文章编号:1674-2974(2024)12-0067-11 DOI:10.16339/j.cnki.hdxbzkb.2024252

基于改进YOLOv8的热轧带钢表面缺陷检测方法

肖科[1†],杨昕宇[1],韩彦峰[1],宋斌[2]
[1. 重庆大学 机械与运载工程学院,重庆 400030;
2. 珞石(山东)机器人集团有限公司,山东 济宁 275312]

摘 要:针对目前热轧带钢表面缺陷检测精度低和效率低的问题,提出了一种基于改进YOLOv8s的目标检测算法.首先,提出了一种基于特征图二次拼接并融入GAM的SPPD模块,提升了模型多尺度信息融合能力.其次,提出了一种融合可变形卷积的特征提取模块DCN-block,以增大模型的感受野,提取完整的缺陷信息.最后,将特征融合网络中的C2f模块替换为BoT(bottleneck transformer)结构,将Transformer中的多头自注意力机制与卷积融合,提升模型的全局位置信息感知能力.实验结果表明,本文提出的算法在NEU-DET数据集上的平均精度均值(mAP)达到了80.5%,较原有的YOLOv8算法提升了5个百分点,同时检测速度达到了83帧/s, 满足实时检测的需求.

关键词:热轧带钢;表面缺陷;目标检测;深度学习
中图分类号:TP391.4 **文献标志码**:A

Surface Defect Detection Method for Hot-rolled Strip Steel Based on Improved YOLOv8

XIAO Ke[1†], YANG Xinyu[1], HAN Yanfeng[1], SONG Bin[2]
[1. College of Mechanical and Vehicle Engineering, Chongqing University, Chongqing 400030, China;
2. ROKAE(Shandong) Robot Group Co., Ltd., Jining 275312, China]

Abstract: A object detection algorithm based on improved YOLOv8s is proposed to address the issues of low accuracy and low efficiency in surface defect detection of hot-rolled strip steel. Firstly, an SPPD module based on feature map secondary stitching and incorporating GAM is proposed, which enhances the model's multi-scale information fusion ability. Secondly, a feature extraction module DCN-block that integrates deformable convolution is proposed to increase the receptive field of the model and extract complete defect information. Finally, the C2f module in the feature fusion network is replaced with a BoT (bottleneck transformer) structure, and the multi-head self-attention mechanism in the Transformer is fused with convolution to enhance the model's global position information perception ability. The experimental results show that the proposed algorithm achieves mean average precision (mAP) of 80.5% on the NEU-DET dataset, which is five percentage points higher than the original YOLOv8 algorithm. At the same time, the detection speed reaches 83 frames per second, meeting the requirements

* **收稿日期**:2023-09-03
基金项目:国家重点研发计划资助项目(2022YFB4702201), National Key Reaearch and Development Program of China (2022YFB4702201);国家自然科学基金资助项目(52375039), National Natural Science Foundation of China(52375039)
作者简介:肖科(1978—),男,重庆人,重庆大学副研究员,博士
† 通信联系人,E-mail: 470567110@qq.com

of real-time detection.

Key words: hot-rolled strip steel; surface defect; object detection; deep learning

钢铁作为我国的基础产业,极大地促进了经济和国防的发展[1].热轧带钢作为重要的钢铁产品,已广泛应用在建筑、化工和机械制造等领域.在实际生产过程中,受设备和工艺的影响,带钢不可避免地会产生表面缺陷,影响产品的外观.此外,热轧带钢常作为结构钢使用,其表面缺陷极易成为应力集中的薄弱部分,对产品的性能产生直接影响[2],存在严重的安全隐患.因此,准确、高效的缺陷检测方法对于带钢生产企业做好品控意义重大.

热轧带钢的表面缺陷检测技术从20世纪70年代初期开始发展,之前主要依赖人工目测发现缺陷,效率低且检测结果受主观因素的影响.后续出现的涡流检测、漏磁检测和红外检测[3],由于速度慢、可靠性低,未能得到大规模的应用.随着电荷耦合器件(CCD)相机的发展,基于传统机器视觉的检测方法相较于以前,在速度和精度上得到了极大的提升.然而,传统的机器学习算法需要该领域专家人工设计,难以提取深层次的图像特征信息[4],且泛化程度低,难以满足现代化钢铁生产企业的实际需求.

2012年之后,随着计算机硬件的发展,深度学习在图像处理领域迅速崛起[5],基于卷积神经网络的热轧带钢缺陷检测方法可以自动提取缺陷的特征信息[6],在检测精度和速度上都已大幅优于传统的机器学习算法.结合深度学习的带钢缺陷检测方法可大致分为基于回归思想的单阶段(one-stage)检测算法和基于区域建议的两阶段(two-stage)检测算法[7].以YOLO[8]系列、SSD[9]和RetinaNet[10]等为主的单阶段检测算法检测速度快但精度略低,以RCNN[11]系列为代表的两阶段检测算法需要先生成候选框,因此检测速度慢但精度略高.目前,关于深度学习算法在热轧带钢表面缺陷检测中的应用,国内外学者已有较多相关研究.邹旺等[12]提出了一种基于YOLOv4-tiny[13]的带钢缺陷检测算法,它结合了多尺度检测与空间注意力机制,在NEU-DET数据集上的检测精度为73.29%,虽然检测速度较快,但精度较低.Zhou等[14]提出了一种残差空间金字塔池化模块(RASPP),它与原特征金字塔池化模块相加构成了双金字塔网络(DPN),加入Faster RCNN后在NED-DET上实验的平均精度为80.93%,但检测速度仅有18.62帧/s. Yu等[15]提出了基于FCOS改进而得的CABF-FCOS算法,用双向特征融合网络代替特征金字塔模块,减少了特征信息的丢失,并在网络中添加了GAM,改进后的网络在NEU-DET数据集上的精度达到了76.88%,检测精度还需提升,但模型参数量较大,难以实现在小型设备上的部署.

综上所述,目前相关缺陷检测算法难以同时满足实际生产环境对检测速度和精度的要求.针对上述问题,本文基于现阶段较先进的YOLOv8缺陷检测网络进行了改进.其主要贡献如下:

1) 提出了一种空间金字塔池化特征图二次拼接(spatial pyramid pooling-double, SPPD)模块,该模块在SPPF的基础上添加卷积分支和GAM,以丰富模型梯度流并增强模型的特征提取能力.

2) 提出了一种融合可变形卷积的特征提取模块DCN-block,将可变形卷积融入Bottleneck,使其在采样时可以自适应物体的形状和尺寸,丰富了模型的感受野,能提取更完整的缺陷信息,训练出来的模型更具有鲁棒性.

3) 引入了BoT(bottleneck transformer)模块替代C2f模块.以卷积为特征提取方式的C2f模块具有平移不变性和局部性,为提升模型的全局感知能力,在特征融合网络PAFPN中使用BoT模块可以更好地提取检测目标的位置和特征信息.

