文章编号:1672-3961(2024)06-0019-10 DOI:10.6040/j.issn.1672-3961.0.2023.220

融合 FasterNet 和 RepVGG 的安全设备佩戴检测方法

张曼,孙凯军*,李翔,孙纪舟
(淮阴工学院计算机与软件工程学院, 江苏 淮安 223003)

摘要:针对工业场景下检测模型参数量多、安全绳易与背景混淆问题,提出一种融合 FasterNet 和 RepVGG 的目标检测算法(you only look once version 5 fasternet-repvgg-cutmix-bifpn, YOLOv5s-FRCB)。通过引入轻量级网络结构 FasterNet Block 和 RepVGG Block 替换 YOLOv5s 部分卷积层,显著减少模型参数量和加快检测速度,满足实时性需求;通过 BiFPN 特征强连通性提升模型特征学习能力;改进 Cutmix 数据增强方法,随机将目标嵌入输入图像,更新标签,缓解标签类别不平衡问题,提高泛化性。在自建安全设备佩戴检测数据集上进行试验,结果表明:YOLOv5s-FRCB mAP 值达到了 96.3%,算法模型内存减少 34%,是一种高效实用的安全设备佩戴检测方法;YOLOv5s-FRCB 能在保证准确率的同时,进一步降低计算量。

关键词:安全设备佩戴检测;目标检测;轻量化;数据增强;实时检测
中图分类号:TP391 **文献标志码**:A
引用格式:张曼,孙凯军,李翔,等. 融合 FasterNet 和 RepVGG 的安全设备佩戴检测方法[J].山东大学学报(工学版),2024,54(6):19-28.
ZHANG Man, SUN Kaijun, LI Xiang, et al. A security device wearing detection method integrating FasterNet and RepVGG[J]. Journal of Shandong University (Engineering Science), 2024, 54(6):19-28.

A security device wearing detection method integrating FasterNet and RepVGG

ZHANG Man, SUN Kaijun*, LI Xiang, SUN Jizhou
(Faculty of Computer and Software, Huaiyin Institute of Technology, Huaian 223003, Jiangsu, China)

Abstract: Addressing the issues of high parameter count in detection models and the ease of safety rope easily confused with the background in industrial scenarios, an object detection algorithm combining fasternet and RepVGG was proposed (you only look once version 5 FasterNet-RepVGG-CutMix-BiFPN, YOLOv5s-FRCB). This algorithm integrated FasterNet and RepVGG by replacing some convolutional layers of YOLOv5s with the lightweight network structures, FasterNet Block and RepVGG Block. This significantly reduced the model′s parameter count and the accelerated detection speed, meeting real-time requirements. The strong connectivity features in BiFPN enhanced the model′s feature learning capacity. An improved Cutmix data augmentation method was adopted, which randomly enhanced targets into input images and updated their labels, alleviating issues of label category imbalance and improving generalization. Experiments were carried on a self-built safety equipment wear detection dataset, the results showed that the mAP value of YOLOv5s-FRCB reached 96.3%, and the algorithm model memory was reduced by 34%, which was an efficient and practical method for safety equipment wear detection; YOLOv5s-FRCB achieved a further reduction in computational complexity while maintaining accuracy.

Keywords: safety equipment wearing detection; object detection; lightweight; data enhancement; real-time detection

0 引言

目标检测旨在从复杂背景中精确定位每个目标位置判断目标所属类别[1]。随着深度学习算法飞速发展,目标检测技术广泛应用于各类工业现场与作业场景。在高危易爆事故作业环境中,如电力设施检修、建筑施工场所及货物装卸作业场

收稿日期:2023-09-08
基金项目:国家自然科学基金资助项目(62002131)
第一作者简介:张曼(1997—),女,贵州安顺人,硕士研究生,主要研究方向为计算机应用。E-mail: zhiman_zm@163.com
*通信作者简介:孙凯军(1981—),男,江苏新沂人,讲师,硕士研究生,主要研究方向为计算机应用。E-mail: sunkaijun_hyit@hotmail.com

景,作业人员需严格遵守规章制度佩戴安全绳、安全帽等安全设备。这些防护设备的佩戴情况通常依靠作业人员自觉性和现场管理人员监督检查。如果作业人员未能正确或及时佩戴安全带、安全帽等必要防护装置,极可能导致不同严重程度的工伤事故或危险事件。开展面向作业人员的安全防护装置佩戴检测研究,具有重要理论价值和实践意义。这种检测技术能够实时监测作业人员防护措施落实情况,有效减少工业事故发生概率。

传统目标检测算法通常依靠手工特征提取方法,结合机器学习算法进行分类[2]。这种方法容易受到天气变化、遮挡和光照强度变化等因素影响,导致检测结果鲁棒性不够。与传统基于手工特征方法相比,深度学习目标检测方法可以同时学习低级和高级图像特征,这种方法已应用于实际工业场景中,可以及时对作业人员安全防护进行检测和监控[3-7]。现有成果中,一些深度检测模型在安全设备佩戴检测方面取得较好效果,但仍存在模型复杂度大、对小目标检测效果较差以及鲁棒性不足的问题。

为解决上述问题,现有成果从以下3个方面进行改进:(1)基于深度学习的目标检测轻量化;(2)基于深度学习的多尺度特征融合;(3)基于深度学习的数据增强。

许多研究者致力于开发轻量级目标检测模型,旨在保持较高检测性能,同时减少模型计算复杂性和内存需求。文献[8]改善两阶段目标检测算法检测速度慢的弊端,将原始YOLO主干网络改为移动网络MobileNet,删除网络最后的全连接层和SoftMax层。上述改进解决了以往任意卷积核都要对所有通道进行卷积操作的弊端,显著提高了网络训练效率。文献[9]减少参数量,将GhostNet与基于CNN的HIS分类器相结合,降低计算成本,实现了一种高效、高性能分类方法,在保证算法性能的同时降低了计算成本。文献[10]将原始网络架构中单一卷积核转变为深度可分离卷积,并利用跳跃式连接以便提取的特征可以跳过部分网络而传递到更深网络层。为进一步提高YOLO算法检测精度,文献[11]将动态锚框机制引入原始检测网络以提高网络检测精度。文献[12]引入DenseNet,增加Bottleneck结构数量,压缩了模型参数量,提高检测速度。但可变性卷积操作复杂,且算法检测精度未得到有效提升。

