

基于生成对抗网络的图像修复

孙 全 曾晓勤

(河海大学计算机与信息学院智能科学与技术研究所 南京 211100)

摘 要 针对现有图像修复算法存在受损区域的形状和大小受限以及修复痕迹明显、修复边缘不连续的问题,文中提出一种基于生成对抗网络的图像修复方法。该方法采用生成对抗网络(Generative Adversarial Networks, GAN)这种新的生成模型作为基本架构,结合 Wasserstein 距离,同时融入条件对抗网络(CGAN)的思想;以破损图像作为附加条件信息,采用对抗损失与内容损失相结合的方式训练网络模型,以修复破损区域。此方法能够修复大多数破损情况下的图像。在 CelebA 和 LFW 两个数据集上的实验结果表明,所提方法能够取得很好的修复效果。

关键词 生成对抗网络,图像修复,对抗学习,Wasserstein 距离

中图分类号 TP391.41 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2018.12.038

Image Inpainting Based on Generative Adversarial Networks

SUN Quan ZENG Xiao-qin

(Institute of Intelligence Science and Technology, College of Computer and Information, Hohai University, Nanjing 211100, China)

Abstract Aiming at the problems of the restricted shape and size of the damaged area, the obvious inpainting tracks and the discontinuous inpainting edge in the existing image inpainting algorithms, this paper proposed an image inpainting method based on generative adversarial networks. In this method, a new generative model named generative adversarial networks(GAN) is used as the basic framework with combining Wasserstein distance and the idea of conditional generative adversarial networks(CGAN). The network receives the damaged image as additional conditional information and combines adversarial loss with content loss to train the network model for restoring pixels of missing areas. This method can be used to repair most of the damages in images. The experimental results on two datasets of CelebA and LFW suggest the capability of this method to obtain good performance.

Keywords Generative adversarial networks(GAN), Image inpainting, Adversarial learning, Wasserstein distance

1 引言

图像修复是指在一张图片中存在局部区域出现破损或空洞,需要利用已有的边缘信息来将这些破损区域进行修复或补充的工作。作为目前图像处理的一种重要方法,图像修复具有重要的现实意义,尤其是在文物修复方面具有重大的历史意义,在资料修复、虚拟现实、图像缩放等方面也具有一定的应用价值。

经典的图像修复方法主要采用局部或非局部的信息来修复缺失部分,而且大多数现有的方法主要针对对张图片进行修复。全变分方法(Total Variation)^[1]考虑了图像的平滑特性,在修复小的缺失区域和消除噪音时很有效。但是此方法仅适用于较小区域的缺失情况,当缺失区域较大时,用此方法将非破损区域的像素扩散到破损区域时就会产生模糊,图像的修复效果较差。块匹配方法(Patch Match)^[2]是另外一种利用已有信息的图像修复方法,它从图像中可获得的部分搜索与缺失块相似的区域来填充缺失区域,获得了较好的修复

效果。但是,此方法太依赖于图像的已知部分,如果已知部分不存在这样的相似块,那么将不能较好地修复缺失区域。

为了能够修复更大缺失区域和任意形状的缺失部分图像,更多的基于学习的方法被提出,这也得益于神经网络和深度学习的不断发展。上下文编码器(Context Encoders)^[3]就是一种采用深度网络来进行图像修复的方法,它利用编解码器作为图像修复的主网络,将破损图像编码后再经过解码器生成缺失区域部分,同时利用对抗损失来使得生成的图像尽可能真实,能够较好地修复缺失区域。但是,该方法存在缺失区域与非缺失区域不连续的问题,有明显的修补痕迹。生成对抗网络(GAN)^[4]是一种新的图像修复方法,以上下文为条件的对抗网络(CC-GAN)^[5]就是其中一种方法,该网络使用编解码器作为生成器,使用 VGG^[6]网络作为判别器,来实现图像修复,同时加入上下文损失项来使得修复后的图像的修复痕迹不那么明显。Yeh 提出的深度生成模型^[7]也利用了对抗网络来实现图像修复,与之前提到的采用对抗网络进行修复的模型不同,该网络使用无破损的图像作为数据集来

到稿日期:2017-11-01 返修日期:2018-01-17 本文受国家重点研发计划项目(2017YFC0405805)资助。

孙 全(1993-),男,硕士生,主要研究方向为神经网络、图像处理,E-mail:905152434@qq.com;曾晓勤(1957-),男,博士,教授,主要研究方向为神经网络、机器学习,E-mail:xzeng@hhu.edu.cn(通信作者)。

