

基于生成对抗网络的图像修复算法^①

黄 健, 韩俊楠

(西安科技大学 通信与信息工程学院, 西安 710600)
通信作者: 韩俊楠, E-mail: 348021464@qq.com



摘 要: 为解决当前基于生成对抗网络的深度学习网络模型在面对较复杂的特征时存在伪影、纹理细节退化等现象, 造成视觉上的欠缺问题, 提出了连贯语义注意力机制与生成对抗网络相结合的图像修复改进算法. 首先, 生成器使用两阶段修复方法, 用门控卷积替代生成对抗网络的普通卷积, 引入残差块解决梯度消失问题, 同时引入连贯语义注意力机制提升生成器对图像中重要信息和结构的关注度; 其次, 判别器使用马尔可夫判别器, 强化网络的判别效果, 将生成器输出结果进行反卷积操作得到最终修复后的图片. 通过修复结果以及图像质量评价指标与基线算法进行对比, 实验结果表明, 该算法对缺失部分进行了更好地预测, 修复效果有了更好的提升.

关键词: 图像修复; 生成对抗网络; 门控卷积; 连贯语义注意

引用格式: 黄健, 韩俊楠. 基于生成对抗网络的图像修复算法. 计算机系统应用, 2023, 32(10): 215-221. <http://www.c-s-a.org.cn/1003-3254/9257.html>

Image Inpainting Algorithm Based on Generative Adversarial Network

HUANG Jian, HAN Jun-Nan

(College of Communication and Information Technology, Xi'an University of Science and Technology, Xi'an 710600, China)

Abstract: The current depth learning network models based on generative adversarial networks encounter artifacts, texture detail degradation, and other phenomena when facing more complex features, which leads to a visual deficiency. In order to solve these problems, an improved image inpainting algorithm combining a generative adversarial network with a coherent semantic attention mechanism is proposed. First of all, the generator adopts a two-stage inpainting method, uses gated convolution to replace the ordinary convolution of the generative adversarial network, and introduces residual blocks to solve the gradient vanishing problem and a coherent semantic attention mechanism to enhance the generator's attention to the important information and structure in the image. Secondly, Markov discriminator is adopted to enhance the network's discrimination effect, and the output results of the generator are processed by deconvolution to get the final repaired image. By comparing the inpainting results and image quality evaluation indicators with the baseline algorithm, the experimental results show that the algorithm can better predict the missing parts and improve the inpainting effect.

Key words: image inpainting; generative adversarial network (GAN); gated convolution; coherent semantic attention

图像修复的应用领域非常广泛, 包括医疗、工业生产、军事、娱乐和卫星通讯等. 但由于在图像的传播扩散或者是保存的过程中, 常有损坏、失真或者受到污染等问题, 从而导致图像信息不完整, 所以, 在图

像处理技术当中, 图像修复技术属于一项十分重要的研究工作. 图像修复技术可用于历史文物修复、旧照片修复、刑侦方面等.

与图像去噪、图像去雾、图像去雨以及图像超分

^① 收稿时间: 2023-03-21; 修改时间: 2023-04-20; 采用时间: 2023-05-06; csa 在线出版时间: 2023-08-22
CNKI 网络首发时间: 2023-08-23

等图像处理技术相对比, 图像修复更强调根据图像内容、细节以及特征纹理进行修复, 在对画面损坏或缺失部分进行补充的同时, 还要确保色彩连贯协调以及质感得到适当延伸, 同时, 在图像存在较大面积的破损时, 还要确保修复之后的图像内容与真实图像的一致性以及主观效果, 使得修复后的图像尽可能还原原始图像并且能够符合人类视觉感官。

在图像修复领域, 基于生成对抗网络的深度学习网络模型展现出其独特的优势, 但是在面对较复杂的特征时往往无法实现很好的修复效果, 存在伪影、纹理细节退化等现象, 造成视觉效果上的欠缺问题。为了解决当前模型对大面积破损图像修复不友好以及修复后的图像存在退化现象等问题, 通过对现有方法进行研究, 改进了图像修复模型。

