



基于生成对抗网络的图像超分辨率方法

包晓安¹, 高春波¹, 张娜¹, 徐璐¹, 吴彪²

(1. 浙江理工大学信息学院, 杭州 310018; 2. 山口大学东亚研究科, 日本山口 753-8514)

摘要: 为了解决生成对抗网络(Generative adversarial network, GAN)训练不稳定问题,降低模型复杂度,加快网络学习速率,提高超分辨率图像的视觉效果和重建速率,提出了一种基于改进生成对抗网络的图像超分辨率方法。该方法以改进的生成对抗网络为模型,通过粗粒度主体内容和细粒度细节边缘结合的方式提取图像特征,利用线性组合的方式重建超分辨率图像,采用 Wasserstein 距离优化生成对抗网络。实验结果表明:该方法能够生成视觉效果良好的超分辨率图像,在 Set5、Set14 等测试集上,其主观视觉评价和客观量化指标(PSNR、SSIM)都优于 SRGAN 方法。该方法通过重新设计网络模型,使得特征提取更为全面,网络训练更加充分,有助于提高超分辨率图像重建速度,提高图像质量。

关键词: 图像超分辨率;生成对抗网络;残差学习;深度学习;图像重建

中图分类号: TP181

文献标志码: A

文章编号: 1673-3851(2019)07-0499-10

Image super-resolution method based on generative adversarial network

BAO Xiaolan¹, GAO Chunbo¹, ZHANG Na¹, XU Lu¹, WU Biao²

(1. School of Information Science and Technology, Zhejiang Sci-Tech University, Hangzhou 310018, China;
2. Department of East Asian Studies, Yamaguchi University, Yamaguchi 753-8514, Japan)

Abstract: To solve the problem of training instability of generative adversarial network, reduce model complexity, and speed up network learning rate, and improve the visual effect and reconstruction rate of super-resolution image, an image super-resolution method based on improved generative adversarial networks is proposed. In the method, improved generative adversarial network is taken as the model, image features are extracted by combining main content of coarse granularity with detail edge of fine granularity, super-resolution images are reconstructed by means of linear combination, and generative adversarial network is optimized via Wasserstein distance. Experimental results show that super-resolution images with advanced visual effect can be generated with this method, and the method is superior to SRGAN in respect of subjective evaluation and objective quantification (PSNR/SSIM) in Set5, Set14 and such test sets. With this method, by redesigning the network model, feature extraction is conducted more comprehensively, and network training is conducted more completely, which helps to improve the speed of super-resolution image reconstruction and image quality.

Key words: image super-resolution; GAN; residual learning; deep learning; image reconstruction

收稿日期: 2018-11-06 网络出版日期: 2019-03-04

基金项目: 国家自然科学基金项目(61502430, 61562015); 广西自然科学基金重点项目(2015GXNSFDA139038); 浙江理工大学 521 人才培养计划

作者简介: 包晓安(1973-), 男, 浙江东阳人, 教授, 硕士, 主要从事软件测试、智能信息处理方面的研究。

通信作者: 张娜, E-mail: zhangna@zstu.edu.cn

0 引言

随着信息技术的发展,低分辨率(Low resolution, LR)图像已经难以满足人们对于高清影像的需求。高分辨率(High resolution, HR)图像的像素密度高,细节清晰,信息丰富,因而能够满足图像分析、图像理解的实际应用需求。在实际应用中,如果成像设备较为低端或受到存储容量等因素的限制,获取到的图像分辨率往往较低。因此,如何提高图像分辨率成为亟需解决的问题。图像超分辨率(Super resolution, SR)重建作为一种提高图像分辨率的技术,在医学影像、卫星遥感、视频监控等众多领域都具有广阔的前景和巨大的应用价值。

目前,图像超分辨率重建方法主要分为三类,分别是基于插值、重建和学习的方法。基于插值(Interpolation based)的方法^[1]假设图像具有连续的灰度值,待插像素的灰度值可以根据邻域像素估计得到。在实际应用中获取的图像往往不符合这种假设,因而重建出的图像容易出现锯齿、模糊等现象。基于重建(Reconstruction based)的方法^[2]通常需要将同一场景的多幅图像进行信息融合,其过程主要包括配准和重建两个步骤。该方法严重依赖高分辨率图像的先验知识,性能表现波动很大,当放大倍数较高时,图像重建质量会迅速下滑,导致图像细节严重丢失。基于学习(Learning based)的方法^[3]通过学习低分辨率图像和高分辨率图像之间的统计学关系,将低分辨率图像映射为高分辨率图像。相较于其他两类方法,基于学习的方法可以更好地表达和利用图像的先验知识,具有更突出的性能,恢复出的图像具有更优秀的视觉效果,是目前的研究热点之一。

深度学习在计算机视觉领域表现出巨大潜力,越来越多的学者开始探索深度学习与图像超分辨率之间的联系。Dong等^[4]提出了一种基于卷积神经网络的图像超分辨率(Super-resolution convolutional neural network, SRCNN)方法,首次将深度学习应用于图像超分辨率领域,在算法性能和重建效果方面都超越了传统的图像超分辨率重建方法。SRCNN方法构建了一种卷积神经网络模型,该模型主要包括特征提取、非线性映射和图像重建三个部分,极大地提升了超分辨率重建图像的量化指标和视觉效果。在此之后,又出现了若干基于深度学习的图像超分辨率方法。Kim等^[5]提出了基于深度递归网络(Deeply-recursive convolutional network, DRCN)的图像超

