

DOI:10.19651/j.cnki.emt.2518158

基于 ARS-YOLOv9s 的 PCB 缺陷检测算法 *

陈枫赟^{1,3} 李鹏^{1,2,3} 张翔凯^{1,3} 于涛^{1,3} 余珺泽^{1,3}

(1. 南京信息工程大学江苏省气象探测与信息处理重点实验室 南京 210044; 2. 江苏省工业环境危害要素监测与评估工程研究中心 无锡 214105; 3. 南京信息工程大学江苏省气象传感网技术工程中心 南京 210044)

摘要: 针对现有印刷电路板缺陷小、种类多、特征不明显的问题,提出了一种基于 ARS-YOLOv9s 的缺陷检测算法,该算法在 YOLOv9s 网络架构基础上进行优化。针对原算法在多尺度特征融合时存在信息丢失的问题,融入 AFPN 对图像进行特征融合从而丰富语义信息;通过在主干网络中引入 iRMB 注意力机制,提升浅层特征中微小缺陷的关注度;针对目标缺陷特征较小的问题,删除大目标检测层并新增微小目标检测层,模型轻量化的同时并提高其检测精度;将原模型损失函数替换为 Shape-IoU 以期改善正负样本不均衡对模型的影响,加速模型收敛。实验结果表明,本文算法 mAP 为 98%,mAP@0.5:0.95 为 68.2%,相较于原 YOLOv9s 分别提升了 2.8%、9.3%,且各个类别缺陷 mAP 均有明显提升,证明了本文算法的有效性。

关键词: YOLOv9s; 小目标检测; 印刷电路板; 特征金字塔网络; 注意力机制**中图分类号:** TP391; TN912 **文献标识码:** A **国家标准学科分类代码:** 520.20

PCB defect detection algorithm based on ARS-YOLOv9s

Chen Fengyun^{1,3} Li Peng^{1,2,3} Zhang Xiangkai^{1,3} Yu Tao^{1,3} Yu Junze^{1,3}

(1. Jiangsu Key Laboratory of Meteorological Observation and Information Processing, Nanjing University of Information Science and Technology, Nanjing 210044, China; 2. Jiangsu Engineering Research Center for Monitoring and Assessment of Industrial Environmental Hazard Elements, Wuxi 214105, China; 3. Jiangsu Meteorological Sensor Network Technology Engineering Center, Nanjing University of Information Science and Technology, Nanjing 210044, China)

Abstract: Aiming at the problems of small, diverse and inconspicuous features of existing printed circuit board defects, a defect detection algorithm based on ARS-YOLOv9s is proposed, which is optimised on the basis of YOLOv9s network architecture. To address the problem of information loss in multi-scale feature fusion in the original algorithm, AFPN is integrated into the image feature fusion so as to enrich the semantic information; by introducing the iRMB attention mechanism in the backbone network, the attention to the tiny defects in the shallow features is improved; to address the problem of the small target defects, the large target detection layer is deleted and a new tiny target detection layer is added, which lightens the model and improves the detection accuracy; the original model loss function is replaced by the loss function, and the original model loss function is replaced by the loss function, and the original model loss function is replaced by the loss function. The original model loss function is replaced by Shape-IoU to improve the impact of positive and negative sample imbalance on the model and accelerate model convergence. The experimental results show that the mAP of this paper's algorithm is 98%, and that of mAP@0.5:0.95 is 68.2%, which is 2.8% and 9.3% higher than that of the original YOLOv9s, respectively, and the mAP of defects of each category is significantly improved, which proves the effectiveness of this paper's algorithm.

Keywords: YOLOv9s; small target detection; printed circuit board; characteristic pyramid network; attention mechanism

0 引言

印刷电路板(printed circuit board, PCB)在电子工业中扮演着核心角色,对电子设备的性能和耐用性具有决定

性影响。PCB 的设计和生产过程变得更为精细和复杂,任何环节的疏漏都可能引发短路、断路和毛刺等质量问题^[1],这些缺陷的检测极具挑战性,人工检测效率低、主观性强,容易造成误差。因此,采用高效的缺陷检测技术对于提高

收稿日期:2025-02-24

* 基金项目:国家自然科学基金(41075115)、江苏省重点研发计划社会发展项目(BE2015692)、无锡市社会发展科技示范工程项目(N20191008)资助

产品质量、降低生产成本非常关键。

在现代工业制造领域,传统的 PCB 缺陷检测手段主要依赖人工目视检查、电气性能测试^[2]、超声扫描和热成像检测等。人工目视检查受限于操作者的技能和视觉能力,其成本高昂以及作业效率不足。电气测试需与元件进行直接电气连接,存在损坏组件的风险,并且成本不菲,无法满足 PCB 高精度检测的需求。超声扫描和热成像检测成本高、速度慢,对表面缺陷不敏感,数据解读起来也需要专业性,不适用于小规模的工程生产线。

基于深度学习的 PCB 缺陷检测技术主要有:单阶段算法和双阶段算法。单阶段算法通过单次检测直接得出结果,而双阶段算法将检测问题分为两个步骤处理。在实践中,YOLO 系列算法作为单阶段检测的代表,因其快速和高效而广泛应用。余翔等^[3]在构建颈部网络时,引入了一个融合多尺度特征的模块,该模块采用了空洞空间和混合注意力机制,在进行目标回归时采用了归一化的 Wasserstein 距离来优化传统的损失函数,增强了模型对小目标的检测能力,但优化过程对超参数敏感。周贤勇等^[4]在 YOLOv7 模型中通过引入 FReLU 视觉激活函数替代 SiLU,集成了 SE 注意力机制,并通过改进的特征融合网络对不同层级的特征进行有效整合,但在处理长文本或复杂图像时可能会丢失一些重要的信息。胡江宇等^[5]在 Faster R-CNN 的基础上,进一步在优化后的 ROI-Pooling 层提取 ROI 特征,添加了过滤特征的金字塔网络,提升了模型在小目标检测精度方面的性能,但其检测速度大大降低。吴兴辉等^[6]在 YOLOv7 算法中采用 FReLU 激活函数替代原有的激活函数,将 CBAM 注意力机制融入网络骨架,但模型的推理速度大幅下降。张莹等^[7]提出的 YOLOPCB 网络通过精简 YOLOv7 主干网络结构,设计跨通道信息连接模块与浅层特征融合模块,采用自适应加权跳层连接策略等简化了网络结构,但因此产生信息丢失问题导致检测精度下降。王军等^[8]通过在 YOLOv8 的 C2f 模块融入 SE 注意力机制增强通道特征权重,在 SPPF 中引入 Basic RFB 扩大感受野,但在轻量化部署场景下资源消耗过高。

