

李晋芳,肖立宝,何明桐,等.基于生成式样本合成的工件缺陷样本数据增强[J].广东工业大学学报,2025,42(3):27-35. doi:10.12052/gdutxb.240056.
Li Jinfang, Xiao Libao, He Mingtong, et al. Method for data augmentation of workpiece defect samples based on generative sample synthesis[J]. Journal of Guangdong University of Technology, 2025, 42(3): 27-35. doi: 10.12052/gdutxb.240056.

基于生成式样本合成的工件缺陷样本数据增强

李晋芳,肖立宝,何明桐,莫建清

(广东工业大学 机电工程学院,广东 广州 510006)

摘要:针对深度学习模型在工业缺陷视觉检测领域中因样本稀缺而难以较好训练的问题,本文提出一种融合生成对抗网络(Generative Adversarial Network, GAN)和基于物理的渲染(Physically Based Rendering, PBR)流程的生成式样本合成方法用于数据增强。该方法以ConSinGAN为缺陷特征扩增模型,并通过引入坐标注意力机制(Coordinate Attention, CA)来优化鉴别器,使其能更精确识别图像中的缺陷特征。同时调整损失函数,引入重构损失与多尺度结构相似度损失的加权组合以缓解小样本训练中的梯度消失问题并提高生成质量。采用PBR流程输出扩增样本,首先为待扩增样本的工件构建三维模型,然后利用泊松融合将扩增的缺陷特征与原始模型贴图融合,最后在虚拟生产环境中通过虚拟相机渲染输出工件缺陷样本。在公共数据集下的实验结果表明该方法可以对给定的工件缺陷小样本进行有效的数据增强。

关键词:数据增强;生成对抗网络;图像生成;样本合成;工件缺陷

中图分类号:TP3-05

文献标志码:A

文章编号:1007-7162(2025)03-0027-09

Method for Data Augmentation of Workpiece Defect Samples Based on Generative Sample Synthesis

Li Jinfang, Xiao Libao, He Mingtong, Mo Jianqing

(School of Electromechanical Engineering, Guangdong University of Technology, Guangzhou 510006, China)

Abstract: To address the problem of severe lack of defect data in workpieces to train the deep-learning-based defect visual detection systems, this paper introduces a generative sample synthesis method that integrates generative adversarial networks (GAN) with a physical-based rendering (PBR) pipeline for data augmentation. The method employs ConSinGAN as the defect feature generation model and enhances the discriminator by incorporating a coordinate attention (CA) mechanism, enabling more precise identification of defect features in images. Additionally, the loss function is adjusted by introducing a weighted combination of reconstruction loss and multi-scale structural similarity loss to alleviate the gradient vanishing in small sample training and improve the quality of generated samples. The PBR pipeline is used to output the augmented samples, which first constructs a 3D model for the workpiece to be augmented, and then use poisson blending to merge the generated defect features with the original model texture. Finally, defect samples of the workpiece are rendered in a simulated production environment using a virtual camera. Experimental results on public datasets demonstrate the effectiveness of the proposed method in augmenting small samples of workpiece defects.

Key words: data augmentation; generative adversarial network; image generation; sample synthesis; workpiece defect

工件表面缺陷如裂纹、划痕等,不仅损害其外观,还可能影响其结构完整性和性能,甚至可能导致

使用安全问题。因此,工件表面缺陷检测是质量控制的关键一环,对提高工件质量和安全性起着至关重要

收稿日期:2024-04-19 录用日期:2024-07-17 网络首发日期:2025-01-08

基金项目:国家重点研发项目(2018YFB1004902);广州市科技计划项目(2023A03J0724);广州市科技计划重点研发项目(202206010130)

作者简介:李晋芳(1975-),女,副教授,博士,主要研究方向为虚拟现实,E-mail:lijinfang@gdut.edu.cn

通信作者:莫建清(1979-),男,讲师,博士,主要研究方向为虚拟现实,E-mail:momolon@gdut.edu.cn

要的作用^[1]。常规人工检测泛用性强,但过度依赖于人的主观判断,对操作人员的经验有较高要求^[2],不适用于大规模应用。随着工厂智能化的不断发展,采用视觉检测系统检测工件表面缺陷是大势所趋。

高效精确的视觉检测系统需要足够的训练样本的支持。但是,随着工艺进步,实际生产中缺陷产生的机率变小,缺陷样本难以获取。此外,缺陷类别具有随机性,导致缺陷样本类别多,而单类缺陷样本偏少。实际生产中缺陷样本的产生特点与视觉检测系统的训练需求是相互冲突的,这阻碍了视觉检测系统的推广应用^[3]。

缺陷小样本数据增强是指在面对数据量有限或特定种类的样本稀缺的情况下,采用各种技术手段来增加样本的多样性和数量,以提高深度学习模型的性能和泛化能力。传统的数据增强方式有基于几何的与非几何的两种方式^[4]。几何数据增强对原始图像数据进行如平移、旋转、缩放、剪切等几何变换操作;非几何数据增强对图像进行随机裁剪、噪声添加、对比度调整等除几何数据外的其他数据变化操作。上述方法在增加数据多样性、提升模型鲁棒性方面取得一定效果,但图像变换操作可能导致信息丢失、引入不真实样本等负面影响,甚至引发模型学习到错误的特征而降低性能。

生成对抗网络(Generative Adversarial Network, GAN)是一种隐式生成模型,通过生成器和鉴别器的对抗训练来学习已有数据的高维特征后,输出具有类似特性的新样本^[5]。使用GAN模型对已有的工件缺陷小样本进行数据增强,可以改善样本分布不均匀、数量不足的问题,为视觉检测系统提供更真实、多样的训练样本^[6]。

