基于 EfficientNetV2 的 PCB 缺陷检测算法

尹嘉超, 吕耀文*, 索科, 黄玺

(长春理工大学光电工程学院　长春　130013)
(lyw@cust.edu.cn)

摘　要: 印刷电路板(PCB)是一种高精密的电子元器件, 其优良与否对电子产品的质量有着重要影响. 但现有的 PCB 缺陷检测算法存在着检测精度不高, 特别是缺陷定位不够精确等问题. 针对以上问题, 提出一种基于 EfficientNetV2 的 PCB 缺陷检测算法. 在 Faster R-CNN 的基础上, 通过选用特征提取能力更强的 EfficientNetV2_M 作为特征提取网络, 同时使用通道注意力机制(ECA)对特征融合网络 FPN 进行优化, 提高了细节信息提取能力. 在北京大学智能机器人开放实验室发布的 PCB 瑕疵数据集上的实验结果表明, 相较于目前检测效果最好的 PCB 缺陷检测算法 LWN-Net, 改进后的缺陷检测算法在 IoU=0.50 时 mAP 由 99.58%提升到 99.66%; 在 IoU=[0.50:0.95]时 mAP 由 52.6%提升到 79.4%. 该网络在提升了 PCB 的检测精度的同时, 解决了缺陷定位不够精确的问题, 实现了高精度的 PCB 缺陷检测, 具有一定的实际意义. 代码已经开源在 https://github.com/ChaO989/Defect_detection.

关键词: 印刷电路板; EfficientNetV2; 缺陷检测; Faster R-CNN; 高效通道注意力
中图法分类号: TP391.41; TN41　　**DOI:** 10.3724/SP.J.1089.2023-00551

PCB Defect Detection Algorithm Based on EfficientNetV2

Yin Jiachao, Lyu Yaowen*, Suo Ke, and Huang Xi

(*Changchun University of Science and Technology, School of Optoelectronic Engineering, Changchun*　130013)

Abstract: Printed circuit boards (PCBs) are crucial electronic components with a significant impact on the quality of electronic products. However, existing PCB defect detection algorithms suffer from limitations in detection accuracy, especially in precise defect localization. To address these challenges, this paper proposes a novel PCB defect detection algorithm based on EfficientNetV2. Building upon the Faster R-CNN architecture, we leverage the powerful feature extraction capabilities of EfficientNetV2_M as the backbone network. Furthermore, we optimize the feature fusion network FPN (feature pyramid network) by incorporating the channel attention mechanism (ECA) to enhance the extraction of intricate details. Experimental results on the PCB defect data set released by Peking University Intelligent Robot Open Laboratory demonstrate that the proposed defect detection algorithm outperforms the state-of-the-art LWN-Net in terms of mAP (mean average precision) at IoU=0.50, with an increase from 99.58% to 99.66%. Moreover, at IoU=[0.50:0.95], the mAP significantly improves from 52.6% to 79.4%. Consequently, the enhanced algorithm achieves highly accurate PCB defect detection while addressing the issue of imprecise defect localization, thus holding substantial practical significance. The code is available at https://github.com/ChaO989/Defect_detection.

Key words: printed circuit board; EfficientNetV2; defect detection; Faster R-CNN; efficient channel attention

---

收稿日期: 2023-08-17; 修回日期: 2024-02-07. **基金项目**: 吉林省自然科学基金(YDZJ202201ZYTS419); 高等学校学科创新引智 111 计划(D21009). **尹嘉超**(1998—), 男, 硕士研究生, 主要研究方向为图像处理、深度学习; **吕耀文**(1987—), 男, 博士, 副教授, 硕士生导师, 论文通信作者, 主要研究方向为光电检测与图像处理; **索科**(1999—), 男, 硕士研究生, 主要研究方向为图像处理; **黄玺**(1998—), 男, 硕士研究生, 主要研究方向为图像处理、深度学习.

随着电子制造业的发展，电子产品趋于多功能化、微型化和智能化. 印刷电路板(printed circuit board，PCB)作为重要的精密零部件，其质量对电子产品的功能和寿命有着重要影响. 然而，PCB 的制作过程复杂且受多种因素影响，可能会出现各种缺陷如 missing_hole(缺孔)，mouse_bite(鼠咬)，open_circuit(开路)，short(短路)，spur(毛刺)和 spurious_copper(余铜)等. 因此，在 PCB 生产过程中准确、高效地进行缺陷检测变得非常必要[1]. 早期的人工目视检测和线上仪器检测方法存在诸多限制[2]，而基于机器视觉的检测算法，尤其是基于深度学习的缺陷检测算法，因其低成本和高效率而成为 PCB 质检的主流方式.

基于深度学习的缺陷检测算法主要分为精度高的两阶段模型和速度快的单阶段模型，其中常用的两阶段模型为 Faster R-CNN，单阶段模型为 YOLO 系列. 2022 年，胡江宇等[3]通过在 Faster R-CNN 中添加过滤特征的金字塔网络增强了模型的适用性，但其平均分类精度为 98.91%，仍有提升空间. 2023 年，孙志超等[4]在 Faster R-CNN 的基础上，将原始的骨干网络 VGG16[5]替换为可变形残差卷积模块，并利用一种伸缩式改进的特征金字塔网络(feature pyramid network，FPN)与原区域建议网络融合，达到了 99.41%的检测精度，但其检测速度较慢，仅为 6.47 FPS，无法满足实时检测的需求. 2024 年，文斌等[6]将 YOLOv3 的骨干网络替换为轻量级网络 LWN-Net，提出的新型 FPN 达到了 99.58%的检测精度，但其在 IoU=[0.50:0.95]时的检测精度为 52.6%，仍有很大的提升空间. 考虑 PCB 表面缺陷复杂多样，部分缺陷与背景相似度过高，对准确率要求较高，并且缺陷的平均大小为 71×68 像素、平均原图占比仅为 0.085%，属于较小目标. 因此，本文选用检测精度更高、能更好地进行小目标检测的 Faster R-CNN 作为 PCB 表面缺陷检测的网络框架.

