

基于改进反向蒸馏网络的电子元器件表面缺陷检测

李少杉^{1,2}, 成建宏^{2,3}, 张 钊^{2,3}, 金 龙¹

(1. 东南大学电气工程学院,南京 210096;2. 贵州航天计量测试技术研究所,贵阳 550009;3. 贵州省基础零部件及装备质量检测与评价全省重点实验室,贵阳 550009)

摘要:针对电子元器件表面缺陷类型复杂、尺寸微小难检测问题,本文提出了一种基于改进反向蒸馏网络的电子元器件表面缺陷检测模型。首先模型采用无监督学习的方法,有效降低了对大量标注数据的依赖;其次,通过运用反向蒸馏网络架构,改变了传统蒸馏网络中教师模型单向指导学生模型的模式,提高了模型在表面缺陷检测任务中的适应性;然后在反向蒸馏网络模型的学生解码器中,引入了感受野注意力卷积模块,以增强模型对微小缺陷的检测能力;最后采用余弦相似度作为损失函数训练学生网络和瓶颈模块。本文采用自建的电子元器件表面缺陷检测数据集进行实验,结果显示本文模型显著提高了检测精度,在接受者操作特征曲线下面积(Area under the receiver operating characteristic curve, AUROC)图像级指数和像素级指数上分别达到了86.7%和89.1%,在每区域重叠率曲线下面积(Area under the per-region-overlap curve, AUPRO)像素级指数上达到了69.4%。

关键词:电子元器件;表面缺陷检测;卷积神经网络;反向蒸馏网络;感受野注意力卷积

中图分类号: TP391.4 文献标志码:A

Surface Defect Detection of Electronic Components Based on Improved Reverse Distillation Network

LI Shaoshan^{1,2}, CHENG Jianhong^{2,3}, ZHANG Zhao^{2,3}, JIN Long¹

(1. School of Electrical Engineering, Southeast University, Nanjing 210096, China;2. Institute of Guizhou Aerospace Measuring and Testing Technology, Guiyang 550009, China;3. Guizhou Provincial Key Laboratory of Quality Inspection and Evaluation for Basic Components and Equipment, Guiyang 550009, China)

Abstract: Addressing the issues of complex surface defect types and minute defect sizes in electronic components that are difficult to detect, this paper proposes an electronic component surface defect detection model based on an improved reverse distillation network. Firstly, the model adopts an unsupervised learning approach, effectively mitigating the dependency on a large amount of annotated data. Secondly, by employing a reverse distillation network architecture, it alters the traditional mode of one-way guidance from the teacher model to the student model in distillation networks, thereby enhancing the model's adaptability in surface defect detection tasks. Furthermore, a receptive field attention convolutional module is introduced into the student decoder of the reverse distillation network to bolster the model's ability to detect minute defects. Finally, the cosine similarity is utilized as the loss function to train the student network and the bottleneck module. Experimental results using a self-constructed dataset for electronic

基金项目:黔科合平台项目(ZSYS[2025]003);黔科合支撑项目([2025]一般 022)。

收稿日期:2024-12-30;修订日期:2025-06-06

component surface defect detection demonstrate significant improvements in detection accuracy, achieving 86.7% and 89.1% in the area under the receiver operating characteristic curve (AUROC) at the image and pixel levels, respectively, and 69.4% in area under the pre-region-overlap curve (AUPRO) at the pixel level.

Key words: electronic components; surface defect detection; convolutional neural network (CNN); reverse distillation network; receptive field attention convolution

引 言

电子元器件种类繁多,其表面缺陷复杂多变。这些缺陷可能在制造、运输及使用过程中的任一环节产生,且可能出现在电子元器件表面封装的各种材质上,包括但不限于金属、塑料、陶瓷以及玻璃等,其中既有微小的斑点状附着物在电子元器件表面,也有长条状的划痕、大面积的腐蚀和凹坑等。此外,金属引脚的缺失或弯曲,以及由于安装不当所引发的结构性问题等,均属于电子元器件表面缺陷的范畴。当前,电子元器件的表面缺陷检测主要依赖于目视检查,但此方法极度依赖技术人员的经验,具有很强的主观性和较低的可靠性。尽管目前已有部分基于传统图像处理技术进行电子元器件表面缺陷检测的方法,但这些方法通常仅限于对特定型号电子元器件中具有一定规律的表面缺陷进行检测。鉴于电子元器件表面缺陷的复杂性,传统的图像处理方法难以实现对各型号电子元器件表面缺陷的全面且有效的检测。研究人员提出了一些基于全监督设置的深度学习算法进行表面缺陷检测的方法,比如唐小煜等^[1]提出基于改进的U-Net和YOLOv5进行绝缘子表面缺陷检测的方法,郝亮等^[2]提出基于改进级联R-CNN的表面缺陷检测方法,这些方法都高度依赖于覆盖全部种类的表面缺陷数据集。但是电子元器件表面缺陷复杂多变,且缺陷样本稀缺,难以构建完备的全监督设置的表面缺陷检测数据集,因而仅依靠正常样本进行训练的、基于无监督设置的深度学习算法更适合覆盖全部类别的电子元器件表面缺陷检测。

目前无监督设置的深度学习表面缺陷检测方法基本可以分为两类:基于特征嵌入的方法和基于重构的方法。基于特征嵌入的方法通过预训练编码器提取正常图片的特征来进行对正常特征的学习,以及和异常特征的对比从而实现表面缺陷检测。基于内存库的方法是基于特征嵌入方法的典型,其通过在大量的正常样本中进行特征提取,建立一个包含大量正常特征的内存库,在检测时提取输入图像的特征并与内存库中的正常特征进行比较从而实现表面缺陷检测。这种方法的代表有2022年提出的PatchCore^[3]和2023年提出的PNI^[4]。但是这样的方法对内存的消耗较大,且过于依赖正常样本,在推理过程中容易出现对正常特征的误判。近年来也有基于标准化流的CS-Flow方法^[5],和基于其改进而来的ValidFlow^[6],但是这些基于规范化流的方法也过于依赖正常特征。基于重构的方法是通过自训练编码器和解码器试图将输入的异常图像重建为正常图像,通过重建图像和输入图像的比较来实现缺陷检测。在基于重构的无监督表面缺陷检测方法中,扩散模型是近期的研究热点,DDPM^[7]和LDMs^[8]是扩散模型的代表。扩散模型用于缺陷检测的原理是对输入图像进行扩散处理,然后以正常图像作为引导进行去噪重建。在此过程中,缺陷区域被视为噪声而不会被重构。通过比较重建后的图像与原始图像,可以实现缺陷的检测。DiAD^[9]、DiffusionAD^[10]及DiffAD^[11]等方法展示了其在表面缺陷检测领域的潜力,但架构复杂、计算资源消耗大是其明显缺点。此外还有研究者提出了基于零样本、少样本的表面缺陷检测模型,如Winclip^[12]和Anomalyclip^[13],这类模型基于大型预训练视觉语言模型(如CLIP^[14]),通过学习与对象无关的文本提示,使得模型能够捕获图像中的一般正常性和异常性,这类模型在数据稀缺和快速部署的情况下具有明显优势,但是其缺陷检测的精确度不佳。

2022年,Deng等^[15]提出的反向蒸馏网络则基于重构原理,不直接将图像输入学生网络,而是先让教师网络提取特征,经瓶颈模块处理后将特征输入学生网络,由学生网络进行特征解码重建,克服了传统蒸馏网络因师生模型结构相似、数据流相似而导致的过度泛化问题。目前已有研究在反向蒸馏网络的基础上进行了改进,如在其基础上增加投影层以降低瓶颈模块异常特征输出的RD++模型^[16]、运用特征匹配技术的ADshift模型^[17]等。

