

多尺度级联网络的墓室壁画数字生成技术

吴萌^{1*}, 任义¹, 王佳²

¹西安建筑科技大学信息与控制工程学院, 陕西 西安 710055;

²陕西历史博物馆, 陕西 西安 710061

摘要 因墓道狭窄,分块揭取的大型墓室壁画在高清采集后会存在周边信息缺失的情况。为了能够重建壁画分块间丢失的数据,提出了一种多尺度级联网络的墓室壁画数字生成技术。首先,对墓室壁画进行大尺度生成;然后,将重建结果输入深度语义小尺度生成网络中进行精细的数字信息生成。在小尺度生成网络引入自注意力机制,增强生成区域与全局信息的关联性,解决生成区域边界的伪影问题。在反馈损失中改进纹理损失,提高重建信息纹理精细度以及壁画生成效果。为加快训练进程、促进梯度反向传播效率,在生成网络中加入跳跃连接。通过消融实验和对比组实验验证,该数字生成技术可提高壁画分块外延信息的纹理匹配率并弱化伪影的影响,在客观指标峰值信噪比和结构相似性上均取得较好的结果。

关键词 墓室壁画; 级联生成网络; 纹理损失函数; 自注意力机制; 数字生成

中图分类号 TP391 文献标志码 A

DOI: 10.3788/LOP213032

Digital Generation Technology for Tomb Murals Based on Multiscale Cascade Network

Wu Meng^{1*}, Ren Yi¹, Wang Jia²

¹School of Information and Control Engineering, Xi'an University of Architecture and Technology, Xi'an 710055, Shaanxi, China;

²Shaanxi History Museum, Xi'an 710061, Shaanxi, China

Abstract Large-scale tomb murals are divided into several blocks by narrow passages. Hence, during high-definition collection, certain information around these blocks may be missing. To address this, a digital generation technology for tomb murals based on a multiscale cascade network is proposed, for reconstructing these lost data between mural blocks. In this approach, tomb murals were first generated on a large scale, and the reconstruction results were subsequently input into a deep semantic small-scale generation network to generate fine digital information. A self-attention mechanism was introduced into the small-scale generation network to enhance the correlation between the generation region and the global information and solve the artifact problem of the generation region boundaries. In terms of the feedback loss, the texture loss and the texture fineness of the reconstructed information are improved, and the mural generation effect is also improved. Jump connections were added to the generation network to accelerate the training process and enhance the efficiency of gradient backpropagation. Based on ablation and comparative group experiments, the proposed digital generation technology can improve the texture matching rate of mural block epitaxial information and reduce the influence of artifacts. This proposed method achieves good objective indexes for the peak signal-to-noise ratio and structural similarity.

Key words tomb mural; cascade generation network; texture loss function; self attention mechanism; digital generation

1 引言

墓室壁画作为我国古代三大壁画(墓室壁画、石窟壁画、寺观壁画)中唯一地下文物,因其墓道出口窄小,

在 20 世纪中期的考古挖掘中^[1]通常采取对整幅大型壁画按照分块揭取的特殊方式进行搬运与保存。如章怀太子墓“马球图”长约 7 m、高约 3 m,被分成宽约 1.5 m 的 5 块壁画^[2],经过激光测距和高清采集^[3]后形成的数

收稿日期: 2021-11-23; 修回日期: 2021-12-17; 录用日期: 2022-01-11; 网络首发日期: 2022-01-21

基金项目: 国家自然科学基金(61701388)、陕西省自然科学基金基础研究计划(2021JM-377)、西安市科技局项目(GXYD10.1)

通信作者: *wumeng@xauat.edu.cn

据总量累计超过 22 GB, 单个壁画分块有 4.4~4.5 GB 的数字图像信息。分块揭取所产生的块间信息缺失导致数字化后的壁画信息间断, 严重影响了游客对整幅壁画的观展和文史专家对历史信息的解读。

古代壁画的数字修复技术可为两大类: 一类是以像素扩散为主的信息生成方法, 如 Karianakis 等^[4]利用形态学算子对德黑兰壁画进行检测, 通过改进 Total variation (TV) 算法对检测到的裂痕进行修复; Jaidilert 等^[5]采用交互式数字修复技术, 通过区域生长形成病害掩码, 然后利用变分法将像素填充到泰国壁画的缺失区域。当壁画存在的病害类型为细长划痕时, 此类方法是有效的, 但随着壁画缺损面积增加则会因高阶偏微分推导产生连续近似值导致画面模糊。另一类是以样本填充为主的信息生成方法, 以 Criminisi^[6]的样本填充算法为基础模型, Purkait 等^[7]提出基于样本相干纹理合成技术, 依据对称性修复印度神庙壁画; 陈永等^[8]引入序贯相似性算法和布谷鸟寻优算法来提高最佳匹配块的搜索精度, 并引入信息熵^[9]改进优先值计算过程, 均在敦煌壁画的修复中表现出较好的效果。此类方法能够修复壁画中间较大的空洞类缺失, 但全局遍历寻优过程以及参考样本不足引起的垃圾块堆积, 导致其并不适用于大型墓室壁画块间连续缺失信息的重建需求。

基于深度学习的修复算法^[10-11]相较于传统 Inpainting 算法能够学习更深层的特征^[12], 不局限于现存有效信息的可利用率, 且能生成新的信息。如 Laube^[13]提出在卷积神经网络 (CNN) 中增加单一网络拓扑结构的方法, 可对高分辨率纹理中的大块缺失区域进行图像修复; 彭广泽等^[14]利用正常曝光和短曝光条件获取两帧条纹图, 通过 CNN 的去噪正则化来实现

条纹图局部高光区域的快速修复。这些方法的主要思路是以神经网络训练填充样本来进行 Inpainting。而 Sabini 等^[15]提出的 Outpainting 数字生成思路, 则以生成对抗网络为基本框架, 能够不局限于图像尺寸向外生成图像信息。

本文以 Outpainting 的思路对壁画进行分块外延性信息的数字生成技术研究, 以 GAN 为基本网络由内向外合理预测壁画分块周边空间内的数字信息, 相比由外向内的图像填充修复更适合因分块揭取丢失大量块间信息的墓室壁画数字重建的需求。针对墓室壁画外延过程中因模型忽略全局特征提取导致的生成边界存在伪影及精度不高等问题, 提出一种多尺度级联网络的墓室壁画数字生成技术。该技术设计一种外延大尺度生成网络与深度语义小尺度生成网络连接的结构, 在深度语义小尺度生成网络中引入自注意力机制获取壁画生成区域与全局信息的关联性, 并引入优化纹理损失函数, 增强生成网络对壁画高频精度信息的生成, 提高墓室壁画的生成质量。

