

基于改进 RT-DETR 的连接器表面缺陷检测算法

邵 铖¹, 殷亚升², 任红光², 张 彪¹, 包建东¹

(1. 南京理工大学 自动化学院, 南京 210094;

2. 海固科技(江苏)有限公司, 江苏 常熟 215555)

摘要: 针对现有连接器外观缺陷检测方法存在多尺度缺陷特征提取能力不足、模型复杂度偏高等问题, 提出一种基于改进 RT-DETR 的连接器表面缺陷检测算法; 为了降低模型参数量, 在主干网络引入 FEBlock 模块; 使用 HS-FPN 网络替换 CCFM 模块进行多尺度特征融合, 增强模型对细微特征的表达能力; 引入 Inner-SIoU 损失函数替换原损失函数, 提高小目标检测精度并强化模型的泛化能力; 在自建连接器缺陷数据集上进行检测试验, 改进后的算法平均检测精度可达 92.5%, FPS 达到 84, 性能指标均优于基准模型和主流的 YOLO 系列模型, 实现了高精度与实时性的平衡, 能够满足工业环境下连接器缺陷实时检测的需求;

关键词: RT-DETR; 连接器外观缺陷; FasterNet; 多尺度特征融合; Inner-SIoU

Defection Algorithm for Connector Surface Defects Based on Improved RT-DETR

SHAO Cheng¹, YIN Yasheng², REN Hongguang², ZHANG Biao¹, BAO Jiandong¹

(1. School of Automation, Nanjing University of Science and Technology, Nanjing 210094, China;

2. Haigu Technology (Jiangsu) Co., Ltd., Changshu 215555, China)

Abstract: Existing connector defect detection methods have the shortages of, insufficient multi-scale defect feature extraction capability and relatively high model complexity. To address this issue, a connector surface defect detection method based on improved real-time detection Transformer (RT-DETR) is proposed. To reduce the number of algorithm parameters, the FEBlock module is introduced in the backbone network. The HS-FPN network is used to replace the CCFM module for multi-scale feature fusion, thereby enhancing the model's ability to express subtle features. The original loss function is replaced by the Inner-SIoU loss function, which enhances detection accuracy for small targets while increasing the model's generalization capability. Experimental results on a self-built connector defect dataset show that the improved algorithm achieves an average detection precision of 92.4%, with the FPS of 84. The proposed method outperforms the baseline model and mainstream YOLO series models in all performance metrics, achieving a balance between high precision and real-time performance, thus meeting the requirements for industrial connector defect detection.

Keywords: RT-DETR; connector surface defect; multi-scale feature fusion; Inner-SIoU

0 引言

随着《中国制造 2025》和“十四五”智能制造发展规划的深入推进, 中国制造业逐步向智能化、高端化

转型, 其中工业产品质量检测技术的突破是实现智能制造的关键环节^[1]。连接器作为电子元器件的关键组成部分之一, 在大批量的生产过程中, 受加工材料和工艺等因素影响, 其注塑外壳表面极易出现凹痕、脏污和毛边

收稿日期:2025-06-10; 修回日期:2025-08-01。

作者简介:邵 铖(2001-), 男, 硕士研究生。

通讯作者:包建东(1979-), 男, 博士, 副教授。

引用格式:邵 铖, 殷亚升, 任红光, 等. 基于改进 RT-DETR 的连接器表面缺陷检测算法[J]. 计算机测量与控制, 2026, 34(5): 86-93.

等质量缺陷,同时其金属插针引脚在装配时也会存在缺失和偏移等情况,从而导致连接器机械强度降低或电气性能受损^[2]。由于连接器缺陷尺寸小,且种类繁多,传统的人工检测方法存在大量的漏检和误检等问题。因此,寻找一种快速且高效的连接器质量检测方法具有极其重大的意义。

目前,主流的实时目标检测算法可以分为 YOLO 系列算法和基于 Transformer^[3] 架构的 RT-DETR^[4] 算法。YOLO 系列算法检测速度快,但需要阈值筛选和非极大值抑制处理。而 RT-DETR 去除了这些后处理步骤,但本身参数量较大,模型不够轻量化。两者在复杂小目标检测上,特征提取困难,容易降低模型的准确率^[5]。

为了解决在小目标缺陷检测领域的难题,众多学者开展相关研究,并取得了一定的研究成果。徐仟祥等^[6]针对车门内拉手表面缺陷目标小、多尺度的问题,其在主干网络中采用并行的膨胀卷积与 CA 注意力机制,增加网络感受野的同时提高网络推理速度。徐凯等^[7]提出了一种基于 YOLOv8 的核工程焊缝缺陷检测方法,其在模型中加入了多尺度特征融合 MSF 模块,同时引入了可变形卷积思想,有效提高了模型在缺陷细节和整体语义特征提取上的能力。董润华^[8]等分别在主干网络和 AIFI 模块引入部分卷积 (PConv, partial conv) 和 Hi-Lo (high-low frequency) 注意力机制,在不影响模型检测精度的同时有效降低了算法的参数数量和计算量。刘国圀^[9]等引入了 CAFMFusion (CAFm-based Fusion Mechanism) 作为新的特征融合策略,提高了模型对不同尺寸目标的检测能力。Wang^[10]等提出了一种基于 yolov7 的木材缺陷检测方法,其引入了 ODCA 模块,显著提升了木材裂纹、虫洞等微小缺陷的特征表达能力。

