

一种利用改进深度图像先验构建的图像降噪模型

徐少平, 李 芬, 陈孝国, 陈晓军, 江顺亮

(南昌大学数学与计算机学院, 江西南昌 330031)

摘要: 为进一步提高深度图像先验(Deep Image Prior, DIP)降噪模型的降噪效果和执行效率,从网络结构、网络输入和 Loss 函数三个方面对其进行改进从而获得了一种改进的深度图像先验(Improved Deep Image Prior, IDIP)降噪模型. 具体地,在网络结构方面,通过新增非线性特征传递路径的方法将原 DIP 模型编码器-解码器(encoder-decoder)架构中相同尺度特征层之间所采用的简单连接改进为复杂连接,有利于特征信息调制与传递从而提高神经网络的非线性映射能力;在网络输入方面,用已具有较高图像质量的初步降噪图像替换随机张量作为网络输入向网络模型提供更为丰富的信息,有利于加快网络收敛速度进而提高执行效率;在 Loss 函数方面,在原噪声图像的基础上新增初步降噪图像作为第二目标图像,有利于提高 Loss 函数的导向能力从而提高降噪效果. 实验结果表明:所提出的 IDIP 降噪模型在各噪声水平值下的降噪性能和执行效率均显著优于原 DIP 模型;与现有的主流降噪方法相比, IDIP 降噪模型也具有更好的降噪效果.

关键词: 深度图像先验; 降噪效果; 执行效率; 复杂连接; 混合 Loss

中图分类号: TP391.4

文献标识码: A

文章编号: 0372-2112(2022)07-1573-06

电子学报 URL: <http://www.ejournal.org.cn>

DOI: 10.12263/DZXB.20211323

An Image Denoising Model Using the Improved Deep Image Prior

XU Shao-ping, LI Fen, CHEN Xiao-guo, CHEN Xiao-jun, JIANG Shun-liang

(School of Mathematics and Computer Sciences, Nanchang University, Nanchang, Jiangxi 330031, China)

Abstract: To further improve the execution efficiency and denoising effect of the deep image prior(DIP) denoising model, an improved deep image prior(IDIP) denoising model was proposed by improving the original DIP from three aspects, network architecture, network input, and Loss function. Specifically, considering the network architecture, the simple connection adopted in the encoder-decoder backbone network was promoted with complex connection by adding nonlinear features transferring path, which brings benefit to the information modulation and transmission of features between encoder and decoder at the same level. In aspect of the network input, the random tensor was replaced by the preliminary denoised image with better image quality, and the preliminary denoised image can provide more abundant information to the network model, accelerating the convergence speed of network and improving the execution efficiency. Regarding the Loss function, the preliminary denoised image was added as the second target image to improve the guidance ability of the loss function, improving denoising effect significantly. Extensive experiments show that, the proposed IDIP denoising model significantly outperforms original one at various noise levels in terms of denoising effect and execution efficiency, and it also has a better performance than other state-of-the-art methods with regard to denoising effect.

Key words: deep image prior; denoising effect; execution efficiency; complex connection; hybrid loss

1 引言

近十年来,深度学习因其强大的非线性映射、特征学习与表达能力被广泛地应用于图像降噪领域,并取得了巨大成功^[1,2]. 以 DnCNN(Denoising Convolutional Neural Network)^[3]、FFDNet(Fast and Flexible Denoising

convolutional neural Network)^[4]和 VDN(Variational Denoising Network)^[5]为代表的基于深度神经网络(Deep Neural Network, DNN)的有监督降噪模型在降噪效果上较传统经典方法体现出显著优势. 然而,基于有监督 DNN 降噪模型要想获得令人满意的降噪效果,必须确保有大量的训练数据集,而且这些训练图像集应与待

降噪的噪声图像 y 在图像内容以及受噪声干扰程度上必须近似. 对于任意给定的噪声图像 y , 如果用于训练模型的图像集与噪声图像在受噪声干扰程度和图像内容上相差较大, DNN 降噪模型很难获得最佳的降噪效果. 尤其是在训练图像采集困难的情况下, 甚至连降噪模型都难以获得. 因此, DNN 降噪模型具有数据依赖 (data dependency) 缺陷.

