

DOI:10.3969/j.issn.1003-5060.2025.09.005

基于改进 YOLOv5s 的烟丝制丝生产线 小目标杂物检测方法

郑银环, 陈恩杰, 吴飞, 张帅彬

(武汉理工大学 机电工程学院, 湖北 武汉 430070)

摘要:文章提出一种基于改进 YOLOv5s 的小目标检测算法,以 YOLOv5s 算法为基础模型,首先将 Focal Loss 和 EIoU Loss 引入模型优化原有的 BCE Loss 和 CIoU Loss,加快模型的收敛速度;其次添加一个目标检测头,提高对小目标杂物的检测精度;最后对比分析不同类型注意力模块对模型的影响,并将坐标注意力引入模型颈部,加强模型对目标关键特征的提取,提高模型的学习能力。基于自制的杂物数据集对模型进行训练,实验结果表明,相较 YOLOv5s 算法,改进后的模型在测试集上的精确率、召回率、平均精度均值(mean average precision, mAP)值分别提高 4.9%、5.5%、7.3%,识别效果更好,满足实际生产中精确性和实时性要求。

关键词:小目标检测; YOLOv5s 算法; 注意力机制; 检测头; 损失函数改进

中图分类号: TP391.4 **文献标志码:** A **文章编号:** 1003-5060(2025)09-1183-09

Small target sundries detection method for tobacco yarn production line based on improved YOLOv5s

ZHENG Yinhan, CHEN Enjie, WU Fei, ZHANG Shuaibin

(School of Mechanical and Electrical Engineering, Wuhan University of Technology, Wuhan 430070, China)

Abstract: A small target detection algorithm based on the improved YOLOv5s is proposed. Firstly, based on YOLOv5s, Focal Loss and EIoU Loss are introduced into the model to optimize the original BCE Loss and CIoU Loss to accelerate the convergence speed of the model. Secondly, a target detection head is added to improve the detection accuracy of small target sundries. Finally, the influence of different types of attention modules on the model is compared and analyzed, and coordinate attention is introduced into the model neck to strengthen the extraction of key features of the target and improve the learning ability of the model. The model is trained based on the self-made sundries dataset, and the experimental results show that compared with YOLOv5s, the accuracy of the improved model on the test set is improved by 4.9%, the recall rate increases by 5.5%, and the mean average precision (mAP) value increases by 7.3%. The model has better recognition effect, which meets the accuracy and real-time requirements in actual production.

Key words: small target detection; YOLOv5s algorithm; attention mechanism; detection head; loss function improvement

在烟丝制丝生产过程中,烟包的拆解是制丝生产的第 1 步,也是极其重要的一步,生产线上采

用工业机械臂自动拆解,容易将纸箱碎片、塑料袋、捆装绳等杂物混杂在烟丝中,且生产线时常需

收稿日期:2023-06-08;修回日期:2023-08-22

基金项目:国家自然科学基金资助项目(52275505)

作者简介:郑银环(1974—),女,湖北钟祥人,博士,武汉理工大学副教授,硕士生导师,通信作者,E-mail:douzyh@163.com;
吴飞(1973—),男,河南叶县人,博士,武汉理工大学教授,硕士生导师。

要维修,维修工人在修理后遗留下螺丝刀、钳子、扳手等工具杂物,没有及时发现并带走,会对后续生产造成很大影响。目前主要依靠人工作业实现杂物的筛检,该方式效率低,且存在杂物混在烟丝中难以被发现的问题,这些漏检的杂物会在后续的环节中被切片、破碎,其中软性杂物不仅会与烟丝混合,影响所制烟丝的质量,还容易引起堵塞;刚性杂物容易导致生产设备的损坏,给生产带来巨大的经济损失。因而为了提高烟丝质量,需要将目标检测技术应用于烟丝制丝生产线上。

随着机器学习、深度学习的发展,目标检测技术开始应用于工业现场监控,实现各种零部件及杂物的检测与识别。文献[1]为了防止电网工人在高空作业时因防具穿戴不当而导致安全事故发生,提出一种轻量级防具检测算法,替换传统的卷积神经网络,解决高空作业时的遮挡问题;文献[2]采用轻量级网络 GhostNet 替代 YOLOv4 的主干特征提取网络,并添加一个预测头,实现工业环境下的发动机金属表面缺陷检测。根据算法的实现原理不同,检测算法大致可分为如下 2 类:① 单阶段目标检测算法,该类算法不是首先提取候选框,而是直接提取网络特征,实现目标的分类与回归,因而检测速度快、实时性好,其代表算法有 SSD^[3]、YOLO 系列^[4]等;② 双阶段目标检测算法,该类算法首先通过候选框生成网络,然后通过分类网络对候选框的目标进行分类,其代表算法有 R-CNN^[5]、SPP-Net^[6]、Faster R-CNN^[7]等,该类算法精度高、效率低,不适应于小目标检测。考虑到制丝生产线传送带实时运行,且检测对象为小目标,本文优先考虑单阶段目标检测算法。

在不同场景下,对小目标的定义不尽相同,目前没有统一标准,常用的 2 种定义方式有基于相对尺度和基于绝对尺度的定义^[8]。基于相对尺度的定义方式是从目标与图像的相对比例角度对小目标进行定义。文献[9]采用相对尺度的定义方式对小目标进行定义,规定目标检测框面积占图像面积的比例不超过 0.58% 为小目标。基于绝对尺度的方式在定义时考虑目标绝对像素尺寸,如目前最常用的 MSCOCO^[10]数据集规定像素小于 32×32 的目标为小目标。

