

DOI:10.3969/j.issn.1003-5060.2026.03.005

基于改进 YOLOv5s-seg 的废钢分类分割算法

王哲¹, 杨文², 范彬彬¹, 顾志远¹, 吴孟武¹

(1. 武汉理工大学汽车工程学院, 湖北武汉 430070; 2. 中冶南方工程技术有限公司, 湖北武汉 430223)

摘要:为解决废钢回收场景中人工分拣效率低和安全隐患高等问题,文章基于改进 YOLOv5s-seg 算法建立废钢分类及实例分割模型,旨在利用计算机视觉技术代替人工分拣。在主干网络中引入 CBAM(convolutional block attention module)注意力机制强调废钢特征,同时用 EIOU(efficient intersection over union)替换原网络中的 CIOU(complete intersection over union)损失函数,加快收敛速度;通过线下数据增强算法对废钢厂采集的图片进行图像增强,并在改进前、后算法构建的两类模型上训练及验证。结果表明:改进的 YOLOv5s-seg 废钢分类及实例分割模型边界框和掩膜的平均检测精度达到 98%、96%,比原模型提高了 5%、3%;全类别分类平均精度高达 96.5%,比原模型提高了 5.8%;改进模型与其他经典的实例分割模型相比也表现出优异的检测性能。该文所提算法具有较好的应用前景。

关键词:钢铁回收料;废钢分类;实例分割;注意力机制;损失函数

中图分类号:TP274.5 **文献标志码:**A **文章编号:**1003-5060(2026)03-0317-08

Steel scrap classification and segmentation algorithm based on improved YOLOv5s-seg

WANG Zhe¹, YANG Wen², FAN Binbin¹, GU Zhiyuan¹, WU Mengwu¹

(1. School of Automotive Engineering, Wuhan University of Technology, Wuhan 430070, China; 2. WISDRI Engineering and Research Incorporation Limited, Wuhan 430223, China)

Abstract: In order to solve the problems of low efficiency and high security risks of manual sorting in steel scrap recycling scenarios, a steel scrap classification and instance segmentation model was established based on the improved YOLOv5s-seg algorithm, aiming to replace manual sorting with computer vision technology. Convolutional block attention module(CBAM) was introduced into the backbone network to emphasize the characteristics of steel scrap. At the same time, the complete intersection over union(CIOU) loss function in the original network was replaced by efficient intersection over union(EIOU) to accelerate the convergence speed. The offline data enhancement algorithm was used to enhance the images collected from the steel scrap mill, and the models before and after the improvement were trained and verified. The results show that the average detection accuracy of the improved YOLOv5s-seg steel scrap classification and instance segmentation model for boundary box and mask reaches 98% and 96%, respectively, which is 5% and 3% higher than that of the original model; the average accuracy for all categories reaches 96.5%, an increase of 5.8% compared with the original model. In addition, the improved model has excellent detection performance compared with other classical instance segmentation models. The proposed algorithm has a good application prospect.

Key words: steel recycling materials; steel scrap classification; instance segmentation; attention mechanism; loss function

收稿日期:2023-11-24;修回日期:2024-04-29

基金项目:国家自然科学基金资助项目(52275370);湖北省重点研发计划资助项目(2022BAD100)

作者简介:王哲(1997—),女,河南鹤壁人,武汉理工大学硕士生;

吴孟武(1985—),男,湖北麻城人,博士,武汉理工大学副教授,博士生导师,通信作者,E-mail:wumw@whut.edu.cn.

废钢作为绿色再生资源,能够大规模代替铁矿石,实现无限循环利用。尽管我国废钢铁资源增长迅速,但钢铁回收料的回收利用率仍相对较低^[1-2]。目前废钢回收主要依赖人工分拣,效率低下且存在安全隐患。随着机器视觉技术的迅猛发展,智能检测算法在工业场景中得到广泛应用,在分拣检测领域也展现出巨大的增长潜力^[3-4]。在废钢回收场景中引入机器视觉技术,通过快速精准识别、检测和分类废钢实现废钢回收的自动化、智能化,可以提高废钢回收利用效率,最大限度地实现废钢资源的循环利用^[5]。