1 基于YOLOv8的改进算法

1.1 模型结构

结合实际生产环境对于模型检测速度和精度的要求,本文基于YOLOv8s模型,提出了改进后的算法架构,如图1所示.YOLOv8目标检测算法是在YOLOv5[16]的基础上改进而得,由骨干网络(backbone)、颈部网络(neck)和结果输出头(head)构成.本文首先将SPPF模块替换为基于特征图二次拼接并融入

GAM 的 SPPD 模块;其次将 Backbone 里面 C2f 中的 Bottleneck 替换为融合可变形的卷积的 DCN-block;最后将 Neck 里面的 C2f 模块替换为融合多头自注意力机制的 BoT 模块,以增强模型的全局感知能力.优化后的网络模型可以实现更准确的缺陷分类和定位,提升了模型的检测性能.

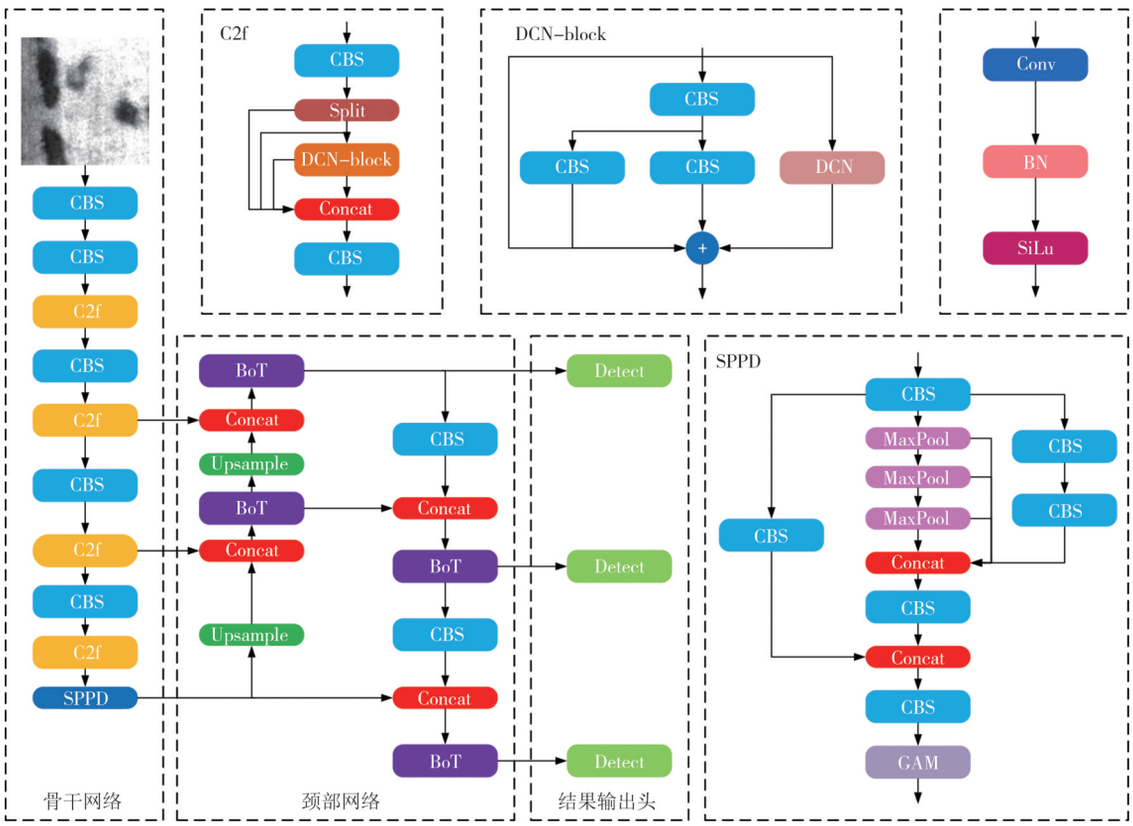


图 1　改进后的 YOLOv8s 网络结构

Fig.1　Improved YOLOv8s network structure

1.2　SPPD

YOLOv8 沿用了 YOLOv5 的快速空间金字塔池化(spatial pyramid pooling-fast, SPPF)模块,SPPF 结构如图 2 所示.输入特征图首先经过 1×1 卷积降维减少参数量,而后通过池化核大小为 5×5 的最大池化操作来融合多尺度特征图信息,最后经 1×1 卷积升维,保持维度的一致性.

热轧带钢数据集有大量的细小缺陷,且存在前景和背景相似的问题,为丰富模型的梯度流,更好地实现对缺陷信息全局和局部特征的融合,本文提出了基于特征图二次拼接并融合 GAM[17]的 SPPD 模块.首先,在原始输入特征图的基础上引入 3×3 卷积分支提取特征,经 1×1 卷积操作后与最大池化后特征图进行拼接.其次,额外加入输入特征图的 1×1 卷积分支,进行二次拼接操作,在不改变图像空间结构的情况下,提升了模型的非线性表达能力.最后,通过 1×1 卷积调整通道数,使空间维度对齐后加入 GAM,

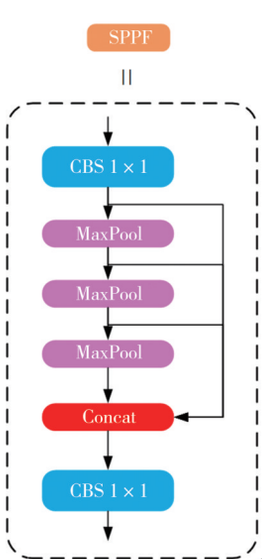


图 2　SPPF 模块

Fig.2　SPPF module

改进后的 SPPD 模块如图 3 所示.SPPD 模块使用类残差结构,增加了网络的深度,融入的 GAM 注意力机制更有利于网络关注缺陷部位.

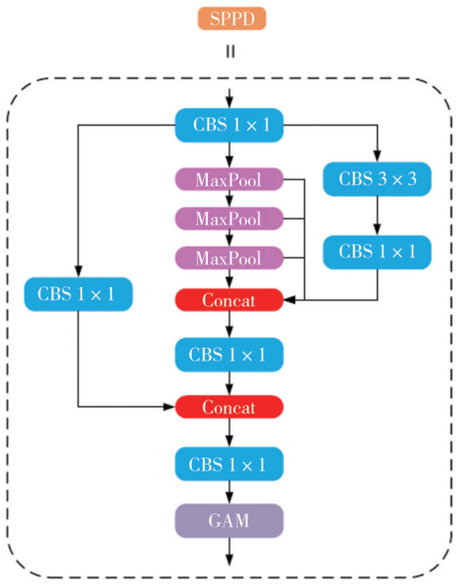


图 3 SPPD 模块

Fig.3 SPPD module

其中,GAM 的结构如图 4 所示,主要由通道注意力和空间注意力组成.相较于以往的注意力机制而言,GAM 可以更好地保留通道和空间维度之间的信息,从而增强跨纬度交互能力.对给定输入特征图 $F_1 \in R^{C\times H\times W}$,其中间状态 F_2 和输出特征图 F_3 可由如下公式计算而得:

$$F_2 = M_c(F_1) \otimes F_1 \tag{1}$$

$$F_3 = M_s(F_2) \otimes F_2 \tag{2}$$

式中:M_c 和 M_s 分别为通道和空间注意图;⊗表示逐元素相乘.

