

# 用于 PCB 缺陷检测的增强上下文信息 Yolov4\_tiny 算法

徐胜军<sup>1</sup>, 胡巧艳<sup>1</sup>, 韩九强<sup>1,2</sup>, 孟月波<sup>1</sup>, 刘光辉<sup>1</sup>, 赵敏华<sup>1</sup>

(1. 西安建筑科技大学 信息与控制工程学院, 西安 710055;

2. 西安交通大学 电信学部, 西安 710061)

**摘要:** 针对现有的 PCB 缺陷检测存在检测精度低、速率慢等问题, 提出一种用于 PCB 缺陷检测的增强上下文信息 Yolov4\_tiny 算法; 该算法首先通过 Transformer 编码单元对特征提取网络深层特征冗余的问题进行优化, 增强网络捕获不同尺度局部特征信息的能力; 然后利用浅层特征增强 PCB 缺陷小目标上下文信息, 提升 FPN 网络对小目标缺陷的表征能力; 最后引入注意力机制对特征提取网络输出的有效特征层加权, 强化目标特征表征能力; 实验结果表明, 该算法对于整体缺陷的平均检测精度的均值 (mAP) 达到 98.70%, 较 Yolov4\_tiny 提升了 3.12%, 实现了 PCB 缺陷精准定位和识别, 满足工业检测的实际需求。

**关键词:** PCB 缺陷检测; Yolov4\_tiny; Transformer; 上下文信息; 注意力机制

## Enhanced Context Information Yolov4\_tiny Algorithm for PCB Defect Detection

XU Shengjun<sup>1</sup>, HU Qiaoyan<sup>1</sup>, HAN Jiuqiang<sup>1,2</sup>, MENG Yuebo<sup>1</sup>, LIU Guanghui<sup>1</sup>, ZHAO Minhua<sup>1</sup>

(1. School of Information and Control Engineering, Xi'an University of Architecture and Technology, Xi'an 710055, China;

2. Department of telecommunications, Xi'an Jiaotong University, Xi'an 710061, China)

**Abstract:** Aiming at the problems of low detection accuracy and slow speed in PCB defect detection, an enhanced context information Yolov4\_tiny algorithm for PCB defect detection is proposed. Firstly, the deep feature redundancy of feature extraction network is optimized by Transformer coding unit to enhance the ability of network to capture the local feature information of the different scales. Then, the shallow features are used to enhance the small target context information of the PCB defects and improve the representation ability of the FPN network for small target defects. Finally, the attention mechanism is introduced to weight the effective feature layer of feature extraction network output to strengthen the ability of target feature representation. The experimental results show that the mean average precision (mAP) of the algorithm for overall defects is up to 98.70%, which is 3.12% higher than that of the Yolov4\_tiny, the accurate positioning and identification of PCB defects are realized, which meet the actual needs of industrial testing.

**Keywords:** PCB defect detection; Yolov4\_tiny; transformer; context information; attention mechanism

## 0 引言

随着我国现代工业技术的不断发展和智能制造体系的建立, 对电子设备的生产效率和性能都带来了新的挑战。受限于当前制作工艺水平和集成电路的复杂性, PCB 在生产过程中不可避免地出现部分印刷错误, 导致电子元件质量和性能受到很大影响。当前, PCB 缺陷检测方法主

要使用人工目视检测<sup>[1]</sup>和电学检测方法, 但是这些方法多存在速率低、成本高、难度高、有损伤等缺点, 难以达到工业生产和检测对速率、精度以及成本的实际需求<sup>[2]</sup>。

近年来, 机器视觉技术飞速发展, 基于机器学习的 PCB 缺陷检测算法相继而出。传统基于图像处理方法多使用特征匹配、相似性度量等方法先定位缺陷位置, 再进行缺陷类型的识别。Liu 等<sup>[3]</sup>基于数学形态学和模式识别的混

收稿日期: 2022-03-11; 修回日期: 2022-04-18。

基金项目: 国家自然科学基金(51678470); 陕西省自然科学基金基础研究计划(2020JM472, 2020JM473, 2019JQ760)。

作者简介: 徐胜军(1976-), 男, 陕西西安人, 博士, 副教授, 主要从事图像处理、模式识别、人工智能及自动化等方向的研究。

胡巧艳(1996-), 女, 河南南阳人, 硕士研究生, 主要从事图像处理、缺陷检测方向的研究。

韩九强(1951-), 男, 陕西咸阳人, 博士, 教授, 主要从事智能检测理论及应用、图像信息融合与模式识别方向的研究。

孟月波(1979-), 女, 陕西西安人, 博士, 副教授, 主要从事机器视觉信息处理与分析、建筑智能化方向的研究。

刘光辉(1976-), 男, 陕西西安人, 博士, 副教授, 主要从事建筑智能化、机器视觉信息处理与分析方向的研究。

赵敏华(1976-), 女, 陕西西安人, 博士, 副教授, 主要从事复杂系统建模、智能建筑与节能方向的研究。

引用格式: 徐胜军, 胡巧艳, 韩九强, 等. 用于 PCB 缺陷检测的增强上下文信息 Yolov4\_tiny 算法[J]. 计算机测量与控制, 2022, 30(9): 80-86, 92.

