

DOI:10.19344/j.cnki.issn1671-5276.2026.03.029

融合动态预测头 YOLOv8 的 PCB 缺陷检测

许智晨^{1,2}, 聂闻², 罗瑶^{1,2}, 钱梁梁^{1,2}

(1. 福州大学 先进制造学院, 福建 泉州 362251; 2. 中国科学院 泉州装备制造研究中心, 福建 泉州 362216)

摘要:针对在制造过程中任何阶段出现的漏焊、破损、开路、短路、毛刺、假铜等 PCB 缺陷, 提出一种改进的 YOLOv8 算法, 旨在提高 PCB 缺陷检测的效率和准确性。通过引入 DCNv3 卷积和 DyHead 动态检测头, 优化模型的特征提取和空间感知能力, 使其适应不同场景下物体的形状和大小变化, 特别是对小尺寸和密集目标的识别。采用 QFL 改进了损失函数, 以增强模型对定位质量的敏感度。实验结果表明: 改进后的算法与原算法相比, 精度得到有效提升, mAP 提升 2.8%。改进后的模型在 PCB 缺陷检测上表现出较高的精度和良好的效率, 相比于其他目标检测算法表现出更高的检测精度和更快的处理速度, 展现了显著的性能提升。

关键词:PCB 缺陷检测; YOLOv8; 动态检测头; 特征提取; 损失函数改进; 小目标识别

中图分类号:TB302.5 **文献标志码:**A **文章编号:**1671-5276(2026)03-0148-06

PCB Defect Detection using YOLOv8 with Dynamic Prediction Head

XU Zhichen^{1,2}, NIE Wen², LUO Yao^{1,2}, QIAN Liangliang^{1,2}

(1. School of Advanced Manufacturing, Fuzhou University, Quanzhou 362251, China;

2. Quanzhou Equipment Manufacturing Institute, Chinese Academy of Sciences, Quanzhou 362216, China)

Abstract:To address PCB defects such as missing solder, breakage, open circuits, short circuits, burrs, and false copper that may occur at any stage of the manufacturing process, an improved YOLOv8 algorithm is proposed to enhance the efficiency and accuracy of PCB defect detection. By introducing DCNv3 convolutions and a DyHead dynamic detection head, the model's feature extraction and spatial perception capabilities are optimized, enabling it to adapt to variations in object shapes and sizes across different scenarios, particularly for recognizing small and densely arranged targets. The loss function is improved using QFL to enhance the model's sensitivity to localization quality. Experimental results show that compared with the original algorithm, the improved version achieves effective gains in precision, with a 2.8% increase in mAP. The enhanced model demonstrates high accuracy and good efficiency in PCB defect detection, outperforming other object detection algorithms in both detection precision and processing speed, exhibiting significant performance improvements.

Keywords:PCB defect detection; YOLOv8; dynamic detection head; feature extraction; loss function improvements; small object recognition

0 引言

印刷电路板 (printed circuit board, PCB) 是现代电子设备的核心组成部分, 其承载着电子元件, 并确保电子信号的正确传输。随着电子行业的快速发展, 对 PCB 的质量要求越来越高, 但出现任何缺陷都可能导致电路板的性能下降甚至出现故障。

PCB 缺陷的类型多样, 包括漏焊、破损、开路、短路、毛刺、假铜等。这些缺陷可能在制造过程中

的任何阶段出现, 因此及时发现并修正这些缺陷对于降低生产成本和提高产品质量至关重要。传统的 PCB 缺陷检测方法主要依赖人工视觉检查或简单的自动化工具, 但这些方法往往耗时、效率低且准确率不高^[1]。

近年来, 深度学习技术在图像处理和目标检测领域取得了显著进展。特别是单阶段目标检测标法 (you only look once, YOLO)^[2] 系列算法以其高速和高准确率在实时目标检测任务中表现出色。YOLOv8 作为该系列的最新版本, 拥有更强

