

基于改进深度卷积生成对抗网络的刺绣图像修复

刘羿漩¹, 葛广英^{2*}, 齐振岭¹, 李振轩¹, 孙福临¹¹聊城大学物理科学与信息工程学院山东省光通信科学与技术重点实验室, 山东 聊城 252059;²聊城大学计算机学院, 山东 聊城 252059

摘要 目前中华传统刺绣工艺传承保护问题中的修复任务主要以人工为主, 修复过程需要大量的人力、物力。随着深度学习的高速发展, 不同类型的刺绣文物损伤可以利用生成对抗网络进行修复。针对上述问题, 提出一种基于改进深度卷积生成对抗网络(DCGAN)的刺绣图像修复方法。首先, 在生成器部分引入空洞卷积层扩大感受野, 并添加卷积注意力机制模块, 在通道与空间 2 个维度增强重要特征的指导作用; 在判别器部分增加全连接层数提升网络解决非线性问题的能力; 在损失函数部分联合均方误差损失与对抗损失通过网络训练相互博弈的过程实现刺绣图像修复。实验结果表明: 引入空洞卷积层与注意力机制提升了网络性能与修复效果, 最终得到修复图像的结构相似性高达 0.955, 能够得到较为自然的刺绣图像修复效果, 可以为专家提供纹理、色彩等信息作为参考辅助后续的修复。

关键词 非遗文化保护; 刺绣图像修复; 生成对抗网络; 卷积神经网络; 空洞卷积; 注意力机制

中图分类号 TP391.4; TP183

文献标志码 A

DOI: 10.3788/LOP223060

Research on Embroidery Image Restoration Based on Improved Deep Convolutional Generative Adversarial Network

Liu Yixuan¹, Ge Guangying^{2*}, Qi Zhenling¹, Li Zhenxuan¹, Sun Fulin¹¹Shandong Provincial Key Laboratory of Optical Communication Science and Technology, School of Physical Sciences and Information Engineering, Liaocheng University, Liaocheng 252059, Shandong, China;²School of Computer Science and Technology, Liaocheng University, Liaocheng 252059, Shandong, China

Abstract Presently, image inpainting in the inheritance and protection of Chinese traditional embroidery often depend on human labor, with considerable work force and material resources. Furthermore, with the rapid development of deep learning, generative adversarial networks can be applied to repair damaged embroidery relics. An embroidery image restoration method based on improved deep convolutional generative adversarial network (DCGAN) is proposed to solve the above problems. In the generator part, dilated convolution is introduced to expand receptive fields; the addition of the convolution attention-mechanism module enhances the guiding role of significant features in two dimensions of channel and space. In the discriminator part, the number of full connection layers are increased to improve the ability of the network to solve nonlinear problems. In the loss function part, the mean square error loss and confrontation loss are combined to realize embroidery image inpainting through the game process of network training. The experimental results show that the dilated convolution and convolution attention mechanism module improves the network performance and repair effect, and the structural similarity of the repaired image is as high as 0.955. This method enables obtaining a more natural embroidery image-restoration effect, which can provide experts with information such as texture and color as a reference to assist subsequent repair.

Key words intangible cultural heritage protection; embroidery image inpainting; generative adversarial network; convolutional neural network; dilated convolution; attention mechanism

1 引言

刺绣作为我国重要的非物质文化遗产之一, 具有

极高的艺术价值, 中国各民族远古时代即“好五色”“重文绣”, 在上古神话和古代文献中早有记述。商周时期, 衣服的“文章”“絺绣”已不足为奇, 缝缀等工艺更是

收稿日期: 2022-11-15; 修回日期: 2022-12-08; 录用日期: 2023-01-06; 网络首发日期: 2023-02-07

基金项目: 中央引导地方科技发展专项(YDZX2017360000283)

通信作者: *ggysd@126.com

古人制衣成衣必不可少的环节^[1]。黄涵煦等^[2]从马王堆汉墓中出土的文物中得出结论:汉代厚葬文化盛行,造型艺术繁荣以及重视神仙思想等社会特点,并从风格与形式 2 个角度分析文物中具有的艺术价值。事实证明,这些文物经过千百年的历史被传承至今,见证了人类社会发展的每一步,具有特殊的价值和重要的意义。

但随着大量历史文物的出土,许多刺绣文物由于年代久远而存在着不同程度的损坏,例如现在位于大英博物馆中,由敦煌藏经洞出土的高近 2.5 m 的巨幅刺绣佛像,目前由于损毁,无法确定其中的佛像是释迦牟尼佛还是凉州佛。

目前文物修复领域主要以人工为主,刺绣文物的修复与其他文物的修复相比更为复杂,除去绣面的分析修复,其中的针法技巧分析、颜料染色复原以及丝线之间间接续缺一不可,通常这种修复需要极其漫长的过程。近年来,随着深度学习技术的发展,针对刺绣修复任务中绣面修复的问题,绣面上破损、脏污以及褪色等不同类型的文物损伤可以尝试利用数字修复技术进行修复,在保护刺绣文物的基础上最大程度恢复其部分原貌,为专家人工修复提供参考,便于进一步研究。

近年来,神经网络在计算机视觉领域大放异彩,其中以生成对抗网络(GAN)^[3]为代表的生成对抗类网络模型,如条件生成对抗网络(CGAN)^[4]、深度卷积生成对抗网络(DCGAN)^[5]、Stack-GAN^[6]等,这些生成对抗类网络模型以半监督的方式训练模型,能够有效提升合成内容的灵活性。针对图像智能修复这一问题,目前基于深度卷积神经网络(DCN)的图像修复方法成为该领域的主流算法,2015 年 Rayford 等将 DCN 与 GAN 相结合提出了 DCGAN 网络结构,利用卷积网络的特征提取能力;2016 年 Pathak 等^[7]提出了一种无监督学习式的网络架构 Context Encoder 算法,基于缺损图像的上下文特征,利用编解码器生成完整图像,引入重建损失与对抗损失两部分提升图像修复效果;2017 年 Iizuka 等^[8]提出了全局与局部结合的图像修复(GL)方法,在全局判别器的基础上引入了局部判别器,双判别器的结构能够从整体和局部 2 个角度判断图像合理性,使图像修复效果得到提升;2019 年 Yu 等^[9]基于 two-stage 网络架构(GIICA)提出 GConv 修复算法,将其中卷积操作替换为门控卷积,对图像中

的有效特征进行选择,生成具有更好修复效果的图像^[10];Nazeri 等^[11]提出了 EdgeConnection(EC)网络模型实现图像修复,EC 网络先修复缺损部分的边缘信息得到完整边缘图,再进一步得到完整的修复图像;刘强等^[12]提出了结合 Squeeze-and-excitation network (SENet)^[13]的密集卷积生成对抗网络的图像修复方法,在 CelebA 数据集上取得了不错的效果。

