

基于尺度多样化改进的钢材表面缺陷检测算法

周孟然^a, 蔡睿^b, 范桃春^b, 王宁^b

安徽理工大学 a. 电气与信息工程学院; b. 力学与光电物理学院, 安徽淮南 232001

摘要:目的 钢材表面缺陷一直是工业生产和安全的难题之一,为解决钢材表面缺陷检测精度低,容易出现误检和漏检的问题,提出了一种基于多尺度特征的钢材表面缺陷检测算法,命名为 GPDN-Yolov5。方法 首先,针对部分缺陷容易与背景相混淆的问题,通过在骨干网络中加入改进后的坐标注意力机制(GJ-CA),加强对缺陷特征的关注。其次,针对钢材表面缺陷具有尺寸大小不一的特点,引入多尺度卷积块 PKI Block,加强了网络的多尺度特征提取能力。然后,采用 Dynamic head 检测头作为改进算法的检测头,增强了检测头的表达能力和对多尺度特征的检测能力。随后,将 NWD(Normalized Wasserstein Distance)损失函数与 WIOU 损失函数加权融合来替代 CIOU 损失函数,解决了原模型专注于简单样本和对低像素缺陷位置偏差敏感的问题。结果 GPDN-Yolov5 算法在 NEU-DET 数据集中的平均精度均值达到 80.5%,比 Yolov5 算法提升了 6.9%,检测结果优于其他经典检测模型。同时,该算法在 GC10-DET 数据集上的平均精度比 Yolov5 提升了 4.7%,这证明了算法在不同数据集上检测效果均有提升。结论 GPDN-Yolov5 算法在精度上明显提高,对钢材表面缺陷检测具有一定价值。

关键词: Yolov5; 缺陷检测; 注意力机制; 多尺度特征; Dynamic head; NWD-WIOU

中图分类号: TP391 **文献标识码:** A **doi:** 10.16055/j.issn.1672-058X.2026.0003.009

An Improved Steel Surface Defect Detection Algorithm Based on Multi-Scale Features

ZHOU Mengran^a, CAI Rui^b, FAN Taochun^b, WANG Ning^b

a. School of Electrical and Information Engineering, b. School of Mechanics, Optoelectronics and Physics, Anhui University of Science and Technology, Huainan 232001, Anhui, China

Abstract: Objective Surface defects in steel remain a critical challenge for industrial production and safety. To address the problems of low accuracy, false detection, and missed detection in steel surface defect detection, a steel surface defect detection algorithm based on multi-scale features, named GPDN-Yolov5, is proposed. **Methods** First, to tackle the problem of certain defects blending with the background, an improved coordinate attention mechanism (GJ-CA) was incorporated into the backbone network to enhance focus on defect features. Second, considering the wide variation in defect sizes, a multi-scale convolutional block (PKI Block) was introduced to strengthen the network's ability to extract multi-scale features. Then, a Dynamic head was applied as the detection head of the improved algorithm, which strengthened the expression ability of the detection head and the detection ability of multi-scale features. Finally, the CIOU loss function was replaced by a weighted combination of the normalized Wasserstein distance (NWD) loss function

收稿日期: 2024-05-30 **修回日期:** 2024-08-18 **文章编号:** 1672-058X(2026)03-0070-11

基金项目: 国家重点研发计划重点专项子课题(2018YFC0604503)。

作者简介: 周孟然(1965—),男,安徽淮南人,博士,教授,从事矿山机电系统监测、光电信息处理和煤矿安全监测监控研究。

通信作者: 蔡睿(1998—),男,安徽马鞍山人,硕士研究生,从事图像处理研究。Email: 1138477297@qq.com。

引用格式: 周孟然,蔡睿,范桃春,等. 基于尺度多样化改进的钢材表面缺陷检测算法[J]. 重庆工商大学学报(自然科学版), 2026, 43(3): 70-80.

Zhou Mengran, Cai Rui, Fan Taochun, et al. An improved steel surface defect detection algorithm based on multi-scale features [J]. Journal of Chongqing Technology and Business University (Natural Science Edition), 2026, 43(3): 70-80.

and the WIOU loss function, thereby addressing the original model's bias toward easy samples and its high sensitivity to positional deviation in low-pixel defects. **Results** The GPDN-Yolov5 algorithm achieves a mean average precision (mAP) of 80.5% on the NEU-DET dataset, representing a 6.9% improvement over the baseline Yolov5 algorithm, and outperforms other classical detection models. Moreover, on the GC10-DET dataset, the algorithm improves the mean average precision by 4.7% over Yolov5, demonstrating its consistent performance gain across different datasets. **Conclusion** The GPDN-Yolov5 algorithm provides a significant improvement in detection accuracy and holds certain value for steel defect detection.

Keywords: Yolov5; defect detection; attention mechanism; multi-scale feature; Dynamic head; NWD-WIOU

钢材在工业生产和日常生活中扮演着至关重要的角色。然而,在生产和运输钢材时,常常会因为各种偶然因素导致钢材表面出现氧化、划痕、夹杂物等缺陷^[1]。这些缺陷通常会对钢材使用性能产生影响,从而导致工业和建筑业产品质量下滑和发生事故的可能性增加。因此,在钢材投入应用前,采用高效的检测方法对钢材表面进行缺陷检测就变得非常有意义^[2]。

随着人工智能技术的发展,人们发现机器视觉技术具有检测准确且应用范围广等特点。国内外研究者对机器视觉技术进行了深入的研究,并将其划分为机器学习方法和深度学习方法^[3]。机器学习方法主要基于手工提取特征,虽然经过了十多年发展,但其识别效果并未得到有效改善。而深度学习算法能够自动提取目标特征,将目标特征融合到模型建立中,比机器学习检测效果更好更快,适用于挑战性较大的钢材表面缺陷检测。

目前,深度学习算法主要分为两大类:一种是以 R-CNN^[4]、Fast-RCNN^[5]、R-FCN^[6]、Faster-RCNN^[7] 为首的两阶段检测算法。它们需要先生成候选框,然后才能进行目标的定位和分类。它们的优点是检测精度较高,但缺点则是两阶段的过程过于复杂,导致检测速度较慢。另一种则是以 Yolo^[8-10] 系列算法、SSD^[11] 算法为首的一阶段检测算法。它们可以直接对目标进行定位和分类,这也使它们的检测速度和计算量优于两阶段算法,但它在某些复杂场景下的检测精度相较于二阶段算法可能有一点差距。这些算法在钢材表面缺陷检测中都得到了广泛应用。

杨溢^[12]将 DFPN 算法和 De Net59 算法相结合对 Faster-RCNN 算法进行了改进,改进后的算法对小尺寸缺陷的检测精度有明显提升。但是该模型对不同尺寸缺陷检测性能一般,在多尺度缺陷上的整体提升并不高,并且检测速度较慢。李鑫等^[13]采用了 Ghost

Bottleneck 模块替代了 Yolov5 中的 C3 模块,并引入了 SE 注意力机制。该算法虽然加快了模型的检测速度,但对于精度提升较小。吕秀丽等^[14]提出了 YOLOv5s-STCD 算法,该算法在骨干网络引入了 Swin Transformer 模块和坐标注意力机制,加强了网络特征提取能力,并针对缺陷尺寸多样化的问题增加了检测层。YOLOv5s-STCD 虽然对钢材表面缺陷检测的平均精度有一定程度的提升,但它对夹杂物这一低像素缺陷的精度比改进前有所下降。

