

基于球形矩匹配与特征判别的图像超分辨率重建

林 静, 黄玉清*, 李磊民

(西南科技大学信息工程学院, 四川 绵阳 621010)

(* 通信作者电子邮箱hyq_851@163.com)

摘要: 由于网络训练不稳定, 基于生成对抗网络(GAN)的图像超分辨率重建存在模式崩溃的现象。针对此问题, 提出了一种基于球形几何矩匹配与特征判别的球面双判别器超分辨率重建网络SDSRGAN, 通过引入几何矩匹配与高频特征判别来改善网络训练的稳定性。首先, 生成器对图像提取特征并通过上采样生成重建图像; 接着, 球面判别器将图像特征映射至高维球面空间, 充分利用特征数据的高阶统计信息; 然后, 在传统判别器的基础上增加特征判别器, 提取图像高频特征, 重建特征高频分量和结构分量两方面; 最后, 对生成器与双判别器进行博弈训练, 提高生成器重建图像质量。实验结果表明, 所提算法能有效收敛, 其网络能够稳定训练, 峰值信噪比(PSNR)为31.28 dB, 结构相似性(SSIM)为0.872, 而与双三次差值、超分辨率残差网络(SRResNet)、加速的卷积神经网络超分辨率(FSRCNN)、基于GAN的单图像超分辨率(SRGAN)和增强型超分辨率生成对抗网络(ESRGAN)算法相比, 所提算法的重建图像具有更加逼真的结构纹理细节。所提算法为基于GAN的图像超分辨率研究提供了球形矩匹配与特征判别的判别方法, 在实际应用中可行且有效。

关键词: 生成对抗网络; 图像超分辨率重建; 高频特征; 双判别器; 模式崩溃

中图分类号: TP391.41; TP389.1 **文献标志码:** A

Image super-resolution reconstruction based on spherical moment matching and feature discrimination

LIN Jing, HUANG Yuqing*, LI Leimin

(School of Information Engineering, Southwest University of Science and Technology, Mianyang Sichuan 621010, China)

Abstract: Due to the instability of network training, the image super-resolution reconstruction based on Generative Adversarial Network (GAN) has a mode collapse phenomenon. To solve this problem, a Spherical double Discriminator Super-Resolution Generative Adversarial Network (SDSRGAN) based on spherical geometric moment matching and feature discrimination was proposed, and the stability of network training was improved by adopting geometric moment matching and discrimination of high-frequency features. First of all, the generator was used to produce a reconstructed image through feature extraction and upsampling. Second, the spherical discriminator was used to map image features to high-dimensional spherical space, so as to make full use of higher-order statistics of feature data. Third, a feature discriminator was added to the traditional discriminator to extract high-frequency features of the image, so as to reconstruct both the characteristic high-frequency component and the structural component. Finally, game training between the generator and double discriminator was carried out to improve the quality of the image reconstructed by the generator. Experimental results show that the proposed algorithm can effectively converge, its network can be stably trained, and has Peak Signal-to-Noise Ratio (PSNR) of 31.28 dB, Structural SIMilarity (SSIM) of 0.872. Compared with Bicubic, Super-Resolution Residual Network (SRResNet), Fast Super-Resolution Convolutional Neural Network (FSRCNN), Super-Resolution using a Generative Adversarial Network (SRGAN), and Enhanced Super-Resolution Generative Adversarial Network (ESRGAN) algorithms, the reconstructed image of the proposed algorithm has more precise structural texture characteristics. The proposed algorithm provides a double discriminant method for spherical moment matching and feature discrimination for the research of image super-resolution based on GAN, which is feasible and effective in practical applications.

Key words: Generative Adversarial Network (GAN); image super-resolution reconstruction; high-frequency feature; double discriminator; mode collapse

0 引言

单幅图像超分辨率(Single Image Super-Resolution, SISR)

重建^[1]是使用单张模糊的低分辨率(Low-Resolution, LR)图像的信息重建出纹理细节丰富、清晰的高分辨率(High-Resolution, HR)图像, 在军事遥感、医学成像和视频监测等领

收稿日期: 2019-12-23; 修回日期: 2020-03-29; 录用日期: 2020-03-30。 基金项目: 国家自然科学基金面上项目(61673220)。

作者简介: 林静(1993—), 女, 四川广安人, 硕士研究生, 主要研究方向: 图像处理、机器视觉; 黄玉清(1962—), 女, 四川绵阳人, 教授, 硕士, 主要研究方向: 图像处理、机器视觉、智能技术; 李磊民(1960—), 男, 四川绵阳人, 教授, 硕士, 主要研究方向: 机器人控制、无线测控、图像处理、机器视觉。

