

基于改进可控扩散模型的缺陷图像生成算法<sup>\*</sup>陈广庆<sup>1</sup> 陈雅惠<sup>1</sup> 周 鹏<sup>1</sup> 刘梓煜<sup>1</sup> 陈玉伦<sup>2</sup>

(1. 山东科技大学机械电子工程学院 青岛 266590; 2. 青岛普华重工机械有限公司 青岛 266400)

**摘 要:** 工业场景下,缺陷工件的获取和标注非常困难,对工件缺陷检测带来极大的阻碍。通过少量真实缺陷样本生成大量缺陷样本,极大地缓解了缺陷样本稀缺的问题,但是现有的缺陷生成方法普遍存在生成缺陷样本的表现真实性差和与掩模对齐性差的问题。针对现有问题,文中提出了一种新颖的可控扩散模型 AnomalyAlign 来生成与掩模高度对齐的逼真工业缺陷图像。AnomalyAlign 在继承文生图大模型 Stable Diffusion 的先验知识基础上,提出了强语义对齐文本提示生成器,通过该生成器获取语义层面上与真实图像更加对齐的文本提示,促进了模型的收敛;同时,AnomalyAlign 还提出了一种缺陷对齐损失来提高生成的缺陷图像和掩模之间的对齐性。通过 MVTec-AD 上的大量实验验证,AnomalyAlign 可以生成与掩模高度对齐的逼真且多样化的缺陷图像,并有效地提升了下游缺陷检测任务的性能。

**关键词:** 工业缺陷检测;可控扩散模型;图像生成;微调大模型

**中图分类号:** TN1      **文献标识码:** A      **国家标准学科分类代码:** 510.4

Defect image generation algorithm based on an improved  
controllable diffusion modelChen Guangqing<sup>1</sup> Chen Yahui<sup>1</sup> Zhou Peng<sup>1</sup> Liu Ziyu<sup>1</sup> Chen Yulun<sup>2</sup>

(1. School of Mechatronic Engineering, Shandong University of Science and Technology, Qingdao 266590, China;

2. Qingdao Puhua Heavy Industry Machinery Co., Ltd., Qingdao 266400, China)

**Abstract:** In industrial settings, the acquisition and annotation of defective workpieces pose significant challenges, severely hindering defect detection efforts. While generating a large number of defective samples from limited real-world samples effectively mitigates the issue of sample scarcity, existing defect generation methods are often constrained by suboptimal visual authenticity and poor alignment with defect masks. To address these limitations, this study introduces AnomalyAlign, a novel controllable diffusion model designed to synthesize highly realistic industrial defect images with precise mask alignment. Leveraging the foundational knowledge of the text-to-image model Stable Diffusion, AnomalyAlign incorporates a semantic-aligned text prompt generator to produce text prompts that achieve closer semantic alignment with real images, thereby accelerating model convergence. Furthermore, the model integrates a defect alignment loss function, which enhances the spatial consistency between generated defect images and their corresponding masks. Extensive experimental validation on the MVTec-AD dataset demonstrates that AnomalyAlign generates defect images with superior realism and diversity, while significantly improving the performance of downstream defect detection tasks.

**Keywords:** industrial defect detection; controllable diffusion model; image generation; fine-tuning large model

## 0 引 言

近年来,工业界普遍采用自动化的流水线来生产工件,这种生产模式极大地提高了生产效率。随着产品数量的增加,人工逐件检测缺陷变得非常困难,工业缺陷检测算法的研究迫在眉睫。工业缺陷检测算法利用计算机视觉、机器

学习和深度学习等技术手段来自动识别工件表面的缺陷区域。它通常用于制造业中,以提高产品质量、降低人工检测成本,并实现高效的质量控制<sup>[1]</sup>。由于自动化生产线性性能优异,缺陷样本很难收集,且人工标注缺陷区域复杂缓慢,这极大地制约有标签数据集的获取,使得训练有监督缺陷检测模型变得困难,给检测工作<sup>[2-4]</sup>带来了巨大挑战。为了

应对这一问题,目前的工业缺陷检测算法主要以无监督学习<sup>[5-6]</sup>和少样本有监督学习<sup>[7-8]</sup>为主。无监督学习通过分析正常数据的特征来检测缺陷情况,而少样本有监督学习则利用少量标注的缺陷样本进行训练,提升检测精度。由于缺乏真实缺陷域的限制和不充分的可辨识性,上述方法在缺陷定位任务上效果欠佳。

为了解决缺陷样本稀缺的问题,许多研究人员通过生成缺陷样本来辅助缺陷检测。一些非深度学习的方法<sup>[5-9]</sup>将少量现有的真实缺陷图像裁剪、粘贴到大量正常图像上来生成新的缺陷图像。这种方法虽然简便,但会破坏粘贴区域边缘的纹理特征,导致生成的缺陷图像在多样性和真实性上表现较差。随着深度学习的兴起,基于生成对抗网络(generative adversarial network, GAN)<sup>[10]</sup>的图像生成模型得到了快速发展,这些模型通过对抗训练生成逼真的缺陷图像。然而,现有基于 GAN 的缺陷图像生成方法<sup>[11-13]</sup>需要大量的样本进行训练,且存在收敛困难的问题。相比于基于 GAN 的生成方法,扩散模型<sup>[14]</sup>在学习数据分布方面表现出更为出色的能力。许多大模型<sup>[15-17]</sup>基于扩散模型进行构建,通过学习海量的输入数据,在图像生成方面展现了强大的性能。为了使大模型更好地符合特定任务的需求,许多微调大模型的方法被提出。AnomalyDiffusion<sup>[18]</sup>通过微调潜在扩散大模型<sup>[15]</sup>来生成缺陷图像,AnomalyDiffusion 采用文本反演<sup>[19]</sup>的方法来间接获取文本,将嵌入编码与图像中的缺陷部分隐式的关联起来,使得缺陷特征和嵌入编码可以耦合在一起,从而实现粗略的个性化编辑。类似于文本反演, DreamBooth<sup>[20]</sup>显式地将罕见单词与图像中的特定物体关联起来,通过增加上下文中的罕见单词,使模型能够识别和生成与这些罕见单词对应的图像特征。相比于通过文本提示与缺陷特征耦合的方法,文本反演和 DreamBooth 是从零开始训练的,因为它们使用的嵌入编码和罕见单词在初始阶段没有语义含义,导致这类方法的收敛效果较差,无法实现精细的图像编辑。SAA+<sup>[21]</sup>使用了笼统的缺陷描述,这种文本获取方法存在很强的主观性。总而言之,现有文本获取方法对大模型先验知识的利用不够充分。

上述缺陷图像生成方法在一定程度上缓解了缺陷样本稀缺的问题,但是生成的缺陷图像仍然存在两个问题,一是生成缺陷图像的表现真实性较差,二是生成缺陷图像与输入掩模之间的对齐性较差。本文提出的 AnomalyAlign 通过强语义对齐文本提示生成器充分利用大模型的先验知识,通过缺陷对齐损失,显式地提高生成的缺陷区域与输入掩模的对齐性,AnomalyAlign 可以生成与掩模高度对齐的逼真缺陷图像,优于现有的缺陷生成模型。

