

# 身份保持约束下的面部图像超分辨率重建方法

田 旭<sup>1</sup>, 刁红军<sup>1</sup>, 凌兴宏<sup>1,2,3</sup>

(1. 苏州大学计算机科学与技术学院, 苏州 215006; 2. 苏州城市学院, 苏州 215104; 3. 吉林大学符号计算与知识工程教育部重点实验室, 长春 130012)

**摘要:** 低分辨率是影响人脸识别精度的重要因素。一种有效方法是使用图像超分辨率技术对低分辨率图像重建,生成超分辨率图像后再对其作人脸识别,从而克服低分辨率面部图像对人脸识别的限制。但是,现有超分辨率方法在重建过程中往往忽略了保持其原始身份信息,这直接影响生成图像的人脸识别结果。针对上述问题,提出了一种身份保持约束下的面部超分辨率重建方法 IPNet,在提高低分辨率面部图像质量的同时,能保持重建后的面部图像身份。IPNet 方法将语义分割网络和面部生成器相结合,通过语义分割网络提取低维隐码和多分辨率空间特征,进而指导面部生成器输出接近于原图的真实面部图像。在此基础上引入人脸识别网络,将身份信息整合到超分辨率方法中,从而约束重建前后的面部图像身份保持一致。实验结果表明,IPNet 方法在超分辨率图像质量和身份保持上均优于其他对比方法。

**关键词:** 超分辨率;人脸识别;语义分割;面部生成器;身份保持

**中图分类号:** TP391      **文献标志码:** A

## Facial Image Super-Resolution Reconstruction Method with Identity Preserving

TIAN Xu<sup>1</sup>, DIAO Hongjun<sup>1</sup>, LING Xinghong<sup>1,2,3</sup>

(1. School of Computer Science and Technology, Soochow University, Suzhou 215006, China; 2. Suzhou City University, Suzhou 215104, China; 3. Key Laboratory of Symbol Computation and Knowledge Engineering of Ministry of Education, Jilin University, Changchun 130012, China)

**Abstract:** Low resolution is an important factor that affects the accuracy of face recognition. To overcome the limitation of low-resolution facial images on face recognition, one effective solution is adopting super-resolution methods to reconstruct low-resolution images and then identify the generated facial images. However, existing super-resolution methods typically fail to consider facial identity preservation during reconstruction, which directly results in poor face recognition performance of reconstructed images. To address the issue mentioned above, this paper proposes a face super-resolution reconstruction method with identity preserving, called IPNet. This method can simultaneously improve the quality of low-resolution facial images and preserve the identity of reconstructed images. IPNet consists of a semantic segmentation network and a face generator. The semantic segmentation network is introduced to extract low-dimensional latent code and multi-resolution spatial features. Then, the extracted features guide the face generator to output super-resolution images similar to the authentic images. Furthermore, we introduce the face

recognition network to integrate the face identity information into the super-resolution model, thus maintaining the identity of reconstructed facial images consistent with original images. Experimental results show that IPNet achieves better results than other comparison methods in terms of both super-resolution image quality and identity preservation, demonstrating effectiveness of the proposed method.

**Key words:** super-resolution; face recognition; semantic segmentation; face generator; identity preserving

## 引 言

很多因素影响人脸识别效果,如图像分辨率低、光照、面部表情、遮挡等。其中,低分辨率是影响人脸识别的一个重要因素,这可能会导致人脸识别准确率降低。在现实场景中,由于物理设备、环境、成像等条件限制,获取到的面部图片分辨率往往比较低,在面部相关的任务中难以达到预期效果。因此,将面部超分辨率模型与人脸识别网络相结合,能够有效克服低分辨率面部图像的限制,在智慧安防、监控寻人、公安追踪等领域发挥重要作用<sup>[1]</sup>。

面部超分辨率重建(Face super-resolution, FSR)是指从模糊的低分辨率(Low resolution, LR)面部图像中重建出清晰的超分辨率(Super resolution, SR)面部图像<sup>[2]</sup>,生成的结果被用于后续基于人脸识别<sup>[3-5]</sup>的相关任务。随着计算机设备和深度学习技术的快速发展,研究者们尝试将深度学习相关方法应用到超分领域,并且取得显著的效果。部分研究工作<sup>[6-8]</sup>通过卷积神经网络直接学习低分辨率图像和高分辨图像之间的端到端的映射关系,可以达到良好的生成质量和速度。

但是,在实践中依然存在以下问题。一方面,目前典型的FSR<sup>[8-11]</sup>方法是利用特定的先验信息(如面部关键点、面部解析图等)指导超分辨率面部重建。先验信息包含面部组成和面部结构信息,精确还原面部细节。但是,先验信息是从输入图像中估计而来,先验信息的效果取决于图像质量。而现实生活中的低分辨率面部图像,往往难以提取有效的面部先验,用于指导面部重建的纹理特征十分有限,限制生成图像的真实性。另一方面,已有的FSR研究工作在生成超分辨率图像时往往不考虑保持其身份信息,没有身份约束的超分辨率方法会导致重建出的面部图像身份不能与原始图像身份保持一致,影响后续人脸识别任务的精度<sup>[12]</sup>。

鉴于上述问题,本文提出了一个身份保持约束<sup>[13]</sup>下的面部图像超分辨率重建方法IPNet。本方法以面部生成器为基础,同时引入语义分割网络,通过多层感知器层(Multilayer perceptron, MLP)<sup>[14]</sup>和空间特征变换层(Spatial feature transform, SFT)<sup>[15]</sup>作特征映射与特征调制,实现面部生成器和语义分割网络的关联。面部生成器使用预训练的面部生成模型,将大量面部信息隐式封装到生成器中作为生成先验。语义分割网络的加入可以有效提取人脸图像的面部信息,其中,编码器能提取面部细节特征,而解码器对生成的多分辨率图像特征分开处理,调制面部生成器后再融合汇总,指导生成器逐步生成接近于原图的真实面部图像。IPNet还引入人脸识别网络,通过比较SR图像以及高分辨率原图(High resolution, HR)的面部特征向量,定义面部身份损失函数实现无标签身份分类。具体而言,通过人脸识别模型提取SR图像和HR图像的面部特征向量,计算这两个向量的相似度并定义身份损失,优化身份损失使两个特征间的距离最小化,相似度最大化,从而保持重建前后面部图像身份的一致性。

本文贡献如下:(1)将面部生成先验信息引入超分辨率任务中,这些生成先验包括面部纹理和颜色信息等,精确指导解码器生成超分辨结果。(2)提出带有残差连接的Res-UNet网络与StyleGAN V2相结合的架构,充分利用生成先验,实现保真度和真实纹理间的平衡。(3)从LR图像中分别提取低维特征和多尺度空间特征,将提取到的两类特征分别通过多层感知器以及特征转换层融入生成对抗网中。

