计算机科学 COMPUTER SCIENCE
http://www.jsjkx.com
DOI:10.11896/jsjkx.200800014

利用生成对抗网络的人脸图像分步补全法

林椹尠[1]　张梦凯[2]　吴成茂[3]　郑兴宁[2]
1 西安邮电大学理学院　西安 710121
2 西安邮电大学通信与信息工程学院　西安 710121
3 西安邮电大学电子工程学院　西安 710121
(lzhx126@126.com)

摘　要　人脸图像修复技术是近年来图像处理领域的研究热点,而人脸图像大面积缺失导致损失语义信息过多,一直是该领域的重点难点问题。针对这一问题,文中提出了一种基于生成对抗网络的图像分步补全算法。将人脸图像修复问题分为两步,设计两个串联的生成对抗网络,首先残缺图像通过预补全网络进行图像的预补全,预补全图像进入增强网络进行特征增强;判别器分别判断预补全图像和增强图像与理想图像的差异性;采用长短时记忆单元连接两部分的信息流,增强信息的传递。然后使用内容损失、对抗损失和全变分损失相结合的损失函数,提高网络的修复效果。最后在 CelebA 数据集上进行实验,结果显示,所提算法相较于对比算法在峰值信噪比指标上提高了 16.84%～22.85%,在结构相似性指标上提高了 10%～12.82%。

关键词:生成对抗网络;人脸图像;图像补全;长短时记忆;深度学习;缺失区域;跳跃连接
中图法分类号　TP183

Face Image Inpainting with Generative Adversarial Network

LIN Zhen-xian[1],ZHANG Meng-kai[2],WU Cheng-mao[3] and ZHENG Xing-ning[2]
1 School of Science,Xi'an University of Posts & Telecommunications,Xi'an 710121,China
2 School of Communication and Information Engineering,Xi'an University of Posts & Telecommunications,Xi'an 710121,China
3 School of Electronic Engineering,Xi'an University of Posts & Telecommunications,Xi'an 710121,China

Abstract　Face image inpainting is a hot topic of image processing research in recent years. Due to the loss of excessive sematic information,it is a difficult problem to inpaint large area missing of face images. Aiming at the problem of inpainting face images, a step-by-step image inpainting algorithm based on generative adversarial network is proposed. Face images inpainting task is divided into two steps. Firstly,face images are completed through the pre-completion network,and pre-completion images is enhanced feature through the enhancement network. The discriminator judges the difference between the pre-completion images,the enhanced images and the ideal image respectively. The long-term memory unit is used to connect the information flow of two parts. Secondly,the adversarial loss,content loss and total variation loss are combined to improve the effectively. Experiments are conducted on CelebA dataset,and this algorithm has an improvement of 16.84%～22.85% in PSNR and 10%～12.82% in SSIM compared with others typical image inpainting algorithms.

Keywords　Generative adversarial networks,Face images,Image inpainting,Long short term memory,Deep learning,Loss areas, Skip connection

图像补全指按照一定规则修复图像中的缺失区域,且补全内容符合图像原始信息和整体结构。然而由于缺失区域信息和背景结构信息的复杂性以及噪声干扰的不确定性,使得图像补全仍然是一个具有挑战性的问题。

传统的图像补全方法有两类,一类是基于像素扩散[1],一类是基于样本[2]。基于扩散的方法是图像像素层面的图像补全算法,将未缺损区域的有效信息扩散到缺损区域。基于扩散的图像补全方法通常使用偏微分方程[3-4],Bertalmio 等[5]提出了方向场扩散图像补全方法,该方法基于等照度线,将残缺区域周围的有效信息由外向内逐渐迭代扩散至缺损区域以补全图像。Levin 等[6]提出基于局部特征直方图的全局图像统计方法对破损图像进行补全,从统计学角度处理补全问题,使用图像特定分布找出给定边界分布最有可能的图像来修复缺失区域。此后发展出总变分[7]、欧拉弹性方程[8]、纳维斯托克斯方程[9]等利用微分方程解决图像补全问题的方法。Li 等[10]将缺失区域划分为子区域,通过扩散的方式对纹理信息进行融合,减少了运行时间,能得到更精确的纹理信息。基于扩散的方法对残缺区域较小的图片效果明显,当缺失区域较

到稿日期:2020-08-03　返修日期:2020-10-29
基金项目:国家自然科学基金(61671377)
This work was supported by the National Natural Science Foundation of China(61671377).
通信作者:张梦凯(zmkdyx@163.com)

