

基于改进增强型超分辨率生成对抗网络的图像超分辨率重建算法

辛元雪², 朱凤婷², 史朋飞^{1,2*}, 杨鑫², 周润康²

¹河海大学江苏省输配电装备技术重点实验室, 江苏 常州 213022;

²河海大学物联网工程学院, 江苏 常州 213022

摘要 针对现有的图像超分辨率重建算法仍存在细节上处理不足的问题, 提出一种基于改进增强型超分辨率生成对抗网络(ESRGAN)的图像超分辨率重建算法。首先, 使用多尺度密集连接模块(MDB)代替密集连接模块(DB), 并且在MDB后添加通道注意力机制, 以调整不同通道的特征响应值, 从而改进了ESRGAN的生成网络模型中的深层信息提取模块。其次, 通过改进ESRGAN模型中的浅层特征提取模块提取低分辨率图像的原始特征, 并通过深层信息提取模块提取低分辨率图像的深度残差特征, 将原始特征和深度残差特征以对应元素相加的方式融合。最后, 利用重建模块完成图像超分辨率重建。在Set5、Set14和BSD100数据集上进行2倍和4倍超分辨率重建测试, 并对所提算法与Bicubic、FSRCNN、ESRGAN方法进行对比。结果显示, 所提算法获得的重建图像边缘更加清晰, 能够提供相对较多的细节, 大大提升了图像的视觉效果。在客观质量评价方面, 所提算法2倍超分辨重建后图像的峰值信噪比(PSNR)平均值相比ESRGAN提高了0.467 dB, 结构相似性(SSIM)平均值提高了0.005; 4倍超分辨重建后图像的PSNR平均值相比ESRGAN提高了0.438 dB, SSIM平均值提高了0.015。

关键词 图像处理; 超分辨率; 生成对抗网络; 注意力机制; 多尺度密集连接

中图分类号 TP391.41

文献标志码 A

doi: 10.3788/LOP202259.0420002

Super-Resolution Reconstruction Algorithm of Images Based on Improved Enhanced Super-Resolution Generative Adversarial Network

Xin Yuanxue², Zhu Fengting², Shi Pengfei^{1,2*}, Yang Xin², Zhou Runkang²

¹Jiangsu Key Laboratory of Power Transmission & Distribution Equipment Technology,

Hohai University, Changzhou, Jiangsu 213022, China;

²College of Internet of Things Engineering, Hohai University, Changzhou, Jiangsu 213022, China

Abstract To solve the problem of insufficient detail processing in the existing image super-resolution reconstruction algorithm, a super-resolution reconstruction algorithm of images based on improved enhanced super-resolution generative adversarial network (ESRGAN) is proposed. Firstly, the deep information extraction module of the improved ESRGAN generation network is improved using multiscale dense block (MDB) instead of dense block (DB), and by adding channel attention mechanism after MDB to adjust the characteristic response values of different channels. Secondly, the shallow feature extraction module of the improved ESRGAN model is used to extract the original features of the low resolution images, and the deep information extraction module is used to extract the depth

收稿日期: 2021-07-21; 修回日期: 2021-08-24; 录用日期: 2021-09-13

基金项目: 国家自然科学基金(61801169)、江苏省输配电装备技术重点实验室自由探索课题(2021JSSPD03)、中央高校基本科研业务费(B210202087)

通信作者: *shipf@hhu.edu.cn

residual features of the low resolution images. The original features and the depth residual features are fused by adding the corresponding elements. Finally, the reconstruction module is used to complete the image super-resolution reconstruction. The proposed algorithm's two and four times super-resolution reconstructions are tested on Set5, Set14, and BSD100 datasets and compared to Bicubic, FSRCNN, and ESRGAN methods. The results show that the proposed algorithm's reconstructed image has a clearer edge, and it can provide more details, which greatly improves the image's visual effect. Compared to ESRGAN, the proposed algorithm improves the average peak signal-to-noise ratio (PSNR) and structural similarity (SSIM) of 2-fold super-resolution reconstruction images by 0.467 dB and 0.005, respectively; At the same time, the proposed algorithm improves the average PSNR and SSIM of 4-fold super-resolution reconstruction images by 0.438 dB and 0.015, respectively.

Key words image processing; super-resolution; generative adversarial network; attention mechanism; multi-scale dense connection

1 引言

图像的超分辨率(SR)重建是一种通过单张或多张低分辨率图像获得高分辨率图像的技术。采用图像超分辨率重建技术对低分辨率图像进行恢复和重建,可以有效改善图像的细节,提升图像的质量。

图像的超分辨率重建算法大致可以分为三类,基于插值的算法^[1]、基于重建的算法^[2]和基于学习的算法^[3]。前两类属于传统的方法,通常存在着图像整体模糊、细节严重缺失等弊端,有较大的局限性。近年来,随着深度学习的发展,基于学习的超分辨率重建技术逐渐成为热点。Dong等^[4]提出超分辨率卷积神经网络(SRCNN),该网络使用三个卷积层进行重建,相比于传统方法,大大提升了重建的速度。Kim等^[5]提出了一种递归循环神经网络(DRCN),该网络采用递归循环和跳跃连接,相比于SRCNN,图像质量得到进一步提升。Lai等^[6]提出fast and accurate image super-resolution with deep Laplacian pyramid network(LapSRN),该网络减轻了图像的伪影问题并降低了计算复杂度。基于深度学习的方法显著提高了重建图像的视觉效果和性能指标,但是对图像的细节和纹理方面效果不佳。近年来,生成对抗网络(GAN)因能够通过判别器学习到比基于像素差异更有意义的损失函数,被广泛应用于超分辨率重建算法当中。Ledig等^[7]首次将GAN应用于单幅图像的超分辨率任务当中,称为SRGAN,能够生成逼真的纹理。胡诗语等^[8]提出了基于密集连接和激励模块的图像SR网络,局部特征和整体特征在密集连接和激励监督后输出至重建网络,进一步提高了生成图像的视觉质量。Wang等^[9]提出SRGAN的增强版本,称为增强型超分辨