## 1 AnomalyAlign 模型

如图 1 所示,ControlNet<sup>[22]</sup>使用附加条件输入来控制预训练的大模型,有效地解决了文生图大模型中仅依赖文

本无法精确控制生成细节的问题。但作为一种通用的微调方法,它在进行异常图像生成时会存在真实性和对齐性较差的问题。

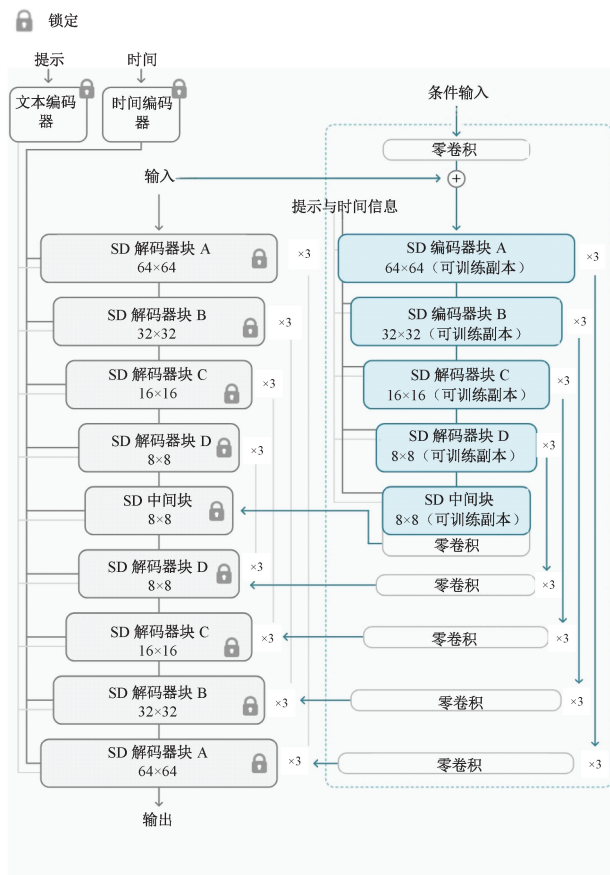


图1 ControlNet 模型整体框架

Fig. 1 Overall framework of the ControlNet model

本研究提出的 AnomalyAlign 模型的整体框架如图 2 所示,AnomalyAlign 开创性的将 ControlNet<sup>[22]</sup>应用到缺陷图像生成领域,此外,AnomalyAlign 在 ControlNet 的基础上提出了两个新颖的模块:强语义对齐文本提示生成器和缺陷对齐损失。强语义对齐文本提示生成器能够生成与输入的真实缺陷图像在嵌入空间中对齐性最好的文本提示,从而充分利用预训练大模型的先验知识,促进模型的快速收敛。缺陷对齐损失对真实掩模和生成缺陷图像的掩模做均方误差(mean squared error, MSE)损失函数,使生成的缺陷区域与输入掩模在对齐性上表现更好。训练 AnomalyAlign 所需的缺陷样本量较少,符合实际工业环境的需求。在推理阶段,只需要输入文本提示和掩模即可生成与输入掩模严格对齐的逼真缺陷图像。

AnomalyAlign 高效地利用了大模型海量的先验知识,可以生成质量高,与输入掩模对齐性好的缺陷图像,从而提升下游工业缺陷检测算法的性能。

### 1.1 强语义对齐文本提示生成器

为了解决现有方法对 Stable Diffusion 先验知识利用

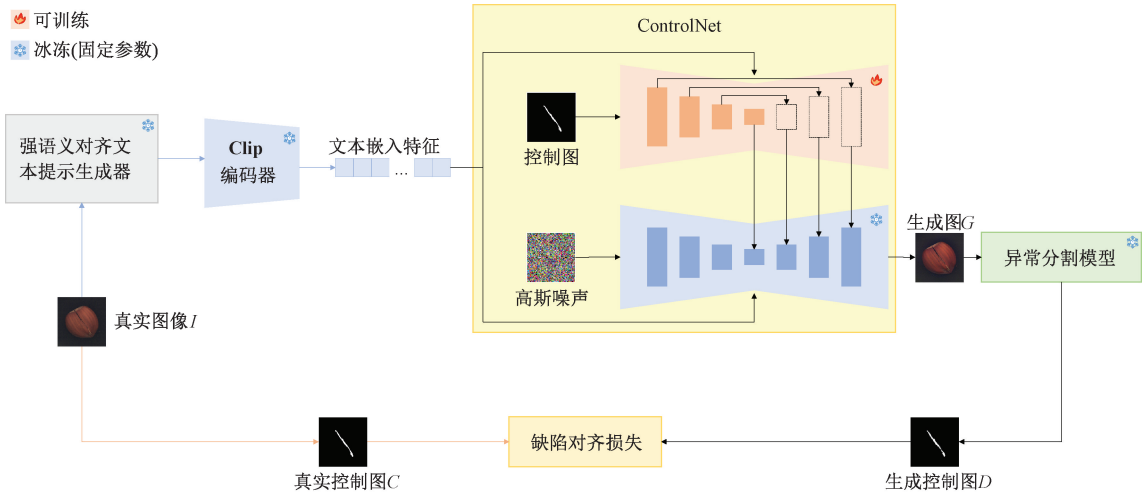


图 2 AnomalyAlign 模型整体框架

Fig. 2 Overall framework of the AnomalyAlign model

不足的问题,AnomalyAlign 借鉴了图 3 中 CLIP 模型强大的图像-文本对齐能力,来获取与输入异常图像在语义上对齐性最好的文本提示。

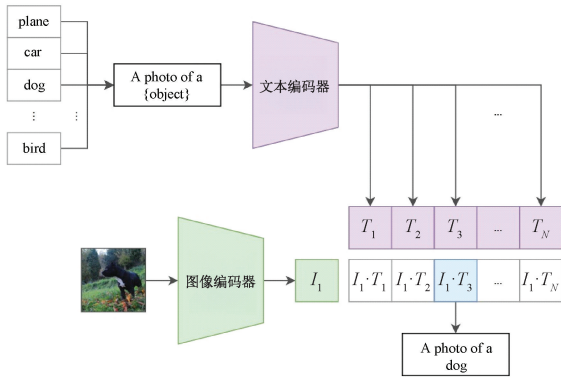


图 3 CLIP 模型推理过程

Fig. 3 Inference process of the CLIP model

AnomalyAlign 提出了一种强语义对齐文本提示生成器来获得包含丰富先验知识的客观文本提示,如图 4 所示。

首先,该模块采用了 SAA+<sup>[21]</sup>的“Anomaly Language Expression as Prompt”方法来获取语义词 $[y_o]$ ,但这样得到的语义词具有较强的主观性,涵盖的缺陷单词不够广泛。因此,该模块通过 WordNet<sup>[23]</sup>拓展语义词,得到同义词集合来构建一个语义词库 $[y_t]$ ,并从 CLIP<sup>[24]</sup>中针对 ImageNet 筛选出了 85 个模板,移除了不适用于缺陷检测的模板,作为句子的主干 $[y_s]$ 。将语义词库 $[y_t]$ 、工件类别和一系列句子主干 $[y_s]$ 组合在一起得到一个包含大量文本描述的文本描述库 $[w_j]$ 。最后,如式(1)所示,通过预训练的 CLIP 图像编码器  $Clip_{img}(\cdot)$  提取输入真实图像的图像嵌入特征  $\mathbf{A}$ ,通过式(2)中 CLIP 文本编码器  $Clip_{txt}(\cdot)$  提取文本描述库 $[w_j]$ 中的每一条文本描述的文本嵌入特征,并在嵌入空间中通过式(3)做余弦相似度比较,获得式(4)

