

光电工程

Opto-Electronic Engineering

中文核心期刊 中国科技核心期刊
Scopus CSCD

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

引用本文:

张润梅, 潘晨飞, 陈梓华, 等. 改进YOLOv8的焊缝表面缺陷检测算法[J]. *光电工程*, 2025, **52**(3): 240296.

Zhang R M, Pan C F, Chen Z H, et al. Improved weld surface defect detection algorithm from YOLOv8[J]. *Opto-Electron Eng*, 2025, **52**(3): 240296.

<https://doi.org/10.12086/oe.2025.240296>

收稿日期: 2024-12-17; 修改日期: 2025-02-24; 录用日期: 2025-02-25

相关论文

快反镜系统滑模复合分层干扰观测补偿控制

罗勇, 刘凯凯, 杨帆, 闻心怡, 黄永梅, 郭珊珊, 任戈, 李涛

光电工程 2023, **50**(4): 220330 doi: [10.12086/oe.2023.220330](https://doi.org/10.12086/oe.2023.220330)

特征扰动频率辨识的自适应倾斜扰动抑制技术

吴红梅, 王琛, 冯念, 文利, 唐涛

光电工程 2023, **50**(10): 230177 doi: [10.12086/oe.2023.230177](https://doi.org/10.12086/oe.2023.230177)

惯性稳定万向架中基于SBG惯导的捷联控制技术

王玉, 边启慧, 廖军, 徐田荣, 唐涛

光电工程 2023, **50**(5): 220238 doi: [10.12086/oe.2023.220238](https://doi.org/10.12086/oe.2023.220238)

更多相关论文见光电期刊集群网站 



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



 OE_Journal



Website

address the problem of mismatched bounding boxes. Experimental results show that the mAP0.5 detection metric of the model in this paper is 7.8% higher than that of the baseline model, and the number of parameters and the amount of computation are reduced by 0.2 M and 0.7 G, respectively.

Keywords: deep learning; defect detection; YOLOv8; Inner-SIoU

1 引言

焊接是一种在工业生产领域广泛应用的技术,而且焊缝表面质量对于产品的耐久性能有着重要影响。因此,在生产过程中对焊缝表面缺陷进行检测具有重要的研究价值^[1]。

近年来,随着制造业和人工智能技术的不断发展,基于卷积神经网络的目标检测在材料表面缺陷检测领域有着长足的发展^[2],特别是对于焊缝表面缺陷的检测,成为相关领域研究的新兴热点之一^[3]。目前目标检测神经网络可分为单阶段网络(one-stage)^[4-6]和双阶段网络(two-stage)^[7-9]。两者比较而言,单阶段网络虽精度略有下降,但推理速度更快,其更符合焊缝表面缺陷检测实际工业需求。

由于工业焊接件缺陷图像中存在纹理特征相似、特征不明显、背景信息复杂的特点,模型难以从复杂的背景中准确地识别缺陷目标,大多数缺陷目标为小尺寸目标,其像素面积小且分布密集,容易出现漏检、误检。在实际工业环境里,焊接前后过程中的焊丝飞溅会造成焊渣等杂质的产生,不同缺陷的存在会对算法的定位和识别分类造成干扰。为了提升YOLO模型对于焊缝表面缺陷的目标精度,文献[10]提出了融合轻量化卷积和注意力机制网络结构(ELAN-PCS)来解决漏检问题,但未能改善其泛化性。文献[11]在下采样阶段引入无参数注意力(SimAM),实现多维信息的交互,但在检测精度和轻量化方面提升不明显。文献[12]改进了PAN模块和SPPF模块,引入了膨胀卷积分支增强不同尺度特征的融合能力,但其损失函数并未更换,导致对于小目标检测效果尚不完善。文献[13]采用由Res2Net和EMA注意力机制组成的REM模块来增强感受野,有效提升检测精度,但检测速度大幅度下降。文献[14]引入可分离注意力(large separable kernel attention, LSKA)模块挤压激励网络(squeeze-and-excitation networks, SE)注意力机制,强化多尺度特征融合能力。针对以上问题,本文选取YOLOv8n为基线网络,做出如下5个改进:首先,

将改进后的注意力机制模块融入特征融合网络模块中,增强模型实时检测时信息的收取效率;然后,在颈部网络中采用轻量化结构化设计降低算法复杂度;接着,更换上采样算子,提高特征图的分辨率和语义信息的传递;改进注意力机制模块,加快推理速度;最后,更换损失函数,进一步提高边框回归的精确度。

2 YOLOv8 模型

YOLOv8是Ultralytics公司于2023年1月发布的目标检测算法,是YOLOv5的继承版本,提供了n、s、m、l和x5个模型,模型体积逐次增大,精度也逐次提升,适用于目标检测、实例分割、姿态估计、目标跟踪和分类等任务。YOLOv8算法引入新功能和特性的模型,成为新一代目标检测任务的理想选择^[5]。相比于当前应用广泛的YOLOv5而言,YOLOv8算法在骨干网络、头部网络、正负样本分配策略、损失计算等多方面做出了诸多改进。本文选取YOLOv8n进行改进,其网络结构主要由三个部分组成,骨干网络(backbone)、颈部网络(neck)和预测网络(head),结构如图1所示。

Backbone主要实现特征提取,用于提取输入图片的特征信息形成特征图。Neck主要负责图像浅层细节特征与深层语义特征的融合,使用上采样操作(upsample)和拼接操作(concat)鉴路径聚合网络(path aggregation network, PANet)和特征金字塔网络(feature pyramid network, FPN)的思想^[6],实现不同深度特征图的信息融合,有效增强网络表达能力。Head主要负责生成目标检测的输出结果,实现对不同尺度的特征目标进行目标分类和定位预测。

3 YOLOv8 网络模型优化

3.1 GD-YOLO 网络模型

本文基于YOLOv8n提出一种焊缝表面缺陷检测算法GD-YOLO,该模型结构如图2所示,改进部分加粗边框展示。首先,将扩张残差模块(dilation-wise residual, DWR)^[17]加入到特征融合网络模块(faster

