

文章编号:1007-2780(2025)10-1532-13

基于改进 RT-DETR 的 PCBA 管脚 焊点缺陷检测方法

李江^{1,2}, 骆炜^{1*}, 陈豪¹, 邓福钦³

(1. 中国科学院福建物质结构研究所, 福建 福州 350108;

2. 中国科学院大学, 北京 100049;

3. 福建星网元智科技有限公司, 福建 福州 350108)

摘要:针对印刷电路板组装件(PCBA)中管脚焊点缺陷因目标尺寸小、分布密集及背景复杂等因素导致检测速度慢、准确率低的问题,本文提出一种基于改进 RT-DETR 的 PCBA 管脚焊点缺陷检测方法。首先,针对 RepC3 在多尺度语义信息融合中缺乏全局感知能力的问题,提出了局部与全局协同增强模块,以提高模型对多尺度特征的全局理解能力。其次,设计了频率感知上采样方法,有效纠正传统插值上采样造成的特征中心偏移问题,增强模型对关键区域的响应能力。最后,引入 RepViT 作为主干网络,在提升特征提取能力的同时有效压缩了模型体积。本文使用在工业生产线采集并自行标注的 PCBA 管脚焊点图像数据集和 PKU-Market-PCB 数据集对所提方法进行评估。在 PCBA 管脚焊点图像数据集上,本文方法与 RT-DETR 相比,模型参数量减少 31.44%,计算量降低 29.21%,mAP₅₀ 提升 9.88%,推理速度提升至 84.74 FPS。在 PKU-Market-PCB 数据集上,mAP₅₀ 和推理速度分别提升了 6.75% 和 23.10%。本文方法在兼顾计算资源与实时性能的基础上显著提升了检测精度与效率,具备良好的工程应用前景和研究价值。

关键词:PCBA; 焊点检测; RT-DETR; 多尺度特征融合; 上采样; RepViT

中图分类号:TP391.4 **文献标识码:**A **doi:**10.37188/CJLCD.2025-0143 **CSTR:**32172.14.CJLCD.2025-0143

PCBA pin solder defect detection method based on improved RT-DETR

LI Jiang^{1,2}, LUO Wei^{1*}, CHEN Hao¹, DENG Fuqin³

(1. *Fujian Institute of Research on the Structure of Matter, Chinese Academy of Sciences,*

Fuzhou 350108, China;

2. *University of Chinese Academy of Sciences, Beijing 100049, China;*

3. *Fujian Xingwang Yuanzhi Technology Co. Ltd., Fuzhou 350108, China)*

Abstract: To address the challenges of the slow detection speed and the low accuracy in pin solder defect

收稿日期:2025-07-04;修订日期:2025-08-01.

基金项目:中国科学院海西研究院 2023 年度“前瞻跨越”计划杰出青年项目(No.CXZX-2023-JQ05);中国科学院海西研究院“福建省创新人才”项目(No.E539UF01)

Supported by 2023 Outstanding Youth Project of the “Foresight and Leap Forward” Program of the Haixi Institute of the Chinese Academy of Sciences (No.CXZX-2023-JQ05); “Fujian Province Innovative Talents” Project of Haixi Institute of the Chinese Academy of Sciences (No.E539UF01)

*通信联系人, E-mail: luowei@fjirsm.ac.cn

detection for printed circuit board assemblies (PCBA), due to small target sizes, dense distributions, and complex backgrounds, this paper proposes an improved RT-DETR-based detection method. First, to overcome the limited global perception capability of the RepC3 module in multi-scale semantic feature fusion, a local and global collaborative enhancement module is introduced to enhance the model's global understanding of multi-scale features. Second, a frequency-aware upsampling module is designed to effectively mitigate feature center offset issues caused by traditional interpolation-based upsampling and to improve the model's responsiveness to key regions. Finally, RepViT is adopted as the backbone network, which improves feature extraction capabilities while significantly reducing model complexity. The proposed method is evaluated on two datasets: a proprietary PCBA pin solder joint image dataset collected and annotated by ourselves on industrial production line and the public PKU-Market-PCB dataset. On the PCBA dataset, compared with the original RT-DETR, the proposed method achieves a 31.44% reduction in model parameters, a 29.21% reduction in computational cost, a 9.88% improvement in mAP_{50} , and an inference speed increase to 84.74 FPS. On the PKU-Market-PCB dataset, it yields gains of 6.75% in mAP_{50} and 23.10% in inference speed. Overall, the proposed method significantly enhances detection accuracy and efficiency while maintaining a balance between computational cost and real-time performance, demonstrating strong potential for practical engineering applications and further research.

Key words: PCBA; solder joint detection; RT-DETR; multi-scale feature fusion; upsampling; RepViT

1 引 言

管脚焊点检测是印刷电路板组装件(Printed Circuit Board Assembly, PCBA)制程中的关键环节,其焊接质量直接关系到产品的性能稳定性。随着电子元器件集成度的不断提高,焊接工艺愈发复杂,PCBA 面板呈现出更高的复杂度。同时,管脚焊点尺寸较小、分布密集,且缺陷类型多样、形态复杂,这些特点使得焊点缺陷检测仍然面临较大的挑战。

目前,主流目标检测方法可分为基于卷积神经网络(CNN)和基于 Transformer 两大类。其中,CNN 方法根据是否需要生成候选框,可进一步分为两阶段和单阶段检测算法。两阶段方法如 RCNN^[1]、Fast RCNN^[2]和 Faster-RCNN^[3],通过候选区域生成与目标识别两步可实现较高的检测精度,但其处理流程复杂,难以满足管脚焊点检测对实时性的要求。相比之下,YOLO^[4]和 SSD^[5]等单阶段检测方法通过直接回归目标位置和类别,具备更高的检测速度,因而更适合在实际工业场景中部署^[6-9]。然而,CNN 固有的局部感受野限制了其对小目标和复杂背景的全局特征建模能力,容易导致误检或漏检。此外,CNN

方法普遍依赖非极大值抑制^[10](Non-Maximum Suppression, NMS)来去除冗余框,而 NMS 对阈值设定较为敏感,易受目标密集度与尺度变化的影响,进而降低模型的检测效率与鲁棒性。因此,基于 CNN 的方法在应对焊点尺寸小、密集度变化大且背景复杂的工业检测场景时,仍存在一定的性能瓶颈。

随着 Transformer 架构^[11]在自然语言处理领域的成功应用,部分研究者开始尝试将其引入到目标检测任务中,以增强模型对全局上下文信息的建模能力,并减少对 NMS 等后处理操作的依赖。DETR^[12]作为首个基于 Transformer 架构的端到端目标检测模型,具备建模长距离依赖关系的能力,并且省略了锚框设计和 NMS 操作,极大简化了检测流程。然而,DETR 存在的收敛速度慢、优化查询困难等问题,导致其难以应用在实时性要求高、计算资源受限的工业场景。为解决上述问题,百度在 2024 年提出 RT-DETR 算法^[13],该算法在保留 Transformer 并行建模优势的基础上,融合轻量化网络结构,设计了高效的多尺度特征处理模块,并引入基于交并比的查询动态选择机制,显著提升了检测效率。RT-DETR 在保证检测精度的同时能够大幅缩减推理时间,