特征融合是一种有效提高目标检测性能的方法,它通过整合多尺度、多层次特征来捕捉目标多样性。近年来,提出了许多特征融合网络,这些方法通过构建金字塔形特征层次结构,将高层次特征与低层次特征进行融合,实现对不同尺度目标精确检测。文献[13]提出了一种新原型对准网络PANet,以更好利用支持集信息;模型精度得到有效提升。文献[14]提出了一种基于NAS-FPN和Faster R-CNN的深度卷积神经网络,用于花卉目标检测、定位和分类;在多个基准数据集上取得良好结果。文献[15]提出了一个带有双向特征金字塔网络多模态融合网络,使用一个单独的编码器来提取四种模态特征,以专注于利用模态之间的复杂关系,在肿瘤分割上取得鲁棒结果。文献[16]为实现跨层次信息融合及提高小目标关注度,分别提出了多层次信息聚合特征金字塔网络结构和基于多尺度通道注意力机制特征融合模块,有效提高了小目标检测效果。

为提高目标检测模型泛化能力和鲁棒性,研究者们引入了各种数据增强技术,包括数据增广(如随机裁剪、翻转、变换等)、对抗性训练、低照度增强等。这些方法通过改变图像外观和内容,增加模型在训练过程中遇到的样本多样性,提高模型在不同场景和条件下的检测性能。文献[17]提出了一种无监督的基于先验领域对抗性对象检测框架,用于在朦胧和下雨条件下进行检测。文献[18]提出了一种在线复制粘贴和混合数据增强方法,以缓解训练期间数据集类别不平衡,有效提高YOLOv5检测模型的分类性能。文献[19]提出一种新的图像自适应IA-YOLO框架,每个图像都可以自适应地增强,获得更好检测性能。文献[20]提出了一种新的技术增强技术Cutmix,通过鼓励网络学习更加鲁棒和可定位特征,提高深度神经网络性能,结果表明,使用Cutmix作为数据增强技术可以提高准确性,适用于多个基准数据集。

本研究在现有工作基础上,针对安全设备佩戴检测中的几个关键问题进行了进一步的改进和优化:(1)开发适于移动终端部署的实时检测模型;(2)提高对小目标如安全绳的检测精度,减少漏检和误检;(3)提升模型在处理遮挡情况下的准确性。

基于这些改进,本研究提出了一种融合FasterNet[21]和RepVGG[22]的安全设备佩戴检测方法YOLOv5s-FRCB。

1 YOLOv5s-FRCB

本研究主要以YOLOv5s网络为基线进行改

进,通过引入 FasterNet 和 RepVGG 结构融入 YOLOv5s 骨干层,进行轻量化改进,降低计算量;将颈部原 Concat 层替换为 BiFPN 特征融合算法,提升对低像素目标的检测效果;通过改进 Cutmix 数据增强方法,提升模型对遮挡的鲁棒性。网络架构如图 1 所示。

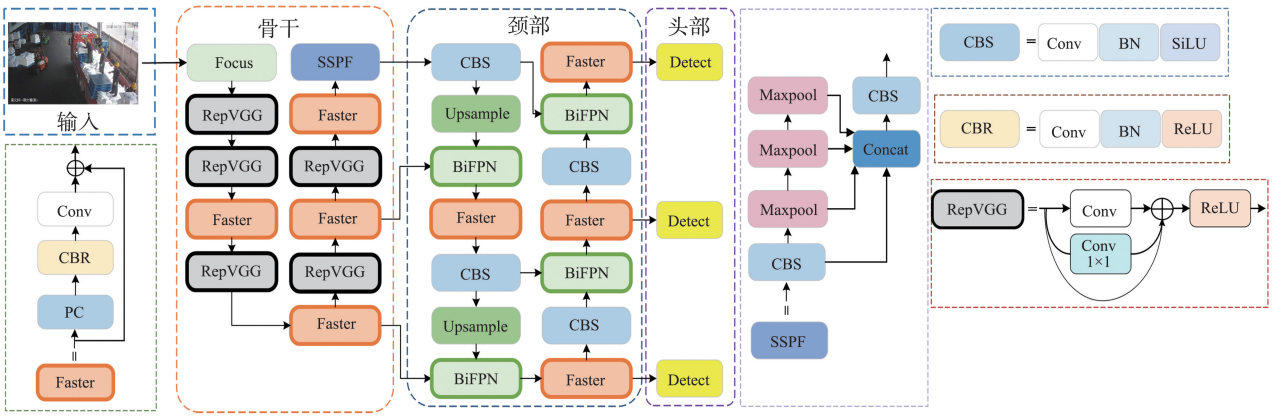


图 1　YOLOv5s-FRCB 网络框架

Fig.1　Network framework of YOLOv5s-FRCB

1.1　YOLOv5 目标检测神经网络

YOLOv5 实现了网络结构系列化,包含 YOLOv5s、YOLOv5m、YOLOv5l、YOLOv5x 4 种网络深度不同的网络结构,YOLOv5s 网络深度和特征图宽度最小,其他 3 个网络深度在其基础上加深加宽。在相同数据集下,YOLOv5s 模型训练时间最短、体积最小且实时高效。本研究在 YOLOv5s 上进行改进,YOLOv5s 模型结构由输入、骨干、颈部和头部 4 个部分组成,结构如图 2 所示。

