

DOI:10.19651/j.cnki.emt.2519279

渐进式融合知识蒸馏的工业异常检测方法^{*}

赵智斌 王浩 朱建光 薛丹

(沈阳工业大学信息科学与工程学院 沈阳 110872)

摘要:为解决知识蒸馏异常检测模型中出现知识冗余的问题,本文提出一种渐进式知识融合蒸馏的工业检测方法。本文采用反向知识蒸馏作为检测模型的基础网络,虽然反向知识蒸馏能阻止异常表征传播至学生网络,但是该方法下学生网络获取的知识不仅复杂且冗余,难以保证学生网络能够重建出对应的浅层表征。为转化教师网络输出的高层次复杂且冗余的信息,提出了渐进式知识融合机制。为转化复杂表征,该机制将基础几何知识逐步迁移至深层语义知识,促进学生网络有效地学习特征表示;为剔除多级教师网络输出的信息中的冗余知识,该机制采用可学习的特征权重分配方式,促进了学生网络重建,实现了模型异常检测能力的提升。该方法在 MVTec AD 数据集上进行实验,实验结果显示,在数据集各个类别下的评价指标 AUROC-image 为 99%,AUROC-pixel 为 97.9%,AUPRO 为 94.8%,超越了绝大部分目前主流的检测模型,验证了方法的有效性和优越性。

关键词: 知识蒸馏;异常检测;渐进式融合

中图分类号: TN98 **文献标识码:** A **国家标准学科分类代码:** 520.2040

Asymptotic fusion knowledge distillation for industrial anomaly detection

Zhao Zhibin Wang Hao Zhu Jianguang Xue Dan

(School of Information Science and Engineering, Shenyang University of Technology, Shenyang 110872, China)

Abstract: To address the issue of knowledge redundancy in knowledge distillation-based anomaly detection models, this paper proposes an industrial detection method based on asymptotic knowledge fusion distillation. This paper adopts reverse knowledge distillation as the backbone network of the detection model. Although reverse knowledge distillation can prevent the propagation of abnormal representations to the student network, the knowledge acquired by the student network under this method is not only complex but also redundant, making it difficult to ensure that the student network can reconstruct the corresponding shallow representations. To transform the high-level, complex, and redundant information output by the teacher network, an asymptotic knowledge fusion mechanism is proposed. For transforming complex representations, this mechanism gradually transfers basic geometric knowledge to deep semantic knowledge, enabling the student network to learn feature representations effectively. For eliminating redundant knowledge from the information output by multi-level teacher networks, this mechanism adopts a learnable feature weight assignment method, which promotes the reconstruction of the student network and improves the anomaly detection capability of the model. Experiments are conducted on the MVTec AD dataset. The results show that the evaluation metrics under all categories of the dataset are 99% for AUROC-image, 97.9% for AUROC-pixel, and 94.8% for AUPRO, which outperform most of the current mainstream detection models, verifying the effectiveness and superiority of the proposed method.

Keywords: knowledge distillation; anomaly detection; asymptotic fusion

0 引言

异常检测是计算机视觉与图像处理领域的重要研究方

向,其核心任务是通过算法自动识别并精确分割输入样本中与正常图像模式存在统计差异的异常区域^[1]。该技术建立有效的特征表示模型,以实现各类表面异常的检测、定

收稿日期:2025-07-03

^{*} 基金项目:辽宁省教育厅基本科研面上项目(JYTMS20231213)、辽宁省科技计划联合计划面上项目(2024-MSLH-349)、国家自然科学基金(62301339)、辽宁省科技厅联合计划项目(2023JH2/101700279)资助

位与分割。异常检测已应用于各种工业检测场景中,如医学影像、产品质检和部件维护等^[2-5]。在工业检测中,由于异常样本稀缺难以获取齐全,这使得主流的监督学习模型无法应用于工业检测。随着无监督学习方法的快速发展及其在计算机视觉领域的关注度持续提升,越来越多的研究人员开始致力于开发适用工业场景的无监督视觉检测模型,以解决异常样本不足这一关键挑战。

在工业异常检测领域,无监督学习模型主要可分为两大类方法:基于特征和基于重建的方法。基于重建的方法是使用正常数据训练生成模型,使得生成模型学习到正常数据的模式和特征^[6-7]。通过训练好的生成模型将缺陷区域重建为正常区域,之后将重建后的图片和原图片作比较,进而检测缺陷。主流的重建模型是生成对抗网络(generate adverse network, GAN)和扩散模型。GAN 不适合重建大面积的异常区域,导致检测精度下降。扩散模型在异常检测任务中需要执行数百至上千次加噪和去噪迭代才能完成图像重建,这一特性导致其计算效率低下。尽管已有研究通过减少迭代步数来优化推理速度,但实验表明步数缩减会直接导致重建质量下降,形成计算效率与检测性能之间的固有矛盾。考虑到工业检测场景对实时性的严格要求,当前基于扩散模型的方法难以满足实际应用需求。因此,从计算效率和实时性两个关键指标评估,扩散模型尚不具备工业级异常检测的适用性。

基于特征的检测模型主要通过计算图片中的异常特征和正常特征分布之间的偏离来检测异常^[8-9]。近年来,基于特征表示的知识蒸馏方法在工业场景的异常检测与定位任务中展现出卓越的检测性能。这类算法模型中包含一对师生网络,教师网络在大规模数据集(如:ImageNet)上进行预训练,学生网络拥有与教师网络同样的或简化的网络结构作为学生网络,并仅在数据集中学习正常样本的特征分布。但同向知识蒸馏模型中教师网络与学生网络结构相似以及模型中相同知识流的限制,致使同向知识蒸馏方式在检测性能上存在一定的局限。Wu 等^[10]设计了具有差异性的师生网络结构,使得同向师生网络的局限性得到缓解。Wang 等^[11]通过特征对比度干扰抑制来扩大前景和背景纹理之间的特征距离,以获得更强大的背景表示。Ding 等^[12]提出一种预训练特征提取与激活学习框架,提高了模型的小样本学习能力和模型可解释性。虽然这些针对同向知识蒸馏网络模型的改进提升了模型对缺陷的识别能力并削弱了背景噪声对检测的影响,但是并没有解决同向知识蒸馏中因相同数据流而导致的检测性能的局限性。近年来,出现了一种新颖的知识蒸馏范式——反向知识蒸馏^[13]。该方法通过编码器-解码器的网络结构,提高师生网络在未知、非正常分布样本上的表示多样性,解决了传统同向知识蒸馏在测试过程中学生网络可能与教师网络无法产生偏移的问题。但是在该框架下,学生网络接收教师网络输出的高复杂度冗余特征信息,并基于此复原包含丰富浅层特征

