

基于扩散模型的中国水墨画风格迁移

陈川, 刘恒

安徽工业大学 计算机科学与技术学院, 安徽 马鞍山 243000

摘要:目的 针对大部分风格迁移方法无法在中国水墨画风格领域实现有效迁移的问题, 提出一种基于改进扩散模型采样算法的中国水墨画风格迁移方法, 以实现高质量水墨画风格迁移效果。方法 使用中国水墨画图像数据集训练去噪 Unet 网络, 使得扩散模型获得真实水墨画图像生成能力; 设计损失驱动的扩散模型采样算法, 先对内容图像加噪得到包含部分结构信息的含噪数据, 然后使用内容损失、风格损失以及语义分离损失, 从内容与风格两个角度引导扩散模型对含噪数据逐步去噪, 最终生成风格化结果。结果 定性定量对比实验结果证明: 所提出的方法在中国水墨画风格迁移任务上相较于其他任意风格迁移方法有更好的性能, 能够生成更高质量的水墨风格化图像。消融实验证明: 在内容损失中使用两种不同的计算方式以及在风格损失中使用均值方差色彩损失是合理且有效的。结论 模型在保持内容结构基本不变的同时生成了合适的真实水墨纹理, 为中国水墨画高质量风格迁移提供了有效方法; 生成结果进一步证明: 在高质量风格迁移的需求下, 使用扩散模型技术有很大的发展空间。

关键词: 风格迁移; 中国水墨画; 扩散模型; 损失引导

中图分类号: J212; TP391.41 **文献标识码:** A **doi:** 10.16055/j.issn.1672-058X.2026.0001.010

Chinese Ink Wash Painting Style Transfer Based on Diffusion Model

CHEN Chuan, LIU Heng

School of Computer Science and Technology, Anhui University of Technology, Maanshan 243000, Anhui, China

Abstract: Objective To address the issue that most style transfer methods fail to achieve effective transfer in the domain of Chinese ink wash painting, a Chinese ink wash painting style transfer method based on an improved diffusion model sampling algorithm is proposed to achieve high-quality ink painting style transfer effects. **Methods** A denoising UNet network was trained on a Chinese ink wash painting image dataset to endow the diffusion model with the capability to generate authentic ink painting images. A loss-driven diffusion model sampling algorithm was designed: first, noise was added to the content image to obtain noisy data containing partial structural information; then, the diffusion model performed step-by-step denoising of this data, guided by content loss, style loss, and semantic separation loss from both content and style perspectives, ultimately generating the stylized result. **Results** Qualitative and quantitative comparative experimental results demonstrated that the proposed method outperforms other arbitrary style transfer methods in Chinese ink wash painting style transfer tasks, producing higher-quality ink wash stylized images. Ablation experiments proved that using two different computation methods for content loss and employing mean-variance color loss in style loss is reasonable and effective. **Conclusion** The model generates appropriate and authentic ink wash textures while preserving the basic content structure, providing an effective method for high-quality Chinese ink wash painting style transfer. The generated results further demonstrate great potential for the development of diffusion model technology under the demand for high-quality style transfer.

Keywords: style transfer; Chinese ink wash painting; diffusion model; loss guidance

收稿日期: 2024-02-28 修回日期: 2024-05-03 文章编号: 1672-058X(2026)01-0080-06

基金项目: 安徽省自然科学基金(2008085MF190)项目资助。

作者简介: 陈川(1999—), 男, 硕士研究生, 从事图像风格迁移研究。

通信作者: 刘恒(1976—), 男, 博士, 教授, 从事计算机视觉与模式识别研究。Email: Hengliuskysy@aliyun.com.

引用格式: 陈川, 刘恒. 基于扩散模型的中国水墨画风格迁移[J]. 重庆工商大学学报(自然科学版), 2026, 43(1): 80-85.

CHEN Chuan, LIU Heng. Chinese ink wash painting style transfer based on diffusion model[J]. Journal of Chongqing Technology and Business University (Natural Science Edition), 2026, 43(1): 80-85.