近年来, Dmitry 等人^[6]基于 DNN 网络提出了一种被称为深度图像先验 (Deep Image Prior, DIP) 的新型降噪模型. DIP 模型采用 U 型编码器-解码器 (encoder-decoder) 网络架构作为骨干网络^[6,7]获取噪声图像 y 中的底层统计特征 (low-level statistics), 以随机生成的张量 (tensor) z 作为网络输入, 将网络输出图像与噪声图像 y (作为目标图像) 之间的均方误差 (Mean Square Error, MSE) 定义为 Loss 函数. 以 Loss 函数值最小化为优化目标, 在在线训练模式下通过不断调整网络的参数值, 并采用早停 (early stopping) 技术^[6]适时终止网络训练从而完成降噪任务, 即基于噪声图像 y 获得关于无失真图像 x 的最佳估计 \hat{x} . DIP 降噪模型仅利用噪声图像自身的信息实现降噪, 并不依赖于特定的图像训练数据集 (实际上无需训练图像), 属于无监督降噪模型. 与主流有监督 DNN 网络模型相比, DIP 降噪模型中的网络参数值是与特定噪声图像相匹配的, 不是固定不变的. 所以, DIP 降噪模型最大的优势在于它没有数据依赖问题, 具有一定的生成能力, 不会像有监督降噪模型那样因使用固定不变的网络参数值而在降噪后图像中引入幻影 (hallucinant) 结构, 且图像中某些局部图像细节可以得到更好地复原和保护. 然而, DIP 降噪模型的降噪效果在整个图像全局上并不高, 其降噪效果甚至有时候差于 BM3D 算法^[8]. 为全面提高 DIP 降噪模型的降噪效果和执行效率, 本文从网络结构、网络输入和 Loss 函数三个方面对其进行改进从而获得一种改进的深度图像先验 (Improved Deep Image Prior, IDIP) 降噪模型.

2 IDIP 降噪模型

2.1 改进措施

本文从以下 3 个方面对 DIP 模型进行改进.

(1) 网络结构. 原 DIP 模型采用了 Skip 连接建立编码器-解码器中相同层级特征之间的联系, 虽然可以减少下采样操作所引起的信息损失^[9], 但 Skip 连接过于简单、描述能力不强. 为了实现编码器-解码器同层级之间更为有效信息传递, IDIP 降噪模型通过新增非线性特征路径构成复杂连接的方式, 允许解码器在利用编码器提取的底层图像特征时, 让图像特征能被更为复杂的非线性过程处理, 使得解码器在解析图像特征时

具有更大的灵活性, 从而为提高网络整体的非线性映射能力打下基础.

(2) 网络输入. 原 DIP 降噪模型的输入是随机生成的张量 z , 随机张量 z 并没有提供任何关于降噪的先验信息, 这使得 DIP 网络模型中参数值调整到达收敛点所需要时间非常漫长 (迭代次数多, 训练时间长). 这启示本文将 DIP 模型的输入 (即随机张量 z) 替换为初步降噪图像 \bar{x} , 初步降噪图像 \bar{x} 采用某个主流降噪方法处理给定的噪声图像后获得, 具有较高的图像质量. 这种改进将原 DIP 模型中随机张量 z 与最佳估计图像 \hat{x} 的映射关系改变为初步降噪图像 \bar{x} 与最佳估计图像 \hat{x} , 使得模型映射难度降低.

(3) Loss 函数. 原 DIP 降噪模型采用噪声图像 y 作为目标图像构建 Loss 函数, 导向能力有限. 为解决该问题, 一个自然的想法就是给 Loss 函数中新增一个子项, 增加一个图像质量较高的图像参与构建 Loss 函数以改善其导向能力. 具体地, 利用某个降噪效果比较好的主流降噪模型先对噪声图像进行降噪, 降噪后所获得图像质量大为改进的图像称为初步降噪图像 \bar{x} (作为第二个目标图像). 初步降噪图像 \bar{x} 的图像质量较噪声图像显著得到改善, 故利用它参与构建 Loss 函数可调整 IDIP 网络模型演进方向, 使得网络输出图像 \hat{x}^i 在图像空间中朝着更为合理的方向演进 (即无失真图像 x). 基于上述改进措施, 所提出 IDIP 降噪模型的体系架构如图 1 所示. 下文就具体实现细节进行详细介绍.

