

DOI:10.19651/j.cnki.emt.2211182

## 改进的双阶段生成对抗数字壁画修复算法\*

张双 杨帆

(河北工业大学电子信息工程学院 天津 300401)

**摘要:** 数字壁画修复是计算机视觉在图像修复领域的一个重要应用。为解决修复过程中存在的模糊、结构紊乱、细节丢失等问题,提出了改进的双阶段生成对抗数字壁画修复模型。首先在第1阶段生成器中设计了特征优化融合策略,将编码器中不同尺度的特征进行优化并在解码器部分按比例融合,减少卷积过程中特征信息的丢失;然后在第2阶段生成器中用空洞残差模块代替空洞卷积,将小膨胀率空洞卷积与残差模块结合,增大感受野的同时减少空洞的累积,有效缓解了修复产生的网格伪影现象。实验结果表明:与其他几种修复算法相比,该方法在敦煌壁画数据集上的视觉效果和客观指标均有明显优势,其中峰值信噪比平均提升了3~5 dB,结构相似度平均提升了2%~6%。

**关键词:** 壁画修复;特征优化融合;空洞残差;生成对抗网络

**中图分类号:** TP391.41 **文献标识码:** A **国家标准学科分类代码:** 510.4050

## Digital mural inpainting model based on improved two-stage generative adversarial network

Zhang Shuang Yang Fan

(School of Electronic and Information Engineering, Hebei University of Technology, Tianjin 300401, China)

**Abstract:** Digital mural inpainting is an important application of computer vision in the field of image inpainting. Digital mural inpainting model based on improved two-stage generative adversarial network was proposed to solve the problems of ambiguity, structure disorder and detail loss in the process of inpainting. Firstly, a feature optimization fusion strategy is designed in the first-stage generator. The features of different scales in the encoder are optimized and fused in the decoder in proportion to reduce the loss of feature information in the convolution process. Then, in the second-stage generator, the dilated residual module is used instead of the dilated convolution process, and the dilated convolution with small expansion rate is combined with the residual module to increase the receptive field and reduce the accumulation of holes, which effectively alleviates the repaired grid artifact phenomenon. The experimental results show that the proposed method has obvious advantages in visual effects and tube indexes on the mural dataset compared with other restoration algorithms, in which the peak signal-to-noise ratio is improved by 3~5 dB on average, and the Structural Similarity is improved by 2%~6% on average.

**Keywords:** mural inpainting; feature optimal fusion; dilated residual network; generative adversarial network

## 0 引言

壁画作为最早的绘画形式,主要存在于石窟、寺观、宫殿、墓室中。而敦煌莫高窟作为世界现存壁画最丰富的石窟之一,具有极高的研究价值。但由于自然风化、地质灾害及人为因素的影响,敦煌石窟内众多壁画都存在空鼓、起甲、酥碱、风化、坍塌等问题<sup>[1]</sup>。为了保存中国古代文明的印记,保护壁画的完整性刻不容缓。传统的人工修复方法存在修复周期长、风险大等问题,因此对敦煌壁画进行数

字化修复成为目前的研究热点。

壁画修复属于图像修复的一个分支应用,而图像修复被广泛应用于对应移除、图像编译、图像修饰等工程领域<sup>[2]</sup>。对壁画进行数字化修复主要通过破损区域周围的有用信息构建出合理的、符合人类视觉效果的内容,并且破损面积的大小与最终修复的质量效果是成负相关的,破损面积越大,修复过程越困难。传统的图像修复方法建立在精密的数学或物理模型之上,具有精密的数学论证,主要分为基于扩散和基于样本块合成两种算法。其中曲率驱动扩散

收稿日期:2022-08-25

\* 基金项目:国家重点研发计划智能机器人专项(2019YFB1312102)、河北省自然科学基金(F2019202364)项目资助

模型是最具代表性的基于扩散的修复方法之一<sup>[3]</sup>,注重修复图像的细节信息,适用于小尺寸破损的修复;基于样本块合成的图像修复算法通过搜寻匹配块的方式来修复破损区域,可以完成大范围破损的修复。应用于壁画修复领域,陈永等<sup>[4]</sup>通过改进 Criminisi 算法优先权的计算方式选择更合适的匹配块,同时提升了敦煌壁画的修复效率。总体上看,传统的图像修复算法适用于图像结构信息基本完整且只存在小范围破损的情况下,而面对复杂的修复任务时缺乏对高层语义信息的理解,无法很好地修复。

