

# 双通道深度图像先验降噪模型

徐少平<sup>1</sup>, 肖楠<sup>1</sup>, 罗洁<sup>2</sup>, 程晓慧<sup>1</sup>, 陈晓军<sup>1</sup>

(1. 南昌大学数学与计算机学院, 江西南昌 330031; 2. 南昌大学附属感染病医院, 江西南昌 330006)

**摘要:** 相对于采用固定网络参数值的有监督深度降噪模型而言, 无监督的深度图像先验 (Deep Image Prior, DIP) 降噪模型更具灵活性和实用性. 然而, DIP 模型的降噪效果远低于有监督降噪模型 (尤其是在处理人工合成噪声图像时). 为进一步提升 DIP 降噪模型的降噪效果, 本文提出了双通道深度图像先验降噪模型. 该降噪模型由噪声图像预处理、在线迭代训练和图像融合 3 个模块组成. 首先, 利用 BM3D 和 FFDNet 两种经典降噪方法对给定的噪声图像进行预处理, 得到 2 张初步降噪图像, 然后, 将原 DIP 单通道逼近目标图像架构拓展为双通道工作模式. 其中, 第一通道以 FFDNet 初步降噪图像和噪声图像为双目标图像, 第二通道则以 BM3D 预处理图像和噪声图像为双目标图像. 在此基础上, 按照标准的 DIP 在线训练方式让 DIP 网络输出图像在两个通道上分别逼近各自的目标图像, 同时依据基于边缘能量定义的伪有参考图像质量评价价值适时终止迭代过程, 从而获得 2 张中间生成图像. 最后, 使用结构化图块分解融合算法将两张中间生成图像融合并作为最终的降噪后图像. 实验数据表明, 在合成噪声图像上, 本文提出的双通道深度图像先验降噪模型在各个噪声水平上显著优于原 DIP 及其他无监督降噪模型 (提升了约 2.2 dB), 甚至逼近和超过了新近提出的主流有监督降噪模型, 这充分表明了本文提出的改进策略的有效性; 在真实噪声图像上, 本文提出的降噪模型优于排名第二的对比降噪方法约 2 dB, 展现出其在实际应用场景下独有的优势.

**关键词:** 深度图像先验; 双通道逼近策略; 预处理图像; 自动迭代终止; 图像质量评价; 图像融合

**基金项目:** 国家自然科学基金 (No.62162043)

**中图分类号:** TP391.4

**文献标识码:** A

**文章编号:** 0372-2112(2024)01-0058-11

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

**DOI:** 10.12263/DZXB.20221012

## Dual-Channel Deep Image Prior for Image Denoising

XU Shao-ping<sup>1</sup>, XIAO Nan<sup>1</sup>, LUO Jie<sup>2</sup>, CHENG Xiao-hui<sup>1</sup>, CHEN Xiao-jun<sup>1</sup>

(1. School of Mathematics and Computer Sciences, Nanchang University, Nanchang, Jiangxi 330031, China;

2. Infectious Disease Hospital Affiliated to Nanchang University, Nanchang, Jiangxi 330006, China)

**Abstract:** Compared with the supervised deep learning-based denoising models adopting fixed network parameter values, the unsupervised deep image prior (DIP) is more flexible and practical than those supervised denoising models. However, the overall performance of the unsupervised DIP model is far lower than those supervised ones, especially when it is easy to obtain training data such as synthetic noisy images. To improve the performance of the DIP model, in this paper we propose a denoising model called dual-channel deep image prior (DCDIP). The DCDIP model consists of three modules: preprocessing, online training, and image fusion. First, two classical denoising methods, i. e., BM3D and FFDNet, are used to preprocess a given noisy image to obtain two corresponding initial denoised images. Then, the original DIP model with single channel approximation target image architecture is expanded to a dual-channel working manner. In the first channel, the initial denoised image obtained with FFDNet and the noisy image are taken as the dual-target images. Similarly, in the second channel, the initial denoised image obtained with BM3D and the noisy image are taken as the dual-target images. On this basis, according to the original DIP online training manner, the output image of the DCDIP is iteratively trained to approach the respective target images in the two channels, and the iterative process is terminated in time with the proposed pseudo reference image quality evaluation index based on the edge energy. As such, we can obtain two intermediate generated images with high quality. Finally, the two intermediate generated images are fused as the final denoised im-

age by using the structural patch decomposition (SPD) fusion algorithm. The experimental results show that the proposed DCDIP significantly outperforms the original DIP model and unsupervised ones by about 2.2 dB across different noise levels. In addition, it even approaches and surpasses the recently proposed supervised denoising models, demonstrating the effectiveness of our improvement strategy. On the real-world noisy image, the proposed DCDIP outperforms the second-ranked competing denoising method by about 2 dB, which verifies its unique advantages in the practical application scenarios. The performance improvement is mainly due to the hybrid use of internal and external image prior-based denoising methods under the dual-channel DIP framework.

**Key words:** deep image prior; dual-channel approximation strategy; initial denoised image; automatic iteration termination; image quality assessment; image fusion

**Foundation Item(s):** National Natural Science Foundation of China (No.62162043)

## 1 引言

图像降噪是图像处理领域中的一项重要基础性工作,其性能直接影响后续图像处理任务能否顺利执行<sup>[1-3]</sup>. 噪声图像通常可以建模为 $y = x + n$ . 其中, $y$ 是噪声图像, $x$ 是无失真图像, $n$ 是噪声. 在研究中, $n$ 通常被假定为满足均值为零的高斯分布 $N(0, \sigma^2)$ ,其中,标准差 $\sigma$ 反映了图像受噪声干扰的严重程度<sup>[4]</sup>. 形式化地,图像降噪问题就是研究如何基于给定的噪声图像 $y$ ,给出关于无失真图像 $x$ 的最佳估计 $\hat{x}$ . 根据所利用的先验信息的不同,现有的图像降噪方法可以大致分为基于图像内部先验和基于外部先验降噪两大类<sup>[5]</sup>.

近年来,随着深度学习技术的快速发展及其在底层视觉处理领域的成功应用,基于深度学习的降噪模型被大量提出<sup>[6]</sup>. 这类降噪模型通常需要先准备海量的(clean-noisy)训练对,经过训练得到降噪网络模型后实现图像降噪,因此可归类为使用外部图像先验的降噪方法. 例如,2017年Zhang等人<sup>[7]</sup>提出的DnCNN(Denoising Convolutional Neural Network)算法可以被认为是这个领域里程碑式的工作. DnCNN降噪模型首次使用深度网络架构(deep architecture)、残差学习(residual learning)和批量归一化(batch normalization)技术将降噪模型的降噪性能大幅提高. 之后,Zhang等人<sup>[8]</sup>在DnCNN网络结构的基础上,通过引入噪声水平映射图(Noise Level Map, NLM)作为模型的辅助输入通道,提出了FFDNet(Fast and Flexible Denoising convolutional neural Network)降噪模型,使其能够仅利用一个降噪模型就可以处理较广噪声水平值范围的降噪问题,且降噪性能还略微比DnCNN有所改进. 近年来,研究者们基于经典的DnCNN和FFDNet降噪模型相继提出了很多改进模型. 例如,Zhang等人<sup>[9]</sup>提出的DRUNet降噪模型利用深度学习和传统的迭代优化方法提升降噪性能. 近年来,为了解决卷积神经网络(Convolutional Neural Network, CNN)视野受限和对输入图像内容适应性差等问题,原本在机器翻译任务中的Transformer结构