上述算法对钢材表面缺陷检测的平均精度有所提升,有的加快了检测速度,但它们又存在着一些缺点。部分算法过于看重轻量化改进,忽略了检测精度的重要性,改进算法的平均精度不高;部分算法专注于小尺寸目标,而忽略了钢材表面缺陷尺寸的多样性;还有部分算法改进后对低像素缺陷精度提升不高。为了解决这些难题,本文提出了一种基于尺度多样化改进的 GPDN-Yolov5 算法。针对部分缺陷容易受背景干扰的问题,本文在骨干网络中加入改进后的 CA 注意力机制。然后,本文还在 SPPF 后加入多尺度卷积块 PKI Block,从而加强模型对各种尺度缺陷的检测能力。对于原检测头表达能力差等问题,本文将 Yolov5 检测头更换为具有较好动态学习能力和表达能力的 Dynamic head 检测头。最后本文将 NWD 损失函数和 WIOU 损失函数结合起来代替 CIOU 损失函数,避免部分像素较低的缺陷被误检和漏检,并解决了模型专注于简单样本的问题。在数据集 NEU-DET 上的实验表明:GPDN-Yolov5 算法对于钢材表面缺陷检测有着良好的检测性能。

1 Yolov5 网络模型与改进

Yolov5 本质上是一个具有回归函数的深度卷积网络。它首先通过骨干网络进行特征提取,而后由颈部

网络对提取的特征进行融合,最后再由头部网络进行目标预测。与其他算法相比,Yolov5 具有速度快、精度高、模型轻便等特点,其网络结构如图 1 所示。

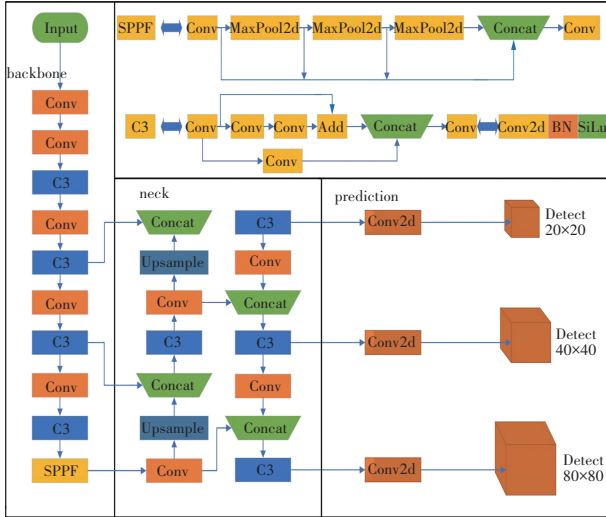


图 1 Yolov5s 网络结构

Fig. 1 The network structure of Yolov5s

针对钢材表面缺陷容易受到背景因素干扰,各个缺陷尺寸相差较大,数据集中缺陷像素较低等问题。本文在骨干网络中加入了改进后的坐标注意力机制^[16]:GJ-CA 注意力机制。在 SPPF 后加入了多尺度卷积块 PKI Block 并将 Yolov5 的 Detect 检测头更换为 Dynamic head^[17] 检测头,最后将 NWD^[18] 损失函数和 WIOU 损失函数进行结合得到 NWD-WIOU 损失函数来代替 CIOU,改进的网络结构如图 2 所示。

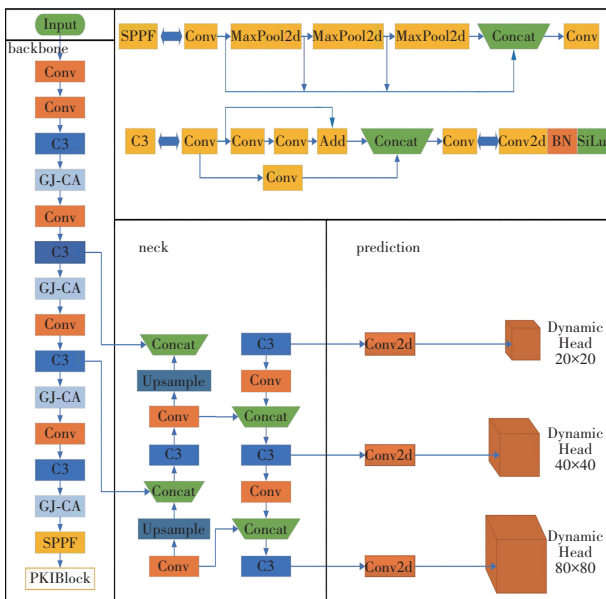


图 2 GPDN-Yolov5 网络结构

Fig. 2 The network structure of GPDN-Yolov5

1.1 改进后的 CA 注意力机制

注意力机制因其轻便和计算量增幅小的特点被广泛应用于深度学习的领域中。在缺陷检测中,它可以减轻背景因素对目标的干扰。常用的注意力机制有:通道注意力机制 SE^[19],混合注意力机制 CBAM^[20],坐标注意力机制 CA 等。SE 注意力机制首先对输入的图片特征进行全部平均池化,然后再将图像特征经过两个全连接层,最后再使用 Sigmoid 函数得到各通道权重值来获取加入注意力机制后的特征图^[21]。它的缺点是不适用于通道数少的场景。CBAM 注意力结合了卷积和注意力机制对输入图像分别进行通道注意力和空间注意力的处理,缺点是考虑的区域有限,忽略了长程依赖问题。CA 注意力将输入特征图分别在宽度和高度两个方向进行全局平均池化,从而获得在宽度和高度两个方向的特征图。它能够准确地捕捉不同通道之间的关系,从而提高特征表达的准确性,并且比 SE 和 CBAM 适用面更广,但是它没有很好地利用较为突出的信息,其原理如图 3 所示。而改进后的 CA 注意力机制(简称 GJ-CA)结合了平均池化和最大池化模式对输入特征层进行池化,既保留了 CA 注意力的原有优势,又可以更好地利用较为突出的信息,并且其能够简便地插入到 Yolov5 的核心模块中。

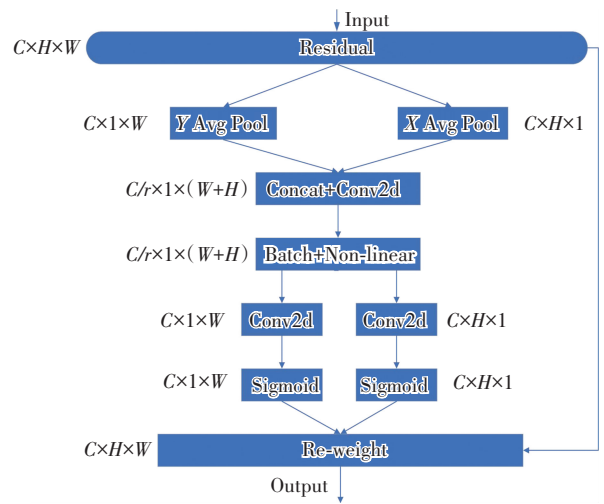


图 3 CA 模块

Fig. 3 CA module

在钢铁缺陷的图像中,针对裂纹、夹杂物、凹面、划痕等中小目标缺陷与背景混杂不清的问题,本文在 Yolov5s 的骨干网络加入了 GJ-CA 注意力,其原理如图 4 所示。从图 4 中不难发现,GJ-CA 比 CA 多出了最大池化模式这一部分。

