



基于投影流网络和嵌入空间损失的纹样风格迁移

朱昱儒^a,侯 珏^{a,b},杨 阳^{a,b},刘 正^{b,c}

(浙江理工大学, a. 服装学院; b. 丝绸文化传承与产品设计数字化技术文旅部重点实验室;

c. 国际时装技术学院, 杭州 311199)

摘 要: 针对编解码器框架造成的图像重建误差、恢复偏差以及图像风格迁移方法存在内容泄露和局部伪影等亟待解决的问题,提出了一种基于投影流网络以及嵌入空间损失的纹样风格迁移模型。通过融合投影流网络与矢量量化变分自编码器实现无偏风格迁移,并对风格特征进行精细编码和匹配,保留更完整图像内容细节,且捕获关键风格特征;设计了一种计算风格化图和风格图的嵌入空间损失,融入总体损失函数以保证风格特征的均匀分布,减少风格差异;以云锦为例进行织物纹样风格的创新设计。结果表明:基于投影流网络和嵌入空间损失的纹样风格迁移的迁移效果,在内容评价指标结构相似性(Structural similarity, SSIM)、内容损失上的分值较对比模型分别提高86.21%、54.29%、20%和32.58%、18.68%、18.99%;风格评价指标 Gram 损失为 4.5×10^{-6} ,较对比模型提高近一倍,表明该方法有效平衡了内容保留与风格迁移的需求,提高风格化效果。该模型在改善内容泄露的同时避免了内容失真和风格过度覆盖,有效捕捉复杂风格特征和色彩层次,促进纹样创新设计,增加风格迁移在纺织服装领域的可能性。

关键词: 风格迁移;嵌入空间损失;矢量量化变分自编码器;内容泄露;投影流;纹样

中图分类号: TS941.26

文献标志码: A

文章编号: 1673-3851(2025)03-0237-09

引文格式: 朱昱儒,侯珏,杨阳,等. 基于投影流网络和嵌入空间损失的纹样风格迁移[J]. 浙江理工大学学报(自然科学),2025,53(2):237-245.

Reference Format: ZHU Yuru, HOU Jue, YANG Yang, et al. Pattern style transfer based on projection flow networks and embedding space loss[J]. Journal of Zhejiang Sci-Tech University, 2025, 53(2): 237-245.

Pattern style transfer based on projection flow networks and embedding space loss

ZHU Yuru^a, HOU Jue^{a,b}, YANG Yang^{a,b}, LIU Zheng^{b,c}

(a. School of Fashion Design & Engineering; b. Key Laboratory of Silk Cultural Heritage and Product Design Digital Technology, Ministry of Culture and Tourism; c. Zhejiang International Institute of Fashion Technology, Zhejiang Sci-Tech University, Hangzhou 311199, China)

Abstract: Aiming at the problems that need to be solved urgently, such as image reconstruction error, recovery biases caused by coded framework, and content leak and local artifacts during the style transfer process, a pattern style transfer model based on projection flow network and embedded space loss was proposed. Firstly, by integrating the projection flow network with the vector quantized-variational autoencoder, unbiased style transfer was realized, and style features were finely coded and matched to retain more complete image content details and capture key style features. Then, a method was designed to calculate the embedded space loss of stylized graph and style graph, and the overall loss function was incorporated to ensure the uniform distribution of style features and reduce style differences. With Yunjin brocade as an example, the innovative design of fabric pattern style was carried out. The results showed that the transfer effect of pattern style transfer based on projection flow network and embedded space loss

收稿日期: 2024-05-13 网络出版日期: 2024-09-13

基金项目: 文化和旅游部重点实验室资助项目(23072247-N)

作者简介: 朱昱儒(2001—),女,河南周口人,硕士研究生,主要从事服装数字化方面的研究。

通信作者: 刘 正, E-mail: koala@zstu.edu.cn

scored higher by 86.21%, 54.29% and 20%, and 32.58%, 18.68% and 18.99%, respectively in terms of the content evaluation metrics of structural similarity (SSIM), and content loss compared to the comparison models; with a Gram loss of 4.5×10^{-6} , the model nearly doubled the performance of comparative models, indicating that this method effectively balanced the needs of content retention and style transfer and improved the stylization effect. The model improves content leak while avoiding content distortion and style over-coverage, effectively captures complex style features and color hierarchies, facilitates innovative design of patterns, and increases the possibility of style transfer in the textile and apparel field.

Key words: style transfer; embedding space loss; vector quantized-variational autoencoder; content leak; projection flow; patterns

0 引言

图像风格迁移技术是将一幅图像的风格特征应用到另一幅图像上的技术,广泛应用于服装设计^[1-3]、纹样创新^[4-6]等场景。风格迁移技术通常基于预训练的卷积神经网络来提取图像的内容特征和风格特征,再由解码器重构这些特征,以生成具有目标风格的新图像^[7]。然而,由于解码器不能实现对内容图像的无损重建,累积的重建误差会逐渐干扰内容细节导致内容泄漏现象,进而造成内容失真和局部伪影等问题^[8]。

风格迁移早期通过非参数化的算法实现^[9],例如:Efros等^[10]提出通过拼接和重组样本纹理以生成新纹理的方法;基于类比思想,Hertzmann等^[11]通过映射两图像间的特征关系生成具有新纹理的图像。上述非参数方法只能提取图像底层特征,在处理具有复杂语义信息的图像时,生成图效果粗糙且费时费力。由于卷积神经网络能够更有效地理解和理解图像的高层语义特征,使得风格迁移不局限于纹理模仿,还能在保持图像内容的同时变换其整体风格,基于深度学习的方法逐渐成为主流。Gatys等^[12]利用卷积神经网络提取图像不同层级的特征,实现对任意图像内容和风格的分离和重组;Johnson等^[13]通过最小化特征空间中的差异实现了实时图像风格迁移。以上固定风格迁移的方法限制了应用的范围和灵活度,与之相较,通用风格迁移不再受限于特定的训练样本,可以实现将任意风格应用于任意内容的风格迁移。Li等^[14]提出白化和着色变换(Whitening and coloring transforms, WCT),不是依赖特定的风格或内容图像,而是通过匹配特征协方差来实现风格迁移,具有风格无关的泛化能力,但对处理具有复杂纹理或细节的艺术风格时具有限制性;Huang等^[15]在前馈神经网络的基础上加入自适应

