

# 基于密集卷积生成对抗网络的图像修复



孟丽莎 任坤 范春奇 黄泷

北京工业大学信息学部 北京 100124

数字社区教育部工程研究中心 北京 100124

城市轨道交通北京实验室 北京 100124

计算智能与智能系统北京市重点实验室 北京 100124

(menglisha508@163.com)

**摘要** 图像修复是一项利用缺损图像中已知信息对缺损区域信息进行估计修复的技术。针对大面积语义信息缺失的图像进行修复时,若训练数据集较小且图像背景相对复杂,则基于生成模型的修复结果常出现模糊、伪影和视觉相似度差等问题。针对上述问题,文中提出了一种基于密集卷积生成对抗网络的图像修复算法。该算法采用生成对抗网络作为图像修复的基本框架。首先,利用密集卷积块构建具有编解码结构的生成网络,不但加强了图像特征的提取,提高了图像修复能力,而且避免了深度增加引起的梯度消失问题。其次,在编码和解码结构之间引入跳跃连接,解决了网络层间信息传递丢失的问题。然后,在网络优化过程中,结合重建损失、对抗损失和 TV 损失来训练网络模型,增强了网络稳定性。最后,分别在 CelebA 和 Car 两个数据集上进行实验,所提算法的修复结果在视觉效果、峰值信噪比 PSNR 和结构相似度 SSIM 3 个方面均优于 3 种代表性图像修复算法,其有效性得到验证。

**关键词:** 生成对抗网络;图像修复;密集卷积块;跳跃连接;损失函数

中图分类号 TP391

## Dense Convolution Generative Adversarial Networks Based Image Inpainting

MENG Li-sha, REN Kun, FAN Chun-qi and HUANG Long

Faculty of Information Technology, Beijing University of Technology, Beijing 100124, China

Engineering Research Center of Digital Community, Ministry of Education, Beijing 100124, China

Beijing Laboratory for Urban Mass Transit, Beijing 100124, China

Beijing Key Laboratory of Computational Intelligence and Intelligent System, Beijing 100124, China

**Abstract** Image inpainting is one technique of reconstruction defect areas by inferring information from the known context of defect images. For semantic image inpainting of large areas, there are still many problems in inpainting algorithms based on generation models, such as blur, artifacts, and poor visual similarity, especially for the complex background images and small datasets. To solve this problem, an image inpainting algorithm based on dense convolution generative adversarial networks is proposed. The generated adversarial network is the basic framework. Firstly, dense convolutional blocks are used to enhance image feature extraction, improve image repair capability, and avoid the problem of gradient disappearance caused by the network depth increasing in the generator network. Secondly, skip connection between the encoding and decoding structures is involved to avoid information transmission lost problems between network layers. After that, a total loss function, composed of the reconstruction loss, adversarial loss and TV loss, is used to optimize the network and enhance network stability. Finally, the proposed algorithm is validated on the CelebA dataset and Car dataset respectively, compared with three typical image inpainting algorithms. The effectiveness of the algorithm is proved in visual perception, peak signal-to-noise ratio (PSNR) and structural similarity (SSIM).

**Keywords** Generative adversarial networks, Image inpainting, Dense convolutional block, Skip connection, Loss function

到稿日期:2019-07-01 返修日期:2019-11-01 本文已加入开放科学计划(OSID),请扫描上方二维码获取补充信息。

基金项目:国家自然科学基金(61803005,61640312,61763037,1305026);北京市自然科学基金(4192011,4172007);北京市教委科学研究计划(KM201310005006)

This work was supported by the National Natural Science Foundation of China (61803005,61640312,61763037,1305026), Beijing Natural Science Foundation (4192011,4172007) and Science Foundation from the Education Commission of Beijing (KM20130005006).

