

正负样本差异特征双径向融合的自监督缺陷检测方法*

高 港 魏利胜 朱圣博

(安徽工程大学电气工程学院 芜湖 241000)

摘要:针对纹理图像表面划痕、裂纹等缺陷不规则、随机分布,导致缺陷检测准确率低的问题,研究一种基于正负样本差异特征双径向融合的自监督缺陷检测方法。首先,采用 Otsu 阈值分割提取图像前景信息,并以 DTD 数据集中的纹理图像或数据增强后的正样本叠加 Perlin 噪声,对正样本图像进行缺陷模拟以合成负样本;然后,利用正负样本经编码器输出的中间特征,计算均方误差进行特征匹配,结合坐标注意力(coordinate attention, CA)和双径向路径聚合网络(path aggregation network, PANet)加强匹配特征的信息融合;最后,将融合特征与编码器输出的低层和高层特征一同输入解码器,优化调整 Focal、L1 和 Dice 损失函数权重,实现对缺陷掩码更精准地预测。实验显示,所提模型在 MVTec AD 数据集纹理类别上的平均图像级、像素级 AUROC 分别达到了 0.995、0.968,相较于其他缺陷检测模型,分类和分割准确率均有提升,表明所提方法在纹理缺陷检测方面的有效性。

关键词:缺陷检测;合成负样本;CA;PANet;加权损失

中图分类号: TP391; TN29

文献标识码: A

国家标准学科分类代码: 520.2040

Self-supervised defect detection based on biradial fusion of differential features between positive and negative samples

Gao Gang Wei Lisheng Zhu Shengbo

(School of Electrical Engineering, Anhui Polytechnic University, Wuhu 241000, China)

Abstract: Aiming at the problem of irregular and random distribution of defects on the surface of texture images, such as scratches and cracks, which leads to low accuracy of defect detection, a self-supervised defect detection method based on the bi-radial fusion of positive and negative sample difference features is proposed. Firstly, Otsu threshold segmentation is used to extract image foreground information, and Perlin noise is superimposed on the data-enhanced positive samples or the texture images, from the DTD dataset, to simulate defects on the positive sample images and synthesize the negative samples. Then, the mean-square error is calculated for feature matching using the intermediate features output from the encoder, while the coordinate attention (CA) and path aggregation network (PANet) are combined to enhance the information fusion of the matched features. Finally, the fused features are input into the decoder together with the low-level and high-level features output from the encoder, and the weights of Focal, L1, and Dice loss functions are optimized and adjusted to realize the prediction of the defective masks more accurately. Experiments show that the average image-level and pixel-level AUROC of the proposed model on the texture category of the MVTec AD dataset reaches 0.995 and 0.968, respectively, which improves the classification and segmentation accuracies compared with the other defect detection models, demonstrating the effectiveness of the proposed method in texture defect detection.

Keywords: defect detection; feature matching; CA; PANet; weighted loss

0 引言

工业场景下,带有纹理条纹产品的制造难免会产生

种类各样的缺陷,比如划痕、裂纹等,给产品的外观和使用带来诸多负面影响。针对纹理图像缺陷的检测与定位,传统人工检测缺陷的方式易受到经验性、疲劳度等主观因素的影响,导致检测准确性不高、效率低、实时性

差^[1]的问题。为了模拟人眼功能,排除主观因素干扰,获取高质量的缺陷检测效果,基于深度学习的计算机视觉被广泛的应用于缺陷检测。因此,如何将深度学习的模型算法更好聚焦于缺陷检测与定位,具有重要的研究意义。

基于深度学习方法的缺陷检测研究工作主要以监督、半监督以及自监督学习方式为主。监督学习因结构简单、理论完整,被应用于缺陷检测研究^[2]。丝织、玻璃、光伏组件等监督模式缺陷检测层出不穷^[3],但是缺陷出现为小概率事件、随机性较强,而且缺陷种类繁多,难以获取大规模数据集,致使监督学习缺陷检测效果泛化能力不强。随后,一系列仅使用正常样本进行训练的半监督和自监督学习模式成为缺陷检测研究的主流。基于自编码器^[4]或生成对抗网络^[5]的重建模型,通过对比正常样本与重建生成的缺陷样本,得到像素级的重建误差,并用于缺陷的检测和定位。如 RIAD 模型,利用自编码器结构在原图像上进行多尺度互补掩码,以覆盖缺陷区域,减少异常区域重建误差^[6]。然而,由于卷积神经网络卓越的特征提取能力^[7],重建网络存在缺陷样本重建生成正常样本的趋势,这使得基于重建的网络模型具有先天劣势,导致检测效果不佳。此外,基于正负样本特征匹配的嵌入模型^[8],利用 ImageNet 预训练网络提取的通用特征,与嵌入多元高斯分布^[9]或存储库^[10]等方式的正样本分布进行比较,进而实现缺陷的检测。其中,PaDiM 模型^[11]通过多元高斯分布嵌入预训练网络提取异常缺陷补丁特征,而 PatchCore 模型^[10]利用正样本特征存储库

进行缺陷检测。但是,直接应用预训练网络提取特征可能导致不匹配问题,计算量和存储空间消耗较大。以 DRAEM^[12]和 CutPaste^[13]为代表的基于缺陷模拟的模型,以丝织纹理图像为异常源,并叠加正样本用于合成缺陷图像或裁剪图像补丁,同时随机粘贴生成缺陷模拟图像,作为模型输入,完成缺陷检测算法学习。然而与实际负样本缺陷相比,模拟缺陷仍不够全面且过于浅显。

针对以上研究仍存在的问题,本研究以融合存储差异特征以及缺陷模拟思想的 MemSeg 网络^[14]为基底模型,探究一种检测效果更优的端到端自监督模型。一方面,在异常模拟中进行阈值分割获取更精确的图像前景,叠加附有纹理图像信息的 Perlin 噪声构建模拟缺陷样本;另一方面,利用负样本特征和正负样本浅层特征的差异信息,通过 CA 注意力^[15]与双径向路径聚合网络 PANet^[16],完成缺陷多尺度特征的融合与高层语义特征表达;在此基础上,探讨 3 种不同损失加权系数对缺陷检测效果的影响。

