2024年6月 第43卷 第6期

DOI:10.19652/j. cnki. femt. 2405955

基于双阶段多尺度生成对抗网络的图像复原方法*

童俊毅¹ 张银胜² 张培琰² 李长帅² 孟祥源² 单慧琳²

(1. 南京信息工程大学电子与信息工程学院 南京 210044;2. 无锡学院电子信息工程学院 无锡 214105)

摘 要:针对人脸图像复原任务中对图像尺度信息利用不足和眼镜结构复原错误的问题,提出一种基于双阶段多尺度生成对 抗网络复原模型。该模型第1阶段引入改进损失的 U-Net 粗重构网络,利用跳连接减少原始图像信息的丢失,融合3种不同 的损失函数提高生成器的重构能力,采用双判别器考虑全局信息和局部信息,并提出一种混合域注意力机制用于关注图像的 空间和通道信息。第2阶段的精修复网络构建了全新的特征增强模块,增强网络对细节信息的提取能力和对结构的表达能 力,引入相对判别器,用于关注生成样本与真实样本之间的相对真实性,提高了生成质量和训练稳定性。实验结果表明,该方 法能够复原各类图像缺失的情况,并能够有效复原佩戴眼镜的人脸图像,与其他方法相比,该方法的峰值信噪比、结构相似性 和感知相似度评估等指标分别提升了 3.81%、2.65%和 0.45%。

关键词:图像复原;生成对抗网络;特征增强;双阶段;U-Net

中图分类号: TP391.4 文献标识码: A 国家标准学科分类代码: 520.6

Image restoration method based on two-stage multi-scale generative adversarial network

Tong Junyi¹ Zhang Yinsheng² Zhang Peiyan² Li Changshuai² Meng Xiangyuan² Shan Huilin² (1. College of Electronic and Information Engineering, Nanjing University of Information Science and Technology, Nanjing 210044, China; 2. School of Electronic Information Engineering, Wuxi University, Wuxi 214105, China)

Abstract: To solve the problem of insufficient use of image scale information and incorrect reconstruction of glasses structure in face image restoration task, a two-stage multi-scale generative adversarial network restoration model is proposed. In the first stage of the model, U-Net coarse reconstruction network with improved loss is introduced, three different loss functions are fused to improve the reconstruction ability of the generator, double discriminator is used to consider the global information and local information, and a mixed domain attention mechanism is proposed to focus on the spatial and channel information of the image. In the second stage, a new feature enhancement module is constructed to enhance the network's ability to extract details and express structures. The experimental results show that this method can recover all kinds of missing images and effectively restore face images wearing glasses. The peak signal-to-noise ratio, structural similarity and perceived similarity evaluation indexes of the method were improved by 3.81%, 2.65% and 0.45%, respectively.

Keywords: image restoration; generate adversarial network; feature enhancement; two-stage; U-Net

0 引 言

图像复原是指利用数字图像未受损区域的已知信息, 进行采集推测,生成语义合理信息去填补缺失的区域。目 前图像复原技术已广泛应用于照片修复、安防刑侦^[1]、图 像去噪去模糊^[2]和文物修复等领域。 传统图像复原技术往往基于插值法和最大似然估计 法等方法进行图像填补,均依赖手工设计特征和算法,难 以适用于不同类型图像。深度学习的发展给图像复原技 术带来了巨大进展。现主流方法基于生成对抗网络^[3-4] (generative adversarial networks,GAN)复原受损图像。 lizuka 等^[5]采用双判别器同时关注局部信息和全局信息,

— 50 — 国外电子测量技术

收稿日期:2024-01-17

^{*} 基金项目:国家自然科学基金(62071240)、2024 年江苏省研究生创新项目(2311082401501)资助

2024年6月 第43卷 第6期

但过度依赖泊松混合技术,复原结果较差。为了充分利 用图像信息,Zeng 等^[6]采用金字塔结构,融合原始图像的 高低级特征,重构出细节纹理更清晰的图像。Quan 等^[7] 和胡雅妮等^[8]引入 U-Net,进一步提取图像的多尺度信 息。马敏等^[9]在多尺度网络中添加自适应机制,增强了 网络的特征表达能力。注意力机制可以有效提高网络对 细节信息的关注能力^[10]。Yu 等^[11]提出上下文注意力来 构建图像语义信息的长期相关性,并采用门控卷积,提出 DFv2 模型。王向军等^[12]针对图像的空间信息和通道信 息,提出一种新的空间通道注意力的连接方式。林晓 等[13]和彭豪等[14]用多尺度特征联合注意力机制,注重不 同尺度特征信息的提取。李海燕等[15]利用空洞卷积和 一致性注意力来关注更大面积的图像信息。Zamir 等^[16]提出监督注意力模块更好地提取了关键信息。随 着 Transformer 的 大 放 异 彩, Dong 等^[17] 结 合 Transformer 的多头注意力机制,更好的提取到了图像的结 构信息,在图像复原任务中取得优异表现,但网络参数 量较大。

