

基于特征融合注意网络的图像超分辨率重建

周登文¹ 马路遥¹ 田金月¹ 孙秀秀¹

摘要 近年来, 基于深度卷积神经网络的单图像超分辨率重建, 取得了显著的进展, 但是, 仍然存在诸如特征利用率低、网络参数量大和重建图像细节纹理模糊等问题. 我们提出了基于特征融合注意网络的单图像超分辨率方法, 网络模型主要包括特征融合子网络和特征注意子网络. 特征融合子网络可以更好地融合不同深度的特征信息, 以及增加跨通道的学习能力; 特征注意子网络则着重关注高频信息, 以增强边缘和纹理. 实验结果表明: 无论是主观视觉效果, 还是客观度量, 我们方法的超分辨率性能明显优于其他代表性的方法.

关键词 单图像超分辨率, 卷积神经网络, 特征融合, 注意网络

引用格式 周登文, 马路遥, 田金月, 孙秀秀. 基于特征融合注意网络的图像超分辨率重建. 自动化学报, 2022, 48(9): 2233-2241

DOI 10.16383/j.aas.c190428

Feature Fusion Attention Network for Image Super-resolution

ZHOU Deng-Wen¹ MA Lu-Yao¹ TIAN Jin-Yue¹ SUN Xiu-Xiu¹

Abstract In recent years, single-image super-resolution (SISR) reconstruction based on deep convolutional neural networks has made significant progress, but there are still problems such as low feature utilization, large number of network parameters and blurred texture of the reconstructed image. We propose a new SISR network based on feature fusion attention mechanism, which mainly consists of a feature fusion sub-network and a feature attention sub-network. The feature fusion sub-network can better fuse feature information of different depths and increase the cross-channel learning ability; the feature attention sub-network focuses on high frequency information to enhance edges and textures. Experimental results demonstrate that the super-resolution performance of our method is significantly better than those of other representative methods in both subjective vision quality and objective metrics.

Key words Single image super-resolution, convolution neural network, feature fusion, attention mechanism

Citation Zhou Deng-Wen, Ma Lu-Yao, Tian Jin-Yue, Sun Xiu-Xiu. Feature fusion attention network for image super-resolution. *Acta Automatica Sinica*, 2022, 48(9): 2233-2241

近年来, 单图像超分辨率 (Single image super-resolution, SISR) 技术是图像处理 and 计算机视觉领域的研究热点, 旨在由单幅低分辨率 (Low-resolution, LR) 图像重建出具有更多细节的高分辨率 (High-resolution, HR) 图像, 它在医学成像^[1]、遥感卫星成像^[2] 和视频监控^[3] 等领域有广泛的应用.

超分辨率技术可大致分为 3 类: 基于插值^[4-5]、基于重建^[6-7] 和基于学习^[8-13] 的方法. 其中, 基于学习的超分辨率算法是目前的主流研究方向. 2014 年, Dong 等^[8] 提出了基于深度卷积神经网络 (Convolutional neural network, CNN) 的 SISR 算法, 称为 SRCNN. 不同于传统的基于学习的方法, SRCNN 直接学习低分辨率 (LR) 和高分辨率 (HR) 图像之

间端到端的映射. 该网络结构虽然简单 (只有 3 层), 但是, 超分辨效果很好. 2016 年, Dong 等^[9] 在 SRCNN 的基础上, 进一步提出了基于沙漏型结构的 FSRCNN 方法. 该方法参数量和计算量都更小, 速度也更快. Kim 等^[10] 利用残差网络 (ResNet)^[14], 加深了网络结构 (20 层), 提出了使用很深 CNN 的 SISR 方法 VDSR (Very deep super-resolution), 取得了更好的效果. Tai 等^[11] 提出了深度递归残差 SISR 网络 DRRN (Deep recursive residual network), 采用了更深的网络结构. 受益于参数共享策略, 相比 VDSR 方法, 参数量更少, 效果更好. 2017 年, Lai 等^[12-13] 提出了基于拉普拉斯图像金字塔的深度残差网络 LapSRN (Laplacian super-resolution network). LapSRN 网络模型包含多级, 每一级完成一次 2 倍上采样操作. 通过逐级上采样和预测残差, 实现图像超分辨率重建. 一次训练模型, 可以完成多个尺度的超分辨率任务. 我们的方法, 主要是受到 LapSRN 方法的启发.

VDSR、DRRN 和 LapSRN 等方法都表明: 网

收稿日期 2019-06-03 录用日期 2019-10-21

Manuscript received June 3, 2019; accepted October 21, 2019

本文责任编辑 黄庆明

Recommended by Associate Editor HUANG Qing-Ming

1. 华北电力大学控制与计算机工程学院 北京 102206

1. School of Control and Computer Engineering, North China Electric Power University, Beijing 102206

络深度对于超分辨率图像重建质量有着至关重要的影响. 以前基于深度 CNN 的超分辨率方法, 大多依赖于简单地叠加卷积层, 构建更深的网络, 以获得性能的改进. 但是, 这会导致对内存和计算能力需求的快速增加.

随着网络深度的增加, 感受野会不断增大, 各网络层提取的特征逐步抽象. 本文认为: 网络模型不同深度提取的特征, 均包含不同的、有利于超分辨率重建的信息, 以前的方法大多忽略了如何充分利用这些信息. 此外, 超分辨率重建的重点是恢复图像边缘和纹理等高频细节, 之前的方法通常对提取出来的特征信息同等对待, 没有重点关注图像边缘和纹理等细节.