本文针对电子元器件种类繁多,其表面缺陷复杂多变的特点,提出了一种改进反向蒸馏网络模型。该方法采用无监督设置,有效减少了数据标注的成本;采用反向蒸馏架构,避免了传统蒸馏网络因师生网络结构过于相似而产生的特征差异消失导致缺陷检测失败的问题;并在其学生解码器每层的残差块中增加了一个感受野注意力卷积(Receptive-field attention convolution, RFACConv)模块^[18],以增强其对正常特征和异常特征重建的差异,从而增强模型对电子元器件表面微小和复杂缺陷的检测能力,为电子元器件表面缺陷检测提供一种更加高效、准确的解决方案。通过实验结果的验证,本文提出的方法在自建的电子元器件表面缺陷检测数据集上显著提高了检测精度,在接受者操作特征曲线下面积(Area under the receiver operating characteristic curve, AUROC)图像级指数和像素级指数上分别达到了86.7%和89.1%,在每区域重叠率曲线下面积(Area under the pre-region-overlap curve, AUPRO)像素级指数上达到了69.4%,展现出其在实际应用中的潜力和优势。

1 本文算法

1.1 传统蒸馏网络和反向蒸馏网络

传统的知识蒸馏网络如图1(a)所示,采用教师编码器-学生编码器的架构,其中学生模型与教师模型的结构相似,并且两者都以原始图像作为输入。其中教师编码器采用预训练的网络并在训练阶段不更新权重,学生模型采用一个和教师编码器相似的网络结构。在训练过程中,教师模型引导学生网络提取图像特征使学生模型在接收正常特征时能够产生与教师模型相近的表示;在推理阶段,当输入包含异常特征的图像时,学生模型的特征表示会与教师模型产生差异,从而实现缺陷检测^[19-21]。

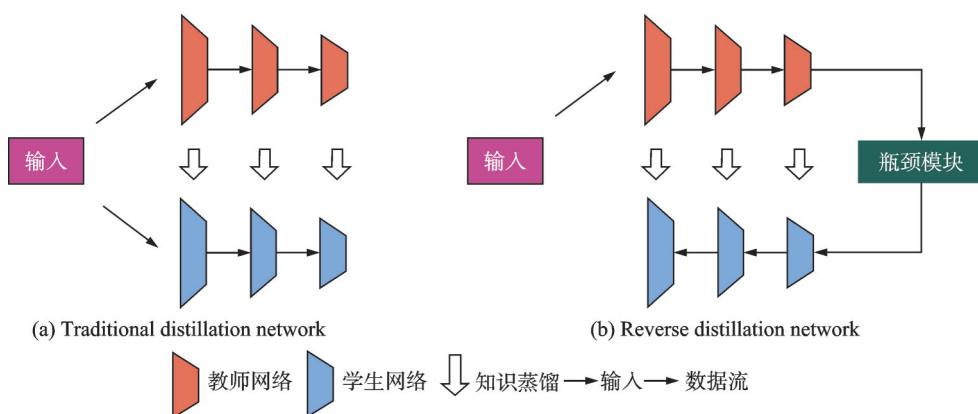


图1 传统-反向蒸馏网络对比图

Fig.1 Comparison of traditional and reverse distillation networks

具体来说,在训练阶段,教师编码器接收训练集中的正常图片 $D_{\text{train}} = \{x_1, x_2, \dots, x_n\}$,并从中提取多尺度特征 $f_T^k = E_T^k(x)$, k 为其对应的层数。这些特征随后与学生编码器从相同输入中提取的多尺度特征 $f_S^k = E_S^k(x)$ 进行对应层的比较。这种比较机制旨在引导学生编码器在接收正常图片时产生与教

师编码器相似的特征表示。如式(1)所示,损失函数设计为不同对应层相似度损失计算得到的异常得分图平均值的总和,其中相似度可以选择余弦相似度、 L_2 相似度、 L_1 相似度等。对于每一层,其损失函数计算为该层每个位置损失函数的平均值,其中 $x_i \in \mathbb{R}^{W \times H \times C}$ 为输入的训练集图片, W, H 为图像的长和宽,单位为像素, C 为图像的图层数。最后的总损失 L_{KD} 为各层损失值之和,表达式为

$$L_{KD} = \sum_{k=1}^K \left\{ \frac{1}{W_k H_k} \sum_{w=1}^{W_k} \sum_{h=1}^{H_k} (L_{FeaD}(f_T^k(w, h), f_S^k(w, h))) \right\} \quad (1)$$

式中: $f_S^k(w, h)$ 为学生解码器 f_S^k 第 k 层特征图在像素坐标 (w, h) 处的特征值; $f_T^k(w, h)$ 为教师编码器 f_T^k 第 k 层特征图在 (w, h) 处的特征值; W_k, H_k 分别为第 k 层特征图的宽和高; L_{FeaD} 为相似度损失函数; L_{KD} 为逐层计算对应特征图所有像素点损失值的均值,并将这些均值累计相加而得; K 为总层数。

在电子元器件表面缺陷检测任务中,传统蒸馏框架因学生模型与教师模型架构高度耦合且数据流高度相似,易引发模型训练异常,致使学生解码器在特征重建时出现过度泛化现象。具体表现为其重构特征与教师的编码器提取的原始特征间差异逐渐消失,导致模型对缺陷区域的鉴别能力退化。由于电子元器件表面缺陷(如划痕、焊点等)通常与正常区域边界模糊,且背景纹理复杂,这种特征同质化问题会进一步加剧检测失效风险。尽管采用轻量化学生网络可通过降低模型复杂度缓解上述问题,但其较弱的特征表达能力限制了图像表面特征的深层次学习,反而牺牲了检测精度。

反向蒸馏网络的结构(如图1(b)所示)由3个模块组成:1个固定的预训练教师编码器、1个用于多尺度特征融合和去除冗余特征的瓶颈模块和1个学生解码器。模型不直接将图像输入学生网络,而是先由教师编码器提取多尺度特征,再将教师网络提取的多尺度特征经瓶颈模块处理后输入学生网络。具体来说,在训练阶段,同样由教师编码器接收训练集中的正常图片 $D_{train} = \{x_1, x_2, \dots, x_n\}$,将教师编码器提取的多尺度特征 $f_T^k = E_T^k(x)$ 输入瓶颈模块进行多尺度特征融合和去冗余操作,将瓶颈模块的输出 ϕ 输入学生解码器得到多尺度特征 $f_S^k = D_S^k(\phi)$,同样如式(1),进行蒸馏损失的计算。反向蒸馏网络改变了传统蒸馏网络中教师模型单向指导学生模型的模式,通过教师特征前置提取与瓶颈约束机制,既保留复杂表面纹理的细节表征能力,又抑制学生网络的过度泛化倾向,特别适用于元件表面缺陷与正常区域边界模糊(如划痕、凹坑等)的检测场景,有效提升了模型对弱缺陷特征的鉴别能力。

在推理阶段,对输入的每张测试图像,通过相似度损失函数在每个对应层 k 上生成异常得分图 M^k ,得到不同尺度的异常得分图,最后将其上采样至输入图像大小,并累加得到最终的异常得分图 S_A ,表达式为

$$S_A = \sum_{k=1}^K \text{Upsample } M^k \quad (2)$$

1.2 基于感受野注意力增强的反向蒸馏网络

传统的反向蒸馏网络,其学生网络采用了简单的和预训练教师编码器对称的结构,本文发现其学生解码器的特征重构能力没有得到充分的发掘,而且在这种相似的结构中,部分与正常区域差别不大的表面缺陷区域仍然可能得到一定程度的重构,从而降低了模型在缺陷识别上的精确度,这一点在电子元器件表面缺陷的检测中尤其明显。本文在传统的反向蒸馏网络模型的基础上进行改进,将RFACConv融入学生解码器,动态优化特征提取,显著增强了对电子元器件表面复杂缺陷的辨识能力。