域应用广泛。超分辨率(Super-Resolution, SR)重建技术主要分为基于插值^[2]、重建^[3]以及学习^[4]的三种方法。基于插值的方法需要图像具有连续灰度值,根据邻域像素点的灰度值估计得到待插点的灰度值。该方法操作简单、时间复杂度低,但在应用中获取的HR图像信息有限,重建图像易产生锯齿或模糊现象。基于重建的方法通过配准等处理将同一场景下采集的多幅图像进行信息融合。该方法借助图像的先验知识进行重建,缓解了基于插值重建图像的锯齿现象,重建效果得到一定改善,但易受设备、场景等影响,当放大倍数过大时由于获取的先验知识较少,重建效果较差。基于学习的方法直接通过学习获得LR图像与HR图像之间的映射关系,通过该关系预测对应HR图像的纹理细节,相较于重建的方法改善了放大倍数的限制,能获得更好的HR图像质量。

随着卷积神经网络(Convolutional Neural Network, CNN)在SISR重建任务中应用,重建图像质量得到了有效改善。Dong等^[5]提出了快速的卷积神经网络超分辨率(Fast Super-Resolution Convolutional Neural Network, FSRCNN)重建,使用反卷积层代替传统双三次插值进行图像上采样,并将卷积层数量增加至8层获取图像信息。该方法无需差值预处理,提高了处理速度。Shi等^[6]提出了亚像素卷积神经网络超分辨率(Efficient Sub-Pixel Convolutional Neural network for image super-resolution, ESPCN)重建算法,首先对LR图像块进行特征提取,再使用亚像素卷积对特征进行上采样,降低了重建过程的计算复杂度,提高了重建的效率。基于CNN的重建算法^[7-9]能获得较高感知评价指标,但是图像总体平滑,缺乏纹理细节。Goodfellow等^[10]首次提出包含生成器和判别器两部分的生成对抗网络(Generative Adversarial Network, GAN)框架,训练时生成器与判别器进行极大极小博弈。Ledig等^[11]将GAN应用SISR重建任务,提出基于GAN的单图像超分辨率网络(photo-realistic single image Super-Resolution using Generative Adversarial Network, SRGAN),把图像重建视为图像生成任务。该方法将LR图像输入生成器生成SR图像,判别器对SR图像和HR图像进行区分,相较于传统重建方法而言能够保留更多细节特征,但是重建图像具有大量伪像。在此基础上,Wang等^[12]改进了SRGAN的网络结构、对抗损失和感知损失三个部分,提出增强型超分辨率生成对抗网络(Enhanced Super-Resolution Generative Adversarial Network, ESRGAN),生成的重建图像不仅具有丰富的纹理细节,而且峰值信噪比(Peak Signal to Noise Ratio, PSNR)和结构相似性(Structural SIMilarity, SSIM)等评价指标较SRGAN明显提高。

传统GAN采用JS(Jensen-Shannon)散度^[13]度量SR图像分布与HR图像分布之间的距离,在判别器训练最优时,优化生成器即最小化SR图像与HR图像分布的JS散度。然而只有当SR图像与HR图像分布完全重叠时JS散度才为0,其余情况均为固定常数 $\log 2$,此时生成器面临梯度消失^[14-15]问题,因此传统GAN存在生成器损失难收敛、网络训练不稳定的问题。并且基于GAN的SISR技术引入了对抗损失,将SR任务视为图像生成,重建图像评价指标值较低。针对此问题,本文提出SDSRGAN(Spherical double Discriminator Super-Resolution Generative Adversarial Network)对球面上HR图像与SR图像特征^[16]进行几何矩匹配,并利用数据的高阶统计信息,通过特征距离间的几何约束实现网络稳定训练,解决模式崩溃问题;同时增加特征判别器^[17],充分利用图像高频特征,生成更多图像相关的结构纹理细节。

1 SDSRGAN 系统架构

SDSRGAN由一个生成器和两个判别器组成,判别器包含球面判别器与特征判别器。首先LR图像经生成器进行特征提取以及图像上采样生成SR图像;然后将SR图像与HR图像分别输入球面判别器与特征判别器,判别图像来源。球面判别器将SR图像与HR图像的特征从一维空间映射至 n 维欧几里得空间,经过几何变换,最终映射到 n 维球面空间,通过计算特征点与以球面北极为中心的几何矩,实现生成器与球面判别器极大极小博弈;特征判别器使用训练好的视觉几何组(Visual Geometry Group, VGG)网络提取SR图像与HR图像的高频特征,并判别高频特征来源。