近年来, 视觉注意机制^[15]在图像分类和超分辨率等应用中取得了成功^[16-19], 它可以重点关注感兴趣的目标, 抑制无用的信息. 为了解决上述问题, 本文提出了基于特征融合注意网络的超分辨率重建方法, 本文的贡献主要包括: 1) 提出了多级特征融合的网络结构, 能够更好地利用网络不同深度的特征信息, 以及增加跨通道的学习能力; 2) 提出了特征注意网络结构, 能够增强图像边缘和纹理等高频信息. 实验结果表明: 本文提出的基于特征融合注意网络的图像超分辨率方法, 无论是主观视觉效果, 还是客观度量, 超分辨率性能均超越了其他代表性的方法.

1 本文方法

本文方法主要是受到 LapSRN 方法的启发, 为了便于介绍本文方法, 首先对 LapSRN 方法作一个简单的介绍. 两级 LapSRN 网络分为两级超分辨率子网络 (每级 11 个卷积层), 每一级超分辨率子网络包含基于 CNN 的特征提取和反卷积上采样 2 个

部分, 以实现 2 倍放大因子的超分辨率重建. LapSRN 方法的两级超分辨率子网络, 可同时重建 $\times 2$ 和 $\times 4$ 倍超分辨率图像.

LapSRN 超分辨率子网络的结构较为简单, 只是常规 3×3 的卷积层的堆叠. 本文特征融合注意网络模型的总体架构类似于 LapSRN, 也由超分辨率子网络组成, 参见图 1. 但是, 在每一级子网络的内部, 有重要的改进. 本文模型的每一级子网络包括 3 个部分: 特征融合子网络、特征注意子网络和图像重建子网络 (2 倍反卷积上采样), 参见图 2. 特征融合子网络由维度转换层、递归卷积块、多通道融合层, 以及全局特征融合层 4 个部分组成. 递归卷积块提取的不同深度特征, 通过多通道融合层, 传递到全局特征融合层, 充分利用不同深度的特征信息和跨通道融合, 以自适应地学习跨通道特征信息之间的相互关系, 增强网络特征信息的选择能力. 特征注意子网络将特征融合子网络输出的特征, 进行高频信息的着重处理, 以增强特征中边缘和纹理等细节信息. 图像重建子网络则使用反卷积上采样层, 生成 2 倍超分辨率图像. 本文的特征融合子网络中包含 8 个递归卷积块, 每个卷积块包含 5 个 3×3

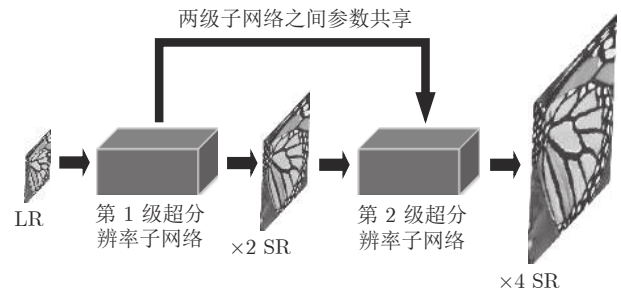


图 1 网络模型的总体结构

Fig.1 Overall structure of network model

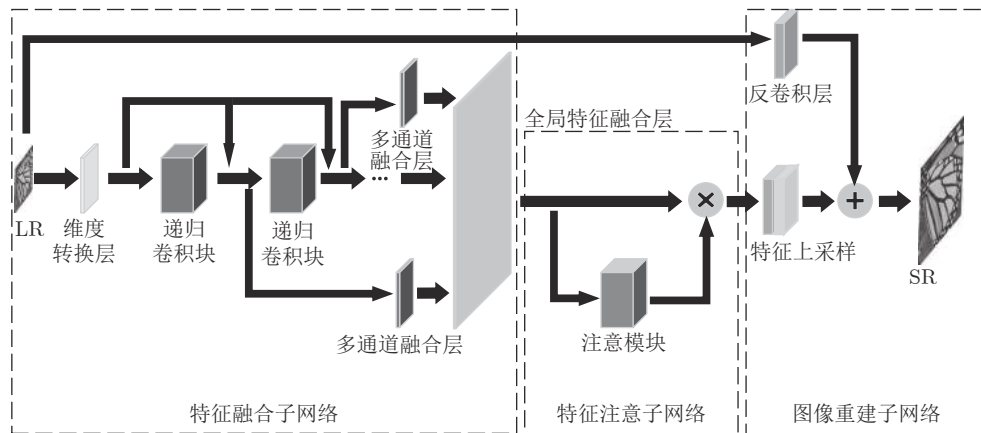


图 2 超分辨率子网络内部结构

Fig.2 Internal structure of super-resolution subnetwork

卷积层. 由于采用了递归结构, 各卷积块之间参数共享, 以及两级超分辨率子网络之间其他参数的共享. 虽然本文的网络模型有 80 个卷积层, 但参数总量大约只有 LapSRN 方法的 3/8. 下面对网络模型的主要部分进行详细的介绍.

1.1 特征融合子网络

本文的特征融合子网络由维度转换层、递归卷积块、多通道融合层, 以及全局特征融合层 4 个部分组成. 维度转换层为一个 3×3 的卷积层, 将输入的 LR 图像转换到高维特征, 并过滤掉一部分低频信息

$$H_0 = F_0(I_{LR}, W_0) \quad (1)$$

其中, I_{LR} 是输入的 LR 图像, F_0 是维度转换层的映射函数, W_0 是 F_0 的权重矩阵参数, H_0 是 F_0 的输出特征. H_0 输入到 8 个顺序连接的递归卷积块, 每个卷积块包含 5 个的卷积层, 参见图 3. 递归卷积块对 H_0 进一步进行特征提取

$$H_i = F_i(H_{i-1}, W_i) + H_0, \quad i = 1, 2, \dots, 8 \quad (2)$$

其中, F_i 是第 i 个递归卷积块的映射函数, W_i 是 F_i 的权重矩阵参数, H_i 是第 i 个递归卷积块的输出.