随着对修复质量要求的提高,多段式修复网络应运而 生。岳焕景等^[18]用两个浅层 U-Net 网络相连接,采用多 尺度输入输出,全面关注图像原始信息。Nazeri等^[19]采用 两段式边缘生成网络,利用受损图像的边缘信息推断出原 始图像的真实结构。邵新茹等^[20]用三段网络共同关注图 像的色彩、边缘和纹理。Yohan 等^[21]利用三段生成对抗 网络,对图像进行渐进式潜在空间扩展,构造了更加具有 鲁棒性的网络。

赵汉理等[22]在平衡网络参数量和性能的基础上,选

理论与方法

用两段式生成对抗网络,第1阶段采用双判别器,同时对局部信息和全局信息进行关注,第2阶段采用基于去模糊 生成对抗网络(deblurring generative adversarial network, Deblur GAN),对粗重构结果进行去模糊操作,但因其未 能充分利用图像多尺度特征,导致遇到人脸戴眼镜等特殊 场景时会生成错误结构信息,修复效果较差。

针对目前在复原大面积被遮挡的人脸图像时,存在对 原始图像信息利用不足、生成细节信息模糊和对戴眼镜人 脸复原时存在结构错误等问题,本文提出一种基于双阶段 多尺度生成对抗网络(two-stage multi-scale generative adversarial network, TMS-GAN)复原方法,利用 U-Net 结 构保留图像的多尺度信息,提出一种基于自注意力的软硬 混合域注意力机制(soft and hard mixed domain attention, SHMD),关注受损图像的关键信息,构造独特的特征增强 模块(residual based feature enhancement module, FER) 加强网络对细节信息的提取能力。实验结果表明,TMS-GAN 可以复原不同遮挡类型下的人脸图像,并能有效复 原被遮挡的佩戴眼镜的人脸图像。

1 网络模型

1.1 总体结构

本文法提出的 TMS-GAN 网络以人脸图像为研究对象,利用粗重构网络对缺失区域进行大致结构的恢复,再 用精修复网络对细节信息进行细化。TMS-GAN 网络的 总体框架如图 1 所示。

由图 1 可知, TMS-GAN 的网络结构由粗重构网络和 精修复网络组成, 粗重构网络由一个生成器 G1 和两个判



图 1 TMS-GAN 总体框架 Fig. 1 Overall frame diagram of TMS-GAN

别器 D1、D2 组成,生成器 G1 生成结构完整的粗重构图 像,由全局判别器 D1 和局部判别器 D2 判断生成图像真 实性。精修复网络由一个生成器 G2 和一个相对判别器 D3 构成,生成器 G2 对粗重构图像进行特征提取,在注意 力机制和特征增强模块的帮助下,更好地捕捉局部范围内 的语义信息,生成更加合理的局部细节,输出最终的复原 图像。

1.2 粗重构网络

网络感受野的大小和对原始图像多尺度信息的利用 程度,会直接影响复原的结果。生成器 G1 采用基于 U-Net 架构网络,可以充分提取图像的多尺度信息。使用跳 连接将提取到的各级特征信息从编码器传播到解码器,可 以保留图像更多的细节信息,利于提高网络对复杂背景图 像的复原能力。同时利用步长为 2 的卷积层代替原始网 络中的上、下采样层,一定程度上减轻了采样过程中信息 丢失的问题,多次下采样操作,使该网络具有较大的感受 野,能提取更多的全局信息,并在每级采样后利用 SHMD 注意力增强特征的空间和通道域信息,使网络能有效复原 存在大面积遮挡情况的图像。具体生成器 G1 结构如图 2 所示,其中,尺寸大小为每层输出大小。





Fig. 2 Coarse reconstruction network generator structure G1

为避免生成对抗网络训练中梯度消失和模式崩溃等 问题,粗重构网络中引入 Wasserstein 距离作为代价函数, 更准确地衡量真实与生成数据分布之间的差异。定义 如下:

$$W(p_r, p_g) = \frac{1}{K} \sup_{\|f\|_L \leq K} (E_{x \sim p_r(x)} [f(x)] -$$

 $E_{x \sim p_g(x)}[f(x)]$ (1) 式中: $p_r \ \pi p_g \ \beta$ 别代表真实分布和生成分布; sup(•) 代 表函数值的上确界; $K \ \beta$ Lipschitz 常数; $|f|| \leq K$ 表示 函数 f 满足 K-Lipschitz 连续; $E \ \beta$ 不同分布下的数学 期望。

