

轻量型 YOLOv5s 车载红外图像目标检测

刘彦磊*, 李孟喆, 王宣宣

(河南师范大学 物理学院, 河南省红外材料光谱测量与应用重点实验室, 河南 新乡 453007)

摘要: 车载红外图像的目标检测是自动驾驶进行道路环境感知的重要方式。针对现有车载红外图像目标检测算法中内存利用率低、计算复杂和检测精度低的问题, 提出了一种改进 YOLOv5s 的轻量型目标检测算法。首先, 将 C3Ghost 和 Ghost 模块引入 YOLOv5s 检测网络, 以降低网络复杂度。其次, 引进 α IoU 损失函数, 以提升目标的定位精度和训练效率。然后, 降低网络结构下采样率, 并利用 KMeans 聚类算法优化先验框大小, 以提高小目标检测能力。最后, 分别在主干网络和颈部引入坐标注意力 (Coordinate Attention, CA) 和空间深度卷积模块进一步优化模型, 提升模型特征的提取能力。实验结果表明, 相对于原 YOLOv5s 算法, 改进算法的模型大小压缩 78.1%, 参数数量和每秒千兆浮点运算数分别减少 84.5% 和 40.5%, 平均检测精度和检测速度分别提升 4.2% 和 10.9%。

关键词: 自动驾驶; 目标检测; 红外图像; 轻量型; YOLOv5s

中图分类号: TP391.4

文献标志码: A

doi: 10.37188/CO.2022-0254

Lightweight YOLOv5s vehicle infrared image target detection

LIU Yan-lei*, LI Meng-zhe, WANG Xuan-xuan

(Henan Key Laboratory of Infrared Materials & Spectrum Measures and Applications,
College of Physics, Henan Normal University, Xinxiang 453007, China)

* Corresponding author, E-mail: liuyanlei@htu.edu.cn

Abstract: Vehicle infrared image target detection is an important way of road environment perception for autonomous driving. However, existing vehicle infrared image target detection algorithms have defects, such as low memory utilization, complex calculation and low detection accuracy. In order to solve the above problems, an improved YOLOv5s lightweight target detection algorithm is proposed. Firstly, the C3Ghost and Ghost modules are introduced into the YOLOv5s detection network to reduce network complexity. Secondly, the α IoU loss function is introduced to improve the positioning accuracy of the target and the networks training efficiency. Then, the subsampling rate of the network structure is reduced and the KMeans clustering algorithm is used to optimize the prior anchor size to improve the ability to detect of small targets. Finally, coordinate attention and spatial depth convolution modules are respectively introduced into the Backbone and Neck to further optimize the model and improve the feature extraction of the model. The experimental results show that compared with the original YOLOv5s algorithm, the improved algorithm can compress the model size by 78.1%, reduce the number of parameters and Giga Floating-point Operations Per Second by

收稿日期: 2022-12-14; 修订日期: 2023-01-06

基金项目: 国家自然科学基金 (No. 61905068)

Supported by National Natural Science Foundation of China (No. 61905068)

84.5% and 40.5% respectively, and improve the mean average precision and detection speed by 4.2% and 10.9%, respectively.

Key words: autonomous driving; target detection; infrared image; lightweight; YOLOv5s

1 引言

随着自动驾驶技术的不断发展,将其作为驾驶员辅助工具或者完全替代驾驶员,可以有效减少交通事故的发生^[1]。目标检测技术能帮助驾驶员或者计算机了解道路情况^[2],在自动驾驶领域扮演着不可或缺的角色。红外图像目标检测以其探测距离远、抗干扰能力强、特征易于提取的优点,被广泛应用于自动驾驶领域^[3-4]。然而,受物体间红外辐射的相互影响和红外成像设备成本的限制,红外图像存在分辨率和对比度低、纹理特征少等缺陷,给红外图像目标检测带来了挑战。因此,在有限资源的硬件设备下,开发高效准确的轻量级车载红外图像目标检测算法具有重要意义。

得益于人工智能技术的不断突破,深度学习算法被应用于车载红外图像目标检测任务中。不同于基于手工特征的传统方法^[5],深度学习算法有着更好的泛化性,主要分为两阶段方法和单阶段方法。两阶段车载红外图像目标检测算法^[6-8]分为两个步骤,先选择出几个可能包含目标的区域,再通过卷积神经网络获得区域内的目标分类,但其存在检测速度较慢的问题。单阶段方法直接通过回归来获取图像中目标的位置和类别^[9],因其优异的检测效率被应用于车载红外图像目标检测的相关研究中。Gong 等人^[10]针对红外图像检测速度无法满足实时性的问题,提出采用轻量级检测网络 YOLOv3-tiny 作为红外图像车辆检测算法,并通过深化网络结构和利用 KMeans 聚类算法优化先验框的方式提升检测精度,但检测网络深度不够,检测精度有待进一步提升。Sun 等人^[11]改进 YOLOv3 的主干网络为 EfficientNet 来加强特征提取,并使用去噪器(DRUNet)来提高输入图像质量,进一步提升算法的检测精度。之后,YOLOv4^[12]和 YOLOv5 相继诞生^[13],目标检测的准确性和速度得到不断提升。Jiang 等人^[14]通过

