

基于残差密集网络的单幅图像超分辨率重建

谢雪晴

(重庆工业职业技术学院信息工程学院 重庆 401120)

摘要 随着数码相机、手机等电子设备的普及,每天都会产生大量的图像,但通常这些图像的分辨率比较低。针对单幅图像超分辨率(Single Image Super-Resolution, SISR)方法性能较低的问题,提出一种基于残差密集网络的单幅图像超分辨率重建方法。将浅层的卷积特征输入到残差密集块,获得全局和局部的特征;对图像进行超分辨率重建,得到清晰的高分辨率图像。为了验证该方法的有效性,在四个公共的数据集 Set5、Set14、B100 和 Urban10 上进行了定性和定量的实验。实验结果表明,该方法能够更好地恢复出高分辨率的图像。

关键词 图像超分辨率 低分辨率 高分辨率 深度学习

中图分类号 TP391 文献标识码 A DOI:10.3969/j.issn.1000-386x.2019.10.038

SUPER-RESOLUTION RECONSTRUCTION OF SINGLE IMAGE BASED ON RESIDUAL-DENSE NETWORK

Xie Xueqing

(College of Information Engineering, Chongqing Industry Polytechnic College, Chongqing 401120, China)

Abstract With the popularity of electronic devices such as digital cameras and mobile phones, a large number of images are produced every day, but usually the resolution of these images is low. Aiming at the problem of low performance of single image super-resolution (SISR) method, this paper proposed a super-resolution reconstruction method for single image based on residual-dense network. I input the convolution feature of shallow layer into residual-dense blocks to obtain global and local features, and reconstructed the image with super resolution to get clear and high resolution image. To verify the effectiveness of this method, this paper conducted qualitative and quantitative experiments on four public data sets Set5, Set14, B100 and Urban10. Experimental results show that the proposed method can better restore high resolution images.

Keywords Image super-resolution Low resolution High resolution Deep learning

0 引言

图像超分辨率重构是数字图像处理领域中的研究热点之一。图像超分辨率是指把分辨率比较低的图像通过某种算法得到分辨率较高的图像,使获得的高分辨率图像具有更加清晰的颜色、纹理等信息。提高低质量图像质量的方法一般有两种:一是改善硬件设备,二是图像超分辨率重建技术。然而在实际应用中,由于比较高的工程成本和制作流程,大多数场合一般不

会采用分辨率较高的相机来采集数字图像。而是利用超分辨率重建技术来得到具有更加丰富的颜色纹理等信息的高分辨率图像。目前,常见的超分辨率重建技术分为三类:基于插值的方法、基于重建的方法、基于学习的方法。

