

文章编号:1007-2780(XXXX)XX-0001-15

## ICR-DETR:面向工业领域的无损检测模型

陈俊<sup>1</sup>, 谢邦天<sup>2</sup>, 梅育青<sup>1</sup>, 范俊<sup>2</sup>, 张艳波<sup>2</sup>, 陈鹏<sup>2\*</sup>

(1. 中国长江电力股份有限公司, 湖北宜昌 443004;

2. 微特技术有限公司, 湖北宜昌 443005)

**摘要:** 钢材表面缺陷的准确高效检测对工业质量控制至关重要。RT-DETR模型虽然在速度和精度之间提供了良好平衡,但在处理复杂、细粒度缺陷时,其固定的特征提取方式和检测精度仍存在局限。为此,本文提出一种基于RT-DETR改进的检测算法——ICR-DETR。首先,采用UniRepLKNet作为骨干网络,显著提升模型对金属表面特征的感知能力;其次,在颈部网络中引入LWN-CS模块,结合可学习小波变换与通道混洗机制,有效增强高频细节建模与特征融合能力;最后,设计Shape-WIoU损失函数,将WIoU的非单调样本加权机制与Shape-IoU融合,综合建模边界框的形状特征与尺度相关性,提升定位精度与鲁棒性。实验在公开的钢材表面缺陷数据集NEU-DET上展开,结果表明ICR-DETR在准确率、召回率与mAP方面分别达到77.92%,75.71%与78.42%,优于现有主流检测算法。此外,为验证模型泛化能力,还在自建的滑触线脱落数据集ICRFD以及重卡裂纹数据集FDMPI上进行了测试,验证了其在复杂工业环境下的实用性和鲁棒性。实验结果表明,所提ICR-DETR能有效支持复杂工业场景下的高精度缺陷检测任务。

**关键词:** 缺陷检测;RT-DETR;滑触线;Shape-IoU;NEU-DET

中图分类号:TP394.1;TH691.9 文献标识码:A doi:10.37188/CJLCD.2026-0024 CSTR:32172.14.CJLCD.2026-0024

## ICR-DETR: A lossless detection model for the industrial sector

CHEN Jun<sup>1</sup>, XIE Bangtian<sup>2</sup>, MEI Yuqing<sup>1</sup>, FAN Jun<sup>2</sup>, ZHANG Yanbo<sup>2</sup>, CHEN Peng<sup>2\*</sup>

(1. China Yangtze Power Co., Ltd, Yichang 443004, China;

2. WEITE TECHNOLOGIES CO., LTD, Yichang 443005, China)

**Abstract:** Accurate and efficient detection of steel surface defects is very important for industrial quality control. Although the RT-DETR model provides a good balance between speed and accuracy, its fixed feature extraction method and detection accuracy are still limited when dealing with complex and fine-grained defects. To this end, this paper proposes an improved detection algorithm based on RT-DETR-ICR-DETR. Firstly, UniRepLKNet is used as the backbone network to significantly improve the model's perception of metal surface features. Secondly, the LWN-CS module is introduced into the neck network, which combines the learnable wavelet transform and channel shuffle mechanism to effectively enhance the high-frequency detail modeling and feature fusion capabilities. Finally, the Shape-WIoU loss function is designed, and the non-monotonic sample weighting mechanism of WIoU is fused with Shape-IoU to comprehensively model the shape feature and scale correlation of the bounding box, so as to improve the

收稿日期:2026-03-03;修订日期:2026-04-13.

基金项目:中国长江电力股份有限公司资助项目(No.Z222302037)

Supported by Project of China Yangtze Power Co., Ltd. (No.Z222302037)

\*通信联系人, E-mail: zhb0213@126.com

positioning accuracy and robustness. The experiment is carried out on the open steel surface defect data set NEU-DET. The results show that ICR-DETR achieves 77.92 %, 75.71 % and 78.42 % in accuracy, recall and mAP, respectively, which is superior to the existing mainstream detection algorithms. In addition, in order to verify the generalization ability of the model, the self-built sliding contact line shedding data set ICRFD and heavy truck crack data set FDMPI were tested to verify its practicability and robustness in complex industrial environments. The experimental results show that the proposed ICR-DETR can effectively support high-precision defect detection tasks in complex industrial scenarios.

**Key words:** defect detection; RT-DETR; sliding contact line; shape-IoU; NEU-DET

## 1 引言

自动化视觉检测(AVI)系统在现代钢铁制造中不可或缺,用于确保产品质量、减少浪费并维持生产效率。传统缺陷检测主要依赖人工目视检查,这种方法效率低下,且受人员经验和主观因素影响较大。

近年来,随着深度学习技术的发展,目标检测在区域提议、特征提取与回归机制方面取得了显著突破<sup>[1]</sup>。许多代表性深度学习模型相继提出并不断完善<sup>[2]</sup>。例如,早期的 AlexNet<sup>[3]</sup>奠定了深度学习在计算机视觉中的基础,以及基于区域提议的经典框架 R-CNN<sup>[4]</sup>,其后续版本如 Fast R-CNN<sup>[5]</sup>和 Faster R-CNN<sup>[6]</sup>在检测效率与准确率方面均表现优异。同时,为满足实时检测需求,一阶段检测方法快速发展,如 YOLO<sup>[7]</sup>系列和 SSD<sup>[8]</sup>等模型,在检测速度与准确率之间达到了良好平衡。随着 Transformer 架构在计算机视觉中的广泛应用,涌现出一批基于该结构的目标检测模型,包括 DETR<sup>[9]</sup>、Vision Transformer (ViT)<sup>[10]</sup>以及 RT-DETR<sup>[11]</sup>,显著拓展了检测算法的设计思路与性能边界。这些模型不仅在多个标准数据集上刷新了性能记录,也在实际应用中表现出卓越的泛化能力和鲁棒性。

得益于深度卷积神经网络(CNNs)<sup>[12]</sup>在图像识别<sup>[13-17]</sup>与数据分析方面的卓越能力,机械磨损监测及故障预警等任务已得到显著优化。在现有的金属表面缺陷检测方法研究中,Zhang 等人<sup>[18]</sup>提出了一种嵌入残差模块的加权双向特征金字塔网络,聚合多尺度特征,增强了对多尺度缺陷的检测性能。Liu 等人<sup>[19]</sup>提出多尺度上下文检测网络,构建了具有不同膨胀率的膨胀卷积并行架构,更好地提取多尺度缺陷信息。Zeng 等

人<sup>[20]</sup>提出了一种实时高效的钢材缺陷检测网络,利用多尺度特征提取模块全面捕获不同尺度缺陷的纹理特征。这些文献提出的方法在一定程度上提高了模型的检测精度,但在多尺度融合和模型计算复杂度方面仍存在一定的不足。

针对上述研究中存在的局限性,为提高模型对金属表面多尺度与相似性缺陷目标的检测能力,解决目标分类不准确和检测精度较低的问题,本文提出 ICR-DETR 工业无损检测模型主要贡献如下:

1. 使用 UniRepLKNet 替换原有骨干网络,以扩大感受野并增强空间模式抽象能力,提升整体模型性能。
2. 在颈部网络中创新地引入改进的 LWN-CS 模块,提升去模糊效果,增强对高频细节信息的恢复能力。
3. 设计一种新的损失函数 Shape-WIoU,以更好建模多维特征之间的相关性与尺度差异,提升目标检测的精度与鲁棒性。

## 2 RT-DETR 网络模型

RT-DETR (Real-Time DEtection Transformer)是由百度提出的一种高效端到端实时目标检测模型,旨在解决传统基于 Transformer 的检测方法存在的计算复杂度高、推理速度慢等问题。该模型通过引入轻量级主干网络提取多尺度特征,并采用动态层级交互模块,实现跨层特征融合,有效减少冗余计算,提升计算效率。在解码阶段,RT-DETR 设计了分组查询注意力机制(Group Query Attention)与并行解码策略(Parallel Decoder Strategy),进一步提升推理效率,并保持端到端训练流程,无需依赖非极大值抑制

(NMS)后处理步骤,简化模型结构的同时提升了检测效率。在特征提取阶段,模型通过主干网络对输入图像进行多层特征编码,随后利用基于CNNs的跨尺度特征融合模块对多尺度特征进行增强,以提升模型的空间建模能力,最终将多尺

