

基于模糊遮蔽与动态推理的生成式工业异常定位模型

吴天月, 张辉, 张邹铨, 唐珺琨

引用本文

吴天月, 张辉, 张邹铨, 唐珺琨. [基于模糊遮蔽与动态推理的生成式工业异常定位模型](#)[J]. 计算机科学, 2023, 50(11A): 230100073-7.

WU Tianyue, ZHANG Hui, ZHANG Zouquan, TANG Junkun. [Generative Industrial Image Abnormal Location Model Based on Fuzzy Masking and Dynamic Inference](#) [J]. Computer Science, 2023, 50(11A): 230100073-7.

相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

Similar articles recommended (Please use Firefox or IE to view the article)

[基于边缘引导的多尺度医学影像分割方法](#)

Medical Image Segmentation Based on Multi-scale Edge Guidance

计算机科学, 2023, 50(11A): 220900059-7. <https://doi.org/10.11896/jsjcx.220900059>

[基于语义注意力的医学图像超分辨率方法](#)

Medical Image Super-resolution Method Based on Semantic Attention

计算机科学, 2023, 50(11A): 221200107-6. <https://doi.org/10.11896/jsjcx.221200107>

[一种基于因果推理的垃圾分类方法](#)

Novel Method for Trash Classification Based on Causal Inference

计算机科学, 2023, 50(11A): 220800218-6. <https://doi.org/10.11896/jsjcx.220800218>

[接诉即办智能派单业务调度算法研究](#)

Study on Scheduling Algorithm of Intelligent Order Dispatching

计算机科学, 2023, 50(11A): 230300029-7. <https://doi.org/10.11896/jsjcx.230300029>

[基于LSTM神经网络的QPSK智能接收机设计](#)

Design of QPSK Intelligent Receiver Based on LSTM Neural Network

计算机科学, 2023, 50(11A): 230200219-5. <https://doi.org/10.11896/jsjcx.230200219>

基于模糊遮蔽与动态推理的生成式工业异常定位模型

吴天月¹ 张辉² 张邹铨¹ 唐珺琨¹

1 长沙理工大学电气与信息工程学院 长沙 410000

2 湖南大学机器人学院 长沙 410000

(yue__wuwuwu@163.com)

摘要 工业生产机械化对工业产品质量检测环节提出了新的要求,需要一种具有高精度、易于移植的异常检测算法来适应生产方式的更新。针对工业生产中,异常样本出现概率低、无法完全预测的固有难题,提出了一种基于模糊遮蔽与动态推理的生成式工业异常定位模型。首先,设计了一个基于随机模糊遮蔽的对比样本生成模块,用于获取高质量的模拟异常图像。同时,利用浅层特征融合路径保留更多的边缘信息,使用损失函数加权使模型更加关注结构相似性,以及使用对比学习的方式使网络获得更好的表示能力。其次,为了缓解生成式模型输出图像模糊的问题,设计了多分支异常动态推理方法,使迭代生成和精准修复两分支相互配合,拉远背景噪声与真实异常间的距离。实验结果表明,所提方法在 MVTec 数据集上取得了 91.42% 的平均定位精度,其中有 12 类达到了前三的异常定位精度,能够较完整地获取异常地位置;对于纹理复杂和背景占比比较大的图像,所提方法仍然保持着较高的指标敏感度,其异常定位性能在近年来提出的生成式检测模型中取得了最佳。

关键词: 工业图像检测;异常定位;深度学习;生成式学习;动态异常推理

中图法分类号 TP3

Generative Industrial Image Abnormal Location Model Based on Fuzzy Masking and Dynamic Inference

WU Tianyue¹, ZHANG Hui², ZHANG Zouquan¹ and TANG Junkun¹

1 School of Electrical & Information Engineering, Changsha University of Science and Technology, Changsha 410000, China

2 School of Robotics, Hunan University, Changsha 410000, China

Abstract The mechanization of industrial production puts forward new requirements for the inspection of industrial product quality, and a high-precision, easy-to-transplant anomaly detection algorithm is required to adapt to the update of production methods. Aiming at the inherent problem of low probability of abnormal samples in industrial production and incomplete prediction, a generative industrial anomaly localization model based on fuzzy masking and dynamic reasoning is proposed. Firstly, a contrast sample generation module based on random fuzzy occlusion is designed to obtain high-quality simulated anomalous images. At the same time, the shallow feature fusion path is used to retain more edge information, the loss-loss function weighting is used to make the model pay more attention to structural similarity, and the contrast learning method is used to make the network obtain better representation ability. Secondly, in order to alleviate the problem of blurred output images of generative models, a multi-branch anomaly dynamic inference method is designed, and the two branches of iterative generation and accurate repair cooperate with each other to widen the distance between background noise and real anomalies. Experimental results show that the proposed method achieves an average localization accuracy of 91.42% on the MVTec dataset, and the top three anomalous localization accuracy are obtained in 12 classes. The location of anomalies can be obtained more completely. For images with complex textures and large backgrounds, it still maintains high index sensitivity, and the average anomaly localization performance has reached the best in published generative detection models published in recent years.

Keywords Industrial image detection, Anomaly location, Deep learning, Generative learning, Dynamic anomaly inference

1 引言

在工业场景中,传统机器视觉受到人工提取特征方式的限制,通常需要利用硬件设备来构建相对稳定的外界环境,

从而保证检测算法的有效性。而深度学习对光源和检测目标一致性的要求相对较低,能够适应场景中的变化。因此,众多研究人员尝试将深度学习方法应用于工业质量检测中。深度学习方法从本质上来说是一种基于数据的方法,通过对已知

基金项目:科技创新 2030—“新一代人工智能”重大项目课题(2021ZD0114503);国家自然科学基金(61971071,62027810);湖南省杰出青年科学基金项目(2021JJ10025);湖南省研究生科研创新项目(CX20210797)

This work was supported by the National Key R&D Program of China (2021ZD0114503), National Natural Science Foundation of China (61971071,62027810), National Science Found for Distinguished Young Scholars of Hunan Province, China (2021JJ10025) and Postgraduate Scientific Research Innovation Project of Hunan Province (CX20210797).

