软件学报 ISSN 1000-9825, CODEN RUXUEW
Journal of Software,2018,29(4):926−934 [doi: 10.13328/j.cnki.jos.005407]
©中国科学院软件研究所版权所有.
E-mail: jos@iscas.ac.cn
http://www.jos.org.cn
Tel: +86-10-62562563

基于深度反卷积神经网络的图像超分辨率算法*

彭亚丽[1,2,3], 张 鲁[1,2], 张 钰[1,2], 刘侍刚[1,2], 郭 敏[2,3]

[1](现代教学技术教育部重点实验室(陕西师范大学),陕西 西安 710062)
[2](陕西省教学信息技术工程实验室(陕西师范大学),陕西 西安 710119)
[3](陕西师范大学 计算机科学学院,陕西 西安 710119)
通讯作者: 张钰, E-mail: flyantzhang@gmail.com

摘 要: 图像超分辨率一直是底层视觉领域的研究热点,现有基于卷积神经网络的方法直接利用传统网络模型,未对图像超分辨率属于回归问题这一本质进行优化,其网络学习能力较弱,训练时间较长,重建图像的质量仍有提升空间.针对这些问题,提出了基于深度反卷积神经网络的图像超分辨率算法,该算法利用反卷积层对低分辨率图像进行上采样处理,再经深度映射消除由反卷积层造成的噪声和伪影现象,使用残差学习降低网络复杂度,同时避免了因网络过深导致的网络退化问题.在 Set 5、Set 14 等测试集中,所提算法的 PSNR、SSIM、IFC 这 3 项评价指标都优于 FSRCNN,重建图像的视觉效果同样验证了该算法出色的性能.

关键词: 卷积神经网络;图像超分辨率;深度映射;上采样
中图法分类号: TP391

中文引用格式: 彭亚丽,张鲁,张钰,刘侍刚,郭敏.基于深度反卷积神经网络的图像超分辨率算法.软件学报,2018,29(4): 926−934. http://www.jos.org.cn/1000-9825/5407.htm

英文引用格式: Peng YL, Zhang L, Zhang Y, Liu SG, Guo M. Deep deconvolution neural network for image super-resolution. Ruan Jian Xue Bao/Journal of Software, 2018,29(4):926−934 (in Chinese). http://www.jos.org.cn/1000-9825/5407.htm

Deep Deconvolution Neural Network for Image Super-Resolution

PENG Ya-Li[1,2,3], ZHANG Lu[1,2], ZHANG Yu[1,2], LIU Shi-Gang[1,2], GUO Min[2,3]

[1](Key Laboratory of Modern Teaching Technology, Ministry of Education (Shaanxi Normal University), Xi’an 710062, China)
[2](Engineering Laboratory of Teaching Information Technology of Shaanxi Province (Shaanxi Normal University), Xi’an 710119, China)
[3](School of Computer Science, Shaanxi Normal University, Xi’an 710119, China)

* 基金项目: 国家自然科学基金(61672333, 61741208, 61402274, 61772325); 陕西省工业科技攻关项目(2016GY-081); 教育部高等教育司产学合作协同育人项目(201701023062); 陕西省自然科学基金(2017JQ6074); 陕西省农业攻关项目(2016NY-176); 陕西省重点科技创新团队计划(2014KTC-18); 教育部科技发展中心“云数融合科教创新”基金(2017A07053); 陕西师范大学学习科学交叉学科培育计划; 中央高校基本科研业务费专项资金(2017CSY024, GK201603091, GK201703054)

Foundation item: National Natural Science Foundation of China (61672333, 61741208, 61402274, 61772325); Key Science and Technology Program of Shaanxi Province of China (2016GY-081); Industry University Cooperative Education Project of Higher Education Department of the Ministry of Education (201701023062); Natural Science Foundation of Shaanxi Province, China (2017JQ6074); Science Research and Development Program of Shaanxi Province of China (2016NY-176); Program of Key Science and Technology Innovation Team in Shaanxi Province (2014KTC-18); Fund for Integration of Cloud Computing and Big Data of Science and Technology Development Center of the Ministry of Education (2017A07053); Interdisciplinary Incubation Project of Learning Science of Shaanxi Normal University; Fundamental Research Funds for the Central Universities (2017CSY024, GK201603091, GK201703054)

本文由“多媒体大数据处理与分析”专题特约编辑赵耀教授、李波教授、华先胜研究员、文继荣教授、蒋刚毅教授、常冬霞副教授推荐.

收稿时间: 2017-04-29; 修改时间: 2017-06-26; 采用时间: 2017-10-13; jos 在线出版时间: 2017-12-01

CNKI 网络优先出版: 2017-12-04 06:46:54, http://kns.cnki.net/kcms/detail/11.2560.TP.20171204.0646.008.html

Abstract: Image super resolution is a research hot spot in the field of low level vision. The existing methods based on convolutional neural network do not optimize the image super resolution as a regression problem. These methods are weak in learning ability and require too much time in training step, also leaving room for improvement in the quality of image reconstruction. To solve above mentioned problems, this article proposes a method based on deep deconvolution neural network, which first upsamples low resolution image by deconvolution layer, and then uses deep mapping to eliminate the noise and artifacts caused by deconvolution layer. The residual learning reduces the network complexity and avoids the network degradation caused by the depth network. In Set 5, Set 14 and other datasets, the presented method performs better than FSRCNN in PSNR, SSIM, IFC and visual.

