

# 融合边缘检测和自注意力的图像修复方法<sup>①</sup>



李维乾, 张晓文

(西安工程大学 计算机科学学院, 西安 710600)

通讯作者: 张晓文, E-mail: blackshaun@163.com

**摘要:** 针对修复后图像边界模糊、图像纹理不清晰、视觉效果差的问题, 提出了一种融合边缘检测和自注意力机制的生成式对抗修复模型. 通过边缘检测可提取出图像的轮廓信息, 避免了修复后边界模糊的问题; 利用自注意力机制能够捕获图像全局信息并生成图像精确细节的能力, 设计出融合自注意力机制的纹理修复网络. 该模型由边缘补全网络和纹理修复网络组成, 首先, 设计的边缘补全网络对受损图像的边缘进行补全, 得到边缘补全图像; 其次, 利用纹理修复网络联合补全的边缘图像对缺失区域的纹理进行精确修复. 在 CelebA 和 Place2 两个图像数据集上对本文所建模型进行了训练和测试. 实验结果表明: 本文所建模型与现有图像修复方法相比, 大幅提高了图像修复的精确度, 且生成的图像更加逼真.

**关键词:** 图像修复; 生成对抗网络; 自注意力机制; 边缘检测; 纹理修复

引用格式: 李维乾, 张晓文. 融合边缘检测和自注意力的图像修复方法. 计算机系统应用, 2021, 30(5): 150-156. <http://www.c-s-a.org.cn/1003-3254/7880.html>

## Combining Edge Detection and Self-Attention for Image Inpainting

LI Wei-Qian, ZHANG Xiao-Wen

(School of Computer Science, Xi'an Polytechnic University, Xi'an 710600, China)

**Abstract:** To address the problems of blurred image boundaries, unclear image texture, and poor visual effect after inpainting, we propose a generative adversarial inpainting model that combines edge detection with self-attention mechanism in this study. Through this model, the contour information of the images can be extracted by edge detection, avoiding the problem of blurred boundaries after inpainting. Since the self-attention mechanism can capture the global information of images and generate precise details, a texture inpainting network incorporating the self-attention mechanism is designed. The proposed model is composed of an edge complement network and a texture inpainting network. First, the designed edge complement network completes the edges of a damaged image to obtain an edge complement image. Secondly, the texture of the missing region is accurately inpainted by the texture inpainting network combining the complemented edge image. Finally, the model proposed in this study is trained and tested on the CelebA and Place2 image datasets. The experimental results show that compared with the existing image inpainting methods, the model can greatly improve the accuracy of image inpainting and generate vivid images.

**Key words:** image inpainting; generation adversarial network; self-attention mechanism; edge detection; texture inpainting

图像修复是计算机视觉领域的一个重要任务. 它在数字艺术品修复、公安刑侦面部修复等种种实际场

景中被广泛应用. 图像修复的核心挑战在于为缺失区域合成视觉逼真和语义合理的像素, 要求合成的像素

<sup>①</sup> 收稿时间: 2020-08-31; 修改时间: 2020-09-23; 采用时间: 2020-09-29; csa 在线出版时间: 2021-04-28

与原像素具有一致性。

传统的图像修复技术有基于结构和纹理两种方法。基于结构的图像修复算法具有代表性的是 Bertalmio 等<sup>[1]</sup>提出的 BSCB 模型和 Shen 等<sup>[2]</sup>提出的基于曲率扩散的修复模型 CDD。基于纹理的修复算法中具有代表性的有 Criminisi 等<sup>[3]</sup>提出的基于 patch 的纹理合成算法。这两种传统的修复算法可以修复小块区域的破损,但是在破损区域越来越大时,修复效果则直线下降,并且修复结果存在图像模糊、结构扭曲、纹理不清晰和视觉不连贯等问题。

