

基于 WGAN 单帧人脸图像超分辨率算法

周传华^{1,2}, 吴幸运¹, 李 鸣¹

(1. 安徽工业大学 管理科学与工程学院, 安徽 马鞍山 243002;
2. 中国科学技术大学 计算机科学与技术学院, 安徽 合肥 230026)

摘 要:针对人脸超分辨率重建中引入的先验知识不够丰富的问题,提出了一种基于 Wasserstein 生成对抗网络(WGAN)的人脸超分辨率重建算法。模型包含生成网络和判别网络,生成网络去除批量规范层并增加残差块数量加深网络深度,判别网络增加特征图通道数并引入了快捷连接优化网络,模型用 Wasserstein 距离代替 KL 散度作为网络的对抗损失,交替训练生成网络和判别网络,生成高分辨率的人脸图像。实验结果表明,相比原始生成对抗网络超分辨率重建算法(SRGAN),所提算法在 MS-Celeb-1M 和 LFW 数据集中峰值信噪比(PSNR)和结构相似性(SSIM)分别提高了 0.26 dB、2% 和 0.31 dB、3%,同时对比最近邻(NN)、双三次插值(Bic)、基于卷积神经网络超分辨率重建(SRCNN)、SRGAN,所提算法在 LFW、MS-Celeb-1M 数据集上均重建出视觉效果更好的人脸图像,证明了该算法的有效性,为人脸超分辨率重建提出了新的解决方案。

关键词:生成对抗网络; Wasserstein 距离; 残差网络; 超分辨率重建; 深度学习

中图分类号: TP391.41

文献标识码: A

文章编号: 1673-629X(2020)09-0029-07

doi: 10.3969/j.issn.1673-629X.2020.09.006

Single Frame Face Images Super-resolution Algorithm Based on WGAN

ZHOU Chuan-hua^{1,2}, WU Xing-yun¹, LI Ming¹

(1. School of Management Science & Engineering, Anhui University of Technology, Maanshan 243002, China;
2. School of Computer Science & Technology, University of Science & Technology of China, Hefei 230026, China)

Abstract: Aiming at lack of prior knowledge in the super-resolution reconstruction of face images, we propose a face super-resolution method based on Wasserstein generative adversarial network (WGAN). A generative net and adversarial net are designed in the model. For generative net, the batch-normalization layer of the standard residual block is removed and the number of residual block is increased to deepen network depth. Also, the number of feature maps is increased and skip connection is used in the discriminative to improve the network performance. The model uses Wasserstein distance instead of KL divergence as network adversarial loss. Alternate training generation network and discriminant network to generate high-quality face images. The experiment shows that compared with SRGAN, the improved algorithm has an increase of 0.26 dB, 2% and 0.31 dB, 3% in peak-signal-to-noise ratio (PSNR) and structural similarity (SSIM) in MS-Celeb-1M and LFW datasets. Compared with nearest neighbor (NN), Bicubic, SRCNN, SRGAN, the improved algorithm can reconstruct the face images with better visual effect. It is proved that the proposed algorithm can effectively improve the resolution of face images and provides a new solution for face super-resolution reconstruction.

Key words: generative adversarial networks; Wasserstein distance; ResNet; super-resolution reconstruction; deep learning

0 引 言

超分辨率重建是指通过技术手段将单帧或多帧低分辨率图像重建成相应的高分辨率图像,超分辨率重建已经在图像处理和计算机视觉方面被广泛应用,如红外图像^[1-2]、医学影像^[3-4]、人脸识别^[5]等场景。人脸超分辨率重建即通过超分辨率重建模型输入一张或

多张低分辨率人脸图像,输出相应的纹理更清晰,视觉效果更好的高分辨率人脸图像,文中主要针对单帧的人脸图像进行超分辨率重建。

由于深度学习^[6]在图像处理和计算机视觉领域展现出的巨大潜力,使得基于深度学习的超分辨率重建技术逐渐成为研究热点。基于深度学习的超分辨率

收稿日期: 2019-09-17

修回日期: 2020-01-20

基金项目: 国家自然科学基金(61702006)

作者简介: 周传华(1965-),男,教授,CCF会员(E8785M),博士,研究方向为机器学习、数据挖掘、智能算法;吴幸运(1995-),男(通信作者),硕士研究生,CCF会员(E1542G),研究方向为机器学习、智能优化。