本文的贡献如下: (1) 修复模型采用两阶段修复结构, 使用门控卷积解决部分像素缺失与大尺度损坏情况下的局部特征精确描述的问题, 结合残差思想解决使用 Sigmoid 函数会出现梯度消失的问题; (2) 在精修复阶段引入连贯语义注意机制, 不仅关注已知区域和未知区域之间的联系, 更能通过缺失区域间语义相关性进行建模, 对生成区域像素资源再次分配, 根据语义信息进行引导修复, 提升语义的连贯性。

1 相关工作

由于深度学习技术^[1]发展迅速, 图像修复也渐渐成为了计算机视觉的热门研究方向。Goodfellow 等人^[2]在 2014 年首次提出 GANs 应用于图像处理方面。Pathak 等人^[3]提出了 Context-Encoder 网络, 将编码—解码网络结构与生成对抗网络相结合, 以对抗损失和图像重构损失作为约束条件提高模型的图像修复能力。Iizuka 等

人^[4]提出了全局和局部判别器网络, 引入扩张卷积增大感受野, 通过双判别器来提升网络对全局信息和局部信息的融合利用, 使得图像修复结果不仅符合全局语义, 整体也更加合理。Liu 等人^[5]提出部分卷积修复方法, 使用 U-Net 网络结构实现图像修复。Yu 等人^[6]提出了一种带有内容感知的前馈生成网络, 将注意力机制引入到深度学习图像修复方法中。Yu 等人^[7]提出了一种基于门控卷积的图像修复算法, 采用两阶段生成结构。先是在生成器的第 1 阶段, 提取图像中的语义信息; 然后在生成器第 2 阶段将粗修复网络的结果作为输入, 细化生成部分的内容, 使生成图像的纹理结构更合理。

2 网络结构

2.1 修复网络体系结构

生成器为两阶段修复, 首先, 第 1 阶段(粗修复阶段)输入被污染的图像以及污染区域的灰度图, 通过网络提取图像中的语义信息; 其次, 第 2 阶段(精修复阶段)输入第 1 阶段的生成结果, 引入连贯语义注意机制更好地对被污染区域以及未被污染区域进行关联, 使修复的纹理结构更加合理。两阶段组成完整的修复流程, 具体修复网络结构如图 1 所示。

由于门控卷积拥有独特的特征筛选能力, 所以在生成器部分可以将其用于替代普通卷积, 提高图像处理的效率。为了进一步提升处理效果, 引入扩张卷积层^[8], 从而拓展了感受范围, 提供更丰富的特征信息, 促进图像修复, 并在模型设计中引入残差思想^[9], 将门控卷积改为残差门控卷积, 防止梯度消失问题。在生成器第 2 阶段, 通过使用连贯语义注意力来将局部特征和全局特征相关联, 能够更好地理解图像的整体结构, 提高图像的修复效果, 解决图像修复细节退化、视觉欠缺问题。

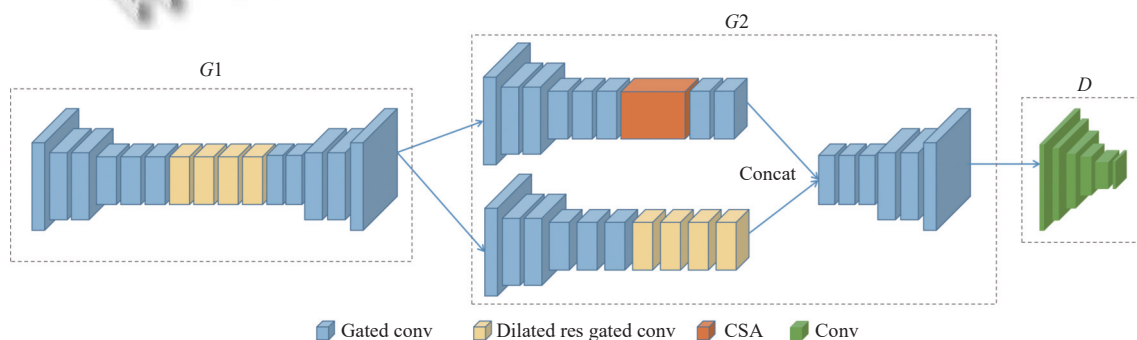


图 1 修复网络体系结构

2.2 生成器

2.2.1 门控卷积

门控卷积^[10]最早出现在 NLP 的任务中,取得了良好的结果, Yu 等人^[7]将其引入到图像修复网络. 门控卷积学习每个通道和每个空间位置的动态特征选择机制,具体操作的计算公式表达如下:

$$Gating_{y,x} = \sum \sum W_g \cdot I \quad (1)$$

$$Feature_{y,x} = \sum \sum W_f \cdot I \quad (2)$$

$$O_{y,x} = \sigma(Feature_{y,x}) \odot (Gating_{y,x}) \quad (3)$$

其中, Sigmoid 激活函数用 σ 表示, 输出门控值范围为 (0, 1). σ 可以表示 ReLU、ELU 和 LeakyReLU 等任意激活函数. W_g 和 W_f 代表两个不同的卷积滤波器. 门控卷积结构如图 2 所示.

