

基于 GAN 先验的图像补全方法^①

卢世杰, 郝文宁, 余晓晗, 于 坤

(中国人民解放军陆军工程大学 指挥控制工程学院, 南京 210007)

通信作者: 郝文宁, E-mail: hwnbox@foxmail.com



摘 要: 设计和利用良好的图像先验知识是解决图像补全问题的重要方式. 生成对抗网络 (GAN) 作为一种优秀的生成式模型, 其生成器可以从大型图像数据集中学习到丰富的图像语义信息, 将预训练 GAN 模型作为图像先验是一种好的选择. 为了利用预训练 GAN 模型更好地解决图像补全问题, 本文在使用多个隐变量的基础上, 在预训练生成器中间层同时对通道和特征图添加自适应权重, 并在训练过程中微调生成器参数. 最后通过图像重建和图像补全实验, 定性和定量分析相结合, 证实了本文提出的方法可以有效地挖掘预训练模型的先验知识, 进而高质量地完成图像补全任务.

关键词: 图像先验; 生成对抗网络; 预训练模型; 图像补全; 深度学习

引用格式: 卢世杰, 郝文宁, 余晓晗, 于坤. 基于 GAN 先验的图像补全方法. 计算机系统应用, 2022, 31(10):397-403. <http://www.c-s-a.org.cn/1003-3254/8705.html>

Image Inpainting Method Based on GAN Prior

LU Shi-Jie, HAO Wen-Ning, YU Xiao-Han, YU Kun

(College of Command & Control Systems, Army Engineering University of PLA, Nanjing 210007, China)

Abstract: Designing and utilizing good image prior knowledge is an important way to enable image inpainting. A generative adversarial network (GAN) is an excellent generative model, and its generator can learn rich image semantic information from large datasets. Thus, it is a good choice to use a pre-trained GAN model as an image prior. Making use of multiple hidden variables, this study adds adaptive weights to the channels and feature maps at the same time in the middle layer of the pre-trained generator and fine-tunes generator parameters in the training process. In this way, the pre-trained GAN model can be used for better image inpainting. Finally, through the contrast experiment of image reconstruction and image inpainting and the combination of qualitative and quantitative analysis, the proposed method is proved effective to mine the prior knowledge of the pre-trained model, thus finishing the task of image inpainting with high quality.

Key words: image prior; generative adversarial network; pre-trained model; image inpainting; deep learning

图像补全是一项极为重要的底层视觉任务, 其目的是根据图像中已有信息还原图像中的缺失部分, 在生活中的各个领域都具有广泛的应用价值. 然而, 图像补全具有不确定性, 一张待修复的图像往往会有多种修复结果, 通常需要添加图像先验知识来约束求解范围. 因此, 设计好的先验知识是完成图像补全任务的关

键. 图像先验描述了自然图像的统计信息, 它在计算机视觉任务中得到了广泛的应用. 传统的图像先验包括马尔可夫随机场^[1-3]、暗通道先验^[4,5]和总变差正则化等. 随着深度学习的发展, 文献 [6] 提出将卷积神经网络作为一种先验信息, 但受限于模型结构和参数随机初始化等问题, 卷积神经网络无法拟合颜色、纹理等

① 收稿时间: 2022-01-04; 修改时间: 2022-01-29; 采用时间: 2022-02-15; csa 在线出版时间: 2022-06-28

更高级的信息. 近年来, 有研究人员发现预训练 GAN 模型中包含了丰富的图像先验知识, 文献 [7] 提出了 mGANprior, 通过使用多个隐变量, 并在生成器中间层加入通道注意力, 实现了较高质量的图像重建, 还将预训练模型应用于图像去噪、超分辨率等图像复原任务, 但由于生成器参数在训练过程中没有被更新, 且在特征融合时每个隐变量对应的特征图被赋予了相同的权重, 导致重建后的图像丢失了大量的细节纹理信息, 并且该模型在图像补全任务中的表现并不理想.