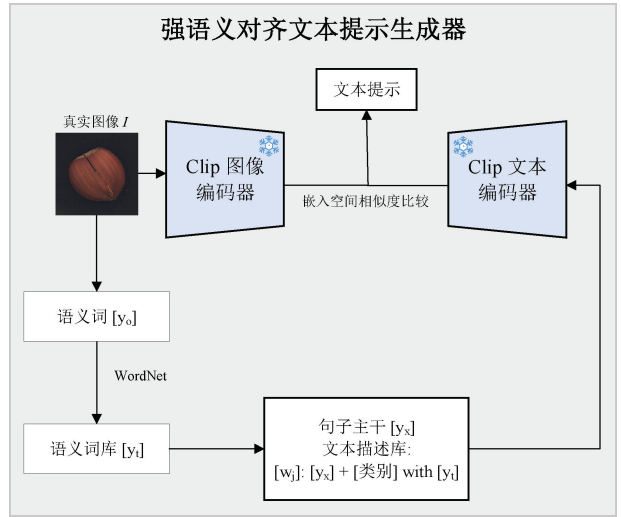


图 4 强语义对齐文本提示生成器结构

Fig. 4 Structure of the strong semantic alignment text prompt generator

中具有丰富先验知识的客观文本提示:

$$\mathbf{A} = Clip_{img}(I) \quad (1)$$

$$\mathbf{B} = Clip_{txt}(w_j) \quad (2)$$

$$Cos(\mathbf{A}, \mathbf{B}) = \frac{\sum_{i=1}^n \mathbf{A}_i \cdot \mathbf{B}_i}{\sqrt{\sum_{i=1}^n \mathbf{A}_i^2} \cdot \sqrt{\sum_{i=1}^n \mathbf{B}_i^2}} \quad (3)$$

$$j^* = \underset{j}{\operatorname{argmax}} \cos(\mathbf{A}, \mathbf{B}) \quad (4)$$

式中:  $I$  是输入的缺陷图像,  $\mathbf{A}$  是经过预训练的 CLIP 图像编码器提取的  $I$  的特征,  $\mathbf{B}$  是经过预训练的 CLIP 文本编码器提取的文本的特征,  $j$  是文本描述库中句子的索引,  $\cos(\mathbf{A}, \mathbf{B})$  是两个嵌入向量之间的余弦相似度,  $j^*$  是最大值对应的文本提示的索引。通过强语义对齐文本提示生成器得到的文本提示可以更好的利用 Stable Diffusion 的先验

知识,有利于模型的收敛,使生成的缺陷图像更加真实。

## 1.2 缺陷对齐损失

为了使生成的缺陷图像与输入的掩模具有更好的对齐性,AnomalyDiffusion<sup>[18]</sup>将掩模注入到交叉注意力图中来隐式地增强生成图像与掩模的对齐度。但是,这种隐式的方法并不能很好的解决对齐性差的问题,生成的缺陷区域仍然和掩模存在较大的差异。

为了更好的提高对齐性,AnomalyAlign 提出了一种缺陷对齐损失,这是一种显式地增强对齐性的方法。如图2所示,首先,通过预训练的缺陷检测模型提取生成图像的掩模,之后将该掩模与输入的掩模计算像素级的均方误差损失,从而可将对齐性量化成数值而作为损失函数使用。缺陷对齐损失  $L_a$  可由式(5)和式(6)获取,具体来说,首先如式(5),通过 AnomalyDiffusion 提供的预训练缺陷检测模型来预测生成图像  $G$  的掩模  $D$ 。其次,通过式(6)计算预测掩模  $D$  和真实掩模  $C$  的均方误差来作为损失函数  $L_a$ 。

$$D = \text{AnoDiff}(G) \quad (5)$$

$$L_a = \text{MSE}(C, D) \quad (6)$$

式中:  $C$  是输入的缺陷图像对应的掩模,  $G$  是生成的图像,  $D$  是通过 AnomalyDiffusion 提供的预训练缺陷检测模型来预测得到的掩模,  $L_a$  是在  $C$  和  $D$  之间做均方误差损失得到的。缺陷对齐损失  $L_a$  需要和式(7)中扩散模型的重构损失  $L_{diff}$  一起使用。

$$L_{diff} = E_{x_0, t, \epsilon_t, \epsilon_v, \epsilon \sim N(0,1)} [\| \epsilon_\theta(x_t, t, c_t, c_v) - \epsilon \|_2^2] \quad (7)$$

$$L = L_{diff} + 0.01 \cdot L_a \quad (8)$$

式中:  $c_v$  是图像条件,  $c_t$  是文本提示,时间步是  $t$ ,  $x_0$  是输入数据,  $\epsilon$  是从高斯分布中采样的噪声映射。总的损失函数  $L$  是  $L_{diff}$  和  $L_a$  的加权和。

训练 AnomalyAlign 分两阶段进行,第1阶段使用式(8)的损失函数,以提高在训练初期生成缺陷区域与输入掩模的对齐性,加快模型收敛。由于掩模  $D$  的预测存在误差,为了防止这种噪声在模型训练后期带来的不利影响,第2阶段只使用式(7)作为损失函数。

## 2 实验设置

### 2.1 数据集

AnomalyAlign 的实验是在 MVTEC-AD 数据集<sup>[25]</sup>上进行的, MVTEC 是一个专门用于工业缺陷检测的标准数据集,包含来自15类不同工业产品的高质量图像,每个类别中的图像分为“正常”和“缺陷”两类。正常图像展示了无缺陷的产品,而缺陷图像则包含各种缺陷,如裂纹、划痕、污染、缺失部件等,总共有73个缺陷类别,缺陷图像中的缺陷类型和位置是不确定的。训练集有3629张,全部为正常图像;测试集有1725张,分为缺陷图像和正常图像。

### 2.2 数据集划分

AnomalyAlign 的训练集来自 MVTEC-AD 中缺陷样本

的前1/3,由于样本较少,为了防止模型很快的过拟合,需要将真实的样本进行随机裁剪、旋转和平移等简单的数据增强操作,将每个真实缺陷样本扩展到100个增强样本,以此用来训练该生成模型。推理生成图像阶段, AnomalyAlign 使用的掩模来自 AnomalyDiffusion 提供的掩模生成方法。AnomalyAlign 为每类缺陷分别生成了1000个样本对,下游缺陷检测任务使用的模型是 AnomalyDiffusion 开源的一个基础 U-Net 模型,训练集由这些生成样本和 MVTEC-AD 中的训练集组成,下游缺陷检测任务测试集中的缺陷样本只使用 MVTEC-AD 中缺陷样本的后2/3。

### 2.3 超参数设置

训练 AnomalyAlign 时,在图像输入网络之前将图像分辨率动态的调整到  $256 \times 256$  像素大小,使用 Adam 优化器更新模型参数,批量大小为8,学习率为  $10^{-5}$ ,训练100个周期,缺陷对齐损失只在前40个周期内使用。下游缺陷检测任务使用的 U-Net 模型来自 AnomalyDiffusion 的开源项目。

### 2.4 量化指标

在图像生成中,初始得分(inception score, IS)旨在评估生成图像的质量和类别多样性。其计算基于生成图像通过预训练分类网络得到的类别分布信息。IS 的取值越大,表明生成图像更加清晰且涵盖了更丰富的类别分布,意味着模型生成的样本质量较高且内容多样性更强。相反,较低的 IS 值可能暗示生成图像模糊,或者生成内容趋于单一,不够丰富多样。簇内成对 LPIPS 距离(intra-cluster pairwise lpips, IC-LPIPS)用于衡量同一类别簇内图像的多样性,较高的 IC-LPIPS 值表明生成图像在视觉特征上存在较大差异,生成结果更加多样,而较低的值可能表示簇内图像过于相似,缺乏多样性。

