

DOI:10.19651/j.cnki.emt.2519713

# LiteSteel-YOLO:小目标低对比度轻量级 钢材缺陷检测网络\*

周迅 李帆 张岩

(青岛科技大学数理学院 青岛 266061)

**摘要:** 钢材缺陷检测对工业质量管控至关重要,但其多尺度、小目标及背景干扰问题制约检测性能。为了提升模型的检测精度与效率,提出一种基于YOLO11改进的缺陷检测网络 LiteSteel-YOLO。首先,设计轻量化多尺度融合结构(C3k2-LMSF),通过多尺度卷积核融合与特征引导机制增强多尺度缺陷感知;其次,提出空间通道感知上采样模块(SCAM),基于通道重组与空间偏移运算机制提升小目标检测的鲁棒性并抑制噪声;最后,构建轻量化高效检测头(Efficient-Head)优化计算效率。实验结果表明,所提出的模型在 NEU-DET 和 GC10-DET 数据集上的 mAP@50 分别达到 81.7% 和 70.7%,相较于原模型 YOLO11 提升了 4.0% 和 2.3%,检测速度为 338 和 530 FPS,有效地提升了对钢材缺陷的检测精度与效率,为工业检测场景提供了解决思路。

**关键词:** 工业缺陷检测;小目标检测;轻量化网络;YOLO11

**中图分类号:** TP391.9;TN911 **文献标识码:** A **国家标准学科分类代码:** 520.6040

## LiteSteel-YOLO: Small target low-contrast lightweight steel defect detection network

Zhou Xun Li Fan Zhang Yan

(School of Mathematics and Physics, Qingdao University of Science and Technology, Qingdao 266061, China)

**Abstract:** Steel defect detection is critical for industrial quality control, yet performance is constrained by multi-scale variations, small targets, and background interference. To enhance the accuracy and efficiency of the detection model, this paper proposes a defect detection network based on an improved version of YOLO11, named LiteSteel-YOLO. First, a Lightweight Multi-Scale Fusion module (C3k2-LMSF) is designed to enhance multi-scale defect perception through fused convolutional kernels and feature guidance mechanisms. Second, a spatial-channel aware upsampling module (SCAM) is proposed, which improves the robustness of small target detection and suppresses noise through channel reorganization and spatial offset operations. Finally, an Efficient-Head detector optimized via structural reconfiguration is introduced to maximize computational efficiency. Experimental results show that the LiteSteel-YOLO receives mAP@50 of 81.7% and 70.7% with inference speed of 338 and 530 FPS on the NEU-DET and GC10-DET datasets (surpassing YOLO11 by 4.0% and 2.3%). The proposed framework enhances the accuracy and efficiency of steel defect detection, providing a solution for industrial inspection scenarios.

**Keywords:** industrial defect detection; small target detection; lightweight network; YOLO11

## 0 引言

钢材缺陷检测在工业质量管控中具有关键意义,直接影响其产品安全性与可靠性。近年来,深度卷积神经网络(convolutional neural networks, CNN)由于其强大的特征提取能力,逐步成为缺陷检测的主流方法。根据算法是否

先提取候选区域,可分为两阶段目标检测算法和单阶段目标检测算法。两阶段目标检测算法主要有 Faster R-CNN<sup>[1]</sup>和 Sparse-RCNN<sup>[2]</sup>等,陆家林等<sup>[3]</sup>在其基础上进行设计了一种基于机器视觉的钢材表面缺陷检测方法,实现了平均检测精度和检测速度的提升。然而,目前单阶段检测器<sup>[4-7]</sup>因速度与精度平衡性优于两阶段方法<sup>[1,8]</sup>而被广泛

收稿日期:2025-08-27

\* 基金项目:山东省自然科学基金(ZR2022MF250)项目资助

采用。其中, RDD-YOLO<sup>[9]</sup> 借助 Res2Net 重新设计主干网络以扩大感受野, 并设计双特征金字塔网络(double feature pyramid network, DFPN)增强颈部的特征表示, 采用解耦头将回归任务与分类任务分离以提高检测精度。在两个典型钢材缺陷数据集的实验表明, RDD-YOLO 取得了良好的综合性能。Zhang 等<sup>[10]</sup> 将拉普拉斯锐化和 YOLOv5s 结合以提取多尺度缺陷特征, 并提出一种嵌入残差模块的改进加权双向特征金字塔网络用于聚合多尺度特征图, 以增强多尺寸缺陷表示的鲁棒性。Yu 等<sup>[11]</sup> 借助卷积块注意力模块(convolution block attention module, CBAM)增强目标特征感知。You 等<sup>[12]</sup> 引入了一种无注意力机制(attention-free mechanism, AM)用于特征权重计算以增强目标区域特征提取能力, 并对 SPPF 模块进行改进, 以扩大感受野并增强目标检测优化。Li 等<sup>[13]</sup> 提出一种 YOLOv8-GhostConv-SEV2 模型, 通过引入 GhostConv 模块优化特征提取, 并利用 SEV2(squeeze-and-excitation version 2, SEV2)注意力机制增强噪声抑制能力。Wang 等<sup>[14]</sup> 通过在头部模块中结合自注意力和图形卷积来提高特征识别能力, 并使用超薄模块结合深度可分离卷积对整个网络架构进行精简, 从而在精度与基线模型相当的同时参数量显著降低。雷超等<sup>[15]</sup> 针对钢板表面缺陷检测精度不足导致的误检频发和漏检严重等问题, 提出了一种基于 YOLOv8n 的改进模型 SGF-YOLOv8n。吴葛等<sup>[16]</sup> 将 Neck 部分换为注意力尺度序列融合框架 ASF, 以提升模型整合多尺度特征信息的能力。胥姜苗等<sup>[17]</sup> 则设计多级特征融合金字塔帮助模型实现跨尺度信息交互能力。徐森等<sup>[18]</sup> 则在网络结构中增加小目标检测头, 利用上下文关系, 提高模型的小目标检测检测能力。Wang 等<sup>[19]</sup> 通过利用自适应多尺度池化方法改进网络对边缘特征的提取能力, 同时引入可变形空间注意力模块增强网络对复杂背景下特征的提取能力。Dang 等<sup>[20]</sup> 为解决复杂缺陷环境中缺陷检测的精度与效率, 设计 FSPPF 结构和 DySample 机制, 为金属表面缺陷检测提供了一种高精度的解决方案。

