

基于ERI-YOLO的钢材表面缺陷检测算法

张胜伟¹, 曹洁^{1,2}, 梁浩鹏¹

(1. 兰州理工大学 计算机与通信学院, 甘肃 兰州 730050;
2. 兰州城市学院 信息工程学院, 甘肃 兰州 730070)

摘要: 针对钢材表面缺陷尺度变化大、分布分散造成的目标错检和漏检问题, 提出一种基于YOLOv7改进的钢材表面缺陷检测算法ERI-YOLO。首先, 通过构建高效的反向残差块(EIRB)改进高效聚合网络(ELAN), 使其兼顾动态全局建模与静态局部信息融合的优势, 从而增强模型对图像特征的提取能力。其次, 设计重校准特征金字塔网络(RCFPN), 在提升多尺度特征融合效果的同时降低了模型的参数量和计算量。最后, 引入动态非单调聚焦机制, 并采用辅助边界框计算损失的方法, 提出Inner-WIoU损失函数, 以提高预测框的精确率, 加速模型的收敛速度。为了验证改进算法的有效性, 在NEU-DET数据集和GC10-DET数据集上进行了实验。实验结果表明, 相较于YOLOv7, 改进算法的平均检测精度分别提升5.7%和4.5%, FPS分别达到94.4帧/s和90.2帧/s, 有效减少了错检和漏检现象。

关键词: 钢材表面缺陷检测; YOLOv7; 反向残差块; 多尺度特征融合; 动态非单调聚焦机制

DOI: 10.11907/rjtk.242019

扫描二维码阅读全文:

中图分类号: TP391.41

文献标识码: A

文章编号: 1672-7800(2026)003-0204-08



Detection Algorithm of Steel Surface Defects Based on ERI-YOLO

ZHANG Shengwei¹, CAO Jie^{1,2}, LIANG Haopeng¹

(1. College of Computer and Communication, Lanzhou University of Technology, Lanzhou 730050, China;
2. College of Information Engineering, Lanzhou City University, Lanzhou 730070, China)

Abstract: Aiming at the problem of wrong detection and missed detection of steel surface defects caused by large scale variation and scattered distribution, an improved steel surface defect detection algorithm ERI-YOLO based on YOLOv7 is proposed. Firstly, by constructing Efficient Inverse Residual Block (EIRB), the Efficient Layer Aggregation Networks (ELAN) is improved, so as to enhance the model's ability to extract features from images by combining dynamic global modeling and static local information fusion. Efficient Layer Aggregation Networks (ELAN) by constructing an Efficient Inverse Residual Block (EIRB), which takes into account the advantages of dynamic global modeling and static local information fusion, so as to enhance the model's ability of extracting image features. Then, the Re-Calibration Feature Pyramid Network (RCFPN) is designed to improve the effect of multi-scale feature fusion while reducing the number of parameters and computation of the model. Finally, the dynamic non-monotonic focusing mechanism is introduced and the Inner-WIoU loss function is proposed by using an auxiliary bounding box to calculate the loss in order to improve the accuracy of the prediction box and accelerate the convergence speed of the model. In order to verify the effectiveness of the improved algorithm, experiments are conducted on the NEU-DET dataset and the GC10-DET dataset. The experimental results show that compared with YOLOv7, the average detection accuracy of the improved algorithm is improved by 5.7% and 4.5%, and the FPS reaches 94.4 frames/s and 90.2 frames/s, which effectively reduces the phenomena of misdetection and omission.

Key Words: steel surface defect detection; YOLOv7; inverse residual block; multi-scale feature fusion; dynamic non-monotonic focusing mechanism

收稿日期: 2025-03-10

基金项目: 国家自然科学基金项目(62463015)

作者简介: 张胜伟(1999-), 男, 兰州理工大学计算机与通信学院硕士研究生, 研究方向为目标检测、目标跟踪; 曹洁(1966-), 女, 兰州理工大学计算机与通信学院教授、博士生导师, 研究方向为智能信息处理、模式识别理论及应用; 梁浩鹏(1995-), 男, 博士, 兰州理工大学计算机与通信学院讲师、硕士生导师, 研究方向为目标检测、故障诊断。本文通讯作者: 曹洁。

0 引言

钢材作为国家基础设施建设和工业现代化推进中不可或缺的基础材料,在国防、航天、建筑等多个领域有着广泛应用。然而,在钢材的生产过程中,其表面常常会出现裂纹、斑块和划痕等各种形式的缺陷,这些缺陷严重影响了产品的质量与安全性。因此,实现高精度的钢材表面缺陷检测已成为钢材生产中至关重要的一步。

随着深度学习理论的不完善和计算机性能的持续提升,基于深度学习的目标检测方法在缺陷检测领域得到了广泛应用。这些方法主要分为两类:一是采用启发式搜索算法生成候选框,将每个候选框作为网络模型的输入,然后利用分类器和线性回归模型完成缺陷分类与定位的双阶段检测算法,这类算法大多以R-CNN系列方法及其改进方法为代表,虽然在精度上表现出色,但其计算量大,检测速度较慢^[1];二是采用网络模型预测候选框的位置、大小和物体类别,将目标检测问题转化为从图像中直接提取检测框、类别概率和置信度的单阶段检测算法。这类算法大多以YOLO系列方法及其改进方法为代表,将缺陷检测视为一个回归任务,在分类的同时直接获取目标位置信息,实现端到端的缺陷检测^[2]。相较于双阶段缺陷检测,单阶段检测方法精度略低,对于小缺陷区域和低分辨率图像检测效果不佳,但其具有更少的参数量,检测过程中节省了大量时间,更符合实时性需求。