通信作者:张辉(zhanghuihy@126.com)

信息的分析来得到目标出现的规律,从而完成下游任务的目标^[1]。但在工业生产环境下,随着生产技术的不断提高,缺陷样本的出现概率不断下降,使缺陷样本收集难度大大提升。同时,与自然图像不同,在工业生产中,缺陷的出现具有随机性,无法保证收集的缺陷覆盖异常样本空间,使得异常数据集搭建难度增加。由于异常样本的稀缺和异常样本收集不完整,使得无论是传统机器视觉算法中的特征分析算法,还是深度学习中的有监督学习,都无法通过大量标记的异常数据进行特征提取和模型训练,在实际实现过程中难以发挥出其真实的性能,导致模型失效。相反,正常样本产出率较高,易于收集且形态特征相比自然图像较为固定^[2-3]。因此,通常以弱监督或无监督学习策略来训练工业场景下的深度学习模型。而在工业生产中,相比区分缺陷种类的方法,判断产品是否存在缺陷的方法更加具有研究价值。因此,通常的做法是将工业场景下的产品表面质量检测任务归类于二分类的异常检测任务。

本文围绕深度学习方法中的自监督生成式模型展开研究,提出了一种基于模糊遮蔽与动态推理的生成式工业异常定位模型。使用正常样本训练检测模型,消除深度学习网络模型对异常数据的依赖。使用生成式模型输入输出可视化对比来定位异常。本文的主要工作如下。

1)通过在训练数据中随机添加具有干扰信息的模糊掩码,使得生成式模型关注上下文信息来实现缺失部分的重构。利用边缘信息融合和结构相似度优化模型训练策略,从而使模型获得更强的表示学习能力。

2)提出了一种多分支异常动态推理方法。利用迭代生成和精准修复两分支之间的相互配合,对两种生成式模型输出情况设计了不同的异常定位方法,用于增强模型对背景噪声的判别能力及模型的通用性。

3)本文方法在 MVTec^[4]数据集上实现了较高的异常定位精度,在 12 类上取得了前三的异常定位精度,平均定位精度达到 91.4%,验证了本文方法的有效性。

2 相关工作

近年来,深度学习方法在工业图像异常检测领域的研究取得了显著的进展。2019 年, MVTec 数据集^[4]的公开为后续工业图像异常检测的研究提供了基础。该数据集提供了 15 类工业样本以及像素级的异常标注。此后,大量国内外研究者针对工业异常检测任务,以无监督异常检测为方法展开

研究。Chen 等^[5]通过超球体最小化策略来训练单阶段异常分类模型,以获取紧凑的特征描述,从而完成异常分类。Han 等^[6]引入自注意力机制来训练模型的抗噪性能,设计深度聚类算法来进行准确的特征聚类分配。Gong 等^[7]和 Park 等^[8]设计了一种记忆增强的自编码器网络,在编码器与解码器之间添加特征“字典”来剔除推理过程中的异常特征。该方法在 MINST 手写数字数据集上取得了较好的检测精度。但是由于输出特征是由特征“字典”加权重组得到的,这导致工业图像输出严重模糊,在工业异常检测任务上的效果不能使人满意。Zhou 等^[9]通过建立结构-纹理对应关系来替代图像-特征间的对应关系,以获取到质量更高的重构图像。Yang 等^[10]则利用流形正则化来挖掘数据中的关键特征,提高图像重构效果。上述方法主要集中在特征重组和关键特征提取方面的改进,但由于高维特征较为抽象,有效特征提取复杂且需要谨慎选择约束,导致模型设计与训练策略的选择较为复杂。

基于图像修复思想的异常检测模型通过将输入图像中的局部区域去除的处理方法,使模型通过周边的特征还原缺失区域,以达到增强生成式模型的代表能力的目的。此类方法只需要对直观的输入输出图像间的相似度进行比较,较上述方法约束选择更为简单。Vitjan 等^[11]首次将图像修复思想运用于工业图像异常定位任务中,使用大小均匀的方块训练 U-Net 模型,以实现异常块的搜索。Li 等^[12]为了减少像素块的数量,采用超像素分割方法划分图像,以提高搜索的效率。Yan 等^[13]使用多尺度栅格条纹遮蔽输入图像来训练生成对抗网络,以匹配不同尺度的异常形状。Wang 等^[14]设计了一个多头生成式模型,在获取上下文信息的同时,保留了局部特征提取路径,两者构成对比关系共同完成决策。但此类方法在推理阶段通常需要在整张图像上搜索包含异常的小块,检测效率受到限制。

知识蒸馏策略的提出给网络训练策略开辟了一条新道路,即用从大型数据集上训练的教师网络中提炼出与下游任务相关的权重分布来构建学生网络的方式替代直接使用数据集训练网络。Salehi 等^[15]与 Xing 等^[16]通过知识蒸馏的训练策略,实现了工业异常定位精度的有效提升。