YOLOv5 算法能够很好地平衡检测精度和检测速度,相较其他单阶段检测算法,整体性能更优。文献[11]为实现复杂环境下的小尺度烟火检测,在 YOLOv5 算法上添加第 4 个检测层,增强

网络对小目标的特征提取能力,同时为了解决烟火因被遮挡而出现漏检的问题,对损失函数进行替换,有效提高对小目标烟火的检测精度;文献[12]从添加卷积 SE 注意力机制、简化 SPPF 模块和多尺度检测 3 个方面对 YOLOv5 进行改进,实现对建筑垃圾的准确分类和识别;文献[13]通过 K 均值聚类算法获取更精确的锚框,并融入小目标检测头和 SE 有效提高小目标的识别精度;文献[14]针对传统煤矿烟火检测存在时间长、误检率高问题,利用 GhostNet 重构 YOLOv5,使得模型更加轻量化,并提出全局上下文特征提取模块 BoT3,帮助模型更好地预测小目标,有效提高煤矿烟火检测的精度;文献[15]为提高雾霾、雨雪等恶劣天气下的小型交通标志的识别率,采用逆向思维、削减特征金字塔深度,并结合 K 均值先验框、全局非极大值抑制等机制对 YOLOv5 进行改进。

本文以 YOLOv5s 算法为基础,构建一种基于改进 YOLOv5s 的小目标检测方法。首先采用焦点损失^[16]可以平衡难易样本、着重训练难分类样本的特点,并结合 EIoU 可以分开考量长、宽对预测框影响的特点,改进原有的 BCE Loss,提高网络的收敛速度;其次针对其在小目标检测上检测能力不足的问题,添加一个检测头,实现对小目标的精确检测;最后对比分析 SE 网络^[17]、空间变换网络(spatial transformer networks, STN)^[18]、卷积注意力模块(convolutional block attention module, CBAM)^[19]、坐标注意力(coordinate attention, CA)^[20]4 种注意力机制的优缺点,综合考量选取最合适的 CA 注意力机制,并将其引入 YOLOv5s 的颈部中,以提高网络对目标特征的学习能力。

1 算法原理

1.1 YOLOv5s 算法原理

作为一种典型的单阶段目标检测算法,YOLOv5s 被广泛应用于目标检测场景,其在上一版本的基础上添加自适应图片缩放,使得检测速度、检测精度均有所提高^[21]。根据网络深度、模型复杂度的不同可以细分为 YOLOv5s、YOLOv5x 等 4 个版本,其中 YOLOv5s 是体积最小的模型,因而其检测速度最快、实时性最好,对运算能力的要求最低,其他几个模型相当于在 YOLOv5s 的基础上按照一定比例拓宽了深度和宽度。YOLOv5s 的网络结构如图 1 所示。

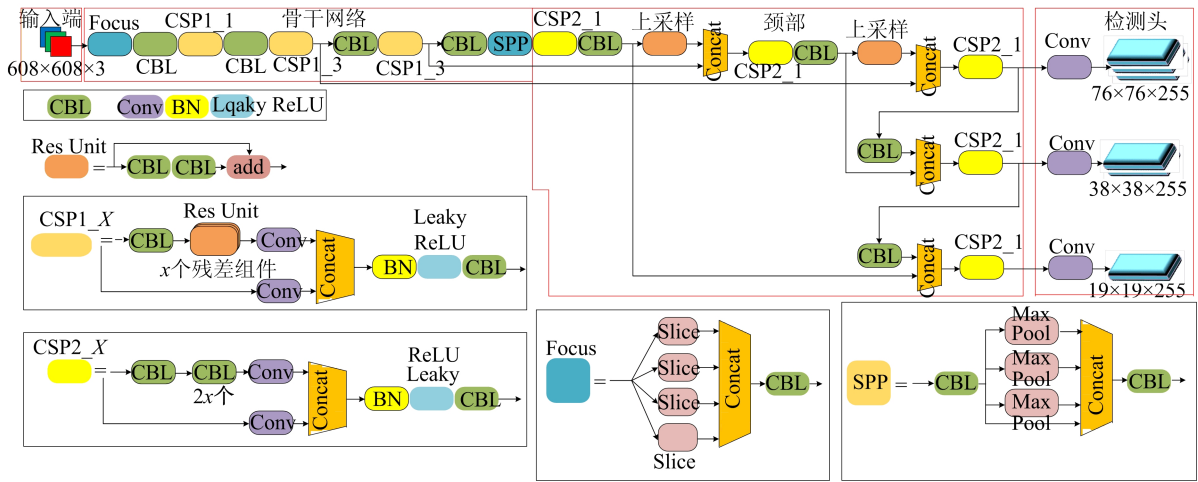


图 1 YOLOv5s 网络结构

从图 1 可以看出, YOLOv5s 的网络结构由输入端、骨干网络(Backbone)、颈部(Neck)、检测头(Detection Head)4 个部分组成,以 Leaky ReLU 作为激活函数,通过 Backbone 进行下采样提取图像特征,并通过 Neck 上采样来更好地融合特征,最终通过 Head 输出预测结果。

输入端采用马赛克数据增强,选取 4 张图片随机拼接,不但丰富了数据集、增加了小目标对象的数量,还使网络鲁棒性更好,有利于小目标对象的检测,且同时计算 4 张图片的数据,能有效提高训练速度,降低对计算机算力的需求。

