

基于特征融合与嵌入的人脸图像盲修复算法

霍智勇¹, 胡山林²

(1. 南京邮电大学教育科学与技术学院, 南京 210003; 2. 南京邮电大学通信与信息工程学院, 南京 210003)

摘要: 人脸图像盲修复是从未知退化中恢复出高质量的人脸图像, 其不适应性往往会造成修复出的图像出现局部纹理缺失或面部成分不匹配的结果, 为此提出基于特征融合与嵌入的人脸图像盲修复算法。通过提取退化输入的面部先验特征, 采用多头交叉注意力进行特征交互融合和全局上下文建模, 将面部先验嵌入预训练生成网络的潜在空间中, 并基于损失函数进行优化, 修复因退化而丢失或损坏的局部纹理, 实现真实性与忠实度之间的平衡。数值实验在3个真实退化图像数据集上进行, 本文方法在客观指标和主观质量上都优于现有方法, 最后的消融实验验证了退化人脸图像盲修复算法的有效性。

关键词: 盲修复; 特征融合; 特征嵌入; 多头注意力; 预训练生成网络

中图分类号: TP391.41 **文献标志码:** A

Blind Face Restoration Algorithm Based on Feature Fusion and Embedding

HUO Zhiyong¹, HU Shanlin²

(1. Department of Educational Science and Technology, Nanjing University of Posts and Telecommunications, Nanjing 210003, China; 2. Department of Communication and Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing 210003, China)

Abstract: Blind face restoration is to recover high quality face from unknown degradation, and the ill-posed problem often results in local texture missing or mismatched facial components for restored images, therefore a degraded blind face restoration algorithm based on feature fusion and embedding optimization is proposed. By extracting face prior features from degraded inputs, using multi-headed cross-attention for feature interaction fusion and global context modeling, embedding facial priors into the latent space of pre-trained generative networks, and carrying out optimization based on loss functions, local textures lost or damaged due to degradation are repaired to achieve a balance between realism and fidelity. Numerical experiments are conducted on three real degraded datasets, which outperform existing methods in terms of objective metrics and subjective quality, and the final ablation experiments validate the effectiveness of the degraded blind face restoration algorithm.

Key words: blind restoration; feature fusion; feature embedding; multi-head attention; pre-trained generative network

引言

人脸图像盲修复(Blind face restoration, BFR)已广泛应用于旧照片翻新^[1-3]、人脸检测^[4-5]和人脸重建^[6-7]等领域。在未知退化(例如下采样、噪声、模糊和压缩等)或多退化组合情况下,一个低质量输入对应多个可能的高质量输出。现有方法主要分为两类:一类为基于先验的方法,包括几何先验^[7]、参考先验^[1]和生成先验^[3,8],这些额外先验被用来对抗其不适应性。如Chen等^[7]利用面部解析图指导恢复,几何先验只能提供有限纹理信息。Li等^[1-2]预构建高质量面部成分特征字典作为参考先验,但字典范围外的面部成分难以恢复。随着生成对抗网络(Generative adversarial networks, GAN)^[9]的快速发展,Meon等^[8]通过训练进行潜在编码的迭代优化,利用预训练GAN作为生成先验来指导图像修复,Wang等^[3]将退化图像逆映射成潜码并在空间特征上进行局部匹配与嵌入。这些方法忽略了人脸图像中全局结构和上下文,导致整体结构不自然并易缺失局部纹理。另一类为基于非先验的方法,旨在设计一个端到端的网络,直接学习低质量和高质量人脸图像之间的映射函数。如Yang等^[10]通过将人脸复原问题转化为语义引导的生成问题,但无法恢复复杂的局部纹理细节。RestoreFormer^[11]使用多头交叉注意力层学习探索上下文信息建模机制,但重建结果的忠实度较低且计算资源要求高。Hu等^[12]对潜在编码中不同语义特征的依赖关系进行关注,实现精确编辑人脸图像的样式。DIFFACE^[13]将输入图像作为条件引入到扩散模型,但可控性较差,导致难以平衡重建结果的质量与忠实度。

