果,实验五用来验证 α -CloU_Loss 函数对算法检测速度提升的能力。

表 4 消融实验结果对比

Tab. 4 Comparison of results of ablation experiments

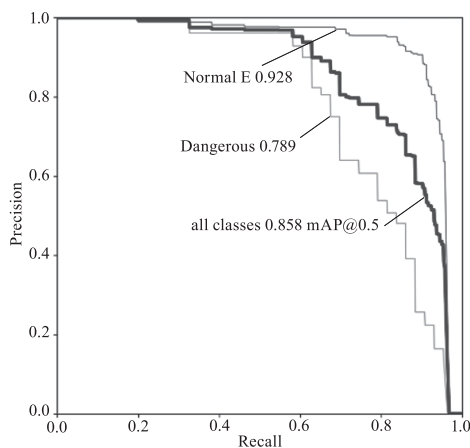
名称	改进部分				mAP@0.5/%
	部分数据增强	跨域迁移学习	SENet	α -CloU	
实验一	×	×	×	×	0.858
实验二	√	×	×	×	0.875
实验三	√	√	×	×	0.885
实验四	√	√	√	×	0.902
实验五	√	√	√	√	0.922

4 结果分析

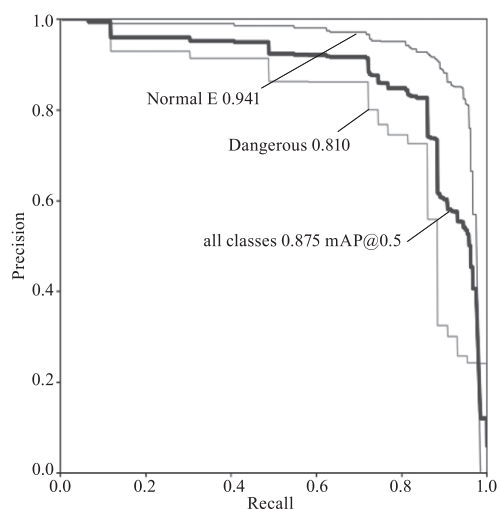
分别从精度分析和效果分析两个方面验证本文提出四种改进方法。对比实验一和实验二,类别一的平均精度上升了 1.3 %,类别二的平均精度上升了 2.1 %,总类别平均精度上升了 1.7 %,验证了数据处理改进方法对 YOLOv5s 模型检测的精度提升的有效性;在实验二到实验五中分别验证了模型改进中三种方法的有效性。对比实验二和实验五,类别一的平均精度上升了 1.8 %,类别二的平均精度上升了 7.4 %,总类别平均精度上升了 4.7 %。同时,改进后的方法能有效降低 YOLOv5s 算法的漏检、误检概率。

4.1 精度分析

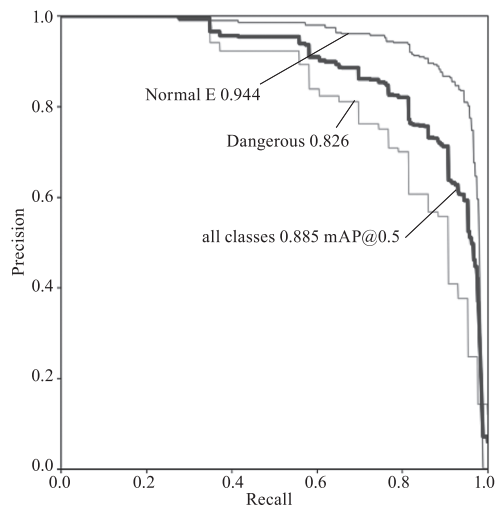
P-R 曲线图能够更为细致的反映各个类别的精度变化情况,实验一到实验五的各类别 P-R 曲线图如 8 所示。