对目标进行识别分类的机器视觉方法分为目标检测算法^[6]和实例分割算法^[7]两类。目前国内对外废金属的自动分类主要借助目标检测算法来实现。在对小目标的有色金属分类中:文献[8-9]采用 DenseNet 分类神经网络对铸铝和锻铝合金进行识别,测试集上的分类精度达 95%,之后又在铸铝和锻铝合金图像基础上加入深度图像和激光诱导击穿光谱图做网络融合,识别准确率增加至 98%;文献[10]提出一种基于多尺度特征的 YOLOv5 铸件检测算法,检测精度达 96.5%;文献[11]提出卷积神经网络和 SEEDS(superpixels extracted via energy-driven sampling)融合算法,实现了铜、铝有色金属废弃物的自动分类;文献[12]采用深度神经网络对散布在传送带上的金属碎片进行分类,比较分析了 AlexNet、GoogleNet、VGGNet 和 ResNet 4 种不同深度网络模型的精度,得出 AlexNet 模型精准度最高;文献[13]提出一种小样本有色金属目标检测模型,在深度网络模型 YOLOv3 的基础上进行改进,实现了对铜和铝的识别。在废钢回收分类方面:文献[14]采用优化的 YOLOv3 网络对轻、中、重型废钢进行分类,精度值由 90.20%提升到 95.05%,但仍无法对重叠目标进行检测;文献[15]提出一种结合注意力与特征融合的深度神经网络模型,将废钢分成 7 个种类来实现对废钢的判级;文献[16]使用 VGG-16、ResNet-50、DenseNet-121、EfficientNet-80 4 种深度神经网络预训练模型对废钢分类数据集进行特征提取及融合,分类准确率高达 99.17%。废钢智能化回收场景中通常需要在废钢分类识别的基础上引导抓钢机或吸盘执行抓取、分拣等作业,此时需要得到废钢的形状和在图像上的具体位置等信息,这就要求对废钢进行精准的实例分割,然而相关研究却很少。

本文提出一种基于改进 YOLOv5s-seg 的废

钢分类分割算法并构建模型,通过线下数据增强方式扩充数据集提高模型的泛化能力,在主干网络中融入 CBAM(convolutional block attention module)^[17]卷积注意力机制,边界框损失函数用 EIOU Loss (efficient intersection over union loss)^[18]替换原算法中的 CIOU Loss (complete intersection over union loss)^[19],加快模型收敛速度,提高对废钢区域检测和分割的准确性。最后,利用废钢厂采集的图片进行模型训练,并对模型的分分类分割效果进行验证。

1 YOLOv5s-seg 算法

YOLOv5s-seg 算法^[20]由 Ultralytics 于 2022 年 8 月公开发布,用于实现对物体的分类、定位、检测、分割,共有 5 个版本,按照算法框架的宽度和深度从小到大排列,依次为 YOLOv5nseg、YOLOv5sseg、YOLOv5mseg、YOLOv5lseg、YOLOv5xseg。随着宽度和深度的增加,基于该算法构建的模型检测速度由快转慢。考虑到实时检测和方便部署,本文选择 YOLOv5s-seg 作为基础算法构建废钢分类分割模型进行后续实验。

YOLOv5s-seg 模型结构如图 1 所示,由主干网络(Backbone)、颈部网络(Neck)、头部网络(Head)3 个部分组成。主干网络用于提取图像中丰富的特征,由 Conv、C3、SPPF(spatial pyramid pooling-fast)组件组成,组件结构如图 2a 所示。图像首先经过一个内核为 6×6 、步长为 2、通道数为 64 并具有批量归一化(batch normalization 2d, BN)层和 SiLU(sigmoid linear unit)激活函数的 Conv 组件进行初始化;其次通过 4 个 C3 模块,每个 C3 模块先通过步长 $s=2$ 的卷积进行降采样,降采样率设置为 2、4、8、16,经过降采样后图像被分成相同通道数的 2 个部分,第 1 部分使用核大小为 3×3 和 1×1 的连续卷积层进行卷积操作,第 2 部分不操作,直接与第 1 部分的最终输出相连接;最后通过 SiLU 函数激活每个 Conv 层,以提高 BN 层后的深度神经网络性能。C3 块的残差块数 m 为 1、2、3、1,通道数为 128、256、512、1 024。C3 网络后提取的特征图传递给 SPPF 模块。SPPF 是对空间金字塔池化(spatial pyramid pooling, SPP)^[21]的改进,由具有 BN 层和 SiLU 激活的 Conv 层和 MaxPool 层组成,它是一种特征增强网络,提取特征图的主要信息并进行拼接,实现了局部特征与全局特征的融合,以减少目标检测的损失。

颈部网络从主干网络中获得相对较浅的特征,然后将它们与更深的语义特征融合。YOLOv5s-seg 是采用特征金字塔网络(feature pyramid network,FPN)^[22]与路径聚合网络(path aggregation network,PAN)^[23]相结合的结构,FPN 网络将主干网络中 SPPF 传出的底层特征进行卷积和上采样,与主干网络中同通道数的特征图做加和操作(Concat),得到双倍维度的图像特征,再将信息传递到如图 2b 所示的 C3 网络中进行输出检测,实现自顶向下捕获强语义特征。PAN 网络通过 $s=2$ 的步幅进行卷积下采样,与剩余特征层堆叠,实现自底向上传达强定位特征。颈部网络最终产生 20×20 、 40×40 、 80×80 3 种