经对比,现有研究中 Faster R-CNN 依赖于区域提议网络来识别候选区域,在小目标上不够敏感,而 YOLO 在检测 PCB 缺陷的诸多研究中不能兼顾检测精度和模型体积的问题。因此,本文将 YOLOv9s 作为基准算法,融合了渐进式金字塔网络(asymptotic feature pyramid network, AFPN)^[9],加入 iRMB(inverted residual mobile block)^[10]注意力机制,增加微小目标检测层并去除大目标检测层,将损失函数替换为 Shape-IoU^[11],提出了一种基于 YOLOv9s 的 PCB 缺陷检测算法 ARS-YOLOv9s。

1 YOLOv9s 网络框架概述与改进

1.1 YOLOv9s 网络框架

YOLOv9^[12]是一种新一代的目标检测算法,它延续了

YOLO(you only look once)^[13]系列的核心理念,以快速且高效的性能在实时检测任务中表现出色。在同团队打造的 YOLOv7 的基础上,YOLOv9 进行了显著的优化,不仅提高了检测的准确性,还加快了处理速度,并且降低了对计算资源的需求。YOLOv9 引入了可编程梯度信息(programmable gradient information, PGI)和广义高效层聚合网络(generalized efficient layer aggregation network, GELAN)两个关键性创新点。PGI 通过辅助可逆分支对梯度信息进行定制化处理,对特定层的梯度进行放大、缩小或抑制,以避免梯度消失或爆炸,如图 1 所示;GELAN 通过更好的特征融合和优化策略,在增强的局部注意力网络(ELAN)的基础上平衡了算法复杂度与检测性能,其结构如图 2 所示。两者的结合优化了深度神经网络训练过程中信息丢失问题,提升了对小目标和在复杂场景下的检测能力,改进的正则化策略帮助模型在训练过程中避免过拟合,提升了模型的鲁棒性。

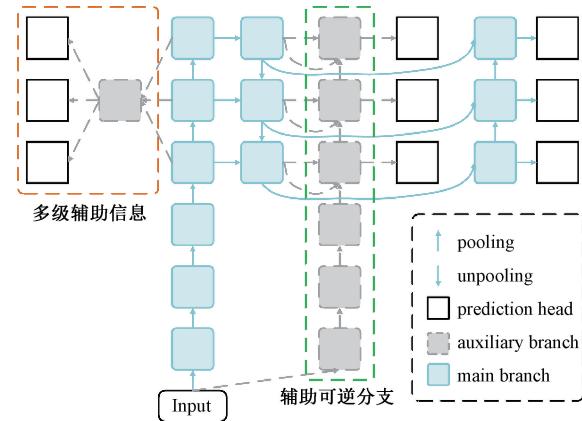


图 1 PGI 结构

Fig. 1 PGI structure

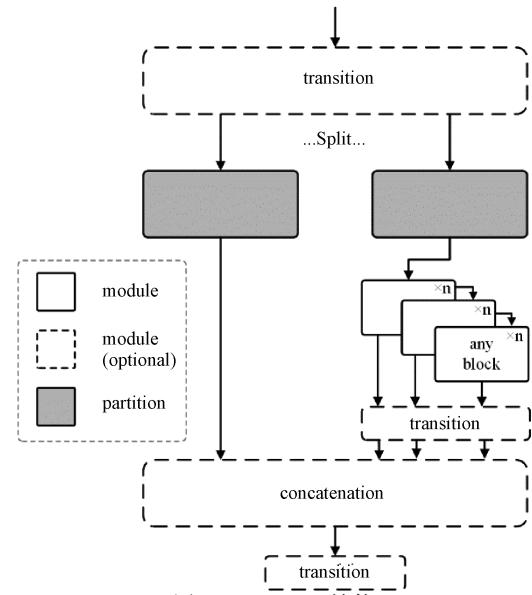


图 2 GELAN 结构

Fig. 2 GELAN structure

YOLOv9s 在维持高效性能的基础上,进一步减少了模型的参数数量和卷积层,同时尽可能保留了 YOLOv9 的特征提取和多尺度特征融合能力,使其更适合在资源受限的环境中使用。网络结构如图 3 所示,输入图像的大小为 $640 \times 640 \times 3$,主干网络中 Conv 和 AConv 是两种不同的卷积层;ELAN1 是一个受 DenseNet^[14]与 ResNet^[15]启发而设计的用于整合模型多层次特征信息的聚合网络;

RepNCSPELAN4 是 YOLOv9 中的特征提取——融合模块,由 Conv 与 RepNCSP 组成。预测网络中 Concat 将不同尺度的特征图在通道维度上进行拼接;Upsample 采用最近邻插值方式进行上采样;SPPELAN 结合了空间金字塔池化(SPP)和 ELAN 模块;辅助分支设计旨在保持网络中梯度的迭代和参数更新,确保信息的完整性,减少了直接集成到主分支所带来的高计算成本。