合识别方法提取 PCB 图像检测模板, 并引入图像畸变检测算法进行阈值分割, 得到标记的缺陷识别结果。Gaidhane 等<sup>[4]</sup>利用图像对称矩阵的秩作为缺陷检测的相似性度量, 在无需计算图像特征的情况下测量待测图像和检测模板的相似性。Li 等<sup>[5]</sup>通过二次匹配误差算法标记目标区域并研究阈值分割图得到可能的缺陷点, 最后结合掩模操作和统计直方图得到缺陷区域。虽然基于传统图像处理的检测方法实现简单、计算复杂度小, 但由于光照、角度、边界等影响导致鲁棒性较差。

深度学习网络具有表达复杂图像特征的强力性能, 通过卷积神经网络自动学习多层次特征, 较传统图像算法具有显著优势<sup>[6]</sup>, 可以运用于 PCB 缺陷的检测问题上。目前主流的深度学习检测算法分为单阶段检测算法和两阶段检测算法。两阶段检测算法首先通过区域建议网络产生候选区域, 然后对其进行分类和边界框回归。Xia 等<sup>[7]</sup>使用结构相似性指数 (SSIM) 代替形态学算子产生目标区域, 设计了一种焦点损耗的轻量级 MobileNet-V3 网络, 在保持缺陷检测精度的前提下提升了检测速度。Ding 等<sup>[8]</sup>基于 Faster R-CNN 提出微小缺陷检测网络 (TDD-Net), 通过 K-means 聚类算法生成锚框, 使用 FPN 网络加强不同特征层的上下文信息。Hu 等<sup>[9]</sup>通过具有特征金字塔网络的 ResNet50 进行特征提取, 结合 ShuffleNetV2 残差单元, 并采用 GARP 生成更精确的锚框。由于两阶段检测算法以高精度为首要追求目标, 故而运行速度较慢。单阶段检测算法在无需候选区域的情况下直接对目标进行类别和位置进行回归预测。Wei 等<sup>[10]</sup>提出利用一个语义提升模块融合不同特征层的上下文信息, 引入注意力机制和 Shuffle 模块加强学习能力和消除融合后的混叠效应。Adibhatla 等<sup>[11]</sup>引入 GoogleNet 改进 Tiny-yolov2 的特征提取网络, 强化缺陷特征表征能力, 提高缺陷的检测精度并降低了错误率。Xin 等<sup>[12]</sup>通过分析 CSPDarkNet53 网络结构的特征分布和 PCB 缺陷的尺寸分布, 微调网络参数增加锚框包含检测目标的概率。李闻等<sup>[13]</sup>使用 DBSCAN+K-means 聚类算法重新生成锚框, 并引入残差单元和 SE Block 提升了特征提取和融合能力。伍济钢等<sup>[14]</sup>使用改进二分 K-means 聚类算法结合交并比损失函数确定锚框, 引用 MobileNetV3 和 Inceptionv3 优化特征提取网络和检测网络, 优化检测速度的同时有效提高对 PCB 小目标缺陷的检测性能。和两阶段检测算法相比, 单阶段检测算法以高速度为首要追求目标, 然而其检测精度较低。

为解决上述问题, 本文提出一种用于 PCB 缺陷检测的增强上下文信息 Yolov4\_tiny 算法。主要贡献如下: 1) 通过引入 Transformer 编码单元改进特征提取网络增强不同局部信息的能力; 2) 利用浅层特征增强上下文信息, 提升小目标缺陷特征的表征能力; 3) 引入注意力机制对有效特征层加权, 强化目标特征信息。

## 1 Yolov4\_tiny 算法

Yolov4\_tiny 算法以 Yolov4<sup>[15]</sup> 算法为基础精简设计的

轻量级算法, 具有更为优异的检测速度。该算法从结构上分为特征提取网络、特征融合网络和预测网络。特征提取网络采用 CSPDarknet-tiny<sup>[16]</sup>, 输入的样本经过数据预处理后利用特征提取网络提取 PCB 图像的浅层特征和深层特征。特征融合网络采用 FPN (feature pyramid networks)<sup>[17]</sup>, 通过卷积和上采样操作自下而上融合特征提取网络输出的两个有效特征层, 加强语义特征传递。预测网络将特征融合网络得到的两个加强特征层通过卷积操作进行分类和回归得到最后的预测结果。Yolov4\_tiny 算法结构图如图 1 所示, 默认输入图像尺寸是  $608 \times 608 \times 3$ , 特征分类为 6, 此时 YOLO Head 输出尺寸为  $19 \times 19 \times 33$  和  $38 \times 38 \times 33$ , 分别对应 3 个先验框, 每个先验框占据 11 维信息, 包括 4 维坐标信息、1 维置信度信息和 6 维类别信息。

## 2 用于 PCB 缺陷检测的增强上下文信息 Yolov4\_tiny 算法

工业领域的 PCB 缺陷检测常存在目标小、特征少等问题, Yolov4\_tiny 作为单阶段检测算法虽然具有检测速度快的特点, 但由于其检测精度较低的问题不能直接应用于工业领域的 PCB 缺陷在线检测。因此提出一种用于 PCB 缺陷检测的增强上下文信息 Yolov4\_tiny 算法。首先, 该算法通过引入 Transformer 编码单元对特征提取网络深层特征冗余的问题进行优化, 增强网络捕获不同尺度 PCB 图像局部特征信息的能力; 然后, 针对 PCB 缺陷的小目标特性, 利用浅层特征增强 PCB 缺陷上下文信息, 提升对小目标缺陷的表征能力; 最后, 引入注意力机制对特征提取网络输出的有效特征层加权, 减轻干扰信息的影响, 强化目标特征信息, 使检测网络更好地适应小目标检测, 所提算法网络结构图如图 2 所示。

### 2.1 基于 Transformer 编码单元的特征提取网络

在 Yolov4\_tiny 算法中, 输入图像通过卷积单元和残差单元操作实现卷积神经网络的特征提取过程。由于轻量级网络对于小目标特征提取存在深层特征的信息冗余特点, Transformer 编码单元<sup>[18]</sup>的残差结构可以提取不同图像区域之间的特征信息, 为此提出使用 Transformer 编码单元替换特征提取网络最后的卷积单元, 不仅可以捕获全局信息和丰富的上下文信息, 从而避免深层特征的过度平滑, 并有效降低计算复杂度。

