



DOI:10.13364/j.issn.1672-6510.20210241

融合光照损失的图像超分辨率生成对抗网络

陈亚瑞¹, 丁文强¹, 徐肖阳¹, 胡世凯¹, 闫潇宁², 许能华²

(1. 天津科技大学人工智能学院, 天津 300457; 2. 深圳市安软科技股份有限公司, 深圳 518131)

摘要: 本文针对在低光照条件下图像分辨率低的问题, 提出一种融合光照损失的图像超分辨率生成对抗网络(image super-resolution generative adversarial network based on light loss, LSRGAN)模型. 该模型通过构建高分辨率-低分辨率图像对, 利用生成器网络、判别器网络进行训练, 实现低光照条件下更好的模型生成图像效果. 该模型的损失函数包括光照损失、结构相似性损失、内容损失和对抗损失. 模型通过构建光照损失函数, 利用 RGB 三原色颜色空间与 YIQ 颜色空间的线性关系计算出图像中的亮度分量, 将图像中的亮度作为损失函数, 更好地恢复低光照条件下的低分辨率图像; 通过增加结构相似性损失, 计算超分辨率图像与真实高分辨率图像之间的结构相似性, 提高生成图像的质量; 内容损失区别于传统的基于像素的损失, 使用 VGG19 网络中的特征映射进行计算, 可以得到更逼真的生成图像; 对抗损失使用判别器网络区分超分辨率图像与真实高分辨率图像, 提高超分辨率图像的视觉效果. 通过在 4 个标准数据集 Set5、Set14、BSDS100 和 Urban100 上设计对比实验, 证明通过增加对光照更加敏感的损失函数, 使该模型在低光照条件下具有更好的模型生成图像效果; 同时通过增加结构相似性损失, 使生成的图像视觉质量更好.

关键词: 单幅图像超分辨率; 生成对抗网络; 深度学习

中图分类号: TP319 文献标志码: A 文章编号: 1672-6510(2022)04-0055-09

Image Super-Resolution Generative Adversarial Network Based on Light Loss

CHEN Yarui¹, DING Wenqiang¹, XU Xiaoyang¹, HU Shikai¹, YAN Xiaoning², XU Nenghua²

(1. College of Artificial Intelligence, Tianjin University of Science & Technology, Tianjin 300457, China;
2. Shenzhen Softsz Co., Ltd., Shenzhen 518131, China)

Abstract: In this article we propose an image super-resolution generative adversarial network based on light loss (LSRGAN) model that incorporates light loss to address the problem of low image resolution in low light situations. This model constructs high resolution-low resolution image pairs, and uses generator network and discriminator network for training to achieve better model generation image effect in low light situations. The loss functions of the model include light loss, structural similarity loss, content loss and adversarial loss. The model constructs the light loss function, calculates the brightness component in the image by using the linear relationship between RGB three primary color space and YIQ color space, and takes the brightness in the image as the loss function to better restore the low resolution image under low light; by increasing the loss of structural similarity, the structural similarity between super-resolution image and real high-resolution image is calculated to improve the quality of generated image; the content loss is different from the traditional pixel based loss. Using the feature mapping in VGG19 network to calculate, a more realistic generated image can be obtained; to combat the loss, the discriminator network is used to distinguish the super-resolution image from the real high-resolution image, so as to promote the visual effect of the super-resolution image. By designing comparative experiments on four standard data sets: Set5, Set14, BSDS100 and Urban100, it is proved that by adding a more sensitive loss function to the light, the model

收稿日期: 2021-11-15; 修回日期: 2022-03-08

基金项目: 天津市研究生科研创新项目(2020YJSZXS31)

作者简介: 陈亚瑞(1982—), 女, 河北邢台人, 副教授, yrchen@tust.edu.cn

has a better model generation image effect in low light situations. At the same time, by increasing the loss of structural similarity, the visual quality of the generated image is better.

Key words: single image super-resolution; generative adversarial network; deep learning

单幅图像超分辨率(single image super-resolution, SSIR)重建是根据一张低分辨率(low resolution, LR)图像恢复出高分辨率(high resolution, HR)图像的过程. SSIR 作为计算机视觉一个非常重要的研究方向,引起了国内外众多学者的广泛关注,其研究成果有大量的商业应用. 超分辨率重建在医学成像、卫星成像、人脸识别、视频监控以及恢复重要历史图像资料等诸多领域发挥了非常重要的作用.

在深度学习没有兴起之前,传统的 SSIR 通常使用插值法^[1-3]. 插值法具有计算简单、计算速度快等优点,但有时会导致生成图像的边缘和细节模糊. 随着深度学习的不断发展,深度学习在单幅图像超分辨率领域的优异表现,促进了单幅图像超分辨率研究的快速发展,各种神经网络的变体也应用于单幅图像超分辨率的研究中,包括卷积神经网络^[4]、对抗神经网络^[5]以及二者的结合等^[6].

超分辨率生成对抗网络(super-resolution generative adversarial networks, SRGAN)^[7]是将图像超分辨率与生成对抗网络结合,生成视觉效果较好的高分辨率图像. 然而,SRGAN 模型的训练数据由人工构造,根据已知的高分辨率图像,利用双三次插值法^[3]进行下采样,得到低分辨率图像. 人工构造的高分辨率-低分辨率(HR-LR)图像对并没有考虑复杂情况下的图像超分辨率问题,例如低光照、模糊、噪声、运动等复杂情况. 这导致了用人工构造的 HR-LR 图像对训练得到的模型在处理人工生成的低分辨率图像时效果很好,但是处理现实生活中低分辨率图像的效果往往不尽如人意. SRGAN 模型的损失更加关注峰值信噪比(peak signal-to-noise ratio, PSNR),而忽略了结构相似性(structural similarity, SSIM),这使生成的高分辨率图像虽然有较好的峰值信噪比,但是视觉效果不佳. 基于此,本文模型在 SRGAN 模型的基础上进行了 3 个方面的改进:首先,构造一个新的损失函数(光照损失),解决低光照条件下的图像超分辨率问题;其次,增加结构相似性损失,平衡生成图像的 PSNR 和 SSIM,使生成图像的视觉效果更好;最后,构造新的训练数据集,在传统低分辨率图像的基础上,通过减少图像的灰度值降低图像的亮度,得到全新的 HR-LR 图像对,此时的 LR 图像不仅是低分辨率而且还是低光照条件.