通信作者:任坤 (renkun@bjut.edu.cn)

## 1 引言

数字图像修复是利用缺损图像的已知区域对丢失信息进行修复,使得修复之后的图像最大化地还原原始图像并且符合人类的视觉感官。2000年,Bertalmio等<sup>[1]</sup>首次提出了数字图像修复的概念。2002年,数字图像修复工作<sup>[2]</sup>在 *Science News* 上发表,受到了广泛关注。作为计算机视觉的底层问题,图像修复技术在图像复原、缩放和超分辨、图像压缩编码、无线图像传输和虚拟现实等方面具有重要的应用价值。

现有的图像修复算法可分为传统算法<sup>[1-3]</sup>和基于深度学习的算法<sup>[5-6]</sup>。传统图像修复算法包括偏微分算法和纹理合成算法。偏微分算法以像素点为修复单位,对小区域缺损的修复效果较好,但是对复杂纹理缺失或大区域缺失的修复效果有待提高。基于样本块的纹理合成算法通过复制已知区域的相似样本信息实现对大区域缺失的修复,适用于自相似性较高的图像,但其计算复杂度较高,修复效率较低。传统图像修复算法由于仅利用原图中存留的周边信息,局限性明显,当已知区域中没有与缺损区域相似的信息时,修复就会出现偏差。

基于将图像修复这样的病态逆问题视为条件生成问题的思想,研究学者将深度卷积神经网络的生成模型自编码器(Autoencoder, AE)和生成对抗网络结合,在一定程度上解决了传统图像修复算法存在的问题。现阶段,基于生成对抗网络(Generative Adversarial Networks, GAN)<sup>[4]</sup>的算法已经成为图像修复的研究热点。Pathak等<sup>[5]</sup>提出的 Context Encoder(CE)算法是生成对抗网络应用于图像修复的开山之作。CE算法将编码-解码器和GAN相结合,适用于对大面积破损图像的修复,优于传统修复算法,但在修复区域周边会产生伪影。Yeh等<sup>[6]</sup>将图像语义分割技术与生成对抗网络相结合,提出了有效的语义图像修复算法。此算法中的生成网络和判别网络均采用5层的浅层网络。对于训练背景比较复杂的小数据集,该网络无法完全学习到特征信息,修复效果有待提高。Zhao等<sup>[7]</sup>利用 Wasserstein 生成对抗网络作为图像修复框架,以 Wasserstein 距离测量生成网络中生成图像和真实图像之间的差距,解决了图像修复网络训练阶段存在的训练梯度消失和模式崩溃的问题。之后,Altelin等<sup>[8]</sup>提出了改进修复网络的损失函数,将对抗损失与图像的结构熵相结合进行网络训练,从而保证了修复图像结构的连贯性。Yu等<sup>[9]</sup>提出了一种带有内容感知层的前馈生成网络,引入粗略到细化的网络结构,同时使用全局和局部判别网络来改善图像修复结果。

目前,基于生成对抗网络的图像修复算法仍存在许多有待解决的问题。首先,生成对抗网络在网络训练中存在不稳定和不易收敛的问题。其次,编码、解码中全连接层会导致解码过程不能准确地恢复相关信息,既不能保证结构连续性,也不能保证纹理一致性。

针对上述问题,本文提出一种基于密集卷积生成对抗网络的图像修复算法。通过在生成网络中引入密集卷积块,加强特征之间深度信息的传递,实现特征重用;取消解码解码中间的瓶颈区——全连接层,直接采用密集卷积块进行连接,保持输入图像中的原始空间信息;同时,在编码器和解码器间引

入跳跃连接(skip-connection)<sup>[10]</sup>来解决结构纹理一致性问题;将重建损失、判别损失、总变分(Total Variation, TV)损失作为优化生成模型的总损失,以此优化生成网络生成图像的细节。

## 2 相关工作

### 2.1 生成对抗网络

GAN自2014年由Goodfellow等<sup>[4]</sup>首次提出,便很快风靡整个人工智能领域。不仅学术界对GAN高度关注,谷歌、Facebook、Open AI等工业界也对其展开了大量的相关研究,其应用遍及图像和视频生成<sup>[11]</sup>、图像风格转换<sup>[12]</sup>、图像修复<sup>[5]</sup>、纹理合成<sup>[13]</sup>和图像超分辨<sup>[14]</sup>等。