2 数字生成技术的基本原理

以 Outpainting 为核心的数字生成技术, 其基本网络框架是 Goodfellow^[16]提出的 GAN 模型。通过“博弈论”引导生成过程, 由互相对抗的生成网络 G 和判别网络 D 组成, 其中生成网络能够从真实样本数据中学习潜在的特征表示并生成新的样本数据。判别网络试图区分是真实数据还是生成的假样本数据, 当无法区分真假数据时, 才能确保生成的数字信息与原始信息保持了高度的一致性, 满足数字生成的效果要求。在这种一直循环的“博弈”下, 生成网络和判别网络的能力将达到最优, 即“纳什均衡”状态^[17]。生成对抗网络原理如图 1 所示。

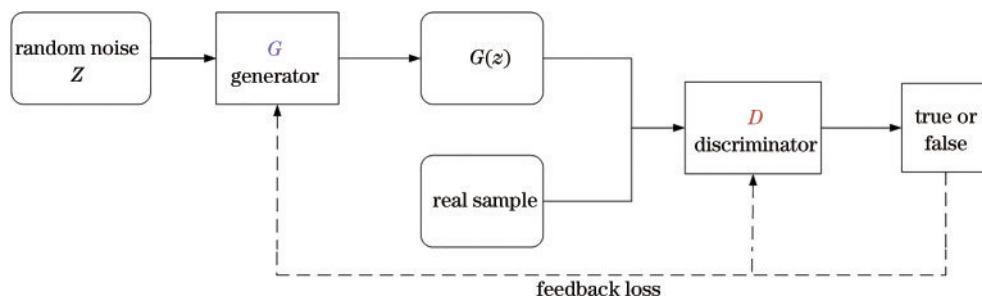


图 1 生成对抗网络原理图

Fig. 1 Schematic diagram of generating countermeasure network

GAN 目标函数的表达式为

$$\min_G \max_D V(D, G) = E_{x \sim P_{\text{data}}(x)} [\ln D(x)] + E_{z \sim P_{z(z)}} \left\{ \ln \left[1 - D[G(z)] \right] \right\}, \quad (1)$$

式中: $V(D, G)$ 为 GAN 优化的目标函数; $E(*)$ 为数学期望; x 为真实数据集 P_{data} 中的真实数据; $P_{z(z)}$ 为生成数据的分布; $D(x)$ 为判别网络判别真实数据的概率大

小; $D[G(z)]$ 为判别网络判别生成数据的概率大小。

生成网络 G 要最小化目标函数来生成更好的数据分布欺骗判别网络, 而判别网络 D 要最大化目标函数, 使 D 能够增强判别真假数据的能力。在最理想的情况下, D 的判别结果应为 1/2, 输入随机数据至判别网络, 不能判别出是否为真实数据集中的真实数据, 也不能判别出是否来自生成网络所生成的假数据。

使用两个级联的生成对抗网络能够从大规模的数据集中学习墓室壁画特征语义信息,并在生成网络和判别网络中分别加入自注意力机制和纹理损失优化,从而对壁画未知区域进行Outpainting。该模型根据已知壁画信息提取特征要素使数字生成的壁画更加逼真,克服传统算法生成大面积区域填充不佳和数字生成信息关联强度较低等问题。

3 多尺度级联网络框架

多尺度级联网络框架如图 2 所示,共由 3 部分构成:外延大尺度生成网络 G_e 、深度语义小尺度生成网络 G_d 、损失优化判别网络 D 。 G_e 对壁画进行外延大尺度数字生成, G_d 对 G_e 生成的大尺度壁画进行更为精细的小尺度数字生成, G_d 中加入自注意力机制,增强生成信息与全局信息的连贯性。损失优化判别网络 D 加入改进的纹理损失对数字生成结果进行优化。 G_e 和 G_d 中使用编码器-译码器结构并加入跳跃连接,有助于

梯度反向传播和加快训练过程。 I_m 为对壁画进行掩码处理后的图像。

3.1 外延大尺度生成网络

外延大尺度生成网络由编码器、解码器两部分组成。编码模块由下采样卷积(Conv)层、批归一化(BN)层和激活函数 Leaky-ReLU 构成。编码阶段使用 6 个下采样卷积层来提取壁画纹理、颜色等特征。将下采样卷积核设置为 4×4 ,步长为 2,每个卷积层后添加 BN 层来提高模型训练速度和训练稳定性。解码模块由上采样反卷积(Deconv)层、BN 层和 ReLU 构成,对编码阶段捕获的潜在特征采用上采样进行壁画分块周边信息生成,上采样卷积核设置为 4×4 ,步长为 2,恢复到起始分辨率大小,其中上采样最后一层使用激活函数 Tanh 输出。下采样、上采样具体结构如图 3 所示。

