

光电工程

Opto-Electronic Engineering

中文核心期刊 中国科技核心期刊
Scopus CSCD

基于YOLOv8-SOE的钢材表面缺陷检测

马玉洁, 曹维清, 张靖

引用本文:

马玉洁, 曹维清, 张靖. 基于YOLOv8-SOE的钢材表面缺陷检测[J]. 光电工程, 2025, 52(5): 250032.

Ma Y J, Cao C Q, Zhang J. Steel surface defect detection based on YOLOv8-SOE[J]. *Opto-Electron Eng*, 2025, 52(5): 250032.

<https://doi.org/10.12086/oe.2025.250032>

收稿日期: 2025-02-13; 修改日期: 2025-04-02; 录用日期: 2025-04-03

相关论文

改进轻量化的FCM-YOLOv8n钢材表面缺陷检测

梁礼明, 陈康泉, 陈林俊, 龙鹏威

光电工程 2025, 52(2): 240280 doi: 10.12086/oe.2025.240280

改进GBS-YOLOv7t的钢材表面缺陷检测

梁礼明, 龙鹏威, 卢宝贺, 李仁杰

光电工程 2024, 51(5): 240044 doi: 10.12086/oe.2024.240044

改进YOLOv8的焊缝表面缺陷检测算法

张润梅, 潘晨飞, 陈梓华, 陈中, 袁彬

光电工程 2025, 52(3): 240296 doi: 10.12086/oe.2025.240296

基于YOLOv8优化改进的太阳能电池片缺陷检测模型

彭自然, 王思远, 肖伸平

光电工程 2024, 51(11): 240220 doi: 10.12086/oe.2024.240220

更多相关论文见光电期刊集群网站 



光电工程
Opto-Electronic Engineering

<http://cn.ojournal.org/oe>



 OE_Journal



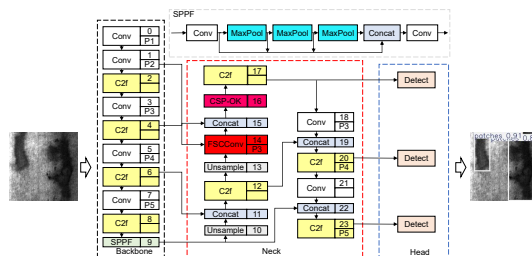
Website

基于 YOLOv8-SOE 的 钢材表面缺陷检测

马玉洁^{1,2}, 曹维清^{1,2*}, 张 靖²

¹安徽工程大学计算机与信息学院, 安徽 芜湖 241000;

²长三角哈特机器人产业技术研究院, 安徽 芜湖 241000



摘要: 为增强钢材表面检测中对小目标缺陷的检测能力, 提出一种改进的 YOLOv8-SOE 模型。该模型通过设计 FSCConv 模块来处理 P2 层特征, 通过压缩 P2 层特征并将其与 P3 层特征深度融合, 有效增强模型对小目标特征的敏感性, 同时避免引入额外检测层带来的计算负担。为进一步优化多尺度特征融合能力, 采用 CSP-OK (cross stage partial omni-kernel) 模块优化多尺度特征融合, 提高不同尺度特征的整合效率。通过引入 SiLU 损失函数优化边界框回归, 进一步提升定位精度。实验结果表明, YOLOv8-SOE 模型在 NEU-DET 数据集上的 mAP 达 80.7%, 较基准模型提升 5.4%, 且在 VOC2012 数据集上具有较强的泛化能力。该模型在提升小目标检测精度的同时, 保持较高的计算效率, 展现出良好的应用潜力。

关键词: YOLOv8; 缺陷检测; 小目标检测; 特征融合; 损失函数

中图分类号: TP391.4

文献标志码: A

马玉洁, 曹维清, 张靖. 基于 YOLOv8-SOE 的钢材表面缺陷检测 [J]. 光电工程, 2025, **52**(5): 250032

Ma Y J, Cao C Q, Zhang J. Steel surface defect detection based on YOLOv8-SOE[J]. *Opto-Electron Eng*, 2025, **52**(5): 250032

Steel surface defect detection based on YOLOv8-SOE

Ma Yujie^{1,2}, Cao Chuqing^{1,2*}, Zhang Jing²

¹School of Computer and Information, Anhui Polytechnic University, Wuhu, Anhui 241000, China;

²Yangtze River Delta HIT Robot Technology Research Institute, Wuhu, Anhui 241000, China

Abstract: In order to improve the detection capability of small target defects in steel surface inspection, an improved YOLOv8-SOE model is proposed. The model processes the P2 layer features by designing the FSCConv module. By compressing the P2 layer features and deeply fusing them with the P3 layer features, the model's sensitivity to small target features is effectively enhanced, while avoiding the computational burden caused by the introduction of additional detection layers. In order to further optimize the multi-scale feature fusion capability, cross stage partial omni-kernel (CSP-OK) module is used to optimize the multi-scale feature fusion, which improves the integration efficiency of features of different scales. The SiLU loss function is introduced to optimize the bounding box regression, which further improves the positioning accuracy. Experimental results show that the mAP of the YOLOv8-SOE model on the NEU-DET dataset achieves 80.7%, which is 5.4% higher than the baseline model, and

收稿日期: 2025-02-13; 修回日期: 2025-04-02; 录用日期: 2025-04-03

基金项目: 中央引导地方科技发展专项高水平新型研发机构定向项目 (S202407a12020270); 安徽省重点研究与开发计划高新领域项目 (202304a05020077)

*通信作者: 曹维清, caochuqing@ahpu.edu.cn。

版权所有©2025 中国科学院光电技术研究所

has good generalization ability on the VOC2012 dataset. While improving the accuracy of small target detection, the model maintains a high computational efficiency and has good application prospects.

Keywords: YOLOv8; defect detection; small object detection; feature fusion; loss function

1 引言

钢材作为全球最重要的工业基础材料之一,其质量直接影响建筑、航空航天、汽车制造等多个关键行业的安全性与经济效益,在现代工业中扮演着至关重要的角色^[1]。生产过程中的表面缺陷会导致产品质量下降、设备磨损和成本增加,因此高效、准确的缺陷检测成为钢材行业的重要需求。传统人工视觉检查效率低、鲁棒性差,难以识别复杂缺陷。机器视觉虽然优化了检测效率和可靠性,但依赖手工设计特征,难以处理大规模、多样化的钢材缺陷检测,尤其在面对不同尺寸、形状和光照条件时会受到明显限制。

随着深度学习技术的快速发展,基于深度学习的目标检测方法取得了显著进步。目标检测模型通常可以分为两大类:两阶段模型和一阶段模型。常见的两阶段目标检测模型包括 R-CNN^[2]、Fast R-CNN^[3] 和 Faster R-CNN^[4] 等。这些方法在工业缺陷检测领域中得到了广泛应用,展现出了极高的精度和效率。Shi 等^[5]提出了一种基于 Faster R-CNN 的改进网络,用于钢材表面缺陷检测,选择新型卷积神经网络 ConvNeXt 作为主干,采用卷积注意力模块(convolutional block attention module, CBAM)结构增强网络的检测能力,将 K-means 聚类算法与区域候选网络(region proposal network, RPN)相结合,以生成更准确的区域建议,但其检测速度仅为 26 f/s,难以满足实时检测的需求。两阶段模型虽然在准确性和速度之间达到了平衡,但在工业应用中,计算效率仍然是一个问题。

单阶段模型的代表则包括 SSD^[6]、RT-DETR^[7]、YOLO^[8-9] (you only look once) 系列和 RetinaNet^[10]。YOLO 系列模型,以其出色的实时检测性能和高精度,迅速成为目标检测领域的主流方法之一。Li 等^[11]提出了一种带钢表面缺陷检测算法 DEW-YOLO,虽然检测精度有所提高,但是与原始模型相比,改进后的模型计算量增加。Yang 等^[12]提出了一种改进的缺陷检测模型 CFE-YOLOv8s,该模型有效地解决了钢材表面缺陷检测的问题,但在应用于其他缺陷检测任务时仍然存在局限性。彭菊红等^[13]提出了一种 YOLOv8n-MDC 的钢材表面缺陷检测算法,对特征提取、目标定位和多尺度检测方面进行了优化,提高了钢材缺陷

