

基于多尺度注意力残差网络的图像超分辨率重建^①



李俊珠¹, 郑 华^{1,2,3,4}, 雷 帅¹, 陈清俊¹, 潘 浩¹

¹(福建师范大学 光电与信息工程学院, 福州 350007)

²(福建师范大学 医学光电科学与技术教育部重点实验室, 福州 350007)

³(福建师范大学 福建省光子技术重点实验室, 福州 350007)

⁴(福建师范大学 福建省光电传感应用工程技术研究中心, 福州 350007)

通信作者: 郑 华, E-mail: hzheng@fjnu.edu.cn

摘 要: 数字图像在传递信息中起着重要的作用, 图像超分辨率技术能丰富图像的细节信息. 针对许多网络对低分辨率图像的有效特征复用不足和参数量过大的问题, 本文结合不同大小的卷积核以及注意力残差机制构建图像超分辨率网络, 用 3 个有差别尺度的卷积层来提取图像的特征, 其中第 2 和第 3 层用小卷积核替代大的卷积核, 对 3 层卷积融合之后引入注意力机制, 最后用传统的 Bicubic 插值直接给网络提供低频信息. 在减小参数量和减轻梯度消失的同时, 让有效的高频信息得到更大的权重且能增强网络之间的非线性表达能力, 这有利于网络训练的迭代收敛. 实验结果表明, 基于多尺度注意力残差网络能够在一定程度上增强图像的重建能力.

关键词: 图像超分辨率重建; 不同大小的卷积核; 注意力残差网络; 卷积神经网络

引用格式: 李俊珠, 郑华, 雷帅, 陈清俊, 潘浩. 基于多尺度注意力残差网络的图像超分辨率重建. 计算机系统应用, 2022, 31(5): 324-330. <http://www.c-s-a.org.cn/1003-3254/8465.html>

Image Super-resolution Reconstruction Based on Multi-scale Attention Residual Network

LI Jun-Zhu¹, ZHENG Hua^{1,2,3,4}, LEI Shuai¹, CHEN Qing-Jun¹, PAN Hao¹

¹(College of Photonic and Electronic Engineering, Fujian Normal University, Fuzhou 350007, China)

²(Key Laboratory of Optoelectronic Science and Technology for Medicine (Ministry of Education), Fujian Normal University, Fuzhou 350007, China)

³(Fujian Provincial Key Laboratory of Photonics Technology, Fujian Normal University, Fuzhou 350007, China)

⁴(Fujian Provincial Engineering Technology Research Center of Photoelectric Sensing Application, Fujian Normal University, Fuzhou 350007, China)

Abstract: Digital images play an important role in information transmission, and image super-resolution technology can enrich image details. To address the problems of insufficient effective feature reuse of low-resolution images and excessive parameters in many networks, this study combines convolution kernels of different sizes and attention residual mechanism to construct the image super-resolution network. Three convolution layers of different scales are used to extract the image features, of which the second and third layers replace the large convolution kernels with small ones, and after the three-layer convolution fusion, the attention mechanism is introduced. Finally, the traditional Bicubic interpolation is used to directly provide low-frequency information for the network. By doing this, while reducing the number of parameters and mitigating the disappearance of gradients, the proposed network can make the effective high-frequency information gain greater weights and can enhance the nonlinear expression ability between the networks, which is conducive to the iterative convergence of network training. Experimental results show that the proposed network can enhance the image reconstruction ability to a certain extent.

Key words: image super-resolution reconstruction; convolution kernel of different sizes; attention residual network;

① 基金项目: 福建省自然科学基金 (2017J01744)

收稿时间: 2021-07-20; 修改时间: 2021-08-18, 2021-08-24; 采用时间: 2021-08-31; csa 在线出版时间: 2022-02-21

convolutional neural network (CNN)

1 引言

图像作为一种重要的信息传递方式, 图像的恢复和复原一直是人们关注的重点领域. 图像超分辨率重建技术在移动数据的传输、卫星遥感成像、小目标检测以及视频监控等领域都有着广泛的应用和重要的研究意义. 在现实生活中, 由于硬件设备条件的原因会导致拍摄出许多低像素的图像. 例如移动数据传输的过程中, 由于设备本身成像的限制以及传输速度的要求, 低分辨率图像的传输可以节约网络移动通信的带宽, 再由用户自己决定是否进行图像超分辨率的重建, 如此既满足了用户的需求又加快了图像的传输速率.

