



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

中国图象学报 图形学报

2019
08
VOL.24

ISSN1006-8961
CN11-3758/TB



灰度图像彩色化 P1249

中国图象图形学报

刊名题字：宋健 月刊（1996年创刊）



第24卷第8期（总第280期）

2019年8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

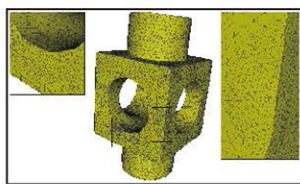
Superintended by Chinese Academy of Sciences
Sponsored by Institute of Remote Sensing and Digital Earth, CAS
China Society of Image and Graphics
Institute of Applied Physics and Computational Mathematics

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



保持特征的联合滤波网格去噪算法(第1292页)



高置信度互补学习的实时目标跟踪(第1315页)



注意力机制改进卷积神经网络的遥感图像目标检测(第1400页)

学者观点

深度卷积神经网络降噪模型的技术瓶颈与研究展望

徐少平, 刘婷云, 林珍玉, 张贵珍, 李崇禧 1207

综述

人脸年龄估计的深度学习方法综述

张珂, 王新胜, 郭玉荣, 苏昱坤, 何颖宣 1215

卷积神经网络在掌纹识别中的性能评估

王海纶, 李书杰, 贾伟, 刘晓平 1231

图像处理和编码

局部自适应的灰度图像彩色化

曹丽琴, 商永星, 刘婷婷, 李治江, 马爱龙 1249

多阶段融合网络的图像超分辨率重建

沈明玉, 俞鹏飞, 汪荣贵, 杨娟, 薛丽霞 1258

融合感知损失的生成式对抗超分辨率算法

杨娟, 李文静, 汪荣贵, 薛丽霞 1270

结合全卷积网络与CycleGAN的图像实例风格迁移

刘哲良, 朱玮, 袁梓洋 1283

保持特征的联合滤波网格去噪算法

黄涛, 曹力, 刘晓平 1292

图像分析和识别

多层语义融合CNN的步态人体语义分割

支双双, 赵庆会, 唐璜 1302

高置信度互补学习的实时目标跟踪

郭伟, 邢宇哲, 曲海成 1315

基于非凸低秩分解判别的叠加线性稀疏人脸识别

叶学义, 罗宵晗, 王鹏, 陈慧云 1327

基于自步学习的鲁棒多样性多视角聚类

唐永强, 张文生 1338

特征注意金字塔调制网络的视频目标分割

汤润发, 宋慧慧, 张开华, 姜斯浩 1349

基于视觉感知的智能扫地机器人的垃圾检测与分类

宁凯, 张东波, 印峰, 肖慧辉 1358

图像理解和计算机视觉

眼球光心标定与距离修正的3维注视点估计

张远辉, 段承杰, 朱俊江, 何雨辰 1369

根据灰度值信息自适应窗口的半全局匹配

黄超, 赵华治 1381

高速公路云台相机的自动标定

李婵, 宋焕生, 武非凡, 王伟, 王璇 1391

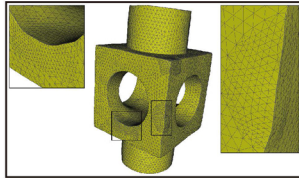
遥感图像处理

注意力机制改进卷积神经网络的遥感图像目标检测

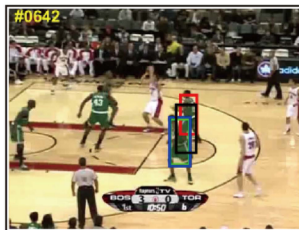
李红艳, 李春庚, 安居白, 任俊丽 1400

CONTENTS

JOURNAL OF IMAGE AND GRAPHICS



Feature preservation with combined filters for mesh denoising(P1292)



Real-time object tracking based on high-confidence complementary learning(P1315)



Attention mechanism improves CNN remote sensing image object detection(P1400)

Scholar View

Main bottlenecks and research prospects of the deep convolutional neural network-based denoising model

Xu Shaoping, Liu Tingyun, Lin Zhenyu, Zhang Guizhen, Li Chongxi 1207

Review

Survey of deep learning methods for face age estimation

Zhang Ke, Wang Xinsheng, Guo Yurong, Su Yukun, He Yingxuan 1215

Performance evaluation of convolutional neural network in palmprint recognition

Wang Hailun, Li Shujie, Jia Wei, Liu Xiaoping 1231

Image Processing and Coding

Novel image colorization of a local adaptive weighted average filter

Cao Liqin, Shang Yongxing, Liu Tingting, Li Zhijiang, Ma Ailong 1249

Image super-resolution reconstruction via deep network based on multi-staged fusion

Shen Mingyu, Yu Pengfei, Wang Ronggui, Yang Juan, Xue Lixia 1258

Generative adversarial network for image super-resolution combining perceptual loss

Yang Juan, Li Wenjing, Wang Ronggui, Xue Lixia 1270

Image instance style transfer combined with fully convolutional network and cycleGAN

Liu Zheliang, Zhu Wei, Yuan Ziyang 1283

Feature preservation with combined filters for mesh denoising

Huang Tao, Cao Li, Liu Xiaoping 1292

Image Analysis and Recognition

Semantic segmentation of gait body with multilayer semantic fusion convolutional neural network

