

# 基于复合卷积神经网络的图像去噪算法<sup>\*</sup>

吕永标 赵建伟 曹飞龙

(中国计量大学 理学院 应用数学系 杭州 310018)

**摘 要** 基于深度学习理论,将图像去噪过程看成神经网络的拟合过程,构造简洁高效的复合卷积神经网络,提出基于复合卷积神经网络的图像去噪算法.算法第 1 阶段由 2 个 2 层的卷积网络构成,分别训练阶段 2 中的 3 层卷积网络中的部分初始卷积核,缩短阶段 2 中网络的训练时间和增强算法的鲁棒性.最后运用阶段 2 中的卷积网络对新的噪声图像进行有效去噪.实验表明文中算法在峰值信噪比、结构相相似度及均方根误差指数上与当前较好的图像去噪算法相当,尤其当噪声加强时效果更佳且训练时间较短.

**关键词** 图像去噪,卷积神经网络,随机梯度下降法

中图法分类号 TN 911.71; TP 183

DOI 10.16451/j.cnki.issn1003-6059.201702001

引用格式 吕永标,赵建伟,曹飞龙.基于复合卷积神经网络的图像去噪算法.模式识别与人工智能,2017,30(2): 97-105.

## Image Denoising Algorithm Based on Composite Convolutional Neural Network

LÜ Yongbiao, ZHAO Jianwei, CAO Feilong

(Department of Applied and Mathematics, College of Sciences, China Jiliang University, Hangzhou 310018)

### ABSTRACT

According to the theory of deep learning, the process of image denoising can be regarded as a fitting process of a neural network. In this paper, an image denoising algorithm based on composite convolutional neural network is proposed through constructing a simple and efficient composite convolutional neural network. The first stage includes two convolutional neural networks with two layers. Some initial convolutional kernels of convolutional neural network with three layers in the second stage are trained by these two networks, respectively. The training time in the second stage is decreased and the robustness of the network is enhanced. Finally, the learned convolutional neural network in the second stage is applied to denoise a new image with noises. Experimental results show that the proposed algorithm is comparable to state of the art image denoising algorithms in peak signal to noise ratio (PSNR), structure similarity, and root mean square error (RMSE). Especially, when the noises get heavier, the proposed algorithm performs better with less training time.

<sup>\*</sup> 国家自然科学基金项目(No.61672477, 61571410)资助

Supported by National Natural Science Foundation of China(No.61672477, 61571410)

收稿日期: 2016-08-30; 修回日期: 2016-12-02; 录用日期: 2016-12-13

Manuscript received August 30, 2016; revised December 2, 2016; accepted December 13, 2016

**Key Words** Image Denoising , Convolutional Neural Network , Stochastic Gradient Descent Method

**Citation** LÜ Y B , ZHAO J W , CAO F L. Image Denoising Algorithm Based on Composite Convolutional Neural Networks. Pattern Recognition and Artificial Intelligence , 2017 , 30( 2) : 97–105.

在现实生活中,图像在数字化和传输过程中常受到成像设备和外部环境噪声干扰等影响,导致采集的图像质量下降。但是在很多图像应用领域,如工业生产、医学研究等,对图像质量要求非常高,而图像具有不可再生性,因此图像去噪成为计算机视觉研究领域中的一个重要的研究方向。

近些年,学者们基于各种理论和技术提出多种图像去噪算法,主要是基于稀疏表示理论<sup>[1-3]</sup>、非局部相似性理论<sup>[4-7]</sup>、机器学习理论<sup>[8-10]</sup>及其它一些基于特定噪声原理的去噪方法<sup>[11-14]</sup>。

基于稀疏表示理论的图像去噪方法是利用图像在过完备字典下的稀疏表示实现去噪功能,而该字典可以由图像数据库的图像块学习得到。Aharon 等<sup>[1]</sup>利用小波变换域下的稀疏表示,提出  $K$ -奇异值分解( $K$ -Singular Value Decomposition,  $K$ -SVD)图像去噪算法。Mairal 等<sup>[2]</sup>统一图像的自相似性和稀疏表示,提高图像去噪效果。Dong 等<sup>[3]</sup>利用图像的局部自相似性,提出非局部集中稀疏表示图像去噪算法。

在一幅图像中,通常会存在某些图案,这些图案在整幅图像的不同位置重复出现,这种性质称为非局部相似性。非局部相似性在图像去噪方面有着重要作用。例如相比传统的局部相似性算法,非局部均值(Non-local Means, NLM)<sup>[4]</sup>效果更好。由于自然图像非局部相似性的广泛存在,基于该性质的图像去噪算法也较多,如块匹配三维滤波方法(Block-Matching and 3D Filtering, BM3D)<sup>[6]</sup>。

图像去噪可以看成是将一幅带有噪声的图像映射到无噪声图像的问题。因此,有些文献将图像去噪过程看作映射的拟合过程,而神经网络具有较强的映射逼近能力,如多层感知器(Multi-layer Perceptron, MLP)<sup>[8]</sup>就是利用多层感知器拟合图像去噪过程,取得一定的去噪效果。近些年,基于深度学习的卷积神经网络在图像处理领域受到广泛关注,并取得显著效果。相比前向神经网络,卷积神经网络使用卷积操作代替神经网络中的内积操作,具有深度提取图像特征的效果,成为研究图像的理想模型,并取得较好效果<sup>[9-10]</sup>。

