

doi: 10.19697/j.cnki.1673-4432.201905008

多尺度卷积神经网络的图像超分辨率重建算法

何一凡, 林熠珉, 林佳敏, 杜晓凤

(厦门理工学院计算机与信息工程学院, 福建 厦门 361024)

[摘 要] 为利用多尺度信息重建超分辨率图像, 提出多尺度卷积神经网络的图像超分辨率重建算法。算法利用不同尺度的卷积核提取图像特征, 为图像重建提供不同大小的邻域信息; 用瓶颈层融合多尺度特征图, 增强网络非线性表示能力, 降低中间层输出的维数, 提高图像的超分辨率重建性能。多个测试集上的实验结果表明, 多尺度卷积神经网络算法优于现有的单幅图像超分辨率方法。

[关键词] 图像重建; 超分辨率算法; 卷积神经网络; 多尺度

[中图分类号] TP391 **[文献标志码]** A **[文章编号]** 1673-4432 (2019) 05-0047-06

图像分辨率是指图像中包含的信息量, 是图像展现细节丰富程度的一个重要衡量指标。高分辨率的图像呈现出更丰富的细节信息, 具有更强的信息表达能力。在实际应用中, 由于成像设备硬件条件的限制和成像环境等因素的影响, 人们很难获得高分辨率的图像。图像超分辨率技术就是以现有的低分辨率图像为输入, 利用软件算法模型重建出高分辨率、高质量的图像^[1]。

基于学习的图像超分辨率重建是一个非常活跃的研究主题, 深度学习这一新兴的工具也被引入到图像重建研究中。在图像超分辨率重建任务中, Dong 等^[2]首次构建基于卷积神经网络(convolutional neural network, CNN)的图像超分辨率重建网络(super-resolution convolutional neural network, SRCNN), 实现了端到端的超分辨率图像恢复。SRCNN 的输入为低分辨率图像, 经过三个卷积层卷积操作后, 输出重建的高分辨率图像。通过大量的样本训练, SRCNN 可以学习到低分辨率图像与高分辨图像之间的映射关系。在图像超分辨率重建时, 有效地利用像素的邻域信息可以提升重建图像的准确度。基于 CNN 的超分辨率重建算法中影响邻域范围的因素主要有两个: 一是卷积层中的卷积核的尺度。较大尺度的卷积核能作用于较大的图像区域, 这就意味着可以获取较大的邻域信息重构图像^{[3]299}。二是由输入到输出所经过的网络路径深度, 重建所经过的卷积层数量越多, 可以利用邻域就越大。Kim 等^[4]通过串联 20 个卷积层, 增大了重建可利用的图像邻域, 提高了重建图像的准确度。近几年, 研究工作主要集中在增加网络深度或设计复杂网络, 以进一步提高图像重建的性能^[5-8]。

虽然深度网络提高了图像超分辨率重建的性能, 但忽略了图像的多尺度表示。自然图像中包含多种类型的特征, 这些不同类型的特征尺度不尽相同。例如, 平滑区域和强边缘区域对应的尺度较大, 而纹理区域的特征对应的尺度稍小^[9-10]。传统 CNN 中卷积层只包含单一尺度卷积核, 当卷积核尺度较大时, CNN 可以捕获大尺度的特征, 而当卷积核尺度较小时, CNN 倾向于检测到较小尺度的细节特征^[3]。因此, 传统的 CNN 无法实现图像的多尺度重建。GoogLeNet 通过不同尺度卷积核提取图像特

收稿日期: 2019-03-05

修回日期: 2019-04-01

基金项目: 厦门理工学院高层次人才项目 (YKJ16018R); 福建省中青年教育科研项目 (JAT170418); 厦门市留学人员科研项目 (厦人社 [2018] 310 号); 福建省自然科学基金项目 (2019J01854)

通信作者: 何一凡, 男, 讲师, 博士, 研究方向为高性能神经网络加速器、数字图像处理等, E-mail: y.he@xmut.edu.cn。

引文格式: 何一凡, 林熠珉, 林佳敏, 等. 多尺度卷积神经网络的图像超分辨率重建算法[J]. 厦门理工学院学报, 2019, 27(5): 47-52.