随着注意力机制在计算机视觉领域被越来越多的探索, 有研究提出了卷积块注意力模型 (CBAM)^[8], 该模型同时在通道和空间上添加注意力机制, 即图片的特征体现在每个通道的每个像素上. CBAM 可以嵌入到目前大部分主流网络中, 在不显著增加计算量和参数数量的同时提升网络模型的特征提取能力. 受此启发, 本文综合利用预训练 GAN 模型的先验知识以及注意力机制在提升模型性能方面的优势, 从而高质量地完成图像补全任务.

综上, 本文的主要工作包括以下几个方面:

(1) 本文在使用多个隐变量的基础上, 在预训练生成器中间层同时对通道和特征图添加自适应权重, 并在训练过程中微调生成器参数.

(2) 使用本文模型进行图像重建实验, 并与其他图像重建方法进行比较, 证实了本文模型可以更好地恢复图像的细节纹理信息.

(3) 将本文模型应用于图像补全任务, 并与其他图像补全方法进行比较, 证实了本文方法可以有效地利用预训练 GAN 模型中的先验知识, 进而更好地完成图像补全任务.

1 相关研究

1.1 生成对抗网络

生成对抗网络 (GAN) 包含两个相互作用的神经网络: 生成器和鉴别器. 生成器的目标是合成类似真实数据的伪数据, 鉴别器的目标是区分真实数据和伪数据, 通过对抗训练使生成器学习从潜在分布到真实数据分布的映射, 在学习了这种非线性映射之后, GAN 就能够从隐空间随机采样的向量中生成逼真的图像. 因其在产生高质量图片方面具备优势, GAN 被应用于多个图像处理任务, 如图像去噪^[9,10], 图像补全^[11,12], 超分辨率^[13,14], 人脸编辑^[15,16]等.

1.2 模型反演

GAN 本质是将采样的噪声向量输入生成器转变为图片, 并不是直接对真实图像进行处理. 若想将预训练 GAN 模型应用于真实图像的处理, 首先需要将图片映射为隐变量, 而后再将隐变量传入预训练生成器, 从而实现图像重建, 该过程称为 GAN 反演. 形式上, 反演问题描述如式 (1):

$$z^* = \operatorname{argmin} L(G(z), x) \quad (1)$$

其中, x 是目标图像, G 为 GAN 生成器, L 表示损失函数. 目前有两种主流方法来实现 GAN 反演, 一种是训练额外的编码器来学习从图像空间到隐空间的映射^[17,18], 另一种是基于优化的方法^[19], 即先将随机初始化的隐变量送入生成器, 再通过反向传播算法最小化重构误差, 进而优化隐变量.

1.3 深度模型先验

图像补全属于图像复原任务的一种, 图像复原任务的目的是将降质图像 (降低分辨率, 加入遮挡等) 恢复为原始图像, 优化公式通常如式 (2) 所示:

$$x^* = \min_x E(x; x_0) + R(x) \quad (2)$$

其中, x 是降质后的图像, x_0 是网络模型生成的图像, $E(x; x_0)$ 是针对具体任务的优化函数, $R(x)$ 是正则化项, 通常表示图像先验信息. 在早期的基于深度学习的图像复原方法中, 通常是通过大量的数据样例的训练学习到 $R(x)$.

文献 [6] 认为随机初始化的卷积神经网络 (CNN) 可以捕获低层级的图像先验信息, 如式 (3)、式 (4) 所示, f 是任意的 CNN 模型, θ 是该模型的权重参数, z 是随机初始化的噪声向量, 通过优化 CNN 模型的参数来获取图像先验, 一旦得到了最优的参数, 把向量 z 输入模型即可得到最优的图像 x , 即使只在单张降质图像上进行反复迭代, 也同样能学习到图像的先验信息, 进而完成图像复原.