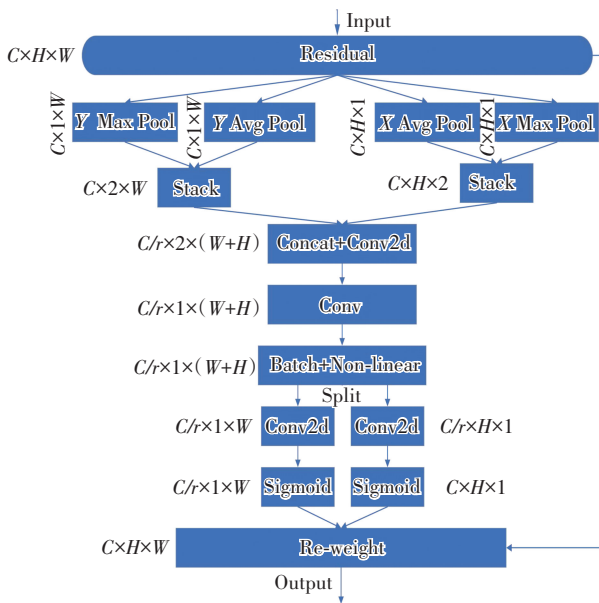


图 4 GJ-CA 模块
Fig. 4 GJ-CA module

GJ-CA 分别沿着高度 h 和宽度 w 方向以两种池化模式(平均池化和最大池化)对输入特征层进行池化,从而生成两个特征层 X 。然后将宽度和高度维度堆叠在一起,以便对相同的维度进行分组,计算过程为式(1)和式(2)。堆叠后,特征层 X 将成为两个具有空间维度的新特征层。然后,通过维度变换和维度叠加得到特征层 X' ,计算过程为式(3)。为了降低计算量,GJ-CA 和 CA 一样设置了一个超参数(衰减率 r)。通过衰减率 r ,卷积核大小为 1×2 或 2×1 、填充为 $(0,0)$ 的卷积运算提取新特征层 X' ,从而生成信息显著的新特征层 X'' ,计算过程为式(4)。最后通过 δ 函数计算特征层中每个特征的权值。再乘以原特征层得到输出的特征层 X_{out} ,可由式(5)表示:

$$X'_h = c(P_{ha}(X), P_{hm}(X)) \quad (1)$$

$$X'_w = c(P_{wa}(X), P_{wm}(X)) \quad (2)$$

$$X' = c(X'_h, X'_w) \quad (3)$$

$$X'' = F(X') \quad (4)$$

$$X_{out} = X * \delta(f(f_1(X'_{split}))) \quad (5)$$

其中, P_{ha} 和 P_{hm} 分别表示沿 h 方向平均池化、最大池化, P_{wa} 和 P_{wm} 则分别指沿 w 方向的平均池化、最大池化。 c 函数表示特征层的叠加, F 则代表卷积核大小为 1×2 或 2×1 填充为 $(0,0)$ 的卷积运算, $*$ 表示逐点乘法, f_1 表示衰减率为 r 的卷积运算, $split$ 函数负责在高度上和宽度上将特征层分为两部分, f 函数表示使用卷积核大小为 1×1 的卷积层和原始维数的输出通道进行卷积操作。 δ 函数则用来计算每个特征的权值。

1.2 加强多尺度特征提取的 PKI Block

由于钢材表面缺陷特征差异和尺寸跨度非常大,

钢材表面缺陷检测的模型需要有着优异的多尺度检测能力。为此,本文通过在骨干网络中的 SPPF 后面加入多尺度卷积块 PKI Block(图 5)来加强网络对多尺度特征的提取能力。PKI Block 主要由 PKI 模块和 CAA 模块组成。它先将 PKI 模块和 CAA 模块的输出进行逐元素求积,然后再将求积结果与 PKI 模块的输出进行逐元素求和,最后通过卷积操作得到结果。PKI 模块的作用是捕获各种尺度纹理特征,它包括一个小核卷积来捕获本地信息和一组并行的深度可分离卷积来捕获不同尺度的上下文信息,最后再由 1×1 的卷积融合这两种信息得到最终的多尺度本地上下文信息,通过这种方式,PKI 模块可以在不损害局部纹理特征完整性的情况下捕获广泛的上下文信息。而 CAA 模块的作用是捕获长距离的上下文信息,增强模型对远距离依赖性的建模能力。它很好地弥补了 PKI 模块专注局部上下文信息的缺点。所以两者相结合组成的 PKI Block 可以加强模型的多尺度特征提取能力。

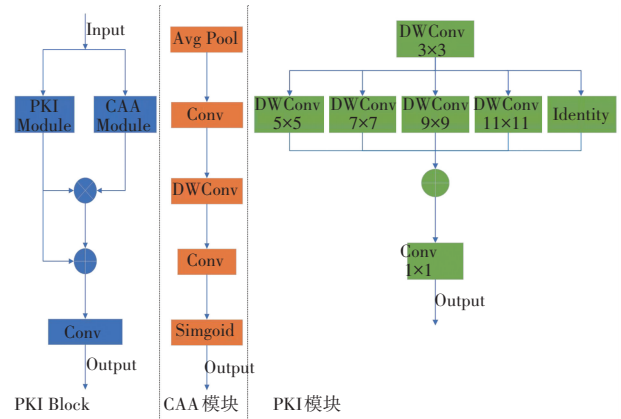


图 5 PKI Block 结构图

Fig. 5 Network structure of the PKI Block

1.3 Dynamic head 检测头

Yolov5 采用的检测头有一些缺点,例如:检测头各个预测位置都是互相独立的,缺少全局视野。仅仅从某一个尺度进行预测,没有考虑其他尺度特征对预测的影响。预测头参数量较小,表达能力一般。而 Dynamic head 有较好的动态学习能力,它选用动态路由机制,很好地融合了上下文信息并动态调整了不同特征层的权重,从而考虑到了其他尺度特征对预测的影响。并且它在不增加计算量的情况下提升了头部表达能力,解决了原检测头表达能力不足的缺点。

为提升检测头的表达能力和多尺度特征检测能力,本文将 Yolov5 检测头更换为 Dynamic head 检测头。Dynamic head 检测头则由多个 Dyhead 模块组成,如图 6 所示。

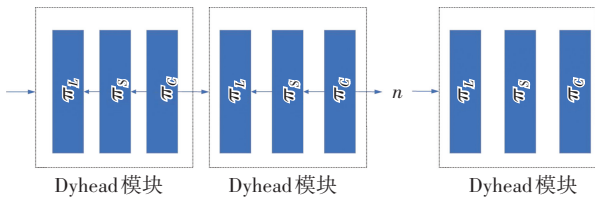


图 6 多个 Dyhead 模块串联

Fig. 6 Multiple Dyhead modules in series

Dyhead 模块将尺度注意力、空间注意力和任务注意力结合在同一个模块内,式(6)为该注意力计算公式:

$$W(F) = \pi_c(\pi_s(\pi_L(F) \cdot F) \cdot F) \cdot F \quad (6)$$

式(6)中, F 是输入的三维张量 $L \times S \times C$, π_c, π_s, π_L 分别代表任务注意力、空间注意力和尺度注意力模块,并分别作用到 F 的 C, S, L 维度上。Dyhead 结构图如图 7 所示。