图像风格迁移技术的任务是将指定的风格样式迁移到内容图像上,实现从内容图像域到风格图像域的转变,在艺术图像生成等多个领域都有广泛的应用。任意风格迁移方法通过神经网络提取色彩、纹理以及局部形状等风格信息,然后将融合模块与内容图像的内容信息相结合,从而生成新的风格化图像。这种生成方式支持将任意风格图像与内容图像作为输入,具备较强的通用性。Gatys 等^[1]首次在风格迁移领域使用神经网络方法,其将内容图像参数化并同时计算内容损失与风格损失,通过迭代参数化内容图像来最小化两种损失,实现了内容信息与风格信息的融合。这开创了通过融合内容信息与风格信息实现风格迁移的技术路线与研究领域,后续的相关工作主要在效率与性能上做优化。Ulyanov 与 Johnson 等^[2-3]针对 Gatys 生成手段效率低下的问题提出使用前馈神经网络实现快速风格迁移,通过训练一个前馈神经网络,以实现在输入的任意内容图像上生成指定的风格纹理。该类方法由于具有较高的生成效率以及更强的实用性逐渐为风格迁移的进一步发展指明方向。后来的 Huang、Li 和 Zhang 等^[4-6]通过设计更好的内容与风格特征提取模块以及特征融合模块,进一步实现了任意内容与风格的信息融合,实现了视觉效果更好的任意风格迁移。这些技术手段的发展大大提高了风格迁移的效果与生成效率,为后续的工业化应用奠定了基础。

随着风格迁移技术的进一步发展,相关研究逐渐发现,基于特征合成方法所生成的图像基本只能迁移一些简单的纹理特征,而且存在模糊及伪影现象。因此,相关研究开始关注如何生成复杂逼真的风格特征以及高质量的风格化图像。在动漫风格领域,Chen Y 与 Chen J 等^[7-8]通过训练生成对抗网络(Generative Adversarial Network, GAN)实现了现实图像到动漫风格的高质量转换,其结果进一步证明了特征合成手段的局限性;在油画风格迁移领域,Kotovenko 等^[9]提出一种基于可微笔触渲染器的风格迁移方法,通过迭代参数化的笔触实现了逼真的油画风格迁移,其生成的油画笔触视觉效果远超之前的方法。在中国水墨画风格迁移领域也有一些早期尝试工作,Xue^[10]通过改进的生成对抗网络实现端对端水墨画合成,但由于水墨画纹理结构复杂且其“写意”的图像内容与现实图像域之间差距过大,该方法的生成结果并不好,而且无法实现水墨风格的个性化控制,存在较大局限性。因此,针对具体风格领域设计特定的生成模型是实现在该风格领域下取得高质量迁移效果的有效途径之一。进一步,中国水墨画风格迁移任务需要一种能够实现高质量风格迁移且支持输入任意内容与风格的生成方法,从而

满足相关的需求。

近年来,Ho 等^[11]的工作在图像生成领域取得了重大进展,利用扩散模型实现图像的高质量生成逐渐成为研究热点。针对 DDPM(Dense Discrete Phase Model)只能无条件生成图像的局限,Chung^[12]、Avrahami^[13]以及 Saharia 等^[14]在可控扩散模型图像生成方面做出了积极尝试,通过显示分类器引导的手段让扩散模型按照输入的条件生成对应的内容,但仍然存在控制力不足的问题,无法满足个性化的生成需求。为了实现更加精细的图像生成控制,Kwon 等^[15]使用基于预训练 DINO Vision Transformer(ViT)的内容与风格损失以及基于 CLIP(Contrastive Language-Image Pre-Training)模型的文本图像一致性损失对扩散模型图像生成进行了复杂的干涉,实现了基于输入文本或输入图片的多模态图像编辑,对风格迁移的发展有较强的借鉴参考价值。

综合对中国水墨画风格迁移方法的研究需求以及扩散模型的强大生成能力,本文提出一种基于改进扩散模型采样算法的风格迁移方法。通过由水墨画数据集训练的去噪网络以及所设计采样算法的有效引导,能够在保留内容图像、信息的同时迁移水墨画风格特征,生成真实感强的水墨风格化图像。基于输入的不同风格图片,模型能够将任意内容图像转换为对应的风格,从而满足个性化要求,相较于 Xue^[10]的方法,有更强的实用性以及更高的风格化效果。定性与定量实验结果证明了本文提出方法在中国水墨画图像风格迁移任务上具有较好的性能,超越了基于特征合成的任意风格迁移方法;消融实验证明:采样算法中所用的引导损失在内容保持方面有较好的表现。

