



JOURNAL OF IMAGE AND GRAPHICS

主办: 中国科学院遥感与数字地球研究所
中国图象图形学学会
北京应用物理与计算数学研究所

中国图象图形学报

2020
07
VOL.25

ISSN1006-8961
CN11-3758/TB





第25卷第7期 (总第291期)
2020年7月16日

中国精品科技期刊
中国国际影响力优秀学术期刊
中国科技核心期刊
中文核心期刊

版权声明

凡向《中国图象图形学报》投稿, 均视为同意在本刊网站及CNKI等全文数据库出版, 所刊载论文已获得著作权人的授权。本刊所有图片均为非商业目的使用, 所有内容, 未经许可, 不得转载或以其他方式使用。

Copyright

All rights reserved by Journal of Image and Graphics, Institute of Remote Sensing and Digital Earth, CAS. The content (including but not limited text, photo, etc) published in this journal is for non-commercial use.

主管单位 中国科学院
主办单位 中国科学院遥感与数字地球研究所
中国图象图形学学会
北京应用物理与计算数学研究所

主 编 顾行发
编辑出版 《中国图象图形学报》编辑出版委员会
通信地址 北京市海淀区北四环西路19号
邮 编 100190
电子信箱 jig@radi.ac.cn
电 话 010-58887035
网 址 www.cjig.cn

广告发布登记号 京朝工商广登字20170218号
总 发 行 北京报刊发行局
订 购 全国各地邮局
海外发行 中国国际图书贸易集团有限公司
(邮政信箱: 北京399信箱 邮编: 100048)
印刷装订 北京科信印刷有限公司

Journal of Image and Graphics

Title inscription: Song Jian | Monthly, Started in 1996

Superintended by Chinese Academy of Sciences
Sponsored by Institute of Remote Sensing and Digital Earth, CAS
China Society of Image and Graphics
Institute of Applied Physics and Computational Mathematics

Editor-in-Chief Gu Xingfa
Editor, Publisher Editorial and Publishing Board of Journal of Image and Graphics
Address No. 19, North 4th Ring Road West, Haidian District, Beijing, P. R. China
Zip code 100190
E-mail jig@radi.ac.cn
Telephone 010-58887035
Website www.cjig.cn

Distributed by Beijing Bureau for Distribution of Newspapers and Journals
Domestic All Local Post Offices in China
Overseas China International Book Trading Corporation
(P.O.Box 399, Beijing 100048, P.R.China)
Printed by Beijing Kexin Printing Co., Ltd.

CN 11-3758/TB
ISSN 1006-8961
CODEN ZTTXFZ

国外发行代号 M1406
国内邮发代号 82-831
国内定价 60.00元



非真实感绘制技术研究现状与展望(第1283页)



图卷积网络下牙齿种子点自动选取(第1481页)



迁移学习下高分快视数据道路快速提取(第1501页)

学者观点

非真实感绘制技术研究现状与展望

钱文华, 曹进德, 徐丹, 吴昊 1283

综述

深度哈希图像检索方法综述

刘颖, 程美, 王富平, 李大湘, 刘伟, 范九伦 1296

严肃游戏中虚拟角色行为建模综述

刘翠娟, 刘箴, 柴艳杰, 刘婷婷, 陈效奕 1318

智能制造中的计算机视觉应用瓶颈问题

雷林建, 孙胜利, 向玉开, 张悦, 刘会凯 1330

图像处理和编码

浅层CNN网络构建的噪声比例估计模型

徐少平, 林珍玉, 李崇禧, 崔燕, 刘蕊蕊 1344

烧结断面火焰图像退化模型及断面图像复原

王福斌, 刘贺飞, 武晨 1356

gyrator变换域的高鲁棒多图像加密算法

王丰, 邵珠宏, 王云飞, 姚启钧, 刘西林 1366

基于透射率修正的湍流模型与动态调整retinex的水下图像增强

汤新雨, 李敏, 徐灵丽, 郝真, 张学武 1380

图像分析和识别

俯视深度头肩序列行人再识别

王新年, 刘春华, 齐国清, 张世强 1393

结合混合池化的双流人脸活体检测网络

汪亚航, 宋晓宁, 吴小俊 1408

增强型灰度图像空间实现虹膜活体检测

刘明康, 王宏民, 李琦, 孙哲南 1421

基于水动力学、分数阶微分及Steger算法的线状目标提取

陈卫卫, 王卫星, 李润青, 蔡晓琰, 郑思凡 1436

时域候选优化的时序动作检测

熊成鑫, 郭丹, 刘学亮 1447

图像理解和计算机视觉

视频中多特征融合人体姿态跟踪

马淼, 李贻斌, 武宪青, 高金凤, 潘海鹏 1459

计算机图形学

插值给定数据点的四次PH曲线构造

方林聪, 阳诚砖, 邸文钰, 刘芳 1473

医学图像处理

图卷积网络下牙齿种子点自动选取

李占利, 孙志浩, 李洪安, 刘童鑫 1481

乳腺超声图像中易混淆困难样本的分类方法

杜章锦, 龚勋, 罗俊, 章哲敏, 杨菲 1490

遥感图像处理

迁移学习下高分快视数据道路快速提取

张军军, 万广通, 张洪群, 李山山, 冯旭祥 1501

CONTENTS

JOURNAL OF IMAGE AND GRAPHICS



Research status and progress of non-photorealistic rendering technology(P1283)



Automatic selection of tooth seed point by graph convolutional network(P1481)



Rapid road extraction from quick view imagery of high-resolution satellites combined with transfer learning(P1501)

Scholar View

Research status and progress of non-photorealistic rendering technology

Qian Wenhua, Cao Jinde, Xu Dan, Wu Hao 1283

Review

Deep Hashing image retrieval methods

Liu Ying, Cheng Mei, Wang Fuping, Li Daxiang, Liu Wei, Fan Jiulun 1296

Virtual character behavior modeling in serious games: a review

Liu Cuijuan, Liu Zhen, Chai Yanjie, Liu Tingting, Chen Xiaoyi 1318

Bottleneck issues of computer vision in intelligent manufacturing

Lei Linjian, Sun Shengli, Xiang Yukai, Zhang Yue, Liu Huikai 1330

Image Processing and Coding

A shallow CNN-based noise ratio estimation model

Xu Shaoping, Lin Zhenyu, Li Chongxi, Cui Yan, Liu Ruirui 1344

Degradation model and restoration of flame image of sintering section

Wang Fubin, Liu Hefei, Wu Chen 1356

Multiple image encryption of high robustness in gyrator transform domain

Wang Feng, Shao Zhuhong, Wang Yunfei, Yao Qijun, Liu Xilin 1366

Underwater image enhancement based on turbulence model corrected by transmittance and dynamically adjusted retinex

Tang Xinyu, Li Min, Xu Lingli, Hao Zhen, Zhang Xuewu 1380

Image Analysis and Recognition

Person re-identification based on top-view depth head and shoulder sequence

Wang Xinnian, Liu Chunhua, Qi Guoqing, Zhang Shiqiang 1393

Two-stream face spoofing detection network combined with hybrid pooling

Wang Yahang, Song Xiaoning, Wu Xiaojun 1408

Enhanced gray-level image space for iris liveness detection

Liu Mingfang, Wang Hongmin, Li Qi, Sun Zhenan 1421

Linear object extraction based on hydrodynamics, fractional differential and Steger algorithm

Chen Weiwei, Wang Weixing, Li Runqing, Cai Xiaoyan, Zheng Sifan 1436

Temporal proposal optimization for temporal action detection

Xiong Chengxin, Guo Dan, Liu Xueliang 1447

Image Understanding and Computer Vision

Human pose tracking based on multi-feature fusion in videos

Ma Miao, Li Yibin, Wu Xianqing, Gao Jinfeng, Pan Haipeng 1459

Computer Graphics

Construction of quartic PH curves via interpolating given points

Fang Lincong, Yang Chengzhuan, Di Wenyu, Liu Fang 1473

Medical Image Processing

Automatic selection of tooth seed point by graph convolutional network

Li Zhanli, Sun Zhihao, Li Hongan, Liu Tongxin 1481

Classification method for samples that are easy to be confused in breast ultrasound images

