

DOI: 10.11830/ISSN.1000-5013.201804064



采用深度学习的快速超分辨率 图像重建方法

张圣祥^{1,2}, 郑力新^{1,2}, 朱建清^{1,2}, 潘书万^{1,2}

(1. 华侨大学 工学院, 福建 泉州 362021;

2. 华侨大学 工业智能化与系统福建省高校工程研究中心, 福建 泉州 362021)

摘要: 为满足实际工业生产需要,提出一种基于深度学习的快速超分辨率图像重建方法.采用一种快速的卷积神经网络结构,使用级联的小卷积核以取得重建速度上的提升,加深卷积网络以取得重建质量上的提升.实验结果表明:在标准的公共数据集上,该算法重建的高分辨率图像在主观视觉感受和客观的图像质量评价(峰值信噪比)上取得较好的效果,且重建时间大大缩短;将算法应用在实际的项目中,能达到阈值分割后准确检测物体的标准,减少企业对高额工业相机的经济开支.

关键词: 超分辨率图像重建;深度学习;卷积神经网络;级联

中图分类号: TP 391 文献标志码: A 文章编号: 1000-5013(2019)02-0245-06

Fast Super-Resolution Image Reconstruction Method Using Deep Learning

ZHANG Shengxiang^{1,2}, ZHENG Lixin^{1,2},
ZHU Jianqing^{1,2}, PAN Shuwan^{1,2}

(1. College of Engineering, Huaqiao University, Quanzhou 362021, China;

2. Industrial Intelligence and System Fujian University Engineering Research Center,
Huaqiao University, Quanzhou 362021, China)

Abstract: In order to meet the needs of actual industrial production, a fast super-resolution image reconstruction method based on deep learning is proposed. We proposed our own convolutional neural network structure, using cascaded small convolution kernels to achieve a higher reconstruction speed, and deepening the convolution network to achieve an improvement in reconstruction quality. The experimental results show that on the standard public dataset, the high-resolution image reconstructed by the our algorithm achieves better results in subjective visual perception and objective image quality evaluation (peak signal-to-noise ratio), at the mean time, the reconstruction time is greatly shortened. The algorithm is applied in projects to solve the problem for accurately detecting objects after threshold segmentation. In this way, it also reduces the high expenses of enterprises for purchasing industrial cameras.

Keywords: super-resolution image reconstruction; deep learning; convolutional neural network; cascade

视觉系统是工业机器人的“眼睛”,从视觉系统中得到图片质量的好坏能直接影响工业生产的效率.近年来,随着计算机硬件运算速度的快速提升,基于深度学习的超分辨率重建方法快速发展起来,它弥

收稿日期: 2018-04-19

通信作者: 郑力新(1967-),男,教授,博士,主要从事光电检测与智能计算的研究. E-mail: zlxgxy@hqu.edu.cn.

基金项目: 国家自然科学基金青年科学基金资助项目(61602191);福建省厦门市科技计划项目(3502Z20173045);福建省泉州市高层次人才创新创业项目(2017G036)

补了工业相机等硬件条件的限制和光学系统的打光偏差,但是,这也提高了对超分辨率重建技术在成像速度和成像质量上的要求.超分辨率重建技术主要分为基于重建的方法和基于学习的方法两大类^[1].基于重建的方法具有一定的约束条件,若大幅度提高分辨率,则其性能会急剧下降,也会出现丢失高频细节、边缘模糊等问题.基于学习的方法应用于单幅图像的超分辨率重建领域^[2],如 Yang 等^[3]提出的稀疏表示法在超分辨率重建领域一直占领着主导地位,但是该算法中,系数正则化效果不明显、字典完备性不强且复原边缘细节质量不高.针对稀疏表示法在物体特征表示不足和字典表示能力不够的问题,研究者们利用高斯金字塔^[4]、拉普拉斯金字塔^[5]、Sreerale 金字塔^[6]和 Contourlet 金字塔^[7]等方法建立图像不同分辨率之间的信息对应样本库.针对传统基于稀疏表示方法存在的问题,张晓燕等^[8]也提出了一些改进的算法,但仍存在一定的局限性,在处理一些细节比较多的图像时,容易造成纹理过重.近年来,深度学习在图像分类及图像分割等任务中的巨大成功也推动了超分辨率重建技术的发展,它弥补了传统浅层学习方法的不足^[9-10].基于此,本文基于深度学习,提出一种快速超分辨率图像重建的方法.