度特征输入Transformer解码器,以建模全局依赖关系并执行目标预测。

为实现模型参数量与计算复杂度之间的平衡,本文选用轻量化的RT-DETR-l作为基准模型,其网络结构如图1所示。

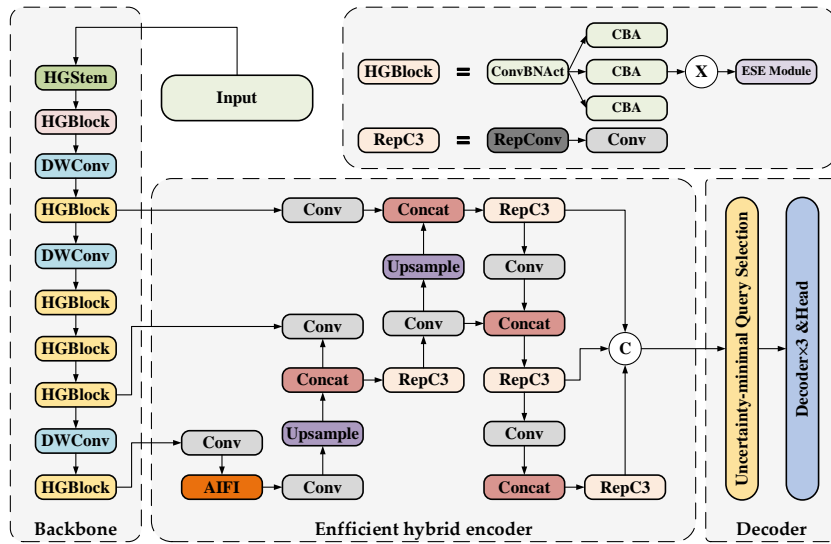


图1 RT-DETR-l模型结构

Fig. 1 Diagram of the RT-DETR-l model structure

### 3 改进算法模型

#### 3.1 改进算法模型

为优化金属表面缺陷检测模型特征提取方

式和检测精度,本文以RT-DETR-l为基线模型进行了改进,提出了一种新的目标检测模型ICR-DETR。具体的模型结构如图2所示。

首先,采用UniRepLKNet替代原有骨干网

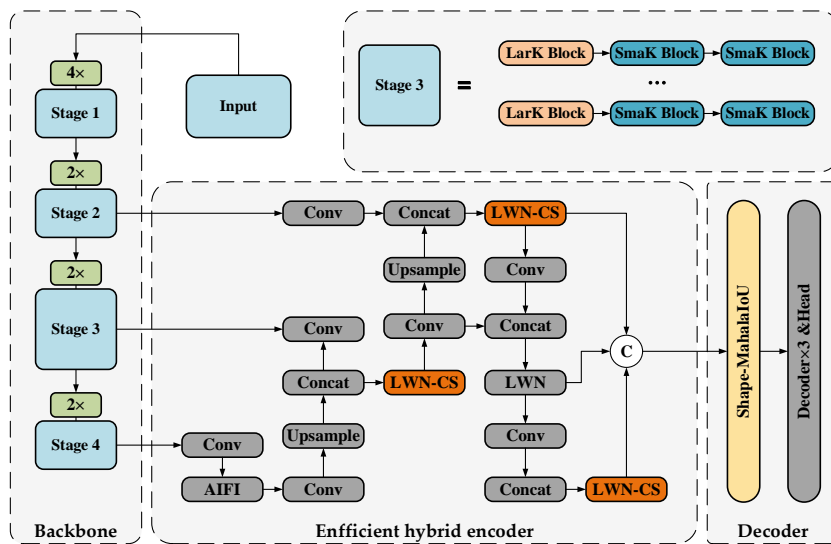


图2 ICR-DETR模型结构图

Fig. 2 Diagram of the ICR-DETR model structure

络,以扩大感受野、增强语义特征表达,提升模型对于结构特征的提取能力;其次,在颈部网络中引入改进的轻量化LWN-CS模块,增强对高频细节信息的恢复能力,从而提升模型的感知能力;最后,设计了一种新的损失函数 Shape-WIoU,通过考虑多维特征之间的相关性与尺度差异,进一步提升检测精度。

### 3.2 优化骨干网络

为提升金属表面缺陷的识别精度,本文将原有骨干网络替换为 UniRepLKNet (Universal Perception Large-Kernel Convolutional Network)<sup>[21]</sup>。该网络架构通过大卷积核与重参数化设计增强了对长距离依赖的建模能力,从而更准确地捕捉金属表面缺陷的边缘细节与整体结构特征。此外,其多尺度特征融合机制提高了对不同尺寸目标的适应性,尤其在小目标检测任务中表现优越。同时,UniRepLKNet具备更强的语义表达能力与鲁棒性,在复杂背景和光照变化等工业环境

中显示出更高的检测稳定性与可靠性。

UniRepLKNet是一种基于CNN的通用感知架构,受 RepLKNet<sup>[22]</sup>启发,采用大感受野卷积核以增强图像空间模式的建模能力。与传统CNN一样,UniRepLKNet通过局部感受野捕捉局部特征,并通过权值共享机制与平移不变性实现高效建模。其层级结构通过多层卷积和下采样操作,逐步构建从低层边缘到高层语义的特征表达。

整体结构上,UniRepLKNet由多个 LarK Block 和 SmaK Block 组成,不同阶段之间通过步长为2的 $3 \times 3$ 稠密卷积进行特征图下采样。其中,LarK Block集成了四个子模块,包括扩张重参数化模块(Dilated Reparam Block)、注意力模块(SE Block)、前馈网络(FFN)与批量归一化层(Batch Normalization);SmaK Block以深度可分离卷积替代了Dilated Reparam Block,保持性能的同时进一步降低计算复杂度。UniRepLKNet的整体结构如图3所示。

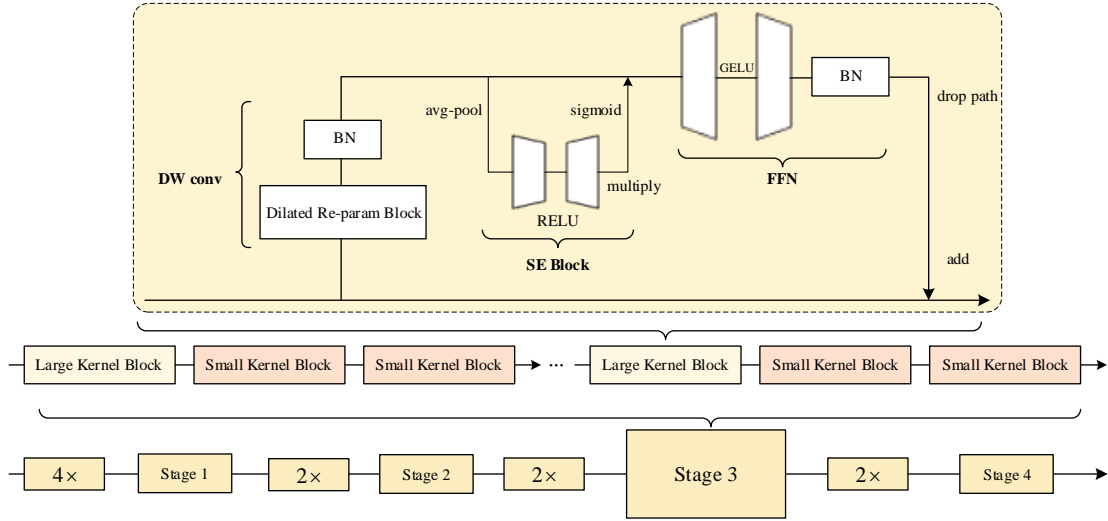


图3 UniRepLKNet网络结构

Fig. 3 The network structure of UniRepLKNet

其中,Dilated Reparam Block是该架构的关键组件,其结构如图4所示。该模块结合膨胀卷积(Dilated Convolution)与结构重参数化技术,实现了高效的大卷积核模拟。具体而言,该模块包含多个不同大小的卷积核分支,分别采用不同的膨胀率。当某一层的卷积核大小为 $k$ ,膨胀率为 $r$ 时,其等效卷积核尺寸为 $K_{eq} = (k-1) \cdot r + 1$ 。通

过多个不同参数的分支组合并使用零填充(zero-padding)对齐输出后融合,即可模拟更大的感受野,相当于实现了一个等效的大卷积核操作。假设有多个卷积核 $\{W_1, W_2, \dots, W_n\}$ ,对应膨胀率 $\{r_1, r_2, \dots, r_n\}$ ,那么最终的等效卷积核可表示为:

$$W_{merged} = \sum_{i=1}^n \text{Pad}(\text{Convert}(W_i, r_i), \text{padding}_i), \quad (1)$$