3 基于模糊遮蔽与动态推理的生成式工业异常定位模型

本文提出的基于模糊遮蔽与动态推理的生成式异常定位模型的整体框架如图 1 所示。

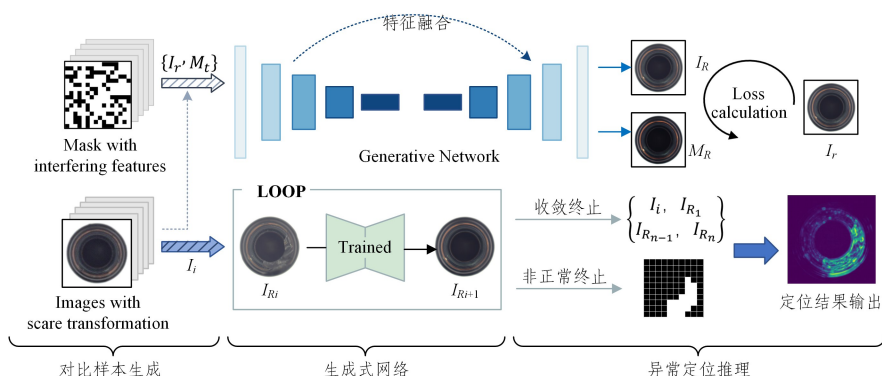


图 1 模型的整体框架

Fig. 1 Overall framework of the proposed model

整个检测模型根据功用可以分为对比样本生成模块、生成式网络模型以及异常推理模块 3 个部分。对比样本生成模块在原始样本中进行随机模糊遮蔽生成;生成式网络模型将输入图像修复为正常样本图像,异常推理模块利用多分支异常定位机制准确且完整地定位出异常区域。

3.1 随机模糊遮蔽的对比样本生成模块

3.1.1 设计动机

在目前基于图像修复思想的异常定位方法中,普遍使用各种形态的全黑掩膜^[14]或者从不相干的蒙版^[17]中获取掩膜。全黑掩膜用于模拟最“苛刻”的实现图像修复的条件,即异常部位的特征完全丢失的情况。而不相干蒙版则用于在正常样本中制造模拟异常来训练网络的定位能力。但是通过观察真实的异常样本发现,真实异常中心区域与正常区域十分相似,存在一定的隐蔽性,如图 2 所示。上述两种方法完全丢弃了遮蔽区域的原始特征,这也导致推理阶段低效率地遍历图像,影响检测效率。因此设计保留部分原始样本特征的模糊掩膜,并且越靠近中心的区域特征的完整度越高。这样能够使训练阶段的检测任务更加贴合生成式任务,实现推理算法的有效配合。

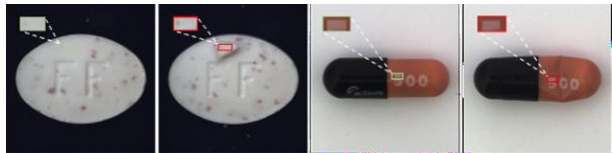


图 2 异常中心区域与正常区域对比图

Fig. 2 Comparison between abnormal central area and normal area

3.1.2 对比样本生成

对比样本生成分为掩膜生成和掩膜覆盖两个部分,具体实现步骤如图 3 所示。首先,为了随机选择训练图像中的遮蔽区域,首先需要将每张图片划分为 $\frac{W}{k} \times \frac{H}{k}$ 个均匀大小的网格,每个网格由 $k \times k$ 个像素组成。其次,为了获取模糊掩膜,对每个网格块进行加权处理,越靠近掩膜中心保留的原始特征越多。最后,掩膜与原始图像叠加构成对比样本,与正常样本一同输入网络模型。

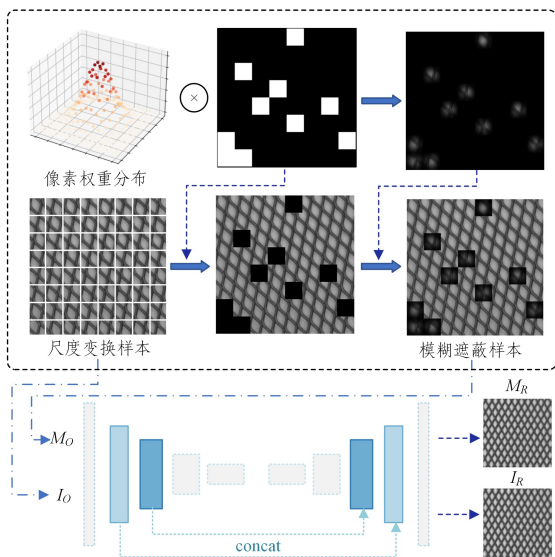


图 3 对比样本生成模块

Fig. 3 Contrastive sample generation module

每次批次开始训练前,模型都会随机选择一组新的网格,被选中的网格为输入图像中被遮蔽的区域。同一批次中,任意一张网格与其他所有网格互补,以此保证网络能够通过图像修复任务获取图像各部位的特征。 k 的取值对最终的模型性能有一定的影响,若 k 过小,则无法对重构模型的表示学习能力进行有效训练,若 k 过大,模型可能无法获取到足够的正常样本特征。本文中掩膜宽度取值为 16。

另外,为了降低浅层的边缘信息在特征压缩与解压的过程中的损失,在网络的第二层与第三层分别添加了一条浅层特征融合路径连接到特征解码器的对应层,特征图通过通道维度进行拼接。

3.2 基于结构相似度优化的生成式网络模型

在生成式异常定位模型中,由于神经网络并非直接输出预测结果,异常判定函数对预测结果有显著影响。均方差损失函数能够对两张图像进行点对点的差异计算,但是对于像素值较低的区域敏感度不够。为了获取更准确的异常评分规则,本文引入了 SSIM 相似度指标^[18],用于量化真实异常与真实样本间的相似度差异,从而获取更精准的异常评价函数。SSIM 相似度以图像小块为基准,综合考虑了多个相似度维度以及相邻像素点之间的关联关系,同时削弱了对像素亮度和颜色的敏感度,加强了对边缘与纹理的对比计算,更符合人眼判定的两张图片的相似度。为了更精准地反映出图像间的差异,我们统计了 MVTec 中的 15 类图像中异常部位 SSIM 损失函数中 3 种相似度之间的变化幅度,如图 4 所示。