率生成对抗网络(ESRGAN),使重构图像具有更好的纹理信息和更清晰的视觉效果。Hu等^[10]针对超分辨率重建任务的实时性,提出了实时超分辨率生成对抗网络(RTSRGAN),可以实时进行图像超分辨率重建。Dou等^[11]使用基于SRGAN的三维卷积层代替二维卷积层,用注意机制处理来自三维卷积层的多重特征,并通过改进生成器损失函数来提高模型的输出,能够高效地进行超分辨率重建任务。Zhang等^[12]提出了一种基于残差密集连接网络(RDN)的方法,该方法融合了多个残差密集块,能够有效地提取特征信息。Shao等^[13]提出了一种基于跨层注意转移机制的多尺度生成对抗网络模型,该模型利用跨层注意力转移模块,使高层语义特征图指导低层语义特征图的填充,保证了修复的视觉和语义一致性。李云红等^[14]提出基于deep super resolution generative adversarial network(DSRGAN)的图像修复与重建方法,加深了DenseNet的网络层数,消除了轮廓模糊和纹理不清晰的现象。彭晏飞等^[15]提出一种融合注意力的生成式对抗网络单图像超分辨率重建算法,构造了注意力卷积神经网络残差块,增加了图像高频信息。陈子涵等^[16]提出一种基于自注意力深度网络的图像超分辨率重建算法,该算法在深度网络的映射过程中引入多个自注意力来强化局部区域和其他位置的依赖关系,使得重建细节更加合理。查体博等^[17]在SRGAN的基础上,使用残差套残差密集块提高了网络提取特征的能力,保留了图像高频细节的同时又避免了伪影的出现。

针对现有算法仍存在伪影、细节处理不足等问题,本文在单幅图像超分辨率(SISR)算法ESRGAN的基础上改进其生成网络结构,提出了一种注意力机制与多尺度密集连接结合的图像超分辨率重建算法。

2 基于改进 ESRGAN 的图像超分辨率重建

2.1 ESRGAN 模型

ESRGAN 以 SRGAN 为基础,在其生成网络的结构上进行两点改进:删除生成网络中的所有批归一化(BN)层;用提出的多级残差密集连接模块(RRDB)替换 SRGAN 模型中的残差模块(RB)。其中,在面向不同的提高峰值信噪比(PSNR)的任务如提高图像超分辨率^[18]和去除图像模糊^[19]中,去

除 BN 层已被证明能够提高模型的性能。RRDB 将多级残差和密集连接相结合,相比于传统 SRGAN 中的残差连接模块,该结构具有更深的层次和更加复杂的结构,同时提升了特征表达能力,能够有效提高模型的性能。此外,为防止训练的不稳定,在残差加入主路径之前,将残差部分提取的特征乘以小于 1 的正常数 β ,完成残差部分缩放。图 1(a)展示了 SRGAN 中所采用的 RB,图 1(b)展示了 ESRGAN 中采用的 RRDB,图 1(c)为 RRDB 中密集连接模块(DB)的具体结构。

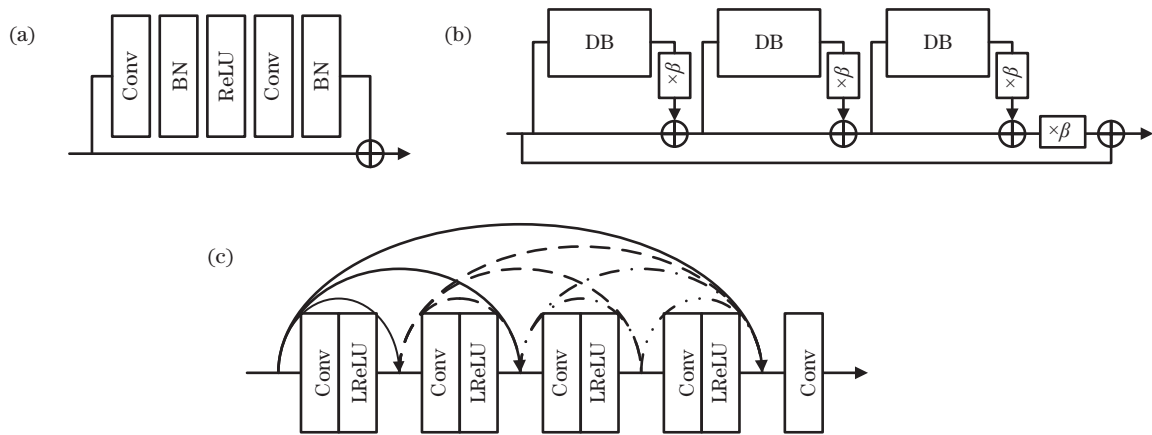


图 1 ESRGAN 特征提取模块结构改进图。(a) RB; (b) RRDB; (c) DB

Fig. 1 Improved structures of ESRGAN feature extraction module. (a) RB; (b) RRDB; (c) DB

ESRGAN 使用了基于相对论平均生成对抗网络(RaGAN)^[20]的判别器,与传统 GAN 中判别器估计输入数据是真实的概率不同,相对论判别器试图预测真实数据图像比生成数据更加真实且自然的概率。同时,RaGAN 的生成器在训练过程中结合了生成数据和真实数据的梯度,而传统 GAN 的生成器在训练过程中仅使用了生成数据的梯度。相对论判别器的数学表达形式为

$$D_{Ra}[\mathbf{y}, G(\mathbf{x})] = \sigma \{ C(\mathbf{y}) - E \{ C[G(\mathbf{x})] \} \} \rightarrow 1, (1)$$

$$D_{Ra}[G(\mathbf{x}), \mathbf{y}] = \sigma \{ C[G(\mathbf{x})] - E \{ C(\mathbf{y}) \} \} \rightarrow 0, (2)$$

式中: D_{Ra} 表示相对平均判别网络; \mathbf{x} 表示生成器输入数据; \mathbf{y} 表示训练集真实数据; σ 表示 Sigmoid 激活函数; $G(\cdot)$ 表示生成器的输出; $C(\cdot)$ 表示未激活判别器的输出; $E(\cdot)$ 表示对小批量中所有数据取均值的操作。当真实图像比合成图像真实自然时, $D_{Ra}(\mathbf{y}, \mathbf{x})$ 的结果趋向于 1,如(1)式所示;若合成图像比真实图像的质量差, $D_{Ra}(\mathbf{x}, \mathbf{y})$ 的结果越接近 0,如(2)式所示。