重建算法主要利用相关卷积神经网络及其改进架构进行图像超分辨率重建。2014 年, Dong 等^[7]率先将卷积神经网络与超分辨率重建相结合, 提出了卷积神经网络超分辨率重建算法 (SRCNN)。该算法先经过双三次插值进行图像上采样, 然后通过三层卷积的神经网络结构学习低分辨率与高分辨率图像间的映射关系, 通过该方法获得的高分辨率图像相对传统方法有较大提高, 但由于神经网络的结构太浅, 难以学习到更深的特征。之后, Dong 等^[8]对其改进, 提出了网络深度更深的 FSRCNN。该网络架构将原先三层的卷积层提升到 8 层, 并在重建网络的最后用反卷积层进行上采样, 代替图像在输入过程中用双三次插值进行上采样, 减少了图像卷积过程的参数, 加快了重建速度并取得了比 SRCNN 更好的重建效果。但 8 层的网络结构深度依然较浅, 对图像的重建效果仍然有限。Kim^[9]制定了深度递归卷积神经网络进行超分辨率重建, 网络相较 SRCNN 通过利用更多的卷积层来增加重建网络的局部感受野, 并进一步减少过多的参数, 重建效果较好。深度更深的网络因为允许建模高度复杂的映射, 能大幅提高训练效果, 但更深的网络结构很难训练^[10]。残差网络 (ResNet)^[11] 和快捷连接 (skip connection) 被应用训练更深层次的网络, 可以提高训练速度, 降低模型训练的内存消耗。Ledig 等人^[12]利用上述训练深层次网络结构的方法, 结合了生成对抗网络 (GAN) 和残差网络, 构建了基于生成对抗网络的超分辨率重建模型 (SRGAN), 实验结果表明 SRGAN 在图像超分辨率重建方面较其他的重建算法生成了视觉效果更好的图像。

将超分辨率重建应用到人脸图像, 可以有效提高人脸图像的清晰度, 增强原始图像的纹理细节信息, 改善图片的质量。鉴于生成对抗网络在其他图像超分辨率重建方面表现出的良好视觉效果, 文中将其应用到人脸图像领域, 并在其基础上针对生成网络, 判别网络和损失函数做出改进, 提出基于 W-GAN 的人脸超分辨率算法 (SRWGAN)。对比其他超分辨率重建算法, 在 4 倍的人脸超分辨率重建中增加了更多的高频信息, 重建出的人脸图像效果更好。

1 相关理论

1.1 生成对抗网络

生成对抗网络启发自二人零和博弈, 由 Goodfellow^[13]在 2014 年提出。该框架包含生成网络和判别网络, 它们的功能分别是生成网络 G 接收噪声 z 生成一个类似真实训练数据的样本, 判别网络 D 判别输入判别器网络的样本来源于生成样本还是真实样本, 通过不断交替地训练生成器和判别器来训练模型,

最后判别器无法辨别样本真伪时, 生成器模型训练完成。其基本架构如图 1 所示。

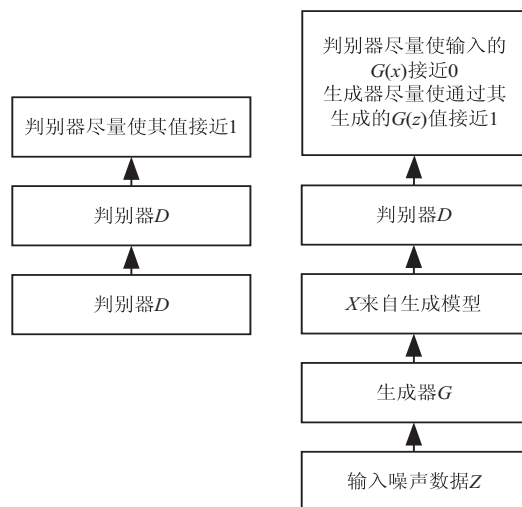


图 1 GAN 基本框架

图 1 所示为 GAN 的基本框架, 通过对抗过程来估计生成模型。该框架包含生成网络和判别网络, 它们的功能分别是生成网络 G 接收噪声 z 生成一个类似真实训练数据的样本, 判别网络 D 判别输入判别器网络的样本来源。具体训练为先将真实样本 x 输入判别网络, 判别网络通过自身的训练, 尽可能输出概率为 1 的值; 其次, 将噪声 z 输入到生成网络生成样本 $G(z)$, 将生成的样本 $G(z)$ 作为判别网络的输入, 判别器网络尽量使 $D(G(z))$ 接近 0, 而生成器尽量使它接近 1。两者相互博弈, 当判别器网络无法准确识别样本来源时达到平衡。

生成对抗网络的训练是通过训练判别网络 D 最大化判别准确率, 同时训练生成网络 G 最小化 $\log(1 - D(G(z)))$, 即训练生成网络 G 和判别网络 D 是关于函数 $V(G, D)$ 的极小极大的博弈:

$$\min_G \max_D V(G, D) = E_{x \sim p_{\text{data}}(x)} [\log D(x)] + E_{z \sim p_D(z)} [\log(1 - D(G(z)))] \quad (1)$$

