

DOI: 10.19650/j.cnki.cjsi.J2312229

ViT 和注意力融合类别不均衡 PCB 缺陷检测方法*

陈俊英, 李朝阳, 席月芸, 刘 冲

(西安建筑科技大学信息与控制工程学院 西安 710055)

摘要:针对实际环境下印刷电路板(PCB)缺陷样本难以收集造成的数据长尾分布和检测精度低以及 ViT 用于检测时计算复杂度高等问题,提出多尺度 ViT 特征提取和注意力特征融合的端到端 PCB 缺陷检测算法。首先结合 ViT 和部分卷积构建多尺度特征提取网络,利用层级多头注意力对不同尺度的特征图执行自适应的注意力操作,使网络能够更好地捕捉局部和全局信息进而增强其特征提取能力,部分卷积可以降低计算开销。其次,基于能量空域抑制的无参数注意力机制将多尺度特征有效融合,提升网络融合特征图的表达能力。最后,引入对类别不均衡敏感的分类函数对网络的损失函数进行改进,增强网络对类别不平衡数据的拟合程度,提高网络的泛化能力。在 3 种不同类型的公开 PCB 数据集上的实验结果表明,所提出的检测算法在 PCB 表面缺陷数据集的平均精度均值(mAP)均有提升,分别为 99.13%、98.67%、99.82%;在类别不均衡的 PCB 缺陷检测任务上,相较于改进前方法,mAP 提升了 11.94%,网络检测速度达到 25 FPS,为 PCB 缺陷的检测提供了一种快速、有效的方法。

关键词: 缺陷检测;印刷电路板;Vision Transformer;注意力机制;多尺度特征提取

中图分类号: TP391 TH862 **文献标识码:** A **国家标准学科分类代码:** 510.40

ViT and attention fusion for class-imbalanced PCB defect detection

Chen Junying, Li Zhaoyang, Xi Yueyun, Liu Chong

(College of Information and Control Engineering, Xi'an University of Architecture and Technology, Xi'an 710055, China)

Abstract: Addressing the challenges of a long-tailed distribution of data and low detection accuracy caused by the difficulty in collecting defect samples for printed circuit boards (PCBs) in real-world environments, as well as the high computational complexity when using Vision Transformer (ViT) for detection, we propose an end-to-end PCB defect detection algorithm that incorporates multi-scale ViT feature extraction and attention feature fusion. Firstly, a multi-scale feature extraction network is constructed by combining ViT and partial convolution. Hierarchical multi-head attention is employed to perform adaptive attention operations on different scales of feature maps, enabling the network to better capture local and global information, thereby enhancing its feature extraction capabilities. Partial convolution is utilized to reduce computational costs. Secondly, a non-parametric attention mechanism based on the energy domain suppression effectively fuses multi-scale features, enhancing the expressive power of the network's fused feature maps. Finally, a classification function sensitive to class imbalance is introduced to improve the loss function of the network, enhancing its fitting ability to imbalanced data and improving generalization. The experimental results on three different types of publicly available PCB datasets indicate that the proposed detection algorithm shows improvement in the mean Average Precision (mAP) for PCB surface defect datasets, with respective values of 99.13%, 98.67%, and 99.82%. In the case of class-imbalanced PCB defect detection tasks, the mAP is improved by 11.94% compared to the previous method, and the network achieves a detection speed of 25 FPS, providing a fast and effective approach for PCB defect detection.

Keywords: defect detection; printed circuit boards; vision transformer; attention mechanism; multi-scale feature extraction

0 引言

印刷电路板(printed circuit board, PCB)在集成电路中被广泛使用。作为承载精密电子元器件的基板,PCB的质量与电子设备质量息息相关。在实际生产过程中,由于PCB生产工艺复杂,其表面极易产生缺陷^[1]。传统的PCB表面缺陷检测主要依靠人工目视法来完成。近年来,针对PCB表面存在的缺陷类型,逐渐出现了功能测试、视觉测试等方法^[2]。其中,人工目视法检测速度慢,需要花费大量的人力物力;功能测试法需要配备专门的测试设备,编程复杂,难以推广应用;视觉测试法对采集的PCB表面图像进行自动光学检测分析,主要包括自动视觉检测技术、自动光学检测技术^[3]和机器视觉检测技术^[4]等。相较于传统的人工检测方法,视觉检测法具有效率高、成本低等特点。随着计算机技术和机器学习技术的迅猛发展,基于深度学习的缺陷检测算法在各种表面缺陷检测场景中得到了广泛应用^[5]。其中大多数算法利用卷积神经网络(convolutional neural networks, CNN)提取图像特征,并以此识别和检测图像中的缺陷^[6]。目前,这种基于深度学习的方法已经成为主流的缺陷检测方法,为缺陷检测任务带来了显著的性能提升。

基于深度学习的缺陷检测方法主要分为有监督学习和无监督学习两类。无监督学习通过发现输入样本中的异常特征来判断样本是否存在缺陷。而在有监督学习中,网络利用已标记的缺陷位置、类别等信息进行训练,以便准确地识别图像中的缺陷位置、类别和缺陷大小^[7]。有监督学习利用已标记样本进行训练通常能取得比无监督学习更高的准确率。目前,基于有监督学习的缺陷目标检测算法主要有以YOLO^[8]、SSD^[9]等为代表的一阶段目标检测算法和以R-CNN^[10]、Faster R-CNN^[11]等为代表的两阶段目标检测算法。陈仁祥等^[12]针对实际环境中噪声干扰使得PCB表面缺陷检测精度低的问题,提出一种基于多注意力的Faster R-CNN印刷电路板表面缺陷检测方法,在特征提取与融合部分分别引入注意力机制获取具有抗噪声干扰的特征表示,获得了较好的检测效果。Luo等^[13]针对柔性PCB表面缺陷的外观相似性提出了解耦两阶段目标检测框架,通过两个子网络增强特征提取能力,并通过嵌入的自注意力机制增加模型对缺陷目标的分类能力,实验结果显示具有较高的检测精度。在一阶段目标检测算法方面,Adibhatla等^[14]提出基于YOLOv5的PCB缺陷检测,减少PCB质检过程中的人力成本和时间成本。这些方法均以CNN为基本操作构建特征提取网络。

近年来,视觉变换(vision transformer, ViT)已在计算机视觉任务中取得了显著的突破^[15]。ViT利用自注意力

机制,能够有效地捕捉图像中不同区域之间的关系,从而实现了强大的全局建模能力。相较于传统的CNN,ViT更适合处理复杂的视觉任务,尤其是当数据规模较大时,其性能更优^[16]。然而,相较于CNN,ViT的自注意力计算复杂度与图像的大小呈平方级增长^[17],这使得它在处理高分辨率图像时计算复杂且对数据要求较高^[18]。为了应对这些问题,Swin-Transformer^[19]引入基于窗口机制的注意力,实现了对输入图像的局部化处理,从而有效降低了Transformer的计算复杂度。此外,Guo等^[20]提出了卷积神经网络迎接视觉变形(convolutional neural networks meet vision transformers, CMT)方法,该方法借助残差结构来减少Transformer的参数量,进一步提高了计算效率。这些改进方法有助于克服原始Transformer模型的局限性,使其更加适应计算机视觉等领域的特定需求。

在视觉缺陷检测任务中,缺陷样本的获取相对困难,这导致可用于训练的数据集规模通常较小,且常呈现出长尾分布的特点;同时视觉缺陷检测任务通常要求模型能够实时处理图像数据。针对上述问题,本文借助ViT的强大特征建模能力和注意力机制的特征融合能力提出了一种改进的PCB缺陷检测方法。首先,提出一种多尺度特征提取网络,该网络融合了ViT和部分卷积的建模能力,增强网络对全局和局部特征的提取。同时,为了提高PCB表面缺陷的检测精度,引入了基于能量空域抑制的无参数注意力机制构建特征融合网络,增强网络的特征分辨能力,提高网络的检测精度。此外,引入了用于处理类别不平衡的分类损失函数,以此构建的综合损失函数有助于提高网络在不同类别上的训练精度,有效应对在实际缺陷检测应用中常见的数据不平衡问题。实验结果验证了所提出的算法在实际PCB缺陷检测任务中具有较高的检测精度和效率,是一种快速、有效的PCB缺陷检测方法。