本文引入卷积神经网络处理图像去噪问题,提出基于复合卷积神经网络的图像去噪算法。一般的卷积神经网络由多层卷积层和池化层构成,由于构造的层

数较多,因此需要调节的参数较多,这导致整个网络训练时间过长,收敛较慢。本文构建一个有效的复合卷积神经网络模型,在简化训练的同时达到和大型卷积神经网络相当的去噪效果。

## 1 基于复合卷积神经网络的图像去噪算法

本文提出的基于复合卷积神经网络的图像去噪算法流程图如图 1 所示,包括 2 个阶段:阶段 1 由 2 个 2 层卷积神经网络组成;阶段 2 由 1 个 3 层卷积神经网络组成,模拟图像去噪过程。其中阶段 2 中第 1 层和第 3 层的初始卷积核分别采用阶段 1 中 2 个训练好的卷积核。

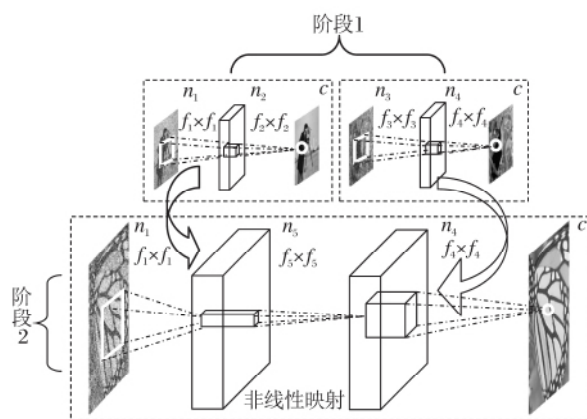


图 1 基于复合卷积神经网络的图像去噪算法流程图

Fig. 1 Framework of image denoising based on composite convolutional neural networks

### 1.1 阶段 1 的网络结构

阶段 1 包含 2 个卷积神经网络,每个卷积神经网络分为 2 个步骤。步骤 1 由卷积核对输入的含噪声图像进行卷积操作,得到特征图谱,步骤 2 针对得到的特征图谱再次利用卷积操作重建清晰的样本图像。

给定 1 组带有噪声的训练图像  $y_1^1, y_2^1, \dots, y_M^1$  和相应的去噪图片  $x_1^1, x_2^1, \dots, x_M^1$ ,  $M$  为训练图像的个数。首先对噪声训练样本进行图像块的提取操作,即对每幅图像  $y_i^1, i = 1, 2, \dots, M$  以步长为  $K^1$  选取对应像素周围的  $k_1 \times k_2$  邻域,令  $Y^1$  表示由  $M$  幅训练图像  $y_1^1, y_2^1, \dots, y_M^1$  组成的特征图谱。

...  $y_M^1$  提取的所有  $k_1 \times k_2$  图像块构成的集合. 对应地, 对每幅目标图像  $x_i^1$  ( $i = 1, 2, \dots, M$ ), 也进行上述相应操作, 可以得到由  $M$  幅目标图像  $x_1^1, x_2^1, \dots, x_M^1$  提供的所有  $k_1 \times k_2$  图像块构成的集合  $X^1$ .

**注1** 在图像块的选取过程中, 若步长  $K^1$  大于或等于邻域尺寸  $k_1$  (一般情况下  $k_1 = k_2$ ) 时, 非叠加取图像块, 否则叠加取块. 每幅图像选取的具体图像块数目与图像本身尺寸大小、步长、邻域尺寸有关. 一般地, 图像块尺寸大小要根据卷积核的大小做出相应改变.

**注2** 每幅训练图像的大小尺寸可以不一致, 即算法可以同时应用于不同尺寸的图像去噪. 另外, 整个训练集和对应的目标集都是经过成对打乱的. 这样的好处是可以使网络在学习过程中不至于产生样本偏向, 即避免在某幅图像或是某些数据集上效果较佳而在另外一些图像或数据集上效果欠佳的情况.

设阶段中第1个卷积神经网络的卷积核数目为  $n_1$ , 尺寸大小为  $f_1 \times f_1$ , 每个训练图像块  $S$  经过卷积操作后, 尺寸大小变为  $(k_1 - f_1 + 1) \times (k_2 - f_1 + 1)$ . 具体操作如下:

$$F_1^1(S) = \max(0, W_1^1 * S + B_1^1). \quad (1)$$

其中:  $*$  表示卷积操作;  $W_1^1$  对应于  $n_1$  个  $c \times f_1 \times f_1$  的卷积核;  $c$  表示输入训练图像块的颜色通道, 如彩色RGB图像中  $c = 3$ , 灰度图像中  $c = 1$ . 经过卷积核  $W_1^1$  对输入图像块进行卷积操作后, 得到  $n_1$  幅特征图谱.  $B_1^1$  对应于1个  $n_1$  维向量, 向量中的第  $i$  个分量加在相应的第  $i$  张特征图谱的每个元素上. 然后经过激活函数  $\max(0, x)^{[15]}$  的作用, 得到第1步的卷积输出  $F_1^1(S)$ .

