

基于改进栈式稀疏去噪自编码器的自适应图像去噪

马红强^{1*}, 马时平^{1**}, 许悦雷¹, 吕超², 朱明明¹

¹空军工程大学航空工程学院, 陕西 西安 710038;

²中国人民解放军 95876 部队, 甘肃 山丹 734100

摘要 针对栈式稀疏去噪自编码器(SSDA)在图像去噪上训练难度大、收敛速度慢和普适性差等问题,提出了一种基于栈式修正降噪自编码器的自适应图像去噪模型。采用线性修正单元作为网络激活函数,以缓解梯度弥散现象;借助残差学习和批归一化进行联合训练,加快收敛速度;而为克服新模型对噪声普适性差等问题,需要对其进行多通道并行训练,充分利用网络挖掘出的潜在数据特征集计算出最优通道权重,并通过训练权重权重预测模型预测出各通道最优权重,从而实现自适应图像去噪。实验结果表明:与目前降噪较好的 BM3D 和 SSDA 方法相比,所提方法不仅在收敛效果上优于 SSDA 方法,而且能够自适应处理未参与训练的噪声,使其具有更好的普适性。

关键词 图像处理; 图像降噪; 批归一化; 残差学习; 自适应性

中图分类号 TP391

文献标识码 A

doi: 10.3788/AOS201838.1010001

Adaptive Image Denoising Based on Improved Stacked Sparse Denoising Auto-Encoder

Ma Hongqiang^{1*}, Ma Shiping^{1**}, Xu Yuelei¹, Lü Chao², Zhu Mingming¹

¹Aeronautics Engineering College, Air Force Engineering University, Xi'an, Shaanxi 710038, China;

²Unit 95876 of PLA, Shandan, Gansu 734100, China

Abstract Aiming at the problems that the stacked sparse denoising auto-encoder(SSDA) is difficult to train on image denoising, such as slow convergence rate and poor universality, an adaptive image denoising model based on stacked rectified denoising auto-encoder is proposed. The rectified linear units is used as a network activation function to alleviate the phenomenon of gradient dispersion. Joint training with the residual learning and batch normalization to accelerate convergence speed. In order to solve the problem of noise poor universality of the new model, it is necessary to carry out the multi-channel parallel training, and make full use of the potential data feature extracted by the network to find the optimal channel weights, and learn to predict optimal column weights via training weight prediction model for realizing adaptive image denoising. The experimental results show that the proposed algorithm is not only better than the SSDA in the convergence effect, but also adaptively processing the non-participating training noise, and has better universality, compared with the current methods of BM3D and SSDA.

Key words image processing; image denoising; batch normalization; residual learning; adaptability

OCIS codes 100.3010; 100.3020; 100.2000; 100.4996

1 引 言

图像去噪一直是计算机视觉领域研究的热点问题,其处理结果的好坏直接影响到后续识别、分割、分析等环节的性能。文献[1]提出了一种基于视觉模型的自然环境下作物识别与导航线提取方法;文献[2]对 RGB-D 图像特征提取和分类方法进行了总结分析。但这些任务大多是针对于正常无噪或含

噪较少的图像开展的,而图像在采集和传递过程中,往往会受到不同程度的噪声污染,导致图像质量严重下降^[3]。因此,研究图像去噪算法和优化算法性能对提高图像质量具有非常重要的指导意义和应用价值。

目前,图像去噪算法主要有:1) 基于小波变换或字典学习的方法。小波变换是把含噪图像变换至小波域,对小波系数进行多层分解,实现信噪分离,

收稿日期: 2018-03-06; 修回日期: 2018-04-14; 录用日期: 2018-05-08

基金项目: 国家自然科学基金(61372167,61379104)

* E-mail: 18049025189@163.com; ** E-mail: 1402543131@qq.com

如基于贝叶斯最小平方-高斯概率混合模型(BLS-GSM)^[4];基于字典的方法是将含噪图像在过完备原子库上进行稀疏表示,从而去除图像中的无用信息,达到去噪目的^[5],如K奇异值分解(KSVD)。2) 基于图像的全局统计特性或其他特性(如自相似)。目前去噪效果较好的三维块匹配(BM3D)算法是通过相似图像块匹配分组生成三维矩阵,再经三维变换域去噪,最后逆变换还原出无噪图像^[6]。徐少平等^[7]结合字典学习和自相似性提出了一种非局部中心化稀疏表示(NCSR)图像去噪方法,虽然该算法对高强度噪声有了较大提升,但其构建用于稀疏表示所需字典时的计算量很大,并且有可能会过度地平滑图像,造成图像失真。3) 基于深度学习的重建方法。近年来,深度学习受到研究人员的广泛关注,并逐渐成为互联网大数据以及人工智能的一个热潮^[8],也为图像去噪提供了新思路。文献^[9]通过训练多层感知器(MLP)学习含噪图像和无噪图像之间的映射关系,实现图像降噪,其去噪效果优于传统大多数去噪算法。文献^[10,11]将栈式稀疏去噪自编码器(SSDA)运用在图像去噪上,取得了良好效果,尤其是当噪声强度较大时,去噪图像相比于KSVD和BM3D,峰值信噪比(PSNR)更高,图像模糊现象也得到了进一步改善。