1 ViT和注意力融合的PCB缺陷检测方法

1.1 缺陷检测网络结构

所提出的PCB缺陷检测算法是基于YOLOX^[21]框架进行改进的单阶段目标检测算法,使用改进ViT多尺度网络进行特征提取,结合SimAM改进颈部融合网络,通过解耦头分类回归子网络进行缺陷目标的位置回归与分类,整体网络框架如下图1所示。改进的ViT主干网络(ViTBackbone)对输入图像进行多尺度特征提取,然后基于特征融合(feature fusion)模块对P3、P4、P5尺度特征进行融合,融合后的输出特征图为P3_OUT、P4_OUT和P5_OUT,随后分别通过解耦头(decoupled head)预测结构得到预测目标的位置回归(Reg)、置信度输出(IoU)和分类输出(Cls)。

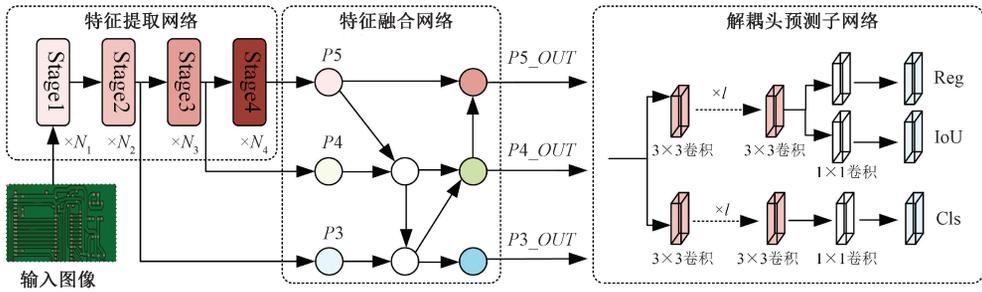


图1 PCB表面缺陷检测网络结构

Fig. 1 Network architecture for PCB surface defect detection

1.2 基于 ViT 和部分卷积的多尺度特征提取网络

FasterViT^[22]是 NVIDIA 结合 ViT 的全局特征建模和 CNN 的局部特征建模能力所提出的面向硬件优化的高效神经网络骨干架构,原始的 FasterViT 采用卷积模块 (Conv

Block) 来实现局部特征提取。为了降低模型在缺陷检测过程中的计算复杂度,本文结合部分卷积^[23] (partial convolution, PConv) 和 FasterViT 提出了一种通用的多尺度特征提取网络 (MulPViT), 具体结构如图 2 所示。

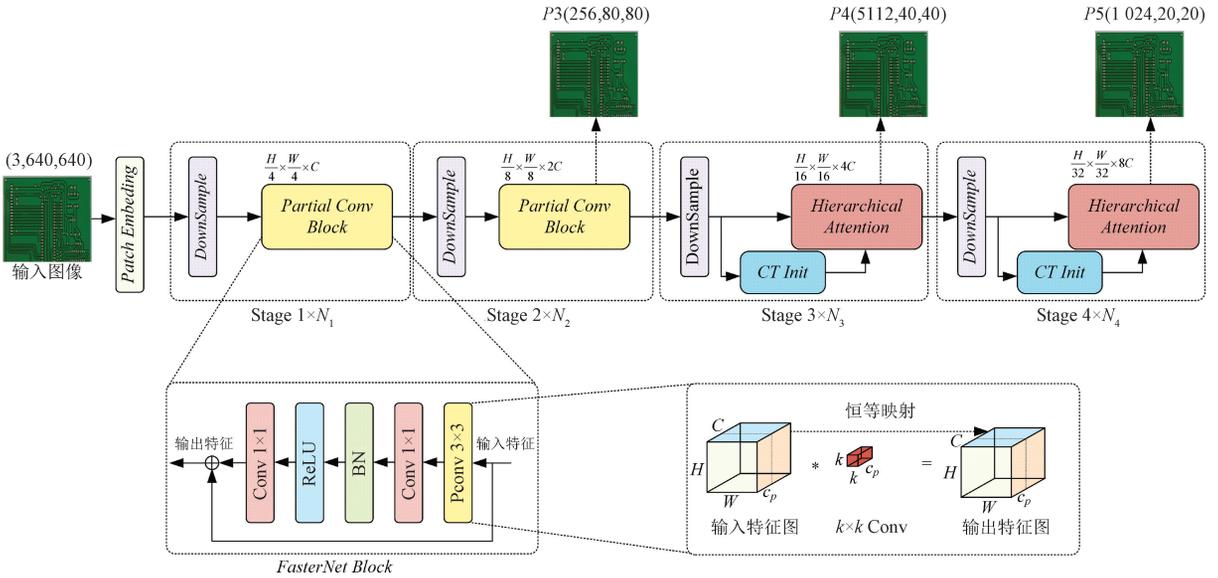


图2 多尺度特征提取网络结构示意图

Fig. 2 Schematic diagram of the multi-scale feature extraction network structure

MulPViT 结合了部分卷积神经网络层与 ViT 的注意力层,旨在实现更高效的特征提取并减小网络模型的规模,从而减少参数量,提高计算效率。为了更好地处理不同大小的目标物体, MulPViT 通过 4 个不同的阶段 (Stage) 提取多尺度特征。其中,将 Stage2、Stage3 和 Stage4 的输出特征图作为主干特征提取网络的多尺度特征。

在 Stage1 和 Stage2 中,主要采用 Partial Conv Block 进行特征的提取,每个 Partial Conv Block 包括两个 1×1 卷积层、一个 ReLU 激活函数、一个 Batch Norm 归一化层和一个 PConv 层,并通过残差连接进行输入和输出的叠加传送。针对卷积神经网络提取的特征图在通道维度上

存在冗余的问题^[24],本文引入了部分卷积来减小计算成本。PConv 的工作原理如图 2 右下角所示。在计算过程中,只对输入特征图的一部分通道进行常规卷积操作,而其余通道直接进行恒等连接,这种操作能进一步优化计算复杂度和内存访问,同时保留特征图的有效信息。假设输入特征图的高度和宽度分别为 H 和 W , 通道数为 c , 将输入特征图连续的 c_p 个通道看作输入特征图的代表, 对前 c_p 个通道采用步长为 1、填充为 $(kernel_size-1)/2$ 的卷积进行特征提取, PConv 的计算复杂度 (FLOPs) 如式(1)所示。

$$FLOPs = H \times W \times k^2 \times c_p^2 \tag{1}$$

定义部分比率 (Partial Ratio) $r = c_p/c$, 当 $r < 1$ 时,

PConv 的 FLOPs 只有常规卷积的 r^2 倍。在内存访问方面, PConv 的内存访问量也会减少, 为常规卷积的 r 倍, 内存访问量 Memory_PConv 可通过下式(2)进行估算:

$$\text{Memory_PConv} = H \times W \times 2c_p + k^2 \times c_p^2 \approx H \times W \times 2c_p \quad (2)$$

在 Stage3 和 Stage4 中, 借鉴了 FasterViT 的设计思路, 采用基于层级注意力(hierarchical attention)的 ViT 结构进行特征提取。图 3 展现了层级注意力的工作原理, 通过引入全局载体令牌(carrier tokens, CTs), 以实现全局信息的汇总和传递; 通过将局部窗口的令牌(local token)与全局 CTs 进行连接来建立局部和全局信息的互动。局部窗口(local window)仅能访问与其对应的 CTs, 从而降低信息交换的计算开销, 提高 ViT 网络在实际工业环境下的缺陷检测效率。通过交替使用全局令牌和局部窗口令牌的策略构建 ViT 的分层注意力网络。这样模型可以同时关注全局信息(通过 CTs)和局部信息。