1 正负样本差异特征双径向融合的自监督缺陷检测方法原理

基于正负样本差异特征双径向融合的缺陷检测继承于存储库与缺陷模拟模型方法,在关注模型精确度的同时,采用参数量较少的 ResNet18 编码器融于 U-Net 架构,以兼顾缺陷检测模型的轻量化性能^[17]。具体模型结构如图 1 所示。

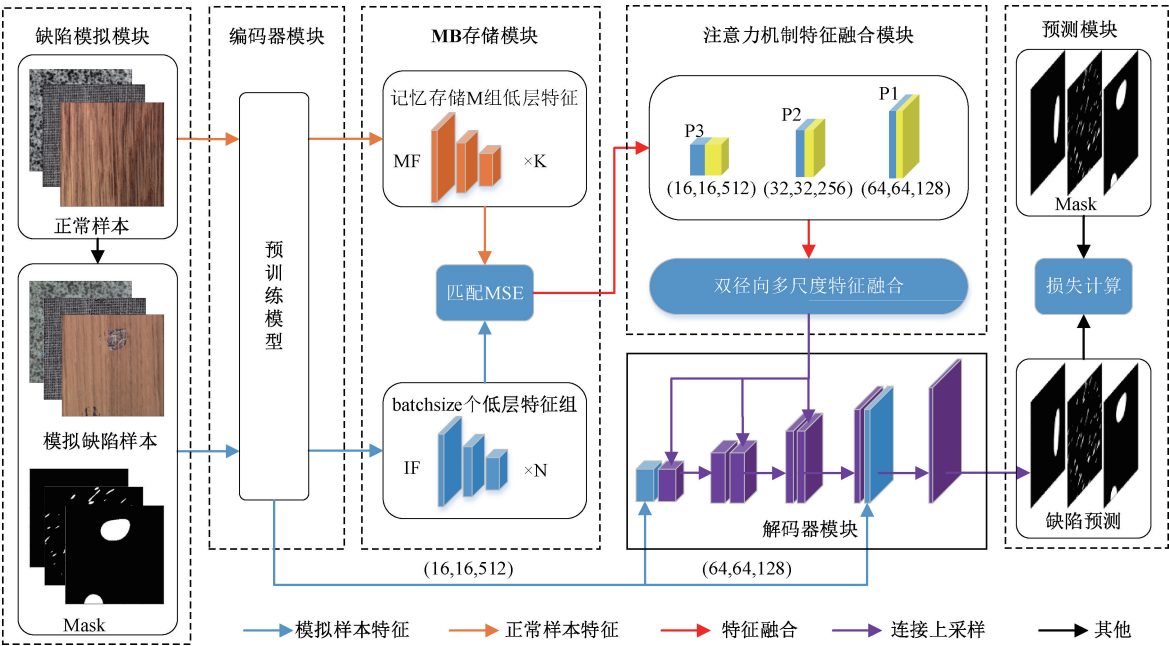


图 1 正负样本差异特征双径向融合的缺陷检测框架

Fig. 1 The defect detection framework for biradial fusion of positive and negative sample difference features

模型包含缺陷模拟、编码器、特征存储、注意力特征融合、解码器以及预测等共6个模块,分别用于实现不同功能。异常模拟模块利用正样本模拟缺陷样本,并与正样本一同作为编码器的输入;编码器模块用于提取正负样本特征,底层和高层特征传入解码器,中间特征传入特征存储模块;特征存储模块利用存储库思想,通过匹配正负样本获取特征差异信息用于特征融合;注意力特征融合模块采用CA注意力和双径向PANet融合多尺度特征;解码器模块利用融合特征与编码器特征经连接和上采样恢复样本图像大小;预测模块计算生成预测图像与掩码Mask的损失指导模型训练。

编码器采用ImageNet预训练的ResNet18网络,并为保持存储库中正负样本低层信息的一致性,冻结前三层预训练模型参数。编码器输出特征从低层到高层的维度依次是: $128 \times 128 \times 64$ 、 $64 \times 64 \times 64$ 、 $32 \times 32 \times 128$ 、 $16 \times 16 \times 256$ 、 $8 \times 8 \times 512$ 。最高层特征经卷积上采样后,将特征维度 $16 \times 16 \times 512$ 传入解码器模块,参与恢复检测图像大小的解码操作。本节将依次介绍基于正负样本差异特征双径向融合缺陷检测模型的其余模块,缺陷模拟模块、特征存储模块、注意力特征融合模块和预测模块。

1.1 缺陷模拟模块

受DRAEM^[12]异常合成思想的启发,本文缺陷模拟模块在基底模型^[14]模拟方法基础上,在掩膜和结构异常源获取过程中添加不同的图像增强策略。如图2所示,利用正样本合成缺陷模拟图像的具体流程分为3个阶段。

首先,在实际生产制造过程中获取的缺陷图像正负样本常含有复杂且无关的背景信息,若不能对背景信息进行有效的剔除,目标物体的缺陷模拟将产生较大误差,进而影响模型训练结果。所以,在缺陷模拟模块的第一阶段,针对背景深浅不同的情况,分别采用叠加正样本或灰度化后二次叠加的简单方式进行图像增强,有效扩大图像目标物体与背景的像素差值,从而利于Otsu自适应阈值分割^[18]获取图像前景信息生成前景掩膜 M_X ;同时使用与正样本大小一致的Perlin噪声,经阈值化和旋转等处理后生成噪声掩膜 M_P ; M_X 与 M_P 按元素相乘得到模拟掩膜 M 。