近年来,为了实现更高效的钢材表面缺陷检测,许多研究者开始将YOLO算法应用到该领域。Kou等^[3]采用无锚点特征选择机制以减少基于锚方法的计算时间并提高模型整体速度。Cheng等^[4]通过引入TBConv增强对不同类型表面缺陷的特征学习能力并利用ECA注意力机制提高模型融合浅层和深层信息的能力。曹义亲等^[5]提出一种改进YOLOv5的钢材表面缺陷检测算法,采用带残差边的SPP_Res特征金字塔结构以加快模型训练速度,增强模型特征提取能力;熊聪等^[6]引入SwinTransformer以捕获全局上下文信息并提取更多差异化特征,同时采用加权双向特征金字塔网络进行跨尺度特征融合。高春艳等^[7]通过融合级联注意力机制和解耦头重新设计YOLOv7检测头,提升复杂场景下模型的表征能力。Chu等^[8]引入了轻量级模块DWR与YOLOv8的特征提取模块C2f相结合,扩大了感受野,使得模型能够获取更广泛的上下文信息。Zhang等^[9]设计了多尺度特征增强块,通过集成多尺度特征融合以增强小目标检测层,提高了检测性能。尽管现有方法在一定程度上提升了钢材表面缺陷检测性能,但仍然存在精度和实时性无法兼顾的难题。此外,由于钢材表面缺陷尺寸小且分布位置分散,很容易发生错检和漏检现象。因此,对于钢材缺陷的检测能力还有待提高。

YOLO系列算法经历了不断的演变和发展,如今已推

出YOLOv8、YOLOv9^[10]、YOLOv10^[11]等最新版本。但是,YOLOv7^[12]凭借其稳定性以及对小尺度缺陷较强的适应性和灵活性,在许多特定任务中仍具有竞争力。Xiao等^[13]研究发现YOLO系列算法的性能提升主要依赖于它们在COCO数据集上的检测结果,而不是针对小目标检测的数据集,较新版本的YOLO算法并非始终表现出卓越的检测性能。Li等^[14]通过实验发现,与YOLO系列更新版本相比,YOLOv7在NEU-DET钢材缺陷检测数据集上的表现更好。通过优化和改进,能够进一步提升其对钢材缺陷的检测效果,满足工业生产中实时性、准确性和鲁棒性的需求。综上所述,本文选取YOLO系列中性能出色的YOLOv7作为基准算法进行改进,提出一种基于YOLOv7改进的钢材表面缺陷检测算法ERI-YOLO。

本文主要贡献如下:①结合PConv和多头自注意力的优势构建高效反向残差块(Efficient Inverse Residual Block, EIRB),利用EIRB对高效聚合网络ELAN进行改进,以提高感受野的扩展速度并增强模型的全局感知能力;②引入选择性边界聚合模块SAB设计一种重校准特征金字塔网络,根据特征图的不同分辨率和内容,自适应地调整特征的权重,从而更好地捕捉目标的多尺度特征;③为了提高模型对困难样本的泛化性能,并降低各尺度缺陷的漏检率,提出损失函数Inner-WIoU,使得模型更加适应各种尺度和类型的缺陷,从而在实际应用中取得更好的检测效果。

1 改进的ERI-YOLO模型

YOLOv7通过引入高效聚合网络、模型重参数化、辅助头训练和正负样本匹配等策略提升了模型性能,但其网络结构仍存在以下缺点:①在特征学习过程中对钢材表面缺陷的识别能力不足,容易导致错检和漏检;②在特征融合过程中忽略了不同分辨率特征之间的相互依赖关系,且网络结构较为复杂,导致模型参数大、检测速度慢;③采用的CIoU损失函数受纵横比的影响大,使其在处理困难样本时泛化能力差。针对以上问题,本文提出一种基于ERI-YOLO的钢材表面缺陷检测方法。首先,通过构建反向残差块重构主干网络,增强模型的特征表示能力;其次,设计重校准特征金字塔,增强边界信息和定位重新校准对象,同时提升模型检测速度;最后,为实现更快且更有效的回归效果,提出损失函数Inner-WIoU。改进后的钢材表面缺陷检测算法ERI-YOLO模型结构如图1所示,主要包括反向残差块重构的主干网络、基于重校准注意力的特征金字塔网络以及头部网络。后续将对模型的主要组成部分进行详细介绍。