检测的精度。Zhang 等^[14]提出了一种 ESI-YOLOv8 的钢材表面缺陷检测方法,该方法融合了大分离卷积注意力机制、空间金字塔池化策略,增加了多尺度特征融合以增强感受野和特征表达能力。但是还需要进一步提升模型检测的精度。Huang 等^[15]设计了一种名为 WFRE-YOLOv8s 的新型单级检测器,用于钢表面缺陷检测,但对于具有龟裂和卷入式刻度的缺陷类别来说,仍然存在检测精度低于其他类别的问题。

综上所述,为了优化钢材表面缺陷检测的检测精度和小目标的检测能力,本文提出了基于 YOLOv8-SOE (small object enhance, SOE) 的钢材表面缺陷检测方法。本文的主要工作如下:

1) 针对小目标检测问题,本文提出了一种 FSCConv (fine-grained spatial compression convolution,) 模块。该模块通过压缩 P2 层特征并将其融入 P3 层,有效增强了小目标检测的敏感性,同时避免了额外检测层带来的计算负担。

2) 为解决多尺度特征融合问题,本文结合跨阶段局部网络与 Omni-Kernel 的思想,设计了 CSP-OK (cross stage partial omni-kernel) 模块。该模块能够高效融合多尺度特征,在提升检测性能的同时避免了添加 P2 层导致的计算复杂度和后处理开销。

3) 在损失函数优化方面,本文引入 SIOU 损失函数,通过改进传统 IoU 在形状和方向处理上的不足,使检测框定位更加精确,适用于不同尺度目标的检测。

2 改进的 YOLOv8 算法

YOLOv8 凭借其改进的特征提取结构和强大的多尺度目标检测能力,在小目标检测和复杂环境中的目标识别中有显著优势。然而,对于钢材表面微小缺陷等小目标检测,常规检测层的分辨率较低,难以有效捕捉细节信息。传统的解决方案通过在检测头中增加 P2 层以增强小目标感知能力,虽然提高了检测精度,但也显著增加了计算开销和推理时间。为解决这一挑战,提出了一种新的目标检测模型——YOLOv8-SOE,专注于提升钢材表面的小目标检测性能。首先,提出 FSCConv 对 P2 层特征进行处理, FSCConv 能

够保留小目标的细节信息, 并且通过扩张卷积捕获多尺度上下文信息。这样能得到小目标信息丰富的 P2 特征, 不仅增强了算法对小目标的感知, 还避免了额外检测层引入的计算开销。接着, 将处理后的 P2 特征与 P3 层特征进行融合, 进一步提升小目标检测效果, 并且计算负担变化不大。为了在特征融合过程中进一步优化特征提取与传递, 结合跨阶段局部网络以及 Omni-Kernel^[16] 提出 CSP-OK 模块, 这种改进有助于减少网络的冗余信息, 并更有效地利用层间特征, 从而使得不同尺度特征的整合更加高效且保留了小目标的细节信息。此外, 为提升边界框预测的精度, 采用 SIOU^[17] 损失函数。与传统的 IoU 损失函数不同, SIOU 不仅关注预测框和真实框的重叠区域, 还综合考虑了两者之间的距离、角度偏差和形状相似性等因素, 从而更全面地评估匹配度, 使得目标定位的精度和鲁棒性显著提升。YOLOv8-SOE 模型的结构图如图 1。

2.1 FSCConv 模块

在钢材表面缺陷检测中, 诸如微裂纹、细小划痕等缺陷往往面积小且分布稀疏, 给检测任务带来了巨大挑战。这类小目标不仅难以在高层特征图上充分表达, 而且传统卷积神经网络 (CNN) 中的步长卷积和池化操作会导致空间信息的丢失, 限制了模型对细微缺陷的捕捉能力。为解决这些问题, 提出了 FSCConv 模块, 将空间信息紧凑地压缩到通道维度中, 从而提高网络对小目标的敏感度, 并减少冗余信

息计算。具体而言, 该模块的核心思想是将输入特征图的空间维度重组并压缩到通道维度。按块采样输入特征图的不同子区域, 并在通道维度上拼接, 可以保留更多细粒度信息。FSCConv 模块的结构图如图 2 所示。给定输入张量 $X \in \mathbb{R}^{C \times H \times W}$, C 表示通道数, H 和 W 分别表示输入的高和宽; 接着把输入张量 X 按照空间位置的奇偶性, 以步长为 2 进行采样, 将其划分为 X_1 、 X_2 、 X_3 、 X_4 四个子张量; 最后将四个子张量在通道维度上进行拼接, 拼接后的输出张量为 $Y \in \mathbb{R}^{4C \times \frac{H}{2} \times \frac{W}{2}}$, 其数学表达式为

$$Y = \text{Concat}(X_1, X_2, X_3, X_4), \quad (1)$$

式中: $\text{Concat}(\cdot)$ 表示拼接操作。经过 $\text{Concat}(\cdot)$ 操作后, 四个子张量的通道维度连接了起来, 在保留输入图像的全部信息的同时, 将空间信息压缩到通道中。

这一重组过程减小了空间维度的大小, 同时在通道中聚合了来自不同空间位置的上下文信息, 避免了步长卷积和池化层可能导致的细节损失。在钢材缺陷检测任务中, 该模块通过保留和提炼低层特征中的小目标信息, 在对小目标缺陷检测中表现出了卓越的提升。本文将 FSCConv 模块用于 P2 层特征的处理, 并与 P3 层特征进行融合。这一设计有效地增强了模型对小目标的感知能力, 同时避免了直接添加 P2 检测层引入的计算负担。经过融合后的特征图包含了多尺度的上下文信息, 在细微缺陷检测中展现出更高的检测精度。

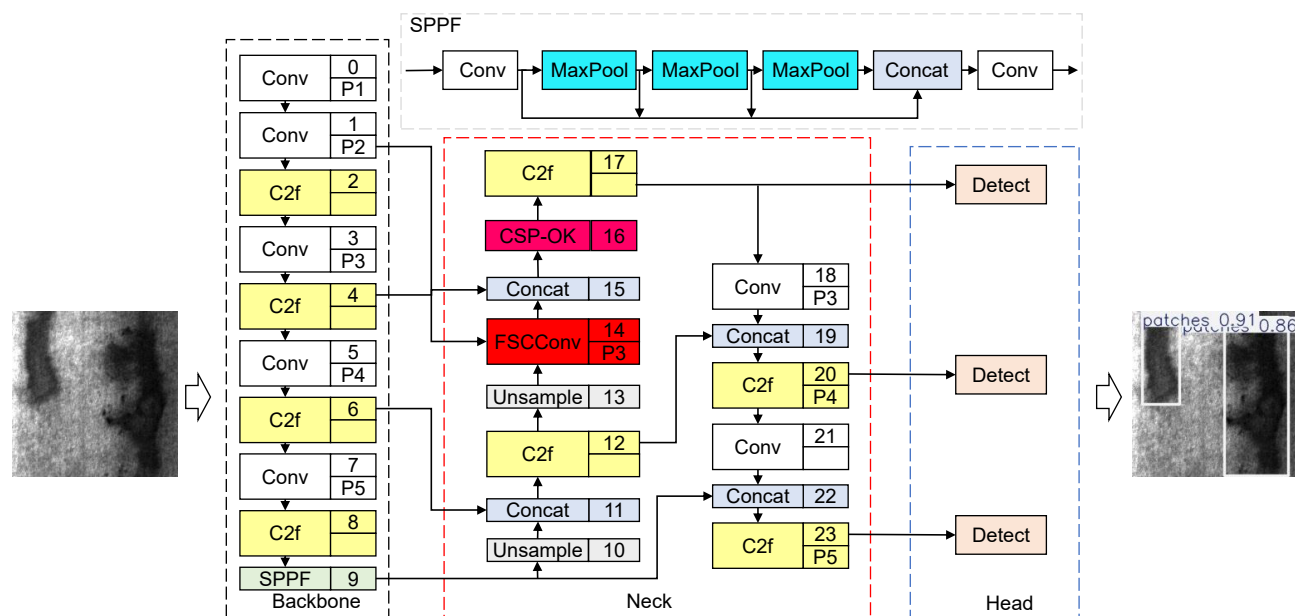


图 1 YOLOv8-SOE 模型结构图

Fig. 1 YOLOv8-SOE model structure diagram

2.2 CSP-OK 模块

为了进一步提升特征整合的效果, 提出了 CSP-OK 模块。CSPNet^[18] 思想的引入, 旨在减少网络中的冗余计算, 并通过分段处理来更有效地利用层间特征, 增强模型的计算效率和特征表征能力。在此基础上, 结合 Omni-Kernel 模块, 使得模型在处理从全局到局部的特征信息时更加高效。Omni-Kernel 由三个分支组成, 分别为全局分支、大分支和局部分支, 旨在高效学习从全局到局部的特征表示。Omni-Kernel 的示意图如图 3 所示。