第2步针对得到的卷积输出  $F_1^1(S)$ , 进行图像重建工作. 再次利用卷积操作把  $n_1$  幅特征图谱重新合并为1幅清晰的输出图像. 最终阶段1中第1个卷积神经网络的输出如下:

$$F^1(S) = W_2^1 * F_1^1(S) + B_2^1. \quad (2)$$

式(2)与式(1)在结构上类似, 但是式(2)中的激活函数取线性函数  $I(t) = t$ . 类似地,  $W_2^1$  对应于  $c$  个  $n_1 \times f_2 \times f_2$  的卷积模板,  $B_2^1$  对应于1个  $c$  维的向量, 得到最终的输出  $F^1(S)$ .

类似地, 阶段1中第2个卷积神经网络和第1个卷积神经网络结构相同, 但卷积核数目和卷积核尺寸大小要根据后面阶段2的网络结构做出相应变化. 调整方式如下: 阶段1中第1个卷积神经网络的第1层卷积核参数和阶段2中第1层的卷积核参数相等, 均

为  $n_1 \times f_1 \times f_1$ . 阶段1中第1个卷积神经网络的第2层卷积核参数  $n_2 \times f_2 \times f_2$ . 则根据具体的训练样本和实验效果设定, 在去噪效果保持不变的情况下逐渐减少卷积核的个数, 直到去噪效果下降为止. 类似地, 阶段1中第2个卷积神经网络的第2层卷积核参数和阶段2中第3层的卷积核参数相等, 均为  $n_4 \times f_4 \times f_4$ . 阶段1中第2个卷积神经网络的第1层参数  $n_3 \times f_3 \times f_3$  也是根据具体训练样本和实验效果设定.

当然, 这里也可以选择阶段1中第2个卷积网络的第1层卷积核  $n_3 \times f_3 \times f_3$  和阶段1中第1个卷积网络的第2层卷积核  $n_2 \times f_2 \times f_2$  作为阶段2中卷积神经网络的第1层和第3层的卷积核的初始值, 并且仍然保留网络的鲁棒性. 因为阶段1中2个卷积网络选取的训练样本不同, 所以以阶段1中学习得到的卷积核作为阶段2中网络的部分初始权值, 有助于提升阶段2中网络的鲁棒性.

## 1.2 阶段1中卷积神经网络的训练

在1.1节中构造阶段1中的2个卷积神经网络. 现在基于随机梯度下降法<sup>[16]</sup>训练2个网络中的权值(卷积核)和偏置向量. 卷积神经网络学习属于监督学习, 对于给定的1组训练样本  $\{(Y_i, X_i)\}_{i=1}^n$ , 选用实际输出与目标输出的平方差作为损失函数, 具体损失函数公式如下:

$$L(\theta) = \frac{1}{n} \sum_{i=1}^n \|F(Y_i; \theta) - X_i\|^2, \quad (3)$$

其中  $\theta$  表示所有权值与偏置值的集合,  $F(Y_i; \theta)$  表示网络对应于图像块  $Y_i$  的实际输出. 在实际操作中采用分批处理方式, 所以并不是  $n$  幅图像块1次训练, 而是1次采用128幅或256幅进行训练.

接着利用随机梯度下降法优化式(3)中的损失函数, 具体的参数迭代公式如下:

$$\Delta_{i+1} = 0.9\Delta_i - \eta \frac{\partial L}{\partial W_i^l},$$

$$W_{i+1}^l = W_i^l + \Delta_{i+1},$$

其中  $l$  为卷积神经网络的层数,  $i$  为权值的迭代次数,

$\eta$  为学习速率,  $\frac{\partial L}{\partial W_i^l}$  为损失函数关于权值的偏导数.

对于阶段1中的2个卷积神经网络, 选择不同的数据库作为训练样本. 2个卷积网络经过随机梯度下降法训练后, 学习得到的权值和偏置作为阶段2中的部分初始权值和偏置, 从而辅助阶段2中网络的学习, 在一定程度上加快阶段2中网络的收敛速度.

## 1.3 阶段2的网络结构

阶段2中的卷积神经网络是整个图像去噪算法

的核心,该阶段的网络利用阶段 1 中训练得到的卷积核作为部分初始卷积核,进一步学习和训练网络,达到较好效果.能这样处理的原因是虽然 2 个阶段的训练样本不同,但是样本所加的噪声和训练目标一致,即 2 个阶段中的网络都是模拟噪声图像到清晰图像之间的映射关系,因此阶段 1 中训练的卷积核具有一定的代表性,在上述映射关系的网络结构中具有一定的通用性.这样,阶段 2 中的卷积神经网络在训练时间上快于单一的卷积神经网络.

阶段 2 中的卷积神经网络包含 3 步.1) 利用阶段 1 中第 1 个卷积神经网络学习得到的第 1 层卷积核对第 2 阶段的输入图像块进行卷积操作,得到 1 组含噪声的特征图谱.2) 利用非线性映射拟合含噪声特征图谱和清晰特征图谱之间的对应关系.这种对应关系由整个网络经过样本训练得到.3) 通过阶段 1 中第 2 个卷积神经网络学习得到的第 2 层卷积核合并清晰的特征图谱为 1 幅清晰的样本图.