1.1 反向残差块改进的ELAN-EB

为了结合卷积神经网络的高效性以模拟短距离依赖,以及Transformer的动态建模能力以学习长距离交互,本文

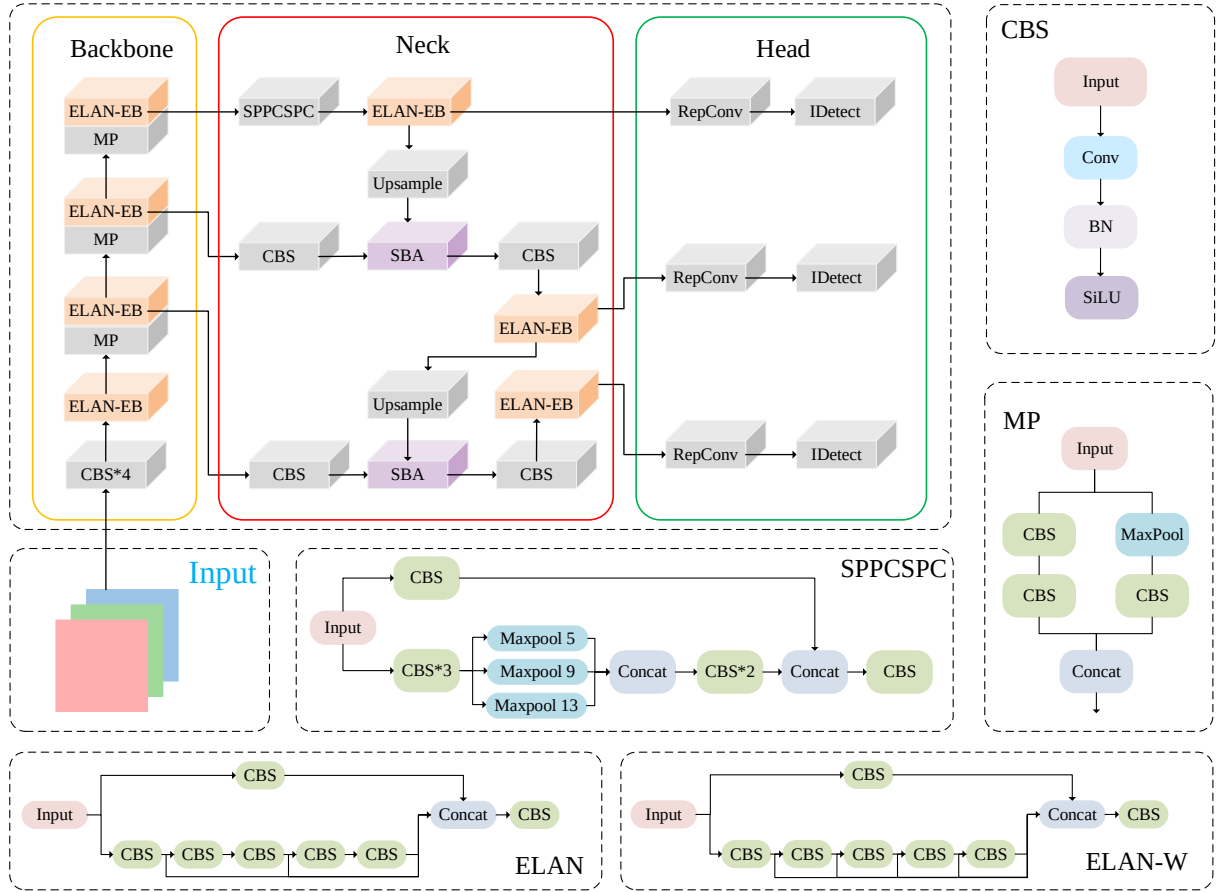


Fig. 1 ERI-YOLO network structure

图1 ERI-YOLO网络结构

借鉴 MobileNetV2^[15]中反向残差结构的思想,提出了一种高效的反向残差块 EIRB,如图2(a)所示。首先,给定输入特征 $X \in \mathbb{R}^{C \times H \times W}$,EIRB 通过 1×1 卷积进行升维操作得到 X_m ;其次,利用 PConv^[16]和多头自注意力增强图像特征得到 X_e ;最后,使用 1×1 卷积调整通道数得到 X_r 。为了避免梯度消失,模块的头尾加入了残差连接。整个过程相关公

式如下:

$$X_m = \text{Conv}_{1 \times 1}(X) \quad (1)$$

$$X_e = (\text{PWConv}(\text{PConv}), \text{Skip}) \\ (\text{MHSA}(X_m)) \quad (2)$$

$$X_r = \text{Conv}_{1 \times 1}(X_e) \quad (3)$$

$$Y = X + X_r \quad (4)$$

PConv对部分输入通道运用常规卷积进行空间特征提取,而对其余通道保持不变,不仅提高了模型精度,还有效减少了计算冗余和内存访问,如图3所示。在使用PConv时,特征通常具有短距离相关性,而多头自注意力具有动态权重和全局感受域,能够带来更多的长距离交互。在反向残差结构中同时采用这两者时,网络会得到更大的感受野,能够更好地建模远距离上下文信息。为了充分有效地利用来自所有通道的信息,进一步在PConv之后引入了逐点卷积(Pointwise Convolution, PWConv),以增强模型对特征图中心位置的关注度。

在YOLOv7主干网络中,ELAN结构由多个CBS模块堆叠而成,忽略了全局上下文信息,导致模型对目标的特征提取存在明显不足。因此,利用反向残差块 EIRB(见图2(a))对ELAN进行改进,将改进后的结构命名为ELAN-EB,如图2(b)所示。ELAN-EB通过两个 1×1 卷积对通道数进行调整,得到两个并行分支,其中一个分支保持不变,

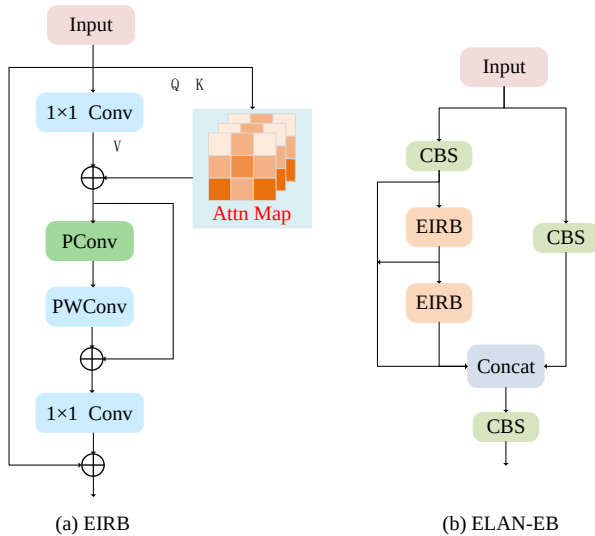


Fig. 2 EIRB and ELAN-EB structure

图2 EIRB和ELAN-EB结构

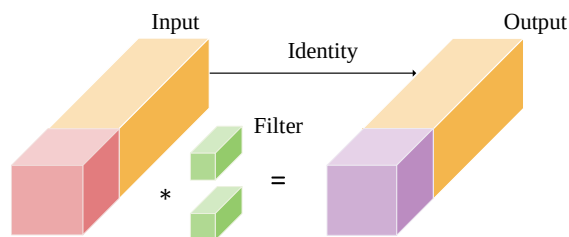


Fig. 3 PConv structure

图3 PConv结构

另一个分支经过两个EIRB残差块进行处理。最后,将所有分支的输出进行拼接。与ELAN相比,ELAN-EB省去了大量的卷积操作,降低了模型复杂度。改进后的模块有效扩大了模型感受野,并且更加关注特征图中的关键信息,增强了模型对钢材表面缺陷目标的特征提取能力。利用ELAN-EB对主干网络进行重构,得益于改进模块的优势,局部特征和全局表示得以最大限度地保留,从而使模型能够获取到更多丰富的语义信息。