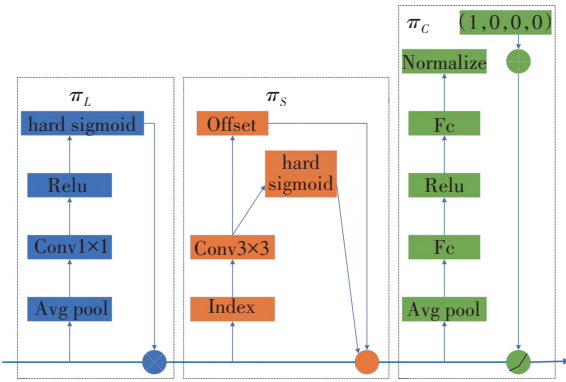


图 7 单个 Dyhead 模块结构图

Fig. 7 Structure of a single DyHead module

1.4 NWD-WIOU 损失函数

本文采用的钢材表面缺陷数据集中的缺陷像素较小,并且它存在难易样本不平衡的问题。而 Yolov5 的 CIOU 损失函数对像素较小的缺陷表现不佳并且它也没有考虑难易样本的平衡问题。针对上述问题,本文选择用 NWD 损失函数加权结合 WIOU 函数(简称 NWD-WIOU)来改进 CIOU 函数,从而达到提升精度的效果。

(1) NWD 损失函数。IOU 主要是用来衡量预测框和真实框的交合程度,属于判断模型好坏的一个指标。但是 IOU 对低像素目标的位置偏差敏感性过高,低像素缺陷轻微的位置偏差就可以引起 IOU 的大幅下降,这就代表 IOU 度量对离散位置偏差的目标尺度会发生变化。从而给缺陷检测带来干扰。因此,IOU 对低像素缺陷检测的效果并不太好,所以本文采用 NWD 损失函数对 CIOU 进行改进。相较于传统的 IOU 度量指标,NWD 采用的度量指标对低像素缺陷聚集产生的重叠具有良好的适应能力,并且它对各个尺度目标的位置

偏差不敏感。同时,NWD 可以应用于单级和多级锚式探测器。这些优点都解决了原 CIOU 函数不适应缺陷像素较低等问题。NWD 损失函数计算过程如下:在对边界框的高斯分布采用二次建模之后,将两个二维高斯分布 $\mu_1 = N(m_1, \Sigma_1)$ 和 $\mu_2 = N(m_2, \Sigma_2)$ 之间的二阶 Wasserstein 距离定义为 $W_2^2(\mu_1, \mu_2) = \|m_1 - m_2\|_2^2 + Tr(\Sigma_1 + \Sigma_2 - 2(\frac{\Sigma_1 \Sigma_2 \Sigma_1}{2})^{\frac{1}{2}})$,经化简可得式(7):

$$W_2^2(\mu_1, \mu_2) = \|m_1 - m_2\|_2^2 + \|\Sigma_1^{\frac{1}{2}} - \Sigma_2^{\frac{1}{2}}\|_F^2 \quad (7)$$

式(7)中, $\|\cdot\|_F$ 表示弗罗贝尼乌斯范数。除此之外,对于边界框建模的高斯分布 N_a 和 N_b ,但 $W_2^2 = (N_a, N_b)$ 是距离度量,不能直接用于相似度,故对其指数归一化得到一个度量叫做归一化的 Wasserstein 距离,化简得:

$$NWD(N_a, N_b) = \exp\left(-\frac{\sqrt{W_2^2(N_a, N_b)}}{c}\right) \quad (c \text{ 是常数})$$

最终 $L_{NWD} = 1 - NWD(N_a, N_b)$ 。

(2) WIOU 损失函数。CIOU 函数并未考虑样本难易不平衡问题,这容易导致模型专注于钢材表面缺陷的简单样本,从而降低精度。因此本文引入 WIOU 损失函数来解决该问题。该损失函数采用动态非单调聚焦的梯度增益分配方法,目的在于降低高质量锚框的竞争劣势,并减少低质量锚框的不利梯度影响,从而平衡简单样本与困难样本的权重,提升模型性能。WIOU 的计算过程可表示为

$$L_{WIOU} = R_{WIOU} L_{IOU} \quad (8)$$

$$R_{WIOU} = \exp\left(\frac{(x - x_{gt})^2 + (y - y_{gt})^2}{(W_g^2 + H_g^2) *}\right) \quad (9)$$

其中, L_{IOU} 是普通损失函数值, R_{WIOU} 表示的是该函数的惩罚项。当 R_{WIOU} 在 $[1, e]$ 区间上时,提升普通锚框的 L_{IOU} ,当 L_{IOU} 在区间 $[0.1]$ 上时,减少高质量锚框的 R_{WIOU} 。 (x, y) 为目标框真实位置, (x_{gt}, y_{gt}) 则是预测框的对应位置, H_g 和 W_g 则代表目标框和预测框组成的最小包围框的高和宽。

因此, 本文将 NWD 和 WIOU 加权结合, 得: $L_{NWD-CIOU} = (1 - w_n) \cdot L_{NWD} + w_n \cdot L_{WIOU}$

以最终的损失函数来代替 CIOU, 其中 w_n 指加权系数,作用是控制两类损失函数的占比, w_n 越大代表 NWD 损失函数占比越小。不同数据集选取的加权系数也是不同的,一般在小目标占比多的数据集选用较小的加权系数值,反之则取较大的加权系数值,本文将在通过实验来寻找合适的系数。

2 实验设计与结果分析

2.1 实验数据集

本次实验采用权威的 NEU-DET 数据集进行实验。该数据集包括划痕(Scratches, Sc)、裂纹(Crazing, Cr)、夹杂物(Inclusion, In)、带麻点的表面,简称麻面(Pitted Surface, PS)、轧制氧化皮(Rolled-in Scale, RS)和斑块(Patches, Pa)六种缺陷。其中部分缺陷图片如图 8 所示。数据集有 1 800 张图片,本文对数据集按 6:2:2 的比例进行均匀随机划分,得到训练集 1 080 张,验证集、测试集各 360 张。

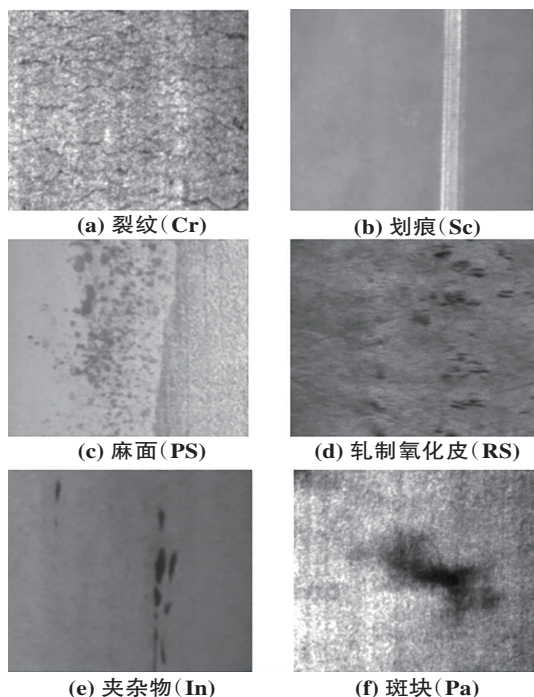


图 8 NEU-DET 六种缺陷示例

Fig. 8 Six examples of defects in NEU-DET

2.2 实验环境与参数配置

本文的实验在 Windows10 系统下进行,深度学习框架为 Pytorch1.12.1,Python 版本为 3.8.13,训练轮次(epoch)为 300 轮,初始学习率是 0.001,衰减系数为 0.000 5;动量为 0.937,批次为 16。实验相关配置和实验参数如表 1 所示。