尺寸的特征图用于头部的预测。

头部网络是检测的最终输出,结构如图 2c 所示。它采用目标检测头 Detect 与全连接神经网络 ProtoNet 相结合输出预测结果。Detect 提取颈部输出的小、中、大特征图的特征,输出废钢类别概率、位置、置信度以及每个边界框所对应的掩膜系数。ProtoNet 是一个具有 SiLU 激活函数由 2D 卷积组成的 3 层网络,其输出原型掩膜,与目标框中的掩膜系数加权求和,从而获得分割结果。头部网络输出结果通过非极大抑制(non-maximum suppression,NMS)映射要素上的锚点来选择合适的废钢目标,并去除冗余信息,最终得到废钢目标的类别、位置、置信度及掩膜信息等。

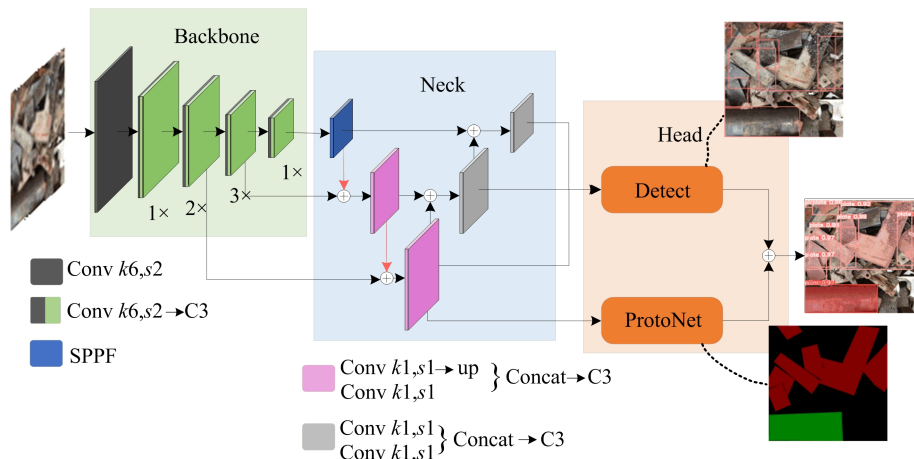


图 1 YOLOv5s-seg 模型结构

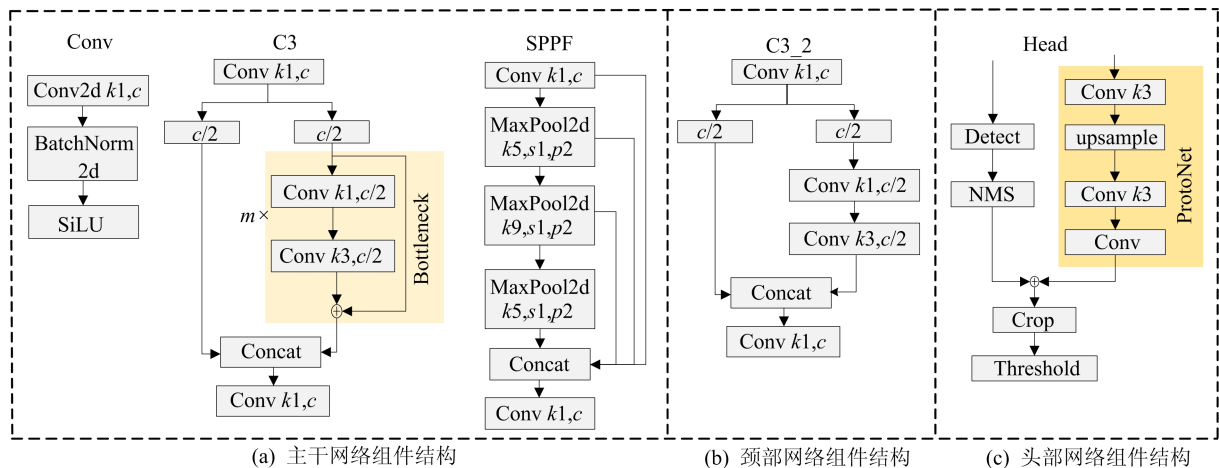


图 2 YOLOv5s-seg 模型组件结构

2 废钢分类分割算法改进

2.1 CBAM 卷积注意力机制

CBAM 卷积注意力机制是一种用于计算机视觉任务的注意力模块,它可以增强卷积神经网络

的表达能力,提高特定通道和空间特征的关注度,从而改善模型的性能。

CBAM 注意力机制的结构图如图 3 所示。 X 为输入特征图, W 、 H 、 C 分别为特征图的宽度、高度及通道数。首先将 X 分别经过基于宽度和高

度的全局最大池化(global max pooling, GMP)和全局平均池化(global average pooling, GAP),得到 2 个通道维度的特征向量,并将其输入到多层感知机(multilayer perceptron, MLP)网络,经过逐元素乘法操作和 sigmoid 激活函数,生成 $1 \times 1 \times C$ 的通道注意力特征图;再将 S_1 与 X 做 element-wise 加和操作,输出特征图 X_1 ;然后将 X_1 作为空间注意力模块的输入,分别经历 2 个基于通道的 GMP 和 GAP,将两者的结果进行 Concat 通道拼接,得到的特征向量再经过 7×7 的卷积操作,将通道数降维到 1,经过 sigmoid 激活函数生成空间注意特征层 S_2 ;最后将 S_2 与 X 做逐元素乘法,生成最终的特征图 X_2 。

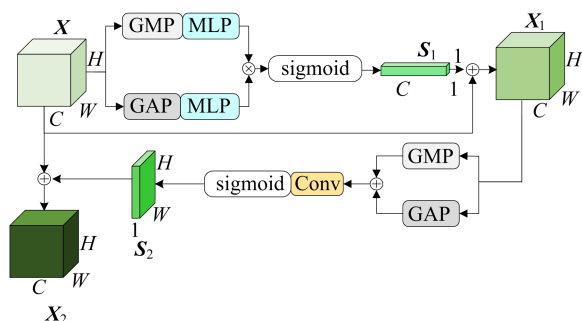


图 3 CBAM 注意力机制结构