Backbone 主要由 Focus 模块、瓶颈模块(cross stage partial, CSP) 2 个部分组成, Focus 主要用于切片操作,可以将每个通道的像素每隔一个取出,使得通道数目变为原来的 4 倍,如将一张 $640 \times 640 \times 3$ 的图像输入 Focus 结构,输出图像尺寸为 $320 \times 320 \times 12$ 。CSP 模块由应用于新版 YOLOv5 中的 C3 模块、用于旧版 YOLOv5 中的 BottleneckCSP 模块 2 个部分组成,两者大致相同,区别是 C3 模块去掉了残差模块后的卷积,并将激活函数换成了 SiLU 函数,如图 2 所示。

Neck 结构由双向特征金字塔^[22](FPN + PAN)组成, FPN 通过增强不同特征层进行特征融合,实现网络在多尺度上进行预测; PAN 在 FPN 的基础上添加从下至上的融合,将较浅层的定位信息传递到较深层,增强网络在多个尺度上的定位能力。

输出端 Head 在进行目标检测时,因为同一位置上会产生大量的候选框,且这些候选框之间可能存在重叠,所以采用非极大值抑制寻找最优的目标边界框,消除冗余的边界框。

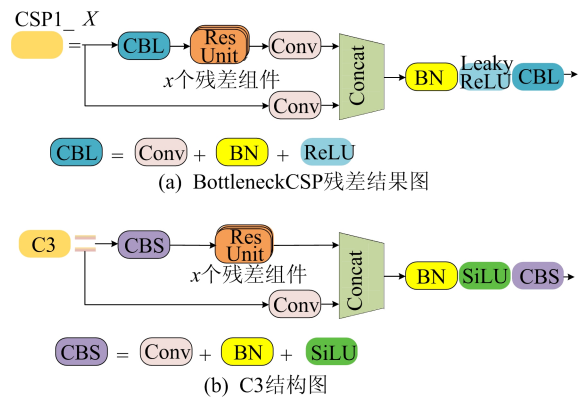


图 2 2 种 CSP 结构

1.2 改进的 YOLOv5s 算法

1.2.1 改进损失函数

YOLOv5 的损失函数由位置损失 (loss-box)、置信度损失 (loss-obj)、分类损失 (loss-cls) 3 个部分组成,损失函数的计算公式为:

$$L_{oss} = l_{oss_box} + l_{oss_obj} + l_{oss_cls} \quad (1)$$

位置损失使用 CIoU 损失,计算公式为:

$$L_{CIoU}(b, b_{gt}) = 1 - C_{IoU} = 1 - \left(I_{IoU} - \frac{d_o^2}{d_c^2} - \frac{v^2}{1 - I_{IoU} + v} \right) \quad (2)$$

其中, $v = \frac{4}{\pi^2} \left(\arctan \frac{w_{gt}}{h_{gt}} - \arctan \frac{b_w}{h_w} \right)^2$ 。

置信度损失与分类损失类似,使用二进制交叉熵损失,计算公式为:

$$L_{obj}(p_o, p_{iou}) = BCE_{obj}^{sig}(p_o, p_{iou}; \omega_{obj}) \quad (3)$$

$$L_{cls}(c^p, c^{gt}) = BCE_{cls}^{sig}(c^p, c^{gt}; \omega_{cls}) \quad (4)$$

其中: I_{IoU} 为预测框与真实框的交并比; d_o 为预测框与真实框中心点的欧式距离; d_c 为预测框与真实框最小外接矩形的对角线距离; v 为宽高比; b_w, h_w 为预测框的宽和高; w_{gt}, h_{gt} 为真实框的宽和

高; p_o 为预测框中目标置信度分数; p_{iou} 为预测框与真实框的交并比; c^p 为预测框中目标类别分数; c^{gt} 为真实框中目标类别分数; ω_{obj} 、 ω_{cls} 分别为预测框中置信度权重和分类权重。

CIoU 虽然考虑了预测框与真实框的重叠面积、中心点距离、宽高比例问题,但式(2)中 v 反映的仅仅是宽高比值大小的差异,不能反映宽、高分别与其置信度的差异,且未能平衡难易样本 这是由于使用单阶段目标检测算法会使得正负样本分配不均衡,产生大量的负样本,易导致负样本因数量优势而淹没掉少量但同样重要的正样本,使得负样本主导损失,最终影响模型的训练效果。因而本文引入 Focal_loss_EIoU 解决该问题,Focal_loss 是一种基于二进制交叉熵的能够动态缩放的交叉熵损失,通过调节动态缩放因子降低训练过程中易区分样本的权重,从而将注意力聚焦在难区分的样本上。EIoU 是在 CIoU 的基础上将宽高的影响拆开,分别计算预测框与真实框的宽和高,其包含重叠损失、中心距离损失、宽高损失 3 个部分,计算公式为:

$$L_{EIoU} = L_{IoU} + L_{dis} + L_{asp} = 1 - I_{IoU} + \frac{\rho^2(b, b^{gt})}{c^2} + \frac{\rho^2(\omega, \omega^{gt})}{c_w^2} + \frac{\rho^2(h, h^{gt})}{c_h^2} \quad (5)$$

1.2.2 添加小目标检测层

YOLOv5s 的检测头采用特征金字塔(FPN)结构,可以用于检测不同尺度的目标,在检测时会将图片缩放到不同尺度得到相应的特征图,进而将不同尺度的特征图进行融合,并在融合的特征图上进行预测,原始的 YOLOv5s 具有 3 个检测头,能够在 8×80 、 4×40 、 2×20 这 3 种尺度的特征图上进行预测,但考虑到本文检测的杂物为小目标,自身携带的信息量小,经过下采样得到的特征图的相对尺度较小,甚至会被筛除掉,严重影响小目标的检测识别率。输入图像的尺寸为 640×640 ,小目标的尺寸为 8×8 ,经过多次下采样后,即使在最大特征图(8×80)上,目标的尺寸也仅有 1×1 ,在其他 2 个特征图上则会消失,因而现有的 YOLOv5s 检测头无法实现对小目标精确检测。