大或者结构复杂时，填充区域模糊且不自然。

基于样本的补全方法利用图像结构的相关性，从已知区域寻找匹配度较高的图像块，将其复制或重新排列来补全残缺区域。这种方法从根本上改变了基于像素操作的图像补全方法面对缺失区域较大或复杂结构时的先天不足。Efros 等[11]首次提出图像块补全图像的思想，从源图像采集纹理块粘贴到目标缺损区域。Criminisi 等[12]利用最优图像块搜索技术，在使用图像块的基础上，兼顾图像的结构和纹理信息的传播，提高了图像的补全效果。Barnes 等[13]提出图像块匹配(PatchMatch)算法，利用随机图像块加速搜索算法，提高了图像补全效率。Wei 等[14]设计了一种新的迭代方法，充分利用了图像的局部一致性，通过将相关区域连接来增强视觉连续性，生成更准确的结构信息，而且对于不同的区域，根据置信度使用不同的处理方式，得到更合理的修复效果。由于基于样本的补全方法合成的纹理信息仅来自输入图像，无法生成原图像中不存在的新内容，对于人脸中眼睛、鼻子等独特的特征无法有效修复，从而限定了算法的通用性。

近年来，深度学习的方法在计算机视觉领域得到广泛应用，并且取得了惊人的成果[15-16]。因此，深度卷积神经网络被用于无监督学习的图像补全问题。Xie 等[17]首次提出将稀疏编码和卷积神经网络相结合，用于解决图像修复任务。Köhler 等[18]使用深度神经网络建立缺损图像与完整图像之间的映射关系来完成图像修复任务，可以生成图像中不存在的内容。2014 年，Goodfellow 等[19]受到博弈论的启发提出生成对抗网络(Generative Adversarial Network，GAN)。GAN 包括生成器和判别器两部分，在图像生成方面具有显著的优势，很快被用于图像补全任务。Pathak 等[20]提出语义编辑器(Context Encoder，CE)，基于 GAN 网络模型，利用对抗损失和均方误差损失共同训练网络，即使针对较大的缺失区域，利用深度卷积神经网络强大的特征提取能力也能得到不错的补全效果。但该方法修复的图像往往缺乏精细的纹理和细节，在补全部分周围会产生可见的伪影，且针对高分辨率图像，网络难以收敛。Yang 等[21]提出基于联合优化框架的多尺度神经图像块合成的高分辨率图像修复算法，通过使用卷积神经网络建模全局内容约束和局部纹理约束，从神经网络的中间层提取特征来合成逼真图像的内容和纹理，但该方法对于复杂的场景不能得到完整的语义信息，且训练速度较慢。Iizuka 等[22]提出全局和局部一致性图像补全法，通过使用经过训练的全局和局部上下文判别器来区分真实图像和补全图像。全局判别器查看整个图像，评估整个图像是否一致，而局部判别器仅查看补全区域，确保生成的图像块的局部一致性。Yu 等[23]引入感知机制，同时在生成模型中加入内容感知层。首先通过重建损失补全缺损信息；然后通过内容感知，将感知到的图像特征当作卷积过滤器，通过卷积从已知的图像信息中匹配最接近的图像；最后利用反卷积恢复缺失部分。Liu 等[24]提出局部卷积的方式，对卷积核增加掩膜操作，在卷积时忽略缺损区域像素，该方法对不规则缺失区域有较好的复原效果。

为了解决人脸图像缺失区域较大时难以补全得到完整语义信息的问题，本文提出了一种基于 GAN 网络的图像分步补全算法。将网络分为预补全网络和增强网络两部分，残缺图像首先输入预补全网络，对图像进行预补全；然后将预补全图像输入增强网络，提高图像的细节信息和语义特征。使用长短时记忆(Long Short-Term Memory，LSTM)[25]单元连接两个网络，以保证信息流的连通。使用内容损失、对抗损失和总变分损失共同优化网络。在 CelebA 数据集上将所提算法与其他算法进行对比实验，以验证本文算法在图像补全中的优势。

1 相关算法

1.1 生成对抗网络

生成对抗网络是深度学习中的一种无监督学习网络模型，其强大的图像生成能力使其在图像合成、图像修补、超分辨率等方面受到广泛应用。网络模型分为生成器(Generator，G)和判别器(Discriminator，D)。生成器模拟真实数据分布情况，以产生相似的样本数据为目的；判别器判断样本是否来自真实数据集。生成对抗网络模型如图 1 所示。