1.1 生成网络

生成网络包含特征提取模块与上采样模块,主要任务是将LR图像重建至高清的、具有细节的SR图像。ESRGAN提出的生成网络结构在SISR重建任务中具有很大优势,因此本文沿用该生成器网络结构。生成网络使用残余剩余块(Residual-in-Residual Dense Block, RRDB)基本单元,结合多层残差网络与密集跳跃连接,在残差网络中仅包含卷积层与泄露的修正线性单元(Leaky Rectified Linear Unit, LReLU),而未使用批量标准化(Batch Normalization, BN)层,提高了网络的泛化能力并且容易训练。采用亚像素卷积进行上采样,每次上采样实现2倍图像放大,执行两次实现4倍比例因子放大。生成器操作如下:

$$I^{SR} = F_c(F_{up}(F_c(I^{LR}) + n * F_{RRDB}(F_c(I^{LR})))) \quad (1)$$

其中: I^{LR} 代表LR图像; I^{SR} 代表SR图像; F_c 代表卷积, F_{RRDB} 代表RRDB基本块; F_{up} 代表上采样。

1.2 球面判别网络

球面判别网络在图像像素层面上判断输入图片来源于SR图像还是HR图像。首先将SR图像与HR图像作为输入经过9个基本块特征提取,每个基本块包含卷积层、BN层以及LReLU激活层,其中卷积层采用3×3尺寸,前8个卷积核通道以每两个卷积呈2的幂次增加,由64个逐渐增加至512个,每两个相同卷积核数量的步长分别为1、2,将图像特征从低维度逐渐映射至更高维度中。使用平均池化(Avg Pooling)去除冗余信息,通过全连接层将图像特征信息映射到一维空间中,最后将一维空间中的特征映射至欧几里得空间,通过几何变换将图像特征映射至球面。网络结构如图1所示。图1中:圆点“•”代表高清图特征,三角形“▲”代表重建图像特征;平行四边形代表欧几里得空间,球面代表球面空间,虚线表示将欧几里得空间中的特征映射至球面; K 代表卷积核尺寸, N 代表卷积核数量, S 代表步长。

将HR图像与SR图像在欧几里得空间中的特征映射至球面,球面的特征点间距离明显短于欧几里得空间中两特征点间距离。基于球面度量HR图像与SR图像特征点之间的几何约束公式如下:

$$D_s(\theta(m), \theta(n)) = \arccos\left(\frac{\|m\|^2\|n\|^2 - \|m\|^2 - \|n\|^2 + 4mn + 1}{(\|m\|^2 + 1)(\|n\|^2 + 1)}\right) \quad (2)$$

其中: D_s 代表距离函数; θ 表示特征点从欧几里得空间映射至

球面; m, n 代表欧几里得空间中图像特征。

在欧几里得空间中特征点间距离可以取到无穷远,但球面空间中两特征点之间距离最大仅能取到以球心为圆心的半

圆弧长 π (半径为单位长)。由于球面上特征点间几何矩受限并且有界,当引入该距离作为判别器目标函数时,作用于球面的全局约束可以使网络稳定训练。

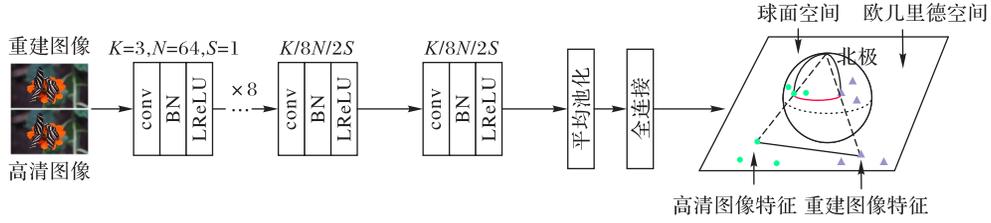


图 1 球面判别器

Fig. 1 Spherical discriminator

1.3 特征判别网络

基于GAN的重建图像引入对抗损失能在一定程度上恢复图像纹理细节,但是生成的图像具有伪像,与真实图像纹理细节有一定差异。这是因为图像高频特征中包含了很多图像纹理相关信息,但并未得到充分利用,因此,重建图像纹理细节大多为高频噪声,而非与图像相关的纹理。

特征判别网络在图像高频特征层面上判断输入图像来源。首先将HR图像和SR图像经训练好的VGG网络提取中间高频特征,并基于提取的特征判别特征图来源,使SR图像高频特征接近于HR图像高频特征。特征判别器结构如图2所示。