生成对抗类网络模型是目前图像合成及修复领域最令人瞩目的研究方向之一^[14-18]。针对常见的街景以及人脸修复等问题,目前基于生成式对抗网络的图像修复算法取得的修复效果已达到较为自然清晰的效果,但针对刺绣图像展开的修复更为复杂,刺绣图像种类繁多,特征复杂且难以提取,需要注重局部区域特征,更要注重整体语义信息。因此,本文提出了一种基于改进 DCGAN 的刺绣图像修复算法,生成器部分利用神经网络优秀的特征提取能力提取特征,针对刺绣图像全局对特征进行提取,引入空洞卷积操作扩大感受野,获取更多的特征信息,充分理解图像整体语义并获取图像更多细节信息;针对修复过程中的局部重要特征提取,与注意力机制相结合,引入卷积注意力模块(CBAM)来关注重要特征所在的区域,优化 GAN 的生成器部分,使修复得到的刺绣图像既能够还原出破损图像的整体结构,又能强调其中重要的特征细节,对判别器也进行了改进,在原始判别器后增加了全连接层用于更好地解决非线性问题,提高了判别器的性能。实验表明所提的算法在刺绣图像修复问题上有非常好的效果,且改进后的模型与改进前的模型相比具有更优越的性能。

2 基本原理

2.1 DCGAN 算法

Goodfellow 等从博弈论中受到启发提出 GAN,该网络的整体结构如图 1 所示。整个结构由生成器 G 和判别器 D 两部分构成。其中,生成器 G 提取输入样本中的特征,获取样本的数据分布,并输出生成样本;将真实样本与生成样本作为一组数据,成对输入到判别器 D 中,经判别器 D 输出后,得到判断输入样本的判别结果,将判别结果反馈给生成器,不断地进行训练和更新网络。输出结果的概率越接近 0.5,得到的图像越

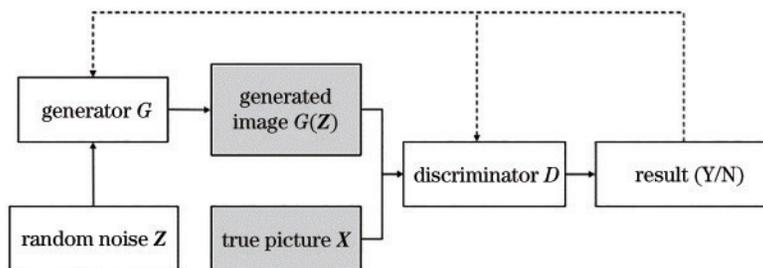


图 1 生成对抗网络

Fig. 1 Generative adversarial network

来越接近真实图像。

DCGAN 将 DCN 与 GAN 相结合,相对于经典的 GAN,DCGAN 在特征提取方式部分进行了优化,有效提高了网络训练的稳定性与生成图像的质量^[19]。在 DCN 结构上采取的优化方式:首先,去掉网络中所有的全连接层,去掉全连接层后的网络为全卷积网络;其次,去掉所有的池化层,在生成器 G 中使用微步幅卷积代替池化层,在判别器 D 中使用步幅卷积代

替池化层;然后生成器 G 和判别器 D 中使用了批量归一化 (BN), DCN 引入 BN 后加速了网络收敛, DCGAN 引入 BN 可以防止生成器将所有样本压缩到一个点,有助于处理不良的初始化而产生的训练问题;最后生成器 G 中使用 ReLU 作为激活函数,最后一层使用 tanh 作为激活函数,而判别器 D 中使用 Leaky ReLU 作为激活函数。DCGAN 网络整体架构如图 2 所示。

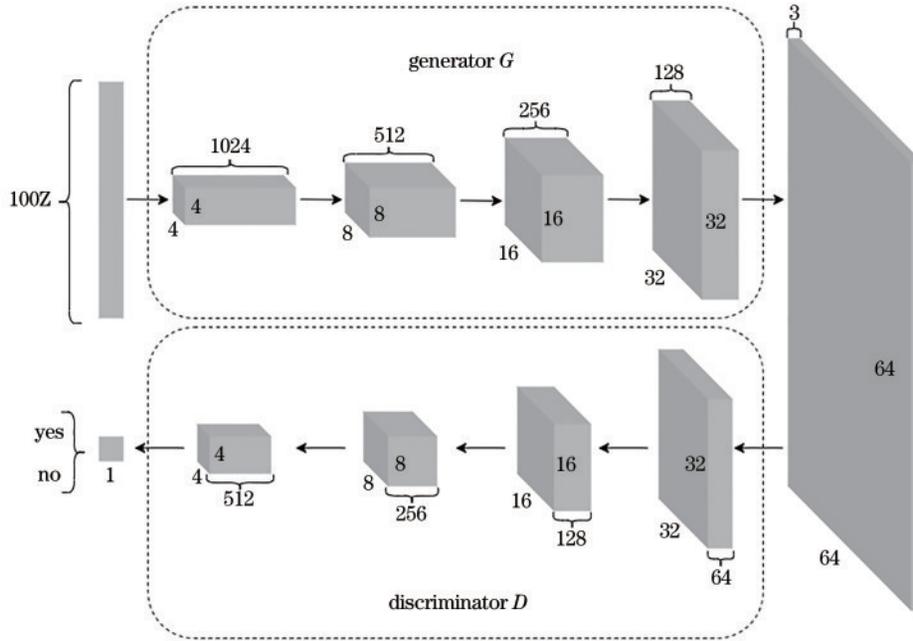


图 2 DCGAN 网络结构

Fig. 2 Network structure of DCGAN

在图像修复问题中,最终生成的缺损部分图像是根据真实图像信息得来的,本质上是求得缺损图像到修复图像之间特征的映射关系^[19]。DCGAN 在求解这种特征映射关系方面表现突出,与 GAN 相比提取特征的能力更强。

2.2 卷积块注意力模块 (CBAM)

注意力机制模拟人类大脑信号处理机制。在快速扫描图像全局后获得需要重点关注的目标区域,然后对该区域投入更多注意力,其本质是一种通过网络自主学习,对不同权重系数进行加权,用来强调感兴趣的区域,抑制不相关区域的机制。

CBAM 是一种结合了通道和空间的注意力机制模块,通过对这两部分特征进行“动态加权”,以达到增强目标特征表达的目的^[20]。CBAM 包括 2 个独立的子模块:分别是通道注意力模块 (CAM) 和空间注意力模块 (SAM)^[21], CAM 的输出是 SAM 的输入特征。CBAM 的整体结构如图 3(a) 所示, CAM、SAM 的结构如图 3(b)、图 3(c) 所示。

将经过 CBAM 得到的特征图与原特征图融合后

作为下一层的输入特征图,用来强调重要特征。

2.3 空洞卷积

空洞卷积首次应用于深度学习中是在图像分割领域^[22],如图 4 所示。

图 4 的空洞卷积操作卷积核大小为 3×3 , 空洞卷积的空洞率分别为 1、2、4。空洞卷积感受野的计算公式如下:

$$f_k = f_{k-1} + (n - 1) \times \prod_{i=1}^{k-1} S_i, \quad (1)$$

式中: f_k 为当前层感受野; f_{k-1} 为上一层感受野; n 为卷积核大小; S_i 为第 i 层卷积步长。空洞卷积通过在滤波器中间填充 0 来扩大核间隔, 设空洞卷积按空洞率 r 进行膨胀, 则膨胀后空洞卷积核大小计算方式如下:

$$m = n + (n - 1) \times (r - 1), \quad (2)$$

式中: m 为当前层卷积核膨胀后的大小; n 为当前层卷积核膨胀前的大小; r 为空洞率, 当 $r = 1$ 时, 空洞卷积即标准卷积。空洞卷积操作虽扩大了感受野, 但并未增加新的参数和计算量。