被引入底层视觉图像处理中. Transformer的自注意力机制能够很好地捕获图像中长距离像素点之间的相互关系,拥有比CNN网络结构更大的处理视野,从而使网络模型的降噪性能得到大幅度提高. 例如,Liang等人<sup>[10]</sup>提出的SwinIR模型就是典型的代表. SwinIR模型的降噪过程可以分为浅层特征提取、深层特征提取、高质量图像重建三大阶段. 在浅层特征提取模块利用卷积层提取浅层特征,并将浅层特征直接传输到重构模块,以保留低频信息. 深层特征提取模块由多个RSTB(Residual Swin Transformer Block)组成. 在RSTB中,多个STL(Swin Transformer Layers)层实现了注意力机制与交叉窗口的交互. SwinIR的降噪性能比基于普通CNN卷积结构构建的降噪模型有较大幅度的提高. Zamir等人<sup>[11]</sup>提出了Restormer(Restoration Transformer)降噪模型. 该降噪模型通过对Transformer的进一步改进,获得了比原Transformer更高的执行效率. 总之,上述各类利用外部先验的降噪模型需要大量的(clean-noisy)训练对才能构建模型,该模型从本质上来讲属于有监督降噪模型. 这些模型的降噪性能受限于训练数据集. 当噪声图像的图像内容或者噪声分布模式显著异于训练集时,上述有监督降噪模型的降噪性能将会有不同程度的下降,这就是所谓的数据依赖(data dependency)或者数据偏见(data bias)问题.

Ulyanov等人<sup>[12]</sup>提出了深度图像先验(Deep Image Prior, DIP)的无监督降噪模型. 该模型开创性地利用深度网络结构本身作为隐式正则约束,基于给定的单张噪声图像(作为目标图像)进行在线迭代训练学习(网络参数在迭代过程中不断更新),在让网络输出图像逼近目标图像过程中实现了获得清晰降噪后图像的目的. 由于仅利用噪声图像本身就可以实现降噪,故其属于无监督降噪模型(主要利用噪声图像内部先验信息),没有数据依赖问题. DIP模型在处理降噪问题时非常灵活,虽然其网络输入是随机生成的,但它通过调整网络参数值,利用深度网络强大的非线性映射能力总

能让网络输出图像逼近给定的噪声图像. 因此, 所获得的降噪模型是所谓的特定图像网络(image-specific network), 与噪声图像相适应, 能在图像某些局部区域很好地保护细节. 然而, 从图像整体的降噪效果来看, DIP降噪模型却落后于有监督降噪模型, 甚至在一些时候(尤其对于合成噪声图像)其降噪性能甚至会低于早期经典的BM3D算法<sup>[13]</sup>. 原因在于, DIP降噪模型本质上属于内部图像先验模型, 可利用的信息比较有限. 由于噪声图像的图像质量一般较差, 故以噪声图像作为目标图像使DIP网络的搜索空间变大, 收敛困难, 且容易出现过拟合(over fitting)问题.

为进一步提升DIP模型的降噪效果且同时保持其灵活与实用性, 本文提出了一种采用双通道的深度图像先验(Dual-Channel Deep Image Prior, DCDIP)降噪模型. 具体的创新工作有:(1)利用具有较高图像质量的预处理图像(利用BM3D<sup>[13]</sup>或者FFDNet<sup>[8]</sup>处理给定的噪声图像获得)作为新的目标图像并配合原噪声图像构成双目标图像, 显著提升DIP的降噪性能;(2)采用双通道工作方式使DIP网络输出图像分别逼近不同的双目标图像, 以新提出的基于图像边缘能量定义的伪有参考图像质量评价(Image Quality Assessment, IQA)值为迭代终止判定依据, 避免DIP模型过拟合问题, 从而解决了约束DIP降噪性能的瓶颈问题, 确保DIP模型在逼近目标图像过程中总能获得具有较好图像质量的中间生成图像(intermediate generated images), 此时中间生成图像的图像质量已优于预处理图像;(3)基于图块的结构化分解SPD(Structural Patch Decomposition)融合算法<sup>[14]</sup>将双通道上分别获得的两张中间生成图像融合作为最终的降噪后图像, 此时降噪后图像的图像质量得到再次提升. 大量实验数据表明, 在合成噪声图像上, 本文提出的DCDIP降噪模型的降噪性能显著优于原DIP模型, 也远远高于其他同类型的无监督降噪模型, 甚至逼近和超过了新近提出的主流有监督降噪模型Restormer. 而在真实噪声图像上, 相对于现有方法, 所提的DCDIP降噪模型有着显著优势, 展现出良好的降噪效果且更具灵活性与实用性.

## 2 相关工作

DIP降噪模型利用U形深度卷积网络(即UNet骨干网络)<sup>[12]</sup>建立从随机输入 $\mathbf{z}$ 到目标图像(噪声图像)之间的非线性映射关系, 通过不断的在线训练学习(调整网络参数)使DIP网络输出图像逼近目标图像, 在经过一定次数的迭代训练后, 终止训练过程并以终止时的网络输出图像作为降噪后图像. 具体地, 在线训练开始时, 网络的输入为随机初始化的 $\mathbf{z}$ , 网络模型参数 $\theta$ 也被随机初始化. 然后进入在线迭代训练过程, 通过构建损

失函数 $E(\mathbf{x}, \mathbf{y})$ 将图像降噪问题转化为求解最小化损失函数问题. 这里,  $\mathbf{x} = f_{\theta}(\mathbf{z})$ 表示随机输入 $\mathbf{z}$ 在网络作用下的输出图像,  $\mathbf{y}$ 则为噪声图像. DIP模型的在线训练过程可以表示为

$$\theta^* = \arg \min_{\theta} E(f_{\theta}(\mathbf{z}), \mathbf{y}) \quad (1)$$

一般来说, 损失函数被定义为均方误差(Mean Square Error, MSE), 即

$$E(f_{\theta}(\mathbf{z}), \mathbf{y}) = \|f_{\theta}(\mathbf{z}) - \mathbf{y}\|^2 \quad (2)$$

DIP网络模型是过参数化(over-parameterization)的, 模型经过充分多次的迭代训练后, 网络输出图像将最终收敛于目标图像, 存在过拟合的问题, 这显然并不是设计者想要的结果, 且也不能实现降噪的目的. Ulyanov等人<sup>[12]</sup>已经证实: DIP网络对噪声有一定的阻抗现象, 即网络输出图像首先逼近的是噪声图像的低频部分, 然后才是高频部分(纹理细节和噪声). 因此, DIP模型采用了所谓的早停(early stopping)技术, 在迭代训练中某个步骤终止训练, 以此时的网络输出图像作为降噪后图像, 此时的网络模型参数记为 $\theta^*$ . 因为不同的噪声图像受噪声干扰的严重程度不同且图像内容也各异, 对于不同的噪声图像, DIP网络最优迭代终止步数是不一样的. 然而, 由于缺乏迭代终止依据, Ulyanov等人<sup>[12]</sup>最终采用固定迭代步数作为迭代终止条件, 具体的迭代步数则以验证图像集上测得的最优值确定. 固定迭代步数作为迭代终止条件虽然有效, 但这以牺牲一定降噪性能为代价, 并不能保证降噪后图像的图像质量总是最优的. 需要说明的是, DIP模型的降噪效果其实并不十分理想, 很多时候甚至赶不上早期经典的BM3D算法. 总之, DIP模型是一种通过逼近目标图像实现降噪的无监督降噪方法, 目标图像在迭代训练过程中起着重要的导向作用. 目标图像从某种意义上来说, 约束着DIP网络输出图像的搜索空间<sup>[15]</sup>. 当噪声图像受噪声干扰比较严重时, 图像质量本身比较低, 因此很难约束网络输出图像在合理图像空间中进行搜索, 所以最终导致所获得的降噪后图像的整体质量比较差.

