



基于生成对抗网络与稳定扩散模型的花卉丝巾图案生成方法

王罕仁^a, 张华熊^b

(浙江理工大学, a. 信息科学与工程学院; b. 计算机科学与技术学院, 杭州 310018)

摘要: 以花卉丝巾图案为研究对象, 结合生成对抗网络与稳定扩散模型, 提出了一种双阶段模型协同生成丝巾图案的快速方法。首先, 构建了基于 SDXL 模型的丝巾图案扩增 workflow, 通过丝巾图案收集、预处理和扩增, 构建了花卉丝巾图案数据集; 然后, 在丝巾图案生成的第一阶段, 将注意力机制融合到 StyleGAN 框架中, 通过集成自注意力与边框注意力机制对现有生成对抗网络进行改进, 构建了 SAB-StyleGAN 模型, 生成基模花卉丝巾图案; 最后, 在丝巾图案生成的第二阶段, 构建了基于 SDXL 模型的图生图 workflow, 将稳定扩散模型的细节刻画能力嫁接至生成对抗网络, 生成兼具可控性与精细度的精绘花卉丝巾图案。实验结果表明, 该方法生成的精绘花卉丝巾图案清晰, FID 值低至 41.25, 与真实样本高度接近。该方法为丝巾图案的快速生成提供了一种高效解决方案, 有助于企业降低设计成本、提升生产效率, 推动时尚产业的数字化转型。

关键词: 丝巾图案; 图案生成方法; 生成对抗网络; 稳定扩散模型; 图生图; 数据扩增

中图分类号: TS106; TP18

文献标志码: A

文章编号: 1673-3851(2025)07-0556-15

引用格式: 王罕仁, 张华熊. 基于生成对抗网络与稳定扩散模型的花卉丝巾图案生成方法[J]. 浙江理工大学学报(自然科学), 2025, 53(4): 556-570.

Reference Format: WANG Hanren, ZHANG Huaxiong. A generative method for floral scarf patterns using GANs and stable diffusion models [J]. Journal of Zhejiang Sci-Tech University, 2025, 53(4): 556-570.

A generative method for floral scarf patterns using GANs and stable diffusion models

WANG Hanren^a, ZHANG Huaxiong^b

(a. School of Information Science and Engineering; b. School of Computer Science and Technology, Zhejiang Sci-Tech University, Hangzhou 310018, China)

Abstract: With floral scarf patterns as the research objects, this study proposed a dual-stage collaborative generation method combining generative adversarial networks (GANs) and stable diffusion models for rapid scarf pattern generation. First, we constructed an SDXL model-based scarf pattern augmentation workflow, establishing a floral scarf pattern dataset through systematic pattern collection, preprocessing, and data augmentation. Subsequently, in the first stage of pattern generation, we improved conventional GANs by integrating both self-attention and border-attention mechanisms into the StyleGAN framework, developing the SAB-StyleGAN model to generate base floral scarf patterns. Finally, in the second stage of pattern generation, we built an image-to-image workflow based on the SDXL model, effectively grafting the detailed rendering capabilities of stable diffusion models onto GANs to produce refined floral scarf patterns with enhanced controllability and precision. Experimental results demonstrated that the generated refined floral scarf patterns exhibited superior clarity, achieving an FID value as low as 41.25, which closely resembled authentic designer samples. This method provides an efficient solution for

收稿日期: 2025-03-03 网络出版日期: 2025-04-29

基金项目: 浙江省“尖兵”“领雁”研发攻关计划项目(2022C01220)

作者简介: 王罕仁(2000—), 男, 江西南昌人, 硕士研究生, 主要从事图像处理方面的研究。

通信作者: 张华熊, E-mail: zhxhz@zstu.edu.cn

rapid scarf pattern generation, significantly reducing enterprise design costs, enhancing production efficiency, and advancing digital transformation in the fashion industry.

Key words: silk scarf pattern; pattern generation method; generative adversarial networks (GANs); stable diffusion models; image-to-image translation; data augmentation

0 引 言

丝巾作为一种经典的配饰,在时尚界占据着重要地位。近年来,随着全球时尚产业的快速发展和用户个性化需求的增加,丝巾图案的设计效率和质量已成为影响产品竞争力的重要因素。然而,当前企业在设计丝巾图案时主要采用人工方式,图案质量严重依赖设计师的经验和创意,而且设计效率低,短时间内难以推出符合流行趋势的产品,满足快速变化的市场;另外,传统设计方法人力成本高,导致企业难以在激烈的市场竞争中维持价格优势,进而影响其市场竞争力。因此,亟需一种丝巾图案的快速设计方法,而基于计算机图像处理的图案生成方法为此问题提供了一种解决思路。

目前常用的图案生成方法主要分为 2 类,一类是基于生成对抗网络(Generative adversarial networks, GANs)的方法,另一类是基于扩散模型(Diffusion models)的方法。在基于生成对抗网络的图案生成方法中,Radford 等^[1]提出了 DCGAN(Deep convolutional generative adversarial networks)模型,通过卷积层结构可以生成稳定图像,但图案质量受到训练数据规模的限制,数据不足时易导致模式崩溃;任雨佳等^[2]也提出了一种基于 DCGAN 的服装款式设计方法,在训练数据不足时生成图案纹理重复混乱。Arjovsky 等^[3]提出了 WGAN(Wasserstein generative adversarial network)模型,该模型可以通过 Wasserstein 距离提升训练稳定性,但仍需充足数据支撑。田乐等^[4]发现,纺织图案具有复杂结构与高频细节,对模型和数据集的要求更为严苛,小数据集易引发特征学习失效和模式崩溃。Wu 等^[5]提出了基于生成对抗网络的 ClothGAN 框架,采用这种方法生成的服装图案即使在充足数据下仍会出现刺绣纹理模糊甚至断裂,难以生成高分辨率细节。刘康等^[6]在基于生成对抗网络的中式婚服设计方法中,采用高质量多标签成对数据对模型进行训练,发现生成对抗网络存在缺陷,生成的图案有边缘锯齿化和模糊等问题。田乐等^[4]通过理论分析发现,生成对抗网络的对抗训练机制本质是对数据分布的低维流形拟合,其在复杂纹理与高分辨率需求下存在

理论局限。综上所述,生成对抗网络在织物图案数据规模有限时,难以适应图案的多样性和复杂性,会出现模式崩溃的问题;即使数据充足,生成对抗网络的架构缺陷仍制约其细节生成能力,难以满足织物图案的高质量要求。

Sohl-Dickstein 等^[7]提出了扩散模型,后经 Ho 等^[8]改进为去噪扩散概率模型(Denoising diffusion probabilistic models)。在基于扩散模型的图案生成方法中,Rombach 等^[9]提出了一种潜空间扩散模型(Latent diffusion model),该模型在降低计算复杂度的同时支持高分辨率图像生成,但其依赖文本隐式引导,易导致生成图案的元素布局和风格随机化。SD(Stable diffusion)及 SDXL(Stable diffusion XL)模型^[10]在图像分辨率与细节质量上取得了进步,但文本描述的模糊性问题仍然存在,难以精准生成符合特定需求的图像。针对此问题,研究者提出通过 Dreambooth^[11]或 LoRA^[12]微调稳定扩散模型参数的方案。然而,此类方法微调的模型参数主要作用于通道注意力层以适配艺术风格^[13-14],却未有效作用于控制空间关系的空间注意力层,导致生成图案存在元素位置偏移、比例失调等结构缺陷。冉二飞等^[15]提出了一种基于稳定扩散模型的蓝印花布单纹样生成方法,结合 ControlNet 控制网络对图案结构进行约束,但该方法依赖人工绘制深度图作为输入,单纯使用微调模型生成的图案仍存在色彩边界模糊与美学表现力不足的问题。综上所述,基于稳定扩散模型的文生图模式因文本引导的模糊性,难以精准控制图案元素的空间布局与风格;而微调技术主要适配于特定艺术风格,并不适配织物图案。

为解决上述问题,本文以花卉丝巾图案为研究对象,结合生成对抗网络与稳定扩散模型,提出了一种双阶段模型协同生成丝巾图案的快速方法。首先,构建丝巾图案扩增 workflow,聚焦丝巾图案中的花卉主题,扩增数据构建花卉丝巾图案数据集。在此基础上,设计双阶段模型协同生成框架。第一阶段,生成具有准确结构和风格的基模花卉丝巾图案;第二阶段,构建图生图 workflow,生成兼顾结构风格准确性与细节清晰度的精绘花卉丝巾图案。在本文中,