改进 YOLOv4-tiny 产生了一种三检测头轻型红外环境感知实时检测网络模型(IEPet),改进算法相比于 YOLOv4-tiny 模型更小,检测精度更高。Wu 等人^[15]利用重参数网络(RepVGG)重建 YOLOv4 主干网络和信道修剪压缩的方法简化模型,使用空间金字塔池化(SPP)获得不同的感受野信息,以提高模型检测的准确性,以适应功耗有限制的嵌入式平台。Xin 等人^[16]基于 YOLOv5s 算法增加了一层小目标检测层,并将卷积注意力模块(Convolutional Block Attention Module, CB-AM)插入主干网络,且用自注意力 Swin Transformer 编码器替换 C3 模块,改善其在复杂环境下漏检率高、误报率高等问题,但增加了检测网络的复杂性。与此同时,参考先进算法 SSD^[17],Dai 等人^[18]提出一种专用于自动驾驶的轻量级红外图像卷积神经网络(TIRNet),以实现可靠和有效的目标检测。该方法主要由自定义 SSD 和残差分支(Residual Branch, RB)组成,解决道路目标检测中速度慢、计算成本高的问题。然而,上述单阶段车载红外图像目标检测算法仍然存在检测模型复杂度高和鲁棒性差等问题,这是部署在内存和计算资源有限的车载计算机上进行红外图像目标检测的重大障碍,在低模型体积下获得高模型检测性能的问题仍有待解决。

针对车载红外图像目标检测存在的问题,本文提出了一种改进 YOLOv5s 的轻量级车载红外目标检测算法,主要研究内容如下:

- (1)引入 C3Ghost 和 Ghost 模块,实现检测模型压缩,以降低计算成本和参数量;
- (2)采用 α IoU_Loss 替换 CIoU_Loss 作为预测框的回归损失函数,通过多个损失函数的对比实验验证其性能;
- (3)改进多尺度特征融合,降低网络深度,减少红外图像的特征信息丢失;
- (4)在主干网络和颈部分别引入坐标注意力(CA)和空间深度卷积模块,并对比了融合不同注意力机制的检测效果。

2 数据集和 YOLO 算法

本文研究基于美国 FLIR 公司发布的用于自动驾驶环境感知的(Advanced Driver Assistance System, ADAS 2.0)红外数据集^[19],下文简称 FLIR 红外数据集。关于汽车和行人的边框面积比的统计数据表明,FLIR 红外数据集主要由中等尺度(汽车)和小尺度(行人)目标组成,这些目标的检出通常比大型对象更重要和困难。因此,本文选择行人和车辆两个类别进行目标检测。

YOLOv5 是 YOLO 系列目标检测算法的代

表,且仍然在不断地更新迭代。如图 1 所示, YOLOv5 v6.1 由主干网络(Backbone)、颈部(Neck)和检测头(Head)3 个主要部分组成。YOLOv5 的 Backbone 部分是 CSPDarknet53,它是受 Darknet53 的跨阶段局部网络(Cross Stage Partial Network, CSPNet)的启发而设计的。在 Neck 部分,采用特征金字塔网络(FPN)加上路径聚合网络(PAN)的双结构进行多层特征融合。在 Head 部分, YOLOv5 模型使用 CIoU_Loss 作为预测框的损失函数,并使用非极大值抑制方法(NMS)筛选得分高的预测框。最后,在 3 个不同大小的特征图上进行多尺度预测,以检测小、中、大对象。

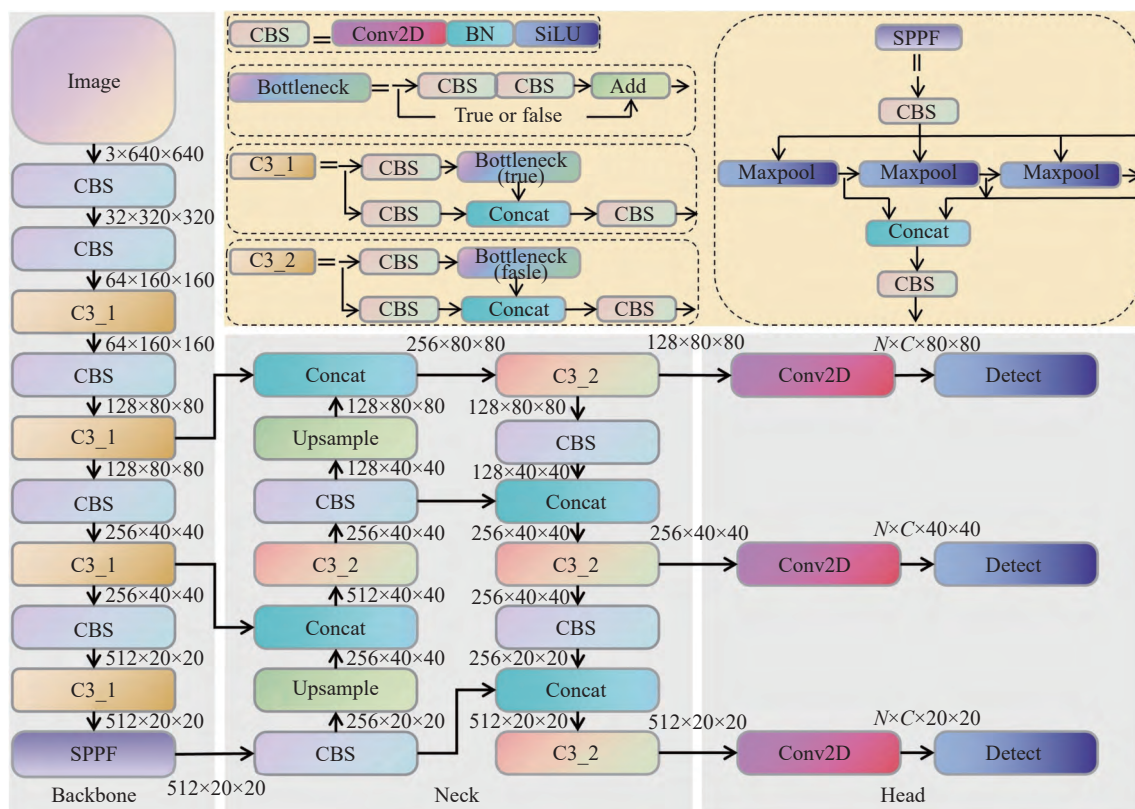


图1 YOLOv5s 算法结构

Fig. 1 YOLOv5s algorithm structure

3 YOLOv5s 算法改进

YOLOv5 算法以其性能和速度著称,易于进行部署,可以很好地满足车载红外图像目标检测的要求。YOLOv5 有 5 种不同的型号,包括

YOLOv5n, YOLOv5s, YOLOv5m, YOLOv5l 和 YOLOv5x, 它们的参数量、体积和检测精度依次增加,检测速度依次降低。YOLOv5n 提取的特征太浅,其他模型的检测速度太慢,综合考虑后,选择 YOLOv5s 作为基线算法,并对其进行了改进,改进后的算法结构如图 2 所示。