Citation: HE Y F, LIN Y M, LIN J M, et al. Multi-scale convolutional neural network for super-resolution imaging [J]. Journal of Xiamen University of Technology, 2019, 27(5): 47-52. (in Chinese)



征,在图像识别任务中获得很好结果^[11]。受 GoogLeNet 的启发,本文提出基于多尺度卷积神经网络的图像超分辨率重建算法。与传统的单一尺度的卷积层不同,多尺度卷积神经网络设计有多尺度模块,其中包含不同尺度的卷积核,用以提取图像的多尺度特征;进而通过瓶颈层融合不同尺度的特征,降低中间层输入,以提高网络的非线性表示能力,最终实现图像的多尺度超分辨率重建。

1 多尺度卷积神经网络算法

本文设计的多尺度卷积神经网络结构如图 1 所示。网络由级联的多尺度模块组成,输入为低分辨率图像,输出为重建的高分辨率图像。

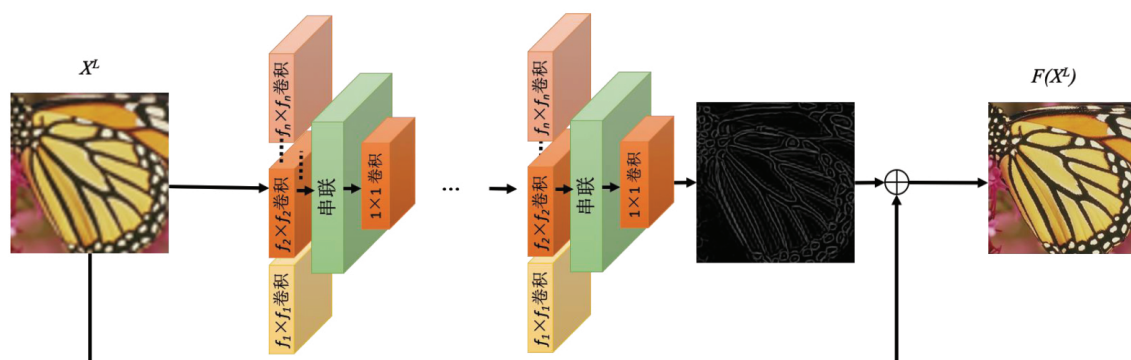


图 1 多尺度卷积神经网络结构

Fig. 1 Architecture of multi-scale CNN

图 1 中,网络的输入为图像 X^L ,通过多尺度模块对 X^L 进行多个不同尺度卷积融合,提取图像的多尺度特征;这些特征图进一步传输到 1×1 卷积层进行融合,然后传入后续模块,最终生成目标图像 $F(X^L)$ 。多个多尺度模块首尾相连,就搭建起了多尺度卷积神经网络。通过训练,网络自我学习到低分辨率图像 X^L 与目标高分辨率图像 X^H 的映射函数: $F(X^L; X^H; \Theta)$, $\Theta = \{W_1, W_2, \dots, W_M; B_1, B_2, \dots, B_M\}$ 。其中 W_l 包括多尺度模块 l 中的所有权值参数, B_l 包括所有的偏置参数, M 表示网络中的多尺度模块个数。网络通过跨层连接实现残差学习以提高网络学习效率。