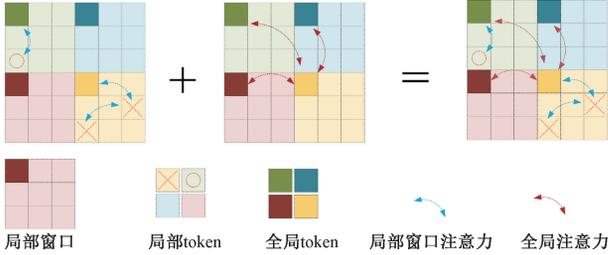


图3 层级注意力工作原理示意图

Fig. 3 Diagram of the working principle of hierarchical attention

在具体实现过程中, 对于输入特征图 $\mathbf{x} \in \mathbb{R}^{H \times W \times d}$, 将其划分为 $n \times n$ 的局部窗口, 如下式(3)所示。

$$\mathbf{x}_1^0 = \text{split}_{n \times n}(\mathbf{x}), n = \frac{H \times W}{k^2} \quad (3)$$

式中: H, W 分别为输入特征图的长度和宽度, d 为输入特征图的通道数, k 为窗口尺寸。通过 3×3 卷积对特征图进行编码构建全局令牌, 如下式(4)所示。

$$\mathbf{x}_c^e = \text{Conv}_{3 \times 3}(\mathbf{x}) \quad (4)$$

然后, 使用平均池化操作将编码后的特征图的尺寸缩减为长度为 L 的全局载体令牌, 用于表示对应窗口特征的汇总信息, 如下式(5)所示。

$$\mathbf{x}_{ct}^e = \text{AvgPool}_{H \times W \rightarrow n^2 L}(\mathbf{x}_c^e) \quad (5)$$

在每个分层注意力模块中, 通过多头自注意力机制更新全局 CTs, 以更好地实现信息交互和特征学习。具体步骤为:

1) 使用多头自注意力机制, 计算每个全局 CTs 与窗口标记之间的相似性。

2) 根据计算得到的注意力权重, 对全局 CTs 进行加权求和, 以获得更新后的 CTs 表示, 计算如式(6)所示。

$$\mathbf{x}_{ct}^{e^1} = \mathbf{x}_{ct}^e + \gamma_1 \odot \text{MHSA}(\text{LN}(\mathbf{x}_{ct}^e)) \quad (6)$$

3) 应用多层感知器(MLP)和 GeLU 激活函数进行后处理, 以使网络能够更好地捕捉上下文信息, 计算过程如式(7)所示。

$$\mathbf{x}_{ct}^{e^2} = \mathbf{x}_{ct}^{e^1} + \gamma_2 \odot \text{MLP}(\text{LN}(\mathbf{x}_{ct}^{e^1})) \quad (7)$$

式中: \mathbf{x}_{ct}^e 表示原始未更新的 CTs, MHSA 表示多头自注意力操作, $\mathbf{x}_{ct}^{e^1}$ 表示通过多头自注意力加权的 CTs。MLP 为通过 GeLU 激活函数和 MLP 引入非线性拟合的操作, 用于进一步处理更新后的 CTs, 增强网络的表示能力。 $\mathbf{x}_{ct}^{e^2}$ 为通过 MLP 引入非线性拟合后的加权 CTs, γ_1 和 γ_2 为可学习的权重参数。LN 为层归一化操作, 用于规范化注意力权重。

接下来通过本地窗口的令牌和全局 CTs 之间的交互作用, 增强网络对长距离和短距离空间特征信息的建模能力。首先, 将本地窗口的令牌和全局 CTs 进行拼接; 然后, 通过多头自注意力操作更新拼接后的令牌, 接着通过 MLP 操作后进一步加权更新以得到最终的令牌。这个交互作用的计算过程如下式(8)、(9)和(10)所示。

$$\mathbf{x}_w^e = \text{Concat}(\mathbf{x}_1^0, \mathbf{x}_{ct}^{e^2}) \quad (8)$$

$$\mathbf{x}_w^{e^1} = \mathbf{x}_w^e + \gamma_3 \odot \text{MHSA}(\text{LN}(\mathbf{x}_w^e)) \quad (9)$$

$$\mathbf{x}_w^{e^2} = \mathbf{x}_w^{e^1} + \gamma_4 \odot \text{MLP}(\text{LN}(\mathbf{x}_w^{e^1})) \quad (10)$$

式中: \mathbf{x}_w^e 为拼接得到的 token, $\mathbf{x}_w^{e^1}$ 为通过多头自注意力加权后的 token, $\mathbf{x}_w^{e^2}$ 为 MLP 处理后进一步加权的 token, γ_3 和 γ_4 也为可学习的权重参数。

通过上述操作, 允许本地窗口的令牌与全局载体令牌之间相互影响, 有助于更好地捕捉图像中的重要特征和上下文信息。然后将更新后的 token 拆分为局部窗口令牌和全局载体令牌, \mathbf{x}_1^0 表示局部窗口令牌, 它包含了 token 向量的前 L 个令牌。 $\mathbf{x}_{ct,1}^e$ 表示全局载体令牌, 它包含了从 token 向量的第 L 个令牌开始的后续令牌, 操作过程如下式(11)所示。

$$(\mathbf{x}_1^0, \mathbf{x}_{ct,1}^e) = \text{split}(\mathbf{x}_w^{e^2}) \quad (11)$$

最终得到的层级注意力输出为(12)式所示。

$$\mathbf{x} = \text{upsample}_{n^2 L \rightarrow H \times W}(\mathbf{x}_{ct,1}^e) + \text{merge}_{n^2 k^2 \rightarrow H \times W}(\mathbf{x}_1^0) \quad (12)$$

为了确保网络在空间维度保留特征的位置信息, 引入了绝对位置偏移, 并采用双层 MLP 将绝对位置偏移信息嵌入到特征中。图 4 展示了层级注意力的计算过程, 将自注意力分为全局自注意力和局部自注意力两个部分。同时, 通过与载体令牌的交互, 全局自注意力和局部自注意力使网络能够更好地融合全局和局部信息, 从而进一步增强网络的特征提取能力。

1.3 基于 SimAM 的特征融合网络

在处理视觉信息时, 人类会格外关注图像中的目标细节特征, 而忽略其他不必要的背景信息干扰。同样, 为了降低在 PCB 表面缺陷检测过程中, 因缺陷目标与背景相似性而产生的干扰问题, 本文创新性地颈部融合

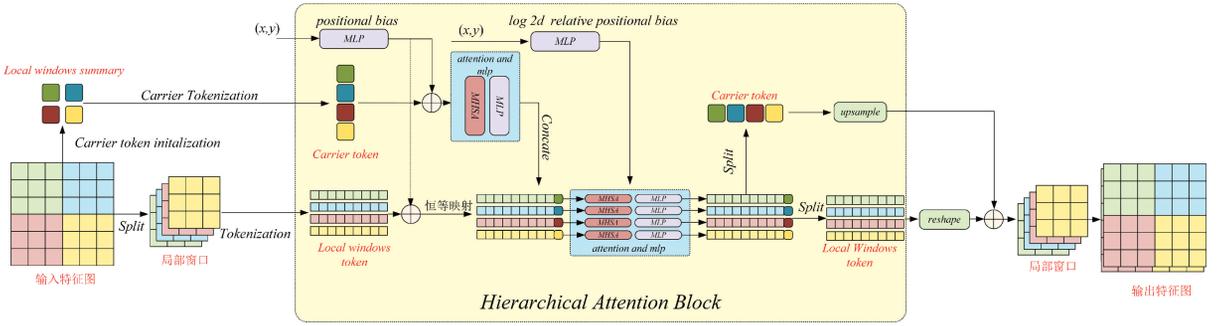


图4 层级注意力计算过程示意图

Fig. 4 Illustration of the computation process of hierarchical attention

网络中引入了一种无参数的3D权重SimAM注意力机制^[25]。该机制与现有的基于通道和空间的注意力不同,其通过能量函数来发掘神经元的重要性,从而在无需额外参数的情况下得到注意力权重值。SimAM基于人脑神经元具有一定的空域抑制效应的特性,给予具有空域抑制效应的神经元更高的权重。对于输入特征图 $X \in \mathbb{R}^{H \times W \times C}$, SimAM的能量函数定义为:

$$e_i(w_i, b_i, y, x_i) = \frac{1}{M-1} \sum_{i=1, i \neq t}^M (-1 - (w_i x_i + b_i))^2 + (1 - (w_i t + b_i))^2 + \lambda w_i^2 \quad (13)$$