本文提出了基于改进反向蒸馏网络的电子元器件表面缺陷检测模型。如图2所示,本文模型以反向蒸馏网络为主干架构,由1个教师编码器、1个瓶颈模块和1个感受野注意力学生解码器组成。图中E1~E3为教师编码器的不同层,分别由3、4、6个ResBlock残差块组成;D1~D3为感受野注意力学生解码器的不同层,由与教师编码器每层对应的3、4、6个感受野注意力残差块组成;M1~M3为师生模型中对应层特征图比较计算得到的异常图,每层对应的异常图大小分别为 64×64 、 32×32 和 16×16 ,最终上采样至大小为 256×256 像素的输入图像。

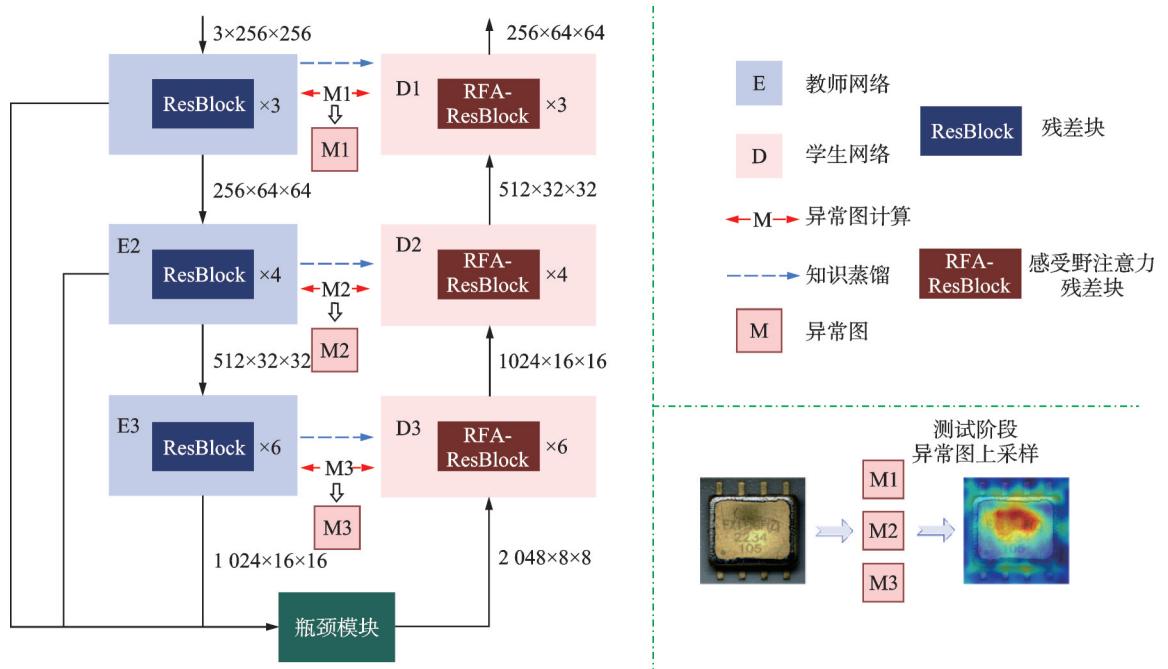


图2 改进的反向蒸馏网络模型架构图

Fig.2 Architecture diagram of improved reverse distillation network model

本文采用一个在 ImageNet 数据集上预训练的 Wide-ResNet 50 网络作为教师编码器。Wide-ResNet 网络通过增加每个残差块中的卷积核数量(即通道数)来增宽网络,而不是简单地增加网络深度,使得其能够在保持或降低参数总数的同时,通过更宽的特征映射来捕获电子元器件表面的细节特征;运用 Dropout 层在训练过程中随机丢弃网络中的一部分神经元(及其连接)来减少神经元之间的复杂共适应性,使网络学习到更加鲁棒的特征表示,能有效应对电子元器件表面缺陷检测中经常出现的光照波动、背景环境差异等干扰因素。

模型提取 Wide-ResNet 50 网络前 3 个残差块输出的多尺度特征,假设输入电子元器件表面图像样本为 $I \in D_s$, D_s 为训练集,教师编码器为 E^k , k 为编码器层数,提取的特征 f_E^k 为

$$f_E^k = E^k(I) \quad k \in \{1, 2, 3\} \quad (3)$$

之后提取到的 3 层多尺度特征被输入到瓶颈模块进行多尺度特征的融合和去冗余操作。本文用 φ 表示输入图像 I 经教师编码器提取和瓶颈模块处理后输出的特征。如图 3 所示,瓶颈模块分为多尺度特征融合(Multi-scale feature fusion, MFF)模块和单类别嵌入(One-class embedding, OCE)特征去冗余模块。特征融合模块用 3×3 的卷积将前两层特征图的尺寸降至第 3 层特征图的尺寸并将其拼接,特征去冗余模块由 3 层瓶颈型 ResBlock 残差块构成,对模型所提取的多尺度特征进行去冗余操作。多尺度特征融合机制配合特征去冗余模块,去除多尺度特征中的冗余信息的同时增强了不同尺度缺陷的识别能力,尤其适合对电子元器件表面常出现的不规则形状缺陷的检测。

最后本文将瓶颈模块处理过的图像特征输入感受野注意力学生解码器,进行电子元器件表面图片多尺度特征的重建。感受野注意力学生解码器采用的是和预训练教师编码器对称的网络架构,使得对应位置的图像特征维度一致,比如当教师编码器使用核大小为 1、步长为 2 的卷积操作时,学生解码器则使用一个核大小为 3 的感受野注意力卷积对每层的输入特征进行处理后,进行核大小为 2、步长同样为

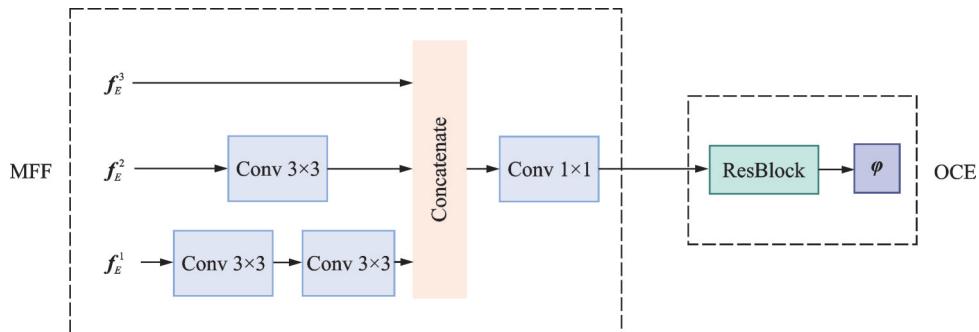


图3 瓶颈模块

Fig.3 Bottleneck module

2的反卷积(或称转置卷积)操作,提取的输出特征 f_D^k 为

$$f_D^k = E^k(\varphi) \quad (4)$$

1.2.1 感受野注意力学生解码器

在电子元器件表面缺陷检测中,缺陷形态多样且表面材质复杂,检测难度显著提升。部分缺陷如微小污渍或气泡,在元器件表面占据极小的面积比例,易被忽略;而划痕类缺陷则可能呈现为细长且不规则的线条,横跨元器件多个部分,甚至延伸至整个表面。此外,金属引脚的弯曲或缺失、部件的细微裂痕等结构性缺陷,也对检测精度提出了更高要求。因此本文在学生解码器中引入了感受野注意力卷积层,增强了模型对微小和复杂缺陷的检测能力。感受野注意力学生解码器采用与Wide-ResNet 50 的前3层对称的结构,其每层分别由3、4、6个感受野注意力ResBlock残差块堆叠而成,其中每层的第一个ResBlock残差块进行上采样操作,感受野注意力ResBlock残差块在每层第一个ResBlock残差块中的 2×2 反卷积前增加了1个 3×3 的感受野注意力模块,在其余ResBlock残差块的 3×3 卷积前增加了一个 3×3 的感受野注意力模块,以增强学生解码器在输入缺陷图像时输出的正常特征和异常特征的差异。