基于插值的方法包括最近邻插值和三次插值,优点是算法复杂度较低,缺点是很容易出现阶梯锯齿状现象以及产生边缘模糊。

基于重建的方法使用了强制约束平滑和下采样技术,使得低分辨率图像和高分辨率图像保持一致性,由

于该方法过分地依赖高分辨率图像的先验知识,当缩小或放大图像的时候,重建质量就会下降,重建的结果就会缺失重要的细节信息。

近年来,深度学习在图像超分辨率重建领域表现出巨大的潜力。文献[1]首次提出基于卷积神经网络的图像超像素重建方法,第一次用深度学习来解决 SISR 问题。该网络共三层卷积层:第一层提取输入的低分辨率图像的特征;第二层将提取的低分辨率图像的特征映射到高分辨率图像的特征;第三层为高分辨率图像的重建。损失函数为简单的 MSE 损失。文献[2]提出了基于学习的图像超分辨率算法,利用样本先验知识来重建图像,相较于其他重建方法有明显的优势,也是近年来研究的热点。本文首先分析了影响图像重建质量的原因,然后对基于深度学习的超分辨率重建算法^[1]提出了两点改进:用随机线性纠正单元(Randomized rectified linear unit, RReLU)来避免原有网络学习中对图像某些重要的信息的过度压缩,同时用 NAG(Nesterov's accelerated gradient)方法来加速网络的收敛,进而避免了在梯度更新时产生较大的震荡。目前,图像超分辨率方法难以同时满足运算速度快和生成的图像高质量这一问题,文献[3]提出一种基于卷积的快速的图像超分辨率方法。此方法首先提取卷积特征,然后利用去卷积层对提取的特征放大膨胀,再以池化层对特征进行浓缩,提炼更鲁棒的特征,进而获得高分辨率图像。文献[4]提出了一种基于 GAN 网络的超分辨率重建。文中,作者论证了 PSNR 不能作为评价超分辨的标准,因为它和人对图片的感受结果不一致,并采用了 MOS 进行评价;另外作者在 loss 函数上进行了改进,更好地恢复出细节的纹理信息。文献[5]提出了用于图像超分辨率任务的新方法,分别是增强深度超分辨率网络 EDSR 和一种新的多尺度深度超分辨率 MDSR,在减小模型大小的同时实现了比当前其他方法更好的性能。文献[6]提出了一个非常深的循环残差网络用于图像超分辨率,这个网络高达 52 层。文献[7]提出了一种多阶段级联残差卷积神经网络模型。该模型分为两阶段来进行超分辨率图像重建,首先重建 2 倍超分辨率图像,然后重建 4 倍超分辨率图像。第一阶段和第二阶段都使用残差层,对两阶段分别构建多任务损失函数,利用第一阶段的损失来指导第二阶段的损失,进而提高网络训练的速度。文献[8]针对经典的基于卷积神经网络的超分辨率(SRCNN)方法存在重建图像纹理结构模糊以及网络

模型训练收敛过慢等问题,提出了一种多通道卷积的图像超分辨率(MCSR)方法。该方法通过增加残差链接,加快了模型收敛速度。同时利用多通道映射来提取更加丰富的特征,进而增强超分辨率重构效果。文献[9]提出了一种维度拉长策略,将模糊和噪声作为输入,可以应对多倍和空间改变的退化模型,提高了实用性。实验结果表明,提出的网络结构在多种多样的退化采样后,依然能生成比较好的结果。近年来,随着深度学习的快速发展,出现了很多基于神经网络的超分辨率算法。深度卷积网络在图像超分辨率重建方面取得了卓越的成就,但是参数数量往往非常巨大。文献[10]提出了一种简洁紧凑型递归残差网络结构。该网络通过局部残差的学习来减轻训练深层网络的困难;引入递归结构保证增加深度的同时控制模型参数数量;采用可调梯度裁剪方法来防止梯度消失。在网络末端使用反卷积层直接上采样图像,进而得到超分辨率图像。

然而,这些方法只是用了某些卷积层的信息,而没有用全部的卷积层的信息。图像中的物体有不同的尺度、视角和长宽比,深度卷积神经网络中的层次特征对于图像重构有很大的帮助。因此,本文提出了一个残差密集网络结构用于图像超分辨率。

1 残差网络和密集网络

在训练的过程中,深度神经网络容易遇到梯度消失问题,在块的归一化中,我们将输入数据由激活函数的收敛区调整到梯度较大的区域,在一定程度上缓解了这种问题。不过,当网络的层数急剧增加时,BP 算法导数的累乘效应很容易让梯度慢慢减小直至消失。为了从根本上解决这个问题,文献[11]提出了一种深度残差网络,简化那些非常深的网络的训练过程,使得层能根据其输入来学习残差函数而非原始函数。实验表明,这些残差网络的优化过程比较简单,能够使得网络结构有更深的层,而且能获得更加高的性能。文献[11]在 ImageNet 数据集上使用了一个 152 层的网络结构来评估所提出的残差网络,虽然它相当于 8 倍深的 VGG 网络,但是在所提出的框架中仍然有非常低的复杂度。这些残差网络的一个组合模型,在 ImageNet 测试集上的错误率仅为 3.57%。此结果在 2015 年的 ILSVRC 分类任务上获得了第一名的好成绩。每一个残差块如图 1 所示。

