

用于去除随机脉冲噪声的两阶段盲卷积降噪模型

徐少平 刘婷云 林珍玉 崔 燕

(南昌大学信息工程学院 南昌 330031)

摘 要 相对于经典的采用逐点检测与复原方式实现的开关型随机脉冲噪声 (Random-Valued Impulse Noise, RVIN) 降噪算法, 基于深度卷积神经网络构建的非开关型 RVIN 降噪模型在降噪效果和执行效率上均有显著优势, 但也存在着固有的数据依赖缺陷, 不能在降噪效果和易用性两个方面同时获得最佳性能. 为此, 以 DnCNN (Denoising Convolutional Neural Network) 深度降噪网络模型架构为设计基础, 提出了一种新的用于去除 RVIN 噪声的两阶段盲卷积降噪 (Two-stage Blind Convolutional Denoising, TBCD) 模型. 在第一阶段, 针对给定的受 0~90% 范围内某个比例 RVIN 噪声干扰的噪声图像, 利用 DnCNN-B (DnCNN for Blind Denoising Task) 盲降噪模型完成初步降噪. 同时, 利用噪声检测模型预测出噪声图像相应的噪声标签, 然后将噪声图像与噪声标签矩阵按位相乘生成稀疏采样图像 (Sparse Sampling Image, SSI). 在第二阶段, 为了进一步提高 DnCNN-B 盲降噪模型所复原的初步降噪图像的质量, 将其与 SSI 图像连接 (concatenate) 后再次输入到预先训练好的双通道图像质量提升模型中获得残差图像, 之后将初步降噪图像减去残差图像得到最终的降噪图像. 与现有的开关型 RVIN 降噪算法相比, 所提出的非开关型 TBCD 模型在各种噪声比例条件下获得的峰值信噪比 (Peak Signal-to-Noise Ratio, PSNR) 指标平均值能高出 5~8 dB, 展现出深度降噪网络模型的显著优势. 而与盲降噪版的 DnCNN-B 和非盲降噪版的 DnCNN-S (DnCNN with Known Specific Noise Ratio) 两个深度降噪模型相比, TBCD 模型复原图像的 PSNR 平均值比 DnCNN-B 高 0.5 dB, 仅比理想实现条件下 (需给定准确的噪声比例值并调用专门训练的降噪模型) 的 DnCNN-S 模型低 0.3 dB. 这表明采用两阶段策略实现的 TBCD 盲卷积降噪模型在降噪效果和易用性两个方面都获得了最佳性能, 很好地解决了深度 RVIN 降噪模型所存在的数据依赖问题.

关键词 随机脉冲噪声; 深度卷积神经网络; 数据依赖; 两阶段; 稀疏采样图像; 降噪效果; 易用性
中图法分类号 TP391 DOI号 10.11897/SP.J.1016.2020.01673

Learning a Two-Stage Blind Convolutional Denoising Model for the Removal of Random-Valued Impulse Noise

XU Shao-Ping LIU Ting-Yun LIN Zhen-Yu CUI Yan

(School of Information Engineering, Nanchang University, Nanchang 330031)

Abstract The traditional switching random-valued impulse noise (RVIN) removal methods normally identify whether the center pixel of a local window is noise or not by comparing its local image statistic with a preset threshold, based on which the identified noise candidates were then suppressed by some noise reduction process. These methods show good denoising effect but at the cost of computational complexity. The deep convolutional neural network (CNN) based non-switching RVIN denoising models have significant advantages in both denoising effect and execution efficiency, in comparison with the classical switching ones that detect and restore RVIN noise pixel-by-pixel. Nonetheless, they cannot obtain best performance with respect to denoising effect and flexibility at the same time due to inherent data

收稿日期: 2019-09-18; 在线发布日期: 2020-02-08. 本课题得到国家自然科学基金项目 (No. 61662044, 61163023, 51765042)、江西省自然科学基金项目 (No. 20171BAB202017) 资助. 徐少平 (通信作者), 博士, 教授, 博士生导师, 主要研究领域为数字图像处理、机器视觉和虚拟手术仿真. E-mail: xushaoping@ncu.edu.cn. 刘婷云, 硕士研究生, 主要研究领域为数字图像处理和机器视觉. 林珍玉, 硕士研究生, 主要研究领域为数字图像处理和机器视觉. 崔 燕, 硕士研究生, 主要研究领域为数字图像处理和机器视觉.

dependency. To this end, a novel two-stage blind convolutional denoising (TBCD) model on the basis of DnCNN(Denoising Convolutional Neural Network) architecture for the removal of RVIN was proposed in this paper. At the first stage, given a noisy image corrupted by RVIN with any noise ratio in the interval [0%, 90%], the DnCNN-B (DnCNN for Blind Denoising Task) model was first applied to removal RVIN noise preliminarily. Meanwhile, a pre-trained noise detection model was exploited to obtain the corresponding noise label matrix of the given noisy image. And then a sparse sampling image was generated by taking dot products between the noisy image and its corresponding noise label matrix. At the second stage, to further improve the quality of the preliminary image restored by the DnCNN-B denoising model at the first stage, the preliminary image and the sparse sampling image were concatenated and fed into a pre-trained dual-channel image quality boosting model to generate a residual image. The final restored latent image can be obtained by subtracting the preliminary restored image from the residual image. The denoising performance of the proposed method was compared with the state-of-the-art switching ones (i.e., PSMF, ROLD-EPR, ASWM, ROR-NLM, MLP-EPR, WCSR, and ALOHA), two CNN-based ones (i.e., DnCNN-B and DnCNN-S), and three regularization-based Gaussian-impulse noise removal methods (i.e., WENSR, WJSR, and LSM-NLR). Experimental results show that the proposed denoising method outperforms all the compared state-of-the-art ones with respect to denoising effect and execution efficiency. Specifically, compared with the existing switching RVIN removal methods, the proposed non-switching TBCD model outperforms them by 5~8dB in terms of peak signal-to-noise ratio (PSNR) across different noise ratios, demonstrating remarkable advantages of the deep CNN-based RVIN denoising model. Compared with the CNN-based DnCNN-B and DnCNN-S (DnCNN with Known Specific Noise Ratio) denoising models, the PSNR performance of TBCD model is about 0.5 dB higher than that of DnCNN-B and 0.3 dB worse than that of DnCNN-S, respectively. Note that the performance of DnCNN-S can only be achieved under ideal circumstances, i.e., the noise ratio of the given noisy image must be accurately measured in advance and the noisy image should also be denoised with the specific pre-trained denoising model accordingly. This indicates that the TBCD model implemented by our two-stage strategy takes competitive advantage in terms of both denoising performance and flexibility, solving the data dependence problem of the non-switching deep RVIN denoising model well. In addition, compared with the mixed Gaussian-impulse noise removal methods, the proposed one exceeds them by at least 5 dB for the case of pure RVIN, and is also more effective than them for the removal of mixed Gaussian-impulse noise. In summary, the proposed denoising method exhibits a competitive performance and shows strong attraction for practical image processing applications.

Keywords random-valued impulse noise; deep convolutional neural network; data dependency; two-stage approach; sparse sampling image; denoising effect; flexibility

1 引 言

数字图像在获取、存储或者在无线通信信道上传输的过程中极易受到脉冲噪声(Impulse Noise, IN)的干扰^[1-3]. 脉冲噪声通常可以分为两类: 固定脉冲噪声(Fixed-Valued Impulse Noise, FVIN)和随机脉冲噪声(Random-Valued Impulse Noise, RVIN)^[4-5]. 相对于 FVIN 噪声而言, RVIN 噪声像素点的亮度值可以是图像像素点亮度值所允许范围内的任意值, 该

值与其邻域内像素点的亮度值往往相差不大, 很难进行区分. 因此, RVIN 噪声的检测与去除更具挑战性^[3-5], 吸引了众多学者的关注. 针对 RVIN 噪声的随机性和不确定性, 研究者们提出了多种 RVIN 降噪算法, 主要包括以下几类:

(1) 基于滤波类. 这类降噪算法主要以中值滤波(Median Filter, MF)衍生算法^[6]为基础, 通过计算当前局部窗口内中值滤波器或其衍生滤波器的输出值与中心像素点亮度值之间的绝对差值, 并将该

绝对差值与预设阈值相比较, 来判定中心像素点是否为噪声点. 对被检测为噪声的像素点, 启动相关降噪程序进行复原, 完成有差别的逐像素点(pixelise)降噪过程, 达到最大限度地保持图像细节的目的. 此类开关型算法实现相对简单, 但保护图像细节的能力有限, 在高噪声比例条件下性能下降非常严重.

(2) 基于图像局部统计值(Local Image Statistic, LIS)类. 例如, Garnett 等人^[7]以中心像素点与其邻域所有像素点之间亮度值的前 m 个最小绝对差值累积和构建 ROAD (Rank-Ordered Absolute Differences) 统计值, 并与预定义的阈值进行比较来确定中心像素点是否为随机脉冲噪声. 针对被判定为噪声的像素点, 利用三边滤波器(trilateral filter)进行降噪并根据像素点的 ROAD 值对其参数进行调整来提升降噪效果. 随后, Dong 等人^[8]通过将 ROAD 统计值作对数变换来扩大噪声点和无噪声像素点之间的差异, 提出了一种称为 ROLD (Rank-Ordered Logarithmic Difference) 的统计特征值. 在此基础上, 利用边缘保持正则化(Edge-Preserving Regularization, EPR)方法完成降噪, 在噪声比例较高时取得了不错的降噪效果. Yu 等人^[9]提出先利用简单加权滤波器对噪声图像进行预处理生成参考图像(reference image), 再基于参考图像构建 RORD (Rank-Ordered Relative Differences) 统计值来提高 RVIN 检测准确性. 为提高噪声检测正确率, 文献[7-9]均采用多次迭代的方式实现, 这些算法的性能严重依赖于迭代过程中阈值的设置, 且在每次迭代过程中受制于某种降噪算法的降噪性能, 需要调整优化的参数较多, 计算效率较低.

(3) 基于改进 LIS 统计值类. 这类 RVIN 降噪算法多以较早提出的 LIS 统计值为基础, 对他们进行必要的改进, 以提高噪声检测正确率, 进而提高降噪效果. 例如, 文献[10]中提出了一种两阶段加权均值滤波器(Weighted Mean Filter with a Two-Phase Noise Detector, WMF-TND)的 RVIN 降噪算法. 该算法先利用 ROD-ROAD(Rank-Ordered Difference of ROAD)和 MEPD(Minimum Edge Pixels Difference)局部统计值提高噪声像素点的检测正确率, 然后用改进的加权均值滤波器完成降噪. Turkmen 将 ROAD 和 ROLD 统计值及他们对应的噪声标签分别作为多层感知网络(Multilayer Perceptron, MLP)的输入和输出, 通过在大量数据集上训练获得噪声检测器^[11]. 一旦噪声检测模型训练完毕, 即可用于判定图像像素点是否为噪声, 快速地完成噪声检测任务,

无需与某种降噪算法绑定迭代执行. 但降噪阶段仍然是采用逐点检测的开关型降噪策略, 故总体执行效率较低.