根据相对论判别器的原理,ESRGAN 中判别器和生成器的损失函数可分别定义为

$$L_D = -E \{ \log D_{Ra}[\mathbf{y}, G(\mathbf{x})] \} - E \{ \log \{ 1 - D_{Ra}[G(\mathbf{x}), \mathbf{y}] \} \}, (3)$$

$$L_G = -E \{ \log \{ 1 - D_{Ra}[\mathbf{y}, G(\mathbf{x})] \} \} - E \{ \log \{ D_{Ra}[G(\mathbf{x}), \mathbf{y}] \} \}. (4)$$

2.2 改进 ESRGAN 模型

2.2.1 生成网络结构

为提高图像的分辨率,增强图像的细节纹理,对 ESRGAN 的生成网络结构进行改进:1) 使用多尺度密集连接模块(MDB)代替原始的 DB;2) 在

MDB 后添加通道注意力机制,以调整不同通道的特征响应值。将结合多尺度密集连接和注意力机制的模块称为 MADB,该模块将多尺度特征融合、密集连接和注意力机制相结合,能有效地从低分辨率图像中学习高频信息。

在改进 ESRGAN 的生成网络结构中,首先,卷积神经网络通过逐层抽象的方式提取目标的特征,若卷积层输出特征上的像素点在输入上所映射区域(感受野)太小,则只能观察到局部特征,若感受野太大,则提取的特征可能包含过多的无用信息。而多尺度特征融合技术结合了不同尺度的特征,能够有效地提取物体及其周围环境的信息。其次,为更有效地利用每一层的特征,加强特征的传递,在卷积网络模型中引入密集连接,对每个卷积层之前的所有输入进行拼接,然后将拼接的特征传递给所有卷积层。密集连接结构的引入使得每一层都能够直接利

用损失函数的梯度信息和输入信息,在一定程度上缓解了梯度消失现象,有助于训练更深的网络。并且在该结构中,每层的输出特征都作为后面所有层的输入,这样的设计是为了防止特征传播过程中信息的丢失,使得特征的利用更加有效。同时密集连接结构的参数量相比于传统的卷积网络少,对过拟合有一定的抑制作用。最后,为进一步提高网络的特征表达能力,所提算法引入注意力机制,根据 MDB 提取的特征的重要性赋予不同的权重,增强了信息提取的能力,提高了网络性能和准确率。改进的 ESRGAN 的生成网络基本模块如图 2 所示。

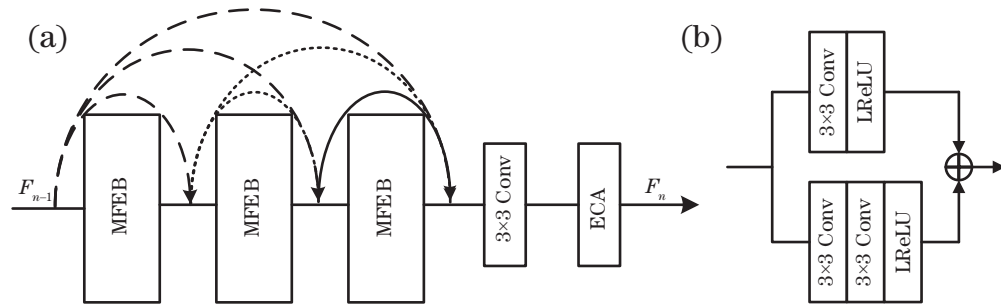


图 2 生成网络基本模块。(a) 带有注意力机制的 MDB(MADB); (b) 多尺度特征提取模块(MFEB)

Fig.2 Basic block of generative network. (a) MDB with attention mechanism (MADB); (b) multi-scale feature extraction block(MFEB)

参照图 2(a),假设 MADB 的输入特征为 F_0 ,MADB 的特征提取过程如下。首先, F_0 通过一个 MFEB,具体结构如图 2(b)所示,该模块采用并行多分支的多尺度特征融合结构,利用大小为 3×3 、卷积层数量分别为 1 和 2 的卷积层提取不同尺度的特征。其中,单个卷积层的感受野为 3,两个串联的

卷积层的感受野为 5。然后将提取的不同尺度的特征按照对应元素相加的方式融合,输出表示为

$$F_1 = f_{3 \times 3}(F_0) + f_{3 \times 3}[f_{3 \times 3}(F_0)], \quad (5)$$

式中: F_1 表示第一个多尺度特征提取模块的输出; $f_{3 \times 3}$ 表示卷积核大小为 3 的卷积运算。同理,后续两个多尺度特征提取模块的输出可表示为

$$F_2 = f_{3 \times 3}[C_a(F_1, F_0)] + f_{3 \times 3}\{f_{3 \times 3}[C_a(F_1, F_0)]\}, \quad (6)$$

$$F_3 = f_{3 \times 3}[C_a(F_2, F_1, F_0)] + f_{3 \times 3}\{f_{3 \times 3}[C_a(F_2, F_1, F_0)]\}, \quad (7)$$

式中: F_2 和 F_3 分别表示第二和第三个多尺度特征提取模块的输出; C_a 表示特征拼接操作。在多尺度特征提取模块的最后,使用卷积核大小为 3 的卷积层实现不同深度信息的融合,同时完成特征降维。

为进一步提升深度学习网络模型的表达能力,在融合特征之后添加注意力机制,通过自适应地调整各个通道的权重,关注重要特征来提升网络性能。然而注意力机制的引入在带来性能提升的同时会不可避免地增加模型的复杂度,考虑到模型的性能和复杂度,在融合特征之后引入计算复杂度低,同时能够保持高性能的通道注意力(ECA)机

制^[21]模块,结构如图 3 所示。该模块采用避免降维的局部跨通道交互策略。考虑高、宽和通道数量分别为 H 、 W 和 C 的输入特征,ECA 模块工作流程如下:首先对每个特征通道独立地使用全局平均池化(GAP),获取每个通道的背景信息,其输出张量的大小为 $1 \times 1 \times C$;然后通过大小为 k 的一维卷积完成跨通道信息的交互,并通过 Sigmoid 函数将交互后的通道信息映射为各特征通道对应的权值,其中一维卷积核的大小 k 表示跨通道交互的范围;最后将各通道与其权值相乘,获得调整权值后的特征。