## 1 相关工作

图像超分辨率是指从低分辨率图像恢复出高分辨率图像的过程,是计算机视觉和图像处理中一类重要的图像处理技术<sup>[12]</sup>。面部超分辨率是图像超分辨率中的一个分支。相比而言,面部图像复杂且敏感,但是面部结构相对固定,在重建过程中可以引入面部先验(如面部关键点、面部解析图等)作为参考,辅助超分辨率模型生成更加真实的图像。还有一些工作通过训练面部生成模型,将生成先验封装到面部生成器中,并将训练好的生成器嵌入到自己的网络结构,充分利用预先训练好的生成器所提供的生成先验。

本节分别从基于面部先验信息、面部生成模型和面部身份保持这3个不同类型的面部超分辨率方法回顾相关的工作。

### 1.1 面部先验信息

与自然图像相比,面部图像包含更多结构化信息,因此将面部先验信息整合到FSR中是一种非常流行且颇有前景的方法<sup>[12]</sup>。

Super-FAN<sup>[16]</sup>通过面部热力图回归整合子网络进行面部关键点定位,然后使用基于GAN的超分辨率网络和一个面部比对网络FAN来约束生成的图像中检测到的面部关键点,保持生成图与原图的一致性。FSRNet<sup>[10]</sup>设计了编码-解码的结构,在编码器中提取图像特征,通过先验信息预测网络估计出关键点热力图和面部解析图,提取的特征和先验信息经过解码器恢复出目标分辨率图像。同时,FSRNet的作者还提出了FSRGAN<sup>[10]</sup>结构,FSRGAN在FSRNet的基础上增加了一个鉴别器网络,优化对抗损失和感知损失生成更加真实的面部图像。DICGAN<sup>[17]</sup>设计了一种深度相互迭代的网络结构,迭代渐进地计算SR图像和面部的特征点,并设计了一个注意力协同模块,通过优化对抗损失来增加图像的细节信息。HiFaceGAN<sup>[18]</sup>是一个具有多个协同抑制和补给(CSR)单元的生成框架,可以在具有分层语义指导的多阶段方式中实现超分辨率重建,每个CSR单元都包含一个抑制模块,通过内容自适应卷积提取分层的语义特征,以指导相应语义内容的补充。

面部先验信息可以精确指导超分辨率方法还原面部细节。但是,由于现实场景中的低分辨率图像退化严重,难以提取有效的面部先验信息,这就限制了这类方法的实用性。

### 1.2 面部生成模型

近些年,许多面部生成模型被提出,如主流的StyleGAN<sup>[19]</sup>、StyleGAN2<sup>[20]</sup>、StarGAN<sup>[21]</sup>等,这些模型能够生成具有高度可变的真实面部图像。因此,越来越多的研究者尝试预训练面部生成模型,学习大量的面部先验特征(如五官特征,面部纹理等),并将训练好的面部生成模型嵌入到自己的网络结构中,通过先验信息指导超分辨率重建任务。第一个提出基于面部生成模型的FSR方法是PULSE<sup>[22]</sup>,该模型将FSR任务作为生成问题,生成高质量的面部图像,约束生成图像在降采样后与原始输入的低分辨率图像保持一致。但是PULSE使用随机采样的低维隐码,使得生成图像失去了重要的面部空间信息。

为了克服这个问题,GLEAN<sup>[23]</sup>、GFP-GAN<sup>[24]</sup>以及GPEN<sup>[25]</sup>等方法相继被提出。这些方法将预训练的面部生成器嵌入到自己的超分辨率模型中,以合并生成先验。为了保持面部真实性,这些模型对LR图像编码来获取低维隐码,而不是PULSE那样使用随机采样的方式。同时,本文方法还从LR面部图像中提取多尺度特征并融合到生成网络中。这样既可以充分利用预先训练的生成器提供的生成先验,又可以维持重要的面部空间信息。

### 1.3 面部身份保持

与人脸结构先验和属性信息相比,包含身份感知细节的面部身份信息更加重要。目前,在超分辨率任务中,常用的一种保持面部身份的设计是使用人脸识别网络定义身份损失,如EIPNet<sup>[1]</sup>、WaSRGAN<sup>[26]</sup>和IPFH<sup>[27]</sup>等。这些方法尝试使用预训练的人脸识别网络并加入自己的超分辨率方法中,将面部身份信息整合到超分辨率模型中,约束重建后的SR图像与真实HR图像的身份保持一致,从而提高后续人脸识别任务的性能。

## 2 模型设计

### 2.1 整体框架

受1.2、1.3节相关工作启发,本文提出了一个身份保持约束下的面部超分辨率方法IPNet。IPNet使用语义分割网络和面部生成器相结合的方式,充分利用生成先验和面部空间信息,确保重建图像与原图面部结构尽可能相似。本文通过引入人脸识别网络,将身份信息整合到面部超分辨率重建任务中,确保重建后的图像身份与原图保持一致。IPNet可以同时改善面部超分辨率模型生成的图像质量和人脸识别效果,从而克服输入图像分辨率对人脸识别的限制。

IPNet包括3部分网络:(1)生成网络,面部超分辨率生成器被用作生成网络,将低分辨率的输入处理成高分辨率的面部图像;(2)对抗网络,生成网络的输出结果作为对抗网络的输入,对抗网络判断图片是真实样本的概率,指导生成器的输出尽可能拟合真实样本;(3)身份保持网络,身份保持网络将生成图片与真实样本作为输入,分别提取面部特征向量并计算其相似度,约束生成图像面部特征与原始身份保持一致。其中,对抗网络和身份保持网络仅在训练阶段优化生成器模型,在推理阶段不再使用。

IPNet对低分辨率图像的处理流程为:首先,低分辨率图像经过生成网络重建出超分辨率图像;然后,对抗网络判断生成图像真实且自然的概率,身份保持网络判断生成的面部图像身份是否与原始图像身份一致。IPNet方法的整体架构如图1所示。