改进的RFA-ResBlock有两种情况,如图4左右两图所示,区别是有无Upsample上采样层。RFA-ResBlock是一种优化的瓶颈型残差块,其核心改进在于上采样ResBlock块的反向 2×2 卷积层前增加一个 3×3 的RFACconv层,在一般ResBlock的 3×3 卷积前也增加了一个 3×3 的RFACconv层。在以往学生解码器采用的瓶颈型残差块中,数据首先通过一个 1×1 的卷积层降低维度,接着采用一个 2×2 的反卷积层进行上采样或者采用一个 3×3 的卷积模块用以处理特征信息,最后再通过另一个 1×1 的卷积层恢复维度。RFA-ResBlock保留了这一基本框架,但在中间的 2×2 反卷积层或 3×3 卷积前增加一个 3×3 的RFACconv层对特征进行处理,旨在利用感受野注意力机制的特殊性质来增强对输入图像特征的提取能力,提升网络的整体性能。本文的替换方法既保留了瓶颈型残差块的高效性,又通过引入RFACconv来进一步优化处理流程。

此外本文对解码器的归一化方法进行微调,将解码器中的归一化方法改为批处理信道归一化(Batch channel normal-

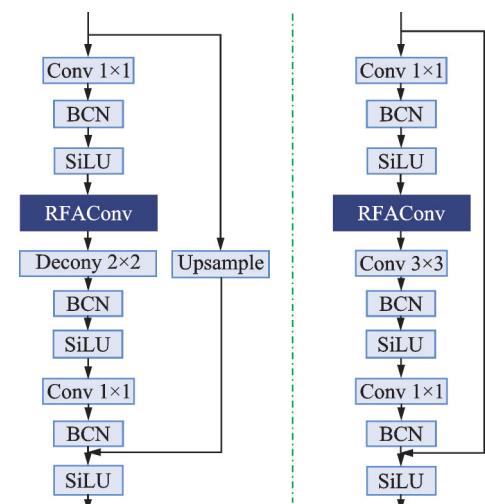


图4 RFA-ResBlock模块

Fig.4 RFA-ResBlock module

ization, BCN)方法^[22],将激活函数改为 SiLU(Sigmoid linear unit)激活函数,以提升感受野注意力学生解码器的重建能力和模型的整体性能。RFACConv 机制如图 5 所示。其机制的核心思想是通过动态生成卷积核权重来增强特征提取能力。在初始化阶段,它首先通过平均池化和分组卷积操作生成权重。随后,这些权重通过 Softmax 进行归一化,并与通过分组卷积、批量归一化和 ReLU 激活函数生成的特征图进行逐元素相乘,以实现动态的特征加权。之后特征图经过重新排列,扩大感受野,并最终通过一个标准的卷积、批量归一化和 ReLU 激活函数进行处理,得到最终的输出特征图。整个过程旨在通过动态调整卷积核权重来提高模型的表达能力和适应性,其表达式如式(5)所示。本文对其归一化方法和激活函数进行调整采用 BCN 和 SiLU 激活函数。

$$F = \text{Softmax}(g^{1 \times 1}(\text{AvgPool}(\mathbf{X}))) \times \text{SiLU}\left(\text{Norm}(g^{k \times k}(\mathbf{X}))\right) \quad (5)$$

式中: $g^{i \times i}$ 表示卷积核大小为 $i \times i$ 的分组卷积, $i=1, 2, \dots, k$; AvgPool 为平均池化; SiLU 为激活函数; Norm 为批通道归一化。

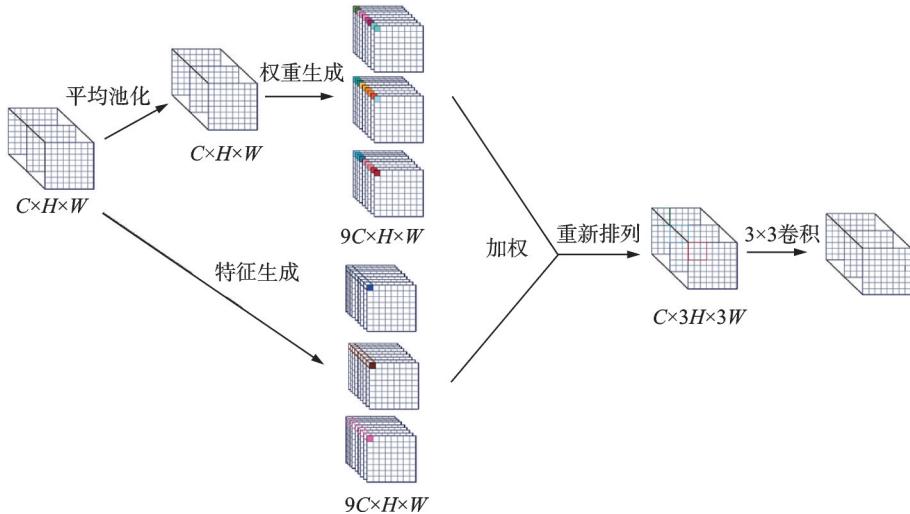


图 5 RFACConv(3×3 卷积)机制示意图

Fig.5 Mechanism diagram of RFACConv (3×3 convolution)

在以往的注意力机制中,通道注意力机制^[23]仅聚焦于通道信息,忽视了位置信息的重要性。CBAM^[24]尝试通过通道上的全局池化来融入位置信息,但这种方式仅能捕获局部信息,无法获取远程相关信息,而空间注意力模块^[25]只解决了 1×1 卷积模块的参数共享问题。RFACConv 解决了卷积核参数共享的问题,允许网络为每个感受野生成特定的权重,从而更灵活地适应不同区域的特征。

1.2.2 损失函数和异常得分

在模型的训练过程中,不对预训练教师解码器进行权重更新,只训练瓶颈模块和感受野注意力学生解码器。采用余弦相似度函数进行计算,师生网络每层对应特征图损失函数的计算过程为

$$\mathbf{M}^k(w, h) = 1 - \frac{\mathbf{f}_D^k(w, h) \cdot (\mathbf{f}_E^k(w, h))^T}{\|\mathbf{f}_D^k(w, h)\| \|\mathbf{f}_E^k(w, h)\|} \quad (6)$$

式中: $\mathbf{f}_E^k(w, h)$ 为本文算法中教师编码器第 k 层特征图在像素坐标 (w, h) 处的特征值; $\mathbf{f}_D^k(w, h)$ 为本文算法学生解码器第 k 层特征图在像素坐标 (w, h) 处的特征值; $\mathbf{M}^k(w, h)$ 为第 k 层相似度损失图在像素坐标 (w, h) 处的相似度损失值。

对式(6)得到的每层相似度损失的平均值进行相加得到最终的损失函数,表达式为

$$L_{KD} = \sum_{k=1}^K \left\{ \frac{1}{W_k H_k} \sum_{w=1}^{W_k} \sum_{h=1}^{H_k} M^k(w, h) \right\} \quad (7)$$

