

# 基于生成对抗网络的女上装图像属性编辑

肖红梅<sup>1</sup>, 王伟珍<sup>\*1,2</sup>, 房媛<sup>3</sup>

(1. 大连工业大学 服装学院, 辽宁 大连 116034; 2. 大连工业大学 服装人因与智能设计研究中心, 辽宁 大连 116034; 3. 大连工业大学 工程训练中心, 辽宁 大连 116034)

**摘要:**为解决当前服装图像属性编辑模型生成图像存在属性缺失或冗余的问题,提出一种基于 Fashion-AttGAN 的优化模型对女上装图像细节进行变换的设计方法;通过优化特征提取网络,将结构相似性损失项加入重构损失,提高生成器的属性编辑能力;使用 CP-VTON 数据集训练,对女上装图像中袖长和颜色的细节进行调整。结果表明,生成图像在袖型连贯性和颜色准确性方面得到提升,改进模型收敛趋势更平稳,重构图像的结构相似性指标提升了 27.4%,峰值信噪比提高了 2.8%。该优化模型有效减少了生成图像的属性冗余和残缺,为服装图像细节变换研究提供参考。

**关键词:** 服装设计;深度学习;图像生成;生成对抗网络;属性编辑

中图分类号:TS 941.26 文献标志码:A 文章编号:2096-1928(2024)01-0042-06

## Image Attribute Editing of Women's Tops Based on Generating Adversarial Networks

XIAO Hongmei<sup>1</sup>, WANG Weizhen<sup>\*1,2</sup>, FANG Yuan<sup>3</sup>

(1. School of Fashion, Dalian Polytechnic University, Dalian 116034, China; 2. Clothing Human Factors and Intelligent Design Research Center, Dalian Polytechnic University, Dalian 116034, China; 3. Engineering Training Center, Dalian Polytechnic University, Dalian 116034, China)

**Abstract:** In order to solve the problem of attributes missing or redundant in the current clothing image attribute editing models of generate images, a design method based on Fashion-AttGAN model was conducted to transform the details of women's tops. This paper optimized feature network and added structure similarity index measure to the reconstructed loss function to improve the attribute editing ability of generator. The CP-VTON dataset was used for training to ultimately achieve fine-grained editing of women's tops sleeve length and color. The experimental results show that the generated image achieves the improvement in sleeve coherence and color accuracy, the improved model is shown to move more smoothly towards convergence trend, the reconstructed image structure similarity index measure realizes the growth of 27.4% and peak signal-to-noise ratio grows by 2.8%. The proposed model reduces attributes missing or redundant in generated images and provides a technical reference for its detail transformation.

**Key words:** fashion design, deep learning, image synthesis, generative adversarial network, attribute editing

款型设计是服装生产过程中核心内容产出的重要环节。深度学习技术在近年不断成熟,其中生

成对抗网络(generated adversarial network, GAN)模型在服装图像生成领域应用广泛。对于服装图像

收稿日期:2023-06-13; 修订日期:2023-11-15。

基金项目:教育部社科规划基金项目(21YJAZH088);辽宁省教育厅高校基本科研重点攻关项目(LJKZZ20220069);辽宁省教育厅项目(1010152);中国纺织工业联合会项目(2021BKJGLX321);辽宁省自然科学基金规划基金项目(2022-BS-263);辽宁省十四五教育科学规划项目(JG21DB054)。