Key words: convolutional neural network; image super-resolution; deep mapping; upsampling

分辨率是评测图像质量的重要指标,高分辨率图像具有高像素密度,包含更多的细节,这些细节在应用中起关键作用.分辨率取决于成像系统和其他因素的影响,图像超分辨率(image super-resolution,简称 SR)是一种克服图像传感器等成像硬件固有限制、改善图像质量的技术,即借助计算机软件从一幅或多幅低分辨率(low resolution,简称 LR)图像中重构高分辨率(high resolution,简称 HR)图像.按照输入图像的数量,SR 可以分为两类:单帧图像超分辨率(single image super resolution,简称 SISR)[1,2],仅利用一张图像提高图像的分辨率;多帧图像超分辨率(multiple image super resolution,简称 MISR)[3],通过给定的同一场景的多张图像重建高分辨率图像.

图像超分辨率作为一种图像复原技术,目的是从退化的低分辨率图像 $Y=HX$(其中,H 是下采样)中恢复潜在的高分辨率图像 $X=F(Y)$.当前,$F(Y)$可分为 3 个研究方向:基于插值(interpolation based)、基于重建(reconstruction based)和基于学习(learning based)的方法.基于插值的方法如最近邻域插值(nearest neighbor interpolation)和双三次插值(bicubic interpolation)[4],这些方法能够有效地增强图像分辨率,但模糊了图像边缘;基于重建的方法旨在重建退化图像丢失的高频信息,如 Yang 等人提出的迭代反向投影(iterative back projection,简称 IBP)[5]算法,该类方法简单且计算量低,但无法处理复杂结构的图像;基于学习的方法认为,低分辨率的图像已经拥有预测其所对应的高分辨率部分的信息,通过对大量样本的学习建立低分辨率图像和高分辨率图像之间的映射关系.典型的基于学习的方法,如 Chang 等人提出的邻域嵌入(neighbor embedding)算法[6]、Yang 等人提出的稀疏编码(sparse coding)算法[7]以及基于卷积神经网络(convolutional neural network,简称 CNN)[8,9]的算法等.目前基于学习的方法已成为 SISR 的主流研究方向,而基于插值和基于重建的方法则作为基于学习方法的辅助处理过程.

本文以基于学习的方法为研究对象,提出了一种基于深度反卷积神经网络的图像超分辨率重建算法,本文算法主要具有以下特点:

(1) 直接利用卷积神经网络建立端到端的映射模型,无需对图像预处理;

(2) 采用残差学习和深度网络结构,有效地提高了网络学习能力,同时降低了训练时间;

(3) 采用 ADAM 替代传统的 SGD 优化方法,并在训练过程中降低学习率,算法性能提升效果明显;

(4) 可以处理复杂结构的图像,在 Set 5、Set 14 等测试集中的客观量化结果和重建图像的视觉效果优于 FRSCNN 等算法.

1 相关工作

卷积神经网络[10]是一种为了处理二维输入数据而设计的多层人工神经网络,基于 CNN 的图像超分辨率算法利用 CNN 在图像处理上的优势:(1) 多维数据可以直接输入网络训练;(2) 局部权值共享的特点可以有效地降低训练参数;(3) 强大的特征学习能力和建模能力.基于 CNN 的算法是当前图像超分辨率的研究热点,如 Dong 等人提出的 SRCNN[8]首先利用双三次插值预处理低分辨率图像,再经 3 层网络训练,包括特征提取(patch extraction & representation)、非线性映射(non-linear mapping)和重建(reconstruction)3 个部分.虽然 SRCNN 成功地将 CNN 引入图像超分辨率问题,但仍有很多不足:如训练时间太长、网络学习能力较弱等.Dong 等人之后又提出了改进算法 FSRCNN[9],该算法舍去了图像预处理阶段,直接利用 CNN 重建高分辨率图像,增加了网络深度并在网络末端加入反卷积层,算法精度和速度都有大幅度提升.

在数字图像处理领域,上采样被广泛用于改善图像的质量,上采样运算借助低分辨率图像中的像素构建高

© 中国科学院软件研究所 http://www.jos.org.cn

分辨图像中更大的区域,同时观测数据中的噪声也会被放大.反卷积(deconvolution)作为一种上采样运算最早出现在 Bengio 等人的工作中[11],在其之后的工作中反卷积被用作一个网络层,通过构建反卷积网络实现了网络的可视化[12].反卷积网络可以学习更抽象的特征,也被应用于语义分割(semantic segmentation)[13]等.

相关文献[14]证实,深度网络可以学习到更多的特征,并且能够促进特征的重复利用,很多计算机视觉问题都得益于较深的网络模型,如目标检测[15]、图像分类[16]等.虽然多数深度网络可以收敛,但会出现退化现象[17,18],即随着网络深度的增加,准确率增长速度会趋于饱和并有所下降[19].Zhang 等人提出的残差学习(residual learning,简称 RL)[20]有效地解决了网络退化问题,ILSVRC 和 COCO 的比赛结果表明,在深度网络中使用残差结构可以降低网络复杂度,促进网络收敛.