为了提供多尺度的感受野, 大分支采用了一种计算代价较低的深度可分离卷积, 其卷积核大小为 $K \times K$, 用于获取更大的感受野。根据参考文献 [16], 当卷积核尺寸从 $K = 3$ 增加到 $K = 63$ 时, 模型的性能显著提升。然而, 为了在保证模型性能的同时降低计算开销和参数量, 在实际应用中选择了 $K = 31$ 的卷积核尺寸。这一尺寸能够在计算效率与感受野覆盖范围之间取得良好平衡, 并在大分支中实现多尺度特征的高效捕获, 确保模型性能不会大幅下降。

全局分支由双域通道注意力模块 (dual-domain

channel attention module, DCAM) 和基于频域的空间注意力模块 (frequency-based spatial attention module, FSAM) 组成。这两个模块协同工作, 以提高模型在大尺寸图像上的表现。下面依次介绍这两个模块的设计与实现。

给定输入特征 $X_{\text{Global}} \in \mathbb{R}^{C \times H \times W}$, 首先采用双域通道注意力模块对其进行频域通道注意力 (frequency channel attention, FCA) 操作, 具体过程为

$$X_{\text{FCA}} = \text{IFFT}(\text{FFT}(X_{\text{Global}}) \otimes W_{1 \times 1}^{\text{FCA}}(\text{GAP}(X_{\text{Global}}))), \quad (2)$$

式中: $\text{FFT}(\cdot)$ 和 $\text{IFFT}(\cdot)$ 分别表示快速傅里叶变换及其逆变换操作; X_{FCA} 、 $W_{1 \times 1}$ 和 $\text{GAP}(\cdot)$ 分别表示频域通道注意力模块的输出、 1×1 卷积层以及全局平均池化; $\otimes$ 表示逐元素乘法操作。基于傅里叶变换的处理, 利用频谱卷积定理, 有效优化和精炼了全局特征。在频谱域完成全局调制后, 进一步将特征传递至空间通道注意力模块中进行处理, 其过程为

$$X_{\text{DCAM}} = X_{\text{FCA}} \otimes W_{1 \times 1}^{\text{SCA}}(\text{GAP}(X_{\text{FCA}})), \quad (3)$$

式中: X_{DCAM} 表示 DCAM 的输出。DCAM 主要在通道维度上以粗粒度增强双域特征。接下来, 在空间维

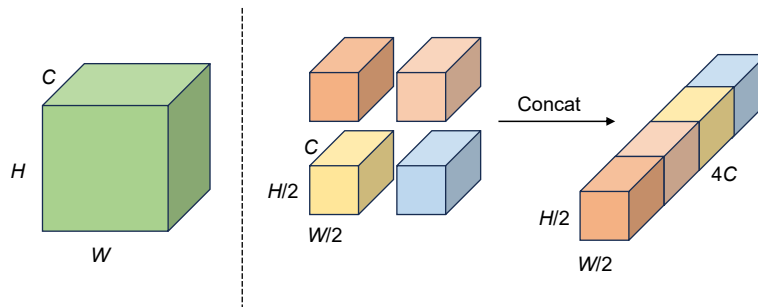


图 2 FSCConv 结构图

Fig. 2 FSCConv structure diagram

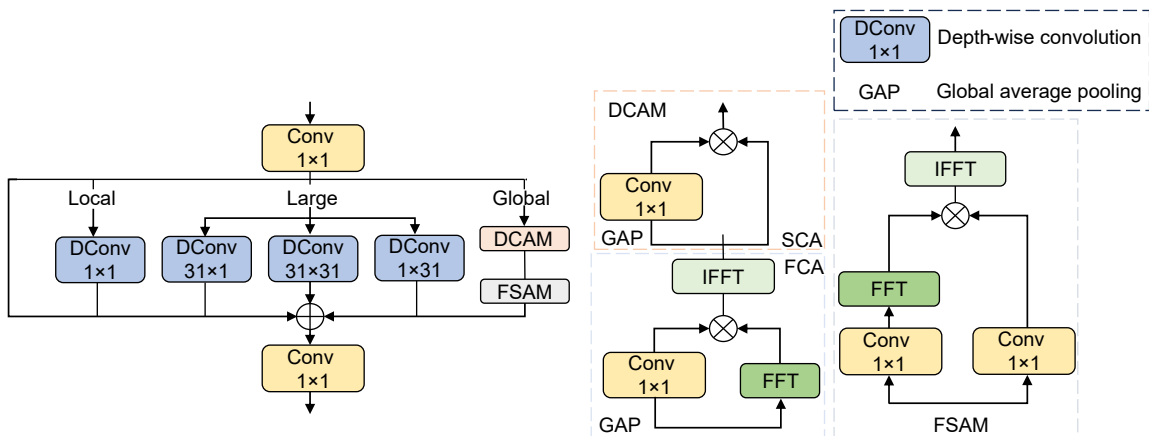


图 3 Omni-Kernel 结构图

Fig. 3 Omni-Kernel structure diagram

度上应用频域注意力模块, 以细粒度水平进一步优化频谱信息, 其过程表示为

$$X_{\text{FSAM}} = \text{IFFT}(\text{FFT}(W_{1 \times 1}^1(X_{\text{DCAM}})) \otimes W_{1 \times 1}^2(X_{\text{DCAM}})), \quad (4)$$

式中: X_{FSAM} 表示 FSAM 的输出。通过这种处理方式, 模型能够关注信息的频率分量, 从而增强对图像细节的捕捉能力, 实现高质量的图像重建。

局部分支采用了 1×1 深度卷积层来处理局部信息。在钢材表面缺陷检测任务中, 不同类型的缺陷通常具有多尺度特性。与大分支和全局分支主要捕获全局结构和大尺度感受野不同, 局部分支在识别细微缺陷特征、增强局部信息处理方面发挥了关键作用。

CSPNet 把特征图划分为两个部分, 其中一部分执行标准的卷积处理, 而另一部分则是直接跳过处理, 最后再将两部分的特征进行融合。这样的结构设计能够显著减少冗余计算。Omni-Kernel 模块具有全局和局部特征提取能力, 在处理复杂场景时表现更加稳健。CSP-OK 模块结合两者的优势, 通过高效的特征分离和多尺度特征提取, 能够精准捕捉钢材表面的微小和复杂缺陷, 同时减少冗余信息的计算, 其结构图如图 4 所示。针对钢材表面缺陷的细粒度特征, CSP-OK 模块有效融合了全局和局部信息, 提升了模型对微小缺陷和大范围缺陷的检测准确性与鲁棒性。