Transformer 编码单元结构图如图 3 所示。特征提取网络的第三个残差单元输出作为 Transformer 编码单元的输入, 先后经过两个残差层。第 1 残差层为多头注意力层 (MultiHead Attention), 由多个自注意力组合形成, 第 2 残差层为前向传播层 (Feed Forward), 由两个线性层和一个非线性层形成。两个残差层都分别经过求和、归一化和随机丢弃操作再进行残差连接, 防止网络退化或者过拟合。

假设输入矩阵为  $X$ , 第 1 残差层输出为  $X_1$ , 第 2 残差层输出即最终输出为  $X_2$ , 计算过程如式 (1~6) 所示。

$$Q = W_Q X, K = W_K X, V = W_V X \quad (1)$$

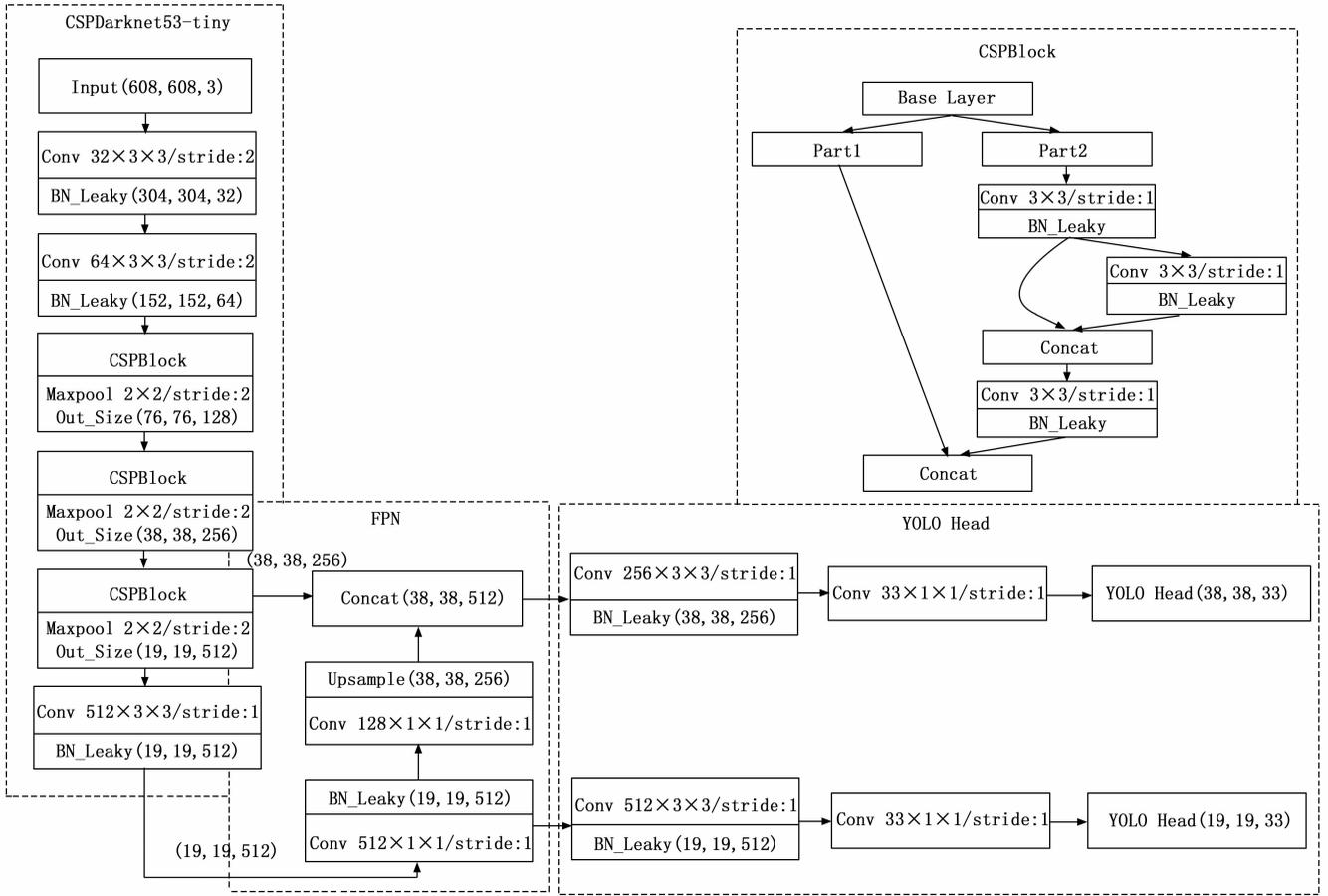


图 1 Yolov4\_tiny 算法网络结构图

$$head = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (2)$$

$$MSH = \text{concat}(head_1, \dots, head_h)W^0 \quad (3)$$

$$X_1 = \text{AND}(MSH + X) \quad (4)$$

$$MLP = \max(0, X_1W_1 + b_1)W_2 + b_2 \quad (5)$$

$$X_2 = \text{AND}(MLP + X_1) \quad (6)$$

式中， $Q$ 、 $K$ 、 $V$  分别表示输入矩阵的 query 矩阵、key 矩阵和 value 矩阵，head 表示单个自注意力的输出， $d_k$  表示  $Q$ 、 $K$  矩阵的向量维度，MSH、MLP 分别表示多头注意力输出和前向传播输出， $W_Q$ 、 $W_K$ 、 $W_V$ 、 $W^0$  分别代表输入矩阵对于 query 矩阵、key 矩阵、value 矩阵的权重矩阵和对多个自注意力进行连接和变换的权重矩阵， $W_1$ 、 $W_2$ 、 $b_1$ 、 $b_2$  表示两个全连接层的权重矩阵和偏置参数，AND 表示求和、归一化和随机丢弃操作。