训练一个深度卷积生成对抗网络(DCGAN)^[8]模型,训练完后利用训练好的模型来寻找与缺失图像最接近的编码以实现图像修复。

相比之前提到的两种利用对抗网络进行图像修复的方法,本文所提方法的创新之处主要体现在以下几个方面。

1) 使用训练更加稳定的 Wasserstein 对抗网络(WGAN)^[9]替代 DCGAN^[8]来训练模型。

2) 不同于 CC-GAN 中生成器使用的编码解码器结构,本文采用带有跳跃连接的能够将低级特征直接连接到网络较高层的 U-Net^[10]网络。

3) 直接使用破损图像训练网络模型,不同于 Yeh 间接地使用无缺失样本完成图像修复。

4) 为了得到更好的修复图像,本文将对抗损失与内容损失相结合来训练模型。

我们在 CelebA 和 LFW 人脸数据集上验证模型的修复效果,并与已有修复方法的修复效果进行对比。实验结果表明,本文提出的修复方法的修复效果较好,甚至优于已有方法。

2 相关内容

2.1 生成对抗网络

GAN^[4]是 Goodfellow 于 2014 年提出的一种训练生成模型的新方法,它的思想来源于二人零和博弈(即二人的利益之和为零,一方的所得正是另一方的所失)^[11]。不同于以往的生成模型,它通过一个判别模型来引导生成模型的训练。网络由两部分组成:生成器 G 和判别器 D 。生成器用于捕获训练数据的分布;判别器相当于一个二分类器,用于判断输入数据是来自真实数据还是生成数据,网络输出的结果代表输入数据来自真实数据的概率。生成对抗网络的基本结构如图 1 所示。

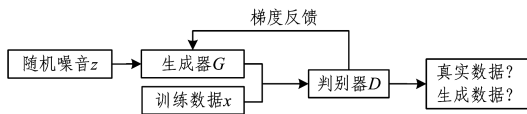


图 1 生成对抗网络模型

Fig. 1 GAN model

在网络模型中,生成器接收一个随机噪音 z 作为输入, z 服从分布 $p_z(z)$,这样生成器便可生成样本 $G(z; \theta_g)$,生成样本服从分布 p_g ,其中 θ_g 为生成器的参数。判别器接收真实数据 x 或者生成器的生成数据 $G(z; \theta_g)$ 作为输入,然后输出 $D(x; \theta_d)$,其中 x 服从分布 $p_r(x)$, θ_d 为判别器的参数, $D(x)$ 代表输入的数据 x 属于真实样本的概率。判别器和生成器同时进行训练,网络的目标是达到一个生成器和判别器的纳什均衡(Nash Equilibrium)^[12],使得判别器的输出结果 $D(x)$ 恒为 $1/2$,即判别器已无法分辨出真实样本和生成样本,基本靠猜测。

判别器的优化目标为尽可能地给真实样本和生成样本分配正确的标签,即调整参数 θ_d 来最大化 $\log(D(x)) + \log(1 - D(G(z)))$;生成器的优化目标为使生成样本尽可能接近真实样本,即调整参数 θ_g 来最小化 $\log(1 - D(G(z)))$ 。整个网络

类似于二人极小极大游戏,代价函数 $V(D, G)$ 如式(1)所示:

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

其中, E 代表求解期望值,下文公式中所有出现的 E 均代表此意。

2.2 Wasserstein 对抗网络

最初提出的对抗网络中的生成器只接收随机噪音作为输入,之后 Mirza 提出了一种叫作条件对抗网络^[13]的对抗模型。该网络接收一个额外信息作为网络的输入,额外信息的加入可以控制生成器的生成模式,能够按照我们的要求做生成任务。加入条件后的网络的代价函数 $V(D, G)$ 如式(2)所示:

$$\min_G \max_D V(D, G) = E_{(x,y) \sim p_r(x,y)} [\log(D(x,y))] + E_{z \sim p_z(z), (x,y) \sim p_r(x,y)} [\log(1 - D(G(z, y), y))] \quad (2)$$

随后,Arjovsky 等又提出了 Wasserstein 对抗网络^[9],该网络是解决对抗网络训练不稳定问题的一种比较彻底的方法,它修改了原始对抗网络中的损失函数,并通过 Earth-Mover 距离来衡量两个分布之间的差距,取代了之前的 JS 散度和 KL 散度。

Earth-Mover 距离可以表示成如下形式:

$$W(p_r, p_g) = \frac{1}{K} \sup_{\|f\|_L \leq K} E_{x \sim p_r} [f(x)] - E_{x \sim p_g} [f(x)] \quad (3)$$

