

基于级联密集网络的轮廓波变换域图像复原*

刘宇男¹, 张姗姗¹, 王春鹏², 李广宇¹, 杨健¹

¹(南京理工大学 计算机科学与工程学院, 江苏 南京 210094)

²(齐鲁工业大学 计算机科学与技术学院, 山东 济南 250353)

通讯作者: 张姗姗, E-mail: shanshan.zhang@njust.edu.cn



摘要: 近年来,卷积神经网络凭借极强的学习能力,在图像复原任务上实现了比传统学习方法更令人满意的结果.但是,由于丢失了重要的纹理细节,这些基于卷积神经网络的方法普遍存在着复原图像过度平滑的缺点.为解决该问题,提出一种基于级联密集型卷积神经网络的轮廓波域图像复原方法,可以应用于单幅图像去噪、超分辨率及 JPEG 解压缩这 3 个经典图像复原任务.首先,构建了一种紧凑的级联密集型网络结构,不但可以充分挖掘和利用不同层次的图像特征,而且解决了由于网络加深带来的长期依赖问题.接着,引入可以稀疏表示图像重要特征的轮廓波变换,分别将低质量图像和重建图像对应的轮廓波子带作为网络的输入和输出,更加有效地恢复出逼真的结构和纹理细节.在标准测试集的实验表明:提出的方法在 3 个图像复原任务上达到了当前最优的性能,不但获得了更高的峰值信噪比和结构相似度,而且在主观的重建图像中包含了更加真实的纹理细节.

关键词: 图像去噪;超分辨率;JPEG 解压缩;轮廓波变换;级联密集型卷积神经网络

中图法分类号: TP391

中文引用格式: 刘宇男,张姗姗,王春鹏,李广宇,杨健.基于级联密集网络的轮廓波变换域图像复原.软件学报,2020,31(12): 3968-3980. <http://www.jos.org.cn/1000-9825/5866.htm>

英文引用格式: Liu YN, Zhang SS, Wang CP, Li GY, Yang J. Image restoration based on cascading dense network in contourlet transform domain. Ruan Jian Xue Bao/Journal of Software, 2020,31(12):3968-3980 (in Chinese). <http://www.jos.org.cn/1000-9825/5866.htm>

Image Restoration Based on Cascading Dense Network in Contourlet Transform Domain

LIU Yu-Nan¹, ZHANG Shan-Shan¹, WANG Chun-Peng², LI Guang-Yu¹, YANG Jian¹

¹(School of Computer Science and Engineering, Nanjing University of Science and Technology, Nanjing 210094, China)

²(School of Computer Science and Technology, Qilu University of Technology, Ji'nan 250353, China)

Abstract: In recent years, due to the powerful learning ability, convolutional neural networks (CNN) have achieved more satisfactory results than conventional learning methods in image restoration tasks. However, these CNN-based methods generally have the disadvantage of producing over-smoothed restored image due to the fact that losing important textural details. In order to solve this problem, this study proposes an image restoration method based on cascaded dense CNN (CDCNN) in contourlet transform, which can be used for three classical image restoration tasks, namely, single image denoising, super resolution, and JPEG decompression. First, this

* 基金项目: 国家自然科学基金(61702262, 61861136011, 61802212, U1713208); 长江学者计划; 江苏省自然科学基金(BK2018 1299); 中央高校基本科研专项基金(30918011322); 中国科学技术协会青年人才托举工程(2018QNRC001); 并行与分布式处理实验室科学技术开放基金(WDZC20195500106)

Foundation item: National Natural Science Foundation of China (61702262, 61861136011, 61802212, U1713208); Program for Changjiang Scholars; Natural Science Foundation of Jiangsu Province (BK20181299); Fundamental Research Funds for the Central Universities (30918011322); Young Elite Scientists Sponsorship Program by CAST (2018QNRC001); Science and Technology on Parallel and Distributed Processing Laboratory (PDL) Open Fund (WDZC20195500106)

收稿时间: 2019-01-10; 采用时间: 2019-05-13; jos 在线出版时间: 2019-08-09

CNKI 网络优先出版: 2019-08-12 12:08:15, <http://kns.cnki.net/kcms/detail/11.2560.TP.20190812.1208.009.html>

study constructs a compact cascading dense network structure, which not only fully exploits and utilizes the different hierarchical features of images, but also solves the problem of the long-term dependency problem as growing the network depth. Next, this study introduces the contourlet transform into CDCNN, which can sparsely represent the important image features. Here, the contourlet subbands of low-quality image and corresponding restored image are used as the input and output of the network respectively, which can recover realistic structure and texture details more effectively. Comprehensive experiments on the standard benchmarks show that the unanimous superiority of the proposed method on all three tasks over the state-of-the-art methods. The proposed method not only obtains higher peak signal-to-noise ratio (PSNR) and structural similarity index (SSIM), but also contains more realistic textural details in the subjective reconstruct images.

Key words: image denoising; super resolution; JPEG decompression; contourlet transform; cascading dense convolutional neural network

作为重要的多媒体资源,图像是人们获取和传播信息的主要载体.但是,传输和存储过程中的很多因素都会造成图像质量的退化,不仅影响了视觉效果,甚至严重干扰了对图像内容的正确理解.为了解决该问题,图像复原技术应运而生,并逐渐成为计算机视觉领域的研究热点问题.

通常情况下,图像的退化模型可以表示为 $x = D(\tilde{x}) + n$ ^[1],这里的 D 和 n 分别表示退化函数和加性噪声.很多图像复原方法都是借助该模型进行反向推演,由已知的低质量图像 x 重建出高质量的目标图像 \tilde{x} .在这些图像复原方法中,超分辨率、去噪和 JPEG 解压缩是研究最广泛的 3 个经典任务.

图像去噪的目标是尽可能准确地去除图像上的噪声,这些噪声通常被认为是标准差为 σ 的加性高斯白噪声.根据贝叶斯理论可知^[2],先验信息可以在图像去噪时发挥重要作用.基于此发现,很多基于先验建模的方法获得了比较理想的去噪结果,比如稀疏表示方法^[3]、非局部自相似方法^[4]和隐马尔可夫树建模方法^[5].但由于涉及非凸优化问题并且需要手工初始化参数,这些方法普遍存在复杂度高、优化困难等缺点.图像超分辨率的目标是根据低分辨率图像重建出高分辨率图像,早期的插值方法^[6]和重建方法^[7]具有处理速度快的优点,但却无法获得令人满意的重建结果.与之相比,基于稀疏编码的学习方法^[8,9]通过字典学习建立出低分辨率和高分辨率图像块间的稀疏关联,使超分辨率重建的性能取得了大幅度提升.但传统学习方法将图像块的提取和融合看成了相互独立的步骤^[10],且超分辨率的性能也极易受到手工初始参数的影响.JPEG 解压缩的目标是去除由于图像 JPEG 压缩造成的块效应^[11],早期的方法主要包括基于滤波的方法^[12]和基于稀疏表示的方法^[13],虽然对压缩图像的边缘起到了锐化作用,但残留的伪影以及生成的平滑区域使得主观复原图像中仍然存在严重失真.