训练过程通过最小化损失函数实现对学生网络的优化。在推理阶段,首要步骤是执行像素级别的异常评分,以实现对异常区域的精确检测。由于学生解码器仅接受正常特征的重建训练,当向模型输入异常特征时,学生解码器重建特征会与教师编码器提取的特征有较大的差距,进而生成与教师编码器不同的表示,根据式(7)对学生-教师模型中对应特征图进行余弦相似度计算,每一对特征图都会生成一张对应的异常分数图 M^k 。随后,将所有生成的异常分数图 M^k 通过双线性上采样方法 Ψ 放大至输入图像的大小,并进行累加,得到精细化的异常分数图 S_{AK} ,异常分数图上每个像素的值都代表其对应位置在原始图像中的异常程度,表达式为

$$S_{AK} = \sum_{k=1}^L \Psi(M^k) \quad (8)$$

具体来说,异常分数图 S_{AK} 的每个像素值反映了该位置在输入图像中可能存在的异常或偏差,值越高表示该区域的异常程度越大,该图能够清晰地展示出原始图像中各个区域的异常分布情况。为了进一步优化结果,对 S_{AK} 应用高斯滤波进行平滑处理。图像级异常分数取 S_{AK} 中的最大值,因为对 S_{AK} 取平均值的方法不适合于缺陷区域较小的图像的异常判别,缺陷区域较小的图像往往异常分数图的平均值会更小,这不符合表面缺陷检测的实际。

2 实验结果与分析

2.1 数据集

本文所采用的电子元器件表面缺陷检测数据集,采集于电子元器件检测现场,涵盖了丰富多样的电子元器件类别。鉴于电子元器件种类繁多且外观特征错综复杂,本文在构建数据集时采取了细致的分类策略,将同一封装的电子元器件表面图片归于一类,总计包含 1928 张图片。采用 Labelme 软件对缺陷图片进行了详尽的逐点标注工作,进而生成了对应的掩码图,以精确标识缺陷区域。将其分为训练集和测试集,训练集由正常样本构成,测试集由部分正常样本和全部缺陷样本构成,测试集中每张缺陷图像都有对应的掩码图。

数据集的构成如表 1 所示,本文将采集到的电子元器件表面图片按照封装方式分为 7 个类别。第 1 类的电子元器件为薄型四侧引脚扁平封装(Low-profile quad flat package,LQFP),具体类别有采用此封装的 D/A 转化器、微控制单元和数字信号处理器等;第 2 类的电子元器件为塑封方形扁平无引脚(Quad flat no-lead,QFN)封装,具体类别有采用此封装的 FPGA、网络芯片、DC/DC、射频芯片、降压芯片、音频放大器、微波开关和微波混频器等;第 3 类的电子元器件为小外形封装(Small outline package,SOP),具体类别有采用此封装的电压基准、收发器、存储器、电压转换器、温度传感器、运算放大器、接口电路、网络变压器、总线收发器、电压调整器、驱动器和逻辑控制器等;第 4 类为金属陶瓷扁平封装(Ceramic flat package,CFP),具体类别有采用此封装的运算放大器、时基、收发器、光电耦合器和比较器等;第 5 类为小外形晶体管(Small outline transistor,SOT)封装,具体类别有采用此封装的中小功率三极管、中小功率场效应管和中小功率二极管等;第 6 类为小外形二极管(Small outline diode,SOD)封装,具体类别有采用此封装的中小功率二极管等;第 7 类为表面贴装器件(Surface-mounted device,SMD)封装,具体类别有采用此封装的稳压器等。图 6 显示了数据集中各个类型的电子元器件的正常和缺陷图片,上行为正常图片,下行为缺陷图片。

表 1 电子元器件表面缺陷检测数据集构成

Table 1 Composition of the electronic component surface defect detection dataset

数据集类别	LQFP	QFN	SOP	CFP	SOT	SOD	SMD
训练集样本数量	258	288	434	107	200	168	69
测试集正常样本数量	21	29	31	13	21	20	13
测试集缺陷样本数量	31	27	31	91	31	26	19

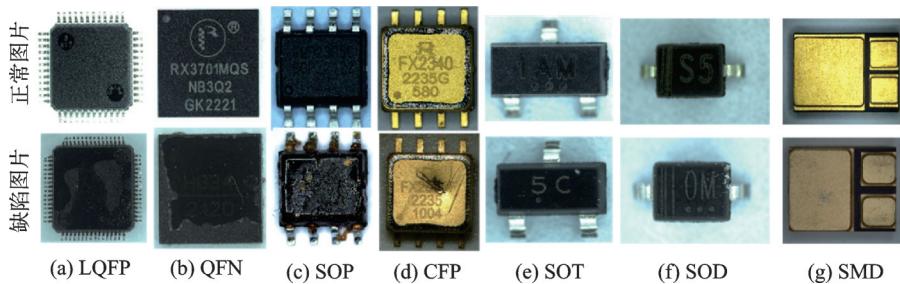


图 6 电子元器件图像示例

Fig.6 Examples of electronic component images

2.2 实验环境

本实验所采用的操作系统为 Linux Ubuntu 操作系统, 使用 Pytorch 2.1.0 作为深度学习框架, CUDA 版本为 12.1。CPU 为 Intel(R) Xeon(R) Gold 6330 CPU @ 2.00 GHz, GPU 为 NVIDIA A100-PCIE-40 GB。

在数据预处理过程中, 将输入图像尺寸统一转换成 256 像素 \times 256 像素, 一次对一个类别进行缺陷检测。在实验中, 本文采用 Wide ResNet-50 的前 3 个残差块作为模型的预训练教师网络。使用一阶矩衰减率为 0.5、二阶矩衰减率为 0.999 的 Adam 优化器。学习率设置为 0.005, Epoch 设置为 300, Batch size 设置为 16。使用 $\sigma = 4$ 的高斯滤波器来平滑异常得分图。

2.3 评价指标

本文采用的评价指标为图像级 AUROC, 像素级 AUROC 和 AUPRO 指数。AUROC 是衡量模型性能的指标, 值域为 0.5 到 1, 值越高表示模型性能越好, 具体表现为真正率-假正率曲线下的面积。图像级 AUROC 评估模型在图像层面区分有、无缺陷的能力, 高值表明模型在此方面准确性高。像素级 AUROC 则侧重于每个像素点的分类准确性, 精确地反映模型细节识别能力。AUPRO 是一个用于评估异常定位性能的指标, 特别关注于每个独立的缺陷区域, 而不是整个图像的像素级分类。在表面缺陷检测任务中, AUPRO 指数不仅评估模型准确识别缺陷的能力, 还强调减少误报的重要性。AUPRO 值高意味着模型能够在保持高精确率的同时, 实现较高的召回率, 从而更全面、准确地评估模型的性能。

2.4 对比实验

为了评估本文所提方法的性能, 本文将图像级 AUROC、像素级 AUROC 和 AUPRO 指数作为评价指标, 在电子元器件表面缺陷检测数据集上与其他先进的表面缺陷检测方法: RD4AD^[15]、RD++^[16]、Patchcore^[3]、DeSTSeg^[26]、SimpleNet^[27] 和 ADShift^[17] 模型进行了对比测试。表 2 为对比实验得到的图像级 AUROC 指数, 表 3 为对比实验得到的像素级 AUROC 指数, 表 4 为对比实验得到的像素级 AUPRO 指数。由于本文建立的电子元器件表面缺陷检测数据集以封装类别进行分类, 在每个类别的训练集中并不一定涵盖所有的电子元器件, 因此测试集中可能出现与训练集特征差异较大的电子元器件。

表2 电子元器件表面缺陷检测数据集上的图像AUROC对比结果

Table 2 Comparison results of image AUROC on the electronic component surface defect detection dataset %