1 模型结构与算法设计

1.1 模型总体结构图

本文提出的风格迁移方法网络模型结构如图 1 所示。输入为内容图像 x_c 与风格图片 x_s 。内容图像经过一次加噪模块以及 T 次去噪模块后转变为风格化图像。去噪模块包括 3 个主要步骤:噪声估计、损失计算以及数据调整。噪声估计由预训练的 Unet 网络完成,损失计算需使用到预训练的 VGG 网络以及预训练的 Vision Transformer(ViT)网络。

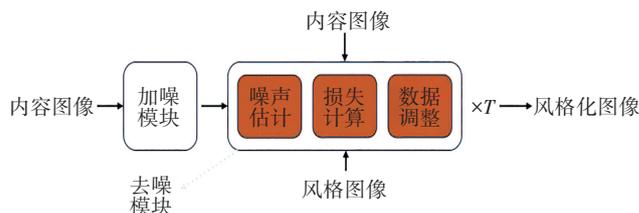


图 1 模型结构图

Fig. 1 Structure diagram of the model

1.2 采样算法

下方采样算法展示了扩散模型通过去噪生成风格化图像的具体过程。

输入:内容图片 x_c , 风格图片 x_s 。

输出:目标风格化图片 x_0 , x_0 同时包含 x_c 的内容信息与 x_s 的风格信息。

$$(1) x_T \sim N(\sqrt{\bar{a}_T} x_c, (1 - \bar{a}_T) I)$$

(2) for $t \in \{T, T-1, \dots, 0\}$ then

$$(3) \varepsilon_t \leftarrow \varepsilon_\theta(x_t, t)$$

$$(4) x'_{t-1} = \frac{1}{\sqrt{a_t}} \left(x_t - \frac{1 - a_t}{\sqrt{1 - \bar{a}_t}} \varepsilon_t \right) + \sigma_t I$$

$$(5) \hat{x}_0(x_t) = \frac{1}{\sqrt{\bar{a}_t}} (x_t - \sqrt{1 - \bar{a}_t} \varepsilon_t)$$

$$(6) \nabla = \nabla_{\text{loss}}(\hat{x}_0(x_t), x_c, x_s)$$

$$(7) x_{t-1} = x'_{t-1} - \nabla$$

(8) end for

采样算法的输入包括内容图片 x_c 与风格图片 x_s 两个部分, ε_θ 是使用中国水墨画数据集预训练好的去噪 Unet 网络。首先使用标准高斯噪声 I 将内容图片进行 T 次加噪, 得到一个包含部分内容结构信息的含噪数据 x_T , 然后从时刻 T 开始迭代去噪。对于任意时刻 $t \in \{T, T-1, \dots, 0\}$, 进行以下步骤:

(1) 使用去噪 Unet 网络 ε_θ 对 x_t 进行噪声估计得到 ε_t 。

(2) 使用估计的噪声进行一次初步的去噪得到 x'_{t-1} 。

(3) 在时刻 t 直接估计最终的干净图像 $\hat{x}_0(x_t)$ 。

(4) 使用 $\hat{x}_0(x_t)$ 与内容图片 x_c 和风格图片 x_s 进行损失计算, 并通过 Pytorch 的自动求导方法得到损失对应于 x_t 的梯度 ∇ 。

(5) 使用梯度 ∇ 对 x'_{t-1} 进行调整, 得到 x_{t-1} 。迭代至 0 时刻后, 得到最终风格化图像 x_0 作为算法的输出。

1.3 内容损失

风格迁移的任务是在保持主要内容结构的基础上进行图像风格的转变, 因此完整捕捉并保留内容图片的内容信息至关重要。常见内容损失(又称内容感知损失)的计算方法基于预训练 VGG 网络提取的全局特征图, 一般认为内容图像与风格化图像的特征图差异越小, 两者在内容结构上的差别越小。然而这种基于特征图的内容表达方式无法将图像内容结构信息与风格信息完全解耦, 所携带的风格信息可能为风格化带来负面影响, 这也正是特征提取模块以及特征融合模块持续被改进的原因之一。

Tumanyan 等^[16]证明, 在预训练的 ViT 网络中, 中

间的多头自注意力层 Keys 包含图像局部块的内容信息, 由 Keys 计算得到的相似性度量矩阵则包含全局的内容结构信息; 最后一层的分类 Token 包含了图像的整体风格表示信息。这为内容信息与风格信息的表达提供了新的计算方式。

针对上述的内容信息与风格信息耦合问题, 本文采用多任务学习思想, 通过同时优化两种不同的内容损失, 促使模型尽可能多地关注共同的内容结构信息, 从而降低其余冗余信息的影响。