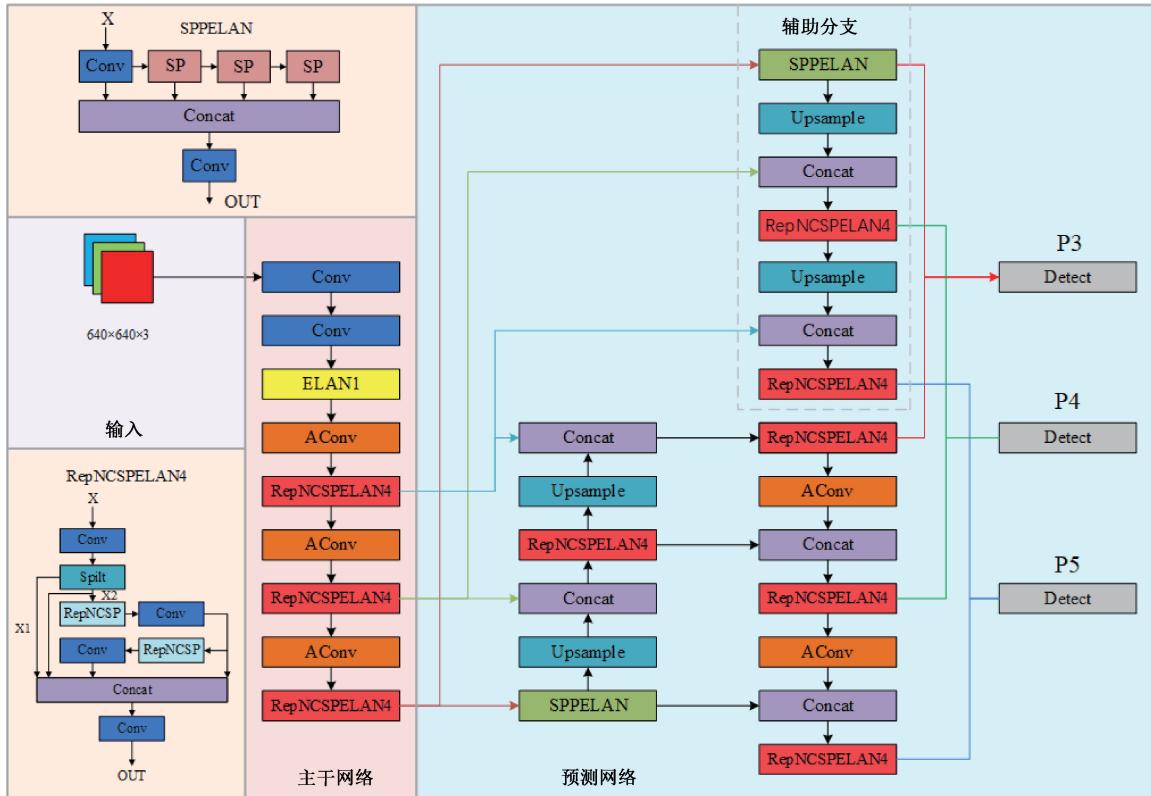


图 3 YOLOv9s 网络结构

Fig. 3 YOLOv9s network architecture

1.2 ARS-YOLOv9s 整体结构

为综合优化 YOLOv9s 模型,提出了一种 ARS-YOLOv9s 的算法,其网络结构如图 4 所示。首先从主干网络选取特定的特征层作为 AFPN 模块的输入,通过逐步融合低层和高层特征以及引入自适应空间融合操作,减少信息丢失,避免不同层次特征融合时的冲突。其次,在主干网络中引入 iRMB 注意力机制,通过其倒置残差块和元移动块的操作克服现有模型的局限性,从而实现信息高效处理。进而,针对 PCB 缺陷特征,去除大目标的检测尺度,保留中小目标检测尺度,并加入更小目标检测尺度。最后,将原网络中的损失函数替换为 Shape-IoU,通过关注边界框本身的状态和比例来计算损失,从而使边界框回归更加准确。

1.3 融合 AFPN 渐进特征金字塔网络

在 YOLOv9s 架构中引入渐进式特征金字塔网络(AFPN)显著增强了对多尺度目标的融合能力。AFPN 通

过自适应特征融合机制,使网络能够针对尺寸较小的目标实现更精准的检测,从而提升模型的检测精度和鲁棒性。此外,AFPN 的自适应学习策略允许根据输入图像的特定内容及检测任务需求,动态调整特征层之间的融合方式,进而增强整体的特征表达能力。

根据图 5 所示,AFPN 从主干网络的 iRMB 模块逐层提取特征,采用渐进式融合策略减小非相邻层间的语义差异。该策略从主干网络中提取的 3 个不同深度特征层 L1、L2 和 L3,由于 L1 与 L3 之间非邻近且存在显著的语义差异,因此 AFPN 通过自上而下的传递路径,利用 L2 层作为中介,将 L2 的语义信息向 L1 层传递。继而,渐进式策略系统地融合 L1 和 L2 的特征,确保细节信息的最大化保留,同时通过自下而上的路径强化特征融合效果。在此基础上,L3 层的特征也按照同样的方法与 L1 和 L2 层进行融合,以实现跨层特征的充分融合。

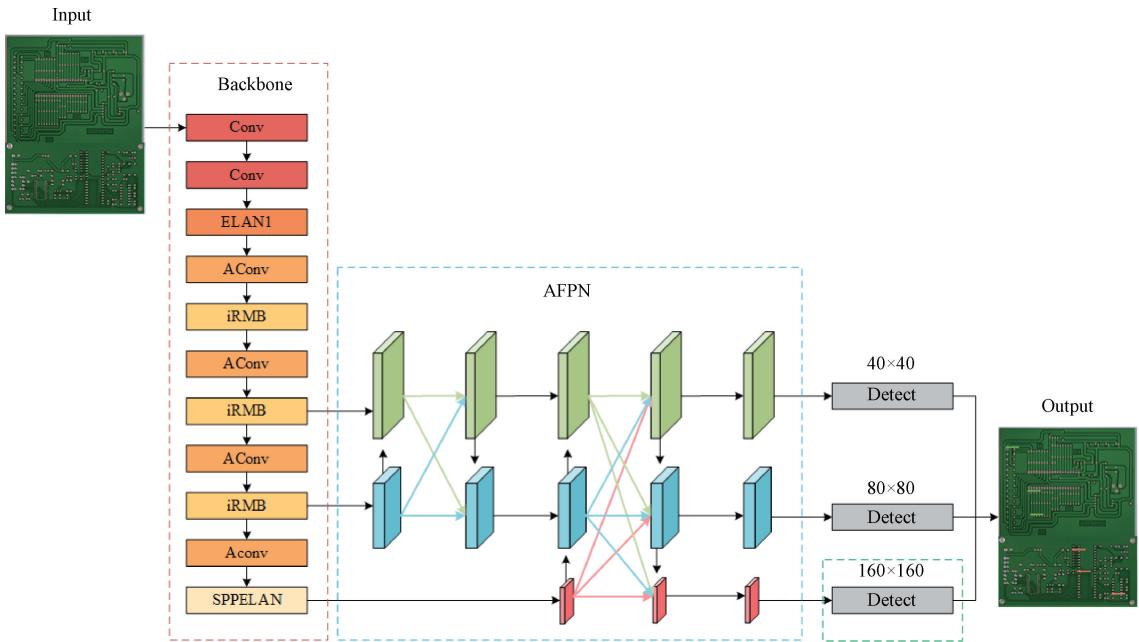


图 4 ARS-YOLOv9s 网络结构

Fig. 4 ARS-YOLOv9s network structure

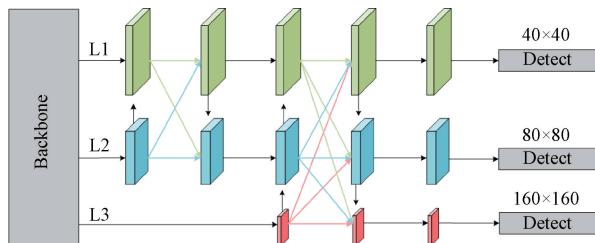


图 5 AFPN 结构

Fig. 5 AFPN structure

图 6 中,黑色箭头代表卷积操作,彩色箭头代表自适应空间融合过程。通过应用 ASFF (adaptively spatial feature fusion)^[16] 算法,在多层次特征融合过程中对各级特征赋予不同的空间权重,以突出重要层级的作用,并减少其他干扰。图 6 表示自适应空间融合运算,其计算公式如式(1)所示。

$$\mathbf{y}_{ij}^l = \alpha_{ij}^l \cdot \mathbf{x}_{ij}^{1 \rightarrow l} + \beta_{ij}^l \cdot \mathbf{x}_{ij}^{2 \rightarrow l} + \gamma_{ij}^l \cdot \mathbf{x}_{ij}^{3 \rightarrow l} \quad (1)$$