外延大尺度生成网络的生成目标是尽可能生成逼真的壁画信息,模型如图 4 所示。

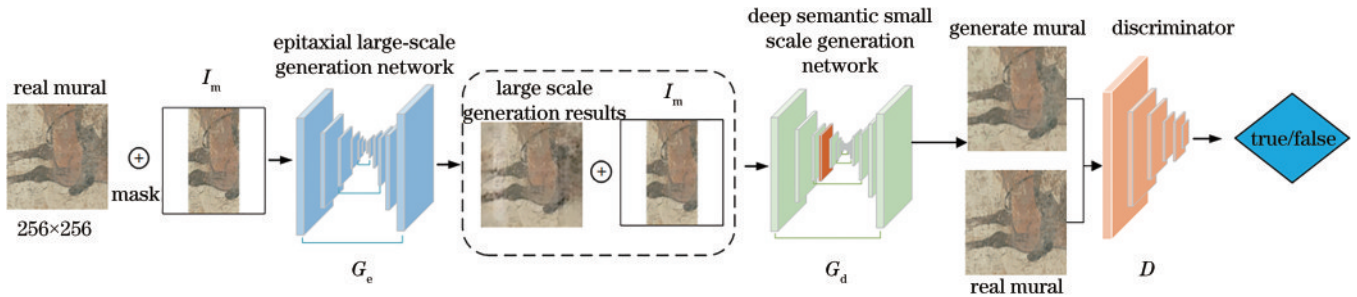


图 2 多尺度级联网络框架

Fig. 2 Multi-scale cascade network framework

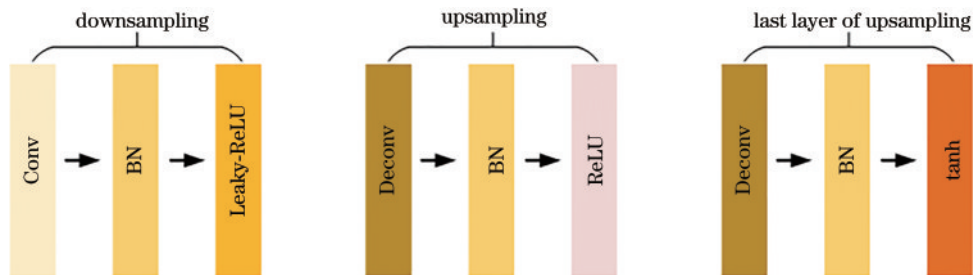


图 3 上采样、下采样具体结构图

Fig. 3 Specific structure diagram of upsampling and downsampling

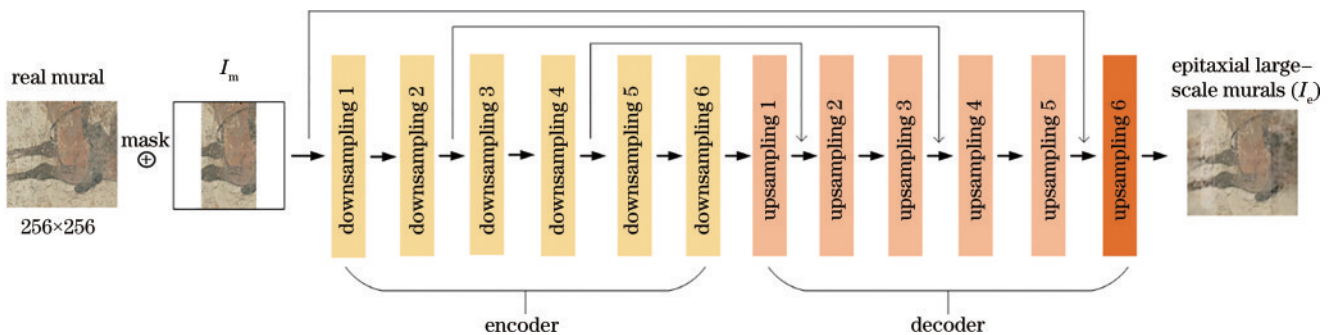


图 4 外延大尺度生成网络

Fig. 4 Epitaxial large-scale generation network

生成过程表达式为

$$I_e = G_e[I_{\text{real}} \odot (1 - M)], \quad (2)$$

式中: I_{real} 为真实壁画; G_e 为外延大尺度生成网络; M 为掩码区域; I_e 为外延大尺度生成网络生成的壁画; \odot 为真实壁画和掩码壁画的像素点乘运算符。但大尺度网络生成的 I_e 高频细节特征缺失量较大, 故将其输出结果作为下一级输入进行深度语义小尺度生成。

3.2 深度语义小尺度生成网络

为增强高频细节特征, 提高墓室壁画生成质量, 保证生成后的像素区域与未缺失区域在空间、视觉上的结构一致性和连续性, 在级联的深度语义小尺度生成网络中设计 8 层对称的编码器-译码器结构。编码阶段使用 8 层卷积核为 4×4 , 步长为 2 的下采样卷积层提取壁画高频细节特征, 加入自注意力机制学习待生成像素点与其他所有像素点的关联性, 使分块壁画能够平

滑过渡到生成区域, 如图 5 所示。解码阶段使用 8 层卷积核为 4×4 , 步长为 2 的上采样 Deconv 层恢复墓室壁画分辨率并生成壁画数字信息。

深度语义小尺度生成网络生成过程表达式为

$$I_{\text{output}} = G_d[I_e \odot (1 - M)], \quad (3)$$

式中: G_d 为深度语义小尺度生成网络; I_{output} 为最终壁画信息生成图。

为了能够提高真实壁画向生成壁画过渡的平滑性与自然性, 引入自注意力机制^[18]。该层能够学习到生成区域的像素点与全局像素点的关联性, 使得生成区域不再依靠相邻位置的像素信息, 而是统筹全局特征信息后计算得出, 捕捉全局信息来增加卷积的感知视野, 克服了真实区域不能向生成区域平滑过渡的问题, 其结构如图 6 所示。

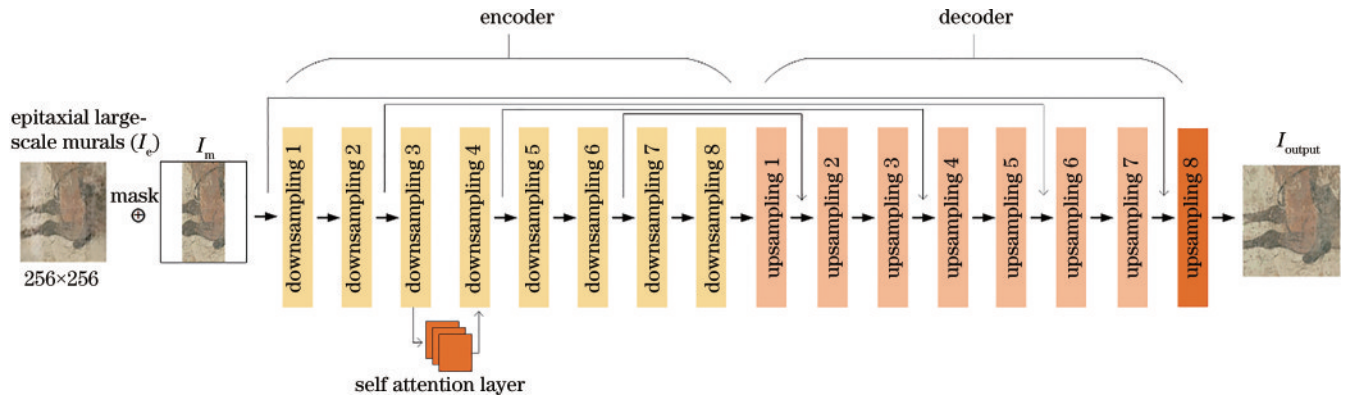


图 5 深度语义小尺度生成网络

Fig. 5 Deep semantic small-scale generation network

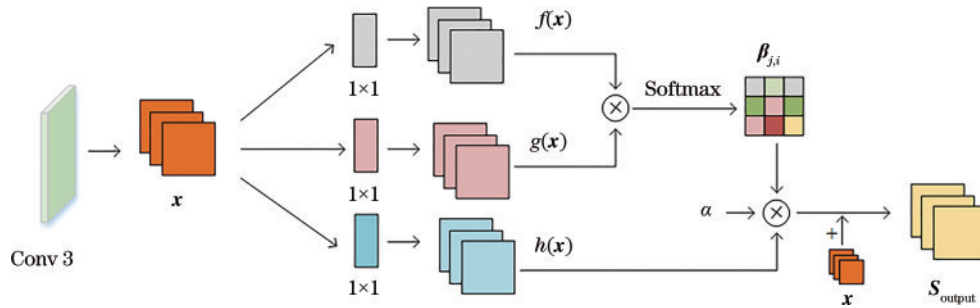


图 6 自注意力机制

Fig. 6 Self attention mechanism

从卷积层 3 中提取的特征图作为自注意力层输入, 通过大小 1×1 的卷积层, 得到特征图 $f(\mathbf{x})$ 、 $g(\mathbf{x})$ 、 $h(\mathbf{x})$; 然后将 $f(\mathbf{x})^T$ 与 $g(\mathbf{x})$ 相乘重构并经过 Softmax 层输出注意力映射特征图, 计算表达式为

$$\beta_{j,i} = \exp(s_{ij}) / \sum_{i=1}^N \exp(s_{ij}), \quad (4)$$

式中: $s_{ij} = f(\mathbf{x})^T g(\mathbf{x})$; $\beta_{j,i}$ 为 $g(\mathbf{x})$ 中的区域位置 j 对 $f(\mathbf{x})$ 中的区域位置 i 的影响作用。