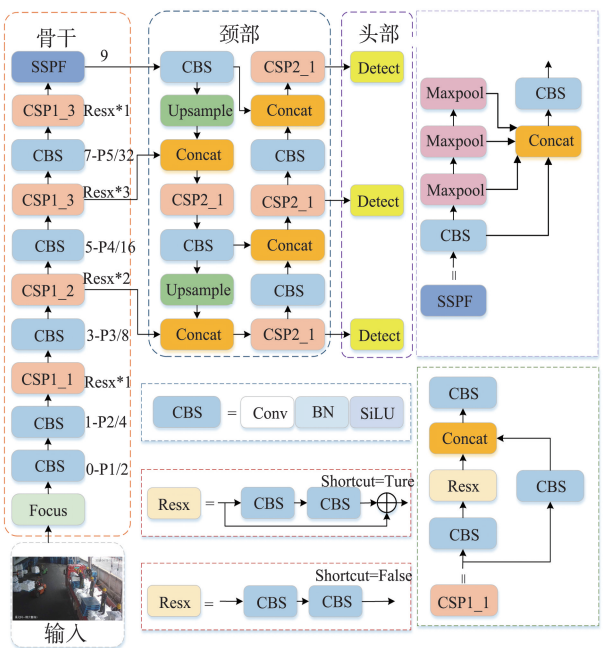


图 2　YOLOv5s-FRCB 网络框架

Fig.2　Network framework of YOLOv5s-FRCB

在输入阶段,通过加入数据增强方式扩展数据集。在骨干阶段,通过 Focus 结构把尺度为 608 像素×608 像素×3 的样本切片拼接成 304 像素×304 像素×12,再经过卷积运算转换成 304 像素×304 像素×32 的特征图。在颈部阶段,使用特征金字塔(feature pyramid network,FPN)和路径聚合网络(pixel aggregation network,PAN)结构,完成不同层的特征融合。

1.2　融合 FasterNet 模块

部分卷积(partial convolution,PConv)是一种新的卷积操作,它利用特征图冗余,只在部分输入通道上应用常规卷积,而其他通道不变。PConv 计算量低于常规卷积,但高于深度可分离卷积和组卷积。PConv 可以充分利用计算资源,在提取空间特征方面也很有效。常规卷积的浮点运算次数(floating point of per second,FLOPs)F_L 为 $h\times w\times k^2\times c^2/s$,PConv 的 FLOPs 为:

$$F_L=h\times w\times k^2\times c_p^2, \tag{1}$$

式中:h、w 分别为输入特征图的高和宽,单位为像素;k 表示卷积核大小;c 表示输入特征图的通道数。对于典型的缩减因子 $r=c_p/c=1/4$,PConv 的 FLOPs 只有常规 Conv 的 1/16。

为了充分有效利用来自所有通道信息,将逐点卷积(pointwise convolution,PWConv)附加到 PConv。如图 3 所示,在 PConv 层后,是 2 个 PWConv层,总体显示为倒置残差快,其中中间层具有扩展的通道数量。YOLOv5s-FRCB 将 FasterNet Block 融入骨干和颈部,旨在提升模型检测速度。

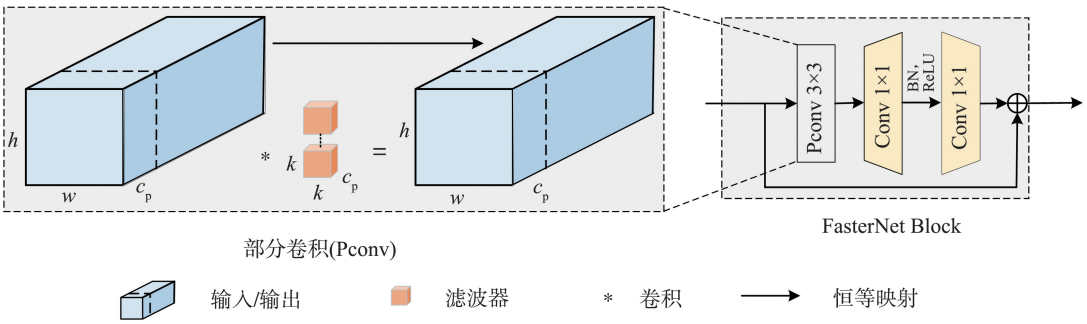


图 3 FasterNet Block 模块
Fig.3 FasterNet Block module

1.3 结构重参数化模块

RepVGG 是 3×3 卷积和 ReLU 组成的简单体系结构,适用于 GPU 和专用推理芯片。如图 4 所示,RepVGG 受 ResNet 启发,应用了结构重参数化方法。该方法在训练期间将一个模块拆分为多个相同或不同模块分支,在推理期间将多个分支模块集成为完全等效的模块。该模块在硬件上能够高效推理,同时保持较好的多尺度特征融合能力。

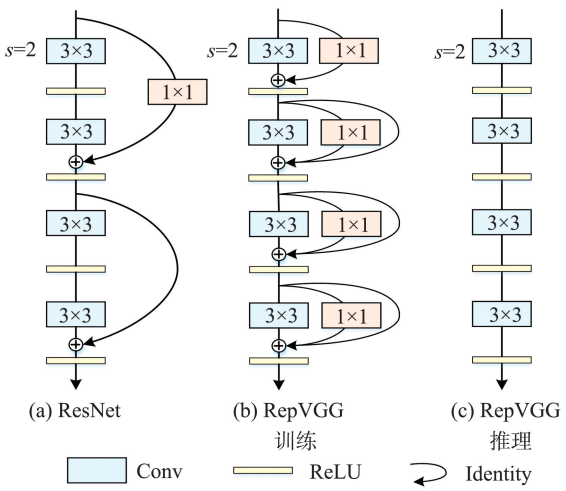


图 4 RepVGG 模块
Fig.4 RepVGG module

ResNet 是一种深度残差网络,通过显式构造快捷分支来建模信息流。这种方法将信息流表示为:

$$y=x+f(x), \tag{2}$$

式中:y 为网络输出,x 为恒等映射,$f(x)$是残差学习部分。当快捷分支 x 与 $f(x)$ 的通道数不匹配时,使用一个 1×1 的卷积 $g(x)$来匹配通道数,将信息流表示为:

$$y=g(x)+f(x)。 \tag{3}$$

RepVGG 是一种借鉴 ResNet 的方法,使用类似的残差学习方式建模信息流。RepVGG 将信息流建模为:

$$y=x+g(x)+f(x), \tag{4}$$

式中:x 是恒等映射,$g(x)$ 是用于匹配通道数的卷积,$f(x)$是残差学习部分。

YOLOv5-FRCB 借鉴 RepVGG 重参数化思想,将 YOLOv5s 结构中原有的 3×3 卷积替换成 RepVGG。

1.4 加权双向特征金子塔网络 BiFPN 引入

YOLOv5 特征融合网络采用 FPN+PAN 网络结构。底层特征图具有强定位特征信息[23],语义特征信息相对较弱,PAN 网络是自底向上融合,可传达强定位特征信息。FPN 网络自顶向下融合,将高层语义特征信息通过上采样方式进行传递融合,传达强语义特征信息。不同输入特征信息具有不同分辨率,它们对融合的输出特征信息做出不同贡献,导致重要特征会被忽略,结构如图 5 所示。