基于深度学习的图像去噪方法与传统方法相比在主观视觉效果与客观评价指标上均有所改善^[12]。但是深度学习是从海量数据中学习数据的特征表达和拟合非线性复杂函数,因此,网络层数很深、参数量巨大、梯度弥散现象严重、耗时较长且调参十分困难,往往依赖于经验知识。同时,深度学习去噪方法由于训练中只添加某单一噪声,网络对该噪声去除表现优异,但是对其他噪声处理效果一般,泛化能力和稳健性差。如若实现其他噪声有效去除,又得耗费大量时间和人力重新训练。

鉴于此,在SSDA的基础上做了如下改进:1)为防止梯度弥散,采用线性修正单元(ReLU)代替SSDA中sigmoid激活函数,并去掉稀疏约束项;2)为便于调参和加快网络收敛速度,采用批归一化(BN)和残差学习的方法;3)对改进后的网络(SRDA)进行多通道训练,以实现自适应图像去噪,提高模型对不同噪声的稳健性。

2 栈式稀疏去噪自编码器

Vincent 等于 2008 年提出了去噪自编码器

(DAE)。DAE是由输入层、隐藏层和输出层组成的3层神经网络,算法流程如图1所示。原始输入数据 x 添加噪声得到干扰输入 \tilde{x} ,通过编码函数 s_f 得到输入的特征表达 h ,再通过解码函数 s_g 将 h 映射至输出层,得到输入数据的重构 y 。

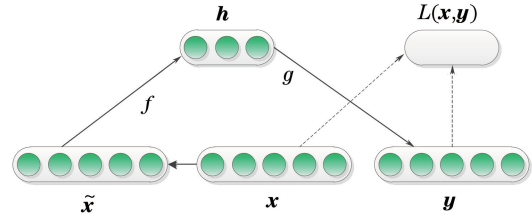


图1 DAE算法流程图

Fig. 1 Algorithm flow of DAE

DAE的编码过程为 $h(\tilde{x}) = s_f(\mathbf{w}\tilde{x} + \mathbf{b})$,解码过程为 $y(\tilde{x}) = s_g(\mathbf{w}'h + \mathbf{b}')$ 。其中, s_f 和 s_g 是非线性激活函数,一般取sigmoid函数,其函数表达式为 $s(z) = [1 + e(-z)]^{-1}$ 。 \mathbf{w} 和 \mathbf{w}' 分别为编码和解码权重, \mathbf{b} 和 \mathbf{b}' 分别为编码和解码偏置。DAE参数 $\theta = \{\mathbf{w}, \mathbf{w}', \mathbf{b}, \mathbf{b}'\}$ 。对于训练样本集 $D = \{\mathbf{x}^{(i)}\}_{i=1}^N$, DAE的损失函数为

$$L_{\text{DAE}}(\theta) = \frac{1}{N} \sum_{i=1}^N \|\mathbf{y}^{(i)} - \mathbf{x}^{(i)}\|_2^2 + \frac{\lambda}{2} (\|\mathbf{w}\|_F^2 + \|\mathbf{w}'\|_F^2), \quad (1)$$

式中 λ 为权重约束项,可以减小权重的幅度,防止过拟合。

后来,学者研究人脑信息处理机制时发现,当图像由视觉进入人脑时,只会刺激少量神经元,大部分神经元仍处于抑制状态^[13]。受此启发,Bengio 通过在误差重构函数中添加一个额外的惩罚因子 KL ,使隐藏层神经元激活度满足一定的稀疏性,即大部分神经元激活值为0。稀疏降噪自编码器(SDA)的损失函数为

$$L_{\text{SDA}}(\theta) = \frac{1}{N} \sum_{i=1}^N \|\mathbf{y}^{(i)} - \mathbf{x}^{(i)}\|_2^2 + \beta \sum_{j=1}^k KL(\hat{\rho}_j \| \rho) + \frac{\lambda}{2} (\|\mathbf{w}\|_F^2 + \|\mathbf{w}'\|_F^2), \quad (2)$$

式中: $KL(\hat{\rho}_j \| \rho) = \rho \log \frac{\rho}{\hat{\rho}_j} + (1 - \rho) \log \frac{1 - \rho}{1 - \hat{\rho}_j}$, 其中 ρ 为稀疏性参数; β 为控制稀疏约束项的权重; $\hat{\rho}_j = \frac{1}{N} \sum_{i=1}^N h_j(\mathbf{x}^{(i)})$ 为隐藏层上第 j 号神经元对应输入数据的平均激活度。