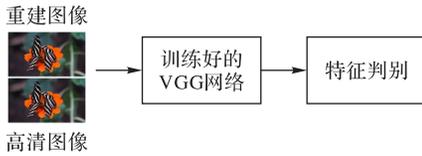


图 2 特征判别器

Fig. 2 Feature discriminator

1.4 网络损失

1.4.1 球面生成对抗损失

本文采用球面上受限的几何矩度量HR图像分布与SR图像分布之间的距离,此时生成器期望最小化球面上SR图像特征点与北极点之间的几何矩距离,判别器期望最小化球面上的HR图像特征与北极点之间的几何矩距离,同时最大化球面上的SR图像特征与北极点之间的几何矩距离,以此区分SR图像与HR图像。生成器与球面判别器之间的极大极小博弈,促进生成器与判别器在像素层面上性能不断优化,最终SR图像具有基于像素层面的HR图像高频细节,如式(3)、(4):

$$L_{Sp_g} = D_s(N, D(I^{SR})) \quad (3)$$

$$L_{Sp_d} = D_s(N, D(I^{HR})) - D_s(N, D(I^{SR})) \quad (4)$$

其中: I^{HR} 代表高清图; I^{SR} 代表重建图像; D_s 代表球面上两点间几何矩; N 代表球面北极中心点; L_{Sp_g} 代表球面生成损失; L_{Sp_d} 代表球面对抗损失。

1.4.2 像素相似性损失

使用像素相似性损失限制SR图像在像素层面上接近HR图像,本文采用均方误差(Mean Squared Error, MSE)损失作为像素相似性损失,计算如下:

$$L_{MSE} = \frac{1}{WHC} \sum_i \sum_j \sum_k (I_{i,j,k}^{HR} - I_{i,j,k}^{SR})^2 \quad (5)$$

其中: W, H, C 代表图像的宽度、高度与通道数; L_{MSE} 代表像素相似性损失。

1.4.3 特征相似性损失

使用像素相似性损失重建的图像能够提高PSNR与SSIM评价指标,但重建图像过于平滑,图像细节不够逼真。将SR图像与HR图像经过预先训练好的VGG网络提取固定层特征图,计算二者图像特征图间的MSE损失,约束SR图像在特征层面上与HR图像具有相似性。

$$L_{F_{MSE}} = \frac{1}{W_m H_m C_m} \sum_i \sum_j \sum_k (\Phi_{i,j,k}^m(I^{HR}) - \Phi_{i,j,k}^m(I^{SR}))^2 \quad (6)$$

其中: Φ 表示特征映射; W_m, H_m, C_m 代表第 m 个特征映射的维数; $L_{F_{MSE}}$ 代表特征相似性损失。

1.4.4 特征结构性损失

利用图像高频特征的结构分量,将SR图像与HR图像经过预先训练好的VGG网络提取第 m 层的图像高频特征后,通过特征判别器在特征层面上判别图像来源。特征生成损失目标是将SR图像特征判别为真,特征对抗损失目标是将SR图像特征判别为假、将HR图像特征判别为真。特征结构性损失对图像感知指标的提高具有显著效果,如式(7)、(8):

$$L_{F_g} = -\lg(\Phi_{i,j,k}^m(I^{SR})) \quad (7)$$

$$L_{F_d} = -\lg(\Phi_{i,j,k}^m(I^{HR})) + \lg(\Phi_{i,j,k}^m(I^{SR})) \quad (8)$$

其中: L_{F_g} 代表特征生成损失; L_{F_d} 代表特征对抗损失。

特征生成损失与特征对抗损失之间的博弈使SR图像与HR图像在特征结构层面上具有相似性,重建图像具有图像相关的高频纹理细节。

1.5 超分辨率重建算法

本文算法SDSRGAN具体实现分为生成网络、球面判别网络以及特征判别网络三大模块。生成网络受ESRGAN启发,采用RRDB作为特征提取块可以提高网络的泛化能力;为了改善网络训练的稳定性,借鉴文献[16]中高维空间的几何矩匹配思想,通过在判别网络中引入基于球面空间的特征判别,实现网络快速收敛;受文献[17]启发,采用基于VGG网络的特征提取,使重建图像具有丰富的纹理细节;采用Adam优化器减少内存需求,提高计算效率。技术路线图如图3所示。SDSRGAN算法流程如算法1所示。

首先,生成器对输入图像基于RRDB模块的剩余残差块进行特征提取,并使用亚像素卷积得到4倍放大的重建图像。然后,经球面判别器的特征提取并映射至高维球面空间,计算基于球面空间的特征间距离;同时图像通过训练好的VGG网络提取高频特征。最后,通过Adam优化器优化球面判别与