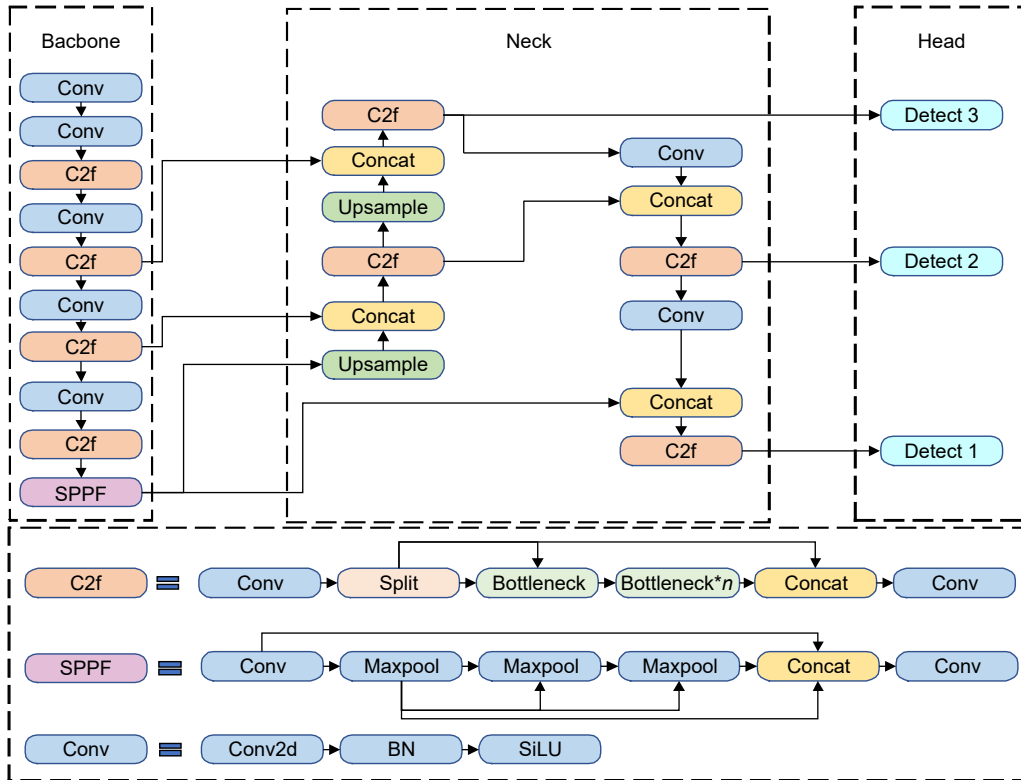


图 1 YOLOv8 网络结构图
Fig. 1 YOLOv8 network structure diagram

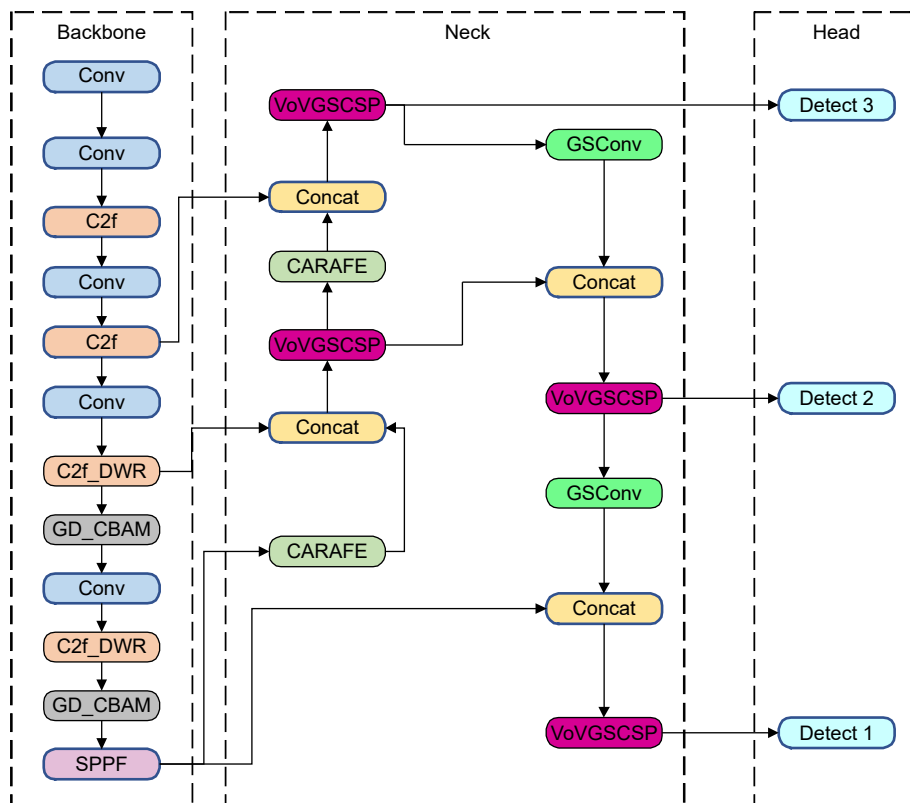


图 2 GD-YOLO 网络结构图
Fig. 2 GD-YOLO network structure diagram

implementation of CSP bottleneck with 2 convolutions, C2f) 中, 增强模型对不同尺度目标的识别能力, 加强特征图的跨层连接和信息传递。然后, 将改进后的注意力机制模块 GD_CBAM 引入 Backbone 主干网络中, 提升模型的识别精度。再者, 在颈部网络中引入轻量级上采样算子 (content-aware ReAssembly of features, CARAFE)^[18] 替换原上采样模块 (upsample), 减少由于上采样所导致的对于特征信息的损失, 增强网络的特征融合能力。接着, 在颈部网络中嵌入轻量化结构 (slim-neck), 利用鬼影混洗卷积 (grouped spatial convolution, GSConv) 和跨级部分网络模块所设计出的多视角组洗牌跨阶部分段网络 (cross-stage partial block, VoV-GSCSP)^[19], 从而提高缺陷检测模型的整体性能。最后, 采用损失函数 Inner-SIoU 计算辅助边框与真实边框之间的损失, 进一步提高边框回归的精度。

3.2 C2f-DWR 模块强化特征提取

传统方法一般会在主干网络提取特征并使用轻量级模块进行特征融合, 然而传统模块的使用存在一些限制, 比如在复杂表面下输入特征映射使得深度大速率扩展卷积难以建立长距离语义连接, 且对不同网络阶段的接受域缺乏合理的设计等。诸如此类问题, 往往会使目标形变特征会变得更加复杂, 难以从复杂特征图中捕获多尺度上下文信息。所以针对颈部网络存在的问题, 提出将扩张残差模块 DWR 与 C2f 模块进行结合以加强网络对不同尺寸缺陷的提取能力, DWR 结构如图 3 所示。

图 3 中 C 表示特征图通道的基数, Conv 表示卷积, DConv 表示深度卷积, $D-n3\times3$ DConv 表示扩张速率为 n 的卷积。DWR 模块采用两步法高效地提取多尺度的上下文信息, 分别为区域残差化 (region residualization) 和语义残差化 (semantic residualization), 并通过多尺度融合生成特征图。

首先, 进行区域残差化, 输入包含图像多尺度信息的特征图 x , 输入的特征图会通过 3×3 的深度卷积层并随之进行批量归一化 (BN) 和 ReLU 激活函数生成区域残差特征 (RR), 可由式 (1) 表示。

$$f_1(x) = \sigma(\text{BN}(\text{Conv}(x))), \quad (1)$$

式中: $f_1(x)$ 表示为区域残差化后的输出; x 表示模块的输入; $\sigma(\cdot)$ 为 ReLU 激活函数; $\text{BN}(\cdot)$ 表示批量归一化操作; $\text{Conv}(\cdot)$ 表示卷积核大小为 3×3 的深度