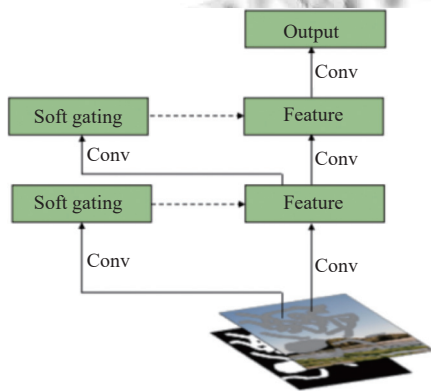


图 2 门控卷积结构图

通过在标准卷积后接入 Sigmoid 激活函数,使得门控卷积可以实现软门控,对当前卷积层的输出结果在进入下一个卷积层之前进行加权.在硬门控中,使用 0 或 1 来进行加权,但在软门控中,使用 0 到 1 进行加权,这样更加灵活,并且这种操作是可学习的.门控卷积特有的参数动态更新机制,对缺失区域进行动态学习,将缺失区域与已知区域进行有效的动态区分,可以更好地解决局部特征精确描述问题.

2.2.2 扩张卷积

Google 团队^[11]提出的扩张卷积 (dilated convolution) 也被称为空洞卷积或膨胀卷积,它是在标准卷积核中注入空洞,通过拓宽图像的感受野来获取图像中更多的特征信息.相比普通卷积操作,扩张卷积多了个扩张率 (dilation rate) 参数,扩张率用来定量衡量卷积核超参数的扩张程度.传统的普通卷积就是扩张率为 1 的

扩张卷积.图 3 为卷积核尺寸是 3×3 , dilation rate 是 2 的扩张卷积示意图.

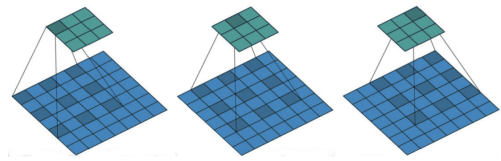


图 3 扩张率为 2 的 3×3 卷积核的扩张卷积示意图

从图 3 可以看出扩张率为 2 的 3×3 卷积核具有与 5×5 卷积核相同的视野,而且扩张卷积只使用了 9 个参数,也就是说,在相同的计算成本下,扩张卷积计算具有更宽的视野,从而能够捕捉更长的依赖关系.

本文将 4 层扩张卷积应用于粗修复阶段的中间部分,扩张卷积层的扩张率自 2 起随层数递增而翻倍增加,将扩张率持续变大的扩张卷积层叠加在一起,以持续增加图像的感受野,从而加强网络对图像整体特征信息的提取和利用.

2.2.3 残差门控卷积

门控卷积的输出门控值通过 Sigmoid 激活函数,输出值在 [0, 1]. 基于此,门控卷积避免了人为设置的更新方法,通过学习每个通道和每个空间位置的动态特征选择机制,由每一个通道进行动态特征选择,让网络自行学习每一个特征值位置.

Sigmoid 激活函数值域为 (0, 1), 由于 Sigmoid 函数值趋近于 0 或 1 时变化平缓,如果隐藏层较多梯度接近 0 时会出现梯度消失问题,在图像修复中导致一些特征无侧重点,引入残差思想,用残差门控卷积替代门控卷积,图 4 为以带扩张的门控卷积为例的残差门控卷积结构.

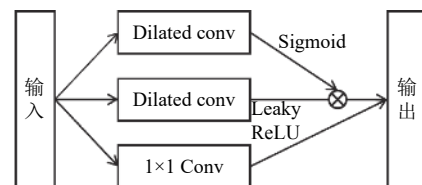


图 4 残差门控卷积结构

从图 4 可以看出,在门控卷积的基础上加上一个 1×1 的残差块构成残差门控卷积,对输入的有效特征信息进行卷积,卷积核大小为 1×1 ,最后将 1×1 卷积的输出添加到原门控卷积输出上,目的是提升对有效像素的使用.