为了进一步提高 Faster R-CNN 的检测精度和检测速度，本文提出一种基于 EfficientNetV2 的 PCB 缺陷检测算法. 首先选用模型更小，特征提取能力更强的 EfficientNetV2_M 作为 Faster R-CNN 的骨干特征提取网络，提高模型的特征提取能力，其次提出了高效特征融合网络(efficient feature pyramid network，EFPN)作为特征融合部分，增强缺陷细节，提高模型的检测效果.

1　改进后的 Faster R-CNN 模型

改进后的 Faster R-CNN 模型结构如图 1 所示，主要由特征提取网络 EfficientNetV2_M，EFPN，候选区域生成网络(region proposal network，RPN)和计算损失，以及进行预测的头部网络组成.

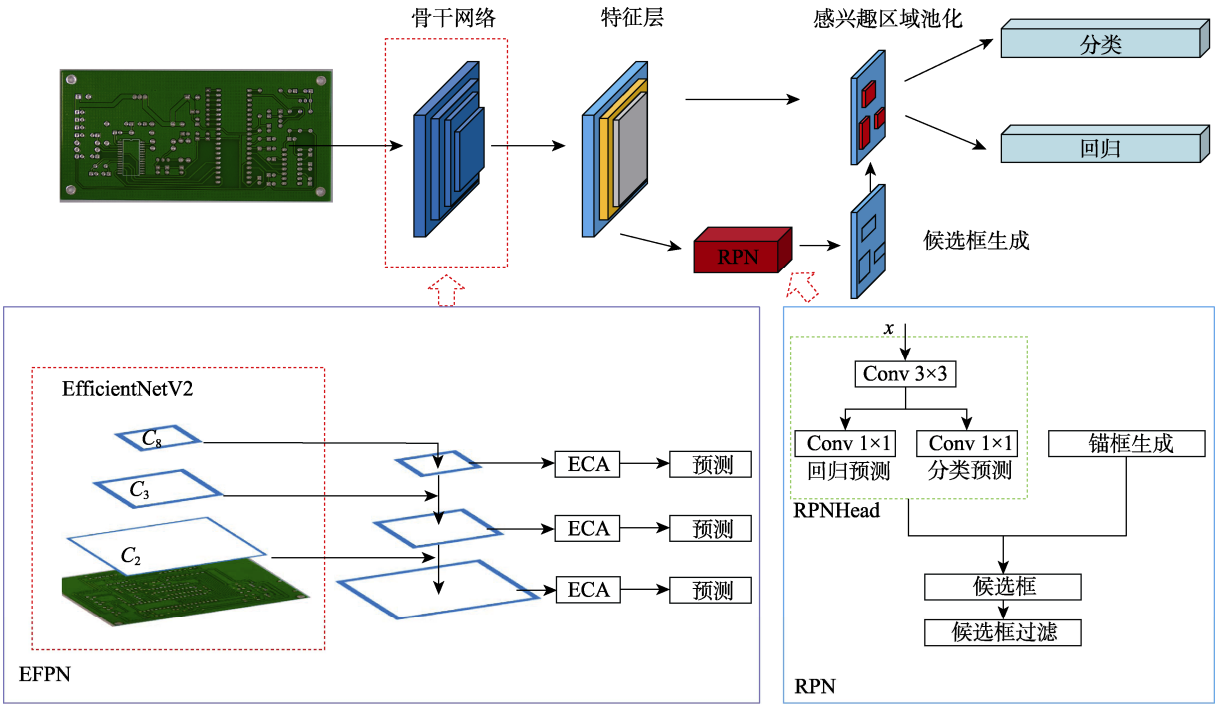


图 1　改进后的 Faster R-CNN 网络结构

改进后的 Faster R-CNN 算法流程图如图 2 所示. 首先将 PCB 缺陷图像传入 EfficientNetV2_M 中进行特征提取; 其次将特征层中的第 C_2, C_3 和 C_8 层送到 EFPN 中进行特征融合; 之后 RPN 根据融合后的特征层生成候选区域, 并将候选区域映射到各个特征层; 再经过感兴趣区域(region of interest, ROI)池化操作, 分别使用交叉熵损失和平滑 L_1 损失(Smooth L_1 loss)计算分类损失与回归损失.

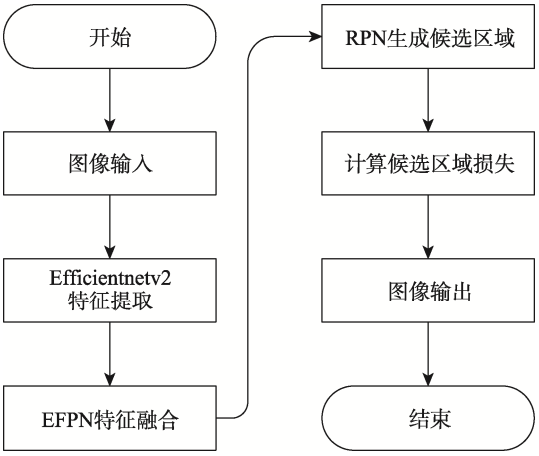


图 2 改进后的 Faster R-CNN 算法流程

1.1 EfficientNetV2

传统的 Faster R-CNN 骨干网络为 VGG16, 模型参数量过大, 为 138M[7], 尽管它在某些任务上表现良好, 但其相对浅层的结构无法有效地提取复杂特征, 导致推理时间较长且检测精度较低, 不能满足对 PCB 的高要求检测需求. 此外, PCB 图像的平均分辨率大小为 2 777×2 138 像素, 选用经典的网络虽然可以实现对复杂特征的提取, 但随着图像尺寸增大, 计算开销逐渐增加, 训练效率也越来越低, 不利于网络后续的重训练与改进. 而 EfficientNetV2_M 采用改进的渐进式学习策略提升训练效率, 在训练过程中通过逐渐增大图像尺寸来加快训练速度, 之后依据图像的大小自适应地调整正则化, 以弥补检测准确率下降的问题. 因此本文选用模型参数量更小、训练效率更高、特征提取能力更强的 EfficientNetV2_M 作为 Faster R-CNN 的骨干网络, 其网络详细参数如表 1 所示. 其中, k 为卷积核大小, SE[8](squeeze and excitation)为压缩和抑制结构.

