

# MEA-YOLO:基于多尺度边缘增强与注意力融合的钢材缺陷检测<sup>\*</sup>

曹万里 刘 强 马圆捷 解峪坤  
(山东理工大学机械工程学院 淄博 255022)

**摘 要:** 钢材表面缺陷检测对工业质量控制具有重要意义,但复杂纹理与多尺度特征增加了检测难度。为此,提出了一种基于多尺度边缘增强与注意力融合的轻量化检测网络(MEA-YOLO)。所提方法以 StarNet 替代 YOLOv11 主干以降低复杂度,设计多尺度边缘信息增强模块结合双域选择机制强化边界与上下文特征,并在检测头中引入空间增强前馈网络提升细粒度识别能力。实验结果表明,所提方法在钢材缺陷检测任务中的 mAP50 达到 74.58%,较 YOLOv11n 提升 2.01%,同时参数量与 GFLOPs 分别减少 14.3%和 10.1%。此外,在 GC10-DET 数据集上的对比实验同样取得了精度与速度的同步提升,进一步证明了模型的跨数据集泛化能力与工业实时检测的适用性。

**关键词:** 钢材缺陷检测;YOLOv11;多尺度边缘增强;注意力机制;轻量化网络

**中图分类号:** TN911 **文献标识码:** A **国家标准学科分类代码:** 510.4050

## MEA-YOLO: Steel surface defect detection via multi-scale edge enhancement and attention fusion

Cao Wanli Liu Qiang Ma Yuanjie Xie Yukun

(School of Mechanical Engineering, Shandong University of Technology, Zibo 255022, China)

**Abstract:** Steel surface defect detection plays a vital role in industrial quality control, yet complex textures and multi-scale characteristics increase the difficulty of accurate detection. To address this, a lightweight detection network named MEA-YOLO, based on multi-scale edge enhancement and attention fusion, is proposed. The approach replaces the YOLOv11 backbone with StarNet to reduce computational complexity, introduces a multi-scale edge information enhancement module combined with a dual-domain selection mechanism to strengthen boundary and contextual features, and incorporates a spatially enhanced feedforward network in the detection head to improve fine-grained recognition. Experimental results show that the proposed method achieves an mAP50 of 74.58% on steel defect detection tasks, outperforming YOLOv11n by 2.01%, while reducing the number of parameters and GFLOPs by 14.3% and 10.1%, respectively. Additional evaluations on the GC10-DET dataset further demonstrate consistent improvements in both accuracy and inference speed, confirming the model's robustness, generalization capability, and suitability for real-time industrial inspection.

**Keywords:** steel defect detection; YOLOv11; multi-scale edge enhancement; attention mechanism; lightweight network

## 0 引 言

钢铁是现代工业中最为重要的基础材料之一,其表面质量直接影响产品的力学性能、耐腐蚀性和后续加工工艺的稳定性。因此,研发高精度、高效率且可工业部署的钢材表面缺陷检测算法,对实现智能化质量管控具有重要意义。

传统钢材表面缺陷检测主要依赖人工检测<sup>[1]</sup>和基于手

工特征的机器视觉方法<sup>[2]</sup>。随着深度学习的发展,基于卷积神经网络(convolutional neural network, CNN)的检测方法在工业视觉检测中得到广泛应用。两阶段检测器方面,Faster R-CNN<sup>[3]</sup>能够实现较高检测精度,但检测速度较慢,难以满足实时需求。冷岳峰等<sup>[4]</sup>提出在特征金字塔网络(feature pyramid networks, FPN)和区域提议网络(region proposal network, RPN)之间加入特征融合与轻量

化通道注意模块,增强对细小缺陷的特征捕获能力。另一类研究基于 RetinaNet 进行了优化,王伟家等<sup>[5]</sup>通过轻量化骨干网络、深度可分离卷积与空间金字塔池化机制提升检测精度并降低参数量,实现生产线在线检测部署。单阶段检测器(you only look once, YOLO)系列以其端到端架构和实时性优势在钢材表面缺陷检测中应用最为广泛<sup>[6]</sup>。乔庆元等<sup>[7]</sup>提出 YOLO-RSI,改进下采样模块和多尺度特征融合结构,并通过 Inner-WIoU 优化回归损失函数,显著降低误检率。刘翰林等<sup>[8]</sup>提出 YOLOv5s-CBC,引入卷积注意力模块(convolutional block attention module, CBAM)与轻量化上采样算子(lightweight upsampling operator, CARAFE)增强特征表达能力。梁礼明等<sup>[9]</sup>的 FMG-YOLOv8s 采用轻量化 FasterNet 骨干和 M-C2f 模块,有效提升检测速度,并改善多尺度特征提取。Tie 等<sup>[10]</sup>提出 LSKA-YOLOv8,引入大核注意力模块并优化特征金字塔结构。梁礼明等<sup>[11]</sup>提出 DCD-YOLOv8n,通过多分支特征聚合与可变形注意力机制实现资源消耗下降并提高平均精度均值(mean average precision, mAP)。Liu 等<sup>[12]</sup>设计 Resilient YOLO 方法,使用轻量卷积模块和坐标注意机制,降低计算开销同时提升检测精度和鲁棒性。除 YOLO 系列外,部分研究探索了其他框架的改进方向。张伟健等<sup>[13]</sup>基于实时检测 Transformer(real-time detection Transformer, RT-DETR)提出轻量化 RepNCSPELAN4 模块与滑动可变焦点损失(slide varifocal loss, SVL),有效提升难样本学习能力并降低参数量和计算量。张航等<sup>[14]</sup>提出 HGnetv2+注意力机制的改进 YOLOv5 算法。Huang 等<sup>[15]</sup>提出 SSA-YOLO,结合卷积挤压与激励机制(convolution squeeze-and-excitation, CSE)、Swin Transformer 和自适应空间特征融合机制(adaptive spatial feature fusion, ASFF)强化多尺度特征交互。冯夫健等<sup>[16]</sup>提出了基于自注意特征融合的检测模型 SFNet,结合 Swin Transformer 与坐标注意力机制,实现多尺度特征高效融合,提升小目标缺陷检测能力。黄志海等<sup>[17]</sup>提出了一种基于多光谱中期融合的 IC 器件缺陷检测方法,在 YOLO 框架中引入双路特征提取与多光谱图像特征融合(multispectral image feature fusion, MIFF),显著提升轻微缺陷的检出性能。

现有基于 YOLO 的改进方法多聚焦于局部结构的替换或单一模块的增强,例如轻量卷积、通道注意力或特征金字塔优化等,虽然在一定程度上提升了检测性能,但仍难以同时满足钢材缺陷检测中对边界细节、尺度一致性与推理效率的综合要求。为解决这些问题,MEA-YOLO 从特征抽取、边缘强化和跨尺度融合 3 个层面构建了较为完整的改进路径。主干网络采用 StarNet,其卷积结构能够以较低计算成本保持底层纹理与结构特征的表达质量;多尺度边缘信息增强模块(multi-scale edge information enhancement, MSEIE)通过尺度调节与差分增强突出弱边界区域,并结合双城选择机制(dual-stream mechanism, DSM)实现空间