近几年,凭借极强的学习能力,卷积神经网络(convolutional neural network,简称 CNN)在图像复原任务上取得了众多令人满意的结果.受到稀疏编码的启发,Dong 等人^[14]提出了 SRCNN 模型,首次将 CNN 应用于图像超分辨率任务,获得了比传统方法更高的重建精度.此后,基于 CNN 的方法逐渐成为了图像复原领域的研究热点.从第 1 个 SRCNN 模型到由 Zhang 等人^[15]提出的 RCAN,网络的卷积层数由 3 层增加到 400 层.网络深度的增加,带来了复原性能的显著提高.与此同时,很多新颖的 CNN 结构设计和训练技巧也被用来进一步提升复原的性能.但这些方法的重建图像普遍存在过度平滑问题,尤其对于退化严重的图像,重建图像丢失了大量的结构和纹理细节.为了解决该问题,很多方法^[16-19]将生成对抗网络^[20]用于图像复原任务,虽然可以生成更加逼真的纹理结构,但通常在客观指标的评估上表现平平,无法与逼真的主观恢复结果相符.而且由于训练不稳定,给模型性能的优化带来了很大困难.作为一种更好的解决办法,一些方法将小波变换与 CNN 结合^[21-23],简单且有效地同时提高了客观评估和主观重建结果.小波变换的高频子带可以稀疏地表示图像的高频(如轮廓、纹理及边缘等)细节,通过 CNN 学习到稀疏的高频信息,可以帮助更好地重建出图像的纹理细节.虽然小波变换在一维信号处理中可以同时捕获信号全局和局部上的重要特征,但当处理二维图像时,小波变换却不能提供最优的稀疏表示^[24],各向同性等缺陷制约了其对图像的稀疏表示能力.

为了解决现有方法的问题,并进一步提升图像复原的性能,本文提出了一种结合级联密集型 CNN 结构和轮廓波变换的图像复原方法.本文的贡献概括如下.

- 本文提出了一种级联的密集型卷积神经网络(cascading dense convolutional neural network,简称 CDCNN):将若干个密集块以级联方式组成主干结构,充分挖掘并利用不同层次上的图像特征.此外,每

个密集块输出的特征图都会作为后面密集块的输入,有效解决了深层网络中的长期依赖问题;

- 引入轮廓波变换,将图像问题的求解从空域转换到变换域.具体地,分别将退化图像和重建图像的分解子带作为 CDCNN 的输入和输出,借助轮廓波变换优秀的稀疏表示能力,通过低频子带捕获图像的全局拓扑结构,高频子带复原出图像更加逼真的结构和纹理细节;
- 本文提出的模型可以应用于图像去噪、超分辨率及 JPEG 解压缩.标准验证集上的实验表明:本方法在 3 个图像复原任务上都达到了目前最优性能,主观结果进一步证明了本文方法在结构和纹理细节重建上的显著优势.

本文第 1 节介绍图像复原的相关工作.第 2 节详细叙述基于 CDCNN 和轮廓波变换结合的图像复原方法.第 3 节为实验评估.最后为总结.

1 图像复原相关工作

与传统方法相比,CNN 凭借极强的学习能力,使图像复原能力得到了显著提升.这部分将对图像复原任务上 CNN 方法的相关工作进行回顾.

作为图像超分辨率领域最早的 CNN 方法,Dong 等人^[10]提出的 SRCNN 模型在重建精度上明显高于传统方法.但由于输入网络前需要对图像进行上采样预处理,造成该方法存在计算复杂度过高的问题.为了提高速度,Dong 等人^[13]和 Shi 等人^[25]分别引入了反卷积层和亚像素卷积层替换上采样预处理,将这两个卷积层作为 CNN 的上采样模块,使大部分的卷积操作可以在低分辨率空间执行.借助上采样模块,计算复杂度大幅度降低,这在此后的很多超分辨率方法中也得到了沿用,例如 DDSR^[26],LapSRN^[27],DBPN^[28]及 IDN^[29]在上采样模块中用了反卷积层,CARN^[30],EDSR^[31],SRMD^[1],MSRN^[32],RCAN^[15]和 RDN^[33]则使用了亚像素卷积.除了上采样模块,特征提取模块是另一个重要组成部分,很多方法的特征提取模块由一系列相同的基础块串联而成^[30-35],每个基础块对特征的挖掘能力将对重建性能起到至关重要的作用.残差块^[36]和密集块^[33,37]是超分辨率 CNN 中比较常用的基础块,此外,Hui 等人^[29]将一种新颖的信息蒸馏块作为基础块,Li 等人^[38]在残差块和密集块^[37]的基础上设计了多尺度残差块作为基础块,Zhang 等人^[33]在密集块基础上引入局部残差学习和局部特征融合,使深层网络的学习更加稳定.在图像识别等高层计算机视觉任务中,GoogLeNet^[36]和 ResNet^[39]的优异表现证明了增加 CNN 深度对提高性能的有效性.显然,这个结论在底层视觉任务中同样适用.除了提高每个基础块能力,增加基础块数量成为了提升性能的有效途径.从最初 3 层卷积的 SRCNN^[14]到 80 层的 MemNet^[34],再到最近提出的 400 层的 RCAN^[15],CNN 深度的增加带来了性能的显著提升.但是,简单地增加深度会带来很多问题,梯度消失和梯度爆炸导致训练不稳定,计算量和参数量的大幅增加导致应用价值降低.为了解决上述问题,很多新颖的 CNN 结构和训练技巧被相继提出和应用,VDSR^[40]引入的残差学习和梯度裁剪策略让训练过程变得更加稳定.DRCN^[41],DRRN^[35]和 CARN^[30]通过引入递归结构,在基本保持性能的情况下大幅度降低网络的参数量.MemNet^[34],MSRN^[32]和 RDN^[33]充分地挖掘和利用图像不同层次的特征,为超分辨率重建提供了更多的线索.

与图像超分辨率领域相同,CNN 在图像去噪及 JPEG 解压缩任务上也取得了丰硕的成果,而且与超分辨率领域的发展有很多相似之处,增加 CNN 网络深度同样带来了重建性能的显著提升.早在 2009 年,Jain 等人^[42]就将一个 4 层卷积的 CNN 结构应用于图像去噪任务,但由于网络深度过浅以及训练数据匮乏等原因,导致早期的很多 CNN 方法^[42-44]并不能获得理想的去噪结果.此后,Zhang 等人^[2]提出了 DnCNN 模型,结合残差学习和批标准化,使去噪能力得到了大幅度提升.Mao 等人^[45]提出了 RED 模型,通过对称的跳跃连接,提高了训练的收敛速度,并进一步提升了去噪质量.为了实现速度和精度更好的平衡,Santhanam 等人^[46]提出了 RBDN 模型,Zhang 等人^[47]引入扩张滤波,提出了 7 层的全卷积网络.Dong 等人^[48]最早提出了 4 层的 ARCNN 模型,用于去除 JPEG 压缩造成的块效应.此后,TNRD^[49],DDCN^[50]和 D3^[11]模型在加深 CNN 结构的同时充分利用了 JPEG 压缩的相关先验知识,大幅度提高了 JPEG 解压缩的能力.