然而上述方法存在共性局限。首先, 过度聚焦浅层-深层特征融合, 忽视通道间信息交互, 导致边缘特征丢失; 其次, 检测头设计依赖复杂结构(如多层堆叠卷积或完全解耦机制), 在提升精度的同时显著增加参数量; 最后, 上述优化难以平衡精度与推理效率的工业需求。

钢材缺陷类型复杂、多尺度特性, 且具有类间相似性和强背景干扰, 导致原生 YOLO 网络难以满足实际需求, 亟需结合任务特性与缺陷特征进行网络结构优化, 以提升模型鲁棒性与检测性能。为应对上述挑战, 且兼顾实际工业部署中对检测精度与推理速度的双重要求, 本文选取 YOLO11 作为基线模型并对其进行改进, 以进一步提升模型检测性能。

针对上述问题, 针对钢材缺陷检测应用, 本文提出一种基于多尺度特征感知轻量化网络 LiteSteel-YOLO, 并在两

种钢板缺陷公开数据集上验证了其有效性, 主要创新为:

1) 提出轻量级多尺度融合模块(C3k2-lightweight multi-scale fusion, C3k2-LMSF), 采用并行多尺度卷积核架构与通道压缩策略, 实现跨尺度特征的高效提取与融合。

2) 设计空间通道感知上采样模块(spatial-channel aware upsampling module, SCAM), 通过通道重组与空间偏移运算建立长程依赖, 引导网络聚焦于更具区分性的区域, 抑制复杂背景纹理对目标检测造成的干扰, 增强缺陷区域特征响应强度。

3) 构建 Efficient-Head 检测头, 基于分组卷积与结构重构优化计算路径, 提高推理速度, 在不增加模型参数的前提下, 保持对多类缺陷的稳定检测能力, 同时提升对定位精度要求高的小目标和长条形缺陷和整体检测精度。

## 1 方 法

### 1.1 LiteSteel-YOLO 网络结构

本文提出的 LiteSteel-YOLO 网络基于 YOLO11<sup>[21]</sup> 进行轻量化优化以满足工业部署的实时性需求, 网络采用主干网络、颈部网络以及检测头这 3 部分组成, 其结构如图 1 所示。LiteSteel-YOLO 将特征输入主干网络执行多尺度特征提取; 随后进入颈部网络进行特征融合, 在逐级上采样过程中, 本文设计 SCAM 对输入特征进行上采样, 以增强网络的空间建模能力和特征之间的信息交互; 最后, 将融合后的特征图送入不同的 Efficient Head 进行检测, 其分组卷积与并行网络的设计, 能够为网络带来更加轻量化和高效的特性。与 YOLO-LSDF<sup>[19]</sup>、FD-YOLO11<sup>[20]</sup> 等网络相比, LiteSteel-YOLO 在结构上进一步平衡了检测精度与推理速度: 前者多通过复杂卷积结构提升精度但牺牲实时性, 而本文通过轻量化多尺度融合和高效上采样策略, 实现了在保证检测性能的同时显著提升推理效率, 从而更契合工业实时检测需求。

### 1.2 C3k2-LMSF 模块

钢材表面缺陷类型具有多尺度的特点, C3k2 中的固定卷积核使得其感受野有限, 无法捕捉细线形状的结构, 容易漏检或者误检; 且其缺乏对模糊边缘的建模机制, 从而导致在强纹理背景下小凹坑特征信息丢失现象。

针对上述不足, 本文通过设计多尺度特征融合模块 C3k2-LMSF 对 C3k2 进行改进, 以此增强其对多尺度缺陷的检测能力与效率, 模块结构如图 2 所示。输入特征首先经过  $1 \times 1$  卷积进行通道调整, 随后将特征复制为两部分, 一部分经过主路径 Bottleneck-LMSF, 另一部分直接跳连至 Concat, 并再次通过逐点卷积调整通道后输出。

为了同时兼顾计算量与检测精度, 本文只在 Backbone 的末端和 Head 的末端使用 C3k2-LMSF 模块。相比深层结构, 浅层网络特征图尺寸较大, 使用 C3k2 模块将降低 FLOPs; 且浅层特征具有高分辨率和细节信息, 过度复杂操作会导致过拟合。而深层特征图分辨率较低, 且需要更大

