

YOLO-POD: 基于多维注意力机制的高精度 PCB微小缺陷检测算法

郭艳¹, 王智文^{2*}, 赵润星¹

(1. 广西科技大学自动化学院, 广西柳州 545616; 2. 广西科技大学电子工程学院, 广西柳州 545006)

摘要: 随着电子设备的广泛应用, 印刷电路板(Printed Circuit Board, PCB)在电子制造行业中具有重要意义。然而, 由于制造过程中的不完美和环境因素的干扰, PCB上可能存在微小的缺陷。因此, 开发高效准确的缺陷检测算法对于确保产品质量至关重要。针对PCB微小缺陷检测问题, 本文提出了一种基于多维注意力机制的高精度PCB微小缺陷检测算法。为降低网络的模型参数量和计算量, 引入部分卷积(Partial Convolution, PConv), 将ELAN(Efficient Layer Aggregation Network)模块设计为更加高效的P-ELAN, 同时, 为增强网络对微小缺陷的特征提取能力, 引入多维注意力机制(Multi-Dimensional Attention Mechanism, MDAM)的全维动态卷积(Omni-dimensional Dynamic Convolution, ODConv)并结合部分卷积, 设计了POD-CSP(Partial ODConv-Cross Stage Partial)和POD-MP(Partial ODConv-Max Pooling)跨阶段部分网络模块, 提出了OD-Neck结构。最后, 本文基于(You Only Look Once version 7, YOLOv7)提出了对小目标更加高效的YOLO-POD模型, 并在训练阶段采用一种新颖的Alpha-SIoU损失函数对网络进行优化。实验结果表明, YOLO-POD的检测精确率和召回率分别达到了98.31%和97.09%, 并在多个指标上取得了领先优势, 尤其是对于更严格的(mean Average Precision at IoU threshold of 0.75, mAP75)指标, 比原始的YOLOv7模型提高28%。验证了YOLO-POD在PCB缺陷检测性能中具有较高的准确性和鲁棒性, 满足高精度的检测要求, 可为PCB制造行业提供有效的检测解决方案。

关键词: 印刷电路板; 小目标缺陷检测; POD-CSP; POD-MP; 全维动态卷积; 多维注意力机制

基金项目: 国家自然科学基金(No. 61962007, No. 62266009); 广西自然科学基金重点项目(No. 2018GXNSF-DA294001); 广西财经大数据重点实验室开放基金(No. FEDOP2022A06)

中图分类号: TP391.4; TN41 **文献标识码:** A **文章编号:** 0372-2112(2024)07-2515-14

电子学报 URL: <http://www.ejournal.org.cn>

DOI: 10.12263/DZXB.20230772

YOLO-POD: High-Precision PCB Tiny-Defect Detection Algorithm Based on Multi-Dimensional Attention Mechanism

GUO Yan¹, WANG Zhi-wen^{2*}, ZHAO Run-xing¹

(1. School of Automation, Guangxi University of Science and Technology, Liuzhou, Guangxi 545616, China;

2. School of Electronic Engineering, Guangxi University of Science and Technology, Liuzhou, Guangxi 545006, China)

Abstract: With the widespread application of electronic devices, printed circuit boards (PCB) hold significant importance in the electronics manufacturing industry. However, due to imperfections in the manufacturing process and interference from environmental factors, tiny defects may in PCB. Therefore, the development of efficient and accurate defect detection algorithms is crucial in ensuring product quality. To address the challenge of detecting tiny defects on PCB, this paper proposes a high-precision PCB tiny defect detection algorithm based on multi-dimensional attention mechanism. To reduce model parameters and computational complexity, partial convolution (PConv) is introduced, and the ELAN module is redesigned as the more efficient P-ELAN. Additionally, to enhance the network's feature extraction capability for tiny defects, the omni-dimensional dynamic convolution (ODConv) based on the multi-dimensional attention mechanism (MDAM) is introduced. By combining partial convolution, the POD-CSP (Partial ODConv-Cross Stage Partial) and POD-MP (Partial ODConv-Max Pooling) cross-stage partial network modules are designed, along with the OD-Neck structure. Finally, based on YOLOv7, a more efficient YOLO-POD model for small object detection is proposed, and the network is op-

timized during the training phase using a novel loss function called Alpha-SIoU. Experimental results demonstrate that YOLO-POD achieves a detection precision of 98.31% and recall rate of 97.09%, exhibiting substantial advantages across multiple metrics. Notably, it achieves a 28% improvement over the original YOLOv7 model, as to more stringent mAP75 metric. These results validate the high accuracy and robustness of YOLO-POD in PCB defect detection, fulfilling the requirements for high-precision detection and providing an effective detection solution for the PCB manufacturing industry.

Key words: PCB; tiny-defect detection; partial odconv-cross stage partial; partial odconv-max pooling; omni-dimensional dynamic convolution; multi-dimensional attention mechanism

Foundation Item(s): National Natural Science Foundation of China (No.61962007, No.62266009); Natural Science Foundation of Guangxi Province (No.2018GXNSFDA294001); Guangxi Key Laboratory of Big Data in Finance and Economics (No.FEDOP2022A06)

1 引言

近年来,随着电子设备的快速发展和广泛应用,印制电路板(Printed Circuit Board, PCB)作为电子产品的重要组成部分,在电子制造行业中扮演着至关重要的角色.然而,在PCB的制造过程中,由于制造设备、材料质量以及环境因素等多种原因,常常会出现微小的缺陷,如漏孔、鼠咬、断路、短路、毛刺和余铜等.因此,PCB缺陷检测成为电子产品生产的重要环节,小目标的检测能力对PCB缺陷检测来说非常重要^[1].由于优秀的工艺,精确的布线和集成电路的快速发展,PCB变得越来越集成化和小型化.这些微小缺陷虽然在肉眼很难察觉,但却可能导致电子产品的性能下降甚至出现故障,严重影响产品的质量和可靠性.为了及时发现和修复这些微小的缺陷,高效准确的缺陷检测算法成为了刻不容缓的需求.目前,深度学习在计算机视觉领域取得了显著的突破,并广泛应用于目标检测任务.其中,基于YOLO(You Only Look Once)系列算法^[2]的目标检测方法以其快速、准确的特点受到了广泛关注.然而,传统的YOLOv7算法在PCB微小缺陷检测上存在一些不足之处.首先,YOLOv7对于微小缺陷的感知能力较弱,往往无法精确地检测到微小缺陷.其次,YOLOv7在处理PCB缺陷特点上的适应性不足,容易出现误检和漏检等问题.对于PCB缺陷检测,近几年基于深度学习的方法在检测精度mAP50上已经实现极高的准确率,但是经过实验测试发现,仅仅考虑该指标的精度是不合理的.PCB中的缺陷大小,占整幅图像很小的一部分,属于小目标检测.主流的一阶段检测框架,虽然可以在mAP50上达到很高的指标,但是在较为严格的指标上(如mAP50:95、mAP75)仍很低,说明仍有较大的改进空间.

针对以上问题,本文提出了一种基于多维注意力机制改进YOLOv7的PCB微小缺陷检测算法,多维注意力机制作为一种有效的特征加权方法,可以自适应地增强模型对感兴趣区域的特征表示.通过引入多维注意力机制,本文旨在提高YOLOv7在微小缺陷检测任务

中的性能,增强模型的感知能力和拟合能力,以满足工业领域对高精度检测的要求.通过大量的实验验证了该模型对PCB微小缺陷检测的有效性.本文的主要贡献如下:

(1)为降低网络的模型参数量和计算量,本文在网络中引入PConv,将ELAN(Efficient Layer Aggregation Network)模块设计为更加高效的P-ELAN.在不降低特征提取能力的前提下,提高了推理效率.

(2)针对PCB缺陷检测中的微小目标,本文使用了一种基于多维注意力机制的ODConv(Omni-dimensional Dynamic Convolution),来增强网络对PCB微小缺陷的特征提取能力.结合部分卷积,通过一次聚合的方法设计了跨阶段部分网络模块POD-CSP(Partial ODConv-Cross Stage Partial)和POD-MP(Partial ODConv-Max Pooling).并使用ODConv进行连接形成更加高效的OD-Neck结构.

(3)本文基于YOLOv7重新设计出对小目标更加高效的YOLO-POD模型来检测PCB缺陷.在训练阶段将AlphaIoU与SIoU结合,通过Alpha-SIoU损失对网络进行优化,实验表明网络对于mAP指标有着较大的性能提升.

2 相关工作

在PCB微小缺陷检测方面,已有许多研究工作致力于提高检测算法的准确性和效率.本节将对相关工作进行综述,包括传统方法和深度学习方法的应用.传统的PCB缺陷检测采用人工视觉检测方法和机器检测方法.常用的PCB缺陷检测方法包括基于规则和特征提取的方法.基于规则的方法通常使用预定义的规则或者阈值来判断是否存在缺陷,但这种方法对于复杂缺陷的检测效果有限.另一类方法是基于特征提取的方法,通过提取图像的纹理、形状和颜色等特征^[3],使用机器学习算法进行分类和检测^[4].然而,由于传统方法对于图像特征的设计和选择高度依赖于领域专家经验,且往往难以充分表达PCB微小缺陷的复杂特征,其检测性能有限.

