

基于深度卷积神经网络的图像去噪研究

李传朋, 秦品乐, 张晋京

(中北大学 计算机与控制工程学院, 太原 030051)

摘 要: 为更有效地去除图像中的噪声, 提出一种深度学习的图像去噪方法。以完整图像作为网络的输入及输出图像, 通过隐含层构成由含噪声图像到去噪图像的非线性映射, 由卷积子网和反卷积子网构成一种对称式的网络结构, 卷积子网学习图像特征, 反卷积子网根据特征图恢复原始图像, 并结合修正线性单元获取更多的纹理细节。以 VOC2012 数据集作为训练集, 使用 Tensorflow 在 GPU 环境下训练网络模型。实验结果表明, 与 GSM, KSVD, CN2, MLP 方法相比, 该方法能更有效地去除图像中的噪声, 获得更高的峰值信噪比, 耗时更短, 视觉效果更佳。

关键词: 卷积神经网络; 图像去噪; 反卷积; 深度学习; 修正线性单元

中文引用格式: 李传朋, 秦品乐, 张晋京. 基于深度卷积神经网络的图像去噪研究 [J]. 计算机工程, 2017, 43(3): 253-260.

英文引用格式: Li Chuanpeng, Qin Pinle, Zhang Jinjing. Research on Image Denoising Based on Deep Convolutional Neural Network [J]. Computer Engineering, 2017, 43(3): 253-260.

Research on Image Denoising
Based on Deep Convolutional Neural Network

LI Chuanpeng, QIN Pinle, ZHANG Jinjing

(College of Computer Science and Control Engineering, North University of China, Taiyuan 030051, China)

【Abstract】 In order to remove the noise in the image more effectively, a method based on deep learning is proposed. The method uses the integrated image as the input and output of the network, and uses hidden layer to compose a nonlinear mapping from the noisy image to denoised image. The network has a symmetric network structure consisting of convolution subnet and deconvolution subnet. Convolution subnet learns about image features, and the deconvolution subnet recovers the original image on the basis of characteristic graph and obtains more texture details combining rectified linear unit. This method uses the VOC2012 data set as the training set and Tensorflow to train the network model in the GPU environment. Experimental result demonstrates that this method can remove the noise in the image more effectively and obtain a higher Peak Signal to Noise Ratio (PSNR). In addition, it takes shorter time, and has better visual effect and perfect practicability.

【Key words】 convolutional neural network; image denoising; deconvolutional; deep learning; rectified linear unit

DOI: 10.3969/j.issn.1000-3428.2017.03.042

0 概述

图像在采集、传输过程中受环境、设备、人为因素等影响不可避免地会引入噪声, 使图像的质量变差, 从而影响图片的可读性, 在处理图像时也会对处理结果造成影响。图像去噪就是为了减少或者消除噪声对图像的影响, 从而获得误差较小的原始图像。目前有多种方法应用于图像去噪, 一种是基于变换域的去噪方法, 如高斯尺度混合模型中的贝叶斯最

小二乘法 (Bayes Least-Squares in Gaussian scale mixture, BLS-GSM)^[1]; 另一种是基于统计信息的去噪方法, 该方法从图像数据中找到一组稀疏线性组合的完备字典, 利用该稀疏编码方式重建图像, 从而达到去噪的效果, K 奇异值分解 (Singular Value Decomposition, SVD)^[2] 就是一种字典训练算法。文献 [3] 提出的基于块匹配的 3D 滤波 (Block-matching and 3D Filtering, BM3D) 去噪方法也是一种统计学习方法, 该方法将图像分为一定大小的块, 将具有相似

基金项目: 山西省自然科学基金 (2015011045)。

作者简介: 李传朋 (1991—), 男, 硕士研究生, 主研方向为深度学习、机器视觉、数字图像处理; 秦品乐, 副教授、博士; 张晋京, 硕士研究生。

收稿日期: 2016-03-31

修回日期: 2016-05-12

E-mail: lichuanpeng08@126.com

特征的块合并成三维数组,使用联合滤波的方法处理三维数组,最后通过逆变换获得去噪后的图像。

K-SVD 等方法使用的是浅层的线性结构,训练字典的方式学习能力有限,国内学者也提出了针对该问题的改进方法^[4-5]。随着深度学习的发展,其在语音、图像、自然语言等领域取得了显著进展^[6],而深层的非线性神经网络结构也适用于图像去噪。目前深度学习中的多种算法已应用到图像去噪中,并取得了较好的效果。如文献[7]提出的一种多层感知机(Multi Layer Perceptron, MLP)去噪网络,该网络由4个全连接层组成,使用VOC2007数据集^[8]和McGill数据集^[9]训练网络,通过滑动窗口的方式对每个窗口内的图像去噪,最后输出去噪图像;栈式稀疏去噪自编码(Stacked Sparse Denoising Auto-encoder, SSDA)^[10]也可应用于去噪以及图像修复,与MLP同样使用滑动窗口的方式对每块图像去噪,然后合成整张图像;深度置信神经网络(Deep Belief Network, DBN)也被用于图像的去噪中^[11],网络的单元个数限制了输入图像的大小,需要以片段的方式输入,这也影响了去噪的效果和效率。此外,文献[12-13]分别将卷积神经网络应用于图像去噪和图像超分辨率处理,并取得了很好的效果,其中文献[12]提出的CN2网络借鉴了全连接的网络结构,使用4层的全卷积结构,并使用相同大小的感受野提取特征,但相同的感受野大小不容易获取不同大小的纹理细节,而且深度的卷积神经网络能够学习出更抽象的特征,这有利于图像的识别和分类,但是过多的抽象特征往往难以整合,这会影响以整个图像作为输出的神经网络的去噪效果。