上述学者使用了不同方法在小目标缺陷检测领域取得了不错的效果,但仍存在对小目标识别检测精度较低、模型参数量较大等问题,小目标缺陷检测的研究仍有进一步探索的空间。针对上述问题,本文提出了一种基于改进 RT-DETR 的连接器表面缺陷检测方法,并命名为 FEHS-DETR。通过实验对比其它的主流算法,改进后的算法能够实现多尺度缺陷下实时性与检测精度的高效平衡,满足工业场景下连接器实时质量检测的需求。本文的主要工作如下:

- 1) 在主干网络中引入 FEBlock 模块, 降低模型参数量与计算量的同时强化对微小缺陷的特征提取能力。
- 2) 在颈部网络中, 使用 HS-FPN 多尺度特征融合网络取代 CCFM 结构, 有效提高模型对多尺度目标的检测能力。
- 3) 使用 Inner-SIoU 损失函数替代 GIoU 损失函数, 提高其在复杂背景下模型的泛化能力与收敛速度。

1 RT-DETR 模型

RT-DETR 是一种实时端到端的目标检测网络, 是对 DETR 模型的进一步发展。针对 DETR 模型难以收敛和计算冗余等问题, 通过多项创新设计实现了高效训练与实时检测的统一。RT-DETR 主要分为 3 个部分, 包括主干网络 (Backbone)、编码器 (Encoder) 和解码器 (Decoder), 其网络结构如图 1 所示。RT-DETR 选用了不同深度的 Resnet^[11] 网络和 HGNetV2 网络并做了相关改进, 有效提高了特征提取能力。因此本文选择 RT-DETR-r18 作为基准模型以达到轻量化与精度的平衡。同时, 针对 DETR 算法编码器的高计算成本和小目标检测精度低的问题, RT-DETR 设计了一种高效混合编码器, 主要由基于注意力机制的尺度内特征交互模块 (AIFI) 和跨尺度特征融合模块 (CCFM) 组成。AIFI 只处理高层特征, 高层特征包含更加丰富的语义信息, 可以更好地捕捉对象间的关系。CCFM 则类似传统 PAFPN^[12] 结构, 通过跨尺度融合以高效处理多尺度特征, 提高多尺度目标的检测能力。最后, 通过 IOU 感知查询和具有辅助预测头的解码器, 实现无后处理的输出, 有效提高检测效率。

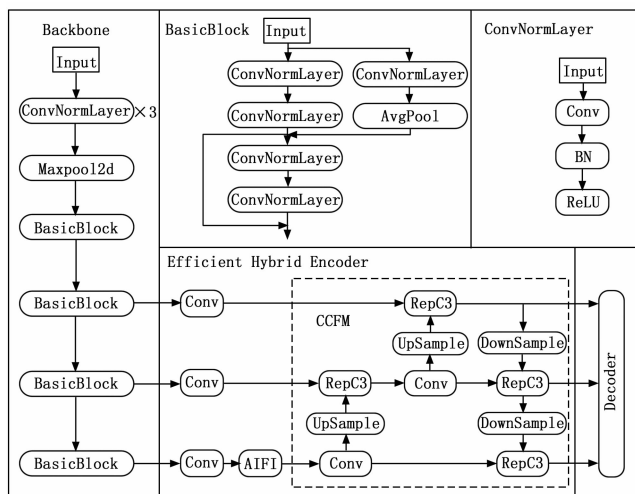


图 1 RT-DETR-r18 网络结构图

2 FEHS-DETR 算法

为了满足工业环境下对连接器小目标缺陷高精度和实时性的要求,本文以 RT-DETR-r18 为基准模型进行改进,并将其命名为 FEHS-DETR,改进后模型结构如图 2 所示。

2.1 FEBlock 模块

在原始 RT-DETR 网络中, 采用 Resnet-18 作为主干特征提取网络, 然而这种网络结构不仅容易在训练时产生不必要的计算和内存访问^[13], 而且在多尺度检测任务中效果也不理想。因此, 本文在原始 BasicBlock 中

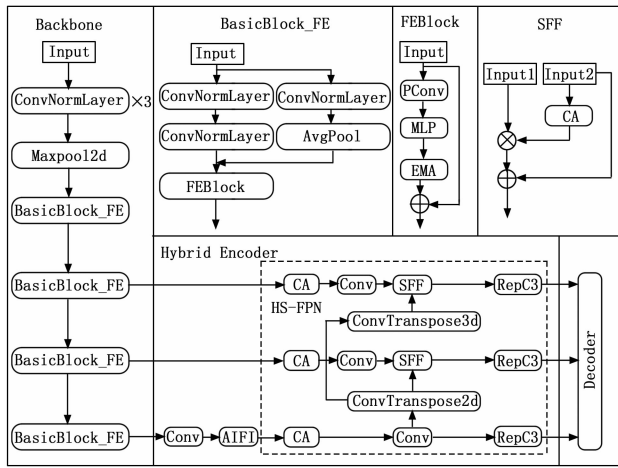


图2 FEHS-DETR 网络结构图

引入 FasterBlock^[14]，并且选择与 FasterNet-T0 相同的主干网络输入通道（即将原始的 64 输入通道降低至 40），最后对其进行改进并融合了 EMA 注意力机制，在降低模型参数数量的同时提高模型对多尺度目标的特征表达能力。