为提高训练过程中的稳定性,在判别器损失中加入梯度惩罚项来满足 Lipschitz 连续条件,判别器损失如下:

 $L_{1(D)} = \min_{G_1} \max_{D1,D2} \{ E_{x \sim p_r(x)} (D_{1,2}(x)) - E_{z \sim p_g(z)} (D_{1,2}(G(z))) + \lambda E_{\hat{x} \sim p_x(\hat{x})} (\| \nabla \hat{x} D_{1,2}(\hat{x}) \|_{p} - 1) \}$ (2) 式中: $\| \cdot \|_{p}$ 表示 p 范数; ∇ 表示梯度算子; λ 表示惩罚项 的系数; $\hat{x} = \vartheta x + (1 - G(z)), \vartheta$ 服从[0,1]范围内的均匀 分布; $p_{\hat{x}}(\cdot)$ 表示从目标数据分布 $p_r(\cdot)$ 和生成数据分布 $p_s(\cdot)$ 中抽样点间的直线均匀抽样; E 为不同分布下的数

2024年6月 第43卷 第6期

学期望; x 和 z 表示真实样本和生成样本。

生成器的损失函数定义为生成样本与真实样本之间 的 Wasserstein 距离的相反数:

 $L_{1(G)} = -E(D_{1,2}(G(z)))$ (3)

为了确保生成图像具有与原图像相似的噪声信号和 相似的颜色差异,粗重构网络中引入了峰值信噪比(peak signal-to-noise ratio, PSNR)损失和结构相似度(structural similarity, SSIM)损失,如式(4)和(5)所示。

$$L_{PSNR} = 10 \lg \left[\frac{\| (x - G_1(x, m) \odot m) \|_2}{(K - 1)^2} \right]$$
(4)

式中:K 表示图像的亮度范围; \lg 函数用于减少欧几里得 波动; $G_1(\cdot)$ 表示生成网络 G_1 的输出;x 表示真实样本; m 表示待修复区域的掩膜。

$$L_{SSIM} = -\frac{2\mu_{a}\mu_{b} + C_{1}}{\mu_{a}^{2} + \mu_{b}^{2} + C_{1}} \times \frac{2\sigma_{ab} + C_{2}}{\sigma_{a}^{2} + \sigma_{b}^{2} + C_{2}}$$
(5)

式中: μ_a 和 μ_b 表示均值; σ_a 和 σ_a 表示标准差; σ_{ab} 表示 a 和 b 的协方差;常数 $C_1 = 0.0001$ 和 $C_2 = 0.0009$ 用于保 证分子分母均不为 0。

此外引入学习感知图像块相似度(learned perceptual image patch similarity, LPIPS)损失,提高生成图像在人类视觉感知上的体验,公式如下:

$$d(x, x_0) = \sum_{l} \frac{1}{H_l W_l} \sum_{h, w} \| w_l \odot (\hat{y}_{hw}^l - \hat{y}_{ohw}^l) \|_2^2$$
(6)

式中: l 为特征通道中第l 层; H 和 W 分别为特征的高和 宽; x 和 x_0 为原图中和生成图像的像素块,其经过预训练 的 VGG 网络进行特征提取,在通道维度上进行单位归一 化,记为 $\hat{y}^l, \hat{y}^l_0 \in \mathbf{R}^{H_l \times W_l \times C_l}$,得到感知特征距离后用权重 W_l 平衡每个层次上的特征距离并加权,得到 LPIPS 损 失值。

生成器 G1 的最终损失函数定义如下:

 $L_{G1} = \lambda L_{1(G)} + \lambda_{psnr} L_{psnr} + \lambda_{ssim} L_{ssim} + \lambda_{lpips} L_{lpips}$ (7) 式中: $\lambda_{\lambda} \lambda_{psnr} \lambda_{ssim} \pi \lambda_{lpips}$ 分别表示用于控制生成器损失、 信噪损失、结构损失和感知损失权重的超参数。

1.3 精修复网络

粗重构阶段重建了结构较为完整的图像,但在局部细节上仍存在纹理模糊等问题。为此,本文基于 Deblur GAN 设计一个新的多尺度精修网络,构造了特征增强模块,在网络层数降低的同时增强了网络的特征提取能力。 生成器 G2 结构如图 3 所示。本文与原 Deblur GAN 的生成器损失一样,定义如下:

$$L_{G2} = \lambda \bar{L}_{2(G)} + \lambda_{lpips} L_{lpips}$$
(8)