表 1 实验的配置与参数

Table 1 Experimental configuration and parameters

名称	配置	参数	参数值
GPU	RTX3080Ti	图片像素	640×640
CPU	AMD EP-YC 9754	学习率	0.01
CUDA	11.3	衰减系数	0.005
Pytorch	1.11.1	批次	16

2.3 评价指标

本次实验通过准确率 Precision(简称 P)、召回率

Recall(简称 R)、平均精度 L_{mAP} 来评价模型性能。准确率和召回率的定义如下:

$$P = \frac{N_{TP}}{N_{TP} + N_{FP}} \quad (10)$$

$$R = \frac{N_{TP}}{N_{TP} + N_{FN}} \quad (11)$$

式(10)、式(11)中, N_{TP} 指的是预测为正样本且预测准确的样本个数, N_{FP} 为预测为正样本但实际上是负样本的样本个数, N_{FN} 是原本为正样本却被分为负样本的样本个数。

而 L_{mAP} 则是对数据集所有种类的 L_{AP} 求均值(L_{AP} 为每个类别预测的准确率,可以表示为 $L_{AP} = \int_0^1 P(R) dR$),表达式如下:

$$L_{mAP} = \frac{\sum_{i=1}^n L_{AP_i}}{n} \quad (12)$$

式(12)中, n 代表数据集中的种类个数,本实验为 6。 $L_{mAP@0.5}$ 是指 IOU 阈值取 0.5 时 L_{mAP} 的值, $L_{mAP@0.5:0.95}$ 表示在区间 $[0.5, 0.95]$ 上以 0.05 为步长取 10 个 IOU 阈值,再分别计算这些阈值下的 L_{mAP} 值后取它们的均值。 L_{mAP} 越高则表示模型在该数据集上的表现越好。

2.4 实验结果与分析

2.4.1 添加注意力机制的对比实验

为了体现 GJ-CA 注意力机制的优越性,本文将添加了 SE、CBAM 和 CA 注意力机制的 Yolov5 在 NEU-DET 上进行了对比实验,实验结果如表 2 所示。从表中可以看出,4 种注意力机制对精度均有提升,GJ-CA 提升效果最好, $L_{mAP@0.5}$ 可达 75.1%,比 Yolov5 高出了 1.5%,比 CA 注意力机制也提升了 0.5%,SE 效果最差, $L_{mAP@0.5}$ 只有 74.0%,仅仅比 Yolov5 上涨了 0.4%。

表 2 添加不同注意力的对比实验

Table 2 Comparative experiments of different

模型	attention mechanisms				/%
	P	R	$L_{mAP@0.5}$	$L_{mAP@0.5:0.95}$	
yolov5	73.2	69.3	73.6	40.7	
SE	73.8	68.6	74.0	40.5	
CBAM	72.5	70.7	74.2	39.8	
CA	73.9	70.1	74.6	40.8	
GJ-CA	74.5	70.4	75.1	41.2	

2.4.2 串联 Dyhead 模块个数对比实验

上一章节提到 Dynamic head 是由多个 Dyhead 模块串联组成。但是不同的串联个数会导致精度的不同。为了研究 Dyhead 模块的有效性和其串联个数对精度的影响,本文在颈部网络后面分别连接不同数量的 Dyhead 模块进行实验,结果如表 3 所示。

表 3 不同个数 Dyhead 模块对比实验
Table 3 Comparative experiment of different numbers of Dyhead modules /%

数量	P	R	$L_{mAP@0.5}$	$L_{mAP@0.5:0.95}$
0	73.2	69.3	73.6	40.7
1	72.8	71.7	74.8	40.9
2	73.5	72.6	76.4	41.7
3	74.4	73.6	77.7	43.3
4	74.0	73.5	77.1	43.4
6	74.3	73.9	77.9	43.6
8	73.2	74.3	76.8	42.1

从表 3 中可以得出串联数为 6 时精度最高, $L_{mAP@0.5}$ 可达 77.9%, 串联个数为 3 的 $L_{mAP@0.5}$ 值也高达 77.7%, 仅仅比 6 个模块串联的情况低 0.2%, 考虑串联 6 个 Dyhead 模块会大大增加模型的计算量(串联 6 个模块的 GFlops 比串联 3 个模块的 GFlops 多出了近 27%), 且没有比串联 3 个 Dyhead 模块提高太多精度, 故本文采用串联 3 个 Dyhead 模块的 Dynamic head 检测头作为改进后的检测头。

2.4.3 NWD-WIOU 损失函数对比实验

(1) 不同加权系数 w_n 对性能的影响。为探究不同加权系数 w_n 对性能的影响, 本文进行了 8 组实验, 分别选取了 8 个不同的 w_n 值在相同的实验条件下来对比精度, 结果如表 4 所示。

从表 4 中可以看出 w_n 为 0.6 时在钢铁缺陷检测中效果最好, $L_{mAP@0.5}$ 高达 76.5%, 比 CIUO 损失函数提升了 2.9%。而取值为 0 时的效果比 CIUO 损失函数更差, $L_{mAP@0.5}$ 只有 73.1%。因此本文取 w_n 的值为 0.6, 以获得较好的模型性能。

(2) 不同损失函数对比实验。为了体现 NWD-CIOU 相较于传统 IOU 损失函数给模型带来的提升, 本文将 NWD-CIOU 与一些传统 IOU 损失函数进行了对比, 实验结果如表 5 所示。

表 4 不同 w_n 对精度影响

Fig. 4 Impact of different values of w_n on accuracy /%

w_n	P	R	$L_{mAP@0.5}$	$L_{mAP@0.5:0.95}$
1.0	73.2	69.3	73.6	40.7
0.8	73.8	71.4	75.3	42.1
0.7	74.4	72.3	75.9	42.4
0.6	74.9	72.2	76.5	42.7
0.5	75.0	71.4	76.2	42.2
0.4	74.4	70.6	75.5	41.4
0.3	73.9	69.9	74.1	40.5
0.0	74.0	68.7	73.1	39.8

表 5 不同损失函数对模型性能的影响

Table 5 Effect of different loss functions on model performance /%

损失函数	P	R	$L_{mAP@0.5}$	$L_{mAP@0.5:0.95}$
CIUO	73.2	69.3	73.6	40.7
SIUO	72.7	69.9	73.9	39.8
GIOU	71.1	66.8	71.8	39.1
WIOU	74.6	70.2	75.1	41.6
NWD-CIOU	73.0	71.5	75.3	41.5
NWD-WIOU	75.1	72.2	76.5	42.7

从表 5 中很容易看出在传统的 IOU 损失函数中 WIOU 效果最好, 它在 IOU 阈值为 0.5 时的 L_{mAP} 比原来的 CIUO 高出了 1.5%。将 NWD 与 WIOU 结合后, 模型的平均精度相较于原来又提升了 2.9%, 比主流的 IOU 损失函数对性能提升效果更好。这说明本文改进的合理性, 引入 NWD-WIOU 大大提升了模型的性能。