作者简介:肖红梅(1999—),女,硕士研究生。

\* 通信作者:王伟珍(1976—),男,副教授,硕士生导师。主要研究方向为服装人因与智能设计。

Email: wz-wang@foxmail.com

区域属性编辑任务,学界分别通过改进模型和训练方式开展相关研究并取得显著效果。其中时尚生成对抗网络 (FashionGAN)<sup>[1]</sup> 模型可以解决深度卷积生成对抗网络 (deep convolution generative adversarial networks, DCGAN)<sup>[2]</sup> 所产生的服装图像结构不一致以及图像模糊的问题,其通过语义分割和文本编码对输入的图像和描述进行处理,并对编辑后的形状加以纹理渲染。Fashion-AttGAN<sup>[3]</sup> 对面部属性编辑生成对抗网络 (attribute generative adversarial networks, AttGAN)<sup>[4]</sup> 模型的损失函数进行改进,使得模型能够编辑服装图像颜色和袖长两个属性。为展现生成模型在产品设计领域的实用性,属性设计生成对抗网络 (design attribute GAN, DA-GAN)<sup>[5]</sup> 在 AttGAN 基础上采用了新的损失函数。属性编辑生成对抗网络 (attribute manipulation generative adversarial networks, AMGAN)<sup>[6]</sup> 模型将条件生成对抗网络 (condition GAN, CGAN) 模型、卷积神经网络 (convolution neural network, CNN) 和类激活图 (class activation maps, CAMs) 算法结合,实现了袖子和领型的精细控制。生成器使用 CAMs 增强注意机制,且在模型中添加属性鉴别器网络,使得模型专注于指定属性的编辑。TailorGAN<sup>[7]</sup> 模型使用自监督训练方式实现服装的属性编辑,使用对抗损失和重建损失实现不同领型和不同袖长之间的转换。虚拟私人裁缝网络 (virtual personal tailoring network, VPTNet)<sup>[8]</sup> 模型将可编辑服装图像属性拓展至衣长,整个模型分为形状生成和色彩纹理渲染两个部分,通过使用变分自编码器生成对抗网络 (variational autoencoder GAN, VAE-GAN) 模型,对输入的语义图和边缘图进行编码,编辑为目标形状后完善细节。

上述模型从不同角度提供了服装款型设计的方法,但对于款型变化的细节之处研究较少,图像的形状编辑效果存在模糊<sup>[9]</sup> 和无关属性 (如图案和纹理) 被修改的问题<sup>[10]</sup>。为解决这些编辑不精准的问题 (如袖子轮廓不完整、领型多余轮廓等),文中提出一种优化 Fashion-AttGAN 模型以提高编辑后的图象质量。

# 1 技术背景

## 1.1 GAN 生成对抗网络

GAN 模型在生成真实自然图像任务中展示出了强大能力,它由生成模型 Generator 和判别模型 Discriminator 组成,其结构如图 1<sup>[11]</sup> 所示。图 1 中将噪音向量  $z$  输入到生成模型  $G$  后生成假图像

$x_{\text{fake}}$ 。将真实图像  $x_{\text{real}}$  和假图像  $x_{\text{fake}}$  输入判别模型,通过分类网络  $D$  识别出输入图像的真假。生成模型和判别模型交替训练,使生成器学习到真实图像数据分布。GAN 的损失函数见式 (1),训练目标为求得损失函数的最小值。

$$\min_G \max_D V(G, D) = E_{x \sim P_{\text{data}}} [D(x)] - E_{z \sim P_z} \{D[G(z)]\}。$$

(1)

式中:  $G$  为生成器;  $D$  为判别器;  $x$  为输入的图像,可为  $x_{\text{real}}$  或者  $x_{\text{fake}}$ ;  $z$  为输入的噪声向量;  $V(G, D)$  为估值函数;  $\min_G V(G, D)$  为  $G$  的目标最小化估值函数;  $\max_D V(G, D)$  为  $D$  的目标最大化估值函数。

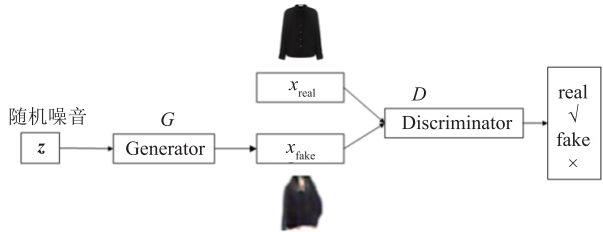


图 1 生成对抗网络结构

Fig.1 Structure of GAN

直接应用该 GAN 模型容易产生模式崩塌,无法完成服装图像属性编辑。Fashion-AttGAN 模型能实现图像属性的精准修改,更适合应用到上装图像属性编辑的研究中。故文中选择 Fashion-AttGAN 模型作为基准进行改进。

