

中图法分类号: TP391.41 文献标识码: A 文章编号: 1006-8961(2025)05-1377-12

论文引用格式: Zhang C, Zhang X J, Zhao Y, Lu J Y and Xie Y Q. 2025. Interactive Thangka image restoration guided by line art. Journal of Image and Graphics, 30(5):1377-1388(张驰, 张效娟, 赵洋, 卢嘉钰, 谢钰麒. 2025. 线稿引导的交互式唐卡图像修复. 中国图象图形学报, 30(5): 1377-1388)[DOI:10.11834/jig.240399]

线稿引导的交互式唐卡图像修复

张驰^{1,2,3}, 张效娟^{1,2,3*}, 赵洋⁴, 卢嘉钰^{1,2,3}, 谢钰麒^{1,2,3}

1. 青海师范大学计算机学院, 西宁 810016; 2. 藏语智能全国重点实验室, 西宁 810016;
3. 青海省非物质文化遗产研究基地, 西宁 810016; 4. 合肥工业大学计算机与信息学院, 合肥 230601

摘要: 目的 唐卡作为人类非物质文化遗产热贡艺术的重要表现形式之一, 承载着重要的历史文化价值。在实地采集过程中发现, 由于保存条件有限, 许多唐卡作品出现裂痕、破损、水渍及霉点等问题, 传统手工的修复方式效率低, 且存在导致唐卡二次受损的风险。此外, 使用传统图像修复方法和基于深度学习的图像自动修复方法修复唐卡时, 往往产生不符合唐卡纹理结构的结果。鉴于此, 提出一种线稿引导的交互式唐卡图像修复网络 LSFNet (an image restoration network that combines line restoration, style and texture restoration, and fine restoration)。方法 该方法由3部分组成, 首先是唐卡艺术家指导的交互式线稿修复, 使得修复的线结构更加接近真实唐卡图像; 其次是风格纹理修复阶段, 通过构建空间风格纹理模块学习唐卡图像整体风格和特征, 并结合通道注意力和全连接层, 捕获全局信息并进行融合, 获得初步修复特征; 最后是精修复阶段, 引入线性注意力模块, 实现全局信息传递, 增强模型对唐卡图像内容的理解能力。结果 以在青海采集的唐卡图像为基础, 创建了唐卡修复数据集, 并通过模拟破损区域, 制作了掩码数据集, 在创建的数据集上进行训练测试。与 DeepFillv2、EdgeConnect、DFNet (deep fusion network)、HiFill 及 T-Former 等图像修复方法进行定量、定性和主观实验对比分析。结果表明, 该方法有良好的修复效果, 在唐卡数据集上的 PSNR (peak signal to noise ratio)、SSIM (structural similarity) 和 LPIPS (learned perceptual image patch similarity) 3个评价指标结果均优于对比方法。与性能第2的模型相比, PSNR 和 SSIM 分别提高 10.55% 和 1.8%, LPIPS 降低 57.98%。此外, 消融实验进一步验证了交互式线稿修复、风格纹理修复和精修复3个模块的有效性。结论 通过采用交互式线稿修补的方法, 能够有效地对破损唐卡图像进行修复, 获得符合唐卡内容风格的修复结果。

关键词: 热贡唐卡; 图像修复; 交互式图像修复; 风格纹理修复; 两阶段交互修复

Interactive Thangka image restoration guided by line art

Zhang Chi^{1,2,3}, Zhang Xiaojuan^{1,2,3*}, Zhao Yang⁴, Lu Jiayu^{1,2,3}, Xie Yuqi^{1,2,3}

1. School of Computer, Qinghai Normal University, Xining 810016, China; 2. State Key Laboratory of Tibetan Intelligence, Xining 810016, China; 3. Qinghai Province Intangible Cultural Heritage Research Base, Xining 810016, China;
4. School of Computer and Information, Hefei University of Technology, Hefei 230601, China

Abstract: Objective Regong art, originating from the Longwu River valley in the Tibetan region of Huangnan, Qinghai Province, has flourished in this area, forming a unique regional artistic style. In 2009, this art was inscribed on the

收稿日期: 2024-07-18; 修回日期: 2024-10-16; 预印本日期: 2024-10-23

* 通信作者: 张效娟 zhxj@qhnu.edu.cn

基金项目: 青海省重点研发与转化计划项目(2021-GX-111); 国家自然科学基金项目(62262056)

Supported by: Key R&D and Transformation Plan Project in Qinghai Province (2021-GX-111); National Natural Science Foundation of China (62262056)

UNESCO Representative List of the Intangible Cultural Heritage of Humanity. Thangka, as one of the most important forms of Regong art, embodies the rich historical and cultural heritage of the Tibetan region, holding substantial historical, cultural, and artistic value. During the field collection process, many Thangka works displayed issues such as cracks, tears, water stains, and mold spots due to poor preservation conditions. However, traditional restoration methods are not only inefficient but also risk causing further damage, making them unsuitable for the proper conservation and development of Regong art. Consequently, conducting research on the restoration of damaged Thangka images is urgently needed. However, attempts to restore Thangka images using current enhancement and restoration algorithms have encountered several issues, such as blurred texture lines and misaligned repairs. These issues arise because the complexity and diversity of Thangka images make it difficult for existing models to capture their unique structural and textural characteristics. **Method** Therefore, an interactive Thangka image restoration network, LSFNet, guided by line draft repair, is proposed to address the aforementioned challenges. This method comprises three parts. First, the interactive line restoration involves collaboration with Thangka artists to guide the restoration of the line structure, ensuring that the restored lines closely resemble those in real Thangka images. Second, the style and texture restoration phase is where an overall style and texture module is constructed to learn the unique characteristics of Thangka images. By integrating channel attention mechanisms and fully connected layers, the module captures global information and synthesizes it into preliminary restoration features. Finally, the refinement restoration phase introduces a linear attention module during the downsampling process. This module captures local and global dependencies, allowing the model to extract features of different scales, further refining the restoration, eliminating restoration traces, and enhancing the overall image quality. PatchGAN is also adopted as the discriminator, dividing the input image into multiple receptive fields and conducting independent binary classification for each receptive field to assess whether it matches the texture characteristics of the target image. This approach effectively enables pixel-level supervision, enhancing the overall image restoration quality. **Result** This paper created a Thangka restoration dataset comprising a total of 25 000 images, collected through field research and data gathering. The Canny algorithm was employed to extract edge line art, resulting in a line art dataset. Additionally, Photoshop tools were used to simulate damage to the Thangka images, generating 5 000 mask maps. Another 1 000 mask maps were sourced from public datasets to enhance restoration performance under various damage conditions, combining them into a final dataset containing 6 000 mask maps. All datasets have a resolution of 256×256 pixels. The proposed method in this paper was trained, tested, and compared with other restoration methods, including DeepFillv2, EdgeConnect, DFNet, HiFill, and T-Former, using the dataset created in this study. Results indicate that the proposed method exhibits strong repair performance, with superior metrics in peak signal-to-noise ratio (PSNR), structural similarity (SSIM), and learned perceptual image patch similarity (LPIPS) on the Tangka dataset compared to other methods. Specifically, compared to the second-best performing model, the proposed method achieved a 10.55% increase in PSNR, a 1.8% increase in SSIM, and a 57.98% reduction in LPIPS. Experimental results demonstrate that the proposed interactive Thangka image restoration method, based on line art repair, can effectively restore damaged Thangka images, producing results that are closer to authentic Thangka images. **Conclusion** This article proposes an interactive line drawing repair method, guided by Thangka artists, to repair damaged areas of the line drawing. Subsequently, a style and texture restoration phase is employed to learn the distinctive style features of Thangka images. Finally, a fine-tuning repair process further optimizes the restoration results. Experimental results demonstrate that this method effectively repairs damaged Thangka images, producing restoration outcomes that conform to the style and content of Thangka art.