1.2 W-GAN

为了解决原始的生成对抗网络不易收敛的问题, Martin^[14]提出了用 Wasserstein 距离代替 KL 或 JS 散度的损失函数训练方法, Wasserstein 距离定义如下:

$$W(P_r, P_g) = \inf_{\gamma \sim \prod(P_r, P_g)} E_{(x, y) \sim \gamma} [\|x - y\|] \quad (2)$$

$\prod(P_r, P_g)$ 是 P_r 和 P_g 组合起来的所有可能的联合分布的集合, 计算该联合分布 γ 下样本对距离的期望值 $E_{(x, y) \sim \gamma} [\|x - y\|]$ 。在所有可能的联合分布中能够对这个期望值取到的下界 $\inf_{\gamma \sim \prod(P_r, P_g)} E_{(x, y) \sim \gamma} [\|x - y\|]$ 。

通过数学变换将 Wasserstein 距离应用到生成对

抗网络得到如下公式:

$$W(P_r, P_g) = \frac{1}{K} \sup_{w: \|f_w\|_L \leq K} E_{x \sim P_r}[f_w(x)] - E_{x \sim P_g}[F_w(x)] \quad (3)$$

要求函数 f_w 的 Lipschitz 常数 $\|f_w\|_L$ 不超过 K 的条件下,对所有可能满足条件的 f_w 取到 $E_{x \sim P_r}[f_w(x)] - E_{x \sim P_g}[F_w(x)]$ 的上界,然后再除以 K ,即近似求解:

$$KW(P_r, P_g) \approx \max_{w: \|f_w\|_L \leq K} E_{x \sim P_r}[f_w(x)] - E_{x \sim P_g}[F_w(x)] \quad (4)$$

通过构造一个含参数 w 、最后一层不是非线性激活层的判别器网络 f_w ,在限制 w 不超过某个范围的前提下,使得 $L = E_{x \sim P_r}[f_w(x)] - E_{x \sim P_g}[F_w(x)]$ 尽可能取到最大,此时 L 就会近似真实分布与生成分布之间的 Wasserstein 距离(忽略常数倍数 K)。同时生成器要近似地最小化 Wasserstein 距离,即最小化 L ,通过整理得到生成对抗网络的两个损失函数: $-E_{x \sim P_r}[f_w(x)]$ (生成损失)和 $E_{x \sim P_g}[f_w(x)] - E_{x \sim P_g}[F_w(x)]$ (判别损失)。

WGAN 相对于原始 GAN,做出如下改进:

- (1) 判别器最后一层去掉 sigmoid;
- (2) 生成器和判别器的 loss 不取 log;
- (3) 每次更新判别器的参数之后把它们的绝对值截断到不超过一个固定常数 c ;
- (4) 不要用基于动量的优化算法。

1.3 残差网络和快捷连接

随着网络深度的加深,模型会出现梯度消失与梯度爆炸问题。为了解决该问题,He 等人^[11]提出了残差网络模型(residual network, ResNet),通过残差学习和恒等映射来优化网络结构,使网络学习能达到更好的效果。ResNet 网络根据网络层数和相关组件的不同可分为 ResNet34、ResNet50、ResNet101 和 ResNet152 等。当前的 ResNet 网络已在网络深度 1 000 层以上仍取得了良好的效果,ResNet 基本组件如图 2 所示。

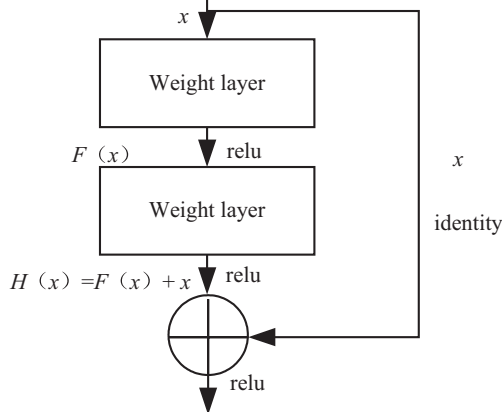


图 2 ResNet 网络基本组件示意图

图 2 曲线部分表示恒等映射 x ,将残差块定义

如下:

$$y = \mathfrak{S}(x, \{W_i\}) + x \quad (5)$$

这里 x 和 y 分别表示输入和输出向量,函数 $\mathfrak{S}(x, \{W_i\})$ 为要学习的残差映射,对于图 2 的示例, $\mathfrak{S} = W_2 \sigma(W_1 x)$ (σ 表示激活函数,并且为了简化符号,这里省略了偏差);通过快捷连接将原始要学习的 $H(x)$ 转换为 $H(x) = \mathfrak{S}(x) + x$,即对 $\mathfrak{S}(x)$ 的学习,使得网络更容易被训练,有效提高了网络的性能。