式中: M 为通道上的神经元数量, λ 为权重系数, t, x_i 分别为目标神经元和其他神经元, $w_i t + b_i$ 和 $w_i x_i + b_i$ 是对目标神经元和其他神经元的线性变换。为使能量函数取得最小值,对(13)式求解可得到其解析解:

$$w_i = -\frac{2(t - \mu_i)}{(t - \mu_i)^2 + 2\sigma_i^2 + 2\lambda} \quad (14)$$

$$\sigma_i^2 = \frac{1}{M-1} \sum_{i=1}^{M-1} (x_i - \mu_i)^2 \quad (15)$$

其中, μ_i 为除目标神经元外的其他神经元的均值, σ_i^2 为其他神经元的方差。为了节省计算量,考虑到每个通道的神经元具有相同的分布,将同一个通道上的所有 μ_i 用下式(16)估算:

$$\mu = \frac{1}{M} \sum_{i=1}^M x_i \quad (16)$$

方差用(17)式近似:

$$\sigma^2 = \frac{1}{M} \sum_{i=1}^M (x_i - \mu)^2 \quad (17)$$

此时每个位置的最小能量可用式(18)表示。

$$e_i^* = \frac{4(\sigma^2 + \lambda)}{(t - \mu)^2 + 2\sigma^2 + 2\lambda} \quad (18)$$

e_i^* 越小,神经元就越具有空域抑制效应,其重要性就越高,因此可将能量函数的倒数再经过Sigmoid函数约束后作为注意力权重,最终经注意力处理后的输出特征 X^* 如下式(19)所示。

$$X^* = \text{sigmoid}\left(\frac{1}{E}\right) \odot X \quad (19)$$

其中, E 是 e_i^* 在所有通道和空间维度的汇总。

引入SimAM的特征融合网络如下图所示,该特征融合网络包含自上而下和自下而上两条特征融合路径。在进行不同尺度的特征融合时,自上而下的路径通过上采样将上一层的特征图增大到下一层相应大小,而自下而上的路径则通过下采样将下一层的特征图缩减到上一层相应大小。接着使用CSP_Layer将特征图划分为两个部分,并将其中一部分的特征与另一部分的特征进行拼接。随后,将经过变换后大小相同的特征图利用SimAM进行加权。最后,将加权后的特征图拼接,生成融合后的特征图。

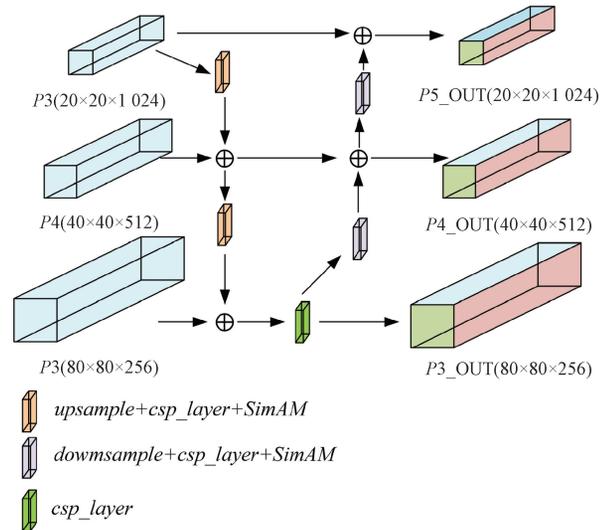


图5 基于SimAM的特征融合网络结构图

Fig. 5 Network architecture diagram for feature fusion based on SimAM

假设主干网络提取的待融合的两个特征图分别为 $p_i \in \mathbb{R}^{B \times H_1 \times W_1 \times C_1}$ 和 $p_{i+1} \in \mathbb{R}^{B \times H_2 \times W_2 \times C_2}$,首先对 p_{i+1} 进行下

采样或上采样,将其变换到 $\mathbf{p}_{i+1}^* \in \mathbb{R}^{B \times H_1 \times W_1 \times C_1}$, 然后通过 SimAM 得到注意力加权后的融合特征图 \mathbf{p}_λ^* , 其计算公式如式(20)所示。

$$\mathbf{p}_\lambda^* = \text{sigmoid}\left(\frac{1}{\mathbf{E}^*}\right) \odot \mathbf{p}_{i+1}^* \quad (20)$$

式中: \mathbf{E}^* 表示特征图 \mathbf{p}_{i+1}^* 的神经元能量在所有通道和空间维度的汇总。最后将注意力处理后的特征图拼接得到输出特征图,即 $\mathbf{p}_{out}^* = \text{concat}(\mathbf{p}_i | \mathbf{p}_\lambda^*)$ 。经过这样拼接融合后的特征包含了缺陷目标更丰富的特征,有助于解耦头进行目标的检测与定位,从而提高 PCB 缺陷检测的准确率。

1.4 引入均衡 Focal Loss 的损失函数设计

在工业环境中,由于缺陷样本出现的几率较低,导致在缺陷检测的实际应用中,能获取的数据集往往规模较小,并呈现出明显的类别不平衡和长尾分布的特点。为解决此类问题,目前通常采用的两种主要方法是重采样^[26]和重加权^[27]。在重采样方法中,针对数据集中的稀有类别,采取过采样的策略,同时对数据集中频繁出现的类别进行欠采样,以实现数据分布的平衡。而在重加权方法中,通过引入不同训练样本的损失函数权重,以解决训练过程中频繁类别对模型训练的主导作用。Focal Loss^[28]函数就是重加权方法的一种,通过引入调节因子来解决训练过程中不同类别样本不平衡的问题。为了进一步解决不同类别正负样本不平衡问题,研究人员提出了均衡 Focal Loss (equalized focal loss, EFL)^[29]。EFL 通过引入与类别有关的调节因子解决不同类别正负样本不平衡问题,对于第 j 类样本来说,损失计算如式(21)所示。

$$EFL(p_i) = -\alpha_i(1-p_i)^{\gamma^j} \log(p_i) \quad (21)$$

式中: $p_i \in [0, 1]$ 表示网络输出的预测置信度分数,参数 α_i 平衡正负样本之间的重要性。 γ^j 为与类别有关的调节因子。

本文结合 EFL 对网络的分类损失进行改进,以期改进后的损失函数提高网络在类别不平衡 PCB 上的缺陷检测性能。引入 EFL 构建的分类损失函数如式(22)所示。

$$L_{cls} = -\frac{1}{N_{obj}} \sum_{i=1}^{N_{obj}} \sum_{j=1}^n \alpha_i(1-p_{i,j})^{\gamma^j} \log(p_{i,j}) \quad (22)$$

其中, N_{obj} 是包含目标的预测框的数量, n 为类别数, $p_{i,j}$ 是第 i 个预测框属于第 j 类的概率。 γ^j 由与类别无关的调节因子系数 γ_b 和与类别有关的调节系数 γ_v^j 构成, γ^j 如式(23)所示。

$$\gamma^j = \gamma_b + \gamma_v^j \quad (23)$$

在学习过程中 γ_v^j 采用梯度引导机制^[30]进行计算,即 $\gamma_v^j = s(1-g^j)$ 。第 j 类的正负样本累计梯度比值 $g^j \in$

$[0, 1]$ 越大,第 j 类就越趋向于平衡训练。 s 决定 EFL 中与类别相关调节因子的大小。使用 EFL 构建分类损失函数有助于处理类别不平衡的问题,提高在稀有类别上的分类性能。

此外为了度量模型对检测框位置的预测与实际标注框位置之间的差异,使用泛化交并比 (generalized intersection over union, GIOU) 计算位置损失,位置损失函数如式(24)所示。

$$L_{reg} = 1 - \frac{1}{N_{obj}} \sum_{i=1}^{N_{obj}} \times \left[\left(\text{IoU}(\hat{B}_i, B_i) - \left| \frac{A^c(\hat{B}_i, B_i) - U(\hat{B}_i, B_i)}{A^c(\hat{B}_i, B_i)} \right| \right) \right] \quad (24)$$