2.2.4 连贯语义注意

图像修复任务仅考虑缺失区域与已知区域之间的

关系是不够的, 因为其忽略了缺失区域的内容, 为了达到更好的图像修复效果, 将连贯语义注意 (coherent semantic attention, CSA) 层^[12]用于图像修复任务重, 不仅分析了缺失区域特征之间的语义相关性, 而且可以从已知区域提取缺失区域的信息, 通过分析远距离的图像信息、缺失区域之间的关系, 可以保证修复结果的局部连续性.

CSA 分为两个阶段实现: 搜索和生成. 图 5 为 CSA 的搜索和生成操作. 每张待修复图片中用 M 表示当前图像受污染部分的信息, \bar{M} 表示图像中可见的可参考信息, 对于 M 中每个 (1×1) 生成的补丁 m_i ($i \in (1 \sim n)$, n 是补丁的数量), CSA 层在已知区域 \bar{M} 中搜索最接近匹配的上下文补丁 \bar{m}_i , 以在搜索过程中初始化 m_i , 然后将 m_i 设置为主要部分, 将所有先前生成的补丁 ($m_{1 \sim i-1}$) 设置为次要部分, 以在生成过程中恢复. 为了衡量这两个部分的权重, 采用以下互相关度量:

$$D_{\max i} = \frac{\langle m_i, \bar{m}_i \rangle}{\|m_i\| \cdot \|\bar{m}_i\|} \quad (4)$$

$$D_{\text{adi}} = \frac{\langle m_i, m_{i-1} \rangle}{\|m_i\| \cdot \|m_{i-1}\|} \quad (5)$$

其中, $D_{\max i}$ 代表 m_i 和上下文区域中最相似的补丁 \bar{m}_i 之间的相似度, D_{adi} 代表两个相邻生成的补丁之间的相似性. $D_{\max i}$ 和 D_{adi} 分别被归一化为上下文补丁部分和所有先前生成的补丁部分的权重.

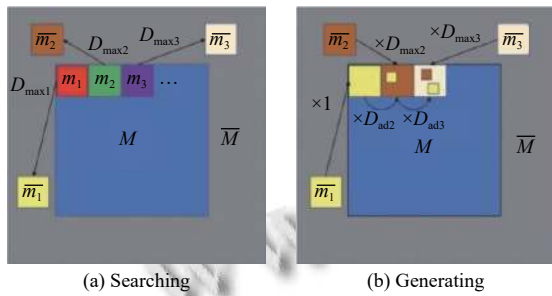


图 5 CSA 层示意图

CSA 的具体步骤为: 首先, 我们在洞 M 中搜索每个生成的补丁 m_i 中最相似的上下文补丁 \bar{m}_i , 并用 m_i 初始化 m_i . 然后, 将先前生成的补丁和最相似的上下文补丁组合起来生成当前补丁.

搜索阶段先提取 \bar{M} 中的补丁, 将补丁重塑为卷积滤波器, 然后在 M 上应用卷积滤波器. 通过这个操作, 我们可以获得一个值向量, 表示 M 中的每个补丁与 \bar{M} 中的所有补丁之间的互相关. 在此基础上, 对于每个

生成的补丁 m_i , 我们用最相似的上下文补丁 \bar{m}_i 对其进行初始化, 并将最大互相关值 $D_{\max i}$ 分配给它以进行下一步操作.

生成阶段将 M 的左上角块作为生成过程的初始块 (在图 5 中用 m_1 标记). 由于 m_1 没有之前的补丁, 所以 $D_{\text{ad}1}$ 为 0, 我们直接将 \bar{m}_1 替换为 m_1 , $m_1 = \bar{m}_1$. 下一个补丁 m_2 有一个前一个补丁 m_1 作为附加参考, 但我们因此将 m_1 视为一个卷积滤波器, 以测量 m_1 和 m_2 之间的互相关度量 $D_{\text{ad}2}$. 然后, 将 $D_{\text{ad}2}$ 和 $D_{\max 2}$ 合并, 分别归一化为 m_1 和 \bar{m}_2 的权重, 生成新的 m_2 值, $m_2 = \frac{D_{\text{ad}2}}{D_{\text{ad}2} + D_{\max 2}} \times m_1 + \frac{D_{\max 2}}{D_{\text{ad}2} + D_{\max 2}} \times \bar{m}_2$.