1 基于深度学习的超分辨率重建方法

基于卷积神经网络的单幅图像重建(SRCNN)算法是深度学习在图像超分辨率重建问题的开山之作^[11].与传统非深度学习的算法相比, SRCNN 算法的优势在于其三层卷积神经网络,在保持高质量图片输出的情况下,能够快速重建高分辨率的图像.虽然 SRCNN 算法使用较大的卷积核可以减少计算量,但每次卷积时还会丢失大量高频信息,从而导致重建质量不高^[12].

在 SRCNN 算法的基础上, Kim 等^[13]借鉴擅长图像分类的 VGG^[14]神经网络结构,提出了极深网络(VDSR)的超分辨率重建方法.采用更深层网络对低分辨率到高分辨率之间的映射关系进行建模,该方法克服了 SRCNN 算法上、下文信息缺乏关联、单一尺度放大等问题.但是该方法却不能实现多尺度的自由放大,且其参数存储和检索也存在较明显的不足,另外,它以牺牲时间为代价,在实时系统或应用到实际项目中有致命的弱点.

Shi 等^[15]提出了一种在低分辨率图像上直接计算卷积得到高分辨率图像的方法(ESPCN).在该方法中,直接将原始低分辨率图像输入网络中,通过两个卷积层后,得到的特征图像大小与输入图像一样的重建图像,其中,核心则是亚卷积层(sub-pixel convolution).该亚卷积层大大减少了图像重建的时间,在图像重建领域,这是 ESPCN 算法立足的最大优势,但相比于 VDSR 的重建方法,ESPCN 则在上、下文信息缺乏关联和物体特征的表达上还略显不足,边缘细节处理有所欠缺.

2 快速超分辨率图像重建的方法

2.1 总体框架

文中算法的总体结构框架主要由 Mini-network 层和亚卷积层^[15]两部分组成,如图 1 所示.各层的参数细节:卷积层 1(Conv1)的卷积核大小和激活函数分别为 $3 \times 3 \times 32$ 和 Tanh;级联层 2(Conv2)的卷积核大小和激活函数分别为 $(3 \times 3 + 3 \times 3) \times 64$ 和 ReLU;卷积层 3(Conv3)的卷积核大小和激活函数分别为 $3 \times 3 \times 32$ 和 Tanh;损失函数为欧式距离损失函数;通道数为 $r \times r$.

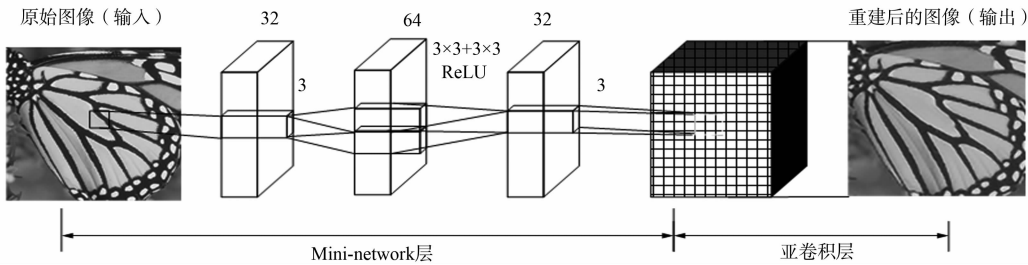


图 1 文中算法的总体结构框架

Fig. 1 Overall structural framework of the algorithm

2.2 重建速度的改进

2.2.1 Mini-network 层 为了缩短超分辨率的时间,在三层卷积网络中使用一个 Mini-network,即两

个级联的 3×3 卷积核. 神经网络中, 大的卷积核可以带来更大的感受野, 但其使用了更多的神经元参数, 同时, 也增加卷积运算的计算量. 由于神经元参数的数量仅与卷积核的大小有关, 因此, 越小的卷积核产生的神经元参数就越少, 用于卷积运算的计算量也就越少. 这里使用 ReLU 激活函数提取非线性特征, 从 ReLU 激活函数的表达式可以看出, 计算机只需要判断输入是否大于零即可输出结果, 所以, 其计算速度非常快.