近年来,GAN在数据增强领域的应用研究逐渐增加。庄昌乾等^[7]提出Defect-GAN网络,通过模拟缺陷生成和缺陷图像重建的过程,利用正常样本来扩增缺陷样本。罗月童等^[8]从提高模型对缺陷区域的注意力的角度改进pix2pixGAN^[9],提出一种工件缺陷数据增强方法。文献^[10-11]分别研究了GAN模型在织物缺陷图像扩增、新冠肺炎CT图像扩增应用领域的有效性。

除此之外,传统的GAN模型如BigGAN^[12]、PGGAN^[13]虽能扩增高质量样本,但它们本身就需要一定规模的样本进行训练,这与工件缺陷的小样本特性是相矛盾的。如何使GAN模型在小样本数据上稳定训练并生成高质量图像,是将GAN广泛应用于数据增强领域的关键问题。

在小样本扩增领域中,SinGAN^[14]是一种新兴的GAN变体,其在训练过程中,针对图像的不同尺度训练对应的生成器网络,以充分学习图像在多尺度上的特征,减少对训练样本数量需求,并且更不容易出现模式崩溃问题,生成的图像能够更好地保留原始图像的特征。ConSinGAN^[15]在两方面对SinGAN模型进行了改进:(1)采取并行训练和多学习率策略以加速训练过程;(2)在不同阶段的生成器输出传递过程中,将SinGAN模型的图像传递改为特征传递,减少了重复计算,同时更好地保留了高维特征信息,提高生成图像质量。彭晏飞等^[16]改进SinGAN网络的生成器中的卷积层,有效提高了模型的生成能力。赵晓枫等^[17]验证了联合SinGAN网络和几何数据增强两种方式对检测算法的数据集进行数据增强,能有效提高检测效果。黄琼男等^[18]在ConSinGAN网络中引入注意力机制,有效提高了样本的扩增质量。

综上所述,基于GAN模型的小样本数据增强方法主要存在以下2个问题:(1)GAN模型的训练仍需一定规模的样本数量,难以应用于工件缺陷类别多、单类样本数量少的场合;(2)现有方法仅是直接在已有缺陷样本上对缺陷的纹理特征进行扩增,没有考虑到缺陷的空间特征和尺度特征等信息。

为解决上述问题,本文以ConSinGAN网络作为缺陷特征扩增模型,结合PBR(Physically Based Rendering)流程与泊松融合算法,提出一种针对工件缺陷小样本的新型数据增强方法。主要创新点包括:(1)提出使用泊松融合算法将ConSinGAN网络扩增的缺陷特征与工件模型进行融合以构建带缺陷的工件模型;(2)引入PBR渲染流程对融合产生的带缺陷工件模型进行渲染成像,进而对原始小样本进行数据增强。

1 本文方法

本文提出一种基于生成式样本合成的工件缺陷样本数据增强方法,由缺陷特征扩增和缺陷样本扩增两阶段组成,其流程如图1所示。缺陷特征从原始数据样本中提取,然后通过改进的ConSinGAN生成网络进行扩增。为了提升模型的鲁棒性和图像生成质量,在ConSinGAN中引入注意力机制,并改进损失函数,以提高模型的训练稳定性。缺陷样本扩增的关键步骤是缺陷特征融合。本文提出利用泊松融合方法将缺陷特征与工件模型相融合,得到带缺陷的工件模型。然后,在虚拟生产环境中对工件模型进行渲染,合成丰富的高质量工件缺陷样本。

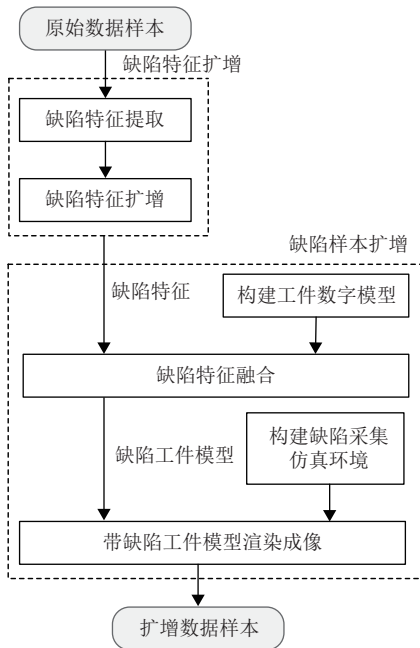


图1 缺陷样本数据增强流程

Fig.1 The data augmentation process of defect samples

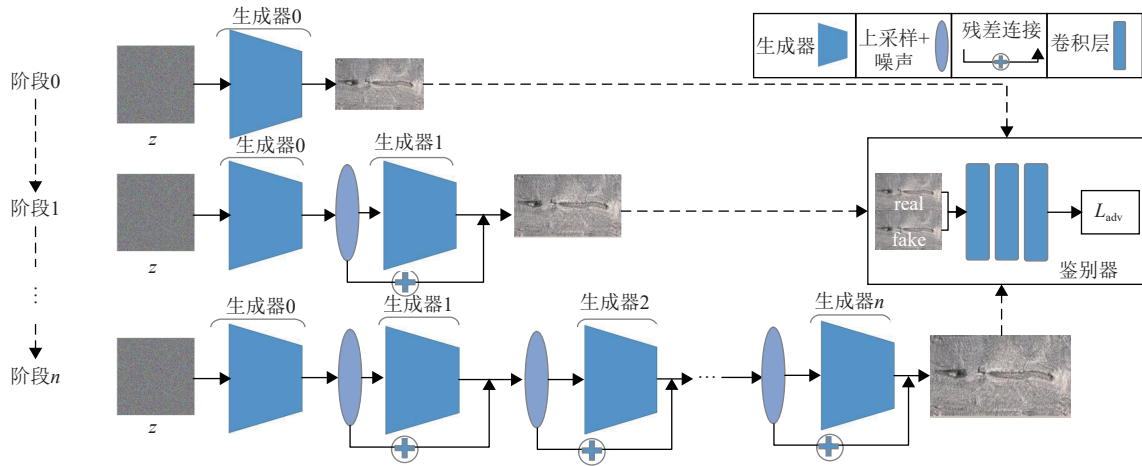


图2 ConSinGAN 网络结构

Fig.2 ConSinGAN structure

注意力机制与GAN模型相结合,能帮助GAN模型更好地学习图像的高维特征,提高模型生成的细节质量^[19]。常见的注意力机制模块有SE(Squeeze and Excitation)^[20]、CBAM(Convolutional Block Attention Module)^[21]和CA (Coordinate Attention)^[22]等。其中,CA模块的结构如图3所示。