卷积。然后进行语义残差化, 分别采用不同通胀率的深度卷积来提取不同尺度的特征, 形成语义残差特征 (SR), 其表达式为

$$f_2(x, d) = D_d \text{DConv}(f_1(x)), \quad (2)$$

式中: $f_2(x, d)$ 代表语义残差化后的输出; $D_d \text{DConv}(\cdot)$ 表示扩张速率为 d 的 3×3 卷积。接着对语义残差特征进行残差连接, 将经过残差连接的特征图通过 BN 层获取语义残差, 并使用 1×1 逐点卷积 PConv 提取空间特征, 表达式为

$$f_3(x) = \sigma(\text{BN}(\text{PConv}(\Gamma_d\{f_2(x, d)\}))), \quad (3)$$

式中: $f_3(x)$ 表示特征融合后的特征输出; $\text{PConv}(\cdot)$ 表示卷积核大小为 1×1 的逐点卷积; $\Gamma_d\{\cdot\}$ 表示连接操作。最后将合并后的特征与输入特征图进行残差连接, 得到输出特征图, 表达式为

$$\text{DWR}(x) = f_1(f_2(f_3)) \oplus x, \quad (4)$$

式中: $\text{DWR}(x)$ 表示扩张残差模块的输出; \oplus 表示相加。

使用 DWR 模块的两步残差化方法, 简化多速率

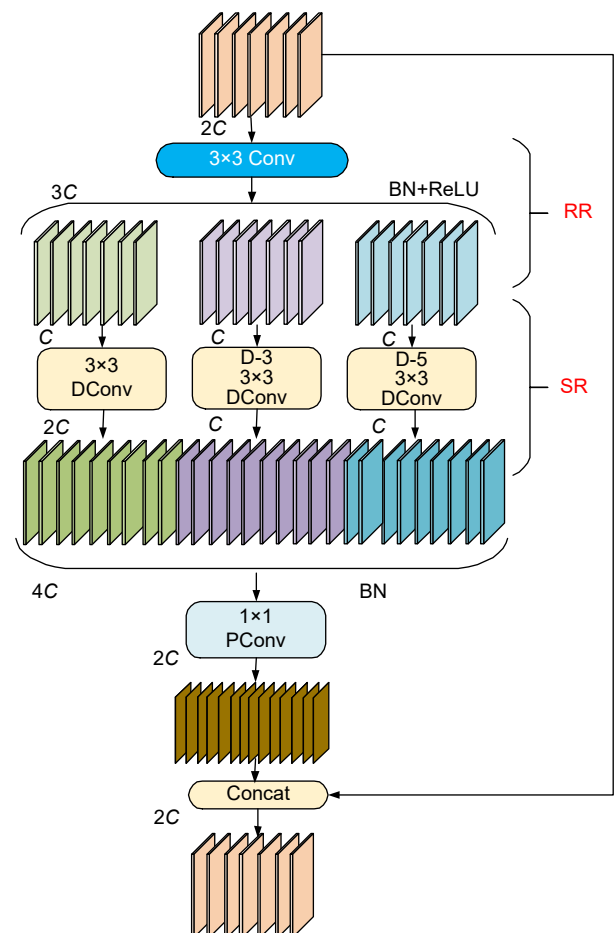


图 3 DWR 模块
Fig. 3 DWR module

深度扩张卷积, 将原本从复杂特征图中获取多尺度信息的过程, 转变为对简洁的特征图进行分析。这种方法使得深度扩张卷积的任务变得简单有序, 从而更有效地提取多尺度的上下文信息。本文通过添加修改后的 C2f_DWR 模块, 帮助模型更有效地融合不同层次的特征, 提高多尺度信息捕获的效率。

3.3 GD_CBAM 改善模型识别精度

卷积块注意力模块 (Convolutional block attention module, CBAM) 通过结合通道注意力模块 (channel attention module) 和空间注意力模块 (spatial attention module), 显著提升了模型的性能。然而, 由于部分焊缝缺陷与其周围环境(如相邻焊缝缺陷)的视觉相似度较高, 对于关键点的定位容易出现误差。为此, 本文在 YOLO v8 模型中引入改进后的卷积注意力模块 GD-CBAM, 通过引入 GD-CBAM 能够增强特征的表达和选择, 使模型更好地聚焦于缺陷的关键点特征, GD-CBAM 结构如图 4 所示。

神经网络的核心优势之一是其能够学习输入数据的复杂、非线性关系。GD-CBAM 注意力机制是在 CBAM 的基础上, 在通道注意力模块的两个全连接层 (FC) 之间加入 Dropout 层, Dropout 层会随机丢弃部分神经元, 使模型在训练过程中不过度依赖某些特征, 增强对新数据的检测能力, 防止过拟合。改进后

的通道注意力模块如图 5 所示, 通道注意力模块旨在通过全局信息聚合, 动态生成通道级别的注意力权重, 从而增强重要特征并抑制不相关特征, 改进后的具体步骤如下。

首先, 设输入特征图为 $X \in \mathbb{R}^{B \times C \times H \times W}$, 其中 B 表示批量大小, C 表示通道数, H 和 W 分别表示特征图的高度和宽度。通过全局最大池化和全局平均池化操作, 将特征图在空间维度上各压缩为一个向量。计算公式为

$$\text{maxpool_out} = \text{MaxPool}(X), \quad (5)$$

$$\text{avgpool_out} = \text{AvgPool}(X), \quad (6)$$

式中: $\text{MaxPool}(\cdot)$ 为全局最大池化操作; $\text{AvgPool}(\cdot)$ 为全局平均池化操作。将上述两个向量分别输入到一个共享的全连接层 (FC)。将第一个全连接层的输出(即降维后的特征)进行非线性变换, 引导第二个全连接层的输出相加, 具体计算过程可表示为

$$\text{maxfc_out} = \text{FC}(\text{Dropout}(\text{ReLU}(\text{maxpool_out}))), \quad (7)$$

$$\text{avgfc_out} = \text{FC}(\text{Dropout}(\text{ReLU}(\text{avgpool_out}))), \quad (8)$$

式中: $\text{FC}(\cdot)$ 为全连接层非线性组合操作; $\text{Dropout}(\cdot)$ 为正则化操作; $\text{ReLU}(\cdot)$ 为非线性拟合操作。然后, 通过 Sigmoid 激活函数生成通道注意力权重, 计算公式为 $\text{channelattention_map} = \sigma_s(\text{maxpool_out} + \text{avgpool_out})$, (9)

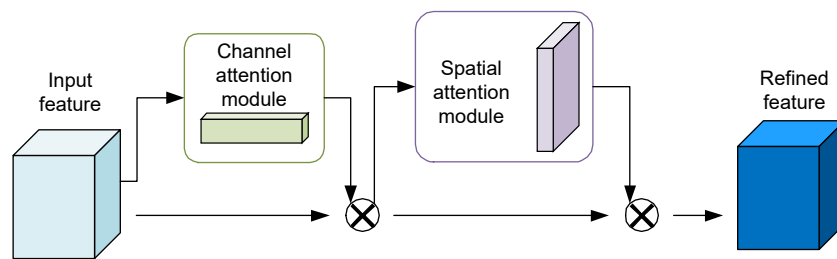


图 4 GD-CBAM 注意力模块

Fig. 4 GD-CBAM attention module

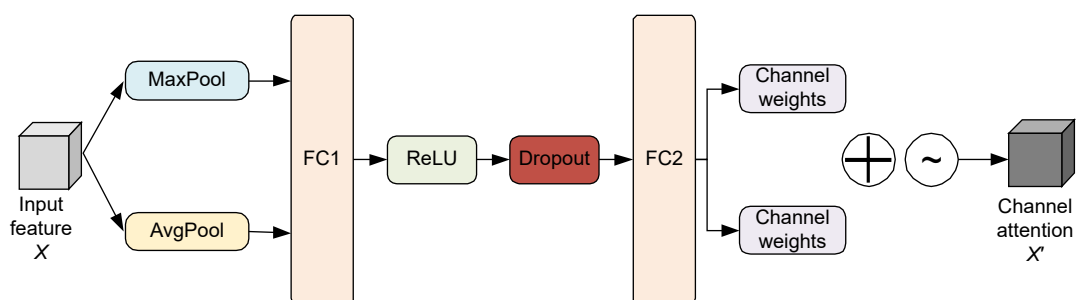


图 5 通道注意力模块

Fig. 5 Channel attention module

式中: $\sigma_s(\cdot)$ 为 Sigmoid 函数。

最后, 将通道注意力权重与输入特征图逐通道相乘, 得到增强后的特征图 X' , 计算公式为

$$X' = \text{channelattention_map} * X. \quad (10)$$

式中: “*” 表示相乘, 是卷积算法中的一步。相较于原注意力模块, 这种非线性变换使得网络能够学习和模拟更加复杂的函数。在模型中 GD-CBAM 被放置在每个特征图层级上的 C2f-DWR 卷积模块操作之后, 通过在特征图层级上使用 GD-CBAM, 帮助网络学习如何根据特征的重要性对它们进行加权, 从而优化模型性能。

