



WMG-GAN: 基于权重图引导的布匹瑕疵 图像生成算法

王迎铭¹, 陈柯烽¹, 潘海鹏^{1,2}, 任 佳^{1,2}

(1. 浙江理工大学信息科学与工程学院, 杭州 310018; 2. 浙江理工大学常山研究院有限公司, 浙江衢州 324299)

摘要: 针对现有方法在重建背景细节和生成图像质量方面存在的不足, 以 CycleGAN 为基础框架, 提出了一种基于权重图引导的布匹瑕疵图像生成算法 WMG-GAN (Weight-map-guided generative adversarial network)。该算法首先通过生成器产生前景权重图和特征权重图, 实现针对前景部分内容的选择性修改, 并完整保留背景细节和结构; 其次, 在判别器中加入 ConvNeXt V2 模块, 增加网络的特征提取能力, 为生成器提供更精确的梯度反馈; 最后, 引入感知学习图像块相似性 (Learned perceptual image patch similarity, LPIPS) 指标, 构建循环一致性损失函数, 以优化生成图像的视觉质量与真实感。在真实布匹瑕疵数据集上的对照实验和消融实验表明, 该算法生成的布匹瑕疵图像相较于传统 CycleGAN, 不仅具有较低的弗雷歇初始距离 (Fréchet inception distance, FID) 和 LPIPS 值, 而且能获得较高的结构相似性指数 (Structural similarity index measure, SSIM) 和峰值信噪比 (Peak signal-to-noise ratio, PSNR)。WMG-GAN 算法可显著提升图像生成质量, 由其生成的图像满足瑕疵检测算法的高精度要求。

关键词: 布匹瑕疵图像生成; 生成对抗网络; CycleGAN; 权重图; ConvNeXt V2

中图分类号: TP183

文献标志码: A

文章编号: 1673-3851 (2026) 01-0114-11

引文格式: 王迎铭, 陈柯烽, 潘海鹏, 等. WMG-GAN: 基于权重图引导的布匹瑕疵图像生成算法[J]. 浙江理工大学学报(自然科学), 2026, 55(1): 114-124.

Reference Format: WANG Yingming, CHEN Kefeng, PAN Haipeng, et al. WMG-GAN: Weight-map-guided fabric defect image generation algorithm[J]. Journal of Zhejiang Sci-Tech University, 2026, 55(1): 114-124.

WMG-GAN: Weight-map-guided fabric defect image generation algorithm

WANG Yingming¹, CHEN Kefeng¹, PAN Haipeng^{1,2}, REN Jia^{1,2}

(1. School of Information Science and Engineering, Zhejiang Sci-Tech University, Hangzhou 310018, China; 2. Changshan Research Institute Co., Ltd. of Zhejiang Sci-Tech University, Quzhou 324299, China)

Abstract: Aiming at the limitations of existing methods in reconstructing background details and achieving high-quality image generation, this paper proposes a weight-map-guided generative adversarial network (WMG-GAN), a fabric defect image generation algorithm based on the CycleGAN framework. The algorithm first employs a generator to produce both a foreground weight map and a feature weight map, enabling selective modification of foreground content while completely preserving background details and structure. Secondly, a ConvNeXt V2 module is incorporated into the discriminator to enhance the network's feature extraction capability, thereby providing more precise gradient feedback to the generator. Finally, the learned perceptual image patch similarity (LPIPS) metric is introduced to construct the cycle-consistency loss function, optimizing the visual quality and realism of the generated images. Comparative

收稿日期: 2025-07-15 网络出版日期: 2025-11-05

基金项目: 浙江省“尖兵”“领雁”研发攻关计划项目(2023C01062)

作者简介: 王迎铭(2001—), 男, 河南商丘人, 硕士研究生, 主要从事机器视觉、图像生成方面的研究。

通信作者: 任 佳, E-mail: jren@zstu.edu.cn

and ablation experiments conducted on a real-world fabric defect dataset demonstrate that the proposed algorithm achieves lower Fréchet inception distance (FID) and LPIPS values, alongside higher structural similarity index measure (SSIM) and peak signal-to-noise ratio (PSNR) compared to the traditional CycleGAN. The WMG-GAN algorithm significantly enhances the quality of generated images, and the resulting images meet the high-precision requirements for automated defect detection algorithms.

Key words: fabric defects generation; generative adversarial network; CycleGAN; weight map; ConvNeXt V2

0 引言

在纺织行业中,布匹质量检测是确保最终产品质量的关键环节之一。随着计算机视觉与机器学习技术的发展,基于图像分析的布匹瑕疵检测系统已成为重要的研究方向^[1]。然而,布匹瑕疵图像的获取面临多重挑战^[2]。首先,布匹瑕疵种类繁多且形态特征各异,大幅增加了图像采集难度;其次,现有瑕疵图像数量有限,特定类型或罕见瑕疵的样本尤为匮乏,这种数据稀缺性严重制约了瑕疵检测网络模型的训练效果与泛化能力。此外,受生产环境复杂性影响,采集图像常存在光照不均、背景干扰等问题,进一步降低了所采集瑕疵图像的可用性。因此,开发一种能够精确控制瑕疵特征并生成高质量图像的方法,对于优化织物质量控制流程、提升瑕疵检测网络模型性能具有重要的现实意义。

在当前图像生成方法中,GAN 网络通过生成器与判别器之间的对抗性训练来优化其生成能力,并能产生具有高视觉逼真度的图像,因而被视为该领域的代表性模型^[3]。然而其训练过程存在明显缺陷。例如,发生模式坍塌 (Mode collapse) 时,生成器仅能输出有限种类的图像,导致生成样本多样性不足;训练过程不稳定,造成生成图像的质量参差不齐^[4]。此外,对于艺术风格迁移^[5]等需要配对数据的任务,样本的获取成本十分高昂。

针对上述问题,目前已有多种解决方案。黄超等^[6]使用 CycleGAN 并通过引入循环一致性损失 (Cycle consistency loss),确保图像经跨域转换并重构后能与原始输入保持高度一致,从而保证转换过程的稳定性;Chan 等^[7]提出的 Pi-GAN 采用具有周期性激活函数的神经表示与体积渲染的方法,实现了高质量的 3D 感知图像合成,同时离散标签、物体关键点、语义分割图及参考图像等附加条件信息可进一步引导图像生成过程;Nirkin 等^[8]提出的 FSGAN 基于循环神经网络实现面部重演,能自适应调整姿态与表情变化,适用于单帧图像与视频序

列处理;Hou 等^[9]分别为前向和反向循环一致性重建过程建立独立的退化分支,并共享同一复原分支进行正则化,使两个退化过程相互约束,以提升模型的鲁棒性与生成质量;Wen 等^[10]提出的 AttenCGAN 则将循环一致性对抗网络与注意力机制相结合,用于合成瑕疵样本并增强特征表达。尽管循环一致性损失为风格迁移开辟了新途径,但其仍存在显著缺陷:在循环一致性约束下,生成器可能过度关注常见的特征,而忽略关键结构细节,导致生成图像虽风格相似却存在细节失真的现象^[11]。这一现象在处理高频信号如复杂户外场景时尤为突出,这是因为背景细节的多样性与复杂性,导致生成算法难以准确重建这些内容^[12]。