为工业部署提供了可行方案。

RT-DETR能够在大多数目标检测场景中表现出优异的性能,但在面对目标尺寸小、分布密集、背景复杂等场景时,仍存在漏检、错检、模型参数量大、推理速度慢等问题。为此,国内外学者从多个角度对该模型进行了深入研究。刘思元等^[14]针对小目标检测效果不佳的问题,在主干网络中引入部分卷积,并结合双向特征金字塔网络与CARAFE上采样算子,增强了多尺度特征融合能力。孙光灵等^[15]设计了二重跨阶段多尺度特征融合模块,以解决密集目标、远景小目标等复杂场景下漏检、误检及检测准确率低的问题。赵子琪等^[16]通过引入微小目标特征融合结构、MCP Block、可学习位置编码及NWD和MPDIoU损失函数,提升了模型的小目标感知能力与收敛速度。宣岁寒等^[17]重新设计特征融合模块,提出可学习位置编码与EMA重参数响应模块,以增强模型对小目标的定位精度与检测效率。Li等^[18]通过多阶段卷积与尺度自适应融合结构提升小目标检测能力,并降低了计算量。Wang等^[19]提出了线性空间金字塔注意力机制,使模型能够感知多尺度信息,并通过信息广播策略促进层级间通信,从而增强其对复杂交通场景的适应性。Yu等^[20]引入部分卷积与门控机制提高局部与全局特征融合效果,并采用HiLo注意力替代传统多头自注意力机制,有效解决了模型在复杂背景下检测准确率低、实时性差的问题。

尽管上述改进方法取得了一定进展,但在算法优化方面仍存在诸多不足。首先,大多数方法在特征融合阶段未能充分利用注意力机制以增强关键区域的响应能力,导致模型在复杂背景下对微小且密集目标容易产生误检或漏检。其次,现有上采样策略普遍采用传统插值方法,容易引起特征中心偏移,进而影响目标的精确定位。此外,主干网络的特征提取能力仍有优化空间,特别是在需要兼顾检测精度与推理效率的嵌入式场景中,现有模型的表现仍不理想。因此,针对RT-DETR在PCBA管脚焊点检测任务中因目标尺寸小、分布密集及背景复杂导致的检测速度慢、准确率低的问题,本文提出一种基于改进RT-DETR模型的管脚焊点缺陷检测方法,在兼顾计算资源与实时性能的基础上显著提升了检

测精度与效率,具备良好的工程应用前景和研究价值。

2 RT-DETR改进算法

本文围绕复杂背景下密集管脚焊点的小目标检测任务,对RT-DETR模型进行了一系列结构性优化,旨在提升检测精度的同时有效压缩模型体积。首先,本文提出了一种局部与全局协同增强模块(Local and Global Collaborative Enhancement Module, LGCEM),用于替换原CCFF模块中的RepC3结构^[21],通过融合局部卷积与多维注意力机制,有效增强特征融合阶段的上下文建模能力和目标感知能力。其次,本文设计了频率感知上采样算法(Frequency-Aware UpSample, FA-UpSample)以缓解特征中心偏移问题,通过引入频域信息重构细粒度特征,细化目标边界信息,从而进一步提高密集小目标的检测效果。最后,本文采用轻量级主干网络RepViT替代原ResNet结构,在强化特征提取能力的基础上显著降低了模型参数量。图1展示了改进模型的整体结构。

2.1 局部与全局协同增强模块

在RT-DETR模型的CCFF结构中,本文采用RepC3模块对通道维度拼接后的多级特征图进行融合。RepC3主要通过并联 1×1 卷积和 3×3 卷积实现多尺度感知建模,在多数目标检测模型中表现良好,具备一定的特征融合能力。然而,该结构主要依赖不同尺度卷积核在空间上的局部感受能力,难以建立跨层级、跨语义空间的有效信息关联^[22],尤其在背景复杂的密集检测任务中,仍然存在多尺度信息利用不充分的问题。

PCBA管脚焊点检测图像常包含大量金属元器件、焊盘、走线等干扰结构,这些区域在纹理、亮度以及形状上与目标焊点高度相似,极易引发误检或漏检。在此类高密度、小目标、强干扰背景的场景下,模型若能通过有效的特征融合机制充分整合不同层级特征图,即同时利用低层特征中的局部细节信息与高层特征中的全局语义信息,将显著增强模型对前景目标的识别能力以及对背景噪声的抑制能力。然而,由于低层与高层特征在感受野大小与语义表达能力上存在显著差异,融合过程不仅需要细节捕捉能力,还

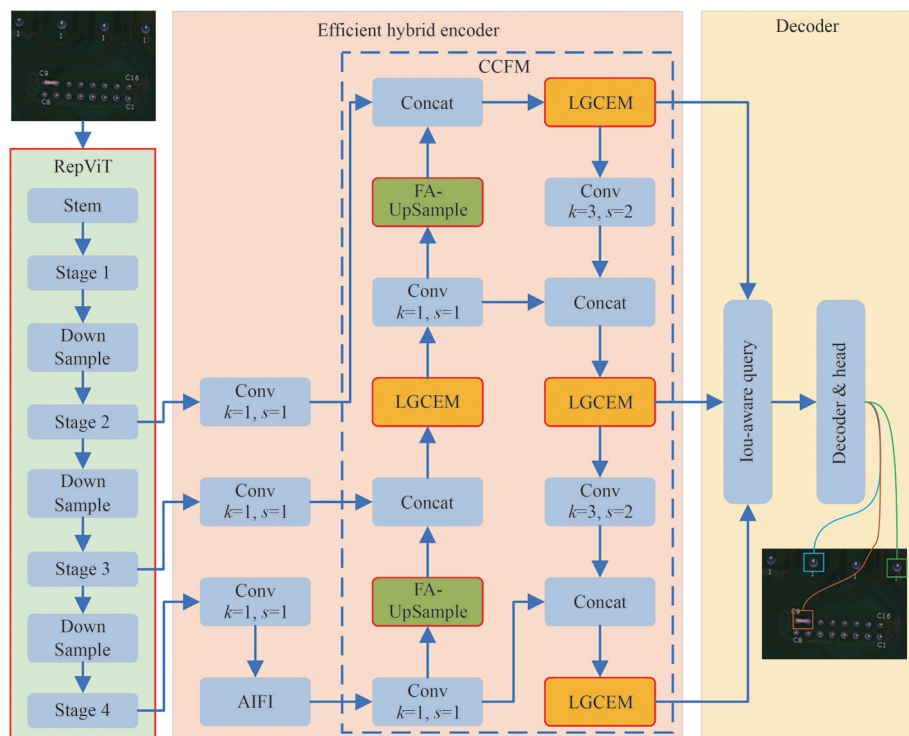


图 1 改进的 RT-DETR 网络结构

Fig. 1 Improved RT-DETR network structure

需具备全局语义建模能力^[23],而 RepC3 结构本质上仍局限于卷积的局部建模范式,难以有效融合多尺度特征间的上下文信息与语义一致性,因而限制了其在复杂检测任务中的表现。

为解决上述多尺度信息融合不充分的问题,本文提出一种局部与全局协同增强模块(LGCEM),用于替换 RT-DETR 中的 RepC3 结构,具体结构如图 2 所示。该模块融合了卷积操作的局部感知能力与注意力机制建模长距离依赖的优势,既保留了局部空间细节信息,也增强了全局语义特征的一致性表达。通过这一改进,模型能够更充分地利用不同尺度特征之间的互补信息,提升对目标与背景的区分能力,从而在复杂背景、小目标密集分布的检测场景中实现更高的检测精度与更强的鲁棒性。