Key words: Regong Thangka; image inpainting; interactive image restoration; style texture restoration; two stage interactive repair

0 引言

热贡艺术作为藏传佛教艺术的重要分支,经过

历史长河的洗礼与磨砺,逐渐形成别具一格的地域艺术风格,在民间美术和建筑艺术等领域有着深厚的历史底蕴和文化价值,于2009年列入联合国教科文组织人类非物质文化遗产代表作名录。

实地调研发现,由于采用天然矿物质颜料,若存放不当,唐卡表面极易出现裂纹;并且对于一些年代久远的寺院,在阴雨天容易出现渗水现象,导致雨水滴到唐卡上,若处理不及时,可能会导致唐卡出现水渍、掉色、发霉等问题,如图1所示。对于这些破损、污染的唐卡,当地人通常采用传统手工修复的方法。修复工作主要依据唐卡上未破损区域的内容纹理,并结合画师的经验进行。然而,能够胜任这一工作的画师屈指可数,因为这不仅需要长时间的经验积累,而且修复效率极低,同时存在一定风险,容易导致唐卡二次受损。

以上问题都严重阻碍了热贡艺术的传播与发展。此外,唐卡承载着热贡地区的历史和文化,记录了当地人生活的方方面面,通过对唐卡的研究,可以更好地了解、保护和传承热贡文化。随着计算机视觉的不断发展,可以采用数字化的方式来对唐卡图像进行修复,实现对热贡艺术的保护和传承。

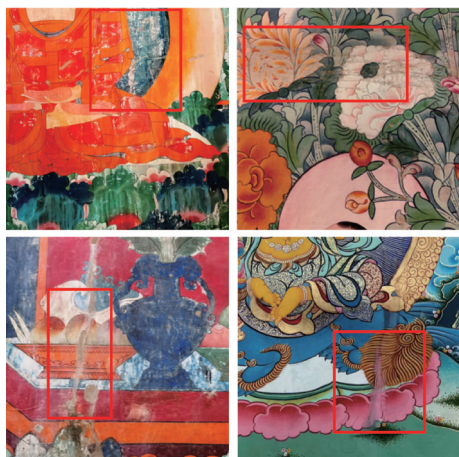


图1 破损污染的唐卡

Fig. 1 Damaged and contaminated Thangka images

传统图像修复的方法主要基于补丁和像素扩散两种方法实现,这两种方法通常是依据破损区域周围的像素对破损区域进行填补,在修复时容易发生修复内容不合理、纹理模糊等问题。

基于深度学习的图像自动修复主要分为3类:1)基于卷积神经网络的方法(Pathak等,2016;Gao和Grauman,2017;Liu等,2018;Yang等,2017;江奎等,2024)。该类方法通常采用编码—解码结构修复图像缺失部分,编码器负责提取图像特征,解码器负责将提取的特征进行重建。2)基于生成对抗网络(generative adversarial network, GAN)的方法(Yu等,

2019;Nazeri等,2019;Hong等,2019;Yi等,2020;兰治等,2023)。GAN网络主要包括一个生成模型G和一个判别模型D。生成模型G通过反卷积神经网络将随机输入转化为图像,判别模型D用于评估生成的图像与真实图像的相似度。在图像修复中,GAN网络可以生成与周围内容相匹配的图像部分,从而修复缺失或损坏的区域。这种方法在低分辨率图像或特定类型的图像修复中可取得较好效果,但对高分辨率图像的修复可能存在困难。3)基于Transformer网络的方法(Deng等,2022;Chen等,2022;刘雨轩等,2023)。这种方法的核心思想是利用Transformer网络结构中的自注意力机制捕捉图像中的长距离依赖关系,从而更准确地修复缺失或损坏的图像区域。

破损区域的复杂纹理结构的重建是图像修复中极具挑战性的问题,因此,研究人员利用手动绘制破损区域纹理结构的方式,指导网络进行修复。Zeng等人(2022)提出SketchEdit方法,用户通过直接在图像上绘制草图,指导网络进行图像局部编辑。Jo和Park(2019)利用SC-FEGAN(face editing generative adversarial network with user's sketch and color)网络,结合用户草图和颜色信息,实现面部图像的精细编辑,通过学习和理解用户输入,自动调整人脸特征,生成高质量、逼真的编辑结果。以上两种方法都聚焦于图像编辑任务,并通过简化用户交互和提升编辑效果来解决现有方法中的不足。SketchEdit通过无掩码编辑和自监督训练实现了更简单的用户工作流程和更好的编辑效果;而SC-FEGAN则通过引入多输入信息和优化网络结构,实现了更精细和高质量的人脸图像编辑。两者在解决思路上的共性在于都注重用户输入的多样性和网络结构的优化,以提升图像编辑的灵活性和效果。但这两种算法很大程度上依赖于输入的草图质量和颜色信息。如果草图过于模糊、不精确或包含大量噪声,可能会影响模型对修改区域的预测和最终编辑效果,且在某些复杂场景下,模型可能无法准确识别或编辑所有期望的区域。

在唐卡图像修复领域,多个研究团队进行了深入探索。王维兰团队专注于破损区域的分割提取(胡文瑾等,2016)与修复(卢小宝和王维兰,2010),通过精确分割与算法创新,实现了破损唐卡的有效修复。刘华明团队采用图像分割技术(刘华明和毕学慧,2013a,b),针对褪色和破损区域进行修复研

究。王威(2016)的研究结合了人眼对颜色的感知,在 Lab 颜色空间内进行修复,提升了图像修复质量。任家乐(2022)采用基于 U-Net 的修复网络结合多尺度注意力机制,充分利用特征信息提高修复效果。刘洁(2020)通过引入空洞卷积减少上下采样次数并取消跳跃连接,减少信息损失。刘丹青(2021)将 Patch Match 与部分卷积方法通过迁移学习应用于唐卡修复。廖年鸿(2022)和廖年鸿等人(2022)提出基于参数与运算轻量化的图像修复模型,通过深度可分离卷积与注意力机制结合,显著减少参数运算,并提升了修复效果。这些研究为唐卡图像修复提供了多种有效的解决方案。

采用目前流行的修复算法对唐卡图像进行修复之后,会出现纹理线条模糊、修复错位等现象。唐卡图像具有丰富的构图和纹理信息,同时蕴含丰富的文化信息,这使得模型无法学习到其独特的结构和纹理特征,难以保证修复结果的合理性。此外唐卡图像破损程度不一,这也要求修复模型需要应对不同类型的破损。针对上述问题,本文提出一个唐卡图像修复网络,该修复网络分为两个阶段,第1阶段为交互式掩码绘制及线稿修补模块;第2阶段分为风格纹理修复部分和精修复部分,其中风格纹理修复部分加入了空间风格模块,用来学习唐卡图像整体风格及纹理特征,进行初步修复;精修复部分引入线性注意力机制,进一步消除修复痕迹,提升图像修复效果。最后基于创建的数据集进行了训练测试和对比实验分析,验证了模型的实用性,并进行消融实验来验证各模块的有效性,结果表明,添加线稿指导