分辨率方法,增加了卷积网络的感受野,相较于SRCNN方法提高了图像重建的效果。Kim等^[6]又提出基于深度卷积网络的图像超分辨率(Very deep super-resolution, VDSR)方法,将残差的思想引入图像超分辨率,加快了网络学习的速度,同时降低了学习成本,证明了通过增加网络深度来提高图像分辨率的有效性。Tai等^[7]提出了基于递归残差网络(Deep recursive residual network, DRRN)的图像超分辨率方法,通过递归的残差结构来增加网络深度,在一定程度上提高了图像分辨率。Lai等^[8]构建了基于拉普拉斯金字塔的卷积神经网络(Laplacian super-resolution network, LapSRN)模型,通过逐步上采样的方式重建超分辨率图像。Ledig等^[9]提出了基于生成对抗网络(Generative adversarial network, GAN)的图像超分辨率(Super-resolution generative adversarial network, SRGAN)方法,采用生成对抗网络的架构,进一步提升了超分辨率图像的视觉效果。

尽管目前深度学习方法在一定程度上提升了图像超分辨率的复原效果,但是仍然存在以下问题^[10]:第一,浅层网络提取图像特征不够充分,模型学习能力不足;第二,随着网络加深,会出现梯度消失的现象,造成网络的不稳定;第三,忽视了超分辨率属于低层视觉任务的本质,冗余的网络层不仅对图像重建没有帮助,还影响了图像重建速率。

本文提出了基于生成对抗网络的图像超分辨率方法。该方法结合图像的高频特征和低频特征,改进了生成对抗网络的结构,挖掘原始低分辨率图像和高分辨率图像之间的映射关系,重新设计网络模型,从而恢复出图像质量更高的超分辨率图像。另外,根据生成对抗网络的特性,本文设计了更为简单的深度学习模型,修改了超分辨率图像的重建方式,删除冗余的网络层,减小模型的规模,加快网络训练和图像生成的速度。

1 生成对抗网络模型

1.1 模型原理

生成对抗网络包括生成模型(Generative model, G)和判别模型(Discriminative model, D)两个部分,训练过程采用二人零和博弈的思想^[11]。在原始GAN中,生成模型将噪声作为输入,输出生成数据。判别模型需要尽可能判别数据来自生成数据还是真实数据,生成模型则需要尽可能提高生成数据的可信度。GAN的损失函数(Loss function)定义如式(1)所示:

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

其中: x 为真实数据, p_{data} 为生成器的数据分布, z 为随机噪声, p_z 为输入噪声的先验概率。GAN 的目的是使得判别器无法判断出输入数据来自生成数据还是真实数据, 本质是使得生成数据和真实数据同分布。

本文提出的基于生成对抗网络的图像超分辨率

重建模型的更新流程如图 1 所示。生成模型 G 将低分辨率图像作为输入, 超分辨率图像作为输出。判别网络 D 将原始高分辨率图像和生成网络 G 生成的超分辨率图像作为输入, 判别结果作为输出。生成网络 G 和判别网络 D 相互对抗, 通过反向传播算法更新模型参数, 当判别网络 D 无法区分输入图像类型是高分辨率图像还是超分辨率图像时, 生成对抗网络就完成了图像重建的任务。

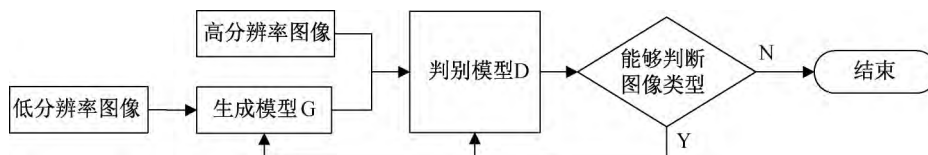


图1 基于生成对抗网络的超分辨率模型更新流程

1.2 模型设计与分析

通过 SRGAN 方法生成的图像具有良好的视觉效果, 但是在实验中仍有以下不足^[12]: a) 直接套用 GAN 框架的结构, 没有考虑 GAN 网络本身的不稳定性; b) 直接套用残差结构到图像超分辨率这样的低层视觉任务中, 可能导致特征提取不够充分, 图像信息单一, 影响重建图像的质量; c) 使用反卷积方式

进行图像重建, 图像生成速率有待提高。针对以上问题, 本文根据 GAN 的特性和低分辨率图像到高分辨率图像之间的映射关系, 重新设计了超分辨率模型的网络结构, 提出了基于改进的生成对抗网络的图像超分辨率方法 SRWGAN (Wasserstein GAN for image super-resolution)。SRWGAN 的网络结构如图 2 所示。