3.4 Inner-SIoU 改进小目标识别精度

定位损失函数在目标检测中拥有着至关重要的作用, 合理的定义将给网络性能带来显著的提升, 在 YOLOv8 模型的边界框回归任务中, 采用 CIoU 损失函数。该损失函数综合考虑了预测边界框与真实边界框之间的中心点距离、重叠区域以及长宽比, 并提高了边界框回归的准确性。然而, 由于 CIoU 损失函数在预测框与真实框在方向上的存在不一致性, 以及 IoU 指标本身可能存在的合理性问题, 在一定程度上限制了模型在焊缝缺陷检测任务中的精度表现。为了克服这一局限性, 本研究引入了一种新的损失函数——Inner-SIoU 损失函数, 以替代原来的 CIoU 损失函数。

Inner-SIoU 是 Inner-IoU^[20] 和 SIoU^[21] 的结合, 它利用辅助包围框 (inner bounding box) 来加速回归过程。SIoU 损失函数具体包含四个部分: 角度成本 (angle cost)、距离成本 (distance cost)、形状成本 (shape cost)、IoU 损失 (IoU cost)。

损失函数 SIoU 计算公式为

$$L_{\text{SIoU}} = 1 - \text{IoU} + \frac{\Delta + \Omega}{2}, \quad (11)$$

式中: IoU 为 IoU 损失; Δ 为距离成本; Ω 为形状成本。对于角度成本, 其表达式为

$$\Lambda = 1 - 2 \times \sin^2 \left(\arcsin \left(\frac{C_h}{\sigma_d} \right) - \frac{\pi}{4} \right), \quad (12)$$

$$\frac{C_h}{\sigma_d} = \sin \alpha, \quad (13)$$

$$\sigma_d = \sqrt{(b_{C_x}^{\text{gt}} - b_{C_x})^2 + (b_{C_y}^{\text{gt}} - b_{C_y})^2}, \quad (14)$$

$$C_h = \max(b_{C_y}^{\text{gt}}, b_{C_y}) - \min(b_{C_y}^{\text{gt}}, b_{C_y}), \quad (15)$$

式中: σ_d 为真实框和预测框中心点的距离; C_h 为真实框和预测框中心点的高度差; $(b_{C_x}^{\text{gt}}, b_{C_y}^{\text{gt}})$ 为真实框的

中心坐标; (b_{C_x}, b_{C_y}) 为预测框的中心坐标; $\max(\cdot)$ 为取最大值操作; $\min(\cdot)$ 为取最小值操作。角度成本如图 6 所示。

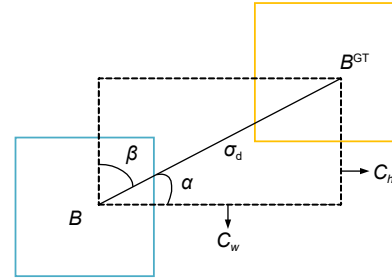


图 6 角度成本
Fig. 6 Angle cost

对于距离成本, 其表达式为

$$\Delta = \sum_{t=x,y} (1 - e^{-(2-\Lambda) \times \rho_t}), \quad (16)$$

$$\begin{cases} \rho_x = \left(\frac{b_{C_x}^{\text{gt}} - b_{C_x}}{C_w} \right)^2 \\ \rho_y = \left(\frac{b_{C_y}^{\text{gt}} - b_{C_y}}{C_h} \right)^2 \end{cases}, \quad (17)$$

式中: C_w 为真实框和预测框最小外接矩形的宽。距离成本如图 7 所示。

对于形状成本, 其表达式为

$$\Omega = \sum_{t=w,h} (1 - e^{-W_t}^\theta), \quad (18)$$

$$\begin{cases} W_w = \frac{|w - w^{\text{gt}}|}{\max(w, w^{\text{gt}})} \\ W_h = \frac{|h - h^{\text{gt}}|}{\max(h, h^{\text{gt}})} \end{cases}, \quad (19)$$

式中: θ 为对形状损失的关注程度; w^{gt} 和 h^{gt} 为真实框的宽、高; w 和 h 为预测框的宽、高。

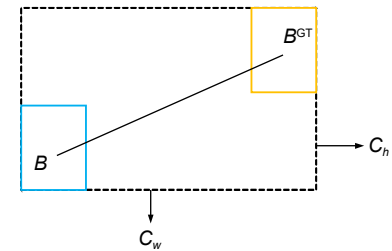


图 7 距离成本
Fig. 7 Distance cost

对于 IoU 损失, 结合图 6、图 7 可表示为

$$\text{IoU} = \frac{|B \cap B^{\text{GT}}|}{|B \cup B^{\text{GT}}|}, \quad (20)$$

式中: $B \cap B^{\text{GT}}$ 为真实框和预测框的交集; $B \cup B^{\text{GT}}$ 为

真实框和预测框的并集。

针对 SIoU 还存在实际使用中泛化性不佳的问题, 引入辅助边界框, 通过引入比例因子 r_a 来控制辅助边界框的大小, 结构如图 8 所示。

当 $r_a \in [0.5, 1)$ 时, 结构如图 8(a) 所示; 当 $r_a \in (1, 1.5]$ 时, 结构如图 8(b) 所示, 其中 (x_c^{gt}, y_c^{gt}) 为真实边界框和辅助边界框的中心坐标, (x_c, y_c) 为预测框和辅助边界框的中心点, w_1 和 h_1 为预测框的宽、高。

Inner-IoU 的过程计算公式为

$$\begin{cases} b_1^{gt} = x_c^{gt} - \frac{w^{gt} * r_a}{2} \\ b_r^{gt} = x_c^{gt} + \frac{w^{gt} * r_a}{2} \end{cases}, \quad (21)$$

$$\begin{cases} b_t^{gt} = y_c^{gt} - \frac{h^{gt} * r_a}{2} \\ b_b^{gt} = y_c^{gt} + \frac{h^{gt} * r_a}{2} \end{cases}. \quad (22)$$

由式 (21) 和式 (22) 可得到四个坐标 (b_1^{gt}, b_t^{gt}) 、 (b_r^{gt}, b_b^{gt}) 、 (b_l^{gt}, b_t^{gt}) 、 (b_r^{gt}, b_b^{gt}) , 这四个坐标为辅助边界框的四个顶点。

$$\begin{cases} b_l = x_c - \frac{w_1 * r_a}{2} \\ b_r = x_c + \frac{w_1 * r_a}{2} \end{cases}, \quad (23)$$

$$\begin{cases} b_t = y_c - \frac{h_1 * r_a}{2} \\ b_b = y_c + \frac{h_1 * r_a}{2} \end{cases}. \quad (24)$$

由式 (23) 和式 (24) 可得到四个坐标 (b_l, b_t) 、 (b_l, b_b) 、 (b_r, b_t) 、 (b_r, b_b) 这四个坐标为预测框的四个