2.4.4 消融实验

为了检验 GPDN-Yolov5 算法的有效性并确定每个改进对性能的影响, 本文在 NEU-DET 数据集上进行了消融实验, 实验结果如表 6 所示。

表 6 消融实验结果

Fig. 6 Results of ablation experiment /%

算法	GJ-CA	PKIBlo-ck	Dyhead 检测头	NDW-WIOU	P	R	$L_{mAP@0.5}$	$L_{mAP@0.5:0.95}$
Yolov5s					73.2	69.3	73.6	40.7
改进 1	✓				74.5	70.4	75.1	41.2
改进 2		✓			75.8	71.8	76.9	42.4
改进 3			✓		74.4	73.6	77.7	43.3
改进 4				✓	75.1	72.2	76.5	42.7
改进 5	✓	✓			76.4	72.1	78.6	42.7
改进 6	✓		✓		74.6	74.3	78.4	43.4
改进 7	✓	✓	✓		76.5	74.7	79.7	44.7
GPDN-Yolov5	✓	✓	✓	✓	76.4	76.6	80.5	45.9

从表 6 中的数据分析可以得出:改进 1 在骨干网络中加入了 GJ-CA 注意力机制,减少了背景因素对目标的干扰,从而提升了模型的性能,改进后的 $L_{mAP@0.5}$ 提升了 1.5%。改进 2 在 SPPF 后加入了 PKI Block 模块,加强了模型多尺度特征提取能力, $L_{mAP@0.5}$ 提升了 3.3%。改进 3 引入了 Dyhead 检测头,它提升了模型检测头的表达能力, $L_{mAP@0.5}$ 提高了 4.1%。改进 4 将 YOLOv5 原损失函数更换为 NWD-WIOU 损失函数, $L_{mAP@0.5}$ 提升了 2.9%。改进 5、改进 6 分别将改进 1 和改进 2,改进 1 和改进 3 进行两两组合,而改进 7 将改进 1、2、3 全部结合。改进 5、改进 6 和改进 7 的 $L_{mAP@0.5}$ 分别提高了 5%、4.8% 和 6.1%,均比前四种改进高,说明了改进的合理性。最终本文将所有改进都加到模型当中,模型准确率为 76.4%,上升了 3.2%,召回率达到 76.6%,上升了 7.3%,IOU 阈值为 0.5 和 0.5 到 0.95 下的平均精度分别达到 80.5% 和 45.9%,提高了 6.9% 和 5.2%。说明 GPDN-YOLOv5 对钢材表面缺陷检测有显著提升。

2.4.5 GPDN-YOLOv5 与其他主流算法对比

为了体现 GPDN-YOLOv5 在钢材表面缺陷检测上的优越性。本文在相同的实验环境下引入了 YOLOv7, YOLOv8, SSD 等主流算法并在 NEU-DET 数据集上进行了对比实验,实验结果如表 7 所示。从表 7 中可以看出,GPDN-YOLOv5 相较于其他主流算法在精度上有显著优势。

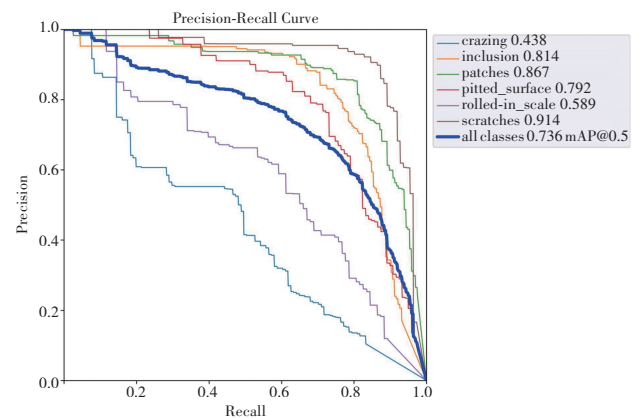
表 7 各算法性能指标对比

Table 7 Performance metric comparison of different algorithms /%

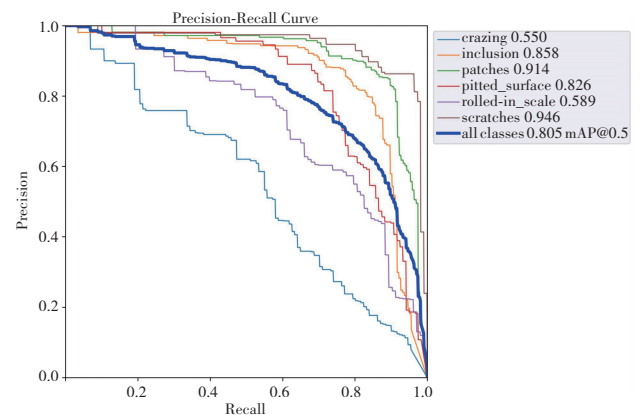
算 法	$L_{mAP@0.5}$	$L_{mAP@0.5:0.95}$
Faster-RCNN	74.5	41.1
SSD	71.8	38.8
YOLOv5	73.6	40.7
YOLOv7	74.7	41.9
YOLOv8	73.9	41.4
YOLOv9	75.4	42.2
RFB-YOLOv5-E	79.4	44.5
GPDN-YOLOv5	80.5	45.9

2.4.6 改进前后模型 P-R 图对比

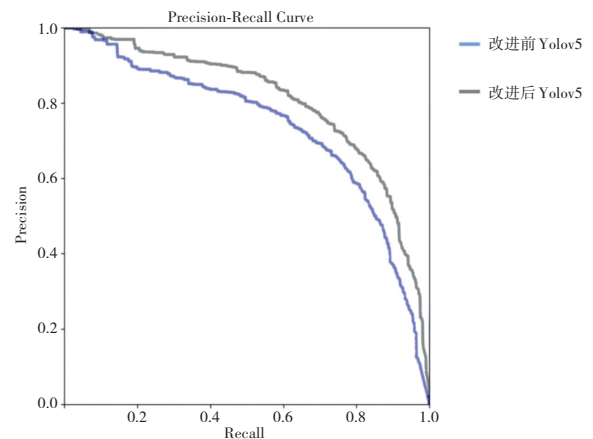
YOLOv5 改进前后算法在 NEU-DET 数据集上实验得出的 P-R 结果如图 9 所示。其中,y 轴代表准确率,x 轴则是召回率,曲线与坐标轴围成的面积则代表是模型的平均精度。



(a) 改进前模型的 P-R 曲线图



(b) 改进后模型的 P-R 曲线图



(c) 改进前后模型的 P-R 曲线对比图

图 9 改进前后模型 P-R 曲线对比

Fig. 9 Comparison of P-R curves of network models before and after improvement

从图 9 中可以看出改进后模型的曲线与坐标轴围成面积明显大于改进前的 YOLOv5 原模型。证明改进后的 YOLOv5 模型对于钢材表面缺陷检测的性能有极大的提升。

除了对比两者的 P-R 曲线图,本文还将改进前后模型的损失曲线进行了对比,如图 10 所示。从图中不难看出改进后 loss 的值下降更为明显并且始终低于改进前 loss。且当训练轮次达到 275 后曲线趋于平稳。