最终输出注意力特征图的计算表达式为

$$S_{\text{output}} = \beta_{j,i} h(\mathbf{x}) \alpha + \mathbf{x}, \quad (5)$$

式中: 注意力映射图 $\beta_{j,i}$ 与 $h(\mathbf{x})$ 、可学习参数 α 相乘并与 \mathbf{x} 相加, 即得到注意力机制的最终结果, 并将其作为下一层卷积的输入。其中, α 在初始阶段设置为 0, 即从相邻局部获取信息, 随着训练生成的进行, 可以给相邻处信息分配更多权重, 逐渐增强生成信息的关联性。

3.3 损失优化判别网络

判别网络模型如图 7 所示, 将生成壁画和真实壁画作为判别网络的输入。使用 IN 层可以将输入特征

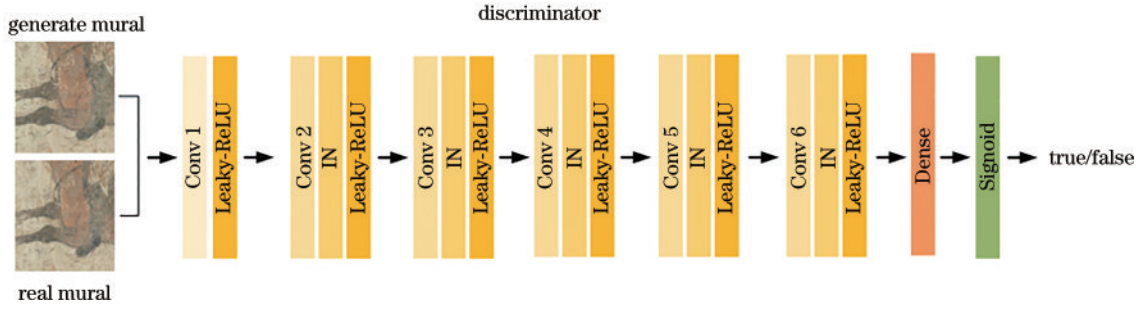


图 7 判别网络模型

Fig. 7 Discriminant network model

进行标准化并提取特征。在迭代过程中与生成网络对抗学习,并通过损失函数调整判别网络参数,使网络能够生成更加真实、自然的数字信息。

该判别网络包括特征提取和线性分类的判别输出,在 Conv 层后添加 BN 层来增强判别网络的稳定性以及避免其梯度消失。为了增加训练速度以及减小过拟合,使用全连接 Dense 层计算特征图的平均像素,并使用 Sigmoid 对输入样本进行判别分类。

4 损失函数优化策略

生成对抗网络的损失函数能够在训练过程中有效提高待生成区域的生成质量,可以将真实壁画特征图与生成壁画特征图进行比较,形成真实壁画 I_{real} 和生成壁画 I_{output} 之间的函数映射关系,能快速学习到目标壁画的高频纹理细节等信息,达到高质量数字生成的目的。为提高生成壁画信息的精度和增强全局内容的一致性,使用如下损失函数来训练网络模型。

1) 优化重构损失

优化重构损失负责优化生成壁画与真实壁画之间的差异,该损失 L_{rec} 使用 L1 范数缩小对应像素间差值并保证待生成区域与原壁画内容上下文信息一致性,其数学表达式为

$$L_{\text{rec}} = \chi_1 \|I_{\text{real}} - I_e\|_1 + \chi_2 \|I_{\text{real}} - I_{\text{output}}\|_1, \quad (6)$$

式中: $\|I_{\text{real}} - I_e\|_1$ 为外延大尺度生成网络生成的壁画 I_e 与真实壁画 I_{real} 之间的损失值; $\|I_{\text{real}} - I_{\text{output}}\|_1$ 为深度语义小尺度生成网络生成的壁画 I_{output} 与真实壁画 I_{real} 之间的损失值; $\chi_1, \chi_2 = 1 - \chi_1$ 为加权超参数。

2) 优化纹理损失

重建损失虽然能够缩小生成壁画与真实壁画之间差异,但对壁画高频纹理细节生成效果较差。将纹理损失函数 $L_{\text{texture-loss}}$ ^[19] 进行优化,在每层卷积输出特征图后都添加纹理损失反馈来增强生成部分的纹理细节匹配度,提取外延大尺度生成网络和深度语义小尺度生成网络 Conv 层输出的特征图来计算 Gram 矩阵。外延大尺度生成网络的纹理损失 $L_{\text{t-e}}$ 和深度语义小尺度生成网络的纹理损失 $L_{\text{t-d}}$ 使用 L2 范数能够提升模型泛化能力、防止过拟合。两个生成网络的纹理损

失数学表达式分别为

$$L_{\text{t-e}} = \|G_{\text{gram}}(L_n(I_{\text{real}})) - G_{\text{gram}}(L_n(G_e(I_{\text{real}} \odot (1-M))))\|_2^2, \quad (7)$$

$$L_{\text{t-d}} = \|G_{\text{gram}}(L_n(I_{\text{real}})) - G_{\text{gram}}(L_n(G_d(I_e \odot (1-M))))\|_2^2, \quad (8)$$

总纹理损失函数表达式为

$$L_{\text{tex}} = \chi_1 L_{\text{t-e}} + \chi_2 L_{\text{t-d}}, \quad (9)$$

式中: $L_{\text{t-e}}$ 为外延大尺度生成网络纹理损失; $L_{\text{t-d}}$ 为深度语义小尺度生成网络纹理损失; G_{gram} 为 Gram 矩阵; L_n 为壁画通过第 n 层 Conv 层的输出,即每层卷积输出特征图后都添加纹理损失反馈; $G_e(I_{\text{real}} \odot (1-M))$ 、 $G_d(I_e \odot (1-M))$ 分别为外延大尺度生成网络和深度语义小尺度生成网络生成的壁画。优化纹理损失提取每个 Conv 层输出纹理特征的差距,消除壁画伪影且能够缩小生成壁画与真实壁画间高维纹理细节差距,增强壁画外延信息真实感。