1 背景

1.1 图像超分辨率

从低分辨率图像估计其对应高分辨率图像的挑战性任务被称作超分辨率. 传统的图像超分辨率方法通常使用插值法,例如双线性插值法^[2]、双三次插值法^[3]、Lanczos^[7]重采样等. 插值法具有计算简单、计算速度快等优点,但是生成的图像往往比较模糊,纹理细节通常比较平滑. 在深度学习兴起之后,传统的图像超分辨率方法逐渐走向了没落,而基于深度学习的图像超分辨率方法迅速发展,研究人员提出了很多新的方法.

基于深度学习的图像超分辨率方法主要包括基于卷积神经网络^[8-12]、基于递归神经网络^[13-15]、基于循环神经网络^[16]、基于对抗神经网络^[5,17]、基于反馈机制^[18-19]、基于通道注意力机制^[20-22]等. 还有一些特殊的图像超分辨率方法,例如针对多尺度超分辨率重建^[23]、基于弱监督学习的超分辨率重建^[24-25]、基于无监督的超分辨率重建^[26-27]、针对盲超分辨率图像重建^[28-30]等.

1.2 基于卷积神经网络的图像超分辨率

由于卷积神经网络在图像分类、识别等领域有着出色的表现,因此 Dong 等^[8]将卷积神经网络(convolutional neural network, CNN)应用到超分辨率重建中,提出了超分辨率重建卷积神经网络(super-resolution convolutional neural network, SRCNN). 虽然 SRCNN 模型只有 3 层神经网络,但是相较于传统的图像超分辨率方法,取得了显著的效果.

由于 CNN 的优异表现,越来越多的研究人员开始尝试用 CNN 实现图像的超分辨率重建,并提出了许多新的模型. 例如, Dong 等^[9]改进了 SRCNN 模型,提出了快速超分辨率重建卷积神经网络(fast super-resolution convolutional neural network, FSRCNN). Kim 等^[4]提出了超分辨率重建极深卷积神经网络(very deep convolutional network for super-resolution, VDSR)模型,相比于 SRCNN 模型, VDSR 模型使用了更深的神经网络,可以提取更多的特征图,使重建后的图像细节更加丰富. Mao 等^[10]提出了极深残差编解码网络(very deep residual encoder-

decoder networks, RED-Net)模型,该模型用到了编码-解码结构.考虑到 DenseNet 模型的优点,Tong 等^[11]提出了 SRDenseNet 网络.Yu 等^[12]在 2018 年提出了 WDSR 模型,并取得了当年 NTIRE 竞赛的冠军.

目前,尽管卷积神经网络在图像超分辨率重建上很难恢复真实的纹理和细节,生成的图像比较模糊,但是其在图像超分辨率领域仍然占据着重要位置,每年都会会有许多基于 CNN 的图像超分辨率模型被提出.

1.3 基于生成对抗网络的图像超分辨率

传统的图像超分辨率方法和基于卷积神经网络的图像超分辨率方法都是以峰值信噪比作为一项重要的评判指标,导致的结果就是随着模型的不断改进,PSNR 越来越大,但是生成的图像在视觉感知上质量较差.这也证实了,即使 PSNR 数值大,图像的感官质量也并不一定就好.为了生成符合人类感官质量的图像,Ledig 等^[5]提出了 SRGAN 模型,网络主体采用对抗生成网络 (generative adversarial net, GAN),损失函数采用感知损失与对抗损失之和.虽然峰值信噪比不是最高,但是 SRGAN 模型生成的图像更加自然清晰,更符合人眼的视觉效果.尽管 SRGAN 模型取得了很好的视觉效果,但是随着网络的加深,SRGAN 模型生成的图像也会在细节上出现模糊,在图像纹理上出现伪影.

基于 SRGAN 出现的问题,Wang 等^[17]对 SRGAN 模型进行了改进,提出了 ESRGAN 模型.ESRGAN 模型对 SRGAN 模型进行了 3 个方面的改进.首先,引入没有批量归一化的 RRDB (residual-in-residual dense block) 作为基本的网络架构单元.其次,使用相对平均 GAN (RaGAN) 改进判别器网络,有助于生成器网络恢复更多的真实纹理细节.最后,提出了一种改进的感知损失,使用激活之前的 VGG 特征代替 SRGAN 模型中激活之后的 VGG 特征,调整之后的感知损失提供了更清晰的图像边缘和视觉上更令人满意的结果.

1.4 创新与改进

不论是传统的图像超分辨率方法,还是基于深度学习的图像超分辨率方法,其关注的研究领域都是从低分辨率图像恢复出高分辨率图像,较少研究现实生活中存在的复杂的低分辨率图像,例如低分辨率图像同时存在低光照、模糊、噪声等情况.这就导致了在用模型处理现实中的低分辨率图像时,生成图像的效果并不好.

本文的融合光照损失的图像超分辨率生成对抗网络 (image super-resolution generative adversarial

network based on light loss, LSRGAN) 模型是在 SRGAN 模型的基础上进行创新:

(1) LSRGAN 模型能够在低光照条件下的图像超分辨率重建.

(2) 提出新的损失为光照损失.利用 RGB 颜色空间与 YIQ 颜色空间的线性关系计算出图像中的亮度分量,将图像中的亮度作为损失函数,更好地恢复低光照条件下的低分辨率图像.