应实例归一化层(Adaptive instance normalization, AdaIN),将内容图像特征与风格图像特征的均值和方差对齐,实现了任意风格的实时风格迁移,同时面临保留内容细节和实现风格迁移之间的平衡难题;Li等^[16]提出了一种基于线性变换矩阵学习的任意风格迁移方法,但对于细腻复杂的风格特征,该方法仍存在伪影的问题。

针对上述问题,本文提出了一种基于投影流网络(Projection flow network, PFN)和嵌入空间损失的纹样风格迁移模型。该模型将投影流网络与矢量量化变分自码器(Vector quantised-variational autoencoder, VQ-VAE)相结合,从而在嵌入空间中实现精细的风格特征匹配;同时设计了一种嵌入空间损失函数,以计算和优化风格信息特征间的差异;最后将模型用于纹样图案创新设计,以云锦为风格、传统山水画为内容进行风格迁移。由于该模型有效地平衡了内容保留与风格迁移的需求,其内容泄漏和局部伪影的问题将得到改善,从而提高风格迁移的质量和在纺织服装领域的实用价值。

1 投影流和嵌入空间损失的风格迁移

图1为内容泄漏现象图,展示了WCT、AdaIN和Avatar-Net风格迁移方法在不同轮次应用下的变化。图1(a)—(b)分别为内容图和风格图,图1(c)体现出在第1轮风格化中,三种方法能较好地保留图像内容;但图1(d)中第20轮风格化后的图像内容几乎完全丧失,难以辨认,这一现象即为内容泄漏。因此,现有风格迁移方法在多次应用后的内容保留能力存在不足,仍需要改进。

1.1 纹样风格迁移模型

为了在克服内容泄漏的同时改善图像迁移效果局部伪影的问题,本文提出了基于投影流网络和嵌入空间损失的纹样风格迁移模型,该模型的

结构图如图 2 所示。模型主要由无偏风格迁移模块和风格增强模块组成,其中无偏风格迁移模块包含 2 个深度学习的网络框架,分别为 PFN 和 AdaIN,风格增强模块包含 1 个深度学习的网络框架,为 VQ-VAE。模型首先通过 PFN 进行前向传播,提取出内容图和风格图的无损深度特征 f_c 和 f_s 作为 AdaIN 的输入;其次通过 AdaIN 得到风格化特征输出 f_{cs} ,再由 PFN 反向传播,将风格化特

征 f_{cs} 重构为风格化图像;然后将风格化图像和风格图像输入 VQ-VAE 中进行编码,分别得到离散的向量,从而计算出二者间的嵌入空间损失 L_{vq} ,并将其添加到总体损失中;最后通过优化嵌入空间损失增加对风格化图像的一致性约束,在尽量减少对内容图像结构性改变的同时,获取更多风格特征,有效融合目标风格,最终生成更精细的迁移效果图。



图 1 内容泄漏现象示例

注:第 1 行为 WCT 风格迁移方法在不同轮次下的效果;第 2 行为 AdaIN 风格迁移方法在不同轮次下的效果;
第 3 行为 Avatar-Net 风格迁移方法在不同轮次下的效果。

1.1.1 无偏风格迁移模块

无偏风格迁移模块基于 ArtFlow 进行设计,由 PFN 与 AdaIN 组成,通过采用 PFN 的正向和反向传播代替编码-解码结构,结合 AdaIN 层进行风格迁移,从而避免累积的图像重建误差与解码器训练偏差所导致的内容泄漏现象,实现无偏风格迁移。

PFN 主要由激活标准化层、可逆 1×1 卷积和加性耦合组成(见图 2),这 3 部分都是可逆的,从而在正向和反向传播过程中达到无偏无损的效果。其中:激活标准化层对每个通道的激活进行仿射变换,通过初始化尺度和偏置参数,在给定初始数据小批量的情况下,使每个通道激活后具有零均值和单位方差,并且通过输入数据预处理使其分布更稳定,有助于提高模型的训练效果和生成样本的质量。可逆 1×1 卷积使用可学习的权重矩阵对特征映射的通道维度进行置换和灵活排列,实现特征映射间的信息交换^[17];加性耦合实现了可逆的非线性变换,相较于仿射耦合层,加性耦合层简化了计算过程,提高了模型训练的稳定性 and 效率。PFN 采用挤压操作

通过将特征沿空间维度分割,再沿通道维度重组,减少了 2 D 特征图的空间大小,通过增加通道维度的维度保留了局部相关性,进一步增强了模型的代表能力。

表 1 为投影流网络主要组成部分,列出了 PFN 的 3 个主要组成部分的函数、逆函数和对数行列式,其中: \mathbf{x}, \mathbf{y} 为形状为 $h \times w \times c$ 的张量,分别表示层的输入和输出; i, j 表示张量上的空间位置; s 和 t 表示仿射变换的尺度和偏置参数,都是可学习的; \mathbf{W} 为 $c \times c$ 的权重矩阵;NN()是输入和输出具有相同形状的神经网络;split()函数沿通道维度将输入张量拆分成两半;concat()函数执行相应的反操作,将输入张量拼接成单个张量。