## 1.2 属性编辑

属性编辑是近年生成式模型关注的热点问题,其可以通过使用特定算法精准控制图像中指定属性的呈现效果,如将女上装图像中的长袖和圆领属性修改为无领和 V 领。对于此类任务,最终目标是在不影响其他信息的情况下控制颜色、领型和袖长等服装属性。文中以女上装图像的袖长和颜色作为研究变量和载体,通过优化 Fashion-AttGAN 模型结构,提高编辑效果的细节真实度与完整性。

Fashion-AttGAN 模型的图像处理流程如图 2 所示。模型包含生成器和判别器两个子网络,生成器部分包括  $G_{\text{enc}}$  编码器和  $G_{\text{dec}}$  解码器,判别器部分包括判别器  $D$  和分类器  $C$ <sup>[3]</sup>。输入图像  $X^a$  表示带有属性  $a$  的真实图像,重构图像  $X^{\hat{a}}$  表示带有属性  $a$  的重构图像,属性编辑图像  $X^{\hat{b}}$  表示带有属性  $b$  的虚假图像。 $a$  和  $b$  为上衣图像的袖长和颜色,均为粗粒度属性,其中袖长属性包括无袖、短袖、盖肩袖、中袖/5 分袖、6 分袖/7 分袖/8 分袖、9 分袖和长袖 7 种细粒度;颜色属性包括红色、绿色、蓝色、紫色、灰色和黑色 6 种细粒度。 $G_{\text{enc}}$  编码器本质上是一个特

征提取网络,负责将输入的信息转化为特征向量;解码器则负责将特征向量转化为目标,将低维度数据转换回原始数据空间。通过这个转换过程的多次迭代,可以学习到有效的数据分布。在训练阶段,向生成器输入真实图像  $X^a$ ,以及  $n$  分类标注的属性标签  $\mathbf{a} = [a_1, a_2, \dots, a_n]$ ,编码器  $G_{\text{enc}}$  根据输入的信息将图像编码为潜在表示  $\mathbf{z}$ 。

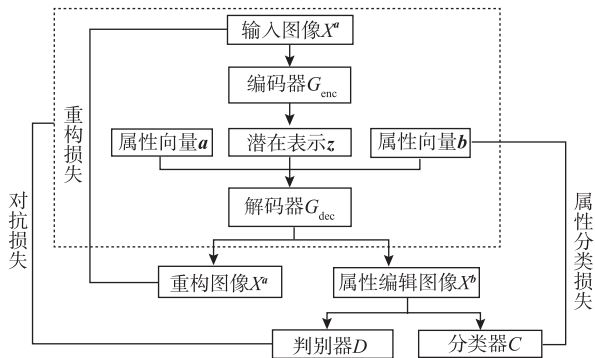


图2 Fashion-AttGAN 流程示意

Fig.2 Process schematic of Fashion-AttGAN

利用解码器  $G_{\text{dec}}$  将输入图像的属性  $\mathbf{a}$  编辑为属性  $\mathbf{b}$ 。具体步骤为:将潜在表示  $\mathbf{z}$  和标注的属性  $\mathbf{b} = [b_1, b_2, \dots, b_n]$  输入给  $G_{\text{dec}}$ ,生成的图像表示为  $X^b$ 。编辑过程为

$$X^b = G_{\text{dec}}(\mathbf{z}, \mathbf{b})。 \quad (2)$$

式中: $X^b$ 为生成的具有属性  $\mathbf{b}$  的图像; $G_{\text{dec}}$ 为生成器中的解码器部分; $\mathbf{z}$ 为潜在表示; $\mathbf{b}$ 为输出图像中目标属性编码。

