利用近清图像空间搜索的深度图像先验降噪模型

徐少平* 熊明海 周常飞
(南昌大学数学与计算机学院 南昌 330031)

摘 要：鉴于深度图像先验(DIP)降噪模型的性能高度依赖于目标图像所确定的搜索空间，该文提出一种新的基于近清图像空间搜索策略的改进降噪模型。首先，使用当前两种主流有监督降噪模型对同一场景下两张噪声图像分别进行降噪，所获得两张降噪后图像称为近清图像；其次，采用随机采样融合法将两张近清图像融合后作为网络输入，同时以两张近清图像替换噪声图像作为双目标图像以更好地约束搜索空间，进而在更为接近参考图像的空间范围内搜索可能的图像作为降噪后图像；最后，将原DIP模型的多尺度UNet网络简化为单尺度模式，同时引入Transformer模块以增强网络对长距离像素点之间的建模能力，从而在保证网络搜索能力的基础上提升模型的执行效率。实验结果表明：所提改进模型在降噪效果和执行效率两个方面显著优于原DIP模型，在降噪效果方面也超过了主流有监督降噪模型。

关键词：深度图像先验；降噪性能；近清图像；随机采样融合；双目标图像；Transformer
中图分类号：TN911.73; TP391.4 **文献标识码**：A **文章编号**：1009-5896(2024)11-4229-07
DOI: 10.11999/JEIT240114

Deep Image Prior Denoising Model Using Relatively Clean Image Space Search

XU Shaoping XIONG Minghai ZHOU Changfei
(*School of Mathematics and Computer Sciences*, *Nanchang University*, *Nanchang* 330031, *China*)

Abstract: Given that the performance of the Deep Image Prior (DIP) denoising model highly depends on the search space determined by the target image, a new improved denoising model called RS-DIP (Relatively clean image Space-based DIP) is proposed by comprehensively improving its network input, backbone network, and loss function.Initially, two state-of-the-art supervised denoising models are employed to preprocess two noisy images from the same scene, which are referred to as relatively clean images. Furthermore, these two relatively clean images are combined as the network input using a random sampling fusion method. At the same time, the noisy images are replaced with two relatively clean images, which serve as dual-target images. This strategy narrows the search space, allowing exploration of potential images that closely resemble the ground-truth image. Finally, the multi-scale U-shaped backbone network in the original DIP model is simplified to a single scale. Additionally, the inclusion of Transformer modules enhances the network's ability to effectively model distant pixels. This augmentation bolsters the model's performance while preserving the network's search capability. Experimental results demonstrate that the proposed denoising model exhibits significant advantages over the original DIP model in terms of both denoising effectiveness and execution efficiency. Moreover, regarding denoising effectiveness, it surpasses mainstream supervised denoising models.

Key words: Deep Image Prior (DIP); Denoising performance; Relatively clean images; Random sampling fusion; Dual-target image; Transformer

1 引言

图像降噪是数字图像处理中一个经典的不适定问题(ill-posed problem)[1,2]，也是长期活跃的研究主题。近十年来，基于深度学习(deep learning)构建的深度神经网络降噪模型在图像降噪性能方面较传统非学习降噪算法有了很大的提升[3–5]。根据训练模型时是否需要参考图像(ground-truth image)，现有深度网络降噪模型可以划分为有监督和无监督类型两种[6]，且大多属于有监督降噪模型。一般来讲，虽然DnCNN[7]，FFDNet[8], SWinIR[9], Restormer[10]

收稿日期：2024-02-28；改回日期：2024-09-07；网络出版：2024-09-29
*通信作者： 徐少平 xushaoping@ncu.edu.cn
基金项目：国家自然科学基金(62162043)
Foundation Item: The National Natural Science Foundation of China (62162043)

