

# 改进的基于卷积神经网络的图像超分辨率方法

王 旺 徐俊武 李颖先

(武汉工程大学计算机科学与工程学院 湖北 武汉 430205)

**摘 要** 图像超分辨率是计算机视觉领域的经典问题。使用深度神经网络来解决图像超分辨率的问题目前得到越来越多的研究学者的关注和青睐。为改善基于卷积神经网络的图像超分辨率方法的图像生成效果,提出一种改进的方法。在神经网络层中加深网络层数,并且针对加深网络可能出现的退化现象引入残差网络结构,并将图像上采样步骤放入网络中。实验表明,在与传统的插值法和原始的基于卷积神经网络方法的对比中,该优化方法生成的图像观感更加锐利清晰、细节丰富,而且无论在峰值信噪比和结构相似性上均有明显提高,验证了该方法的有效性。

**关键词** 超分辨率 深度学习 卷积神经网络

中图分类号 TP391

文献标识码 A

DOI: 10.3969/j.issn.1000-386x.2019.03.039

## AN IMPROVED IMAGE SUPER-RESOLUTION METHOD BASED ON CONVOLUTION NEURAL NETWORK

Wang Wang Xu Junwu Li Yingxian

(School of Computer Science and Engineering, Wuhan Institute of Technology, Wuhan 430205, Hubei, China)

**Abstract** Image super-resolution is a classical problem in the field of computer vision. The use of depth neural network to solve the problem of image super-resolution has attracted more and more attention and favor of researchers. In order to improve the image generation effect of image super-resolution method based on convolution neural network, an improved method was proposed. It mainly increased the number of network layers in the neural network layer, introduced residual network structure for the degradation phenomenon that might occur in the deepening network, and put the image sampling steps into the network. Experiments show that compared with the traditional interpolation method and the original method based on convolutional neural network, the optimization method generates sharper and clear image with richer details. And it has significant improvement both in peak signal-to-noise ratio and structural similarity, which verifies the effectiveness of this method.

**Keywords** Super-resolution Deep learning Convolution neural network

## 0 引 言

图像的超分辨率 SR(Super Resolution)技术,是通过软件算法来把低分辨率 LR(Low Resolution)的图像转换成为高分辨率 HR(High Resolution)图像的技术<sup>[1]</sup>。目前在许多重要的领域都十分需要这门技术的支持,例如在医学图像处理领域,高分辨率图像可以帮助医生掌握病情,看到人眼不易察觉到的病灶,可以更

好地针对治疗<sup>[2]</sup>。

SR技术最早可以追溯到20世纪60年代,提出者为Harris<sup>[3]</sup>。当时的主流方法是插值法,之后随着技术的发展,有新的方法被研究者提出,主要归纳为基于重建的方法和基于学习的方法<sup>[4]</sup>。插值法具有最简单的计算过程和最低的复杂度,常见的有邻插值法<sup>[5]</sup>、双线性插值法<sup>[6]</sup>和双三次插值法<sup>[7]</sup>;基于重建的方法在一段时间内颇为热门,经典的有迭代反投影法<sup>[8]</sup>、凸集投影法<sup>[9]</sup>和最大后验概率估计法<sup>[10]</sup>,这种方法最大

收稿日期:2018-10-11。王旺,硕士生,主研领域:图像处理。徐俊武,副教授。李颖先,硕士生。

的问题在于放大倍数有限,生成效果也不及后来者;基于学习的方法分为基于浅层网络的学习方法,主要有流形学习<sup>[11]</sup>和稀疏表示<sup>[12]</sup>;另一种基于学习的方法就是基于深度网络的方法。

将深度学习应用于图像超分辨率领域,最早由Dong等<sup>[13]</sup>提出,也就是经典的基于卷积神经网络的超分辨率方法(SRCNN),之后不断地有学者进行优化和改进。