LGCEM 由局部建模与全局建模两个阶段构成。局部建模阶段设计为多分支卷积结构,增强空间细节建模能力。输入特征首先经过 1×1 卷积进行通道调整并降低计算复杂度,随后送入两个并行分支:一支通过 1×1 卷积建模通道间语义关系,另一支采用 3×3 卷积提取局部空间信息。两分支输出通过逐元素相加进行融合,实现

通道信息与空间信息的协同建模。该结构重复堆叠两次,并引入 SiLU 激活函数以提升非线性表达能力。此外,将经 1×1 卷积调整后的输入与堆叠输出直接相加以建立残差连接,从而保证特征完整性。

为避免局部建模与全局建模过程中的信息冗余或冲突,本文通过串行结构对两阶段的输入、输出进行了显式解耦,并通过残差连接与注意力加权的方式进行融合。局部建模阶段更侧重于细节信息的提取与纹理边缘的保留,全局建模阶段则专注于跨空间依赖关系与语义一致性的增强。两阶段在功能上互为补充,前者提供准确的局部定位依据,后者提供稳定的上下文判别能力。最终,通过注意力加权融合与残差连接,确保局部与全局信息能够协同发挥作用而非相互干扰,提升模型整体特征表达的完整性与判别性。

全局建模阶段主要构建长距离依赖关系,提升低层语义与高层语义之间的一致性。考虑传统注意力机制的计算开销较大,本文借鉴 MobileViT^[24]与 CoTNet^[25]的设计思想,采用基于卷积的轻量注意力机制模拟查询(Query)、键(Key)、值(Value)的生成过程。该阶段包含 3 个功能分支:第一分

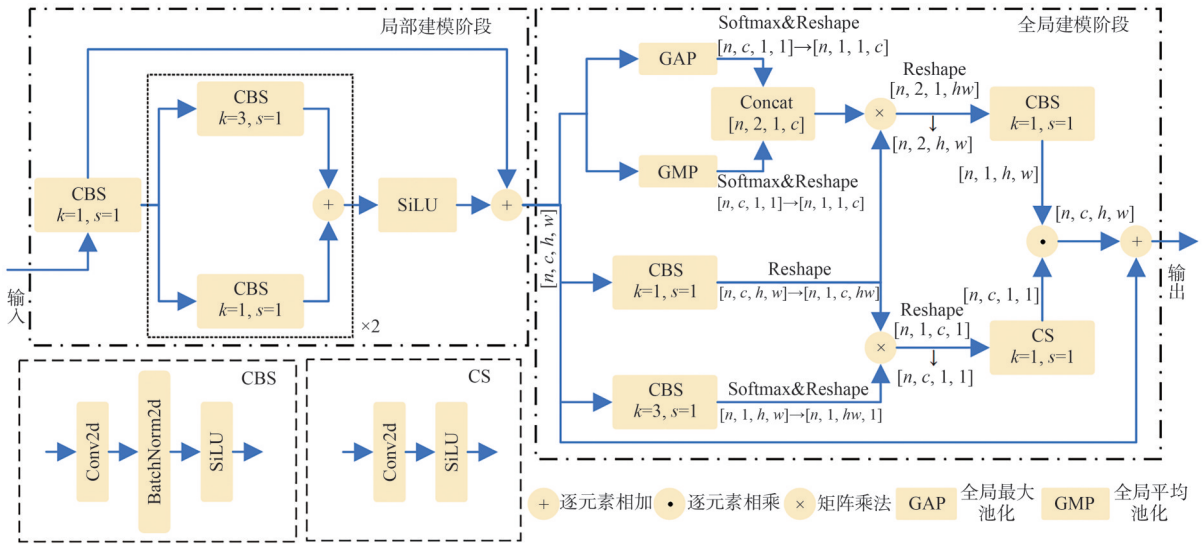


图 2 LGCEM网络结构

Fig. 2 LGCEM network structure

支通过全局平均池化和最大池化提取通道注意力,第二分支以 1×1 卷积生成值(Value)特征,第三分支通过 3×3 卷积提取空间上下文信息并经归一化形成空间注意力。通道注意力与空间注意力分别与值(Value)特征进行加权融合,生成跨通道与跨空间维度的全局响应,最后通过逐元素乘法将两种注意力结果融合,实现多维度全局信息的联合建模。

2.2 频率感知上采样

在RT-DETR框架中,默认使用最邻近插值完成上采样操作。虽然该方法简洁高效,但它通过直接将单个像素值复制到多个位置进行尺寸扩展,并没有充分考虑特征重建过程中特征空间对应区域的一致性^[26],进而导致特征中心偏移的问题^[27]。

RT-DETR将主干网络提取的具有高级语义信息的P3、P4、P5级特征图作为解码器的输入,如图3所示。目标在P3级特征图中以橙色区域表示,经过两次跨步卷积下采样后,分别在P4级和P5级特征图中映射为蓝色区域和绿色区域。P5级特征图经最邻近插值上采样得到分辨率与P4级特征图一致的P4'级特征图后,目标映射的绿色区域在P3级特征图中的感受野相比蓝色区域更显著地偏离了真实的目标位置,表现为明显的特征中心偏移现象。

尽管最邻近插值在计算效率和实现简便性

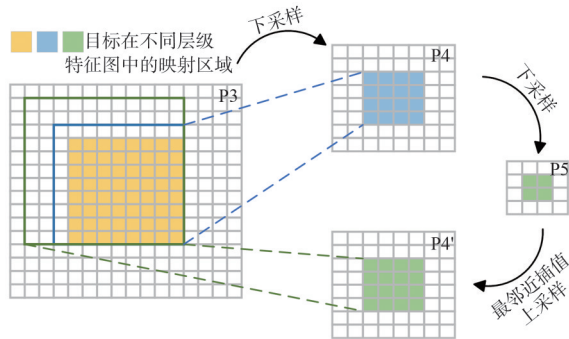


图 3 特征中心偏移示意图

Fig. 3 Feature center offset diagram

方面具有一定优势,但其引入的不准确边界会对空间特征的一致性造成破坏,特别是在目标小、预测密集的管脚焊点缺陷检测任务中,这种特征中心偏移会严重削弱模型对关键目标区域的响应能力,从而导致误检或漏检问题。因此,空间对齐误差成为限制密集小目标检测性能提升的重要因素。为了解决这一问题,本文在FreqFusion^[28]框架的基础上设计了可以缓解特征中心偏移的频率感知上采样方法(FA-UpSample),其网络结构如图4所示。FreqFusion是一种创新的特征融合方法,通过频域分解方法将融合特征中的高频与低频成分分别处理,解决了传统融合方法中类别内不一致性和边界模糊问题,显著提升了模型在密集图像预测任务中的表现。