等有监督深度降噪模型的降噪效果与执行效率相对于传统非学习降噪算法有显著优势。但在实际应用中，完成有监督降噪模型的训练必须准备大量的噪声-清晰图像对，而采集清晰图像(参考图像)在很多情况下是非常耗时的(有时甚至是不可能完成的任务)，这导致所训练的有监督降噪模型存在所谓的数据偏见问题(data bias)，极大限制了它们在真实噪声图像(real-world noisy image)降噪中的效果。因此，为了解决有监督降噪模型必须依赖参考图像作为目标图像才能完成模型训练的问题，各类无监督降噪模型被相继提出。其中，比较具有代表性的工作是Lehtinen等人[11]在2018年提出的被称为N2N (Noise2Noise)的降噪模型。N2N模型基于统计学原理，使用两张经过不同独立噪声干扰的图像分别作为网络输入和目标图像。在训练噪声图像样本足够多的情况下，能实现较好的降噪效果。随后，Krull等人[12]在2019年提出了一种被称为N2V(Noise2Void)的降噪模型。N2V降噪模型属于自监督训练(self-supervised training)方法，它通过像素点随机掩盖(mask)技术获得了可用于模型训练的图像对(training pairs)。N2V模型能够直接在噪声的图像上完成训练而无需噪声-噪声图像对，具有更为广泛的应用范围。然而，N2V模型在随机掩盖掉像素点时会导致一些信息的丢失，从而降低了模型的降噪效果。近期，由Huang等人[13]提出的Ne2Ne(Neighbor2Neighbor)也是一种简单而有效的无监督图像降噪方法，其通过随机邻域子采样器(random neighbor sub-sampler)在无需参考图像情况下就能生成训练模型所需要的图像对，并构建额外的正则化损失函数以提高模型的降噪性能。总之，尽管上述这些无监督降噪模型无需参考图像就可以完成模型的训练，但它们仍需大量的噪声图像样本进行训练，故它们在一定程度上同样存在着有监督降噪模型的数据偏向问题，即待降噪的噪声图像一旦与训练集合中的图像在图像内容或者噪声分布模式存在较大差异时，其降噪效果将不可避免的下降。

2018年，由Lempitsky等人[14]提出的深度图像先验(Deep Image Prior, DIP)降噪模型在无监督降噪方面影响深远。有别于N2N, N2V, Ne2Ne等无监督降噪模型，DIP严格来讲属于基于单图像的自监督(self-supervised)模型，仅需待降噪图像自身一张图像就能够完成模型训练(降噪)。DIP模型在降噪过程中能针对给定的待降噪图像，生成与之相适应的特定降噪网络(image-specific denoising network)，它的网络参数值是根据给定待降噪图像的图像内容和噪声分布，通过训练而确定的(每一次降噪所用参数值均不同)。因此，与有监督(或无监督)降噪模型相比，自监督的DIP模型最大的优势在于它没有数据偏向的问题，具有更好的灵活性(自适应能力)，这种灵活性比各类有监督、无监督降噪模型更具优势(尤其在处理真实噪声图像时)。然而，DIP降噪模型的降噪效果还有很大提升空间。

鉴于同一场景下多张噪声图像之间在不同的局部区域具有互补性且易于获得，为进一步提升DIP降噪模型在真实噪声图像上的降噪效果和执行效率，本文从网络输入、损失函数和网络结构3个方面对其进行全面改进，提出了一种新的基于近清图像空间搜索策略的无监督深度图像先验改进模型(Relatively clean image Space search based DIP, RS-DIP)。RS-DIP模型总体上仍工作在DIP框架下，即在让网络输出图像逼近目标图像的迭代过程中实现降噪，但通过将网络输入和目标图像均改为近清图像而获得更高的降噪效果和执行效率。由于近清图像具有较高的图像质量，这种改进策略使得RS-DIP网络模型输出图像的搜索范围更小且输入图像与目标图像之间的映射难度显著降低，从而能保障提升DIP模型的降噪效果和执行效率。

2 相关工作

DIP深度图像先验是一种单图像自监督降噪模型，它通过深度神经网络结构自身捕获低级图像的统计先验信息建立隐式约束条件，仅使用一张噪声图像从零开始训练完成降噪任务。DIP采用随机向量z作为网络输入，噪声图像$\boldsymbol{y}$作为目标图像，网络训练以网络输出图像与目标图像之间的均方根误差(Mean Square Error, MSE)作为损失函数值。形式化地，DIP模型的降噪过程可以描述为

$$\theta^* = \arg\min_{\theta} \mathrm{MSE}(\boldsymbol{f}_\theta(\boldsymbol{z}); \boldsymbol{y}),\quad \boldsymbol{I}_{\mathrm{Stop}} = \boldsymbol{f}_{\theta^*}(\boldsymbol{z}) \tag{1}$$

即让深度网络$\boldsymbol{f}$通过不断调整其网络参数值θ，使得网络输出图像$\boldsymbol{f}_\theta(\boldsymbol{z})$不断逼近目标图像$\boldsymbol{y}$。理论上来讲，DIP降噪模型经过充分多的迭代训练，网络最终将收敛于目标图像，但DIP利用深度神经网络对噪声图像中噪声信息具备高阻抗(high impedance)特性[14](即网络输出图像会优先重构图像中的信号部分)。故可通过早停 (early stopping) 技术实现降噪，以早停点的网络输出图像$\boldsymbol{I}_{\mathrm{Stop}} = \boldsymbol{f}_{\theta^*}(\boldsymbol{z})$作为降噪后图像(作为参考图像的最佳估计)。简而言之，网络训练的过程其实就是降噪的过程。