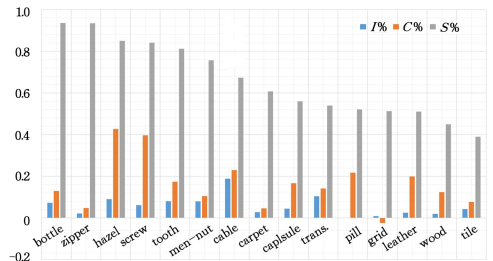


图 4 异常样本损失柱状图

Fig. 4 Histogram of abnormal sample loss

通过观察各相似度间的数值和比值,并结合每类图像本身的特征后发现,结构相似度 ($S(x, y)$) 在各类异常中的变化幅度最大,并且以其数值大小作为依据,可以将 15 类图像划分为 3 类具有鲜明特征的图像,分别为具有重复性背景纹理的图像(I类),共 8 类;具有标准的形状特征且背景占比较大的图像(II类),共 6 类;第三类则较为特殊,包含上述两类图像的特征,我们将其称之为混合图像,以 cable 类为代表。根据统计值,对 SSIM 相似度的 3 个部分进行加权,改进后的 SSIM 函数如式(4)所示,式(1)~式(3)为各部分相似度计算函数,其中 C_1 取 0.01^2 , C_2 取 0.03^2 , C_3 取 0.015^2 。

$$I(x, y) = \frac{2 \mu_x \mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1} \quad (1)$$

$$C(x, y) = \frac{2 \sigma_x \sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2} \quad (2)$$

$$S(x, y) = \frac{2 \sigma_{xy} + C_3}{\sigma_x + \sigma_y + C_3} \quad (3)$$

$$\hat{L}_{SSIM} = 1 - [I(x, y)^\alpha \times C(x, y) \times S(x, y)^\beta]^\frac{1}{\gamma} \quad (4)$$

根据统计值, α, β, γ 的取值如表 1 所列。为了避免乘方造

成相似度分布区间过窄的问题,取相似度值的 $\frac{1}{3}$ 次幂。加权后的损失函数能够更加关注影响异常判断的相似度,对输出结果进行精准评分。

表1 加权参数取值表

Class	α	β	γ
I	0.2	0.5	2.3
II	0.1	0.8	2.1
III	0.3	0.6	2.1

同时,为了调和评价函数的敏感度,我们也保留了均方差损失函数。通过结合上述两种损失函数,最终用于训练的损失函数与异常评价函数如式(5)所示。

$$L = [L_{MSE}(I_i, I_R) + \hat{L}_{SSIM}(I_i, I_R)] + \lambda [L_{MSE}(I_i, M_R) + \hat{L}_{SSIM}(I_i, M_R)] \quad (5)$$

其中,第一部分为对生成结果的约束;第二部分为对修复结果的约束。通过损失函数能够同时约束模型的修复与重构效果,为后续动态推理模块的实现提供基础。

3.3 异常动态推理模型

生成式模型的输出是进行异常定位的基础。为了获取更真实的输出图像和更完整的异常定位,本文设计了一个两分支异常动态推理算法。首先对网络输出稳定程度进行评估,根据输出稳定性激活重构分支或修复分支。重构分支利用连续重构来获取更真实的输出,而修复分支则利用局部修复获取更完整的异常定位。算法的执行步骤如算法1所示。稳定性判定以相邻两次重构图像的均方差相似度为标准。同时,为了避免迭代进入死循环,设置了最大迭代次数10次。

算法1 异常迭代推理算法

输入:迭代次数 N ,收敛条件 m ,评价函数 $\text{Sim}(x, y)$,修复算法 $\text{Model}(I)$,输入图像 $I_i \in \mathbb{R}^{H \times W \times C}$,图块大小 $k //$ 将图像切分成大小为 $k \times k$ 的小块

输出:异常定位结果 $\text{Score}_{\text{abnor}}$

1. 初始化: $I_{R_n} \leftarrow I_i / * I_i$ 为输入图像 $*$; $\text{flag} \leftarrow \text{False} / * \text{flag}$ 为收敛标志符 $*$;
2. for $i \leftarrow 1$ to N do $*$ N 为最大迭代次数 $*$ /
3. $I_{R_{n+1}} = \text{Model}(I_{R_n})$
4. if $\text{Sim}(I_{R_n}, I_{R_{n+1}}) \leq m$ then
5. $\text{flag} \leftarrow \text{True}$; break;
6. end for
7. if flag is True then
8. $\text{back} = (1 - \text{Sim}(I_{R_1}, I_i)) + (1 - \text{Sim}(I_{R_n}, I_{R_{n+1}}))$;
9. $\text{Score}_{\text{abnor}} = (1 - \text{Sim}(I_i, \text{Model}(I_m \times I_i))) - \text{back}$;
10. else
11. $\text{Mas } k_0 = 1 - \text{Sim}(I_{R_1}, I_i)$, $I_m = \text{ones}(H, W, C)$;
12. Spilt $\text{Mas } k_0$ to $k \times k$ patches p_1, p_2, \dots, p_{N_k} ;
13. for each patch p do
14. if $p \neq 0$ then $I_m[(x, y) \in p] = 0$; end if
15. end for
16. $\text{Score}_{\text{abnor}} = (1 - \text{Sim}(I_i, \text{Model}(I_m \times I_i)))$;
17. end if
18. $\text{Score}_{\text{abnor}} = (1 - \text{Sim}(I_i, \text{Model}(I_m \times I_i)))$;
19. return $\text{Score}_{\text{abnor}}$

当重构图像之间的相似度正常收敛时,最后两次的输出 $I_{R_{n-1}}$ 和 I_{R_n} 可以认为是正常样本的输入与输出图像,而两者之

间的相似度差值则可被认为是重构中的特征损失。同时,背景中的噪声作为容易被重构为正常的区域,在多次重构中显示出稳定的相似度,由输入原始图片 I_i 、第一次重构图片 I_{R_1} 之间的差值得出。上述提到的两个重构误差将在最终的输出之前被剔除。