的内容,这一过程存在显著的表征学习局限,致使学生网络难以完整捕捉教师网络输出的正常样本特征分布规律,最终导致其建模能力受到严重制约。虽然,已有针对反向知识蒸馏框架而做出的改进工作^[14],使得模型具有较强的缺陷特征识别能力,但并没有解决反向知识蒸馏中存在复杂冗余信息的问题,进而导致模型无法识别微小且稀疏的缺陷区域。

为解决上述问题,本文提出了一种渐进式融合知识蒸馏的工业异常检测方法。具体而言,采用反向知识蒸馏作为模型框架,用于阻止异常表征传入学生网络,防止学生网络与教师网络产生偏移,提升知识蒸馏效果;同时,为转化教师网络输出的高层次语义信息和冗余知识,设计了一种渐进式知识融合机制(asymptotic knowledge fusion, AKF),该机制逐步将浅层几何特征精准迁移至语义相近的深层特征空间,实现了基础几何表征与高级语义信息的有机统一,这种跨层特征协同机制有效提升了学生网络的浅层表征重构能力;之后采用可学习的特征权重分配方式,实现了多层级特征的智能融合,并通过动态调节各层特征贡献度,显著抑制了深层特征的冗余信息,从而构建出更具判别性的融合特征表示。

1 相关方法

非监督学习异常检测包括基于重建的方法和基于特征的方法。基于重建的方法主要包括 GAN、变分自编码器(variational autoencoder, VAE)和扩散模型等。基于重建的方法通过将异常区域重建为正常区域,进而和原图像进行逐个像素对比检测出异常区域。Yan 等^[15]提出了一种基于语义上下文的无监督异常检测网络,通过从正常样本中学习语义上下文来实现异常检测。Wyatt 等^[16]首次将扩散模型应用在医疗异常检测领域上,该方法通过引入简型噪声代替高斯噪声,使得扩散模型的检测效果得到提升。He 等^[17]提出了一种多类别异常检测框架,该方法引入了扩散模型去除异常特征,通过潜空间语义引导网络和空间感知特征融合模块促进图像重建和提升重建精度。尽管这些方法具备较强的缺陷定位能力,但仍存在一定的局限性,如 GAN 的训练过程易发散,扩散模型不能平衡检测时长和检测精度之间的问题,而 VAE 在像素级别的重建上精度较低,使得这些方法并不适合实际的工业检测场景。

基于特征的方法是通过将图片信息转化到特征空间进行异常检测。Roth 等^[18]提出了一种基于图像块的异常检测模型,该方法通过预训练深度神经网络提取多尺度特征,并利用内存库(memory bank)存储正常样本的特征分布,实现对复杂工业场景中细微缺陷的高精度检测,但该方法对设备硬件要求过高,需要较大的内存空间。魏金洋等^[19]提出了一种改进深度特征重建的无监督缺陷检测方法,在特征重建算法中引入空洞金字塔池化,在不增加参数数量的情况下多尺度捕捉上下文信息,提高模型对不同尺寸缺陷

的检测能力。然而这类方法对工业检测设备的计算性能要求过高且模型参数量较大,不适用于工业检测。

知识蒸馏是近年来提出的一种具有轻量化参数的缺陷检测模型,该模型网络结构简洁且实时性高能满足工业缺陷检测的需求。冯志远等^[20]提出一种基于判别式增强的自监督蒸馏学习方法,通过学习正常样本和伪异常样本的特征得到更加符合真实异常分布的异常分割决策边界。Rudolph等^[21]提出了非对称知识蒸馏异常检测方法,该方法采用可逆归一化流对教师网络进行密度估计,采用常规前馈网络作为学生网络。这些方法虽具有轻量的参数和较高的实时性,但这些方法所使用的同向知识蒸馏框架中网络结构相似和数据流相同,限制了同向网络的检测能力。

反向知识蒸馏框架通过编码器-解码器的网络结构,在该

网络结构中,知识从教师的深层提取到浅层,即首先将高级的语义知识传授给学生,提高师生网络在未知、非分布样本上的表示多样性。这一结构避免了网络结构相似和数据流相同所导致的限制^[13]。但该框架中教师网络输出的复杂冗余信息阻碍了学生网络的正常学习无异常样本的特征空间。