GAN是一种深度生成模型,学习真实样本数据的概率分布规律,并利用学到的概率分布映射生成类似的内容。生成对抗网络由生成网络和判别网络构成,如图1所示。生成网络代表从随机分布到样本集中目标数据分布的映射关系,它以随机噪声作为源数据,通过训练不断地模拟逼近训练库中样本的数据分布,从而生成与样本集具有相同数据分布的看似真实的虚假图像。生成网络的输入是一个随机变量 $z$ ,输出是具有训练样本分布规律的尽可能以假乱真的生成数据 $G(z)$ 。判别网络是一个二分类器,它的输入是真实数据 $x$ 或者生成数据 $G(z)$ ,通过学习到的特征判断输入数据的真假,根据输出的概率值区分输入图像是真实图像还是生成图像。

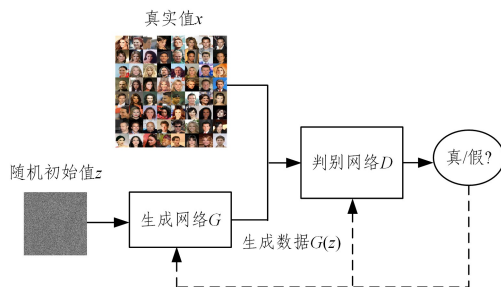


图1 GAN模型

Fig.1 Model of GAN

GAN的核心思想来源于纳什均衡。生成网络和判别网络相互对抗并迭代优化,使性能不断提高,直到判别网络无法判别数据来源时,网络达到最优。优化过程是一个极大极小问题,即:

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

其中, $G$ 表示生成网络, $D$ 表示判别网络, $E[\cdot]$ 表示数学期望, $V$ 表示目标函数, $x$ 表示样本, $p_r$ 表示真实样本的概率分布, $p_g$ 表示生成样本的概率分布。

相比于其他深度生成模型,如变分自编码器<sup>[16]</sup>和自回归模型<sup>[17]</sup>,GAN只进行反向传播计算,不需要复杂的马尔可夫链<sup>[18]</sup>,并且不依赖于假设的近似分布模型的好坏,可以渐进地逼近任意概率分布。如果判别网络训练良好,那么生成网络可以很好地学习到训练样本的分布。但在训练过程中,GAN需要达到纳什均衡,模型仍然存在训练难、稳定性差、模型易崩坍等问题。

### 2.2 DensNet 模型

网络深度是深度卷积神经网络的重要参量。尽管增加深度能够提升网络性能,但随着层级的加深,网络训练时的梯度消失问题也会愈加严重。密集网络 DenseNet<sup>[15]</sup> 的提出很好地解决了这个问题。DenseNet 是一种具有密集连接的卷积神经网络,其核心思想是跨层连接,网络中每一层都会以前面所有层作为输入,以保证层间最大程度的信息传输。

DenseNet 一般由多个密集卷积块和过渡层组成,其中最重要的部分为密集卷积块。密集卷积块以  $x_0$  作为输入,任何

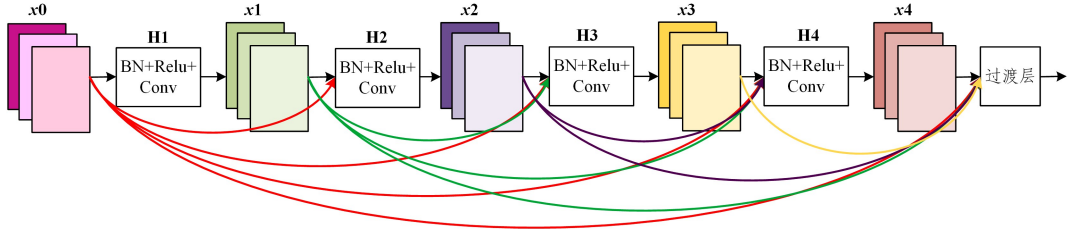


图 2 密集卷积块的结构

Fig. 2 Structure of dense block

### 3 基于密集卷积生成对抗网络的图像修复

基于上述分析,本文将 DenseNet 引入生成对抗中,利用密集卷积块构建生成网络进行图像修复,算法框架如图 3 所示。本文算法以残缺图像作为生成网络的输入,通过编码-解码器输出生成图像,然后通过判别网络进行判别,利用对抗学习算法优化生成网络;训练完成后,生成网络可输出与缺损图像相似的完整图像,将生成图像中对应缺损区域的信息填充到待修复的图像中,实现图像修复。