其中,  $\text{Convert}(W_i, r_i)$  表示将膨胀卷积核  $W_i$  转换为非膨胀形式;  $\text{Pad}(\cdot, \text{padding}_i)$  表示对转换后的卷积核进行零填充以对齐尺寸。通过这一设计,

网络在训练阶段具有较强的表达能力,而在推理阶段可通过重参数化将多个分支融合为一个标准卷积操作,从而显著提升推理效率。

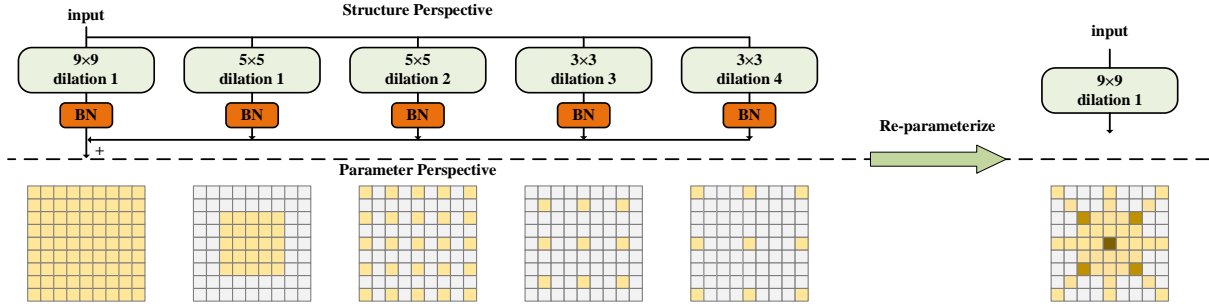


图4 Dilated Reparam Block的网络结构

Fig. 4 The network architecture of Dilated Reparam Block

### 3.3 优化特征提取

为了更有效地提取金属表面缺陷的图像特征,本文受MLWNet<sup>[23]</sup>启发,对Learnable Wavelet Node (LWN) 模块进行改进,提出了集成通道混洗机制的LWN-CS模块。

LWN模块的核心思想是引入可学习的离散小波变换(Discrete Wavelet Transform, DWT),以增强图像去模糊算法对高频细节的建模能力。与传统使用固定基函数的小波变换不同,LWN模块支持端到端训练,使得其小波滤波器能够根据输入图像的分布与特征空间实现自适应学习,进而更加精准地捕捉图像中的多尺度、多方向细节信息。

对于一维离散小波变换(1D-DWT)。对于输入的离散信号  $t$ , 给定小波函数  $\phi_{j,k}(t) = 2^{j/2} \psi(2^j t - k)$ , 其中  $j$  是缩放因子,  $k$  是时间因子, 以及缩放函数  $\phi_{j_0,k}(t) = 2^{j_0/2} \phi(2^{j_0} t - k)$ , 在  $j_0$  对输入信号  $t$  的分解表示为:

$$f(t) = \sum_{j > j_0} \sum_k d_{j,k} \phi_{j,k}(t) + \sum_k c_{j_0,k} \phi_{j_0,k}(t), \quad (2)$$

其中  $d_{j,k} = \langle f, \phi_{j,k} \rangle$  表示高频细节系数,  $c_{j_0,k} = \langle f, \phi_{j_0,k} \rangle$  表示低频近似系数, 从而使输入信号能够逐步分解为丰富的小波域信号。将其应用于DNN时, 引入低频卷积核分析滤波器  $\vec{a}_0[k] = \langle 1/\sqrt{2} \phi(t/2), \phi(t-k) \rangle$  和 高频卷积核分析滤波器  $\vec{a}_1[k] = \langle 1/\sqrt{2} \psi(t/2), \phi(t-k) \rangle$ , 则多尺度分解可写为:

$$\begin{aligned} c_{j+1,p} &= \sum_k \vec{a}_0[k-2p] c_{j,k}, \\ d_{j+1,p} &= \sum_k \vec{a}_1[k-2p] d_{j,k}. \end{aligned} \quad (3)$$

根据上述公式, 原始信号在小波域上的分解可以被视为通过特定滤波器进行递归卷积, 步长为2。类似地, 逆小波变换可通过两个合成滤波器  $s_0[k], s_1[k]$  实现转置卷积重建。通过采用二维离散小波变换(2D-DWT)可有效提取图像中的轴向高频信息, 并设计可学习的方法以适应不同数据分布和特征层次。图5(b)展示了二维小波卷积核的构造方式, 其表达式为:

$$\begin{aligned} \mathcal{F}_{ll} &= \vec{a}_0 \times \vec{a}_0^T, \mathcal{F}_{lh} = \vec{a}_0 \times \vec{a}_1^T, \\ \mathcal{F}_{hl} &= \vec{a}_1 \times \vec{a}_0^T, \mathcal{F}_{hh} = \vec{a}_1 \times \vec{a}_1^T, \\ K_w &= \text{cat}(\mathcal{F}_{ll}, \mathcal{F}_{lh}, \mathcal{F}_{hl}, \mathcal{F}_{hh}), \end{aligned} \quad (4)$$

其中  $a_0$  与  $a_1$  为可学习滤波器;  $F_{ll}, F_{lh}, F_{hl}$  和  $F_{hh}$  是从向量外积中获得的低频、水平高频、垂直高频和对角高频卷积算子; 小波核  $K_w$  由上述卷积算子拼接而成。

如图5(a)所示, 将可学习的小波变换实现为一组卷积操作。输入特征图  $X_{in} \in (C, H, W)$ , 其小波域中的投影  $X_{out} \in (4C, H/2, W/2)$  (包括低频、水平高频、垂直高频和对角高频分量) 通过小波卷积生成。逆卷积核的构造过程与前向传播类似, 两者互为逆过程。

为防止小波核退化为普通卷积, 本文引入完美重建条件作为正则约束。对于复数  $z \in \mathbb{C}$ , 给定应用于  $z$  变换的滤波器  $\vec{x}$ , 其  $z$  变换可以表示为

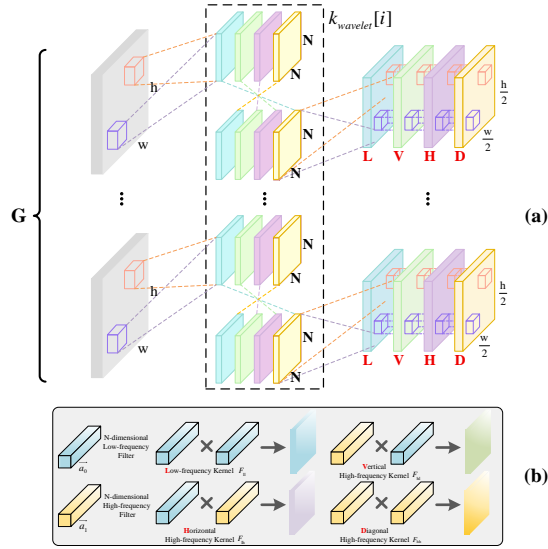


图5 LWN-CS

Fig. 5 LWN-CS

$X(z) = \sum_{n \in \mathbb{Z}} \vec{x}(n) z^{-n}$ 。然后  $\vec{a}_0, \vec{a}_1, \vec{s}_0$  和  $\vec{s}_1$  可以分别表示为其对应  $z$  变换  $A_0, A_1, S_0$  和  $S_1$ , 完美重建需满足:

$$A_0(-z)S_0(z) + A_1(-z)S_1(z) = 0, \quad (5)$$

$$A_0(z)S_0(z) + A_1(z)S_1(z) = 2. \quad (6)$$

上述混叠消除条件用于抑制由下采样引起的混叠效应, 确保频域信息的完整性。LWN 模块由小波分解、深度卷积、 $1 \times 1$  通道调整及逆小波重建构成, 输出尺寸与输入一致。

在此基础上, 本文在多层小波分解中引入通道混洗 (Channel Shuffle), 如图 5(a) 所示。打乱特征通道顺序, 增强特征间的信息流动, 提升特征表达能力, 从而显著改善模型性能与泛化能力。