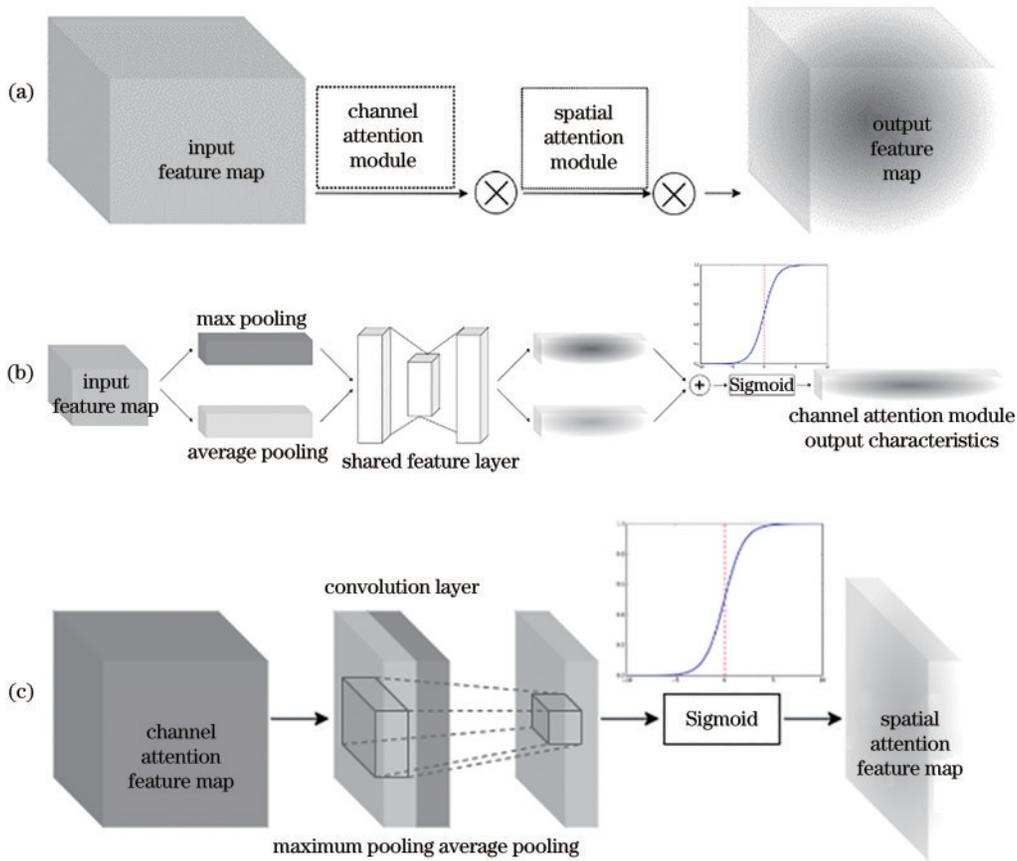


图 3 CBAM 结构。(a) CBAM 模块整体结构; (b) CBAM 中的 CAM 结构; (c) CBAM 中的 SAM 结构
 Fig. 3 CBAM structure. (a) Overall structure of CBAM; (b) CAM structure in CBAM; (c) SAM structure in CBAM

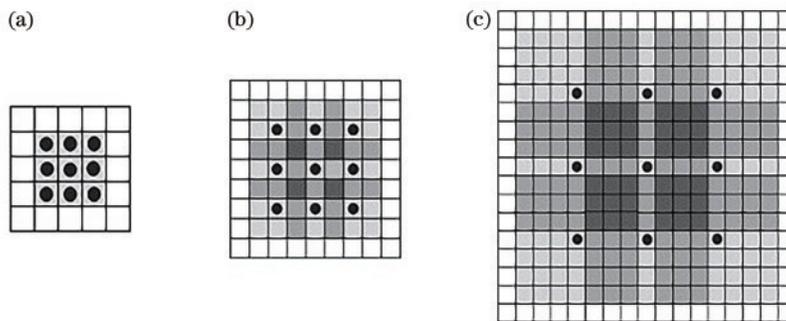


图 4 空洞卷积示意图。(a) 空洞率为 1; (b) 空洞率为 2; (c) 空洞率为 4
 Fig. 4 Dilated convolution diagram. (a) Void rate is 1; (b) void rate is 2; (c) void rate is 4

3 改进 DCGAN 的图像修复模型

提出的基于改进 DCGAN 的刺绣图像修复模型, 整体结构如图 5 所示。在训练过程中, 模型的网络参数通过不断对抗博弈、更新优化, 当整个模型达到稳定后, 截取生成图像的缺失区域与原图像合并成为修复图像, 得到修复结果。

3.1 生成器改进

生成器 G 结合 DCN 对样本进行特征提取, 输入为破损图像, 输出为修复图像。改进后的 DCGAN 生成器由卷积块、空洞卷积块与 CBAM 构成, 在空洞卷积块中引入 4 个空洞率分别为 2、4、8、16 的空洞卷积层,

扩大了感受野, 获取刺绣图像更多的特征信息。若只采用单一且空洞率较小的空洞卷积, 会造成卷积核不连续, 进而破坏损失信息连续性, 因此设置连续的空洞卷积层时不能采用相同的空洞率。根据文献[23]空洞率设计建议, 一般设计连续空洞卷积时, 第一个膨胀率一般是从 1 开始, 文献[23]还给出建议: 多个空洞率的公约数不能大于 1。设定送入网络的图像大小为 $256 \text{ pixel} \times 256 \text{ pixel}$, 因此设计 4 个分别为 2、4、8 和 16 的空洞率, 构成连续空洞卷积层, 当空洞率为 1 时空洞卷积即为标准卷积, 因此生成器的 Conv 6、Conv 7、Conv 8、Conv 9 和 Conv 10 可看作空洞率分别为 1、2、4、8 和 16 的连续空洞卷积层, 空洞率公约数为 1, 符合