$$\theta^* = \operatorname{argmin} E(f_\theta(z); x_0) \quad (3)$$

$$x^* = f_{\theta^*}(z) \quad (4)$$

CNN 模型结构虽然可以获取到图像的底层统计特性, 但由于其模型参数由随机初始化得到, 因此无法拟合到图像更高级的信息, 如颜色, 纹理等. 生成对抗网络 (GAN) 作为一种很好的自然图像流形逼近器, 具有拟合自然图像分布的能力, 通过从大型图像数据集

学习,其生成器部分可以捕捉到丰富的高层图像语义信息,包括颜色、空间相干性、纹理等高级概念. SinGAN^[20]利用目标图像的图像补丁(patch)来学习图像内部统计信息,来实现各种图像编辑或恢复效果.文献^[21]表明预训练的GAN模型捕获了丰富自然图像的统计数据作为先验知识.文献[7]利用多个隐变量实现GAN反演,并利用GAN捕获的先验知识完成多种图像处理任务.

2 模型架构

本文使用预训练PGGAN的生成器作为基础模型架构,PGGAN(progressive GAN)^[22]采用一种渐进式的训练策略,在训练过程中逐渐向生成器和判别器网络中添加新的层,以此对越来越细粒度的细节进行建模,提高了训练速度和稳定性.文献[7]提出利用多个隐变量重建图像,每个隐变量负责重建图像的一块区域,并在生成器中间层加入通道注意力.文献[8]提出了一种轻量级的注意力模块,将注意力同时用到了通道和空间两个维度上,在ResNet等经典结构上展现了强大的

特征提取能力.

2.1 模型结构

本文在使用多个隐变量的基础上,提出了在预训练PGGAN模型中同时对通道和特征图添加自适应权重的方法,模型结构如图1所示,将预训练生成器 G 在第 k 层切分为两部分,即 G_1 和 G_2 , T_n 表示 N 个隐变量经过生成器 G_1 后形成的 N 个特征图. α_n 是一个 M 维向量,表示特征图中对应通道的重要性, M 表示生成器第 k 层输出通道的个数, β 是一个 N 维向量,表示 N 个中间层特征图在合成图像过程中的重要性权重,通过给中间层特征图赋予自适应权重,可以充分发挥不同隐变量的作用.图像重建过程分为以下两个阶段,式(5)表示 N 个隐变量经过第1个生成器生成 N 个中间特征图,式(6)表示经过特征加权融合后传入第2个生成器生成最终的图片.

$$T'_n = G_1(z_n) \odot \alpha_n \quad (5)$$

$$x^{gt} = G_2\left(\sum_{n=1}^N T'_n \times \beta_n\right) \quad (6)$$

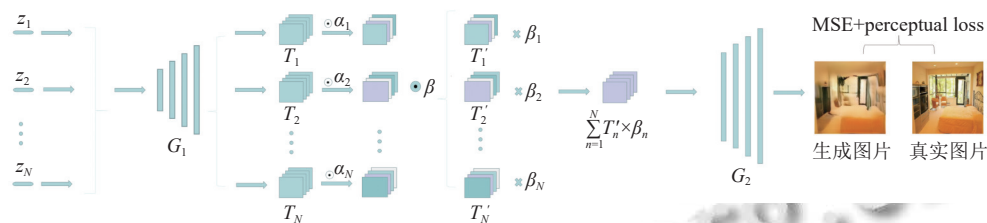


图1 图像重建模型

2.2 优化目标

利用预训练GAN生成器重建一张目标图片,此前基于优化的反演方法是固定预训练生成器,仅优化隐变量.然而,这种优化方式难以完成复杂纹理图片的重建.此外,目标图片与训练集图片的分布可能存在较大的差异,GAN生成器不可避免地受到训练数据分布的限制,从而难以重建训练集以外的图片.为了更精确的重建目标图像,本文在优化自适应权重 α 、 β 及隐变量 z 的同时,微调生成器参数 θ ,以减小图像近似流型和真实流型之间的差距,通过这种过参数化的方式提高了图像重建质量,待优化参数及优化过程如式(7):

$$\{z_n^*\}_{n=1}^N, \beta^*, \{\alpha_n^*\}_{n=1}^N, \theta^* = \underset{\{z_n\}_{n=1}^N, \beta, \{\alpha_n\}_{n=1}^N, \theta}{\operatorname{argmin}} L(x^{gt}, x) \quad (7)$$