SSDA是将多个SDA堆叠而成的一种深度神经网络,各SDA之间,前一个SDA隐藏层的激活值

作为下一个 SDA 的输入。图 2 所示为 2 个隐藏层的 SSDA 网络结构。

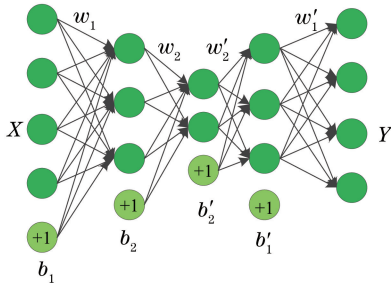


图 2 SSDA 网络结构示意图
Fig. 2 Network diagram of SSDA

目前,SSDA 仍采用 Hinton 提出的贪婪逐层训练法,可以一定程度内缓解梯弥散现象。即先单独训练每一个 SDA,得到网络最优权重后,再将这些权重作为 SSDA 网络权重的初始值,并通过反向传播(BP)算法整体微调,直至得出整个网络的最优参数。SSDA 在微调阶段的损失函数为

$$L_{SSDA}(\theta) = \frac{1}{N} \sum_{i=1}^N \| \mathbf{y}^{(i)} - \mathbf{x}^{(i)} \|_2^2 + \frac{\lambda}{2} \sum_{l=1}^l (\| \mathbf{w}^{(l)} \|_F^2 + \| \mathbf{w}'^{(l)} \|_F^2), \quad (3)$$

式中 l 为 SDA 堆叠的个数,在此阶段无稀疏约束项,这是由于在预训练各 SDA 阶段已经包含在内。

3 改进栈式稀疏去噪自编码器

3.1 栈式修正降噪自编码器

3.1.1 激活函数选择

梯度弥散究其根本原因仍在于残差。SSDA 中采用 sigmoid 函数,由于其导数值 f' 恒小于 1,而残差从输出层向底层反向传播过程中,每经过一层,其残差乘以 f' ,这样残差经过多层后,深度神经网络因底层残差过小而无法更新网络权重。残差表达式为

$$\boldsymbol{\delta}^{(n_l)} = -(\mathbf{y} - \mathbf{a}^{(n_l)}) \cdot f'(\mathbf{z}^{(n_l)}), \quad (4)$$

$$\boldsymbol{\delta}^{(l)} = -(\mathbf{W}^{(l)})^T \boldsymbol{\delta}^{(l+1)} \cdot f'(\mathbf{z}^{(l)}), \quad (5)$$

$$\nabla \mathbf{W}^{(l)} = \boldsymbol{\delta}^{(l+1)} \cdot (\mathbf{a}^{(l)})^T, \quad (6)$$

式中: $\boldsymbol{\delta}^{(n_l)}$ 为输出层残差; $\mathbf{a}^{(n_l)}$ 为输出层实际输出; \mathbf{y} 为输出层的期望输出; $\boldsymbol{\delta}^{(l)}$ 为隐层残差; \mathbf{z} 为激活函数的输入; $\mathbf{a}^{(l)}$ 为隐层输出; $\nabla \mathbf{W}$ 为损失函数对网络权重的梯度。

ReLU 的函数表达式为 $\max(0, x)$, 当输入数据 $x < 0$ 时,输出为 0,满足稀疏性;当 $x > 0$ 时,输出等于输入,且导数 $f' = 1$,由(4)~(6)式可看出残差并不会随着层数增多而出现过小的情况,权重则能够很好地

更新,有效缓解了梯度弥散。在经验规则初始化网络权重后,sigmoid 函数近乎 50% 的神经元被激活,不符合神经科学的研究,而 ReLU 会稀疏掉负值,且能够动态调节,是更接近生物学的激活模型^[14]。

3.1.2 残差学习和批归一化

残差学习起初是为解决图像分类精准度下降问题而提出的,假设网络输入为 μ ,要拟合的函数映射(即输出)为 $H(\mu)$,可以定义另外一个残差映射 $F(\mu)$ 为 $H(\mu) - \mu$,则原始的函数映射 $H(\mu)$ 可以表示为 $F(\mu) + \mu$,在网络模型中是通过短连接实现,该短连接既没有增加额外参数又没有增加额外计算,同时由于残差相比于 $H(\mu)$ 会减少很多特征,则更容易找到最优结果^[15],因此可以加快网络训练速度。

假设 \mathbf{x} 是无噪图像, $\tilde{\mathbf{x}}$ 是其对应的含噪图像, $\tilde{\mathbf{x}}$ 和 \mathbf{x} 的残差 \mathbf{n} 即为图像中的噪声,三者关系可表示为 $\tilde{\mathbf{x}} = \mathbf{x} + \mathbf{n}$ 。传统方法大多数致力于实现 $\tilde{\mathbf{x}}$ 到 \mathbf{x} 的端到端映射,本文提出的网络学习是 $\tilde{\mathbf{x}}$ 到 \mathbf{n} 的映射,即寻找一种映射 f 使得代价函数最小化为

$$\operatorname{argmin}_f \| f(\tilde{\mathbf{x}}) - \mathbf{n} \|_2^2. \quad (7)$$