在传统基于生成先验算法和Transformer^[14]的启发下,本文提取退化输入的多分辨率空间特征,以及潜在特征向量作为面部先验,利用多头交叉注意力实现对长距离特征依赖的全局上下文建模,在不同区域之间建立关联,对每个区域自适应地选择相关的上下文进行修复,通过引入多尺度特征融合模块,在不同尺度上进行面部先验融合,并基于损失函数进行优化,从而更好地处理图像中的细节和全局结构,最后将空间特征嵌入预训练GAN潜在空间中,实现真实性与保真度之间的平衡。

1 算法概述

本文算法总体框架如图1所示。图中 H 表示输入图像的分辨率高度, W 表示输入图像的分辨率宽度, C 表示输入图像的特征通道数;VGG(Visual geometry group)表示一种深度卷积神经网络模型;FFT(Fast Fourier transform)即为快速傅里叶变换; $W+$ 潜在编码由 W_1, W_2, \dots, W_{16} 等共16个 1×512

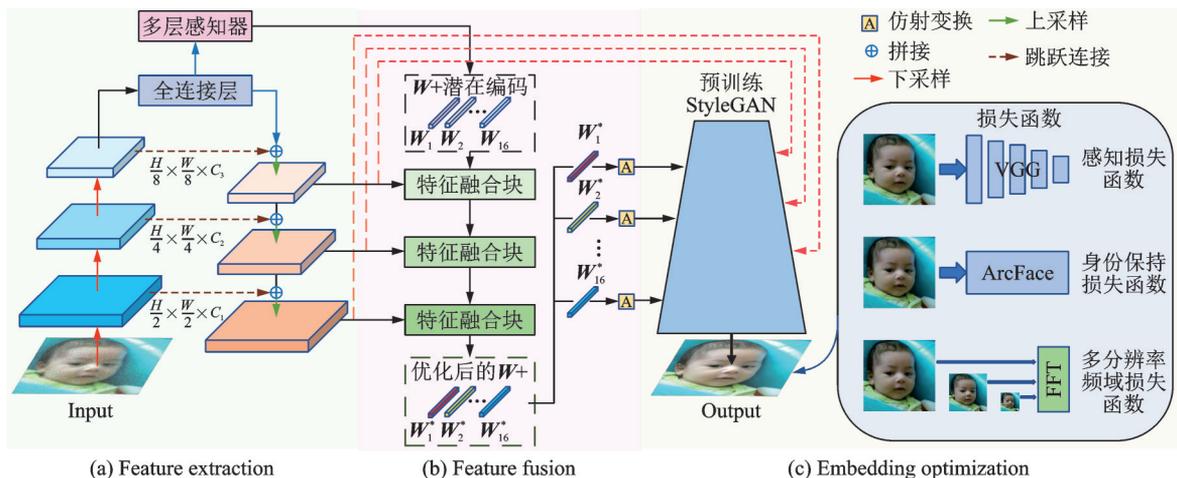


图1 本文算法整体框架图

Fig.1 Overview of framework of the proposed algorithm

维度大小的潜在特征向量组成。首先由编-解码网络生成潜在特征向量和从粗到细的多分辨率空间特征作为面部先验信息,随后通过多级式特征融合块进行从粗到细的面部先验融合,实现全局上下文建模, $W+$ 潜码经过迭代优化后送入预训练 StyleGAN^[15-16]生成器中,空间特征通过空间特征变换(Spatial feature transform, SFT)^[17]对应尺度地嵌入生成器的潜在空间中,最后生成器对 $W+$ 潜码进行解码以生成高质量的重建图像。