基于 CNN 方法使图像复原的能力得到了显著提升,但这些方法的主要缺陷在于:由于丢失了大量纹理和结构细节,造成重建图像存在过度平滑的问题.为了解决这个问题,很多基于生成对抗网络的方法^[30,31,38]开始应用

于图像复原任务,虽然这些方法生成的图像包含了更多逼真的纹理细节,但训练的不稳定性给性能优化带来了新的难题.此外,基于生成对抗网络方法生成的图像通常会出现比较严重的失真,导致这些方法通常提供了较差的客观评估结果.作为一种简单的解决方案,小波变换与 CNN 结合的方法^[21-23]被引入到图像复原任务中.小波变换提供的高频子带可以稀疏地捕获到不同方向下的图像特征,借助这个优势,这些方法将空域的非线性映射转换到小波变换域,简单有效地实现了图像复原在主观和客观评估结果上的同时优化.但是,小波变换的主要缺陷在于:当处理二维图像及更高维信号时,小波变换并不能提供最优的稀疏表示^[24],高频子带用“点”拟合的边缘和轮廓会出现不连续问题,各向同性的特点也限制了小波变换对图像的稀疏表示能力.

2 本文方法

为了更好地解决图像复原任务,本文提出了一种基于级联密集型的轮廓波域的图像复原方法.首先介绍级联密集型卷积神经网络结构,然后对轮廓波变换及其与 CDCNN 结构结合进行详细阐述.

2.1 级联密集型卷积神经网络

如图 1 所示,CDCNN 包括两个部分,分别为浅层特征提取模块和深层特征提取模块:浅层特征提取模块中包含了串联的两个卷积层;深层特征提取模块由若干个密集块(dense block,简称 DB)级联而成,每个密集块结构如图 2 所示.图 2 的结构图中,每个 DB 包含 4 层卷积,与残差块(residual block,简称 RB)^[36]相比,每个 DB 中的卷积层连接更加紧密,通过通道拼接操作,使其中每个卷积层的输出都作为后面卷积层的输入,实现了信息的最大化流通^[38].在 CDCNN 中,位于不同深度的 DB 可以充分地挖掘出不同层次的图像特征,而且每个 DB 的输入是浅层特征提取模块以及前面所有 DB 输出特征图的拼接结果.这样的设计使得每个 DB 挖掘出来的不同层次特征都会作用于后面 DB,不但为后面 DB 提供了更多的特征挖掘线索,而且解决了深度 CNN 中存在的长期依赖问题.

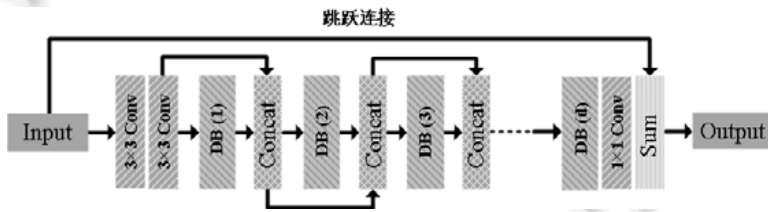


Fig.1 Structure of the proposed cascading dense convolutional neural network (CDCNN)

图 1 级联密集型网络结构

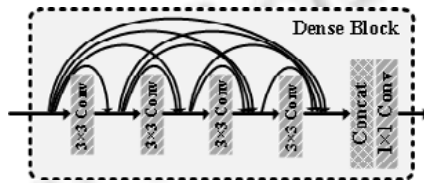


Fig.2 Structure of the proposed dense block (DB)

图 2 密集块结构

假设退化图像和重建的目标图像分别表示为 I^{LR} 和 I^{HR} ,我们 CDCNN 的目的是解决下面问题:

$$\hat{\theta} = \arg \min_{\theta} \frac{1}{N} \sum_{i=1}^N L(F_{CDN}(I_i^{LR}), I_i^{HR}) \tag{1}$$

其中, N 表示训练样本个数, $F_{CDN}(I^{LR})$ 表示通过 CDCNN 重建的复原图像, θ 表示卷积的参数集合, L 表示目标函数.Hui 等人^[29]实验性地证明了:在图像复原任务上先用 L1 损失函数训练,再用 L2 损失函数优化参数,可以达到更好的重建结果.为了避免引入额外的训练技巧,本文训练过程将只使用 L1 损失函数.

退化图像 I^{LR} 经过浅层特征提取模块后,输出的特征图可以表示为 $F_{shallow}$,在其后面跟随 D 个级联 DB.令 F_{d_output} 表示第 $d(d=1,2,\dots,D)$ 个 DB 的输出,则第 d 个 DB 的输入可以表示为

$$F_{d_input}=[F_{shallow},F_{1_output},F_{2_output},\dots,F_{d-1_output}] \quad (2)$$

其中, $[\cdot]$ 表示通道拼接操作.如图 1 中,CDCNN 中引入了跳跃连接,通过残差学习即得到重建图像 $F_{CDN}(I^{LR})$:

$$F_{CDN}(I^{LR})=F_{final}(F_{D_output})+I^{LR} \quad (3)$$

其中, F_{final} 表示连接在最后一个 DB 后的卷积函数,用于改变输出的通道数.

2.2 轮廓波变换域的 CDCNN

与小波变换相比,轮廓波变换为图像提供了更好的稀疏表示^[24].本文提出的方法利用一级轮廓波变换和 CNN 结构结合,图 3 中给出了轮廓波变换和小波变换的高频子带融合示意图,即将同一尺度不同方向的高频子带按元素相加.为了直观显示,将高频子带系数扩大 20 倍.如图 3 所示:小波变换的高频系数不够连续,而且丢失了很多细节信息,例如蝴蝶图像中一些条纹的边缘;与之相比,轮廓波变换高频子带则可以更好地捕获到图像的高频信息(如边缘、轮廓及纹理等细节).

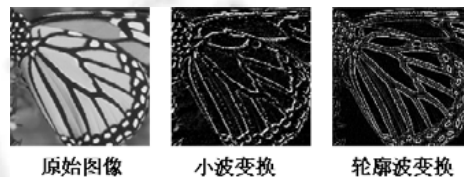


Fig.3 Comparison of contourlet transform and wavelet coefficients

图 3 轮廓波变换和小波变换系数对比

本文在 CDCNN 的基础上,引入轮廓波变换去恢复更多的高频细节,将传统 CNN 空域像素的预测问题转变为轮廓波变换域系数的预测问题,如图 4 所示.首先对退化图像进行一级轮廓波变换,获得一个低频子带和 4 个高频子带:低频子带为近似系数,捕获了输入图像的全局拓扑关系;高频子带为细节系数,稀疏地表示了图像不同方向上的高频信息.然后,分别将低频子带和高频子带输入到两个 CDCNN 中.最后,将两个 CDCNN 的输出子带进行组合,通过逆轮廓波变换即可得到复原图像.这种方法简单且有效,而且同样适用于其他基于 CNN 的图像复原方法.