1.1 多尺度模块

多尺度模块由两个部分组成:第一部分执行多尺度卷积并产生一组特征图;第二部分通过 1×1 卷积在串联的特征图上执行跨通道融合。

1) 多尺度卷积。设 Y 为第 l 层的输入数据,对 Y 进行尺度为 l_i 的卷积和激活操作得到特征图 $F_{li}(Y) = \max(0, W_{li} * Y + B_{li})$, W_{li} 表示第 l 层的第 i 种尺度的权值矩阵, B_{li} 表示第 l 层的第 i 种的偏置, $*$ 为卷积操作, \max 为激活函数。多尺度卷积操作得到的结果为特征图集合 $\{F_{l_1}(Y), F_{l_2}(Y), \dots, F_{l_s}(Y)\}$, s 为尺度数目。设 $\{f_1, f_2, \dots, f_s\}$ 表示有 s 个尺度的模块,该模块包括 $f_1 \times f_1, f_2 \times f_2, \dots, f_s \times f_s$ 的 s 个尺度卷积层。例如 $\{9, 5, 3\}$ 表示有 3 个尺度的模块,其中包括 $9 \times 9, 5 \times 5, 3 \times 3$ 的 3 个尺度卷积。

2) 跨通道融合。假设 $\{f_1, f_2, \dots, f_s\}$ 中每种尺度的卷积核为 m 个,卷积结果串联得到通道数 $m \times s$ 特征图。在计算资源有限的情况下,过大的通道数会消耗过多计算资源。利用 1×1 卷积层可以融合特征并降低特征图维度。具体来说,设 1×1 卷积层中卷积核的个数为 m 。多尺度卷积特征图集合中的每组特征图串联拼接在一起得到 $F_l'(Y) = [F_{l_1}, F_{l_2}, \dots, F_{l_s}]$,然后利用 1×1 卷积 $F_l'(Y)$ 得到第 l 层的融合输出:

$$F_l(Y) = \max(0, W_{1 \times 1} * F_l'(Y) + B_{1 \times 1}) \quad (1)$$

将单一尺度卷积层替换为多尺度模块,多尺度模块就可以灵活地嵌入到传统卷积神经网络中,也就建立了多尺度卷积神经网络。

1.2 网络训练

1) 损失函数。给定低分辨率图像 X_N^L 作为网络输入, 将 X_N^L 对应的高分辨率图像 X_N^H 作为网络输出, 损失函数定义为

$$E(\theta) = \frac{1}{N} \sum_{n=1}^N \|F(X_N^L; \theta) - X_N^H\|^2. \quad (2)$$

式(2)中: n 表示第 N 张图像; N 是训练集中样本的数量; $\|\cdot\|^2$ 表示 $E(\theta)$ 矩阵范数的平方。

2) 训练。模型的训练就是通过最小化损失函数来估计映射函数 F 中参数 θ 的最优取值 $\hat{\theta}$ 。损失函数的最小化由梯度下降算法和标准的反向传播算法实现^[12]。训练中学习率设为 10^{-3} 。

1.3 模型分析

图像重建依赖于可利用的邻域信息, 在 CNN 网络中, 重建所依赖的邻域大小与网络的感受野相关, 而 CNN 网络感受野尺度由各卷积层的滤波器尺度所决定。假设共 L 层卷积层, 卷积层 l 的滤波器尺度为 $f_l \times f_l$, 则第 l 层的感受野大小为

$$r_l = r_{l+1} - 1 + f_l; \quad i \in 1, 2, \dots, L; \quad r_{L+1} = 1. \quad (3)$$

式(3)中: f_l 表示第 l 层的卷积核尺度; r_{L+1} 表示输出层的感受野。SRCNN 中 CNN 模型的网络结构为 9-5-5, 对应到图像超分辨率中, 重建像素可利用的图像邻域信息大小为 17×17 。设多尺度卷积神经网络 {5, 9, 13}-5-5, 即网络由一个多尺度模块两个单尺度模块串联组成。其中多尺度模块包括 3 个尺度的卷积核, 5×5 , 9×9 和 13×13 , 单尺度卷积层的卷积核为 5×5 。因此网络输出层的有 3 个尺度的感受野: 13×13 , 17×17 和 21×21 , 从而为图像重建提供多种尺度的邻域信息。设 {5, 9, 13} 中每种尺度的卷积核为 64 个, 卷积结果串联得到的特征图通道数为 64×3 。在跨通道融合中, 1×1 卷积层的卷积核为 64 个, 即卷积核的维度为 $(64 \times 3) \times 1 \times 1 \times 64$, 通过 1×1 卷积后得到的特征图通道数为 64。这样不但有效地减小了传入后续网络的特征图通道数, 还增加了网络的深度, 进而提高了网络非线性映射能力。