改进的 YOLOv5s 网络结构如图 3 所示。

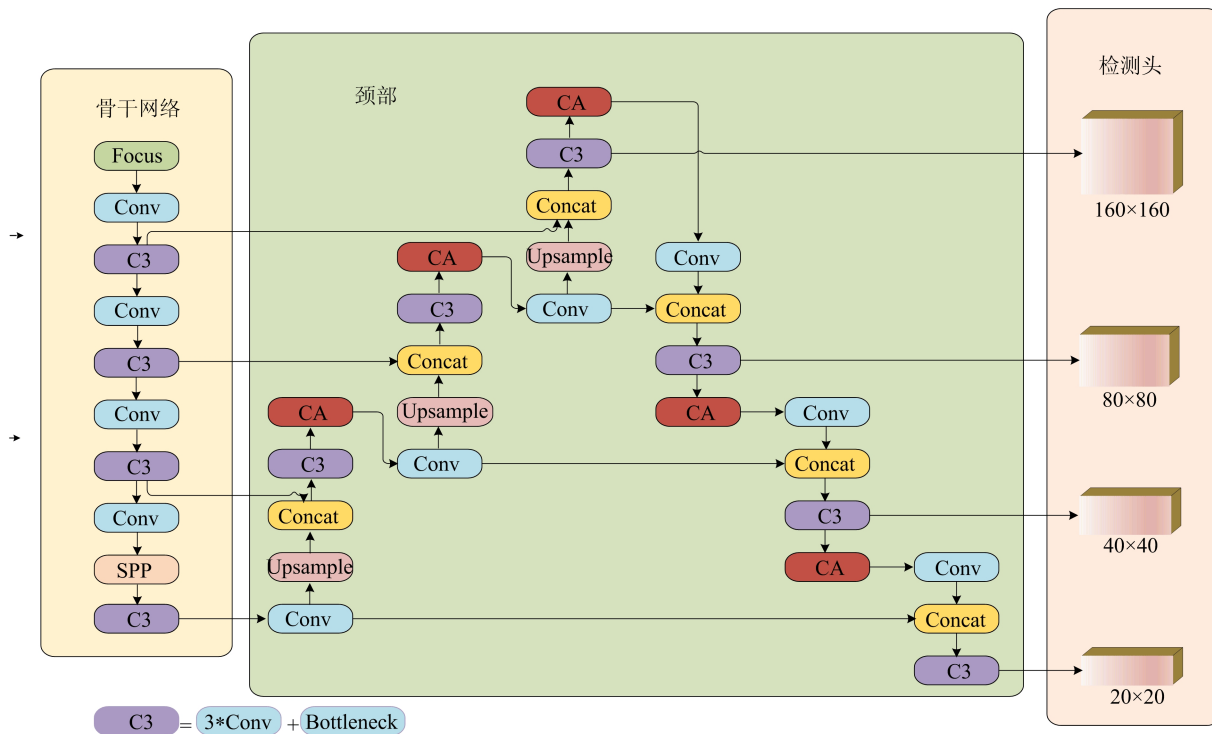


图 3 改进的 YOLOv5s 网络结构

相较于原始的 YOLOv5s,新增一个专用于小目标的检测头,虽然以增大参数计算量为代价,但能够有效提高小目标检测精度,具体改进为在保留原本 3 个检测头的基础上,扩充一层卷积层得到 160×160 大小的特征图。

1.2.3 添加注意力机制

注意力机制自被提出以来,逐渐应用于各个领域,如计算机视觉、统计学习、自然语言处理等。按注意力关注域的不同主要可以分为通道注意力模块(channel attention module, CAM)、空

间注意力模块(spatial attention module, SAM)、混合注意力模块(mix attention module, MAM) 3类,其典型代表有 SE 注意力、STN 注意力、CBAM 注意力、CA 注意力等^[23]。

SE 注意力模块属于通道注意力模块中的一种,通过学习自动获取每一个特征通道的权重,并依据这些权重提升特征并抑制对当前任务不重要的特征,但却忽略了空间特征,SE 注意力模块如图 4 所示。

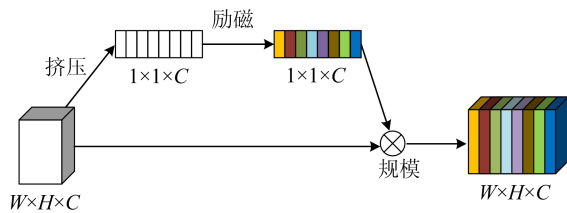


图 4 SE 注意力模块

STN 作为空间注意力模块的典型代表,能够通过学习的方式对输入数据进行任意空间变换,以增强网络对图像旋转、变形等几何变换的鲁棒性,且 STN 可以结合其他深度学习模型使用,从

而提高整个系统的性能,STN 注意力模块如图 5 所示。

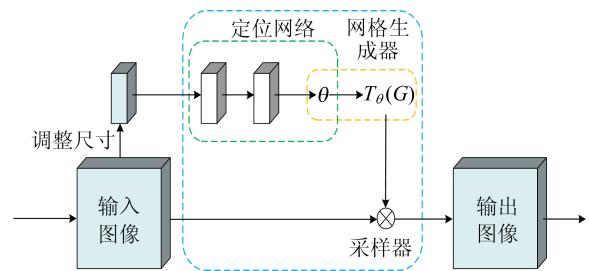


图 5 STN 注意力模块

CBAM 注意力机制属于混合注意力模块中的一种,通过提取通道特征的方式与 SE 注意力模块类似,与之不同的是在 SE 注意力模块的基础上添加了空间结构,将通道注意力模块提取的特征作为空间注意力模块的输入,与 SE 相比,能够在通道中引入空间的位置信息,性能上有了明显提升,但也存在一定的局限性,即通过这种方式只能获得局部范围的信息,而无法从较远距离提取特征,CBAM 注意力模块如图 6 所示。

