



DOI:10.12404/j.issn.1671-1815.2403865

引用格式:马荣恒,喻春雨,童亦新.基于残差密集网络与注意力机制的图像降噪[J].科学技术与工程,2025,25(9):3795-3805.

Ma Rongheng, Yu Chunyu, Tong Yixin. Image denoising based on residual dense networks and attention mechanism[J]. Science Technology and Engineering, 2025, 25(9): 3795-3805.

基于残差密集网络与注意力机制的图像降噪

马荣恒,喻春雨*,童亦新

(南京邮电大学电子与光学工程学院、柔性电子(未来技术)学院,南京 210023)

摘要 针对基于卷积神经网络图像降噪模型采用简单编码器-解码器结构而导致图像降噪性能差的问题,提出一种基于残差密集网络与注意力机制的残差密集图像降噪网络(residual dense image denoising network, RDIDNet)。首先,利用全局残差块增强网络模型的非线性映射能力;其次,引入双元素卷积注意力模块以实现 RDIDNet 模型解码过程中的自适应特征融合;最后,将 RDIDNet 降噪模型和 14 种代表性降噪方法进行对比,并进行消融实验,验证在基准模型上使用 RDU Sub-Network、DE-CAM、PSNRLoss 进行网络优化的有效性。实验结果表明,在 Set12 数据集、BSD68 数据集中,RDIDNet 在峰值信噪比(peak signal to noise ratio, PSNR)、结构相似性(structural similarity, SSIM)指标上相比传统经典方法 BM3D 分别平均提高 1.03 dB 和 0.027 5;比基于 Vision Transformers 架构的 SwinIR 分别平均提高 0.03 dB 和 0.001 4;比基于 CNN 的最新降噪方法 NHNet 分别平均提高 0.22 dB 和 0.008 9。RDIDNet 降噪网络更关注低频信息、模型训练更稳定,在有效消除图像噪声的同时能有效保留图像细节纹理,具有较好的表现。

关键词 图像降噪;深度学习;残差网络;注意力机制

中图分类号 TP391.4; 文献标志码 A

Image Denoising Based on Residual Dense Networks and Attention Mechanism

MA Rong-heng, YU Chun-yu*, TONG Yi-xin

(College of Electronic and Optical Engineering & College of Flexible Electronics (Future Technology),
Nanjing University of Posts and Telecommunications, Nanjing 210023, China)

[Abstract] In order to solve the problem of poor image denoising performance caused by the simple encoder-decoder structure of the convolutional neural network image denoising model, a residual dense image denoising network (RDIDNet) based on the residual dense network and attention mechanism was proposed. Firstly, the global residual block was used to enhance the nonlinear mapping ability of the network model. Secondly, the double-element convolutional attention module was introduced to realize the adaptive feature fusion in the decoding process of RDIDNet model. Finally, the RDIDNet denoising model was compared with 14 representative denoising methods, and ablation experiments were conducted to verify the effectiveness of using RDU Sub Network, DE-CAM, and PSNRLoss for network optimization on the benchmark model. The experimental results show that in the Set12 dataset and BSD68 dataset, RDIDNet improves the peak signal to noise ratio (PSNR) and structural similarity (SSIM) metrics by an average of 1.03 dB and 0.027 5, respectively, compared to the traditional classical method BM3D. Compared to SwinIR based on Vision Transformers architecture, the average improvement is 0.03 dB and 0.001 4, respectively. Compared to the latest CNN based denoising method NHNet, it has an average improvement of 0.22 dB and 0.008 9. The RDIDNet denoising network focuses more on low-frequency information and has more stable model training. It can effectively eliminate image noise while preserving image details and textures, and has good performance.