近年来,深度学习技术的研究发展使得越来越多的研究者通过卷积神经网络提取深层语义信息,图像修复的视觉效果和客观指标有了飞跃式的提升。目前基于深度学习图像修复的关键是如何理解图像的结构信息和深层语义信息使破损区域与非破损区域保持一致性。2016年,Pathak等<sup>[5]</sup>提出了上下文编码器,首次将生成对抗网络(generative adversarial network, GAN)<sup>[6]</sup>引入图像修复领域,自此开启了基于生成对抗图像修复的先河。为了解决局部模糊的问题,Iizuka等<sup>[7]</sup>引入了全局鉴别器,既保证了局部修复的真实性,又保证了全局语义信息的一致性。Wang等<sup>[8]</sup>利用并行机制,通过不同尺寸的卷积核提取图像的多尺度特征,增强了网络的表达能力,图像修复的质量有了显著提升。Pepsi机制<sup>[9-10]</sup>在并行网络中采用了共享编码器的网络结构,在节约计算资源的同时获得了较为精细的纹理细节。Liu等<sup>[11]</sup>利用注意力机制来挖掘图像深层信息的特征关联,解决了图像修复中色差、扭曲等问题。Xie等<sup>[12]</sup>提出了可学习的双向注意力机制,可以学习不规则的掩膜形状,修复不规则的破损图像。Zeng等<sup>[13]</sup>通过金字塔形式的编码器对图像缺失区域进行分层级补全,在图像层面以及特征层面都保持了连贯性。Li等<sup>[14]</sup>提出了区域识别模块和循环特征推理模块结合的方法将生成的特征

图融合,修复细节更完整。Peng等<sup>[15]</sup>将语义和纹理信息分别训练学习,通过特征损失来约束纹理和语义信息的一致性。而Liao等<sup>[16]</sup>则将分割和修复相结合,利用语义和纹理信息的一致性来补全空洞信息。

上述研究的优势基本都体现在公共数据集上,但将这些方法直接应用于壁画修复会忽略壁画的独特性,往往会出现细节丢失、填充区域与壁画风格不一致等问题,针对此问题,本文提出了改进的双阶段生成对抗壁画修复模型,通过不同尺度特征优化融合以及空洞残差来提升壁画修复的纹理细节、减少网格伪影现象。

## 1 实验模型

在壁画修复任务中,充分利用完好区域的语义和纹理信息是修复壁画缺失区域信息的关键。而目前图像修复算法对壁画缺乏语义理解,同时存在纹理细节修复不足、数据缺乏等问题。因此为了降低模型复杂度的同时充分利用壁画的潜在特征信息,我们提出了一种新的壁画修复模型,设计思路如下:

1)设计了连续下采样的双阶段生成对抗壁画修复模型,降低了模型的复杂度。

2)设计了多尺度特征优化融合(multi-scale feature extraction, MSFF)策略,将编码器中不同尺度的特征进行优化并在解码器部分按比例融合,在保证网络稀疏卷积的同时利用了密集矩阵的高计算性能,提取更丰富的特征信息。

3)用空洞残差模块(dilated residual module, DRM)作为编码-解码之间的连接层,扩大感受野的同时减少了空洞累积。

网络结构如图1所示,样本需经过两个不同优化策略的生成器,带有遮挡的壁画传入第1阶段生成器只是进行

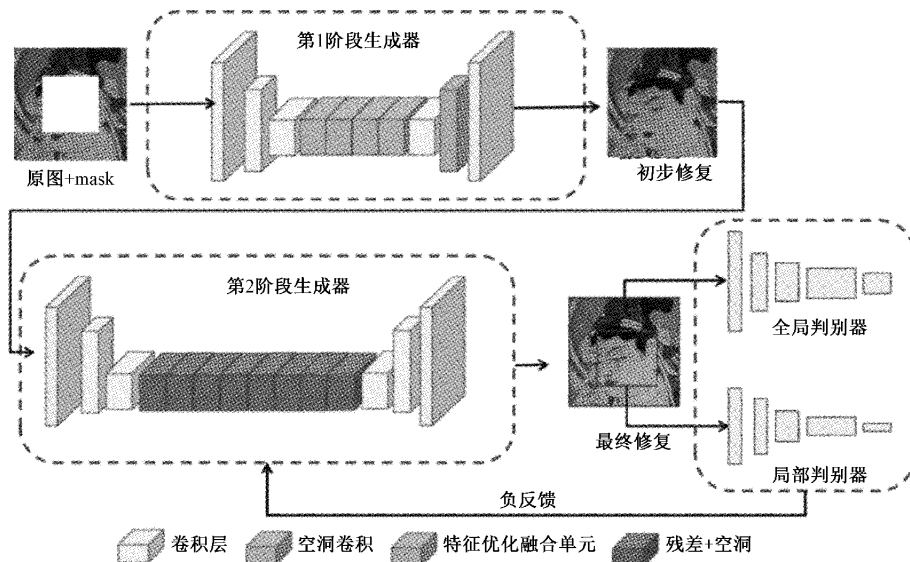


图1 改进的双阶段生成对抗网络结构

简单的轮廓和部分细节的修复,第 2 阶段生成器将一阶段生成器的修复结果进一步优化,产生合理的壁画修复图像。将输出壁画补全部分与标签图像的相应区域共同输入到局部判别器,同时将修复后的完整壁画与标签图像输入到全局判别器,最终将两判别器的判别结果同时反馈到第 2 阶段生成器优化修复的效果。