其中, $\|f\|_L \leq K$ 表示函数 f 需要满足 Lipschitz 条件限制。Lipschitz 条件限制的简单理解就是限制一个连续函数的最大局部变动幅度小于常数 K 。我们用一个神经网络来表示函数 $f(x)$,即对应我们的判别器网络。对此限制的施加通过给判别器的损失函数添加一个梯度惩罚项来实现,即 $E_{x \sim p_x} [\|\nabla_x D(x)\|_p - 1]^2$,其中 x 从生成样本和真实样本之间的样本分布 p_x 上进行随机采样。

Wasserstein 对抗网络中生成器的损失函数为:

$$-E_{x \sim p_g} [D(x)] \quad (4)$$

判别器的损失函数为:

$$E_{x \sim p_g} [D(x)] - E_{x \sim p_r} [D(x)] + \lambda E_{x \sim p_x} [\|\nabla_x D(x)\|_p - 1]^2 \quad (5)$$

其中, λ 为平衡不同损失项的权重系数。

3 网络结构

3.1 生成器

从最初的使用多层感知机(MLP)解决 MNIST 手写体数据集的生成^[4],到后来使用深度卷积网络构造生成器来实现更加复杂的图像生成任务^[8],生成器的网络结构随着处理图像复杂度的变化也在不断变化。

对于图像修复问题,图像可能存在局部像素点的像素破损的情况,我们只需要对图像中的这些点进行修复即可,而图片的底层结构还是应该保持一致,也就是说输入与输出之间可能共享一些低级信息,这些信息可能有利于提高输出图像的质量。考虑到这一点,采用带有跳跃连接(Skip Connection)的编码解码器来将低级特征直接传送到网络的更高层

上,即将编码器低级特征与对应的解码器层得到的高级特征进行拼接后作为下一层网络的输入,这样可以在修复图像的同时最大可能地保持图像的底层结构。本实验中,我们采用 Encoder-Decoder 的一个变体即 U-Net^[10]来构造生成器,同时生成器不直接输入随机噪音,而在网络中间层通过 dropout 引入噪音。整个生成器的执行流程为:将输入送至编码器中,编码器对其进行编码后将得到的编码输入到解码器中,解码器再对其做解码工作以生成修复后的图像。生成器的整体结

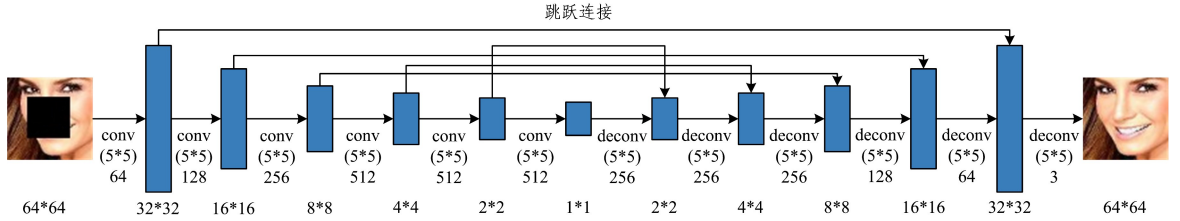


图 2 生成器的网络结构

Fig. 2 Network structure of generator

3.2 判别器

判别器的主要功能在于与生成器形成对抗学习,不断提升自己的能力,以提高识别出真假样本的准确率。本文中判别器由一系列的卷积层构成,结构如图 3 所示,图中 conv 代表卷积层,full-connection 代表全相连层。

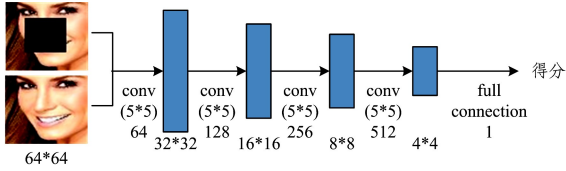


图 3 判别器的网络结构

Fig. 3 Network structure of discriminator

判别器有两种类型的输入,即生成器生成的图像以及真实图像,网络结构中为输入真实样本(图 3 中下方的图像)的情况。同时我们采用条件对抗网络的思想,以破损图像(图 3 中上方的图像)为条件,通过简单的通道拼接来将条件与原输入拼接作为判别器的最终输入。网络输出一个标量,来表示判别器对输入的评价。

在判别器中加入破损图像作为条件的目的在于,只有当生成样本足够真实并与破损图像匹配时,我们才给其更高的分数,即要求修复图像与原图尽量保持一致。因为我们使用 Wasserstein 距离来衡量两个分布之间的差距,需要对判别器的每一个输入样本进行独立的 Lipschitz 限制,所以在判别器的所有层中均不采用批量规范化^[14],这样可以避免同一个 batch 中的不同样本之间存在相互依赖。同时,去掉原网络结构中最后一层的 sigmoid 激活函数。