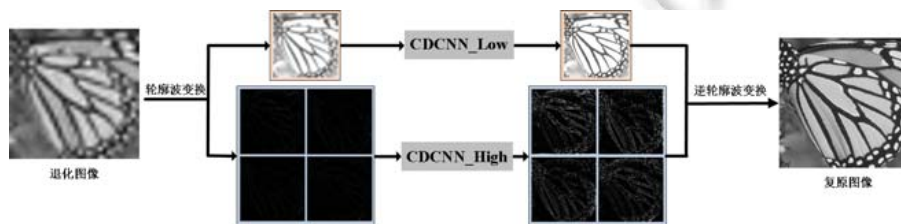


Fig.4 Prediction of contourlet transform sub-bands

图 4 轮廓波变换子带的预测

3 实验

这部分将从客观评估指标和主观重建效果两个方面衡量图像的复原能力.训练图像来自 DIV2K^[51]中包含 800 张图像的原始训练集,且均未经过任何数据扩充;测试图像来自于 Set5,Set14 和 BSD100 这 3 个经典的测试集.采用大小为 48×48 、步长为 36 的滑窗在退化图像和高质量图像上选取出对应的图像块.然后对每个图像块进行一级轮廓波变换,得到训练样本对.初始学习率设置为 0.00 01,每 50 个 epoch 降低一半,在 Tesla P40 上训练 150 个 epoch 可达到收敛.

3.1 模型简化测试

本节将对 CDCNN 结构设计和轮廓波变换的有效性进行验证.

- 网络深度选择实验

基于 CNN 的图像复原模型中,增加网络深度对提升性能起到了至关重要的作用,但同时也带来参数量的增加,影响了模型的应用价值.这里,我们将给出 CDCNN 深度与图像复原性能的关系,从中选取出合理的结构设计.表 1 中给出了 BSD100 数据集上不同深度 CDCNN 的去噪结果以及参数量,BSD 测试图像分别添加 0 均值标准差为 50 的加性高斯噪声.这里,每个卷积层的卷积核为 3×3,滤波个数为 32,A 表示 CDCNN 中密集块的个数,B 表示每个密集块中卷积层的个数.可以看出:虽然增加网络深度可以带来复原性能的不不断提升,但当达到一定深度时,性能的提升开始变慢,而参数量却仍在持续快速增加.为了使复原性能和网络参数量达到很好的平衡,本文提出的 CDCNN 中包含 7 个密集块,每个密集块中包含 6 个卷积层,剩余实验部分均采用这种网络结构设计.

Table 1 Comparison results of different network depths on BSD100

表 1 BSD100 上不同网络深度对比结果

PSNR(dB)/Parameters(K)	A=6	A=7	A=8
B=4	25.90/555	26.03/679	26.14/812
B=6	26.19/1063	26.35/1272	26.42/1490
B=8	26.45/1793	26.52/2123	26.60/2463

图 5 给出了浅层以及深层特征提取模块中每个密集块输出特征图的平均值,以及深层特征提取模块中每个密集块输出特征图的平均值.可以看出:随着深度增加,深层特征图的高频细节越来越清晰.在 CDCNN 中,每个密集块位于网络中的不同深度,可以很好地捕获到不同层次的图像特征,而且底层模块输出的特征将用于更高层特征的挖掘,不但为高层特征的挖掘提供了更多线索,还有效地解决了长期依赖问题.

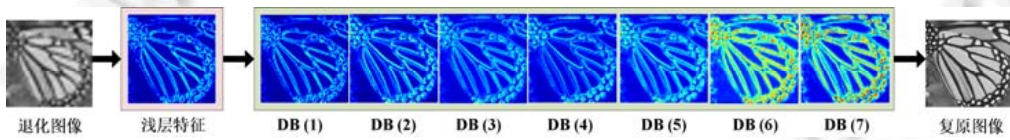


Fig.5 Average feature maps

图 5 平均特征图

- 轮廓波变换有效性证明

本文在 CDCNN 的基础上引入轮廓波变换,可以有效地提升纹理细节的恢复能力.这里,将对 BSD100 数据集上退化图像的去噪、JPEG 解压缩和超分辨率进行测试.图 6 中给出了 3 种任务复原结果,选用包含 7 个 DB 的 CDCNN 结构作为基准,分别与小波变换和轮廓波变换结合,结合方式如图 4 所示.

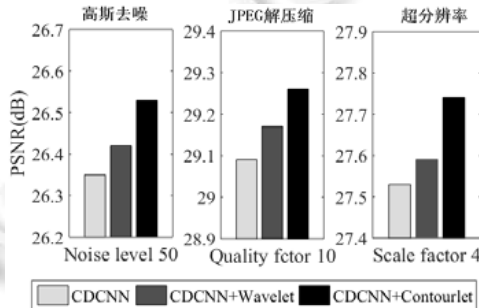


Fig.6 Effectiveness of contourlet transform prediction

图 6 轮廓波变换预测的有效性

在 3 个图像复原任务中,高斯噪声为均值为 0 标准差为 50 的加性噪声,JPEG 压缩的质量因子为 10,超分辨率的尺度因子为 4.结合 CDCNN 和轮廓波变换的模型在 3 种任务的性能明显优于其他两种模型,证明了轮廓波

变换与 CDCNN 结合的有效性.

3.2 客观评估

我们将选择 PSNR 和 SSIM 指标^[52]对本文提出方法进行客观评估.首先,表 2 给出了图像去噪结果,其中:BM3D^[3]和 PGPD^[4]为传统基于先验的图像去噪方法;RED^[45],DnCNN^[2],MemNet^[34]以及本文提出方法属于基于 CNN 的方法.表 2 中对比方法均采用与其对应的开源代码进行测试,“-/-”表示该方法没有处理对应高斯噪声浓度的模型.表 2 中测试图像为 BSD100 中的 100 幅彩色图像,每幅图像转化灰度图像后添加均值为 0,标准差分别为 30,50,70 的高斯白噪声.可以看出:CNN 方法的去噪能力强于传统基于先验估计的方法,且本文方法在不同高斯噪声浓度下均展现了最优的复原性能.