在缺陷检测中,曲线下面积(area under the receiver operating characteristic curve, AUROC)衡量模型区分正负样本的能力,其取值接近1.0表示模型对缺陷检测表现优异,接近0.5则说明效果接近随机猜测。平均准确率(average precision, AP)综合评估模型在不同阈值下对正样本的检测性能,取值越高表示模型能在多个阈值下保持较好的精度和召回率。F1-max 得分用于权衡准确率(precision)和召回率(recall),较高的 F1-max 值表示模型在正样本检测上达到了较好的平衡,既能减少漏检,又能控制误报,而较低的值说明模型在这两个方面表现不佳。这些指标的取值大小共同反映了生成图像的质量、多样性以及缺陷检测模型的准确性和鲁棒性。

## 3 实验结果与分析

### 3.1 不同缺陷生成模型对比实验

缺陷生成方法的相关实验分为两部分进行,第1部分是生成质量的比较;第2部分是缺陷生成模型在下游检测



任务上的性能比较,具体来说,使用各种缺陷生成模型分别生成缺陷样本,并使用经典的 U-Net 模型进行训练。

1)生成质量

为了直观的展示 AnomalyAlign 的生成质量,图 5 展示了 AnomalyAlign 与 DFMGAN 和 AnomalyDiffusion 的可视化比较。对于“Hazelnut”类别,DFMGAN 生成的缺陷表观真实性很差,AnomalyDiffusion 生成的缺陷区域与掩膜的对齐性差,AnomalyAlign 生成了与真实缺陷最相似的表观和与掩膜对齐性最好的缺陷区域。对于其他类别同样存在 DFMGAN 生成的缺陷表观真实性差,AnomalyDiffusion 生成的缺陷区域与掩膜的对齐性差的问题,相反,AnomalyAlign 生成的缺陷表观最真实,与输入的掩膜对齐性最好。

表 1 展示了 AnomalyAlign 与 Crop&Paste<sup>[9]</sup>、SDGAN<sup>[11]</sup>、DefectGAN<sup>[12]</sup>、DFMGAN<sup>[13]</sup>、AnomalyDiffusion<sup>[18]</sup> 和 ControlNet<sup>[22]</sup> 生成质量的量化比较。对于 IS 值,可以看出 AnomalyAlign 在 IS 平均值上表现最优,特别地 AnomalyAlign 在“pill”和“toothbrush”工件种类上比现有的最优模型 AnomalyDiffusion 分别提高了近一倍和一倍以上,证明了 AnomalyAlign 生成的图像质量最好,真实性最高。对于 IC-LPIPS 值,AnomalyAlign 在平均值上表现

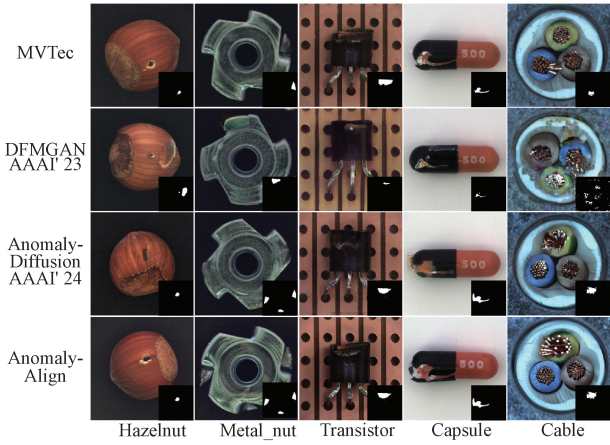


图 5 不同方法生成结果的可视化比较  
Fig. 5 Visual cmparison of generation results by different methods

出最好的效果,尤其对于“pill”和“toothbrush”工件, AnomalyAlign 比 DFMGAN 和 AnomalyDiffusion 提高了一倍以上,证明了 AnomalyAlign 生成的图像多样性最好。定量实验显示 AnomalyAlign 在 IS 和 IC-LPIPS 上取得了最好的性能,充分证明了利用 AnomalyAlign 生成的缺陷图像在真实性和多样性方面优于现有的缺陷生成方法。

表 1 生成质量的量化结果

Table 1 Quantitative results of generation quality

| 种类         | ControlNet |      | SDGAN       |      | Defect-GAN  |      | DFMGAN |             | Anomaly-Diffusion |             | Anomaly-Align |             |
|------------|------------|------|-------------|------|-------------|------|--------|-------------|-------------------|-------------|---------------|-------------|
|            | IS         | IC-L | IS          | IC-L | IS          | IC-L | IS     | IC-L        | IS                | IC-L        | IS            | IC-L        |
| bottle     | 1.52       | 0.11 | 1.57        | 0.06 | 1.39        | 0.07 | 1.62   | 0.12        | 1.25              | 0.12        | <b>1.69</b>   | <b>0.18</b> |
| cable      | 1.62       | 0.30 | <b>1.89</b> | 0.19 | 1.70        | 0.22 | 1.96   | 0.25        | 1.66              | 0.27        | 1.83          | <b>0.31</b> |
| capsule    | 1.11       | 0.16 | 1.49        | 0.03 | 1.59        | 0.04 | 1.59   | 0.11        | 1.47              | 0.06        | <b>1.72</b>   | <b>0.18</b> |
| carpet     | 1.13       | 0.19 | 1.18        | 0.11 | <b>1.24</b> | 0.12 | 1.23   | 0.13        | 0.92              | 0.09        | 1.01          | <b>0.21</b> |
| grid       | 2.26       | 0.31 | 1.95        | 0.10 | 2.01        | 0.12 | 1.97   | 0.13        | 2.62              | 0.28        | <b>2.65</b>   | <b>0.44</b> |
| hazelnut   | 2.14       | 0.18 | 1.85        | 0.16 | 1.87        | 0.19 | 1.93   | 0.24        | 1.89              | 0.19        | <b>2.33</b>   | <b>0.20</b> |
| leather    | 1.21       | 0.12 | 2.04        | 0.12 | <b>2.12</b> | 0.14 | 2.06   | 0.17        | 1.36              | 0.10        | 1.41          | <b>0.18</b> |
| metal_nut  | 1.48       | 0.21 | 1.45        | 0.28 | 1.47        | 0.30 | 1.49   | <b>0.32</b> | <b>1.96</b>       | 0.27        | 1.67          | 0.26        |
| pill       | 2.48       | 0.32 | 1.61        | 0.07 | 1.61        | 0.10 | 1.63   | 0.16        | 1.33              | 0.11        | <b>2.51</b>   | <b>0.35</b> |
| screw      | 1.55       | 0.25 | 1.17        | 0.10 | 1.19        | 0.12 | 1.12   | 0.14        | 1.31              | <b>0.35</b> | <b>1.61</b>   | 0.29        |
| tile       | 1.59       | 0.26 | <b>2.53</b> | 0.21 | 2.35        | 0.22 | 2.39   | 0.22        | 1.94              | 0.25        | 1.72          | <b>0.29</b> |
| toothbrush | 2.12       | 0.48 | 1.78        | 0.03 | 1.85        | 0.03 | 1.82   | 0.18        | 1.37              | 0.10        | <b>3.07</b>   | <b>0.53</b> |
| transistor | 1.59       | 0.29 | <b>1.76</b> | 0.13 | 1.47        | 0.13 | 1.64   | 0.25        | 1.51              | 0.19        | 1.70          | <b>0.34</b> |
| wood       | 2.16       | 0.27 | 2.12        | 0.25 | 2.19        | 0.29 | 2.12   | <b>0.35</b> | 1.92              | 0.22        | <b>2.31</b>   | 0.33        |
| zipper     | 1.86       | 0.22 | 1.25        | 0.10 | 1.25        | 0.10 | 1.29   | <b>0.27</b> | 1.42              | 0.11        | <b>2.04</b>   | 0.26        |
| 平均值        | 1.72       | 0.24 | 1.71        | 0.13 | 1.69        | 0.15 | 1.72   | 0.20        | 1.60              | 0.18        | <b>1.95</b>   | <b>0.29</b> |