而对于背景纹理难以重构的类别,迭代重构会加剧输出图像的模糊程度,引起模型崩溃。针对此情况,在迭代生成分支之外,设计了一个图像修复分支,仅对可疑区域进行局部精准修复。首先,通过 I_{R_1} 和 I_i 以获取初步的异常可疑区域。再根据可疑区域标定包含可疑区域的网格,用于图像修复。当异常区域的干扰特征被排除,模型就能够更好地修复缺失区域,从而提高异常定位的精度。图像局部精准修复过程如图5所示。

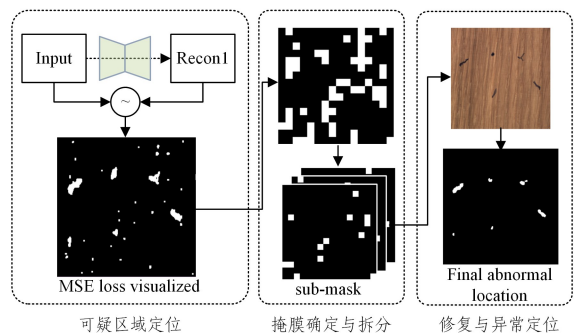


图5 精准修复分支的实现过程

Fig. 5 Implementation process of inpainting branch

4 实验结果与分析

4.1 数据集与实验环境设置

本次实验以 MVTec^[4] 数据集为对象进行相关实验。MVTec 数据集是由 MVTec 公司于 2019 年公开的用于工业场景下异常检测的数据集,并且对异常样本进行了像素级的标注。数据集中 15 类工业产品图像被预先分为训练集和测试集,其中训练集均为正常样本,共 3629 张,测试集包含正常样本 467 张以及各类异常样本共 1332 张。由于本文中检测任务为异常定位任务,因此将所有异常样本设置为统一“abnormal”类。

本文中所有实验均部署在 Ubuntu 16.04 操作系统上,在 GPU 环境下运行,使用显卡型号为 NVIDIA GeForce GTX1080(显存 4GB),深度学习框架配置为 pytorch1.11,python3.8。

在本次实验中,所有输入图像均被缩放为 256×256 大小进行训练,训练批次设置为 6,初始学习率设置为 0.0002,训练次数设置为 200,使用 Adam 优化器进行优化,并在训练过程中采用学习率指数衰减策略。

4.2 评价指标

本文采用 AUROC 曲线作为评价模型异常定位精度的指标。AUCROC 曲线是利用不同阈值下 recall(召回率)和 precision(准确率)的变化趋势所连成的曲线,对召回率和准确性进行了综合评价,能够较为准确地反映出模型的检测精度。两者的计算式如式(6)、式(7)所示。

$$\text{Recall} = \frac{TP}{TP + FN} \quad (6)$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad (7)$$

其中, TP 表示实际与预测均为异常的样本, FN 表示实际为异常但被判定为正常的样本, FP 表示实际为正常但被判定为异常的样本。

4.3 对比实验

为了验证本文方法在工业异常定位任务上的有效性,我们选择了经典的生成式模型以及近年来提出的基于生成式模型的异常定位方法进行对比实验。其中 OCGAN^[19] 为采用生成对抗网策略的自监督网络; R-VAE^[20] 使用 VAE 结构作为生成网络, 并采用循环生成策略定位异常; MemAE^[7] 为针对特征层进行改进的检测模型; CAE 为使用 ResNet34 搭建的自编码器生成式网络; SSM^[21], SMIAD^[12] 和 GP^[14] 为基于图像修复思想

的代表方法。MKD^[15] 为使用知识蒸馏进行模型训练的方法。

表 2 列出了上述模型在 MVTEc 数据集上的异常定位精度, 加粗数据为该类的最高精度, 下划线数据为排名前三的精度。实验数据显示, 本文方法在异常定位精度上优于其他方法, 在 6 个类别上达到最优, 在 12 个类别上排名前三。图 6 给出了不同方法下的 AUROC 曲线, 红色曲线为本文方法。所提方法在 multi-class 图像上保持了召回率和准确率的均衡性, 在工业异常定位任务中能够有效地定位异常, 并且具有通用性。此外, 通过观察 MKD 方法的实验数据与 AUROC 曲线发现, 此方法在第 I 类图像上取得了优秀的检测精度, 显示出了其在异常检测任务中的巨大潜力。

表 2 MVTEc 数据集下各生成式异常检测模型的检测结果

Table 2 Results of different anomaly location models on MVTEc dataset

类别	OCGAN	R-VAE	MemAE	SSM	CAE	SMIAD	MKD	GP	OURS
bottle	56.70	90.710	75.620	92.40	90.46	86	<u>92.77</u>	93	94.24
cable	56.40	70.45	73.55	72.30	79.56	92	79.59	94	<u>86.47</u>
capsule	63.70	63.18	82.72	87.70	85.58	<u>93</u>	95.63	90	<u>91.95</u>
carpet	54.60	40.45	53.70	82.60	65.92	<u>88</u>	<u>94.82</u>	96	82.97
grid	65.20	77.94	63.51	71.60	89.35	<u>97</u>	<u>89.42</u>	78	97.20
hazelnut	84.10	82.64	88.51	92.30	<u>96.79</u>	<u>97</u>	95.02	84	97.60
leather	74.90	51.12	64.35	<u>95.70</u>	83.52	86	98.07	90	88.06
metal_nut	53.40	48.48	72.96	87.20	85.49	92	82.16	<u>91</u>	94.61
pill	59.60	76.02	82.66	81.30	<u>92.70</u>	92	90.43	<u>93</u>	95.10
screw	70.80	41.60	91.65	65.20	93.21	<u>96</u>	94.92	<u>96</u>	97.27
tile	59.20	54.22	55.77	55.40	69.75	62	79.22	<u>80</u>	<u>87.74</u>
toothbrush	76.30	73.91	82.63	97.90	94.62	<u>96</u>	95.55	<u>96</u>	<u>97.00</u>
transistor	58.20	74.71	69.82	64.50	67.39	<u>85</u>	73.11	100	<u>88.22</u>
wood	65.50	94.74	64.03	56.10	62.66	80	<u>83.23</u>	81	81.82
zipper	62.40	70.33	71.76	90.40	81.82	90	<u>93.72</u>	99	<u>91.01</u>
AVE	64.07	67.37	72.88	79.51	82.59	88.8	<u>89.18</u>	<u>90.73</u>	91.42