特征判别损失,并更新判别器参数,通过双判别器的判别结果,优化生成损失并更新生成器参数。

算法 1 SDSRGAN 算法。

- 1) 初始化:Batchsize 大小 a , 当前迭代次数 z , 完全迭代次数 Z , 已完成的批次次数 p , 训练集图片总数与每批次图片数的比值 P , 球面与特征判别损失权重系数 $\alpha = \beta = 0.5$, 生成损失权重系数 $\gamma = 10^{-2}$;
- 2) for z to Z do
- 3) for p in P do
- 4) 从训练集中取出每批次样本数为 a 的样本 $\{x_1, x_2, \dots, x_a\}$ 作为输入数据集;
- 5) 将输入数据集输入生成器, 得到重建图像 $\{G(x_1), G(x_2), \dots, G(x_a)\}$;
- 6) 使用式(4)计算球面对抗损失 $L_{Sp,d}$, 使用式(8)计算特征对抗损失 $L_{F,d}$;
- 7) 使用 Adam 优化器更新球面判别网络和特征判别网络的参数 $\omega_1, \omega_2: \nabla_{\omega_1, \omega_2} [\alpha L_{Sp,d} + \beta L_{F,d}]$;
- 8) 将输入数据 $\{x_1, x_2, \dots, x_a\}$ 输入生成器, 得到重建图像, $\{G(x_1), G(x_2), \dots, G(x_a)\}$;
- 9) 根据球面判别器与特征判别器的判别结果, 依次使用式(3)、(5)、(6)、(7)计算球面生成损失 $L_{Sp,g}$ 、像素相似性损失 L_{MSE} 、特征相似性损失 $L_{F,MSE}$ 和特征生成损失 $L_{F,g}$;
- 10) 使用 Adam 优化器更新生成网络参数 φ : $\nabla_{\varphi} [L_{Sp,g} + L_{F,g} + \gamma(L_{F,MSE} + L_{MSE})]$;
- 11) end for
- 12) end for

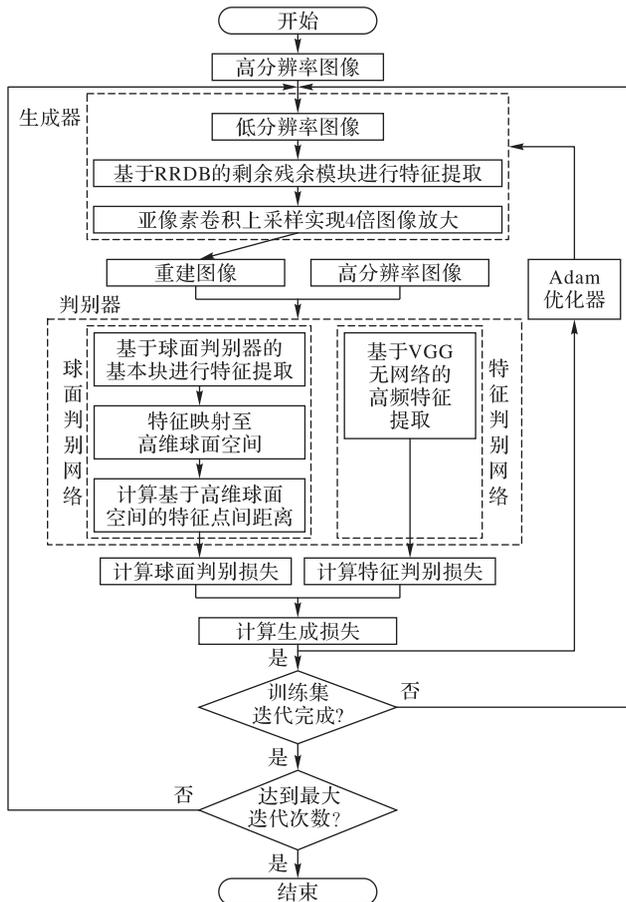


图3 SDSRGAN 算法流程

Fig. 3 Flowchart of SDSRGAN algorithm

2 实验与结果分析

2.1 实验环境

实验平台是 Linux 下 pytorch1.0 框架 Python 语言, 处理器为 Intel Xeon CPU E5-2650 v3 @ 2.30 GHz, 内存是 283 GB, 显卡为 GTX 1080Ti, 显存为 48 GB。

2.2 数据集

预训练和训练数据集均采用由 800 张 HR 图像组成的 DIV2K 数据集, 使用 Matlab 中采样因子为 4 的双三次内核对 HR 图像进行下采样获得 LR 图像。采用广泛使用的 Set5、Set14、BSDS100 和 Urban100 等基准数据集作为测试集。

2.3 网络训练

生成器剩余残余基本块数量为 16, 训练分为预训练及训练两个阶段。预训练过程中, 采用网络插值^[12]方法消除重建过程中产生的噪声。预训练分为两步: 首先采用像素相似性损失训练以 PSNR 为导向的模型, 在此基础上, 采用球面生成损失训练基于 GAN 的模型进行调优, 其中初始学习率为 2×10^{-4} 。训练过程中, 生成器与判别器均使用 Adam 优化器, 超参数设置为 $\beta_1 = 0.9, \beta_2 = 0.99$ 。初始学习率为 10^{-4} , 并在损失停止降低时减少 1/10 学习率, 直至学习率降至 10^{-6} , 训练周期为 500 000。