随着现代技术的发展, 越来越多的人参与图像超分辨率的研究. 在图像的超分辨率重建算法受到广泛地关注之后, 陆续有研究人员提出了基于插值的超分辨率重建方法, 基于重建的超分辨率重建方法, 基于深度学习的超分辨率重建方法. 其中基于插值的方法虽然重建速度快且算法比较简单, 但是容易丢失高频信息, 适合小倍数图像的重建; 基于重建的方法相对前者要好, 但是从常规来说, 加深网络的深度的确能够提高重建精度, 但是随着卷积神经网络深度的增加, 也面临着计算复杂度加深和内存消耗的问题. 单图像超分辨率重建的目的就是把分辨率低的图像通过一些技术方法恢复出与之相对应的高分辨率图像, 这是个没有唯一解的问题. 如何在不增加网络参数量的同时还能提高重建图像的精度, 这是很多研究者都在思索的问题^[1-5].

本文所提出的图像超分辨率算法, 使用多尺度残差模块充分利用低分辨率图像本身的特征, 再分别用两个、3个 3×3 的卷积替代 5×5 、 7×7 的卷积来减小参数量. 在模块中也引入了通道注意力机制, 它可以自适应地学习特征权重, 赋予高频信息更大的权重, 使网络更注重传递有效的特征信息, 增强网络的学习能力. 通过实验证明, 本文结合不同大小的卷积核以及注意力残差机制构建的单图像超分辨率重建网络, 在精度上和视觉效果上都取得了一定的提升.

2 基于深度学习的图像超分辨率技术

2.1 深度学习在图像超分辨率的应用

目前, 深度学习的快速发展使得图像超分辨率领域得到了重大的突破. SRCNN^[1]运用卷积神经网络来

学习低分辨率到高分辨率图像之间的映射关系, 而FSRCNN^[2]在网络重建部分加入反卷积层的方式来替代Bicubic插值下采样, 有效的提升了网络的训练速度. 但由于FSRCNN算法卷积层数较少, 且相邻的卷积之间缺乏相关性, 重建效果不是很理想. VDSR^[3]通过对网络层数的加深, 增大了网络层的感受野, 还利用残差学习的方式加速网络收敛的速度. SRDenseNet^[4]的稠密块结构将每一层的特征都传递给后面的每一层, 特征的重复利用能够减轻梯度消失且加强了特征的传播. RCAN^[5]网络将通道注意力机制(channel attention, CA)^[6]引入, 让网络自适应地学习特征信息的权重, 赋予高频有效信息更大的权重, 提升了神经网络的表达能力. MemNet^[7]利用递归单元的内存块建立长期记忆, 以及使用门控单元控制不同的网络模块在输出的权重, 自适应地形成长期持续的记忆. IDN^[8]提出的信息蒸馏块包含了增强单元和压缩单元, 对低分辨率图像的轮廓增强输入并将特征像素映射压缩. MSRN^[9]使用3种不同大小的卷积核, 对初始输入的低分辨率图像进行反复的特征信息提取, 提高网络重建的性能和加快网络收敛的速度. MWRN^[10]使用多窗口残差网络改变了卷积核大小, 使得网络性能和参数量也有了一定的提升. TTSR^[11]鼓励低分辨率图像和参考图像进行联合学习, 通过CA机制调整通道的特征来传递图像的纹理特征.

2.2 通道注意力机制

由于卷积神经网络有着十分强大的非线性表达能力, 它在图像超分辨率领域近10年来得到广泛的应用. 为了使得网络给予高频有效信息更多关注, 忽略与网络数据中无关的信息, 注意力机制重新被应用在计算机视觉领域. 自2018年Hu等人在SENet^[6]中提出CA机制之后, 就被广泛应用于深度学习的网络中, 虽然增加了一些参数量, 但是性能却有了比较大的提升.