EfficientNet 是 Google 团队 2019 年提出的一个图像分类网络, 具有很强的特征提取能力[9]. 与传统网络单一地调整网络宽度、深度和分辨率以提升精度不同的是, 其提出了一种多维度混合的放缩模型. 首先固定可用的网络资源为 1, 用网格搜索法确定网络宽度、深度和分辨率的最佳分配比例; 之后随着网络资源的增加, 等比例地对宽度、深度和分辨率进行放大, 以保证深度、宽度和分辨率的最佳平衡比. 该放缩模型使 EfficientNet 同时兼顾了速度与精度. 而 EfficientNetV2 在保证了速度与精度的情况下, 还提升了训练速度[10], 其主要采用 Fused-MBConv 和 MBConv 这 2 种结构, 具体结构如图 3 所示.

表 1 EfficientNetV2_M 网络参数

阶段	操作	步长	通道数	层数
0	Conv3×3	2	24	1
1	Fused-MBConv1, k3×3	1	24	3
2	Fused-MBConv4, k3×3	2	48	5
3	Fused-MBConv4, k3×3	2	80	5
4	MBConv4, k3×3, SE0.25	2	160	7
5	MBConv6, k3×3, SE0.25	1	176	14
6	MBConv6, k3×3, SE0.25	2	304	18
7	MBConv6, k3×3, SE0.25	1	512	5
8	Conv1×1&Pooling&FC		1 280	1

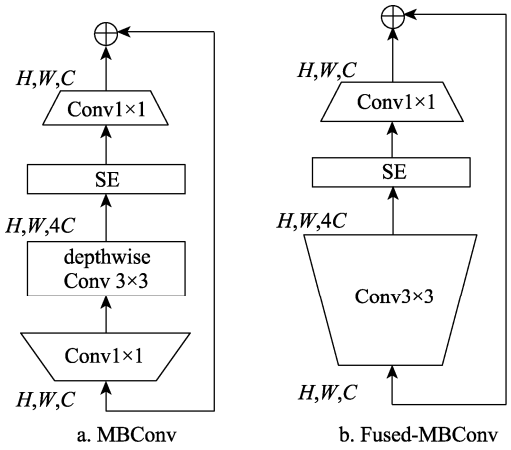


图 3 MBConv 与 Fused-MBConv 对比图

EfficientNet 中使用的是 MBConv 结构, 虽然在参数与 FLOPs 上都少于 Fused-MBConv, 但因无法充分地利用现有的加速器, 使得实际训练速度没有结构简单的 Fused-MBConv 快[11]. 但是, 如果将所有 MBConv 结构都替换为 Fused-MBConv 结构, 会明显增加参数量以及 FLOPs, 训练速度也会降低. 因此 Tan 等[10]使用神经架构搜索(neural architecture search, NAS)技术去搜索 MBConv 与 Fused-MBConv 的最佳组合.

为验证骨干网络的有效性, 使用 Grad-CAM[12]对骨干网络更换前后进行可视化分析, 结果如图 4 所示.

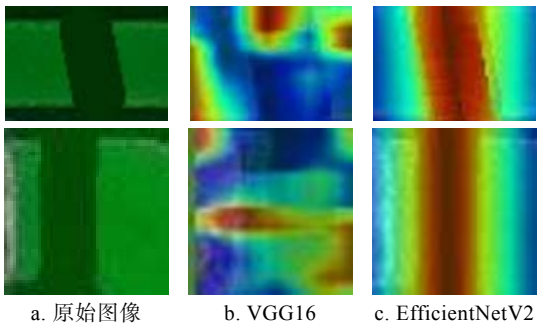


图 4　Grad-CAM 可视化结果

如图 4b 所示，更换骨干网络前进行可视化时，其热力图关注点(颜色越趋近于红色，其关注程度越强)会有一些偏差，关注范围过于分散；如图 4c 所示，在选用 EfficeientNetV2 作为骨干网络时模型的关注位置基本覆盖了缺陷位置，主要集中在缺陷周围．对比结果表明，选用 EfficientNetV2 作为特征提取网络时，网络的特征提取能力更强，模型也更加关注于主要表达特征的通道．

1.2　EFPN

在深度学习中，网络越深，其含有的语义信息也越丰富，但随之而来的是细节信息的缺失，在 FPN[13]中也是如此．FPN 每层均具备多层特征信息[14]，但不同层级之间的特征存在着语义鸿沟，它们直接进行融合会导致大量的信息冗余和冲突[15]，不利于 PCB 的缺陷检测．因此，在进行特征融合时，需要精选适用的特征层．在 Faster R-CNN 的 FPN 中，不同特征层具有不同的优势．低层特征能够保留更多底层细节信息，更适合用于小目标检测；而高层特征包含更多语义信息，更适合用于大目标检测．为了确定哪些特征层最适合本文的 PCB 缺陷检测任务，需要分析 PCB 缺陷数据集中缺陷的尺寸分布情况，具体分布如图 5 所示．这样的统计分析能够帮助优化特征层的选择，以更好地满足本文的目标需求．

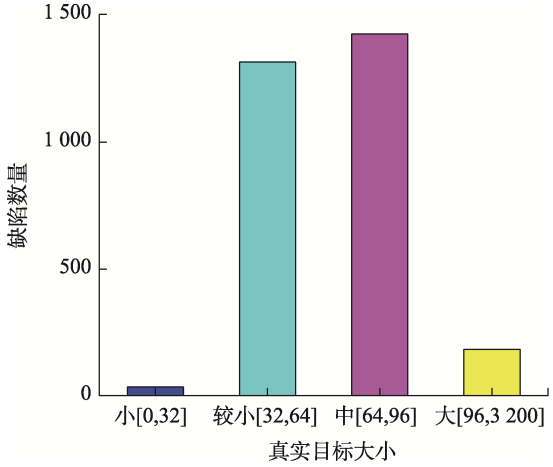


图 5　PCB 数据集 GT 分布

由图 5 可以得出，PCB 数据集的真实目标大小主要为较小目标，适合采用低层特征进行检测，因此选用低层特征 C_2 与 C_3 进行融合；同时为了获取更多的语义信息，也将高层特征 C_8 加入到了融合中．本文在此基础上提出 EFPN，其具体结构如图 6 所示．首先选用 EfficientNetV2 中能获得更多细节特征的低层特征 C_2 和 C_3，与含有更多语义信息的 C_8 进行特征融合，并且在融合后添加了高效通道注意力(efficient channel attention, ECA)模块[16]，以提高有用信息在特征融合中的占比，并甄选出有用的信息用于检测．与原 FPN 相比，改进后的模型缩短了顶层与底层之间的距离，减少了特征融合的次数，在一定程度上避免了因融合造成的细节信息丢失，提高了模型对小目标的检测能力．同时，添加的 ECA 模块能够更加关注对特征表达