域与通道域的联合筛选,从而改善复杂背景下细粒度缺陷的可分性;检测头中引入的空间增强融合网络(spatial-enhanced fusion network, SEFN)依据空间依赖建模原理增强跨尺度语义交互,使高层语义特征能够更有效地补充细节特征。三者形成由低层到高层的递进式特征增强链路,与多数仅在局部结构上进行替换的 YOLO 改进方法相比,具有更强的整体协同性与表达能力。实验结果表明,该方法在钢材表面缺陷数据集上兼具较高精度与实时性,验证了其在工业在线检测场景中的应用价值。

## 1 方 法

### 1.1 网络总体结构

本文提出的多尺度边缘增强与注意力融合钢材缺陷检测网络(MEA-YOLO)基于 YOLOv11 框架构建,在保持端到端检测架构的基础上,通过骨干网络轻量化、特征提取增强以及多尺度特征交互优化,实现了钢材缺陷检测任务中精度与实时性的兼顾。网络整体结构如图 1 所示。主要由骨干网络(backbone)、特征融合网络(neck)和检测头(head)3 部分组成。

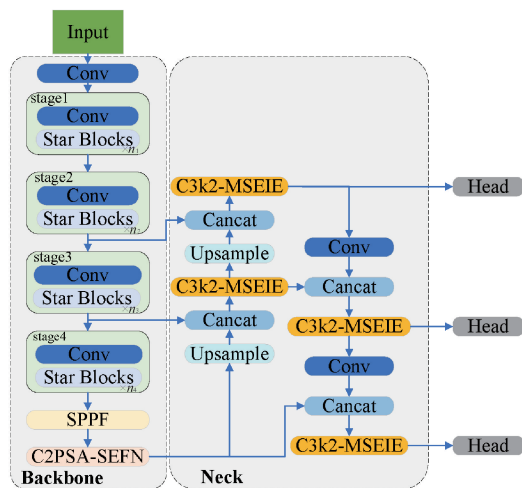


图 1 MEA-YOLO 网络总体结构

Fig. 1 Overall architecture of the MEA-YOLO network

#### 1) 轻量化骨干网络

骨干部分采用轻量化的 StarNet 结构替代 YOLOv11 原有骨干,以显著降低参数量与 GFLOPs。StarNet 由多个 Star Blocks 组成,每个阶段均包含一层卷积与若干个 Star Block 堆叠,用于提取不同层次的语义与空间特征。随着网络的加深,特征图在空间分辨率逐步下降的同时,通道数逐步增加,从而获得多尺度特征表示。

#### 2) 特征融合与多尺度边缘信息增强

Neck 部分针对不同尺度特征图引入改进的 C3k2-MSEIE 模块,替代原有 C3k2 结构,以强化边缘与细粒度缺陷特征的表达能力。该模块结合多尺度边缘提取与 DSM,在保留目标结构轮廓的同时有效抑制背景干扰。上采样

(upsample)与特征拼接(concat)操作确保了高分辨率细节信息与低分辨率语义信息的充分融合,从而提升不同尺寸缺陷的感知能力。

### 3) 高效特征交互与检测头

在最高层语义特征与融合特征之间,引入 C2PSA-SEFN 替换原有的 C2PSA 结构,实现跨尺度特征的高效融合与全局上下文增强。最终,不同尺度的特征图分别输入 3 个检测头(head),预测目标位置、类别及置信度。该多尺度检测机制可同时兼顾钢材表面微小缺陷与大范围缺陷的检测需求。

## 1.2 骨干网络

为兼顾钢材表面缺陷检测任务中的实时性与检测精度,本文在 YOLOv11 框架的骨干网络部分引入了轻量化且高效的 StarNet 结构<sup>[18]</sup>。骨干网络主要由 4 个阶段(Stage1~Stage4)组成,每一阶段均包含卷积层与若干个核心模块 Star Blocks。通过逐级下采样,骨干网络能够提取从浅层纹理细节到深层语义信息的多尺度特征,为后续的特征融合与检测提供充分的特征表达能力。

Star Blocks 是骨干网络的核心构建单元,其结构如图 2 所示,由深度可分离卷积(depthwise separable convolution, DWConv)、全连接层(fully connected, FC)、批归一化(batch normalization, BN)以及非线性激活函数(rectified linear unit 6, ReLU6)组成。该模块的设计思想在于利用通道间的关系提升特征表达能力,并通过轻量化操作降低计算开销。首先,输入特征经过深度可分离卷积提取空间特征,并通过批归一化稳定特征分布。随后,特征向量通过全连接层进行通道映射,其中一条分支经过 ReLU6 激活,另一条分支保持线性传递,二者相乘实现特征的高维映射。最后,融合后的特征通过另一全连接层和深度可分离卷积恢复空间信息并输出。

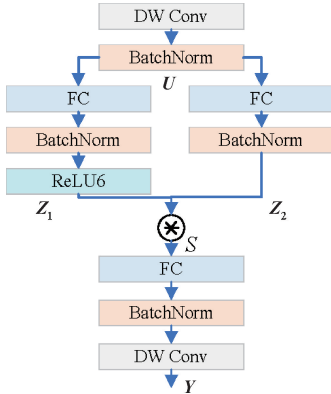


图 2 Star Blocks 模块结构

Fig. 2 Structure of the Star Blocks module

设输入特征图为  $\mathbf{X} \in \mathbb{R}^{C \times H \times W}$ , 则 Star Blocks 的计算过程可表示为:

$$\mathbf{U} = \text{BN}(\text{DWConv}(\mathbf{X})) \quad (1)$$

$$\mathbf{Z}_1 = \text{ReLU6}(\text{BN}_1(\text{FC}_1(\mathbf{U}))) \quad (2)$$

$$\mathbf{Z}_2 = \text{BN}_2(\text{FC}_2(\mathbf{U})) \quad (3)$$

$$\mathbf{S} = \mathbf{Z}_1 \odot \mathbf{Z}_2 \quad (4)$$

$$\mathbf{Y} = \text{DWConv}(\text{BN}_3(\text{FC}_3(\mathbf{S}))) \quad (5)$$

其中,  $\mathbf{U}$  表示中间特征,  $\mathbf{Z}_1$  表示第 1 分支特征,  $\mathbf{Z}_2$  表示第 2 分支特征,  $\mathbf{S}$  表示两分支特征逐元素相乘后的特征,  $\mathbf{Y}$  表示最终输出特征,  $\text{DWConv}(\cdot)$  表示深度可分离卷积,  $\text{BN}(\cdot)$  表示批归一化,  $\text{FC}(\cdot)$  表示通道全连接操作,  $\text{ReLU6}(\cdot)$  为激活函数,  $\odot$  为逐元素乘法。