后,修复结果的纹理特征更接近真实唐卡图像,优于其他 SOTA(state of the art)方法。

本文主要工作和贡献如下:1)由于全自动修复方法无法准确修复纹理结构,本文通过唐卡艺术家的专业指导,对破损唐卡图像的线稿进行修补,在修补完整的线稿基础上进行修复。2)提出一种风格纹理修复和精修复结合的修复网络,并取得良好的修复效果。其中风格纹理修复中的空间风格注意力机制能够有效学习唐卡图像的风格和纹理,对破损区域进行初步修复。精修复网络能有效捕获局部与全局依赖关系,解决不同尺度特征间关联问题,实现全局信息传递,增强模型对唐卡图像内容的理解能力。3)采用在青海省黄南州同仁县实地调研采集到的唐卡图像,构建了唐卡修复数据集 25 000 幅,通过 Canny 算法提取边缘,创建了对应的唐卡线稿数据集,通过模拟破损污染区域,创建了掩码数据集 5 000 幅。

1 本文方法

本文提出的修复网络 LSFNet(an image restoration network that combines line restoration, style and texture restoration, and fine restoration)包含以下 3 个模块:线稿修补、风格纹理修复和精修复,如图 2 所示。第 1 个模块为线稿修补,根据唐卡破损区域绘制掩码,生成对应线稿图,并在唐卡非遗传承人或工艺美术师等唐卡艺术家的指导下,对线稿破损区域进行修补;第 2 个模块为风格纹理修复,以恢复唐卡

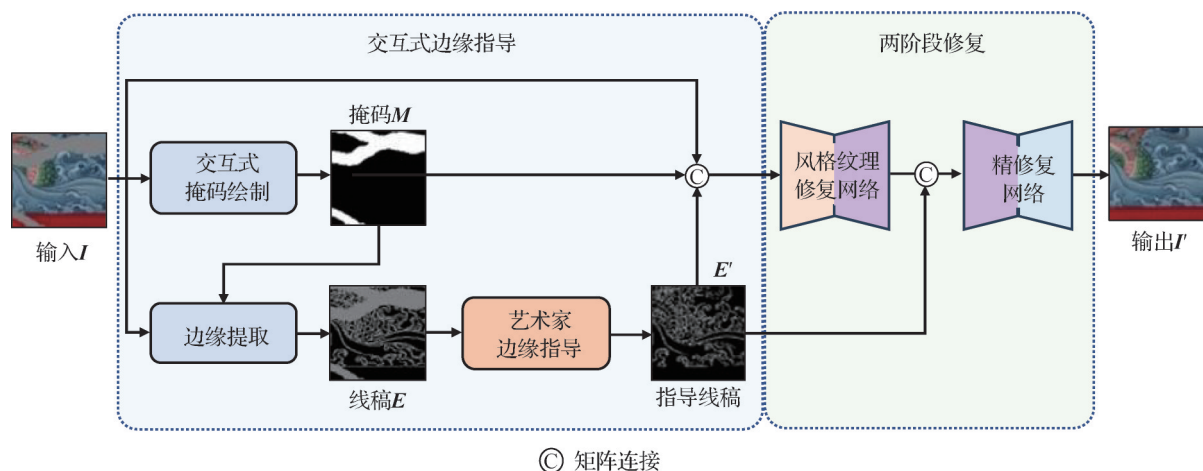


图2 LSFNet网络整体结构图

Fig. 2 Overall structure diagram of LSFNet network

和纹理,对破损区域进行初步修复,不仅能恢复轮廓,还能提高模型性能和精度。结合通道注意力和全连接层,能捕获全局信息并融合特征,更精准预测掩码区域内容,提升修复效果。通过该网络得到初步修复特征 I_1 ,为后续处理奠定良好基础。

1.2.2 精修复

精修复部分网络结构如图4所示。首先将线稿图 E' 与初步修复特征 I_1 进行特征融合,再输入至精修复模块,以提升修复质量,使纹理细腻,并消除修复痕迹,得到修复结果 I_2 。下采样阶段添加 mGAtten 模块(Deng 等,2022),该模块采用线性注意力机制,能有效捕获局部与全局依赖关系,解决了不同尺度特征间关联的问题,实现全局信息传递,增强了模型对唐卡图像内容的理解能力。这种结合不仅保留了更多图像细节和上下文信息,还提升了模型鲁棒性,允许模型自适应调整特征关联程

度,学习更有用的特征表示,进一步提升修复效果。另外,编码部分则设计由深度可分离卷积和普通卷积相结合,这种混合方式显著减少了计算量,从而加快了模型的计算速度,使得整个修复过程更加高效。

随着网络深度的增加,模型表达能力增强,但易遭遇梯度消失或梯度爆炸问题。为此,在跳跃连接中添加残差块,实现跨层直接连接,有效缓解这些问题,避免深度网络性能下降。残差块不仅确保信号顺畅传递,还使网络能融合不同尺度下的图像信息,获得更丰富准确的图像表征。残差学习有助于融合多层次特征,对处理复杂的唐卡图像特征至关重要。此外,残差块的引入还提升了模型性能,加快了训练速度,使网络在修复唐卡图像时更高效学习和优化。最终得到的修复图像 I' 表示为

$$I' = I \times M + I_2 \times (1 - M) \quad (3)$$

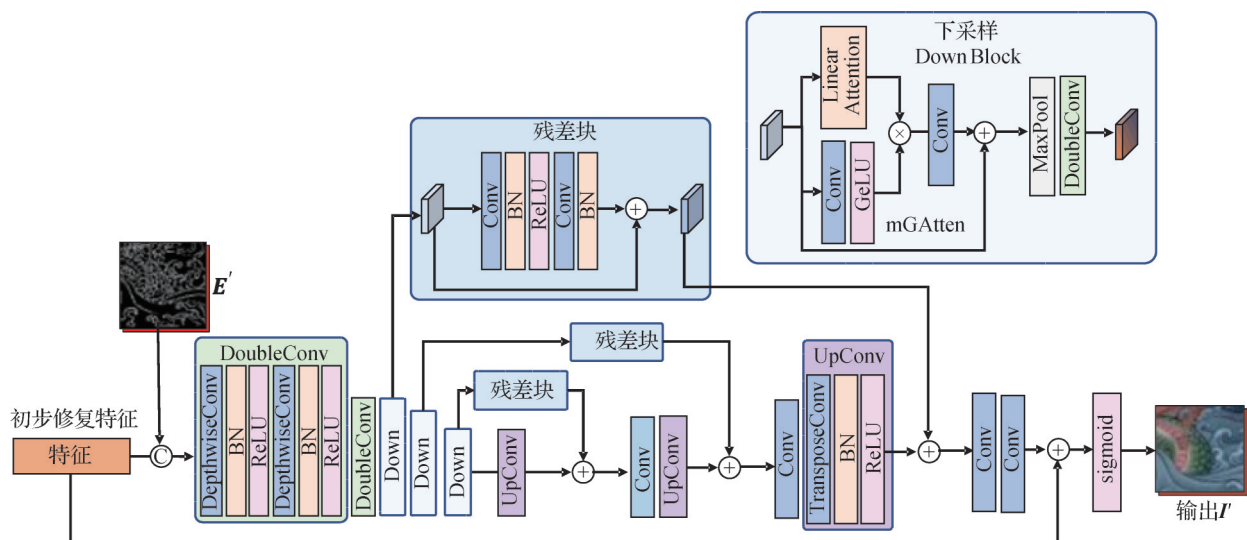


图4 精修复结构图

Fig. 4 Fine repair structure diagram

1.2.3 判别器

判别器部分选择 PatchGAN(Isola 等,2017)作为核心,这是一种基于卷积神经网络的判别模型,专为图像生成和风格迁移任务设计。PatchGAN 完全由卷积层构成,其输出是一个 $n \times n$ 矩阵, n 为矩阵长度,每个元素代表输入图像中对应感受野的真实或虚假概率。这种设计使 PatchGAN 成为一种精细的纹理样式损失,能够关注到图像局部纹理细节的同时,也考虑到图像的全局特征,确保生成图像的高分辨率和细腻纹理,使修复后的结果在整体风格等方面保持一致,实现良好的修复效果。此外, Patch-