由于废钢回收场景杂乱、废钢形状尺寸差异大并且存在遮挡问题,为此将 CBAM 注意力机制引入至 YOLOv5s-seg 算法的主干网络 C3 模块中。具体地,将 CBAM 注意力机制插入在 Bottleneck 跳跃连接之前的位置,如图 4 所示。

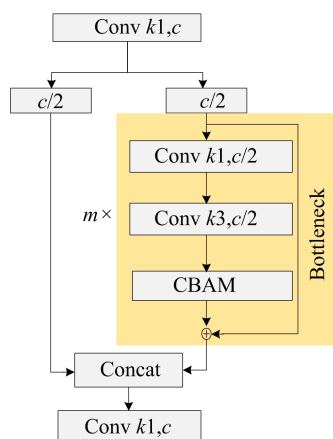


图 4 C3_CBAM 网络结构

模型在跳跃连接之前可以根据 CBAM 注意力机制来相应调整 C3 模块中的特征表示,根据废钢特征选择性地强调特定的通道和空间位置,有助于

更好地捕捉图像中废钢的重要信息。本文将主干网络中的 C3 模块全部替换为 C3-CBAM 模块。

2.2 损失函数

为了衡量模型预测质量并显示预测值与实际值之间的差距,本文置信度损失、分类损失及掩膜损失采用二进制交叉熵(binary cross entropy, BCE)损失函数计算,而边界框损失函数替换原 YOLOv5-seg 模型中的 CIOU Loss,采用收敛速度更快、回归精度更高、对样本不均衡情况更加友好的 EIOU Loss 来度量预测框与真实框之间的位置误差。模型总损失为 4 类损失的平均值,即

$$L_{\text{oss}} = \lambda_1 L_{\text{box}} + \lambda_2 L_{\text{obj}} + \lambda_3 L_{\text{cls}} + \lambda_4 L_{\text{seg}} \quad (1)$$

其中, $\lambda_i (i=1, 2, 3, 4)$ 为平衡系数。

位置损失为:

$$L_{\text{box}} = L_{\text{IOU}} + L_{\text{dis}} + L_{\text{asp}} = 1 - \text{IOU} + \frac{\rho^2(\mathbf{b}, \mathbf{b}_{\text{gt}})}{\omega_c^2 + h_c^2} + \frac{\rho^2(\omega, \omega_{\text{gt}})}{\omega_c^2} + \frac{\rho^2(h, h_{\text{gt}})}{h_c^2} \quad (2)$$

其中: L_{IOU} 、 L_{dis} 、 L_{asp} 分别为交并比损失、距离损失和方向损失; \mathbf{b} 和 \mathbf{b}_{gt} 、 ω 和 ω_{gt} 、 h 和 h_{gt} 分别为预测框和真实框的中心点坐标、宽度和高度; $\rho(\cdot)$ 为欧氏距离; ω_c 、 h_c 分别为覆盖真实框与预测框的最小边界框的宽度和高度。

