

5-25-2022

Murals Super-resolution Reconstruction with the Stable Enhanced Generative Adversarial Network

Jianfang Cao

1.College of Computer Science and Technology, Taiyuan University of Science and Technology, Taiyuan 030024, China;2.Department of Computer, Xinzhou Teachers University, Xinzhou 034000, China;;
kcxdj122@126.com

Yiming Jia

1.College of Computer Science and Technology, Taiyuan University of Science and Technology, Taiyuan 030024, China;

Minmin Yan

1.College of Computer Science and Technology, Taiyuan University of Science and Technology, Taiyuan 030024, China;

Xiaodong Tian

1.College of Computer Science and Technology, Taiyuan University of Science and Technology, Taiyuan 030024, China;

Follow this and additional works at: <https://dc-china-simulation.researchcommons.org/journal>



Part of the Artificial Intelligence and Robotics Commons, Computer Engineering Commons, Numerical Analysis and Scientific Computing Commons, Operations Research, Systems Engineering and Industrial Engineering Commons, and the Systems Science Commons

This Paper is brought to you for free and open access by Journal of System Simulation. It has been accepted for inclusion in Journal of System Simulation by an authorized editor of Journal of System Simulation.

Murals Super-resolution Reconstruction with the Stable Enhanced Generative Adversarial Network

Abstract

Abstract: Aiming at the problems of low resolution and unclear texture details of ancient murals, which led to insufficient viewing of murals and low research value, a stable enhanced super-resolution generative adversarial networks (SESRGAN) reconstruction algorithm is proposed. *Based on the generative adversarial network, the generative network uses dense residual blocks to extract mural features, and uses the visual geometry group (VGG) network as the basic framework of the discriminating network to determine the authenticity of the input mural, and introduces perception loss, content loss and penalty loss to jointly optimize the model.* Experimental results show that, compared with other related super-resolution algorithms, the peak signal-to-noise ratio (PSNR) is improved by 0.4~2.62 dB on average, the structural similarity is improved by 0.013~0.027, and the subjective perception evaluation is also improved.

Keywords

ancient mural, super-resolution reconstruction, generation adversarial network, dense residual block, penalty loss

Recommended Citation

Jianfang Cao, Yiming Jia, Minmin Yan, Xiaodong Tian. Murals Super-resolution Reconstruction with the Stable Enhanced Generative Adversarial Network[J]. Journal of System Simulation, 2022, 34(5): 1076-1089.

稳定增强生成对抗网络在壁画的超分辨率重建

曹建芳^{1,2}, 贾一鸣¹, 闫敏敏¹, 田晓东¹

(1. 太原科技大学 计算机科学与技术学院, 山西 太原 030024; 2. 忻州师范学院 计算机系, 山西 忻州 034000)

摘要: 针对古代壁画分辨率低、纹理细节模糊不清导致壁画观赏性不足和研究价值不高的问题, 提出了一种稳定增强生成对抗网络的超分辨率重建算法(stable enhanced super-resolution generative adversarial networks, SESRGAN)。以生成对抗网络为基础框架, 生成网络采用密集残差块提取壁画特征, 使用VGG(visual geometry group)网络作为判别网络的基本框架判断输入壁画的真假, 引入感知损失、内容损失和惩罚损失三个损失共同优化模型。实验结果表明, 与其他相关的超分辨率算法进行比较, 峰值信噪比平均提高了0.4~2.62 dB, 结构相似性提高了0.013~0.027, 主观感知评估也有提高。

关键词: 古代壁画; 超分辨率重建; 生成对抗网络; 密集残差块; 惩罚损失

中图分类号: TP391.47

文献标志码: A

文章编号: 1004-731X(2022)05-1076-14

DOI: 10.16182/j.issn1004731x.joss.20-0989

Murals Super-resolution Reconstruction with the Stable Enhanced Generative Adversarial Network

Cao Jianfang^{1,2}, Jia Yiming¹, Yan Minmin¹, Tian Xiaodong¹

(1. College of Computer Science and Technology, Taiyuan University of Science and Technology, Taiyuan 030024, China;

2. Department of Computer, Xinzhou Teachers University, Xinzhou 034000, China)

Abstract: Aiming at the problems of low resolution and unclear texture details of ancient murals, which led to insufficient viewing of murals and low research value, a stable enhanced super-resolution generative adversarial networks (SESRGAN) reconstruction algorithm is proposed. Based on the generative adversarial network, the generative network uses dense residual blocks to extract mural features, and uses the visual geometry group (VGG) network as the basic framework of the discriminating network to determine the authenticity of the input mural, and introduces perception loss, content loss and penalty loss to jointly optimize the model. Experimental results show that, compared with other related super-resolution algorithms, the peak signal-to-noise ratio (PSNR) is improved by 0.4~2.62 dB on average, the structural similarity is improved by 0.013~0.027, and the subjective perception evaluation is also improved.

Keywords: ancient mural; super-resolution reconstruction; generation adversarial network; dense residual block; penalty loss

引言

中国古代壁画曾经有辉煌历史, 历经数千年数朝代积累沉淀, 出现过如敦煌壁画那样至今仍然令

人叹为观止的艺术经典。一直以来, 对中国古代壁画的艺术研究多数立足于对传统绘画技术的发掘、探究和总结。重现作品昔日的风采成为了研究的基础。由于古代壁画经过历史上的人为破坏和大自然

收稿日期: 2020-12-10

修回日期: 2021-03-10

基金项目: 山西省高等学校人文社会科学重点研究基地项目(20190130)

第一作者: 曹建芳(1976-), 女, 博士, 教授, 研究方向为数字图像处理\大数据技术。E-mail: kcxj122@126.com

的洗礼之后,原本平整清晰的图像已是被一层薄雾所遮盖,壁画纹理细节已经变得模糊不清,部分壁画已破损,现存的古代壁画产生了不同程度和种类病害,要对现存的壁画进行预防性保护和修复。师晓波^[1]等针对古代墓葬壁画中不同形态的裂缝,提出基于样本块修复算法对壁画裂缝进行修补,Wang^[2]等提出利用最有方向算法对壁画受损区域及进行修复。曹建芳^[3]等针对壁画起甲、脱离等问题,使用生成对抗网络对纹理结构强、缺失区域面积较大的壁画进行修复。其中一项重要的方法是对现有的古代壁画进行超分辨率重建,超分辨率重建是一种图像处理的技术,可以提高图像的识别能力和识别精度。将一组低分辨率图像作为输入,经过程序最终产生单幅高质量、高分辨率图像。马强^[4]等利用基于加权核范数稀疏矩阵模型降低了壁画重建的伪影,但是此算法比传统稀疏矩阵表示方法计算成本要高;徐志刚^[5]等针对壁画边缘锯齿效应的问题将注意力网络和残差网络融合对壁画进行重建,壁画在纹理细节方面也有了提升,但是增加了网络训练时间并且容易使网络陷入局部最优。

超分辨率算法可以分为基于插值的方法、基于重建的方法和基于学习的方法。基于插值的方法是较早提出来的相对简单的算法。通过计算低分辨率图像和高分辨率图像之间的配准关系,高分辨率图像由合适的插值算法插值得到。常见的算法有双线性插值^[6]、双三次插值(bicubic interpolation, BI)^[7]和最近邻插值^[8]。但此方法是适应性比较差,处理场景比较单一,往往会出现边缘模糊和细节丢失的问题。基于重建的方法是在得到配准关系的基础上,得出高分辨率和低分辨率图像之间的依赖关系作为先验知识,然后利用其重建出高分辨率的图像。常见的算法有迭代反向投影法^[9],凸集投影法^[10]。基于学习的方法是近年来超分辨率算法研究的热点,也是本文使用的方法。它采用大量的高清超分辨率图片和对应的低清分辨率图片对设计的模型进行不断的训练,从而使模型不仅能够恢复清晰的图像,而且使图像具有更多高频细节的内容。