在第二阶段,从DTD数据集^[19]中选取图片作为异常模拟的低层纹理,即纹理异常源 Y_S 。同时,添加原始图像的纹理信息,使异常模拟合成图像更趋近实际缺陷图像。为此,对正样本进行亮度、对比度、饱和度、锐度、色调以及转置等方面的增强,并在增强后分成 8×8 栅格随机排列,重新组合成与正样本大小一致的图像,生成结构异常源 Y_J 。然后,引入透明因子 β 平衡正样本 X 和异常源 Y 数据,并与模拟掩膜 M 融合,生成带有异常源噪声信息的前景图像 Z_J 或 Z_S ,增加模拟缺陷与实际情况的

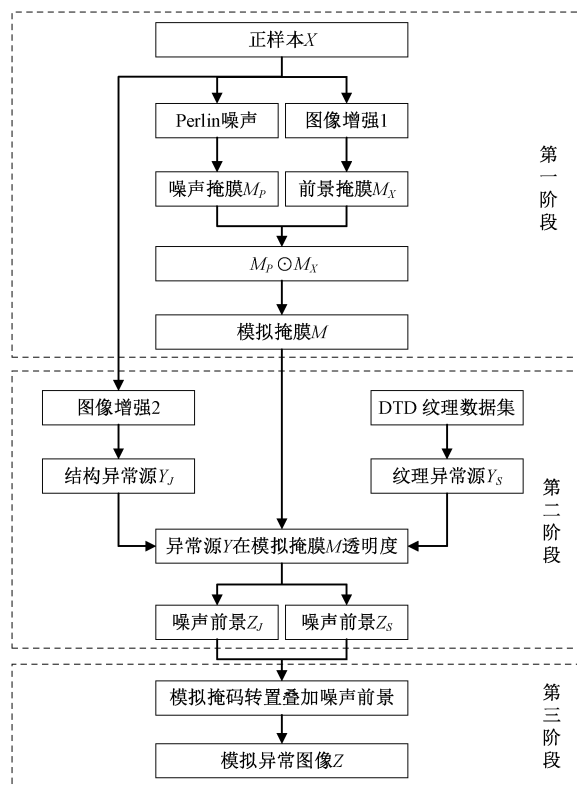


图2 合成缺陷图像流程图

Fig. 2 The flow chart for synthesizing defect images

相似性,具体公式如下:

$$\begin{cases} Z_J = \beta(M \odot Y_J) + (1 - \beta)(M \odot X) \\ Z_S = \beta(M \odot Y_S) + (1 - \beta)(M \odot X) \end{cases} \quad (1)$$

式中: $\odot$ 表示按元素相乘,透明因子 β 在区间 $[0.15, 1]$ 随机取值,可以视为对训练数据集的一种数据增强,扩展数据集的多样性。

在第三阶段,经转置后的模拟掩膜 M 与正样本按元素相乘,得到将前景像素值置零的图像,在此基础上,叠加异常源噪声前景图像 Z_J 或 Z_S ,使0像素点与噪声前景像素点相加,获取模拟缺陷图像 Z ,如式(2)所示:

$$\begin{cases} Z = (1 - M) \odot X + Z_J \\ Z = (1 - M) \odot X + Z_S \end{cases} \quad (2)$$

经过以上3个阶段的缺陷模拟过程,合成模型训练所需的缺陷图像。选取MVTec AD数据集^[20]中的carpet和tile纹理类别,对添加纹理异常源的缺陷模拟各阶段掩膜和图像进行展示,如图3所示。

1.2 特征存储模块

结合存储库嵌入的思想,特征存储模块对正样本模式形成记忆,在模型训练和推理过程中计算输入特征与正样本模式的差异,并将这种差异融入注意力机制和多尺度特征融合模块,最终将多方处理的差异特征传递至解码器。本文采取一种简单有效的均方与差方误差结合

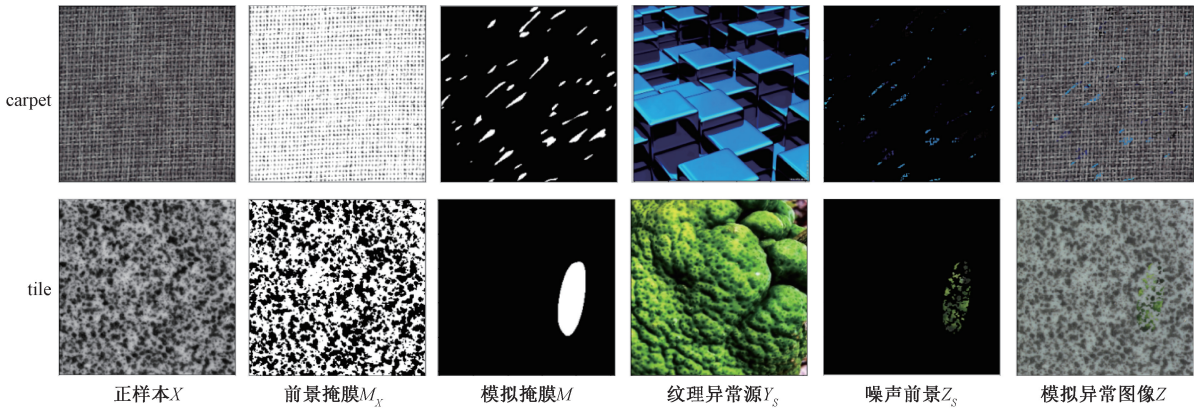


图 3 缺陷模拟图像生成展示

Fig. 3 The display of defect simulation image generation

的方式提取差异特征。

首先,正样本模式应当反映样本中有效细节信息,去除抽象化的信息碎片,同时约束高分辨率以保证检测速率。因此,本研究从训练集的每个类别中随机选取 K 个正样本,经预训练模型提取中间 3 层特征作为 K 组记忆存储特征,每组特征维度信息分别为: $64 \times 64 \times 64 \times K$ 、 $32 \times 32 \times 128 \times K$ 、 $16 \times 16 \times 256 \times K$ 。在训练和推理过程中,经编码器提取输入图像特征后,通过计算 3 个中间层特征与所有正样本记忆特征的均方误差,获得输入图像与存储信息的差异信息。针对输入的 1 张图片和存储库中的第 k 组正样本特征,差异信息 DI_k 由下式计算:

$$DI_k = \frac{1}{64^2} \sum_{i=0}^{64^2} [MF_k(64) - IF(64)]^2 + \frac{1}{32^2} \sum_{i=0}^{32^2} [MF_k(32) - IF(32)]^2 + \frac{1}{16^2} \sum_{i=0}^{16^2} [MF_k(16) - IF(16)]^2 \quad (3)$$