二进制交叉熵损失函数对矩阵逐元素进行计算的公式为:

$$L_{\text{BCE}} = -L \ln P - (1 - L) \ln(1 - P) \quad (3)$$

其中: L 为标签矩阵; P 为预测矩阵。在计算不同类型的损失时, L 、 P 即为该类型矩阵。置信度、分类和掩膜损失的计算公式分别为:

$$L_{\text{obj}} = \frac{a}{N_T} \sum L_{\text{BCE}} + \frac{1-a}{N_F} \sum L_{\text{BCE}} \quad (4)$$

$$L_{\text{cls}} = \frac{1}{N_T} \sum L_{\text{BCE}} \quad (5)$$

$$L_{\text{seg}} = \frac{1}{N_{\text{mask}}} \sum_{i=1}^N \frac{1}{A_i} \sum L_{\text{BCE}} \quad (6)$$

其中: a 为平衡系数; N_T 为正样本的数量; N_F 为负样本的数量; N_{mask} 为预测的掩膜数量; A_i 为第 i 个预测掩膜对应的区域面积。

3 试验及结果分析

3.1 原始数据集

本文废钢数据集采集于某废钢回收加工企业,将废钢类别划分为柱型(Pillar)、板型(Plate)、结构件(Structure)、气瓶(GC)、灭火器(FE)5 个类别。每张图片包括多种类别的废钢,共 700 张。图片采用 labelme 工具进行标注,生成含有类别信息和物体轮廓坐标的 json 文件。由于废钢图

像样本不足,为了使模型充分学习到不同类型的废钢特征,本文废钢数据集按照 9 : 1 的比例随机划分训练集和验证集,并对其进行数据增强。为保证模型推理的精准度,精简数据集制作过程,另外拍摄 3 张废钢图片作为验证集。

3.2 线下图像增强

数据集数量有限可能会导致模型出现过拟合或欠拟合现象,无法充分学习到废钢的特征。因此本文采用镜像、旋转、加噪声、图像锐化、Copy-Paste^[24]的方法扩充数据集。具体分为以下 2 种方式:① 为了解决总体数据样本不足的问题,采用传统的数据增强方法包括镜像、旋转、加噪声以及图像锐化,对原始废钢图片和对应的 json 文件随机地同时使用 1 种或多种方法扩充;② 为了让实例分割模型更加充分地学习不同类型的废钢特征,采用 Copy-Paste 方法对含有结构件和气瓶的图片进行数据扩充。Copy-Paste 数据增广方法的核心是将一个图像中精确分割的实例掩码随机复制粘贴到另一个图像中,从而生成新的训练数据。这种方法直接进行像素级的粘贴操作,而无需考虑实例位置的视觉联系。通过掩膜融合公式,可以在不同的位置粘贴不同的实例,从而产生多样化的训练数据,即

$$I = I_1\alpha + I_2(1 - \alpha) \quad (7)$$

其中: I 为新图像; I_1 为粘贴图像; I_2 为主图像; α 为实例掩码(mask)。

为了保证数据集不被污染,将由于镜像导致的废钢特征消失、旋转导致的标注文件中轮廓超出图片范围以及 Copy-Paste 导致的实例掩码互相遮挡等影响废钢特征判断的图片及 json 文件删除。最终,数据集增强前、后图片和废钢标签的数量见表 1 所列。

表 1 数据集增强前、后数量统计

废钢类别	标签名称	原数据集/张	原标签数	增强后数据集/张	增强后标签数
柱型	Pillar	233	617	3 057	8 061
板型	Plate	212	516	2 852	6 804
结构件	Structure	86	390	1 130	5 106
灭火器	FE	137	266	1 781	3 458
气瓶	GC	79	297	1 138	4 937

增强后的数据集也为最终的废钢数据集,包括原始的以及通过 2 种增强方法得到的图片和 json 文件。

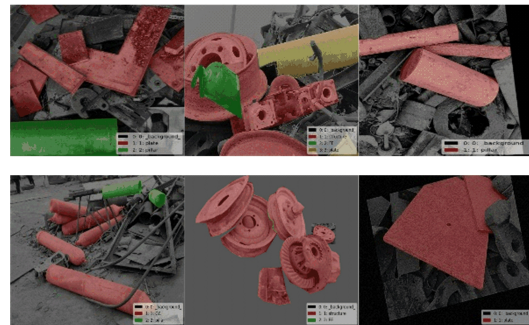
废钢数据集的示例图像如图 5 所示。图 5

中:图 5a 第 1 列为原始图像,另 2 列为经过数据增强后的图像;图 5b 为图 5a 的 json 文件可视化图像。

线下图像增强后的数据集提供了更加多样化的样本,有助于改善模型的泛化能力,从而在实际环境中取得更稳定的预测效果。



(a) 废钢数据集图像



(b) 数据集的可视化图像

图 5 废钢数据集示例图像

3.3 试验环境

本试验在 Ubuntu 20.04.2 LTS 操作系统中完成,CPU 型号为 Intel(R) Xeon(R) CPU E5-2680 v4 @ 2.40 GHz,CUDA 版本 12.0,显卡为 NVIDIA GeForce RTX 2080 Ti,使用 PyTorch 2.0 深度学习框架搭建模型,采用 Python 语言编程,Visual Studio Code 编码。模型的图片输入均为 640×640 像素,采用预训练权重训练,训练迭代次数为 300 次,bachsize 设置为 16。