Table 2 Average PSNR(dB)/SSIM results of different methods for image denoising on BSD100

表 2 BSD100 上不同方法图像去噪的平均 PSNR 和 SSIM 结果

标准差	PSNR					
	BM3D	PGPD	RED	DnCNN	MemNet	本文方法
$\sigma=30$	25.966	27.611	28.227	28.144	28.235	28.678
$\sigma=50$	23.807	25.580	26.094	26.047	26.197	26.531
$\sigma=70$	22.585	24.395	24.843	-/-	24.949	25.230
SSIM						
$\sigma=30$	0.763 5	0.755 5	0.792 1	0.785 8	0.791 9	0.803 7
$\sigma=50$	0.672 5	0.669 3	0.705 6	0.699 3	0.714 9	0.723 2
$\sigma=70$	0.609 8	0.614 7	0.648 5	-/-	0.657 5	0.664 2

表 3 给出了 BSD100 数据集上,图像 JPEG 解压缩结果,对比方法均为基于 CNN 方法.JPEG 质量因子分别设置为 5,10,20.对比方法 DnCNN^[2],MemNet^[34],RED^[45],TNRD^[49]结果来自于对应方法的开源代码,这些方法未提供 JPEG 质量因子为 5 时的复原模型,在表 3 中用“-/-”表示.通过表 3 看出:当 JPEG 质量因子为 10 和 20 时,本文方法的复原结果明显高于其他方法;当受到更严重 JPEG 压缩(质量因子为 5)时,本文方法仍然可以实现较高的 PSNR 和 SSIM.

Table 3 Average PSNR(dB)/SSIM results of different JPEG decompression methods on BSD100

表 3 BSD100 上不同图像 JPEG 解压缩方法的平均 PSNR 和 SSIM 结果

质量因子	PSNR					
	JPEG	RED	TNRD	DnCNN	MemNet	本文方法
5	25.254	-/-	-/-	-/-	-/-	26.792
10	27.576	28.864	28.823	28.843	29.049	29.260
20	29.725	31.053	31.012	31.054	31.241	31.417
SSIM						
5	0.642 6	-/-	-/-	-/-	-/-	0.698 2
10	0.747 8	0.784 1	0.782 6	0.783 1	0.788 5	0.791 2
20	0.831 7	0.857 0	0.856 1	0.857 8	0.860 4	0.866 7

表 4 给出了 Set5,Set14 及 BSD100 数据集上,不同超分辨率方法的 PSNR 和 SSIM,对比方法包括基于 CNN 方法 DnCNN^[2],MemNet^[34],DRRN^[35],RED^[45]以及基于对抗神经网络的方法 SRGAN^[17].由表 4 可以看出,本文方法在 2 倍和 4 倍尺度的超分辨率任务上均提供了最优的重建性能.

Table 4 Average PSNR(dB)/SSIM results of different methods

表 4 图像超分辨率方法的 PSNR 和 SSIM 结果

数据集	尺度	PSNR							本文方法
		Bicubic	RED	DnCNN	SRGAN	DRRN	MemNet	IDN	
Set 5	×2	33.66	37.66	37.58	-/-	37.74	37.78	37.83	38.07
	×4	28.42	31.51	31.40	29.40	31.68	31.74	31.82	32.22
Set 14	×2	30.24	32.94	33.03	-/-	33.23	33.28	33.30	33.66
	×4	26.00	27.86	28.04	26.02	28.21	28.26	28.25	28.70
BSD100	×2	29.56	31.99	31.90	-/-	32.05	32.08	32.08	32.28
	×4	25.96	27.40	27.29	25.16	27.38	27.40	27.41	27.74

Table 4 Average PSNR(dB)/SSIM results of different methods (Continued)

表 4 图像超分辨率方法的 PSNR 和 SSIM 结果(续)

数据集	尺度	SSIM							本文方法
		Bicubic	RED	DnCNN	SRGAN	DRRN	MemNet	IDN	
Set 5	×2	0.929 9	0.959 9	0.959 0	—/—	0.959 1	0.959 7	0.960 0	0.961 1
	×4	0.810 4	0.886 9	0.884 5	0.847 2	0.888 8	0.889 3	0.890 3	0.894 6
Set 14	×2	0.868 8	0.914 4	0.912 8	—/—	0.913 6	0.914 2	0.914 8	0.917 5
	×4	0.702 7	0.771 8	0.767 2	0.739 7	0.772 1	0.772 3	0.773 0	0.777 8
BSD100	×2	0.843 1	0.897 4	0.896 1	—/—	0.897 3	0.897 8	0.898 5	0.902 3
	×4	0.667 5	0.729 0	0.725 3	0.668 8	0.728 4	0.728 1	0.729 7	0.736 1

3.3 主观复原结果

这部分将通过主观结果进一步证明本文方法在恢复高频细节上的显著优势.图 7 给出了图像去噪结果:图 7(f)为原始图像;在图 7(f)上添加标准差为 30 的高斯噪声得到图 7(a);图 7(b)~图 7(d)分别为 DnCNN^[2], MemNet^[34]及 RED^[45]的去噪结果,这些方法虽然去除了大部分高斯噪声,但在一些区域出现了比较严重的失真情况,如图中方框中给出的局部放大区域;如图 7(e)所示,本文方法的结果不但取得了更高的 PSNR 和 SSIM,而且在局部细节的恢复能力上明显强于其他方法,如方框中人眼睛部分的结构及树木区域的纹理细节.



Fig.7 Comparisons results of image denoising (Gaussian noise level is 30)

图 7 图像去噪的对比结果(高斯噪声等级为 30)

图 8 给出图像 JPEG 解压缩结果:图 8(f)为原始图像;对图 8(f)进行质量因子为 10 的 JPEG 压缩得到图 8(a);图 8(b)~图 8(d)分别为 DnCNN^[2]及 MemNet^[34],RED^[45]的 JPEG 解压缩结果,这些方法虽然在一定程度上解决了 JPEG 压缩造成的块效应,但在复原一些高频信息复杂的区域时却丢失了大量细节,如图中方框中给出的局部放大区域;本文方法的结果如图 8(e)所示,不但去除 JPEG 伪影的能力强于其他方法,而且还展现出了最优结构及纹理复原能力,如方框中桌布网格结构及裤子的线条边缘.

图 9 给出 4 倍尺度上图像超分辨率的结果:图 9(h)为原始图像;图 9(a)~图 9(f)分别为双三次插值上采样, DnCNN^[2],IDN^[29],MemNet^[34],DRRN^[35]以及 RED^[45]的结果,这些方法虽然恢复出了部分高频细节,但对于纹理分布密集的区域却出现了比较严重的失真,如图中方框中给出的局部放大区域;与其他方法相比,本文方法结果中的斑马条纹更加接近于原始高分辨率图像.如表 4 所示,基于对抗神经网络的方法 SRGAN^[17]在客观评价指标

PSNR 和 SSIM 上均明显低于其他方法,但由于 SRGAN 提出的初衷是为了重建出视觉上更加逼真的超分辨率图像,主观的视觉质量通常会优于其他方法.图 10 中给出了本文方法与 SRGAN 的对比结果,可以看出:本文方法不但获得了更高的 PSNR 和 SSIM,而且在主观复原质量方面也强于 SRGAN,图像超分辨率性能的优势在主观视觉质量和客观评价指标上实现了统一.