## 1.3 C3k2-MSEIE 模块

为了进一步提升钢材缺陷检测在多尺度目标与细微边缘结构上的表征能力,本文对原始 C3K2 模块进行了重构,将其中的骨干卷积分支替换为 MSEIE。如图 3 所示, MSEIE 首先利用多尺度池化与轻量卷积建立不同尺度的上下文表征,以适应钢材缺陷在尺寸、形态和纹理复杂度上的显著差异;随后在各尺度分支输出上引入边缘增强单元,使网络能够显式强调强度突变区域,从而提升对弱纹理、模糊边界缺陷的感知能力。

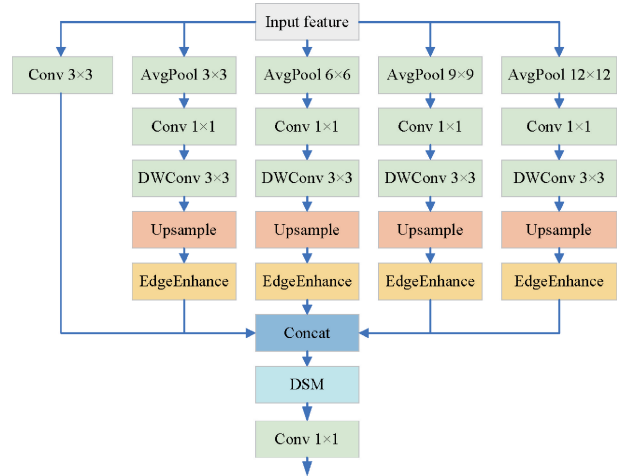


图 3 C3k2-MSEIE 模块结构

Fig. 3 Structure of the C3k2-MSEIE module

在多尺度特征聚合后,模块进一步通过  $\text{DSM}^{[19]}$  进行特征筛选。 $\text{DSM}$  由空间域与通道域两个并行路径组成,前者突出显著位置区域,后者强调语义相关通道,从而抑制冗余背景并强化缺陷特征的辨识度。通过“多尺度特征捕获—边缘强化—双域选择—跨尺度融合”这一协同流程, MSEIE 能够显著增强改进后 C3k2 的细节敏感性与跨尺度建模能力,为后续检测头提供更具判别性的表征。

MSEIE 由 4 个主要部分组成:多尺度特征分支:通过多种感受野的池化操作( $3 \times 3, 6 \times 6, 9 \times 9, 12 \times 12$ )获取不同尺度的上下文信息;边缘增强模块:利用特征差分与卷积操作强化边缘信息; $\text{DSM}$ :综合空间域与通道域的特征选择能力;特征融合:通过特征拼接与  $1 \times 1$  卷积进行融合降维。输入特征  $\mathbf{X} \in \mathbb{R}^{C \times H \times W}$  经多尺度分支处理可表示为:

$$F_i = U_i(\text{DWConv}_{3 \times 3}(\text{Conv}_{1 \times 1}(\text{Pool}_{s_i}(\mathbf{X}))))$$

$$s_i \in \{3, 6, 9, 12\} \quad (6)$$

其中,  $\text{Pool}_{s_i}$  表示尺度为  $s_i$  的平均池化操作,  $\text{Conv}_{1 \times 1}$  与  $\text{DWConv}_{3 \times 3}$  分别为逐点卷积与深度可分离卷积,  $U_i(\cdot)$  表示上采样操作。

所有分支特征经边缘增强模块处理后与原始卷积分支特征  $F_0$  融合:

$$\mathbf{F}_{\text{concat}} = \text{Concat}(\mathbf{F}_0, E(\mathbf{F}_3), E(\mathbf{F}_6), E(\mathbf{F}_9), E(\mathbf{F}_{12})) \quad (7)$$

其中,  $E(\cdot)$  表示 Edge Enhance 操作。

边缘增强模块(edge enhance, EE)结构如图 4 所示,旨在通过局部差分操作突出缺陷的边界信息。输入特征  $\mathbf{X}$  首先进行平均池化得到低频特征  $\mathbf{X}_{\text{avg}}$ , 再与原特征相减以获取高频残差  $\mathbf{R}$ :

$$\mathbf{R} = \mathbf{X} - \text{AvgPool}_{3 \times 3}(\mathbf{X}) \quad (8)$$

随后使用  $3 \times 3$  卷积提取边缘模式并与原特征相加形成增强输出:

$$\mathbf{F}_{\text{edge}} = \text{Conv}_{3 \times 3}(\mathbf{R}) + \mathbf{X} \quad (9)$$

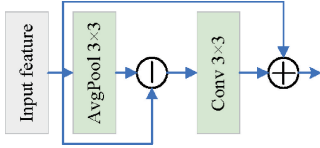


图 4 边缘增强模块结构

Fig. 4 Architecture of the edge enhancement module

DSM 结构如图 5 所示,由空间选择模块(spatial selection module, SSM)与特征选择模块(feature selection module, FSM)组成。SSM 通过多尺度卷积( $3 \times 3, 5 \times 5, 7 \times 7$ )及池化操作提取不同尺度的空间特征,并采用逐点乘法与逐元素加法实现特征融合:

$$\mathbf{T}_1 = (\text{DWConv}_{7 \times 7}(\text{DWConv}_{5 \times 5}(\mathbf{X}))) \odot$$

$$(\text{Conv}_{3 \times 3}(\text{MaxPool}(\mathbf{X}) + \text{AvgPool}(\mathbf{X}))) \quad (10)$$

$$\mathbf{F}_{\text{ssm}} = \mathbf{T}_1 + \text{DWConv}_{3 \times 3}(\mathbf{X}) \quad (11)$$

FSM 则在通道维度进行均值池化并广播操作,对各通道进行重要性建模。首先将  $\mathbf{F}_{\text{ssm}}$  与其均值差进行逐元素相减,再与原特征逐元素相乘,最后通过残差相加实现融合:

$$\mathbf{D} = \mathbf{F}_{\text{ssm}} - \text{Broadcast}(\text{Mean}(\mathbf{F}_{\text{ssm}})) \quad (12)$$

最终输出为:

$$\mathbf{F}_{\text{out}} = \mathbf{D} \odot \mathbf{F}_{\text{ssm}} + \mathbf{F}_{\text{ssm}} \quad (13)$$