基金项目:国家自然科学基金项目 (41072232); 国家重点研发计划项目 (2021YFC3001300); 福建省科学院科学技术合作计划项目 (STS, 2022T3051); 江西省杰出青年基金项目 (20212ACB214005)

第一作者简介:许智晨 (1998—), 男, 福建福州人, 硕士研究生, 研究方向为机械与电子相关图像处理, gesila_grief@foxmail.com。

通信作者简介:聂闻 (1982—), 男, 重庆人, 研究员, 博士生导师, 研究方向为图像处理嵌入式电子, wen.nie@vip.tom.com。

的性能和更好的泛化能力^[3]。因此,本研究旨在探索如何将 YOLOv8 算法改进并应用于 PCB 缺陷检测,以期实现更高效、更准确的自动化检测。

由于印刷电路板自身体积不大,PCB 上的缺陷及一些元器件都是极小的,尤其对于 PCB 上一些线路缺陷如开路、短路等都是毫米级缺陷,其表面缺陷具有目标小和类型较多的问题^[4-5]。为此本研究对 YOLOv8 算法进行改进,使其更适合 PCB 缺陷检测的需求,能较好地进行 PCB 小缺陷的检测。具体目标包括:1) 调整 YOLOv8 的网络结构,以适应 PCB 图像的特点;2) 优化损失函数和训练策略,以提高缺陷检测的准确率和稳定性;3) 在 PCB 数据集上验证改进算法的性能,并与其他检测方法进行对比分析。通过实现这些目标,本研究希望为 PCB 缺陷检测领域提供一种高效、可靠的深度学习解决方案。

1 YOLOv8 改进

1.1 标准 YOLOv8 模型

YOLOv8 的网络结构如图 1 所示。主要由 3 部分组成:骨干网络(backbone)、颈部网络(neck)和预测输出头(head)。在这 3 部分中,骨干网络负责利用卷积技术从输入的 RGB 图像中提取深层特征。这些特征包含了图像中的重要信息,为后续的目标检测提供了基础。颈部网络则主要承担着整合骨干网络提取特征的任务,采用特征金字塔网络(FPN)的结构,通过将不同层级的特征进行融合,形成更加丰富和细致的特征表示。预测输出头是负责具体的目标检测任务,它利用 3 种不同尺度的检测器来对图像中的目标进行分类和定位。这种多尺度的检测机制使得 YOLOv8 能够同时检测不同大小的目标,从而提高了检测的准确性。