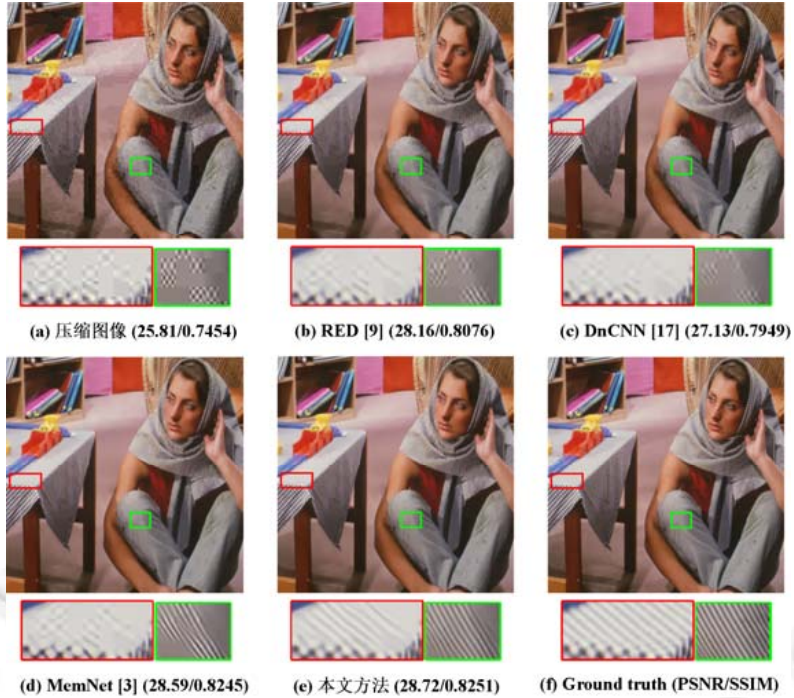


Fig.8 Comparisons results of image JPEG decompression (quality factor is 10)

图 8 图像 JPEG 解压缩对比结果(质量因子为 10)

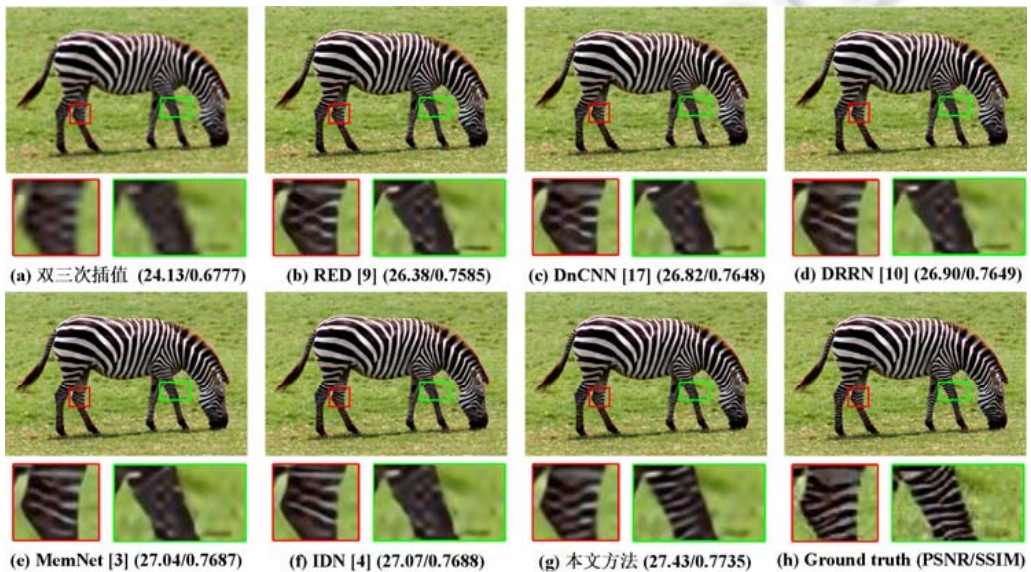


Fig.9 Qualitative results of different image super resolution methods (scale factor is 4)

图 9 不同图像超分辨率方法的定性结果(缩放因子为 4)



Fig.10 Qualitative results of different methods (scale factor is 4)

图 10 不同方法的定性结果(缩放因子为 4)

4 总 结

本文提出了一种结合级联密集神经网络和轮廓波变换的图像复原方法,该方法通过连接更加紧密的卷积神经网络,充分地挖掘和利用了不同层次的图像特征,并且有效地解决了深层网络中存在的长期依赖问题.在此基础上,本文引入轮廓波变换,提高了结构及纹理等高频细节的复原能力.标准测试集上的测试结果表明:本文方法在图像去噪、JPEG 解压缩以及超分辨率这 3 个复原任务上均展现了最优的性能,不但获得了更好的客观评估结果,而且主观的复原图像中包含了更加逼真的结构及纹理细节.

References:

- [1] Zhang K, Zuo WM, Zhang L. Learning a single convolutional super-resolution network for multiple degradations. In: Proc. of the IEEE Int'l Conf. on Computer Vision and Pattern Recognition (CVPR). 2018. 3262–3271. [doi: 10.1109/CVPR.2018.00344]
- [2] Zhang K, Zuo WM, Chen YJ, Meng DY, Zhang L. Beyond a Gaussian denoiser: Residual learning of deep CNN for image denoising. IEEE Trans. on Image Processing (TIP), 2017,26(7):3142–3155. [doi: 10.1109/TIP.2017.2662206]
- [3] Dabov K, Foi A, Katkovnik V, Egiazarian KO. Image denoising by sparse 3-d transform-domain collaborative filtering. IEEE Trans. on Image Processing (TIP), 2007,16(8):2080–2095. [doi: 10.1109/TIP.2007.901238]
- [4] Xu J, Zhang L, Zuo WM, Zhang D, Feng XC. Patch group based nonlocal self-similarity prior learning for image denoising. In: Proc. of the IEEE Int'l Conf. on Computer Vision (ICCV). 2015. 244–252. [doi: 10.1109/ICCV.2015.36]
- [5] Wang XH, Zhu YH, Lv F, Su X, Song CM. Cauchy distribution NSST-HMT model and its applications in image denoising. Chinese Journal of Computers, 2018,41(11):2496–2508 (in Chinese with English abstract). <http://cjc.ict.ac.cn/qwjs/view.asp?id=5087> [doi: 10.11897/SP.J.1016.2018.02496]
- [6] Zhou F, Yang WM, Liao QM. Interpolation-Based image super-resolution using multisurface fitting. IEEE Trans. on Image Processing (TIP), 2012,21(7):3312–3318. [doi: 10.1109/TIP.2012.2189576]
- [7] Chantas GK, Galatsanos NP, Woods NA. Super-Resolution based on fast registration and maximum a posteriori reconstruction. IEEE Trans. on Image Processing (TIP), 2007,16(7):1821–1830. [doi: 10.1109/TIP.2007.896664]
- [8] Li M, Cheng J, Le X, Luo HM. Super-Resolution based on sparse dictionary coding. Ruan Jian Xue Bao/Journal of Software, 2013, 23(5):1315–1324 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/3989.htm> [doi: 10.3724/SP.J.1001.2012.03989]