#### 1.4 C2PSA-SEFN 模块

在检测头阶段,为进一步提升特征表达能力与跨通道信息的交互效率,本文对原始 C2PSA 模块中的前馈网络(feed forward network, FFN)进行了功能增强,将其替换为受 SEM-Net<sup>[20]</sup> 启发的空间增强前馈网络 SEFN。与传统 FFN 仅依赖逐点卷积与非线性映射不同,SEFN 在设计上引入了空间分支与主干分支的双路径结构:主干分支负责局部语义的线性变换与通道重排,而空间分支通过降采样、

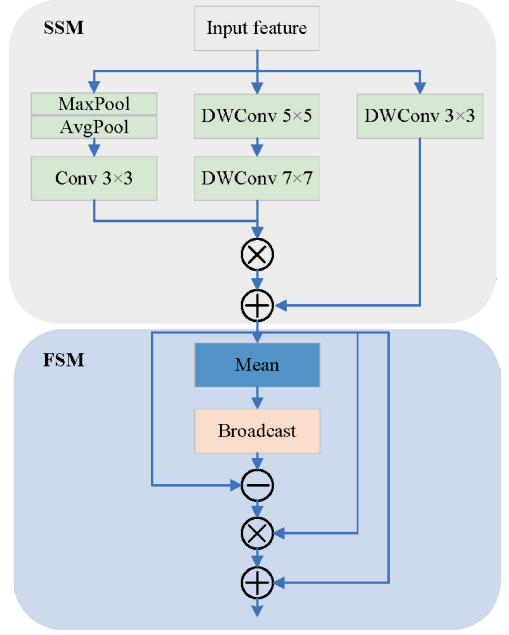


图 5 DSM 模块结构

Fig. 5 Architecture of the DSM module

局部卷积与上采样操作提取跨区域的空间依赖关系,从而弥补传统 FFN 在空间维度缺乏显式建模的局限。

改进后的 PSA 模块结构如图 6 所示,其整体处理流程包括特征压缩、注意力增强以及基于 SEFN 的特征重构 3 个阶段。通过在 FFN 位置引入空间信息建模机制,使 PSA 模块能够在保持轻量化特性的同时提升对细粒度缺陷区域的响应能力,并强化跨尺度特征的融合效果,有助于进一步提升检测头在复杂钢材表面缺陷场景中的识别能力。

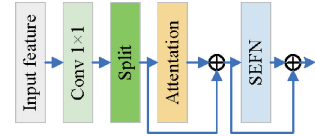


图 6 改进后的 PSA 模块结构

Fig. 6 Improved PSA module architecture

输入特征  $\mathbf{X} \in \mathbb{R}^{C \times H \times W}$  首先经过  $1 \times 1$  卷积进行通道扩增,得到高维嵌入表示  $\mathbf{X}'$ , 以方便后续拆分:

$$\mathbf{X}' = \text{Conv}_{1 \times 1}(\mathbf{X}) \quad (14)$$

特征  $\mathbf{X}'$  被分为两支,一支进入注意力分支,通过构建跨通道依赖关系生成加权重数  $A$ ,另一支则直接保留原始特征以增强残差信息:

$$\mathbf{Y}_1 = A \oplus \mathbf{X}' \quad (15)$$

在 FFN 阶段,引入了 SEFN 模块,如图 7 所示,结合了深度可分离卷积、多尺度空间上下采样和 GELU,以更高效地融合局部与全局特征。

主分支首先执行  $1 \times 1$  卷积与  $3 \times 3$  深度可分离卷积:

$$\mathbf{F}_1 = \text{DWConv}_{3 \times 3}(\text{Conv}_{1 \times 1}(\mathbf{Y}_1)) \quad (16)$$

辅助分支通过  $2 \times 2$  平均池化提取低分辨率全局特征,

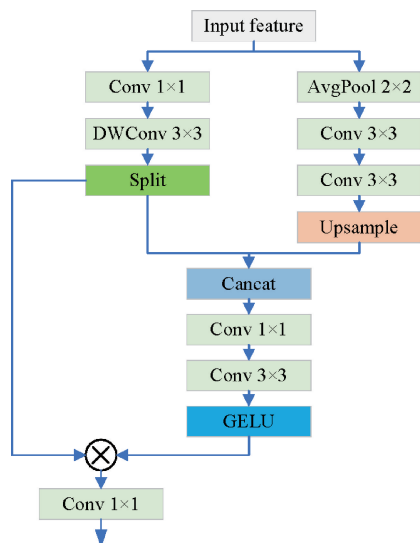


图 7 SEFN 模块结构

Fig. 7 Structure of the SEFN module

再经过两层  $3 \times 3$  卷积后上采样回原尺度:

$$\mathbf{F}_2 = \text{Up}(\text{Conv}_{3 \times 3}(\text{Conv}_{3 \times 3}(\text{AvgPool}_{2 \times 2}(\mathbf{Y}_1)))) \quad (17)$$

将两分支特征拼接并融合,经过  $1 \times 1$  卷积、 $3 \times 3$  卷积及 GELU 激活得到融合特征  $\mathbf{F}$ :

$$\mathbf{F} = \text{GELU}(\text{Conv}_{3 \times 3}(\text{Conv}_{1 \times 1}(\text{Concat}[\mathbf{F}_1, \mathbf{F}_2]))) \quad (18)$$

最后,  $\mathbf{F}$  与主分支输入进行逐元素相乘,实现通道间的特征再校准:

$$\mathbf{Z} = \mathbf{F} \odot \mathbf{F}_1 \quad (19)$$

通过引入 SEFN,改进后的 PSA 模块能够在保持较低计算复杂度的同时,显著提升通道间的特征选择性与跨空间的上下文融合能力,为后续检测任务提供更加判别性的多尺度特征表示。

## 2 实验结果与分析

### 2.1 数据集

本文实验基于公开的 NEU-DET 数据集,以验证所提方法在钢材表面缺陷检测任务中的有效性。该数据集由东北大学构建,专门面向热轧钢带表面质量检测场景,涵盖工业生产中 6 类典型缺陷:裂纹(crazing, Cr)、氧化皮(rolled-in scale, Rs)、麻点(pitted surface, Ps)、划痕(scratches, Sc)、夹杂(inclusion, In)以及斑块(patches, Pa)。每类样本数量均为 300 张,整体共计 1 800 张图像,保证了数据的均衡性与代表性。所有图像的分辨率统一为  $200 \text{ pixel} \times 200 \text{ pixel}$ ,便于模型在相同尺度下进行训练与评估。在数据集划分方面,本文遵循 8:1:1 的比例,将数据随机划分为训练集、验证集和测试集,并保持各类别在 3 个子集中的分布一致,以避免因类别不平衡导致的评估偏差。