GAN 能够减少运算量,提高运行速度,在不牺牲性能的前提下,加快模型的训练。最后, PatchGAN 能够适应不同类型的图像修复任务,适合唐卡图像的修复工作,可以根据绘画风格来调整参数,以获得最佳修复效果。

1.3 损失函数

1.3.1 重建损失

重建损失指的是输出图像 I_{out} 与真实图像 I_{gt} 之间的绝对差值的平均值,其表示为

$$L_{re} = \left\| I_{out} - I_{gt} \right\|_1 \quad (4)$$

1.3.2 风格损失

风格损失的核心作用在于确保修复后的图像在风格上与原始图像高度契合。利用Gram矩阵来精确衡量不同特征图之间的细微差异,从而确保图像间的风格一致性。通过细致地比较待修复图像与原始图像在多个不同层级的特征图之间的差异,确保修复后的图像在风格层面与原始图像保持一致。风格损失为

$$L_{\text{style}} = E_j \left[\left\| \mathbf{G}_j^\phi(I_{\text{out}}) - \mathbf{G}_j^\phi(I_{\text{gt}}) \right\|_1 \right] \quad (5)$$

式中, E_j 表示对不同的特征层 j 的损失进行平均, 得到平均值, \mathbf{G}_j^ϕ 表示由相应的激活函数 ϕ_j 构成的 $C_j \times C_j$ 的Gram矩阵, 具体为

$$\mathbf{G}_j^\phi(x)_{c,c'} = \frac{1}{C_j H_j W_j} \sum_{h=1}^{H_j} \sum_{w=1}^{W_j} \phi_j(x)_{h,w,c} \phi_j(x)_{h,w,c'} \quad (6)$$

式中, h 表示特征图高度索引, w 表示特征图宽度索引, c 和 c' 表示特征图的通道索引, C_j, H_j, W_j 分别表示在ImageNet上预训练的VGG19(Visual Geometry Group)第 j 层上特征图的通道数、高度和宽度, $\phi_j(x)_{(h,w,c)}$ 表示在VGG19第 j 层第 (h, w, c) 位置的输出。

1.3.3 感知损失

感知损失致力于提升图像的感知质量, 而并非仅局限于像素级别的差异。在图像修复过程中, 感知损失发挥了至关重要的作用, 其能够有效地保留图像中丰富的细节和精致的纹理, 从而显著提升修复结果的视觉效果。感知损失为

$$L_{\text{per}} = E \left[\sum_i \frac{1}{N_i} \left\| \phi_i(I_{\text{out}}) - \phi_i(I_{\text{gt}}) \right\|_1 \right] \quad (7)$$

式中, N_i 表示第 i 层特征图的元素数量, ϕ_i 是在VGG19第 i 层的激活函数。

1.3.4 对抗损失

对抗损失表示为

$$L_{\text{adv}} = E_{I_{\text{gt}}} \left[\log D(I_{\text{gt}}) \right] + E_{I_{\text{out}}} \left[\log [1 - D(I_{\text{out}})] \right] \quad (8)$$

式中, $E_{I_{\text{gt}}}$ 和 $E_{I_{\text{out}}}$ 分别表示对真实图像 I_{gt} 和生成图像 I_{out} 的期望, $D(\cdot)$ 表示鉴别器的输出。

1.3.5 总损失

综上所述, 总损失函数表示为

$$L_{\text{total}} = \lambda_a L_{\text{adv}} + \lambda_b L_{\text{re}} + \lambda_c L_{\text{per}} + \lambda_d L_{\text{style}} \quad (9)$$

根据Deng等人(2022)的方法, 并结合实验结果, 最终权重设置为 $\lambda_a = 0.1, \lambda_b = \lambda_c = 1, \lambda_d = 250$ 。

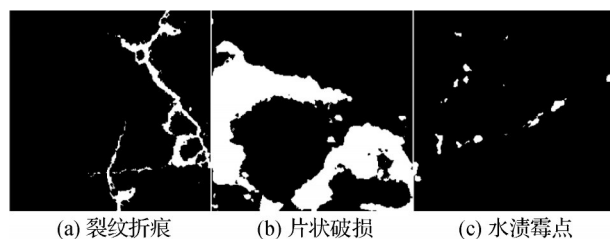
2 实验

2.1 数据集

本实验使用的唐卡数据来自在青海省黄南州同仁市实地调研采集的唐卡图像, 并从拍摄的图像中挑选300幅成像质量较高的图像, 随后进行裁剪、修正, 去除多余部分, 之后将图像剪切成尺寸为 256×256 像素, 共得到25 000幅图像, 按8:1:1的比例将这些图像分为训练集、测试集和验证集。在训练阶段, 由于数据量较大, 若逐个绘制线稿图, 工作量巨大, 且可能出现偏差, 因此采用Canny算法对裁剪后的图像进行边缘提取, 得到对应的边缘线稿数据集。在实际应用阶段, 则由唐卡专家进行指导绘制。

与普通掩码数据集不同, 唐卡图像通常存在以下几种破损、污染情况: 裂纹状折痕、点状破损、片状破损、水渍及霉点等。本文采用Photoshop模拟这几种破损形状, 制作掩码数据集并裁剪成尺寸为 256×256 像素的掩码图, 如图5所示, 共得到5 000幅掩码图。同时, 从公共掩码数据集(Liu等, 2018)中再选取1 000幅, 共计6 000幅掩码图, 按5:1随机分为训练集和测试集。

为验证本文方法的泛化性, 采用公共数据集CelebA-Dialog(Jiang等, 2021)进行泛化实验, 该数据集共包含30 000幅人脸图像, 本文随机选取15 000幅, 按8:1:1的比例将这些图像分为训练集、测试集和验证集。



(a) 裂纹折痕 (b) 片状破损 (c) 水渍霉点

图5 掩码图

Fig. 5 Masks ((a) cracks and creases; (b) flaky damage; (c) water stained mold spots)

2.2 实验设置

本实验采用有监督训练模式, 生成器学习率为 10^{-4} , 鉴别器学习率为 10^{-5} , batchsize为8, 采用Adam优化器。本实验均在云服务器上完成, 硬件环境: Intel (R) Xeon (R) Gold 6330 CPU, NVIDIA RTX

3090 GPU 24 GB; 软件环境: Python 3.8, PyTorch 1.8.1, CUDA 11.1。

采用唐卡数据集进行图像修复对比实验,对比方法包括近年来几种先进的图像修复方法: DeepFillv2 (Yu 等, 2019)、EdgeConnect (Nazeri 等, 2019)、DFNet (Hong 等, 2019)、HiFill (Yi 等, 2020) 及 T-Former (Deng 等, 2022), 以验证本文方法在唐卡修复上的可靠性和优越性。