2 本文算法

传统的基于 CNN 的方法大多先经卷积层进行特征提取,再通过改变卷积核和网络层实现不同的应用,SRCNN 和 FSRCNN 也都借鉴了这种设计思路.虽然 FSRCNN 以更深的网络结构实现对 SRCNN 的改进,在实验中我们仍发现了 FSRCNN 的几点不足:(1) 直接使用传统网络模型,未对图像超分辨率属于回归问题这一本质进行优化;(2) 加深了网络却没有考虑到网络退化问题,且需要过多的训练时间;(3) 仍然使用传统优化方法,网络性能仍有提升空间.针对这些问题,我们从低分辨率图像和高分辨率图像映射的角度出发,重新设计了网络结构,并选择了更先进的 ADAM 优化方法,通过研究,本文提出基于深度反卷积神经网络的图像超分辨率算法(deep deconvolution neural network for image super resolution,简称 DDSR),DDSR 的网络结构仅由上采样和深度映射两部分构成,如图 1 所示.

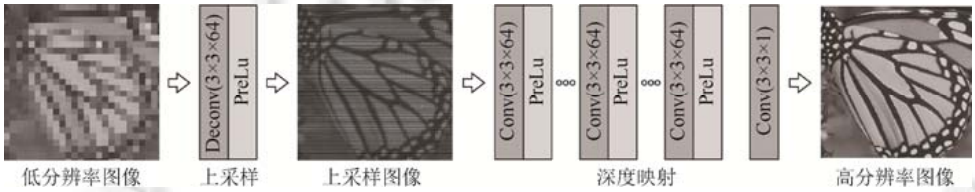

Fig.1 The architecture of the proposed DDSR network

图 1 本文算法(DDSR)的网络结构

2.1 上采样

SRCNN 借助双三次差值对图像预处理,性能低且运算速度慢,而直接利用反卷积层对低分辨率图像上采样不仅省去了预处理阶段,还可以学习到更多的特征.不同于传统网络先经卷积层进行特征提取,本文算法的网络结构是先通过反卷积层进行上采样操作.上采样结构如图 2 所示,低分辨率图像经零填充(灰色像素为低分辨率图像,白色像素为零填充)与反卷积核作用得到上采样图像.

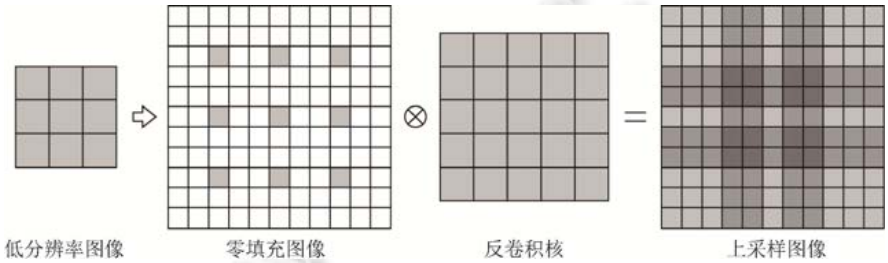

Fig.2 The structure of upsampling

图 2 上采样结构

$$F_1(Y)=\max(W_1\times Y+B_1,0)+a_j\min(0,W_1\times Y+B_1) \tag{1}$$

上式中,Y 表示低分辨率输入图像,参数 W_1 和 B_1 分别是上采样的反卷积核(权重矩阵)和偏置量,$F_1(Y)$表示上

© 中国科学院软件研究所 http://www.jos.org.cn

采样图像.这里的反卷积可以视为卷积的逆运算,步长设置为 k(反卷积层参数 stride)输出图像实现 k 倍放大.本文分别对 3 种尺寸的反卷积核进行测试,包括 3×3×64、5×5×64、7×7×64($f\times f\times n$,其中,f 为卷积核大小,n 为卷积核的数量),实验结果见表 1(深度映射为 18 层,图像放大 3 倍),7×7×64 的卷积核 PSNR 最高,但仅比 3×3×64 高 0.04dB(Set 5)和 0.05dB(Set 14),而重建图像所需时间相差不多,考虑到在网络深度相同的情况下较小的卷积核所需的训练时间更少,因此,上采样结构最终选用尺寸为 3×3×64 的反卷积核.

Table 1 The average PSNR and run time of three different filters

表 1 3 种反卷积核的平均 PSNR 和平均运行时间测试结果

Datasets	3×3×64		5×5×64		7×7×64	
	PSNR (dB)	Time (s)	PSNR (dB)	Time (s)	PSNR (dB)	Time (s)
Set 5	33.32	0.40	33.33	0.41	33.36	0.40
Set 14	29.57	0.39	29.57	0.40	29.62	0.42

虽然反卷积可以有效地提高图像分辨率,但会产生不均匀的伪影现象,即当反卷积核移动的步长小于边长时,会出现反卷积核与输入数据作用区域上的重叠,并且这种重叠在二维数据中更为极端,表现为不同幅度的“棋盘”状图案(如图 1 中的上采样图像所示),图 3 所示为低分辨率图像经上采样的可视化图像.为了得到更加清晰的重建图像,我们在上采样后加入深度映射结构,用来消除噪声和抑制伪影现象.

Fig.3 The output image of upsampling

图 3 上采样结构的可视化图像

2.2 深度映射

文献[17]分析了网络深度与卷积核大小间的关系,即用较小的卷积核替代较大的卷积核,同时增加网络深度可以有效地提高网络的特征学习能力,相关实验结果表明,当时间复杂度相同时,网络深度比卷积核大小更重要.然而,随着网络的加深,会出现准确率增长饱和甚至下降的现象,同时网络训练时间也随之增加.为了避免出现这些问题,深度映射采用带有残差的网络结构,残差学习认为,直接利用深度网络去学习潜在的恒等映射关系 $H(Y)=F_1(Y)$比较困难,如果将网络转化为 $H(Y)=F_1(Y)+F_n(Y)$(其中,$F_1(Y)$为上采样输出图像,$F_n(Y)$为第 N 层深度映射输出图像),则残差函数为 $F_n(Y)=H(Y)-F_1(Y)$,令 $F_n(Y)=0$ 得到恒等映射,相比之下,利用深度网络学习残差更加简单.对于同等深度的网络,残差网络收敛更快且更容易优化,我们将残差学习引入到图像超分辨率问题,与大多数残差网络不同,DDSR 的深度映射结构仅使用单个残差单元.