图1左侧是由 C 个通道, 大小均为 $H\times W$ 的特征图组成, 假设特征图组 $X = [X_1, X_2, \dots, X_i, \dots, X_C]$. 在压缩(squeeze)操作中, 对每一通道输入首先进行全局平均池化(global average pooling, GAP), 每一个通道得到一个特征标量, 那么 C 个通道进行操作之后就得到 C 个特征标量, 如式(1)所示^[6].

$$Z_C = H_{GAP}(x_C) = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W x_C(i, j) \quad (1)$$

其中, Z 为通道描述符, Z_C 为通道描述符的第 C 个元素, H_{GAP} 为全局平均池化, $x_C(i, j)$ 表示输入位置为 (i, j) 的特征图.

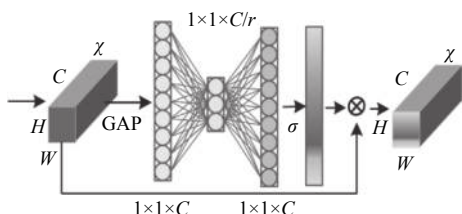


图1 通道注意力机制示意图

经过 GAP 操作之后引入门结构机制, 使用两个全连接层为每个通道形成相应的权重系数, 先将 C 个通道压缩成 C/r 个通道, 再恢复成 C 个通道, 两个全连接层中的瓶颈结构可以降低模型的复杂度, 如式 (2)^[6]:

$$S = F_{ex}(z, w) = \delta(g(z, w)) = \sigma W_u \sigma(W_D Z) \quad (2)$$

其中, F_{ex} 表示参数 w 形成权重的操作, S 为统计量, δ 表示 Sigmoid 函数, σ 表示激活函数 ReLU, W_D 代表下采样的权重值, W_u 代表上采样的权重值.

对特征进行激励 (excitation) 操作之后再行特征的融合操作, 用特征权重乘以输入特征得到新的特征 X' , 如式 (3) 所示^[6]:

$$X' = S_C \times X_C \quad (3)$$

其中, S_C 是对应 X_C 的权重, X_C 表示第 C 个特征图.

3 基于多尺度注意力残差网络的图像超分辨率算法

由于网络深度和层数增加可以提升图像重建的效果, 但是网络深度的增加会使得网络难以训练且很难收敛. 针对特征复用不足且参数量过大的问题, 本文提出基于多尺度注意力残差的图像超分辨率重建算法.

3.1 网络框架

如图 2 所示, 多尺度注意力残差网络的单图像超分辨率主要包括浅层特征提取、深层特征提取、特征融合、特征重建 4 个方面. 首先经过 Bicubic 插值得到低分辨率图像 (LR), 在我们对 I^{LR} 进行浅层的特征提取时, 要先经过一个 3×3 的卷积层提取浅层的图像特征, 再使用激活函数 ReLU 增强网络的非线性表达. 浅层

特征提取如式 (4) 所示^[1]:

$$F_0 = H(I^{LR}) = \sigma(W_{3 \times 3}^1 * I^{LR} + b^1) \quad (4)$$

其中, H 代表浅特征的提取, σ 代表激活函数, $W_{3 \times 3}^l$, b^l 代表的是第 l 个 3×3 卷积的权重和偏移量. 在经过 n 个多尺度注意力残差模块提取特征之后, 得到深层特征 F_{LR} , 如式 (5) 所示^[1]:

$$F_{LR} = w * [M_1, M_2, M_3, \dots, M_n] + b \quad (5)$$

其中, $[M_1, M_2, M_3, \dots, M_n]$ 代表特征连接操作, F_{LR} 代表深层特征的提取. 接着进行特征重建, 先经过一个卷积层, 去掉分层特征融合中冗余的部分, 再经过一个亚像素卷积层, 对像素重新洗牌, 将图片扩大 r 倍, 达到所需的尺寸. 最后经过 3×3 的卷积层进行 I^{LR} 到 I^{HR} 之间的映射, 过程如式 (6) 所示^[1]:

$$I^{HR} = \sigma [W_{3 \times 3}^a \times PS(F_{LR}) + b^a] \quad (6)$$

其中, $PS(x)$ 表示亚像素卷积函数, a 为网络中第 a 个卷积层, I^{HR} 代表重建出来的高清图像, 再将原始输入的低分辨率图像进行 Bicubic 插值得到的图像添加到 I^{HR} 上, 最终生成高分辨率效果图 I^{SR} .