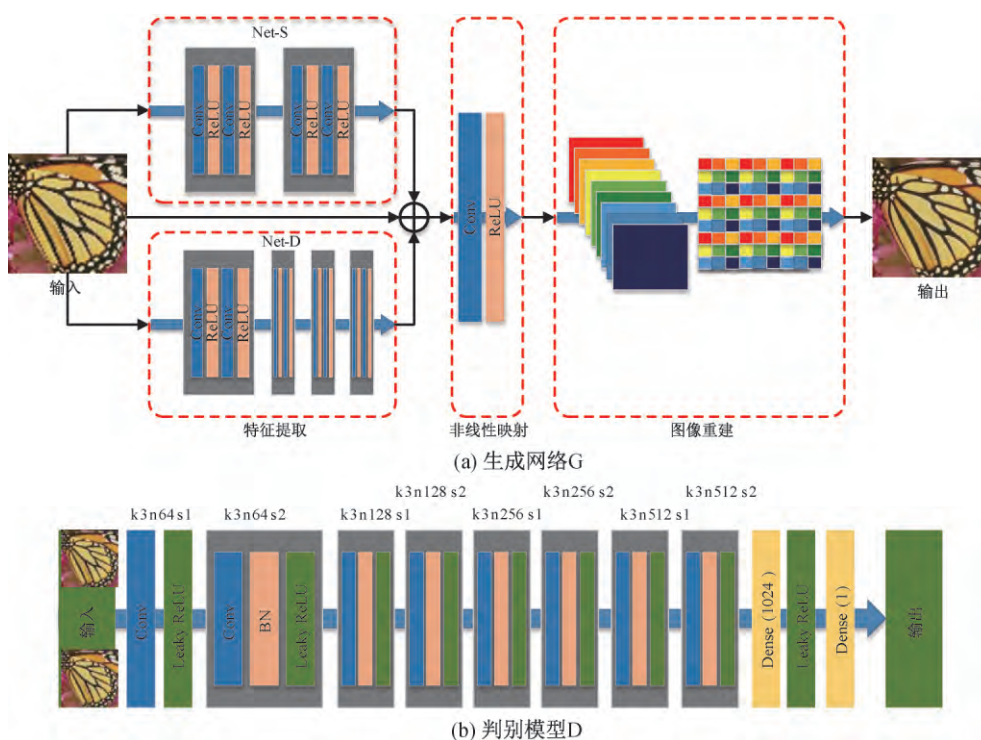


图2 SRWGAN 网络模型

图 2(a) 为改进的生成网络模型, 其中: Net-S 表示低频结构网络, Net-D 表示高频细节网络。生成模型将低分辨率图像作为输入, 通过 Net-S 和 Net-D 提取图像特征, 然后经过非线性映射和图像重建, 输出

超分辨率图像。图 2(b) 为改进的判别网络模型, 该模型与原始 SRGAN 的判别模型结构相同, 只在最后一层去除了 sigmoid 函数。

2 图像超分辨率方法

2.1 特征提取

低层视觉任务通常涉及两个部分:低频轮廓和高频细节^[13]。而 SRGAN 方法忽略了低层视觉任务的复杂性,导致提取到的图像特征较为单一,无法

拟合色彩纹理多变的图像。本文采用两条分支提取图像特征,同时处理高频细节信息和低频结构信息,低频结构网络(Net-S)提取低频结构特征,高频细节网络(Net-D)提取高频细节特征,然后利用残差结构对两方面的信息进行融合,从而得到完整图像特征。SRWGAN 特征提取过程如图 3 所示。

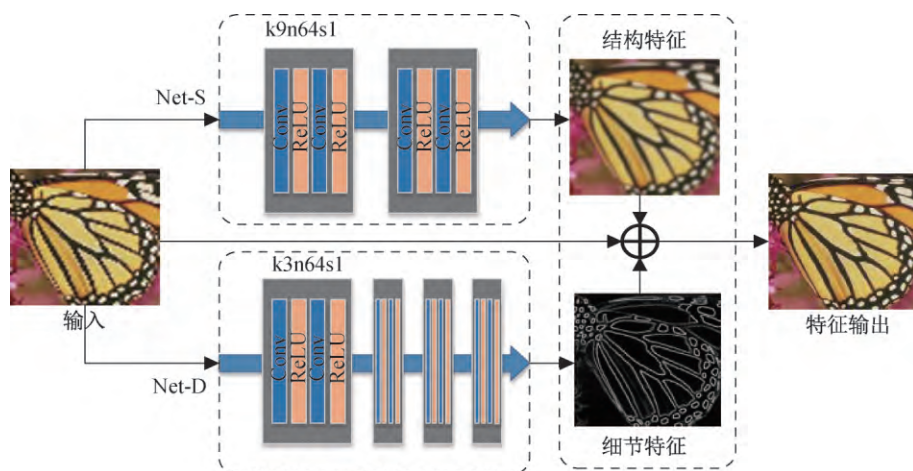


图3 SRWGAN 特征提取过程

如图 3 所示,低频结构网络采用 4 层卷积网络提取结构特征,每层包含一个卷积核数量为 64,大小为 9×9 的卷积网络和一个 ReLU 激活函数。低频结构网络特征提取过程如式(2)所示:

$$F_l(Y_{l-1}) = \max(0, W_l \cdot Y_{l-1} + b_l) \quad (2)$$

其中: l 为当前网络层数,取值范围为 $[1, 4]$; F_l 为当前卷积层的特征提取操作; W_l 为当前卷积层的参数; b_l 为偏置项; Y_{l-1} 为上一层的输出,即当前层的输入,当 $l=1$ 时, Y_0 为输入的低分辨率图像。卷积运算采用大小为 9×9 的滤波器,以此扩大特征提取的感知范围,抽象出图像的低频特征。高频细节网络的网络结构与低频结构网络类似,但是采用 8 层卷积网络,并将卷积核大小设置为 3×3 。通过更深的卷积层和较小的卷积核,能够为网络提供更深的感受野,同时将图像块进行更细致的分解,以此提高对图像细节信息的处理能力。当低频结构网络和高频细节网络各自完成特征提取操作时,对各自提取到的信息进行融合,信息融合过程如式(3)所示:

$$Y = X + \mu h_s(X) + \nu h_d(X) \quad (3)$$

其中: μ 为低频结构网络的非负权重, ν 为高频细节网络的非负权重, X 为输入的低分辨率图像, h_s 为低频结构网络的特征提取函数, h_d 为高频细节网络的特征提取函数, Y 为细节信息和结构信息融合后的特征图像。

在 SRGAN 中,每个残差块中都有批归一化

(Batch normalization, BN)^[14] 的处理过程。批归一化是通过均值和方差的方式对卷积后的数据进行白化处理、重构变换的过程,能够增强网络对数据的学习能力,在一定程度上可以避免梯度消失,加快网络收敛。批归一化处理适用于图像分类问题,但并不适用于图像超分辨率问题。批归一化白化中间特征的方式破坏了原始空间的表征,在重建过程中需要模型的部分层或者部分参数来恢复这种表征。同样多的参数,包含 BN 层的网络需要部分参数恢复这部分特征,导致网络收敛速度降低,loss 值波动变大。此外,BN 层增加了内存和计算负担,导致训练速度下降,所以在改进的网络中去除了 BN 层。