因此,所提出的网络用残差学习,输入为含噪图像,输出为提取出的噪声,两者相减便可得到去噪后的图像 $\mathbf{x} = \tilde{\mathbf{x}} - f(\tilde{\mathbf{x}})$ 。

神经网络学习过程本质上就是学习数据特征分布,传统方法在训练前只对输入数据做归一化处理,却忽略了中间层数据分布情况。然而输入数据经过非线性运算后其特征分布已发生变化,并且随着深度网络的多层运算之后,数据分布的变化将越来越大。因此,每一层网络又得重新学习新的数据分布,大大降低了网络训练速度。

BN 算法就是为解决上述问题而提出,对每一中间层的输入数据进行批归一化处理,减少网络中间数据分布的改变对神经网络参数训练的影响,从而提高神经网络稳定性^[16]。假设网络某一中间层输入数据集为 $\{\mathbf{x}_1, \dots, \mathbf{x}_n\}$,其算法过程为

$$\boldsymbol{\mu} = \frac{1}{m} \sum_{i=1}^m \mathbf{x}_i, \quad (8)$$

$$\sigma = \frac{1}{m} \sum_{i=1}^m (\mathbf{x}_i - \boldsymbol{\mu})^2, \quad (9)$$

式中 $\boldsymbol{\mu}$ 和 σ 分别为批次均值和方差,再将数据归一化为均值为 0、方差为 1 的数据 $\hat{\mathbf{x}}_i$,

$$\hat{\mathbf{x}}_i = \frac{\mathbf{x}_i - \boldsymbol{\mu}}{\sqrt{\sigma^2 + \epsilon}}, \quad (10)$$

式中 ϵ 是为避免方差 $\sigma = 0$ 时分式不成立而设置的常量。为避免数据归一化破坏特征分布,需要通过

重构变换来恢复原始的特征分布。

$$z_i = \text{BN}_{\alpha, \beta}(\hat{x}_i) = \alpha \hat{x}_i + \beta, \quad (11)$$

式中 α 和 β 为因重构变换而引入的重构参数，需要通过网络训练得到。

文献[17]中提出，随机梯度下降法对于深度网络训练简单高效，但是需要人为地选择参数，而参数的选择对训练结果至关重要，以至于大量时间都浪费在调参上。BN 的引入大大降低了网络训练的难度，不需要刻意调整参数也能取得比较不错的效果。

综上所述，提出了 SRDA 网络，网络模型如图 3 所示。为缓解梯度弥散，采用 ReLU 作为激活函数，并去掉了 SSDA 损失函数中的稀疏约束项，

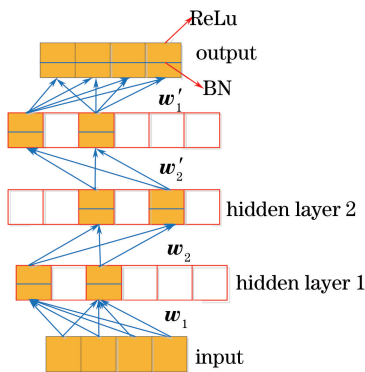


图 3 SRDA 网络结构示意图

Fig. 3 Network diagram of SRDA

采用直接训练的方式；为加快网络收敛速度，利用残差学习和批归一化进行联合训练。

栈式修正降噪自编码器的损失函数为

$$L_{\text{SRDA}}(\theta) = \frac{1}{N} \sum_{i=1}^N \| \mathbf{y}^{(i)} - (\hat{\mathbf{x}}^{(i)} - \mathbf{x}^{(i)}) \|_2^2 + \frac{\lambda}{2} \sum_{l=1}^L (\| \mathbf{w}^{(l)} \|_F^2 + \| \mathbf{w}'^{(l)} \|_F^2). \quad (12)$$

3.2 自适应栈式修正降噪自编码器

基于深度学习的图像去噪方法近年来取得了不错的效果，但是该方法只能去除训练样本特定类型特定强度的噪声，一旦测试样本与训练样本不同，模型降噪能力将大大减弱。因此，对 SRDA 进一步改进，提出了自适应栈式修正降噪自编码器 (ASRDA) 网络模型，能够去除训练中未出现的噪声，算法的普适性和稳健性得到了提高。

ASRDA 是多个通道 SRDA 的线性组合。每个通道 SRDA 采取某单一噪声训练学习出最优特征，然后将这些特征全部输入权重预测模型，得到每个通道 SRDA 的权重，最后将每个 SRDA 的输出线性组合起来，从而得到最终的降噪图像。网络模型如图 4 所示。