3.2 多尺度注意力残差模块

本文提出了多尺度注意力残差模块 (multi-scale attention residual block, MCRB), 每个模块在残差块的基础上使用 3 个通道自适应检测尺度不同的特征图像, 再对 3 层卷积层提取的特征进行特征融合, 尽可能的提取初始的特征图像, 加强了特征复用. 模块的卷积后面加入了激活函数 ReLU, 以增强每个模块的非线性表达能力. 通道注意力机制的引入使得网络对高频有效信息赋予更大的权重, 对低频无效的信息进行忽略.

在不降低模型性能的前提下, 用两个 3×3 的卷积核代替一个 5×5 的卷积核, 3 个 3×3 的卷积核代替一个 7×7 的卷积. 如图 3 所示, 一个像素经过两个 3×3 的卷积核之后会产生 5×5 个与之相关联的像素点. 用小尺寸卷积核堆叠的卷积层来替代大卷积核的卷积层, 不仅能够维持感受野大小不变, 且每个卷积层中含有的激活函数也能增强网络的非线性表达能力.

如图 4 所示, 在 MCRB 模块中, 让输入特征通过 3 层有差别大小的卷积层来进行特征提取, 之后再把 3 层特征融合, 并通过一个 1×1 的卷积核和通道注意力机制分别进行特征过滤和不同权重的赋予, 最后在该过程中加入跳跃连接. 它的具体表示方法如式 (7):

$$M_{m+1} = F_{CA}\{W_{1 \times 1}^l \times [Q_1, Q_2, Q_3] + b^l\} + M_m \quad (7)$$

其中, $[Q_1, Q_2, Q_3]$ 表示3层卷积层融合操作, 变量 $W_{1 \times 1}^l$

和 b^l 表示在同一层的权重张量和偏移张量, F_{CA} 表示对所提取的特征进行特征注意力机制操作, 而 M_m 表示上一个模块的输出结果.

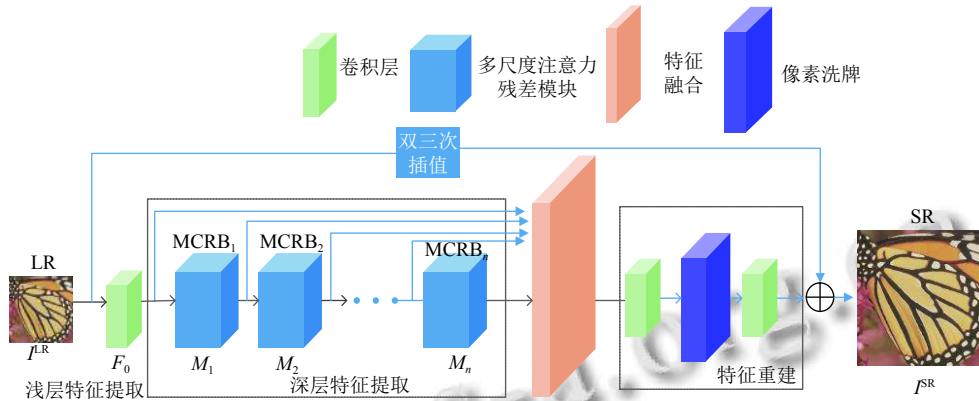


图2 MCRN 模型网络框架

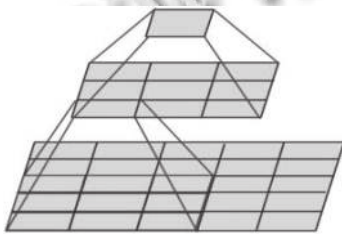


图3 两个3×3代替5×5示意图

在网络中, 如果只是一味的加深网络的深度和宽度, 会使得网络重建的性能有一定的提升, 但同时也增加了大量的参数量, 使得训练时间大幅增加, 模型难以收敛. 本文提出的MCRB模块, 使用3级卷积层对特征图进行特征提取, 充分复用原始特征和加强网络的非线性表达能力. 再对每一层卷积提取的信息进行特征融合, 引入注意力机制对网络的特征权重进行赋值, 增强高频有效信息的传播, 过滤掉低频无效的信息.