上述现象限制了循环一致性损失在风格迁移和图像编辑等领域的应用,现有算法大多无法捕捉图像转换任务中的高阶语义信息,导致背景与前景的视觉元素相互干扰。针对上述问题,众多学者尝试通过引入外部约束或先验知识来增强模型的高阶语义理解能力。孙增国等^[13]使用 EMA 注意力保留关键特征和位置信息,使网络获得更多位置信息并学习到输入数据的空间关系;Wu 等^[14]基于 Retinex 理论简化了 CycleGAN 的低光照增强问题,利用 CycleGAN 合成的配对图像指导 Retinex 分解网络训练,使网络能自适应学习低光照图像增强并提升复杂光照场景的处理能力;Zhou 等^[15]通过共享编码器捕获两域间的共同分布实现跨域图像转换,获得了更真实的转换效果;Yang 等^[16]通过量化不同抽象层次的语义信息,探究网络各层表征与输出图像语义间的因果关系,并据此构建了因果干预模块,在生成过程中强化关键语义特征,从而使输出图像更加符合语义约束。上述方法有效提高了生成图像的质量,但通常需要额外的网络或数据来实现。

此外,现有生成算法普遍注重生成器的图像合成能力,而忽视了判别器的关键作用。GAN 的训练过程本质上是一个零和博弈的过程,生成器与判别

器通过对抗训练实现性能上的同步优化,判别器为生成器的优化提供所需要的梯度信息^[17]。当生成器过度拟合训练数据中的常见模式时,会忽略图像的高层结构与细节特征,导致输出结果模糊失真,缺乏真实感;更关键的是,生成器依赖判别器的反馈来学习数据分布,若判别器无法提供准确反馈,生成器在新数据上将难以产生高质量图像^[18]。

现有算法虽然解决了非成对训练数据生成图像的问题,但生成过程中图像主体特征时常发生改变,无法映射所需的特征关系。这是由于现有算法大多采用无监督训练方式,使得网络无法保证对应关系的转换,对目标域的特征信息和位置信息不够敏感。为了解决上述生成算法无法重建背景细节、生成结果模糊、判别器性能差导致图像质量偏低等问题,本文提出了一种新型图像生成算法 WMG-GAN (Weight map guided-generative adversarial network)。该算法基于 CycleGAN 框架进行改进,首先由生成器生成前景区权重图与特征权重图,提

取源域与目标域中最具区分性的语义部分;然后,为进一步增强判别器对图像细节特征的提取能力,采用一种由权重图引导的判别器架构,并在其中集成了 ConvNeXt V2 模块^[19],显著提升其对真假图像差异的识别精度,促使生成器产出更逼真的细节;最后,引入感知学习图像块相似性 (Learned perceptual image patch similarity, LPIPS) 指标,构建循环一致性损失函数,以优化视觉质量。本文在课题组采集的布匹瑕疵数据集上进行实验,以验证该算法所生成图像的可用性。

1 WMG-GAN 算法介绍

WMG-GAN 的生成框架如图 1 所示。首先生成器 $G_{X \rightarrow Y}$ 生成前景权重图 F_{y2} 和特征权重图 F_{y1} , 二者与输入图像融合以获取生成图像 G_y ; 接着 G_y 经过生成器 $G_{Y \rightarrow X}$ 生成其对应的前景权重图 F_{x2} 和特征权重图 F_{x1} , 进而转化为重建图像 R_x ; 最后将生成图像 x 与重建图像 R_x 进行对比以获取 LPIPS 损失。

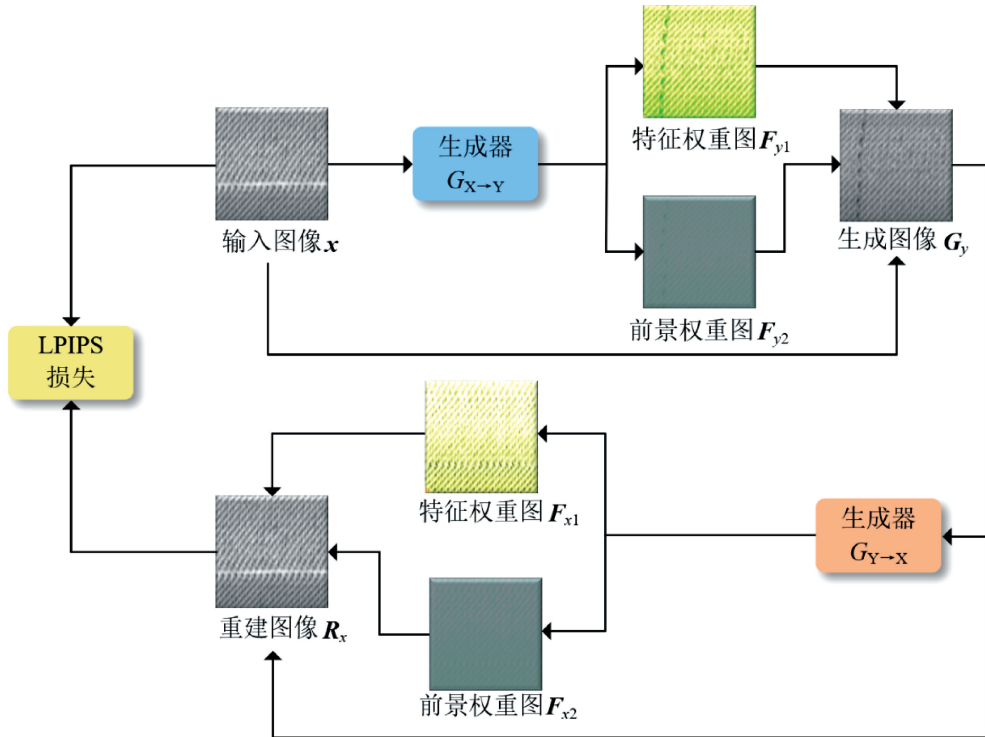


图 1 WMG-GAN 生成框架示意图

1.1 CycleGAN

传统生成算法通常需要成对的训练数据,然而在实际的图像采集,很难获得具有不同风格的图像对。CycleGAN 通过引入循环一致性损失解决了这个问题。CycleGAN 的整体结构如图 2 所示,其主体结构由生成器和判别器两部分组成,通过生成

器实现 X 域和 Y 域之间的相互映射,最后再由判别器将原始输入图像和生成图像进行比较,以获得循环一致性损失。域是指不同风格且无法建立一一对应关系的图像集合,本文设置了 3 个图像集合,其中 X 域表示纬向瑕疵图像集合, Y 域表示经向瑕疵图像集合, Z 域表示无瑕疵图像集合。