2 实验及分析

2.1 测试样本及环境

实验样本来自 Yang 等^[13]的超分辨率训练图库中 91 张图、伯克利图像分割库 (BSD)^[14]中 200 张图。

训练集的构建主要有以下 3 个步骤: (1) 对训练库中的每一张 32×32 的高分辨图像进行双三次下采样, 缩放到原图 $1/s$ 大小, 得到一系列低分辨小图 (当 s 为 3 时, 图像大小为 33×33); (2) 将小图双三次上采样 s 倍, 生成低分辨率图像, 这样就得到一系列高分辨率图像和低分辨率图像; (3) 以 10×10 间隔采样 32×32 大小的图像块作为样本 (在放大 3×3 倍时, 图像块大小为 33×33)。

在 291 张图中共采样 3 000 000 个图像块作为样本库。训练中每批次样本数为 32 个; 每个卷积层初始化权值是方差为 10^{-2} 、均值为 0 的随机值。因为人眼对亮度通道最为敏感, 在处理彩色图像时, 将图像从 RGB 颜色空间转换到 YCbCr 颜色空间, 相关操作都在亮度通道上进行, 针对 Cb 和 Cr 通道, 在重建时直接使用双三次插值算法进行放大。

客观评价标准采用峰值信噪比 (peak signal-to-noise ratio, PSNR) 和结构相似度 (structure similarity, SSIM) 两种指标评价不同超分辨率算法的重建性能。网络性能越好, 则 PSNR 和 SSIM 的值越高。

测试库: 测试库为通用测试库 Set 5、Set 14 及 BSD 库中的测试图像 100 张。

硬件环境: Nvidia GTX 960, Intel i7 6700, 16 G 内存。

软件环境: Matlab 2014 Rb, matconvnet^[15]。

所有的对比实验都在同样的软硬件环境下进行。

2.2 参数设置对性能的影响

通过串联多个多尺度模块可以搭建不同深度的卷积神经网络, 表 1 给出不同深度网络的 PSNR 性能评价和参数量。

表 1 在 Set 5 和 Set 14 测试库放大 4×4 倍不同深度网络的 PSNR 结果

Table 1 PSNR values of networks of different depth on Set 5 and Set 14 with upscaling factor 4.

不同深度网络	含义	PSNR		网络参数量/个
		Set 5	Set 14	
{5, 9, 13} - 5-5	多尺度模块加两个单一尺度卷积层	30.56	27.67	133 888
{5, 9, 13} - {3, 5} - 5-5	两个多尺度模块加两个单一尺度卷积层	30.77	27.83	285 440
{3, 5, 7} - {3, 5} - {3, 5} - 5-5	3 个多尺度模块加两个单一尺度卷积层	30.78	27.86	424 704

由表 1 可见, 当网络加深后重建的效果比浅层网络有所提升; 然而更深的网络导致参数大量增长。现实应用中要根据实际需求和计算资源设计网络深度。

2.3 对比实验及结果分析

为了验证基于多尺度卷积神经网络在超分辨率重建算法时的有效性, 在测试集图像上进行 2×2 倍、3×3 倍和 4×4 倍的超分辨率实验。对比实验中, 有双三次插值 (BiCubic) 的经典算法、基于学习的算法 A+^[16]、基于卷积神经网络算法的 SRCNN^[3]、基于稀疏编码网络 (CSCN)^[17] 和快速重建网络 (FSRCNN)^[18]。对比实验中多尺度卷积神经网络采用 {5, 9, 13} - {3, 5} - 5-5 并引入集成预测^[19] 进行后处理。集成预测时, 图像的 90° 旋转变换仅针对像素位置进行变换, 低分辨率的图像 X_L 进行旋转得到 $X_1^L, X_2^L, \dots, X_4^L$ 输入网络重建图像, 取 90° 旋转逆变换后的高分辨率图像 $F(X_1^L), F(X_2^L), \dots, F(X_4^L)$ 的均值, 得到最终的高分辨率图像 $F(X^L)$:

$$F(X^L) = \frac{1}{4} \sum_{i=1}^4 F(X_i^L) \quad (4)$$

表 2 是算法在 Set 5、Set 14 和 BSD 300 上的对比结果。

表 2 在 Set 5、Set 14 和 BSD 300 上的 PSNR、SSIM 对比结果

Table 2 Results of PSNR and SSIM on Set 5, Set 14 and BSD 300

算法	评价方法	Set 5			Set 14			BSD 300		
		放大 2 倍	放大 3 倍	放大 4 倍	放大 2 倍	放大 3 倍	放大 4 倍	放大 2 倍	放大 3 倍	放大 4 倍
Bicubic	PSNR	33.66	30.39	28.42	30.33	27.61	26.06	29.56	27.21	25.96
	SSIM	0.929 9	0.868 1	0.810 4	0.869 4	0.775 2	0.704 2	0.843 1	0.738 4	0.667 4
A+	PSNR	35.95	32.07	29.83	31.84	28.71	27.02	30.78	28.18	26.77
	SSIM	0.950 8	0.898 4	0.846 5	0.901 7	0.812 1	0.742 1	0.877 3	0.780 8	0.708 5
SRCNN	PSNR	36.66	32.75	30.49	32.55	29.39	27.60	31.31	28.37	26.87
	SSIM	0.954 2	0.909 0	0.862 8	0.907 3	0.822 5	0.752 9	0.887 2	0.785 2	0.709 3
CSCN	PSNR	36.93	33.10	30.86	32.56	29.41	27.64	31.40	28.50	27.03*
	SSIM	0.955 2	0.914 4	0.873 2	0.907 4	0.823 8	0.757 8	0.888 4	0.788 5	0.716 1
FSRCNN	PSNR	37.00	33.16	30.71	32.63	29.43	27.59	31.50	28.52	26.96
	SSIM	0.955 8	0.914 0	0.865 7	0.908 8	0.824 2	0.753 5	0.890 6	0.789 3	0.712 8
本文算法	PSNR	37.26	33.48	30.95	32.89	29.67	27.83	31.56	28.53	26.99
	SSIM	0.957 7	0.919 5	0.875 9	0.914 1	0.828 4	0.760 2	0.892 5	0.792 4	0.716 3
本文算法	PSNR	37.43*	33.66*	31.05*	33.01*	29.78*	27.90*	31.64*	28.59*	27.02
集成预测	SSIM	0.958 3*	0.921 1*	0.879 0*	0.911 4*	0.830 0*	0.762 0*	0.893 5*	0.793 6*	0.717 5*

注: * 表示最好结果

3 个测试库中包括人物、动物、建筑等种类丰富的图像。这些图像中包含不同尺度的图像特征, 包括动物皮毛纹理、建筑外观、天空及草地等。与 SRCNN 和 FSRCNN 算法相比, 多尺度卷积神经网络重建的结果在两个客观评价指标上都有不同程度的提升。相对 SRCNN, 提升平均超过 0.2 dB。在 BSD300 放大 4×4 倍的结果略低于 CSCN 算法 0.04 dB。在结构相似度的比较上, 多尺度卷积神经网络

在3个测试库都得到了提升。客观评价结果表明多尺度卷积神经网络有效地恢复了丢失的高频信息,并且整个结果的结构性也保持的比较好。在加入集成预测后,PSNR和SSIM有了进一步的提升。图2是Head图的4×4倍超分辨率重建对比结果。