FA-UpSample 保留了FreqFusion中对高级

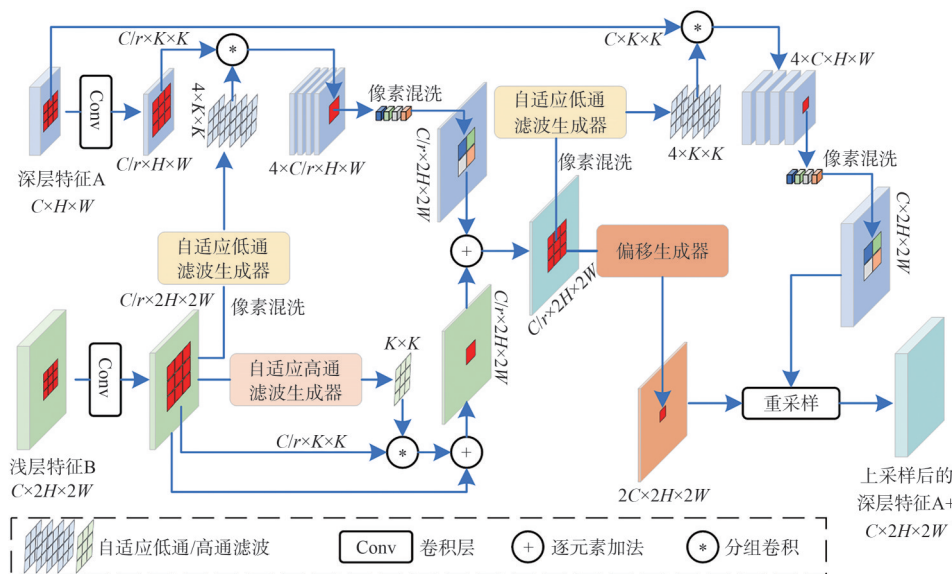


图 4 FA-UpSample 网络结构

Fig. 4 FA-UpSample network structure

特征的两阶段增强策略,同时精简了对低级特征的频域冗余调制过程,有效降低了计算开销。具体而言,初次融合后的中间特征仅利用自适应低通滤波器提取的引导信息对高级特征进行再次调制,既保留了低频主干信息,也避免了不必要的高频干扰。随后,偏移生成器对调制后的特征图进行空间对齐修正,以缓解由传统插值上采样导致的特征中心偏移问题,从而输出更具语义一致性的重建特征图。偏移生成器基于局部区域特征间的相似性估计空间偏移量,通过余弦相似度计算生成引导图,在上采样过程中动态调整采样位置,从而实现针对不同区域的精细空间校准。这一过程从原理上避免了传统插值中“一对多”像素复制带来的空间误差,实现了对目标结构中心的有效对齐。

不同于传统插值方法仅依赖深层特征自身进行空间扩展,FA-UpSample 进一步引入浅层特征作为结构先验,将其作为补充输入与深层语义信息协同参与上采样过程。浅层特征中保留了更多原始边界与结构信息,其空间分布更接近输入图像的真实结构。通过将其作为先验信息引入偏移生成器,FA-UpSample 能够更准确地感知各位置的实际边缘走向,从而辅助深层语义特征完成精确重采样,有效缓解特征中心偏移问题。通过结合浅层特征的高频边界信息与深层特征的语义表达,FA-UpSample 能够输出具有更清晰

边缘轮廓且语义一致性更强的上采样特征。此外,该模块引入的自适应高通滤波器有效弥补了浅层特征中缺失的结构细节,提升了边界恢复能力与小目标的定位精度。

2.3 轻量级 RepViT 主干网络

在 RT-DETR 模型中,采用 ResNet 网络作为特征提取网络。尽管该网络在通用目标检测模型中表现良好,但其固定大小的卷积核和有限的感受野限制了其对复杂背景和小尺寸目标的建模能力,从而在处理具有复杂背景的小目标任务时表现欠佳,尤其是在处理 PCBA 管脚焊点检测这类具有背景复杂、目标密集、尺寸微小等特点的任务时,其性能严重影响模型整体的检测精度。与此同时,虽然近年来涌现的基于 Transformer 构建的主干网络(如 Swin Transformer^[29])在特征建模方面具有优势,能够捕捉图像中的长距离依赖关系,但其庞大的计算资源需求和较慢的推理速度使其难以适用于 PCBA 产线等对实时性和资源约束较为敏感的工业应用场景。为了解决这些问题,本文引入轻量化的 RepViT^[30]作为特征提取网络。RepViT 主要包括 Stem、Stage 和 Down-Sample 三个关键模块,其网络结构如图 5 所示。

RepViT 融合了 Vision Transformers 的全局建模优势和卷积神经网络在局部感知与高效特征提取方面的优点,具备良好的精度与效率平衡,其核心设计包括结构重参数化机制,在训练

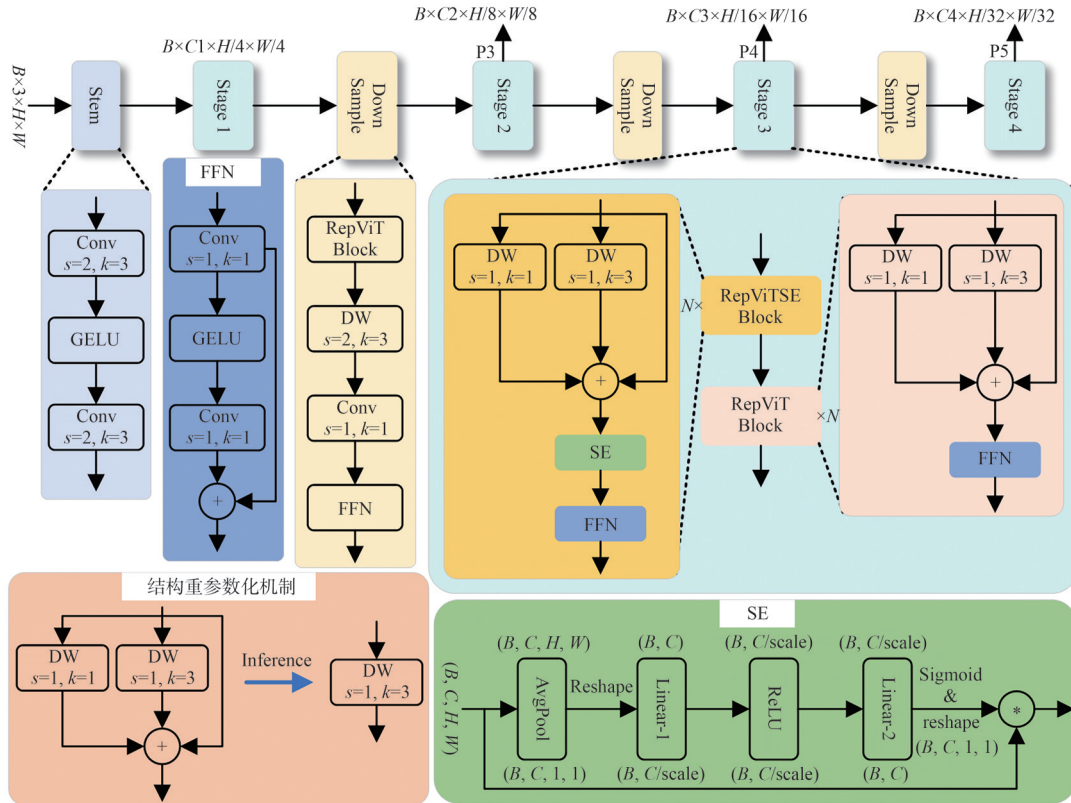


图5 RepViT网络结构

Fig. 5 RepViT network structure

阶段通过多分支增强特征表达,而在推理阶段则简化为等效的单分支网络,从而显著提升模型的推理速度。此外,RepViT还采用了MetaFormer框架,通过显式分离Token mixer与Channel mixer模块,增强模型对多尺度特征的建模能力以及提升其在复杂场景下的适应性。因此,基于RepViT的模型不仅可以在小目标密集检测任务中表现出色,还能在资源受限设备上实现快速部署,满足管脚焊点检测对精度与实时性的要求。