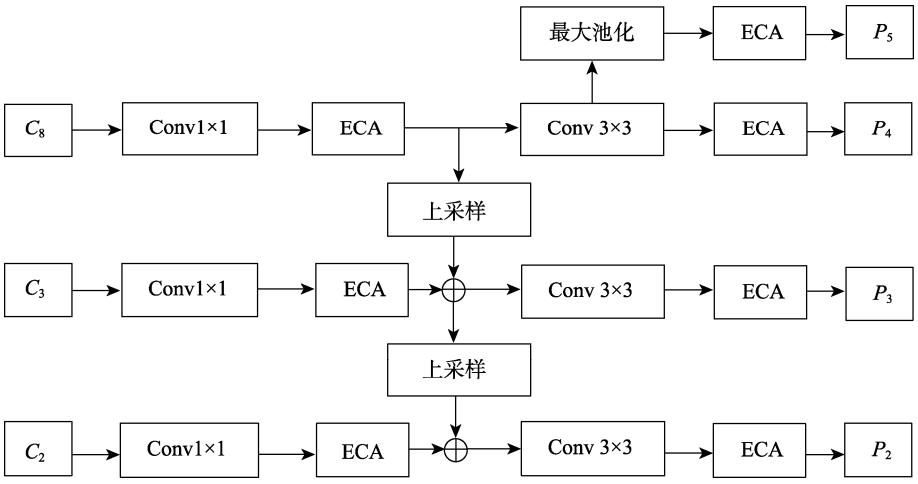


图 6　EFPN 结构图

更有效的通道，并对不主要的特征进行抑制，使得改进后的结构不仅具备更准确的表达能力，而且可以提高模型的抗噪声能力.

ECA 模块最早是 SE 结构的基础上提出的避免降维、高效捕获跨通道交互的注意力模块[17]，其降低了 SE 模块的降维操作对预测能力的副作用，并指出捕捉所有通道之间的依赖关系效率低且不是必要的，其结构如图 7 所示. 其中，χ 为输入特征图 $\tilde{\chi}$ 为输出特征图, GAP 为全局平均池化.

首先将输入特征图进行全局平均池化，特征图从[H, W, C]的矩阵变成[1, 1, C]的向量；其次根据特征图的通道数计算得到自适应的一维卷积核大小

$$k=\varphi(C)=\left|\frac{\mathrm{lb}(C)}{\gamma}+\frac{b}{\gamma}\right|_{\mathrm{odd}} \tag{1}$$

其中，$|t|_{\mathrm{odd}}$ 表示 t 的最接近奇数；C 为输入特征通道数，当 $\gamma=2$, b=1 时取得最佳 k 值. 之后通过卷积核大小为 k 的一维卷积得到特征图每个通道的权重，并将归一化权重和原输入特征图逐通道相乘，生成加权后的特征图. ECA 模块具体结构如图 7 所示.

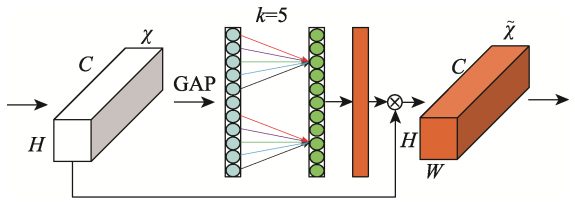


图 7 ECA 模块结构图

1.3 损失函数

改进后的 Faster R-CNN 模型损失函数主要由分类损失和定位损失 2 部分组成. 分类损失选用的交叉熵损失函数能够很好地衡量模型预测概率分布与实际标签之间的差异，特别适用于多类别分类问题；定位损失使用的是 Smooth L_1 损失，可以帮助提高模型在回归问题中的稳定性，并减少异常值对损失函数的影响.

1.3.1 交叉熵损失函数

在深度学习中，交叉熵损失经常与 Softmax 激活函数一起使用，通过比较模型预测的概率分布与实际标签之间的差异来度量模型的性能[18]. 交叉熵损失的定义为

$$H(y,\hat{y})-\sum_{i=1}^{n}y_i\ln\hat{y}_i \tag{2}$$

其中，y_i 表示样本真实分布的概率；$\hat{y}_i$ 表示样本预测概率. 在训练过程中，模型通过最小化交叉熵损失来调整参数，使得模型的预测更接近于实际标签，从而提高分类准确性.

1.3.2 Smooth L_1 损失

Smooth L_1 损失是一种用于回归问题的损失函数，常用于目标检测和物体定位任务中. Smooth L_1 损失在较小的误差范围内采用 L_1 损失函数，在较大的误差范围内采用 L_2 损失函数[19]. 这种设计使得 Smooth L_1 损失在面对离群点时比均方误差更稳定. Smooth L_1 损失的定义为

$$\text{Smooth } L_1(x)=\begin{cases}0.5x^2, & |x|<1\\ |x|-0.5, & |x|\geqslant 1\end{cases} \tag{3}$$

其中，x 表示预测值和真实值之间的差异. 与传统的均方差损失函数相比，它能够更好地处理异常值，并提高模型的稳定性和鲁棒性.

2 实 验

2.1 实验环境以及数据

本实验所采取的硬件环境为 Intel(R) Xeon(R) Platinum 8255C CPU，内存 43 GB，显卡为 RTX 3090，显存 24 GB，使用 CUDA 11.6 和 cuDNN 8.3.2 进行 GPU 加速；软件环境为 Python3.8，Opencv 和 Pytorch 1.12.1.

实验数据集采用北京大学智能机器人开放实验室发布的 PCB 瑕疵数据集[20]，后经过镜像、旋转等操作将数据集扩充到 10 354 幅图像，其中训练集、验证集、测试集的比例为 8∶1∶1，扩充后的数据集中各缺陷实际分布情况如表 2 所示.

