

# 基于 YOLOv8-NFMC 的带钢表面缺陷检测算法<sup>\*</sup>

朱成杰 刘乐乐 朱洪波

(安徽理工大学电气与信息工程学院 淮南 232001)

**摘要:**针对 YOLOv8 算法在应用于带钢表面缺陷检测时存在漏检和错检等问题,提出了一种改进 YOLOv8 算法。针对数据集中的小目标的标签,在原损失 CIOU 的基础上加入标准化高斯瓦瑟斯坦距离(normalized Gaussian Wasserstein distance, NWD),提升模型对小目标缺陷的检测能力;采用聚焦调制(focal modulation)替换 YOLOv8 模型的空间池化金字塔(spatial pyramid pooling-fast, SPPF),在轻量化的同时,提高多尺度特征的表达能能力;采用移动翻转瓶颈卷积(mobile inverted bottleneck conv, MBConv)替换 C2f 中的 Conv 构建新模块 C2f-MB,同时使用 C2f-MB 替换原有的 C2f 模块,增强特征表达能能力和多尺度特征融合能力;在主干部分加入卷积块注意力机制(convolutional block attention module, CBAM)来抑制背景干扰,能更好捕获全局信息,提升了主干部分的特征提取能力。实验结果表明,改进后的 YOLOv8 算法在计算量下降的同时, mAP@0.5 提高了 3%,对漏检和错检等问题有明显改善。

**关键词:**带钢表面缺陷; NWD; Focal Modulation; MBConv; 注意力机制

**中图分类号:** TP391.41 **文献标识码:** A **国家标准学科分类代码:** 520.2040

## Surface defect detection algorithm for strip steel based on YOLOv8-NFMC

Zhu Chengjie Liu Lele Zhu Hongbo

(School of Electrical and Information Engineering, Anhui University of Science and Technology, Huainan 232001, China)

**Abstract:** Aiming at the problems of leakage and wrong detection when the YOLOv8 algorithm is applied to the surface defect detection of strip steel, an improved YOLOv8 algorithm is proposed. For the labels of small targets in the dataset, normalized Gaussian Wasserstein distance (NWD) is added on top of the original lossy CIOU, which improves the model's ability to detect defects of small targets. Focal Modulation is used to replace the spatial pooling pyramid of the YOLOv8 model, which improves the expression ability of multi-scale features while lightweighting. Mobile inverted bottleneck Conv (MBConv) is used to replace the Conv in C2f to construct a new module C2f-MB, and at the same time replace the original C2f-MB with C2f-MB. MB to replace the original C2f module with C2f-MB, which enhances the feature expression ability and multi-scale feature fusion ability. the convolutional block attention module (CBAM) is added in the backbone part to suppress the background interference, which can better capture the global information and improve the feature extraction ability of the backbone part. Experiment results show that the improved YOLOv8 algorithm improves mAP@0.5 by 3% while decreasing the computation amount, which significantly improves the problems of missed detection and wrong detection.

**Keywords:** steel surface defects; NWD; Focal Modulation; MB Conv; attention mechanism

### 0 引言

在工业生产中,钢材作为一种重要的结构材料,在各个领域扮演着不可或缺的角色,在生产工业钢材和加工过

程中可能出现各种缺陷,因此缺陷检测技术至关重要。产品缺陷检测技术是指对检测样本的表面斑点、凹坑、划痕、色差、缺损和内部结构等缺陷进行检测,获得检测样本表面或内部的缺陷深度、大小、轮廓、缺陷类别等相关

收稿日期:2024-03-03

<sup>\*</sup> 基金项目:国家自然科学基金(62003001)项目资助

信息<sup>[1]</sup>。

随着深度学习的不断发展,深度卷积神经网络在目标检测领域中的应用愈加广泛,现已被应用于农业、交通和医学等众多领域。深度学习凭借强大的表征和建模能力,使得目标检测的效率大大提升<sup>[2]</sup>。深度学习中的目标检测模型主要分为两种:一阶段算法如 YOLO (you only look once)<sup>[3]</sup> 系列等直接预测目标框和类别;二阶段算法如 Faster-RCNN<sup>[4]</sup> 先生成候选区域再进行分类与回归。不管是一阶段算法还是二阶段算法,很多学者利用不同的深度学习模型应用于钢材的缺陷检测,提高了整体的检测精度和工作效率。文献[5]提出了轻量级带钢表面缺陷检测模型 YOLOv8-VSC,采用 VanillaNet 作为特征提取网络,引入 SPD 模块加速推理速度,使用 CARAFE 算子提高特征融合质量,降低了参数量和计算量的同时提升了检测精度。文献[6]提出一种小样本驱动的训练样本生成方法解决钢板制造行业数据匮乏问题,同时结合改进的 YOLOv7 算法,优化 ELAN 模块、引入 ACmix 注意力模块和 WIoU 损失函数,显著提升了钢板缺陷检测性。文献[7]提出了轻量级 DCN-YOLO 模型,结合了可形变卷积网络与 YOLOv5,利用 DSConv 和 ECA 模块降低计算复杂度,提高模型灵敏度与泛化能力,为资源受限的边缘设备提供了高精度的表面缺陷检测解决方案。文献[8]对 YOLOv5 模型进行改进,通过引入 Transformer 自注意力机制、采用 T-BiFPN 网络结构和 RepVGG 轻量化网络,优化了速度与精度的平衡。文献[9]在建立了一整套无酸除鳞带钢表面缺陷数据集的基础上对 YOLOv3 模型进行改进得到了 DF-YOLO 检测模型,并且通过引入深度可分离卷积降低参数量后可以实现带钢表面缺陷的实时检测。文献[10]提出一种基于改进 YOLOX 的钢材表面缺陷检测算法,采用 Swin Transformer 模块捕获全局上下文信息,引入加权双向特征金字塔网络进行跨尺度特征融合,并改进了损失函数,提升了准确率和实用性。对于二阶段算法,也有一些学者进行了研究,文献[11-12]都基于 Faster R-CNN 模型对其优化,并基于此进行带钢表面的缺陷检测。