为了进一步验证模型在更复杂工业场景下的泛化能力,本文引入 GC10-DET 数据集作为补充测试。该数据集

专用于钢铁表面质量检测任务,包含 10 类典型缺陷:冲孔(Pu)、焊缝(WI)、新月形缝隙(Cg)、水斑(water spot)、油斑(Os)、丝斑(Ss)、夹杂物(In)、轧坑(Rp)、折痕(Cr)以及腰部折痕(Wf)。GC10-DET 数据集规模大于 NEU-DET,总计收录 2 312 张缺陷图像,图像分辨率较高,来源于不同生产批次与光照环境,具有良好的样本多样性与代表性。为保证训练规范一致性,本文按照 8:1:1 的比例将数据集划分为训练集、验证集与测试集,并确保各类别样本在不同子集中分布均衡。

### 2.2 实验环境

本实验在 Windows 10 操作系统环境下完成,硬件平台配备 NVIDIA RTX 5070 GPU,深度学习框架为 PyTorch 2.7.1 并结合 CUDA 12.8 提供加速支持。所使用的检测模型为基于 Ultralytics YOLO 框架的自定义结构。在训练参数方面,模型输入尺寸固定为  $640 \times 640$ ,优化器选用 SGD,初始学习率设为 0.01,批量大小为 32,总训练轮次为 200。训练过程中启用了 AMP 混合精度机制以提升计算效率和显存利用率。在数据增强策略上,除采用 Rand Augment 自动增强机制外,还引入 Mosaic 拼接增强、水平翻转与颜色扰动等方式,以提升模型的泛化性能和对复杂场景的鲁棒性。

### 2.3 评价指标

为全面评估所提方法的性能与效率,本文从精度指标与复杂度指标两个方面进行考量,具体包括:

交并比(intersection over union, IoU):在目标检测任务中,该指标用于衡量预测框与真实框的重叠程度:

$$\text{IoU} = \frac{|B_p \cap B_{gt}|}{|B_p \cup B_{gt}|} \quad (20)$$

其中,  $B_p$  表示预测框,  $B_{gt}$  表示真实框。以不同的置信度阈值绘制精确率-召回率曲线(PR 曲线),其下的面积即为平均精度(average precision, AP):

$$\text{AP} = \int_0^1 P(R) dR \quad (21)$$

平均精度均值(mean average precision, mAP):针对多类别任务,采用各类别 AP 的平均值作为综合评价指标:

$$\text{mAP} = \frac{1}{N} \sum_{i=1}^N \text{AP}_i \quad (22)$$

其中,  $N$  为类别数,  $\text{AP}_i$  为第  $i$  类的平均精度。本文重点使用 mAP50 作为核心指标,即 IoU 阈值设为 0.5 时的平均精度。

在复杂度指标方面,参数量(parameters)用于衡量模型的存储规模,而 GFLOPs(giga floating point operations)则反映模型在一次前向推理过程中所需的计算量。二者能够从存储与运算两个维度评价模型的实际开销。

### 2.4 消融实验

为验证所提模块的有效性,本文在 YOLOv11n 基础模型上逐步引入 StarNet、C3k2-MSEIE 与 C2PSA-SEFN

模块开展消融实验。结果如表 1 所示,基准 YOLOv11n 的 mAP 为 72.57%,参数量与计算量分别为 6.3 M 和 2.58 GFLOPs;当引入 StarNet 后,mAP 提升至 72.76%,同时参数量与计算量显著下降至 5.0 M 和 1.94 GFLOPs,体现了该主干网络的轻量化优势。在此基础上加入 C3k2-MSEIE 模块后,mAP 进一步提升至 73.58%,验证了其在增强特征表达能力方面的作用;继续引入 C2PSA-SEFN

模块后,mAP 达到 74.58%,较基准模型整体提升 2.01%,并在保持参数量与计算量低于原始模型的前提下取得最优性能。综上,各改进模块均对性能提升起到积极作用,其中 StarNet 贡献于复杂度下降,C3k2-MSEIE 增强了特征建模能力,而 C2PSA-SEFN 则在特征融合与注意力机制方面发挥关键作用,三者结合实现了精度与效率的兼顾。

表 1 消融实验结果对比  
Table 1 Ablation study results

| 组别 | StarNet | C3k2-MSEIE | C2PSA-SEFN | mAP/% | GFLOPs | Param/M |
|----|---------|------------|------------|-------|--------|---------|
| 1  |         |            |            | 72.57 | 2.58   | 6.3     |
| 2  | ✓       |            |            | 72.76 | 1.94   | 5.0     |
| 3  | ✓       | ✓          |            | 73.58 | 1.98   | 5.2     |
| 4  | ✓       | ✓          | ✓          | 74.58 | 2.32   | 5.4     |

此外,为分析各模块在不同缺陷类型上的具体贡献,本文对 NEU-DET6 类缺陷分别计算了 AP 变化情况。结果表明:StarNet 对 Pa 与 Ps 等纹理相对规则的缺陷提升更明显;C3k2-MSEIE 对 Cr 与 Rs 等边缘模糊、尺度变化显著的缺陷改善突出;SEFN 在 In 等纹理复杂类上带来更稳定的增益,说明 3 模块在结构特征、边缘特征与空间语义增强方面具有互补性。为验证各模块带来的性能提升是否具有统计意义,本文对 4 组模型在相同训练配置下进行了多次独立训练,并对 mAP50 指标采用配对 t 检验进行显著性分析。结果显示,在引入 C3k2-MSEIE 与 C2PSA-SEFN 后,模型相较于基准 YOLOv11n 的性能提升均达到

统计显著水平( $p<0.05$ )。该结果表明,性能增益并非源于随机波动,而是由所提出模块的稳定改进效果所驱动。

如图 8 所示,NEU-DET 数据集中 6 类典型缺陷的检测可视化结果。第 1 行为原始图像,第 2 行为 YOLOv11n 的检测结果,第 3 行为本文提出方法的检测结果。可以观察到,YOLOv11n 在第 1、2、5 列图像中出现漏检现象,未能完整识别出目标缺陷;而在第 3、4、6 列图像中则出现过检问题,将部分非缺陷区域错误识别为缺陷。相比之下,本文方法能够有效克服上述问题,显著降低了漏检率,还减少了虚警现象,整体检测结果更加精确与稳定,体现了所提出改进策略在钢材表面缺陷检测任务中的有效性。