作为一种单图像自监督图像降噪方法，DIP通过生成与噪声图像相适应的特定降噪网络从而较有监督模型更适于处理各类真实噪声图像(具有不同的图像内容和噪声分布)，在细节保留和噪声去除之

间能够实现更好的平衡，更具灵活性。然而，DIP降噪模型的性能高度依赖于目标图像所确定的搜索空间。以噪声图像本身作为目标图像(图像质量通常比较低，与参考图像之间的距离较远)，基于目标图像所构建的损失函数缺乏导引能力，从而使得网络输出图像在逼近目标图像过程中不易于收敛到离参考图像更近的位置。此外，在优化过程中也需要对网络参数进行大量的迭代更新操作(通常需要经过数千步)才能使得网络输出图像具有较高的图像质量，导致执行效率偏低。总之，DIP降噪模型在降噪效果和执行效率两个方面都有进一步提升的空间。

3　RS-DIP降噪模型

3.1　改进策略

为增强目标图像的约束能力，提出以下改进策略：(1)同时利用两个具有较高图像质量的近清图像作为双目标图像构建混合损失函数可增强原DIP模型单目标图像的导向能力。如图1所示，图中$\boldsymbol{x}$为参考图像，$\boldsymbol{y}$为噪声图像，$\boldsymbol{I}_{\text{avg}}$为近清图像$\boldsymbol{I}_1$和$\boldsymbol{I}_2$的均值图像。从图中可知RS-DIP方法停止点$\boldsymbol{I}_{\text{Stop}}^{\text{RS-DIP}}$与参考图像$\boldsymbol{x}$的距离显著小于DIP方法停止点$\boldsymbol{I}_{\text{Stop}}^{\text{DIP}}$。首先利用主流降噪方法所获得的近清图像$\boldsymbol{I}_1$和$\boldsymbol{I}_2$有利于约束搜索空间，从而确保网络输出图像能够收敛到更为合理的位置。与此同时，考虑到DIP网络模型训练通常需要经过相当多的迭代更新才能使得网络输出图像达到较好降噪效果(这是因为DIP网络模型的输入是随机张量，但随机张量$\boldsymbol{z}$与目标图像在图像空间中相距较远，使得网络参数值调整到收敛点的时间较长)。(2) 利用近清图像替换随机张量作为DIP网络的输入，从而使得网络输入和输出之间的映射难度降低，为简化模型打下

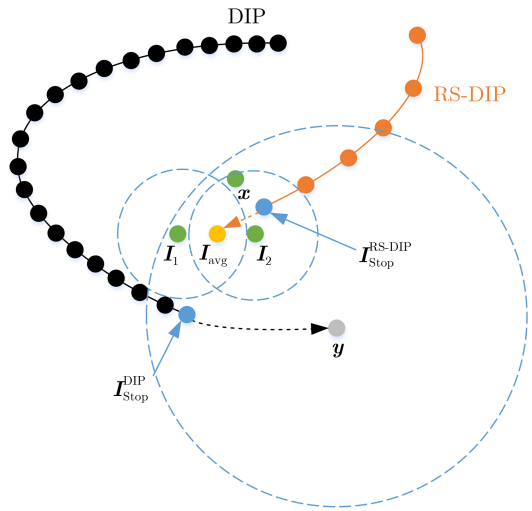


图 1 RS-DIP降噪模型降噪原理的可视化示意图

基础。(3)将原DIP网络的多尺度UNet骨干网络简化为单尺度(即将原网络的上下采样和Skip连接操作全部去除)。然而，这样的简化策略使得网络对长距离像素点之间的建模能力存在一定程度的削弱。为此，提出在网络中引入Transformer模块以解决网络简化后所导致的对长距离像素点建模能力减弱的问题。