Chang 等^[11]利用局部线性嵌入流形学习的思想提出了基于领域嵌入的超分辨率重建算法,Yang 等^[12]利用稀疏信号提出了基于稀疏表示的重建算法。Dong 等^[13]提出基于卷积神经网络的超分辨率重建 SRCNN(super-resolution convolutional neural network),SRCNN 是深度学习在超分辨率上的开山之作,首次使用双三次差值将低分辨率图像放大至原图像的尺寸,仅仅用三层卷积和非线性映射最终输出产生高分辨率的图像。将三层卷积结构解释成 3 个步骤:图像块的提取和特征表示,特征表示的非线性映射和最终的重建工作。Dong 等^[14]提出 FSRCNN(fast SRCNN),在网络上将 SRCNN 进行了改进,在最后一层使用反卷积层放大尺寸,使用更小的卷积核和更多的映射层,同时添加了收缩层和扩张层,并且对数据集进行了数据增强。Tong 等^[15]提出 SRDensNet,采用的稠密块将每一层的特征输入给之后的所有层,使所有层特征串联起来,这样的结构减轻了梯度相似的问题,加强了特征的传递,减少了参数的数量。Ledig C 等^[16]等提出基于生成对抗网络的超分辨率重建 SRGAN (super-resolution generative adversarial networks),将生成对抗网络应用在解决超分辨率的问题中,利用感知损失和对抗损失减轻了图像丢失高频细节、图像平滑的问题,使人有较好的视觉感受,但是得到的峰值信噪比不是很高。Arjovsky 等^[17]针对生成对抗网络(generative adversarial network, GAN)训练不稳定的问题,提出了 WGAN(wasserstein GAN)理论,采用 wasserstein 距离代替(jensen-shannon, JS)散度来衡量真实图像与重建图像之间的距离,由于使用权重剪枝策略,还是会出现梯度爆炸和梯度消失等问题。Wang 等^[18]对 SRGAN 做出了改进,生成网络采用密集残差块进行特征提取,使用 relativistic discriminator 作为判别器,使生成器生成更加真实的纹理细节,保留更多的图像特征信息,但仍存在较多伪影的问题。

上述算法在公开数据集上有良好表现,但是由于壁画数据集缺乏,不足以支撑起网络的训练深

度,并且在训练网络模型的过程中不稳定导致壁画色彩特征信息提取不完整、纹理细节缺失、亮度不一致、重建效果不理想等问题。因此本文针对上述问题提出了 SESRGAN(stable enhanced SRGAN)算法,通过结合残差网络和密集网络的特征学习方式,堆叠多个密集残差块、去除批归一化层(batch normalization, BN)增强模型的性能,充分捕获壁画色彩图像的特征信息来提高模型的泛化能力和鲁棒性;通过引入感知损失、对抗损失和内容损失共同优化网络,在细节方面采用梯度惩罚理论 WGAN-GP(WGAN-gradient penalty)^[19]进一步优化对抗损失;受袁功霖等^[20]二次迁移学习的启发,本文将公开数据集和壁画数据集进行联合训练,以获得较好的初始权重值训练模型,最终提高了壁画重建的质量和清晰度。

1 基本理论

1.1 生成对抗网络

GAN 是近年来比较热门的一种深度学习的模型,是在复杂分布上无监督学习最具前景的方法之一。生成网络首先捕捉图像中随机噪声的分布,然后通过这些噪声的分布生成一个类似于真实数据的样本作为判别网络的输入,判别网络作用是估计一个样本来自于训练数据的概率大小,从而去判断输入的数据是真实数据还是生成样本。在训练网络的过程中,生成网络不断提高生成真实样本的能力以骗过判别网络,而判别网络不断提高识别样本真伪的能力。通过生成网络和判别网络的互相博弈中不断学习从而优化模型,随着时间的推移,生成网络和对抗网络不断训练优化最终 2 个网络达到一个动态的平衡:生成样本近似于真实样本的分布,判别网络对给定的样本判别概率为 0.5。生成对抗网络损失公式为

$$\min_G \max_D V(D, G) = E_{x \sim p_x} [\ln D(x)] + E_{z \sim p_z} [\ln(1 - D(G(z)))] \quad (1)$$

式中: x 为真实图像数据; z 为噪声图像数据, p_x 为真实图像数据的概率分布; p_z 为生成数据的概率

分布; $G(z)$ 为重建后的图像数据; $D(x)$ 为判别网络判断真实图像数据是否真实的概率; $D(G(z))$ 为判别网络判断重建后图像数据是否真实的概率; $\ln D(x)$ 为判别网络对真实图像数据的判断; $\ln(1 - D(G(z)))$ 为对重建后图像数据的判断。生成对抗网络的结构示意图如图 1 所示。

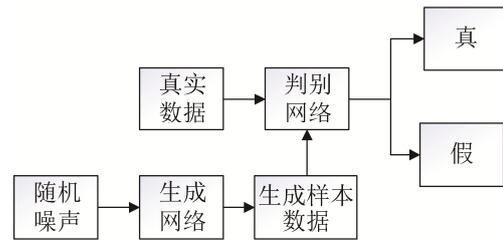


图 1 生成对抗网络结构图
Fig. 1 Structure of GAN

1.2 残差网络和密集网络

残差网络(ResNet)是为了解决卷积神经网络深度变深时网络性能饱和、退化,容易出现梯度弥散或者消失并且训练精度和测试精度会下降的问题,残差网络结构图如图 2 所示,假设上一层的输出即本层的输入为 x ,本层的残差为 $F(x)$ 将下一层的输出为 $H(x)$;通过跳跃连接(skip connection)的方法将 x 传递给下一层 $H(x)$ 。那么 $H(x)=F(x)+x$,这一层的残差就可以表示为 $F(x)=H(x)-x$,这种残差结构可以在比较深的网络中提升性能,本文引入残差块作为生成网络的一部分。

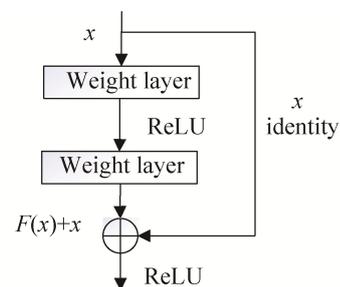


图 2 残差网络结构图
Fig. 2 Structure of ResNet

密集网络(DenseNet)致力于从特征重用的角度来提升网络的性能,是一种具有密集连接的卷积神经网络并且任意两层之间都有着紧密的联系。

DenseNet 的每一层都收到来自所有先前层输入, 然后将其特征映射输出到所有的后续层, 因此提升了信息和梯度在网络中的传输效率。这种结构的好处在于可以实现特征的重用, 可以缓解梯度消失, 加强特征传播, 大幅减少了参数与数量。DenseNet 网络结构如图 3 所示, CONV 表示卷积层, ReLU 表示 ReLU 激活函数层, BN 表示批归一化层。