通过上述的分析可知,基于深度学习的带钢缺陷检测方法研究又很多,但是很多都致力于模型最终的体积和检测精度,针对一些缺陷的错检,漏检的问题研究甚少,本文针对 YOLOv8 原模型应对于带钢表面缺陷检测任务时出现错检,漏检的情况提出了一种改进的 YOLOv8 的算法。首先,针对带钢表面缺陷存在小目标缺陷,通过在原损失中加入标准化高斯瓦瑟斯坦距离(normalized gaussian wasserstein distance, NWD)来提升对小目标标签的检测能力,使模型提高对小目标的适应性,其次,使用聚焦调制(Focal Modulation)来替换空间池化金字塔(spatial pyramid pooling-fast, SPPF)结构来增强多尺度的表达能力,再使用移动翻转瓶颈卷积(mobile inverted bottleneck conv, MBCConv)来替换 C2f 中的 Conv,并使用改进后的

C2f-MB 替换模型中的 C2f,在轻量的同时增强特征表达和融合能力,最后在主干部分引入卷积块注意力机制(convolutional block attention module, CBAM),来提高主干部分的特征提取的能力。改进后的算法在 NEU-DET 数据集上进行验证,证实了改进后的算法在带钢表面缺陷检测方面的可行性。

## 1 YOLOv8n 检测模型及改进

YOLO 系列算法是一系列基于深度学习的目标检测算法,该系列算法通过将目标检测任务转化为一个回归问题,实现了在图像或视频中同时检测出多个对象的位置和类别,YOLOv8 是这个系列当前最新的模型。

YOLOv8 的结构原理可以分为特征提取和目标检测两个部分。YOLOv8n 网络模型主要包含了输入、主干网络、颈部和头部 4 个部分。YOLOv8 网络模型的主干网络部分主要是由 Conv, SPPF 和 C2f 构成。其中 C2f 模块参考了 C3 模块和 ELAN 模块的设计思想, C2f 结构有着更多的残差连接,所以梯度流更为丰富。Head 部分为目前主流的解耦头结构(decoupled-head),将分类和检测头分离。同时采用的是无锚(anchor-free)模型,即直接预测目标的中心点和宽高比例,而不是预测 Anchor 框的位置和大小。针对目前带钢缺陷检测的现状本文提出了一种改进的 YOLOv8n 的网络模型,其改进后算法的结构如图 1 所示。

### 1.1 NWD 损失函数

NWD<sup>[13]</sup> 是一种基于 Wasserstein 距离的归一化高斯距离度量,在目标检测中用于衡量预测边界框和真实目标边界框之间的相似性。它通过将边界框建模为二维高斯分布,计算它们之间的距离,并使用归一化操作来降低尺寸和间距等因素的影响。NWD 的计算方法是基于将预测边界框和真实目标边界框建模为二维高斯分布。通过建模,可以将边界框表示为具有均值和协方差矩阵的高斯分布。然后, NWD 利用这两个高斯分布之间的 Wasserstein 距离来衡量它们之间的相似度,该方法的 Wasserstein 距离计算公式为:

$$NWD(N_a, N_b) = \exp\left(-\frac{\sqrt{W_2^2(N_a, N_b)}}{C}\right) \quad (1)$$

式中:  $C$  是与数据集密切相关的常数;  $W_2^2(N_a, N_b)$  是一个权重参数,用于衡量真实框损失与预测框损失之间的关系。

在带钢缺陷的数据集中,存在部分的小目标标签, NWD 对不同尺度的目标不敏感,因此更适合测量小目标之间的相似性,但是又考虑到将 YOLOv8 的 CIoU 全部换成 NWD 来计算损失,模型的整体收敛速度会变慢,所以最后将 NWD 添加到原模型算法损失函数中去,来提升模型对小目标的检测能力。通过定义一个权重来将 CIoU 和 NWD 进行结合,从而更准确地评估目标检测结果之间的差异和相似性。