Du Zhangjin, Gong Xun, Luo Jun, Zhang Zheming, Yang Fei 1490

Remote Sensing Image Processing

Rapid road extraction from quick view imagery of high-resolution satellites with transfer learning

Zhang Junjun, Wan Guangtong, Zhang Hongqun, Li Shanshan, Feng Xuxiang 1501

中图法分类号: TP391 文献标识码: A 文章编号: 1006-8961(2020)07-1344-12

论文引用格式: Xu S P, Lin Z Y, Li C X, Cui Y and Liu R R. 2020. A shallow CNN-based noise ratio estimation model. Journal of Image and Graphics, 25(07): 1344-1355 (徐少平, 林珍玉, 李崇禧, 崔燕, 刘蕊蕊. 2020. 浅层 CNN 网络构建的噪声比例估计模型. 中国图象图形学报, 25(07): 1344-1355) [DOI:10.11834/jig.190472]

浅层 CNN 网络构建的噪声比例估计模型

徐少平, 林珍玉, 李崇禧, 崔燕, 刘蕊蕊

南昌大学信息工程学院, 南昌 330031

摘要: **目的** 利用深度卷积神经网络 (deep convolutional neural network, DCNN) 构建的非开关型随机脉冲噪声 (random-valued impulse noise, RVIN) 降噪模型在降噪效果和执行效率上均比主流的开关型 RVIN 降噪算法更有优势, 但在实际应用中, 这类基于训练 (数据驱动) 的降噪模型, 其性能却受制于能否对待降噪图像受噪声干扰的严重程度进行准确的测定 (即存在数据依赖问题)。为此, 提出了一种基于浅层卷积神经网络的快速 RVIN 噪声比例预测 (noise ratio estimation, NRE) 模型。**方法** 该预测模型的主要任务是检测待降噪图像中的噪声比例值并将其作为反映图像受噪声干扰严重程度的指标, 依据 NRE 预测模型的检测结果可以自适应调用相应预先训练好的特定区间 DCNN 降噪模型, 从而快速且高质量地完成图像降噪任务。**结果** 分别在 10 幅常用图像和 50 幅纹理图像两个测试集上进行测试, 并与现有的主流 RVIN 降噪算法中的检测模块进行对比。在常用图像测试集上, 本文所提出的 NRE 预测模型的预测准确性最高。相比于噪声比例预测精度排名第 2 的算法, NRE 预测模型在噪声比例预测值均方根误差上低 0.6% ~ 2.4%。在 50 幅纹理图像测试集上, NRE 模型的均方根误差波动范围最小, 表明其稳定性最好。通过在 1 幅大小为 512 × 512 像素图像上的总体平均执行时间来比较各个算法执行效率的优劣, NRE 模型执行时间仅为 0.02 s。实验数据表明: 所提出的 NRE 预测模型在受各种不同噪声比例干扰的自然图像上均可以快速而稳定地测定图像中受 RVIN 噪声干扰的严重程度, 非盲的 DCNN 降噪模型与其联用后即可无缝地转化为盲降噪算法。**结论** 本文 RVIN 噪声比例预测模型在各个噪声比例下具有鲁棒的预测准确性, 与基于 DCNN 的非开关型 RVIN 深度降噪模型配合使用能妥善解决 DCNN 网络模型固有的数据依赖问题。

关键词: 随机脉冲噪声; 噪声比例估计; 浅层卷积神经网络; 非逐点模式; 执行效率; 盲降噪

A shallow CNN-based noise ratio estimation model

Xu Shaoping, Lin Zhenyu, Li Chongxi, Cui Yan, Liu Ruirui

School of Information Engineering, Nanchang University, Nanchang 330031, China

Abstract: **Objective** Nonswitching random-valued impulse noise (RVIN) denoisers built with deep convolution neural networks (DCNNs) have many advantages compared with the mainstream switching RVIN removal algorithms in terms of denoising effect and execution efficiency. However, the performance of training-based (data-driven) denoisers in practical applications experiences inaccurate measurement of the distortion level of a given image to be denoised (data dependency problem). A fast noise ratio estimation (NRE) model based on shallow CNN (SCNN) was proposed in this study.

Method The noise ratio reflecting the distortion level of a given noisy image was estimated using the proposed NRE model. On the basis of the estimated noise ratio, the corresponding DCNN-based denoiser trained at a specific interval of noise ratio

收稿日期: 2019-09-16; 修回日期: 2019-12-14; 预印本日期: 2019-12-21

基金项目: 国家自然科学基金项目 (61662044, 61163023, 51765042); 江西省自然科学基金项目 (20171BAB202017)

Supported by: National Natural Science Foundation of China (61662044, 61163023, 51765042); Natural Science Foundation of Jiangxi Province, China (20171BAB202017)

os can be adaptively exploited to efficiently remove RVIN with high denoised image quality. **Result** Comparison experiments were conducted to test the validity of the proposed NRE model from three aspects, namely, estimation accuracy, denoising effect, and execution efficiency. We utilized the NRE model to estimate the noise ratios of given noisy images and compared the results with the existing classical RVIN noise reduction algorithms (including PSMF (progressive switching median filter), ROLD-EPR (rank-ordered logarithmic difference edge-preserving regularization), ASWM (adaptive switching median), ROR-NLM (robust outlyingness ratio nonlocal means), MLP-EPR (multilayer perceptron edge-preserving regularization)) to verify its estimation accuracy. Considering that these competing algorithms detect noisy pixels in a pixelwise manner, the number of pixels identified as noise was divided by the total number of pixels in the entire image, and the detection results were transformed as noise ratio to facilitate comparison with the proposed NRE model. Two image sets were used, where the first image set included Lena, House, Peppers, Couple, Hill, Barbara, Boat, Man, Cameraman, and Monarch images, and the second image set was randomly selected from the Business Structure Database. For a fair comparison, all the competing algorithms were implemented on MATLAB 2017b and conducted on the same hardware platform. First, noise ratio detection was performed on the first image set, and Lena, Boat, and House images corrupted with different noise ratios were selected for demonstration. Regarding the estimation accuracy of noise ratio, PSMF, ROLD-EPR, and MLP-EPR algorithms are low. The estimation accuracy of the ASWM algorithm is high at high noise ratios, and the prediction accuracy of the ROR-NLM algorithm is high at medium and low noise ratios. The performance of the proposed NRE model consistently ranks in top two. The root mean square error (RMSE) between the estimation results and the ground-truth noise ratios was used to verify the stability of the proposed NRE model. On the first image set, the NRE model outperforms the second-best algorithm by 0.6% - 2.4% in terms of RMSE. In the second image set, the RMSE of the proposed NRE model on 50 images is the smallest, indicating that its stability is the best. Although the estimation accuracies of several switching RVIN removal algorithms are higher than the proposed NRE prediction model in some cases, their execution time is extremely long. We applied different ratios of RVIN noise (10%, 20%, 30%, 40%, 50% and 60%) to Lena image with size of 512×512 pixels to test the denoising effect and used the pretrained NRE model to estimate the noise ratios for the noisy images. A DnCNN-S (DnCNN for specific ratio range) noise reduction model was used to remove the noise in accordance with the estimation results. Experimental results show that the denoising results with the estimated noise ratios are similar to the ground truth values, indicating that the estimation accuracy of the proposed NRE model is satisfactory. As the preprocessing module of denoising algorithms, the execution efficiency of noise detector is used as an evaluation index that determines the execution performance of the entire denoising algorithm. Only the MLP-EPR algorithm and the NRE + DnCNN-S denoising algorithm proposed in this paper can be clearly divided into two stages, namely, noise detection and noise reduction. Therefore, we provided the execution time of the noise detection module for MLP-EPR and NRE + DnCNN-S algorithms in addition to the average execution time. Regarding the execution time of the noise detection module, the MLP-EPR algorithm needs 0.07 s, whereas the proposed NRE model only needs 0.02 s, indicating that the efficiency of the NRE model is relatively high. Experimental results show that the proposed SCNN-based NRE model can quickly and stably measure the distortion level of a given noisy image corrupted by RVIN under different noise ratios. Any nonblind DCNN-based denoiser combined with the proposed NRE model can be seamlessly converted into a blind version. **Conclusion** Extensive experiments show that the estimation accuracy of the proposed CNN-based NRE model is robust under a wide range of noise ratios. DCNN's inherent data dependence problem can be properly solved when the proposed NRE model is combined with DCNN-based non-switching RVIN deep denoising models.