[Keywords] image denoising; deep learning; residual networks; attention mechanism

图像降噪是提升图像视觉效果的必要手段,是图像处理的经典研究课题。近年来,计算机水平的迅速发展促使人工智能得到革命性突破,人们开始将深度学习应用到图像降噪方面,同时卷积神经网络

(convolutional neural networks, CNN)在图像恢复中取得的优秀成绩也使其成为图像降噪的主流网络。在传统的图像降噪中,人们大都是采用空间域滤波或变换域滤波的方式进行降噪,基于空间域去

收稿日期:2024-05-24 修订日期:2024-12-27

基金项目:南京邮电大学校企合作项目(2018 外 002, 2019 外 157)

第一作者:马荣恒(2000—),男,汉族,山东泰安人,硕士研究生。研究方向:计算机图像处理。E-mail:1647237056@qq.com。

*通信作者:喻春雨(1976—),女,汉族,辽宁沈阳人,博士,副教授。研究方向:图像处理。E-mail:yucy@njupt.edu.cn。

噪的方法一般会先选定一个模板,将图像中的像素点与模板相计算,最后得到处理后的图像,这类方法计算速度较快,但图像细节信息丢失较严重。基于变换域去噪方法主要有傅里叶变换去噪和小波变换去噪,首先将图像和噪声转换到指定的变换域,再根据噪声系数和图像系数进行筛选性的滤除,最后经过逆变换得到去噪后的图像^[1]。2010年,季维勋^[2]提出了一种新的基于奇异值分解(singular value decomposition, SVD)图像降噪方法,通过截断奇异值降噪和分块降噪取得了较奇异值降噪更好的降噪结果,但仍存在图像边缘模糊的问题。2017年,Zhang等^[3]提出降噪卷积神经网络(denoising convolutional neural network, DnCNN),通过引入残差学习和批量归一化提高网络降噪性能,在去除高斯噪声方面超过了大部分传统降噪方法;2018年,Zhang等^[4]提出快速灵活的降噪神经网络(fast and flexible denoising neural network, FFDNet),它的结构与DnCNN一致,区别在于将含噪声图像下采样后的多张子图像作为输入,使消除高斯噪声的能力增强。2019年,Guo等^[5]提出了面向真实场景照片的卷积盲降噪神经网络(convolutional blind denoising neural network, CBDNet),该网络将噪声估计子网络与非盲降噪子网络结合,提高了网络性能与泛化能力。同年,白瑞君等^[6]提出了一种改进的深度卷积神经网络,降低了DnCNN网络的参数量,使得网络训练更平稳。2021年,Zhang等^[7]在具有Denoiser先验的即插即用图像恢复研究中结合ResNet和U-Net各自优点,提出了深度降噪网络(plug-and-play image restoration with deep denoiser prior, DPiR),它以噪声水平映射为输入,能够处理程度更广的噪声。2021年,Liang等^[8]提出了基于Swin Transformer的图像重建网络(Swin Transformer based image restoration, SwinIR),第一次将Vision Transformer(ViT)引入降噪模型中,利用其对图像特征高效提取,超越了以往基于CNN降噪模型的降噪效果。2022年,Wang等^[9]提出了一种新的自监督降噪框架Blind2Unblind,通过全局感知掩模映射器和重新可见损失实现全局感知、盲点可见,改善了盲点驱动降噪方法的信息损失问题。同年,姜旭等^[10]提出了一种改进的U形神经网络,设计了一种边缘提取模块,注重图像的边缘细节和纹理特征。同年,Zheng等^[11]提出了一种新的混合降噪卷积神经网络(hybrid CNN for image denoising, HDCNN),它使用多个深度网络训练图像降噪模型,获得了更好的降噪效果。同年,Huang等^[12]提出小波启发的图像降噪可逆网络(wavelet-inspired invertible network

for image denoising, WINNet),它结合基于小波方法和基于学习方法的优点,提高了模型的泛化能力、可解释性。同年,周联敏等^[13]提出了一种结合迁移学习的真实图像去噪算法,提高了面对真实噪声时网络结构的鲁棒性,获得了较好的图像视觉效果。2023年,Lin等^[14]提出一种隐式深度降噪先验(implicit deep denoiser prior, IDDP)的正则化方法,将网络的输出视为一个“先验”,在“重新加噪”后再次对其进行降噪,使网络在没有见过任何干净图像的情况下也能学习降噪。同年,Zhang等^[15]提出了核基图像重建网络(kernel basis network, KBNNet)设计了一种核基注意(KBA)模块,通过引入可学习的核基来为获取图像的空间聚集信息完成图像降噪。