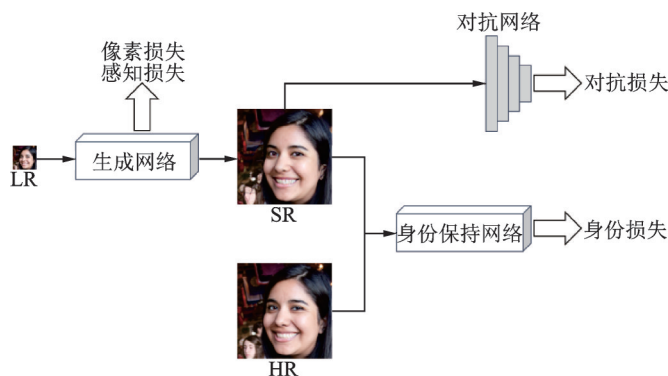


图1 IPNet架构图

Fig.1 Architecture of IPNet

IPNet方法包括:(1)度量像素差异的像素损失;(2)描述风格差异的感知损失;(3)恢复真实纹理的对抗损失;(4)面部身份损失共4个损失函数。

### 2.2 生成网络

生成网络由模糊上采样模块(Coarse upsample module, CUM)、语义分割网络和面部生成器共3个



部分构成。通过MLP<sup>[14]</sup>作特征映射以及SFT<sup>[15]</sup>作特征调制,实现语义分割和面部生成器两个模块关联。生成网络的总体架构如图2所示,其中,低分辨率图像、高分辨率原图和超分辨率生成图像分别表示成 $I_{LR}$ 、 $I_{HR}$ 和 $I_{SR}$ 。

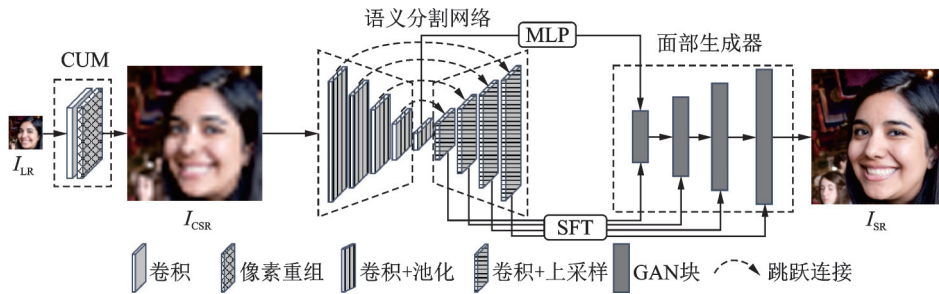


图2 面部超分辨率生成器结构

Fig.2 Model structure of face super-resolution

具体而言,通过CUM模块将 $I_{LR}$ 预处理成目标分辨率尺寸,语义分割模块提取两类特征:低维隐码和多分辨率空间特征。低维隐码通过MLP层映射到面部生成器空间中,得到中间卷积特征。多分辨率空间特征通过SFT层调制相应分辨率的中间特征,多轮迭代后生成最终结果 $I_{SR}$ 。

(1) 模糊上采样模块

低分辨率面部图像难以直接提取有效的面部特征,因此在将特征图输入到语义分割网络前,需要将输入的低分辨率图像预处理成目标分辨率大小。本文通过模糊上采样模块生成模糊超分辨率面部图像 $I_{CSR}$ 。模糊上采样模块由一个卷积层和一个像素重组层(Pixel shuffle)构成,将低分辨率的特征图通过卷积和多通道间的重组得到高分辨率的特征图<sup>[28]</sup>。对于输入的低分辨率图像 $I_{LR}$ ,输出的模糊高分辨率图像 $I_{CSR}$ 公式表示为

$$I_{CSR} = f_{CUM}(I_{LR}) \tag{1}$$

式中: $f_{CUM}$ 表示模糊上采样操作,得到的模糊特征图 $I_{CSR}$ 作为后续语义分割网络的输入。

(2) 语义分割网络

语义分割是像素级别的图像分类任务,能够从图像中分割出感兴趣的区域。就面部图片而言,主要的面部组件有眼睛、鼻子、嘴巴等器官,加入语义分割模块能够有效提取这些面部信息,从而在图像重建过程中能更深层次地推断出面部高级特征。

本文使用Res-Unet<sup>[29]</sup>作为语义分割网络,Res-Unet是Unet<sup>[30]</sup>的一种变体,在Unet的基础上引入了残差连接的思想。Res-Unet结构图如图3所示。

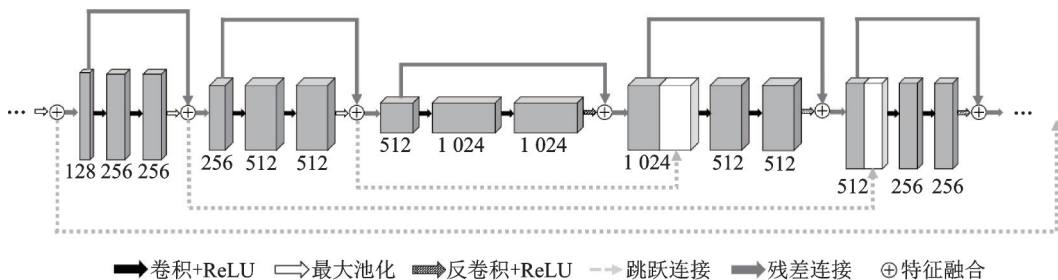


图3 Res-Unet结构图

Fig.3 Architecture of Res-Unet

Res-Unet结构延续了Unet的U形对称结构,在上采样和下采样的每一层块中引入残差连接。具体做法是添加一条从两次卷积的输入到输出的连接,并做一次卷积操作。相比单纯使用Unet结构,这种残差结构思想改善了网络的梯度流通,避免网络退化,加速网络收敛,并且能避免因感受野扩大导致的面部细节丢失问题。Res-Unet共经过4次下采样获取语义信息,4次上采样恢复原始分辨率。通过跳跃连接的方式,融合不同分辨率特征使上采样的特征图包含更多的语义信息,确保上采样准确恢复出原始分辨率的特征图。下采样的瓶颈层提取 $F_{\text{latent}}$ 特征,上采样的各层提取多分辨率空间特征序列 $F_{\text{spatial}}$ ,公式表示为

$$F_{\text{latent}}, F_{\text{spatial}} = f_{\text{Res-Unet}}(I_{\text{CSR}}) \quad (2)$$

式中 $f_{\text{Res-Unet}}$ 表示语义分割网络提取特征操作。这两类特征被用作后续的生成器任务。 $F_{\text{latent}}$ 通过MLP映射到生成器潜在空间, $F_{\text{spatial}}$ 用于调制生成器中间卷积特征。

(3) 特征映射网络

给定输入图像的编码特征向量 $F_{\text{latent}}$ ,通过非线性映射网络将输入向量映射到中间隐码 $W$ ,以更好地保持面部信息。特征映射网络结构如图4所示。

