



DOI:10.13364/j.issn.1672-6510.20210142

数字出版日期: 2021-11-12; 数字出版网址: <http://kns.cnki.net/kcms/detail/12.1355.N.20211110.1654.001.html>

基于误差反馈和面部后先验信息的人脸超分辨率重建

杨巨成¹, 左美然¹, 魏峰², 孙笑¹, 白亚欣¹, 王嫻¹, 陈亚瑞¹

(1. 天津科技大学人工智能学院, 天津 300457; 2. 天津科技大学机械工程学院, 天津 300222)

摘要: 人脸超分辨率方法对于低分辨率输入重建很有成效, 这些方法往往通过一次或有限次上采样操作提高分辨率, 从而构建一种基于前馈结构的网络模型, 但这些方法不能完全解决低分辨率到高分辨率人脸图像之间的映射关系, 因此生成的图像不够清晰, 视觉感知质量差. 针对上述问题, 本文提出一种基于误差反馈和面部后先验信息的人脸超分辨率重建方法 (face super-resolution using error feedback and facial posterior, EFBNet), 以获得清晰、逼真的人脸图像. 通过一组上下采样操作计算误差, 然后反馈误差调整特征的中间表示. 这可以理解成一种自校正过程, 借助采样操作提供的误差迭代地改变最终结果. 此外, 设计了一种包含通道级的面部注意力损失和相对判别对抗损失的优化目标, 面部注意力损失聚焦于由超分辨率结果预测的精确的后先验信息周围的面部细节, 对比由低分辨率输入或中间特征预测的先验信息, 该方法可以减少失真, 进一步提高生成图像的质量. 相对判别对抗损失有助于学习更清晰的边缘和纹理, 从而提高图像的清晰度. 实验结果表明, 该方法可以获得更好的重建结果.

关键词: 人脸超分辨率; 误差反馈; 面部注意力; 生成对抗网络

中图分类号: TP391

文献标志码: A

文章编号: 1672-6510(2022)02-0035-08

Face Super-Resolution Using Error Feedback and Facial Posterior

YANG Jucheng¹, ZUO Meiran¹, WEI Feng², SUN Xiao¹, BAI Yaxin¹, WANG Yuan¹, CHEN Yaru¹

(1. College of Artificial Intelligence, Tianjin University of Science & Technology, Tianjin 300457, China;

2. College of Mechanical Engineering, Tianjin University of Science & Technology, Tianjin 300222, China)

Abstract: Face super-resolution methods are very effective for low-resolution input reconstruction. These methods often increase the resolution through one or a limited number of up-sampling operations to build a network model based on a feed-forward structure. However, these methods cannot completely solve the mapping relationship between low-resolution and high-resolution face images, so the generated images are not clear enough and the visual perception quality is poor. To solve these problems, in this article we propose a face super-resolution using error feedback and facial posterior (EFBNet) to obtain sharp and realistic face images. Specifically, we calculate the error by a set of up and down sampling operations and then tune intermediate representations of features by feeding back the error. This method can be understood as a self-correction procedure which feeds an error to the sampling operation and iteratively changes the final results. Furthermore, we design a novel optimization objective that involves facial attention loss and relative discriminator adversarial loss. Facial attention loss focuses on the facial details around the accurate posterior information predicted by the super-resolution results. Compared with the prior information predicted by low-resolution input or intermediate features, this method can reduce distortion and further improve the quality of the generated image. The relative discriminator adversarial loss helps to learn sharper edges and textures to improve the clarity of the image. Experimental results indicate that the proposed method can obtain better reconstruction results.

Key words: face super-resolution; error feedback; facial attention; generative adversarial network

收稿日期: 2021-06-12; 修回日期: 2021-08-17

基金项目: 国家自然科学基金资助项目 (61976156, 61702367); 天津市教委科研项目 (2017KJ033)

作者简介: 杨巨成 (1980—), 男, 湖北天门人, 教授, jcyang@tust.edu.cn

人脸超分辨率(face super-resolution, FSR)是一种低水平的计算机视觉任务,其目标是从低分辨率输入中恢复高分辨率人脸图像,提高后续人脸图像任务的性能,如人脸对齐^[1]、人脸检测^[2-3]和人脸识别^[4-6]。随着深度学习的发展,尽管人脸超分辨率方法提供了很多解决方案,但是这些方法依然存在以下两个问题。

一方面,目前典型的方法^[7-8]利用前馈结构,从输入的低分辨率图像计算一系列中间特征,通过一次或有限次上采样操作直接增加图像分辨率到最终高分辨率结果。SRCNN^[9]对输入的低分辨率图像直接进行插值上采样到最终尺寸,然后再进行特征提取和重建。FSRNet^[10]设计了编码解码的结构,在编码器中对低分辨率输入进行一次上采样到中间表示,然后通过解码器对提取的特征进一步上采样到最终尺寸。此外,LapSRN^[11]和EDSR^[12]利用多个上采样子网络,逐步对低分辨率输入进行上采样。上述这些方法采用前馈结构,然而人类视觉系统更倾向于使用反馈连接指导任务^[13]。因此,这些网络由于缺乏反馈难以充分表示低分辨率输入到高分辨率结果的映射,特别是在放大倍数较大(如8倍放大)的情况下导致生成图像存在模糊、纹理粗糙等问题。