(3) 将图像的结构相似性作为损失函数添加到原有的损失中,这有利于生成更加真实、自然的图像,提高图像的感官质量.

(4) 构建新的训练数据集.先将高分辨率图像进行变黑处理,得到人造的低光照图像.然后使用双三次插值法进行下采样,得到低光照条件下的低分辨率图像.

2 本文方法

2.1 模型损失函数

LSRGAN 模型的网络结构如图 1 所示.

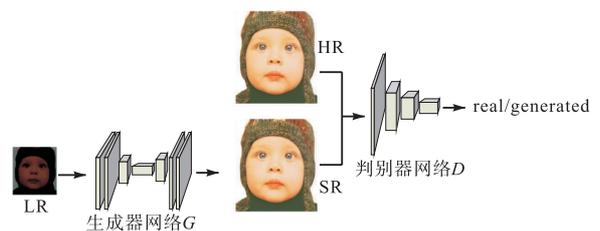


图 1 网络结构图

Fig. 1 Structure diagram of network

LSRGAN 模型在对抗过程中的目标函数为

$$\min_{G_{\theta}} \max_{D_{\phi}} V(D_{\phi}, G_{\theta}) = E_{I^{HR} \sim P_{data}(I^{HR})} [\log \log D_{\phi}(I^{HR})] - E_{I^{LR} \sim P_{data}(I^{LR})} [\log \log (1 - D_{\phi}(G_{\theta}(I^{LR})))]$$
 (1)

其中: G 为生成器网络, θ 为生成器网络的网络参数; D 为判别器网络, ϕ 为判别器网络的网络参数; E 为期望, I^{HR} 为真实高分辨率图像, I^{LR} 为低分辨率图像; $P_{data}(I^{HR})$ 为真实高分辨率图像的分布, $P_{data}(I^{LR})$ 为低分辨率图像的分布; $G(I^{LR})$ 为生成器网络生成的超分辨率图像, $D(I^{HR})$ 为当输入图像为真实高分辨率图像时判别真伪, $D(G(I^{LR}))$ 为输入图像为超分辨率图像时判别真伪.

LSRGAN 模型生成器网络的全部损失函数为

$$L = \alpha l_{light}^{SR} + \beta l_{VGG}^{SR} + \gamma l_{SSIM}^{SR} + \delta l_D^{SR}$$
 (2)

其中: α 、 β 、 γ 、 δ 为模型的超参数, $I_{\text{light}}^{\text{SR}}$ 为光照损失, $I_{\text{VGG}}^{\text{SR}}$ 为内容损失, $I_{\text{SSIM}}^{\text{SR}}$ 为结构相似性损失, I_{D}^{SR} 为对抗损失。

光照损失 $I_{\text{light}}^{\text{SR}}$ 是通过把 RGB 颜色空间转换到 YIQ 颜色空间, 然后计算图像中的亮度分量, 将图像亮度作为损失函数。相较于其他颜色空间, YIQ 颜色空间具有能将图像中的亮度分量分离提取的优点, 并且 YIQ 颜色空间与 RGB 颜色空间之间是线性变换的关系, 计算量小, 聚类特性也较好, 可以适应光照强度不断变化的场合, 所以 YIQ 颜色空间能够有效地用于彩色图像处理。光照损失可以更好地恢复低光照条件下的低分辨率图像。

光照损失 $I_{\text{light}}^{\text{SR}}$ 的表达式为

$$I_{\text{light}}^{\text{SR}} = (Y(I^{\text{SR}}) - Y(I^{\text{HR}}))^2 \quad (3)$$

其中: I^{SR} 为超分辨率图像; I^{HR} 为高分辨率图像; Y 为图像亮度, $Y = 0.299R + 0.587G + 0.144B$, R 、 G 、 B 分别为图像的 RGB 三通道值。

内容损失 $I_{\text{VGG}}^{\text{SR}}$ 区别于传统的基于像素的内容损失, 在给定的 VGG19 网络中, 选取第 9 层池化层之前的第 8 层卷积层(激活层之后)的特征映射。使用重建图像 $G_{\theta}(I^{\text{LR}})$ 的特征表示和真实高分辨率图像 I^{HR} 之间的欧氏距离定义内容损失, 其中 F 表示 VGG19 神经网络。

$$I_{\text{VGG}}^{\text{SR}} = \|F(I^{\text{HR}}) - F(G_{\theta}(I^{\text{LR}}))\|^2 \quad (4)$$

相较于基于像素的内容损失, 基于式(4)内容损失生成的图像具有更好的视觉效果。

结构相似性损失 $I_{\text{SSIM}}^{\text{SR}}$ 通过计算超分辨率图像与真实高分辨率图像之间的结构相似性, 提高生成图像的质量。结构相似性的表达式为

$$I_{\text{SSIM}}^{\text{SR}} = \text{SSIM}(I^{\text{SR}}, I^{\text{HR}}) = \frac{(2\mu_{I^{\text{SR}}}\mu_{I^{\text{HR}}} + C_1)(2\sigma_{I^{\text{SR}}I^{\text{HR}}} + C_2)}{(\mu_{I^{\text{SR}}}^2 + \mu_{I^{\text{HR}}}^2 + C_1)(\sigma_{I^{\text{SR}}}^2 + \sigma_{I^{\text{HR}}}^2 + C_2)} \quad (5)$$

其中: I^{SR} 为生成的超分辨率图像, I^{HR} 为真实的高分辨率图像; $\mu_{I^{\text{SR}}}$ 为 I^{SR} 的平均值, $\mu_{I^{\text{HR}}}$ 为 I^{HR} 的平均值; $\sigma_{I^{\text{SR}}}^2$ 为 I^{SR} 的方差, $\sigma_{I^{\text{HR}}}^2$ 为 I^{HR} 的方差, $\sigma_{I^{\text{SR}}I^{\text{HR}}}$ 为 I^{SR} 和 I^{HR} 的协方差; $C_1 = (k_1l)^2$, $C_2 = (k_2l)^2$, C_1 和 C_2 为维持稳定的常数, l 为像素值的动态范围, $k_1 = 0.01$, $k_2 = 0.03$ 。结构相似性的范围为 0 ~ 1, 当两张图像一模一样时, SSIM 的值等于 1。