时间复杂度 $T=O(N^2 \cdot K^2 \cdot F)$. 其中: T 表示时间复杂度; N 是输入图像的大小; K 是卷积核大小; F 是过滤器的数量. 卷积公式 $\text{Output}=(I-K+2P)/S+1$. 其中: I 为输入图像大小; K 是卷积核大小; P 表示填充, 即在图像之间添加额外的零层, 使输出图像与输入大小相同; S 表示原始图像中水平和垂直方向上滤波器的步长. 为便于讨论, 假定 $P=0, S=1$. 根据时间复杂度和卷积, 将文中算法中的 Mini-network 和 ESPCN 算法中的 5×5 卷积核进行性能比较, 结果如表 1 所示.

表 1 Mini-network 和的 5×5 卷积核的性能比较
Tab. 1 Performance comparison between Mini-network and 5×5 convolution kernel

项目	卷积核(5×5)	Mini-network($3 \times 3+3 \times 3$)
时间复杂度	$O(N^2 \times (5 \times 5 \times 64))=O(1\ 600 \times N^2)$	$O(N^2 \times (3 \times 3+3 \times 3) \times 64)=O(1\ 152 \times N^2)$
参数量	$5 \times 5+1=26$	$3 \times 3+1+3 \times 3+1=20$
计算量	$25(I-4)^2$	$9(I-2)2+9(I-4)^2$

由表 1 可知: Mini-network 在 3 种性能上都有优势; 在计算量比较中, 当 $I>10$ 时, Mini-network 在卷积乘法和加法两个方面的计算量都少于 5×5 模型.

2.2.2 亚卷积层 在卷积层中, 将低分辨率的图像直接作为整个网络的输入, 在不同滤波器 W_s 中, 有卷积核大小为 k_s 的参数在低分辨率空间中被激活后的数量正好是 r^2 . 在激活模式的对应位置上, 刚好激活 $(k_s/r)^2$ 个大小的权重. 然后, 将 $\text{mod}(x, r), \text{mod}(y, r)$ (x, y 是高分辨率空间中的输出像素坐标) 的结果重新排列, 排列后的位置表示为元素所在位置. 在最后一层, 使用一个亚像素卷积层重建一个图像, 即周期性的混洗操作 (periodic shuffling operator). 周期性的混洗操作将从亚卷积层中输出的高 H , 宽 W , 通道数 C , 即 $H \times W \times C \times r^2$ 的张量元素重新映射为形状 $rH \times rW \times C$ 的张量. 由于周期性混洗操作直接将低分辨率特征图重建成高分辨率图像, 节省了卷积运算的计算成本, 因此, 与其他卷积反卷积操作相比, 其花费的时间更少. 这个周期性混洗算子操作可以描述为

$$\text{SR}(P)_{x,y,C} = P_{\lfloor \frac{x}{r} \rfloor, \lfloor \frac{y}{r} \rfloor, C \cdot r \cdot \text{mod}(y,r) + C \cdot \text{mod}(x,r) \cdot r}$$

2.2.3 隐含层 深度神经网络的体系结构由多层参数化非线性模块组成, 每个模块的参数都经过学习. 添加的隐藏层越多, 网络学习的功能就越多. 此外, 与 ESPCN 算法相比, 文中算法增加了一个新的层, 用以利用神经网络内部的高频信息.

在 ESPCN 算法的网络结构中, 卷积核的数量分别为 64 和 32, 这两层应用了 Tanh 激活函数. 隐藏层由 3 个卷积层组成. 众所周知, 当添加更多卷积神经网络层时, 可以从图像中提取更多特征. 文中算法提出的隐含层网络结构为: Conv1 表示隐藏层中的第一层. 然后, Conv1 和 Conv3 包含相同的卷积核数 (32), 相同大小的卷积核 (3×3) 和相同的激活函数 (Tanh). 在 Mini-network 的结构中, 涉及 2 个卷积内核级联, 即大小为 3×3 的卷积内核和 ReLU 激活函数.