本文在充分考虑去噪速度和去噪效果的基础上,构建一种用于图像去噪的卷积神经网络,该网络由卷积层、反卷积层^[14-15]、ReLU层^[16]组成,并由卷积子网和反卷积子网构成一种对称式的网络。该网络以完整的含噪图像作为输入,以去噪图像作为输出,构成由噪声图像到原始图像的非线性映射。本文以VOC2012数据集^[17]作为训练集,使用Tensorflow^[18]在GPU上训练网络参数。

1 卷积神经网络

卷积神经网络是一种人工神经网络,目前已成为图像识别、语音分析、自然语言处理等领域的研究热点。卷积神经网络是一种权值共享的网络结构,与深度学习中的MLP、DBN等网络相比,具有较少的可调参数,降低了学习复杂度;卷积神经网络在处理二维图像时,对平移、倾斜、缩放或其他形式的形变具有高度不变性。卷积神经网络的优良特性归因于其网络结构,图1为一种卷积神经网络的典型结构,

该网络由输入层、卷积层、池化层、全连接层及输出层组成。网络输入二维图像,卷积层和池化层分别提取和映射图像特征,针对分类任务的网络,一般加入全连接层学习全局特征,输出层的神经单元由网络用途和网络结构决定。

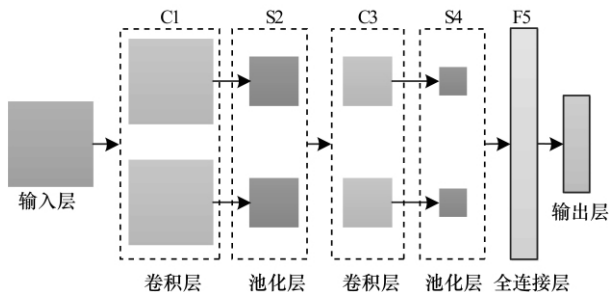


图1 典型卷积神经网络

卷积神经网络隐含层一般由特征提取层和特征映射层组成,特征提取时,神经单元与前一层的局部感受野连接,并提取局部特征,将每个特征图的特征进行组合即可得到全局特征,卷积神经网络的神经单元无需与上一层所有神经单元连接,从而减少了可调参数的数目。这种共享权值的方式在图像分类中能够有效地减少训练权值的数目,这对针对图像去噪的卷积神经网络十分有利,将会大幅减少图像去噪的训练时间,同时训练完网络后也会减少去噪处理的耗时。

2 基于深度卷积神经网络的图像去噪

图像去噪的目的是从含有被噪声腐蚀的图像中恢复图像的原始信息,该过程可以由下式表示^[19]:

$$g(x, y) = f(x, y) + \eta(x, y) \quad (1)$$

其中, $g(x, y)$ 为含有噪声的二维图像; $f(x, y)$ 为无噪声的原始图像; $\eta(x, y)$ 为添加的噪声。原始图像经过加性噪声项处理,退化为含噪图像。图像去噪就是为了获得原始图像的一个估计 $\hat{f}(x, y)$, 并使其尽可能接近原始图像。

通过先验知识学习输入图像中的噪声特征在去噪过程中至关重要,MLP等方法均是通过大量的原始图像和噪声图像学习噪声特征。MLP、SSDA以及DBN将图像的去噪任务划分到多个片段中,将二维图像片段转化为一维向量,训练过程和去噪过程均需要输入一维向量,输出同样也是一维向量,输出向量转化为二维图像片段后拼接即可得到去噪后的图片。这种网络结构存在2个主要的弊端:1) 输入输出的神经单元个数直接限制了网络能处理图像片段大小,同时也将噪声特征限制在片段中,全局特征无法获取;2) 网络中采用了多个全连接层,这很大程度地增加了网络可调参数的数量,对学习效率会造成影响。CN2设计了4层的卷积结构,每个卷积层有24个特征图,

并借鉴了全连接的网络结构,这种全卷积的网络结构充分考虑了特征提取在去噪中的作用,但是过多的抽象特征不容易整合。本文构建了一种对称式的卷积神经网络,通过调节卷积的核大小和特征图数量的方式更大程度地学习含噪声图像特征,此外网络加入了反卷积子网,充分整合卷积子网学习到的抽象特征,与卷积子网产生映射,从而使整个网络构成了由含噪图像到去噪图像的非线性映射。