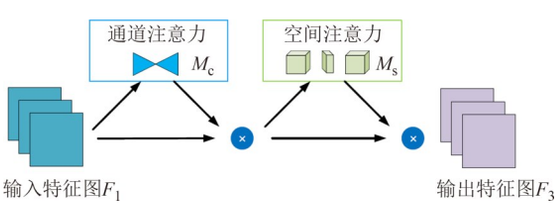


图 4 GAM

Fig.4 GAM

其中,通道注意力子模块如图 5 所示,先将输入的特征图 $C\times W\times H$ 通道重排为 $W\times H\times C$,接着输入一个两层的多层感知机(MLP)实现跨维通道信息的融合,最后将维度还原后经过 Sigmoid 函数输出通道注意力图.

空间注意力子模块如图 6 所示,使用两个 7×7 大小的卷积核融合空间信息,同时为减少计算量,先将通道特征维度由 C 降维到 C/r,再通过卷积还原通道数,最后经过 Sigmoid 激活函数输出空间注意力图.

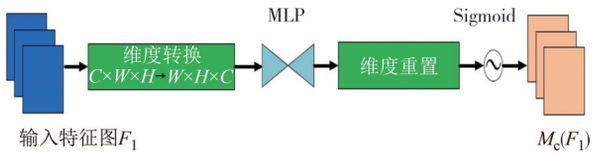


图 5 通道注意力子模块

Fig.5 Channel attention submodule

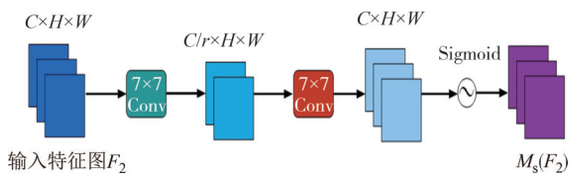
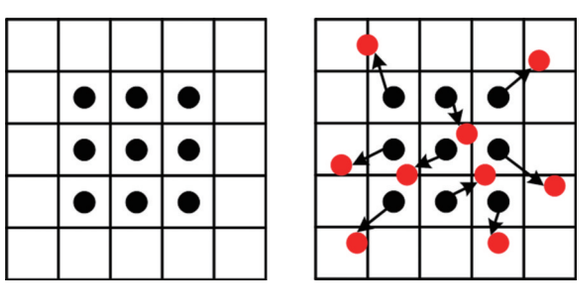


图 6 空间注意力子模块

Fig.6 Spatial attention submodule

1.3 DCN-block

带钢的表面缺陷呈现出不规则的形态,传统的卷积结构是固定的,限制了特征提取过程中感受野的大小,给缺陷检测带来了困难.Dai 等[18]提出了一种可变形卷积,通过对标准卷积施加可学习的偏移量来扩大卷积核的采样区域,使检测网络可以在动态学习的过程中拟合完整的缺陷信息.此外,可变形卷积增大了模型的感受野,让网络能更好地学习远程空间关系,提升模型的鲁棒性.标准卷积和可变形卷积的采样过程如图 7 所示,从中可以看出可变形卷积能够适应检测目标形状的变化.



(a)标准卷积 (b)可变形卷积

图 7 标准卷积和可变形卷积采样对比

Fig.7 Comparison of standard convolutional and deformable convolutional sampling

二维可变形卷积采样过程可简述如下:先用常规卷积核 R 对于输入特征图 x 进行采样,再对采样值施加权重 w,最后依次相加求和.以 3×3 大小的卷积核为例,可定义为

$$R = \{(-1,-1),(-1,0),\cdots,(0,1),(1,1)\}$$

在常规卷积的采样过程中,对于输出特征图 y 上的每个位置 p_0,有

$$y(p_0) = \sum_{p_n \in R} w(p_n)\cdot x(p_0 + p_n) \tag{3}$$

式中:p_n 为 R 中的采样点.

在可变形卷积中,对常规卷积核施以偏移量 Δp_n,其中 $\{\Delta p_n | n = 1, 2, \cdots, N\}$, $N = |R|$. 因此,可变形卷积的公式可表达为

$$y(p_0) = \sum_{p_n \in R} w(p_n) \cdot x(p_0 + p_n + \Delta p_n) \tag{4}$$

本文对原有的 Bottleneck 模块进行改进,提出了一种结合可变形卷积的 DCN-block 模块,其结构对比如图 8 所示. YOLOv8 的 Bottleneck 模块采用两次 3×3 卷积提取缺陷信息,并添加残差连接防止梯度消失. 本文提出的 DCN-block 在此基础上在初次 3×3 卷积操作后添加 1×1 卷积分支,并额外增加 3×3 可变形卷积分支,提升模型的感受野和采样能力. DCN-block 较原 Bottleneck 模块,增强了模型对不规则缺陷的提取能力,使得网络可以获取更加完整的缺陷特征信息,从而提升模型的检测性能.

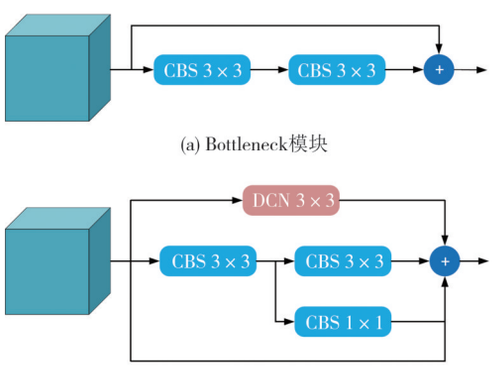


图 8　Bottleneck 和 DCN-block 结构对比

Fig.8　Comparison of Bottleneck and DCN-block structure

1.4　BoT

热轧带钢缺陷的背景复杂,且同类缺陷也存在形状差异,容易导致模型误检和漏检. YOLOv8 采用的 C2f 特征提取模块由卷积构成,对于特征层的感知受限于局部,缺乏全局和远程感知能力,为提升模型的特征提取能力,本文引入 BoT[19] 模块.

BoT 模块由 ResNet 中的 Bottleneck 模块改进而来,其结构分别如图 9 所示. 其核心思想是将 Transformer 与 CNN 融合[20],相较于 Bottleneck 模块,BoT 模块将原本的 3×3 卷积替换为多头自注意力(multi-head self-attention, MHSA)机制,融合了 MHSA 的全局关注和 CNN 的局部感知能力,通过结合全局和局部的热轧带钢缺陷信息,能够解决卷积过程中缺陷信息的丢失问题,提升了网络对于小缺陷的检出率. 考虑到颈部网络的特征图分辨率较低,可以节约计算资源且深层特征图蕴含丰富的语义信息,本文将其中的 C2f 模块替换为 BoT 模块,可以捕获更丰富的全局特征,提升模型区分缺陷和背景信息的能力.