3.4 评价指标

为了评估多标签检测分类器和掩膜分割的性能,选取精准率(Precision) P 、召回率(Recall) R 、全类别的平均精度(mean average precision, mAP) m_{AP} 、损失函数作为模型的评价指标。就类别而言,被分类器判断正确的废钢记作正例 T_P ,实际为负例的废钢被错误的预测为正例的废钢记作 F_P ,实际为正例的废钢被错误的预测为负例的废钢记作为 F_N ;就掩膜分割而言,以预测边界框

与真实标记边界多边形的重叠是否超过 IOU 阈值为基准划分正、负例,划分标准与上述类别一致。精准率 P 表示为正确检测到的废钢与废钢检测总量的比率,精准率越高,表示类别判断错误的数量越少。召回率 R 是正确检测到的废钢与数据集中废钢总数的比率;平均召回率 R_A 是所有类别召回率的平均值,平均召回率越高,表示目标的漏检率越低。精准率 P 、召回率 R 及全类别平均精度 m_{AP} 的计算公式分别为:

$$P = \frac{T_P}{T_P + F_P} \quad (8)$$

$$R = \frac{T_P}{T_P + F_N} \quad (9)$$

$$m_{AP} = \frac{1}{N} \sum_{k=1}^N \int_0^1 P(R) dR \quad (10)$$

3.5 试验结果

在废钢分类数据集上对改进前、后的 YOLOv5s-seg 进行训练,并开展分类分割对比试验。改进前、后的废钢实例分割模型在验证集上的评价结果见表 2 所列。

由表 2 可知:虽然改进后模型的边界框和掩膜在精准率上提升幅度微小,但在召回率上有明显的提升,分别比原模型提高了 7%、5%;与此同时,在 $I_{OU}=0.5$ 时,边界框和掩膜的平均检测精度 ($m_{AP}@0.5$) 达 98%、96%,比原模型提升了 5%、3%,废钢检测正确率进一步增加。

表 2 模型改进前、后验证集的评价结果

模型	边界框			掩膜		
	P	R	$m_{AP}@0.5$	P	R	$m_{AP}@0.5$
改进前	0.95	0.88	0.93	0.96	0.87	0.93
改进后	0.96	0.95	0.98	0.96	0.92	0.96

改进前、后模型的混淆矩阵如图 6 所示。图 6 中,A、B、C、D、E 分别表示废钢类别为柱型、板型、结构件、灭火器、气瓶、背景。

比较图 6a 与图 6b 可知:改进后的模型正确预测类别的数量占比有明显提升,柱型、板型、结构件、灭火器、气瓶的分类精度分别为 95%、98%、95%、91%、99%,比原模型分别提高了 0%、2%、14%、11%、2%;全类别分类平均精度达 96.5%,比原模型提高了 5.8%。

废钢数据集在改进前、后的模型上训练,得出在验证集上的损失曲线,如图 7 所示。

图 7 中:横轴表示训练迭代次数;纵轴表示验证集上的损失值。从图 7 可以看出,改进后的模

型收敛更快,且改进模型的位置损失、掩膜损失、置信度损失、分类损失比原模型分别下降了 7.25%、5.80%、6.90%、92.17%。因此,改进模型在各项性能上取得了显著改善,这也使得模型能够更准确地定位目标位置,生成更精确的掩膜,并提供更可靠的置信度估计和分类结果。

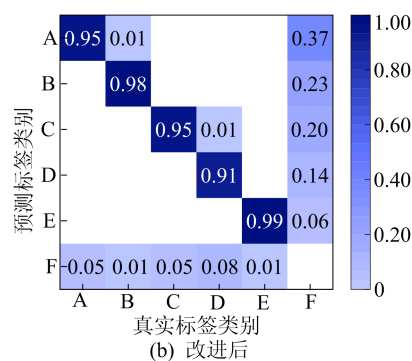
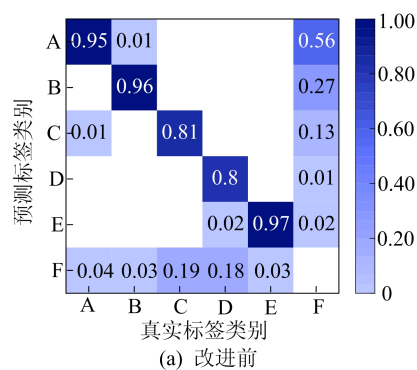