其中, \mathbf{y}_{ij}^l 表示通过多层次特征自适应空间融合得到的特征向量, $\mathbf{x}_{ij}^{n \rightarrow l}$ 表示从 n 层到 l 层位置 (i, j) 的特征向量, $\alpha_{ij}^l, \beta_{ij}^l, \gamma_{ij}^l$ 代表三层特征在 l 层的空间权重,三者关系需满足 $\alpha_{ij}^l + \beta_{ij}^l + \gamma_{ij}^l = 1$ 。

此外,AFPN 策略为了减少直接融合 L1 与 L3 时带来的信息损失,因此优先考虑融合语义信息相近的 L1 和 L2 层。鉴于 L2 与 L3 层的相对较小的语义差异,该策略进一步缩小了 L1 与 L3 间的语义差距,从而提高了特征融合的精确性与效率。这种层间的语义一致性优化,不仅增强了特征表达的连贯性,而且提高了整体网络在 PCB 缺陷检测任务中的准确性。

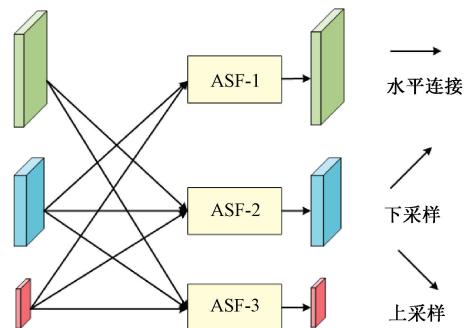


图 6 自适应空间融合

Fig. 6 Adaptive spatial fusion

1.4 iRMB 注意力机制

PCB 缺陷检测属于小目标检测任务,原模型 YOLOv9s 通过不同尺度的检测头检测不同大小的目标,针对小目标检测任务而言,赋予不同尺度的检测头相同的权重会导致计算资源浪费,因此,在主干网络中引入一种 iRMB 注意力机制来实现模型对小目标权重的增加,以提高检测精度。

iRMB 借鉴了 MobileNetv2^[17] 中的反向残差结构 (IRB),并融合了 Transformer^[18] 架构中的核心机制——多头自注意力 (MHSA) 和前馈神经网络 (FFN),提炼出了一种通用的 Meta Mobile Block (MMB),该模块通过参数化扩展比 λ 和高效的操作符 F 来具体化不同的变体。MMB 能够揭示 IRB、MHSA 和 FFN 之间的内在联系,并且可以被看作是轻量级 Transformer 的集中体。如图 7 所示通过将 MobileNetv2 的 IRB 结构与 Transformer 的 MHSA 和 FFN 进行概念化,用一个 MMB 来归纳并表示

这些不同的架构,即通过扩展比 λ 和高效算子 F 来实现多

样化的模块实例,实现了高效的特征提取和信息流动。

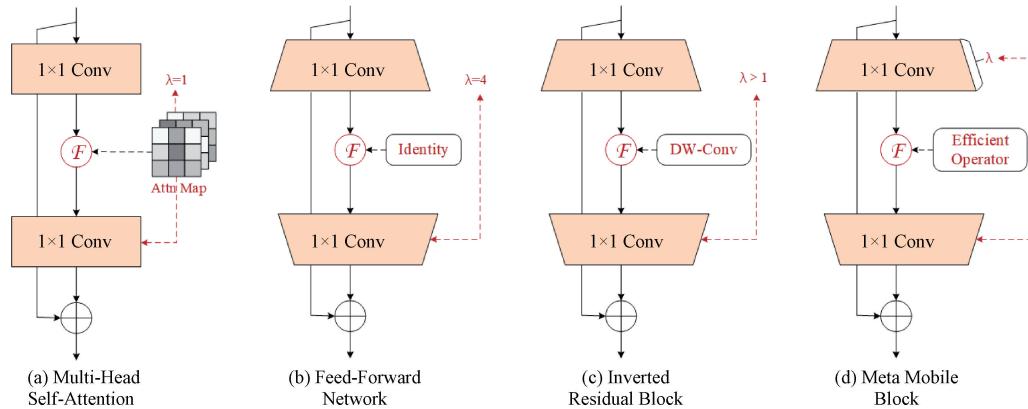


图 7 MMB 模块示意图

Fig. 7 MMB module schematic diagram

以图像输入 X 为例,MMB 首先使用输出/输入比为 λ 的扩展 MLP_e ,扩展通道维数,如式(2)所示。

$$X_e = MLP_e(X) (\in \mathbb{R}^{\lambda C \times H \times W}) \quad (2)$$

中间操作符 F 进一步增强图像特征,例如恒等操作符、静态卷积、动态 MHSA 等。考虑到 MMB 适用于高效的网络设计,将 F 作为高效操作符的概念,如式(3)所示。

$$X_f = F(X_e) (\in \mathbb{R}^{\lambda C \times H \times W}) \quad (3)$$

最后,一个反向输入/输出比等于 λ 的收缩 MLP_s ,来收缩通道尺寸,如式(4)所示。

$$X_s = MLP_s(X_f) (\in \mathbb{R}^{C \times H \times W}) \quad (4)$$

其中,使用残差连接来获得最终的输出,如式(5)所示。

$$Y = X + X_s (\in \mathbb{R}^{C \times H \times W}) \quad (5)$$

如图 8 所示,基于 MMB,设计了一个反向残差移动块(iRMB),吸收了 CNN 架构的效率来建模局部特征和 Transformer 架构动态建模的能力来学习长距离交互。iRMB 中的 F 被建模为级联的 EW-MHSA 和 DW-Conv 卷积,如式(6)所示。

$$F(\cdot) = (DW-Conv, Skip)(EW - MHSA(\cdot)) \quad (6)$$

该注意力机制通过特定的设计减少了参数量和计算量,简化了模型结构的同时还保持了模型的高效性。

1.5 目标检测层优化

YOLOv9s 原网络 Head 部分包含 3 个输出层 P3、P4、P5,步长分别为 8、16、32,当输入图像大小为 640×640 时,3 个输出层检测尺度分别对应 80×80 、 40×40 、 20×20 。而 PCB 缺陷检测的目标都相对较小,P5 层 20×20 的尺度感受野较大,模型在训练过程中不断卷积会丢失小目标的像素信息,该层深度较深占用了大量计算资源,故去除 P5 层保留 P3、P4 层并增加一个 160×160 的 P2 检测层,图像经主干网络特征提取后生成浅层且高分辨率的特征图 C2,提高了对微小目标检测的敏感度。改进后的三尺度检测结