2)缺陷检测与定位

图 6 定性地比较了 AnomalyAlign 与 AnomalyDiffusion 缺陷定位的结果,可以看出 AnomalyAlign 在缺陷定位上的可视化结果比现有最优模型 AnomalyDiffusion 具有更

高的置信度。特别是对于“Screw”类别,与 AnomalyDiffusion 相比,AnomalyAlign 分割出的缺陷区域与真实标签的重叠度更高,避免了 AnomalyDiffusion 中可能出现过度和不足的分割现象。总的来说,AnomalyAlign

在精确度上表现出更好的性能。

表 2 定量地比较了 AnomalyAlign 与现有的缺陷生成方法 DRAEM<sup>[5]</sup>、DFMGAN<sup>[13]</sup>、AnomalyDiffusion<sup>[18]</sup> 和 ControlNet<sup>[22]</sup> 的像素级 (判断每个像素是否为缺陷像素) 指标,结果显示,AnomalyAlign 在缺陷定位任务中取得了最优的定量结果,尤其是在 AP 指标上,提升幅度显著。具体而言,AnomalyAlign 的 AP 指标比 DRAEM、DFMGAN 和 AnomalyDiffusion 分别提高了 25.1%、16.5% 和 5.4%。这一结果充分证明了 AnomalyAlign 在缺陷定位任务中的优越性,表明该方法在定位缺陷区域方面,明显优于其他现有的生成式缺陷检测方法,具有更强的实际应用潜力。

表 3 是图像级 (判断整张图像是否为缺陷图像) 指标的定量实验结果,可以看出 AnomalyAlign 在缺陷检测任务中的表现出众,尤其是 AUROC 指标提升显著,比 DRAEM、DFMGAN 和 AnomlayDiffusion 分别提升了 4.3%、11.7% 和 1.3%,充分证明了 AnomalyAlign 在缺陷

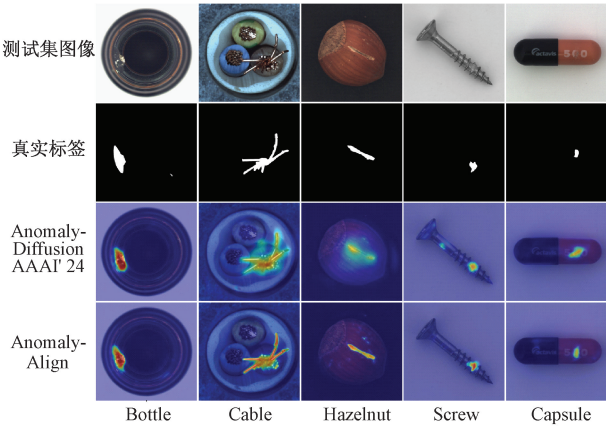


图 6 使用 AnomalyDiffusion 和 AnomalyAlign 生成的样本训练 U-Net 进行缺陷定位的定性比较

Fig. 6 Qualitative comparison of defect localization using samples generated by AnomalyDiffusion and AnomalyAlign for training U-Net

表 2 缺陷生成方法的像素级缺陷定位结果

Table 2 Pixel-level defect localization results of defect generation methods

| 种类         | DRAEM |      |        | DFMGAN |             |             | ControlNet |      |        | AnomalyDiffusion |             |             | AnomalyAlign |             |             |
|------------|-------|------|--------|--------|-------------|-------------|------------|------|--------|------------------|-------------|-------------|--------------|-------------|-------------|
|            | AUC   | AP   | F1-max | AUC    | AP          | F1-max      | AUC        | AP   | F1-max | AUC              | AP          | F1-max      | AUC          | AP          | F1-max      |
| bottle     | 96.7  | 80.2 | 74.0   | 98.9   | 90.2        | 83.9        | 96.7       | 84.1 | 81.6   | 99.4             | <b>92.7</b> | <b>86.0</b> | <b>99.4</b>  | 92.3        | 85.7        |
| cable      | 80.3  | 21.8 | 28.3   | 97.2   | 81.0        | 75.4        | 89.9       | 70.0 | 70.2   | <b>98.4</b>      | <b>82.6</b> | <b>75.4</b> | 93.8         | 77.1        | 74.9        |
| capsule    | 76.2  | 25.5 | 32.1   | 79.2   | 26.0        | 35.0        | 91.7       | 38.4 | 47.2   | <b>97.1</b>      | <b>41.7</b> | 42.9        | 95.3         | 41.5        | <b>50.4</b> |
| carpet     | 92.6  | 43.0 | 41.9   | 90.6   | 33.4        | 38.1        | 97.1       | 77.4 | 72.4   | 94.4             | 58.7        | 57.3        | <b>99.2</b>  | <b>84.3</b> | <b>76.0</b> |
| grid       | 99.1  | 59.3 | 58.7   | 75.2   | 14.3        | 20.5        | 95.3       | 46.6 | 48.2   | <b>98.2</b>      | 44.9        | 46.4        | 98.1         | <b>49.6</b> | <b>50.2</b> |
| hazelnut   | 98.8  | 73.6 | 68.5   | 99.7   | 95.2        | 89.5        | 97.3       | 86.1 | 86.4   | 98.8             | 73.0        | 66.5        | <b>99.7</b>  | <b>94.6</b> | <b>90.2</b> |
| leather    | 98.5  | 67.6 | 65.0   | 98.5   | 68.7        | 66.7        | 94.4       | 70.4 | 64.1   | <b>99.8</b>      | 73.0        | 68.3        | 99.7         | <b>84.3</b> | <b>76.9</b> |
| metal_nut  | 96.9  | 84.2 | 74.5   | 99.3   | 98.1        | 94.5        | 94.7       | 78.5 | 64.8   | 99.6             | <b>98.1</b> | <b>92.2</b> | <b>99.6</b>  | 95.4        | 91.7        |
| pill       | 95.8  | 45.3 | 53.0   | 81.2   | 67.8        | 72.6        | 97.4       | 76.2 | 73.9   | <b>99.7</b>      | <b>94.9</b> | <b>87.7</b> | 99.6         | 92.6        | 86.8        |
| screw      | 91.0  | 30.1 | 35.7   | 58.8   | 2.2         | 5.3         | 91.1       | 36.8 | 35.1   | 92.8             | 24.8        | 34.4        | <b>96.6</b>  | <b>43.4</b> | <b>43.1</b> |
| tile       | 98.5  | 93.2 | 87.8   | 99.5   | <b>97.1</b> | <b>91.6</b> | 95.1       | 80.0 | 81.4   | 99.7             | 96.2        | 88.7        | <b>99.7</b>  | 94.7        | 90.9        |
| toothbrush | 93.8  | 29.5 | 28.4   | 96.4   | <b>75.9</b> | <b>72.6</b> | 95.4       | 69.1 | 66.8   | <b>98.8</b>      | 68.9        | 70.3        | 98.2         | 73.8        | 69.1        |
| transistor | 76.5  | 31.7 | 24.2   | 96.2   | 81.2        | 77.0        | 95.4       | 76.1 | 66.3   | <b>99.4</b>      | <b>92.3</b> | <b>84.5</b> | 98.9         | 86.8        | 78.8        |
| wood       | 98.8  | 87.8 | 80.9   | 95.3   | 70.7        | 65.8        | 98.4       | 76.2 | 76.8   | 98.4             | 77.6        | 70.0        | <b>99.3</b>  | <b>89.2</b> | <b>82.4</b> |
| zipper     | 93.4  | 65.4 | 64.7   | 92.9   | 65.6        | 64.9        | 94.9       | 72.7 | 68.4   | <b>99.6</b>      | 87.5        | 79.2        | 99.5         | <b>88.1</b> | <b>81.1</b> |
| 平均值        | 92.2  | 54.1 | 53.1   | 90.0   | 62.7        | 62.1        | 95.0       | 69.2 | 66.9   | 98.3             | 73.8        | 70.0        | <b>98.4</b>  | <b>79.2</b> | <b>75.2</b> |