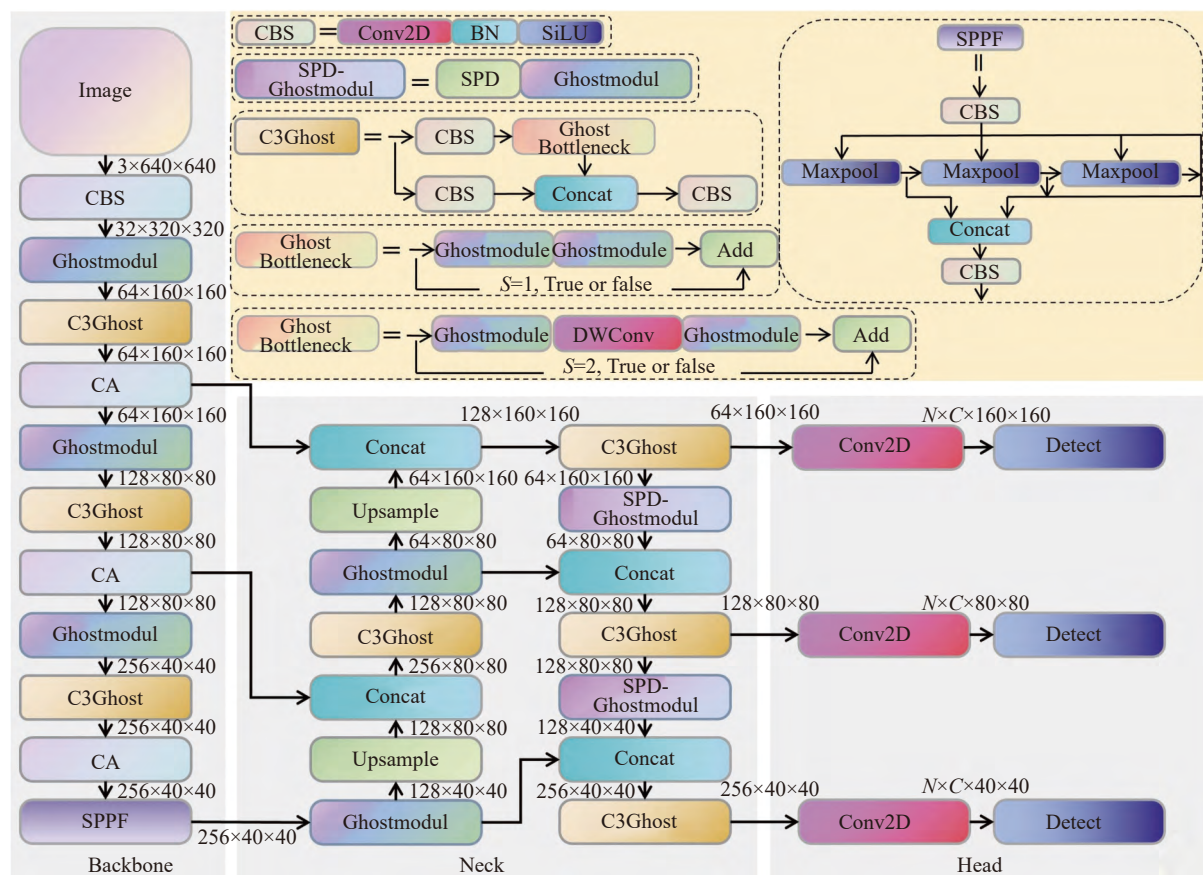


图 2 改进 YOLOv5s 算法结构

Fig. 2 Improved YOLOv5s algorithm structure

3.1 轻量化网络构建

为了使红外图像目标检测网络更加高效,减少模型的参数量和计算复杂度,突破车载计算机内存和计算资源的限制。本文引入了轻量化卷积模块 Ghost 和 C3Ghost^[20],替换了原算法中的卷积模块 CBS 和 C3 模块,具体结构如图 2 所示,其中 C3Ghost 是用 Ghost 模块替换了 C3 模块瓶颈层 Bottleneck 中的普通卷积而产生的。如图 3(a)所示,给定输入,进行普通卷积,将输出固有特征图,而 Ghost 模块将深度神经网络中的普通卷积

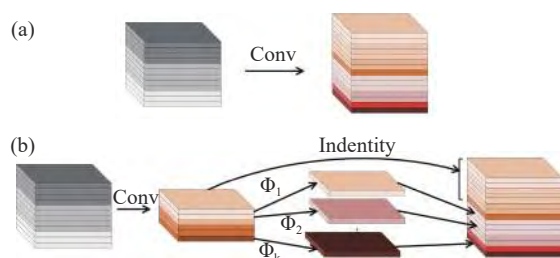
分为两部分,如图 3(b)所示。Ghost 模块先对输入特征图进行普通卷积输出固有特征图;然后,对得到的固有特征进行线性分组卷积,最终的特征图由固有特征和 Ghost 特征拼接而成。

3.2 损失函数的选择

损失函数反映了检测算法性能的好坏,可以帮助神经网络快速收敛。由于红外图像具有信噪比低的特点,而 $\alpha\text{IoU_Loss}$ ^[21] 对于噪声又有很强的鲁棒性,所以本文将 CIoU_Loss 改进为 $\alpha\text{IoU_Loss}$,使用功率参数 α 自适应地调整 IoU 目标的损失和梯度,从而进一步提升预测框和真实框的回归精度。经实验验证, $\alpha=3$ 时在大多数情况下表现良好,并且不会引入额外的参数或增加训练推断时间。 $\alpha\text{IoU_Loss}$ 表示为:

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

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

图 3 (a)普通卷积和(b)Ghost 卷积(Φ 为线性操作)Fig. 3 (a) Ordinary convolution and (b) Ghost convolution (Φ is a linear operation)

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

其中, IoU 为预测框和真实框的交并比, b 表示预测框中心点, b^{gt} 表示真实框中心点, ρ 是欧几里得距离, c 是预测框和真实框最小矩形的对角线长度, v 为纵横比, β 为正权衡参数, w^{gt} 和 h^{gt} 是真实框的宽度和高度, w 和 h 是预测框的宽度与高度。

3.3 多尺度特征融合改进

在车载红外图像目标检测过程中, 更好的检测小目标意味着可以更早期地探测远方的路况, 从而提升行车安全性。YOLOv5s 原算法在 Backbone 部分进行了 5 次下采样, 分别为 $2\times$ 、 $4\times$ 、 $8\times$ 、 $16\times$ 和 $32\times$, 如图 1 所示。当输入图像为 640×640 时, Head 部分对应小、中和大对象检测层的特征图大小分别为 80×80 、 40×40 和 20×20 。由于红外图像相对于可见光图像分辨率低, 导致红外图像有更多的小目标且特征信息较少。因此, 考虑通过剔除 $32\times$ 下采样, 降低采样率, 以避免过深的卷积神经网络导致的红外图像细节丢失, 并生成一层新的小目标检测层, 新结构如图 2 所示。最终, 算法在 Backbone 部分进行了 4 次下采样, 其中 $4\times$ 、 $8\times$ 和 $16\times$ 下采样特征图被送入 Neck 部分进行多尺度融合。当输入图像为 640×640 时, Head 部分对应于小、中和大对象检测的特征图大小分别为 160×160 、 80×80 和 40×40 。相比于四检测头方法^[16]通过增加上采样获得 160×160 尺度的小目标检测层, 本文方法避免了检测网络复杂度和推理时间的增加。

此外, 由于数据集和尺度特征的改变, 原有的先验框大小已经不适合改进后的算法。在本文算法中, 将预测框中真实框的召回比例 (Best Possible Recall, BPR) 设置为 0.99, 当 BPR 值小于 0.99 时, 重新生成先验框, 原算法 BPR 限制值为 0.98。此时, 基于 KMeans 聚类算法^[13]获得了一组新先验框, 如表 1 所示, BPR 值高达 0.999, 优于默认

表 1 优化后先验框大小

Tab. 1 Optimized prior anchor size

特征图尺度	160×160	80×80	40×40
感受野大小	小	中	大
	[6,8]	[14,37]	[35,94]
先验框	[7,19]	[31,26]	[96,68]
	[15,13]	[50,37]	[154,145]

先验框。

3.4 Backbone 和 Neck 改进

当神经网络处理大量的信息输入时, 注意力机制会使网络迅速集中在一些需要处理的关键信息上。坐标注意力 CA^[22] 兼顾了图像的通道信息和空间信息。为了提升车载红外图像目标检测算法的特征提取能力, 使模型能更准确地定位和识别感兴趣的对象, 本文引入轻量化注意力机制 CA, 嵌入 Backbone 部分, CA 的具体结构如图 4 所示。

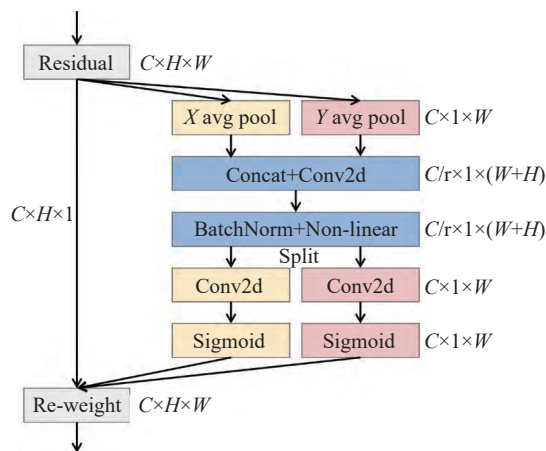


图 4 CA 结构

Fig. 4 CA structure

现有卷积神经网络中不可避免地要进行跨步卷积和池化, 这在一定程度上导致了红外图像特征信息的丢失。空间深度卷积 (SPD-Conv)^[23] 由空间深度层和非跨步卷积层组成, 有助于消除跨步卷积和池化引起的特征信息丢失问题。因此, 本文在权衡运算参数量和检测精度的情况下, 将 Neck 部分 PAN 结构中的下采样卷积更换为空间深度卷积。图 5 给出了当比例因子 $scale=2$ 时的空间深度卷积示意图, 其中, S 为图像矩阵形状参数, $s=1$ 中的 s 为卷积比。Backbone 和 Neck 部分的改进结构如图 2 所示。

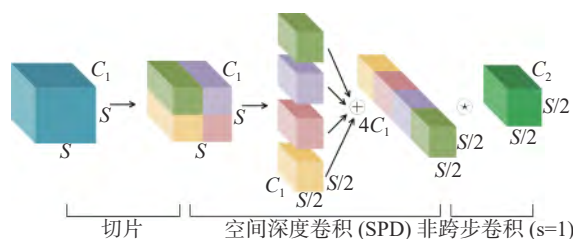


图 5 空间深度卷积 (Scale=2)

Fig. 5 SPD-Conv (Scale=2)

4 实验结果及分析

4.1 准备工作

本文遍历了整个 FLIR 红外数据集的对象,合并汽车类和其他小型商用车辆为汽车类,合并行人和骑自行车者为行人类,剔除其他类别标签。然后,将清洗后的 FLIR 红外图像调整到相同的大小以进行训练和验证,最终 FLIR 红外数据集总共包括 10249/1097 张训练/验证红外图像。