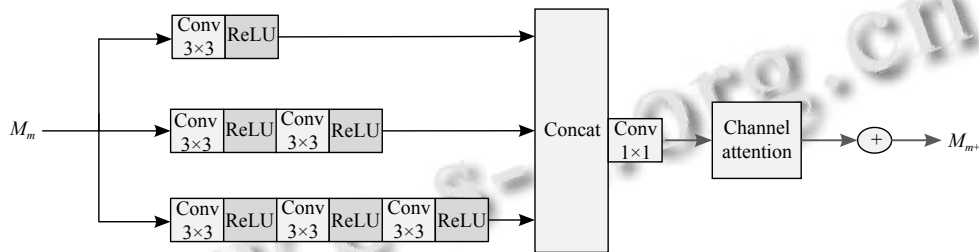


图4 MCRB 模块工作原理

4 实验

4.1 实验环境和数据集

本文的训练集是采用图像超分辨率算法常用高质量的图像数据集DIV2K, 它包含1 000张高质量图像数据集中的前800张作为我们实验的训练集, 另外200张分别作为验证集和测试集. 实验测试集还包括Set5^[12]、Set14^[13]、B100^[14]、Urban100^[15]、Manga109^[16].

实验环境基于Ubuntu 16.04系统, Python 3.8深度学习开源框架PyTorch 1.8.1、NVIDIA 1080Ti、CUDA Version 10.1上完成. 在训练的每个批次中, 随

机地选择16个LR色块, 它的输入图像剪裁块的大小为48×48, 采用损失函数、ADAM优化器^[17]进行优化. 我们将学习率的初始值设置为 $Lr=1E-4$, 每200个epoch就减半, 模型训练完成1 000个epoch大概需要36 h.

4.2 评价指标

对于重建的效果数值上, 本文主要采用峰值信噪比(PSNR)和结构相似度(SSIM)^[18]来作为评估网络重建的性能.

(1) PSNR

PSNR 主要是通过误差敏感对图像进行评估,若一个 HR 图像的大小为 $m \times n$,重建得到的图像为 S ,则 PSNR 的定义如式 (8):

$$PSNR = 10 \lg \left(\frac{Max_H^2}{\frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [S(i, j) - X(i, j)]^2} \right) \quad (8)$$

其中, Max_H^2 一般取 255, dB 为 PSNR 的单位, PSNR 的数值一般是越大代表图像效果越好。

(2) SSIM

结构相似度是通过输入的 LR 和 HR 两张图片进行结构的对比,取值范围是 [0, 1],公式如式 (9):

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + c_1)(\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \quad (9)$$

其中, x 代表没有经过压缩原始无失真的图像, y 代表经过重建恢复出来的图像, μ_x 、 μ_y 、 σ_x 、 σ_y 分别代表的

是 x 、 y 的平均灰度值和标准差, σ_{xy} 是 x 和 y 的协方差, c_1 、 c_2 都为常数。

4.3 实验结果分析

在本文实验中,我们与目前一些主流的、效果较好的网络模型进行对比,包括 SRCNN、LapSRN^[19]、DRRN^[20]、MSRN. 分别在数据集 Set5、Set14、B100、Urban100、Manga109 上进行测试. 在测试数据时,我们将原始的 RGB 通道的图像转换到 $YCbCr$ 通道的色彩空间上,对 Y 通道进行 PSNR 和 SSIM 的数值计算。

表 1 是实验测试的结果,从表中可以看出在放大倍数为 4 的时候测试集 Manga109 PSNR 的值为 30.98 dB,相比 MSRN 增加 0.41 dB. 从客观的评价指标上可以看出,本文提出的方法优于目前一些较好的网络. 本文在 MSRN 的基础上增加了一条 Bicubic 插值上采样的路径,用小卷积核 3×3 代替较大的 5×5 和 7×7 卷积核,并引入了注意力机制,在控制网络参数量的同时获取更多的特征,从而使得 SR 性能提升。