式中: $A^c(\hat{B}_i, B_i)$ 为包围第 i 个预测框和真实框的最小矩形框的面积, $U(\hat{B}_i, B_i)$ 为第 i 个预测框和真实框并集的面积, $\text{IoU}(\hat{B}_i, B_i)$ 为第 i 个预测框和真实框的交并比。

为了评估模型检测结果的置信度,采用二元交叉熵损失函数计算置信度损失。计算过程中将模型输出的置信度通过 Sigmoid 激活函数得到预测框有目标的概率。置信度损失函数如式(25)所示。

$$L_{obj} = -\frac{1}{N_{obj}} \sum_{i=1}^{N_{obj}} (y_i \log(p_i) + (1-y_i) \log(1-p_i)) \quad (25)$$

其中, y_i 为二元标签, p_i 为第 i 个预测框含有目标的概率。

最终设计的综合损失函数包括检测框分类类别损失、位置损失以及置信度损失,如下式(26)所示。

$$L_{oss} = L_{obj} + \lambda_1 L_{cls} + \lambda_2 L_{reg} \quad (26)$$

其中, λ_1, λ_2 是用于调节三种类型损失权重的重要超参数。综合损失函数确保在训练网络时能够充分反映和权衡分类准确率、目标位置和检测结果可信度这三方面因素。借助该损失函数来指导网络优化参数,提高模型在类别不平衡数据集上的缺陷检测精度。

2 实验与结果分析

2.1 数据集简介

本文采用的第一个数据集是北京大学智能机器人实验室公开的 PCB 缺陷检测数据集^[31] (PKU-Market-PCB)。该数据集共 693 张图像,包含缺孔、开路、短路、鼠咬、假铜和骨刺等缺陷,各种缺陷示例见图 6。为防止网络过拟合,本文对原始数据集进行了旋转、添加噪声、翻转等预处理。处理后的数据集总共有 4 158 张 PCB 图像,包含 15 916 处缺陷实例,其中缺孔类缺陷有 2 725 例、鼠咬缺陷 2 650 例、开路缺陷 2 593 例、短路缺陷 2 605 例、骨刺缺陷 2 622 例、假铜类缺陷有 2 727 例。训

练过程中,按照 8:1:1 的比例将数据集划分为训练集、验证集和测试集。

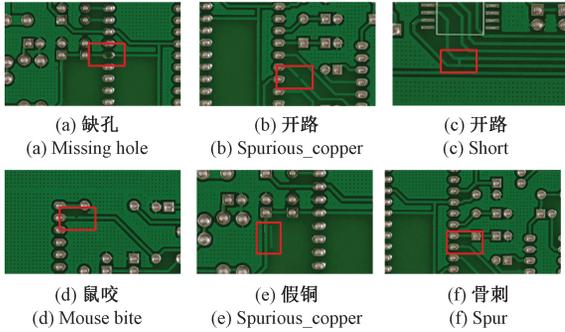


图6 PCB表面缺陷示例

Fig. 6 Examples of PCB surface defects

2.2 评价指标

本文采用了平均检测精度均值 (mean average precision, mAP) 和模型每秒处理的图像数 (frames per second, FPS) 这两项指标,以全面评估模型的性能。mAP 的计算涉及到所有类别的准确率-召回率 (P-R) 曲线的综合。在每个类别上绘制一个 P-R 曲线,曲线下的面积即为平均检测精度 (AP)。然后求所有类别上 AP 的平均值得到 mAP。AP 和 mAP 的计算公式可表示为(27)和(28)。

$$AP = \int_0^1 p(r) dr \quad (27)$$

$$mAP = \frac{\sum_{i=1}^N AP_i}{N} \quad (28)$$

式中: $p(r)$ 为 P-R 曲线, N 为缺陷类别数。

模型 FPS 作为评价模型检测速度的指标,定义为网络在给定单位时间内能够处理的最大图像样本数目,其计算公式可表示为式(29)。

$$FPS = \frac{N}{step_time} \quad (29)$$

式中: N 为一个批次实例图像的数量, $step_time$ 为处理完一个批次数据所耗费的时间。

2.3 实验环境与超参数设置

本文实验环境配置如下:操作系统为 Ubuntu20.04, CPU 为 Intel(R) Xeon(R) Gold6240 @ 2.60 GHz, GPU 为 4 张 NVIDIA GeForce RTX 2080Ti 显卡,显卡显存 12 G;基于 PyTorch 深度学习框架搭建算法,PyTorch 版本为 1.11.0, CUDA 版本为 10.2, cuDNN 版本为 8.2.0。

网络超参数设置为:输入图像大小为 640×640 , 批次大小为 4, 使用 SGD 优化器, 初始学习率设置为 1×10^{-2} , 采用余弦退火算法进行学习率衰减, 总的训练 Epoch 轮次设置为 150 轮, 优化器动量系数设置为 0.937, 权重衰减系数为 5×10^{-4} , 特征提取 4 个阶段窗口尺寸分别设置

为 7×7 、 7×7 、 20×20 、 20×20 。在构建均衡 Focal Loss 时, 通过实验将平衡正负样本之间重要性的参数 $\alpha_i = 0.25$, 类别无关的调节因子系数 γ_i 设置为 2.0。在计算类别有关调节系数过程中, 类别相关调节因子 s 的大小设置为 8.0。通过实验将平衡三种损失的系数 λ_1 设为 1, λ_2 设为 5.0, 从而使模型的训练效果达到最优。网络训练的损失函数与 mAP 变化曲线如下图 7 所示, 实线线条表示基线模型 YOLOX 在训练过程中的 Loss 曲线和 mAP 的变化情况, 点划线表示在 YOLOX 中加入 MulPViT 特征提取模块和 SimAM 特征融合模块后的 Loss 曲线和 mAP 曲线, 短点曲线是在上述情况下再引入 EFL 后的 Loss 和 mAP 变换曲线。

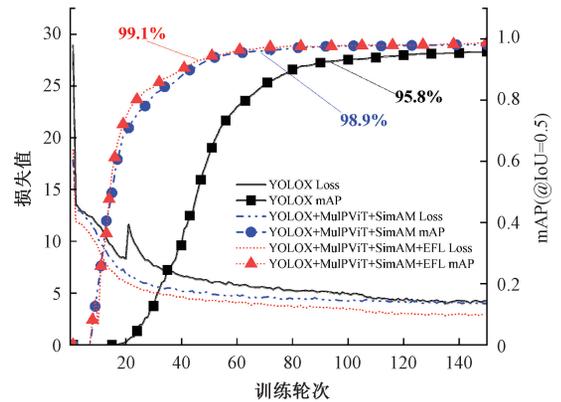


图7 网络训练过程中损失函数值与 mAP 曲线

Fig. 7 Loss Function and mAP Curve during Network Training

由图 7 可见, 在 YOLOX 基础上使用 MulPViT 特征提取模块和 SimAM 特征融合模块后, 网络训练过程中的 mAP 值有明显的提升; 当迭代次数超过 120 轮时, 网络的损失函数值趋于稳定, 与此同时, mAP 值也呈现出稳定的趋势。值得注意的是, 所提出的改进方法在训练过程中能够更快地达到相应的精度水平, 显著超越了基线模型的平均检测精度。此外, 通过使用引入 EFL 后的损失函数, 损失值下降的趋势更明显, 网络的检测精度也有提升。结果表明了改进特征提取网络和特征融合网络的有效性以及损失函数的重要作用。这种改进方法不仅能够显著提升模型性能, 还有助于提高模型训练的稳定性。

2.4 与其他方法的比较分析

为了验证本文所提出方法的有效性与先进性, 在 PKU-Market-PCB 数据集、上海交通大学黄晓霖教授开源的数据集 Deep PCB^[32] 以及 PCBA 缺陷检测项目数据集^[33] 上, 将本文所提方法与现有的 PCB 检测算法 MAFaster RCNN^[12]、TDD-Net^[34] 和目前具有 SOTA 性能的一阶段目标检测算法 MobileNetV3-YOLO^[1]、YOLOX^[21]、YOLOv3^[35]、AT-YOLO^[36]、SW-YOLOX^[37] 等方法进行对比, 实验结果如表 1 所示。

表 1 不同数据集上检测算法的评价指标值
Table 1 Evaluation metric values for different detection algorithms