为了保持目标属性以外的图像语义不变,还需进行重构学习。将潜在表示  $\mathbf{z}$  和属性  $\mathbf{a} = [a_1, a_2, \dots, a_n]$  输入  $G_{\text{dec}}$ ,生成图像  $X^a$ 。重构过程为

$$X^a = G_{\text{dec}}(\mathbf{z}, \mathbf{a})。 \quad (3)$$

式中: $X^a$ 为生成的具有属性  $\mathbf{a}$  的图像; $\mathbf{a}$ 为输入图像中的属性编码。

在生成器网络部分,使用一次  $G_{\text{enc}}$  生成潜在表示  $\mathbf{z}$ ,使用两次  $G_{\text{dec}}$  生成  $X^b$  和  $X^a$ ;在判别器网络部分,将  $X^b$  放进  $D$  判别器和  $C$  分类器,  $D$  进行粗粒度的真假判别,  $C$  进行细粒度的属性判别。同时,将  $X^b$  和  $X^a$  放入  $D$  判别器进行图像真伪判别的博弈训练。

在测试阶段,只进行图像的生成,判别器网络不参与,将  $X^a$  编辑为  $X^b$  的过程为

$$X^b = G_{\text{dec}}[G_{\text{enc}}(X^a), \mathbf{b}]。 \quad (4)$$

式中: $X^a$ 为真实的具有属性  $\mathbf{a}$  的图像; $G_{\text{enc}}$ 为生成器中的编码器部分。

由图2 流程示意以及式(1)~式(4)得出 Fashion-AttGAN 的损失函数包括属性分类损失、重构学习损失和对抗学习损失<sup>[3]</sup>。

## 2 女上装图像属性编辑方法

在现有的研究中,由于模型提取特征能力有限,加之数据集分布不均等,生成图像存在袖型、领型的残缺或冗余问题,因此,选取 Fashion-AttGAN 模型进行改进,通过优化模型结构和损失函数实现服装设计的细节要求。

### 2.1 特征提取网络优化

原始 Fashion-AttGAN 中编码器使用5层特征提取网络,从结果来看,并不能满足服装图像的特征提取要求,其生成的袖长存在残缺或冗余,颜色变化受到输入图像自身色彩影响较大。针对这些问题,改进模型增加了生成器和判别器的网络层数,同时使用了U-net结构,U-net通过拼接的方式实现服装图像中颜色以及袖长的特征融合。改进后,生成器的特征提取能力增强。优化后的网络结构如图3所示。由图3可知,分类器和判别器共享前7层网络,卷积层进行卷积操作,反卷积层进行转置卷积操作。

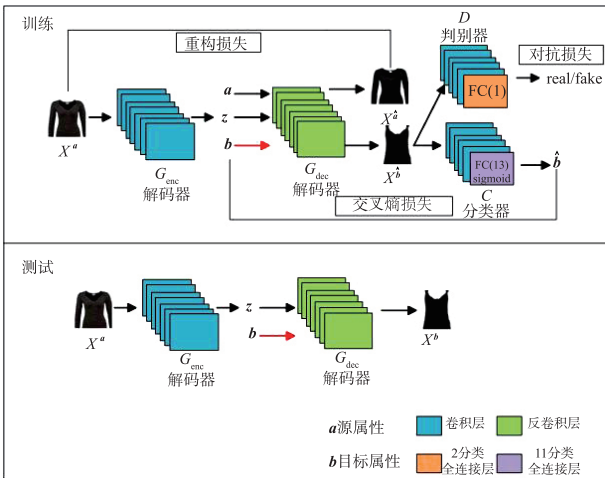


图3 改进 Fashion-AttGAN 模型结构

Fig.3 Improved Fashion-AttGAN structure

### 2.2 损失函数优化

使用合适的损失函数是最大化 GAN 生成能力的关键步骤之一。在图3 模型训练中,模型的损失函数被分为重构损失、对抗损失和交叉熵损失3部分,分别命名为  $L_{\text{rec}}$ ,  $L_{\text{adv}}$  和  $L_{\text{cls}}$ ,最终目标是通过训练使得损失函数达到最小值。重构损失保证除指定属性以外的其他属性编辑前后保持一致;对抗损失保证生成图像的真实性;交叉熵损失鉴别生成图像是否满足输入的属性要求。