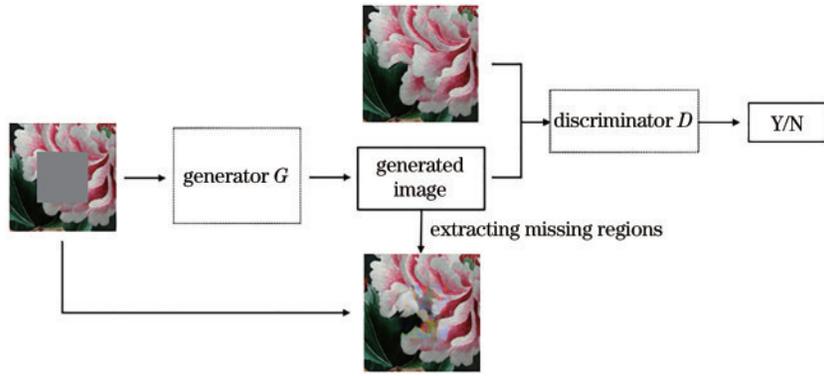


图 5 改进 DCGAN 的图像修复模型

Fig. 5 Improved DCGAN image inpainting model

建议。

此外,将模型与注意力机制相结合,在首个卷积层与最后一个卷积层后引入 CBAM,对通道和空间两个维度进行“动态加权”,从通道和空间两个维度来增

大重要特征的权值,使生成器生成的图像更符合人类视觉聚焦式特点,即选择性关注重要特征,达到优化生成对抗网络的目的。改进后的生成器 G 模型结构,如图 6 所示。

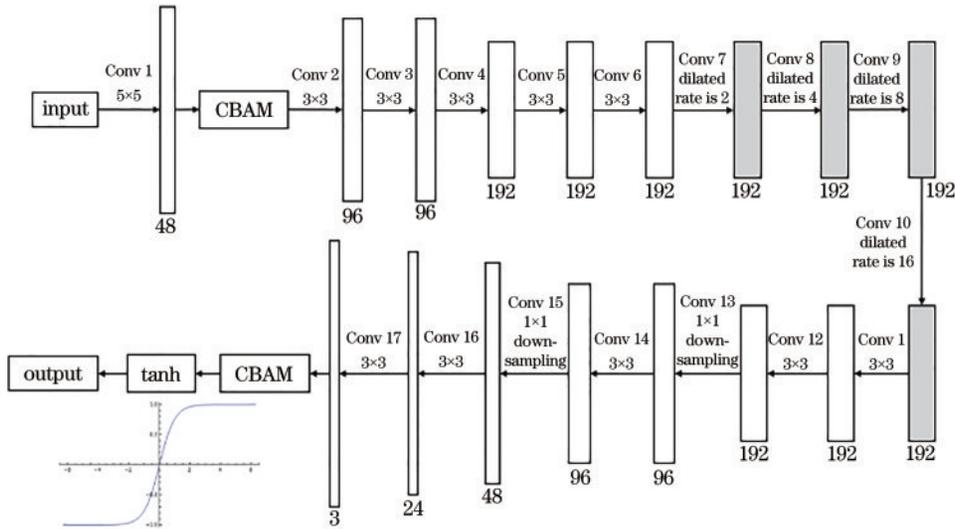


图 6 改进 DCGAN 的生成器模型结构

Fig. 6 Improved generator model structure of DCGAN

3.2 判别器改进

判别器 D 在原判别器后新增 2 个全连接层,共 3 个全连接层相互连接,更好地解决非线性问题来提高判

别器的性能^[24]。判别器最终进行一个二分类判别,改进后判别器 D 的模型结构,如图 7 所示。

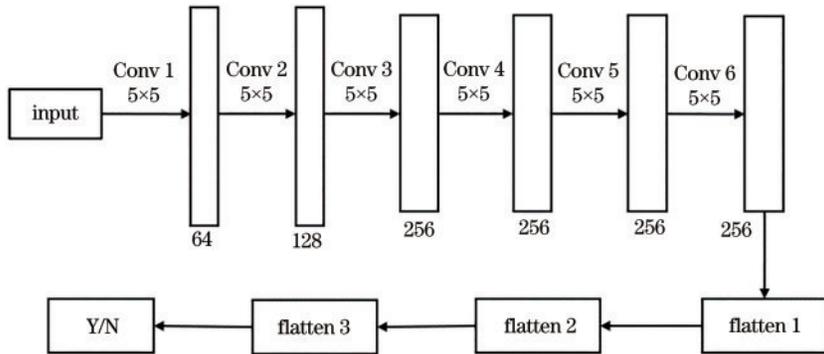


图 7 改进 DCGAN 的判别器模型结构

Fig. 7 Improved discriminator model structure of DCGAN

3.3 损失函数设计

损失函数的设计通常能够决定整体网络的最终性能,采用文献[25]的损失函数进行设计。

在训练过程中,将均方误差(MSE)损失与对抗损失两部分联合起来共同作为网络的总损失,增加迭代次数以提升图像的修复效果。

MSE 损失又称二次损失、L2 损失函数,表示预测值与目标值之间差值的平方和,在训练过程中,根据 MSE 损失不断迭代来更新网络参数,使生成器 G 可以生成与原图具有相似结构的图像,MSE 损失作为损失函数中最重要的组成部分,其反映预测值与生成值之间的差距,即生成图像与真实图像的差距,可以用来衡量模型的优劣。

设有 n 个训练数据 x_i , 每个训练数据 x_i 的真实输出为 y_i , 模型对训练数据 x_i 的预测值为 g_i , 生成网络运算表示为 $G(x_i)$, 则模型在 n 个训练数据下产生的 MSE 损失 L_{MSE} 为

$$\begin{cases} g_i = G(x_i) \\ L_{\text{MSE}} = \frac{1}{n} \sum_{i=1}^n (y_i - g_i)^2 \end{cases} \quad (3)$$

对抗损失用于加强生成网络和判别网络的博弈过程,提高生成图像的质量,使得结果更接近真实图像。 $D(x)$ 表示对真实样本进行判别,其值越接近 1 越好,而对生成图像而言,判别器的判别结果 $D[G(z)]$ 越接近 0 越好^[25]。对抗损失 L_{adv} 为

$$L_{\text{adv}} = \min_G \max_D E_{x \sim P_{\text{data}(x)}} [\text{lb} D(x)] + E_{z \sim P_{\text{data}(z)}} \left\{ \text{lb} \left[1 - D[G(z)] \right] \right\} \quad (4)$$

总损失函数 L_{Loss} 为

$$L_{\text{Loss}} = \lambda_{\text{MSE}} L_{\text{MSE}} + \lambda_{\text{adv}} L_{\text{adv}}, \quad (5)$$

式中: λ_{MSE} 、 λ_{adv} 分别为 MSE 损失与对抗损失对应的损失项权重。通过生成器和判别器的交替迭代优化,使判别器的判别结果达到最优。

生成器与判别器的 L_{MSE} 、 L_{adv} 的部分数值,如表 1 所示。

表 1 L_{MSE} 和 L_{adv} 的部分数值
Table 1 Partial values of L_{MSE} and L_{adv}

Epoch / 10 ⁴	L_{MSE}	$L_{\text{adv}}(G)$	$L_{\text{adv}}(D)$
0	1.506	0.793	1.227
10	2.063	0.439	0.490
20	2.822	0.372	0.322
30	2.278	0.273	0.286
40	2.165	0.157	0.329

L_{MSE} 与 L_{adv} 都是模型优化的途径,由表 1 可知, L_{MSE} 呈先增大后减小的趋势,体现了生成器 G 的生成样本与原始样本的区别,但 GAN 类网络训练过程不易收敛,容易产生巨大的梯度不稳定现象,因此在前