## 1 相关工作

### 1.1 基于卷积神经网络的图像超分辨率方法

基于卷积神经网络的超分辨率(SRCNN)方法是最先将深度学习引入图像超分辨率重建的问题。超分辨率方法将一张LR图像放大为一张HR图像,尺寸的变化造成像素点总量的变化,其中最难的挑战是如何在新增的像素位置进行值的填充。SRCNN使LR图像能通过一定的算法升为HR图像,主要是两者之间存在“共同特征”,所以在SRCNN中,将超分辨率过程分为三个阶段<sup>[14]</sup>:

(1) 特征提取。此阶段就是对LR图像进行特征提取和特征表示,利用卷积网络的性质提取图像块的特征,公式如下:

$$F_1(Y) = \max(0, W_1 \times Y + B_1) \quad (1)$$

(2) 非线性映射。将第一阶段提取的 $n_1$ 维特征映射至 $n_2$ 维,公式如下:

$$F_2(Y) = \max(0, W_2 \times F_1(Y) + B_2) \quad (2)$$

(3) 重建。这个阶段是将第二阶段映射后的特征恢复为HR图像,公式如下:

$$F(Y) = W_3 + F_2(Y) + B_3 \quad (3)$$

式中: $W$ 和 $B$ 分别代表卷积模板和偏置参数。

SRCNN的网络模型结构如图1所示。文献[13]只使用了3层简单的卷积神经网络,分别进行上述三个阶段。

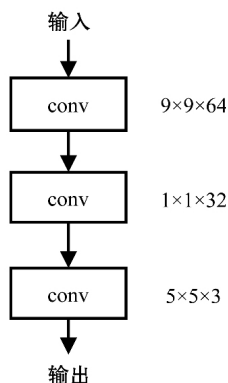


图1 SRCNN网络模型结构图

在SRCNN中,使用MSE作为损失函数,在图像输入前需要使用双三次插值放大至目标尺寸。此时虽然图像尺寸达到了,但是仍然称之为低分辨率图像,然后再进行输入。输出的是最终重建高分辨率图像。

### 1.2 优化相关技术

(1) 亚像素卷积层。亚像素卷积层<sup>[15]</sup>在输入原始低分辨率图像之后,经过卷积之后,得到通道数为 $r^2$ 的与输入图像大小一样的特征图像。然后将每个像素的 $r^2$ 通道再排列,成为 $r \times r$ 的区域,则每个像素都成为 $r \times r$ 大小,从而大小为 $H \times W \times r^2$ 的特征图像被重新排列成 $rH \times rW$ 的高分辨率图像。需要注意的是亚像素卷积层的输入维度为输出维度前 $r^2$ 倍,且 $r$ 为放大倍数。

(2) 残差网络。通过加深网络的深度可以很好地提高网络的性能,但是可能会出现退化问题,随着网络层数的增加,在训练集上的准确率却饱和甚至下降。深度残差网络<sup>[16]</sup>的出现主要是为了退化的问题。

出现退化的原因,主要是因为神经网络在反向传播过程中要不断地传播梯度,而当网络层数加深的时候,梯度在传播过程中会逐渐消失<sup>[17]</sup>。

残差网络解决退化现象的原理如图2所示,将某一层的输入 $x$ 直接传到后几层的输出中。这种方式被称为捷径连接,也就是最终的输出结果 $H(x)$ 为原本的输出 $F(x)$ 加上前几层的输入 $x$ ,即:

$$H(x) = F(x) + x \quad (4)$$

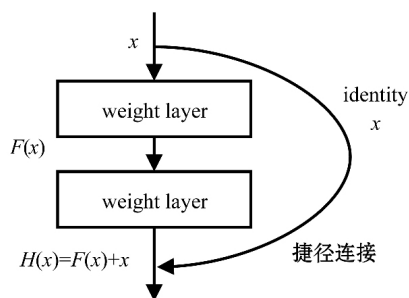


图2 残差网络示意图

这样做的好处在于,当 $F(x)$ 由于退化问题变为0的时候,那么至少输出 $H(x)$ 还是有值的。当 $H(x) = x$ 的时候,被称为恒等映射,由于还有输出,那么梯度还可以继续传递。

同时在残差块中也会加入批量归一化BN<sup>[18]</sup>层,添加BN的最大优点是,减少梯度消失,加快收敛速度。根据研究表明,BN层需要添加在网络层后激活函数前<sup>[19]</sup>效果最佳。