1.1 特征融合设计

受到 Transformer 在图像修复^[2,11]和图像编辑^[12,18]任务中取得巨大成功的启发,注意力机制有效解决了卷积运算中感受野有限导致无法对长距离特征依赖关系建模的问题,RestoreFormer^[10]通过交叉注意力进行多尺度局部-全局的特征交互,充分利用长距离上下文。Hu等^[12]关注 $W+$ 潜在在编码中不同语义特征的依赖关系,实现精确控制人脸图像生成。如图2所示,先将多层感知器生成的 16×512 维的 $W+$ 潜在编码视为初始查询 Q 标记,通过多头自注意力进行 $W+$ 潜码之间不同语义特征的正则化,经过自注意力运算后通过线性投影再次作为查询标记,将提取的多分辨率空间特征作为键 K 和值 V ,多头交叉注意力将空间特征与 $W+$ 潜码中语义特征进行查询和键值对之间的交互融合,实现 $W+$ 潜码的一次迭代优化,在潜在空间中寻找出最匹配的 $W+$ 潜码。

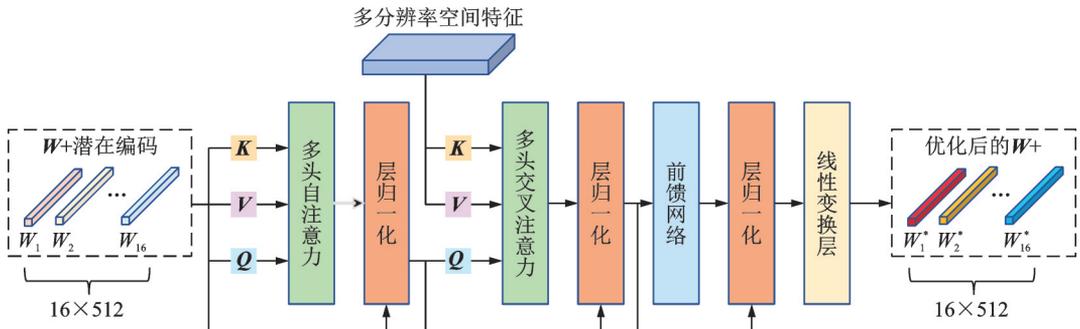


图2 特征融合块内部结构

Fig.2 Internal structure of the feature fusion block

为了进行特征尺度从粗到细的修复,采用了3层不同分辨率大小的特征融合层,让 $W+$ 潜码能从输入图像中吸收足够的局部细节。其中多头交叉注意力表达如式(1~3)所示。

$$Q = X_q W_Q^{crs}, K = X_q W_K^{crs}, V = X_q W_V^{crs} \quad (1)$$

$$\text{Attention}(Q, K, V) = \text{Softmax}\left(\frac{QK^T}{\sqrt{d}}\right) \cdot V \quad (2)$$

$$\text{MHA}(Q, K, V) = [\text{Attention}(Q, K, V)] \cdot W^O \quad (3)$$

式中: Q 为查询矩阵; K 为键值矩阵; V 为值矩阵; q 为512维的查询标记; $X_q \in \mathbf{R}^{16 \times 512}$ 为查询标记的集和; $W_Q^{crs}, W_K^{crs}, W_V^{crs}, W^O \in \mathbf{R}^{512 \times 512}$ 都为可学习映射矩阵; d 代表特征维度且等于 $512/H$, H 为注意力头数量。

1.2 嵌入优化设计

文献[6,8]证明了只使用低维潜在编码难以准确指导人脸修复,无论是采用不同种类编码器或是不同的潜在空间,重建结果仍难以在真实性和忠实度之间达到平衡,因此Wang等^[3]将输入图像的空间特征嵌入预训练GAN潜在空间作为额外的约束条件,因为潜在编码的高级语义特征难以完全地控制

纹理和结构等低级空间特征,就像难用一句话准确描述出一张图像内所有的空间结构信息,最后利用预训练GAN内丰富的先验(例如几何、面部纹理和颜色等)作为指导,在输入的空间特征约束下进行面部细节修复。嵌入优化过程如图3所示,预训练GAN生成器先对 $W+$ 潜在编码进行解码生成中间特征,将多分辨率空间特征通过SFT对GAN的中间特征进行调制,保留输入的空间信息对于几何结构修复非常重要,准确对应的局部特征能保持结果的忠实度,而GAN的中间特征内丰富的先验信息对于保持真实性很重要,因此按特征通道维度进行平均分割,然后按通道维度进行拼接操作,分别用于保持忠实度和真实性,并可以按照任务需求进行不同比例的混合以实现忠实度和真实性之间的良好平衡。其中多分辨率空间特征大小是从 $8 \times 8 \times 256$ 到 $512 \times 512 \times 16$,共进行7次不同特征尺度的嵌入,在每个分辨率尺度上通过2个卷积层从多分辨率空间特征中生成一对仿射参数 (γ, β) ,利用卷积运算的平移不变性对GAN中间特征进行调制,如式(4,5)所示。