为解决反向知识蒸馏框架中的复杂且冗余的信息,本文方法结合渐进式知识融合机制和单类嵌入机制转化冗余信息并增强模型对微小缺陷的识别能力。

2 方法

本文方法采用反向知识蒸馏作为模型框架^[13],并设计了一种渐进式知识融合机制来消除教师网络输出的冗余信息。本文方法如图1所示。

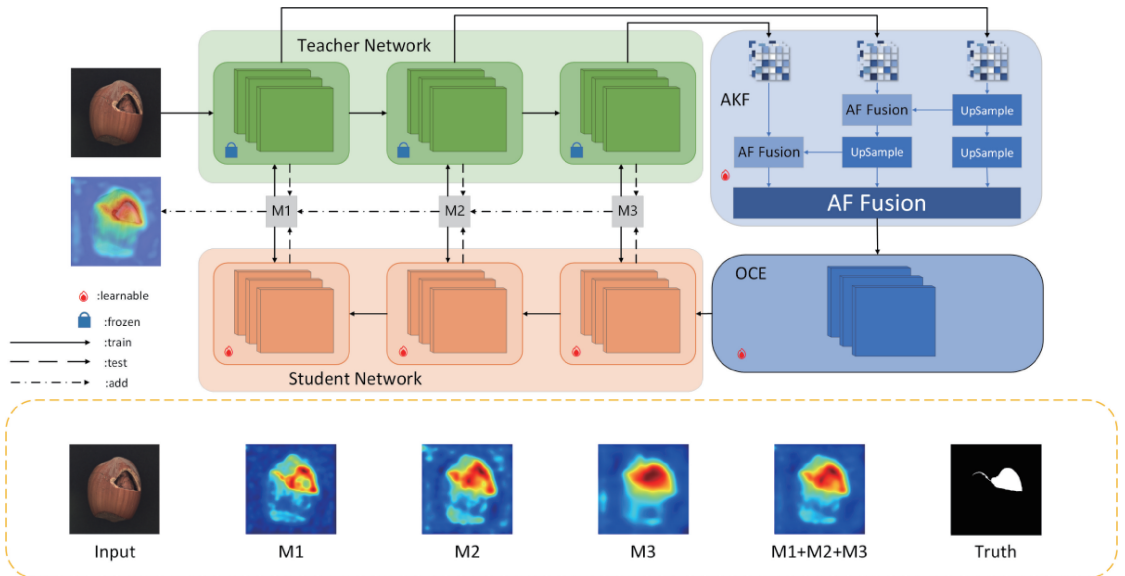


图1 渐进式融合知识蒸馏的工业异常检测模型

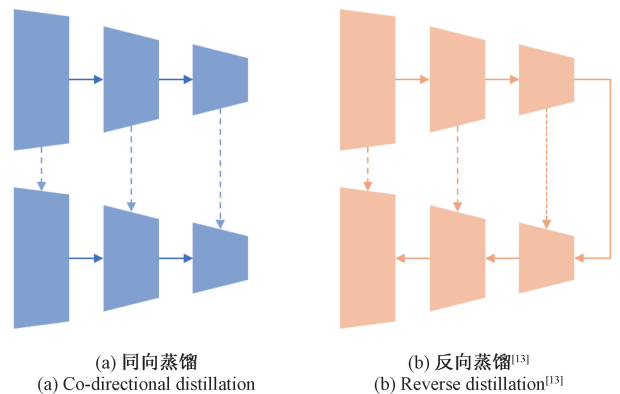
Fig. 1 Industrial anomaly detection model based on asymptotic knowledge fusion distillation

2.1 反向知识蒸馏

在知识蒸馏(knowledge distillation, KD)框架中,知识通过师生(teacher-student, T-S)网络对进行传递。在无监督异常检测场景下,由于学生网络仅接触正常样本训练,当输入异常时,其输出的表征将与教师网络产生差异,进而检测出异常区域。但是在同向知识蒸馏框架下,学生网络和教师网络具有相同结构和相同方向的数据流,该框架在检测过程中可能导致模型对异常表征的差异消失,学生网络无法和教师网络产生偏移,从而导致异常检测失效^[13]。

为避免上述问题,本文采用反向知识蒸馏框架^[13]。相比于同向知识蒸馏,反向知识蒸馏将学生网络设计为与教师网络对称且反向的解码器结构,将预训练教师网络输出的知识反向输入到学生网络。该框架提升了模型对正常样本知识分布外的表征多样性。同时,这种结构有助于去除学生网络对异常表征的响应,增强了教师网络和学生网

络之间的差异,提升了模型针对异常的检测能力。现有知识蒸馏结构如图2所示。



(a) 同向蒸馏

(a) Co-directional distillation

(b) 反向蒸馏^[13]

(b) Reverse distillation^[13]

图2 现有知识蒸馏结构

Fig. 2 Existing knowledge distillation structure

2.2 渐进式知识融合机制

在目标检测领域, Yang 等^[22]提出了一种新型特征融合结构, 该结构通过浅层特征和深层特征双向融合来避免浅层特征在模型中流失。该方法能够有效地融合模型需要的特征, 但是其直接应用于反向知识蒸馏框架中时, 会加剧教师网络和学术网络之间特征信息复杂度高且冗余的问题。由此在文献[22]所设计的方法基础上, 提出了一种新型的渐进式知识融合机制 (asymptotic knowledge fusion, AKF)。渐进式知识融合机制如图 3(a) 所示。AKF 在教师网络自底向上的特征提取过程中, 采用渐进式方式单向整合低层特征、高层特征及顶层特征。具体而言, AKF 将浅层特征和与其相邻的深层特征逐渐单向加权融合。由于, 顶层特征和底层特征之间的差异要大于顶

层特征和深层特征, 这导致非相邻层级特征直接融合效果不佳。因此, 直接使用各层特征进行特征融合会造成学生网络学习不充分。并且在特征融合过程中, 不同层级特征在特定位置上可能存在不同物体特征的矛盾冲突, 这也导致逐个元素相加的融合方式会造成学生网络训练不稳定。而 AKF 采用知识逐层迁移的方式, 这种融合过程使得代表简单几何图形的知识迁移到融合后的深层知识, 进而使深层知识同时具有高级语义信息和基础特征, 从而缓解了依据深层语义特征信息难以解构浅层信息的问题。区别于文献[22]中的双向特征融合方式, AKF 方法通过单向渐进式的浅层到深层特征融合机制, 在保持特征表达能力的同时, 避免了文献[22]中双向特征融合方法可能加剧层级间特征信息冗余问题。

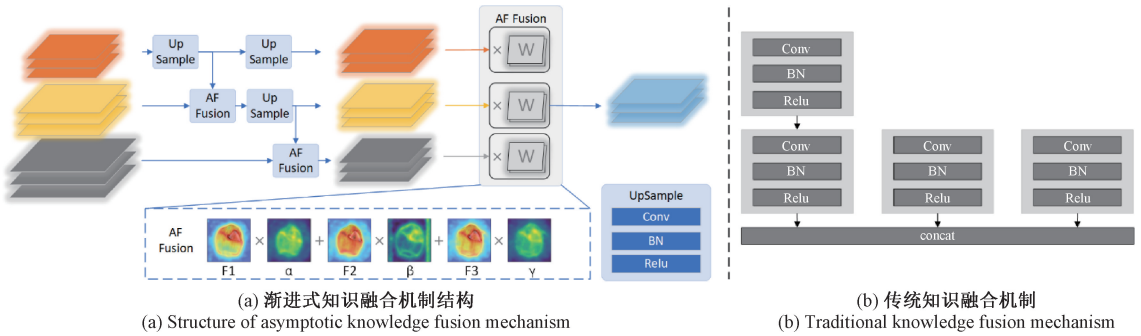


图 3 渐进式知识融合机制和传统知识融合机制的网络结构

Fig. 3 Network structures of asymptotic knowledge fusion mechanism and traditional knowledge fusion mechanism

传统多尺度融合机制其本质是通过 Concat 操作将特征提取网络不同层级输出的特征信息有机整合, 从而突破单一尺度特征的局限性。然而, 各个层级的知识直接聚合会出现多个代表同一含义的冗余知识, 并且可能将一些有利于学生网络训练的特征信息去除, 进而导致模型检测效果变差。同时, 该结构对所有空间位置采用统一融合策略, 导致浅层高分辨率细节与深层语义信息的冲突未被有效解决。为解决上述问题, AKF 机制通过为每个层级的特征图的每个像素位置分配独立的可训练权重参数, 使模型能够根据目标的空间分布与尺度特征, 动态地调整各层级特征的贡献度。该机制促进了该位置目标感知最有效的特征层级主导信息表达, 抑制了冗余知识对学生网络训练结果的影响, 同时也避免了传统方法中“有益特征被稀释、有害特征被保留”的问题。