2.3 重建质量的改进

分析比较 3 种基于深度学习的重建算法, SRCNN 算法虽然结构简单, 但其重建质量远超过传统训练字典的重建方法; 而 VDSR 重建算法增加了很多隐含层, 以增加对高频信息的捕获能力, 但它是以牺牲重建的时间为代价, 重建质量的提高伴随着重建时间的增加; ESPCN 算法以速度取胜, 在重建质量上超过了初代的 SRCNN 算法, 但在整体的结构上可以做出一定的改良, 吸取 VDSR 算法的经验, 在神经网络的隐含层上增加并调整合适的参数, 在增加网络层的同时, 保证速度不会慢太多. 针对以上 3 种算法的优缺点, 采用各个算法的优点去弥补其相应的缺点, 再通过调节参数, 构成文中算法的整个神经网络框架.

通过实验证明, 文中提出的网络结构大大提高了超分辨率重建的质量. SRCNN 算法^[11]首先使用预处理操作, 即先将原图像放大 r 倍, 这样在卷积神经网络中加入池化层可以减少计算量, 而不会对原图

像的细节信息产生太大的影响. 而文中算法是直接将低分辨率的图像作为输入, 如果加入池化层, 会直接减少低分辨率图像中的特征点, 严重影响图片重建的质量, 因此, 文中算法不使用池化层. 单幅图像的高频细节信息与过滤器提取的感受野有关, 在 Mini-network 中, 应确保 2 个级联 3×3 卷积核与 ESPCN 算法中 5×5 模型所提取的感受野大小相同. 此外, 还添加一个新的卷积层以增强对低分辨率图像中高频信息的提取. 深层神经网络结构由多层参数化非线性模块组成, 每个模块的参数都需要学习, 添加的隐藏层越多, 网络结构学到的特征就越多, 重建出来图片的细节信息也就越多, 重建质量也得到了相应的提升.

3 实验结果

实验环境包括硬件设备和软件配置, 测试用的计算机配置为 Intel[®] Core(TM) i7-6700CPU@3.40 GHz, GPU 为 NVIDIA GeForce GTX 1050 Ti. 实验平台搭载的操作系统是 64 位 Windows 7, Caffe, Matlab R2016a, CUDA Toolkit v8.0 和 Anaconda 2.

3.1 公共数据集上测试

在测试评价期间, 选择使用广泛的 5 种公开数据集 (Set5, Set14, BSD300, BSD500 和 SuperTexture) 进行测试. 将文中算法与以速度著称的 ESPCN 算法比较重建时间, 同时, 与传统的 Bicubic 算法比较重建质量, 采用峰值信噪比 (R_{PSN}) 值评判重建质量的好坏. 统计 3 种算法在 5 种数据集上的重建质量和重建时间 (t), 实验结果如表 2 所示. 由表 2 可知: 在同样的数据集上, 文中算法重建的图像质量远高于传统的 Bicubic 算法, 且与基于深度学习的 ESPCN 算法相比, 峰值信噪比提高了 0.37 dB; 在算法运行时间上, 文中算法在各个数据集上的平均速度比 ESPCN 算法快了 50 ms.

表 2 文中算法与 ESPCN 算法的参数对比

Tab. 2 Parameters comparison between the algorithm of this paper and ESPCN algorithm

数据集	放大倍数	R_{PSN}/dB			t/s	
		Bicubic 算法	ESPCN 算法	文中算法	ESPCN 算法	文中算法
Set5	3	32.611	35.086	35.456	0.365	0.220
Set14	3	32.256	35.379	35.735	0.236	0.251
BSD300	3	36.258	39.197	39.764	0.277	0.201
BSD500	3	32.898	35.131	35.309	0.241	0.217
SuperTexture	3	23.031	28.901	29.368	0.217	0.187

3 种算法在数据集 Set5, Set14, BSD300 上的重建质量效果图, 如图 2~4 所示.