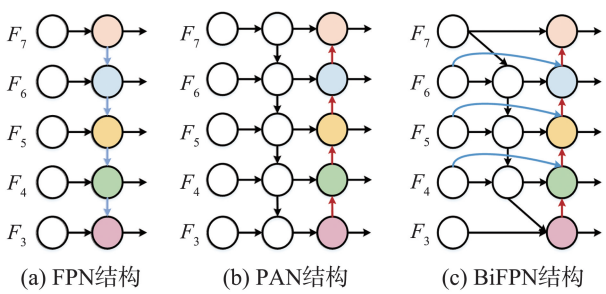


图 5 FPN、PAN、BiFPN 网络结构图
Fig.5 Network structure diagram of FPN, PAN and BiFPN

研究引入 BiFPN 加权双向特征金字塔网络学习不同输入特征信息的重要性,对重要特征信息进行融合。BiFPN 计算公式为:

$$F_6^{\text{td}}=\text{Conv}\left(\frac{w_1\cdot F_6^{\text{in}}+w_2\cdot \text{Resize}(F_7^{\text{in}})}{w_1+w_2+\varepsilon}\right), \tag{5}$$

$$F_6^{\text{out}}=\text{Conv}\left(\frac{w_1'\cdot F_6^{\text{in}}+w_2'\cdot F_6^{\text{td}}+w_3'\cdot \text{Resize}(F_5^{\text{out}})}{w_1'+w_2'+w_3'+\varepsilon}\right), \tag{6}$$

式中:Conv 为卷积计算;Resize 用于匹配分辨率的采样操作;F_6^{in} 为自顶向下路径上第六级输入特征;F_6^{td}为自顶向下路径上第六级中间特征;F_6^{out} 为自底向上路径上第 6 层输出特征;w_i为每条路径上的权重;ε 为防止数值不稳定设置的较小学习率。

1.5 Cutmix 方法改进

Cutmix 是一种用于图像分类任务的数据增强

方法,通过将训练图像中的一部分区域剪切并粘贴到其他图像中,生成新的训练样本,有效提高模型泛化能力和鲁棒性,同时也能防止过拟合。本研究在 Cutmix 的思想上进行改进,提出一种新方法 S-Cutmix(sampling cutmix)。预先从数据集中提取目标图像,将目标图像作为嵌入图像嵌入到训练图像中,解决图像标签类别数量均衡问题。利用标签文件将图像中的目标区域剪切出来,作为嵌入图像样本集,随机嵌入原图像中,设置不可与原图像中标签区域重合,避免遮挡对精度产生影响。

S-Cutmix 具体数据增强流程如图 6 所示:从数据集中提取目标图像,作为嵌入图像样本集;随机从数据集选择训练图像和对应标签;在训练图像中随机采样矩形区域;用嵌入图像替换训练图像中的采样区域,形成新的混合图像;对应更新标签中的采样区域标签,形成混合标签;将混合图片和混合标签作为模型输入,用于训练。



(a) 嵌入图像样本集　(b) 训练图像　(c) 混合图像

图 6　数据增强示意图

Fig.6　Schematic diagram of data enhancemen

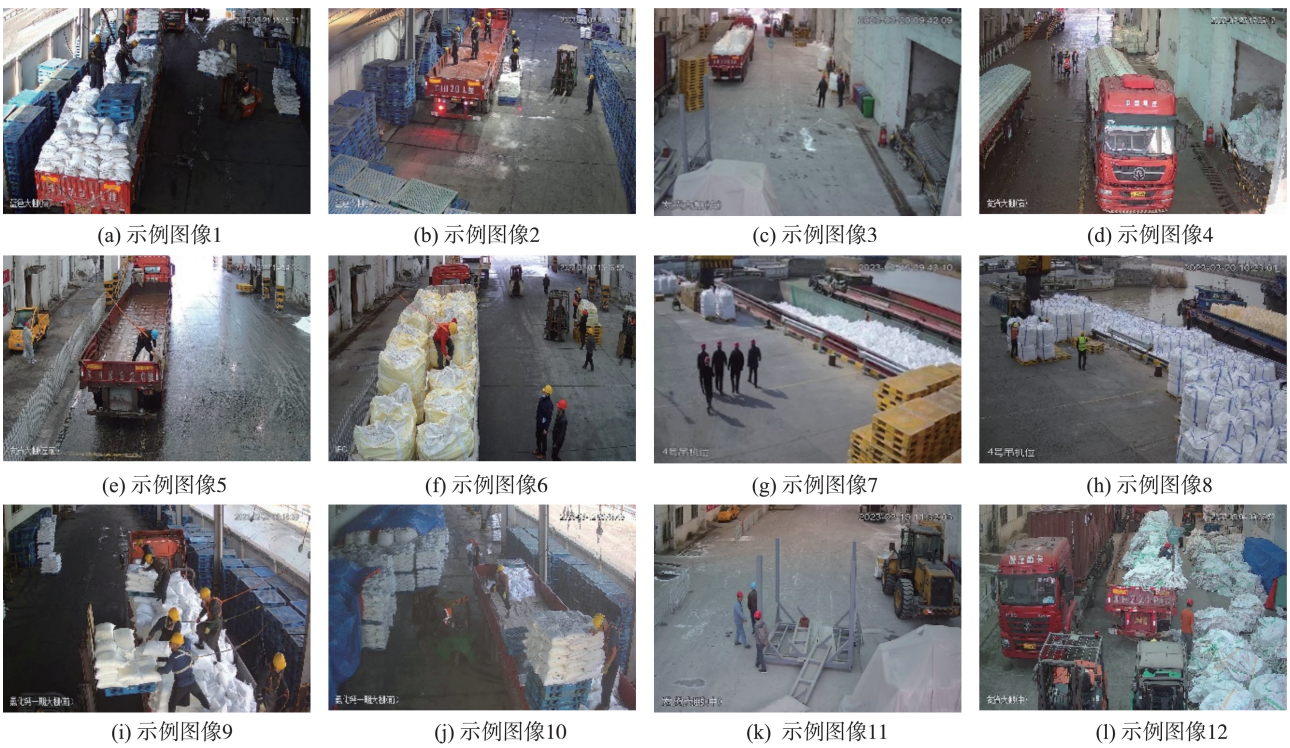
在训练过程中,模型在混合图像上同时学习来自训练图像和嵌入图像的特征,这使模型可以学习到更加泛化的特征表示,提高模型对目标的学习能力。训练数据中将两个图像信息同时输入到模型,这也可以防止模型过拟合于单一类别。