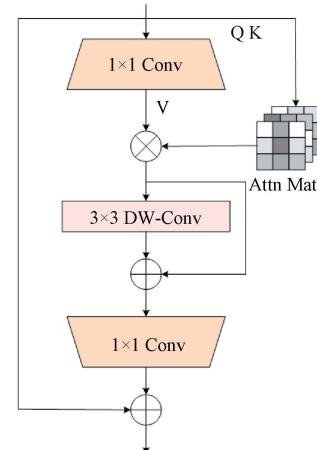


图 8 iRMB 结构

Fig. 8 iRMB structure

构如图 9 所示。

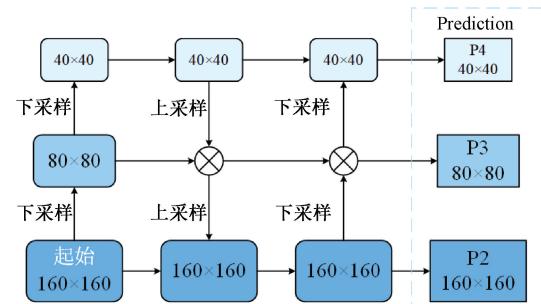


图 9 检测层优化示意图

Fig. 9 Schematic diagram of detection layer optimization

1.6 边界框回归损失函数改进

损失函数用于评估模型预测结果与实际结果之间的偏差,其值越低表示预测结果的准确性越高,即预测值与真实值越接近。YOLOv9s 采用的损失函数由分类损失和边界框回归损失构成,其中边界框回归损失融合了控制点

损失(distribution focal loss, DFL)^[19]与连续 IoU 损失(CIoU Loss)^[20]。DFL 通过将边界框参数化为距离先验框 4 个边的距离来改进回归精度,CIoU Loss 在传统 IoU 的基础上衡量两个矩形框的重叠面积,综合考虑中心点距离和宽高比的一致性,通过引入中心点距离惩罚项和形状损失项,有效地提高了模型训练的稳定性和收敛速度,进一步提升了目标检测的精度。其公式如式(7)~(9)所示。

$$L_{\text{CIoU}} = 1 - \text{IoU} + \frac{\rho^2}{d^2} + \alpha\nu \quad (7)$$

$$\nu = \frac{4}{\pi^2} \left(\arctan \frac{w^{gt}}{h^{gt}} - \arctan \frac{w}{h} \right)^2 \quad (8)$$

$$\alpha = \frac{\nu}{(1 - \text{IoU}) + \nu} \quad (9)$$

式中: IoU 是预测框和真实框的交集与并集的比值, ρ 是预测框和真实框中心点之间的欧氏距离, d 是框的对角线距离,用于标准化中心点距离, ν 是宽高比惩罚项,它衡量预测框和真实框宽高比的不一致性, α 是一个用于平衡宽高比惩罚项的权重,通常是一个根据 IoU 动态调整的系数。

然而,CIoU 存在宽高比定义不明确、对质量差的回归样本过于敏感导致正负样本不均衡等局限性,过于注重边界框的精准性并且忽视了边框形状、尺度等固有属性对回归的影响,从而限制了模型的优化和泛化能力以及可能引起算法过拟合。因此,为解决上述问题引入了一种关注边界框自身形状和尺寸的边界框回归方法,即 Shape-IoU 损失函数。该损失函数通过专注于边界框的自身形状、尺度和比例来计算损失,有效地利用相似度度量和梯度信息,优化模型训练过程,从而提高目标检测的精度和定位的准确度,增强了其在边界框回归中的准确性和鲁棒性。

Shape-IoU 损失函数如图 10 所示,粉框表示真实框,蓝框表示预测框, b 和 b^{gt} 分别表示预测框和真实框的中心点, w^{gt}, h^{gt} 表示真实框的宽度和高度, w, h 分别预测框的宽度和高度, x_c^{gt} 和 y_c^{gt} 是真实框中心点的坐标, x_c 和 y_c 是预测框中心点的坐标。

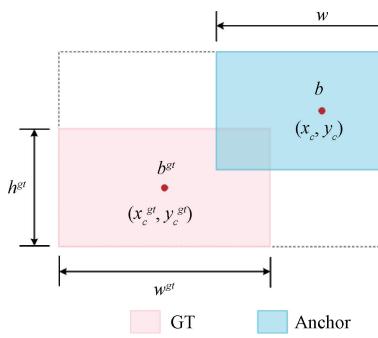


图 10 Shape-IoU 示意图
Fig. 10 Shape-IoU schematic

其计算公式如式(10)~(15)所示。

$$\text{IoU} = \frac{|B \cap B^{gt}|}{|B \cup B^{gt}|} \quad (10)$$

$$ww = \frac{2 \times (w^{gt})^{scale}}{(w^{gt})^{scale} + (h^{gt})^{scale}} \quad (11)$$

$$hh = \frac{2 \times (h^{gt})^{scale}}{(w^{gt})^{scale} + (h^{gt})^{scale}} \quad (12)$$

$$\text{distance}^{shape} = hh \times \frac{(x_c - x_c^{gt})^2}{c^2} + \\ ww \times \frac{(y_c - y_c^{gt})^2}{c^2} \quad (13)$$

$$\Omega^{shape} = \sum_{t=w,h} (1 - e^{-\omega_t})^\theta, \theta = 4 \quad (14)$$

$$\begin{cases} \omega_w = hh \times \frac{|w - w^{gt}|}{\max(w, w^{gt})} \\ \omega_h = ww \times \frac{|h - h^{gt}|}{\max(h, h^{gt})} \end{cases} \quad (15)$$

该损失函数由 3 个成本函数(cost function)组成:损失函数 IoU 、距离损失函数 distance^{shape} 和形状损失函数 Ω^{shape} 。其中 ww 与 hh 分别为水平方向与竖直方向的权重系数,其取值与 GT 框的形状有关; $scale$ 为尺度因子,与数据集中目标的大小有关,在小目标数据集中,通常取 [0, 1.5]^[21];参数 θ 用于调节模型对形状损失的重视程度,以防止过分关注形状损失而忽视预测框的定位精度,通过遗传算法得知该参数的最优值为 4; ω_t 为宽/高损失; c 为包含 GT 框和预测框的最小框的对角线距离;最终边界框回归损失如式(16)所示。

$$L_{\text{Shape-IoU}} = 1 - \text{IoU} + \text{distance}^{shape} + 0.5 \times \Omega^{shape} \quad (16)$$