2.3 定量比较

本实验采用了3种评价指标: 峰值信噪比 (peak signal to noise ratio, PSNR)、结构相似性 (structural similarity, SSIM) 和 LPIPS (learned perceptual image patch similarity)。PSNR 用于表示信号的最大可能功率与其表示精度的破坏性噪声功率之间的比率, 值越大, 则表示生成的图像与原图越相近。SSIM 是一种衡量两幅图像中亮度、对比度和结构3个特征相似度的指标, 值越大, 则说明图像失真越小。LPIPS 是通过比较两幅图像在深度神经网络特征空间中的距离, 来评估它们的感知相似性, 值越小, 则表示生成的图像与原图更相似。

表1和表2列出了不同方法在唐卡数据集和 CelebA-Dialog 数据集上的 PSNR、SSIM 和 LPIPS 值。通过数据分析可以发现, 在两个数据集上, 本文方法在3个指标上均取得最好结果, 证明了本文方法在唐卡数据集上获得了很好的修复效果, 同时验证了本文方法的泛化性。

表1 不同方法在唐卡数据集上的 PSNR、SSIM、LPIPS 值

Table 1 The PSNR, SSIM, and LPIPS values of different methods on the Thangka dataset

| 方法 | PSNR/dB | SSIM | LPIPS |
|-------------|-----------------|----------------|----------------|
| DeepFillv2 | 30.480 1 | 0.943 4 | 0.043 6 |
| EdgeConnect | 30.604 5 | 0.946 7 | 0.039 9 |
| DFNet | 30.362 0 | 0.941 7 | 0.053 0 |
| HiFill | 27.401 2 | 0.912 5 | 0.065 6 |
| T-Former | <u>31.475 5</u> | <u>0.951 3</u> | <u>0.037 6</u> |
| LSFNet(本文) | 34.796 7 | 0.968 5 | 0.015 8 |

注: 加粗、下划线字体分别表示各列最优、次优结果。

2.4 定性比较

图6展示了不同方法在唐卡数据集上的修复结果。可以看出, DeepFillv2 在第3、4幅图的修复中出现明显模糊、纹理不清晰现象, 在第5幅图中出现明

表2 不同方法在 CelebA-Dialog 数据集上的 PSNR、SSIM、LPIPS 值

Table 2 The PSNR, SSIM, and LPIPS values of different methods on the CelebA-Dialog dataset

| 方法 | PSNR/dB | SSIM | LPIPS |
|-------------|-----------------|----------------|----------------|
| DeepFillv2 | 31.278 6 | <u>0.946 5</u> | 0.046 2 |
| EdgeConnect | 30.918 0 | 0.937 7 | 0.056 6 |
| DFNet | <u>31.563 6</u> | 0.942 6 | 0.057 9 |
| HiFill | 27.641 3 | 0.913 5 | 0.070 1 |
| T-Former | 30.820 2 | 0.944 7 | <u>0.042 2</u> |
| LSFNet(本文) | 34.117 4 | 0.966 1 | 0.015 4 |

注: 加粗、下划线字体分别表示各列最优、次优结果。

显修复痕迹。在相同图像的结果中可以看出, EdgeConnect 同样出现了修复痕迹重、细节模糊的现象, 主要由于该算法采用自动修复线稿的方法, 与自然图像不同的是, 唐卡图像边缘风格多样, 自动修复无法准确修复线稿边缘。DFNet 和 HiFill 有明显修复痕迹, 且修复部分不符合唐卡图像。T-Former 在第3、4、6幅图的修复中出现明显模糊、纹理不清晰的现象。本文方法由于采用交互式线稿修补来约束修复结果, 因此, 对比没有添加线稿约束的算法, 本文的修复结果纹理连贯、清晰, 没有明显修复痕迹, 视觉效果更佳; 由于 EdgeConnect 采用自动线稿修复的方法, 可能会产生不合理的线稿修复结果, 因而导致修复效果不佳, 本文采用交互的方式, 因此修复结果更接近真实唐卡图像。

图7展示了不同方法在 CelebA-Dialog 数据集上的修复结果。可以看出, DeepFillv2 有明显修复痕迹, EdgeConnect 和 DFNet 在第3~5幅图的修复中出现明显噪点, 在第5幅图中, 眼睛部分纹理不清晰, HiFill 和 T-Former 在修复中出现明显模糊现象, 本文方法的修复结果没有明显修复痕迹, 由于采用交互的修复方法, 使得修复后的细节能够更加接近真实图像。并且由于二阶段精修复处理, 使得修复后的图像没有明显修复痕迹, 修复区域与原图衔接自然。但注意到, 算法处理的部分细节还有待提高, 例如第4幅图中, 眼部的修复出现了颜色不一致, 这是由于目前算法交互部分仅提供准确的线条信息指导, 并没有直接引入额外的色彩指导。