映射网络由多层感知器层MLP<sup>[14]</sup>,即8个全连接层构成。公式表示为

$$W = f_{\text{MLP}}(F_{\text{latent}}) \quad (3)$$

式中 $f_{\text{MLP}}$ 表示MLP映射网络。映射得到的中间隐码 $W$ 通过后续生成器中每个卷积层,卷积计算对应分辨率下的中间卷积特征。

(4) 面部生成器

通过训练面部生成模型,将大量面部信息隐式封装到生成模型中,本文采用StyleGAN2<sup>[20]</sup>模型作为面部生成器。StyleGAN2具备生成多变的真实面部图像的能力,可以精确还原五官、面部纹理和颜色等面部细节。生成面部图像时,从预训练面部生成模型的卷积权重中获取面部分布,即获取生成先验。部署生成先验的典型方法是将输入图像映射到中间隐码,然后通过预训练的生成器得到相应的面部输出。这种方法在视觉上逼真,但是低维度的中间隐码不足以精确指导图像生成,面部细节保真度较低。因此,本文模型没有直接生成最终图像,而是通过中间隐码 $W$ 生成中间卷积特征 $F_{\text{GAN}}$ ,多分辨率空间特征 $F_{\text{spatial}}$ 进一步对其调制以获得更准确的面部细节。生成中间卷积特征公式为

$$F_{\text{GAN}} = f_{\text{StyleGAN2}}(W) \quad (4)$$

式中 $f_{\text{StyleGAN2}}$ 指隐码 $W$ 通过预训练StyleGAN2的每个卷积层,为每个分辨率尺寸生成中间卷积特征 $F_{\text{GAN}}$ 。

(5) 空间特征调制

本文使用多分辨率空间特征序列 $F_{\text{spatial}}$ 调制中间卷积特征 $F_{\text{GAN}}$ 。面部超分辨率重建需要局部特征来保持保真度,而 $F_{\text{spatial}}$ 中保留重要的面部空间信息,能确保生成图片面部细节的真实性,并在面部的不同空间位置进行自适应恢复。SFT<sup>[15]</sup>结构如图5所示。

对于每个分辨率下的面部空间特征 $F_{\text{spatial}}$ ,通过两次卷积计算得到仿射变换参数对 $(\gamma, \beta)$ , $\gamma$ 和 $\beta$ 分别表示仿射变换的尺度和平移参数。仿射变换参数通过缩放和移动中间卷积特征 $F_{\text{GAN}}$ 来维持面部空间信息,并作为下一轮输入。多轮迭代后得到的最终生成图更接近原图。公式为

$$\gamma, \beta = f_{\text{Conv}}(F_{\text{spatial}}) \quad (5)$$

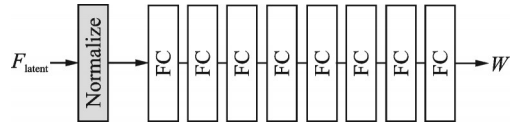


图4 MLP结构图

Fig.4 Architecture of MLP

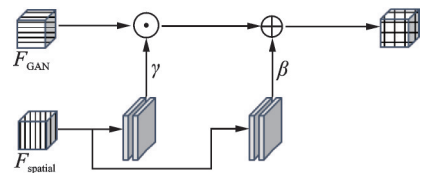


图5 SFT结构图

Fig.5 Architecture of SFT

$$F_{\text{output}} = f_{\text{SFT}}(F_{\text{GAN}}, \gamma, \beta) = \gamma \odot F_{\text{GAN}} + \beta \quad (6)$$

式中:  $f_{\text{Conv}}$  表示通过两次卷积分别得到仿射变换参数对  $\gamma$  和  $\beta$  的过程;  $f_{\text{SFT}}$  表示中间卷积特征  $F_{\text{GAN}}$  的调制函数;  $\odot$  和  $+$  分别表示对  $F_{\text{GAN}}$  特征的缩放和移动过程。

### (6) 像素损失和感知损失

给定一组低分辨率面部图像  $I_{\text{LR}}$  和对应的高分辨率原图  $I_{\text{HR}}$ , 经过 FSR 模型重建得到  $I_{\text{SR}}$ 。使用最常见的 L1 损失度量生成图片  $I_{\text{SR}}$  和目标图片  $I_{\text{HR}}$  的像素级差异, 定义像素损失, 公式如下

$$L_{\text{pixel}} = \lambda_{\text{pixel}} \sum_{x=1}^w \sum_{y=1}^h \|I_{\text{SR}}(x, y) - I_{\text{HR}}(x, y)\|_1 \quad (7)$$

式中:  $w$ 、 $h$  分别表示目标面部图像的宽和高,  $x$  表示图像上对应像素位置的横坐标,  $y$  表示对应像素位置的纵坐标,  $\lambda_{\text{pixel}}$  表示像素损失的权重。虽然优化 L1 损失函数可以获得较高的 PSNR 值, 但是也会难以避免地产生细节上的模糊、伪影。为了解决这个问题, 本文使用感知损失以获得视觉上更清晰的图像, 将预训练的 VGG-19<sup>[31]</sup> 网络作为特征提取器, 卷积提取生成图  $I_{\text{SR}}$  和原图  $I_{\text{HR}}$  的风格特征向量, 计算两组特征的相似度, 用于定义感知损失, 公式为

$$L_{\text{per}} = \lambda_{\text{per}} \|\varphi(I_{\text{SR}}) - \varphi(I_{\text{HR}})\|_1 \quad (8)$$

式中  $\varphi$  表示预训练的 VGG-19 网络, 用作图像的特征提取器,  $\lambda_{\text{per}}$  表示感知损失的权重。

## 2.3 对抗网络

对抗网络将图像作为输入, 判断该输入是真实图像还是生成的虚假图像。使用对抗网络是为了将生成器输出的超分辨率图像  $I_{\text{SR}}$  与真实图像  $I_{\text{HR}}$  区分开来, 定义对抗损失  $L_{\text{adv}}$ 。与 StyleGAN2<sup>[20]</sup> 类似, 本文采用 logistic 损失计算对抗损失  $L_{\text{adv}}$ 。公式表示为

$$L_{\text{adv}} = -\lambda_{\text{adv}} E \left[ \lg(D(I_{\text{HR}})) + \lg(1 - D(I_{\text{SR}})) \right] \quad (9)$$

式中:  $D$  表示 StyleGAN2 中的鉴别器,  $\lambda_{\text{adv}}$  表示对抗损失的权重。通过优化目标函数  $L_{\text{adv}}$ , 生成让鉴别器网络难以辨别的超分辨率样本。