客观评价: 采用 PSNR 和 SSIM 作为客观评价指标。PSNR 基于图像像素点间的误差反映图像相似性: PSNR 越大, 图像相似性越高; PSNR 越小, 图像相似性越低。计算公式如下:

$$PSNR = 10 \times \lg \left(\frac{255 * WHC}{\sum_i \sum_j \sum_k (I_{i,j,k}^{HR} - I_{i,j,k}^{SR})^2} \right) \quad (9)$$

其中: I^{HR} 代表 HR 图像; I^{SR} 代表 SR 图像; W, H, C 代表图像宽度、高度和通道数。

SSIM 从亮度、对比度、结构三个方面进行对比, SSIM 取值范围是 $[-1, 1]$, SSIM 越接近 1, 两张图像结构越相似。计算公式如下:

$$SSIM(x, y) = \frac{(2\mu_x \mu_y + c1)(2\sigma_{xy} + c2)}{(\mu_x^2 + \mu_y^2 + c1)(\sigma_x^2 + \sigma_y^2 + c2)} \quad (10)$$

其中: x 代表 SR 图像; y 代表 HR 图像; μ 代表均值, σ_x, σ_y 代表 x, y 的方差; σ_{xy} 代表 x, y 的协方差。

2.4 结果分析

为了展示本文算法的重建效果以及损失收敛情况, 与基于深度学习的重建方法以及传统重建算法进行对比。其中超分辨率残差网络 (Super-Resolution Residual Network, SRResNet) 在 SISR 重建任务中具有良好的可移植性, 可用于基于 GAN 的图像超分辨率生成器部分; SRGAN 作为 GAN 应用于图像重建的开山之作, 关注人类视觉主观评价; ESRGAN 作为 SRGAN 的改进算法, 对重建任务的理论研究以及实际应用具有突出贡献, 并获得了 ECCV2018 PRIM-SR 比赛冠军。为保证数据公平性, SDSRGAN 框架中生成器基本块数量与 ESRGAN、SRGAN 以及 SRResNet 一致。图 4 为 SDSRGAN 与 ESRGAN 算法在训练过程中生成器损失曲线变化; 图 5 为 PSNR 结果趋势对比, 测试数据集为 Set5。表 1 为分别在 Set5、Set14、BSDS100 以及 Urban100 数据集上对 Bicubic、SRResNet、SRGAN、ESRGAN、SDSRGAN 几种算法进行重建的 PSNR 与 SSIM 数据统计均值; 图 6、7 为重建图像中纹理细节丰富区域的细节对比和具体 PSNR、SSIM, 它们分别选取的是 Set5 中

“baby”图片和 Urban100 中“img_047”图片。除图 4 外,其余结果均采用图像 Y 通道的 PSNR 和 SSIM。

由图 4 可知,训练时 SRGAN 和 ESRGAN 生成器损失处于 0 至 0.1 范围内持续震荡,无明显收敛迹象。SDSRGAN 生成器损失仅在 1×10^5 处相对不稳定,但总体损失值由 0.15 收敛至 0~0.05 范围,最后趋于平缓。这说明 SDSRGAN 的生成器损失收敛效果优于 SRGAN 和 ESRGAN。

从图 5 所示的 PSNR 曲线趋势可知,随着迭代次数增加,SRGAN、ESRGAN、SDSRGAN 三种算法的 PSNR 值增加,且增加趋势逐渐平缓。其中 SDSRGAN 的 PSNR 明显高于另外两种算法,说明 SDSRGAN 改善了基于 GAN 的重建图像 PSNR。

由表 1 的数据统计结果不难看出,与 Bicubic、SRGAN、ESRGAN 相比,SDSRGAN 重建图像的 PSNR 和 SSIM 在几种基准数据集上均取得最优值,且 SDSRGAN 比性能最好的 ESRGAN 算法的 PSNR 高 0.80 dB、SSIM 高 0.031。这表明在基于 GAN 的重建算法中 SDSRGAN 算法的 PSNR、SSIM 评价指标表现优越。相较 SRResNet 算法,虽然 SDSRGAN 算法的

评价指标较低,但是重建图像却更具纹理细节,具体参考图 6、7 的细节对比图。

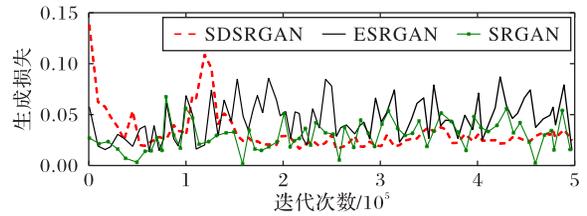


图 4 生成器损失曲线

Fig. 4 Generator loss graph

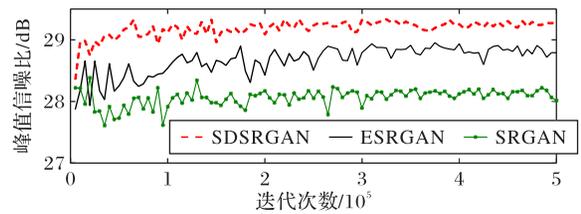


图 5 PSNR 曲线

Fig. 5 PSNR trend graph

表 1 基准数据集上重建结果的 PSNR 和 SSIM 统计

Tab. 1 Statistics of PSNR and SSIM of the reconstruction results on the benchmark dataset