检测对象	RD4AD ^[15]	RD++ ^[16]	Patchcore ^[3]	DeSTSeg ^[26]	SimpleNet ^[27]	ADShift ^[17]	本文模型
LQFP	76.3	81.2	57	70.7	72.8	58.3	88.2
QFN	75.4	81.6	81.5	77.6	82.8	68.7	80.5
SOP	99.2	99.7	95.2	92.0	95.4	96.8	98.3
CFP	97.9	92.9	99.0	87.8	99.8	95.2	95.3
SOT	76.2	77.6	59.1	52.1	66.5	60.2	80.8
SOD	88.3	86.5	66.5	48.9	78.9	83.7	85.2
SMD	76.1	82.2	60.7	66.8	81.8	69.2	78.9
Mean	84.2	85.9	74.1	68.0	82.6	76.0	86.7

表3 电子元器件表面缺陷检测数据集上的像素AUROC对比结果

Table 3 Comparison results of pixel-wise AUROC on the electronic component surface defect detection dataset %

检测对象	RD4AD ^[15]	RD++ ^[16]	Patchcore ^[3]	DeSTSeg ^[26]	SimpleNet ^[27]	ADShift ^[17]	本文模型
LQFP	97.5	96.9	96.8	96.4	92.8	95.7	96.4
QFN	84.2	85.8	80.3	84.2	78.0	82.4	86.8
SOP	93.8	93.7	92.8	92.7	88.6	91.1	94.6
CFP	83.4	82.3	71.3	73.8	70.9	87.6	85.8
SOT	81.4	76.5	68.8	75.1	69.8	78.6	81.8
SOD	90.0	90.7	86.6	87.8	73.3	85.1	89.7
SMD	88.1	82.0	83.6	84.3	80.2	87.2	88.4
Mean	88.3	86.8	82.9	84.9	79.1	86.8	89.1

表4 电子元器件表面缺陷检测数据集上的像素AUPRO对比结果

Table 4 Comparison results of pixel-wise AUPRO on the electronic component surface defect detection dataset %

检测对象	RD4AD ^[15]	RD++ ^[16]	DeSTSeg ^[26]	SimpleNet ^[27]	ADShift ^[17]	本文模型
LQFP	83.1	81.5	70.7	47.5	77.9	76.4
QFN	75.2	68.5	82.4	42.1	71.6	73.1
SOP	63.4	62.9	74.5	54.2	59.8	64.5
CFP	46.8	39.4	31.0	21.2	53.4	48.7
SOT	77.2	75.4	76.2	43.3	74.1	77.6
SOD	69.4	75.5	60.3	37.1	63.3	73.3
SMD	70.9	60.7	60.3	55.9	69.4	72.1
Mean	69.4	66.3	65.1	43.0	67.1	69.4

这种特征差异导致模型在缺陷检测时难以准确识别正常特征,从而增加了对正常特征的误判率,这就导致在缺陷检测时模型的AUPRO指数往往低于像素级AUROC指数;而本文采用异常图的最大值作为图像级异常得分,如果对正常图片中的像素级误判较多就容易导致较高的图像级异常得分,使得在图像级异常识别中出现误判,即将正常图片误判为异常图片,进而导致图像级的AUROC值较低。电

子元器件表面缺陷数据集需要在日常的电子元器件表面缺陷检测中继续收集电子元器件表面图片样本,进一步完善数据集。

在对比的先进方法中,SimpleNet模型的特征提取器在提取某个位置的特征时会聚合其领域内的特征,增强了其在图像级缺陷识别时应对局部异常误判的鲁棒性,从而在部分类别上获得相对较高的图像级AUROC指数,但是这也会导致一定程度的空间信息损失,在缺陷检测时难以准确识别出图像中的具体异常区域,所以其像素级分数不佳。本文所提出的改进反向知识蒸馏网络模型在图像级AUROC指数、像素级AUROC指数和AUPRO指数上超越了其他先进模型,取得了86.7%的图像级AUROC、89.1%的像素级AUROC和69.4%的AUPRO指数。

图7展示了对电子元器件表面缺陷检测数据集中各个类别进行测试后得到的可视化缺陷热力图,其中第1,3行为经过预处理的256像素×256像素的输入图片,第2,4行为生成的可视化缺陷热力图。

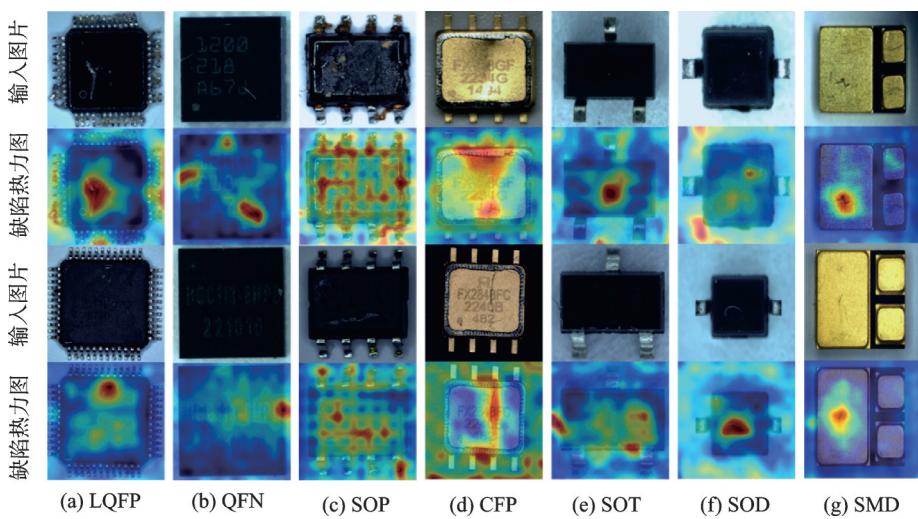


图7 可视化缺陷热力图示例

Fig.7 Defect heatmap visualization examples

图8展示了改进的反向蒸馏模型在7个不同类别的训练过程中,损失值(Loss)随训练轮次(Epoch)的变化趋势。随着训练的进行,损失值从较高的初始值逐步降低,最终稳定在一个较低的水平,这清晰地反映了本文模型在学习和优化过程中的进步。训练初期,损失值迅速下降,表明本文模型能够迅速适应并调整其参数以更好地拟合数据。随后,下降速度逐渐放缓,但本文模型仍持续改进,直至大约250轮训练后,损失值趋于稳定,这标志着本文模型已接近收敛状态,其性能也趋于稳定。图9展示了改进的反向蒸馏模型在各个类别上训练得到的ROC曲线。