4 损失函数

4.1 判别器的损失函数

在众多不同结构的对抗网络中,判别器的损失函数基本都保持一致,本文所使用的损失函数形式也一样。对于判别器而言,一共接收两种可能的输入:1)真实样本,对于真实样本而言,判别器的目的是给予其更高的评分,即 $E_{x \sim p_r(x)} D(x)$

足够大;2)生成样本,对于生成样本而言,判别器的目的是给予其较低的评分,即 $E_{x \sim p_g(x)} D(x)$ 足够小。因为我们采用 Wasserstein 距离来衡量生成样本分布与真实样本分布之间的差距,所以还需要对判别器进行 Lipschitz 限制,即附加判别器对于其输入的梯度惩罚 $E_{x \sim p_g} [\|\nabla_x D(x)\|_p - 1]^2$ 。由于在整个样本空间上实施 Lipschitz 限制是不太现实的,因此 WGAN 的提出者选择生成样本的集中区域、真实样本的集中区域以及它们之间的这部分数据, p_x 代表的就是这一部分的样本分布情况。结合以上两个部分,判别器的损失函数为:

$$E_{z \sim p_z(z)} D(G(z)) - E_{x \sim p_r(x)} D(x) + \lambda_1 E_{x \sim p_x} [\|\nabla_x D(x)\|_p - 1]^2 \quad (6)$$

由于我们在判别器中加入了条件,因此损失函数的形式对应修改为:

$$E_{z \sim p_z(z), (x, y) \sim p_r(x, y)} D(G(z, y)) - E_{(x, y) \sim p_r(x, y)} D(x, y) + \lambda_1 E_{(x, y) \sim p_r(x, y), x_1 \sim p_x} [\|\nabla_{x_1} D(x_1, y)\|_p - 1]^2 \quad (7)$$

其中, λ_1 为平衡不同损失项的权重系数。

4.2 生成器的损失函数

生成器的生成样本要足够真实,这样判别器才能以更大的概率将其判定为真实样本。同样,我们希望生成器生成的样本与真实样本足够相似。基于此考虑,我们在对抗损失的基础上加入了内容损失,并将这两部分损失结合在一起作为最后的损失函数。下面将详细介绍这两部分损失函数的形式。

4.2.1 对抗损失

与以往的对抗网络一样,生成器希望自己生成的样本能够得到判别器最大的认可,即将其生成的样本输入到判别器中,判别器能够给予更高的评分。同时,我们采用破损图像作为条件加入到生成器的输入中,因此对抗损失为:

$$-E_{z \sim p_z(z), (x, y) \sim p_r(x, y)} D(G(z, y), y) \quad (8)$$

4.2.2 内容损失

对于图像修复问题,我们希望修复后的破损区域能够尽可能真实,同时未破损的区域能够保持原样。基于此考虑,我们使用 L1 距离来衡量生成器生成的图像的未破损区域与真实样本中未破损区域的差别。假设 M 为图像掩膜,其与输入

图像具有同样大小的尺寸,其中值为0的部分代表破损像素点,值为1的部分代表未破损像素点,则定义内容损失为如下的形式:

$$E_{z \sim p_z(z), (x,y) \sim p_r(x,y)} [\|M \odot (x - G(z, y))\|_1] \quad (9)$$

其中,符号 \odot 代表两个矩阵的对应元素相乘。

最终,结合以上两部分损失函数,生成器的损失函数的形式为:

$$-E_{z \sim p_z(z), (x,y) \sim p_r(x,y)} D(G(z, y), y) + \lambda_2 E_{z \sim p_z(z), (x,y) \sim p_r(x,y)} [\|M \odot (x - G(z, y))\|_1] \quad (10)$$

其中, λ_2 为平衡不同损失项的权重系数。

5 实验

5.1 数据集的预处理

CelebA 是一个由 200000 张左右的人脸图像组成的数据集,数据集中的图像尺寸均为 178×218 。我们在做人脸修复时,只需要人脸区域,因此使用基于 Haar 特征^[15]、采用 Ada-boost 算法^[16]训练的人脸检测器从一张图像中检测出人脸区域并分割,同时将其归一化到 64×64 的尺寸。实验中选取处理后的 100000 张图片作为训练集来训练网络,同时为了区别于训练数据,对 LFW 数据集中的图像做了同样的处理,并将其作为测试集。数据集中单张图片的处理流程如图 4 所示。