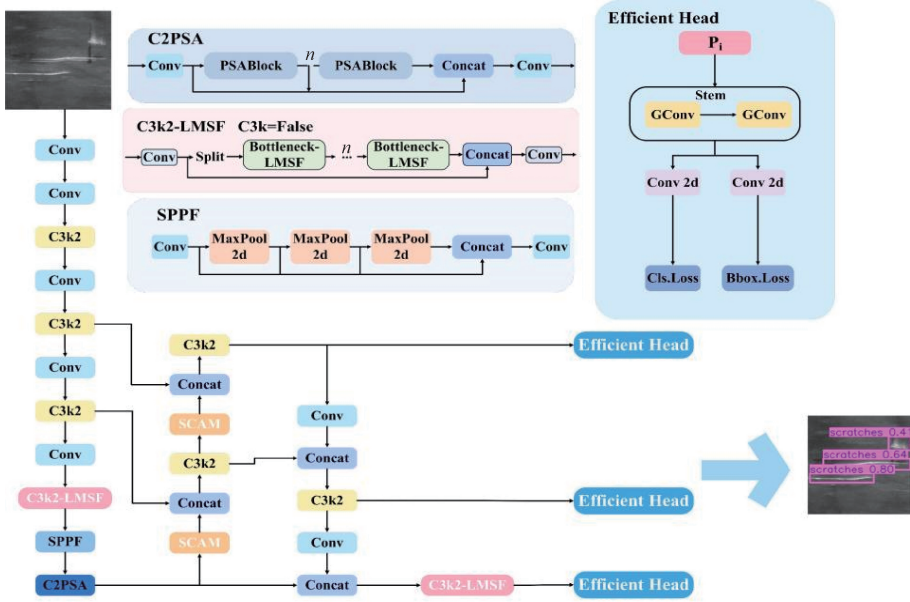


图 1 LiteSteel-YOLO 网络主要结构

Fig.1 LiteSteel-YOLO network main structure

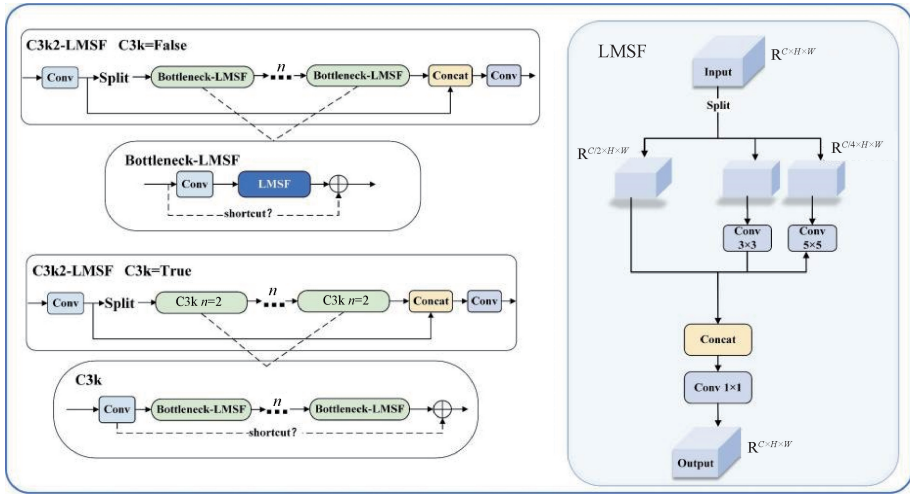


图 2 C3k2-LMSF 模块结构

Fig.2 C3k2-LMSF module structure

的感受野, C3k2-LMSF 模块的多尺度检测能力可实现效率与精度的良好平衡。

相较于传统的 C3k2 模块, 本文提出的 C3k2-LMSF 通过融合多尺度卷积核和特征指导机制, 有效提升了网络对多尺度缺陷特征的提取能力与检测效率。与现有改进 YOLO11 网络(如 FD-YOLO11<sup>[20]</sup>)相比, FD-YOLO11 虽然通过自校准卷积结合 C3k2 结构在精度上有所提升, 但检测速度显著下降, 难以满足实时性场景的需求; 而 C3k2-LMSF 在保持较高检测速度的同时, 实现了多尺度特征的增强, 从而在精度与效率之间取得更优平衡。C3k2-LMSF 中的核心模块 LMSF 流程如下: 对于输入特征  $\mathbf{X} \in \mathbf{R}^{B \times C \times H \times W}$ , 通过 Split 操作按照通道维度将其分为两部分

$\mathbf{X}_{cheap} = \mathbf{R}^{B \times \frac{C}{2} \times H \times W}$  和  $\mathbf{X}_{group} = \mathbf{R}^{B \times \frac{C}{2} \times H \times W}$ , 随后将  $\mathbf{X}_{group}$  部分

依据卷积核数量进行重排列得到  $\mathbf{X}_g \in \mathbf{R}^{B \times \frac{C}{2g} \times H \times W \times g}$ , 使其分别经过不同卷积核大小的卷积操作, 并于 cheap 分支的特征进行 Concat 操作, 经过逐点卷积得到最终输出结果, 即:

$$\mathbf{X}_{cheap}, \mathbf{X}_{group} = Split(\mathbf{X}, dim = C) \quad (1)$$

$$\mathbf{Y}_i = Conv(\mathbf{X}_g^{(i)}), i = 1, 2, \dots, g \quad (2)$$

$$\mathbf{Y} = Concat(\mathbf{Y}_1, \mathbf{Y}_2, \dots, \mathbf{Y}_g), \mathbf{Y} \in \mathbf{R}^{B \times \frac{C}{2} \times H \times W} \quad (3)$$

$$\mathbf{Z} = Concat(\mathbf{X}_{cheap}, \mathbf{Z}), \mathbf{Z} \in \mathbf{R}^{B \times C \times H \times W} \quad (4)$$

$$\hat{\mathbf{Z}} = Conv_{1 \times 1}(\mathbf{Z}) \quad (5)$$

式中:  $B$  为  $Batchsize$ ;  $C$  为通道数;  $H$  为高度;  $W$  为深度;  $k_i$  为第  $i$  个分支的卷积核尺寸;  $\hat{Z}$  为最终 LMSF 模块的输出结果。

### 1.3 SCAM 模块

为提升网络在上采样过程中的特征信息处理能力, 本文设计一种空间-通道感知模块 SCAM, 其结构如图 3 所示, SCAM 核心在于通过对通道分割与空间平移操作增强特征的表达能力, 可以帮助网络实现复杂的特征交互效果, 在保证网络轻量化的同时显著提升特征表达能力。SCAM 整体流程如下: 首先, 输入特征  $\mathbf{X} \in \mathbf{R}^{B \times C \times H \times W}$  经过 2 倍上采样操作提升空间分辨率, 并经过深度可分离卷积实现局部空间信息提取, 其次, 通过 Channel Shuffle 操作使得通道被重新分组和打乱, 以此实现跨通道的特征混合, 规避通道间信息相互独立的弊端, 提升特征融合的效果:

$$\mathbf{X}_{reshape} = Reshape\left(\mathbf{X}_{dwc}; B; G; \frac{C}{G}; 2H; 2W\right) \quad (6)$$

$$\mathbf{X}_{transposed} = Transposed\left(\mathbf{X}_{resaped}, \dim = 1, 2\right) \quad (7)$$

$$\mathbf{X}_{shuffle} = Reshape\left(\mathbf{X}_{transposed}; B; C; 2H; 2W\right) \quad (8)$$