Key words: random-valued impulse noise (RVIN); noise ratio estimation (NRE); shallow convolutional neural network (SCNN); non-pointwise mode; execution efficiency; blind denoising

0 引言

数字图像在获取、存储或者在无线通信信道上

传输过程中极易受到脉冲噪声 (impulse noise, IN) 的干扰 (Chen 和 Wu, 2001; Jin 和 Ye, 2018; Liu 等, 2015; Turkmen, 2016; Xu 等, 2018)。通常脉冲噪声可以分为两类: 固定脉冲噪声 (fixed-valued impulse

noise, FVIN) 和随机脉冲噪声(random-valued impulse noise, RVIN)(Dong 等, 2007; Garnett 等, 2005)。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, 很难区分。因此, RVIN 噪声的检测与去除更具挑战性(Chen 和 Wu, 2001; Dong 等, 2007; Garnett 等, 2005), 本文专注于 RVIN 降噪问题。

早期的降噪算法主要以中值滤波器(median filter, MF)及其衍生算法(Chen 和 Wu, 2001)为主, 通过将局部窗口内中值滤波器(或其衍生滤波器)的输出值替换为中心像素点的亮度值, 简单快速地实现图像的复原。这类算法的计算效率很高, 但完全不考虑中心像素点是否为噪声, 无差别地对图像所有像素点进行处理, 容易导致复原图像的纹理细节模糊(尤其是在 RVIN 噪声比例较高的情况下)。为此, Garnett 等人(2005)以中心像素点与局部窗口其他像素点之间亮度值的前 m 个最小绝对差值累积和作为 ROAD(rank-ordered absolute differences)统计值, 并将其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。针对检测出的噪声点, 使用一个三边滤波器(trilateral filter)对其进行降噪。为了扩大噪声点和正常像素点之间的差异, Dong 等人(2007)对 ROAD 统计值作对数变换提出了 ROLD(rank-ordered logarithmic difference)统计特征值, 并利用边缘保持正则化(edge-preserving regularization, EPR)方法对判断为噪声的像素点进行复原。该算法在处理中低比例噪声时比较有效, 但在高比例噪声下, 降噪效果仍然有待提高。为了进一步提高噪声检测正确率, Liu 等人(2015)在 ROAD 检测器的基础上提出了 ROD-ROAD(rank-ordered difference of ROAD)统计特征值来检测图像中的噪声点, 然后使用 MEPD(minimum edge pixels difference)局部统计值将正常的边缘像素点从检测出的噪声点中剔除, 最后用加权均值滤波器对最终判断为噪声的像素点降噪。近期, Turkmen(2016)将像素点的 ROAD 和 ROLD 统计值作为多层感知网络(multilayer perceptron, MLP)的输入, 对应的噪声标签(noise label)作为输出训练了一个噪声检测器。该噪声检测模型能够直接判定图像中所有的像素点是否为噪声, 无需与某种降噪算法绑定迭代执行即可高效率地完成噪声检测任务, 但噪声像素点的检

测正确率仍有待提高, 影响到后续降噪算法的复原效果。总之, 现有的各类开关型 RVIN 噪声检测和降噪算法所采用的逐点检测与复原方式的降噪策略, 其在降噪效果和执行效率方面的性能严重受制于 RVIN 噪声点检测模块, 现有技术目前还很难实现对图像中所有的 RVIN 噪声点进行准确的定位与检测, 漏检(miss detection)和错检(false detection)在所难免, 而后继的降噪模块则完全基于 RVIN 噪声点检测模块的检测结果决定是否启动图像复原过程, 故降噪效果受到较大限制。

由卷积(convolution)、修正线性单元(rectifier linear unit, ReLU)(Rawat 和 Wang, 2017)、池化(pooling)等操作构成的深度卷积神经网络(deep convolutional neural network, DCNN)(常亮 等, 2016; 时永刚 等, 2018; 周锦峰 等, 2019)能很好地模拟人眼视觉特性, 具有强大的非线性映射能力。研究者们将深度卷积神经网络结构引入图像降噪领域, DnCNN(denoising convolutional neural networks)(Zhang 等, 2017)、FFDNet(fast and flexible denoising convolutional neural network)(Zhang 等, 2018)、UD-Net(universal denoising networks)(Lefkimmatis, 2018)、CBD-Net(convolutional blind denoising network)(Guo 等, 2019)、VST-Net(variance-stabilizing transformation inspired network)(Zhang 等, 2019)等基于 CNN 技术设计的降噪模型获得了比传统方法更好的性能。以用于去除高斯噪声(Gaussian noise, GN)的 DnCNN 网络为例, 在降噪效果方面, DnCNN 降噪网络所采用的残差网络结构使得网络收敛比较容易; 执行效率方面, DnCNN 降噪网络中最耗费计算资源的卷积计算在 GPU(graphics processing unit)硬件支撑下具有极高的执行效率。较各类深度降噪模型, DnCNN 降噪模型在降噪效果和执行效率两方面的综合性能比较好, 常被用来作为对比算法。其实, 通过改变训练数据对 TP(training pair)中的输入图像(让它们受 RVIN 噪声干扰), DnCNN 降噪模型可以扩展用于处理 RVIN 噪声。与主流的开关型 RVIN 降噪算法相比, 这种基于 DCNN 技术构建的端到端(end to end)非开关型 RVIN 降噪模型理论上可获得更佳的降噪效果和执行效率。然而, 由于基于深度卷积神经网络构建的非开关型降噪模型采用的基于训练(数据驱动)的实现策略, 存在数据依赖的问题, 只有当训练 DCNN

模型所使用的图像集合与待降噪图像受到噪声干扰严重程度类似时才能获得最佳的降噪效果。传统的开关型算法主要是先利用噪声检测器判断图像每一个像素点是否为噪声,然后针对噪声点进行复原。由于 RVIN 噪声的随机性,噪声图像对应的噪声标签并不适用于在大量数据集上训练降噪模型。事实上,噪声比例(noise ratio)大小是衡量噪声图像受噪声干扰严重程度的简单而有效的指标,即受 RVIN 噪声干扰的像素点的个数占图像中所有像素点的比例值。因此,为将 DCNN 技术引入到 RVIN 降噪算法中来,构建噪声比例预测(noise ratio estimation, NRE)模型成为需要解决的首要问题。

本文提出了一种基于浅层卷积神经网络(shallow convolutional neural network, SCNN)的 NRE 噪声比例预测模型。该算法以噪声比例值作为衡量图像受噪声干扰程度的预测指标,通过构建 SCNN 网络并在大量训练集上经过训练后获得 NRE 预测模型。将所提出的 NRE 预测模型与 DnCNN 深度降噪模型结合使用,所构建的非开关型降噪算法可以根据待降噪图像受 RVIN 噪声干扰严重程度自适应调用相应预先训练的特定区间深度降噪模型(训练图像集中的噪声比例值变化范围限制在某个较窄的范围内),从而高效率地完成高质量的降噪效果。NRE 预测模型可将现有的非盲的 DnCNN 深度降噪模型无缝地转换为盲降噪模型,提高了它的实用性,确保在降噪效果和执行效率上超过现有主流的开关

