

## 基于潜在特征重构和注意力机制的人脸图像修复

张苒苒, 刘 恒

安徽工业大学 计算机科学与技术学院, 安徽 马鞍山 243000

**摘要:**本研究针对现有图像修复方法不能有效地分离结构和纹理信息,修复结果往往会出现边界模糊、结构扭曲等伪影问题,提出了基于潜在特征重构和注意力机制的人脸图像修复方法。人脸图像修复方法分为两阶段,第一阶段,通过结构重建器网络提取样式向量,按照 StyleGAN 所述的原理分为粗尺度特征、中尺度特征和精细特征三组,插入到预先训练好的 StyleGAN 生成器中,产生初步的修复结果;第二阶段通过构建纹理生成网络并使用上下文注意力机制,注意力分数由注意力计算模块计算,注意力转移模块根据较高级别特征图和注意力分数来填充较低级别特征图中的对应缺失区域,以细化上一阶段初步的人脸修复结果。在 CelebA-HQ 数据集上的训练并进行测试,本文的方法在定量和定性分析两个方面均优于现有方法。因此,基于潜在特征重构和注意力机制的人脸图像修复方法能够有效地修复缺损人脸图像,大大减少了边界过度平滑和存在纹理伪影的问题。

**关键词:**图像修复; 结构重建; 纹理生成; 注意力

中图分类号:TP183 文献标识码:A doi:10.16055/j.issn.1672-058X.2023.0002.011

### Face Image Restoration Based on Latent Feature Reconstruction and Attention Mechanism

ZHANG Tongtong, LIU Heng

School of Computer Science and Technology, Anhui University of Technology, Anhui Maanshan 243000, China

**Abstract:** In this study, a face image restoration method based on latent feature reconstruction and attention mechanism was proposed to address the problem that existing image restoration methods cannot effectively separate structure and texture information, and the restoration results often show artifacts such as blurred boundaries and distorted structures. The face image restoration method was divided into two stages. In the first stage, the style vectors were extracted through the structural reconstruction network and divided into three groups of coarse-scale features, medium-scale features, and fine features according to the principles described by StyleGAN, which were inserted into the pre-trained StyleGAN generator to produce initial restoration results. In the second stage, by building a texture generation network and using a contextual attention mechanism, the attention score was calculated by the attention calculation module, and the attention transfer module filled in the corresponding missing regions in the lower-level feature images based on the higher level feature images and the attention scores to refine the initial face restoration results from the previous stage. Trained and tested on the CelebA-HQ dataset, the method in this paper outperformed existing methods in both quantitative and qualitative analysis. Thus, the face image restoration method based on latent feature reconstruction and attention mechanism can effectively repair defective face images, greatly reducing the problems of excessive smooth boundaries and the presence of texture artifacts.

**Keywords:** image restoration; structural reconstruction; texture generation; attention

收稿日期:2021-12-31 修回日期:2022-01-06 文章编号:1672-058X(2023)02-0073-06

基金项目:安徽省自然科学基金资助(2008085MF190);安徽省高校协同创新项目(GXXT-2019-018).

作者简介:张苒苒(1996—),女,安徽黄山人,硕士研究生,从事计算机视觉研究.

引用格式:张苒苒,刘恒.基于潜在特征重构和注意力机制的人脸图像修复[J].重庆工商大学学报(自然科学版),2023,40(2):73—78.

ZHANG Tongtong, LIU Heng. Face image restoration based on latent feature reconstruction and attention mechanism[J]. Journal of Chongqing Technology and Business University (Natural Science Edition), 2023, 40(2): 73—78.

