

文章编号:1672-3961(2025)04-0040-08

DOI:10.6040/j.issn.1672-3961.0.2024.183

基于自监督卷积和无参数注意力机制的工业品表面缺陷检测

周群颖,隋家成,张继*,王洪元

(常州大学计算机与人工智能学院,江苏常州 213164)

摘要:目前,基于深度学习的工业品表面缺陷检测存在着负样本数据集少、样本易受复杂工业环境影响无法提取有效特征、标注数据集需要耗费大量人工成本的问题。为解决上述问题,本研究提出一个基于知识蒸馏的工业品表面缺陷检测模型,并在此模型中加入自监督预测卷积模块和无参数注意力机制。该模型将教师网络所学习到的丰富特征知识传递给学生网络,有效提高模型特征表达能力并对缺陷进行像素级定位。该模型在 MVTec-AD 数据集进行试验,并与各类试验方法的结果进行对比,检测指标和定位指标在不同模型曲线下面积 A_{ROC} 上有所提升,证明该方法可以有效提高模型的检测和定位能力。

关键词:工业品表面缺陷检测;深度学习;知识蒸馏;自监督预测;注意力机制

中图分类号:TP391

文献标志码:A

引用格式:周群颖,隋家成,张继,等.基于自监督卷积和无参数注意力机制的工业品表面缺陷检测[J].山东大学学报(工学版),2025,55(4):40-47.

ZHOU Qunying, SUI Jiacheng, ZHANG Ji, et al. Industrial product surface defect detection based on self supervised convolution and parameter free attention mechanism[J]. Journal of Shandong University (Engineering Science), 2025, 55(4):40-47.

Industrial product surface defect detection based on self supervised convolution and parameter free attention mechanism

ZHOU Qunying, SUI Jiacheng, ZHANG Ji*, WANG Hongyuan

(College of Computer Science and Artificial Intelligence, Changzhou University, Changzhou 213164, Jiangsu, China)

Abstract: Currently, surface defect detection in industrial products using deep learning faced several challenges. These included a limited number of negative sample datasets, susceptibility to complex industrial environments that hindered effective feature extraction, and high labor cost for labeling datasets. In order to solve the above problems, a knowledge distillation-based surface defect detection method for industrial products was proposed. The model incorporated a self-supervised predictive convolution module and a parameter-free attention mechanism. It transferred the rich feature knowledge learned by the teacher network to the student network. This approach effectively improved the model feature expression and localized defects at the pixel level. The model was experimented on the MVTec-AD dataset, and the comparison results with the state-of-the-art experimental methods showed that its detection and localization metrics were improved on A_{ROC} . The results proved that the method could improve the model's detection and localization capabilities.

Keywords: surface defect detection of industrial products; deep learning; knowledge distillation; self supervised prediction; attention mechanism

收稿日期:2024-07-23

基金项目:江苏省研究生科研创新资助项目(KYCX22_3061)

第一作者简介:周群颖(1999—),女,江苏南京人,硕士研究生,主要研究方向为计算机视觉。E-mail:zqy349105910@163.com

* 通信作者简介:张继(1981—),男,江苏常州人,副教授,硕士生导师,硕士,主要研究方向为计算机视觉、人工智能。

E-mail:zhangji@cczu.edu.cn

0 引言

在实际工业品生产中,由于工作环境及现有技术的不足会出现各类具有不同缺陷的工业品。这类工业品美观性差,严重影响产品质量和使用寿命^[1]。因此,工业品表面缺陷检测是提升工业品性能的重要手段和提高工业生产技术的重要环节^[2]。传统的工业品表面缺陷检测采用人工检测的方式,结果受人工检测的主观性影响,同时还存在着效率低、成本高等问题,不能满足现代工业发展的要求^[3]。针对存在的不足,基于深度学习的工业品缺陷检测技术不断发展并投入到工业生产中^[4]。基于有监督的工业品表面缺陷检测方法因符合传统缺陷检测方法理念而最先被研发^[5],但其训练样本必须被大量标记,不符合工业品缺陷检测的要求。基于无监督的工业品表面缺陷检测方法^[6]可以直接使用正常样本训练模型,分析位置缺陷的特征并自动分类,但结果也易受复杂环境的影响。基于弱监督的工业品表面缺陷检测方法通过采用图像级标注来获取缺陷定位的效果,能够在降低成本的同时获得更好的检测能力,但其发展并不成熟,研究时间较短,模型研究少^[7-8]。

针对上述工业缺陷检测的不足,本研究提出一个基于自监督预测卷积模块和无参数注意力机制的无监督知识蒸馏模型(multiresolution knowledge distillation for anomaly detection based on convolution and attention mechanism, MKDA-CA)。该模型通过教师网络向学生网络传递知识实现知识迁移,主要贡献如下。