表 2 PCB 数据集缺陷分布表

缺陷类型	图像数量	缺陷数量
缺孔	1 716	7 425
鼠咬	1 716	7 356
开路	1 733	7 207
短路	1 739	7 362
毛刺	1 713	7 288
余铜	1 737	7 534
合计	10 354	44 172

数据集共有缺孔、鼠咬、开路、短路、毛刺和余铜 6 类缺陷，各种缺陷如图 8 所示. 因为选用的是平均分辨率为 2 777×2 138 像素的未经过裁剪的原图像，所以检测速度相对于其他进行剪裁过的图像(600×600 像素)会稍慢一点，但直接使用原图像进行检测符合实际工业生产流程，检测结果更符合实际.

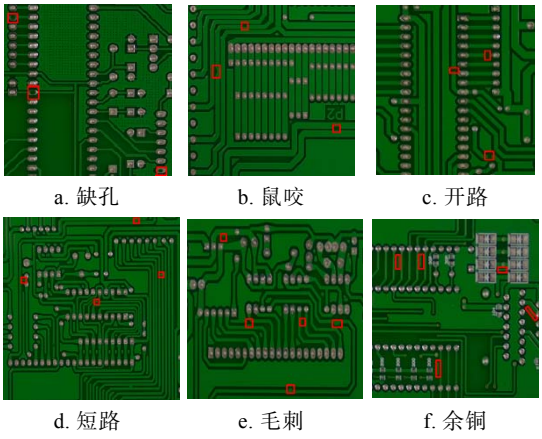


图 8　PCB 缺陷图像

2.2　实验结果与分析

2.2.1　模型训练

为使模型有更快的收敛速度，采用迁移学习的思想，载入 Pytorch 官方 EfficientNetV2_M 的预训练权重作为模型的初始化权重，训练 50 个 epoch 后模型基本收敛，损失不再下降. 预测效果最佳的损失曲线以及评估指标平均精度均值(mean verage precision, mAP)、召回率如图 9 所示.

在图 9 所示的召回率曲线中，只有开路与短路的召回率没有达到 100%存在漏检现象. 分析其原因发现，通过在测试集中找出漏检目标找到漏检原因主要有 2 个：(1) 缺陷位置在两孔之间，与两孔边缘连在一起较难区分如图 10a 所示；(2) 真实锚框较小，而检测框较大导致 2 个锚框的交并比(intersection over union, IoU)较低，如图 10b 所示.

2.2.2　不同模块性能对比

为了更好地评估本文改进的缺陷检测算法对于 PCB 缺陷目标的检测性能，对本文所提出的模块进行消融实验，结果如表 3 所示. 其中，IoU=0.50 时，平均召回率(mRecall)为 6 个类别召回率(Recall)取平均算出；IoU=[0.50:0.95]时，其值由 coco 工具箱计算得出. 实验 2 选用 EfficientNetV2_M 作为 Faster R-CNN 的骨干网络，比实验 1 的初始网络