## 3 DCDIP降噪模型

### 3.1 基本思想

由上文可知, 导致DIP降噪模型整体降噪效果较差的原因在于: 目标图像的图像质量比较低和缺乏自动迭代终止判定依据. 为此, 本文提出以下改进策略.

(1)双目标图像. 本文将目标图像由给定噪声图像替换为由某个主流降噪方法处理过后的降噪后图像(称为预处理图像). 预处理图像含有的噪声相较噪声

图像已经大大减少,图像质量更高,在图像空间(image space)中距离无失真图像也更近.因此,以预处理图像作为新的目标图像,可以很好地约束搜索空间,更有利于引导网络参数的更新调整.需要特别说明的是,其实噪声图像并非完全无用,毕竟在原DIP模型框架下仅用噪声图像也能在很大程度上实现降噪.另外,在预处理图像中,因图像降噪算法设计实现的同时也会不可避免地导致图像某些区域存在过度光滑的情况,破坏原始图像细节信息.而噪声图像虽然受噪声图像干扰,图像中相当多的区域仍有可能并未被噪声严重干扰,甚至未受干扰.这样,噪声图像正好与预处理图像形成互补,故本文最终选择使预处理图像配合噪声图像同时构成双目标图像.但随着噪声水平值的提升,噪声图像所占的权重应适当下调.双目标图像逼近方案可以获得比单目标图像更好的降噪效果,相关实验参见第4.2节的消融实验.

(2)双通道逼近.截至目前,尚不存在单个降噪方法能在各种噪声和图像内容条件下具有一致性压倒其他降噪方法的性能,但各个降噪方法之间存在互补性<sup>[16,17]</sup>.为此,本文考虑基于DIP模型且利用现有两个降噪方法之间的互补性构建更为鲁棒的双通道DCDIP降噪方法.具体地,将网络的输出单通道改为双通道模式,同时学习逼近来自两个不同主流降噪方法所处理的预处理图像(外加噪声图像),这样两个逼近通道各自所获得的输出图像(称为中间生成图像)具有较高的图像质量且相互之间还有互补性,这样将两张中间生成图像融合后可以得到更好的降噪效果,性能也更为鲁棒.

(3)伪有参考图像质量客观评价指标.Ulyanov等人<sup>[12]</sup>采用固定迭代次数作为迭代终止条件并不能保证DIP模型总能获得最佳的降噪效果.为解决这个问题,本文基于图像边缘能量定义提出一种新的伪有参考图像质量评价客观指标,通过监控输出图像的图像质量评价价值变化情况,在其值出现极大值时自动终止DIP模型迭代训练过程,从而确保输出图像的图像质量.

### 3.2 执行流程

如图1所示,DCDIP的整体流程共分为预处理图像生成、在线迭代训练和图像融合3个阶段.

(1)预处理图像生成阶段.本文提出从外部图像先验和内部图像先验两大类降噪方法中各选择一个经典降噪方法生成2张预处理图像.具体地,使用FFDNet与BM3D降噪方法各对噪声图像进行预处理,生成相应的初步降噪图像.之前的研究工作<sup>[17,18]</sup>表明,FFDNet与BM3D之间具有很好的互补性且两者的预处理图像融合后的图像质量能达到和超过当前主流降噪方法的水平.原因在于:FFDNet是基于深度学习的降噪模型,

是利用海量的外部训练图像获得的,模型利用的是外部图像先验信息.它对自然图像中常见内容有很好的处理能力,但对图像中较为少见的内容(训练图像集中没有的图像内容)则处理能力较弱.而BM3D基于图像的非局部相似性与稀疏性,利用内部图像信息,对特定待降噪图像中少见但重复出现的内容具有较好的处理效果<sup>[17]</sup>.因此,本文分别从内、外两种图像先验降噪方法中提出经典的FFDNet和BM3D用于生成预处理图像,将这两张预处理图像作为迭代训练阶段中的目标图像.

(2)在线迭代训练阶段.本文采用标准的UNet网络<sup>[12]</sup>作为骨干网络,网络输入为随机变量 $z$ ,目标图像为第一阶段生成的初步降噪图像与噪声图像.具体地,本文将DIP网络的输出拓展为2个逼近通道,在这两个通道上同时完成标准的DIP迭代逼近过程.其中一个通道上的双目标图像为FFDNet预处理图像 $x_{\text{FFDNet}}$ 和噪声图像;另外一个通道上的双目标图像为BM3D预处理图像 $x_{\text{BM3D}}$ 和噪声图像.在线迭代训练过程中,基于迭代终止条件(伪有参考图像质量值极大值时)最终在两个通道上分别输出1张中间生成图像,共获得两张中间生成图像.需要说明的是,理论上,获得两张中间生成图像无需将DIP网络的输出层扩展为双通道工作模式,仅需要单独执行DIP两次(每次以不同的预处理图像和噪声图像作为双目标图像).但DIP降噪模型本身的执行时间比较长,单独执行DIP两次迭代训练的过程会造成执行时间近乎翻倍.故本文所采用的双通道工作模式可以较小的时间代价同时获得两张中间生成图像,具体见第4.2节的消融实验.

(3)融合阶段.将两个通道上所获的两张中间生成图像通过图像融合方法获得最终的降噪图像.具体地,采用原本用于多曝光图像融合(Multi-Exposure image Fusion,MEF)的SPD融合算法完成融合<sup>[14]</sup>.两张中间生成图像的互补性体现在局部对应位置像素点之间存在的细微差别.与其他融合方法相比,SPD算法能更有效地保护这种图像的局部结构信息,故被选用.