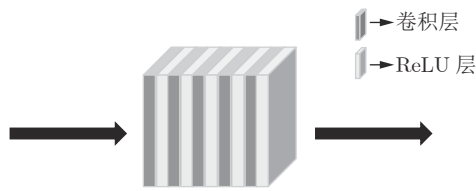


图 3 递归卷积块的结构

Fig.3 Structure of recursive convolutional block

多通道融合层包含 8 个 1×1 的卷积层, 分别对 8 个递归卷积块的输出特征 $H_i, i = 1, 2, \dots, 8$, 进行多通道融合和联通

$$H_{M,i} = F_{M,i}(H_i, W_{M,i}), \quad i = 1, 2, \dots, 8 \quad (3)$$

其中, $F_{M,i}$ 是第 i 个多通道融合层的映射函数, $W_{M,i}$ 是 $F_{M,i}$ 的权重矩阵参数, $H_{M,i}$ 是第 i 个多通道融合层的输出特征.

全局特征融合层将多通道融合层的输出特征 $H_{M,i}$ 进行拼接, 生成特征 H_M , 再通过 2 个 1×1 的卷积层, 进一步融合和降维, 以提高不同深度特征的利用率

$$H_G = F_G(H_M, W_G) \quad (4)$$

其中, F_G 是全局特征融合层的映射函数, W_G 是 F_G 的权重矩阵参数, H_G 是全局特征融合层的输出.

多通道融合层和全局特征融合层的相互结合, 使得本文的网络模型能够联通不同深度、不同通道

特征, 以更好地学习特征之间的相互关系.

1.2 特征注意子网络

注意力机制旨在将更多的注意力集中在感兴趣的信息上. 在图像超分辨率重建中, 我们更感兴趣图像的高频信息, 即图像的边缘和纹理等信息. 合适的网络结构是实现注意机制的关键. 本文的特征注意子网络包含一个注意模块, 由 7×7 的最大池化层、三个 3×3 的卷积层和一个 S 形 (sigmoid) 非线性激活层组成, 参见图 4. 7×7 的大感受野最大池化层, 可以更多地保留特征的纹理信息, 过滤一部分平滑的低频信息. S 形非线性激活函数的输出限制在 0 到 1 之间, 减少对特征信息的扰动, 并能够进一步抑制不重要的信息. 特征注意子网络得到一个与输入特征尺寸相同的增强矩阵 A , A 与输入特征 (即特征融合子网络的输出) H_G 逐元素相乘, 对 H_G 进行信息增强

$$A = R(P_{\max}(H_G), W_R) \quad (5)$$

$$H_A = A \times H_G \quad (6)$$

其中, A 是注意模块的输出, R 是注意模块中卷积块待学习的映射函数, W_R 是 R 的权重矩阵参数, P_{\max} 是最大池化层函数, H_A 是增强后的特征, (\cdot) 是逐元素相乘运算.

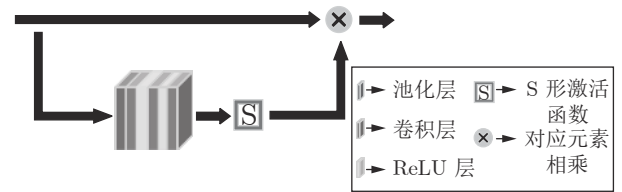


图 4 特征注意子网络结构

Fig.4 Structure of feature attention network

1.3 图像重建子网络

我们使用与 LapSRN 方法类似的图像重建子网络, 在每一级超分辨率子网络的末端进行一次 2 倍超分辨率重建. 将运用注意力机制增强后的特征 H_A 和 LR 图像分别进行反卷积上采样, 并逐元素相加, 生成重建的 2 倍超分辨率图像

$$I_{SR} = T_L(I_{LR}, W_{TL}) + F_D(T_H(H_A, W_{TH}), W_D) \quad (7)$$

其中, I_{SR} 是重建的超分辨率图像, I_{LR} 是输入的 LR 图像, T_L 是 LR 图像上采样反卷积层待学习的映射函数, W_{TL} 是 T_L 的权重矩阵参数, T_H 是特征 H_A 上采样反卷积层待学习的映射函数, W_{TH} 是 T_H 的权重矩阵参数, D 为降维卷积的映射函数, W_D 是 D 的权重矩阵参数.

1.4 跳跃连接

残差网络在计算机视觉任务中应用广泛,其跳跃连接通常源自前一层的输出.在特征融合子网络中,采用了同源跳跃连接结构,如图5所示.该结构将维度转换层的输出 H_0 传递到每个递归卷积块的输入中,每个递归卷积块的输入是上一个块的输出与 H_0 的和.这样可以向网络深层传递更多的浅层特征信息,也可以有效地减轻梯度消失和爆炸问题.

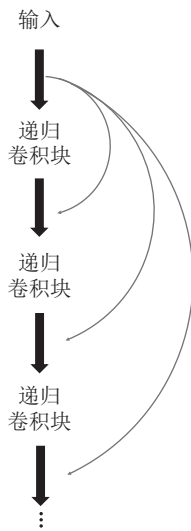


图5 跳跃连接结构

Fig.5 Structure of skip link

1.5 损失函数

本文采用了与 LapSRN 相同的 L1 损失函数^[13]

$$Loss = \frac{1}{N} \sum_{i=1}^N \sum_{l=1}^L \sqrt{(h_l^i - y_l^i)^2 + \epsilon^2} \quad (8)$$

其中, N 为每个训练批次的图像数量, L 为整个网络包含的子网络数. h_l^i 为高分辨率图像, y_l^i 为对应的重建图像, ϵ 设置为 10^{-6} .