在 MADB 模块的最后,多尺度特征经 ECA 模

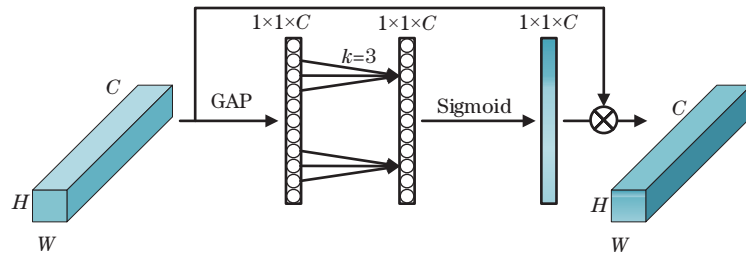


图 3 ECA 模块结构

Fig. 3 Structure of the ECA module

块调整后的输出可以表示为

$$F_{out} = ca \left\{ f_{3 \times 3} \left[C_a \left(F_3, F_2, F_1, F_0 \right) \right] \right\}, \quad (8)$$

式中: $ca(\cdot)$ 表示通过 ECA 通道注意力机制调整不同通道特征权重的操作。本文对生成网络结构的改进主要集中在深度残差信息提取模块。在生成网络中,没有采用 BN 策略。这是由于 BN 层在训练过程中使用均值和方差对输入进行归一化处理,并在推理过程中使用整个训练集估计的均值和方差,当训练集和测试集数据的统计特性相差较大时,BN 层的引入反而会限制网络的泛化能力,降低模型的鲁棒性。同时,去除 BN 层后,也减小了计算的复杂

度和内存的使用量。此外,采取了 ESRGAN 中多级残差模块的设计,并使用所设计的带有注意力机制的 MDB 代替 ESRGAN 中的 DB。其中,残差结构的引入简化了学习的过程,加强了梯度的传播,有利于训练层次更深的网络。图 4 为提出的多级残差 MADB 模块 (RR-MADB)。

最终,改进后的生成网络结构如图 5 所示。其中,RR-MADB 的输入为 64 通道,输出为 32 通道,在生成网络中共使用 18 个 RR-MADB 模块。此外,上采样层采用最近邻插值操作,同时生成网络中均使用大小为 3 的卷积核。

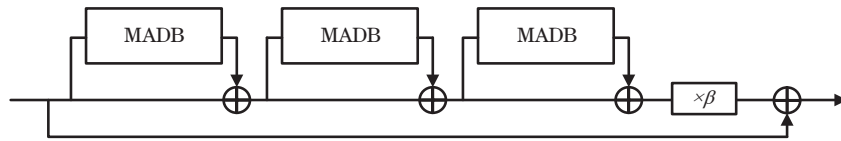


图 4 RR-MADB 结构

Fig. 4 RR-MADB structure

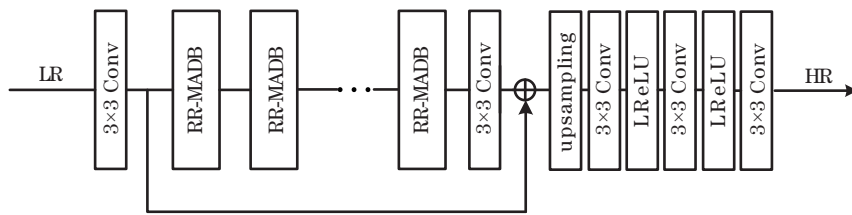


图 5 生成网络结构

Fig. 5 Generative network structure

2.2.2 判别网络结构

为合理地估计目标超分辨率图像比生成网络生成的超分辨率图像更加真实且自然的概率,在对抗训练中采用图 6 所示的判别网络结构,其中 c/s 表示卷积核数量/卷积步长。其中判别网络的输入为生成网络输出图像和对应的目标图像,而判别网络的输出为目标图像比生成图像更加真实且自然的概率。

对于判别网络结构的设计,遵循 DCGAN^[22] 设计的原则:使用 Leaky ReLU (LReLU) 激活函数和 BN 策略,同时避免使用池化操作。设计的判别网络共包含 8 个卷积层,所有卷积层均采用大小为 3×3 的卷积核。卷积层中的卷积核数量成倍增长,从 64 逐渐增加到 512。该判别网络采用卷积步长分别为 1 和 2 的卷积层交替构成,每当特征数量增加 1 倍时,使用步长为 2 的卷积层降低图像的分辨率。在

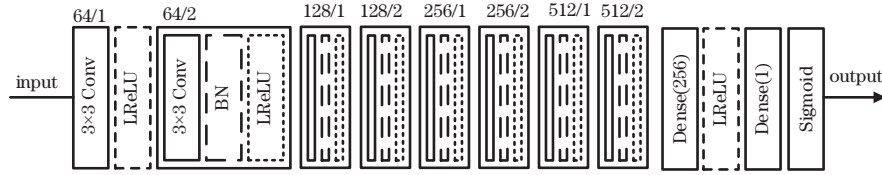


图 6 判别网络结构

Fig. 6 Discriminant network structure

卷积层获得的高级特征之后通过两个全连接层 (Dense) 和 Sigmoid 激活函数获得最终的概率。

2.2.3 损失函数选择

为提升生成网络重建图像的质量,所设计的生成网络使用多个损失函数线性组合的形式,公式为

$$L_G = L_{\text{per}} + \lambda_1 L_{L1} + \lambda_2 L_G^{\text{Ra}}, \quad (9)$$

式中: L_{per} 为感知损失函数; L_{L1} 为内容损失函数; L_G^{Ra} 为对抗损失函数; λ_1 和 λ_2 分别为各损失函数的加权系数。

感知损失函数对生成网络的性能至关重要,不同于基于像素差异的损失函数,感知损失通过惩罚预训练网络提取的高维特征的差异来提升生成网络输出图像的质量。选择预训练 VGG19 网络^[23]构造更加有效的感知损失函数。具体地,选择预训练 VGG19 第 5 个最大池化层之前的第 4 个卷积层提取的特征,通过最小化生成网络输出图像和目标图像特征的欧氏距离来提高输出图像的视觉感知质量。