### 3.3 双通道逼近

如图1所示,在中间图像生成阶段,本文使用的骨干网络为UNet网络.网络的输入是一张与噪声图像相同大小且随机初始化的图像 $z$ .输入图像 $z$ 经过5次下采样后经过跳跃连接(skip connection)进入上采样过程.在网络右侧,包括5次上采样过程,在每次上采样过程结束后,对应相同层次的下采样图像块会经过跳跃连接与上采样输出图像进行连接,共同作为输入进入下一个上采样过程.在UNet网络图中蓝色矩形之间箭头为跳跃连接,跳跃连接由卷积层(convolutional layer)、批量正则化层(batch normalization layer)、激活层

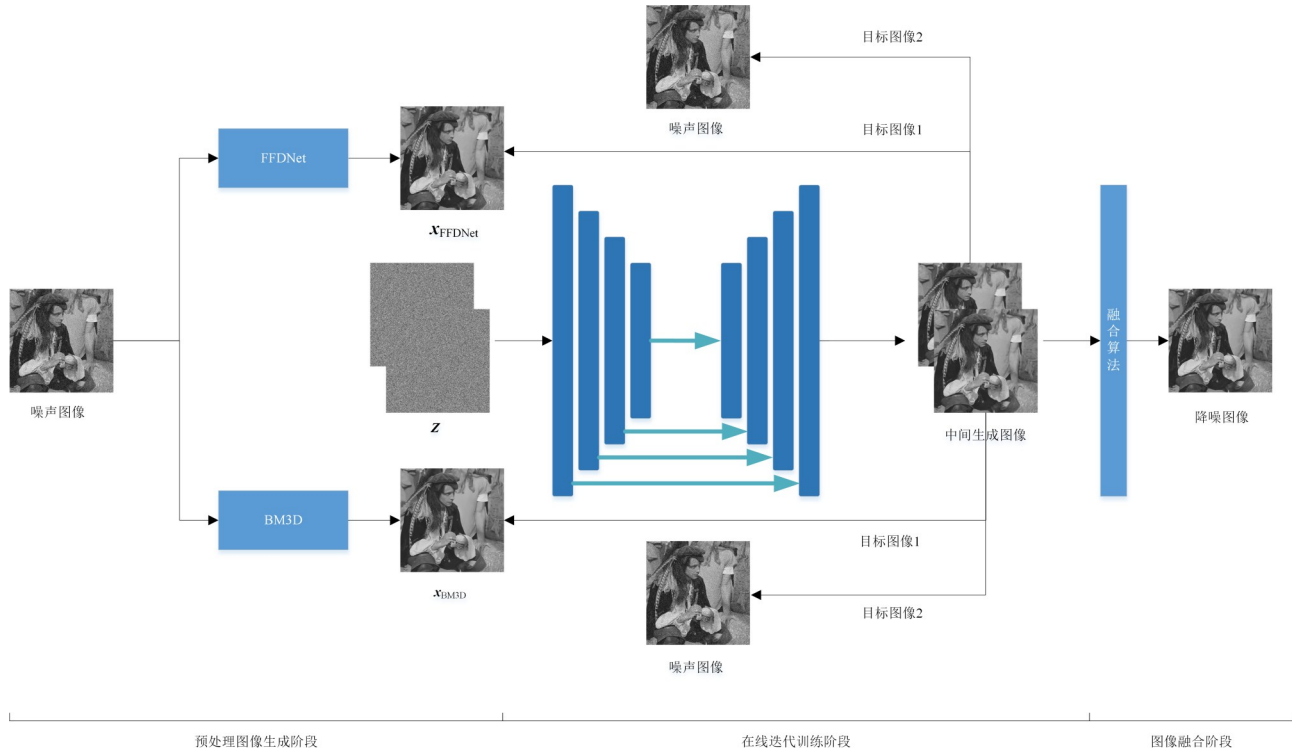


图1 所提出的DCIDIP降噪模型整体执行流程框图

(activation layer)组成. 与原DIP最大的区别在于网络输出为2个通道,分别逼近不同的目标图像(图1以灰度图像为例,故通道数为2,彩色图像时为6通道). 为训练DCIDIP网络,在第1个通道上定义的Loss函数为

$$L_{\text{FFDNet}} = \|f^1_{\theta}(\mathbf{z}) - \mathbf{x}_{\text{FFDNet}}\|^2 + \omega \|f^1_{\theta}(\mathbf{z}) - \mathbf{y}\|^2 \quad (3)$$

其中,  $f^1_{\theta}(\mathbf{z})$ 表示在第一通道上网络输出的中间生成图像,FFDNet预处理图像  $\mathbf{x}_{\text{FFDNet}}$  和噪声图像  $\mathbf{y}$  构成了双目标图像. 由式(3)可知,损失函数  $L_{\text{FFDNet}}$  包括两个部分. 第一部分为网络输出图像  $f^1_{\theta}(\mathbf{z})$  与预处理图像  $\mathbf{x}_{\text{FFDNet}}$  的均方误差,第二部分为网络输出图像  $f^1_{\theta}(\mathbf{z})$  与噪声图像  $\mathbf{y}$  的均方误差. 这里仍然保留噪声图像作为目标图像,主要目的在于最大限度地利用噪声图像中未受噪声干扰的信息. 然而,随着噪声水平值的升高,噪声图像中可以利用的未受噪声干扰的信息逐渐减少. 因此设计了权重系数  $\omega$  以调节第二部分损失函数占整体损失函数的权重,权重系数  $\omega$  被定义为

$$\omega = \begin{cases} \frac{(100 - \sigma)}{100}, & \sigma < 100 \\ 0, & \sigma \geq 100 \end{cases} \quad (4)$$

其中,  $\sigma$  表示噪声水平值. 类似地,在以BM3D预处理图像为目标图像的另一个通道上,损失函数定义为

$$L_{\text{BM3D}} = \|f^2_{\theta}(\mathbf{z}) - \mathbf{x}_{\text{BM3D}}\|^2 + \omega \|f^2_{\theta}(\mathbf{z}) - \mathbf{y}\|^2 \quad (5)$$

其中,  $f^2_{\theta}(\mathbf{z})$  表示在第二通道上获得的中间生成图像,BM3D预处理的图像  $\mathbf{x}_{\text{BM3D}}$  和噪声图像  $\mathbf{y}$  构成了双目标

图像. 这样,总的Loss函数定义为

$$L_{\text{total}} = L_{\text{FFDNet}} + L_{\text{BM3D}} \quad (6)$$

### 3.4 迭代训练自动终止

本文基于图像边缘强度定义的伪有参考图像质量评价价值来确定DIP网络输出图像在逼近目标图像的迭代过程中某一步是否达到了最佳图像质量,这要求先计算给定图像的边缘强度. 首先,利用Daubechies 9/7滤波器将图像分解为LL,HL,LH和HH子带滤波图像(字母组合中第一个字母表示水平滤波类型,第二个字母表示垂直方向上的滤波类型,其中L表示低通滤波,H表示高通滤波)<sup>[19,20]</sup>,共在3个尺度上分解图像. 然后,通过计算子带中小波系数能量来度量图像边缘强度. 以计算LH子带的能量为例:

$$E[\text{LH}_n] = \log \left( 1 + \frac{1}{N_n} \sum_{i,j} [\text{LH}_n(i,j)]^2 \right) \quad (7)$$

其中,  $E[\text{LH}_n]$  表示尺度  $n$  上LH子带的能量,  $N_n$  表示了  $\text{LH}_n$  子带中小波系数的个数,  $i$  和  $j$  表示像素点在图像中的位置.  $\text{HL}_n$  与  $\text{HH}_n$  子带上的计算方法与  $\text{LH}_n$  类似,这样尺度  $n$  上总能量的计算方法为

$$E_n = (1 - \alpha) \left( \frac{E[\text{LH}_n] + E[\text{HL}_n]}{2} \right) + \alpha E[\text{HH}_n] \quad (8)$$

其中,权重系数  $\alpha$  被设置为0.8是为了突出水平和垂直方向上均为高通滤波的HH子带重要性. 最终,整个图像的边缘强度指标  $S$  可以由各个尺度上的能量非线性