在 Transformer 编码单元自注意力的计算过程中，网络的各个单元格会搜索与其他单元格的潜在关系将其关联起来，进而得出丰富的、相关的特征信息。通过两个残差层的操作，不仅挖掘了特征的代表性潜能，也增加了捕获不同局部信息的能力。

## 2.2 增强上下文信息的 FPN 特征融合网络

在目标检测网络中，被检测目标与周围目标和背景存

在丰富的上下文信息，本文通过增强上下文信息来提高特征的代表性能力。在 Yolov4\_tiny 算法中，特征提取网络将进行下采样后的  $38 \times 38$  和  $19 \times 19$  的特征层作为有效特征层进行加强特征融合。对于这种 PCB 缺陷的小目标检测问题而言，两个有效特征层都属于较深层特征，缺乏浅层特征表示，不利于目标定位，容易出现漏检或误检现象，影响检测精确度。故此本文从特征提取网络获取  $76 \times 76$  的浅层特征层，提升 FPN 网络对小目标缺陷的代表性能力，增强 PCB 缺陷小目标上下文信息。改进后的算法如图 2 所示，特征融合网络输出的  $38 \times 38$  加强特征层继续进行卷积和上采样操作得到  $76 \times 76$  的特征层与特征提取网络输出的  $76 \times 76$  有效特征层融合得到具有浅层特征的加强特征层。为了减少参数数量和计算量，本文去除了特征融合网络输出的适用于大目标的  $19 \times 19$  检测层。改进后的 FPN 网络特征融合层可以获得加强的浅层输出，从而增强网络对小目标缺陷的代表性能力。

## 2.3 通道空间注意力模块

在特征提取过程中，通道维度和空间维度不可避免会出现一些干扰信息，导致网络性能不佳。为了在特征融合时能高效利用特征提取网络提取到的特征信息，本文对特征提取网络输出的 3 个有效特征层分别添加一个注意力机制

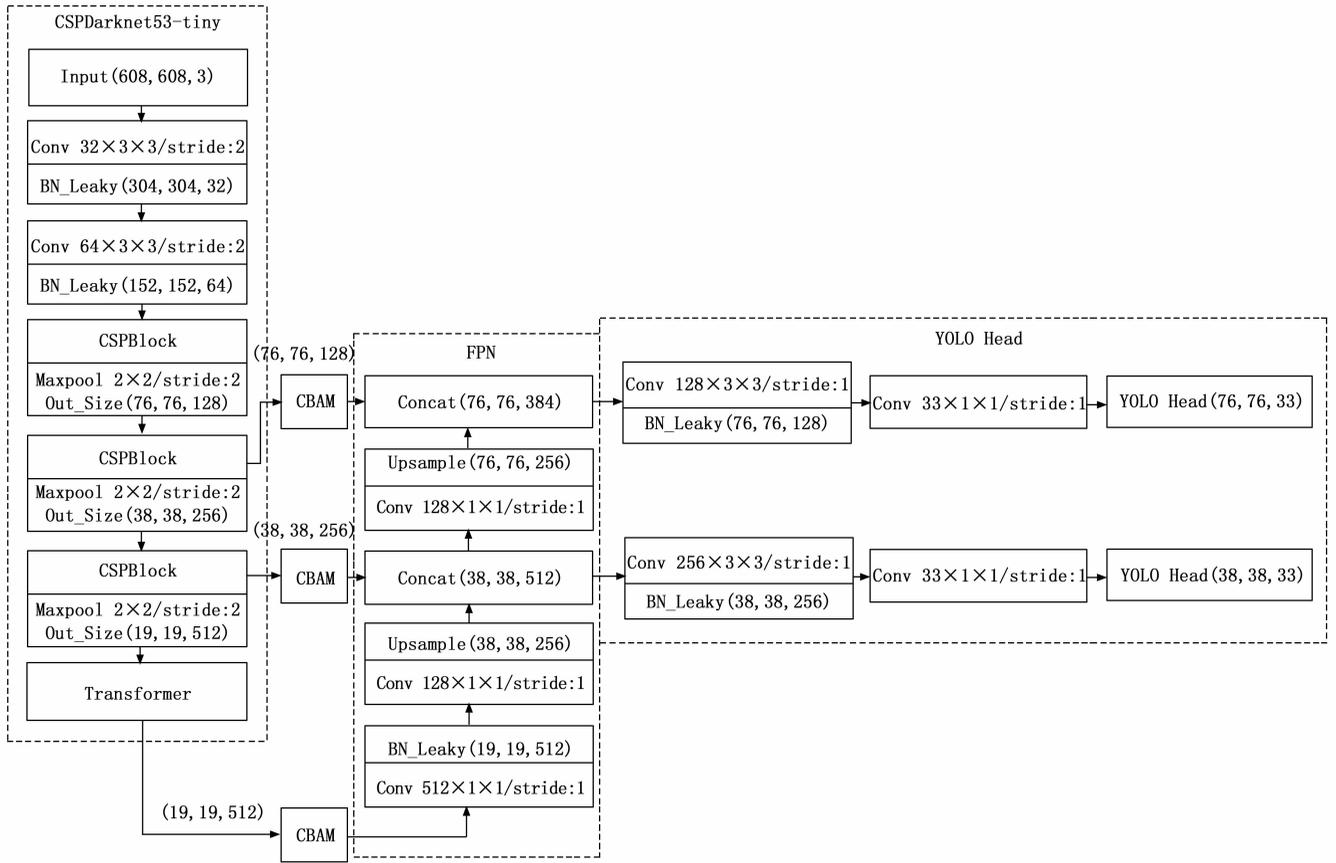


图 2 所提算法网络结构图

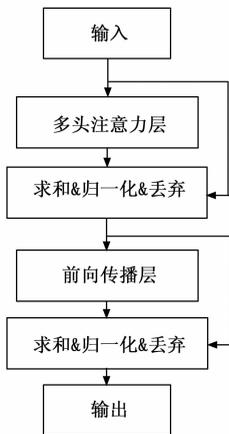


图 3 Transformer 编码单元结构图

CBAM 模块 (convolutional block attention module)<sup>[19]</sup>, 分别从通道维度和空间维度依次推断注意力映射, 并与输入特征层相乘以进行自适应特征优化, 减轻干扰信息的影响, 使模型捕获更多关键的特征信息。