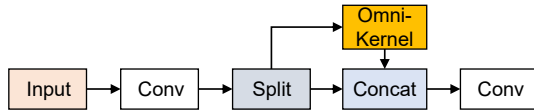


图 4 CSP-OK 结构图

Fig. 4 CSP-OK structure diagram

2.3 SIoU 损失函数

YOLOv8 将 CIoU^[19] 作为默认损失计算方法, 虽

然 CIoU 在损失计算中引入了形状因素, 即帧的长宽对检测损失的影响, 但它未能顾及角度对结果的影响。SIoU 不仅关注边界框之间的重叠区域, 还引入了距离、角度偏差和形状相似性等因素, 从而更加全面地评估预测框与真实框的匹配度。钢材缺陷往往具有不规则的形状和不同的方向, 传统的 IoU 损失函数在处理这些旋转、细长或异形缺陷时效果有限。而 SIoU 通过综合考虑边界框的距离、角度偏差和形状相似性, 克服了这些局限。通过引入 SIoU 损失函数, 模型能够在提升检测精度的同时, 减少因旋转和形状不规则带来的误差, 从而在钢材缺陷检测中表现出更强的鲁棒性和准确性。因此, 本文提出使用 SIoU 来替代 CIoU, 以加快两个预测框之间距离的收敛速度, 从而减少推理时间与计算资源消耗。图 5 展示了角度成本在损失函数中所起的作用。SIoU 不仅考虑几何参数的成本, 还考虑预测框与实际框之间的角度成本, 计算角度的公式为

$$\begin{cases} \text{Angle}_{\text{Cost}} = \cos\left(2 \arcsin(\sin \alpha) - \frac{\pi}{2}\right) \\ \sin \alpha = \begin{cases} \sin \alpha_2, & \sin \alpha_2 > \frac{\pi}{4} \\ \sin \alpha_1, & \text{otherwise} \end{cases} \\ \sin \alpha_1 = \frac{c_h}{s_{ch}} \\ \sin \alpha_2 = \frac{c_w}{s_{cw}} \end{cases}, \quad (5)$$

$$\begin{aligned} c_w &= \max(x_2^{b_1}, x_2^{b_2}) - \min(x_1^{b_1}, x_1^{b_2}) \\ c_h &= \max(y_2^{b_1}, y_2^{b_2}) - \min(y_1^{b_1}, y_1^{b_2}) \\ s_{cw} &= \frac{(x_1^{b_2} + x_2^{b_2}) - (x_1^{b_1} + x_2^{b_1})}{2} \\ s_{ch} &= \frac{(y_1^{b_2} + y_2^{b_2}) - (y_1^{b_1} + y_2^{b_1})}{2} \end{aligned}$$

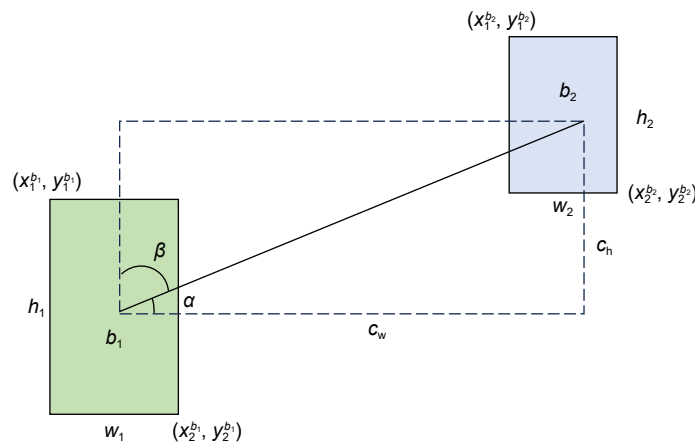


图 5 角度成本在损失函数中的作用

Fig. 5 Role of angular cost in loss function

式中: c_w 和 c_h 分别为真实框和预测框中心点的宽度差和高度差, s_{cw} 和 s_{ch} 分别为真实框和预测框中心的水平和垂直位移; $(x_1^{b_1}, y_1^{b_1})$ 和 $(x_2^{b_1}, y_2^{b_1})$ 为预测框的左上角和右下角的坐标, $(x_1^{b_2}, y_1^{b_2})$ 和 $(x_2^{b_2}, y_2^{b_2})$ 为真实框的左上角和右下角的坐标; b_1 为预测框; b_2 为真实框。

SIoU 的距离成本的定义公式为

$$\begin{cases} Distance_{Cost} = 2 - e^{-\gamma p_x} - e^{-\gamma p_y} \\ \rho_x = \left(\frac{s_{cw}}{c_w} \right)^2 \\ \rho_y = \left(\frac{s_{ch}}{c_h} \right)^2 \\ \gamma = 2 - Angle_{Cost} \end{cases}, \quad (6)$$

式中: ρ_x 和 ρ_y 分别为中心点水平和垂直偏移的平方误差。

形状成本的公式为

$$\begin{cases} Shape_{Cost} = (1 - e^{-\omega_w})^\theta + (1 - e^{-\omega_h})^\theta \\ \omega_w = \frac{|w_1 - w_2|}{\max(w_1, w_2)} \\ \omega_h = \frac{|h_1 - h_2|}{\max(h_1, h_2)} \end{cases}, \quad (7)$$

式中: (w_1, h_1) 和 (w_2, h_2) 分别表示预测框和真实框的宽和高; θ 表示对形状成本的关注。根据文献 [20] 计算出 θ 在 2 和 6 之间, 在本文中, 选择中位数 $\theta = 4$ 。

IoU 损失的公式为

$$IoU = \frac{|b_1 \cap b_2|}{|b_1 \cup b_2|}. \quad (8)$$

SIoU 的最终计算公式为

$$L_{SIoU} = 1 - IoU + \frac{Distance_{Cost} + Shape_{Cost}}{2}. \quad (9)$$

3 实验结果与分析

3.1 实验准备

3.1.1 数据集

本文采用了东北大学开源的 NEU-DET^[21-23] 钢材表面缺陷数据集。该数据集涵盖 6 种典型缺陷类型: 裂纹 (crazing, Cr)、夹杂物 (inclusion, In)、斑点 (patches, Pa)、麻面 (pitted-Surface, Ps)、轧入氧化皮 (rolled-in-Scale, Rs) 以及划痕 (scratches, Sc), 如图 6 所示。数据集中共有 1800 张图片, 分辨率均为 200×200 , 每类包含 300 张。数据集按照 8:1:1 的比例随机划分为训练集 (1440 张)、验证集 (180 张) 和测试集 (180 张)。

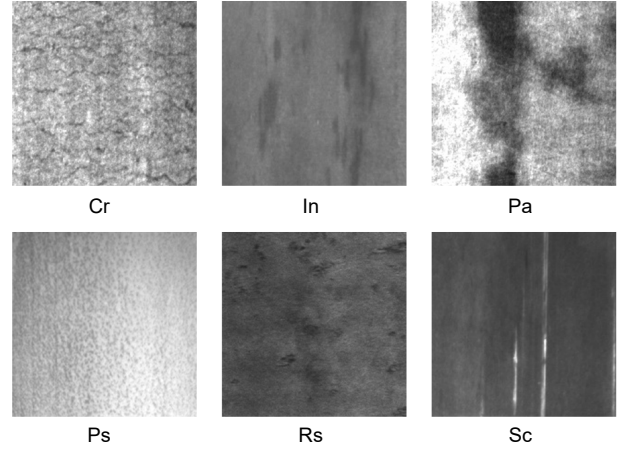


图 6 钢材表面缺陷示例图

Fig. 6 Sample diagram of steel surface defects

3.1.2 实验环境与参数设置

本实验在 NVIDIA GeForce RTX 4060 Laptop GPU 和 Intel(R) Core(TM) i7-13620H CPU 上实现。实验采用的操作系统为 Windows 11, 深度学习框架为 PyTorch 2.0.0, 编程语言为 Python 3.8。本实验的参数设置基本采用 YOLOv8 官方推荐, 输入图像尺寸为 640×640 , 初始学习率为 0.01, 使用随机梯度优化器, IoU 设置为 0.5, batch 为 16, epoch 为 200。