感知损失函数表示为

$$L_{\text{per}} = E \left\{ \left\| \varphi[G(\mathbf{x})] - \varphi(\mathbf{y}) \right\|_2^2 \right\}, \quad (10)$$

式中: \mathbf{x} 为低分辨率图像; \mathbf{y} 为训练集中与 \mathbf{x} 相对应的超分辨率图像; $\varphi(\cdot)$ 表示通过 VGG19 网络提取的高级特征。

为减小生成网络输出图像与真实超分辨率图像信息像素差异,通过缩小生成图像与训练集真实图像的 L1 距离来提高像素级别的相似度,内容损失函数可表示为

$$L_{L1} = E \left[\left\| G(\mathbf{x}) - \mathbf{y} \right\|_1 \right]. \quad (11)$$

对抗损失与 ESRGAN 采用相同的相对论生成对抗网络原理,该判别器估计目标图像比由生成网络生成的图像更加真实且自然的概率。具体地,采用 RaGAN,从平均角度估计真实数据比合成数据更加真实且自然的概率。判别网络的损失函数定义为

$$L_D = -E \left\{ \log D_{\text{Ra}}[\mathbf{y}, G(\mathbf{x})] \right\} - E \left\{ \log \left\{ 1 - D_{\text{Ra}}[G(\mathbf{x}), \mathbf{y}] \right\} \right\}. \quad (12)$$

相应地,生成网络的对抗损失函数定义为

$$L_G = -E \left\{ \log \left\{ 1 - D_{\text{Ra}}[\mathbf{y}, G(\mathbf{x})] \right\} \right\} - E \left\{ \log \left\{ D_{\text{Ra}}[G(\mathbf{x}), \mathbf{y}] \right\} \right\}, \quad (13)$$

式中: $D_{\text{Ra}}(A, B)$ 表示输入 A 比输入 B 更加真实且自然的概率。

3 实验对比与结果分析

本实验主要针对低分辨率图像的 2 倍和 4 倍超分辨率重建。为验证所提算法的有效性,对所提算法与 ESRGAN 算法以及其他具有代表性的单幅图像超分辨率算法进行对比,并从定量和定性两方面对实验结果进行对比分析。

3.1 实验数据及训练

实验在 Linux 操作系统,深度学习框架 Pytorch 下实现。实验中的训练数据集均采用公开的超分辨率数据集 DIV2K^[24]。DIV2K 共有 1000 张 2000 分辨率的高质量图像,共有 800 张训练图像,100 张验

证图像和 100 测试图像。测试时使用 Set5^[25]、Set14^[26]、BSD100^[27]数据集。其中,Set5 包含 5 张动植物图像,Set14 包含 14 张动植物图像,包含比 Set5 中的图像更多的细节信息,BSD100 包含 100 张测试图像,包含丰富的自然景色和人工景物。这三个数据集元素丰富,常用于图像超分辨率任务的性能测试。在训练网络模型时,针对数据集数据较少的情况,使用数据增强技术,将图像随机水平或垂直翻转,然后随机裁剪图像获得 128×128 的高质量图像块。训练集和测试集中的低分辨率图像由随机裁剪的高质量图像块通过 Python 语言调用 Python Image Library 库函数实现双三次插值(Bicubic)获得。

训练过程分为两个阶段。首先,以生成器输出图像与超分辨率图像的 L1 距离为损失函数,训练

一个面向峰值信噪比 (PSNR) 的模型。初始学习率设置为 2×10^{-4} , 并且每进行 10^5 次迭代, 学习率衰减为原来的 0.5。然后, 使用训练好的面向 PSNR 的模型初始化生成器。生成器和判别器初始学习率设置为 1×10^{-4} , 并且每进行 10^5 次迭代, 学习率衰减为原来的 0.5。带有逐像素损失的预训练能够提高基于 GAN 方法生成的超分辨率图像的视觉效果, 这是因为该策略能够避免生成器陷入局部最优, 并且以面向 PSNR 的模型初始化生成器, 判别器将接收到质量相对较高的超分辨率图像而不是极端的质量较差的图像, 这有利于对抗训练将重点集中在细节纹理上。根据训练过程, 最终获得两个模型: 一个侧重减少像素差异, 关注客观评价指标 PSNR; 另一个侧重于减少感知差异, 旨在提高图像的视觉效果。

模型训练时使用 Adam 优化器, 其中, 一阶动量项 β_1 设置为 0.9, 二阶动量项 β_2 设置为 0.999。训练以小批量方式进行, batchsize 设置为 8。优化生成网络时, 损失函数的权重 λ_1 和 λ_2 分别设置为 1×10^{-2} 和 5×10^{-5} 。

此外, 主要针对低分辨率图像的 2 倍和 4 倍超

分辨率重建, 生成网络中采用 18 个 MADB 提取图像的深度残差信息, 整个生成网络的总参数量为 15.66×10^6 。

3.2 实验结果与分析

将所提改进后的算法在公开测试集 Set5、Set14、BSD100 上与双三次插值 (Bicubic)、FSRCNN^[3]、ESRGAN 进行对比。其中, ESRGAN 的生成网络中采用 23 个 RRDB, 总参数量为 16.70×10^6 。此外, 将训练阶段一中面向 PSNR 的模型作为对比之一, 记作 Ours-L1。

3.2.1 客观指标

选取 PSNR 和结构相似性 (SSIM) 作为图像的客观评价指标。PSNR 定义为能量峰值信号与清晰图像和噪声图像能量差均值之比, 公式为

$$R_{\text{PSN}} = 10 \log_{10} \frac{M_{\text{max}}^2}{E_{\text{MS}}}, \quad (14)$$

式中: M_{max} 表示能量峰值信号, 对于数字图像, 取值为 255; E_{MS} 表示重建图像与目标超分辨率图像的均方误差。PSNR 通过计算两幅图像对应像素点间的误差来评价重建图像的质量, 在超分辨率任务中使用最为广泛, 该指标的值越大, 说明图像的失真越小。

$$S_{\text{SSIM}} = \left[(2\mu_y\mu_x + c_1)(2\sigma_{yx} + c_2) \right] / \left[(\mu_y^2 + \mu_x^2 + c_1)(\sigma_{yx}^2 + \sigma_x^2 + c_2) \right], \quad (15)$$