2.2 非线性映射和图像重建

原始低分辨率图像经过特征提取网络,输出与原始图像相同分辨率的特征图像。首先将特征图像输入卷积神经网络进行非线性映射,输出 $r \times r$ (r 为图像的目标放大倍数)个像素子通道;然后将每个通道通过线性组合方式生成一个 $r \times r$ 的像素区域,该区域对应于高分辨率图像的像素子块,从而将 $r \times r \times H \times W \times C$ (H 为图像的高, W 为图像的宽, C 为图像的通道数)的特征图像重建为 $r \times H \times r \times W \times C$ 的高分辨率图像。插值过程被隐含的包含在图像特征提取的过程中,只在最后一层通过线性组合重建图像。特征提取和非线性映射的卷积运算在低分辨率图像上进行,图像的超分辨率重建是一个线

性的过程,网络训练和图像生成的时间复杂度会大幅降低。非线性映射与图像重建过程如式(4)所示:

$$\mathbf{I}_{\text{SR}} = f_{\text{SR}}(f_{\text{mapping}}(\mathbf{Y})) \quad (4)$$

其中: \mathbf{I}_{SR} 为超分辨率图像, f_{mapping} 表示低分辨率图像到特征图像的特征映射, f_{SR} 表示重建超分辨率图像操作。非线性映射如式(5)所示。

$$f_{\text{mapping}}(\mathbf{Y}) = \max(0, \mathbf{W} \cdot \mathbf{Y} + \mathbf{b}) \quad (5)$$

其中:参数 \mathbf{W} 为非线性映射层的卷积核,其大小为 3×3 ,个数为 $r \times r$, \mathbf{b} 为偏置项。经过卷积运算,得到 $r \times r$ 个与输入图像相同分辨率的特征图像。重建过程如式(6)所示:

$$f_{\text{SR}}(\mathbf{T})_{w,h,c} = \mathbf{T}_{\lfloor h/r, \lfloor w/r, C \times r \bmod(h,r) + C \times \bmod(w,r) + c} \quad (6)$$

其中: h 为图像的高, w 为图像的宽, r 为图像的分辨率放大倍数, \mathbf{T} 为超分辨率特征图像集合。本层输入的图像为 $r \times r$ 个 $H \times W \times C$ 的低分辨率图像,经过线性组合,输出为超分辨率图像。

2.3 网络优化

GAN采用JS散度(Jessen-Shannon divergence, JS)^[15]计算生成数据和真实数据的距离,但是由于JS散度的不连续性,导致判别网络训练不稳定。因此,本文通过Wasserstein距离优化SRWGAN模型,以此提高图像重建的稳定性。Wasserstein距离^[16]定义如式(7)所示:

$$W_L(P_r, P_g) = \inf_{\gamma \in \Pi(P_r, P_g)} E_{(x,y) \sim \gamma} [\|x - y\|] \quad (7)$$

其中: P_r 为真实样本分布, P_g 为生成样本分布, y 为生成数据, $\Pi(P_r, P_g)$ 为所有联合分布 $\gamma(x, y)$ 的集合。对于任意可能的联合分布 γ ,从 $(x, y) \sim \gamma$ 中采样得到真实样本 x 、生成样本 y 、样本距离 $\|x - y\|$ 和距离的期望值 $E_{(x,y) \sim \gamma} [\|x - y\|]$ 。相比于JS散度,无论生成数据和真实数据的分布是否发生重叠,Wasserstein距离都可以反映它们的相似性,避免了梯度消失问题。为了求解Wasserstein距离,通过Lipschitz连续^[17]将Wasserstein距离变换为式(8)的形式:

$$W_L(P_r, P_g) \approx \max_{\theta: \|f_{\theta}\|_L \leq 1} E_{x \sim P_r} [f_{\theta}(x)] - E_{x \sim P_g} [f_{\theta}(x)] \quad (8)$$

其中: θ 为判别器参数, f_{θ} 表示判别器网络, L 为高分辨率图像和生成图像分布之间的Wasserstein距离。因此,优化SRWGAN模型的任务可以转换为:构建一个带参数 θ 的拟合函数 f_{θ} 来最大化Wasserstein距离 $W_L(P_r, P_g)$ 。原始SRGAN方法属于二分类任

务,判别器采用sigmoid作为损失函数,本文的SRWGAN方法属于回归任务,去掉了最后一层sigmoid函数,采用判别器 f_{θ} 近似拟合Wasserstein距离,最小化损失函数。损失函数的值越小,生成图像与高分辨率图像越相近。判别器D的loss函数和生成器G的loss函数如式(9)~(10)所示。

$$E_{x \sim P_g} [f_{\theta}(x)] - E_{x \sim P_r} [f_{\theta}(x)] = \nabla_{\theta} \left[\frac{1}{m} \sum_{i=1}^m f_{\theta}((\mathbf{I}_{\text{HR}})_i) - \frac{1}{m} \sum_{i=1}^m f_{\theta}(G((\mathbf{I}_{\text{LR}})_i)) \right] \quad (9)$$

$$- E_{x \sim P_g} [f_{\theta}(x)] = \nabla_{\theta} \left[\frac{1}{m} \sum_{i=1}^m f_{\theta}(G((\mathbf{I}_{\text{LR}})_i)) \right] \quad (10)$$