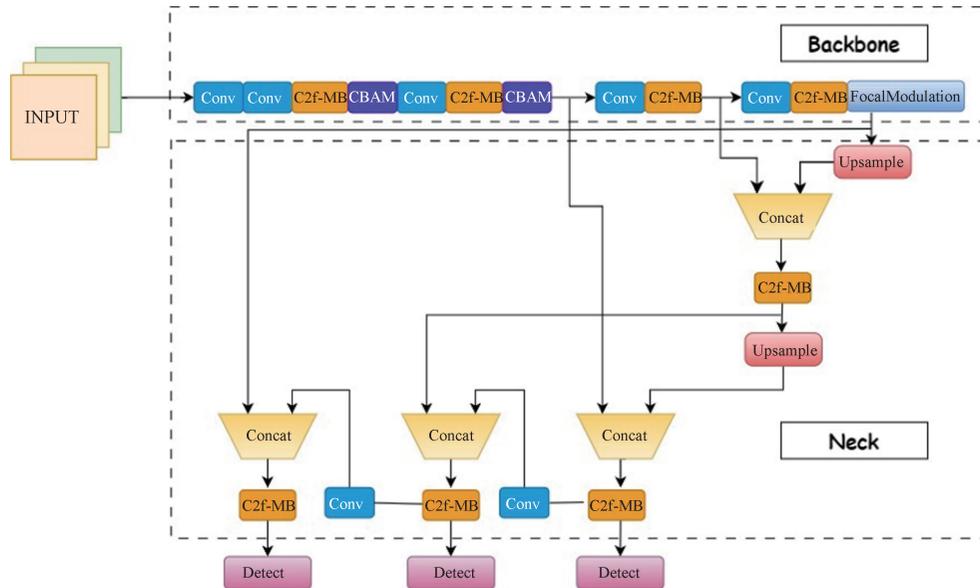


图1 改进的 YOLOv8n 结构  
Fig.1 Improved YOLOv8n structure

### 1.2 Focal Modulation 模块

Focal Modulation<sup>[14]</sup> 由 3 部分组成,其整体过程如图 2(a)所示。第 1 部分为焦点上下文(focal contextualization),焦点上下文使用一系列深度卷积来编码不同范围内的视觉上下文信息。这样的操作可以捕捉从近处到远处的视觉特征,从而允许网络在不同层次上理解图像内容。通过这种方式,Focal Modulation 的输出关注了多尺度的上下文,网络能够在聚合上下文信息时保持对局部细

节的敏感性,并增强对全局结构的认识。

第 2 部分为门控聚合(gated aggregation),门控聚合过程如图 2(b)所示,门控机制在深度学习中常用于控制信息流,它通常用于决定哪些信息应该被传递,哪些应该被阻断。在 Focal Modulation 中门控聚合的目的是为每个查询令牌(即处理中的数据单元)选择性地聚合上下文信息,这意味着网络能够决定哪些特定的上下文信息对于当前处理的查询令牌是重要的,从而专注于那些最相关的

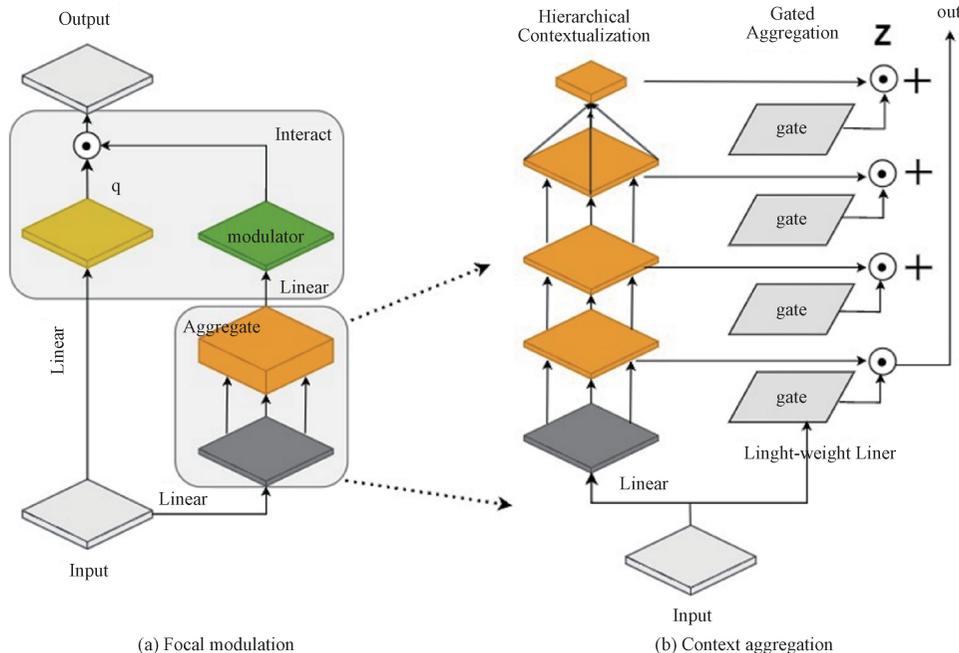


图2 聚焦调制结构和门控聚合结构  
Fig.2 Focal modulation diagram and context aggregation diagram