2.1 网络结构

去噪卷积神经网络的设计充分考虑网络在含噪图像与去噪图像的关系,即网络实际是一种由含噪图像到去噪图像的非线性映射关系。图2是本文构建的一种用于去噪的卷积神经网络,包括输入输出层、卷积子网和反卷积子网。含噪声的二维图像(单通道)从输入层输入,网络隐含层均由特征图组成,网络输入层不限制图像的大小,可以输入任意宽高的图像。输入图像经对称式的隐含层网络,在输出

层给出与输入图像同等宽高的二维图像(单通道)。网络隐含层由多个卷积层和反卷积层组成,图像输入网络后,使用 (5×5) 的卷积核对图像进行卷积操作,为每个特征图添加偏置,形成C1层,包含32幅特征图,需要训练的参数个数为 $5 \times 5 \times 1 \times 32 + 32$;然后减小卷积核大小到 (3×3) ,增加特征图数量,并与上一层的特征图全连接,卷积操作后形成含有64个特征图的C2层,需要训练的参数个数为 $3 \times 3 \times 32 \times 64 + 64$;C2层与64个大小为 (1×1) 的核进行卷积,对应生成64个特征图,需要训练的参数个数为 $1 \times 1 \times 64 \times 64 + 64$;D1层经 (3×3) 的反卷积操作,生成D2层,包含32个特征图,需要训练的参数个数为 $3 \times 3 \times 64 \times 32 + 32$;增大卷积核大小为 (5×5) ,将输出特征图大小设置为1,需要训练的参数个数为 $5 \times 5 \times 32 \times 1 + 1$,此时反卷积输出的结果就是网络输出层输出的去噪图像。

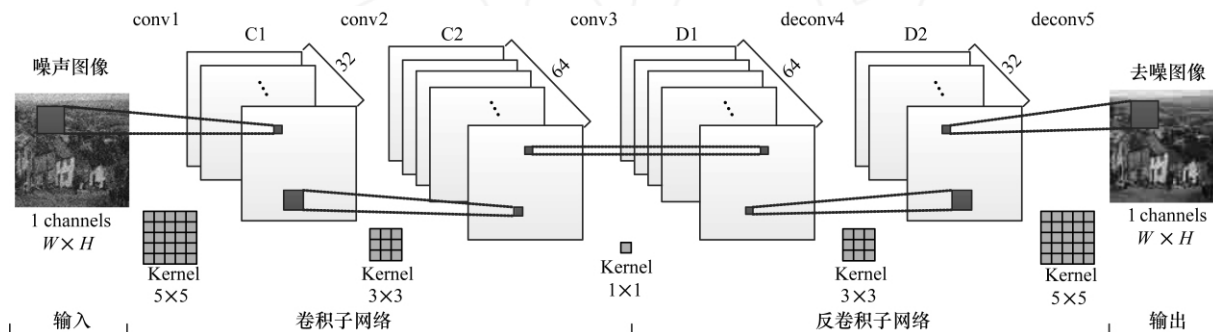


图2 去噪卷积神经网络

该网络具有以下特点:

1) 输入和输出。本文所提出的网络接受原始图像作为输入,并经过隐含层产生一个完整的图像作为输出,原始图像直接输入网络避免了复杂耗时的图像预处理过程,输出层输出的是去噪后的图像,该图像与原始图像大小一致,从而减少了以片段形式输出所需要的拼接工作。

2) 非线性映射。本文期望通过卷积神经网络构造一种由含噪图像到去噪图像的非线性映射,卷积神经网络的过滤器能够完成从特征图到另一特征图的映射。本文使用不同大小的过滤器构造了一种对称式的网络结构,可以有效地共享权值,同时做到图像到图像的非线性映射。

3) 学习。通过批量原始图像的学习方式是低效的,本文采用片段式的学习方式。首先无噪声的图像数据集经过加噪处理获得网络所需的含噪图像集,并从该数据集中随机采样,获得片段大小相同的训练数据。该网络是一种有监督学习的网络,由训练时输入含噪的原始图像片段和网络输出计算网络损失,并通过随机梯度优化网络参数。

2.2 卷积层

在卷积层,特征图与 (5×5) , (3×3) 等大小的卷积核直接进行卷积运算会减小下一层特征图的大小,因此考虑为特征图扩展边界,将每个特征图上下各扩展 $(kernel\ size - 1) / 2$ 行0像素,左右各扩展 $(kernel\ size - 1) / 2$ 列0像素,如图3所示。卷积核的中心自图像左上角滑动,滑动步长设置为1个像素,滑动范围为 $W \times H$, W 和 H 为图像的宽和高。