与SE、CBAM机制相比,CA具有更小计算量,还能在考虑局部结构信息的情况下,捕获空间位置和通道之间的关系来实现更为精细的全局上下文建模。这些特点有助于强化网络在图像扩增任务中的细节表达能力。因此,本文改进ConSinGAN模型,在

1.1 缺陷特征扩增

1.1.1 注意力机制引入

ConSinGAN模型的架构如图2所示。模型训练由多阶段组成。首先训练一个低分辨率生成器(见图2中的阶段0)。当阶段*i*的训练收敛后,增加一个新的生成器*i+1*。新生成器以前一阶段生成器的输出特征为基础,引入额外的噪声作为输入,继续训练。重复*n*次,直至生成目标分辨率的图像。

ConSinGAN模型通过多阶段、多分辨率的并行训练方法,结合特征传播,自适应学习率调整等策略,可以从单张图像生成高质量且多样的图像变体,适用于工件缺陷特征图像的生成,以扩充工件缺陷样本。然而,在输入的缺陷工件图像中,缺陷区域面积通常占比较小,ConSinGAN模型中的鉴别器可能无法有效捕捉缺陷特征,从而导致生成器在训练过程中倾向于优先学习简单且覆盖面广的背景特征。因此,如何提升鉴别器识别缺陷特征的能力,是提高缺陷特征扩增质量的关键。

其鉴别器中引入CA模块,以提高鉴别器对缺陷特征的鉴别能力。CA模块能够有效捕捉缺陷特征在输入图像中的空间关系和结构信息,包括位置和尺度信息等,从而使鉴别器能够更精确地识别和分析图像中的缺陷特征,提高判别能力,进而更有效地监督生成器训练,提升模型的鲁棒性和图像生成质量。图4展示了引入CA模块后的鉴别器架构。

1.1.2 损失函数改进

ConSinGAN的初始损失函数包含对抗损失和重建损失,定义为

$$\min_{G_n} \max_{D_n} L_{adv}(G_n, D_n) + \alpha L_{rec}(G_n)$$

式中, G 为生成器; D 为鉴别器; n 表示训练阶段; L_{adv} 为对抗损失, 采用的是 WGAN-GP^[23] 对抗损失函数; L_{rec} 表示重构损失; α 为缩放系数。

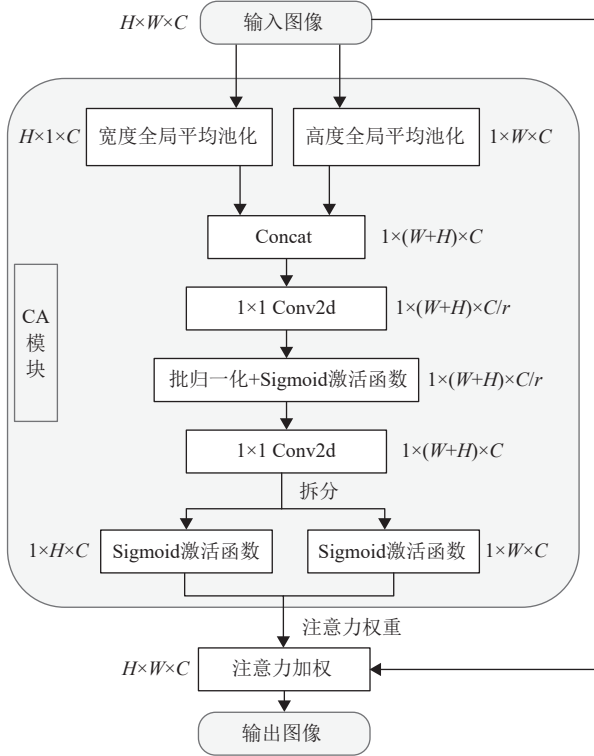


图3 CA注意力机制工作方式

Fig.3 The basic flowchart of CA attention mechanism

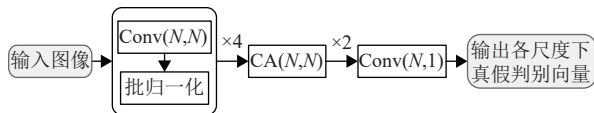


图4 CA鉴别器架构

Fig.4 CA discriminator architecture

针对工件缺陷小样本数量缺乏和分布不均匀的问题, 本文通过改进重构损失, 并引入多尺度结构相似度(Multi-Scale Structural Similarity, MS-SSIM)损失来约束模型训练, 提高GAN模型的训练稳定性和生成的图像质量。改进后的损失函数定义为

$$L_{total} = L_{adv} + \lambda_1 L_{rec} + \lambda_2 L_{MS-SSIM} \quad (1)$$

式中, L_{total} 表示总损失函数; L_{rec} 表示重构损失; $L_{MS-SSIM}$ 表示多尺度结构相似度损失; λ_1 和 λ_2 是平衡2个损失函数的超参数。

重构损失用于约束生成图像与真实图像的相似度。原始模型采用均方损失(Mean Squared Error, MSE)作为重构损失函数, 但其对异常值比较敏感, 可能会导致模型在训练过程中出现不稳定的情况。因此, 本文将重构损失函数从MSE损失更改为