给定 1 组带有噪声的训练图片  $y_1^2, y_2^2, \dots, y_N^2$  和相应的目标图片  $x_1^2, x_2^2, \dots, x_N^2$ , 分别对它们进行大小为  $k_3 \times k_4$ 、步长为  $K^2$  的图像块提取操作,得到的图像块集合分别记为  $Y^2$  和  $X^2$ .

由阶段 1 训练得到的卷积核  $\widehat{W}_1^1$  和偏置  $\widehat{B}_1^1$  对训练的图像块  $S$  进行卷积操作,得到  $n_1$  幅特征图谱  $F_1^2(S)$ :

$$F_1^2(S) = \max(0, \widehat{W}_1^1 * S + \widehat{B}_1^1).$$

再利用非线性映射拟合含噪声特征图谱和清晰特征图谱之间的对应关系.这里应用  $n_5$  个大小为  $1 \times 1$  的卷积核  $Y$  以实现.第 2 层的输出如下:

$$F_2^2(S) = \max(0, W_2^2 * F_1^2(S) + B_2^2),$$

其中,  $W_2^2$  为  $n_5 \times 1 \times 1$  的权值,  $B_2^2$  为  $n_5$  维的偏置. Dong 等<sup>[17]</sup>证实利用类似结构的卷积神经网络与稀疏表示法在一定程度上存在关联,但是卷积神经网络能取得更好效果.

由非线性映射得到清晰图谱  $F_2^2(S)$  后,利用阶段 1 学习得到的卷积核重建图像:

$$F_3^2(S) = \widehat{W}_4^1 * F_2^2(S) + \widehat{B}_4^1,$$

其中,  $\widehat{W}_4^1$  为阶段 1 中第 2 个卷积神经网络学习得到的  $c$  个大小为  $n_4 \times f_4 \times f_4$  的重建卷积核,  $\widehat{B}_4^1$  为  $c$  维向量.经卷积重建得到的  $F_3^2(S)$  是最终的去噪图像块.

对于阶段 2 中的卷积神经网络,只需要训练第 2 层中的权值和偏置,这里同样采用 1.2 节中的随机梯度下降方法进行训练.

## 2 实验及结果分析

本节对比基于复合卷积神经网络的图像去噪算法与目前常用的图像去噪算法,对比算法包括基于稀疏表示理论的 K-SVD<sup>[1]</sup> 和非局部集中稀疏表示 (Nonlocally Centralized Sparse Representation, NCSR)<sup>[3]</sup>、基于非局部相似性的 NLM<sup>[4]</sup>、基于机器学习理论的 MLP<sup>[8]</sup> 与期望的补丁对数似然估计 (Expected Patch Log Likelihood, EPLL)<sup>[18]</sup>.

本文的图像去噪实验都是在 Matlab 2014b 下操作进行,处理器为 Intel CORE i3 4 GB 内存.训练卷积神经网络的数据库为 Berkeley Segmentation Dataset (BSD)<sup>[19]</sup>.训练阶段 1 中第 1 个卷积网络使用的样本为 BSD500 中 Train 文件夹下的 100 幅随机图像(共有 200 幅),训练阶段 1 中第 2 个卷积网络使用的数据为 BSD500 中 Train 文件夹下的另外 100 幅图像.训练阶段 2 中的卷积网络使用的数据为 BSD500 中 Val 文件夹下的 100 幅图像.同时选用图像去噪领域常用的 12 张标准图像作为测试图像.使用峰值信噪比 (Peak Signal to Noise Ratio, PSNR) 和结构相似度 (Structural Similarity, SSIM) 作为去噪评价指标.

### 2.1 网络参数设置

在整个卷积神经网络中,需要设定的参数主要是阶段 1 和阶段 2 中卷积核的个数和图像块尺寸大小.针对具体实验,设置各个参数如下.在阶段 1 中第 1 个卷积神经网络中,取步长  $K^1 = 12$  图像块尺寸大小

$$k_1 \times k_2 = 33 \times 33,$$

第 1 层卷积核个数及尺寸大小

$$n_1 \times f_1 \times f_1 = 64 \times 9 \times 9,$$

第 2 层卷积核个数及尺寸大小

$$n_2 \times f_2 \times f_2 = 16 \times 5 \times 5.$$

阶段 1 中第 2 个卷积神经网络中的第 1 层卷积核数目及尺寸大小

$$n_3 \times f_3 \times f_3 = 48 \times 7 \times 7,$$

第 2 层卷积核及尺寸大小

$$n_4 \times f_4 \times f_4 = 32 \times 5 \times 5.$$

阶段 2 中第 1 层卷积核与阶段 1 中第 1 个神经网络的第 1 层卷积核相同,即

$$n_1 \times f_1 \times f_1 = 64 \times 9 \times 9,$$

第 2 层为卷积核形成的非线性映射,其中

$$n_5 \times f_5 \times f_5 = 32 \times 1 \times 1,$$

第 3 层为重建卷积核,与阶段 1 中第 2 个卷积神经网络的第 2 层的卷积核结构相同,即

$$n_4 \times f_4 \times f_4 = 32 \times 5 \times 5.$$

2.2 高斯白噪声去噪实验

在图像去噪实验中,12 幅标准测试图像添加标

准差  $\sigma = 25$  的高斯白噪声,然后运用各种图像去噪算法进行去噪,实验结果如表 1 所示.表中黑体数字为该测试图像在所列算法中取得最好的测试结果.