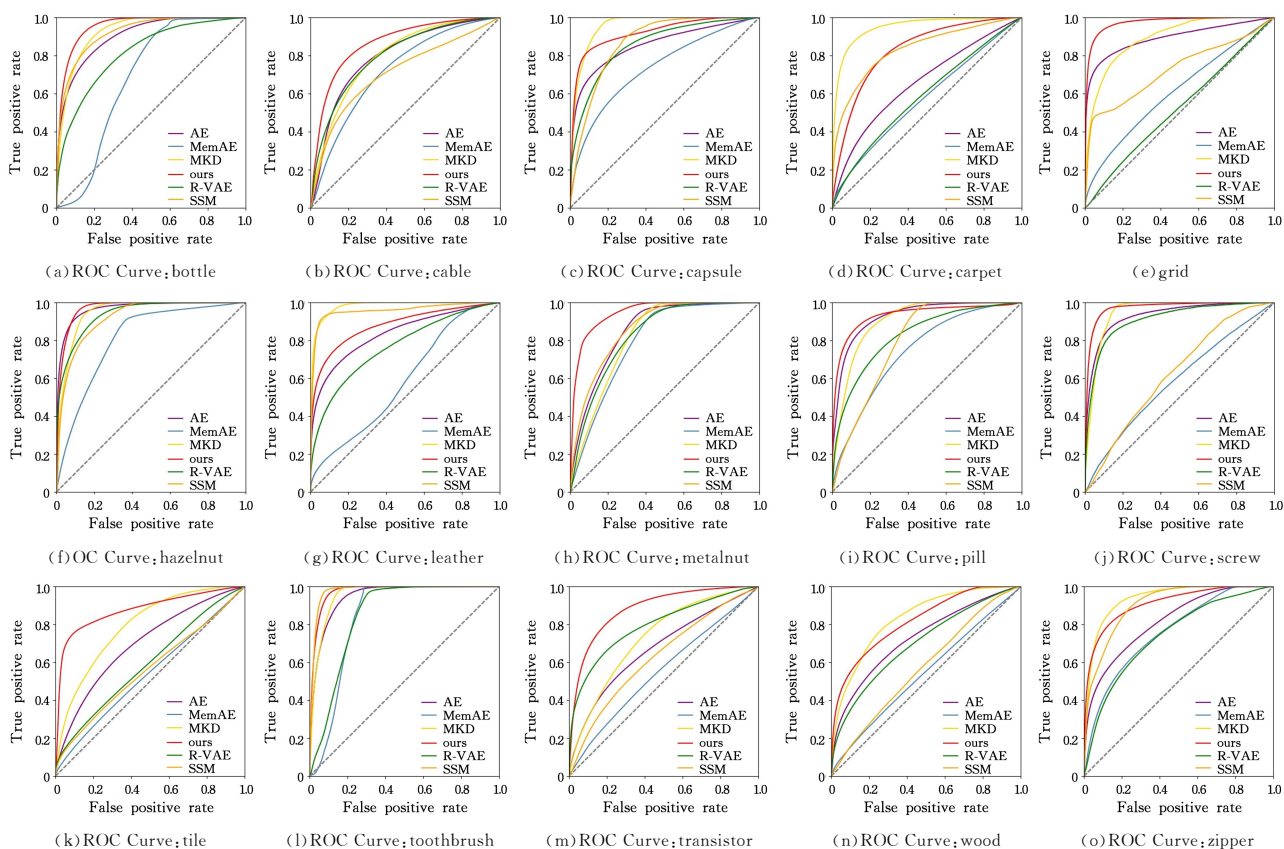


图 6 对比实验的 AUROC 曲线图(电子版为彩图)

Fig. 6 AUROC curve of comparative experiments

4.4 消融实验

为了验证各模块对生成式检测模型性能的改进效果,对本文提出的3个改进模块进行分模块评估。消融实验结果如表3所列,其中加粗数据为最优值。验证方法有效性的基线模型选择由ResNet34网络搭建的自编码器。添加的模块中(1)表示使用模糊遮蔽的对比样本生成模块,(2)表示添加了浅层特征融合与结构相似度优化的生成式模型,(3)表示多分支异常动态推理方法。为了测试所提出网络模型的通用性,对3个类别的提升精度分别进行测试。

表3 消融实验的结果
Table 3 Results of ablation experiment

(1)	(2)	(3)	I	II	III
—	—	—	77.11	90.4	79.56
✓			81.47	90.83	82.18
✓	✓		81.95	92.61	84.21
✓	✓	✓	89.13	95.29	86.47

在添加了模糊遮蔽后,第I类与第III类图像的异常定位精度提升幅度较大,这是由于第II类图像中背景的占比较大,使得遮蔽区域有较大可能落在背景区域,而其他两类中有效面积占比大,因此改善效果明显。添加了浅层特征融合与符合损失函数后,II、III类异常定位精度提升了2%,这说明边缘特征与结构特征被模型有效利用,能够提高模型对具有边缘特征图像异常定位的准确性;最后,在添加多分支异常推理模块后,三类图像均有大幅提升。这说明设计的多分支动态推理算法能够与生成式网络模型进行良好的适配,并且能够进一步发挥深度学习方法的优势,获取更高精度的异常定位精度。