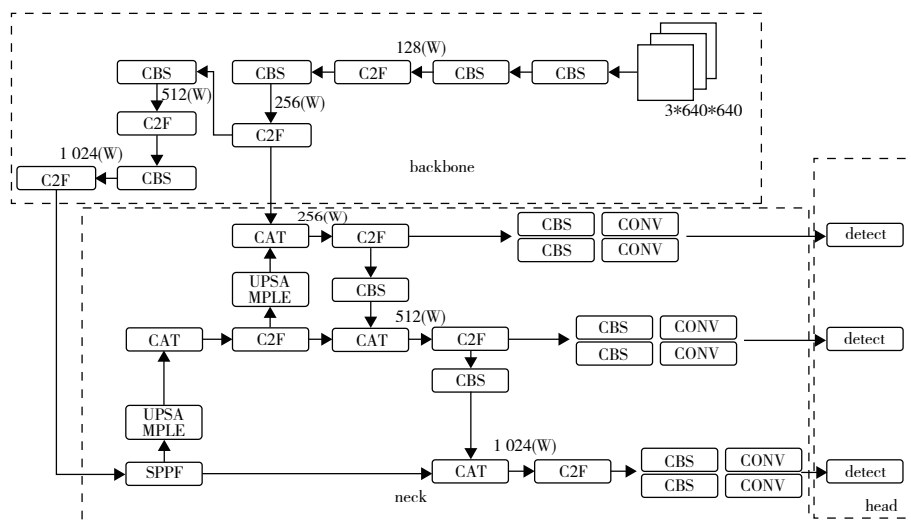


图 1 YOLOv8 原网络结构

1.2 改进 YOLOv8 模型

为了提高 PCB 缺陷检测的效率和准确性,以及对一些小缺陷目标的检测,本文通过引入可变形卷积网络(deformable strip convolution, DCNv3)和 DyHead 动态检测头对原 YOLOv8 网络结构进行改进。其中将骨干网络中的 C2F 模块改进为 DCNv3 模块,原 YOLOv8 的检测头,改为 DyHead 检测头,并采用改进了的质量焦点损失函数(Quality Focal Loss, QFL),优化了模型的特征提取和空间感知能力,以及目标的分类和定位,特别是对小尺寸和密集目标的识别。改进后的网络如图 2 所示。

1.3 DCNv3 卷积

PCB 板表面缺陷特征通常表现为小尺寸且类型多样化,而在 YOLOv8 的骨干网络中,C2F 模块作为残差特征学习的核心部件,通过增加跨层连接的分支,不仅丰富了模型的梯度流,还形成了具有强大特征表示能力的网络模块。然而,C2F 模块及其内部的 Bottleneck 结构(如图 3 所示)中涉及的传统卷积运算在提取小尺寸目标特征方面存在一定的局限性。为了克服这一挑战并增强模型的特征提取能力,本研究引入了一种新型的可变形卷积网络,即 DCNv3,以提升对 PCB 板表面缺陷的检测性能。