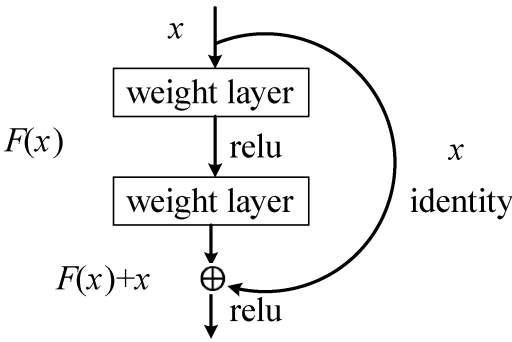


图1 残差块

其中: x 是输入。假设 $H(x)$ 是映射函数,我们不期望得到近似的 $H(x)$,而是把它们近似为一个残差函数 $F(x):F(x):=H(x)-x$ 。因此,原始的函数 $H(x):=F(x)+x$ 。

对每个堆栈层都应用残差学习,残差块可以被表示为如下形式:

$$y = F(x, \{W_i\}) + x \tag{1}$$

式中: x 和 y 分别为输入和输出;函数 $F(x, \{W_i\})$ 表示学习到的残差映射。对于图 1 所示的残差块,总共为两层, $F = W_2\sigma(W_1x)$,其中, σ 为 ReLU。

根据最近的成果来看,如果神经网络的每一层到输入层和输出层采用更短的连接,那么网络结构就可以设计得更深、更准确,并且训练起来也更加高效。文献[12]根据这个现象,提出了密集的卷积神经网络(Dense Convolutional Network, DenseNet),该网络以前馈的方式把每一层与其他层分别相连接。传统的 L 层卷积网络总共有 L 个连接,而 DenseNet 的任一层不仅与相邻层有连接,而且与它随后的所有层都有直接连接,所以该网络有 $\frac{L(L+1)}{2}$ 个直接连接。DenseNet 有

如下几个令人信服的优点:(1) 有效解决了梯度消失问题;(2) 增强了特征传播能力;(3) DenseNet 把网络中的每一层都直接与其前面的层相连接,实现了特征的重复利用;(4) DenseNet 把网络中的每一层设计得非常窄,即只学习很少的特征图,最极端情况就是每一层只学习一个特征图,这样可以降低冗余性。

2 方法设计

2.1 整体流程

本文提出了一种残差密集网络结构用于图像超分辨率,整体流程图如图 2 所示。该网络结构包括三部分:浅层的特征提取,残差密集块以及上采样操作。

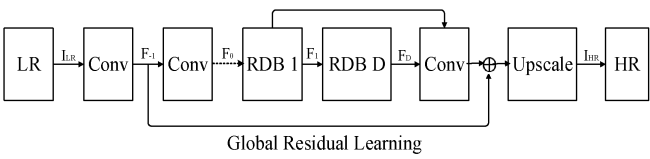


图2 残差密集网络结构

我们用两层卷积层来提取浅层的特征,第一层卷积层提取输入的低分辨率图像(Low Resolution, LR)的特征,表示为:

$$F_{-1} = H_{sf1}(I_{LR}) \tag{2}$$

式中: $H_{sf1}(\cdot)$ 表示卷积操作。 F_{-1} 可以进一步用于浅层特征的提取,因此,可以依次类推得到 F_0 :

$$F_0 = H_{sf2}(F_{-1}) \tag{3}$$

式中: $H_{sf2}(\cdot)$ 表示第二个特征提取层的卷积操作。特征 F_0 作为残差密集块的输入。假设我们有 D 个残差密集块。第 d 层残差密集块的输入 F_d 可以表示为:

$$F_d = H_{RDB,d}(F_{d-1}) = H_{RDB,d}(\cdots H_{RDB,0}(F_0)) \tag{4}$$

式中: $H_{RDB,d}(\cdot)$ 表示第 d 个残差密集块(Residual Dense Block, RDB)的操作,它可以表示为复合函数运算,例如卷积层和 Rectified Linear Units(ReLU)层。最后,所有的特征可以表示:

$$F_{all} = H(F_{-1}, F_0, F_1, \cdots, F_D) \tag{5}$$

在提取了局部和全局的特征 F_{all} 之后,我们用一个上采样的网络,来得到高分辨率图像(High Resolution, HR)。

2.2 残差密集块

残差密集网络结构包括密集的连接层、局部特征融合层,如图 3 所示。

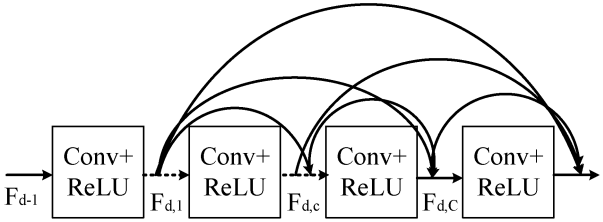


图3 残差密集块

令 F_{d-1} 和 F_d 分别表示第 d 个残差密集块的输入和输出。那么,第 d 个残差密集块中的第 c 个卷积层的输出可以表示为:

$$F_{d,c} = \sigma(W_{d,c}[F_{d-1}, F_{d,1}, \cdots, F_{d,c-1}]) \tag{6}$$

式中: σ 为 ReLU 激活函数, $W_{d,c}$ 是第 c 个卷积层的权重。

本文所提出的残差密集网络生成高质量的超分辨率图像的方法,充分利用了原始低分辨率图像的每一层的特征。残差密集块不仅可以通过连续记忆机制从

前一个残差密集块读取状态,还可以通过局部密集连接充分利用其中的所有层的信息。然后通过局部特征融合自适应地保留累积的特征。此外,还利用全局残差学习,将浅层特征和深层特征结合在一起,从原始 LR 图像中得到全局密集特征。

3 实 验

3.1 数据集和评价指标

为了评估本文提出的方法的有效性,本文训练模型用文献[1]提到的 91 幅图像。对于测试,本文用四个基准的数据集:Set5、Set14、B100 和 Urban100。Set5 共包括 5 幅图像,Set14 包括 14 幅图像。

图像超分辨率重构常用的客观评价指标主要包括峰值信噪比(Peak Signal to Noise Ratio, PSNR)和结构相似性(Structural Similarity Index Method, SSIM)。

峰值信噪比是一种简单且广泛使用的 SISR 测量方法,它通过均方差(MSE)进行定义。PSNR 可以定义为:

$$PSNR = 10 \cdot \lg \left(\frac{MAX^2}{MSE} \right) \tag{7}$$

式中:MAX 表示图像颜色的最大数值,8 bit 图像的最大取值为 255。MSE 为均方差,定义为:

$$MSE = \frac{1}{mn} \sum_{i=1}^n \sum_{j=1}^m \|K(i,j) - I(i,j)\|^2 \tag{8}$$

式中:I 和 K 分别是原始图像和处理后的图像,图像的大小为 $m \times n$ 。

SSIM 评价指标采用更加直接的方法来比较重建图像和参考图像的结构,表示为:

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

式中: μ_x 和 μ_y 表示图像所有像素的平均值, σ_x 和 σ_y 表示图像像素值的方差, C_1 和 C_2 为常数。

PSNR 值越大,说明效果越好;SSIM 值越小,说明效果越好。

3.2 实验结果与分析

对于每一个超分辨率的尺度($\times 2$ 、 $\times 3$ 和 $\times 4$),本文分别训练独立的模型。将本文方法与 SRCNN^[1]方法、SRGAN^[4]方法、SRRRN^[6]方法、SRMulti^[9]方法分别在数据集 Set5、Set14、B100 和 Urban100 上进行对比,实验结果如表 1 – 表 4 所示。可以看出,本文方法更加有效。