其中: \mathbf{I}_{LR} 为低分辨率图像, \mathbf{I}_{HR} 为高分辨率图像, m 为批尺寸。

本文超分辨率方法主要流程如下:

- 1) 开始;
- 2) 初始化参数,令学习率 $a = 0.00005$,裁剪参数 $c = 0.01$,单次迭代次数 $n_{\text{critic}} = 5$, $m = 64$;
- 3) 定义生成器参数 ω ;
- 4) 开始训练;
- 5) 开始 n_{critic} 次迭代;
- 6) 从高分辨率图像中采样 m 张图像 $\{(\mathbf{I}_{\text{HR}})_i\}$;
- 7) 从超分辨率图像中采样 m 张图像 $\{(\mathbf{I}_{\text{SR}})_i\}$;
- 8) 根据式(9)计算判别损失 g_{θ} ;
- 9) 通过 RMSPProp ^[18]算法更新参数 $\theta \leftarrow \theta + a \cdot \text{RMSPProp}(\theta, g_{\theta})$;
- 10) 裁剪参数 $\theta \leftarrow \theta + \text{clip}(\theta, -c, c)$;
- 11) 结束迭代过程;
- 12) 从低分辨率图像中采样 m 张图像 $\{(\mathbf{I}_{\text{LR}})_i\}$;
- 13) 根据式(10)计算生成损失 g_{ω} ;
- 14) 更新参数 $\omega \leftarrow \omega + a \cdot \text{RMSPProp}(\omega, g_{\omega})$;
- 15) 结束训练;
- 16) 保存模型参数;
- 17) 结束。

3 仿真实验与结果分析

3.1 实验环境

网络训练采用的数据集为BSD200(共200张)、General100(共100张)和T91(共91张),共391张不同的彩色图像,对以上图像进行降采样处理,得到相应的低分辨率图像。将图像裁剪为步长为3,大小为 7×7 的图像块,得到381073张图像以扩充训练集。用于测试的数据集都是通用数据集,

分别为 BSD100(共 100 张)、Set14(共 14 张)和 Set5(共 5 张)。

实验的软件环境:Tensorflow1.2 深度学习框架。用于训练和测试的硬件配置:Ubuntu 16.04 操作系统,NVIDIA GeForce GTX 1080 Ti 显卡,Intel(R) Core(TM) i7-4790 K CPU @4.00 GHz 处理器,Kingstom DDR332 GB 内存。

现有超分辨率图像质量评价指标构建主要包括主观评价和客观量化。主观评价是指通过人眼观察超分辨率图像,评定重建图像的质量。客观量化是指定量分析重建图像的数学指标。本文采用图像超分辨问题普遍的两种量化方式,即峰值信噪比(Peak signal to noise ratio,PSNR)^[19]和结构相似

性(Structural similarity index method,SSIM)^[19]。重建图像的 PSNR 值越高,表明超分辨率图像失真越少,图像质量越高;超分辨率图像和高分辨率图像的 SSIM 值越趋近于 1,表明超分辨率图像与高分辨率图像越相似。

3.2 图像重建实验分析与对比

3.2.1 BN 层对比验证

为了验证 BN 层对算法性能的影响,实验分别对去除 BN 层和保留 BN 层的 SRGAN 模型和 SRWGAN 模型进行 300 次迭代,在每次迭代完成后记录模型的 loss 值,并且计算模型生成 4 倍图像的平均 PSNR 值和平均 SSIM 值。实验结果如表 1 和图 4 所示。

表 1 Batch Normalization 层对网络性能影响

模型	数据集	放大倍数	保留 BN 层			去除 BN 层		
			PSNR/dB	SSIM	运行时间/s	PSNR/dB	SSIM	运行时间/s
SRGAN	Set5	4	31.63	0.884	0.20	31.68	0.885	0.19
	Set14	4	28.03	0.768	0.41	28.10	0.770	0.40
	BSD100	4	27.10	0.725	0.20	27.21	0.728	0.18
SRWGAN	Set5	4	31.70	0.889	0.11	31.79	0.890	0.08
	Set14	4	28.21	0.772	0.16	28.36	0.778	0.12
	BSD100	4	27.32	0.731	0.11	27.49	0.739	0.07