2.2 网络结构改进

如图 1 中复杂连接图例所示, E 和 C 分别表示扩展 (expanding) 和联接 (concatenation) 操作. 为了建立同层 $\mathbf{m} \in \mathbb{R}^{C \times H \times W}$ 特征与 $\hat{\mathbf{m}} \in \mathbb{R}^{C \times H \times W}$ 特征之间的复杂连接, 分为以下三个步骤: (1) 通过式 (1) 提取每个通道的统计信息 $f_{\text{pool}}(\cdot)$; (2) 通过式 (2) 中的全连接层 $f_{\text{FC}}(\cdot)$ 使用提取的 $\hat{\mathbf{x}}_{\text{mean}}$ 来挖掘特征通道间相互的关系; (3) 通过式 (3) 将式 (2) 提取的特征与原 DIP 模型简单直通路联合并经卷积层融合后建立复杂连接, 式 (1)~(3) 的具体定义为:

$$\hat{\mathbf{x}}_{\text{mean}} = f_{\text{pool}}(\mathbf{m}) \quad (1)$$

$$\boldsymbol{\rho} = f_{\text{FC}}(\hat{\mathbf{x}}_{\text{mean}}) \quad (2)$$

$$\hat{\mathbf{m}} = \text{Conv}(C(E(\boldsymbol{\rho}), \mathbf{m})) \quad (3)$$

其中, $f_{\text{pool}}(\cdot)$ 表示平均池化操作, $f_{\text{FC}}(\cdot)$ 表示完全连接层, $E(\cdot)$ 表示将 $\boldsymbol{\rho}$ 的空间维扩展到 \mathbf{m} 维, $C(\cdot)$ 是连接操作, $\text{Conv}(\cdot)$ 表示卷积层. 从图 1 中给出的复杂连接实现中可以看出: 复杂连接实际上是在原来简单连接的路径之外, 又拓展了一个非线性映射路径, 是增强扩展版本. 与原 DIP 模型相比, 在 IDIP 降噪模型中使用复杂连接可以缓解特征尺度变化而导致的信息丢失问题, 模