(4) 基于正则化类. 2014 年, Liu 等人^[12]采用 Laplacian 金字塔框架实现了图像的多尺度表示, 然后利用混合图拉普拉斯正则化(Hybrid Graph Laplacian Regularization, HGLR)模型从粗到精渐进地复原各个尺度空间的图像, 并且利用图像局部光滑特性和非局部的自相似性最大限度地复原噪声图像中的边缘和纹理细节. Jiang 等人^[13]提出一种被称为 WESNR(Weighted Encoding with Sparse Nonlocal Regularization)的可同时去除高斯-脉冲混合噪声的降噪算法. 该算法中的脉冲噪声检测采用加权编码方式隐式实现, 而基于图像稀疏和非局部自相似性的先验知识被集成到正则项中以提高加权编码的稳定性. 之后, Chen 等人^[14]提出了一种称为 WCSR (Weighted Couple Sparse Representation)的脉冲降噪算法. 随后, Deka 等人^[15]利用灰色关联理论(Grey Relational Theory, GRT)和稀疏表示技术区分脉冲噪声和无失真像素点, 级联使用稀疏表示和全变分正则化实现了脉冲噪声的去除. 2017 年, 文献[16]通过对非局部相似的图块进行联合编码, 提出了一种用于去除高斯-脉冲混合噪声的加权联合稀疏表示(Weighted Joint Sparse Representation, WJSR)模型, 该模型依赖 W-SOMP (Weighted Simultaneous Orthogonal Matching Pursuit)算法求解, 执行效率较低. 同年, Huang 等人^[17]基于自然图像所具有的非局部自相似和低秩特性, 利用拉普拉斯尺度混合(Laplacian Scale Mixture, LSM)模型对噪声进行建模, 采用非局部低秩正则化(Nonlocal Low-Rank Regularization, NLR)技术实现了高斯-脉冲混合噪声的去除. 近期, Jin 等人^[3]将脉冲噪声建模为稀疏分量, 基于矩阵补全(matrix completion)并利用稀疏和低秩分解 Hankel 结构化矩阵技术提出了一种工作在图块级上的 ALOHA (Annihilating Filter-Based Low-Rank Hankel Matrix)脉冲噪声降噪方法. 该算法在低比例噪声下能获得良好的降噪效果, 但在高比例噪声条件下需要用其他噪声检测算法替换内置的稀疏离群点检测(sparse outlier detection)模块, 否则将导致其降噪性能急剧下降. 此外, 该算法核心模块虽然已经利用 GPU(Graphic Processing Unit)硬件加速, 但由于采用迭代型的 ADMM(Alternating Direction Method of Multiplier)算法求解稀疏低秩分解问题导致其执行效率仍然比较低.

上述各类 RVIN 降噪算法虽然都取得了一定的

成绩,但仍然存在以下 2 个问题:(1) 降噪效果依赖噪声检测器.大多数 RVIN 降噪算法都内嵌有显式或隐式的噪声检测模块.对于漏检(miss detection)的像素点,RVIN 降噪算法不会进行处理,导致噪声残留;误检(false detection)的像素点则会因执行了降噪过程,导致图像细节模糊.因此,上述两种情况均会影响到开关型 RVIN 降噪算法的降噪性能,尤其是在高噪声比例条件下.噪声比例越高,无失真像素点的个数越少,噪声检测的正确性难以保证.很多依赖无失真像素复原噪声像素点的传统算法,在这种情况下的降噪效果会急剧下降;(2) 执行效率偏低.为了获得较好的噪声检测正确性和降噪效果,大多数算法都设计了复杂的迭代处理步骤,所需执行时间长,且执行时间会随着噪声强度的增加而显著增加.由于 RVIN 降噪算法往往是其他图像处理算法的前置模块,过长的执行时间势必拖累整个图像处理任务的执行效率.因此,采用迭代执行方式的 RVIN 降噪算法在那些对执行时间要求比较高的应用中存在较大局限性.

近年来,DnCNN(Denoising Convolutional Neural Network)^[18]、FFDNet(Fast and Flexible Denoising Convolutional Neural Network)^[19]、UDNet(Universal Denoising Network)^[20]、CBDNet(Convolutional Blind Denoising Network)^[21]等深度网络降噪模型充分利用非线性映射能力强大的卷积神经网络(Convolutional Neural Network, CNN)技术^[22-24]提取图像中蕴含的特征信息,且因软、硬件技术的发展在执行效率方面相较于大多数传统降噪方法更具有优势.例如,Zhang 等人^[18]提出的高斯噪声(Gaussian noise)降噪卷积神经网络 DnCNN 表现出了比传统高斯降噪算法更好的降噪效果和执行效率.从具体的实现过程来看,DnCNN 模型采用图块级的非开关型降噪策略,它的输入为噪声图像,输出为对应的残差图像(噪声图像与无失真图像之间的差值图像),将噪声图像减去残差图像即可得到复原图像.该模型降噪过程的核心计算代价为卷积操作,没有分支判断的处理过程,也无需迭代执行,执行效率很高.但是,DnCNN 模型要想获得最佳的降噪性能,需要针对特定的噪声水平值(反映高斯噪声严重程度的指标)训练多个专用的非盲(non-blind)降噪模型,导致其灵活性和实用性在一定程度上受到限制.为了克服这一缺陷,文献[19]在 DnCNN 框架的基础上进行了修改,通过引入噪声水平映射图(noise level map)作为模型的辅助输入通道,获得了一个降噪能力相当(稍微优于 DnCNN 模型)、但

灵活性和执行效率更高的 FFDNet 降噪模型.噪声水平映射图的引入使得噪声水平值参数从网络模型参数中剥离出来,这样仅需训练 1 个降噪模型即可处理全局噪声水平范围内的高斯噪声,灵活性更好.尽管如此,FFDNet 模型需要在执行降噪过程之前对噪声图像的噪声水平值进行估计,然后才能以该值为基础生成噪声水平图,故仍然属于非盲降噪算法.总之,以 DnCNN 和 FFDNet 为代表的基于数据驱动的非盲降噪模型利用了卷积神经网络强大的建模能力,直接在大量数据上训练获得端到端(end to end)的降噪模型,且在 GPU 硬件支持下具有极高的执行效率.但这类模型获得最佳降噪效果的前提条件是待降噪图像与用于训练降噪模型的图像受到噪声干扰的严重程度相近,存在数据依赖的不足.换句话说,当输入图像受噪声干扰的程度与训练模型所用噪声图像存在较大差距时,他们的降噪性能将会有不同程度的下降.

理论上讲,只要将训练模型的训练对(Training Pairs, TP)中的输入图像由受高斯噪声干扰改为受随机脉冲噪声干扰,DnCNN 高斯降噪网络即可用于去除 RVIN 噪声.这种基于深度卷积网络技术构建的 RVIN 降噪模型将获得比传统逐点开关型 RVIN 降噪算法更高的执行效率和更佳的降噪效果,是非常具有潜力的解决方案.但受基于训练的卷积神经网络降噪技术路线本身的限制(即数据依赖),这种方法仍然需要以牺牲一定降噪效果为代价才能实现盲降噪.其原因在于:为了实现 RVIN 噪声盲降噪,TP 训练对的输入数据中必须包含多种受不同噪声比例值(反映 RVIN 噪声严重程度)的 RVIN 噪声干扰的训练图像,这样训练所获得的降噪模型 DnCNN-B(DnCNN for Blind Denoising Task)才能对任意噪声比例值的噪声图像进行降噪.虽然 DnCNN-B 模型具有良好的灵活性,但因网络拟合难度加大,模型的降噪效果也必然要低于专门针对某一个特定噪声比例值训练的 DnCNN-S(DnCNN with Known Specific Noise Ratio)降噪模型.

为了能够既获得 DnCNN-S 非盲降噪模型的降噪性能又具有 DnCNN-B 盲降噪模型的灵活性,本文在 FFDNet 模型引入噪声水平映射图设计思想的启发下,以 DnCNN 深度卷积神经网络为基础实现了一种用于去除 RVIN 噪声的两阶段盲卷积降噪(Two-stage Blind Convolutional Denoising, TBCD)模型.具体地,在第一阶段中,针对给定的受任意比例 RVIN 噪声干扰的噪声图像,利用 DnCNN-B 降噪模型完成降噪获得 1 张初步降噪图像.同时,

利用预训练好的噪声检测模型预测出给定噪声图像各像素点的噪声标签. 为了利用噪声标签的指示性, 将噪声图像与噪声标签矩阵相乘, 得到 1 张稀疏采样图像 (Sparse Sampling Image, SSI). SSI 图像中噪声点对应位置处的亮度值设置为 0, 而非噪声像素点亮度值与噪声图像对应位置的亮度值相同. 在第二阶段中, 将第一阶段获得的初步降噪图像与 SSI 稀疏采样图像联立, 同时输入到预先训练好的双通道图像质量提升模型中获得残差图像, 将初步降噪图像减去残差图像即可得到最终降噪图像. TBCD 模型两个阶段中所使用到的 3 个子模型均采用与 DnCNN 卷积神经网络相同的核心架构实现. 一旦模型训练完成, 无需用户显式地设置任何参数就能够在任意给定的噪声图像上完成降噪, 属于盲降噪模型. 实验数据表明: TBCD 模型的降噪性能显著优于传统的主流算法. 而相对于基于深度学习构成的降噪模型, TBCD 模型在获得了 DnCNN-B 盲降噪模型灵活性的同时, 其降噪性能高于 Dn-

CNN-B 模型, 并接近于 DnCNN-S 非盲降噪模型在理想条件下获得的降噪水平, 较好地解决了深度 RVIN 降噪模型所存在的数据依赖问题.

2 相关工作

2.1 DnCNN降噪模型

如图 1 所示, DnCNN 去噪卷积神经网络^[18]的隐含层共有 d 层: 第 1 层使用卷积层(Convolutional Layer, Conv)对输入图像进行处理, 并利用 ReLU (Rectifier Linear Unit)激活函数^[25]为网络提供非线性能力, 共采用了 64 个 $3 \times 3 \times 1$ 的滤波器生成 64 个特征映射图; 第 2 ~ $(d-1)$ 层则在卷积层与 ReLU 之间增加了批归一化(Batch Normalization, BN)操作, 主要用于减少内部协变量转移(internal covariate shift)^[26]对网络参数选取的影响, 加快网络模型的收敛速度, 采用的是 64 个大小为 $3 \times 3 \times 64$ 的滤波器; 最后一层通过 1 个单独的卷积操作(滤波器大小为 $3 \times 3 \times 64$)得到输出结果.

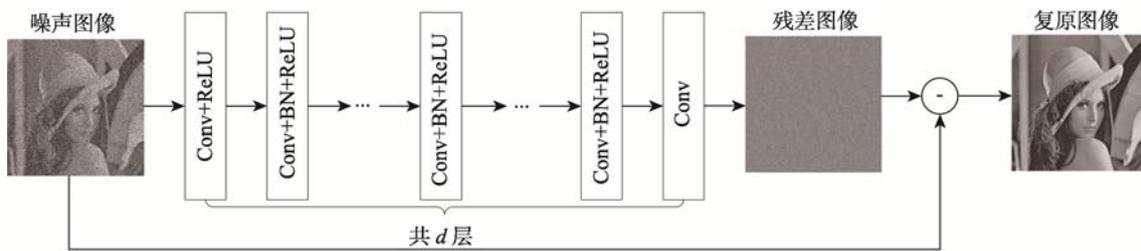


图 1 DnCNN 网络模型的组成架构

与通过学习映射函数 $\varphi(\mathbf{y})$ 直接将输入噪声图像 \mathbf{y} ($\mathbf{y} = \mathbf{x} + \mathbf{v}$, \mathbf{x} 为无失真图像, \mathbf{v} 是噪声)映射为无失真图像 \mathbf{x} 的最佳估计 $\hat{\mathbf{x}}$ 的普通训练模式不同, DnCNN 模型以噪声图像 \mathbf{y} 与原始图像 \mathbf{x} 之间的残差映射图(residual mapping) $\mathbf{v} = \mathbf{y} - \mathbf{x}$ 作为学习的目标值. 对 N 张原始无失真图像添加噪声, 以噪声图像 \mathbf{y} 及其对应的残差图像 \mathbf{v} 分别作为网络模型的输入和输出进行训练, 利用残差学习规则学习残差映射使得 $\mathcal{R}(\mathbf{y}; \Theta) \approx \mathbf{v}$. 在训练模型的过程中, 主要是通过最小化残差图像 \mathbf{v} 与估计残差图像 $\mathcal{R}(\mathbf{y}; \Theta)$ 之间的平均均方误差

$$\ell(\Theta) = \frac{1}{2N} \sum_{i=1}^N \|\mathcal{R}(\mathbf{y}_i; \Theta) - (\mathbf{y}_i - \mathbf{x}_i)\|_{\text{F}}^2 \quad (1)$$

来调整网络模型参数 Θ . 其中, \mathbf{y}_i 和 \mathbf{x}_i 分别表示第 i ($i = 1, 2, \dots, N$) 张噪声图像及其对应的无失真图像. 在应用 DnCNN 模型降噪时, 将噪声图像 \mathbf{y} 减去网络模型所预测的残差映射图 $\mathcal{R}(\mathbf{y})$ 即可得到原始图

像的估计图像 $\hat{\mathbf{x}} = \mathbf{y} - \mathcal{R}(\mathbf{y}; \Theta)$. DnCNN 深度降噪模型存在数据依赖问题, 这意味着要想获得最佳的降噪效果, 必须针对不同的噪声水平值训练多个 DnCNN-S 降噪模型才能满足要求, 不便于训练和使用. 而仅训练一个适用于各种严重级别噪声的 DnCNN-B 模型又不能保证其降噪质量是最佳的.