为关注图像的局部细节信息,提出一种基于自注意力 的软硬混合域注意力机制 SHMD 同时关注图像的空间与 通道信息。将大小为 $C \times H \times W$ 的特征 A 经过卷积层 后,通过重塑和转置操作得到特征 $B(大小为 N \times C, 其中$ $N = H \times W$),再将特征 A 经过卷积层后通过重塑操作得

2024年6月 第43卷 第6期



Fig. 3 Fine-repair network generator structure G2

到特征 C 和特征 D,将特征 B 与特征 C 进行矩阵相乘和 归一化操作得空间特征图 S。设置一个自学习的权重因 子∂(初始为 0),将提取到的特征 D 与空间特征图 S 进行 相乘,重塑为原始大小尺寸的特征后,在权重因子∂的影响 下与原始输入特征 A 进行相加融合。在通道域上不需要 关注空间信息,因此舍去卷积操作,直接对特征 A 进行与 空间域类似操作,再将通道域和空间域的特征信息分别计 算软注意力和硬注意力,即分别取通道上的加权平均值和 最大值,再对二者进行融合操作,得到最终的特征信息。 SHMD注意力机制增强了网络对空间域特征和通道域特 征的提取能力,提高了网络对细节信息的敏感性,从而达 到更加精细的修复结果。SHMD注意力机制结构如图 4 所示。





1.4 特征增强模块

为了加强生成器 G2 对图像特征的提取能力,本文提 出一种基于残差的特征增强模块 FER,对提取到的特征 进行增强。为了增强网络对特征的表达能力,先将提取到 的特征经过最大池化操作和平均池化操作后,相加得到新 的特征信息 f_1 ;再将原始特征经过由卷积层构成的残差 组,进行重提取,得到更深层的特征信息 f_2 ;最终为了进 一步锐化特征,再次将原始特征信息进行最大池化操作, 取特征图通道域上像素最大值,得到更锐化的特征信息 f_3 ,将 f_3 与之前两种特征信息 f_1 、 f_2 进行融合拼接,利用 卷积层将其转变为大小为 1*xHxW* 的权重特征图,再将原始的特征按最终权重进行重新学习,提高网络对关键特征的自主关注能力。FER 结构如图 5 所示。



Fig. 5 Residual based feature enhancement module FER

1.5 相对判别器

为使生成器逐渐逼近真实数据分布,减少生成样本与 真实数据之间的偏差,在精修复网络中采用相对判别器, 同时平衡生成器和判别器之间的竞争关系,减轻训练不稳 定的问题。

$$\bar{L}_{2(D_3)} = \begin{cases} sigmoid(C(x) - E_{x_f \sim \varrho}C(x_f)) \\ sigmoid(C(x) - E_{x_r \sim \varrho}C(x_r)) \end{cases}$$
(9)

式中: C(•)为鉴别器在激活函数之前的输出; E 为数学 期望值; x_f和x_f分别为生成样本和真实样本。相对判别 器结构如图 6 所示。



Fig. 6 Relative discriminator structure diagram

2 实验结果与分析

2.1 实验数据及参数

实验采用公开人脸图像数据集 CelebA HQ 和 seeprettyface 戴眼镜人脸数据集上进行实验。随机共选取 30 000 张作为训练集,4 000 张为测试集,2 000 张为验证 集。训练过程中使用大小为 128×128 的正方形掩码矩 阵,模拟图像被大面积遮挡的情况。具体实验配置信息如 表1所示。

2.2 评估标准

为了更加科学直观地对比,本文除主流图像复原评估 指标 PSNR 和 SSIM 外,还引入了 LPIPS 指标来评估生成 图像是否更符合人类高级视觉标准。数值上,PSNR 值越 大,SSIM 值大,LPIPS 值越小,代表生成的图像更加合理 真实。



Table 1	Configuration information
配置名称	参数
CPU	Intel(R) Core(TM) i9-12900KF
GPU	NVIDIA GeForce RTX3090
操作系统	Windows10
框架	Pytorch1. 12. 0
Batch Size	16

主 1

2.3 对比实验

为了比较本文方法的优越性,本文使用其他4种主流 图像复原算法进行对比实验,包括 GLCIC 算法^[5]、LGNet 算法^[7]、DFv2 算法^[11]和 EdgeConnect 算法^[19]。本文方法 与上述4种方法的对比可视化结果如图7所示,为更好地 观察,给出了局部放大图。