3) 对抗损失

使用对抗损失函数,使生成壁画输入到判别网络中能够大概率被判定为真实壁画,并提高模型收敛速度。其中,外延大尺度生成网络对抗损失函数表达式为

$$L_e = E[\lg D(I_{\text{real}})] + E[\lg(1 - D(G_e(I_{\text{real}} \odot (1-M))))], \quad (10)$$

深度语义小尺度生成网络对抗损失函数表达式为

$$L_{\text{deep}} = E[\lg D(I_{\text{real}})] + E[\lg(1 - D(G_d(I_e \odot (1-M))))], \quad (11)$$

总对抗损失函数表达式为

$$L_{\text{adv}} = \chi_1 L_e + \chi_2 L_{\text{deep}}, \quad (12)$$

式中: $\lg D(I_{\text{real}})$ 为判别网络将真实壁画判别为真的概率; $\lg(1 - D(G_e(I_{\text{real}} \odot (1-M))))$ 、 $\lg(1 - D(G_d(I_e \odot (1-M))))$ 为判别网络将生成壁画判别为假的概率; $E(*)$ 为函数期望值。在生成对抗训练中,生成网络和判别网络更新反馈损失,逐步使训练网络达到“纳什均衡”状态,此时网络性能达到最优。

5 实验测试与分析

为了验证多尺度级联网络数字生成技术的有效

性,在唐墓室马球图壁画数据集下进行消融实验和对比实验,并与文献[15]的DCGAN网络与文献[20]的WGAN-GP网络进行数字生成结果对比。评价指标使用峰值信噪比^[21](Peak signal-to-noise ratio, PSNR)和结构相似度^[22](Structural similarity, SSIM)来进行客观评价。其中,PSNR用来衡量两幅图片对应像素的误差,值越大表示壁画信息生成效果越好;SSIM可计算生成壁画与真实壁画之间的亮度、对比度及结构间的相似性,其值越接近于1表示两张壁画的相似度越高。实验平台操作系统为64位Windows 10;Python 3.6;PyTorch 1.2。硬件配置:GPU NVIDIA GeForce RTX 3090;显卡内存40 GB;CPU AMD Ryzen 9 5900X 12-Core Processor@3.70 GHz;内存32 GB。

5.1 数据集来源

墓室壁画现存量少,且年代、绘画风格、表现的历史内容以及保存状况有很大差异,同时考虑到壁画内容信息较为单一、地表沉降导致的地仗断裂错位等因素,现存壁画纹理信息结构较为复杂,因此对残存壁画有效信息的采集以及预处理方法变得尤为重要。

以唐章怀太子墓“马球图”为例,使用瑞士的Sinar P2型大画幅技术相机和Sinar 75LV数字后背进行数据采集,色域模型为Adobe RGB。采集过程使用施耐德APO镜头,德国巴赫聚光灯照明,按照同一行的拍摄方向从左至右依次拍摄,相邻两张图像的重叠范围在40%~50%之内。拍摄全程为非接触式,分镜头平行采集,最终获取了每块数据量在10 GB以上的高清壁画图像。采集后的壁画分割成像素大小为256×256的壁画分块,并筛选出744张纹理细节较为丰富的壁画样本。

为避免壁画分块的外延性信息在数字生成过程中受到训练样本的病害影响,采用课题组前期Image Inpainting数字修复技术^[23-24]的成果对训练样本进行病害清除修复,修复结果如图8所示。

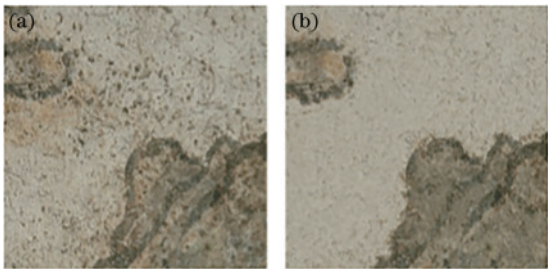


图8 马球图壁画修复前后对比图。(a)原始壁画;(b)修复壁画
Fig. 8 Comparison of polo murals before and after restoration.
(a) Original mural; (b) restored mural

为了增强模型的泛化性能和解决数据量较少出现的过拟合问题,使用数据增广技术^[25]使数据集尽可能多样化,其主要方法有翻转、旋转、剪切、平移、对比度、色彩抖动、风格迁移^[26]。增广后的有效数据量达到7440张,其中:训练数据5952张、测试数据744张、验

证数据744张。在训练过程中,经过反复训练将学习率设置为0.0003,Batchsize为8,使用Adam优化器,训练步骤epoch为200。

5.2 超参数设置实验

为了加快模型收敛速度,分别设置外延大尺度生成网络损失超参数 χ_1 和深度语义小尺度生成网络损失超参数 $\chi_2 = 1 - \chi_1$ 。因为深度语义小尺度生成网络在整个生成网络性能中起关键作用,所以当 $\chi_1 < \chi_2$ 时收敛速度会更快,但也不宜过大。当 χ_2 趋近于1时,外延大尺度生成网络损失反馈为0,得不到反馈损失权重信息,无法进行壁画数字信息生成。设置7组实验,得到不同取值下模型收敛情况,如表1所示。

表1 不同取值下模型收敛情况

χ_1	χ_2	Epoch
0.7	0.3	188
0.5	0.5	166
0.4	0.6	143
0.3	0.7	133
0.2	0.8	110
0.1	0.9	Non convergence
0.001	0.999	Non convergence

由表1可知:当 $\chi_1 = 0.2$ 、 $\chi_2 = 0.8$ 时,模型收敛速度较快,收敛曲线如图9所示。

图9中:train G_loss为训练集生成网络损失;train D_loss为训练集判别网络损失;test G_loss为测试集生成网络损失;test D_loss为测试集判别网络损失。当G_loss与D_loss达到“纳什均衡”状态时模型训练效果最佳。由图9可知,在epoch为110时,模型达到“纳什均衡”状态,故将超参数 χ_1 设置为0.2、 χ_2 设置为0.8。

5.3 自注意力机制生成实验

为了能捕获全局特征像素的依赖关系,增强感受视野和获得上下文信息,将自注意力机制放在编码器阶段来指导解码器生成全局信息相关联的图像。因引入自注意力机制前需要较为丰富的特征图信息,选取第2层之后的特征图进行实验,且编码器下采样次数越多图像分辨率越低,提取特征越不明显。所以设计3组对比实验,即第2层与第3层、第3层与第4层、第4层与第5层分别加入自注意力机制并输出通道特征,如图10所示。

由图10可知,第2层与第3层通道特征图添加自注意力机制前后对比效果不明显,这是因为引入自注意力机制前由于没有提取足够特征,自注意力表达相关性较弱。第4层与第5层后分辨率较低,提取特征不明显,引入自注意力后生成效果不佳。而在第3层与第4层内引入自注意力机制,通道特征图的特征区域信息表现较为明显,呈现出多区域间高亮状态,有利于网络捕获更多的特征信息并增强特征信息间的关联性。为能客观评价在不同层间添加自