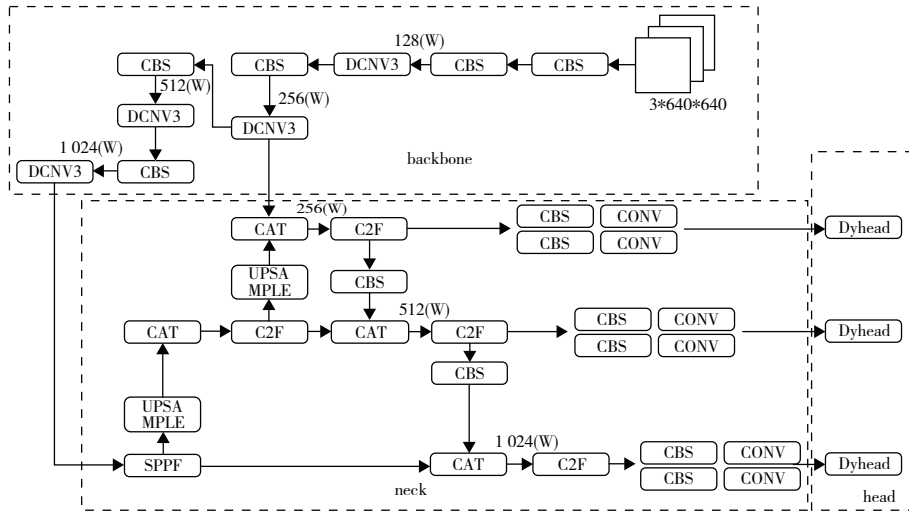


图2 改进式 YOLOv8 网络结构

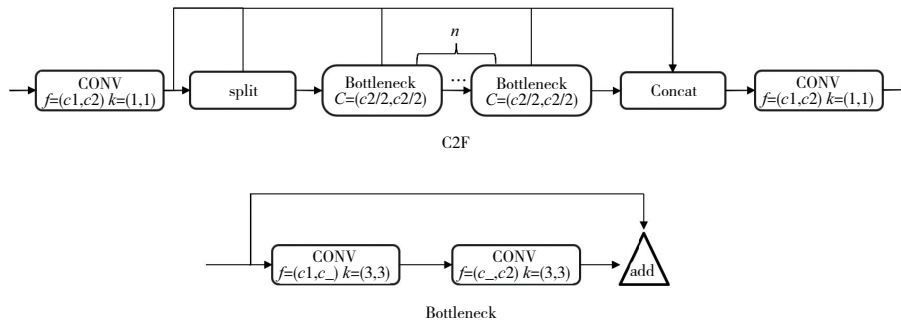


图3 C2F 和 Bottleneck 模块结构

DCNv3 是在原始可变形卷积 (DCN)^[6] 基础上的一种改进,它借鉴了深度可分离卷积的思想,采用与位置无关的权重替代分组权重,实现不同采样点之间的投影权重共享。此外,DCNv3 将空间聚合过程分成若干组,每个组具有独立的采样偏移量,使得卷积核能够更灵活地适应不同场景下物体的形状和大小变化。核心算子为稀疏动态卷积,它通过输入相关的信息为条件,实现自适应空间聚合^[7]。通过权重共享、引入多组机制和归一化处理,DCNv3 在提取关键目标特征的同时,减少了无关上下文信息的干扰,从而提高了模型在小目标检测等任务中的性能。DCNv3 的数学表达式如下

$$y(p_0) = \sum_{g=1}^G \sum_{k=1}^K w_g m_{gk} x_g(p_0 + p_k + \Delta p_{gk}) \quad (1)$$

式中: G 为聚合组数; w_g 为每组内的共享投影权重; m_{gk} 为第 g 个组的第 k 个采样点的调制标量; x_g 为切片的输入特征图; Δp_{gk} 为 p_k 的偏移。

1.4 DyHead 检测头

在 PCB 缺陷检测的场景中,图像往往包含

密集且小尺寸的目标,这些目标容易发生遮挡,且细节难以识别。因此,需要缺陷检测算法具备全面的感知能力。尽管 YOLOv8 采用了解耦的头结构,将目标检测和分类任务分离,但在针对 PCB 缺陷检测的场景中,其对小目标的检测和定位效果仍有待提升。为了解决这一问题,本研究引入了一种新型的动态检测头 DyHead^[8],旨在增强模型在不同尺度、空间和任务感知方面的能力。

DyHead 通过集成尺度感知注意力机制、空间感知注意力机制和任务感知注意力机制,分别作用于不同的特征层级、空间位置和输出通道,从而显著提升模型对目标的表征能力。这使得模型在检测小尺寸目标时更加灵活和精确。具体来说,DyHead 从骨干网络 (backbone) 接收输入向量 (L, S, C),并将其分别传递到 3 个增强感知的注意力模块中,其结构如图 4 所示,以实现针对不同感知维度的优化。这种设计使得 DyHead 能够有效地适应 PCB 缺陷检测等复杂场景,提高模型的检测性能。

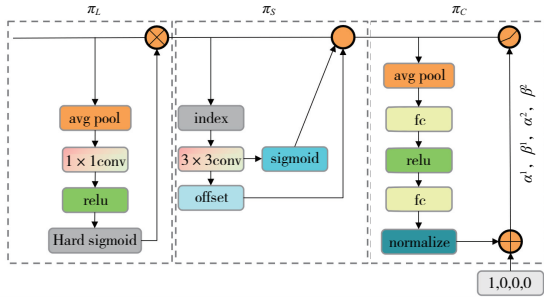


图4 DyHead 模块结构

尺度感知注意力通过池化和卷积操作融合不同尺度的特征信息。

$$\pi_L(F) \cdot F = \sigma \left(f \left(\frac{1}{SC} \sum_{sc} F \right) \right) \cdot F \quad (2)$$

式中 f 为 1×1 卷积层线性函数, 激活函数 hard-sigmoid 公式为

$$\sigma(x) = \max \left(0, \min \left(1, \frac{x+1}{2} \right) \right) \quad (3)$$

空间感知注意力则通过采用可变形卷积, 使得注意力机制能够学习到稀疏化的表示, 进而聚合同一空间内的不同特征。

$$\pi_S(F) \cdot F = \frac{1}{L} \sum_{l=1}^L \sum_{k=1}^K w_{l,k} \cdot F(l; p_k + \Delta p_k; c) \cdot \Delta m_k \quad (4)$$

式中: k 为系数采样位置数量; $p_k + \Delta p_k$ 为通过自学习的空间偏移移动的位置; Δp_k 为由可变形卷积学习而来; Δm_k 为 Δp_k 自学习的重要性。