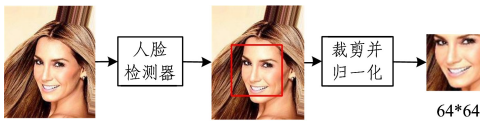


图 4 数据集中单张图片的处理流程

Fig. 4 Processing flow of single picture in datasets

5.2 训练细节

实验中,我们在 CelebA 的部分数据集上迭代训练 10 个周期,每一个周期迭代训练 25000 次,每一次交替训练判别器和生成器各一次,采用文献^[17]中的对抗网络算法进行训练。将每一个 batch 的大小设置为 4,参数 λ_1 设置为 10,参数 λ_2 设置为 100。训练中,网络的学习速率设置为 0.0001,采用 Adam^[18]优化算法,其中参数 beta1 设置为 0.5, beta2 设置为 0.9。最后,我们交换训练集和测试集来验证模型的健壮性。但是,由于 LFW 数据集只有 10000 张左右的图片,相对于之前用于训练的 CelebA 数据集只有 1/10,因此我们在 LFW 数据集上迭代训练 40 个周期,且每一个周期迭代训练 2500 次,其他参数保持一致。

5.3 修复图像的后处理

考虑到生成器生成的非破损区域并不可能与原图 100% 相同,如果直接取生成器生成的整体图像作为修复后的图像,原本图像中未破损的区域也会被修改,因此只取破损区域,并与原图中未破损区域进行混合,从而得到最终的修复完成图像。

Yeh 等^[7]还对最后的修复结果进行了泊松图像编辑^[19],以进行图像混合。泊松图像编辑是一种能够让两张图片完全且自然地进行无缝混合的操作。该方法通过求解泊松方程来实现混合区域的插值处理,使得处理过的图像能够完美地混合在一起。

如果不加这一混合操作,Yeh 等修复的破损区域与非破损区域的拼接痕迹会十分明显;但本文模型不存在该问题,直接进行拼接的边缘就很平滑。各方法对图像的后处理效果对比如图 5 所示,从左到右分别是原图、破损图片、Yeh 的未加泊松混合处理的修复图片、Yeh 的添加泊松混合处理的修复图片和利用本文方法直接得到的修复图片。从图 5 中可以看出,本文修复方法不存在修复痕迹明显的情况。



图 5 图像的后处理对比图

Fig. 5 Post-processing comparison of images

5.4 实验结果分析

首先通过训练过程中损失函数的变化情况来分析对抗网络一直存在的训练不稳定的问题,然后通过是否添加内容损失项来比较修复效果的好坏,最后与 Yeh 等^[7]使用的方法进行对比来分析修复效果,同时对破损区域较大和破损区域形状不规则的情况进行实验。

5.4.1 训练算法稳定性的分析

基于 KL 散度和 JS 散度来衡量生成样本与真实样本分布差距的对抗网络训练方法一直存在训练不稳定的问题,典型的模型就是深度卷积生成对抗网络^[6]。为了解决这一问题,本文采用 Wasserstein 距离来衡量两个分布之间的差距。训练过程中,判别器损失函数的变化情况如图 6 所示。

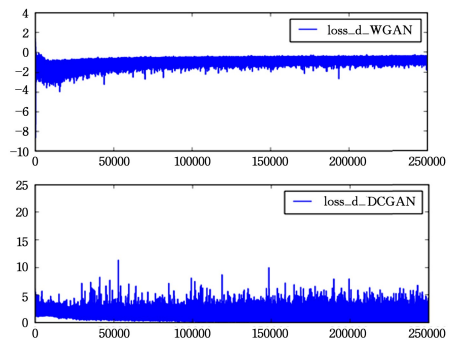


图 6 判别器损失函数的变化情况

Fig. 6 Change of discriminator loss function

从损失函数的变化情况中可以看出,相比于 WGAN,采用 DCGAN 网络模型的训练过程中出现了训练不稳定的现象,在整个训练过程中,损失函数经常出现忽高忽低的情况,由此可见,使用 WGAN 训练模型较为稳定。同时,我们从图 6 中也可以看出,大约迭代训练 100000 次时,WGAN 损失函数已基本保持不变,网络已基本处于收敛状态,这两点也充分证明了 WGAN 模型的训练稳定性和快速收敛性。

5.4.2 内容损失项重要性的分析

内容损失项的作用在于使得生成的图像与真实样本尽可能接近。为了验证此项损失函数对图像修复效果的影响,将参数分别设置为 0 和 100 来比较实验效果,结果如图 7 所示。图中从左到右分别是未破损图像、破损图像、包含内容损失的

修复结果和不包含内容损失的修复结果。



图 7 添加内容损失与否的对比

Fig. 7 Comparison of adding content loss or not

从图 7 可以看出,相比于添加内容损失项,不添加内容损失项的图像修复效果较差,这也验证了我们之前选择添加此损失项的用意。在后面的实验中,我们均加上内容损失项来完成修复工作。