基模花卉丝巾图案是指由生成对抗网络生成的图案,其特点为结构准确、清晰度较低、细节不够丰富完整;而精绘花卉丝巾图案则是指稳定扩散模型生成的图案,具有高清晰度和丰富的细节。该方法整合了两类模型优势,避免了单一模型的局限,为丝巾图案的快速生成提供了一种高效解决方案;同时,该方法有助于降低企业设计成本、提高生产效率,为时尚产业的数字化转型提供技术支持。

1 方法设计

1.1 方法基本流程

本文提出的丝巾图案生成方法的基本流程如图1所示,主要包括数据集构建、基模花卉丝巾图案

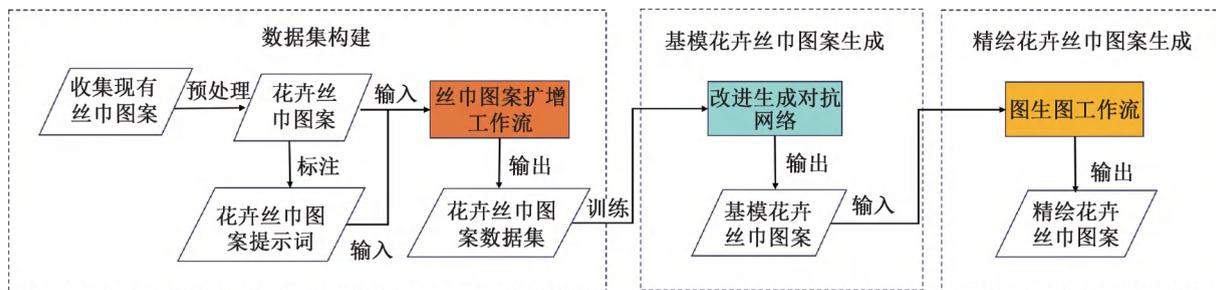


图1 本文方法的基本流程

1.2 数据集构建

1.2.1 丝巾图案收集

基于京东、天猫、淘宝三大电商平台公开的丝巾品类销售榜单,整合形成涵盖绣娘、万事利、上海故事、宝石蝶等十大头部品牌的样本池。通过平台内置搜索引擎,以“丝巾”“方巾”“围巾”为关键词,对这些品牌的官方店铺进行全店铺商品检索,系统收集丝巾图案。为确保数据质量,对收集的丝巾图案进行人工筛选,剔除主题不符样本,得到一批花卉丝巾图案。

1.2.2 丝巾图案预处理

在花卉丝巾图案收集完成后,本文依据生产工艺对图案样本的严格要求,通过人工筛选剔除不符合要求的样本,以保证丝巾图案正面平铺、画面平整、无遮挡物、无折叠痕迹及无环境阴影干扰,满足

生成、精绘花卉丝巾图案生成等3个部分。首先,通过丝巾图案扩增 workflow 构建数据集,解决丝巾数据不足的问题。收集现有丝巾图案,并进行标准化预处理,得到一组花卉丝巾图案,并以此作为基准数据集;根据该基准数据集,进行图案标注,得到花卉丝巾图案提示词;构建基于 SDXL 模型的丝巾图案扩增 workflow;将基准数据集和提示词一起输入图案扩增 workflow,得到花卉丝巾图案数据集。然后,使用花卉丝巾图案数据集,训练改进的生成对抗网络 SAB-StyleGAN,得到具有准确构图和风格的基模花卉丝巾图案。最后,构建基于 SDXL 模型的图生图 workflow,并使用该 workflow 对基模花卉丝巾图案进行图生图处理,得到精绘花卉丝巾图案。

生产工艺要求。

为进一步提升数据集的训练适用性,为后续数据增强与模型训练奠定基础,本文对图像进行标准化处理,具体步骤为:对色彩空间进行标准化转换,消除采集差异;统一尺寸,采用自适应插值算法将所有图像调整为同一像素分辨率,保持图案比例一致。

1.2.3 花卉丝巾图案扩增

1.2.3.1 花卉丝巾图案扩增提示词

选择花卉丝巾图案作为基准,利用图像标签器 WD14 对花卉丝巾图案进行标注,获取标注结果后,提取描述丝巾内容的高频提示词。将这些提示词作为花卉丝巾图案的关键提示词,并结合通用的质量正向提示词和质量反向提示词,得到完整的花卉丝巾图案提示词,见表1。

表1 完整的花卉丝巾图案提示词表

提示词类型	提示词
花卉丝巾提示词	flowers, leaf, border, simple background, no human, silk
质量正向提示词	masterpiece, best quality, ultra_detailed, highres, absurdres
质量反向提示词	worst quality, low quality, ugly, poorly drawn hands, poorly drawn feet, poorly drawn face, out of frame, mutation, mutated, extra limbs, extra legs, extra arms, disfigured, deformed, cross-eye, blurry, bad art, bad anatomy, blurred, text, watermark

1.2.3.2 丝巾图案扩增 workflow 构建

SDXL 模型是一个先进的两阶段级联扩散模

型,包含基础模型和精细化模型两个部分。基础模型与基础 Stable diffusion 模型类似,具有文本到图

像(txt2img)、图像到图像(img2img)和图像修复(inpainting)等功能。精细化模型具有优化图像细节的功能。

本文以 SDXL 模型为基础构建丝巾图案扩增 workflow, SDXL 模型的文本生成图像流程如图 2 所示。SDXL 模型采用 CLIP 文本编码器(CLIP text encoder),将文本(Prompt)转化为文本特征,这个过

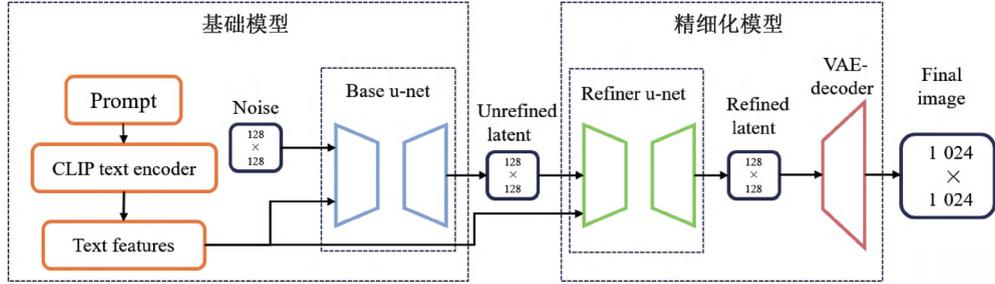


图 2 SDXL 模型的文生图流程

根据花卉主题丝巾的关键特征提示词和质量提示词,通过 SDXL 模型的文生图技术,已基本完成丝巾图案的扩增任务。单纯的文生图存在生成随机性过强、图案美观度和逻辑性较差等问题。为了解决这些问题,本文在 SDXL 模型文生图的基础上,采用 ComfyUI 构建了一个结合 IP-Adapter-Plus 节点的专用 workflow 进行丝巾图案扩增。

基于 ComfyUI 的丝巾图案扩增 workflow 如图 3

程也称文本嵌入(Text embedding),以引导 Base u-net 网络在潜在空间逐步去除噪声(Noise)。初步去噪后,将潜在特征(Unrefined latent)输入精细化模型,使用 Refiner u-net 网络进行精细化去噪,以提升图像质量。最终优化后的精细化特征(Refined latent)通过 VAE 解码器(VAE-decoder)还原至像素空间,得到最终图像(Final image)。

所示。该 workflow 以花卉丝巾图案(Floral scarf pattern)、花卉纹样图案(Floral motif pattern)、正向提示词(Positive prompt)和反向提示词(Negative prompt)作为输入,生成新的花卉丝巾图案(New floral scarf pattern)。本文选用花卉丝巾图案作为构图参考,花卉纹样图案作为内容素材和风格参考,花卉纹样图案来自网络收集,共 1000 张。