在 Fashion-AttGAN 模型结构中,重构损失使用了  $L_1$  损失函数优化生成器网络,  $L_1$  损失函数即平均绝对误差。  $L_1$  损失函数将重构图像与输入图像进行比较,计算平均绝对误差,所得值越小,代表重构图像与输入图像越相似,但这种像素级别的还原会导致图像丢失部分信息,使模型学习到的数据分布存在偏差,生成的图像模糊不清。为优化这一问题,文中研究在重构损失部分加上了 (structure similarity index measure, SSIM) 结构相似性<sup>[12]</sup> 损失。SSIM 损失函数用来评估生成图像与真实图像之间的相似度,可通过图像的对比度、亮度和结构相关性计算得到。SSIM 的取值范围是  $[0, 1]$ , 数值越接近 1, 表示生成图像与输入图像之间的相似度越高,计算公式为

$$L_{\text{SSIM}}(x^a, x^{\hat{a}}) = \frac{(2\mu_{\chi a}\mu_{\chi \hat{a}}\mu_{\chi \hat{a}} + C_1)(2\sigma_{\chi a\chi \hat{a}} + C_2)}{(\mu_{\chi a}^2 + \mu_{\chi \hat{a}}^2 + C_1)(\sigma_{\chi a}^2 + \sigma_{\chi \hat{a}}^2 + C_2)} \quad (5)$$

式中:  $x^a$  和  $x^{\hat{a}}$  分别为待比较的两幅图像;  $\mu_{\chi a}$  和  $\mu_{\chi \hat{a}}$  分别为两幅图像的均值;  $\sigma_{\chi a}$  和  $\sigma_{\chi \hat{a}}$  分别为图像的标准差;  $\sigma_{\chi a\chi \hat{a}}$  为图像的协方差;  $C_1$  和  $C_2$  是为了避免分母为  $\mu_{\chi a}^2 + \mu_{\chi \hat{a}}^2$  和  $\sigma_{\chi a}^2 + \sigma_{\chi \hat{a}}^2$  接近 0 时的不稳定性添加的常数项<sup>[12]</sup>。

重构损失可表示为

$$\min_{G_{\text{dec}}, G_{\text{enc}}} L_{\text{rec}} = E_{x^a \sim P_{\text{data}}} [\|x^a - x^{\hat{a}}\|_1] + L_{\text{SSIM}}(x^a, x^{\hat{a}}) \quad (6)$$

式中:  $L_{\text{rec}}$  为生成器部分的重构损失;  $\min_{G_{\text{dec}}, G_{\text{enc}}} L_{\text{rec}}$  为编码器  $G_{\text{enc}}$  和解码器  $G_{\text{dec}}$  训练目标最小化的重构损失函数;  $E$  为后续函数的数学期望值;  $x^a \sim P_{\text{data}}$  为  $x^a$  服从真实数据的概率分布  $P_{\text{data}}$ 。

对抗损失  $L_{\text{adv}}$  使用 WGAN-GP 的方法,由  $L_{\text{adv}_G}$  和  $L_{\text{adv}_D}$  组成,即

$$\min_{G_{\text{dec}}, G_{\text{enc}}} L_{\text{adv}_G} = -E_{x^a \sim P_{\text{data}}, b \sim P_{\text{attr}}} [D(x^{\hat{b}})] \quad (7)$$

式中:  $L_{\text{adv}_G}$  为对抗损失函数中生成器的部分;  $x^{\hat{b}}$  为生成的具有属性  $b$  的图像;  $b \sim P_{\text{attr}}$  为  $b$  服从真实数据的概率分布  $P_{\text{attr}}$ ;  $D$  为判别器。

$$\min_D L_{\text{adv}_D} = -E_{x^a \sim P_{\text{data}}} D(x^a) + E_{x^a \sim P_{\text{data}}, b \sim P_{\text{attr}}} [D(x^{\hat{b}})] \quad (8)$$

交叉熵损失也称为属性分类损失,使用交叉熵计算得到,在改进模型中,属性分类损失由  $L_{\text{cls}_G}$  和  $L_{\text{cls}_C}$  组成,分别计算属性  $a$  和  $b$  的分类损失。属性分类损失表达式分别为

$$\min_C L_{\text{cls}_C} = E_{x^a \sim P_{\text{data}}} [l_G(x^a, a)] \quad (9)$$

$$\min_{G_{\text{dec}}, G_{\text{enc}}} L_{\text{cls}_G} = E_{x^a \sim P_{\text{data}}, b \sim P_{\text{attr}}} [l_G(x^{\hat{b}}, b)] \quad (10)$$