随着深度学习的兴起,基于深度学习的方法在 PCB 缺陷检测领域取得了显著的进展^[5]。目前,基于深度学习的目标检测算法可以分为单阶段网络和两阶段网络。单阶段网络包括(Single Shot Detector, SSD)^[6]和 YOLO^[7] 算法家族(YOLO9000^[8]、YOLOv3^[9]、YOLOv4^[10]、YOLOv5、YOLOv6^[11]、YOLOv7^[12])。相比之下,两阶段网络包括卷积神经网络区域(Region-CNN, R-CNN)^[13]系列算法(Fast R-CNN^[14]和 Faster R-CNN^[15], Mask R-CNN^[16]等),后者是 R-CNN 的改进版本。这两种网络的主要区别在于,单阶段网络在特征提取后直接预测网络中缺陷的位置和类别,而两阶段网络首先生成可能包含缺陷的提案,然后进行检测过程。具体来说,两阶段网络生成可能包含缺陷特征不同大小的候选框,然后执行目标检测以预测缺陷类别和位置。但是由于产生的候选框太多,检测速度较慢。另一方面,单阶段网络在单个网络中同时进行训练和检测,不需要明确的区域建议,因此检测速度更快。

由于计算技术的进步,计算机能够更容易处理复杂的操作,导致神经网络的快速发展,其中包括大量的深度神经网络。在 PCB 缺陷检测领域,许多学者已经应用了深度学习技术。Huang 等^[17]通过密集连接所有前后层以实现特征重用,以更少的参数和计算成本实现了更好的性能。Huang 等^[18]通过设计一种前馈连接各层的卷积神经网络,提高了检测精度和效率。与传统的机器视觉方法相比,深度学习算法具有更强的非线性能力、更高的鲁棒性,适用于更复杂的场景。Che 等^[19]使用 ResNet50 作为主干网络,网络的检测准确率提高到 96.65%。Ding 等^[20]设计了专门针对 PCB 微小缺陷的检测网络 TDD-net,该网络采用多尺度融合策略,并采用在线硬样本挖掘,来提高 ROI 建议的确定性,检测准确率达到 98.90%。Peng 等^[21]设计了基于 CNN 的增强 Inception-ResNet-v2 模型。该模型建立基于全局信息比对信道进行滤波和连接通道的信道滤波模块,通过数据增强、批量归一化和学习率调整,准确率可达 94.8%。Chen 等^[22]提出了一种轻量级模型,采用 AI 深度学习模型对(Automated Optical Inspection, AOI)机检测到的缺陷产品进行重新判断,降低 AOI 机的误判率。Xin 等^[23]通过分析 CSPDarkNet53 结构层的特征分布,在数据预处理和输入阶段,自动细分图像,提高锚点中包含检测目标的概率。改进的 YOLOv4 算法 mAP 提高到 96.88%。Wang 等^[24]提出了一种轻量级模型,在 YOLOv5 骨干中使用 ShuffleNetV2 结构,准确率达到 95%。

其中,YOLOv7 作为一种经典的单阶段目标检测算法,广泛应用于多个领域。YOLOv7 算法通过将目标检测任务转化为回归问题,实现了端到端的检测与定位。然而,在 PCB 微小缺陷检测问题上,传统的 YOLOv7 算

法存在一些局限性。首先,由于 PCB 微小缺陷的尺寸小且表现复杂,传统 YOLOv7 算法往往无法精确检测。其次,传统 YOLOv7 对于 PCB 图像特征的建模能力不足,对于复杂的缺陷表现容易出现误检测和漏检等问题。为了克服这些问题,近年来,研究人员提出了各种改进的深度学习算法来提高 PCB 微小缺陷检测的准确性和鲁棒性。Yang 等^[25]提出了 SwinV2_TDD 模块,将激活函数改为 Mish,并引入(Magnification Factor Shuffle Attention, MFSA)机制,平均精度达到了 98.74%。Wu 等^[26]结合多种数据增强方法,建立 DA-YOLOv7 模型,为复杂条件下作物的检测和收获提供了理论参考。Chen 等^[27]提出了一种 SAS-FPN 模块,结合了注意机制、属性卷积以及 SiLU 损失函数,使模型能够关注重要信息,忽略无关信息,AP 值提高到 98.60%。Wang 等^[28]将 YOLOv7tiny 主干网络中的传统卷积模块和 ELAN 模块替换为可变形卷积,将主特征提取网络与可变形卷积融合,并引入 SE(Squeeze and Excitation)注意机制,提高其在复杂环境中提取小缺陷特征的能力。Liu 等^[29]采用 k-means++ 算法代替 K-means 获取锚盒,提高模型精度。

尽管在 PCB 微小缺陷检测领域的研究已经取得了一些进展,但仍存在着一些挑战和问题。首先,由于 PCB 微小缺陷的多样性和复杂性,如何提取有效的特征、建模缺陷的特征表示仍然是一个亟待解决的问题。其次,在标注数据方面,由于 PCB 缺陷数据获取较为困难和耗时,标注数据的数量和质量对于算法的性能起着重要作用。为了克服这些挑战和问题,本文提出了一种基于多维注意力机制改进 YOLOv7 的 PCB 微小缺陷检测算法。通过多种数据增强方式和训练优化策略,本文旨在提高 PCB 微小缺陷检测算法的准确性和鲁棒性。接下来,将详细介绍本文提出算法的设计与实验结果及分析。

3 方法

3.1 整体网络结构

本文提出的 YOLO-POD 模型的总体框架如图 1 所示。在 YOLO-POD 模型中,首先为了降低模型的大小,引入 P-ELAN 结构对 YOLOv7 中原有的 ELAN 网络进行了改进。其次,对于 PCB 这种小目标的缺陷,通过多维注意力机制^[30]将传统卷积方式改为 ODCConv,精确地提取了小目标的特征信息。并结合 PConv 提出 POD-MP,重新构造了主干网络。最后,在颈部网络中,将 PConv 和 ODCConv 结合,通过一次聚合的方法设计了跨阶段部分网络模块 POD-CSP,并使用动态卷积进行连接形成更加灵活的 OD-Neck,加快了网络的收敛速度,同时有效提高了检测精度。

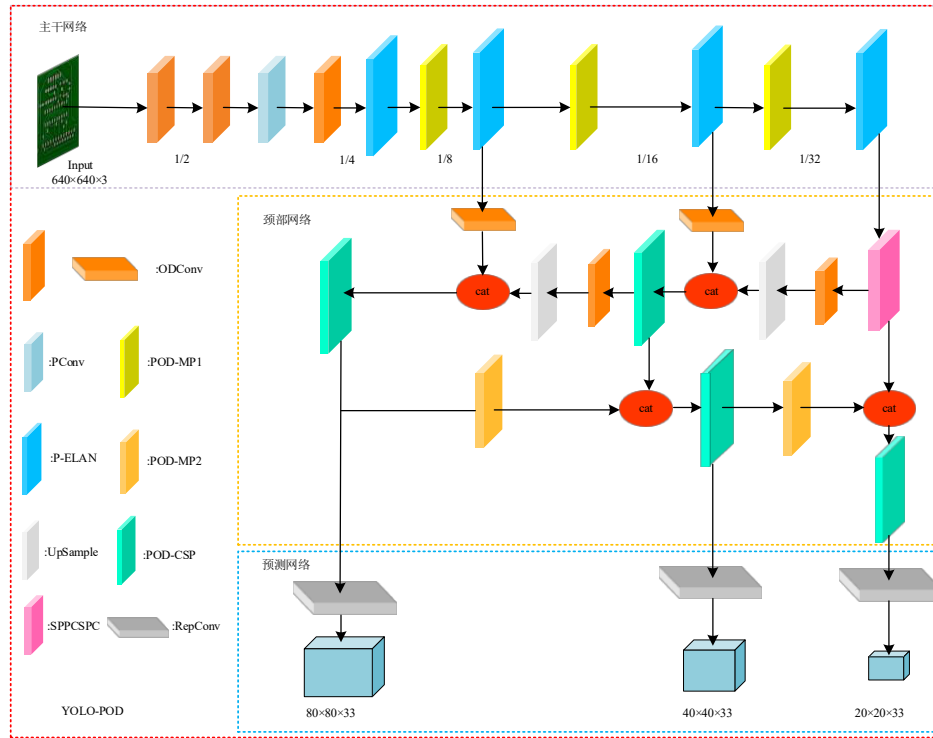


图1 网络结构图

3.2 多维注意力机制的全维动态卷积

由于PCB中的缺陷属于微小缺陷,在网络中采用普通的卷积方式很难精确地提取这些特征.普遍的解决方法是在网络中引入不同的注意力机制,虽然这些方法可以提高检测精度,但是有着一定的局限性^[31].空间注意力机制^[32]在目标检测中被广泛应用.它用于在图像或特征图的空间维度上进行加权处理,使模型能够更集中地关注图像中的特定区域或位置.这对于定位目标的位置以及抑制背景干扰非常有用,可以实现检测多尺度的目标.通道注意力机制^[33]用于调整输入数据在不同通道之间的权重,以便更集中地关注重要的通道信息.可以帮助模型学习到更具有区分性的通道表示,从而提高目标分类和检测的性能.动态卷积^[34]将注意力机制应用于卷积核中学习 n 个卷积核的线性组合,并对其输入依赖的关注进行加权,可以有效地提升推理精度.但是这些注意力机制仅仅考虑了单个维度,没能从不同角度去关注,效果会受到一定限制.