表 1 不同方法的 PSNR 和 SSIM 的对比

方法	Scale	Set5		Set14		B100		Urban100		Manga109	
		PSNR (dB)	SSIM	PSNR (dB)	SSIM	PSNR (dB)	SSIM	PSNR (dB)	SSIM	PSNR (dB)	SSIM
Bicubic	×2	33.66	0.929 9	30.24	0.868 8	29.56	0.843 1	26.88	0.840 3	30.30	0.933 9
SRCNN	×2	36.66	0.952 4	32.42	0.906 3	31.36	0.887 9	29.50	0.894 6	35.74	0.966 1
VDSR	×2	37.53	0.958 7	33.03	0.912 4	31.90	0.896 0	30.76	0.914 0	37.22	0.972 9
LapSRN	×2	37.52	0.959 0	33.08	0.913 0	31.80	0.895 0	30.41	0.910 0	37.27	0.974 0
DRRN	×2	37.74	0.959 1	33.23	0.913 6	32.05	0.897 3	31.23	0.918 8	37.92	0.976 0
MSRN	×2	38.08	0.960 7	33.70	0.918 6	32.23	0.900 2	32.29	0.930 3	38.69	0.977 2
本文	×2	38.31	0.962 1	34.02	0.921 4	32.37	0.902 6	32.75	0.934 3	39.05	0.997 7
Bicubic	×3	30.39	0.868 2	27.55	0.774 2	27.21	0.738 5	24.46	0.734 9	26.95	0.855 6
SRCNN	×3	32.75	0.909 0	29.28	0.820 9	28.41	0.786 3	26.24	0.798 9	30.59	0.910 7
VDSR	×3	33.66	0.921 3	29.77	0.831 4	28.82	0.797 6	27.14	0.827 9	32.01	0.931 0
LapSRN	×3	33.82	0.920 7	29.89	0.834 9	28.82	0.795 0	27.07	0.829 8	32.21	0.931 8
DRRN	×3	34.03	0.924 4	29.96	0.834 9	28.95	0.800 4	27.53	0.837 8	32.74	0.939 0
MSRN	×3	34.46	0.927 8	30.41	0.843 7	29.15	0.806 4	28.33	0.856 1	33.95	0.946 6
本文	×3	34.77	0.930 0	30.65	0.846 1	29.25	0.808 9	28.71	0.858 8	34.02	0.950 5
Bicubic	×4	28.42	0.810 4	26.00	0.702 7	25.96	0.667 5	23.14	0.657 7	24.89	0.786 6
SRCNN	×4	30.48	0.862 8	27.50	0.751 3	26.90	0.710 1	24.52	0.722 1	27.58	0.855 5
VDSR	×4	31.35	0.883 0	28.02	0.768 0	27.26	0.726 0	25.18	0.754 0	28.83	0.887 0
LapSRN	×4	31.54	0.885 0	28.19	0.772 0	27.32	0.728 0	25.21	0.756 0	29.09	0.884 5
DRRN	×4	31.68	0.888 8	28.21	0.772 0	27.38	0.728 4	25.44	0.763 8	29.46	0.896 0
MSRN	×4	32.26	0.896 0	28.63	0.783 6	27.61	0.738 0	26.22	0.791 1	30.57	0.910 3
本文	×4	32.50	0.898 9	28.82	0.786 6	27.74	0.740 8	26.52	0.798 6	30.98	0.919 3

图 5、图 6 展示了数据集 Set5、Set14 在 ×4 放大倍数下的视觉重建效果图,左侧部分是原始的高清图像,右侧是 SR 重建图像针对左侧局部区域放大的对比展示图. 通过放大细节图,我们可以看出,仅通过 Bicubic 插值法对图像进行 ×4 倍采样的图像十分模糊,

睫毛细节和书本的纹路也难以观察到. 在图 5 中睫毛的高频信息更丰富. 对于图 6 的重建,我们放大了书架上横放的书本,其他一些经典算法的 SR 重建使得书本纹路紊乱,而本文算法重建细节效果比较好,图像更光滑和清晰。