$$(\gamma, \beta) = \text{Conv}(F_{\text{spatial}}) \quad (4)$$

$$F_{\text{modulate}} = \text{SFT}(F_{\text{GAN}} | \gamma, \beta) = \gamma \odot F_{\text{GAN}} + \beta \quad (5)$$

式中: F_{spatial} 为多分辨率空间特征; γ 和 β 为仿射参数; F_{GAN} 为GAN中间特征; \odot 为逐像素乘积运算; F_{modulate} 为调制后的输出特征。

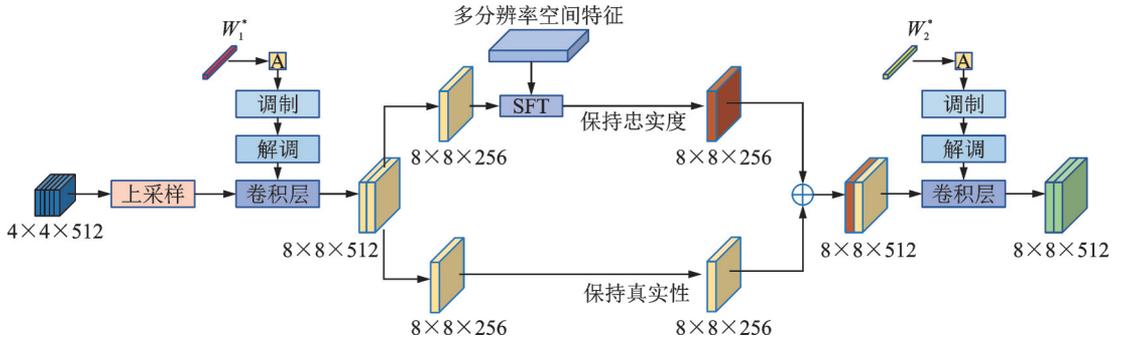


图3 嵌入优化过程

Fig.3 Embedding optimization process

1.3 损失函数设计

传统的像素级损失函数可能无法捕捉到人眼对图像质量的主观感知,为了度量图像在结构、纹理和感知上的差异,采用广泛认可的工作^[19]进行感知损失函数的构建,对VGG-19生成的中间特征进行L1损失计算作为感知损失,如式(6)所示。

$$L_{\text{per}} = \left\| \Phi(\hat{I}) - \Phi(I) \right\|_{L1} \quad (6)$$

式中: $\hat{I}, I \in \mathbb{R}^{H \times W \times C}$ 分别为重建图像和输入图像; $\Phi(I)$ 表示使用VGG-19网络对图像 I 处理,对经过LeakyReLU激活函数前的conv1_2、conv2_2、conv3_2、conv4_2和conv5_2层输出特征进行计算。

由于BFR的不适定性,单个输入可能对应多个输出导致忠实度低,为了增强身份一致性,目前流行方法^[3,11]广泛采用预训练ArcFace^[20]作为人脸识别模型,它能捕获身份信息中最突出的特征,并通过特征匹配作为身份损失函数,如式(7)所示。

$$L_{\text{id}} = \left\| R(\hat{I}) - R(I) \right\|_{L1} \quad (7)$$

式中 R 代表人脸特征提取器,即采用的预训练ArcFace模型。

输入图像由于退化易丢失或损坏大部分的高频信息,而这些往往是恢复图像中结构细节和纹理的

关键,为了弥补输入图像中高频信息的缺失,采用基于FFT的频率域损失函数进行优化,从空域转换为频域进行处理,可以更好地恢复纹理细节和结构信息,其中文献[21]通过在多分辨率特征的频率域进行处理发现拥有更高的鲁棒性,多分辨率频率重建(Multi-resolution frequency reconstruction, MRFR)损失函数如式(8)所示。

$$L_{MRFR} = \sum_{j=1}^J \frac{1}{t_j} \left\| \mathcal{F}(\hat{D}_j) - \mathcal{F}(D_j) \right\|_{L_1} \quad (8)$$