Smooth L1损失。

Smooth L1损失是一种基于L1范数的平滑函数, 相比于MSE损失, 对异常值响应更加平滑。损失定义如式(2)所示。

$$L_s(y_i - \hat{y}_i) = \frac{1}{c} \sum_{i=1}^c z_i$$

$$z_i = \begin{cases} \frac{1}{2} (y_i - \hat{y}_i)^2, & |y_i - \hat{y}_i| < 1 \\ |y_i - \hat{y}_i| - \frac{1}{2}, & |y_i - \hat{y}_i| \geq 1 \end{cases} \quad (2)$$

式中, y_i 表示真实的图像; \hat{y}_i 表示生成的图像; c 表示图像中像素的数量; z_i 是一个平滑函数, 用于计算每个像素点的损失值。当 $|y_i - \hat{y}_i| < 1$ 时, 使用MSE误差作为损失函数。当 $|y_i - \hat{y}_i| \geq 1$ 时, 使用绝对值误差作为损失函数, 此函数在平衡模型训练速度与稳定性方面具备一定优势^[24]。

在缺陷工件图像中, 背景通常相对单一、特征较少或很弱, 缺陷与背景之间对比度较低, 这些因素可能导致模型在训练时发生模式崩溃。因此, 在损失函数中引入MS-SSIM损失, 以提高模型训练的稳定性^[25]。 $L_{MS-SSIM}$ 损失的定义见式(3)。

$$L_{MS-SSIM} = \prod_{m=1}^M [U(X, Y)^{\beta_m} C(X, Y)^{\gamma_m}]$$

$$U(X, Y) = \frac{2\mu_X\mu_Y + c_1}{\mu_X^2 + \mu_Y^2 + c_1}$$

$$C(X, Y) = \frac{2\sigma_{XY} + c_2}{\sigma_X^2 + \sigma_Y^2 + c_2} \quad (3)$$

式中, X, Y 分别表示生成器产生的图片和同尺寸缩小的原图; M 表示不同尺度的数量; μ_X 和 μ_Y 分别表示 X, Y 的平均值; σ_X 和 σ_Y 为 X, Y 的方差; σ_{XY} 为协方差; $c_1 = (k_1 l)^2$, $c_2 = (k_2 l)^2$ 是用于维持稳定的常数, 其中 l 表示图像的动态范围, k_1 和 k_2 分别设置为0.01和0.03; β_m 和 γ_m 表示两项之间的相对重要性。

1.2 缺陷样本扩增

1.2.1 缺陷提取

工业缺陷样本通常具有背景单一、缺陷面积相对较小、特征不够明显的特点。此外, 使用生成网络进行数据扩增时, 随着训练图像的像素数量增加, 训练时间会延长, 硬件要求也随之提升。针对这些问题, 本方法在扩增缺陷特征时, 仅对缺陷区域进行处理, 通过裁剪无关背景, 从原始图像中提取缺陷部分。此策略不仅可以加快训练速度, 还能提高生成网络对缺陷特征的学习效果, 从而增强网络的稳定性。

1.2.2 工件三维模型构建

为了模拟真实缺陷样本采集过程, 首先需要得

到工件的数字模型。

PBR渲染流程是常用的输出逼真渲染图像的工作流程。PBR渲染流程主要包括几何建模、材质定义、光源设置、贴图计算、合成绘制5个工作环节^[26]。具体而言,通过材质制作软件,如SP(Substance Painter)等,事先将光照与贴图材质的交互后效果计算出来,并将其效果烘焙为贴图以赋予模型。在渲染引擎调用模型渲染成像时,通过采样贴图的方式代替实时计算光照效果,在渲染质量与渲染速度两方面取得平衡。

几何建模是PBR流程中的关键步骤,常见的建模方法包括多边形建模和扫描建模等。针对本文中工件数字模型的构建,由于工件需要用于生产加工,通常已具备相应的工业三维模型。目前,诸如3DMax和Blender等常用的三维建模软件,都提供了将工业三维模型转换为适用于渲染的网格模型的功能。为了确保工件模型与真实工件一致,本文采用了这种方法。然而,需要注意的是,使用该方法生成的模型在法线方向、纹理映射等方面可能存在偏差,因此需要进行手动调整,以确保后续获得准确的纹理贴图。

1.2.3 缺陷融合

缺陷特征与工件模型纹理贴图融合过程包含3个关键步骤:指定融合位置、计算融合中心点像素坐标和使用泊松融合算法融合图像。

在指定融合位置步骤中,首先在工件模型上选择目标融合位置点 P ,随后通过遍历模型的面片并计算该点与每个面片的距离,选取距离最小的面片作为包含该点的面片 t_0 。

下一步需要计算点 P 在模型纹理贴图上的像素坐标。设面片 t_0 的3个顶点的坐标为

$$P_i(x_i, y_i, z_i), i \in (1, 2, 3)$$

通过式(4)计算出点 P 在面片 t_0 中的重心坐标 (w_1, w_2, w_3) 。

$$\begin{aligned} w_1 &= \frac{S(P, P_2, P_3)}{S(P_1, P_2, P_3)}, \\ w_2 &= \frac{S(P, P_1, P_3)}{S(P_1, P_2, P_3)}, \\ w_3 &= \frac{S(P, P_1, P_2)}{S(P_1, P_2, P_3)} \end{aligned} \quad (4)$$

式中, $S(A, B, C)$ 表示以点 A, B, C 为顶点的三角面元的面积,通过式(5)计算。

$$S(A, B, C) = \frac{1}{2} |(B-A) \times (C-A)| \quad (5)$$

随后将点 P 重心坐标转化为模型纹理贴图UV坐

标系中的UV坐标。设面片 t_0 的3个顶点的UV坐标为

$$P_{UV,i}(u_i, v_i), i \in (1, 2, 3)$$

通过式(6)计算点 P 的UV坐标 (u, v) :

$$\begin{aligned} u &= w_1 u_1 + w_2 u_2 + w_3 u_3, \\ v &= w_1 v_1 + w_2 v_2 + w_3 v_3 \end{aligned} \quad (6)$$