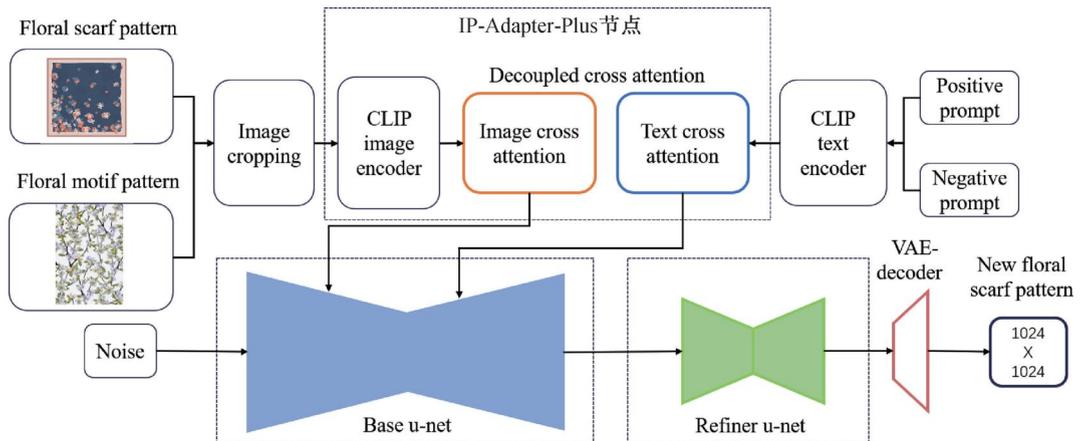


图 3 丝巾图案扩增 workflow 示意图

IP-Adapter-Plus 节点是一个专为 SDXL 图像生成流程设计的 ComfyUI 组件,它能够在该过程中整合额外的图像特征作为参考,以提升图像生成的多样性和精准度。该节点通过预训练的 CLIP 图像编码器(CLIP image encoder)提取输入图像特征,采用解耦交叉注意力机制(Decoupled cross attention)结合文本特征和图像特征,构建文本与图像双通道独立处理的注意力层结构。解耦交叉注意力可用式(1)表示:

$$\mathbf{Z} = \text{Softmax}\left(\frac{\mathbf{Q}\mathbf{K}^T}{\sqrt{d}}\right)\mathbf{V} + \text{Softmax}\left(\frac{\mathbf{Q}'(\mathbf{K}')^T}{\sqrt{d}}\right)\mathbf{V}' \quad (1)$$

其中: \mathbf{Z} 表示解耦交叉注意力的输出,是图像交叉注意力的输出与文本交叉注意力的输出相加; \mathbf{Q} 、 \mathbf{K} 、 \mathbf{V} 分别表示来自文本特征的查询向量、键向量和值向量; \mathbf{Q}' 、 \mathbf{K}' 、 \mathbf{V}' 分别表示来自图像特征的查询向量、键向量和值向量; d 表示键向量的维度。

文本特征维持原始文本交叉注意力层(Text

cross attention)的参数与计算路径,而图像特征则通过解耦交叉注意力机制新构建的图像交叉注意力层(Image cross attention)进行处理。该技术将解耦的图像交叉注意力层和文本交叉注意力层一起动态插入 SDXL 模型 Base u-net 网络,形成双流注意力协同工作机制。这种架构使得 SDXL 模型在生成过程中能够同步解析文本提示词的语义约束与参考图像的视觉特征。

图 3 所示的丝巾图案扩增 workflow 具体工作流程如下:一方面将输入的正向提示词和反向提示词通过 CLIP 文本编码器(CLIP text encoder)转换为文本特征;另一方面将输入的花卉丝巾图案和花卉纹样图案经过图像裁剪(Image cropping)后,输入 IP-Adapter-Plus 节点。使用 IP-Adapter-Plus 节点,提取输入图像的图像特征,借助解耦的交叉注意力机制,将图像特征的交叉注意力层和文本特征的交叉注意力层插入 Base u-net 网络中。使用双重特征共同引导 Base u-net 网络,在潜在空间对噪声初步去噪。在完成初步去噪之后,由 workflow 中专门用于去噪的 Refiner u-net 网络进一步处理,经过 VAE 解码器还原至像素空间,得到最终的丝巾图案。

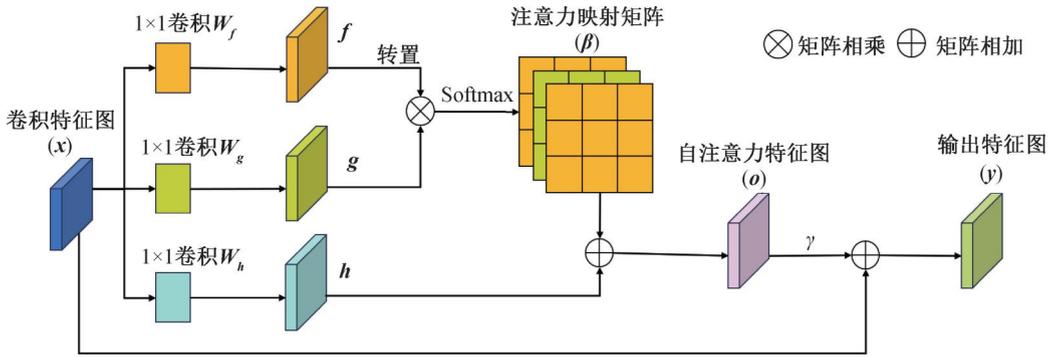


图 4 自注意力机制示意图

边框注意力是一种本文自定义的简化空间注意力机制,用于提升图像边框部分的生成质量。边框注意力与空间注意力的主要区别在于权重的生成方式不同。空间注意力的权重是动态生成的,通常通过神经网络学习得到,能够根据输入图像的不同自动调整。而边框注意力的权重是预先定义的固定权重,不随输入图像的变化而变化,固定权重将注意力集中在图像的边缘部分,从而增强丝巾边框的生成质量。

边框注意力的工作原理示意图如图 5 所示。首先,从神经网络的某一层中提取卷积特征图 x ,接着定义一个边框注意力映射 W_b ,该边框注意力映射

1.3 基模花卉丝巾图案生成

为了生成结构完整和风格准确的丝巾图案,本文选取 StyleGAN (Style-based generative adversarial network)作为基础模型,对此模型进行改进,集成了自注意力机制和边框注意力机制。这一改进使模型能更专注于图案的关键细节和图案边界,显著提高了基模花卉丝巾图案的结构完整度。

1.3.1 注意力机制简介

自注意力机制是一种提升生成模型性能的技术,它能够增强模型对图像细节和结构特征的捕捉能力,从而提升生成图像的清晰度和准确性。其工作原理如图 4 所示。该机制通过 3 个并行的 1×1 卷积层 W_f 、 W_g 和 W_h ,将输入特征图 x 映射为三个子空间:查询向量 f 、键向量 g 和值向量 h 。其中 f 与转置后的 g 进行矩阵相乘,经 Softmax 归一化生成注意力特征图 β ,该权重揭示了空间位置间的依赖关系。然后通过 β 与 h 的加权融合实现特征重构,得到自注意力特征 o 。自注意力特征 o 与卷积特征图 x 引入可学习参数 γ 加权相加,得到最终输出特征图 y 。这种设计使模型能突破卷积核的局部感受野限制,从而捕捉像素间的结构关联,提升生成图像的细节连贯性与几何合理性。

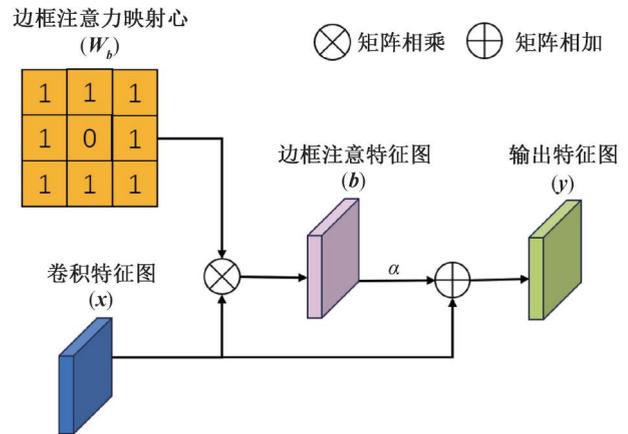


图 5 边框注意力机制示意图

在中心 80% 区域的值为 0, 在边缘 20% 区域的值为 1。此映射与原始卷积特征图 x 相乘, 得到聚焦于边框部分的边框注意力特征图 b 。边框注意力特征图 b 乘以一个可学习的缩放参数 α , 加上原始卷积特征图 x , 得到最终输出特征图 y 。缩放参数 α 用于调节边框注意力特征对最终输出结果的影响程度。