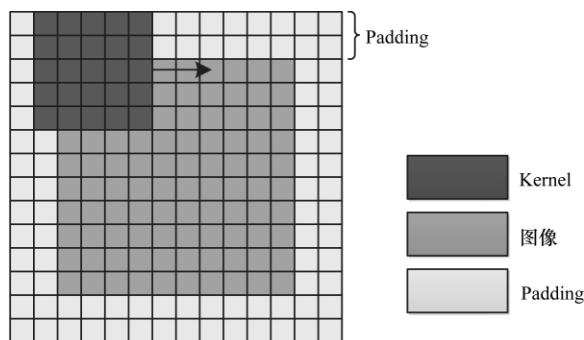


图3 卷积神经网络边界扩展

卷积层的计算形式为:

$$Y_j^l = f(\sum_{i \in M_j} W_{ij}^l * Y_j^{l-1} + B^l) \quad (2)$$

其中, l 表示第几层; Y_j^l 为第 l 层的第 j 个特征图; W 为卷积核; B 为第 j 个特征图对应的偏置; $*$ 为卷积操作。卷积核 W 相当于 n^l 个过滤器, 过滤器大小为 $k^l \times k^l$, k^l 一般取 1, 3, 5 等。 $f(\cdot)$ 函数为激活函数, 采用修正线性单元 (Rectified Linear Unit, ReLU) 激活函数, 形式为:

$$f(x) = \max(0, x) \quad (3)$$

其中, x 是特征图中的任意神经单元, ReLU 是一种非线性的激活函数, 与神经元信号激励原理类似, 具有单侧抑制、兴奋边界宽阔和稀疏激活的特点。卷积过程如图 4 所示。

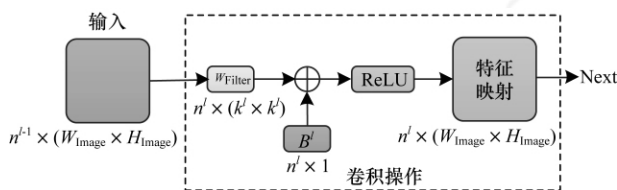


图 4 卷积过程

2.3 反卷积层

卷积层的作用是在噪声图像中学习噪声特征, 并经卷积核生成多幅特征图, 这些特征在识别和分类中是至关重要的, 但是过多的抽象特征往往难以整合, 这会影响以整个图像作为输出的神经网络的去噪效果。而反卷积层的目的就是通过特征图获取去噪的输出图像, 与文献 [14] 所提出的 Deconvolutional Network 不同的是, 本文的反卷积实际上是卷积的逆过程, 其过程如图 5 所示。卷积通过一个过滤器窗口将含有多个激活单元的区域连接到一个激活单元, 而反卷积则是将一个激活单元连接到多个输出单元。经过学习的反卷积核能够在卷积层输出的特征图中重建原始图像。多层的反卷积神经网络能够从不同的特征图中获取更多的图像细节, 从而更好地恢复图像。

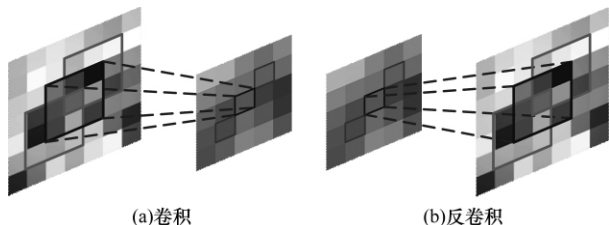


图 5 卷积和反卷积过程

2.4 训练

构建好网络后, 需要对网络进行训练, 以优化网络的可调参数, 包括各层的卷积核 W 和偏置 B 。峰

值信噪比 (Peak Signal to Noise Ratio, PSNR) 能够客观地评价网络的去噪效果, 而峰值信噪比的大小直接与图像之间的均方误差 (Mean Squared Error, MSE) 有关, 因此本文使用均方误差作为网络的代价函数, 代价函数为:

$$L = \frac{1}{n} \sum_{i=1}^n \left(\frac{1}{wh} \sum_{j=1}^w \sum_{k=1}^h \| D_i(j, k) - X_i(j, k) \|^2 \right) \quad (4)$$

其中, n 为每个训练批次的样本数量; w 和 h 分别为每个样本的宽高; D 为经网络去噪后的图像; X 为无噪声的原始图像。在网络训练过程中, 每次输入网络 n 个样本, 经卷积网络输出后与对应的 n 张无噪图像计算代价函数 L , 然后使用标准反向传播 (Back Propagation, BP) [20] 算法逐层反向传播误差, 并更新各层的卷积核权重和偏置。权值更新方式如下:

$$W_{t+1}^l = W_t^l + \Delta_{t+1}^l \quad (5)$$

$$\Delta_{t+1}^l = -\eta \frac{\partial L}{\partial W_t^l} \quad (6)$$

其中, t 为批次; l 代表第几个卷积层; η 为第 l 层的学习率。误差对该层权值的偏导数乘以一个负学习率就是该层权值的更新值, 其中学习率通常取 $10^{-4} \sim 10^{-5}$ 。权值在网络训练开始前需要初始化, 并使每层权值 W 的初值符合均值为 0、标准差为 0.001 的高斯分布, 偏置初始值都设置为 0。

3 实验与结果分析

3.1 实验平台与数据集

卷积神经网络的训练过程需要大量的矩阵运算, 图形处理单元 (Graphics Processing Unit, GPU) 比中央处理单元 (Central Processing Unit, CPU) 更适合这样的操作, 因此在 GPU 上实现卷积神经网络的去噪训练, 使用的 GPU 为 NVIDIA Tesla K40。同时使用 Tensorflow 深度学习框架进行实验。

本文训练所采用的图像来自 VOC2012, 部分图像如图 6(a) 所示, 随机选取 15 000 张图像用作训练集。VOC2012 的所有图像均为彩色图像, 灰度图像去噪网络需要首先将彩色图像先转化为灰度图, 并添加噪声, 添加噪声的部分图像如图 6(b) 所示。