## 2 改进的SRCNN方法

本文提出基于SRCNN的图像超分辨率的优化方

法,主要在 SRCNN 的基础上做出如下修改:

1) 在输入方面,本方法在最后以两层亚像素卷积层<sup>[20]</sup>的形式放大图像尺寸,因此可以直接将原始低清图像作为输入,而无需像 SRCNN 那样先以双三次插值的方式放大尺寸再输入。

2) 本方法还改变特征维数,使用更小的卷积核和使用更多的映射层, SRCNN 总共只有 3 层网络,映射层只有一层,加深网络可以有效提升网络能力。

3) 本方法由于并非在网络外部进行放大图像的操作,所以更具有灵活性。当需要不同的上采样倍率时,只需要微调后面的上采样层就可以了,前面的映射层不变。

在 SRCNN 方法中,超分辨率的工作是在高分辨率的基础上完成的,这样并不是最优的办法,而且计算量很大。如果可以在网络内部进行超分辨率,通过训练学习将低分辨率特征映射到高分辨率输出,这样就能够减轻计算量,并且效果也更优。

在本方法中,网络模型可以分为 3 个部分:

1) 特征提取。SRCNN 中的特征提取是对插值后的高分辨率图像进行提取,卷积核大小为  $9 \times 9$ ; 本方法是直接对原始的低分辨率进行操作,因此可以选择小一些,设置为  $3 \times 3$ 。

2) 非线性映射。由于感受野大,能够表现得更好,在 SRCNN 中采用的是  $5 \times 5$  的卷积核,但是卷积核大的计算量会比较大,用两个串联的  $3 \times 3$  的卷积核可以替代一个  $5 \times 5$  的。同时两个串联的小卷积核需要的参数数量为 18,而  $5 \times 5$  的参数数量为 25,所以本方法使用多个  $3 \times 3$  的小卷积层,兼顾计算量和网络能力。

3) 上采样。使用亚像素卷积层进行上采样。

具体流程描述如下:

1) 输入图像张量,维度为 3;

2) 进入卷积层,输入维度为 3,输出维度为 64,卷积核为  $3 \times 3$ ,步长为 1,之后接入激活函数 ReLU;

3) 进入残差块,内部含有 4 个小残差块和一个卷积层,每个小残差块内部含有两个卷积层,每层卷积之后都加入 BN 层,激活函数为 ReLU;

4) 进入卷积层,输入维度为 64,输出维度为 256,卷积核为  $1 \times 1$ ,步长为 1,之后接入 BN 层,再激活函数 ReLU;

5) 上采样卷积之前需要计算维度,先使用卷积层将维度升到一定高度再进行上采样,该卷积层输入维度为 64,输出维度 256,卷积核为  $3 \times 3$ ,步长为 1,之后再进行上采样层,输出维度 64,ReLU 激活;

6) 再重复步骤 5,两次上采样将图像尺寸提升 4 倍;