常规的卷积层有一个静态卷积核,它应用于所有的输入样本,用式(1)来计算:

$$y = g(\mathbf{W}^T \mathbf{x} + \mathbf{b}) \quad (1)$$

其中, \mathbf{W} 和 \mathbf{b} 为固定权重矩阵和偏置向量.对于动态卷积,它使用 n 个卷积核的线性组合,不同的输入通过权重 $\tilde{\mathbf{W}}^T(\mathbf{x})$ 和偏置 $\tilde{\mathbf{b}}(\mathbf{x})$ 不断变换来实现注意力机制动态加权,用式(2)来计算:

$$y = g(\tilde{\mathbf{W}}^T(\mathbf{x})\mathbf{x} + \tilde{\mathbf{b}}(\mathbf{x})) \quad (2)$$

全维动态卷积是一种新的动态卷积设计^[35].ODConv利用一种多维注意力机制,它弥补了单一的卷积核维度的局限性,通过沿核空间的四个维度(卷积核数、卷积核的空间大小、输入通道数和输出通道数)并行学习卷积核的不同注意,并逐步将这些注意应用到相应的卷积核上,可以大大增强CNN卷积操作的特征提取能力,有效提高小目标检测的多维注意力机制.ODConv的结构如图2所示,首先通过全局平均池化将输入 \mathbf{x} 压缩成长度为 c_{in} 的特征向量,再通过(Fully Connected,FC)层和ReLU激活函数以 $r=1/16$ 的映射率来降低模型的复杂度.之后分别映射到四个注意力的分支头,每个分支都有一个FC层和Sigmoid或Softmax激活函数来分别生成输入通道权值、卷积核权值、核空间权值以及输出通道权值.

根据式(2)中动态卷积的定义,本文重新定义和解释权重 $\tilde{\mathbf{W}}^T(\mathbf{x})$ 和偏置 $\tilde{\mathbf{b}}(\mathbf{x})$,ODConv四种维度的权值用式(3)和式(4)计算:

$$\tilde{\mathbf{W}}(\mathbf{x}) = \sum_{k=1}^K \pi_{wk}(\mathbf{x}) \cdot \pi_{fk}(\mathbf{x}) \cdot \pi_{ck}(\mathbf{x}) \cdot \pi_{sk}(\mathbf{x}) \tilde{\mathbf{W}}_k \quad (3)$$

$$\tilde{\mathbf{b}}(\mathbf{x}) = \sum_{k=1}^K \pi_{wk}(\mathbf{x}) \cdot \pi_{fk}(\mathbf{x}) \cdot \pi_{ck}(\mathbf{x}) \cdot \pi_{sk}(\mathbf{x}) \tilde{\mathbf{b}}_k \quad (4)$$

其中, $\pi_{wk}(\mathbf{x})$ 为softmax函数为整个卷积核分配的的关注权值; $\pi_{fk}(\mathbf{x})$ 、 $\pi_{ck}(\mathbf{x})$ 和 $\pi_{sk}(\mathbf{x})$ 代表sigmoid函数; π_k 为第 k

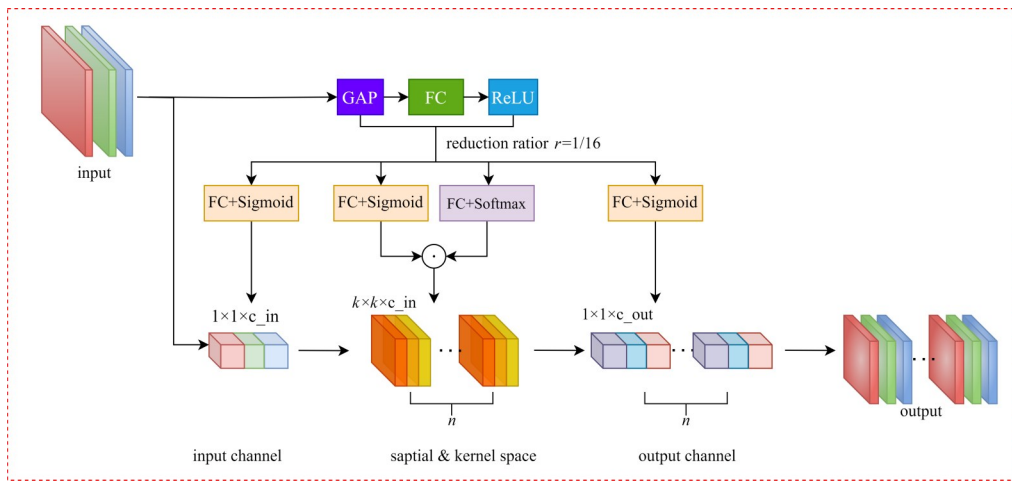


图2 多维注意力机制的ODConv结构图

个线性函数 $\tilde{W}_k^T x + \tilde{b}_k$ 的注意权值。 $\pi_{fk}(x)$ 和 $\pi_{ck}(x)$ 分别为卷积滤波器的 c_{out} 和 c_{in} 通道分配不同的权值； $\pi_{sk}(x)$ 在 $k \times k$ 个空间位置为卷积参数分配不同的注意权值。总权重 $\tilde{W}(x)$ 和偏置 $\tilde{b}(x)$ 是输入的函数，其中四个不同的权值不是固定的，而是随着输入 x 的不同而变化，它们代表了给定输入下线性模型 $\{\tilde{W}_k^T x + \tilde{b}_k\}$ 的最优集合。这四种类型分别以位置、通道、滤波器和核的顺序将它们累加到卷积核 W_i 中，使得卷积操作在所有空间位置、输入通道、滤波器和输入上都是不同的，为捕获丰富的上下文线索提供了性能保证。因此，ODConv 可以显著增强 CNN 基本卷积操作的特征提取能力。在网络中使用 ODConv，在降低网络参数的前提下，提高了网络的特征学习能力，对于小目标缺陷特征提取更加友好，可以有效提高检测精度。

3.3 更高效的卷积

本文引入 FasterNet^[36] 中的 PConv。由于不同通道之间的特征图有很高的相似性，传统的卷积方式提取特征，会造成计算和内存的冗余。如图 3 所示，PConv 利用特征映射中的冗余性^[37]，对少数几个输入通道应用标准的卷积，而其余的通道保持不变。通过该方法可以保证和标准卷积具有相同特征提取能力的前提下，提

高推理速度。

为了可以连续的访问内存，可以将通道中的第一个或最后一个连续的 c_p 通道作为特征映射来计算，其余的 $c - c_p$ 通道则保持不变。因此在不损失通用性的情况下，输入和输出特征映射具有相同数量的通道。可以用式(5)得出 PConv 的 FLOPs：

$$FLOPs = h \times w \times k^2 \times c_p^2 \quad (5)$$

通过设置映射部分比 r 可以改变 FLOPs 的大小，本文采用映射部分比为 $c_p : c = 1 : 4$ ，因此可以得出 PConv 的 FLOPs 为普通卷积的 1/16。同时，通过式(6)可计算出 PConv 的内存访问量(MAC)约为普通卷积的 1/4：

$$MAC = h \times w \times 2c_p + k^2 \times c_p^2 \approx h \times w \times 2c_p \quad (6)$$

通过上式的推理分析，可以知道 PConv 获得比常规卷积有着更低的 FLOPs，实现更快的推理速度。其原理类似于 GroupConv，只对部分组进行操作。但 PConv 的操作未影响其他组的通道信息，可以防止运算符过度访问内存，进而提高计算效率。因此，这种新颖的 PConv 可以作为一种有竞争力的替代方案，它减少了计算冗余以及内存访问的数量^[38]，更有效地提取空间特征。

3.4 主干网络设计

YOLOv7 网络中采用了高效的 ELAN 结构，该结构通过控制最短和最长的梯度路径使网络学习更多的特征，具有较强的鲁棒性和收敛速度。本文在 YOLOv7 原始的 ELAN 结构上将最长的分支结构的四个普通卷积替换为 PConv，提出新的 P-ELAN。与原始 ELAN 相比，P-ELAN 具有更快的速度，且不会影响准确性。其模块结构如图 4 所示。

在下采样部分，通常采用一个卷积核大小为 3×3 、步长为 2 的卷积或者一个步长为 2 的最大池化。这两种方法，不能适应不同的尺度变化，会丢失部分小尺度的

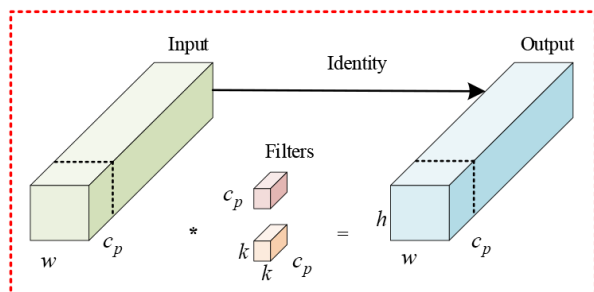


图3 Partial Conv示意图

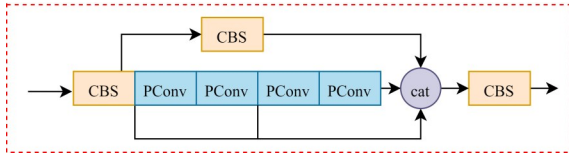


图4 Partial-ELAN(P-ELAN)模块结构图

特征以及位置信息。YOLOv7采用了这两种组合方式来进行过渡,虽然该方法有助于缓解上述不足,但是通过普通的卷积方式来处理小目标的特征无法达到自适应的目的,会损失其特征的细节。为此通过ODConv和PConv的结合,本文设计出一种POD-MP过渡结构,来提高下采样的特征精度。其模块结构如图5所示。