1.2 重校准特征金字塔

YOLOv7采用的路径聚合网络对特征金字塔的特征融合策略进行了一定程度的优化,从而提高了目标检测效果。然而,由于输入特征的分辨率不同,路径聚合网络简单的融合策略会忽视不同分辨率特征之间的依赖关系,并导致特征冗余。此外,路径聚合网络增加了自底向上的路径,浅层特征信息会因网络层的加深而丢失,在钢材表面缺陷检测任务中,钢材表面缺陷包含的特征信息较少,更容易导致对缺陷的错检和漏检,而且额外的路径增加了计算复杂度和网络参数。针对上述问题,本文设计了一种重校准特征金字塔网络(Re-Calibration Feature Pyramid Network, RCFPN),如图4所示。首先,通过引入选择性边界聚合模块SBA,加强浅层特征和深层特征之间的交互能力^[17]。接着,将原模型中的ELAN-W替换为本文提出的ELAN-EB,以增强网络感知远距离位置信息和学习局部特征的能力。同时,去除原网络中自底向上的路径,以减少网络层深度,进一步降低了模型复杂度。

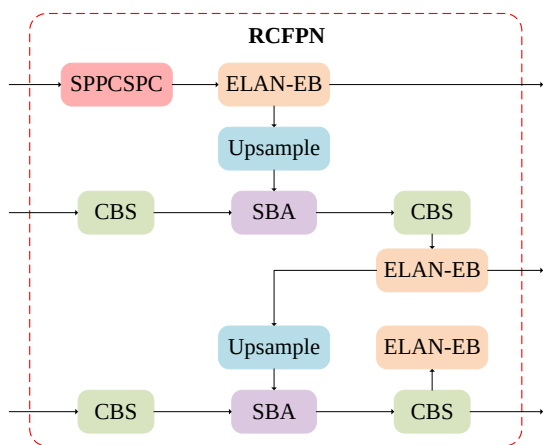


Fig. 4 Re-calibration feature pyramid network structure

图4 重校准特征金字塔网络结构

虽然浅层特征的语义信息较少,但其细节丰富且边界更分明。相比之下,深层特征则蕴含了丰富的语义信息。SBA模块通过聚合来自浅层特征的边界特征与深层特征的语义信息,从而有效地保留边界细节并重新校准钢材缺陷的位置,进一步提升了多尺度特征融合效果,如图5所示。具体而言,该模块首先将浅层与深层特征信息以不同的方式输入到两个重校准单元RAU中,RAU根据特征图的不同分辨率和内容,自适应地调整特征的权重,从而更好地捕捉目标的多尺度特征。接着,将两个RAU块的输出沿通道维度拼接,并输入一个具有批处理归一化和ReLU激活函数的3×3卷积层,以实现不同特征的鲁棒组合,并细化粗糙特征。RAU的处理过程表示为:

$$F'_h = W_\theta(F_h), F'_l = W_\phi(F_l) \quad (5)$$

$$RAU(F_h, F_l) = F_h' \odot F_h + F_l' \odot F_l \odot (\textcircled{R}(F_h')) + F_h \quad (6)$$

其中, $F_h \in \mathbb{R}^{C \times H_h \times W_h}$ 、 $F_l \in \mathbb{R}^{C \times H_l \times W_l}$ 分别为高分辨率输入特征和低分辨率输入特征。输入特征经过两个线性映射和 Sigmoid 激活函数 $W_\theta(\cdot)$ 、 $W_\phi(\cdot)$, 得到特征映射 F_h' 和 F_l' 。 $\textcircled{R}$ 是通过减去 F_h' 特征的反向操作,旨在将不精确和粗糙的估计细化为精确且完整的预测图。

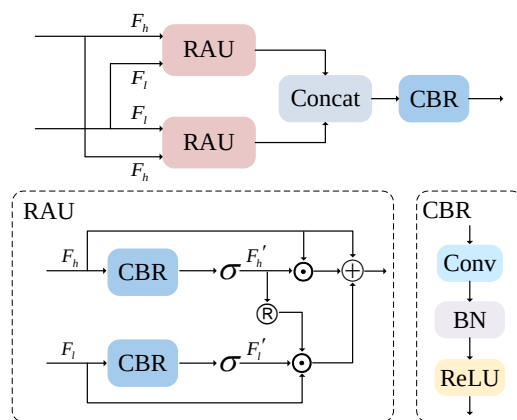


Fig. 5 SBA structure

图5 SBA结构

RCFPN通过高分辨率和低分辨率特征之间的双向融合机制,加强了网络通道间的信息传递,能够更好地融合多个位置和细节信息,解决了位置信息严重丢失的问题,在提高模型对钢材表面缺陷关注度的同时降低了模型参数量,加快了模型检测速度,相比于路径聚合网络更适用于钢材表面缺陷检测。

1.3 损失函数改进

YOLOv7的定位损失采用了CIoU损失函数进行计算,而分类损失和置信度损失则由二元交叉熵损失函数计算。其中,定位损失CIoU采用预测框与真实框宽和高的比值。然而,当两者的宽高比相同时,其惩罚项为0,函数存在一定的模糊性,得不到稳定表达,从而导致收敛速度较慢,并且CIoU存在处理难易样本不平衡的问题。为了解决上述问题,本文结合基于辅助边界框的损失函数 Inner-IoU^[18]

与采用动态非单调聚焦机制的边界框损失函数 WIoUv3^[19],提出了一种新的损失函数 Inner-WIoU,以替代 CIoU 作为改进算法模型的边界框损失函数。