训练所使用的输入输出数据均为图像片段, 本文使用滑动窗口的方式产生图像片段, 滑动窗口大小为 $f_{patch} \times f_{patch}$, 步长取 $\lfloor w_{patch}/2 \rfloor$ 。对所有图像按照此种方式产生图像片段, 生成含有大量含噪图像片段和无噪图像片段的训练集, 为了使网络学习能力更强, 本文还将训练集进行乱序操作。图 6(c) 为无噪图像片段集的部分图像, 图 6(d) 为图 6(c) 中无噪图像添加噪声后的图像。

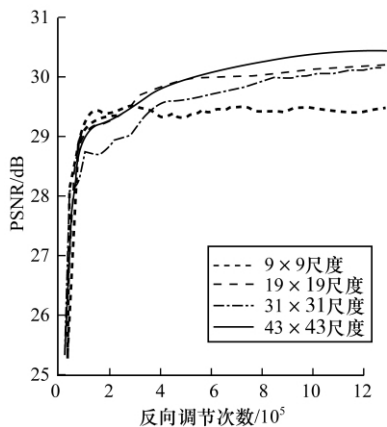


图6 训练集

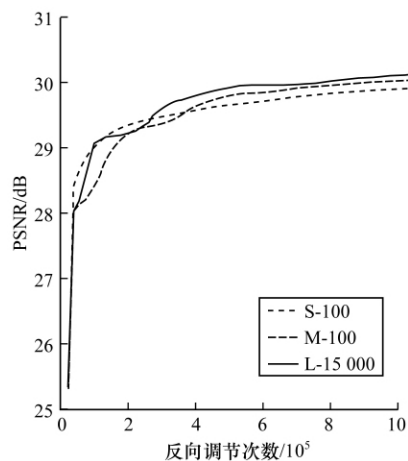
3.2 训练数据对去噪性能的影响

训练数据能影响图像的去噪效果,为了对比训练数据对去噪性能的影响,本文针对训练图像片段大小和训练图像数目这3种影响因素设计了实验。

为了对比训练片段大小对网络学习能力的影响,以多种尺度的训练集训练网络,包括 19×19 的小尺度训练集以及 43×43 的大尺度训练集。训练输入样本噪声均采用 $\sigma = 25$ 的高斯噪声。为了显示网络在不同数据集下的去噪表现,使用经 $\sigma = 25$ 的高斯噪声处理的 Lena 图像作为测试图像。图7显示了多种尺度训练集训练过程的峰值信噪比。由图7可知,在网络训练初期,小尺度的训练集因为包含的特征较少,所以学习效率较高,但随着反向调节数目的增多,大尺度的训练集能够学习到更多的特征,PSNR 一直呈上升趋势,好于小尺度训练集的 PSNR。训练图像片段在一定范围内越大,所包含的噪声特征越多,去噪效果更好,但网络达到稳定所需要的训练周期也越多,从而需要更长的训练时间。

图7 不同尺度训练集的训练过程($\sigma = 25$)

数据集的大小同样会对去噪性能产生影响。本文使用不同大小的数据集训练网络,包括含有100张图像的小数据集(S-100)、含有1000张图像的中等数据集(M-1000)以及含有15000张图像的大数据集(L-15000)。各种训练集所使用的图像均在VOC2012数据集中随机抽取,并且均使用 $\sigma = 25$ 的高斯噪声处理。图8显示了3种数据集训练过程的峰值信噪比。可以看出,训练样本的多少会对去噪性能产生影响,小的数据集更容易学习到样本的特征,但能学习出的特征有限,而大的数据集虽然训练初期学习效率较低,但是经过多个周期和大量样本学习后,去噪效果好于小数据集所训练的网络。

图8 不同大小训练集的训练过程($\sigma = 25$)

3.3 网络模型对去噪性能的影响

深度学习中增加网络的深度往往能够使网络学习到更多的细节,得到更优的效果,例如 AlexNet^[21]和 GoogLeNet^[22]均采用了多层的卷积神经网络,在图像识别领域取得了突破性进展。深度卷积神经网络能够学习更抽象的数据特征,因此深层次的网络应该更有利于图像噪声的去除。基于这个构想,本文构建了几种不同的神经网络,用以比较网络模型对去噪性能的影响。

为了对比各种网络模型对去噪效果的影响,使用峰值信噪比(Peak Signal to Noise Ratio, PSNR)和去噪耗时作为评价标准。同时使用10张图像(lena, barbara, boat, couple, fingerprint, hill, man, house, montage, peppers)作为测试图像,向网络输入添加高斯噪声($\sigma = 25$)的测试图像,获得网络输出的去噪图像,通过计算这10张图像的平均峰值信噪比和平均耗时的方式比较各种网络的去噪效果。实验环境为:(MacBook Pro, Intel Core i5, Python 2.7, Tensorflow 0.5)。表1显示了本文构建的各种卷积神经网络的网络结构以及去噪性能对比,其中训练周期为10,训练样本为5000张图像。