20 万次迭代中出现 L_{MSE} 振荡现象,取部分 L_{MSE} 呈现先增大后减小的趋势,最终收敛稳定于 2.0 附近,表示生成样本与原始样本较为相似,此时修复图像与原图像的峰值信噪比(PSNR)非常大,也进一步说明修复效果较好。而表 1 中 L_{adv} 生成器部分逐渐减小的同时,判别器部分也在不断减小,但会出现小幅震荡现象,生成器部分与判别器部分的 L_{adv} 下降也意味着生成器生成的图像与原样本越来越相似,但基于生成网络的博弈特性,损失值处于一种波动状态,因此判别器部分的 L_{adv} 选取部分数据后呈现出先减小后又增大的趋势,由表 1 可知,最后对抗总损失 L_{adv} 收敛于 0.5 附近,表明此时生成器与判别器达到平衡。

4 实验结果及分析

4.1 实验环境与评价指标

实验平台硬件环境的中央处理器(CPU)为 Intel (R) Xeon(R) Silver 4210R CPU @2.40 GHz,显存为 8 G,图形处理器(GPU)为 NVIDIA Quadro RTX 4000。软件环境使用 Win10 操作系统,PyCharm 版本为 2021,PyThon 版本为 3.7,TensorFlow 深度学习框架版本为 1.14 与图像处理库为 OpenCV4.5。

刺绣图像数据集来源于网络,筛选出 8 类中华传统刺绣,包括中国四大名绣:苏绣、湘绣、粤绣、蜀绣以及部分少数民族刺绣:哈密刺绣、水族马尾绣、西秦刺绣和其他类别,其中其他类别内的刺绣还包含京绣、鲁绣、盘金绣以及十字绣等。刺绣图像数据集共 763 张,训练集与测试集的比例约为 9:1,即训练集共 687 张刺绣图像,测试集共 76 张图像。训练集中的图像作为样本图像送入网络进行训练,送入网络的图像大小统一为 256 pixel × 256 pixel,训练完成后得到刺绣修复模型。在刺绣图像数据的测试集上分别添加规则矩形区域掩模和随机不规则区域掩模进行测试,实验结果评价指标取 76 张测试图像的平均值。

实验选取的评价指标使用 PSNR 和结构相似性(SSIM)^[26]以及模型复杂度(包括参数量与浮点运算计算量)。PSNR 是峰值信号的能量与噪声的平均能量之比,其定义为

$$P_{\text{PSNR}} = 10 \lg \left(M_{\text{Maxvalue}}^2 / L_{\text{MSE}} \right), \quad (6)$$

式中: M_{Maxvalue} 为存储的最大位数 $2^b - 1$, b 表示位数; L_{MSE} 为均方误差; P_{PSNR} 的值越高,说明图像质量越好。SSIM 评价方法是在 2001 年被提出用于衡量电视、电影或其他数字图像、视频的主观感受质量的一种方法,SSIM 算法在设计上考虑人眼视觉特性,比传统方式更符合人眼视觉感知^[27]。从自然图像高度结构化特征出发,通过亮度、对比度、结构 3 个方面估计感知结构信息的变化。

SSIM 基于不同窗口进行计算,设窗口 x 、 y 的大小为 $N \times N$ (通常取 8×8)。SSIM 定义为

$$c_1 = (k_1 L)^2$$

$$c_2 = (k_2 L)^2$$

$$SSIM(x, y) = \frac{2(\mu_x \mu_y + c_1)(2\sigma_{xy} + c_2)}{(\sigma_x^2 + \sigma_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)}, \quad (7)$$

式中: μ_x 为 x 的平均值; μ_y 为 y 的平均值; σ_x^2 为 x 的方差; σ_y^2 为 y 的方差; σ_{xy} 为 x 和 y 的协方差; L 为像素的动态范围; k_1, k_2 为默认值 ($k_1 = 0.01, k_2 = 0.03$); c_1, c_2 为大于 0 的常数值 (防止式中出现除 0 的异常)。SSIM 数值范围通常为 0~1 之间, 其数值越大, 表明修复后图像与原图像越相似, 修复质量越好。

对比模型复杂度评价指标包括参数量与浮点运算计算量。参数量为模型中所有带参数的层的权重参数总量, 参数量大小决定模型训练占用设备显存的量, 可用于形容模型大小, 卷积层参数量 P_{Params} 为

$$P_{Params} = \sum_{l=1}^D K_l^2 \cdot C_{l-1} \cdot C_l + \sum_{l=1}^D M_l^2 \cdot C_l, \quad (8)$$

式中: K_l 为第 l 层卷积核的大小; C_l 为输入第 l 层的特征图通道数; M_l 为第 l 层特征图的大小。最后得出共计 D 层网络的参数量 P_{Params} 。

浮点运算量为实际网络训练过程中加减乘除的计算次数, 其中卷积操作所占比重最大, 结果与输入图像大小也有关系, 可用于形容模型的复杂程度, 卷积层浮点运算量 F_{flops} 为

$$F_{flops} = \sum_{l=1}^D M_l^2 \cdot K_l^2 \cdot C_{l-1} \cdot C_l. \quad (9)$$

4.2 实验结果分析

本次实验刺绣图像数据集中, 样本图像只有 700 多幅, 训练样本较少, 所以在实验过程中, 训练轮次设

为 100, 批大小设置为 4, 每一个训练轮次迭代训练 4000 次, 训练过程参考文献[28]的生成器与判别器交替训练 GAN 算法的训练过程。训练过程采用了 Adam 优化算法, 优化器权重 $\beta_1 = 0.5, \beta_2 = 0.9$, 将网络学习率设为 0.0002, 学习率更新策略选用指数衰减, 衰减底数 Gamma 设为 0.99, 损失函数权重 λ_{MSE} 设为 0.9, λ_{adv} 设为 0.1, 模型的损失函数权重参数对初始值设定具有一定的依赖性, 若改变 λ_{MSE} 与 λ_{adv} 会导致精度波动。

4.2.1 不同算法实验结果对比分析

使用 CE、GL、GConv、EdgeConnect、DCGAN 以及本文方法对具有规则矩形遮挡区域面积为 64×64 的刺绣进行了修复效果对比。这些模型分别为:

1) CE: Context Encoder 算法主要利用缺损区域周围信息与原图像全局信息, 生成缺损部分的图像信息, 依赖上下文特征生成与之匹配的信息。

2) GL: Iizuka 等于 2017 年, 在 DCGAN 基础上对判别器进行优化并提出 GL 算法, 在原判别器部分引入全局判别器判别整个图像语义信息, 并引入局部判别器判别修复区域语义信息。

3) GConv: Yu 等于 2018 年, 在 Two-stages 网络架构(GIICA)基础上用门控卷积代替其中的卷积操作, 提升图像修复效果。

4) EC: Nazeri 等于 2019 年提出 EdgeConnect 算法, 选用高斯滤波器标准差为 2 的 Canny 边缘检测得到刺绣图像的二值边缘图像, 在第一阶段先针对缺损区域修复图像边缘信息, 然后再根据第一阶段修复得到的完整边缘图与破损图像送入网络生成最终修复图像, 其修复过程如图 8 所示。