数据集	方法	骨干网络	mAP (@ IoU = 0.5) / %	FPS
PKU-Market-PCB	Faster RCNN	ResNet-50	93.61	9.03
	SSD	VGG16	89.23	94.17
	YOLOv3	DarkNet-53	96.53	59.12
	YOLOX	CSPDarkNet53	95.84	84.05
	TDD-Net ^[34]	ResNet-101	96.40	-
	MAFaster RCNN ^[12]	ResNeSt-50	97.90	-
	AT-YOLO ^[36]	ResNeSt-50	98.40	-
	MobileNetV3-YOLOv4-Inceptionv3 ^[1]	MobileNetV3	99.10	43.01
	Swin-Transformer-YOLOX ^[37]	Swin-Transformer	93.86	17.66
	本文方法	MulPViT	99.13	25.97
Deep PCB ^[32]	Faster RCNN	ResNet-50	96.71	10.89
	YOLOv4	DarkNet-53	91.40	29.20
	文献[38]	CSPGohstNet	98.30	49.30
	文献[32]	-	98.60	62.00
	本文方法	MulPViT	98.67	25.10
PCBA ^[33]	Faster RCNN	ResNet-50	95.83	9.01
	YOLOv3	DarkNet-53	96.69	61.27
	YOLOX	DarkNet-53	98.45	54.10
	MobileNetV3-YOLOv4	MobileNetV3	99.26	50.56
	Swin-Transformer-YOLOX ^[37]	Swin-Transformer	99.49	16.12
	本文方法	MulPViT	99.82	27.70

由表 1 可知,本文所提出的方法在 PCB 检测上的 mAP 达到了 99.13%,较基线模型 YOLOX 提高了 3.3%,提升了对 PCB 表面缺陷的检测性能。同时所提出的方法在检测速度方面,同时模型的 FPS 达到了 25.97,满足实际工业 PCB 缺陷检测的实时性要求。

在 PKU-Market-PCB 数据集上,与经典的 Faster RCNN、SSD 和 YOLOv3 等目标检测算法相比,本方法在检测精度上均表现更为出色。需要指出的是,相对于 SSD 而言,本方法检测速度较慢,这是因为 SSD 采用了较浅的 VGG16 网络进行特征提取,从而导致检测精度不高,但同时也带来了更快的检测速度。与 YOLOv3 相比,本方法的平均检测精度提高了 2.6%。虽然单次检测速度略有降低,但所提出方法的 FPS 达到了 25.97 帧,满足实际工业检测的实时性要求。与其他用于 PCB 表面缺陷检测的方法相比,本方法的检测性能也有所提升。相对于 TDD-Net 和 MAFaster RCNN,本方法的平均检测精度分别提高了 2.7% 和 1.2%。与 AT-YOLO 相比,本方法的检测精度提高了 0.7%。与目前具有最好 mAP 指标值的轻量级方法 MobileNetV3-YOLOv4-Inceptionv3 相比,本

研究方法在检测精度上进一步提升,取得了最高的 mAP,此外,尽管检测速度略有降低,但仍满足工业应用的实时检测需求。与最近提出的具有 SOTA 性能的小目标检测算法 Swin-Transformer-YOLOX 相比,本方法在检测精度和速度方面均有较大提升。值得注意的是,该算法和本文算法均以 Transformer 的改进方法为骨干网络,对比结果表明了本文特征提取网络更加有效。

在 Deep PCB 数据集上,相较于其他四种目标检测算法,所提出的算法在检测精度上均有提升,具体而言,相较于 YOLOv4,提升幅度达到了 7.27%;与文献[36]中针对 PCB 缺陷检测的方法相比,检测精度提升了 0.37%。为进一步验证所提出算法的优越性,在 PCBA 数据集上进行了对比实验,所提出算法的平均检测精度达到了 99.82%,相较于基线模型 YOLOX,检测精度提高了 1.37%;同时,与 SOTA 的 MobileNetV3-YOLOv4 和 Swin-Transformer-YOLOX 相比,所提出方法的检测精度也有所提升。

总的来说,综合考虑本算法与当前用于 PCB 缺陷检测和目标检测算法的对比结果,本方法在检测精度和检

测速度方面均具备较高的性能水平。

2.5 消融实验

为验证本文所提出检测模型的合理性,进行了一系列消融实验。在消融实验过程中,主要考虑本研究提出的不同改进对模型性能的提升效果。所提出的 PCB

缺陷检测算法主要包括以下关键部分:1) 基于分层注意力和部分卷积的多尺度特征提取网络 MulPViT;2) 基于无参数注意力 SimAM 的颈部特征融合网络;3) 引入均衡 Focal Loss 的损失函数改进。在 PKU-Market-PCB 数据集上进行了消融实验,下表 2 给出了消融实验结果。

表 2 消融实验结果
Table 2 Results of ablation experiments

方法	mAP (@ IoU=0.5)/%	AP (@ IoU=0.5)/%					
		穿孔	鼠咬	开路	短路	骨刺	假铜
YOLOX	95.83	99.44	94.78	96.44	97.71	92.72	93.90
YOLOX+EFL	95.88	99.82	94.10	95.59	97.62	92.17	95.98
YOLOX+SimAM	96.14	99.35	95.92	96.65	98.20	90.86	95.84
YOLOX+MulPViT	98.43	98.80	98.00	98.70	98.21	97.54	99.31
YOLOX+SimAM+EFL	95.99	99.82	94.07	95.13	97.60	93.00	96.29
YOLOX+MulPViT+EFL	98.19	99.79	97.87	98.00	99.02	95.97	98.50
YOLOX+MulPViT+SimAM	98.96	100.00	99.07	98.05	98.54	98.77	98.89
YOLOX+MulPViT+SimAM+EFL	99.13	100.00	97.88	98.20	99.61	99.22	99.89

由上表 2 可知,在 YOLOX 基础上,相对于单独引入 SimAM 和 EFL,引入 MulPViT 后 mAP 提升最多,为 2.6%。在颈部特征融合网络中加入 SimAM 后,各类别上的 AP 值均在 98% 之上,mAP 再次提升了 0.53%。继续在网络中引入 EFL 损失函数后,mAP 又提升了 0.17%。上述实验结果表明,本文提出的算法改进措施对 PCB 表面缺陷的平均精度均值提升均有一定的作用,MulPViT 特征提取网络对于提升缺陷检测准确率的作用最为突出;基于 SimAM 的改进特征融合网络则有效提高了各类缺陷检测中模型的稳定性;基于 EFL 改进的损失函数进一步提高了网络的检测准确率。

2.6 类别不均衡数据集上的实验结果分析

为了验证本方法在具有长尾分布特性的检测任务上的性能,在类别均衡的数据集 PKU-Market-PCB 上进行一系列样本的生成和删减操作,让其更接近工业应用场景中缺陷样本不均衡且呈现长尾分布的现象。处理后的数据集包含 1 693 张 PCB 图像,其中穿孔缺陷实例 2 402 个、鼠咬缺陷 1 409 个、开路缺陷 720 例、短路缺陷 468 例、骨刺缺陷 352 例、假铜缺陷 505 例,稀有类和频繁出现类之间的数量差达到了 8 倍,处理前后各类别实例数量分布如图 8 所示。处理后的各类别实例数量不均衡,数据呈现长尾分布,极大地增加了网络的构建和训练难度。在此数据集上模型的消融实验结果如表 3 所示。