式中: 64、32 和 16 分别表示不同特征图尺度大小; $MF_k(64)$ 、 $MF_k(32)$ 、 $MF_k(16)$ 分别表示第 k 组不同尺度的存储特征, $IF(64)$ 、 $IF(32)$ 、 $IF(16)$ 为输入图像不同尺度特征。

其次,计算存储特征与输入图像特征差方损失的最小值。 K 组存储特征对应 K 组差异信息,遍历获取最小的差异信息,得到此差异信息所对应的存储特征作为最佳存储特征。该差异信息反映出输入图像与最佳存储样本具有最高的相似度。最佳差异信息则由最佳存储特征和输入图像特征计算每个像素点差方损失构成,该差异信息在某一区域的值越大,在输入图像相应的区域为缺陷的概率越大。输入图像特征 IF 与记忆存储特征 MF 的最佳差异特征计算所示。

$$DI_{best} = \begin{cases} [MF_{best}(64) - IF(64)]^2, & c = 64 \\ [MF_{best}(32) - IF(32)]^2, & c = 128 \\ [MF_{best}(16) - IF(16)]^2, & c = 256 \end{cases} \quad (4)$$

式中: c 为特征通道数; $MF_{best}(64)$ 、 $MF_{best}(32)$ 、 $MF_{best}(16)$ 为记忆存储的不同尺度的最佳特征; DI_{best} 为最佳差异特征,维度分别为 $64 \times 64 \times 64$ 、 $32 \times 32 \times 128$ 、 $16 \times 16 \times 256$ 。

最后,最佳差异特征 DI_{best} 与输入图像特征 IF 在通道维度上连接,将 $64 \times 64 \times 128$ 、 $32 \times 32 \times 256$ 、 $16 \times 16 \times 512$ 的 P_1 、 P_2 、 P_3 多尺度特征送入注意力机制特征融合模块。

1.3 注意力特征融合模块

经过特征存储模块,获取由输入特征 IF 和最佳差异特征 DI_{best} 连接而得的多尺度特征 P_n 。直接使用多尺度特征 P_n 将会产生特征冗余的问题,而且会增大模型计算量,导致推理速度减慢。为缓解这一问题,本文在 PANet 等路径聚合^[16]方法的启发下,采用一种利用 CA 注意力和双路径聚合的方法融合多尺度特征,组成注意力特征融合模块,结构如图 4 所示。

输入注意力特征融合模块的多尺度特征 P_n ,利用整合通道信息和方向位置信息的 CA 注意力,分别获取 P_1 、 P_2 及 P_3 不同通道的关联信息,并加权不同维度特征。由于 CA 注意力的输入输出不改变特征维度大小,如图 5 所示,因而在此基础上使用 3×3 卷积与注意力输出按元素相乘,得到特征聚焦后的多尺度特征维度分别为: $64 \times 64 \times 128$ 、 $32 \times 32 \times 256$ 、 $16 \times 16 \times 512$ 。然后每个尺度特征经过 2 层 3×3 卷积实现通道减半,进一步提高特征表达能力,卷积后特征尺度分别为: $64 \times 64 \times 64$ 、 $32 \times 32 \times 128$ 、 $16 \times 16 \times 256$ 。接着,进行双径向特征融合操作:使用双线性上采样或下采样对不同尺度大小的特征图进行高宽对齐,使用 3×3 卷积对通道数量进行 2 倍压缩或扩充,保持高层到低层或低层到高层特征方向聚合的维度信息一致,并对聚合特征执行按元素相加运算,实现高层到低层或低层到高层的特征聚合。经过双路径聚合操作后,不同尺度特征的维度信息分别为: $64 \times 64 \times 64$ 、 $32 \times 32 \times 128$ 、 $16 \times 16 \times 256$ 。

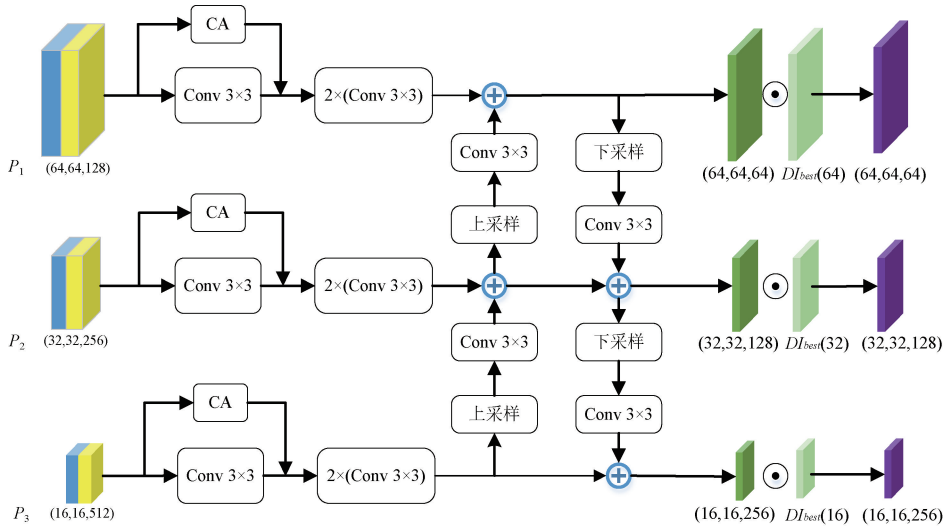


图4 注意力特征融合模块结构图

Fig. 4 The structure diagram of the attention feature fusion module

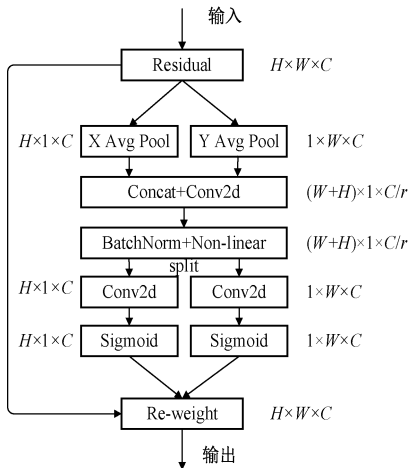


图5 CA 注意力机制结构图

Fig. 5 The structure diagram of the CA attention mechanism

最后,为充分利用均方与方差损失计算获取的差异信息,更好地分辨和定位缺陷区域,采用空间注意力特征图与路径聚合特征图按元素相乘的方法,将缺陷位置信息相关的最佳差异特征再次融合。不同尺度的空间特征计算如式(5)所示。