Inner-IoU 引入尺度因子比,对辅助边界框大小进行调节,如图 6 所示。该损失函数通过使用较小尺度的辅助边界框计算 IoU 损失,从而有助于更好地回归高 IoU 样本,进而加速收敛过程。相反,使用更大尺度的辅助边界框计算 IoU,则能够有效促进低 IoU 样本的回归过程。具体计算公式如下:

$$b_l^{gt} = x_c^{gt} - \frac{w^{gt} * ratio}{2}, b_r^{gt} = x_c^{gt} + \frac{w^{gt} * ratio}{2} \quad (7)$$

$$b_t^{gt} = y_c^{gt} - \frac{h^{gt} * ratio}{2}, b_b^{gt} = y_c^{gt} + \frac{h^{gt} * ratio}{2} \quad (8)$$

$$b_l = x_c - \frac{w * ratio}{2}, b_r = x_c + \frac{w * ratio}{2} \quad (9)$$

$$b_t = y_c - \frac{h * ratio}{2}, b_b = y_c + \frac{h * ratio}{2} \quad (10)$$

$$inter = (\min(b_r^{gt}, b_r) - \max(b_l^{gt}, b_l)) * (\min(b_b^{gt}, b_b) - \max(b_t^{gt}, b_t)) \quad (11)$$

$$union = (w^{gt} * h^{gt}) * (ratio)^2 + (w * h) * (ratio)^2 - inter \quad (12)$$

$$IoU^{inner} = \frac{inter}{union} \quad (13)$$

$$L_{Inner-IoU} = 1 - IoU^{inner} \quad (14)$$

其中, (x_c^{gt}, y_c^{gt}) 分别代表真实框和内部真实框的中心点, (x_c, y_c) 分别代表预测框和内部预测框的中心点。 w^{gt} 和 h^{gt} 分别代表真实框的宽高, w 和 h 分别代表预测框的宽高。 $ratio$ 是尺度因子比,通常在 $[0.5, 1.5]$ 范围内。

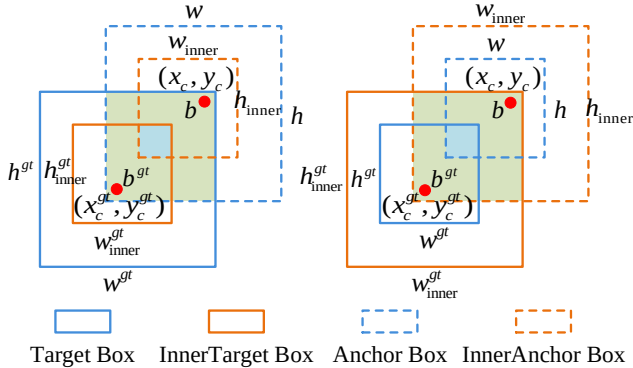


Fig. 6 Diagram of Inner-IoU

图 6 Inner-IoU 示意图

WIoUv3 采用动态非单调聚焦机制以屏蔽训练过程中的诸多负面影响。通过权衡低质量样本与高质量样本的比例,解决了小目标模糊不清、重叠遮挡目标难以检测的问题,从而进一步提高模型精度和模型检测整体性能。具体公式如下:

$$\mathcal{R}_{WIoU} = \exp\left(\frac{(x - x_{gt})^2 + (y - y_{gt})^2}{(W_g^2 + H_g^2)^*}\right) \quad (15)$$

$$L_{WIoUv1} = R_{WIoU} L_{IoU} \quad (16)$$

$$L_{WIoUv3} = r L_{WIoUv1}, r = \frac{\beta}{\delta \alpha^{\beta - \delta}} \quad (17)$$

本文结合上述两种损失函数的优势,提出 Inner-WIoU 损失函数,如式 (18) 所示。Inner-WIoU 一方面通过对离群度较大的预测框分配较小的梯度增益,以减少低质量样本训练产生的有害梯度,从而提升模型对小尺度缺陷的定位能力与回归精度;另一方面,通过引入辅助边界框,弥补了现有损失函数在不同数据集上泛化能力较弱且收敛速度较慢的不足。

$$L_{Inner-WIoU} = \exp\left(\frac{(x - x_{gt})^2 + (y - y_{gt})^2}{(W_g^2 + H_g^2)^*}\right) (1 - IoU^{inner}) \frac{\beta}{\delta \alpha^{\beta - \delta}} \quad (18)$$

其中, (x, y) 为预测框的中心点, (x_{gt}, y_{gt}) 为真实框的中心点。 w_g 和 h_g 是封闭锚框的最小尺寸。 r 为梯度增益, β 为离群度, α 和 δ 是可以调整以适应不同模型的超参数。

2 实验与结果分析

2.1 实验环境与参数设置

本文实验环境如表 1 所示。在训练过程中,采用 SGD 优化器对模型进行优化,动量设置为 0.937,迭代批次为 16,初始学习率设为 0.01,权重衰减设为 0.0005,输入图片大小为 640×640 ,训练轮数为 200。

Table 1 Experimental environment

表 1 实验环境

实验环境	配置
操作系统	CentOS Linux release 7.9.2009
CPU	Intel(R) Xeon(R) Platinum 8358P@2.60 GHz
GPU	Z100L
显存	32 GB
框架	Pytorch 1.10.0
加速环境	Cuda 10.2.89

2.2 数据集

本文采用东北大学提供的钢材表面缺陷数据集 NEU-DET 和 GC10-DET 数据集作为实验数据,验证改进方法的有效性^[20]。NEU-DET 数据集共有 1 800 张图片,图片分辨率为 200×200 像素。该数据集包含了 6 种典型的钢材表面缺陷,分别为:裂缝(Crazing)、夹杂(Inclusion)、斑块(Patches)、麻点(Pitted-Surface)、氧化皮(Rolled-in-Scale)和划痕(Scratches),每种缺陷有 300 个样本,每种缺陷的典型样本如图 7 所示。

GC10-DET 是在真实工业场景中收集的表面缺陷数据集,共有 2 294 张图片,包含 10 种类型钢板表面缺陷,分别为:冲孔(Punching-hole)、焊缝(Welding-line)、月形缝隙(Crescent-gap)、水斑(Water-spot)、油斑(Oil-spot)、丝斑(Silk-spot)、夹杂(Inclusion)、轧坑(Rolled-pit)、折痕(Crease)、腰部折痕(Waist-folding)。实验中,将数据集按照 8:1:1 划分成训练集、验证集、测试集。