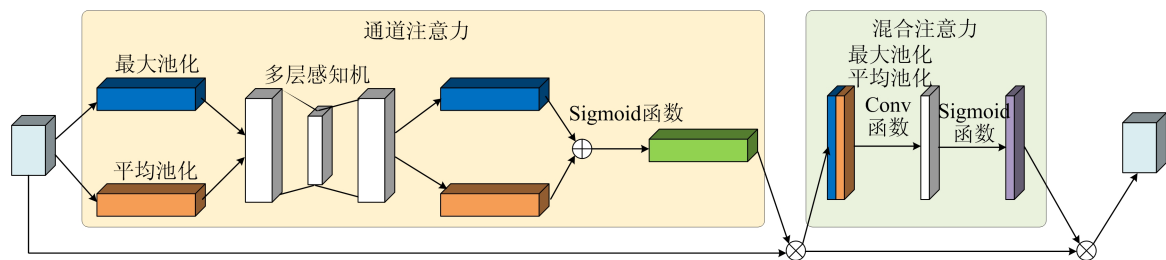


图 6 CBAM 注意力模块

CA 注意力模块虽同 CBAM 一样属于混合注意力模块,但该注意力模块通过将空间中的位置信息嵌入到通道注意力中,使得网络能够在获得更大区域信息的同时,避免产生更大的代价,CA 注意力模块如图 7 所示。

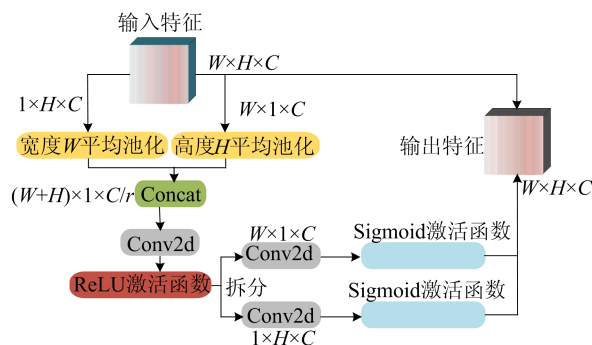


图 7 CA 注意力模块

2 数据集

2.1 数据采集

本文采用海康威视 DS-2DC4423IW-DE 型摄像头拍摄制作数据集,拍摄时摄像头位于烟丝生产线上方约 8 m 处,共采集到 1 174 张尺寸为 1 280 x 720 的原始图像,包含碎纸板、塑料袋、抹布、螺丝刀、扳手等 8 类杂物。

2.2 数据定义

小目标广义上是指尺寸较小、面积较小的目标物体,但在实际中小目标的定义往往取决于具体的应用场景,考虑到本文所用摄像头拍摄的图像尺寸为 1 280 x 720,并结合实际生产中杂物在摄像头画面中的大小,采用相对尺度的方式对不同类型的小目标杂物进行定义,设定目标检测框

面积占图像面积的比例不超过 0.5% 为小目标。

2.3 数据增强

数据增强是强化小目标检测的一个有效手段,其作用在于能够丰富数据集,提高模型的泛化能力和鲁棒性,还能防止因图像数据量小而使模型在训练时产生过拟合现象,常用的数据增强方式有随机旋转、裁剪、缩放、模糊处理、二值化处理、添加噪声、改变亮度等,数据增强如图 8 所示。经过数据增强后,共得到 5 870 张图像,按照 8:1:1 的比例划分为训练集、验证集、测试集,处理完后的训练集、验证集、测试集分别为 4 696、587、587 张图像。

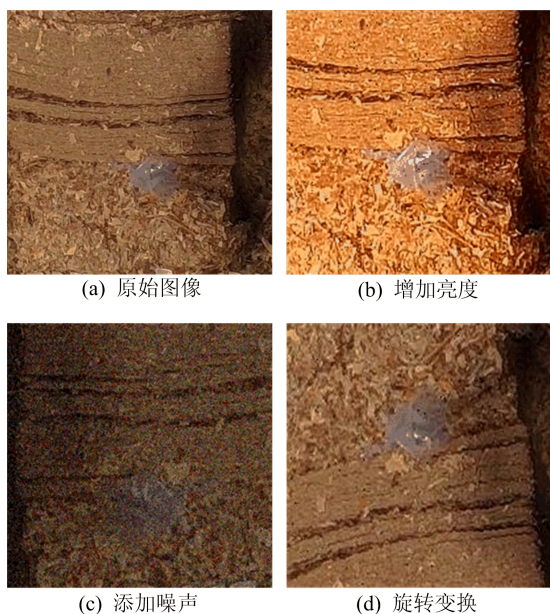


图 8 数据增强方式效果图

3 实验结果与分析

3.1 实验环境

本文的实验环境包括软件环境、硬件环境 2 个部分,软件环境在 Windows 11 操作系统上进行,使用的 CUDA 版本为 11.5,编程语言为 Python 3.9,使用的深度学习框架为 PyTorch 1.12.0,实验环境软件参数见表 1 所列。

表 1 实验环境软件参数

环境名称	版本
操作系统	Windows 11 操作系统
CUDA	11.5
Cudnn	8.3.2
Python	3.9
PyTorch	1.12.0

硬件环境主要是笔记本电脑,电脑的 CPU

为 AMD Ryzen 76800H@3.2 GHz,运行内存为 16 GiB,GPU 为 NVIDIA GeForce RTX 3060,运行显存为 6 GiB,实验环境硬件配置见表 2 所列。