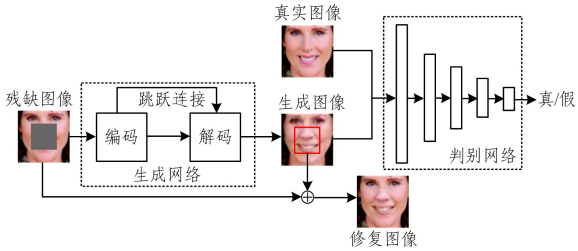


图 3 本文算法的框架图

Fig. 3 Frame diagram of proposed algorithm

两层之间都有直接连接,  $x_1, x_2, x_3, x_4$  的输入都是前面所有层输出的并集,单个块内层与层之间利用变换函数连接,变换函数包括 Batch Normalization(BN), ReLU 和 Convolution (Conv),其结构如图 2 所示。各密集卷积块之间利用过渡层连接。DenseNet 中的密集卷积块可以综合利用浅层特征,由于大量的特征被复用,网络中使用少量的卷积块就能学习到大量的特征,因此其对小数据集的训练具有良好的抗拟合性能,同时还能节省网络中的参数量和计算量,缓解梯度消失问题。

### 3.1 生成网络

生成网络由编码器、解码器和跳跃连接构成。由于 DenseNet 模型的密集卷积块中任何两层之间都是直接相连,有利于强化特征传播并且支持特征重用,在很大程度上解决了信息丢失问题,进而提高了修复区域内容的一致性。

编码器以带有中心矩形掩码的模拟残缺图像作为输入,逐步将图像空间维度减半。在编码器中,首先用步长为 2 的  $4 \times 4$  卷积核输入图像进行特征提取,再用 4 个密集卷积块、4 个卷积和池化层进行下采样,将待修复图像逐层抽象为隐层表征。解码器是编码器的反过程,包含 4 个密集卷积块和 4 个反卷积层,通过隐层表征生成预测图像,最后利用反卷积上采样得到原始图像大小的修复图像。

这里,取消传统编码-解码结构之间的全连接层,直接利用密集卷积块进行连接,以减少图像特征的丢失。同时,引入 U-Net 网络的编码-解码器之间的跳跃连接,在高层解码阶段直接利用对应分辨率的低层编码阶段的输出,为解码器补充编码器下采样损失掉的部分结构特征信息,强化生成网络的结构预测能力。

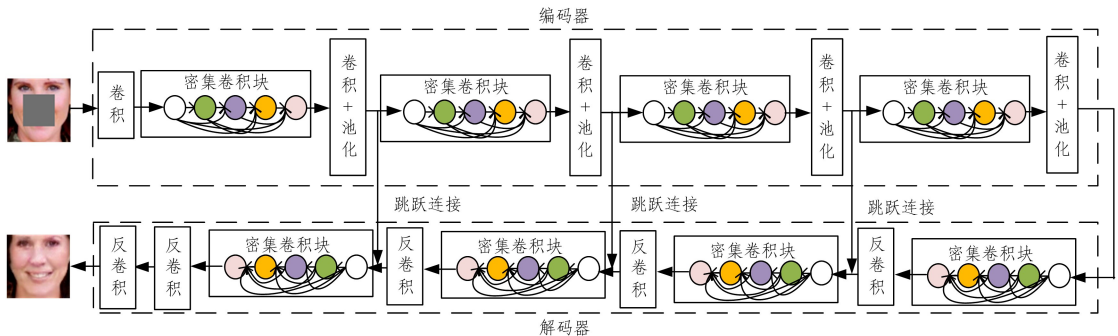


图 4 生成网络的结构

Fig. 4 Structure of generative network

### 3.2 判别网络

损失函数设计是优化网络的关键。判别网络是一个根据

输入图像的提取特征进行分类的二分类器。这里,选用 DC-GAN<sup>[19]</sup> 的判别网络,包括 5 个步长为 2、 $4 \times 4$  卷积核的卷积