模型训练的硬件环境为 Intel(R) Core(TM) i7-12700KF CPU 和 RTX 3060Ti GPU,软件环境为 Python3.6、CUDA11.0、Pytorch1.7.1。本文的检测模型从零开始培训,不使用迁移学习,因此,可以检验每个算法的真实学习能力,学习 FLIR 红外数据集最真实的特征。与此相对应的,增加训练轮次到 400epoch。送入深度学习网络训练的图像大小设置为 640×640,训练优化器为随机梯度优化器(SGD),IoU 阈值为 0.5,动量设为 0.937,初始学习率为 0.001,权重衰减设置为 0.0005,批大小(Batch size)设置为 16。此外,改进算法 Mosaic、MixUp 和 Copy-Paste 数据增强^[13]的权重分别设置为 1.0、0.1 和 0.1,来丰富数据集,提高模型的检测精度和鲁棒性,如图 6 所示,其他未提及的超参数与 YOLOv5s 中的默认参数相同。

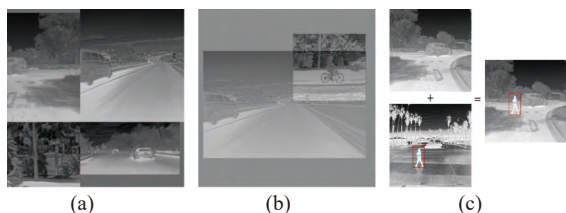


图 6 数据增强结果。(a)Mosaic 增强;(b)MixUp 增强;(c)Copy-Paste 增强

Fig. 6 Data augmentation results. (a) Mosaic augmentation; (b) MixUp augmentation; (c) Copy-Paste augmentation

4.2 评价标准

准确率(Precision rate, P)表示检测出的阳性样本中真实阳性样本的比例,用来评估车载红外图像目标预测是否准确。召回率(Recall rate, R)表示检测出的真实阳性样本在总的真实阳性样本中的比例,反映目标是否全部被发现。平均精度(Average Precision, AP)表示由 P 和 R 曲线和坐

标轴包围的区域的面积。平均检测精度(mean Average Precision, mAP)是所有类别计算的平均精度 AP 的平均值,是用来评价模型精度的综合性参数。具体如式(4)~式(7)所示,其中, tp 为真阳性, fp 为假阳性, fn 为假阴性, N 为类别数量。

$$P = \frac{tp}{tp + fp} \quad (4)$$

$$R = \frac{tp}{tp + fn} \quad (5)$$

$$AP = \int_0^1 PdR \quad (6)$$

$$mAP = \frac{\sum_{i=1}^N AP_i}{N} \quad (7)$$

此外, FPS 表示每秒检测的图像帧数,用来衡量检测速度; GFLOPs 和 Params 分别表示每秒千兆浮点计算量和模型参数量,数值越小,模型或算法的复杂性越低; Size 表示模型所占的内存大小, t 表示模型训练完成所耗费的时间。

4.3 对比实验

为验证改进后的 YOLOv5s 在 FLIR 红外数据集上的综合检测能力,本文设计了 6 个实验进行对比讨论,以验证所提出的各改进方法对 YOLOv5s 模型性能的影响。

实验 1 比较了原始 YOLOv5s 算法和轻量化后的 YOLOv5s 算法的性能,对比结果在表 2 中列出。如表 2 所示, YOLOv5s 引入 Ghost 模块和 C3Ghost 后,能够达到使检测网络轻量化的目的。虽然改进模型 YOLOv5s-G 的平均检测精度有略微下降,但是模型的参数量和计算量分别减少了 47.6% 和 49.4%,模型大小压缩了 45.5%,算法的训练速度也提升了 38.0%。

通过引入 α IoU_Loss 作为改进算法预测框的回归损失函数,加快检测网络的收敛速度,增强检测网络对有噪声的红外图像的鲁棒性。设计实验 2 对当前较先进的回归损失函数进行对比实验,实验对比结果如表 3 所示。 YOLOv5-G 算法默认使用 CIoU_Loss 作为回归损失函数,训练时间为 30.25 h,训练后的平均检测精度为 77.5%。使用 EIoU_Loss、SIoU_Loss 和 α IoU_Loss 相比使用 CIoU_Loss,训练时间分别缩短了 19.6%、

表 2 YOLOv5s 和 YOLOv5s-G 轻量化性能对比

Tab. 2 Performance comparison of lightweight for YOLOv5s and YOLOv5s-G

Model	<i>t</i> /hours	Size/MB	Params/M	GFLOPs	<i>P</i> (%)	<i>R</i> (%)	mAP(%)	FPS
YOLOv5s	48.77	13.70	7.02	15.8	87.1	69.8	80.8	119
YOLOv5s-G	30.25	7.46	3.68	8.0	86.1	66.3	77.5	137

18.6% 和 22.3%, 平均检测精度相比于 CIoU_Loss 分别提升 1.4%、0.3% 和 2.3%。由此可见, YOLOv5s-G- α IoU 在 FLIR 红外数据集上有更好的表现, 能加速收敛, 并提升检测精度。

表 3 不同损失函数性能对比

Tab. 3 Performance comparison of different loss functions

Model	<i>t</i> /hours	<i>P</i> (%)	<i>R</i> (%)	mAP(%)	FPS
YOLOv5s-G	30.25	86.1	66.3	77.5	137
YOLOv5s-G-EIoU	24.31	84.5	68.7	78.9	141
YOLOv5s-G-SIoU	24.62	85.8	67.2	77.8	139
YOLOv5s-G- α IoU	23.50	85.9	69.3	79.8	147

实验 3 对本文设计的降低下采样的 YOLOv5s-G2- α IoU 算法和四检测头算法 YOLOv5s-G1- α IoU 进行对比实验, 以验证改进模型对小目标的检测能力, 结果如表 4 所示。可以看到, 四检测头算