7) 最后以  $1 \times 1$  卷积核降维到 3,激活函数改为 Tanh,输出大尺寸图像。

网络流程见图 3。

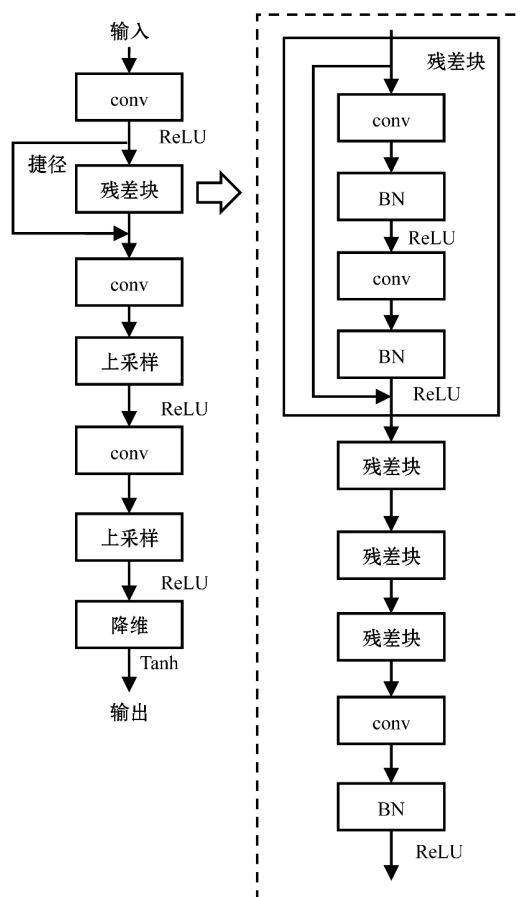


图3 本文方法网络流程示意图

在本文方法中,损失函数依然使用 MSE,优化器选用 Adam。

### 3 实验

#### 3.1 实验设计及评价标准

实验采用 DIV2K 数据集的图像作为训练集,训练完成后使用测试图像进行重建生成高分辨率图像用于查看效果。

由于实验中的低分辨率图像是从高分辨率图像进行  $1/4$  倍降采样得到,所以本实验有原始高分辨率图像作为对照。同时一起列出用双三次插值法重建图像;还另外用相同的训练集和训练次数训练 SRCNN 方法、ESPCN<sup>[20]</sup> 方法和 FSRCNN<sup>[21]</sup> 方法并将测试图像生成结果一起列出。图 4 和图 5 分别展示这三种方法重建生成图像的结果和原始图像,都是截取了关键部分,以便查看细节部分。