图 6 模型改进前、后的混淆矩阵

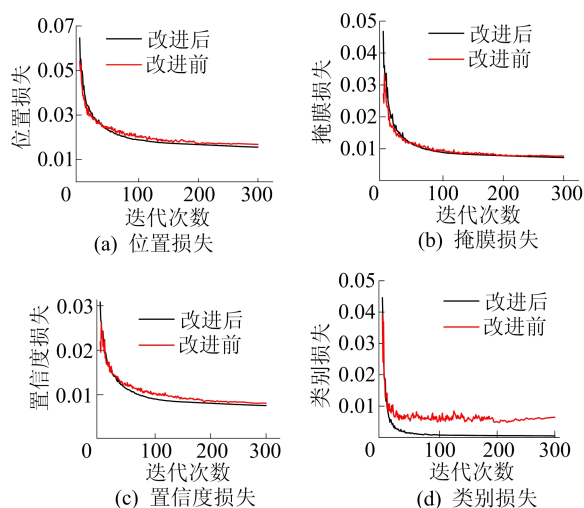


图 7 模型改进前、后的损失曲线

为了验证模型改进前、后对废钢预测的实际效果,对废钢测试集进行推理,部分推理结果如图 8 所示。从图 8 可以看出,改进后的模型与原模型相比可以检测出更多的目标,较好

地弥补了原模型存在的漏检、错检及分割精度低等不足。由此验证了本文建立的 YOLOv5s-

seg-CBAM 模型在废钢智能分类与实例分割应用中的可行性。



图 8 模型改进前、后废钢的预测效果图

3.6 本文模型与其他模型检测结果比较

为了证明本文改进的 YOLOv5s-seg 模型在废钢分类分割上的优越性,采用上述废钢数据集对其他经典的二阶段 Mask RCNN^[25]实例分割模型及一阶段 SOLO^[26]、YOLOACT^[27]实例分割模型进行训练,实验参数 batchsize 设置为 8,训练迭代次数同样为 300 次。4 种模型在测试集上的检测精度比较见表 3 所列。由于 SOLO 模型不生成边界框,因此边界框精度不参与比较。

表 3 不同模型检测结果比较

模型	掩膜		边界框	
	$m_{AP}@50$	R_A	$m_{AP}@50$	R_A
Mask RCNN	0.87	0.80	0.88	0.83
SOLO	0.86	0.77		
YOLOACT	0.89	0.81	0.91	0.82
改进模型	0.98	0.92	0.98	0.95

由表 3 可知,在对废钢的分类分割上,改进的 YOLOv5s-seg 模型比 Mask RCNN、SOLO、YO-

LOACT 模型具有更高的检测精度和召回率,废钢的分类效果优异。

4 结 论

针对目前钢铁回收料种类多、场景杂乱、人工分拣效率低、安全隐患大等问题,本文提出一种基于改进 YOLOv5s-seg 的废钢分类分割算法并构建了模型。通过线下数据增强算法提高模型的泛化能力,在 C3 模块中引入 CBAM 注意力机制,用 EIOU 替换原模型中的 CIOU 损失函数,提高模型的精准度并增加模型的鲁棒性。试验结果表明, YOLOv5s-seg-CBAM 模型在废钢识别中表现出更高的检测精度,改进后的模型对柱型、板型、结构件、灭火器、气瓶这 5 类废钢分类精度达到了 95%、98%、95%、91%、99%;边界框和掩膜的平均检测精度达到了 98%、96%,比原模型提高了 5%、3%;全类别分类平均精度高达 96.5%,比原模型提高了 5.8%。改进后的模型收敛更快,且改进模型的位置损失、掩膜损失、置信度损

失和分类损失比原模型分别下降了 7.25%、5.80%、6.90%、92.17%。在与其他经典实例分割模型比较中,改进 YOLOv5s-seg 模型也表现出优异的性能,得到更优异的检测效果。本文提出并建立的 YOLOv5s-seg-CBAM 算法及模型为实现废钢的智能化高效回收利用提供了技术支撑,具有较好的指导意义。