(1)基于教师网络与学生网络构建知识蒸馏模型,提高模型对缺陷特征的泛化能力并实现知识迁移。

(2)加入自监督预测卷积模块,预测全局信息达到增强特征的效果,有利于分辨缺陷特征并对缺陷区域进行像素级定位。

(3)加入无参数注意力机制,综合了通道和空间两个维度的特征信息,增强了模型对重要特征的提取能力,解决过拟合问题的同时提高缺陷检测的效率。

1 研究背景

1.1 知识蒸馏

2006年,文献[9]提出一种模型压缩方式。为

了达到压缩目的,使用许多复杂的学习器训练出伪标签,利用初始数据和伪数据训练小型网络,然后将来自教师网络的信息转移到学生网络中,不仅提高了模型的运行速度,还保证了学生网络的精确度。由此,具有模型压缩功能的知识蒸馏方法发展起来^[10]。文献[11]首次基于教师网络的输出层提出知识蒸馏的概念,并验证方法的有效性。该方法使用一种可以应用于任何网络体系结构的激活函数输出知识提取方法,通过教师网络诱导学生网络进行训练,获得相应的全连接层输出并计算学生网络的损失函数,最后利用反向传播算法优化学生网络。该方法将集成中的知识压缩成更容易部署的单一模型,重点在于设计维度变换函数,保证学生网络和教师网络维度相同,减少在维度变化过程中特征的损失,适用于具有多层次高抽象的特征样本。文献[12]将知识蒸馏技术与度量学习相结合,对教师网络的中间层进行知识蒸馏,获得教师网络和学生网络中每个中间层输出的特征图,利用两个损失函数使学生网络和教师网络拥有相同的特征分布,增大异类特征的差距。该方法不仅让学生网络接收到教师网络的特征结果,还学习教师网络模型的特征输出过程,从根本上提高学生网络的特征学习能力。基于关系层的知识蒸馏可以从角度和距离对样本进行分类,重点在于探索特征映射关系。文献[13]提出一种以关系层知识蒸馏为基础的模型,摒弃传统蒸馏中教师网络和学生网络特征对应的关系,保留教师网络中间层样本的特征相似性,提高了知识蒸馏的鲁棒性。

1.2 注意力机制

注意力机制在计算机视觉领域中发挥着重要作用,其模仿人类视觉系统对于外部信息的处理能力,能够对不同的输入图像进行加权处理,自适应地调整权值以调整模型对每个输入数据的关注程度,注意力放在显著区域和关键信息上,过滤其他无关的外部信息,从而提高模型特征提取的能力^[14]。计算机视觉中的注意力机制按作用区域可分为以下4类:通道注意力机制、空间注意力机制、时间注意力机制、分支注意力机制。本研究着重介绍通道注意力机制和空间注意力机制。

(1)通道注意力机制。通道注意力的目的为通过网络学习的方式获取每个特征通道的权重,从而得到不同通道之间的相关性。文献[15]提出通道

注意力的概念,在卷积运算的基础上,提出挤压激励网络,使其通过在局部感受野内的通道信息来提取信息特征,具有良好的泛化能力,以最小的计算成本带来显著的性能改进。文献[16]基于挤压激励网络进行改进,提出了轻量级高效通道模块(efficient channel attention network, ECA),实现局部跨通道的信息交互,利用最少的参数克服了性能和复杂性权衡的矛盾,提高各类深度卷积网络性能。

(2)空间注意力机制。空间注意力将输入图像中的空间信息通过空间转换到另一个空间中,在保留关键特征信息的同时将位置信息加权输出,从而达到提高特征表达能力的目的。文献[17]提出一个用于前馈卷积网络的注意力模块(convolutional block attention module, CBAM),在通道和空间维度上推断注意力,并通过自适应操作进行特征细化,提高模型分类检测的性能。

2 本研究方法

2.1 总体框架

知识蒸馏模型主要分为教师网络和学生网络两个模块,其中教师网络使用了一种深度卷积神经网络模型(visual geometry group 16, VGG16)^[18]在

ImageNet^[19]上的预训练网络,可以在精度较高的情况下,将提前训练好的相关参数导入到更紧凑的网络,有利于提升迁移学习效率。学生网络选择了更简易的深度卷积神经网络模型(visual geometry group, VGG),有利于提升学习的效率。

在训练过程中,将正常图像样本分别输入到教师网络和学生网络中。教师网络的每个卷积层都有相应数量的卷积核对输入的特征图像进行特征提取,然后将教师网络中间特征作为真实标签,将学生网络中的中间特征作为预测标签,最后将真实标签和预测标签输出。教师网络总共分为5个卷积块,由于第一个卷积块的特征传递的知识较少,因此模型只选用了后四个卷积块的最后一层作为特征提取层,并在中间层训练学生网络。为了增强特征的提取能力,在教师网络的第三个特征传递层增加了无参数注意力机制,在第五个特征传递层前加入了自监督预测卷积模块,提升了模型对缺陷特征的检测能力。最后,利用损失函数计算教师网络和学生网络的输出差异,如果存在差异就表示输入数据具有缺陷,再通过阈值检测对缺陷区域进行像素级定位,即可得到输出结果。该模型整体框架如图1所示。