对抗损失 I_{D}^{SR} 是基于判别器网络, 判别器网络用于区分真实的高分辨率图像和生成的高分辨率图像, 对抗损失可以有效提高生成图像的视觉效果, 其表达

式为

$$I_{\text{D}}^{\text{SR}} = -\log D(G(I^{\text{LR}})) \quad (6)$$

其中: G 为生成器网络, D 为判别器网络, $G(I^{\text{LR}})$ 为生成器网络生成的超分辨率图像, $D(G(I^{\text{LR}}))$ 为输入图像为超分辨率图像时判别真伪。

LSRGAN 模型在训练过程中, 根据式(1)中的目标函数, 对生成器网络和判别器网络进行交替优化, 解决对抗最小-最大问题, 得到最优的生成器网络的网络参数 θ 和判别器网络的网络参数 φ 。

2.2 生成器网络结构

生成器网络包括特征提取单元、深度残差单元以及上采样单元 3 个主要部分。生成器网络的结构示意图如图 2 所示, 其中特征提取单元包括 64 个卷积核大小为 3×3 的卷积层和 ReLU 层, 低分辨率图像先经过该卷积层进行卷积特征提取, 然后通过 ReLU 函数进行激活。深度残差单元包括 12 个残差块, 用于进一步对特征提取单元提取到的特征图进行特征学习并提取高频信息。每个残差块内部依次为卷积层、批标准化 BN 层、ReLU 层、卷积层、批标准化 BN 层和残差边。卷积层进行特征图的特征提取, 批标准化 BN 层对特征图的特征进行标准化, 以减少模型的计算量。用 ReLU 函数进行激活, 给学习到的特征图增加非线性元素。残差边将输入和输出进行叠加操作, 避免因网络层数太多而导致梯度消失的问题。上采样单元由 1 个反卷积层和 ReLU 层组成, 用于根据提取的高频信息将低分辨率图像进行指定倍数的放大 [本文是将低分辨率(LR)图像的长和宽均放大 3 倍], 从而得到超分辨率图像。

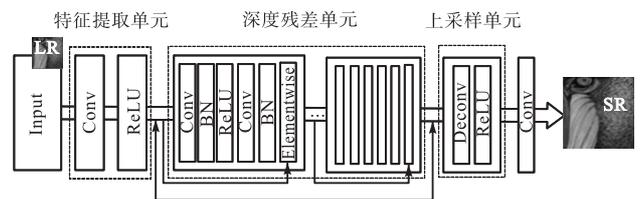


图 2 生成器网络的结构示意图

Fig. 2 Structure diagram of generation network

2.3 判别器网络结构

判别器网络用于区分超分辨率图像和真实的高分辨率图像, 促进生成器网络生成效果更好的超分辨率图像。判别器网络的结构示意图如图 3 所示。判别器网络由卷积块和全连接层组成。其中卷积块由卷积核大小为 3×3 的卷积层、LeakyReLU 层以及批标准化 BN 层组成, 用于不断地提取输入图像的特征。

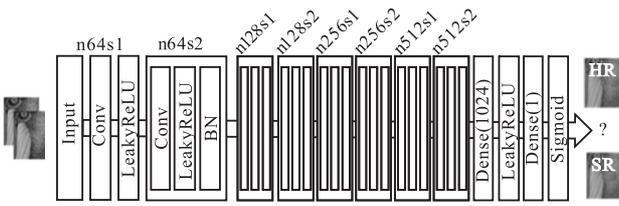


图 3 判别器网络的结构示意图

Fig. 3 Structure diagram of discrimination network

输入至判别器网络的图像先经过卷积核大小为 3×3 、通道数为 64 的卷积层进行特征提取, 然后经过 LeakyReLU 函数进行激活. 通过 7 个组成相同的卷积块进一步提取图像特征, 随着特征图大小的不断减少, 通道数不断增多, 分别为 64、128、128、256、256、512、512, 将神经网络学习到的特征存储到通道中. 最后经过全连接层进行分类, Sigmoid 层计算出输入图像是真实高分辨率图像的概率.

3 实验

3.1 实验设置

实验在 NVIDIA GeForce RTX 3090 GPU 上训练所有网络, 所有实验均在低分辨率和高分辨率图像之间以 3 倍的尺度因子执行. 为了公平比较, 所有实验的 PSNR 和 SSIM 度量均使用 TensorFlow 封装好的方法计算. 优化器为 Adam, 学习率为 0.0002, 数据的批量处理大小为 16, 更新迭代 13000 次.

训练数据来自数据集 BSDS200^[31]、General-100^[9]、DIV2K^[32]和 Flickr2K^[33]中随机采样的高分辨率图像, 训练数据集详细信息见表 1. BSDS200 数据集是 BSDS500 数据集的训练子集, 由伯克利大学提供, 常用于图像分割、轮廓检测和图像超分辨率重建. General-100 数据集有 100 张图像, 里面包含人物、动物、日常景象等. DIV2K 数据集有 1000 张高分辨率图像, 其中 800 张用于训练, 100 张用于验证, 100 张用于测试. Flickr2K 数据集有 2650 张高分辨率图像, 里面包含人物、动物、日常景象等.

测试数据来自 Set5^[34]、Set14^[35]、BSDS100^[36]和 Urban100^[19]数据集, 测试数据集详细信息见表 2. Set5 数据集有 5 张图像, 里面含有人物、动物、昆虫等. Set14 数据集有 14 张图像, 里面含有人物、动物、自然景象等. BSDS100 数据集有 100 张图像, 是 BSDS500 数据集的测试子集. Urban100 数据集有 100 张图像, 里面包含各种建筑物.