### 1.1 基于特征优化融合的一阶段生成器

一阶段生成器仅对壁画遮挡区域进行粗颗粒度的修复,因此为了简化网络结构,生成器均采用三层卷积完成编码部分,但连续的下采样会造成特征信息的丢失。因此为了在不增加网络深度的同时增加网络特征提取的能力,提出了特征优化融合策略,实现对编码器中不同尺度的特征进行优化并在解码器阶段与高层语义信息融合。特征优化融合网络如图 2 所示,由优化单元和融合单元两部分组成。

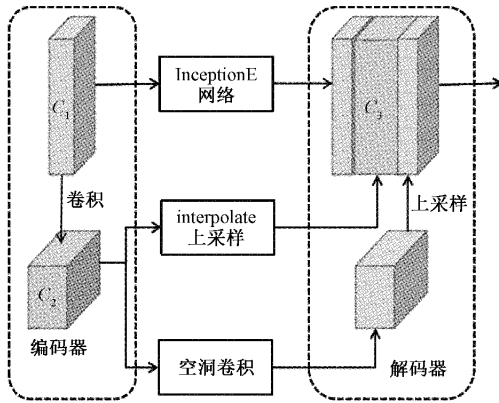


图 2 特征优化融合结构图

优化单元对不同尺度的特征分别进行优化,融合单元将每层优化后的结果在解码器部分进行整合。假设浅层特征  $F_1$  的尺寸大小  $\frac{H}{2} \times \frac{W}{2} \times C_1$ , 在优化单元中经过 Inception 网络输出特征  $F'_1$  尺寸为  $\frac{H}{2} \times \frac{W}{2} \times 4C_1$ 。由于深层特征  $F_2$  的分辨率仅为  $F_1$  特征的一半,所以优化单元直接对  $F_2$  进行上采样,使其与  $F_1$  的分辨率保持一致,输出特征  $F'_2$  的尺寸为  $\frac{H}{2} \times \frac{W}{2} \times C_2$ 。融合单元将浅层级特征  $F_1$  优化后的特征  $F'_1$ 、深层特征  $F_2$  上采样后的特征  $F'_2$  按通道数线性叠加,叠加比例 2:1。通过  $1 \times 1$  卷积实现不同优化单元的占比。

$F_1$  特征优化单元的网络结构如表 1 所示,其中包含 4 个分支,分别以不同的感受野进行特征提取,获取多个不同尺度的信息。每个分支均采用  $1 \times 1$  卷积减少通道数实现信息的聚合,branch1、branch2、branch3 的感受野分别为 1、3、5,而卷积核大小仅有 1 和 3,其中 branch3 用两个  $3 \times 3$  的小卷积核叠加来代替一个  $5 \times 5$  的卷积核,减少显存占用。同时为了减少网络的计算量,branch2 和 branch3 将

$3 \times 3$  卷积拆分成非对称卷积  $1 \times 3$  和  $3 \times 1$  叠加的形式。相比于主干网络中的稀疏卷积  $F_1$  特征优化单元的密集卷积可以提升网络的计算性能,实现不同尺度信息的整合。

表 1  $F_1$  特征优网络

分支	Layer	Kernal	channels	padding
branch1	Conv1	1	64	0
	Conv2_1	1	32	0
branch2	Conv2_2a	$1 \times 3$	32	(0,1)
	Conv2_2b	$3 \times 1$	32	(1,0)
Concat(conv2_2a(conv2_1), Conv2_2b(Conv2_1))				
branch3	Conv3_1	1	32	0
	Conv3_2	3	32	1
	Conv3_3a	$1 \times 3$	32	(0,1)
	Conv3_3b	$3 \times 1$	32	(1,0)
Concat(conv3_3a(conv3_2), Conv3_3b(Conv3_2))				
branch4	avg_pool	3	—	1
	Conv4	1	64	0
Concat(branch1, branch2, branch3, branch4)				