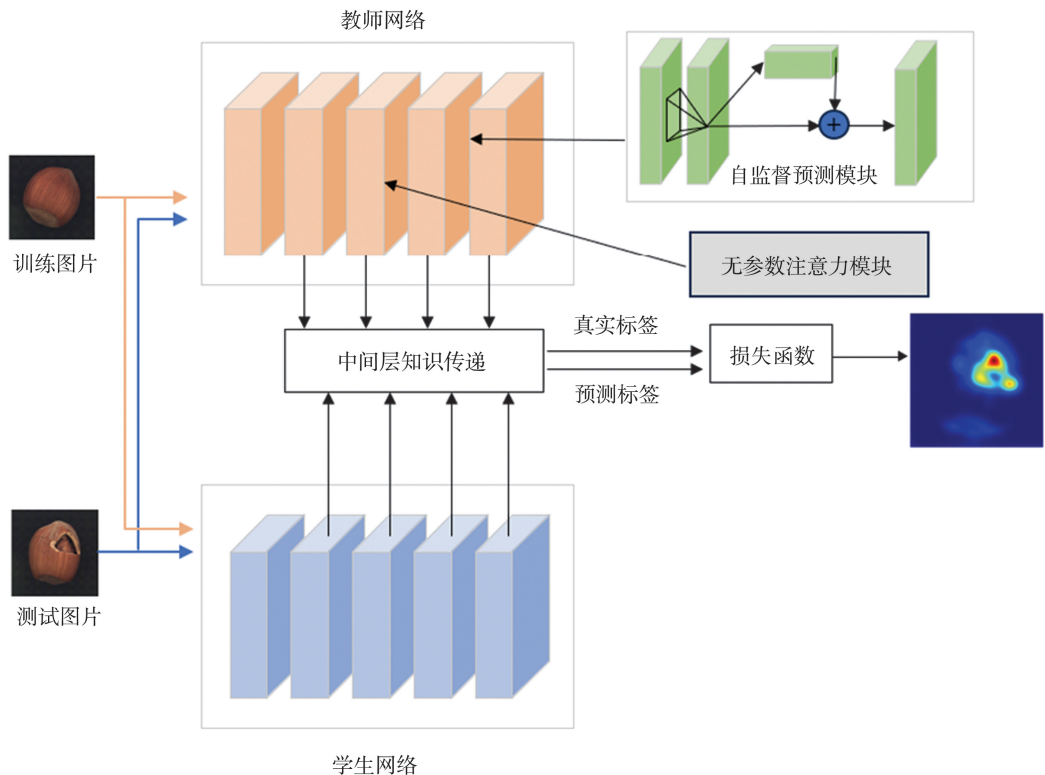


图1 知识蒸馏模型框架图

Fig.1 Framework diagram of knowledge distillation model

本章通过两个组合的损失函数计算样本的特征距离,第一个损失函数为

$$L_1 = \sum_{i=1}^{N_D} \frac{1}{N_i} \sum_{j=1}^{N_i} (a_T^D(j) - a_C^D(j))^2, \quad (1)$$

第二个损失函数为

$$L_2 = 1 - \sum_i \frac{\text{vec}(a_T^D)^T \cdot \text{vec}(a_C^D)}{||\text{vec}(a_T^D)|| \cdot ||\text{vec}(a_C^D)||}, \quad (2)$$

两个损失函数组合为

$$L_{\text{total}} = L_1 + \lambda L_2, \quad (3)$$

式中, L_1 为利用欧氏距离表示教师网络与学生网络激活值之间的距离, N_D 为第 i 层关键层神经元的数量, a_C^D 为克隆网络的激活值, N_i 为第 i 个神经元数量, L_2 为利用方向损失函数表示学生网络激活值与教师网络激活值之间的距离, $\text{vec}(x)$ 为向量化函数, L_{total} 表示两个损失函数的组合结果, λ 为两个组成项的平衡因子。

在本研究的方法中,教师网络通过知识蒸馏引导学生网络的学习。训练教师网络学习丰富的特征表示和复杂的类间关系。再将教师网络的输出(软标签)、中间特征图等知识提取出来。这些信息不仅包含最终预测结果,还包括对特征空间的理解。接着,通过设计包含蒸馏损失和传统分类损失的联合损失函数,加强对特征的提取,使学生网络在训练过程中不仅模仿教师网络的预测分布,还学习其特征提取方式和类间关系。蒸馏损失通过调整参数,强调软标签的细微差别,使学生网络在学习过程中关注到类间的微小差异。最终,学生网络能够从教师网络中继承知识,实现高效的特征学习和准确的缺陷检测,从而提高模型的检测精度和鲁棒性。