检测任务中的有效性。

3.2 不同缺陷检测模型对比实验

为了进一步验证 AnomalyAlign 的先进性,将 AnomalyAlign 与现有的一些先进的缺陷检测方法 DRAEM<sup>[5]</sup>、DevNet<sup>[7]</sup>、DRA<sup>[8]</sup>、CFA<sup>[26]</sup>、RD4AD<sup>[27]</sup>、PatchCore<sup>[28]</sup>、KDAC<sup>[29]</sup>、SSPCAB<sup>[30]</sup> 和 AnomalyDiffusion<sup>[18]</sup>, 进行了一项对比实验,具体来说,DRAEM、CFA、RD4AD、PatchCore、KDAC 和 SSPCAB 是无监督缺陷检测算法,

DevNet 和 DRA 是有监督异常检测算法, AnomalyDiffusion 和 AnomalyAlign 是缺陷生成方法,它们通过生成大量训练样本并在一个经典的 U-Net 模型上进行地训练。表 4 是这些方法在像素级缺陷定位上的性能 (AUROC/AP), 可以看出相比于这些模型, AnomalyAlign 取得效果较好,特别是对于 AP 指标,比传统的有监督缺陷检测算法 DevNet 提升了 29.9%,比现有的最优算法 AnomalyDiffusion 提升了 5.4%。

表 3 缺陷生成方法的图像级缺陷检测结果

Table 3 Image-level defect detection results of defect generation methods

| 种类         | DRAEM        |              |              | DFMGAN |              |        | ControlNet |       |        | AnomalyDiffusion |              |              | AnomalyAlign |              |              |
|------------|--------------|--------------|--------------|--------|--------------|--------|------------|-------|--------|------------------|--------------|--------------|--------------|--------------|--------------|
|            | AUC          | AP           | F1-max       | AUC    | AP           | F1-max | AUC        | AP    | F1-max | AUC              | AP           | F1-max       | AUC          | AP           | F1-max       |
| bottle     | 99.3         | 99.8         | 98.9         | 99.3   | 99.8         | 97.7   | 99.5       | 99.4  | 99.1   | 100.0            | 100.0        | 100.0        | <b>100.0</b> | <b>100.0</b> | <b>100.0</b> |
| cable      | 72.1         | 83.2         | 79.2         | 95.9   | 97.8         | 93.8   | 97.7       | 96.4  | 95.6   | <b>99.4</b>      | <b>99.3</b>  | 95.8         | 98.9         | 98.9         | <b>97.5</b>  |
| capsule    | 93.2         | 98.7         | 94.0         | 92.8   | 98.5         | 94.5   | 94.9       | 96.2  | 92.3   | 95.2             | 98.5         | 93.9         | <b>96.7</b>  | <b>98.9</b>  | <b>95.4</b>  |
| carpet     | <b>95.3</b>  | <b>98.7</b>  | <b>93.4</b>  | 67.9   | 87.9         | 87.3   | 90.4       | 94.1  | 89.9   | 87.7             | 93.9         | 89.1         | 94.1         | 97.6         | 92.6         |
| grid       | 99.8         | 99.9         | 98.8         | 73.0   | 90.4         | 85.4   | 100.0      | 100.0 | 99.9   | 98.2             | 99.3         | 98.6         | <b>100.0</b> | <b>100.0</b> | <b>99.9</b>  |
| hazelnut   | 100.0        | 100.0        | <b>100.0</b> | 99.9   | <b>100.0</b> | 99.0   | 99.1       | 99.0  | 98.6   | 98.7             | 98.8         | 96.8         | <b>100.0</b> | 99.9         | 99.9         |
| leather    | 100.0        | 100.0        | 100.0        | 99.9   | 100.0        | 99.2   | 99.9       | 100.0 | 100.0  | <b>100.0</b>     | 100.0        | 100.0        | 99.9         | <b>100.0</b> | <b>100.0</b> |
| metal_nut  | 97.8         | 99.6         | 97.6         | 99.3   | 99.8         | 99.2   | 99.1       | 98.2  | 96.4   | 100.0            | 100.0        | <b>100.0</b> | <b>100.0</b> | <b>100.0</b> | 99.9         |
| pill       | 94.4         | 98.9         | 95.8         | 68.7   | 91.7         | 91.4   | 97.9       | 98.5  | 96.1   | 99.3             | 99.7         | 97.9         | <b>99.7</b>  | <b>99.9</b>  | <b>98.9</b>  |
| screw      | 88.5         | 96.3         | 89.3         | 22.3   | 64.7         | 85.3   | 89.2       | 94.1  | 88.5   | 86.1             | 93.8         | 85.3         | <b>95.5</b>  | <b>98.0</b>  | <b>92.4</b>  |
| tile       | 100.0        | 100.0        | 100.0        | 100.0  | 100.0        | 100.0  | 100.0      | 100.0 | 99.9   | 100.0            | 100.0        | <b>100.0</b> | <b>100.0</b> | <b>100.0</b> | 99.9         |
| toothbrush | 99.4         | 99.8         | 97.6         | 100.0  | 100.0        | 100.0  | 99.1       | 99.4  | 97.7   | <b>100.0</b>     | <b>100.0</b> | <b>100.0</b> | 99.6         | 99.9         | 97.7         |
| transistor | 79.6         | 80.5         | 71.4         | 90.8   | 92.5         | 88.9   | 97.2       | 98.4  | 96.3   | 99.7             | 99.5         | 98.1         | <b>99.8</b>  | <b>99.7</b>  | <b>98.3</b>  |
| wood       | <b>100.0</b> | <b>100.0</b> | <b>100.0</b> | 98.4   | 99.4         | 98.8   | 97.2       | 98.9  | 97.1   | 99.4             | 99.7         | 97.7         | 99.5         | 99.7         | 98.7         |
| zipper     | 100.0        | 100.0        | 100.0        | 99.7   | 99.9         | 99.4   | 99.8       | 99.5  | 98.2   | 100.0            | 100.0        | 100.0        | <b>100.0</b> | <b>100.0</b> | <b>100.0</b> |
| 平均值        | 94.6         | 97.0         | 94.4         | 87.2   | 94.8         | 94.7   | 97.4       | 98.1  | 96.4   | 97.6             | 98.8         | 96.9         | <b>98.9</b>  | <b>99.5</b>  | <b>98.1</b>  |