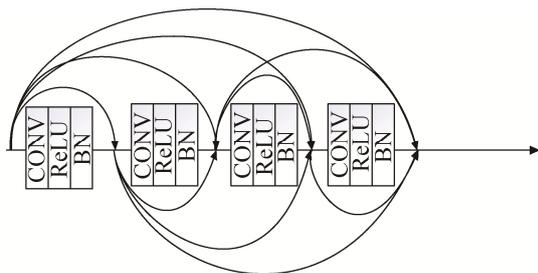


图 3 密集网络结构图

Fig. 3 Structure of DenseNet

1.3 梯度惩罚理论

梯度惩罚理论 WGAN-GP^[19]是针对 WGAN 存在的问题提出来的, WGAN 在真实的实验过程中依旧存在着训练困难、收敛速度慢的问题, 相比较传统 GAN 在实验上提升不是很明显。WGAN-GP 指出了 WGAN 存在问题的原因, WGAN 在处理 Lipschitz 限制条件的过程中直接采用了权重裁剪 (weight clipping), 每当更新完一次判别器的参数之后, 就检查判别器的所有参数的绝对值有没有超过一个阈值 n , 有的话就把这些参数裁剪回 $[-n, n]$ 范围内。通过在训练过程中保证判别器的所有参数有界, 就保证了判别器不能对 2 个略微不同的样本在判别上不会差异过大, 从而间接实现了 Lipschitz 限制, 但是这样做导致了绝大部分的 weight 趋近于两个极端。因此 WGAN-GP 做了如下改进: 使用梯度惩罚 (gradient penalty) 取代了权重裁剪, 采用 Adam 优化器取代 RMSprop 优化器, 在生成图像上增加高斯噪声。实验结果表明, 该方法相比 WGAN 而言, 能够更稳定训练 GAN 模型, 几乎不需要反复调整超参数, 并且使模型快速的收敛并生成质量较好的壁画。

2 稳定增强的生成对抗网络算法

由于壁画注重色彩协调和搭配, 背景基调淡雅, 但是人物色彩浓重富有变化、不同种类壁画的内容和其色彩地位存在差异等特点, 结合生成对抗网络的特性, 从网络结构设计和损失函数的引入做出以下改进:

(1) 在生成网络中使用含有残差缩放的密集残差块代替原生成网络中的残差块对图像的深层特征充分提取壁画色彩和内容信息, 并且去除密集块中的 BN 层以获得更多的壁画特征。

(2) 本文在判别网络的设计中使用谱归一化层 (spectral normalization, SN) 代替原网络中的 BN 层, 去除 Sigmoid 层并引入 WGAN-GP 理论优化对抗损失, 增加网络训练的稳定性, 避免了模型训练崩溃。

(3) 在计算感知损失时并且使用激活前而不是激活后的特征信息进行计算, 为壁画的亮度一致性和局部纹理细节恢复提供更强的监督。

2.1 网络总体结构设计

为了更好的重建高分辨率图像, 本文设计的稳定增强生成对抗网络结构 (SESARGAN) 如下图 4 所示。生成网络将低分辨率壁画 LR 作为输入, 经过密集残差网络提取特征, 然后通过上采样和卷积重建图像后输出超分辨率重建壁画 SR, 将 SR 与高分辨率壁画 HR 真值一起输入判别网络, 最后判别网络负责判别输入图像的真假。

2.1.1 残差缩放的密集残差块

为了能提取壁画不同深度的色彩信息和内容特征, 更好的恢复图片质量, 对生成网络作了如下调整: 使用含有残差缩放的密集残差块 (dense residual block with residual scaling, RS-RDB) 代替原有的残差块 (RB) 对输入的图像进行深层特征提取, 由于更多的层和更多的连接总会提高网络的性能。所以结合多残差网络和密集连接加深了网络的深度, RS-RDB 如图 5 所示。为了能提取更多图像特征的同时又防止过深的网络容易造成过拟合, 因此使用了 23 个 RS-RDB, 有利于获取于恢复纹理去

除噪声。1 个 RS-RDB 由 3 个密集块(dense block, DB)组成, 并且引入残差缩放在其中, 将残差乘以一个(0,1)中的数值 β , 增加训练深层网络的稳定性, 其中 Dense Block 由 4 个卷积层 CONV 和 4 个带泄露的激活函数 Leaky ReLU 构成, 为了壁画网络模型的一致性、稳定性, 去除 Dense Block 模块中的 BN 层以减少计算复杂度和内存占用。对于不同的基于峰值信噪比任务中比如超分辨率和去模糊来

说, 去掉 BN 层已经被证明会提高模型的效果和减少计算复杂度, 并且 BN 层会倾向于在壁画中生成不好的伪影, 限制了模型的泛化能力。DB 如图 6 所示。通过上述的密集残差块将输入的壁画特征提取出来, 然后生成高分辨率图像, 使用两层亚向素卷积层(Pixel shuffler)来放大图像的尺寸, 最后使用一个 3×3 的卷积层输出 3 通道的高分辨率图像。