2.2 自监督预测卷积模块

在无监督缺陷检测中,很多特征信息会被屏蔽,需将重构误差作为异常分数来预测被屏蔽的特征,以增强重构功能。由于知识蒸馏模型在训练过程中只接受正常样本,所以当知识蒸馏模型进入模型测试阶段时,面对异常样本,会造成重构误差,导致模型训练结果不准确。而自监督预测卷积模块利用其特征预测功能,通过自监督的方式进行训练,该模块可以在神经网络的任何卷积层进行特征屏蔽预测,并具有良好的适应性,可用于大部分的缺陷检测模型,同时提高缺陷检测模型的性能。

自监督预测卷积模块由一个屏蔽卷积层和一个通道注意力机制模块(SE注意力机制模块)组成,屏蔽卷积层结构如图2所示。该自监督预测卷积模块选取感受野四个角落部分的像素点并忽视

剩下的像素点,将4个部分的像素点进行卷积操作并相加后得到一个值,用该值代表该感受野的特征表示。

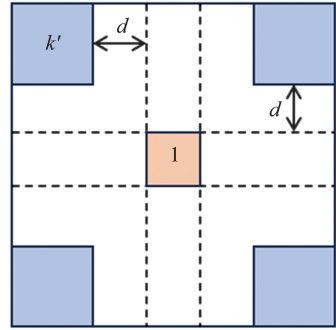


图2 屏蔽卷积层示意图

Fig.2 Schematic diagram of shielding convolutional layers

该卷积操作的计算公式为

$$k = 2k' + 2d + 1, \quad (4)$$

式中: k 为感受野的特征值; k' 为选中像素点的区域大小; d 为感受野边缘到中心点的距离,感受野大小设定为1。经过屏蔽卷积层得到的结果将导入SE注意力机制模块,通过全局性信息加强重构。

SE注意力机制模块的流程图如图3所示,该模型包含压缩模块和激励模块,图中 W 为特征图的宽度, H 为特征图的高度, C 为通道数量。SE注意力机制模块可以根据屏蔽卷积层输出的特征缩放模型中的每个通道,其非线性的处理方式不仅可以提高自监督预测卷积模块的重构能力,还能够关注全局信息,放大正常样本和异常样本的差距。因此,在模型中加入自监督预测卷积模块,能够提高缺陷检测模型的异常检测效果。

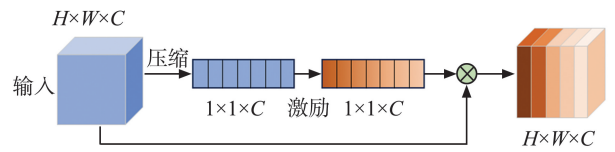


图3 SE注意力机制流程图

Fig.3 Flowchart of SE attention mechanism

2.3 无参数注意力机制模块

注意力机制可以极大地提高缺陷检测的性能,目前对注意力机制的研究更着重于构建复杂模块以提升卷积神经网络的特征表达能力,然而现有的注意力机制模块存在着不足:各类复杂因素决定了注意力机制模块缺乏灵活性;注意力机制模块只能根据通道维度或空间维度处理特征信息,维度的单一性限制神经网络发现更多特征的能力。

无参数注意力机制就能很好的解决上述问题,该注意力机制可以灵活地被使用在各个模块中,且同时考虑到通道维度和空间维度,平等对待两个维度中的

各特征信息,在不向原始神经网络添加任何参数的情况下推断出特征映射的三维注意力权重,提高了缺陷特征信息的检测能力,其结构如图4所示。

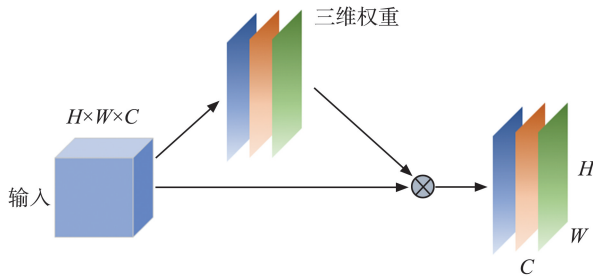


图4 无参数注意力机制流程图

Fig.4 Flowchart of non parametric attention mechanism

在对神经元的研究发现,具有丰富信息的神经元会表现出与周围神经元不同的放电模式,激活神经元因可以抑制周围神经元的放电模式而具有更高的重要性。无参数注意力模块重视发掘每个神经元的重要性,通过不断学习判别更多的神经元,为每一个神经元分配唯一权重。为了正确评估神经元的重要性,该模块运用了一个新的能量函数,能量函数得到的值越低,该神经元的重要性越高。能量函数为

$$e_i = \frac{1}{M-1} \sum_{i=0}^{M-1} [-1 - (w_i x_i + b_i)]^2 + [1 - (w_i t + b_i)]^2 + \lambda w_i^2, \quad (5)$$