以图 2 中的  $76 \times 76 \times 128$  尺度的特征层为例, 记作  $F^{76 \times 76 \times 128}$ 。CBAM 注意力模块先对输入特征层  $F^{76 \times 76 \times 128}$  从通道维度推理出一维的通道注意力映射  $M_c^{1 \times 1 \times 128}$ , 再将输入特征层  $F$  与通道注意力映射  $M_c^{1 \times 1 \times 128}$  相乘得到  $F_1$ , 之后对  $F_1$  从空间维度推理出二维的空间注意力映射  $M_s^{76 \times 76 \times 1}$ , 最后将  $F_1$  与空间注意力映射  $M_s^{76 \times 76 \times 1}$  相乘得到最终的输出, 示意图

如图 4 所示。

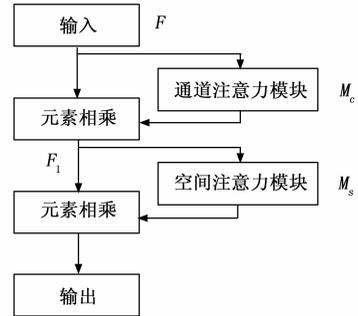


图 4 CBAM 模块结构图

其中, 通道注意力模块强调特征图的目标信息, 从通道维度对目标信息进行加权处理, 计算过程如式 (7) 所示。先使用最大池化和平均池化从空间维度对特征图  $F^{76 \times 76 \times 128}$  进行压缩, 得到两个  $1 \times 1 \times 128$  的特征描述, 再使用由 MLP (multi-layer perceptron) 组成的共享网络对这两个不同的特征描述进行计算并进行激活得到  $M_c(F)$ 。

$$M_c(F) = \sigma(MLP(AP(F)) + MLP(MP(F))) \quad (7)$$

$$M_s(F) = \sigma(f^{7 \times 7}([AP(F); MP(F)])) \quad (8)$$

式中,  $AP$ 、 $MP$  分别表示平均池化操作、最大池化操作,  $f^{7 \times 7}$  表示核大小为  $7 \times 7$  的卷积操作,  $\sigma$  表示 Sigmoid 激活函数。

空间注意力模块强调特征图的位置信息, 从空间维度对于目标信息进行加权处理, 计算过程如式 (8) 所示。先

使用最大池化和平均池化从通道维度对特征图  $F^{76 \times 76 \times 128}$  进行压缩, 得到两个  $76 \times 76 \times 1$  的特征描述, 使用 concat 将两个特征描述合并, 进一步使用  $7 \times 7$  的卷积核进行卷积操作并进行激活生成  $M_s(F)$ 。

### 2.4 损失函数

YOLO Head 输出的特征层通过解码得到预测框输出参数即坐标参数、置信度参数和类别参数, 通过学习三项参数逐步确定检测目标的坐标和类别, 置信度参数用于删除低于置信度阈值的预测框。因此损失函数主要由坐标误差  $L_{loc}$ 、置信度误差  $L_{conf}$  和类别误差  $L_{class}$  组成。

$$Loss = L_{loc} + L_{conf} + L_{class} \quad (9)$$

坐标误差采用 CIOU (Complete-IoU) 损失函数, 综合考虑预测框与真实框的中心距离、重合度以及宽高比, 计算方式如公式 (10) 所示。

$$L_{loc} = 1 - IOU + \frac{\rho^2(b, b^{gt})}{c^2} + \frac{16}{\pi^2} \frac{(\arctan \frac{w^{gt}}{h^{gt}} - \arctan \frac{w}{h})^2}{1 - IOU + \frac{4}{\pi^2} (\arctan \frac{w^{gt}}{h^{gt}} - \arctan \frac{w}{h})^2} \quad (10)$$

式中,  $IOU$  表示预测框和真实框的交并比,  $w^{gt}$ 、 $h^{gt}$ 、 $w$ 、 $h$  分别表示真实框和预测框的宽度和高度,  $\rho^2(b, b^{gt})$  表示真实框和预测框的中心点的欧式距离,  $c$  表示包含预测框和真实框的最小框的对角线距离。

置信度误差采用交叉熵损失函数, 计算方式如公式 (11) 所示。

$$L_{conf} = \sum_{i=0}^{K \times K} \sum_{j=0}^M I_{ij}^{obj} [\bar{C}_i \log(C_i) + (1 - \bar{C}_i) \log(1 - C_i)] - \lambda_{noobj} \sum_{i=0}^{K \times K} \sum_{j=0}^M (1 - I_{ij}^{obj}) [\bar{C}_i \log(C_i) + (1 - \bar{C}_i) \log(1 - C_i)] \quad (11)$$

式中,  $K$  表示特征图的网格尺寸,  $M$  表示每个网格的预测个数,  $I_{ij}^{obj}$  表示第  $i$  个网格的第  $j$  个预测框是否负责检测该目标 (1 或 0),  $\bar{C}_i$ 、 $C_i$  分别表示置信度的真实值和预测值,  $\lambda_{noobj}$  表示不包含缺陷的置信度损失权重系数。

类别误差采用交叉熵损失函数, 计算方式如公式 (12) 所示。

$$L_{class} = \sum_{i=0}^{K \times K} I_{ij}^{obj} \sum_{c \in classes} [\bar{p}_i(c) \log(p_i(c)) + (1 - \bar{p}_i(c)) \log(1 - p_i(c))] \quad (12)$$