Zhi Shuangshuang, Zhao Qinghui, Tang Jin 1302

Real-time object tracking based on high-confidence complementary learning

Guo Wei, Xing Yuzhe, Qu Haicheng 1315

Face recognition with superposed linear sparse representation based on discriminative nonconvex low-rank matrix decomposition

Ye Xueyi, Luo Xiaohan, Wang Peng, Chen Huiyun 1327

Robust and diverse multi-view clustering based on self-paced learning

Tang Yongqiang, Zhang Wensheng 1338

Video object segmentation via feature attention pyramid modulating network

Tang Runfa, Song Huihui, Zhang Kaihua, Jiang Sihao 1349

Garbage detection and classification of intelligent sweeping robot based on visual perception

Ning Kai, Zhang Dongbo, Yin Feng, Xiao Huihui 1358

Image Understanding and Computer Vision

3D gaze estimation using eyeball optical center calibration and distance correction

Zhang Yuanhui, Duan Chengjie, Zhu Junjiang, He Yuchen 1369

Semi-global stereo matching with adaptive window based on grayscale value

Huang Chao, Zhao Huazhi 1381

Auto-calibration of the PTZ camera on the highway

Li Chan, Song Huansheng, Wu Feifan, Wang Wei, Wang Xuan 1391

Remote Sensing Image Processing

Attention mechanism improves CNN remote sensing image object detection

Li Hongyan, Li Chungeng, An Jubai, Ren Junli 1400

中图法分类号: TP391.41 文献标识码: A 文章编号: 1006-8961(2019)08-1207-08

论文引用格式: Xu S P, Liu T Y, Lin Z Y, Zhang G Z, Li C X. Main bottlenecks and research prospects of the deep convolutional neural network-based denoising model[J]. Journal of Image and Graphics, 2019, 24(08): 1207-1214. [徐少平, 刘婷云, 林珍玉, 张贵珍, 李崇禧. 深度卷积神经网络降噪模型的技术瓶颈与研究展望[J]. 中国图象图形学报, 2019, 24(08): 1207-1214.] [DOI:10.11834/jig.190165]

深度卷积神经网络降噪模型的技术瓶颈与研究展望

徐少平, 刘婷云, 林珍玉, 张贵珍, 李崇禧

南昌大学信息工程学院, 南昌 330031

摘要: 现有的深度卷积神经网络(DCNN)图像降噪模型受其技术路线内在固有特性的制约, 降噪性能仍然有待进一步改进。为了推动现有DCNN图像降噪模型技术的发展, 需要正视并及时解决制约其进一步完善的瓶颈问题。本文简要概述了传统的基于自然图像非局部自相似性、稀疏性和低秩性这3种先验知识设计的图像降噪算法的技术路线特点和优缺点, 从传统图像降噪算法存在的问题中引出基于DCNN构建图像降噪模型的技术优势, 并梳理并总结了DCNN降噪模型未来的发展瓶颈, 就相应的解决方案(研究方向)进行详细讨论。通过深入分析发现, 可以从扩大卷积核的感受野、降低网络参数与训练集之间的依赖关系以及充分利用DCNN网络的建模能力这三个角度入手, 突破现有基于数据驱动的DCNN降噪模型的瓶颈制约, 把图像降噪算法的研究水平推向新的高度。

关键词: 综述; 图像降噪; 深度卷积神经网络; 瓶颈问题; 感受野; 数据依赖; 参数空间

Main bottlenecks and research prospects of the deep convolutional neural network-based denoising model

Xu Shaoping, Liu Tingyun, Lin Zhenyu, Zhang Guizhen, Li Chongxi

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

Abstract: As a representative technique for deep learning, the deep convolution neural network (DCNN) with strong feature learning and nonlinear mapping ability in the field of digital image processing offers a novel opportunity for image denoising research. DCNN-based denoising models show significant advantages over traditional methods in terms of their denoising effect and execution efficiency. However, most of the existing image denoising models are driven by data. Given their inherent restrictions, the denoising performance of these models can be further improved. To promote the development of existing image denoising technologies, some key challenges that restrict their further improvement must be analyzed and addressed. We first summarize the core ideas of traditional image denoising algorithms based on three types of prior knowledge of the natural image, namely, non-local self-similarity, sparsity, and low rank, and then analyze the advantages and disadvantages of these algorithms. Image denoising algorithms modeled with prior knowledge can flexibly deal with distorted images under different noise levels. Unfortunately, they demonstrate the following limitations: 1) the limited hand-crafted image priors are not enough to describe all changes in the image structure, thereby limiting the denoising ability of these algorithms; 2) most of the traditional image denoising algorithms iteratively solve their objective functions, thereby resulting in a high computational complexity; and 3) the optimal solution of the objective function needs to adjust several parameters