虽然基于CNN的降噪方法已经取得巨大进展,但大多是基于由几组重复卷积层组成的单个深度网络,或者采用简单的编码解码结构,这样会造成图像的纹理细节过于平滑^[16]。尽管徐少平等^[17]采用复杂连接替代简单跳跃连接(skip connection),缓解了模型在特征尺度变化时信息丢失的问题,但对低频特征的关注依旧不足。

现提出残差密集图像降噪网络(residual dense image denoising network, RDIDNet),以经典的编码器-解码器结构U-Net为基准模型,利用全局残差块(global residual block, GRB)替代编码器-解码器中简单的单层卷积,以增强模型的非线性映射能力,通过双元素卷积注意力模块(double element-convolution attention module, DE-CAM)代替U-Net中同尺度特征间的简单跳跃连接,在进行同尺度特征间的特征传递时,进行多尺度特征的自适应融合,以期加强低频信息的利用,降低空间信息的丢失率,提高网络降噪性能。

1 RDID 降噪神经网络

1.1 RDID 网络构成

如图1所示,RDIDNet网络结构是由特征域模块、重构模块、作为特征提取模块的残差密集U形子网络(residual dense U-shaped sub-network, RDU Sub-Network)和特征融合模块DE-CAM组成。RDIDNet网络降噪过程如下:首先,噪声图通过特征域模块将原来1通道或3通道的噪声图像扩展至64通道;然后将扩展后的 X_0 送入RDU Sub-Network进行初步降噪;再将初步降噪后的特征图与 X_0 在DE-CAM中进行特征融合,输出为 X_1 。如此重复操作4次后输出特征图 X_4 ,并通过重构模块将64通道的 X_4 重构为干净的1通道灰度图像或3通道RGB图像,从而完成降噪。

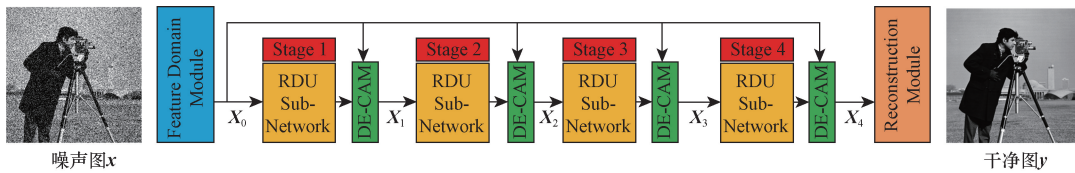


图1 RDID网络结构

Fig. 1 RDID network structure

1.1.1 特征域模块与重建模块

如图2所示,特征域模块通过扩展原图像的通道,将原图像从空间域转换到特征域,以利于后续特征提取获得更精细的图像特征;而重建模块则是对扩展后的特征图像进行复原,使其重建转换到空间域。

如图2(a)所示,特征域模块是由一个卷积核为 3×3 的卷积层与一个残差单元构成,而残差单元包含两个卷积核为 3×3 的卷积层和一个ReLU激活层。前者将输入图像 x 扩展为通道数64的特征域图像特征 X ,后者用来生成初始元素的特征权重 X_0 ,实现模块的稳定性和非线性。运算公式如下。

$$\begin{cases} X = \text{Conv}_{3 \times 3}(x) \\ X_0 = \text{Conv}_{3 \times 3}\{\delta[\text{Conv}_{3 \times 3}(X)]\} \end{cases} \quad (1)$$