式中:  $Reshape(\mathbf{X}; \dots)$  为将张量按照给定维度进行变形;  $Transposed(\mathbf{X}, \dim = 1, 2)$  是指将第 1 维度和第 2 维度进行交换, 即  $\mathbf{X}_{transposed} \in \left(B, \frac{C}{G}, G, 2H, 2W\right)$ 。输入特征经过通道混洗操作后, 再通过 Channel Shuffle 操作将其沿着通道维度划分为 4 组, 分别在水平方向和垂直方向进行平移; 平移后的特征通过通道拼接操作重新组合, 以实现跨通道的信息交互与局部特征增强, 即:

$$[\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3, \mathbf{X}_4] = Split(\mathbf{X}_{shuffle}, chunks = 4, \dim = 1) \quad (9)$$

$$\mathbf{X}'_1 = Roll(\mathbf{X}_1, +1, \dim = 2) \quad (\text{向上平移}) \quad (10)$$

$$\mathbf{X}'_2 = Roll(\mathbf{X}_2, -1, \dim = 2) \quad (\text{向下平移}) \quad (11)$$

$$\mathbf{X}'_3 = Roll(\mathbf{X}_3, +1, \dim = 3) \quad (\text{向右平移}) \quad (12)$$

$$\mathbf{X}'_4 = Roll(\mathbf{X}_4, -1, \dim = 3) \quad (\text{向左平移}) \quad (13)$$

式中:  $\mathbf{X}'_i = Roll(\mathbf{X}_i, s_i, \dim = d_i)$ ,  $s_i \in \mathbf{Z}, d_i \in \{2, 3\}$  是指将特征  $\mathbf{X}_i$  沿着  $d_i$  维度平移  $s_i$  个步长, 向下和向右平移记为正, 向上和向左平移记为负。将 4 组平移后的特征拼接为一个完整的张量:

$$\mathbf{X}_{mix} = Concat([\mathbf{X}'_1, \mathbf{X}'_2, \mathbf{X}'_3, \mathbf{X}'_4], \dim = 1) \quad (14)$$

最后通过逐点卷积整合通道维度信息, 即:

$$\mathbf{Y} = Conv_{1 \times 1}(\mathbf{X}_{mix}) \quad (15)$$

本文提出 SCAM 模块通过通道混洗与通道移位操作, 以解决上采样过程中信息交互不足的问题, 强化了通道间和空间维度的信息交互能力, 帮助网络建立更加丰富的特征表示。相比简单的上采样操作以及 YOLO-LSDI 中的 LDConv 空间维度建模, SCAM 在保持计算开销低的同时显著提升了上采样特征的质量, 从而增强了网络的表达能力和检测精度。

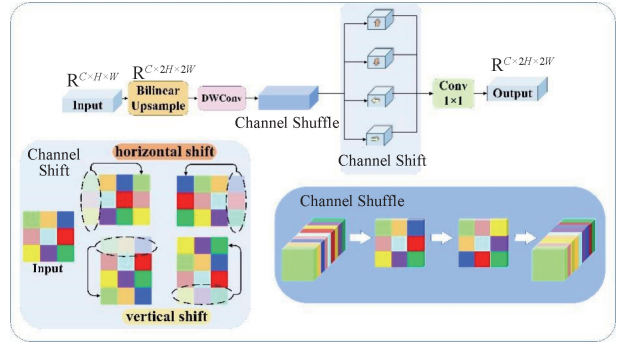


图 3 SCAM 模块结构

Fig. 3 SCAM module structure

### 1.4 Efficient Head 检测头

YOLO11 的检测头包含定位和分类两个分支, 每个分支包含多个串联的卷积块。该结构具有良好的任务解耦性, 但存在早期特征共享缺失问题, 导致表达能力受限, 且两个独立分支容易导致参数量增大。针对该问题, 本文提出一种更为高效和轻量的检测头 Efficient Head, 其结构如图 4 所示, 区别于常规检测头中采用的独立卷积分支或全卷积操作, 其通过设计两层连续的  $3 \times 3$  分组卷积对输入特征图进行统一的编码: 即在前两层卷积中对特征进行统一建模, 在第 1 层实现局部特征聚合任务, 在第 2 层进一步增强跨组交互。

具体而言, 将每层输入特征  $\mathbf{X} \in \mathbf{R}^{C_i \times H \times W}$  的通道数除以 16 得到每层卷积的分组数, 并将其经过分组卷积处理, 在保证局部感受野的同时, 一定程度上减少计算量和参数量:

$$Groups_i = \frac{C_i}{16}, i = 0, 1, 2 \quad (16)$$

$$\mathbf{F}_1 = GConv_{3 \times 3}(\mathbf{X}) \quad (17)$$

$$\mathbf{F}_2 = GConv_{3 \times 3}(\mathbf{F}_1) \quad (18)$$

式中:  $Groups_i$  表示第  $i$  层特征的分组数;  $C_i$  为第  $i$  层的通道数;  $GConv_{3 \times 3}$  为分组卷积操作;  $\mathbf{F}_i$  为经过第  $i$  个分组卷积处理后的特征图。随后, 通过分组卷积处理后的特征信息分别进入分类分支和回归分支, 实现任务解耦。

在分类分支中, 首先利用逐点卷积对输入特征进行通道映射, 随后结合 Sigmoid 函数生成目标类别的概率分布, 即:

$$Cls = \sigma(Conv_{1 \times 1}(\mathbf{F}_i)) \quad (19)$$

式中:  $\sigma$  为 Sigmoid 激活函数;  $Cls$  为每个空间位置的类别预测结果。在边界框回归分支中, 同样利用逐点卷积生成候选框的边界分布, 此时生成的结果为每个空间位置对应的 4 个边界的离散概率分布, 再利用分布焦点损失函数 (distribution focal loss, DFL) 将离散分布转化为连续值, 得到最终的回归边界框坐标, 即:

$$Bbox = DFL(Conv_{1 \times 1}(\mathbf{F}_i)) \quad (20)$$

式中:  $Bbox$  表示每个位置的边界框预测结果。

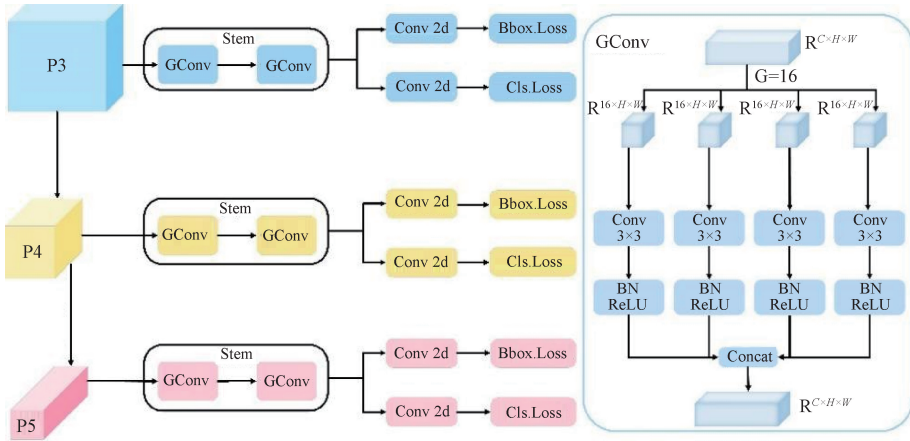


图 4 Efficient Head 结构

Fig. 4 Efficient Head structure

Efficient Head 通过分组卷积与分类和定位并行分支的结构设计,使得每一组仅负责少量特征通道,以实现任务特征的精细分配,提高检测稳定性,从而在准确性和计算效率之间取得良好的平衡,便于工业环境下部署。

## 2 实验与结果分析

### 2.1 数据集

为验证本文方法的有效性和泛化能力,在 NEU-DET 和 GC10-DET 两个公开钢材表面缺陷数据集分别进行了消融实验和与 SOTA 方法的对比实验。

1)NEU-DET 数据集:包含热轧钢材表面 6 种典型缺陷,总计 18 00 张图像,每类缺陷各 300 张。实验中,按照 8:2 的比例将数据集划分为 40 张训练图像和 360 张测试图像。

2)GC10-DET 数据集:包含 10 种钢材表面缺陷类型,总计 2 280 张图像。实验中,按照 8:2 比例划分为 1 835 张训练图像和 459 张测试图像。

### 2.2 实验环境

本文实验硬件环境为: Intel(R) Core(TM) i5-12400 (16 GB 内存)、NVIDIA GeForce RTX 4060(8 GB 内存),使用 PyTorch 深度学习框架来完成实验。PyTorch 版本为 2.0.0,CUDA 版本为 11.8。在网络训练过程中,采用 SGD 优化器,训练时期为 300 个,批量大小为 4。动量和权重衰减分别设置为 0.937 和 0.000 5,初始学习率为 0.01。

### 2.3 评价指标

为全面评估模型性能,采用精确度(Precision)、召回率(Recall)、平均精度(AP)和均值平均精度(mAP)等广泛使用的目标检测指标其表达式如式(21)~(24)所示。

$$precision = \frac{TP}{TP + FP} \quad (21)$$

$$recall = \frac{TP}{TP + FN} \quad (22)$$

$$AP = \int_0^1 P(R) dR \quad (23)$$

$$mAP = \frac{1}{n} \sum_{i=1}^n AP_i \quad (24)$$

式中:  $TP$  表示真阳性,即模型正确检测到的正样本;  $FP$  表示假阳性,即模型错误的将负样本标记为正样本;  $FN$  表示假阴性,即模型未能检测到的真实样本;  $n$  表示缺陷类别的数量;  $precision$  为精确度,用于衡量模型预测结果的准确性,即用于预测“正样本中多少为真”;  $recall$  为召回率,用于衡量模型捕捉正样本的能力,即“真实样本中有多少被找到”;  $AP$  用于综合反映模型对某种类别的检测能力,  $mAP$  则用于反映模型的整体性能。

此外,为评估模型的计算效率和复杂度,本文同时将参数量、FPS 以及计算量纳入评价指标中。考虑到部分指标会受网络中所含卷积的类型、数量以及硬件环境的影响,本文以  $mAP$  作为主要评价指标。

### 2.4 消融实验结果及讨论

为系统验证本文提出的模块在钢材表面缺陷检测中的有效性,本文以 YOLO11n 为基线模型进行消融实验。实验通过逐步将各模块集成到基线模型中,并和基线模型对比其  $mAP$  和计算量等指标,以量化各模块的贡献。消融实验结果如表 1 所示。

模型 I 在 Baseline 的基础上改进并引入 C3k2-LMSF 模块,提升模型的多尺度缺陷检测能力。相比 Baseline 的结果,模型 I 的  $mAP@50$  增加了 2.3%,参数量下降了 3.4%,检测速度提升了 5.6%。模型 II 在 Baseline 的基础上对上采样结构进行改进,如表 1 所示,模型 II 的  $mAP@50$  相比 Baseline 增加了 0.9%。模型 III 通过将原网络检测头改进为 Efficient Head 以解决原检测头参数冗余、信息流不充分问题,模型 III 的  $mAP@50$  相比 Baseline 提升了 0.7%,同时参数量下降了 10%,计算量下降了 20%,在保证检测精度的同时实现了轻量化的效果。

将 C3k2-LMSF 和 SCAM 同时加入 Baseline 得到模型

表 1 基于 NEU-DET 的消融实验

Table 1 Ablation experiment based on NEU-DET

模型	C3k2-LMSF	SCAM	Efficient Head	mAP@50/%	参数量/M	计算量/G	FPS
Baseline				77.7	2.58	6.3	318.77
I	✓			80.0	2.49	6.2	336.69
II		✓		78.6	2.67	6.8	320.76
III			✓	78.4	2.31	5.1	371.18
IV	✓	✓		79.5	2.57	6.7	314.78
V	✓		✓	80.4	2.22	5.1	368.34
VI		✓	✓	78.7	2.40	5.6	343.35
VII	✓	✓	✓	81.7	2.30	5.5	338.63