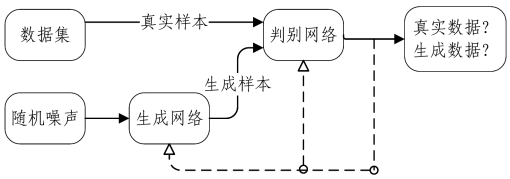


图 1 生成对抗网络模型

Fig. 1 Model of GAN

在训练过程中，生成器 G 的目标是生成接近真实样本的生成样本，以此来“迷惑”判别器 D。D 的目标是要判断输入的样本是生成的样本还是真实样本。这样，G 和 D 构成了一个动态的“博弈”，在相互对抗中提升各自的性能，这就是生成对抗网络的基本思想。两者通过不断地博弈学习，最终，G 生成的样本能够“以假乱真”，D 无法区分输入的样本是真实样本还是由生成器生成的样本，此时生成器达到最好状态。

1.2 长短时记忆

将图像补全任务分步处理，不同阶段间的信息传递是至关重要的。LSTM 能够缓解长序列训练过程中梯度消失或梯度爆炸的问题，通过 LSTM 单元连接相邻的网络，将上一阶段中的有用信息传递到下一阶段，从而稳定训练过程。

LSTM 的结构如图 2 所示，其中 C_t 为 LSTM 单元 t 时刻的细胞状态(cell state)，H_t 为 LSTM 单元 t 时刻的隐藏状态(hidden state)。此外，f_t，i_t，g_t 和 o_t 分别是遗忘门、输入门、细胞门和输出门。

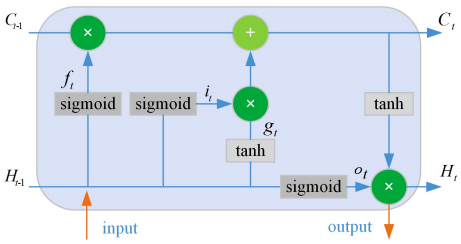


图 2 LSTM 结构

Fig. 2 Structure of LSTM

LSTM 首先决定细胞状态需要丢弃哪些信息，这需要由遗忘门来决定，输入通过 sigmoid 函数转换为 0 到 1 之间的值，0 表示丢弃，1 表示保留。遗忘门 f_t 的数学表达式为：

$$f_t = \mathrm{sigmoid}(w_{if} x_t + b_{if} + w_{hf} h_{t-1} + b_{hf}) \tag{1}$$

其次，决定给细胞状态添加哪些新的信息。首先通过输入门来决定更新的信息；然后通过细胞门得到新的候选细胞信息，这些信息可能会被更新到细胞状态中。输入门 i_t 和细胞门 g_t 的数学表达式为：

$$i_t = \mathrm{sigmoid}(w_{ii} x_t + b_{ii} + w_{hi} h_{t-1} + b_{hi}) \tag{2}$$

$$g_t = \tanh(w_{ig} x_t + b_{ig} + w_{hc} h_{t-1} + b_{hg}) \tag{3}$$

然后，更新旧的细胞状态。新的状态由两部分构成，一部分是通过遗忘门丢掉的部分长期记忆，另一部分是通过输入门添加的当前时刻新增的信息。细胞状态 C_t 的数学表达式为：

$$C_t = f_t * C_{t-1} + i_t * g_t \tag{4}$$

接着，输出门基于细胞状态决定具体的输出。输出门 o_t 的数学表达式为：

$$o_t = \mathrm{sigmoid}(w_{io} x_t + b_{io} + w_{ho} h_{t-1} + b_{ho}) \tag{5}$$

最后，根据新的细胞状态与输出门得到的判断条件更新隐藏状态，并作为最终的输出。隐藏状态 H_t 的数学表达式为：

$$H_t = o_t * \tanh(C_t) \tag{6}$$

$$output_t = H_t \tag{7}$$

其中，w 和 b 为可调参数，在训练时进行优化。

2 补全网络结构

补全网络模型由两个结构相同的子网络组成，每个子网络都是包括生成器和判别器的生成对抗网络。其中生成器是一个 U-Net 网络，判别器是全卷积网络。本文算法的网络结构如图 3 所示。