- [9] Yang J, Wright J, Huang TS, Ma Y. Image super-resolution via sparse representation. *IEEE Trans. on Image Processing (TIP)*, 2010,19(11):2861–2873. [doi: 10.1109/TIP.2010.2050625]
- [10] Dong C, Loy CC, He KM, Tang XO. Image super-resolution using deep convolutional networks. *IEEE Trans. on Pattern Analysis and Machine Intelligence (T-PAMI)*, 2016,38(2):295–307. [doi: 10.1109/TPAMI.2015.2439281]
- [11] Wang ZY, Liu D, Chang SY, Ling Q, Yang YZ, Huang TS. D3: Deep dual-domain based fast restoration of JPEG-compressed image. In: *Proc. of the IEEE Int'l Conf. on Computer Vision and Pattern Recognition (CVPR)*. 2016. 2764–2772. [doi: 10.1109/CVPR.2016.302]
- [12] Lee K, Kim DS, Kim T. Regression-Based prediction for blocking artifact reduction in jpeg-compressed images. *IEEE Trans. on Image Processing (TIP)*, 2005,14(1):36–48. [doi: 10.1109/TIP.2004.838699]
- [13] Chang H, Ng MK, Zeng T. Reducing artifacts in JPEG decompression via a learned dictionary. *IEEE Trans. on Signal Processing (TSP)*, 2014,62(3):718–728. [doi: 10.1109/TSP.2013.2290508]
- [14] Dong C, Loy CC, He KM, Tang XO. Learning a deep convolutional network for image super-resolution. In: *Proc. of the European Conf. on Computer Vision (ECCV)*. 2014. 184–199. [doi: 10.1007/978-3-319-10593-2_13]
- [15] Zhang YL, Li KP, Li K, Wang LC, Zhong BN, Fu Y. Image super-resolution using very deep residual channel attention networks. In: *Proc. of the European Conf. on Computer Vision (ECCV)*. 2018. 294–310. [doi: 10.1007/978-3-030-01234-2_18]
- [16] Wang XT, Yu K, Dong C, Loy CC. Recovering realistic texture in image super-resolution by deep spatial feature transform. In: *Proc. of the IEEE Int'l Conf. on Computer Vision and Pattern Recognition (CVPR)*. 2018. 606–615. [doi: 10.1109/CVPR.2018.00070]
- [17] Ledig C, Theis L, Huszar F, Caballero J, Cunningham A. Photo-Realistic single image super-resolution using a generative adversarial network. In: *Proc. of the IEEE Int'l Conf. on Computer Vision and Pattern Recognition (CVPR)*. 2017. 105–114. [doi: 10.1109/CVPR.2017.19]
- [18] Galteri L, Seidenari L, Bertini M, Bimbo AD. Deep generative adversarial compression artifact removal. In: *Proc. of the IEEE Int'l Conf. on Computer Vision (ICCV)*. 2017. 4836–4845. [doi: 10.1109/ICCV.2017.517]
- [19] Guo J, Chao HY. One-to-Many network for visually pleasing compression artifacts reduction. In: *Proc. of the IEEE Int'l Conf. on Computer Vision and Pattern Recognition (CVPR)*. 2017. 4867–4876. [doi: 10.1109/CVPR.2017.517]
- [20] Goodfellow I, Jean PA, Mirza M, Xu B, Warde-Farley D, Ozair S, Courville A, Bengio Y. Generative adversarial nets. In: *Proc. of the Advances in Neural Information Processing Systems (NIPS)*. 2014. 2672–2680.
- [21] Guo T, Mousavi HS, Vu TH, Monga V. Deep wavelet prediction for image super-resolution. In: *Proc. of the IEEE Int'l Conf. on Computer Vision and Pattern Recognition Workshops (CVPRW)*. 2017. 1100–1109. [doi: 10.1109/CVPRW.2017.148]
- [22] Huang HB, He R, Sun ZN, Tan TN. Wavelet-SRNET: A wavelet-based CNN for multi-scale face super resolution. In: *Proc. of the IEEE Int'l Conf. on Computer Vision (ICCV)*. 2017. 1698–1706. [doi: 10.1109/ICCV.2017.187]
- [23] Liu PJ, Zhang HZ, Zhang K, Lin L, Zuo WM. Multi-Level wavelet-CNN for image restoration. In: *Proc. of the IEEE Int'l Conf. on Computer Vision and Pattern Recognition Workshops (CVPRW)*. 2018. 886–895. [doi: 10.1109/CVPRW.2018.00121]
- [24] Do MN, Vetterli M. The contourlet transform: An efficient directional multiresolution image representation. *IEEE Trans. on Image Processing (TIP)*, 2005,14(12):2091–2106. [doi: 10.1109/TIP.2005.859376]
- [25] Shi WZ, Caballero J, Huszar F, Totz J, Aitken AP, Bishop R, Rueckert D, Wang Z. Real-Time single image and video super-resolution using an efficient sub-pixel convolutional neural network. In: *Proc. of the IEEE Int'l Conf. on Computer Vision and Pattern Recognition (CVPR)*. 2016. 1874–1883. [doi: 10.1109/CVPR.2016.207]
- [26] Peng YL, Zhang L, Zhang Y, Liu SG, Guo M. Deep deconvolution neural network for image super-resolution. *Ruan Jian Xue Bao/ Journal of Software*, 2018,29(4):926–934 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/5407.htm> [doi: 10.13328/j.cnki.jos.005407]
- [27] Lai WS, Huang JB, Ahuja N, Yang MH. Deep laplacian pyramid networks for fast and accurate super-resolution. In: *Proc. of the IEEE Int'l Conf. on Computer Vision and Pattern Recognition (CVPR)*. 2017. 5835–5843. [doi: 10.1109/CVPR.2017.618]
- [28] Haris M, Shakhnarovich G, Ukita N. Deep back-projection network for super-resolution. In: *Proc. of the IEEE Int'l Conf. on Computer Vision and Pattern Recognition (CVPR)*. 2018. 1664–1673. [doi: 10.1109/CVPR.2018.00179]