AdaIN 具有无偏性^[18],其核心是自适应实例归一化层,通过将内容图像特征的均值和方差与风格图像的均值和方差对齐,在特征空间中进行风格迁移。假定 \mathbf{x}, \mathbf{y} 分别表示输入的内容图与风格图经 PFN 正向传播所得到的特征张量,AdaIN 层计算公式如下:

AdaIN(\mathbf{x}, \mathbf{y}) = $\sigma(\mathbf{y}) \left(\frac{\mathbf{x} - \mu(\mathbf{x})}{\sigma(\mathbf{x})} \right) + \mu(\mathbf{y})$ (1)

其中: $\sigma(\cdot)$ 为标准差, $\mu(\cdot)$ 为均值。式(1)表示直接

用对应特征通道上风格图 \mathbf{y} 所计算得到的均值和方差替换 BN、IN 中学习得到的仿射参数。

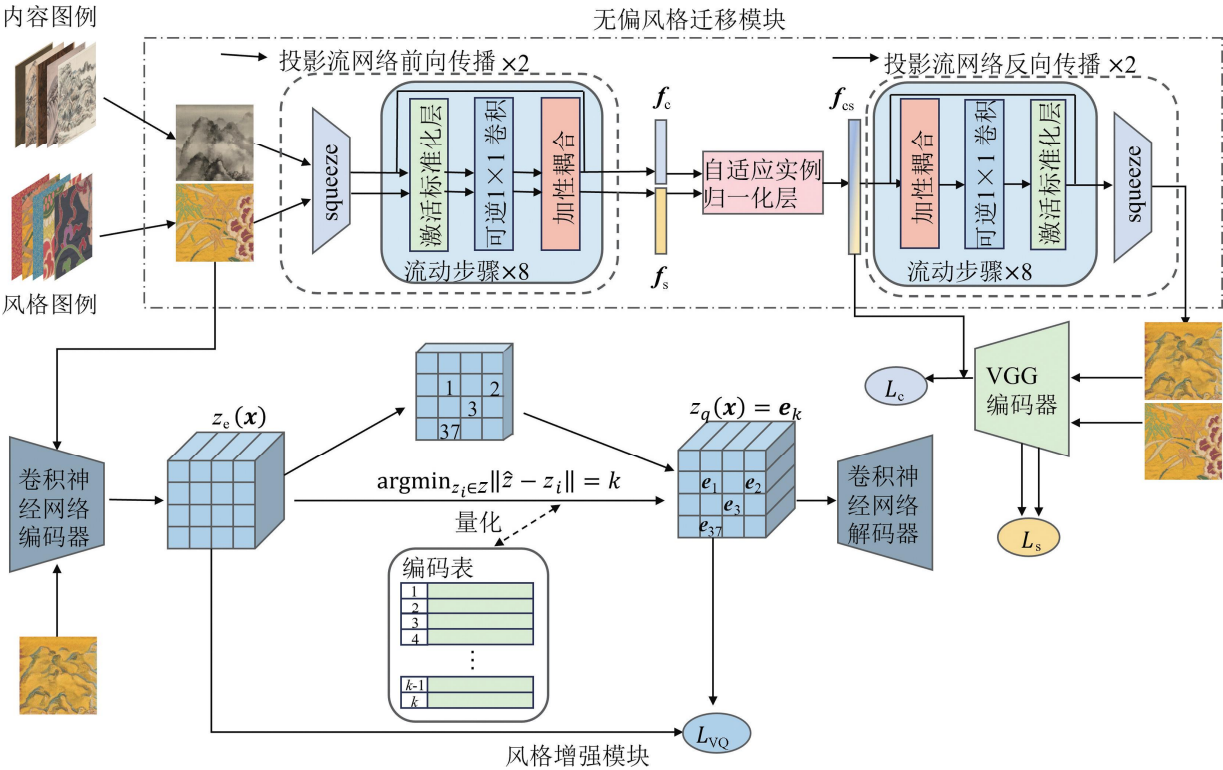


图 2 基于投影流网络和嵌入空间损失的纹样风格迁移模型结构示意图

表 1 投影流网络主要组成部分

组件	函数	逆函数	对数行列式
激活标准化层	$y_{i,j} = s \odot x_{i,j} + t$	$x_{i,j} = (y_{i,j} - t) / s$	$h \cdot w \cdot \text{sum}(\log s)$
可逆 1×1 卷积 $\mathbf{W}: [c \times c]$	$y_{i,j} = \mathbf{W}x_{i,j}$	$x_{i,j} = \mathbf{W}^{-1}y_{i,j}$	$h \cdot w \cdot \log \det(\mathbf{W}) $
加性耦合	$\mathbf{x}_a, \mathbf{x}_b = \text{split}(\mathbf{x}),$ $\mathbf{y}_b = \text{NN}(\mathbf{x}_a) + \mathbf{x}_b,$ $\mathbf{y} = \text{concat}(\mathbf{x}_a, \mathbf{x}_b)$	$\mathbf{y}_a, \mathbf{y}_b = \text{split}(\mathbf{y}),$ $\mathbf{x}_a = \mathbf{y}_a,$ $\mathbf{x}_b = \mathbf{y}_b - \text{NN}(\mathbf{y}_a),$ $\mathbf{x} = \text{concat}(\mathbf{x}_a, \mathbf{x}_b)$	0

1.1.2 风格增强模块

将无偏风格迁移模块所输出的风格化图与风格图输入到风格增强模块,增加嵌入空间损失项(见图2)。风格增强模块可看作 VQ-VAE 优化模块,主要由编码器、量化层和解码器组成^[19]。VQ-VAE 是结合变分自编码器^[20]和量化方法的一种生成模型,不同于变分自编码器将输入数据映到一个连续的潜在空间中,VQ-VAE 将潜在空间离散化为有限的离散代码本,以获得更紧凑和离散的表示。