## 2.4 身份保持网络

超分辨率模型在重建图像质量方面已经获得了很逼真的效果, 但是这些模型在重建过程中往往忽略了重要的面部身份信息, 重建后的图像容易丢失其身份, 导致后续人脸识别效果下降。于是, 本文通过整合人脸识别网络来增强低分辨率和高分辨率面部图像的身份一致性。

本文选用 ArcFace<sup>[5]</sup> 作为身份保持网络, 该模型在开源人脸识别算法中性能更加突出。ArcFace 是一种类似 resnet<sup>[32]</sup> 的卷积神经网络模型, 通过加性角度间隔损失函数 (Additive angular margin loss, ArcFace Loss) 进行训练, 可以有效增强特征向量的分辨能力。ArcFace Loss 是对传统 softmax 损失函数的改进, 将类间的距离映射到超球面的间距, 直接在角度空间中最大化分类界限。ArcFace 中大量实验证实其效果始终优于其他人脸识别方法<sup>[5]</sup>。

与定义感知损失的方法类似, 本文使用预训练的 ArcFace 作为特征提取器提取面部特征, 计算两组特征的余弦距离衡量其相似度, 定义身份损失。通过优化目标函数  $L_{\text{id}}$ , 使生成图和原图在面部特征空间上尽可能相近。身份损失公式表示为

$$L_{\text{id}} = \lambda_{\text{id}} \|\psi(I_{\text{SR}}) - \psi(I_{\text{HR}})\|_1 \quad (10)$$

式中:  $\psi$  表示面部特征提取器, 即预训练的 ArcFace 模型;  $\lambda_{\text{id}}$  代表身份损失的权重。

## 2.5 总损失函数

用于训练 IPNet 的总损失为上述几个损失函数的总和, 即

$$L_{\text{total}} = L_{\text{pixel}} + L_{\text{per}} + L_{\text{adv}} + L_{\text{id}} \quad (11)$$

实验中,各个损失函数权重分别设置为 $\lambda_{\text{pixel}} = 0.1, \lambda_{\text{per}} = 1, \lambda_{\text{adv}} = 0.1, \lambda_{\text{id}} = 20$ 。

## 2.6 算法描述

具体的算法流程如算法1所示。

**算法1** 身份保持约束下的超分辨率重建方法 IPNet

输入:低分辨率图像 $I_{\text{LR}}$ ,高分辨率原图 $I_{\text{HR}}$ 。

输出:超分辨率重建图像 $I_{\text{SR}}$ 。

① 对低分辨率图像模糊上采样到指定分辨率图像 $I_{\text{CSR}}: I_{\text{CSR}} \leftarrow \text{CUM}(I_{\text{LR}})$

② Res-UNet 网络从 $I_{\text{CSR}}$ 中提取低维隐码 $F_{\text{latent}}$ 和多分辨率空间特征 $F_{\text{spatial}}$ :  
 $F_{\text{latent}}, F_{\text{spatial}} \leftarrow \text{Res-UNet}(I_{\text{CSR}})$

③ 通过MLP层将 $F_{\text{latent}}$ 映射到中间隐码 $W$ 上: $W \leftarrow \text{MLP}(F_{\text{latent}})$

④ StyleGAN2对 $W$ 卷积生成最近的中间卷积特征 $F_{\text{GAN}}: F_{\text{GAN}} \leftarrow \text{StyleGAN2}(W)$

⑤ 使用多分辨率空间特征 $F_{\text{spatial}}$ 调制对应分辨率的中间卷积特征 $F_{\text{GAN}}$ 。对于每个分辨率空间特征 $f$ ,首先对 $f$ 卷积提取仿射变换参数 $\gamma$ 和 $\beta$ ,之后使用 $\gamma$ 和 $\beta$ 对 $F_{\text{GAN}}$ 缩放和平移:

For each  $f$  in  $F_{\text{spatial}}$ :

$\gamma, \beta \leftarrow \text{Conv}(f)$

$F_{\text{GAN}} \leftarrow F_{\text{GAN}} \odot \gamma + \beta$

$F_{\text{GAN}} \leftarrow \text{StyleGAN2}(F_{\text{GAN}})$

end For

⑥ 将最终调制结果 $F_{\text{GAN}}$ 的特征图转换为RGB图像,输出超分辨率图像 $I_{\text{SR}}$ 。

## 3 实验与分析

### 3.1 环境与设置

本文方法在开源高清面部数据集FFHQ<sup>[24]</sup>上进行充分的实验。FFHQ数据集有70 000张分辨率为 $1024 \times 1024$ 高清面部图片,将所有图片通过双三次插值下采样至 $128 \times 128$ 的分辨率作为高分辨率原图 $I_{\text{HR}}$ 。将处理后的70 000张图片随机分割成两部分,63 000张图片用作训练集,7 000张图片用作测试集。低分辨率图像通过高斯模糊下采样(Blur-down, BD)降质方法获得,采用以下降质模型模拟现实场景中的低分辨率面部图像

$$I_{\text{LR}} = (I_{\text{HR}} \otimes k) \downarrow_r + n \quad (12)$$

首先将高质量图像 $I_{\text{HR}}$ 通过高斯模糊核 $k$ 进行卷积,然后进行比例因子为 $r$ 的降采样,再对图像中加入高斯白噪声 $n$ ,模拟现实场景中低质量、低分辨率图像。对于每张训练图片,分别从 $[0.2:10], [0:15]$ 中随机抽样 $k$ 和 $n$ 的值, $r$ 设置为4表示当前 $128 \times 128$ 的 $I_{\text{HR}}$ 图像降采样成 $32 \times 32$ 的低分辨率图 $I_{\text{LR}}$ 。

同时,在LFW<sup>[33]</sup>数据集上验证本文模型对身份保持的有效性。LFW数据集来源于生活中的自然场景,由于光照、遮挡等因素影响,图像上模糊干扰严重,在生成低分辨率图时不需要再对它降质。LFW数据集以人物姓名命名图片,其中有423个人对应的面部图像不低于5张,用这些人的所有图片作为评估人脸识别性能的测试集,共提取出5 985张图片。从每个人图片中随机采样一张,生成包含423张图片的底库并将其降采样成 $32 \times 32$ 低分辨率图像。低分辨率图像作为超分辨率模型的输入,重建后的超分辨率图像与测试集中剩余的5 562张图片作面部匹配,验证超分辨率模型对身份保持的有效性。