2　试验设计与结果分析

为评估模型有效性,在自制数据集 SEWD(safety equipment wear dataset, SEWD)上,针对安全帽和安全绳检测任务,对本研究进行验证,针对各模块进行消融试验分析。

2.1　数据集

试验数据集采用自制数据集 SEWD,公开于 https://github.com/zmwblzzc/SEWQ。数据集包含人、头、安全绳、安全帽、卡车 5 个类别。数据集采集自不同场景、不同天气状况下的实时监控,共 4 335 张图像,40 443 个标签,其中 3 900 张图像为训练集,455 张图像为测试集。数据集部分图像如图 7 所示,原始图像标签如图 8 所示。



(a) 示例图像1　(b) 示例图像2　(c) 示例图像3　(d) 示例图像4

(e) 示例图像5　(f) 示例图像6　(g) 示例图像7　(h) 示例图像8

(i) 示例图像9　(j) 示例图像10　(k) 示例图像11　(l) 示例图像12



(m) 示例图像13 (n) 示例图像14 (o) 示例图像15 (p) 示例图像16

图 7 数据集图像示例

Fig.7 Sample images from the dataset



图 8 原始图像标签示例

Fig.8 Sample of original image labels

2.2 试验环境

本试验在 Ubuntu 18.04 操作系统上,使用 AMD Ryzen 75800H 3.20 GHz CPU 和 NVIDIA GeForce RTX 3090 6 GB GPU 进行训练,系统内存 16 GB。采用 Python 3.7编程语言,GPU 加速使用CUDA11.7 和 CUDNNv8.2.1,基于Pytorch1.8.1框架进行深度学习。训练参数设置如下:输入图像大小为640 像素×640 像素,训练周期(Epochs)为 300,初始学习率为 0.01,批处理量(Batchsize)为 32。

2.3 模型算法评估指标

本研究从精确率 P、召回率 R、平均精度均值 M_{AP}、参数量 P_{AR}、模型大小 M_O及每秒所执行的浮点运算次数 F_L对模型准确性进行评估。

P 反映模型正确检测出目标的概率:

$$P=\frac{T_P}{T_P+F_P}。\tag{7}$$

式中:T_P 为模型正确地将正样本预测为正样本的情况数,F_P 为模型错误地将负样本预测为正样本的情况数,F_N 为模型错误地将正样本预测为负样本的情况数。

R 反映模型检测出全部目标的概率:

$$R=\frac{T_P}{T_P+F_N}。\tag{8}$$

A_P是模型在单个检测对象上的精度:

$$A_P=\int_0^1 P(R)\,\mathrm{d}R。\tag{9}$$

M_{AP}在目标检测任务中指计算所有类别的 A_P求平均,可以衡量模型在所有类别上的检测能力:

$$M_{AP}=\sum_{n=1}^{N}A_{PN}\Big/N,\tag{10}$$

式中:N 表示目标检测类别,A_{PN} 表示第 N 类的 A_P值。

P_{AR}指神经网络中可学习的参数总数,包含所有层的权重和偏置。M_O 是整个模型在存储中的占用空间。

2.4 试验结果与分析

本研究从参数量对比、数据增强对比、消融试验以及与其他算法对比进行试验说明 YOLOv5s-FRCB 算法的先进性及优越性,通过可视化结果直观展示改进所取得成效。

2.4.1 参数量对比

将 FasterNet Block(F)和 RepVGG Block(R)两个模块融合于 YOLOv5 骨干网络,减少参数量,优化模型,达到提速目的。由表 1 可以看出,加入 FasterNetBlock 后 P_{AR}减少 1.23×10^6,同时 F_L 有所降低;加入 RepVGG Block 模块,P_{AR} 也减少 0.86×10^6,但 F_L 并未减低;同时融合 2 个模块,P_{AR} 和 F_L 都维持在较低的水平。本研究在保持检测精度的同时能够使网络更加轻量化。

表 1 参数量及 FLOPs 对比

Table 1 Comparison of the number of parameters and FLOPs

模型	$P_{AR}/10^6$	$F_L/10^9$
YOLOv5s	7.02	15.8
YOLOv5s+F	5.79	12.7
YOLOv5s+R	6.16	14.4
YOLOv5s+FR	4.52	12.0

2.4.2 数据增强对比试验

表 2 为 YOLOv5s 模型训练结果,分析表 2 所有训练类别 P、R、M_{AP}可以看出:安全绳类别在 3 种评价指标下都是最低;而从标签数来看,安全绳类别标签数量除车以外最少,只有 6 502 个标签,类别分布不平衡。因此针对安全绳类别进行数据增强,增加安全绳类别标签数量,其余类别数量保持不变。在原训练集每一张图像上都随机嵌入一张安全绳图像,将安全绳标签数从 6 502 增加至 10 403 个标签,进行训练。