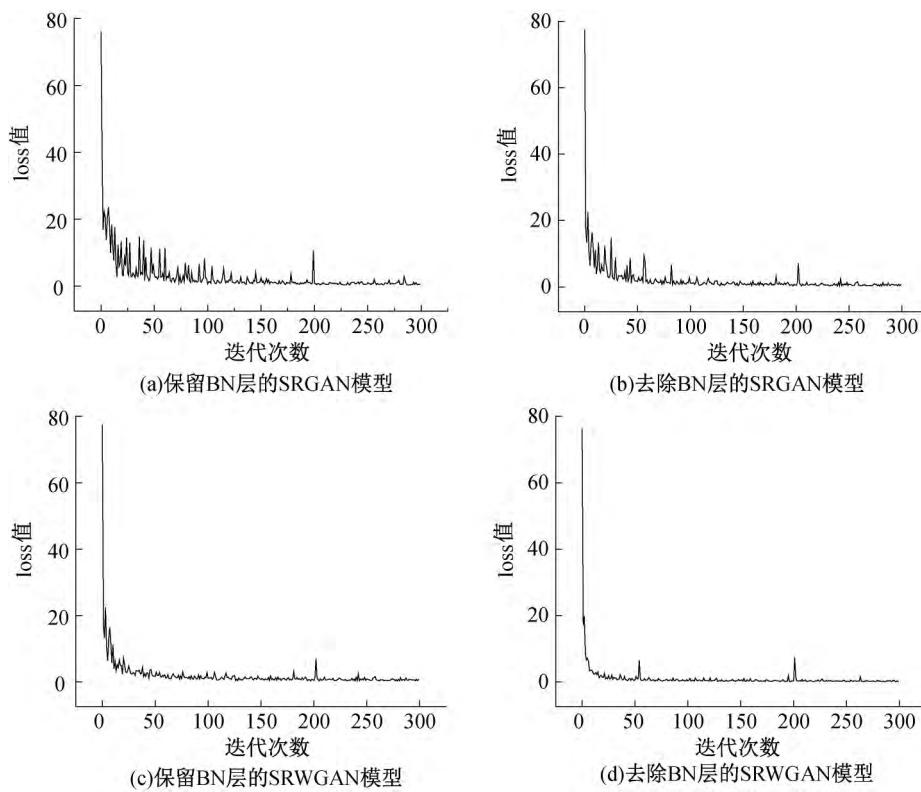


图 4 损失值随迭代次数变化曲线

表 1 反映了图像生成质量受 BN 层影响的情况,其中去除 BN 层的网络模型重建图像的平均

PSNR 和平均 SSIM 值均有所提升,而且去除 BN 层的 SRWGAN 模型的图像生成时间均有所下降。

图4显示了模型的loss值随迭代次数的变化情况,其loss值的迭代曲线总体趋于下降。对于保留BN层的SRGAN模型,其loss值在128次迭代后趋于稳定状态,在221次迭代后出现最小值0.4130。对于去除BN层的SRGAN模型,其loss值在97次迭代后趋于稳定状态,在145次迭代后出现最小值0.4042。对于保留BN层的SRWGAN模型,其loss值在38次迭代后趋于稳定状态,在128次迭代后出现最小值0.3983。对于本文的SRWGAN模型,其loss值在21次迭代后趋于稳定状态,在76次迭代后出现最小值0.2708。相对于保留BN层的网络模型,本文的SRWGAN方法保留了原始图像

的空间表征,加快了网络训练的收敛速度,而且网络训练更为稳定,更早达到饱和状态。实验结果表明,去除BN层有助于加快图像生成速度,提高图像生成质量。

3.2.2 框架修改对比验证

将原始SRGAN模型作为对比,对本文提出的改进模型实验分析,实验包括4个部分:a)原始SRGAN模型,b)改进特征提取网络的SRGAN模型,c)改进图像重建方式的SRGAN模型,d)改进优化函数的SRGAN模型。分别对4种模型进行300次迭代,记录损失函数随时间变化曲线。实验结果如图5所示。

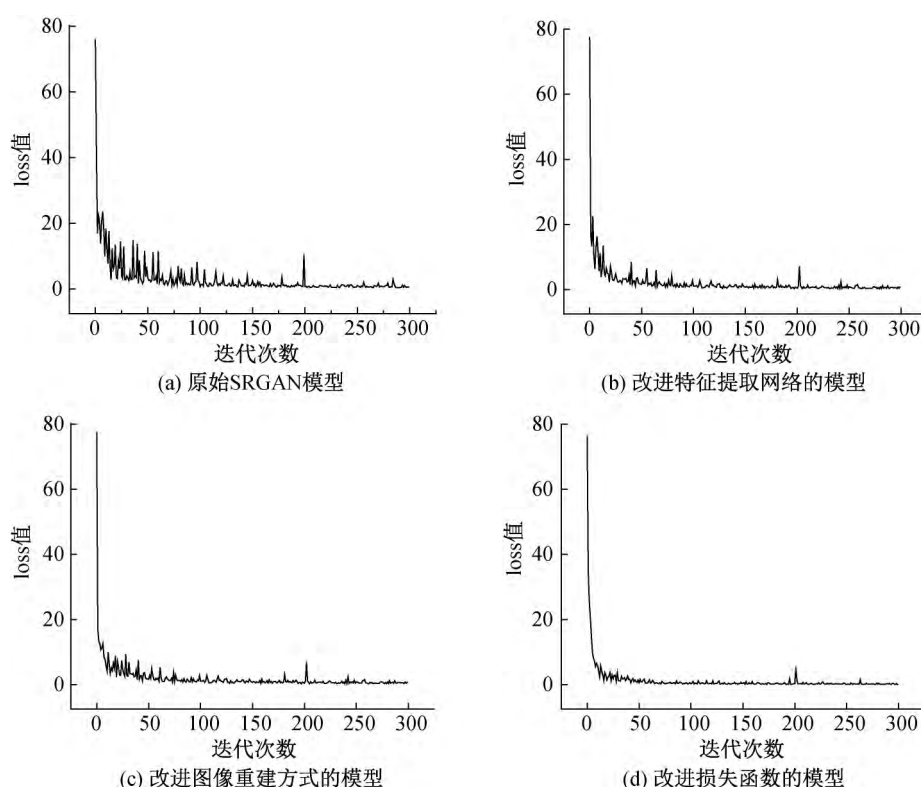


图5 损失函数随迭代次数变化曲线

图5显示了改进模型的loss值随迭代次数的变化。对于原始SRGAN模型,其损失函数的训练波动巨大,需要多次迭代才能达到收敛状态。对于改进特征提取网络的模型,虽然训练过程中存在一定的波动,但是损失函数的曲线相对稳定。对于改进图像重建方式的模型,前期能够迅速达到收敛状态,网络训练过程相对平缓。对于改进损失函数的模型,其loss值在训练过程中的波动很小,损失函数曲线最为平滑,需要较少的训练即可得到最优解。以上实验结果表明,本文提出的改进方案能够有效提高网络性能,高频细节和低频结构融合的特征提取方式能够提升图像信息的完整性,线性重建的方