式中, e_i 为能量函数值, w_i 和 b_i 分别为变换的权重和偏置值, x_i 为标量, M 为通道上神经元的数量。

3 试验与结果分析

3.1 数据集和评估方法

本研究使用的是模仿工业生产场景的 MVTEC-AD 数据集^[20]。该数据集包含 5 354 张不同类型的高分辨率彩色图像,分辨率为 700×700 或 $1\,024 \times 1\,024$,数据集分为训练集和测试集。其中训练集共有 3 629 张正常样本图片,测试集共有近 2 000 张正常样本和 70 多种不同类型的缺陷样本。该数据集包括 15 个不同工业类别,其中有 5 类为纹理类图像、10 类为物体类图像。

本试验与以往无监督缺陷检测采用相似指标,利用不同模型曲线下面积 A_{ROC} 作为评估指标。曲线下的面积越接近 1,表示模型性能越好、缺陷的检测及定位效果越好^[21]。

3.2 试验环境

本研究代码基于 Ubuntu16.04、Cuda10、Python 3.6 的操作系统,硬件配置为 1 块 GPU: GTX 2080Ti

(显存:11 GB)。本研究利用 ImageNet 预训练权重参数对模型初始化,教师网络采用 VGG16 网络,学生网络使用 VGG 的简化模型。在训练数据集时,使用 Adam 优化器来优化模型,学习率为 0.001,批量大小为 64 字节,损失函数的参数设置为 0.5,训练周期为 601 次。

3.3 比较试验

为了验证本研究方法的有效性,通过与以下方法进行对比试验:对抗性视觉不规则检测方法 (adversarial visual irregularity detection, AVID)^[22]、无监督异常检测方法 (unsupervised anomaly detection with generative adversarial networks to guide marker discovery, AnoGAN)^[23]、利用 SSIM 对图像结构信息进行重构的自编码器改进无监督缺陷分割的方法 (improving unsupervised defect segmentation by applying structural similarity to autoencoders, AE-SSIM)^[24]、深度分类方法 (deep one-class classification, DSVDD)^[25]、正态数据流形异常定位投影方法 (iterative energy-based projection on a normal data manifold for anomaly localization, VAE)^[26]、多分辨率知识提取方法 (multiresolution knowledge distillation for anomaly detection, MKDA)^[27]、纳米纤维材料异常检测方法 (anomaly detection in nanofibrous materials by cnn-based self-similarity, CNN)^[28]。

本研究以 A_{ROC} 为评价指标,分析该知识蒸馏模型在检测和定位两方面的效果,列出了 MVTEC-AD 数据集内的 15 个工业类别的检测结果。缺陷定位热力图如图 5 所示,各方法对比结果如表 1 和表 2 所示。表中加黑数据为各数据集的平均值。

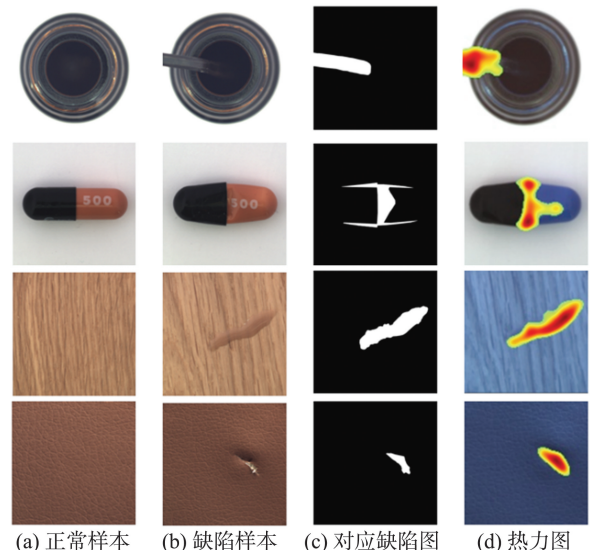


图5 缺陷定位热力图

Fig.5 Defect localization heatmap

表1 基于MVTec-AD数据集的 A_{ROC} 检测分数对比
Table 1 Comparison of A_{ROC} detection scores based on the MVTec-AD dataset