顶点。

$$I = (\min(b_r^{gt}, b_r) - \max(b_l^{gt}, b_l)) * (\min(b_b^{gt}, b_b) - \max(b_t^{gt}, b_t)), \quad (25)$$

式中: I 为辅助边界框与辅助预测框的相交区域, 即图 8 中的紫色区域。

$$U = (w^{gt} * h^{gt}) * r_a^2 + (w_1 * h_1) * r_a^2 - I, \quad (26)$$

式中: U 为真实边界框与预测框的相交区域, 即图 8 中的蓝色区域。

由式 (25) 和式 (26) 可得 Inner-IoU 的交并比, 其表达式为

$$IoU^{inner} = \frac{I}{U}. \quad (27)$$

综上, 损失函数 Inner-SIoU 的计算公式为

$$L_{Inner-SIoU} = L_{SIoU} + IoU - IoU^{inner}. \quad (28)$$

4 实验结果与分析

4.1 实验环境及评价指标

实验硬件配置为 Intel (R) Core (TM) i5-13400 处理器, Nvidia RTX 4060ti (8 G) 显卡。软件系统采用 Windows 11 系统, 配套安装 CUDA 11.3, Pytorch 1.10, 编程语言是 Python 3.9.19, 所有输入图像的分辨率统一转换为 640 pixel×640 pixel。模型的训练周期为 300 个 Epoch、每个批次的训练样本数量 (batch size) 为 16、初始学习率 (learning rate) 设置为 0.01。

本文选取平均精度均值 (mAP)、平均精度 (AP)、召回率 (recall, R)、用准确率 (precision, P) 模型的

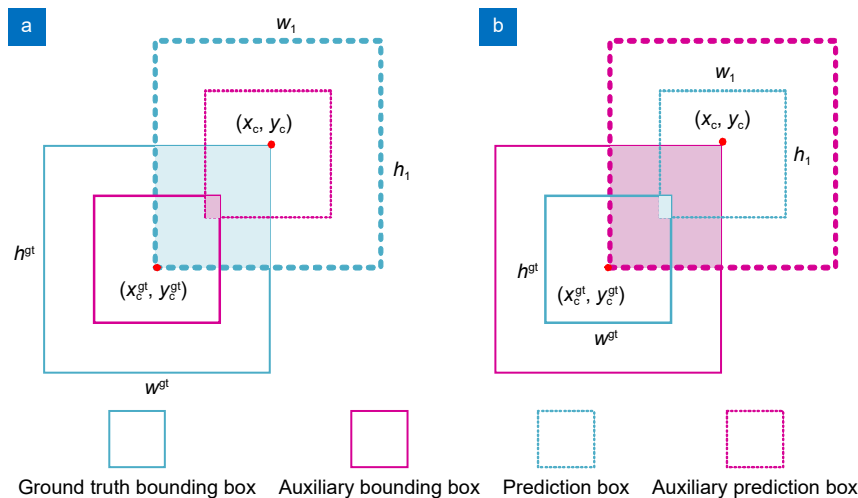


图 8 Inner-IoU 损失函数

Fig. 8 Inner-IoU loss function

参数量 (parameters)、模型计算量 (FLOPs)、帧率 (FPS) 作为性能评估指标。相关指标计算公式为

$$P = \frac{TP}{TP + FP}, \quad (29)$$

$$R = \frac{TP}{TP + FN}, \quad (30)$$

$$AP = \int_0^1 P(R) dR, \quad (31)$$

$$mAP = \frac{\sum_{i=1}^n AP(i)}{n}, \quad (32)$$

$$FPS = \frac{1}{Time}, \quad (33)$$

式中: TP 表示模型正确识别出缺陷的数量; FP 表示模型错误识别出非缺陷的数量; FN 表示模型未能识别出缺陷的数量; n 表示检测缺陷的类别数; AP 表示 PR 曲线下的面积; FPS 表示每秒模型处理的图像数量, 可描述模型检测的理论速度; $Time$ 表示处理每一帧图像所需的时间。 AP 的值反映了模型的平均准确率, mAP 是 IoU 阈值取 0.5 时计算出来的值, AP 和 $mAP0.5$ 的值越大, 代表模型检测性能越好。FLOPs 表示使用每秒浮点运算次数, FLOPs 越大表示占用的计算资源越多。

4.2 数据集

本文采用 Roboflow 网站开源的钢板焊缝缺陷检测数据集 Welding Defect Test-V2 和东北大学开源的带钢表面缺陷数据集 NEU-DET。

Welding Defect Test-V2 数据集共有 4 类缺陷, 分别是几何缺陷 (geometric defect)、非融合缺陷 (non-fusion defect)、孔隙 (porosity)、飞溅 (spatters), 共计 3866 张, 将这其中的 2706 张划分为训练集, 773 张划分为验证集, 387 张划分为测试集。

NEU-DET 数据集由划痕 (scratches, Sc)、裂纹 (crazing, Cr)、麻点 (pitted surface, Ps)、夹杂 (inclusion, In)、斑块 (patches, Pa)、压入氧化铁皮 (rolled in scale, Rs) 6 类缺陷组成。数据集总共包含 1800 张图像, 每类缺陷各 300 张, 其中训练集为 1440 张, 验证集为 180 张, 测试集 180 张。

4.3 损失函数优化实验

为验证不同 IoU 损失函数对模型性能的影响, 本文分别使用 CIoU、SIoU、GIoU^[22]、EIoU^[23]、DIoU^[24]、WIoU^[25] 和 Inner-SIoU 作为边界框回归损失函数在

Roboflow 开源数据集 Welding Defect Test-V2 上进行损失函数优化实验, 实验结果如表 1 所示。Inner-SIoU 在 $mAP0.5$ 方面的 CIoU 提升了 0.7%, 然而, 选择 SIoU、EIoU、DIoU 以及 GIoU 训练后, 模型的 $mAP0.5$ 值有所下降; 在召回率方面 Inner-SIoU 和 WIoU 表现最优。因此, 选用 Inner-SIoU 作为改进的损失函数能够更准确地预测缺陷目标的位置和大小。

表 1 损失函数对比

Table 1 Loss function comparison

IoU loss function	$mAP0.5/\%$	$R/\%$	Params/M
CIoU	94.6	89.3	2.89
SIoU	92.9	85.1	2.89
GIoU	94.0	88.9	2.89
EIoU	93.3	90.3	2.89
DIoU	93.8	91.7	2.89
WIoU	95.0	89.4	2.89
Inner-SIoU	95.3	89.4	2.89

4.4 消融试验

为了验证本文所提出的各项模块对于模型的提升效果, 在 Roboflow 开源数据集 Welding Defect Test-V2 上进行消融实验, 设计了 7 组消融实验, 实验结果如表 2 所示。

根据表 2 的实验结果, 改进的注意力机制模块 GD-CBAM 有效提高模型检测缺陷的精确率, 降低了误判风险。加入 CARAFE 上采样算子, 其自适应内核更好地实现上采样操作, 检测精度为 94.2%, 相比基线模型提升了 6.7 个百分点, 计算量增加了 0.3 G, 参数量增加 5%。C2f-DWR 模块对模型召回率提升最为显著, $mAP0.5$ 提升 7.6 个百分点, 通过多分支结构扩张感受野捕捉丰富的上下文信息, 降低了缺陷的漏检率。然后将颈部 C2f 结构中替换为轻量化卷积神经网络 GSConv 和 VoV-GSCSP 后, 模型参数量减少了 8%, FLOPs 也有所降低, 表明 GSConv 结构和 VoV-GSCSP 结构能够最大限度地减少特征图中的冗余, 节约了计算成本。最后加入 Inner-SIoU 损失函数提升小目标检测精度, 检测精度达到 95.3%。