FasterNet 特征提取网络的高效性主要得益于其设计了一种全新的 PConv 卷积模块，通过模拟人类视觉系统的注意力聚焦机制，仅对部分关键通道进行深度处理，实现了感受野扩展与计算效率的平衡。当卷积率 $r=1/4$ 时，其计算量仅为常规卷积的 $1/16$ ，有效减少了计算冗余和内存访问。在每个基础模块 FasterNet Block 中，使用 1 个 PConv 卷积以降低计算量，随后使用 2 个 1×1 的卷积捕捉更复杂的空间特征，增强模型表达能力，如图 3 所示。

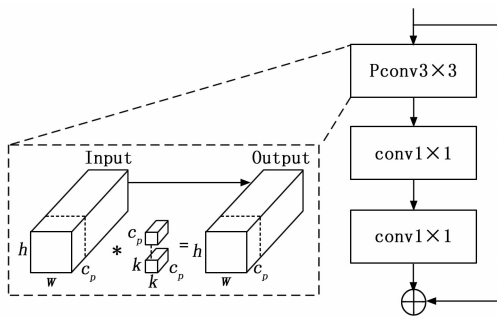


图3 Fasternet Block 与 PConv 模块结构图

由于工厂环境光线和复杂背景等因素的干扰，模型对缺陷的判定需要依赖更多的局部信息，同时连接器外观缺陷的多尺度也对模型的特征提取能力有着更高的要求。因此，在 FasterNetBlock 中引入 EMA 注意力机制^[15]，在保留 FasterNet 网络快速高效的同时，弥补其对多尺度目标特征提取能力不足的缺点。EMA 机制的核心思想是通过指数衰减的加权平均，动态调整特征图

中不同位置或通道的注意力权重，使模型更加关注关键信息并减少噪声干扰^[16]，具体模型结构如图 4 所示。首先将输入特征图 X 沿通道维度 c 划分为 G 组子特征图，即 $Groups = [x_0, x_1, \dots, x_{G-1}]$ ，且每个子特征图的维度为 $R^{(c/G) \times h \times w}$ ，然后并行使用水平、垂直平均池化与 3×3 卷积三条路径来提取分组特征图的注意力权重，前两条路径经过 concat 与 1×1 卷积后，能够聚焦于小目标的特征信息，而 3×3 的卷积分支能够捕获更大感受野的多尺度信息，从而完成多尺度特征提取。最后三条路径的输出通过跨分支特征交互融合，实现多尺度特征的动态加权与互补。

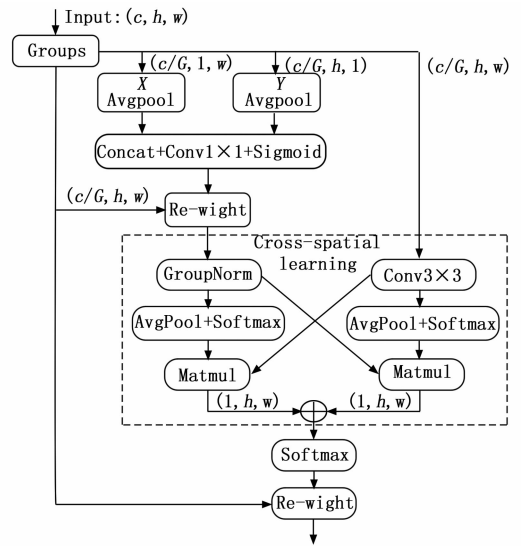


图4 EMA 注意力机制模块结构图

融合后的模块 FEBlock 模块结构如图 5 所示，该模块首先通过 PConv 从输入图像中提取多尺度特征空间，并保持相对较低的计算开销。然后借助 EMA 注意力机制对卷积输出进行加权，自动调整对于不同尺度信息的权重，从而在降低模型参数数量的同时，提高连接器微小缺陷特征的提取能力。

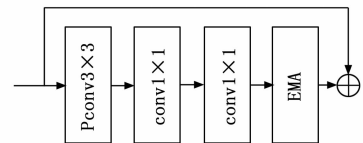


图5 FEBlock 模块结构图

2.2 HS-FPN 多尺度特征融合网络

在 RT-DETR 网络中，Neck 部分的原始跨尺度特征融合模块 (CCFM) 虽然通过整合不同尺度的特征提升了多尺度目标检测能力，但其由于采用固定方式 (Hadamard 乘法或像素级求和) 融合相邻的层级特征，缺乏动态的权重调整机制，导致对复杂场景的适应性不

足^[17]。同时,由于未引入注意力筛选机制,低级特征中的噪声可能干扰融合过程,降低了对细微特征的敏感度。

为解决以上问题,本文利用 HS-FPN^[18]对 CCFM 模块进行优化,以实现更加高效的多尺度特征融合,其网络结构如图 6 所示。Backbone 部分中 S3、S4、S5 代表 3 个不同尺度的特征图,首先由特征选择模块进行动态特征筛选,通过 CA (Channel Attention) 通道注意力机制生成权重,以高级语义特征为引导过滤低级特征的噪声,保留关键信息,并通过逐点相加实现语义增强。筛选过后利用选择性特征融合机制并结合空间特征筛选 (SFF),自适应融合不同层级的定位精度 (低级特征) 与语义信息,有助于检测连接器缺陷的细微特征,从而提高模型的多尺度检测能力。