## 1 引言

图像修复是指对图像的损坏区域(也称为缺失区域或孔洞)进行修补的过程,产生完整的、视觉上可信的图像。图像修复被广泛应用于图像处理领域,包括物体移除、照片修复和编辑照片内容。具体地说,该任务侧重于恢复给定掩模的缺损图像,掩模表示缺失的区域。图像修复任务的主要挑战之一是生成准确的图像结构和逼真的纹理细节。

图像修复的方法大致可以分为两类,第一种是传统方法,另一种是基于深度网络的方法。传统的图像修复方法是从周围借用内容或纹理来填充缺损区域,包括基于扩散和基于块的方法。基于扩散的方法通过将局部像素强度的高阶导数从外部传递到受损区域<sup>[1,2]</sup>,或者通过最小化修复模型的高阶偏微分方程来重建目标区域<sup>[3]</sup>。基于块的算法,它通过借用周围区域的纹理来填补缺损区域<sup>[4-6]</sup>,从掩模边界上的像素最相似的背景区域搜索图像块来逐步填充掩模内的像素。虽然与基于扩散的方法相比,基于块的方法可以在大掩模的情况下生成更合理的结果,但当背景信息复杂时,基于块的方法可能会产生线条断开等不合理的修复结果。人脸图像包含着丰富的细节信息,因此传统方法中仅用背景区域的图案信息找到缺失区域的像素信息的方法很难修复损坏区域。

近年来,基于深度网络的方法被提出,利用卷积编解码器网络将修复问题描述为一个条件图像生成问题,编码器学习图像的潜在特征表示,然后解码器利用该表示对丢失的内容进行推理。PATHAR<sup>[7]</sup>提出了上下文编码器方法,这是一种开创性的方法。文献<sup>[8]</sup>通过引入生成对抗性网络来产生更接近现实的结果。基于深度网络的方法可以通过从深层学习语义信息来生成合理的结果。但是,这些方法不能有效地分离结构和纹理信息,它们无法有效地利用结构(如边缘和低频信息),从而生成不合理的结果,其中包括伪影、扭曲的结构和模糊的纹理。

为了解决上述问题,人们提出了一些两阶段模型。这些方法在第一阶段重建缺损区域的结构,在第二阶段利用初步的重建信息生成纹理,得到最终修复好的结果。Yu<sup>[9]</sup>提出的方法中使用真实图像合成模糊图像作为修复缺失区域的指导,但是真实图像包含高频纹理,这些不相关的细节可能会误导结构重建。SONG<sup>[10]</sup>使用图像分割图作为结构信息来指导像素的生成。但是,语义标签相似的区域会有不同的纹理,比如,同一建筑物的窗户和墙壁。NAZERI<sup>[11]</sup>提出的方法由一个边缘生成器和一个图像完成网络组成,图像完成网络利用边缘生成器生成的边缘图像作为结构指导进行

图像修复,但边缘提取器会丢弃太多有用的信息(例如图像颜色),因此很难生成生动的纹理。REN<sup>[12]</sup>提出了结构流,通过应用结构流来构建掩模和有效区域之间的关联,以生成生动的纹理。虽然这些方法可以在一定程度上获取全局信息,但信息搜索的范围是有限的。

基于以上研究,本文提出了一种新的基于潜在特征重构和注意力机制两阶段的人脸图像修复模型,第一阶段产生初步的修复结果,第二阶段通过添加纹理对第一阶段初步修复结果进行细化。利用PSP<sup>[13]</sup>生成网络,直接生成一系列重构的风格特征向量,输入到预先训练好的StyleGAN生成器<sup>[14-15]</sup>中,得到目标图像。本文通过引入掩模内外区域的 $L_1$ 损失,将其应用在图像修复任务中,作为第一阶段的结构重建器网络。在第二阶段,构建纹理生成网络并使用上下文注意力机制,根据较高级别特征图和注意力得分来填充较低级别特征图中的对应缺失区域,以细化上一阶段产生的初步人脸修复结果。

综上,本文提出了一种新的基于潜在特征重构和注意力机制两阶段的人脸图像修复模型,第一阶段通过重构风格特征向量并输入到StyleGAN生成器中,产生初步的修复结果,第二阶段通过添加纹理对第一阶段初步修复结果进行细化,得到最终的人脸修复结果。通过主客观实验,与几种常见的方法进行比较,结果表明,本文的方法能够达到更好的修复效果。此外,本文还进行了网络结构消融研究,以验证本文提出模型方法的有效性。

## 2 网络模型架构

图1是网络的总体结构图。该方法由结构重建器和纹理生成器两部分组成。输入缺损图像,结构重建器填充缺损图像的掩模区域。生成的拼接图片送到纹理生成器,得到修复好的图像。最终得到的输出结果掩模内是修复好的图像,掩模外的区域是原始输入的副本。