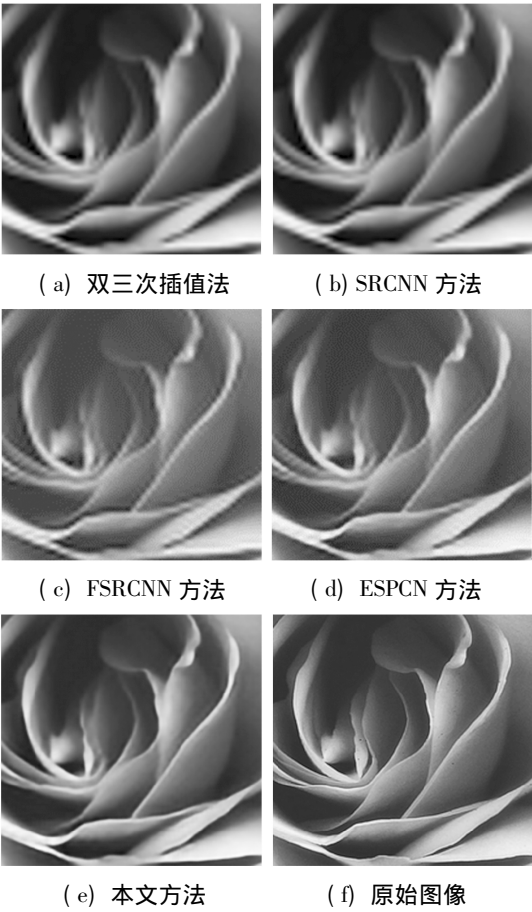


图 4 测试图像 1

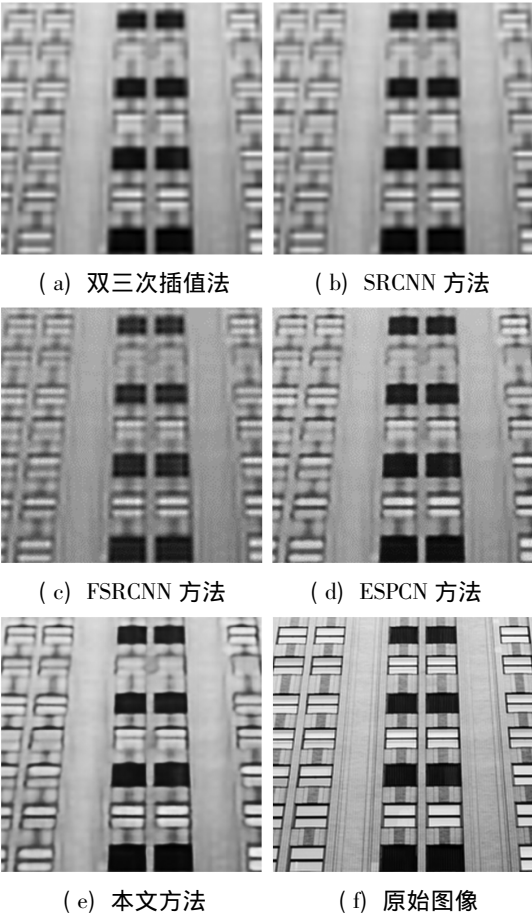


图 5 测试图像 2

本实验评价采用主观评价和客观评价共同分析生成图像质量。其中主观评价就是在图像的视觉效果上感受图像的效果,对比各种方法生成的图像之间的差异;客观评价以峰值信噪比 PSNR 和结构相似性 SSIM 来进行衡量。

(1) 峰值信噪比。峰值信噪比是重建的高分辨率图像真实高分辨率图像之间的像素差值的量度,可以用来衡量被处理后图像的质量,单位为 dB。定义如下:

$$PSNR = 10 \lg \left( \frac{255^2}{MSE} \right) \tag{5}$$

$$MSE = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} \|I(i,j) - K(i,j)\|^2 \tag{6}$$

式中:  $I$  为真实图像  $K$  为生成图像  $m$  和  $n$  分别为图像和长和宽。

(2) 结构相似性。结构相似性为两幅图像结构相似度的度量的标准,取值范围为  $[0, 1]$ ,值越大就越表示两幅图像结构相似程度越高。定义如下:

$$SSIM(x,y) = \frac{4\mu_x\mu_y\sigma_{xy}}{(\mu_x^2 + \mu_y^2)(\sigma_x^2 + \sigma_y^2)} \tag{7}$$

式中:  $\mu_x$  和  $\mu_y$  分别是  $x$  和  $y$  的平均值;  $\sigma_x^2$  和  $\sigma_y^2$  分别是  $x$  和  $y$  的方差;  $\sigma_{xy}$  是  $x$  和  $y$  的协方差。

表 1 和表 2 分别列出了 4 组测试图像的生成图像和原始图像进行计算的 PSNR 和 SSIM 的值。

表 1 图像重建结果的 PSNR

使用方法	图像 1	图像 2	图像 3	图像 4
插值法	22.55	17.57	19.81	21.37
SRCNN	22.58	17.53	19.73	21.41
FSRCNN	21.30	16.10	20.13	20.52
ESPCN	23.87	18.05	19.67	21.46
本文方法	26.52	19.46	20.67	21.64
原图像	$\infty$	$\infty$	$\infty$	$\infty$

表 2 图像重建结果的 SSIM

使用方法	图像 1	图像 2	图像 3	图像 4
插值法	0.949 3	0.815 3	0.936 3	0.958 7
SRCNN	0.949 5	0.812 7	0.936 0	0.964 2
FSRCNN	0.926 9	0.760 7	0.931 5	0.948 9
ESPCN	0.962 7	0.843 9	0.940 7	0.961 4
本文方法	0.980 8	0.893 0	0.946 8	0.958 9
原图像	1	1	1	1

3.2 实验结果分析

从图 4 和图 5 的实验结果可以看出,使用本文方

法生成的图像略优于其他方法。如图 4 的花瓣的边缘,图 4(a)和(b)不够锐利清晰,(c)和(d)也比较模糊,且花瓣层次不明显,有一些波纹,(e)花瓣边缘则较为清晰,花瓣层次也比较明显,整体观感会比较舒适。图 5 的建筑外墙,图 5(a)、(b)和(c)的外墙窗户十分模糊,窗户的边缘线条不清晰,(d)的外墙观感稍可,但是整体观感还是比较模糊,(e)的观感最为舒适纯净,外墙窗户边缘较为清晰,细节较为丰富,处理效果更胜一筹。其中 SRCNN、FSRCNN 和 ESPCN 三种方法和本文方法使用了相同的训练集和训练次数,在同样的训练条件下,本文的优化方法是有一定的优化效果的。这在表 1 和表 2 中的数据也能得到证明,在计算出的 PSNR 和 SSIM 值中,可以看出使用本文方法在客观评价方面,较其他的方法高出不少。