表 1 卷积神经网络结构及去噪性能($\sigma=25$)

层数	卷积核	特征图数	平均 PSNR/ dB	平均耗时/s
4	$3 \times 1 \times 3$	$1 \times 32 \times 32 \times 1$	26.821 3	0.137 7
4	$5 \times 1 \times 5$	$1 \times 32 \times 32 \times 1$	27.174 5	0.222 9
4	$5 \times 1 \times 5$	$1 \times 64 \times 64 \times 1$	28.729 5	0.348 8
5	$5 \times 3 \times 3 \times 5$	$1 \times 32 \times 64 \times 32 \times 1$	29.170 4	0.995 3
6	$5 \times 3 \times 1 \times 3 \times 5$	$1 \times 32 \times 64 \times 64 \times 32 \times 1$	29.574 9	1.156 1

由表 1 可以看出卷积层数的增加会提高图像的去噪效果,但是会以相应的时间损耗作为代价。4 层的卷积神经网络处理一张 512×512 的图像通常只需要 0.3 s,而 6 层的卷积神经网络则需要 1.5 s。在训练网络的过程中差距表现得更加明显,相同的训练周期,6 层的卷积神经网络要比 4 层的卷积神经网络多消耗几个小时甚至几天的时间。另一方面,卷积核的大小以及特征图的数量同样会影响网络的去噪性能。图 9 为 6 层卷积神经网络第 1 个卷积层的特征图,是由 32 个 5×5 的卷积核卷积形成的,由图中可以看出每个卷积核能学习不同的特征。因此在一定程度上增加特征图的数量会更大程度地学习特征,达到更优的去噪效果。图 10 所示的图像是 6 层卷积神经网络第 1 个反卷积层的输出特征图,与第 1 个卷积层输出相比,该层的特征图包含的噪声特征明显变少,全局特征更加明显。这也表明了随着卷积神经网络层数的增加,神经网络去噪性能会在一定程度上提高。

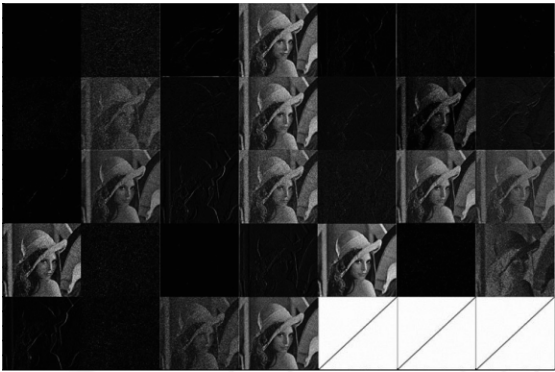


图 9 第 1 个卷积层输出的特征图

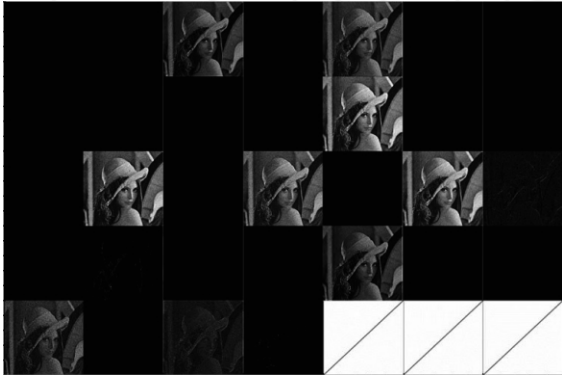


图 10 第 1 个反卷积层输出的特征图

3.4 对比实验分析

为了验证本文所提出的基于深度卷积神经网络去噪方法的有效性,选用 10 幅测试图像进行仿真实验,并将去噪结果与现有性能优越的算法 GSM,KSVD,CN2,MLP 进行对比。在本次仿真实验中,测试图像添加的噪声为 $\sigma=25$ 的高斯白噪声,所有算法均使用 CPU 进行运算处理。使用峰值信噪比 (PSNR) 和处理耗时这 2 种评价指标衡量去噪效果。实验环境为: MacBook Pro, Intel Core i5, Matlab R2013a。表 2 为 GSM,KSVD,CN2,MLP 和本文方法对测试图像的去噪峰值信噪比。表 3 为各方法对 10 幅测试图像去噪的平均 PSNR 和平均耗时。

表 2 不同方法去噪 PSNR 比较($\sigma=25$) dB

图像	噪声	GSM	KSVD	CN2	MLP	本文方法
House	20.23	31.63	32.21	31.32	32.48	32.50
Montage	20.77	29.35	30.53	30.80	31.34	31.85
Peppers	20.32	29.39	29.63	29.55	30.12	30.49
Lena	20.22	31.27	31.31	31.01	32.09	31.94
Barbara	20.29	27.81	29.51	27.40	29.15	29.07
Boat	20.26	29.30	29.23	29.13	29.84	29.85
Couple	20.25	28.85	28.67	28.87	29.56	29.58
Fingerprint	20.26	27.16	27.23	26.28	27.49	27.45
Hill	20.26	29.32	29.22	29.17	29.79	29.69
Man	20.23	29.16	29.08	29.14	29.75	29.71