2 实验结果

2.1 训练和测试数据集

不同的基于学习的 SISR 方法,使用不同的训练集. SRCNN 方法使用 ImageNet 训练集^[8], FSR-CNN 方法使用 General-100 训练集^[9], Yang 等^[20]使用 91 幅训练图像. LapSRN 方法使用 Berkeley Segmentation Dataset^[21] 中的 200 幅图像,以及 Yang 等^[20]方法中的 91 幅图像,共 291 幅图像作为训练集,并通过缩放、旋转等方式,扩充了训练集.由于本文的方法与 LapSRN 方法的网络模型总体

架构类似,因此也采用了与其相同的训练数据集和训练集增扩方法.

本文使用了 Set5^[22]、Set14^[23] 和 BSD100^[24] 标准测试数据集,分别包含 5、14 和 100 个高分辨率 (HR) 原图像.

2.2 训练数据集预处理和训练参数设置

本文使用不同的下采样因子,下采样原高分辨率 (HR) 训练图像,获得对应的低分辨率 (LR) 图像,并且通过缩放和旋转增扩了训练图像:首先,对训练图像以 0.5 和 0.7 的比例进行缩放;接着,对所有训练图像进行 $[90^\circ, 180^\circ, 270^\circ]$ 的旋转.所有训练图像都裁剪为 96×96 的图像片 (块), LR 和 HR 训练图像片对共 123 384 个.本文在 DIV2K^[25] 数据集中随机选取 20 幅图像,也进行了增扩处理,获得 2 688 个 LR 和 HR 图像片对,作为验证图像.本文网络模型中每一级超分辨率子网络参数设置可参见表 1.其中, H 、 W 是输入的 LR 图像的高度和宽度, Conv 是卷积层, LReLU 是渗漏、修正的非线性层, MaxPool 是最大池化层, ConvT 是反卷积层.

表 1 每一级超分辨率子网络的参数设置
Table 1 Parameter setting of each level of super-resolution sub-network

网络组件名称	组件内容及滤波器尺寸	输入尺寸	输出尺寸
维度转换层	Conv 3×3	$H \times W \times 3$	$H \times W \times 64$
递归卷积块×5	Conv 3×3	$H \times W \times 64$	$H \times W \times 64$
	LReLU	$H \times W \times 64$	$H \times W \times 64$
多通道特征融合层×8	Conv 1×1	$H \times W \times 64$	$H \times W \times 64$
全局特征融合层	Conv 1×1	$H \times W \times 512$	$H \times W \times 64$
	Conv 1×1	$H \times W \times 64$	$H \times W \times 64$
特征注意网络	MaxPool 7×7	$H \times W \times 64$	$H \times W \times 64$
	Conv 3×3	$H \times W \times 64$	$H \times W \times 32$
	LReLU	$H \times W \times 32$	$H \times W \times 32$
	Conv 3×3	$H \times W \times 32$	$H \times W \times 32$
	LReLU	$H \times W \times 32$	$H \times W \times 32$
特征上采样层	Conv 3×3	$H \times W \times 32$	$H \times W \times 64$
	ConvT 4×4	$H \times W \times 64$	$2H \times 2W \times 64$
	Conv 3×3	$2H \times 2W \times 64$	$2H \times 2W \times 3$
LR图像上采样层	ConvT 4×4	$H \times W \times 3$	$2H \times 2W \times 3$

网络模型的其他训练参数设置:每批 64 个训练图像对,全部训练图像对共 1 928 批,作为一个训练周期 (Epoch),共训练 100 个周期;训练优化算法: Adam^[26];初始学习率: 10^{-4} ,每迭代 20 个周期,学习率衰减一半;卷积滤波器初始化采用 MSRA 算法^[27];反卷积滤波器初始化采用均值为 0、标准差

为 0.001 的高斯分布^[9]. 在 Intel Core i7-4790K 4.00 GHz CPU, NVIDIA GTX Titan X GPU (12 GB 内存) 上, 使用 PyTorch 1.0 深度学习框架, 作为训练平台.

2.3 模型分析

本文在 Set5 和 Set14 标准测试数据集上进行 4 倍超分辨率实验. 分析特征融合子网络和特征注意力网络对超分辨率性能的影响, 参见表 2. 我们给出了不同模型变种的平均峰值信噪比 (Peak signal to noise ratio, PSNR)^[28] 和模型参数量. 可以看出: 递归结构可以在对网络性能影响不大 (PSNR 最大相差 0.01dB) 的情况下, 有效减少参数量, 不使用递归结构, 参数量将急剧增加 (大约增加到 7 倍). 特征融合子网络和特征注意力网络均可以提升网络的超分辨率性能, 两个子网络并用, 效果最好.

表 2 不同变种的网络模型×4 超分辨率, 在 Set5、Set14 数据集上的平均峰值信噪比 (dB) 及参数量

Table 2 Average PSNR (dB) and number of parameters of different super-resolution network models for scale factor ×4, on Set5 and Set14 datasets

特征融合子网络	特征注意力网络	递归结构	参数量 (k)	Set5	Set14
×	×	×	1 566	31.55	28.21
×	×	√	222	31.54	28.21
×	√	√	232	31.65	28.27
√	×	√	274	31.79	28.33
√	√	×	1 628	31.83	28.38
√	√	√	284	31.85	28.39

本文用 NVIDIA Titan X (单卡), 训练 100 个迭代周期 (Epoch), 大约需要 2 天. 网络收敛曲线参见图 6, 纵坐标是损失值, 横坐标是训练迭代周期数. “train loss” 是训练损失收敛曲线, “val loss” 是验证损失曲线. 当训练损失和验证损失趋向于稳定时, 停止训练.