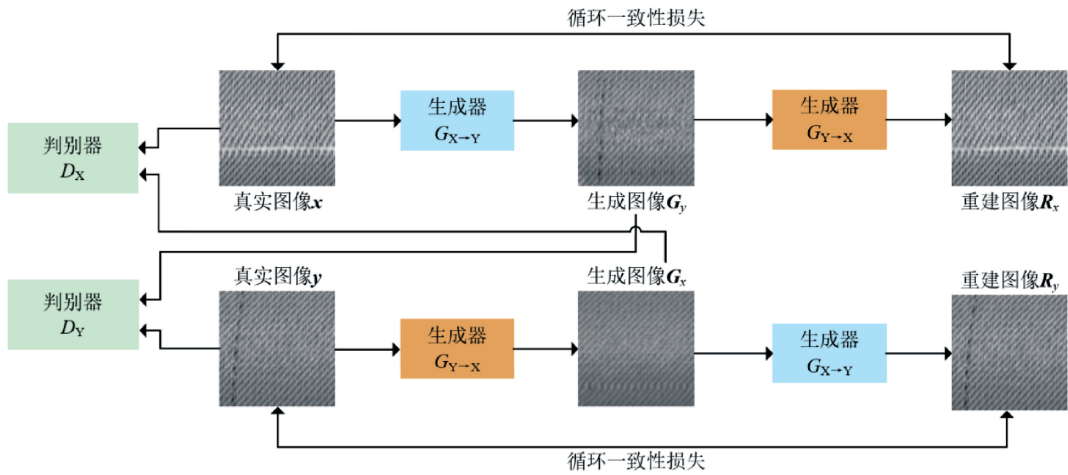


图 2 CycleGAN 模型结构图

1.2 WMG-GAN 生成器

WMG-GAN 的生成器采用混合结构, 将自编码器与跳跃连接相结合。自编码器是一种无监督学习方法, 通过对数据进行编码和解码以保留其结构信息; 跳跃连接通过促进低层特征在不同网络层之间的传递, 以缓解梯度消失的问题, 从而增强特征的保留与融合。WMG-GAN 的生成器结构

如图 3 所示, 主要由编码器和解码器两部分组成, 具体操作为: 输入图像先经过镜像填充, 然后经过 6 个卷积核大小为 4、步长为 2、填充层为 1 的卷积模块进行特征提取, 每经过 2 次特征提取后由归一化层进行特征整合, 接着经过残差块以缓解梯度消失现象, 最后使用反卷积操作还原生成图像。

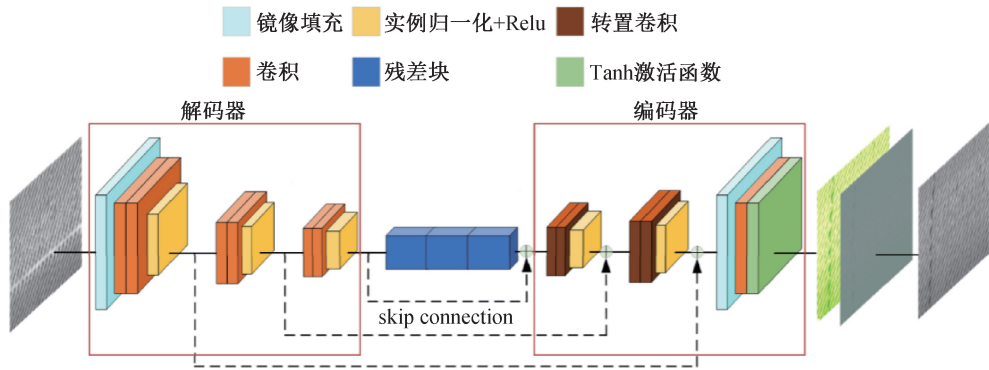


图 3 WMG-GAN 生成器结构图

为了学习 X 域和 Y 域中最具区分性的语义对象, 同时最小化对背景部分无关内容的修改, WMG-GAN 的生成器内部集成了一个注意力模块。在瑕疵生成任务中, 能否做到保留源域的背景和目标域的前景非常关键。该算法的目的是学习 X 域和 Y 域之间的两个映射, 即 $G_{X \rightarrow Y}: \mathbf{x} \rightarrow (\mathbf{F}_{y1}, \mathbf{F}_{y2}) \rightarrow \mathbf{G}_y$ 和 $G_{Y \rightarrow X}: \mathbf{y} \rightarrow (\mathbf{F}_{x1}, \mathbf{F}_{x2}) \rightarrow \mathbf{G}_x$ 。前景权重图 \mathbf{F}_{x2} 和 \mathbf{F}_{y2} 可定位前景区域的最显著部分; 特征权重图 \mathbf{F}_{x1} 和 \mathbf{F}_{y1} 通过量化每个像素的重要性来明确风格转换的具体范围。这两种权重图共同作用, 确保转换仅针对关键语义内容, 从而保持背景不变。最后将输入图像 \mathbf{x} 、特征权重图 \mathbf{F}_{y1} 、前景权重图 \mathbf{F}_{x1} 进行融合以获取生成图像 \mathbf{G}_y 。 \mathbf{G}_y 可用式(1)表示:

$$\mathbf{G}_y = \mathbf{x} \times (1 - \mathbf{F}_{y2}) + \mathbf{F}_{y1} \times \mathbf{F}_{y2} \quad (1)$$

其中: $\mathbf{x} \times (1 - \mathbf{F}_{y2})$ 表示生成图像保留了源域图像的背景; $\mathbf{F}_{y1} \times \mathbf{F}_{y2}$ 表示网络着重关注目标域图像的前景, 即需要做出改变的瑕疵区域。

1.3 WMG-GAN 判别器

传统 CycleGAN 的判别器采用 PatchGAN 结构, 该结构将输入图像划分为多个局部图像块, 并对每个块进行独立的真伪判别, 给出每一块来自真实图像的置信度, 从而使判别器具备局部细节的鉴别能力。PatchGAN 结构虽能有效提升判别器对局部细节的鉴别能力, 但其对局部区域的过度关注会导致生成图像在全局结构上缺乏一致性与连贯性。为了提高传统 PatchGAN 判别器对背景细节区域的判别能力, WMG-GAN 的判别器由权重图引导, 并引入 ConvNeXt V2 模块, 以提高判别器对生成图像和真实

图像的判别能力,并增强网络提取特征的能力。

1.3.1 ConvNeXt V2

ConvNeXt V2 是一种先进的卷积神经网络架构,作为 ConvNeXt 的升级版,它在继承原有模型融合卷积与 Transformer 优势的基础上,引入了全新的设计思路,以提升模型性能与效率。其核心机制类似于一种注意力机制,能够对特征进行重新标定,从而增强网络的特征表示能力。具体而言,ConvNeXt V2 在 ConvNeXt 的基础上引入了全局响应归一化(Global response normalization)层,以增强通道间的特征竞争机制,从而提升特征的选择性与对比度。GRN 层通过全局归一化处理每个通道的响应强度,旨在强化对任务关键的特征通道,同时抑制不重要的通道,进而优化网络整体的特征表