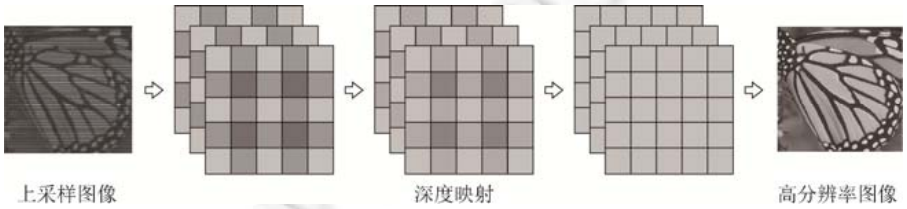

Fig.4 The structure of deep mapping

图 4 深度映射结构

$$F_i(Y)=\max(W_i\times F_1(Y)+B_i,0)+a_j\min(0,W_i\times F_1(Y)+B_i) \tag{2}$$

上式中,参数 W_i 和 $B_i(i=2,3,4,\ldots)$分别表示深度映射中各层的卷积核和偏置量,$F_i(Y)(i=2,3,4,\ldots)$表示每个卷

© 中国科学院软件研究所 http://www.jos.org.cn

积层的输出图像.为了确保经过卷积层的图像大小不变,每个卷积层都进行零填充(如 3×3 大小的卷积核,参数 padding 设置为 1).深度映射结构如图 4 所示,受 VGG-Net[16]的启发,深度映射由许多相同的卷积层组成,卷积核尺寸为 3×3×64,最后一层为 3×3×1 输出重建图像的 Y 通道.我们分别对有残差的深度映射结构和没有残差的深度映射结构进行测试,同时对比 ADAM(adaptive moment estimation)[21]和 SGD(stochastic gradient descent)两种优化方法,实验结果如图 5 所示(Set 5 数据集,深度映射为 18 层,图像放大 3 倍).ADAM 和 SGD 两种优化方法都能保证网络收敛,但 ADAM 性能更优,比 SGD 分别高 0.61dB 和 0.63dB(RL).如 ADAM+RL 和 SGD+RL 迭代曲线所示(图 5 中实线),残差学习可以避免震荡,使网络更快地收敛,且比没有残差结构的网络平均 PSNR 高 0.09dB (ADAM)和 0.07dB(SGD).

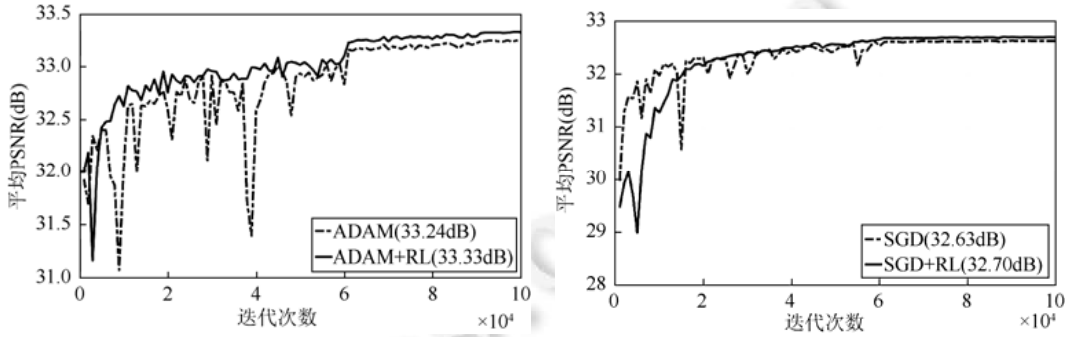

Fig.5 The results of average PSNR with the number of iterations: ADAM (left) and SGD (right)

图 5 平均 PSNR 随迭代次数变化曲线图:ADAM(左)和 SGD(右)优化算法

2.3 激活函数和损失函数

DDSR 各层激活函数采用 PReLU(parametric rectified linear unit)[22]:$f(x_j)=\max(x_j,0)+a_j \min(0,x_j)$,即带修正参数的 ReLU[23]激活函数,相比 ReLU 激活函数,PReLU 激活函数仅增加少量计算量便实现了更高的准确率,并且可以避免 ReLU 造成的 dead features 现象[12].

重建图像与同尺寸高分辨率图像的均方误差(mean squared error,简称 MSE)作为损失函数(loss function):

$$L(\theta)=\frac{1}{n}\sum_{k=1}^{n}\left\|F(Y_k;\theta)-X_k\right\|^2,$$

其中,$F(Y_k;\theta)$ 和 X_k 分别为重建图像和原始高分辨率图像.通过最小化 $L(\theta)$ 经反向传播修正各层参数 $\theta=\{W_1, W_i, B_1, B_i\}$,得到最佳映射模型.