加权获得:

$$S = \sum_{n=1}^3 2^{3-n} E_n \quad (9)$$

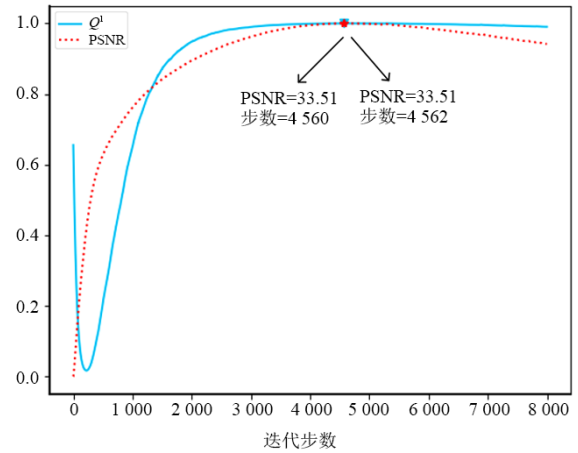
理论上讲,图像质量评价方法根据是否有参考图像(reference image)参与可以分为无参考和有参考两种实现方法.就本文图像降噪任务而言,参考图像正是求解的目标,故无参考图像质量方法最贴近应用实际,但其构造的方法和过程比较复杂;有参考图像质量评价实现方法则相对简单,但需要提供无失真的参考图像.为此,本文将初步降噪图像(已经具有较高的图像质量)认定为伪无失真的参考图像,其图像边缘强度值与无失真图像近似.这样就可以利用式(9)(类似于常用的Structural SIMilarity(SSIM)指标定义形式<sup>[21]</sup>)定义DIP模型输出图像的伪有参考图像质量指标 $Q^b$ :

$$Q^b = \frac{2S(\mathbf{x}_{\text{FFDNet}})S(f_{\theta}^b(\mathbf{z})) + C}{S(\mathbf{x}_{\text{FFDNet}})^2 + S(f_{\theta}^b(\mathbf{z}))^2 + C} \quad (10)$$

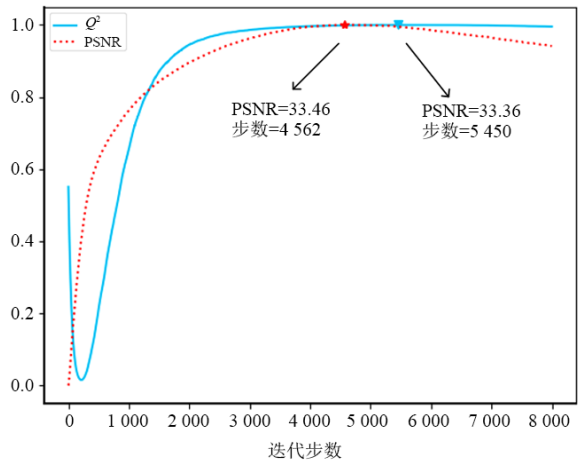
其中, $\mathbf{x}_{\text{FFDNet}}$ 表示利用FFDNet模型获得的预处理图像, $f_{\theta}^b(\mathbf{z})$ 表示在迭代训练过程中第**b**个通道上( $b=1$ 表示以FFDNet预处理图像和噪声图像为双目标图像的第一逼近通道, $b=2$ 表示以BM3d预处理图像和噪声图像为目标图像的第二逼近通道)生成的中间生成图像, $C$ 为可以忽略的任意小常数,具体值设置为0.000 1,函数 $S(\cdot)$ 表示给定图像的边缘强度值(基于式(10)进行计算).理论上讲,在整个迭代过程中, $Q^b$ 值在刚开始时并不稳定,属于Burn-in阶段(大概需3 000步,即小于3 000步不必计算 $Q^b$ 值).随着迭代步数增加,网络输出图像的质量逐渐提高,也越来越接近无失真图像,这时 $Q^b$ 逐渐逼近1,其最大值所对应的迭代步骤即可认为是网络输出图像达到最优的位置.

为了验证 $Q^b$ 是否能够作为迭代终止判断依据,通过追踪Lena噪声图像(受 $\sigma=25$ 噪声干扰)一次完整的降噪过程(共迭代8 000次),记录每一步迭代的输出图像的 $Q^b$ 值(该值范围为0~1),同时也将输出图像与无失真图像的峰值信噪比(Peak Signal to Noise Ratio, PSNR)经过归一化后呈现在一张图上,其曲线如图2所示.

由图2(a)可知,在一个通道上, $Q^1$ 值在取最大值自动终止时的迭代步数为4 560步,所对应的PSNR值为33.51 dB,而实际上最优步数为4 562步,输出图像最优PSNR值也为33.51 dB,两者在PSNR值上无差别;如图2(b)所示,在另外一个通道上, $Q^2$ 值取最大值自动终止时的迭代步数为5 450步,所对应的PSNR值为33.36 dB,而实际上最优步数为4 562步,对应图像的PSNR值为33.46 dB,两者在PSNR值上相差0.1 dB.由此可见,利用 $Q^b$ 值取极大值时作为迭代终止条件所获得的图像质量基本上能保证与实际能获得的最优降噪



(a)  $Q^1$  曲线图



(b)  $Q^2$  曲线图

图2  $Q^b$ 值曲线与归一化PSNR曲线的对应关系

图像相差不大.需要说明的是,一般来讲,在双通道配置下, $Q^1$ 和 $Q^2$ 同步达到极大值的可能性不大,故所提的DCDIP方法需要在2个通道上相继符合终止条件后,整个迭代训练过程才会最终终止.

## 4 试验结果与分析

### 4.1 测试配置

为了验证所提出DCDIP双通道深度图像先验降噪模型的降噪性能,将其与基于外部图像先验类的降噪模型(包括DnCNN<sup>[7]</sup>,FFDNet<sup>[8]</sup>,IRCNN<sup>[22]</sup>,DRUNet<sup>[9]</sup>,DAGL<sup>[23]</sup>,SwinIR<sup>[10]</sup>和Restormer<sup>[11]</sup>)和基于内部图像先验类的降噪方法(包括BM3D<sup>[8]</sup>,N2V<sup>[24]</sup>和DIP<sup>[12]</sup>)共10种降噪方法进行比较.测试数据集为权威文献中广泛采用的Set5<sup>[12]</sup>,Set12<sup>[7]</sup>,BSD68<sup>[25]</sup>,Kodak24<sup>[11]</sup>这4个基准测试数据库.对于合成噪声降噪实验,人工添加高斯噪声(噪声水平值同权威文献中广泛采用的15,25和

50 共 3 个级别)到数据库中的无失真图像中,从而获得合成噪声图像.对于真实噪声图像降噪,则直接使用 Nam 真实噪声图像集<sup>[26]</sup>中的噪声图像进行测试.

## 4.2 消融实验

首先,为了验证使用不同目标图像对 DIP 模型降噪性能的影响,共设计了 4 种配置条件下的对比实验,测试数据集为 Set12.如表 1 所示,实验 1 的配置其实是原 DIP 降噪模型,即以噪声图像为目标图像;实验 2 以 FFDNet 预处理图像为目标图像;实验 3 同时以 FFDNet 预处理图像和噪声图像为双目标图像;实验 4 的网络输出包含两个通道,其中一个通道以 FFDNet 预处理图像和噪声图像为目标图像,另一个通道以 BM3D 预处理图像和噪声图像为双目标图像,该配置即所提的 DCDIP 模型.