信息。这种方法提高了模型的效率和性能,因为它减少了不必要信息的处理,同时增强了对关键特征的关注。

第3部分是逐元素仿射变换,在深度学习中,逐元素的仿射变换通常指的是对每个元素进行线性变换,这种变换可以被描述为  $y = ax + b$ , 其中  $x$  是输入,  $y$  是输出,  $a$  和  $b$  是权重的偏差。在 Focal Modulation 中,逐元素仿射变换的作用是将聚合后的信息注入到每个查询令牌中,通过逐元素仿射变换,模型能够更细致地调整每个查询令牌的特征,根据上下文信息来增强或抑制某些特征。这种精细的调整机制允许网络更好地适应复杂的视觉场景,提高对细节的捕捉能力,从而提升了模型在各种视觉任务中的性能。

使用 Focal Modulation 替换 YOLOv8 中的 SPPF 结

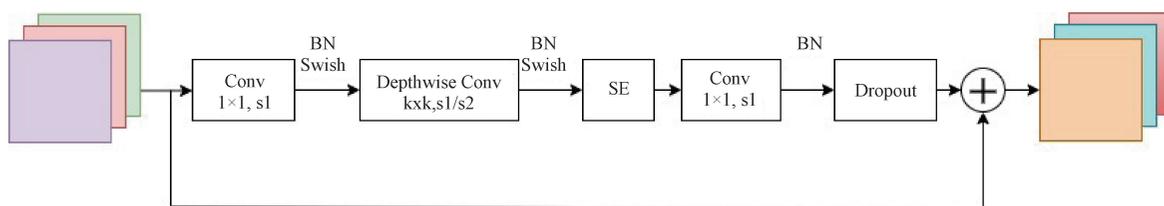


图3 MBConv 结构

Fig. 3 MBConv structure diagram

其中 BN(batch normalization)层可以提供给网络批标准化,Swish 函数可以给网络带来更大的非线性,提高网络的拟合能力。MBConv 相比于传统卷积,采用了 DWConv,其次采用了“倒瓶颈”的结构,特征经历了升维和降维两个步骤,提高模型的学习能力。

MBConv 中融合的 SE 模块<sup>[16]</sup>是一种用于增强特征表达的机制,使得网络能够更加关注信息量最大的通道特征,并抑制那些不重要的通道特征。通过引入 SE 模块,可以提高网络对于重要特征的感知能力,增强网络的表达能力。SE 模块首先将输入的特征图经过全局平均池化操作,将其压缩成一个  $1 \times 1 \times C$  的向量,其中  $C$  表示通道数。然后,这个向量通过一个激励函数进行处理,得到各通道的权重向量。这个权重向量表示了每个通道对于最终输出的重要程度。最后,将权重向量与原始特征图进行逐通道相乘,得到最终的输出结果。这样就实现了对特征图的通道权重的有效分配,使得网络能够更加关注重要的特征信息。在本文中,使用 MBConv 来改进 YOLOv8 中的 C2f 结构,通过使用 MBConv 替换 C2f 中的 Bottleneck 中的 Conv 构建新的结构 C2f-MB,改进后 C2f 结构如图 4 所示。

MBConv 中的 DWConv 可以将每个通道的卷积操作分离开来并且 MBConv 是在 DWConv 基础上增加了升维和降维两个步骤,从而改进后的 C2f 结构减少卷积的计算量和参数数量,进一步提高模型的学习能力和表达能力。同时 MB Conv 中融合的 SE 模块可以增强特征的表达能力,通过自适应地对通道进行加权处理,提高模型对于重

构可以提高多尺度特征的表达能力,通过使用深度卷积来编码不同范围内的空间上下文,捕捉从近处到远处的视觉特征,能够更好地学习到不同尺度的特征,增强对全局结构的认知,同时通过门控聚合的操作选择性地聚合上下文信息,专注于那些最相关的信息,从而提升缺陷检测任务的性能。

### 1.3 改进的 C2f 模块

MBConv<sup>[15]</sup>结构主要由一个  $1 \times 1$  的普通卷积来进行升维,然后经过一个深度卷积(depthwise convolution, DWConv)、压缩激励(squeeze excitation, SE)模块进一步提取特征,通过一个  $1 \times 1$  卷积降维后通过一个 Dropout 层输出,其结构如图 3 所示。

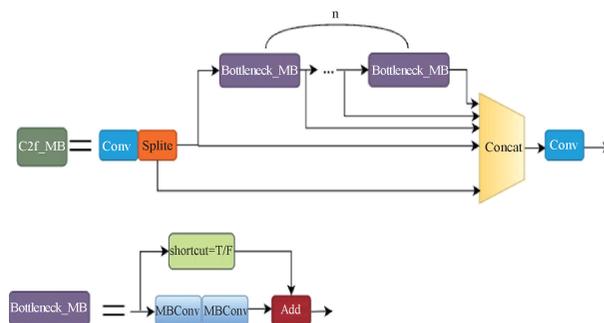


图4 C2f-MB 结构

Fig. 4 C2f-MB structure diagram

要特征的感知能力,也进一步提高了改进后 C2f 模块的性能,增强特征表达能力和多尺度特征融合能力。

### 1.4 CBAM 注意力机制模块

深度学习中的注意力机制是一种仿生概念,模拟人类对于观察对象内容的特定部分进行重点关注的过程,通过将注意力聚焦到对象信息的位置来提升信息获取的能力<sup>[17]</sup>。