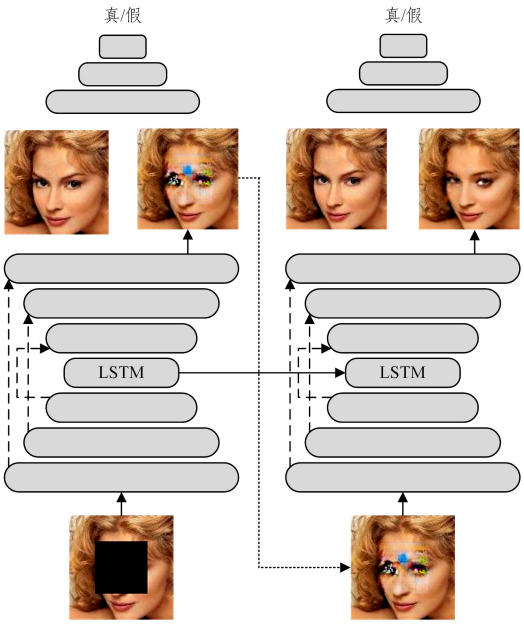


图 3 本文算法的框架图

Fig. 3 Frame diagram of proposed algorithm

2.1 生成器

生成器的目的是生成复原图像。现在大多数方法使用自编码器及其变体复原网络[26]。但是，Isola 等[27]发现 U-Net 网络[28]在图像生成方面有更强大的能力，并且使用 U-Net 在图像变换任务中取得了不错的效果。因此，本文使用 U-Net 作为生成器。

生成器结构如图 4 所示，其分为下采样和上采样两部分。其中左侧一列为输入图像的大小，卷积块下面数字为卷积核数量。缺损图像输入预补全网络的生成器中，首先进入下采样部分，下采样中卷积核大小为 5×5，使用大核能够有更大的感受野，从而得到更准确的结构信息；下采样部分最后一层是全连接层，通过全连接层得到 1 024 维向量，接着输入 LSTM 单元中；在 LSTM 单元中分别通过输入门、遗忘门、细胞门和输出门，得到 1 024 维的输出；将全连接层的输出与 LSTM 的输出拼接成 2 048 维的特征向量，输入到上采样部分。上采样中卷积核大小为 5×5，每一层的输入是上一层的输出与下采样对应层输出的拼接。上采样的最后一层恢复得到原图分辨率。

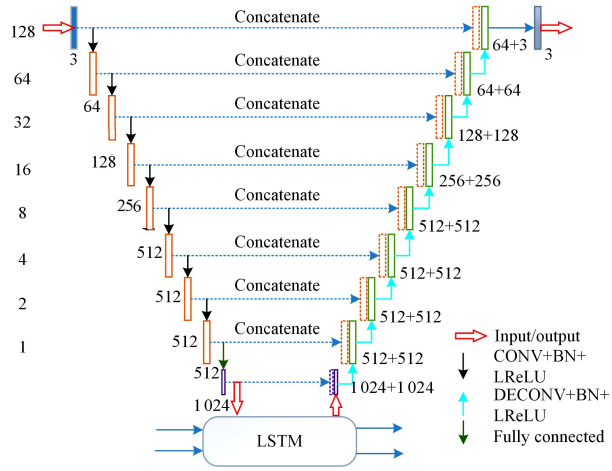


图 4 生成器结构(电子版为彩色)

Fig. 4 Structure of generator

预补全网络的输出同时输入到预补全网络的判别器和增强网络的生成器中。增强网络的生成器与预补全网络的结构相同，不同的是增强网络中卷积核大小为 3×3，小的卷积核能够得到更好的细节信息。增强网络的 LSTM 单元的输入为预补全网络 LSTM 单元的输出，以此来连接两个网络。

2.2 判别器

判别器用于估计输入的图片是真实样本的概率，如果输入来自真实的训练样本，则输出概率较高；如果输入是生成样本，则输出概率较低。判别器是一个全卷积网络，遵循 Radford 等[29]的设计样式，其网络参数如表 1 所列。

表 1 判别器网络参数

Table 1 Parameters of discriminator

网络层	核数量	核大小	步长	输出大小
输入图像	—	—	—	128×128×6
卷积层	32	3	1	128×128×32
卷积层	32	3	2	64×64×32
卷积层	64	3	1	64×64×64
卷积层	64	3	2	32×32×64
卷积层	128	3	1	32×32×128
卷积层	128	3	2	16×16×128
卷积层	256	3	1	16×16×256
卷积层	256	3	2	8×8×256
卷积层	512	3	1	8×8×512
卷积层	512	3	2	4×4×512
全连接层	1 024	—	—	1 024
全连接层	1	—	—	1
sigmoid	—	—	—	1