### 3.4 改进损失函数

在目标检测任务中, 边界框回归的质量直接决定了模型的定位精度。传统的 IoU (Intersection over Union) 及其主流变体主要通过几何距离或纵横比约束来优化边界框, 但往往忽略了样本质量差异对梯度更新的影响。Zhang 等人提出的 Shape-IoU<sup>[24]</sup> 通过引入形状敏感的加权项和指数衰减函数, 有效提升了对细长形变目标的定位能力。

然而, Shape-IoU 对所有回归样本采用统一的优化策略, 导致模型容易被低质量样本产生梯度干扰, 从而限制其整体性能。为此, 本文引

入 Wise-IoU (WIoU) v3<sup>[25]</sup> 的动态非单调聚焦机制, 提出一种新的边界框回归损失函数 Shape-WIoU。

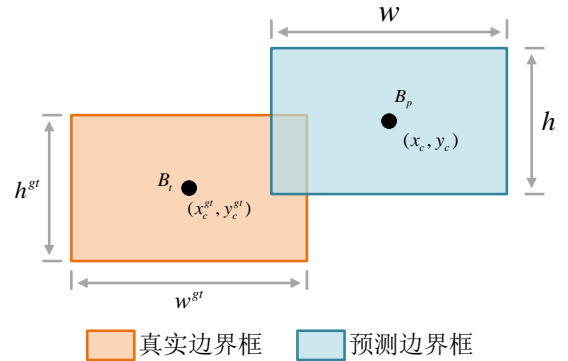


图6 真实框与边界框示意图

Fig. 6 Schematic diagram of the ground truth box and the prediction box

Shape-WIoU 的具体定义如下。首先, 预测边界框为  $B = (x, y, w, h)$ , 真实边界框为  $B_{gt} = (x_{gt}, y_{gt}, w_{gt}, h_{gt})$ 。Shape-IoU 中的形状敏感的加权项  $D_s$  和指数衰减函数  $\Omega$  分别为:

$$D_s = \frac{(x - x_{gt})^2}{c^2} + \frac{(y - y_{gt})^2}{c^2} + \frac{(w - w_{gt})^2}{c_w^2} + \frac{(h - h_{gt})^2}{c_h^2}, \quad (7)$$

$$\Omega = \sum_{t=w, h} (1 - e^{-\omega_t})^\theta, \quad \theta = 4, \quad (8)$$

$$\begin{cases} \omega_w = \frac{|\omega - \omega^{gt}|}{\max(\omega, \omega^{gt})} \\ \omega_h = \frac{|h - h^{gt}|}{\max(h, h^{gt})} \end{cases}, \quad (9)$$

其中  $\omega_w$  和  $\omega_h$  是基于目标宽高的权重系数;  $c$  表示最小外接矩形对角线长度。

WIoU v3 的动态聚焦系数  $r$  定义为:

$$r = \frac{\beta}{\delta \alpha^{\beta - \delta}}, \quad \beta = \frac{L_{IoU}}{\text{mean}(L_{IoU})}, \quad (10)$$

其中  $\alpha = 1.9$ ,  $\delta = 3$ 。  $\beta$  衡量样本异常程度, 中等  $\beta$  时  $r$  最大, 用来聚焦普通样本, 过小或过大则降低梯度权重。最终 Shape-WIoU 损失函数为:

$$L_{\text{Shape-WIoU}} = (1 - IoU + D_s + 0.5 \times \Omega) \times r, \quad (11)$$

其中  $IoU$  为预测框与真实框的标准交并比。该设计在保持 Shape-IoU 对形状与尺度敏感性优势的同时, 引入了 WIoU 的非单调样本加权机制,

$D_s$  和  $\Omega$  显式惩罚中心偏移与宽高差异, 强化对细长目标的定位约束,  $r$  根据样本质量自适应调整梯度幅度, 避免被噪声或极端样本带偏, 整体不引入任何额外可学习参数, 保持模型结构简洁。

## 4 实验与分析

### 4.1 数据集介绍

#### 4.1.1 NEU-DET 数据集

NEU-DET<sup>[26]</sup> 是东北大学收集的热轧带钢表面缺陷数据集, 包含 6 种缺陷类型, 每种缺陷有 300 张图像, 共 1 800 张图像。NEU-DET 数据集缺陷类型示例如图 7 所示, 分别为夹杂物、裂纹、点蚀面、轧制鳞片、划痕和斑块, NEU-DET 数据集为金属表面缺陷检测的研究提供了重要的数据支持。

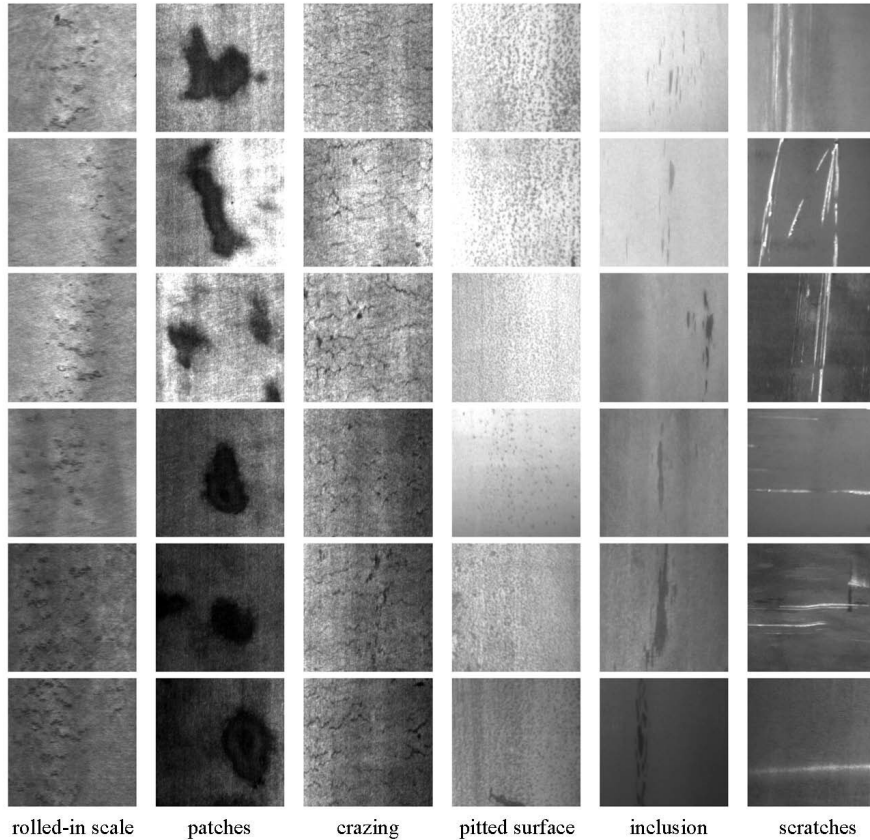


图7 NEU-DET数据集图例<sup>[26]</sup>

Fig. 7 NEU-DET

#### 4.1.2 ICRFD 数据集

本研究使用自建数据集对改进模型进行泛

化验证。原始图像数据采集自某水电站摄像头, 涵盖由 4 个不同视角拍摄的工业滑触线图像, 从

而构建了一个具有高质量和代表性的滑触线脱落检测图像数据集。该数据集包含 1 286 张影像,并按照 0.5:2.0:7.5 的比例划分为测试集、验证集和训练集。图像包括了多种复杂场景,如不同光照条件、遮挡情况以及多目标干扰等,具有较强的挑战性和代表性。所有图像均通过 LabelImg 软件进行了精细标注,标注结果保存为 TXT 格式,可直接用于后续模型训练。

为确保数据集具备更强的泛化能力并提升模型对复杂场景的适应性,本文引入了三种典型的数据增强方法:旋转、模糊和噪声添加。旋转增强模拟了滑触线在不同角度下的成像变化,有效扩展了视角多样性;模糊增强模拟水电站潮湿环境对摄像头成像质量的影响;而噪声添加则模拟了因传感器干扰而产生的图像噪点。这些增强手段通过随机方式对图像进行扰动,显著提升了样本的多样性与模型的鲁棒性。最终,经增强处理后的图像总数扩展至 5 144 张。本文将构建的滑触线数据集命名为 ICRFD,其部分图像如图 8 所示。