2 实验结果与分析

2.1 实验环境搭建与参数设置

1) 实验环境搭建

本文实验环境如表 1 所示。

表 1 实验环境配置

Table 1 Experimental environment configuration

名称	参数
CPU	Intel-i5-13600KF
GPU	NVIDIA GeForce RTX 4060Ti 16 G
RAM	32 G
操作系统	Windows 10
编程语言	Python 3.9.19
学习框架	Pytorch 1.12
GPU 加速库	CUDA 11.6

2) 参数设置

为了确保实验变量的一致性,统一采 640×640 的输入尺寸,采用随机梯度下降(stochastic gradient descent, SGD)^[22] 优化器进行参数的优化,模型训练迭代次数

epochs 为 300 次,优化器权重衰减为 0.0005,学习动量为 0.937,batch size 大小设置为 16,学习率大小采用动态调整策略,前 100 批次设置 0.01,后 100 批次设置 0.001,参数优化采用随机梯度下降算法。

2.2 数据集准备

本文实验所使用的数据集为北京大学智能机器人开放实验室公开的 PCB 缺陷数据集^[23],缺陷类型如表 2 所示分为 6 类:漏孔(missing_hole)、鼠咬(mouse_bite)、开路(open_circuit)、短路(short)、毛刺(spur)以及余铜(spurious_copper)。

表 2 PCB 缺陷检测类型

Table 2 Types of PCB defect detection

缺陷类型	实物图	检测重点
缺孔		焊盘内有孔径不同的洞
鼠咬		线路缺口不规则
开路		线路断开
短路		邻近线路导通
毛刺		线路有尖刺凸起
余铜		两线路间有多余铜线

由于该数据集只有 693 张缺陷图像,样本数量较少,容易产生过拟合,故本文通过随机翻转、灰度化、添加噪声、高斯模糊、锐化等数据增强方式将数据集扩增到 7 623 张图片,数量如图 11 所示。将处理后的数据集按 8:1:1 的比例随机划分训练集、验证集与测试集。

2.3 评估指标

实验采用计算量、参数量、平均精度均值和精确率与召回率的调和平均数 F1 分值作为评价指标,所有指标都

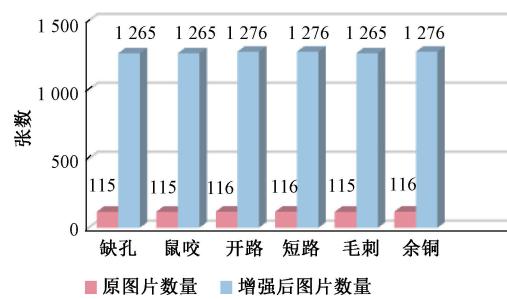


图 11 PCB 数据集数量

Fig. 11 Number of PCB datasets

是数值越大,检测性能越好。

精确率 P (Precision) 表示在检测出的全部目标中识别正确的比例,计算公式如式(17)所示。

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

召回率 R (Recall) 表示正确预测为正的占全部实际为正的比例,计算公式如式(18)所示。

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

其中, TP (true positive) 为将实际为正的样本正确预测为正的数量; FP (false positive) 为错误地将实际为负的样本分类为正的数量; FN (false negative) 为未能检测到实际为正的样本。

平均精度值(average precision, AP) 是 P-R 曲线的积分,即曲线与坐标系所围面积,其公式为:

$$AP = \int_0^1 P(R) dR \quad (19)$$

平均精度均值(mean average precision, mAP) 是 n 个不同类别的 AP 和的平均值,反映了网络整体缺陷检测的准确性,公式如下:

$$mAP = \frac{1}{n} \sum_{i=0}^n AP_i \quad (20)$$

$mAP@0.5$ 是 IoU 阈值为 0.5 时的 mAP, $mAP@0.5:0.95$ 为不同 IoU 阈值(从 0.5~0.95,步长 0.05)上的平均 mAP。

F1 分值表示精准率 P 与召回率 R 的谐波平均值,如式(21)所示。

$$F1 = 2 \frac{PR}{P + R} \quad (21)$$

2.4 实验结果对比分析

1) 优化策略的消融实验

为了验证各模块对 YOLOv9s 改进的有效性,本文在原算法的基础上设计消融实验:将 AFPN、iRMB、优化后的检测层、Shape-Iou 逐步叠加到 YOLOv9s 中,原算法作为对照组,来验证不同模块对网络性能的提升。消融结果如表 3 所示。

由表 3 可知,在 YOLOv9s 的基础上引入渐进式特征

表 3 消融实验结果对比
Table 3 Comparison of ablation experiment results

YOLOv9s	AFPN	iRMB	检测层优化	Shape-Iou	mAP@0.5/%	mAP@0.5:0.95/%	F1 值/%	Params/M
√					95.2	58.9	94.9	9.7
√	√				96.4	64.3	95.8	9.6
√	√	√			97.5	66.1	97.2	9.4
√	√	√	√		97.8	67.8	97.5	7.2
√	√	√	√	√	98.0	68.2	97.8	7.3

金字塔网络,通过自适应空间融合操作来处理不同层次特征融合时的冲突,mAP@0.5提升了1.2%,mAP@0.5:0.95提升了5.4%,F1分值提高了0.9%,参数量与原算法相比几乎不变。在此基础上将主干网络中的RepNCSPELAN4模块替换为iRMB注意力机制模块,赋予小目标更高的权重来解决PCB缺陷小的问题,mAP@0.5比YOLOv9s算法提高了2.3%,mAP@0.5:0.95提升了7.2%,F1分值提高了2.3%,参数量比原算法略微减少。接着,在融合了AFPN和iRMB模块后进行检测层的优化,mAP@0.5比YOLOv9s算法提高了2.6%,mAP@0.5:0.95提升了8.9%,F1分值提高了2.6%,参数量显著减少了2.5M。最后,把改进后的算法中的损失函数替换为Shape-Iou,mAP@0.5比YOLOv9s算法提高了2.8%,mAP@0.5:0.95提升了9.3%,F1分值提高了2.9%,参数量比改进后的算法没替换损失函数前增加了0.1M。综上所述,本文改进后的算法相比YOLOv9s平均精度均值提升了2.8%,参数量因检测层的优化而减少了24.7%,在模型参数量不增多的同时实现了精确度的提升。

2) 对比实验

为进一步验证本文所提算法的优越性,在相同的实验环境和数据集下,将ARS-YOLOv9s与文献[5]、SSD、YOLOv5s、文献[4]、YOLOv8s、YOLOX-s 6种算法进行对比实验,结果如表4所示。