式中: μ_x, μ_y 分别表示生成图像 \mathbf{x} 和目标图像的均值; σ_x, σ_y 分别表示 \mathbf{x} 和 \mathbf{y} 的方差; σ_{yx} 表示 \mathbf{x} 与 \mathbf{y} 的协方差; c_1 和 c_2 为避免分母为零的常量。

表 1 和表 2 列出了各超分辨率算法在公开测试集 Set5、Set14、BSD100 上进行 2 倍和 4 倍超分辨率重建的客观评价结果。从表 1、2 可以发现, 所提算法所得的客观指标不是最高, 这是因为 Ours 是基

于对抗损失和感知损失进行训练的, 而不是以 PSNR、SSIM 为导向进行训练的, 主要关注图像的细节恢复, 并且 PSNR 和 SSIM 不代表图像的真实视觉效果^[17]。Ours-L1 是由像素级损失训练的面向 PSNR 的模型, 所得的客观指标虽均为最大, 但重建图像的视觉效果容易产生模糊, 远不如 Ours 算法。

表 1 各算法 2 倍和 4 倍重建结果 PSNR 比较

Table 1 Comparison of PSNR values of the 2× and 4× reconstructed results of each algorithm

Dataset	Scale	Bicubic	FSRCNN	ESRGAN	Ours-L1	Ours
Set5	2×	31.787	33.404	32.637	35.758	33.193
	4×	26.690	27.569	27.928	30.347	28.629
Set14	2×	28.298	29.487	28.787	31.387	29.481
	4×	24.235	24.833	24.513	26.614	24.786
BSD100	2×	26.725	26.763	25.801	26.964	25.953
	4×	23.701	23.730	22.478	23.872	22.836

表 2 各算法 2 倍和 4 倍重建结果 SSIM 比较

Table 2 Comparison of SSIM values of the 2× and 4× reconstructed results of each algorithm

Dataset	Scale	Bicubic	FSRCNN	ESRGAN	Ours-L1	Ours
Set5	2×	0.908	0.922	0.904	0.943	0.906

续表

Dataset	Scale	Bicubic	FSRCNN	ESRGAN	Ours-L1	Ours
Set14	4×	0.769	0.785	0.799	0.863	0.818
	2×	0.809	0.865	0.837	0.892	0.845
	4×	0.660	0.683	0.660	0.745	0.670
BSD100	2×	0.736	0.755	0.725	0.759	0.730
	4×	0.572	0.580	0.525	0.590	0.541

3.2.2 主观效果

为了更加直观地感受所提改进的ESRGAN算法重建图像的视觉效果和以像素级差异为损失函数的缺陷,图7和图8分别展示了各超分辨率算法针对

Set5和Set14数据集部分图像2倍和4倍超分辨率重建的结果。其中图7(a)和图8(a)中的方框分别表示对其进行2倍和4倍超分辨率重建的区域,图7(b)和图8(b)表示作为参考的高分辨率图像。

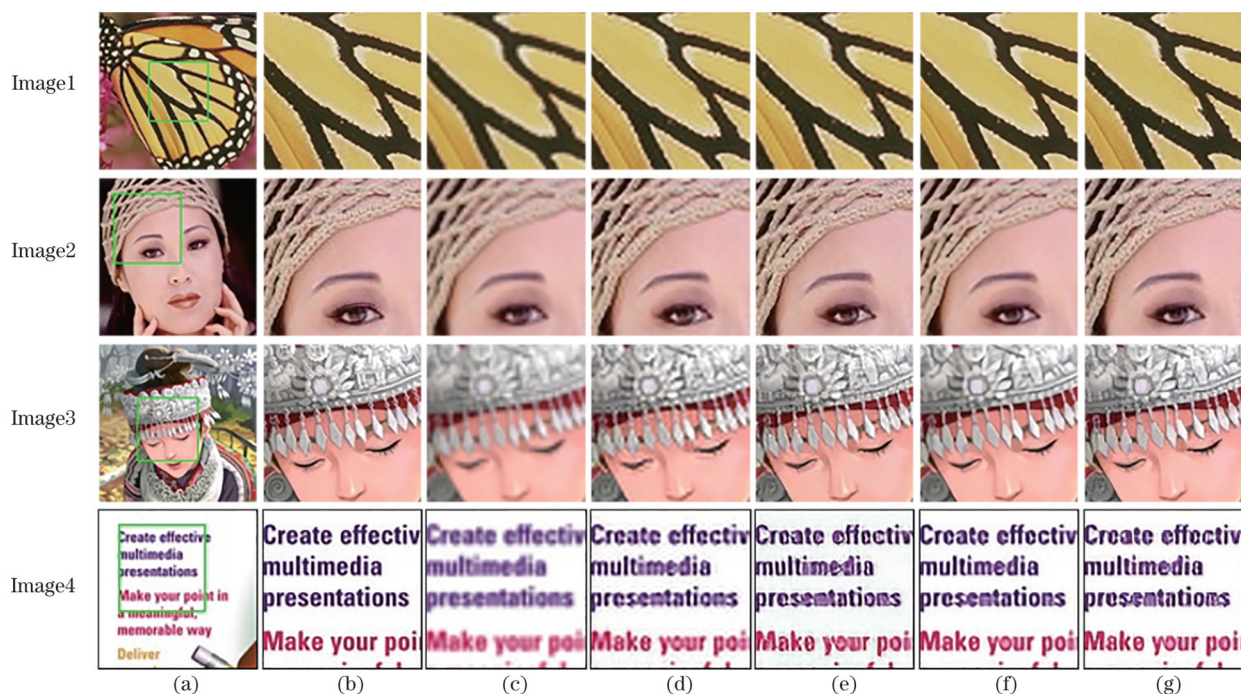


图7 各算法2倍超分辨率重建结果。(a)原图;(b)高分辨率图像;(c)Bicubic;(d)FSRCNN;(e)ESRGAN;(f)Ours-L1;(g)Ours