其中, $f_n, W_n \in R^{C \times H \times W}$ 分别表示教师网络各个特征层次的特征信息和对应的权重, n 代表教师网络的第 n 层输出, N 代表教师网络共有 N 层。 f'_n 代表相邻层特征加权融合后的结果, $\overline{f'_n}$ 表示 UpSample 块调整特征尺寸后的输出的, f_{out} 表示 AKF 机制输出的结果。

2.3 单类嵌入机制

为进一步将教师网络输出的复杂且冗余的特征信息转化为易于学生网络训练的信息, 在渐进式知识融合机制之后引入了残差块作为单类嵌入机制 (one-class embedding, OCE)^[13]。该模块通过将深层特征信息映射到浅层并且去除冗余的特征信息, 使得学生网络能够更好的学习到正常样本的特征空间。

3 实验结果及分析

3.1 实验环境和参数配置

本文所有实验均在单张显卡 RTX4090GPU(24 G 显存)上运行。batch_size 设置为 32。训练阶段使用 Adam 优化器, 初始学习率设置 learning_rate=0.0001, 训练轮数为 100 个 Epoch。

模型配置中, 选取 wide_resnet_50_2 为特征提取网络, 教师网络为在 ImageNet 数据集上训练好的网络, 学生网络与教师网络结构相同, 为未经过训练的初始化网络。

选择 wide_resnet_50_2 特征提取网络的前3层作为教师网络和学生网络的多尺度输出特征。同时将教师网络和学生网络各层级之间的余弦相似度作为模型的损失函数,使得学生网络能充分地学习到正常样本之间的特征分布。输入图像经预处理调整为 256 pixel×256 pixel。

3.2 数据集及评价指标

本文所提出的算法需要在小、有限的区域中检测出异常。因此,为了更好的实现这一目标,本文选用 MVTec AD 数据集,模拟真实世界的工业检测场景^[1,23]。MVTec AD 数据集包含 5 354 张来自不同领域的 5 种纹理和 10 种物体的高分辨率图像,一共有 73 种类型的缺陷,为每张缺陷图像(一共有 1 888 张)提供了像素级精确的标签。数据集如图 4 所示。

在异常检测领域,准确评估算法性能是至关重要的,其中检测准确性是核心考量因素之一^[24-25]。本文选用 AUROC(area under the receiver operating characteristic curve)和 PRO(per-region-overlap)作为模型的评价指标。AUROC 为 AUC-ROC 曲线下面积,为目前异常检测领域最常用的评价指标,适用于二分类问题,AUC-ROC 值越大,反应模型预测正确率越高。

PRO 为 PR 曲线下面积。PR 曲线反映精确率和召回率的关系。PRO 用于评估模型在图像中定位异常的能力。

3.3 对比实验

本文方法与同为师生网络的异常检测方法(包括 RD^[13], US^[26], MKD^[27], EMSF^[14], MFR^[28], EFAD^[11] 和 PFEAL^[12],其中 RD 和 EMSF 同为反向知识蒸馏网络,MFR、EFAD 和 PFEAL 为同向知识蒸馏网络,这些检测模型均为近年所提出的优秀检测模型)和其他的异常检测方法(包括 DRAEM^[29], PaDiM^[8] 和 EdgRec^[30])进行了对比实验。

各类异常检测方法的 AUROC-image 检测指标如表 1 所示。相较于基准网络 RD,本文方法提升了 0.5%,并在 6 个类别上 AUROC-image 的分数达到了 100%,验证了本文方法的优越性。本文方法在各个类别上的 AUROC-image 平均值达到了 99.0% 超越绝大部分缺陷检测方法,并且在纹理类别上超越了绝大多数其他方法,这也验证了本文提出的渐进式空间特征融合机制的有效性。

各类检测方法的 AUROC-pixel 和 PRO 检测指标如表 2 所示。在这两项指标上,本文方法的均值为 97.9% 和 94.8% 优于绝大多数其他缺陷检测方法。其中在 PRO 这一指标上,本文方法有 7 个类别达到了最优和次优,相较于基准方法 RD 提升了 0.9%。在多个富含微小细节的物体类别(如螺丝钉、拉链等)上,本文方法优于同为反向知识蒸馏的检测网络 EMSF,这证明了本文方法在面对异常时识别效果明显,并且在含有细微缺陷的类别上的检测效果要优于其他同类型异常检测方法。

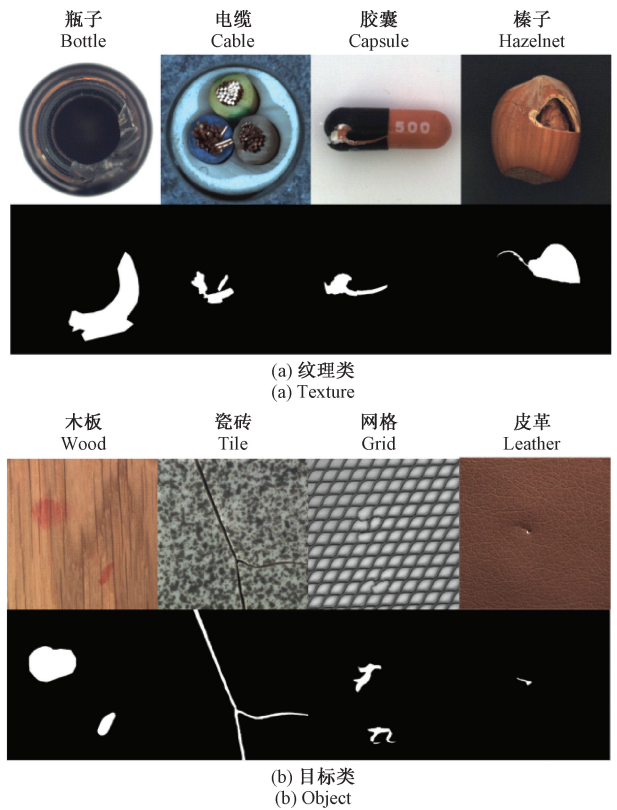


图 4 MVTec AD 样例

Fig. 4 MVTec AD samples

3.4 消融实验

为了验证本文方法中所提到的各个模块的有效性,本文进行了消融实验进行对比。表 3 比较了 4 种情况:1)去除 AKF 和 OCE,仅使用传统知识融合;2)使用渐进式知识融合机制和 OCE;3)去除 OCE,仅使用 AKF;4)使用 AKF 和 OCE。在实验过程中为保证尺度统一,先后对特征信息进行卷积核为 3×3 的卷积操作, Batch_Normalization 归一化和 ReLU 激活。