IV, 如表 1 实验结果所示, 相比模型 I 和 II 其 mAP@50 其 mAP@50 分别下降 0.5% 和上升 0.9%。相比模型 I 其 mAP@50 有所下降的原因在于 C3k2-LMSF 已经有效地整合了多尺度的信息, SCAM 的加入进一步强化了此操作, 出现了特征权重的竞争或者冗余, 导致个别类别的检测结果有所波动。但是相比 Baseline, 模型 IV 的 mAP@50 上升了 1.8%。

将 C3k2-LMSF 和 Efficient Head 同时加入 Baseline 得到模型 V, 由表 1 所示实验结果可以看出, 相比分别单独使用上述模块该组合 mAP@50 分别提高 0.4% 和 1.8%。此外, 模型 V 获得了消融实验中最小的参数量 (2.22 M) 和最高的推理速度 (368.34 FPS)。上述提升得益于 C3k2-LMSF 模块的多分支结构同时兼顾了浅层特征和深层语义信息, 提升了对多尺度缺陷的检测能力; 同时, 轻量化检测头较好地保留特征信息。

将 SCAM 和 Efficient Head 两个模块引入 Baseline 得到模型 VI。由表 1 实验结果可得, 相比分别单独使用上述模块该组合 mAP@50 提升了 0.1% 和 0.3%, 在参数量与推理速度方面均得到了平衡。

最后, 将 3 个模块都加入 Baseline 的本文模型 VII 的 mAP@50 相比基准网络提升了 4%, 参数量下降 10%, 计算量下降 12.7%。由于斑块缺陷形态纹理相对简单, 进一步的结构特征增强对其增益较小, 而纹理复杂、边界模糊以及多尺度特征的强化对其带来干扰, 导致精度有 2.4% 的小幅波动。基于以上消融实验结果, 证明本文模型具有较好的适应性和泛化能力。

## 2.5 对比实验结果及讨论

为进一步验证本文模型在钢材缺陷检测任务中的性能, 选取主流目标检测方法进行对比试验 (结果如表 2 所示), 包括 YOLOv8n、YOLO11n<sup>[21]</sup>、YOLOv10n<sup>[22]</sup> 等。所有实验均采用相同的训练配置在公开数据集 NEU-DET 和 GC10-DET 上进行以保证公平性。

分析表 2 结果可知, LiteSteel-YOLO 在钢材缺陷检测中展现显著优势, 其 mAP@50 达到了 81.7%, 在精度上优于多数轻量级网络如 YOLOv8n 等, 相较于 YOLO-LSDI

表 2 基于 NEU-DET 数据集对比实验结果

Table 2 Comparative experimental results on the NEU-DET dataset

类型	mAP@50/ %	参数量/ M	计算量/ G	FPS
YOLOv8n	72.8	3.2	8.9	215.2
YOLOv10n	74.2	2.7	8.4	172.4
YOLO11n	77.7	2.6	6.4	343.3
文献[14]	81.6	4.4	8.7	83.6
文献[17]	78.1	2.2	7.1	—
YOLO-LSDI <sup>[19]</sup>	83	2.7	6.1	162.1
FD-YOLO11 <sup>[20]</sup>	81.1	—	—	67.9
本文	81.70	2.30	5.50	338.63

仅低 1.3%, 此外, LiteSteel-YOLO 运算量为 5.5 GFLOPs, 参数量为 2.3 M, 且实测推理速度达到 338.63 FPS, 综合来看, LiteSteel-YOLO 在精度、模型轻量化与推理速度之间达成了良好平衡, 具有较强的工程适配性与部署价值。

本文将近两年 YOLO 主流模型和本文模型在 NEU-DET 数据集上的对比实验进行了可视化分析, 结果如图 5 所示。

通过图 5 可视化结果可以得出本文模型能清晰、完整地将目标 (如轧入氧化皮, 划痕, 夹杂物) 与背景分离, 显著降低背景噪声影响, 提高检测置信度并减少漏检。尤其对于裂纹缺陷, 本文模型检测置信度显著高于 YOLO11n。此外, 为了验证本文所提出的方法的泛化能力, 在更具挑战性的 GC10-DET 数据集上进行了对比实验, 实验结果如表 3 所示, 本文模型同样取得了较优的 mAP@50 70.7%, 相比 YOLO11n (68.4%) 提高 2.3%, 验证了模型的泛化能力。与 YOLOv8n (65.4%) 相比提升 5.3%, 并略优于文献[15] (70.4%), 仅次于 FD-YOLO11<sup>[20]</sup> (71.3%)。该结果表明本文方法在不同数据分布下仍能保持稳健的检测性能, 证明了较好的泛化能力。