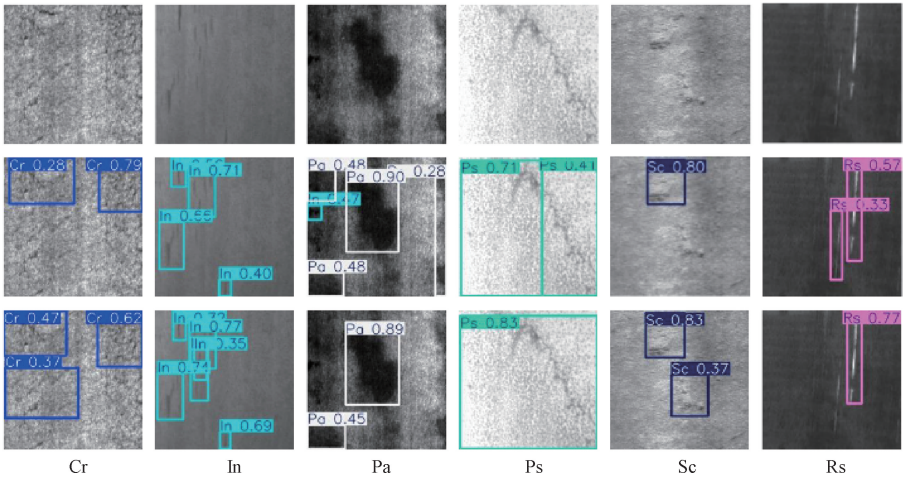


图 8 不同方法在 6 类缺陷上的检测可视化对比  
Fig. 8 Visualization comparison of six defect types using different methods

如图 9 所示,训练过程中的收敛情况,其中图 9(a)为训练集损失曲线对比,图 9(b)为验证集损失曲线对比,可以看出所提方法在训练与验证阶段均收敛更快且整体损失更低。

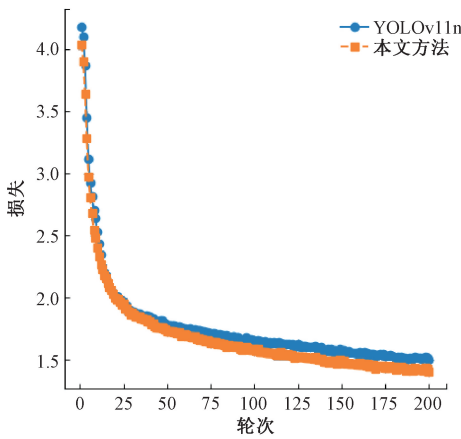
如图 10 所示,模型在验证集上的精度表现,其中图 10(a)为 mAP50 的收敛对比,图 10(b)为 PR 曲线对比,结

果表明所提方法在不同召回率下均能保持更高的精确率。

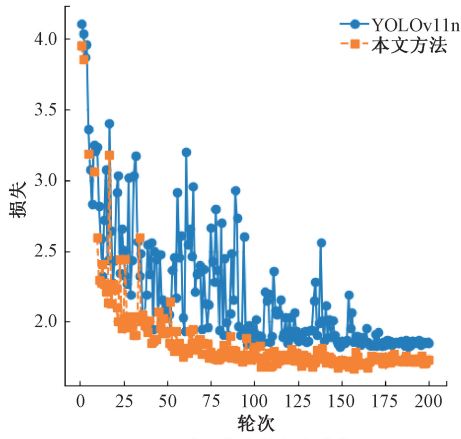
2.5 对比实验

1)NEU-DET 数据集

在对比实验中,本研究将所提出的方法与主流检测器 SSD、Faster R-CNN 以及不同版本的 YOLO 系列模型进行了系统评估。精度对比如表 2 所示。



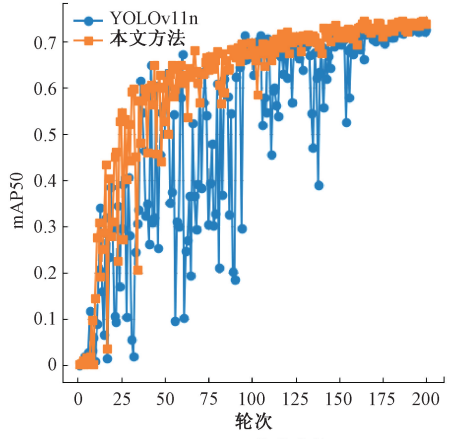
(a) 训练集上的损失曲线  
(a) Loss curve on the training set



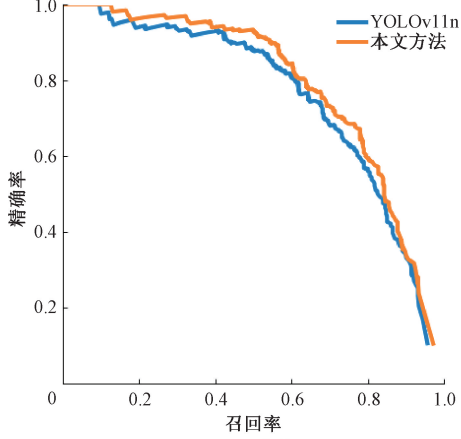
(b) 验证集上的损失曲线  
(b) Loss curve on the validation set

图 9 训练与验证损失曲线对比

Fig. 9 Training and validation loss curves



(a) mAP50收敛曲线  
(a) Convergence curve of mAP50



(b) PR曲线  
(b) PR curve

图 10 验证集精度对比

Fig. 10 Validation performance comparison

表 2 不同检测模型在钢材缺陷数据集上的分类精度对比

Table 2 Classification accuracy comparison of different detection models on the steel defect dataset

%

| 方法         | AP    |       |       |       |       |       | mAP   |
|------------|-------|-------|-------|-------|-------|-------|-------|
|            | Cr    | In    | Pa    | Ps    | Rs    | Sc    |       |
| SSD        | 27.63 | 70.63 | 83.17 | 79.64 | 53.16 | 81.87 | 66.02 |
| FasterRCNN | 36.68 | 75.73 | 91.69 | 83.59 | 60.49 | 80.78 | 71.49 |
| YOLOv5n    | 27.39 | 72.84 | 91.26 | 81.29 | 63.10 | 87.65 | 70.59 |
| YOLOv8n    | 36.09 | 71.59 | 90.46 | 82.04 | 62.04 | 88.03 | 71.71 |
| YOLOv11n   | 34.26 | 73.92 | 87.17 | 83.31 | 64.84 | 91.91 | 72.57 |
| YOLOv11s   | 35.79 | 74.39 | 91.60 | 83.21 | 62.41 | 94.14 | 73.59 |
| 本文方法       | 38.67 | 76.27 | 91.92 | 83.87 | 64.77 | 91.99 | 74.58 |

传统 SSD 在大多数类别上的检测精度明显不足，mAP50 仅为 66.02%，难以满足复杂场景需求。Faster R-CNN 在整体精度上有所提升，mAP50 达到 71.49%，但其

在类别 Cr 上的表现依然受限。YOLO 系列模型则在速度与精度之间取得了较好平衡，其中 YOLOv11n 的 mAP50 为 72.57%，YOLOv11s 则进一步提升至 73.59%。相比之

下,本文方法在各类别上均表现出较强的检测能力,在 Cr 与 In 等难点类别上显著优于其他模型,最终实现了 74.58%的 mAP50,取得了最优结果。

如图 11 所示,不同检测模型在数据集上的性能对比,包括精确率、召回率以及 mAP50。可以看出,所提方法在 3 项指标上均取得改进,在精确率和 mAP50 上较其他方法有明显提升,体现了其在保证检测精度的同时具备更强的鲁棒性。