收稿日期: 2019-04-29; 修回日期: 2019-05-21; 预印本日期: 2019-05-28

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)

adjust several parameters manually according to the actual situation. Based on the above problems, we point out that the technical advantage of the DCNN-based denoising model lies in its strong nonlinear approximation supported by a graphics processing unit. The inherent characteristics of the DCNN-based denoising model are then analyzed, the bottleneck problems that restrict their future development are presented, and the possible solutions (research directions) to these problems are discussed in detail. **Result** A thorough analysis reveals many bottleneck problems in data-driven DCNN-based denoising models that need to be solved, including: 1) the small receptive field of the DCNN network that limits the range of image feature representation and the ability to fully utilize the priors contained in natural images; 2) the strong dependence of DCNN-based model parameters on the training dataset, that is, the optimal denoising effect can only be obtained if the distortion level of the observed image is close to that of the training images; and 3) the training set cannot be easily constructed and the denoising model cannot be easily trained to denoise an image if both the noise type and level of noisy image are unknown. To solve these problems, we expand the receptive field of the convolution kernel, weaken the dependency between the network parameters and the training set, or fully utilize the modeling ability of the DCNN network. Therefore, the bottlenecks of the existing DCNN denoising models can be addressed, and research on image denoising algorithms can move to a higher level. **Conclusion** In this paper, the technical advantages and development bottlenecks of DCNN in the field of image denoising are summarized, and some future research directions for the image denoising method are proposed. This paper should be of interest to readers in the area of image denoising.

Key words: review; image denoising; deep convolutional neural network (DCNN); bottleneck problem; receptive field; data dependencies; parameter space

0 引言

一般来讲,图像受到噪声干扰后的退化模型可以定义为

$$\mathbf{y} = \mathbf{x} + \mathbf{n} \quad (1)$$

式中, \mathbf{y} 是噪声图像, \mathbf{x} 为原始无失真图像, \mathbf{n} 为噪声。在现有的图像降噪研究中,通常假设噪声 \mathbf{n} 为服从零均值正态高斯分布 $N(0, \sigma^2)$ 的加性高斯噪声^[1-3]。在高斯噪声模型中, σ 是反映噪声严重程度的噪声水平值(noise level),是各类降噪算法中重要参数^[4-5]。基于式(1)的定义,图像降噪问题可以描述为基于噪声图像 \mathbf{y} 获得关于无失真图像 \mathbf{x} 的最佳估计 $\hat{\mathbf{x}}$ (或者关于噪声 \mathbf{n} 的最佳估计 $\hat{\mathbf{n}}$,两者是同一问题)。图像降噪作为一个典型的欠约束优化(unconstrained optimization)问题,近40年来研究者们采用不同技术路线提出了很多解决方案^[6-12]。早期,高斯高/低通滤波器^[6-7]常用于图像降噪任务。它们简单地利用相邻像素点亮度值的非线性组合复原噪声像素点,不可避免地会造成图像边缘细节模糊化。为了更好地保护图像细节,Buades等人^[13]提出的NLM(non-local means)算法利用自然图像的非局部自相似性(NSS)实现了较高的图像质量修复效果。但NLM算法及其之后提出的多种变种算法均

需要在图像空域内搜索大量与待复原图块相似的图块,导致执行时间较长,这在一定程度上限制了这类算法实际的使用范围。为了提高NLM算法的执行效率,Dabov等人^[14]基于NSS自相似特性和稀疏表示(sparse representation)技术,实现了一种降噪效果和执行效率俱佳且具有里程碑意义的被称为BM3D(block-matching and 3D filtering)的降噪算法,该算法常被用做基准测试对比算法。之后,为了从噪声图像 \mathbf{y} 中最佳估计出图像 $\hat{\mathbf{x}}$,研究者们陆续提出了许多策略来提高算法的降噪效果。这些降噪算法通常利用了关于无失真自然图像的先验知识来约束解的空间,通过求解目标函数

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x}} \left\{ \frac{1}{2} \|\mathbf{y} - \mathbf{x}\|_2^2 + \lambda \Phi(\mathbf{x}) \right\} \quad (2)$$

最优值的方式实现降噪,即求解由保真项 $\frac{1}{2} \|\mathbf{y} - \mathbf{x}\|_2^2$ 、正则化项 $\Phi(\mathbf{x})$ 以及正则化参数 λ 组成的目标优化函数极小值问题。这里,保真项用于保证估计图像与噪声图像在内容上保持一致,而正则化项则用于控制估计图像与原始图像之间的估计误差。依据正则化项所使用的自然图像先验知识的不同,基于稀疏表示^[15-17]和基于低秩最小化(low-rank minimization)^[18-19]的图像降噪算法相继被提出。这两类降噪算法在图像降噪效果方面获得了较大的提高,但也都存在以下3个方面的局限性:

1) 基于自然图像稀疏表示和低秩所构建的图像先验知识不足以描述所有复杂的图像结构变化, 没有充分挖掘出图像中的先验信息, 限制了图像的降噪效果;

2) 目标函数的求解通常以复杂的迭代优化过程实现, 使得算法时间复杂度较高;

3) 式(2)所定义的目标函数一般是非凸的(non-convex), 具体实现时需要手动调整若干个参数才能获得最优的结果。

近年来, 深度学习(deep learning)技术因其强大的特征学习和非线性映射能力在图像降噪领域取得了巨大成功^[20-29], 其中尤以基于深度卷积神经网络(DCNN)构建的图像降噪模型发展迅速。基于DCNN构建的降噪模型^[28]通过在大量噪声图像—无失真图像训练数据集上以最小化估计图像 \hat{x} 与原始无失真图像 x 之间的损失函数 $\mathcal{L}(\hat{x}, x)$ 为目标驱动, 学习并调整DCNN网络降噪模型 $\mathcal{F}(y; \Theta)$ 中的网络参数 Θ 实现图像降噪任务, 可形式化表示为

$$\hat{x} = \mathcal{F}(y; \Theta) \quad \text{s. t.} \quad \min_{\Theta} \mathcal{L}(\hat{x}, x) \quad (3)$$