近年来随着硬件设备等计算能力的不断提升,以及深度学习技术在图像翻译<sup>[4,5]</sup>、图像超分辨率<sup>[6]</sup>、图像修复<sup>[7-9]</sup>等计算机视觉领域的迅速发展,采用深度学习技术的修复方法能够捕获图像的高层语义信息,与传统的修复方法相比,具有良好的修复效果。现有基于深度学习的修复方法包括卷积神经网络的方法和生成式对抗网络 (Generative Adversarial Networks, GAN) 的方法,其中基于 GAN 的方法在图像修复领域的应用比较出色。该技术采用对抗训练的方式来帮助模型生成需要的样本图像,采用 GAN 方法能够提高数据质量,提升学习效率,但存在修复区域与周围区域语义不相关、边界模糊、训练不稳定等问题,因此相关学者对基于 GAN 网络的图像修复方法进行了改进。如 Pathak 等<sup>[10]</sup>提出的 Context Encoder 模型,该模型使用对抗损失训练一个上下文编码器预测图像的受损区域,但使用该方法修复的图像比较模糊且不一致; Iizuka 等<sup>[11]</sup>提出了双鉴别器的生成对抗模型,该模型通过增设局部鉴别器跟踪受损区域保证了图像的一致性; Liu 等<sup>[12]</sup>提出的连贯性语义注意力方法,通过搜索已知区域中最相似的特征 patch 填满缺失区域,增加 patch 判别器进行一致性优化,保证了全局一致性;为了得到符合语义的图像, Yu 等<sup>[13]</sup>提出了基于上下文注意力的两阶段模型,加入上下文注意力模块,实现提取局部区域的有效特征,但是它不能处理任意形状的受损区域。还有一些利用其它技术的研究,如 Nazeri 等<sup>[14]</sup>提出了基于边缘连接的修复模型,它应用随机生成二值掩码解决了不能处理任意形状受损区域的问题,通过边缘连接网络生成图像边缘信息,有效解决了图像边界模糊的问题; Ren 等<sup>[15]</sup>提出的结构感知修复模型,通过边缘图像来表示图像的全局结构,虽然基于边缘特征的方法有效地描绘了轮廓信息,但是它有时无法准确地描绘高度纹理化区域。

尽管上述方法在图像修复任务中表现了出色的性能,但还存在以下不足:首先图像的高频纹理信息无法准确修复;其次虽然保证了图像在视觉上的一致性和连续性,但是有些区域的边界和纹理存在修复模糊的现象。

因此,针对修复后图像边界模糊、纹理不清晰、视觉效果差的问题,本文提出了融合边缘检测和自注意力机制的生成对抗修复模型。该模型由边缘补全网络和纹理修复网络构成,首先利用边缘补全网络填补图像的边缘,得到边缘补全图像;其次通过纹理修复网络联合边缘补全图像修复缺失的纹理信息。由于卷积神经网络主要处理局部区域的特征,仅使用卷积神经网络很难建立长相关性模型,而自注意力机制可以处理图像中全局范围的特征,可以建立长范围、多层次的依赖关系。因此本文将卷积神经网络与自注意力机制相结合设计出纹理修复网络来生成缺失区域的纹理信息,最终使模型生成的图像与原图像在边缘、纹理结构、视觉上高度逼真。

## 1 生成式对抗网络

生成式对抗网络由 Goodfellow 于 2014 年提出<sup>[16]</sup>。GAN 由一个生成器 Generator 和一个鉴别器 Discriminator 组成,生成器  $G$  的输入是一个随机噪声  $z$ ,通过这个噪声来生成“以假乱真”的图像  $G(z)$ 。鉴别器  $D$  的输入是真实图像和生成图像的混合,它的目的是判别输入图像的真实性。网络通过对抗训练让生成器和鉴别器同时得到增强,训练时,生成器  $G$  用来生成逼真的“假”图像去混淆  $D$  的判断,鉴别器  $D$  要判别出生成器  $G$  生成的图像为假,同时给生成器一个反馈来更新网络的参数,使生成器生成质量更高的图像,鉴别器的判别能力也会随之提高。GAN 的基本网络结构如图 1 所示。

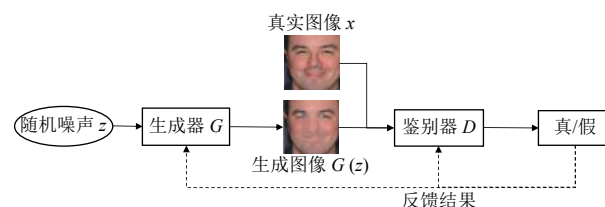


图 1 GAN 的基本网络结构

生成对抗网络的目标是最小化生成器生成的图像与真实图像概率分布之间的差距。因此生成式对抗网络的目标函数可以表示为:

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

式中,  $G$  和  $D$  分别表示生成器与鉴别器的可微函数,  $x$  是真实数据样本,  $z$  是随机噪声矢量,  $G(z)$  是生成器的生成数据.  $E$  代表数学期望值. 在同时训练生成器和鉴别器的情况下, GAN 训练的目标是令判别器  $D$  的输出  $D(G(z))$  趋近于 0,  $D(x)$  趋近 1, 然后将结果反馈给生成网络, 生成器调整自身参数, 使生成器下一次生成更加逼真的图像. 训练中一次次迭代使生成器和鉴别器最终达到一个“平衡点”, 即生成器生成的图片已经“以假乱真”, 令鉴别器无法判别图像的真伪.