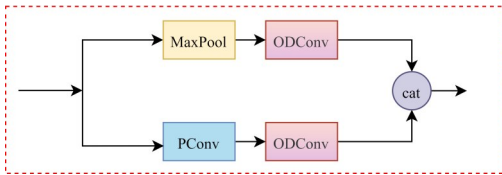


图5 POD-MP模块结构图

为了能更好地捕获PCB中的缺陷信息,本文将P-ELAN与POD-MP应用在原始的YOLOv7结构中,重新设计了一种适应于微小目标的主干网络。该主干网络结合了多维注意力机制的ODConv和PConv的优势,网络不仅减少了参数和计算量,而且增强了小目标特征的提取能力。

3.5 颈部网络设计

在YOLO系列的算法中,不同的特征大小的分支需要通过模块进行融合,虽然不同网络所用的结构有所不同,但其本质上是属于CSP结构。CSP结构通过将不同分支的特征图分割和融合,有效地增加了网络的感知能力和表达能力,更好地学习到特征,从而提高模型在复杂任务上的性能。尽管CSP结构作为一种有效的特征融合^[39]方法,但其在通道维度上的信息融合,可能无法充分利用特征图的空间信息。在某些情况下,特定的局部信息可能更有利于目标检测,并且特征图的分割和融合可能会造成边界信息的丢失。这不利于目标检测任务,因为目标的准确边界信息对于检测框的精确定位很重要。尤其是对于小目标检测,无法充分利用这些信息。因此本文通过在CSP结构中引入ODConv和PConv,命名为POD-CSP,来弥补原始的CSP结构的不足,其模块结构如图6所示。

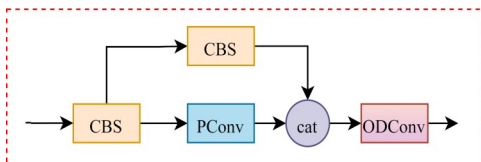


图6 POD-CSP模块结构图

采用POD-CSP和POD-MP结构,并使用动态卷积进行连接形成更加灵活的OD-Neck结构。来自不同分支的特征拼接后,会经过POD-CSP模块进行充分的融合。在两个下采样阶段使用POD-MP模块来弥补信息的丢失,有效提高了检测精度。

3.6 Alpha-SIoU损失

YOLOv7在训练时采用了CIoU损失函数进行优化。虽然CIoU、GIoU、DIoU这些损失函数常用于模型的训练优化,但它们依赖于边界框回归的度量聚合(预测框与真实框的距离、重叠面积和长宽比)。而SIoU损失函数^[40]考虑了两者之间的方向不匹配,并重新定义了乘法项,这加快了网络的收敛。SIoU损失包括角度代价、距离代价、形状代价和IoU代价。角度代价用式(7)来计算:

$$A = 1 - 2 \times \sin^2 \left(\arcsin \left(\frac{c_h}{\sigma} \right) - \frac{\pi}{4} \right) \quad (7)$$

$$\sigma = \sqrt{(b_{c_x}^{\text{gt}} - b_{c_x})^2 + (b_{c_y}^{\text{gt}} - b_{c_y})^2} \quad (8)$$

$$c_h = \max(b_{c_y}^{\text{gt}}, b_{c_y}) - \min(b_{c_y}^{\text{gt}}, b_{c_y}) \quad (9)$$

其中, c_h/σ 表示角度 α 的正弦值; (b_{c_x}, b_{c_y}) 和 $(b_{c_x}^{\text{gt}}, b_{c_y}^{\text{gt}})$ 分别为检测框和真实框的中心点坐标。通过式(8)可以计算出两个中心之间的距离 σ ,通过式(9)可以计算出两个中心之间的垂直距离 c_h 。

距离代价用式(10)和式(11)来计算:

$$\Delta = \sum_{l=x,y} (1 - e^{-\Delta \rho_l}) \quad (10)$$

$$\rho_x = \left(\frac{b_{c_x}^{\text{gt}} - b_{c_x}}{c_w} \right)^2, \rho_y = \left(\frac{b_{c_y}^{\text{gt}} - b_{c_y}}{c_h} \right)^2 \quad (11)$$

其中, ρ_x 和 ρ_y 分别为距离损失的平方。当角度 α 接近0时,会退化为距离损失;相反当角度 α 接近 $\pi/4$ 时,角度损失增大。

形状代价用式(12)和式(13)来计算:

$$\Omega = \sum_{w,h} (1 - e^{-\omega_{w,h}})^\theta \quad (12)$$

$$\omega_w = \frac{|w - w^{\text{gt}}|}{\max(w, w^{\text{gt}})}, \omega_h = \frac{|h - h^{\text{gt}}|}{\max(h, h^{\text{gt}})} \quad (13)$$

其中, w 和 h 为预测框的宽和高; w^{gt} 和 h^{gt} 为真实框的宽和高; ω_w 、 ω_h 和 θ 定义了形状的代价, θ 默认设置为4。最终可以用式(14)和式(15)来定义SIoU损失:

$$L_{\text{box}} = 1 - \text{IoU} + \frac{\Delta + \Omega}{2} \quad (14)$$

$$\text{IoU} = \frac{|B \cap B^{\text{GT}}|}{|B \cup B^{\text{GT}}|} \quad (15)$$

为了追求更高的检测精度,本文在SIoU的基础上加入了Alpha调制^[41],最终用式(16)来计算Alpha-SIoU。

$$L_{\alpha-SIoU} = 1 - (L_{IoU})^\alpha + \left(\frac{\Delta + \Omega}{2}\right)^\alpha \quad (16)$$

其中,提高 α 的大小可以增加高IoU对象的损失和梯度,从而可以提高边界框的回归精度.当 α 为1时,该公式退化为SIoU损失.

4 实验结果与分析

在本节中,描述了实验环境的配置、测试数据集和优化之后算法的实验结果分析以及与主流算法对比实验的结果.本文具体的实验软硬件环境配置如表1所示.

表1 实验软硬件环境配置

名称	相关配置
操作系统	Windows10 64位
CPU	Intel(R)Core(TM)i7-12700K CPU@3.60 GHz
GPU	NVIDIA GeForce RTX 3090 24 GB
Pytorch	1.10.0
CUDA 版本	CUDA 10.2

4.1 数据增强

本文实验采用北京大学智能机器人公开实验室的开源PCB缺陷检测数据集,该数据集包括六种常见缺陷,分别为漏孔、鼠咬、断路、短路、毛刺和余铜.该数据集包含693幅图像,不同缺陷示例图像如图7所示.数据集中的每幅图像包含3~5个微小缺陷,如图7中红色小框框出来的部分.

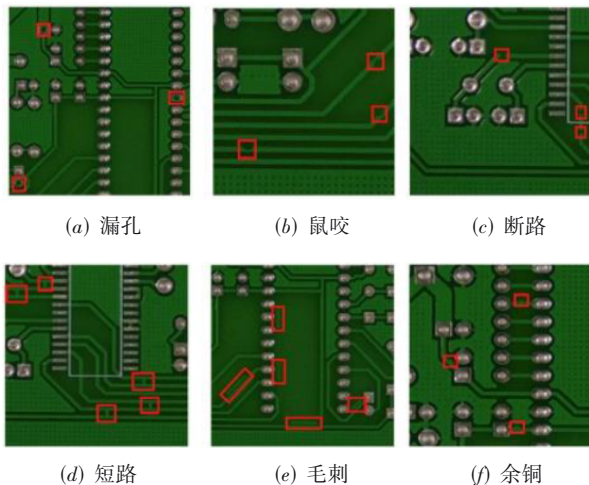


图7 不同缺陷的实例

由于原始的数据集样本量较少,不足以使网络有着高精度和泛化能力.因此在训练之前,数据增强是必不可少的.为了能使模型可以应对不同环境的干扰,提高鲁棒性,本文除了使用垂直、镜像、曝光、旋转这些简

单的数据增强方法外^[42],额外使用了高斯噪声、动态模糊、随机填充像素块等多种方式,并将这些方法进行了随机组合来生成更多的数据样本.数据增强的一些示例图像如图8所示.通过这些方法本文重新扩充并构建了新的数据集,总共9 222幅图像.新的数据集可以使网络有着更强的鲁棒性和泛化能力.