由图 7 可知,GLCIC 算法虽能完成粗略的人脸重构 阶段,但生成图像质量较差,结构混乱;EdgeConnect算法



图 7 不同模型复原结果对比实验

Fig. 7 Comparison experiment of restoration results of different models

生成的图像较为模糊;DFv2算法虽然引入了注意力机制, 但是在整体的人脸复原上仍然缺少面部细节信息;LGNet 算法采用三段式网络,注重提取图片的多尺度信息,复原 效果相较于之前较好,但网络参数量较大,且在细节处会 结构混乱。本文方法不仅生成了完整的人脸图像,并且在 细节之处也有很好的表现,定量结果如表2所示。

表 2 不同算法在数据集上的评估指标 Table 2 Evaluation indexes of different algorithms on data sets

对比算法	峰值信噪比/dB	SSIM	LPIPS
EdgeConect	21.32	0.902 7	0.089 2
GLCIC	20.98	0.880 9	0.078 2
DFv2	23.32	0.897 1	0.067 5
LGNet	24.81	0.917 3	0.050 8
本文	28.62	0.943 8	0.046 3

为了证明本文方法的泛化能力,本文选取不同形式的

掩码对原图进行遮盖并复原。引入 PSNR、SSIM 和 LPIPS 评估指标来做定量实验,结果如表 3 所示。

表 3 不同算法在不同掩码下平均修复结果评估指标

Table 3 Average repair result evaluation indexes of different

algorithms	under	different	masks	

峰值信噪比/dB	SSIM	LPIPS
25.32	0.940 7	0.049 6
24.71	0.938 1	0.069 7
23.83	0.925 3	0.078 1
26.70	0.932 9	0.056 3
29.17	0.952 3	0.038 7
	峰值信噪比/dB 25.32 24.71 23.83 26.70 29.17	峰值信噪比/dBSSIM25.320.940724.710.938123.830.925326.700.932929.170.9523

对于修复小面积缺失区域图像时,EdegeConect 表现 良好,其根据图像边缘线来生成图像,对于小面积缺失场 景有很好的应用效果;而对于不规则缺失区域,GLCIC、 DFv2和LGNet算法均有一定程度提升,但远不如本文方

2024年6月 第43卷 第6期

理论与方法

法,本文方法在面对不同图像缺失区域修复的情况时,复原 表现仍然优秀,在 PSNR、SSIM 和 LPIPS 评估指标上分别 达到了 29.17、0.9523和0.0387。修复结果如图 8 所示。



图 8 面对不同缺失区域时的复原情况 Fig. 8 Recovery in the face of different missing areas

针对人脸图像复原中存在戴眼镜难以正确修复的问题,本文利用基于自注意力的混合域注意力 SHMD 与特征增强模块 FER 对提取的特征信息重点关注,更好提取了图像的关键信息,捕捉图像的原始细节结构,针对文献[22]提出的复原戴眼镜人脸图像困难的问题有很好的解决效果。定量结果如表 4 所示。

表 4 不同算法在针对戴眼镜人脸修复评估指标 Table 4 Evaluation indexes of different algorithms for facial repair with glasses

	factal repair with glasses							
对比算法	峰值信噪比/dB	SSIM	LPIPS					
EdgeConect	21.19	0.906 4	0.076 7					
GLCIC	20.26	0.882 6	0.071 5					
DFv2	23.03	0.887 5	0.063 9					
LGNet	23.84	0.911 4	0.057 2					
本文	25.23	0.938 9	0.0511					

本文方法在针对戴眼镜人脸复原时,在 PSNR、SSIM 和 LPIPS 评估指标上分别达到了 25.23、0.938 9 和 0.0511。虽然在峰值信噪比指标上略低,但在结构相似 性和感知视觉指标上仍有良好表现,高于其他算法,产生 了结构完整,信息正确的完整图像,解决了文献[22]网络 的不足之处。因 GLCIC 和 EdgeConect 方法结果较差,仅 展示与 DFv2 和 LGNet 方法对比的结果图,如图 9 所示。



图 9 针对戴眼镜人脸缺失复原对比 Fig. 9 Reconstructed comparison of facial deletion with glasses