3 实验结果与分析

3.1 实验环境与超参数设置

本研究在 Windows 11 操作系统下进行,硬件平台包括 Intel Core i9-12900K CPU 和 NVIDIA GeForce RTX 4090 GPU。深度学习框架为 PyTorch 2.0.1,并使用 CUDA 11.8 作为 GPU 计算平台。RT-DETR 原始模型建议训练 72 轮,为提升模型收敛效果,本研究将训练周期扩展至 216 轮。批量大小设置为 8,其余超参数保持默认

配置,即采用 AdamW 优化器,初始学习率为 0.000 1,权重衰减率为 0.000 1。

3.2 实验数据集

本研究实验阶段使用的 PCBA 管脚焊点图像数据集由本文研究团队从实际生产线采集,该数据集包含 1 753 张分辨率为 $4\,024 \times 3\,036$ 的焊点图像。根据生产中常见的焊点缺陷类型,该数据集共标注了 5 类目标,分别为正常(Normal)、跪脚(Bent Lead)、多锡(Excessive Solder)、连焊(Solder Bridge)以及少锡(Insufficient Solder),部分典型图像样本如图 6 所示。数据集被随机划分为训练集、验证集和测试集,比例为 8:1:1,分别包含 1 404 张、175 张与 175 张图像。

为进一步验证所提改进方法的有效性及鲁棒性,本文在 PKU-Market-PCB 数据集^[31]上进行了消融实验。PKU-Market-PCB 是一个公开的合成 PCB 数据集,由北京大学智能机器人开放实验室发布,其中包含 1 386 张图像以及 6 种缺陷(漏孔、鼠咬、开路、短路、杂散、杂铜),用于图像检测、分类和配准任务。本文选取了其中适用于

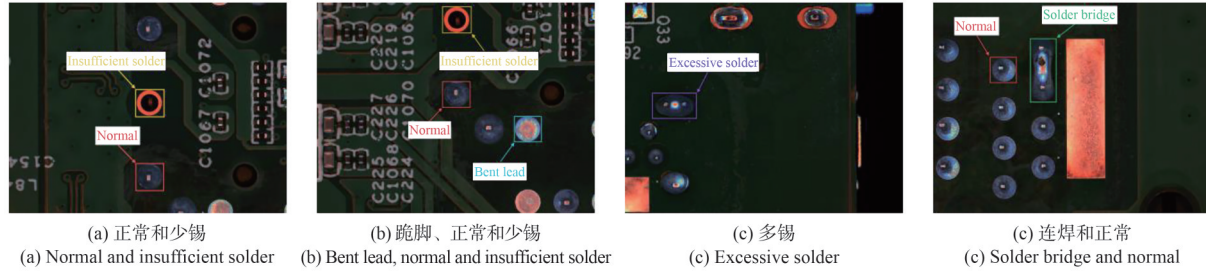


图6 PCBA 管脚焊点图像示例

Fig. 6 Example of PCBA pin solder joint image

目标检测任务的 693 张图像,并随机选择 593 张图像作为训练集,100 张图像作为验证集。相关图像示例及详细信息可在北京大学智能机器人开放实验室官方网站获取。

值得说明的是,PKU-Market-PCB 数据集所包含的缺陷类型并不在本文讨论范围内,但相关缺陷检测具有背景复杂、目标尺寸小、缺陷分布较为密集等特点,与本文所探讨的 PCBA 管脚焊点缺陷检测有一定相似度,故仅在 3.4 节消融实验中作为辅助数据集使用。

3.3 评价指标

为全面评估本文所提出的 PCBA 管脚焊点缺陷检测模型的实际效果,本文采用了多项关键性能指标对其进行评价。这些指标包括平均精度均值(mean Average Precision, mAP)、每秒千兆浮点运算数(GFLOPs)、参数量(Params)以及每秒帧数(FPS)。其中,mAP 用于评估模型在不同类别上的检测性能,是评价模型性能的综合性指标,具体定义如下:

$$P = \frac{TP}{TP + FP}, \quad (1)$$

$$R = \frac{TP}{TP + FN}, \quad (2)$$

$$AP = \int_0^1 P(R) dR, \quad (3)$$

$$mAP = \frac{\sum_{i=0}^N AP_i}{N}, \quad (4)$$

式中: P 代表精确率(Precision); R 代表召回率(Recall); TP (True Positive)表示被正确识别的目标数量; FN (False Negative)表示未被检测出的目标数量,即漏检; FP (False Positive)表示被错误识别为目标的数量,即误检; N 表示目标类别的总数。

GFLOPs 和 Params 是评估模型计算复杂度和资源开销的关键指标。其中,GFLOPs 表示模型在前向传播过程中所需执行的浮点运算次数,能够反映模型的计算开销;Params 则代表模型中需要训练和存储的参数总量,是衡量模型复杂度与存储需求的重要依据。此外,模型的实时性可以通过每秒处理的图像帧数(FPS)进行评估,FPS 越高,表示模型在实际部署中的推理速度越快。通过综合比较 GFLOPs、Params 与 FPS,可以分析模型改进对计算资源的消耗情况,确保检测精度与轻量化、实时性的平衡。

3.4 消融实验

本文对 RT-DETR 模型进行了逐步优化,依次引入局部与全局协同增强模块(LGCEM)、频率感知上采样模块(FA-UpSample),并将主干网络替换为轻量化的 RepViT,以提高其在复杂背景下对密集小尺寸焊点缺陷的检测能力。为了系统评估各项改进对检测性能的具体贡献,本文在 PCBA 管脚焊点图像数据集和 PKU-Market-PCB 数据集上开展了消融实验,相关实验结果如表 1 和表 2 所示,其中“√”表示该模型版本包含对应模块。

通过表 1 可知,在 PCBA 管脚焊点图像数据集上,将原有的 RepC3 结构替换为 LGCEM 后,模型检测精度显著提升,mAP₅₀ 由 74.9% 提升至 79.6%,mAP₇₅ 由 74.0% 提升至 79.2%,mAP_{50:95} 从 58.8% 提升至 63.8%,表明本文所设计的 LGCEM 在充分整合多尺度特征图中的局部细节与全局语义信息方面效果显著,有效提升了特征融合质量。同时,模型的计算量和参数量都有明显下降,分别下降 8.17% 和 8.33%;推理速度小幅提升。这说明 LGCEM 相较于 RepC3 结构可以降低模型复杂度,同时提升检测精度。进一步引

表 1 在 PCBA 管脚焊点图像数据集上的消融实验

Tab. 1 Ablation experiment on PCBA pin solder joint image dataset

RT-DETR	LGCEM	FA-UpSample	RepViT	mAP ₅₀ /%	mAP ₇₅ /%	mAP _{50:95} /%	GFLOPs	Params/M	FPS
✓				74.9	74.0	58.8	63.92	49.80	73.60
✓	✓			79.6	79.2	63.8	58.70	45.65	74.09
✓	✓	✓		81.0	80.0	64.9	59.22	45.65	72.98
✓	✓	✓	✓	82.3	80.5	66.6	43.82	35.26	84.74

注:粗体数字代表最优值

表 2 在 PKU-Market-PCB 数据集上的消融实验

Tab. 2 Ablation experiments on the PKU-Market-PCB dataset

RT-DETR	LGCEM	FA-UpSample	RepViT	mAP ₅₀ /%	mAP ₇₅ /%	mAP _{50:95} /%	GFLOPs	Params/M	FPS
✓				91.8	46.9	50.1	63.92	49.80	97.16
✓	✓			93.0	45.7	50.0	58.70	45.65	113.16
✓	✓	✓		93.3	50.1	51.1	59.22	45.65	103.32
✓	✓	✓	✓	98.0	52.5	53.1	43.82	35.26	119.60