5.4.3 跳跃连接重要性的分析

前文已从理论上分析了使用跳跃连接的好处,为验证跳跃连接在修复效果上的重要性,对生成器分别采用无跳跃连接的 Encoder-decoder 和带跳跃连接的 U-Net 进行对比,结果如图 8 所示。图中从左到右分别为原图、破损图像、使用跳跃连接的修复结果和未使用跳跃连接的修复结果。

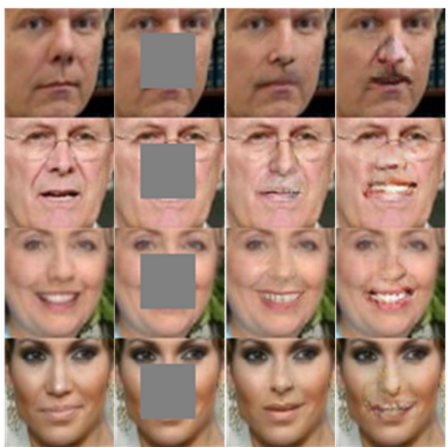


图 8 添加跳跃连接与否的对比

Fig. 8 Comparison of adding skip connection or not

从图 8 中可以看出,跳跃连接的加入对修复效果的影响显著,这也验证了跳跃连接在本文模型中对于图像修复质量的重要性。

5.4.4 修复效果的对比分析

为了分析本文模型所做的图像修复的效果,将其与 Yeh 等使用的修复方法进行对比分析。

首先,使用峰值信号比(PSNR)定量分析文中方法修复的图像与 Yeh 等修复的图像以及原图的差距。PSNR 是一种评价图片质量的标准,虽然肉眼看到的越清晰的图片所对应

的 PSNR 不一定越高,但其依旧可以作为评价修复效果的一个参考标准。

从 LFW 数据集中随机选取 32 张图片进行测试,分别计算 3 种情况下对应图片的 PSNR 值。为了更好地比较 32 组 PSNR 值,图 9 中将结果直观呈现出来。从图 9 中可以看出,不管是本文方法修复的图像还是 Yeh 等修复的图像,都与原图具有相似的 PSNR 值,在某些图像上甚至有更大的 PSNR 值,这说明本文方法所修复的图像与原图较为接近。

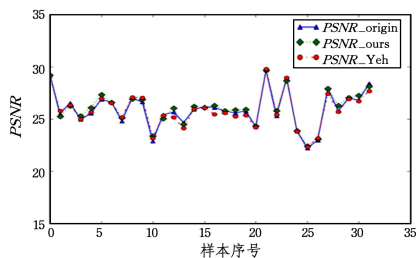


图 9 PSNR 值的对比

Fig. 9 Comparison of PSNR

此外,图 10 定性分析了对比效果。图中从左到右分别为未破损图像、破损图像、本文方法的修复结果、Yeh 等的修复结果。



图 10 中心破损修复的对比

Fig. 10 Comparison of center missing inpainting

从图 10 的第二行与第三行可以明显看出修复效果的差距,第四列中间破损区域的修复结果都出现了模糊情况。在人脸角度不大的情况下,两种模型都可以得到较好的修复结果,说明本文所使用的方法在某种程度上优于 Yeh 等的修复方法。

同样,为了测试本文提出的模型对破损区域较大以及破损区域形状不规则的图像的修复效果,分别取破损区域为整张图片的 1/2 以及随机选择破损点并选取 LFW 数据集中的其他图片作为测试本来进行实验,修复结果如图 11 和图 12 所示。图中从左到右分别为原图、破损图像、本文方法的修复结果、Yeh 等的修复结果。从图 11 中可以看出,相对于破损区域为中心较小区域,破损区域较大情况下的修复效果出现了修复模糊的情况,但是相对于 Yeh 等的修复区域,本文方法依然得到了较好的修复效果。同时,从图 12 中也可以看出,对于破损区域形状不规则的情况,本文所提方法也得