示能力与模型性能。

ConvNeXt V2 结构图如图 4 所示。GRN 层的操作为先计算全局响应,对于一张输入特征图 $\mathbf{P} \in \mathbf{R}^{H \times W \times C}$ (H, W, C 分别表示特征图的高度、宽度和通道数),计算每个通道的全局响应 $r_i = \sum_{h=1}^H \sum_{w=1}^W p_{h,w,i}^2$ (h, w 和 i 分别表示该层通道特征图的高度、宽度和维度),即该通道所有像素值的平方和;再计算归一化因子,对全局响应进行归一化处理,得到每个通道的归一化因子 $s_i = \frac{r_i}{\sqrt{\sum_{j=1}^C r_j + \sigma}}$,其中 σ 是一个很小的常数,用于防止除零错误;最后将归一化因子应用于原始特征图,得到归一化后的特征图,即 $\mathbf{P}'_{h,w,i} = \mathbf{P}_{h,w,i} \times s_i$ 。

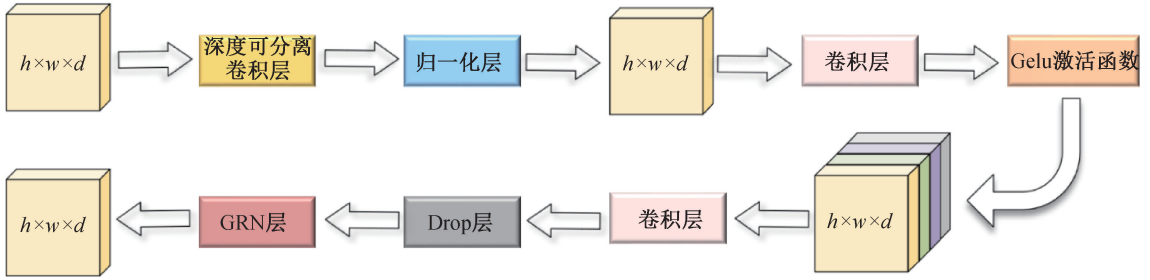


图 4 ConvNeXt V2 结构图

1.3.2 基于权重图引导的判别器

WMG-GAN 的判别器在结构上与 PatchGAN 相似,但它不止接收真实图像和生成图像,还将权重图 \mathbf{F}_{x1} 和 \mathbf{F}_{y2} 作为输入。例如, D_Y 的输入是生成图像对 $(\mathbf{F}_{y1}, G_{X \rightarrow Y}(x))$ 和真实图像对 (\mathbf{F}_{y1}, y) 。

WMG-GAN 的判别器集成了 ConvNeXt V2 模块,以改善特征间的竞争,提高特征的选择性和对比

度。当背景中有噪声或不规则的纹理,传统 PatchGAN 可能会过度关注这些区域,导致生成的图像受到背景噪声的影响,但加入 ConvNeXt V2 模块后,判别器能更好地关注关键区域,减少背景噪声的干扰,使判别器能够更精确地识别图像中的重要区域,从而提高对生成图像质量的评估准确性,促使网络生成更精细的图像。WMG-GAN 的判别器结构如图 5 所示。

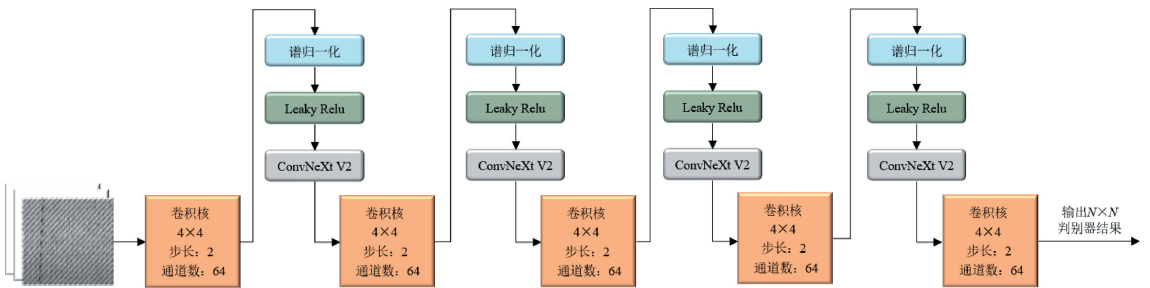


图 5 WMG-GAN 判别器结构图

1.4 损失函数

1.4.1 LPIPS 损失

循环一致性损失通过计算图像在循环重建后与原始图像之间的差异来实现,通常采用 L_1 和 L_2 范数进行度量,然而这些像素级度量难以有效感知图

像质量的变化。为优化循环转换效果并提高网络对图像质量变化的感知能力,本文在 WMG-GAN 中采用 LPIPS 指标来构建循环一致性损失。

LPIPS 通过预训练的神经网络(如 VGG、AlexNet)提取的高级特征来衡量图像间的相似

性。与传统基于像素级差异(如 L_1 或 L_2 损失)的方法不同, LPIPS 能够更好地捕捉人类视觉系

统对图像变化的感知。LPIPS 的具体网络结构如图 6 所示。

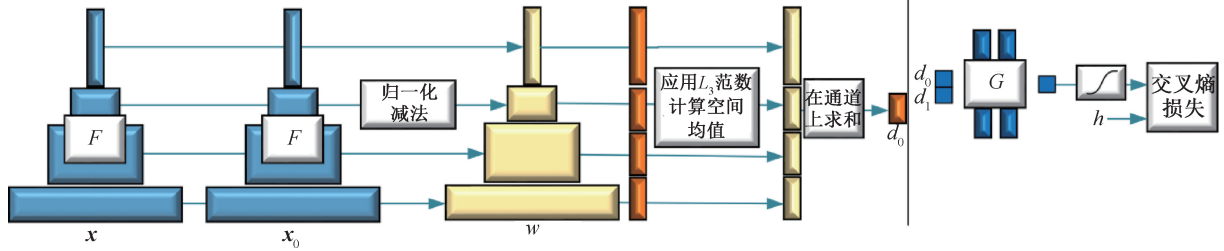


图 6 LPIPS 网络结构图

LPIPS 计算方法为: 首先将所选取的一组图像 x 和 x_0 送入网络 F 中进行特征提取, 计算二者在不同通道中的特征的距离; 接着在不同的卷积层中提取特征堆栈, 并对不同通道中的特征堆栈进行归一化; 随后使用 L_2 范数计算距离, 并在空间中求平均值, 在通道上求和计算出 d_0 ; 最后将 d_0 和 d_1 (由预训练模型计算出的真实值) 传到含有两个 32 个通道的 ReLU 全连接层、一个单通道全连接层和一个 sigmoid 层的模型中进行拟合, 用于将原始距离映射至与人类感知一致的 LPIPS 值。 d_0 计算公式可用式(2)表示:

$$d_0 = d(x, x_0) = \sum_L \frac{1}{H_L \times W_L} \sum_{h,w} \|\mathbf{q}(\hat{\mathbf{b}}_{hw}^L - \hat{\mathbf{b}}_{0hw}^L)\|^2 \quad (2)$$

其中: L 表示预训练神经网络中特征层的层数; H 和 W 表示特征图的高度和宽度; h 和 w 表示特征图上的空间位置坐标; $\hat{\mathbf{b}}_{hw}^L$ 和 $\hat{\mathbf{b}}_{0hw}^L$ 表示位于 (h, w) 处的归一化特征向量; \mathbf{q} 表示与该层特征图通道数相同的可学习权重向量。