### 1.2 基于空洞残差优化的二阶段生成器

二阶段生成器是对一阶段生成器结果的精细化修复,包含编码器、连接层、解码器三部分。其中编码器将输入图片映射于隐藏空间提取特征矩阵,而连接层是特征矩阵的最终表达,解码器对特征矩阵升维恢复原始图像。因此连接层对于图像信息的恢复至关重要。通常的做法是利用 4 层不同膨胀率(分别为 2,4,8,16)的空洞卷积作为连接层提升感受野,通过更大范围的孔外信息构建壁画缺失区域。但连续的空洞卷积带来“0”填充的累积造成网格伪影现象。以单个空洞卷积为例,卷积核大小为  $3 \times 3$  且膨胀率  $r=2$  的感受野大小为  $5 \times 5$ ,然而实际参与计算的有效像素仅有 9 个,其余的均以“0”填充,导致空洞卷积的特征映射是稀疏分布的。假设采用连续的空洞卷积层,若级联的空洞卷积的膨胀率具有公因子(例如 2,4,8,16 的公因子为 2),则网格伪影会一直传播下去。本文设计了空洞残差模块作为连接层解决这个问题。如图 3 所示,每个残差块<sup>[17]</sup>均采用了小膨胀率( $r=2$ )的空洞卷积,为了达到大膨胀率空洞卷积感受野同样的效果,使用了 8 个残差块级联的形式保证在不降低空间分辨率的基础上增加了感受野,与此同时为了防止网格伪影的传播,每层空洞卷积之后均加入了相同尺寸的传统卷积将空洞卷积隔离。

### 1.3 损失函数

实验训练过程中涉及到的损失函数主要包括重构损失和对抗损失两种。其中重构损失仅通过生成器的生成作用约束,负责衡量真实图像与缺失图像之间的距离。与以往不同的是,本文并非将局部鉴别器和全局鉴别器提取的特征融合,而是分开对生成图像进行优化。因此对抗损

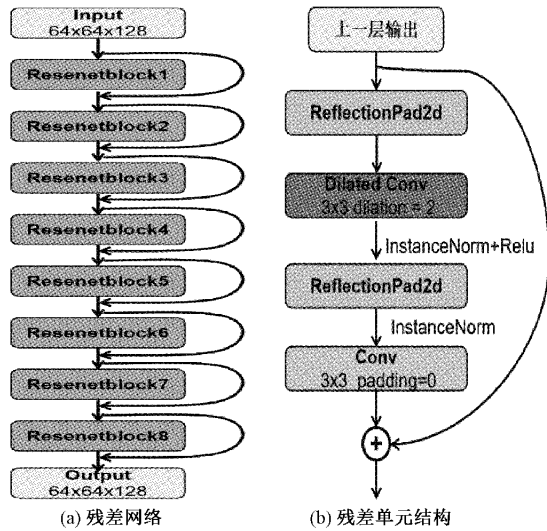


图 3 连接层结构

失分为局部对抗损失和全局对抗损失,局部对抗损失负责约束生成缺失区域与真实图像对应区域特征图的一致性;全局对抗损失负责约束生成器生成图像与真实图像特征图的整体一致性。

一阶段生成器的损失函数仅包括重构损失,利用编码-解码结构生成粗粒度的修复结果,为了保证网络的稳定性,选用 ELU 作为激活函数。若用  $a$  代表输入的原图,  $M_a$  代表二进制掩码,二者经过哈达玛运算得到输入的带掩膜的破损壁画图像,破损壁画由式(1)表示,其中  $\tilde{a}_1$  代表输入的待修复图像,经过一阶段生成器后的输出由式(2)表示,由式(2)推理得一阶段生成器的重构损失函数表达式(3),其中  $\odot$  表示像素间的乘积,而  $\cdot$  代表欧几里得范式。

$$\tilde{a}_1 = a \odot (1 - M_a) \quad (1)$$

$$I_{\alpha_1} = G_1(\tilde{a}_1) \quad (2)$$

$$L_1(a, M_a) = \| M_a \odot (a - I_{\alpha_1}) \| \quad (3)$$

二阶段生成器是在一阶段生成器修复结果的基础上进行精细化修复,其损失函数除了重构损失之外,还包括对抗损失。二阶段生成器同样选用 ELU 作为激活函数,则二阶段生成器的输入、输出以及重构损失函数分别表示为式(4)~(6):

$$\tilde{a}_2 = I_{\alpha_1} \odot M_a + a(1 - M_a) \quad (4)$$

$$I_{\alpha_2} = G_1(\tilde{a}_2) \quad (5)$$

$$L_2(a, M_a) = \| M_a \odot (a - I_{\alpha_2}) \| \quad (6)$$

壁画修复仅使用重构损失会造成缺失区域的模糊填充,因此引入了对抗损失。由于在传统的 GAN 网络中目标函数采用的是 JS 散度衡量两个分布的相似度。当判别器训练精度较高时,JS 散度值为常数  $2\log 2$ ,若采用梯度下降法训练样本时生成器的梯度为 0,此时生成器无法迭代更新,GAN 将出现训练不稳定的现象。因此本文采用