最后通过式(7)获得点 P 在模型纹理贴图上的像素坐标 $(x_{\text{pixel}}, y_{\text{pixel}})$,即缺陷特征图像融合的中心点位置。设模型纹理贴图的尺寸为 (h, w) , h 表示贴图高度, w 表示贴图宽度。

$$\begin{aligned} x_{\text{pixel}} &= uw, \\ y_{\text{pixel}} &= vh \end{aligned} \quad (7)$$

在将缺陷特征图像与模型纹理贴图在指定位置融合的过程中,直接将缺陷特征图像替换在模型纹理贴图会使叠加区域的像素信息丢失,且由于缺陷图像的背景与融合位置背景可能存在显著差异,这将导致明显的融合边界,从而使得融合后的带缺陷工件模型与真实缺陷工件之间产生较大差异。为了解决这些问题,本文在缺陷融合过程中采用了泊松融合算法,以实现更加无缝自然的图像融合效果。

泊松融合算法被广泛应用于图像拼接、图像融合等图像处理工作。即使源图像和目标图像的纹理、光照或颜色有所差异,也可以在保留源图像梯度信息的前提下实现无缝融合,消除融合痕迹^[27]。

假设 f^* 是在非待融合区域的已知标量函数, f 是待融合区域 Ω 内部定义的未知标量函数,在 Ω 上的 f 的最简单的插值为

$$\begin{aligned} \min_f \iint_{\Omega} M^2, f|_{\partial\Omega} &= f^*|_{\partial\Omega}, \\ M &= (\nabla f_x - \mathbf{v}_x)^2 + (\nabla f_y - \mathbf{v}_y)^2 \end{aligned} \quad (8)$$

式中, $\nabla = \left[\frac{\partial}{\partial x}, \frac{\partial}{\partial y} \right]$ 是梯度算子, \mathbf{v} 是向量场。在Dirichlet边界条件下,式(8)的解与方程(9)的唯一解相同为

$$\Delta f = \text{div}(\mathbf{v}), f|_{\partial\Omega} = f^*|_{\partial\Omega} \quad (9)$$

式中, Δ 是拉普拉斯算子, $\text{div}(\mathbf{v})$ 表示 \mathbf{v} 的散度。分别对图像的3个通道建立方程进行求解,就可以获得所有待融合区域的像素点的RGB值,实现图像无缝融合。

1.2.4 缺陷样本扩增

生成的缺陷样本是通过带缺陷纹理的三维工件模型渲染而来。图像渲染流程包含以下几个关键步

骤:建模、场景布置、光照、渲染成像和后期处理。本方法中建模步骤所需模型即为融合缺陷后的工件模型。正确的场景设置和光照是提高渲染真实度的关键。参考缺陷样本图像,反复调整场景、摄像机视角,设置光源的类型、颜色、位置和强度等,尽可能恢复工件的实际生产和检测环境。在渲染引擎中,如Unity3D等,完成上述场景、摄像机和光源的布置后,应用PBR渲染引擎计算输出缺陷样本图像中每个像

素点的颜色和明暗。

构建好渲染环境后,通过改变融合的缺陷样式、模型的位姿、相机位姿、相机成像参数、光照信息等属性,可以得到不同缺陷特征在不同分辨率、视角、光照条件下产生的缺陷样本,从缺陷类别、单类样本数量、不同视角和分辨率等多个维度对工件缺陷样本进行增强。图5描述了从扩增缺陷特征到扩增工件缺陷样本的整体流程。

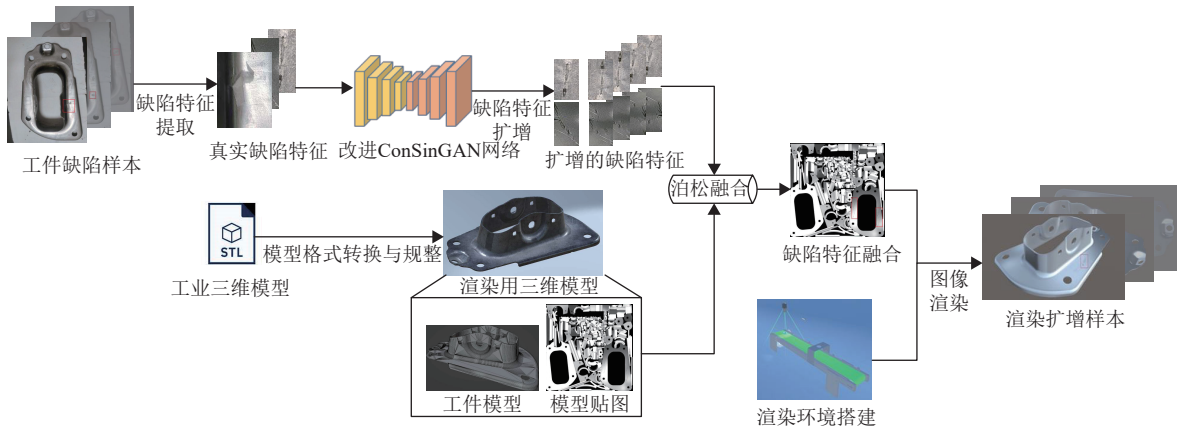


图5 样本渲染扩增流程

Fig.5 Sample rendering amplification process

2 实验结果及讨论

为了验证本文改进后的ConsinGAN模型在缺陷特征扩增方面的效果,采用该网络对自采集的冲压缺陷数据集进行了扩增处理,并进行了对比实验以验证所采取的改进策略的有效性。在公开的瓷砖缺陷数据集上应用了本文所述的小样本数据增强方法,结果证明本文方法所增强的样本与实际采集样本具有高度相似性。