Fig. 7 2× super-resolution reconstruction results of each algorithm. (a) Original images; (b) high resolution images; (c) Bicubic; (d) FSRCNN; (e) ESRGAN; (f) Ours-L1; (g) Ours

从图7、8中可以观察到:Bicubic算法得到的超分辨率图像伴有一定程度的模糊现象,并且模糊程度随着放大倍率的增大而加深,给观察者呈现较差的视觉效果;FSRCNN算法重建图像的质量整体上优于Bicubic,该算法以基于像素差异的损失函数优化网络模型,虽然在部分测试结果中获得了较高的客观指标,但重建后得到的超分辨率图像在边缘区域存在模糊,且这种现象在4倍重建的结果中更为明显,如图8中Image1的帽子和Image3花蕊的细节纹理几乎全部丢失;相比FSRCNN,同样基于像素级损失的Ours-L1算法采

用了学习能力更加强大的特征提取网络结构,获得重建质量相对较高的结果,然而以像素级损失优化的模型恢复细节的能力较差,Ours-L1算法重建后的图像存在细节模糊问题,如图7中Image2的帽子和Image4中的字母,尤其在图8中Image1和Image4的4倍超分辨率结果中,重建图像边缘区域显得尤为平滑;区别于FSRCNN和Ours-L1,ESRGAN和Ours以对抗损失和感知损失相结合的优化策略恢复了比较清晰的边缘,重建后图像的细节也更加丰富,如图7中Image1蝴蝶翅膀纹理区域,Ours重建的纹理比较清晰,而ESRGAN

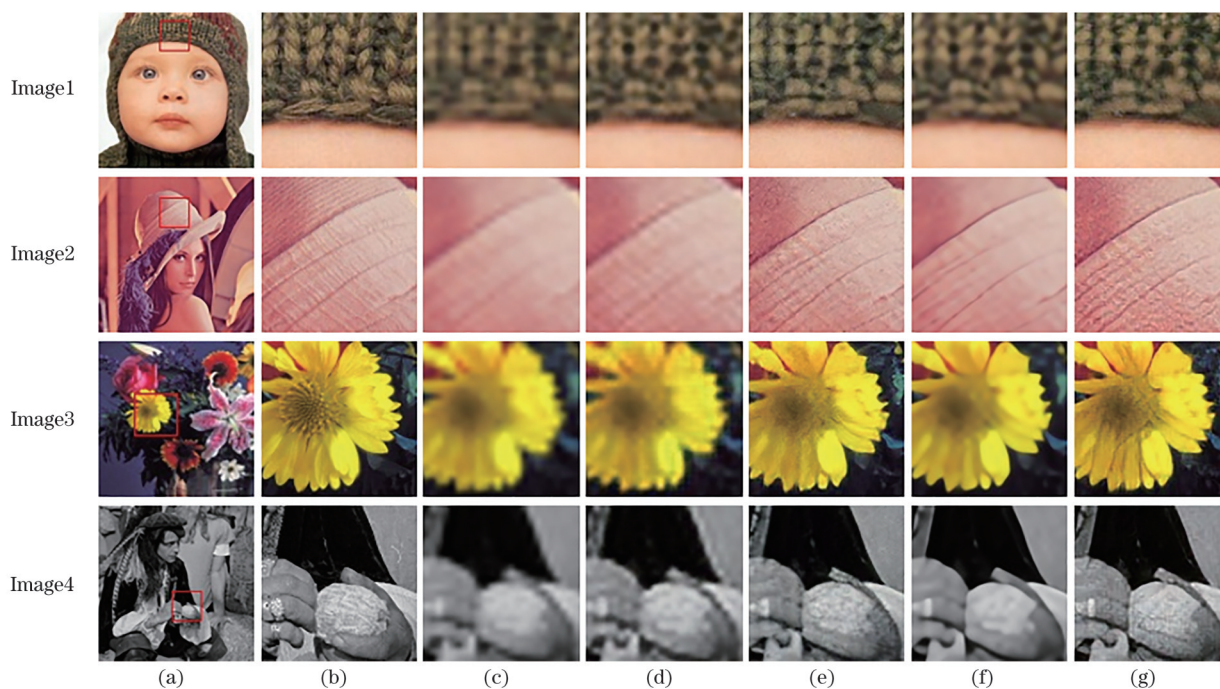


图 8 各算法 4 倍超分辨率重建结果。(a) 原图; (b) 高分辨率图像; (c) Bicubic; (d) FSRCNN; (e) ESRGAN; (f) Ours-L1; (g) Ours

Fig. 8 $4\times$ super-resolution reconstruction results of each algorithm. (a) Original images; (b) high resolution images; (c) Bicubic; (d) FSRCNN; (e) ESRGAN; (f) Ours-L1; (g) Ours

在纹理区域存在伪影,在图 8 中 Image3 和 Image4 的重建结果中,Ours 重建的结果相对于 ESRGAN 获得了更加清晰的边缘。

经上述分析,所提改进过后的算法在公开测试集上获得了最好的重建质量,同时 ESRGAN 也获得了不错的重建效果。但值得注意的是,相对于 ESRGAN,所提改进后的生成网络使用了更加少的参数量。其中,ESRGAN 生成网络的参数量约为 16.70×10^6 ,而所提算法约为 15.66×10^6 。

4 结 论

在 ESRGAN 的基础上,提出了结合注意力机制与多尺度密集连接的图像超分辨率重建算法,重点在于解决低分辨率图像细节不清晰的问题。在 ESRGAN 的生成网络模型的深层信息提取模块中引入多尺度特征融合、密集连接和注意力机制,能有效地从低分辨率图像中学习高频信息,提升学习效率。将所提算法与其他典型的超分辨率算法进行对比,实验结果显示,所提算法得到的重构图像纹理细节更清晰,大大提升了图像的视觉效果,图像重建客观评价指标 (PSNR、SSIM) 相比于 ESRGAN 也有所提高。

参 考 文 献

- [1] Zhu S Y, Zeng B, Liu G H, et al. Image interpolation based on non-local geometric similarities [C]//2015 IEEE International Conference on Multimedia and Expo (ICME), June 29-July 3, 2015, Turin, Italy. New York: IEEE Press, 2015: 16213035.
- [2] Papyan V, Elad M. Multi-scale patch-based image restoration[J]. IEEE Transactions on Image Processing, 2016, 25(1): 249-261.
- [3] Dong C, Loy C C, Tang X O. Accelerating the super-resolution convolutional neural network[M]//Leibe B, Matas J, Sebe N, et al. Computer vision-ECCV 2016. Lecture notes in computer science. Cham: Springer, 2016, 9906: 391-407.
- [4] Dong C, Loy C C, He K M, 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]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 1637-1645.