3.1.3 评估指标

使用以下评价指标对模型进行评估, 分别是均值平均精度 (mean average precision, mAP)、精确率 (precision, P)、召回率 (recall, R)、FPS (frames per second)、计算量 (floating point operations, FLOPs)。mAP 通过计算所有类别的 AP 值的平均值来评估检测精度; FPS 表示模型每秒可处理的图像数量, 用于衡量检测速度; FLOPs 通常用于衡量模型的复杂性和计算开销, 可通过分析模型的层次结构计算出来。相关公式为

$$AP = \int_0^1 P(R) dR, \quad (10)$$

$$mAP = \frac{1}{N} \sum_{i=1}^N AP_i, \quad (11)$$

$$P = \frac{TP}{TP + FP}, \quad (12)$$

$$R = \frac{TP}{TP + FN}, \quad (13)$$

$$FPS = \frac{n}{T}, \quad (14)$$

式中: N 表示类别总数; AP_i 表示第 i 个类别的平均

精度; TP 表示真正例; FP 表示假正例; FN 表示假负例; n 表示总处理的帧数; T 表示总耗时。

3.2 改进模块实验

3.2.1 FSCConv 模块实验

为了验证 FSCConv 模块在不同位置替换 YOLOv8n 结构时网络性能的变化, 在 NEU-DET 数据集上进行了实验对比, 实验结果如表 1 所示。实验分别在两种不同的设置下进行, 其中 A 是在整个模型中将对应的卷积层替换为 FSCConv, 而 B 则仅在检测头部第二次上采样后的下一层引入 FSCConv。从实验结果可以看出, 方案 A 在准确率上表现更佳, 达到了 0.827, 但在计算量和推理速度方面表现较差, 计算量增加至 11.6 G, 推理速度降低至 112.2 f/s。相比之下, 方案 B 在召回率和 mAP 方面表现最优, 召回率达到 0.768, mAP 提升至 79.4%。同时, 方案 B 的推理速度相对较快, 计算复杂度也较为适中, 这表明方案 B 在检测精度和推理速度之间取得了更好的平衡。基于上述分析, 最终选择了方案 B 作为 FSCConv 模块的引入方式, 有效增强了小目标的检测能力, 同时避免了全局替换带来的计算负担和信息损失, 在保证较高检测精度的同时, 兼顾计算效率和推理速度。

3.2.2 CSP-OK 模块实验

为了评估 CSP-OK 模块在不同位置替换 YOLOv8n 结构对网络性能的影响, 在 NEU-DET 数据集上进行了对比实验, 实验结果如表 2 所示。具体而言, 实验设置包括两种方案: 方案 C 将检测头中第二个 C2f 层替换为 CSP-OK 模块; 方案 D 在检测头的第二个 Concat 操作后新增一层 CSP-OK 模块。对比不同结构配置下的检测性能, 分析 CSP-OK 模块在优化特征融

合与提升检测精度方面的影响。与基线模型相比, 方案 C 和方案 D 均提升了检测精度。但是方案 C 的 FLOPs 增幅超过 140%, 推理速度从 218.2 f/s 下降至 118.8 f/s, 表明该方案虽然提升了检测性能, 但计算开销显著增加。方案 D 的 FLOPs 增加至 9.8 G, 相较于方案 C 计算开销较小, 推理速度为 116.6 f/s。在 C2f 位置引入 CSP-OK 模块, 每个尺度的卷积核都必须处理较原始的特征, 这会导致不必要的计算开销, 增加 GFLOPs。相较之下, 在 Concat 之后引入 CSP-OK 模块, 则该模块能够作用于经过深度卷积和特征提取后的高层特征图。此时, 特征已经过不同层次的提取, 并通过 Concat 操作融合了多个尺度的信息, 包含更丰富的上下文信息和小目标细节。在这一阶段使用 CSP-OK 模块, 不仅能够进一步优化多尺度特征的传递, 还能减少冗余计算, 增强特征表达能力。同时, 由于 CSP 结构能够分块处理特征流, 减少重复计算, 结合 Omni-Kernel 可进一步增强不同感受野的信息交互, 从而在保证检测精度的同时控制计算开销, 使 GFLOPs 仅小幅增加。本文选择了方案 D 作为 CSP-OK 模块的引入方式, 方案 D 在提升精度的同时保持了较低的计算复杂度, 相较于方案 C 具有更好的计算效率和推理速度。

3.3 消融实验

为了评估 YOLOv8-SOE 相对于 YOLOv8n 算法的优化效果, 本文设计了逐步的消融实验。实验共分为 8 组, 均在 NEU-DET 训练集上训练, 验证集上验证。为确保实验准确性, 每一组实验训练过程均采用相同参数。为了直观展示 YOLOv8-SOE 在小目标上的效果, 图 7 展示了 YOLOv8-SOE 和 YOLOv8n 在

表 1 FSCConv 在不同位置的实验结果
Table 1 Experiment results on FSCConv at different positions

Scheme	Precision	Recall	FLOPs/G	FPS	mAP/%
Baseline model	0.759	0.722	8.1	218.2	75.3
A	0.827	0.677	11.6	112.2	77.0
B	0.754	0.768	9.1	187.9	79.4

表 2 CSP-OK 模块在不同位置的实验结果
Table 2 Experiment results on CSP-OK module at different positions

Scheme	Precision	Recall	FLOPs/G	FPS	mAP/%
Baseline model	0.759	0.722	8.1	218.2	75.3
C	0.766	0.779	19.9	118.8	79.1
D	0.762	0.743	9.8	116.6	79.1

相同数据上的特征热力图。消融实验结果如表 3 所示。在引入 FSCConv 模块后, 相较于基准模型, 其 mAP 提升了 3.9%, 漏检率 (r_m) 减少了 1.15%, 误检率 (r_f) 减少了 0.38%。进一步引入 SIoU 损失函数后, mAP 提升至 79.4%, 误检率略下降至 20.48%, 漏检率进一步降低至 24.89%。引入 CSP-OK 模块后, 虽然计算量增加, 推理速度有所下降, 误检率有所增加, 但是 mAP 较基准模型提升了 3.8%, 漏检率减少了 1.61%, 进一步引入 SIoU 损失函数后, mAP 提升至 79.3%, 推理速度变快, 漏检率降低至 21.58%。引入 CSP-OK 模块后, 模型能够提取更加丰富的多尺度特征, 然而, 过于复杂的特征融合可能导致模型在不同缺陷类别之间产生混淆, 从而增加误检率。在仅引入 SIoU 损失函数时, mAP 相较于基准模型提升了 2.1%, 计算量保持不变, 但是误检率有所增加。这表明 SIoU 在优化目标框回归性能和提高检测精度方面

表现出色, 同时对计算负担影响较小。但是 SIoU 损失函数过于强调目标的形状、角度和边界的精确度, 会导致边界框对细节变化更加敏感。当模型对这些细节过于敏感时, 可能会将一些非缺陷区域误判为缺陷。当 FSCConv 和 CSP-OK 模块组合使用时, mAP 提升了 5.1%。接着在引入 FSCConv 和 CSP-OK 模块的基础上, 引入 SIoU 损失函数, mAP 又提升了 0.3%。

在同时引入 FSCConv 和 CSP-OK 模块时, FSCConv 通过压缩特征图, 将空间信息有效地压缩到通道维度, 从而减少了冗余信息的计算并提升了特征表达能力。CSP-OK 模块通过更高效特征融合和提取方式, 在保证检测精度的同时优化了特征的计算过程, 提高了特征提取的效率。这两者的结合增强了模型的检测能力, FLOPs 维持在 11.8 G, 推理速度达到 172.0 f/s, 确保了计算量和推理速度的平衡, 同时有效提高了对小目标和复杂缺陷的识别精度。