ASRDA 网络训练有 3 个阶段，分别是训练多通道 SRDA，寻找理想通道权重和通道权重预测模型训练。多通道 SRDA 训练方法如 2.1 节所述。

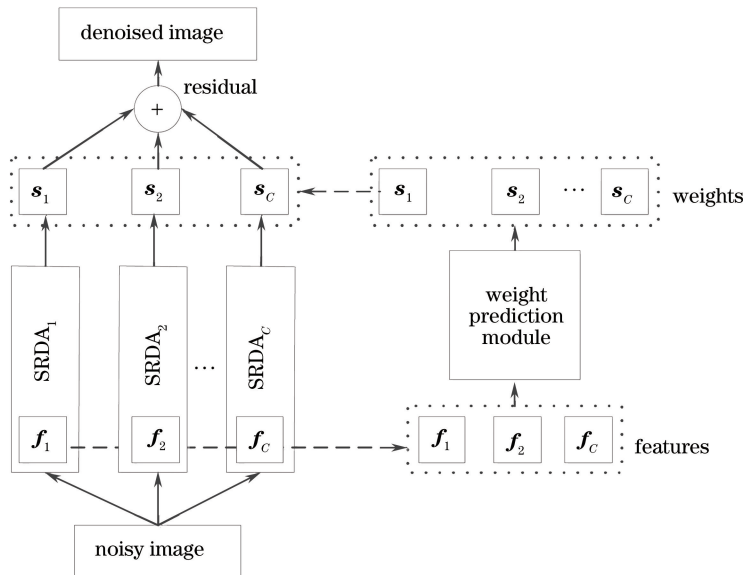


图 4 ASRDA 网络模型

Fig. 4 Network model of ASRDA

3.2.1 寻找最优通道权重

一旦多通道 SRDA 训练完毕，充分利用 SRDA 网络的特征学习能力，将多通道 SRDA 的隐层激活

值构造一个新的特征训练集。具体而言，对 SRDA 通道 c ，隐藏层激活值 $[\mathbf{h}^{(1)}; \dots; \mathbf{h}^{(l)}]$ 组成特征向量 f_c ，然后 f_1, f_2, \dots, f_c 组成整个特征训练

集 $\phi = [f_1; \dots; f_c]$ 。 i 为 SRDA 堆叠的个数, C 为通道的个数。

此外,对于每幅输入图像,每个通道都会产生一个输出, $\hat{Y} = [y_1, \dots, y_c]$ 表示所有通道输出, $n = [n_1, \dots, n_c]$ 表示对应通道的期望噪声。为了找到自适应栈式修正降噪自编码器的最优通道权重,通过最小化的损失函数为

$$\operatorname{argmin} \left(\frac{1}{2} \|\hat{Y}s - n\|^2 \right), \quad (13)$$

$$0 \leq s_c \leq 1, \forall c, \quad (14)$$

$$1 - \delta \leq \sum_{c=1}^C s_c \leq 1 + \delta, \quad (15)$$

式中 s 为通道权重向量; s_c 为通道 c 的权重。(15) 式可以避免某些通道权重受权重总和的影响而出现过大或过小的现象。虽然通道权重总和为 1 更符合人的直观感受,但是如果能够给予一定的灵活性,网络性能会稍微提高。本文中,取 $\delta = 0.04$ 。

3.2.2 训练权重预测模型

最后一个阶段是权重预测模型的训练。由于径向基函数(RBF)网络能够逼近任意非线性函数,可以处理系统内的难以解析的规律性,具有良好的泛化能力^[18]。鉴于此,权重预测模型利用 RBF 网络通过输入 2.2.1 节所述的特征训练集,来拟合最优通道权重。

经过上述 3 个阶段,自适应栈式修正降噪自编码器训练完成。测试时当输入一幅任意种类含噪图像 x 时,多通道联合输出噪声的逼近值 \hat{Y} ,同时计算出所有通道的隐层特征集 ϕ ,将 ϕ 输入权重预测模型,网络会根据数据深层特征自动预测出各通道最优权重 s^* 。该模型最终降噪后的图像为 $x - \hat{Y}s^*$ 。

4 实验结果与分析

4.1 实验数据

实验数据均取自图像去噪领域常用的自然图像数据集。训练样本选自文献[19]中的 400 幅 180 pixel×180 pixel 的 Train400,由于深度神经网络的训练需要海量数据样本,为了丰富训练数据和提高模型的泛化能力,实验采用数据增强的方式对原始数据库进行扩增。即对 Train400 分别进行了 90°、180°旋转变换以及 270°旋转变换和上下翻转,总计 1600 幅图像作为最终的训练样本。同时,为及时验证训练模型性能,实验选取 BSD68 数据库,总计 68 幅 321 pixel×481 pixel 的自然图像作为验证