内容损失包括 VGG 内容损失以及 ViT 内容损失两部分, 公式表达如下:

$$l_e = \alpha_1 l_{\text{vggc}} + \alpha_2 l_{\text{vite}}$$

$$l_{\text{vggc}}(x_c, x) = \| \text{VGG}(x_c)_i - \text{VGG}(x)_i \|_2$$

$$l_{\text{vite}}(x_c, x) = \| S^l(x_c) - S^l(x) \|_2$$

$$[S^l(x)]_{i,j} = \cos(k_i^l(x), k_j^l(x)) \quad (1)$$

其中, l_{vggc} 中的 $\text{VGG}(x)_i$ 为图片 x 在 VGG 网络中第 i 层的特征图, x_c 是内容图像, x 是在中间时刻 t 估计的最终图像 $\hat{x}_0(x_t)$ 的简称, $S^l(x)$ 是自相似性度量矩阵, $[S^l(x)]_{i,j}$ 为自相似性度量矩阵 $S^l(x)$ 中第 i 行第 j 列的值, $\cos(\cdot)$ 为余弦相似度计算函数, $k_i^l(x)$ 为图像 x 在 ViT 网络的第 l 层多头自注意力层的第 i 个 Key。具体来说, VGG 网络中使用第 4_1 层特征图, ViT 网络中使用第 11 层多头自注意力模块。

另外, 参照文献[15]的做法, 计算图像块对比损失, 增加风格化图像内部的区分度, 以提高内容信息的质量。图像块对比损失表达如下:

$$l_{\text{patch}} = -\alpha_3 \sum \log\left(\frac{p}{p+q}\right)$$

$$p = \exp(\cos(k_i^l(x_c), k_i^l(x))/\tau)$$

$$q = \sum_{i \neq j} \exp(\cos(k_i^l(x_c), k_j^l(x))/\tau) \quad (2)$$

其中, τ 是温度系数。

1.4 风格损失

本文使用 ViT 风格损失传递风格信息, 通过风格损失的引导, 模型能够通过去噪生成与风格纹理类似的纹理细节, 实现高质量中国水墨画的风格迁移。风格损失表达如下:

$$l_{\text{vis}}(x_s, x) = \beta_1 \| t_{[\text{CLS}]}^l(x_s) - t_{[\text{CLS}]}^l(x) \|_2 \quad (3)$$

其中, $t_{[\text{CLS}]}^l(x)$ 为 ViT 网络中最后一层分类 Token。

由于在模型生成实践中, 存在风格损失对色彩生成约束不足的问题, 本文引入额外的色彩损失约束。文献[15]使用像素级的 L_1 损失作为色彩约束, 但同时也引入了风格图像的内容结构信息, 为风格化效果带来负面影响。针对这个问题, 本文使用风格图像色彩

的统计信息(均值与方差)进行色彩约束,从而避免引入风格图像的空间信息,规避了 L_1 损失存在的问题。风格色彩损失表达如下:

$$l_{\text{col}}(x_s, x) = \beta_2 (\| \mu(x_s) - \mu(x) \|_2 + \| \sigma(x_s) - \sigma(x) \|_2) \quad (4)$$

其中, $\mu(x)$ 与 $\sigma(x)$ 分别表示图像 x 在 RGB 色彩空间中的均值与方差。

1.5 语义分离损失

语义分离损失计算迭代去噪过程中连续两次估计的最终图像在风格上的差异,通过最大化该损失能够促使模型生成与上一时刻不同的风格特征,从而加速风格图像的生成。语义分离损失表达如下:

$$l_{\text{sem}}(x_t, x_{t+1}) = -\gamma \| t_{[\text{CLS}]}^t(\hat{x}_0(x_t)) - t_{[\text{CLS}]}^{t+1}(\hat{x}_0(x_{t+1})) \|_2 \quad (5)$$

其中, $\hat{x}_0(x_t)$ 与 $\hat{x}_0(x_{t+1})$ 分别是在 t 时刻以及 $t+1$ 时刻对最终干净风格化图像的估计。

1.6 总损失

本文在采样算法中用来引导的总损失如式(6)所示:

$$l_{\text{total}} = l_c + l_{\text{patch}} + l_{\text{vits}} + l_{\text{col}} + l_{\text{sem}} \quad (6)$$