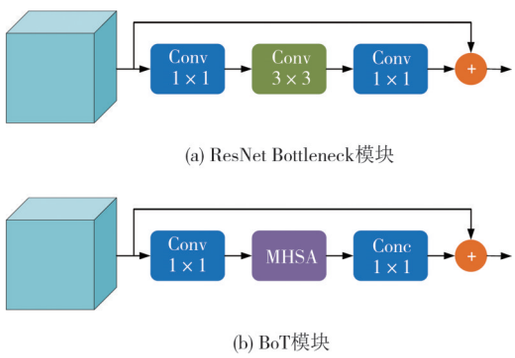


图 9　BoT 和 ResNet Bottleneck 结构对比

Fig.9　Comparison of BoT and ResNet Bottleneck structure

BoT 模块实际使用 4 个头的多头自注意力机制,多头自注意力机制的输出分别由单头自注意力机制计算的结果矩阵拼接而得,本文将其简化绘制的多头自注意力机制结构如图 10 所示. 其中,输入 x 代表高、宽和通道数分别为 H、W、d 的特征图,$\boldsymbol{q}$、$\boldsymbol{k}$、$\boldsymbol{v}$ 分别表示查询编码、键编码和值编码,由 x 与不同的权值矩阵 $\boldsymbol{W}_Q$、$\boldsymbol{W}_K$ 和 $\boldsymbol{W}_V$ 相乘而得. 图片的位置编码为 $\boldsymbol{r}$,由高度和宽度的相对位置编码 $\boldsymbol{R}_h$ 和 $\boldsymbol{R}_w$ 相加而得. 综上,多头自注意力机制输出的计算公式为

$$\text{Attention}(\boldsymbol{q}, \boldsymbol{k}, \boldsymbol{v}, \boldsymbol{r}) = \text{Softmax}(\boldsymbol{q}\boldsymbol{k}^{\mathrm{T}} + \boldsymbol{q}\boldsymbol{r}^{\mathrm{T}})\boldsymbol{v} \tag{5}$$

式中:$\boldsymbol{q} = \boldsymbol{W}_Q x$, $\boldsymbol{k} = \boldsymbol{W}_K x$, $\boldsymbol{v} = \boldsymbol{W}_V x$, $\boldsymbol{r} = \boldsymbol{R}_h + \boldsymbol{R}_w$.

2　实验结果与分析

2.1　实验数据集及环境设置

本文采用东北大学的 NEU-DET 热轧带钢缺陷数据集作为实验对象,包括六类典型的热轧带钢表面缺陷,即开裂(crazing, Cr)、夹杂(inclusion, In)、斑块(patches, Pa)、点蚀(pitted surface, PS)、氧化铁皮压入(rolled-in scale, RS)和划痕(scratches, Sc),每类缺陷的图片都有 300 张,共 1 800 张,图像的原始分辨率为 200 像素×200 像素,其缺陷样本示例如图 11 所示.

实验的操作系统环境为 Ubuntu 20.04.1,计算机的硬件配置为 Inter(R) Corei9-10900x CPU, NVIDIA GeForce GTX3090 GPU,内存为 32 GB. 所有实验的软件环境基于 PyTorch1.10.1, Python 版本为 3.7,并利用 CUDA 11.4 和 cuDNN8.5 实现 GPU 训练加速.

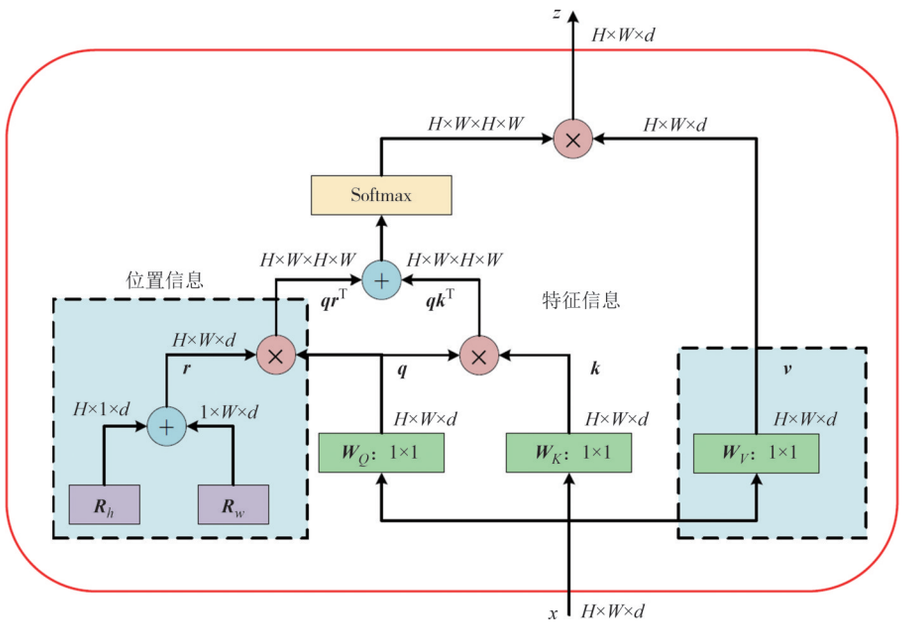


图 10 MHSA 结构

Fig.10 MHSA structure

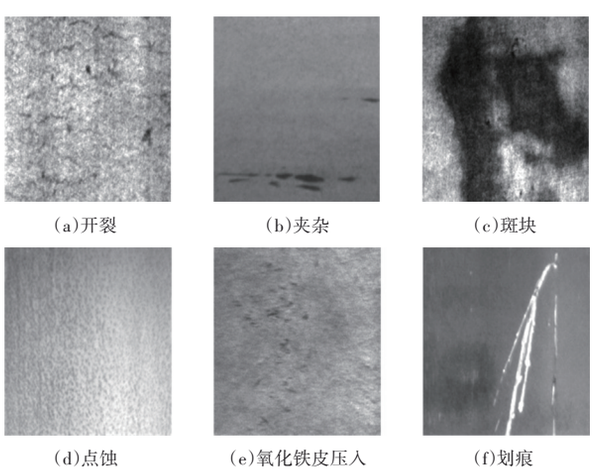


图 11 NEU-DET 六类带钢缺陷示例

Fig.11 Examples of six types of strip steel defects in NEU-DET

实验时将 NEU-DET 数据集的 1 800 张图片按照 8∶2 的比例随机划分，得到训练集 1 440 张，测试集 360 张. 采用 SGD 优化器，设置初始学习率为 0.01，动量系数为 0.937，权重衰减为 0.000 5，训练时 batch size 设置为 16，共训练 150 个 epoch.

2.2 模型评价指标

本文采用目标检测任务中常用的平均精度均值(mean average precision，mAP)作为模型检测精度的评价指标，可由精确率(P)和召回率(R)计算得到. 精确率的定义为模型预测为正样本中实际为正样本的比例，召回率是指模型预测正确的正样本数与总的正样本数之比，计算公式分别如下所示：

$$P = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}} \tag{6}$$

$$R = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} \tag{7}$$

式中：TP(true positive)为将正样本预测为正样本的数量；FP(false positive)为将负样本预测为正样本的数量；FN(true negative)为将正样本预测为负样本的数量.