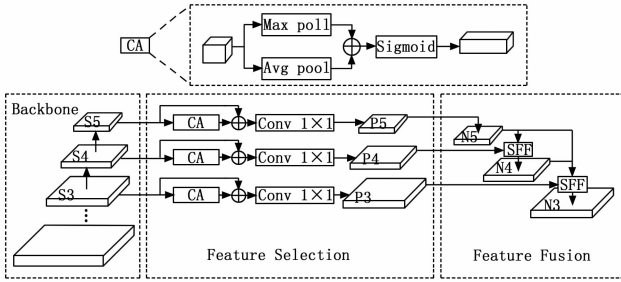


图 6 HS-FPN 模块结构图

在特征选择模块中,不同尺度的特征图首先通过 CA 注意力模块进行特征加权处理,然后经过 1×1 的卷积进行降维处理,最后得到新的特征图 P_3 、 P_4 、 P_5 。该模块的核心是 CA 坐标注意力机制,其通过全局平均池化与全局最大池化的组合同时捕获特征图的通道关系与空间位置信息,从而显著提升模型在复杂场景下的特征定位与小目标检测能力。

传统的特征融合模块直接将不同尺度的特征图经过上、下采样后的结果相加,这种分层堆叠的融合方式容易导致特征冗余,甚至同一目标的特征发生偏移,难以有效调和两者的互补性。为了解决该问题,HS-FPN 特征融合模块中创新地加入了选择性特征融合 (SFF) 模块,利用高级特征的语义信息筛选低级特征的细节信息,从而避免传统特征融合中因简单相加导致的信息冗余或丢失问题,模块结构如图 7 所示。该模块首先利用转置卷积与双线性插值方法调整高层特征的维度,完成高层与低层特征维度的统一。然后通过 CA 注意力模块将高层特征转换为相应的注意力权重,用于筛选低级特征中的关键区域。最后,加权后的低级特征与调整后的高级特征进行特征融合,这一过程保留了语义信息与细节信息的最佳平衡,从而有效提高模型对连接器多尺度缺陷的识别能力。

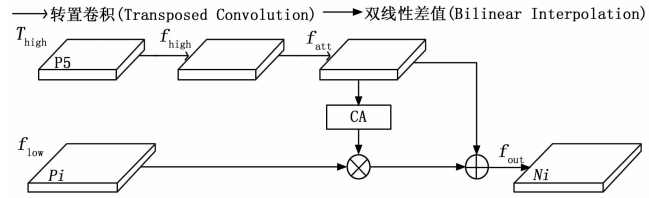


图 7 SFF 结构示意图

2.3 Inner-SIoU 损失函数

在目标检测任务中,边界框回归 (BBR, bounding box regression) 损失函数是模型训练的核心组件之一,将直接影响模型的最终表现和泛化能力。RT-DETR 中采用了 GIOU^[19]作为边界损失函数,它在传统 IoU 的基础上引入了最小外接矩形和差集区域的概念,解决了当预测框与真实框完全无重叠时, IoU 无法有效优化的问题。然而,在连接器多尺度缺陷检测任务中, GIOU 未能考虑预测框与真实框之间的角度偏差,不能充分体现预测框与目标框的匹配程度。而 SiOU^[20]损失函数通过引入角度和形状的几何信息,增强对缺陷预测框与目标框在中心点偏差与形状相似度方面的惩罚能力,从而有效提高了对小尺寸缺陷目标识别的能力, SiOU 具体计算原理示意图如图 8 所示。

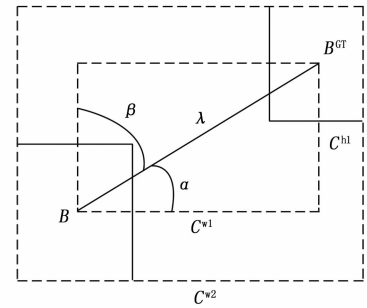


图 8 SiOU 损失函数计算原理图

SiOU 损失函数分别由角度成本、距离成本、形状成本和 IoU 损失组成。角度成本的核心目的是减少边界框回归过程中的自由度,通过约束预测框与真实框的方向对齐,优化模型训练的动态路径,当预测框和真实框的角度越大,则角度损失成本也越高,从而引导梯度方向,其计算公式如下:

$$L_{\theta} = 1 - 2 \times \sin^2 \left(\alpha - \frac{\pi}{4} \right) \quad (1)$$

$$\alpha = \arcsin \left(\frac{C_{h1}}{\lambda} \right) \quad (2)$$

$$\lambda = \sqrt{(b_{C_c}^{gt} - b_{C_c})^2 + (b_{C_c}^{gt} - b_{C_c})^2} \quad (3)$$

$$C_{h1} = \max(b_{C_c}^{gt}, b_{C_c}) - \min(b_{C_c}^{gt}, b_{C_c}) \quad (4)$$

式中, L_{θ} 为角度成本, λ 为真实框与预测框中心点距离, C_{h1} 为真实框与预测框中心点的高度差; $(b_{C_c}^{gt}, b_{C_c}^{gt})$ 为真实框的中心点坐标, (b_{C_c}, b_{C_c}) 为预测框的中心点坐标。