图 8 Edgeconnect 算法修复过程

Fig. 8 Repair process of Edgeconnect algorithm

5) DCGAN: 将 DCN 与 GAN 相结合进行图像修复。

不同算法在刺绣图像测试集上的修复效果, 如图 9 所示。

由图 9 可知, 所提的修复算法较其他算法具有更好的效果, 第 3 行采取 CE 算法修复的刺绣图像只能大致还原出图像的色彩信息, 不能还原图像细节, 上下文语义并不连贯, 其余算法都能够大致恢复出图像的整体结构, 但本文方法对图像中的细节恢复效果更好, 例如有本文方法还原出了第 3 列的粉色花朵上叶片的

叶脉, 而且修复效果也还原出上面叶片比下面叶片的颜色较深这一色彩特点, 说明本文方法对破损区域的纹理特征与颜色信息都能进行较好的恢复, 能够大致还原图像的原貌。实验的客观评价选取 PSNR、SSIM 以及模型复杂度 (包括参数量以及浮点运算计算量), 如表 2 所示。

由表 2 可知, 与其他方法相比, 本文方法的 PSNR、SSIM 都是最高的。

从模型复杂度的角度分析, GConv 与 EdgeConnect 算法的参数量都达到百兆以上, 浮点运算量在对比算

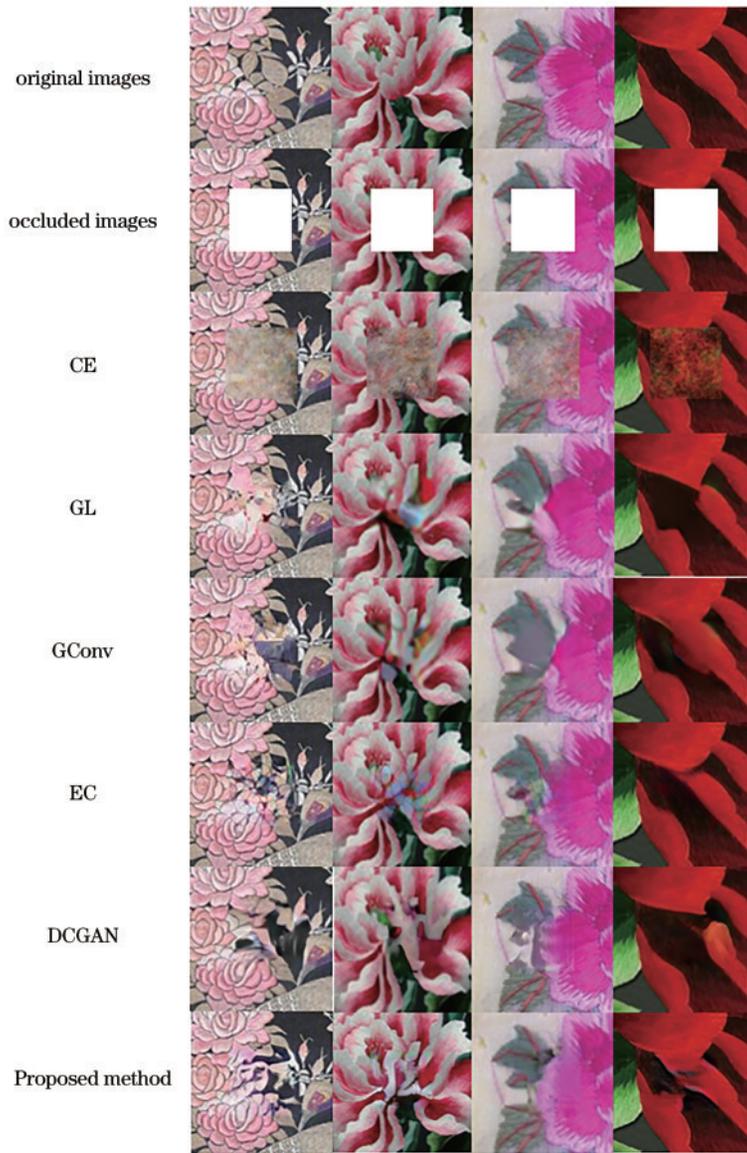


图 9 各方法修复效果对比图

Fig. 9 Comparison chart of repair effect of each method

表 2 不同算法修复规则矩形缺损区域的客观评价分数

Table 2 Objective evaluation scores of different algorithms for repairing regular rectangular defect areas

Algorithm	PSNR / dB	SSIM	Parameter quantity / MB	FLOPs / GB
CE	19.3027	0.7187	68.35	16.69
GL	19.8657	0.7436	66.07	53.11
GConv	21.0095	0.7806	100.33	63.59
EdgeConnect	20.8723	0.7859	103.75	91.03
DCGAN	20.2844	0.6883	88.67	28.74
Proposed method	21.3007	0.7863	96.55	34.40

法中也较大,在训练及测试时可能会增大内存占用或延长程序初始化的时间,因此在 EdgeConnect 与本文方法修复图像的 PSNR 与 SSIM 相差不大的情况下,优先考虑参数量以及浮点运算计算量较少的算法。

4.2.2 CBAM 与空洞卷积操作的影响

为验证空洞卷积层与 CBAM 对模型的影响,进一步在自建刺绣图像数据集上进行对比实验,对比未引入空洞卷积层与 CBAM、引入空洞卷积层未引入 CBAM、引入空洞卷积层与 CBAM 这 3 种算法在原始图像具有随机遮挡情况下的修复效果,随机遮挡面积占图像总面积的 10%,结果如图 10 所示。

由图 10 可知,第 4 列为未使用空洞卷积层和 CBAM 的模型结构的修复结果,对较细线条式的遮挡有较好的修复效果,但缺损面积较宽、较大时,无法保持良好的上下文一致性;第 5 列为引入 4 层空洞卷积层但未使用 CBAM 结构的修复结果,可以观察到修复出来的图像能够表达出大部分的图像语义信息,但仍有色彩不均的问题;第 6 列为本文方法的修复结果,CBAM 对 DCN 提取到的浅层特征,从通道与空间 2 个维度上赋予重要的特征较大的权重,加强重要特征的关注意