式中: D 、 \hat{D} 分别为输入图像和重建图像提取的多分辨率空间特征; j 代表多分辨率空间特征中的第 j 层; t_j 为生成的特征图累计总层数; \mathcal{F} 代表快速傅里叶变换操作。

综上所述,总的损失函数定义如式(9)所示。

$$L_{total} = \lambda_{per} L_{per} + \lambda_{id} L_{id} + \lambda_{MRFR} L_{MRFR} \quad (9)$$

式中 λ_{per} 、 λ_{id} 和 λ_{MRFR} 分别为对应的损失函数权重系数,预设比例为1:10:1。

2 实验与分析

2.1 实验环境设置

数值实验使用FFHQ数据集^[15]进行训练,构建训练数据时遵循文献[1]的退化模型进行处理来合成退化人脸数据集。在真实世界收集的3种不同退化程度数据集CelebChild-Test(轻度)^[1]、LFW-Test(中度)^[22]和WebPhoto-Test(重度)^[3]上评估算法鲁棒性。使用FID(Fréchet inception distance)^[23]、NIQE(Natural image quality evaluator)^[24]和LPIPS(Learned perceptual image patch similarity)^[25]评估修复结果。FID衡量生成图像与真实图像之间的Fréchet距离,NIQE用于衡量生成图像的感知质量,LPIPS用于表示输入图像与生成图像的相似度。实验计算平台为Quadro RTX 8000显卡,平台环境为Ubuntu18.04系统和Pytorch版本1.7的深度学习框架。训练模型时,设置训练批量大小为8,采用Adam优化器训练网络,总共进行500 000次迭代,设置初始学习率为 $2e^{-3}$ 。

2.2 实验结果与分析

2.2.1 真实世界数据集的人脸图像盲修复结果对比

数值实验在3种不同真实世界数据集上进行,与PSFRGAN^[7]、DFDNet^[1]、DMDNet^[2]、GFPGAN^[3]、PULSE^[8]和DIFFACE^[12]六种算法进行了对比,结果如图4~6所示。