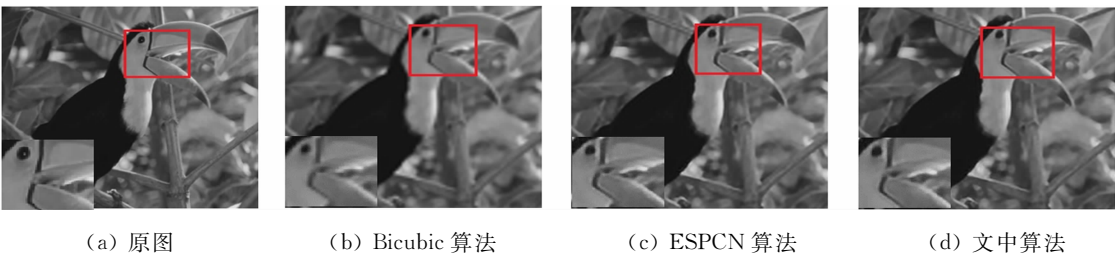


图 2 在 Set5 数据集上的重建结果

Fig. 2 Reconstruction results on the Set5 data set

3.2 算法的应用

以实际生产项目出发, 使用提出的超分辨率重建算法降低使用工业相机的成本, 从一个低分辨率的图像重建出满足实际生产要求的高分辨率图像. 具体有以下 3 个步骤.

步骤 1 将低分辨率的图像作为输入, 直接输入网络.

步骤 2 得到重建后的图像, 进行边缘检测: 滤波, 增强和检测. 使用大津算法^[16]将图像阈值化, 采用霍夫变换^[17]检测直线.

步骤 3 检测出待检测的物体.

重建前的图像分割过程, 如图 5 所示. 在不改变实验阈值的条件下, 输入像素 213 px \times 160 px 的低



(a) 原图 (b) Bicubic 算法 (c) ESPCN 算法 (d) 文中算法

图 3 在 Set14 数据集上的重建结果

Fig. 3 Reconstruction results on the Set14 data set



(a) 原图 (b) Bicubic 算法 (c) ESPCN 算法 (d) 文中算法

图 4 在 BSD300 数据集上的重建结果

Fig. 4 Reconstruction results on the BSD300 data set

分辨率的原始图像(图 5(a)),在阈值分割时出现了明显的边缘不闭合现象(图 5(b)),料盘中的小方块右下角出现缺口,使得在检测物体时,目标丢失(图 5(c)).

重建后的图像分割过程,如图 6 所示. 将像素为 $213\text{ px}\times160\text{ px}$ 的低分辨率图像(图 6(a)),输入网络训练后重建出像素大小为 $615\text{ px}\times456\text{ px}$ 的清晰图(图 6(b)),阈值分割后出现完整的小方块闭合图像,顺利完成物体的精确定位(图 6(d)).



(a) 原图 (b) 阈值分割图 (c) 目标丢失图

图 5 重建前的图像分割过程

Fig. 5 Image segmentation process before reconstruction



(a) 原图 (b) 清晰图 (c) 阈值分割图 (d) 精准检测图

图 6 重建后的图像分割过程

Fig. 6 Image segmentation process after reconstruction

4 结论

分析并总结 3 种基于深度学习的重建算法,调整合适的参数和网络结构以达到图像重建的最优效果. 根据在公共数据上测试的结果,文中算法比 ESPCN 算法的峰值信噪比提升了 0.37 dB ,在图像重建的质量上有一定程度的提升;在重建时间上,文中算法相比 ESPCN 算法快了 50 ms . 在工程的实际应用中,文中算法只需要 0.37 s 就可以完成图像重建,远远小于 SRCNN 卷积神经网络算法的 4.30 s . 因此,所提算法的网络结构在速度上有较大的优势,能满足实际工业场景中实时应用的要求. 值得一提的是,文中算法在实际应用中取得了很好的效果,大大减小了对昂贵工业相机的依赖,使用低分辨率的相机就可以满足实际生产中的应用需求,为企业节省了一大笔经济开支.

然而,文中算法仍存在问题,在物体的纹理处理上还是过于平滑,如遇到车轮的纹路等重建问

题时,还不能很好地提取出特征,不能将网络结构设计地更深,并提取出更为明显的体征.因此,每层卷积层参数的优化和改进将是以后研究的重点.