2 仿真实验与结果分析

2.1 训练数据与实验环境

本文训练数据集为公开的中国水墨画图像训练集,包含 2 195 张图像。模型训练部署在两张 RTX 3090 GPU 上。图像大小被设为 $256 * 256$,扩散模型的总步长为 1 000,加噪步长设为 120,损失权重分别为 $\alpha_1 = 10, \alpha_2 = 1\ 500, \alpha_3 = 200, \beta_1 = 3\ 500, \beta_2 = 8\ 000, \gamma = 100$ 。

2.2 定性对比实验

图 2 展示了与 Safin^[17]、Micro^[18] 以及 EFDM^[19] 风格迁移方法定性对比的实验结果。图 2 中风格图像的纹理属于密集线条型的“牛毛皴”,是水墨山水画的典型纹理之一。纹理迁移是风格迁移的一个主要目标,但在图 2 的 4 种风格化结果中,只有本文的方法 Our 将这种密集线条型纹理生成出来;而其他 3 种方法尽管生成了清晰的风格化图像,但是其纹理和“牛毛皴”纹理相差甚远。在内容结构上,Our 与其他 3 种方法都生成了较为清晰的山体结构,但是从水墨画图像的角度出发,这 3 种方法所生成的图像边缘过于锋利并且与背景天空的区分过于鲜明,存在较为明显的“照片感”;而 Our 的边缘以及与背景的组合则有更加真实的水墨画质感。从图像内山体的色彩进行分析,Our 的山体色彩与风格图片基本一致,而另外 3 种结果的山体色彩

过淡,与风格图像中黝黑的山体色彩不一致。这是因为 Safin、Micro 以及 EFDM 都是基于特征融合实现图像生成的,这种方式需要同时协调内容特征以及风格特征,从而在很大程度上制约了风格纹理的生成,因而难以生成真实性较高的纹理模式(尤其是水墨画山体纹理)。本文方法是通过迭代去噪,从噪声数据中逐步生成内容结构与纹理细节,而预训练的 Unet 去噪网络已经事先学习到真实水墨画所具备的域特征,所以本文提出的方法在生成真实水墨画纹理方面体现出较大的优势。

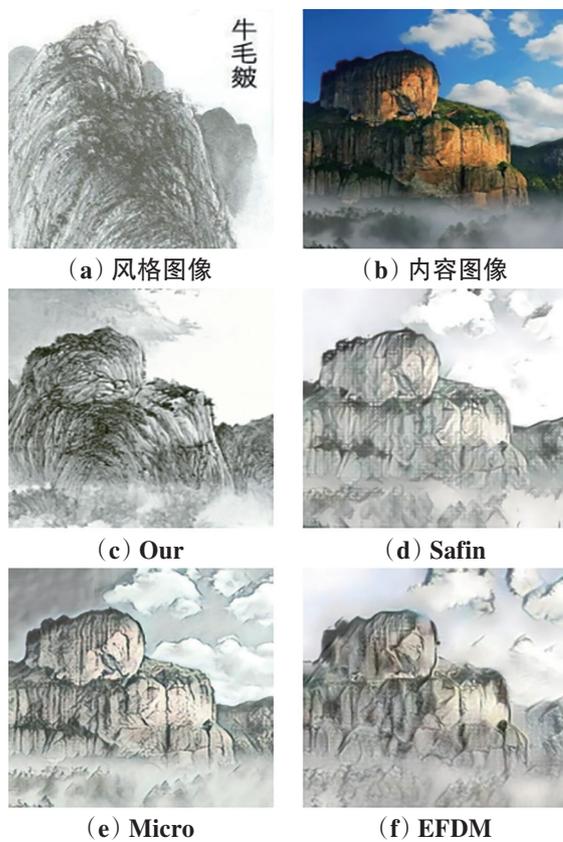


图 2 风格化视觉效果定性对比图

Fig. 2 Qualitative comparison of stylized visual effects

2.3 定量对比实验

为了进一步定量对比模型的性能,选择 ViT 内容损失 l_{vite} 与风格色彩损失 l_{col} 作为评价指标,观察其在内容保持与色彩迁移亮度方面的表现。随机选择 18 张水墨风格图像以及 6 张真实内容图像构成若干组合作为各模型的输入,计算各模型风格化输出在指标上的平均表现,比较结果如表 1 所示。 l_{vite} 衡量的是风格化图像与内容图像之间内容结构的差异程度,损失越小,说明风格化图像的内容保真度越高; l_{col} 衡量风格化图像与风格图像之间色彩亮度的差异,损失越小,说明两者亮度差异越小。