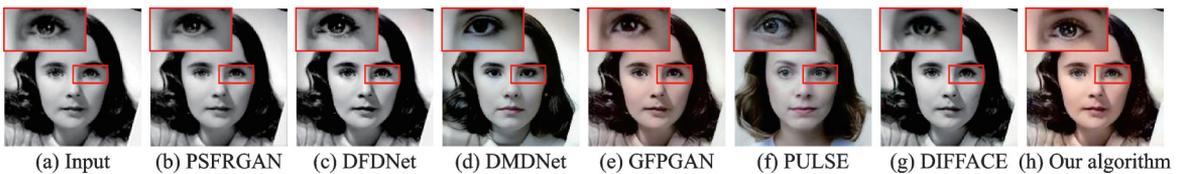


图4 CelebChild-Test数据集的主观比较

Fig.4 Subjective comparison of the CelebChild-Test dataset

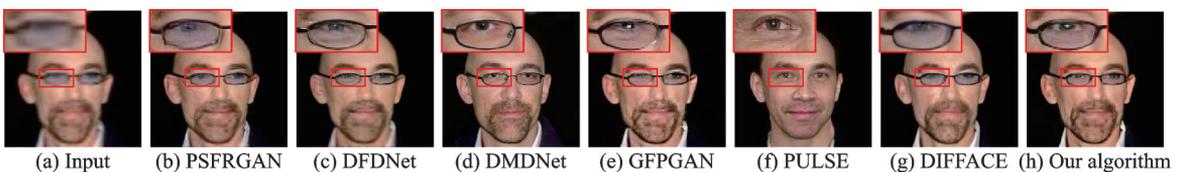


图5 LFW-Test数据集的主观比较

Fig.5 Subjective comparison of the LFW-Test dataset

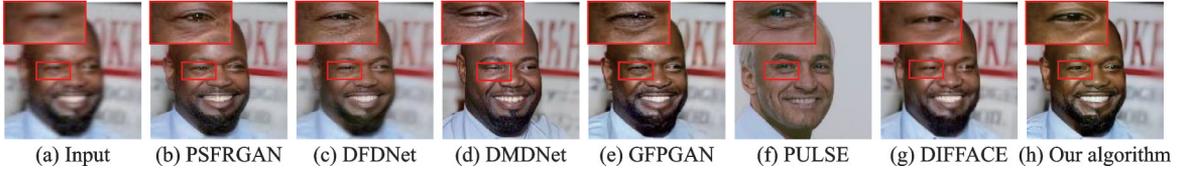


图6 WebPhoto-Test数据集的主观比较

Fig.6 Subjective comparison of the WebPhoto-Test dataset

由图4可以看出,图4(h)因为利用了生成先验中的丰富面部先验信息,不仅逼真地修复图4(a)黑白照片局部(例如眼睛、嘴巴和头发纹理)的纹理细节信息,还进行了颜色增强;DFDNet和DMDNet(面部成分字典)和PSFRGAN(解析图)因其先验中不存在而无法进行精细的局部细节修复和颜色增强,PULSE难以找到准确对应的 $W+$ 潜在编码,导致重建图像的忠实度低。

由图5,6可以看出,在复杂退化图5(a)和图6(a)的人脸图像时,由于使用多头交叉注意力进行了全局上下文建模,图5(h)和图6(h)重建图像不仅在局部区域修复出了丰富的纹理细节,还拥有更自然的整体面部结构。GFPGAN强调局部区域的匹配而忽略了全局上下文信息,导致图5(e)和图6(e)重建图像的左右眼光方向不一致,整体结构不自然。DIFFACE由于低质量输入导致去噪后容易产生不同结果,会生成额外的纹理细节,难以平衡图5(g)和图6(g)重建图像的真实性和保真度。

实验结果比较如表1所示,除了FID指标数值在CelebChild-Test和WebPhoto-Test数据集上高于PULSE,其余均低于其他比较算法。但如图4~6中(f)图所示,PULSE修复的人脸图像拥有高视觉感知质量(即FID数值小),但没有保留足够的身份特征(即NIQE数值较大),由于本文算法进行了面部不同层次先验特征的融合和嵌入,能保留更多的身份信息从而提高忠实度,拥有更小的FID和NIQE,代表本文算法修复的人脸图像拥有更高的视觉质量和保真度。

表1 LFW-Test、CelebChild-Test和WebPhoto-Test数据集上的实验结果对比

Table 1 Numerical result comparison on LFW-Test, CelebChild-Test, and WebPhoto-Test datasets

Method	LFW-Test		CelebChild-Test		WebPhoto-Test	
	FID ↓	NIQE ↓	FID ↓	NIQE ↓	FID ↓	NIQE ↓
PSFRGAN	51.89	5.096	107.40	4.804	88.45	5.582
DFDNet	62.57	4.026	111.55	4.414	100.68	5.293
DMDNet	51.25	4.239	109.34	4.382	92.62	4.715
GFPGAN	49.96	3.882	111.78	4.349	87.35	4.144
PULSE	64.86	5.097	102.74	5.225	86.45	5.146
DIFFACE	50.18	4.057	110.53	4.637	88.16	4.527
Our algorithm	49.36	3.754	107.74	4.329	87.03	4.019