由表 3 可知,YOLOX 在稀有类别的检测精度较低,特别是对于假铜类别的 AP 仅为 57.38%,骨刺类别的

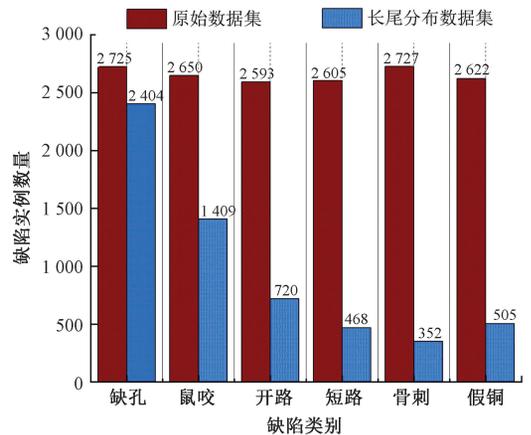


图 8 各类别缺陷实例数量对比图

Fig. 8 Comparison of the number of instances for each defect class

AP 值也仅为 67.55%,全类别的 mAP 为 77.16%。通过引入 MulPViT 特征提取网络和 SimAM 构建的特征融合网络,进一步提取和融合了缺陷目标的细节特征,mAP 提高了 8.68%。相应地在各类缺陷上的平均精度均有所提升,尤其在实例较少的假铜缺陷和骨刺缺陷上 AP 指标值有了较大提升,分别达到 69.59% 和 79.30%。特别地,对于短路缺陷,模型的检测精度 AP 值提升了 18.98%。然而,在实例较少的类别上的平均精度仍较低,通过进一步引入基于均衡 Focal Loss 的损失函数,在假铜缺陷上的平均精度提高了 7.01%,其他类别上的

表 3 类别不平衡 PCB 数据集上的实验结果
Table 3 Experimental results on class-imbalanced PCB dataset

方法	mAP(@ IoU=0.5)/%	AP(@ IoU=0.5)/%					
		穿孔	鼠咬	开路	短路	骨刺	假铜
YOLOX	77.16	99.62	83.68	82.33	72.41	67.55	57.38
YOLOX+MulPViT+SimAM	85.84	99.83	89.41	85.59	91.39	79.30	69.59
YOLOX+MulPViT+SimAM+EFL	89.10	99.92	92.11	92.16	91.41	82.41	76.60

平均检测精度也都有不同程度的提升。这种损失函数设计方法通过平衡不同类别的损失,显著提高了稀有类别的检测精度,相对于损失函数改进前,mAP 提高了 3.26%。本文算法相较于 YOLOX 算法的 mAP 提高了 11.94%。实验结果验证了本文所提出的网络结构和损失函数能显著提升类别不平衡 PCB 缺陷检测的性能。

2.7 检测结果可视化对比

为了更加直观地展现本文所提算法的先进性,以 PKU-Market-PCB 为例,在相同的硬件环境下,分别使用基线模型 YOLOX 和本文方法进行检测,前向推理过程中

均采用非极大值抑制算法,IoU 阈值为 0.5,检测结果的置信度阈值为 0.5,结果如图 9 所示,从上到下依次为真实值(Ground Truth)、模型 YOLOX 的可视化检测效果和本文方法的检测效果。从图 9 可知,针对不同的缺陷类型,本文所提出的算法可以准确检测出 PCB 上存在的缺陷目标。基线模型 YOLOX 在开路缺陷和骨刺缺陷检测过程中出现了漏检情况,而本文所提出的 PCB 缺陷检测方法没有出现缺陷的漏检与错检。且通过实验可知,本文所提出的方法针对 PCB 表面缺陷检测任务的检测速度达到 25 fps。综合上述分析,本文所提出的方法针对 PCB 缺陷可以达到较好的检测效果。

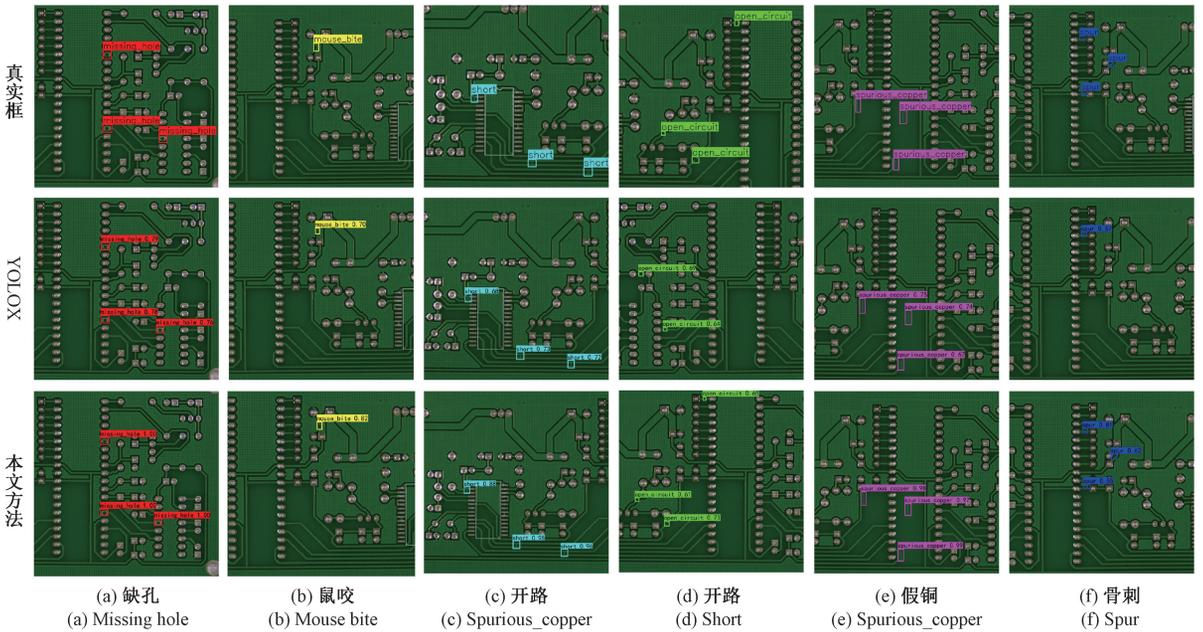


图 9 PCB 表面缺陷检测结果可视化

Fig.9 Visualization of PCB surface defect detection results

3 结 论

本文针对类别不平衡 PCB 缺陷检测中目标特征的有效提取和计算复杂度高这些核心问题,提出了集特征提取、特征融合和损失函数全面改进的 PCB 缺陷检测新

算法,并通过实例验证了本文方法的有效性,得到的主要结论有:基于分层多头注意力机制和部分卷积等构建的多尺度 ViT 特征提取网络和基于 SimAM 注意力机制设计的特征融合网络能有效增强网络的特征表达能力,提高了网络的检测精度,特别是在类别不平衡数据集上的检测精度提升更为明显;采用均衡 Focal Loss 分类函数、

基于 GIOU 的位置损失函数和二元交叉熵的置信度综合构建的损失函数也有利于提高缺陷检测精度,特别在样本较少缺陷类别上的检测精度提升更为显著;所提方法在检测精度上较改进前方法进一步提升的同时,仍具有较高的实时效率,能满足工业应用的实时性要求。该方法对于其他相似工业应用领域的缺陷检测具有参考意义。