WGAN<sup>[18]</sup>作为训练模型,用 Wasserstein 距离取代 JS 散度,通过引入惩罚项来满足 Lipschitz 连续性条件。如式(7),其中  $P_\gamma, P_g$  表示两种分布,  $\gamma$  表示的是两种分布可能的联合分布的集合,  $x, y$  表示的是在联合分布中采样得到的真实样本和生成器样本。

$$W(P_\gamma \| P_g) = \inf_{\gamma \in \Pi(P_\gamma, P_g)} E_{(x,y) \sim \gamma} [\|x - y\|] \quad (7)$$

$$L_{adv} = E_{a \sim P_r} [D(a)] - E_{I_{\alpha_2} \sim P_g} [D(I_{\alpha_2})] + \lambda E_{\hat{a} \sim P_a} [( \|\nabla_{\hat{a}} D(\hat{a}) \odot (1 - M_a)\|_2 - 1 )^2] \quad (8)$$

对抗损失函数如式(8)所示,其中  $I_{\alpha_2}$  表示二阶段生成器的输出;  $M_a$  代表二进制掩码,0 代表的是缺失区域,1 代表完好区域;  $P_r, P_g$  分别表示真实数据与生成数据,  $\hat{a}$  代表两分布的随机采样值,  $\hat{a}$  服从分布  $P_a$ 。

## 2 实验结果与分析

### 2.1 数据集与数据增强

实验使用收集整合的敦煌壁画作为数据集对模型进行评估。原敦煌数据集包含 5 400 幅壁画,通过剪裁设置每张壁画的大小均为  $256 \times 256$ ,实验使用 RGB 三通道的图像,并选用 4 800 张壁画作为训练集,600 张壁画作为测试集。由于敦煌壁画难以获取,所以训练数据有限。为了防止过拟合,对采集的壁画数据进行剪裁、缩放、水平翻转、亮度变换等数据增强的方法对数据集进行扩充,扩充后训练集约为 34 000 幅壁画。

### 2.2 实验策略与评价指标

考虑到壁画修复本质是一个病态问题,所以采用人为添加中心掩膜与人为添加随机掩膜两种训练策略。训练在 Ubuntu20.04 的系统中运行,采用的是 Python3.8 环境搭载 PyTorch 为开发框架。硬件环境为 Intel Xeon Silver 4114 CPU @ 2.2 GHz, RAM 64 GB,训练中使用的是 NVIDIA GeForce GTX 1080 Ti GPU,显存为 32 GB。批尺寸设置为 16,学习率设为  $10^{-4}$ ,并采用 Adam 作为优化算法对参数进行优化。采用两种评价指标对网络修复性能进行评估,其中包括峰值信噪比 (peak signal-to-noise ratio, PSNR) 和结构相似性 (structural similarity, SSIM),分别由式(9)、(10)表示。

$$PSNR = 20 \lg \left( \frac{2^n - 1}{MSE} \right) \quad (9)$$

$$SSIM = \frac{(2\mu_a \mu_{I_{\alpha_2}} + c_1) (2\sigma_{a I_{\alpha_2}} + c_2)}{(\mu_a^2 + \mu_{I_{\alpha_2}}^2 + c_1) (\sigma_a^2 + \sigma_{I_{\alpha_2}}^2 + c_2)} \quad (10)$$

式(9)中, MSE 表示的是当前图像与真实图像之间的均方误差,  $n$  为每像素的比特数,一般取 8,即像素的灰度阶数为 256。PSNR 的单位是 dB,衡量的是图像之间对应像素点之间的误差。因此 PSNR 值越高代表图像的失真



越少。式(10)中 $\mu_a, \mu_{I_{\alpha_2}}$ 分别表示真实图像 $a$ 和修复图像 $I_{\alpha_2}$ 的均值,用作亮度估计; $\sigma_a^2, \sigma_{I_{\alpha_2}}^2$ 分别表示图像 $a, I_{\alpha_2}$ 的方差,用作对比度估计; $\sigma_{aI_{\alpha_2}}$ 代表图像 $a$ 和 $I_{\alpha_2}$ 的协方差,作为结构相似程度的估计。

### 2.3 实验结果与分析

#### 1) 人为添加中心缺失实验结果

将本文方法使用敦煌壁画数据集训练得到的模型应用到其测试集上,中心缺失实验中掩膜尺寸大小为 $128 \times 128$ 。同时与两种经典算法 CA<sup>[19]</sup>、EDFE<sup>[20]</sup>进行对比分析。