到了较好的修复效果,这也充分证明了本文模型对各种破损情况的通用性。



图 11 左半边破损修复的对比

Fig. 11 Comparison of left missing inpainting



图 12 随机破损修复的对比

Fig. 12 Comparison of random missing inpainting

图 13 给出交换训练集与测试集后对 CelebA 数据集中人脸的修复效果,从中可以看出本文所提模型的健壮性。

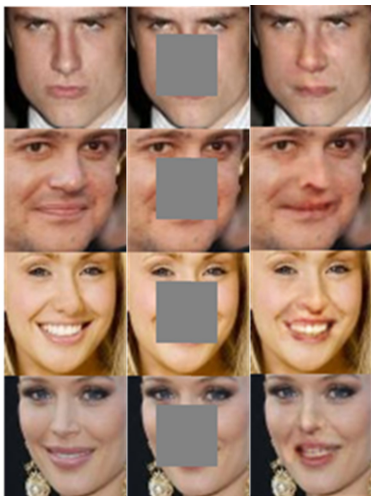


图 13 交换数据集的修复结果

Fig. 13 Inpainting results of changing datasets

修复工作,其对缺失区域的大小以及缺失区域的形状没有太多限制,同时能够实现较好的修复效果。实验结果证明了所提出的新对抗网络的生成模型在图像修复上具有重要作用。该方法虽然在缺失区域较大的情况下已经得到不错的效果,但是对于一些复杂结构的修复还是存在明显的模糊现象,如何处理这一问题有待于进一步研究。同时,类似于图像修复,文本修复也是一项非常有意义的工作,但是文本是离散的,在某种程度上不同于图像,因此处理离散文本的修复也将是下一步研究的重点。

参 考 文 献

- [1] CHAN T F, SHEN J. Mathematical Models for Local Nontexture Inpaintings [J]. Siam Journal on Applied Mathematics, 2002, 62(3): 1019-1043.
- [2] BARNES C, SHECHTMAN E, FINKELSTEIN A, et al. Patch-Match: a randomized correspondence algorithm for structural image editing [J]. Acm Transactions on Graphics, 2009, 28(3): 1-11.
- [3] PATHAK D, KRAHENBUHL P, DONAHUE J, et al. Context Encoders: Feature Learning by Inpainting [C] // Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas: IEEE Computer Society, 2016: 2536-2544.
- [4] GOODFELLOW I J, POUGET-ABADIE J, MIRZA M, et al. Generative Adversarial Networks [C] // Proceedings of the Conference on Advances in Neural Information Processing Systems. Montreal, Canada: Curran Associates, 2014: 2672-2680.
- [5] DENTON E, GROSS S, FERGUS R. Semi-Supervised Learning with Context-Conditional Generative Adversarial Networks [J]. arXiv preprint arXiv: 1611. 06430, 2016.
- [6] SIMONYAN K, ZISSERMAN A. Very Deep Convolutional Networks for Large-Scale Image Recognition [J]. arXiv preprint arXiv: 1409. 1556, 2014.
- [7] YE H R A, CHEN C, LIM T Y, et al. Semantic Image Inpainting with Deep Generative Models [C] // Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, HI, USA: IEEE Computer Society, 2017: 6882-6890.
- [8] RADFORD A, METZ L, CHINTALA S. Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks [J]. arXiv preprint arXiv: 1511. 06434, 2015.
- [9] ARJOVSKY M, CHINTALA S, BOTTOU L. Wasserstein gan [J]. arXiv preprint arXiv: 1701. 07875, 2017.
- [10] RONNEBERGER O, FISCHER P, BROX T. U-Net: Convolutional Networks for Biomedical Image Segmentation [C] // Proceedings of the 18th International Conference on Medical Image Computing and Computer-Assisted Intervention. Munich: Springer Cham, 2015: 234-241.
- [11] WANG K F, GOU C, DUAN Y J, et al. Generative adversarial networks: the state of the art and beyond [J]. Acta Automatica Sinica, 2017, 43(3): 321-332. (in Chinese)
王坤峰, 苟超, 段艳杰, 等. 生成式对抗网络 GAN 的研究进展与展望 [J]. 自动化学报, 2017, 43(3): 321-332.