另一方面,人脸超分辨率方法的解空间随着放大倍数的增加呈指数增长,所以研究者们尝试使用强大的人脸先验约束生成高质量的人脸图像。PMGMSAN^[14]使用预先训练的网络提取面部成分或解析映射图,并将先验信息输入到后续网络中以恢复超分辨率结果。然而,PMGMSAN直接从低分辨率输入中提取先验信息,使得获得准确的先验信息成为一个困难和挑战。多阶段人脸超分辨率方法^[15]则将先验估计分支嵌入到超分辨率重建分支中。首先对低分辨率输入进行初步采样,然后利用中间特征估计先验信息以进行后续重建,并通过 L_2 损失约束估计的先验信息以及由高清人脸图像生成的先验信息。由于上述方法从低分辨率输入、中间结果或特征中提取人脸先验信息,导致先验信息不准确,并且使用像素级先验损失函数如 L_2 作为约束条件,因此难以提供强有力的约束,使得生成的图像模糊,感知质量较差。

鉴于上述问题,本文提出一种基于误差反馈和面部后先验信息的人脸超分辨率重建方法(face super-resolution using error feedback and facial posterior, EFBNet),并通过一个包含通道级的面部注意力损失和相对判别器对抗损失的优化目标训练网络。首先,采用一系列上下采样操作捕获低分辨率输入到高分

辨率结果之间的相互依赖性,提取多粒度特征,并提出误差反馈机制优化提取到的多粒度特征。通过上采样操作计算生成高维特征图与输入之间的误差,通过下采样操作计算生成低维特征图与输入之间的误差。误差反馈机制可以在早期自校正特征提取,并引导网络提取与面部细节相关区域的特征,以充分表示低分辨率输入到高分辨率结果的映射,生成高质量的人脸图像。然后,为了得到精确的先验信息,采用面部先验信息提取和人脸超分辨率重建相互促进的结构,从超分辨率结果而不是低分辨率输入或中间特征提取精确的先验信息,即后先验信息。先验估计分支利用重建的图像提取后先验信息,精确的后先验信息进而促进生成更高质量的人脸图像。最后,为了借助精确的后先验信息进一步提供强有力的约束,设计了一个新的优化目标,它在 L_2 损失的基础上引入通道级的面部注意力损失和相对判别对抗损失。面部注意力损失聚焦于后先验信息周围区域的特征,相对判别对抗损失专注于锐化的边缘,生成内容精确、纹理逼真的超分辨率结果。在标准数据集上的实验表明,本方法在定量指标上达到了当前最优性能,可视化结果进一步表明该方法在清晰度、失真程度和纹理细节等方面具有显著优势。

1 模型设计

1.1 网络结构

提出了一种基于生成对抗网络的使用误差反馈和面部后先验信息的人脸超分辨率方法。其中输入的低分辨率人脸图像、超分辨率结果和高清人脸图像分别表示为 I_{LR} 、 I_{SR} 和 I_{HR} 。生成器由两个分支组成:基于误差反馈的迭代上下采样重建分支(简称重建分支)和生成后先验信息的人脸关键点估计分支(简称关键点估计分支)。判别器采用ESRGAN^[8]中的结构。整个网络以端到端的方式进行训练。基于误差反馈和面部后先验信息的人脸超分辨率模型如图1所示。

重建分支包含3个部分:粗上采样模块(CUM)、基于误差反馈的迭代上下采样模块(EFSM)和全局跳跃连接。全局跳跃连接绕过粗上采样模块和基于误差反馈的迭代上下采样模块,为结果提供一个插值上采样的图像。因此,粗上采样模块和基于误差反馈的迭代上下采样模块用于恢复 I_{LR} 的残差图像 I_{res} 。

由于从非常低分辨率的输入中获得先验信息不精确,因此首先通过粗上采样模块恢复粗糙的上采样结果。该模块由卷积层和像素重组层组成,输出为

$$F_1 = f_{\text{CUM}}(I_{\text{LR}}) \quad (1)$$

其中 f_{CUM} 表示粗上采样操作. F_1 与关键点估计分支的输出 F_{landmark} 拼接在一起作为基于误差反馈的迭代上下采样模块的输入. F_{concat} 是拼接的结果.

$$F_{\text{concat}} = f_{\text{concat}}(F_1, F_{\text{landmark}}) \quad (2)$$

其中 f_{concat} 表示拼接操作.