表 2 实验环境硬件配置

硬件环境	配置
CPU	AMD Ryzen 7-6800H@3.2 GHz
GPU	NVIDIA GeForce RTX 3060
运行内存/GiB	16
运行显存/GiB	6

3.2 模型评价指标

本文以精确率 P 、召回率 R 、平均精度平均值 (mean average precision, mAP) 为模型的评价指标,对 YOLOv5s 基准模型及改进后模型进行性能评估,计算公式为:

$$P = T_P / (T_P + F_P) \quad (6)$$

$$R = T_P / (T_P + F_N) \quad (7)$$

$$P_{mAP} = \left(\sum_{i=1}^n A_{P_i} \right) / n \quad (8)$$

其中, $A_P = \sum_{k=1}^n P(k) \Delta r(k)$ 。

3.3 消融实验验证

基于所制的杂物数据集,通过消融实验来验证各种改进策略的合理性,模型训练输入图像的尺寸均为 640×640 ,迭代训练次数均为 300 次,同时启用早停策略,若连续验证 5 次,模型性能没有提升,则提前终止训练。

3.3.1 损失函数验证

YOLOv5s 自带的损失函数 BCE_Loss_CIoU 和改进后的损失函数 Focal_loss_EIoU 的损失值变化曲线如图 9 所示。从图 9 可以看出:在位置损失方面,2 种损失函数的曲线大致重合,但改进后的 Focal_loss_EIoU 损失曲线更加稳定,波动更小;在置信度损失和分类损失方面,改进后的损失函数具有更低的损失值,表明模型对目标预测得更加准确。

BCE_loss_CIoU、Focal_loss_EIoU 召回率变化曲线如图 10 所示。从图 10 可以看出,在模型训练初期,相同训练次数下 Focal_loss_EIoU 的召回率更高,改进前的损失函数在训练第 6 个迭代左右达到收敛,而改进后的损失函数在第 2 个迭代达到收敛,表明 Focal_loss_EIoU 能够加快模型的收敛速度,且 Focal_loss_EIoU 加强了对难分类样本的训练,因而召回率、mAP 均有一定程度的提高,改进了损失函数的实验 2 相较改进前的实验 1,召回率、mAP 分别提高 2.0%、

1.8%,充分验证了改进模型的有效性,因而本文余下的实验,均在 Focal_loss_EIoU 基础上进行。

添加小目标检测头模型的精确率和召回率均有一定提高。实验 3 与实验 2 相比,精确

率、召回率、mAP 分别提高 3.4%、1.5%、3.0%,充分说明添加小目标检测头能够有效提升对小目标杂物的检测,同时也验证改进模型的有效性。

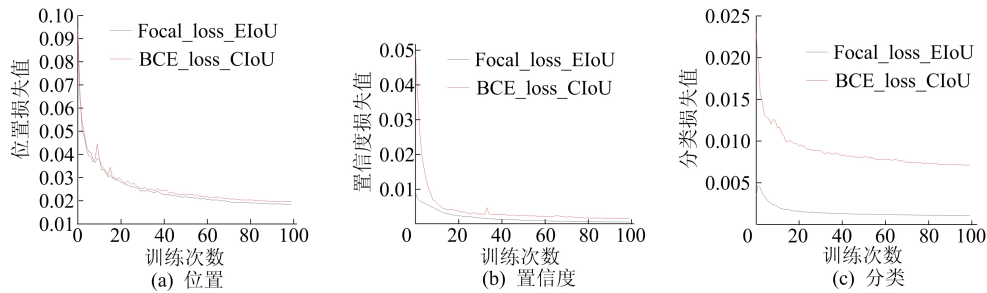


图 9 BCE_loss_CIoU、Focal_loss_EIoU 损失值变化曲线

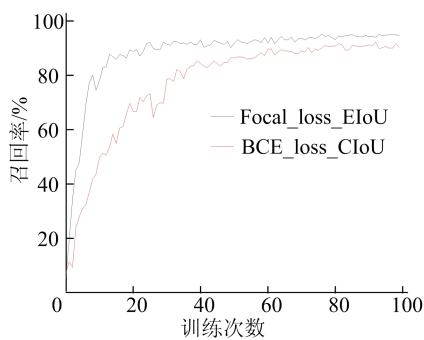


图 10 BCE_loss_CIoU、Focal_loss_EIoU 召回率变化曲线

3.3.2 小目标检测头验证

添加小目标检测头后,对比分析 2 种模型的评价指标,添加检测头前、后精确率、召回率变化曲线如图 11 所示。

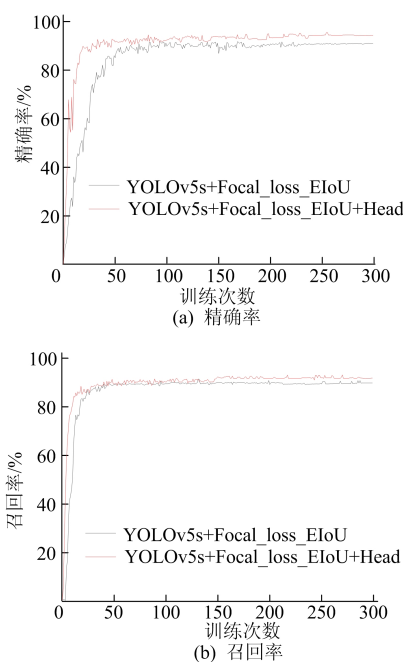


图 11 添加检测头前、后精确率和召回率变化曲线

3.3.3 注意力模块验证

为探究通道注意力模块、空间注意力模块和混合注意力模块对模型性能的影响,本文在原始 YOLO v5s 的颈部分别添加 SE、STN、CBAM、CA 注意力模块,经过 300 轮迭代,4 种注意力模块的性能对比见表 3 所列。