型 RVIN 降噪算法。

1 DnCNN 算法

Zhang 等人(2017)对前馈(feed-forward)卷积神经网络(即 VGG(visual geometry group)网络(Simonyan 和 Zisserman, 2019))的结构和深度进行了调整,设计了一个适用于图像降噪任务的 DnCNN 去噪卷积神经网络模型。为了获得与输入图像大小一致的输出, DnCNN 降噪模型删除了 VGG 模型中的池化层,仅使用卷积层对输入图像进行处理,并结合 ReLU(rectifier linear unit)激活函数、残差学习(residual learning)以及批归一化(batch normalization, BN)(Ioffe 和 Szegedy, 2019)技术来进一步加快网络的训练速度并提高模型的降噪效果。如图 1 所示, DnCNN 整个网络的隐含层共有 d 层:第 1 层为卷积层加激活函数(Conv + ReLU),采用 64 个 $3 \times 3 \times 1$ 的滤波器生成 64 个特征映射图;第 $2 \sim (d-1)$ 层则在卷积层与 ReLU 之间增加了 BN 操作(Conv + ReLU + BN),采用的是大小为 $3 \times 3 \times 64$ 的滤波器;最后一层通过 1 个单独的卷积操作得到输出结果。与通过学习映射函数 $\varphi(y) = x$ 直接将输入噪声图像 $y = x + v$ 映射为无失真图像 x 的普通方式不同, DnCNN 模型以噪声图像 y 与原始图像 x 之间的残差映射图(residual mapping) $v = y - x$ 作为学习的目标值。对 N 幅原始无失真图像添加噪声,以噪声图像 y 及其对应的残差图像 v 分别作为网络模型的

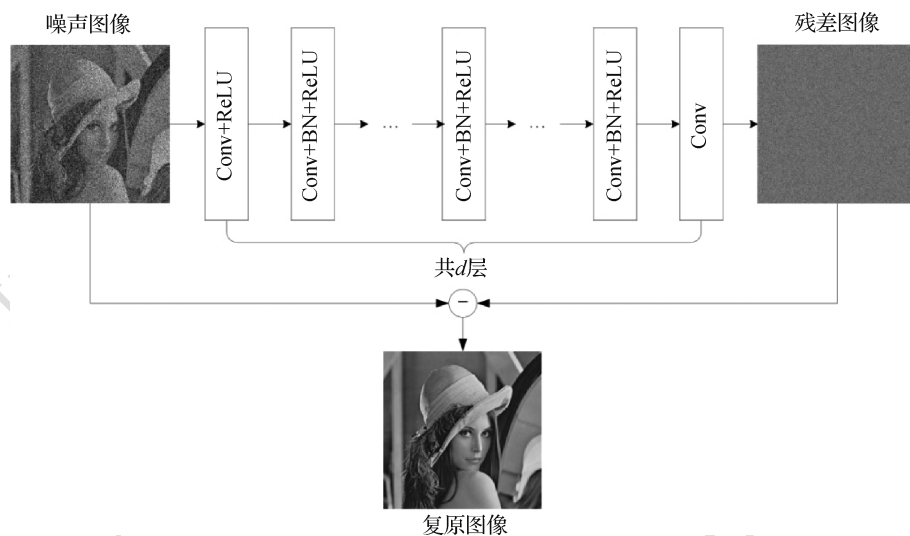


图 1 DnCNN 降噪模型的网络架构

Fig. 1 The network architecture of DnCNN denoising model

输入和输出对进行训练,利用残差学习规则学习残差映射使得 $R(y; \Theta) \approx v$ 。在训练模型的过程中,主要是通过控制真实残差图像 v 与估计残差图像 $R(y; \Theta)$ 之间的平均均方误差

$$\ell(\Theta) = \frac{1}{2N} \sum_{i=1}^N \|R(y_i; \Theta) - (y_i - x_i)\|_F^2 \quad (1)$$

来调整网络模型参数 Θ 。式中, y_i 和 x_i 分别表示第 i ($i = 1, 2, \dots, N$) 幅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 y 减去网络模型所预测的残差映射图 $R(y)$ 即可得到关于原始图像 x 的估计图像 $\hat{x} = y - R(y)$ 。DnCNN 降噪模型原本是针对高斯噪声设计的,其网络架构其实还可用于处理更多噪声类型的降噪问题,将 TP 对中的高斯噪声图像换成 RVIN 噪声图像,即可以通过训练获得可以完成 RVIN 降噪的深度 DnCNN 降噪模型。在实现上,可以让训练图像集中噪声图像受到 RVIN 噪声干扰比例值均衡分布在 $[0, 60\%]$ 范围内,这样只需要训练一个降噪模型用于降噪即可。但是这种方法所获得的降噪模型的降噪效果不佳,DnCNN 降噪模型目前的网络结构从大量训练数据中学习先验知识的能力还有一定局限性,存在数据依赖的缺陷。即用于训练模型的图像集合受噪声干扰的严重程度必须与待降噪图像差别不大,降噪效果才会比较好。

2 NRE 预测模型

2.1 存在问题

如图 2 所示,对于一幅受 30% RVIN 噪声干扰的 Lena 噪声图像来说,当使用在 30% 比例训练图像集上训练的 DnCNN 模型进行降噪,所获得降噪后图像的峰值信噪比(peak signal to noise ratio, PSNR)值为 36.73 dB,而使用 20% 和 40% 噪声图像集训练的 DnCNN 降噪模型时的 PSNR 值仅能达到 33.79 dB 和 35.66 dB。从图 2 中各个子图的局部放大图来看,图 2(c)与原图最接近(测试图像中的噪声比例值 p_{test} 与降噪模型训练图像集合中所添加的噪声比例值 p_{train} 相等),而图 2(d)中存在残留的噪声没有处理(使用了比噪声图像中噪声比例值更低的降噪模型),图 2(e)则将图像中的细节模糊化了(使用了比噪声图像中噪声比例值更高的降噪模型)。这表明 p_{test} 和 p_{train} 值不匹配时,降噪效果会存在不同程度的下降。其本质原因在于:DnCNN 降噪模型是基于数据驱动的,必须保证训练图像集中的图像受 RVIN 噪声干扰的严重程度与待降噪图像近似才可以获得最佳降噪效果,存在数据依赖问题。因此,DnCNN 降噪模型在实际使用中必须在 NRE 预测模型的支撑下才能有效工作。

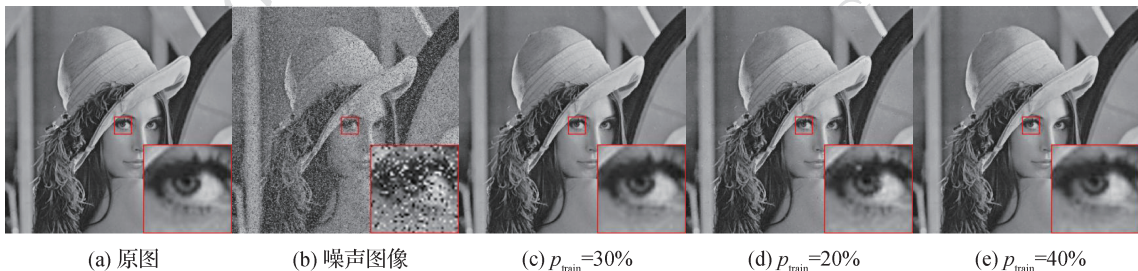


图 2 用不同 p_{train} 训练的 DnCNN 降噪模型复原噪声图像($p_{\text{test}} = 30\%$)的降噪效果对比

Fig. 2 Denoising effect comparison of DnCNN denoising models trained with different p_{train} for noisy image corrupted with $p_{\text{test}} = 30\%$

((a) clean image; (b) noisy image; (c) $p_{\text{train}} = 30\%$; (d) $p_{\text{train}} = 20\%$; (e) $p_{\text{train}} = 40\%$)