表 1 目标图像组合方式

实验编号	通道 1	通道 2	目标图像设置
实验 1	Noisy	—	噪声图像
实验 2	FFDNet	—	预处理图像
实验 3	FFDNet+Noisy	—	预处理与噪声图像
实验 4	FFDNet+Noisy	BM3D+Noisy	双通道

由表 2 可知,把目标图像由噪声图像更换成 FFDNet 预处理图像(即实验 2),DIP 降噪性能得到了大幅度提升.这表明提升目标图像的图像质量对提高降噪效果有利,但是所获得的降噪后图像的质量尚没有超越预处理图像本身的质量(为便于比较,最后一列给出了 FFDNet 预处理图像的 PSNR 值).原因是,在以 FFDNet 预处理图像为目标图像时,网络的输出最终会与目标图像几乎相同,无法提高降噪图像质量.相较于实验 1 与实验 2 的单目标图像学习,实验 3 采取双目标学习的方法,在 PSNR 数值上的表现要更好一些,且已经全面超越了 FFDNet 预处理图像的 PSNR 值(尤其是低噪声水平值条件下).这充分说明,噪声图像包含了一些可利用的降噪信息,故作为目标图像有利于提高降噪效果.在实验 4 中,将网络的输出由单通道改为双通道,在 2 个通道中同时学习不同的预处理图像,并通过监控  $Q^b$  最大值实现迭代自动终止,以避免过拟合,最后再使用融合算法将输出图像进行融合.这不仅可以让网络学习到更多的信息,还能够利用预处理图像中互补的信息,所以最终实验 4 中的 PSNR 值最高,且明显高于最后一列 FFDNet 预处理图像的 PSNR 值.由表 2 可以看出,使用双目标图像学习以及双通道逼近方法后,DCDIP 所获得的降噪后图像的 PSNR 值相比单目标学习、单通道的原 DIP 方法有了显著提高,大约提升了 2 dB.

其次,为了验证损失函数中权重参数  $\omega$  对降噪性能的影响,在 BSD68 数据集上分别使用不改变损失函

表 2 在 Set12 数据集上不同试验所获得的 PSNR 均值 单位: dB

$\sigma$	实验 1	实验 2	实验 3	实验 4	FFDNet
15	31.35	32.07	33.07	33.55	32.77
25	28.92	29.54	30.41	31.04	30.46
50	25.56	26.72	27.41	27.58	27.35
平均值	28.61	29.44	30.30	30.72	30.19

数中噪声部分权重(即  $\omega=1$ )与改变损失函数中噪声部分权重(即  $\omega$  值可变)两种策略在噪声水平值为 15, 25 和 50 下进行实验,结果如表 3 所示,加粗数据表示最优结果.由表 3 可知,在 3 个噪声水平值下采用可变  $\omega$  值情况下的 PSNR 性能相比固定  $\omega$  值分别高出约 0.36 dB.原因在于:随着噪声严重程度的提高,噪声图像中所蕴含的未被噪声干扰的像素点逐渐减少,故参数  $\omega$  应逐渐降低,直至为 0.因此,在损失函数中采用可变  $\omega$  值能最大限度地利用噪声图像中未被噪声干扰信息,同时排除噪声对网络输出图像的影响.

表 3 改变损失函数的噪声部分权重所获得的 PSNR 均值 单位: dB

$\omega$ 值	噪声水平值			
	15	25	50	平均值
固定值	32.71	30.04	26.54	29.76
可变值	<b>32.93</b>	<b>30.28</b>	<b>27.14</b>	<b>30.12</b>

再次,为了分析双通道逼近和 SPD 融合两个阶段对 DCDIP 降噪性能的影响,在 Set12, BSD68 和 Nam 数据集上进行了测试,结果如表 4 所示,加粗数据表示最优结果.其中,SPD 行的数据是在获得 BM3D 和 FFDNet 预处理图像后,直接利用 SPD 算法将它们融合所得到的结果.由表 4 可知,在噪声水平值比较低时,直接利用 SPD 融合算法将预处理图像(此时图像质量比较好)融合能获得较好的结果;而在噪声水平值较高时,所获得的预处理图像图片质量不太好时,使用双通道逼近策略对预处理图像进行调整,获得中间生成图像后再使用 SPD 融合(即本文策略),获得的图片质量会得到明显提高.需要说明的是,在处理真实噪声图像时,由于预处理降噪方法通常都无法对真实图像进行很好的降噪处理,故直接使用 SPD 算法融合所获得的图像质量偏低,而本文策略则能更为鲁棒地保证最终的降噪效果.

最后,为了测试增加通道数和计算迭代自动终止指标对执行时间的影响,本文将  $512 \times 512$  Lena 图像(噪声水平值为 25)各种配置条件下的执行时间进行了比较,结果如表 5 所示.具体地,当迭代步数超过 3 000 步时,每隔 10 步在各个通道上进行一次  $Q^b$  指标计算.表 5 中,第一行迭代次数表示各种配置条件下算法自动迭代终止时的迭代步数,第二行数据则为对应所需要的执行时间(括号里面的数据为不含计算  $Q^b$  迭代终止指

标,即 DIP 网络纯迭代的执行时间).从表 5 中可以看出:

(1)第二行数据与对应括号里面的数据差距很小,表明所提的迭代终止指标  $Q^b$  的计算时间非常少,几乎可以忽略.这是因为该迭代终止指标核心计算为小波变换,在成熟软件包的支撑下具有非常高的执行效率.

(2)本文提出的 DCDIP 最终执行时间(约 398 s)大概为两次单通道执行时间(约 707 s)的 56%.所以 DCDIP 模型基本与单通道 DIP 模型执行时间相当,并没有显著增加执

行时间.这是因为 DCDIP 仅仅是将网络输出的最后一层由单通道改为双通道,而计算 MSE 损失函数所增加的计算量相对网络参数更新的计算量占比很小.

(3)DCDIP 混合 Loss 函数最终在 5 450 步终止(取两通道最大值).其中,FFDNet+噪声图像单通道在 4 560 步终止,而 BM3D+噪声图像单通道在 5 450 步终止.需要说明的是,这里 DCDIP 所采用的混合 Loss 函数(式 6),与单独采用式(3)和式(5)所定义的单通道 Loss 数更新网络参数使  $Q^b$  达到最优值所需的迭代步数有少许不同.