在基于误差反馈的迭代上下采样模块中采用循环结构, 其可以展开为 T 次迭代. 第 t 次迭代 ($t \in T$) 中 EFSM 接收 F_{concat} 和上一次迭代的输出 F_{EFSM}^{t-1} . F_{EFSM}^T 表示 EFSM 的最终输出, 即

$$F_{\text{EFSM}}^T = f_{\text{EFSM}}(F_{\text{EFSM}}^{t-1}, F_{\text{concat}}) \quad (3)$$

其中 f_{EFSM} 表示 EFSM 的操作.

为了与全局跳跃连接提供的插值上采样图像匹配, 使用卷积操作得到残差图像 I_{res} , 即

$$I_{\text{res}} = f_{\text{conv}}(F_{\text{EFSM}}^T) \quad (4)$$

其中 f_{conv} 表示卷积操作.

最终的输出 I_{SR} 为

$$I_{\text{SR}} = I_{\text{res}} + f_{\text{up}}(I_{\text{LR}}) \quad (5)$$

其中 f_{up} 表示通过全局跳跃连接提供的插值上采样操作.

关键点估计分支的体系结构遵循文献[15].

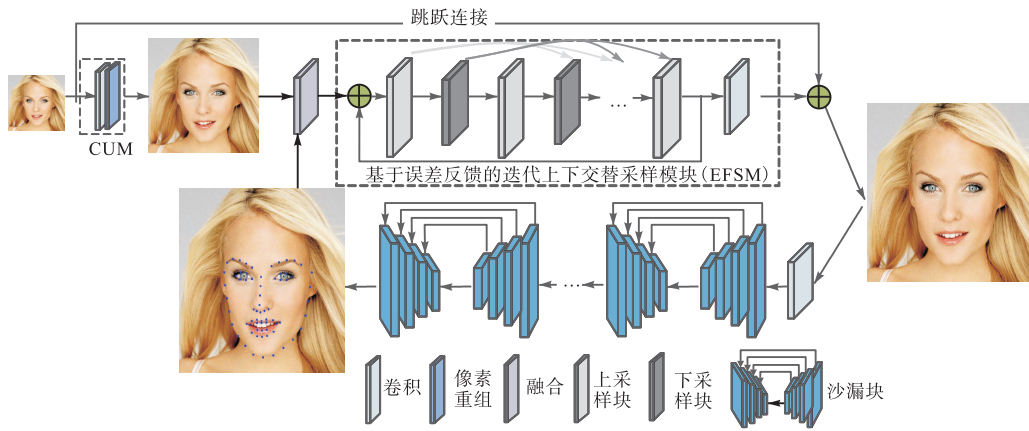


图 1 基于误差反馈和面部后先验信息的人脸超分辨率模型

Fig. 1 Face super-resolution using error feedback and facial posterior information

1.2 基于误差反馈的迭代上下采样模块

基于误差反馈的迭代上下采样模块由一系列上下采样块组成, 共 N 个采样块, 在低分辨率和高分辨率特征图之间交替. 此外, 在此模块中引入密集连接充分利用特征. 先前所有下采样(上采样)块的输出串联起来作为上采样(下采样)块的输入, 如图 1 所示. 这种连接能够有效地利用各种高分辨率成分产生理想的结果.

在第 n 个上采样块中: 先前 $n-1$ 个下采样块计算出的低分辨率特征图 $[L_1, \dots, L_{n-1}]$ 被拼接为 \bar{L}_{n-1} , 并作为第 n 个上采样块的输入 ($n \in N$), 结构图如图 2(a) 所示, 其定义为

$$\begin{aligned} \text{up} : H'_n &= f_{\text{deconv1}}(\bar{L}_{n-1}) \\ \text{down} : L'_n &= f_{\text{conv}}(H'_n) \\ \text{error} : e_n &= L'_n - L_{n-1} \\ \text{error up} : H_{e_n} &= f_{\text{deconv2}}(e_n) \\ \text{output} : H_n &= H'_n + H_{e_n} \end{aligned} \quad (6)$$

其中: f_{deconv1} 和 f_{deconv2} 表示 s 倍的上采样操作, f_{conv} 表

示 s 倍的下采样操作.

第 n 个上采样块将 \bar{L}_{n-1} 映射到中间高分辨率集合 H'_n , 然后再将其映射回低分辨率集合 L'_n . 产生的 L'_n 和 L_{n-1} 之间的误差再次映射到高分辨率, 从而产生新的中间误差集合 H_{e_n} . 通过将两个中间集合相加得到该块的最终输出 H_n .

下采样块的定义与上采样块非常相似, 但它将高分辨率特征图映射到低分辨率空间, 如图 2(b) 所示, 定义为

$$\begin{aligned} \text{down} : L'_n &= f_{\text{conv}}(\bar{H}_{n-1}) \\ \text{up} : H'_n &= f_{\text{deconv}}(L'_n) \\ \text{error} : e_n &= H'_n - H_{n-1} \\ \text{error down} : L_{e_n} &= f_{\text{conv2}}(e_n) \\ \text{output} : L_n &= L'_n + L_{e_n} \end{aligned} \quad (7)$$

这些采样块可以理解为一种自我校正过程, 该过程将误差反馈到采样层, 并通过反馈误差迭代地修改中间表示, 有助于引导网络更关注与面部细节有关的区域.