表 1 训练数据集详细信息

Tab. 1 Details of training dataset

数据集名称	图像数量/张	图像格式	图像描述	平均分辨率
BSDS200	200	JPG	BSDS500 的训练子集	(432, 370)
General-100	100	BMP	人物、动物、日常景象等	(435, 381)
DIV2K	800	PNG	人物、动物、日常景象等	(2 040, 1 848)
Flickr2K	2 650	PNG	人物、动物、日常景象等	(2 040, 1 356)

表 2 测试数据集详细信息

Tab. 2 Details of test dataset

数据集名称	图像数量/张	图像格式	图像描述	平均分辨率
Set5	5	PNG	人物、动物、昆虫等	(313, 336)
Set14	14	PNG	人物、动物、自然景象	(492, 446)
BSDS100	100	JPG	BSDS500 的测试子集	(432, 370)
Urban100	100	PNG	建筑物	(984, 797)

3.2 构建新的训练数据集

对随机采样得到的高分辨率图像进行随机裁剪, 裁剪成 $224 \text{ 像素} \times 224 \text{ 像素}$ 大小的高分辨率图像; 挑选裁剪后的图像, 剔除不符合要求的图像; 最终得到 3428 张高分辨率图像. 对裁剪、挑选后得到的 3428 张高分辨率图像先得到其灰度图, 然后通过减少灰度图的灰度值减少图像的亮度, 接着使用双三次插值法进行下采样, 得到低光照低分辨率图像. 这一步可以用公式表示为

$$I^{\text{LR}} = f(I^{\text{HR}}) \downarrow_s \quad (7)$$

式中: I^{LR} 为所求的低光照低分辨率图像; I^{HR} 为高分辨率图像; \downarrow_s 表示下采样; s 为下采样的尺度因子; $f(\cdot)$ 为用于将高分辨率图像进行变黑处理的函数, 其中 $f(I^{\text{HR}}) = \alpha \cdot I^{\text{HR}} + \gamma \cdot \text{ones}(\text{size}(I^{\text{HR}}))$, α 为超参数, γ 为黑度参数.

3.3 精度对比

对比 Bilinear、SRCNN、SRFBN^[19]、SRGAN 和 LSRGAN 这 5 个模型在黑度参数 γ 分别为 -70、-130 和 -160 条件下的精度, 即对比 PSNR 和 SSIM.

首先在构造的训练数据集(在数据集 BSDS200、General-100、DIV2K 和 Flickr2K 中随机采样)上进行训练, 然后在 4 个测试数据集 BSDS100、Set14、Set5 和 Urban100 上进行测试. 测试时, 先对测试数据进行 3.2 节中的变黑处理和下采样处理, 然后对处理后的测试数据进行测试, 实验结果见表 3 和图 4.

表 3 定量结果实验对比

Tab. 3 Quantitative results of experimental comparison

测试集	模型	$\gamma = -70$		$\gamma = -130$		$\gamma = -160$	
		PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Set5	Bilinear	12.675 70	0.409 46	8.736 59	0.182 87	7.677 83	0.108 37
	SRCNN	22.648 90	0.692 12	17.675 21	0.559 31	15.918 97	0.501 62
	SRFBN	18.587 15	0.556 98	18.492 12	0.562 82	10.395 60	0.454 84
	SRGAN	20.850 35	0.674 45	18.014 77	0.567 99	12.990 87	0.490 99
	LSRGAN	22.850 21	0.693 56	17.873 21	0.580 33	16.637 20	0.523 58
Set14	Bilinear	11.909 16	0.407 01	7.960 13	0.171 85	6.922 53	0.096 70
	SRCNN	22.688 40	0.672 14	18.161 91	0.548 57	16.134 50	0.482 09
	SRFBN	23.060 97	0.654 21	18.645 37	0.542 67	16.396 49	0.482 26
	SRGAN	21.329 22	0.635 57	18.368 35	0.539 76	14.075 90	0.469 46
	LSRGAN	21.912 76	0.640 39	18.065 05	0.549 26	16.428 42	0.486 55
BSDS100	Bilinear	12.022 78	0.332 57	8.546 69	0.110 30	7.758 06	0.058 04
	SRCNN	23.536 05	0.661 75	18.937 10	0.525 69	17.148 31	0.455 15
	SRFBN	23.722 83	0.638 77	19.290 88	0.514 18	16.996 31	0.455 74
	SRGAN	22.024 23	0.619 69	19.365 42	0.504 67	14.163 98	0.441 99
	LSRGAN	22.778 41	0.627 33	18.769 85	0.520 79	17.165 47	0.463 39
Urban100	Bilinear	11.871 37	0.404 84	8.076 82	0.180 32	7.037 93	0.107 54
	SRCNN	20.734 60	0.630 20	17.052 77	0.508 40	15.462 73	0.456 72
	SRFBN	20.808 67	0.626 33	17.353 42	0.504 15	16.089 34	0.489 82
	SRGAN	19.904 13	0.638 29	17.692 79	0.538 81	13.850 05	0.468 80
	LSRGAN	20.815 60	0.649 85	17.462 72	0.549 06	16.277 33	0.496 87