- [6] Lai W S, Huang J B, Ahuja N, et al. Deep Laplacian pyramid networks for fast and accurate super-resolution[C] //2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 5835-5843.
- [7] Ledig C, Theis L, Huszár F, et al. Photo-realistic single image super-resolution using a generative adversarial network[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 105-114.
- [8] Hu S Y, Wang G D, Zhao Y, et al. Image super-resolution network based on dense connection and squeeze module[J]. *Laser & Optoelectronics Progress*, 2019, 56(20): 201005.
胡诗语, 王国栋, 赵毅, 等. 基于密集连接与激励模块的图像超分辨率网络[J]. *激光与光电子学进展*, 2019, 56(20): 201005.
- [9] Wang X T, Yu K, Wu S X, et al. ESRGAN: enhanced super-resolution generative adversarial networks[M] //Leal-Taixé L, Roth S. *Computer vision-ECCV 2018 workshops. Lecture notes in computer science*. Cham: Springer, 2019, 11133: 63-79.
- [10] Hu X Y, Liu X J, Wang Z C, et al. RTSRGAN: real-time super-resolution generative adversarial networks[C]//2019 Seventh International Conference on Advanced Cloud and Big Data (CBD), September 21-22, 2019, Suzhou, China. New York: IEEE Press, 2019: 321-326.
- [11] Dou X Y, Li C Y, Shi Q, et al. Super-resolution for hyperspectral remote sensing images based on the 3D attention-SRGAN network[J]. *Remote Sensing*, 2020, 12(7): 1204.
- [12] Zhang Y L, Tian Y P, Kong Y, et al. Residual dense network for image restoration[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2021, 43(7): 2480-2495.
- [13] Shao M W, Zhang W T, Zuo W M, et al. Multi-scale generative adversarial inpainting network based on cross-layer attention transfer mechanism[J]. *Knowledge-Based Systems*, 2020, 196: 105778.
- [14] Li Y H, Mu X, Zhu Y L, et al. Super resolution image restoration and reconstruction of deep generative countermeasure network[J]. *Journal of Xi'an University of Technology*, 2021, 35(4): 1-8.
李云红, 穆兴, 朱耀麟, 等. 深度生成式对抗网络的超分辨率图像修复与重建[J]. *西安工程大学学报*, 2021, 35(4): 1-8.
- [15] Peng Y F, Zhang P J, Gao Y, et al. Attention fusion generative adversarial network for single-image super-resolution reconstruction[J]. *Laser & Optoelectronics Progress*, 2021, 58(20): 2010012.
彭晏飞, 张平甲, 高艺, 等. 融合注意力的生成式对抗网络单图像超分辨率重建[J]. *激光与光电子学进展*: 58(20): 2010012.
- [16] Chen Z H, Wu H B, Pei H D, et al. Image super-resolution reconstruction method based on self-attention deep network[J]. *Laser & Optoelectronics Progress*, 2021, 58(4): 0410013.
陈子涵, 吴浩博, 裴浩东, 等. 基于自注意力深度网络的图像超分辨率重建方法[J]. *激光与光电子学进展*, 2021, 58(4): 0410013.
- [17] Zha T B, Luo L, Yang K, et al. Image reconstruction algorithm based on improved super-resolution generative adversarial network[J]. *Laser & Optoelectronics Progress*, 2021, 58(8): 0810005.
查体博, 罗林, 杨凯, 等. 基于改进超分辨率生成对抗网络的图像重建算法[J]. *激光与光电子学进展*, 2021, 58(8): 0810005.
- [18] Lim B, Son S, Kim H, et al. Enhanced deep residual networks for single image super-resolution [C]//2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 1132-1140.
- [19] Nah S, Kim T H, Lee K M. Deep multi-scale convolutional neural network for dynamic scene deblurring[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 257-265.
- [20] Jolicœur-Martineau A. The relativistic discriminator: a key element missing from standard GAN[EB/OL]. (2018-07-02)[2020-07-02]. <https://arxiv.org/abs/1807.00734>.
- [21] Wang Q L, Wu B G, Zhu P F, et al. ECA-net: efficient channel attention for deep convolutional neural networks[C] //2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 13-19, 2020, Seattle, WA, USA. New York: IEEE Press, 2020: 11531-11539.
- [22] Radford A, Metz L, Chintala, et al. Unsupervised representation learning with deep convolutional generative adversarial networks[C]//4th International Conference on Learning Representations, ICLR 2016, May 2-4, 2016, San Juan, Puerto Rico. [S.l.:

- s.n.], 2016.
- [23] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition[C]//3rd International Conference on Learning Representations, {ICLR} 2015, May 7-9, 2015, San Diego, CA, USA. [S.l.: s.n.], 2015.
- [24] Agustsson E, Timofte R. NTIRE 2017 challenge on single image super-resolution: dataset and study[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 1122-1131.
- [25] Bevilacqua M, Roumy A, Guillemot C, et al. Low-complexity single-image super-resolution based on nonnegative neighbor embedding[C]//Proceedings of the British Machine Vision Conference 2012, September 3-7, 2012, Surrey, UK. London: British Machine Vision Association, 2012: 135.1-135.10.
- [26] Zeyde R, Elad M, Protter M. On single image scale-up using sparse-representations[M]//Boissonnat J D, Chenin P, Cohen A, et al. Curves and surfaces. Lecture notes in computer science. Heidelberg: Springer, 2012, 6920: 711-730.
- [27] Martin D, Fowlkes C, Tal D, et al. A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics[C] //Proceedings Eighth IEEE International Conference on Computer Vision, July 7-14, 2001, Vancouver, BC, Canada. New York: IEEE Press, 2001: 416-423.