由表 3 的结果可知,AKF 逐渐将深层特征中学生网络难以学习的特征稀释,并将其逐步融入到易于学生网络学习的浅层特征,再结合 OCE 进一步转化特征,这使得输入到学生网络的特征信息不再冗余并易于重构。而单一的传统知识融合机制和 OCE 虽然能一定程度上转化冗余信息,但是深层特征中学生网络难以学习的特征所占的权重依然较大,使得学生网络的训练效果不佳。

由图 5 的可视化效果分析,本文方法在各个类别上的缺陷识别效果均优于传统知识融合,并能够更加准确的识别出缺陷区域。在螺丝钉这个类别上,螺丝纹上的异常缺陷和正常螺纹之间的差异较小,传统知识融合难以从正常的螺纹中识别出异常螺纹。从传统知识融合识别坚果的效果分析,传统知识融合容易将坚果的果蒂识别为异常区域,而本文方法能够正确的分割出异常区域。在药片这一类上,传统知识融合易受药片上的正常斑点和背景噪声影

表 1 各类别异常检测的 AUROC-image 分数

Table 1 AUROC-image scores for anomaly detection across categories

| 类别 | US | MKD | PaDiM | EdgRec | DRAEM | RD | EMSF | MFRAD | EFAD | PFEAL | 本文 | |
|-----|-----|------|-------|--------|--------------|-------|-------|-------|-------|-------|-------|--------------|
| 纹理类 | 地毯 | 91.6 | 79.3 | 99.8 | 97.4 | 97.0 | 98.9 | 99.9 | 99.2 | 99.42 | 100.0 | 100.0 |
| | 网格 | 81.0 | 78.0 | 96.7 | 99.7 | 99.9 | 100.0 | 97.6 | 99.8 | — | 98.3 | 99.5 |
| | 皮革 | 88.2 | 95.1 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | — | 100.0 | 100.0 |
| | 瓷砖 | 99.1 | 91.6 | 98.1 | 100.0 | 99.6 | 99.3 | 99.4 | 99.9 | — | 99.9 | 99.7 |
| | 木板 | 97.7 | 94.3 | 99.2 | 94.0 | 99.1 | 99.2 | 99.4 | 98.7 | — | 99.8 | 99.6 |
| 对象类 | 瓶子 | 99.0 | 99.4 | 99.9 | 100.0 | 99.2 | 100.0 | 99.8 | 100.0 | — | 100.0 | 100.0 |
| | 电缆 | 86.2 | 89.2 | 92.7 | 97.9 | 91.8 | 95.0 | 96.5 | 99.8 | — | 97.8 | 98.9 |
| | 胶囊 | 86.1 | 80.5 | 91.3 | 95.5 | 98.5 | 96.3 | 95.9 | 98.7 | — | 97.6 | 96.4 |
| | 榛子 | 93.1 | 98.4 | 92.0 | 98.4 | 100.0 | 99.9 | 100.0 | 100.0 | — | 99.3 | 100.0 |
| | 螺母 | 82.0 | 73.6 | 98.7 | 97.3 | 98.7 | 100.0 | 97.8 | 100.0 | — | 96.2 | 100.0 |
| | 药片 | 87.9 | 82.7 | 93.3 | 99.0 | 98.9 | 96.6 | 97.2 | 97.1 | — | 96.4 | 98.3 |
| | 螺丝钉 | 54.9 | 83.3 | 85.8 | 89.9 | 93.9 | 97.0 | 91.7 | 95.1 | — | 96.3 | 96.2 |
| | 牙刷 | 95.3 | 92.2 | 96.1 | 100.0 | 100.0 | 99.5 | 99.8 | 94.4 | — | 100.0 | 100.0 |
| | 三极管 | 81.8 | 85.6 | 97.4 | 99.8 | 93.1 | 96.7 | 99.3 | 99.0 | — | 98.9 | 98.7 |
| | 拉链 | 91.9 | 93.2 | 90.3 | 98.3 | 100.0 | 98.5 | 96.9 | 98.4 | — | 99.4 | 97.4 |
| | 均值 | 87.7 | 87.8 | 95.5 | 97.8 | 98.0 | 98.5 | 98.1 | 98.7 | — | 98.7 | 99.0 |