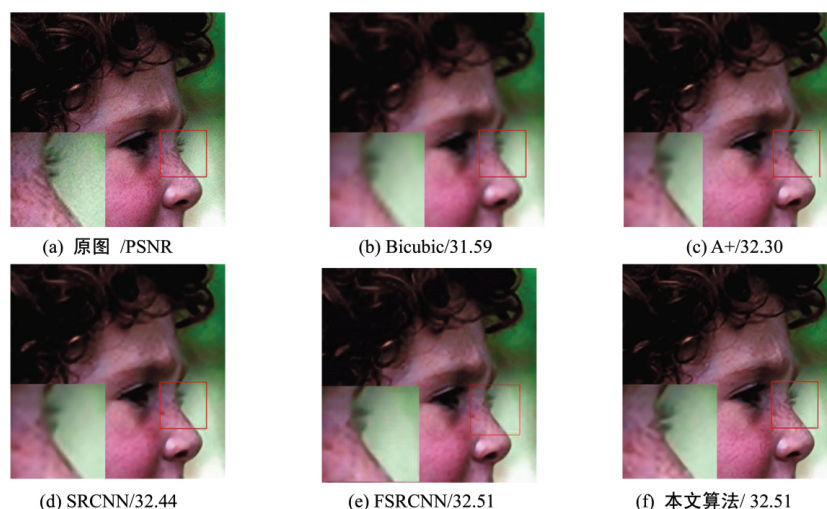


图2 Head图的4×4倍超分辨率重建对比结果

Fig. 2 Result of head with different methods with upscaling factor 4

由图2可见,传统的双三次插值和A+算法重建的图像在纹理处重建效果较差;SRCNN视觉质量有一定的提升,但在一些细节部分的重建效果依然比较模糊;FSRCNN算法重建的图像在PSNR评价上与本文算法接近。但从视觉质量上来看,本文算法所重建出的图像中睫毛处、雀斑的边缘部分更加锐利,在细节和纹理重建的效果更加清晰。

3 结论

针对图像超分辨率重建,提出多尺度卷积神经网络的图像超分辨率算法。设计的多尺度模块在卷积层利用不同尺度的卷积核提取图像的多尺度特征,更好地利用图像多尺度的邻域信息重建图像;利用1×1卷积融合多尺度特征图,不但提升网络的非线性表示,还降低了特征维度。多尺度模块可以灵活地嵌入基于CNN的超分辨率算法中,提高网络的多尺度表示能力。在多个测试库上的实验结果表明,多尺度卷积神经网络算法提升了图像重建性能,主客观重建质量均优于现有的单一尺度卷积神经网络超分辨率算法。

[参考文献]

- [1] NASROLLAHI K, MOESLUND T B. Super-resolution: a comprehensive survey [J]. Machine Vision and Applications, 2014, 25(6): 1 423-1 468.
- [2] DONG C, LOY C C, HE K, et al. Learning a deep convolutional network for image super-resolution [C] // European Conference on Computer Vision. [S.L.]: Springer, 2014: 184-199.
- [3] 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.
- [4] KIM J, LEE J K, LEE K M. Accurate image super-resolution using very deep convolutional networks [C/OL] // 2016 IEEE Conference on Computer Vision and Pattern Recognition. 2016: 1-9 [2019-02-28]. <https://arxiv.org/pdf/1511.04587v1.pdf>.
- [5] KIM J, LEE J K, LEE K M. Deeply-recursive convolutional network for image super-resolution [C/OL] // 2016 IEEE Conference on Computer Vision and Pattern Recognition. 2016: 1-9 [2019-02-28]. <https://arxiv.org/pdf/1511.04491.pdf>.
- [6] LIM B, SON S, KIM H, et al. Enhanced deep residual networks for single image super-resolution [C/OL] // 2017 IEEE