平均精度(average precision，AP)的定义为不同召回率下精确率的平均值，常用于评价单类别的检测精度. mAP 则通过计算所有类别的平均精度均值来评估多目标检测任务模型的整体性能. 对 n 类缺陷检测而言，可通过如下公式计算 AP 和 mAP 值：

$$\mathrm{AP} = \int_0^1 P(R)\mathrm{d}R \tag{8}$$

$$\mathrm{mAP} = \frac{\sum_{i=0}^{n} \mathrm{AP}(i)}{n} \tag{9}$$

模型对图片的处理过程主要包括图像预处理(pretreatment)、模型推理(inference)和非极大值抑制(NMS). 为衡量网络模型对图像的处理速度，本文引入 FPS(frames per second)指标来评价模型的实时检测能力. FPS 的定义为模型每秒钟分析处理的图像帧数，可由如下公式计算而得：

$$\mathrm{FPS} = \frac{1\,000}{t_{\mathrm{Pretreatment}} + t_{\mathrm{Inference}} + t_{\mathrm{NMS}}} \tag{10}$$

2.3 实验结果

为探究模型大小对检测精度的影响,本文基于上述实验环境,对YOLOv8的5个检测模型YOLOv8n~YOLOv8x在NEU-DET热轧带钢表面缺陷数据集上进行了实验,实验结果如表1所示.由结果可知,随着网络模型结构的深度和宽度的依次增大,网络的检测速度呈依次递减的趋势,检测的mAP则呈现先上升后下降的趋势.这是因为在一定范围内随着网络结构的加深,模型可以学习到更为丰富的特征信息,从而提升了模型的检测能力;如果继续加深模型,可能会造成浅层特征信息的丢失,使模型退化,导致检测精度反而下降.随着模型的依次增大,YOLOv8l的mAP达到最大值,为77.50%,但检测速度下降到了70帧/s,继续使用最深的YOLOv8x检测模型时,网络的mAP开始下降到76.33%,检测速度下降到最低,为61帧/s.综合考虑检测速度和精度,因为YOLOv8s在保证较高的mAP的同时有较快的检测速度,所以本文选取YOLOv8s作为基础模型进行后续实验.

表1 YOLOv8各模型的实验结果

Tab.1 Experimental results of YOLOv8 models

模型	AP/%						mAP/%	FPS/(帧·s^{-1})
	Cr	In	Pa	PS	RS	Sc		
YOLOv8n	40	82	92	76	63	95	74.67	116
YOLOv8s	45	81	91	78	64	94	75.50	106
YOLOv8m	47	83	92	80	64	94	76.67	85
YOLOv8l	50	84	94	73	67	97	77.50	70
YOLOv8x	47	83	90	76	66	96	76.33	61

为证明本文提出的SPPD模块、DCN-block模块和引入BoT模块对模型性能提升的有效性,通过消融实验来直观展现模型精度变化情况,便于后续评估各项改进点对模型检测能力的影响.本实验以YOLOv8s为基准算法,实验结果如表2所示.

表2 消融实验结果

Tab.2 Results of ablation experiment

模型	AP/%						mAP/%	FPS/(帧·s^{-1})
	Cr	In	Pa	PS	RS	Sc		
YOLOv8s	45	81	91	78	64	94	75.50	106
YOLOv8s+SPPD	45	82	93	80	64	96	76.67	102
YOLOv8s+SPPD+DCN-block	53	84	93	81	68	93	78.67	88
YOLOv8s+SPPD+DCN-block+BoT	58	87	91	81	72	94	80.50	83

根据表2可知,原始的YOLOv8s模型的mAP为75.50%,将SPPF模块替换为SPPD模块后,模型的mAP提升了1.17个百分点,证明了基于特征图二次拼接并融入GAM后可以使网络更关注缺陷信息,从而提升检测精度,且相较于基准模型,检测速度只出现了轻微下降.进一步加入融入可变形卷积的DCN模块后,mAP提升最多,上升了2个百分点,证明了DCN-block模块可以更好拟合热轧带钢表面缺陷形状,显著增强模型对于不规则缺陷的特征提取能力.最后在模型的颈部引入BoT模块后,mAP提升了1.83个百分点,证明了BoT模块通过结合全局和局部的热轧带钢缺陷信息,提升了模型对缺陷检出的正确率.通过上述方式改进后的YOLOv8模型的mAP较基线模型总共提升了5个百分点,检测速度达到了83帧/s,满足实时检测的要求.

此外,为探究各改进模型对不同大小缺陷的检测效果,分别列出小、中、大目标的检测结果,如表3所示.整体来看,使用提出的SPPD、DCN-block、BoT模块对不同大小目标的平均检测精度均有提升,这也证明提出的各改进模块的有效性.其中,使用BoT对小目标的平均检测精度提升最为明显,较基础模型提升了11.2个百分点,尽管对大目标的平均检测精度略微下降0.1个百分点,但模型的总体精度达到最大值.

为体现本文所提出的SPPD模块里融合的GAM对模型精度提升的优越性,以改进后(即引入SPPD、

表3 不同改进方法的mAP@0.5:0.95结果

Tab.3 Results of mAP@0.5:0.95 for different methods

模型	mAP/%			
	小目标	中目标	大目标	平均
YOLOv8s	32.30	36.90	53.60	41.7
YOLOv8s+SPPD	33.90	37.30	53.60	42.6
YOLOv8s+SPPD+DCN-block	37.80	38.80	54.30	42.8
YOLOv8s+SPPD+DCN-block+BoT	43.50	39.10	53.50	44.1

DCN-block 和 BoT 模块)的 YOLOv8s 网络为基准,并分别将 SPPD 模块的 GAM 替换为目前主流的注意力机制 SE[21]、CBAM[22]和 CA[23],最后在前文划分好的 NEU-DET 热轧带钢表面缺陷数据集上进行验证并记录模型平均精度均值的变化.实验结果如表 4 所示,分析表中数据可知,将 GAM 替换为其他注意力机制后,模型的 mAP 均出现了不同程度的下降,因此将 GAM 融入 SPPD 模块更适用于改进后的模型.

表 4 引入不同注意力机制的 SPPD 模块对比试验

Tab.4 Comparison experiment of introducing different attention mechanisms into SPPD modules

模型	AP%						mAP/%	FPS/(帧·s^{-1})
	Cr	In	Pa	PS	RS	Sc		
YOLOv8s+DCN-block+BoT+SPPD(SE)	56	85	90	82	68	94	79.17	86
YOLOv8s+DCN-block+BoT+SPPD(CBAM)	49	84	89	78	74	93	77.83	84
YOLOv8s+DCN-block+BoT+SPPD(CA)	54	85	91	77	70	92	78.17	80
YOLOv8s+DCN-block+BoT+SPPD(GAM)	58	87	91	81	72	94	80.50	83

为直观呈现 GAM 对网络关注区域的影响,本文绘制了特征提取网络的注意力热力图,如图 12 所示.通过对比分析热力图中信息后可得,原算法对于热轧带钢缺陷部分的关注域较小,不利于网络获取完整的缺陷信息,而融入 GAM 的 SPPD 模块提升了模型的特征提取能力,在抑制非缺陷区域的同时,让网络更加集中关注目标区域.从而验证了 SPPD 模块可以在背景复杂的环境下增强网络对目标区域的关注能力,能够充分提取和利用缺陷部分的特征和位置信息,进一步提升网络的检测精度.