由图 9 可知, DFv2 方法在复原镜片颜色时发生错误, 且对于暗部细节复原能力较差, 镜框与眼睛边界较为模 糊; LGNet 方法虽能很好复原结构, 但整体色彩有偏差, 且复原部分噪点较高;本文方法能较好完成戴眼镜人脸图 像的复原工作, 但仍存在一些问题, 如图 9 中的输入图像 1,本文方法虽能很好复原镜框的结构和镜片颜色, 但对于 复原图像中的高级视觉特征较为困难, 例如原图中镜面的 反光现象, 本文方法复原结果仍与原始图像有一定差距。

2.4 消融实验

为了验证本文方法中各模块对整体网络复原性能的 影响,对粗重构网络和精修复网络分别进行消融实验,在 此过程中除消融项外,其他结构和参数保持不变。

1) 粗重构网络消融实验

粗重构网络消融实验结果如表 5 所示,标记、/表示模型在训练过程中包含了对应项。策略 1 表示采用原生成 对抗网络的结构,仅引人 3 种损失函数;策略 2~9 采用基 于 U-Net 的网络结构,不同在于对比了 3 类损失函数和 SHMD 注意力机制对结果的影响。由策略 2 可知,仅使 用原本损失函数生成图像的 SSIM 指标和 LPIPS 指标都 很低,远不能满足现今图像复原的要求;策略 3 在策略 2

Table 5	Results of ablat	tion experi	ments between	modules of th	ne coarse reconstruction	network	
U-No	t SHMD	I	I	I	終店信品 レ/dB	SSIM	

消融策略	U-Net	SHMD	L_{psnr}	L_{ssim}	L_{lipis}	峰值信噪比/dB	SSIM	LPIPS
策略1			\checkmark	\checkmark	\checkmark	19.063 34	0.887 96	0.096 5
策略 2	\checkmark					17.117 03	0.840 75	0.138 4
策略 3	\checkmark		\checkmark			18.728 45	0.832 74	0.109 1
策略 4	\checkmark			\checkmark		19.263 96	0.856 63	0.086 6
策略 5	\checkmark				\checkmark	18.524 17	0.848 91	0.136 3
策略 6	\checkmark		\checkmark	\checkmark	\checkmark	19.920 13	0.852 71	0.115 9
策略7	\checkmark	\checkmark	\checkmark		\checkmark	19.025 81	0.845 03	0.153 7
策略8	\checkmark	\checkmark		\checkmark	\checkmark	20.377 61	0.871 12	0.099 3
策略9	\checkmark	\checkmark	\checkmark	\checkmark		19.629 13	0.869 97	0.342 7
本文	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	21.933 68	0.912 56	0.076 8

的基础上加上 PNSR 损失后,生成图像的 PSNR 上升到 18.728,但 LPIPS 同样上升到 0.1384,表明生成图像峰值 信噪比虽有所提升,但在人类视觉上的感知效果变差;策 略4在策略2基础上加入 SSIM 损失后, PSNR 和 SSIM 指标提升到 19.263 和 0.856,表明 SSIM 损失引入不仅能 提高生成图像与原图的结构相似性,也能有效消除噪点; 策略 5 加入 LPIPS 损失后,整体评价指标反不如策略 4; 策略6证明了3种损失函数组合后效果最好,为体现改进 损失的优越性,给出不同损失函数、相同训练批次大小下 的复原结果,如图 10 所示。



本文指生

图 10 不同损失、相同训练批次模型复原对比 Fig. 10 Comparison of model restoration with different losses and the same training batch

由图 10 可知,使用原损失函数的模型复原结果较差,

2024年6月 第43卷 第6期

存在明显噪点,且头发区域信息混乱,说明其训练时未曾 较好收敛:策略 9 中则是引 PSNR 损失和 SSIM 损失后的 复原结果,可看出头发区域较为清晰,但在眼睑周围仍存 在虚影,与本文采用完整损失训练出的模型存在一定差 距。策略 7~9 共同证明了 SHMD 注意力机制在不同损 失函数组合下的有效性,本文所用策略在 PSNR、SSIM 和 LPIPS 指标上均达到最优,分别为 21.933、0.912 和 0.0768.