表 1 6 种算法的实验结果对比  
Table 1 Experimental result comparison of 6 algorithms

图像	K-SVD		NCSR		NLM		MLP		EPLL		本文算法	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Barbara	29.61	0.8498	<b>30.63</b>	<b>0.8856</b>	28.09	0.7803	29.54	0.8640	28.61	0.8464	26.32	0.8705
Boat	29.30	0.7707	29.69	0.7919	27.93	0.7131	<b>29.93</b>	0.8011	29.69	0.7944	29.08	<b>0.8865</b>
Cameraman	28.98	0.8357	29.38	0.8501	27.75	0.7543	<b>29.54</b>	<b>0.8590</b>	29.22	0.8523	27.87	0.8220
Couple	28.90	0.7769	29.45	0.8058	27.47	0.7153	29.70	0.8151	<b>29.50</b>	0.8069	28.50	<b>0.8777</b>
Fingerprint	27.28	0.8944	<b>27.85</b>	0.9103	26.04	0.8680	27.67	0.9109	27.20	0.9029	26.82	<b>0.9541</b>
Hill	29.14	0.7369	29.64	0.7700	28.06	0.6907	<b>29.76</b>	0.7734	29.54	0.7651	29.08	<b>0.8689</b>
House	32.17	0.8475	<b>33.01</b>	<b>0.8620</b>	30.09	0.7563	32.47	0.8586	32.19	0.8547	31.54	0.8260
Lena	31.37	0.8442	31.96	0.8621	29.67	0.7498	<b>32.34</b>	0.8676	31.77	0.8557	31.51	<b>0.8945</b>
Man	29.11	0.7785	29.54	0.8004	28.08	0.7162	<b>29.82</b>	0.8113	29.61	0.8041	29.17	<b>0.8791</b>
Monarch	29.06	0.8897	29.65	0.9033	27.42	0.8087	<b>29.74</b>	<b>0.9040</b>	29.56	0.9024	28.80	0.8762
Peppers	29.82	0.8600	30.01	0.8708	28.03	0.7761	30.28	0.8745	30.24	0.8750	<b>31.85</b>	<b>0.9057</b>
Straw	25.78	0.8508	<b>26.26</b>	0.8723	24.62	0.8149	26.24	<b>0.8742</b>	25.83	0.8609	24.64	0.7806

由表 1 可以看出,本文算法在 PSNR 和 SSIM 上可以与上述算法相当.这是因为卷积神经网络是以图像块为单位进行卷积操作,相关的优化求解也是以此为基础进行处理.相比许多以恢复图像精确像素值为基础的算法不同,复合卷积神经网络更偏向于 1 个图像块的像素值及整体像素误差.

图 2 给出表 1 中 6 种图像去噪算法对测试图像 Boat 和 Man 的去噪效果,并在特定部位进行放大处理以便观察.从图 2 可以看出,本文算法效果完全可以和当前一些通用的去噪算法相当.

为了进一步体现本文算法对图像去噪的效果,又随机选取 BSD500 中 Test 文件夹下的 100 幅图像作为测试图像进行实验,并根据不同噪声参数和评价指标刻画算法效果.由于 MLP 和 NLM 在  $\sigma = 50$  或  $\sigma = 100$  时无对应模型,使用  $\sigma = 25$  时的模型效果较差,所以换成文献 [9] 中的第一个卷积神经网络(The First Convolutional Network, CN1) 代替.



(b) 噪声图  
(b) Noise images



(c) NLM



(a) 原图  
(a) Original images



(d) K-SVD





(e) EPLL



(f) NCSR



(g) MLP



(h) 本文算法

(h) The proposed algorithm

图 2 6 种去噪算法在 2 幅测试图像上的视觉对比图

Fig.2 Visual images comparison of 6 denoising algorithms on 2 test images

5 种算法的去噪效果如图 3 所示。

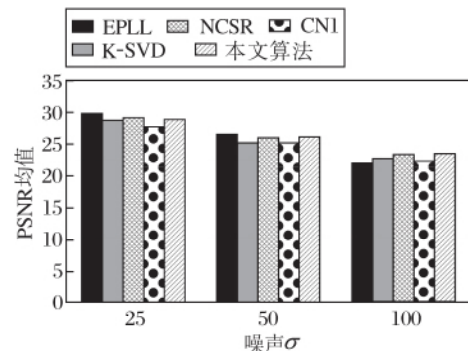
从图 3(a) (条形图越高越好) 可以看出, 当噪声级别  $\sigma = 25$  时, 本文算法和其它算法效果相当, 处于

中间水平; 当噪声程度加剧到  $\sigma = 50$  时, 本文算法已经能达到在对比算法中并列第一的效果(条形图高度并列高); 而当噪声程度加剧到  $\sigma = 100$  时, 本文算法已经超过所有的对比算法(条形图高度最高)。