3.2　骨干网络

RS-DIP降噪模型的骨干网络如图2所示。首先，该网络使用标准卷积操作提取图像特征，然后使用核心组件Transformer模块(串联使用2个)来强化长距离像素之间的描述能力。单个Transformer模块由两个部分组成：(1)多头注意力机制。如图2所示，多头注意力模块是在跨通道上而不是在跨空间维度上应用自注意力机制，具有线性复杂度。此外，它还引入了深度卷积来强调局部上下文，之后再计算特征协方差以生成全局注意力图。该模块能够聚合局部和非局部像素交互，能有效地处理长距离像素点之间的关系。(2)前馈网络。与一般的前馈网络不同之处在于门控机制和深度卷积。门控机制是通过两个并行的线性转换层的逐元素乘积来实现的，其中一个线性转换层使用GELU(Gaussian Error Linear Unit)非线性激活函数[15]。深度卷积用于编码邻近像素位置的信息，有助于修复图像的局部细节。最后，一个卷积层用来将通道数降低(降低到1或者3)，可输出灰度或者彩色图像。由于RS-DIP降噪模型训练用的输入图像与目标图像均是近清图像，因此网络结构不必设计得太复杂，浅层的网络就能够达到较好的非线性映射效果。同时，较少的参数量使得模型具有较高的执行效率。

3.3　网络输入

RS-DIP降噪模型的网络输入来源于两张近清图像$\boldsymbol{I}_1$和$\boldsymbol{I}_2$，但非直接使用。具体地，在每次迭代时先随机生成掩码矩阵$\boldsymbol{m}$和$\boldsymbol{m}_{\text{r}}$($\boldsymbol{m}$的互补矩阵)，随后按照式(2)、式(3)的定义生成两张采样融合图像$\boldsymbol{I}_{\text{s}}^1$和$\boldsymbol{I}_{\text{s}}^2$。最后，两张采样融合图像$\boldsymbol{I}_{\text{s}}^1$和$\boldsymbol{I}_{\text{s}}^2$经Concatenate操作后作为网络输入。

$$\boldsymbol{I}_{\text{s}}^1 = (\boldsymbol{I}_1 \odot \boldsymbol{m}) + (\boldsymbol{I}_2 \odot \boldsymbol{m}_{\text{r}}) \tag{2}$$

$$\boldsymbol{I}_{\text{s}}^2 = (\boldsymbol{I}_1 \odot \boldsymbol{m}_{\text{r}}) + (\boldsymbol{I}_2 \odot \boldsymbol{m}) \tag{3}$$

RS-DIP网络模型利用随机采样融合策略生成网络输入图像具有以下优点：(1)能向网络提供更为丰富的信息。图像$\boldsymbol{I}_{\text{s}}^1$和$\boldsymbol{I}_{\text{s}}^2$是在近清图像$\boldsymbol{I}_1$和$\boldsymbol{I}_2$基础上随机采样融合生成的，其图像质量很大程度上已经接近参考图像。相较于DIP模型使用的随机张量$\boldsymbol{z}$，能够向RS-DIP网络提供丰富的信息，使得