4.5 不同目标检测算法对比实验

为了全面评价本文所提出模型在焊缝缺陷检测中的有效性, 将 GD-YOLO 模型与各 YOLO 目标检测模型进行对比, 实验结果如表 3 所示。

改进后的算法 GD-YOLO 在 $mAP0.5$ 上相比于基线 YOLOv8n 算法提高了 7.8 个百分点。通过与传统

检测算法 Faster R-CNN 对比本文模型在各项指标上均具有更好的表现。而在与 YOLOv3-tiny、YOLOv5n、YOLOv6n、YOLOv9t、YOLOv10n、RT-DETR、YOLOv11n 算法相比, 指标 mAP0.5 分别提升了 1.9%、9.6%、10.3%、7.9%、1.0%、4.6%、1.1%, 涨幅明显; 在参数和计算量方面, 相对于 YOLOv3-tiny, 分别减少 9.24 M、11.6 G; 相对于参数量较低的 YOLOv5n 算法, 参数量和计算量相仿; 与基线模型相比分别减少了 0.2 M、0.7 G。因此证明本文算法拥有低计算量, 高精确度的特点, 具备良好的检测性能, 并且算法每秒帧数可达到 257 f/s, 满足工业生产场景下实时监测的需求。表 3 对比结果说明进一步验证了本文算法在焊缝表面缺陷检测任务中的性能优势。

4.6 模型检测效果展示

图 9 展示了 YOLOv8 算法和 GD-YOLO 算法在 Roboflow 开源数据集 Welding Defect Test-V2 上的精确-召回率 (P - R) 对比曲线。

图 10 展示了对比实验中检测精度较高的 4 种算

法在测试集上的检测效果。由图 10 可知, 相比于其他算法, 本文提出的 GD-YOLO 能够更准确地检测出小尺寸缺陷, 有效地减少了漏检、误检的情况。同时, 从图中目标框上方的置信度得分可以看出, GD-YOLO 的分类结果更具有可靠性。

4.7 模型泛化性实验

为了验证 GD-YOLO 算法的泛化能力, 本文采用 NEU-DET 数据集进行了相应的泛化实验。在本次实验中对比了 YOLOv5m、YOLOv7、YOLOv8s、YOLOv8n、YOLOv9t、YOLOv11n、RT-DETR 以及先进的 YOLO 改进模型 WFRE-YOLOv8s^[26], 实验环境以及参数配置均遵循 4.1 节中的描述, 以确保实验的可靠性和一致性。实验结果已汇总在表 4。

通过表 4 将算法检测结果进行对比可以发现, 本文模型在 NEU-DET 数据集上 mAP 可达到 78.8%, 与 YOLOv5m、YOLOv7、YOLOv8s、YOLOv9t、YOLOv11n、RT-DETR 模型相比, 指标 mAP 分别提升 4.0%、5.1%、4.1%、5.0%、0.9%、2.1%。与本文

表 2 消融实验结果

Table 2 Ablation experiments results

C2f-DWR	CARAFE	GSCConv	Inner-SIoU	GD-CBAM	mAP0.5/%	P /%	R /%	FLOPs/G	Params/M	FPS
					87.5	85.5	81.4	8.1	3.01	247
√					95.1	91.1	88.9	8.0	2.95	286
	√				94.2	90.6	88.7	8.4	3.15	132
		√			94.4	89.9	90.1	7.3	2.79	273
				√	90.9	89.1	88.3	8.1	3.01	259
√				√	95.9	92.0	89.5	11.6	3.35	318
√			√	√	96.5	92.3	89.9	11.6	3.35	322
√	√	√		√	94.6	91.2	89.3	7.4	2.89	243
√	√	√	√	√	95.3	90.7	89.4	7.4	2.89	257

表 3 对比实验

Table 3 Results of comparison experiments

Model	mAP0.5/%	P /%	R /%	FLOPs/G	Params/M	FPS
YOLOv8n(baseline)	87.5	85.5	81.4	8.1	3.01	247
Faster R-CNN	79.8	83.1	77.9	33.3	41.20	95
YOLOv3-tiny	93.4	93.0	88.1	19.0	12.13	242
YOLOv5n	85.7	84.5	78.7	7.1	2.51	282
YOLOv6n	85.0	84.2	78.1	11.8	4.23	285
YOLOv9t	87.4	85.2	80.1	7.6	1.97	229
YOLOv10n	94.3	88.3	88.6	6.5	2.26	292
RT-DETR	90.7	89.5	82.9	110.2	32.01	117
YOLOv11n	94.2	89.1	89.6	6.3	2.58	307
Ours	95.3	90.7	89.4	7.4	2.89	257

基线模型 YOLOv8n 相比, 指标 mAP 提升 1.5%。虽然 WFRE-YOLOv8s 比本模型平均精度高 0.6%、召回率高 0.2%, 但其准确率下降 1.6%, 而且其模型计算量过大, 对实际工业生产场景下的检测终端性能要求较高。以上结果证明, 在同一计算量的情况下, 本文模型拥有更好的性能, 由此说明 GD-YOLO 算法不仅在焊缝数据集上对具有较好的检测能力, 在 NEU-DET 数据集上也具有良好的表现, 进一步证明改进

的有效性。

5 结论

本文提出了一种改进 YOLOv8 模型 GD-YOLO。通过引入高效的特征提取模块, 显著增强了模型实时检测信息的能力。同时, 对原有注意力机制模块进行改进, 使模型更好地聚焦于关键点特征, 并引入 Slim-Neck 结构化设计和 CAFARE 上采样算子, 辅助提高