表 4 采用直接融合策略与本文策略在降噪性能上的对比

单位:dB

数据集	Set12			BSD68			Nam
噪声水平	15	25	50	15	25	50	真实噪声
BM3D	32.36	29.96	26.70	31.38	28.82	25.81	40.52
FFDNet	32.77	30.46	27.35	31.91	29.44	26.47	40.87
SPD	<b>33.90</b>	<b>31.39</b>	27.80	<b>33.01</b>	29.99	26.52	40.78
DCDIP	33.76	31.36	<b>28.01</b>	32.93	<b>30.28</b>	27.14	<b>41.97</b>

表 5 Lena 图像上不同配置下执行时间比较

方案	FFDNet+噪声图像单通道	BM3D+噪声图像单通道	总和	DCDIP
迭代次数	4 480	5 200	9 680	Max(4560,5450)
执行时间/s	327.079(320.378)	380.291(367.797)	707.370(688.174)	397.900(385.938)

### 4.3 合成噪声图像降噪

#### 4.3.1 客观指标比较

首先,分别在 Set12, BSD68 和 Kodak24 数据集中无失真图像上分别施加噪声水平值为 15, 25 和 50 的高斯噪声,计算各噪声水平值下各降噪方法所获得的降噪后图像的 PSNR 均值,结果如表 6 所示,最优值加粗表示.从表 6 可以看出,在所有数据集上,DCDIP 在各个噪声水平值下 PSNR 值优于所有的对比方法.总体来说,DCDIP 在各噪声水平值下的降噪效果要优于其所采用的预处理方法(BM3D 与 FFDNet),并且显著优于内部图像先验类的无监督学习降噪方法(N2V 和 DIP).与新近提出的 Restormer 有监督降噪方法相比,DCDIP 也具有更优的性能.需要说明的是,表 6 中的实验数据是在人工合成噪声图像下完成的,Restormer 有监督降噪模型因有大量训练数据而占据优势.在真实噪声图像上,DCDIP 方法的性能则能超越更多.

其次,对 Set5 数据集中的彩色图像分别施加了噪声水平值为 15, 25 和 50 的高斯噪声,结果如表 7 所示. DCDIP 在各个噪声水平值总体上显著优于内部图像先验类的 N2V 和 DIP 无监督学习降噪方法(2.4~2.8 dB),优于 DRUNet,与 SwinIR 持平,稍弱于 Restormer 有监督降噪模型(0.05 dB).虽然 DCDIP 模型要稍弱于新近提

出的 Restormer 有监督模型,但是还是显著优于 BM3D 和 FFDNet 预处理方法(0.7~1.1 dB),这说明所提的 DCDIP 改进措施是有效的.

#### 4.3.2 主观比较

为了直观地分析 DCDIP 降噪模型的视觉效果,利用各个降噪方法对施加了高斯噪声水平值为 50 的 Man 图像进行降噪,整体降噪后的图像及对应的局部(袖口处)方法区域的视觉效果如图 3 所示.从图 3 放大的子图可以看出,对于具有复杂内容的袖口区域, DnCNN, FFDNet, DRUNet 等基于外部图像先验的有监督降噪模型普遍存在过度光滑的情况,基于内部自相似性的 BM3D 算法处理效果也不好,而基于内部图像先验的 N2V, DIP 和 DCDIP 则体现出比较大的优势.相对而言,DCDIP 的复原效果最接近无失真图像,表现出最佳的降噪能力与细节保持能力.通过在灰度的主观比较可以发现,DCDIP 方法对图像细节的保护能力很强,较原 DIP 方法已有较大幅度的提升,其降噪性能显著优于其他降噪方法,甚至达到和超过了新近提出的主流有监督降噪方法.

### 4.4 真实噪声图像降噪

为了验证 DCDIP 的泛化能力和实用性,在 Nam 真实噪声图像集完成了测试<sup>[26]</sup>.由于 Nam 是彩色图像集合,故

表 6 灰度图像上各方法对人工合成高斯噪声的降噪效果对比

单位: dB

方法	Set12				BSD68				Kodak24			
	15	25	50	Avg	15	25	50	Avg	15	25	50	Avg
BM3D	32.36	29.96	26.70	29.67	31.38	28.82	25.81	28.67	31.91	29.42	26.36	29.23
DnCNN	32.67	30.35	27.18	30.07	31.88	29.39	26.40	29.22	32.43	30.00	27.00	29.81
FFDNet	32.77	30.46	27.35	30.19	31.91	29.44	26.47	29.27	32.46	30.06	27.10	29.87
IRCNN	32.76	30.37	27.12	30.08	31.90	29.38	26.36	29.21	32.44	29.97	26.94	29.78
DRUNet	33.25	30.94	27.90	30.70	32.15	29.68	26.76	29.53	32.83	30.43	27.49	30.25
DAGL	33.19	30.86	27.75	30.60	32.10	29.64	26.68	29.47	32.80	30.38	27.36	30.18
SwinIR	33.36	31.01	27.91	30.76	32.20	29.71	26.76	29.56	32.92	30.49	27.50	30.30
Restormer	33.35	31.04	<b>28.01</b>	30.80	32.18	29.72	26.79	29.56	32.91	30.51	27.57	30.33
N2V	30.83	29.26	26.46	28.85	29.60	28.07	25.69	27.78	30.36	28.80	26.23	28.46
DIP	31.35	28.92	25.56	28.61	30.37	27.85	24.59	27.60	30.86	28.49	25.36	28.24
DCDIP	<b>33.76</b>	<b>31.36</b>	<b>28.01</b>	<b>31.04</b>	<b>32.93</b>	<b>30.28</b>	<b>27.14</b>	<b>30.12</b>	<b>32.98</b>	<b>30.63</b>	<b>27.65</b>	<b>30.42</b>

表 7 彩色图像上各方法对人工合成高斯噪声的降噪效果对比

单位: dB

$\sigma$	BM3D	DnCNN	FFDNet	DRUNet	SwinIR	Restormer	N2V	DIP	DCDIP
15	34.03	34.29	34.31	34.89	35.06	35.02	31.90	32.84	34.99
25	31.62	32.12	32.12	32.71	32.81	32.87	30.73	30.24	32.87
50	28.66	29.25	29.27	29.84	29.89	30.03	27.45	25.84	29.92
平均值	31.44	31.89	31.90	32.48	32.59	<b>32.64</b>	30.03	29.64	32.59

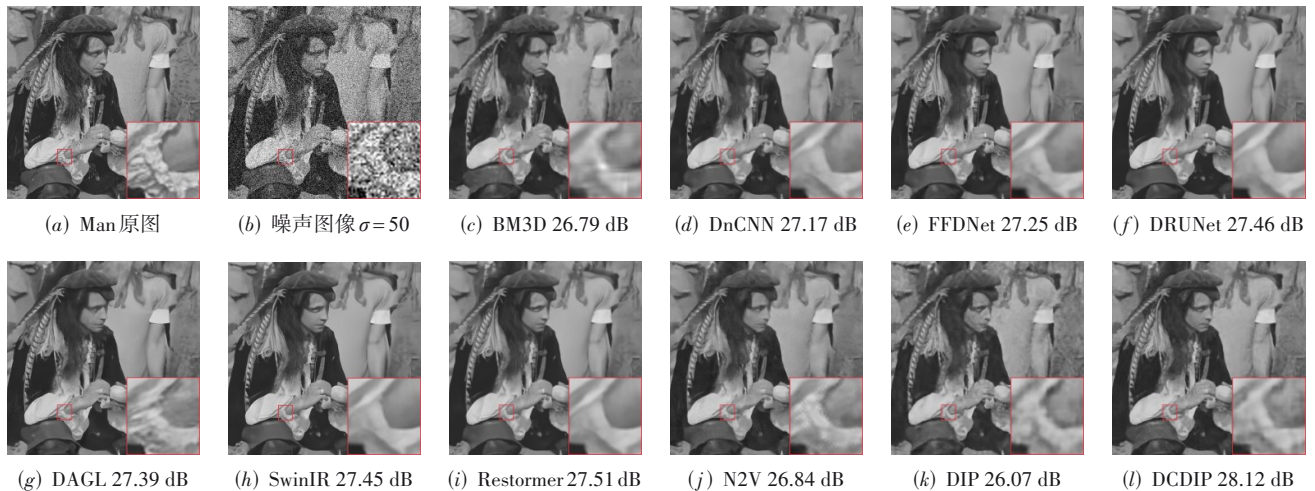


图 3 各降噪方法对 Man 噪声图像降噪的效果对比

选用 BM3D<sup>[8]</sup>, FFDNet<sup>[8]</sup>, VNet (Variational Denoising Network)<sup>[27]</sup>, Restormer<sup>[11]</sup> 和 DIP<sup>[12]</sup> 彩色图像处理方法进行对比. 由于 Nam 图像分辨率非常高, 故将随机从 Nam 数据集图像上裁切下来的大小为  $512 \times 512$  的图像块作为测

试图像. 图 4 展示了各对比方法在 Nam 数据集 17 号图块上降噪结果的视觉效果对比. 由图 4 可以看出, DCDIP 对图片的下半部分红色区域噪声的抑制效果较其他方法有明显优势 (其他方法仍然残留有微弱噪声未得到处理).