2 基于 W-GAN 的人脸超分辨率重建

2.1 改进思想

文中的网络结构基于原始的 SRGAN 进行改进,针对其生成网络、判别网络和损失函数分别进行改进。

2.1.1 生成网络的改进

原始 SRGAN 模型的生成器网络是包含 16 个残差块的 ResNet 网络,每个残差块包含卷积层、BN 层、激活函数(ReLU),并通过跳跃连接构建残差块,最后通过两层亚像素卷积层进行上采样。其中,BN 层在构建识别任务的相关卷积神经网络模型中被广泛使用,通过使用一个 batch 数据的均值和方差对特征进行归一化,能有效对深层次的网络进行训练,防止梯度消失和梯度爆炸。但在超分辨率重建任务中模型对图像 Scale 比较敏感,BN 层会使特征标准化,同时限制了模型的泛化能力,增加了 CPU 的负担,相关实验^[15]也证明去掉 BN 层的相关神经网络超分辨率重建模型获得了更好的收敛效果,并减少计算复杂度。因此文中在构建生成器网络时去除了残差块中的 BN 层,同时,深度更深的网络能学习更多的高频信息^[16],引入更多的先验知识。因此,生成网络模型在原有 16 层网络深度的基础上进一步加深网络深度,构建 32 层残差块的生成模型。

2.1.2 判别网络的改进

判别网络以原始的 VGG19 网络作为判别网络,用来判别输入判别网络的图像是真实样本图像(HR)还是生成网络重建出的图像(SR)。但 VGG19 的模型结构较为复杂,深度较深,且特征图通道数较少,模型在训练过程中难以收敛。因此,文中设计判别网络模型时较原有模型增加了特征图通道数^[15],并引入快捷连接,同时基于 Wasserstein 距离作为对抗损失函数,判别网络由原来二分类任务变成近似拟合 Wasserstein 距离,变成回归任务。因此,判别网络最后去除 Sigmoid 层。

2.1.3 损失函数的改进

SRGAN 采用原始生成对抗网络的损失函数 KL 散度作为对抗损失,在训练时存在难以收敛、训练不稳定等问题,因此文中采用 Wasserstein 距离代替原始生

成对抗网络的损失函数 KL 散度作为对抗损失,加快网络收敛,并使用训练 W-GAN 的方式进行训练。

2.2 基于 W-GAN 人脸超分辨率重建网络结构

文中采用 W-GAN 来进行人脸超分辨率重建,并采用 2.1 的改进思想设计具体的生成网络、判别网络和损失函数。

生成器网络结构如图 3(a)所示,输入低分辨率图像,先进入卷积层进行卷积操作(卷积层包含 64 个 9×9 的滤波器,通道数为 3,步长为 1),再使用 PreLU 进行非线性映射,然后依次进入 32 个相同的残差块网络进行训练,每个残差块去除了 BN 层,包含两个 $3 \times 3 \times 3 \times 64$ 的卷积层和 PreLU 激活层,并使用 Skip

Connection 加快网络训练,用两次亚像素卷积进行图像的上采样增加图片分辨率,最后通过卷积层输出重建的高分辨率图像。

为了判别生成的 SR 样本是否和真实样本足够相似,训练判别网络 D ,其网络架构如图 3(b)所示。输入生成的 SR 样本,先经过 6 层卷积(包含 64 个 4×4 的卷积核,3 个通道)抽取图片特征,并借鉴 VGGNet,卷积层的通道数以 2 倍逐层增加,由最初的 64 个通道逐层增加到第 6 层的 2 048 个;接着通过 1×1 卷积进行降维,然后通过两层卷积并使用 Skip Connection 加快网络训练,再将图片输入 Flatten 层进行维度压平,最后经 Dense 层输出结果。