3 实 验

实验训练数据集为 General-100[9](共 100 张)+91 images[24](共 91 张)+BSD 100[25](共 100 张),共 291 张不同的彩色图像,通过将每张图像裁剪为更小的图像扩充训练集(例如图像大小为 7×7,步长为 4,可以得到 252 032 张图像),用于测试的数据集为 Set 5[26](共 5 张)、Set 14[27](共 14 张)、BSD 200[25](共 200 张),皆为通用数据集.为了与现有图像超分辨算法作对比,本文仅对图像的 Y 通道(YCrCb 色彩空间,其中,Y 通道表示亮度)进行训练和测试.

DDSR 采用 ADAM 优化方法,参数 beta 1 和 beta 2 分别设置为 0.9 和 0.999.本文的实验结果皆为 100 000 次迭代(约 50 个 epoch),初始学习率为 1e-3,共下降 2 次(其中,epoch=1~30,学习率为 1e-3;epoch=31~45,学习率为 1e-4;epoch=45~50,学习率为 1e-5).每次实验需要训练约 3 个小时.

实验环境为基于 Windows 10 操作系统的 Caffe[28]深度学习框架,该框架支持 GPU 运算,配置 NVIDIA CUDA 8.0+cuDNN-V5.1 深度学习库加速 GPU 运算;用于测试的软件为 Matlab 2017a;用于训练和测试的硬件配置为:Intel Xeon E3-1230 v5 处理器,Nvidia GeForce GTX 1080 Ti 显卡,Kingston DDR4 64GB 内存.

© 中国科学院软件研究所 http://www.jos.org.cn

现有图像超分辨率的评测标准包括主观评价和客观量化,主观评价即人眼观测输出图像,评定图像的质量.目前最普遍的客观量化方法包括:峰值信噪比(peak signal to noise ratio,简称 PSNR),两幅图像间的 PSNR 值(单位 dB)越高,则重建图像相对于高分辨率图像失真得越少;结构相似性(structural similarity index method,简称 SSIM)[29],SSIM 值越趋近于 1,则两幅图相似度越高;信息保真标准(information fidelity criterion,简称 IFC)[30]的判断方法与 PSNR 相同.

3.1 实验数据

为了验证网络深度对本文算法性能的影响,我们首先对 6 层、12 层和 18 层的深度映射进行测试,实验结果见表 2,以增加少量时间复杂度为代价(图像平均重建时间受实验软件和硬件的影响,此处仅作参考),对于不同测试集,本文算法的平均 PSNR 和 SSIM 随网络深度的增加而提高.由图 6 所示的迭代曲线可见,通过改变学习率可以有效地促进网络收敛,在 60 000 次迭代处(黑色方框)学习率由 1e-3 降为 1e-4,平均 PSNR 提高 0.13dB;在 90 000 次迭代处学习率由 1e-4 降为 1e-5,平均 PSNR 提高 0.03dB.SSIM 评价指标可以得到相同的结论.

Table 2 The average PSNR (dB), SSIM and run time (s) of different DDSR

表 2 不同网络深度的平均 PSNR(dB)、平均 SSIM 和平均运行时间(s)测试结果

Datasets	Upscaling factor	DDSR 6			DDSR 12			DDSR 18		
		PSNR	SSIM	Time	PSNR	SSIM	Time	PSNR	SSIM	Time
Set 5	X3	33.09	0.914 9	0.26	33.23	0.916 3	0.29	33.33	0.917 4	0.38
Set 14	X3	29.41	0.827 2	0.20	29.50	0.828 7	0.28	29.57	0.829 7	0.37
BSD 200	X3	29.08	0.813 3	0.18	29.13	0.814 6	0.28	29.17	0.815 5	0.39

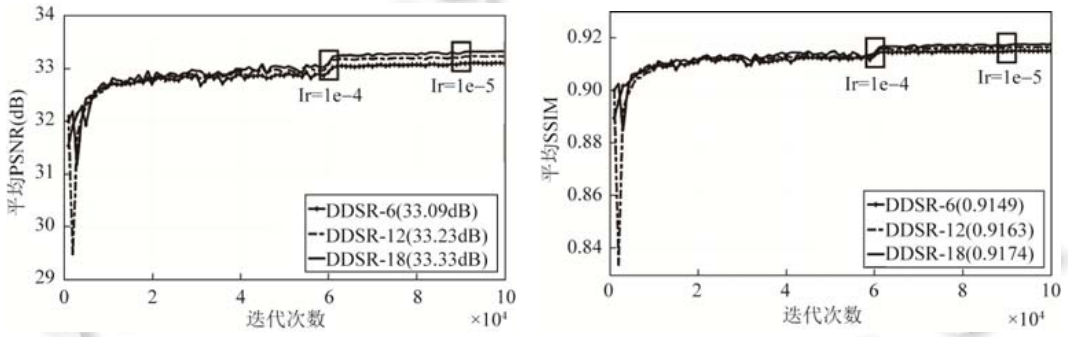

Fig.6 The results of average PSNR (left) and SSIM (right) with the number of iterations

图 6 平均 PSNR(左)和平均 SSIM(右)随迭代次数变化曲线图

根据上述实验结果,我们以 18 层深度映射的 DDSR 与现有的图像超分辨率算法进行对比,包括 Bicubic、A+[31]、SRF[32]、SRCNN[8]和 FSRCNN[9]这 5 种算法,每种算法分别对图像放大 2 倍、3 倍、4 倍,采用 PSNR、SSIM、IFC 这 3 种图像质量评价方法,测试集为 Set 5、Set 14、BSD 200.实验结果见表 3,本文算法的 PSNR 和 SSIM 优于其他算法,但在 IFC 评价指标上略低于 SRF 算法.