综上所述, 从 m_1 到 m_n , 生成过程可以概括为:

$$m_1 = \bar{m}_1, D_{\text{ad}1} = 0 \quad (6)$$

$$m_{i \in (2 \sim n)} = \frac{D_{\text{adi}}}{D_{\text{adi}} + D_{\max i}} \times m_i + \frac{D_{\max i}}{D_{\text{adi}} + D_{\max i}} \times \bar{m}_i \quad (7)$$

如式 (7) 所示, 生成操作是一个迭代过程, 每个 m_i 包含 m_i 和 $m_{1 \sim i-1}$ 的信息, 当我们计算 m_i 和 m_{i-1} 之间的 D_{adi} 时, 都考虑了 m_i 和 $m_{1 \sim i-1}$ 之间的相关性. 并且由于 D_{adi} 值的范围是 0 到 1, 因此当前生成的补丁与之前生成的补丁之间的相关性随着距离的增加而降低. 基于式 (6), 我们得到了一个注意力图 A_i , 它记录了 m_i 的 $\frac{D_{\max i}}{D_{\text{adi}} + D_{\max i}}$ 和 $\frac{D_{\max i}}{D_{\text{adi}} + D_{\max i}} \times A_{i-1}$, 然后 A_1 到 A_n 形成一个注意力矩阵, 最后将 \bar{M} 中提取的补丁重新用作反卷积滤波器以重建 M . CSA 算法具体流程如算法 1 所示.

算法 1. CSA 算法流程

输入: 当前批次的特征图集 F_{in}
输出: 重构的特征图 F_{out}

- 1) 搜索:
- 2) 将 \bar{M} 重塑为卷积滤波器并将其应用于 M ;
- 3) 用式 (4) 计算 $D_{\max i}$ 并得到 \bar{m}_i ;
- 4) 用 \bar{m}_i 初始化 m_i ;
- 5) 结束搜索.
- 6) 生成:
- 7) for $i=1 \rightarrow n$ do;
- 8) 用式 (5) 计算 D_{adi} ;
- 9) 用式 (7) 得到 m_i 的注意力图 A_i ;
- 10) end for;
- 11) 将 A_1 到 A_n 结合得到一个注意力矩阵;
- 12) 重用 \bar{M} 作为反卷积得到 F_{out} ;
- 13) 结束生成.
- 14) 返回 F_{out} .

2.3 判别器

马尔可夫判别器 (PatchGAN)^[13] 一种仅在补丁范围内进行“惩罚”的鉴别器, 如图 6 所示, PatchGAN 在输入上滑动一个 $N \times N$ 的窗口, 并生成一个分数, 输出矩阵的每个值表示对应的图像补丁是真实的还是人为生成的概率。

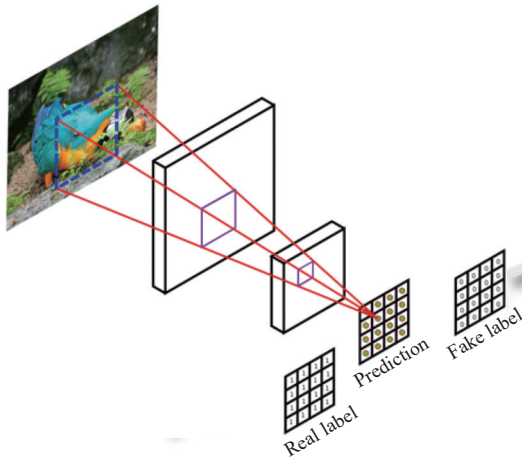


图 6 马尔可夫判别器

谱归一化马尔可夫判别器 (spectral normalized PatchGAN, SN-PatchGAN) 用于训练自由形式的图像修复网络, 它直接对卷积判别器的输出特征图中的每个点应用 GAN 损失, 从而产生高质量的修复效果. 采用光谱归一化马尔可夫判别器作为模型判别器, 通过对图像修复网络特征图进行检验, 将修复后的图像与真实图像进行判别, 并采用谱归一化方法对 GANs 的训练进行进一步的稳定, 网络通过判别器来进行对抗训练, 进而提升生成器的生成效果。