2.2 FFDNet降噪模型

为提高 DnCNN 降噪模型的灵活性, Zhang 等人^[19]提出了一种被称为 FFDNet 的快速灵活卷积神经网络降噪模型, 其网络结构图如图 2 所示. 不同于 DnCNN 模型采用残差学习技术间接获得清晰图像, FFDNet 模型以噪声图像作为输入, 以对应的无失真图像作为输出进行训练. 模型在使用时直接输出降噪后的图像. 在网络核心架构方面, 两者都采用了“Conv+BN+ReLU”的组合结构. 为了进一步提高 DnCNN 模型的执行效率和灵活性, FFDNet 模型做了以下 2 处改进: (1) 在进行非线性

性映射之前, 采用下采样算子将大小为 $W \times H \times C$ 的输入图像 x 下采样生成对应的 4 张大小为 $W/2 \times H/2 \times 4C$ 子图像 ($C=1$ 时为灰度图像, 而 $C=3$ 时为彩色图像), 在最后一层则采用上采样将子图像重构成大小为 $W \times H \times C$ 的降噪后图像 \hat{x} . 上、下采样层的引入有助于让 FFDNet 网络具有更大的处理视野, 提高执行速度, 但是对提高降噪性能贡献不大. (2) 引入了 1 张大小为 $W/2 \times H/2$ 的噪声水平映射图 M 作为输入信息 (M 图像中所有像素点的亮度值均为噪声图像的噪声水平值 σ), 将其与下采样子图像联接成一个大小为 $W/2 \times H/2 \times (4C+1)$ 的张量作为网络的输入. 噪声水平映射图的引入使得

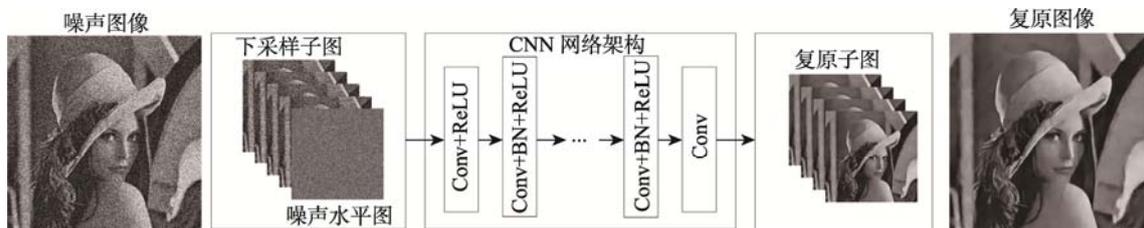


图 2 FFDNet 网络模型的组成架构

2.3 存在的问题

由引言可知: DnCNN 降噪模型原本是用来处理高斯噪声的, 仅需要在训练模型时将 TP 训练对中的输入图像更换为受 RVIN 噪声干扰的图像, 即可将其改造为能处理 RVIN 噪声的降噪模型. 然而, DnCNN 和 FFDNet 这类基于数据驱动的深度卷积网络降噪模型存在数据依赖的缺陷. 限于篇幅, 下面以 DnCNN 降噪模型为例来说明这个问题. 具体地, 对同一失真图像集合分别添加噪声比例值 $r_{\text{train}} = 20\%$, 30% , 40% 的随机脉冲噪声构成 3 个训练集, 训练出 3 个对应的专用 DnCNN-S 降噪模型. 利用这 3 个 DnCNN-S 模型依次对噪声比例值为 $r_{\text{test}} = 30\%$ 的 Lena 噪声图像进行降噪处理, 视觉效果对比如图 3 所示. 从主观视觉上来看: 图 3(c)、图 3(d)和图 3(e) 这 3 张降噪图像的整体视觉效果差异比较小, 这是因

单个 FFDNet 降噪模型就能够处理不同噪声水平的噪声, 灵活度更高. FFDNet 降噪模型可以形式化地描述为 $\hat{x} = \varphi(y, M; \Theta)$. 总体上, FFDNet 网络可比 DnCNN 网络的深度更浅、感受野更大、可学习的网络参数 (learnable parameters) 更少. FFDNet 模型比 DnCNN 模型的灵活性要好一些, 仅需要训练 1 个降噪模型即可对具有各种噪声水平值的噪声图像进行降噪. 但是在实际降噪过程中, 由于噪声水平映射图 M 中的噪声水平值 σ 仍然要由用户准确设置才能获得令人满意的降噪效果, 故本质上 FFDNet 降噪模型还是属于非盲降噪算法.

为降噪后图像的 PSNR (Peak Signal-to-Noise Ratio) 值在 38dB 以上且图像是整体缩小观察的. 然而, 从局部细节放大图中的眼珠和下睫毛对比可以看出, 图 3(c)中仍然残留了少数噪声点未被处理, 图 3(e)的边缘细节部分过于平滑, 而图 3(d)的纹理细节保持相对良好. 图 3(c)的 PSNR 值为 38.36dB, 图 3(d)的 PSNR 值达到了 40.38dB, 图 3(e)的 PSNR 质量值为 38.76dB, 使用 $r_{\text{train}} = 30\%$ 训练的降噪模型对 $r_{\text{test}} = 30\%$ 图像的降噪效果最好. 由上述实验数据可知: 当测试图像的噪声比例值 r_{test} 与训练图像集合中所添加的噪声比例值 r_{train} 相等时, 才可以获得最好的降噪效果; 当 r_{test} 和 r_{train} 不匹配时, 降噪效果则会存在不同程度的下降. 总之, 以 DnCNN 为代表的基于训练策略实现的 RVIN 降噪模型存在固有的数据依赖缺陷.

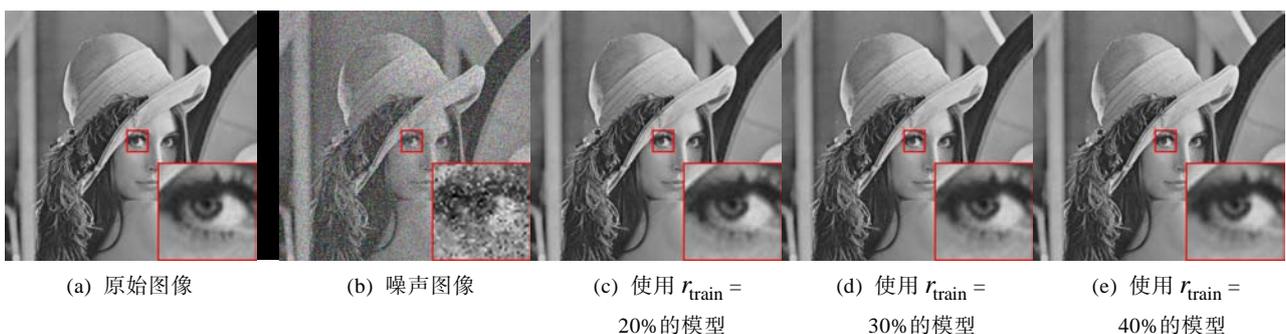


图 3 特定 DnCNN 降噪模型处理不同比例 RVIN 噪声图像的降噪效果对比

3 两阶段盲卷积降噪模型

3.1 基本思想与实现框架

鉴于基于数据驱动的深度降噪模型固有的数据依赖缺陷, DnCNN 降噪模型获得最佳降噪效果的前提是待降噪图像与训练图像受噪声干扰严重程度一致, 这意味着盲降噪版的 DnCNN-B 模型的降噪能力必然要弱于非盲版的 DnCNN-S 模型. 此外, 由 FFDNet 降噪模型可知, 如果在 TP 训练对中的输入数据中增加一个反映噪声图像受 RVIN 噪声干扰情况的辅助通道数据, 那么只需训练一个降噪模型即可处理全局范围内的噪声. 为获得降噪效果和灵活

性俱佳的降噪算法, 受 FFDNet 降噪模型通过引入辅助通道提升模型性能的启发, 以 DnCNN 网络架构为核心基础提出一种称为 TBCD 的两阶段 RVIN 噪声盲降噪模型. TBCD 模型的核心设计思想是: 通过设计一个 RVIN 噪声检测模型预测出给定噪声图像相应的 SSI 稀疏采样图像作为辅助指示信息, 与 DnCNN-B 模型输出的初步降噪图像连接(concatenate)后输入到预先训练的双通道图像质量提升模型中进一步降噪, 以便在实现盲降噪的同时, 其降噪效果也可以接近 DnCNN-S 非盲降噪模型所能达到的水平, 使其在实际应用系统中更具灵活性和实用性. TBCD 模型的具体执行过程如图 4 所示:

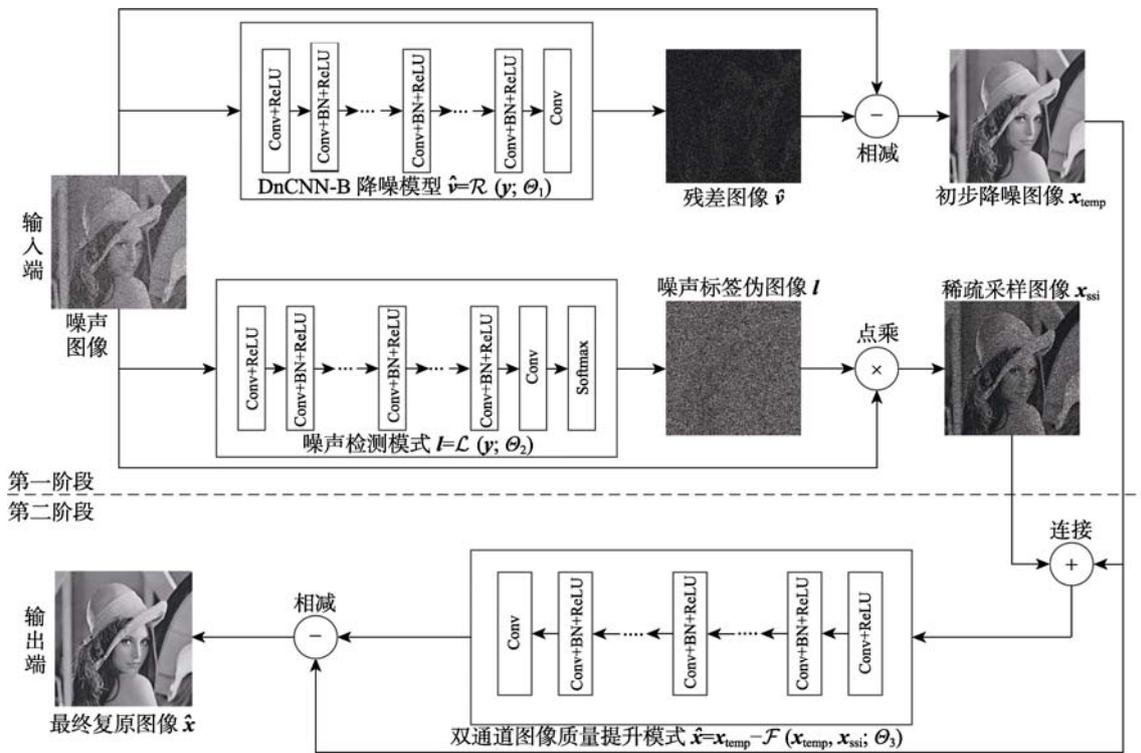


图 4 TBCD 两阶段 RVIN 噪声盲卷积降噪模型的执行流程框图

首先, 采用文献[18]所提出的 DnCNN 深层卷积神经网络架构在全局噪声比例范围(本文以噪声比例值作为度量噪声图像受 RVIN 噪声干扰的程度指标, 考虑随机脉冲噪声的比例值变化范围为[0, 90%], 在该区间内训练一个用于去除 RVIN 噪声的 DnCNN-B 盲降噪模型 $\hat{y} = \mathcal{R}(y; \Theta_1)$, 输入给定的待降噪图像 y , 得到对应的残差图像 \hat{y} 作为输出, 将噪声图像减去残差图像获得对应的初步降噪图像 $x_{temp} = y - \hat{y}$. 另外, 采用相似的深层卷积神经网络架构训练了一个噪声标签预测模型 $l = \mathcal{L}(y; \Theta_2)$, 用于获得噪声图像中每个像素点 $y_{i,j}$ 对应的噪声标签

$l_{i,j}$. 将噪声图像与对应的噪声标签矩阵 l 点乘获得一张稀疏采样图像 $x_{ssi} = y \times l$. 然后, 以稀疏采样图像 x_{ssi} 作为噪声图像 y 中正常像素点的辅助指示信息, 将第一阶段初步复原的图像 x_{temp} 与该稀疏采样图像 x_{ssi} 连接, 作为双通道图像质量提升模型 $\mathcal{F}(x_{temp}, x_{ssi}; \Theta_3)$ 的输入, 进一步提高初步降噪图像 x_{temp} 的图像质量. 双通道图像质量提升模型 $\mathcal{F}(\cdot)$ 同样采用了残差学习技术, 故将初步降噪图像 x_{temp} 与双通道降噪模型 $\mathcal{F}(\cdot)$ 输出的残差图像相减, 即可获得最终的复原图像 $\hat{x} = x_{temp} - \mathcal{F}(x_{temp}, x_{ssi}; \Theta_3)$. 需要说明的是: 双通道图像质量提升模型 $\mathcal{F}(\cdot)$ 其实

与 FFDNet 模型的架构类似,但与 FFDNet 模型直接使用噪声图像不同的是,本文使用由 DnCNN-B 模型得到的初步降噪后的图像(图像质量远远优于噪声图像).这种处理方式可以让模型在训练的时候速度更快,且残差学习技术的使用保证了最终复原图像(初步输出图像减去残差图像)的降噪效果更好.另外,本文设计的 SSI 稀疏采样图像的作用与 FFDNet 模型中的噪声水平映射图的作用是类似的,主要用于提高降噪模型的灵活性,达到仅需训练一个双通道图像质量提升模型即可在全局噪声比例范围内实现盲降噪的目的.

所提出的 TBCD 盲卷积降噪模型的优点有:(1)降噪效果好.在第二阶段执行之前利用噪声检测模型较为准确地判定待降噪图像中噪声像素点位置,初步降噪图像在 SSI 稀疏采样图像辅助指示下同时输入到预先训练的双通道图像质量提升模型 $\mathcal{F}(\cdot)$ 中,使得最终的图像降噪效果得到较大幅度提升.对于任意给定的受 RVIN 噪声干扰的待降噪图像,所提出的 RVIN 降噪模型能够高质量地对其进行复原,盲降噪效果接近于 DnCNN-S 非盲降噪模型的水平.(2)执行效率高.TBCD 模型中所设计的 3 个子模型全部采用 CNN 网络模型实现,没有传统开关型 RVIN 降噪模型反复迭代执行的过程,且受 GPU 硬件支持,故执行效率非常高,能适用于那些对执行时间有严格要求的图像处理系统.(3)使用方便.TBCD 模型中的 3 个子网络模型一旦训练完成,整个降噪过程无需人工参与,且对待降噪图像所受噪声比例干扰没有任何限制,属于盲降噪方法.

3.2 DnCNN-B 降噪子模型

为了去除图像中的 RVIN 噪声,首先需要在全局噪声比例范围内([0, 90%])训练一个 DnCNN-B 盲降噪模型.具体地,从 BSD(Berkeley Segmentation Dataset)数据库^[27]中选取 100 张原始无失真图像,对每张图像分别施加比例为 0~90%(间隔 1%)的随机脉冲噪声生成噪声图像集合.然后将每张噪声图像减去对应的无失真图像,得到相应的残差图像集合.以噪声图像集合和残差图像集合分别作为 DnCNN-B 降噪模型的输入和输出构成 TP 训练对,保持 DnCNN-B 模型的中间结构及网络参数与文献[18]一致,采用随机梯度下降(Stochastic Gradient Descent, SGD)训练算法得到最终的适用于去除 RVIN 噪声的 DnCNN-B 模型.

由上文可知,DnCNN 模型需要在测试图像与训练图像受噪声干扰严重程度近似时才能获得最佳的

降噪效果,那么在全局范围内训练的 DnCNN-B 模型的降噪性能会劣于针对特定噪声水平训练的专用 DnCNN-S 模型(该结论在高斯噪声条件下的证明已在文献[18]中给出).在 RVIN 噪声条件下,为了说明这一点,使用 DnCNN-S 模型和 DnCNN-B 模型分别对 50 张在不同噪声比例下 RVIN 噪声图像(与训练图像不重复)进行降噪,以降噪后图像的 PSNR 平均值作为图像质量评价指标,实验数据列于表 1 中.从表 1 可知,在各个 RVIN 噪声比例下,DnCNN-B 模型所获得的结果相对 DnCNN-S 模型而言要差一些,故需要对复原图像质量进一步提升.这说明 DnCNN-B 模型在获得使用灵活性的同时,图像降噪效果下降了.

表 1 DnCNN-S 模型和 DnCNN-B 模型在 50 张图像上的降噪效果对比(单位: dB)

对比算法	噪声比例								
	10%	20%	30%	40%	50%	60%	70%	80%	90%
DnCNN-S	39.87	35.64	33.13	30.38	28.81	26.86	24.77	22.92	20.45
DnCNN-B	37.71	34.35	31.93	29.85	27.94	26.13	24.32	22.48	19.86

3.3 噪声检测子模型

FFDNet 降噪模型的性能相对于 DnCNN 模型有所提升的关键原因之一在于噪声水平映射图的引入,也就是充分利用了高斯噪声的噪声水平值作为图像受噪声干扰的先验知识.为了在 RVIN 噪声条件下提供与噪声水平映射图类似的功能,根据 RVIN 脉冲噪声特点,本文设计了一种能够反映 RVIN 噪声信息的 SSI 稀疏采样图像 x_{ssi} ,将其作为网络的辅助通道数据.噪声标签是传统开关型 RVIN 降噪算法用来判定图像中某个像素点是否需要复原的依据,本文用“0”标记噪声点,用“1”标记正常像素点.为了更好地适用于 CNN 网络,本文并不直接使用简单的由“0”和“1”构成的矩阵作为最终的映射图,而是使用噪声标签矩阵与噪声图像相乘所得到的 SSI 稀疏采样图像来提供先验信息.这种做法使得无失真像素点的亮度值保持不变,而所有判定为噪声的像素点其值全都置为 0,有利于后续模型的训练.

如图 3 所示,所提出的噪声检测模型仍然采用深层卷积神经网络技术实现从噪声图像 y 到噪声标签 l 的映射(预测).该模型在结构上其实与 DnCNN-B 非常类似,仅在模型的最后增加了一个 Softmax 层来实现噪声像素点的判定.为训练噪声检测模型,从 BSD 数据库^[27]中随机选取 100 张原始无失真图像,对每张图像分别施加比例为 0~90%(间隔 1%)

的 RVIN 噪声得到噪声图像集合, 将噪声图像与原始无失真图像相减获得对应的标签图像作为训练目标. 最后, 以噪声图像集合作为深层卷积神经网络 (网络参数与文献[18]一致) 的输入, 对应的噪声标签矩阵集合作为输出, 训练得到噪声标签检测模型. 一旦模型训练好后, 即可用于获取任意给定噪声图像的稀疏采样图像.

稀疏采样图像 x_{ssi} 作为后续双通道图像质量提升模型 $\mathcal{F}(\cdot)$ 的关键辅助指示输入数据, 其预测正确性对算法降噪效果的提升至关重要. 为了验证所提出的噪声检测模型 $\mathcal{L}(y; \Theta_2)$ 的检测正确率, 使用预先训练好的噪声检测模型分别对大小为 512×512 的、受不同比例 RVIN 噪声 (10%~90%, 间隔 10%) 干扰的 Lena、Barbara 和 Boat 图像中的噪声像素点进行检测, 统计该模型在每张噪声图像上的误检数、

漏检数以及检测正确率 (检测正确率 = $1 - (\text{误检数} + \text{漏检数}) / \text{像素点总数}$), 实验数据如表 2 所示. 由表 2 可知, 所提出的噪声检测模型在 10% 低比例 RVIN 噪声的条件下检测正确率最高能达到 99%, 在 90% 高比例 RVIN 噪声条件下的检测正确率也能达到 91% 以上, 展现了令人满意的预测正确性 (Barbara 图像的纹理细节非常丰富, 检测难度非常大, 故检测准确率要稍微低一些). 由文献[2]可知, 对于 1 张大小为 512×512 的受 40% RVIN 噪声干扰的 Lena 图像, 传统的开关型随机脉冲噪声检测算法对其检测正确率最高只能达到 94%, 而本文所提出的检测模型却可以达到 97.5%, 相当于多正确检测出了将近 8000 个像素点. 这表明所提出的深度噪声检测模型能够为后续图像质量提升模型 $\mathcal{F}(x_{temp}, x_{ssi}; \Theta_3)$ 提供更为准确的辅助指示信息.

表 2 所提出的噪声检测模型在不同图像上检测结果比较

对比图像	检测结果	噪声比例								
		10%	20%	30%	40%	50%	60%	70%	80%	90%
Lena	误检数	91	281	517	893	1622	2817	4718	8889	14027
	漏检数	917	2136	3591	5688	8007	10503	12425	11738	6280
	检测正确率	99.62%	99.08%	98.43%	97.49%	96.33%	94.92%	93.46%	92.13%	92.25%
Barbara	误检数	220	547	1014	1944	3900	7157	12634	18401	18122
	漏检数	2557	5977	10224	14660	18369	20837	18756	11733	4483
	检测正确率	98.94%	97.51%	95.71%	93.67%	91.51%	89.32%	88.03%	88.50%	91.38%
Boat	误检数	174	538	1103	1846	3160	5125	8471	13277	17870
	漏检数	1440	3500	5868	8972	12132	15280	16564	13959	4800
	检测正确率	99.38%	98.46%	97.34%	95.87%	94.17%	92.22%	90.45%	89.61%	91.35%

表 2 中的关于噪声检测模型检测正确率的数据仅从数值上反映了检测结果的正确性, 为了进一步验证该噪声检测模型检测结果在空间分布上的正确性, 根据噪声检测结果采用简单的三角插值函数对受比例为 40% 的 RVIN 噪声干扰的 Lena 图像进行插值复原 (即利用标记为 “1” 的无失真像素点对标记为 “0” 的像素点亮度值进行插值). 由图 5(d) 可以

看出, 所获得的插值后图像与原始无失真图像在视觉上的差异很小 (尽管图像放大观察后仍然可以发现少量位置处存在显著的噪声), 插值后图像的 PSNR 值也达到了 35dB, 这意味着所提出的噪声检测模型的检查结果在图像空域分布上也是有所保证的, 故基于器生成的 SSI 稀疏采样图像能为后继的图像质量提升模型提供有益的指示性信息.