Table 3 The average PSNR (dB), SSIM, IFC of different image super-resolution methods

表 3 不同图像超分辨率算法的平均 PSNR(dB)、SSIM、IFC 的对比结果

算法	放大倍数	2			3			4		
	数据集	Set 5	Set 14	BSD 200	Set 5	Set 14	BSD 200	Set 5	Set 14	BSD 200
Bicubic	PSNR	33.66	30.23	29.70	30.39	27.54	27.26	28.42	26.00	25.97
	SSIM	0.93	0.87	0.86	0.87	0.77	0.76	0.81	0.70	0.69
	IFC	6.10	6.09	5.70	3.10	3.41	3.19	2.35	2.23	2.04
A+[31]	PSNR	36.55	32.28	31.44	32.59	29.13	28.36	30.28	27.32	26.83
	SSIM	0.95	0.91	0.90	0.91	0.82	0.81	0.86	0.75	0.74
	IFC	8.48	8.11	7.49	4.84	4.45	4.07	3.26	2.74	2.62
SRF[32]	PSNR	36.87	32.51	31.65	32.71	29.23	28.45	30.35	27.41	26.89
	SSIM	0.96	0.91	0.91	0.91	0.82	0.81	0.86	0.75	0.74
	IFC	8.63	8.22	7.60	4.90	4.49	4.11	3.26	2.94	2.62

© 中国科学院软件研究所 http://www.jos.org.cn

Table 3 The average PSNR (dB), SSIM, IFC of different image super-resolution methods (Continued)

表 3 不同图像超分辨率算法的平均 PSNR(dB)、SSIM、IFC 的对比结果(续)

算法	放大倍数	2			3			4		
	数据集	Set 5	Set 14	BSD 200	Set 5	Set 14	BSD 200	Set 5	Set 14	BSD 200
SRCNN[8]	PSNR	36.34	32.18	31.38	32.39	29.00	28.28	30.09	27.20	26.73
	SSIM	0.95	0.90	0.93	0.90	0.81	0.80	0.85	0.74	0.73
	IFC	7.54	7.22	6.80	4.25	3.96	3.67	2.86	2.60	2.37
FSRCNN[9]	PSNR	37.00	32.63	31.80	33.16	29.43	28.60	30.71	27.59	26.98
	SSIM	0.96	0.91	0.91	0.91	0.82	0.81	0.87	0.75	0.74
	IFC	8.06	7.71	7.25	4.88	4.47	4.11	3.01	2.70	2.41
DDSR	PSNR	37.10	32.65	32.35	33.33	29.57	29.17	31.25	28.02	27.60
	SSIM	0.96	0.91	0.91	0.92	0.83	0.82	0.88	0.77	0.75
	IFC	8.22	7.92	7.12	4.83	4.42	3.83	3.39	3.02	2.53

3.2 重建图像

用于测试的 4 幅图像分别为 12003(BSD 200)、monarch(Set 14)、bird(Set 5)和 comic(Set 14),这里仅展示 Bicubic、FSRCNN、DDSR 这 3 种算法的彩色重建图像,其中,Y 通道由 DDSR 重建,Cr 通道(红色与亮度值的差异)和 Cb 通道(蓝色与亮度值的差异)利用双三次差值重建.放大图像 3 倍效果如图 7 所示,本文算法的重建图像视觉效果最好,最接近原图.如图像 monarch 和 comic,DDSR 的重建图像边缘最清晰;图像 bird,DDSR 的重建图像噪声最少;但对于图像 12003,DDSR 和 FSRCNN 相差不多,改善效果并不明显.