表 4 与现有缺陷检测模型的比较

Table 4 Comparison with existing defect detection models

| 种类         | KDAD      | DRAEM                     | SSPCAB            | CFA               | RD4AD             | PatchCore | DevNet    | DRA       | Anomaly-Diffusion         | Anomaly-Align             |
|------------|-----------|---------------------------|-------------------|-------------------|-------------------|-----------|-----------|-----------|---------------------------|---------------------------|
| bottle     | 94.7/50.5 | 99.1/88.5                 | 98.9/88.6         | 98.9/50.9         | 98.8/51.0         | 97.6/75.0 | 96.7/67.9 | 91.7/41.5 | 99.4/ <b>92.7</b>         | <b>99.4</b> /92.3         |
| cable      | 79.2/11.6 | 94.8/61.4                 | 93.1/52.1         | 98.4/79.8         | <b>98.8</b> /77.0 | 96.8/65.9 | 97.9/67.6 | 86.1/34.8 | 98.4/ <b>82.6</b>         | 93.8/77.1                 |
| capsule    | 96.3/9.9  | 97.6/47.9                 | 90.4/48.7         | 98.9/ <b>71.1</b> | <b>99.0</b> /60.5 | 98.6/46.6 | 91.1/46.6 | 88.5/11.0 | 97.1/41.7                 | 95.3/41.5                 |
| carpet     | 91.5/45.8 | 96.3/62.5                 | 92.3/49.1         | 99.1/47.7         | <b>99.4</b> /46.0 | 98.7/65.0 | 94.6/19.6 | 98.2/54.0 | 94.4/58.7                 | 99.2/ <b>84.3</b>         |
| grid       | 89.0/7.6  | 99.5/53.2                 | <b>99.6</b> /58.2 | 98.6/ <b>82.9</b> | 98.0/75.4         | 97.2/23.6 | 90.2/44.9 | 86.2/28.6 | 98.2/44.9                 | 98.1/49.6                 |
| hazelnut   | 95.0/34.2 | 99.5/88.1                 | 99.6/94.5         | 98.5/80.2         | 94.2/57.2         | 97.6/55.2 | 76.9/46.8 | 88.8/20.3 | 98.8/73.0                 | <b>99.7</b> / <b>94.6</b> |
| leather    | 98.2/26.7 | 98.8/68.5                 | 97.2/60.3         | 96.2/60.9         | 96.6/53.5         | 98.9/43.4 | 94.3/66.2 | 97.2/5.1  | <b>99.8</b> /73.0         | 99.7/ <b>84.3</b>         |
| metal nut  | 81.7/30.6 | 98.7/91.6                 | 99.3/95.1         | 98.6/74.6         | 97.3/53.8         | 97.5/86.6 | 93.3/57.4 | 80.3/30.6 | 99.6/ <b>98.1</b>         | 99.6/95.4                 |
| pill       | 90.1/23.1 | 97.7/44.8                 | 96.5/48.1         | 98.8/67.9         | 98.4/58.1         | 97.0/75.9 | 98.9/79.9 | 79.6/22.1 | <b>99.7</b> / <b>94.9</b> | 99.6/92.6                 |
| screw      | 95.4/5.9  | <b>99.7</b> / <b>72.9</b> | 99.1/62.0         | 98.7/61.4         | 99.1/51.8         | 98.7/34.2 | 66.5/21.1 | 51.0/5.1  | 92.8/24.8                 | 96.6/43.4                 |
| tile       | 78.6/26.7 | 99.4/ <b>96.4</b>         | 99.2/96.3         | 98.6/92.6         | 97.4/78.2         | 94.9/56.0 | 88.7/63.9 | 91.0/54.4 | 99.7/96.2                 | <b>99.7</b> /94.7         |
| toothbrush | 95.6/20.0 | 97.3/49.2                 | 97.5/38.9         | 98.4/61.7         | <b>99.0</b> /63.1 | 97.6/37.1 | 96.3/52.4 | 74.5/4.8  | 98.8/68.9                 | 98.2/ <b>73.8</b>         |
| transistor | 76.0/25.9 | 92.2/56.0                 | 85.3/36.5         | 98.6/82.9         | <b>99.6</b> /50.3 | 91.8/66.7 | 55.2/4.4  | 79.3/11.2 | 99.4/ <b>92.3</b>         | 98.9/86.8                 |
| wood       | 88.3/24.7 | 97.6/81.6                 | 97.2/77.1         | 97.6/25.6         | 99.3/39.1         | 95.7/54.3 | 93.1/47.9 | 82.9/21.0 | 98.4/77.6                 | <b>99.3</b> / <b>89.2</b> |
| zipper     | 95.1/30.5 | 98.6/73.6                 | 98.1/78.2         | 95.9/53.9         | <b>99.7</b> /52.7 | 98.5/63.1 | 92.4/53.1 | 96.8/42.3 | 99.6/87.5                 | 99.5/ <b>88.1</b>         |
| 平均值        | 89.6/24.9 | 97.7/69.0                 | 96.2/65.5         | 98.3/66.3         | 98.3/57.8         | 97.1/56.6 | 86.4/49.3 | 84.8/25.7 | 98.3/73.8                 | <b>98.4</b> / <b>79.2</b> |

3.3 消融实验

该部分设计了消融实验来验证 AnomalyAlign 各个组件的有效性。该实验为每个缺陷类生成 1 000 个缺陷样本用于实验。没有使用强语义对齐文本提示生成器和缺陷对齐损失的实验 a 中,对于文本提示的获取采用了

DreamBooth<sup>[21]</sup>的方法,即给每个真实的缺陷样本分配一个罕见单词,并将这个罕见单词和句子主干合并作为文本提示;对于损失函数只使用式(7)。没有使用缺陷对齐损失的实验 b 中,损失函数只使用式(7)。图 7 展示了消融实验的可视化结果,图中线条标示了输入掩膜的边缘。通过

对比实验 a~c 组的生成结果,可以直观观察到各新模块的具体作用。具体而言,a 组是仅使用 ControlNet 生成的缺陷图像,其表观真实性和与输入掩膜的对齐性均表现较差;在 a 组的基础上引入强语义对齐的文本提示生成器后,b 组生成的缺陷图像显著提升了表观的真实性;进一步在 b 组的基础上加入缺陷对齐损失后,c 组生成的图像在异常区域与输入掩膜的对齐性上得到了明显改善。这表明每个新增模块都对生成质量的提升起到了关键作用。

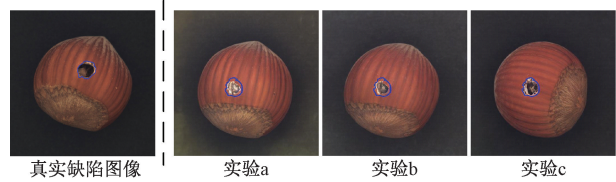


图7 消融实验的可视化比较

Fig.7 Visual comparison of ablation experiments

表5中定量的比较了像素级缺陷定位结果,可见,强语义对齐文本提示生成器和缺陷对齐损失都是 AnomalyAlign 不可或缺的一部分,它们共同组成了功能强大的新模型。