2.5 消融实验

为了验证本文在瓶颈模块中采用的感受野注意力机制的有效性,本文进行了消融实验。首先用基础的反向蒸馏模型RD4AD^[15]对电子元器件表面缺陷数据集进行

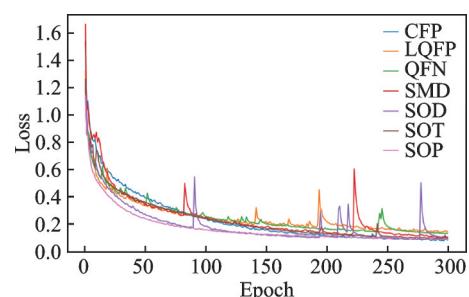


图8 损失值变化图

Fig.8 Loss value variation diagram

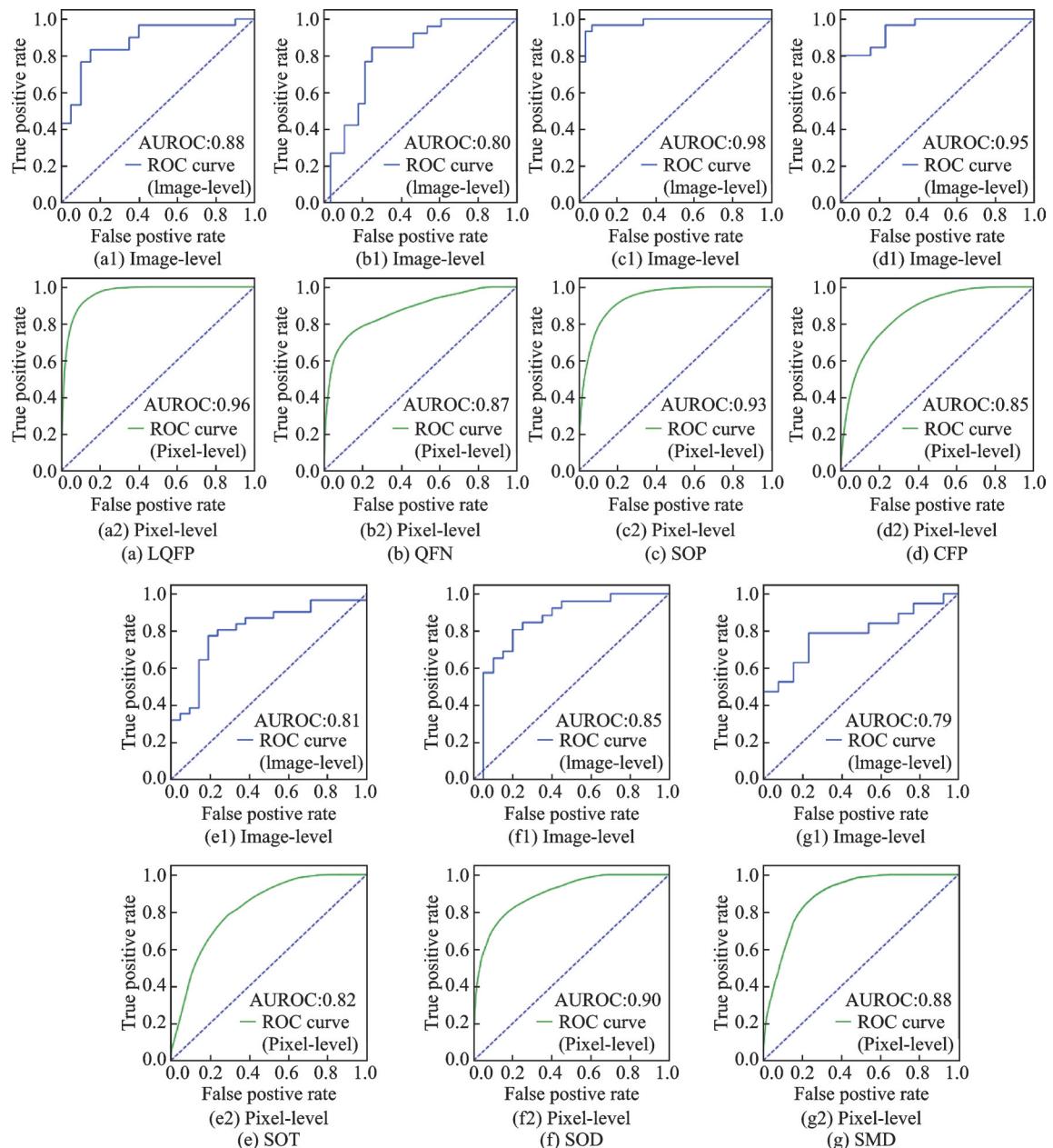


图9 本文模型在7个类别上训练得到的ROC曲线

Fig.9 ROC curves obtained by training the proposed model on seven categories

了测试,接着测试了加入感受野注意力卷积的情况,即与本文提出的改进反向蒸馏网络模型进行对比,结果如表5所示。由表5可见,相对于原始的反向知识蒸馏网络,本文提出的改进反向知识蒸馏模型在图像AUROC和像素AUROC分数上有了明显的提升,分别提升了2.5%和0.8%。图10展示了本文模型与RD4AD模型在电子元器件缺陷检测可视化图上的对比。通过对比分析,可以清晰地观察到本文模型在识别与正常区域相似的缺陷区域方面展现出了显著的优势。这一优势主要得益于本文模型中引入的感受野注意力机制,该机制有效地增强了模型对缺陷区域的检测能力,从而提高了整体检测的

准确性和鲁棒性。

本文还使用不同的预训练编码器 ResNet34、ResNet50 和 WideResNet50 进行了对比测试,结果如表 6 所示。由表 6 可见,越广越深的网络架构往往具有更强的表现能力,能够更有效地检测和识别异常,WideResNet50 网络在图像级和像素级的缺陷检测中明显优于 ResNet34 和 ResNet50 网络,故选择 WideResNet50 网络作为本文改进反向蒸馏网络的预训练编码器。

此外,本文将根据不同层的特征图生成的异常分数图 M_k 通过双线性上采样方法 Ψ 放大至输入图像的大小后,分别对 3 张异常分数图进行累加和累乘的融合方式得到最终的特征图,进行缺陷图像的检测,结果如表 7 所示。由表 7 可见,使用累加方法融合不同尺寸的异常分数图在像素级指标上展现出了显著的优势。这种方法通过线性累积不同尺度的异常信息,有效地保留了更多的细节特征,进而显著提升了缺陷检测的精确性。累加的方法在精确捕捉和表征各种尺度异常方面优于累乘的方法,使得其在对细节要求较高的电子元器件检测实际应用场景中具有优势,最后本文选择累加的方法进行多尺度异常分数图的融合。

3 结束语

本文针对电子元器件表面缺陷复杂多变的特点,提出了一种改进的反向蒸馏网络。该模型采用无监督设置,降低了数据集标注成本;采用的反向蒸馏网络架构解决了传统蒸馏网络模型中出现的特征差异消失问题;提出的感受野注意力学生解码器增大了正常特征和异常特征的差异,从而增强了模型对电子元器件表面微小和复杂缺陷的检测精度。实验结果表明,该模型在电子元器件表面缺陷检测数据集上表现优异,提高了检测精度。本文展示了该方法的创新性和实用性,模型在电子元器件表面缺陷检测中具有明显优势。目前针对电子元器件表面缺陷检测的研究并不多,下一步需要建立更加完备的电子元器件表面缺陷检测数据集,而且电子元器件表面缺陷复杂,对结构性缺陷和微小缺陷检测的精度提高问题仍然需要继续研究。

表 5 电子元器件表面缺陷检测数据集上的消融实验

Table 5 Ablation study on the electronic component surface defect detection dataset

指标	RD4AD	本文模型	%
图像 AUROC	84.2	86.7	
像素 AUROC	88.3	89.1	
像素 AUPRO	69.4	69.4	

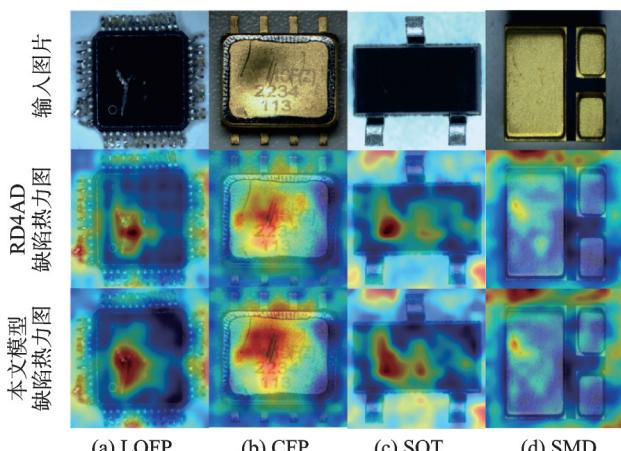


图 10 消融实验可视化图对比

Fig.10 Comparison of ablation study visualizations

表 6 不同预训练编码器的定量比较

Table 6 Quantitative comparison of different pre-trained encoders

指标	ResNet34	ResNet50	WideResNet50	%
图像 AUROC	85.5	83.5	86.7	
像素 AUROC	87.7	87.2	89.1	
像素 AUPRO	67.0	67.5	69.4	