相似性损失函数 L_{LPIPS} 可用式(3)表示:

$$L_{LPIPS}(x, x_0, x_1, h) = -h \log G(d(x, x_0), d(x, x_1)) - (1-h) \log(1 - G(d(x, x_0), d(x, x_1))) \quad (3)$$

其中: x_1 表示所选取的生成图像。

1.4.2 损失函数

除了 LPIPS 损失以外, 该算法还使用了传统对抗性损失、基于权重图引导的对抗性损失。

在对抗性训练中, 生成器与判别器进行一种最小最大博弈。判别器的目标是最大化其区分真伪的能力, 即对真实图像赋予接近 1 的概率, 对生成图像赋予接近 0 的概率。与之相对, 生成器的目标则是学习一个从域 X 到域 Y 的映射, 生成足以“欺骗”判别器的图像, 诱导其将这些生成图像误判为真实图像。因此, 对于映射 $G_{X \rightarrow Y}$, 对抗性损失可用式(4)

表示:

$$L_{GAN}(G_{X \rightarrow Y}, D_Y) = E_{y \sim p_{data}(y), x \sim p_{data}(x)} ((\log D_Y(y)) + (\log(1 - D_Y(G_{X \rightarrow Y}(x)))))) \quad (4)$$

对于另一个映射 $G_{Y \rightarrow X}$, 其对抗性损失可用式(5)表示:

$$L_{GAN}(G_{Y \rightarrow X}, D_X) = E_{x \sim p_{data}(x), y \sim p_{data}(y)} ((\log D_X(x)) + (\log(1 - D_X(G_{Y \rightarrow X}(y)))))) \quad (5)$$

为了尽可能地保持背景不变, 引入了权重图引导的对抗性损失, 以保持真实图像和生成图像所生成的特征权重图尽可能一致, 使关键信息得以保留。对于映射 $G_{X \rightarrow Y}$, 权重图引导的对抗性损失可用式(6)表示:

$$L_F(G_{X \rightarrow Y}, D_Y) = E_{y \sim p_{data}(y), x \sim p_{data}(x)} ((\log D_Y(\mathbf{F}_{y1}, \mathbf{y})) + (\log(1 - D_Y(\mathbf{F}_{y1}, G_{X \rightarrow Y}(x)))))) \quad (6)$$

其中: D_Y 的任务是区分和比较生成图像对 $(\mathbf{F}_{y1}, G_{X \rightarrow Y}(x))$ 和真实图像对 $(\mathbf{F}_{y1}, \mathbf{y})$ 。

最终完整损失函数可用式(7)表示:

$$L = L_{GAN} + L_F + L_{LPIPS} \quad (7)$$

2 实验结果与分析

2.1 环境配置和数据集

实验环境配置如下: 操作系统为 Windows 11 (64 位), 硬件平台搭载 Intel Core i5-12600KF CPU 与 NVIDIA GeForce RTX 3090 GPU, 运行内存为 24 GiB; 软件开发环境为 PyCharm, 采用基于 Python 3.8 与 PyTorch 1.8.1 的深度学习框架, 并行计算环境为 CUDA 11.10。

参数设置如下: 所有算法模型都在相同的数据集上训练 400 个 epoch, batch_size 设置为 16, 初始学习率为 0.01, 动量设置为 0.937, 使用随机梯度下降法进行优化。

本文所用数据集采集自某纺织企业织布车间。图像采集系统位于出布口, 由 6 台并行摆放的面阵 CCD 相机组成, 可实时拍摄布料图像, 原始图像尺

寸为 3072×96 像素。为满足布匹瑕疵实时监测的需求,并避免直接处理极端尺寸图像导致经向特征信息丢失的问题,本文采用了一种基于滑动窗口的图像预处理方法。该方法将原始图像切割成多个 96×96 像素的子图像用于训练。子图像高度与原始图像保持一致(96 像素);在宽度方向上,采用步长为 3 像素的滑动窗口进行分割,即相邻子图像间有 93 像素的重叠。最终,每张原始图像被划分为 33 个子图像。

数据集含有毛纬、带纬、白档子、曲纤、双经和断经共 6 种类型的瑕疵图以及正常图像样本,其中:正常样本 145 张,毛纬 291 张,带纬 299 张,白档子 182 张,曲纤 115 张,双经和断经各 139 张。考虑到布匹由经线和纬线两组线组纺织而成,为了进行机器故障的定位,将瑕疵样本按照瑕疵产生的原因分别标注为“纬向瑕疵图像”(放置在 X 域)和“经向瑕疵图像”(放置在 Y 域)两类,将正常样本标注为“正常图像”(放置在 Z 域)。然后使用数据增强技术,如旋转、噪声添加、亮度调整等方法进行数据扩充,扩充后的数据集中共含有图像样本 1459 张,每个域的图像分为训练集和测试集:X 域有 887 幅训练图像和 70 幅测试图像;Y 域有 278 幅训练图像和 49 幅测试图像;Z 域有 145 幅训练图像和 30 幅测试图像。不同域图像示例如图 7 所示。

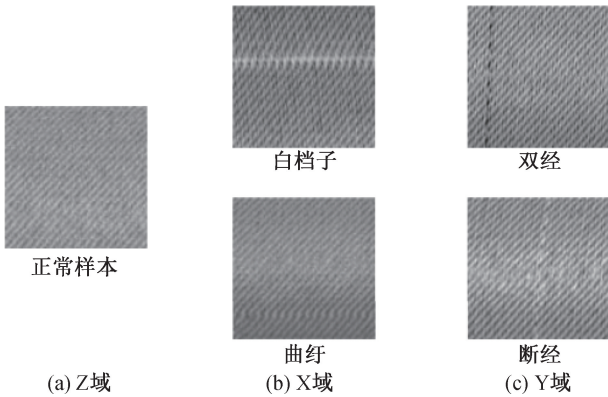


图 7 不同域图像示例

2.2 评价指标

本文使用算法评估指标有弗雷歇初始距离(Fréchet inception distance, FID)、峰值信噪比(Peak signal-to-noise ratio, PSNR)、结构相似性指数(Structural similarity index measure, SSIM)、LPIPS。FID通过计算生成图像与真实图像在特征空间中的均值和协方差矩阵,进而度量两者多元高斯分布之间的距离,以评估生成图像的质量。FID值越低,表示生成图像与真实图像在视觉上的差异

越小,生成图像的质量越高。其计算公式可写为:

$$M_{\text{FID}} = \|\nu_g - \nu_r\|_2^2 + \text{tr}(\mathbf{N}_g + \mathbf{N}_r - 2(\mathbf{N}_g \mathbf{N}_r)^{\frac{1}{2}}) \quad (8)$$

其中: M_{FID} 表示两组图像间的 FID 值; $\|\nu_g - \nu_r\|_2^2$ 表示均值向量差的平方的 L_2 范数; tr 表示矩阵的迹,即矩阵对角线元素的和; $(\mathbf{N}_g \mathbf{N}_r)^{\frac{1}{2}}$ 表示协方差矩阵 \mathbf{N}_g 和 \mathbf{N}_r 的乘积的平方根。