集。测试集选取了 12 幅 256 pixel×256 pixel 的细节纹理较为丰富的标准测试集 Set12。

4.2 收敛效果分析

为验证模型 SRDA 的有效性,本文对 SRDA 和 SSDA 进行对比实验。SRDA 和 SSDA 的层数均为 3 层,网路输入为 80000 张 8 pixel×8 pixel 的图像块,隐藏层神经元个数均为 40, $\lambda = 10^{-4}$ 、 $\beta = 10^{-2}$ 、 $\rho = 0.01$ 。实验采用动量梯度下降法进行训练,批样本大小为 1000, epoch 设为 10,动量项为 [0.9 0.95 0.99],每 3 个 epoch 更新一次动量项,学习率的初始值设为 0.1,并且每经过 3 个 epoch 后缩小 10 倍,总共需要 800 次迭代。实验结果如图 5 所示。

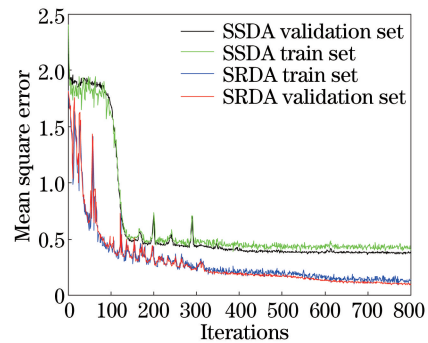


图 5 收敛效果对比

Fig. 5 Comparison of convergence effects

由此可知,SRDA 相比于 SSDA 网络可以更好地拟合样本数据。无论是训练集还是验证集,SRDA 均能达到更低的均方误差,且所需迭代次数明显少于 SSDA。均方误差反映了降噪后图像与原始图像之间的差距,间接验证了 SRDA 网络降噪后图像保留了更多原始图像信息,具有更好的降噪性能。而两个网络验证集均方误差曲线和训练集均方误差曲线基本一致,也进一步说明了模型超参数选择的合理性,网络泛化能力较好。

4.3 自适应去噪

为了验证所提出的 ASRDA 算法的普适性,以及更好地训练 ASRDA 模型,实验对训练样本集分别添加高斯噪声、乘性噪声和椒盐噪声并且每种噪声有 6 个程度,总计需训练 18 个通道 SRDA。其中,高斯噪声的方差依次为 0.02、0.06、0.10、0.14、0.18 和 0.22,乘性噪声的密度依次为 0.05、0.10、0.15、0.20、0.25、0.30,椒盐噪声密度和乘性噪声密度相同。同时为便于实验对比分析,分别用同种类型所有噪声程度的噪声图像作为训练样本训练模型,记为 G-SRDA、S-SRDA 和 S&P-SRDA,分别对应高斯噪

声、乘性噪声和椒盐噪声训练的 SRDA 网络模型。MC-SRDA 是指训练 18 个通道,但每个通道的权重相等。在测试 ASRDA 模型性能时,对 Set12 数据集分别添加高斯噪声、乘性噪声、椒盐噪声、泊松噪声和均匀分布噪声。其中,每种噪声均有两个程度,依次是 $G1:0.01$ 、 $G2:0.07$ 、 $S1:0.1$ 、 $S2:0.15$ 、 $S\&P1:0.1$ 、 $S\&P2:0.2$ 、 $P1:0.5$ 、 $P2:1.5$ 、 $U1:0.2$ 、 $U2:0.4$ 。

在主观评价中,实验选取 Set12 数据集中 3 幅典型图像为例说明本文算法的去噪效果。实验结果如图 6 所示,由上至下分别添加了方差为 0.01 的加

性高斯白噪声、密度为 0.2 的椒盐噪声和密度为 1.5 的泊松噪声。其中,G-SRDA 和 BM3D 对高斯噪声去除效果较好,降噪后图像边缘纹理等细节特征保留较好,但是对椒盐噪声去除效果不理想,仍残留很多噪声;对泊松噪声有一定去除作用,会使图像非噪声像素点过于平滑,细节信息丢失。S&P-SRDA 虽然对这 3 种噪声图像能够去除,但是降噪图像的质量仍有待提高。相比于其他算法,本文算法 ASRDA 对这 3 种噪声都有效,降噪图像轮廓清晰,纹理细节信息保留更加完整。

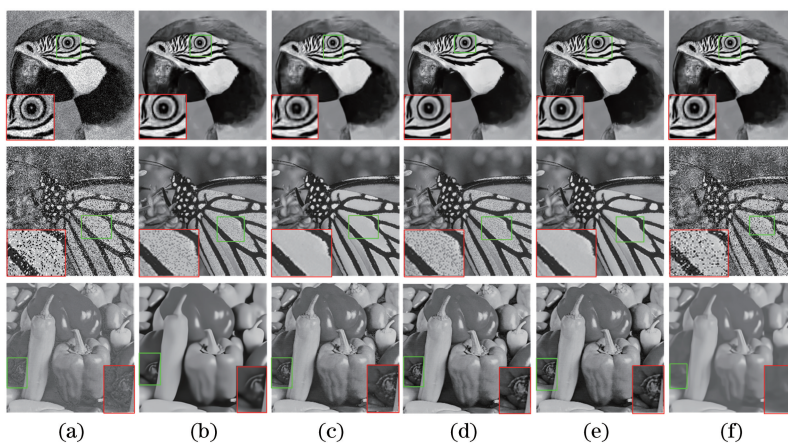


图 6 不同算法的去噪效果对比。(a) Noisy;(b) G-SRDA;(c) S&P-SRDA;(d) MC-SRDA;(e) ASRDA;(f) BM3D
Fig. 6 Comparison of denoising performance of different algorithms. (a) Noisy; (b) G-SRDA; (c) S&P-SRDA; (d) MC-SRDA; (e) ASRDA; (f) BM3D