2.4 实验结果分析

我们在 Set5、Set14 和 BSD100 标准测试集上, 验证 2 倍、4 倍和 8 倍超分辨率性能. 图像质量客观度量标准采用共同使用的 PSNR 和结构相似度 (Structural similarity, SSIM)^[28], PSNR 和 SSIM 值越大, 表示图像质量越高. 我们也与一些代表性的 SISR 方法进行了比较, 包括经典的双三次插值方法 Bicubic^[29], Huang 等^[30] 基于自相似的 SISR 方法 SelfEx, Dong 等^[8] 使用深度卷积网络的 SISR 方法 SRCNN, Kim 等^[10] 使用深度 CNN 的 SISR 方法 VDSR, Tai 等^[11] 深度递归残差网的 SISR 方法

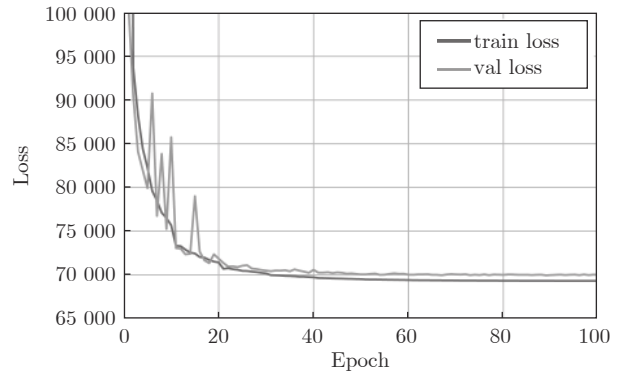


图 6 网络收敛曲线 (“train loss” 是训练损失收敛曲线, “val loss” 是验证损失曲线)

Fig.6 Network convergence curves (“train loss” is the training loss convergence curve; “val loss” is the validation loss curve)

DRRN, Lai 等拉普拉斯金字塔网络的 SISR 方法 LapSRN^[12] 和 Ms-LapSRN^[13] 等. Bicubic 方法使用 MATLAB interp2 函数实现; 其他方法的实现均来自作者公开的源代码, 使用原方法论文中的参数. 2 倍、4 倍模型使用原作者预训练模型, 8 倍模型用作者源代码训练得到. 计算结果参见表 3.

一个已知的事实是: PSNR 和 SSIM 等客观度量方法与人的主观视觉并不是完全一致的. 本文方法与其他比较的方法, 在 Set5、Set14 和 BSD100 测试数据集上部分图像的 4 倍超分辨率的视觉效果 (参见图 7 ~ 9), 图 10 是 8 倍超分辨率的结果. 为了便于观察, 对比的结果图像进行了局部裁剪和放大. 第 1 行和第 3 行是选择的裁剪区域, 第 2 行和第 4 行是对红色方框中重点区域的放大. 各个比较的方法标注在图像的下方. 本文方法的超分辨结果明显优于其他比较的方法, 更好地恢复了图像的边缘和纹理等细节, 是清晰可视的. 尤其是图 8 中花瓶的纹理和鱼的眼睛, 图 7 中的文字和马腿部花纹.

基于深度 CNN 的 6 个 SISR 方法在 BSD100 测试数据集上, 4 倍超分辨率的平均 PSNR 和参数量的对比, 参见图 11. SRCNN 的网络模型只有 3 层, 故参数量最小, 但是其 PSNR 性能最低; 本文方法的平均 PSNR 最高, 参数量却大约只有 LapSRN 的 1/3, 与 DRRN 和 Ms-LapSRN 方法, 大致相当.

基于深度 CNN 的 6 个 SISR 方法在 BSD100 测试数据集上, 4 倍超分辨率的平均运行时间对比, 参见图 12. 本文方法的速度略低于 LapSRN、Ms-LapSRN 和 SRCNN, 但明显快于 VDSR 与 DRRN 方法. LapSRN 方法最快, 比本文方法也仅快大约 0.06 s, 本文方法比 VDSR 大约快 0.1 s, 比 DRRN 大约快 0.8 s.

表 3 在 Set5、Set14 和 BSD100 测试数据集上, 各种超分辨率方法的 $\times 2$ 、 $\times 4$ 和 $\times 8$ 超分辨率的平均 PSNR (dB) 和 SSIM

Table 3 Average PSNR (dB)/SSIMs of various SISR methods for scale factor $\times 2$, $\times 4$, and $\times 8$ on Set5, Set14 and BSD100

测试集	放大倍数	Bicubic ^[28] PSNR/SSIM	SelfEx ^[29] PSNR/SSIM	SRCNN ^[8] PSNR/SSIM	LapSRN ^[12] PSNR/SSIM	VDSR ^[10] PSNR/SSIM	DRRN ^[11] PSNR/SSIM	MsLapSRN ^[13] PSNR/SSIM	本文方法 PSNR/SSIM
Set5	2	33.66/0.929	36.49/0.953	36.66/0.954	37.52/0.959	37.53/0.959	37.74/0.959	37.78/0.960	37.91/0.969
Set14	2	30.24/0.868	32.22/0.903	32.42/0.906	33.08/0.913	33.05/0.913	33.23/0.914	33.28/0.915	33.42/0.925
BSD100	2	29.56/0.841	31.18/0.885	31.36/0.887	31.80/0.895	31.90/0.896	32.05/0.897	32.05/0.898	32.19/0.904
Set5	4	28.42/0.810	30.31/0.861	30.48/0.862	31.54/0.885	31.35/0.883	31.68/0.888	31.74/0.889	31.85/0.908
Set14	4	26.00/0.702	27.40/0.751	27.49/0.753	28.19/0.772	28.02/0.768	28.21/0.772	28.26/0.774	28.39/0.789
BSD100	4	25.96/0.667	26.84/0.710	26.90/0.711	27.29/0.727	27.32/0.726	27.38/0.728	27.43/0.731	27.50/0.748
Set5	8	24.40/0.658	25.49/0.703	25.33/0.690	26.15/0.738	25.93/0.724	26.18/0.738	26.34/0.752	26.40/0.755
Set14	8	23.10/0.566	23.92/0.601	23.76/0.591	24.35/0.620	24.26/0.614	24.42/0.622	24.57/0.629	24.60/0.631
BSD100	8	23.67/0.548	24.19/0.568	24.13/0.566	24.54/0.586	24.49/0.583	24.59/0.587	24.65/0.591	24.72/0.602