表 2 各类别异常检测的 AUROC-pixel 和 PRO 分数

Table 2 AUROC-pixel and PRO scores for anomaly detection across categories

| 类别 | US | MKD | PaDiM | EdgRec | DRAEM | RD | EMSF | MFR | EFAD | PFEAL | 本文 | |
|-----|-----|--------|--------|-----------|-----------|--------|-----------|------------------|-----------|---------|-----|-------------------|
| 纹理类 | 地毯 | —/87.9 | 95.6/— | 99.1/96.2 | 99.4/96.9 | 95.5/— | 98.9/97.0 | 99.3/97.8 | —/— | 98.2/82 | —/— | 99.3/ 97.9 |
| | 网格 | —/95.2 | 91.8/— | 97.3/94.6 | 99.2/97.0 | 99.7/— | 99.3/97.6 | 96.3/98.1 | —/— | —/— | —/— | 98.7/97.3 |
| | 皮革 | —/94.5 | 98.1/— | 99.2/97.8 | 99.7/98.8 | 98.6/— | 99.4/99.1 | 99.5/99.2 | —/— | —/— | —/— | 99.5/ 99.2 |
| | 瓷砖 | —/94.6 | 82.8/— | 94.1/86.0 | 98.6/96.3 | 99.2/— | 95.6/90.6 | 95.5/92.2 | —/— | —/— | —/— | 95.4/90.1 |
| | 木板 | —/91.1 | 84.8/— | 94.9/91.1 | 91.4/77.5 | 96.4/— | 95.3/90.9 | 95.7/92.9 | —/— | —/— | —/— | 95.5/ 93.3 |
| 对象类 | 瓶子 | —/93.1 | 96.3/— | 98.3/94.8 | 98.3/94.3 | 96.4/— | 98.7/96.6 | 98.7/96.5 | —/— | —/— | —/— | 98.9/96.9 |
| | 电缆 | —/81.8 | 82.4/— | 98.3/94.8 | 97.7/88.7 | 94.7/— | 97.4/91.0 | 98.2/93.3 | —/— | —/— | —/— | 98.3/93.8 |
| | 胶囊 | —/96.8 | 95.9/— | 98.5/93.5 | 95.2/82.2 | 94.3/— | 98.7/95.8 | 98.6/94.7 | —/— | —/— | —/— | 98.7/95.6 |
| | 榛子 | —/96.5 | 94.6/— | 98.2/92.6 | 99.4/95.4 | 99.7/— | 98.9/95.5 | 99.2/96.4 | —/— | —/— | —/— | 99.2/ 95.8 |
| | 螺母 | —/94.2 | 86.4/— | 97.2/85.6 | 98.0/91.2 | 99.5/— | 97.3/92.3 | 99.8/92.1 | —/— | —/— | —/— | 97.4/91.6 |
| | 药片 | —/96.1 | 89.6/— | 95.7/82.7 | 98.7/96.1 | 97.6/— | 98.2/96.4 | 98.0/96.9 | —/— | —/— | —/— | 98.2/ 96.9 |
| | 螺丝钉 | —/94.2 | 96.0/— | 98.5/94.4 | 97.7/89.3 | 97.6/— | 99.6/98.2 | 99.1/96.4 | —/— | —/— | —/— | 99.5/ 98.3 |
| | 牙刷 | —/93.3 | 96.1/— | 98.8/93.1 | 99.2/94.9 | 98.1/— | 99.1/94.5 | 99.0/93.4 | —/— | —/— | —/— | 98.7/93.9 |
| | 三极管 | —/66.6 | 76.5/— | 97.5/84.5 | 94.3/87.7 | 90.9/— | 92.5/78.0 | 96.5/84.0 | —/— | —/— | —/— | 93.3/86.5 |
| | 拉链 | —/95.1 | 93.9/— | 98.5/95.9 | 98.7/96.0 | 98.8/— | 98.2/95.4 | 97.3/93.1 | —/— | —/— | —/— | 97.9/94.7 |
| | 均值 | —/91.4 | 90.7/— | 97.5/92.1 | 97.7/92.1 | 97.3/— | 97.8/93.9 | 98.3/95.0 | 98.0/94.3 | —/— | —/— | 97.9/ 94.8 |

表 3 消融实验结果

Table 3 Results of ablation experiments %

| 方法 | 评价指标 | | |
|------------|-------------|-------------|-------|
| | AUROC-image | AUROC-pixel | AUPRO |
| 传统知识融合 | 98.5 | 97.7 | 94.2 |
| 传统知识融合+OCE | 98.5 | 97.8 | 93.9 |
| AKF | 98.7 | 97.9 | 94.4 |
| AKF+OCE | 99.0 | 97.9 | 94.8 |

响进而误判,而本文方法能够规避斑点对模型识别的作用,提取异常信息进而精准识别缺陷。从总体分析得,本文中的方法能够规避不重要信息的影响,使得学生网络能够学习到正常样本之间的特征分布,进而识别异常缺陷。

图 6 分别展示了表 3 中的 4 种情况下网络间的异常分数图 M1、M2、M3 和特征融合 Fusion 之后的可视化结果。从图 6 可以看到,本文方法在各个特征层次上的识别效果均优于其他方法。在螺母这一类别上,本文方法在 M1 和 M2 层次的识别效果最好,背景噪声对模型的异常识别几乎没有产生影响。在 Fusion 这一结果来看,本文方法优于传统方法。因此,实验证明了训练中渐进式知识融合的作用优于传统特征融合,并且渐进式知识融合机制和 OCE 机制达到了互补的效果。从整体上看,渐进式知识融合能够通过过滤各个特征层次上的异常特征,使得模型能够更加关注异常区域,提升检测精度。

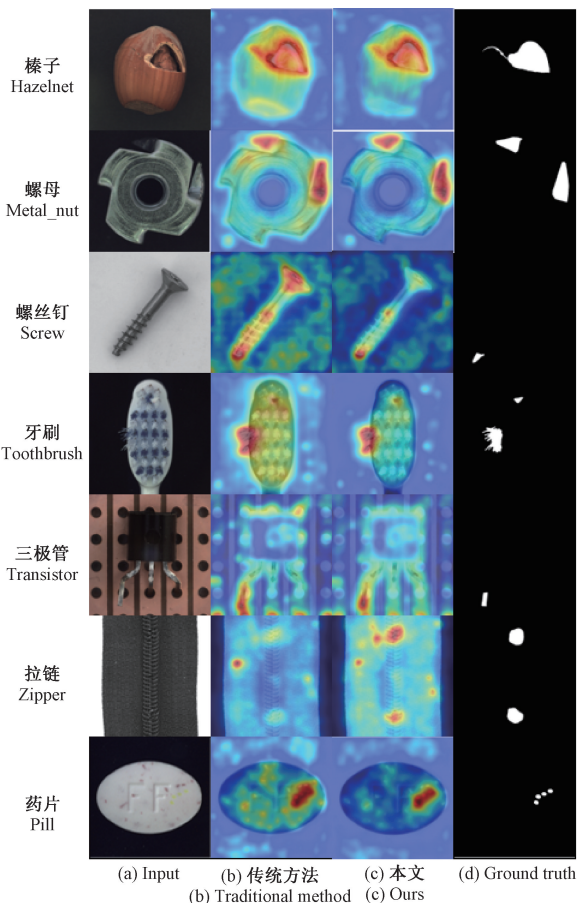
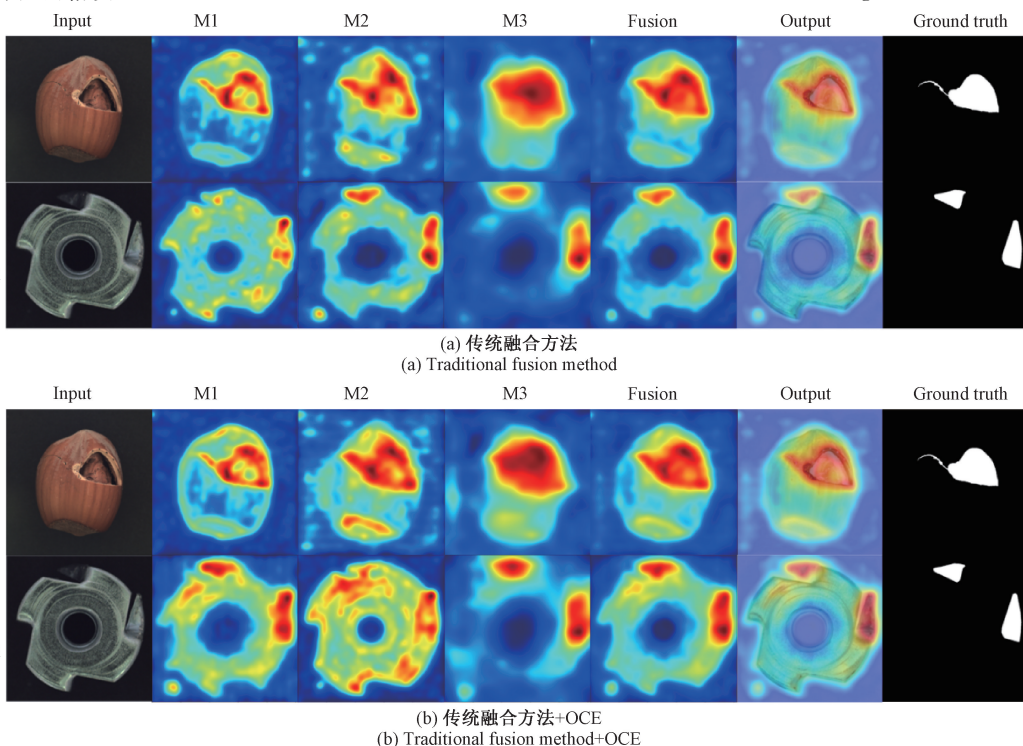