如图 4 所示,图 4(a)表示原始的壁画图像,图 4(b)列表示中心缺失的破损图,图 4(c)~(e)分别表示经 CA、EDFE、本文所提算法输出的壁画修复结果。其中 CA 算法仅采用粗化-细化网络相结合的方式,因此修复的壁画细节(例如眉毛、鼻子、眼睛)连续性差,且面部线条不流畅,甚至会出现错误填充结构紊乱的现象。而 EDFE 算法采用的是编码-解码结构并在特征空间进行特征均衡,修复后的壁画较平滑,但细节修复不完善,对于复杂结构同样存在修复不彻底的问题。本文所提算法在第一阶段生成器中采用了多尺度特征优化融合策略,可以更好的修复出缺失区域的轮廓,在第二阶段生成器中改变了连接层的结构,增强了编码器与解码器的关联性,减少了棋盘伪影现象。在视觉效果方面,可以看出本文所提算法不仅可以修复鼻子、眼睛等特征信息,像头饰这样复杂的纹理细节也能恢复出来,并且基本不存在伪影和局部色差。

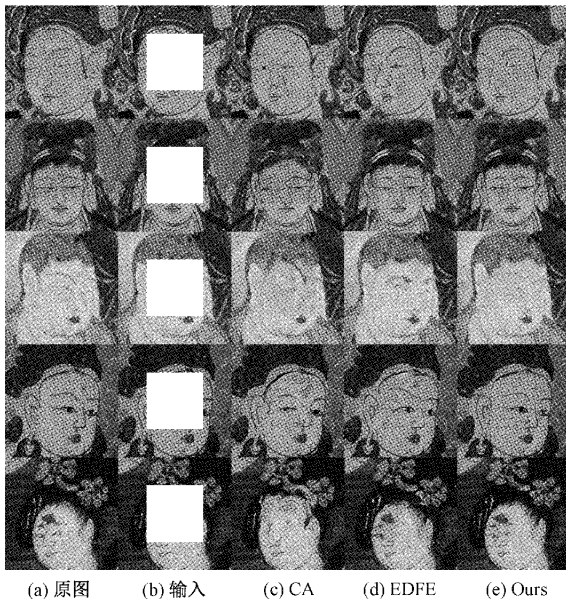


图 4 不同算法的修复效果

客观指标如表 2、3 所示,可以看出相比于其他两种算法,本文所提算法的 PSNR 平均提升了 3~5 dB, SSIM 平均提升了 2%~6%。说明多尺度特征优化融合策略以及

空洞残差模块不仅提升了视觉效果,同时也大幅度提升了相应的客观指标。表 4 统计了单幅壁画的平均修复时间,可以看出本文算法的平均修复时间相比于 CA 算法仅增加了 0.26 s,从而说明本文在保证修复效率的基础上最大程度提升了性能。

表 2 不同算法的峰值信噪比

图像	CA	EDFE	本文
Image1	31.527 3	32.193 0	<b>36.258 8</b>
Image2	28.096 4	31.687 7	<b>35.226 7</b>
Image3	31.785 8	32.923 7	<b>35.390 6</b>
Image4	31.040 1	31.406 8	<b>35.520 5</b>
Image5	32.002 5	32.621 2	<b>34.913 4</b>

表 3 不同算法的结构相似性

图像	CA	EDFE	本文
Image1	0.897 5	0.909 0	<b>0.950 6</b>
Image2	0.787 3	0.865 4	<b>0.920 7</b>
Image3	0.872 5	0.901 8	<b>0.946 6</b>
Image4	0.870 6	0.922 5	<b>0.960 4</b>
Image5	0.788 5	0.862 7	<b>0.919 3</b>