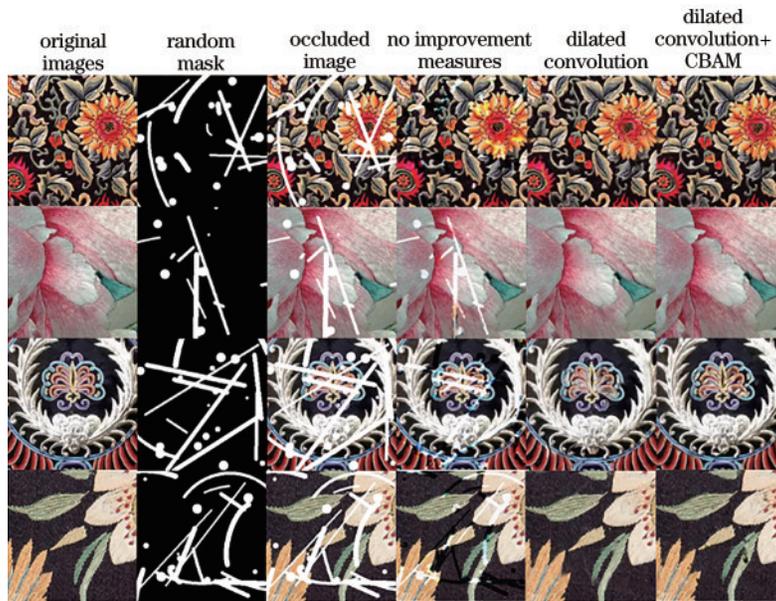


图 10 算法改进前后修复效果对比图

Fig. 10 Comparison chart of repair effect before and after algorithm improvement

度,使得到的修复图像能够更好地恢复出原始图像的信息,基本解决色彩不均等问题。实验结果证明:引入

空洞卷积层与CBAM对图像修复产生正面作用,不同情况下PSNR与SSIM的对比情况,如表3所示。

表 3 算法改进前后修复随机缺损区域的客观评价分数

Table 3 Objective evaluation score of random defect area before and after algorithm improvement

Algorithm	PSNR /dB	SSIM	Parameter quantity /MB	FLOPs /GB
No improvement measures	20.55680	0.87229	96.55	34.40
Dilated convolution	22.70757	0.92859	98.23	52.27
Dilated convolution+CBMA	24.18276	0.95521	98.75	52.89

由表3可知,虽然模型的参数量以及浮点运算计算量在增加,但PSNR与SSIM也有所提升,引入空洞卷积与CBAM之后,以2.28%的参数量提升来换取17.64%的PSNR提升与9.51%的SSIM提升,进一步证明引入CBAM可以对修复中的特征提取起到指导作用,引入空洞卷积层与CBAM可以加强生成器的特征提取能力,提高模型的性能并提升图像修复效果。

为了更加直观地说明引入注意力机制能使网络模

型选择性地关注重要特征,选择基于上述3种不同算法得到的生成图像进行热力图对比实验,根据生成图像在空间上的特征权重大小绘制出了热力矩阵,并将热力矩阵与原始输入图像叠加,这样便于直观地观察特征在空间上的重要程度。本次实验选取了4张刺绣图像作为样本,选择了2张颜色偏深与2张颜色偏浅的刺绣图像,并且这4张刺绣图像分别属于不同类别,得到的热力矩阵及叠加后的热力图,如图11所示。

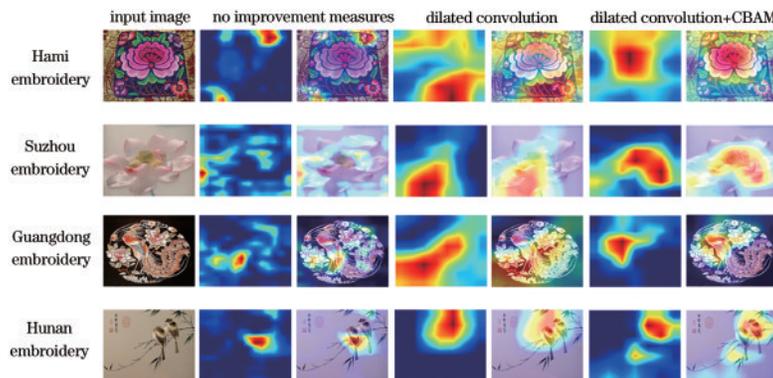


图 11 算法改进前后的热力图矩阵及热力图

Fig. 11 heat map matrix and heat map before and after algorithm improvement

图 11 中第 1 列为原图像,第 2、4、6 列为选用不同算法后根据在空间上的特征权重大小绘制出的热力矩阵,第 3、5、7 列为选用不同算法后得到的热力矩阵与原图叠加后生成的热力图。由图 11 可知,改进前空间上的重要特征作用并不明显,在引入了空洞卷积后,重要特征可以很好地凸显出来,而在引入注意力机制后,重要特征更为突出。以第 3 行粤绣刺绣图像为例,关注的特征应为图像中间圆形刺绣部分,在未采取改进措施时,特征的权重差别没有十分明显;在引入空洞卷积后,重要特征几乎落在了整个图像的中心;在引入 CBAM 后,重要特征的权值增大且重要特征的作用部分更为集中,较大权值的特征集中于刺绣图像中凤凰的头部部分,进一步说明注意力模块可以选择性关注重要特征。

5 结 论

针对非物质文化遗产保护领域中的刺绣修复问题,提出了一个基于改进 DCGAN 的刺绣图像修复算法。利用空洞卷积层扩大感受野,提升生成器的特征提取能力,引入 CBAM 来加强重要特征的指导作用,并在判别器后新增全连接层,用于提升网络解决非线性问题的能力,损失函数部分联合 MSE 损失与对抗损失构建网络模型,提高生成图像的质量。从主观效果上来看,本文算法对比现阶段其他算法在刺绣图像数据集上的修复效果更好,可以恢复出清晰的结构纹理与自然的颜色信息,从客观评价指标来看,本文算法具有较高的 PSNR、SSIM(PSNR 最高达 24.183、SSIM 最高达 0.955),进一步表明本文算法的有效性。

刺绣的表现形式精巧细致且多为立体,在对刺绣文物的修复过程中必要时需对针脚以及绣法进行还原复现,因此下一步应寻找相应的解决方法,如引入 3D 相机拍摄数据集进行训练,进一步提升修复效果,提升图像的分辨率;此外,中华刺绣历史源远流长种类繁多,同一种类的刺绣可能具有很多相似的特点,例如水族马尾绣底色多为黑色,刺绣丝线多以墨绿色与紫色为主,蜀绣则多以软缎为主要原料,色彩多为淡雅,再与刺绣分类应用相结合,针对不同种类进行对应类别的修复,可以提升刺绣图像的修复效果。