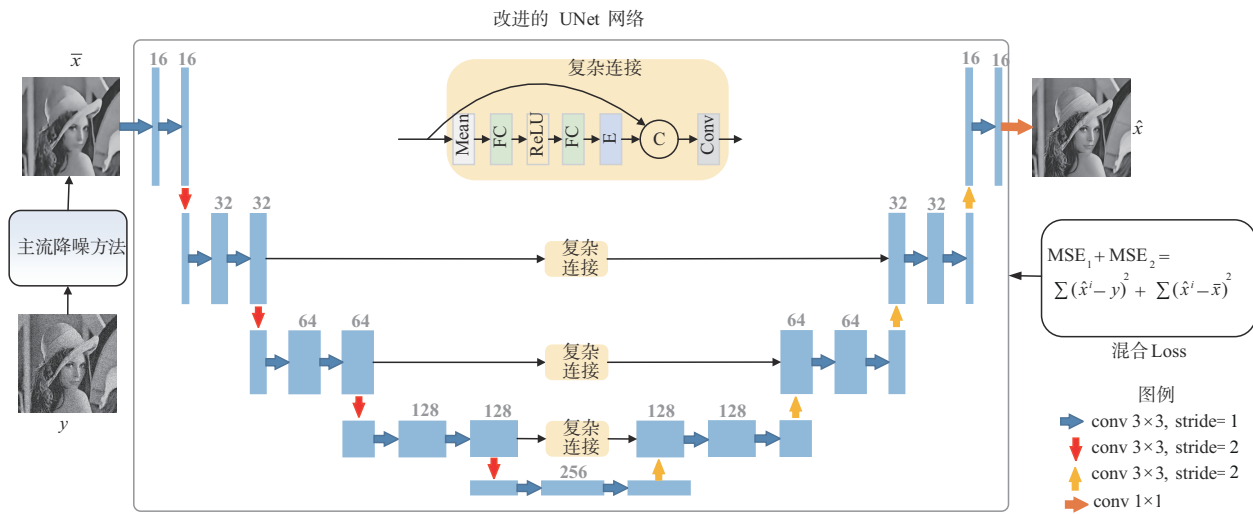


图1 IDIP降噪模型框图

型所需要参数数量仅需增加约2.4%.

为了分析引入复杂连接技术对IDIP降噪模型所带来益处,在10张测试图像集上完成了对比实验(图像受 $\sigma=30$ 中等严重程度噪声干扰),改进后的模型记为IDIP-1(网络的输入和Loss函数的设置与原DIP模型相同).由表1可知:IDIP-1模型的降噪性能在9张图像上

的降噪效果都优于DIP降噪模型,尤其是在图像内容比较复杂Barbara图像上,性能提升达到0.41 dB. 尽管如此, IDIP-1降噪模型在降噪效果上相对于对DIP模型而言,在均值和标准差两项指标上总体提升幅度并不是特别大,这表明对网络结构的改进有一定的效果,为后续网络的进一步改进打下了很好的基础.

表1 骨干网络的改进对降噪性能的影响(dB)

模型	House	Peppers	Barbara	Boat	Cameraman	Couple	Hill	Lena	Man	Monarch	平均值	标准差
DIP	30.84	28.00	24.38	27.31	26.81	27.25	27.89	29.99	27.56	27.27	27.73	1.66
IDIP-1	30.92	28.09	24.79	27.43	26.96	27.26	28.00	30.03	27.61	26.82	27.79	1.61

2.3 网络输入改进

初步降噪图像 \bar{x} 在IDIP模型中作为网络输入替换了随机张量 z . 为分析这种改进对IDIP降噪模型所带来的益处,在常用测试图像上完成了测试. 改进后的模型记为IDIP-2(即网络的输入为预处理图像,骨干网络中采用复杂连接, Loss函数与DIP模型相同). 由表2可知: IDIP-2模型的降噪性能显著优于DIP降噪模型, PSNR均值提升了1 dB以上,同时标准差指标也大幅度减少了; IDIP-2模型的降噪效果已经达到稍优于主流FFDNet降噪模型的水平(提升了0.08 dB),这充分说明了利用初步降噪后图像 \bar{x} 作为网络的输入可以有效提高降噪性能.

2.4 Loss函数改进

IDIP网络采用了以下混合Loss函数:

表2 网络输入和Loss函数的改进对降噪性能的影响(dB)

模型	House	Peppers	Barbara	Boat	Cameraman	Couple	Hill	Lena	Man	Monarch	平均值	标准差
DIP	30.84	28.00	24.38	27.31	26.81	27.25	27.89	29.99	27.56	27.27	27.73	1.66
FFDNet	32.37	29.63	28.70	29.46	28.73	29.22	29.38	31.87	29.11	28.70	29.72	1.25
IDIP-2	32.40	29.59	29.09	29.65	28.40	29.32	29.46	32.06	29.34	28.75	29.80	1.27
IDIP	32.99	30.34	29.40	30.18	29.24	29.76	29.91	32.56	29.67	29.45	30.35	1.26

$$\text{Loss} = \text{MSE}_1 + \text{MSE}_2 = \sum (\hat{x}^i - y)^2 + \sum (\hat{x}^i - \bar{x})^2 \quad (4)$$

其中, \hat{x}^i 表示迭代第 i 次时IDIP网络的输出图像, y 代表噪声图像(第一个目标图像), \bar{x} 表示经过FFDNet模型处理后所获得的初步降噪后图像(即作为第二目标图像). 为了分析混合Loss函数的改进对IDIP模型所带来的益处,将网络输入设置为初步降噪图像,骨干网络采用复杂连接, Loss函数采用混合Loss函数(上述配置即为本文所提出的IDIP模型, Loss函数仅采用噪声目标图像的模型为IDIP-2模型). 由表2可知, IDIP降噪模型的降噪性能明显优于IDIP-2降噪模型, PSNR均值进一步提高了0.45 dB, 标准差也减少了0.01, 此时IDIP降噪模型的降噪效果已经显著超过主流FFDNet降噪模型(作为对照模型), 同时标准差也排名第二, 表明使用初步降噪图像构造混合Loss函数能有效提高降噪效果.