式中:  $L_{\text{cls}_C}$  为分类损失函数中分类器的部分;  $l_G$  为交叉熵函数;  $a$  为输入图像中的属性编码;  $L_{\text{cls}_G}$  为分类损失函数中判别器的部分;  $b$  为输出图像中目标属性编码。

因此,改进后的 Fashion-AttGAN 中总损失函数包括  $L_{\text{cls}_G}$  和  $L_{\text{cls}_C}$  组成的判别器损失函数、 $L_{\text{rec}}$  和  $L_{\text{adv}_G}$  组成的生成器损失函数和  $L_{\text{cls}_G}$  组成的生成器中解码器额外损失函数。

判别器损失函数表达式为

$$\min_{D, C} L_{\text{dis}, \text{cls}} = \lambda_2 L_{\text{cls}_C} + L_{\text{adv}_D} \quad (11)$$

式中:  $L_{\text{dis}, \text{cls}}$  为判别器损失函数;  $\lambda_2$  为超参数,需要在训练过程中确定;  $L_{\text{adv}_D}$  为对抗损失函数中判别器部分。

生成器网络损失函数表达式为

$$\min_{G_{\text{dec}}, G_{\text{enc}}} L_{\text{enc}, \text{dec}} = \lambda_1 L_{\text{rec}} + L_{\text{adv}_G} \quad (12)$$

式中:  $L_{\text{enc}, \text{dec}}$  为生成器中编码器和解码器共同的损失函数;  $\lambda_1$  为超参数,需要在训练过程中确定。

生成器中解码器的额外损失函数表达式为

$$\min_{G_{\text{dec}}} L_{\text{dec}} = \lambda_3 L_{\text{cls}_G} \quad (13)$$

式中:  $L_{\text{dec}}$  为生成解码器的损失函数;  $\lambda_3$  为超参数,需要在训练过程中确定。

## 3 实验及结果

### 3.1 数据集

实验使用特性保留的虚拟试穿网络 (characteristic-preserving virtual try-on network, CP\_VTON)<sup>[13]</sup> 数据集进行训练。数据集中包含 14 221 张女装上衣图像,每一张图像都有对应的属性向量,分别标注了袖长、衣领和颜色 3 种粗粒度属性。文中实验选取袖长和颜色 2 种粗粒度属性作为实验对象,其中袖长有 7 种细粒度,颜色有 6 种细粒度,所以实验中总计使用了 13 种细粒度属性。

### 3.2 搭建环境

使用 PyCharm 作为集成开发环境, TensorFlow 作为上衣图像属性编辑的学习框架。实验程序运行在 NVIDIA GeForce RTX 4080 显卡的计算机中,其内存为 128G,显卡内存为 16G,处理器为 i7。

### 3.3 参数设定

经过反复实验,式(11)~式(13)中超参数  $\lambda_1$  设定为 100,控制重构损失权重;  $\lambda_2$  设定为 1,控制输入图像属性  $a$  分类损失权重;  $\lambda_3$  设定为 3,控制目标属性  $b$  分类损失权重。迭代次数 epoch 代表整个数据集被送入到网络完成一轮前向传播和反向传播的次数, batch\_size 是每次 epoch 从训练集中随机



抽取的部分样本进行一次前向传播和反向传播的样本数。实验中 epoch 设置为 200, batch\_size 设置为 32。

3.4 结果与分析

将改进模型与 Fashion-AttGAN 模型进行比较,使用 CP-VTON 数据集,实现袖长和颜色的编辑。

在袖型编辑效果方面,改进模型的袖型结构连贯性和颜色的编辑效果优于原始 Fashion-AttGAN 模型。如对“9 分袖”这个属性的编辑效果进行改进,改进后的模型相较于 Fashion-AttGAN 模型残缺部分减少。袖长编辑效果对比如图 4 所示。