表 4 单幅壁画在不同算法下平均修复时间

算法	时间/s
CA	<b>2.93</b>
EDFE	3.46
<b>Ours</b>	3.19

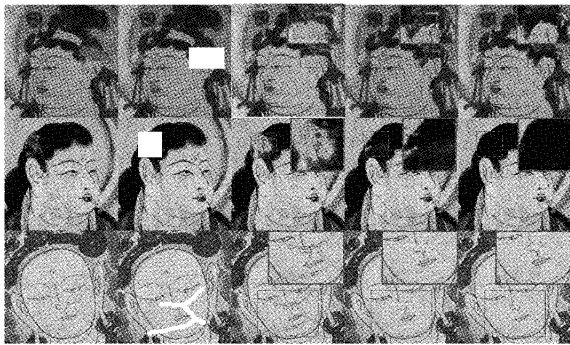
#### 2) 真实破损壁画修复结果

为了证实本方法具有良好的泛化性,用训练好的模型(采用随机添加矩形掩膜策略)修复壁画真实存在的破损区域,存在破损的壁画数据未参与训练,3 种方法得到的测试结果如图 5 所示,其中图 5(a)表示带有破损的原始壁画,图 5(b)表示在壁画破损处添加遮挡后的图像,图 5(c)、(d)、(e)分别表示经 CA、EDFE、本文所提算法输出的壁画修复结果。真实壁画的破损区域位置、大小、形状均存在不确定性且破损的类型多种多样,以下 3 幅壁画分别存在墙皮脱落、掉色、裂痕的问题。CA、EDFE 两种算法的结果均不能完整清晰的恢复出破损区域的内容,出现了填充断层、色差等问题。而本方法对于破损区域的填充内容更符合人类视觉效果。

#### 3) 消融实验

为了评估多尺度特征融合策略和空洞残差作为连接层对模型性能的影响,在设计基线模型基础上分别改变特征融合的方式和连接层的卷积方式,对前后的视觉效果和客观指标进行可行性分析。

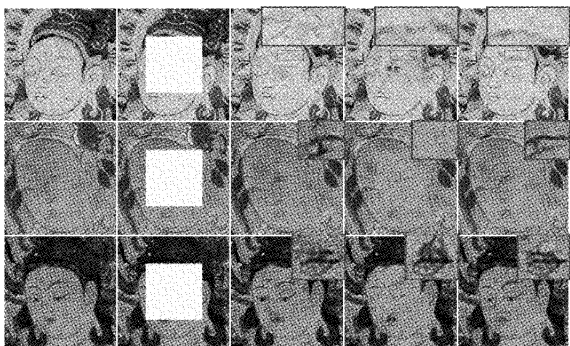
视觉效果如图 6 所示,其中图 6(a)表示原始的壁画图



(a)原图 (b)输入 (c)CA (d)EDFE (e)Ours

图 5 不同算法对真实破损壁画的修复效果

像,图 6(b)表示中心缺失的破损图,图 6(c)表示在基线模型的基础上仅采用多尺度特征融合策略的修复结果,图 6(d)表示在基线模型的基础上仅增加空洞残差模块的修复结果,图 6(e)列表示本文算法的最终修复结果。由此可以看出在第一阶段生成器使用特征优化融合策略后的纹理细节修复基本完整,这是由于特征优化融合可以获取更加丰富的特征信息,因此恢复出的壁画细节更完善;在第 2 阶段生成器中,将连接层中的 4 层空洞卷积替换为空洞残差模块,用残差和小膨胀率空洞卷积结合的方式,在扩大感受野的同时减少了“0”填充的累积,因此测试结果相比于前者棋盘伪影明显减少。



(a)原图 (b)输入 (c)基线+MSFF (d)基线+DRN (e)Ours

图 6 消融实验视觉效果

客观指标如表 5 和 6 所示,多尺度特征优化融合策略的 PSNR 和 SSIM 均高于空洞残差模块,其中 PSNR 值差异不大,说明两种方式均可有效提升像素间的平均差异,但对于 SSIM 值前者明显大于后者,说明多尺度特征融合更注重结构信息,但因此忽略了修复区域的清晰度,而空

表 5 消融实验的峰值信噪比

图像	基线+MSFF	基线+DRN	本文
Image1	35.180 8	34.828 1	<b>35.777 0</b>
Image2	34.371 7	34.184 3	<b>35.779 7</b>
Image3	34.992 1	34.418 9	<b>35.330 9</b>

表 6 消融实验的结构相似性

图像	基线+MSFF	基线+DRN	本文
Image1	0.911 8	0.874 9	<b>0.953 3</b>
Image2	0.955 6	0.949 0	<b>0.968 4</b>
Image3	0.953 8	0.929 9	<b>0.970 7</b>

洞残差模块弥补了这一问题。由此可以证明:特征融合策略在一定程度上可以提升修复细节和各项客观指标,而空洞残差卷积可以减少棋盘伪影现象,提高清晰度。因此两者结合可以达到更好的修复效果。

### 3 结 论

本文针对壁画的特殊性提出了一种全新的双阶段生成对抗修复网络,使用不同尺度特征优化融合模块将编码器特征采用不同策略优化后在解码器线性叠加直至恢复原始壁画的第一阶段模型结束,可以有效地提取更丰富的特征信息,提升纹理细节的修复效果,解决壁画修复中细节丢失、结构紊乱等问题。同时还提出了空洞残差模块作为编码-解码器的连接层,通过小膨胀率空洞卷积与残差网络结合,在增加感受野的同时减少空洞的累积,因此可以有效的减少网格伪影现象。实验表明,通过采取不同尺度特征优化融合和空洞卷积两种方式,在壁画数据集上的视觉效果和客观指标均优于几种主流的算法。