法 YOLOv5s-G1- α IoU 检测性能的增加是以增加参数量和计算复杂度为代价的, 而 YOLOv5s-G2- α IoU 则通过简单的降低采样率, 进一步对内存压缩了 63.4%, 参数量和计算量分别减少 74.2% 和 10%, 检测速度和平均检测精度也分别得到了 4.8% 和 3.1% 的提升。因此, 本文选择 YOLOv5s-G2- α IoU 结构。此外, KMeans 聚类算法针对 FLIR 红外数据集生成的新先验框, 与默认先验框的实验效果相比, YOLOv5s-G2- α IoU-KMeans 算法准确率增加了 1%, 平均检测精度增加了 0.1%。

为了进一步提升算法的检测精度, 本文将注意力机制嵌入算法的 Backbone 部分, 以增强模型的特征提取能力。实验 4 对比了主流的注意力机制嵌入算法的效果。如表 5 所示, 注意力 CA 有着最好的检测精度, YOLOv5s-G2- α IoU-KMeans-CA 以参数量和计算量的微量增加, 获得了平均检测精度增加了 1.3%。

表 4 多尺度融合性能对比

Tab. 4 Performance comparison of multi-scale fusion

Model	<i>t</i> /hours	Size/MB	Params/M	GFLOPs	<i>P</i> (%)	<i>R</i> (%)	mAP(%)	FPS
YOLOv5s-G- α IoU	23.50	7.46	3.68	8.0	85.9	69.3	79.8	147
YOLOv5s-G1- α IoU	26.89	8.60	3.75	9.6	86.0	73.6	83.6	125
YOLOv5s-G2- α IoU	24.56	2.73	0.95	7.2	84.5	72.8	82.9	154
YOLOv5s-G2- α IoU-KMeans	25.62	2.73	0.95	7.2	85.5	72.4	83	154

表 5 不同注意力机制性能对比

Tab. 5 Performance comparison of different attention mechanisms

Model	<i>t</i> /hours	Size/MB	Params/M	GFLOPs	<i>P</i> (%)	<i>R</i> (%)	mAP(%)	FPS
YOLOv5s-G2- α IoU-KMeans	25.62	2.73	0.95	7.2	85.5	72.4	83	154
YOLOv5s-G2- α IoU-KMeans-SE	30.95	2.75	0.96	7.2	86.0	73.5	84.1	149
YOLOv5s-G2- α IoU-KMeans-ECA	26.06	2.73	0.95	7.2	85.5	73.8	84.2	145
YOLOv5s-G2- α IoU-KMeans-CBAM	28.21	2.76	0.96	7.3	85.7	73.4	84	135
YOLOv5s-G2- α IoU-KMeans-CA	28.62	2.76	0.96	7.3	86.6	73.6	84.3	139

通过将模型的 Neck 部分 PAN 结构中的下采样方式改进为空间深度卷积, 来避免跨步卷积和池化导致的红外图像不同程度的特征信息丢失问题, 实验 5 的对比结果如表 6 所示。可见, 与普通

的下采样卷积相比, 经空间深度卷积改进下采样卷积的 YOLOv5s-G2- α IoU-KMeans-CA-SPD 算法, 准确率增加了 0.8%, 召回率增加了 1%, 平均检测精度增加了 0.7%。

表 6 空间深度卷积效果

Tab. 6 SPD-Conv effect

Model	t/hours	Size/MB	Params/M	GFLOPs	P(%)	R(%)	mAP(%)	FPS
YOLOv5s-G2- α IoU-Kmeans-CA	28.62	2.76	0.96	7.3	86.6	73.6	84.3	139
YOLOv5s-G2- α IoU-Kmeans-CA-SPD	30.28	3.0	1.09	9.4	87.4	74.6	85.0	132

实验 6 基于 FLIR 红外数据集对本文改进算法与先进算法 SSD、EfficientDet、YOLOv4+GhostNet、YOLOv5-MobileNetV3、YOLOv3-tiny、YOLOv4-tiny、YOLOv5n、YOLOv6-N、YOLOv7-tiny 和原算法 YOLOv5s 进行比较, 如表 7 所示。可以看到, 本文提出的轻量型算法相对于原算法模型大小压缩了 78.1%, 参数量减少了 84.5%, 平

均检测精度达到了 85%, 检测速度为 132 FPS; 相对于其他轻量化算法更有利于移动设备的模型移植。在模型体积最小和参数量最少的情况下, 平均检测精度最高, 且检测速度相对于原算法有进一步提升, 达到了检测精度和检测速度的均衡效果, 可以保证优越的车载检测性能。

表 7 与其他先进算法对比

Tab. 7 Comparison with other advanced algorithms

Model	Size/MB	Params/M	GFLOPs	P(%)	R(%)	mAP(%)	FPS
SSD	186.0	23.70	115.7	68.9	55.7	63.2	88
EfficientDet	302.0	39.40	107.5	72.8	58.4	67.8	52
YOLOv4+GhostNet	150.3	39.30	25.6	81.1	66.9	77.7	112
YOLOv5-MobileNetV3	7.9	4.0	9.3	83.7	67.5	76.9	128
YOLOv3-tiny	16.6	8.67	12.9	79.3	54.9	62.9	175
YOLOv4-tiny	12.9	6.27	16.2	78.9	57.3	67.2	149
YOLOv5n	3.7	1.76	5.1	83.6	66.1	76.6	164
YOLOv6-N	9.3	4.30	11.1	84.8	71.5	80.3	208
YOLOv7-tiny	12.3	6.02	13.2	84.2	74.7	83.6	143
YOLOv5s	13.7	7.02	15.8	87.1	69.8	80.8	119
proposed in this paper	3.0	1.09	9.4	87.4	74.6	85.0	132

图 7 显示了本文改进算法和 YOLO 系列的轻量化版本在复杂场景(昏暗、强光、多目标遮挡等)下的检测效果。对比可发现, 红外图像目标检测抗干扰能力强, 不受复杂环境的影响。根据图 7 中昏暗和车灯眩光两场景下各算法的对比效果可知, 本文改进算法目标检测的置信水平得分