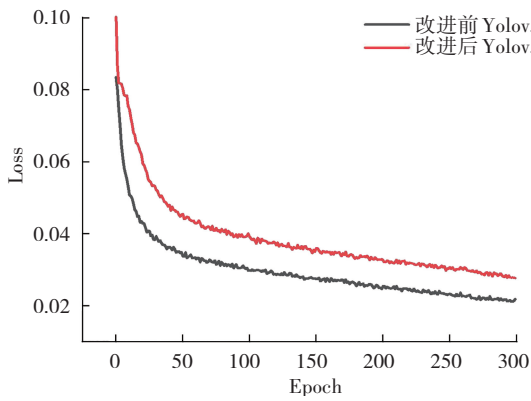
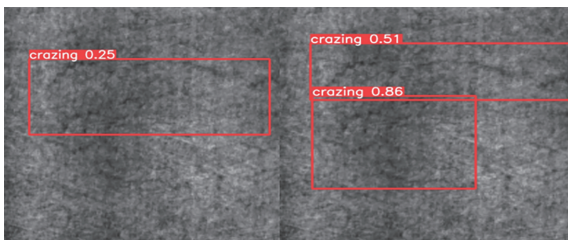


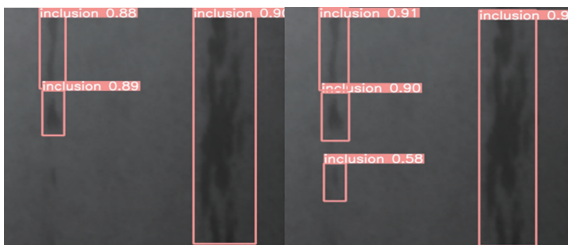
图 10 Loss 曲线对比

Fig. 10 Comparison of Loss curves

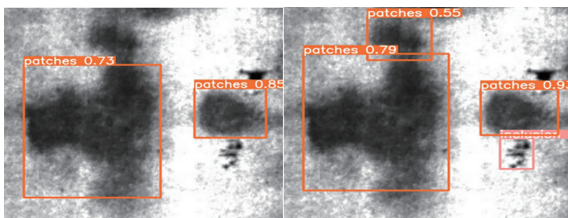
为进一步展示 GPDN-YOLOv5 算法的预测效果,本文把部分检测结果和原 YOLOv5 算法的检测结果进行了可视化对比,如图 11 所示,图 11 中左边是改进前的预测结果,右边是改进后的预测结果。



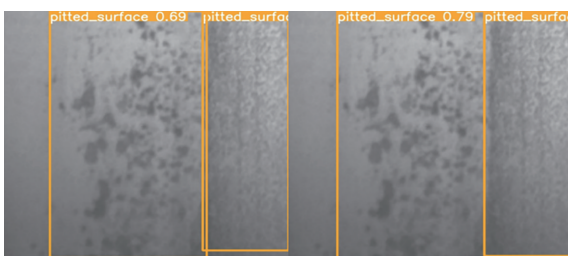
(a) 改进前后 crazing 检测效果对比



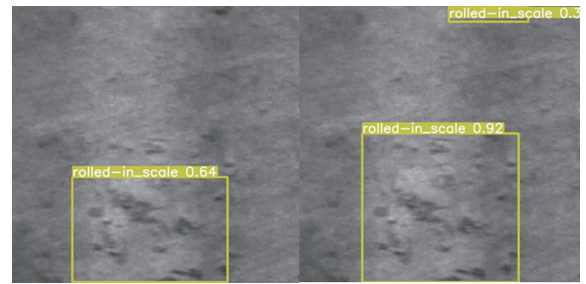
(b) 改进前后 inclusion 检测效果对比



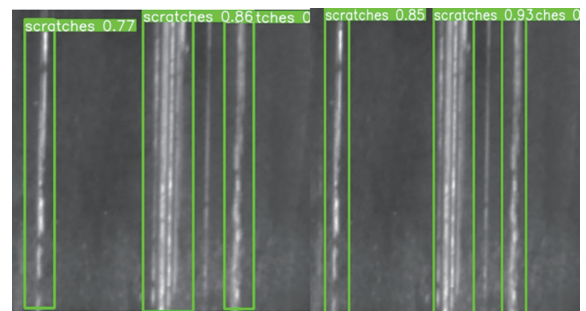
(c) 改进前后 patches 检测效果对比



(d) 改进前后麻面 (PS) 检测效果对比



(e) 改进前后氧化皮 (RS) 检测效果对比



(f) 改进前后划痕 (Sc) 检测结果对比

图 11 改进前后检测效果对比

Fig. 11 Comparison of detection results before and after improvement

从图 11 中的可视化对比可以得出, GPDN-YOLOv5 算法可以正确地检测出原 YOLOv5 算法中误检和漏检的缺陷,并且 GPDN-YOLOv5 算法对于六种缺陷的置信度均有一定程度的提高。这说明 GPDN-YOLOv5 算法对各种不同尺度的缺陷检测的精度都有很大提升。

2.4.7 算法的通用性对比实验

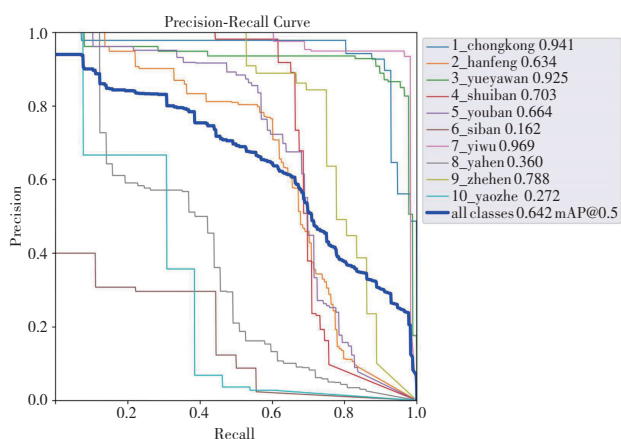
为了证明 GPDN-YOLOv5 算法在其他钢材表面数据集上同样有较好的提升,本文在 GC10-DET 数据集上将 GPDN-YOLOv5 和其他主流算法进行了对比,实验结果如表 8 所示,从表 8 可以看出, GPDN-YOLOv5 在该数据集上的表现优于其他主流算法,说明该算法在除 NEU-DET 以外的钢材表面缺陷数据集上也有不错的表现。

表 8 各算法在 GC10-DET 上的表现

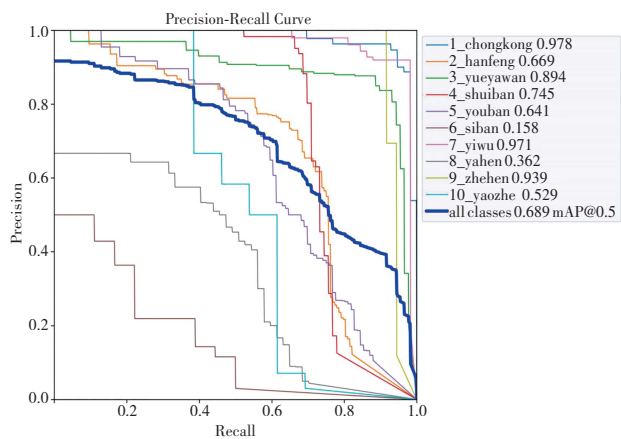
Table 8 Performance of different algorithms on the GC10-DET dataset