表 3 不同方法去噪平均性能比较($\sigma=25$)

性能评价因素	GSM	KSVD	CN2	MLP	本文方法
平均 PSNR/dB	29.32	29.66	29.27	30.16	30.21
平均耗时/s	4.20	96.37	4.03	14.22	3.87

从表 2、表 3 可以看出,本文所提出的方法在 $\sigma=25$ 的噪声水平下 PSNR 优于 GSM,KSVD,CN,略好于多层感知机网络 (MLP),而就处理相同的图片所耗时间相比,本文所提出的方法耗时最少。CN2 去噪耗时相对于 GSM 和 KSVD 较少,其使用的是 4 层 24 个特征图的神经网络,去噪时只需输入含噪声的二维图像即可,所以耗时较短。本文利用调整卷积核大小和特征图数量的方式构造对称式的卷积神经网络,所以去噪耗时较 CN2 少,而本文所使用的反卷积层在噪声去除方面有较大的作用,取得了较好的去噪效果。MLP 利用 4 层全连接的神经网络处理图像的片段,然后将所有处理后的图像片段拼接融合,形成去噪图像,因此耗时很大。而本文方法输入输出均为整个图像,对图像的大小也没有限制,能够在取得峰值信噪比较高的情况下所使用的处理时间较短。

图11为GSM,KSVD以及本文所提的基于深度卷积神经网络去噪方法对灰度图的去噪效果对比,3种去噪模型的噪声水平均设置为 $\sigma=25$ 的高斯白噪声,可以看出,3种去噪方法均能有效去除图像中的噪声。观察局部图像可发现,基于深度卷积神经网络的去噪方法能够得到更清晰的边缘并且能够恢复更多的纹理细节,这表明该方法较常见的GSM及KSVD方法去噪视觉效果更好。



图11 去噪效果比较

本文所提出的深度卷积神经网络同样适用于其他噪声级别的图像去噪。图12为各种方法的去噪效果比较,由图可知,基于深度学习的图像去噪方法对训练数据依赖性很强,经 $\sigma=25$ 的训练集训练后,对 $\sigma=25$ 的噪声图像处理效果很好,但对其他级别的噪声并不适用。经过不同级别噪声训练的深度卷积神经网络能有效地处理不同级别的噪声,去噪效果好于GSM和KSVD; $\sigma < 50$ 时,本文方法与MLP去噪效果相近, $\sigma > 50$ 时MLP去噪效果更优,这是因为噪声过大,特征不易学习;MLP网络所需训练的参数有 2×10^7 个,训练样本个数为 5×10^7 ,而本文所提

出的网络所需要训练的参数仅有几个,训练样本个数为 1×10^6 个,所以噪声较大时去噪效果不如MLP。增加卷积神经网络层数、调整卷积核大小及数目以及增加训练样本个数会增强网络的去噪效果,但是会以更长的训练时间以及更多的处理时间作为代价。

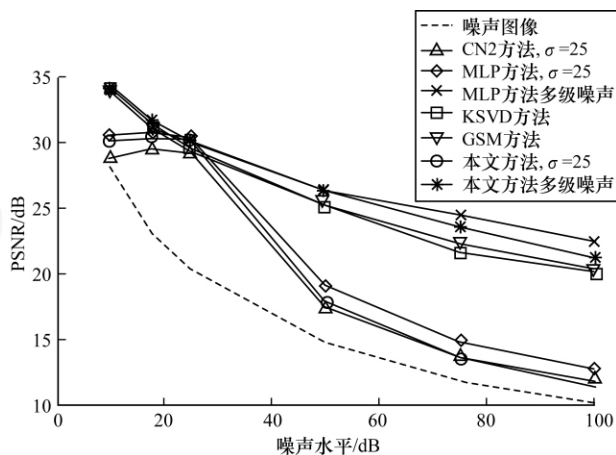


图12 各方法在不同噪声级别下的去噪效果

3.5 椒盐噪声

本文所提出的深度卷积神经网络也可用于其他类型的噪声,如椒盐噪声。椒盐噪声也称双极脉冲噪声,在图像中表现为随机出现的黑点或白点。去除椒盐噪声比较有效的方法是中值滤波,MLP也提出了去除椒盐噪声的网络。本文使用添加20%椒盐噪声的2500张图像制作 5×10^5 个训练样本,网络初始参数与针对高斯噪声的网络相同,训练周期为25。图13为中值滤波、MLP及本文方法对20%椒盐噪声的去噪效果,由图可以看出,本文所提出的方法能够有效地去除椒盐噪声,并且具有很高的峰值信噪比,去噪效果最佳。



图13 各方法对椒盐噪声的去噪效果

4 结束语

本文提出了一种基于深度卷积神经网络的图像去噪方法,并构建了一种对称式的网络结构。该网络由输入层、卷积子网、反卷积子网和输出层构成,输入层直接将含噪的二维图像输入到神经网络,卷积子网充分利用卷积神经网络的特征学习能力,通过调整卷积核大小、特征图数量和隐含层数目的方式调节网络结构,从而学习出更多的特征,反卷积子网整合图像特征,恢复原始图像信息,并由输出层输出完整的去噪图像,形成了一种由含噪图像到去噪图像的非线性映射。本文通过实验分析了训练数据集以及模型本身对去噪效果的影响,实验结果表明,大量训练样本和大尺寸训练片段能够取得更好的去噪效果,卷积神经网络层数的增加也有助于提高去噪效果。通过与常用去噪方法及其他深度学习去噪方法对比,显示出本文所构建的深度卷积神经网络不仅能取得较好的去噪效果,而且耗时更少。