参考文献

- [1] 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.
WU J G, CHENG Y, SHAO J, et al. A defect detection method for PCB based on the improved YOLOv4[J]. Chinese Journal of Scientific Instrument, 2021, 42(10): 171-178.
- [2] 赵朗月,吴一全. 基于机器视觉的表面缺陷检测方法研究进展[J]. 仪器仪表学报, 2022, 43(1):198-219.
ZHAO L Y, WU Y Q. Research progress of surface defect detection methods based on machine vision[J]. Journal of Instrumentation, 2022, 43(1):198-219.
- [3] 卢荣胜,吴昂,张腾达,等. 自动光学(视觉)检测技术及其在缺陷检测中的应用综述[J]. 光学学报, 2018, 38(8):23-58.
LU R SH, WU ANG, ZHANG T D, et al. Review on automated optical (vision) inspection and its applications in defect detection [J]. Acta Optica Sinica, 2018, 38(8):23-58.
- [4] 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.
DONG J Y, LYU W T, BAO X M, et al. Research progress of the PCB surface defect detection method based on machine vision [J]. Journal of Zhejiang Sci-Tech University, 2021, 45(3):379-389.
- [5] ANNABY M H, FOU DA Y M, RUSHDI M A. Improved normalized cross-correlation for defect detection in printed-circuit boards [J]. IEEE Transactions on Semiconductor Manufacturing, 2019, 32(2):199-211.
- [6] 刘艳菊,王秋霖,赵开峰,等. 基于卷积神经网络的热轧钢条表面实时缺陷检测[J]. 仪器仪表学报, 2021, 42(12):211-219.
LIU Y J, WANG Q J, ZHAO K F, et al. Real-time defect detection of hot rolling steel bar based on convolution neural network [J]. Chinese Journal of Scientific Instrument, 2021, 42(12):211-219.
- [7] 陶显,侯伟,徐德. 基于深度学习的表面缺陷检测方法综述[J]. 自动化学报, 2021, 47(5):1017-1034.
TAO X, HOU W, XU D. A survey of surface defect detection methods based on deep learning [J]. Acta Optica Sinica, 2021, 47(5):1017-1034.
- [8] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: Unified, real-time object detection[C]. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas: IEEE, 2016: 779-788.
- [9] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single shot multibox detector[J]. IEEE European Conference on Computer Vision, 2016, 10: 21-37.
- [10] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation [C]. 2014 IEEE Conference on Computer Vision and Pattern Recognition, 2014: 580-587.
- [11] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6):1137-1149.
- [12] 陈仁祥,詹赞,胡小林,等. 基于多注意力 Faster RCNN 的噪声干扰下印刷电路板缺陷检测[J]. 仪器仪表学报, 2021, 42(12):167-174.
CHEN R X, ZHAN Z, HU X L, et al. Printed circuit board defect detection based on the multi-attentive faster RCNN under noise interference[J]. Chinese Journal of Scientific Instrument, 2021, 42(12):167-174.
- [13] LUO J, YANG Z, LI S, et al. FPCB surface defect detection: A decoupled two-stage object detection framework [J]. IEEE Transactions on Instrumentation and Measurement, 2021, 70: 1-11.
- [14] ADIBHATLA V A, CHIH H C, HSU C, et al. Applying deep learning to defect detection in printed circuit boards via a newest model of you-only-look-once [J]. Mathematical Biosciences and Engineering: MBE, 2021, 18(4):4411-4428.
- [15] DOSOVITSKIY A, BEYER L, KOLESNIKOV A, et al. An image is worth 16x16 words: Transformers for image

- recognition at scale [J]. International Conference on Learning Representations, ICLR, 2021, DOI: org/10.48550/arXiv.2010.11929.
- [16] 田永林,王雨桐,王建功,等. 视觉Transformer研究的关键问题:现状及展望[J]. 自动化学报, 2022, 48(4):957-979.
- TIAN Y L, WANG Y T, WANG J G, et al. Key problems and progress of vision Transformers: The state of the art and prospects[J]. ACTA OPTICA SINICA, 2022, 48(4):957-979.
- [17] 彭雨彤,梁凤梅. 融合CNN和ViT的乳腺超声图像肿瘤分割方法[J/OL]. 智能系统学报, DOI:10.11992/tis.202304046.
- PENG Y T, LIANG F M. A tumor segmentation method for breast ultrasound images incorporating CNN and ViT [J/OL]. CAAI Transactions on Intelligent Systems, DOI:10.11992/tis.202304046.
- [18] 徐昊,郭黎,李润泽. 基于紧凑型Vision Transformer的细粒度视觉分类研究[J]. 控制与决策, 2024, 39(3):893-900.
- XU H, GUO L, LI R Z. Fine-grained visual classification based on computer Vision Transformer[J]. Control and Decision, 2024, 39(3):893-900.
- [19] LIU Z, LIN Y T, CAO Y, et al. Swin Transformer: Hierarchical Vision Transformer using shifted windows[C]. IEEE/CVF International Conference on Computer Vision (ICCV), 2021: 9992-10002.
- [20] GUO J Y, HAN K, WU H, et al. CMT: Convolutional neural networks meet vision transformers [C]. 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2021: 12165-12175.
- [21] GE Z, LIU S, WANG F, et al. YOLOX: Exceeding YOLO series in 2021[J]. ArXiv Preprint, 2021, ArXiv: 2107.08430.
- [22] HATAMIZADEH A, HEINRICH G, YIN H X, et al. FasterViT: Fast vision transformers with hierarchical attention [J]. ArXiv Preprint, 2023, ArXiv: 2306.06189.
- [23] 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), 2023: 12021-12031.
- [24] HAN K, WANG Y H, TIAN Q, et al. Ghostnet: More features from cheap operations [C]. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2020: 1577-1586.
- [25] YANG L X, ZHANG R Y, LI L, et al. SimAM: A simple, parameter-free attention module for convolutional neural networks [C]. International Conference on Machine Learning, 2021:11863-11874.
- [26] BUDA M, MAKI A, MAZUROWSKI A. A systematic study of the class imbalance problem in convolutional neural networks [J]. Neural Networks, 2018, 106: 249-259.
- [27] CUI Y, JIA M L, LIN T Y, et al. Class-balanced loss based on effective number of samples [C]. 2019 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2019: 9260-9269.
- [28] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection [C]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 42(2): 318-327.
- [29] LI B, YAO Y Q, TAN J R, et al. Equalized focal loss for dense long-tailed object detection [C]. 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2022: 6980-6989.
- [30] TAN J R, LU X, ZHANG G, et al. Equalization loss v2: A new gradient balance approach for long-tailed object detection [C]. 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2021: 1685-1694.
- [31] HUANG W, WEI P, ZHANG M, et al. HRIPCB: A challenging dataset for PCB defects detection and classification [J]. Journal of Engineering, 2020, 13: 303-309.
- [32] HE F, TANG S, HUANG X, et al. A real-time PCB defect detector based on supervised and semi-supervised learning [C]. Proceedings, 28th European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning, 2020: 527-532.
- [33] PCBA缺陷检测项目数据集 [Z/OL]. (2021-11-11) [2024-01-24]. <https://aistudio.baidu.com/dataset-detail/179997/0>. PCBA defect detection project dataset [Z/OL]. (2021-11-11) [2024-01-24]. <https://>

aistudio.baidu.com/datasetdetail/179997/0.

- [34] 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.
- [35] RDEMON J, FARHADI A. YOLOv3: An incremental improvement [J]. *ArXiv Preprint*, 2018, ArXiv: 1804.02767.
- [36] 卞佰成, 陈田, 吴入军, 等. 基于改进 YOLOv3 的印刷电路板缺陷检测算法 [J]. *浙江大学学报(工学版)*, 2023, 57(4): 735-743.
- BIAN B CH, CHEN T, WU R J, et al. Improved YOLOv3-based defect detection algorithm for printed circuit board [J]. *Journal of Zhejiang University (Engineering Science)*, 2023, 57(4): 735-743.
- [37] DING J W, LI W, PEI L, et al. Sw-YOLOX: An anchor-free detector based transformer for sea surface object detection [J]. *Expert Systems with Applications*, 2023, 217: 119560.
- [38] 谢翔, 肖金球, 汪俞成, 等. 基于改进 YOLOv5s 的 Deep PCB 缺陷检测算法研究 [J]. *微电子学与计算机*, 2023, 40(7): 1-9.
- XIE X, XIAO J Q, WANG Y CH, et al. Deep PCB defect detection based on improved YOLOv5s algorithm [J]. *Microelectronics & Computer*, 2023, 40(7): 1-9.

作者简介



陈俊英 (通信作者), 2001 年于西安建筑科技大学获得学士学位, 分别于 2004 年和 2010 年于西安交通大学获得硕士学位和博士学位, 现为西安建筑科技大学信息与控制工程学院副教授, 主要研究方向为计算机视觉及机器学习。

E-mail: chenjy@xauat.edu.cn

Chen Junying (Corresponding author) received B. Sc. degree from Xi'an University of Architecture and Technology in 2001, M. Sc. degree from Xi'an Jiaotong University in 2004, and Ph. D. degree from the same university in 2010. She is currently an associate professor at the College of Information and Control Engineering, Xi'an University of Architecture and Technology. Her main research interests include computer vision and machine learning.



李朝阳, 2022 年于西安建筑科技大学获得学士学位, 现为西安建筑科技大学信息与控制工程学院硕士研究生, 主要研究方向为计算机视觉和缺陷检测。

E-mail: nicholas@xauat.edu.cn

Li Zhaoyang received his B. Sc. degree from Xi'an University of Architecture and Technology in 2022. He is currently a M. Sc. candidate in the School of Information and Control Engineering at Xi'an University of Architecture and Technology. His main research interests include computer vision and defect detection.