2.2.2 消融实验

为了验证提出算法的有效性,将在3个方面进行消融实验比较:(1)只使用U型网络的解码器部分进行潜在特征提取并去除用于保持局部高频细节的MRFR损失函数,简称w/o MRFR;(2)去除特征融合模块(Feature fusion block,FFB),即不对全局上下文建模和进行 $W+$ 潜在编码迭代优化,简称w/o FFB;(3)不使用预训练GAN作为生成先验,简称w/o GP,通过FID、NIQE和LPIPS指标来衡量改变对算法性能的影响。主观结果对比如图7所示,其中图7(b)中去除MRFR后导致局部(如眼睛和嘴巴)区域的高频细节变少,而图7(c)中去除FFB后重建图像的整体结构不自然且局部细节信息丢失,图7(d)不使用生成先验导致重建图像的质量严重下降。

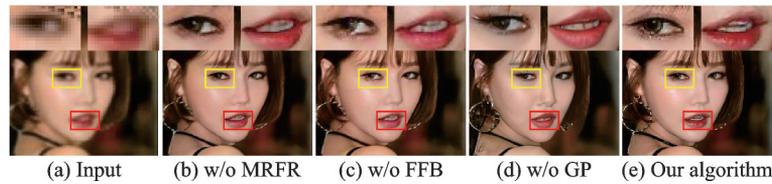


图7 消融实验主观结果

Fig.7 Subjective results of ablation experiments

客观指标结果比较如表2所示,相对于本文完整算法,通过 LPIPS、FID 和 NIQE 指标结果的对比,w/o MRFR 分别下降了 2.4%、15.1% 和 2.2%,表明 MRFR 损失有效保留重建图像的细节信息,w/o FFB 分别下降了 16.0%、20.9% 和 5.2%,表明进行面部先验特征融合和全局上下文建模有效提高忠实度和面部组成整体结构,w/o GP 分别下降了 29.5%、47.2% 和 12.5%,这表明预训练 GAN 作为生成先验有效地指导了修复,对缺失或损坏的局部细节进行补充,从而提高重建图像的感知质量。

表2 消融实验的数值结果

Table 2 Numerical results of ablation experiments

Mehod	LPIPS ↓	FID ↓	NIQE ↓
w/o MRFR	0.366 9	49.96	4.346
w/o FFB	0.415 7	52.43	4.475
w/o GP	0.464 1	63.84	4.784
Our algorithm	0.358 3	43.37	4.253

3 结束语

本文提出了一种基于特征融合与嵌入优化的人脸图像盲修复算法,将预训练 StyleGAN 作为生成先验,为修复提供了丰富的面部特征信息指导,运用多头交叉注意力将空间特征与 $W+$ 潜在在编码进行特征交互融合,并通过 $W+$ 潜码迭代优化,在潜在空间中找到更准确的 $W+$ 潜码,利用全局上下文建模使得重建图像拥有更自然的面部结构和更丰富的局部细节。数值实验结果表明,本文算法在客观指标和主观质量上都优于现有方法。未来将进一步研究如何利用扩散模型作为生成先验或者利用多模态特征(如文本和图像、音频和图像)进行指导图像修复,以适应图像退化更复杂的场景和需求。

参考文献:

- [1] LI X, CHEN C, ZHOU S, et al. Blind face restoration via deep multi-scale component dictionaries[C]//Proceedings of Computer Vision—ECCV 2020: 16th European Conference. Glasgow, UK: Springer International Publishing, 2020: 399-415.
- [2] LI X, ZHANG S, ZHOU S, et al. Learning dual memory dictionaries for blind face restoration[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022, 45(5): 1-13.
- [3] WANG X, LI Y, ZHANG H, et al. Towards real-world blind face restoration with generative facial prior[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. [S.l.]: IEEE, 2021: 9168-9178.
- [4] 刘峰, 居昊, 干宗良. 基于全卷积孪生神经网络的复杂监控场景下前景提取方法[J]. 南京航空航天大学学报, 2021, 53(5): 743-750.
LIU Feng, JU Hao, GAN Zongliang. Foreground extraction method in complex monitoring scenarios based on fully convolutional twin neural network[J]. Journal of Nanjing University of Aeronautics & Astronautics, 2021, 53(5): 743-750.
- [5] 高凌飞, 王海龙, 王海涛, 等. 基于轻量级卷积神经网络的人证比对[J]. 南京航空航天大学学报, 2021, 53(5): 751-758.
GAO Lingfei, WANG Hailong, WANG Haitao, et al. Personnel and witness comparison based on lightweight convolutional neural network[J]. Journal of Nanjing University of Aeronautics & Astronautics, 2021, 53(5): 751-758.
- [6] ABDAL R, QIN Y, WONKA P. Image2StyleGAN: How to embed images into the StyleGAN latent space?[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. [S.l.]: IEEE, 2019: 4432-4441.
- [7] CHEN C, LI X, YANG L, et al. Progressive semantic-aware style transformation for blind face restoration[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. [S.l.]: IEEE, 2021: 11896-11905.