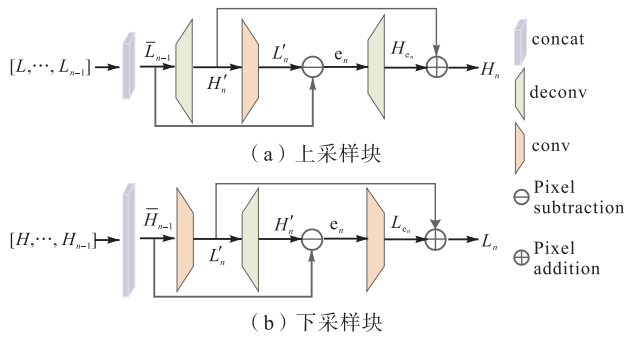


图 2 上采样块和下采样块

Fig. 2 Up-sampling block and down-sampling block

2 优化目标

为了提供强有力的约束,设计优化目标,该优化目标包括关键点损失、面部注意力损失、像素损失、感知损失和相对判别对抗损失.

关键点损失 (L_{landmark}): 在关键点估计分支,使用 L_{landmark} 计算在像素级最小化 I_{SR} 的关键点热图 L_{SR} 和 I_{HR} 的关键点热图 L_{HR} 之间的距离.

$$L_{\text{landmark}} = \frac{1}{N} \sum_{i=1}^N \|L_{\text{HR}} - L_{\text{SR}}\|_2^2 \quad (8)$$

面部注意力损失 ($L_{\text{attention}}$): 除关键点损失外,还使用面部注意力损失关注预测的关键点周围区域的面部细节. 这使关键点估计分支自适应地引导与关键点有关的特征,而不必过多关注那些特征较少的区域.

$$L_{\text{attention}} = M_{\text{max}} \times |I_{\text{HR}} - I_{\text{SR}}|_1 \quad (9)$$

其中 M_{max} 为 L_{HR} 的通道最大值.

像素损失 (L_{pixel}): 使用 L_2 损失作为 I_{HR} 和 I_{SR} 之间的像素损失.

$$L_{\text{pixel}} = \frac{1}{N} \sum_{i=1}^N \|I_{\text{HR}} - I_{\text{SR}}\|_2^2 \quad (10)$$

感知损失 (L_{perc}): 应用感知损失增强超分辨率图像的感知质量^[10,16]. 采用预训练的人脸识别模型 VGG19^[17] 提取图像特征. 该损失通过减小 I_{HR} 和 I_{SR} 特征之间的欧氏距离提高感知相似度,其中 $\phi(I_{\text{SR}})$ 和 $\phi(I_{\text{HR}})$ 分别是通过 VGG19 网络提取的图像特征.

$$L_{\text{perc}} = \frac{1}{N} \sum_{i=1}^N |\phi(I_{\text{HR}}) - \phi(I_{\text{SR}})|_1 \quad (11)$$

相对判别对抗损失 (L_{D}): 不同于标准判别器 D 估计一个输入图像是真实和伪造的概率,本文采用 RaGAN^[18] 的思想试图预测真实图像比伪造的相对更真实的概率.

$$L_{\text{D}} = -E[\log(D_{\text{Ra}}(I_{\text{HR}}, I_{\text{SR}}))] - E[\log(1 - D_{\text{Ra}}(I_{\text{SR}}, I_{\text{HR}}))] \quad (12)$$

其中 D_{Ra} 表示 RaGAN 的判别器.

生成器试图欺骗判别器并最小化对抗损失 L_{adv} .

$$L_{\text{adv}} = -E[\log(1 - D_{\text{Ra}}(I_{\text{HR}}, I_{\text{SR}}))] - E[\log(D_{\text{Ra}}(I_{\text{SR}}, I_{\text{HR}}))] \quad (13)$$

最终,生成器的损失函数定义为

$$L_{\text{G}} = \alpha L_{\text{pixel}} + \beta L_{\text{landmark}} + \gamma L_{\text{perc}} + \mu L_{\text{attention}} + \omega L_{\text{adv}} \quad (14)$$

其中 $\alpha, \beta, \gamma, \mu, \omega$ 均为模型的训练参数.

3 实验与分析

3.1 实验细节

数据集: 模型在两个常用的人脸数据集 Helen^[19] 和 CelebA^[20] 进行了充分的实验. 对于这两个数据集,先使用 OpenFace^[21-22] 检测 I_{HR} 的 68 个关键点. 根据检测的关键点,在每幅图像中裁剪正方形区域以删除背景,并将像素大小调整为 128×128 ,而无需任何预对齐. 然后,将这些高清人脸图像采样到 16×16 的低分辨率输入. 对于 CelebA 数据集,使用 168 854 张图像进行训练,使用 1 000 张图像进行测试. 对于 Helen 数据集,使用 2 005 张图像进行训练,使用 50 张图像进行测试.