将理想图像和生成器的生成图像拼接作为判别器输入。判别器全部使用卷积核大小为 3×3 的卷积层，每个卷积层后面跟随一个 LReLU 激活函数；卷积层输出依次进入 1024 维的全连接层（Fully Connected Layer，FC）和 1 维的全连接层，最后通过 sigmoid 输出判别结果。

3 损失函数

为了使网络的训练更稳定，生成的图像效果更好，本文的损失函数包括内容损失、对抗损失和 TV 损失。本文的损失函数如下：

$$L=\alpha L_a+\beta L_C+\lambda L_{tv} \tag{8}$$

其中，α，β，λ 为损失函数的权重。

3.1 内容损失

内容损失通过对比生成图像和真实图像的每个像素值，保证最后的生成图像能够接近理想值。内容损失一般使用 L_1 损失或者 L_2 损失。然而研究发现[30]，由于 L_2 损失对异常值更加敏感，使用 L_1 损失可以减少重构图像中斑点状的人工痕迹，并减小误差，因此本文使用 L_1 损失作为内容损失。但是这里对 L_1 损失做了一些修改：

$$L_c=\|G(x_i)-\hat{x}_i\odot M\|_1 \tag{9}$$

其中，$\hat{x}$为真实图像，x 为残缺图像，$G(x)$为生成图像，M 为缺损块（缺损部分像素为 0，其他部分像素为 1）。

3.2 对抗损失

由于像素间的误差影响，内容损失函数将在所有解的平均值处收敛[31]，这会导致复原图像过于平滑，丢失锐利的边缘信息，而模糊依然存在。因此我们把内容损失和对抗损失结合起来共同作为损失函数。

在 GAN 网络中，生成器的目的是生成尽量真实的图片来迷惑判别器，判别器的目的是尽量判别出图片是来自生成器还是样本集。这样生成器和判别器在一个“零和游戏”中相互博弈。对抗损失的定义如下：

$$\begin{aligned}L_a&=\min_G\max_D V(D,G)\\&=E[\ln D(\hat{x}\odot M)]+E[\ln(1-D(G(x_i)))]\end{aligned} \tag{10}$$

其中，$\hat{x}$为真实图像，x 为残缺图像，G 为生成器，D 为判别器，M 为残缺块。

3.3 TV 损失

总变差损失（Total Variation，TV）[32]用于图像去噪和去模糊，但是，由于它可以促进生成图像中的空间平滑性，所以也被广泛应用于图像补全问题中。TV 损失的定义如下：

$$L_{tv}=\sum_{m,n}((x_{m,n+1}-x_{m,n})^2+(x_{m+1,n}-x_{m,n})^2) \tag{11}$$

其中，m，n 为图像中像素的坐标。

4 实验分析

4.1 实验数据集处理

实验使用的数据集为 Nah 等[33]公开的 CelebA，其中包括超过 20 万幅人脸图像。通过拉普拉斯算子筛选出清晰的图像 10000 幅，重新设置图像大小为 128×128 像素，并在图像中间区域设置 64×64 像素大小的遮挡区域，结合理想图像，制作训练集 8000 幅图像，测试集 2000 幅图像。

4.2 实验设置

实验中批处理大小为 16，设置 $\alpha=100$，$\beta=1$，$\lambda=100$。使用 ADAM 算法优化损失函数，设置网络初始学习率为 0.0002，衰减系数为 0.5。

实验中设置迭代周期为 50，每个周期遍历训练数据集，同时交替训练判别器和生成器。

4.3 评价指标

为了客观评价复原图像的效果，实验采用峰值信噪比（Peak Signal-to-Noise Ratio，PSNR）[34]和结构相似性（Structural Similarity，SSIM）[35]两种评价指标。PSNR 基于像素点间的误差来评价图像质量，*PSNR* 值越高表明图像失真越小。PSNR 的定义如下：

$$PSNR=20\log_{10}\left(\frac{2^n-1}{\sqrt{M}}\right) \tag{12}$$

其中，n 为每个像素的比特数，通常像素为 8 位比特即取 $n=8$；M 为两幅灰度图像的均方误差。

SSIM 以小像素块为单位，同时对比图像块的亮度、对比度和结构，*SSIM* 值越接近 1，表示两图像越接近。SSIM 的定义如下：

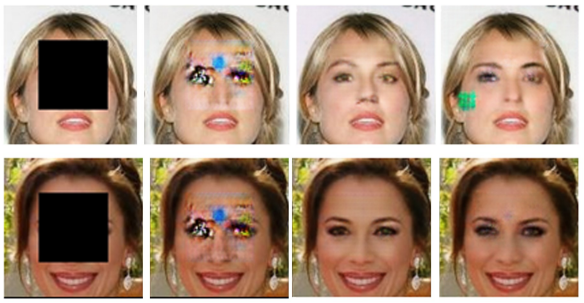
$$SSIM=\frac{(2\mu_x\mu_y+C_1)(2\sigma_{xy}+C_2)}{(\mu_x^2+\mu_y^2+C_1)(\sigma_x^2+\sigma_y^2+C_2)} \tag{13}$$