参 考 文 献

- [1] 邓启耀. 中国传统刺绣工艺的多重文化遗产价值[J]. 重庆三峡学院学报, 2018, 34(1): 39-49.
Deng Q Y. The multi-culture heritage value of Chinese traditional embroidery craftsmanship[J]. Journal of Chongqing Three Gorges University, 2018, 34(1): 39-49.
- [2] 黄涵煦, 王志亮. 汉代马王堆 T 形帛画的艺术价值与美学意义[J]. 河北科技大学学报(社会科学版), 2022, 22(2): 59-63.
Huang H X, Wang Z L. Artistic value and aesthetic significance of the T-shaped silks painting of Mawangdui tombs of Han dynasty[J]. Journal of Hebei University of Science and Technology (Social Sciences), 2022, 22(2): 59-63.
- [3] Goodfellow I, Pouget-Abadie J, Mirza M, et al. Generative adversarial nets [C]//Annual Conference on Neural Information Processing Systems 2014, December 8-13, 2014, Montreal, Quebec, Canada. Copenhagen: MLR Press, 2014: 2672-2680.
- [4] Mirza M, Osindero S. Conditional generative adversarial nets[EB/OL]. (2014-11-06) [2022-10-08]. <https://arxiv.org/abs/1411.1784>.
- [5] Radford A, Metz L, Chintala S. Unsupervised representation learning with deep convolutional generative adversarial networks[EB/OL]. (2015-11-19) [2022-10-08]. <https://arxiv.org/abs/1511.06434>.
- [6] Zhang H, Xu T, Li H S, et al. StackGAN: text to photo-realistic image synthesis with stacked generative adversarial networks[C]//2017 IEEE International Conference on Computer Vision (ICCV), October 22-29, 2017, Venice, Italy. New York: IEEE Press, 2017: 5908-5916.
- [7] Pathak D, Krähenbühl P, Donahue J, et al. Context encoders: feature learning by inpainting[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 2536-2544.
- [8] Iizuka S, Simo-Serra E, Ishikawa H. Globally and locally consistent image completion[J]. ACM Transactions on Graphics, 2017, 36(4): 1-14.
- [9] Yu J H, Lin Z, Yang J M, et al. Generative image inpainting with contextual attention[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 18-23, 2018, Salt Lake City, UT, USA. New York: IEEE Press, 2018: 5505-5514.
- [10] Yu J H, Lin Z, Yang J M, et al. Free-form image inpainting with gated convolution[C]//2019 IEEE/CVF International Conference on Computer Vision (ICCV), October 27-November 2, 2019, Seoul, Korea (South). New York: IEEE Press, 2020: 4470-4479.
- [11] Nazeri K, Ng E, Joseph T, et al. EdgeConnect: generative image inpainting with adversarial edge learning [EB/OL]. (2019-01-01) [2022-10-08]. <https://arxiv.org/abs/1901.00212>.
- [12] 刘强, 张道畅. 结合 SENet 的密集卷积生成对抗网络图像修复方法[J]. 小型微型计算机系统, 2022, 43(5): 1056-1060.
Liu Q, Zhang D C. Dense convolution generative adversarial network image inpainting method with SENet [J]. Journal of Chinese Computer Systems, 2022, 43(5): 1056-1060.
- [13] Hu J, Shen L, Sun G. Squeeze-and-excitation networks [C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 18-23, 2018, Salt Lake City, UT, USA. New York: IEEE Press, 2018: 7132-7141.
- [14] 高文超, 任圣博, 田驰, 等. 多层次生成对抗网络的动画头像生成方法研究[J]. 计算机工程与应用, 2022, 58

- (9): 230-237.
- Gao W C, Ren S B, Tian C, et al. Research on method of animated avatar generation based on multi-level generative adversarial networks[J]. *Computer Engineering and Applications*, 2022, 58(9): 230-237.
- [15] 罗敦浪, 蒋旻, 袁琳君, 等. 基于条件生成对抗网络的图像着色研究[J]. *计算机工程与应用*, 2021, 57(13): 193-198.
- Luo D L, Jiang M, Yuan L J, et al. Research on image coloring based on conditional generative adversarial network[J]. *Computer Engineering and Applications*, 2021, 57(13): 193-198.
- [16] 侯向丹, 刘昊然, 刘洪普. 基于卷积自编码生成式对抗网络的高分辨率破损图像修复[J]. *中国图象图形学报*, 2022, 27(5): 1645-1656.
- Hou X D, Liu H R, Liu H P. High-resolution damaged images restoration based on convolutional auto-encoder generative adversarial network[J]. *Journal of Image and Graphics*, 2022, 27(5): 1645-1656.
- [17] 李海燕, 吴自莹, 郭磊, 等. 基于混合空洞卷积网络的多鉴别器图像修复[J]. *华中科技大学学报(自然科学版)*, 2021, 49(3): 40-45.
- Li H Y, Wu Z Y, Guo L, et al. Multi-discriminator image inpainting algorithm based on hybrid dilated convolution network[J]. *Journal of Huazhong University of Science and Technology (Natural Science Edition)*, 2021, 49(3): 40-45.
- [18] 陈紫柠, 张宏怡, 曾念寅, 等. 融合注意力机制的模糊图像多尺度复原[J]. *中国图象图形学报*, 2022, 27(5): 1682-1696.
- Chen Z N, Zhang H Y, Zeng N Y, et al. Attention mechanism embedded multi-scale restoration method for blurred image[J]. *Journal of Image and Graphics*, 2022, 27(5): 1682-1696.
- [19] 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.
- Li Y L, Gao Y, Yan J L, et al. Image inpainting methods based on deep neural networks: a review[J]. *Chinese Journal of Computers*, 2021, 44(11): 2295-2316.
- [20] Woo S, Park J, Lee J Y, et al. CBAM: convolutional block attention module[M]//Ferrari V, Hebert M, Sminchisescu C, et al. *Computer vision-ECCV 2018*. Lecture notes in computer science. Cham: Springer, 2018, 11211: 3-19.[LinkOut]
- [21] 黄友文, 唐欣, 周斌. 结合双注意力和结构相似度量的图像超分辨率重建网络[J]. *液晶与显示*, 2022, 37(3): 367-375.
- Huang Y W, Tang X, Zhou B. Image super-resolution reconstruction network with dual attention and structural similarity measure[J]. *Chinese Journal of Liquid Crystals and Displays*, 2022, 37(3): 367-375.
- [22] Huang Y W, Tang X, Zhou B. Image super-resolution reconstruction network with dual attention and structural similarity measure[J]. *Chinese Journal of Liquid Crystals and Displays*, 2022, 37(3): 367-375.
- [23] Wang P Q, Chen P F, Yuan Y, et al. Understanding convolution for semantic segmentation[C]//2018 IEEE Winter Conference on Applications of Computer Vision (WACV), March 12-15, 2018, Lake Tahoe, NV, USA. New York: IEEE Press, 2018: 1451-1460.
- [24] 周泽津, 王昊, 张小琴, 等. 基于Xception-TD的中华传统刺绣分类模型构建[J]. *数据分析与知识发现*, 2022, 6(S1): 338-347.
- Zhou Z Y, Wang H, Zhang X Q, et al. Classification model for Chinese traditional embroidery based on Xception-TD[J]. *Data Analysis and Knowledge Discovery*, 2022, 6(S1): 338-347.
- [25] 刘洁. 基于深度学习的唐卡图像修复系统的研究与实现[D]. 银川: 宁夏大学, 2020.
- Liu J. Research and implementation of thang-ga image inpainting system based on deep learning[D]. Yinchuan: Ningxia University, 2020.
- [26] Wang Z, Bovik A C, Sheikh H R, et al. Image quality assessment: from error visibility to structural similarity[J]. *IEEE Transactions on Image Processing*, 2004, 13(4): 600-612.
- [27] 康金梦. 基于对抗神经网络的敦煌壁画修复研究[D]. 天津: 天津大学, 2019.
- Kang J M. Research on murals inpainting based on generative adversarial networks[D]. Tianjin: Tianjin University, 2019.
- [28] Razavian A S, Azizpour H, Sullivan J, et al. CNN features off-the-shelf: an astounding baseline for recognition[C]//2014 IEEE Conference on Computer Vision and Pattern Recognition Workshops, June 23-28, 2014, Columbus, OH, USA. New York: IEEE Press, 2014: 512-519.