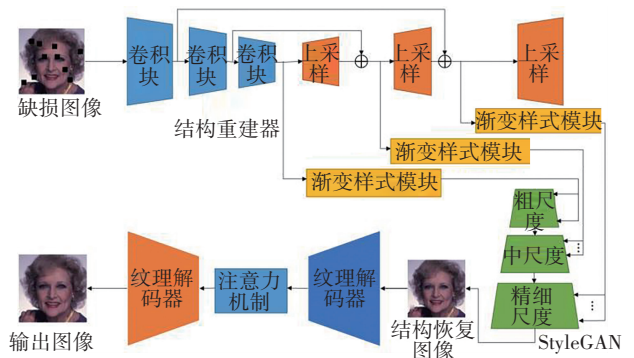


图1 总体网络结构图

Fig. 1 The diagram of the overall network structure

### 3 结构重建器

#### 3.1 模型架构

图像修复的一个核心挑战是保持全局语义结构,所以本方法在第一阶段设计了一个结构重建器来生成恢复图像的全局结构。

在 StyleGAN<sup>[14-15]</sup>中,作者已经证明了不同的样式向量对应于不同层次的特征,大致可以分为粗尺度特征、中尺度特征和精细特征三组。即底层样式向量控制姿态、面部形状,本文称为粗尺度特征。中层样式向量控制较精细的面部特征如眼睛的睁开或闭合、头发的样式,本文称为中尺度特征。高层样式向量控制肤色如头发颜色、背景色等,本文称为精细特征。StyleGAN 通过不同尺度的特征控制了图像不同的高级属性的生成,从粗尺度特征(脸型)到精细的特征(头发颜色),不会影响其他的层级特征。

结构重建器中编码器网络使用了 ResNet 主干上的一个标准特征金字塔结构<sup>[16]</sup>,按照 StyleGAN 所述的分三组提取了 18 个样式向量,其中样式向量 0、样式向量 1 和样式向量 2 是由小尺寸特征图生成,样式向量 3、样式向量 4、样式向量 5 和样式向量 6 是从中尺寸特征图中生成,样式向量 7 到样式向量 17 这 11 个特征向量是从最大的特征图生成。将生成的 18 个样式向量插入到预先训练好的 StyleGAN 生成器中,得到初步的恢复结果。如图 1 所示,渐变样式模块是一个小型的卷积网络,由一组步长为 2 的卷积加一个 LeakReLU 激活函数构成。

#### 3.2 结构重建损失函数

编码器用五种类别的损失函数来训练结构重建网络,获得初步修复结果。

首先,给定掩模  $M$ 、修复结果  $I_r$  和真实图像  $I_{gt}$ ,本文使用  $L_2$  范数来计算重建损失:

$$L_2 = \|I_{gt} - I_r\|_2$$

为了了解感知相似性,本文还使用 LPIPS<sup>[17]</sup> 损失,与标准的感知损失<sup>[18]</sup> 相比,LPIPS 损失能更好地保持图像质量。具体感知损失计算如下:

$$L_{s-prec} = \|F(I_{gt}) - F(I_r)\|_2$$

其中, $F(\cdot)$  表示感知特征提取器。

人脸生成任务的主要挑战之一是在输入和输出图像之间保持人脸身份的能力。身份保留是人脸重建任务的关键部分,由于前述损失函数对面部身份的保持不那么敏感,本文引入了一个专用的识别损失来度量输出图像与其真实图像之间的余弦相似性,作为身份识别损失。

$$L_{ID} = 1 - \langle R(I_{gt}), R(I_r) \rangle$$

其中, $R(\cdot)$  是用于人脸识别的预先训练的 ArcFace<sup>[19]</sup> 网络。

另外,本文使用了图像修复的两种经典损失,并将生成图像和真实图像的掩模内像素和掩模外像素的  $L_1$  损失分别定义为  $L_{hole}$  和  $L_{valid}$ :