图 5 本文方法和传统知识融合的可视化结果
Fig. 5 Visualization results of the proposed method and traditional knowledge fusion



(b) 传统融合方法+OCE
(b) Traditional fusion method+OCE

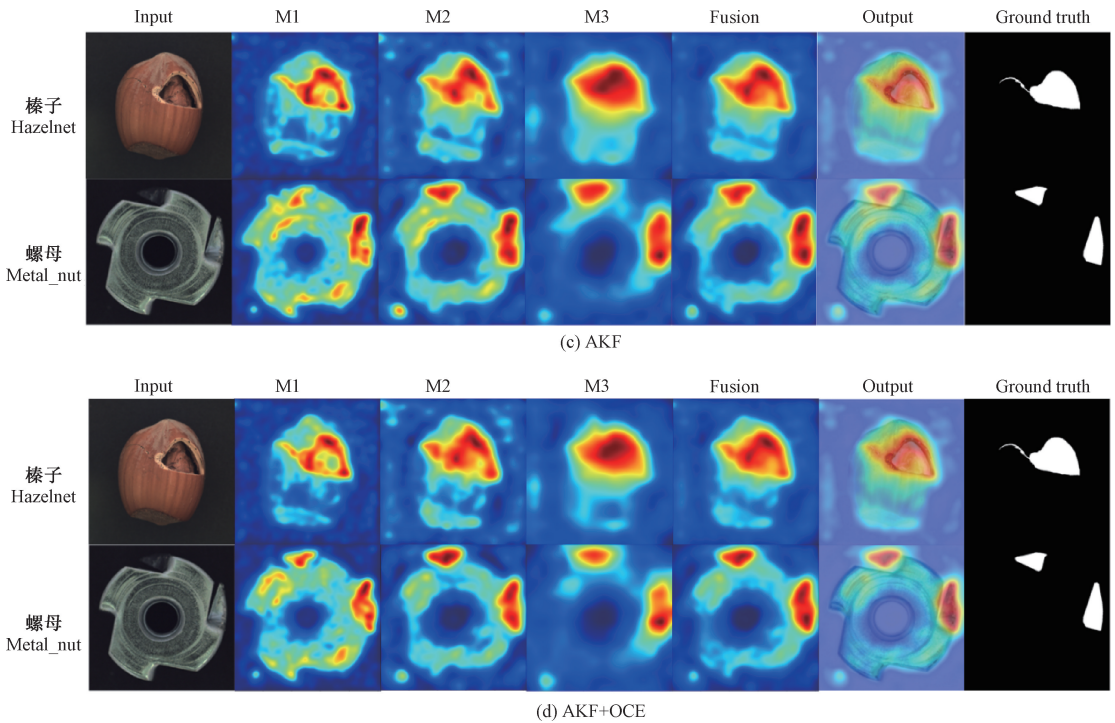


图 6 传统知识融合方法和 AKF 的消融实验可视化结果

Fig. 6 Visualized ablation results of traditional knowledge fusion and AKF

4 结 论

传统的知识蒸馏模型在训练过程中,学生网络难以学习复杂且冗余的特征信息。为了解决这一问题,本文提出一种基于渐进式知识融合的知识蒸馏工业异常检测方法。本文在反向知识蒸馏模型的基础上引入渐进式知识融合机制增强了瓶颈层处理教师网络输出的冗余且复杂的特征信息的能力。同时使用单类嵌入机制,该方法能够将复杂特征转化为能够促进学术网络重构的浅层特征,从而有效地融合了来自不同层级的特征,提升了检测的准确性和鲁棒性。

实验结果表明,本文所提出的算法在 MVTEC AD 数据集上超越了多数主流的检测模型。在 MVTEC AD 数据集上,模型在图像级和像素级的多个评价指标上均达到了优异的检测性能,通过消融实验,进一步验证了模型中各部分设计的有效性。

参考文献

- [1] BERGMANN P, FAUSER M, SATTLEGGER D, et al. MVTEC AD—A comprehensive real-world dataset for unsupervised anomaly detection [C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019: 9592-9600.
- [2] NI X F, MA Z J, LIU J W, et al. Attention network for rail surface defect detection via consistency of intersection-over-union (IOU)-guided center-point

estimation [J]. IEEE Transactions on Industrial Informatics, 2022, 18(3): 1694-1705.