进行数据增强后,模型训练结果如表 3 所示,从

表 2 和表 3 可以看出,使用 S-Cutmix 方法,可以有效解决标签类别不均衡问题,增强后训练结果在 P、R、M_{AP}上均有提高,模型 M_{AP}由增强前95.5%提高至96.0%。从图 9(a)和(b)来看,使用本研究数据增强方法,每个类别误检率均有所减低。

表 2　数据增强前各类别检测结果

Table 2　Detection results for each category before data augmentation

类别	标签数/个	P/%	R/%	M_{AP}/%
人	8 519	93.5	95.6	97.4
车	5 420	98.5	98.8	99.0
安全帽	10 190	96.2	91.9	95.6
安全绳	6 502	92.2	84.6	89.3
头	9 806	94.2	94.2	96.3
平均	8 088	94.9	93.0	95.5

表 3　数据增强后各类别检测结果

Table 3　Detection results for each category after data augmentation

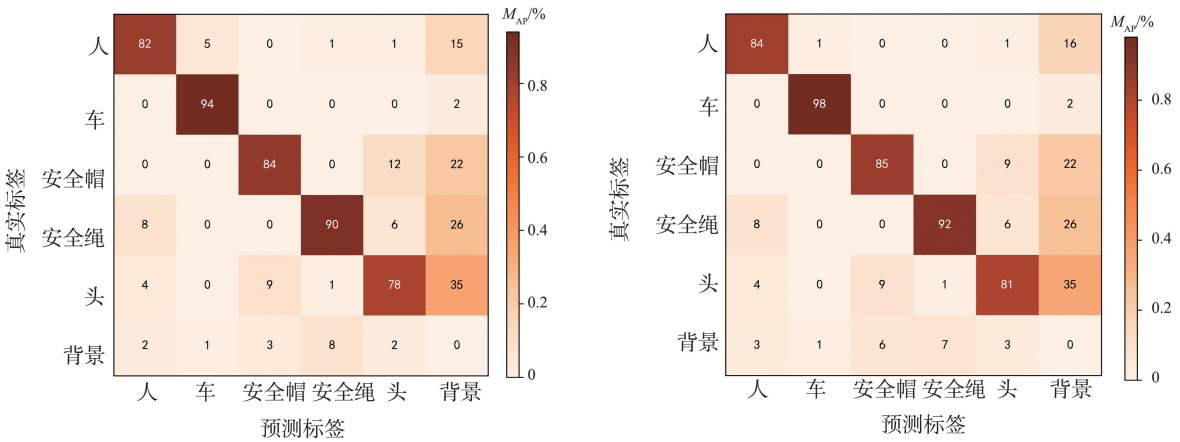
类别	标签数/个	P/%	R/%	M_{AP}/%
人	8 519	93.8	97.0	98.4
车	5 420	98.8	99.4	99.1
安全帽	10 190	96.5	93.2	96.5
安全绳	10 403	90.7	84.3	88.8
头	9 806	95.7	94.1	97.3
平均	8 953	**95.1**	**93.6**	**96.0**

2.4.3　消融试验

对试验结果进行消融试验比较,从表 4 中可以看出:相比于原始 YOLOv5s 算法,引入 FasterNet Block 和 RepVGG Block 均可以降低模型参数量,同时引入 2 个模块 P_{AR}降低至 4.52×10^6,对模型检测精度有影响,减低了 0.2%;引入 S-Cuxmix(C)对检测精度提高最为明显,M_{AP}提高了 1.2%;引入 BiFPN(B)精度提高 0.1%,对 P_{AR}有较小的提高;YOLOv5s-FRCB 算法较原始 YOLOv5s 在应用数据集上精度提高了 0.6%,F_L 降低了 3.8×10^9,模型体积缩小了 5 MB,能够在减低参数量的同时,保持较高检测精度。

2.4.4　YOLOv5-FRCB 与其他算法对比

将 YOLOv5s-FRCB 与其他 4 个目标检测经典模型训练结果进行比较,如表 5 所示。从表 5 可知:YOLOv5s-FRCB 模型的 P、R、M_{AP}明显优于其他参比模型。YOLOv5s-FRCB 模型的 P_{AR}、F_L 和 M_O 均优于其他 4 个模型,具有更轻量化、高效的特点,能够在计算资源有限情况下实现更快速、更准确的目标检测任务。



(a) 数据增强前混淆矩阵　　(b) 数据增强后混淆矩阵

图 9　数据增强前后混淆矩阵

Fig.9　Confusion matrix before and after data augmentation

表 4　消融试验结果

Table 4　Ablation experiment results

算法模型	FasterNet	RepVGG	Cutmix	BiFPN	M_{AP}/%	$P_{AR}/10^6$	$F_L/10^9$	M_O/ MB
YOLOv5s					95.7	7.02	15.8	14.5
YOLOv5s-F	√				95.8	5.79	12.7	12.0
YOLOv5s-R		√			95.6	6.16	14.2	12.7
YOLOv5s-C			√		**96.9**	7.02	15.8	14.5
YOLOv5s-B				√	95.8	7.09	16.0	14.6
YOLOv5s-FR	√	√			95.5	**4.52**	12.0	**9.5**
YOLOv5s-FRC	√	√	√		96.0	**4.52**	**12.0**	**9.5**
YOLOv5s-FRCB	√	√	√	√	96.3	4.55	12.1	**9.5**

注:“√”表示引入该方法,黑体加粗数字表示最优结果。

表 5 YOLOv5s-FRCB 与其他经典模型对比

Table 5 Comparison of YOLOv5s-FRCB with other classical models

算法模型	P/%	R/%	M_{AP}/%	$P_{AR}/10^6$	$F_L/10^9$	M_O/MB
SSD	80.3	78.1	80.9	13.76	15.4	110.1
Faster RCNN	88.7	82.4	89.5	41.38	134.5	659.3
Deformable DETR	90.3	87.5	91.4	38.30	122.2	389.4
YOLOv7-tiny	95.6	93.1	95.1	6.02	13.2	12.0
YOLOv5s-FRCB	**96.8**	**95.9**	**96.3**	**4.55**	**12.1**	**9.5**