层,且每个卷积层之后都使用 Batch Normalization 和 Leaky Relu。最后一层是一个二分类的全连接层。

### 3.3 损失函数

本文将重建损失、对抗损失和 TV 损失相结合作为训练网络的总损失,以提高网络稳定性和修复效果。联合损失函数的定义为:

$$L = \lambda_{\text{rec}} L_{\text{rec}} + \lambda_{\text{adv}} L_{\text{adv}} + \lambda_{\text{TV}} L_{\text{TV}} \quad (2)$$

其中, $\lambda_{\text{rec}}$ 、 $\lambda_{\text{adv}}$ 和 $\lambda_{\text{TV}}$ 分别表示对应损失项的权重。

重建损失<sup>[5]</sup> $L_{\text{rec}}$ 能够捕获缺失区域的整体结构和关于其上下文的一致性,作为主要损失函数的优化生成网络。这里,采用模拟缺失图像保留区域与生成图像相应区域之差的 $L_2$ 范数作为重建损失,即:

$$L_{\text{rec}} = \| (x - G(M \odot x)) \odot M \|_2 \quad (3)$$

其中, $x$ 表示输入的真实图像; $G$ 表示生成网络; $M$ 是二进制掩码,值为1的部分代表保留区域,值为0的部分代表缺失区域; $\odot$ 代表对应元素相乘。

对抗损失<sup>[4]</sup> $L_{\text{adv}}$ 是用于辅助生成网络拟合数据分布的,使生成图像看起来真实,并具有从分布中选择特定模式的作用。这里,对抗损失直接采用判别网络的目标函数,定义如下:

$$L_{\text{adv}} = \log(D(x)) + \log(1 - D(G(M \odot X))) \quad (4)$$

TV 损失<sup>[20]</sup> $L_{\text{TV}}$ 能够约束图像强度的平滑性,使输出图像比较平滑,用于解决生成图像的修复区域与残缺图像叠加时周边存在伪影的问题。其定义如下:

$$L_{\text{TV}} = \sum_{i,j} |G_{i+1,j}(M \odot X) - G_{i,j}(M \odot X)| + |G_{i,j+1}(M \odot X) - G_{i,j}(M \odot X)| \quad (5)$$

其中, $i$ 和 $j$ 表示像素点的坐标。

## 4 实验验证

### 4.1 实验参数设置

实验计算平台为主频 3.6 GHz, Intel Core i7-6850 CPU 和 NVIDIA 1080Ti GPU。软件环境为 Ubuntu 16.04, Python3 和 Tensorflow 1.3.0。

实验在不同数据集上迭代 50 个周期,并且根据硬件内存将 batchsize 的大小设置为 64。为了平衡不同损失函数的影响,借鉴文献<sup>[5]</sup>的参数设置,重建损失和对抗损失的权重分别取 $\lambda_{\text{rec}} = 0.999$ , $\lambda_{\text{adv}} = 0.001$ ,TV 损失权重为 $\lambda_{\text{TV}} = 1 \times 10^{-6}$ 。在生成对抗网络的训练阶段,优化网络采用学习率为 0.0002 的 Adam 算法,其中参数 $\beta_1$ 设置为 0.5, $\beta_2$ 设置为 0.9。

### 4.2 实验数据

本文在 CelebA 人脸数据集和 Car 数据集上对所提算法的有效性进行了实验验证。其中,CelebA 数据集是一个包括 202599 张名人面孔图像、存储容量为 1.8 GB 的数据集。这里选取 200000 张图像进行网络训练,使用其余 2599 张图像作为测试样本。实验中所有图像均利用 Openface 进行预处理,将人脸主体信息(如鼻子、眼睛、嘴巴)检测并截取出来,转换为 $64 \times 64$ 的样本。Car 数据集是一个收集了 16185 张车

辆图像的小数据集,存储容量为 28.1 MB。本文选取其中 14000 幅图像作为训练集,其余图像作为测试图像,图像被预处理为 $64 \times 64$ 的样本。