VQ-VAE 定义了一个潜在嵌入空间 $\mathbf{e} \in \mathbf{R}^{K \times D}$,其中: K 是离散潜在嵌入空间的大小, D 是每个潜在嵌入向量 \mathbf{e}_i 的维度,存在 K 个嵌入向量 $\mathbf{e}_i \in \mathbf{R}^D, i \in 1, 2, \dots, K$ 。首先将图像 x 传入编码器得到连续的编码向量 $\mathbf{z}_e(\mathbf{x})$;其次量化层将 $\mathbf{z}_e(\mathbf{x})$ 映射到离散的潜在嵌入空间中,并通过最邻近搜索

将其量化为 K 个向量之一;最后将嵌入空间的最近邻向量 $\mathbf{z}_q(\mathbf{x})$ 作为解码器的输入映射回原始图像空间,从而得到重构图像 $\hat{\mathbf{x}}$ 。上述计算过程可以用式(2)表示:

$$\begin{cases} \mathbf{z}_e(\mathbf{x}) = \text{Encoder}(\mathbf{x}); \\ \mathbf{z}_q(\mathbf{x}) = \mathbf{e}_k, k = \arg\min_j \|\mathbf{z}_e(\mathbf{x}) - \mathbf{e}_j\|_2; \\ \hat{\mathbf{x}} = \text{Decoder}(\mathbf{z}_q(\mathbf{x})) \end{cases} \quad (2)$$

VQ-VAE 的优化目标可以总结为 2 部分:重建损失和嵌入空间损失,分别用来优化编码器、解码器以及嵌入空间。其整体损失函数可以用式(3)表示:

$$L = \log(\mathbf{x} | \mathbf{z}_q(\mathbf{x})) + \|\text{sg}[\mathbf{z}_e(\mathbf{x})] - \mathbf{e}\|_2^2 + \delta \|\mathbf{z}_e(\mathbf{x}) - \text{sg}[\mathbf{e}]\|_2^2 \quad (3)$$

其中:第一部分 $\log(\mathbf{x} | \mathbf{z}_q(\mathbf{x}))$ 为重建损失,是输入图像和重建图像的均方误差,衡量了模型对输入数据的还原能力,用于优化模型的编码器以及解码

器。第二、三部分为嵌入空间损失,衡量了模型对数据的表示能力,其中: $\| \text{sg} [\mathbf{z}_e(\mathbf{x})] - \mathbf{e} \|_2^2$ 为码本损失,用于优化嵌入空间; $\delta \| \mathbf{z}_e(\mathbf{x}) - \text{sg} [\mathbf{e}] \|_2^2$ 为编码器损失,用于约束编码器的输出和嵌入空间保持一致,避免编码器输出产生较大变动。第四部分 sg 表示停止梯度,通过分开计算嵌入更新和编码器更新,嵌入和编码器以不同的速度优化。

1.2 损失函数

1.2.1 内容损失和风格损失

内容损失为风格化图像特征与目标特征表示之间的欧氏距离。在计算内容损失时,为了更好地实现内容和风格的分离,控制二者之间的平衡,使用自适应实例归一化层的输出 \mathbf{f}_{cs} 作为内容目标,而不是直接使用内容图像的常用特征响应。内容损失 L_C 可以用式(4)表示:

$$L_C = \| \mathbf{F}(\mathbf{I}_{\text{cs}}) - \mathbf{f}_{\text{cs}} \|_2 \quad (4)$$

其中: \mathbf{I}_{cs} 表示生成的风格化图像, \mathbf{F} 表示预训练好的 VGG-19 编码器。

风格损失是根据特征层特征表示的均值和标准差之间的欧氏距离计算得到的,用来衡量风格化图像和风格图像之间的差异。由于低层卷积网络重点获取图像的纹理、线条等底层特征,高层卷积网络重点获取图像的轮廓、语义信息等高级特征,选取等权重的 relu1_1 、 relu2_1 、 relu3_1 、 relu4_1 层来计算风格损失^[12-13,15],以平衡不同层次特征对风格化图像的影响,均衡各层特征图在风格迁移中所发挥的作用,避免某一层特征过度主导风格迁移的结果。风格损失 L_S 可以用式(5)表示:

$$L_S = \sum_{i=1}^L \| \mu(\boldsymbol{\phi}_i(\mathbf{I}_{\text{cs}})) - \mu(\boldsymbol{\phi}_i(\mathbf{I}_s)) \|_2 + \sum_{i=1}^L \| \sigma(\boldsymbol{\phi}_i(\mathbf{I}_{\text{cs}})) - \sigma(\boldsymbol{\phi}_i(\mathbf{I}_s)) \|_2 \quad (5)$$

其中: \mathbf{I}_s 表示输入的风格图像, μ 、 σ 分别表示特征图的均值和标准差, $\boldsymbol{\phi}$ 表示 VGG-19 编码器输出的计算风格损失的特征映射, i 表示当前层。

1.2.2 嵌入空间损失

为了使生成图在保留更多内容信息,避免内容泄漏的同时,更好地捕捉和应用风格信息,增加对风格化图像的一致性约束,本文引入了风格增强模块。通过将风格化图像和风格图像在 VQ-VAE 的嵌入空间进行编码,计算得到嵌入和其对应的编码器输出的均方误差,从而增加了嵌入空间损失 L_{vq} ,该损失可以用式(6)表示:

$$L_{\text{vq}} = \| \text{sg} [\mathbf{I}_{\text{cs}}] - \mathbf{I}_s \|_2^2 + \delta \| \mathbf{I}_{\text{cs}} - \text{sg} [\mathbf{I}_s] \|_2^2 \quad (6)$$