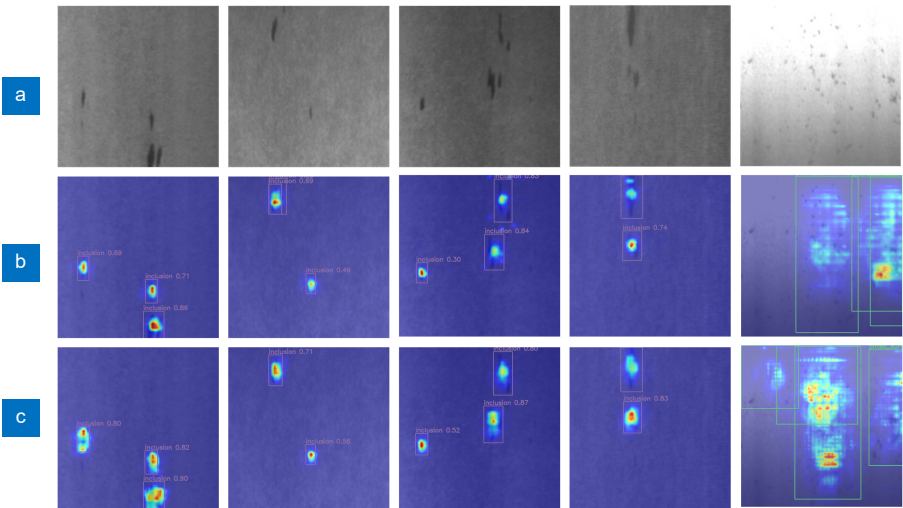


图 7 特征热力图。(a) 原始图像; (b) YOLOv8n 特征热力图; (c) YOLOv8-SOE 特征热力图
Fig. 7 Feature heatmap. (a) Original image; (b) YOLOv8n feature heatmap; (c) YOLOv8-SOE feature heatmap

表 3 消融实验结果

Table 3 Results of ablation experiments									
FSCConv	CSP-OK	SIoU	Precision	Recall	FLOPs/G	FPS	mAP/%	r_m /%	r_f /%
—	—	—	0.759	0.722	8.1	218.2	75.3	27.33	21.78
√	—	—	0.760	0.762	9.1	185.3	79.2	26.18	21.40
√	—	√	0.754	0.768	9.1	187.9	79.4	24.89	20.48
—	√	—	0.762	0.743	9.8	116.6	79.1	25.72	23.79
—	√	√	0.731	0.786	9.8	137.2	79.3	21.58	26.89
—	—	√	0.738	0.761	8.1	215.7	77.4	23.88	26.21
√	√	—	0.783	0.746	11.8	164.0	80.4	24.93	19.41
√	√	√	0.739	0.767	11.8	172.0	80.7	21.48	23.79

3.4 对比实验

3.4.1 主流模型对比实验

为了验证所提算法的有效性, 在相同的 NEU-DET 数据集上, 对比了几种主流检测方法在 AP、mAP、GFLOPs 和 FPS 方面的表现, 对比算法包括 Faster-RCNN、RT-DETR、YOLOv5s、YOLOv7、YOLOX、YOLOv8n、YOLOv10n, 此外再选择钢材表面缺陷检测算法 GBS-YOLOv7t^[24]、PIC2f-YOLO^[25]、RFB-YOLO^[26] 和参考文献 [27] 进行性能对比, 以客观展示所提模型的优越性。算法性能对比结果如表 4 所示, 为了更清晰地展示各个模型的性能差异, 在表格中对每个性能指标的最高值进行了加粗处理。

根据表 4 的结果分析可知, 所提改进模型在检测性能和计算效率方面均表现出显著优势。在准确率方面, 改进模型在 Cr 和 Sc 类别上取得了最高的平均精度, 其中 Cr 类别的平均精度提升幅度最大, 相较于原始 YOLOv8n 模型提高了 15.9%, 所提模型的 mAP 达到 80.7%, 是所有类别中的最高值。此外, 相较于 Faster-RCNN 提升了 4.2%, 相较于 RT-DETR 提升 8.2%, 比 YOLOv5s 高 4.9%, 与 YOLOv7 相比提升 3.0%, 比 YOLOX 高 14.1%, 比原始 YOLOv8n 提升

5.4%, 比 YOLOv10n 高 9.4%。在计算复杂度方面, 改进模型表现同样优异, 其 FLOPs 仅为 11.8 G, 略高于 YOLOv8n、YOLOv10n 和 PIC2f-YOLO, 但仍保持在较低水平。同时, 所提模型的检测速度达到 172 f/s, 能够在高效推理的基础上维持较高的检测速度。与同样应用于钢材表面缺陷检测的模型 GBS-YOLOv7t、PIC2f-YOLO、RFB-YOLO 和参考文献 [27] 对比, 所提改进模型的 mAP 分别高出 7.8%、2.7%、1.5% 和 6.6%。尽管 RFB-YOLOv5-E 在 Pa 和 Ps 类别上取得最高平均精度, 但所提改进模型在整体检测速度和计算复杂度上均优于该模型, 充分展现其在检测精度与计算效率之间的卓越平衡。

3.4.2 鲁棒性对比实验

为了验证所提改进算法在复杂环境下的鲁棒性, 对原始数据集进行了随机亮度调整、模糊处理和加噪实验, 以模拟工业场景中可能出现的不同光照环境、运动模糊以及噪声干扰。亮度调整包括增加和降低亮度 ($\pm 30\%$); 模糊处理采用高斯模糊以模拟成像过程中可能的模糊现象; 加噪实验则采用高斯噪声, 模拟工业摄像头在恶劣环境下拍摄时的信号干扰。亮度调整、模糊处理和加噪后的部分数据如图 8 和图 9

表 4 对比实验结果
Table 4 Results of comparative experiments

Model	AP/%						FLOPs/G	FPS	mAP/%
	Cr	In	Pa	Ps	Rs	Sc			
Faster-RCNN	37.2	84.1	89.7	82.3	72.7	93.2	134.0	34.200	76.5
RT-DETR	44.5	77.9	89.0	67.5	56.1	92.2	57.0	48.100	72.5
YOLOv5s	53.4	79.8	93.2	81.1	51.5	96.0	15.8	74.505	75.8
YOLOv7	57.2	84.6	91.1	84.9	54.8	93.3	103.2	65.790	77.7
YOLOX	30.2	77.5	85.2	75.2	39.0	88.8	13.32	43.200	66.6
YOLOv8n	44.1	75.7	91.9	82.7	61.0	96.3	8.1	218.200	75.3
YOLOv10n	42.2	78.2	88.9	76.9	55.3	86.5	6.5	192.500	71.3
GBS-YOLOv7t ^[24]	32.7	69.6	92.4	96.5	57.7	88.6	—	104.1	72.9
PIC2f-YOLO ^[25]	—	—	—	—	—	—	10.6	80	78.0
RFB-YOLOv5-E ^[26]	52.1	75.5	95.5	97.3	62.3	92.4	22.4	122	79.2
Ref. [27]	39.5	86.0	92.1	78.9	62.6	85.3	12.2	—	74.1
Ours	60.0	85.6	93.1	84.8	62.2	98.2	11.8	172.0	80.7

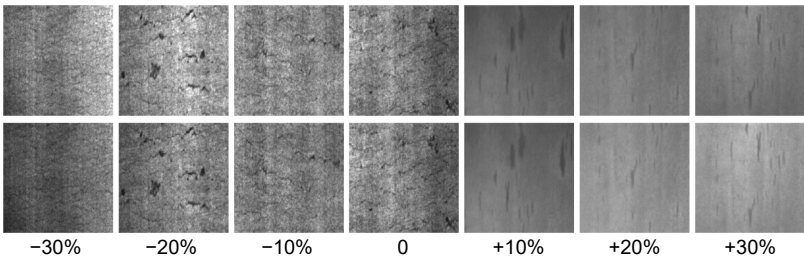


图 8 亮度调整后的示例图
Fig. 8 Sample images after brightness adjustment

所示。

表 5 和表 6 展示了在 NEU-DET 数据集和随机亮度调整、模糊处理和加噪后的数据集上, YOLOv8n 与所提改进模型的性能对比。根据表 5、表 6 的结果分析可知, 在随机亮度调整、模糊处理和加噪后的数据集上的检测性能相较于原始 NEU-DET 数据集有所下降, 这说明光照变化、模糊和噪声对缺陷检测模型有一定的影响, 使得检测精度和召回率下降。但是改进模型在亮度调整后的数据集上的准确率为 0.757, 召回率为 0.745, mAP 为 79.6%, 整体优于 YOLOv8n 在此数据集上的表现。在模糊处理和加噪后的数据集上, 改进模型在综合处理后的数据集上表现优异, Precision 提升至 0.789, Recall 为 0.726, mAP 达到 78.0%, 优于 YOLOv8n 在原始数据集上的表现。从实验结果来看, 无论是在光照变化还是模糊与噪声干扰的环境下, 所提改进模型在各项指标上均优于 YOLOv8n, 展现出更强的鲁棒性。这表明所提改进模型在复杂工业环境下具有更稳定的检测能力和更强的抗干扰能力, 能够更可靠地应对不同光照条件、噪声干扰和模糊影响。