2.5 主观评价

鉴于唐卡图像的特殊性, 直接使用自然图像上

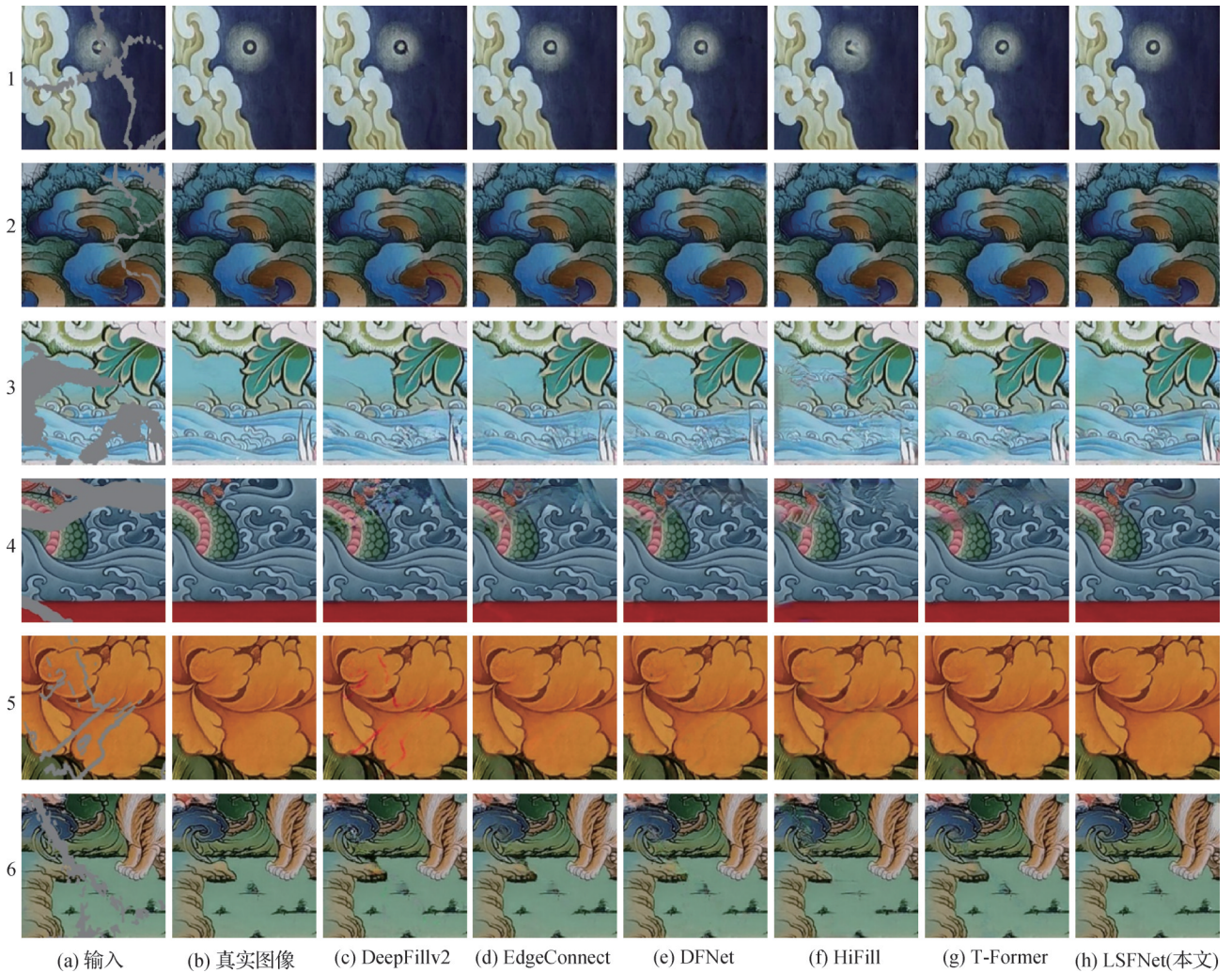


图6 唐卡数据集可视化对比结果图

Fig. 6 The visualization comparison result graphs of the Thangka dataset

((a)input; (b)ground truth; (c)DeepFillv2; (d)EdgeConnect; (e)DFNet; (f)HiFill; (g)T-Former; (h)LSFNet(ours))

的通用客观指标评判增强后的效果并不准确。本文设计了双激励主观评价实验,并邀请了3组不同层次的参与者进行评分,每组5人,实验组1为未接触过唐卡图像的参与者,实验组2为对唐卡有一定了解的参与者,实验组3为唐卡方面的专家。对修复后结果从以下3个方面进行评分并取均值:1)纹理是否连贯;2)图像结构内容是否完整;3)整体修复效果。以5分为满分,计算组内的平均得分得到每组结果。主观评价结果如表3所示,分数越高代表修复效果越好。实验结果表明,本文方法的设计是有效的,通过专业唐卡艺术家对线稿的指导绘制修复,使得本文的唐卡修复结果在纹理、结构和完整性上明显优于其他方法。

2.6 消融实验

为了验证交互式线稿修补、风格纹理修复及精

修复3个模块的有效性,本文进行消融实验,分别在去除交互式线稿修补、去除风格纹理修复和去除精修复3种情况进行实验。消融实验在同一参数、硬件和软件条件下进行。

表4为消融实验的定量结果,显示了3种情况下唐卡数据集上的PSNR、SSIM和LPIPS结果。添加交互式线稿修复后,PSNR和SSIM分别提升2.5482 dB和0.013, LPIPS降低50%;添加风格纹理修复后,PSNR和SSIM分别提升0.9173 dB和0.0041, LPIPS降低23.8%;添加精修复后,PSNR和SSIM分别提升4.9254 dB和0.021, LPIPS降低66.7%。从实验结果可以看出,添加交互式线稿修补、风格纹理修复及精修复之后,修复后的图像质量有明显提升。

图8显示了不同组合的修复结果图。可以发

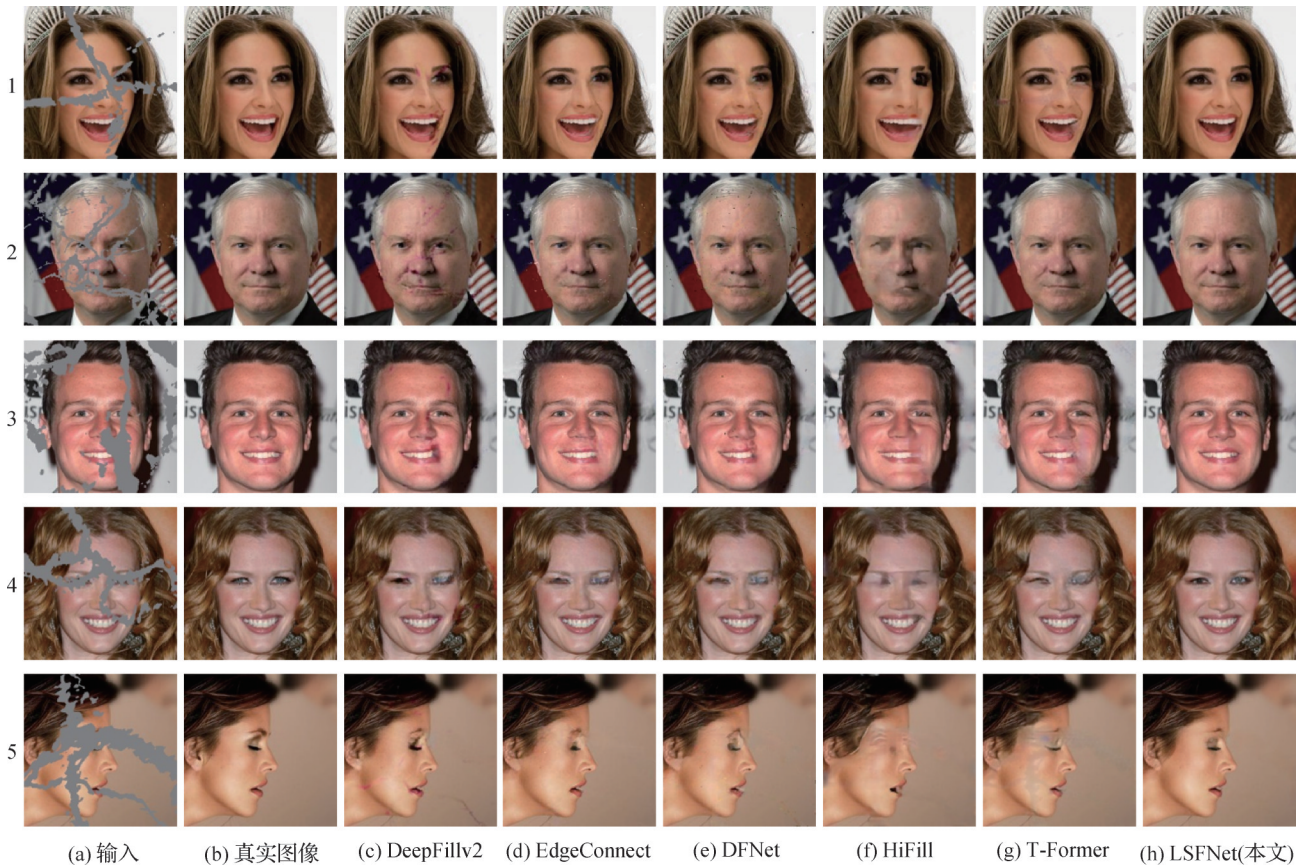


图7 CelebA-Dialog数据集可视化对比结果图

Fig. 7 The visualization comparison result graphs of the CelebA-Dialog dataset

((a)input; (b)ground truth; (c)DeepFillv2; (d)EdgeConnect; (e)DFNet; (f)HiFill; (g)T-Former; (h)LSFNet(ours))