## 2 融合边缘检测和自注意力的修复模型

为了解决图像修复技术中存在的边界模糊、纹理

不清晰、视觉效果差的问题, 本文提出了融合边缘检测和自注意力机制的生成对抗修复模型. 如图 2 所示, 整个模型包含边缘补全网络和纹理修复网络两个部分. 首先检测真实图像和受损图像的边缘, 得到其边缘图像, 以真实图像的边缘为标签, 受损图像及其边缘图像为输入训练边缘补全网络, 补全缺失区域的轮廓信息, 得到边缘补全图像; 其次, 以边缘补全图像和受损图像作为输入, 对融合注意力机制的纹理修复网络进行训练, 使生成器  $G_2$  生成缺失区域的精确纹理信息, 从而使纹理修复网络完成缺失区域纹理的精确修复; 最后将边缘补全网络的生成器  $G_1$  和纹理修复网络的生成器  $G_2$  进行级联, 以受损图像及其边缘图像作为模型的输入, 以真实图像为标签进行联合训练, 实现端到端的图像修复.

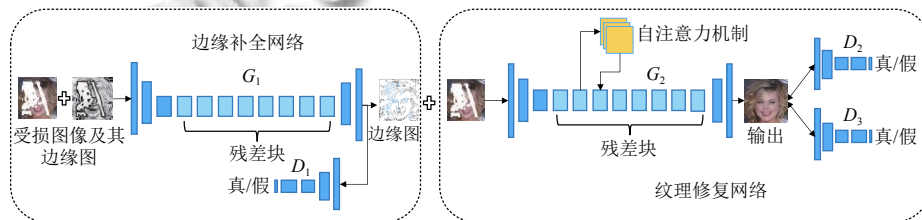


图2 融合边缘检测和自注意力机制的生成对抗修复模型

### 2.1 边缘补全

为了在图像修复中提供清晰的边界和结构, 本文只提取物体的轮廓信息. 传统的基于 Sobel 和 Canny 的检测方法只考虑颜色、光线和梯度局部变化. 在复杂的场景中, 它们是无法满足需要的. 因此本文选择整体嵌套边缘检测 (Holistically-nested Edge Detection, HED)<sup>[17]</sup> 方法. 它在处理的是整个图像, 并能够提取出图像的高层边界信息. 本文使用已经训练好的 HED 模型提取图像的边缘.

图像边缘的生成分为两个步骤, 第一是从受损图像中提取边缘信息, 第二是补全缺失的边缘. 如图 2 左半部分所示, 边缘补全网络包含一个生成器  $G_1$  和一个鉴别器  $D_1$ . 生成器  $G_1$  由对图像进行两次下采样的编码器, 随后的 8 个残差块和将图像上采样回原始大小的解码器组成. 边缘补全网络的输入由受损图像及其边缘图像组成, 通过卷积和标准化处理, 提取出图像深层次纹理边缘信息, 通过对边缘特征的学习训练, 最终让生成器  $G_1$  预测出缺失区域的边缘. 为了提取深层次的边缘纹理特征, 在卷积和反卷积之间增加了残差块,

避免了由于网络深度过深而导致的梯度弥散问题.

鉴别器  $D_1$  是判别边缘修复是否准确的网络. 本文利用 HED 的边缘检测方法提取出真实图像的边缘信息, 将其作为网络训练的正样本, 然后, 将生成器  $G_1$  生成的样本作为负样本和正样本进行混合输入鉴别器  $D_1$ , 判别其真实性, 通过不断优化损失函数来提升鉴别器  $D_1$  的判别能力, 生成器  $G_1$  的生成能力随着鉴别器  $D_1$  判别能力的提升而不断上升, 促使生成器  $G_1$  生成的边缘图像与原图的边缘图像趋于一致.

$I_{gt}$  表示真实图像, 其边缘图用  $B_{gt}$  表示. 在边缘补全网络中, 使用受损图像  $I_{in}$  及其边缘图像作为输入, 其边缘图像为  $\tilde{B}_{gt} = B_{gt} \odot (1 - M)$ , 其中  $M$  为受损图像  $I_{in}$  的掩码, 该变量为一个二值矩阵 (缺失区域为 1, 背景为 0); 在此,  $\odot$  表示矩阵乘法. 将受损图像  $I_{in}$  和它的边缘图像  $\tilde{B}_{gt}$  作为生成器  $G_1$  的输入, 那么边缘补全网络生成的缺失区域的边缘图像是:

$$B_{pred} = G_1(I_{in}, \tilde{B}_{gt}, M) \quad (2)$$

其中,  $B_{pred}$  为边缘补全图像. 鉴别器  $D_1$  以  $B_{gt}$  和  $B_{pred}$  作

为输入来判断边缘补全图像是否真实. 用对抗损失作为目标函数对网络进行训练, 通过优化目标函数不断提升鉴别器和生成器的能力来实现图像边缘信息的补全.