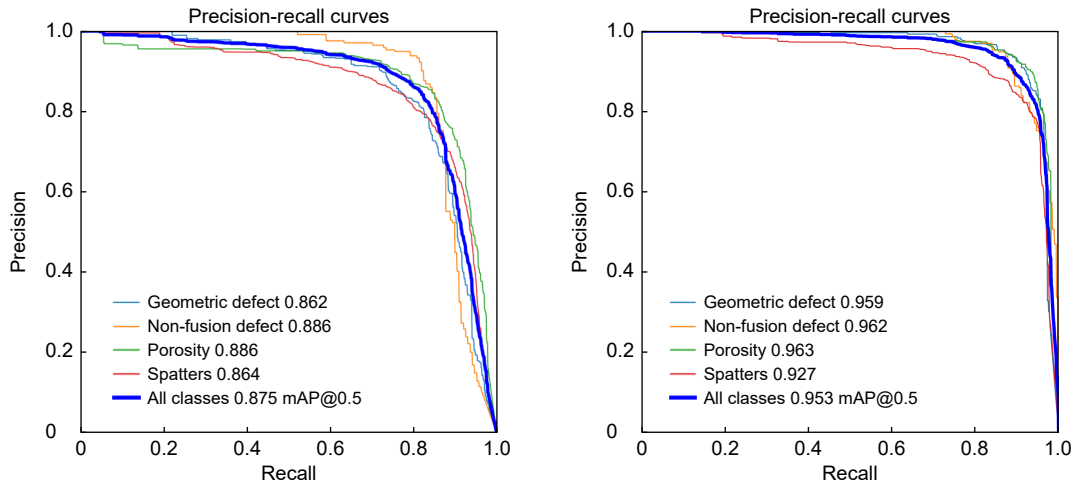


图 9 改进前后 P-R 曲线图

Fig. 9 P-R curves before and after improvement

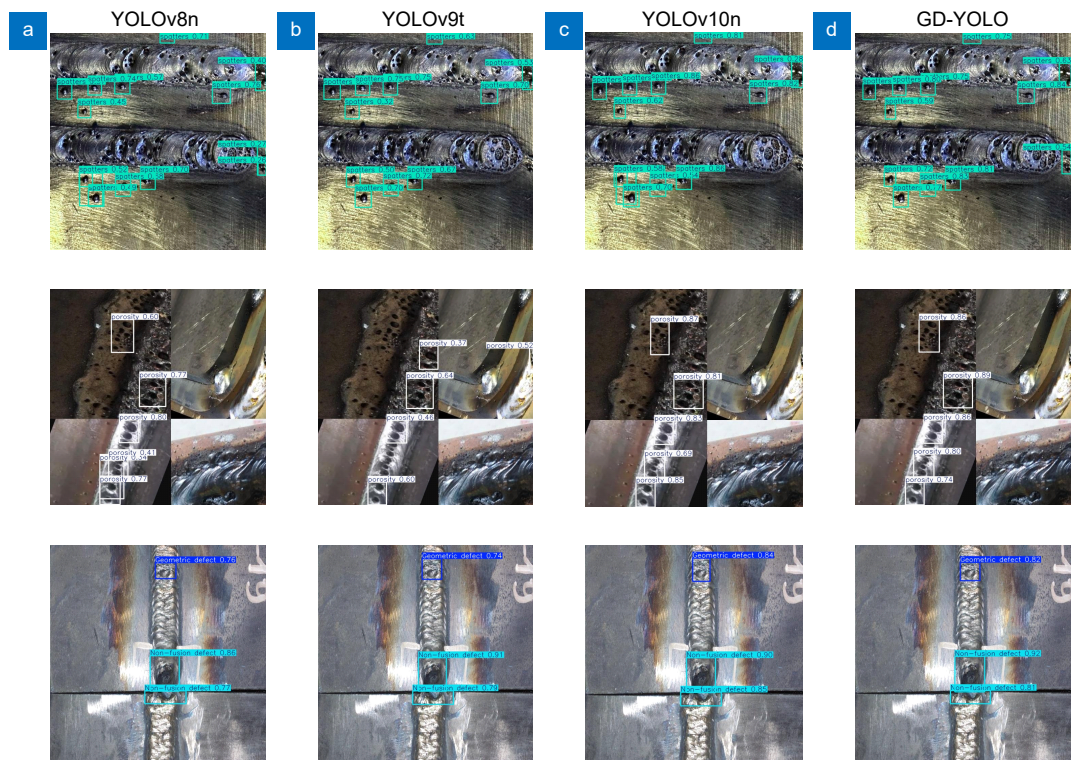


图 10 检测效果对比

Fig. 10 Comparison of detection effects

表 4 NEU-DET 数据集对比实验结果
Table 4 Experimental results of NEU-DET dataset

Model	mAP@0.5/%	P/%	R/%	FLOPs/G
YOLOv5m	74.8	71.2	69.0	48.0
YOLOv7	73.7	66.3	68.6	104.7
YOLOv8s	74.7	69.1	69.0	28.4
YOLOv8n	77.3	71.9	70.0	8.1
YOLOv9t	73.8	69.7	70.3	7.6
YOLOv11n	77.9	73.9	74.6	6.5
RT-DETR	76.7	72.4	70.2	110.2
WFRE-YOLOv8s	79.4	73.6	75.9	32.6
Ours	78.8	75.2	75.7	7.4

特征图的分辨和语义信息的传递。最后, 加入 Inner-SIoU 损失函数, 进一步提高边框回归的精度。实验表明, GD-YOLO 能更好地应用于焊缝表面缺陷检测任务中。但本文模型计算量和参数量方面还存在改进空间, 未来的研究方向在于如何部署在资源有限的移动端或嵌入式设备, 以期算法在焊缝缺陷检测任务中展现出更高的应用潜力和价值。

利益冲突: 所有作者声明无利益冲突

参考文献

- 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.
- Su H, Zhang J B, Zhang B H, et al. Review of surface defect inspection based on visual perception[J]. *Comput Integr Manuf Syst*, 2023, 29(1): 169–191.
苏虎, 张家斌, 张博豪, 等. 基于视觉感知的表面缺陷检测综述[J]. *计算机集成制造系统*, 2023, 29(1): 169–191.
- 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.
- Liu W, Anguelov D, Erhan D, et al. SSD: single shot multibox detector[C]//*Proceedings of the 14th European Conference on Computer Vision*, Amsterdam, 2016: 21–37.
https://doi.org/10.1007/978-3-319-46448-0_2.
- Redmon J. You only look once: unified, real-time object detection[C]//*Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition*, 2016: 779–788.
<https://doi.org/10.1109/CVPR.2016.91>.
- Wang C, Bochkovskiy A, Liao H M. YOLOv7: trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[C]//*Proceedings of 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022: 7464–7475.
<https://doi.org/10.1109/CVPR52729.2023.00721>.
- Girshick R. Fast R-CNN[C]//*Proceedings of 2015 IEEE International Conference on Computer Vision*, Santiago, 2015: 1440–1448. <https://doi.org/10.1109/ICCV.2015.169>.
- He K M, Gkioxari G, Dollár P, et al. Mask R-CNN[C]//*Proceedings of 2017 IEEE International Conference on Computer Vision*, 2017: 2980–2988.
<https://doi.org/10.1109/ICCV.2017.322>.
- Sun X D, Wu P C, Hoi S C H. Face detection using deep learning: an improved faster RCNN approach[J]. *Neurocomputing*, 2018, 299: 42–50.
- Mu C Y, Li C, Ma X, et al. Improved YOLOv7-tiny lightweight large casting weld defect detection[J]. *Modular Mach Tool Autom Manuf Tech*, 2024, (7): 156–160.
穆春阳, 李闯, 马行, 等. 改进 YOLOv7-tiny 的轻量化大型铸件焊缝缺陷检测[J]. *组合机床与自动化加工技术*, 2024, (7): 156–160.
- Li Q P, Miao H B, Li Z W, et al. Weld surface defect detection based on the improved YOLOv7-tiny algorithm[J]. *Weld Technol*, 2024, 53(7): 123–126.
李其鹏, 缪海波, 李志文, 等. 基于改进 YOLOv7-tiny 焊缝表面缺陷检测算法[J]. *焊接技术*, 2024, 53(7): 123–126.
- Song J S. Research on weld defect detection based on improved Yolov8[J]. *China Plant Engineering*, 2024, (14): 196–199.
宋杰三. 基于改进 Yolov8 的焊缝缺陷检测研究[J]. *中国设备工程*, 2024, (14): 196–199.
- Zhang E Z, Li L, Wang J H, et al. Improved YOLOv8-based quality inspection method for ship plate welds[J]. *J Shaanxi Univ Sci Technol*, 2024, 42(6): 172–179.
张恩召, 李磊, 汪建华, 等. 基于改进 YOLOv8 的船板焊缝质量检测方法[J]. *陕西科技大学学报*, 2024, 42(6): 172–179.
- Li X W, He S Q, He D Q, et al. Lightweight detection method for weld defects of metro train based on improved YOLOv8[J]. *J Guangxi Univ (Nat Sci Ed)*, 2024, 49(3): 540–552.
李先旺, 贺岁球, 贺德强, 等. 基于改进 YOLOv8 的地铁列车焊缝缺陷轻量化检测方法[J]. *广西大学学报(自然科学版)*, 2024, 49(3): 540–552.
- Wu L, Chu Y K, Yang H C, et al. Sim-YOLOv8 object detection model for DR image defects in aluminum alloy welds[J]. *Chin J Lasers*, 2024, 51(16): 1602103.
吴磊, 储钰昆, 杨洪刚, 等. 面向铝合金焊缝 DR 图像缺陷的 Sim-YOLOv8 目标检测模型[J]. *中国激光*, 2024, 51(16): 1602103.
- Lin T Y, Dollár P, Girshick R, et al. Feature pyramid networks for object detection[C]//*Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition*, Honolulu, 2017: 936–944. <https://doi.org/10.1109/CVPR.2017.106>.
- Wei H R, Liu X, Xu S C, et al. DWRSeg: rethinking efficient acquisition of multi-scale contextual information for real-time semantic segmentation[Z]. arXiv: 2212.01173, 2022.
<https://arxiv.org/abs/2212.01173>.