参考文献

- [1] Javier P, Vasily S, Wainwright M J, et al. Image Denoising Using Scale Mixtures of Gaussians in the Wavelet Domain [J]. IEEE Transactions on Image Processing, 2003, 12 (11): 1338-1351.
- [2] Michael E, Michal A. Image Denoising via Sparse and Redundant Representations over Learned Dictionaries [J]. IEEE Transactions on Image Processing, 2007, 15 (12): 3736-3745.
- [3] Dabov K, Foi A, Katkovnik V, et al. Image Denoising by Sparse 3-D Transform-domain Collaborative Filtering [J]. IEEE Transactions on Image Processing, 2007, 16 (8): 2080-2095.
- [4] 杨爱萍, 田玉针, 何宇清, 等. 基于改进 K-SVD 和非局部正则化的图像去噪 [J]. 计算机工程, 2015, 41 (5): 249-253.
- [5] 王智文, 李绍滋. 基于多元统计模型的分形小波自适应图像去噪 [J]. 计算机学报, 2014, 37 (6): 1380-1389.
- [6] 余凯, 贾磊, 陈雨强, 等. 深度学习的昨天、今天和明天 [J]. 计算机研究与发展, 2013, 50 (9): 1799-1804.
- [7] Harmeling S. Image Denoising: Can Plain Neural Networks Compete with BM3D? [C] // Proceedings of IEEE Conference on Computer Vision & Pattern Recognition. Washington D. C., USA: IEEE Press, 2012: 2392-2399.
- [8] Everingham M, Winn J. The Pascal Visual Object Classes Challenge [J]. International Journal of Computer Vision, 2010, 88 (2): 303-338.
- [9] Andriana O, Kingdom F A A. A Biologically Inspired Algorithm for the Recovery of Shading and Reflectance Images [J]. Perception, 2004, 33 (12): 1463-1473.
- [10] Xie Junyuan, Xu Linli, Chen Enhong, et al. Image Denoising and Inpainting with Deep Neural Networks [C] // Proceedings of the 25th International Conference on Neural Information Processing Systems. New York, USA: ACM Press, 2012: 350-358.
- [11] Keyvanrad M A, Pezeshki M, Homayounpour M A. Deep Belief Networks for Image Denoising [C] // Proceedings of International Conference on Learning Representations. Berlin, Germany: Springer, 2014.
- [12] Jain V, Seung H S. Natural Image Denoising with Convolutional Networks [C] // Proceedings of Conference on Neural Information Processing Systems. Berlin, Germany: Springer, 2008: 769-776.
- [13] Dong Chao, Loy C C, He Kaiming, et al. Image Super-resolution Using Deep Convolutional Networks [J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2016, 38 (2): 295-307.
- [14] Zeiler M, Krishnan D, Taylor G, et al. Deconvolutional Network [C] // Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Washington D. C., USA: IEEE Press, 2010: 2528-2535.
- [15] Noh Hyeon-woo, Hong Seung-hoon, Han Boh-yung. Learning Deconvolution Network for Semantic Segmentation [C] // Proceedings of IEEE International Conference on Computer Vision. Washington D. C., USA: IEEE Press, 2015.
- [16] Nair V, Hinton G E. Rectified Linear Units Improve Restricted Boltzmann Machines [C] // Proceedings of the 27th International Conference on Machine Learning. Haifa, Israel [s. n.], 2010: 807-814.
- [17] Everingham M, Gool L, Williams C K I, et al. The Pascal Visual Object Classes Challenge [EB/OL]. (2012-09-03). <http://www.pascal-network.org/challenges/VOC/voc2012/workshop/index.html>.
- [18] Abadi M, Agarwal A, Barham P, et al. Large-scale Machine Learning on Heterogeneous Systems [EB/OL]. (2015-12-12). <http://download.tensorflow.org/paper/whitepaper2015.pdf>.
- [19] Gonzalez R C, Woods R E. 数据图像处理 [M]. 3 版. 阮秋琦, 阮宇智, 译. 北京: 电子工业出版社, 2012.
- [20] Lecun Y, Boser B, Denker J S, et al. Backpropagation Applied to Handwritten Zip Code Recognition [J]. Neural Computation, 1989, 1 (4): 541-551.
- [21] Krizhevsky A, Sutskever I, Hinton G E. Image Net Classification with Deep Convolutional Neural Networks [C] // Proceedings of the 25th International Conference on Neural Information Processing Systems. New York, USA: ACM Press, 2012: 1097-1105.
- [22] Szegedy C, Liu Wei, Jia Yangqing, et al. Going Deeper with Convolutions [C] // Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Washington D. C., USA: IEEE Press, 2014: 1-9.

编辑 顾逸斐