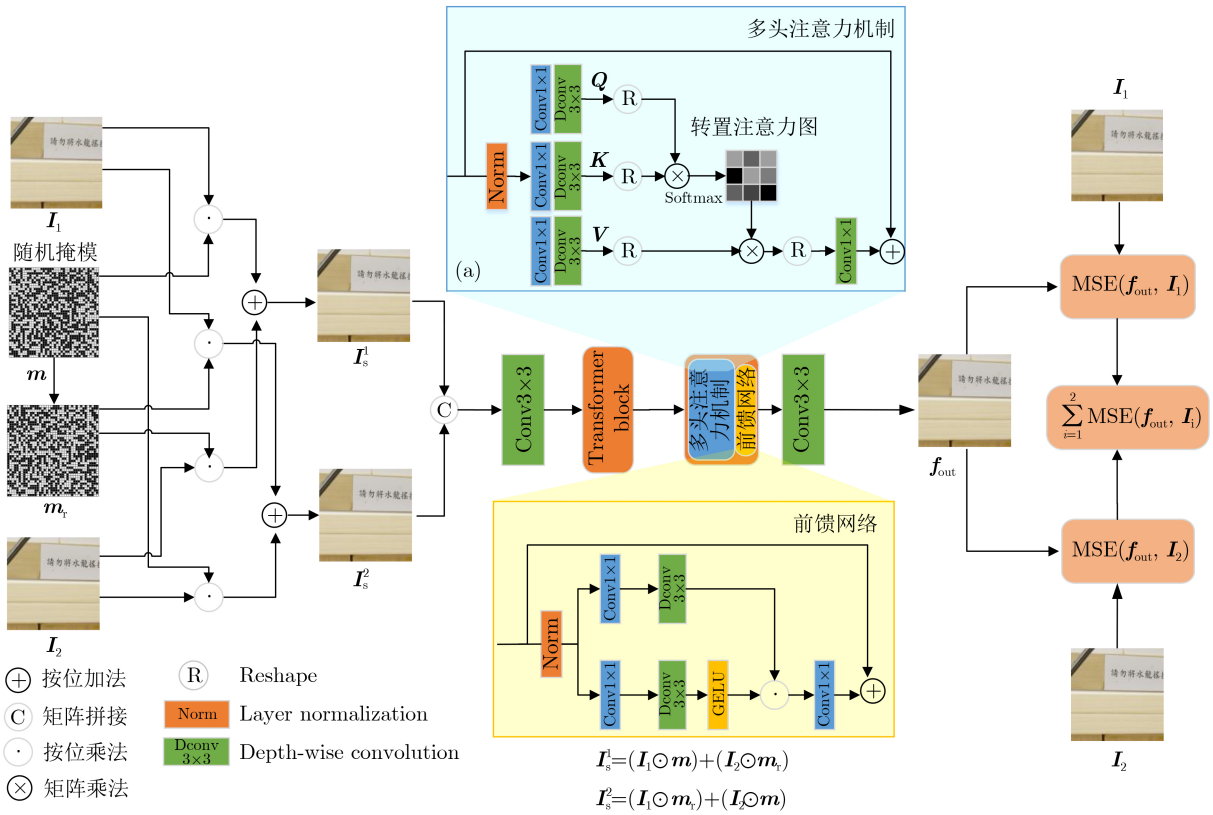


图 2 RS-DIP网络模型框架图

RS-DIP完成非线性映射的难度显著降低。因此，RS-DIP的网络参数调整也将更为容易，有利于加快网络收敛速度，减少训练迭代次数，从而提高RS-DIP降噪模型的执行效率。(2)能使网络模型更为鲁棒。由式(2)、式(3)可知：在每一次迭代步骤中，用于生成融合图像$\boldsymbol{I}_{\rm s}^1$和$\boldsymbol{I}_{\rm s}^2$的掩码矩阵$\boldsymbol{m}$均是不同的，这意味着每次迭代时所用的图像均是不同的(存在细微的差别)。从网络输入的宏观角度来看，融合图像$\boldsymbol{I}_{\rm s}^1$和$\boldsymbol{I}_{\rm s}^2$可以看成由某个固定图像叠加细微的随机扰动构成，这种细微扰动使得训练后所获得的网络模型具有更好的鲁棒性，从而能获得降噪效果更好的输出图像(具体实验数据参见第4节中消融实验)。

3.4 损失函数

RS-DIP降噪模型的损失函数包括两个部分

$$L=\mathrm{MSE}(\boldsymbol{f}_{\rm out},\boldsymbol{I}_1)+\mathrm{MSE}(\boldsymbol{f}_{\rm out},\boldsymbol{I}_2) \tag{4}$$

其中，$\boldsymbol{f}_{\rm out}$表示RS-DIP网络模型的输出图像，$\boldsymbol{I}_1$和$\boldsymbol{I}_2$分别是噪声图像经主流降噪方法(可以是1种或者2种不同的降噪模型)处理后所获得的近清图像。

4 实验结果与分析

4.1 测试环境配置

为了全面评价所提出的RS-DIP降噪模型整体的降噪性能，将其与DnCNN, FFDNet, SwinIR, Restormer, DAGL和DIP共6种主流降噪模型在3个真实噪声基准测试图像集上进行对比实验：(1)PolyU数据集[16]。同一场景下提供100张噪声图像以及1张参考图像。(2)NIND(Natural Image Noise Dataset)数据集[17]。同一场景由使用不同ISO值的相机获得多张噪声图像(大于等于4张)，提供1张参考图像。(3)SIDD(Smartphone Image Denoising Dataset)数据集[18]。同一场景提供2张噪声图像和1张参考图像。以上所有的真实噪声测试数据集合中同一场景下均提供2张以上的噪声图像。参与比较所有的模型都在相同硬件平台(英特尔(R)酷睿(TM)i7-11700 CPU@2.5 GHz, 32 GB RAM和NVIDIA Geforce RTX 3090 GPU)和软件环境(Windows10操作系统)上运行。