2.3 损失函数

本文使用MSE和感知损失作为距离度量,二者分别表示像素级重建误差以及把两幅图像分别输入到VGG等深度网络中提取的感知特征^[23]之间的距离,如式(8)所示,其中 ϕ 表示感知特征提取器,本文使用VGG16作为特征提取器.

$$L(x_1, x_2) = \|x_1 - x_2\|_2^2 + \|\phi(x_1), \phi(x_2)\|_1 \quad (8)$$

3 实验分析

本文先使用LSUN数据集上预训练的PGGAN模型进行图像重建实验,在此基础上利用预训练GAN模型中的先验知识完成图像补全任务.

为了定量评估图像的重建质量,本文使用峰值信

噪比 (PSNR) 和结构相似性 (SSIM) 两种图像评价指标。PSNR 是目前应用最为广泛的一种图像评价指标, 它通过均方差 (MSE) 进行定义。两个 $m \times n$ 单色图像 I 和 K , 如果一个为另外一个的噪声近似, 那么它们的均方差定义如式 (9) 所示, 峰值信噪比的定义如式 (10) 所示, MAX_I 表示图像中可能出现的最大像素值, 它通过计算对应像素点间的误差来衡量目标图片与重建图片之间的相似度, 数值越大表示失真越小, SSIM 是一种全参考的图像质量评价指标, 它分别从亮度、对比度、结构 3 方面度量图像相似性, 如式 (10) 所示, μ_x, μ_y 分别表示图像 x 和 y 的均值, σ_x, σ_y 分别表示图像 x 和 y 的方差, σ_{xy} 表示图像 x 和 y 的协方差。SSIM 的取值范围为 $[0, 1]$, 值越大表示补全的图像与原图越相似。

$$MSE = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} \|I(i, j) - K(i, j)\|^2 \quad (9)$$

$$PSNR = 10 \times \log_{10} \left(\frac{MAX_I^2}{MSE} \right) = 20 \times \log_{10} \left(\frac{MAX_I}{\sqrt{MSE}} \right) \quad (10)$$

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + c_1)(\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \quad (11)$$

3.1 图像重建

本文分别使用 LSUN 卧室、LSUN 教堂数据集上预训练的 PGGAN 模型上进行实验, 对于每个模型, 随机选取对应数据集中 100 张图像进行重建, 最终结果取平均。将本文方法与两种基于 GAN 反演的图像重建方法进行比较, mGANprior 模型仅在通道层面添加自适应权重, 且在优化过程中固定生成器参数。idinvert 采用基于学习的反演方法, 使用编码器初始化隐变量。表 1 显示了定量比较结果, 本文在 LSUN 卧室、LSUN 教堂数据集上预训练的两个 PGGAN 模型上, 无论是像素级 (PSNR) 还是全参考的评价指标 (SSIM), 都优于其他两种方法。图 2 显示了定性比较结果, 可看出传统多隐变量先验方法 mGANprior 无法恢复目标图像的细节纹理信息, 这是因为生成器参数在训练过程中没有被更新, 且在特征融合时每个隐变量对应的特征图被赋予了相同的权重, 而事实上每个隐变量在图像重建中的贡献程度应是不同的。训练域导向编码器的方法 idinvert 将图像编码为隐变量, 再送入生成器实现图像重建, 该方法重建的图像虽然在人眼感知层面有着较好的效果, 但在布局和结构上与目标图像差别较大。相

比之下, 使用本文方法重建的图像, 在保持原图像布局和结构的同时, 增添了更多的细节信息, 如卧室图片中的抽屉, 枕头, 教堂图片中建筑物的纹理以及路面的标线等。