3 实验结果及其分析

3.1 测试环境配置

测试图像集合分别为:(1)常用图像集.由各个文献中被广泛使用的10张测试图像组成,包括House、Peppers、Barbara、Boat、Cameraman、Couple、Hill、Lena、Man和Monarch;(2)BSD图像集,由BSD数据库^[10]随机挑选中的50张图像组成.所有算法都在相同硬件平台(Intel(R) Xeon(R) CPU E5-1603 v4 @ 2.80 GHz RAM 16 GB)和软件环境(Windows10)上运行.

3.2 网络输入的选择

IDIP模型选择利用某个主流降噪方法对噪声图像降噪后所获得的初步降噪图像作为网络的输入来替换DIP降噪模型的随机张量 z .同时,初步降噪图像也作为第二目标图像参与构建Loss函数.为分析使用主流降

噪方法的不同对IDIP模型性能所带来的影响,本文选择了代表性的DnCNN^[3]、BM3D^[8]、NCSR^[11]、WNNM^[12]和FFDNet^[4]5种主流降噪方法参与比较,各种情况下的模型记为IDIP(DnCNN)、IDIP(BM3D)、IDIP(NCSR)、IDIP(WNNM)和IDIP(FFDNet).由表3可知,IDIP降噪模型使用不同的主流降噪方法获得初步降噪图像对最终的降噪效果会产生一定影响(最大差距可达0.18 dB,但是相对于网络输入和Loss函数改进而言幅度要小很多).根据实验数据,使用WNNM和DnCNN也能获得相近的降噪效果.但是,考虑到WNNM算法自身执行效率偏低,而DnCNN降噪模型的通用性存在一定的缺陷(相对于FFDNet而言),故本文最终选定利用FFDNet获得初步降噪图像.下文的实验数据均是在此配置条件下获得的.

表3 在噪声水平 $\sigma=30$ 下不同初步降噪图像对降噪性能的影响(dB)

模型	House	Peppers	Barbara	Boat	Cameraman	Couple	Hill	Lena	Man	Monarch	平均值
IDIP(DnCNN)	32.91	30.41	29.30	30.13	29.23	29.61	29.98	32.58	29.69	29.59	30.34
IDIP(BM3D)	32.63	30.12	29.83	29.86	29.06	29.46	29.82	32.25	29.38	29.33	30.17
IDIP(NCSR)	32.87	29.92	29.82	29.84	29.16	29.40	29.80	32.19	29.44	29.39	30.18
IDIP(WNNM)	33.07	30.16	30.26	30.07	29.14	29.51	29.96	32.28	29.51	29.38	30.33
IDIP(FFDNet)	32.99	30.34	29.40	30.18	29.24	29.76	29.91	32.56	29.67	29.45	30.35

3.3 降噪效果

首先,在图2所示的降噪效果视觉对比中,在所有参与比较的方法中,仅有IDIP和DIP降噪模型很好地保留了图像中较弱的边缘细节(框内箭头所指处),而其他对比方法几乎已经将Lena图像中微弱的

边缘细节模糊化了,这表明IDIP和DIP降噪模型这类基于深度图像先验构建的降噪模型更善于利用图像自身的内部结构信息实现更好的降噪效果,而从图像整体降噪效果而言,IDIP降噪模型要远优于DIP模型.