4.2 消融实验

理论上来讲，RS-DIP降噪模型可以使用任意两种不同的降噪模型组合来生成近清图像。为测试不同的组合对降噪性能的影响，从DnCNN, FFDNet, DAGL, SwinIR和Restormer中选择两种降噪模型进行组合并在PolyU图像集上进行了测试，实验结果列在表1中。由表1可知，SwinIR和Restormer降噪模型组合生成近清图像能获得最佳的降噪效果，这表明由SwinIR和Restormer降噪模型生成的近清图像之间具有较好的互补性，能更好地约束搜索空

表 1　不同降噪模型组合对降噪性能的影响(dB)

算法组合	Restormer+ DnCNN	Restormer+ FFDNet	Restormer+ DAGL	DAGL+ DnCNN	DAGL+ FFDNet	DAGL+ SwinIR	Restormer+ SwinIR
1	36.33	36.29	35.79	36.19	35.77	36.08	**36.44**
2	36.74	36.64	36.06	36.43	36.14	36.48	**36.84**
3	35.25	34.78	34.78	34.78	**35.57**	35.28	35.03
4	31.33	31.49	31.48	31.55	31.56	**31.62**	31.58
5	**38.41**	37.00	37.57	38.21	36.86	37.30	37.31
6	38.51	40.11	38.53	38.40	39.72	39.82	**40.23**
7	30.19	30.27	30.02	29.58	29.78	29.77	**30.32**
8	34.52	34.69	34.17	33.78	33.98	34.03	**34.70**
9	**35.83**	35.41	34.79	34.53	34.81	34.80	35.50
10	**36.65**	36.38	36.04	36.14	35.85	35.94	36.48
平均值	35.27	35.31	34.98	34.96	35.00	35.11	**35.44**

间范围，让RS-DIP降噪模型收敛到离参考图像更近的位置。

在每次迭代时，RS-DIP降噪模型均会对Restormer和SwinIR模型生成的近清图像进行随机采样融合，从而获得存在一定细微扰动的输入图像。为了验证随机采样融合方法是否能对RS-DIP降噪模型的性能带来益处，在直接使用Restormer和SwinIR生成近清图像作为网络输入的配置下进行了测试(即省略采样融合操作，模型命名为RS-DIP-1)。骨干网络和损失函数与RS-DIP模型完全相同，实验仍然在PolyU图像集上完成，实验结果列在表2中。由表2可知，RS-DIP模型的平均降噪性能稍优于RS-DIP-1模型，具体到单个图像上最高能提升0.1 dB左右，这说明随机采样融合策略提升了模型的鲁棒性，从而能在一定程度上提升模型的降噪效果，这主要得益于Transformer结构的引入。

RS-DIP降噪模型的骨干网络为简化后的单尺度简化网络(但含Transformer模块)，为了验证采用简化骨干网络对模型降噪性能的影响，将其与使用多尺度UNet骨干网络(命名为RS-DIP-2)在PolyU图像集上进行对比测试，实验结果列在表3中。由表3可知：RS-DIP降噪模型的平均降噪性能优于RS-DIP-2模型，单张图像上最多能超约0.3dB。这说明RS-DIP所采用的含Transformer模块的单尺度简化网络能够获得与原多尺度UNet更好的结果。

4.3　客观评价

为了全面客观地验证所提出RS-DIP降噪模型的降噪效果，使用各个对比模型在PolyU,NIND和SIDD共3个真实噪声数据集上完成了对比实验。采用PSNR值(Peak Signal to Noise Ratio)作为各个对比模型的客观评价指标[19]，实验结果列在表4中。

表 2　随机采样操作对模型降噪性能的影响比较(dB)

对比算法	RS-DIP-1	RS-DIP
1	36.42	**36.44**
2	**36.90**	36.84
3	34.93	**35.03**
4	**31.59**	31.58
5	37.18	**37.31**
6	40.14	**40.23**
7	30.30	**30.32**
8	34.70	**34.70**
9	35.49	**35.50**
10	36.36	**36.48**
平均值	35.40	**35.44**

表 3　简化骨干网络对模型降噪性能的影响(dB)

对比算法	RS-DIP-2	RS-DIP
1	36.22	**36.44**
2	36.79	**36.84**
3	**35.07**	35.03
4	31.48	**31.58**
5	36.83	**37.31**
6	40.21	**40.23**
7	30.31	**30.32**
8	34.54	**34.70**
9	35.43	**35.50**
10	36.20	**36.48**
平均值	35.31	**35.44**

由表4可知：在PolyU,NIND和SIDD图像集上，RS-DIP降噪模型的降噪效果在所有噪声水平值上显著优于其他对比模型(包括用于生成近清图像的

降噪模型)。其中，在PolyU,NIND和SIDD图像集上PSNR均值超过排名第2的降噪模型分别为1 dB, 0.3 dB和0.4 dB。

表 4 各对比方法在各真实噪声数据集上所获得的PSNR值比较(dB)