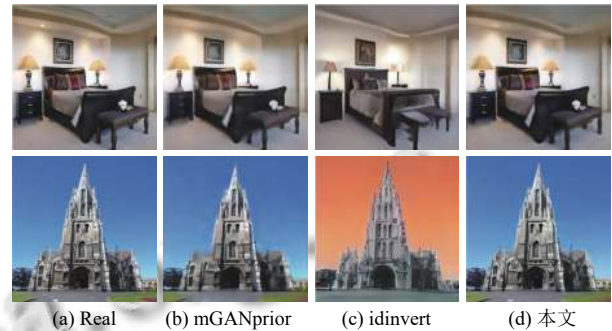


图 2 不同图像重建方法的定性比较

表 1 不同图像重建方法的定量比较

方法	PSNR		SSIM	
	卧室	教堂	卧室	教堂
mGANprior	25.31	22.17	0.8396	0.7914
idinvert	16.38	7.41	0.5361	0.4355
本文	26.06	22.65	0.8431	0.8018

3.2 图像补全

高质量的图像重建可以使得预训练 GAN 模型作为一种先验知识应用于图像处理任务, 本文利用预训练 GAN 模型的先验知识完成图像补全任务。图像补全的目标函数如式 (12) 所示, 已知完整图像 x , 模型生成的图像 x^{gt} 与二值掩模 m , \otimes 表示两个矩阵对应元素相乘, 在重建过程中只重建未加入遮挡的部分:

$$L = L(x^{gt} \otimes m, x \otimes m) \quad (12)$$

预训练 GAN 模型可以利用自身的先验知识自动填充缺失部分的像素。图 3 展现了更为直观的补全过程, 模型采用与图像重建实验相同的网络结构, 不同之处在于计算损失的方式, 首先给目标图像添加随机遮挡, 再给生成器生成的图片添加相同的遮挡, 随后计算二者之间的误差, 最后通过梯度下降算法更新待优化参数, 经过一定次数的迭代, 生成器便可生成完整无缺失的图像。

本文使用 30 个隐变量, 并在生成器第 8 层进行特征加权融合, 将本文方法与两种图像补全的方法作比较, 随机选取 LSUN 卧室和 LSUN 教堂数据集中各 100 张图像进行图像补全实验, 表 2 显示了定量比较结

果,可看出本文方法优于其他两种方法.图4显示了定性比较结果,多隐变量先验方法 mGANprior 在生成器第4层进行特征融合,虽然可以填补缺失部分的像素,却使得整张图像变得模糊,将 CNN 结构作为先验知

识 DIP 的补全效果也并不理想,补全后的区域与背景的语义信息不一致.相比之下,本文方法更有效的利用了预训练 GAN 模型的先验知识,实现了较高质量的图像补全.

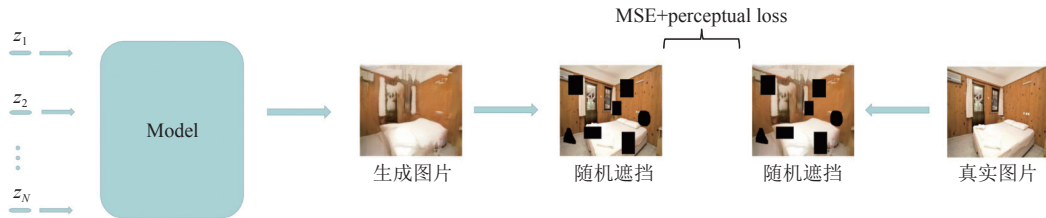


图3 图像补全模型

表2 不同图像补全方法的定量比较

方法	PSNR		SSIM	
	卧室	教堂	卧室	教堂
mGANprior	18.36	19.55	0.5662	0.6042
DIP	19.01	19.17	0.6740	0.7048
本文	21.85	22.21	0.7231	0.7521