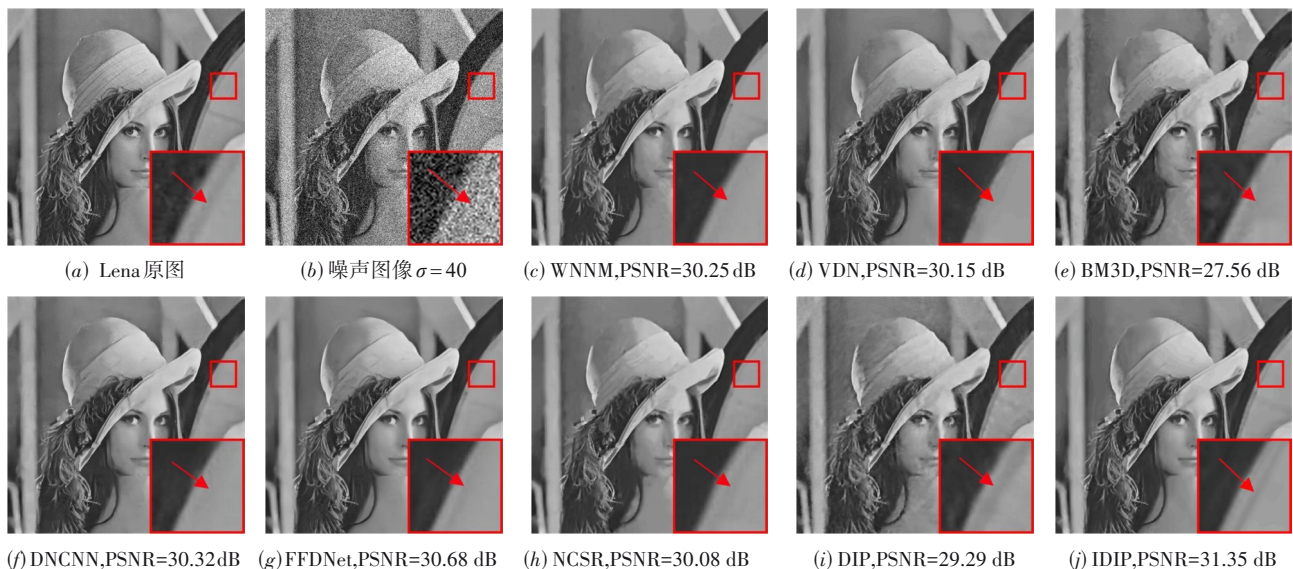


图2 各对比方法在Lena图像上的降噪效果对比

其次,由表 4 和表 5 可知:在常用和 BSD 测试图像集合上, IDIP 降噪模型在各噪声水平值 σ 下的降噪效果均显著优于各对比方法,且总体上是稳健的.

表 4 各对比方法在常用图像上所获得的 PSNR 均值比较 (dB)

σ	10	20	30	40	50	60	平均值
FFDNet	34.83	31.40	29.72	28.50	27.60	26.85	29.82
BM3D	34.82	31.44	29.59	28.13	27.54	26.37	29.65
NCSR	34.81	31.38	29.43	28.06	27.02	26.08	29.46
WNNM	34.94	31.61	29.80	28.48	27.51	26.68	29.84
DnCNN	34.94	31.69	29.83	28.52	27.56	26.65	29.87
DIP	32.41	29.32	27.73	26.60	25.56	24.77	27.73
IDIP	35.68	32.13	30.35	29.16	28.13	27.22	30.45

3.4 执行效率

由表 6 可知:达到 DIP 降噪模型所获得的最优 PSNR 值(表 6 中第一行数据,各图像受噪声水平值为

表 6 DIP 与 IDIP 降噪模型执行迭代次数和执行时间的比较

图像	Peppers	Barbara	Boat	Hill	Lena
PSNR	26.67 dB	23.37 dB	26.34 dB	26.94 dB	29.29 dB
DIP	1 048(117.80 s)	2 132(466.80 s)	2 041(444.95 s)	1 802(397.20 s)	1 775(387.42 s)
IDIP	245(34.38 s)	159(40.25 s)	203(48.55 s)	245(59.40 s)	255(60.69 s)

4 结束语

本文对原 DIP 降噪模型的网络结构、网络输入和 Loss 函数三个方面进行了全面改进,所获得的 IDIP 降噪模型在降噪效果和执行效率两个方面均显著优于原 DIP 降噪模型.相对于其他当前主流的降噪方法来说, IDIP 降噪模型在降噪效果上有显著优势.对于 IDIP 模型执行效率相对于主流降噪方法仍然偏低的问题,未来可考虑通过研究快速初始化网络参数值的方法,在 Meta-Transfer Learning 技术^[13]支撑下让 IDIP 降噪模型经过少量迭代训练达到收敛,从而进一步提升执行效率.