1.3.2 SAB-StyleGAN 构建

本文将注意力机制融合到 StyleGAN 框架中, 提出了一种基于注意力的 StyleGAN 模型, 即 SAB-StyleGAN, 其结构图如图 6 所示。

典型的 StyleGAN 的生成流程以解耦式生成架构为核心, 通过潜变量映射与渐进式合成实现高质量图像生成。其工作流程为: 首先, 映射网络接收来自潜空间 Z 的随机向量 z 作为输入, 经过像素级归一化 (Normalize) 消除幅值差异后, 通过 8 层全连接网络 FC 构建的非线性网络进行非线性变换, 将其映射至中间潜空间 W , 得到中间潜变量 w 。随后, w 通过仿射变换 A 转化为风格向量, 用于控制后续生成网络的风格属性。生成网络从固定的 $4 \times 4 \times 512$ 维常数张量 (Const $4 \times 4 \times 512$) 开始, 在每一层级执行以下操作: 向特征图添加经广播运算 B , 调整维度的高斯噪声 (Noise) 以增强细节多样性; 通过自适应实例规范化 (AdaIN) 将风格向量注入特征图, 利用 3×3 卷积 (Conv 3×3) 进行特征变换, 并循环执行一次“噪声注入-AdaIN 风格控制”的双阶段调节。随着网络深度增加, 通过上采样 (Upsample) 逐步提升分辨率 ($4 \times 4 \rightarrow 8 \times 8 \rightarrow \dots \rightarrow 256 \times 256$), 最终输出高清合成图像。与此同时, 判别器采用反向对称结

构, 通过堆叠卷积块 (ConvBlock) 对生成虚假图像与真实图像 (real/fake image) 进行多尺度特征提取, 逐步下采样至 1×1 , 然后经全连接层 FC 输出二值判别结果 D 。整个模型通过对抗训练机制迭代优化: 生成器力图合成足以欺骗判别器的逼真丝巾图案, 而判别器持续提升对真伪图像的鉴别能力, 直至两者达到动态平衡 (纳什均衡)。此时模型的生成器可输出与真实数据分布高度一致的图案。该流程通过分离风格控制与细节生成, 实现了对丝巾花型结构稳定性与纹理复杂度的精准调控。

本文为增强 StyleGAN 在图像生成中的细节表现力和风格精确度, 在生成网络的多个维度的最后一层自适应实例规范化后引入了自注意力机制, 分别是 4×4 、 32×32 和 128×128 这几个特征维度。该机制通过在不同层级集成自注意力模块 (self-attention), 显著提升了模型对图像局部特征的敏感性和捕捉能力。自注意力机制通过特征图间的点积运算和随后的 Softmax 归一化处理, 帮助模型深入理解特征间的复杂关系, 有效提高生成图像的整体质量。

在丝巾设计中, 边框不仅是装饰元素, 更是强化图案视觉效果和风格表达的关键。丝巾边框的精细处理对图像的整体自然度和视觉完整性至关重要。在生成网络的最后一层自适应实例规范化之后和判别器的第一层卷积块之后, 本文分别集成了边框注意力 (border-attention) 模块。该模块引导模型关注于图案边界部分, 针对图像边界进行优化。使用前文构建的花卉丝巾图案数据集对 SAB-StyleGAN 进行训练, 得到基模花卉丝巾图案。

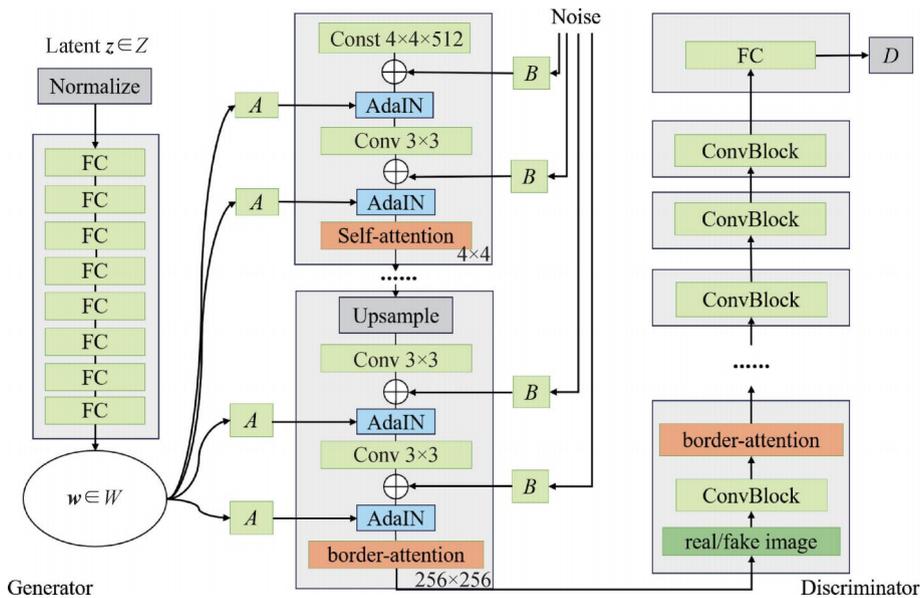


图 6 SAB-StyleGAN 网络结构示意图

1.4 精绘花卉丝巾图案生成

改进后的 StyleGAN 模型成功生成了基模花卉丝巾图案,该基模花卉丝巾图案虽然已有丝巾图案的基本结构和风格,但清晰度不足,细节也有待完善。为了进一步提升图案的清晰度和细节完整度,本文提出了一个结合 IP-Adapter-Plus 节点的图生图 workflows,将基模花卉丝巾图案转化为高质量的精绘花卉丝巾图案, workflows 示意图如图 7 所示。

图生图 workflows 从输入一个基模花卉丝巾图案 (Base-generated floral scarf pattern) 开始,一方面使用 VAE 编码器 (VAE-encoder) 对输入的基模花

卉丝巾图案进行编码,将其转换为潜在空间表示,并添加噪声,得到噪声图像。另一方面使用 IP-Adapter-Plus 节点处理基模花卉丝巾图案,经由 CLIP 图像编码器处理,得到基模花卉丝巾图案的图像特征;使用 CLIP 文本编码器处理正向提示词和反向提示词得到文本特征;再由 IP-Adapter-Plus 节点中的解耦交叉注意力结合文本特征和图像特征,共同引导 Base u-net 网络,对噪声图像进行初步去噪。最后使用 Refiner u-net 网络进一步去噪,经过 VAE 解码器还原至像素空间,得到精绘花卉丝巾图案 (Refined-generated floral scarf pattern)。

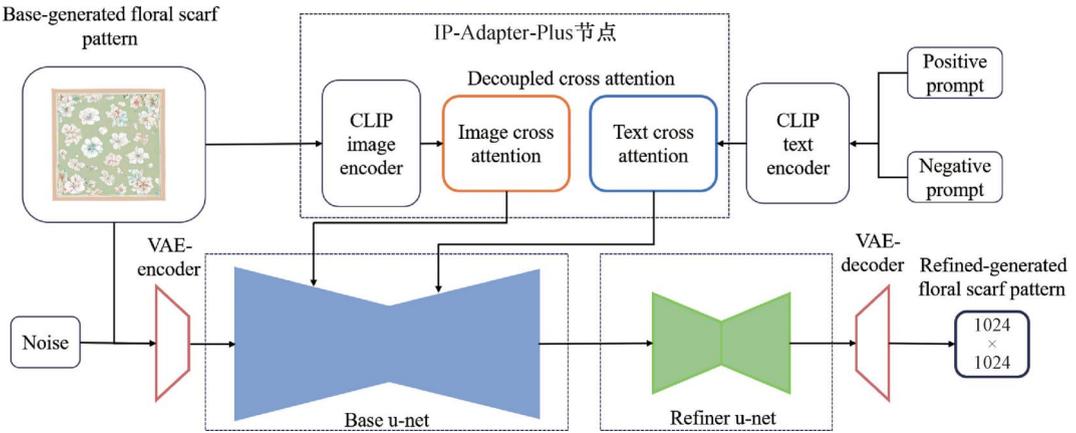


图7 图生图 workflows 示意图

2 实验结果

本文搭建了实验平台,硬件采用配备高性能运算能力的深度学习工作站,具有 128 GB 内存和 NVIDIA RTX 3090 Ti 24 GB GPU;软件采用 Windows 10 64 位操作系统,PyTorch 1.8.1 深度学习框架,编程语言为 Python 3.8,计算平台为 CUDA 11.3。

2.1 评估指标

本文采用 FID、SSIM、PSNR、MBM 和 LPIPS 等 5 个指标综合评估生成图像的质量。

SSIM (Structural similarity index measure, 结构相似性):用于衡量生成图像与参考图的局部相似性,其值范围在 0~1 之间,越接近 1 表示相似性越高。其计算公式如下:

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

其中: μ_x 和 μ_y 表示图像 x 和图像 y 的像素亮度均值; σ_x^2 和 σ_y^2 表示图像 x 和 y 的像素值方差; σ_{xy} 表示图像 x 和 y 的协方差; c_1 和 c_2 是为了避免除零错

误而加的小常数。

PSNR (Peak signal-to-noise ratio, 峰值信噪比):用于比较生成图像与参考图的失真程度,值越高表明表示图像质量越好,失真越小。具体计算公式如下:

$$PSNR = 20 \cdot \log_{10} \left(\frac{MAX}{MSE} \right) \quad (3)$$

其中: MAX 表示图像的最大像素值 (对于 8 位图像,通常是 255); MSE 表示两幅图像的均方误差。

MBM (Mean blur metric, 平均模糊度):用于评估图像清晰度,数值越低表示图像边缘和细节保留越好。具体计算公式如下:

$$MBM = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N |\nabla \mathbf{I}(i, j)| \quad (4)$$

其中: \mathbf{I} 代表图像; $\nabla \mathbf{I}(i, j)$ 表示图像在位置 (i, j) 的梯度; M 和 N 表示图像的行数和列数。

FID (Fréchet inception distance, 弗雷谢特初始距离):用于衡量生成图案与训练集的全局相似度。FID 值越低,表示两个图像分布越接近,图像间的相似度越高,生成图像的质量也越佳。具体计算公式如下:

$$FID = \|\mu_r - \mu_g\|^2 + \text{tr}[\Sigma_r + \Sigma_g - 2(\Sigma_r \Sigma_g)^{1/2}] \quad (5)$$

其中: r 代表真实图像; g 代表生成图像; μ_r 和 μ_g 代表图像 r 和 g 的特征向量的均值; Σ_r 和 Σ_g 代表图像 r 和 g 的特征向量的协方差矩阵; tr 表示矩阵的迹。

LPIPS (Learned perceptual image patch similarity, 学习型感知图像块相似度): 用于评估图像之间的感知差异, 结果在 $0 \sim 1$ 之间。其中 0 表示两幅图像完全一致, 1 表示两幅图像在语义感知层面完全不同。具体计算公式如下:

$$LPIPS(x, y) =$$

$$\sum_l \frac{1}{H_l W_l} \sum_{h=1}^{H_l} \sum_{w=1}^{W_l} \|w_l \odot (\phi_l(x) - \phi_l(y))\|_2^2 \quad (6)$$

其中: x 和 y 分别代表输入的两幅图像, l 代表预训练网络的特征层索引, $\phi_l(x)$ 代表图像 x 第 l 层的特征, $\phi_l(y)$ 代表图像 y 第 l 层的特征, w_l 代表第 l 层通道的权重, H_l 和 W_l 分别代表第 l 层特征的高度与宽度。

2.2 数据集构建结果

2.2.1 丝巾图案数据预处理结果

经过预处理流程, 最终得到一组高质量花卉丝巾图案, 共 176 张, 样例如图 8 所示。由图 8 可知: 真实花卉丝巾图案具有清晰的方形边框, 花卉元素均匀的分布在丝巾图案。该组花卉丝巾图案保证了视觉完整性和主题一致性, 为后续的图案扩增任务提供了优质基准数据。



图 8 预处理后的花卉丝巾图案样例

2.2.2 花卉丝巾图案扩增实验结果

经由人工操作该丝巾图案扩增 workflow, 选定花卉丝巾图案, 输入花卉纹样图案和提示词进行图案扩增。人工调节两类输入图案的参考权重配比, 完成花卉主题丝巾图案扩增, 去除其中不符合丝巾图案要求的部分。该扩增方法不仅丰富了图案的多样性, 也保证了生成丝巾图案的主题和构图与真实图

像一致。本文生成的丝巾图案在视觉上与市场的热销丝巾构图保持一致, 同时也融入了新的设计元素。扩增的丝巾图案样例如图 9 所示, 其中扩增花卉丝巾图案的方形边框清晰可见, 花卉元素分布均匀, 整体色彩搭配协调, 艺术美观度高。

扩增后的图案与扩增前的花卉主题丝巾图案组合, 得到花卉丝巾图案数据集, 通过这种方法, 构建了包含 1000 张图案的花卉丝巾图案数据集。



图 9 扩增的花卉丝巾图案样例

本文通过 LPIPS 指标来评估扩增数据集的多样性, 通过 FID 和 SSIM 指标来评估扩增数据集与原始数据的分布差异。其中 SSIM 值计算时以 workflow 输入的原始花卉丝巾图案作为基准, 以 workflow 输出的扩增图案作为比较对象。花卉丝巾图案扩增量化分析结果如表 2 所示, 从表 2 可以看出: 扩增数据集的 LPIPS 值为 0.4126, 属于中等感知差异范围, 表明数据集中的图像在人类视觉系统中既有可识别的差异, 又保留了一定的相似性; 对于花卉丝巾图案来说, 适度的形态与色彩变化避免了设计同质化, 严谨的构图规则有利于产品商业落地。扩增数据集与原始数据的 FID 值为 75.22, SSIM 值为 0.4396, 表明扩增数据集与原始数据有一定的分布相似度, 且结构框架基本相似。综上可知, 扩增数据集与原始数据既有相似度也有多样性, 有利于后续生成高质量丝巾图案。

表 2 花卉丝巾图案扩增量化分析

评价指标	指标值
FID	75.22
LPIPS	0.4126
SSIM	0.4396

2.3 基模花卉丝巾图案生成结果

2.3.1 未扩增花卉丝巾图案数据集结果

使用未扩增的花卉丝巾图案数据集训练生成对

抗网络 StyleGAN, 训练完成后, 生成基模花卉丝巾图案, 结果如图 10 所示。图 10 显示: 176 张的小规模丝巾图案训练集难以有效支持生成对抗网络的稳定训练, 具体表现为生成的基模花卉丝巾图案质量有重大缺陷, 出现模式崩溃; 图案结构和色彩严重失真, 模糊且缺失细节。



图 10 未扩增花卉丝巾图案数据集实验结果样例

2.3.2 扩增花卉丝巾图案数据集实验结果

在扩增后得到的 1000 张花卉丝巾图案数据集的基础上, 分别采用 DCGAN、WGAN、StyleGAN、StyleGAN-XL、StyleGAN&self-Attention、StyleGAN&border-Attention、SAB-StyleGAN 模型进行训练。DCGAN、WGAN、StyleGAN 和 StyleGAN-XL 是经典的生成对抗网络模型, StyleGAN&self-Attention 是在 StyleGAN 模型的基础上添加自注意力机制后得到的模型, StyleGAN&border-Attention 是在 StyleGAN 模型的基础上添加边框注意力机制后得到的模型, SAB-StyleGAN 是在 StyleGAN 模型基础上添加自注意力机制和边框注

意力机制后得到的模型。

在模型训练过程中, 采用批量规模 (Batch size) 为 4 的样本输入配置, 持续进行 1000 个完整训练周期 (Epoch)。优化器选用 Adam 方法, 其动量参数设置为 $\beta_1 = 0, \beta_2 = 0.99$, 初始学习率 (Learning rate) 固定为 0.0025, 并引入零动量自适应优化策略 (Zero-momentum adaptive optimization) 动态调节学习率。训练完成后, 各自生成 1000 张基模花卉丝巾图案, 用于后续对比生成效果和图生图处理。

2.3.2.1 对比实验

花卉丝巾图案的生成效果能直观地反映了不同模型的性能差异。不同模型生成的基模花卉丝巾图案如图 11 所示。由图 11 可见: WGAN 模型和 DCGAN 模型生成的基模花卉丝巾图案存在纹理模糊、色彩不协调结构失真等明显问题; StyleGAN 模型生成的基模花卉丝巾图案较为真实, 但在色彩还原和图像结构上仍有失真, 未能准确捕捉真实图像的细微特征; StyleGAN-XL 模型生成的基模花卉丝巾图案边框连贯清晰, 色彩还原, 但在花卉结构上出现严重扭曲, 花卉边缘出现羽化现象; SAB-StyleGAN 模型 SAB-StyleGAN 模型生成的基模花卉丝巾图案表现最佳, 结构和风格都与真实图像最为接近, 视觉效果自然, 边框连贯清晰。