## 2.2 纹理修复

如图2右半部分所示, 纹理修复网络由生成器  $G_2$  和双鉴别器  $D_2$ 、 $D_3$  组成, 生成器  $G_2$  由对图像进行两次下采样的编码器, 引入自注意力机制的残差块和将图像上采样回原始大小的解码器组成, 其输入是边缘补全图像和受损图像的结合. 该网络将自注意力机制与卷积神经网络进行结合来提取输入图像深层次特征. 自注意力机制会利用所有位置的特征来生成图像某一位置的细节. 通过自注意力机制计算得到输入图像的注意力特征图, 然后将注意力特征图与提取的卷积特征图进行连接输入到网络中继续进行训练, 通过不断优化模型提升生成器的能力, 最后促使生成器生成与原图像趋于一致的图像.

鉴别器由全局鉴别器  $D_2$  和局部鉴别器  $D_3$  构成双鉴别器机制. 全局鉴别器  $D_2$  的作用是判别修复后图像的整体真实性, 它的作用集中在图像的全部区域, 它监督生成器生成视觉一致的图像; 局部鉴别器  $D_3$  用于判别图像受损区域的真实性. 鉴别器的输入为生成器的生成图像和真实图像的混合, 通过优化对抗损失和重构损失函数提升鉴别器的判别能力, 让鉴别器与生成器进行抗衡, 来迫使生成器生成与原图纹理细节一致的图像.

纹理修复网络是一个加入自注意力机制的生成式对抗网络. 生成器  $G_2$  用于生成生动的图像纹理信息. 鉴别器  $D_2$  和  $D_3$  分别用于判别生成的图像全局和局部区域是否为真. 将边缘补全图像  $B_{\text{pred}}$  和受损图像  $I_{\text{in}}$  一起作为纹理修复网络的输入, 最终生成的完整图像是:

$$I_{\text{out}} = G_2(I_{\text{in}}, B_{\text{pred}}, M) \quad (3)$$

其中,  $I_{\text{out}}$  为纹理修复网络的输出结果, 它是缺失区域的完整修复结果.

## 2.3 自注意力机制

在图像修复任务中, 最终目标是训练一个生成器网络, 使其生成与真实图像高层纹理细节尽可能相似的图像. 为了实现这一点, 引入自注意力机制, 它可以更好地利用全局特征信息对图像进行重建.

卷积操作主要处理局部区域间的信息, 因此多个卷积层在建立全局依赖性上计算效率非常低. 自注意

力机制提供了全局的感受野, 它会学习某个像素点和其它所有位置像素点的关系, 使得每个位置的生成不再是仅仅依赖这个位置临近的其他位置, 而是引入了距离更远且具有更大信息量的位置, 每个位置的特征都由其它所有位置的加权和来计算得到. 因此, 本文借鉴文献[18]思路, 在纹理生成网络的残差块中引入自注意力机制, 构建自注意力机制的生成式对抗网络, 通过该网络获取图像的全局特征信息, 从而更好地重建出图像的纹理细节, 如图3所示.

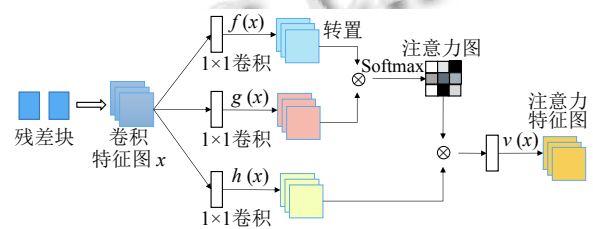


图3 自注意力机制

图3中, 从残差块第二层中提取的卷积特征图  $x$  分别经过两个  $1 \times 1$  的卷积层变换为两个特征空间  $f(x)$ ,  $g(x)$  去计算注意力图.  $f(x)$  用来提取像素特征,  $g(x)$  用来提取全局特征信息, 其中  $f(x) = W_f x$ ,  $g(x) = W_g x$ .