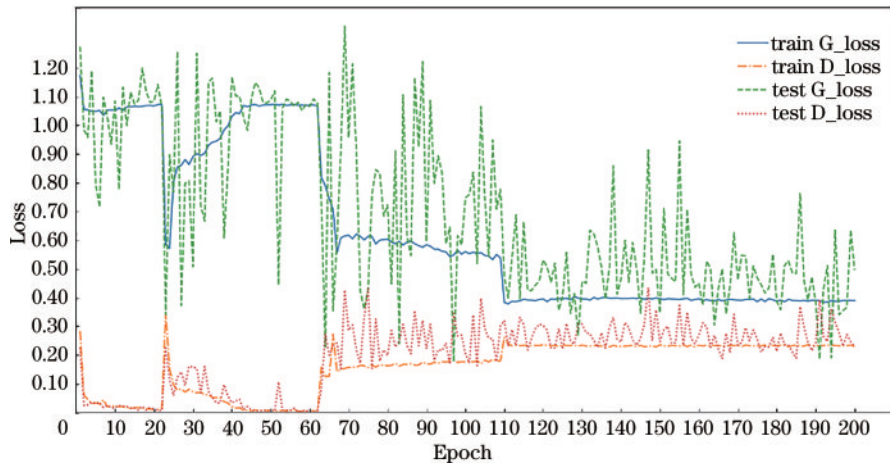


图 9 Loss 收敛曲线图

Fig. 9 Loss convergence curve

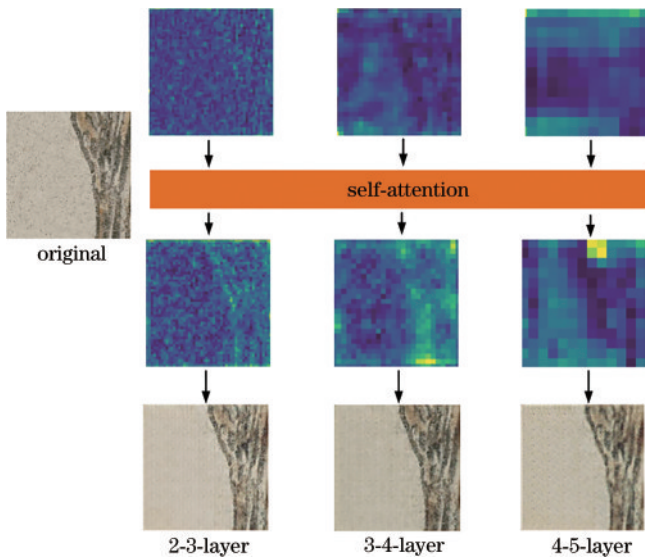


图 10 通道特征对比图

Fig. 10 Comparison diagram of channel characteristics

及 SSIM 作为评价指标,不同层间生成结果的 SSIM 和 PSNR 如表 2 所示。

表 2 不同层间的生成结果分析

Table 2 Analysis of generation results between different layers

Location of self attention mechanism	SSIM	PSNR /dB
2-3-layer	0.74	31.02
3-4-layer	0.84	31.86
4-5-layer	0.69	20.09

由表 2 可知,第 3 层与第 4 层的 PSNR 和 SSIM 结果均高于第 2 层与第 3 层和 4 层与第 5 层,验证了对图 10 的实验结果,故在设计网络时将自注意力机制放在下采样第 3 块和第 4 块中间。

5.4 消融实验

为体现多尺度级联网络的壁画外延数字信息生成效果,进行了消融实验,选取类型分别为树木、山、墓道的 3 幅壁画进行数字生成实验,实验结果如图 11 所示。

注意力机制后壁画外延信息的生成效果,采用 PSNR

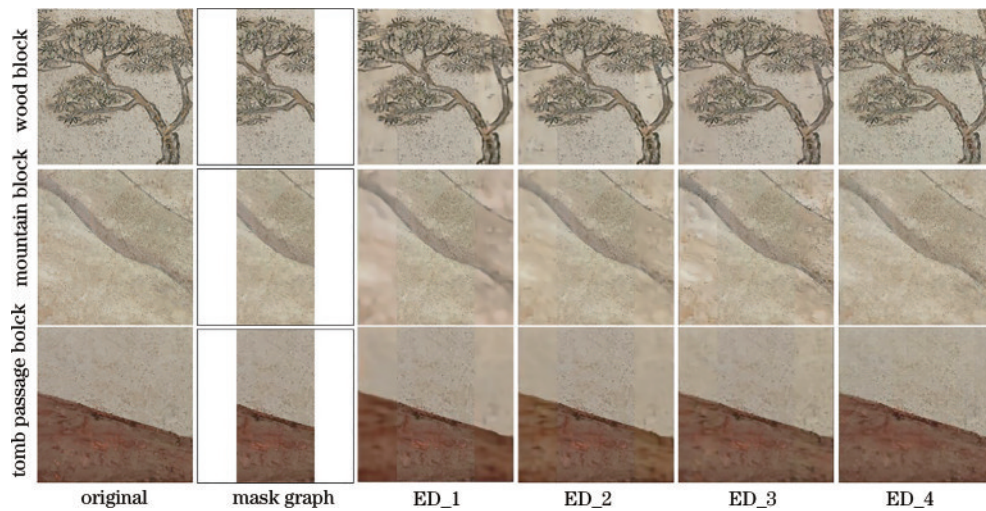


图 11 消融实验

Fig. 11 Ablation experiment

图 11 中:original 为原壁画,mask graph 为掩码壁画,由消融实验可以看出 ED_1 网络生成结果外延效果较差,生成边界线明显,整体不连贯,且纹理细节信息较为模糊。ED_2 引入纹理损失后,生成部分高频纹理细节质量提升,但仍存在边界伪影残留。ED_3 引入改进纹理损失能够生成高频纹理细节,边界伪影效果进一步弱化。ED_4 相比于其他未改进网络具有更好的生成结果,高频纹理细节丰富且边界伪影处过渡平滑,在人眼视觉层面违和度最小。网络与损失的组合名称如表 3 所示。

表 3 网络与损失组合名称
Table 3 Network and loss group name

Name	Network
ED_1	$ED + L_{rec} + L_{adv}$
ED_2	$ED + L_{rec} + L_{adv} + L_{texture-loss}$
ED_3	$ED + L_{rec} + L_{adv} + L_{tex}$
ED_4	$ED + L_{rec} + L_{adv} + L_{tex} + self-attention$

表 3 中:ED 为外延大尺度生成网络和深度语义小尺度生成网络的级联; L_{rec} 为重构损失; L_{adv} 为对抗损失; $L_{texture-loss}$ 为纹理损失; L_{tex} 为优化纹理损失;self-attention 为自注意力机制。为能客观评价壁画的外延信息生成效果,实验采用 PSNR 及 SSIM 作为评价指标进行评价,其结果如表 4 所示。