实验设置和训练设置:本次实验采用Adam优化器<sup>[34]</sup>训练模型,使用颜色抖动、水平翻转方式来增



强数据,对上面提到的训练集总计进行了80万次迭代。初始学习率设置成 $2 \times 10^{-3}$ ,在第60万次、70万次迭代时分别更新成 $1 \times 10^{-3}$ 、 $5 \times 10^{-4}$ 。本文实验在Ubuntu16.04 64位操作系统上完成,使用PyTorch开源框架,在显存为24 GB的3 090上训练并测试模型。

### 3.2 对比实验

在FFHQ测试集上,将本文提出的方法与其他主流超分辨率方法进行对比,包括:ESRGAN<sup>[35]</sup>、HiFaceGAN<sup>[18]</sup>、PULSE<sup>[22]</sup>、Super-FAN<sup>[16]</sup>、VDSR<sup>[36]</sup>、SRCNN<sup>[6]</sup>、SRGAN<sup>[37]</sup>、GPEN<sup>[25]</sup>,并加入了双三次插值(Bicubic)的比较。所有对比方法都使用官方提供的代码,并且在本文使用的训练集上重新训练,以便进行公平的比较。

评估模型时,使用广泛使用的图像生成质量指标FID、像素级度量PSNR和SSIM以及感知度量LPIPS。使用预训练的FaceNet<sup>[4]</sup>模型计算生成图与原图的面部特征距离Dis,值越小表示生成图在面部特征空间上越接近于原图。使用人脸识别精确率Precision、召回率Recall和F1值度量超分辨率模型身份保持性能。

#### 3.2.1 与现有方法对比

为了验证IPNet方法在面部超分辨率重建以及重建前后身份保持任务上的有效性,分别在两个常用面部数据集FFHQ和LFW上与其他方法进行对比实验。

**FFHQ-TEST 测试集的评估** 将本文模型与其他对比模型进行定量的比较,对比结果如表1所示。其中标粗的结果表示几组实验中的最优值,PSNR、SSIM越高越好,LPIPS、FID、Dis越低越好。从表1可以看出,当放大倍数为4时,本文模型在LPIPS、FID、Dis三项指标上远远优于其他模型。最低的LPIPS值表明本文重建的面部图像从感知上更接近于原图像,最低的FID值意味着重建图像具有最好的图片质量和多样性。同时,得益于身份损失和SFT结构的使用,本文模型还保持了较高的人脸识别性能,最低的Dis值表明生成图与原图在面部特征空间上距离最小。在像素级指标PSNR和SSIM上,本文模型也有较好的表现。

与其他模型可视化结果对比如图6所示,本文方法生成的面部样貌更加真实自然,五官结构完整,相比之下更加清晰。在面部细节方面,本文方法的面部纹理更加丰富(如眼睛、牙齿、头发等细节)并且

表1 IPNet与其他对比方法在FFHQ-TEST和LFW-TEST测试集上的结果比较

Table 1 Comparison of IPNet with other methods on FFHQ-TEST and LFW-TEST test datasets

方法	FFHQ-TEST					LFW-TEST		
	PSNR	SSIM	LPIPS	FID	Dis	精确率	召回率	F <sub>1</sub> 值
Bicubic	25.117 1	0.768 4	0.286 4	102.872 4	0.201 60	0.909 6	0.679 5	0.777 9
ESRGAN <sup>[35]</sup>	24.331 5	0.748 5	0.158 7	17.209 5	0.140 80	0.909 5	0.557 7	0.691 5
HiFaceGAN <sup>[18]</sup>	25.860 8	0.800 9	0.154 3	24.213 0	0.076 60	0.880 7	0.615 4	0.724 5
PULSE <sup>[22]</sup>	19.663 3	0.526 9	0.385 3	63.949 9	0.612 60	0.000 0	0.000 0	0.000 0
Super-FAN <sup>[16]</sup>	22.542 5	0.667 9	0.222 3	39.789 4	0.167 70	0.895 2	0.569 4	0.696 1
VDSR <sup>[36]</sup>	<b>27.025 4</b>	<b>0.828 8</b>	0.195 5	55.648 1	0.104 50	0.904 5	0.696 6	0.787 0
SRCNN <sup>[6]</sup>	25.877 2	0.792 2	0.247 5	71.316 0	0.176 90	0.913 7	0.679 5	0.779 4
SRGAN <sup>[37]</sup>	22.496 1	0.660 4	0.209 1	33.817 5	0.149 90	0.845 6	0.568 3	0.679 7
GPEN <sup>[25]</sup>	24.607 1	0.756 0	0.170 8	18.777 8	0.070 58	0.893 8	0.639 9	0.745 9
IPNet	26.119 2	0.809 3	<b>0.122 5</b>	<b>14.585 3</b>	<b>0.064 60</b>	<b>0.914 0</b>	<b>0.699 7</b>	<b>0.792 7</b>
GT	∞	1.000 0	0.000 0	0.000 0	0.000 00	0.909 8	0.772 3	0.835 4



图6 不同模型在 FFHQ-TEST 测试集上视觉效果对比

Fig.6 Comparison of visual effects of different models on FFHQ-TEST dataset

失真度更小。

尽管在像素级指标 PSNR 和 SSIM 上,本文模型未能超过 VDSR 模型,但就生成图像视觉效果而言,本文模型生成图片质量要明显更优。这一对比结果也说明了感知类指标在粗尺度上与人的感知相关,细尺度上相关性不强,而像素级指标则与它相反。目前,很多超分辨率模型在定量评估时仅仅使用像素级指标,这种评估方法不能很好地反映人眼主观感受,需要配合感知类指标综合评估。

**LFW-TEST 测试集的评估** 在 LFW-TEST 测试集上验证本文模型对面部身份保持的有效性,将重建后的超分辨率图像与 LFW 中的测试集做面部匹配。人脸识别模型选用预训练的 FaceNet<sup>[4]</sup>模型,特征提取器分别对生成图与对比图提取面部特征,使用余弦相似度描述两幅面部图像的相关程度。其中,面部匹配阈值设置为 0.3,余弦相似度低于这个阈值则认为面部匹配成功,超出阈值则匹配失败,根据匹配结果统计出 TP、FN、TN、FP 四个基础指标。统计结果的线性组合计算精确率、召回率和 F1 值。各指标对比结果如表 1 中所示,本文模型在这 3 项指标上均达到了最优值,精确率、召回率、F1 值验证了本文模型在超分辨率任务上对面部身份保持的有效性。并且,本文方法的生成图片在精确率指标上超过原图(Groundtruth, GT)的值,这是因为 LFW 数据集源自真实场景,图片质量低,而部分生成图片比原图质量更高,在精确率指标上取得了更好的效果。