$$L_{hole} = \|(1-M) \odot (I_{gt} - I_r)\|_1$$

$$L_{valid} = \|(M) \odot (I_{gt} - I_r)\|_1$$

总而言之,结构重建器的总损失函数表示为

$$L_S = L_2 + 0.8 * L_{s-prec} + 0.1 * L_{ID} + 6 * L_{hole} + L_{valid}$$

### 4 纹理生成器

#### 4.1 模型架构

图 2 说明了纹理生成器的总体流程。它使用结构重建器得到的初步修复图像  $I_r$  作为输入,注意力分数由注意力计算模块计算,注意力转移模块根据较高级别特征图和注意力得分来填充较低级别特征图中的对应缺失区域。最后,将得到的结果添加到上采样的修复结果中,掩模外区域复制原始输入。

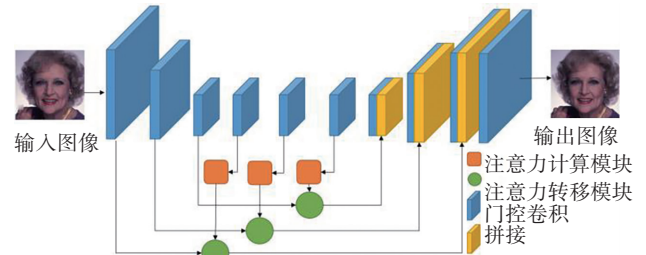


图 2 纹理生成器结构图

Fig. 2 The structure of texture generator

#### 4.2 上下文注意力机制

注意力计算模块先将第  $l$  层的特征图  $P$  划分为小块,计算掩模区域内部块和外部块之间的余弦相似度:

$$c_{i,j}^l = \left( \frac{p_i^l}{\|p_i^l\|}, \frac{p_j^l}{\|p_j^l\|} \right)$$

其中, $l$  表示是层号, $l \in 4, 5, 6$ ,  $p_i^l$  是从第  $l$  层的特征图  $P$  在掩模外区域提取的第  $i$  个块,  $p_j^l$  是从第  $l$  层的特征图  $P$  在掩模内区域的提取的第  $j$  个块。然后对相似性分数应用 Softmax,以获得每个块的注意力分数:

$$s_{i,j}^l = \frac{e^{c_{i,j}^l}}{\sum_{i=1}^N e^{c_{i,j}^l}}$$

其中, $N$  是掩模区域外部块的数量。

注意力转移模块在从特征图获得注意力分数之后,可以用注意力分数加权的上下文块来填充较低级别特征图中的对应缺失区域:

$$p_j^1 = \sum_{i=1}^N s_{i,j}^6 p_i^1$$

$$p_j^2 = \sum_{i=1}^N s_{i,j}^5 p_i^2$$

$$p_j^3 = \sum_{i=1}^N s_{i,j}^4 p_i^3$$

其中,  $p_i^l$  是从第  $l$  层的特征图  $P$  在掩模外区域提取的第  $i$  个块,  $p_j^l$  是从第  $l$  层的特征图  $P$  在掩模内区域的提取的第  $j$  个块,  $l$  表示是层号,  $l \in [1, 2, 3]$ 。  $N$  是缺失区域外部块的数量。在计算完所有掩模区域内的块后, 最终可以得到掩模区域的特征图。如图 2 所示, 得到的特征图送到解码器中对应的层, 与上一层生成的特征图进行拼接, 再送入卷积中, 最终得到修复的结果。

### 4.3 门控卷积

在图像特征提取过程中, 普通卷积是将网络输入的每个像素全部都视为是有效的。而对于修复任务来说, 掩模内区域的像素在图像提取特征的过程中应该被视为无效像素, 掩模内区域的像素应该被视为有效像素。LIU<sup>[20]</sup> 表明: 对缺损图像的掩模内区域直接采用普通卷积的方式, 将其当成和其他区域的像素一样处理, 随着网络层加深, 必然会导致修复结果产生边界模糊和颜色不一致等情况。部分卷积的提出使得卷积结果仅依赖于有效像素, 有效提高了非规则掩模上图像修复的质量。但仍然存在一些问题, 无论前一层二进制掩模有多少个有效像素, 下一层的掩模的像素值都将设置 1。同时, 网络的每个层中的所有通道都共享同一个掩模, 这限制了灵活性。因此在本文的纹理生成器中引入了门控卷积<sup>[21]</sup>。门控卷积的输出可以表示为:

$$Gating = conv(W_g, I)$$

$$Feature = conv(W_f, I)$$

$$O = \sigma(Gating) \odot \Psi(Feature)$$

其中,  $I$  是输入的特征图,  $\sigma$  和  $\Psi$  分别是 Sigmoid 和 ELU 激活函数。  $\odot$  是哈达玛积, 最终输出值在  $[0, 1]$  范围内。  $W_g$  和  $W_f$  是两组不同的卷积滤波器。

### 4.4 纹理生成器损失函数

纹理生成网络用三种类型的损失函数进行训练。首先, 给定掩模  $M$ 、结果  $I_{out}$  和真实图像  $I_{gt}$ , 本文使用像素  $L_1$  损失:

$$L_1 = \|I_{gt} - I_{out}\|_1$$

本文使用 WGAN-GP 损失作为对抗损失<sup>[22]</sup>, 这加强了纹理生成器的全局一致性。用  $L_d$  定义的损失训练鉴别器。

$$L_d = \mathbb{E}_{\tilde{x} \in P_g} [D(\tilde{x})] - \mathbb{E}_{x \in P_g} [D(x)]$$

其中,  $D(\cdot)$  是鉴别器输出。  $x$  是真实图像,  $\tilde{x}$  是最后的

输出图像。  $P_g$ 、 $P_r$  分别是它们的对应分布。

$$L_{adv} = \mathbb{E}_{\tilde{x} \in P_g} [D(\tilde{x})]$$

本文还使用了感知损失, 其中  $P(\cdot)$  表示基于 VGG16 的感知特征图提取网络。

$$L_{l-prec} = \|P(I_{gt}) - P(I_r)\|_2$$

总之, 本文的纹理生成器的总损失函数表示为

$$L_T = 10 * L_{l-prec} + L_{adv} + 100 * L_1$$

## 5 实验结果与分析

本节提供了定性、定量实验结果及比较分析。其中实验配置如下: 训练所需的人脸数据集为 CelebA-HQ 数据集<sup>[23]</sup>; 掩模数据集为正方形掩模, 数量为 10, 缺损图像的样式如图 1 所示; 实验的图形显卡平台为 NVIDIA 2080TI GPU; 结构重建器训练期间的优化器为 Adam, 学习率恒定设置为  $10e-3$ ; 纹理生成器训练期间的优化器为 Adam, 初始学习率设置为  $10e-4$ , 每 100 000 步衰减一次, 基数为 0.96; 结构重建器批大小设置为 2, 纹理生成器批大小设置为 1; 输入训练图片大小为  $256 \times 256$ 。

### 5.1 定性结果及分析

图 3 中展示了与其他方法比较的定性比较结果: 第一列为 PCONV<sup>[20]</sup> 的结果, 第二列为 RFR<sup>[24]</sup> 的结果, 第三列为本文的方法输出的结果, 第四列为原始图像。如图 3 所示, 实验结果来看, 与这些方法相比, 本文的方法的结果能得到的图像具有更合理的结构和更逼真的纹理。

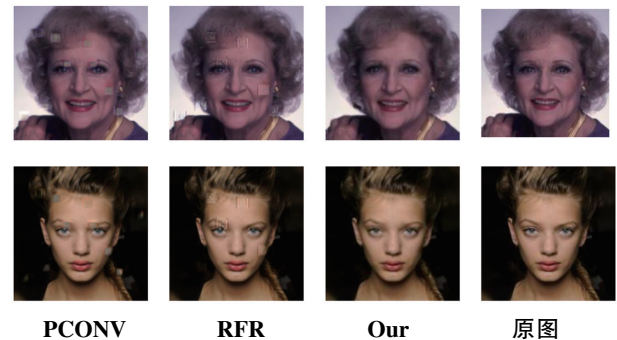


图 3 与其他方法的对比

Fig. 3 Comparisons with other methods

### 5.2 定量结果及分析

借鉴以往的图像修复方法, 本文使用峰值信噪比 (PSNR) 和结构相似度 (SSIM)<sup>[25]</sup> 等常用的评价指标来度量修复结果与真实图片之间的相似度。图像修复方法应该专注于产生逼真的结果, 而不仅仅是近似真实图片。因此, 本文还使用 Fréchet 初始距离分数 (FID)<sup>[26]</sup>, 从图像的计算机视觉特征的统计方面的相似度来衡量生成图像和真实图像的相似度, 该指标与人类的判断是一致的。