表 7 不同多尺度异常分数图融合方法的比较

Table 7 Comparison of different multi-scale anomaly score map fusion methods

指标	累乘	累加	%
图像 AUROC	86.7	86.7	
像素 AUROC	88.9	89.1	
像素 AUPRO	69.1	69.4	

参考文献：

- [1] 唐小煜, 熊浩良, 黄锐珊, 等. 基于改进的 U-Net 和 YOLOv5 的绝缘子掩模获取与缺陷检测[J]. 数据采集与处理, 2021, 36(5): 1041-1049.
TANG Xiaoyu, XIONG Haoliang, HUANG Ruishan, et al. Insulator mask acquisition and defect detection based on improved U-Net and YOLOv5[J]. *Journal of Data Acquisition and Processing*, 2021, 36(5): 1041-1049.
- [2] 郝亮, 周诗洋, 莫允扬, 等. 基于改进级联 R-CNN 的钢材带状碳化物检测与分级[J]. 数据采集与处理, 2024, 39(5): 1228-1239.
HAO Liang, ZHOU Shiyang, MO Yunyang, et al. Detection and classification of banded carbide in steel based on improved cascade R-CNN[J]. *Journal of Data Acquisition and Processing*, 2024, 39(5): 1228-1239.
- [3] ROTH K, PEMULA L, ZEPEDA J, et al. Towards total recall in industrial anomaly detection[C]//Proceedings of 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). New Orleans, LA, USA: IEEE, 2022: 14298-14308.
- [4] BAE J, LEE J H, KIM S. PNI: Industrial anomaly detection using position and neighborhood information[C]//Proceedings of 2023 IEEE/CVF International Conference on Computer Vision (ICCV). Paris, France: IEEE, 2024: 6350-6360.
- [5] RUDOLPH M, WEHRBEIN T, ROSENHAHN B, et al. Fully convolutional cross-scale-flows for image-based defect detection[C]//Proceedings of 2022 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV). Waikoloa, HI, USA: IEEE, 2022: 1829-1838.
- [6] 张兰尧, 陈晓玲, 张达敏, 等. ValidFlow: 基于标准化流的无监督图像缺陷检测[J]. 数据采集与处理, 2023, 38(6): 1445-1457.
ZHANG Lanyao, CHEN Xiaoling, ZHANG Damin, et al. ValidFlow: Unsupervised image defect detection based on normalizing flows[J]. *Journal of Data Acquisition and Processing*, 2023, 38(6): 1445-1457.
- [7] HO J, JAIN A, ABBEEL P. Denoising diffusion probabilistic models[EB/OL]. (2020-06-15). <https://arxiv.org/abs/2006.11239>.
- [8] ROMBACH R, BLATTMANN A, LORENZ D, et al. High-resolution image synthesis with latent diffusion models[C]//Proceedings of 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). New Orleans, LA, USA: IEEE, 2022: 10674-10685.
- [9] HE H, HE H, ZHANG J, et al. A diffusion-based framework for multi-class anomaly detection[C]//Proceedings of the Thirty-Eighth AAAI Conference on Artificial Intelligence and the Thirty-Sixth Conference on Innovative Applications of Artificial Intelligence and Fourteenth Symposium on Educational Advances in Artificial Intelligence. [S.l.]: ACM, 2024: 8472-8480.
- [10] ZHANG H, WANG Z, ZENG D, et al. DiffusionAD: Norm-guided one-step denoising diffusion for anomaly detection[EB/OL]. (2023-03-18). <https://arxiv.org/abs/2303.08730>.
- [11] ZHANG X, LI N, LI J, et al. Unsupervised surface anomaly detection with diffusion probabilistic model[C]//Proceedings of 2023 IEEE/CVF International Conference on Computer Vision (ICCV). Paris, France: IEEE, 2024: 6759-6768.
- [12] JEONG J, ZOU Y, KIM T, et al. WinCLIP: Zero-/few-shot anomaly classification and segmentation[C]//Proceedings of 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Vancouver, BC, Canada: IEEE, 2023: 19606-19616.
- [13] ZHOU Q, PANG G, TIAN Y, et al. AnomalyCLIP: Object-agnostic prompt learning for zero-shot anomaly detection[EB/OL]. (2023-10-10). <https://arxiv.org/abs/2310.18961>.
- [14] RADFORD A, KIM J W, HALLACY C, et al. Learning transferable visual models from natural language supervision[C]//Proceedings of International Conference on Machine Learning. [S.l.]: [s.n.], 2021.
- [15] DENG H, LI X. Anomaly detection via reverse distillation from one-class embedding[C]//Proceedings of 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). New Orleans, LA, USA: IEEE, 2022: 9727-9736.
- [16] TIEN T D, NGUYEN A T, TRAN N H, et al. Revisiting reverse distillation for anomaly detection[C]//Proceedings of 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Vancouver, BC, Canada: IEEE, 2023:

24511-24520.

- [17] CAO T, ZHU J, PANG G. Anomaly detection under distribution shift[C]//Proceedings of 2023 IEEE/CVF International Conference on Computer Vision (ICCV). Paris, France: IEEE, 2024: 6488-6500.
- [18] ZHANG X, LIU C, YANG D, et al. RFACConv: Innovating spatial attention and standard convolutional operation[EB/OL]. (2023-04-09). <https://arxiv.org/abs/2304.03198>.
- [19] GONG D, LIU L, LE V, et al. Memorizing normality to detect anomaly: Memory-augmented deep autoencoder for unsupervised anomaly detection[C]//Proceedings of 2019 IEEE/CVF International Conference on Computer Vision (ICCV). Seoul, Korea: IEEE, 2020: 1705-1714.
- [20] SALEHI M, SADJADI N, BASELIZADEH S, et al. Multiresolution knowledge distillation for anomaly detection[C]// Proceedings of 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Nashville, TN, USA: IEEE, 2021: 14897-14907.
- [21] WANG G, HAN S, DING E, et al. Student-teacher feature pyramid matching for anomaly detection[EB/OL]. (2021-03-15). <https://arxiv.org/abs/2103.04257>.
- [22] KHALED A. BCN: Batch channel normalization for image classification[C]//Proceedings of Pattern Recognition. Cham: Springer, 2025: 295-308.
- [23] HU J, SHEN L, SUN G. Squeeze-and-excitation networks[C]// Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, UT, USA: IEEE, 2018: 7132-7141.
- [24] WOO S, PARK J, LEE J Y, et al. CBAM: Convolutional block attention module[C]//Proceedings of Computer Vision—ECCV 2018. Cham: Springer, 2018: 3-19.
- [25] HOU Q, ZHOU D, FENG J. Coordinate attention for efficient mobile network design[C]//Proceedings of 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Nashville, TN, USA: IEEE, 2021: 13708-13717.
- [26] ZHANG X, LI S, LI X, et al. DeSTSeg: Segmentation guided denoising student-teacher for anomaly detection[C]// Proceedings of 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Vancouver, BC, Canada: IEEE, 2023: 3914-3923.
- [27] LIU Z, ZHOU Y, XU Y, et al. SimpleNet: A simple network for image anomaly detection and localization[C]//Proceedings of 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Vancouver, BC, Canada: IEEE, 2023: 20402-20411.

作者简介:



李少杉(1998-),男,硕士研究生,研究方向:电子元器件表面缺陷检测技术,E-mail: lishaoshan@seu.edu.cn。



成建宏(1992-),通信作者,男,博士,高级工程师,研究方向:计算机视觉、工业缺陷检测等,E-mail: jianhong_cheng@csu.edu.cn。



张钊(1995-),男,工程师,研究方向:元器件失效分析技术。



金龙(1965-),通信作者,男,教授,博士生导师,研究方向:微特电机设计和控制、大功率电力电子技术、机器人技术等,E-mail: jinlong@seu.edu.cn。

(编辑:张黄群)