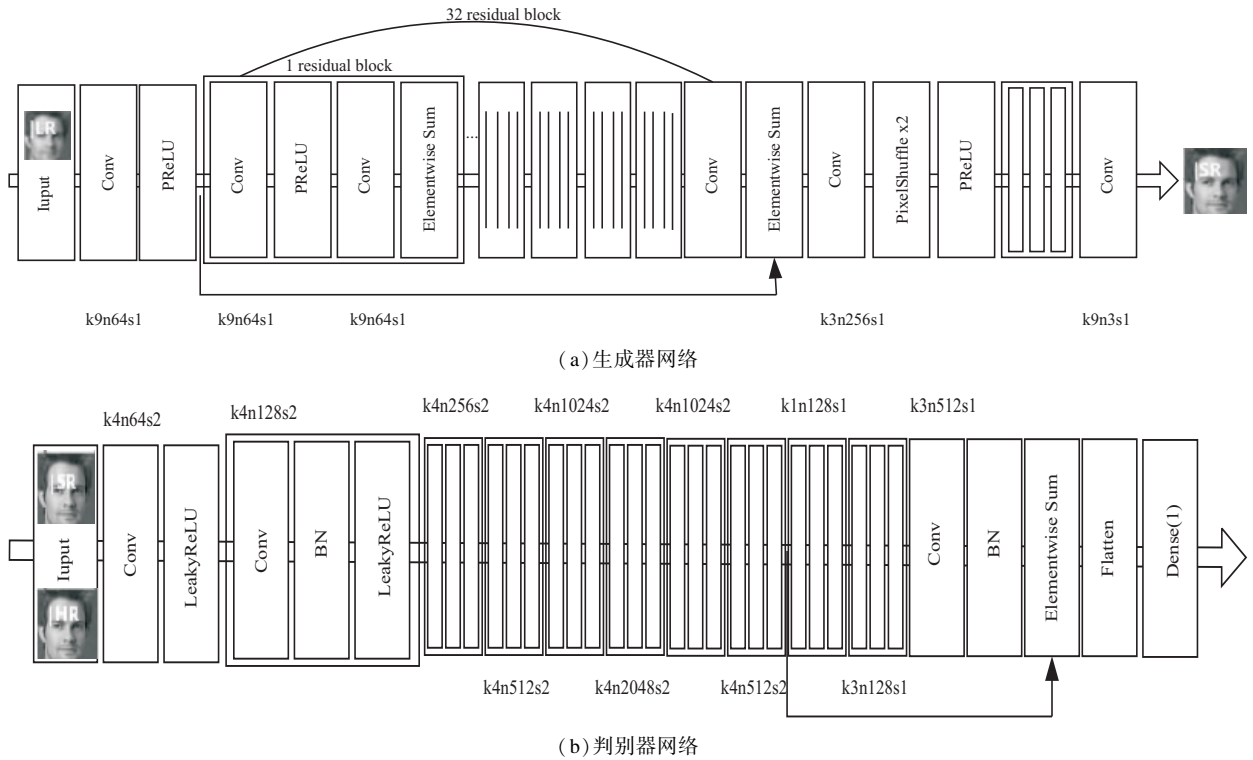


图 3 SRWGAN 网络结构

2.3 损失函数

设计生成网络和判别网络的损失函数,通过交替训练生成网络和判别网络获得更好的重建效果,损失函数如下:

$$\text{loss} = \{\min(l_G^{\text{SR}}), \min(l_D^{\text{SR}})\} \quad (6)$$

优化损失函数即优化生成网络的损失函数和判别网络的损失函数。

2.3.1 生成网络的损失函数

鉴于 W-GAN 相对 GAN 易收敛且生成的样本效果更好,文中选择 Wasserstein 距离作为对抗损失,同时像素的 MSE 损失通常与 PSNR 相关,优化 MSE 能实现更高的 PSNR。综合考虑,文中的生成损失函数采用像素 MSE 损失加上对抗损失,公式如下:

$$l_G^{\text{SR}} = l_{\text{MSE}}^{\text{SR}} + 10^{-3} l_{\text{Gen}}^{\text{SR}} \quad (7)$$

像素损失 $l_{\text{MSE}}^{\text{SR}}$ 为:

$$l_{\text{MSE}}^{\text{SR}} = \frac{1}{r^2 WH} \sum_{x=1}^{rW} \sum_{y=1}^{rH} (I_{x,y}^{\text{HR}} - G_{\theta G}(I^{\text{LR}})_{x,y})^2 \quad (8)$$

其中, W, H 分别表示原始低分辨率图像的长度和高度, r 表示放大的倍数, $I_{x,y}^{\text{HR}} - G_{\theta G}(I^{\text{LR}})_{x,y}$ 表示单位面积生成的高分辨率图像和原始高清图像间像素损失。对抗损失 $l_{\text{Gen}}^{\text{SR}}$ 是基于所有训练样本鉴别器的 Wasserstein 距离定义的,将生成的图像输入判别网络,通过判别网络近似生成高分辨率图像和原始高清图像的 Wasserstein 距离。判别网络输出的近似 Wasserstein 距离即是生成网络的对抗损失。

$$l_{\text{Gen}}^{\text{SR}} = \sum_{n=1}^N -D_{\theta D}(G_{\theta G}(I^{\text{LR}})) \quad (9)$$