PSNR 通过计算原始图像与处理后图像之间的均方误差来衡量图像失真程度,其计算公式如式(9)所示:

$$M_{\text{PSNR}} = 10 \times \lg \frac{V^2}{E_{\text{MS}}} \quad (9)$$

其中: M_{PSNR} 表示两组图像间的 PSNR 值; E_{MS} 表示两张图像的均方误差; V 表示图像像素可取到的最大值。PSNR 计算简单且结果直观,但其主要反映像素级别的差异,往往无法完全对应人类视觉系统的感知质量。因此,在实际应用中,常需结合 SSIM 或 LPIPS 等更复杂的感知质量指标,以提供更全面的评价。

SSIM 是一种衡量两幅图像结构相似性的指标,它通过比较图像的亮度、对比度和结构信息来计算相似度。该指标因考虑了人类视觉系统的感知特性,在处理复杂纹理和结构的图像时表现尤为出色,其计算公式可写为:

$$M_{\text{SSIM}}(\mathbf{x}, \mathbf{y}) = \frac{(2\mu_x \mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \quad (10)$$

其中: $M_{\text{SSIM}}(\mathbf{x}, \mathbf{y})$ 表示两组图像间的 SSIM 值; μ_x 和 μ_y 分别表示输入图像 \mathbf{x} 和 \mathbf{y} 的平均值; σ_x^2 和 σ_y^2 分别表示图像 \mathbf{x} 和 \mathbf{y} 的方差; σ_{xy} 表示图像 \mathbf{x} 和 \mathbf{y} 的协方差; c_1 和 c_2 表示用来稳定分母的小常数。

LPIPS 通过加权整合深度网络不同层的特征以计算相似度,从而更准确地模拟人类视觉系统的感知特性。本文以 VGG-16 作为预训练的特征提取网络。

2.3 消融实验

为了系统地评估每个组件对算法性能的影响,本文进行了消融实验。对基础 CycleGAN 为基准进行多组改进以构建对照模型,将仅改进生成器的算法命名为 G-CycleGAN;将 ConvNeXt V2 模块加入到 G-CycleGAN 判别器中,该算法命名为 GD-CycleGAN。对生成图像和真实图像进行 FID、SSIM、PSNR、LPIPS 值计算,计算结果如表 1—表 3 所示。

表 1 纬向瑕疵图像(X域)消融实验结果

算法	FID↓	SSIM↑	PSNR↑	LPIPS↓
CycleGAN	57.3320	0.0853	14.9017	0.4499
G-CycleGAN	34.8531	0.1082	16.8159	0.2058
GD-CycleGAN	16.1625	0.1527	15.8284	0.1993
WMG-GAN	15.6935	0.4784	31.1031	0.0799

表 2 经向瑕疵图像(Y域)消融实验结果

算法	FID↓	SSIM↑	PSNR↑	LPIPS↓
CycleGAN	38.0585	0.1687	17.3518	0.1290
G-CycleGAN	35.2029	0.1055	19.3670	0.1131
GD-CycleGAN	24.0220	0.2364	20.3105	0.0854
WMG-GAN	20.9138	0.2489	20.8625	0.0849

表 3 无瑕疵图像(Z域)消融实验结果

算法	FID↓	SSIM↑	PSNR↑	LPIPS↓
CycleGAN	57.3320	0.0853	14.9017	0.4499
G-CycleGAN	60.6342	0.0720	14.1419	0.2064
GD-CycleGAN	49.4156	0.1508	19.5663	0.1079
WMG-GAN	20.8715	0.1722	20.2126	0.1209

从表 1—表 3 中可以看出,当仅对生成器进行改进时,X 域图像中的 FID、LPIPS 分别降低了 39.21%、54.26%,SSIM、PSNR 分别提高了 26.85%、12.85%;Y 域图像中的 FID、SSIM、LPIPS 分别降低了 7.51%、37.46%、12.33%,PSNR 提高了 11.61%;Z 域图像中的 FID 值提高了 5.76%,SSIM、PSNR、LPIPS 分别降低 15.59%、5.09%、54.12%。这表明只改进生成器可能会导致生成图像的质量下降,并在某些生成图像中出现整体结构信息的丢失。这是由于传统 CycleGAN 的判别器特征提取能力较弱,难以捕捉图像中的细微差异和高层特征,导致算法无法获得准确的反馈信息。此外,该算法依赖权重图来区分前景和背景区域,如果判别器无法准确识别这些区域,则生成的图像可能会出现背景与前景混淆的问题,进而导致图像细节不够丰富、真实感较差。当增强判别器的能力后,X 域图像中的 FID、PSNR、LPIPS 分别降低了 53.63%、5.78%、3.16%,SSIM 提高了 41.13%;Y 域图像中的 FID、LPIPS 分别降低了 31.76%、24.49%,SSIM 和 PSNR 分别提高了 124.88%和 4.87%;Z 域图像中 FID、LPIPS 分别降低 18.51%、47.72%,SSIM 和 PSNR 分别提高 109.44%和 38.36%。这表明生成图像的质量得到了显著提升。这是因为通过增强判别器对图像细节和结构信息的捕捉能力,判别器网络能够为生成器提供更加准确的反馈信息,从而实现更高质量的图像生成效果。

此外,在循环一致性损失中引入 LPIPS 指标进一步提升了算法的性能:X 域图像中的 FID、LPIPS 分别降低了 2.91%、59.19%,SSIM、PSNR 分别提高了 213.29%、96.51%;Y 域图像中的 FID、LPIPS 分别降低了 12.94%、0.59%,SSIM、PSNR 分别提高了 5.29%、2.72%;Z 域图像中 FID 值降低 57.76%,SSIM、PSNR 和 LPIPS 分别提高了 14.19%、3.31%和 12.05%。这是由于 LPIPS 强调感知相似性,有助于生成不仅在结构上相似,而且在视觉感知上更加真实的图像。

2.4 与主流算法的对比实验

在评估 WMG-GAN 用于图像生成的性能时,本文将该算法与目前主流的生成算法进行了对比,对比算法包括 MUINT^[20]、DRIT^[21]、StarGAN-v2^[22]、CycleGAN 和 WMG-GAN。不同算法所生成的布匹瑕疵图像的评价指标如表 4—表 6 所示。

表 4 纬向瑕疵图像(X域)对比结果

算法	FID↓	SSIM↑	PSNR↑	LPIPS↓
MUINT	180.5731	0.0883	19.9801	0.4023
DRIT	87.2472	0.1124	20.5113	0.4512
StarGAN-v2	45.3983	0.0981	22.5377	0.2247
CycleGAN	17.5134	0.1155	16.1352	0.1865
WMG-GAN	15.6935	0.4784	31.1031	0.0799

表 5 经向瑕疵图像(Y域)对比结果

算法	FID↓	SSIM↑	PSNR↑	LPIPS↓
MUINT	103.3276	0.1050	15.8623	0.3827
DRIT	67.1428	0.0815	17.8924	0.1044
StarGAN-v2	39.2174	0.1349	19.4789	0.2654
CycleGAN	38.0585	0.1687	17.3518	0.1290
WMG-GAN	20.9138	0.2489	20.8625	0.0849