数据集	PolyU	NIND	SIDD
BM3D	33.90	32.87	38.00
DnCNN	33.28	31.33	32.84
FFDNet	34.25	32.72	37.54
DAGL	33.44	32.10	42.47
Restormer*	33.47	35.20	44.17
SwinIR*	34.45	33.84	37.24
DIP	34.43	33.79	39.42
RS-DIP	**35.44**	**35.46**	**44.60**

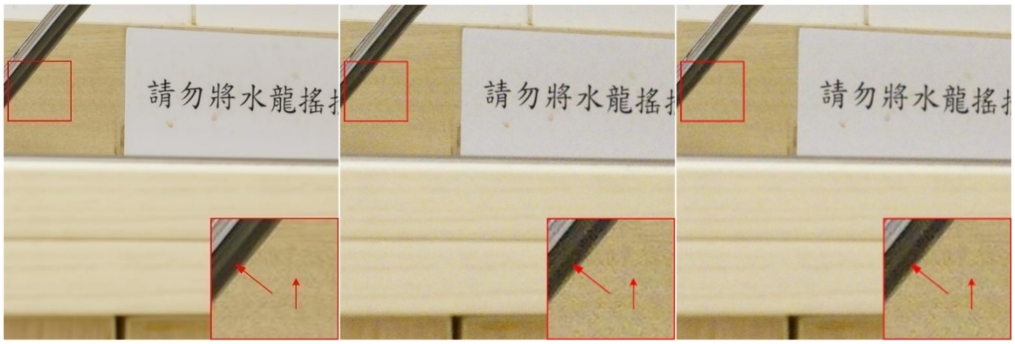
*表示该模型被用于处理噪声图像，在NIND和PolyU数据集上分别使用Restormer和SwinIR作为预处理算法处理不同的噪声图像，而在SIDD数据集上仅使用Restormer作为预处理算法处理两张噪声图像。

4.4 主观评价

如图3所示，只有本文提出的RS-DIP模型在去除棱形柱状物区域噪声的同时保留了背景上的纹理细节，而其他模型并没有兼顾到这两点。RS-DIP降噪模型不仅有效去除了噪声，而且保留了清晰的纹理和细节，具有更好的视觉效果。

5 总结

本文提出一种基于近清图像空间搜索策略的深度图像先验改进模型(即RS-DIP模型)，该模型在处理噪声分布复杂的真实噪声图像集合上较原DIP模型获得了更高的降噪效果和执行效率。目前RS-DIP网络参数值随机初始化的策略使得网络模型需要较多次数的迭代更新才能调整到最佳值，因此未来可考虑研究快速初始化网络参数的方法，如在Meta-Transfer Learning技术[20]支撑下让RS-DIP



(a) 原图 (b) 噪声图像 (c) DnCNN, PSNR=35.24 dB

(d) FFDNet, PSNR=39.53 dB (e) DAGL, PSNR=36.32 dB (f) SwinIR, PSNR=39.73 dB

(g) Restormer, PSNR=37.47 dB (h) DIP, PSNR=39.26 dB (i) 本文, PSNR=40.23 dB

图 3 各对比方法在PolyU图像上的降噪效果对比

降噪模型经过少量迭代训练即可达到收敛，从而进一步提升效率。

参 考 文 献

[1] JIN Xin, ZHANG Li, SHAN Chaowei, *et al*. Dual prior learning for blind and blended image restoration[J]. *IEEE Transactions on Image Processing*, 2022, 31: 1042–1056. doi: 10.1109/TIP.2021.3135482.

[2] 白勇强, 禹晶, 李一秾, 等. 基于深度先验的盲图像去模糊算法[J]. 电子学报, 2023, 51(4): 1050–1067. doi: 10.12263/DZXB.20211483.
BAI Yongqiang, YU Jing, LI Yinong *et al*. Deep prior-based blind image deblurring[J]. *Acta Electronica Sinica*, 2023, 51(4): 1050–1067. doi: 10.12263/DZXB.20211483.

[3] 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.

[4] GOYAL B, DOGRA A, AGRAWAL S, *et al*. Image denoising review: From classical to state-of-the-art approaches[J]. *Information Fusion*, 2020, 55: 220–244. doi: 10.1016/j.inffus.2019.09.003.