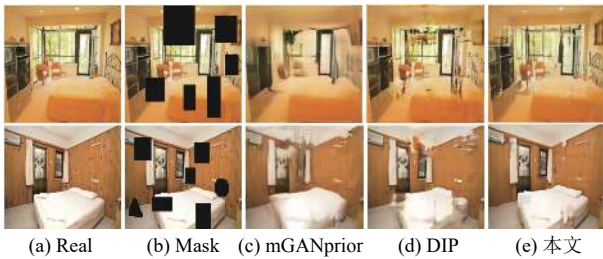


图4 不同图像补全方法的定性比较

3.3 消融实验

本文方法在图像补全任务中展现出一定的优越性,主要得益于3个部分:预训练的GAN模型,针对目标图像的生成器微调,对中间层特征图添加自适应权重.为了体现不同部分的作用,本文进行了消融实验.首先将随机初始化权重和使用预训练权重两种方式进行对比,以此证明预训练模型中的先验知识对图像补全效果的影响.结果如图5所示,实验结果表明,权重随机初始化的GAN模型难以捕捉丰富的图像语义知识,只能重建出目标图像的轮廓和结构,而加入预训练权重后的生成器则可以作为一种先验知识引导图像重建,生成的图像包含丰富的细节纹理信息.

为了探究模型微调对图像补全效果的影响,本文对生成器微调和未微调两种方式进行了对比实验,如图6所示,当利用卧室数据集上预训练的GAN模型完

成房屋图像的补全时,如果不微调生成器参数,生成的图片会存在伪影,原因在于待补全图片与训练集图片的分布存在较大的差异,而本文在优化隐变量和自适应权重的同时,微调了生成器参数,在一定程度上减小了生成器模拟分布和真实图像分布之间的差距,增强了图像的真实感.



图5 权重不同初始化方式效果对比



图6 不同优化策略的效果对比

本文同时给通道和特征图赋予自适应权重,为了体现给特征图赋权重的作用,本文进行了对比实验,如图7所示,实验结果表明,仅给通道加权的模型在细节

补全方面效果不太理想,如图中的窗子出现了模糊和截断,这是因为在特征融合时,每个隐变量生成的中间层特征图被赋予了相同的权重,本文给特征图加入自适应权重后,效果有一定的提升,原因在于每个隐变量在图像补全任务中的贡献程度是不同的,通过给中间层特征图添加自适应权重,可以充分发挥每个隐变量的作用,更好地利用已知区域的信息填补未知区域。



图7 不同加权方式的效果对比

4 总结与展望

为了有效地利用预训练 GAN 模型中的先验知识完成图像补全任务,本文在使用多个隐变量的基础上,在生成器中间层同时对通道和特征图添加自适应权重,并在训练过程中微调生成器参数.实验结果表明,本文方法可以更高质量地完成图像补全任务,优于其他深度模型先验方法.但现有工作仍存在一些不足,有待进一步研究.

(1) 利用预训练 GAN 模型的先验知识完成图像补全任务,虽然不需要额外的训练数据,但是需要每张图片在线自适应,训练模型的时间转移到了优化参数上面,这就导致优化过程过于复杂缓慢,如何缩短优化时间仍是一个待解决的问题.

(2) 本文方法只能针对降质类型确定的图像,对于降质类型未知的图像,利用模型先验知识完成图像复原较为困难,仍值得进一步探索.