距离成本核心理念是,随着预测框与真实框之间角度差的增加,距离误差对于总体损失的贡献应显著降低,这是为了促使预测框在空间位置上更紧密地贴近真实框。其计算公式如下,当角度对齐好时,模型对中心点距离误差容忍度更低,惩罚更强。反之,对距离误差的惩罚减弱:

$$L_{\Delta} = \sum_{t=x,y} [1 - e^{-(2-L_{\theta}) \times \rho_t}] \quad (5)$$

$$\rho_x = \left(\frac{b_{c_x}^{gt} - b_{c_x}}{C_{w2}} \right)^2 \quad (6)$$

$$\rho_y = \left(\frac{b_{c_y}^{gt} - b_{c_y}}{C_{h2}} \right)^2 \quad (7)$$

式中, C_{w2} 和 C_{h2} 分别为真实框与预测框最小外接矩形的宽和高。

形状成本旨在通过量化预测框与真实框在宽度和高度上的差异,引导模型更精确地调整预测框的形状,使其与目标物体的实际形状更匹配,其计算公式如下:

$$L_{\Omega} = \sum_{t=x,y} (1 - e^{-W_t})^{\gamma} \quad (8)$$

$$W_w = \frac{|\omega - \omega^{gt}|}{\max(\omega, \omega^{gt})} \quad (9)$$

$$W_h = \frac{|h - h^{gt}|}{\max(h, h^{gt})} \quad (10)$$

式中, γ 是对形状损失的关注程度的超参数, ω^{gt} , h^{gt} , ω , h 为真实框和预测框的宽和高。

IoU 损失强调真实框与预测框的非重叠部分,与传统 IoU 的计算公式相同。因此, $SIoU$ 损失函数计算公式如下:

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

针对传统 $SIoU$ 损失函数在极端尺度敏感度不足和复杂场景下回归路径冗余的问题,引入 Inner- IoU ^[21] 的思想改进 $SIoU$ 损失函数。Inner- $SIoU$ 通过辅助边界框的尺度控制机制优化形状成本对小目标检测的适应性,以进一步增强对小目标检测能力和边界框回归效率。其具体原理是使用较小尺度的辅助框计算 IoU 损失,模型聚焦于边界框内部的核心区域从而加速高 IoU 样本的收敛,即比例因子 $r_a < 1$ 。反之,使用较大尺度的辅助框计算 IoU 损失,扩展回归范围,能够提升低 IoU 样本回归效果,即 $r_a > 1$ 。辅助真实框的 4 个顶点坐标计算公式如下:

$$\begin{cases} b_1^{gt} = b_{c_x}^{gt} - \frac{\omega^{gt} * r_a}{2} \\ b_r^{gt} = b_{c_x}^{gt} + \frac{\omega^{gt} * r_a}{2} \\ b_t^{gt} = b_{c_y}^{gt} - \frac{h^{gt} * r_a}{2} \\ b_b^{gt} = b_{c_y}^{gt} + \frac{h^{gt} * r_a}{2} \end{cases} \quad (12)$$

式中, (b_1^{gt}, b_b^{gt}) 、 (b_1^{gt}, b_r^{gt}) 、 (b_r^{gt}, b_t^{gt}) 、 (b_r^{gt}, b_b^{gt}) 分别为辅助真实框的坐标, r_a 为影响因子超参数,通常取 $[0.5, 1.5]$ 。

辅助预测框的 4 个顶点坐标计算公式如下:

$$\begin{cases} b_1 = b_{c_x} - \frac{\omega * r_a}{2} \\ b_r = b_{c_x} + \frac{\omega * r_a}{2} \\ b_t = b_{c_y} - \frac{h * r_a}{2} \\ b_b = b_{c_y} + \frac{h * r_a}{2} \end{cases} \quad (13)$$

式中, (b_1, b_b) 、 (b_1, b_r) 、 (b_r, b_t) 、 (b_r, b_b) 分别为辅助预测框的坐标。

则辅助真实框与预测框的相交区域的宽 w_1 和高 h_1 计算公式如下:

$$\begin{cases} w_1 = [\min(b_r^{gt}, b_r) - \max(b_1^{gt}, b_1)] \\ h_1 = [\min(b_b^{gt}, b_b) - \max(b_t^{gt}, b_t)] \end{cases} \quad (14)$$

则 IoU_{Inner} 计算公式如下:

$$I = w_1 * h_1 \quad (15)$$

$$U = (\omega^{gt} * h^{gt}) * r_a^2 + (\omega * h) * r_a^2 - I \quad (16)$$

$$IoU_{Inner} = \frac{I}{U} \quad (17)$$

式中, I 辅助真实框与预测框的相交区域面积, U 为辅助真实框与预测框并集区域面积。

则 Inner- $SIoU$ 损失函数的计算公式如下:

$$L_{Inner-SIoU} = 1 - IoU_{Inner} + \frac{L_{\Delta} + L_{\Omega}}{2} \quad (18)$$

3 实验设计

3.1 数据集构建

本文使用的数据集来自 500 W 工业相机在现场设备进行多角度拍摄,经筛选后总共得到 820 张含有缺陷的图像,缺陷种类为脏污、凹痕和毛边,同时还需要检测出连接器的金属插针(通过检测出的金属插针的数量判断是否缺少)。随后利用 Labelimg 工具对收集到的缺陷图片进行精确地标注,考虑到连接器缺陷数据集样本相对较少,难以完全覆盖工况场景下连接器的各种缺陷类型。因此,为了提高模型的泛化能力和鲁棒性,本文对标注后的图像进行了一系列的预处理和图像增强步骤,包括但不限于翻转、镜像、缩放、亮度调整和添加随机噪声等数据增强方法,最终生成了 4 100 张可用于训练的图片,并按照 8:1:1 的比例将数据集分别划分为训练集、验证集和测试集,保证了模型训练的有效性。数据集具体的组成信息如表 1 所示。