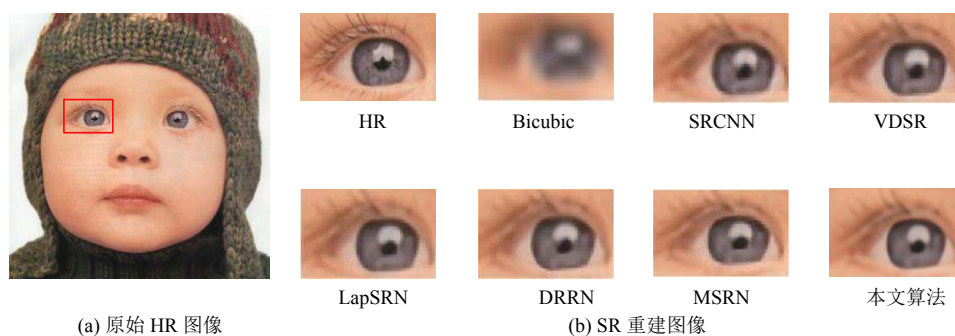


图5 数据集 Set5 中图像“baby”重建对比图

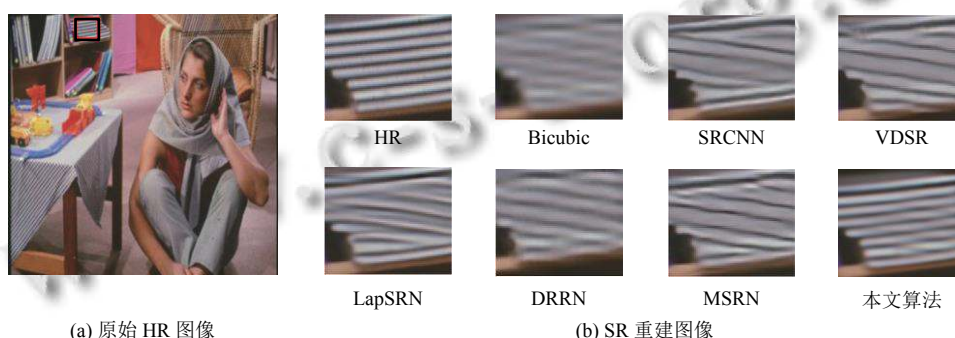


图6 数据集 Set14 中图像“barbara”重建对比图

4.4 不同数量残差块分析

本文展示的 MCRN 网络模型使用了多尺度注意力残差模块来对图像进行重建. 通过实验表明, 评价指标 *PSNR* 和 *SSIM* 都有所提升, 图像视觉的细节重建效果较好. 随着多尺度注意力残差模块数量的增加, 评价指标 *PSNR* 也会增加, 当残差块为 24 的时候, *PSNR* 达到 32.50 dB 且参数量也只有 5.8 M. 由图 7 可见, 在 12-24 个残差块的这个区间, 性能增加的最多. 所以本文选择了 24 个残差块对图像进行重建, 它平均重建一张图片需要 0.25 s.

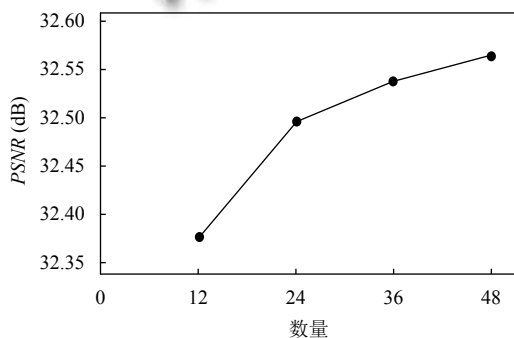


图7 Set5 在×4 倍不同 MCRB 数量性能对比

5 结语

本文提出多尺度注意力残差网络主要用 3 个不同大小的卷积层进行特征提取再进行融合, 以多个小卷积核替代大卷积核并引入注意力机制. 在控制参数数量的同时取得不错的重建效果, 不仅有效地缓解了梯度消失和梯度爆炸的问题, 网络模块的非线性表达能力也得到了增强. 注意力机制的引入, 使得通道赋予不同的权重, 有利于高频有效信息的传播和过滤掉网络中冗余的部分. 此外本文用 Bicubic 插值直接给重建图像提供低频信息, 此操作对网络前期训练优化迭代的稳定有益. 我们通过实验可以看出网络整体性能的提升, 也验证了网络的有效性.