注:黑体加粗数字表示最优结果。

将 YOLOv5s-FRCB 与其他轻量级网络改进的模型进行比较,其中参比的 3 种方法 MobileNetV3、ShuffleNetV2、GhostNet 表示将 YOLOv5s 主干网络分别替换为 MobileNetV3、ShuffleNetV2、GhostNet 网络。以此检测本算法的性能以及优势,试验结果如表 6 及图 10 所示。

表 6 YOLOv5-FRCB 与其他轻量化主干算法对比

Table 6 Comparison of YOLOv5-FRCB with other lightweight backbone algorithms

算法模型	P/%	R/%	M_{AP}/%	$P_{AR}/10^6$	$F_L/10^9$	M_O/MB
MobileNetV3	93.8	93.3	94.0	5.03	11.3	10.0
ShuffleNetV2	93.4	92.9	93.6	3.80	**8.0**	7.7
GhostNet	94.0	94.3	94.6	**3.69**	8.1	**7.5**
YOLOv5s	96.0	95.4	95.7	7.02	15.8	14.5
YOLOv5s-FRCB	**96.8**	**95.9**	**96.3**	4.55	12.1	9.5

注:黑体加粗数字表示最优结果。

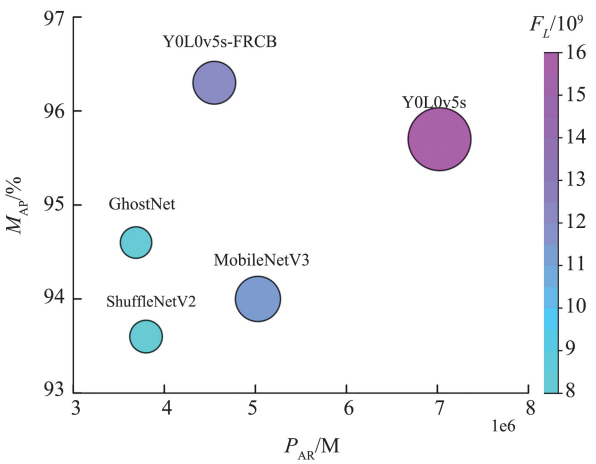


图 10 模型性能对比散点图

Fig.10 Scatter plot comparing models

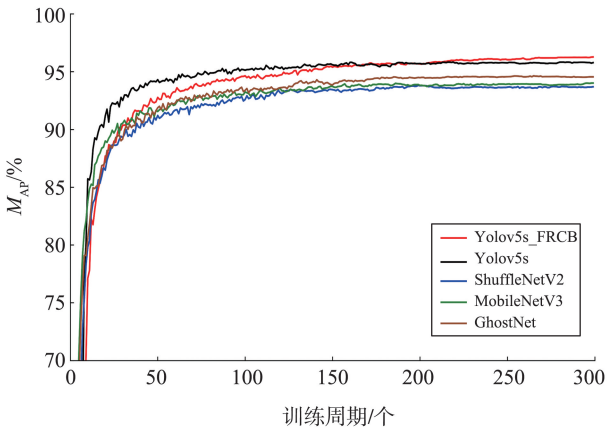


图 11 YOLOv5s-FRCB 与其他模型训练对比

Fig.11 Training comparison of YOLOv5s-FRCB with other models

与 YOLOv5s 相比,YOLOv5s-FRCB 的 P_{AR} 大幅下降,F_L 下降,模型精度提升,说明 YOLOv5s-FRCB 的拓展结构有效地减小了模型复杂度,还提高了检测精度。与其他轻量级模型(MobileNetV3、ShuffleNetV2、GhostNet)相比较,YOLOv5s-FRCB 在 P、R、M_{AP}、P_{AR}、M_O 上都优于 MobileNetV3,精度高于 ShuffleNetV2 和 GhostNet。从图 11 可以看出,虽然轻量化改进对模型收敛速度造成一定影响,使得各轻量化改进模型收敛速度都慢于 YOLOv5s,YOLOv5s-FRCB 较其他模型有明显的精度优势。

2.4.5 可视化分析

图 12~15 展示了在不同场景下安全帽和安全绳的检测结果,红圈区域为漏检目标。具体来说,图 12 和图 13 分别为采用原始 YOLOv5 网络和改进后的 YOLOv5s-FRCB 网络检测安全帽结果。与 YOLOv5s 对比,YOLOv5s-FRCB 网络在检测准确度方面表现更优,并且降低了漏检情况。图 14 和图 15 分别展示了原始 YOLOv5 网络和 YOLOv5s-FRCB 网络在检测安全绳任务上的表现。YOLOv5s-FRCB 网络在提高检测精度的同时,成功解决了大量漏检问题。



(a) 安全帽检测结果1　(b) 安全帽检测结果2　(c) 安全帽检测结果3　(d) 安全帽检测结果4

图 12　YOLOv5s 安全帽检测结果可视化

Fig.12　Visualization of YOLOv5s helmet detection results



(a) 安全帽检测结果1　(b) 安全帽检测结果2　(c) 安全帽检测结果3　(d) 安全帽检测结果4

图 13　YOLOv5s-FRCB 安全帽检测结果可视化

Fig.13　Visualization of YOLOv5s-FRCB helmet detection results



(a) 安全绳检测结果1　(b) 安全绳检测结果2　(c) 安全绳检测结果3　(d) 安全绳检测结果4

图 14　YOLOv5s 安全绳检测结果可视化

Fig.14　Visualization of YOLOv5s safety rope detection results



(a) 安全绳检测结果1　(b) 安全绳检测结果2　(c) 安全绳检测结果3　(d) 安全绳检测结果4

图 15　YOLOv5s-FRCB 安全绳检测结果可视化

Fig.15　Visualization of YOLOv5s-FRCB safety rope detection results

3　结束语

针对实际工业环境中目标检测算法模型规模较大且实时性较差的问题,本研究提出了一种改进的 YOLOv5s-FRCB 算法。通过引入轻量化模块 FasterNet Block 和 RepVGG Block,成功降低了模型参数量并显著提高了检测速度。同时,通过 S-Cutmix 数据增强方法和引入 BiFPN,进一步提高了模型检测精度。

试验结果表明,在自制安全设备数据集 SEWD 上,改进后 YOLOv5s-FRCB 算法精度得到有效提升,且参数量减少。这意味着在保持高度轻量化的同时,算法明显提高了检测精度。与其他主流轻量化检测算法相比,本研究提出的算法在 FLOPs 和参数数量方面具有明显优势,保持了较高检测精度。

参考文献:

[1] DHILLON A, VERMA G K. Convolutional neural network: a review of models, methodologies and applications to object detection[J]. Progress in Artificial Intelligence, 2020, 9(2):85-112.