在 GC10-DET 数据集上的可视化结果如图 6 所示, 结果证明, 在光照条件不足或者存在背景干扰的复杂场景下

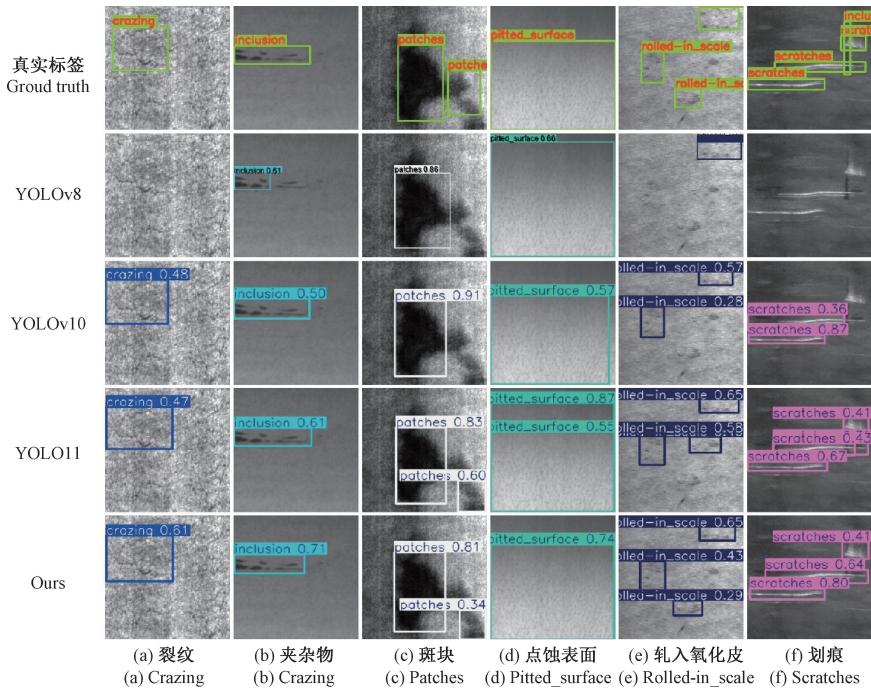


图 5 LiteSteel-YOLO 在 NEU-DET 数据集上的可视化结果

Fig. 5 Visualization results of LiteSteel-YOLO on the NEU-DET dataset

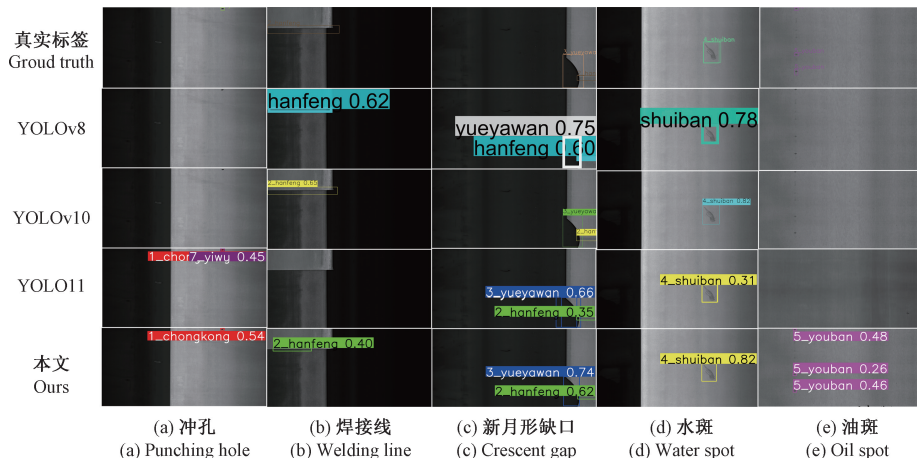
表 3 基于 GC10-DET 数据集对比实验结果

Table 3 Comparative experimental results on the GC10-DET dataset

模型	mAP@50/	参数量/	计算量/	FPS
	%	M	G	
YOLOv8n	65.4	3.1	8.8	243
YOLOv10n	64.5	2.7	8.4	172.4
YOLO11n	68.4	2.6	6.4	569.8
文献[14]	70.4	4.4	8.7	83.6
文献[17]	65.2	2.2	7.1	—
YOLO-LSDI <sup>[19]</sup>	66.5	2.6	6.1	156.3
FD-YOLO11 <sup>[20]</sup>	71.3	—	—	66.7
本文	70.7	2.3	5.5	530.3

(如轧坑(Rp), 夹杂物(In)), 本文模型也能够更加精准地定位缺陷。

综上, 本文模型在 NEU-DET 和 GC10-DET 数据集上均取得了较高的 mAP@50(81.7% 和 70.7%), 证明其在钢材表面缺陷检测任务中相比主流模型具有显著优势; 在多数缺陷类别上的优异表现证明模型对多尺度缺陷的泛化能力。同时, LiteSteel-YOLO 计算效率高、模型参数量小(约 2.3 M)、计算量低(5.5 GFLOPs)、推理速度较快(338.63 FPS, 530.3 FPS), 具备优异的工业部署可行性。在缺陷差异较大的 GC10-DET 数据集上取得最佳性能, 验证了模型对不同场景、不同缺陷类型的泛化能力。可视化结果表明, LiteSteel-YOLO 在复杂背景、光照变化、噪声干扰等挑战下, 仍能保持精准的缺陷定位和识别能力, 具有较好的鲁棒性。



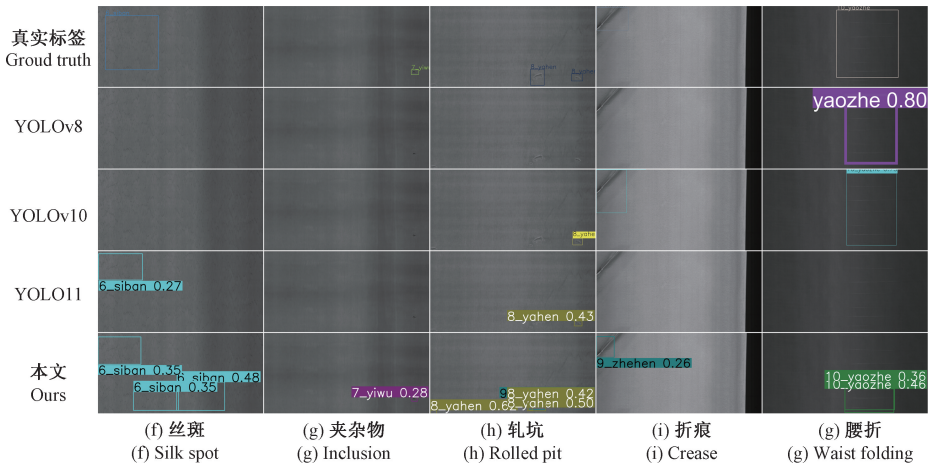


图 6 LiteSteel-YOLO 在 GC-10 数据集上的可视化结果

Fig. 6 Visualization results of LiteSteel-YOLO on the GC-10 dataset

### 3 结 论

本文提出了一种兼顾轻量化设计与检测精度的钢材缺陷检测方法 LiteSteel-YOLO。该方法融合了多尺度语义增强、背景信息抑制以及高效的检测头机制。为应对钢材缺陷存在不同尺寸、不同形态的检测难点,提出具有多尺度卷积融合和特征引导机制的 C3k2-LMSF 模块,有效提升网络对不同缺陷的感知能力。通过 SCAM 结构引入通道重组和空间偏移运算,在保持上采样语义连续性的同时,增强了目标区域的响应能力并抑制了背景噪声,有效提升了小目标检测和边界分割的鲁棒性。基于轻量化设计和解耦策略的检测头,既降低了网络的参数量与计算量,还确保了良好的检测精度。实验结果表明, LiteSteel-YOLO 在公开数据集上的 mAP@50 分别达到 81.7% 和 70.7%,较基线模型 YOLO11 提升了 4% 和 2.3%,检测速度上分别达到 338.63 与 530.3 FPS。该模型在检测精度、效率与轻量化之间实现了良好的平衡,展现出较好的综合应用潜力。未来本研究将重点优化模型在嵌入式设备上的部署效率,提升资源利用率;同时,针对复杂、多样化缺陷,增强模型的鲁棒性和泛化能力,以满足实际工业检测中的多场景应用需求。