在客观评价中,使用峰值信噪比 (PSNR) 和结构相似度 (SSIM) 来评价去噪图像的质量。PSNR 反映图像的降噪程度,SSIM 反映图像结构信息的完整性。表 1 所示为不同算法对不同噪声类型的去噪性能。由此可知,G-SRDA、S&P-SRDA 和 S-SRDA 等算法在去除与之相对应的噪声类型时,PSNR 和 SSIM 值较高,而针对其他噪声类型时,二

者值都较低。本文算法 ASRDA 除了在去除乘性噪声上性能略差于 S-SRDA,无论是在训练中出现的噪声类型(如高斯、椒盐噪声),还是未参与训练的噪声(如泊松、均匀分布噪声),其 PSNR 和 SSIM 值整体上都高于其他算法,同时,为验证本文算法的有效性,与去噪效果较好的 BM3D 算法进行实验对比,可以发现,BM3D 算法只对高斯噪声有效,而对其

表 1 Set12 数据集上不同算法实验结果平均 PSNR/SSIM

Table 1 Average PSNR/SSIM of the experimental results of different algorithms on the Set12 data set

Noisy type	G-SRDA	S&P-SRDA	S-SRDA	MC-SRDA	ASRDA	BM3D
G1	27.63/0.8156	25.39/0.7504	25.41/0.7561	27.44/0.8075	29.87/ 0.8517	29.88 /0.8486
G2	25.39/0.7233	23.13/0.6194	22.78/0.6003	25.38/0.7171	25.44/0.7321	25.39/0.7232
S&P1	28.27/0.8406	34.90/0.9851	23.9/0.6795	26.62/0.8002	36.87/0.9863	19.50/0.4326
S&P2	25.75/0.7998	31.66/0.9661	22.92/0.6284	23.79/0.7347	33.66/0.9711	13.71/0.1646
S1	23.87/0.6851	24.45/0.6971	28.58/0.8411	24.91/0.7029	28.22/0.8396	19.40/0.4743
S2	23.6/0.6596	23.85/0.6644	27.75/ 0.8147	24.24/0.6713	27.77/0.8136	16.64/0.3769
P1	23.79/0.6845	24.44/0.7420	23.90/0.6877	24.26/0.6720	28.83/0.8479	27.13/0.8289
P2	23.4/0.6520	23.77/0.6561	22.92/0.6310	23.40/0.6463	25.63/0.7517	20.61/0.7454
U1	23.65/0.6770	24.11/0.6738	25.05/0.6739	24.14/ 0.6815	25.05/0.6669	16.98/0.3098
U2	19.28/0.6368	19.41/0.6366	20.21/0.6500	19.88/0.6443	21.10/0.6610	14.13/0.3036
Average	24.46/0.7210	25.51/0.7391	24.80/0.7014	24.35/0.7083	28.25/0.8122	20.34/0.5208

他噪声类型的两种客观指标值都低于 ASRDA。

综上所述, ASRDA 具有更好的去噪效果, 其既能够满足在主观视觉系统下有所提升, 又能够在客观评价指标上实现提高。同时, 算法克服了大多降噪算法稳健性差的缺点, 在未获取测试图像噪声类型先验的情况下, 也能够自适应处理一定范围内不同类型的噪声。

5 结 论

在深入剖析 SSDA 网络模型基础上, 针对其收敛速度慢、普适性差等问题提出了一种基于栈式修正降噪自编码器的自适应图像去噪方法。该模型采用 ReLU 作为激活函数, 自行引入稀疏性, 并有效缓解了梯度弥散; BN 和残差学习的联合使用降低了网络训练难度, 提高了训练速度和图像降噪性能; 此外, 为克服 SRDA 模型普适性差等问题, 对其进行进一步改进, 充分利用网络特征学习能力进行多通道并行联合自适应训练。实验结果表明: 所提方法不仅在收敛速度上快于 SSDA, 而且在主观视觉和客观评价指标上均好于其他方法, 算法的普适性也得到了提高, 具有一定的借鉴意义和实用价值。今后将对卷积神经网络进行自适应训练, 进一步提高算法的图像降噪性能。