### 参考文献

- [1] 杨晓东. 探索古代石窟壁画保存现状以及修复与保护[J]. 工业建筑, 2021, 51(11): 230-231.
- [2] 李月龙, 高云, 闫家良, 等. 基于深度神经网络的图像缺损修复方法综述[J]. 计算机学报, 2021, 44(11): 2295-2316.
- [3] 刘泽鑫, 万旺根. 基于曲率驱动扩散与改进型稀疏表示的图像修复算法[J]. 电子测量技术, 2018, 41(11): 103-107.
- [4] 陈永, 陈锦, 艾亚鹏, 等. 基于边缘缺失重构和改进优先权的壁画图像修复[J]. 激光与光电子学进展, 2021, 58(24): 231-240.
- [5] PATHAK D, KRAHENBUHL P, DONAHUE J, et al. Context encoders: Feature learning by inpainting[C]. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 2536-2544, DOI: 10.1109/cvpr.2016.278.
- [6] GOODFELLOW I, POUGET-ABADIE J, MIRZA M, et al. Generative adversarial networks[J]. Communications of the ACM, 2020, 63(11): 139-144.
- [7] IIZUKA S, SIMO-SERRA E, ISHIKAWA H. Globally and locally consistent image completion[J]. ACM Transactions on Graphics(ToG), 2017, 36(4): 1-14.

- [8] WANG Y, TAO X, QI X, et al. Image inpainting via generative multi-column convolutional neural networks[C]. In Proceedings of the 32nd International Conference on Neural Information Processing Systems, 2018: 329-338.
- [9] SAGONG M, SHIN Y, KIM S, et al. Pepsi: Fast image inpainting with parallel decoding network[C]. 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), IEEE Computer Society, 2019: 11352-11360, DOI: 10.1109/cvpr.2019.01162.
- [10] SHIN Y G, SAGONG M, YEO Y, et al. Pepsi++: Fast and lightweight network for image inpainting[J]. IEEE transactions on neural Networks and Learning Systems, 2020, 32(1): 252-265.
- [11] LIU H, JIANG B, XIAO Y, et al. Coherent semantic attention for image inpainting[C]. 2019 IEEE/CVF International Conference on Computer Vision (ICCV), IEEE Computer Society, 2019: 4169-4178, DOI: 10.1109/iccv.2019.00427.
- [12] XIE C, LIU S, LI C, et al. Image inpainting with learnable bidirectional attention maps[C]. 2019 IEEE/CVF International Conference on Computer Vision (ICCV), IEEE Computer Society, 2019: 8857-8866, DOI: 10.1109/iccv.2019.00895.
- [13] ZENG Y, LIN Z, YANG J, et al. High-resolution image inpainting with iterative confidence feedback and guided upsampling[C]. In European Conference on Computer Vision, Springer, Cham, 2020: 1-17.
- [14] LI J, WANG N, ZHANG L, et al. Recurrent feature reasoning for image inpainting[C]. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), IEEE Computer Society, 2020: 7757-7765, DOI: 10.1109/cvpr42600.2020.00778.
- [15] PENG J, LIU D, XU S, et al. Generating diverse structure for image inpainting with hierarchical VQ-VAE[C]. 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, 2021: 10770-10779.
- [16] LIAO L, XIAO J, WANG Z, et al. Image inpainting guided by coherence priors of semantics and textures[C]. 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), IEEE Computer Society, 2021: 6535-6544.
- [17] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 770-778, DOI: 10.1109/cvpr.2016.90.
- [18] ARJOVSKY M, CHINTALA S, BOTTOU L. Wasserstein generative adversarial networks[C]. In International Conference on Machine Learning, PMLR, 2017: 214-223.
- [19] YU J, LIN Z, YANG J, et al. Generative image inpainting with contextual attention[C]. 31st Meeting of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, CVPR 2018, IEEE Computer Society, 2018: 5505-5514, DOI: 10.1109/cvpr.2018.00577.
- [20] LIU H, JIANG B, SONG Y, et al. Rethinking image inpainting via a mutual encoder-decoder with feature equalizations[C]. European Conference on Computer Vision, Springer, Cham, 2020: 725-741.

#### 作者简介

张双, 硕士研究生, 主要研究方向为图像修复与图像处理。

E-mail: zszsnk@163.com

杨帆, 博士, 教授, 博士生导师, 主要研究方向为电子电路与计算机视觉。

E-mail: yangfan@hebut.edu.cn