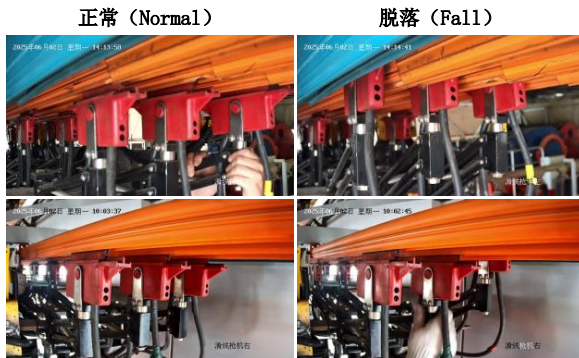


图 8 ICRFD 展示图

Fig. 8 ICRFD display images

#### 4.1.3 FDMPI 数据集

FDMPI 数据集来源于某锻造有限公司重型卡车转向节生产流水线的探伤车间,通过相机拍摄缺陷图片。初始数据集包含 362 张图片,数量不足以满足神经网络训练需求。因此,进行了标注和数据增强,包括旋转、翻转和亮度调整。旋转角度在  $-20^\circ$  到  $20^\circ$  之间,翻转包括上下和左右翻转,比例各为 0.5,以提升模型的检测性能和鲁棒性。亮度调整范围为原始亮度的 0.8 倍至 1.2 倍,以模拟不同光照条件,增强模型适应性。数

据增强后,训练集扩展至 3 206 张图片,包含 11 700 个缺陷目标;验证集包含 1 416 张图片,共 3 792 个缺陷目标,所有目标均为“defect”类别。部分带标签的缺陷数据如图 9 所示。

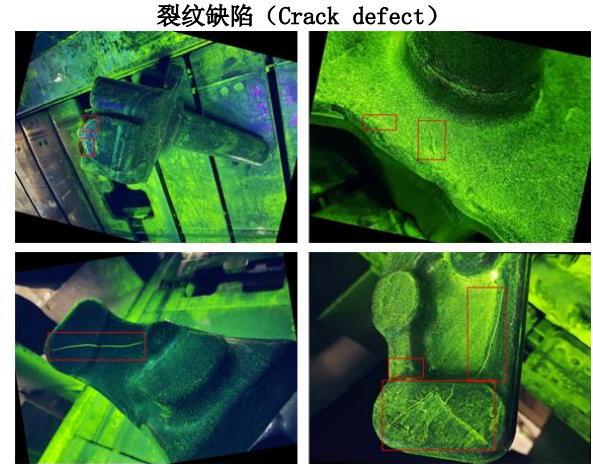


图 9 重卡裂纹数据集内部分图片

Fig. 9 Partial images in the heavy-duty truck crack dataset

#### 4.2 训练环境及参数设置

本文对金属外观缺陷检测的实验使用深度学习 PyTorch 框架,在 Ubuntu24.04 操作系统上运行,硬件配置为:CPU 使用 AMD EPYC 9754, GPU 使用 NVIDIA RTX 3090,24G 内存;软件配置为:Python 版本是 3.11,框架使用 Pytorch2.6.2+CUDA11.8。

训练轮次设置为 500 轮,patience 设置为 50, batch\_size 设置为 16,workers 设置为 4。

#### 4.3 评估指标

在计算机视觉中,评估算法性能的主要指标包括准确率 (Precision, 用  $P$  表示)、召回率 (Recall, 用  $R$  表示),综合指标平均精度均值 (Mean Average Precision, 用  $P_{mAP}$  表示),以及每秒帧数 FPS。式中:  $TP$  表示预测对的正样本、 $FP$  表示预测错的正样本、 $TN$  预测对的负样本和  $FN$  表示预测错的负样本。本文所有 mAP 均指 mAP。具体公式如下:

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

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

$$P_{AP} = \int_0^1 P(R) dR, \quad (14)$$

$$P_{mAP} = \frac{1}{c} \sum_{j=1}^c AP_j, \quad (15)$$

$$FPS = \frac{1000}{preprocess + inference + postprocess}. \quad (16)$$

#### 4.4 实验结果与分析

##### 4.4.1 对比实验

为全面验证所提出ICR-DETR模型在实际目标检测任务中的性能表现,本文选取当前主流的目标检测算法作为对比对象,涵盖轻量级模型与中高精度模型两个类别。所有模型均在同一硬件与数据集条件下进行测试,以确保评估的客观性与一致性。

表1 对比实验结果

Tab. 1 Results of comparative experiment

算法	参数量	计算量	P%	R%	mAP%
YOLOv5m	21.20	48.93	75.17	67.24	73.87
YOLOv8m	25.92	78.97	75.73	70.15	75.83
YOLOv9c	25.30	102.10	75.83	69.26	74.97
RT-DETR	66.10	108.00	72.35	69.12	72.33
RT-DETR-r18	20.10	57.00	70.12	67.70	71.79
文献 <sup>[27]</sup>	3.6	—	79.2	71.7	77.6
文献 <sup>[28]</sup>	—	40.6	—	—	72.9
ICR-DETR	81.50	185.00	77.92	75.71	78.42

实验结果如表1所示,YOLOv5m与YOLOv8m虽然在效率与精度之间取得了较好的平衡,mAP分别达到73.87%和75.83%,但由于卷积感受野的固有限制,其对金属外观缺陷的定位能力仍显不足,召回率分别为67.24%和70.15%。YOLOv9c引入了可编程梯度信息,取得了75.83%的精确率和74.97%的mAP,但其召回率仅为69.26%,导致在实际检测中容易出现漏检目标的问题。

重点对比RT-DETR系列模型与本文ICR-DETR。RT-DETR作为本文的基线模型,凭借端到端的注意力机制取得了72.33%的mAP,但其精确率(72.35%)和召回率(69.12%)均表现平平,计算量达108GFLOPs。其轻量级版本RT-DETR-r18虽将参数量降至20.10MB、计算量降至57GFLOPs,但mAP进一步下滑至71.79%,召回率仅为67.70%,在检测灵敏度方

面做出了明显妥协。

与文献<sup>[27]</sup>相比,其是根据YOLOv8n改进模型,参数很低,仅为3.6MB,mAP为77.6,与本文mAP相比,低了0.82%。与文献<sup>[28]</sup>相比,其与本文一样做泛化类模型,但是将NEU-DET作为泛化实验,其mAP仅为72.9%,说明在泛化类模型中,模型精度提升难度较大。

相比之下,本文提出的ICR-DETR模型在NEU-DET检测任务中显著优于RT-DETR系列。该模型取得了78.42%的mAP,相较于基线RT-DETR(72.33%)大幅提升了6.09个百分点。同时,ICR-DETR实现了最高的精确率(77.92%)和召回率(75.71%),分别超出RT-DETR 5.57和6.59个百分点,有效解决了原模型在召回率偏低、漏检率较高方面的问题。尽管ICR-DETR的参数量为81.5MB、计算量为185GFLOPs,高于RT-DETR,但其在检测精度上的巨大提升充分验证了本文所设计模块的有效性与实用性,满足工业场景下对低漏检率和高置信度的要求。

##### 4.4.2 消融实验

为了验证每一步改动的有效性,对相关创新模块进行消融实验。如表2所示。实验围绕三项主要改进展开,分别为引入UniRepLKNet作为主干网络结构、在特征融合阶段引入LWN-CS模块,以及在损失函数层面引入Shape-WIoU损失。实验通过逐步添加不同模块,构建多个子模型(标记为A至F),并最终在全部引入的基础上形成最终模型(标记为“ICR-DETR”),以系统分析各个模块对检测性能、计算资源及运行效率的影响。

首先,模型A单独引入UniRepLKNet将骨干网络替换为具备大感受野和高精度的UniRepLK结构。该改进增加1.5MB的参数量和24GFLOPs的计算量,使mAP大幅提升3.14个百分点,达到75.47%,表明大卷积核骨干网络对金属表面缺陷的特征提取具有显著增益。模型B仅引入LWN-CS模块显著增强了跨尺度的上下文建模能力。尽管参数量增加12.6MB,计算量上升59GFLOPs,mAP获得4.56个百分点的绝对增益,达到76.89%,验证了该模块在多尺度缺陷检测中的关键作用。模型C单独采用

Shape-WIoU 损失函数在不改变模型结构的前提下,通过动态聚焦机制与形状感知的边界框回归优化,使 mAP 提升 0.79 个百分点,达到 73.12%,验证了其在提升细长线状目标定位精度方面的有效性。