3.2 实验环境及参数配置

为了避免不同环境对算法性能评估的影响,本文所

表 1 数据集信息			
缺陷种类	训练集/张	验证集/张	测试集/张
脏污	864	108	108
凹痕	824	103	103
毛边	1008	126	126
缺针	584	73	73
总计	3280	410	410

有实验均使用相同的实验环境配置，具体如表 2 所示。模型训练的超参数为：输入图片的分辨率调整为 640 × 640，批次大小（batchsize）设置为 8，线程数设置为 8，优化器使用 AdamW 优化器，初始学习率（Learning rate）为 0.000 1，同时不加载任何预训练权重，并固定随机种子 *seed* = 42，以保证实验的可复现性，共训练 150 个轮次（epoch）。

表 2 实验环境	
名称	参数
操作系统	Windows 11
CPU	Intel(R)Core(TM)i7-14700HX
GPU	NVIDIA GeForce RTX 4070 Laptop
Python	3.9.7
pytorch	1.12.1
CUDA	12.5

3.3 模型评价指标

为了客观评估实时目标检测器的性能表现，本文分别从检测精度与实时检测性能两个维度展开综合评价。在检测精度方面，采用精确度 *P*（Precision）、召回率 *R*（Recall）和平均精确均值 *mAP*（mean Average Precision）来衡量模型对目标位置与类别的识别准确度。计算公式为：

$$P = \frac{TP}{TP + FP}$$

(19)

$$R = \frac{TP}{TP + FN}$$

(20)

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

(21)

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

(22)

式中，*TP* 即目标为正样本且预测正确，*FN* 为目标样本且预测错误，*FP* 为目标为负样本且预测正确。*AP* 和 *mAP* 分别为单一和所有类别的平均精度，*n* 为检测类别的总数。

在实时检测性能方面，通过采用参数量（Params）、浮点计算量（FLOPs）和帧数（FPS）来衡量模型大小与推理速度。其中，参数量和浮点计算量越小，模型越轻量化。FPS 越大，模型推理速度越快，则

实时性越强。

4 实验与结果分析

4.1 消融实验

为验证本文所提出的各个改进模块对模型检测效果提升的有效性，以 RT-DETR-r18 为基准模型进行消融实验，结果如表 3 所示，其中 A、B、C、D 分别代表引入 FasterNetBlock、FE_Block、HS-FPN 和 Inner-SIoU 损失函数。轻量化模块 FasterNetBlock 的引入，使得参数量和 GFLOPs 分别降低了 44.2% 和 40.4%，FPS 达到 96 帧，但也导致平均精度下降了 1.2%，表明其多尺度特征提取能力不足。而当引入本文设计的融合 EMA 注意力机制的创新模块 FE_Block 后，平均精度提升了 1.0%，已经与基准模型大致相同，验证了 EMA 注意力机制的有效性。当引入 HS-FPN 替代原始 RT-DETR 中的 CCFM 模块后，*mAP*₅₀ 提升了 0.4%，模型参数量进一步降低了 0.7 M，验证了 HS-FPN 高效的多尺度特征融合机制，但由于其中注意力机制的运算，导致 GFLOPs 略微提高，FPS 降低至 84 帧。最后，引入 Inner-SIoU 边界框损失函数取代原始的 GIoU 损失函数后，平均精度进一步提升了 0.6%。

表 3 消融实验结果											
A	B	C	D	<i>P</i> /%	<i>R</i> /%	<i>mAP</i> ₅₀ /%	<i>mAP</i> _{50~95} /%	Params/M	GFLOPs	FPS	
—	—	—	—	92.9	87.8	91.7	62.7	19.9	57.2	82	
✓	—	—	—	92.4	85.7	90.5	58.3	11.1	34.1	96	
✓	✓	—	—	92.5	87.6	91.5	61.7	11.2	34.2	87	
✓	✓	✓	—	92.2	88.5	91.9	62.8	10.5	37.6	84	
✓	✓	✓	✓	93.1	89.3	92.5	63.2	10.5	37.6	84	

综上所述，本文引入的所有改进策略与方法均使模型的检测性能有不同程度的提升，相比于基准模型，本文设计的 FEHS-DETR 在精度和轻量化方面表现更优，能够有效的检测到连接器多尺度缺陷，同时其轻量化的设计更适合工业场景下的部署。

4.2 损失函数有效性验证

为了验证 Inner-SIoU 损失函数的有效性，本文在其它改进的基础上，引入当前其它主流的 IoU 进行对比验证，包括 CIoU、DIoU、SIoU 和 WIoU，具体对比结果如表 4 所示。当替换为 CIoU 和 DIoU 时，模型平均精度有所下降。当替换为 SIoU 和 WIoU 时，模型在精度或平均精度方面有小幅度的提升。当更换为 Inner-SIoU 时，模型的精度、召回率和平均精度均达到最高。因此，选用 Inner-SIoU 能够更准确的预测多尺度缺陷的位置和大小。