表 6 无瑕疵图像(Z域)对比结果

算法	FID↓	SSIM↑	PSNR↑	LPIPS↓
MUINT	92.7513	0.1024	15.5016	0.3750
DRIT	40.5561	0.1455	19.9127	0.2716
StarGAN-v2	23.9178	0.0886	17.5512	0.1673
CycleGAN	57.3320	0.0853	14.9017	0.4499
WMG-GAN	20.8715	0.1722	20.2126	0.1209

从表 4—表 6 可以看出,本文算法在生成各种瑕疵图像方面表现出良好的性能。WMG-GAN 在 X 域图像中的 FID 值分别比 MUINT、DRIT、StarGAN-v2、CycleGAN 低 77.49%、48.54%、12.74%、63.59%;在 Y 域图像中的 FID 值分别比 MUINT、DRIT、StarGAN-v2、CycleGAN 低 79.76%、68.85%、46.67%、45.05%;在 Z 域图像中的 FID 值分别比 MUINT、DRIT、StarGAN-v2、CycleGAN 低 77.51%、

48.54%、12.74%、63.01%。这表明生成的瑕疵图像在质量和多样性方面与真实的瑕疵图像高度相似。在 SSIM 方面,WMG-GAN 生成图像的质量也取得较大提升,在 X 域图像中的 SSIM 值分别比 MUINT、DRIT、StarGan-v2、CycleGAN 高出 441.79%、325.63%、387.67%、314.22%;在 Y 域图像中的 SSIM 值分别比 MUINT、DRIT、StarGan-v2、CycleGAN 高出 137.05%、205.41%、84.51%、47.54%;在 Z 域图像中 SSIM 值分别比 MUINT、DRIT、StarGan-v2、CycleGAN 高出 68.16%、18.35%、94.36%、101.88%。这表明生成的瑕疵图像和原始瑕疵图像之间的结构相似性非常高,表明该算法在保留结构信息方面的有效性。此外,WMG-GAN 生成图像的 PSNR 值也有明显提升,在 X 域图像对比结果中分别比 MUINT、DRIT、StarGan-v2、CycleGAN 高 30.39%、1.51%、15.16%、35.64%;在 Y 域图像对比结果中分别比 MUINT、DRIT、StarGan-v2、CycleGAN 高 31.52%、16.61%、7.17%、20.23%;在 Z 域图像对比结果中分别比 MUINT、DRIT、StarGan-v2、CycleGAN 高 30.39%、1.51%、15.16%、35.64%。这说明了生成的瑕疵图像的重建质量好、失真较低,证明了该算法在图像生成过程中的精度和稳定性。

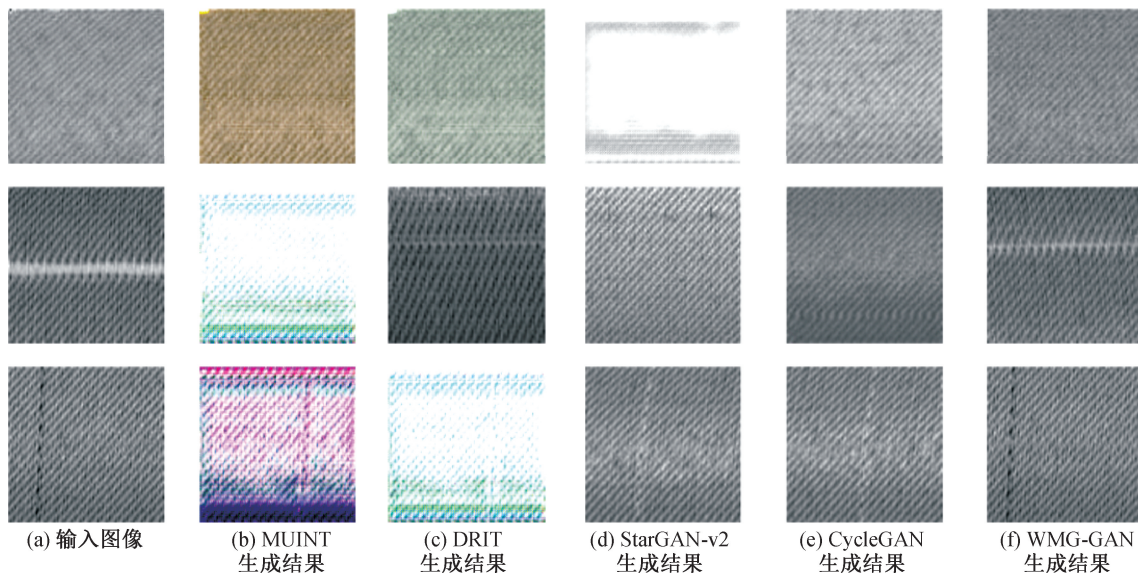


图 8 不同算法的生成图像

2.5 目标检测算法对比实验

为了验证 WMG-GAN 所生成的布匹瑕疵图像的可用性,以 900 张真实布匹瑕疵图像构成的数据集对主流目标检测算法进行了训练,然后在真实瑕疵图像的数据集中加入 300 张 WMG-GAN 生成的布匹瑕疵图像,对算法进行混合训练,进一步验证所

WMG-GAN 获得了较低的 LPIPS 值,在 X 域图像中分别比 MUINT、DRIT、StarGan-v2、CycleGAN 降低了 80.14%、82.29%、64.44%、57.16%;在 Y 域图像中分别比 MUINT、DRIT、StarGan-v2、CycleGAN 降低了 77.82%、18.68%、68.01%、34.19%;在 Z 域图像中分别比 MUINT、DRIT、StarGan-v2、CycleGAN 降低了 67.76%、55.49%、27.73%、73.13%。这表明生成的瑕疵图像和原始图像之间具有较高的感知相似性,进一步增强了算法生成高质量瑕疵图像的能力。

为直观地展示和比较不同生成算法瑕疵图像的生成能力,本文将上述算法在 X、Y 和 Z 域中生成结果做了可视化展示,如图 8 所示。从图 8 中可以看出:MUINT 生成的图像相对模糊,缺乏明显的瑕疵特征;DRIT 生成的图像存在伪影,瑕疵痕迹不够清晰;StarGan-v2 所生成图像中背景的细节纹理纠缠在一起;CycleGAN 生成的图像对比前者有显著改善,但背景颜色往往过暗,瑕疵也不够明显,这可能源于 CycleGAN 倾向于将背景纹理迁移到前景中,导致前景与背景信息混合;相比之下,WMG-GAN 生成的图像质量较高,便于识别瑕疵类别。可视化对比结果进一步验证了 WMG-GAN 在生成高质量布匹瑕疵图像方面的优越性。

生成图像的可用性,结果如表 7 所示。由表 7 可知:使用由真实布匹瑕疵图像和生成的布匹瑕疵图像组成的混合数据集对主流目标检测算法进行训练,各个算法的 mAP 均有较大提升,SSD、YOLOv4-tiny、YOLOv5s、YOLOv7-tiny 各提升 3.2%、1.3%、3.9%、8.3%,进一步印证了 WMG-GAN 生成的布