任务感知注意力通过动态地激活或关闭特征通道, 根据不同的任务需求选择合适的特征子集, 从而提高模型的适应性和效率。

$$\pi_C(F) \cdot F = \max(\alpha^1(F) \cdot F_C + \beta^1(F), \alpha^2(F) \cdot F_C \beta^2(F)) \quad (5)$$

其中学习控制激活阈值的超函数

$$\theta(\cdot) = [\alpha^1, \beta^1, \alpha^2, \beta^2]^T \quad (6)$$

最后, 通过 3 个模块不断堆叠结合, 从而得到增强后的输出结果。

1.5 损失函数改进

在 YOLOv8 中, 边界框回归损失的计算采用了 CIoU 和 DFL 损失的结合方法。具体而言, DFL 部分以交叉熵形式计算边界框与标签之间的损失概率。然后, 将边界框的分布概率转换为预测框, 并通过 CIoU 损失来优化预测框与真实框之间的匹配度, 从而实现了对预测框的整体优化。然而, CIoU 损失在描述边界框纵横比的相对值时存在一定的不精确性, 且未能有效处理不同难易程度

样本之间的平衡问题。

为了解决这些问题, 本研究引入了一种新的 QFL^[9], 这是一种专为目标检测任务设计的改进型损失函数。QFL 的核心思想是将目标的定位质量(如边界框与真实对象的 IoU 得分)直接融合到分类损失中, 形成一个统一的度量。这种设计不仅能够有效解决分类和定位之间的一致性问题, 而且使得模型在训练过程中更加关注那些难以定位或分类的样本, 从而提高检测性能。通过引入 QFL, YOLOv8 在处理 PCB 缺陷检测等任务时能够实现更准确的目标检测和定位。

QFL 数学公式如下

$$QFL(\sigma) = -|y - \sigma|^\beta (1 - y) \log(1 - \sigma) + y \log(\sigma) \quad (7)$$

式中: y 为 0~1 的质量标签; σ 为预测。QFL 的全局最小解即是 $\sigma = y$ 。这样交叉熵部分变为完整的交叉熵, 同时调节因子变为距离绝对值的幂次函数。

2 实验结果与分析

2.1 训练设备和参数设置

实验以 Linux 为操作系统, 采用 Xeon(R) 8255C CPU 和 Nvidia RTX 3090 GPU 的工作站进行训练。具体的实验环境细节如表 1 所示。在模型训练阶段, 我们采用随机梯度下降进行优化, 初始学习率设定为 0.01, 输入图像为 640×640 的尺寸, 批量大小设为 32, 训练轮次 300。这些参数的设定旨在确保模型的训练效果和稳定性, 同时考虑到计算资源的有效利用。

表1 训练设备配置图

名称	配置
中央处理器	Xeon(R) 8255C CPU @ 2.50 GHz
内存	32 GB
显卡	GeForce RTX 3090 24G
系统环境	Linux
框架	PyTorch 2.0.0
CUDA	CUDA 11.8

2.2 实验数据集

本研究中采用的实验数据集来自某大学智能机器人开放实验室, 该数据集专门针对 PCB 缺陷检测, 涵盖了 6 种常见的 PCB 缺陷类型, 包括漏孔(missing hole)、鼠咬(mouse bite)、开路(open circuit)、短路(short)、毛刺(spur)和假铜(spurious copper), 如图 5 所示。每张图像中可能包含多个