表 4 不同 IOU 对比结果

IoU 类别	$P/\%$	$R/\%$	$mAP50/\%$	$mAP50\sim95/\%$
GIoU	92.2	88.5	91.9	62.3
CIoU	92.1	88.1	91.6	62.2
DIoU	92.4	88.1	91.8	62.8
SIoU	92.7	88.0	91.9	62.9
WIoU	92.5	88.7	92.0	62.5
Inner-SIoU	93.1	89.3	92.5	63.2

4.3 主流模型对比实验

为进一步评估本文改进模型在连接器多尺度缺陷检测任务的性能，在相同的数据集及实验条件下，将本改进算法分别与当前主流的实时目标检测模型 YOLOv5m、YOLOv5s、YOLOv8m、YOLOv8s、YOLOv3-tiny 进行对比实验，具体实验结果如表 5 所示。

表 5 主流算法对比实验结果

算法	$P/\%$	$R/\%$	$mAP50/\%$	$mAP50\sim95/\%$	Params/M	GFLOPs	FPS
YOLOv5m	92.6	88.0	92.1	63.8	25.1	64.0	88
YOLOv5s	92.5	85.5	90.3	60.9	9.1	23.8	128
YOLOv8m	92.9	89.0	92.3	65.4	25.9	78.7	79
YOLOv8s	91.8	86.7	91.2	62.1	11.1	28.4	123
YOLOv3-tiny	89.9	79.8	85.8	54.6	12.1	18.9	147
RT-DETR-r18	93.0	87.8	91.7	62.7	19.9	57.2	82
FEHS-DETR	93.1	89.3	92.5	63.2	10.5	37.6	84

实验结果表明，本文提出的改进模型 FEHS-DETR，在精确度、召回率和平均检测精度上均优于 YOLO 系列模型算法。其中，平均检测精度 $mAP50$ 分别比基准模型、YOLOv5m、YOLOv5s、YOLOv8m、YOLOv8s、YOLOv3-tiny 提升了 0.8%、0.4%、2.2%、0.2%、1.3%和 6.7%，说明其在模型检测性能方面优于其它主流实时检测模型。在模型轻量化方面，改进模型的参数量为 10.5 M，仅次于 YOLOv5s。同时其相比于基准模型、YOLOv5m 和 YOLOv8m，参数量分别降低了 47.2%、58.2%、59.5%。检测速度 FPS 达到了 84，高于基准模型，仅次于 YOLOv5s 和 YOLOv8s。表明本文算法在保持较高性能的同时具有更加轻量级的模型结构，适合在资源有限的工业场景下应用。

4.4 可视化结果分析

为更直观的对比改进前后算法在本文数据集上的检测效果，本文分别对基准模型 RT-DETR 和改进后的模型 FEHS-DETR 在测试集上的检测效果进行对比。图 9 (a) 分别为 3 张待检测图片，每张图片若干个缺陷区域。图 9 (b) 和 (c) 分别表示经原算法 RT-DETR 和改进算法 FEHS-DETR 的检测效果图。由 3 张图片的检

测结果图可知，改进后的模型 FEHS-DETR 对于金属插针引脚和各种类别缺陷（尤其是微小缺陷如毛边和凹痕）的置信度分数均高于原模型 RT-DETR。由第二张图片可知，改进后的模型不容易被复杂的工业背景所干扰，同时对于缺陷的定位也更加的准确，预测框紧密贴合缺陷区域。由第三张图片可知，改进后的模型检测出了原模型所漏检的脏污缺陷，对于小目标缺陷特征提取能力强。综上所述，改进后的算法在复杂工业背景的干扰下仍然具有良好的检测性能，能够更加充分的表达和预测连接器缺陷目标，同时也对小尺寸目标检测效果提升显著。

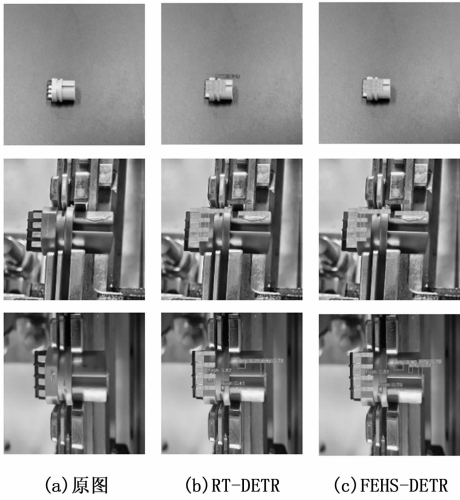


图 9 算法改进前后结果对比

为了更加直观的观察模型学习到的特征情况，本文通过 Grad-CAM 可视化分析技术，对基准模型 RT-DETR 和 FEHS-DETR 进行热力图分析，结果如图 10 所示。改进后的模型相较于基准模型展现出更优的注意力分布和检测性能，基准模型虽然注意力分布主要集中在连接器本身，但对缺陷区域并未有足够的权重。而改进模型不仅对背景的注意力更少，更是精准的将注意力分布集中在连接器的缺陷区域，即改进模型能够关注到毛边和小划痕缺陷具体位置，并且不容易被噪音和背景所干扰。