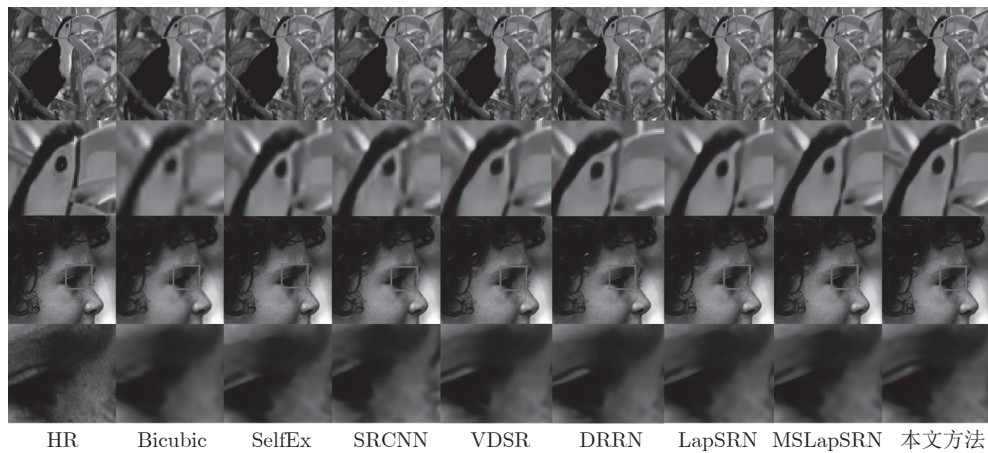


图 7 在 Set5 测试数据集上, 2 个测试图像 $\times 4$ 超分辨率结果对比

Fig. 7 A comparison of super-resolution results of two test images in Set5 for scale factor $\times 4$



图 8 在 Set14 测试数据集上, 2 个测试图像 $\times 4$ 超分辨率结果对比

Fig. 8 A comparison of super-resolution results of two test images in Set14 for scale factor $\times 4$

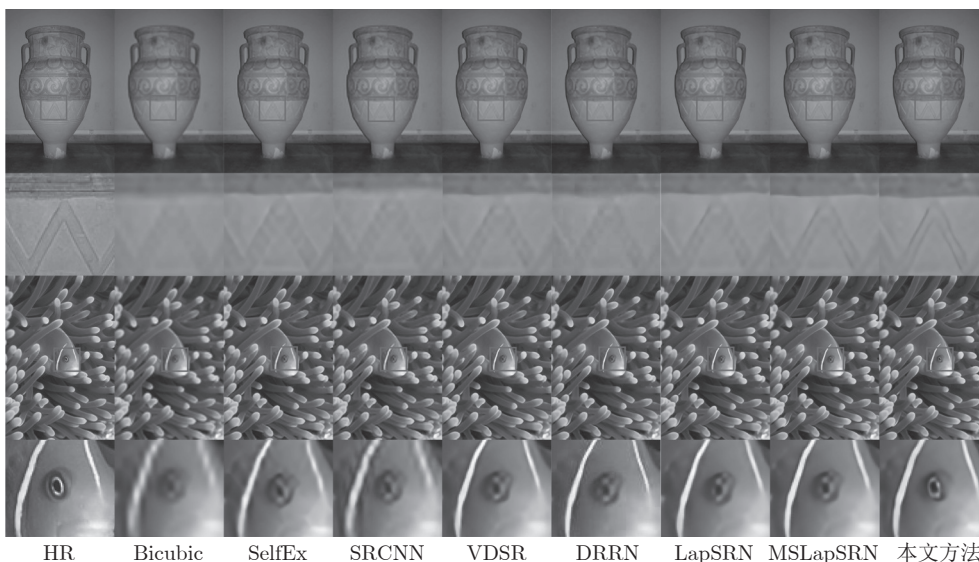


图 9 在 BSD100 测试数据集上, 2 个测试图像×4 超分辨率结果对比

Fig.9 A comparison of super-resolution results of two test images in BSD100 for scale factor ×4

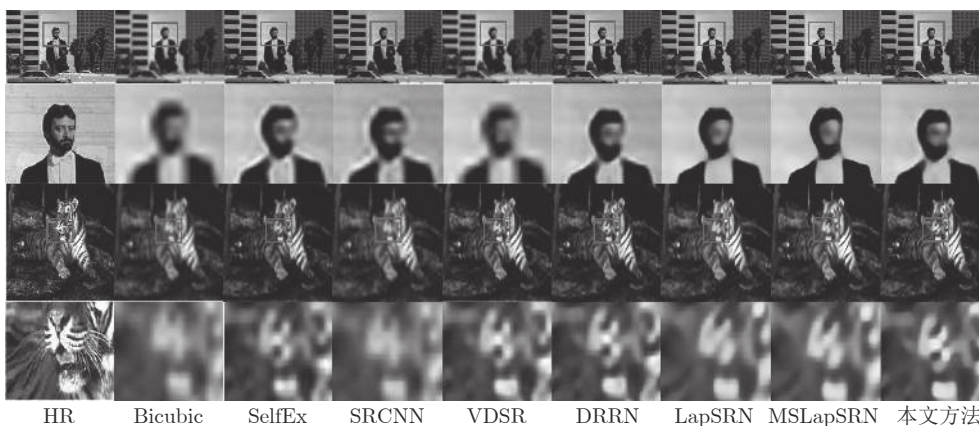


图 10 在 BSD100 测试数据集上, 2 个测试图像×8 超分辨率结果对比

Fig.10 A comparison of super-resolution results of two test images in BSD100 for scale factor ×8