表 1 本文方法和其他先进的方法在 Set5 数据集上的比较

尺度	指标	SRCNN	SRGAN	SRRRN	SRMulti	本文方法
$\times 2$	PSNR	36.65	36.23	36.98	36.91	37.24
	SSIM	0.954	0.941	0.962	0.959	0.971
$\times 3$	PSNR	33.15	33.04	33.31	33.41	33.56
	SSIM	0.913	0.910	0.923	0.924	0.943
$\times 4$	PSNR	30.49	30.26	30.64	30.62	30.95
	SSIM	0.862	0.859	0.868	0.870	0.882

表 2 本文方法和其他先进的方法在 Set14 数据集上的比较

尺度	指标	SRCNN	SRGAN	SRRRN	SRMulti	本文方法
$\times 2$	PSNR	32.29	32.25	32.35	32.43	32.68
	SSIM	0.903	0.896	0.905	0.901	0.910
$\times 3$	PSNR	29.41	29.35	29.57	29.51	29.84
	SSIM	0.823	0.821	0.826	0.828	0.831
$\times 4$	PSNR	27.61	27.53	27.80	27.89	27.96
	SSIM	0.754	0.750	0.759	0.756	0.769

表 3 本文方法和其他先进的方法在 B100 数据集上的比较

尺度	指标	SRCNN	SRGAN	SRRRN	SRMulti	本文方法
$\times 2$	PSNR	31.22	31.21	31.29	31.33	31.45
	SSIM	0.887	0.885	0.889	0.892	0.899
$\times 3$	PSNR	28.41	28.32	28.47	28.52	28.67
	SSIM	0.787	0.786	0.789	0.792	0.806
$\times 4$	PSNR	26.91	26.89	26.93	26.96	27.12
	SSIM	0.712	0.711	0.715	0.719	0.723

表 4 本文方法和其他先进的方法在 Urban100 数据集上的比较

尺度	指标	SRCNN	SRGAN	SRRRN	SRMulti	本文方法
$\times 2$	PSNR	29.52	29.48	29.58	29.67	29.78
	SSIM	0.895	0.891	0.897	0.906	0.916
$\times 3$	PSNR	26.24	26.12	26.28	26.38	26.45
	SSIM	0.800	0.796	0.802	0.812	0.819
$\times 4$	PSNR	24.53	24.43	24.56	24.62	24.82
	SSIM	0.724	0.721	0.729	0.730	0.733

本文除了对提出的方法进行定量的比较,还进行了定性的比较,如图 4 所示,为了得到更好的视觉效果,对标出的长方形框内的区域进行放大。图 4 中第一列是原始的图像,第二至第五列分别为比较先进的超分辨率方法,最后一列是本文提出的方法。可以看出,本文提出的方法可以更好地恢复出细节,进而得到更加清晰的高分辨率图像。

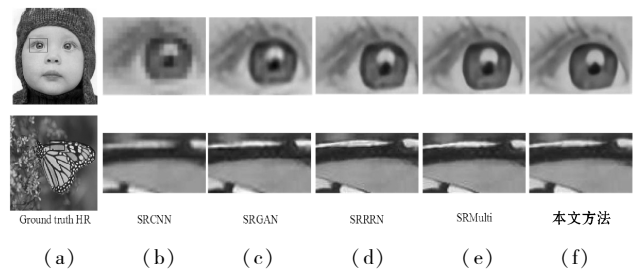


图4 超分辨率的定性比较

4 结 语

本文提出了一种残差密集深度网络用于图像超分辨率,该方法充分利用了每一层的特征,得到了纹理等细节更清晰的超分辨率图像。在四个公共的数据集上进行了定量的实验,用常用的评价指标进行评价,实验结果表明,本文的方法更有效。此外,本文进行了定性的比较,通过直观的视觉观察可以看出,本文的方法相比其他相对先进的方法可以得到更加清晰的高分辨率图像。

参 考 文 献

[1] 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,2014,38(2):295-307.