表 1 定量对比结果

Table 1 Results of quantitative comparison

指 标	风格迁移模型			
	Our	Safin	Micro	EFDM
l_{vite}	0.050	0.055	0.053	0.069
l_{col}	25.4	6.5	8.5	21.3

从表 1 可以看到:在指标 l_{vite} 上,本文提出的方法数值最低,表现最佳,相较于另外 3 种模型表现出最好的内容结构保持能力。指标 l_{col} 计算的是图像整体明暗程度的差异。在这一点上,本文方法表现不佳,这说明在整体色彩表达方面,本文工作有进一步提升的空间。这是因为在扩散模型逐步去噪生成图像的过程中,模型会根据含噪数据自身特征调整图像,与损失函数的优化方向产生一定程度的偏移,从而降低了模型在色彩迁移方面的性能。而另外 3 种基于特征合成的生成图像,在风格特征中包含了直接的色彩信息,这使得它们能够轻松生成色彩更为准确的风格化图像。需要注意的是, l_{col} 通过均值与方差两个数值反映了风格化图像与目标风格图像之间的亮度差异,而不是具体的水墨色彩分布差异,因此无法证明整体风格色彩迁移性能的好坏,只有有限参考价值。尽管米志鹏^[20]的工作使用用户满意度与运行时间作为定量评价指标,但却因其主观性过强,仍然无法实现通过定量评价结果充分比较模型性能优劣这一目的。因此,如何在采样算法中加强损失函数的引导作用,从而生成更加接近目标风格色彩的风格化图像,是未来的一个重点研究方向。更进一步,探索出能够充分合理证明风格迁移模型整体性能的定量评价指标也是格外重要的一个研究方向。

2.4 消融实验

为了进一步证明在优化 ViT 内容损失的同时优化 VGG 内容损失(l_{vggc})的必要性,本文针对 l_{vggc} 进行了消融,实验结果如图 3 所示。图中可以看到:Our 的生成结果充分保留了山体原本的崎岖特征,湖中倒影也与原图十分相近。而在不使用 l_{vggc} 的情况下,生成图像的轮廓变得过于圆滑,湖中的倒影也与原图相比发生了较大变化,存在较为明显的内容失真。这是因为 ViT 内容损失包含的内容结构信息不足,导致其在内容指导方面的引导作用不够,所以额外使用 VGG 内容损失作为补充,能够加强两者共同的内容信息在采样过程中的指导作用,从而生成内容结构更加准确的结构化图像。



图 3 不使用 VGG 的内容损失

Fig. 3 Without VGG content loss

本文也对风格色彩损失进行了比较,将均值方差色彩损失替换为 L_1 色彩损失,结果如图 4 所示。从图中可以看到:使用均值方差色彩损失的 Our 生成了干净的天空以及清晰的山体结构;而在使用 L_1 损失时,所生成风格化图像错误地受到风格图像内容的影响,使得右上原本干净的天空变得格外复杂,内容发生较大变化。这是因为 L_1 损失的计算方式是像素级的,这会促使风格图像与风格化图像在所有像素上一致,从而引入不必要的风格图像内容信息。而本文所用的是色彩均值与方差,是统计信息,能够避免 L_1 损失所带来的负面作用。然而该统计信息只是两个数字,很难较为全面地指导图像的色彩迁移。



图 4 替换为 L1 的色彩损失

Fig. 4 Replacement with L1 color loss

3 结 论

本文研究了基于扩散模型的中国水墨画风格迁移方法,提出一个用于引导风格化图像生成的采样算法。提出的采样算法包括内容损失函数、风格损失函数以及语义分离损失,这些损失函数从内容以及风格两个

角度共同引导网络模型生成目标水墨风格化图像。实验从定性以及定量两个角度将本文方法的生成结果与其他模型的生成结果进行对比,结果显示:所提出的方法能够在保持内容结构的基础上生成真实感强的水墨画纹理,在中国水墨画风格迁移领域具有优越的性能。消融实验证明了采样算法中所用损失的合理性与必要性。本方法进一步证明了在风格迁移任务中使用基于扩散模型的方法能够生成高质量的风格化图像,对未来的相关工作提供了一定的参考价值。另外,本文所提出的方法存在损失引导约束力不足的问题,在未来的工作中,研究者将进一步优化采样算法,提升色彩迁移的精准度以及风格化图像的视觉效果。