实验设置和训练设置: 本文提出的 EFBNet 模型在两个分支之间进行了 4 次迭代. 由于 Helen 数据集的图像数量很少,因此对训练图像执行数据增强,训练图像随机旋转 $90^\circ, 180^\circ, 270^\circ$ 并水平翻转. 模型中的训练参数分别设置为 $\alpha = 1, \beta = 0.1, \gamma = 0.1, \mu = 0.1$ 和 $\omega = 0.05$. 该模型由 ADAM 优化器^[23] 训练,其中 $\beta_1 = 0.9, \beta_2 = 0.999$ 和 $\varepsilon = 1 \times 10^{-8}$. 初始学习率是 1×10^{-4} ,并在 $1 \times 10^4, 2 \times 10^4, 4 \times 10^4$ 次迭代后逐次减半. 实验基于 Pytorch^[24] 在 NVIDIA TITAN RTX (24 G) 上实现.

评估指标: 与之前的模型相似^[9-10,16], 本文使用常用的评估指标峰值信噪比 (PSNR) 和结构相似度 (SSIM)^[25] 评估重建性能.

3.2 与现有方法比较

在两个常用人脸数据集 CelebA 和 Helen 上将本文提出的方法与最新方法进行了比较.

对 CelebA 数据集的评估: 将本文提出的 EFBNet 与其他方法在定性和定量上进行比较,其中包括一般的超分辨率方法 (SRCNN^[9], SRGAN^[16] 和 EDSR^[26])

和人脸超分辨率方法 (FSRNet^[10]和 SAAN^[27]). 所有模型都使用相同的训练集进行训练. 如表 1 所示, 当放大倍数为 8 时, EFBNet 在 PSNR 和 SSIM 方面均明显优于其他方法. 图 3 为在 CelebA 数据集上与现有其他方法的可视化对比, 其中 GT 代表高清人脸图像. 由此可见, EFBNet 借助误差反馈及时进行自校正处理, 获得了更好的推理结果. 精确的后先验信息和强有力的监督进一步优化面部的纹理细节, 提高图像的视觉保真度. 此外, 在图 3 中, EFBNet 不仅能够重建清晰的面部结构, 而且具有很强的鲁棒性, 对于姿势和旋转变大的人脸具有接近真实的恢复效果. 这主要得益于人脸先验估计使用超分辨率结果生成更准确的后先验信息, 从而鼓励重建分支进一步增强人脸图像质量.

表 1 在 CelebA 数据集的定量对比

Tab. 1 Quantitative comparison on CelebA datasets

方法	PSNR/dB	SSIM
SRCNN	25.66	0.721 0
SRGAN	25.57	0.721 3
EDSR	25.26	0.707 8
FSRNet	26.01	0.721 9
SAAN	26.21	0.756 2
EFBNet	27.34	0.796 5

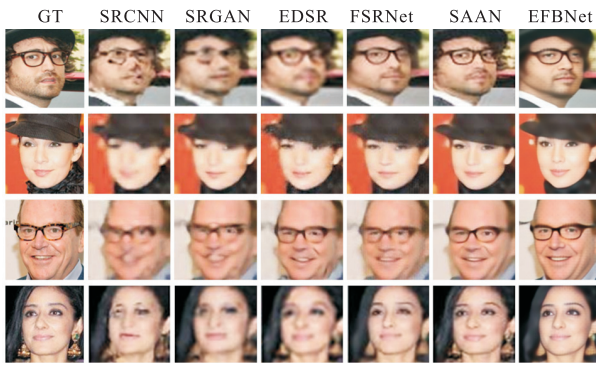


图 3 在 CelebA 数据集上与现有其他方法的可视化对比
Fig. 3 Visual comparison with state-of-the-art methods on CelebA datasets

对 Helen 数据集的评估: 当放大倍数为 8 时, 将本文提出的 EFBNet 与其他人脸超分辨率方法^[10,28-29]进行了定性和定量的比较, 结果如表 2 和图 4 所示, 其中图 4 中 LR 代表输入的低分辨率人脸图像, GT 代表高清人脸图像. 从表 2 中可以看出, EFBNet 在 Helen 测试集上实现了最高性能, 比现有的人脸超分辨率方法 (FSRNet) 高 2.98 dB. 与其他方法的可视化结果对比可看出, 受益于误差反馈和提出的优化目标, EFBNet 生成的人脸图像五官结构完整且更加清

晰. 由于引入面部注意力损失, 其引导网络学习更真实的面部细节, 故相较于其他方法, 在眼睛、嘴巴等部位的纹理细节信息更加丰富且失真程度小.