2)精修复网络消融实验

精修复网络消融实验结果如表6所示,标记√表示模 型在训练过程中包含了对应项。策略1~4表示采用原网 络的结构;策略1在原网络结构在引入了 SHMD 注意力 和 FER 特征增强模块后,生成图像在 PSNR、SSIM 和 LPIPS 指标上达到 23.733、0.921 和 0.065 1,远高于第 1 阶段的最佳复原结果,与策略 2~4 共同证明了 SHMD 注 意力、FER 特征增强模块和相对判别器共同组合后网络 的复原能力最佳;策略 5~8 均采用了本文提出的 U-Net 结构,不同之处在于 SHMD 注意力机制、特征增强模块和 相对判别器的引入对最终修复结果的影响。其中,策略 5 结果反而比不上策略1,证明了只提高判别器性能是不可 行的;策略 6 则证明了相对判别器的必要性;策略 7~8 共 同证明了 SHMD 注意力机制与特征增强模块 FER 组合 是最优解。本文策略的结果在 PSNR、SSIM 和 LPIPS 指 标上均达到最优,分别为 28.621、0.943 和 0.046 3。

Table 0 Results of ablation experiments between modules of the precision repair network							
消融策略	U-Net	SHMD	FER	相对判别器	峰值信噪比/dB	SSIM	LPIPS
策略1		\checkmark	\checkmark		23.733 51	0.921 05	0.065 1
策略 2		\checkmark		\checkmark	22.810 13	0.928 82	0.072 5
策略 3		\checkmark	\checkmark	\checkmark	24.374 10	0.934 73	0.057 4
策略 4			\checkmark	\checkmark	23.575 27	0.931 84	0.066 3
策略 5	\checkmark			\checkmark	23.123 95	0.926 63	0.068 1
策略 6	\checkmark	\checkmark	\checkmark		26.759 81	0.940 56	0.049 2
策略 7	\checkmark	\checkmark		\checkmark	23.770 32	0.920 81	0.067 9
策略 8	\checkmark		\checkmark	\checkmark	23.309 67	0.923 49	0.056 1
本文	\checkmark	\checkmark	\checkmark	\checkmark	28.62171	0.943 86	0.046 3

	10	伯修女网站侠妖问	· *		
Table 6	Deculta of oblation	avanimente hetween	modulos of the		nonain notward

主 4 转换有网络带机间的谐动实际结果

3 结 论

本文提出一种基于双阶段多尺度生成对抗网络的人脸 复原方法,有效复原了不同缺失情况下的人脸图像,成功解 决了其他方法在复原佩戴眼镜的缺失人脸图像任务中失败 的问题。本文方法在 CelebA HQ 和 seeprettyface 数据集上 测试,最终复原图像在 PSNR、SSIM 和 LPIPS 指标上分别 达到 28.62、0.943 8 和 0.046 3。为了模拟图像不同损坏情 况下的复原情况,采用多种不规则掩码空洞修复测试,最终 平均结果在 PSNR、SSIM 和 LPIPS 评估指标上比其他复原 方法相比分别提升了 3.81%, 2.65%和 0.45%。

参考文献

[1] 刘颖,张艺轩,佘建初,等,人脸去遮挡新技术研究综 述[J]. 计算机科学与探索, 2021, 15(10): 1773-1794.

LIU Y, ZHANG Y X, SHE J CH, et al. Research review on new techniques of face occlusion [J]. Exploration of Computer Science and Technology, 2021,15(10):1773-1794.

[2] ZAMIR S W, ARORA A, KHAN S, et al. Restormer: Efficient transformer for high-resolution

■ 理 论 与 方 法

2024年6月 第43卷 第6期

image restoration[C]. 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022: 5718-5729.

- [3] GOODFELLOW I, POUGET-ABADIE J, MIRZA M, et al. Generative adversarial networks [J]. Communications of the ACM, 2020, 63(11): 139-144.
- [4] 孙晶晶,张艳艳,高超,等. 基于 DeblurGAN 的运动模 糊图像盲复原算法研究[J]. 电子测量技术,2022, 45(22):112-119.
 SUN J J, ZHANG Y Y, GAO CH, et al. Research on blind restoration Algorithm of motion blur image

based on DeblurGAN [J]. Electronic Measurement Technology, 2022, 45(22):112-119.

- [5] IIZUKA S, SIMO-SERRA E, ISHIKAWA H. Globally and locally consistent image completion[J]. ACM Transactions on Graphics (TOG),2017,36(4): 1-14.
- [6] ZENG Y,FU J,CHAO H, et al. Learning pyramidcontext encoder network for high-quality image inpainting [C]. 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2019:1486-149.
- [7] QUAN W Z, ZHANG R S, ZHANG Y, et al. Image inpainting with local and global refinement[J]. IEEE Transactions on Image Processing, 2022, 31: 2405-2420.
- [8] 胡雅妮,李光亚,韩晓东,等.基于双判别生成对抗网络的壁画图像虚拟修复[J].国外电子测量技术, 2022,41(6):14-19.

HU Y, LI G Y, HAN X D, et al. Virtual restoration of mural image based on double discriminant generative adversarial network [J]. Foreign Electronic Measurement Technology, 2022, 41(6):14-19.