对比图 11 和图 10 中 StyleGAN 模型生成的花卉丝巾图案可以发现, 扩增数据集后生成质量显著提升, 表明丝巾图案扩增工作流有效。

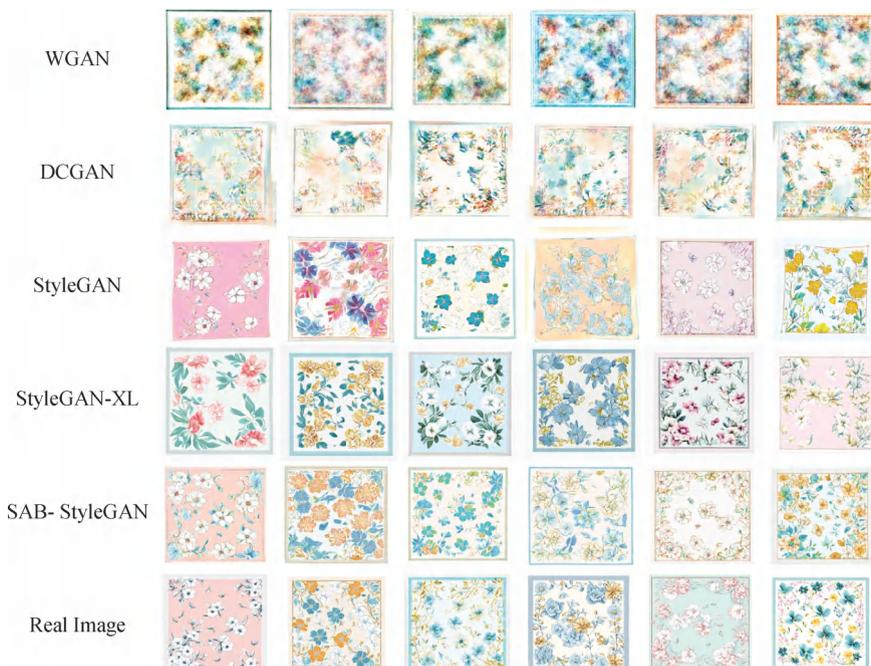


图 11 不同模型生成的基模花卉丝巾图案

表 3 为不同生成模型在生成基模花卉丝巾图案时的 FID 值。DCGAN 的 FID 值为 130.80, WGAN 的 FID 值为 255.56, 表明这两个模型生成图像质量较低; StyleGAN-XL 模型的 FID 值为 61.29, StyleGAN 模型的 FID 值为 62.21, 而 SAB-StyleGAN 模型的 FID 值为 46.71, 相比于基线模型 StyleGAN 降低了 24.91%, 表明图像生成质量提升, 添加双重注意力机制有效。综上所述, SAB-StyleGAN 模型效果最佳。

表 3 不同生成模型 FID 值

评价指标	FID 值
DCGAN	130.80
WGAN	255.56
StyleGAN	62.21
StyleGAN-XL	61.29
SAB-StyleGAN	46.71

2.3.2.2 消融实验

为直观展示双重注意力机制对模型生成效果的影响, 本文对模型改进前后的生成效果进行了可视化对比, 结果如图 12 所示。图 12 显示: StyleGAN 模型生成的花卉丝巾图案在纹理和结构上存在一定失真, 丝巾图案边框扭曲; StyleGAN&self-Attention 模型通过引入自注意力机制强化全局特征建模, 生成的花卉丝巾图案质量显著提升, 结构上接近真实图像, 纹理细腻, 色彩渐变自然, 但在丝巾图案边缘的处理上仍有不足, 丝巾边框凹凸不平且线条模糊断裂; StyleGAN&border-Attention 模型在生成网络和判别器中集成了边框注意力模块, 更加专注于图像的边框区域, 生成的花卉丝巾图案在边缘的处理上有所提升, 边框连贯清晰; SAB-StyleGAN 模型融合双重注意力机制, 生成的花卉丝巾图案表现最佳, 结构和风格都与真实图像最为接近, 视觉效果自然, 纹理细腻, 色彩渐变自然, 边框连贯清晰。



图 12 不同注意力机制增强的 StyleGAN 模型生成的花卉丝巾图案

表 4 为添加双重注意力机制的不同生成模型在生成基模花卉丝巾图案时的 FID 值, StyleGAN 模型的 FID 值为 62.21, StyleGAN&border-attention 模型的 FID 值为 58.37, 相较于基线模型 StyleGAN 降低了 3.84, 表明图像生成质量提升, 添加的边框注意力有效; StyleGAN&self-attention 模型的 FID 值为 49.24, 相较于基线模型 StyleGAN 降低了 12.97, 表明图像生成质量提升, 添加的自注意力模块有效; SAB-StyleGAN 模型的 FID 值为 46.71, 相比于基线模型 StyleGAN 降低了 15.5, 表明图像生成质量提升, 同时添加双重注意力机制有效。

表 4 基于双重注意力机制的不同生成模型 FID 值

评价指标	FID 值
StyleGAN	62.21
StyleGAN&border-attention	58.37
StyleGAN&self-attention	49.24
SAB-StyleGAN	46.71

2.4 精绘花卉丝巾图案生成实验结果

精绘花卉丝巾图案的生成实验结果包括两部分, 分别是客观量化实验和主观评估实验。这两种实验从客观的图像质量和主观的图像美感两方面, 综合评估精绘花卉丝巾图案。

2.4.1 客观评估实验

尽管 SAB-StyleGAN 模型生成的丝巾图案结构和风格都与真实图像接近,但仍有模糊和细节缺失的问题,需要进一步进行图生图处理,优化图案以得到精绘花卉丝巾图案。本文对比了不同方法生成的精绘花卉丝巾图案(见图 13),其中, SAB-StyleGAN、img2img、img2img&IP-Adapter-Plus、Dreambooth 和 SDXL 分别表示 SAB-StyleGAN 生成的基模花卉丝巾图案、基础图生图工作流生成的精绘花卉丝巾图案、集成 IP-Adapter-Plus 节点的图生图工作流所生成的精绘花卉丝巾图案、利用

Dreambooth 技术微调后的 SDXL 模型所生成的精绘花卉丝巾图案,以及 SDXL 模型生成的精绘花卉丝巾图案。

图 14 为两种不同图生图处理工作流的效果对比图。SAB-StyleGAN 生成的基模花卉丝巾图案仅用于对比。Dreambooth 是一种用于 SDXL 模型的微调技术,允许用户通过少量特定主题图像个性化预训练的扩散模型,适用于将自定义元素融入模型中,生成具有特定特征的图像。利用该技术,本文使用花卉丝巾图案数据集对 SDXL 模型进行微调,得到了一个能够快速生成花卉丝巾图案的微调 SDXL 模型。

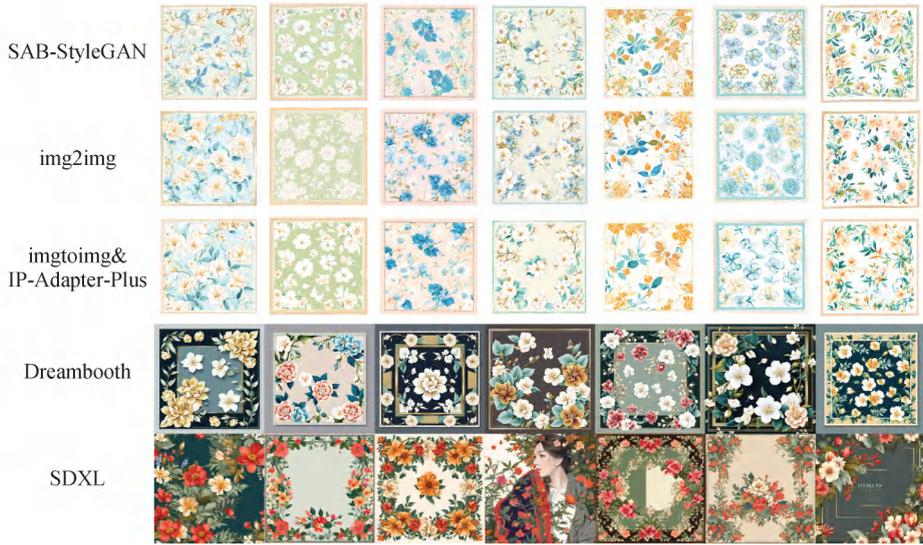


图 13 不同方法生成的精绘花卉丝巾图案



图 14 两种不同图生图工作流效果

由图 13 和图 14 可知:采用 SDXL 生成的精绘花卉丝巾图案色调单一,结构混乱,无明显的边框,且生成图案不稳定,时常出现一些人像和文字;采用 Dreambooth 技术微调的 SDXL 模型生成的精绘花

卉丝巾图案与训练数据集有较大偏差,图案风格过于写实,缺少艺术感和装饰性,色彩搭配冲突,视觉观感单调,边框呈现显著的遮挡性破坏;SAB-StyleGAN 模型生成的基模花卉丝巾图案结构合