表 3 不同注意力模块性能对比 %

模块	P	R	mAP@0.5
SE	93.1	85.3	84.5
STN	93.4	84.5	83.7
CBAM	94.3	83.1	85.3
CA	93.9	87.6	88.9

从表 3 可以看出:添加 4 种注意力模块后,CBAM 和 CA 注意力精确率较高,这是由于两者均属于混合注意力模块,能够综合考量通道信息和空间信息的重要性;而 SE 和 STN 注意力模块只能单独考量通道信息或空间信息的重要性;CA 注意力的 mAP 值高于 CBAM,这是由于 CBAM 的空间模块只能获取局部范围的信息,而忽略了其他位置的信息。

本文选取综合性能最优的 CA 注意力模块,并将其添加在 YOLOv5s 颈部的 5 个 C3 层,训练 300 次后与仅改进损失函数的模型进行对比,结果发现模型的精确率有一定提升,实验结果对比见表 4 所列。添加 CA 注意力模块前、后精确率变化曲线如图 12 所示。

与只改进了损失函数的实验 2 相比,添加 CA 注意力模块的实验 4 的精确率提高 2.6%,召回率略有下降,mAP 值略有提升,这是由于添加注意力模块后使得模型更加聚焦于当前任务更为关键的部分,而降低了对其他信息的关注度,甚至

忽略了部分信息,因而召回率有一定程度的降低,实验结果表明,引入 CA 注意力模块对于杂物特

征的提取有明显提升,验证了在 YOLOv5s 中添加 CA 注意力模块的合理性。

表 4 实验结果对比

实验编号	模型	P	R	mAP@0.5
1	YOLOv5s (BCE_loss_CIoU)	91.5	88.1	87.3
2	YOLOv5s+Focal_loss_EIoU	90.8	90.1	89.1
3	YOLOv5s+Focal_loss_EIoU+Head	94.2	91.6	92.4
4	YOLOv5s+Focal_loss_EIoU+CA	93.4	89.7	89.5

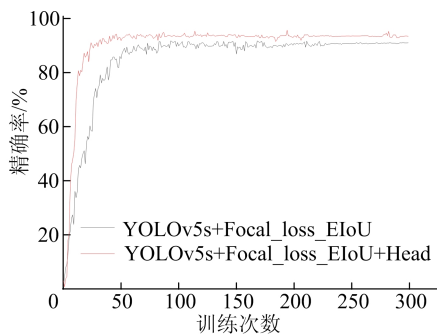


图 12 添加 CA 注意力模块前、后精确率变化曲线

3.3.4 改进 YOLOv5s 验证

将改进后的 YOLOv5s 与原始 YOLOv5s 的性能进行对比,结果如图 13 所示,改进后 YOLOv5s 的精确率、召回率、mAP 皆优于 YOLOv5s,由表 4 可知,相较于 YOLOv5s,改进后模型的精确率、召回率、mAP 分别提高 4.9%、5.5%、7.3%。

基于本文改进后的模型,杂物检测效果如图 14 所示。由图 14 可知,改进后的模型能够准确识别出烟丝背景下的塑料袋、碎纸板等杂物。

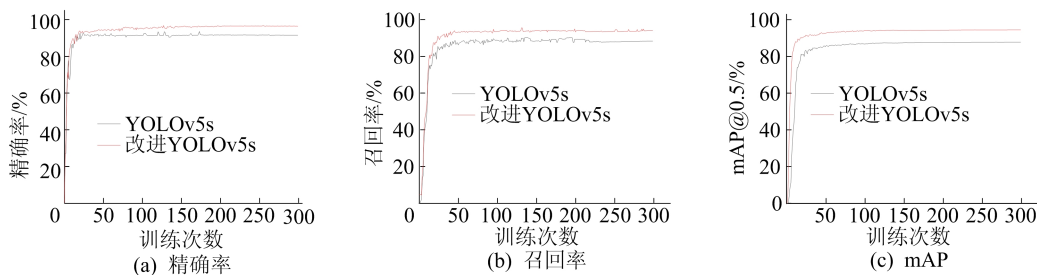


图 13 改进前、后 YOLOv5s 的精确率、召回率、mAP 变化曲线

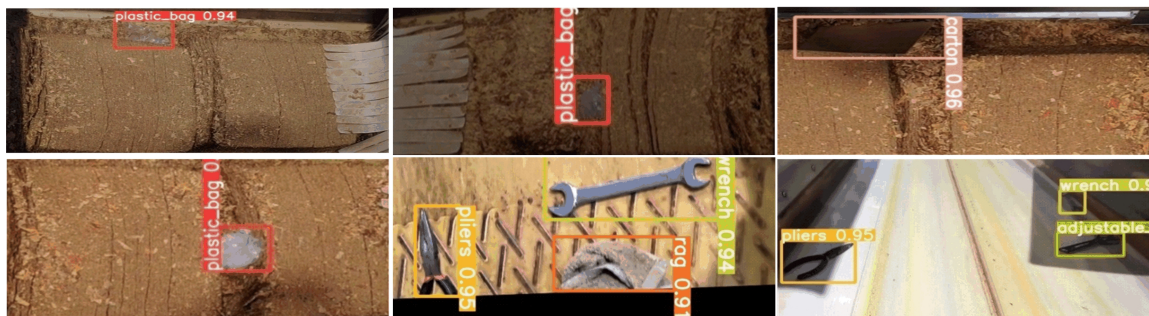


图 14 杂物检测效果

目前 YOLOv8s 已经发布,且 YOLOv6 与 YOLOv7 存在重复,为丰富对比实验,验证模型

的可行性,本文将改进后的模型与 YOLOv7、YOLOv8s 模型性能进行对比,结果见表 5 所列。

表 5 模型性能对比

模型	P/%	R/%	mAP@0.5/%	内存/MiB	帧数/(帧/s)
YOLOv5s	91.5	88.1	87.3	13.9	67
改进 YOLOv5s	96.4	93.6	94.6	15.6	62
YOLOv7	96.9	92.4	93.3	74.8	27
YOLOv8s	96.7	93.1	94.2	21.7	36