2.3 评价指标

本文采用准确率(Precision)、召回率(Recall)、模型参

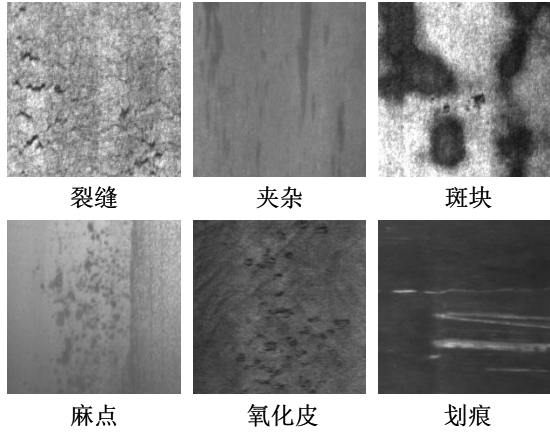


Fig. 7 Examples of deficiencies in each category of the NEU-DET dataset

图7 NEU-DET数据集各类缺陷示例

数量(Params)、模型计算量(GFLOPs)、所有类别平均精度均值(mean Average Precision, mAP)、每秒检测图片帧数(Frame Per Second, FPS)以衡量模型性能。这些指标的计算公式如下:

$$P = \frac{TP}{TP + FP} \quad (19)$$

$$R = \frac{TP}{TP + FN} \quad (20)$$

$$AP = \int_0^1 p(r) dr \quad (21)$$

$$mAP = \frac{1}{k} \sum_{i=1}^k AP_i \quad (22)$$

其中,TP表示模型检测结果为正的正样本,FP表示模型检测结果为正的负样本,FN为模型检测结果为负的正样本。平均精度(AP)代表模型对于某类缺陷的检测准确度,通过P-R图曲线下面积计算得出,平均精度均值是指所有类别的平均AP值,在多类别目标检测中用于衡量识别精度。

2.4 消融实验

为了验证本文各改进点的有效性,在NEU-DET数据集上进行了消融实验。首先,构建了高效残差块EIRB对ELAN模块进行改进,并通过改进的ELAN-EB重构主干网络;其次,设计了重校准特征金字塔对颈部网络进行更新;最后,提出了损失函数Inner-WIoU。针对上述改进,设计了7组实验,结果如表2所示,其中“√”表示使用了对应的改进方法,“-”表示未使用改进方法。

由实验1可知,加入ELAN-EB后,mAP提升了3.2%,表明ELAN-EB能够兼顾动态全局建模和静态局部信息融合的优势,从而减少错误背景信息干扰,增强主干网络的特征提取能力,并且ELAN-EB省去了原模型中大量的卷积操作,使得模型计算量降低了18.8 G。由实验2可知,加入重校准特征金字塔RCFPN后,mAP提升了1.5%,表明RCFPN通过特征间的双向融合机制,能够加强网络通道间的信息传递,提高网络的多尺度特征融合能力。此外,RCFPN去除了原网络中自底向上的路径,并将原模型中的

ELAN-W替换为ELAN-EB,使得模型参数量和计算量分别下降了3.1 M和13.4 G,从而提高模型检测速度。

同时,由实验3可知,加入损失函数Inner-WIoU后,网络模型收敛速度加快,使得mAP提升了0.9%,FPS提升了1.5 帧/s。Inner-WIoU的提出弥补了现有损失函数在钢材表面缺陷检测任务中泛化能力较弱且收敛速度较慢的问题。如图8所示,本文提出的损失函数相比YOLOv7使用的损失函数CIoU收敛速度更快,证明了Inner-WIoU的有效性。

Table 2 Results of ablation experiment

表2 消融实验结果

实验	ELAN-EB	RCFPN	Inner-WIoU	mAP@0.5/%	Params/M	GFLOPs	FPS
YOLOv7	-	-	-	75.6	37.2	105.2	114.2
1	√	-	-	78.8	39.7	86.4	74.6
2	-	√	-	77.1	34.1	91.8	97.8
3	-	-	√	76.5	37.2	105.2	115.7
4	√	√	-	80.8	36.8	75.6	92.5
5	√	-	√	79.0	39.7	86.4	76.7
6	-	√	√	77.6	34.1	91.8	99.1
7	√	√	√	81.3	36.8	75.6	94.4

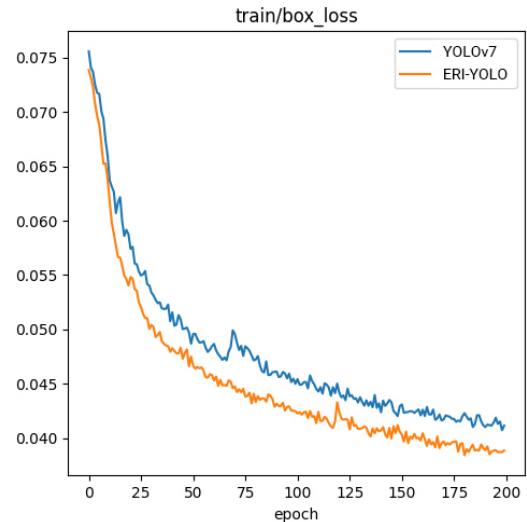


Fig. 8 Loss function comparison

图8 损失函数比较

最后,通过实验4—实验7验证了改进方法的可行性。本文对原模型进行ELAN-EB、RCFPN和Inner-WIoU共3处改进后,网络模型相比基准模型mAP提升了5.7%,参数量和计算量分别下降了0.4 M和29.6 G,处理速度达到了94.4 帧/s,能够满足实际工业环境需求。

2.5 对比实验

为了进一步验证本文算法的有效性,将改进算法与SSD^[21]、Faster-RCNN、YOLOX^[22]、YOLOv5、YOLOv7、YOLOv8、YOLOv9、YOLOv10这些主流的目标检测算法以及其他文献提出的改进算法CDN-YOLOv7、YOLO-SDS、GDM-YOLO进行了对比实验,结果如表3所示。