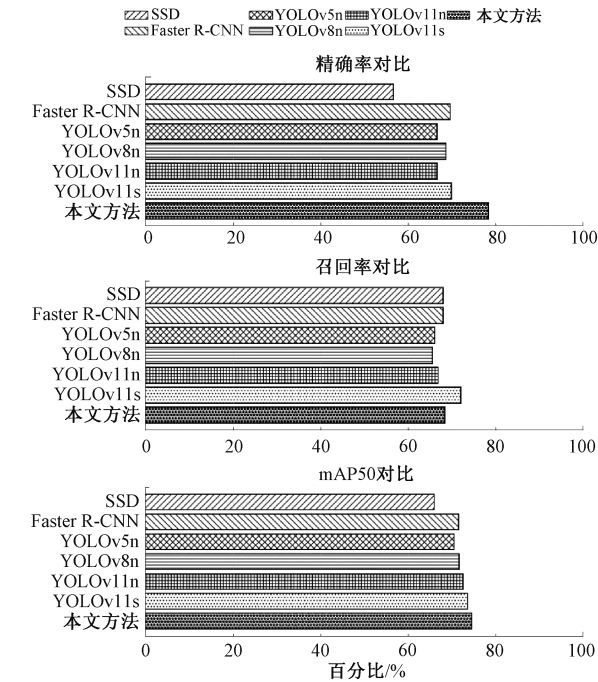


图 11 不同方法在精确率、召回率、mAP50 上的对比  
Fig. 11 Comparison of different methods in precision, recall, and mAP50

复杂度分析如表 3 所示,进一步验证了所提方法的有效性。SSD 与 Faster R-CNN 尽管在部分类别上具有一定

表 3 各检测模型的复杂度对比

| 方法          | mAP/% | GFLOPs | Parameters/M |
|-------------|-------|--------|--------------|
| SSD         | 66.02 | 30.68  | 24.4         |
| Faster RCNN | 71.49 | 37.54  | 28.3         |
| YOLOv5n     | 70.59 | 2.23   | 5.9          |
| YOLOv8n     | 71.71 | 3.18   | 8.9          |
| YOLOv11n    | 72.57 | 2.58   | 6.3          |
| YOLOv11s    | 73.59 | 9.42   | 21.3         |
| 本文方法        | 74.58 | 2.32   | 5.4          |

精度,但其计算量与参数规模过大,限制了其在轻量化场景中的应用。YOLO 系列轻量化模型在参数与计算量上优势明显,但在精度上仍存在一定差距。相比之下,本文方法在仅需 2.32 GFLOPs 与 5.4 M 参数的条件下,即实现了 74.58%的 mAP50,较 YOLOv11n 提升 2.01%,且计算开销与参数量均更低,充分体现了高效性与实用性。

2)GC10-DET 数据集

为进一步验证模型的泛化性能,本文在钢材表面缺陷大规模数据集 GC10-DET 上进行了补充实验。该数据集包含工业生产中 10 类典型表面缺陷,纹理差异大、成像噪声强、目标尺度跨度明显,对检测模型的特征表达能力与稳定性提出了更高要求。

性能对比结果如表 4 所示。经典方法 SSD 与 Faster R-CNN 受限于浅层特征表达或推理速度,mAP50 分别为 64.99%和 68.27%,难以兼顾精度与效率。轻量化 YOLO 系列在实时性方面表现突出,其中 YOLOv11n 取得 69.85%的 mAP50 与 213 fps,但在复杂缺陷边界上仍存在一定漏检情况;YOLOv11s 虽具有更高精度,但因模型规模较大,其推理速度明显下降。相比之下,本文提出的 MEA-YOLO 在 GC10-DET 上实现了 71.28%的 mAP50 与 237 fps 的推理速度,在所有对比模型中同时获得更高精度与更优实时性。其性能提升主要得益于 StarNet 主干降低冗余计算量,多尺度边缘增强模块 MSEIE 强化细粒度边界特征,以及 SEFN 提升跨尺度特征融合能力。

表 4 不同模型在 GC10-DET 数据集上的检测结果

Table 4 Detection performance of different models on the GC10-DET dataset

| 方法          | mAP/% | FPS |
|-------------|-------|-----|
| SSD         | 64.99 | 21  |
| Faster RCNN | 68.27 | 16  |
| YOLOv5n     | 69.36 | 231 |
| YOLOv8n     | 67.83 | 173 |
| YOLOv11n    | 69.85 | 213 |
| YOLOv11s    | 68.32 | 71  |
| 本文方法        | 71.28 | 237 |

可视化结果如图 12 所示,YOLOv11n 在多类缺陷上出现重复框及漏检等问题,反映其对弱边缘与低对比度缺陷的判别不足。相比之下,MEA-YOLO 有效消除了重框并减少漏检,边界定位更准确,充分体现了多尺度边缘增强与注意力融合机制带来的识别稳定性提升。

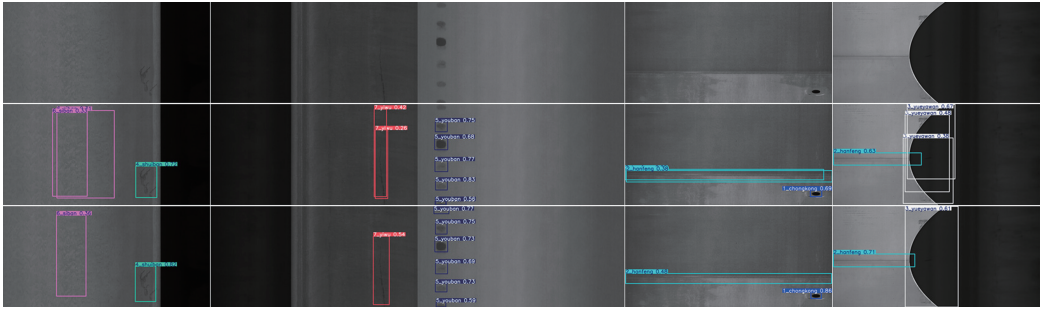


图 12 不同方法在 GC10-DET 数据集上的检测可视化对比

Fig. 12 Visualization comparison on the GC10-DET dataset using different methods

3 结 论

本研究围绕钢材表面缺陷的自动化检测,提出了一种基于多尺度边缘增强与注意力融合的轻量化检测网络(MEA-YOLO)。实验结果表明,该方法在保持较低参数量和计算复杂度的同时,能够有效提升对复杂纹理和多尺度缺陷的检测精度,mAP50 较 YOLOv11n 提高 2.01%,参数量与 GFLOPs 分别下降 14.3%和 10.1%。研究结果表明,多尺度边缘特征与注意力机制的融合在缺陷检测中具有普遍适用性和实际价值。与已有工作相比,本方法兼顾精度与效率,更契合实时工业检测需求。然而,对于极小缺陷样本不足及复杂背景下的误检问题,仍有进一步改进空间。后续研究可在数据多样性扩展、模型结构优化及少样本检测等方面展开,以进一步提升方法的鲁棒性与通用性。

参考文献

[1] MORDIA R, VERMA A K. Visual techniques for defects detection in steel products: A comparative study [J]. Engineering Failure Analysis, 2022, 134: 106047.