### 4.3 实验结果及分析

#### 1) 稳定性分析

将本文算法与无密集卷积结构的编码-解码算法进行对比。如图 4 所示,无密集卷积结构是指将生成网络中的所有密集卷积块去掉,编码-解码结构只包含卷积和池化层。生成网络的损失曲线如图 5 所示。无密集卷积算法在整个训练过程中的损失值在 0.025 附近变化,变化范围为 $-0.005 \sim 0.005$ ,震荡相对较大。而本文算法在训练过程中迭代 20000 次时,损失值下降到 0.013 附近,之后在 $-0.002 \sim 0.002$ 变化,比无密集卷积块的网络损失值减小一半,并且随着迭代次数的增加,网络基本处于收敛状态。可见,密集卷积网络能够提升网络的稳定性。

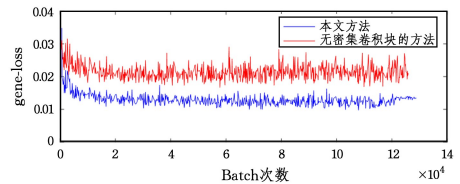


图 5 生成网络的训练收敛过程

Fig. 5 Convergence process of generative network training

#### 2) 定性分析

将本文算法与自动编码器、Context Encoder(CE)<sup>[5]</sup>和 Semantic Inpainting(SI)<sup>[6]</sup>这类代表性算法进行对比实验,验证其有效性。图 6 给出了各算法在 CelebA 数据集上的修复结果。图 6(c)为自动编码器修复图,可观察到修复区域较为模糊。图 6(d)为 CE 算法修复图,可以看出修复区域与原图存在明显的差异,一致性较差。图 6(e)为 SI 修复图,SI 算法后处理的泊松融合解决了伪影问题,但是修复的图像细节有所缺失,且修复的相似度也有待提高。图 6(f)为本文算法修复图,可以看出清晰度较好,无明显模糊和边缘效应,与原图像的相似度也明显优于对比算法。

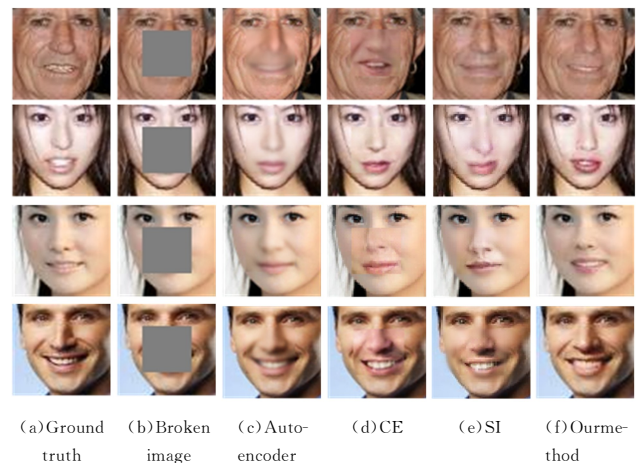


图 6 CelebA 数据集上的修复效果

Fig. 6 Inpainting results on CelebA dataset

Car 数据集的背景信息相对复杂且数据集较小,对修复算法具有一定的挑战。图 7 给出了 Car 数据集上应用不同算法修复的实验结果。图 7(c)为自动编码器的修复结果,图像特别模糊,无结构信息。图 7(d)为 CE 算法的修复结果,该算法能够学习数据集中的相关信息并且还还原缺损信息,但是连接处存在伪影和扭曲现象。从图 7(e)可以看到,SI 算法无法准确学习到周边图像信息,修复区域内车的结构出现了严重偏差。从图 7(f)可以看到,本文算法能更好地获取缺失区域的语义信息,结构较相似,纹理相似度也比较高。

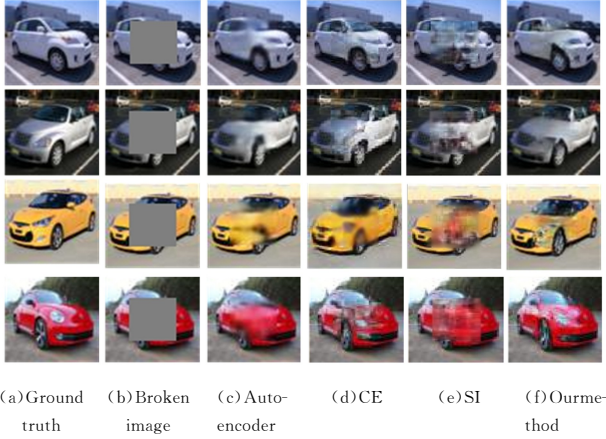


图 7 Car 数据集上的修复结果

Fig. 7 Inpainting results on Car dataset

### 3) 定量分析

本文采用平均  $L_2$  损失、峰值信噪比 (Peak Signal to Noise Ratio, PSNR) 和结构相似度 (Structural Similarity Index, SSIM) 3 个指标对上述修复算法的结果进行定量分析。