[2] 胡长胜,詹曙,吴从中. 基于深度特征学习的图像超分辨率重建[J]. 自动化学报, 2017, 43(5):814-821.

[3] 孙超,吕俊伟,李健伟,等. 基于去卷积的快速图像超分辨率方法[J]. 光学学报, 2017(12):142-152.

[4] Ledig C, Theis L, Huszar 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). IEEE Computer Society, 2017.

[5] 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). IEEE Computer Society, 2017.

[6] Tai Y, Yang J, Liu X. Image super-resolution via deep recursive residual network [C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE Computer Society, 2017.

[7] 王一宁,秦品乐,李传朋,等. 基于残差神经网络的图像超分辨率改进算法[J]. 计算机应用,2018(1):246-254.

[8] 李云飞,符冉迪,金炜,等. 多通道卷积的图像超分辨率方法[J]. 中国图象图形学报,2018,22(12):1690-1700.

[9] Zhang K, Zuo W, Zhang L. Learning a single convolutional super-resolution network for multiple degradations [C]//2018 IEEE Conference on Computer Vision and Pattern Rec-

ognition(CVPR). IEEE Computer Society, 2018.

[10] 周登文,赵丽娟,段然,等. 基于递归残差网络的图像超分辨率重建[J/OL]. 自动化学报:1-9[2019-01-23]. <https://doi.org/10.16383/j.aas.c180334>.

[11] He K, Zhang X, Ren S, et al. Deep residual learning for image recognition [C]//IEEE Conference on Computer Vision and Pattern Recognition. 2016.

[12] Gao H, Zhuang L, Maaten L V D, et al. Densely connected convolutional networks[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition(CVPR). 2017.

(上接第 189 页)

[13] Kim Y, Jernite Y, Sontag D, et al. Character-Aware Neural Language Models [C]//Association for the Advancement of Artificial Intelligence, 2016.

[14] Bengio Y, Ducharme R, Vincent P, et al. A Neural Probabilistic Language Model[C]. Journal of machine learning research, 2003, 3: 1137-1155.

[15] Yepes A J, MacKinlay A. NER for Medical Entities in Twitter using Sequence to Sequence Neural Networks[C]//Proceedings of Australasian Language Technology Association Workshop, 2016: 138-142.

[16] Tomori S, Ninomiya T, Mori S. Domain Specific Named Entity Recognition Referring to the Real World by Deep Neural Networks[C]//Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics, 2016: 236-242.

[17] Cotterell R, Duh K. Low-Resource Named Entity Recognition with Cross-Lingual, Character-Level Neural Conditional Random Fields[C]//Proceedings of the The 8th International Joint Conference on Natural Language Processing, 2017: 91-96.

[18] Li P H, Dong R P, Wang Y S, et al. Leveraging Linguistic Structures for Named Entity Recognition with Recursive Neural Networks[C]//Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing, 2017: 2664-2669.

[19] He H F, Sun X. F-Score Driven Max Margin Neural Network for Named Entity Recognition in Chinese Social Media[C]// Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics; Volume 2, 2017: 713-718.

[20] Gridach M. Character-Aware Neural Networks for Arabic Named Entity Recognition for Social Media [C]//Proceedings of the 6th Workshop on South and Southeast Asian Natural Language Processing, 2016: 23-32.

[21] Démoncourt F, Lee J Y, Szolovits P. NeuroNER: An Easy-to-use Program for Named-entity Recognition based on Neural Networks[C]//Proceedings of the 2017 EMNLP System Demonstrations, 2017: 97-102.