基于DCNN的降噪模型依赖网络结构隐式地学习图像中的先验知识, 展现出了强大的图像先验知识建模能力(非线性映射能力), 能够避免基于稀疏和低秩优化降噪模型中正则项 $\Phi(x)$ 设计困难的问题, 进一步提高了复原图像的质量。此外, 受高性能图形处理单元(GPU)并行计算技术的支持, 基于DCNN网络进行图像降噪能够获得极高的执行效率。基于上述优势, 深度学习方法现已成为近年来图像降噪领域的研究热点之一, 具有广泛的应用前景^[24, 28]。然而, 这类本质上基于数据驱动(data-driven)的降噪模型, 其性能也同时受其所采用技术路线内在固有特性的制约, 存在若干瓶颈问题亟需进一步改进。

1 现存问题及研究展望

基于数据驱动的技术路线使得DCNN降噪模型可以获得远胜于其他主流降噪算法的降噪效果, 但同时也存在以下瓶颈问题亟待解决:

1) 图像特征提取范围小的问题。DCNN网络依赖大量的卷积核提取图像特征用于后续功能模块处理, 然而现有的DCNN网络的卷积核大小通常都设置得比较小, 即所谓的感受野(receptive field)非常

局限, 导致获取的图像特征表征范围有限, 不能充分利用自然图像中广泛存在的NSS特性提高降噪效果。

2) 待降噪图像与降噪模型的相互匹配问题。DCNN降噪模型要想获得最佳的降噪效果, 必须根据图像降质(失真)的严重程度选用预先训练好的特定DCNN降噪模型才能最大程度地发挥其降噪能力。其本质的原因在于通过训练获得的降噪模型的网络参数 Θ 与图像数据集之间存在较强的依赖关系。

3) 训练数据缺乏问题。现有的DCNN降噪模型常常将噪声模型简化为加性高斯噪声, 然而, 在现实噪声图像中的噪声往往并不严格符合高斯分布。在待降噪图像噪声信息未知的情况下, 构建训练数据集非常困难, 导致无法训练降噪模型完成降噪任务。为了解决这些问题, 未来可能的解决方案(研究方向)详细描述如下。

1.1 扩大卷积核的感受野

自然图像在内容上存在显著的NSS非局部自相似性, 研究者利用这个特性或将这个特性与图像的稀疏和低秩特性结合, 实现了很多降噪性能不错的降噪算法^[16-18, 30-33]。

目前DCNN网络中的卷积核通常都设置得非常小, 仅仅能够提取图像中非常有限的局部特征信息, 其处理视野有待进一步扩大。为扩大DCNN网络的感受野, 常规的策略是增大卷积核尺寸和增加网络深度。然而, 增大卷积核的尺寸会引入大量的网络参数并导致网络训练难度增大, 同时执行效率也将降低。另外, 设计层次非常深的神经网络结构容易出现过拟合和非凸优化问题。最近提出的膨胀卷积(dilated convolution)^[34]、沙漏状结构(hourglass-shaped architectures)^[35]等技术可以达到增大感受野的目的。其实, 引入图像的NSS非局部自相似性是扩大DCNN网络卷积视野更好的方式, 有望获得更高的性能提升。

最近这方面探索性的工作有: Ahn等人^[36]提出了一种基于NSS先验知识的块匹配卷积神经网络(BMCNN)算法。该算法首先采用现有的降噪算法(如3维块匹配算法(BM3D)、去噪卷积神经网络(DnCNN)^[24])对噪声图像进行预处理获得对应的参考图像, 然后从噪声图像及参考图像中找到相似的图块, 并将它们堆叠成与BM3D算法类似的3维

矩阵,最后通过 DnCNN 网络架构完成图块降噪。BMCNN 算法可以看做是传统基于非局部自相似性的算法的扩展,其以数据驱动方式训练的框架能够学习更精确的降噪模型。与现有的各种 DCNN 的降噪模型相比,在图像结构比较丰富的自然图像上能获得更好的降噪性能。然而,BMCNN 算法的性能受限于获取参考图像的降噪算法,所获得的参考图像与原始无失真图像毕竟尚有差距,而且也只使用了 1 幅参考图像,从而会影响到相似图块搜索的正确性,在一定程度上制约了其性能。类似地,在文献[37]中,Cruz 等人将基于卷积神经网络的降噪模型与基于非局部自相似性的传统降噪算法级联组合使用。总之,上述两个工作初步实现了在 DCNN 网络中结合 NSS 非局部自相似性,进一步提高了 DCNN 网络模型的降噪效果,但是上述工作还存在一定的发展空间。