式中,  $c$  表示检测到的目标所属类别,  $\bar{p}_i(c)$ 、 $p_i(c)$  分别表示网格  $i$  所预测的缺陷类别概率和真实的类别概率。

## 3 实验与分析

### 3.1 实验环境与数据集

深度学习实验平台系统配置为 Ubuntu 18.04, CPU 为 Inter (R) Core (TM) i9-9900CPU @2.30 GHz, GPU 为 NVIDIA GeForce GTX 2080Ti, python 版本为 3.7.2, pytorch 版本为 1.2.0, Cuda 版本为 10.1。

本实验采用北京大学智能机器人开放实验室的公开

PCB 缺陷数据集<sup>[8]</sup>进行网络的训练和测试。该数据集包含 10 668 张 PCB 缺陷图像及其对应的标注文件。图像缺陷类型分别为漏焊、缺口、开路、短路、毛刺和余铜, 每张图像包含 1 到 2 类缺陷。整个数据集按照 8 : 2 的比例随机划分为训练集和测试集。

### 3.2 训练参数设置与评价指标

对比实验基于迁移学习的思想, 使用 Yolov4\_tiny 算法在 COCO 数据集训练过的模型作为预训练模型, 训练分为冻结阶段和解冻阶段, 设置冻结阶段的迭代次数为 50, 批处理大小为 48, 学习率为 0.001, 解冻阶段的迭代次数为 50, 批处理大小为 24, 学习率为 0.000 1。为了使得先验框尺寸与缺陷尺寸更加匹配, 利用 K-means 聚类算法对先验框重新聚类, 得到适合  $76 \times 76$  特征层的先验框尺寸为 [13, 12]、[13, 17]、[16, 17], 适合  $38 \times 38$  特征层的先验框尺寸为 [21, 22]、[27, 14]、[31, 28]。

为了评估算法的检测性能, 本文采用准确率 (Precision)、召回率 (Recall)、均值平均精度 (mAP)、检测时间 (Time) 以及模型参数规模 (Size) 作为评价指标对算法进行测试和分析。Precision、Recall 和 mAP 各类指标的计算公式如下。

$$Precision = \frac{TP}{TP + FP} \quad (13)$$

$$Recall = \frac{TP}{TP + FN} \quad (14)$$

$$AP = \int_0^1 P(R) dR \quad (15)$$

$$mAP = \frac{\sum_{i=1}^c AP_i}{c} \quad (16)$$

式中,  $TP$ 、 $FP$ 、 $FN$  分别为正样本且预测正确数、正样本且预测错误数、负样本且预测错误数。其中预测值为正样本指 IOU 大于 0.5 的预测值的分类结果。

### 3.3 消融实验结果与分析

为了分析 Transformer 编码单元、改进的 FPN 网络以及 CBAM 注意力模块分别对所提算法的提升效果, 以 Yolov4\_tiny 为基础架构针对本文所提出的改进方案进行消融实验。各消融算法的相关参数设置和训练策略均相同, 测试结果如表 1 所示。

表 1 消融性能对比实验 %

算法	Precision	Recall	mAP
Yolov4_tiny	94.49	93.14	95.58
Yolov4_tiny + Transformer	95.24	94.50	96.67
Yolov4_tiny + 改进 FPN	96.49	94.39	97.20
Yolov4_tiny + CBAM	95.01	93.80	96.20
本文算法	97.22	97.20	98.70

实验结果表明, 以 Yolov4\_tiny 为基础在特征提取网络引入 Transformer 编码单元对数据集进行训练, 较 Yolov4\_tiny 算法 Precision 提升 0.75%, Recall 提升

1.36%, mAP 提升 1.09%; 以 Yolov4\_tiny 为基础增加浅层特征映射来改进 FPN 特征融合网络对数据集进行训练, 较 Yolov4\_tiny 算法 Precision 提升 2.00%, Recall 提升 1.25%, mAP 提升 1.62%; 以 Yolov4\_tiny 为基础在特征融合网络前引入 CBAM 注意力模块对数据集进行训练, 较 Yolov4\_tiny 算法 Precision 提升 0.52%, Recall 提升 0.66%, mAP 提升 0.62%; 以 Yolov4\_tiny 为基础加入 3 个模块后, 本文算法达到了最佳检测性能, 较 Yolov4\_tiny 算法 Precision 提升 2.73%, Recall 提升 4.06%, mAP 提升 3.12%。因此, 整体改进后的算法检测精度均优于单个改进后的算法。综上, 通过消融实验可以证明本文所提出的改进方案对于 Yolov4\_tiny 算法在 PCB 缺陷检测任务中的检测性能有所提升。

图 5 显示了原始 Yolov4\_tiny 算法和本文算法在训练期间的损失曲线。无论是前半段的冻结阶段还是后半段的解冻阶段, 本文算法的训练集损失和测试集损失曲线都比原始 Yolov4\_tiny 算法具有加速收敛的趋势, 且损失值更低。