[2] ZHAO Z Q, ZHENG P, XU S, et al. Object detection with deep learning: a review[J]. IEEE transactions on neural networks and learning systems, 2019, 30(11):

3212-3232.

[3] LI Z, XIE W, ZHAGN L, et al. Toward efficient safety helmet detection based on YoloV5 with hierarchical positive sample selection and box density filtering[J]. IEEE Transactions on Instrumentation and Measurement, 2022, 71:1-14.

[4] FAROOQ M U, BHUTTO M A, KAZI A K. Real-time safety helmet detection using Yolov5 at construction sites [J]. Intelligent Automation & Soft Computing, 2023, 36 (1):911-927.

[5] SONG R, WANG Z. RBFPDet: an anchor-free helmet wearing detection method [J]. Applied Intelligence, 2023, 53(5): 5013-5028.

[6] 张锦,屈佩琪,孙程,等. 基于改进 YOLOv5 的安全帽佩戴检测算法[J]. 计算机应用,2022,42(4):1292-1300.
ZHANG Jin, QU Peiqi, SUN Cheng, et al. Helmet wearing detection algorithm based on improved YOLOv5[J]. Journal of Computer Applications, 2022, 42 (4): 1292-1300.

[7] 田坤,李冠,赵卫东. 基于 YOLO 和极限学习机的驾驶员安全带检测模型研究[J].计算机应用与软件,2019, 36(11):196-201.
TIAN Kun, LI Guan, ZHAO Weidong. Research on driver seat belt detection model based on YOLO and extreme learning machine[J]. Computer Applications and Software, 2019, 36(11):196-201.

[8] WAN L, LIU R, SUN L, et al. UAV swarmbased radar signal sorting via multi-source data fusion: a deep transfer learning framework[J]. Information Fusion, 2022, 78: 90-101.

[9] PAOLETTI M E, HAUT J M, PEREIRA N S, et al. Ghostnet for hyperspectral image classification[J]. IEEE Transactions on Geoscience and Remote Sensing, 2021, 59(12):10378-10393.

[10] HUANG T Y, LEE M C, YANG C H, et al. YOLO-ORE: adeep learning-aided object recognition approach for radar systems[J]. IEEE Transactions on Vehicular Technology, 2022, 72:5715-5731.

[11] FAN S, LIANG X, HUANG W, et al. Real-time defects detection for apple sorting using NIR cameras with pruning-based YOLOv4 network[J]. Computers and Electronics in Agriculture, 2022, 193:106715.

[12] GAI R, CHEN N, YUAN H. A detection algorithm for cherry fruits based on the improved YOLO-v4 model[J]. Neural Computing and Applications, 2023, 35 (19): 13895-13906.

[13] WANG K, LIEW J H, ZOU Y, et al. Panet: few-shot image semantic segmentation with prototype alignment [C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. Seoul, Korea:IEEE, 2019: 9197-9206.

[14] PATEL I, PATEL S. An optimized deep learning model for flower classification using NAS-FPN and faster R-CNN[J]. International Journal of Scientific & Technology Research, 2020, 9(3):5308-5318.

[15] SYAZWANY N S, NAM J H, LEE S C. MM-BiFPN: multi-modality fusion network with Bi-FPN for MRI brain tumor segmentation[J]. IEEE Access, 2021, 9: 160708-160720.

[16] 谢椿辉,吴金明,徐怀宇. 改进 YOLOv5 的无人机影像小目标检测算法[J].计算机工程与应用, 2023, 58 (9):1-11.
XIE Chunhui, WU Jinming, XU Huaiyu. An improved small target detection algorithm for UAV image based on YOLOv5 [J]. Computer Engineering and Applications, 2023, 58(9):1-11.

[17] SINDAGI V A, OZA P, YASARLA R, et al. Prior-based domain adaptive object detection for hazy and rainy conditions[C]// Computer Vision-ECCV 2020: 16th European Conference. Glasgow, UK: Springer, 2020:763-780.

[18] KIM J H, LIM N, PARK Y W, et al. Object detection and classification based on YOLO-V5 with improved maritime dataset[J]. Journal of Marine Science and Engineering, 2022, 10(3):377.

[19] LIU W, REN G, YU R, et al. Image-adaptive YOLO for object detection in adverse weather conditions[C]// Proceedings of the AAAI Conference on Artificial Intelligence. Palo Alto, USA: IEEE, 2022, 36 (2): 1792-1800.

[20] YUN S, HAN D, OH S J, et al. Cutmix: Regularization strategy to train strong classifiers with localizable features[C]// Proceedings of the IEEE/CVF International Conference on Computer Vision. Seoul, Korea: IEEE, 2019:6023-6032

[21] CHEN J, KAO S, HE H, et al. Run, don't walk: chasing higher FLOPS for faster neural networks[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Vancouver, Canada, IEEE, 2023:12021-12031.

[22] DONG X, ZHANG X, MA N, et al. Repvgg: Making vgg-style convnets great again[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville, USA: IEEE, 2021: 13733-13742.

[23] WAN G, FANG H, WANG D, et al. Ceramic tile surface defect detection based on deep learning [J]. Ceramics International, 2022, 48(8):11085-11093.

(编辑:陈燕)