图 4 袖长编辑效果对比

Fig. 4 Sleeve length editing comparison

颜色花纹编辑效果对比如图 5 所示。由图 5 可知,改进后模型生成的颜色更深,而原始模型受输入图像颜色影响较大。由图 5 可以看出,改进后模型的花纹保留得更加完整。在图 5 第 2 行中,输入的服装图像带有花纹,原始 Fashion-AttGAN 模型在进行属性编辑时,花纹信息有很大损失;在改进模型中添加 SSIM 损失能最大程度保留与目标属性无关的其他属性,因此在改进模型结果中花纹信息最大程度得以保留。



图 5 颜色花纹编辑效果对比

Fig. 5 Color and pattern editing comparison

重构图像效果对比如图 6 所示。相较于原始 Fashion-AttGAN 模型,改进后模型生成的重建图像清晰度更优。对于服装内轮廓的阴影部分还原得更加真实,如图 6 第 1 行输入图像服装内轮廓中有

白色标签,在原始 Fashion-AttGAN 模型重构图像中,此标签相对模糊,在改进模型中更加真实地还原了内轮廓以及标签。



图 6 重构图像效果对比

Fig. 6 Reconstruction image comparison

Fashion-AttGAN 重构损失趋势和改进模型的重构损失趋势如图 7 所示。由图 7 可知,改进模型的网络训练收敛趋势更加平稳,而原始模型出现了大幅度的震荡。

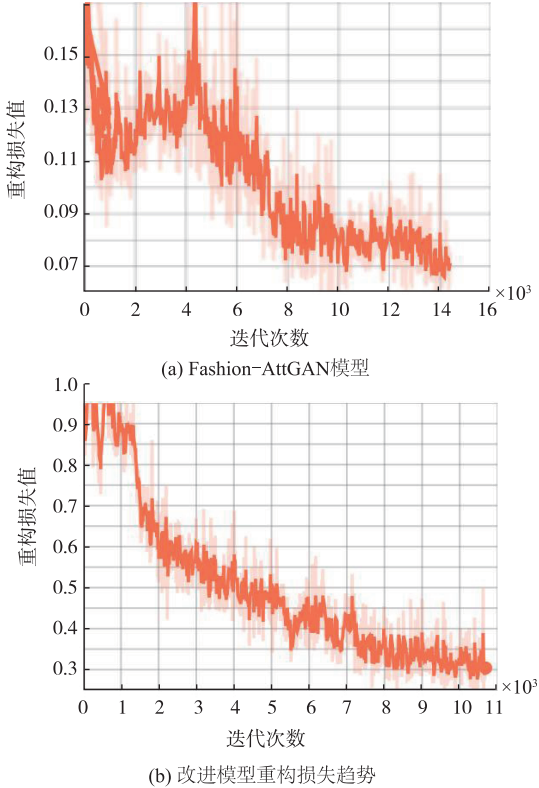


图 7 改进模型重构损失趋势

Fig. 7 Improved Fashion-AttGAN reconstruction loss graph

为了证明使用 SSIM 损失的有效性,对改进后的模型和 Fashion-AttGAN 模型生成的重构图像的结构相似性指标 SSIM 和峰值信噪比 PSNR (peak signal-to-noise ratio) 进行了数据对比。SSIM 是衡量两幅图像相似度的指标,其取值范围为[0,1],数值越大代表重构图像越接近输入图像。PSNR 是信号

最大功率与信号噪声功率之比,用于量化经过影像压缩后图像或视频的重建质量,值越大代表生成图像质量越好。

将改进前后的各 20 张重构图像以及原始图像输入程序,并测量生成图像的 SSIM 和 PSNR,比较其均值。具体结果见表 1。

表 1 重构图像质量对比

Tab.1 Evaluation of image reconstruction quality		
对比指标	SSIM	PSNR
Fashion-AttGAN	0.475 6	30.149 1
改进后模型	0.606 0	31.010 1