$$\beta_{j,i} = \frac{\exp(s_{ij})}{\sum_{i=1}^N \exp(s_{ij})} \quad (4)$$

其中,  $s_{ij} = f(x_i)^T g(x_j)$ .  $\beta_{j,i}$  表示模型在合成第  $j$  个区域时对第  $i$  个位置的关注程度.  $N$  表示前一层特征图的特征位置数. 注意力特征图的输出是  $o = (o_1, o_2, \dots, o_j, \dots, o_N)$ .

$$o_j = v \left( \sum_{i=1}^N \beta_{j,i} h(x_i) \right) \quad (5)$$

其中,  $h(x_i) = W_h x_i$ ,  $v(x_i) = W_v x_i$ . 此外再将自注意力机制的输出乘以一个比例参数, 并添加到输入自注意力机制的卷积特征图中, 因此, 最终输出为:

$$y_i = \gamma o_i + x_i \quad (6)$$

其中,  $y_i$  是自注意力机制的最终输出结果, 将其作为下一个残差块的输入.  $\gamma$  是一个可学习的标量, 它的初始值为 0, 引入  $\gamma$  让网络首先依赖局部相邻的线索, 然后逐渐学习给非局部线索分配更多的权重, 这样做的原因是先让模型学习简单的任务, 然后逐渐增加任务的复杂性.

## 2.4 损失函数

传统的生成式对抗网络在训练时使用对抗损失作

为目标函数,但存在训练不稳定的问题.本文使用上下文注意力模型<sup>[13]</sup>中基于 WGAN-GP 的修改版本作为训练网络的目标函数,促使网络生成视觉逼真的图像. WGAN-GP 损失在图像生成任务方面的性能比现有的 GAN 损失性能强.

WGAN 使用 EM 距离  $W(P_r, P_g)$  来比较真实数据和生成数据的分布. WGAN 的目标函数是:

$$\min_G \max_{D \in \mathcal{d}} E_{x \sim P_r} [D(x)] - E_{\tilde{x} \sim P_g} [D(\tilde{x})] \quad (7)$$

其中,  $\mathcal{d}$  是 1-Lipschitz 函数的集合,  $P_r$  是真实数据的分布,  $P_g$  是  $\tilde{x} = G(z)$  的模型分布,  $z$  是生成器的输入. 对于图像修复来说,只需要修复图像缺失的区域,所以梯度惩罚项为:

$$\lambda E_{\hat{x} \sim P_{\hat{x}}} (\|\nabla_{\hat{x}} D(\hat{x}) \odot (1 - M)\|_2 - 1)^2 \quad (8)$$

这里  $\hat{x}$  是从分布  $P_g$  和  $P_r$  采样的点之间的直线中采样的.

(1) 边缘补全的损失. 本文将 WGAN-GP 损失应用到模型的两个部分中,对边缘补全网络来说,它的对抗损失定义为:

$$L_1 = E_{\tilde{x} \sim P_{B_{pred}}} [D(\tilde{x})] - E_{x \sim P_{B_{gt}}} [D(x)] + \lambda E_{\hat{x} \sim P_{\hat{x}}} (\|\nabla_{\hat{x}} D(\hat{x}) \odot (1 - M)\|_2 - 1)^2 \quad (9)$$

其中,  $P_{B_{pred}}$  是边缘补全图像的数据分布,  $P_{B_{gt}}$  是真实边缘图像数据的分布,  $\hat{x}$  是从生成数据分布和真实数据分布采样的点之间的直线中采样的.

(2) 纹理修复的损失. 纹理修复网络的输出为修复结果  $I_{out}$ . 它的对抗损失定义为:

$$L_2 = E_{\tilde{x} \sim P_{I_{out}}} [D(\tilde{x})] - E_{x \sim P_{I_{gt}}} [D(x)] + \lambda E_{\hat{x} \sim P_{\hat{x}}} (\|\nabla_{\hat{x}} D(\hat{x}) \odot (1 - M)\|_2 - 1)^2 \quad (10)$$

其中,  $P_{I_{out}}$  是生成的修复图像的分布,  $P_{I_{gt}}$  是真实图像的分布.

WGAN-GP 损失和  $I_1$  重构损失联合使用时更容易训练,优化过程更稳定,因此在纹理修复网络中使用重构损失和 WGAN-GP 损失组合作为目标函数对网络进行训练. 纹理修复网络的重构损失被定义为:

$$L_{(l_1,2)} = \frac{1}{N} \sum_{n=1}^N \frac{1}{WHC} \|I_{out} - I_{gt}\| \quad (11)$$

其中,  $N$  是样本的数量,  $W, H, C$  分别是图像的宽度、高度和通道数. 那么纹理修复网络训练的目标损失函数为:

$$L_{G_2} = \lambda_1 L_2 + \lambda_2 L_{(l_1,2)} \quad (12)$$

其中,  $\lambda_1$  和  $\lambda_2$  分别是重构损失和对抗损失的参数,在实验中,  $\lambda_1 = 1, \lambda_2 = 4$ .