表3 主观评价结果

Table 3 Subjective evaluation results

| 方法 | 实验组 1 | 实验组 2 | 实验组 3 | 加权平均 |
|-------------|-------|-------|-------|--------------|
| DeepFillv2 | 3.20 | 3.15 | 3.00 | 3.065 |
| EdgeConnect | 3.40 | 3.40 | 3.25 | 3.310 |
| DFNet | 3.15 | 3.20 | 3.05 | 3.105 |
| HiFill | 3.10 | 3.00 | 2.85 | 2.920 |
| T-Former | 3.60 | 3.65 | 3.55 | <u>3.585</u> |
| LSFNet(本文) | 3.95 | 3.80 | 3.85 | 3.845 |

注:加粗、下划线字体分别表示最优、次优结果。

现,未添加线稿约束,修复结果有明显的模糊、纹理不清晰等现象,而添加之后,由于线稿约束,修复的结果纹理清晰,且不存在不符合唐卡图像的其他特征。线稿加风格纹理修复的方法则有明显的类似掩码形状的修复痕迹,但纹理已接近真实唐卡图像。而线稿加精修复则没有明显痕迹,但部分纹理特征却不及线稿加风格纹理修复的方法。由此可以发现,添加的线稿可以约束修复结果,使得其纹理特征更接近真实唐卡纹理,风格纹理修复部分可以学习

表4 消融实验定量结果

Table 4 Quantitative results of ablation experiments

| 实验编号 | 交互式线稿修补 | 风格纹理修复 | 精修复 | PSNR /dB | SSIM | LPIPS |
|------|---------|--------|-----|-----------------|----------------|----------------|
| 1 | - | √ | √ | 32.248 5 | 0.955 5 | 0.032 4 |
| 2 | √ | √ | - | 29.871 3 | 0.947 5 | 0.048 3 |
| 3 | √ | - | √ | 33.879 4 | 0.964 4 | 0.020 5 |
| 4 | √ | √ | √ | 34.796 7 | 0.968 5 | 0.015 8 |

注:加粗字体表示各列最优结果,“√”表示添加该模块,“-”表示未添加。

到唐卡的风格纹理特征,精修复则进一步提高了修复质量,消除修复痕迹,3个模块的结合使得最终修复好的唐卡图像内容纹理清晰,没有修复痕迹,从而验证了各模块的有效性。

3 结论

现有的传统图像修复方法和基于深度学习的图

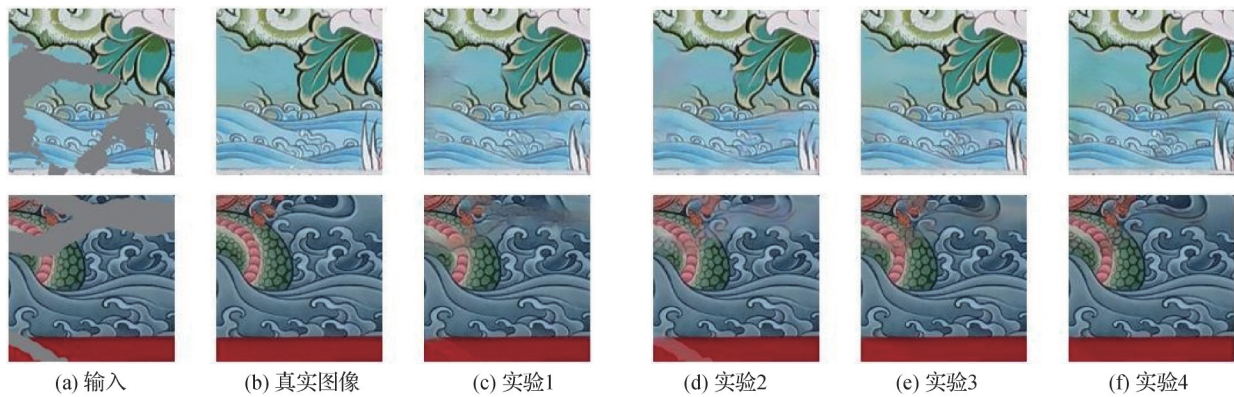


图8 消融实验可视化结果图

Fig. 8 Visualization results of ablation experiments

((a) input; (b) ground truth; (c) experiment 1; (d) experiment 2; (e) experiment 3; (f) experiment 4)

像自动修复方法在修复唐卡时,往往产生不符合唐卡纹理结构的不合理结果。唐卡图像丰富的构图和纹理信息使得模型无法学习到其独特的结构和纹理特征,难以保证修复结果的合理性。因此,本文设计了一个基于线稿修补的交互式唐卡图像修复网络。在唐卡艺术家的专业指导下,对破损线稿进行精准修复,确保线稿结构的完整性,保留其独特的艺术风格。添加空间风格模块的风格纹理修复网络,学习唐卡图像整体风格及纹理特征,能对破损唐卡图像进行初步修复。通过加入线性注意力模块的精修复网络,对修复结果进一步细化,达到了更佳的图像修复效果。实验对比与消融实验均验证了本文方法的优越性,不仅能有效恢复破损唐卡,还确保了修复效果的艺术性与真实性,修复效果更加接近真实唐卡图像,对于非物质文化遗产保护与传承具有重大意义。

但目前本文方法仍存在一些不足,在未来的研究工作中还需不断改进。在数据集构建方面,唐卡图像采集困难,难以采集到高质量的数字图像,导致数据样本过少,网络学习到的唐卡图像特征有限。因此未来需要对唐卡数据集进一步扩展,提高数据集质量,并进一步研究大面积破损区域的修复方法。

参考文献 (References)

- Chen Z, Zhang Y L, Gu J J, Zhang Y B, Kong L H and Yuan X. 2022. Cross aggregation transformer for image restoration//Proceedings of the 36th International Conference on Neural Information Processing Systems. New Orleans, USA: Curran Associates Inc.: #1847 [DOI: 10.5555/3600270.3602117]
- Deng Y, Hui S Q, Zhou S P, Meng D Y and Wang J J. 2022. T-former: an efficient transformer for image inpainting//Proceedings of the 30th ACM International Conference on Multimedia. Lisboa, Portugal: ACM: 6559-6568 [DOI: 10.1145/3503161.3548446]
- Gao R H and Grauman K. 2017. On-demand learning for deep image restoration//Proceedings of 2017 IEEE International Conference on Computer Vision. Venice, Italy: IEEE: 1095-1104 [DOI: 10.1109/ICCV.2017.124]
- Hong X, Xiong P F, Ji R H and Fan H Q. 2019. Deep fusion network for image completion//Proceedings of the 27th ACM International Conference on Multimedia. Nice, France: ACM: 2033-2042 [DOI: 10.1145/3343031.3351002]
- Hu W J, Wang W L and Liu Z M. 2016. Damaged region segmentation of Thangka based on domain knowledge and multi-feature. Journal of Central South University (Science and Technology), 47(7): 2326-2333 (胡文瑾, 王维兰, 刘仲民. 2016. 结合领域知识和多特征表示的唐卡破损区域分割算法. 中南大学学报(自然科学版), 47(7): 2326-2333) [DOI: 10.11817/j.issn.1672-7207.2016.07.021]
- Isola P, Zhu J Y, Zhou T H and Efros A A. 2017. Image-to-image translation with conditional adversarial networks//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, USA: IEEE: 5967-5976 [DOI: 10.1109/CVPR.2017.632]
- Jiang K, Jia X M, Huang W X, Wang W B, Wang Z and Jiang J J. 2024. Dynamic association learning of self-attention and convolution in image restoration. Journal of Image and Graphics, 29(4): 890-907 (江奎, 贾雪梅, 黄文心, 王文兵, 王正, 江俊君. 2024. 图像复原中自注意力和卷积的动态关联学习. 中国图象图形学报, 29(4): 890-907) [DOI: 10.11834/jig.230323]
- Jiang Y M, Huang Z Q, Pan X G, Loy C C and Liu Z W. 2021. Talk-to-edit: fine-grained facial editing via dialog//Proceedings of 2021 IEEE/CVF International Conference on Computer Vision. Montreal, Canada: IEEE: 13779-13788 [DOI: 10.1109/ICCV48922.2021.01354]
- Jo Y and Park J. 2019. SC-FEGAN: face editing generative adversarial network with user's sketch and color//Proceedings of 2019 IEEE/CVF International Conference on Computer Vision. Seoul, Korea