缺陷实例。

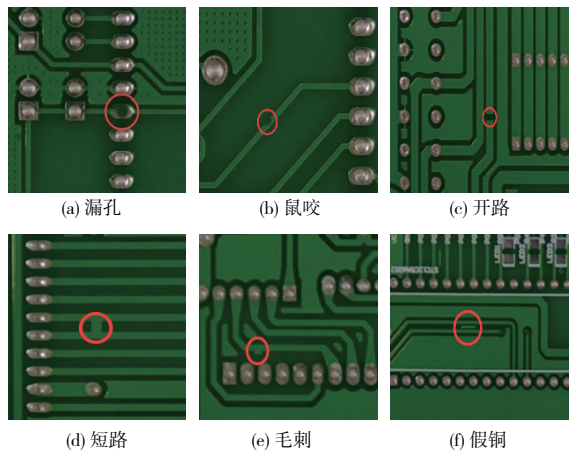


图 5 数据集缺陷展示

在深度学习领域,模型的性能很大程度上依赖于大量标注数据的质量。数据集中存在的噪声和错误标注往往会对模型产生不利影响,导致特征提取和学习过程受到干扰。特别是当数据量有限时,模型容易发生过拟合现象、过度拟合噪声和无关信息,从而降低了模型的泛化能力和准确率。本研究提出为实际 PCB 场景下使用的数据增强策略,包括添加高斯噪声、模糊处理、随机旋转以及进行均值方差缩放等操作。这些增强方法旨在提高模型对各种干扰和变化的适应性,从而增强其鲁棒性和泛化能力。这些丰富的数据为本研究的模型训练和测试提供了良好的基础。最终,整个训练集图片 3 383 张的具体信息如表 2 所示。将整个数据集划分为训练集、验证集和测试集,比例分别为 7 : 2 : 1。这样的划分旨在确保模型在训练过程中能够充分学习特征,并在验证集和测试集上进行有效的性能评估。

表 2 数据集信息 单位:张

数据集类型	现场数据
漏孔	587
鼠咬	582
开路	572
短路	561
毛刺(Spur)	578
假铜	503

2.3 模型评估指标

在本研究中,我们采用了多个性能指标来评估所提出模型的效果,包括平均精度 (mAP)、帧率、模型占用存储空间大小 (model size)、和浮点

运算次数 (FLOPs)。

平均精度 (mAP) 是衡量目标检测模型性能的常用指标,它考虑了不同置信度阈值下的精确度 (P) 和召回率 (R),反映了模型整体的检测精度。

$$mAP = \frac{1}{n} \sum_{i=1}^n \left(\int_0^1 P(R) dR \right) \tag{8}$$

帧率:表示模型每秒能够处理的帧数,用于评估模型在实际应用中的检测速度。帧率越高,模型的实时性能越好。

模型占用存储空间大小 (model size):表示模型在磁盘上所占用的空间大小,单位为 MB。这一指标反映了模型的轻量化程度。

浮点运算次数 (FLOPs):表示模型进行一次前向推理所需的浮点运算次数,用于衡量模型的计算复杂度。FLOPs 越低,模型的运算效率越高。

通过这些评价指标,我们可以全面地了解模型在精度、效率和轻量化方面的性能,从而为模型的优化和应用提供依据。

2.4 消融实验比较

为验证本文算法改进的有效性,以原始模型 YOLOv8 为基线模型,并以 mAP@ 0.5、帧率、model size、FLOPs 作为评价指标,对模块顺序添加,“√”表示在模型中使用了该方法,“×”表示未使用该方法,可以观察到在引入了 DyHead 检测头后,精度得到较大幅度的提升,但计算成本增加较多,并使帧率降低。再通过 DCNv3 模块替换骨干网络中的 C2F 模块后,有效的降低了计算成本, FLOPs 降低 1.3,并略微增加了精度。最后将损失函数改进为 QFL 后,在对计算成本没有太大影响下,进一步提升检测精度。最后本文改进算法对比原 YOLOv8 网络精度提升 2.8%,帧率相差不多,虽然 FLOPs 有少量增加,但是整体精度提升较大,在精度和计算成本中找到了很好的平衡点,在具有高检测精度的同时拥有较高的计算效率。消融实验结果如表 3 所示。

表 3 消融实验结果

DyHead	DCNv3	QFL	mAP@ 0.5/ %	帧率/ (帧/s)	model size/MB	FLOPs/ GFLOPs
×	×	×	95.1	138.9	6.0	8.8
√	×	×	97.0	92.7	6.2	10.8
√	√	×	97.2	131.6	5.8	9.5
√	√	√	97.9	133.6	5.8	9.5