从表5可以看出,YOLOv7、YOLOv8s在精确率上略高于改进的YOLOv5s,这是由于YOLOv7、YOLOv8s模型更复杂,学习能力更强,但其所占内存较大,推理速度较慢,不满足生产线上实时性检测的需求,而本文改进的YOLOv5s能够兼顾检测实时性和识别精确率,因此本文模型在烟厂制丝生产线环境下具有一定的可行性和优越性。

4 结 论

针对烟厂小目标杂物难以检测的问题,本文提出一种改进的YOLOv5s检测模型,该模型以目标检测的通用模型YOLOv5s为基础模型,在检测头部分增加一个专用于小目标的检测头,有效提高了对小目标杂物的检测精度;在模型颈部添加CA注意力模块,增强了模型对目标杂物特征的提取能力;使用更加灵活的损失函数,着重关注难分类样本,有效提高了模型的收敛速度,改进后模型的精确率、召回率、mAP均有所提升,证明本文改进策略的合理性。模型添加了一个检测头,使得模型参数增多,导致模型的存储空间有一定程度的增大,因此下一步的研究工作是在保证模型精度的同时对模型进行轻量化处理。

[参 考 文 献]

- [1] 戴云周,路红,纪陈阳. 基于改进YOLO v5的电网工人防具检测[J]. 南京工程学院学报(自然科学版), 2023, 21(1): 33-38.
- [2] 叶卓勋,刘妹琴,张森林. 基于轻量化深度学习网络的工业环境小目标缺陷检测[J]. 控制与决策, 2023, 38(5): 1231-1238.
- [3] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot multibox detector[J]. Computer Vision-ECCV, 2016, 9: 21-37.
- [4] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection[C]//Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. [S. l. : s. n.], 2016: 779-788.
- [5] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]//2014 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). [S. l. : s. n.], 2014: 580-587.
- [6] HE K M, ZHANG X Y, REN S Q, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37(9): 1904-1916.
- [7] REN S Q, HE K M, GIRSHICK R, et al. Faster R-CNN: towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2017, 39(6): 1137-1149.
- [8] 高新波,莫梦竟成,汪海涛,等. 小目标检测研究进展[J]. 数据采集与处理, 2021, 36(3): 391-417.
- [9] CHEN C Y, LIU M Y, TUZEL O, et al. R-CNN for small object detection[J]. Computer Vision-ACCV, 2016, 11(3): 214-230.
- [10] LIN T Y, MAIRE M, BELONGIE S, et al. Microsoft COCO: common objects in context[J]. Lecture Notes in Computer Science, 2014, 8693(1): 740-755.
- [11] 温秀兰,焦良葆,李子康,等. 复杂环境下小尺度烟火目标检测研究[J]. 南京信息工程大学学报, 2023, 15(6): 676-683.
- [12] ZHOU Q H, LIU H S, QIU Y H, et al. Object detection for construction waste based on an improved YOLOv5 model[J]. Sustainability, 2022, 15: 681.
- [13] YANG R J, LI W F, SHANG X N, et al. KPE-YOLOv5: an improved small target detection algorithm based on YOLOv5[J]. Electronics, 2023, 12(4): 817.
- [14] 刘春霞,李超,潘理虎,等. 改进YOLOv5s的煤矿烟火检测算法[J]. 计算机工程与应用, 2023, 59(17): 286-294.
- [15] 尹靖涵,瞿绍军,姚泽楷,等. 基于YOLOv5的雾霾天气下交通标志识别模型[J]. 计算机应用, 2022, 42(9): 2876-2884.
- [16] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2017, 99: 2999-3007.
- [17] HU J, SHEN L, ALBANIE S, et al. Squeeze-and-excitation networks[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). [S. l. : s. n.], 2018: 18-23.
- [18] JADERBERG M, SIMONYAN K, ZISSERMAN A, et al. Spatial transformer networks[C]//Advances in Neural Information Processing Systems. Cambridge, Mass. : MIT Press, 2015: 2017-2025.
- [19] WOO S, PARK J, LEE J Y, et al. CBAM: convolutional block attention module[J]. Lecture Notes in Computer Science, 2018, 11211: 3-19.
- [20] BIN H Q, QUAN Z D, SHI F J. Coordinate attention for efficient mobile network design [C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). [S. l. : s. n.], 2021: 20-25.
- [21] 马琳琳,马建新,韩佳芳,等. 基于YOLOv5s目标检测算法的研究[J]. 电脑知识与技术, 2021, 17(23): 100-103.
- [22] LIN T Y, DOLLAR P, GIRSHICK R, et al. Feature pyramid networks for object detection[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). [S. l. : s. n.], 2017: 21-26.
- [23] 李丹阳. 基于机器学习的交通标志检测与识别方法研究[D]. 郑州: 中原工学院, 2022.

(责任编辑 张 镛)