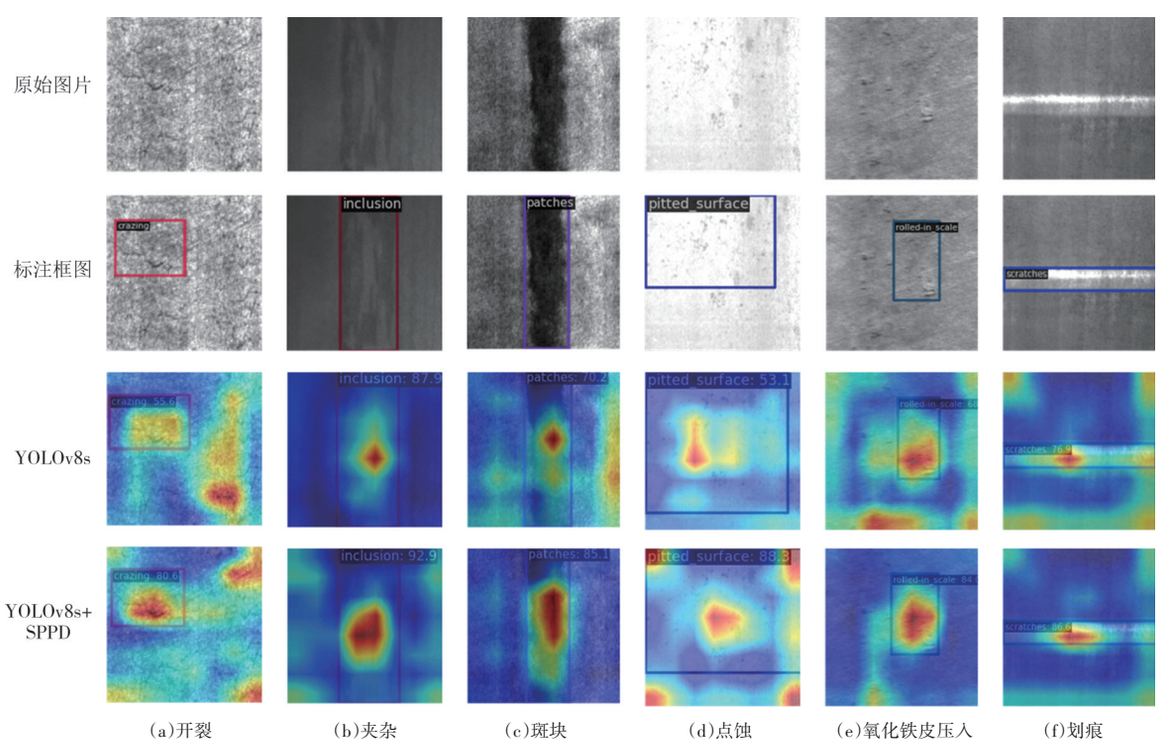


图 12 加入 SPPD 热力图对比结果

Fig.12 Heat map comparison results after adding SPPD

2.4 算法对比实验

为进一步验证改进 YOLOv8s 算法的性能,本文将其与目标检测常用的主流算法,如 SSD、RetinaNet、YOLOv5、YOLOX[24]、YOLOv7[25]和 Faster RCNN 进行对比试验,训练轮次都设定为 150 个 epoch,并保持实验环境一致,记录得到的实验结果如图 13 所示,具体数据如表 5 所示.

由图 13 可知,本文提出的基于 YOLOv8s 的改进算法相较于现有算法更好地平衡了检测速度与精度.由表 5 数据可得,改进的 YOLOv8s 算法的 mAP 达到了 80.50%,相较于原始的 YOLOv8s 算法,mAP 提升了 5 个百分点;FPS 则由改进前的 106 帧/s 下降到

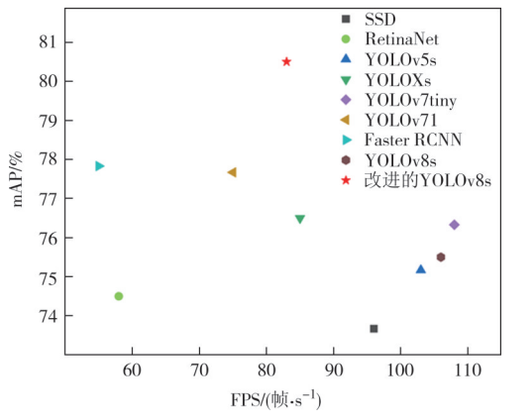


图 13　不同模型精度-速度对比图

Fig.13　Comparison chart of precision-speed for different models

表 5　改进 YOLOv8s 与其他算法实验结果对比

Tab.5　Comparison of experimental results between improved YOLOv8s and other algorithms

模型	AP/%						mAP/%	FPS/(帧·s⁻¹)
	Cr	In	Pa	PS	RS	Sc		
SSD	60	70	89	74	72	77	73.67	96
RetinaNet	55	73	94	84	72	69	74.50	58
YOLOv5s	34	83	94	74	73	93	75.17	103
YOLOXs	38	84	92	78	75	92	76.50	85
YOLOv7tiny	46	83	95	74	66	94	76.33	108
YOLOv7l	47	86	92	80	67	94	77.67	75
Faster RCNN	49	82	93	85	65	93	77.83	55
YOLOv8s	45	81	91	78	64	94	75.50	106
改进的YOLOv8s	58	87	91	81	72	94	80.50	83

83 帧/s,但仍满足实时检测的要求.其中,Faster RCNN 相较于 SSD、RetinaNet 及 YOLO 系列的目标检测算法,在检测精度上达到了最大值,mAP 为 77.83%.这也充分证明了两阶段算法较单阶段算法在检测精度上存在显著优势,但由于双阶段算法需要先生成候选区域再检测,因此Faster RCNN在检测速度上仅有55帧/s,远低于其他的单阶段检测算法.

对比实验结果可知,不同模型对不同种类缺陷的检测精度各有优劣.其中RetinaNet对斑块(Pa)类大目标缺陷有着较好的检测效果,但对于背景复杂的划痕(Sc)类细小目标缺陷的检测效果最差.原因在于RetinaNet在特征提取的卷积过程中,小目标特征信息逐渐退化消失且易受到背景噪声的干扰.与此类似的是SSD算法没有真正将高维和低维特征图融合,导致没有充分利用低维特征图的小目标缺陷信息,同样导致SSD算法对于划痕(Sc)类缺陷的检出效果不佳.相较于改进前的YOLOv8s算法,改进后算法的mAP有明显提升,特别是对于检出率最低的开裂(Cr)类缺陷及背景复杂的氧化铁皮压入(RS)类缺陷的检测精度有着大幅提升,证明了改进后的网络有着更强的缺陷特征信息的提取和融合能力.