参考文献

- [1] LI J X, GUO X B, LU G M, et al. DRPL: Deep regression pair learning for multi-focus image fusion[J]. IEEE Transactions on Image Processing, 2020, 29(3): 4816-4831.
- [2] WANG C H, HOU B C, SHI J, et al. Uniformity evaluation of temperature field in an oven based on image processing[J]. IEEE Access, 2020, 8(1): 10243-10253.
- [3] 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.

表 5 各对比方法在 50 张 BSD 图像上所获得的 PSNR 均值比较 (dB)

σ	10	20	30	40	50	60	平均值
FFDNet	33.61	30.02	28.16	26.93	26.04	25.34	28.25
BM3D	33.41	29.56	27.60	26.26	25.40	24.73	27.83
NCSR	33.42	29.58	27.60	26.27	25.35	24.59	27.80
WNNM	33.53	29.73	27.81	26.54	25.63	24.92	28.03
DnCNN	33.69	30.06	28.14	26.87	25.99	25.23	28.33
DIP	31.27	27.97	26.22	25.09	24.18	23.43	26.36
IDIP	35.02	30.80	28.84	27.56	26.55	25.76	29.09

$\sigma=40$ 的高斯噪声干扰), IDIP 模型所需的迭代次数和执行时间均比 DIP 模型减少了 80% 左右,远小于 DIP 模型,执行效率得到显著提升.需要说明的是:相对于 DIP 模型,尽管 IDIP 模型的执行效率获得了显著提升,但是相对于当前主流的有监督降噪模型而言,由于网络参数值仍然需要较多次数的迭代更新才能调整到最佳值, IDIP 降噪模型的执行效率仍然处于劣势.

- [4] 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.
- [5] YUE Z S, YONG H W, ZHAO Q, et al. Variational denoising network: Toward blind noise modeling and removal [C]//Advances in Neural Information Processing Systems. Vancouver: Neural Information Processing Systems Foundation, 2019: 1-12.
- [6] DMITRY U, VEDALDI A, VICTOR L. Deep image prior [J]. International Journal of Computer Vision, 2020, 128(8): 1867-1888.
- [7] VIGNESH R, BALAMURALI M, SHARATH M S, et al. Deep network for capacitive ECG denoising[C]//Medical Measurements and Applications. Istanbul: IEEE, 2019:1-6.
- [8] 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.
- [9] FU J, DONG J B, ZHAO F. A deep learning reconstruction framework for differential phase-contrast computed tomography with incomplete data[J]. IEEE Transactions on Image Processing, 2020, 29(1): 2190-2202.
- [10] ARBELAEZ P, MAIRE M, FOWLKES C, et al. Contour

- detection and hierarchical image segmentation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2011, 33(5): 898-916.
- [11] 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.
- [12] GU S H, ZHANG L, ZUO W M, et al. Weighted nuclear norm minimization with application to image denoising [C]//Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Columbus: IEEE Computer Society, 2014: 2862-2869.
- [13] SOH J W, CHO S, CHO N I. Meta-transfer learning for zero-shot super-resolution[C]//Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Seattle: IEEE Computer Society, 2020: 3513-3522.

作者简介



徐少平 男,1976年5月出生于江西省九江市. 博士,南昌大学数学与计算机学院计算机科学与技术系教授,博士生导师. 主要研究方向为图形图像处理、机器视觉、虚拟手术仿真等.
E-mail: xushaoping@ncu.edu.cn



李 芬 女,1995年8月出生于江西省南昌市. 现为南昌大学硕士研究生,主要研究方向为图像处理与计算机视觉.
E-mail: 411014519034@email.ncu.edu.cn