$$\begin{cases} S(16) = \frac{1}{256} \sum_{i=1}^{256} DI_{best}(16)_i \\ S(32) = \frac{1}{128} \left[\sum_{i=1}^{128} DI_{best}(32)_i \right] \odot S(16)^{up} \\ S(64) = \frac{1}{64} \left[\sum_{i=1}^{64} DI_{best}(64)_i \right] \odot S(32)^{up} \end{cases} \quad (5)$$

式中: $S(16)$ 、 $S(32)$ 、 $S(64)$ 表示高宽为 16、32、64 的空间注意力特征图; 256、128 及 64 表示不同尺度大小特征图的通道数; $DI_{best}(16)_i$ 、 $DI_{best}(32)_i$ 、 $DI_{best}(64)_i$ 表示不同

尺度的第 i 个通道特征图; $S(16)^{up}$ 、 $S(32)^{up}$ 表示对 16×16 、 32×32 的空间注意力特征图进行 2 倍的双线性差值上采样,保持特征信息的完整性。

1.4 预测模块

预测模块针对解码还原的图像进行阈值分割,以生成预测掩码,并将预测掩码与模拟生成的掩码进行损失计算,量化预测结果,指导模型优化方向。预测模块的损失函数通过对 L1 损失、Focal 损失以及 Dice 损失加权获取。其中,L1 损失与 L2 损失相比,可以保留更多的边缘信息;Focal 损失有助于缓解预测图像中正常区域与缺陷区域的不平衡问题,使模型训练更多地集中在难以训练的样本上,进而提高损失分割精度;Dice 损失常用于图像分割任务,能够对少数类别的模型预测结果进行更好的评估,同时具有一定边缘细节信息保留能力。利用预测分割图像 $\hat{M}$ 与模拟异常掩码图像 M 分别计算上述 3 种损失,公式如下:

$$L_{l1} = |M - \hat{M}| \quad (6)$$

$$L_f = -\alpha_i (1 - p_i)^{\gamma} \log(p_i) \quad (7)$$

$$L_d = 1 - \frac{2|M \cap \hat{M}|}{|M| + |\hat{M}|} \quad (8)$$

式中: L_{l1} 、 L_f 、 L_d 分别表示 L1 损失、Focal 损失及 Dice 损失; p_i 表示像素预测概率值,若 p 表示像素类别 0 或 1 的预测概率,当某像素对应模拟掩码 M 该位置像素为 1 时 $p_i = p$,当某像素对应模拟掩码 M 该位置像素为 0 时 $p_i = 1 - p$ 。

结合 3 种损失的优点,对以上 3 种损失赋权值,组成加权损失,在训练过程中通过不断学习降低该损失,优化模型。

$$L = \lambda_l L_l + \lambda_f L_f + \lambda_d L_d \quad (9)$$

式中: L 为模型总的加权损失, λ_l 、 λ_f 、 λ_d 分别表示 L1 损失、Focal 损失及 Dice 损失加权超参数, 且满足条件 $\lambda_l + \lambda_f + \lambda_d = 1$ 。

2 实验与分析

2.1 实验数据集

MVTec AD 数据集是 2019 年由 MVTec 公司团队开发并提供的一个用于表面缺陷检测和分类的计算机视觉数据集。该数据集包含丝织纹理 (Textures) 和物体 (Objects) 两个大类, 本实验选取的纹理类别包含地毯 (carpet)、栅格 (grid)、皮革 (leather)、瓷砖 (tile)、木材 (wood) 共 5 个类别, 每个类别约 350 张图片, 共 1 781 张图片, 其中训练与验证集 1 266 张图片, 测试集 515 张。图像分辨率以 1 024×1 024 为主, 为适应网络模型训练, 将输入图像大小统一为 256×256。

2.2 实验配置与测评标准

本文实验采用的硬件配置为 AMD Ryzen 7 5800H (3.20 GHz) 16 GB 处理器, NVIDIA GeForce RTX 3070 8GB 显卡; 操作系统为 Windows 10, 软件平台为 PyCharm; 使用深度学习框架为 Pytorch1.10.0, Python 版本为 3.8, 在 CUDA11.3 和 cuDNN8.21 上进行加速训练。实验参数的设置: batch size 为 8; 训练迭代次数为 5 000; 存储特征组数 K 取 30; Perlin 噪声阈值为 0.5; 优化器为 AdamW; 训练的学习率采用余弦退火预热重启策略, 设置最大初始学习率 0.003, 最小学习率 0.000 1, 预热比重 0.1, 衰减权重 0.000 5。

实验以接受者操作特征曲线线下面积 (area under the receiver operating characteristic curve, AUROC) 为评价指标, 其取值范围在 0~1, 数值越大表示模型越好。AUROC 表示 ROC 曲线线下面积的大小, 而 ROC 曲线是以真阳性率 (true positive rate, TPR) 为纵坐标, 假阳性率 (false positive rate, FPR) 为横坐标的曲线, TPR 和 FPR 如式 (10)、(11) 所示。

$$TPR = \frac{TP}{TP + FN} \quad (10)$$

$$FPR = \frac{FP}{FP + TN} \quad (11)$$

式中: TP 为真正例, FN 为假反例, FP 为假正例, TN 为真反例。

图像级 AUROC 综合考虑模型在多个类别上的分类能力, 提供一个全局的性能度量; 而像素级 AUROC 考虑图像中每个像素点的分类结果, 在像素级别上评估模型对不同缺陷的分割能力。实验分别从图像级别和像素级