2.2 解决方案

为实现一个预测正确性和执行效率俱佳的 NRE 预测模型,本文在 CNN 的技术框架下,通过构建一个 SCNN 网络并以大量噪声图像及对应添加的 RVIN 噪声比例值作为该 SCNN 网络的 TP 训练对,完成训练后即可获得一个 NRE 预测模型。本文在图块级上实现对图像噪声比例的预测。具体地,首先通过随机采样的方式从图像中抽取若干个不同位

置的样本图块,然后用所训练 NRE 模型预测出各个样本图块的噪声比例值,最后将所有采样图块预测结果的均值作为整幅图像的噪声水平估计值。

2.3 网络结构

如图 3 所示,文中所构建的 NRE 噪声比例预测模型的输入为 $d \times d$ 大小的噪声图块,经过多个卷积层(每一卷积层中均使用 $3 \times 3 \times 16$ 的卷积核生成中间特征映射图)与 ReLU 激活函数提取特征并结合

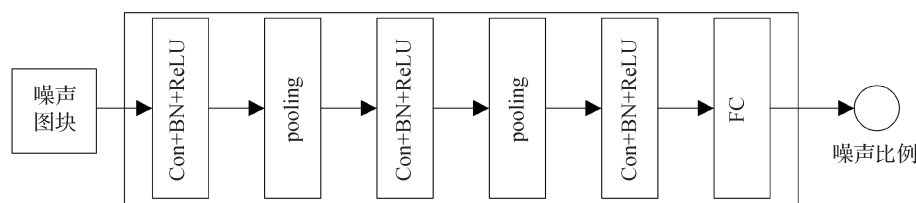


图3 基于浅层 CNN 网络的 NRE 预测模型

Fig. 3 The NRE prediction model based on shallow CNN network

池化(pooling)和全连接(fully connected layers, FC)层降维后得到最终的噪声比例值。理论上所使用的卷积层数越多,网络的映射能力越强,但同时执行时间也将显著增加。由表1可知:所提出的 NRE 噪声比例预测模型(使用3层卷积)在预测值与真实值的均方根误差(root mean squared error, RMSE)平均

值已经控制在1%以内,完全满足调用 DnCNN 模型的精度要求。而所构建的 NRE 噪声比例预测模型作为 DnCNN 随机脉冲降噪模型的前置支撑模块,其执行效率是优先考虑因素,以避免拖累降噪算法的整体性能。故综合考虑预测模型的预测准确性和执行效率,本文使用3个卷积层构建 NRE 模型。

表1 不同训练图块数下 NRE 预测模型预测结果与真值之间的 RMSE 值比较

Table 1 RMSE comparison between NRE models and ground-truths with different number of training patches

图像数/幅	噪声比例真值/%						均值
	10	20	30	40	50	60	
20	0.63	0.96	1.17	1.28	0.97	2.70	1.19
40	0.62	0.71	0.89	1.02	1.01	2.60	1.04
60	0.49	0.58	0.75	0.99	0.88	2.60	0.96
80	0.60	0.73	0.85	1.14	0.94	2.23	1.01
100	0.62	0.93	1.18	1.61	1.63	0.79	1.05
120	0.52	0.56	0.72	0.85	0.81	2.01	0.85
140	0.96	1.08	1.26	1.41	1.68	1.69	1.28
160	0.31	0.51	0.64	0.84	0.91	1.21	0.85

注:加粗字体表示最优结果。

为训练 NRE 预测模型,首先对大量原始无失真图像施加不同比例的随机脉冲噪声,在每幅噪声图像上随机提取若干图块及其噪声比例值(同一幅图像上所有图块的噪声比例值相同)构成训练集合。具体地,从 BSD(Berkeley segmentation data)数据库(Arbeláez 等,2011)中选取若干幅原始无失真图像,对每幅图像施加各个级别(范围为0~60%,间隔1%)的随机脉冲噪声,然后从每幅噪声图像中随机提取200个大小为 $d \times d$ 的图块。所提取的图块及其对应的噪声比例值构成训练的 TP 对,分别作为预测模型的输入和输出。一旦基于卷积神经网络的通用噪声比例检测模型训练好后,即可用于检测任意给定图块中的噪声比例。

使用多少个训练图块训练 NRE 预测模型会对其预测精度产生一定的影响,为测试训练图块个数对噪声比例检测正确率的影响并确定最佳的训练图像数量(每幅图像中提取200个图块),在不同训练图像数量下训练多个 NRE 模型后,利用这些预测模型对随机选取的100幅 BSD 纹理图像(与训练集中图像不同),在各个噪声比例下进行预测,计算各个 NRE 模型预测结果与真实噪声比例值之间的 RMSE 值以及均值,结果列在表1中。由表1可知:随着训练图块个数不断增加,NRE 模型的预测准确性总体呈上升趋势(RMSE 均值趋于变小)。其中,在训练图像数为120时,所获得的 RMSE 均值为最低值。虽然图像数量为160时,预测精度能达到最小值,鉴

于 120 幅图像(相当于 1 464 000 个图块)上训练的 NRE 预测模型已达到预测性能最佳状态,本文最终选定在 120 幅图像上训练模型。

2.4 参数设置

对于工作在图块级的 NRE 预测模型来说,输入图块的大小是非常重要的参数。为了测试图块大小对噪声比例检测正确率的影响并确定该参数的最佳值,在不同图块大小参数配置下训练多个

预测模型,并且利用得到的预测模型对从 BSD 数据库中随机选取不同于训练图像的 100 幅图像进行测试,计算不同图块大小配置下模型预测结果与真实噪声比例值之间的均方根误差,结果见表 2。由表 2 可知,在各个噪声比例下,28 × 28 像素大小的图块所训练的 NRE 模型在各个噪声比例下所获得预测结果的 RMSE 指标值最优次数最多,且 RMSE 均值最低,故本文将图块大小设置为 28 × 28 像素。

表 2 在不同图块大小配置下 NRE 预测模型的 RMSE 值比较
Table 2 RMSE comparison of the NRE models with different patch sizes

图块大小/像素	噪声比例/%							平均值
	5	10	20	30	40	50	60	
20 × 20	3.81	2.07	1.44	1.36	1.58	1.44	4.34	2.29
28 × 28	2.45	0.64	1.15	1.24	1.46	1.59	2.44	1.57
30 × 30	2.20	0.76	1.46	1.52	1.83	2.07	5.18	2.15
40 × 40	2.23	0.96	1.21	1.53	1.57	1.60	3.75	1.84

注:加粗字体表示每列最优结果。

由于自然图像中的内容一般都较为复杂,其实很难依靠从图像中提取的单个图块准确地预测出整幅图像的噪声比例值,不可避免地存在过估计和欠估计的情况。为此,随机从图像中取若干个图块并对它们进行噪声比例预测,以所有图块噪声比例预测值的均值作为整幅图像噪声比例的最终预测值。在这种预测的策略下,选用图块个数的多少对最终预测结果具有一定影响。为了测试图块个数对噪声比例检测正确率的影响并确定该参数的最佳值,随机选

取 100 幅 BSD 图像(与训练集中图像不同)作为测试图像,在各个噪声比例下计算不同图块个数配置下 NRE 模型预测结果与真实噪声比例值之间的 RMSE 均值,结果见表 3。由表 3 可知:随着图块个数的增加,NRE 模型的预测准确率会相应提升,但其提升的幅度越来越小;而随着图块个数的增加,NRE 模型的执行时间却会急剧增加,影响后续的图像降噪模型的性能。因此,综合考虑 NRE 检测器预测结果的准确率和执行效率,本文将图块个数参数设置为 20。

表 3 在不同图块个数配置下 NRE 预测模型的 RMSE 值比较
Table 3 RMSE comparison of the NRE models with different number of patches