表 4 不同算法实验对比

Table 4 Comparison of different algorithm experiments

算法	mAP@0.5/%	FPS/fps	Params/M	GFLOPs
文献[5]	78.2	20	107.1	—
SSD	67.8	42	99.5	156
YOLOv5s	87.3	156	8.9	16.5
文献[4]	95.9	—	36.9	—
YOLOv8s	96.1	118	11.4	28.6
YOLOX-s	94.5	132	9.3	26.8
ARS-YOLOv9s	98.0	164	7.2	23.9

由表4可知,本文改进后的ARS-YOLOv9s算法

mAP@0.5/%为98%,相较于其他算法中检测精度最好的YOLOv8s提高了1.9%;在检测速度方面,该算法的FPS值为164,相较于检测速度最快的YOLOv5s提高了5.1%;ARS-YOLOv9s的参数量为7.2M,与YOLOv5s相比减少了1.7M;计算量虽然比最小的YOLOv5s大,但与YOLOv8s、YOLOX-s这些精度差异小的算法相比有所减少。综上所述,本文提出的算法在各检测指标方面都有不错的提升,综合地优化了检测能力。验证了该算法在PCB缺陷检测任务中的卓越性能,能够高效且准确地识别缺陷。

如图12所示为本文算法与改进前YOLOv9s的P-R曲线对比图,优化后的算法在P-R曲线下面积上超越了原始的YOLOv9s,代表更高的召回率和检测精度,这表明算法在PCB缺陷检测方面的性能得到了增强,同时准确度也有所提升。

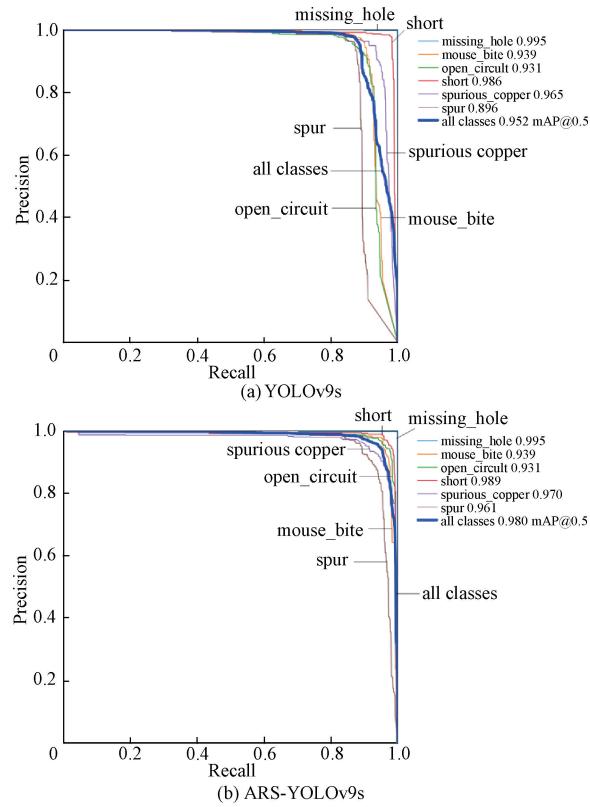


图 12 P-R 曲线对比图

Fig. 12 P-R curve comparison chart

3) 模型验证分析

为了直观展示 ARS-YOLOv9s 算法在缺陷检测方面的性能,本文选取了改进前后的模型,并对 6 种不同类型缺陷的检测效果进行了可视化对比。相关的检测结果分别呈现在图 13 中。在预测框上标注了预测的缺陷类别和置信度,通过对比可以发现,本文提出的算法在整体置信度上超过了 YOLOv9s 算法,特别是在缺孔、短路和余铜缺陷的检测上表现更为出色。

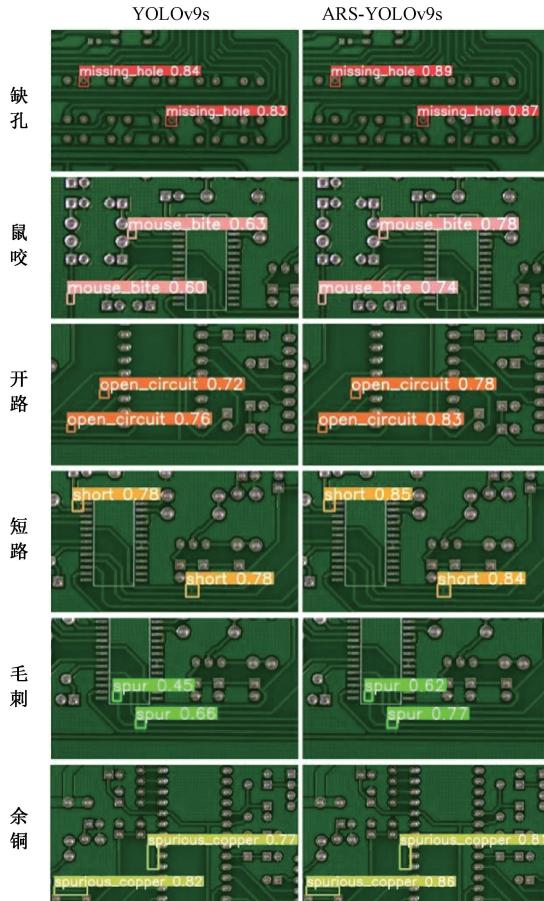


图 13 YOLOv9s 与本文算法各类缺陷检测效果对比

Fig. 13 Comparison between YOLOv9s and this paper's algorithm for detecting various types of defects

由于鼠咬、开路和毛刺这 3 类缺陷在视觉上具有较高的相似性,且其特征细节较为细微,这导致在网络训练过程中容易丢失关键的轮廓信息,从而引发分类错误和定位不准确,降低了检测的准确率。然而,本文算法有效地解决了这些问题,减少了误检和漏检的情况,显著提升了 YOLOv9s 算法在 PCB 缺陷检测方面的性能。

3 结 论

为进一步提高 PCB 缺陷检测的精度与速度,提出了一种基于 ARS-YOLOv9s 的 PCB 缺陷检测方法。在网络结构中引入 AFPN 以缩小非相邻层间的语义差异,减少信息

丢失,增强特征表达的连贯性;针对微小缺陷特征信息在深度网络特征提取不充分的问题,引入 iRMB 注意力机制模块,从而增加模型检测小目标的权重;通过删除大目标检测层 P5,新增微小目标检测层 P2 并调整网络结构来保持模型的轻量化;选取 Shape-IoU 以解决由于宽高比定义不明确、对质量差的回归样本过于敏感从而导致正负样本不均衡等局限性的问题。经验证,本算法在检测精度、速度和鲁棒性方面均超越了其他算法。该算法的轻量化设计满足工业小型计算机系统对 PCB 板进行实时检测的需求,对提升印刷电路板工业生产检测环节的效率和质量具有重大意义。未来的研究将采集更多的缺陷样本加入训练集,深入探讨更复杂的缺陷类型、小样本学习问题,以及在复杂环境下的 PCB 表面缺陷检测。