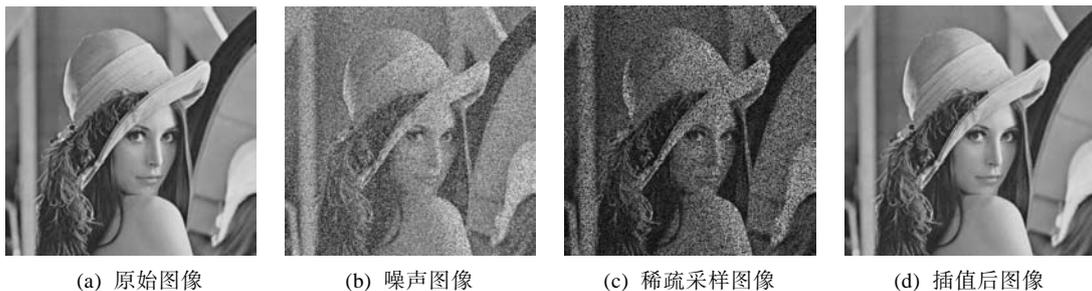


图 5 采用插值函数复原稀疏采样图像的效果

3.4 双通道图像质量提升子模型

在得到待降噪图像 y 对应的稀疏采样图像 x_{ssi} 之后, 将其作为辅助信息和初步降噪图像 x_{temp} 共同输入到双通道的深度卷积神经网络当中, 以进一步提高初步降噪图像的图像质量. 为训练双通道图像质量提升模型, 同样地在 BSD 数据库中选取 100 张原始无失真图像, 并对每张图像施加 0~90%, 间隔为 1% 的 RVIN 噪声, 然后使用 DnCNN-B 模型对每张噪声图像进行复原得到初步降噪图像集合, 再使用噪声标签映射模型获取对应的噪声标签矩阵集合, 并将每张噪声图像与对应的噪声标签矩阵相点乘生成稀疏采样图像集合, 然后将每张稀疏采样图像与相应的第一阶段获得的初步降噪图像相连接 (concatenate), 构成第二阶段降噪网络的输入, 以初步降噪图像 x_{temp} 与原始无失真图像 x 之间的差作为对应的输出, 这样通过训练可以获得双通道质量提升模型 $\mathcal{F}(\cdot)$. 该网络的结构部件、深度以及网络结构参数与文献[18]基本一致, 仅在第 1 个卷积层使用 64 个大小为 $3 \times 3 \times 2$ 的滤波器, 以便处理两个通道的数据. 模型训练完成之后, 将该模型的输出结果与初步降噪图像 x_{temp} 相减即可获得最终的高质量复原图像 \hat{x} .

3.5 各子模型在 TBCD 模型中的作用

所提出的 TBCD 降噪模型主要由 DnCNN-B 降噪模型、噪声检测模型以及双通道图像质量提升模型共 3 个子模型构成. 为了验证每个子模型在整个 TBCD 降噪模型中所起作用的大小, 通过减少某个模块并重新训练降噪模型完成了一组消融实验 (ablation experiment). 具体地, (1) 去除 DnCNN-B 降噪模型, 保留噪声检测模型. 这样直接以稀疏采样图像 x_{ssi} 作为图像质量提升模型 (工作在单通道模式) 输入, 对应的无失真图像 x 作为输出, 通过训练可得到一个降噪模型, 记为 TBCD-S; (2) 保留 DnCNN-B 降噪模型, 去除噪声检测模型. 这样以初

步降噪图像 x_{temp} 和无失真图像 x 分别作为图像质量提升模型 (工作在单通道模式) 的输入和输出训练一个降噪模型, 记为 TBCD-T; (3) 去除 DnCNN-B 降噪模型, 保留噪声检测模型, 以噪声图像 y 和稀疏采样图像 x_{ssi} 共同作为图像质量提升模型 (工作在双通道模式) 的输入, 记为 TBCD-NS. 这样, 分别使用 DnCNN-B、TBCD-S、TBCD-T、TBCD-NS 和 TBCD 共 5 个模型对相同的测试图像集 (包含 50 张与训练集不同的图像, 每张图像均受 10%~90%, 间隔 10% 的 RVIN 噪声干扰) 进行降噪, 以各个比例下 50 张复原图像的 PSNR 均值作为评价模型降噪效果整体性能指标, 实验数据如表 3 所示.

由表 3 可知: TBCD-T 模型比 TBCD-S 模型的性能指标普遍要高一些, 这说明对于第二阶段的图像质量提升模型来说, 输入图像 x_{temp} 比图像 x_{ssi} 所能提供的有价值的信息更多, 更有利于提高 TBCD 模型的整体降噪效果. 但是 TBCD-T 模型比 TBCD 模型性能略差, 说明图像 x_{ssi} 能为图像 x_{temp} 提供很好的指示信息, 两者结合使用有利于提高 TBCD 模型的整体降噪效果. 另外, 从采用噪声图像 y 和稀疏采样图像 x_{ssi} 作为图像质量提升模型输入的 TBCD-NS 模型的性能数据来看, 其降噪效果随着噪声比例的增加严重下降, 这表明在缺失图像 x_{temp} 情况下, 即便有图像 x_{ssi} 帮助, TBCD 模型的性能也会比较差. 此外, TBCD-T 模型所得到的结果比 DnCNN-B 模型高 0.1 dB~0.5 dB. 这表明尽管 TBCD-T 模型中没有使用图像 x_{ssi} , 第二阶段的图像质量提升模型仍然可以进一步提升第一阶段的降噪效果. 上述分析表明: 使用图像 x_{temp} 而不是噪声图像 y 作为双通道图像质量提升模型的输入图像可以使得 TBCD 模型的降噪性能得到较大提高. 这是因为 x_{temp} 图像本身的图像质量已经相对于噪声图像 y 得到显著提升, 有利于第二阶段图像质量提升模型的训练. 同时, 稀疏采样图像 x_{ssi} 中仅保留了被判定为正常像素点

表 3 DnCNN-B 降噪子模型、噪声检测子模型和图像质量提升子模型对 TBCD 模型降噪性能的影响 (单位: dB)

模型	噪声比例									平均值
	10%	20%	30%	40%	50%	60%	70%	80%	90%	
DnCNN-B	36.63	33.30	30.91	28.97	27.19	25.54	23.84	22.06	19.40	27.54
TBCD-S	36.34	32.35	29.62	27.50	25.70	24.09	22.51	20.41	15.48	26.00
TBCD-T	36.78	33.48	31.10	29.18	27.40	25.75	24.07	22.35	19.98	27.79
TBCD-NS	35.50	28.92	22.10	16.22	12.22	9.83	7.99	6.17	4.86	15.98
TBCD	37.66	34.00	31.44	29.41	27.57	25.87	24.15	22.38	19.88	28.04

注: 最优的数据加粗表示. 在 TBCD-[X][X]命名规则中, 后缀为单个字母或者两个字母分别表示图像质量提升子模型采样的单通道或者双通道输入结构. 其中, 后缀含有字母 N 的表示模型的输入包含噪声图像 y , 后缀字母含有 S 表示模型输入包含稀疏采样图像 x_{ssi} , 后缀字母含有 T 的表示模型输入包括初步降噪图像 x_{temp} .

的信息,具有很好的指示作用,配合图像 x_{temp} 使用后能有效地提高 TBCD 模型的降噪性能.另外,双通道图像质量提升模型具有进一步提升图像质量的作用.综上所述,所提出的 TBCD 降噪模型良好的性能是由 DnCNN-B 降噪模型、噪声检测模型和双通道质量提升模型 3 个子模型共同协作的结果,缺少其中任意一个模块都会导致 TBCD 模型的降噪效果不同程度的下降.

4 实验结果及其分析

4.1 环境介绍

为了验证所提出的 TBCD 盲卷积降噪模型的性能,将其与 7 个经典主流 RVIN 噪声降噪算法(包括 PSMF (Progressive Switching Median Filter) [28]、ROLD-EPR^[8]、ASWM (Adaptive Switching Median Filter) [29]、ROR-NLM (Combining Robust Outlyingness Ratio with Nonlocal Means) [30]、MLP-EPR^[11]、WCSR^[14]和 ALOHA^[3])、3 个基于正则化的

高斯-脉冲混合噪声降噪算法(包括 WENSR^[13]、WJSR^[16]和 LSM-NLR^[17])以及 2 个基于深度学习的 DnCNN-B 和 DnCNN-S 模型共 12 个算法进行对比.所有算法都运行在相同的硬件平台(Intel(R) Xeon(R) CPU E5-1603 v4 @ 2.80GHz RAM 16GB)上.具体地,TBCD、DnCNN-B 和 DnCNN-S 模型在 GPU 环境(显卡为 NVIDIA Quadro M4000,深度学习框架为 PyTorch 1.0.0)下运行,PSMF、ROLD-EPR、ASWM、ROR-NLM、MLP-EPR、WCSR、WENSR、WJSR 和 LSM-NLR 算法不受 GPU 支持,则在 CPU 环境(编程环境为 Matlab R2017b)下运行.对比实验所采用的测试集包括 2 个.其中一个是由各个文献中广泛采用的 10 张测试图像构成的常用图像集,如图 6 所示;另外一个是由文献[27]中的 50 张图像(与训练图像不重复)组成的纹理图像集,部分图像如图 7 所示.该图像集中的图像具有丰富的纹理细节特征,特别适用于测试各类 RVIN 降噪算法的鲁棒性.