Table 3 Results of comparison experiment
表3 对比实验结果

数据集	网络模型	P/%	R/%	mAP@0.5/%	Params/M	GFLOPs	FPS
NEU-DET	SSD	63.6	71.3	71.5	24.4	88.4	47.6
	Faster R-CNN	69.1	73.5	73.9	136.8	183.6	20.7
	YOLOX	69.0	71.7	72.0	8.9	25.3	71.2
	YOLOv5	67.6	68.8	70.3	46.2	108.4	76.3
	YOLOv7	73.4	72.1	75.6	37.2	105.2	114.2
	YOLOv8	75.8	72.9	78.3	43.7	165.7	80.7
	YOLOv9	79.7	71.0	79.3	60.8	266.2	58.0
	YOLOv10	83.2	76.4	76.7	16.5	63.4	125.6
	CDN-YOLOv7	81.0	72.2	80.3	73.4	—	60.8
	YOLO-SDS	76.5	71.0	77.7	2.0	5.3	168.0
GC10-DET	GDM-YOLO	—	—	79.3	9.0	28.1	163.0
	Ours	82.9	76.3	81.3	36.8	75.6	94.4
	SSD	62.8	63.6	64.7	24.4	88.4	30.8
	Faster R-CNN	59.8	65.6	62.7	136.8	183.6	12.1
	YOLOX	69.6	60.3	65.3	8.9	25.3	68.6
	YOLOv5	61.5	63.4	64.2	46.2	108.4	70.1
	YOLOv7	63.0	68.7	66.9	37.2	105.3	116.5
	YOLOv8	67.6	63.1	68.3	43.7	165.7	80.7
	YOLOv9	68.9	65.6	69.8	60.8	266.2	52.0
	YOLOv10	71.7	64.8	69.1	16.5	63.4	118.0
	CDN-YOLOv7	52.3	59.2	56.7	73.4	—	—
	Ours	71.2	63.4	71.4	36.9	75.6	90.2

由实验结果可知,SSD的检测精度相对较低,主要是因为它使用低层级特征检测小尺寸缺陷,而这些特征卷积层数较少,可能导致一些目标被错检或漏检。双阶段算法Faster R-CNN参数量大、检测速度过慢,无法满足实时性要求。YOLOX采用IoU损失函数,虽然检测速度较快、模型参数量较少,但其检测精度较低,因为无法精确计算预测框与真实框的重合度。YOLOv5相比其他算法检测精度稍低,主要原因是其特征数据在不同维度之间的关联性不足,导致重要信息的权重未能得到充分体现,在深度卷积过程中可能会逐渐丢失,从而影响对小尺度缺陷的有效特征提取。尽管YOLOv8具有较快的检测速度,但其参数量较大,检测精度仍有待进一步提升。YOLOv9相比其他YOLO系列方法参数量大、检测速度慢,并且对于小尺度缺陷检测效果不佳。YOLOv10的性能高度依赖于训练数据的质量和多样性,钢材缺陷检测任务中训练数据较少,模型泛化能力会受到严重影响。此外,与其他文献提出的改进算法相比,本文改进算法在精确率、召回率、平均检测精度和FPS指标上均表现出显著优势。CDN-YOLOv7虽然检测精度较高,但其参数量大,并且检测速度远低于本文算法。YOLO-SDS和GDM-YOLO检测速度比本文算法更快,但其检测精度低于本文算法,没有实现精度和速度的平衡。

本文算法在NEU-DET和GC10-DET数据集上的平均检测精度分别达到了81.3%和71.4%,参数量分别降低至36.8 M和36.9 M,FPS分别达到94.4 帧/s和90.2 帧/s。综合各项评价指标来看,本文改进算法ERI-YOLO与众多主流

检测算法和改进算法相比表现更为出色,合理地平衡了检测精度与实时性,进一步说明了改进算法的可行性和有效性。

2.6 可视化分析

为了更加直观地对比本文算法与基准模型的检测效果,分别在NEU-DET和GC10-DET数据集上对原始算法YOLOv7和改进算法ERI-YOLO的检测结果进行了比较,如图9和图10所示。