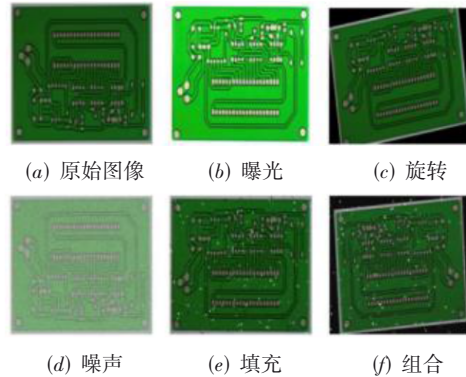


图8 数据增强效果实例

本文按照6:2:2的比例随机将其划分为训练集、测试集和验证集,其中5 534幅图像组成训练集,1 844幅图像组成测试集,1 844幅图像组成验证集.经过数据增强方法得到了PCB表面缺陷数据集中各个类别的缺陷数量如表2所示.

表2 数据增强前后的各类别缺陷数量

缺陷类型	原始图片数量	原始缺陷数量	增强缺陷数量
漏孔	115	497	5 330
鼠咬	115	492	5 288
断路	116	482	5 115
短路	116	491	5 398
毛刺	115	488	5 228
余铜	116	503	5 501
总计	693	2 953	31 860

4.2 模型评价指标

本文从精确率、召回率、 F_1 分数、平均精度等多个方面评价PCB表面缺陷检测网络的性能.在实验中,建立了IoU^[43]阈值.当IoU阈值越高,当前训练模型对于这个框的结果越认可.本文取IoU=0.5,如果预测框与真实框的IoU>0.5,那么此预测框归为TP,如果0<IoU≤0.5,那么此预测框归为FP,如果IoU<0,那么此预测框归为FN.因此,TP表示实际为正样本,检测正确;FP表示实际为负样本,检测错误;TN表示实际为负样本,检测正确;FN表示实际为正样本,检测错误.各个指标的具体计算如下.

(1) 精确率

精确率可定义为识别图像中真阳性样本的比例,

也叫查准率,即正确预测检测到PCB缺陷,以百分比表示,可用Precision(P)表示,精确度用式(17)来计算:

$$P = \frac{TP}{TP+FP} \times 100\% \quad (17)$$

(2) 召回率

召回率定义为测试集中识别正确目标的所有阳性样本的比例,也叫查全率,即预测为缺陷比实际缺陷,可用Recall(R)表示,召回率用式(18)来计算:

$$R = \frac{TP}{TP+FN} \times 100\% \quad (18)$$

(3) F_1 分数

F_1 分数(F_1 -score)是对精确率和召回率的整体评价, F_1 -score用式(19)来计算:

$$F_1 = \frac{2PR}{P+R} \quad (19)$$

(4) 平均精度和平均精度均值

仅凭 R 或 P 的值不能客观反映检测结果的质量. 因此,需要结合这两个评价指标来衡量算法的性能. 使用不同 P 和 R 值的点的组合可以绘制 P-R 曲线. 根据 P-R 曲线,计算每个 R 值对应的 P 值,PR 曲线下的面积就是平均精度(Average Precision, AP)的值,AP 的值越高表示分类器的性能越好. 平均精度均值(mean Average

Precision, mAP)表示所有类别 AP 的平均值,即对每个类的 AP 再求平均,得到的就是 mAP 的值, mAP 的大小一定在 $[0, 1]$ 区间,越大越好. 该指标是目标检测算法中最重要的一个.

AP 用式(20)来计算:

$$AP = \int_0^1 \text{Precision}(\text{Recall})d\text{Recall} \quad (20)$$

mAP 用式(21)来计算, C 是图像数据集中所有类别的数量.

$$mAP = \frac{1}{C} \int_0^1 \text{Precision}(\text{Recall})d\text{Recall} \quad (21)$$

4.3 消融实验

本文选择采用 Alpha-SIoU 作为最终网络训练的损失函数,通过在原始的 YOLOv7 模型上进行了多组实验. 表 3 展示了不同损失函数以及 α 值下的结果,通过表中数据可以得出使用 Siou 损失可以有效提高网络的精度. 使用 Alpha-SIoU 损失 mAP 可以进一步提高,但是网络的精确率却急剧下降. 通过分析可以知道, α 值增大了高 IoU 对象的损失和梯度,使得网络在预测时将大量的背景识别为正样本,导致精确率下降. 所以本文在下面的实验中选择最佳 $\alpha=1.5$ 的 Alpha-SIoU 损失函数进行训练.

表 3 损失函数优化

基准模型	损失函数	mAP50:95	mAP50	mAP75	F_1	$P/\%$	$R/\%$
YOLOv7	CIoU	0.464	0.950	0.348	0.935	98.01	89.63
	SIoU	0.488	0.956	0.411	0.950	97.69	92.33
	Alpha-SIoU($\alpha=1.5$)	0.516	0.957	0.486	0.935	89.73	98.41
	Alpha-SIoU($\alpha=2$)	0.522	0.956	0.499	0.920	87.51	98.42
	Alpha-SIoU($\alpha=2.5$)	0.523	0.956	0.502	0.880	80.21	98.73
	Alpha-SIoU($\alpha=3$)	0.500	0.944	0.450	0.680	54.05	98.89

为了验证本文提出的模型结构的有效性,本文使用 YOLOv7 作为基准模型进行消融实验,来评估本文提出的 YOLO-POD 在各模块下的表现效果. 通过控制增减各个模块,本文共进行了七组实验,其中第 1 组是在原始 YOLOv7 模型下的实验结果. 之后便在此基础模型上改进其部分结构进行测试,检验每一步的改进效果,消融实验结果如表 4、表 5 所示. 在第 2 组实验中,我们单独使用 P-ELAN 结构后各项 mAP 指标分别提高了 5.1%、2.2% 和 11.8%, F_1 和 R 分别提高了 3.3%、5.76%. 同时模型的计算量减少了 26.663G, FPS 达到了最快. 这是因为 P-ELAN 采用了更加高效的 PConv,可以有效降低模型的 FLOPs,提高推理速度. 在第 3 组实验中,我们将 MDAM 应用在主干网络上增强模型对特征的提取能力. 可以看到各项 mAP 指标分别提高了 10.4%、2.6% 和 25.4%, F_1 和 R 分别提高了 4.1%、7.27%. 采用多维注意力机制的卷积方式可以更

准确地捕获缺陷的特征信息. 在第 4 组实验中,我们在网络的 Neck 部分采用了本文提出的 OD-Neck 结构,各项 mAP 指标分别提高了 11.5%、2.8% 和 28.3%, F_1 和 R 分别提高了 4.1%、7.42%. 其中 mAP50:95、mAP75 和 P 达到了最好的效果,验证了 OD-Neck 有效性,但是模型的推理速度却不理想. 在第 5 组和第 6 组实验中,我们依次将各个模块应用在网络当中,可以看到模型的参数量和计算量进一步减少,推理速度取得了提升.

然而从第 6 组实验中可以看到, YOLO-POD 的各项指标却不如第 4 组实验中单独使用 OD-Neck 模块有效,特别是在 mAP75 指标上性能下降了 3.1%,这是因为原始模型中 CIoU 损失函数存在的局限性. 由表 3 的实验结果可以看到,在相同训练条件下, Alpha-SIoU 对 mAP75 和 Recall 指标有很大的效果提升. Alpha-SIoU 不仅增加高 IoU 对象的损失和梯度,也加快了回归速

表 4 消融实验:精度对比

序号	P-ELAN	MDAM	ODNeck	Alpha-SIoU	mAP50:95	mAP50	mAP75	F_1	P/%	R/%
1	×	×	×	×	0.464	0.950	0.348	0.935	98.01	89.63
2	√	×	×	×	0.515	0.972	0.466	0.968	98.32	95.39
3	×	√	×	×	0.568	0.976	0.602	0.976	98.44	96.90
4	×	×	√	×	0.579	0.978	0.631	0.976	98.51	97.05
5	√	√	×	×	0.568	0.978	0.602	0.975	98.16	96.69
6	√	√	√	×	0.566	0.978	0.600	0.976	98.31	96.93
7	√	√	√	√	0.576	0.980	0.629	0.977	98.31	97.09

表 5 消融实验:参数量与计算量对比

序号	P-ELAN	MDAM	ODNeck	Alpha-SIoU	FLOPs/G	Params/M	FPS/(f/s)
1	×	×	×	×	105.199	37.222	28
2	√	×	×	×	78.536	32.677	37
3	×	√	×	×	85.071	33.169	25
4	×	×	√	×	89.975	30.467	29
5	√	√	×	×	58.408	28.625	33
6	√	√	√	×	41.721	21.026	35
7	√	√	√	√	41.721	21.026	35

度,从而提高了边界框的回归精度.从第7组实验可以看到,Alpha-SIoU可以有效优化高精度检测模型的性能.最终优化的YOLO-POD模型的各项mAP指标相比原始的YOLOv7模型提高了11.2%、3%和28.1%.对于严格的mAP75指标有着很大的提升,同时 F_1 、精确率和召回率分别提高了0.3%、4.2%和7.46%.由表5可以得出,YOLO-POD的模型参数量大小为21.026 M,减少了43.5%;计算量为41.721 G,减少了60.3%;FPS为35 f/s,加快了7 f/s.验证了本文提出的YOLO-POD模型可以有效提高PCB微小缺陷的检测精度和速度.