进一步地,两个模块的联合使用可带来协同增益。模型 D 在基线基础上将 mAP 提升 4.88 个百分点,达到 77.21%,计算量增至 150GFLOPs。模型 E 增加 2.6MB 参数量和 27GFLOPs 计算量,mAP 提升 5.73 个百分点至 78.06%,体现出 Shape-WIoU 与强骨干网络的深度融合优势,这也是所有模型中 mAP 最高的组合。而模型 FmAP 达到 77.43%,较基线提升 5.10 个百分点,计算开销最大,表明跨尺度建模与几何感知回归的联合优化对性能提升具有显著贡献。

最终,完整模型 ICR-DETR(集成全部三个组件)在参数量为 81.5 MB、计算量为 185GFLOPs 的条件下,取得 78.42% 的 mAP,较基线提升 6.09 个百分点。值得注意的是,从模型 F 到 ICR-DETR,通过引入 UniRepLKNet 替代原骨干网络,带来 0.99 个百分点的 mAP 增益,同时参数量增加 2.3MB,计算量增加 10GFLOPs,显示出 UniRepLKNet 在高负载模型中仍具有可观的边际效益。

综上所述,三项改进均对检测性能具有正向贡献,其中模型 E 的 UniRepLKNet 与 Shape-WIoU 组合带来最大性能增益,而三者协同构成的 ICR-DETR 在综合性能上达到了最优平衡。这些改进共同构成了 ICR-DETR 高效工业检测能力的核心基础。

表 2 消融实验结果

Tab. 2 Resultsofablationexperiment

模型	改进			参数量/ MB	计算量/ GFLOPs	P%	R%	mAP%
	UniRepLKNet	LWN-CS	Shape-WIoU					
base				66.10	108.00	72.35	69.12	72.33
A	✓			67.6	132	73.18	69.95	75.47
B		✓		78.7	167	75.52	74.36	76.89
C			✓	66.1	108	74.35	69.12	73.12
D	✓	✓		73.5	150	75.64	72.08	77.21
E	✓		✓	68.7	135	73.89	70.43	78.06
F		✓	✓	79.2	175	76.87	75.12	77.43
ICR-DETR	✓	✓	✓	81.5	185	77.92	75.71	78.42

#### 4.4.3 不同损失函数对比实验

为验证本文所采用损失函数组合的有效性,并探究其对金属外观缺陷检测性能的具体贡献,本小节设计了消融实验。实验固定采用 UniRepLKNet 作为骨干网络及完整的检测架构,仅系统性地替换分类与回归任务中的损失函数组件。我们选取了目标检测领域中具有代表性的损失函数进行对比

所有对比模型均在 NEU-DET 数据集上,使用统一的训练策略、优化器及超参数进行端到端训练。由于本实验仅改变损失函数的计算逻辑,而不引入任何新的可学习参数或改变网络结构,因此所有模型的参数量与计算量保持完全一致。评估指标包括精度、召回率和平均精

度均值。

如表 3 所示,在不改变模型架构的前提下,损失函数的设计对最终性能有决定性影响。本文设计的 Shape-WIoU 取得了最优结果。该损失函数极大地提高了定位精度,不仅将 mAP% 推高至 78.42%,同时在精度和召回率上也实现了全面领先。该实验充分证明了本文损失函数选择的科学性与优越性,是实现工业无损检测的关键因素之一。

#### 4.4.4 定性分析

为进一步验证所提出 ICR-DETR 模型在实际应用场景中的有效性与鲁棒性,本文选取基准模型 RT-DETR 和改进模型 ICR-DETR,在相同测试集条件下进行了可视化对比实验,分别加载各自训练得到的最优权重模型,对目标进行检测

表3 不同损失函数对比实验结果

Tab. 3 Comparative experimental results of different loss functions

损失函数	参数量	计算量	P%	R%	mAP%
InnerIoU	81.5	185	73.56	70.87	74.12
FocalerIoU	81.5	185	76.89	72.34	77.23
MPDIoU	81.5	185	73.98	71.23	74.56
Inner-MPDIoU	81.5	185	75.12	72.67	75.67
Focaler-MPDIoU	81.5	185	75.45	73.12	76.01
Inner-ShapeIOU	81.5	185	74.67	72.89	75.34
Shape-WIoU	81.5	185	77.92	75.71	78.42

与可视化展示。

实验结果如图10所示。ICR-DETR模型的预测置信度明显高于RT-DETR,检测边界更加清晰、贴合度更高。在复杂背景环境下,RT-DE-

TR常出现漏检与误检现象,尤其在多目标区域时,模型稳定性较差。而ICR-DETR则展现了出色的鲁棒性,能够准确识别出图像中的所有目标,且几乎未出现误判情况。

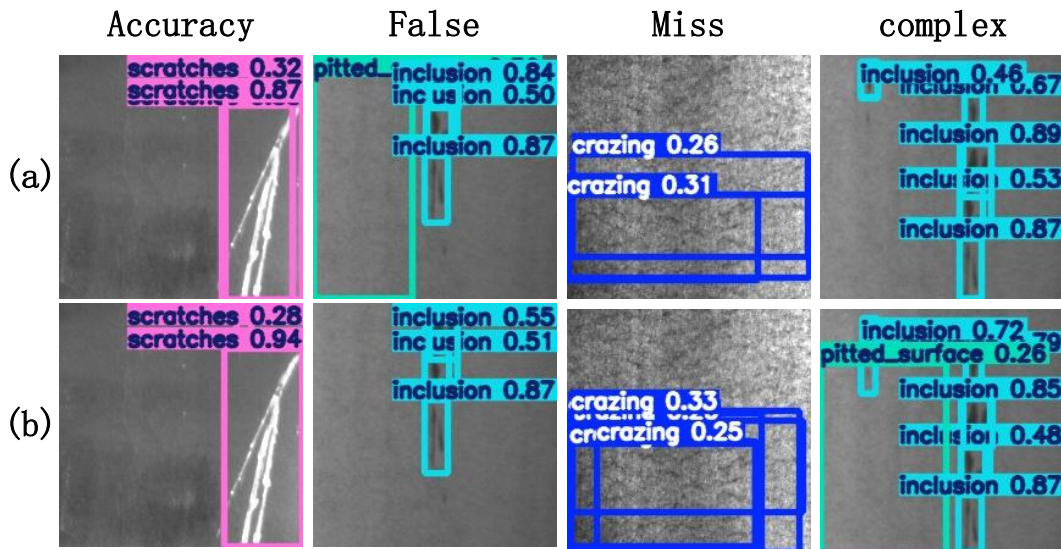


图10 NEU-DET定性分析结果(a) RT-DETR,(b) ICR-DETR。

Fig. 10 Qualitative analysis result of NEU-DET (a) RT-DETR and (b) ICR-DETR

#### 4.4.5 泛化实验

为验证所提方法在未见工业场景下的泛化能力,我们在自建滑触线脱落泛化数据集ICRFD以及重卡转向节裂纹数据集FDMPI上开展跨条件评估,实验结果如下表所示。

实验结果表明,本文提出的方法在性能上显著优于当前主流的目标检测器。在ICRFD数据集上,该方法的平均精度均值(mAP)达到了98.43%,相较于性能次优的YOLOv9c模型(94.86%)提升了3.57%,相较于RT-DETR-r18模型(94.72%)提升了3.71%。尤为突出的是,

ICR-DETR同时实现了97.64%的精确率和94.21%的召回率,分别比YOLOv9c高出6.81和8.34%,充分证明该方法在复杂工况下依然能够保持高置信度的检测能力并实现低漏检率。

ICR-DETR在FDMPI数据集上同样展现了卓越的跨域适应能力。尽管所有模型在迁移至FDMPI后性能均有不同程度的下降,但ICR-DETR仍以87.61%的mAP%稳居榜首,领先于次优模型RT-DETR-r18(86.15%)1.46%。更为关键的是,面对分布差异带来的挑战,ICR-DE-

表 4 泛化实验结果

Tab. 4 Results of generalization experiment

算法	参数量	计算量	ICRFD			FDMPI		
			P%	R%	mAP%	P%	R%	mAP%
YOLOv5m	21.20	48.93	89.37	87.68	92.16	85.66	79.42	85.73
YOLOv8m	25.92	78.97	87.59	84.32	91.64	84.75	78.32	83.8
YOLOv9c	25.30	102.10	90.83	85.87	94.86	86.45	80.16	85.24
RT-DETR	66.10	108.00	89.53	86.32	93.23	85.94	79.54	84.6
RT-DETR-r18	20.10	57.00	91.57	85.24	94.72	87.12	81.35	86.15
ICR-DETR	81.50	185.00	97.64	94.21	98.43	88.25	86.96	87.61