参考文献:

- [1] 曾凯,丁世飞.图像超分辨率重建的研究进展[J].计算机工程与应用,2017,53(16):29-35. DOI:10.3778/j.issn.1002-8331.1705-0097.
- [2] YANG C Y, MA Chao, YANG M H. Single-image super-resolution: A benchmark[C]//European Conference on Computer Vision. Switzerland:Springer,2014:372-386. DOI:10.1007/978-3-319-10593-2_25.
- [3] YANG Jianchao, WRIGHT J, HUANG T, *et al.* Image super-resolution as sparse representation of raw image patches[C]//IEEE Conference on Computer Vision and Pattern Recognition. Anchorage:IEEE Press,2008:1-8. DOI:10.1109/CVPR.2008.4587647.
- [4] HE Hu, KONDI L P. A regularization framework for joint blur estimation and super-resolution of video sequences [C]//Proceedings of the 2005 IEEE International Conference on Image Processing. Genova:IEEE Press,2005:329-332. DOI:10.1109/ICIP.2005.1530395.
- [5] LI Xiaoguang, LAM K M, QIU G P, *et al.* Examplebased image super-resolution with class-specific predictors[J]. Journal of Visual Communication and Image Representation,2009,20(5):312-322. DOI:10.1016/j.jvcir.2009.03.008.
- [6] SU Congyong, ZHUANG Yueting, LI Huang, *et al.* Steerable pyramid-based face hallucination[J]. Pattern Recognition,2005,38(6):813-824. DOI:10.1016/j.patcog.2004.11.007.
- [7] JIJI C V, CHAUDHURI S. Single-frame image super-resolution through contourlet learning[C]//EURASIP Journal on Advances in Signal Processing. New York:Springer,2006:235. DOI:10.1155/ASP/2006/73767.
- [8] 张晓燕,秦龙龙,钱渊,等.一种改进的稀疏表示超分辨率重建算法[J].重庆邮电大学学报(自然科学版),2016,28(3):400-405. DOI:10.3979/j.issn.1673-825X.2016.03.020.
- [9] FREEMAN W T, PASZTOR E C, CARMICHAEL O T. Learning low-level vision[J]. International Journal of Computer Vision,2000,40(1):25-47. DOI:10.1109/ICCV.1999.790414.
- [10] FREEMAN W T, JONE T R, PASZTOR E C. Example-based super-resolution[J]. IEEE Computer Graphics and Applications,2002,22(2):56-65. DOI:10.1109/38.988747.
- [11] DONG Chao, CHEN C L, HE Kaiming, *et al.* Image super-resolution using deep convolutional networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,2016,38(2):295-307. DOI:10.1109/TPAMI.2015.2439281.
- [12] 刘晨羽,蒋云飞,李学明.基于卷积神经网络的单幅图像超分辨率重建算法[J].计算机辅助设计与图形学学报,2017,29(9):1643-1649. DOI:10.3969/j.issn.1003-9775.2017.09.007.
- [13] KIM J, LEE J K, LEE K M. Accurate image super-resolution using very deep convolutional networks[C]//IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas:IEEE Press,2016:1646-1654. DOI:10.1109/CVPR.2016.182.
- [14] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[J/OL]. Computer Science,2014.[2017-06-10]. <http://arxiv.org/pdf/1409.1556>.
- [15] SHI Wenzhe, CABALLERO J, HUSZAR F, *et al.* Real-time single image and video super-resolution using an efficient subpixel convolutional neural network[C]//IEEE Conference on Computer Vision and Pattern Recognitionhttp. Las Vegas:IEEE Press,2016:1874-1883. DOI:10.1109/CVPR.2016.207.
- [16] OTSU N. A threshold selection method from gray-level histograms[J]. Automatica,1979,9(1):62-66. DOI:10.1109/TSMC.1979.4310076.
- [17] XU Xiaohong, WU Zhihui, CHEN Yu, *et al.* Plant root spatial distribution measurements based on the hough transformation[J]. Neurocomputing,2014,145:209-220. DOI:10.1016/j.neucom.2014.05.041.

(责任编辑:黄晓楠

英文审校:吴逢铁)