2.5 不同算法性能对比实验

为验证改进算法的有效性以及更好的评估性能,本文引入另外 4 种目标检测算法 FasterR-CNN^[10-11]、SSD^[12]、YOLOv3、YOLOv5,并使用 model size、FLOPs、mAP 和帧率 4 个指标来评估性能,实验结果如表 4 所示。可以看出,本文改进后的模型,在各个指标下都表现出了优异的性能。

表 4 本文改进算法与各种算法性能对比

网络模型	mAP/%	帧率/ (帧/s)	model size/MB	FLOPs/ GFLOPs
FasterR-CNN	93.9	21	108	336
SSD	89.3	47	92	61
Yolov3	89.7	49	32	58
Yolov5	93.5	88	15.4	15.8
Yolov8	95.1	138.9	6.0	8.8
改进模型	97.9	133.6	5.8	9.5

2.6 改进后的算法展示

原本在 YOLOv8 中出现误检、错检、漏检的目标,在本文改进后的算法中得到很好的检测。改进后的算法可以检测到更多的缺陷而且错检率极低。综上分析,本文改进后模型漏检率低,误检率低,准确率高。

3 结语

本研究成功地将 YOLOv8 算法改进应用于 PCB 缺陷检测领域,通过引入 DCNv3 卷积、DyHead 检测头,以及使用 QFL 对损失函数进行改进,显著提高了模型对小缺陷场景的适应性和识别准确率。通过消融实验结果验证了改进策略的有效性,并与其他目标检测算法进行对比,整体优于其他算法,为 PCB 缺陷检测提供了一种高效、可靠的深度学习解决方案。未来工作将进一步探索算法的优化和应用,以满足工业生产中对高精度自动化检测的需求。

参考文献:

[1] 吴一全,赵朗月,苑玉彬,等. 基于机器视觉的 PCB 缺陷检测算法研究现状及展望[J]. 仪器仪表学报, 2022,43(8):1-17.

[2] 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 (CVPR). Las Vegas, NV, USA. IEEE: 2016:779-788.

[3] TERVEN J, CORDOVA-ESPARZA D. A comprehensive review of YOLO: from YOLOv1 to YOLOv8 and beyond [J]. arXiv preprint arXiv:2304.00501, 2023.

[4] 范国华. 基于旋转目标检测的 PCB 加工图样缺陷检测[J]. 光电子技术,2021,41(3):203.

[5] 何国忠,梁宇. 基于卷积神经网络的 PCB 缺陷检测[J]. 图学学报,2022,43(1):21-27.

[6] DAI J F, QI H Z, XIONG Y W, et al. Deformable convolutional networks [C]//2017 IEEE International Conference on Computer Vision (ICCV). Venice, Italy: IEEE,2017:764-773.

[7] WANG W H, DAI J F, CHEN Z, et al. InternImage: exploring large - scale vision foundation models with deformable convolutions [C]//2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Vancouver, BC, Canada: IEEE, 2023: 14408-14419.

[8] DAI X Y, CHEN Y P, XIAO B, et al. Dynamic head: unifying object detection heads with attentions [C]// 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Nashville, TN, USA: IEEE,2021:7369-7378.

[9] LI X, WANG W, WU L, et al. Generalized focal loss: learning qualified and distributed bounding boxes for dense object detection [C]// Advances in Neural Information Processing Systems, 2020: 21002-21012.

[10] 孙书魁,范菁,孙中强等. 基于深度学习的图像数据增强研究综述[J]. 计算机科学, 2024, 51(1): 150-167.

[11] CHEN X, GUPTA A. An implementation of faster rcnn with study for region sampling [J]. arXiv preprint arXiv:1702.02138, 2017.

[12] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single shot MultiBox detector[M]// Computer Vision-ECCV 2016. ChamSpringer International Publishing 2016: 21-37.

收稿日期:2024-02-23