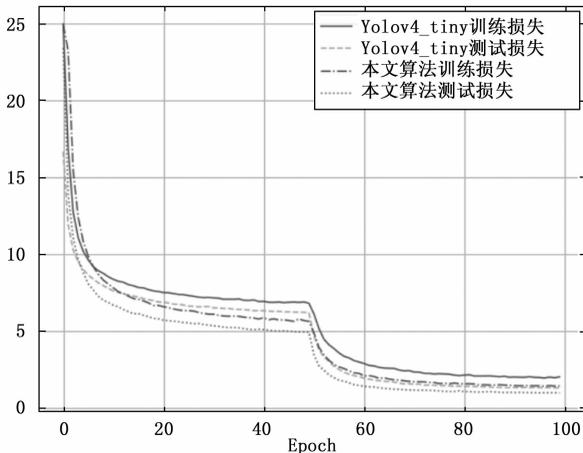


图 5 Yolov4\_tiny 算法和本文算法检测模型的损失曲线

### 3.4 不同检测算法性能结果与分析

为了进一步验证本文算法的优越性, 选取 3 种主流目标检测算法同本文算法进行对比, 对比算法包括以 Resnet50 为主干网络的两阶段目标检测算法 Faster RCNN<sup>[20]</sup>和以 VGG16 为主干网络的单阶段目标检测算法 SSD<sup>[21]</sup>、Yolov3 目标检测算法<sup>[22]</sup>、用于 PCB 缺陷检测的改进版 Yolov3 算法<sup>[13]</sup>、轻量级目标检测算法 Yolov4\_tiny。针对不同算法的 Precision、Recall、mAP、Time、Size 评价指标的对比结果如表 2 所示。

从衡量算法精确率 Precision、查全率 Recall 来看, Faster RCNN 算法使用候选框机制遍历目标时能够较为全面的检测到目标, 但因此也存在无缺陷区域, 导致 Recall 指标表现优异, Precision 指标较低; SSD 算法网络结构拥有多尺度分类器, 能较好地兼顾提取不同尺寸目标, 又因为网络深度较低, 特征提取不够充分, 导致 Precision 指标表现优异, Recall 指标较低; Yolov3 算法和改进版 Yolov3 算法通过特征网络提取深层特征与浅层特征并进行加强融

合, 但对部分缺陷仍存在检测难度, Precision 和 Recall 指标略显不足; 本文算法在 Yolov4\_tiny 算法的基础上增加特征提取网络输出的有效特征层, 将浅层特征加权后传递到特征融合模块中, 保留了小目标的特征信息, 两类指标分别提升 2.73%、4.06%。

表 2 不同检测算法性能对比实验

算法	Precision/%	Recall/%	mAP/%	Time/ms	Size/MB
Faster RCNN	36.28	88.95	61.14	565.85	108.84
SSD	97.20	18.38	89.20	90.73	93.14
Yolov3	76.21	73.19	74.70	136.00	263.50
改进 Yolov3	88.96	85.63	87.60	139.00	263.55
Yolov4_tiny	94.49	93.14	95.58	51.72	22.50
本文算法	97.22	97.20	98.70	81.78	18.57

从衡量算法整体性能的 mAP 来看, Faster RCNN 算法、SSD 算法、Yolov3 算法、改进 Yolov3 算法在 PCB 缺陷检测过程中错误较多, mAP 指标处于 90% 以下; Yolov4\_tiny 算法表现较好, 达到了 95.58%; 本文算法相对于 Yolov4\_tiny 算法提升了 3.12%, 达到了最优的性能指标。

从衡量算法部署应用可行性的 Time 和模型参数规模 Size 来看, Faster RCNN 算法、Yolov3 算法、改进版 Yolov3 算法的检测时间较大, 难以保证实时性; SSD 算法、Yolov4\_tiny 算法和本文算法的检测时间分别是 90.73 ms、51.72 ms 和 81.78 ms, 可以较好的保证在线检测实时性的要求, 但本文算法较 Faster RCNN 算法、SSD 算法、Yolov3 算法、改进版 Yolov3 算法和 Yolov4\_tiny 算法的模型参数规模分别减少 90.27 MB、74.57 MB、244.93 MB、244.98 MB 和 3.93 MB。综上所述, 考虑到定量评价的检测精度、检测速度以及模型参数规模等指标, 本文算法优于其它算法。

### 3.5 不同检测算法检测结果与分析

为直观评价本文算法在 PCB 缺陷数据集中检测效果的优越性, 分别采用 Faster RCNN、SSD、Yolov4\_tiny 和本文所提算法进行对比实验, 在测试结果中选择 5 种具有代表性的测试图对比。图 6 所示为测试原图。为更直观进行缺陷检测结果对比, 图 7 中仅放大展示图 6 方框部分的测试结果, 自上向下分别为毛刺 (spur)、缺口 (mouse\_bite)、短路 (short)、余铜 (spurious\_copper) 和漏焊 (missing\_hole) 缺陷检测, 自左向右分别为 Faster RCNN、SSD、Yolov4\_tiny 和本文算法的检测结果以及实际标注结果。

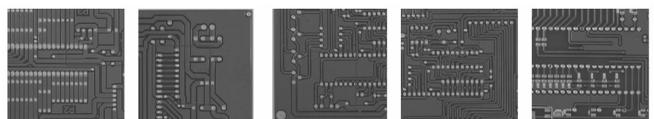


图 6 测试原图

由图中对比结果可以看出, Faster RCNN 算法的候选框机制虽然可以将大部分缺陷检测出来, 但是对于同一个目标存在较为严重的重叠现象, 如图 7 的第 1、2、4、5 行