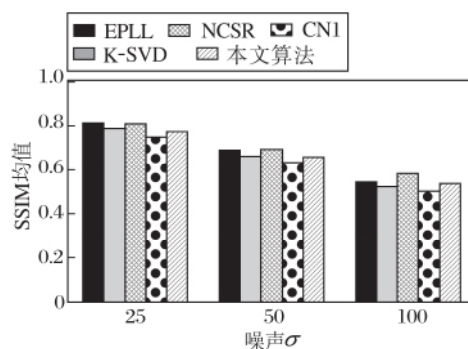
从图 3(b) (条形图越高越好) 可以看出, 在各个级别的噪声上, 本文算法均处于一般程度, 即能和其它算法相当, 却无法超越。

从图 3(c) (条形图越低越好) 可以看出, 当噪声级别  $\sigma = 25$  时, 本文算法指标一般(条形图高度处于中间); 当  $\sigma$  上升到 50, 本文算法的指标降低为同类算法的最优水平(条形图高度处于并列最低); 当  $\sigma$  上升到最剧烈的 100 时, 本文算法的均方根误差 (Root Mean Square Error, RMSE) 均值效果达到最好(条形图降至同类方法的最低值)。

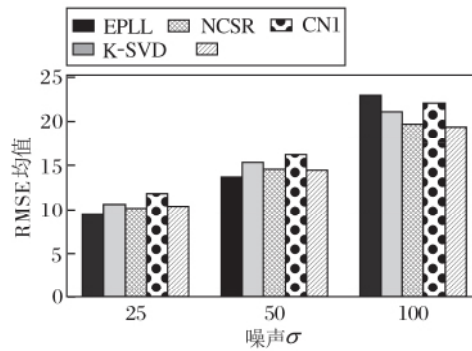
无论是在 PSNR、SSIM, 还是在 RMSE 指标上, 本文算法都可以和当前较好的算法相当。因为本文算法是以 2 幅图像的像素值误差为目标函数进行优化, 故在 PSNR 和 RMSE 上效果较好。另外, 当噪声程度加剧时, 由于本文算法运用 2 个阶段学习去噪关系, 因此效果更优。



(a) PSNR



(b) SSIM



(c) RMSE

图3 100幅测试图片去噪效果对比

Fig.3 Comparison of denoising results on 100 test images

### 2.3 算法运行时间分析

由于本文构建的网络是全卷积结构,为了更好地分析算法的时间优势,根据 He 等<sup>[20]</sup>的卷积网络计算复杂度:

$$O\left(\sum_{l=1}^d n_{l-1} s_l^2 n_l m_l^2\right),$$

其中  $d$  为卷积层的层数,  $n_l$  为第  $l$  层卷积核的数目,  $s_l^2$  为卷积核的空间尺寸大小,  $m_l^2$  为输出特征图谱的空间尺寸大小。

一般卷积网络的算法时间主要由上述参数决定,但阶段2中网络的部分卷积核采用第1阶段中训练好的卷积核,故  $l=1,3$  时算法的实际计算复杂度会大幅下降。另外,阶段1中的2个网络并行,所以在阶段1中只需花费1个网络的训练时间。故本文网络在训练时间上短于相应的一般卷积网络。

为了说明本文算法的训练时间短于普通的卷积神经网络(直接训练的3层卷积网络),现从算法效率和时间上给出解释。由于本文训练实验部分是在 C++ 环境下的 caffe<sup>[21]</sup> 框架下运行,而现在大多数算法都是在 Matlab 环境下训练,所以不具有可比性。为了更准确地说明实现效果,把超分辨率卷积神经网络(Super-Resolution Convolutional Neural Network, SRCNN)<sup>[17]</sup>的3层卷积网络应用到去噪上(只改变输入图像,其它均不变),然后与 SRCNN 进行对比。实验结果如图4所示。

为了更好地观察效果,图中右上方对原图进行一定程度的局部放大。从图中可以明显看出,本文算法收敛较快,且收敛后较稳定,震荡幅度较小。以0.7为阈值线(水平虚线)对比算法时间,本文算法在

300次迭代后就能较稳定地在其上下波动,500次迭代后首次低于阈值。而文献[17]算法在1200次到1600次的迭代中依然波动较大,1800次迭代后才能逐渐稳定,到1900次迭代后才首次低于阈值。本文算法在阶段1的训练时间为580s,阶段2500次迭代的训练时间为455s,总时间为1035s。文献[17]算法1900次迭代的训练时间为1784s。因此,本文算法比直接构造的卷积神经网络在时间上加快72%。