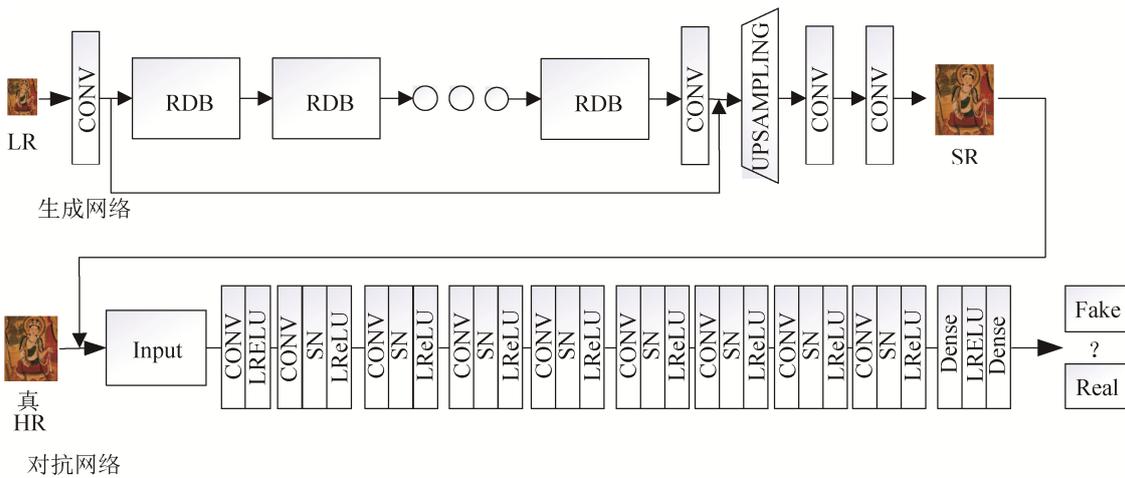


图 4 稳定增强生成对抗网络结构图
Fig. 4 Structure of ESERGAN

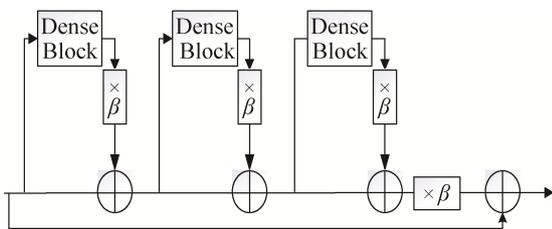


图 5 含有残差缩放的密集残差块
Fig. 5 Dense residual block with residual scaling

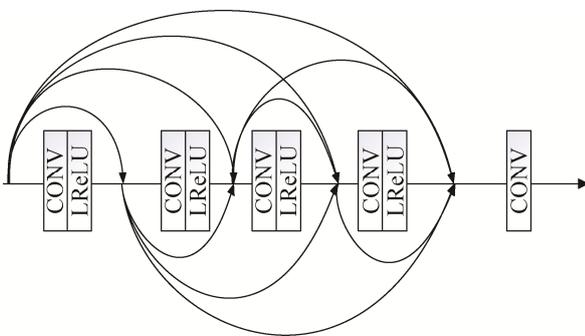


图 6 密集块
Fig. 6 Dense block

2.1.2 判别网络的设计

判别网络首先使用了 64 通道卷积层对输入的壁画进行浅层提取, 然后使用了 8 个卷积层, 为了更好的满足 Lipschitz 约束, 每个卷积层 CONV 后面归一化层使用 SN 层而不是 BN 层, 即将参数矩阵除以它的谱矩阵, 不会破坏矩阵内部参数关系。谱归一化的作用是对判别网络中结构参数进行限制, 类似引入一个新的正则化项, 防止网络训练过程中总朝着一个方向训练, 有助于生成网络训练达到更好的效果, 属于过程参数的优化, 因此在模型的可行性证明中对 GAN 的目标函数无影响。采用 Leaky Relu 作为激活函数, 由于用 Wasserstein 距离代替 JS 散度来度量壁画真实样本和壁画生成样本之间的距离属于回归任务, 所以本文做出如下调整: 在最后没有使用原始判别网络设计的 Sigmoid 激活函数, 而是使用 2 个全连接层直接输出图像真实的概率、判别网络信息表如表 1 所示。

表 1 判别网络详情信息表
Table 1 Adversarial Network information

Name	Type	Kernel	Stride	Padding	Output
CONV0_0	CONV	3×3	1×1	1	64
CONV1_0	CONV	4×4	2×2	1	64
CONV1_1	CONV	3×3	1×1	1	64
CONV2_0	CONV	4×4	2×2	1	128
CONV2_1	CONV	3×3	1×1	1	128
CONV3_0	CONV	4×4	2×2	1	256
CONV3_1	CONV	3×3	1×1	1	256
CONV4_0	CONV	4×4	2×2	1	512
CONV4_1	CONV	3×3	1×1	1	512
FC0	FC	-	-	-	100
FC1	FC	-	-	-	1

2.2 损失函数

损失函数用于判断模型重建图像后的好坏, 为了使重建后的图像有更真实的视觉效果, 本文使用感知损失和内容损失对生成网络优化, 使用对抗损失优化对抗网络, 最终损失函数计算公式(2)可表达如下, 其中 λ_1 、 λ_2 表示用来平衡不同损失的系数。

$$l_G = l_{\text{MSE}} + \lambda_1 l_{\text{VGG}} + \lambda_2 l_{\text{adv}} \quad (2)$$

2.2.1 内容损失

为了保证重建后的图像和原始图像信息的准确性, 本文采用的是 MSE(mean squared error) 损失作为生成网络的内容损失。像素级别的损失计算的是生成图像像素与真实图像像素的空间误差, 像素级别的 MSE 损失计算公式(3)如下:

$$l_{\text{MSE}} = \frac{1}{r^2 WH} \sum_{x=1}^{rW} \sum_{y=1}^{rH} (I_{x,y}^{\text{HR}} - G_{\theta_G}(I^{\text{LR}})_{x,y})^2 \quad (3)$$

式中: r 为图像的缩放因子; W , H 为图像的宽度和高度; $I_{x,y}^{\text{HR}} - G_{\theta_G}(I^{\text{LR}})_{x,y}$ 为真实图像像素与生成图像像素之间的欧式距离。

2.2.2 感知损失

MSE 损失能够使重建的图形极高的峰值信噪比, 但是通常缺乏高频内容, 对图像的细节起到了模糊的作用。所以本文在生成网络中又引入 VGG (visual geometry group) 网络中的感知损失, 并且使用激活前而不是激活后的特征信息进行计算。因为壁画人物亮度较明显, 但是壁画背景亮度偏暗, 使

用激活后的特征信息监督较弱, 计算会造成壁画背景信息特征的缺失。VGG 损失是将生成的高分辨率图像和真实的高分辨率图像送入 VGG19 网络进行特征提取, 然后再提取的特征图上再使用均方根误差计算欧氏距离。VGG 损失的计算公式为

$$l_{\text{VGG}/i,j} = \frac{1}{W_{i,j} H_{i,j}} \sum_{x=1}^{W_{i,j}} \sum_{y=1}^{H_{i,j}} (\varphi_{i,j}(I^{\text{HR}}) - \varphi_{i,j} G_{\theta_G}(I^{\text{LR}})_{x,y})^2 \quad (4)$$

式中: $\varphi_{i,j}$ 为从 VGG19 网络的第 j 层卷积(激活前)的第 i 个池化层之间获取的特征图, 将 VGG 损失定义为重构图像的特征表示 $G_{\theta_G}(I^{\text{LR}})$ 与参考图像 I^{HR} 之间的欧氏距离; $W_{i,j}$ 和 $H_{i,j}$ 为 VGG 网络中对应特征图的维数。

2.2.3 对抗损失

为了使模型训练更稳定, 引入文献[19]WGAN-GP 方法, 在 WGAN 基础上进一步完善模型的目标函数, 对判别网络做梯度惩罚, WGAN 的损失函数计算公式为

$$L(D) = E_{x \sim p_r}[D(x)] - E_{x \sim p_g}[D(x)] \quad (5)$$

式中: p_r 为真实数据的分布; p_g 为生成数据的分布。

通过添加 Lipschitz 限制使判别器 $D(x)$ 梯度不大于一个有限的常数 K , 如式(6)所示:

$$\|\nabla_x D(x)\| \leq K \quad (6)$$

当输入样本变化后不出现大幅度的变化。一般将 K 设置为 1, 通过和原来的判别器损失加权合并, 即可得到 WGAN-GP 判别器的对抗损失为

$$l_{\text{adv}} = -E_{x \sim p_r}[D(x)] + E_{x \sim p_g}[D(x)] + \lambda E_{x \sim p_x}[(\|\nabla_x D(x)\|_p - 1)^2] \quad (7)$$

式中: p_x 为真实数据和生成数据随机取样所得的分布。

2.3 SESRGAN 算法流程描述

由于较小的初始化可以帮助网络训练的更优秀, 在 SESRGAN 的生成网络使用 DIV2K 和 Flickr2K 高清数据集预训练一个生成器, 在此基础上继续对网络进行训练。为了清晰的展现出 SESRGAN 算法实现的有效性, 算法具体流程描述如下:

输入：训练数据集由成对的图像组成，包含低分辨率图像数据集 LR 和与之对应的高分辨率图像数据集 HR。

输出：生成的网络 G 和对抗网络 D 的模型。

Step 1: 将 LR 图像输入到生成网络 G 中，经过重建后输出得到生成的高分辨率图像 HR_i 。

Step 2: 计算 HR 图像与 HR_i 图像之间的均方误差 MSE 并更新生成网络的参数。

Step 3: 重复 step1~2 共 n_1 次，得到并保存预训练的生成网络模型 G_1 。

Step 4: 将 LR 图像输入到 G_1 ，输出得到生成的高分辨率图像 HR_i 。

Step 5: 将 HR_i 和对应的 HR 输入到对抗网络 D 中，计算对抗损失 l_{adv} ，更新对抗网络 D 的参数。

Step 6: 将 HR_i 和对应的 HR 输入预训练的 VGG 网络，然后使用激活前的特征值计算感知损失 l_{VGG} 。

Step 7: 计算 HR_i 和对应的 HR 之间的内容损失 l_{MSE} 。

Step 8: 计算生成网络的总损失 l_G ，更新并保存生成网络 G 和对抗网络 D 的模型。

Step 9: 迭代 step 3~8 共 n_2 次，不断更新并保存生成网络的 G 和对抗网络 D 的模型。

SERGAN 算法训练流程图如图 7 所示。

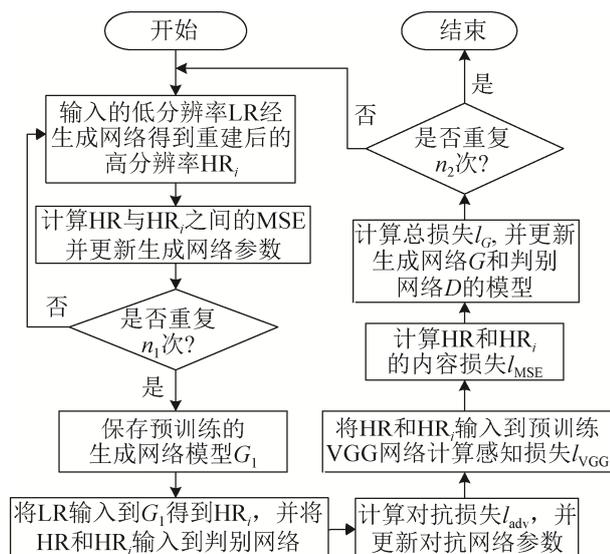


图7 SERGAN 算法训练流程图
Fig. 7 Training flow of SESRGAN

3 实验与结果分析

3.1 实验环境与实验设计

本实验所搭建的硬件环境：CPU 为 Inter Core i7-7700K、内存为 16GB、显卡为 NVIDIA GeForce GTX 1080Ti；所搭建的软件环境：CUDA 版本为 9.0、cuDNN 版本为 7.0、windows10 操作系统；使用 python3.6、采用 pytorch 框架编写测试实验。使用的编译软件为 pycharm2019_3.1_x64。

本实验所用到的训练数据集有 800 张 DIV2K 图像、2 650 张 Flickr2K 图像和 90 张壁画。测试数据集有 30 张壁画。本文的壁画数据集由不同风格和类型的壁画组成。首先对数据集进行扩充：使用翻转，旋转 90° 方法进行数据增强如图 8 所示，本文的高分辨率图像与低分辨率图像的下采样因子为 4。为了加快训练中的 IO 速度，将低分辨率图像裁剪成 120×120 大小的尺寸、高分辨率图像裁剪成 480×480 尺寸送入本文模型。采用的优化器是 Adam 优化器， $\beta_1=0.9$ ， $\beta_2=0.99$ ，设置初始学习率为 0.000 1，学习率衰减参数为 0.5，在生成网络模型中引入了残差缩放策略，其中 $\beta=0.2$ 。结合文献[20]二次迁移学习的思想，使用 DIV2K 数据集在密集残差网络上预训练一个生成器，用来初始化参数以获得更好的质量和更快的收敛速度，然后再将 Flickr2K 图像输入到网络中训练，最后将壁画数据集送入网络继续进行训练，交替更新生成器和判别器模型直至收敛，最后对得出的结果进行以下的实验分析。

3.2 评价指标

为了证明算法的有效性，本文采用常用的客观评价指标有峰值信噪比(peak signal to noise ratio, PSNR)，结构相似性(structural similarity index method, SSIM)2 个指标判断图像重建质量的好坏。两幅图像之间的 PSNR 值(单位 dB)越高代表重建图像与高分辨率图像之间的失真越少，即重建后的效果就越理想；SSIM 则是从亮度、对比度和结构三方面去判断重建后图像的相似度，SSIM 值越接近 1，代表两幅图像相似度越高，更加符合大众的

视觉感官效果。PSNR 的具体计算表达式如式(8)所示, SSIM 的具体计算表达式如式(9)所示。

$$PSNR = 10 \times \lg \frac{225^2 WH}{\sum_{i=1}^W \sum_{j=1}^H [\bar{x}(i, j) - x(i, j)]^2} \quad (8)$$

$$SSIM(X, Y) = \frac{(2\mu_X\mu_Y + C_1)(2\sigma_{XY} + C_2)}{(\mu_X^2 + \mu_Y^2 + C_1)(\sigma_X^2 + \sigma_Y^2 + C_2)} \quad (9)$$

式中: W, H 为图像的宽度和高度; $\bar{x}(i, j) - \bar{x}(i, j)$ 为两幅图像像素点之间的欧式距离; μ_X, μ_Y 分别为图像 X, Y 的均值; σ_X, σ_Y 分别为图像 X, Y 的标准差; σ_{XY} 表示图像 X, Y 的协方差; C_1, C_2 为常数。

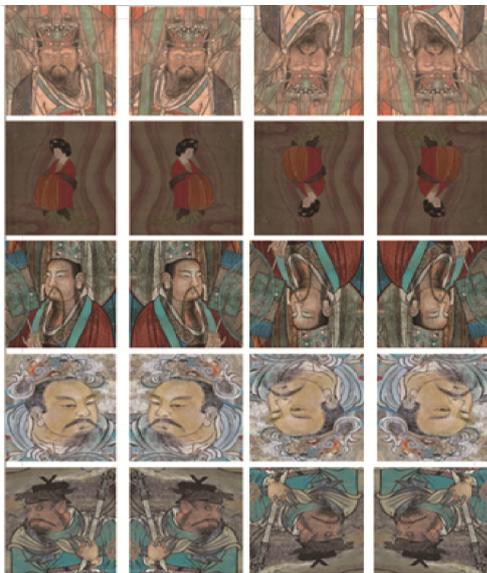


图 8 SESRGAN 算法训练流程图
Fig. 8 Training flow of SESRGAN

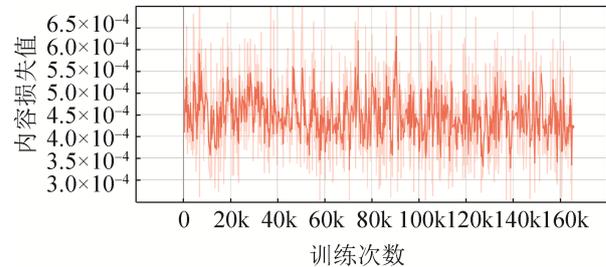
目前经常使用的主观评价标准是平均主观意见(mean opinion score, MOS), 即用 5 个等级(很差、差、中等、良好、优秀)对图像进行打分。针对 MOS 值, 50 位评估人员从图像的整体感官性和局部细节纹理对各个算法进行打分, 分值取值范围 1-5 分, 统计各个评估人员对图像打分的结果, 取其平均值作为 MOS 值的大小评价算法性能。