由上述 DCNN 网络降噪模型的发展态势可以看出:将自然图像的 NSS 特性集成到 DCNN 模型中,有效地增大 DCNN 网络的感受野是大势所趋,而将传统的基于 NSS 的降噪技术完全无缝地集成到 DCNN 框架下才能最大限度地利用各自的优势。近期,Yang 等人^[38]就传统降噪技术 CNN 网络化的问题进行了探索,他们将 BM3D 算法的执行过程等价展开为一个卷积神经网络结构,提出了一种称为 BM3D-Net 的降噪模型。BM3D-Net 网络模型依次由提取层(extraction)、卷积层、非线性变换层、卷积层和聚集层(aggregation)堆叠而成。其中,提取层对应于 BM3D 中的块匹配操作(block-matching),卷积、非线性变换和卷积这 3 层实现了 BM3D 算法中的 3 维小波变换收缩(3D wavelet transform shrinkage),聚集层则实现了 BM3D 中的块聚集(patch group aggregation)操作。BM3D-Net 算法获得了比传统 BM3D 算法更佳的降噪性能,且能够在 GPU 上执行,从而可以获得极高的执行效率。但该网络模型仅包含 2 个卷积层,其非线性映射能力有待进一步加强。尽管如此,文献[38]的工作表明:传统的基于 NSS 特性的降噪算法的处理过程完全可以在网络化和模块化后作为嵌入网络模块(ENB)集成到现有的 DCNN 网络中。ENB 模块可以看做卷积神经网络结构中类似于卷积层的一种新型网络结构,将其集成到 DCNN 网络架构中可增大 DCNN 网络卷积处理视野,实现一个端到端(end-to-end)的非

局部处理网络,从而进一步提高现有 DCNN 网络的降噪能力。ENB 模块具体的设计与实现将是一个未来可以积极探索的研究方向。

1.2 降低网络参数 Θ 与训练集之间的依赖关系

基于数据驱动的 DCNN 降噪模型充分利用了从训练集大量数据中学习到的先验知识,其降噪能力相比现有主流降噪算法而言提升非常显著。然而,通过训练获得的先验知识泛化能力较弱,只有当用于训练降噪模型的图像集合与待降噪图像在降质程度(高斯噪声模型时即为噪声水平值)近似时才能获得最好的降噪效果。换句话说,现有的 DCNN 网络中的网络参数 Θ 与待降噪图像降质程度存在依赖关系,这样必须对待降噪图像中的噪声水平值进行准确估计,才能调用与之相匹配的预先训练好的降噪模型实现最佳的降噪效果。在这种情况下,要想对任意给定的噪声图像完成降噪,必须训练多个针对特定噪声水平值范围专门训练的降噪模型,导致在训练模型和使用模型方面存在困难。针对这一技术缺陷,文献[28]通过构建一张大小与图像相同的噪声水平映射图 M (每个像素点的亮度值都为噪声水平值 σ) 与待降噪图像联结(concatenate)后作为网络的输入,训练所得的 FFDNet(fast and flexible denoising convolutional neural network)降噪模型可以形式化地表示为 $\hat{x} = \mathcal{F}(y, M; \Theta)$ 。噪声水平映射图的引入给网络增加了新的参数变量(以辅助通道的形式体现),使得训练所获得的模型参数 Θ 不再受噪声水平值 σ 的影响,这样只需训练一个降噪模型就可以灵活处理不同的噪声水平。当然,FFDNet 降噪模型仍然属于非盲降噪模型,使用该模型的时候需要给出准确的噪声水平值 σ (通过 σ 构成噪声水平映射图 M)。然而,FFDNet 模型中的噪声水平映射图 M 的每个元素都统一被设定为噪声水平值,处理过于简单,没有考虑在图像内部噪声水平值存在一定程度变化的实际情况,且仅适用于高斯降噪。未来可以通过学习噪声的分布特征,针对不同像素位置设置不同的噪声水平值,获得更为精准的噪声水平映射图 M ,从而使得训练的 DCNN 降噪模型的处理能力更强。另外,目前的工作只考虑了将噪声图像的噪声水平值作为辅助通道信息,没有考虑将图像结构信息作为辅助信息。为此,未来可将反映图像内容的局部结构边缘信息以映射图 E 的形式抽取出来作为第 2 种辅助通道信息,这样可

以使得训练所获得的降噪模型形式化地表示为 $\hat{x} = \mathcal{F}(y, \mathbf{M}, \mathbf{E}; \Theta)$, 模型中的参数 Θ 将不再受失真程度和图像边缘结构的影响, 从而可以进一步提高模型的适应能力。传统的 DCNN 降噪模型仅以待降噪图像作为输入, 在网络不同层次上提取的图像特征也会在不同程度上呈现出噪声水平映射图 \mathbf{M} 和图像结构信息映射图 \mathbf{E} 类似的信息, 但这种模式是单输入通道的, 且需要使用更多的网络层数(参数), 导致执行效率下降。

因此, 在噪声图像的基础上通过联结噪声水平映射图 \mathbf{M} 和图像结构信息映射图 \mathbf{E} 之后, 采用多通道模式向 DCNN 网络提供这些信息, 有助于降低 DCNN 网络的深度和加快训练速度, 训练所得到的 DCNN 降噪模型的降噪能力也将显著增强。其本质上就是通过利用噪声水平映射图 \mathbf{M} 和图像结构信息映射图 \mathbf{E} 降低预训练模型中的参数 Θ 与图像数据集之间的依赖关系, 使得模型的适应能力更强、通用性更好, 目前这方面的工作亟待开展。