注：其中加粗的数字表示最优指标值

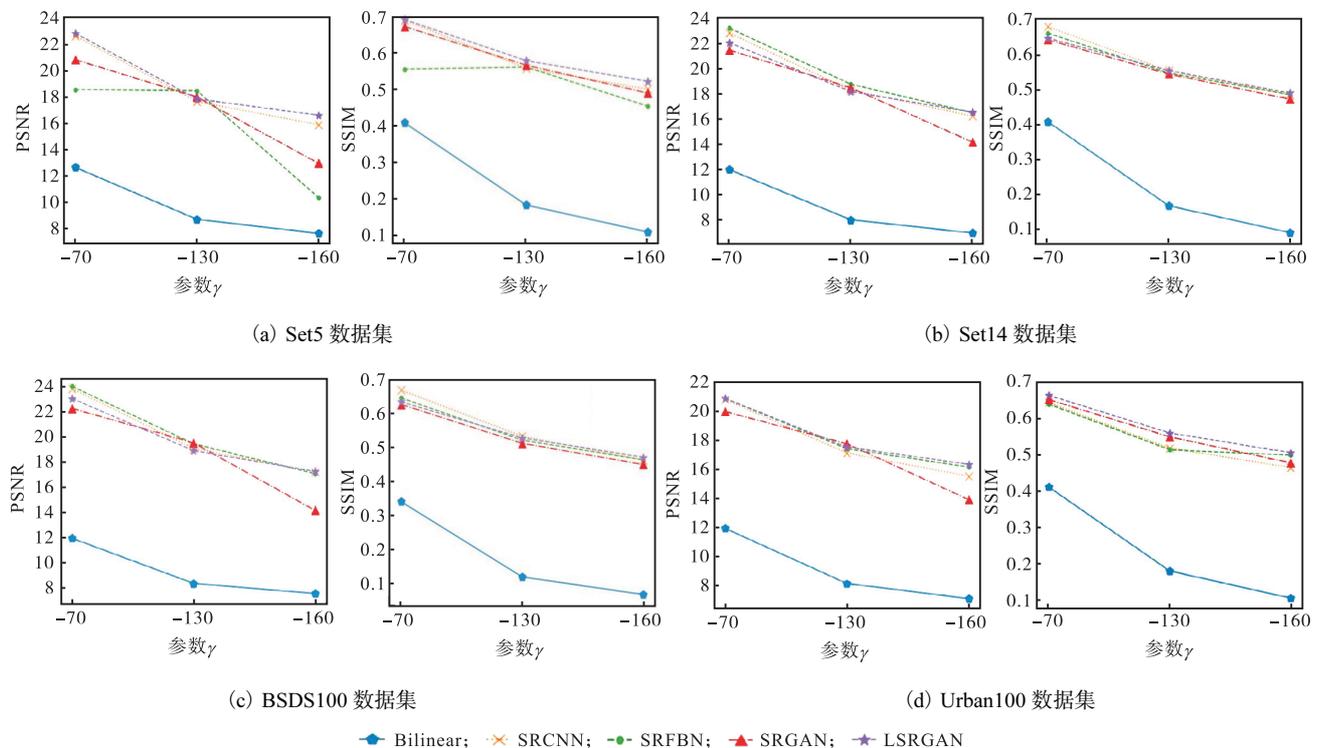


图 4 定性结果实验对比

Fig. 4 Qualitative results of experimental comparison

由表 3 可知：黑度参数 γ 的值越小，图像的亮度就越低，图像超分辨率重建的难度越大。当黑度参数 $\gamma = -70$ 时，测试图像的亮度改变不大，亮度减小的程

度对测试图像的影响较小，此时 LSRGAN 模型在 4 个测试数据集上的 8 个指标中取得了 4 个最好的结果，SRFBN 模型取得了 2 个最好的结果，SRCNN 模

型取得了 2 个最好的结果. 当黑度参数 $\gamma=-130$ 时, 测试图像的亮度改变较大, 亮度减小的程度对测试图像的影响较大, 此时 LSRGAN 模型在 4 个测试数据集上的 8 个指标中取得了 3 个最好的结果, SRFBN 模型取得了 2 个最好的结果, SRGAN 模型取得了 2 个最好的结果, SRCNN 模型取得了 1 个最好的结果. 当黑度参数 $\gamma=-160$ 时, 测试图像的亮度改变很大, 亮度减小的程度对测试图像的影响很大, 此时仅凭肉眼很难分辨出测试图像的内容. 在这种情况下, LSRGAN 模型展现出了绝对的优势, 在 4 个测试集上的 8 个指标全部取得了最好的定量结果. 综上所述, LSRGAN 模型在整体上表现最好, 且随着图像亮度的降低, LSRGAN 模型的生成效果越来越好; 当图像的亮度很低且影响到图像内容的观测时, 光照损失发挥了很重要的作用, 此时 LSRGAN 模型具有很好的生成效果.

由图 4 可知: 随着黑度参数 γ 的值逐渐减小, 图



图 5 公共测试集 Set14 的模型生成图像效果展示

Fig. 5 Model generation image effect display of public test Set14

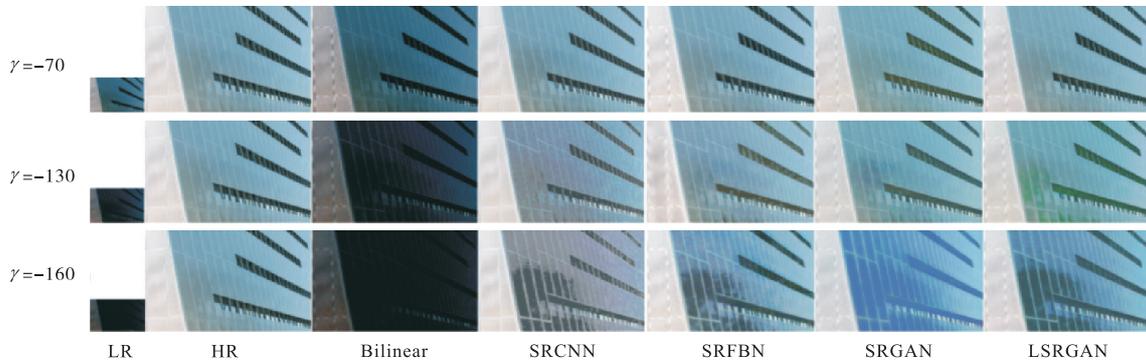


图 6 公共测试集 Urban100 的模型生成图像效果展示

Fig. 6 Model generation image effect display of public test Urban100

根据图 5 和图 6 可知: 随着黑度参数 γ 的值越来越小, 低分辨率图像的亮度越来越低; 当 $\gamma=-160$ 时,

像的亮度逐渐降低, 图像超分辨率重建的难度逐渐增大, 模型的生成能力逐渐减弱. 此时, LSRGAN 模型的折线斜率变化最小, 这表明 LSRGAN 模型的生成能力受图像亮度降低的影响最小. 因此可以得出结论, 当图像逐渐变暗时, LSRGAN 模型受图像亮度降低的影响最小, LSRGAN 模型依然具有很好的模型生成能力, 可以生成视觉效果较好的图像.