### 3 实验结果及分析

#### 3.1 数据集

为了验证本文图像修复方法的效果,本文实验中用 Place2 和 CelebA 数据集来训练和评估本文的两阶段模型, CelebA 人脸数据集包含 10177 个名人身份的 202599 张人脸图像. Place2 是一个场景图片数据集,它包含 1 千万张图片,400 多个不同种类的场景. 将掩码设置为不规则覆盖,覆盖范围设置整幅图像的 10-20%.

#### 3.2 参数设置

本文的实验平台为 Windows10 系统、Python3.6 和 PyTorch1.0.0 结合的编程环境. CPU 信息为 Intel(R) Core(TM) i7-8550U CPU, 主频 2.0 GHz, 内存为 16 GB, NVIDIA Titan XP GPU. 在训练中使用  $256 \times 256$  的图像, batch size 设置为 32. 使用 Adam 优化器<sup>[19]</sup> 来优化模型, 它的参数设置为  $\beta_1 = 0.9$  和  $\beta_2 = 0.999$ . 本文设置学习率为  $10^{-3}$  来训练  $G_1$  和  $G_2$  生成器直到收敛为止, 鉴别器以生成器学习率的 0.1 倍进行训练.

#### 3.3 实验结果及分析

为了分析本文方法在多个场景的应用效果,在 Place2 和 CelebA 数据集上进行效果评估. 并与近年来基于生成式对抗网络具有代表性的图像修复方法进行比较(上下文注意力的生成式对抗模型<sup>[13]</sup>和边缘连接的生成式对抗模型<sup>[14]</sup>). 本文用峰值信噪比 (PSNR) 和结构相似性指数 (SSIM) 对实验结果进行定量评估,以此来表示修复后图像和真实图像的差距.

峰值信噪比 PSNR 是一种评价图像的客观标准,它直接衡量像素值的差异,数值越大表示失真越小. 假设两张图像分别是  $X_1$  和  $X_2$ , 计算公式如下:

$$PSNR(X_1, X_2) = 10 \lg \left( \frac{MAX_{X_1}^2}{MSE} \right) \quad (13)$$

其中,  $MAX_{X_1}^2$  为图片可能的最大像素值,这里每个像素都由 8 位二进制表示,所以最大像素取 256,  $MSE$  表示  $X_1$  和  $X_2$  的均方误差,  $MSE$  的计算公式如下:

$$MSE(X_1, X_2) = \frac{1}{H \times W} \sum_{i=0}^H \sum_{j=0}^W (X_1(i, j) - X_2(i, j))^2 \quad (14)$$

结构相似性指数 (SSIM) 是衡量图像相似度的指标. 它是一个 0-1 之间的数,这个数越大,表示修复图

像和真实图像之间越相似, 图像更逼真. 假设两张图像分别是 $X_1$ 和 $X_2$ , 计算公式如下:

$$SSIM(X_1, X_2) = \frac{(2\mu_{X_1}\mu_{X_2} + c_1)(2\sigma_{X_1X_2} + c_2)}{(\mu_{X_1}^2 + \mu_{X_2}^2 + c_1)(\sigma_{X_1}^2 + \sigma_{X_2}^2 + c_2)} \quad (15)$$

其中,  $\mu_{X_1}$ 和 $\mu_{X_2}$ 分别为 $X_1$ 和 $X_2$ 的均值,  $\sigma_{X_1}^2$ 和 $\sigma_{X_2}^2$ 分别为 $X_1$ 和 $X_2$ 的方差,  $\sigma_{X_1X_2}$ 为 $X_1$ 和 $X_2$ 的协方差,  $c_1$ 和 $c_2$ 为两个常数, 避免除零.

本文的模型在 CelebA 和 Place2 两个数据集上进行训练和评估, 通过大量数据对模型进行训练, 最后进行测试并计算该模型在不同数据集上的 PSNR 和 SSIM 平均值, 并与基于上下文注意力的生成式对抗模型和基于边缘连接的生成式对抗模型进行比较, 以上 3 种修复方法在数据集上的 PSNR 和 SSIM 平均数值对比如表 1、表 2 所示.