4.5 超参数分析

为了使各改进模块间能够实现有效的配合,我们对改进部分的超参数进行了仔细的选择。首先对遮蔽单元的大小进行了选择,我们选择{8,16,32}这3种大小的遮蔽单元进行实验,其中当取值为16时获得最优平均性能。此外,针对损失函数中两部分的比重进行权衡,我们选择0.1:1,1:1,1:0.1这3组进行实验,实验发现当取值为1:0.1时取得最优平均性能。参数分析实验结果如表4所列,“/”用于区分类别,加粗数据为最优精度。可以看到,I、III类图像随 k 值变化有明显规律,但第I类图像变化不大,这是由于I类图像中各处特征较为接近,对遮蔽区域并不敏感。此外,通过 λ 参数实验发现,修复损失作为比重较小的调节损失函数时模型取得了更高的精度。这是由于推理过程中是以生成图像的整体质量作为算法基础,因此当生成损失占主导时,训练出的模型能够更好地与推理算法匹配。

表4 超参数分析实验的结果

Table 4 Results of hyperparameter experiment

	$\lambda=10$	$\lambda=1$	$\lambda=0.1$
$k=8$	88.47/91.04/86.00	88.56/92.03/86.51	88.97/94.98/86.06
$k=16$	88.78/91.18/85.79	89.05/92.16/ 86.55	89.13 /95.29/86.47
$k=32$	88.99/91.33/85.89	88.99/93.21/86.36	89.01/ 95.34 /86.16

为了分析3.2节中提出的基于结构相似度指标划分工业图像方法的鲁棒性,本文进行了参数实验。通过交换加权参数来测试不同参数对检测结果的影响。图6给出了参数实验结果。通过分析实验结果发现,对于70%的类别而言,基于数据的分类结果与最终测试结果相对应,但在其他四类(tile, wood, toothbrush 和 zipper)中在第III类图像对应的权重上取得了最佳效果。我们认为这是由于纹理的随机性与表征复杂所引起的。但我们认为所给出的权重仍然具有参考价值,因为在toothbrush和zipper的图像表征中,同时存在清晰的轮廓与明显的纹理。在缺少异常数据无法进行数据分析的情况下,会更偏向于将其归类于第三类图像,符合最终参数的测试结果。

wood, toothbrush 和 zipper)中在第III类图像对应的权重上取得了最佳效果。我们认为这是由于纹理的随机性与表征复杂所引起的。但我们认为所给出的权重仍然具有参考价值,因为在toothbrush和zipper的图像表征中,同时存在清晰的轮廓与明显的纹理。在缺少异常数据无法进行数据分析的情况下,会更偏向于将其归类于第三类图像,符合最终参数的测试结果。

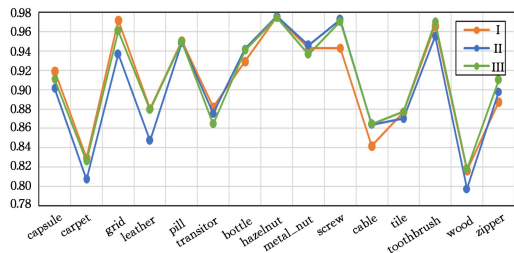


图6 加权参数分析实验的结果

Fig. 6 Results of weight hyperparameter experiments

4.6 可视化分析

为了更加直观地显示出模型的异常定位性能,我们对模型异常定位结果进行了可视化显示。实验结果如图7所示。其中,绿色标注为异常真实值,红色标注为异常预测值。

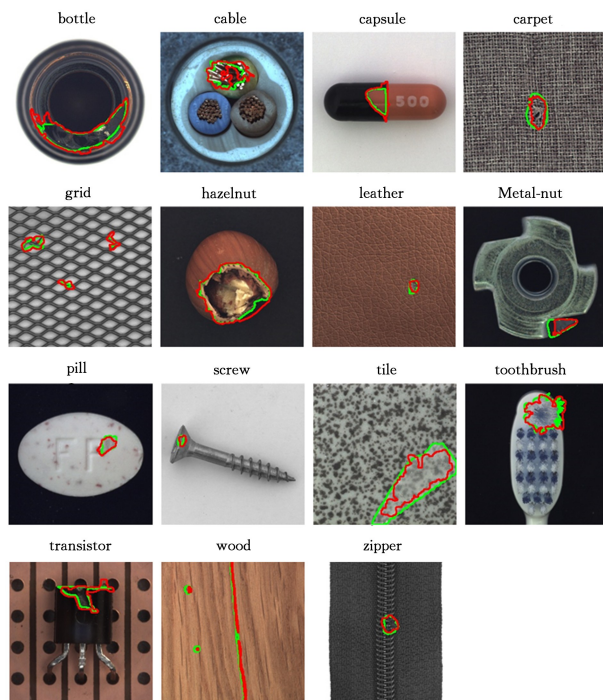


图7 异常定位结果可视化(电子版为彩图)

Fig. 7 Visualization of anomaly location results

从bottle, transistor和zipper类中可以看出,所提方法在物体本身像素值较低的情况下仍然能够完整地定位异常,保持了较好的异常敏感度。但在“cable”类中,由于电缆部位细节复杂,生成式模型输出图像中还原较为模糊,导致部分区域漏检的情况。同时,通过tile和toothbrush类的检测结果可以推断出,在异常内部与正常样本相似的情况下,所提方法仍然能够实现完整的异常检出,这对后续异常严重程度的判定具有指导意义。最后,所提方法对于较小异常也显示出了精准的定位性能,异常检出效果良好。

结束语 针对工业生产中,异常样本出现概率低、无法完

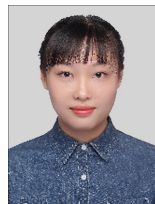
全预测的固有难题,提出了一种基于模糊遮蔽与动态推理的生成式工业异常定位模型,用于实现工业图像异常定位。通过对图像输入、网络模型和推理算法3部分的改进,提升了生成式模型在工业异常图像上的异常定位能力,在MVTec数据集上显示出了良好的异常定位效果。