在网络中分别使用P-ELAN、MDAM和OD-Neck模型的参数量和计算量都有所减小,精度并没有下降.这

是因为本文在网络中的多个模块使用了PConv,在保持网络提取特征能力的前提下,缩短了模型的推理时间,减少参数量和计算量.同时采用多维注意力机制的全维动态卷积,有效提取了微小目标缺陷的特征信息,有效地提高了检测精度.

4.4 算法对比

为了验证YOLO-POD模型检测PCB缺陷的优势,本文采用目前七种主流的目标检测网络模型,SSD、CenterNet、YOLOv4、YOLOv5、YOLOX、YOLOv7和YOLOv8进行性能对比实验与分析.所有深度学习检测算法采用相同的训练和测试数据集,模型的输入图像大小设置为640×640.不同算法在PCB缺陷检测中的各评价指标的对比实验结果如表6所示.

表 6 不同算法的性能对比

方法	mAP50:95	mAP50	mAP75	F_1	P/%	R/%	FLOPs/G	Params/M
SSD	0.460	0.945	0.344	0.840	98.70	74.32	61.254	24.280
CenterNet	0.450	0.930	0.336	0.892	97.93	82.24	70.217	32.665
YOLOX	0.487	0.974	0.372	0.970	97.48	96.56	26.768	8.940
YOLOv4	0.420	0.935	0.258	0.928	97.64	88.87	59.989	63.965
YOLOv5	0.495	0.952	0.426	0.945	98.40	91.22	16.520	7.077
YOLOv7	0.464	0.950	0.348	0.935	98.01	89.63	105.199	37.222
YOLOv8	0.533	0.965	0.514	0.927	95.67	98.90	28.817	11.167
YOLO-POD	0.576	0.980	0.629	0.977	98.31	97.09	41.721	21.026

与其他算法对比,本文提出的YOLO-POD在多个指标上取得的领先的优势.SSD虽然有着最高的精确率,但召回率却很低,说明有许多正样本存在漏检情况.YOLO-POD同时兼顾了精确率和召回率,因此有着较高的 F_1 值.虽然与其他算法对比在mAP50指标上都

有着较高的精度,在追求高精度的检测结果上需要考虑多个指标.YOLO-POD在mAP50:95指标上与最好的YOLOv8相比提高了4.3%,特别是在mAP75指标上提高了11.5%,有着较大的领先优势,说明模型可以满足高精度的检测要求.对比计算量和参数量,我们可以发

现 YOLO-POD 的计算量和参数量分别为 41.721 G、21.026 M, 大于 YOLOv5、YOLOX 和 YOLOv8, 但低于其他四种算法。虽然 YOLOv5 和 YOLOX 具有较小的参数量, 但其检测性能不如 YOLO-POD。YOLO-POD 在各项 mAP 指标上取得了最好结果, 特别是在较为严格的 mAP75 指标上, 这表明了本模型相对于其他模型在高精度检测上具有优越性。相较于原始的 YOLOv7 模型实现了更快的推理速度, 这足以满足在现实场景下的检测需求。

为了进一步分析验证本文提出的 YOLO-POD 算法对小目标缺陷的检测性能, 本文专门选择了针对小目

标检测模型的算法。对比了 RFLA^[44] (FCOS-RFLA、Faster R-CNN-RFLA)、DHRec^[45] (ATSS-DHRec、FCOS-DHRec) 和 S2A-Net^[46] 这三类小目标检测模型的五种算法。如表 7 所示, 为不同算法在 PCB 缺陷检测中的各评价指标的对比实验结果。YOLO-POD 在各项 mAP 指标都达到了最好的效果, 并且模型的参数量和计算量是最小的。这些小目标检测模型都是采用 ResNet50 作为主干网络, 主要对网络训练时标签和检测框分配的方式进行改进。因此模型的计算量和参数量没有明显的改变, 相反, 本文提出的方法主要针对模型结构进行优化, 在高精度检测的同时, 可以实现较快的推理速度。

表 7 不同小目标算法的性能对比

方法	mAP50:95	mAP50	mAP75	FLOPs/G	Params/M
FCOS-RFLA	0.551	0.923	0.598	78.71	31.85
Faster R-CNN-RFLA	0.564	0.959	0.603	78.14	41.15
ATSS-DHRec	0.535	0.967	0.531	80.92	31.93
FCOS-DHRec	0.563	0.926	0.623	80.18	32.02
S2A-Net	0.527	0.945	0.572	81.42	38.64
YOLO-POD	0.576	0.980	0.629	41.721	21.026

为更好体现出算法的检测结果, 本文统计了 YOLO-POD 在六种缺陷类别上的平均精度并与原 YOLOv7 进行对比, 实验对比结果如图 9 所示。本文改进的方法针对 PCB 六类缺陷的检测平均精度值分别为漏孔 99.91%、鼠咬 98.42%、断路 98.15%、短路 99.48%、毛刺 96.7%、余铜 98.73%, 相比原 YOLOv7 方法对于每类缺陷的检测精度都有一定的提升, 其中对于毛刺缺陷的检测精度提升最多, 平均精度值提升了 5.2%。由此可以得出, 所提出方法相比原 YOLOv7 方法, 在 PCB 六类缺陷的检测中, 对于 PCB 这类缺陷的微小性、位置不固定性, 以及在检测中容易产生漏检的情况, 均能够得到很好解决。

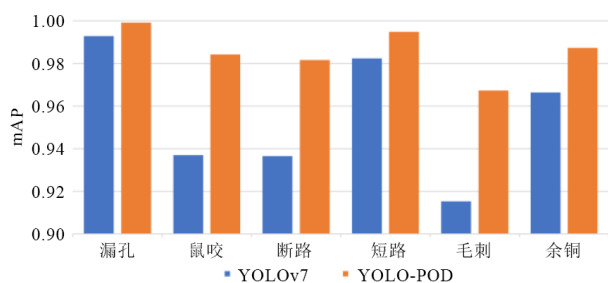


图 9 六种缺陷的平均精度对比图

实验结果显示, 本文的算法在微小缺陷检测方面表现出了优势。与其他算法相比, 本文的算法具有更高的 mAP 和召回率, 以及较低的误检率。这进一步验证了本文提出的基于多维注意力机制改进 YOLOv7 的 PCB 微小缺陷检测算法的有效性和优越性。综上所述,

所提出方法的综合性能是最优的, 能够在保持较高性能的同时, 减少参数量和计算量, 为后续满足工业现场的 PCB 缺陷检测提供方法基础。

本文实验采用了北京大学智能机器人公开实验室的开源 PCB 缺陷检测数据集, 因此, 我们针对这个数据集对不同 PCB 检测算法的性能进行了对比。从表 8 可以看出, 在近几年研究中有较多的方法都是基于 YOLO 系列算法进行了改进。本文在改进 YOLOv7 模型的同时, 采用多种数据增强的组合方式, 在较低训练样本的情况下提高了 PCB 缺陷的检测精度。

表 8 北京大学公共实验室 PCB 数据集的目标检测算法对比

算法	年份	基准模型	训练集图片数量	mAP50/%
文献[23]	2021	YOLOv4	—	96.88
文献[47]	2023	YOLOv5	6 653	95.97
文献[48]	2023	YOLOv7	8 534	97.50
本文	—	YOLOv7	5 534	98.00

4.5 PCB 微小缺陷检测可视化结果分析

为验证所提出方法的可行性, 本文开展了随机测试实验。从测试集的每类缺陷图像中各随机选取一幅图像, 利用所提出 YOLO-POD 网络模型和其他主流网络进行模拟检测。高效的 PCB 表面缺陷检测网络不仅要求预测精度高, 而且要求提取特征的有效性。本文可视化了鼠咬、余铜两个典型缺陷的 PCB 表面特征提取能力, 检测结果和可视化效果如图 10 所示。为了进一步验证 YOLO-POD 对 PCB 表面缺陷特征提取的真实效果, 本文在原始图像的基础上分别加入高斯噪声和椒

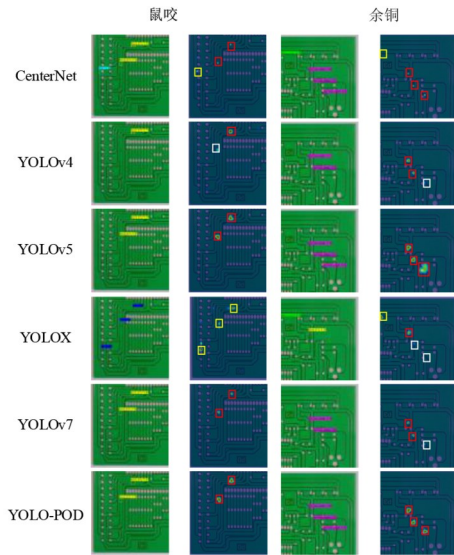


图 10 检测效果对比图

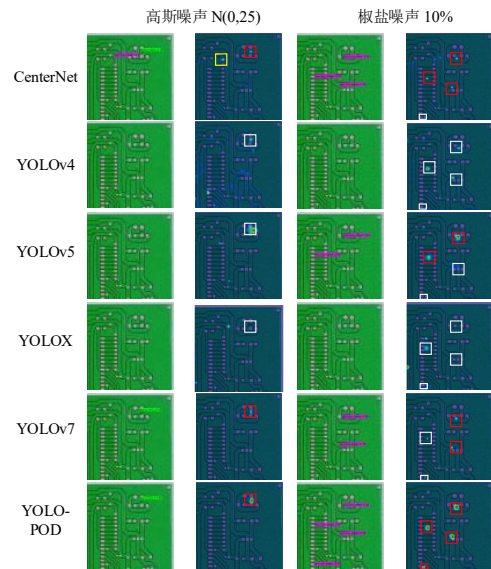


图 11 加入不同噪声的检测效果对比图

盐噪声,测试干扰噪声对不同模型性能的影响,检测结果和可视化效果如图 11 所示.