[2] ZHANG J W, WANG H Y, TIAN Y, et al. An accurate fuzzy measure-based detection method for various types of defects on strip steel surfaces [J]. Computers in Industry, 2020, 122: 103231.

[3] REN SH Q, HE K M, GIRSHICK R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2017, 39 (6): 1137-1149.

[4] 冷岳峰,刘正,徐宝祯,等. 改进 Faster R-CNN 的钢材表面缺陷检测[J]. 机械科学与技术, 2025, 44(1): 75-83.

LENG Y F, LIU ZH, XU B Y, et al. Detection of steel surface defect based on improved Faster R-CNN [J]. Mechanical Science and Technology for Aerospace Engineering, 2025, 44(1): 75-83.

[5] 王伟家,张宇,王京华,等. 基于改进 RetinaNet 的轻量化钢材表面缺陷检测算法[J]. 模式识别与人工智能, 2024, 37(8): 692-702.

WANG W J, ZHANG Y, WANG J H, et al. Lightweight steel surface defect detection algorithm based on improved RetinaNet[J]. Pattern Recognition and Artificial Intelligence, 2024, 37(8): 692-702.

[6] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: Unified, real-time object detection[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2016: 779-788.

[7] 乔庆元,骆晓玲,程焕新,等. 基于 YOLO-RSI 算法的钢材表面缺陷检测[J]. 组合机床与自动化加工技术, 2025(5): 145-149,155.

QIAO Q Y, LUO X L, CHENG H X, et al. Steel surface defect detection based on the YOLO-RSI algorithm [J]. Combined Machine Tools and Automated Processing Technology, 2025 (5): 145-149,155.

[8] 刘翰林,刘凌云,李超凡,等. 基于 YOLOv5s-CBC 的钢材表面缺陷检测[J]. 组合机床与自动化加工技术, 2024(10): 155-159.

LIU H L, LIU L Y, LI CH F, et al. Steel surface defect detection based on YOLOv5s-CBC[J]. Modular Machine Tool & Automatic Manufacturing Technique, 2024(10): 155-159.

[9] 梁礼明,龙鹏威,李俞霖. 改进轻量高效 FMG-YOLOv8s 的钢材表面缺陷检测算法[J]. 计算机工程与应用, 2025, 61(3): 84-93.

LIANG L M, LONG P W, LI Y L. Improved lightweight and efficient FMG-YOLOv8s for steel surface defect detection[J]. Computer Engineering and Applications, 2025, 61(3): 84-93.

[10] TIE J, ZHU CH G, ZHENG L, et al. LSKA-YOLOv8: A lightweight steel surface defect detection algorithm based on YOLOv8 improvement [J]. Alexandria Engineering Journal, 2024, 109: 201-212.

- [11] 梁礼明, 陈康泉, 钟奕, 等. DCD-YOLOv8n: 一种高效的钢材表面缺陷检测算法[J]. 计算机工程与应用, 2025, 61(7): 117-127.  
LIANG L M, CHEN K Q, ZHONG Y, et al. DCD-YOLOv8n: Efficient algorithm for steel surface defect detection[J]. Computer Engineering and Applications, 2025, 61(7): 117-127.
- [12] LIU L J, ZHANG Y, KARIMI H R. Resilient machine learning for steel surface defect detection based on lightweight convolution [J]. The International Journal of Advanced Manufacturing Technology, 2024, 134: 4639-4650.
- [13] 张伟健, 徐向英, 章永龙. 改进 RT-DETR 的钢材表面缺陷检测算法[J/OL]. 机械科学与技术, 1-9[2025-08-19]. <https://doi.org/10.13433/j.cnki.1003-8728.20250009>.  
ZHANG W J, XU X Y, ZHANG Y L. Steel surface defect detection algorithm based on improved RT-DETR[J/OL]. Mechanical Science and Technology for Aerospace Engineering, 1-9[2025-08-19]. <https://doi.org/10.13433/j.cnki.1003-8728.20250009>.
- [14] 张航, 周毅, 邱宇峰. 融合 HGnetv2 和注意力机制的钢材表面缺陷检测方法[J]. 电子测量与仪器学报, 2025, 39(1): 36-49.  
ZHANG H, ZHOU Y, QIU Y F. Detection method of steel surface defects with fusion of HGnetv2 and attention mechanism [J]. Journal of Electronic Measurement and Instrumentation, 2025, 39(1): 36-49.
- [15] HUANG X H, ZHU J H, HUO Y. SSA-YOLO: An improved YOLO for hot-rolled strip steel surface defect detection [J]. IEEE Transactions on Instrumentation and Measurement, 2024, 73: 1-17.
- [16] 冯夫健, 罗太维, 谭棉, 等. 基于自注意特征融合的钢材表面小目标缺陷检测[J]. 电子测量技术, 2024, 47(19): 172-180.  
FENG F J, LUO T W, TAN M, et al. Defect detection of small targets on steel surface based on self-attention feature fusion[J]. Electronic Measurement Technology, 2024, 47(19): 172-180.
- [17] 黄志海, 邓耀华, 吴光栋. IC 器件表面缺陷多光谱图像特征融合检测方法[J]. 仪器仪表学报, 2024, 45(9): 24-33.  
HUANG ZH H, DENG Y H, WU G D. Multispectral image feature fusion method for detecting surface defects in IC devices [J]. Chinese Journal of Scientific Instrument, 2024, 45(9): 24-33.
- [18] MA X, DAI X Y, BAI Y, et al. Rewrite the stars[J]. ArXiv preprint arXiv:2403.19967, 2024.
- [19] CUI Y N, REN W Q, CAO X CH, et al. Focal network for image restoration [C]. IEEE/CVF International Conference on Computer Vision, 2023.
- [20] CHEN SH, ZHANG H ZH, ATAPOUR-ABARGHOUEI A, et al. SEM-Net: Efficient pixel modelling for image inpainting with spatially enhanced SSM[C]. 2025 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV), 2025: 461-471.

## 作者简介

曹万里, 硕士研究生, 主要研究方向为机器视觉和深度学习。

E-mail: kehanliedu@163.com

刘强(通信作者), 博士, 副教授, 硕士生导师, 主要研究方向为复合型机器人、深度学习和人机协作机器人。

E-mail: lqiang@sdu.edu.cn