由表 4 可知,本文方法的 PSNR 提高了 0.70~

表 4 不同方法生成结果对比

Table 4 Comparison of generated by different methods

Method	Wood block		Mountain block		Tomb passage block	
	PSNR / dB	SSIM	PSNR / dB	SSIM	PSNR / dB	SSIM
ED_1	25.29	0.69	27.61	0.62	27.86	0.70
ED_2	28.23	0.73	29.69	0.71	30.40	0.77
ED_3	30.66	0.80	31.74	0.83	31.91	0.82
ED_4	31.03	0.85	32.31	0.86	32.82	0.88

1.94 dB,SSIM 提高了 0.06~0.13。说明本文方法能兼顾全局信息进行壁画的数字生成,解决了壁画外延边界伪影问题,在保证外延生成信息高频纹理细节不丢失的情况下同时具有视觉效果上的连贯性。

5.5 对比实验

为进一步体现多尺度级联网络墓室壁画数字生成技术的优越性,本文的对比实验选取 4 幅马球图进行生成效果对比,生成结果如图 12 所示。由生成结果得:WGAN-GP 能够重建出生成区域大体轮廓,但纹理生成精度较低且存在伪影残留问题;DCGAN 使用空洞卷积进行信息重建,提高了数字生成精度,但在视觉效果上存在色彩缺失及边缘伪影残留问题;本文算法相比其他两种算法虽在色彩部分存在较小偏差,但整体生成效果违和度低且外延边界过渡平滑,同时边

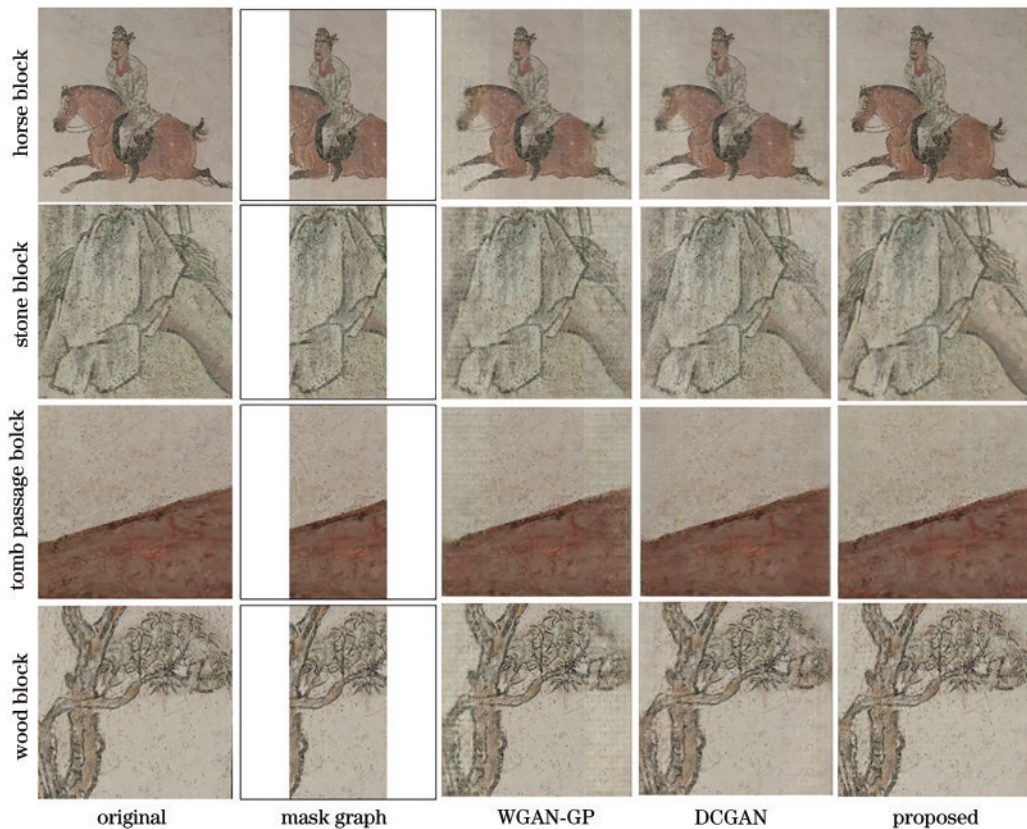


图 12 不同算法生成结果对比

Fig. 12 Comparison of generation results of different algorithms

沿噪声点可被滤除,如墓道左上角小部分噪声被滤除。

为客观评价生成效果,使用 PSNR 和 SSIM 对生成结果进行评价,结果如表 5 所示。

由表 5 可知,本文算法较 WGAN-GP 的 PSNR 提

高了 4.81~6.87 dB, SSIM 提高了 0.16~0.23, 较 DCGAN 算法的 PSNR 提高了 1.32~2.50 dB, SSIM 提高了 0.06~0.12, 说明本文算法在壁画数字信息生成上可取得较好效果。

表 5 不同算法生成结果对比

Table 5 Comparison of generation results of different algorithms

Method	Horse block		Stone block		Tomb passage block		Wood block	
	PSNR /dB	SSIM	PSNR /dB	SSIM	PSNR /dB	SSIM	PSNR /dB	SSIM
WGAN-GP	27.35	0.68	24.87	0.57	26.36	0.65	27.85	0.64
DCGAN	30.16	0.78	29.69	0.72	31.59	0.80	29.47	0.72
Proposed	32.66	0.84	31.74	0.80	32.91	0.88	31.66	0.84

6 结 论

针对墓室壁画因揭取过程导致的壁画分块画面连续缺失问题,建立了一种多尺度级联网络的墓室壁画数字生成模型。生成网络中采用两个级联递进关系的多尺度网络指导壁画数字信息生成,并优化损失函数。实验结果表明:该技术外延结果在主观视觉层面可达到生成信息细节纹理丰富、外延边界过渡平滑的效果;在自建墓室壁画数据集上的客观评价指标 PSNR 及 SSIM 皆优于其他算法,为整幅大型壁画数字化展示提供可行方案,增强优秀传统文化的宣传力度。未来有望借助更多同年代古代墓室壁画样本进行数据集的推广,提高模型泛化性,进一步提升数字外延效果。