2.3.2 判别网络的损失函数

判别网络仅考虑对抗损失作为损失函数,分别将

原始高清图像和生成的高分辨率图像输入判别网络,将判别网络分别输出的近似 Wasserstein 距离相加作为判别网络的损失函数。公式如下:

$$l_D^{\text{SR}} = l_{\text{Adv}}^{\text{SR}} = \sum_{n=1}^N D_{\theta_D}(I^{\text{HR}}) - \sum_{n=1}^N D_{\theta_D}(G_{\theta_G}(I^{\text{LR}})) \quad (10)$$

3 实验

3.1 数据

文中使用了进行处理的 MS-Celeb-1M 和 LFW 人脸数据集作为实验数据集,MS-Celeb-1M 是微软提供的世界上最大的人脸公开数据集,主要用于人脸检测识别等任务,包含 100 多万名人的 1 000 万多张人脸照片。实验中,随机选取 3 万个名人的 3 万张人脸作为数据集,并按照 8:2 的比例,选取 24 000 张人脸作为训练集,6 000 张人脸作为测试集。LFW 人脸数据集是由美国马萨诸塞州立大学阿默斯特分校计算机视觉实验室整理完成的数据库,主要从互联网上搜集图像,用来研究非受限情况下的人脸识别问题,一共含有 13 000 多张人脸图像。实验中通过数据预处理剔除重复和不满要求的人脸图像,最终选取 6 000 张人脸作为数据集,并选取 4 800 张人脸图像作为训练集,1 200 张人脸图像作为测试集。

人脸具有五官相对固定的特征,而神经网络对不同位置提取的特征是不同的,因此,固定相应的人脸区域,结合神经网络的局部感知性,可以优化网络训练,加快收敛。文中结合人脸检测对齐算法,针对以上预处理过的 MS-Celeb-1M 和 LFW 人脸数据集通过 OpenCV 和 dlib 返回对齐的人脸图像,具体就是通过人脸检测返回人脸关键点,然后旋转平移进行人脸对齐,并截取 96 * 96 的人脸图像,处理的示例如图 4 所示。

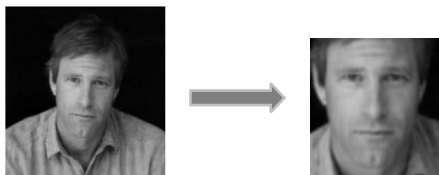


图4 人脸检测对齐



(a) MS-Celeb-1M 数据集各方法重建效果对比

预处理过的 MS-Celeb-1M 和 LFW 数据集经过上述处理组成 HR 人脸数据集,之后对图像进行下采样(进行 4 倍下采样),形成对应的 LR 人脸数据集。

3.2 训练细节和参数

实验环境:Linux 系统环境;python3.6;CUDA10.0+ CUDNN7.60+Tensorflow1.13+Keras2.2+;32 G 内存, RTX2080 GPU(8 G 显存)。

训练过程中为了避免 SRWGAN 陷入局部最优解,首先使用 SRResNet 网络在处理过的 LFW 人脸数据集上进行训练,实验迭代 1 000 000 次,batchsize 设置为 16,学习率初始化为 1×10^{-4} ,每隔 500 000 次学习率变为原来的 0.1 倍,学习动量设置为 0.9,采用 Adam 优化算法更新网络权重。在训练 SRWGAN 时,利用训练好的 SRResNet 网络作为初始化生成模型,交替训练生成模型和判别模型,batchsize 设置为 16,初始学习率为 1×10^{-5} ,迭代 100 000 次,每隔 50 000 次将学习率变为原来的 0.1 倍,同时使用 RMSProp 优化算法进行训练更新网络权重。

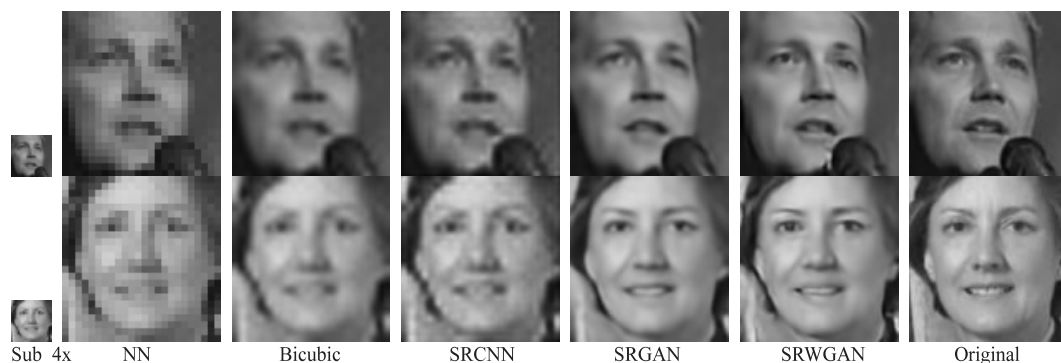
3.3 质量评价

超分辨率重建的算法性能可以通过超分辨率图片质量进行评价,超分辨率质量可分为主观评价和客观评价。主观评价方法主要是通过主观的视觉感受评价生成图片的质量。客观评价主要通过数据量化图片的指标,评价重建图像质量。文中选用常用的峰值信噪比(PSNR)^[17]和结构相似性(SSIM)^[18]作为客观评价的评价指标。