2.5　模型检测效果可视化与分析

为直观呈现基于YOLOv8s改进后算法相较于原算法的优越性,本文将改进前后模型在热轧带钢缺陷数据集上的检测效果可视化,并与原始图片和真实的标注框图进行对比,得到的结果如图14所示,并列出改进前后模型对各缺陷类别定位的平均IoU,如表6所示.

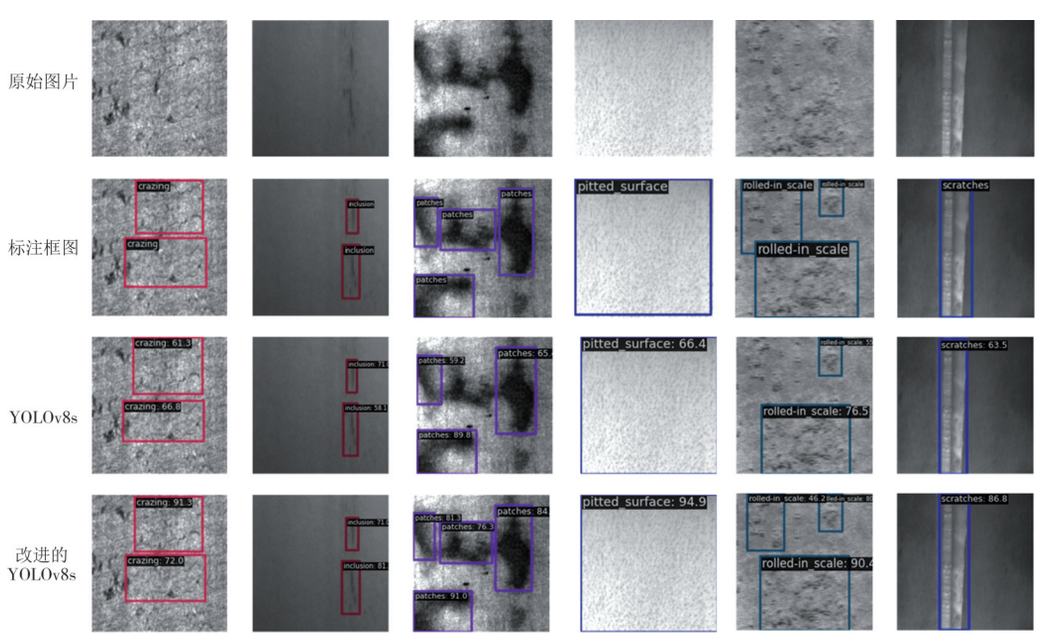


图 14　改进前后模型检测效果可视化

Fig.14　Visualization of model detection effects before and after improvement

表6 改进前后IoU@0.5结果

Tab.6 Results of IoU@0.5 before and after improvement

模型	IoU					
	Cr	In	Pa	PS	RS	Sc
YOLOv8s	0.72	0.82	0.85	0.86	0.74	0.83
改进YOLOv8s	0.79	0.85	0.89	0.87	0.80	0.88

从缺陷检出率的角度分析,改进前的YOLOv8s算法在斑块(Pa)和氧化铁皮压入(RS)两类缺陷中均存在漏检的情况,而改进后的算法则将所有缺陷都检测了出来.从分类置信度的角度分析,改进后的YOLOv8s算法较改进前而言,对各缺陷的分类置信度都有提高.从缺陷定位的准确性分析,由表6可知改进后的YOLOv8s算法较改进前,各类缺陷的IoU均有提升,表明了模型对缺陷的定位精度更高.从图14的检测结果也可以看出,改进后的YOLOv8s算法对于开裂(Cr)、夹杂(In)、斑块(Pa)和划痕(Sc)类缺陷的定位也都更接近真实的标注框图.

2.6 缺陷数据集分割实验

YOLOv8作为视觉多任务集成框架,可对图像进行检测、分割、分类和姿态估计.在实现热轧带钢的缺陷检测后,利用YOLOv8s-seg算法实现对NEU-DET热轧带钢数据集的缺陷分割任务.YOLOv8s-seg是YOLOv8的实例分割模型,网络设计基于YOLACT实现,主要由掩码预测分支和目标检测分支并行组成.掩码预测分支使用全卷积网络结构生成实例的原型掩码,目标检测分支可实现对目标的位置、类别和掩码系数的预测,通过快速非极大值抑制后,将2个并行分支计算得到的原型掩码和掩码系数进行线性组合得到最终的分割结果.

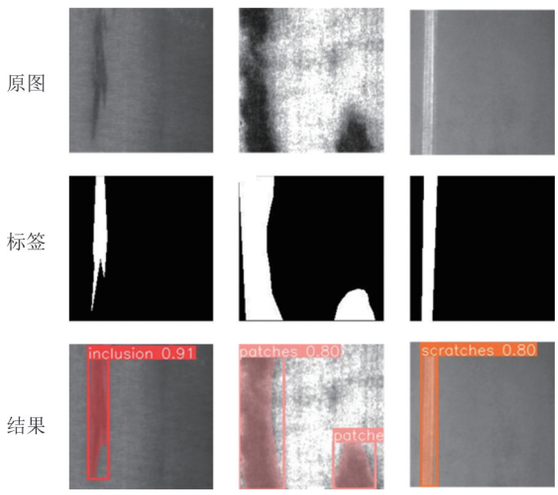
东北大学的NEU-DET热轧带钢缺陷数据集缺乏分割标签,考虑到开裂(Cr)、点蚀(PS)和氧化铁皮压入(RS)类缺陷难以实现准确地分割标注,借助开源标注工具LabelMe对夹杂(In)、斑块(Pa)和划痕(Sc)三类缺陷进行标注,共得到900张图片,按照8:2的比例随机划分,其中训练集720张,测试集180张.

本文以分割任务的平均精度(AP)、平均精度均值(mAP)和平均交并比(mIoU)作为实例分割掩模的评价指标,基于前文实验环境进行训练,得到的实验结果如表7所示,分割结果如图15所示.