图像的特征提取可视化结果以热图的形式呈现,用暗红色突出的区域是网络预测缺陷图像类别的主要基础,颜色越深,说明效果越好,反之效果越差.在图 10 和图 11 的热力图中,用白色框表示漏检缺陷,黄色框表示误检缺陷,红色框表示真实缺陷.图 10 使用原始图像,图 11 使用分布为 $N(0, 25)$ 的高斯噪声和 10% 的椒盐噪声进行可视化分析.通过对比可以看到,在图 10 中,所提出的方法将所有缺陷全部检出,不存在漏检的情况,和真实标签值所标记的缺陷位置也全部相同.在鼠咬缺陷特征的提取中,CenterNet、YOLOv4、YOLOX 均出现误检或者漏检,YOLOv5、和 YOLOv7 虽然都能够检测到相应的缺陷,但提取的鼠咬特征颜色都不如 YOLO-POD 深.在图 11 加入不同噪声的情况下,本文所提出的方法依旧将所有缺陷全部检出,不存在误检、漏检的情况.在加入高斯噪声的特征提取中,CenterNet

算法出现了误检,YOLOv4、YOLOv5、YOLOX 均出现漏检的情况,虽然 YOLOv5 在特征提取中关注到缺陷的信息,但其明显受到背景信息的干扰,导致无法聚焦于缺陷位置.YOLOv7 虽然都能够检测到相应的缺陷,但提取的特征颜色比 YOLO-POD 更暗一些,导致检测的置信度较低.在加入椒盐噪声后,其他算法受到的背景干扰变得更为明显.CenterNet、YOLOv4、YOLOv5、YOLOX、YOLOv7 均出现漏检的情况.YOLO-POD 在一定程度上可以有效地抑制背景噪声的干扰,验证了 MDAM 对微小缺陷特征提取的有效性.总体而言,在 YOLO-POD 检测结果中,缺陷特征点的红色较为饱和,背景特征点较为一致.这些现象说明,YOLO-POD 在最大程度上抑制了背景的干扰,同时促进了对小缺陷特征的识别.可以看出 YOLO-POD 达到了最满意的 PCB 表面缺陷特征提取效果.不同算法在各个类别的误检率如表 9 所示.

表 9 不同算法在各个类别的误检率

方法	类别					
	漏孔	鼠咬	断路	短路	毛刺	余铜
SSD	0.01	0.20	0.20	0.08	0.22	0.13
CenterNet	0.03	0.19	0.22	0.09	0.23	0.20
YOLOv4	0.04	0.16	0.16	0.09	0.22	0.15
YOLOv5	0.01	0.18	0.14	0.03	0.14	0.10
YOLOX	0.01	0.08	0.06	0.03	0.11	0.05
YOLOv7	0.02	0.18	0.15	0.06	0.21	0.13
YOLO-POD	0	0.07	0.04	0.02	0.11	0.04

由表 9 可以看出,本文改进的方法针对 PCB 六类缺陷的误检率都有一定的降低,其中对于漏孔缺陷误检

率为 0,对于断路、鼠咬缺陷的误检率降低最为明显.相比之下,所提出方法提高了 PCB 缺陷检测中对于检测

缺陷目标的个数和位置的准确度,能更好地适应PCB缺陷的微小性和位置不固定性,验证了所提出的方法的可行性.

综上所述,本文提出的算法对检测网络进行了显著的改进,在PCB微小缺陷检测方面的各项评估指标上也均表现出优势,与传统方法和其他基准算法相比都取得了更好的结果.因此,本文的算法具有潜在的应用前景,可为PCB制造行业提供一种高效准确的缺陷检测解决方案.

4.6 算法在不同小目标数据集的性能分析

为验证YOLO-POD在不同小目标数据集上的检测效果,我们分别在VisDrone2019和TinyPerson数据集上对模型的性能进行了实验.VisDrone2019是由无人机在拥挤场景下拍摄的小目标数据集,主要标注了行人、不同种类交通工具共十种类型的目标.该数据集包含不同的光照场景,标注的目标较为密集.TinyPerson属于微小目标的行人数据集,主要分为海上行人和陆地行人两类.同时,我们采用目前主流的目标检测网络模型,YOLOv7、YOLOv8以及本文提出的YOLO-POD进行性能对比与分析.不同算法在VisDrone2019和TinyPerson数据集上的性能对比如表10和表11所示.

表10 不同算法在VisDrone2019数据集上的性能对比

方法	mAP50:95	mAP50	mAP75
YOLOv7	0.216	0.298	0.214
YOLOv8	0.222	0.309	0.226
YOLO-POD	0.230	0.336	0.234

表11 不同算法在TinyPerson数据集上的性能对比

方法	mAP50:95	mAP50	mAP75
YOLOv7	0.144	0.240	0.126
YOLOv8	0.146	0.251	0.128
YOLO-POD	0.157	0.245	0.154

在VisDrone2019数据集上,选择mAP50:95、mAP50和mAP75作为本次评估指标,从表10可以看出,YOLO-POD检测的mAP50:95、mAP50、mAP75为23%、33.6%和23.4%,比标准YOLOv7分别提高1.4%、3.8%和2%.比YOLOv8分别提高0.8%、2.7%和0.4%.在TinyPerson数据集上,所选择的评估指标与VisDrone2019数据集上的评估指标相同,从表11可以看出,YOLO-POD检测的mAP50:95、mAP75分别比YOLOv7提高1.3%、2.8%,比YOLOv8提高1.1%、2.6%,mAP50结果高于YOLOv7却低于YOLOv8.所提出的YOLO-POD在指标mAP50上之所以表现不佳,这是由于TinyPerson数据集里的行人比较密集,这是一项困难的小尺寸目标检测任务.对于其他指标却有所提高,表明总体表现有所改善.因此,在上述两个公共数据集上的实验验证了所提出的

YOLO-POD的有效性,提高了小目标的检测精度,同时,YOLO-POD在多个数据集上的通用能力也得到了验证.

5 结论

本文基于YOLOv7提出了一种高精度的PCB微小缺陷检测模型YOLO-POD.首先结合PConv,将ELAN模块设计为更加高效的P-ELAN,有效降低了模型的参数和计算量.其次,基于多维注意力机制的ODConv可以更好地提取微小目标的特征信息.本文使用PConv与ODConv设计了POD-MP和POD-CSP模块,并提出了OD-Neck结构.实验证明了该模块可以提高对微小目标的检测精度,减少背景信息对目标特征的干扰.最后,YOLO-POD在训练时采用了Alpha-SIoU损失函数,可以进一步提高检测精度.实验结果表明,所提出的YOLO-POD模型在PCB表面缺陷检测中的性能优于目前最先进的目标检测模型,有着较低的误检率、较高的精确率和召回率.并且在mAP50:95和mAP75指标上有着明显的精度提升,证明了该方法对PCB缺陷高精度检测的有效性.

综上所述,本文的算法在PCB微小缺陷检测方面取得了令人满意的结果,并为未来的研究和实际应用提供了有价值的参考.同时,我们也在VisDrone2019和TinyPerson数据集上进行了实验,验证了本文改进模型在小目标检测任务上的通用性.然而,必须指出的是,小目标缺陷检测仍然是目标检测发展的主要挑战,可进一步优化和扩展,以应对不同场景和更复杂的缺陷检测任务.因此,未来的研究工作将致力于持续提升小目标检测的性能.