1.3 充分利用 DCNN 网络的建模能力

现有基于数据驱动的 DCNN 降噪模型都需要在相当庞大的数据集上完成训练才可以投入使用。在具体实践中, 常常仅有待降噪图像一幅图像可以利用, 而且其噪声分布是未知的, 并不属于常用且简化的高斯模型。这意味着对于这种待降噪图像, 无法事先生成合适的噪声图像训练集, 训练降噪模型也无从谈起, 故在这种情况下基于训练策略的 DCNN 降噪模型将失效。

为解决训练数据缺乏的问题, 关键需要研究如何仅依赖单幅的噪声图像获得它的噪声分布模型, 有了噪声分布模型就可以实现在选定的无失真图像集合上加入根据噪声分布模型生成的噪声信号, 从而生成用于训练的噪声图像训练集。然而仅依赖单张的噪声图像获得它的噪声分布模型是非常困难的, 这是个典型的欠约束问题。幸运的是自然图像像素点亮度值的分布是非常有规律的, 而且噪声图像的分布模型显著异于正常的无失真图像。目前, 利用解析的方法获得噪声图像的分布模型效果并不理想。考虑到神经网络强大的非线性逼近能力, 已有一些研究学者基于神经网络技术开展了这方面的工作。例如, Chen 等人^[39] 所提出的 GCBD (GAN-CNN based blind denoiser) 算法试图利用具有复杂分布估计能力的生成对抗网络 (GAN) 隐式地学习给

定噪声图像的噪声分布模型。然后, 基于噪声模型构造噪声图像训练数据集。最后, 再利用文献[24]提出的 DnCNN 网络结构训练降噪模型实现盲降噪任务。Chen 等人提出的 GCBD 算法可以直接基于给定的单幅噪声图像完成降噪任务, 克服了显式定义未知噪声模型的困难, 避免了模型对图像训练集的依赖, 其所采用的技术路线对未来的研究工作具有很好的借鉴价值。然而, 该方法仅局限于对加性噪声有效, 在噪声模型方面仍然存在一定的限制。未来需要将 GCBD 模型扩展到能够处理更为真实的噪声 (real-world noise), 以增强降噪模型的鲁棒性和普适性。

鉴于 DCNN 强大的非线性生成(映射)能力, 研究者未来还可以尝试另外一种完全摆脱训练数据集的降噪策略^[40]。也就是说, 基于随机初始化网络参数的 DCNN 网络和给定的单张噪声图像, 通过搜索 DCNN 的参数空间 Θ 获得复原图像。这种方法可以让 DCNN 降噪模型完全摆脱对训练数据的依赖, 仅仅利用调整 DCNN 网络参数空间 Θ 参数值的方法获得所谓的基于 DIP (deep image prior) 先验知识从而实现降噪, 这是一个非常有前途的 DCNN 网络发展方向。目前采用 DIP 深度图像先验知识的降噪模型, 其网络参数 Θ 仅对特定的噪声图像有效, 仅有一次使用性(不同的噪声图像需要搜索不同的参数空间), 且在降噪过程中需要大量的迭代过程才能获得, 未来提高其执行效率是很好的研究方向, 从而改善它的实用性。

2 结 论

综上所述, 基于 DCNN 的降噪模型近几年已经在图像降噪领域获得了巨大的成功, 然而这种基于数据驱动的降噪模型在发展过程中也遇到了其本身技术特点所带来的结构性瓶颈问题亟需解决。

对于 DCNN 的小感受野问题, 可以将基于 NSS 特性实现的降噪算法处理过程网络化后集成到 DCNN 来解决; 对于降噪模型的网络参数 Θ 与图像数据集有严重的依赖关系问题, 可以利用噪声水平映射图 \mathbf{M} 和图像结构信息映射图 \mathbf{E} 作为网络的多通道输入信息来降低网络参数 Θ 对噪声图像的依赖关系, 使得模型的适应能力更强, 通用性更好; 对于在缺乏训练数据的条件下, 可以利用 DCNN 网络

的强大生成能力,直接通过调整网络参数实现对噪声建模,从而获得大量数据训练数据集实现降噪的目的。未来还可以基于给定噪声图像直接通过调整网络参数实现图像降噪。这样使得 DCNN 降噪模型不再需要训练数据即可实现降噪,拓展了应用范围。总之,若突破了上述瓶颈问题,将有望把基于 DCNN 网络的图像降噪技术推向新的研究高度。

参考文献 (References)