## 4 结 语

本文提出一种改进的基于卷积神经网络的图像超分辨率方法。主要在经典的 SRCNN 方法的基础上进行了改进,即加深了网络层并且引入残差网络避免退化问题,而且将上采样过程改为在网络内部进行。这样可以加强网络能力并且减轻计算量,最终得到效果更好的图像,而且训练次数要求更低。实验结果表明,本文方法从主观和客观角度,生成图像均好于双三次插值法和 SRCNN 方法,证明本文提出的改进是有效的。

## 参 考 文 献

- [1] Park S C, Min K P, Kang M G. Super-resolution image reconstruction: a technical overview [J]. IEEE Signal Processing Magazine, 2003, 20(3): 21-36.
- [2] 韩小虎. 基于深度学习的图像超分辨率算法研究[D]. 开封: 河南大学, 2016.
- [3] 黄冬冬. 基于深度学习的图像超分辨率重建算法研究[D]. 马鞍山: 安徽工业大学, 2017.
- [4] 首照宇, 廖敏璐, 陈利霞. 改进的基于稀疏表示的图像超分辨率重建算法[J]. 计算机应用与软件, 2014, 31(4): 201-204.
- [5] Schultz R R, Stevenson R L. A Bayesian approach to image expansion for improved definition. [J]. Image Processing IEEE Transactions on, 1994, 3(3): 233-242.
- [6] Hou H, Andrews H. Cubic splines for image interpolation and digital filtering [J]. IEEE Transactions on Acoustics Speech & Signal Processing, 1978, 26(6): 508-517.
- [7] Li X, Orchard M T. New edge-directed interpolation. [J]. IEEE Trans Image Process, 2001, 10(10): 1521-1527.
- [8] Irani M, Peleg S. Improving resolution by image registration [J]. Cvgip Graphical Models & Image Processing, 1991, 53(3): 231-239.
- [9] Stark H, Oskoui P. High-resolution image recovery from image-plane arrays, using convex projections [J]. Journal of the Optical Society of America A Optics & Image Science, 1989, 6(11): 1715.
- [10] Schultz R R, Stevenson R L. Improved definition video frame enhancement [C]//International Conference on Acoustics, Speech, and Signal Processing. IEEE, 1995, 4: 2169-2172.
- [11] Roweis S T, Saul L K. Nonlinear Dimensionality Reduction by Locally Linear Embedding [J]. Science, 2000, 290(5500): 2323-2326.
- [12] Yang J, Wright J, Huang T, et al. Image super-resolution as sparse representation of raw image patches [C]//2008 IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2008: 1-8.
- [13] Dong C, Loy C C, He K, et al. Image super-resolution using deep convolutional networks [J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2016, 38(2): 295-307.
- [14] 王学文. 基于学习的图像超分辨率算法研究[D]. 武汉: 华中科技大学, 2016.
- [15] 龙法宁, 朱晓姝, 胡春娇. 基于深层卷积网络的单幅图像超分辨率重建模型[J]. 广西科学, 2017, 24(3): 231-235.
- [16] He K, Zhang X, Ren S, et al. Deep residual learning for image recognition [C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016: 770-778.
- [17] 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, 38(1): 246-254.
- [18] Ioffe S, Szegedy C. Batch normalization: accelerating deep network training by reducing internal covariate shift [C]//International Conference on International Conference on Machine Learning. JMLR. org, 2015: 448-456.
- [19] 王晓斌, 黄金杰, 刘文举. 基于优化卷积神经网络结构的交通标志识别[J]. 计算机应用, 2017, 37(2): 530-534.
- [20] Shi W, Caballero J, Huszár F, et al. Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network [C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016: 1874-1883.
- [21] Dong C, Chen C L, Tang X. Accelerating the super-resolution convolutional neural network [C]//Proceedings of the 14th European Conference on Computer Vision, ECCV 2016, 2016: 391-407.