- (South): IEEE: 1745-1753 [DOI: 10.1109/ICCV.2019.00183]
- Lan Z, Yan C P, Li H and Zheng Y D. 2023. HDA-GAN: hybrid dual attention generative adversarial network for image inpainting. *Journal of Image and Graphics*, 28(11): 3440-3452 (兰治, 严彩萍, 李红, 郑雅丹. 2023. 混合双注意力机制生成对抗网络的图像修复模型. *中国图象图形学报*, 28(11): 3440-3452) [DOI: 10.11834/jig.220919]
- Liao N H. 2022. A Deep-Learning Research and Application on Thangka Inpainting Based on Params and Computational Methods. Xining: Qinghai Normal University (廖年鸿. 2022. 基于参数与运算轻量化的深度学习唐卡图像修复研究与应用. 西宁: 青海师范大学) [DOI: 10.27778/d.cnki.gqhzy.2022.000214]
- Liao N H, Zhang X J, Peng C Y and Fan H. 2022. LRGAN: a computational lightweight image inpainting neural network. *Scientia Sinica Technologica*, 52(3): 447-457 (廖年鸿, 张效娟, 彭春燕, 范虹. 2022. LRGAN: 一种运算轻量化图像修复网络. *中国科学: 技术科学*), 52(3): 447-457 [DOI: 10.1360/SST-2021-0462]
- Liu D Q. 2021. Research and Application of Thangka Image Inpainting Based on Deep Learning. Xining: Qinghai Normal University (刘丹青. 2021. 基于深度学习的唐卡图像修复研究与应用. 西宁: 青海师范大学) [DOI: 10.27778/d.cnki.gqhzy.2021.000176]
- Liu G L, Reda F A, Shih K J, Wang T C, Tao A and Catanzaro B. 2018. Image inpainting for irregular holes using partial convolutions//*Proceedings of the 15th European Conference on Computer Vision*. Munich, Germany: Springer: 89-105 [DOI: 10.1007/978-3-030-01252-6_6]
- Liu H M and Bi X H. 2013a. Damaged regions segmentation on Thangka for missing colors. *Computer Engineering and Applications*, 49(4): 219-223 (刘华明, 毕学慧. 2013a. 色彩脱落型唐卡破损区域的分割. *计算机工程与应用*, 49(4): 219-223) [DOI: 10.3778/j.issn.1002-8331.1110-0249]
- Liu H M and Bi X H. 2013b. Segmentation of faded regions in Thangka images. *Journal of Computer-Aided Design and Computer Graphics*, 25(2): 241-247, 257 (刘华明, 毕学慧. 2013b. 唐卡褪色区域分割. *计算机辅助设计与图形学报*, 25(2): 241-247, 257) [DOI: 10.3969/j.issn.1003-9775.2013.02.015]
- Liu J. 2020. Research and Implementation of Thang-ga Image Inpainting System Based on Deep Learning. Yinchuan: Ningxia University (刘洁. 2020. 基于深度学习的唐卡图像修复系统的研究与实现. 银川: 宁夏大学) [DOI: 10.27257/d.cnki.gnxhc.2020.000240]
- Liu Y X, Zhao Q J, Pan F, Gao D G and Pubu D. 2023. Structure prior guided text image inpainting model. *Journal of Image and Graphics*, 28(12): 3699-3712 (刘雨轩, 赵启军, 潘帆, 高定国, 普布旦增. 2023. 结构先验指导的文本图像修复模型. *中国图象图形学报*, 28(12): 3699-3712) [DOI: 10.11834/jig.220960]
- Lu X B and Wang W L. 2010. Improved exemplar-based inpainting algorithm for broken Thangka images. *Journal of Computer Applications*, 30(4): 943-946, 952 (卢小宝, 王维兰. 2010. 基于样本块的破损唐卡图像修复算法的改进. *计算机应用*, 30(4): 943-946, 952)
- Nazeri K, Ng E, Joseph T, Qureshi F Z and Ebrahimi M. 2019. Edge-Connect: generative image inpainting with adversarial edge learning [EB/OL]. [2024-07-03]. <https://arxiv.org/pdf/1901.00212.pdf>
- Pathak D, Krahenbuhl P, Donahue J, Darrell T and Efros A A. 2016. Context encoders: feature learning by inpainting//*Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition*. Las Vegas, USA: IEEE: 2536-2544 [DOI: 10.1109/CVPR.2016.278]
- Ren J L. 2022. Thangka Image Inpainting Research Based on Multi-Scale Attention Mechanism and Edge Constraints. Yinchuan: Ningxia University (任家乐. 2022. 基于多尺度注意力机制和边缘约束的唐卡图像修复研究. 银川: 宁夏大学) [DOI: 10.27257/d.cnki.gnxhc.2022.001353]
- Wang W. 2016. Study on Exemplar-Based Thangka Image Inpainting Technologies. Chengdu: Southwest Jiaotong University (王威. 2016. 基于样本块的唐卡图像修复技术研究. 成都: 西南交通大学)
- Yang C, Lu X, Lin Z, Shechtman E, Wang O and Li H. 2017. High-resolution image inpainting using multi-scale neural patch synthesis//*Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition*. Honolulu, USA: IEEE: 4076-4084 [DOI: 10.1109/CVPR.2017.434]
- Yi Z L, Tang Q, Azizi S, Jang D and Xu Z. 2020. Contextual residual aggregation for ultra high-resolution image inpainting//*Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Seattle, USA: IEEE: 7505-7514 [DOI: 10.1109/CVPR42600.2020.00753]
- Yu J H, Lin Z, Yang J M, Shen X H, Lu X and Huang T. 2019. Free-form image inpainting with gated convolution//*Proceedings of 2019 IEEE/CVF International Conference on Computer Vision*. Seoul, Korea (South): IEEE: 4470-4479 [DOI: 10.1109/ICCV.2019.00457]
- Zeng Y, Lin Z and Patel V M. 2022. SketchEdit: mask-free local image manipulation with partial sketches//*Proceedings of 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. New Orleans, USA: IEEE: 5941-5951 [DOI: 10.1109/CVPR52688.2022.00586]

作者简介

张驰,男,硕士研究生,主要研究方向为非物质文化遗产数字化保护和计算机视觉。E-mail:756629946@qq.com

张效娟,通信作者,女,教授,主要研究方向为非物质文化遗产数字化保护和计算机视觉。E-mail:zhxj@qhnu.edu.cn

赵洋,男,副研究员,主要研究方向为视频图像处理、计算机视觉和人工智能。E-mail:yzhao@hfut.edu.cn

卢嘉钰,女,硕士研究生,主要研究方向为非物质文化遗产数字化保护和计算机视觉。E-mail:1564103034@qq.com

谢钰麒,男,硕士研究生,主要研究方向为非物质文化遗产数字化保护和计算机视觉。E-mail:captainbro@163.com