[9] 马敏,梁雅蓉.基于多尺度自适应特征聚合网络的 ECT 图像重建[J].仪器仪表学报,2023,44(6):264-272.

MA M, LIANG Y R. ECT image reconstruction based on multi-scale adaptive feature aggregation network [J]. Chinese Journal of Scientific Instrument, 2019,44(6):264-272.

[10] 陈广秋,温奇璋,尹文卿,等.用于红外与可见光图像 融合的注意力残差密集融合网络[J].电子测量与仪器学报,2023,37(8):182-193.
CHEN G Q, WEN Q ZH, YIN W Q, et al. Attention residual dense fusion network for infrared and visible image fusion [J]. Chinese Journal of Electronic Measurement and Instrumentation, 2023,

37(8):182-193.

- [11] YU J H, LIN Z, YANG J M, et al. Free-form image inpainting with gated convolution[C]. Proceedings of the IEEE/CVF International Conference on Computer Vision, 2019: 4471-4480.
- [12] 王向军,欧阳文森.多尺度循环注意力网络运动模糊
 图像复原方法[J].红外与激光工程,2022,51(6):460-468.
 WANG X J, OUYANG W S. Multi-scale cyclic

attention network motion fuzzy image restoration method [J]. Infrared and Laser Engineering, 2022, 51(6):460-468.

- [13] 林晓,周云翔,李大志,等.利用多尺度特征联合注意 力模型的图像修复[J].计算机辅助设计与图形学学 报,2022,34(8):1260-1271.
 LIN X, ZHOU Y X, LI D ZH, et al. Image restoration using multi-scale feature combined attention model [J]. Journal of Computer-Aided Design and Graphics,2022,34(8):1260-1271.
- [14] 彭豪,李晓明.利用金字塔空间注意力与特征推理的 图像修复[J].计算机辅助设计与图形学学报,2023, 35(1):87-98.

PENG H, LI X M. Image restoration using pyramid space attention and feature reasoning [J]. Journal of Computer-Aided Design and Graphics, 2023, 35(1): 87-98.

- [15] 李海燕,尹浩林,钟杏苑,等. 基于改进知识一致性注意力的图像修复算法[J].华中科技大学学报(自然科学版) 2023,51(9):110-117.
 LI H Y, YIN H L, ZHONG X Y, et al. Image restoration algorithm based on improved knowledge consistency attention [J]. Journal of Huazhong University of Science and Technology (Natural Science Edition) 2023, 51(9):110-117.
- [16] ZAMIR S W, ARORA A, KHAN S, et al. Multi-Stage progressive image restoration [C]. 2021 IEEE/ CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2021: 14816-14826.
- [17] DONG Q L,CAO C J, FU Y W, et al. Incremental transformer structure enhanced image inpainting with masking positional encoding [J]. 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022: 11348-11358.
- [18] 岳焕景,廖磊,杨敬钰.基于双重自注意力机制的人脸 图像修复[J].湖南大学学报(自然科学版) 2023, 50(8):32-41.

YUE H J, LIAO L, YANG J Y. Face image restoration based on Double self-attention mechanism [J]. Journal of Hunan University (Natural Science Edition) 2023,

2024年6月 第43卷 第6期

50(8):32-41.

- [19] NAZERI K, NG E, JOSEPH T, et al. EdgeConnect: Structure guided image inpainting using edge prediction [C]. 2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCV), 2019;3265-3274.
- [20] 邵新茹,叶海良,杨冰,等.基于三阶段生成网络的图 像修复[J].模式识别与人工智能,2022,35(12):1047-1063.

SHAO X R, YE H L, YANG B, et al. Image restoration based on three-stage generation network [J]. Pattern Recognition and Artificial Intelligence, 2022, 35(12):1047-1063.

[21] YOHAN P G, LALONDE J F. Robust unsupervised StyleGAN image restoration [C]. 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR),2023; 22292-22301. [22] 赵汉理,刘影,卢望龙,等.基于感知去模糊的高效人 脸图像复原算法[J].计算机辅助设计与图形学学报, 2022,34(9):1420-1431.

> ZHAO H L, LIU Y, LU W L, et al. Efficient Face image restoration algorithm based on perceptual deblurring [J]. Journal of Computer-Aided Design and Graphics, 2022, 34(9):1420-1431.

作者简介

童俊毅,硕士研究生,主要研究方向为图像处理,人工 智能。

E-mail: Ttooong@163. com

张银胜(通信作者),博士,副教授,主要研究方向为图 像处理,人工智能。

E-mail:yorkzhang@nuist.edu.cn