参考文献

- [1] LING Q, ISA N A M. Printed circuit board defect detection methods based on image processing, machine learning and deep learning: A survey [J]. IEEE Access, 2023, 11: 15921-15944.
- [2] ZHENG J F, SUN X P, ZHOU H X, et al. Printed circuit boards defect detection method based on improved fully convolutional networks [J]. IEEE Access, 2022, 10: 109908-109918.
- [3] 余翔, 靳闪闪, 杨路. 一种改进多尺度特征融合的交通标志识别算法 [J]. 电讯技术, 2024, 64(12): 1955-1962.
YU X, JIN SH SH, YANG L. An improved traffic sign recognition algorithm with multi-scale feature fusion [J]. Telecommunication Engineering, 2024, 64(12): 1955-1962.
- [4] 周贤勇, 祝俊辉, 王圆, 等. 基于改进 YOLOv7 算法的 PCB 裸板缺陷检测 [J]. 无线电工程, 2023, 53(12): 2791-2797.
ZHOU X Y, ZHU J H, WANG Y, et al. Defect detection of PCB bare boards based on improved YOLOv7 algorithm [J]. Radio Engineering, 2023, 53(12): 2791-2797.
- [5] 胡江宇, 贾树林, 马双宝. 基于改进级联 Faster RCNN 的 PCB 表面缺陷检测算法 [J]. 仪表技术与传感器, 2022(7): 106-110, 126.
HU J Y, JIA SH L, MA SH B. PCB surface defect detection algorithm based on improved cascade faster RCNN [J]. Instrumentation Technology and Sensors, 2022(7): 106-110, 126.
- [6] 吴兴辉, 何赟泽, 周辉, 等. 改进 YOLO v7 算法下的监控水域环境人员识别研究 [J]. 电子测量与仪器学报, 2023, 37(5): 20-27.
WU X H, HE Y Z, ZHOU H, et al. Research on

- personnel identification in monitored aquatic environments using improved YOLO v7 algorithm[J]. Journal of Electronic Measurement and Instrumentation, 2023, 37(5): 20-27.
- [7] 张莹, 邓华宣, 王耀南, 等. 基于多通道特征融合学习的印制电路板小目标缺陷检测[J]. 仪器仪表学报, 2024, 45(5): 10-19.
ZHANG Y, DENG H X, WANG Y N, et al. Small target defect detection of printed circuit boards based on multi-channel feature fusion learning[J]. Chinese Journal of Scientific Instrument, 2024, 45(5): 10-19.
- [8] 王军, 伍毅, 陈正超. 基于 SMT-YOLOv8 的 PCB 缺陷检测研究[J]. 电子测量技术, 2024, 47(11): 131-137.
WANG J, WU Y, CHEN ZH CH. Research on PCB defect detection based on SMT-YOLOv8 [J]. Electronic Measurement Technology, 2024, 47(11): 131-137.
- [9] YANG G Y, LEI J, ZHU Z K, et al. AFPN: Asymptotic feature pyramid network for object detection[C]. 2023 IEEE International Conference on Systems, Man, and Cybernetics (SMC), 2023: 2184-2189.
- [10] ZHANG J N, LI X T, LI J, et al. Rethinking mobile block for efficient attention-based models[C]. 2023 IEEE/CVF International Conference on Computer Vision (ICCV). IEEE Computer Society, 2023: 1389-1400.
- [11] ZHANG H, ZHANG SH J. Shape-IoU: More accurate metric considering bounding box shape and scale[J]. ArXiv preprint arXiv:2312.17663, 2023.
- [12] WANG C Y, YEH I H, LIAO H Y M. YOLOv9: Learning what you want to learn using programmable gradient information[C]. Computer Vision - ECCV 2024;18th European Conference, 2024: 1-21.
- [13] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: Unified, real-time object detection[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2016: 779-788.
- [14] HUANG G, LIU Z, VAN DER MAATEN L, et al. Densely connected convolutional networks[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2017: 4700-4708.
- [15] HE K M, ZHANG X Y, REN SH Q, et al. Deep residual learning for image recognition [C]. IEEE Conference on Computer Vision and Pattern Recognition, 2016: 770-778.
- [16] LIU S T, HUANG D, WANG Y H. Learning spatial fusion for single-shot object detection [J]. ArXiv preprint arXiv:1911.09516, 2019.
- [17] SANDLER M, HOWARD A, ZHU M, et al. Mobilenetv2: Inverted residuals and linear bottlenecks[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2018: 4510-4520.
- [18] VASWANI A, SHAZER N, PARMAR N, et al. Attention is all you need [C]. 31st International Conference on Neural Information Processing Systems (NIPS'17), 2017: 6000-6010.
- [19] LI X, WANG W H, WU L J, et al. Generalized focal loss: Learning qualified and distributed bounding boxes for dense object detection [J]. Advances in Neural Information Processing Systems, 2020, 33: 21002-21012.
- [20] ZHENG ZH H, WANG P, REN D W, et al. Enhancing geometric factors in model learning and inference for object detection and instance segmentation[J]. IEEE Transactions on Cybernetics, 2021, 52(8): 8574-8586.
- [21] 刘凡凡, 朱成梅, 赵娜娜, 等. 基于多模态融合的遥感小目标检测[J]. 激光与光电子学进展, 2024, 61(24): 296-305.
LIU F F, ZHU CH M, ZHAO N N, et al. Remote sensing small target detection based on multimodal fusion[J]. Laser & Optoelectronics Progress, 2024, 61(24): 296-305.
- [22] SU D, JIANG Q H, LIU EN H, et al. Improving optimizers by Runge-Kutta method: A case study of SGD and Adam [C]. 2024 12th International Conference on Intelligent Control and Information Processing(ICICIP). IEEE, 2024: 47-53.
- [23] 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.

作者简介

- 陈枫赟,硕士研究生,主要研究方向为机器视觉、图像处理。
E-mail: 772771670@qq.com
- 李鹏(通信作者),教授,博士生导师,主要研究方向为超声成像、通信信号处理、北斗定位、图像处理等。
E-mail: lipengnust@163.com
- 张翔凯,硕士研究生,主要研究方向为机器视觉。
E-mail: 583385427@qq.com
- 于涛,硕士研究生,主要研究方向为视觉控制。
E-mail: 2544831633@qq.com
- 余珺泽,硕士研究生,主要研究方向为机器视觉。
E-mail: 947953946@qq.com