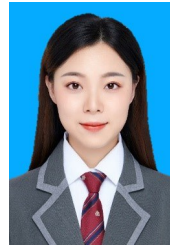
参考文献

- [1] CHEN S Y, CHENG T H, FANG J M, et al. TinyDet: Accurately detecting small objects within 1 GFLOPs[J]. Science China Information Sciences, 2022, 66(1): 119102.
- [2] GE Z, LIU S T, WANG F, et al. YOLOX: Exceeding YOLO series in 2021[EB/OL]. (2021-07-18) [2023-08-01]. <http://arxiv.org/abs/2107.08430>.
- [3] CARINI A, RAMPONI G, MATHEWS V J. Nonlinear signal and image processing—A special issue in honour of Giovanni L. Sicuranza on his seventy-fifth birthday[J]. EURASIP Journal on Advances in Signal Processing, 2017, 2017(1): 28.
- [4] ZHU Y, LING Z G, ZHANG Y Q. Research progress and prospect of machine vision technology[J]. Journal of Graphics, 2020, 41(6): 871-890.
- [5] MALYUSKIN O, FUSCO V. Super-resolution defect char-

- acterization using microwave near-field resonance reflectometry and cross-correlation image processing[J]. *Sensing and Imaging*, 2017, 18(1): 7.
- [6] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single shot MultiBox detector[C]//European Conference on Computer Vision. Cham: Springer, 2016: 21-37.
- [7] 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). Piscataway: IEEE, 2016: 779-788.
- [8] REDMON J, FARHADI A. YOLO9000: Better, faster, stronger[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2017: 6517-6525.
- [9] REDMON J, FARHADI A. YOLOv3: An incremental improvement[EB/OL]. (2018-04-08) [2023-08-01]. <http://arxiv.org/abs/1804.02767>.
- [10] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. YOLOv4: Optimal speed and accuracy of object detection[EB/OL]. (2020-04-23)[2023-08-01]. <http://arxiv.org/abs/2004.10934>.
- [11] LI C Y, LI L L, JIANG H L, et al. YOLOv6: A single-stage object detection framework for industrial applications[EB/OL]. (2022-09-07) [2023-08-01]. <https://arxiv.org/abs/2209.02976>.
- [12] 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 (CVPR). Piscataway: IEEE, 2023: 7464-7475.
- [13] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]//Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition. New York: ACM, 2014: 580-587.
- [14] GIRSHICK R. Fast R-CNN[C]//2015 IEEE International Conference on Computer Vision (ICCV). Piscataway: IEEE, 2015: 1440-1448.
- [15] REN S Q, HE K M, GIRSHICK R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks[C]//NIPS'15: Proceedings of the 28th International Conference on Neural Information Processing Systems - Volume 1. New York: ACM, 2015: 91-99.
- [16] HE K M, GKIOXARI G, DOLLÁR P, et al. Mask R-CNN[C]//2017 IEEE International Conference on Computer Vision (ICCV). Piscataway: IEEE, 2017: 2980-2988.
- [17] HUANG G, LIU Z, VAN DER MAATEN L, et al. Deeply connected convolutional networks[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2017: 2261-2269.
- [18] HUANG W B, WEI P. A PCB dataset for defects detection and classification[EB/OL]. (2019-01-24) [2023-08-01]. <https://arxiv.org/abs/1901.08204>.
- [19] CHE X, ZHU W Z. Tile surface defect detection based on improved faster R-CNN[J]. *Academic Journal of Science and Technology*, 2023, 7(1): 184-190.
- [20] DING R W, DAI L H, LI G P, et al. TDD-net: A tiny defect detection network for printed circuit boards[J]. *CAAI Transactions on Intelligence Technology*, 2019, 4(2): 110-116.
- [21] PENG C, LIU Y K, YUAN X P, et al. Research of image recognition method based on enhanced inception-ResNetV2[J]. *Multimedia Tools and Applications*, 2022, 81(24): 34345-34365.
- [22] CHEN I C, HWANG R C, HUANG H C. PCB defect detection based on deep learning algorithm[J]. *Processes*, 2023, 11(3): 775.
- [23] XIN H J, CHEN Z B, WANG B Y. PCB electronic component defect detection method based on improved YOLOv4 algorithm[J]. *Journal of Physics: Conference Series*, 2021, 1827(1): 012167.
- [24] WANG S, LU H, LU D, et al. PCB board defect detection based on lightweight artificial neural network[J]. *Instrumentation Science & Technology*, 2022, 50(5): 98-104.
- [25] YANG Y J, KANG H Y. An enhanced detection method of PCB defect based on improved YOLOv7[J]. *Electronics*, 2023, 12(9): 2120.
- [26] WU D L, JIANG S, ZHAO E L, et al. Detection of camellia oleifera fruit in complex scenes by using YOLOv7 and data augmentation[J]. *Applied Sciences*, 2022, 12(22): 11318.
- [27] CHEN Z, LIU C, FILARETOV V, et al. Multi-scale ship detection algorithm based on YOLOv7 for complex scene SAR images[J]. *Remote Sensing*, 2023, 15(8): 2071.
- [28] WANG F H, JIANG J, CHEN Y, et al. Rapid detection of Yunnan Xiaomila based on lightweight YOLOv7 algorithm[J]. *Frontiers in Plant Science*, 2023, 14: 1200144.
- [29] LIU K Y, SUN Q, SUN D M, et al. Underwater target detection based on improved YOLOv7[J]. *Journal of Marine Science and Engineering*, 2023, 11(3): 677.
- [30] 王莹, 王晶, 高岚, 等. 一种注意力机制优化方法及硬件加速设计[J]. *电子学报*, 2023, 51(4): 1021-1029.
- WANG Y, WANG J, GAO L, et al. An improved attention mechanism algorithm model and hardware acceleration

- tion design method[J]. Acta Electronica Sinica, 2023, 51(4): 1021-1029. (in Chinese)
- [31] WOO S, PARK J, LEE J Y, et al. CBAM: Convolutional block attention module[C]//European Conference on Computer Vision. Cham: Springer, 2018: 3-19.
- [32] JADERBERG M, SIMONYAN K, ZISSERMAN A, et al. Spatial transformer networks[EB/OL]. (2015-06-05) [2023-08-01]. <http://arxiv.org/abs/1506.02025>.
- [33] HU J, SHEN L, SUN G. Squeeze-and-excitation networks[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2018: 7132-7141.
- [34] CHEN Y P, DAI X Y, LIU M C, et al. Dynamic convolution: Attention over convolution kernels[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2020: 11027-11036.
- [35] LI C, ZHOU A, YAO A. Omni-dimensional dynamic convolution[EB/OL]. (2022-09-16) [2023-08-01]. <https://arxiv.org/abs/2209.07947>.
- [36] CHEN J R, KAO S H, HE H, et al. Run, don't walk: Chasing higher FLOPS for faster neural networks[C]//2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2023: 12021-12031.
- [37] ZHANG Q L, JIANG Z Q, LU Q S, et al. Split to be slim: An overlooked redundancy in vanilla convolution[EB/OL]. (2020-06-22)[2023-08-01]. <http://arxiv.org/abs/2006.12085>.
- [38] HAN K, WANG Y H, TIAN Q, et al. GhostNet: More features from cheap operations[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2020: 1577-1586.
- [39] 王硕, 王坚, 王亚男, 等. 一种基于特征融合的恶意代码快速检测方法[J]. 电子学报, 2023, 51(1): 57-66.
- WANG S, WANG J, WANG Y N, et al. A fast malicious code detection method based on feature fusion[J]. Acta Electronica Sinica, 2023, 51(1): 57-66. (in Chinese)
- [40] GEVORGYAN Z. Siou loss: More powerful learning for bounding box regression[EB/OL]. (2022-05-25)[2023-08-01]. <https://arxiv.org/abs/2205.12740>.
- [41] HE J B, ERFANI S, MA X J, et al. Alpha-IoU: A family of power intersection over union losses for bounding box regression[EB/OL]. (2021-10-26)[2023-08-01]. <http://arxiv.org/abs/2110.13675>.
- [42] GUO D J, QIU B, LIU Y P, et al. Supernova detection based on multi-scale fusion faster RCNN[C]//2021 6th International Conference on Intelligent Computing and Signal Processing (ICSP). Piscataway: IEEE, 2021: 985-988.
- [43] YU J H, JIANG Y N, WANG Z Y, et al. UnitBox: An advanced object detection network[C]//Proceedings of the 24th ACM International Conference on Multimedia. New York: ACM, 2016: 516-520.
- [44] XU C, WANG J W, YANG W, et al. RFLA: Gaussian receptive field based label assignment for tiny object detection[C]//Computer Vision - ECCV 2022: 17th European Conference, Part IX. New York: ACM, 2022: 526-543.
- [45] NIE G T, HUANG H. Multi-oriented object detection in aerial images with double horizontal rectangles[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2023, 45(4): 4932-4944.
- [46] HAN J M, DING J, LI J, et al. Align deep features for oriented object detection[J]. IEEE Transactions on Geoscience and Remote Sensing, 2022, 60: 5602511.
- [47] TANG J L, LIU S B, ZHAO D X, et al. PCB-YOLO: An improved detection algorithm of PCB surface defects based on YOLOv5[J]. Sustainability, 2023, 15(7): 5963.
- [48] CHEN B Y, DANG Z C. Fast PCB defect detection method based on FasterNet backbone network and CBAM attention mechanism integrated with feature fusion module in improved YOLOv7[J]. IEEE Access, 2023, 11: 95092-95103.

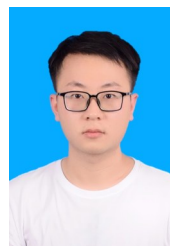
作者简介



郭艳 女, 1997年7月出生于山西省忻州市. 现为广西科技大学硕士研究生. 研究方向为图像处理和目标检测.
E-mail: 1655807007@qq.com



王智文 男, 1969年1月出生于湖南省邵阳市. 现为广西科技大学教授. 主要研究领域为机器学习与计算机视觉、移动目标检测与识别. 主持与参与国家自然科学基金项目及广西自然科学基金项目等25项, 在国内外公开发表研究论文173篇, 被SCI或EI收录论文63篇. 出版专著3部.
E-mail: wzw69@gxust.edu.cn



赵润星 男, 1998年10月出生于河南省洛阳市. 现为广西科技大学硕士研究生. 主要研究领域为图像处理和目标跟踪.
E-mail: zrx_1998@163.com