式能够加快图像的生成速率,Wasserstein距离优化模型的方式能够提高网络训练的稳定性。

3.2.3 图像重建对比验证

对本文所提SRWGAN模型进行实验分析,实验选取近年来的经典超分辨率方法进行对比,参与对比的方法包括Bicubic、SRCNN、VDSR、DRCN、SRGAN以及本文的SRWGAN方法。分别记录每个模型的平均图像生成时间、重建图像的平均PSNR值和平均SSIM值,并展示生成图像的局部放大图像。表2—表4显示了Bicubic、SRCNN、VDSR、DRCN、SRGAN本文的SRWGAN模型在不同放大倍数下重建图像的PSNR、SSIM和生成

时间。在相同测试集和相等放大倍数的情况下,本文所提的 SRWGAN 方法能够用最短的时间重建出 PSNR 和 SSIM 值较高的超分辨率图像。图 6 显示了以上几种方法在不同放大倍数下的部分重建图像。当前显示图片为 butterfly (Set5) 和 baboon

(Set14),放大倍数分别为 2 倍、3 倍和 4 倍。由重建图像可以看出,本文所提的 SRWGAN 方法重建出的图像边缘最为清晰,纹理更加清楚。以上实验结果表明,SRWGAN 能够降低模型时间复杂度,重建出视觉效果良好的超分辨率图像。

表 2 不同超分辨率方法的平均 PSNR 值

数据集	放大倍数	超分辨率方法的平均 PSNR/dB					
		Bicubic	SRCNN	VDSR	DRCN	SRGAN	SRWGAN
Set5	2	33.69	36.60	37.53	37.63	37.74	37.78
	3	30.41	32.62	33.67	33.83	33.82	34.08
	4	28.43	30.32	31.35	31.54	31.54	31.79
Set14	2	30.25	32.32	33.05	33.06	33.08	33.09
	3	27.55	29.15	29.78	29.77	29.58	29.75
	4	26.01	27.34	28.02	28.03	28.03	28.36
BSD100	2	29.57	31.24	31.90	31.85	31.95	32.03
	3	27.22	28.31	28.83	28.82	28.74	28.94
	4	25.97	26.83	27.29	27.24	27.10	27.49

表 3 不同超分辨率方法的平均 SSIM 值

数据集	放大倍数	超分辨率方法平均的 SSIM 值					
		Bicubic	SRCNN	VDSR	DRCN	SRGAN	SRWGAN
Set5	2	0.931	0.955	0.959	0.959	0.950	0.956
	3	0.869	0.909	0.921	0.912	0.910	0.923
	4	0.811	0.863	0.883	0.884	0.879	0.890
Set14	2	0.870	0.908	0.913	0.912	0.913	0.917
	3	0.775	0.823	0.832	0.836	0.832	0.838
	4	0.704	0.753	0.768	0.768	0.770	0.778
BSD100	2	0.844	0.889	0.896	0.898	0.892	0.896
	3	0.741	0.786	0.799	0.797	0.801	0.812
	4	0.670	0.712	0.726	0.725	0.723	0.739

表 4 不同超分辨率方法的平均图像生成时间

数据集	放大倍数	超分辨率方法的平均图像生成时间/s					
		Bicubic	SRCNN	VDSR	DRCN	SRGAN	SRWGAN
Set5	2	0.00	2.18	0.12	0.64	0.20	0.08
	3	0.00	2.24	0.14	0.45	0.21	0.09
	4	0.00	2.11	0.15	0.33	0.20	0.08
Set14	2	0.00	4.34	0.26	1.12	0.42	0.11
	3	0.00	4.45	0.27	0.81	0.43	0.12
	4	0.00	4.36	0.25	0.63	0.41	0.13
BSD100	2	0.00	2.57	0.17	0.84	0.19	0.07
	3	0.00	2.58	0.20	0.65	0.22	0.07
	4	0.00	2.59	0.20	0.46	0.21	0.06