3.4.3 泛化性对比实验

为了验证所提改进算法的泛化性, 在 VOC2012

数据集^[28]上进行了实验, 并与原始 YOLOv8n 算法的性能进行了对比。VOC2012 数据集包含 17125 张图像, 覆盖了 20 类目标。在数据集划分方面, 按照 8:1:1 的比例, 将数据集分为训练集、验证集和测试集。整个实验过程中超参数设置一致, 确保不同算法之间的结果具有可比性。实验结果如表 7 所示。所提改进算法仅 Recall 略低于 YOLOv5s, 其余结果均优于 YOLOv5s、YOLOX、YOLOv10n、YOLOv8n 算法的检测结果, 特别是在精确率和平均精度上有显著提升。所提改进算法在整体性能上相较于原始模型取得了显著提升。具体而言, 平均精度由 YOLOv8n 的 63.0% 提高至 69.4%, 提升了 6.4%。同时, 精确率由 0.689 提升至 0.754, 增幅达 6.5%。上述实验结果充分验证了改进模型在不同类型目标检测任务中具有较优的鲁棒性和较强的泛化能力。

4 结 论

所提 YOLOv8-SOE 结构在钢材表面细微缺陷检测中展现了出色的性能, 特别是在小目标的检测方面。通过引入 FSCConv 和 CSP-OK 模块, 实现了对不同尺度特征的有效融合, 确保了小目标的细节信息得以保留。同时, SIOU 损失函数的引入进一步增强了模

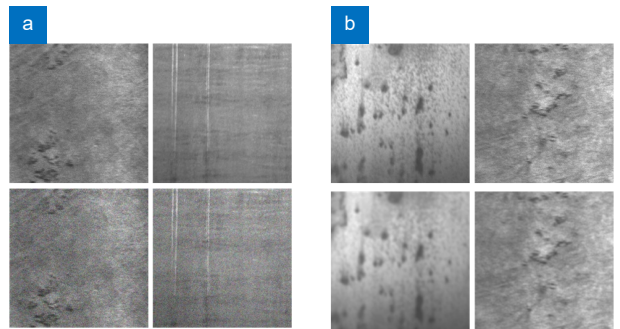


图 9 (a) 高斯噪声处理后的图像示例图; (b) 高斯模糊处理后的图像示例图

Fig. 9 (a) Sample images processed by Gaussian noise; (b) Sample images after Gaussian blur processing

表 6 在模糊与噪声环境下的对比实验结果

Table 6 Comparative experimental results in fuzzy and noisy environments

Dataset	Model	Precision	Recall	mAP/%
NEU-DET	YOLOv8n	0.759	0.722	75.3
	Ours	0.739	0.767	80.7
NEU-DET-augmented	YOLOv8n	0.644	0.688	72.7
	Ours	0.789	0.726	78.0

表 5 在不同光照环境下的对比实验结果

Table 5 Comparative experimental results under different lighting environments

Dataset	Model	Precision	Recall	mAP/%
NEU-DET	YOLOv8n	0.759	0.722	75.3
	Ours	0.739	0.767	80.7
NEU-DET-brightness	YOLOv8n	0.730	0.672	74.9
	Ours	0.757	0.745	79.6

表 7 在 VOC2012 数据集上的对比实验结果

Table 7 Comparative experiment results on VOC2012 dataset

Model	Precision	Recall	mAP/%
YOLOv5s	0.724	0.646	68.2
YOLOX	0.715	0.608	61.6
YOLOv10n	0.702	0.536	60.3
YOLOv8n	0.689	0.569	63.0
Ours	0.754	0.621	69.4

型对不同形状目标的检测能力。实验结果表明, 该模型在 NEU-DET 数据集上的 mAP 达到了 80.7%, 相较于 YOLOv8n 提升了 5.4%。此外, 为了验证所提模型的泛化能力, 进一步在 VOC2012 数据集上进行了测试。结果表明, 该模型不仅在钢材缺陷检测任务中表现出色, 在更广泛的检测场景中同样展现出良好的泛化能力。尽管该模型有效平衡了检测精度与计算成本, 但在处理高分辨率图像时, 推理速度和计算资源的消耗仍存在优化空间。未来工作可以进一步探讨如何在保持高检测精度的前提下, 提升模型的计算效率与推理速度。