2.1 ConSinGAN模型训练

实验硬件环境采用Windows10操作系统下AMD Ryzen 7 5800H与NVIDIA GeForce GTX 3060 Laptop,软件环境包括Python 3.8、Pytorch 1.8.1和CUDA 11.1。ConSinGAN模型参数与文献[15]使用的相同,主要超参数有:训练尺度和学习率收缩率分别为6和0.1,卷积核的尺寸和通道数为3和64,生成器的训练阶段数为3。

2.2 对比实验

2.2.1 生成模型对比

从广东佛山一汽车冲压件工厂采集其冲压工件缺陷样本构建缺陷数据集,部分样本如图6所示。

使用SIFID(Single Image Fréchet Inception

Distance) 和PSNR(Peak Signal to Noise Ratio) 指标对各个模型生成效果进行评估和比较。



图6 部分数据样本展示

Fig.6 Some data samples

SIFID是FID(Fr chet Inception Distance) 针对单图像扩增情况下的改进,用于计算真实图像和扩增图像多尺度下的平均FID指标。SIFID指标越低,说明扩增图像与真实图像的特征分布越接近,图像质量越高。

PSNR指标反映真实和扩增图像在像素级别的差异,PSNR值越大表示生成图像的失真程度越小,

生成质量越高。

分别使用SinGAN、ConSinGAN和本文所述改进网络进行缺陷扩增实验,实验效果如图7所示。

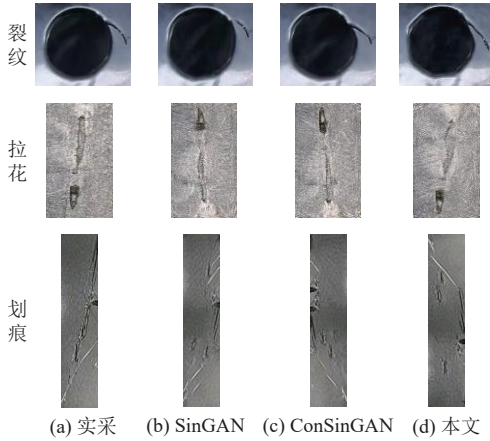


图7 缺陷扩增效果
Fig.7 The performance of the defect generation

为了进行定量计算,从包含裂纹、拉花和划痕3类样本的图像数据集中随机选取了100张样本。评估结果如表1所示。

表1 不同模型指标对比

Table 1 Performance comparison of different models

方法	SIFID			PSNR		
	划痕	拉花	裂纹	划痕	拉花	裂纹
SinGAN	0.762	0.753	0.682	33.51	33.36	33.21
ConSinGAN	0.324	0.337	0.272	34.43	34.23	33.94
本文	0.143	0.143	0.192	35.21	35.19	34.97