别对缺陷检测结果进行评估, 更加综合全面地评价模型的性能。

2.3 实验结果分析

为探究 Focal、L1 以及 Dice 不同加权系数对模型检测精度的影响, 在 MVTec AD 数据集的 carpet 类别上进行不同损失搭配下模型训练的实验, 实验以 0.2 为损失权重最小单位, 共采取 20 种损失加权分配策略。实验每迭代 100 次记录一次图像级、像素级 AUROC, 分别对比不同加权方式下, 第 2 300、3 100、5 000 次迭代以及最佳迭代次数的评价指标, 其中 2 300 和 3 100 是 carpet 类别模型训练过程中精度易掉点的迭代次数。实验结果如表 1 所示。

由表 1 的实验结果可知, 单一地使用 Focal、L1 或 Dice 损失对模型进行训练, 均不能取得很好的效果。仅使用 L1 或 Focal 损失, 模型在取得最佳指标后, 会在第 2 300 和 3 100 次迭代记录中降点; 而仅使用 Dice 损失, 模型虽然可以很早地达到最佳指标, 但在后续的迭代记录中性能指标下降, 同时一直维持 0.500 的低图像级 AUROC 水平。在多种损失加权实验中, 若以最早获得最佳性能为主要衡量指标, 则 L1、Focal 以及 Dice 损失权重分别为 0.6、0、0.4 时, 1400 次迭代为最佳, 图像级、像素级 AUROC 分别达到 0.999 与 0.950, 并且在后续迭代的记录中均保持 0.908 以上; 若综合衡量各阶段指标, 则 L1、Focal 以及 Dice 损失的权重分别为 0、0.2、0.8 时, 第 2 100 次迭代取得最佳性能, 图像级、像素级 AUROC 分别达到 1.000 与 0.985, 且后续迭代指标均保持在 0.923 以上。

实验分别选取 L1、Focal 以及 Dice 损失权重为 0、0.2、0.8, 并在原始模型上分别添加不同的改进点, 进行消融实验, 实验结果如表 2 所示。

实验结果表明, 缺陷模拟阶段的数据增强与注意力特征融合阶段的双径向特征, 均能有效地提高模型区分缺陷正负样本的能力, 图像和像素级 AUROC 分别较原始模型提高了 0.005、0.017 和 0.013、0.024, 而结合缺陷模拟图像增强与双径向特征融合模型在图像和像素级 AUROC 分别达到 0.995 和 0.968, 相较原始模型提升了 0.028 和 0.049。

2.4 对比实验

为进一步探究正负样本差异特征双径向融合检测模型对纹理图像裂纹、划痕等缺陷的检测效果, 本文模型在 MVTec AD 数据集纹理类别的 5 个目标上进行训练, 获取图像级和像素级 AUROC 数据, 并与 SPADE^[8]、PaDiM^[11]、CutPaste^[13] 等经典缺陷检测算法进行对比。实验数据如表 3、4 所示, 各个模型检测效果对比如图 6 所示。

表 1 不同权重损失搭配的性能指标对比

Table 1 The comparison of performance indicators for different weight loss combinations					
L1 + Focal + Dice	第 2 300 次	第 3 100 次	第 5 000 次	最佳值	最佳迭代次数
1+0+0	0.683,0.768	0.917,0.856	0.983 ,0.918	0.958,0.955	2 200
0.8+0.2+0	0.938,0.915	0.916,0.884	0.940,0.916	0.979,0.926	1 800
0.8+0+0.2	0.912,0.928	0.816,0.880	0.923,0.942	0.987,0.944	2 100
0.6+0.4+0	0.846,0.926	0.790,0.866	0.910,0.936	0.998,0.966	2 100
0.6+0.2+0.2	0.810,0.719	0.954,0.881	0.965,0.942	0.982,0.971	2 100
0.6+0+0.4	0.913,0.908	0.929,0.928	0.966,0.950	0.999,0.950	1 400
0.4+0.6+0	0.756,0.881	0.842,0.925	0.866,0.948	0.974,0.954	1 700
0.4+0.4+0.2	0.861,0.893	0.724,0.825	0.919,0.937	0.991,0.967	2 100
0.4+0.2+0.4	0.752,0.803	0.891,0.885	0.928,0.951	0.997,0.957	3 800
0.4+0+0.6	0.758,0.791	0.904,0.834	0.974,0.912	0.976,0.965	1 300
0.2+0.8+0	0.631,0.742	0.831,0.924	0.777,0.946	0.954,0.976	2 100
0.2+0.6+0.2	0.651,0.867	0.902,0.918	0.859,0.955	0.988,0.967	1 100
0.2+0.4+0.4	0.873,0.904	0.889,0.936	0.934,0.966	0.984,0.977	2 100
0.2+0.2+0.6	0.807,0.882	0.831,0.862	0.963,0.956	0.983,0.975	1 300
0.2+0+0.8	0.955,0.946	0.933,0.943	0.949,0.965	1.000 ,0.973	2 100
0+1+0	0.774,0.859	0.879,0.883	0.814,0.837	0.984,0.974	2 100
0+0.8+0.2	0.909,0.936	0.911,0.927	0.955,0.948	0.986,0.974	2 100
0+0.6+0.4	0.830,0.906	0.917,0.936	0.964, 0.969	0.993,0.976	2 100
0+0.4+0.6	0.901,0.928	0.944, 0.962	0.946,0.960	0.983,0.980	2 100
0+0.2+0.8	0.957 ,0.923	0.981 ,0.957	0.980, 0.965	1.000 , 0.985	2 100
0+0+1	0.500, 0.952	0.500,0.949	0.500,0.936	0.997,0.950	500