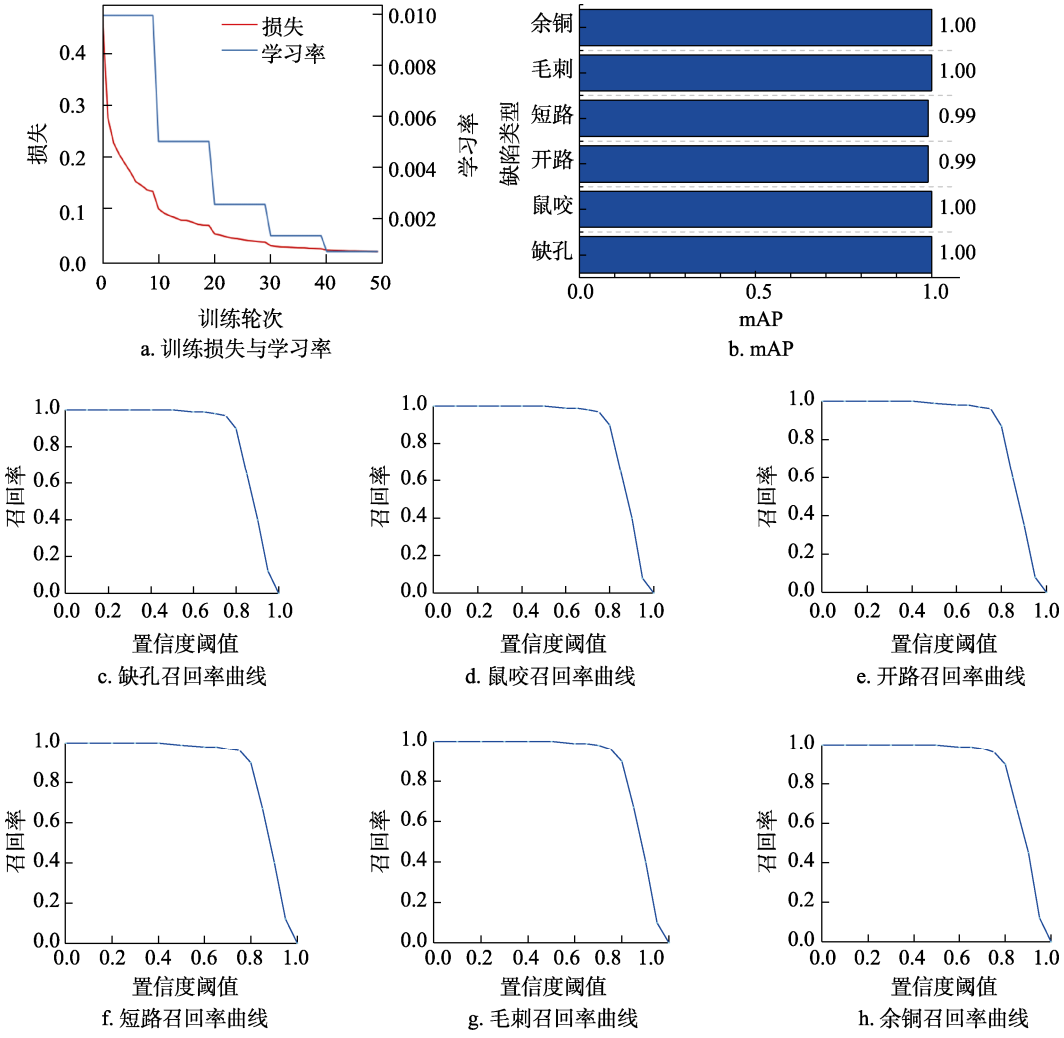


图 9　对不同缺陷目标检测的损失曲线及评价指标对比



a. 开路漏检 b. 短路漏检

图 10 PCB 漏检图像

VGG16 在 IoU=0.50 时 mAP 提升了 0.26 个百分点, mRecall 提升了 1.26 个百分点; 在 IoU=[0.50:0.95]时, mAP 提升了 2.4 个百分点, mRecall 提升了 1.7 个百分点, 从而验证了 EfficientNetV2_M 的有效性. 实验 3 使用最为常见的 FPN 进行特征融合, 比未进行特征融合的实验 2 在 IoU=0.50 时 mAP 提升了 0.94 个百分点, mRecall 提升了 0.53 个百分点; 在 IoU=[0.50:0.95]时 mAP 提升了 10.5 个百分点, mRecall 提升了 9.3 个百分点. 实验 4 使用 EFPN 进行特征融合, 比实验 3 在 IoU=0.50 时 mAP 提升了 0.35 个百分点, mRecall 提升了 0.02 个百分点; 在 IoU=[0.50:0.95]时 mAP 提升了 12.2 个百分点, mRecall 提升了 10.6 个百分点, 进一步提高了检测精度, 从而验证了 EFPN 的有效性.

2.2.3 不同骨干网络性能对比

在相同的实验条件下, 选取不同的经典骨干网络进行对比, 其 mAP 与每秒帧数(frames per second, FPS)如表 4 所示. 虽然本文缺陷检测算法的 FPS 略低于 ResNet50[21], MobileNetV3[22]与 GhostNetv2[23], 但 mAP 有较大提升, 相比于 4 种对比算法, 其在 IoU=[0.50:0.95]时 mAP 分别提升 14.2 个百分点, 20.6 个百分点和 31.3 个百分点; 在 IoU=0.50 时 mAP 分别提升 1.14 个百分点, 2.44 个百分点和 8.56 个百分点. 当选取 EfficientNet 作为骨干时, mAP 与 FPS 均低于本文网络, 即其性能低于本文的骨干网络. 综上所述, 与其他骨干网络相比, 本文使用 EfficientNetV2_M 骨干网络提取特征, 模型的检测精度更高, 效果更好.

表 3 消融实验性能对比

实验	VGG16[5]	EfficientNetV2_M	EFPN	FPN	IoU=0.50		IoU=[0.50:0.95]	
					mAP/%	mRecall/%	mAP/%	mRecall/%
1	√				98.11	98.14	54.3	62.7
2		√			98.37	99.40	56.7	64.4
3		√		√	99.31	99.93	67.2	73.7
4		√	√		**99.66**	**99.95**	**79.4**	**84.3**

注. 粗体表示最优值.

表 4 不同骨干网络性能对比

骨干网络	mAP/%		FPS/(帧·s^{-1})
	IoU=0.50	IoU=[0.50:0.95]	
ResNet50[21]	98.52	65.2	16.75
MobileNetV3[22]	97.22	58.8	**17.03**
EfficientNet[9]	98.90	61.7	14.39
GhostNetv2[23]	91.10	48.1	16.90
本文	**99.66**	**79.4**	15.48

注. 粗体表示最优值.

2.2.4 EFPN 中不同层融合结果对比

在 EFPN 中, 融合不同特征层会导致不同的检测效果. 为了更深入地评估参与特征融合的各特征层的准确性, 本文在保持其他网络结构不变的情况下进行了对比实验. 选择不同的特征层组合进行融合, 实验结果如表 5 所示. 根据对比结果, 得出, 挑选的特征层组合 C_2, C_3 和 C_8 最适用于 PCB 数据集.

表 5 不同特征层融合对比

特征层组合	mAP/%	
	IoU=0.50	IoU=[0.50:0.95]
2, 4, 8	99.03	73.6
3, 4, 8	99.01	71.3
2, 3, 7	99.20	77.8
本文	**99.66**	**79.4**

注. 粗体表示最优值.

2.2.5 相同缺陷场景不同检测算法对比

为了验证改进后的 Faster R-CNN 在 PCB 检测算法中的性能, 在相同数据集上与其他算法[3,6,17,24]进行实验对比, 当 IoU=0.50 和 IoU=[0.50:0.95]时, 所得数据如表 6 所示.

YOLO-J 算法通过注意力机制和改进的 PANet (path aggregation network)结构提升了小目标的检测能力, 但是检测精度却低于本文算法. 改进 YOLOv7 使用 FasterNet[25]作为骨干网络, 并且使用部分卷积能够更有效地提取空间特征, 同时, 通

表 6　相同缺陷场景下不同检测算法对比

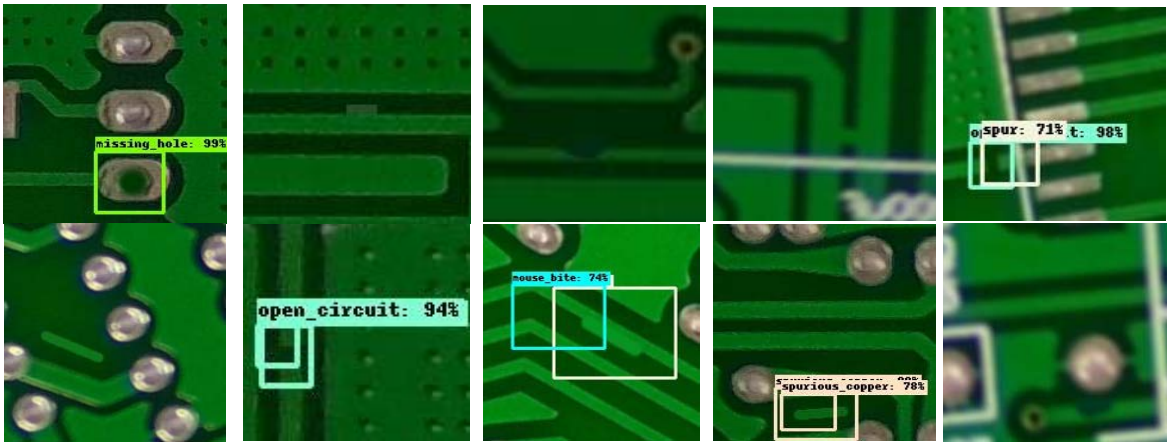
算法	mAP/%	
	IoU=0.50	IoU=[0.50:0.95]
YOLO-J[17]	99.02	51.70
改进 YOLOv7[24]	97.50	54.70
LWN-Net[6]	99.58	52.60
级联 Faster RCNN[3]	98.91	78.21
YOLOv8	99.40	77.50
本文	**99.66**	**79.40**