生成的具有干扰信息的模拟异常样本与原始样本构成对比样本对,实现对生成式网络的对比训练,从而提高模型的代表能力。在生成式模型中加入浅层特征融合路径,保留更多边缘信息,加入结构相似度优化约束条件,使网络对纹理复杂和背景占比较大的图像仍然保持着较高的指标敏感度,获取更高质量的生成式输出。最后,设计了一个多分支异常推理算法,利用两个分支之间的互补,建立完整的异常判定机制,提高生成式模型的鲁棒性,对背景噪点有更为准确的判断。

本文方法在多类工业图像上均保持了较好的异常定位性能,这体现了所提方法有效地提升了生成式模型的异常定位性能。但目前生成式方法往往缺少对高维特征的利用。理论上来说,生成式模型生成输出图像的过程可以看成异常特征被“消除”的过程。此外,从对比实验中可以看出,知识蒸馏方法在工业异常检测任务中展示出了巨大的潜力。在后续的工作中,将进一步研究如何有效利用高维特征信息以及改进训练策略,以实现高效率的工业异常定位。

参考文献

- [1] TAO X, GONG X, ZHANG X, et al. Deep Learning for Unsupervised Anomaly Localization in Industrial Images: A Survey [J]. IEEE Transactions on Instrumentation and Measurement, 2022, 71: 1-21.
- [2] LV C, SHEN F, ZHANG F. Review of Image Anomaly Detection [J]. Acta Automatica Sinica, 2022, 48(6): 1402-1428.
- [3] CHOI J, KIM C. Unsupervised detection of surface defects: A two-step approach [C] // 2012 19th IEEE International Conference on Image Processing. 2012: 1037-1040.
- [4] BERGMANN P, FAUSER M, SATTLEGGER D, et al. MVTec AD—A comprehensive real-world dataset for unsupervised anomaly detection [C] // Proceedings of the IEEE/CVF Conference on CVPR. 2019: 9592-9600.
- [5] CHEN Y, TIAN Y, PANG G, et al. Deep One-Class Classification via Interpolated Gaussian Descriptor [J/OL]. (2022-05-24) [2022-07-31]. <https://arxiv.org/pdf/2101.10043v5.pdf>.
- [6] HAN J, CHENG J F, LI Y, et al. Self-supervised Deep Clustering Algorithm Based on Self-attention [J]. Computer Science, 2022, 49(3): 134-143.
- [7] GONG D, LIU L, LE V, et al. Memorizing normality to detect anomaly: Memory-augmented deep autoencoder for unsupervised anomaly detection [C] // Proceedings of the IEEE International Conference on Computer Vision. 2019: 1705-1714.
- [8] PARK H, NOH J, HAM B. Learning memory-guided normality for anomaly detection [C] // Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2020: 14372-14381.
- [9] ZHOU K, LI J, XIAO Y, et al. Memorizing Structure-Texture Correspondence for Image Anomaly Detection [J]. IEEE Transactions on Neural Networks and Learning Systems, 2021, 33(6): 2335-2349.
- [10] YANG L, JIANG A L, QIANG Y. Structure Preserving Unsupervised Feature Selection Based on Autoencoder and Manifold Regularization [J]. Computer Science, 2021, 48(8): 53-59.
- [11] VZ A, MK A, DS A. Reconstruction by inpainting for visual anomaly detection [J]. Pattern Recognition, 2021, 112(2): 107706.
- [12] LI Z, LI N, JIANG K, et al. Superpixel masking and inpainting for self-supervised anomaly detection [C] // British Machine Vision Conference. 2020: 7-10.
- [13] YAN X, ZHANG H, XU X, et al. Learning semantic context from normal samples for unsupervised anomaly detection [C] // Proceedings of the AAAI Conference on Artificial Intelligence. 2021: 3110-3118.
- [14] WANG S, WU L, CUI L, et al. Glancing at the Patch: Anomaly Localization with Global and Local Feature Comparison [C] // Computer Vision and Pattern Recognition. 2021: 254-263.
- [15] SALEHI M, SADJADI N, BASELIZADEH S, et al. Multiresolution Knowledge Distillation for Anomaly Detection [C] // Computer Vision and Pattern Recognition. 2021: 14897-14907.
- [16] XING P, JIANG X, TANG J H, et al. Feature Consistent Restricted Distillation Learning for Visual Anomaly Detection [J/OL]. Journal of Software. (2021-10-11) [2022-07-31]. <http://jos.org.cn/jos/article/abstract/Lf051>.
- [17] LI C L, SOHN K, YOON J, et al. Cutpaste: Self-Supervised Learning for Anomaly Detection and Localization [C] // Proceedings of the IEEE International Conference on Computer Vision. Nashville: IEEE Press, 2021: 9664-9674.
- [18] AHORÉ, ZIOU D. Image quality metrics: PSNR vs. SSIM [C] // 20th International Conference on Pattern Recognition (ICPR 2010). IEEE Computer Society, 2010: 2366-2369.
- [19] PERERA P, NALLAPATI R, BING X, OCGAN; One-Class Novelty Detection Using GANs With Constrained Latent Representations [C] // 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2019: 2898-2906.
- [20] DEHAENE D, FRIGO O, COMBEXELLE S, et al. Iterative energy-based projection on a normal data manifold for anomaly localization [J/OL]. (2020-12-10) [2022-10-25]. <https://arxiv.org/abs/2002.03734v1>.
- [21] HUANG C Q, XU Q W, WANG Y F, et al. Self-Supervised Masking for Unsupervised Anomaly Detection and Localization [J]. IEEE Transactions on Multimedia. doi: 10.1109/TMM.2022.3175611.



WU Tianyue, born in 1999, postgraduate, is a member of China Computer Federation. Her main research interests include image processing, self-supervised learning, and anomaly detection.



ZHANG Hui, born in 1983, Ph.D., professor. His main research interests are machine vision, sparse representation, and visual tracking.