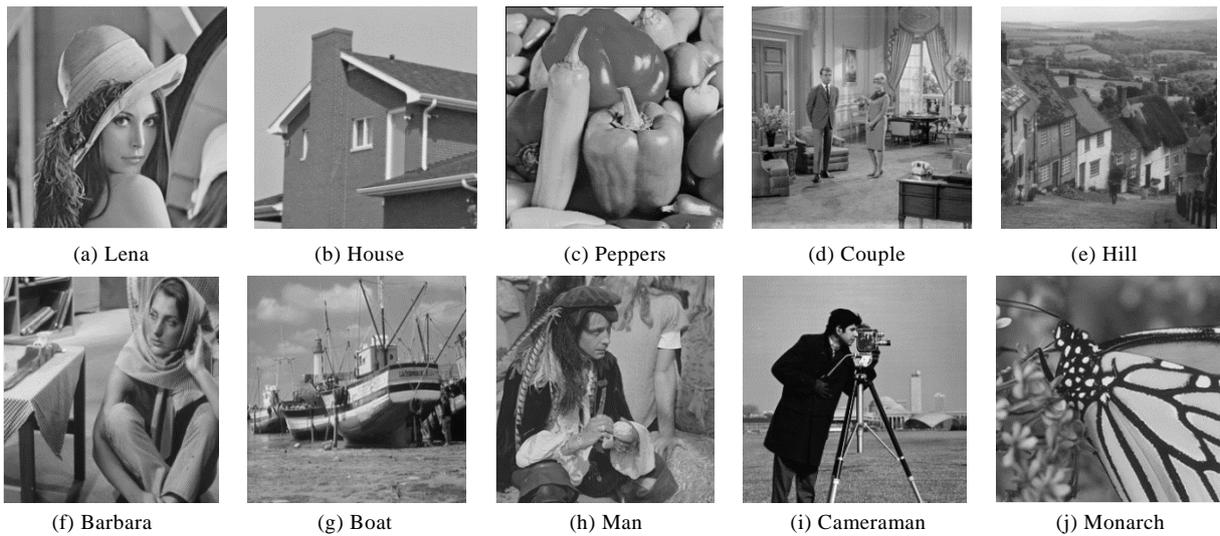


图 6 各类文献中常用的图像集合



图 7 从文献[27]中随机挑选的部分测试图像

4.2 常用图像上的降噪效果

首先,对图 6 所示的 10 张图像分别添加不同比例的随机脉冲噪声(10%~90%,间隔 10%),用各个降噪方法对噪声图像进行降噪并计算每张降噪后图像的 PSNR、SSIM (Structural Similarity)^[31]和 FSIM (Feature Similarity)^[32]值,实验数据列在表 4~表 6 中.限于篇幅,本文只给出各个对比算法在 Barbara 图像上的具体实验数据,以及各个噪声比例下 10 张图像的 PSNR 均值.从表 4~表 6 可知:在 PSNR、SSIM 和 FSIM 客观评价指标上,所提出的 TBCD 盲降噪模型在各个比例下的降噪性能明显优于传统的各个对比降噪算法,也优于基于深度学习的 DnCNN-B,接近于最优的 DnCNN-S 非盲降噪模型(在 SSIM 和 FSIM 指标上, TBCD 模型有时也会优于 DnCNN-S).所提出的 TBCD 模型在低、中、高噪声比例条件下表现得比较稳健,降噪能力随噪声比例的增加下降得最为缓慢.另外,从表 7 所反映出的各个降噪算法在 10 张图像上的平均降噪性能来看, TBCD 盲降噪模型在各个噪声比例条件下的平均性能排名第 2,仅次于 DnCNN-S 模型.需要说明的是: DnCNN-S 模型的实验数据是在理想条件下获得的,在实际应用情况下是很难满足的.换句话说, TBCD 盲降噪模型在所有参与对比降噪算法中降噪性能其实是最佳的.

其次,为了更为直观地展示所提出 RVIN 降噪模型的实际降噪效果,用各个对比算法对受 40% RVIN 噪声干扰的 Lena 图像进行降噪,获得的降噪后图像如图 8 所示.为了更好地对比各个算法的边缘保持能力,选定 Lena 图像左眼局部区域放大后进行比较.仔细观察后可以发现: ROLD-EPR、

ASWM、ROR-NLM 和 MLP-EPR 算法处理后图像的眼珠外部轮廓较为模糊,而基于 CNN 网络模型构建的 DnCNN-B、DnCNN-S 模型以及所提出的 TBCD 盲降噪模型对图像的轮廓保持相对较好,其中 TBCD 算法对边缘的处理效果最好,其所获得的 PSNR 值也很好地反映了这一点.

最后,为了进一步测试 TBCD 盲降噪模型处理高比例 RVIN 噪声和复杂图像内容的的能力,对具有丰富纹理结构的 Barbara 图像施加了 80%高比例的 RVIN 噪声,获得的复原图像如图 9 所示.在所有参与对比的算法中, WENSR、DnCNN-B、DnCNN-S 和 TBCD 算法复原图像的质量相对较高(PSNR 值大于 20dB),图像中的人物和细节基本可以分辨.从图像局部区域的放大图可以看出: TBCD 盲降噪模型在 4 个算法中保留了最为清晰的图像边缘信息.总之,从视觉主观评价的角度来看,图 8~9 充分体现了 TBCD 盲降噪模型具有良好的处理 RVIN 噪声的降噪能力和边缘细节保持能力,在高比例噪声条件下的处理能力也很强,体现了基于深度学习技术构建盲卷积降噪网络模型的优势.

4.3 纹理图像上的降噪效果

限于篇幅,这里仅以各个对比算法在纹理图像集上获得的图像 PSNR 均值作为对比数据,列在表 8 中.从表 8 可知:各个对比算法在纹理图像集上获得的结果平均值有所下降(因为测试图像更为困难),但所提出的 TBCD 盲降噪模型显著优于传统的 RVIN 降噪算法(平均值能高出 5 dB~8 dB),超过 DnCNN-B 模型 0.5dB,低于针对特定噪声比例训练的非盲 DnCNN-S 模型 0.3 dB(仅能在理想条件下获得),其实际的降噪效果是最佳的.

表 4 各个降噪算法在 Barbara 图像所获得降噪后图像的 PSNR 值的比较(单位: dB)

噪声比例	对比算法												
	PSMF	ROLD-EPR	ASWM	ROR-NLM	MLP-EPR	WCSR	ALOHA	WENSR	WJSR	LSM-NLR	DnCNN-B	DnCNN-S	TBCD
10%	25.85	28.93	25.09	24.21	29.76	30.00	37.55	29.82	36.17	28.40	35.23	39.66	36.09
20%	24.69	26.21	24.54	24.01	26.49	27.60	33.40	27.68	35.49	25.50	31.66	35.86	32.22
30%	23.37	24.70	24.03	23.67	24.62	25.78	31.21	25.83	33.62	23.91	29.11	33.59	29.47
40%	21.71	23.69	23.43	23.17	23.40	23.96	26.72	23.74	29.22	22.88	27.08	28.96	27.37
50%	19.79	23.15	22.52	22.20	22.94	22.64	24.61	22.88	28.37	22.11	25.27	26.99	25.48
60%	17.69	22.66	20.89	20.36	22.18	21.53	20.09	22.13	26.90	21.59	23.85	24.35	23.99
70%	15.76	21.85	18.32	18.03	21.40	17.47	16.89	21.39	24.20	20.86	22.90	23.14	23.07
80%	14.01	20.23	15.55	15.86	19.41	11.32	14.73	20.28	20.25	19.53	21.76	22.05	22.05
90%	12.61	16.14	13.35	13.98	14.12	7.48	13.25	16.94	16.31	16.08	19.31	19.91	19.70

注:排名前 2 的数据加粗表示.

表 5 各个降噪算法在 Barbara 图像所获得降噪后图像的 SSIM 值的比较

噪声比例	对比算法												
	PSMF	ROLD-EPR	ASWM	ROR-NLM	MLP-EPR	WCSR	ALOHA	WENSR	WJSR	LSM-NLR	DnCNN-B	DnCNN-S	TBCD
10%	0.8543	0.8999	0.8455	0.8198	0.9378	0.9064	0.9848	0.9128	0.9406	0.8844	0.9834	0.9937	0.9885
20%	0.7730	0.8346	0.8208	0.8024	0.8835	0.8691	0.9535	0.8901	0.9387	0.8258	0.9651	0.9842	0.9729
30%	0.6721	0.7976	0.7941	0.7740	0.8295	0.8250	0.9232	0.8528	0.9206	0.7644	0.9372	0.9726	0.9482
40%	0.5493	0.7634	0.7580	0.7228	0.7618	0.7868	0.7895	0.7753	0.8428	0.7263	0.8974	0.9253	0.9120
50%	0.4218	0.7196	0.6941	0.6254	0.7110	0.7049	0.7421	0.6936	0.8273	0.6665	0.8406	0.8895	0.8592
60%	0.3037	0.6722	0.5754	0.4883	0.6488	0.6536	0.5130	0.6463	0.8011	0.6207	0.7670	0.8001	0.7898
70%	0.2068	0.6006	0.4117	0.3577	0.5752	0.5845	0.2875	0.6121	0.7368	0.5635	0.6947	0.7141	0.7238
80%	0.1313	0.4815	0.2379	0.2611	0.4606	0.3889	0.1448	0.5482	0.6125	0.4636	0.6007	0.6317	0.6392
90%	0.0748	0.3009	0.1229	0.1745	0.3515	0.1539	0.0809	0.4072	0.4656	0.2821	0.4630	0.5042	0.5019

注：排名前 2 的数据加粗表示。

表 6 各个降噪算法在 Barbara 图像所获得降噪后图像的 FSIM 值的比较

噪声比例	对比算法												
	PSMF	ROLD-EPR	ASWM	ROR-NLM	MLP-EPR	WCSR	ALOHA	WENSR	WJSR	LSM-NLR	DnCNN-B	DnCNN-S	TBCD
10%	0.9705	0.9793	0.9672	0.9660	0.9841	0.9827	0.9956	0.9779	0.9861	0.9775	0.9931	0.9974	0.9944
20%	0.9535	0.9633	0.9602	0.9610	0.9679	0.9709	0.9880	0.9685	0.9833	0.9605	0.9860	0.9941	0.9878
30%	0.9240	0.9512	0.9523	0.9524	0.9516	0.9562	0.9780	0.9581	0.9759	0.9447	0.9774	0.9899	0.9796
40%	0.8725	0.9412	0.9391	0.9342	0.9366	0.9429	0.9393	0.9417	0.9349	0.9333	0.9660	0.9745	0.9690
50%	0.7989	0.9325	0.9084	0.8866	0.9289	0.9253	0.9096	0.9226	0.9248	0.9184	0.9532	0.9647	0.9571
60%	0.7087	0.9137	0.8379	0.7976	0.9030	0.9010	0.8288	0.8921	0.9084	0.8957	0.9365	0.9438	0.9420
70%	0.6246	0.8722	0.7222	0.7084	0.8642	0.8318	0.7483	0.8647	0.8650	0.8618	0.9158	0.9243	0.9241
80%	0.5546	0.7851	0.5991	0.6403	0.7542	0.6335	0.6763	0.8094	0.7693	0.7903	0.8708	0.8880	0.8841
90%	0.5148	0.6607	0.5237	0.5887	0.5809	0.4810	0.6198	0.6967	0.6562	0.6774	0.7733	0.7888	0.7827

注：排名前 2 的数据加粗表示。

表 7 各降噪算法在 10 张常用图像不同噪声比例下所获得的 PSNR 均值比较（单位：dB）

对比算法	噪声比例										均值
	10%	20%	30%	40%	50%	60%	70%	80%	90%		
PSMF	30.82	28.15	25.71	23.30	20.86	18.48	16.40	14.58	13.06	21.26	
ROLD-EPR	31.90	29.72	28.31	27.18	26.15	24.90	23.29	20.68	16.60	25.41	
ASWM	30.60	29.47	28.32	27.02	25.23	22.59	19.39	16.36	13.91	23.65	
ROR-NLM	28.73	27.97	27.02	25.82	24.06	21.67	19.02	16.65	14.71	22.85	
MLP-EPR	31.29	29.01	27.74	26.56	25.52	24.24	22.55	20.13	14.83	24.65	
WCSR	31.20	29.04	27.58	26.35	25.31	24.17	21.57	13.60	8.33	23.02	
ALOHA	35.19	31.36	29.56	26.03	24.32	20.67	17.56	15.38	13.85	23.77	
WENSR	30.55	28.78	27.46	26.41	25.45	24.46	22.97	20.73	17.16	24.88	
WJSR	31.91	30.38	29.32	27.77	26.08	24.23	21.79	18.55	15.58	25.07	
LSM-NLR	29.80	27.63	26.17	25.36	24.37	23.45	22.02	19.97	16.71	23.94	
DnCNN-B	39.43	35.94	33.42	31.13	29.06	26.97	24.81	22.33	19.28	29.15	
DnCNN-S	42.37	37.99	35.44	32.01	30.19	27.80	25.46	23.00	19.82	30.45	
TBCD	40.85	36.92	34.21	31.75	29.60	27.41	25.26	22.76	19.77	29.84	