表 2 消融实验结果

Table 2 The results of the ablation test				
方法	第 1 组	第 2 组	第 3 组	第 4 组
原始模型	√	√	√	√
缺陷模拟数据增强		√		√
双径向特征			√	√
图像级 AUROC	0.967	0.972	0.984	0.995
像素级 AUROC	0.919	0.932	0.943	0.968

表 3 与经典模型缺陷检测图像级 AUROC 对比表

Table 3 The table of the comparison for defect detection image level AUROC between the model and the classical models					
类别	文献[8]	文献[11]	文献[13]	文献[14]	本文方法
carpet	0.928	0.984	0.929	0.873	0.977
grid	0.473	0.898	0.946	0.972	0.997
leather	0.954	0.988	0.909	1.000	1.000
tile	0.965	0.959	0.978	1.000	1.000
wood	0.959	0.990	0.965	0.991	0.999
平均	0.856	0.964	0.945	0.967	0.995

表 4 与经典模型缺陷检测像素级 AUROC 对比表

Table 4 The table of the comparison for defect detection pixel level AUROC between the model and the classical models					
类别	文献[8]	文献[11]	文献[13]	文献[14]	本文方法
carpet	0.663	0.988	0.926	0.901	0.948
grid	0.566	0.936	0.962	0.864	0.963
leather	0.615	0.990	0.974	0.944	0.987
tile	0.632	0.917	0.914	0.983	0.991
wood	0.644	0.900	0.908	0.949	0.949
平均	0.624	0.946	0.937	0.919	0.968

由图 6 可知,本文方法对纹理图像裂纹、划痕等缺陷的分割定位效果最为清晰。由表 2 和 3 可知,本文方法在诸多纹理缺陷检测效果最优,平均图像级和像素级 AUROC 分别达到 0.995 和 0.968,为所有对比模型最优值,表明改进模型在纹理缺陷分类与分割上效果最佳,整体上具有良好的缺陷检测效果。

3 结 论

为了避免监督学习缺陷样本难以获取与学习泛化能力弱的问题,本文研究一种融合异常模拟、存储库差异信息和预训练模型的方法,在异常模拟中引入利于分割图像背景的数据增强,在存储库模块采用均方、差方获取差异特征,使用双径向特征进行多尺度融合,实现了仅利用无缺陷正样本进行训练的高精度自监督缺陷检测。在此基础上,采用 MVTec AD 纹理数据集实验研究 Focal、L1 和 Dice 损失对模型训练的影响,分别选取 0、0.2、0.8 的损失权重,进行模型消融和对比实验,结果显示所研究方法在图像级和像素级的平均指标上分别达到 0.995、0.968,相较原模型分别提高了 0.028 和 0.049,实现了缺陷检测与分割的良好效果。

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.
TANG B, KONG J Y, WU SH Q. Review of surface defect detection based on machine vision [J]. Journal of Image and Graphics, 2017, 22(12): 1640-1663.

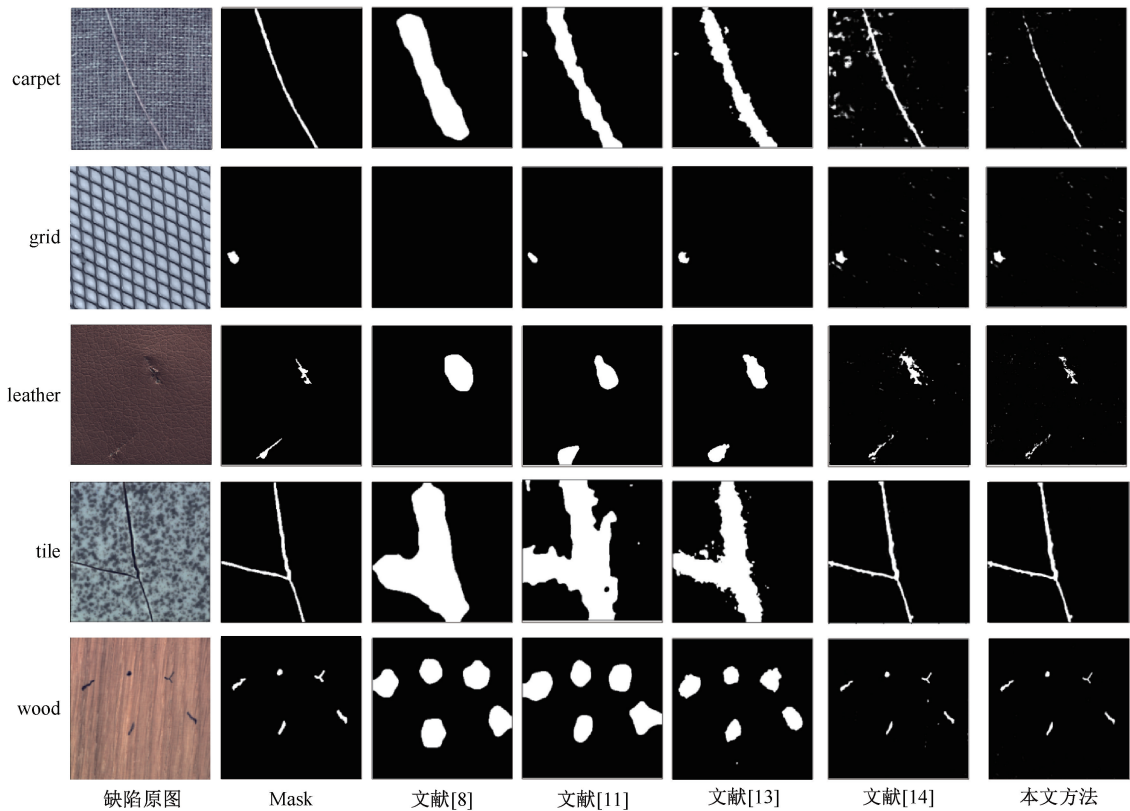


图 6 缺陷检测效果对比图

Fig. 6 The comparison diagrams of defect detection effect

[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]. Chinese Journal of Scientific Instrument, 2022, 43(1): 198-219.