3.3 实验结果分析

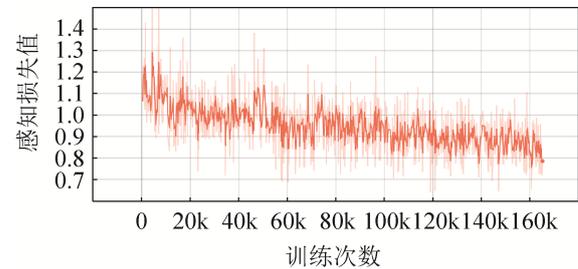
3.3.1 模型训练损失及分析

损失函数是衡量网络结构好坏的一种工具。

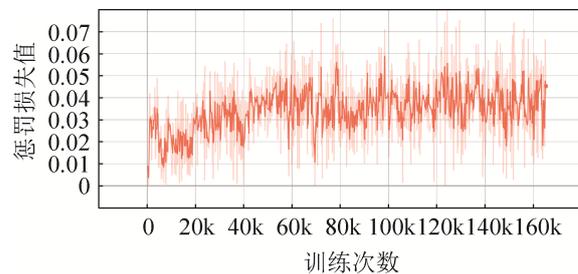
SESRGAN 算法在训练过程中, 通过不断更新计算每个参数来实现梯度优化, 训练了 160×10^3 次之后, 使参数达到最优值, 进而最小化损失函数。各种损失如图 9 所示。



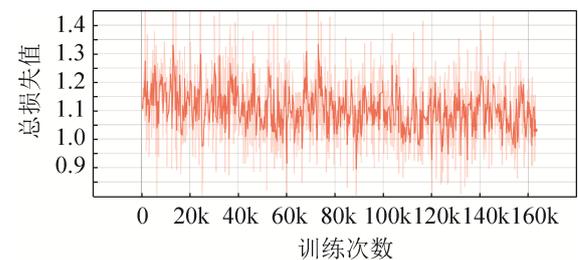
(a) 内容损失



(b) 感知损失



(c) 惩罚损失



(d) 总损失

图 9 网络训练中的 4 种不同损失图

Fig. 9 Four different losses during network training

从图 9 可以看到, 内容损失稳定在 $4 \times 10^{-4} \sim 5.5 \times 10^{-4}$ 之间, 这表明生成壁画的像素点充分逼近真实图像之间的像素点, 没有过多的像素误差; 感

知损失在 1.2~0.8 内稳定下降, 最终稳定在 0.9 附近, 表明生成壁画的特征提取后与真实壁画的特征提取后的误差较小, 证明该生成模型能够较好的提取图像特征; 对抗损失稳定在 0.03~0.04 范围内, 判别器损失周期性的从较高的损失出发然后下降, 是因为生成器训练时判别器锁定, 导致判别器的能力下降; 然后生成器锁定, 判别器训练不断优化使得损失下降, 表明判别器在不断优化最终训练成熟, 能够较好的判断是否为生成壁画或真实壁画, 总损失也随着训练过程的进行稳定收敛。

3.3.2 SESRGAN 算法重建壁画效果及分析

为了验证算法重建古代壁画的效果, 从重建好的放大 4 倍高分辨率壁画中随机选取 4 张与对应的低分辨率壁画进行对比观察, 如图 10 所示。



图 10 不同风格的壁画重建效果图

Fig. 10 Reconstruction renderings of different styles of ancient murals

从图 10 可以看出, 各幅低分辨率古代壁画经过本文算法重建后纹理细节恢复的效果较为理想。究其原因使用较深的网络提取高频的特征, 并且使用预训练好的 VGG 网络在激活前而不是在激活后提取特征信息。观察低分辨率图像, 其中保留了较多细节的低分辨率图像重建后较为清晰, 从亮度、对比度等都有较好的一致性; 另外从细节较少的低分辨率图像重建后可以看见图像恢复的较为完整, 细节方面也不逊色于其他重建后的壁画。重建后的图像的 PSNR 和 SSIM 如表 2 所示, 表 2 的结果也可以看出, 重建后的壁画也得到了良好的 PSNR、SSIM 值。从

以上几方面证明了本文算法有较好的稳定性。

表 2 重建壁画后的 SSIM 和 PSNR
Table 2 PSNR and SSIM of the reconstruction mural

图像	PSNR	SSIM
a	28.71	0.876
b	32.03	0.782
c	32.03	0.782
d	32.03	0.782

3.3.3 与经典算法对比实验及分析

为了保证实验的对比效果, 选取了 6 种不同风格类型的壁画各一张作为代表进行对比, 分别如图 11 所示。使用现有的 BI 算法、文献[4]、文献[5]、SRGAN 算法和 ESRGAN 算法进行了对比实验。为了更清晰的观察到对比效果, 在 6 幅壁画图片分别选取了局部重建放大的细节如图 12 所示。

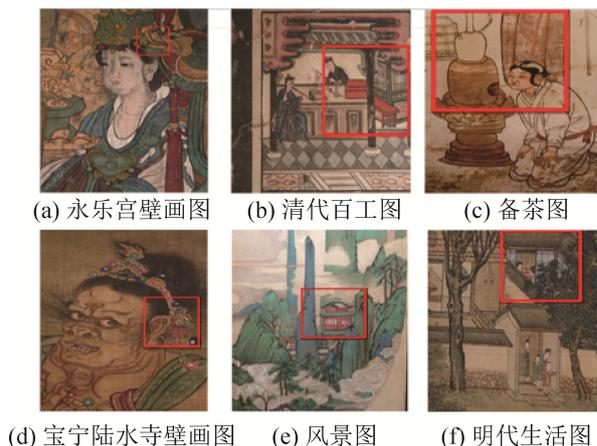


图 11 不同风格的壁画图

Fig. 11 Ancient murals in different styles

从图 12 中可以看出, 在永乐宫壁画图和清代百工图中各个算法恢复细节较好, 但是亮度方面 BI、SRGAN、文献[4]不如文献[5]、ESRGAN 和本文算法 SESRGAN; 在备茶图中, 观察壁画人像的脸部, SESRGAN 算法重建的图像较平滑并且噪音信息去除相比其他算法效果较好; 在明代保定寺水路图中, 观察壁画人像下面的配饰可看出, 其他算法纹理细节部分缺失, SESRGAN 保留了较多的纹理细节; 风景图和明代生活图中, 由于 BI 算法未考虑到整个图像的特征信息导致缺少高频细节; 文

献[4]采用传统基于稀疏信号的重建算法提取特征能力不足, 部分细节不够完善; 文献[5]虽然解决了边缘锯齿问题, 但是亮度不如 SESRGAN; SRGAN 算法采用相对比较浅的网络提取的特征信息不足, 所以细节纹理相比 ESRGAN 算法和文本算法较差, ESRGAN 恢复效果较好, 但是其引入了一些不愉快的噪音。综上所述, SESRGAN 算法的重建效果在细节方面均优于其他算法, BI 算法在各种风格的壁画中亮度和细节恢复方面都不如其他算法; 文献[4]对于去除壁画伪影效果较好, 但是重建

出壁画细节尚有不足; 文献[5]对于边缘锯齿问题处理较好, 但是亮度和细节方面不如 ESRGAN、SESRGAN; SRGAN 算法恢复了部分高频信息, 由于这两种算法采用的网络层数较少, 所以没有学习到一些高频的细节, 边缘的锐化严重; ESRGAN 算法恢复细节整体效果比前两种效果好, 但是增加了较多不愉快的伪影和噪声信息。相比之下本实验重建后的效果在纹理细节方面、整体亮度和清晰度都有了一定的提高。



图 12 不同算法下重建的壁画图

Fig. 12 Mural images reconstructed under different algorithms

50 名评估人员对图 12 中不同风格古代壁画的重建效果的主观打分如图 13 所示。在通过评估人员评估和讨论后, 一致认为本文算法的整体感官性相比较其他算法更加符合人类的视觉感受, 在亮度