TR保持了极高的召回率(86.96%),显著优于YOLOv5m(79.42%)和YOLOv8m(78.32%),差距分别达到7.54和8.64%。这一结果有力地证明了本文方法并未过拟合于源数据集,而是学习到了更具鲁棒性的特征表示,从而在未知场景中有效抑制了漏检现象。

值得注意的是,尽管YOLOv9c和RT-DETR-r18在参数量较低(分别为25.30M和20.10M)的情况下取得了接近95%的mAP,但两者的召回率在泛化场景中均表现不足(FDMPI上低于82%),表明它们对域外目标的敏感性仍有欠缺。相比之下,本文方法虽然计算开销较高,达到185.00GFLOPs,但其带来的性能增益远超计算成本的增加。例如,与计算量为108.00GFLOPs的RT-DETR模型相比,ICR-DETR在计算量增加71.3%的情况下,却在ICRFD上带来了5.20%的mAP增益,并在FDMPI上保持了最大的性能领先优势。这充分体现出本文所设计的UniRepLKNet、LWN-CS和S-MIoU等模块在提升特征鲁棒性与几何感知能力方面的显著泛化优势,使其成为对检测可靠性要求极高场景下的

理想选择。

#### 4.4.6 泛化实验定性分析

为验证所提ICR-DETR模型在泛化场景下的有效性与鲁棒性,本文在相同测试集上对基准模型RT-DETR与改进模型ICR-DETR进行了可视化对比实验,分别加载各自训练所得最优权重进行推理。

结果如图11所示:在单目标检测中,ICR-DETR的预测置信度更高,边界框贴合更紧密;在复杂背景、遮挡或低对比度条件下,RT-DETR易出现漏检与误检,而ICR-DETR仍能稳定、准确地识别所有目标,表现出更强的鲁棒性。此外,在多目标、多尺度及复杂背景场景下,ICR-DETR始终保持高检测精度,说明其改进结构在特征表达、边界定位与上下文建模方面具有显著优势;相比之下,RT-DETR在边缘模糊或目标密集区域的检测性能明显受限。

图12展示了模型在锻件数据集上的泛化表现。面对域外场景中光照不均与纹理差异的挑战,RT-DETR出现了明显的微小缺陷漏检及背景误报,定位框也较为松散;而ICR-DETR凭借

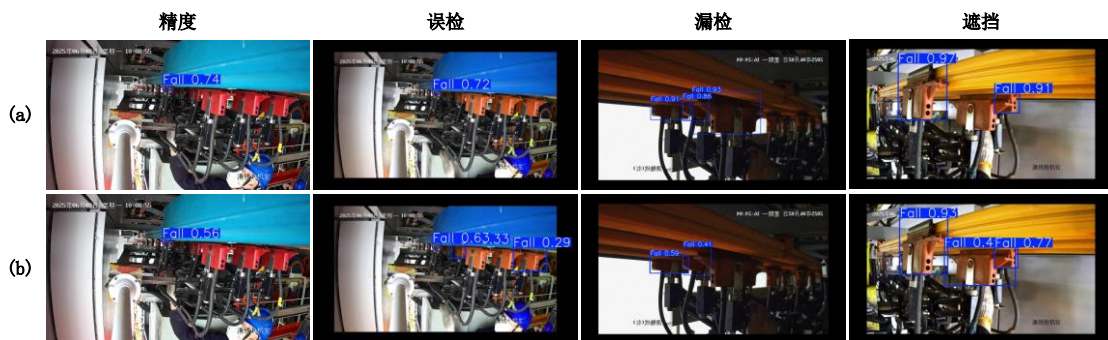


图 11 ICRFD 定性分析结果(a) ICR-DETR,(b) RT-DETR。

Fig. 11 Qualitative analysis result of ICRFD (a) ICR-DETR and RT-DETR

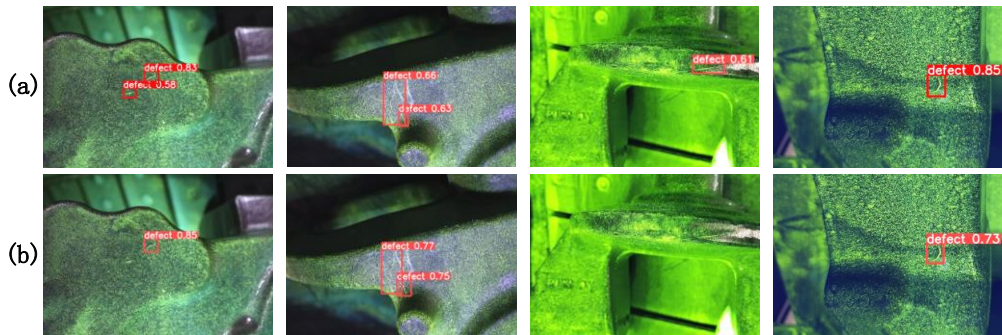


图12 锻件数据集定性分析结果(a) ICR-DETR, (b) RT-DETR。

Fig. 12 Qualitative analysis result of FDMPI (a) ICR-DETR and RT-DETR

更强的特征鲁棒性,不仅精准捕捉到所有细微目标,且边界回归更加严密,有效抑制了跨域干扰。这进一步证实了本文方法未过拟合,而是学习到了通用的本质特征,具备优异的跨场景适应能力。

因此ICR-DETR不仅在量化评估指标上显著优于对比模型,在ICRFD数据集的可视化检测效果同样表现优异,充分满足工业滑触线脱落检测对高精度的实际需求。

## 5 结 论

本文针对金属外观缺陷检测任务中存在的检测精度低,误检漏检率高等问题,提出了一种基于RT-DETR改进的目标检测算法ICR-DETR。首先,引入UniRepLKNet作为骨干网络,利用其大感受野与多尺度特征融合能力,有效增强了对滑触线边缘和结构特征的感知;其次,在颈部网络中集成LWN-CS模块,通过可学习小波

变换结合通道混洗机制,提升了模型在模糊与复杂背景中的高频细节建模能力;最后,设计Shape-WIoU损失函数,将马氏距离引入边界框相似性度量中,精准建模多维特征的尺度与相关性,从而提升边界定位的精度与鲁棒性。

实验在公开的NEU-DET数据集上开展,ICR-DETR在准确率、召回率和mAP等指标上均显著优于RT-DETR及多种主流检测模型,验证了所提方法的有效性与先进性。此外,模型在自建数据集ICRFD上亦取得了良好的检测性能,展示出较强的跨场景泛化能力。综上所述,本文以NEU-DET数据集上的多类别钢材表面缺陷检测为主任务,所提ICR-DETR的各项改进均围绕提升该任务性能展开;在自建数据集ICRFD和FDMPI上验证模型在跨场景、跨任务下的泛化能力。ICR-DETR不仅在精度、稳定性方面达到工业级水平,也为工业无损智能化检测提供了可推广的解决方案,具有良好的工程应用前景与研究价值。

## 参 考 文 献:

- [1] HAN G R, ERYILMAZ M, GONCHAROV A, *et al.* Deep learning-enhanced dual-mode multiplexed optical sensor for point-of-care diagnostics of cardiovascular diseases[J]. *Light: Science & Applications*, 2026, 15(1): 190. DOI: 10.1038/s41377-026-02275-9.
- [2] JALALI B, MACPHEE C. VEVid: Vision Enhancement via Virtual diffraction and coherent Detection[J]. *eLight*, 2022, 2(1): 24. DOI: 10.1186/s43593-022-00034-y.
- [3] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks[J/OL]. *Communications of the ACM*, 2017, 60(6): 84-90. DOI: 10.1145/3065386.
- [4] GIRSHICK R, DONAHUE J, DARRELL T, *et al.* Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation[C/OL]//2014 IEEE Conference on Computer Vision and Pattern Recognition. Columbus, OH, USA: IEEE, 2014: 580-587. <http://ieeexplore.ieee.org/document/6909475/>. DOI: 10.1109/