参考文献

- 1 Dong C, Loy CC, He KM, *et al.* Image super-resolution using deep convolutional networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2016, 38(2): 295-307. [doi: 10.1109/TPAMI.2015.2439281]
- 2 Dong C, Loy CC, Tang XO. Accelerating the super-resolution convolutional neural network. *Proceedings of the 14th European Conference on Computer Vision*. Amsterdam:

- Springer, 2016. 391–407.
- 3 Kim J, Lee JK, Lee KM. Accurate image super-resolution using very deep convolutional networks. *Proceeding of 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Las Vegas: IEEE, 2016. 1646–1654.
 - 4 Zhang YL, Tian YP, Kong Y, *et al.* Residual dense network for image super-resolution. *Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Salt Lake City: IEEE, 2018. 2472–2481.
 - 5 Zhang YL, Li KP, Li K, *et al.* Image super-resolution using very deep residual channel attention networks. *Proceeding of the 15th European Conference on Computer Vision*. Munich: Springer, 2018. 294–310.
 - 6 Hu J, Shen L, Sun G. Squeeze-and-excitation networks. *Proceeding of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Salt Lake City: IEEE, 2018. 7132–7141.
 - 7 Tai Y, Yang J, Liu XM, *et al.* MemNet: A persistent memory network for image restoration. *Proceedings of the 2017 IEEE International Conference on Computer Vision (ICCV)*. Venice: IEEE, 2017. 4549–4557.
 - 8 Hui Z, Wang XM, Gao XB. Fast and accurate single image super-resolution via information distillation network. *Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Salt Lake City: IEEE, 2018. 723–731.
 - 9 Li JC, Fang FM, Mei KF, *et al.* Multi-scale Residual network for image super-resolution. *Proceedings of the 15th European Conference on Computer Vision*. Munich: Springer, 2018. 527–542.
 - 10 肖雅敏, 张家晨, 冯铁. 基于多窗口残差网络的单图像超分辨率重建. *计算机工程*, 2021, 47(2): 293–299, 306.
 - 11 Yang FZ, Yang H, Fu JL, *et al.* Learning texture transformer network for image super-resolution. *Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Seattle: IEEE, 2020. 5790–5799.
 - 12 Bevilacqua M, Roumy A, Guillemot C, *et al.* Low-complexity single-image super-resolution based on nonnegative neighbor embedding. *Proceedings of the 23rd British Machine Vision Conference*. Surrey: BMVA Press, 2012. 135.1–135.10.
 - 13 Zeyde R, Elad M, Protter M. On single image scale-up using sparse-representations. *Proceedings of the 7th International Conference on Curves and Surfaces*. Avignon: Springer, 2010. 711–730.
 - 14 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. *Proceedings of the 8th IEEE International Conference on Computer Vision*. Vancouver: IEEE, 2001. 416–423.
 - 15 Huang JB, Singh A, Ahuja N. Single image super-resolution from transformed self-exemplars. *Proceedings of 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Boston: IEEE, 2015. 5197–5206.
 - 16 Matsui Y, Ito K, Aramaki Y, *et al.* Sketch-based manga retrieval using Manga109 dataset. *Multimedia Tools and Applications*, 2017, 76(20): 21811–21838. [doi: [10.1007/s11042-016-4020-z](https://doi.org/10.1007/s11042-016-4020-z)]
 - 17 Kingma DP, Ba LJ. Adam: A method for stochastic optimization. arXiv: 1412.6980, 2014.
 - 18 Horé A, Ziou D. Image quality metrics: PSNR vs. SSIM. *2010 20th International Conference on Pattern Recognition*. Istanbul: IEEE, 2010. 2366–2369.
 - 19 Lai WS, Huang JB, Ahuja N, *et al.* Deep Laplacian pyramid networks for fast and accurate super-resolution. *Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Honolulu: IEEE, 2017. 5835–5843.
 - 20 Kim J, Lee JK, Lee KM. Deeply-recursive convolutional network for image super-resolution. *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Las Vegas: IEEE, 2016. 1637–1645.