参考文献(References):

- [1] GATYS L A, ECKER A S, BETHGE M. Image style transfer using convolutional neural networks[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2016: 2414–2423.
- [2] ULYANOV D, LEBEDEV V, VEDALDI A, et al. Texture networks: feed-forward synthesis of textures and stylized images[C]//Proceedings of the 33rd International Conference on Machine Learning. New York: ACM, 2016: 1349–1357.
- [3] JOHNSON J, ALAHI A, LI F F. Perceptual losses for real-time style transfer and super-resolution[C]//Computer Vision-ECCV 2016. Cham: Springer, 2016: 694–711.
- [4] HUANG X, BELONGIE S. Arbitrary style transfer in real-time with adaptive instance normalization[C]//Proceedings of the IEEE International Conference on Computer Vision. Piscataway: IEEE Press, 2017: 1510–1519.
- [5] LI Y, FANG C, YANG J, et al. Universal style transfer via feature transforms[C]//Proceedings of the 31st International Conference on Neural Information Processing Systems. New York: ACM, 2017: 386–396.
- [6] ZHANG Y, TANG F, DONG W, et al. Domain enhanced arbitrary image style transfer via contrastive learning [C]//Proceedings of the Special Interest Group on Computer Graphics and Interactive Techniques Conference Proceedings. 2022: 1–8.
- [7] CHEN Y, LAI Y K, LIU Y J. CartoonGAN: generative adversarial networks for photo cartoonization[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2018: 9465–9474.
- [8] CHEN J, LIU G, CHEN X. AnimeGAN: a novel lightweight GAN for photo animation[C]//Artificial Intelligence Algorithms and Applications. Springer: Springer Singapore, 2020: 242–256.
- [9] KOTOVENKO D, WRIGHT M, HEIMBRECHT A, et al. Rethinking style transfer: From pixels to parameterized brushstrokes[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2021: 12191–12200.
- [10] XUE A. End-to-end Chinese landscape painting creation using generative adversarial networks[C]//Proceedings of the IEEE Winter Conference on Applications of Computer Vision. Piscataway: IEEE Press, 2021: 3863–3871.
- [11] HO J, JAIN A, ABBEEL P. Denoising diffusion probabilistic models[C]//Advances in Neural Information Processing Systems, 2020, 33: 6840–6851.
- [12] CHUNG H, SIM B, RYU D, et al. Improving diffusion models for inverse problems using manifold constraints[C]//Proceedings of the 36th International Conference on Neural Information Processing Systems. New York: ACM, 2022, 35: 683–696.
- [13] AVRAHAMI O, LISCHINSKI D, FRIED O. Blended diffusion for text-driven editing of natural images[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2022: 187–197.
- [14] SAHARIA C, CHAN W, SAXENA S, et al. Photorealistic text-to-image diffusion models with deep language understanding[J]. Advances in Neural Information Processing Systems, 2022, 35: 36479–36494.
- [15] KWON G, YE J C. Diffusion-based Image Translation using disentangled style and content representation[C]//The Eleventh International Conference on Learning Representations. 2022.
- [16] TUMANYAN N, BAR-TAL O, BAGON S, et al. Splicing ViT features for semantic appearance transfer[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2022: 738–747.
- [17] SINGH A, HINGANE S, GONG X, et al. SAFIN: Arbitrary style transfer with self-attentive factorized instance normalization[C]//Proceedings of the IEEE International Conference on Multimedia and Expo. Piscataway: IEEE Press, 2021: 1–6.
- [18] WANG Z, ZHAO L, ZUO Z, et al. MicroAST: Towards super-fast ultra-resolution arbitrary style transfer[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2023, 37(3): 2742–2750.
- [19] ZHANG Y, LI M, LI R, et al. Exact feature distribution matching for arbitrary style transfer and domain generalization[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE Press, 2022: 8025–8035.
- [20] 米志鹏. 基于风格特征分布匹配的图像风格迁移[J]. 重庆工商大学学报(自然科学版), 2023, 40(2): 51–56.
MI Zhi-peng. Image style transfer based on the distribution matching of the style features[J]. Journal of Chongqing Technology and Business University (Natural Science Edition), 2023, 40(2): 51–56.

责任编辑:李翠薇