表7 分割结果

Tab.7 Results of segmentation

模型	AP/%			mAP/%	mIoU
	In	Pa	Sc		
YOLOv8s	63.60	74.70	76.20	71.50	0.58



(a)开裂 (b)夹杂 (c)斑块

图15 缺陷分割结果

Fig.15 Segmentation results of defects

3 结 论

本文针对热轧带钢表面缺陷检测任务,为进一步提升缺陷的检测精度和速度,提出了一种基于YOLOv8s的改进算法.首先,提出了基于特征图二次拼接并融入GAM的SPPD模块,在丰富模型梯度流的同时可以使得模型更关注缺陷位置.其次,针对热轧带钢表面缺陷形态不规则和常规卷积存在感受野受限的问题,对原有的Bottleneck模块进行改进,提出了一种结合可变形卷积的特征提取模块DCN-block,从而使检测网络可以动态学习到完整的缺陷信息.最后,热轧带钢数据集背景复杂且差异较大,全卷积的C2f特征提取模块缺乏全局和远程感知能力,本文通过在模型颈部引入BoT模块来捕获更丰富的全局特征,提升了模型区分缺陷和背景信息的能力.实验结果显示,本文所提出的YOLOv8s改进模型在NEU-DET测试集上的mAP达到了80.5%,较原模型提升了5个百分点,且检测速度达到了83帧/s.综上所述,本文所提出的改进算法较好地平衡了检测精度和速度,更符合热轧带钢生产企业的实际需求.

参考文献

[1] 李毅仁,邢奕,孙宇佳,等. 钢铁工业低碳绿色发展路径与实践[J]. 工程科学学报,2023,45(9):1583-1591.
LI Y R,XING Y,SUN Y J,et al. Low-carbon green development path and practice of the iron and steel industry[J]. Chinese

Journal of Engineering,2023,45(9):1583-1591.(in Chinese)

[2] 周李洪,龚金科,李兵. 基于稀疏表示的车用带钢表面图像信息修复[J]. 湖南大学学报(自然科学版),2021,48(8):141-148.
ZHOU L H,GONG J K,LI B. Image information restoration of automotive strip steel surface based on sparse representation[J]. Journal of Hunan University (Natural Sciences),2021,48(8):141-148.(in Chinese)

[3] 李桂元.无损检测技术在钢棒探伤方面的应用[J].冶金与材料,2022,42(1):89-90.

[4] 慕君林,马博,王云飞,等.基于深度学习的农作物病虫害检测算法综述[J].农业机械学报,2023,54(增刊2):301-313.
MU J L,MA B,WANG Y F,et al. Review of crop disease and pest detection algorithms based on deep learning [J]. Transactions of the Chinese Society for Agricultural Machinery,2023,54(Sup.2):301-313.(in Chinese)

[5] 赵立新,邢润哲,白银光,等. 深度学习在目标检测的研究综述[J]. 科学技术与工程,2021,21(30):12787-12795.
ZHAO L X,XING R Z,BAI Y G,et al. Review on survey of deep learning in target detection [J]. Science Technology and Engineering,2021,21(30):12787-12795.(in Chinese)

[6] 张睿,高美蓉,傅留虎,等. 基于多域多尺度深度特征自适应融合的焊缝缺陷检测研究[J]. 振动与冲击,2023,42(17):294-305.
ZHANG R,GAO M R,FU L H,et al. Weld defect detection based on adaptive fusion of multi-domain and multi-scale deep features[J]. Journal of Vibration and Shock,2023,42(17):294-305.(in Chinese)

[7] 付苗苗,邓森磊,张德贤. 深度神经网络图像目标检测算法综述[J]. 计算机系统应用,2022,31(7):35-45.
FU M M,DENG M L,ZHANG D X. Survey on deep neural network image target detection algorithms[J]. Computer Systems and Applications,2022,31(7):35-45.(in Chinese)

[8] REDMON J,DIVVALA S,GIRSHICK R,et al. You only look once: unified, real-time object detection [C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas,NV,USA. IEEE,2016:779-788.

[9] LIU W,ANGUELOV D,ERHAN D,et al. SSD: single shot MultiBox detector [M]//Lecture Notes in Computer Science. Cham:Springer International Publishing,2016:21-37.

[10] LIN T Y,GOYAL P,GIRSHICK R,et al. Focal loss for dense object detection [C]//2017 IEEE International Conference on Computer Vision (ICCV). Venice,Italy. IEEE,2017:2999-3007.

[11] 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. Columbus,OH,USA. IEEE,2014:580-587.

[12] 邹旺,吉畅. 一种改进YOLOv4-tiny的带钢表面缺陷实时检测方法[J]. 机械科学与技术,2023,42(6):883-889.
ZOU W,JI C. Real-time detection method of surface defects of hot-rolled strip *via* improved YOLOv4-tiny model [J]. Mechanical Science and Technology for Aerospace Engineering,2023,42(6):883-889.(in Chinese)

[13] BOCHKOVSKIY A,WANG C Y,LIAO H Y M. YOLOv4: optimal speed and accuracy of object detection[EB/OL]. 2020:2004.10934. https://arxiv.org/abs/2004.10934v1.

[14] ZHOU X W,WEI M G,LI Q L,et al. Surface defect detection of steel strip with double pyramid network[J]. Applied Sciences,2023,13(2):1054.

[15] YU J B,CHENG X,LI Q F. Surface defect detection of steel strips based on anchor-free network with channel attention and bidirectional feature fusion [J]. IEEE Transactions on Instrumentation and Measurement,2021,71:5000710.

[16] JOCHER G,NISHIMURA K,MINEEVA T,et al. YOLOv5 [EB/OL]. (2020-06-26)[2023-06-02]. https://github.com/ultralytics/yolov5.

[17] LIU Y C,SHAO Z R,HOFFMANN N.Global attention mechanism: retain information to enhance channel-spatial interactions[EB/OL]. 2021:2112.05561. https://arxiv.org/abs/2112.05561v1.

[18] DAI J F,QI H Z,XIONG Y W,et al. Deformable convolutional networks[C]//2017 IEEE International Conference on Computer Vision (ICCV). Venice,Italy. IEEE,2017:764-773.

[19] SRINIVAS A,LIN T Y,PARMAR N,et al. Bottleneck transformers for visual recognition [C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Nashville,TN,USA. IEEE,2021:16514-16524.

[20] 石争浩,李成建,周亮,等. Transformer驱动的图像分类研究进展[J]. 中国图象图形学报,2023,28(9):2661-2692.
SHI Z H,LI C J,ZHOU L,et al. Survey on Transformer for image classification[J]. Journal of Image and Graphics,2023,28(9):2661-2692.(in Chinese)

[21] HU J,SHEN L,ALBANIE S,et al. Squeeze-and-excitation networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,2020,42(8):2011-2023.

[22] WOO S,PARK J,LEE J Y,et al. CBAM: convolutional block attention module [M]//Lecture Notes in Computer Science. Cham:Springer International Publishing,2018:3-19.

[23] HOU Q B,ZHOU D Q,FENG J S. Coordinate attention for efficient mobile network design[C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Nashville,TN,USA. IEEE,2021:13708-13717.

[24] GE Z,LIU S T,WANG F,et al. YOLOX: exceeding YOLO series in 2021[EB/OL]. (2021-07-18)[2023-6-10]. https://arxiv.org/abs/2107.08430.

[25] WANG C Y,BOCHKOVSKIY A,LIAO H Y M. YOLOv7: trainable bag-of-freebies sets new state-of-the-art for real-time object detectors [C]//2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Vancouver,BC,Canada. IEEE,2023:7464-7475.