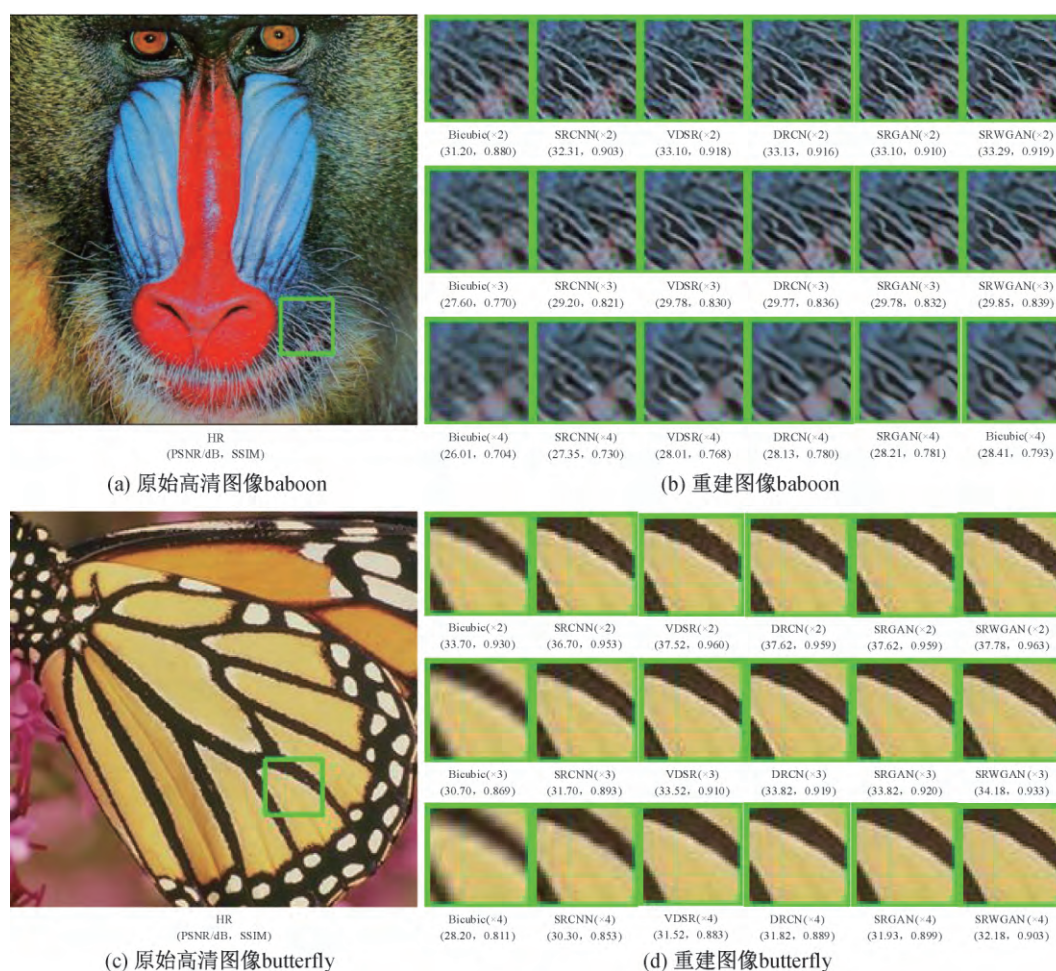


图6 不同超分辨率方法的重建图像

4 结 语

本文提出了一种基于生成对抗网络的图像超分辨率方法,该方法通过图像的高频结构特征和低频细节特征融合,使深度网络提取的特征更为全面;利用线性的方式对图像进行重建,有效提高了图像重建的速度;采用 Wasserstein 距离优化 GAN 网络,保证了算法运行的稳定性;去除 BN 层可以有效提升算法性能。通过 Set5、Set14 和 BSD100 数据集对所提算法进行验证,实验结果表明,SRWGAN 方法可以有效降低时间复杂度,提升网络训练的稳定性,其重建图像的 PSNR、SSIM 值和视觉效果均优于 SRGAN。

由于本文采用高频细节网络和低频结构网络两条分支提取图像特征,可能会出现对称解,从而造成网络训练的不稳定,后续将研究解决对称解造成的梯度消失问题,提高模型学习效率。

参考文献:

[1] Zhou F, Yang W, Liao Q. Interpolation-based image super-resolution using multisurface fitting [J]. IEEE

Transactions on Image Processing, 2012, 21(7): 3312-3318.

- [2] 曾凯,丁世飞. 图像超分辨率重建的研究进展[J]. 计算机工程与应用, 2017, 53(16): 29-35.
- [3] 胡长胜, 詹曙, 吴从中. 基于深度特征学习的图像超分辨率重建[J]. 自动化学报, 2017, 43(5): 814-821.
- [4] 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.
- [5] Kim J, Lee J K, Lee K M. Deeply recursive convolutional network for image super-resolution[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas. IEEE, 2016: 1637-1645.
- [6] 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.
- [7] Tai Y, Yang J, Liu X. Image super-resolution via deep recursive residual network [C]// Proceedings of the

- IEEE Conference on Computer Vision and Pattern Recognition, Hawaii, IEEE, 2017: 2790-2798.
- [8] Lai W S, Huang J B, Ahuja N, et al. Deep Laplacian pyramid networks for fast and accurate super-resolution[C]//IEEE Conference on Computer Vision and Pattern Recognition, Hawaii, IEEE, 2017: 5835-5843.
- [9] Ledig C, Theis L, Huszár F, et al. Photo-realistic single image super-resolution using a generative adversarial network[C]//IEEE Conference on Computer Vision and Pattern Recognition, Hawaii, IEEE, 2017: 105-114.
- [10] 孙旭, 李晓光, 李嘉锋, 等. 基于深度学习的图像超分辨率复原研究进展[J]. 自动化学报, 2017, 43(5): 697-709.
- [11] Goodfellow I, Pouget-Abadie J, Mirza M, et al. Generative adversarial nets[C]//Advances in Neural Information Processing Systems. Cambridge: MIT Press, 2014: 2672-2680.
- [12] 王一宁, 秦品乐, 李传朋, 等. 基于残差神经网络的图像超分辨率改进算法[J]. 计算机应用, 2018, 38(1): 246-254.
- [13] Pan J, Liu S, Sun D, et al. Learning dual convolutional neural networks for low-level vision[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, IEEE, 2018: 3070-3079.
- [14] 麻旋, 戴曙光. 基于残差网络的图像超分辨率算法改进研究[J]. 软件导刊, 2018, 17(4): 91-93.
- [15] 王耀杰, 钮可, 杨晓元. 基于生成对抗网络的信息隐藏方案[J]. 计算机应用, 2018, 38(10): 2923-2928.
- [16] Arjovsky M, Chintala S, Bottou L. Wasserstein generative adversarial networks [C]//International Conference on Machine Learning, Sydney: PMLR, 2017: 214-223.
- [17] 杨延涛. L_p 空间中 Lipschitz 强单调算子方程解的迭代算法[J]. 浙江大学学报(理学版), 2018, 45(4): 405-408.
- [18] Andrychowicz M, Denil M, Gomez S, et al. Learning to learn by gradient descent by gradient descent[C]//Advances in Neural Information Processing Systems. Cambridge: MIT Press, 2016: 3981-3989.
- [19] Hore A, Ziou D. Image quality metrics: PSNR vs. SSIM [C]// International Conference on Pattern Recognition, Istanbul, IEEE, 2010: 2366-2369.

(责任编辑: 康 锋)