对于 CelebA 数据集测试图像,本文算法与自动编码器、CE 算法和 SI 算法的修复结果如表 1 所列。对比表 1 的实验结果可知,本文算法的平均  $L_2$  损失、PSNR 和 SSIM 都优于其他算法,说明在生成网络使用密集卷积和跳跃连接,能有效提升生成网络的预测能力。

表 1 不同算法的 PSNR 和 SSIM

Table 1 PSNR and SSIM of different methods

| Algorithm   | Mean $L_2$ loss/% | PSNR    | SSIM   |
|-------------|-------------------|---------|--------|
| Autoencoder | 7.89              | 16.2319 | 0.6513 |
| CE          | 3.25              | 21.1547 | 0.7751 |
| SI          | 2.96              | 18.5438 | 0.6629 |
| Our method  | 2.74              | 23.3119 | 0.7949 |

### 4) 损失函数的影响

针对不同损失函数对网络性能的影响进行对比实验,结果如图 8 所示。只使用重建损失  $L_{rec}$  时,图 8(c)的修复区域过于平滑;当采用重建损失  $L_{rec}$  和对抗损失  $L_{adv}$  时,图 8(d)的修复区域细节相对清晰,但修复区域仍有不同程度的结构扭曲现象;引入 TV 损失后,3 个损失函数同时作用,图 8(e)的修复结果明显较好,相似度良好。表 2 列出了不同损失函数的 PSNR 和 SSIM 定量分析结果。3 个损失函数相结合,不仅能获得相似度较高的修复效果,也能取得良好的数值分析结果。

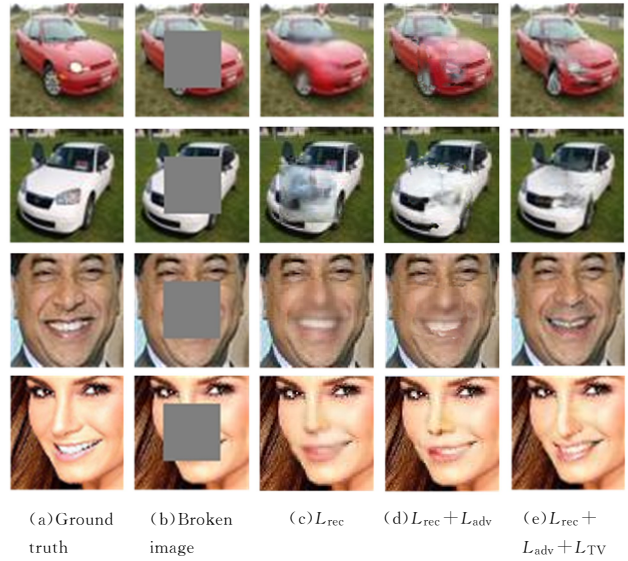


图 8 不同损失函数的修复结果

Fig. 8 Inpainting results of different loss functions

表 2 不同损失函数的 PSNR 和 SSIM

Table 2 PSNR and SSIM of different loss functions

| algorithm                                | Mean $L_2$ loss/% | PSNR    | SSIM   |
|--|-------------------|---------|--------|
| Our method, $L_{rec}$                    | 3.05              | 21.1464 | 0.7751 |
| Our method, $L_{rec} + L_{adv}$          | 2.95              | 21.6677 | 0.7991 |
| Our method, $L_{rec} + L_{adv} + L_{TV}$ | 2.74              | 22.0737 | 0.8097 |