图块个数	噪声比例/%						均值	平均时间/ms
	10	20	30	40	50	60		
5	0.83	1.24	1.87	2.06	2.26	1.32	1.48	4.88
10	0.79	1.01	1.42	1.86	1.97	0.93	1.23	7.38
15	0.70	0.98	1.30	1.79	1.67	0.89	1.13	9.81
20	0.62	1.05	1.34	1.56	1.62	0.78	1.09	12.57
25	0.67	0.86	1.20	1.53	1.70	0.83	1.05	16.16
30	0.61	0.86	1.16	1.54	1.74	0.64	1.02	18.67

注:加粗字体表示每列最优结果。

2.5 优点分析

所提出的 NRE 预测模型具有以下优点:1)执行

效率高。在 SCNN 网络上实现的噪声比例预测模型工作在图块级上,不但可以加快网络的训练速

度,而且在执行噪声比例检测任务时效率也高,相对于深度的 DnCNN 主降噪模块来说,执行时间占比很少。2) 预测准确率高。NRE 预测模型虽然为浅层 CNN 网络,但各层总共仍有 64 个卷积核,可以提取大量的特征映射图,而 pooling 和 FC 层则可以将这些特征值映射图准确性映射为噪声比例值。其建模能力远超开关型 LIS (local image statistic) 统计值方法,避免了开关型检测方法中阈值设置困难的问题。3) 鲁棒性好。所提出的噪声比例值估计方法采用了随机采样方法,可以减弱单个图块预测结果偏高或者偏低的情况,更好地减少图像背景内容对预测结果的影响,使得预测结果更为鲁棒。

3 实验结果及其分析

3.1 测试环境和方法介绍

为测试所提出 NRE 预测模型的性能,使用 NRE 模型对施加了不同噪声比例 RVIN 噪声的噪声图像进行噪声比例值的预测,并与现有的主流 RVIN 降噪算法(包括 PSMF (progressive switching median filter) (Wang 和 Zhang, 1999)、ROLD-EPR (rank-ordered logarithmic difference edge-preserving regularization) (Dong 等, 2007)、ASWM (adaptive switching median) (Akkoul 等, 2010)、ROR-NLM (robust outlyingness ratio nonlocal means) (Xiong 和 Yin, 2012) 和

MLP-EPR (multilayer perceptron edge-preserving regularization) (Turkmen, 2016)) 中的检测模块进行对比。由于这些对比算法都是基于逐点检测像素点的方式实现的,故将各算法检测出被认定为噪声像素点的个数除以整幅图像的像素点总数,从而将检测结果转化为噪声比例值后再比较。为测试 NRE 预测模型结合 DnCNN 降噪模型后实际的降噪效果,将 $[0, 60\%]$ 噪声比例范围每间隔 5% 划分为一个子区间范围,共训练 12 个区间 DnCNN 模型。对于某个特定区间 DnCNN 模型 (DnCNN for specific ratio range, DnCNN-S), 选择大量无失真图像对其添加特定噪声比例区间范围的 RVIN 噪声并与相应的无失真图像构成 TP 训练对,添加的 RVIN 噪声比例值在区间范围内从最小值 $p_{\min}$ 到最大值 $p_{\max}$ 以 0.5% 递增。为保证在各个区间邻接处的预测准确性,噪声比例实际训练范围为 $[p_{\min} - 1\%, p_{\max} + 1\%]$ 。本文训练的 DnCNN 区间降噪模型参数设置与文献 (Zhang 等, 2017) 相同。实验的硬件平台为: Intel (R) Xeon (R) CPU E5-1603 v4 @ 2.80 GHz (NVIDIA Quadro M4000 GPU) 16 GB RAM; 软件环境为: Windows 10 操作系统, MATLAB R2017b 编程环境。

3.2 噪声比例检测比较

首先,在如图 4 常用的图像集合上进行实验,数据列在表 4 中(限于篇幅,仅给出 3 幅图像在 10%、30%、50% 和 60% 的 RVIN 噪声干扰下的实验数

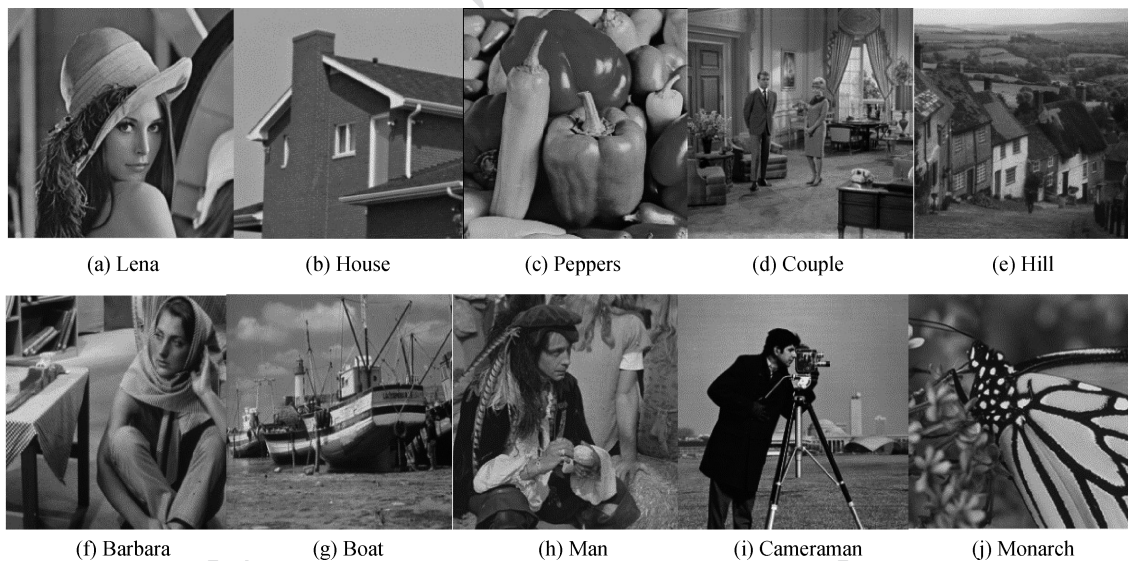


图4 各类文献中常用的图像集合

Fig.4 Commonly used images in the literature

((a) Lena; (b) House; (c) Peppers; (d) Couple; (e) Hill; (f) Barbara; (g) Boat; (h) Man; (i) Cameraman; (j) Monarch)

表 4 各个降噪算法在添加不同比例噪声的 3 幅图像上的预测结果与真值之间的对比

Table 4 Comparison between estimated results and ground truths of different methods on three images with different noise ratios

算法	噪声比例/%											
	10			30			50			60		
	Lena	Boat	House	Lena	Boat	House	Lena	Boat	House	Lena	Boat	House
ASWM	8.27	5.03	6.18	29.67	27.12	27.83	51.50	50.12	50.23	63.53	62.47	62.45
ROR-NLM	10.77	13.48	11.42	28.06	29.56	29.05	44.26	45.21	44.76	52.31	53.16	52.35
PSMF	7.22	9.41	7.47	21.18	22.85	21.40	33.88	35.51	34.30	40.23	41.66	39.97
ROLD-EPR	8.12	7.59	8.11	25.77	24.69	26.33	44.27	42.89	44.70	54.18	52.48	53.38
MLP-EPR	7.95	7.67	7.90	25.77	25.91	26.19	45.45	46.90	46.11	55.60	57.36	55.76
NRE(本文)	10.78	11.01	11.03	30.53	30.57	30.14	49.84	52.10	49.24	58.00	57.75	58.04

注:加粗字体表示预测结果接近真实噪声比例的前两名。

据,其他图像的实验结果类似)。在各噪声比例下,预测结果最接近真实噪声比例值的前两个算法加粗显示。从表 4 中可以看出,对于图像噪声比例值的预测准确性来说,PSMF 算法、ROLD-EPR 算法以及 MLP-EPR 算法较低,ASWM 算法在高噪声水平条件下的预测准确度比较高,而 ROR-NLM 算法在中、低噪声水平下表现较好。仅有本文 NRE 检测器在各