### 参考文献

- [1] 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 and Machine Intelligence, 2017, 39(6): 1137-1149.
- [2] SUN P Z, ZHANG R F, JIANG Y, et al. Sparse R-CNN: End-to-end object detection with learnable proposals [C]. 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021: 14449-14458.

- [3] 陆家林,程颖,冯赛,等. 基于机器视觉的钢材表面缺陷检测[J]. 机电工程技术, 2022, 51(7): 159-163.  
LU J L, CHENG Y, FENG S, et al. Steel surface defect detection based on machine vision [J]. Mechanical & Electrical Engineering Technology, 2022, 51(7): 159-163.
- [4] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: Unified, real-time object detection[C]. 2016 IEEE Conference on Computer Vision and Pattern Recognition, 2016: 779-788.
- [5] REDMON J, FARHADI A. YOLOv3: An incremental improvement[J]. ArXiv preprint arXiv: 1804.02767, 2018.
- [6] 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]. 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023: 7464-7475.
- [7] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single shot multibox detector [C]. European Conference on Computer Vision, 2016: 21-37.
- [8] HE K M, GKIOXARI G, DOLLÁR P, et al. Mask R-CNN[C]. 2017 IEEE International Conference on Computer Vision, 2017: 2980-2988.
- [9] ZHAO CH, SHU X, YAN X, et al. RDD-YOLO: A modified YOLO for detection of steel surface defects[J]. Measurement, 2023, 214: 112776.
- [10] ZHANG Y ZH, WANG W J, LI ZH M, et al. Development of a cross-scale weighted feature fusion network for hot-rolled steel surface defect detection[J]. Engineering Applications of Artificial Intelligence, 2023, 117: 105628.
- [11] YU J L, SHI X G, WANG W H, et al. LCG-YOLO:

- A real-time surface defect detection method for metal components[J]. IEEE Access, 2024, 12: 41436-41451.
- [12] YOU C ZH, KONG H ZH. Improved steel surface defect detection algorithm based on YOLOv8 [J]. IEEE Access, 2024, 12: 99570-99577.
- [13] LI ZH CH, XIAO L X, SHEN M L, et al. A lightweight YOLOv8-based model with squeeze-and-excitation version 2 for crack detection of pipelines[J]. Applied Soft Computing, 2025, 177: 113260.
- [14] WANG G Q, ZHANG CH ZH, CHEN M S, et al. A high-accuracy and lightweight detector based on a graph convolution network for strip surface defect detection [J]. Advanced Engineering Informatics, 2024, 59: 102280.
- [15] 雷超,陈德基,孙家栋,等. 基于 SGF-YOLO 的钢板缺陷检测方法[J]. 电子测量技术, 2025, 48(21): 215-225.  
LEI CH, CHEN D J, SUN J D, et al. Steel plate defect detection method based on SGF-YOLO [J]. Electronic Measurement Technology, 2025, 48(21): 215-225.
- [16] 吴葛,朱宇凡,叶天成. 改进 YOLOv10 的钢材表面缺陷检测方法[J]. 电子测量技术, 2025, 48(4): 158-168.  
WU G, ZHU Y F, YE T CH. Improved YOLOv10 for steel surface defect detection [J]. Electronic Measurement Technology, 2025, 48(4): 158-168.
- [17] 胥姜苗,曹爽,管海燕. 基于改进 YOLOv8 的轻量化钢材表面缺陷检测方法[J]. 电子测量技术, 2025, 48(24):138-147.  
XU J M, CAO SH, GUAN H Y. A lightweight steel surface defect detection method based on improved YOLOv8[J]. Electronic Measurement Technology, 2025, 48(24):138-147.
- [18] 徐森,徐畅,花小鹏,等. 改进 YOLOv8n 的热轧钢材表面缺陷检测[J]. 哈尔滨工程大学学报, 2025, 46(12):2596-2603,2630.  
XU S, XU CH, HUA X P, et al. Improved YOLOv8n for hot-rolled steel surface defect detection[J]. Journal of Harbin Engineering University, 2025, 46(12):2596-2603, 2630.
- [19] WANG F Q, JIANG X B, HAN Y, et al. YOLO-LSDI: An enhanced algorithm for steel surface defect detection using a YOLOv11 network[J]. Electronics, 2025, 14(13): 2576.
- [20] DANG Z CH, WANG X SH. FD-YOLO11: A feature-enhanced deep learning model for steel surface defect detection[J]. IEEE Access, 2025, 13: 63981-63993.
- [21] KHANAM R, HUSSAIN M. YOLOv11: An overview of the key architectural enhancements [J]. ArXiv preprint arXiv:2410.17725, 2024.
- [22] WANG AO, CHEN H, LIU L H, et al. YOLOv10: Real-time end-to-end object detection [J]. ArXiv preprint arXiv:2405.14458, 2024.

### 作者简介

**周迅**, 硕士研究生, 主要研究方向为计算机视觉、缺陷检测。

E-mail:zhouxunn\_n@163.com

**李帆**, 硕士研究生, 主要研究方向为计算机视觉、缺陷检测。

E-mail:lf\_2426923409@163.com

**张岩**(通信作者), 博士, 教授, 硕士生导师, 主要研究方向为无损检测、计算机视觉、深度学习。

E-mail:zy@qust.edu.cn