2.4 损失函数

训练模型的损失函数包括像素级 $L1$ 重建损失以及 SN-PatchGAN 损失两个损失项, 用来平衡这两个损失项的超参数为 1:1. 为了区分输入的真假, 以较链损失 (hinge loss) 作为目标函数:

$$L_{D^{sn}} = E_{x \sim P_{data}(x)} [ReLU(1 - D^{sn}(x))] + E_{z \sim P_z(z)} [ReLU(1 - D^{sn}(z))] \quad (8)$$

$$L_G = -E_{z \sim P_z(z)} [D^{sn}(G(z))] \quad (9)$$

其中, D^{sn} 表示谱归一化判别器, G 是带有不完整图像 z 的图像修复网络。

由于谱归一化判别器中已经编码了类似的补丁级

别信息, 所以这里的损失函数不再使用感知损失。

3 图像质量评价指标

文献 [6] 中提到图像修复缺乏良好的定量评价指标. 本文使用 $L1$ 、 $L2$ 损失函数、峰值信噪比 (PSNR) 以及结构相似度 (SSIM) 这 4 种指标来评估图像修复的结果。

$L1$ 损失函数被用于最小化误差, 即真实图像与修复图像间全部绝对差的总和。

$$L1 = \frac{1}{n} \sum_{i=1}^n |y_{true} - y_{predicted}| \quad (10)$$

$L2$ 损失函数用于最小化误差, 该误差是真实图像和修复图像所有平方差的总和。

$$L2 = \frac{1}{m} \sum_{i=1}^m (y_{true} - y_{predicted})^2 \quad (11)$$

峰值信噪比 (PSNR) 是评价图像质量的一个指标, PSNR 值越大, 图像失真越少。

$$PSNR = 10 \log_{10} \left(\frac{(2^{bits} - 1)^2}{m \cdot L2} \right) \quad (12)$$

结构相似性 (SSIM) 是衡量真实图像和生成图像相似度的指标, SSIM 值越大代表图像越相似。

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

其中, μ_x 是 x 的平均值, μ_y 是 y 的平均值, σ_x^2 是 x 的方差, σ_y^2 是 y 的方差, σ_{xy} 是 x 和 y 的协方差. $C_1 = (k_1 L)^2$, $C_2 = (k_2 L)^2$ 是维持稳定的常数. L 是像素值的动态范围. $k_1 = 0.01$, $k_2 = 0.03$ 。

4 实验及结果分析

4.1 实验数据

实验使用的数据集是 Places365 数据集^[14], 包含数百个实际场景的综合场景, 能够更好地结合各个场景的图像修复, 从中选取 30 000 张图片作为本文实验数据集. 矩形掩码为随机生成破损区域的方式, 随机掩码使用的是文献 [7] 中提出的。

4.2 实验设置

本文模型使用 PyTorch 深度学习框架. 同时使用 CUDA 11.0 和 CuDNN 8.0 加速模型训练, 硬件设备配置表如表 1 所示. 模型选择 Adam 优化器, 学习率大小为 0.0001, 总计迭代 30 万次。

表1 硬件设备配置表

| 设备名称 | 型号 |
|---------|---------------------|
| 显卡 | RTX 3080 |
| 处理器 | AMD7302×2 (32核64线程) |
| 系统内存 | 512 GB |
| 系统类型及版本 | Ubuntu 20.04 |

4.3 实验结果对比与分析

为验证算法的有效性, 将本文提出的算法通过矩形掩码以及不规则掩码在图像质量评价指标上与基线进行对比, 具体结果如表2所示。

由表2可以看出, 当掩码为矩形掩码时, 本文算法在 $L1$ 指标上提升了 0.14%, 在 $L2$ 指标上提升了 0.04%, $PSNR$ 提升了 0.41 dB, $SSIM$ 提升了 0.75%; 掩码为不规则掩码时, 本文算法 $L1$ 提升了 0.3%, $L2$ 提升了 0.05%,