- 徐晓丹,姚明海,刘华文,等. 基于 kNN 的多标签分类预处理方法[J]. 计算机科学,2015,42(5):106-108,131.
- [5] ZHANG Y F, HU G Q, ZHANG X F, et al. An algorithm study on tongue color recognition of patients[J]. Beijing Biomedical Engineering,2016,35(1):7-11. (in Chinese)
张艺凡,胡广芹,张新峰,等. 基于支持向量机的痤疮患者舌苔色识别算法研究[J]. 北京生物医学工程,2016,35(1):7-11.
- [6] WANG Y G, YANG J, ZHOU Y, et al. Tongue Image Color Recognition in Traditional Chinese Medicine[J]. Journal of Biomedical Engineering,2005,22(6):1116-1120. (in Chinese)
王永刚,杨杰,周越,等. 中医舌象颜色识别的研究[J]. 生物医学工程杂志,2005,22(6):1116-1120.
- [7] YANG Z H, ZHANG D P, LI N M. Kernel False-Colour Transformation and Line Extraction for Fissured Tongue Image[J]. Journal of Computer-Aided Design and Computer Graphics,2010,22(5):771-776. (in Chinese)
杨朝辉,张大鹏,李乃民. 裂纹舌图像的核假彩色变换及其纹线提取[J]. 计算机辅助设计与图形学学报,2010,22(5):771-776.
- [8] ZHUMU L M, LU P, XIA C M, et al. Research on Douglas-Peucker Method in Feature Extraction from 55 Cases of Tooth-Marked Tongue Image[J]. Chinese Archives of Traditional Chinese Medicine,2014,32(9):2138-2140. (in Chinese)
朱穆朗玛,陆萍,夏春明,等. 基于道格拉斯-普克法提取 55 例齿痕舌图像特征研究[J]. 中华中医药学刊,2014,32(9):2138-2140.
- [9] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single shot multibox detector [C] // Proceedings of European Conference on Computer Vision. Amsterdam: Springer International Publishing,2016:21-37.
- [10] JONATHAN L, EVAN S, TREVOR D. Fully convolutional networks for semantic segmentation[C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Boston: IEEE,2015:3431-3440.
- [11] SUN Y, WANG X G, TANG X O. Deep Learning Face Representation from Predicting 10,000 Classes[C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Columbus: IEEE,2014:1891-1898.
- [12] 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. Columbus: IEEE,2014:580-587.
- [13] GIRSHICK R. Fast R-CNN[C] // Proceedings of the IEEE Conference on International Conference on Computer Vision. Boston: IEEE,2015:1440-1448.
- [14] REN S Q, HE K M, GIRSHICK R, et al. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence,2017,39(6):1137-1149.
- [15] HOOCHANG S, ROTH H R, GAO M, et al. Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning [J]. IEEE Transactions on Medical Imaging,2016,35(5):1285-1298.
- [16] DONAHUE J, JIA Y Q, VINYALS O, et al. DeCAF: A Deep Convolutional Activation Feature for Generic Visual Recognition [C] // Proceedings of the 31st International Conference on Machine Learning. 2014: I-647-I-655.
- [17] 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 (CV PRW). 2014:512-519.
- [18] YOSINSKI J, CLUNE J, BENGIO Y, et al. How transferable are features in deep neural networks? [C] // Proceedings of the IEEE Conference on Neural Information Processing Systems. Montreal: MIT Press,2014:3320-3328.
- [19] DENG J, DONG W, SOCHER R, et al. ImageNet: A large-scale hierarchical image database[C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Miami: IEEE,2009:248-255.
- [20] KAN H X, ZHANG L Y, DONG C W. A Tongue Image Recognition Method Based on Type 2 Diabetes Traditional Chinese Medicine Syndrome Classification [J]. Chinese Journal of Biomedical Engineering,2016,35(6):658-664. (in Chinese)
阙红星,张璐瑶,董昌武. 一种 2 型糖尿病中医证型的舌图像识别方法[J]. 中国生物医学工程学报,2016,35(6):658-664.

(上接第 234 页)

- [12] RATLIFF L J, BURDEN S A, SASTRY S S. Characterization and computation of local Nash equilibria in continuous games [C] // Proceedings of the 2013 51st Annual Allerton Conference on Communication, Control, and Computing. Allerton: IEEE Computer Society,2013:917-924.
- [13] MIRZA M, OSINDERO S. Conditional Generative Adversarial Nets[J]. arXiv preprint arXiv:1411.1784,2014.
- [14] IOFFE S, SZEGEDY C. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift[C] // Proceedings of the 32nd International Conference on Machine Learning. Lille, France: International Machine Learning Society,2015:448-456.
- [15] TREFNY J, MATAS J. Extended Set of Local Binary Patterns for Rapid Object Detection[C] // Proceedings of the 15th Computer Vision Winter Workshop. Nove Hradky, Czech Republic: CVWW,2010:1-7.
- [16] VIOLA P, JONES M. Rapid object detection using a boosted cascade of simple features[C] // Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Kauai: IEEE Computer Society,2001:511-518.
- [17] GULRAJANI I, AHMED F, ARJOVSKY M, et al. Improved Training of Wasserstein GANs [C] // Proceedings of Annual Conference on Neural Information Processing Systems. Long Beach, CA, USA,2017:5769-5779.
- [18] KINGMA D P, BA J. Adam: A Method for Stochastic Optimization[J]. arXiv preprint arXiv:1412.6980,2014.
- [19] GANGNET M, BLAKE A. Poisson image editing [C] // Proceedings of the ACM SIGGRAPH. New York: Association for Computing Machinery,2003:313-318.