综上所述,总体损失函数为内容损失 L_C 、风格损失 L_S 以及嵌入空间损失 L_{vq} 的加权组合,可以用式(7)表示:

$$L = \alpha L_C + \beta L_S + \gamma L_{\text{vq}} \quad (7)$$

其中: α 、 β 和 γ 均为权重参数,表示所对应损失在整个损失中的权重占比。

2 实验与结果分析

2.1 实验设置

本文使用 Pytorch 作为风格迁移的深度学习框架进行模型的训练和测试,模型训练过程使用单张 NVIDIA GeForce RTX 3080Ti 显卡。实验以 MS-COCO 数据集作为内容图像, WikiArt 数据集作为风格图像。由于数据集中图像尺寸不一致,在训练中首先将输入图像的大小统一调整为 512×512 像素,然后随机裁剪为 256×256 像素。模型最大迭代次数设置为 160000 次,批大小设置为 2,学习率设置为 1×10^{-4} ,衰减率设置为 5×10^{-5} 。将内容损失 L_C 、风格损失 L_S 以及嵌入空间损失 L_{vq} 作为模型的训练目标,采用 Adam 优化算法进行对网络进行优化,其中 L_C 、 L_S 和 L_{vq} 的损失权重分别被设置为 0.1、1.0 和 0.1。

2.2 风格迁移结果

为验证模型在图像风格迁移中的有效性,将本文模型与 AdaIN、LinearWCT、ArtFlow 进行对比实验,得到主客观评价结果。

2.2.1 主观评价

风格迁移效果对比图如图 3 所示,其中:风格图为 WikiArt 数据集中选取的画作,示例图像如图 3(a)所示;内容图为 MS-COCO 数据集中选取的图像,示例图像如图 3(b)所示;图 3(c)~(f)分别为 AdaIN、LinearWCT、ArtFlow 和本文方法生成的风格迁移效果图。

从图 3(c)中可以看出:AdaIN 生成的风格化图像虽然大致保留了内容结构,但存在局部纹理细节缺失、色彩混杂等问题。这是由于 AdaIN 忽略了局部特征分布的复杂性,且 AdaIN 生成的图像(c)依赖于风格图和内容图的高层特征,导致生成图细节层面的纹理细节被忽略或过度简化;通过调整均值和方差对复杂色彩分布的风格图进行风格迁移时,可能导致色彩分布的混杂不清。LinearWCT 生成的图像(d)表现出局部边缘模糊和伪影的迁移现象,产生这种现象的原因是 LinearWCT 的白化和着色过程,可能会在图像边缘产生不自然的效果,导致视

觉上的伪影。ArtFlow 生成的风格化图像(e)相比前两种迁移模型表现出更好的迁移效果,但仍存在局部伪影、内容信息丢失等问题。本文优化模型利用投影流进行无偏风格迁移,并在此基础上增加了嵌入空间损失,由本方法生成的图像(f)可以看出,

猫咪以及人物脸部细节更加清晰丰富,建筑结构保留完整、边缘和背景清晰光滑,表明优化模型在有效保护内容信息的同时,能够更细腻地处理不同区域的特征分布,更细致地调整色彩分布和纹理细节。