3.4 生成图像对比

3.4.1 公共测试集

为了进一步展示 5 种模型的生成效果, 在公共测试集 Set14 中随机选择 1 张图片, 在黑度参数 γ 分别为 -70、-130 和 -160 时进行生成测试, 生成结果如图 5 所示.

在公共测试集 Urban100 中随机选择 1 张图片, 在黑度参数 γ 分别为 -70、-130 和 -160 时进行生成测试, 生成结果如图 6 所示.

图像几乎全黑,此时图像中包含的有用信息很少,将其恢复成高分辨率图像是十分困难的。

从生成图像的效果可以看出,LSRGAN 模型可以在低光照条件下的图像超分辨率重建,并且生成的图像具有较好的视觉效果。

3.4.2 真实图像

在模型的训练和测试时,训练数据和测试数据的

低光照条件都是人为构造的,如此得到的低光照图像是线性的,整个图像的亮度一起降低。为了进一步验证模型的生成能力,收集了现实生活中的低光照图像进行测试。将现实生活中的低光照图像进行3倍下采样处理,这里不需要进行变黑处理;然后将处理后的图像输入模型进行测试,生成结果如图7所示。



图7 现实生活中的低光照图像的模型生成图像效果展示

Fig. 7 Model generation image effect display of low light image in real life

由图7可知:现实生活中的低光照图像情况比较复杂,可能存在局部亮度低的问题,并且导致图像质量不高的原因是多方面的。在这种情况下,LSRGAN 模型仍然具有较好的图像超分辨率和恢复光照的能力,可以处理现实生活中的低光照图像。

4 结论

本文提出了一种融合光照损失的图像超分辨率生成对抗网络 LSRGAN 模型,并在实验中展示了模型定性指标 PSNR 和 SSIM 的值及模型生成图像的效果。与经典 SRGAN 模型的对比实验显示,该模型不仅能够实现一般情况下的图像超分辨率重建,而且可以在低光照条件下的图像超分辨率重建,并取得了较好的实验效果。理论分析及实验结果表明,LSRGAN 模型中的光照损失和结构相似性损失可以帮助恢复图像的光照强度,具有较好的恢复效果,同时使生成的图像具有较好的视觉效果。

参考文献:

- [1] COVER T, HART P. Nearest neighbor pattern classification[J]. IEEE Transactions on information theory, 1967, 13(1): 21-27.
- [2] KIRKLAND E J. Bilinear interpolation[EB/OL]. [2021-11-10]. https://doi.org/10.1007/978-1-4419-6533-2_12.
- [3] KEYS R G. Cubic convolution interpolation for digital image processing[J]. IEEE Transactions on acoustics, speech, and signal processing, 2003, 29(6): 1153-1160.
- [4] KIM J, LEE J K, LEE K M. Accurate image super-resolution using very deep convolutional networks[C]//IEEE. 2016 IEEE Conference on Computer Vision and Pattern Recognition(CVPR). New York: IEEE, 2016: 1646-1654.
- [5] LEDIG C, THEIS L, HUSZÁR F, et al. Photo-realistic single image super-resolution using a generative adversarial network[C]//IEEE. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). New York: IEEE, 2017: 4681-4690.
- [6] IGNATOV A, KOBYSHEV N, TIMOFTE R, et al. WESPE: weakly supervised photo enhancer for digital cameras[C]//IEEE. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). New York: IEEE, 2018: 691-700.
- [7] DUCHON C E. Lanczos filtering in one and two dimensions[J]. Journal of applied meteorology, 1979, 18(8): 1016-1022.
- [8] 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.
- [9] DONG C, LOY C C, TANG X. Accelerating the super-resolution convolutional neural network[EB/OL]. [2021-11-10]. <https://arxiv.org/pdf/1608.00367.pdf>.
- [10] MAO X J, SHEN C, YANG Y B. Image restoration using convolutional auto-encoders with symmetric skip connections[EB/OL]. [2021-11-10]. <https://arxiv.org/abs/1606.08921v3>.
- [11] TONG T, LI G, LIU X, et al. Image super-resolution using dense skip connections[C]//IEEE. 2017 IEEE International Conference on Computer Vision (ICCV). New York: IEEE, 2017: 4799-4807.