个噪声比例上都进入了前 2 名,很多情况下离噪声比例真值最接近。

为了进一步测试所提出的 NRE 噪声比例检测器的稳定性,分别在常用测试图像和纹理图像测试集合(从 BSD 数据库中随机挑选的不同于训练图像的 50 幅图像)进行测试,并计算预测结果与真实值的均方根误差,将实验结果数据列在表 5 和表 6 中。

表 5 各噪声检测器在不同噪声水平下的 10 幅常用图像下预测情况的 RMSE 值比较

Table 5 RMSE comparison on ten commonly used images with different noise ratios

算法	噪声比例/%					
	10	20	30	40	50	60
ASWM	5.88	4.81	3.58	2.21	1.21	2.63
ROR-NLM	4.61	3.06	1.93	2.63	4.44	6.26
PSMF	1.77	3.96	7.15	10.66	14.39	18.53
ROLD-EPR	2.53	4.20	5.38	6.08	6.39	6.99
MLP-EPR	2.03	3.35	3.75	4.16	3.28	2.99
NRE(本文)	1.17	0.62	0.81	0.98	0.72	1.85

注:加粗字体表示均方根误差结果的前 2 名。

表 6 各噪声检测器在不同噪声水平下的 50 幅纹理图像集下预测情况的 RMSE 值比较

Table 6 RMSE comparison on 50 texture images corrupted with different noise ratios

算法	噪声比例/%					
	10	20	30	40	50	60
ASWM	4.13	3.48	2.77	1.99	1.16	0.94
ROR-NLM	8.76	6.46	4.47	3.32	3.86	5.27
PSMF	4.51	4.40	6.73	10.07	13.71	17.51
ROLD-EPR	2.98	4.98	6.44	7.19	7.60	7.59
MLP-EPR	1.84	3.03	3.41	3.73	3.61	3.62
NRE(本文)	1.94	1.34	1.52	1.82	1.38	2.43

注:加粗字体表示均方根误差结果的前 2 名。

表5和表6中的数据表明,本文 NRE 噪声比例预测器的排名进入前2名。需要说明的是,虽然某些开关型降噪算法的预测准确性在某些情况下高于所提出的 NRE 预测模型,但是其执行时间都非常长。

3.3 降噪效果比较

为了测试 NRE 预测器的预测准确性对后续降噪模块的影响,将根据噪声比例真值(Truth)调用 DnCNN-S 降噪模型(即 Truth + DnCNN-S)完成降噪的方法列为对比算法;另外,为了验证使用 NRE 预测器的必要性,盲 DnCNN 降噪模型(记为 DnCNN-B)也被列为对比算法。所谓 DnCNN-B 降噪模型是指在 $[0, 60\%]$ 噪声比例范围内仅训练一个全局的 DnCNN 降噪模型。对于任意给定的噪声图像,均仅用该全局模型进行降噪。DnCNN 降噪模型3种不同的使用方式在各个比例 Lena 噪声图像

上所获得的 PSNR 值列在表7中。从表7可知: NRE + DnCNN-S 降噪方法的降噪性能与 DnCNN-B 相比得到了较大提升(提高 $0.2 \sim 1.1$ dB),这说明所提出的 NRE 模型与基于 DnCNN 的非开关型 RVIN 深度降噪模型配合使用后妥善解决了 DnCNN 网络模型固有的数据依赖问题(限于篇幅,这里仅提供 DnCNN 深度降噪模型的测试数据。采用 U2Net (Lefkimiatis, 2018)、CBD-Net (Guo 等, 2019)等其他类型的 DCNN 网络结构实现的降噪模型,其降噪效果没有 DnCNN 模型好,但结论是相同的:即针对特定噪声比例范围训练的降噪模型其降噪效果要高于相应的盲降噪模型)。另外 NRE + DnCNN-S 与 Truth + DnCNN-S 的降噪效果相差不大,也印证了 NRE 模型的预测值基本上与真值一致。

表7 DnCNN 模型在不同工作模式下在 Lena 图像上降噪效果(PSNR)比较
Table 7 Performance (PSNR) comparison of different DnCNN models on Lena image

算法	噪声比例/%						/dB
	10	20	30	40	50	60	
Truth + DnCNN-S	38.04	36.36	34.12	32.47	30.35	28.44	
NRE + DnCNN-S	38.15	36.31	34.12	32.42	30.40	28.30	
DnCNN-B	37.79	35.35	33.44	31.62	29.77	27.30	

其次,为了更加全面地验证 NRE 预测器与 DnCNN 降噪模型配合使用后的降噪效果,表8给出各个降噪算法在更为困难的纹理图像集合中各个图像上所获得 PSNR 均值,并且计算 PSNR 值的标准

差(standard deviation)作为反映算法稳定性的指标。由表8可知: NRE + DnCNN-S 模式对噪声图像进行降噪所获得的图像质量是最优的,且波动幅度不大,说明鲁棒性最好。另外, NRE + DnCNN-S 比 DnCNN-S

表8 各降噪算法在50幅纹理图像不同噪声比例下所获得的 PSNR 均值和标准差比较
Table 8 Performance comparison of the competing algorithms in terms of PSNR and its standard deviation on 50 texture images

噪声比例/%	算法							/dB
	ASWM	ROR-NLM	PSMF	ROLD-EPR	MLP-EPR	DnCNN-B	NRE + DnCNN	
10	28.79 ± 4.3	26.89 ± 4.7	18.78 ± 0.8	30.03 ± 4.2	30.03 ± 3.5	32.70 ± 3.15	32.76 ± 2.7	
20	27.88 ± 4.0	26.31 ± 4.4	16.03 ± 0.8	28.01 ± 4.1	27.66 ± 3.8	30.19 ± 3.20	30.62 ± 3.1	
30	26.86 ± 3.6	25.54 ± 4.0	14.49 ± 0.7	26.69 ± 4.1	26.30 ± 3.8	28.29 ± 3.18	28.23 ± 2.8	
40	25.71 ± 3.4	24.55 ± 3.7	13.48 ± 0.6	25.70 ± 4.0	25.22 ± 3.7	26.69 ± 3.17	27.21 ± 3.3	
50	24.05 ± 3.0	22.95 ± 3.2	12.81 ± 0.5	24.80 ± 3.9	24.29 ± 3.7	25.25 ± 3.08	25.69 ± 3.3	
60	21.55 ± 2.6	20.72 ± 2.8	12.37 ± 0.4	23.83 ± 3.6	23.19 ± 3.3	23.73 ± 2.83	24.30 ± 3.3	

注:加粗字体表示每行最优结果。

NN-B 的降噪效果更优,说明本文 NRE 预测器与 DnCNN-S 联合使用后,解决了 DnCNN 等深度卷积网络降噪模型固有的数据依赖问题。

3.4 执行效率

表 9 给出了各个降噪算法在 512×512 像素大小图像上的平均执行时间。在参与对比的算法中,由于仅有 MLP-EPR 算法和所提出的 NRE + DnCNN-S 降噪算法可明确划分为噪声检测和降噪两个阶段并测量执行时间,故单独给出噪声检测模块的执行时间。由表 9 可知:在所有算法中,所提出的 NRE + DnCNN-S 降噪算法的执行时间仅比 PSMF 和 MLP-EPR 算法要长一些。结合表 8 可知,PSMF 和 MLP-EPR 算法的降噪效果比较差。故所提出 NRE + DnCNN-S 降噪方法的降噪效果和执行时间两个方面的综合性能是最优的。单就噪声检测模块的执行时间而言,MLP-EPR 算法的噪声检测部分执行时间为 0.07 s,本文 NRE 模型执行时间仅为 0.02 s,表明 NRE 模型执行效率是相当高的。其原因在于 NRE 预测模型仅仅需要多点随机采样提取图像中 20 个 28×28 像素大小的图块(并不需要处理图像中所有的图块数据),经 3 层浅层卷积网络处理后即可获得噪声比例的估计值。