[5] 周建新, 周凤祺. 基于改进协同量子粒子群的小波去噪分析研究[J]. 电光与控制, 2022, 29(1): 47–50. doi: 10.3969/j.issn.1671-637X.2022.01.010.
ZHOU Jianxin and ZHOU Fengqi. Wavelet denoising analysis based on cooperative quantum-behaved particle swarm optimization[J]. *Electronics Optics & Control*, 2022, 29(1): 47–50. doi: 10.3969/j.issn.1671-637X.2022.01.010.

[6] 曾理, 熊西林, 陈伟. 低剂量CT图像降噪的深度图像先验的目标偏移加速算法[J]. 电子与信息学报, 2023, 45(6): 2188–2196. doi: 10.11999/JEIT220551.
ZENG Li, XIONG Xilin, and CHEN Wei. Deep image prior acceleration method for target offset in low-dose CT images denoising[J]. *Journal of Electronics & Information Technology*, 2023, 45(6): 2188–2196. doi: 10.11999/JEIT220551.

[7] ZHANG Kai, ZUO Wangmeng, CHEN Yunjin, *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.

[8] ZHANG Kai, ZUO Wangmeng, and ZHANG Lei. 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.

[9] GUO Shi, YAN Zifei, ZHANG Kai, *et al*. Toward convolutional blind denoising of real photographs[C]. 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, USA, 2019: 1712–1722. doi: 10.1109/CVPR.2019.00181.

[10] ANWAR S and BARNES N. Real image denoising with feature attention[C]. 2019 IEEE/CVF International Conference on Computer Vision, Seoul, Korea (South), 2019: 3155–3164. doi: 10.1109/ICCV.2019.00325.

[11] LEHTINEN J, MUNKBERG J, HASSELGREN J, *et al*. Noise2Noise: Learning image restoration without clean data[C]. Proceedings of the 35th International Conference on Machine Learning, Stockholm, Sweden, 2018: 2965–2974.

[12] KRULL A, BUCHHOLZ T O, and JUG F. Noise2void-learning denoising from single noisy images[C]. 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, USA, 2019: 2124–2132. doi: 10.1109/CVPR.2019.00223.

[13] HUANG Tao, LI Songjiang, JIA Xu, *et al*. Neighbor2Neighbor: Self-supervised denoising from single noisy images[C]. 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Nashville, USA, 2021: 14776–14785. doi: 10.1109/CVPR46437.2021.01454.

[14] LEMPITSKY V, VEDALDI A, and ULYANOV D. Deep image prior[C]. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, USA, 2018: 9446–9454. doi: 10.1109/CVPR.2018.00984.

[15] HENDRYCKS D and GIMPEL K. Gaussian error linear units (GELUs)[EB/OL]. https://arxiv.org/abs/1606.08415, 2023.

[16] XU Jun, LI Hui, LIANG Zhetong, *et al*. Real-world noisy image denoising: A new benchmark[EB/OL]. https://arxiv.org/abs/1804.02603, 2018.

[17] YANG Yu, XU Haifeng, QI Bin, *et al*. Stroke screening data modeling based on openEHR and NINDS Stroke CDE[C]. 2020 IEEE International Conference on BioInformatics and BioMedicine (BIBM), Seoul, Korea (South), 2020: 2147–2152. doi: 10.1109/BIBM49941.2020.9313127.

[18] ABDELHAMED A, LIN S, and BROWN M S. A high-quality denoising dataset for smartphone cameras[C]. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, USA, 2018: 1692–1700. doi: 10.1109/CVPR.2018.00182.

[19] 徐胜军, 杨华, 李明海, 等. 基于双频域特征聚合的低照度图像增强[J]. 光电工程, 2023, 50(12): 230225. doi: 10.12086/oee.2023.230225.
XU Shengjun, YANG Hua, LI Minghai, *et al*. Low-light image enhancement based on dual-frequency domain feature aggregation[J]. *Opto-Electronic Engineering*, 2023, 50(12): 230225. doi: 10.12086/oee.2023.230225.

[20] SOH J W, CHO S, and CHO N I. Meta-transfer learning for zero-shot super-resolution[C]. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, USA, 2020: 3513–3522. doi: 10.1109/CVPR42600.2020.00357.

徐少平：男，博士，教授，博士生导师，研究方向为数字图像处理、机器视觉、虚拟手术模拟.

熊明海：男，硕士生，研究方向为数字图像处理、机器视觉.

周常飞：男，硕士生，研究方向为数字图像处理、机器视觉.

责任编辑：马秀强