理,风格美观,但稍显模糊,细节不足;img2img 工作流在基模花卉丝巾图案的基础上进行图生图处理,生成的精绘花卉丝巾图案清晰度高,但风格和构图与输入图案存在差异,整体视觉观感上杂乱无章;img2img&IP-Adapter-Plus 工作流生成的精绘花卉丝巾图案在风格和构图上更贴近输入图像,该工作流应用 IP-Adapter-Plus 节点的图像风格参考能力,结合图生图本身的色彩、构图参考能力,进行图生图处理,大大提高了基模花卉丝巾图案的清晰度,补全了图案缺失的细节,在色彩、纹理、图案布局及整体视觉效果上展现出更高的一致性,效果最佳。

表 5 展示了不同精绘花卉丝巾图案生成方法的量化指标。其中 SSIM 值和 PSNR 值的计算以 SAB-StyleGAN 模型生成的基模花卉丝巾图案为基准图案,img2img 工作流和 img2img&IP-Adapter-Plus 工作流生成的精绘花卉丝巾图案作为对比图案。其他方法并无对应图案,无法计算 SSIM 值和 PSNR 值。

SDXL 模型的 FID 值为 225.02,表明该方法生成的花卉丝巾图案与训练集花卉丝巾图案存在较大

差异,效果较差。Dreambooth 微调方法的 FID 值达到 168.37,表明该方法生成的花卉丝巾图案与训练集花卉丝巾图案存在较大差异,效果较差。img2img 工作流的 FID 值为 63.14,相比于 SAB-StyleGAN 模型增加了 16.4,表明该工作流生成的精绘花卉丝巾图案质量降低。img2img&IP-Adapter-Plus 工作流结合 IP-Adapter-Plus 节点,参考输入图案风格,FID 值为 41.25,相较于 SAB-StyleGAN 模型降低了 5.5,降幅为 11.69%,表明该工作流生成的精绘花卉丝巾图案与训练集更加接近,图案质量提高。

观察 SSIM 值和 PSNR 值可知,img2img&IP-Adapter-Plus 工作流相比于 img2img 工作流的 SSIM 值提高了 0.0271,PSNR 值提高了 0.38,表明加入 IP-Adapter-Plus 节点的工作流能够更好地保留了输入基模花卉丝巾图案的构图和风格,有更高的视觉相似度。MBM 值表明,两种图生图处理都使得图案清晰度得到显著提升。综合对比,img2img&IP-Adapter-Plus 工作流的效果最佳。

表 5 不同精绘花卉丝巾图案生成方法量化指标

评价指标	FID 值	SSIM 值	PSNR 值	MBM 值
SDXL	225.02			646
Dreambooth	168.37			725
SAB-StyleGAN	46.71			4244
img2img	63.14	0.5104	17.67	570
img2img&IP-Adapter-Plus	41.25	0.5375	18.05	633

2.4.2 主观评估实验

为了从美学表现、满足丝巾图案需求等角度综合评估各种精绘花卉丝巾图案生成方法的性能,本文设计了一份包含 5 个问题的问卷,分别对应前文所分析的 5 种精绘花卉丝巾图案生成方法。该问卷从艺术美观度与丝巾规范度两个维度,对精绘花卉丝巾图案进行主观评估。艺术美观度是指生成的图案在视觉上的美感,给人带来愉悦感和吸引力的程度。丝巾规范度是指生成的图案在构图和边框设计上与真实丝巾图案的符合程度,用于衡量生成图案是否遵循真实丝巾图案的设计规范和特征,高丝巾规范度意味着生成的图案在视觉上更接近真实的丝巾设计,符合实际的使用标准。问卷的评分采用 5 分制,1 分表示最低评价,5 分表示最高评价。此次问卷调查共回收有效问卷 118 份,统计结果如图 15 所示。由图 15 可知,SAB-StyleGAN 模型在艺术美观度和丝巾规范度上表现均衡,两个维度的 4 分和 5 分选择人数占比均超过 80%,显示出

较高的认可度。img2img 工作流虽然 4 分以上的评分占比也较高,但 5 分评价的比例较 SAB-StyleGAN 模型降低,整体认可度不如 SAB-StyleGAN 模型。img2img&IP-Adapter-Plus 工作流表现最优,艺术美观度 71 人给出 5 分,丝巾规范度 75 人选择 5 分,是最受认可的方案。Dreambooth 微调方法评分呈现中间集中趋势,2 分和 3 分的选择人数占比较高,显示出两极分化较弱但整体认可度偏低的特点。SDXL 模型表现最差,两个维度均出现 1 分和 2 分占比过高的问题,明显落后于其他方法。

表 6 展示了不同方法生成的精绘花卉丝巾图案的量化评分,评分由所有问卷结果的平均计算得出。SDXL 模型在艺术美观度和丝巾规范度上的得分较低,表明其生成的图案在艺术美观度和丝巾规范度方面表现不佳。Dreambooth 微调方法在艺术美观度和丝巾规范度上的表现较 SDXL 有所提升,但在所有方法中仍处于较低水平,表明其生成的图案在

艺术美观度和丝巾规范度方面表现不佳。SAB-StyleGAN 模型在艺术美观度和丝巾规范度上均取得了较高的分数,表明其生成的图案在艺术美观度和符合真实丝巾图案构图及边框方面表现优异。img2img 工作流在丝巾规范度上表现良好,但在图生图处理后破坏了 SAB-StyleGAN 模型所生成图案的原有风格,在艺术美观度上略低于 SAB-

StyleGAN。img2img&IP-Adapter-Plus 工作流在艺术美观度和丝巾规范度上均取得了最高的分数,表明其在生成具有高艺术美感、符合规范的精绘花卉丝巾图案方面具有显著优势,能够同时满足艺术美观度和丝巾规范度的要求。综上可知,在这些方法中 img2img&IP-Adapter-Plus 工作流生成的精绘花卉丝巾图案最佳。