- [8] MENON S, DAMIAN A, HU S, et al. Pulse: Self-supervised photo upsampling via latent space exploration of generative models[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. [S.l.]: IEEE, 2020: 2437-2445.
- [9] GOODFELLOW I J, POUGET-ABADIE J, MIRZA M, et al. Generative adversarial networks[EB/OL]. (2014-06-10)[2023-03-10]. <http://arxiv.org/abs/1406.2661>.
- [10] YANG L, WANG S, MA S, et al. HifaceGAN: Face renovation via collaborative suppression and replenishment[C]//Proceedings of the 28th ACM International Conference on Multimedia. [S.l.]: ACM, 2020: 1551-1560.
- [11] WANG Z, ZHANG J, CHEN R, et al. RestoreFormer: High-quality blind face restoration from undegraded key-value pairs [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. [S.l.]: IEEE, 2022: 17512-17521.
- [12] HU X, HUANG Q, SHI Z, et al. Style transformer for image inversion and editing[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. [S.l.]: IEEE, 2022: 11337-11346.
- [13] YUE Z, LOY C C. DIFFACE: Blind face restoration with diffused error contraction[EB/OL]. (2022-12-13)[2023-03-10]. <https://arxiv.org/abs/2212.06512>.
- [14] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[C]//Proceedings of Conference on Neural Information Processing Systems. Long Beach, CA, USA: [s.n.], 2017.
- [15] 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. [S.l.]: IEEE, 2019: 4401-4410.
- [16] 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. [S.l.]: IEEE, 2020: 8110-8119.
- [17] WANG X, YU K, DONG C, et al. Recovering realistic texture in image super-resolution by deep spatial feature transform [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. [S.l.]: IEEE, 2018: 606-615.
- [18] LIU Z, LIN Y, CAO Y, et al. Swin transformer: Hierarchical vision transformer using shifted windows[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. [S.l.]: IEEE, 2021: 10012-10022.
- [19] JOHNSON J, ALAHI A, FEI-FEI L. Perceptual losses for real-time style transfer and super-resolution[C]//Proceedings of Computer Vision—ECCV. [S.l.]: Springer International Publishing, 2016: 694-711.
- [20] 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. [S.l.]: IEEE, 2019: 4690-4699.
- [21] CHO S J, JI S W, HONG J P, et al. Rethinking coarse-to-fine approach in single image deblurring[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. [S.l.]: IEEE, 2021: 4641-4650.
- [22] HUANG G B, MATTAR M, BERG T, et al. Labeled faces in the wild: A database for studying face recognition in unconstrained environments[C]//Proceedings of Workshop on Faces in Real-Life Images: Detection, Alignment, and Recognition. Marseille, France: [s.n.], 2008.
- [23] 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. DOI: 10.48550/arXiv.1706.08500.
- [24] MITTAL A, SOUNDARARAJAN R, BOVIK A C. Making a “completely blind” image quality analyzer[J]. *IEEE Signal Processing Letters*, 2012, 20(3): 209-212.
- [25] 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. [S.l.]: IEEE, 2018: 586-595.

作者简介:



霍智勇(1973-),男,教授,
研究方向:数字图像处理,
E-mail:huozhy@njupt.edu.cn。



胡山林(1998-),通信作者,
男,硕士研究生,研究方
向:数字图像处理,E-mail:
1666503978@qq.com。