- Conference on Computer Vision and Pattern Recognition Workshops. IEEE Computer Society, 2017: 136-144 [2019-02-28].
https://arxiv.org/pdf/1707.02921.pdf.
- [7] HARIS M, SHAKHNAROVICH G, UKITA N, et al. Deep back-projection networks for super-resolution [J]. Computer Vision and Pattern Recognition, 2018: 1 664-1 673. doi: 10.1109/CVPR.2018.00179.
- [8] TIMOFTE R, GU S, WU J, et al. NTIRE 2018 challenge on single image super-resolution: methods and results [C] // 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops. IEEE Computer Society, 2017: 852-863. doi: 10.1109/CVPRW.2017.150.
- [9] ADELSON E H, SIMONCELLI E P, FREEMAN W T. Pyramids and multiscale representations [C] // Proc European Conf. Visual Percep. Cambridge: Cambridge University Press, 1991: 3-16.
- [10] MAIRAL J, SAPIRO G, ELAD M. Learning multiscale sparse representations for image and video restoration [J]. Multiscale Modeling & Simulation, 2008, 7(1): 214-241.
- [11] SZEGEDY C, LIU W, JIA Y, et al. Going deeper with convolutions [C] // 2015 IEEE Conference on Computer Vision and Pattern Recognition. IEEE Computer Society, 2015: 1-9.
- [12] LECUN Y, BOTTOU L, BENGIO Y, et al. Gradient-based learning applied to document recognition [J]. Proceedings of the IEEE, 1998, 86(11): 2 278-2 324.
- [13] YANG J, WRIGHT J, HUANG T S, et al. Image super-resolution via sparse representation [J]. IEEE Transactions on Image Processing, 2010, 19(11): 2 861-2 873.
- [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 [C] // Proceedings Eighth IEEE International Conference on Computer Vision. IEEE, 2001, 2: 416-423. doi: 10.1109/ICCV.2001.937655.
- [15] VEDALDI A, LENC K. Matconvnet: convolutional neural networks for matlab [C] // Proceedings of the 23rd ACM International Conference on Multimedia. ACM, 2015: 689-692. doi: 10.1145/2733373.2807412.
- [16] TIMOFTE R, DE SMET V, VAN GOOL L. A+: Adjusted anchored neighborhood regression for fast super-resolution [C] // Asian Conference on Computer Vision, 2014: 111-126 [2019-02-28]. doi: 10.1007/978-3-319-16817-3_8.
- [17] WANG Z, LIU D, YANG J, et al. Deep networks for image super-resolution with sparse prior [C/OL] // Proceedings of the IEEE International Conference on Computer Vision. 2015: 370-378 [2019-02-28]. https://arxiv.org/pdf/1507.08905v4.pdf.
- [18] DONG C, LOY C C, TANG X. Accelerating the super-resolution convolutional neural network [C] // European Conference on Computer Vision, 2016: 391-407. doi: 10.1007/978-3-319-46475-6_25.
- [19] TIMOFTE R, ROTHE R, VAN GOOL L. Seven ways to improve example-based single image super resolution [C] // 2016 IEEE Conference on Computer Vision and Pattern Recognition, 2016: 1 865-1 873. doi: 10.1109/CVPR.2016.206.

Multi-Scale Convolutional Neural Network for Super-Resolution Imaging

HE Yifan, LIN Yimin, LIN Jiamin, DU Xiaofeng

(School of Computer & Information Engineering, Xiamen University of Technology, Xiamen 361024, China)

Abstract: To learn multi-scale information for single super-resolution imaging, this paper proposed a multi-scale convolutional neural network for super-resolution imaging reconstruction. Different convolution kernels were introduced to extract image features, thus providing different sizes of image context for image reconstruction. The bottleneck layer performed fusion of multi-scale information, improved the non-linear representation, and reduced the dimensionality of the output of the intermediate layer. Experimental results of several test sets demonstrate the multi-scale convolutional neural network method outperforms many state-of-the-art single super-resolution imaging methods.

Key words: image reconstruction; super-resolution imaging algorithm; convolutional neural network; multi-scale

(责任编辑 宋 静)