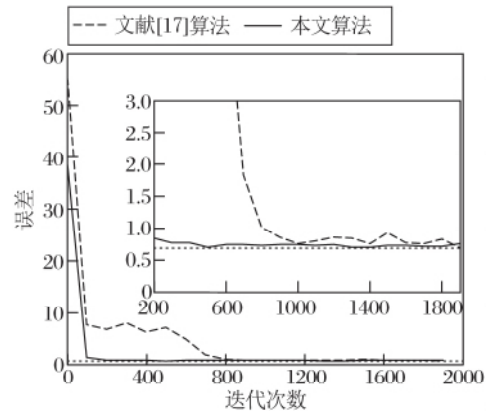


图4 2种算法的误差对比

Fig.4 Error comparison of 2 methods

### 2.4 参数调整及分析

阶段1中2个卷积神经网络的参数设置和调整见1.1节。在阶段2中卷积神经网络的训练中,刚开始以图像块大小为

$$k_1 \times k_2 = 11 \times 11,$$

卷积核模板为

$$n_1 \times f_1 \times f_1 = 256 \times 9 \times 9,$$

非线性映射层为

$$n_5 \times f_5 \times f_5 = 128 \times 1 \times 1$$

的3层卷积网络进行实验,但是效果不理想。然后调整图像块的大小,发现图像块大小为

$$k_1 \times k_2 = 33 \times 33$$

时较符合样本特征。但即使调整图像块大小使其较优,损失误差还是过大。

通过在非线性映射层之后又新增1层卷积层

$$n_4 \times f_4 \times f_4 = 64 \times 7 \times 7$$

以调整实验效果,发现效果可以接受,但是训练时间过长。为了解决此问题,在保持效果不变的情况下通过减少卷积核数目和模板大小加快收敛速度,发现卷积核尺寸

$$n_1 \times f_1 \times f_1 = 64 \times 9 \times 9,$$

$$n_5 \times f_5 \times f_5 = 32 \times 1 \times 1,$$

$$n_4 \times f_4 \times f_4 = 32 \times 5 \times 5$$

时效果较好,如果继续缩减会导致效果下降.因此最后经过参数调整后,采用这个配置.

### 2.5 椒盐噪声去噪实验

在图像去噪领域中,绝大多数算法都是基于高斯噪声去噪.一旦噪声类型改变,原有算法模型失效.本文算法具有较强的泛化性,当噪声类型变为椒盐噪声时,依旧可以达到较好的效果.椒盐噪声是指图像(像素值范围为 0~255)中的像素值有 50% 的概率突变为 0 或 255,这些突变的像素值个数占总像素的一定比例( $\rho$ ),这里取  $\rho=20\%$ .

表 2 为本文算法在 12 幅标准测试图像上的实验效果.图 5 为本文算法在 Lena 图像的视觉去噪效果.当  $\rho=20\%$  时,根据文献[8],MLP 在 Lena 图像上的 PSNR 值为 34.50,效果差于本文算法.

表 2 本文算法在椒盐噪声上的实验结果

Table 2 Experimental results of the proposed method on the image with salt & pepper noises

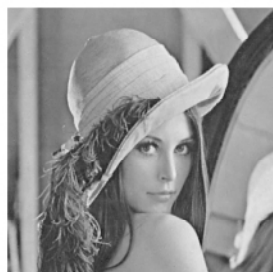
图像	PSNR	SSIM
Barbara	28.24	0.9375
Boat	31.58	0.9638
Cameraman	27.38	0.8707
Couple	31.43	0.9657
Fingerprint	30.43	0.9876
Hill	32.88	0.9655
House	33.63	0.8964
Lena	34.54	0.9655
Man	32.53	0.9653
Monarch	29.36	0.9300
Peppers	31.43	0.9145
Straw	26.73	0.8899

PSNR=12.40 SSIM=0.2731



(a) 噪声图像  
(a) The noise image

PSNR=34.54 SSIM=0.9655



(b) 复原图像  
(b) The recovered image

图 5 本文算法在椒盐噪声上的视觉图

Fig.5 Visual image of the proposed method on the image with salt & pepper noises

## 3 结束语

本文基于深度学习理论,提出基于复合卷积神经网络的图像去噪算法.利用 3 层卷积神经网络拟合带噪声图像和去噪图像之间的映射关系.由于部分初始卷积核经过阶段 1 预先训练学习得到,因此相比其它卷积神经网络,网络结构相对简洁.实验表明,本文算法在峰值信噪比、结构相似度及均方根误差指标上都可以和当前较好的算法相当,并在算法时间上占优.在下一步工作中,将会根据网络结构特性,在加深网络深度时优化网络的训练速度和效果.