注．粗体表示最优值．

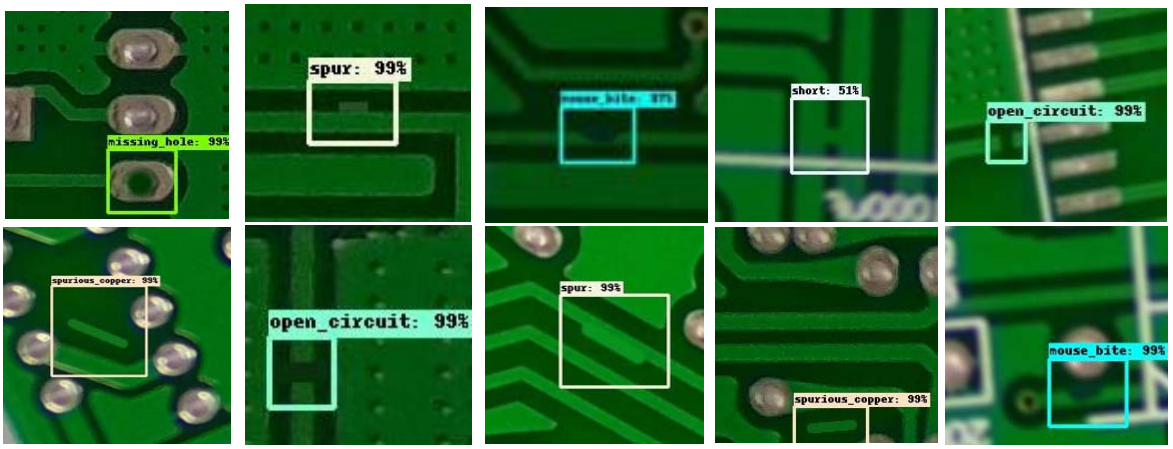
过减少冗余计算提高了检测速度，并将 CBAM (convolutional block attention module)与特征融合模块相结合，使模型能够更加关注相关特征，提高检测准确性；但其在 IoU=0.50 与 IoU=[0.50:0.95]时的检测精度均低于本文算法．LWN-Net 算法将轻量级特征增强网络作为 YOLOv3 的骨干网络保证了模型的轻量化，同时引入权重分配机制和新型特征金字塔网络，提升了模型精度．但其在 IoU=[0.50:0.95]上的平均精度较低，为 52.6%，相对本文算法而言定位不够精确．级联 Faster RCNN 选用轻量网络 MobileNetV3 作为骨干网络，并使用过滤 FPN 与多尺度特征融合进行优化，但其在 IoU=0.50 和 IoU=[0.50:0.95]时的检测精度均低于本文算法．YOLOv8 将 YOLOv5 骨干网络中的 C3(CSP bottleneck with 3 convolutions)结构换成了梯度流更丰富的 C2f(faster implementation of CSP bottleneck with 2 convolutions)结构，大幅度提高了模型性能，其中 CSP(cross stage partial)能够强化 CNN 的学习能力．同时 YOLOv8 还采用了 anchor- free 的思想来提升小目标的检测精度，虽然最终的检测结果与本文算法相近，但整体上仍然稍差于本文算法．综上所述，本文算法具有较高的检测精度，并且满足实时检测需求，其性能优于其他对比网络．

2.2.6　检测结果对比

从未参与训练的原始数据集中随机挑选 PCB 缺陷图像，在相同环境下分别使用改进后的 Faster R-CNN 和以 VGG16 为骨干网络的原始 Faster R-CNN 对它们进行检测，其对比结果如图 11 所示．



a. 原始 Faster R-CNN



b. 改进后的 Faster R-CNN

图 11　不同网络对 PCB 检测结果对比图

原始的 Faster R-CNN 对缺陷检测精度不高, 检测毛刺、鼠咬、短路、余铜等缺陷时存在漏检现象; 检测开路和余铜缺陷时存在重复检测误检的现象, 严重影响检测结果. 改进后的 Faster R-CNN 不仅没有出现漏检现象, 还解决了重复检测的问题, 同时在其他缺陷检测的缺陷的精度均优于原始的 Faster R-CNN. 综上所述, 本文算法检测精度更高, 在工业生产缺陷检测中能达到更好的检测效果.

3 结 语

本文提出了一种基于 EfficientNetV2 与 EFPN 的 Faster R-CNN 目标检测算法, 用于 PCB 缺陷目标检测研究中. 实验结果表明, 与原始 Faster R-CNN、级联 Faster RCNN, YOLOv7 与 LWN-Net 等目前主流的 PCB 缺陷检测算法相比, 本文算法在检测精度上均有提升. 在检测时间方面, 虽然本文的检测速度相对较低, FPS 仅为 15.48 帧/s, 但可满足实时检测需求, 能够适应实际工业生产中 PCB 缺陷目标的高效精准检测任务. 未来将进一步提高模型算法效率, 期望能够在提高模型检测速度同时提高检测精度, 并解决短路与开路的漏检问题.