- [12] YU J, FAN Y, YANG J, et al. Wide activation for efficient and accurate image super-resolution[EB/OL]. [2021-11-10]. <https://arxiv.org/pdf/1808.08718.pdf>.
- [13] KIM J, LEE J K, LEE K M. Deeply-recursive convolutional network for image super-resolution[C]//IEEE. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). New York: IEEE, 2016: 1637-1645.
- [14] TAI Y, YANG J, LIU X. Image super-resolution via deep recursive residual network[C]//IEEE. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). New York: IEEE, 2017: 3147-3155.
- [15] 周登文, 赵丽娟, 段然, 等. 基于递归残差网络的图像超分辨率重建[J]. 自动化学报, 2019, 45(6): 1157-1165.
- [16] HAN W, CHANG S, LIU D, et al. Image super-resolution via dual-state recurrent networks[C]//IEEE. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New York: IEEE, 2018: 1654-1663.
- [17] WANG X, YU K, WU S, et al. ESRGAN: enhanced super-resolution generative adversarial networks[EB/OL]. [2021-11-10]. <https://arxiv.org/pdf/1809.00219.pdf>.
- [18] HARIS M, SHAKHAROVICH G, UKITA N. Deep back-projection networks for super-resolution[C]//IEEE. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New York: IEEE, 2018: 1664-1673.
- [19] LI Z, YANG J, LIU Z, et al. Feedback network for image super-resolution[C]//IEEE. 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). New York: IEEE, 2019: 3867-3876.
- [20] ZHANG Y, LI K, LI K, et al. Image super-resolution using very deep residual channel attention networks[EB/OL]. [2021-11-10]. <https://arxiv.org/abs/1807.02758v2>.
- [21] DAI T, CAI J, ZHANG Y, et al. Second-order attention network for single image super-resolution[C]//IEEE. 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). New York: IEEE, 2019: 11065-11074.
- [22] 周登文, 马路遥, 田金月, 等. 基于特征融合注意网络的图像超分辨率重建[J/OL]. 自动化学报: 1-9 [2021-11-15]. <https://doi.org/10.16383/j.aas.c190428>.
- [23] LAI W S, HUANG J B, AHUJAN, 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.
- [24] BULAT A, YANG J, TZIMIROPOULOS G. To learn image super-resolution, use a GAN to learn how to do image degradation first[EB/OL]. [2021-11-10]. <https://arxiv.org/pdf/1807.11458.pdf>.
- [25] YUAN Y, LIU S, ZHANG J, et al. Unsupervised image super-resolution using cycle-in-cycle generative adversarial networks[C]//IEEE. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). New York: IEEE, 2018: 701-710.
- [26] SHOCHER A, COHEN N, IRANI M. "Zero-shot" super-resolution using deep internal learning[C]//IEEE. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New York: IEEE, 2018: 3118-3126.
- [27] LEMPITSKY V, VEDALDI A, ULYANOV D. Deep image prior[C]//IEEE. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New York: IEEE, 2018: 9446-9454.
- [28] ZHANG K, ZUO W, ZHANG L. Learning a single convolutional super-resolution network for multiple degradations[C]//IEEE. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New York: IEEE, 2018: 3262-3271.
- [29] GU J, LU H, ZUO W, et al. Blind super-resolution with iterative kernel correction[C]//IEEE. 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). New York: IEEE, 2019: 1604-1613.
- [30] ZHANG K, ZUO W, ZHANG L. Deep plug-and-play super-resolution for arbitrary blur kernels[C]//IEEE. 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). New York: IEEE, 2019: 1671-1681.
- [31] ARBELAEZ P, MAIRE M, FOWLKES C, et al. Contour detection and hierarchical image segmentation[J]. IEEE Transactions on pattern analysis and machine intelligence, 2010, 33(5): 898-916.
- [32] AGUSTSSON E, TIMOFTE R. NTIRE 2017 challenge on single image super-resolution: dataset and study[C]//IEEE. 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). New York: IEEE, 2017: 126-135.
- [33] TIMOFTE R, AGUSTSSON E, VAN GOOL L, et al.

- 计算机工程与应用, 2019, 55(3): 1-9.
- [12] 陈佛计, 朱枫, 吴清潇, 等. 生成对抗网络及其在图像生成中的应用研究综述[J]. 计算机学报, 2021, 44(2): 347-369.
- [13] 鲍建敏. 基于生成对抗网络的图像合成[D]. 合肥: 中国科学技术大学, 2019.
- [14] QIN Z, LIU Z, ZHU P, et al. A GAN-based image synthesis method for skin lesion classification[J]. Computer methods and programs in biomedicine, 2020, 195: 105568.
- [15] LIU S, GUO H, HU J, et al. A novel data augmentation scheme for pedestrian detection with attribute preserving GAN[J]. Neurocomputing, 2020, 401: 123-132.
- [16] 王金娜, 苏杰, 杨凯, 等. 基于 Cycle-GAN 的绝缘子图像生成方法[J]. 广东电力, 2020, 33(1): 100-108.
- [17] 王雷雷. 基于 GAN 的 SAR 图像生成研究[D]. 成都: 电子科技大学, 2019.
- [18] 王光陞. 基于卷积神经网络的鸟类细粒度识别方法研究[D]. 北京: 华北电力大学, 2018.
- [19] 翁雨辰. 基于深度学习模型的细粒度鸟类识别[D]. 北京: 北京林业大学, 2018.
- [20] 李春燕, 王思军, 张庆东. 天津大黄堡湿地鸟类资源现状及保护对策[J]. 天津农林科技, 2013(5): 29-30.
- [21] 梁俊杰, 韦舰晶, 蒋正锋. 生成对抗网络 GAN 综述[J]. 计算机科学与探索, 2020, 14(1): 1-17.
- [22] 陈文兵, 管正雄, 陈允杰. 基于条件生成式对抗网络的数据增强方法[J]. 计算机应用, 2018, 38(11): 3305-3311.

责任编辑: 郎婧

(上接第 63 页)

- NTIRE 2017 challenge on single image super-resolution: methods and results[C]//IEEE. 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). New York: IEEE, 2017: 114-125.
- [34] BEVILACQUA M, ROUMY A, GUILLEMOT C, et al. Low-complexity single-image super-resolution based on nonnegative neighbor embedding[EB/OL]. [2021-11-10]. <http://www.bmva.org/bmvc/2012/BMVC/paper135/abstract135.pdf>.
- [35] ZEYDE R, ELAD M, PROTTER M. On single image scale-up using sparse-representations[EB/OL]. [2021-11-10]. <http://citeseerx.ist.psu.edu/viewdoc/download?jsessionid=566DF490D3BE2902B5B8C6E700A61776?doi=10.1.1.310.3418&rep=rep1&type=pdf>.
- [36] HUANG J B, SINGH A, AHUJA N. Single image super-resolution from transformed self-exemplars[C]// IEEE. 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). New York: IEEE, 2015: 5197-5206.

责任编辑: 郎婧