## 参 考 文 献

- [1] AHARON M, ELAD M, BRUCKSTEIN A. The K-SVD: An Algorithm for Designing Overcomplete Dictionaries for Sparse Representation. *IEEE Transactions on Signal Processing*, 2006, 54(11): 4311–4322.
- [2] MAIRAL J, BACH F, PONCE J, *et al.* Non-local Sparse Models for Image Restoration // *Proc of the 12th IEEE International Conference on Computer Vision*. Washington, USA: IEEE, 2009: 2272–2279.
- [3] DONG W S, ZHANG L, SHI G M, *et al.* Nonlocally Centralized Sparse Representation for Image Restoration. *IEEE Transactions on Image Processing*, 2012, 22(4): 1620–1630.
- [4] BUADES A, COLL B, MOREL J M. A Non-local Algorithm for Image Denoising // *Proc of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. Washington, USA: IEEE, 2005, II: 60–65.
- [5] DAI J J, AU O C, FANG L, *et al.* Multichannel Nonlocal Means Fusion for Color Image Denoising. *IEEE Transactions on Circuits and Systems for Video Technology*, 2013, 23(11): 1873–1886.
- [6] DABOV K, FOI A, KATKOVNIK V, *et al.* Image Denoising by Sparse 3-D Transform-Domain Collaborative Filtering. *IEEE Transactions on Image Processing*, 2007, 16(8): 2080–2095.
- [7] JI H, LIU C Q, SHEN Z W, *et al.* Robust Video Denoising Using Low Rank Matrix Completion // *Proc of the IEEE Conference on Computer Vision and Pattern Recognition*. Washington, USA: IEEE, 2010: 1791–1798.
- [8] BURGER H C, SCHULER C J, HARMEILING S. Image Denoising: Can Plain Neural Networks Compete with BM3D? // *Proc of the IEEE Conference on Computer Vision and Pattern Recognition*. Washington, USA: IEEE, 2012: 2392–2399.
- [9] JAIN V, SEUNG H S. Natural Image Denoising with Convolutional Networks // *Proc of the 22nd International Conference on Neural Information Processing Systems*. New York, USA: Curran Associates Inc., 2008: 769–776.
- [10] XIE J Y, XU L L, CHEN E H. Image Denoising and Inpainting with Deep Neural Networks // *Proc of the 26th Annual Conference on Neural Information Processing Systems*. Berlin, Germany: Springer, 2012: 350–358.



- [11] 牛和明, 杜茜, 张建勋. 一种自适应全变分图像去噪算法. 模式识别与人工智能, 2011, 24(6): 798-803.  
(NIU H M, DU Q, ZHANG J X. An Algorithm of Adaptive Total Variation Image Denoising. Pattern Recognition and Artificial Intelligence, 2011, 24(6): 798-803.)
- [12] 纪建, 许双星, 李晓. 基于形态成分分析和 Contourlet 变换的自适应阈值图像去噪方法. 模式识别与人工智能, 2014, 27(6): 561-568.  
(JI J, XU S X, LI X. An Adaptive Thresholding Image Denoising Method Based on Morphological Component Analysis and Contourlet Transform. Pattern Recognition and Artificial Intelligence, 2014, 27(6): 561-568.)
- [13] 金良海, 李德华. 基于噪声检测的图像去噪算法. 模式识别与人工智能, 2008, 21(3): 298-302.  
(JIN L H, LI D H. An Image Denoising Algorithm Based on Noise Detection. Pattern Recognition and Artificial Intelligence, 2008, 21(3): 298-302.)
- [14] 王相海, 刘倩倩, 张爱迪, 等. 曲线拟合确定阈值的非抽取小波贝叶斯图像去噪方法. 模式识别与人工智能, 2016, 26(4): 322-331.  
(WANG X H, LIU X Q, ZHANG A D, et al. Undecimated Wavelet Bayesian Image Denoising Method with Its Threshold Determined by Curve Fitting. Pattern Recognition and Artificial Intelligence, 2016, 26(4): 322-331.)
- [15] NAIR V, HINTON G E. Rectified Linear Units Improve Restricted Boltzmann Machines // Proc of the 27th International Conference on Machine Learning. Madison, USA: Omnipress, 2010: 807-814.
- [16] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. Imagenet Classification with Deep Convolutional Neural Networks // PEREIRA F, BURGESS C J C, BOTOU L, et al., eds. Advances in Neural Information Processing Systems 25. Cambridge, USA: The MIT Press, 2012: 1097-1105.
- [17] DONG C, LOY C C, HE K M, et al. Image Super-Resolution Using Deep Convolutional Networks. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2016, 38(2): 295-307.
- [18] ZORAN D, WEISS Y. From Learning Models of Natural Image Patches to Whole Image Restoration // Proc of the IEEE International Conference on Computer Vision. Washington, USA: IEEE, 2011: 479-486.
- [19] MARTIN D, FOWLKES C, TAL D, et al. A Database of Human Segmented Natural Images and Its Application to Evaluating Segmentation Algorithms and Measuring Ecological Statistics. Report, UCB/CSD-1-1133. Berkeley, USA: University of California, Berkeley, 2001.
- [20] HE K M, SUN J. Convolutional Neural Networks at Constrained Time Cost // Proc of the IEEE Conference on Computer Vision and Pattern Recognition. Washington, USA: IEEE, 2015: 5353-5360.
- [21] JIA Y Q, SHELHAMER E, DONAHUE J, et al. Caffe: Convolutional Architecture for Fast Feature Embedding // Proc of the 22nd ACM International Conference on Multimedia. New York, USA: ACM, 2014: 675-678.

## 作者简介



吕永标,男,1992年生,硕士研究生,主要研究方向为机器学习、图像处理等.E-mail: mtlyb@gmail.com.

(LÜ Yongbiao, born in 1992, master student. His research interests include machine learning and image processing.)



赵建伟,女,1977年生,博士,教授,主要研究方向为机器学习、图像处理等.E-mail: zhaojw@amss.ac.cn.

(ZHAO Jianwei, born in 1977, Ph.D., professor. Her research interests include machine learning and image processing.)



曹飞龙(通讯作者),男,1965年生,博士,教授,主要研究方向为智能计算、图像处理等.E-mail: feilongcao@gmail.com.

(CAO Feilong( Corresponding author), born in 1965, Ph.D., professor. His research interests include intelligence computing and image processing.)