和图像平滑度方面也有较好的效果。在局部纹理细节方面, 评估人员在对比了生成的壁画和现有的壁画之后, 认为该算法重建效果较为理想, 在低分辨率图像中恢复了较多准确的细节, 具有一定的研究

价值。通过比较在整体观感性和局部细节纹理上的评分也都均优于其他算法，因此提出的该算法在主观打分方面也有较好的说服力。

表 3 给出了不同算法在 6 种不同类型的局部壁画上的平均 PSNR、SSIM，在 PSNR 和 SSIM 指标中与其他算法比较下均有了一定的提高，PSNR 平均提高了 0.4~1.37 dB，SSIM 平均提高了 0.013~0.027，证明了 SESRGAN 算法能够对壁画进行预防性保护工作，也提升了壁画的观赏性和可研究性价值。

BI 算法、文献[4]、文献[5]缺少捕捉高频信息的能力，导致 PSNR 和 SSIM 值较低；由于本文生成网络的 RS-RDB 中去除了 BN 层并且采用较深的网络进行特征提取，因此相比 SRGAN 算法和

ESRGAN 算法而言，能够较好的捕捉高频信息、拥有较高的 PSNR 和 SSIM 值。

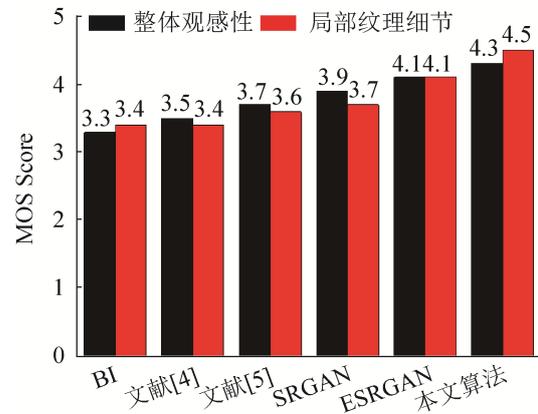


图 13 重建后壁画的主观评分图
Fig. 13 Subjective scoring of the reconstruction effect of ancient murals

表 3 不同算法下各个风格的壁画 PSNR、SSIM 值对比表

Table 3 Comparison of PSNR、SSIM values of various styles of mural images under different algorithms

图像	BI	文献[4]	文献[5]	SRGAN	ESRGAN	ESERGAN
a	30.58/0.827	30.78/0.824	30.97/0.831	31.24/0.843	32.06/0.862	32.53/0.859
b	28.01/0.792	28.24/0.799	28.16/0.795	28.39/0.806	28.60/0.809	29.20/0.824
c	22.69/0.741	23.96/0.743	24.84/0.745	25.17/0.749	26.20/0.751	27.03/0.776
d	25.48/0.739	26.54/0.740	27.29/0.751	26.59/0.743	28.42/0.746	28.06/0.764
e	24.41/0.764	25.16/0.766	25.75/0.771	25.81/0.767	26.54/0.774	27.29/0.786
f	23.59/0.731	24.88/0.730	25.43/0.733	25.66/0.734	26.28/0.738	26.39/0.743
Avg	25.79/0.765	26.59/0.767	27.07/0.771	27.14/0.773	28.01/0.779	28.41/0.792

3.3.4 BN 层和 SN 层对判别网络的影响分析

为了证明 BN 层能更好的训练网络，将含有 BN 层和含有 SN 层的判别网络分别进行训练，观察判别网络的 loss 值和 IS(inception score)值分别来评价两种归一化对判别网络的影响，IS 值是基于一个称为 Inception Net 的图片分类网络，它能对输入的图片进行分类，并得出其属于某一类的概率大小，如果图片生成的足够清晰，inception Net 分类的更准确，那么也就是说 IS 值也就越大；与此对应的生成器若能够生成多样的图像，那么生成的图像经过 inception Net 算法分类后的各类别分布应该是均匀的。BN 层和 SN 层损失训练对比图如图 14 所示，选取 3.3.3 节中 6 类不同风格的壁画作为测

评指标，得到的 IS 值如表 4 所示。

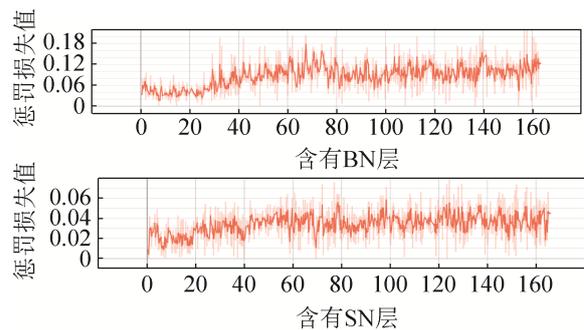


图 14 含有 BN 层和含有 SN 层的损失对比图
Fig. 14 Comparison of loss between BN layer and SN layer

从图 14 可以看出，使用了 BN 层的判别网络损失在 0.04~0.14 区间波动，而使用了 SN 层的判别网络稳定在 0.02~0.05 区间波动，壁画经过 BN

层其色彩分布都会被归一化, 会破坏图像原本的对比度信息, 影响了网络输出的质量, 而使用 SN 层则不会破坏, 并且有助于满足 Lipschitz 约束, 降低了训练的难度, 从而使判别网络的训练更加稳定, 效果更好。

表 4 含有 BN 层和含有 SN 层的 IS 值对比
Table 4 Comparison of IS between BN layer and SN layer

图像	含有 BN 层	含有 SN 层
a	7.31	9.54
b	6.47	8.25
c	7.06	7.03
d	5.36	6.49
e	5.45	7.24
f	6.17	8.49

从表 4 可以看出含有 BN 层判别网络训练后的 IS 得分普遍低于含有 SN 层判别网络训练后的 IS 得分, 说明含有 SN 层判别网络能够帮助整个网络训练稳定, 使生成网络生成的壁画更加清晰、真实。

3.3.5 激活前特征图与激活后特征图对比实验分析

随机选取一幅壁画对其激活前特征图和激活后特征图进行提取如图 15 所示, 从图中可以看出, 激活前提取到的壁画特征较清晰, 而激活后提取到的特征只能模糊看到人物的外形轮廓。因为随着网络层数不断的加深, 激活后的特征图监督效果越来越弱并且是已经稀疏化的特征, 稀疏的特征会导致网络性能变差, 而激活前的特征图保留了更详细的细节, 因此可以带来更强的监督。



图 15 激活前与激活后特征提取对比图

Fig. 15 Comparison of feature extraction before and after activation

为了证明激活前提取特征能为壁画提供更好的亮度, 选取激活前与激活后生成的壁画作对比如

图 16 所示。观察图 16 壁画中的窗户可以看出, 激活前提取特征的窗户清晰、明亮, 而激活后提取特征的窗户偏暗淡, 因此使用激活前计算感知损失能够为壁画一致性提供更好的监督。