为了充分验证文中方法的有效性,分别在 MS-Celeb-1M 和 LFW 人脸数据集进行了 4 倍的人脸超分辨率重建,分别从主观评价和客观评价两方面进行对比,综合评估该算法的有效性。

3.3.1 主观效果

分别用 NN^[19]、Bicubic^[20]、SRCNN^[7]、SRGAN、SRWGAN 对 MS-Celeb-1M 和 LFW 人脸数据集进行了 4 倍的人脸超分辨率重建,实验结果如图 5(a)、图 5(b)所示。



(b) LFW 数据集各方法重建效果对比

图 5 各方法超分辨率重建效果对比

图 5 为 MS-Celeb-1M 和 LFW 测试集中的人脸图像的超分辨率重建主观视觉效果图(其中上面两个为 MS-Celeb-1M 测试集中的人脸图像,下面两个为 LFW 测试集中的人脸图像)。可以发现对人脸图像进行超分辨率重建时,基于深度学习的 SRGAN 算法和文中所提算法重建出的图像视觉效果较好。具体来说,NN 重建出的人脸图像棋格现象严重,效果最差,Bicubic 和 SRCNN 算法重建出的人脸图像整体看起来模糊,虽然 SRCNN 算法在图像锐度方面有所提升,但清晰度和图像细节方面的表现不足;SRGAN 算法和

文中提出的改进算法都重建处理了视觉效果出色的人脸图像,但 SRGAN 算法在纹理细节方面处理的不够细腻,文中所提算法在视觉上表现出了更逼真的纹理细节。综上所述,文中算法重建表现出了具有最好视觉效果的人脸图像。

3.3.2 客观效果

为了客观评价所提 SRWGAN 算法与其他超分辨率重建的效果,分别计算 MS-Celeb-1M 和 LFW 人脸数据集中重建的人脸图像与原始高清人脸图像的 PSNR 和 SSIM,所得结果如表 1 所示。

表 1 MS-Celeb-1M 数据集各算法评价指标

| 算法模型 | | NN | Bic | SRCNN | SRGAN | SRWGAN |
|-------------|------|-------|-------|-------|-------|--------|
| MS-Celeb-1M | PSNR | 25.04 | 27.09 | 28.01 | 27.06 | 27.32 |
| | SSIM | 0.77 | 0.86 | 0.88 | 0.90 | 0.92 |

表 2 LFW 数据集各算法评价指标

| 算法模型 | | NN | Bic | SRCNN | SRGAN | SRWGAN |
|------|------|-------|-------|-------|-------|--------|
| Lfw | PSNR | 25.11 | 27.70 | 27.75 | 25.88 | 26.19 |
| | SSIM | 0.76 | 0.85 | 0.86 | 0.88 | 0.91 |

由表 1 和表 2 各超分辨率重建算法重建出的图像与原始图像的 PSNR 和 SSIM 指标值可以看出,对于客观评价指标 SSIM,文中算法获得了最高的 SSIM 值,并且 NN、Bicubic、SRCNN、SRGAN、SRWGAN 重建出图像的 SSIM 值依次提高。同时,对于评价指标 PSNR,文中算法的 PSNR 值低于 Bicubic 和 SRCNN 算法,但高于 NN 和 SRGAN 算法。相较 SRGAN 算法,文中算法在客观评价指标上均有所提高,其中,在 MS-Celeb-1M_LR 人脸数据集中 PSNR 和 SSIM 分别提高了 0.26 和 2%,在 LFW_LR 人脸数据集中 PSNR 和 SSIM 分别提高了 0.31 和 3%。

3.3.3 综合评价

虽然 PSNR 广泛应用于图像的客观评价,但其有很大的局限性,最近的大多数实验和文献^[21]都显示了其作为评价指标往往数值无法与视觉感官的图像质量相符合,有些情况,PSNR 较高的反而不如 PSNR 值低

的视觉效果出色。针对此,文中采用客观评价指标 PSNR、SSIM 和主观视觉效果综合的评价方法。

在重建出的人脸图像的主观视觉效果上,SRWGAN 重建出了纹理细节效果更出色的人脸图像。在客观评价上,SRWGAN 算法获得了最好的 SSIM 评价指标数值。PSNR 评价指标数值相较 SRGAN 也有所提高,因此,提出的基于 SRWGAN 的超分辨率算法能重建出效果更好的人脸图像,且效果优于 NN、Bicubic、SRCNN、SRGAN 算法。