从表1可以看出,无论是在像素差异还是特征分布差异方面,本文的改进模型对3类缺陷的扩增效果都更好。结果表明本文所提改进模型在提高生成图像质量方面取得一定的成效。

2.2.2 注意力机制对比

为对比SE、CBAM、CA这三种注意力机制引入后对ConSinGAN网络训练速度和生成质量的影响,设置对比实验,结果如图8所示。

从图8可以看出,与SE和CBAM相比,应用CA注意力模块的损失函数收敛得更快,有效提高了模型的训练收敛速度。

2.2.3 重构损失对比

分别应用MSE损失、Smooth L1损失和L1损失作为重构损失对模型进行训练,得到收敛曲线如图9所示。

图9的结果表明,以Smooth L1作为重构损失,训练过程中损失下降更加平滑,对梯度变化更敏感,有效提高了模型的训练稳定性。

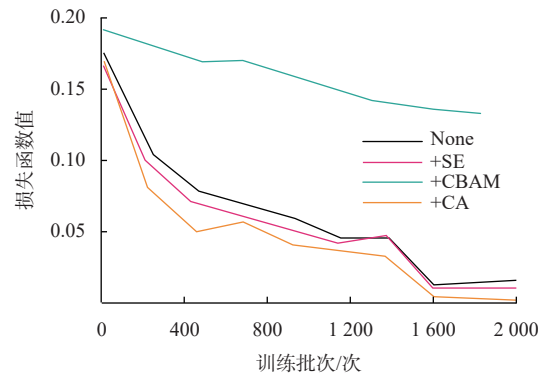


图8 注意力机制对比

Fig.8 Comparison of different attention mechanisms

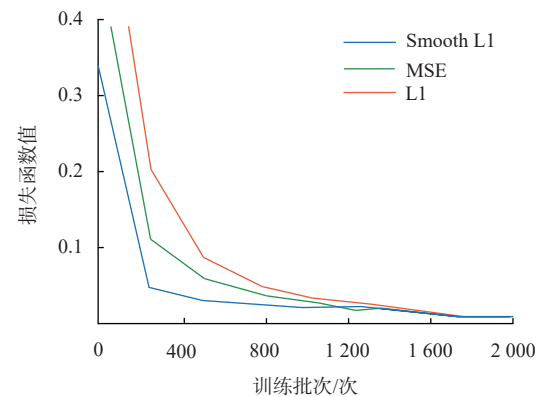


图9 重构损失对比

Fig.9 Comparison of different reconstruction losses

2.3 缺陷样本增强效果实验

瓷砖划痕缺陷检测是一种常见的视觉任务,本节以瓷砖缺陷样本数据增强为例,进行小样本数据增强效果评估实验。

2.3.1 评估指标

评估从颜色、纹理、形状和结构相似度4个方面对比真实图像与扩增图像。

颜色评估通过比较两个图像的颜色直方图,量化对比在颜色分布上的差异程度。验证扩增样本与真实样本的整体相似度。

纹理评估采用基于Gabor滤波器的纹理提取相似度计算方法,以验证扩增样本对真实缺陷特征的纹理复现能力。

形状评估比较扩增样本与真实样本之间的形状相似度,以验证通过数字模型进行渲染成像是否可以保留与真实工件相同的物理特征,渲染扩增缺陷流程是否具备合理性。

结构相似度评估通过亮度、对比度和结构3方面的信息来验证两幅图像之间的结构相似度。

2.3.2 实验结果

采用2021年广东工业制造创新大赛智能-算法赛提供的瓷砖检测数据集,该数据集包含瓷砖划伤,记号笔迹等缺陷共24 000张图像。

为模拟工件缺陷小样本特点,从记号笔迹类缺陷中选取300张样本,其中100张为原始RGB彩色图,用于对比颜色相似度,100张为弱照明灰度图,100张

为强照明灰度图,用于对比在不同照明情况下对扩增质量的影响。通过OpenCV的图像截取功能,根据数据集中给定的缺陷所在像素区域分割出每张样本的缺陷区域,使用上文所述生成模型,对每个样本的缺陷特征扩增3次,扩增缺陷特征与无缺陷样本模型进行融合,在虚拟采集环境下进行缺陷样本渲染生成。结果如图10所示。

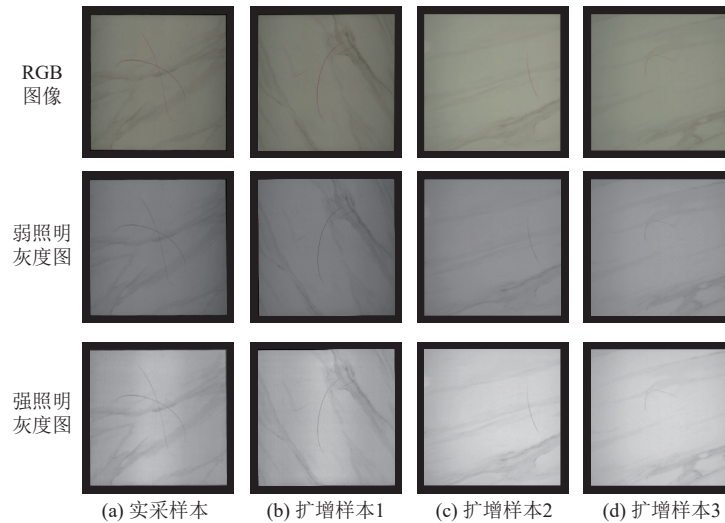


图10 瓷砖缺陷扩增效果

Fig.10 The visual performance of the tile defect amplification

对扩增图像与真实图像计算上述指标差异,统一用相似度百分比表达,实验结果如表2所示。

表2 图像相似度评估结果
Table 2 Image similarity evaluation results %

图像类型	颜色相似度	纹理相似度	形状相似度	结构相似度
RGB彩色图	92.6	97.3	98.5	95.8
强照明灰度图		97.8	98.1	96.2
弱照明灰度图		96.9	98.3	95.1

从表2中可知,由于在充足光照下,图像的纹理会变得更加清晰,所以强照明下纹理相似度最高。由于工件的渲染模型是根据工件的制造模型更改而来,其具备与真实工件相同的物理特征,3种情况下的形状相似度均为较高值。强弱照明时形状相似度相近,说明渲染时的照明对特征的保留影响较小。3种情况下的结构相似度均为较高值,说明扩增样本与真实样本高度相似。

综上所述,根据颜色、纹理、形状和结构相似度的比较,表明本文方法可以在保留样本图像中工件物理特征的情况下,很好地对缺陷特征进行样式泛化和数量扩增,有效地对缺陷样本数据进行多维度的数据增强。

3 结论

本文提出了一种基于生成式样本合成的工件缺陷样本数据增强方法,以缓解工业视觉检测技术在实际生产中由于工件缺陷样本数据匮乏而无法大规模推广应用的问题。本文方法通过引入注意力机制、改进损失函数的方式,提高了生成模型在小样本情况下训练的稳定性。同时利用PBR渲染流程实现工件缺陷样本的样本扩增。实验结果表明,本文方法扩增的样本与真实样本高度相似,可以用于工件缺陷小样本的数据增强。需要注意的是,目前本文方法在缺陷融合环节中需要手动指定缺陷融合位置,下一步将研究如何提升自动化水平。另外,目前本文方法仅用于工件的缺陷样本扩增,下一步考虑所提方法的推广应用。

参考文献:

- [1] 王慧菁, 杨长辉, 吕庆. 基于机器视觉的金属表面缺陷检测方法综述[J]. 微纳电子与智能制造, 2022, 4(4): 71-81. WANG H J, YANG C H, LYU Q. Review of metal surface defect detection methods based on machine vision[J]. Mirco/Nano Electronics and Intelligent Manufacturing, 2022, 4(4): 71-81.
- [2] 金映谷, 张涛, 杨亚宁, 等. 基于深度学习的产品缺陷检