[3]

曾秀云,陆华才,吕禾丰. 基于改进 Faster R-CNN 的棉布包装缺陷检测的方法研究[J]. 电子测量与仪器学报, 2022, 36(4) : 179-186.

ZENG X Y, LU H C, LYU H F. Research on cotton packaging defect detection method based on improved Faster R-CNN [J]. Journal of Electronic Measurement and Instrumentation, 2022, 36(4) : 179-186.

[4]

杨文元. 多标记学习自编码网络无监督维数约简[J]. 智能系统学报, 2018, 13(5) : 808-817.

YANG W Y. Unsupervised dimensionality reduction of multi-label learning via autoencoder networks [J]. CAAI Transactions on Intelligent Systems, 2018, 13 (5) : 808-817.

[5]

王坤峰,苟超,段艳杰,等. 生成式对抗网络 GAN 的研究进展与展望[J]. 自动化学报, 2017, 43 (3) : 321-332.

WANG K F, GOU CH, DUAN Y J, et al. Generative adversarial networks: the state of the art and beyond [J]. Acta Automatica Sinica, 2017, 43(3) : 321-332.

[6]

ZAVRTANIK V, KRISTAN M, SKOČAJ D. Reconstruction by inpainting for visual anomaly detection [J]. Pattern Recognition, 2021, 112: 107706.

[7]

郑远攀,李广阳,李晔. 深度学习在图像识别中的应用研究综述[J]. 计算机工程与应用, 2019, 55 (12) : 20-36.

ZHENG Y P, LI G Y, LI Y. Survey of application of deep learning in image recognition [J]. Computer Engineering and Applications, 2019, 55(12) : 20-36.

[8]

COHEN N, HOSHEN Y. Sub-image anomaly detection with deep pyramid correspondences [J]. ArXiv Preprint, 2020, ArXiv:2005. 02357.

[9]

杨贺雅,邢纹硕,向鑫,等. 基于多元高斯分布异常检测模型的 MMC 子模块开路故障诊断方法[J]. 电工技术学报, 2023, 38(10) : 2744-2756.

YANG H Y, XING W SH, XIANG X, et al. A sub-module open-circuit fault detection and location strategy for modular multilevel converters based on multivariate gaussian distribution [J]. Transactions of China Electrotechnical Society, 2023, 38(10) : 2744-2756.

- [10] ROTH K, PEMULA L, ZEPEDA J, et al. Towards total recall in industrial anomaly detection [C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022: 14318-14328.
- [11] 李亚舟, 曹江涛, 姬晓飞. 融合改进 Padim 建模和 ResNet 网络的喷涂质量检测算法[J]. 电子测量与仪器学报, 2022, 36(11): 91-97.
LI Y ZH, CAO J T, JI X F. Spraying quality detection algorithm by fusing improved Padim modeling with ResNet network [J]. Journal of Electronic Measurement and Instrumentation, 2022, 36(11): 91-97.
- [12] ZAVRTANIK V, KRISTAN M, SKOČAJ D. Draem-a discriminatively trained reconstruction embedding for surface anomaly detection [C]. Proceedings of the IEEE/CVF International Conference on Computer Vision, 2021: 8330-8339.
- [13] LI C L, SOHN K, YOON J, et al. Cutpaste: Self-supervised learning for anomaly detection and localization [C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021: 9664-9674.
- [14] YANG M, WU P, FENG H. MemSeg: A semi-supervised method for image surface defect detection using differences and commonalities [J]. Engineering Applications of Artificial Intelligence, 2023, 119: 105835.
- [15] HOU Q, ZHOU D, FENG J. Coordinate attention for efficient mobile network design [C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021: 13713-13722.
- [16] LIU S, QI L, QIN H, et al. Path aggregation network for instance segmentation [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 8759-8768.
- [17] 肖东, 韩晨, 范文强. 基于 U-net 和 ResNet 的图像缺陷检测 [J]. 计算机与数字工程, 2022, 50(8): 1791-1794.
XIAO D, HAN CH, FAN W Q. Image defect detection based on U-net and ResNet [J]. Computer & Digital Engineering, 2022, 50(8): 1791-1794.
- [18] 黄梦涛, 连一鑫. 基于改进 Canny 算子的锂电池极片表面缺陷检测 [J]. 仪器仪表学报, 2021, 42(10):

199-209.

HUANG M T, LIAN Y X. Lithium battery electrode plate surface defect detection based on improved Canny operator [J]. Chinese Journal of Scientific Instrument, 2021, 42(10): 199-209.

- [19] CIMPOI M, MAJI S, KOKKINOS I, et al. Describing textures in the wild [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2014: 3606-3613.
- [20] BERGMANN P, FAUSER M, SATTLEGGGER D, et al. MVTec AD—A comprehensive real-world dataset for unsupervised anomaly detection [C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019: 9592-9600.

作者简介



高港, 2020 年于安徽工程大学获得学士学位, 现安徽工程大学硕士研究生, 主要研究方向为深度学习缺陷检测。

E-mail: 2210320106@stu.ahpu.edu.cn

Gao Gang received his B. Sc. degree in 2020 from Anhui Polytechnic University. Now he is a M. Sc. candidate at Anhui Polytechnic University. His research interests include deep learning and defect detection.



魏利胜 (通信作者), 2001 年于安徽工程大学获得学士学位, 2004 年于中国航天科工集团 061 基地获得硕士学位, 2009 年于上海大学获得博士学位, 现为安徽工程大学教授、硕士生导师, 主要研究方向为图像识别与应用、嵌入式仪器仪表及系统、智能化网络控制理论、系统和仿真。

E-mail: lshwei_11@163.com

Wei Lisheng (Corresponding author) received his B. Sc. degree in 2001 from Anhui Polytechnic University, received his M. Sc. degree in 2004 from China Aerospace Science and Industry Corporation 061 Base and received his Ph. D. degree in 2009 from Shanghai University. Now he is a professor and master's supervisor at Anhui Polytechnic University. His research interests include image recognition and application, embedded instrumentation and system, intelligent network control theory, system and simulation.