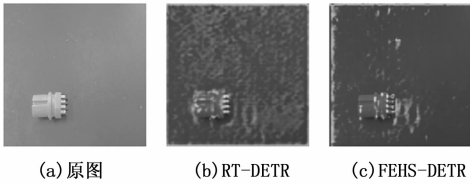


图 10 热力图可视化对比

5 结束语

针对目前连接器外观缺陷检测算法精度低、模型

轻量化程度不高等问题,本文以 RT-DETR-r18 为基准模型,提出了一种改进的小目标缺陷实时检测算法 FEHS-DETR。具体而言,首先使用本文创新模块 FE-Block 改进原主干网络中的 BasicBlock,以降低模型复杂度。然后使用 HS-FPN 多尺度特征融合模块,显著提升模型对小目标的空间定位精度和语义特征表达能力。最后,使用 Inner-SIoU 损失函数替换 GIoU 损失函数以提高模型的泛化能力。实验结果表明,改进算法在拥有更少参数数量的同时,平均精度超过大多数主流实时检测算法,符合连接器外观缺陷实时检测任务的需求。

但是在工况场景下,比如亮度过低、噪声干扰、设备背景复杂等诸多外部因素仍会对本算法产生一定的影响,从而降低检测精度。因此,未来的工作将继续丰富连接器外观缺陷数据集,以提升模型的鲁棒性。然后在保持模型高精度的同时进一步轻量化模型,并且选择合适的剪枝方法降低模型的参数数量和计算量,以适应工业小型算力硬件的缺陷检测需求。

参考文献:

- [1] 文 波. 智能制造技术在工业自动化中的应用分析 [J]. 河南科技, 2021, 40 (28): 58-60.
- [2] 柳 江, 李圣强. 电连接器 IP 防护设计要点 [J]. 机电元件, 2024, 44 (4): 7-11.
- [3] CARION N, MASSA F, SYNNAEVE G, et al. End-to-end object detection with transformers [C] //European conference on computer vision, Cham: Springer International Publishing, 2020: 213-229.
- [4] ZHAO Y, LV W, XU S, et al. Detrs beat yolos on real-time object detection [C] //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2024: 16965-16974.
- [5] ZHANG H, HAO C, SONG W, et al. Adaptive slicing-aided hyper inference for small object detection in high-resolution remote sensing images [J]. Remote Sensing, 2023, 15 (5): 1249.
- [6] 徐仟祥, 曾 勇, 卢 倩, 等. 基于改进 RT-DETR 的车门内拉手表面缺陷检测方法 [J]. 电子测量技术, 2024, 47 (18): 172-181.
- [7] 徐 凯, 武 刚, 刘健鹏, 等. 基于 DMF-YOLOv8n 模型的核电工程焊缝缺陷检测 [J]. 施工技术 (中英文), 2025, 54 (9): 1-8.
- [8] 董润华, 常 青, 孔鹏伟, 等. FDH-DETR 工况场景工人行为及火灾检测算法 [J]. 电子测量技术, 2025, 48 (3): 145-153.
- [9] 刘团圉, 刘淑娟, 哈妮克孜·伊拉洪, 等. Trans-YOLO: 基于 RT-DETR 解码器改进 YOLOv8 的红外小目标检测 [J]. 激光与光电子学进展, 2025, 62 (10): 378-389.
- [10] WANG R, LIANG F, WANG B, et al. ODCA-YOLO: An omni-dynamic convolution coordinate attention-based YOLO for wood defect detection [J]. Forests, 2023, 14 (9): 1885.
- [11] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition [C] //Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 770-778.
- [12] LIU S, QI L, QIN H, et al. Path aggregation network for instance segmentation [C] //Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 8759-8768.
- [13] HOWARD A G, ZHU M, CHEN B, et al. MobileNets: Efficient convolutional neural networks for mobile vision applications [EB/OL]. <https://arxiv.org/abs/1704.04861>.
- [14] CHEN J, KAO S, HE H, et al. Run, don't walk: chasing higher FLOPS for faster neural networks [C] //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023: 12021-12031.
- [15] OUYANG D, HE S, ZHANG G, et al. Efficient multi-scale attention module with cross-spatial learning [C] //ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), IEEE, 2023: 1-5.
- [16] 皮大能, 周 琦, 杨一群, 等. 基于改进 YOLOv8 的水面垃圾检测模型研究 [J]. 湖北师范大学学报 (自然科学版), 2025, 45 (1): 27-34.
- [17] 赵 博, 柳苏纯, 张巍朋, 等. 面向樱桃番茄采摘识别的轻量化 Transformer 架构优化研究 [J]. 农业机械学报, 2024, 55 (10): 62-71.
- [18] CHEN Y, ZHANG C, CHEN B, et al. Accurate leukocyte detection based on deformable-DETR and multi-level feature fusion for aiding diagnosis of blood diseases [J]. Computers in Biology and Medicine, 2024, 170: 107917.
- [19] REZATOFIGHI H, TSOI N, GWAK J Y, et al. Generalized intersection over union: a metric and a loss for bounding box regression [C] //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019: 658-666.
- [20] GEVORGYAN Z. SLOU loss: More powerful learning for bounding box regression [J]. Arxiv Preprint Arxiv: 2205.12740, 2022.
- [21] ZHANG H, XU C, ZHANG S. Inner-iou: more effective intersection over union loss with auxiliary bounding box [J]. Arxiv Preprint Arxiv: 2311.02877, 2023.