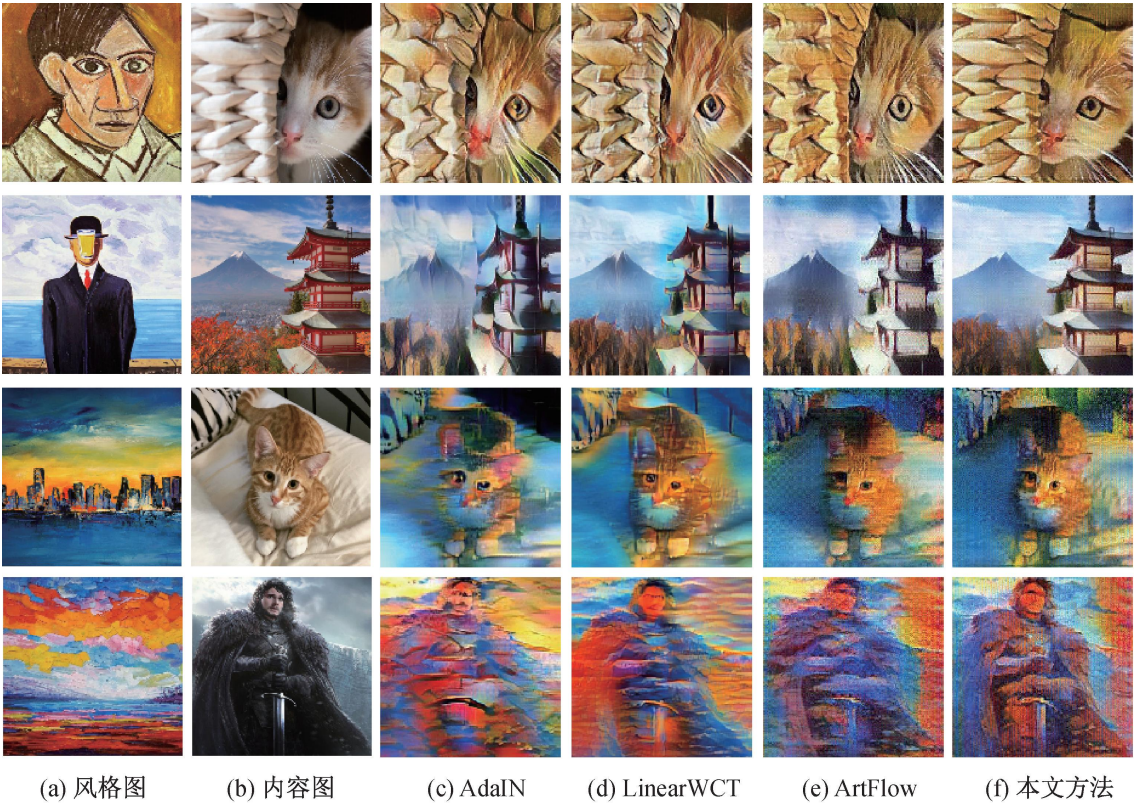


图 3 风格迁移效果示例图像

2.2.2 客观评价

SSIM、内容损失和 Gram 矩阵是衡量风格迁移的重要客观量化指标^[21]。SSIM 和内容损失是内容度量标准,用于评估风格迁移中保留内容信息细节的能力,其中:内容损失通过计算原始内容图像和生成图像在某一层特征映射上的差异来评估两个图像间的相似度,其值越小,表示生成图保留的原始内容特征越多,迁移效果越好;Gram 矩阵均方误差是风格度量标准,通过特征映射计算来评估生成图像和目标风格图像在风格特征上的相似度,其值越小,表示生成图像与目标风格图像的风格越相似,风格化的效果和质量越好(见表 2)。

表 2 客观评价指标对比

模型	结构相似性	内容损失	Gram 损失
AdaIN	0.29	3.10	1.27×10^{-3}
LinearWCT	0.35	2.57	9.3×10^{-4}
ArtFlow	0.45	2.58	9.8×10^{-4}
本文模型	0.54	2.09	4.5×10^{-6}

SSIM 基于图像中提取的亮度、对比度及结构等关键特征来评估两幅图像相似度,其中像素均值用于比较两个图像的亮度;像素方差用于比较两个

图像的对比度;像素协方差用于比较两个图像的结构。SSIM 的范围为 0 到 1,值越大表示图像质量越好,两幅图像相似性越高,SSIM 的计算公式为:

$$SSIM(m,n)=\frac{(2\mu_m\mu_n+c_1)(2\sigma_{mn}+c_2)}{(\mu_m^2+\mu_n^2+c_1)(\sigma_m^2+\sigma_n^2+c_2)} \tag{8}$$

其中: μ_m 、 μ_n 和 σ_m 、 σ_n 分别表示图像 m 和 n 的均值和方差; σ_{mn} 表示图像 m 和 n 之间的协方差; c_1 、 c_2 为常数。

由表 2 可知:在内容评价指标上,本文模型的 SSIM 值较对比模型分别提高了 86.21%、54.29% 和 20%,并且内容损失较对比模型分别提高了 32.58%、18.68% 和 18.99%,风格评价指标 Gram 矩阵均方误差为 4.5×10^{-6} ,较对比模型提高了近一倍,进一步验证了基于 PFN 和嵌入空间损失的纹样风格迁移模型能在保留更多内容信息的同时生成更精细的效果图。

2.3 内容泄漏

如果发生内容泄漏现象,在进行多次风格迁移后,原始图像的内容信息会逐渐丢失。在训练过程中,当损失不再显著下降,生成图像与内容图像间的

SSIM 开始趋于稳定且视觉效果满意时,可认为达到预期轮次。将 ArtFlow 与本文模型分别进行多轮风格迁移,选取第 1 轮、第 10 轮和第 20 轮迁移结果进行对比,获得内容泄漏现象对比图,示例图像如图 4 所示。图 4(a)~(b)分别为内容图和风格图,图 4(c)~(d)分别为 ArtFlow 迁移结果和本文方法迁移结果。从图 4 中可以看出:通过使用 PFN 代替基于 VGG19 的编码-解码结构,2 种方法都成功避

免了内容泄漏,其中 ArtFlow 模型第 1 轮、第 10 轮和第 20 轮的迁移结果与内容图间的 SSIM 值分别为 0.6480、0.6479、0.6483,本文模型第 1 轮、第 10 轮和第 20 轮的迁移结果与内容图间的 SSIM 值分别为 0.7314、0.7307、0.7311,这表明本文模型通过利用 VQ-VAE 增加嵌入空间损失,进一步提高了风格迁移的质量,呈现出更清晰、不失真的图像。