4 结束语

针对 4 倍人脸的超分辨率重建提出了基于 WGAN 的人脸超分辨率重建算法,该算法由生成网络生成高分辨率人脸图像,由判别网络近似原始高分辨率图像和生成图像的 Wasserstein 距离,并通过生成图像和原始图像的内容损失以及网络结构的对抗损失构建

损失函数进行网络训练优化。最后在 MS-Celeb-1M 和 LFW 人脸数据集上与其他超分辨重建算法的重建结果对比表现出所提算法的优越性。下一步的研究工作将考虑多帧人脸图像的超分辨率重建,以探索更高效的超分辨率重建算法。

参考文献:

- [1] 李方彪,何 昕,魏仲慧,等.生成式对抗神经网络的多帧红外图像超分辨率重建[J].红外与激光工程,2018,47(2):17-24.
- [2] 孙 刚.基于 MAP 的红外图像超分辨率重建算法研究[D].西安:西安电子科技大学,2008.
- [3] 高 媛,刘 志,秦品乐,等.基于深度残差生成对抗网络的医学影像超分辨率算法[J].计算机应用,2018,38(9):2689-2695.
- [4] 白福栋.基于稀疏理论的医学图像超分辨率重建算法的研究与应用[D].杭州:浙江大学,2016.
- [5] 贾 洁.基于生成对抗网络的人脸超分辨率重建及识别[D].成都:电子科技大学,2018.
- [6] ZHANG W, SHIH F Y, HU S, et al. A visual secret sharing scheme based on improved local binary pattern[J]. International Journal of Pattern Recognition and Artificial Intelligence, 2018, 32(6):1850017. 1-1850017. 19.
- [7] DONG C, LOY C C, HE K, et al. Learning a deep convolutional network for image super-resolution[C]//European conference on computer vision. Cham: Springer, 2014:184-199.
- [8] DONG C, LOY C C, TANG X. Accelerating the super-resolution convolutional neural network[C]//European conference on computer vision. Cham: Springer, 2016:391-407.
- [9] KIM J, LEE J K, LEE K M. Deeply-recursive convolutional network for image super-resolution[C]//Proceedings of the 2016 IEEE conference on computer vision and pattern recognition. Washington, DC: IEEE, 2016:1637-1645.
- [10] SZEGEDY C, LIU W, JIA Y, et al. Going deeper with convolutions[C]//Proceedings of the 2015 IEEE conference on computer vision and pattern recognition. Washington, DC: IEEE, 2015:1-9.
- [11] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]//Proceedings of the 2016 IEEE conference on computer vision and pattern recognition. Washington, DC: IEEE, 2016:770-778.
- [12] LEDIG C, THEIS L, HUSZAR F, et al. Photo-realistic single image super-resolution using a generative adversarial network[C]//Proceedings of the 2017 IEEE conference on computer vision and pattern recognition. Washington, DC: IEEE, 2017:105-114.
- [13] GOODFELLOW I, POUGET-ABADIE J, MIRZA M, et al. Generative adversarial nets[C]//Proceedings of the 27th international conference on neural information processing systems. Cambridge, MA: MIT Press, 2014:2672-2680.
- [14] ARJOVSKY M, CHINTALA S, BOTTOU L. Wasserstein gan[J]. Neural Information Processing Systems Foundation, 2017(12):5768-5778.
- [15] LIM B, SON S, KIM H, et al. Enhanced deep residual networks for single image super-resolution[C]//Proceedings of the 2017 IEEE conference on computer vision and pattern recognition workshops. Washington, DC: IEEE, 2017:1132-1140.
- [16] SUN J, XU Z, SHUM H Y. Image super-resolution using gradient profile prior[C]//Proceedings of the 2008 IEEE conference on computer vision and pattern recognition. Washington, DC: IEEE, 2008:1-8.
- [17] 佟雨兵,张其善,祁云平.基于 PSNR 与 SSIM 联合的图像质量评价模型[J].中国图象图形学报,2006,11(12):1758-1763.
- [18] 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.
- [19] UNSER M. Splines: a perfect fit for signal and image processing[J]. IEEE Signal Processing Magazine, 1999, 16(6):22-38.
- [20] RAMPONI G. Warped distance for space-variant linear image interpolation[J]. IEEE Transactions on Image Processing, 1999, 8(5):629-639.
- [21] 龙法宁,朱晓姝,胡春娇.基于深层卷积网络的单幅图像超分辨率重建模型[J].广西科学,2017,24(3):231-235.