表 2 在 Helen 数据集的定量对比

Tab. 2 Quantitative comparison on Helen datasets

方法	PSNR/dB	SSIM
FSRNet	24.13	0.668 8
DICNet	26.73	0.795 5
SPARNet	26.97	0.802 6
EFBNet	27.11	0.804 5

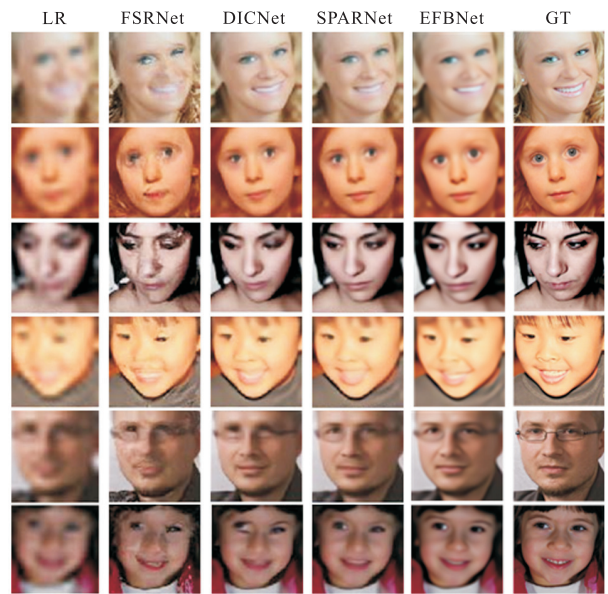


图 4 在 Helen 数据集上与现有方法的可视化对比
Fig. 4 Visual comparison with state-of-the-art methods on Helen datasets

3.3 对比实验

误差反馈的影响: 为了说明误差反馈机制的有效性, 当放大倍数为 8 时, 对比有和没有误差反馈机制 (命名为 EFBNet_L) 对 PSNR 和 SSIM 的影响. 分别用卷积和反卷积操作替换上采样块和下采样块, 则基于误差反馈的迭代上下采样模块类似于先前的超分辨率方法^[30]. 其可视化比较如图 5 所示.

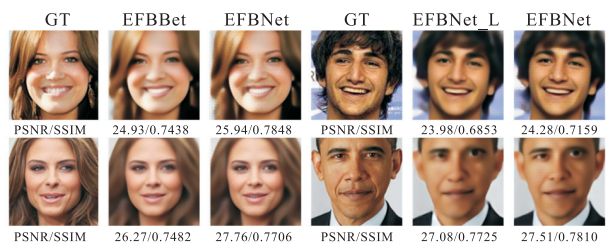


图 5 误差反馈机制对实验结果的影响
Fig. 5 Visual results using error feedback mechanism

由图 5 可知, 本文提出的 EFBNet 能够保留像素

级精度,同时提高面部图像的感知质量.原因是误差反馈将引导特征提取关注与面部细节有关的区域,以优化当前预测.因此,它可以实现从低分辨率输入到超分辨率结果的较优映射,以生成更好细节的人脸图像.

损失函数的影响:3个实验分析面部注意力损失和相对判别对抗损失对 Helen 数据集的影响,分别命名为 model 1、model 2 和 model 3(本文的优化目标).表 3 为使用不同损失函数的定量对比.

表 3 使用不同损失函数的定量对比

Tab.3 Quantitative results of using different loss functions

模型	model 1	model 2	model 3
L_{pixel}	√	√	√
$L_{landmark}$	√	√	√
L_{perc}	√	√	√
$L_{attention}$	×	√	√
L_{adv}	标准对抗损失	标准对抗损失	相对判别对抗损失
PSNR/dB	26.84	27.08	27.11
SSIM	0.784 2	0.795 1	0.804 5

注:“√”表示相应的损失存在

由表 3 可知, model 2 的性能优于 model 1. 这是由于面部注意力损失使网络在通道级更关注面部的关键点区域,从而产生更清晰的面部图像且失真小.此外,相对判别对抗损失有助于学习更锐利的边缘和更详细的纹理.因此, model 3 较 model 2 添加相对判别对抗损失后,定量指标几乎没有改善,但是面部图像更加逼真自然,尤其是纹理细节方面(图 6).总之, model 3 不仅着眼于面部关键点区域以减少失真,而且迫使网络恢复更具结构意义的面部细节以得到清晰逼真的结果.