算法	GC10-DET dataset	
	$L_{mAP@0.5}$	$L_{mAP@0.5 : 0.95}$
Faster-RCNN	62.8	29.7
SSD	63.9	31.5
YOLOv5	64.2	31.4
YOLOv7	65.1	32.8
YOLOv8	64.3	32.9
YOLOv9	64.6	32.2
文献[14]	67.6	33.9
GPDN-YOLOv5	68.9	35.7

GC10-DET 数据集共有 3 570 张图片,包含虫孔、焊缝、月牙湾、水斑、折痕等 10 种 NEU-DET 数据集中没有的缺陷。为了更直观地对比改进前后的性能,验证 GPDN-Yolov5 的通用性,本文绘制了两者的 PR 曲线图(图 12),从图 12 中不难看出改进后的 PR 曲线与坐标轴围成的面积更大,并且改进后大多数缺陷的精度都有所提升。这说明了 GPDN-Yolov5 不仅对 NEU-DET 的六种缺陷的检测精度有所提升,还对其他一些常见缺陷的检测精度有所提升。



(a) 改进前 P-R 曲线图



(b) 改进后 P-R 曲线图

图 12 改进前后模型 P-R 曲线对比图

Fig. 12 Comparison of P-R curves of network models before and after improvement

3 结 论

钢材表面缺陷检测问题一直是工业和目标检测领域的热点问题之一。本文针对 NEU-DET 数据集中部分缺陷易受背景干扰和各个缺陷之间尺寸差异较大等问题,提出了一种基于多尺度特征提取的钢材表面缺

陷检测算法:GPDN-Yolov5。该算法相较于 Yolov5 的改进体现在以下四个方面,首先在骨干网络中加入 GJ-CA 注意力机制,减少了背景对缺陷的干扰。其次,在 SPPF 后加入 PKI Block 模块,加强网络的多尺度提取能力。随后,将 Yolov5 检测头更换为 Dyhead 检测头,加强了检测头的动态学习能力和表达能力。最后,用 NWD 损失函数和 WIOU 损失函数加权结合来代替 CIUO,从而解决算法专注于简单样本和对低像素缺陷偏差敏感的问题。实验结果表明:GPDN-Yolov5 在提升精度的同时也减少了算法的误检和漏检情况,比当今主流算法有着更优异的性能。

本文提出的 GPDN-Yolov5 算法虽然大大提升了钢材表面缺陷的检测精度,但算法的计算量却比原来略有提升,未来的研究可以进一步对算法进行轻量化,降低算法的计算量并加快算法的检测速度,保证算法既可以快速检测又能保持高精度。

参考文献(References):

[1] 黄如兵, 赵涟漪. 基于 YOLOv4 卷积神经网络的钢材缺陷检测系统[J]. 南阳理工学院学报, 2022, 14(4): 83-87, 98.
HUANG Ru-bing, ZHAO Lian-yi. Steel defect detection system based on YOLOv4 convolutional neural network[J]. Journal of Nanyang Institute of Technology, 2022, 14(4): 83-87, 98.

[2] 朱宏平. 基于卷积神经网络的钢表面缺陷检测方法[J]. 工业控制计算机, 2020, 33(8): 83-84.
ZHU Hong-ping. Detection method of steel surface defects based on convolution neural network[J]. Industrial Control Computer, 2020, 33(8): 83-84.

[3] 宋育斌, 孔维宾, 陈希, 等. 钢材表面缺陷检测研究综述[J]. 软件导刊, 2024, 23(3): 203-211.
SONG Yu-bin, KONG Wei-bin, CHEN Xi, et al. Survey of steel surface defect detection research[J]. Software Guide, 2024, 23(3): 203-211.

[4] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]//Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition. New York: ACM, 2014: 580-587.

[5] HE K, ZHANG X, REN S, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition[C]//Proceedings of

- the IEEE Transactions on Pattern Analysis and Machine Intelligence. Piscataway: IEEE Press, 2015: 1904–1916.
- [6] DAI J F, YI L, He K M, et al. R-FCN: Object detection via region-based fully convolutional networks[C]//In Proceedings of the 30th International Conference on Neural Information Processing Systems (NIPS' 16). New York: Curran Associates Inc, 2016: 379–387.
- [7] 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.
- [8] 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. Piscataway: IEEE Press, 2016: 779–788.
- [9] 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]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2023: 7464–7475.
- [10] REDMON J, FARHADI A. YOLO9000: Better, faster, stronger[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2017: 6517–6525.
- [11] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single shot multibox detector[C]//Computer Vision-ECCV 2016: 14th European Conference, Amsterdam, October 11–14, 2016. Cham: Springer International Publishing, 2016.
- [12] 杨溢. 基于深度学习的钢材表面缺陷识别方法研究[D]. 武汉: 华中科技大学, 2020: 32–41.
YANG Yi. Research on Steel Surface Defect Recognition Method based on Deep Learning [D]. Wuhan: Huazhong University of Science and Technology, 2020: 32–41.
- [13] 李鑫, 汪诚, 李彬, 等. 改进 YOLOv5 的钢材表面缺陷检测算法[J]. 空军工程大学学报(自然科学版), 2022, 23(2): 26–33.
LI Xin, WANG Cheng, LI Bin, et al. Steel surface defect detection algorithm based on improved YOLOv5[J]. Journal of Air Force Engineering University (Natural Science Edition), 2022, 23(2): 26–33.
- [14] 吕秀丽, 卢海滨, 侯春光, 等. 改进 YOLOv5s 的钢材表面缺陷检测算法[J]. 化工自动化及仪表, 2024, 51(2): 301–309.
LV Xiu-li, LU Hai-bin, HOU Chun-guang, et al. Improved YOLOv5 algorithm for steel surface defect detection[J]. Control and Instruments in Chemical Industry, 2024, 51(2): 301–309.
- [15] 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, 2020, 69(4): 1493–1504.
- [16] HOU Q, ZHOU D, FENG J. Coordinate attention for efficient mobile network design[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2021: 13708–13717.
- [17] DAI X, CHEN Y, XIAO B, et al. Dynamic head: Unifying object detection heads with attentions[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2021: 7369–7378.
- [18] 付锦焱, 张自嘉, 孙伟, 等. 改进 YOLOv8 的航拍图像小目标检测算法[J]. 计算机工程与应用, 2024, 60(6): 100–109.
FU Jin-yi, ZHANG Zi-jia, SUN Wei, et al. Improved YOLOv8 small target detection algorithm in aerial images [J]. Computer Engineering and Applications, 2024, 60(6): 100–109.
- [19] HU J, SHEN L, SUN G. Squeeze-and-excitation networks[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2018: 7132–7141.
- [20] WOO S, PARK J, LEE J Y, et al. CBAM: Convolutional block attention module[C]//Computer Vision-ECCV 2018. Cham: Springer International Publishing, 2018: 3–19.
- [21] 张学锋, 王子琦, 汤亚玲. 基于 YOLO-CDF 神经网络的安全帽检测[J]. 重庆工商大学学报(自然科学版), 2022, 39(4): 32–41.
ZHANG Xue-feng, WANG Zi-qi, TANG Ya-ling. Helmet detection based on YOLO-CDF neural network[J]. Journal of Chongqing Technology and Business University (Natural Science Edition), 2022, 39(4): 32–41.

责任编辑:陈 芳