参 考 文 献

- [1] 杨芝荣. 用土办法揭取古代壁画试验成功[J]. 文物参考资料, 1958(10): 15.
Yang Z R. The experiment of uncovering ancient frescoes with earth method was successful[J]. Cultural Relics, 1958(10): 15.
- [2] 陕西省博物馆、乾县文教局唐墓发掘组. 唐章怀太子墓发掘简报[J]. 文物, 1972(7): 13-25, 68.
Shaanxi Provincial Museum and Tang tomb excavation group of Xi County Bureau of culture and education. Excavation of prince Zhang Huai's tomb in Tang dynasty [J]. Cultural Relics, 1972(7): 13-25, 68.
- [3] 王建荣. 唐墓壁画高精度影像数字化采集技术与应用[J]. 文博, 2011(3): 66-70.
Wang J R. Digital acquisition technology and application of high-precision images of Tang tomb murals[J]. Relics and Museology, 2011(3): 66-70.
- [4] Karianakis N, Maragos P. An integrated system for digital restoration of prehistoric Theran wall paintings [C]//2013 18th International Conference on Digital Signal Processing (DSP), July 1-3, 2013, Fira, Greece. New York: IEEE Press, 2013.
- [5] Jaidilert S, Farooque G. Crack detection and images inpainting method for Thai mural painting images[C]//2018 IEEE 3rd International Conference on Image, Vision and Computing, June 27-29, 2018, Chongqing, China. New York: IEEE Press, 2018: 143-148.
- [6] Criminisi A, Perez P, Toyama K. Region filling and object removal by exemplar-based image inpainting[J]. IEEE Transactions on Image Processing, 2004, 13(9): 1200-1212.
- [7] Purkait P, Ghorai M, Samanta S, et al. A patch-based constrained inpainting for damaged mural images[M]//Mallik A, Chaudhury S, Chandru V, et al. Digital Hampi: preserving Indian cultural heritage. Singapore: Springer, 2017: 205-223.
- [8] 陈永, 陶美凤, 艾亚鹏, 等. 基于 Gabor 变换和组稀疏表示的敦煌壁画修复算法[J]. 激光与光电子学进展, 2020, 57(22): 221015.
Chen Y, Tao M F, Ai Y P, et al. Algorithm for Dunhuang mural inpainting based on Gabor transform and group sparse representation[J]. Laser & Optoelectronics Progress, 2020, 57(22): 221015.
- [9] 陈永, 陈锦, 艾亚鹏, 等. 基于序贯相似性和布谷鸟寻优的敦煌壁画修复算法[J]. 激光与光电子学进展, 2020, 57(20): 201020.
Chen Y, Chen J, Ai Y P, et al. Dunhuang mural inpainting algorithm based on sequential similarity detection and cuckoo optimization[J]. Laser & Optoelectronics Progress, 2020, 57(20): 201020.
- [10] Pathak D, Krähenbühl P, Donahue J, et al. Context encoders: feature learning by inpainting[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition, June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 2536-2544.
- [11] Iizuka S, Simo-Serra E, Ishikawa H. Globally and locally consistent image completion[J]. ACM Transactions on Graphics, 2017, 36(4): 107.
- [12] Paunwala C N. Image inpainting evolution: a survey[M]//Encyclopedia of image processing. Boston: CRC Press, 2018: 319-333.
- [13] Laube P, Grunwald M, Franz M O, et al. Image inpainting for high-resolution textures using CNN texture synthesis[EB/OL]. (2017-12-08) [2021-02-05]. <https://arxiv.org/abs/1712.03111>.
- [14] 彭广泽, 陈文静. 基于卷积神经网络去噪正则化的条纹图修复[J]. 光学学报, 2020, 40(18): 1810002.
Peng G Z, Chen W J. Fringe pattern inpainting based on convolutional neural network denoising regularization[J].

- Acta Optica Sinica, 2020, 40(18): 1810002.
- [15] Sabini M, Rusak G. Painting outside the box: image outpainting with GANs[EB/OL]. (2018-08-25)[2021-02-05]. <https://arxiv.org/abs/1808.08483>.
- [16] Goodfellow I J, Pouget-Abadie J, Mirza M, et al. Generative adversarial networks[EB/OL]. (2014-06-10)[2021-02-05]. <https://arxiv.org/abs/1406.2661>.
- [17] 陈俊周, 王娟, 龚勋. 基于级联生成对抗网络的人脸图像修复[J]. 电子科技大学学报, 2019, 48(6): 910-917.
Chen J Z, Wang J, Gong X. Face image inpainting using cascaded generative adversarial networks[J]. Journal of University of Electronic Science and Technology of China, 2019, 48(6): 910-917.
- [18] Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need[EB/OL]. (2017-06-12)[2021-02-05]. <https://arxiv.org/abs/1706.03762>.
- [19] Risser E, Wilmot P, Barnes C. Stable and controllable neural texture synthesis and style transfer using histogram losses[EB/OL]. (2017-01-31)[2021-02-03]. <https://arxiv.org/abs/1701.08893>.
- [20] Gulrajani I, Ahmed F, Arjovsky M, et al. Improved training of Wasserstein GANs[EB/OL]. (2017-03-31)[2021-02-05]. <https://arxiv.org/abs/1704.00028>.
- [21] 朱秀昌, 唐贵进. 可视质量评价方法综述[J]. 南京邮电大学学报(自然科学版), 2018, 38(4): 1-11.
Zhu X C, Tang G J. Overview of visual quality assessment methods[J]. Journal of Nanjing University of Posts and Telecommunications (Natural Science Edition), 2018, 38(4): 1-11.
- [22] 文义红, 杨凯, 李波. 静态图像压缩评估方法测评[J]. 中国空间科学技术, 2010, 30(1): 72-77.
Wen Y H, Yang K, Li B. Performance test for image quality assessment[J]. Chinese Space Science and Technology, 2010, 30(1): 72-77.
- [23] 李彩艳, 王慧琴, 吴萌, 等. 唐墓室壁画泥斑病害自动标定及虚拟修复[J]. 计算机工程与应用, 2016, 52(15): 233-236.
Li C Y, Wang H Q, Wu M, et al. Automatic recognition and virtual restoration of mud spot disease of Tang dynasty tomb murals image[J]. Computer Engineering and Applications, 2016, 52(15): 233-236.
- [24] 申婧妮, 王慧琴, 吴萌, 等. MCA 分解的唐墓室壁画修复算法[J]. 计算机科学与探索, 2017, 11(11): 1826-1836.
Shen J N, Wang H Q, Wu M, et al. Tang dynasty tomb murals inpainting algorithm of MCA decomposition[J]. Journal of Frontiers of Computer Science and Technology, 2017, 11(11): 1826-1836.
- [25] 马崇昇, 唐娉, 赵理君, 等. 深度学习图像数据增广方法研究综述[J]. 中国图象图形学报, 2021, 26(3): 487-502.
Ma D A, Tang P, Zhao L J, et al. Review of data augmentation for image in deep learning[J]. Journal of Image and Graphics, 2021, 26(3): 487-502.
- [26] Park D Y, Lee K H. Arbitrary style transfer with style-attentional networks[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 15-20, 2019, Long Beach, CA, USA. New York: IEEE Press, 2019: 5880-5888.