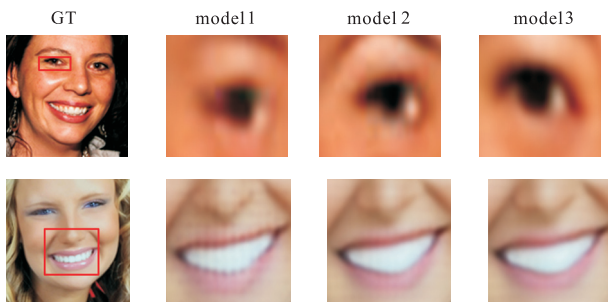


图 6 损失函数对实验结果的影响

Fig. 6 Visual comparisons for showing the effects of different loss function

不同先验信息的影响:与 FSRNext^[10]类似,本文使用人脸对齐模型广泛使用的度量指标归一化均方根误差(NRMSE),探究先验信息提取顺序对超分辨

率图像质量的影响.较低的 NRMSE 值表示更好的对齐精度和更高的 SR 图像质量. FSRNet 在中间特征中提取面部关键点和面部解析映射图指导超分辨率重建,而 SAAN 从输入和中间结果中自适应地提取面部语义信息作为先验信息.不同的先验信息对超分辨率图像质量的影响见表 4.由表 4 可以看出, EFBNet 方法均优于其他方法.虽然其他方法也使用面部先验,例如面部关键点和面部解析映射图等,但先验信息是从输入的低分辨率图像或中间特征中估计的.因此,这种面部先验只能为重建过程提供有限的指导,导致恢复结果存在不完整的、模糊的面部结构.不同的是,本文提出的 EFBNet 使用生成的超分辨率结果估计先验信息,即后先验信息,以便为重建提供更准确的指导信息,得到更清晰的超分辨率结果.

表 4 不同的先验信息对超分辨率图像质量的影响

Tab.4 Influence of different prior information on the quality of super-resolution images

方法	Bicubic	FSRNet	SAAN	EFBNet
NRMSE	0.338 5	0.146 3	0.123 0	0.109 6

4 结 语

本文提出的 EFBNet 方法是基于生成对抗网络使用误差反馈和面部后先验信息指导网络生成高质量的人脸图像,并通过一个新的优化目标训练网络.在该方法中,人脸超分辨率重建和人脸后先验估计相互促进,以产生准确的后先验约束.误差反馈机制帮助重建分支迭代的优化中间表示.优化目标中的面部注意力损失,其基于强大的后先验信息在通道级约束解空间和相对判别对抗损失,其专注于恢复清晰的边缘纹理,驱动网络生成高分辨率的、更好面部纹理细节的图像.在公开的人脸数据集上进行的人脸超分辨率的实验显示出本文方法的显著优势.

致谢:本论文还获得了天津市企业科技特派员项目(20YDTPJC00560)和天津市研究生科研创新项目(人工智能专项)(2020YJSZXB11)的资助.

参考文献:

- [1] JOURABLOO A, YE M, LIU X, et al. Pose-invariant face alignment with a single CNN[C]//IEEE. Proceedings of the IEEE International Conference on Computer Vision. New York: IEEE, 2017: 3200-3209.
- [2] 赵军艳, 降爱莲, 强彦. YOLOv3 融合图像超分辨率重建的鲁棒人脸检测[J/OL]. 计算机工程与应用: 1-9[2021-06-11]. <http://kns.cnki.net/kcms/detail/11.2127>.

- TP.20210420.1527.087.html.
- [3] TZIMIROPOULOS G. Project-out cascaded regression with an application to face alignment[C]//IEEE. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. New York:IEEE,2015:3659–3667.
- [4] 徐润昊,程吉祥,李志丹,等. 基于循环生成对抗网络的含遮挡人脸识别[J/OL]. 计算机工程:1–12[2021–06–11]. <https://doi.org/10.19678/j.issn.1000-3428.0061124>.
- [5] LI Y,LIU S,YANG J,et al. Generative face completion[EB/OL]. [2021–06–11]. <https://arxiv.org/pdf/1704.05838.pdf>.
- [6] ZHU C,ZHENG Y,LUU K,et al. CMS-RCNN: contextual multi-scale region-based CNN for unconstrained face detection[EB/OL]. [2021–06–11]. <https://arxiv.org/abs/1606.05413>.
- [7] YU X,FERNANDO B,HARTLEY R,et al. Super-resolving very low-resolution face images with supplementary attributes[C]//IEEE. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New York:IEEE,2018:908–917.
- [8] WANG X,YU K,WU S,et al. ESRGAN: enhanced super-resolution generative adversarial networks[C]//ECCV. Computer Vision-ECCV 2018 Workshops. Berlin:Springer,2018:63–79.
- [9] 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,2015,38(2):295–307.
- [10] CHEN Y,TAI Y,LIU X,et al. FSRNet: end-to-end learning face super-resolution with facial priors[C]//IEEE. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New York:IEEE,2018:2492–2501.
- [11] LAI W S,HUANG J B,AHUJA N,et al. Deep laplacian pyramid networks for fast and accurate super-resolution[C]//IEEE. 2017 IEEE Conference on Computer Vision and Pattern Recognition(CVPR). New York:IEEE,2017:5835–5843.
- [12] LIM B,SON S,KIM H,et al. Enhanced deep residual networks for single image super-resolution[C]//IEEE. 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops(CVPRW). New York:IEEE,2017:136–144.
- [13] 楼鑫杰,李小薪,刘志勇. 基于反馈机制的图像超分辨率算法[J/OL]. 计算机工程:1–8[2021–06–11]. <https://doi.org/10.19678/j.issn.1000-3428.0059901>.
- [14] WANG C,ZHONG Z,JIANG J,et al. Parsing map guided multi-scale attention network for face hallucination[C]//IEEE. ICASSP 2020–2020 IEEE International Conference on Acoustics, Speech and Signal Processing(ICASSP). New York:IEEE,2020:2518–2522.
- [15] ZHANG Y,WU Y,CHEN L. MSFSR: a multi-stage face super-resolution with accurate facial representation via enhanced facial boundaries[C]//IEEE. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. New York:IEEE,2020:504–505.
- [16] LEDIG C,THEIS L,HUSZÁR F,et al. Photo-realistic single image super-resolution using a generative adversarial network[C]//IEEE. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. New York:IEEE,2017:4681–4690.
- [17] SIMONYAN K,ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[EB/OL]. [2021–06–11]. <http://www.arxiv.org/pdf/1409.1556.pdf>.
- [18] JOLICOEUR-MARTINEAU A. The relativistic discriminator: a key element missing from standard GAN[EB/OL]. [2021–06–11]. <https://arxiv.org/pdf/1807.00734.pdf>.
- [19] LE V,BRANDT J,LIN Z,et al. Interactive facial feature localization[C]//ECCV. European Conference on Computer Vision. Berlin:Springer,2012:679–692.
- [20] LIU Z,LUO P,WANG X,et al. Deep learning face attributes in the wild[C]//IEEE. Proceedings of the IEEE International Conference on Computer Vision. New York:IEEE,2015:3730–3738.
- [21] BALTRUŠAITIS T,ROBINSON P,MORENCY L P. Openface: an open source facial behavior analysis toolkit[C]//IEEE. 2016 IEEE Winter Conference on Applications of Computer Vision(WACV). New York:IEEE,2016:1–10.
- [22] BALTRUSAITIS T,ZADEH A,LIM Y C,et al. Openface 2.0: facial behavior analysis toolkit[C]//IEEE. 2018 13th IEEE International Conference on Automatic Face & Gesture Recognition(FG 2018). New York:IEEE,2018:59–66.
- [23] DA K. A method for stochastic optimization[EB/OL].