- [3] CARRERA D, MANGANINI F, BORACCHI G, et al. Defect detection in sem images of nanofibrous materials [J]. IEEE Transactions on Industrial Informatics, 2017, 13(2): 551-561.
- [4] GONG D, LIU L Q, LE V, et al. Memorizing normality to detect anomaly: Memory-augmented deep autoencoder for unsupervised anomaly detection[C]. IEEE/CVF International Conference on Computer Vision, 2019: 1705-1714.
- [5] 张彪, 荀荣科, 许家忠. 基于深度学习的工业轴承缺陷检测算法研究 [J]. 仪器仪表学报, 2025, 46(4): 136-149.
ZHANG B, XUN R K, XU J ZH. Research on industrial bearing defect detection algorithm based on deep learning [J]. Chinese Journal of Scientific Instrument, 2025, 46(4): 136-149.
- [6] RISTEA N C, MADAN N, IONESCU R T, et al. Self-supervised predictive convolutional attentive block for anomaly detection[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022: 13576-13586.
- [7] ZAVRTANIK V, KRISTAN M, SKOČAJ D. Reconstruction by inpainting for visual anomaly detection [J]. Pattern Recognition, 2021,

- 112: 107706.
- [8] DEFARD T, SETKOV A, LOESCH A, et al. PaDiM: A patch distribution modeling framework for anomaly detection and localization [C]. Pattern Recognition, ICPR International Workshops and Challenges. Cham: Springer International Publishing, 2021: 475-489.
- [9] CAI Y X, LIANG D K, LUO D L, et al. A discrepancy aware framework for robust anomaly detection [J]. IEEE Transactions on Industrial Informatics, 2024, 20(3): 3986-3995.
- [10] WU Q W, LI H, TIAN CH Y, et al. AEKD: Unsupervised auto-encoder knowledge distillation for industrial anomaly detection [J]. Journal of Manufacturing Systems, 2024, 73: 159-169.
- [11] WANG J B, CHENG J Y, GAO C, et al. Enhanced fabric defect detection with feature contrast interference suppression [J]. IEEE Transactions on Instrumentation and Measurement, 2025, 74: 1-12.
- [12] DING F Q, LI B, LIU H, et al. PFEAL: A novel framework for image anomaly detection using pre-trained feature extraction and activation learning [J]. IEEE Transactions on Emerging Topics in Computational Intelligence, 2026, 10(1): 336-347.
- [13] DENG H Q, LI X Y. Anomaly detection via reverse distillation from one-class embedding [C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022: 9737-9746.
- [14] TONG G X, LI Q Q, SONG Y. Enhanced multi-scale features mutual mapping fusion based on reverse knowledge distillation for industrial anomaly detection and localization [J]. IEEE Transactions on Big Data, 2024, 10(4): 498-513.
- [15] YAN X D, ZHANG H D, XU X M, et al. Learning semantic context from normal samples for unsupervised anomaly detection [J]. AAAI Conference on Artificial Intelligence, 2021, 35(4): 3110-3118.
- [16] WYATT J, LEACH A, SCHMON S M, et al. AnoDDPM: Anomaly detection with denoising diffusion probabilistic models using simplex noise [C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022: 650-656.
- [17] HE H Y, ZHANG J N, CHEN H X, et al. A diffusion-based framework for multi-class anomaly detection [J]. AAAI Conference on Artificial Intelligence, 2024, 38(8): 8472-8480.
- [18] ROTH K, PEMULA L, ZEPEDA J, et al. Towards total recall in industrial anomaly detection [C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022: 14318-14328.
- [19] 魏金洋, 苑明哲, 曹飞道, 等. 基于注意力机制和空洞金字塔池化的缺陷检测 [J]. 信息与控制, 2024, 53(5): 662-672.
- WEI J Y, YUAN M ZH, CAO F D, et al. Defect detection based on attention mechanism and dilated pyramid pooling [J]. Information and Control, 2024, 53(5): 662-672.
- [20] 冯志远, 陈莹. 基于判别式增强的蒸馏学习自监督缺陷检测 [J]. 系统仿真学报, 2025, 37(6): 1499-1511.
- FENG ZH Y, CHEN Y. Self-supervised defect detection with distillation learning based on discriminative enhancement [J]. Journal of System Simulation, 2025, 37(6): 1499-1511.
- [21] RUDOLPH M, WEHRBEIN T, ROSENHAHN B, et al. Asymmetric student-teacher networks for industrial anomaly detection [C]. IEEE/CVF Winter Conference on Applications of Computer Vision, 2023: 2592-2602.
- [22] YANG G Y, LEI J, TIAN H, et al. Asymptotic feature pyramid network for labeling pixels and regions [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2024, 34(9): 7820-7829.
- [23] 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.
- [24] 宋晔, 吴一全. 基于无人机航拍的风力发电机叶片表面缺陷检测综述 [J]. 仪器仪表学报, 2024, 45(10): 1-25.
- SONG Y, WU Y Q. Review of surface defect detection of wind turbine blades based on UAV aerial photography [J]. Chinese Journal of Scientific Instrument, 2024, 45(10): 1-25.
- [25] 王海群, 王文科, 于海峰. 改进 YOLOv10n 的输电线路部件缺陷检测算法 [J]. 电子测量技术, 2025, 48(10): 62-72.
- WANG H Q, WANG W K, YU H F. Improved YOLOv10n-based transmission line component defect detection algorithm [J]. Electronic Measurement Technology, 2025, 48(10): 62-72.
- [26] BERGMANN P, FAUSER M, SATTLEGER D, et al. Uninformed students: Student-teacher anomaly detection with discriminative latent embeddings [C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020: 4183-4192.

- [27] SALEHI M, SADJADI N, BASELIZADEH S, et al. Multiresolution knowledge distillation for anomaly detection[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021: 14902-14912.
- [28] GU H CH, LI G Y, LIU ZH. Masked feature regeneration based asymmetric student-teacher network for anomaly detection[J]. Multimedia Tools and Applications, 2024, 83(42): 90573-90594.
- [29] ZAVRTANIK V, KRISTAN M, SKOČAJ D. DRAEM-A discriminatively trained reconstruction embedding for surface anomaly detection[C]. IEEE/CVF International Conference on Computer Vision, 2021: 8330-8339.
- [30] LIU T K, LI B, ZHAO ZH, et al. Reconstruction from edge image combined with color and gradient difference for industrial surface anomaly detection[J]. Advanced Engineering Informatics, 2025, 64: 103064.

作者简介

赵智斌, 本科, 主要研究方向为图像处理。

E-mail: zhaozb168@163.com

王浩(通信作者), 博士, 讲师, 主要研究方向为图像处理。

E-mail: travis@sut.edu.cn

朱建光, 博士, 教授, 主要研究方向为故障诊断。

E-mail: zhujg@sut.edu.cn

薛丹, 博士, 副教授, 主要研究方向为机器视觉。

E-mail: x2d2003@163.com