参 考 文 献

- [1] Meng Q K, He J, Qiu R C, *et al.* Crop recognition and navigation line detection in natural environment based on machine vision [J]. *Acta Optica Sinica*, 2014, 34(7): 0715002.
孟庆宽, 何洁, 仇瑞承, 等. 基于机器视觉的自然环境下作物行识别与导航线提取[J]. *光学学报*, 2014, 34(7): 0715002.
- [2] Tu S Q, Xue Y J, Liang Y, *et al.* Review on RGB-D image classification [J]. *Laser & Optoelectronics Progress*, 2016, 53(6): 060003.
涂淑琴, 薛月菊, 梁云, 等. RGB-D 图像分类方法研究综述[J]. *激光与光电子学进展*, 2016, 53(6): 060003.
- [3] Yu L Q, Qin Y L, Zhang X S. Denoising of strong noisy image via gradient reweighted non-local averaging over learned dictionaries [J]. *Laser & Optoelectronics Progress*, 2016, 53(11): 111002.
余临倩, 覃亚丽, 张晓帅. 基于字典学习的梯度重权非局部平均的强噪声图像去噪[J]. *激光与光电子学进展*, 2016, 53(11): 111002.
- [4] Portilla J, Strela V, 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.
- [5] Yan X F, Yang B, Zhang W, *et al.* An improved denoising algorithm of feather and down image based on KSVD [C] // *Proceedings of IEEE International Conference on Information Technology in Medicine and Education*, 2016: 419-423.
- [6] 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.
- [7] Xu S P, Zhang X Q, Jiang Y N, *et al.* Noise level estimation based on local means and its application to the blind BM3D denoising algorithm [J]. *Journal of Image and Graphics*, 2017, 22(4): 422-434.
徐少平, 张兴强, 姜尹楠, 等. 局部均值噪声估计的盲 3 维滤波降噪算法[J]. *中国图象图形学报*, 2017, 22(4): 422-434.
- [8] Yu K, Jia L, Chen Y Q, *et al.* Deep learning: yesterday, today, and tomorrow [J]. *Journal of Computer Research and Development*, 2013, 50(9): 1799-1804.
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.
- [9] Harmeling S, Schuler C J, Burger H C. Image denoising: Can plain neural networks compete with BM3D? [C] // *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 2012: 2392-2399.
- [10] Xie J Y, Xu L L, Chen E H. Image denoising and inpainting with deep neural networks [C] // *Proceedings of the 25th International Conference on neural information processing systems*, 2012: 341-349.
- [11] Li H M. Deep learning for image denoising [J]. *International Journal of Signal Processing, Image Processing and Pattern Recognition*, 2014, 7(3): 171-180.
- [12] Zhang Y G, Yi B S, Wu C Y, *et al.* Low-dose CT image denoising method based on convolutional neural network [J]. *Acta Optica Sinica*, 2018, 38(4): 0410003.
章云港, 易本顺, 吴晨玥, 等. 基于卷积神经网络的低剂量 CT 图像去噪方法[J]. *光学学报*, 2018, 38(4): 0410003.
- [13] Xu J, Xiang L, Liu Q S, *et al.* Stacked sparse autoencoder (SSAE) for nuclei detection on breast cancer histopathology images [J]. *IEEE Transactions on Medical Imaging*, 2016, 35: 119-130.
- [14] Glorot X, Bordes A, Bengio Y, *et al.* Deep sparse

- rectifier neural networks [J]. *Journal of Machine Research*, 2010: 315-323.
- [15] Lu Y S, Li Y X, Liu B, *et al.* Hyperspectral data haze monitoring based on deep residual network [J]. *Acta Optica Sinica*, 2017, 37(11): 1128001.
陆永帅, 李元祥, 刘波, 等. 基于深度残差网络的高光谱遥感数据霾监测 [J]. *光学学报*, 2017, 37(11): 1128001.
- [16] Zhu W, Qu J Y, Wu R B. Straight convolutional neural networks algorithm based on batch normalization for image classification [J]. *Journal of Computer-Aided Design & Computer Graphics*, 2017, 29(9): 1650-1657.
朱威, 屈景怡, 吴仁彪. 结合批归一化的直通卷积神经网络图像分类算法 [J]. *计算机辅助设计与图形学学报*, 2017, 29(9): 1650-1657.
- [17] Ioffe S, Szegedy C. Batch normalization: accelerating deep network training by reducing internal covariate shift [C] // *Proceedings of 32nd International Conference on Machine Learning*, 2015: 448-456.
- [18] Gai J F, Zhao G R, Song C. Model predictive control based on linearization and neural network approach [J]. *Systems Engineering and Electronics*, 2015, 37(2): 394-399.
盖俊峰, 赵国荣, 宋超. 基于线性近似和神经网络逼近的模型预测控制 [J]. *系统工程与电子技术*, 2015, 37(2): 394-399.
- [19] Zhang K, Zuo W M, Chen Y J, *et al.* Beyond a gaussian denoiser: residual learning of deep CNN for image denoising [J]. *IEEE Transactions on Image Processing*, 2017, 26(7): 3142-3155.