模型	PSNR/dB				SSIM			
	Set5	Set14	BSDS100	Urban100	Set5	Set14	BSDS100	Urban100
Bicubic	28.43	25.89	25.98	23.14	0.810	0.705	0.669	0.659
SRResNet	32.24	28.28	27.59	26.15	0.895	0.770	0.737	0.788
SRGAN	29.96	26.36	25.59	24.37	0.850	0.702	0.656	0.729
ESRGAN	30.71	26.37	25.91	24.88	0.852	0.698	0.670	0.747
SDSRGAN	31.28	27.34	26.71	25.75	0.872	0.738	0.705	0.777

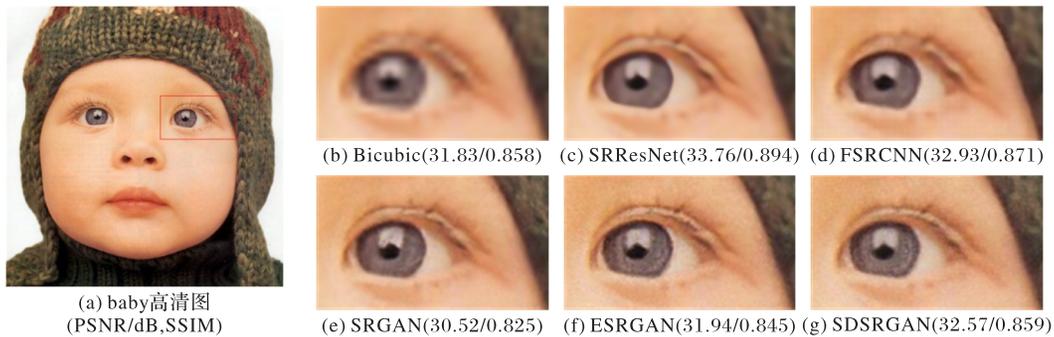


图 6 Set5 数据集中“baby”图像细节对比

Fig. 6 Comparison of “baby” image details in Set5 dataset

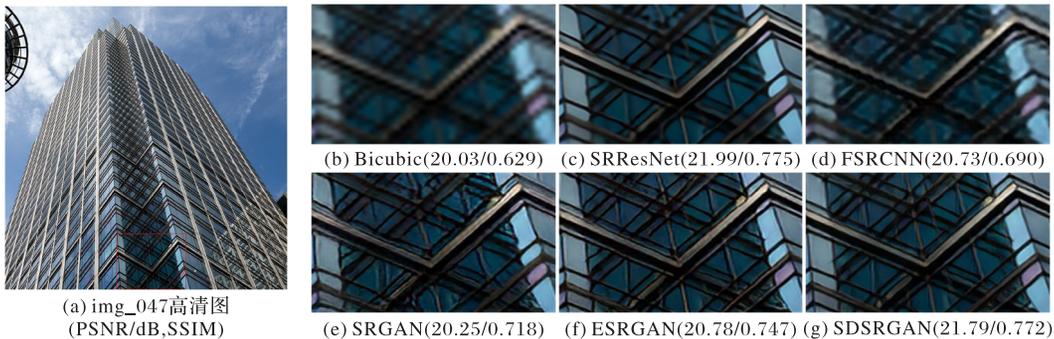


图 7 Urban100 数据集中“img_047”图像细节对比

Fig. 7 Comparison of “img_047” image details in Urban100 dataset

对比几种重建图像的纹理细节,相较未引入 GAN 的网络,FSRCNN 和 SRResNet 重建图像具有较高的 PSNR 和 SSIM

值,但是就人类视觉感官而言,缺乏纹理细节;并且图像结构越复杂,FSRCNN 和 SRResNet 优势越小,对于简单图像

“baby”, SRResNet、FSRCNN 分别比 SDSRGAN 的 PSNR 高 1.19 dB、0.36 dB, SSIM 高 0.035、0.012。但对于复杂图像 “img_047”, SRResNet 只比 SDSRGAN 的 PSNR 高 0.2 dB、SSIM 高 0.003, FSRCNN 反而比 SDSRGAN 的 PSNR 低 1.06 dB、SSIM 低 0.082。与基于 GAN 的重建算法相比, ESRGAN 和 SDSRGAN 具有较丰富的纹理细节, 但是 SDSRGAN 重建了图像相关的纹理细节, ESRGAN 重建纹理细节丰富, 但具有一定伪像。因此得出结论, SDSRGAN 具有较优的 PSNR 和 SSIM 评价指标, 且具有与图像相关的逼真纹理细节。