**结束语** 受密集卷积网络和 U-Net 的启发,本文提出了一种基于密集卷积生成对抗网络的图像修复算法。

本算法利用密集卷积块构建编码-解码结构的生成网络,引入层级间的跳跃连接,不仅提高了网络的稳定性,也增强了生成模型的信息预测能力;引入 TV 损失,与重建损失和判别损失共同优化生成网络,增强了图像细节修复能力。本文在 CelebA 人脸数据集和小样本量 Car 数据集上分别进行了定性和定量的对比实验,验证了本文算法的有效性和优越性。本文算法虽然针对背景相对复杂、训练样本少的数据集可以得到较好的修复结果,但是由于训练过程是针对中心区域的缺损进行修复,因此对随机位置缺损的图像修复效果仍有待改进。多级生成网络和双判别网络将是进一步研究的方向。

### 参考文献

- [1] BERTALMIO M, SAPIRO G, CASELLES V, et al. Image inpainting[C]//Conference on Computer Graphics and Interactive Techniques. ACM Press/Addison-Wesley Publishing Company, 2000:417-424.
- [2] PETERSON I. Filling in Blanks: Automating the Inpainting of a Picture's Missing Pieces [J]. Science News, 2002, 161(19): 299-300.
- [3] CRIMINISI A, PEREZ P, TOYAMA K. Region Filling and Object Removal by Exemplar-based Image Inpainting [J]. IEEE Transactions on Image Processing, 2004, 13(9): 1200-1212.
- [4] GOODFELLOW I, POUGET-ABADIE J, MIRZA M, et al. Generative Adversarial Nets[C]//International Conference on Neural Information Processing Systems. 2014:2672-2680.

- [5] 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.
- [6] YEHR A, CHEN C, LIM T Y, et al. Semantic Image Inpainting with Deep Generative Models [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2017:6882-6890.
- [7] ZHAO G, LIU J, JIANG J, et al. A Deep Cascade of Neural Networks for Image Inpainting, Deblurring and Denoising [J]. Multimedia Tools and Applications, 2018, 77(22):29589-29604.
- [8] ALTINEL F, OZAY M, OKATANI T. Deep Structured Energy-Based Image Inpainting [C] // 2018 24th International Conference on Pattern Recognition (ICPR), Beijing, 2018:423-428.
- [9] YU J, LIN Z, YANG J, et al. Generative Image Inpainting with Contextual Attention [C] // 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2018:5505-5514.
- [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. Cham: Springer, 2015:234-241.
- [11] DOSOVIKIY A, BROX T. Generating Images with Perceptual Similarity Metrics based on Deep Networks [C] // Advances in Neural Information Processing Systems. 2016:658-666.
- [12] ISOLA P, ZHU J Y, ZHOU T, et al. Image-to-Image Translation with Conditional Adversarial Networks [C] // Computer Vision and Pattern Recognition. 2016:5967-5976.
- [13] LI C, WAND M. Precomputed Real-Time Texture Synthesis with Markovian Generative Adversarial Networks [C] // European Conference on Computer Vision. 2016:702-716.
- [14] LEDIG C, THEIS L, HUSZAR F, CABALLERO J, et al. Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network [C] // 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, HI, 2017:105-114.
- [15] HUANG G, LIU Z, MAATEN L V D, et al. Densely Connected Convolutional Networks [C] // IEEE Conference on Computer Vision and Pattern Recognition. IEEE Computer Society, 2017:2261-2269.
- [16] KINGMA D P, WELING M. Auto-Encoding Variational Bayes [C] // Proceedings of the International Conference on Learning Representations (ICLR). 2014.
- [17] AKAIKE H. Fitting Autoregressive Models for Prediction [J]. Annals of the Institute of Statistical Mathematics, 1969, 21(1):243-247.
- [18] HINTON G E, OSINDEROS S, TEH Y W. A Fast Learning Algorithm for Deep Belief Nets [J]. Neural Computation, 2006, 18(7):1527-1554.
- [19] RADFORD A, METZ L, CHINTALA S. Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks [J]. arXiv:1511.06434, 2015.
- [20] JAVANMARDI M, SAJJADI M, LIU T, et al. Unsupervised Total Variation Loss for Semi-supervised Deep Learning of Semantic Segmentation [J]. arXiv:1605.01368v3, 2016.



**MENG Li-sha**, born in 1992, postgraduate. Her main research interests include deep learning and computer vision.



**REN Kun**, born in 1973, Ph.D, lecturer. Her main research interests include deep learning and computer vision.