匹瑕疵图像数据质量高、种类多,可以提高目标检测算法对布匹瑕疵的检测能力和泛化性能。

表 7 生成数据加入前后的目标检测算法性能对比

算法	mAP/%	
	原始数据集	加入生成图像的数据集
SSD	90.5	93.7(3.2%↑)
YOLOv4-tiny	94.8	96.1(1.3%↑)
YOLOv5s	93.6	97.5(3.9%↑)
YOLOv7-tiny	88.1	96.4(8.3%↑)

以上实验结果表明,相较于主流算法,WMG-GAN一方面将前景区域和背景区分开来,在修改瑕疵细节的同时不会对源域背景产生较大影响;另一方面提升了判别器对特征的提取能力,进一步提高了生成图像的质量。该算法大大提高了布匹瑕疵数据的生成能力,使得不常见的或者不容易搜集的布匹瑕疵得以高质量生成,扩充后的数据集对训练目标检测算法的泛化性和检测精度也有较大提升。这表明该算法不仅在布匹瑕疵图像生成任务中具备出色的生成能力和可控性,还可为后续的布匹瑕疵检测任务提供了高质量、多样化的瑕疵图像数据。这种方式极大丰富了数据来源,降低了数据采集成本,为布匹瑕疵检测领域的增强提供了新的思路。

3 结 语

针对布匹瑕疵图像生成过程中出现无法重建背景细节、瑕疵特征不明显的问题,本文提出了一种基于权重图引导的高质量图像生成算法 WMG-GAN。该算法基于两种权重图来区分前景和背景,使其能在保留源域背景的同时修改目标域的重点内容;在判别器中集成 ConvNeXt V2 模块增强了网络捕捉细节和结构信息的能力,从而生成更加真实和高质量的图像。

实验结果表明,WMG-GAN 在图像质量和真实性方面超越了传统的 CycleGAN 及其他变体。WMG-GAN 获得较低 FID 值,表明生成图像和真实图像之间有紧密的语义相似性;获得较高的 SSIM 和 PSNR 值,保证了生成图像中结构细节的保留和较低的失真;此外,较低的 LPIPS 值表明生成图像与真实图像在感知上具有高度相似性。WMG-GAN 在非配对图像转换任务中的表现,表明其在扩展布匹瑕疵数据集以及改进瑕疵检测算法性能和泛化方面具有较好的潜力。通过生成多样化的瑕疵图像,WMG-GAN 可促进布匹自动检查系统的进步,满足实际工业现场的应用需求,具有较好的推

广性。

目前本文算法在生成布匹图像时仍存在一定的局限性,如生成特征不可控等。未来的工作将在纹理与背景特征的可控性优化方面开展进一步研究,以提升算法的泛化能力和生成质量。

参考文献:

- [1] 徐勇, 乔茹飞. 基于 CycleGAN 的脏污图像数据增强方法研究[J]. 信息技术与信息化, 2024, (12): 155-158.
- [2] 郭华. 基于生成对抗网络的图像修复技术研究[J]. 长江信息通信, 2024, 37(12): 64-66.
- [3] Kumar A, Soni S, Chauhan S, et al. Navigating the realm of generative models: GANs, diffusion, limitations, and future prospects: A review [C]//Proceedings of Fifth International Conference on Computing, Communications, and Cyber-Security. Singapore: Springer, 2024: 301-319.
- [4] Gui J, Sun Z, Wen Y, et al. A review on generative adversarial networks: Algorithms, theory, and applications [J]. IEEE Transactions on Knowledge and Data Engineering, 2023, 35(4): 3313-3332.
- [5] 林志坤, 许建龙, 包晓安. 基于 STGAN 的人脸属性编辑改进模型[J]. 浙江理工大学学报(自然科学), 2023, 49(3): 285-292.
- [6] 黄超, 胡勤友, 黄子硕. 基于改进 CycleGAN 的水上图像去雾算法[J]. 上海海事大学学报, 2025, 46(1): 17-22, 111.
- [7] Chan E R, Monteiro M, Kellnhofer P, et al. Pi-GAN: Periodic implicit generative adversarial networks for 3D-aware image synthesis[C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). June 20-25, 2021, Nashville, TN, USA. IEEE, 2021: 5795-5805.
- [8] Nirkin Y, Keller Y, Hassner T. FSGANv2: Improved subject agnostic face swapping and reenactment[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2023, 45(1): 560-575.
- [9] Hou H, Xu J, Hou Y K, et al. Semi-cycled generative adversarial networks for real-world face super-resolution[J]. IEEE Transactions on Image Processing, 2023, 32: 1184-1199.
- [10] Wen L, Wang Y, Li X Y. A new cycle-consistent adversarial networks with attention mechanism for surface defect classification with small samples[J]. IEEE Transactions on Industrial Informatics, 2022, 18(12): 8988-8998.
- [11] 包晓安, 高春波, 张娜, 等. 基于生成对抗网络的图像超分辨率方法[J]. 浙江理工大学学报(自然科学版), 2019, 41(4): 499-508.
- [12] 汤健, 郭海涛, 夏恒, 等. 面向工业过程的图像生成及其应用研究综述[J]. 自动化学报, 2024, 50(2): 211-240.
- [13] 孙增国, 彭学俊, 刘慧霞, 等. 基于自注意力机制和 CycleGAN 的高分三号 ScanSAR 图像的扇贝效应抑制[J]. 光子·激光, 2023, 34(12): 1279-1287.
- [14] Wu K, Huang J, Ma Y, et al. Cycle-retinex: Unpaired low-light image enhancement via retinex-inline CycleGAN [J]. IEEE Transactions on Multimedia, 2024, 26: 1213-1228.

- [15] Zhou Y F, Jiang R H, Wu X, et al. BranchGAN: Unsupervised mutual image-to-image transfer with a single encoder and dual decoders [J]. *IEEE Transactions on Multimedia*, 2019, 21(12): 3136-3149.
- [16] Yang C, Shen Y, Zhou B. Semantic hierarchy emerges in deep generative representations for scene synthesis[J]. *International Journal of Computer Vision*, 2021, 129(5): 1451-1466.
- [17] 张曦, 库少平. 基于生成对抗网络的人脸超分辨率重建方法[J]. *吉林大学学报(工学版)*, 2025, 55(1): 333-338.
- [18] Sloboda T, Hudec L, Benešová W. xAI-CycleGAN, a cycle-consistent generative assistive network[C]//*Computer Vision Systems. ICVS 2023*. Cham: Springer, 2023: 403-411.
- [19] Woo S, Debnath S, Hu R, et al. ConvNeXt V2: Co-designing and scaling convnets with masked autoencoders [EB/OL]. (2023-01-02) [2025-07-15]. <https://arxiv.org/abs/2301.00808>.
- [20] Jabeen S, Li X, Amin M S, et al. A review on methods and applications in multimodal deep learning [J]. *ACM Transactions on Multimedia Computing, Communications and Applications*, 2023, 19(2s): 1-41.
- [21] Li H, Wang L, Liu J. A review of deep learning-based image style transfer research[J]. *The Imaging Science Journal*, 2025, 73(4): 504-526.
- [22] Choi Y, Uh Y, Yoo J, et al. StarGAN v2: Diverse image synthesis for multiple domains [C] // *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. June 13-19, 2020. Seattle, WA, USA. IEEE, 2020: 8185-8194.

(责任编辑:康 锋)