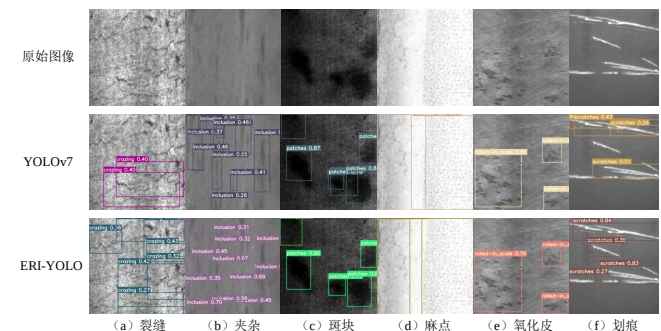


Fig. 9 Detection results on NEU-DET dataset before and after model improvement

图9 模型改进前后在NEU-DET数据集上的检测结果

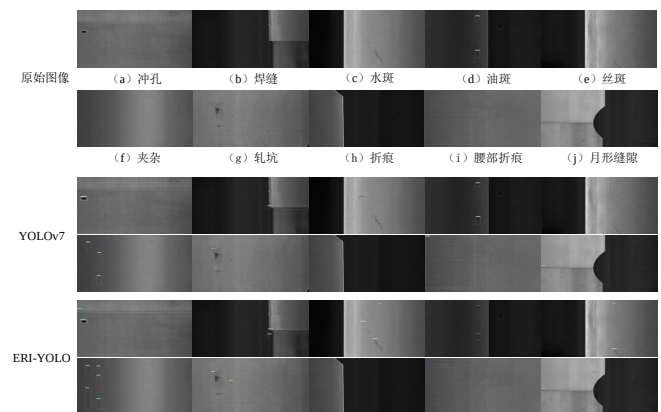


Fig. 10 Detection results on GC10-DET dataset before and after model improvement

图10 模型改进前后在GC10-DET数据集上的检测结果

由图9(a)裂缝和(f)划痕的检测对比结果可以看出,当钢材表面图像中缺陷过于密集或者目标较小时,YOLOv7存在漏检现象。而改进后的模型ERI-YOLO能够准确识别出图像中的小尺度目标,其对于各类缺陷的检测效果更佳,解决了YOLOv7对于钢材表面缺陷的误检与漏检概率,以及对小目标缺陷关注度不足的问题。从图10的检测结果可以看出,ERI-YOLO虽然在降低漏检和误检率方面表现出色,显著提高了模型稳定性和泛化能力,但对于(d)油斑、(f)夹杂等图像中存在的极少部分微小尺度缺陷检测效果仍不理想。由于这些缺陷在图像中占据的像素太少,分辨率过低,从而导致模型无法有效识别这些不易察觉的微小缺陷。因此,在未来工作中,对于微小缺陷的检测效果仍然有待提升。

3 结语

本文提出了一种基于ERI-YOLO的钢材表面缺陷检测方法,在实时推理的前提下保证了钢材缺陷的高精度检测。首先,通过构建反向残差块重构主干网络,使得模型能够更有效地建模远距离上下文信息,从而增强了全局感知的能力;其次,设计了重校准特征金字塔,提高了模型的特征融合能力;最后,提出了Inner-WIoU损失函数,进一步提升了检测速度和准确率。实验结果表明,本文方法在钢材表面缺陷检测数据集NEU-DET和GC10-DET上的平均检测精度分别达到了81.3%和71.4%,FPS分别达到94.4帧/s和90.2帧/s,在钢材表面缺陷检测任务中展现出良好的综合性能。然而,由于极少部分微小缺陷远小于常规缺陷,导致其有效信息不足。因此,模型在微小缺陷的检测精度方面仍存在提升空间。未来工作中,将着力提升微小缺陷检测效果。

参考文献:

- [1] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation [C]//Proceedings of the IEEE Conference On Computer Vision And Pattern Recognition, 2014:580-587.
- [2] 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.
- [3] KOU X, LIU S, CHENG K, et al. Development of a YOLO-V3-based model for detecting defects on steel strip surface [J]. Measurement, 2021, 182: 109454.
- [4] CHENG Y, WANG S. Improvements to YOLOv4 for steel surface defect detection [C]//2022 5th International Conference on Intelligent Autonomous Systems (ICoIAS), 2022:48-53.
- [5] CAO Y Q, WU M L, XU L. Steel surface defect detection based on improved YOLOv5 algorithm [J]. Journal of Graphics, 2023, 44 (2): 335-345.
曹义亲,伍铭林,徐露. 基于改进YOLOv5算法的钢材表面缺陷检测[J]. 图学学报, 2023, 44(2): 335-345.
- [6] XIONG C, YU A N, GAO X H, et al. Steel surface defect detection algorithm based on improved YOLOX [J]. Electronic Measurement Technology, 2023, 46(9): 151-157.
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.
- [7] GAO C Y, QIN S, LI M H, et al. Research on steel surface defect detection with improved YOLOv7 algorithm [J]. Computer Engineering and Applications, 2024, 60(7): 282-291.
- [8] 高春艳,秦荣,李满宏,等. 改进YOLOv7算法的钢材表面缺陷检测研究[J]. 计算机工程与应用, 2024, 60(7): 282-291.
- [9] CHU Y, YU X, RONG X. A lightweight strip steel surface defect detection network based on improved YOLOv8 [J]. Sensors, 2024, 24(19): 6495.
- [10] ZHANG T, PANG H, JIANG C. GDM-YOLO: a model for steel surface defect detection based on YOLOv8s [J]. IEEE Access, 2024, 12: 148817-148825.
- [11] WANG C Y, YEH I H, LIAO H Y M. YOLOv9: learning what you want to learn using programmable gradient information [DB/OL]. <https://arxiv.org/abs/2402.13616>, 2024.
- [12] WANG A, CHEN H, LIU L, et al. YOLOv10: real-time end-to-end object detection [DB/OL]. <https://arxiv.org/abs/2405.14458>, 2024.
- [13] 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, 2023: 7464-7475.
- [14] XIAO Y, DI N. SOD-YOLO: a lightweight small object detection framework [J]. Scientific Reports, 2024, 14(1): 25624.
- [15] LI L, ZHANG R, XIE T, et al. Experimental design of steel surface defect detection based on MSFE-YOLO—an improved YOLOv5 algorithm with multi-scale feature extraction [J]. Electronics, 2024, 13(18): 3783.
- [16] SANDLER M, HOWARD A, ZHU M, et al. Mobilenetv2: inverted residuals and linear bottlenecks [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 4510-4520.
- [17] CHEN J, KAO S, HE H, et al. Run, don't walk: chasing higher FLOPS for faster neural networks [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023: 12021-12031.
- [18] TANG F, XU Z, HUANG Q, et al. DuAT: dual-aggregation transformer network for medical image segmentation [C]//Proceedings of the Chinese Conference on Pattern Recognition and Computer Vision, 2023: 343-356.
- [19] ZHANG H, XU C, ZHANG S. Inner-IoU: more effective intersection over union loss with auxiliary bounding box [DB/OL]. <https://arxiv.org/abs/2311.02877>, 2023.
- [20] TONG Z, CHEN Y, XU Z, et al. Wise-IoU: bounding box regression loss with dynamic focusing mechanism [DB/OL]. <https://arxiv.org/abs/2301.10051>, 2023.
- [21] 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.
- [22] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot multibox detector [C]//Proceedings of the 14th European Conference on Computer Vision, 2016: 21-37.
- [23] GE Z, LIU S, WANG F, et al. YOLOX: exceeding YOLO series in 2021 [DB/OL]. <https://arxiv.org/abs/2107.08430>, 2021.

(责任编辑:孙娟)