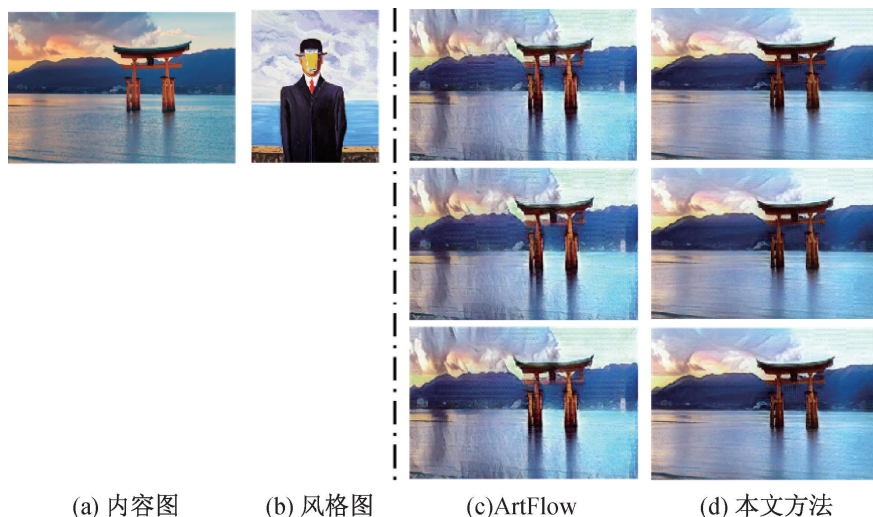


图 4 内容泄漏现象对比图

注:第 1 行为第 1 轮的风格迁移结果;第 2 行为第 10 轮的风格迁移结果;第 3 行为第 20 轮的风格迁移结果。

2.4 以云锦风格为例的纺织品纹样创新设计

为了验证模型在处理我国传统织物的独特风格特征和复杂色彩信息方面的能力,本文构建了一个包含中国传统山水画^[22]和南京云锦面料的特色数据集,将山水画作为内容图,云锦作为风格图对比多种风格迁移模型的性能。图 5 为纺织品纹样风格迁移效果对比图,风格图为图 5(a)中的云锦面料,内容图为图 5(b)中的山水画。AdaIN 和 LinearWCT 在使用色彩分布复杂的云锦风格图进行风格迁移时,不能细致地表达原风格图的色彩信息,由于 AdaIN 算法限制,其生成的迁移图像(图 5(c))无法将云锦风格均匀地应用到内容图像的不同纹理和颜色上,造成一些灰色斑块,影响视觉质量。LinearWCT 生成的迁移图像(图 5(d))显示,在处理对比度较高、原始内容图像中边缘不清晰或者有阴影效果的区域时,难以保持颜色边界,导致颜色“溢出”的问题。ArtFlow 生成的迁移图像(图 5(e))表现出该模型不能很好地平衡好内容结构与色彩风格的关系,导致色彩和纹理的不均匀匹配。本文方法生成的迁移图像(图 5(f))显示:本文模型通过结合 VQ-VAE 实现对云锦风格特征的精细编码,有效捕捉云锦纹样的复杂细节和色彩层次,使风格化图像中的色彩分布更加贴近原风格图像,保证了风格特征的

一致性;通过优化嵌入空间损失,不仅促进了整体风格的均匀迁移,还针对云锦风格中易产生局部伪影的区域进行细致调整,提高了视觉一致性。由于云锦或者其他织锦的风格具有高度的独特性和复杂性,传统方法往往难以准确表达其丰富的色彩,迁移结果表明本文模型在风格特征捕捉的精确性、内容与风格的平衡处理以及改善局部伪影等方面具有优越性。

3 结 论

本文提出了一种基于投影流网络和嵌入空间损失的纹样风格迁移模型,并在 MS-COCO、WikiArt 以及云锦和山水画数据集上进行风格迁移,实现纹样创新设计,主要结论如下:

a) 利用投影流进行无偏风格迁移,再引入矢量量化变分自编码器将风格信息编码到离散的代码本中,通过计算嵌入空间损失函数,对网络进一步优化,解决了传统风格迁移模型内容结构失真和色彩纹理表现不佳等问题。

b) 在 MS-COCO 和 WikiArt 数据集上的主客观评价结果表明,基于投影流网络和嵌入空间损失的纹样风格迁移模型能够捕捉和保留更多风格信息,获得内容结构保留更加完整、色彩纹理更加精细的效果图,有效地提高了图像风格迁移的视觉质量。