表 1 不同修复方法在 CelebA 上的 PSNR 和 SSIM 平均指数对比

方法	PSNR	SSIM
上下文注意力	19.7	0.824
边缘连接	25.28	0.846
本文方法	26.13	0.852

表 2 不同修复方法在 Place2 上的 PSNR 和 SSIM 平均指数对比

方法	PSNR	SSIM
上下文注意力	18.91	0.802
边缘连接	23.46	0.834
本文方法	24.06	0.841

由表 1 和表 2 可以看出本文方法的 PSNR 和 SSIM 平均值比其他两种方法的平均值高, 实验证明本文的算法在 CelebA 和 Place2 数据集上的修复结果要优于其余两种修复方法.

本文的方法在 Place2 和 CelebA 数据集上的修复结果如图 4 所示, 由图可以看出修复后得到的图像和真实图像在颜色、结构、风格上基本一致, 且图像的纹理细节构建比较准确和清晰; 缺失区域与背景区域之间边界自然清晰, 视觉效果一致, 没有修复过的痕迹. 这表明本文的方法对于图像修复任务有明显的效果.

不同修复算法在 Place2 和 CelebA 数据集上的修复结果如图 5 所示, 由图 5 可以看出, 在 CelebA 数据集的图像中, 对比人面部表情以及五官的修复效果来看, 本文的方法修复效果最好; 可以看到图 5(c)(上下文注意力的方法)中, 缺失区域的修复在视觉上比较模糊;

在图 5(d)(边缘连接的方法)中, 可以明显看到修复的结果出现了偏差, 从图中可以看出修复后的人脸面部稍微扭曲, 并且出现五官与原图偏差的现象.

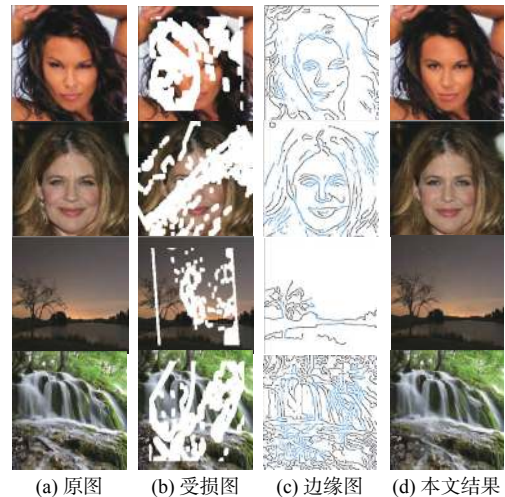


图 4 本文的修复方法结果图

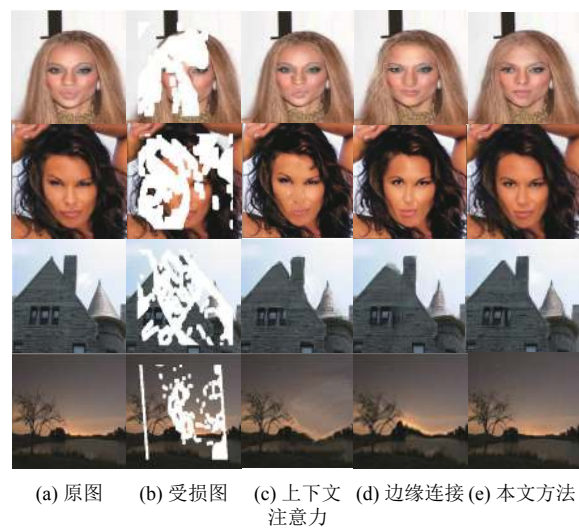


图 5 不同方法修复结果对比图

在 Place2 数据集的图像中, 从房屋窗户和圆锥建筑来看, 本文的方法修复效果比较好, 窗户和圆锥建筑的基本轮廓还有纹理与原图都是最接近的; 图 5(c)中, 可以看到屋顶和烟囱都比较模糊, 圆锥建筑修复结果比较好; 在图 5(d)中, 可以看到窗户和圆锥建筑基本轮廓不清晰, 修复结果比较模糊. 最后一幅图像中, 上下文注意力和边缘连接的修复视觉上比较模糊, 且河水出现了扭曲, 说明修复出现了偏差. 总体来说, 本文方法的修复结果最佳, 在一定程度上缓解了图像边界模糊的现象, 视觉效果上比较一致、连贯, 缺失区域的纹

理修复准确且清晰。

#### 4 结束语

本文提出了一种由边缘到精细的两阶段模型,引入边缘检测和自注意力机制来解决图像修复任务中存在的视觉效果差、边界模糊和纹理不清晰等问题。本文基于生成式对抗网络的模型,引入边缘检测得到受损图像的边缘补全图,在纹理修复网络中加入自注意力机制,利用自注意力具有全局视野的特点,让纹理修复网络生成了更加精确的纹理图像。通过实验结果表明,本文的方法在不同场景图像的修复任务中,能够生成边界清晰,纹理精确,视觉连贯的图像,表现了出色的性能。