$PSNR$ 提升了 0.55 dB, $SSIM$ 提升了 1.23%。可以得出, 本文的算法在各项指标上都优于基线算法。

表2 图像质量评价指标对比

| Mask | Method | $L1$ (%) | $L2$ (%) | $PSNR$ (dB) | $SSIM$ |
|------------------|-------------|----------|----------|-------------|--------|
| Rectangular mask | Deepfill V2 | 8.6 | 2.0 | 24.63 | 0.8540 |
| | 本文算法 | 8.46 | 1.96 | 25.04 | 0.8615 |
| Free-form mask | Deepfill V2 | 9.1 | 1.6 | 25.47 | 0.8790 |
| | 本文算法 | 8.8 | 1.55 | 26.02 | 0.8913 |

除了客观的评价指标, 也能从实际的图像修复结果中反映图像的修复效果, 本文算法的修复结果与基线的对比如图7所示, 从左到右依次为受损图像, 基线模型结果, 本文算法的修复结果以及真实图像。前2行为矩形掩码, 后3行为随机掩码。通过对比图6的修复结果可以发现本文算法在图像纹理、边缘细节上的处理都是优于基线的。



图7 修复结果对比

5 结论与展望

本文提出了一种基于生成对抗网络的图像修复模型,在生成器阶段采用门控卷积网络,使用两阶段修复法,在第1阶段加入扩张卷积并引入残差思想,对图像进行粗修复,在第2阶段引入连贯语义注意对生成区域像素资源再次分配,根据语义进行引导修复,从而提升语义的连贯性.使用谱归一化马尔可夫判别器给生成器提供精细的反馈,从而提高图像的修复质量.通过实验表明,本文提出的图像修复模型修复效果良好,使修复区域内容更丰富,纹理更逼真.

本文希望在下一步工作中,能够对图像缺失部分的关键特征信息进行更好地研究以及得到多重修复结果,达到更好的修复效果.

参考文献

- 1 赵露露,沈玲,洪日昌.图像修复研究进展综述.计算机科学,2021,48(3):14–26. [doi: 10.11896/jsjx.210100048]
- 2 Goodfellow IJ, Pouget-Abadie J, Mirza M, *et al.* Generative adversarial nets. Proceedings of the 27th International Conference on Neural Information Processing Systems. Montreal: MIT Press, 2014. 2672–2680.
- 3 Pathak D, Krähenbühl P, Donahue J, *et al.* Context encoders: Feature learning by inpainting. Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas: IEEE, 2016. 2536–2544.
- 4 Iizuka S, Simo-Serra E, Ishikawa H. Globally and locally consistent image completion. ACM Transactions on Graphics, 2017, 36(4): 107.
- 5 Liu GL, Reda FA, Shih KJ, *et al.* Image inpainting for irregular holes using partial convolutions. Proceedings of the 15th European Conference on Computer Vision. Munich: Springer, 2018. 89–105.
- 6 Yu JH, Lin Z, Yang JM, *et al.* Generative image inpainting with contextual attention. Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018. 5505–5514.
- 7 Yu JH, Lin Z, Yang JM, *et al.* Free-form image inpainting with gated convolution. Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision. Seoul: IEEE, 2019. 4470–4479.
- 8 Yu F, Koltun V. Multi-scale context aggregation by dilated convolutions. arXiv:1511.07122, 2015.
- 9 Yi ZL, Tang Q, Azizi S, *et al.* Contextual residual aggregation for ultra high-resolution image inpainting. Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle: IEEE, 2020. 7505–7514.
- 10 Dauphin YN, Fan A, Auli M, *et al.* Language modeling with gated convolutional networks. Proceedings of the 34th International Conference on Machine Learning. Sydney: PMLR, 2017. 933–941.
- 11 Chen LC, Papandreou G, Kokkinos I, *et al.* DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2018, 40(4): 834–848. [doi: 10.1109/TPAMI.2017.2699184]
- 12 Liu HY, Jiang B, Xiao Y, *et al.* Coherent semantic attention for image inpainting. Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision. Seoul: IEEE, 2019. 4169–4178.
- 13 Demir U, Unal G. Patch-based image inpainting with generative adversarial networks. arXiv:1803.07422, 2018.
- 14 Zhou BL, Lapedriza A, Khosla A, *et al.* Places: A 10 million image database for scene recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2018, 40(6): 1452–1464. [doi: 10.1109/TPAMI.2017.2723009]

(校对责编:牛欣悦)