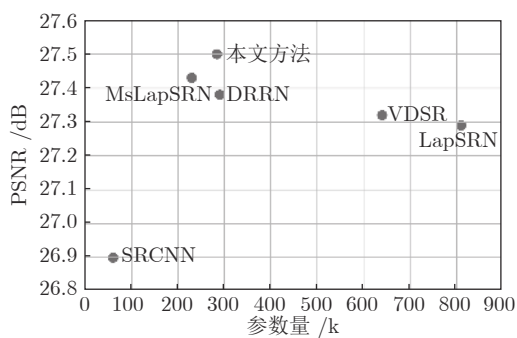


图 11 6 个基于深度 CNN 的方法, 在 BSD100 数据集上×4 超分辨率的平均 PSNR 和参数量对比

Fig.11 Number of parameters and average PSNR of six methods based on depth CNN, on the BSD100 for scale factor ×4

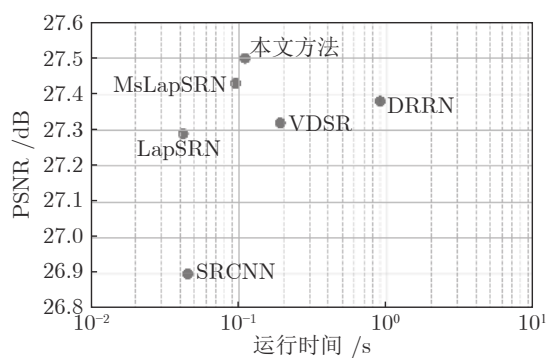


图 12 6 个基于深度 CNN 的方法, 在 BSD100 数据集上×4 超分辨率的平均运行时间对比

Fig.12 A comparison of running times of six methods based on depth CNN, on the BSD100 for scale factor ×4

3 结束语

针对之前基于深度 CNN 的 SISR 方法中参数量大, 图像边缘和纹理恢复效果不好等问题, 受基于拉普拉斯图像金字塔的深度残差网络方法 (LapSRN) 的启发, 提出了基于特征融合注意网络的超分辨率重建方法. 我们保留了 LapSRN 方法分级重建的优点, 采用了递归和其他参数共享策略, 使参数量急剧下降, 大约只有 LapSRN 方法的 3/8. 由于采用了多级特征融合和特征注意机制, 使得本文的网络模型, 能够充分融合和利用不同深度、不同通道的特征, 重点增强边缘和纹理等高频信息, 超分辨率效果显著提升. 实验结果也证实了本文方法的良好性能, 在进行比较的代表性方法中, 本文方法更好地平衡了性能、模型的复杂度和运行时间等因素. 未来我们将计划进一步优化和改进我们的方法.

References

- Oktay O, Bai W J, Lee M, Guerrero R, Kamnitsas K, Caballero J, et al. Multi-input cardiac image super-resolution using convolutional neural networks. In: Proceedings of the 19th International Conference on Medical Image Computing and Computer-assisted Intervention. Athens, Greece: Springer, 2016. 246–254
- Luo Y M, Zhou L G, Wang S, Wang Z Y. Video satellite imagery super resolution via convolutional neural networks. *IEEE Geoscience and Remote Sensing Letters*, 2017, **14**(12): 2398–2402
- Rasti P, Uiboupin T, Escalera S, Anbarjafari G. Convolutional neural network super resolution for face recognition in surveillance monitoring. In: Proceedings of the 9th International Conference on Articulated Motion and Deformable Objects. Palma de Mallorca, Spain: Springer, 2016. 175–184
- Harris J L. Diffraction and resolving power. *Journal of the Optical Society of America*, 1964, **54**(7): 931–936
- Lu Zhi-Fang, Zhong Bao-Jiang. Image interpolation with predicted gradients. *Acta Automatica Sinica*, 2018, **44**(6): 1072–1085 (陆志芳, 钟宝江. 基于预测梯度的图像插值算法. *自动化学报*, 2018, **44**(6): 1072–1085)
- Xiong Jiao-Jiao, Lu Hong-Yang, Zhang Ming-Hui, Liu Qie-Gen. Convolutional sparse coding in gradient domain for MRI reconstruction. *Acta Automatica Sinica*, 2017, **43**(10): 1841–1849 (熊娇娇, 卢红阳, 张明辉, 刘且根. 基于梯度域的卷积稀疏编码磁共振成像重建. *自动化学报*, 2017, **43**(10): 1841–1849)
- Sun Jing, Yuan Qiang-Qiang, Li Ji-Wei, Zhou Chun-Ping, Shen Huan-Feng. License plate image super-resolution based on intensity-gradient prior combination. *Journal of Image and Graphics*, 2018, **23**(6): 802–813 (孙京, 袁强强, 李冀玮, 周春平, 沈焕锋. 亮度—梯度联合约束的车牌图像超分辨率重建. *中国图象图形学报*, 2018, **23**(6): 802–813)
- Dong C, Loy C C, He K M, Tang X O. Learning a deep convolutional network for image super-resolution. In: Proceedings of the 13th European Conference on Computer Vision (ECCV). Zurich, Switzerland: Springer, 2014. 184–199
- Dong C, Loy C C, Tang X O. Accelerating the super-resolution convolutional neural network. In: Proceedings of the 14th European Conference on Computer Vision (ECCV). Amsterdam, the Netherlands: Springer, 2016. 391–407
- Kim J, Lee J K, Lee K M. Accurate image super-resolution using very deep convolutional networks. In: Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, USA: IEEE, 2016. 1646–1654
- Tai Y, Yang J, Liu X M. Image super-resolution via deep recursive residual network. In: Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, USA: IEEE, 2017. 2790–2798
- Lai W S, Huang J B, Ahuja N, Yang M H. Deep Laplacian pyramid networks for fast and accurate super-resolution. In: Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, USA: IEEE, 2017. 5835–5843
- Lai W S, Huang J B, Ahuja N, Yang M H. Fast and accurate image super-resolution with deep Laplacian pyramid networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2019, **41**(11): 2599–2613
- He K M, Zhang X Y, Ren S Q, Sun J. Deep residual learning for image recognition. In: Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, USA: IEEE, 2016. 770–778
- Koch C, Ullman S. Shifts in selective visual attention: Towards the underlying neural circuitry. *Matters of Intelligence: Conceptual Structures in Cognitive Neuroscience*. Dordrecht: Springer, 1987. 115–141
- Mnih V, Heess N, Graves A, Kavukcuoglu K. Recurrent models of visual attention. In: Proceedings of the 27th International Conference on Neural Information Processing Systems. Montreal, Canada: MIT Press, 2014. 2204–2212
- Wang F, Jiang M Q, Qian C, Yang S, Li C, Zhang H G, et al. Residual attention network for image classification. In: Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, USA: IEEE, 2017. 6450–6458
- Zhang Q S, Wu Y N, Zhu S C. Interpretable convolutional neural networks. In: Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA: IEEE, 2018. 8827–8836
- Zhang Y L, Li K P, Li K, Wang L C, Zhong B N, Fu Y. Image super-resolution using very deep residual channel attention networks. In: Proceedings of the 15th European Conference on Computer Vision (ECCV). Munich, Germany: Springer, 2018. 294–310
- Yang J C, Wright J, Huang T S, Ma Y. Image super-resolution via sparse representation. *IEEE Transactions on Image Processing*, 2010, **19**(11): 2861–2873
- Martin D, Fowlkes C, Tal D, Malik J. A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. In: Proceedings of the 8th IEEE International Conference on Computer Vision. Vancouver, Canada: IEEE, 2002. 416–423
- Bevilacqua M, Roumy A, Guillemot C, Morel M L M. Neighbor embedding based single-image super-resolution using semi-non-negative matrix factorization. In: Proceedings of the 2012 IEEE International Conference on Acoustics, Speech and Signal Pro-