由表 1 可以看出,相较于原始 Fashion-AttGAN 模型,改进后模型 SSIM 从 0.475 6 变化为 0.606 0,上升 27.4%,SSIM 值越大代表重构图像越接近输入图像,由此验证了添加 SSIM 损失函数的有效性;改进后模型的 PSNR 评分相较于原始 Fashion-AttGAN 提升 2.8%,PSNR 值越大代表生成图像质量越好。

## 4 结 语

文中以女上装的袖长和颜色作为研究变量和载体,通过优化 Fashion-AttGAN 模型结构,增强特征提取能力,增加 SSIM 损失提高图像的清晰度以提升生成指定属性的能力,从而提高编辑效果的细节真实性与完整性。最终实验表明,改进后的 Fashion-AttGAN 模型在袖长和颜色的编辑效果以及图像的清晰度上优于 Fashion-AttGAN 模型,重构图像的 SSIM 提升 27.4%,PSNR 值提高 2.8%,为服装图像细节变化研究提供技术参考。后续将继续进行参数简化和模型优化,减少训练时长。

### 参考文献:

[ 1 ] ZHU S Z, FIDLER S, URTASUN R, et al. Be your own prada: fashion synthesis with structural coherence[ C ]//2017 IEEE International Conference on Computer Vision (ICCV). Venice: IEEE, 2017: 1689-1697.

[ 2 ] RADFORD A, METZ L, CHINTALAS S. Unsupervised representation learning with deep convolutional generative adversarial networks[ EB/OL]. (2016-01-07) [2023-05-28]. <https://arxiv.org/pdf/1511.06434.pdf>.

[ 3 ] PING Q, WU B, DING W Y, et al. Fashion-AttGAN: attribute-aware fashion editing with multi-objective GAN

[ C ]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). Long Beach: IEEE, 2019: 323-325.

[ 4 ] HE Z L, ZUO W M, KAN M N, et al. AttGAN: facial attribute editing by only changing what you want[ J]. IEEE Transactions on Image Processing, 2019, 28(11): 5464-5478.

[ 5 ] YUAN C X, MOGHADDAM M. Attribute-aware generative design with generative adversarial networks[ J]. IEEE Access, 2020(8): 190710-190721.

[ 6 ] AK K, LIM J H, THAM J, et al. Attribute manipulation generative adversarial networks for fashion images[ C ]//2019 IEEE/CVF International Conference on Computer Vision (ICCV). Seoul: IEEE, 2019: 10540-10549.

[ 7 ] CHEN L L, TIAN J, LI G, et al. TailorGAN: making user-defined fashion designs[ C ]//2020 IEEE Winter Conference on Applications of Computer Vision (WACV). Snowmass: IEEE, 2020: 3241-3250.

[ 8 ] KWON Y, PETRANGELI S, KIM D, et al. Tailor me: an editing network for fashion attribute shape manipulation[ C ]//2022 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV). Waikoloa: IEEE, 2022: 3142-3151.

[ 9 ] LIU M, DING Y K, XIA M, et al. STGAN: a unified selective transfer network for arbitrary image attribute editing[ C ]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Long Beach: IEEE, 2019: 3668-3677.

[ 10 ] ZHANG G, KAN M N, SHAN S G, et al. Generative adversarial network with spatial attention for face attribute editing[ C ]//Computer Vision-ECCV 2018: 15th European Conference. Munich: ACM, 2018: 422-437.

[ 11 ] CRESWELL A, WHITE T, DUMOULIN V, et al. Generative adversarial networks: an overview[ J]. IEEE Signal Processing Magazine, 2018, 35(1): 53-65.

[ 12 ] WANG Z, BOVIK A C, SHEIKH H R, et al. Image quality assessment: from error visibility to structural similarity[ J]. IEEE Transactions on Image Processing: a Publication of the IEEE Signal Processing Society, 2004, 13(4): 600-612.

[ 13 ] WANG B C, ZHENG H B, LIANG X D, et al. Toward characteristic-preserving image-based virtual try-on network[ C ]//Computer Vision-ECCV 2018: 15th European Conference. Munich: ACM, 2018: 607-623.

(责任编辑:张 雪)