其中，x，y 为输入图像；μ 为图像的灰度平均值；σ 为图像的标准差；σ_{xy} 为两幅图像的协方差；C_1 和 C_2 为常数，通常取 $C_1=6.5025$，$C_2=58.5225$。

4.4 结果分析

4.4.1 生成器个数对图像修复效果的影响

图 5 对比了使用不同复原步骤对图像补全效果的影响，表 2 列出了不同步骤复原结果的量化指标。从图 5 可以看出，使用单生成器一步复原缺损图像，复原部分有明显的噪声，人脸特征复原不明显；使用 3 个生成器分 3 步复原人脸图像，能够基本复原人脸特征，但可能出现明显的色块或斑点；而通过两步复原图像，不仅能够复原人脸特征，而且复原的图像完整，效果逼近原图。因此本文将图像复原分为两步。



(a)遮挡图　(b)一步　(c)两步　(d)三步

图 5　不同复原步骤对图像补全效果的影响

Fig. 5　Effect of different steps on image restoration

表 2　不同步骤复原结果的量化指标

Table 2　Quantitative index of different steps of image restoration

	一步	两步	三步
PSNR	21.45	26.51	25.82
SSIM	0.81	0.88	0.87

4.4.2 LSTM 对图像复原的影响

图 6 对比了 LSTM 结构对图像复原的影响。可以看出，由于通过多层 UNet 网络导致输入信息被干扰，不使用 LSTM 结构的修复图像有明显的痕迹，且复原部分比较模糊，而使用 LSTM 结构的修复图像完整清晰。

(a)遮挡图 (b)无 LSTM (c)有 LSTM

图 6 LSTM 对图像复原的影响

Fig. 6 Effect of LSTM on image restoration

4.4.3 不同算法的补全效果对比

图 7 展示了文献[23,36]中的方法与本文算法在 CelebA 数据集上修复图像的对比结果。可以明显看出，文献[23]的方法虽能生成完整的人脸，但是人脸的特征不符合原图；文献[36]的方法对于大面积的残缺块的补全效果不明显，不能得到一幅完整的人脸图像，复原的人脸与原图特征相差较大；而本文算法能够完整地复原人脸，且基本符合原图人脸的语义特征。

(a)原图 (b)遮挡图 (c)文献[23] (d)文献[36] (e)本文

图 7 CelebA 数据集上的修复效果

Fig. 7 Inpainting results on CelebA dataset

表 3 列出了不同算法的图像质量评价结果。可以看出，本文算法的量化指标明显优于对比算法。

表 3 CelebA 数据集上修复结果的 *PSNR* 和 *SSIM*

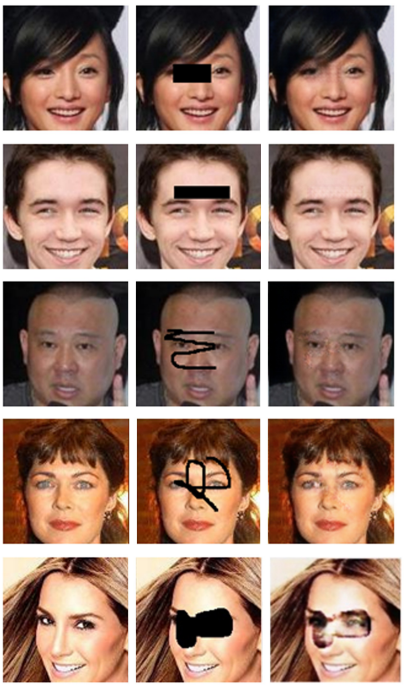
Table 3 *PSNR* and *SSIM* of inpainting results on CelebA dataset

人脸图像	文献[23]	文献[36]	本文算法
PSNR	22.69	21.58	26.51
SSIM	0.80	0.78	0.88

4.4.4 算法局限性分析

本文算法对于中心矩形区域残缺的图像有较好的补全效果，能够完整地复原人脸特征，并符合图像语义信息，但通过实验发现本文算法仍有局限性。

图 8 展示了本文算法对于不规则区域的补全效果。可以看出，对于较小的残缺区域，即使残缺部分是不规则的，本文算法也能够较好地修复。然而对于较大的不规则残缺区域，本文算法的补全效果并不好，不能得到完整的图像特征，且在补全区域有明显“黑边”。