- CVPR. 2014. 81.
- [5] GIRSHICK R. Fast R-CNN [C/OL]//2015 IEEE International Conference on Computer Vision (ICCV). Santiago, Chile: IEEE, 2015: 1440-1448. <http://ieeexplore.ieee.org/document/7410526/>. DOI: 10.1109/ICCV.2015.169.
- [6] REN S, HE K, GIRSHICK R, *et al.* Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks[J/OL]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 39(6):1137-1149. DOI: 10.1109/TPAMI.2016.2577031.
- [7] REDMON J, DIVVALA S, GIRSHICK R, *et al.* You Only Look Once: Unified, Real-Time Object Detection[C/OL]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, NV, USA: IEEE, 2016:779-788. <http://ieeexplore.ieee.org/document/7780460/>. DOI:10.1109/CVPR.2016.91.
- [8] LIU W, ANGUELOV D, ERHAN D, *et al.* SSD: Single Shot MultiBox Detector[M/OL]//LEIBE B, MATAS J, SEBE N, *et al.* Computer Vision - ECCV 2016: v. 9905. Cham: Springer International Publishing, 2016: 21-37. [http://link.springer.com/10.1007/978-3-319-46448-0\\_2](http://link.springer.com/10.1007/978-3-319-46448-0_2). DOI:10.1007/978-3-319-46448-0\_2.
- [9] CARION N, MASSA F, SYNNAEVE G, *et al.* End-to-End Object Detection with Transformers [M/OL]//VEDALDI A, BISCHOF H, BROX T, *et al.* Computer Vision - ECCV 2020: v. 12346. Cham: Springer International Publishing, 2020: 213-229. [https://link.springer.com/10.1007/978-3-030-58452-8\\_13](https://link.springer.com/10.1007/978-3-030-58452-8_13). DOI: 10.1007/978-3-030-58452-8\_13.
- [10] DOSOVITSKIY A, BEYER L, KOLESNIKOV A, *et al.* An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale[PP/OL]. V2. *arXiv(2020)[2025-06-24]*. <https://arxiv.org/abs/2010.11929>. DOI: 10.48550/ARXIV.2010.11929.
- [11] ZHAO Y, LV W, XU S, *et al.* DETRs Beat YOLOs on Real-time Object Detection[C/OL]//2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle, WA, USA: IEEE, 2024: 16965-16974. <https://ieeexplore.ieee.org/document/10657220/>. DOI:10.1109/CVPR52733.2024.01605.
- [12] LECUN Y, BOTTOU L, BENGIO Y, *et al.* Gradient-based learning applied to document recognition[J/OL]. *Proceedings of the IEEE*, 1998, 86(11):2278-2324. DOI:10.1109/5.726791.
- [13] 林珊玲, 彭雪玲, 王栋, 等. 多尺度增强特征融合的钢表面缺陷目标检测[J]. *光学精密工程*, 2024, 32(07):1075-1086.
- LIN Shanling, PENG Xueling, WANG Dong, *et al.* Object detection of steel surface defect based on multi-scale enhanced feature fusion[J]. *Optics and Precision Engineering*, 2024, 32(07):1075-1086.
- [14] 梁礼明, 龙鹏威, 冯耀, 等. 改进轻量化 VTG-YOLOv7-tiny 的钢材表面缺陷检测[J]. *光学精密工程*, 2024, 32(08):1227-1240.
- LIANG Liming, LONG Pengwei, FENG Yao, *et al.* Improving the lightweight VTG-YOLOv7-tiny for steel surface defect detection[J]. *Optics and Precision Engineering*, 2024, 32(08):1227-1240.
- [15] 乔健, 陈能达, 伍雁雄, 等. 融合注意力机制的金属锅圆柱表面缺陷检测[J]. *光学精密工程*, 2023, 31(03):404-416.
- QIAO Jian, CHEN Nengda, WU Yanxiong, *et al.* Defect detection of cylindrical surface of metal pot combining attention mechanism[J]. *Optics and Precision Engineering*, 2023, 31(03):404-416.
- [16] 杨彦萍, 高军伟, 刘兆龙, 等. 基于改进 YOLO11 的印刷电路板表面缺陷检测算法[J]. *液晶与显示*, 2025, 40(10):1545-1556.
- YANG Yanping, GAO Junwei, LIU Zhaolong, *et al.* Surface defect detection on printed circuit boards based on improved YOLO11[J]. *Chinese Journal of Liquid Crystals and Displays*, 2025, 40(10): 1545-1556.
- [17] 陈顺龙, 廖映华, 林峰, 等. 基于改进 YOLOv8n 的液晶屏 Mura 缺陷检测[J]. *液晶与显示*, 2025, 40(03):439-447.
- CHEN Shunlong, LIAO Yinghua, LIN Feng, *et al.* Mura defect detection of LCD screen based on improved YOLOv8n[J]. *Chinese Journal of Liquid Crystals and Displays*, 2025, 40(3): 439-447.
- [18] ZHANG Y, WANG W, LI Z, *et al.* Development of a cross-scale weighted feature fusion network for hot-rolled steel surface defect detection[J/OL]. *Engineering Applications of Artificial Intelligence*, 2023, 117: 105628. DOI: 10.1016/j.engappai.2022.105628.

- [19] LIU R, HUANG M, GAO Z, *et al.* MSC-DNet: An efficient detector with multi-scale context for defect detection on strip steel surface[J/OL]. *Measurement*, 2023, 209: 112467. DOI:10.1016/j.measurement.2023.112467.
- [20] ZENG Q, WEI D, ZOU M. Rtsds: a real-time and efficient method for detecting surface defects in strip steel[J/OL]. *Journal of Real-Time Image Processing*, 2024, 21(4): 117. DOI:10.1007/s11554-024-01497-7.
- [21] DING X, ZHANG Y, GE Y, *et al.* UniRepLKNet: A Universal Perception Large-Kernel ConvNet for Audio, Video, Point Cloud, Time-Series and Image Recognition[C/OL]//2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle, WA, USA: IEEE, 2024: 5513-5524. <https://ieeexplore.ieee.org/document/10655060/>. DOI:10.1109/CVPR52733.2024.00527.
- [22] DING X, ZHANG X, HAN J, *et al.* Scaling Up Your Kernels to  $31 \times 31$ : Revisiting Large Kernel Design in CNNs[C/OL]//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). New Orleans, LA, USA: IEEE, 2022: 11953-11965. <https://ieeexplore.ieee.org/document/9880273/>. DOI: 10.1109/CVPR52688.2022.01166.
- [23] GAO X, QIU T, ZHANG X, *et al.* Efficient Multi-scale Network with Learnable Discrete Wavelet Transform for Blind Motion Deblurring[PP/OL]. *arXiv* (2024-03-13) [2025-06-30]. <http://arxiv.org/abs/2401.00027>. DOI: 10.48550/arXiv.2401.00027.
- [24] ZHANG H, ZHANG S. Shape-IoU: More Accurate Metric considering Bounding Box Shape and Scale[PP/OL]. V2. *arXiv* (2023) [2025-07-07]. <https://arxiv.org/abs/2312.17663>. DOI:10.48550/ARXIV.2312.17663.
- [25] TONG Z, CHEN Y, XU Z, *et al.* Wise-IoU: Bounding Box Regression Loss with Dynamic Focusing Mechanism[PP/OL]. V3. *arXiv* (2023) [2026-04-09]. <https://arxiv.org/abs/2301.10051>. DOI:10.48550/ARXIV.2301.10051.
- [26] SONG K, YAN Y. A noise robust method based on completed local binary patterns for hot-rolled steel strip surface defects[J/OL]. *Applied Surface Science*, 2013, 285: 858-864. DOI:10.1016/j.apsusc.2013.09.002.
- [27] 游紫绒, 梁志剑. 基于YOLOv8-CEW的钢材表面缺陷检测方法研究[J]. *电子测量技术*: 1-16.  
YouZiLong, LiangZhijian. Research on steel surface defect detection method based on YOLOv8-CEW [J]. *Electronic Measurement Technology*: 1-16.
- [28] 周得辉, 赵军, 程进峰. 基于RT-DETR的轴承表面微小缺陷检测算法[J]. *计算机应用*, 2025, 45(06): 1987-1997.  
ZHOUDehui, ZHAOJun, CHENGJinfeng. Tiny defect detection algorithm for bearing surface based on RT-DETR [J]. *Journal of Computer Applications*, 2025, 45(6): 1987-1997.

#### 作者简介:



陈俊,男,学士,高级工程师,2009年毕业于武汉大学,主要研究方向为水电站起重金结设备运维管理、嵌入式系统。E-mail: chen\_jun6@ctg.com.cn



陈鹏,男,硕士,2023年毕业于湖北工业大学,主要研究方向为嵌入式系统、电气工程、计算机视觉。E-mail: zhb0213@126.com