PSNR 和 SSIM 越大,表明图片失真越小,修复结果越好。FID 越小,表明生成图像具有较高的质量和多样性。从表 1 中可以观察到:在 Celeba-HQ 数据集中,此办法评估结果最好。这表明本文的方法可以更好地处理人脸图像的修复问题。

表 1 定量分析

Table 1 Quantitative analysis

度量因子	方 法		
	PCONV	RFR	Our
PSNR	33.058	32.252	<b>36.733</b>
SSIM	0.966	0.948	<b>0.977</b>
FID	67.956	106.312	<b>15.077</b>

### 5.3 消融实验

本文给出了提出模型的两种变体,并在 CelebA-HQ 数据集上进行了消融实验,使用的是方形掩模。与以往的图像修复工作一样,采用 PSNR、SSIM 和 FID 来衡量模型。

变体一是只使用了纹理生成器,且纹理生成器中的注意力分数只来自第六层,如图 4 所示,变体二是加入了结构重建器,纹理生成器中的注意力分数只来自第六层。与本文的方法的区别在于:本文方法的纹理生成器中的注意力来自四到六这三层。

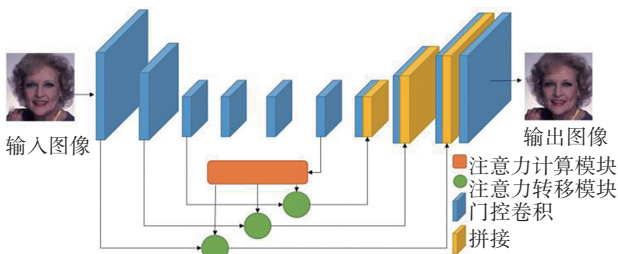


图 4 变体一示意图

Fig. 4 The Schematic diagram of variant one

图 5 是两种变体和本文方法以及原图定性的比较。表 2 展示了定量的结果,变体一和变体二的结果比较可以看出结构重建器的必要性,变体二和本文的方法的结果比较可以看出用多层特征图计算注意力分数可以得到更好的结果。

表 2 消融实验定量分析

Table 2 Quantitative analysis of ablation experiment

度量因子	方 法		
	变体一	变体二	Our
PSNR	34.364	36.239	<b>36.733</b>
SSIM	0.959	0.973	<b>0.977</b>
FID	33.357	20.534	<b>15.077</b>



图 5 消融实验结果的对比

Fig. 5 The Comparison of the results of ablation experiments

## 6 结论与展望

本文提出了一种新的基于潜在特征重构和注意力机制的两阶段人脸图像修复模型。首先,使用结构重建网络来生成初步的人脸修复结果。之后,纹理生成器通过添加纹理来细化修复。实验结果显示,相较于现有的图像修复方法,本文的方法在定量和定性两个方面均有不错的结果,有效改善了之前的方法易产生边界过度平滑或纹理伪影的问题,提升了图像的修复质量。