cessing (ICASSP). Kyoto, Japan: IEEE, 2012. 1289–1292

- 23 Zeyde R, Elad M, Protter M. On single image scale-up using sparse-representations. In: Proceedings of the 7th International Conference on Curves and Surfaces. Avignon, France: Springer, 2010. 711–730
- 24 Arbeláez P, Maire M, Fowlkes C, Malik J. Contour detection and hierarchical image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2011, **33**(5): 898–916
- 25 Timofte R, Agustsson E, van Gool L, Yang M H, Zhang L, Lim B, et al. NTIRE 2017 challenge on single image super-resolution: Methods and results. In: Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). Honolulu, USA: IEEE, 2017. 1110–1121
- 26 Kingma D P, Ba J. Adam: A method for stochastic optimization. In: Proceedings of the 3rd International Conference on Learning Representations (ICLR). San Diego, USA, 2015.
- 27 He K M, Zhang X Y, Ren S Q, Sun J. Delving deep into rectifiers: Surpassing human-level performance on ImageNet classification. In: Proceedings of the 2015 IEEE International Conference on Computer Vision (ICCV). Santiago, Chile: IEEE, 2015. 1026–1034
- 28 Wang Z, Bovik A C, Sheikh H R, Simoncelli E P. Image quality assessment: From error visibility to structural similarity. *IEEE Transactions on Image Processing*, 2004, **13**(4): 600–612
- 29 Keys R G. Cubic convolution interpolation for digital image processing. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 1981, **29**(6): 1153–1160
- 30 Huang J B, Singh A, Ahuja N. Single image super-resolution from transformed self-exemplars. In: Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Boston, USA: IEEE, 2015. 5197–5206



周登文 华北电力大学控制与计算机工程学院教授。主要研究方向为图像去噪, 图像去马赛克, 图像插值和图像超分辨率。本文通信作者。

E-mail: zdw@ncepu.edu.cn

(ZHOU Deng-Wen Professor at the School of Control and Computer

Engineering, North China Electric Power University. His research interest covers image denoising, image demosaicing, image interpolation, and image super-resolution. Corresponding author of this paper.)



马路遥 华北电力大学控制与计算机工程学院硕士研究生。2015 年获得华北电力大学核科学与工程学院学士学位。主要研究方向为计算机视觉和深度学习。E-mail: maly@163.com

(MA Lu-Yao Master student at the School of Control and Computer

Engineering, North China Electric Power University. He received his bachelor degree from the School of Nuclear Science and Engineering, North China Electric Power University in 2015. His research interest covers computer vision and deep learning.)



田金月 华北电力大学控制与计算机工程学院硕士研究生。2017 年获得北华航天工业学院计算机与遥感信息技术学院学士学位。主要研究方向为计算机视觉和深度学习。

E-mail: tjytjy20220810@163.com

(TIAN Jin-Yue Master student at

the School of Control and Computer Engineering, North China Electric Power University. She received her bachelor degree from the School of Computer and Remote Sensing Information Technology, North China Institute of Aerospace Engineering in 2017. Her research interest covers computer vision and deep learning.)



孙秀秀 华北电力大学控制与计算机工程学院硕士研究生。2017 年获得青岛大学计算机科学技术学院学士学位。主要研究方向为计算机视觉和深度学习。

E-mail: 18810805256@163.com

(SUN Xiu-Xiu Master student at

the School of Control and Computer Engineering, North China Electric Power University. She received her bachelor degree from the College of Computer Science and Technology, Qingdao University in 2017. Her research interest covers computer vision and deep learning.)