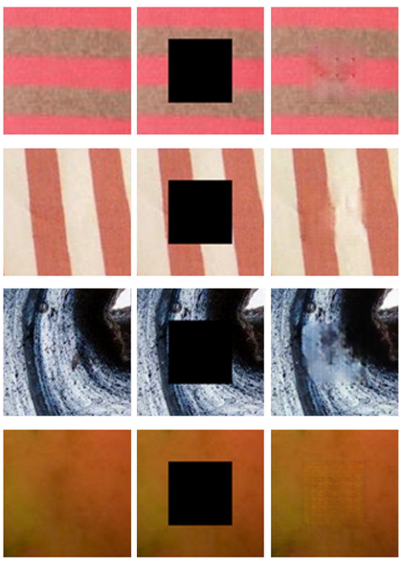


(a)原图 (b)残缺图像 (c)补全图像

图 8 对不规则区域的修复效果

Fig. 8 Inpainting results to irregular area

图 9 给出了本文算法在 DTD 纹理数据库[37]中的修复效果。可以看出，对于纹理数据，本文算法能够基本复原图像中损失的信息，但是由于 TV 损失，使得复原部分模糊，而且存在明显的修复痕迹。



(a)原图 (b)残缺图像 (c)补全图像

图 9 DTD 数据集上的修复效果

Fig. 9 Inpainting results on DTD dataset

结束语 针对人脸图像大面积缺损的问题，本文提出了

一种基于生成对抗网络的图像分步补全方法。本文所提网络由预补全网络和增强网络两个结构相同的子网络组成，子网络是以 U-Net 为生成器、以全卷积网络为判别器的生成对抗网络。缺损图像输入网络中，首先通过预补全网络得到预补全图像，然后预补全图像通过增强网络得到补全图像。在 CelebA 数据集上进行实验，通过分析对比可知，本文方法能够在保证语义信息的情况下对规则的损失区域进行更好的修复。

然而，本文方法修复对象单一、修复模型不具有通用型、对于不规则损失区域的修复仍有局限性。在未来的模型训练中，我们将尝试加入多种类、多样式的训练数据，以提高模型的泛化性能。

参 考 文 献

[1] HE Y T，TANG X H，ZHANG Y，et al. Improved Criminisi algorithm based on structure tensor[J]. Journal of Image and Graphics，2018，23(10)：1492-1507.

[2] LAN X L，LIU H X，YAO H B. Improved image inpainting algorithm based on texture blocks and gradient feature[J]. Computer Engineering and Applications，2018，54(20)：172-177.

[3] KRHENBÜHL P，KOLTUN V. Efficient Inference in Fully Connected CRFs with Gaussian Edge Potentials[J]. Advances in Neural Information Processing Systems，2012，24(2011)：109-117.

[4] XU L M，WU Y J，ZHANG B. Image Inpainting Algorithm Based on Adaptive High Order Variation in Eight Neighbors [J]. Journal of Graphics，2017，38(4)：558-565.

[5] BERTALMIO M，SAPIRO G，CASELLES V，et al. Image inpainting[C]// Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques. ACM Press/ Addison-Wesley Publishing Co.，2000：417-424.

[6] LEVIN A，ZOMET A，WEISS Y. Learning how to inpaint from global mage statistics[C]// IEEE In International Conference on Computer Vision. 2003：305-312.

[7] SHEN J，CHAN T F. Mathematical Models for Local Nontexture Inpaintings[J]. SIAM Journal on Applied Mathematics，2001，62：1019-1043.

[8] SHEN J，KANG S H，CHAN T F. Euler's Elastica and Curvature-Based Inpainting[J]. SIAM Journal on Applied Mathematics，2002，63(2)：564-592.

[9] BOUREAU Y L，BACH F，LECUN Y，et al. Learning mid-level features for recognition[C]// 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. IEEE，2010，2559-2566.

[10] LI S，ZHAO M. Image inpainting with salient structure completion and texture propagation[J]. Pattern Recognition Letters，2011，32(9)：1256-1266.

[11] EFROS A A，LEUNG T K. Texture synthesis by non-parametric sampling[C]// Proceedings of the Seventh IEEE International Conference on Computer Vision. 1999，2：1033-1038.