参考文献

- Zhu SC, Mumford D. Prior learning and Gibbs reaction-diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1997, 19(11): 1236–1250. [doi: 10.1109/34.632983]
- Geman S, Geman D. Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1984, PAMI-6(6): 721–741. [doi: 10.1109/TPAMI.1984.4767596]
- He KM, Sun J, Tang XO. Single image haze removal using dark channel prior. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2011, 33(12): 2341–2353. [doi: 10.1109/TPAMI.2010.168]
- Rudin LI, Osher S, Fatemi E. Nonlinear total variation based noise removal algorithms. *Physica D: Nonlinear Phenomena*, 1992, 60(1–4): 259–268. [doi: 10.1016/0167-2789(92)90242-F]
- Zhang K, Zuo WM, Gu SH, *et al.* Learning deep CNN denoiser prior for image restoration. *Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition*. Honolulu: IEEE, 2017. 2808–2817.
- Lempitsky V, Vedaldi A, Ulyanov D. Deep image prior. *Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Salt Lake City: IEEE, 2018. 9446–9454.
- Gu JJ, Shen YJ, Zhou BL. Image processing using multi-code GAN prior. *Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Seattle: IEEE, 2020. 3009–3018.
- Woo S, Park J, Lee JY, *et al.* CBAM: Convolutional block attention module. *Proceedings of the 15th European Conference on Computer Vision*. Munich: Springer, 2018. 3–19.
- Chen JW, Chen JW, Chao HY, *et al.* Image blind denoising with generative adversarial network based noise modeling. *Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Salt Lake City: IEEE, 2018. 3155–3164.
- Kim DW, Chung JR, Jung SW. GRDN: Grouped residual dense network for real image denoising and GAN-based real-world noise modeling. *Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*. Long Beach: IEEE, 2019. 2086–2094.
- Yeh RA, Chen C, Lim TY, *et al.* Semantic image inpainting with deep generative models. *Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition*. Honolulu: IEEE, 2017. 6882–6890.
- Yu JH, Lin Z, Yang JM, *et al.* Generative image inpainting with contextual attention. *Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Salt Lake City: IEEE, 2018. 5505–5514.
- Ledig C, Theis L, Huszár F, *et al.* Photo-realistic single image super-resolution using a generative adversarial network. *Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition*. Honolulu: IEEE,

2017. 105–114.
- 14 Wang XT, Yu K, Wu SX, *et al.* ESRGAN: Enhanced super-resolution generative adversarial networks. Proceedings of Computer Vision. Munich: Springer, 2018. 63–79.
- 15 Lample G, Zeghidour N, Usunier N, *et al.* Fader networks: Manipulating images by sliding attributes. Proceedings of Neural Information Processing Systems 30. Long Beach: NIPS, 2017. 5967–5976.
- 16 Shen YJ, Luo P, Yan JJ, *et al.* FaceID-GAN: Learning a symmetry three-player GAN for identity-preserving face synthesis. Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018. 821–830.
- 17 Perarnau G, van de Weijer J, Raducanu B, *et al.* Invertible conditional GANs for image editing. arXiv: 1611.06355, 2016.
- 18 Zhu JY, Krähenbühl P, Shechtman E, *et al.* Generative visual manipulation on the natural image manifold. Proceedings of the 14th European Conference on Computer Vision. Amsterdam: Springer, 2016. 597–613.
- 19 Lipton ZC, Tripathi S. Precise recovery of latent vectors from generative adversarial networks. Proceedings of the 5th International Conference on Learning Representations. Toulon: ICLR, 2017.
- 20 Shaham RT, Dekel T, Michaeli T. SinGAN: Learning a generative model from a single natural image. Proceedings of 2019 IEEE/CVF International Conference on Computer Vision. Seoul: IEEE, 2019. 4569–4579.
- 21 Pan XG, Zhan XH, Dai B, *et al.* Exploiting deep generative prior for versatile image restoration and manipulation. Proceedings of the 16th European Conference on Computer Vision. Glasgow: Springer, 2020. 262–277.
- 22 Karras T, Aila T, Laine S, *et al.* Progressive growing of GANs for improved quality, stability, and variation. Proceedings of the 6th International Conference on Learning Representations. Vancouver: ICLR, 2018.
- 23 Johnson J, Alahi A, Li FF. Perceptual losses for real-time style transfer and super-resolution. Proceedings of the 14th European Conference on Computer Vision. Amsterdam: Springer, 2016. 694–711.

(校对责编: 牛欣悦)