利益冲突: 所有作者声明无利益冲突

参考文献

- [1] Luo Q W, Fang X X, Liu L, et al. Automated visual defect detection for flat steel surface: a survey[J]. *IEEE Trans Instrum Meas*, 2020, **69**(3): 626–644.
- [2] 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. <https://doi.org/10.1109/CVPR.2014.81>.
- [3] Girshick R. Fast R-CNN[C]//*Proceedings of the IEEE International Conference on Computer Vision*, 2015: 1440–1448. <https://doi.org/10.1109/ICCV.2015.169>.
- [4] Ren S Q, He K M, Girshick R, et al. Faster R-CNN: towards real-time object detection with region proposal networks[J]. *IEEE Trans Patt Anal Mach Intell*, 2017, **39**(6): 1137–1149.
- [5] Shi X C, Zhou S K, Tai Y C, et al. An improved faster R-CNN for steel surface defect detection[C]//*2022 IEEE 24th International Workshop on Multimedia Signal Processing (MMSP)*, 2022: 1–5. <https://doi.org/10.1109/MMSP55362.2022.9949350>.
- [6] 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. https://doi.org/10.1007/978-3-319-46448-0_2.
- [7] Zhao Y A, Lv W Y, Xu S L, et al. DETRs beat YOLOs on real-time object detection[C]//*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2024: 16965–16974. <https://doi.org/10.1109/CVPR52733.2024.01605>.
- [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*, 2016: 779–788. <https://doi.org/10.1109/CVPR.2016.91>.
- [9] Redmon J, Farhadi A. YOLO9000: better, faster, stronger[C]//*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017: 7263–7271. <https://doi.org/10.1109/CVPR.2017.690>.
- [10] Lin T Y, Goyal P, Girshick R, et al. Focal loss for dense object detection[C]//*Proceedings of the IEEE International Conference on Computer Vision*, 2017: 2980–2988. <https://doi.org/10.1109/ICCV.2017.324>.
- [11] Li J J, Chen M X. DEW-YOLO: an efficient algorithm for steel surface defect detection[J]. *Appl Sci*, 2024, **14**(12): 5171.
- [12] Yang S X, Xie Y, Wu J Q, et al. CFE-YOLOv8s: improved YOLOv8s for steel surface defect detection[J]. *Electronics*, 2024, **13**(14): 2771.
- [13] Peng J H, Zhang C, Gao Q, et al. Steel defect detection based on improved YOLOv8 algorithm[J/OL]. *Comput Eng*, 1–9. <https://doi.org/10.19678/j.issn.1000-3428.00EC0069283>.
彭菊红, 张弛, 高谦, 等. 基于改进的 YOLOv8 算法的钢材缺陷检测[J/OL]. *计算机工程*, 1–9. <https://doi.org/10.19678/j.issn.1000-3428.00EC0069283>.
- [14] Zhang X R, Wang Y L, Fang H S. Steel surface defect detection algorithm based on ESI-YOLOv8[J]. *Mater Res Express*, 2024, **11**(5): 056509.
- [15] Huang Y, Tan W Z, Li L, et al. WFRE-YOLOv8s: a new type of defect detector for steel surfaces[J]. *Coatings*, 2023, **13**(12): 2011.
- [16] Cui Y N, Ren W Q, Knoll A. Omni-kernel network for image restoration[C]//*Proceedings of the 38th AAAI Conference on Artificial Intelligence*, 2024: 1426–1434. <https://doi.org/10.1609/aaai.v38i2.27907>.
- [17] Gevorgyan Z. SloU loss: more powerful learning for bounding box regression[Z]. arXiv: 2205.12740, 2022. <https://arxiv.org/abs/2205.12740>.
- [18] Wang C Y, Liao H Y M, Wu Y H, et al. CSPNet: a new backbone that can enhance learning capability of CNN[C]//*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, 2020: 390–391. <https://doi.org/10.1109/CVPRW50498.2020.00203>.
- [19] Zheng Z H, Wang P, Ren D W, et al. Enhancing geometric factors in model learning and inference for object detection and instance segmentation[J]. *IEEE Trans Cybern*, 2022, **52**(8): 8574–8586.
- [20] Tsai F J, Peng Y T, Lin Y Y, et al. Stripformer: strip transformer for fast image deblurring[C]//*Proceedings of the 17th European Conference on Computer Vision*, 2022: 146–162. https://doi.org/10.1007/978-3-031-19800-7_9.
- [21] Bao Y Q, Song K C, Liu J, et al. Triplet-graph reasoning network for few-shot metal generic surface defect segmentation[J]. *IEEE Trans Instrum Meas*, 2021, **70**: 5011111.
- [22] Song K C, Yan Y H. A noise robust method based on completed local binary patterns for hot-rolled steel strip surface defects[J]. *Appl Surf Sci*, 2013, **285**: 858–864.
- [23] He Y, Song K C, Meng Q G, et al. An end-to-end steel surface defect detection approach via fusing multiple hierarchical features[J]. *IEEE Trans Instrum Meas*, 2020, **69**(4): 1493–1504.
- [24] Liang L M, Long P W, Lu B H, et al. Improvement of GBS-YOLOv7t for steel surface defect detection[J]. *Opto-Electron Eng*, 2024, **51**(5): 240044.
梁礼明, 龙鹏威, 卢宝贺, 等. 改进 GBS-YOLOv7t 的钢材表面缺陷检测[J]. *光电工程*, 2024, **51**(5): 240044.
- [25] Hu Y L, Yang J, Xu C Y, et al. PIC2f-YOLO: a lightweight method for the detection of metal surface defects[J]. *Opto-Electron Eng*, 2025, **52**(1): 240250.
胡依伦, 杨俊, 许聪源, 等. PIC2f-YOLO: 金属表面缺陷检测轻量化方法[J]. *光电工程*, 2025, **52**(1): 240250.
- [26] Huang S Q, Huang J G. Improved steel defect detection method based on enhanced fusion of RFB and YOLOv5 features[J]. *Comput Eng*, 2025, **51**(4): 249–260.
黄硕清, 黄金贵. 基于 RFB 和 YOLOv5 特征增强融合改进的钢材缺陷检测方法[J]. *计算机工程*, 2025, **51**(4): 249–260.

- [27] Yang L S, Li M J, Hu J W, et al. Strip steel surface defect detection algorithm based on improved YOLOv7-tiny[J]. *Comput Eng*, 2025, 51(1): 208–215.
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]. *计算机工程*, 2025, 51(1): 208–215.
[28] Everingham M, Eslami S M A, Van Gool L, et al. The pascal visual object classes challenge: A retrospective[J]. *Int J Comput Vis*, 2015, 111(1): 98–136.

作者简介



马玉洁 (2000-), 女, 硕士研究生, 主要研究方向为计算机视觉、缺陷检测等。

E-mail: 2230922101@stu.ahpu.edu.cn



【通信作者】曹维清 (1982-), 男, 博士, 高级工程师, 主要研究方向为智能机器人视觉感知、人机交互技术、机器人群智能技术等。参与或主持国家重点研发计划、国家自然科学基金、安徽省重点研发计划等项目等 15 项。共发表 SCI/EI 论文 20 余篇, 获发明专利授权 35 项。

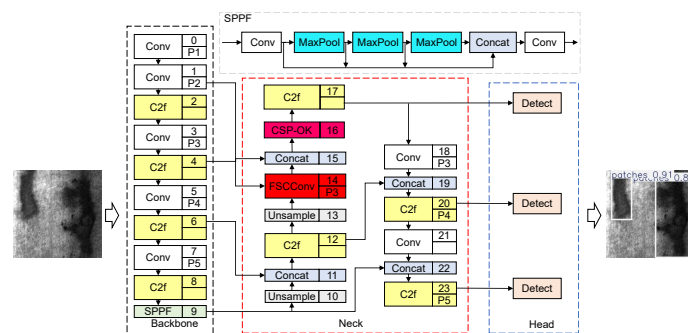
E-mail: caochuqing@ahpu.edu.cn



扫描二维码, 获取PDF全文

Steel surface defect detection based on YOLOv8-SOE

Ma Yujie^{1,2}, Cao Chuqing^{1,2*}, Zhang Jing²



YOLOv8-SOE model structure diagram

Overview: In industrial applications such as steel surface defect detection, small target detection remains a challenging task due to the limited resolution of conventional detection layers, making it difficult to capture fine-grained defect details. Although YOLOv8 has shown remarkable performance in multi-scale target detection and complex environments, the model struggles with small target detection, especially when it comes to tiny defects on steel surfaces. Traditional solutions to enhance small target perception, such as adding additional detection heads like the P2 layer, often lead to increased computational overhead and inference time. To address this issue, this paper proposes an improved YOLOv8-SOE model that specifically enhances the detection performance for small defects on steel surfaces. The YOLOv8-SOE model incorporates several innovations aimed at improving both detection accuracy and computational efficiency. First, a novel feature processing module, FSCConv, is introduced to handle the P2 layer features. FSCConv leverages dilated convolutions to capture multi-scale contextual information while preserving fine details of small targets. This approach enhances small target perception without the need for additional detection layers, thus avoiding the computational burden typically associated with such modifications. Next, the processed P2 features are fused with P3 layer features, improving small target detection further without incurring significant computational costs. To optimize the feature fusion process, a cross-stage local network combined with Omni-Kernel (CSP-OK) is proposed. CSP-OK primarily leverages the CSPNet approach to reduce redundant gradient computations and integrates the Omni-Kernel to prevent the repetitive extraction of similar information at different layers, thereby improving information utilization efficiency. This optimization reduces redundant information computations and effectively utilizes inter-layer features, resulting in a more efficient and detailed integration of multi-scale information. In addition, the model uses the SIOU loss function for bounding box regression. This loss function not only considers the overlap between the predicted and ground truth boxes but also incorporates their distance, angular deviation, and shape similarity. By integrating these factors, the SIOU loss function provides a more comprehensive optimization strategy, thereby improving the accuracy of target localization. Experimental results demonstrate that the YOLOv8-SOE model achieves a mean average precision (mAP) of 80.7% on the NEU-DET dataset, a 5.4% improvement over the baseline YOLOv8 model. The model also exhibits excellent generalization ability on the VOC2012 dataset. Overall, the proposed YOLOv8-SOE model significantly enhances small target detection precision while maintaining high computational efficiency, making it a promising solution for real-world industrial defect detection applications.

Ma Y J, Cao C Q, Zhang J. Steel surface defect detection based on YOLOv8-SOE[J]. *Opto-Electron Eng*, 2025, 52(5): 250032; DOI: 10.12086/oe.2025.250032

Foundation item: Special Project for the Central Government to Guide Local Science and Technology Development-High-Level New R&D Institutions Fund (S202407a12020270), Anhui Province Key Research and Development Plan-High-tech Field Fund(202304a05020077)

¹School of Computer and Information, Anhui Polytechnic University, Wuhu, Anhui 241000, China; ²Yangtze River Delta HIT Robot Technology Research Institute, Wuhu, Anhui 241000, China

* E-mail: caochuqing@ahpu.edu.cn