注:粗体数字代表最优值

入 FA-UpSample 后,检测性能继续提升,mAP₅₀、mAP₇₅ 和 mAP_{50:95} 分别提升至 81.0%、80.0% 和 64.9%。尽管 FA-UpSample 与传统的最近邻插值方法在参数量方面相当,但由于其具备频率感知和分解能力,可以有效缓解传统插值上采样方法引起的特征中心偏移问题,因而可以提高模型检测精度。计算量虽略有上升,但推理速度仅小幅下降,在可接受范围内。最后,采用 RepViT 替代原模型的 ResNet50 主干网络,模型整体性能进一步优化。最终的 mAP₅₀ 提升至 82.3%,mAP₇₅ 提升至 80.5%,mAP_{50:95} 提升至 66.6%。与此同时,模型的计算量与参数量分别降低了 31.44% 和 29.21%,推理速度 FPS 提升至 84.74。该结果充分证明了 RepViT 具备优异的特征提取能力和模型压缩效果,能够在提升检测精度的同时显著降低计算开销,是焊点缺陷检测任务中理想的主干网络。

通过表 2 可知,在 PKU-Market-PCB 数据集上,本文在逐步引入 LGCEM、FA-UpSample 和 RepViT 后,模型性能得到不同程度的提升,其各项指标的变化趋势与在 PCBA 管脚焊点数据集上取得的效果基本相近,进一步验证了本文所提方法在不同类型的 PCB 缺陷检测任务中具有良好的通用性。特别是在引入 RepViT 后,模型的检测精度得到明显提升,并且实现了更优的计算

效率,充分证明了本文所提出改进的有效性和对不同数据集的适用性。

3.5 对比实验

为了评估本文改进模型在 PCBA 管脚焊点检测任务中的综合性能,本文选取了其他 9 种具有代表性的目标检测算法进行对比实验,包括 SSD、Faster-RCNN、Deformable DETR、RT-DETR-R34、YOLOv9-c、YOLOv10-l 以及 Ultralytics 团队开源的 YOLOv5-mu、YOLOv8-m 和 YOLO11-l。所有模型均在相同的训练环境与统一数据集上进行训练,并使用各自官方或论文推荐的超参数,确保实验结果的可比性与公正性。

对比实验结果如表 3 所示,本文方法在多个关键性能指标上均表现出显著优势。在检测精度方面,本文模型在 mAP₅₀ 与 mAP₇₅ 上分别达到 82.3% 与 80.5%,在综合指标 mAP_{50:95} 上达到 66.6%,不仅显著优于传统检测器 SSD(54.1%) 与 Faster R-CNN(60.4%),也优于所有 YOLO 系列最新版本,包括 YOLOv10-l(65.4%) 与 YOLO11-l(63.0%),表明本文方法在小目标定位和复杂结构识别任务中具备更强的鲁棒性和语义建模能力。在模型复杂度方面,本文模型的参数量为 35.26M,虽然略高于 YOLOv8-m(25.86M)、YOLOv9-c(25.53M)、YOLOv10-l(24.40M)等轻量模型,但显著低于 Deformable

表 3 在 PCBA 管脚焊点数据集上的对比实验
Tab. 3 Comparative experiment on PCBA pin solder joint dataset

Models	Input size	mAP ₅₀ /%	mAP ₇₅ /%	mAP _{50:95} /%	GFLOPs	Params/M	FPS
SSD ^[5]	300×300	77.7	72.7	54.1	24.28	30.63	44.63
Faster-RCNN ^[3]	800×800	74.7	73.7	60.4	41.37	105.00	26.90
Deformable DETR ^[32]	640×640	75.4	29.8	38.3	40.10	94.21	19.62
YOLOv5-mu	640×640	78.0	—	64.0	25.07	64.40	92.07
YOLOv8-m	640×640	80.2	—	65.4	25.86	79.10	87.72
YOLOv9-c ^[33]	640×640	79.5	—	65.6	25.53	103.70	76.92
YOLOv10-l ^[34]	640×640	80.4	—	65.4	24.40	120.30	67.11
YOLO11-l	640×640	79.7	—	63.0	25.30	86.90	50.06
RT-DETR-R34 ^[13]	640×640	77.8	77.1	63.1	33.08	43.98	73.25
本文算法	640×640	82.3	80.5	66.6	35.26	43.82	84.74

注:粗体数字代表最优值

DETR(40.10M)和Faster R-CNN(41.37M),在保持结构紧凑的同时实现了更高的检测精度。在计算开销方面,本文模型的GFLOPs为43.82,接近RT-DETR-R34(43.98),远低于YOLOv9-c(103.70)、YOLOv10-l(120.30)等高精度模型,显示出良好的可部署性。在推理速度方面,本文模型在单张图像上的推理帧率为84.74 FPS,接近YOLOv5-mu(92.07 FPS)与YOLOv8-m(87.72 FPS),高于YOLO11-l(50.06 FPS)和Deformable DETR(19.62 FPS),展现出优异的实时性与部署效率。因此,本文提出的改进模型在保持较低参数量和计算开销的前提下,实现了更高的检测精度与推理速度,在PCBA管脚焊点缺陷这类对微小目标敏感且场景复杂的工

业检测任务中展现出更强的实用性与工程应用价值。

为了更加直观地展示不同模型在PCBA管脚焊点数据集上的检测性能差异,本文选取了一张检测难度较高的测试图像作为模型输入,输出各模型的检测结果图,如图7所示。该输入图像的焊点数量较多,且存在强烈的背景干扰因素,因而更能真实反映各模型在复杂场景下的鲁棒性与泛化能力。从图7检测结果可见,本文方法、YOLOv8-m、YOLOv10-l和YOLO11-l在该图像上均能准确识别所有焊点目标,表现出良好的检测完整性与背景抑制能力。而Faster R-CNN、YOLOv5-mu、YOLOv9-c与RT-DETR-R34虽然同样检测出了全部焊点目标,但均存在不同程

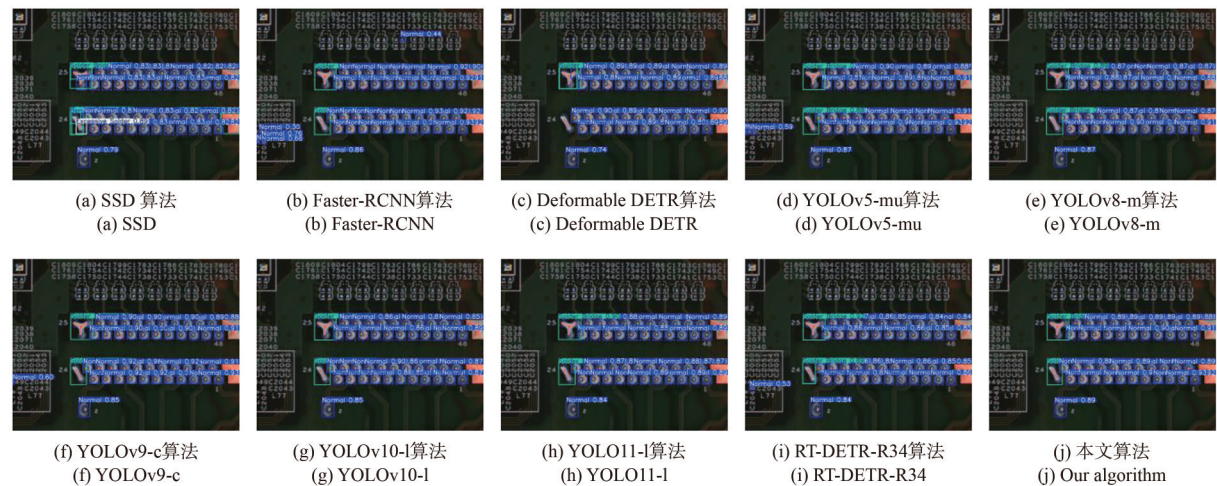


图 7 各模型在 PCBA 管脚焊点测试图像上的检测结果
Fig. 7 Detection results of each model on the PCBA pin solder joint test image