表 9 各随机脉冲降噪算法在一幅 512×512 像素的图像上平均执行时间对比

Table 9 The average execution time of different denoising algorithms on an image of size 512×512 pixels

算法	执行时间 /s
ASWM	70.2
ROR-NLM	15.3
PSMF	0.4
ROLD-EPR	7.8
MLP-EPR	1.7(0.07)
NRE + DnCNN-S	2.9(0.02)

注:括号中数字为噪声检测模块的执行时间。

4 结论

本文提出了一种基于浅层卷积神经网络的 NRE 预测模型,搭配非开关型的 DnCNN 深度降噪模型使用后,可将其无缝地转化为盲降噪算法,其在

降噪效果和执行效率两方面的性能显著优于现有主流的开关型的 RVIN 降噪算法。取得优异性能的原因在于:NRE 预测模型采用浅层的 CNN 网络结构确保了其执行效率,而多点随机采样测试机制则保证了它的预测准确性。作为独立的前置模块,一旦训练完成,则无需设置任何参数,且可与任意基于 DCNN 的非盲降噪模型配合使用,具有很好的兼容性。

未来的研究可在现有工作的基础上,考虑将 NRE 模型与 DnCNN 模型集成,使得仅训练一个降噪模型即可完成盲降噪任务,进一步提高模型在使用上的便捷性。

参考文献 (References)

- Akkoul S, Ledee R, Leconge R and Harba R. 2010. A new adaptive switching median filter. *IEEE Signal Processing Letters*, 17(6): 587-590 [DOI: 10.1109/LSP.2010.2048646]
- Arbeláez P, Maire M, Fowlkes C and Malik J. 2011. Contour detection and hierarchical image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(5): 898-916 [DOI: 10.1109/TPAMI.2010.161]
- Chang L, Deng X M, Zhou M Q, Wu Z K, Yuan Y, Yang S and Wang H A. 2016. Convolutional neural networks in image understanding. *Acta Automatica Sinica*, 42(9): 1300-1312 (常亮, 邓小明, 周明全, 武仲科, 袁野, 杨硕, 王宏安. 2016. 图像理解中的卷积神经网络. *自动化学报*, 42(9): 1300-1312) [DOI: 10.16383/j.aas.2016.c150800]
- Chen T and Wu H R. 2001. Adaptive impulse detection using center-weighted median filters. *IEEE Signal Processing Letters*, 8(1): 1-3 [DOI: 10.1109/97.889633]
- Dong Y Q, Chan R H and Xu S F. 2007. A detection statistic for random-valued impulse noise. *IEEE Transactions on Image Processing*, 16(4): 1112-1120 [DOI: 10.1109/TIP.2006.891348]
- Garnett R, Huegerich T, Chui C and He W J. 2005. A universal noise removal algorithm with an impulse detector. *IEEE Transactions on Image Processing*, 14(11): 1747-1754 [DOI: 10.1109/TIP.2005.857261]
- Guo S, Yan Z F, Zhang K, Zuo W M and Zhang L. 2019. Toward convolutional blind denoising of real photographs//*Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Long Beach, CA, USA: IEEE: 1713-1722 [DOI: 10.1109/CVPR.2019.00181]
- Ioffe S and Szegedy C. 2019. Batch normalization: accelerating deep network training by reducing internal covariate shift[EB/OL]. [2019-07-15]. <https://arxiv.org/pdf/1502.03167.pdf>

- Jin K H and Ye J C. 2018. Sparse and low-rank decomposition of a Hankel structured matrix for impulse noise removal. *IEEE Transactions on Image Processing*, 27(3): 1448-1461 [DOI: 10.1109/TIP.2017.2771471]
- Lefkimiatis S. 2018. Universal denoising networks: a novel CNN architecture for image denoising//*Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Salt Lake City, UT, USA: IEEE: 3204-3213 [DOI: 10.1109/CVPR.2018.00338]
- Liu L C, Chen C L P, Zhou Y C and You X G. 2015. A new weighted mean filter with a two-phase detector for removing impulse noise. *Information Sciences*, 315: 1-16 [DOI: 10.1016/j.ins.2015.03.067]
- Rawat W and Wang Z H. 2017. Deep convolutional neural networks for image classification: a comprehensive review. *Neural Computation*, 29(9): 2352-2449 [DOI: 10.1162/NECO_a_00990]
- Shi Y G, Hao H Y and Liu Z W. 2018. Cascaded convolutional neural network based hippocampus subfields segmentation. *Journal of Image and Graphics*, 23(1): 74-83 (时永刚, 郝华胤, 刘志文. 2018. 串行处理卷积神经网络的海马子区分割. *中国图象图形学报*, 23(1): 74-83) [DOI: 10.11834/jig.170334]
- Simonyan K and Zisserman A. 2019. Very deep convolutional networks for large-scale image recognition [EB/OL]. [2019-07-15]. <https://arxiv.org/pdf/1409.1556.pdf>
- Turkmen I. 2016. The ANN based detector to remove random-valued impulse noise in images. *Journal of Visual Communication and Image Representation*, 34: 28-36 [DOI: 10.1016/j.jvcir.2015.10.011]
- Wang Z and Zhang D. 1999. Progressive switching median filter for the removal of impulse noise from highly corrupted images. *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, 46(1): 78-80 [DOI: 10.1109/82.749102]
- Xiong B and Yin Z P. 2012. A universal denoising framework with a new impulse detector and nonlocal means. *IEEE Transactions on Image Processing*, 21(4): 1663-1675 [DOI: 10.1109/TIP.2011.2172804]
- Xu Q, Li Y H, Guo Y J, Wu S and Sbert M. 2018. Random-valued impulse noise removal using adaptive ranked-ordered impulse detector. *Journal of Electronic Imaging*, 27(1): #013001 [DOI: 10.1117/1.JEI.27.1.013001]
- Zhang K, Zuo W M, Chen Y J, Meng D Y and Zhang L. 2017. Beyond a Gaussian denoiser: residual learning of deep CNN for image denoising. *IEEE Transactions on Image Processing*, 26(7): 3142-3155 [DOI: 10.1109/TIP.2017.2662206]
- Zhang K, Zuo W M and Zhang L. 2018. FFDNet: toward a fast and flexible solution for CNN-based image denoising. *IEEE Transactions on Image Processing*, 27(9): 4608-4622 [DOI: 10.1109/TIP.2018.2839891]
- Zhang M H, Zhang F Q, Liu Q G and Wang S S. 2019. VST-Net: variance-stabilizing transformation inspired network for Poisson denoising. *Journal of Visual Communication and Image Representation*, 62: 12-22 [DOI: 10.1016/j.jvcir.2019.04.011]
- Zhou J F, Ye S R and Wang H. 2019. Text sentiment classification based on deep convolutional neural network model. *Computer Engineering*, 45(3): 300-308 (周锦峰, 叶施仁, 王晖. 2019. 基于深度卷积神经网络模型的文本情感分类. *计算机工程*, 45(3): 300-308) [DOI: 10.19678/j.issn.1000-3428.0050043]

作者简介



徐少平, 1976年生, 男, 教授, 博士生导师, 主要研究方向为图形图像处理技术、机器视觉、虚拟手术仿真。

E-mail: xushaoping@ncu.edu.cn

林珍玉, 女, 硕士研究生, 主要研究方向为图形图像处理技术、机器视觉。E-mail: 401030918076@mail.ncu.edu.cn
李崇禧, 男, 硕士研究生, 主要研究方向为图形图像处理技术、机器视觉。E-mail: 406130917315@mail.ncu.edu.cn
崔燕, 女, 硕士研究生, 主要研究方向为图形图像处理技术、机器视觉。E-mail: 401030919004@mail.ncu.edu.cn
刘蕊蕊, 女, 硕士研究生, 主要研究方向为图形图像处理技术、机器视觉。E-mail: 411014519042@mail.ncu.edu.cn