更高。而图 7 中强光和有多目标两场景下的检测效果表明, 本文改进算法相对于其他算法对小目标检测更准确, 不存在漏检现象。因此, 本文提出的改进算法在同类算法中有一定的优越性, 能够为该算法在车载终端设备中的应用提供技术参考, 促进自动驾驶的发展。

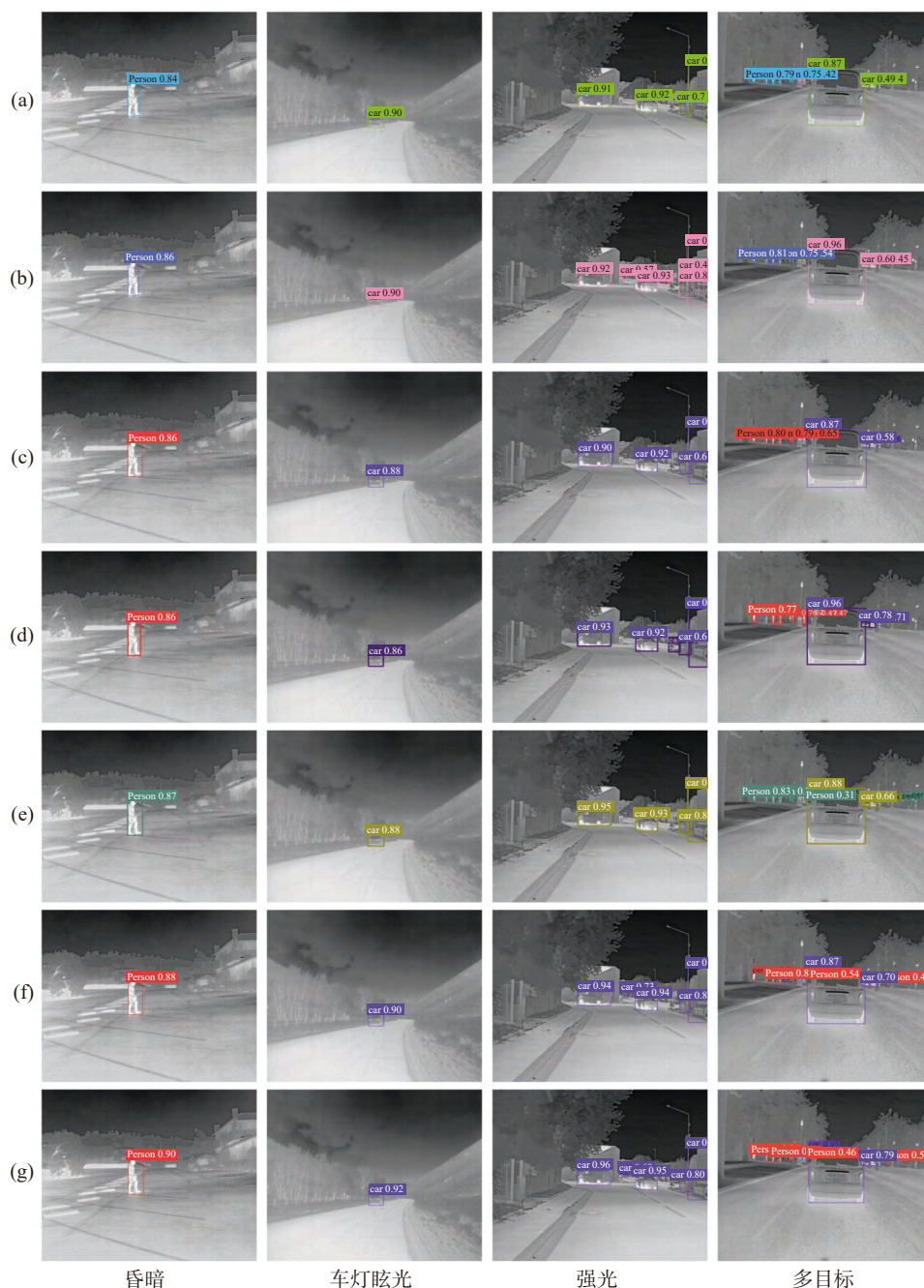


图 7 几种不同算法的检测效果。(a) YOLOv3-tiny; (b) YOLOv4-tiny; (c) YOLOv5n; (d) YOLOv6-N; (e) YOLO7-tiny; (f) YOLO5s; (g) 本文算法

Fig. 7 Detection results of different algorithms. (a) YOLOv3-tiny; (b) YOLOv4-tiny; (c) YOLOv5n; (d) YOLOv6-N; (e) YOLO7-tiny; (f) YOLO5s; (g) proposed in this paper

5 结 论

本文提出了一种改进 YOLOv5s 的轻量型车载红外图像目标检测算法。该算法基于 FLIR 红外数据集训练网络模型来检测行人和车辆。对改进算法和其他轻量化算法及原算法分别进行了比较。实验结果表明:改进算法相较于其它算法

在模型结构和检测精度方面有明显的优势。改进后算法的模型体积仅 3.0 MB, 参数量仅 1.09 M, 平均检测精度 mAP 为 85.0%, 检测速度高达 132 FPS, 有效解决了现有车载红外图像目标检测算法中内存利用率低、计算复杂和检测精度低的问题。本文研究工作能够实现车载红外图像的快速、准确检测, 可以为自动驾驶中的环境感知研究提供参考。

参考文献:

- [1] MUHAMMAD K, ULLAH A, LLORET J, *et al.*. Deep learning for safe autonomous driving: current challenges and future directions[J]. *IEEE Transactions on Intelligent Transportation Systems*, 2021, 22(7): 4316-4336.
- [2] TAKUMI K, WATANABE K, HA Q SH, *et al.*. Multispectral object detection for autonomous vehicles[C]. *Proceedings of the on Thematic Workshops of ACM Multimedia 2017*, ACM, 2017: 35-43.
- [3] CHOI Y, KIM N, HWANG S, *et al.*. KAIST multi-spectral day/night data set for autonomous and assisted driving[J]. *IEEE Transactions on Intelligent Transportation Systems*, 2018, 19(3): 934-948.
- [4] LIU Q, ZHUANG J J, MA J. Robust and fast pedestrian detection method for far-infrared automotive driving assistance systems[J]. *Infrared Physics & Technology*, 2013, 60: 288-299.
- [5] 任凤雷, 周海波, 杨璐, 等. 基于双注意力机制的车道线检测[J]. *中国光学 (中英文)*, 2023, 16(3): 645-653.
- REN F L, ZHOU H B, YANG L, *et al.*. Lane detection based on dual attention mechanism[J]. *Chinese Optics*, 2023, 16(3): 645-653. (in Chinese)
- [6] WANG H, CAI Y F, CHEN X B, *et al.*. Night-time vehicle sensing in far infrared image with deep learning[J]. *Journal of Sensors*, 2016, 2016: 3403451.
- [7] GALARZA-BRAVO M A, FLORES-CALERO M J. Pedestrian detection at night based on faster R-CNN and far infrared images[C]. *Proceedings of the 11th International Conference on Intelligent Robotics and Applications*, Springer, 2018: 335-345.
- [8] CHEN Y F, XIE H, SHIN H. Multi - layer fusion techniques using a CNN for multispectral pedestrian detection[J]. *IET Computer Vision*, 2018, 12(8): 1179-1187.
- [9] 王驰, 于明坤, 杨辰烨, 等. 抛撒地雷的夜视智能探测方法研究[J]. *中国光学*, 2021, 14(5): 1202-1211.
- WANG CH, YU M K, YANG CH Y, *et al.*. Night vision intelligent detection method of scatterable landmines[J]. *Chinese Optics*, 2021, 14(5): 1202-1211. (in Chinese)
- [10] GONG J, ZHAO J H, LI F, *et al.*. Vehicle detection in thermal images with an improved yolov3-tiny[C]. *Proceedings of 2020 IEEE International Conference on Power, Intelligent Computing and Systems*, IEEE, 2020: 253-256.
- [11] SUN M Y, ZHANG H CH, HUANG Z L, *et al.*. Road infrared target detection with I - YOLO[J]. *IET Image Processing*, 2022, 16(1): 92-101.
- [12] 吴海滨, 魏喜盈, 刘美红, 等. 结合空洞卷积和迁移学习改进YOLOv4的X光安检危险品检测[J]. *中国光学*, 2021, 14(6): 1417-1425.
- WU H B, WEI X Y, LIU M H, *et al.*. Improved YOLOv4 for dangerous goods detection in X-ray inspection combined with atrous convolution and transfer learning[J]. *Chinese Optics*, 2021, 14(6): 1417-1425. (in Chinese)
- [13] 张印辉, 庄宏, 何自芬, 等. 氨气泄漏混洗自注意力轻量化红外检测[J]. *中国光学 (中英文)*, 2023, 16(3): 607-619.
- ZHANG Y H, ZHUANG H, HE Z F, *et al.*. Lightweight infrared detection of ammonia leakage using shuffle and self-attention[J]. *Chinese Optics*, 2023, 16(3): 607-619. (in Chinese)
- [14] JIANG X H, CAI W, YANG ZH Y, *et al.*. IEPet: a lightweight multiscale infrared environmental perception network[J]. *Journal of Physics:Conference Series*, 2021, 2078: 012063.
- [15] WU ZH L, WANG X, CHEN CH. Research on lightweight infrared pedestrian detection model algorithm for embedded Platform[J]. *Security and Communication Networks*, 2021, 2021: 1549772.
- [16] XIN X L, PAN F, WANG J CH, *et al.*. SwinT-YOLOv5s: improved YOLOv5s for vehicle-mounted infrared target detection[C]. *Proceedings of the 41st Chinese Control Conference (CCC)*, IEEE, 2022: 7326-7331.
- [17] ZHAI SH P, SHANG D R, WANG SH H, *et al.*. DF-SSD: an improved SSD object detection algorithm based on DenseNet and feature fusion[J]. *IEEE Access*, 2020, 8: 24344-24357.
- [18] DAI X R, YUAN X, WEI X Y. TIRNet: object detection in thermal infrared images for autonomous driving[J]. *Applied Intelligence*, 2021, 51(3): 1244-1261.
- [19] 2022. FREE FLIR Thermal Dataset for Algorithm Training. [Online]. Available: <https://www.flir.com/oem/adas/adas-dataset-form>.
- [20] CAO M L, FU H, ZHU J Y, *et al.*. Lightweight tea bud recognition network integrating GhostNet and YOLOv5[J]. *Mathematical Biosciences and Engineering*, 2022, 19(12): 12897-12914.
- [21] HE J B, ERFANI S M, MA X J, *et al.*. Alpha-IOU: a family of power intersection over union losses for bounding box

- regression[C]. *Proceedings of the 34th Advances in Neural Information Processing Systems*, 2021.
- [22] ZHA M F, QIAN W B, YI W L, *et al.*. A lightweight YOLOv4-Based forestry pest detection method using coordinate attention and feature fusion[J]. *Entropy*, 2021, 23(12): 1587.
- [23] SUNKARA R, LUO T. No more strided convolutions or pooling: a new CNN building block for low-resolution images and small objects[C]. *Proceedings of the Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, Springer, 2022: 443-459.

作者简介:



刘彦磊(1986—), 男, 河南中牟人, 博士, 讲师, 2011 年, 2014 年于河南师范大学分别获得学士、硕士学位, 2018 年 6 月于北京理工大学获得博士学位, 主要从事红外光谱测量及应用技术方面的研究。E-mail: liuyanlei@htu.edu.cn