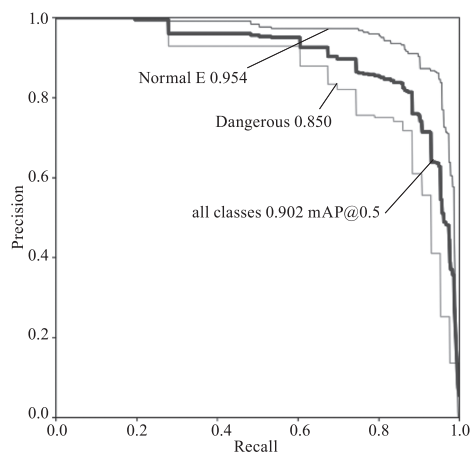
(a) 实验一 P-R 曲线图



(b) 实验二 P-R 曲线图



(c) 实验三 P-R 曲线图



(d) 实验四 P-R 曲线图

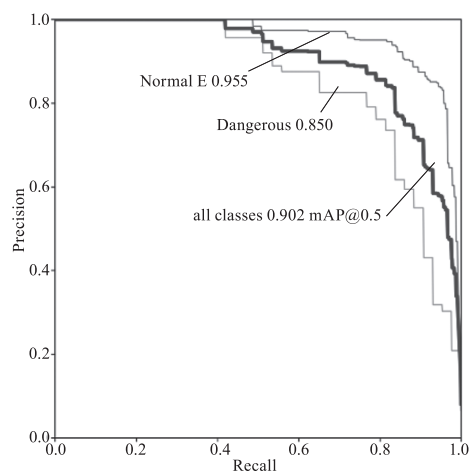
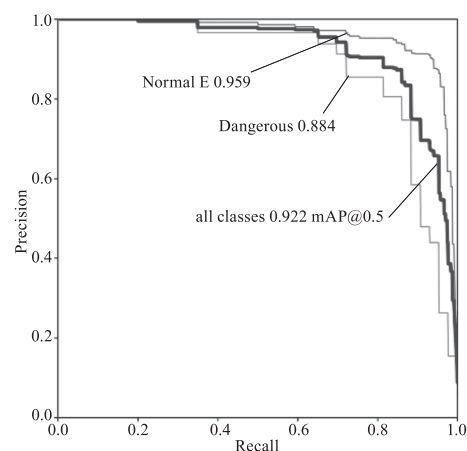
(e) 实验五 P-R 曲线图 ($\alpha = 2$)(f) 实验六 P-R 曲线图 ($\alpha = 3$)

图 8 P-R 曲线对比

Fig. 8 P-R curve comparison

实验一使用数据集一对 YOLOv5s 模型进行训练所得的各类别的 P-R 曲线图如图 8(a) 所示, 类别一的平均精度为 0.928, 类别二的平均精度为 0.789, 总类别平均精度为 0.858。

实验二使用数据集二对 YOLOv5s 模型进行训

练所得的各类别 P-R 曲线如图 8(b) 所示, 类别一的平均精度为 0.941, 类别二的平均精度为 0.810, 总类别平均精度为 0.875, 结果明显优于实验一。在 300 轮次的训练下, 低召回率范围的精度变化较大, 体现出了小样本数据集对 YOLOv5s 模型进行训练造成的不稳定性。

实验三在实验二的基础上在网络中的 Input 部分预加载网络权重训练, 各类别 P-R 曲线图如图 8(c) 所示, 有效改善了低召回率范围下精度变化较大的问题, 类别一的平均精度为 0.944, 类别二的平均精度为 0.826, 总类别平均精度为 0.885。

实验四以实验三为基础, 在网络中的 Backbone 部分插入 SENet 模块, 通过通道赋予权重的方式抑制用处不大的特征图层, 进一步提升检测精度。各类别的 P-R 曲线图如图 8(d) 所示, 类别一的平均精度为 0.954, 精度提升 1%; 类别二的平均精度为 0.850, 精度提升 2.4%; 总类别平均精度为 0.902, 精度提升 1.7%。

实验五以实验四为基础, 在网络中的 Prediction 部分使用 α -CIoU_Loss 替换网络中原有的 GIoU_Loss, CIoU 与 GIoU 相比能够加速网络中梯度的更新速率, α 参数的引入提高了高 IoU 预测框的相对损失权重, 这一改进使低 IoU 目标被检测到的更少, 造成模型在 AP50 的性能下降, 但可以通过在高 AP 时的提升进行弥补, 因此总体的检测精度是上升的, 不同的是样本中具有更多真阳性和假阳性目标框。由于 α 参数可自由选取为大于零的实数, 分别将其赋值为 2 和 3, 图 8 中的 (e)、(f) 两图分别为 α 取 2 和 3 的 P-R 曲线图, 可以看到, 当 α 值取 3 时比 α 取 2 时的效果更好。

4.2 效果分析

图 9 展示了改进后 YOLOv5s 检测效果图 9(b) 与改进前 YOLOv5s 检测效果图 9(a)。图 9 中的 1、2 组图像显示出改进后的网络可以有效的降低漏检的频率, 图 9 第 1 组 (b) 图中成功检测出了 (a) 图中未能检测到的目标; 第 2 组 (b) 图中成功检测出了 (a) 图中未能检测到的被遮挡目标。图 9 中的 3、4 组图像显示出改进后的网络可以有效的降低误检的频率, 第 3 组 (a) 图中将骑自行车的行人误检为单人驾驶电动自行车行为, 在 (b) 图中没有出现误检现象; 第 4 组 (a) 图中将被汽车遮挡住的行人误检

为单人驾驶电动自行车行为,在(b)图中没有出现误检现象。检测结果表明,改进后的 YOLOv5s 模型能够很好的应用于红外图像中电动自行车驾驶行为检测。



图 9 检测效果对比

Fig. 9 Comparison of detection effects

5 结 论

本文针对夜间交通环境下红外摄像机拍摄图像质量不高造成的目标边缘模糊,复杂背景下目标自身、目标与目标之间、目标与其他目标之间相互遮挡造成的漏检、误检问题,以及小样本数据集训练效果不佳的问题,网络梯度更新速率慢、高 IoU 锚框的预测数量少的问题,提出一种基于 YOLOv5s 改进算法。通过消融实验的方式,将改进后的 YOLOv5s 应用于红外图像的目标检测问题,实现了夜间道路下的两种电动自行车驾驶行为检测,并有效降低了红外目标漏检、误检的概率,提升了检测精度。