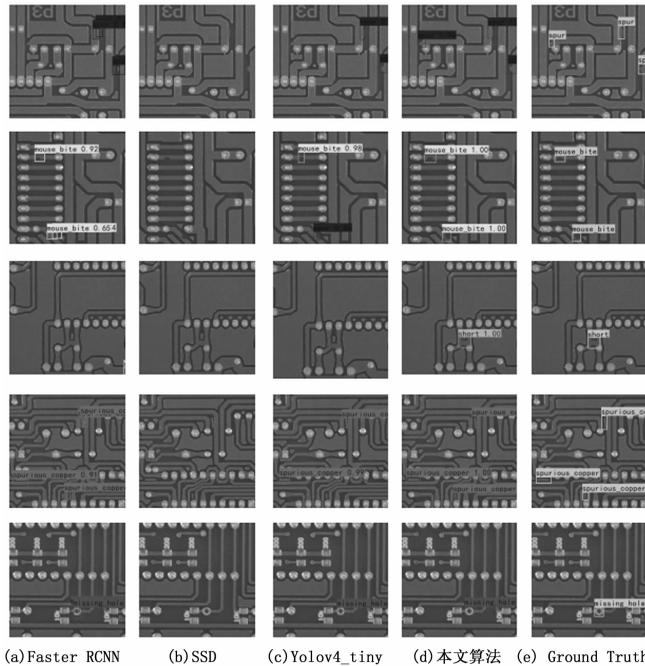


图 7 检测结果对比

所示, 同一个目标出现多个检测框, 且第 3 行出现漏检、多检现象; SSD 算法由于网络深度不足对特征的提取效果不佳, 存在大量的漏检现象, 对于图 7 中的检测目标均未检测出; YOLOv4\_tiny 算法对于毛刺、缺口、余铜和漏焊这类尺寸较小且不易检测的目标存在一定程度上的漏检现象, 如图 7 的第 1、2、3、4 行所示, 且第 2 行中毛刺和缺口两类目标识别不清; 本文算法可以将所有目标都正确地检测到, 且检测结果得分最高, 具有更加优异的检测效果。综上所述, 同其它对比算法相比, 本文所提算法的检测精度和检测速度表现最为优异。

#### 4 结束语

针对 PCB 缺陷检测问题, 提出了一种用于 PCB 缺陷检测的增强上下文信息 YOLOv4\_tiny 算法, 能够在保证实时性的基础上对 PCB 缺陷进行精准定位和分类。首先, 本文选用高实时性的 YOLOv4\_tiny 算法作为基本框架, 确保了缺陷检测的准确性和快速性; 其次引入 Transformer 编码单元来改进特征提取网络, 增强了不同局部信息的能力, 并利用注意力机制对特征提取网络的有效特征层加权, 强化了目标特征信息。考虑到 PCB 缺陷目标尺寸小、特征信息少的问题, 本文增加特征提取网络输出的有效特征层, 利用浅层特征增强 PCB 缺陷小目标上下文信息, 并去掉了冗余的检测层。实验结果表明本文算法可以满足工业上 PCB 缺陷检测的实际需求。

#### 参考文献:

- [1] 汤勃, 孔建益, 伍世虔. 机器视觉表面缺陷检测综述 [J]. 中国图象图形学报, 2017, 22 (12): 1640-1663.
- [2] 黄亚伟. PCB 裸板缺陷光学检测技术的研究与应用 [D]. 南

京: 南京理工大学, 2018.

- [3] LIU Z, QU B. Machine vision based Online detection of PCB Defect [J]. Microprocessors and Microsystems, 2021, 82 (9): 103807.
- [4] GAIDHANE V H, HOTE Y V, SINGH V. An efficient similarity measure approach for PCB surface defect detection [J]. Pattern Analysis and Applications, 2018, 21 (1): 277-289.
- [5] LI R, XUE B, ZHAO K, et al. PCB Defect Recognition and Elimination Based on Secondary Error and Statistical Histogram [C] //2020 International Wireless Communications and Mobile Computing (IWCMC). IEEE, 2020: 1981-1984.
- [6] 李红光, 于若男, 丁文锐. 基于深度学习的小目标检测研究进展 [J]. 航空学报, 2021, 42 (7): 107-125.
- [7] XIA B, CAO J, WANG C. SSIM-NET: Real-time PCB Defect Detection Based on SSIM and MobileNet-V3 [C] //2019 2nd World Conference on Mechanical Engineering and Intelligent Manufacturing (WCMEIM). November 22-24, 2019, Shanghai, China; IEEE, 2019: 756-759.
- [8] DING R, DAI L, LI G, et al. TDD-net: a tiny defect detection network for printed circuit boards [J]. CAAI Transactions on Intelligence Technology, 2019, 4 (2): 110-116.
- [9] HU B, WANG J. Detection of PCB surface defects with improved faster-RCNN and feature pyramid network [J]. IEEE Access, 2020, 8: 108335-108345.
- [10] SHI W, LU Z, WU W, et al. Single-shot detector with enriched semantics for PCB tiny defect detection [J]. The Journal of Engineering, 2020, 2020 (13): 366-372.
- [11] ADIBHATLA V A, CHIH H C, HSU C C, et al. Defect detection in printed circuit boards using you-only-look-once convolutional neural networks [J]. Electronics, 2020, 9 (9): 1547.
- [12] XIN H, CHEN Z, Wang B. PCB Electronic Component Defect Detection Method based on Improved YOLOv4 Algorithm [J]. Journal of Physics: Conference Series, 2021, 1827 (1): 012167.
- [13] 李闻, 李小春, 闫昊雷. 基于改进 YOLO v3 的 PCB 缺陷检测 [J]. 光电与控制, 2022, 29 (4): 106-111.
- [14] 伍济钢, 成远, 邵俊, 等. 基于改进 YOLOv4 算法的 PCB 缺陷检测研究 [J]. 仪器仪表学报, 2021, 42 (10): 171-178.
- [15] BOCHKOVSKIY A, WANG C Y, LIAO H. YOLOv4: Optimal Speed and Accuracy of Object Detection [EB/OL]. [2020-06-05]. <http://arxiv.org/abs/2004.10934>.
- [16] WANG C Y, LIAO H Y M, WU Y H, et al. CSPNet: A new backbone that can enhance learning capability of CNN [C] // Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops. June 14-19, Seattle, WA, USA; IEEE, 2020: 390-391.
- [17] LIN T Y, DOLLAR P, GIRSHICK R, et al. Feature pyramid networks for object detection [C] //2017 IEEE Conference on Computer Vision and Pattern Recognition. July 21-26, 2017, Honolulu, HI, USA. New York; IEEE Press, 2017: 936-944.

(下转第 92 页)