CBAM 模块<sup>[18]</sup>主要由通道注意力和空间注意力两部分组成,对输入进来的特征层分别进行通道注意力模块和空间注意力模块的处理,是一种用于增强特征表达能力的注意力模块,CBAM 模块的结构如图 5 所示。CBAM 模块是一种简单高效的注意力模块,也是一个轻量级的通用注意力模块,可以方便地加入到前馈卷积神经网络中,它沿着两个独立的维度(通道和空间)依次推断注意力图,然

后将注意力图乘以输入特征图以进行自适应特征细化。

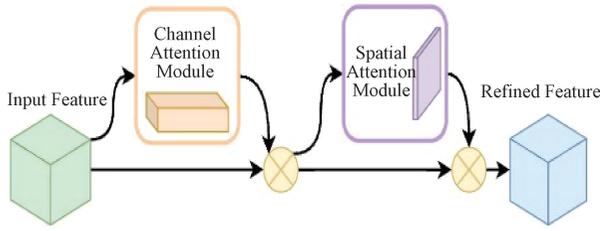


图5 CBAM结构

Fig. 5 CBAM structure diagram

本文的带钢缺陷数据集为灰度图像,图像的背景复杂,部分缺陷相似度高,具有不同尺度、形状和纹理等特征。这些冗余信息会干扰缺陷的检测,降低检测算法的准确性,CBAM模块在通道注意力机制中引入全连接,并通过全连接降维,有利于提取更重要的信息。本文在主干部分加入CBAM模块来抑制背景干扰,提高模型的特征提取能力,更好地捕获全局范围内的重要特征。

## 2 实验结果与分析

### 2.1 实验环境与参数设定

本文研究时的环境配置如下: Intel(R) Xeon(R) CPU E5-2686 v4, 显卡 308-10 G, CUDA 版本 11.7.0, OS 为 Ubuntu 20.04.5 LTS, PyTorch 版本为 1.13.1, CUDA 版本为 11.7.0, Python 版本 3.8.0, 设置初始学习率 lr0 为 0.01, 动量(momentum)为 0.937, 优化器为 Adamw, Batch Size 为 32, Works 为 4, IoU 为 0.5, 图片大小为 640×640。

### 2.2 数据集与评估标准

本文所做实验的数据集采用是东北大学发布的 NEU-DET 数据集, NEU-DET 数据集包含了裂纹(crazing)、夹杂(inclusion)、斑块(patchs)、麻点(pitted\_surface)、压入氧化铁(rolled-in\_scale)及划痕(scratches)这6类带钢表面缺陷各300张,共计1800张,考虑到数据集的样本数较少,本文将1800张图片按8:2的比例将1440张作为训练集,360作为验证集,各类缺陷图如图6所示。

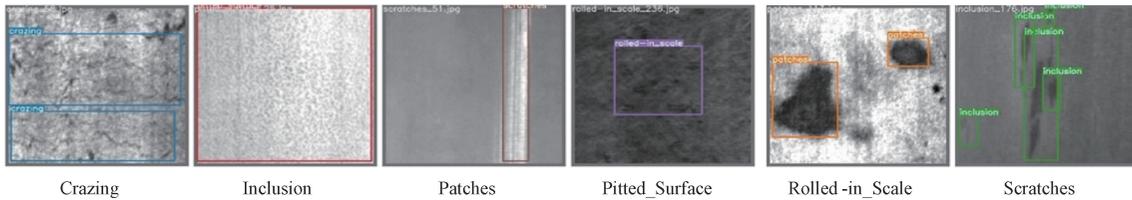


图6 各类缺陷图

Fig. 6 Diagram of various defects

本文的训练模型指标评估从精确率(precision, P)、召回率(recall, R)、平均精度均值(mean average precision, mAP)、参数量(Parameters)、计算量(GFLOPs)、帧率对训练模型进行指标评估。其中, mAP@0.5 (IoU 阈值大于 0.5 的 mAP) 反映的是模型的精确率随召回率的变化趋势, 该指标的数值越高代表模型越容易在高召回率下保持

这高精确率。

### 2.3 实验结果及消融实验

为保证改进前后的公平性和改进效果的显著性,改进前后均未载入预训练模型,改进前后6类缺陷精度对比如图7所示。由图7可知,除了麻点和压入氧化铁的平均精度(AP)有所轻微下降外,其余4类的缺陷AP都有所提