单位:%

类别	A_{ROC}						
	AVID	AE-SSIM	AnoGAN	DSVDD	VAE	MKDA	MKDA-CA
瓶子	88.2	88.1	69.7	86.4	86.1	99.4	99.5
栗子	86.3	54.5	50.4	71.6	74.4	98.1	98.7
胶囊	85.8	61.2	58.3	69.9	86.8	80.2	81.7
螺母	63.1	54.6	50.1	75.6	78.9	73.2	73.4
皮革	58.6	46.4	52.2	73.1	71.3	94.8	98.6
药片	86.2	60.3	62.7	77.4	80.7	82.1	86.9
木头	83.6	83.4	68.9	87.8	89.6	94.1	96.0
地毯	70.5	67.1	49.4	54.4	67.8	78.6	82.2
瓷砖	66.4	52.2	51.1	81.6	81.4	90.9	92.9
铁丝网	59.9	69.4	51.5	59.3	83.3	77.4	88.5
电缆	64.1	61.3	53.4	71.8	56.7	88.8	90.1
晶体管	58.2	52.8	67.6	65.9	70.8	85.3	85.4
牙刷	73.3	74.7	57.1	70.4	89.5	91.8	95.6
螺丝	66.5	51.5	35.0	64.8	71.2	82.7	96.9
弹簧	84.7	80.6	59.6	74.7	67.3	92.9	95.7
平均	73.5	63.4	55.4	72.6	77.5	87.4	90.9

表2 基于MVTec-AD数据集的 A_{ROC} 定位分数对比
Table 2 Comparison of A_{ROC} localization scores based on the MVTec-AD dataset

单位:%

类别	A_{ROC}						
	AVID	AE-SSIM	AnoGAN	CNN	VAE	MKDA	MKDA-CA
瓶子	86.2	93.7	86.9	78.8	92.2	95.2	94.1
栗子	95.3	97.1	87.4	72.9	97.6	94.6	96.3
胶囊	88.6	94.5	84.2	84.4	91.7	95.6	95.7
螺母	86.0	89.4	76.5	82.6	90.7	82.5	82.9
皮革	78.8	78.0	64.7	87.5	92.5	98.0	98.6
药片	91.4	91.4	87.9	68.8	93.0	89.6	91.5
木头	73.3	73.5	62.1	91.7	83.8	84.8	88.9
地毯	59.5	87.1	54.4	72.8	73.5	95.6	95.7
瓷砖	51.4	59.9	50.7	93.4	65.4	81.6	81.7
铁丝网	90.9	94.5	58.8	59.8	96.1	89.0	91.4
电缆	86.7	82.7	78.6	79.6	91.0	80.9	81.6
晶体管	86.2	90.2	80.8	66.5	91.9	76.1	72.0
牙刷	93.1	92.8	90.2	77.7	98.5	96.0	96.1
螺丝	96.0	96.4	80.5	87.3	94.5	95.6	97.1
弹簧	77.3	88.2	78.1	76.8	86.9	93.9	95.8
平均	82.4	87.4	74.5	78.6	89.3	89.9	90.7

MKDA方法是基于知识蒸馏进行缺陷检测的效果最好且最新的方法,因此本研究 and MKDA 进行比较。从表1可以看出改进后的知识蒸馏模型在MVTec-AD数据集的15个不同类别缺陷上,平均缺陷检测分数比MKDA方法提升3.5%,在单个类别的数据集上平均缺陷检测分数最高提升14.2%。其平均缺陷定位分数比MKDA方法提升0.8%,在单个类别的数据集上其平均缺陷检测分数最高提升4.1%。该试验证明,在知识蒸馏模型中加入了自监督预测卷积模块和无参数注意力机制模块

后,可以更好地保留特征信息,提升缺陷检测模型的检测效果,缺陷定位能力也有很大提升。综合以上分析,可以看出本研究的方法具有优越性。

3.4 有效性试验

为了验证本研究使用的无参数注意力模块能有效提高该模型的缺陷检测和定位能力,将无参数注意力模块与SE注意力机制^[15]、ECA注意力机制^[16]、MS-CAM注意力机制^[29]三种方法进行对比试验,以 A_{ROC} 检测分数作为试验指标,比较MVTec-AD数据集的15个不同工业缺陷类型的

平均精确度,结果如图6所示。从结果可知,在知识蒸馏模型加入了无参数注意力模块后,其缺陷检测的 A_{ROC} 值高于其他注意力机制。这说明自监督预测卷积模块的重构功能更加贴合基于知识蒸馏的无监督模型,使学生网络可以更好地学习教师网络传递的正样本知识,在测试阶段提高区分正常样本和异常样本的能力,最终提高了检测缺陷的能力。