图4 各对比方法对真实噪声图像降噪的效果比较

## 5 结束语

本文提出了一种被称为DCDIP的双通道深度图像先验降噪模型,主要工作有:(1)利用具有较高图像质量的预处理图像作为新的目标图像并配合原噪声图像构成双目标图像,显著提升原DIP的降噪性能;(2)利用 $Q^b$ 值监控迭代训练过程,确保DCDIP模型在逼近双目标图像过程中总能获得具有较优图像质量的中间生成图像;(3)采用双通道工作方式使DCDIP网络输出图像分别逼近不同的双目标图像,再通过SPD图像融合技术将双通道上分别获得的中间生成图像融合,再次提升了降噪性能.在未来的工作中,可利用MAML(Model Agnostic Meta-Learning)<sup>[28]</sup>并结合网络参数压缩技术进一步提高DCDIP模型的执行效率.

### 参考文献

- [1] JIA X X, LIU S Y, FENG X C, et al. FOCNet: A fractional optimal control network for image denoising[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2019: 6047-6056.
- [2] GUO S, YAN Z F, ZHANG K, et al. Toward convolutional blind denoising of real photographs[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2019: 1712-1722.
- [3] CHANG M, LI Q, FENG H J, et al. Spatial-adaptive network for single image denoising[C]//2020 European Conference on Computer Vision (ECCV). Cham: Springer, 2020: 171-187.
- [4] REN C, HE X H, WANG C C, et al. Adaptive consistency prior based deep network for image denoising[C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2021: 8592-8602.
- [5] XU J, ZHANG L, ZHANG D. External prior guided internal prior learning for real-world noisy image denoising[J]. IEEE Transactions on Image Processing, 2018, 27(6): 2996-3010.
- [6] 盖杉, 鲍中运. 基于深度学习的高噪声图像去噪算法[J]. 自动化学报, 2020, 46(12): 2672-2680.  
GAI S, BAO Z Y. High noise image denoising algorithm based on deep learning[J]. Acta Automatica Sinica, 2020, 46(12): 2672-2680. (in Chinese)
- [7] 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.
- [8] 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.
- [9] ZHANG K, LI Y W, ZUO W M, et al. Plug-and-play image restoration with deep denoiser prior[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022, 44(10): 6360-6376.
- [10] LIANG J Y, CAO J Z, SUN G L, et al. SwinIR: Image

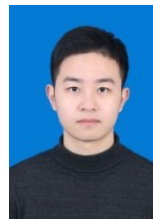
- restoration using swin transformer[C]//2021 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW). Piscataway: IEEE, 2021: 1833-1844.
- [11] ZAMIR S W, ARORA A, KHAN S, et al. Restormer: Efficient transformer for high-resolution image restoration [EB/OL]. (2021-11-18)[2022-08-29]. <https://arxiv.org/abs/2111.09881>.
- [12] ULYANOV D, VEDALDI A, LEMPITSKY V. Deep image prior[J]. International Journal of Computer Vision, 2020, 128(7): 1867-1888.
- [13] 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.
- [14] MA K D, LI H, YONG H W, et al. Robust multi-exposure image fusion: A structural patch decomposition approach[J]. IEEE Transactions on Image Processing, 2017, 26(5): 2519-2532.
- [15] MEINHARDT T, MOELLER M, HAZIRBAS C, et al. Learning proximal operators: Using denoising networks for regularizing inverse imaging problems[C]//2017 IEEE International Conference on Computer Vision (ICCV). Piscataway: IEEE, 2017: 1799-1808.
- [16] CHOI J H, ELGENDY O A, CHAN S H. Optimal combination of image denoisers[J]. IEEE Transactions on Image Processing, 2019, 28(8): 4016-4031.
- [17] YU L J, LUO J, XU S P, et al. An unsupervised weight map generative network for pixel-level combination of image denoisers[J]. Applied. Sciences, 2022, 12(12): 6227.
- [18] LUO J Y, XU S P, LI C X. A fast denoising fusion network using internal and external priors[J]. Signal, Image and Video Processing, 2021, 15(6): 1275-1283.
- [19] VU P V, CHANDLER D M. A fast wavelet-based algorithm for global and local image sharpness estimation[J]. IEEE Signal Processing Letters, 2012, 19(7): 423-426.
- [20] 周先春, 吴婷, 石兰芳, 等. 一种基于曲率变分正则化的小波变换图像去噪方法[J]. 电子学报, 2018, 46(3): 621-628. ZHOU X C, WU T, SHI L F, et al. A kind of wavelet transform image denoising method based on curvature variation regularization[J]. Acta Electronica Sinica, 2018, 46(3): 621-628. (in Chinese)
- [21] WANG Z, BOVIK A C, SHEIKH H R, et al. Image quality assessment: From error visibility to structural similarity [J]. IEEE Transactions on Image Processing, 2004, 13(4): 600-612.
- [22] ZHANG K, ZUO W M, GU S H, et al. Learning deep CNN denoiser prior for image restoration[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2017: 2808-2817.
- [23] MOU C, ZHANG J, WU Z Y. Dynamic attentive graph learning for image restoration[C]//2021 IEEE/CVF International Conference on Computer Vision (ICCV). Piscataway: IEEE, 2021: 4308-4317.
- [24] KRULL A, BUCHHOLZ T O, JUG F. Noise2Void - Learning denoising from single noisy images[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2019: 2124-2132.
- [25] MARTIN D, FOWLKES C, TAL D, et al. A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics[C]//Proceedings Eighth IEEE International Conference on Computer Vision. Piscataway: IEEE, 2001: 416-423.
- [26] NAM S, HWANG Y, MATSUSHITA Y, et al. A holistic approach to cross-channel image noise modeling and its application to image denoising[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2016: 1683-1691.
- [27] 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: NeurIPS, 2019: 1-12.
- [28] ZHANG K, XIE M Y, GOR M, et al. Metadip: Accelerating deep image prior with meta-learning[EB/OL]. (2022-09-18)[2022-12-27]. <https://arxiv.org/abs/2209.08452>.

#### 作者简介



徐少平 男, 1976年5月生, 江西九江人. 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



肖楠 男, 1999年7月生, 江西吉安人. 现为南昌大学硕士研究生. 主要研究方向为图像处理与深度学习.

E-mail: 406100210085@email.ncu.edu.cn