3 结语

本文提出了一个双判别器生成对抗网络进行图像端到端超分辨率重建。在判别网络中将图像特征映射至高维空间并提取高频特征, 快速收敛了生成损失, 避免了模式崩溃问题。实验结果表明, 本文算法在不同数据集上重建性能具有普适性, 提高了重建图像质量。本文算法针对固定比例因子的图像重建, 网络灵活性较低, 未来的研究重点是针对上采样模块进行改进, 使网络适用于实际场景中任意比例因子的图像重建。

参考文献 (References)

- [1] 张良培, 沈焕锋, 张洪艳, 等. 图像超分辨率重建[M]. 北京: 科学出版社, 2012: 3-11. (ZHANG L P, SHEN H F, ZHANG H Y, et al. Image Super-Resolution Reconstruction [M]. Beijing: Science Press, 2012: 3-11)
- [2] 谢颂华, 陈黎, 聂晖. 基于联合插值—恢复的超分辨率图像盲复原[J]. 计算机应用, 2010, 30(2): 341-343, 347. (XIE S H, CHEN L, NIE H. Blind image super-resolution restoration based on joint interpolation-restoration scheme [J]. Journal of Computer Applications, 2010, 30(2): 341-343, 347.)
- [3] SEEMA R, BAILEY K. Multi-frame image super-resolution by interpolation and iterative backward projection [C]// Proceedings of the 2nd International Conference on Signal Processing and Communication. Piscataway: IEEE, 2019: 36-40.
- [4] 陈晨, 赵建伟, 曹飞龙. 基于四通卷积稀疏编码的图像超分辨率重建方法[J]. 计算机应用, 2018, 38(6): 1777-1783. (CHEN C, ZHAO J W, CAO F L. Image super-resolution reconstruction method based on four-channel convolutional sparse coding [J]. Journal of Computer Applications, 2018, 38(6): 1777-1783.)
- [5] DONG C, LOY C C, TANG X. Accelerating the super-resolution convolutional neural network [C]// Proceedings of the 14th European Conference on Computer Vision, LNCS 9906. Cham: Springer, 2016: 391-407.
- [6] SHI W, CABALLERO J, HUSZÁR F, et al. Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network [C]// Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2016: 1874-1883.
- [7] DONG C, LOY C C, HE K, et al. Learning a deep convolutional neural network for image super-resolution [C]// Proceedings of the 13th European Conference on Computer Vision, LNCS 8692. Cham: Springer, 2014: 184-199.
- [8] 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. Piscataway: IEEE, 2016: 1637-1645.
- [9] LAI W S, HUANG J B, AHUJA N, et al. Deep Laplacian pyramid networks for fast and accurate super-resolution [C]// Proceedings of the IEEE 2017 Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2017: 5835-5843.
- [10] 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: MIT Press, 2014: 2672-2680.
- [11] LEDIG C, THEIS L, HUSZÁR 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. Piscataway: IEEE, 2017: 105-114.
- [12] WANG X, YU K, WU S, et al. ESRGAN: enhanced super-resolution generative adversarial networks [C]// Proceedings of the 2018 European Conference on Computer Vision, LNCS 11133. Cham: Springer, 2018: 63-79.
- [13] FUGLEDE B, TOPSOE F. Jensen-Shannon divergence and Hilbert space embedding [C]// Proceedings of the 2004 International Symposium on Information Theory. Piscataway: IEEE, 2004; No. 1365067.
- [14] GULRAJANI I, AHMED F, ARJOVSKY M, et al. Improved training of wasserstein GANs [C]// Proceedings of the 31st International Conference on Neural Information Processing Systems. Red Hook: Curran Associates Inc., 2017: 5767-5777.
- [15] LUO Y, ZHANG S, ZHENG W, et al. WGAN domain adaptation for EEG-based emotion recognition [C]// Proceedings of the 25th International Conference on Neural Information Processing, LNCS 11305. Cham: Springer, 2018: 275-286.
- [16] PARK S W, KWON J. Sphere generative adversarial network based on geometric moment matching [C]// Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2019: 4287-4296.
- [17] PARK S J, SON H, CHO S, et al. SRFeat: single image super-resolution with feature discrimination [C]// Proceedings of the 15th European Conference on Computer Vision, LNCS 11220. Cham: Springer, 2018: 445-471.

This work is partially supported by the Surface Program of National Natural Science Foundation of China (61673220).

LIN Jing, born in 1993, M. S. candidate. Her research interests include image processing, machine vision.

HUANG Yuqing, born in 1962, M. S., professor. Her research interests include image processing, machine vision, intelligent technology.

LI Leimin, born in 1960, M. S., professor. His research interests include robot control, wireless measurement and control, image processing, machine vision.