- [2021-06-11]. <https://arxiv.org/pdf/1412.6980.pdf>.
- [24] PASZKE A, GROSS S, CHINTALA S, et al. Automatic differentiation in pytorch[EB/OL]. [2021-06-11]. <https://openreview.net/pdf?id=BJJsrnfCZ>.
- [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] LIM B, SON S, KIM H, et al. Enhanced deep residual networks for single image super-resolution[C]//IEEE. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops. New York: IEEE, 2017: 136-144.
- [27] ZHAO T, ZHANG C. SAAN: semantic attention adaptation network for face super-resolution[C]//IEEE. 2020 IEEE International Conference on Multimedia and Expo (ICME). New York: IEEE, 2020: 1-6.
- [28] MA C, JIANG Z, RAO Y, et al. Deep face super-resolution with iterative collaboration between attentive recovery and landmark estimation[C]//IEEE. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. New York: IEEE, 2020: 5569-5578.
- [29] CHEN C, GONG D, WANG H, et al. Learning spatial attention for face super-resolution[J]. IEEE Transactions on image processing, 2020, 30: 1219-1231.
- [30] LI Z, YANG J, LIU Z, et al. Feedback network for image super-resolution[C]//IEEE. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. New York: IEEE, 2019: 3867-3876.

责任编辑: 郎婧

(上接第6页)

- [14] STROBEL G A. Endophytes as sources of bioactive products[J]. Microbes and infection, 2003, 5(6): 535-544.
- [15] 刘洋露, 张杰, 杨志荣, 等. 金银花中产绿原酸的内生细菌的分离与鉴定[J]. 四川大学学报(自然科学版), 2014, 51(3): 604-608.
- [16] 冷慕婵, 董诚明, 乔毅琳, 等. 金银花中产绿原酸内生真菌的分离与鉴定[J]. 河南农业科学, 2014, 43(7): 107-111.
- [17] WEBER R W S, STENGER E, MEFFERT A, et al. Brefeldin A production by *Phoma medicaginis* in dead pre-colonized plant tissue: a strategy for habitat conquest[J]. Mycological research, 2004, 108(6): 662-671.
- [18] NAKAMURA S, YANG C S, SAKON N, et al. Direct metagenomic detection of viral pathogens in nasal and fecal specimens using an unbiased high-throughput sequencing approach[J]. PLOS ONE, 2009, 4(1): e4219.
- [19] 康兴娇, 贾招闪, 申红妙, 等. 内生枯草芽孢杆菌 CN181 发酵条件优化[J]. 山东农业大学学报(自然科学版), 2016, 47(5): 647-653.
- [20] MAGDALENA J S, AGNIESZKA Z G. Analysis of antioxidant activity, chlorogenic acid, and rutin content of camellia sinensis infusions using response surface methodology optimization[J]. Food analytical methods, 2014, 7(10): 2033-2041.

责任编辑: 郎婧