表5 消融实验的定量结果

Table 5 Quantitative results of ablation experiments

| 实验 | 方法               |            | AUC  | AP   | F1-max |
|----|------------------|------------|------|------|--------|
|    | 强语义对齐<br>文本提示生成器 | 缺陷<br>对齐损失 |      |      |        |
| a  | ×                | ×          | 95.0 | 69.2 | 66.9   |
| b  | ✓                | ×          | 97.8 | 75.4 | 72.3   |
| c  | ✓                | ✓          | 98.4 | 79.2 | 75.2   |

4 结 论

针对现有缺陷生成方法生成图像真实性和对齐性差的问题,提出了一种新颖的可控扩散模型 AnomalyAlign 来生成与掩膜高度对齐的逼真工业缺陷图像。首先,该方法以 ControlNet 为基础,继承了大模型丰富的先验知识。同时,提出了一种强语义对齐文本提示生成器来获得在嵌入空间中最对齐的图像-文本对,从而充分利用微调对象的先验信息,加快了模型的收敛,提高了生成质量,在像素级 AUROC、AP 和 F1-max 上分别提高了 2.8%、6.2% 和 5.4%。此外,缺陷对齐损失显式地使生成的缺陷区域与输入的缺陷区域更加对齐,提高了对齐性,在像素级 AUROC、AP 和 F1-max 上分别提高了 0.6%、3.8% 和 2.9%。在 MVTec 数据集上的大量实验证明了 AnomalyAlign 的先进性,特别对于像素级 AP 和 F1-max 指标,比 AnomalyDiffusion 分别提高了 5.4% 和 5.2%。未来 AnomalyAlign 将致力于生成更高分辨率的图像并探

索使用更少的计算资源的方法。

参考文献

[1] 王震洲,杨榕,宿景芳.改进YOLOv8的烟盒缺陷检测算法[J].电子测量技术,2024,47(13):110-119.  
WANG ZH ZH, YANG R, SU J F. Improved YOLOv8 algorithm for cigarette box defect detection [J]. Electronic Measurement Technology, 2024, 47(13): 110-119.

[2] 李金涛,周兴林,尹雨飞,等.基于TAS-YOLO的道路表面缺陷检测[J].电子测量技术,2024,47(13):148-156.  
LI J T, ZHOU X L, YIN Y F, et al. Road surface defect detection based on TAS-YOLO[J]. Electronic Measurement Technology, 2024, 47(13): 148-156.

[3] 王德伟,刘小芳.基于改进YOLOv8n的带钢表面缺陷检测方法[J].国外电子测量技术,2024,43(7):158-169.  
WANG D W, LIU X F. Surface defect detection method for steel strip based on improved YOLOv8n[J]. Foreign Electronic Measurement Technology, 2024, 43(7): 158-169.

[4] 孙铁强,魏光辉,宋超,等.基于YOLO的多模态钢轨表面缺陷检测方法[J].电子测量技术,2024,47(21):72-81.  
SUN T Q, WEI G W, SONG CH, et al. Multi-modal rail surface defect detection method based on YOLO[J]. Electronic Measurement Technology, 2024, 47(21): 72-81.

[5] 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.

[6] ZHOU Q, PANG G, TIAN Y, et al. Anomalyclip: Object-agnostic prompt learning for zero-shot anomaly detection[J]. ArXiv preprint arXiv:2310.18961, 2023.

[7] PANG G, DING C, SHEN C, et al. Explainable deep few-shot anomaly detection with deviation networks[J]. ArXiv preprint arXiv:2108.00462, 2021.

[8] DING C, PANG G, SHEN C. Catching both gray and black swans: Open-set supervised anomaly detection[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022: 7388-7398.

[9] LIN D, CAO Y, ZHU W, et al. Few-shot defect segmentation leveraging abundant defect-free training samples through normal background regularization and crop-and-paste operation[C]. 2021 IEEE International Conference on Multimedia and Expo(ICME), 2021: 1-



- 6.
- [10] GOODFELLOW I, POUGET-ABADIE J, MIRZA M, et al. Generative adversarial networks [J]. Communications of the ACM, 2020, 63 (11): 139-144.
- [11] NIU S, LI B, WANG X, et al. Defect image sample generation with GAN for improving defect recognition[J]. IEEE Transactions on Automation Science and Engineering, 2020, 17(3): 1611-1622.
- [12] ZHANG G, CUI K, HUNG T Y, et al. Defect-GAN: High-fidelity defect synthesis for automated defect inspection[C]. IEEE/CVF Winter Conference on Applications of Computer Vision, 2021: 2524-2534.
- [13] DUAN Y, HONG Y, NIU L, et al. Few-shot defect image generation via defect-aware feature manipulation[J]. AAAI Conference on Artificial Intelligence, 2023, 37 (1): 571-578.
- [14] NICHOL A Q, DHARIWAL P. Improved denoising diffusion probabilistic models[C]. 38th International Conference on Machine Learning. PMLR, 2021: 8162-8171.
- [15] ROMBACH R, BLATTMANN A, LORENZ D, et al. High-resolution image synthesis with latent diffusion models [C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022: 10684-10695.
- [16] RAMESH A, DHARIWAL P, NICHOL A, et al. Hierarchical text-conditional image generation with clip latents [J]. ArXiv preprint arXiv: 2204.06125, 2022.
- [17] SAHARIA C, CHAN W, SAXENA S, et al. Photorealistic text-to-image diffusion models with deep language understanding [J]. Advances in Neural Information Processing Systems, 2022, 35: 36479-36494.
- [18] HU T, ZHANG J, YI R, et al. AnomalyDiffusion: Few-shot anomaly image generation with diffusion model[J]. AAAI Conference on Artificial Intelligence, 2024, 38(8): 8526-8534.
- [19] GAL R, ALALUF Y, ATZMON Y, et al. An image is worth one word: Personalizing text-to-image generation using textual inversion[J]. ArXiv preprint arXiv:2208.01618, 2022.
- [20] RUIZ N, LI Y, JAMPANI V, et al. DreamBooth: Fine tuning text-to-image diffusion models for subject-driven generation [C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023: 22500-22510.
- [21] CAO Y, XU X, SUN C, et al. Segment any anomaly without training via hybrid prompt regularization[J]. ArXiv preprint arXiv:2305.10724, 2023.
- [22] ZHANG L, RAO A, AGRAWALA M. Adding conditional control to text-to-image diffusion models[C]. IEEE/CVF International Conference on Computer Vision, 2023: 3836-3847.
- [23] FELLBAUM C, BROWN K. Wordnet and Wordnets [M]. Elsevier: Encyclopedia of Language and Linguistics, 2005: 2-665.
- [24] RADFORD A, KIM J W, HALLACY C, et al. Learning transferable visual models from natural language supervision [C]. 38th International Conference on Machine Learning. PMLR, 2021: 8748-8763.
- [25] 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.
- [26] LEE S, LEE S, SONG B C. CFA: Coupled-hypersphere-based feature adaptation for target-oriented anomaly localization [J]. IEEE Access, 2022, 10: 78446-78454.
- [27] DENG H, LI X. Anomaly detection via reverse distillation from one-class embedding[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022: 9737-9746.
- [28] 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.
- [29] 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.
- [30] 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.

## 作者简介

陈广庆(通信作者), 副教授, 主要研究方向为智能制造、工业异常检测。

E-mail: chgq75@sdust.edu.cn

陈雅惠, 硕士研究生, 主要研究方向为机器视觉、缺陷检测。

E-mail: 2628815144@qq.com