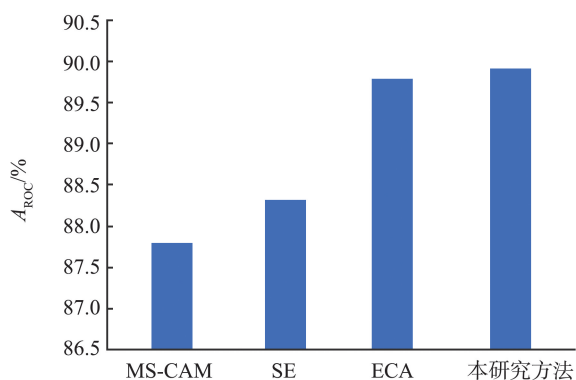


图6 各注意力机制的结果对比

Fig.6 Comparison of results of various attention mechanism

自监督预测卷积模块适用于绝大部分的缺陷检测模型,能够提升对缺陷特征提取能力。为了分析自监督预测卷积模块在本研究模型中的作用,在MVTec-AD数据集的15个不同缺陷类别上进行试验,结果如图7所示。从图7可以看出,在知识蒸馏模型加入自监督预测卷积模块后,对上文15个不同缺陷类别的检测效果总体有所提高,尤其是在螺丝缺陷的数据集上检测分数提升最高,提高约10.8%。该试验表明在知识蒸馏模型上加入自监督预测卷积模块可以提高缺陷特征的表达能力,从而提高了模型缺陷检测能力,证明了该模块的有效性。

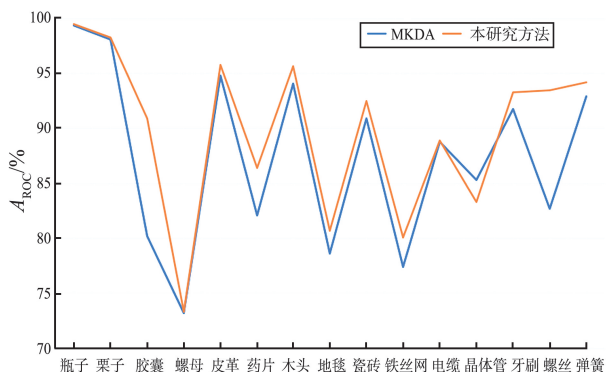


图7 自监督预测卷积模块应用效果

Fig.7 Application effect of self supervised predictive convolution module

为了证明本研究方法具有更好的缺陷检测能力,对本研究模型进行消融试验。分别对只加入

了自监督预测卷积模块、只加入无参数注意力模块、加入自监督预测卷积模块和无参数注意力模块的知识蒸馏模型,以 A_{ROC} 检测分数作为试验指标,对MVTec-AD数据集的15个不同工业缺陷类型的检测平均精确度进行比较,消融结果如图8所示。从图8可以看出,在模型只加入自监督预测卷积模块和只加入无参数注意力模块时, A_{ROC} 检测分数值略微提升。然而在模型中加入自监督预测卷积模块和无参数注意力模块时,该知识蒸馏模型的 A_{ROC} 值达到90.9%,取得比之前更好的检测效果。证明这两个模块在模型中有很好的融合性,均可提高模型的性能,在保证效率的情况下提高了缺陷检测的能力,证明本方法的有效性。

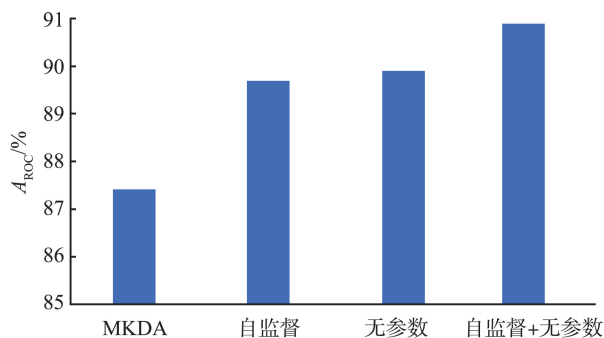


图8 消融试验结果

Fig.8 Figure of ablation experiment results

4 结论

本研究提出一种基于知识蒸馏模型的工业品表面缺陷检测方法。为了提升特征提取能力和缺陷定位能力,在提高模型缺陷检测精确度的情况下保证工业品表面缺陷检测的效率,在模型中加入了具有重构功能的自监督预测卷积模块和能够在通道维度和空间维度加强信息的无参数注意力机制。最后在MVTec-AD数据集的15个不同工业缺陷上验证该方法,通过和先进方法的比较和消融试验,证明了改进后的缺陷检测模型取得更好的检测和定位效果。在之后的工作中,可以针对以下的这几个方面进行研究:(1)采用合适的方法扩大缺陷样本或者采用不需要缺陷样本的无监督缺陷检测方法或弱监督缺陷检测方法进行检测,增强模型泛化性。(2)在研究缺陷检测方法时,不仅要考虑到如何提升算法对缺陷分类定位的能力,还要考虑到适当减少缺陷检测的参数数量,提高缺陷检测的效率。(3)要发挥工业品表面缺陷检测模型的实际应用价值,将模型投入到实际工业生产中去。

参考文献:

- [1] 孙博言,王洪元,刘乾,等. 基于多尺度和注意力机制的混合监督金属表面缺陷检测[J]. 智能系统学报, 2023, 18(4): 886-893.
SUN Boyan, WANG Hongyuan, LIU Qian, et al. Hybrid supervised metal surface defect detection based on multi-scale and attention[J]. CAAI Transactions on Intelligent Systems, 2023, 18(4): 886-893.
- [2] 李健,李华,胡翔坤,等. 基于深度学习的表面缺陷检测技术研究进展[J]. 计算机集成制造系统, 2024, 30(3): 774-790.
LI Jian, LI Hua, HU Xiangkun, et al. Research progress of surface defect detection technology based on deep learning[J]. Computer Integrated Manufacturing System, 2024, 30(3): 774-790.
- [3] CHUA L O. CNN: a paradigm for complexity [M]. Santiago, Chile: World Scientific Publishing Co. Pte Ltd., 1999: 529-837.
- [4] NG H F. Automatic thresholding for defect detection[J]. Pattern Recognition Letters, 2006, 27(14): 1644-1649.
- [5] BROMLEY J, GUYON I, LECUN Y, et al. Signature verification using a "siamese" time delay neural network [J]. Advances in Neural Information Processing Systems, 1993, 7(4): 669-688.
- [6] WANG Y S, YAO H X, ZHAO S C. Auto-encoder based dimensionality reduction [J]. Neurocomputing, 2016, 184: 232-242.
- [7] ZHENG X Q, WANG H C, CHEN J, et al. A generic semi-supervised deep learning-based approach for automated surface inspection[J]. IEEE Access, 2020, 8: 114088-114099.
- [8] 成科扬,丁杨柳,詹永照,等. 基于回顾蒸馏学习的无监督工业品缺陷检测方法[J]. 南京大学学报(自然科学版), 2022, 58(6): 1030-1040.
CHENG Keyang, DING Yangliu, ZHAN Yongzhao, et al. Unsupervised industrial product defect detection method based on retrospective distillation learning [J]. Journal of Nanjing University (Natural Science Edition), 2022, 58(6): 1030-1040.
- [9] BUCILUĂ C, CARUANA R, NICULESCU-MIZIL A. Model compression[C]//Proceedings of the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. Philadelphia, USA. ACM, 2006: 535-541.
- [10] HINTON G, VINYALS O, DEAN J. Distilling the knowledge in a neural network[EB/OL]. (2015-03-09) [2024-03-12]. <https://arxiv.org/abs/1503.02531>
- [11] GOU J P, YU B S, MAYBANK S J, et al. Knowledge distillation: a survey [J]. International Journal of Computer Vision, 2021, 129(6): 1789-1819.
- [12] YU L, YAZICI V O, LIU X L, et al. Learning metrics from teachers: compact networks for image embedding [C]//Knowledge Distillation: A Survey. Long Beach, USA: IEEE, 2019: 2907-2916.
- [13] PARK W, KIM D, LU Y, et al. Relational knowledge distillation [C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Long Beach, USA: IEEE, 2019: 3962-3971.
- [14] CHEN M H, LAINA I, VEDALDI A. Training-free layout control with cross-attention guidance [C]//2024 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV). Waikoloa, Hawaii, USA: IEEE, 2024: 5331-5341.
- [15] HU J, SHEN L, SUN G. Squeeze-and-excitation networks [C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA: IEEE, 2018: 7132-7141.
- [16] WANG Q L, WU B G, ZHU P F, et al. ECA-net: efficient channel attention for deep convolutional neural networks [C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2020. Seattle, WA, USA: IEEE, 2020: 13-19.
- [17] WOO S, PARK J, LEE J Y, et al. CBAM: convolutional block attention module [C]// Computer Vision-ECCV 2018. Munich, Germany: Springer International Publishing, 2018: 3-19.
- [18] LI X, WANG W H, HU X L, et al. Selective kernel networks [C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Long Beach, USA: IEEE, 2019: 510-519.
- [19] DENG J, DONG W, SOCHER R, et al. ImageNet: a large-scale hierarchical image database [C]//2009 IEEE Conference on Computer Vision and Pattern Recognition. Miami, USA: IEEE, 2009: 248-255.
- [20] BERGMANN P, FAUSER M, SATTLEGGER D, et al. MVTEC AD: a comprehensive real-world dataset for unsupervised anomaly detection [C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Long Beach, USA: IEEE, 2019: 9592-9600.
- [21] HENDRYCKS D, GIMPEL K, NOVELLO P, et al. A baseline for detecting misclassified and out-of-distribution examples in neural networks [EB/OL]. (2016-10-07) [2024-03-12]. <https://arxiv.org/abs/1610.02136v3>
- [22] SABOKROU M, POURREZA M, FAYYAZ M, et al. AVID: adversarial visual irregularity detection [C]// Computer Vision - ACCV 2018. Perth, Australia: Springer International Publishing, 2019: 488-505.