注：排名前 2 的数据加粗表示。

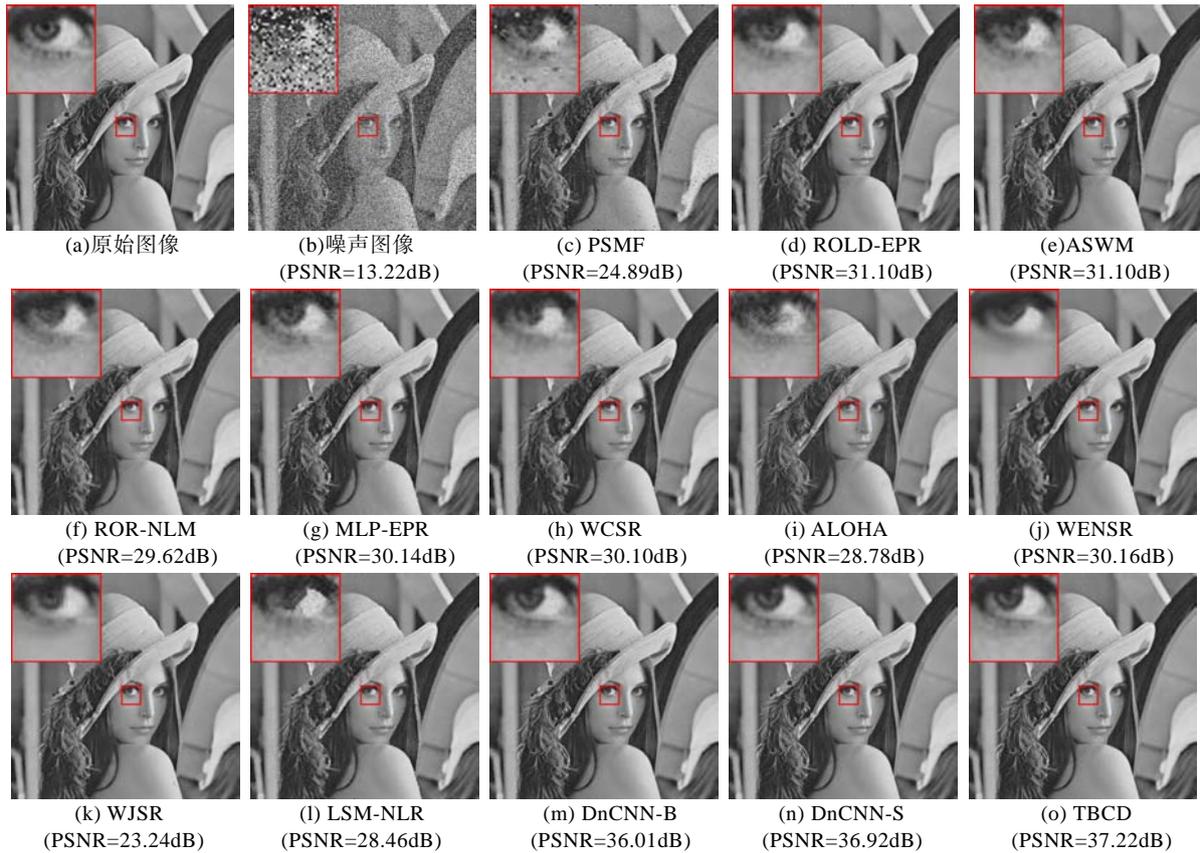


图 8 各算法对噪声比例为 40% 的 Lena 噪声图像降噪后结果的视觉对比

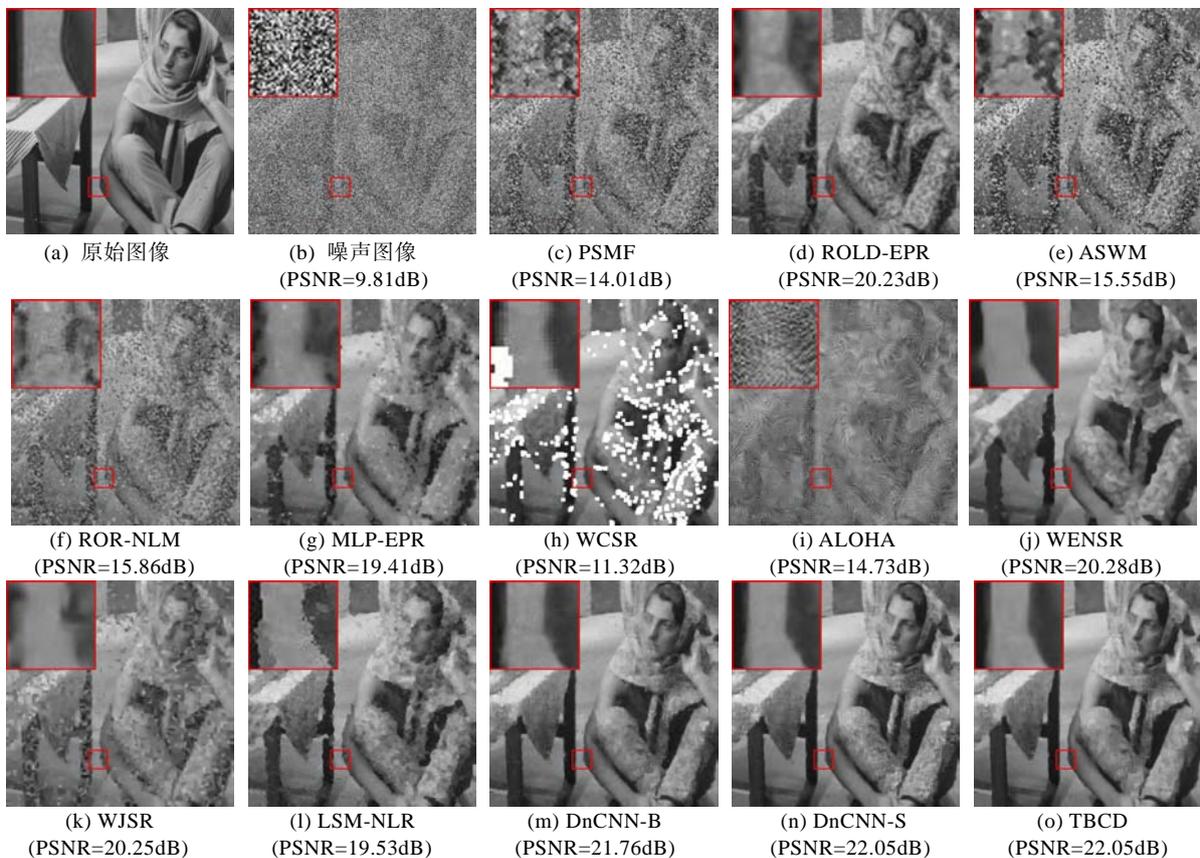


图 9 各算法对噪声比例为 80% 的 Barbara 噪声图像降噪后结果的视觉对比

表 8 各降噪算法在 50 张纹理图像不同噪声比例下所获得的 PSNR 均值比较 (单位: dB)

对比算法	噪声比例									平均值
	10%	20%	30%	40%	50%	60%	70%	80%	90%	
PSMF	28.85	26.77	24.66	22.43	20.11	17.86	15.83	14.10	12.65	20.36
ROLD-EPR	29.98	27.97	26.64	25.66	24.77	23.78	22.43	20.07	16.11	24.16
ASWM	28.25	27.38	26.45	25.31	23.76	21.35	18.55	15.74	13.45	22.25
ROR-NLM	26.88	26.30	25.56	24.55	22.95	20.70	18.26	16.04	14.22	21.72
MLP-EPR	30.03	27.65	26.28	25.21	24.28	23.21	21.72	19.57	14.35	23.59
WCSR	29.55	27.46	26.06	24.96	24.02	23.09	20.68	13.95	8.45	22.02
ALOHA	31.77	29.02	26.87	23.96	22.26	19.28	16.69	14.74	13.31	21.99
WENSR	28.02	26.85	25.82	24.84	24.10	23.31	22.16	20.46	16.93	23.61
WJSR	29.30	27.26	25.79	23.18	22.59	21.76	20.08	17.65	15.23	22.54
LSM-NLR	28.62	26.66	25.38	24.37	23.51	22.68	21.53	19.76	16.34	23.21
DnCNN-B	36.63	33.30	30.91	28.97	27.19	25.54	23.84	22.06	19.40	27.54
DnCNN-S	38.57	34.44	32.04	29.48	27.92	26.09	24.26	22.53	20.07	28.38
TBCD	37.66	34.00	31.44	29.41	27.57	25.87	24.15	22.38	19.88	28.04

注：排名前 2 的数据加粗表示。

4.4 泛化能力

所提出的 TBCD 盲降噪模型的降噪框架对其他噪声类型具有一定的泛化处理能力。为了测试这种能力，选择常见的高斯-脉冲混合噪声进行了实验。相对于单一的高斯噪声或者脉冲噪声而言，高斯-脉冲混合噪声属于强噪声，对图像的破坏模式更为复杂，而且复原难度也更大。具体地，同时对 BSD 数据库中随机挑选的 50 张图像施加不同级别的高斯噪声（噪声水平值 σ 在 0~30 之间）和随机脉冲噪声（噪声比例值 r 在 0~50% 之间）构成训练图像集。虽然这里高斯噪声的噪声水平值最大为 30，随机脉冲噪声比例最大为 50%，但在混合噪声条件下对图像的破坏能力远超单一类型的噪声，属于强噪声。构建完训练集后，不改变 TBCD 网络架构和训练参数配置（仅将网络模型的输入图像改为高斯-脉冲混合噪声图像）重新训练用于处理高斯-脉冲噪声的 TBCD 新模型。然后对 Barbara、Couple、House 这 3 张常用图像分别施加 $\sigma=15$ 、25 的高斯噪声和 $r=25\%$ 、45% 的随机脉冲噪声用于测试（测试条件与文献[16]相同）。将所训练的新 TBCD 降噪模型与 ROR-NLM、WENSR、WJSR 和 LSM-NLR 这 4 种能处理高斯-脉冲混合降噪算法进行对比，测试数据列于表 9 中。由表 9 可知：所提出的 TBCD 盲降噪模型亦能够有效地处理高斯-脉冲混合噪声，获得的 PSNR 值明显优于基于正则化方法实现的传统算法（高 1dB~2dB），这表明所提出用于去除随机脉冲噪声的 TBCD 盲降噪模型也适用于高斯-脉冲混合噪声。需要说明的是：对于常见的高斯噪声，所提出

的 TBCD 盲降噪模型也可以适用，其中双通道图像质量提升子模型起到了关键作用。限于篇幅，这里不再提供具体数据。

表 9 各算法去除高斯-脉冲混合噪声的性能对比 (单位: dB)

图像	σ	r	对比算法				
			ROR-NLM	WENSR	WJSR	LSM-NLR	TBCD
Barbara	15	25%	23.84	25.69	24.68	25.22	28.05
		45%	22.76	23.27	22.84	23.24	25.07
	25	25%	23.85	25.30	23.45	25.16	26.19
		45%	22.74	22.34	21.87	22.38	24.01
Couple	15	25%	27.14	28.20	27.88	28.39	30.00
		45%	24.81	25.80	24.98	25.75	27.64
	25	25%	27.18	26.67	25.73	26.77	27.89
		45%	24.76	24.32	23.09	24.13	25.92
House	15	25%	31.11	32.03	31.30	31.82	33.10
		45%	27.28	28.70	27.73	28.53	30.81
	25	25%	30.77	29.34	28.51	30.04	30.96
		45%	27.12	26.94	25.09	26.04	28.90

注：最好结果加粗表示。

4.5 执行效率

为了测试各个对比降噪算法的执行效率，以 1 张受 40% RVIN 噪声污染的 Lena 图像（大小为 512×512）为例，记录各个算法 10 次复原该噪声图像的平均执行时间并列于表 10 中。由表 10 可知，在 GPU 环境下运行的 DnCNN-B、DnCNN-S 和 TBCD 模型其执行效率明显高于在 CPU 环境下运行的 PSMF、ROLD-EPR、ASWM、ROR-NLM、

MLP-EPR、WCSR、WENSR、WJSR 和 LSM-NLR 算法. 虽然 TBCD 模型的执行时间比 DnCNN-B 和 DnCNN-S 的要长一些, 但由上文降噪效果对比实验可知, 所付出的代价是值得的. 总之, 本文提出 TBCD 盲降噪模型具有出众的降噪能力, 同时具有相对较高的执行效率, 综合优势相对于其他对比算法更高.