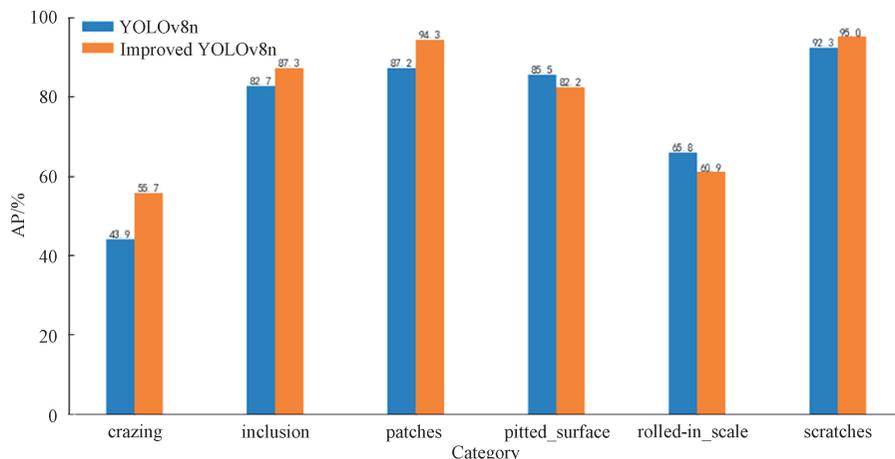


图7 改进前后各类缺陷精度对比

Fig. 7 Comparison of the accuracy of various types of defects before and after improvement

升,其中裂纹提升最大,提高了 11.8%。改进后的模型在参数量上有轻微的提升,但是计算量降低了 12.2%,改进前 mAP@0.5 为 76.2%,改进后的 mAP@0.5 为 79.2%,较原模型的 mAP@0.5 提高了 3%。

本文通过 4 个改进方案对网络模型进行了改进,为探

究 4 个改进方案对研究结构的影响对此进行了消融实验,实验结果如表 1 所示。改进 1 为将 NWD 加入到原损失中,改进 2 为使用 Focal Modulation 替换 YOLOv8 中的 SPPF 结构,改进 3 为使用 C2f-MB 结构替换模型中的 C2f 结构,改进 4 为在主干部分加入 CBAM。

表 1 消融实验

Table 1 The ablation experiment

改进 1	改进 2	改进 3	改进 4	Layers	Parameters	浮点数/GFLOPs	mAP@0.5/%	帧率/fps
×	×	×	×	225	3 012 018	8.2	76.2	72.99
√	×	×	×	225	3 012 018	8.2	77.4	71.94
√	√	×	×	230	3 120 053	8.3	78.1	75.19
√	√	√	×	310	3 468 629	7.2	78.6	67.11
√	√	√	√	326	3 474 041	7.2	79.2	62.89

从上述数据可以看出,在每个改进加入到模型后 mAP@0.5 都有所提升,改善模型的检测性能。改进后的模型较原模型对比在参数量上升了 15.7%,计算量下降了 12.2%,mAP@0.5 提高 3%。改进 1 对模型的精度提升最为明显,通过将 NWD 引入到 CIoU 中来提升对小目标的检测性能,mAP@0.5 提高 1.2%。改进 1 和改进 2 组合时,计算量稍微增加但是帧率增加了 3%。改进 3 的使用虽然提升了一部分的参数,但是计算量下降了 12.2%,mAP@0.5 提高 2.4%。在 4 个改进组合的时候在计算量降低了 12.2%,mAP@0.5 提高 3%,帧率达到了 62.89 fps,改进后的模型通过优化损失函数、增强特征提取能力和增强了模型对目标物体的表达和感知能力来提升对小目标的更精确的检测,进而整体提升模型的检测性能。

## 2.4 不同算法对比

为了验证改进后的 YOLOv8n 的模型的性能,与当前的几种主流检测模型文献[7]、文献[19]、YOLOv3<sup>[20]</sup>、YOLOv5<sup>[21]</sup>和 YOLOv7<sup>[22]</sup>对一些指标进行对比,在和本文同样的数据集下进行改进后的模型进行实验对比,结果如表 2 所示。

综合分析上述数据,改进的 YOLOv8n 的 mAP@0.5 达到了 79.2%,与其他目标检测算法相比有效提升了缺陷检测的精度。其中,对于文献[7]提高 2.7%,对于文献[19]提高 5.1%,对于 YOLOv3 模型提高 3.8%,YOLOv3-Tiny 提高 8.8%,YOLOv5n 模型提高 2.7%,YOLOv5s 模型提高 3.6%,YOLOv7 模型提高 3.8%,YOLOv7-tiny 提高 8%。在计算量方面,本文方法的计算量和 YOLOv5n 模型的计算量相同,比其他模型的参数量都要少。在参数量方面,除 YOLOv5n 模型比本文算法少外,其余算法的参数量都比本文算法要大。改进后的 YOLOv8n 模型在模型大小、计算资源消耗和性能之间达

表 2 不同模型对比实验

Table 2 Comparative experiments with different models

模型	Parameters /( $\times 10^6$ )	浮点数 /GFLOPs	mAP@0.5/%
文献[7]	16.3	39.5	76.5
文献[19]	5.7	12.2	74.1
YOLOv3	103.7	283.0	75.2
YOLOv3-Tiny	12.13	19.0	70.6
YOLOv5n	2.51	7.2	76.6
YOLOv5s	9.12	24.1	75.6
YOLOv7	37.22	105.2	75.4
YOLOv7-Tiny	6.03	12.3	71.2
本文方法	3.47	7.2	79.2