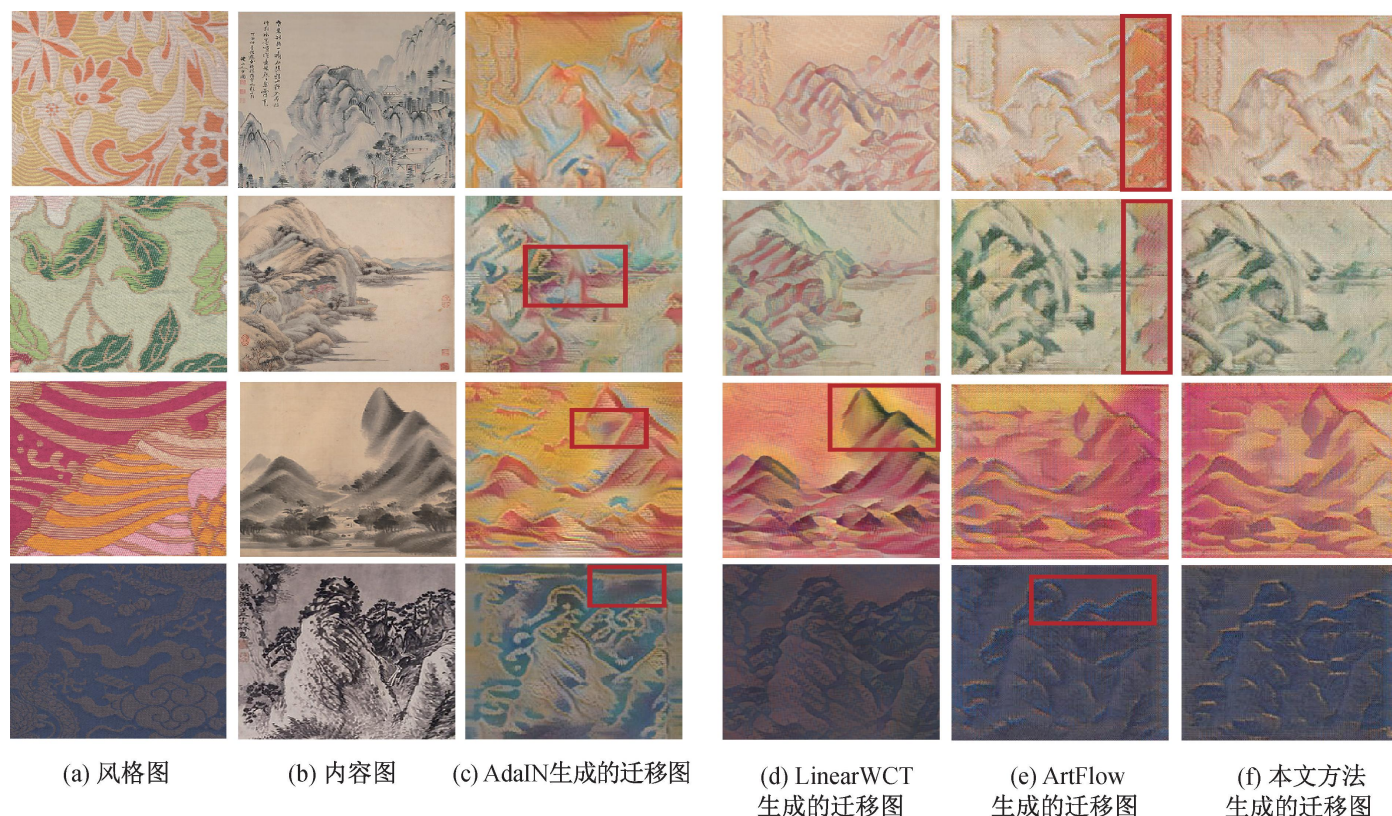


图5 纺织品纹样风格迁移效果对比

c)在云锦和山水画数据集上的迁移结果验证了本文模型在纺织服装领域实际场景的应用中具有优越性,能有效促进纹样创新设计。

本文提出的基于投影流网络和嵌入空间损失的纹样风格迁移模型在纺织服装领域显示出一定的应用前景,但相比真实的纺织品依然存在差距。后续将进一步研究风格迁移中纹理效果的提升方法,以生成更具真实感和质感的迁移效果图,进而拓展其在纺织服装领域的应用。

参考文献:

- [1] 程鹏飞, 王伟珍, 房媛. 基于卷积神经网络的风格迁移泳装图案设计[J]. 丝绸, 2023, 60(3): 97-104.
- [2] 王伟珍, 张功. 基于 CycleGAN 的服装图像混搭风格迁移[J]. 现代纺织技术, 2023, 31(4): 250-258.
- [3] 蔡兴泉, 李治均, 奚梦瑶, 等. 基于神经网络的手绘服饰图纹上色及风格迁移[J]. 系统仿真学报, 2023, 35(3): 604-615.
- [4] 侯宇康, 吕健, 刘翔, 等. 基于神经风格迁移网络的民族图案创新方法[J]. 图学学报, 2020, 41(4): 606-613.
- [5] 邱雪琳, 孙迎, 刘正, 等. 基于色彩优化的云锦局部风格迁移[J]. 丝绸, 2022, 59(12): 64-70.
- [6] 姚琳涵, 张颖, 姚岚, 等. 基于多尺度纹理合成的刺绣风格迁移模型[J]. 纺织学报, 2023, 44(9): 84-90.
- [7] 李慧, 万晓霞. 深度卷积神经网络下的图像风格迁移算法[J]. 计算机工程与应用, 2020, 56(2): 176-183.
- [8] 聂雄锋, 王俊英, 董方敏, 等. 融合注意力机制的多模态动漫风格迁移方法[J]. 计算机工程与应用, 2023, 59(15): 223-234.
- [9] 陈淑环, 韦玉科, 徐乐, 等. 基于深度学习的图像风格迁移研究

综述[J]. 计算机应用研究, 2019, 36(8): 2250-2255.

- [10] Efros A A, Freeman W T. Image quilting for texture synthesis and transfer[C]//Proceedings of the 28th annual conference on Computer graphics and interactive techniques. New York, NY, USA: ACM, 2001: 341-346.
- [11] Hertzmann A, Jacobs C E, Oliver N, et al. Image analogies [C]//Proceedings of the 28th annual conference on Computer graphics and interactive techniques. New York, NY, USA: ACM, 2001: 327-340.
- [12] Gatys L A, Ecker A S, Bethge M. A neural algorithm of artistic style[EB/OL]. (2015-09-02) [2024-06-25]. <https://arxiv.org/pdf/1508.06576>.
- [13] Johnson J, Alahi A, Li F F. Perceptual losses for real-time style transfer and super-resolution[C]//European Conference on Computer Vision. Cham: Springer, 2016: 694-711.
- [14] Li Y J, Fang C, Yang J M, et al. Universal style transfer via feature transforms[C]// Proceedings of the 31st International Conference on Neural Information Processing Systems. Red Hook, NY, USA: Curran Associates Inc, 2017: 385-395.
- [15] Huang X, Belongie S. Arbitrary style transfer in real-time with adaptive instance normalization [C]// Proceedings of the IEEE International Conference on Computer Vision. Venice, Italy: IEEE, 2017: 1510-1519.
- [16] Li X T, Liu S F, Kautz J, et al. Learning linear transformations for fast arbitrary style transfer [EB/OL]. (2018-08-14) [2024-06-25]. <https://arxiv.org/pdf/1808.04537>.
- [17] Kingma D P, Dhariwal P. Glow: generative flow with invertible 1×1 convolutions[C]//Proceedings of the 32nd International Conference on Neural Information Processing

- Systems. Red Hook, NY, USA: Curran Associates Inc, 2018: 10236-10245.
- [18] An J, Huang S Y, Song Y B, et al. ArtFlow: unbiased image style transfer via reversible neural flows[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville, TN, USA: IEEE, 2021: 862-871.
- [19] Van Den Oord A, Vinyals O, Kavukcuoglu K. Neural discrete representation learning [C] // Proceedings of the 31st International Conference on Neural Information Processing Systems. Red Hook, NY, USA: Curran Associates Inc, 2017: 6309-6318.
- [20] Doersch C. Tutorial on variational autoencoders[EB/OL]. (2021-01-03) [2024-06-25]. <https://arxiv.org/pdf/1606.05908>.
- [21] Yoo J, Uh Y, Chun S, et al. Photorealistic style transfer via wavelet transforms [C] // Proceedings of the IEEE/CVF International Conference on Computer Vision. Seoul, Korea (South): IEEE, 2019: 9035-9044.
- [22] Xue A. End-to-end Chinese landscape painting creation using generative adversarial networks[C]//Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision. Waikoloa, HI, USA: IEEE, 2021: 3862-3870.

(责任编辑:张会巍)