图 16 激活前与激活后壁画对比图

Fig. 16 Comparison of mural before and after activation

4 结论

本文针对古代壁画模糊不清晰、分辨率低导致观赏性不高和研究价值不足的问题, 本文提出了 SESRGAN 算法对古代壁画进行重建, 以生成对抗网络框架, 在生成网络使用 RS-RDB 进行深层特征的提取, 有利于获取更多的特征信息, 然后引入 WAN-GP 理论改进对抗损失, 采用 Wasserstein 距离计算重建图像与真实图像之间的距离, 对判别网络做出梯度惩罚, 增强网络训练的稳定性。使用激活前的特征计算感知损失, 有助于恢复更多的纹理细节和更令人愉悦的亮度, 采用二次迁移的训练方法训练模型, 最终得到一个稳定增强的壁画重建模型。在与其他相关的超分辨率算法相比, 提出算法的 PSNR 值提高了 0.4~2.62 dB, SSIM 提高了 0.013~0.027。在主观评价方面整体感官性的 MOS 值为 4.1, 局部细节纹理的 MOS 值为 4.3, 相较于其他算法拥有更高的评分值。该模型在整体壁画和局部区域壁画都拥有较好的重建效果, 重建后高分辨率的壁画清晰、亮度较好, 保留了丰富的纹理细节, 并且壁画的观赏性和研究价值有了一定的提高, 也进一步保护了古代壁画的流失。

实验不足之处在于只在壁画放大 4 倍下进行超分辨率重建, 没有在多尺度下进行壁画的超分辨率重建工作, 未将全局特征信息和局部特征信息充

分利用,部分重建后的纹理细节恢复的不完整,并且模型训练周期长。下一步的主要工作有:①收集更多不同风格的壁画数据集以适应生成网络的深度,增加模型的泛化性和稳定性;②对壁画进行多尺度的超分辨率重建工作,展示出不同尺度下壁画的艺术性和研究价值;③精简网络结构以减少训练的时间,并将全局特征与局部特征融合训练以获取更多的细节纹理特征,恢复出质量高的壁画。

参考文献:

- [1] 师晓波. 毁损古墓葬壁画虚拟修复技术研究[D]. 太原: 中北大学, 2018.
Shi Xiaobo. Research on Virtual Repair Technology of Damaged Ancient Tomb Murals[D]. Taiyuan: North University of China, 2018.
- [2] Wang H. Inpainting of Potala Palace Murals Based on Sparse Representation[C]// International Conference on Biomedical Engineering and Informatics, USA: IEEE, 2015: 737-741
- [3] 曹建芳, 张自邦, 赵爱迪. 增强一致性生成对抗网络在壁画修复上的应用[J]. 计算机辅助设计与图形学学报, 2020, 32(8): 1315-1323.
Cao Jianfang, Zhang Zibang, Zhao Aidi. The Application of Enhanced Consistency Generative Adversarial Network in Mural Repair[J]. Journal of Computer Aided Design and Graphics, 2020, 32(8): 1315-1323.
- [4] 马强. 基于稀疏表示的单幅彩色图像超分辨率重建研究[D]. 兰州: 兰州理工大学, 2019.
Ma Qiang. Research on Super-resolution Reconstruction of a Single Color Image Based on Sparse Representation [D]. Lanzhou: Lanzhou University of Technology, 2019.
- [5] 徐志刚, 闫娟娟, 朱红蕾. 基于多尺度残差注意力网络的壁画图像超分辨率重建算法[J]. 激光与光电子学进展, 2020, 57(16): 152-159.
Xu Zhigang, Yan Juanjuan, Zhu Honglei. Super-resolution Reconstruction Algorithm of Mural Image Based on Multi-Scale Residual Attention Network[J]. Progress in Laser and Optoelectronics, 2020, 57(16): 152-159.
- [6] 张洋. 基于双线性插值法的图像缩放算法的设计与实现[J]. 电子设计工程, 2016, 24(3): 169-170, 175.
Zhang Yang. Design and Implementation of Image Scaling Algorithm Based on Bilinear Interpolation[J]. Electronic Design Engineering, 2016, 24(3): 169-170, 175.
- [7] 练秋生, 张伟. 基于图像块分类稀疏表示的超分辨率重构算法[J]. 电子学报, 2012, 40(5): 920-925.
Lian Qiusheng, Zhang Wei. Super-resolution Reconstruction Algorithm Based on Sparse Representation of Image Block Classification[J]. Electronic Journal, 2012, 40(5): 920-925.
- [8] 钟宝江, 陆志芳. 图像插值技术综述[J]. 数据采集与处理, 2016, 31(6): 1083-1096.
Zhong Baojiang, Lu Zhifang. Overview of Image Interpolation Technology[J]. Data Collection and Processing, 31(6): 1083-1096.
- [9] 宋佳伟, 徐煜明. 基于小波变换和迭代反向投影的超分辨率算法[J]. 计算机技术与发展, 2015, 25(2): 74-77.
Song Jiawei, Xu Yuming. Super-Resolution Algorithm Based on Wavelet Transform and Iterative Back Projection[J]. Computer Technology and Development, 2015, 25(2): 74-77.
- [10] 张砚, 徐昆. 基于外存和凸集投影法的遥感图像超分辨率方法[J]. 清华大学学报(自然科学版), 2010, 50(10): 1743-1746, 1756.
Zhang Yan, Xu Kun. Remote Sensing Image Super-Resolution Method Based on External Memory and Convex Set Projection Method[J]. Journal of Tsinghua University (Natural Science Edition), 2010, 50(10): 1743-1746, 1756.
- [11] Chang H, Yeung D Y, Xiong Y. Super-Resolution Through Neighbor Embedding[C]// IEEE Conference on Computer Vision and Pattern Recognition. USA: IEEE, 2004: 275-282.
- [12] Yang J, Wright J, Huang T S, et al. Image Super-Resolution via Sparse Representation[J]. IEEE Transactions on Image Processing (S1057-7149). 2010, 19(11): 2861-2873.
- [13] Dong C, Loy C. Image Super-Resolution Using Deep Convolutional Networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence (S0612-8828), 2016, 38(2): 295-307.
- [14] Dong C, Chen C L, Tang X. Accelerating the Superresolution Convolutional Neural Network[C]// 14th European Conference on Computer Vision. Berlin: Springer International Publishing, 2016: 391-407.
- [15] Tong T, Li G, Liu X, et al. Image Super-Resolution Using Dense Skip Connections[C]// 2017 IEEE International Conference on Computer Vision. New York: IEEE Computer Society, 2017: 4809-4817.
- [16] Ledig C, Theis L, Huszar F, et al. Photo-Realistic

- Single Image Super-Resolution Using a Generative Adversarial Network[C]// International Conference on Computer Vision and Pattern Recognition(CVPR). New York: IEEE Computer Society, 356-379.
- [17] Hui D L, Gu X F. Wasserstein GAN With Quadratic Transport Cost[C]// 2019 IEEE/CVF International Conference on Computer Vision. Seoul, ROK: IEEE, 2019: 4831-4840.
- [18] Wang X, Yu K, Wu S, et al. Esrgan: Enhanced Super-Resolution Generative Adversarial Networks [C]// European Conference on Computer Vision. Munich, Germany: ECCV, 2018: 63-79.
- [19] Jin Q M, Lin R H. Enhanced Generative Model of Signaling Data Based on WGAN-GP and ACGAN[J]. IEEE Systems Journal (S1932-8184), 2020, 14(3): 3289-3330.
- [20] 袁功霖, 侯静, 尹奎英. 基于迁移学习与图像增强的夜间航拍车辆识别方法[J]. 计算机辅助设计与图形学学报, 2019, 31(3): 467-473.
Yuan Gonglin, Hou Jing, Yin Kuiying. Night Aerial Vehicle Recognition Method Based on Transfer Learning and Image Enhancement[J]. Journal of Computer Aided Design and Graphics, 2019, 31(3): 467-473.