### 3.2.2 模型收敛性分析

为了验证 IPNet 方法的收敛性,本文与其他基于生成器结构的超分辨率方法作对比研究,比较各个模型在 FFHQ 数据集上训练时多个阶段的 PSNR 收敛情况。模型收敛性对比图如图 7 所示,从图中可以看出,相较于 SRGAN 和 ESRGAN 模型,本文模型训练过程更加平稳,几乎没有振荡的情况。当迭代到 60 轮次的时候,模型基本趋近于收敛状态,并且 PSNR 始终保持较好的效果。而 SRGAN、ESRGAN 在训练时会出现较大幅度振荡,收敛速度慢,并且收敛时 PSNR 指标的表现效果不如本文方法。这几组对比实验充分验证了 IPNet 方法的收敛性。使用 Res-UNet 结合预训练的面部生成器结构,有助于模型效果提升,加快网络的收敛能力,并且在训练过程中表现更加稳定。

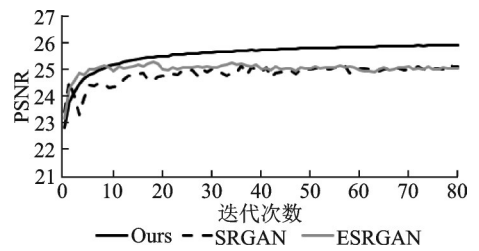


图7 模型收敛性对比

Fig.7 Model convergence comparison

### 3.3 消融实验

在 FFHQ 和 LFW 测试集上完成消融实验,研究各个模块的影响,消融结果如表 2 所示,可视化结果

如图8所示。其中,表2和图8中的编号对应本文消融实验去除的模块:(1)身份损失,(2)SFT调制,(3)Res-UNet,(4)生成先验,(5)感知损失。

表2 在FFHQ-TEST和LFW-TEST测试集上消融实验的结果  
Table 2 Ablation experimental results on FFHQ-TEST and LFW-TEST test datasets

方法	FFHQ-TEST					LFW-TEST		
	PSNR	SSIM	LPIPS	FID	Dis	精确率	召回率	F1值
(1)无身份损失	26.062 2	0.807 9	0.123 3	14.983 2	0.074 1	0.908 2	0.697 9	0.789 3
(2)无SFT调制	15.672 5	0.420 6	0.344 5	52.909 9	0.401 9	0.000 0	0.000 0	0.000 0
(3)无残差连接的Unet	26.073 6	0.808 2	0.123 7	14.731 8	0.064 6	0.906 3	0.693 6	0.785 8
(4)无生成先验	26.058 9	0.808 1	0.123 8	14.785 4	0.064 9	0.913 9	0.699 1	0.792 2
(5)无感知损失	24.659 5	0.763 3	0.168 4	22.922 2	0.069 0	0.913 8	0.698 1	0.791 5
IPNet(Ours)	26.119 2	0.809 3	0.122 5	14.585 3	0.064 6	0.914 0	0.699 7	0.792 7



图8 消融实验可视化结果

Fig.8 Visual results of ablation experiments

#### (1) 身份损失的影响

本文通过对比加入身份损失前后模型在各项评估指标上的表现,研究身份损失对超分辨率重建的影响。从表2可以看出,模型加入身份损失生成的面部图像质量要优于未加身份损失的模型。并且在Dis、精确率、召回率和F1值这4项指标上均达到最优,说明了加入身份损失的超分辨率重建图像在面部特征空间上更接近原图,在重建高分辨率图像过程中能更好地保持面部身份。这一对比结果说明身份损失为FSR任务提供了高层次的面部身份约束,有利于面部超分辨率图像重建。

#### (2) SFT的影响

本文分别比较了模型去除和加入空间调制层SFT前后的对比效果。当去除空间调制层时,即仅使用 $F_{\text{latent}}$ 作特征映射, $F_{\text{spatial}}$ 不对其进行空间调制。从生成图效果对比中可以看出,不考虑面部空间信息的超分辨率模型生成效果要远远低于加入空间信息调制的生成效果,并且重建出的图像不能保持面部身份。表2的实验对比结果证明了本模型加入SFT层的有效性,使用多分辨率空间特征对中间卷积特征进行调制对于超分辨率重建任务至关重要。

#### (3) 语义分割的影响

受残差连接思想的启发,将Unet的每一个子模块分别替换为具有残差连接的形式。相较于Unet结构,Res-UNet结构能避免因感受野扩大导致的面部细节丢失问题。从表2可以看出,使用带有残差连接的Res-UNet作为语义分割网络,重建出的面部图片质量要优于不带残差的Unet。

#### (4) 生成先验的影响

通过预训练StyleGAN2模型,封装大量生成先验,为本文模型生成超分辨率图像过程提供丰富多

变的GAN特征。移除预训练的StyleGAN2模型,以达到消除生成先验的目的。从表2的实验可以看出,消除生成先验后的模型,在各指标上均出现下降。

#### (5) 感知损失的影响

VGG网络分别对真实图片和生成图片进行卷积,得到两个特征并计算比较相似度,通过训练优化使得生成图片在高层信息上与原图接近。从表2可以看出,去除感知损失的模型在图像感知相似度指标LPIPS表现下降严重。

## 4 结束语

本文提出了一种身份保持约束下的面部图像超分辨率方法IPNet,克服低分辨率情况下人脸识别性能下降的问题。该方法利用语义分割模块提取低分辨率图像中有效的面部特征,通过MLP层映射特征以及SFT层调制特征,指导面部生成器重建出高分辨率图像。本文还引入人脸识别网络,将面部身份信息整合到生成器中以实现重建后的生成图像与原图身份保持一致。实验结果表明,本文方法在超分辨率图像质量和身份保持上都展现出其优越性。