度的误检现象,部分背景区域被错误识别为焊点,其中 Faster R-CNN 的误检数量最多,显示出其在复杂背景条件下的特征判别能力存在一定局限。检测效果相对较差的是 SSD 与 Deformable DETR。Deformable DETR 虽在该图像中未出现明显的误检区域,但漏检了一个实际存在的焊点缺陷,反映出其在该类缺陷类型上的学习能力和敏感性不足。SSD 的表现最为逊色,不仅存在多个重复检测框,还将一个连桥缺陷误识别为多锡,同时误将两个背景区域识别为连桥缺陷,整体检测稳定性和准确性较差。

4 可视化分析

为了进一步验证各模块对模型关注能力的提升效果,本文对引入不同模块的模型版本在测试图像上的特征响应进行了热力图可视化分析,结果如图 8 所示。RT-DETR 原始模型(图 8(a))在未引入任何改进模块时,热力图中响应区域较为分散,部分焊点缺陷未能被有效激活,反映出

模型在小目标检测任务中关注能力不足。引入 LGCEM 后的模型(图 8(b))在焊点区域的整体响应显著增强,说明该模块在融合局部细节与全局语义特征方面具有良好表现,有助于提升多尺度特征的融合质量。然而,由于缺乏有效的上采样机制,多个目标的边界存在一定程度的模糊,影响了精确定位能力。在此基础上,加入 FA-UpSample 后(图 8(c)),特征响应更加聚焦于真实缺陷位置,焊点的响应中心更加清晰,边缘轮廓也得到明显改善,展现出较强的位置定位能力与边界感知能力。最终模型(图 8(d))在引入轻量化 RepViT 的基础上,融合上述所有模块,在所有焊点区域均实现了高度集中且强度显著的响应,缺陷边界清晰完整,非目标区域的响应有效抑制,整体热力图呈现出良好的聚焦性。上述可视化结果充分验证了所提出的结构优化方法在增强模型对目标关注能力方面的有效性,证明本文方法在面对复杂背景与密集小目标场景时,具备更高的鲁棒性与误检抑制能力,有效提升了焊点缺陷检测的可靠性。

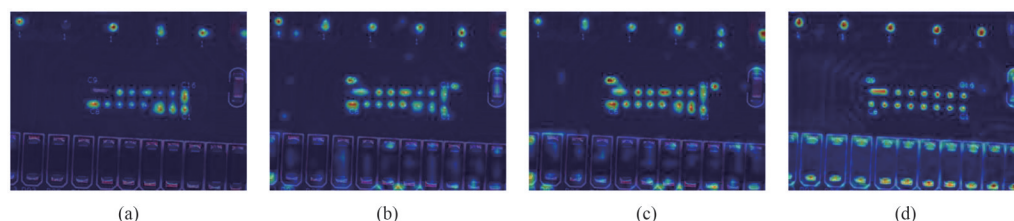


图 8 基于各项改进方法的模型热力图效果对比。(a) RT-DETR 原始模型;(b) 在 RT-DETR 中引入 LGCEM;(c) 在 RT-DETR 中引入 LGCEM 与 FA-UpSample;(d) 在 RT-DETR 中引入 LGCEM、FA-UpSample 与 RepViT。

Fig. 8 Comparison of model heat map effects based on various improvement methods. (a) Baseline RT-DETR architecture; (b) RT-DETR with LGCEM; (c) RT-DETR with LGCEM and FA-UpSample; (d) RT-DETR with LGCEM, FA-UpSample, and RepViT.

5 结 论

本文针对 PCBA 管脚焊点因目标尺寸小、分布密集及背景复杂等因素导致检测精度低、推理速度慢的问题,对 RT-DETR 模型进行了结构优化,并在自制数据集与公开数据集上开展了系统的实验验证。实验结果表明,与原始 RT-DETR 模型相比,本文方法在 PCBA 管脚焊点图像数据集上的 mAP_{50} 提高了 9.88%,达到 82.3%,推理速度提升至 84.74 FPS,参数量减少 31.44%,计

算量降低 29.21%,充分证明了所提方法在检测精度、效率与模型规模之间取得了良好的平衡。为进一步验证改进模型的通用性,本文在 PKU-Market-PCB 数据集上进行了测试。实验结果显示, mAP_{50} 与推理速度分别提升了 6.75% 和 23.10%,表明该方法在其他复杂工业缺陷检测场景中也具有较强的适应能力。综上所述,本文方法在多个指标上取得了显著优化,满足了实际生产线对 PCBA 管脚焊点质量检测的高性能需求,具有较好的部署价值与应用前景。

参 考 文 献:

- [1] GIRSHICK R, DONAHUE J, DARRELL T, *et al.* Rich feature hierarchies for accurate object detection and semantic segmentation [C]//*Proceedings of 2014 IEEE Conference on Computer Vision and Pattern Recognition*. Columbus: IEEE, 2014: 580-587.
- [2] GIRSHICK R. Fast R-CNN [C]//*Proceedings of 2015 IEEE International Conference on Computer Vision*. Santiago: IEEE, 2015: 1440-1448.
- [3] REN S 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.
- [4] REDMON J, DIVVALA S, GIRSHICK R, *et al.* You only look once: unified, real-time object detection [C]//*Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition*. Las Vegas: IEEE, 2016: 779-788.
- [5] LIU W, ANGUELOV D, ERHAN D, *et al.* SSD: single shot MultiBox detector [C]//*Proceedings of the 14th European Conference on Computer Vision*. Amsterdam: Springer, 2016: 21-37.
- [6] 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(8): 1227-1240.
- LIANG L M, LONG P W, FENG Y, *et al.* Improving the lightweight VTG-YOLOv7-tiny for steel surface defect detection [J]. *Optics and Precision Engineering*, 2024, 32(8): 1227-1240. (in Chinese)
- [7] 张润梅, 潘晨飞, 陈梓华, 等. 改进 YOLOv8 的焊缝表面缺陷检测算法[J]. *光电工程*, 2025, 52(3): 240296.
- ZAHNG R M, PAN C F, CHEN Z H, *et al.* Improved weld surface defect detection algorithm from YOLOv8 [J]. *Opto-Electronic Engineering*, 2025, 52(3): 240296. (in Chinese)
- [8] 王驰, 沈晨, 黄庆, 等. 夜间动物图像自监督学习增强与检测方法[J]. *中国光学(中英文)*, 2024, 17(5): 1087-1097.
- WANG C, SHEN C, HUANG Q, *et al.* Self-supervised learning enhancement and detection methods for nocturnal animal images [J]. *Chinese Optics*, 2024, 17(5): 1087-1097. (in Chinese)
- [9] 陈名渝, 谢玥辰, 吕雄涛, 等. 冷轧钢表面与内部缺陷检测研究[J]. *中国光学(中英文)*, 2024, 17(4): 823-833.
- CHEN M Y, XIE Y C, LV X T, *et al.* Detection of surface and internal defects in cold rolled steel [J]. *Chinese Optics*, 2024, 17(4): 823-833. (in Chinese)
- [10] VIOLA P, JONES M. Rapid object detection using a boosted cascade of simple features [C]//*Proceedings of 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. Kauai: IEEE, 2001: 511-518.
- [11] VASWANI A, SHAZEER N, PARMAR N, *et al.* Attention is all you need [C]//*Proceedings of the 31st International Conference on Neural Information Processing Systems*. Long Beach: Curran Associates Inc., 2017: 6000-6010.
- [12] CARION N, MASSA F, SYNNAEVE G, *et al.* End-to-end object detection with transformers [C]//*Proceedings of the 16th European Conference on Computer Vision*. Glasgow: Springer, 2020: 213-229.
- [13] ZHAO Y, LV W Y, XU S L, *et al.* DETRs beat YOLOs on real-time object detection [C]//*Proceedings of 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Seattle: IEEE, 2024: 16965-16974.
- [14] 刘思元, 高凯, 雍龙泉. 改进 RT-DETR 的航拍小目标检测算法[J]. *计算机工程与应用*, 2025, 61(4): 272-281.
- LIU S Y, GAO K, YONG L Q. Improved RT-DETR algorithm for aerial small object detection [J]. *Computer Engineering and Applications*, 2025, 61(4): 272-281. (in Chinese)
- [15] 孙光灵, 王薪博, 李艳秋. 改进 RT-DETR 的双轮车头盔检测算法[J]. *电子测量与仪器学报*, 2025, 39(4): 62-73.
- SUN G L, WANG X B, LI Y Q. Improved helmet detection algorithm for two-wheeled vehicles of RT-DETR [J]. *Journal of Electronic Measurement and Instrumentation*, 2025, 39(4): 62-73. (in Chinese)
- [16] 赵子琪, 李卫东, 李晓娟. 基于改进实时 Transformer 的航拍图像小目标检测算法[J]. *科学技术与工程*, 2025, 25(13): 5527-5534.
- ZHAO Z Q, LI W D, LI X J. Small target detection algorithm in aerial images based on improved RT-DETR [J]. *Science Technology and Engineering*, 2025, 25(13): 5527-5534. (in Chinese)
- [17] 宣岁寒, 罗印升, 宋伟. 基于改进 RT-DETR 的航拍图像小目标检测算法[J]. *电光与控制*, 2025, 32(4): 44-51.
- XUAN S H, LUO Y S, SONG W. A small target detection algorithm in aerial images based on improved RT-DETR [J]. *Electronics Optics & Control*, 2025, 32(4): 44-51. (in Chinese)
- [18] JI L, HUANG C H, LI H W, *et al.* MS-DETR: a real-time multi-scale detection transformer for PCB defect detection [J]. *Signal, Image and Video Processing*, 2025, 19(3): 203.

- [19] WANG Y H, XU S B, WANG P, *et al.* Vehicle detection algorithm based on improved RT-DETR [J]. *The Journal of Supercomputing*, 2025, 81(1): 290.
- [20] YU C H, CHEN X W. Railway rutting defects detection based on improved RT-DETR [J]. *Journal of Real-Time Image Processing*, 2024, 21(4): 146.
- [21] 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]//*Proceedings of 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Vancouver: IEEE, 2023: 7464-7475.
- [22] 张志豪, 厉小润, 陈淑涵. 基于改进 YOLO11 的无人机航拍图像小目标检测算法[J]. *液晶与显示*, 2025, 40(6): 915-930.
- ZHANG Z H, LI X R, CHEN S H. Small object detection algorithm in UAV aerial images based on improved YOLO11 [J]. *Chinese Journal of Liquid Crystals and Displays*, 2025, 40(6): 915-930. (in Chinese)
- [23] CHEN Z L, LU S N. CAF-YOLO: a robust framework for multi-scale lesion detection in biomedical imagery [C]//*Proceedings of the ICASSP 2025—2025 IEEE International Conference on Acoustics, Speech and Signal Processing*. Hyderabad: IEEE, 2025: 1-5.
- [24] MEHTA S, RASTEGARI M. MobileViT: light-weight, general-purpose, and mobile-friendly vision transformer [C]//*Proceedings of the 10th International Conference on Learning Representations*. Online: OpenReview.net, 2022: 1-13.
- [25] LI Y H, YAO T, PAN Y W, *et al.* Contextual transformer networks for visual recognition [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2023, 45(2): 1489-1500.
- [26] 彭雪玲, 林珊玲, 林志贤, 等. 改进的 YOLOv5s 太阳能电池片缺陷检测算法[J]. *液晶与显示*, 2024, 39(2): 237-247.
- PENG X L, LIN S L, LIN Z X, *et al.* Defect detection algorithm of improved YOLOv5s solar cell [J]. *Chinese Journal of Liquid Crystals and Displays*, 2024, 39(2): 237-247. (in Chinese)
- [27] CHEN S Y, CHENG T H, FANG J M, *et al.* TinyDet: accurate small object detection in lightweight generic detectors [J/OL]. *arXiv*, 2023: 2304.03428.
- [28] CHEN L W, FU Y, GU L, *et al.* Frequency-aware feature fusion for dense image prediction [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2024, 46(12): 10763-10780.
- [29] LIU Z, LIN Y T, CAO Y, *et al.* Swin transformer: hierarchical vision transformer using shifted windows [C]//*Proceedings of 2021 IEEE/CVF International Conference on Computer Vision*. Montreal: IEEE, 2021: 10012-10022.
- [30] WANG A, CHEN H, LIN Z J, *et al.* Rep ViT: revisiting mobile CNN from ViT perspective [C]//*Proceedings of 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Seattle: IEEE, 2024: 15909-15920.
- [31] HUANG W B, WEI P. A PCB dataset for defects detection and classification [J/OL]. *arXiv*, 2019: 1901.08204.
- [32] ZHAO X Z, SU W J, LU L W, *et al.* Deformable DETR: deformable transformers for end-to-end object detection [C]//*Proceedings of the 9th International Conference on Learning Representations*. Online: OpenReview.net, 2021: 1-11.
- [33] WANG C Y, YEH I H, LIAO H Y M. YOLOv9: learning what you want to learn using programmable gradient information [C]//*Proceedings of the 18th European Conference on Computer Vision*. Milan: Springer, 2025: 1-21.
- [34] WANG A, CHEN H, LIU L H, *et al.* YOLOv10: real-time end-to-end object detection [C]//*Proceedings of the 38th Conference on Neural Information Processing Systems*. Vancouver: Curran Associates Inc., 2024: 1-15.

作者简介:



李江, 男, 硕士研究生, 2019 年于海南大学获得学士学位, 主要从事图像处理技术在工业检测上的应用研究。
E-mail: lijia19@mails.ucas.ac.cn



骆炜, 男, 博士, 高级工程师, 2023 年于德国斯图加特大学获得博士学位, 主要从事智能机器人以及图像处理技术方面的研究。E-mail: luowei@fjirsm.ac.cn