#### 参考文献

- 1 Bertalmio M, Sapiro G, Caselles V, *et al.* Image inpainting. Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques. New Orleans, LA, USA. 2000. 417–424.
- 2 Shen JH, Chan TF. Mathematical models for local nontexture inpaintings. *SIAM Journal on Applied Mathematics*, 2002, 62(3): 1019–1043. [doi: [10.1137/S0036139900368844](https://doi.org/10.1137/S0036139900368844)]
- 3 Criminisi A, Pérez P, Toyama K. Region filling and object removal by exemplar-based image inpainting. *IEEE Transactions on Image Processing*, 2004, 13(9): 1200–1212. [doi: [10.1109/TIP.2004.833105](https://doi.org/10.1109/TIP.2004.833105)]
- 4 Zhu JY, Park T, Isola P, *et al.* Unpaired 19-7880-to-image translation using cycle-consistent adversarial networks. Proceedings of 2017 IEEE International Conference on Computer Vision. Venice, Italy. 2017. 2242–2251.
- 5 Isola P, Zhu JY, Zhou TH, *et al.* 19-7880-to-image translation with conditional adversarial networks. Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, HI, USA. 2017. 5967–5976.
- 6 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, HI, USA. 2017. 105–114.
- 7 李天成, 何嘉. 一种基于生成对抗网络的图像修复算法. *计算机应用与软件*, 2019, 36(12): 195–200, 267. [doi: [10.3969/j.issn.1000-386x.2019.12.031](https://doi.org/10.3969/j.issn.1000-386x.2019.12.031)]
- 8 曹仰杰, 贾丽丽, 陈永霞, 等. 生成式对抗网络及其计算机视觉应用研究综述. *中国图象图形学报*, 2018, 23(10): 1433–1449. [doi: [10.11834/jig.180103](https://doi.org/10.11834/jig.180103)]
- 9 杨柳, 王敏, 林竹. 基于新生成器结构的图像修复方法. *计算机系统应用*, 2020, 29(1): 158–163. [doi: [10.15888/j.cnki.csa.007234](https://doi.org/10.15888/j.cnki.csa.007234)]
- 10 Pathak D, Krähenbühl P, Donahue J, *et al.* Context encoders: Feature learning by inpainting. Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, NV, USA. 2016. 2536–2544. [doi: [10.1109/cvpr.2016.278](https://doi.org/10.1109/cvpr.2016.278)]
- 11 Iizuka S, Simo-Serra E, Ishikawa H. Globally and locally consistent image completion. *ACM Transactions on Graphics*, 2017, 36(4): 107. [doi: [10.1145/3072959.3073659](https://doi.org/10.1145/3072959.3073659)]
- 12 Liu HY, Jiang B, Xiao Y, *et al.* Coherent semantic attention for image inpainting. Proceedings of 2019 IEEE/CVF International Conference on Computer Vision. Seoul, Republic of Korea. 2019. 4169–4178. [doi: [10.1109/iccw.2019.00427](https://doi.org/10.1109/iccw.2019.00427)]
- 13 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, UT, USA. 2018. 5505–5514.
- 14 Nazari K, Ng E, Joseph T, *et al.* Edgeconnect: Generative image inpainting with adversarial edge learning. arXiv: 1901.00212, 2019.
- 15 Ren YR, Yu XM, Zhang RN, *et al.* Structureflow: Image inpainting via structure-aware appearance flow. Proceedings of 2019 IEEE/CVF International Conference on Computer Vision. Seoul, Republic of Korea. 2019. 181–190. [doi: [10.1109/iccw.2019.00027](https://doi.org/10.1109/iccw.2019.00027)]
- 16 Goodfellow IJ, Pouget-Abadie J, Mirza M, *et al.* Generative adversarial nets. Proceedings of the 27th International Conference on Neural Information Processing Systems. Montreal, QC, Canada. 2014. 2672–2680.
- 17 Xie SN, Tu ZW. Holistically-nested edge detection. Proceedings of 2015 IEEE International Conference on Computer Vision. Santiago, Chile. 2015. 1395–1403.
- 18 Zhang H, Goodfellow I, Metaxas D, *et al.* Self-attention generative adversarial networks. Proceedings of the 36th International Conference on Machine Learning. Long Beach, CA, USA. 2019. 7354–7363.
- 19 Kingma DP, Ba J. Adam: A method for stochastic optimization. arXiv: 1412.6980, 2014.