改进后的 YOLOv5s 模型大小为 15.3 MB,比原

算法增加约 8%,检测单幅红外图像平均耗时 16.5 ms,比原算法检测速度提升约 7.6%,具备较好的移植能力和检测速度,未来可运用于城市智能交通和精细管理中。

参考文献:

- [1] Zhu Xinyu, Chu Zhaoming, Zhu Jian'an, et al. Analysis on the causes and countermeasures of electric bicycle traffic accidents in China[J]. Urban Traffic, 2021, 19(6): 64 - 70. (in Chinese)
朱新宇, 褚昭明, 朱建安, 等. 中国电动自行车交通事故析因及对策建议[J]. 城市交通, 2021, 19(6): 64 - 70.
- [2] Xiao Yuqing, Yang Huimin. Research on application of object detection algorithm in traffic scene[J]. Computer Engineering and Applications, 2021, 57(6): 30 - 41. (in Chinese)
肖雨晴, 杨慧敏. 目标检测算法在交通场景中应用综述[J]. 计算机工程与应用, 2021, 57(6): 30 - 41.
- [3] Alexey Bochkovskiy, Wang Chien-Yao, Liao Hong-Yuan-Mark. Yolov4: optimal speed and accuracy of object detection[J]. arXiv Preprint arXiv:2004.10934, 2020.
- [4] Duan Huijun, Wang Zhigang, Wang Yan. Two-channel salient target recognition algorithm based on improved YOLO network[J]. Laser & Infrared, 2020, 50(11): 1370 - 1378. (in Chinese)
段辉军, 王志刚, 王彦. 基于改进 YOLO 网络的双通道显著性目标识别算法[J]. 激光与红外, 2020, 50(11): 1370 - 1378.
- [5] Liu Zhijia, Wang Xuan, Zhao Jinbo, et al. Infrared image target detection based on YOLO algorithm[J]. Laser & Infrared, 2020, 50(12): 1512 - 1520. (in Chinese)
刘智嘉, 汪璇, 赵金博, 等. 基于 YOLO 算法的红外图像目标检测的改进方法[J]. 激光与红外, 2020, 50(12): 1512 - 1520.
- [6] Tan Kangxia, Ping Peng, Qin Wenhui. Infrared image pedestrian detection based on YOLO model[J]. Laser & Infrared, 2018, 48(11): 1436 - 1442. (in Chinese)
谭康霞, 平鹏, 秦文虎. 基于 YOLO 模型的红外图像行人检测方法[J]. 激光与红外, 2018, 48(11): 1436 - 1442.
- [7] Wang Yongping, Zhang Hongmin, Peng Chuang, et al. Abnormal hot spot target detection method of high voltage switchgear based on YOLOv3[J]. Infrared Technology, 2020, 42(10): 983 - 987. (in Chinese)

- 王永平,张红民,彭闯,等. 基于 YOLOv3 的高压开关设备异常发热点目标检测方法[J]. 红外技术,2020,42(10):983-987.
- [8] Feng Yixiong,Zhao Bin,Zheng Hao, et al. Real-time detection of shaft surface defects based on integrated transfer learning[J]. Computer Integrated Manufacturing Systems, 2019,25(12):3199-3208. (in Chinese)
冯毅雄,赵彬,郑浩,等. 集成迁移学习的轴件表面缺陷实时检测[J]. 计算机集成制造系统,2019,25(12):3199-3208.
- [9] Jiang Wenping, JIANG Zhencun. Research on YOLOv5 initial fire detection based on multiple transfer learning [J]. Fire Science and Technology, 2021, 40(1): 109-112. (in Chinese)
蒋文萍,蒋珍存. 基于多重迁移学习的 YOLOv5 初期火灾探测研究[J]. 消防科学与技术, 2021, 40(1): 109-112.
- [10] Wang Zhuo, Wang Jian, Wang Xiaoxiong, et al. Light-weight real-time apple detection method based on improved YOLOv4[J]. Transactions of the Chinese Society for Agricultural Machinery, 1-9. (in Chinese)
王卓,王健,王枭雄,等. 基于改进 YOLOv4 的轻量级苹果实时检测方法[J]. 农业机械学报,2022,53(8):294-302.
- [11] Jie Hu, Shen Li, Sun Gang. Squeeze-and-excitation networks [C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2018.
- [12] Hamid RezaTofighi, Tsoi Nathan, Gwak JunYoung, et al. Generalized intersection over union: a metric and a loss for bounding box regression [C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2019.
- [13] Zhaohui Zheng, Wang Ping, Liu Wei, et al. Distance-IoU loss: faster and better learning for bounding box regression [J/OL]. <https://arxiv.org/pdf/1911.08287.pdf>.
- [14] Jiabo He, Erfani Sarah, Ma Xingjun, et al. Alpha-IoU: a family of power intersection over union losses for bounding box regression[J]. Advances in Neural Information Processing Systems, 2021, 34.
- [15] Joseph Redmon, Farhadi Ali. Yolov3: an incremental improvement[J]. arXiv Preprint arXiv:1804.02767, 2018.