参考文献(References):

[1] Qian Wanming, Zhu Hongping, Zhu Hongzhi, *et al*. Research on defect detection of PCB bare board based on adaptive weighted feature fusion[J]. Journal of Electronic Measurement and Instrumentation, 2022, 36(10): 92-99(in Chinese)
(钱万明, 朱红萍, 朱泓知, 等. 基于自适应加权特征融合的 PCB 裸板缺陷检测研究[J]. 电子测量与仪器学报, 2022, 36(10): 92-99)

[2] Dong Jingyi, Lv Wentao, Bao Xiaomin, *et al*. Research progress of the PCB surface defect detection method based on machine vision[J]. Journal of Zhejiang Sci-Tech University (Natural Sciences Edition), 2021, 45(3): 379-389(in Chinese)
(董静毅, 吕文涛, 包晓敏, 等. 基于机器视觉的 PCB 板表面缺陷检测方法研究进展[J]. 浙江理工大学学报(自然科学版), 2021, 45(3): 379-389)

[3] Hu Jiangyu, Jia Shulin, Ma Shuangbao. PCB surface defect detection algorithm based on improved cascaded faster RCNN[J]. Instrument Technique and Sensor, 2022(7): 106-110+126(in Chinese)
(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)

[4] Sun Zhichao, Wang Bo, Zhang Xiaoling. PCB defect detection based on deformable residual convolution and scalable feature pyramid algorithm[J]. Telecommunication Engineering, 2023, 63(6): 798-805(in Chinese)
(孙志超, 王博, 张晓玲. 基于可变形残差卷积与伸缩式特征金字塔算法的 PCB 缺陷检测[J]. 电讯技术, 2023, 63(6): 798-805)

[5] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition[OL]. [2023-08-17]. https://ora.ox.ac.uk/objects/uuid:60713f18-a6d1-4d97-8f45-b60ad8aebbce

[6] Wen Bin, Hu Hui, Yang Chao. PCB defect detection based on LWN-Net algorithm[J]. Computer Integrated Manufacturing Systems, 2024, 30(2): 496-507 (in Chinese)
(文斌, 胡晖, 杨超. 基于 LWN-Net 的 PCB 缺陷检测算法[J]. 计算机集成制造系统, 2024, 30(2): 496-507)

[7] Ren S Q, He K M, Girshick R, *et al*. Faster R-CNN: towards real-time object detection with region proposal networks[C] // Proceedings of the 28th International Conference on Neural Information Processing Systems. Cambridge: MIT Press, 2015: 91-99

[8] Hu J, Shen L, Sun G. Squeeze-and-excitation networks[C] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2018: 7132-7141

[9] Tan M X, Le Q V. EfficientNet: rethinking model scaling for convolutional neural networks[C]//Proceedings of the 36th International Conference on Machine Learning. New York: PMLR, 2019: 6105-6114

[10] Tan M X, Le Q V. EfficientNetV2: smaller models and faster training[C] //Proceedings of the 38th International Conference on Machine Learning. New York: PMLR, 2021: 10096-10106

[11] Li Yong, Wang Yang, Fang Xia, *et al*. Micro-motor armature surface defect detection method based on improved BiFPN[J]. Machine Tool & Hydraulics, 2022, 50(6): 1-8(in Chinese)
(李勇, 王杨, 方夏, 等. 基于改进 BiFPN 的微特电机电枢表面缺陷检测方法[J]. 机床与液压, 2022, 50(6): 1-8)

[12] Selvaraju R R, Cogswell M, Das A, *et al*. Grad-CAM: visual explanations from deep networks via gradient-based localization[C] //Proceedings of the IEEE International Conference on Computer Vision. Los Alamitos: IEEE Computer Society Press, 2017: 618-626

[13] Lin T Y, Dollár P, Girshick R, *et al*. Feature pyramid networks for object detection[C] //Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2017: 936-944

[14] Qiao Meiying, Shi Jianke, Li Bingfeng, *et al*. Enhanced FPN underwater small target detection with improved loss function[J]. Journal of Computer-Aided Des ign & Computer Graphics, 2023, 35(4): 525-537(in Chinese)
(乔美英, 史建柯, 李冰锋, 等. 改进损失函数的增强型 FPN 水下小目标检测[J]. 计算机辅助设计与图形学学报, 2023, 35(4): 525-537)

[15] Guo C X, Fan B, Zhang Q, *et al*. AugFPN: improving multi-scale feature learning for object detection[C] //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2020: 12592-12601

[16] Wang Q L, Wu B G, Zhu P F, *et al*. ECA-net: efficient channel attention for deep convolutional neural networks[C] //Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Socie-

ty Press, 2020: 11531-11539

[17] Su Jia, Jia Xinyu, Hou Weimin. YOLO-J based PCB defect detection algorithm[J]. Computer Integrated Manufacturing Systems, 2024, 30(11): 3984-3998(in Chinese)
(苏佳, 贾欣雨, 侯卫民. 基于YOLO-J的PCB缺陷检测算法[J]. 计算机集成制造系统, 2024, 30(11): 3984-3998)

[18] Hao Yaxian, Sun Yanrui. Comparative study of loss functions for handwritten digit recognition[J]. Electronic Technology & Software Engineering, 2022(6): 203-206(in Chinese)
(郝雅娴, 孙艳蕊. 基于手写体数字识别的损失函数对比研究[J]. 电子技术与软件工程, 2022(6): 203-206)

[19] Chen Xiaocong. Improved bounding box regression loss function based on Smooth L_1[J]. College Mathematics, 2021, 37(5): 18-23(in Chinese)
(陈孝聪. 基于 Smooth L_1 改进的边框回归损失函数[J]. 大学数学, 2021, 37(5): 18-23)

[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] He K M, Zhang X Y, Ren S Q, *et al*. Deep residual learning for image recognition[C] //Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2016: 770-778

[22] Howard A, Sandler M, Chen B, *et al*. Searching for MobileNetV3[C] //Proceedings of the IEEE/CVF International Conference on Computer Vision. Los Alamitos: IEEE Computer Society Press, 2019: 1314-1324

[23] Tang Y H, Han K, Guo J Y, *et al*. GhostNetv2: enhance cheap operation with long-range attention[C] //Proceedings of the 36th International Conference on Neural Information Processing Systems. New York: Curran Associates, Inc, 2022: 9969-9982

[24] 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

[25] Chen J, Kao S, He H, *et al*. Run, Don't Walk: chasing higher FLOPS for faster neural networks[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2023: 12021-12031