- [29] Hui Z, Wang XM, Gao XB. Fast and accurate single image super-resolution via information distillation network. In: Proc. of the IEEE Int'l Conf. on Computer Vision and Pattern Recognition (CVPR). 2018. 723–731. [doi: 10.1109/CVPR.2018.00082]
- [30] Ahn N, Kang B, Sohn KA. Fast, accurate, and lightweight super-resolution with cascading residual network. In: Proc. of the European Conf. on Computer Vision (ECCV). 2018. 256–272. [doi: 10.1007/978-3-030-01249-6_16]
- [31] Lim B, Son S, Kim H, Nah S, Lee KM. Enhanced deep residual networks for single image super-resolution. In: Proc. of the IEEE Int'l Conf. on Computer Vision and Pattern Recognition Workshops (CVPRW). 2017. 1132–1140. [doi: 10.1109/CVPRW.2017.151]
- [32] Li JC, Fang FM, Mei KF, Zhang GX. Multi-Scale residual network for image super-resolution. In: Proc. of the European Conf. on Computer Vision (ECCV). 2018. 527–542. [doi: 10.1007/978-3-030-01237-3_32]
- [33] Zhang YL, Tian YP, Kong Y, Zhong BN, Fu Y. Residual dense network for image super-resolution. In: Proc. of the IEEE Int'l Conf. on Computer Vision and Pattern Recognition (CVPR). 2018. 2472–2481. [doi: 10.1109/CVPR.2018.00262]
- [34] Tai Y, Yang J, Liu XM, Xu CY. MemNet: A persistent memory network for image restoration. In: Proc. of the IEEE Int'l Conf. on Computer Vision (ICCV). 2017. 4549–4557. [doi: 10.1109/ICCV.2017.486]
- [35] Tai Y, Yang J, Liu XM. Image super-resolution via recursive residual network. In: Proc. of the IEEE Int'l Conf. on Computer Vision and Pattern Recognition (CVPR). 2017. 2790–2798. [doi: 10.1109/CVPR.2017.298]
- [36] He KM, Zhang XY, Ren SQ, Sun J. Deep residual learning for image recognition. In: Proc. of the IEEE Int'l Conf. on Computer Vision and Pattern Recognition (CVPR). 2016. 770–778. [doi: 10.1109/CVPR.2016.90]
- [37] Tong T, Li Gen, Liu XJ, Gao QQ. Image super-resolution using dense skip connections. In: Proc. of the IEEE Int'l Conf. on Computer Vision (ICCV). 2017. 4809–4817. [doi: 10.1109/ICCV.2017.514]
- [38] Huang G, Liu Z, Maaten LVD, Weinberger KQ. Densely connected convolutional networks. In: Proc. of the IEEE Int'l Conf. on Computer Vision and Pattern Recognition (CVPR). 2017. 2261–2269. [doi: 10.1109/CVPR.2017.243]
- [39] Szegedy C, Liu W, Jia YQ, Sermaner P, Reed S, Anguelov D, Erhan D, Vanhoucke V, Rabinovich A. Going deeper with convolutions. In: Proc. of the IEEE Int'l Conf. on Computer Vision and Pattern Recognition (CVPR). 2015. 1–9. [doi: 10.1109/CVPR.2015.7298594]
- [40] Kim J, Lee JK, Lee KM. Accurate image super-resolution using very deep convolutional networks. In: Proc. of the IEEE Int'l Conf. on Computer Vision and Pattern Recognition (CVPR). 2016. 1646–1654. [doi: 10.1109/CVPR.2016.182]
- [41] Kim J, Lee JK, Lee KM. Deeply-Recursive convolutional network for image super-resolution. In: Proc. of the IEEE Int'l Conf. on Computer Vision and Pattern Recognition (CVPR). 2016. 1637–1645. [doi: 10.1109/CVPR.2016.181]
- [42] Jain V, Seung S. Natural image denoising with convolutional networks. In: Proc. of the Advances in Neural Information Processing Systems (NIPS). 2009. 769–776.
- [43] Agostinelli F, Anderson MR, Lee H. Adaptive multi-column deep neural networks with application to robust image denoising. In: Proc. of the Advances in Neural Information Processing Systems (NIPS). 2013. 1493–1501.
- [44] Xie J, Xu L, Chen E. Image denoising and inpainting with deep neural networks. In: Proc. of the Advances in Neural Information Processing Systems (NIPS). 2012. 341–349.
- [45] Mao XJ, Shen CH, Yang YB. Image restoration using very deep convolutional encoder-decoder networks with symmetric skip connections. In: Proc. of the Advances in Neural Information Processing Systems (NIPS). 2016. 2802–2810.
- [46] Santhanam V, Morariu VI, Davis LS. Generalized deep image to image regression. In: Proc. of the IEEE Int'l Conf. on Computer Vision and Pattern Recognition (CVPR). 2017. 5395–5405. [doi: 10.1109/CVPR.2017.573]
- [47] Zhang K, Zuo W, Gu S, Zhang L. Learning deep CNN denoiser prior for image restoration. In: Proc. of the IEEE Int'l Conf. on Computer Vision and Pattern Recognition (CVPR). 2017. 2808–2817. [doi: 10.1109/CVPR.2017.300]
- [48] Dong C, Deng YB, Loy CC, Tang XO. Compression artifacts reduction by a deep convolutional network. In: Proc. of the IEEE Int'l Conf. on Computer Vision (ICCV). 2015. 576–584. [doi: 10.1109/ICCV.2015.73]
- [49] Chen Y, Pock T. Trainable nonlinear reaction diffusion: A flexible framework for fast and effective image restoration. IEEE Trans. on Pattern Analysis and Machine Intelligence (T-PAMI), 2017,39(6):1256–1272. [doi: 10.1109/TPAMI.2016.2596743]
- [50] Guo J, Chao HY. Building dual-domain representations for compression artifacts reduction. In: Proc. of the European Conf. on Computer Vision (ECCV). 2016.628–644. [doi: 10.1007/978-3-319-46448-0_38]

- [51] Agustsson E, Timofte R. Ntire 2017 challenge on single image super-resolution: Dataset and study. In: Proc. of the IEEE Int'l Conf. on Computer Vision and Pattern Recognition Workshops (CVPRW). 2017. 1122–1131. [doi: 10.1109/CVPRW.2017.150]
- [52] Wang Z, Bovik AC, Sheikh HR, Simoncelli EP. Image quality assessment: from error visibility to structural similarity. IEEE Trans. on Image Processing (TIP), 2004,13(4):600–612. [doi: 10.1109/TIP.2003.819861]

附中文参考文献:

- [5] 王向海,朱毅欢,吕芳,苏欣,宋传鸣.基于 Cauchy 分布的非下采样 Shearlet HMT 模型及其图像去噪应用.计算机学报,2018,41(11): 2496–2508. <http://cjc.ict.ac.cn/qwjs/view.asp?id=5087> [doi: 10.11897/SP.J.1016.2018.02496]
- [8] 李民,程建,乐翔,罗环敏.稀疏字典编码超分辨率重建.软件学报,2013,23(5):1315–1324. <http://www.jos.org.cn/1000-9825/3989.htm> [doi: 10.3724/SP.J.1001.2012.03989]
- [26] 彭亚丽,张鲁,张钰,刘侍刚,郭敏.基于深度反卷积神经网络的图像超分辨率算法.软件学报,2018,29(4):926–934.<http://www.jos.org.cn/1000-9825/5407.htm> [doi: 10.13328/j.cnki.jos.005407]



刘宇男(1991—),男,博士生,主要研究领域为计算机视觉,图像处理.



李广宇(1985—),男,博士,讲师,CCF 专业会员,主要研究领域为模式识别.



张姗姗(1987—),女,博士,教授,博士生导师,CCF 专业会员,主要研究领域为计算机视觉,模式识别.



杨健(1973—),男,博士,教授,博士生导师,主要研究领域为模式识别,计算机视觉,机器学习.



王春鹏(1989—),男,博士,副教授,主要研究领域为图像处理,机器学习.