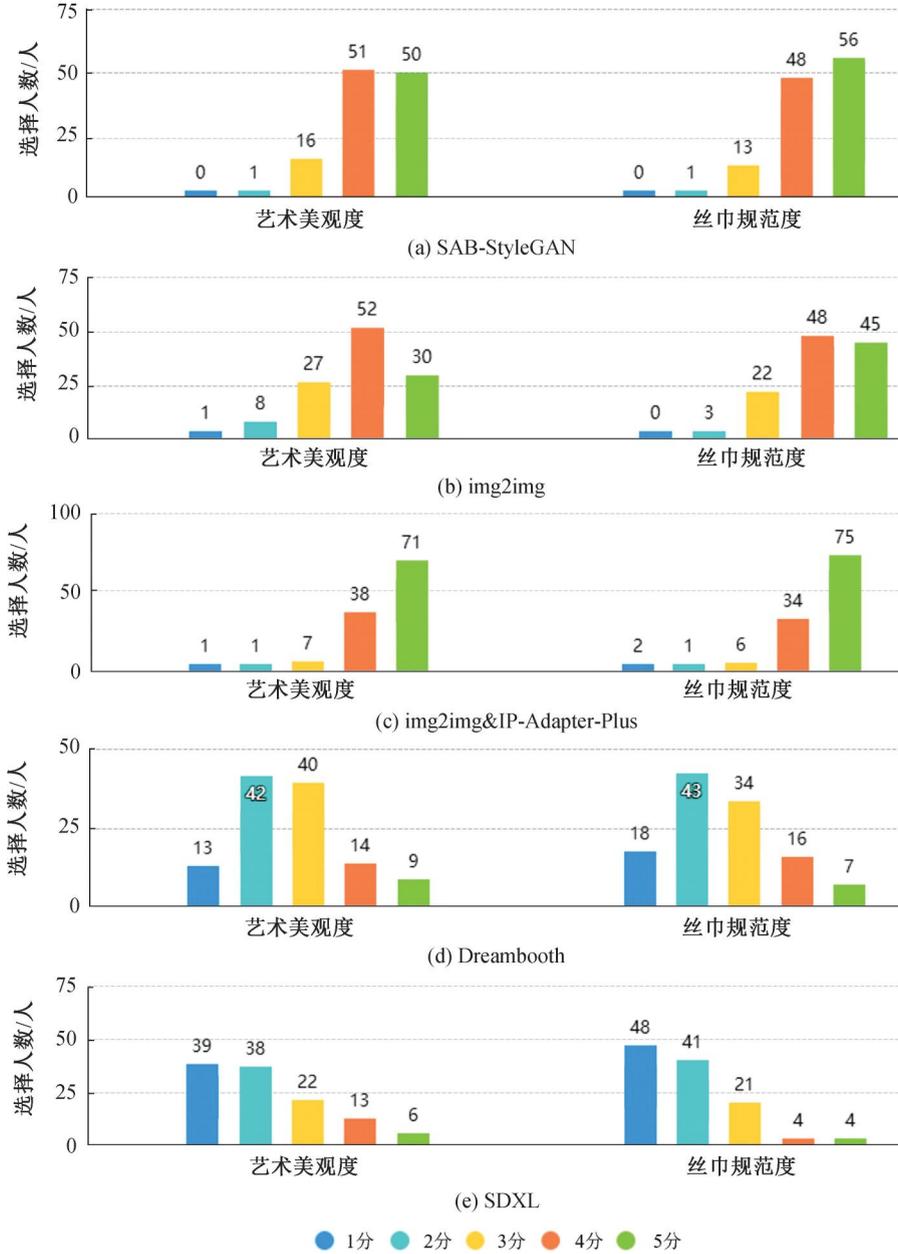


图 15 问卷结果统计

表 6 不同精绘花卉丝巾图案生成方法量化指标

评价指标	艺术美观度评分	丝巾规范度评分
SDXL	2.23	1.94
Dreambooth	2.69	2.58
SAB-StyleGAN	4.27	4.35
img2img	3.86	4.14
img2img&IP-Adapter-Plus	4.50	4.52

2.5 计算效率与硬件成本分析

为了评估本文方法的计算效率和硬件成本,记录了相同硬件环境下不同方法的训练时间、推理时间和显存占用。基模花卉丝巾图案生成方法的计算效率与硬件成本见表 7,精绘花卉丝巾图案生成方法的计算效率与硬件成本见表 8。其中,精绘花卉丝巾图案

生成方法除 Dreambooth 外,均采用预训练的 SDXL 模型,无训练时间。训练时间受到超参数和训练数据的影响,并不完全固定,因此仅作为参考。

表 7 显示:SAB-StyleGAN 模型在训练时间、推理时间、显存占用 3 个方面都比 StyleGAN-XL 模型表现更优异,表明 SAB-StyleGAN 模型在计算效率和硬件成本上均优于 StyleGAN-XL 模型;SAB-StyleGAN 模型的训练时间和推理时间在基准模型 StyleGAN 的基础上略微增加,显存占用不变;训练时间增加 5 h,增加了 11.9%,推理时间增加 0.002 s,增加了 6%,增加幅度较小。综上所述,本文采用的 SAB-StyleGAN 模型在计算效率方面和硬件成本方面符合实际应用要求。

表 7 基模花卉丝巾图案生成方法的计算效率与硬件成本

评价指标	训练时间/h	推理时间/s	显存占用/GB
DCGAN	6	0.012	2.4
WGAN	14	0.015	4.7
StyleGAN	42	0.031	3.0
StyleGAN-XL	49	0.051	3.8
SAB-StyleGAN	47	0.033	3.0

表 8 显示:img2img&IP-Adapter-Plus 工作流与 img2img 工作流相比,推理时间减少 2 s,降幅 20%,显存占用提高 0.8 GB,增加了 4.8%;img2img&IP-Adapter-Plus 工作流与 SDXL 模型和 Dreambooth 微调方法相比,推理时间减少 1 s,降幅 10%,显存占用提高 1 GB,涨幅 6.1%;综上所述可以发现,与其他 3 种方法相比,本文采用的 img2img&IP-Adapter-Plus 工作流推理时间明显降低,占用显存轻微提高,在计算效率方面和硬件成本方面符合实际应用要求。

表 8 精绘花卉丝巾图案生成方法的计算效率与硬件成本

评价指标	训练时间/h	推理时间/s	显存占用/GB
SDXL		9	16.3
Dreambooth	5	9	16.3
img2img		10	16.5
img2img&IP-Adapter-Plus		8	17.3

3 结 语

为了提高丝巾图案设计效率、降低人工成本,本文提出了一种结合生成对抗网络和稳定扩散模型的花卉丝巾图案生成方法。首先提出一种数据扩增工作流,工作流创新性的融合真实丝巾图案的图像特征和文本特征,构建了具有多样性的花卉丝巾数据

集。然后在此数据集基础上,设计了“双阶段模型协同生成”框架,形成“粗生成—细加工”的完整流程,最后生成精绘花卉丝巾图案。

实验结果表明,该方法的 FID 值低至 41.25,生成的精绘花卉丝巾图案清晰且质量优异,与真实丝巾图案高度接近。该方法的艺术美观度平均得分为 4.5 分,生成的精绘花卉丝巾图案具有较强的艺术美感。该方法的丝巾规范度平均得分为 4.52,生成图案遵循真实丝巾图案的设计规范。该方法相较于其他精绘花卉丝巾图案生成方法推理时间明显降低,占用显存轻微提高,在计算效率方面和硬件成本方面符合实际应用要求。该方法为丝巾图案的快速生成提供了一种高效高质的解决方案,有助于降低企业的设计成本、提高生产效率,进而推动时尚产业的数字化转型。

本文主要聚焦于花卉主题的丝巾图案生成,在图案多样性和应用普适性方面有待后续研究,如几何抽象、文化符号等多元主题丝巾图案的生成,或跨品类的纺织图案生成。

参考文献:

- [1] Radford A, Metz L, Chintala S, et al. Unsupervised representation learning with deep convolutional generative adversarial networks [EB/OL]. (2016-01-07) [2025-03-03]. <https://arxiv.org/abs/1511.06434v2>.
- [2] 任雨佳,陈璐,陈郁.基于 DCGAN 算法的服装款式交互设计[J].毛纺科技,2022,50(3):93-97.
- [3] Arjovsky M, Chintala S, Bottou L. Wasserstein generative adversarial networks[C]// Proceedings of the 34th International Conference on Machine Learning. Sydney, Australia. PMLR, 2017: 214-223.
- [4] 田乐,祝双武,王茹,等.生成对抗网络及其在纺织行业中的应用[J].纺织科技进展,2023(11):1-7.
- [5] Wu Q, Zhu B, Yong B, et al. ClothGAN: Generation of fashionable Dunhuang clothes using generative adversarial networks[J]. Connection Science, 2021, 33(2): 341-358.
- [6] 刘康,马浩然,邢乐.基于生成对抗网络的中式婚服设计[J].服装学报,2024,9(3):208-214.
- [7] Sohl-Dickstein J, Weiss E, Maheswaranathan N, et al. Deep unsupervised learning using nonequilibrium thermodynamics[C]// Proceedings of the 32nd International Conference on Machine Learning. Lille, France. PMLR, 2015: 2256-2265.
- [8] Ho J, Jain A N, Abbeel P. Denoising diffusion probabilistic models[C]// Proceedings of the 34th International Conference on Neural Information Processing Systems. Red Hook, NY: Curran Associates, Inc., 2020: 6840-6851.
- [9] Rombach R, Blattmann A, Lorenz D, et al. High-resolution

- image synthesis with latent diffusion models[C]//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). June 18-24, 2022, New Orleans, LA, USA. IEEE, 2022: 10674-10685.
- [10] Podell D, English Z, Lacey K, et al. SDXL: Improving latent diffusion models for high-resolution image synthesis [EB/OL]. (2023-07-04) [2025-03-03]. <https://arxiv.org/abs/2307.01952v1>.
- [11] Ruiz N, Li Y Z, Jampani V, et al. DreamBooth: fine tuning text-to-image diffusion models for subject-driven generation[C]//2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). June 17-24, 2023, Vancouver, BC, Canada. IEEE, 2023: 22500-22510.
- [12] Hu E J, Shen Y, Wallis P, et al. LoRA: Low-rank adaptation of large language models [EB/OL]. (2021-10-16) [2025-03-03]. <https://arxiv.org/abs/2106.09685?context=cs.LG>.
- [13] 黄诗敏. STABLE DIFFUSION(SD)技术在广绣创作中的应用[D]. 广州: 广州大学, 2024: 30-64.
- [14] 李钢. 基于改进 CycleGAN 和稳定扩散模型的陶瓷图案生成算法研究[D]. 景德镇: 景德镇陶瓷大学, 2024: 45-56.
- [15] 冉二飞, 贾小军, 王子祥, 等. 应用稳定扩散模型的蓝印花布单纹样自动生成[J]. 现代纺织技术, 2024, 32(12): 48-57.

(责任编辑: 康 锋)