到了良好的平衡。其具有较少的参数和计算量,同时在 mAP@0.5 指标上表现出更高的目标检测性能。在实际部署中,改进后的 YOLOv8n 模型能够在保证高效运行的同时,提供更为准确的目标检测结果,满足了现实场景中对于精度和效率兼顾的需求。

## 2.5 检测效果对比

6 种类别缺陷样本改进前后的检测如图 8 所示,每个标注框上方有标签类别和置信度。改进前后的检测图对比后可以得到如下两个结论:1)原 YOLOv8n 模型对于一些缺陷存在漏检,错检的情况,改进后的 YOLOv8n 模型对于缺陷的检测更加充分,一定程度上减少了漏检,错检的情况;2)同一张缺陷图片上的缺陷标签在改进的 YOLOv8n 模型检测后的置信度更高,总体的检测性能更强。

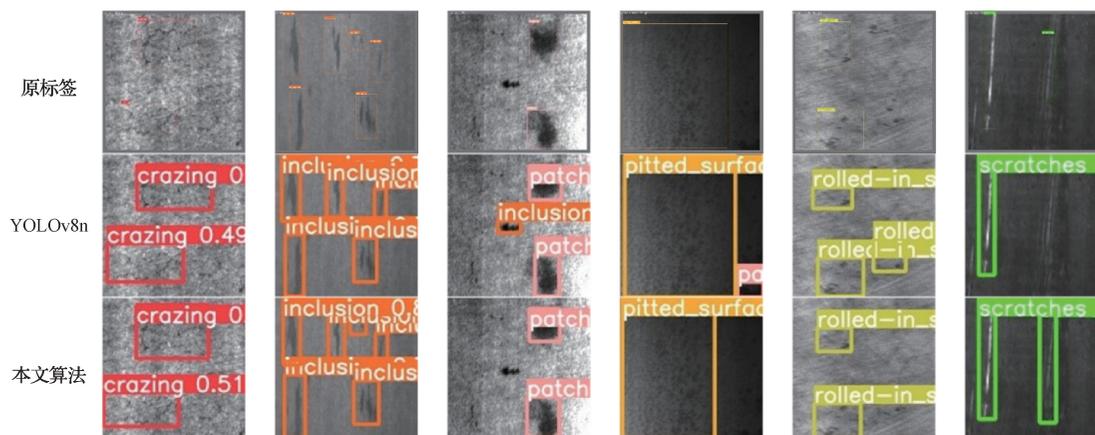


图8 检测效果对比

Fig.8 Comparison chart of detection effect

### 3 结论

本文提出了一种改进的YOLOv8n算法模型应用于工业钢铁表面缺陷检测。首先,在原损失函数CIoU中加入NWD来计算损失,提升对小目标的检测性能;其次Focal Modulation替换YOLOv8中的SPPF结构,提升多尺度的特征表达能力;然后使用C2f-MB结构替换模型中的C2f结构,在降低计算量的同时增加重要特征的权重。最后,引入CBAM注意力机制,来增强主干网络的特征提取能力。改进后的模型在计算量下降12.2%的情况下,模型整体的性能获得提升。但是最后训练的模型中,部分类别精度存在偏低的问题,可能是因为训练集和验证集标签数目太少的的原因。后续将继续优化网络结构,增强数据集,进一步提升模型的检测能力。

#### 参考文献

- [1] 李少波,杨静,王铮,等. 缺陷检测技术的发展与应用研究综述[J]. 自动化学报,2020,46(11):2319-2336.  
LI SH B, YANG J, WANG ZH, et al. Review of development and application of defect detection technology [J]. Acta Automatica Sinica, 2020, 46(11):2319-2336.
- [2] 侯学良,单腾飞,薛靖国. 深度学习的目标检测典型算法及其应用现状分析[J]. 国外电子测量技术,2022,41(6):165-174.  
HOU X L, SHAN T F, XUE J G. Analysis of typical target detection algorithm based on deep learning and its application status [J]. Foreign Electronic Measurement Technology, 2022, 41(6):165-174.
- [3] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: Unified, real-time object detection[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 779-788.
- [4] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6): 1137-1149.
- [5] 王春梅,刘欢. YOLOv8-VSC:一种轻量级的带钢表面缺陷检测算法[J]. 计算机科学与探索,2024,18(1):151-160.  
WANG CH M, LIU H. YOLOv8-VSC: A lightweight algorithm for strip surface defect detection [J]. Journal of Frontiers of Computer Science and Technology, 2024, 18(1):151-160.
- [6] 窦智,胡晨光,李庆华,等. 改进YOLOv7的小样本钢板表面缺陷检测算法[J]. 计算机工程与应用,2023,59(23):283-289.  
DOU ZH, HU CH G, LI Q H, et al. Improved YOLOv7 algorithm for small sample steel plate surface defect detection [J]. Computer Engineering and Applications, 2023, 59(23):283-289.
- [7] 卢俊哲,张毓怡,刘世鹏,等. 面向复杂环境中带钢表面缺陷检测的轻量级DCN YOLO[J]. 计算机工程与应用,2023,59(15):318-328.  
LU J ZH, ZHANG CH Y, LIU SH P, et al. Lightweight DCN-YOLO for strip surface defect detection in complex environments [J]. Computer Engineering and Applications, 2023, 59(15): 318-328.
- [8] 马燕婷,赵红东,阎超,等. 改进YOLOv5网络的带钢表面缺陷检测方法[J]. 电子测量与仪器学报,2022,36(8):150-157.  
MA Y T, ZHAO H D, YAN CH, et al. Strip steel surface defect detection method by improved