- 测方法综述[J]. 大连民族大学学报, 2020, 22(5): 420-427.
- JIN Y G, ZHANG T, YANG Y N, *et al.* Review of product defect detection methods based on deep learning[J]. Journal of Dalian Minzu University, 2020, 22(5): 420-427.
- [3] 罗东亮, 蔡雨萱, 杨子豪, 等. 工业缺陷检测深度学习方法综述[J]. 中国科学: 信息科学, 2022, 52(6): 1002-1039.
- LUO D L, CAI Y X, YANG Z H, *et al.* Survey on industrial defect detection with deep learning[J]. *Scientia Sinica Informationis*, 2022, 52(6): 1002-1039
- [4] MUMUNI A, MUMUNI F. Data augmentation: a comprehensive survey of modern approaches[J]. *Array*, 2022, 16: 100258.
- [5] GOODFELLOW I, POPUGET-ABADIE J, MIRZA M, *et al.* Generative adversarial networks[J]. *Communications of the ACM*, 2020, 63(11): 139-144.
- [6] 孙书魁, 范菁, 孙中强, 等. 基于深度学习的图像数据增强研究综述[J]. 计算机科学, 2024, 51(1): 150-167.
- SUN S K, FAN J, SUN Z Q, *et al.* Survey of image data augmentation techniques based on deep learning[J]. *Computer Science*, 2024, 51(1): 150-167.
- [7] 庄昌乾, 李璟文. 基于YOLOv5和生成对抗网络的塑料标签缺陷检测[J]. 计算机测量与控制, 2023, 31(7): 91-98.
- ZHUANG C Q, LI J W. Industrial defect detection of plastic labels based on YOLOv5 and generative adversarial networks[J]. *Computer Measurement & Control*, 2023, 31(7): 91-98.
- [8] 罗月童, 段昶, 江佩峰, 等. 一种基于pix2pix改进的工业缺陷数据增强方法[J]. 计算机工程与科学, 2022, 44(12): 2206-2212.
- LUO Y T, DUAN C, JIANG P F, *et al.* An improved industrial defect data augmentation method based on pix2pix[J]. *Computer Engineering & Science*, 2022, 44(12): 2206-2212.
- [9] PHILLIP I, ZHU J Y, ZHOU T H, *et al.* Image-to-image translation with conditional adversarial networks[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu: IEEE, 2017: 5967-5976.
- [10] LIU J H, WANG C Y, SU H, *et al.* Multistage GAN for fabric defect detection[J]. *IEEE Transactions on Image Processing*, 2019, 29: 3388-3400.
- [11] 闫艺丹, 孙君顶, 姚冲, 等. 基于生成对抗网络的CT图像数据扩增[J]. 计算机系统应用, 2022, 31(12): 78-86.
- YAN Y D, SUN J D, YAO C, *et al.* CT image data amplification based on generative adversarial network[J]. *Computer Systems & Applications*, 2022, 31(12): 78-86.
- [12] BROCK A, DONAHUE J, SIMONYAN K. Large scale GAN training for high fidelity natural image synthesis [EB/OL]. arXiv: 1809.11096(2019-02-25) [2024-03-16]. <https://doi.org/10.48550/arXiv.1809.11096>.
- [13] KARRAS T, AILA T, LAINE S, *et al.* Progressive growing of GANs for improved quality, stability, and variation [EB/OL]. arXiv: 1710.10196(2018-02-26) [2024-03-16]. <https://doi.org/10.48550/arXiv.1710.10196>.
- [14] TAMAR R S, TALI D, TOMER M. SinGAN: learning a generative model from a single natural image[C]//2019 IEEE/CVF International Conference on Computer Vision (ICCV). Seoul: IEEE, 2019: 4569-4579.
- [15] HINZ T, FISHER M, WANG O, *et al.* Improved techniques for training single-image gans[C]//2021 IEEE Winter Conference on Applications of Computer Vision (WACV). Waikoloa: IEEE, 2021: 1299-1308.
- [16] 彭晏飞, 邓佳楠, 王刚. 基于改进SinGAN的遥感图像数据增强方法[J]. 液晶与显示, 2023, 38(3): 387-396.
- PENG Y F, DENG J N, WANG G. Remote sensing image data enhancement based on improved SinGAN[J]. *Chinese Journal of Liquid Crystals and Displays*, 2023, 38(3): 387-396.
- [17] 赵晓枫, 夏玉婷, 徐叶斌, 等. 地面红外目标数据联合增强方法[J]. 激光与红外, 2023, 53(7): 1117-1124.
- ZHAO X F, XIA Y T, XU Y B, *et al.* Joint data augmentation method for ground infrared target[J]. *Laser & Infrared*, 2023, 53(7): 1117-1124.
- [18] 黄琼男, 朱卫纲, 刘渊, 等. 基于多尺度GAN网络的SAR舰船目标扩充[J]. 兵工自动化, 2022, 41(7): 47-52.
- HUANG Q N, ZHU W G, LIU Y, *et al.* SAR ship target expansion based on multiscale GAN network[J]. *Ordnance Industry Automation*, 2022, 41(7): 47-52.
- [19] ZHANG H, GOODFELLOW I, METAXAS D, *et al.* Self-attention generative adversarial networks[J]. *Proceedings of Machine Learning Research*, 2019, 97: 7354-7363.
- [20] HU J, SHEN L, SUN G, *et al.* Squeeze-and-excitation networks[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018: 7132-7141.
- [21] WOO S, PARK J, LEE J Y, *et al.* CBAM: convolutional block attention module[C]//Computer Vision - ECCV 2018: 15th European Conference. Munich: Springer, 2018: 3-19.
- [22] HOU Q B, ZHOU D Q, FENG J S. Coordinate attention for efficient mobile network design[C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Nashville: IEEE, 2021: 13708-13717.
- [23] ARJOVSKY M, CHINTALA S, BOTTOU L. Wasserstein generative adversarial networks[J]. *Proceedings of Machine Learning Research*, 2017, 70: 214-223.
- [24] GIRSHICK R. Fast R-CNN[C]//2015 IEEE International Conference on Computer Vision (ICCV). Santiago: IEEE, 2015: 1440-1448.
- [25] WANG Z, SIMONCELLI E P, BOVIK A C. Multiscale structural similarity for image quality assessment[C]//The Thirty-Seventh Asilomar Conference on Signals, Systems & Computers. Pacific Grove: IEEE, 2003: 1398-1402.
- [26] 姜滢稳, 江自昊. PBR方法在虚拟现实技术中的研究与应用[J]. 黑河学院学报, 2022, 13(6): 178-180.
- JIANG Y W, JIANG Z H. On the application of PBR method in virtual reality technology[J]. *Journal of Heihe University*, 2022, 13(6): 178-180.
- [27] PEREZ P, GANGNET M, BLAKE A. Poisson image editing[J]. *ACM Transactions on Graphics*, 2003, 22(3): 313-318.
- (责任编辑: 王威娜 英文审核: 费伦科)