表 10 各算法上执行时间对比 (单位: s)

算法	执行时间	算法	执行时间
PSMF	0.5	WENSR	47.7
ROLD-EPR	5.4	WJSR	1486.5
ASWM	88.1	LSM-NLR	262.2
ROR-NLM	41.5	DnCNN-B	0.0151
MLP-EPR	2.6	DnCNN-S	0.0144
WCSR	1079.6	TBCD	0.0306
ALOHA	1897.1		

5 结 语

本文提出了一种被称为 TBCD 的两阶段 RVIN 盲卷积降噪模型. 所提出的 TBCD 降噪模型在盲降噪工作模式下的降噪效果显著优于主流的非开关型 RVIN 降噪算法以及基于正则化技术实现的 RVIN 降噪算法, 超过了 DnCNN-B 盲降噪模型的降噪性能, 接近于理想条件下实现的 DnCNN-S 非盲降噪模型, 摆脱了深度降噪模型固有的数据依赖缺陷. 这主要得益于噪声标签检测模型的引入, 基于该模型生成的 SSI 图像的作用类似于 FFDNet 降噪模型中噪声水平映射图. 基于此, 利用 DnCNN-B 深度网络模型获得的初步降噪图像在 SSI 图像指示下共同输入到双通道图像质量提升模型中, 使得最终获得的降噪图像在质量上比初步降噪图像更高.

参 考 文 献

- [1] Khaw H Y, Soon F C, Chuah J H, et al. High-density impulse noise detection and removal using deep convolutional neural network with particle swarm optimization. *IET Image Processing*, 2019, 13(2): 365-374
- [2] Xu S P, Zhang G Z, Hu L Y, et al. Convolutional neural network-based detector for random-valued impulse noise. *Journal of Electronic Imaging*, 2018, 27(5): #050501. DOI: 10.1117/1.JEI.27.5.050501
- [3] Jin K H, Ye J C. Sparse and low-rank decomposition of a Hankel structured matrix for impulse noise removal. *IEEE Transactions on Image Processing*, 2018, 27(3): 1448-1461
- [4] Xu Q, Li Y H, Guo Y J, et al. Random-valued impulse noise removal using adaptive ranked-ordered impulse detector. *Journal of Electronic Imaging*, 2018, 27(1): #013001. DOI: 10.1117/1.JEI.27.1.013001
- [5] Mafi M, Rajaei H, Cabrerizo M, et al. A robust edge detection approach in the presence of high impulse noise intensity through switching adaptive median and fixed weighted mean filtering. *IEEE Transactions on Image Processing*, 2018, 27(11): 5475-5490
- [6] Chen T, Wu H R. Adaptive impulse detection using center-weighted median filters. *IEEE Signal Processing Letters*, 2001, 8(1): 1-3
- [7] Garnett R, Huegerich T, Chui C, et al. A universal noise removal algorithm with an impulse detector. *IEEE Transactions on Image Processing*, 2005, 14(11): 1747-1754
- [8] Dong Y Q, Chan R H, Xu S F. A detection statistic for random-valued impulse noise. *IEEE Transactions on Image Processing*, 2007, 16(4): 1112-1120
- [9] Yu H C, Zhao L, Wang H X. An efficient procedure for removing random-valued impulse noise in images. *IEEE Signal Processing Letters*, 2008, 15(12): 922-925
- [10] Liu L C, Chen P, Zhou Y C, et al. A new weighted mean filter with a two-phase detector for removing impulse noise. *Information Sciences*, 2015, 315: 4052-4057. DOI: 10.1016/j.ins.2015.03.067
- [11] Turkmen I. The ANN based detector to remove random-valued impulse noise in images. *Journal of Visual Communication & Image Representation*, 2016, 34(1): 28-36
- [12] Liu X M, Zhai D M, Zhao D B, et al. Progressive image denoising through hybrid graph Laplacian regularization: A unified framework. *IEEE Transactions on Image Processing*, 2014, 23(4): 1491-1503
- [13] Jiang J L, Zhang L, Yang J. Mixed noise removal by weighted encoding with sparse nonlocal regularization. *IEEE Transactions on Image Processing*, 2014, 23(6): 2651-2662
- [14] Chen C L P, Liu L C, Chen L, et al. Weighted couple sparse representation with classified regularization for impulse noise removal. *IEEE Transactions on Image Processing*, 2015, 24(11): 4014-4026.
- [15] Deka B, Handique M, Datta S. Sparse regularization method for the detection and removal of random-valued impulse noise. *Multimedia Tools and Applications*, 2017, 76(5): 6355-6388
- [16] Liu L C, Chen L, Chen C L P, et al. Weighted joint sparse representation for removing mixed noise in image. *IEEE Transactions on Image Processing*, 2017, 47(3): 600-611
- [17] Huang T, Dong W S, Xie X M, et al. Mixed noise removal via Laplacian scale mixture modeling and nonlocal low-rank approximation. *IEEE Transactions on Image Processing*, 2017, 26(7): 3171-3186
- [18] Zhang K, Zuo W M, Chen Y J, et al. Beyond a Gaussian denoiser: Residual learning of deep CNN for image denoising. *IEEE Transactions on Image Processing*, 2016, 26(7): 3142-3155
- [19] Zhang K, Zuo W M, Zhang L. FFDNet: Toward a fast and flexible solution for CNN based image denoising. *IEEE Transactions on Image Processing*, 2018, 27(9): 4608-4622
- [20] Lefkimmatis S. Universal denoising networks: A novel CNN architecture for image denoising. *Proceedings of the 2018*

- IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA, 2018: 3204-3213
- [21] Chen J W, Chen J W, Chao H Y, et al. Image blind denoising with generative adversarial network based noise modeling// Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA, 2018: 3155-3164
- [22] Zhu-Hui, Qin Pin-Le. U-Net pulmonary nodule detection algorithm based on multi-scale feature structure. *Computer Engineering*, 2019, 45(4): 254-261 (in Chinese)
(朱辉, 秦品乐. 基于多尺度特征结构的 U-Net 肺结节检测算法. *计算机工程*, 2019, 45(4): 254-261)
- [23] Ren Jing-Jing, Fang Xian-Yong, Chen Shang-Wen, et al. Image deblurring based on fast convolutional neural networks. *Journal of Computer-Aided Design and Computer Graphics*, 2017, 29(8): 1444-1456 (in Chinese)
(任静静, 方贤勇, 陈尚文, 等. 基于快速卷积神经网络的图像去模糊. *计算机辅助设计与图形学学报*, 2017, 29(8): 1444-1456)
- [24] Chang Liang, Deng Xiao-Ming, Zhou Ming-Quan, et al. Convolutional neural networks in image understanding. *Acta Automatica Sinica*, 2016, 42(9): 1300-1312 (in Chinese)
(常亮, 邓小明, 周明全, 等. 图像理解中的卷积神经网络. *自动化学报*, 2016, 42(9): 1300-1312)
- [25] Rawat W, Wang Z H. Deep convolutional neural networks for image classification: A comprehensive review. *Neural Computation*, 2017, 29(9): 2352-2449
- [26] Li Y H, Wang N Y, Shi J P, et al. Adaptive batch normalization for practical domain adaptation. *Pattern Recognition*, 2016, 80: 109-117 DOI : 10.1016/j.patcog.2018.03.005
- [27] Arbelaez P, Maire M, Fowlkes C, et al. Contour detection and hierarchical image segmentation. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 2011, 33(5): 898-916
- [28] Zhang Z, Zhang D. Progressive switching median filter for the removal of impulse noise highly corrupted images. *IEEE Transactions on Circuits and System II Analog Digit Signal Processing*. 1999, 46(1): 78-80
- [29] Akkoul S, Lédée R, Leconge R, et al. A new adaptive switching median filter. *IEEE Signal Processing Letters*, 2010, 17(6): 587-590
- [30] Xiong B, Yin Z P. A universal denoising framework with a new impulse detector and nonlocal means. *IEEE Transactions on Image Processing*, 2012, 21(4): 1663-1675
- [31] Zhou W, Simoncelli E P, Bovik A C. Multiscale structural similarity for image quality assessment// Proceedings of the 37th Asilomar Conference on Signals, Systems and Computers. Pacific Grove, USA, 2003: 1398-1402
- [32] Zhang L, Zhang L, Mou X Q, et al. FSIM: A feature similarity index for image quality assessment. *IEEE Transactions on Image Processing*, 2011, 20(8): 2378-2386



XU Shao-Ping, Ph.D., professor, Ph.D supervisor. His main research interests include digital image processing, machine vision, and virtual surgery simulation.

LIU Ting-Yun, master candidate. Her main research interests include

digital image processing and machine vision.

LIU Zhen-Yu, master candidate. Her main research interests include digital image processing and machine vision.

CUI Yan, master candidate. Her main research interests include digital image processing and machine vision.

Background

The research is devoted to removing random-valued impulse noise (RVIN) in images. As only some of the pixels in a noisy image are corrupted by RVIN and the others are noise-free, traditional RVIN denoising algorithms adopt so-called switching strategy that first detects RVIN and then removes the detected RVIN noise. Obviously, the denoising performance of switching-based RVIN denoising methods is subject to the accuracy and execution efficiency of the noise detection. In recent years, deep convolutional neural network (DCNN) based denoising methods achieve great success regarding denoising effect and has superiority over traditional ones in execution efficiency. For instance, the state-of-art Gaussian denoising method DnCNN (denoising convolutional neural network) shows stronger denoising

ability and faster execution speed than the well-known denoising algorithms such as BM3D. However, due to the inherent data dependency limitation, the DCNN-based denoising models can only achieve blind denoising at the cost of the denoising effect. The denoising ability of the blind DnCNN denoising model (DnCNN-B) that is trained across the entire range of noise ratios is weaker than that of its non-blind version with known specific noise ratio (namely, DnCNN-S), although DnCNN-B is more flexible than DnCNN-S. To obtain a novel denoising model that is of the denoising effect of the DnCNN-S denoising model and the flexibility of the DnCNN-B model, a two-stage blind convolutional denoising (TBCD) model on the basis of the DnCNN architecture for the removal of RVIN was proposed

in this work. Our work is also inspired by the Gaussian denoising model FFDNet (fast and flexible denoising convolutional neural network) that separates the Gaussian noise level from the network parameters. Specifically, a sparse sampling image (SSI) that reflects RVIN distribution characteristic of the given noisy image was generated as auxiliary indication information by a pre-trained noise detection model. And then a preliminary denoised image restored by the DnCNN-B denoising model and the SSI image were concatenated and fed into a pre-trained dual-channel image quality boosting sub-model to obtain better denoising effect. The proposed TBCD model exceeds

blind DnCNN-B denoising model in denoising effect with high flexibility and close to non-blind DnCNN-S denoising model, which solves the data dependence problem of the non-switching deep RVIN denoising model well. In addition, our TBCD model implemented on GPU can meet well with those image processing systems with critical time requirements.

This work is supported by the National Natural Science Foundation of China (Grant No. 61662044, 61163023, and 51765042), and the Natural Science Foundation of Jiangxi Province of China (20171-AB202017).