参 考 文 献

- [1] 上官方钦, 郦秀萍, 周继程, 等. 中国废钢资源发展战略研究[J]. 钢铁, 2020, 55(6): 8-14.
- [2] 崔志峰, 上官方钦, 王方杰, 等. 2022—2060 年中国废钢资源量分析预测[J]. 钢铁, 2023, 58(6): 126-133.
- [3] 卢宏涛, 罗沐昆. 基于深度学习的计算机视觉研究新进展[J]. 数据采集与处理, 2022, 37(2): 247-278.
- [4] 周婧, 秦伦明. 废钢资源发展现状及废钢分类研究[J]. 科技创新与应用, 2022, 12(11): 88-90.
- [5] 刘文仲. 中国钢铁工业智能制造现状及思考[J]. 中国冶金, 2020, 30(6): 1-7.
- [6] 曹家乐, 李亚利, 孙汉卿, 等. 基于深度学习的视觉目标检测技术综述[J]. 中国图象图形学报, 2022, 27(6): 1697-1722.
- [7] 卢健, 贾旭瑞, 周健, 等. 基于深度学习的三维点云分割综述[J]. 控制与决策, 2023, 38(3): 595-611.
- [8] DIAZ-ROMERO D J, EYNDE S V D, STERKENS W, et al. Simultaneous mass estimation and class classification of scrap metals using deep learning[J]. Resources, Conservation and Recycling, 2022, 181: 106272.
- [9] DIAZ-ROMERO D J, EYNDE S V D, ZAPLANA I, et al. Classification of aluminum scrap by laser induced breakdown spectroscopy (LIBS) and RGB+D image fusion using deep learning approaches[J]. Resources, Conservation and Recycling, 2023, 190: 106865.
- [10] 王成军, 薛玉珂, 杨超宇. 基于多尺度特征的 YOLOv5 铸件自动检测算法研究[J]. 制造技术与机床, 2023(10): 130-136.
- [11] LI Y F, QIN X P, ZHANG Z Y, et al. A robust identification method for nonferrous metal scraps based on deep learning and superpixel optimization[J]. Waste Management & Research, 2021, 39(4): 573-583.
- [12] DANG T L, CAO T, HOSHIBO Y. Classification of metal objects using deep neural networks in waste processing line[J]. International Journal of Innovative Computing, Information and Control, 2019, 15(5): 1901-1912.
- [13] CHEN S, HU Z L, WANG C, et al. Research on the process of small sample non-ferrous metal recognition and separation based on deep learning[J]. Waste Management, 2021, 126: 266-273.
- [14] 段盛晴. 基于光学图像 YOLO 算法的废钢识别分类与统计[D]. 大连: 大连理工大学, 2021.
- [15] 肖鹏程, 徐文广, 张妍, 等. 基于 SE 注意力机制的废钢分类评级方法[J]. 工程科学学报, 2023, 45(8): 1342-1352.
- [16] 周婧, 秦伦明. 基于 VA-Expo-WA 集成迁移学习的废钢分类[J]. 信息技术, 2022(5): 18-24.
- [17] WOO S, PARK J, LEE J, et al. CBAM: convolutional block attention module[C]//Proceedings of the European Conference on Computer Vision (ECCV). [S. l.]: ECVA, 2018: 3-19.
- [18] ZHENG Z, WANG P, LIU W, et al. Focal and efficient IOU loss for accurate bounding box regression[J]. Neurocomputing, 2022, 506: 146-157.
- [19] ZHENG Z H, WANG P, LIU W, et al. Distance-IoU loss: faster and better learning for bounding box regression[C]//Proceedings of the AAAI Conference on Artificial Intelligence. [S. l.]: AAAI, 2020, 34(7): 12993-13000.
- [20] JOCHER G, CHAURASIA A, STOKEN A, et al. Ultralytics/Yolov5:v7.0-Yolov5 sota realtime instance segmentation[J]. Zenodo, 2022. DOI:10.5281/ZENODO.4154370.
- [21] HE K, ZHANG X, REN S, 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.
- [22] LIN T Y, DOLLAR P, GIRSHICK R, et al. Feature pyramid networks for object detection[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. [S. l.]: IEEE, 2017: 2117-2125.
- [23] LIU S, QI L, QIN H, et al. Path aggregation network for instance segmentation[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. [S. l.]: IEEE, 2018: 8759-8768.
- [24] GOLNAZ G, CUI Y, SRINIVAS A, et al. Simple copy-paste is a strong data augmentation method for instance segmentation[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. [S. l.]: IEEE, 2021: 2918-2928.
- [25] HE K, GKIOXARI G, DOLLAR P, et al. Mask R-CNN[C]//Proceedings of the IEEE International Conference on Computer Vision. [S. l.]: IEEE, 2017: 2961-2969.
- [26] WANG X, KONG T, SHEN C, et al. Solo: segmenting objects by locations[C]//Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). [S. l.]: IEEE, 2020: 649-665.
- [27] BOLYA D, ZHOU C, XIAO F, et al. Yolact: real-time instance segmentation[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. [S. l.]: IEEE, 2019: 9157-9166.

(责任编辑 胡亚敏)