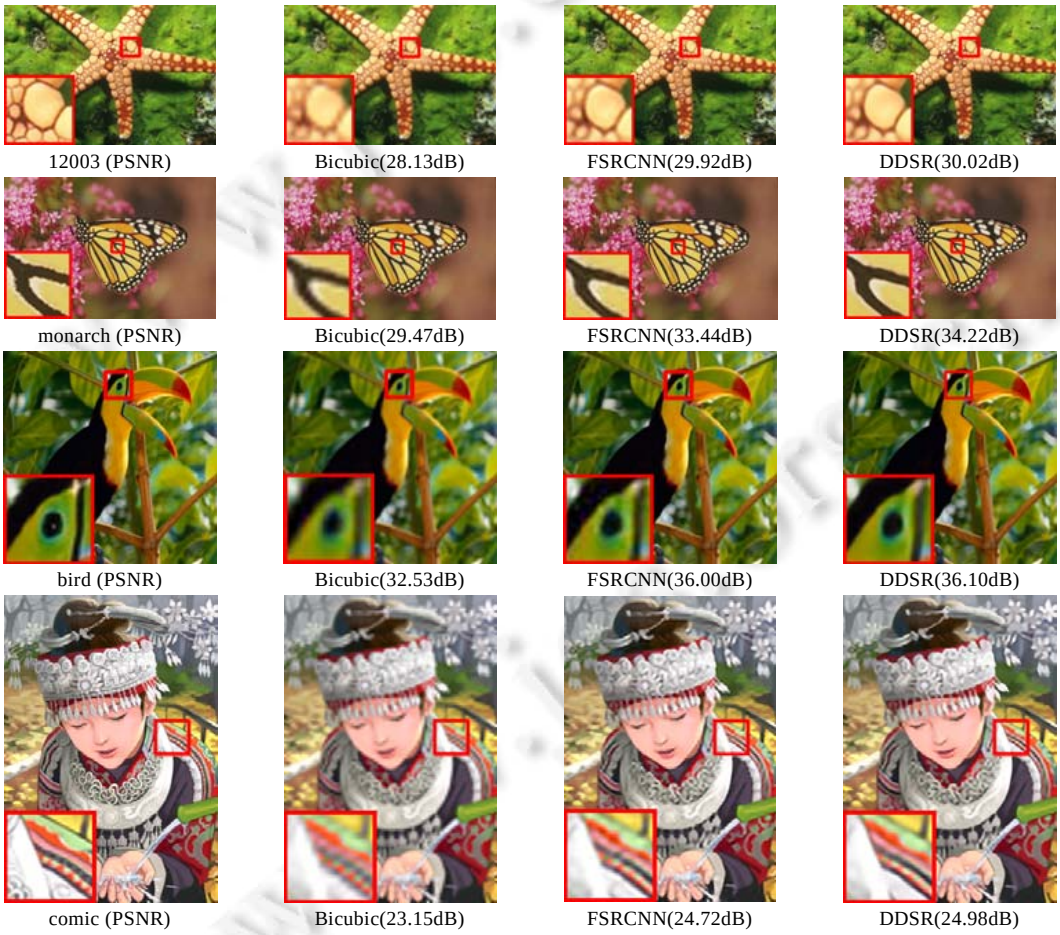

Fig.7 The reconstruction images (X3) of different image super-resolution methoods

图 7 不同超分辨率算法放大图像 3 倍对比

© 中国科学院软件研究所 http://www.jos.org.cn

4 总 结

本文提出了一种基于深度反卷积神经网络的图像超分辨率算法,该算法借助反卷积层对低分辨率图像上采样,对于上采样造成的噪声和伪影现象,利用带有残差的深度映射进行消除,我们对 SGD 和 ADAM 两种优化方法进行测试,实验结果表明,采用 ADAM 优化方法,并在训练过程中降低学习率可以有效地促进网络收敛,本文算法的 PSNR、SSIM、IFC 这 3 项评价指标和重建图像的视觉效果皆优于 FSRCNN.下一步,我们会以提高算法精确度、降低时间复杂度作为研究重点,并将本文算法扩展至其他图像复原问题.

References:

[1] Zhang K, Tao D, Gao X, Li X, Li J. Coarse-to-Fine learning for single-image super-resolution. IEEE Trans. on Neural Networks and Learning Systems, 2017,28(5):1109–1122. [doi: 10.1109/TNNLS.2015.2511069]

[2] Zeng K, Yu J, Wang R, Li C, Tao D. Coupled deep autoencoder for single image super-resolution. IEEE Trans. on Cybernetics, 2017,47(1):27–37. [doi: 10.1109/TCYB.2015.2501373]

[3] Chen C, Liang H, Zhao S, Lyu Z, Sarem M. A novel multi-image super-resolution reconstruction method using anisotropic fractional order adaptive norm. The Visual Computer, 2015,31(9):1217–1231. [doi: 10.1007/s00371-014-1007-5]

[4] Francesc A. A nonlinear algorithm for monotone piecewise bicubic interpolation. Applied Mathematics and Computation, 2016,272: 100–113. [doi: 10.1016/j.amc.2015.08.027]

[5] Yang X, Zhang Y, Zhou D, Yang R. An improved iterative back projection algorithm based on ringing artifacts suppression. Neurocomputing, 2015,162:171–179. [doi: 10.1016/j.neucom.2015.03.055]

[6] Fang B, Huang Z, Li Y, Wang Y. υ-Support vector machine based on discriminant sparse neighborhood preserving embedding. Pattern Analysis and Applications, 2017,20(4):1077–1089. [doi: 10.1007/s10044-016-0547-x]

[7] Yang J, Wang Z, Lin Z, Huang T. Coupled dictionary training for image super-resolution. IEEE Trans. on Image Processing, 2012, 21(8):3467–3478. [doi: 10.1109/TIP.2012.2192127]

[8] Dong C, Loy CC, He K, Tang X. Image super-resolution using deep convolutional networks. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2016,38(2):295–307. [doi: 10.1109/TPAMI.2015.2439281]

[9] Dong C, Loy CC, Tang X. Accelerating the super-resolution convolutional neural network. In: Proc. of the European Conf. on Computer Vision. Springer-Verlag, 2016. 391–407. [doi: 10.1007/978-3-319-46475-6_25]

[10] LeCun Y, Bengio Y, Hinton GE. Deep learning. Nature, 2015,521:436–444. [doi: 10.1038/nature14539]

[11] Bengio Y, Courville A, Vincent P. Representation learning: A review and new perspectives. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2013,35(8):1798–1828. [doi: 10.1109/TPAMI.2013.50]

[12] Fakhry A, Zeng T, Ji S. Residual deconvolutional networks for brain electron microscopy image segmentation. IEEE Trans. on Medical Imaging, 2017,36(2):447–456. [doi: 10.1109/TMI.2016.2613019]

[13] Wei Y, Liang X, Chen Y, Shen X, Cheng MM, Feng JS, Zhao Y, Yan SC. STC: A simple to complex framework for weakly-supervised semantic segmentation. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2017,39(11):2314–2320. [doi: 10.1109/TPAMI.2016.2636150]