- [1] Chen L, Yang G, Chen Z Y, et al. Linearized Bregman iteration algorithm for matrix completion with structural noise[J]. Chinese Journal of Computers, 2015, 38 (7): 1357-1371. [陈蕾, 杨庚, 陈正宇, 等. 基于线性 Bregman 迭代的结构化噪声矩阵补全算法[J]. 计算机学报, 2015, 38 (7): 1357-1371.] [DOI: 10.11897/SP.J.1016.2015.01357]
- [2] Shen J M, Li J F, Dai W Z. No-reference image quality assessment based on structure information and luminance statistics[J]. Acta Electronica Sinica, 2016, 44(4): 804-812. [沈军民, 李俊峰, 戴文战. 结合结构信息和亮度统计的无参考图像质量评价[J]. 电子学报, 2016, 44 (4): 804-812.] [DOI: 10.3969/j.issn.0372-2112.2016.04.008]
- [3] Xu S P, Zeng X X, Tang Y L, et al. A noise level estimation algorithm using prior knowledge of similar images[J]. Journal of Computer Research and Development, 2018, 55 (12): 2741-2752. [徐少平, 曾小霞, 唐祎玲, 等. 基于多图像先验知识的噪声水平评估算法[J]. 计算机研究与发展, 2018, 55(12): 2741-2752.] [DOI: 10.7544/issn1000-1239.2018.20170336]
- [4] Xu S P, Zeng X X, Tang Y L. Fast noise level estimation algorithm based on two-stage support vector regression[J]. Journal of Computer-Aided Design & Computer Graphics, 2018, 30 (3): 447-458. [徐少平, 曾小霞, 唐祎玲. 基于两阶段支持向量回归的快速噪声水平估计算法[J]. 计算机辅助设计与图形学学报, 2018, 30 (3): 447-458.] [DOI: 10.3724/SP.J.1089.2018.16422]
- [5] Gong X C, Li Z M. An image denoising method using deep asymmetrical skip connection [J]. Journal of Computer-Aided Design & Computer Graphics, 2019, 31 (2): 295-302. [公绪超, 李宗民. 深度非对称跳跃连接的图像降噪方法[J]. 计算机辅助设计与图形学学报, 2019, 31 (2): 295-302.] [DOI: 10.3724/SP.J.1089.2019.17167]
- [6] Smith S M, Brady J M. SUSAN—a new approach to low level image processing[J]. International Journal of Computer Vision, 1997; 23(1): 45-78. [DOI: 10.1023/a:1007963824710]
- [7] Tomasi C, Manduchi R. Bilateral filtering for gray and color images [C]//Proceedings of the 6th International Conference on Computer Vision. Bombay, India: IEEE, 1998: 839-846. [DOI: 10.1109/iccv.1998.710815]
- [8] Aharon M, Elad M, Bruckstein A. K-SVD: an algorithm for designing overcomplete dictionaries for sparse representation [J]. IEEE Transactions on Signal Processing, 2006, 54 (11): 4311-4322. [DOI: 10.1109/tsp.2006.881199]
- [9] Zhang L, Dong W S, Zhang D, et al. Two-stage image denoising by principal component analysis with local pixel grouping [J]. Pattern Recognition, 2010, 43 (4): 1531-1549. [DOI: 10.1016/j.patcog.2009.09.023]
- [10] Tang C, Xu J L, Zhou Z G. Strong noise image-denoising algorithm based on improved curvature filters [J]. Journal of Image and Graphics, 2019, 24 (3): 346-356. [汤成, 许建龙, 周志光. 改进的曲率滤波强噪声图像去噪方法[J]. 中国图象图形学报, 2019, 24 (3): 346-356.] [DOI: 10.11834/jig.180302]
- [11] 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.] [DOI: 10.11834/jig.20170402]
- [12] Remez T, Litany O, Giryes R, et al. Class-aware fully convolutional Gaussian and Poisson denoising [J]. IEEE Transactions on Image Processing, 2018, 27 (11): 5707-5722. [DOI: 10.1109/TIP.2018.2859044]
- [13] Buades A, Coll B, Morel J M. A non-local algorithm for image denoising [C]//Proceedings of 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. San Diego, CA, USA: IEEE, 2005: 60-65. [DOI: 10.1109/CVPR.2005.38]
- [14] 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. [DOI: 10.1109/TIP.2007.901238]
- [15] Mairal J, Bach F, Ponce J, et al. Non-local sparse models for image restoration [C]//Proceedings of the 12th International Conference on Computer Vision. Kyoto, Japan: IEEE, 2009: 2272-2279. [DOI: 10.1109/ICCV.2009.5459452]
- [16] Dong W S, Zhang L, Shi G M, et al. Nonlocally centralized sparse representation for image restoration [J]. IEEE Transactions on Image Processing, 2013, 22 (4): 1620-1630. [DOI: 10.1109/TIP.2012.2235847]
- [17] Zhang J, Zhao D B, Gao W. Group-based sparse representation for image restoration [J]. IEEE Transactions on Image Processing, 2014, 23 (8): 3336-3351. [DOI: 10.1109/TIP.2014.2323127]
- [18] Dong W S, Shi G M, Li X. Nonlocal image restoration with bilateral variance estimation: a low-rank approach [J]. IEEE Trans-