### 参考文献:

- [1] KIM J, LI G, YUN I, et al. Edge and identity preserving network for face super-resolution[J]. *Neurocomputing*, 2021, 446: 11-22.
- [2] 薛杉,朱虹,王婧,等. 低分辨率人脸图像的迭代标签传播识别算法[J]. *模式识别与人工智能*, 2018, 31(7): 602-611.  
XUE Shan, ZHU Hong, WANG Jing, et al. Low-resolution face recognition based on recursive label propagation algorithm[J]. *Pattern Recognition and Artificial Intelligence*, 2018, 31(7): 602-611.
- [3] TAHERKHANI F, NASRABADI N M, DAWSON J. A deep face identification network enhanced by facial attributes prediction[C]//*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*. Salt Lake City: IEEE, 2018: 553-560.
- [4] JOSE E, GREESHMA M, HARIDAS M T P, et al. Face recognition based surveillance system using FaceNet and MTCNN on Jetson TX2[C]//*Proceedings of the 2019 5th International Conference on Advanced Computing & Communication Systems (ICACCS)*. Coimbatore: IEEE, 2019: 608-613.
- [5] 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*. Long Beach: IEEE, 2019: 4690-4699.
- [6] DONG C, LOY C C, HE K, et al. Image super-resolution using deep convolutional networks[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2016, 38(2): 295-307.
- [7] 汪荣贵,汪庆辉,杨娟,等. 融合特征分类和独立字典训练的超分辨率重建[J]. *光电工程*, 2018, 45(1): 1-10.  
WANG Ronggui, WANG Qinghui, YANG Juan, et al. Image super-resolution reconstruction by fusing feature classification and independent dictionary training[J]. *Opto-Electronic Engineering*, 2018, 45(1): 1-10.
- [8] LIM B, SON S, KIM H, et al. Enhanced deep residual networks for single image super-resolution[C]//*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*. Honolulu: IEEE, 2017: 136-144.
- [9] 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*. Nashville: IEEE, 2021: 11896-11905.
- [10] CHEN Y, TAI Y, LIU X, et al. Fsmet: End-to-end learning face super-resolution with facial priors[C]//*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Salt Lake City: IEEE, 2018: 2492-2501.
- [11] YU X, FERNANDO B, GHANEM B, et al. Face super-resolution guided by facial component heatmaps[C]//*Proceedings of the European Conference on Computer Vision (ECCV)*. Munich: Springer, 2018: 217-233.



- [12] WANG Z, CHEN J, HOI S C H. Deep learning for image super-resolution: A survey[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2020, 99: 3365-3387.
- [13] 王旭东, 卫红权, 高超, 等. 身份保持约束下的人脸图像补全[J]. *网络与信息安全学报*, 2018, 4(8): 71-76.  
WANG Xudong, WEI Hongquan, GAO Chao, et al. Identity preserving face completion with generative adversarial networks [J]. *Chinese Journal of Network and Information Security*, 2018, 4(8): 71-76.
- [14] ZHU J, SHEN Y, ZHAO D, et al. In-domain GAN inversion for real image editing[C]//*Proceedings of European Conference on Computer Vision*. Glasgow: Springer, 2020: 592-608.
- [15] 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*. Salt Lake City: IEEE, 2018: 606-615.
- [16] BULAT A, TZIMIROPOULOS G. Super-FAN: Integrated facial landmark localization and super-resolution of real-world low resolution faces in arbitrary poses with GANs[C]//*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Salt Lake City: IEEE, 2018: 109-117.
- [17] MA C, JIANG Z, RAO Y, et al. Deep face super-resolution with iterative collaboration between attentive recovery and landmark estimation[C]//*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Seattle: IEEE, 2020: 5569-5578.
- [18] 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*. Seattle: ACM, 2020: 1551-1560.
- [19] 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*. Long Beach: IEEE, 2019: 4401-4410.
- [20] 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*. Seattle: IEEE, 2020: 8110-8119.
- [21] CHOI Y, CHOI M, KIM M, et al. StarGAN: Unified generative adversarial networks for multi-domain image-to-image translation[C]//*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Salt Lake City: IEEE, 2018: 8789-8797.
- [22] 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*. Seattle: IEEE, 2020: 2437-2445.
- [23] CHAN K C K, WANG X, XU X, et al. GLEAN: Generative latent bank for large-factor image super-resolution[C]//*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Nashville: IEEE, 2021: 14245-14254.
- [24] 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*. Nashville: IEEE, 2021: 9168-9178.
- [25] YANG T, REN P, XIE X, et al. GAN prior embedded network for blind face restoration in the wild[C]//*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Nashville: IEEE, 2021: 672-681.
- [26] HUANG H, HE R, SUN Z, et al. Wavelet domain generative adversarial network for multi-scale face hallucination[J]. *International Journal of Computer Vision*, 2019, 127(6): 763-784.
- [27] LAI S C, HE C H, LAM K M. Low-resolution face recognition based on identity-preserved face hallucination[C]//*Proceedings of the 2019 IEEE International Conference on Image Proceedings (ICIP)*. Taipei, China: IEEE, 2019: 1173-1177.
- [28] 孙忠凡, 周正华, 赵建伟. 基于空间元学习的放大任意倍的超分辨率重建方法[J]. *计算机应用*, 2020, 40(12): 3471-3477.  
SUN Zhongfan, ZHOU Zhenghua, ZHAO Jianwei. Super-resolution reconstruction method with arbitrary magnification based on spatial meta-learning[J]. *Journal of Computer Applications*, 2020, 40(12): 3471-3477.
- [29] XIAO X, LIAN S, LUO Z, et al. Weighted res-unet for high-quality retina vessel segmentation[C]//*Proceedings of the 9th International Conference on Information Technology in Medicine and Education (ITME)*. Hangzhou: IEEE, 2018: 327-331.

- [30] RONNEBERGER O, FISCHER P, BROX T. U-net: Convolutional networks for biomedical image segmentation[C]// Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention. Munich: Springer, 2015: 234-241.
- [31] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[C]//Proceedings of the International Conference on Learning Representations. San Diego: [s.n.], 2015.
- [32] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Las Vegas: IEEE, 2016: 770-778.
- [33] 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: Springer, 2008: 1-14.
- [34] 徐龙飞, 郁进明. 不同优化器在高斯噪声下对LR性能影响的研究[J]. 计算机技术与发展, 2020(3): 7-12.  
XU Longfei, YU Jinming. Study on influence of different optimizers on performance of LR under gaussian noise[J]. Computer Technology and Development, 2020(3): 7-12.
- [35] WANG X, YU K, WU S, et al. Enhanced super-resolution generative adversarial networks[C]//Proceedings of the European Conference on Computer Vision. Munich: Springer, 2018: 63-79.
- [36] KIM J, LEE J K, LEE K M. Accurate image super-resolution using very deep convolutional networks[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas: IEEE, 2016: 1646-1654.
- [37] LEDIG C, THEIS L, HUSZAR F, et al. Photo-realistic single image super-resolution using a generative adversarial network [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Honolulu: IEEE, 2017: 4681-4690.

## 作者简介:



田旭(1998-),男,硕士研究生,研究方向:人脸图像超分辨率、单帧图像超分辨率方法等, E-mail: 20205227055@stu.suda.edu.cn。



刁红军(1978-),男,讲师,研究方向:图形图像处理、中文信息处理等。



凌兴宏(1968-),通信作者,男,副教授,硕士生导师,研究方向:机器学习, E-mail: lingxinghong@suda.edu.cn。

(编辑:夏道家)