[14] Du B, Wang Z, Zhang L, Zhang L, Tao D. Robust and discriminative labeling for multi-label active learning based on maximum correntropy criterion. IEEE Trans. on Image Processing, 2017,26(4):1694–1707. [doi: 10.1109/TIP.2017.2651372]

[15] Li J, Mei X, Prokhorov D, Tao D. Deep neural network for structural prediction and lane detection in traffic scene. IEEE Trans. on Neural Networks and Learning Systems, 2017,28(3):690–703. [doi: 10.1109/TNNLS.2016.2522428]

[16] Qu Y, Lin L, Shen F, Lu C, Wu Y, Xie Y, Tao D. Joint hierarchical category structure learning and large-scale image classification. IEEE Trans. on Image Processing, 2017,26(9):4331–4346. [doi: 10.1109/TIP.2016.2615423]

[17] Du B, Xiong W, Wu J, Zhang L, Zhang L, Tao D. Stacked convolutional denoising auto-encoders for feature representation. IEEE Trans. on Cybernetics, 2017,47(4):1017–1027. [doi: 10.1109/TCYB.2016.2536638]

[18] Shelhamer E, Long J, Darrell T. Fully convolutional networks for semantic segmentation. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2017,39(4):640–651. [doi: 10.1109/TPAMI.2016.2572683]

© 中国科学院软件研究所 http://www.jos.org.cn

[19] Gong C, Tao D, Liu W, Liu L, Yang J. Label propagation via teaching-to-learn and learning-to-teach. IEEE Trans. on Neural Networks and Learning Systems, 2017,28(6):1452–1465. [doi: 10.1109/TNNLS.2016.2514360]

[20] Zhang K, Zuo W, Chen Y, Meng D, Zhang L. Beyond a Gaussian denoiser: Residual learning of deep CNN for image denoising. IEEE Trans. on Image Processing, 2017,26(7):3142–3155. [doi: 10.1109/TIP.2017.2662206]

[21] Liu T, Tao D, Song M, Maybank S. Algorithm-Dependent generalization bounds for multi-task learning. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2017,39(2):227–241. [doi: 10.1109/TPAMI.2016.2544314]

[22] Xu C., Liu T, Tao D, Xu C. Local rademacher complexity for multi-label learning. IEEE Trans. on Image Processing, 2016,25(3): 1495–1507. [doi: 10.1109/TIP.2016.2524207]

[23] Gong C, Liu T, Tao D, Fu K, Tu E, Yang J. Deformed graph Laplacian for semi-supervised learning, IEEE Trans. on Neural Networks and Learning Systems, 2015,26(10):2261–2274. [doi: 10.1109/TNNLS.2014.2376936]

[24] Yang J, Wright J, Huang TS, Ma Y. Image super-resolution via sparse representation. IEEE Trans. on Image Processing, 2010, 19(11):2861–2873. [doi: 10.1109/TIP.2010.2050625]

[25] Martin D, Fowlkes C, Tal D, Jitendra Malik J. Learning to detect natural image boundaries using local brightness, color, and texture cues. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2004,26(5):530–549. [doi: 10.1109/TPAMI.2004. 1273918]

[26] Bevilacqua M, Roumy A, Guillemot C, Morel M. Single-Image super-resolution via linear mapping of interpolated self-examples. IEEE Trans. on Image Processing, 2014,23(12):5334–5347. [doi: 10.1109/TIP.2014.2364116]

[27] Zhang K, Gao X, Tao D, Li X. Single image super-resolution with non-local means and steering kernel regression. IEEE Trans. on Image Processing, 2012,21(11):4544–4556. [doi: 10.1109/TIP.2012.2208977]

[28] Jaderberg M, Simonyan K, Vedaldi A, Zisserman A. Reading text in the wild with convolutional neural networks. Int'l Journal of Computer Vision, 2016,116(1):1–20. [doi: 10.1007/s11263-015-0823-z]

[29] Ye Y, Shan J, Bruzzone L, Shen L. Robust registration of multimodal remote sensing images based on structural similarity. IEEE Trans. on Geoscience and Remote Sensing, 2017,55(5):2941–2958. [doi: 10.1109/TGRS.2017.2656380]

[30] Kuo TY, Su PC, Tsai CM. Improved visual information fidelity based on sensitivity characteristics of digital images. Journal of Visual Communication and Image Representation, 2016,40:76–84. [doi: 10.1016/j.jvcir.2016.06.010]

[31] Timofte R, De Smet V, Van Gool L. A+: Adjusted anchored neighborhood regression for fast super-resolution. In: Proc. of the Asian Conf. on Computer Vision. Springer-Verlag, 2015. 111–126. [doi: 10.1007/978-3-319-16817-3_8]

[32] Schulter S, Leistner C, Bischof H. Fast and accurate image upscaling with super-resolution forests. In: Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition. IEEE, 2015. 3791–3799. [doi: 10.1109/CVPR.2015.7299003]

彭亚丽(1979－),女,陕西石泉人,博士,副教授,主要研究领域为模式识别.

张鲁(1994－),男,硕士,主要研究领域为深度学习,图像复原(图像超分辨率,图像去噪).

张钰(1982－),男,博士,讲师,主要研究领域为计算机视觉,人工智能.

刘侍刚(1973－),男,博士,教授,CCF 专业会员,主要研究领域为计算机视觉.

郭敏(1964－),女,博士,教授,博士生导师,主要研究领域为计算机视觉,图像处理.

© 中国科学院软件研究所 http://www.jos.org.cn