- actions on Image Processing, 2013, 22(2): 700-711. [DOI: 10.1109/TIP.2012.2221729]
- [19] Gu S H, Zhang L, Zuo W M, et al. Weighted nuclear norm minimization with application to image denoising [C]//Proceedings of 2014 IEEE Conference on Computer Vision and Pattern Recognition. Columbus, OH, USA: IEEE, 2014: 2862-2869. [DOI: 10.1109/CVPR.2014.366]
- [20] 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. Lake Tahoe, Nevada: ACM, 2012: 341-349.
- [21] Burger H C, Schuler C J, Harmeling S. Image denoising: can plain neural networks compete with BM3D? [C]//Proceedings of 2012 IEEE Conference on Computer Vision and Pattern Recognition. Providence, RI, USA: IEEE, 2012: 2392-2399. [DOI: 10.1109/CVPR.2012.6247952]
- [22] Santhanam V, Morariu V I, Davis L S. Generalized deep image to image regression [C]//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, HI, USA: IEEE, 2017: 5395-5405. [DOI: 10.1109/CVPR.2017.573]
- [23] Zhang K, Zuo W M, Gu S H, et al. Learning deep CNN denoiser prior for image restoration [C]//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, HI, USA: IEEE, 2017: 2808-2817. [DOI: 10.1109/CVPR.2017.300]
- [24] 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. [DOI: 10.1109/TIP.2017.2662206]
- [25] Du B, Xiong W, Wu J, et al. Stacked convolutional denoising auto-encoders for feature representation [J]. IEEE Transactions on Cybernetics, 2017, 47(4): 1017-1027. [DOI: 10.1109/TCYB.2016.2536638]
- [26] Wu C Z, Chen X, Ji D, et al. Image denoising via residual network based on perceptual loss [J]. Journal of Image and Graphics, 2018, 23(10): 1483-1491. [吴从中, 陈曦, 季栋, 等. 结合深度残差学习和感知损失的图像去噪 [J]. 中国图象图形学报, 2018, 23(10): 1483-1491.] [DOI: 10.11834/jig.180069]
- [27] Zhang J S, Zhang Y, Bai L F, et al. Lossless-constraint denoising based auto-encoders [J]. Signal Processing: Image Communication, 2018, 63: 92-99. [DOI: 10.1016/j.image.2018.02.002]
- [28] Zhang K, Zuo W M, Zhang L. FFDNet: toward a fast and flexible solution for CNN-based image denoising [J]. IEEE Transactions on Image Processing, 2018, 27(9): 4608-4622. [DOI: 10.1109/TIP.2018.2839891]
- [29] Plötz T, Roth S. Neural nearest neighbors networks [EB/OL]. 2018-10-30 [2019-04-04]. <http://arxiv.org/pdf/1810.12575.pdf>.
- [30] Liu H F, Xiong R Q, Ma S W, et al. Non-local extension of total variation regularization for image restoration [C]//Proceedings of 2014 IEEE International Symposium on Circuits and Systems. Melbourne VIC, Australia: IEEE, 2014: 1102-1105. [DOI: 10.1109/ISCAS.2014.6865332]
- [31] Xu J, Zhang L, Zuo W M, et al. Patch group based nonlocal self-similarity prior learning for image denoising [C]//Proceedings of 2015 IEEE International Conference on Computer Vision. Santiago, Chile: IEEE, 2015: 244-252. [DOI: 10.1109/ICCV.2015.36]
- [32] Xu S P, Yang X H, Jiang S L. A fast nonlocally centralized sparse representation algorithm for image denoising [J]. Signal Processing, 2017, 131: 99-112. [DOI: 10.1016/j.sigpro.2016.08.006]
- [33] Jin K H, Ye J C. Sparse and low-rank decomposition of a Hankel structured matrix for impulse noise removal [J]. IEEE Transactions on Image Processing, 2018, 27(3): 1448-1461. [DOI: 10.1109/TIP.2017.2771471]
- [34] Yu F, Koltun V. Multi-scale context aggregation by dilated convolutions [EB/OL]. 2016-04-30 [2019-04-04]. <http://arxiv.org/pdf/1511.07122.pdf>.
- [35] Shelhamer E, Long J, Darrell T. Fully convolutional networks for semantic segmentation [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(4): 640-651. [DOI: 10.1109/TPAMI.2016.2572683]
- [36] Ahn B, Cho N I. Block-matching convolutional neural network for image denoising [EB/OL]. 2017-04-03 [2019-04-04]. <http://arxiv.org/pdf/1704.00524.pdf>.
- [37] Cruz C, Foi A, Katkovnik V, et al. Nonlocality-reinforced convolutional neural networks for image denoising [J]. IEEE Signal Processing Letters, 2018, 25(8): 1216-1220. [DOI: 10.1109/LSP.2018.2850222]
- [38] Yang D, Sun J. BM3D-Net: a convolutional neural network for transform-domain collaborative filtering [J]. IEEE Signal Processing Letters, 2018, 25(1): 55-59. [DOI: 10.1109/lsp.2017.2768660]
- [39] Chen J W, Chen J W, Chao H Y, et al. Image blind denoising with generative adversarial network based noise modeling [C]//Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, UT, USA: IEEE, 2018: 3155-3164. [DOI: 10.1109/CVPR.2018.00333]
- [40] Ulyanov D, Vedaldi A, Lempitsky V. Deep image prior [EB/OL]. 2018-04-05 [2019-04-04]. <http://arxiv.org/pdf/1711.10925.pdf>.

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

刘婷云, 女, 硕士研究生, 主要研究方向为图形图像处理和机
器视觉。E-mail: 416114517210@email.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和机
器视觉。E-mail: 401030918076@email.ncu.edu.cn

张贵珍, 女, 硕士研究生, 主要研究方向为图形图像处理和机
器视觉。E-mail: 406130917331@email.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和机
器视觉。E-mail: 406130917315@email.ncu.edu.cn