本方法聚焦于小型掩模的缺损人脸图像进行修复。在未来,我们的研究还将集中在大尺寸掩膜图像修复上,使研究内容更加实用化。

### 参考文献(References):

- [1] BERTALMIO M, BERTOZZI A L, SAPIRO G. Navier-stokes, fluid dynamics, and image and video inpainting[C]// Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. IEEE, 2001, 1: 355—362.
- [2] LI F, PI L, ZENG T. Explicit coherence enhancing filter with spatial adaptive elliptical kernel[J]. IEEE Signal Processing Letters, 2012, 19(9): 555—558.
- [3] CHAN T F, SHEN J. Nontexture inpainting by curvature-driven diffusions [J]. Journal of Visual Communication and Image Representation, 2001, 12(4): 436—449.
- [4] DRORI I, COHEN-OR D, YESHURUN H. Fragment-based image completion[M]//America: ACM SIGGRAPH, 2003.
- [5] WILCZKOWIAK M, BROSTOW G J, TORDOFF B, et al. Hole filling through photomontage [C]//Proceedings of the British Machine Vision Conference. IEEE, 2005: 492—501.
- [6] EFROS A A, FREEMAN W T. Image quilting for texture synthesis and transfer [C]//Proceedings of the 28th Annual Conference on Computer Graphics and Interactive Techniques.

- IEEE, 2001: 341—346.
- [7] 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. IEEE, 2016: 2536—2544.
- [8] YE H R A, CHEN C, YIAN LIM T, et al. Semantic image inpainting with deep generative models [C]//Proceedings of the IEEE conference on computer vision and pattern recognition. IEEE, 2017: 5485—5493.
- [9] YU J, LIN Z, YANG J, et al. Generative image inpainting with contextual attention [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2018: 5505—5514.
- [10] SONG Y, YANG C, SHEN Y, et al. Spg-net: segmentation prediction and guidance network for image inpainting [J]. arXiv e-prints, 2018, 5: 03356.
- [11] NAZERI K, NG E, JOSEPH T, et al. Edgeconnect: generative image inpainting with adversarial edge learning[J]. arXiv e-prints, 1901, 1: 00212.
- [12] REN Y, YU X, ZHANG R, et al. Structureflow: image inpainting via structure-aware appearance flow [C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. IEEE, 2019: 181—190.
- [13] RICHARDSON E, ALALUF Y, PATASHNIK O, et al. Encoding in style: a stylegan encoder for image-to-image translation[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. IEEE, 2021: 2287—2296.
- [14] KARRAS T, LAINE S, AILA T. A style-based generator architecture for generative adversarial networks [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. IEEE, 2019: 4401—4410.
- [15] KARRAS T, LAINE S, AITTALA M, et al. Analyzing and improving the image quality of stylegan [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. IEEE, 2020: 8110—8119.
- [16] PARK D Y, LEE K H. Arbitrary style transfer with style-attentional networks [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. IEEE, 2019: 5880—5888.
- [17] ZHANG R, ISOLA P, EFROS A A, et al. The unreasonable effectiveness of deep features as a perceptual metric [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2018: 586—595.
- [18] JOHNSON J, ALAHI A, FEI-FEI L. Perceptual losses for real-time style transfer and super-resolution [C]//European Conference on Computer Vision. Springer, Cham, 2016: 694—711.
- [19] DENG J, GUO J, XUE N, et al. Arcface: additive angular margin loss for deep face recognition [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. IEEE, 2019: 4690—4699.
- [20] LIU G, REDA F A, SHIH K J, et al. Image inpainting for irregular holes using partial convolutions [C]//Proceedings of the European Conference on Computer Vision (ECCV). IEEE, 2018: 85—100.
- [21] YU J, LIN Z, YANG J, et al. Free-form image inpainting with gated convolution [C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. IEEE, 2019: 4471—4480.
- [22] GULRAJANI I, AHMED F, ARJOVSKY M, et al. Improved training of Wasserstein GANs [J]. arXiv e-prints, 2017, 4: 0028.
- [23] KARRAS T, AILA T, LAINE S, et al. Progressive growing of gans for improved quality, stability, and variation [J]. arXiv e-prints, 2017, 10: 10196.
- [24] LI J, WANG N, ZHANG L, et al. Recurrent feature reasoning for image inpainting [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. IEEE, 2020: 7760—7768.
- [25] WANG Z, BOVIK A C, SHEIKH H R, et al. Image quality assessment: from error visibility to structural similarity [J]. IEEE Transactions on Image Processing, 2004, 13(4): 600—612.
- [26] HEUSEL M, RAMSAUER H, UNTERTHINER T, et al. GANs trained by a two time-scale update rule converge to a local Nash equilibrium [J]. Advances in Neural Information Processing Systems, 2017: 6626—6637.