- YOLOv5 network [J]. Journal of Electronic Measurement and Instrumentation, 2022, 36(8): 150-157.
- [9] 宋紫棋, 毛桂庭, 侯奔. 基于 YOLO 网络的无酸除鳞带钢表面缺陷实时检测[J]. 矿冶工程, 2022, 42(2): 136-140.
- SONG Z Q, MAO G T, HOU B. Real-time detection of surface defects of acid-free descaled steel strip based on YOLO network [J]. Mining and Metallurgical Engineering, 2022, 42(2): 136-140.
- [10] 熊聪, 于安宁, 高兴华, 等. 基于改进 YOLOX 的钢材表面缺陷检测算法[J]. 电子测量技术, 2023, 46(9): 151-157.
- XIONG C, YU AN N, GAO X H, et al. Steel surface defect detection algorithm based on improved YOLOX [J]. Electronic Measurement Technology, 2023, 46(9): 151-157.
- [11] HAIYUN W, JIANPING W, FUHUA L. Study on surface defect detection of metal sheet and strip using faster R-CNN with multilevel feature[J]. Mechanical Science and Technology for Aerospace Engineering, 2021, 40(2): 262-269.
- [12] HE Y, SONG K, MENG Q, et al. An end-to-end steel surface defect detection approach via fusing multiple hierarchical features[J]. IEEE Transactions on Instrumentation and Measurement, 2019, 69(4): 1493-1504.
- [13] WANG J, XU C, YANG W, et al. A normalized Gaussian Wasserstein distance for tiny object detection[J]. ArXiv preprint, arXiv: 2110. 13389, 2021.
- [14] YANG J, LI C, DAI X, et al. Focal modulation networks [J]. Advances in Neural Information Processing Systems, 2022, 35: 4203-4217.
- [15] TAN M, LE Q. Efficientnet: Rethinking model scaling for convolutional neural networks [C]. International Conference on Machine Learning, 2019: 6105-6114.
- [16] HU J, SHEN L, SUN G. Squeeze-and-excitation networks[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 7132-7141.
- [17] 孔子迁, 邓蕾, 汤宝平, 等. 基于时频融合和注意力机制的深度学习行星齿轮箱故障诊断方法[J]. 仪器仪表学报, 2019, 40(6): 221-227.
- KONG Z Q, DENG L, TANG B P, et al. Fault diagnosis of planetary gearbox based on deep learning with time-frequency fusion and attention mechanism[J]. Chinese Journal of Scientific Instrument, 2019, 40(6): 221-227.
- [18] WOO S, PARK J, LEE J Y, et al. CBAM: Convolutional block attention module[M]. FERRARI V, HEBERT M, SMINCHISESCU C, et al. Computer Vision-ECCV 2018. Lecture Notes in Computer Science. Cham: Springer, 2018, 11211: 3-19.
- [19] 阳丽莎, 李茂军, 胡建文, 等. 基于改进 Yolov7-tiny 的带钢表面缺陷检测算法[J/OL]. 计算机工程, 1-9 [2024-06-13]. <https://doi.org/10.19678/j.issn.1000-3428.0068397>.
- YANG L SH, LI M J, HU J W, et al. Strip steel surface defect detection algorithm based on improved Yolov7-tiny [J/OL]. Computer Engineering, 1-9 [2024-06-13]. <https://doi.org/10.19678/j.issn.1000-3428.0068397>.
- [20] REDMON J, FARHADI A. YOLOv3: An incremental improvement[J]. ArXiv preprint, arxiv: 1804. 02767, 2018.
- [21] JOCHER G. YOLOv5 [EB/OL]. [2020-08-09]. <https://github.com/ultralytics/yolov5>.
- [22] WANG C Y, BOCHKOVSKIV A, LIAO H Y M. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023: 7464-7475.

## 作者简介

朱成杰, 副教授, 主要研究方向为嵌入式系统与应用、图像处理与分类算法。

E-mail: ahhbzcj@126.com

刘乐乐(通信作者), 硕士研究生, 主要研究方向为机器视觉、深度学习。

E-mail: 307408983@qq.com

朱洪波, 副教授, 主要研究方向为机器人运动规划及控制、移动机器人与 WSN 协同感知及协作定位。

E-mail: hbzhu@aust.edu.cn