[12] CRIMINISI A，PÉREZ P，TOYAMA K. Region filling and object removal by exemplar-based image inpainting [J]. IEEE Transactions on Image Processing，2004，13(9)：1200-1212.

[13] BARNES C，SHECHTMAN E，FINKELSTEIN A，et al. PatchMatch：A randomized correspondence algorithm for structural image editing[C]// ACM Transactions on Graphics (ToG). ACM，2009：1-11.

[14] WEI Y，LIU S. Domain-based structure-aware image inpainting [J]. Signal Image & Video Processing，2016，10(5)：911-919.

[15] GIRSHICK R，DONAHUE J，DARRELL T，et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2014：580-587.

[16] LONG J，SHELHAMER E，DARRELL T. Fully convolutional networks for semantic segmentation[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2015：3431-3440.

[17] XIE J，XU L，CHEN E. Image denoising and inpainting with deep neural networks[C]// Advances in Neural Information Processing Systems. 2012：341-349.

[18] KÖHLER R，SCHULER C，SCHÖLKOPF B，et al. Mask. specific inpainting with deep neural networks[C]// German Conference on Pattern Recognition. Cham：Springer，2014：523-534.

[19] GOODFELLOW I，POUGET-ABADIE J，MIRZAM，et al. Generative adversarial nets[C]// Advances in Neural Information Processing Systems. 2014：2672-2680.

[20] PATHAK D，KRAHENBUHL P，DONAHUE J，et al. Context encoders：Feature learning by inpainting[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2016：2536-2544.

[21] YANG C，LU X，LIN Z，et al. High-resolution image inpainting using multi-scale neural patch synthesis[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2017：6721-6729.

[22] IIZUKA S S，SERRA E，SHIKAWA H. Globally and locally consistent image completion[J]. ACM Transactions on Graphics (TOG)，2017，36(4)：107：1-2.

[23] YU J H，ZHE L，YANG J M，et al. Generative Image Inpainting with Contextual Attention [C]// 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2018：5505-5514.

[24] LIU G，REDA F A，SHIH K J，et al. Image inpainting for irregular holes using partial convolutions [C]// Proceedings of the European Conference on Computer Vision (ECCV). 2018：85-100.

[25] HOCHREITER S，SCHMIDHUBER J. Long Short-Term Memory[J]. Neural Computation，1997，9(8)：1735-1780.

[26] YEH R A，CHEN C，LIM T Y，et al. Semantic Image Inpainting with Deep Generative Models [C]// CVPR 2016. 2016：5485-5493.

[27] ISOLA P，ZHU J Y，ZHOU T H，et al. Image-to-image translation with conditional adversarial networks[C]// IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2017.

[28] RONNEBERGER O，FISCHER P，BROX T. U-Net：Convolutional Networks for Biomedical Image Segmentation[C]// International Conference on Medical Image Computing and Computer-assisted Intervention. Cham：Springer，2015：234-241.

[29] RADFORD A, METZ L, CHINTALA S. Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks [C]// Proceedings of the International Conference on Learning Representations (ICLR). 2016.

[30] ZHAO H, GALLO O, FROSIO I, et al. Loss functions for image restoration with neural networks [J]. IEEE Transactions on Computational Imaging, 2016, 3(1): 47-57.

[31] WHYTE O, SIVIC J, ZISSERMAN A, et al. Non-uniform deblurring for shaken images[J]. International Journal of Computer Vision, 2012, 98(2): 168-186.

[32] RUDIN L I, OSHER S, FATEMI E. Nonlinear total variation based noise removal algorithms[J]. Physica D Nonlinear Phenomena, 1992, 60(1/2/3/4): 259-268.

[33] NAH S, KIM T H, LEE K M. Deep multi-scale convolutional neural network for dynamic scene deblurring[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2017: 3883-3891.

[34] TONG Y B, ZHANG Q S, QI Y P. Image Quality Assessing by Combining PSNR with SSIM [J]. Journal of Image and Graphics, 2006, 11(12), 1758-1763.

[35] 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.

[36] ZHENG C X, CHAM T J, CAI J. Pluralistic Image Completion [C]// 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2019: 1438-1447.

[37] CIMPOI M, MAJI S, KOKKINOS I, et al. Describing Textures in the Wild[C]// 2014 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2014: 3606-3613.

LIN Zhen-xian, born in 1969, Ph.D, professor. Her main research interests include wavelet theory and its application in signal and image.

ZHANG Meng-kai, born in 1995, postgraduate. His main research interests include deep learning and computer vision.