- [18] Wang J, Chen K, Xu R, et al. CARAFE: content-aware ReAssembly of features[C]//*Proceedings of 2019 IEEE/CVF International Conference on Computer Vision*, 2019: 3007–3016. <https://doi.org/10.1109/ICCV.2019.00310>.
- [19] Li H L, Li J, Wei H B, et al. Slim-neck by GSConv: a lightweight-design for real-time detector architectures[J]. *J Real Time Image Process*, 2022, 21(3): 62.
- [20] Zhang H, Xu C, Zhang S J. Inner-IoU: more effective intersection over union loss with auxiliary bounding box[Z]. arXiv: 2311.02877, 2023. <https://arxiv.org/abs/2311.02877>.
- [21] Gevorgyan Z. SloU loss: more powerful learning for bounding box regression[Z]. arXiv: 2205.12740, 2022. <https://arxiv.org/abs/2205.12740>.
- [22] Rezatofghi H, Tsoi N, Gwak J, et al. Generalized intersection over union: a metric and a loss for bounding box regression[C]//*Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019: 658–666. <https://doi.org/10.1109/CVPR.2019.00075>.
- [23] Zhang Y F, Ren W Q, Zhang Z, et al. Focal and efficient IOU loss for accurate bounding box regression[J]. *Neurocomputing*, 2022, 506: 146–157.
- [24] Zheng Z H, Wang P, Liu W, et al. Distance-IoU loss: faster and better learning for bounding box regression[C]//*Proceedings of the 34th AAAI Conference on Artificial Intelligence*, 2020: 12993–13000. <https://doi.org/10.1609/aaai.v34i07.6999>.
- [25] Tong Z J, Chen Y H, Xu Z W, et al. Wise-IoU: bounding box regression loss with dynamic focusing mechanism[Z]. arXiv: 2301.10051, 2023. <https://arxiv.org/abs/2301.10051>.
- [26] Huang Y, Tan W, Li L, et al. WFRE-YOLOv8s: a new type of defect detector for steel surfaces[J]. *Coatings*, 2023, 13(12): 2010.

作者简介



张润梅 (1971-), 女, 博士, 教授, 主要研究方向为机器人技术、图像处理。

E-mail: zrmahjzu11@163.com



【通信作者】陈梓华 (1993-), 男, 博士, 讲师, 主要研究方向为智能协同控制, 机器视觉。

E-mail: czh1619@ahjzu.edu.cn



潘晨飞 (2000-), 男, 硕士研究生, 主要研究方向为深度学习、图像处理。

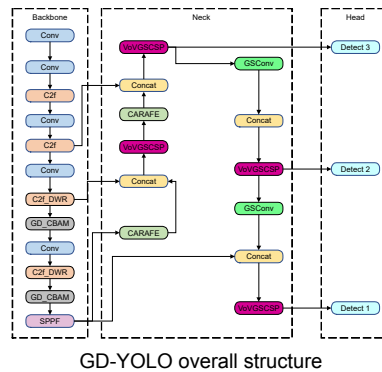
E-mail: 1344980647@qq.com



扫描二维码, 获取PDF全文

Improved weld surface defect detection algorithm from YOLOv8

Zhang Runmei¹, Pan Chenfei², Chen Zihua^{1*}, Chen Zhong¹, Yuan Bin¹



GD-YOLO overall structure

Overview: The welding quality of steel thin plates is of paramount importance in modern industrial production, with its widespread use in sectors such as intelligent manufacturing, industrial construction, and aerospace. Weld defects can lead to significant safety and performance issues, making the detection of these flaws a critical task. In light of this, this paper introduces an enhanced YOLOv8-based algorithm for weld surface defect detection, named GD-YOLO. The development of GD-YOLO begins with the introduction of a novel feature extraction module, C2f-DWR, which integrates the lightweight attention mechanism DWR with the feature extraction module C2f. This replacement of the original C2f module is designed to boost the model's efficiency in gathering information during real-time detection. The integration of the Slim-Neck structure into the neck network further reduces the algorithm's complexity and enhances its ability to detect the rough edges of defects. The algorithm also incorporates the upsampling operator CAFARE in the feature fusion stage, replacing the traditional Upsample operator. This change is aimed at increasing the resolution of feature maps and the transmission of semantic information, which is vital for accurate defect detection. Additionally, the improved attention mechanism module GD-CBAM is incorporated into the backbone network of YOLOv8, which not only accelerates the inference speed but also ensures that the model remains lightweight and efficient. To address the common issue of mismatch between the true and predicted bounding boxes, the GD-YOLO model employs the Inner-SIOU bounding box regression loss function. This function is specifically designed to minimize the discrepancies between the actual defect locations and the model's predictions. Empirical evidence from experiments demonstrates that the proposed GD-YOLO model outperforms the original YOLOv8 by 7.8% in the mAP0.5 detection metric, a significant improvement in accuracy. Moreover, the model shows a reduction of 0.2 M in parameter quantity and 0.7 G in computational load, making it more efficient than its predecessor. Compared to other target defect detection models, GD-YOLO exhibits a clear advantage in terms of detection accuracy. Ablation experiments conducted to validate the effectiveness of each module within the GD-YOLO framework confirm that each component contributes positively to the overall performance of the model. Furthermore, generalization experiments substantiate the improved algorithm's ability to perform well across different datasets, indicating its robustness and versatility in various real-world applications. In conclusion, GD-YOLO represents a significant advancement in steel thin plate weld defect detection, offering a more accurate, efficient, and reliable solution for industrial quality control.

Zhang R M, Pan C F, Chen Z H, et al. Improved weld surface defect detection algorithm from YOLOv8[J]. *Opto-Electron Eng*, 2025, 52(3): 240296; DOI: 10.12086/oe.2025.240296

Foundation item: Anhui Provincial Natural Science Projects in Universities (2024AH050234), Open Research Project of Anhui Simulation Design and Modern Manufacture Engineering Technology Research Center (SGCZXZD2101), Provincial and Ministerial Key Laboratory of Chang'an University (300102254501-202412), and Anhui University Natural Science Key Research Project (2024AH050245)

¹School of Mechanical and Electrical Engineering, Anhui Jianzhu University, Hefei, Anhui 230601, China; ²School of Electronic and Information Engineering, Anhui Jianzhu University, Hefei, Anhui 230601, China

* E-mail: czh1619@ahjzu.edu.cn