式(1)中: $\text{Conv}_{3 \times 3}$ 为卷积核为 3×3 的卷积层; δ 为ReLU激活函数。

重建模块如图2(b)所示,其构成与特征域模块相似,但排列顺序不同,旨在将特征域图像特征 X 还原为干净的空间域图像 y ,运算公式为

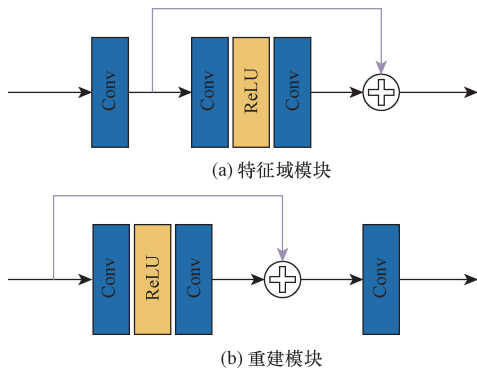


图2 特征域模块与重建模块

Fig. 2 Feature domain module and reconstruction module

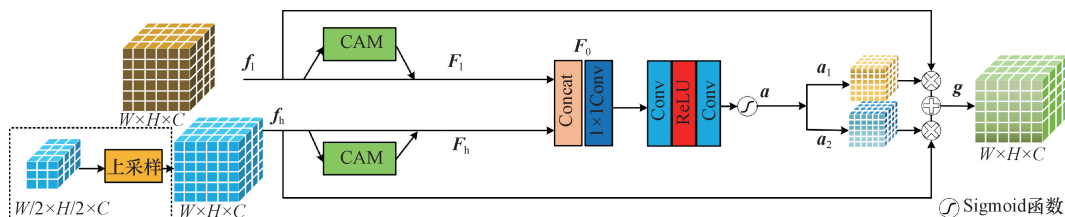


图3 DE-CAM 模块

Fig. 3 DE-CAM Module

投稿网址:www.stae.com.cn

$$\begin{cases} X = \text{Conv}_{3 \times 3}\{\delta[\text{Conv}_{3 \times 3}(X4)]\} \\ y = \text{Conv}_{3 \times 3}(X) \end{cases} \quad (2)$$

1.1.2 DE-CAM 模块结构

为加强模型对于不同尺度卷积特征的学习能力,设计了如图3所示的特征融合模块DE-CAM。该模块具有低等级特征 f_l 和高等级特征 f_h 两个输入和一个输出 g 。其中通道注意力模块^[18](channel attention module, CAM)如图4所示。若以特征图 $F \in \mathbf{R}^{C \times W \times H}$ 为输入,那么特征图 F 将分别进行平均池化(average-pooling, AvgPool)和最大池化(max-pooling, MaxPool),从而生成两个不同的特征通道信息 F_{avg}^c 和 $F_{\text{max}}^c \in \mathbf{R}^{C \times 1 \times 1}$ 。然后将它们送入包含一个隐藏层的多层感知器(multilayer perceptron, MLP)中生成 $M_c(F) \in \mathbf{R}^{C \times 1 \times 1}$,其中隐藏层的激活大小为 $\mathbf{R}^{C/r \times 1 \times 1}$, r 为衰减率,运算表达式为

$$M_c(F) = \text{sigmoid}[\text{MLP}(F_{\text{avg}}^c) + \text{MLP}(F_{\text{max}}^c)] \quad (3)$$

式(3)中:MLP为多层感知器;sigmoid为激活函数。

通过式(3),低等级特征 f_l 和高等级特征 f_h 输入CAM模块后将分别得到细化特征 F_l, F_h ,即

$$\begin{cases} F_l = M_c(f_l) \times f_l \\ F_h = M_c(f_h) \times f_h \end{cases} \quad (4)$$

将 F_l, F_h 在Concat层拼接,得到特征 $F_0 = [F_l, F_h]$, $F_0 \in \mathbf{R}^{2C \times W \times H}$ 。再将 F_0 通过卷积核为 1×1 的卷积层降低维度,然后通过两个卷积核和一个ReLU层生成初始元素的特征权重,最后通过sigmoid函数生成归一化的权重张量 a 。运算过程如下。

$$a = \sigma[\text{Conv}_{3 \times 3}(\delta\{\text{Conv}_{3 \times 3}[\text{Conv}_{1 \times 1}(F_0)]\})] \quad (5)$$

式(5)中: $\text{Conv}_{1 \times 1}$ 为积核为 1×1 的卷积层; σ 为sigmoid函数。

低等级特征 f_l 和高等级特征 f_h 的权重张量分别为 a_1, a_2 , 其中 $a_1 = a, a_2 = (1 - a)$ 。由此, DE-CAM 模块的输出表达式为

$$g = a_1 \otimes f_l + a_2 \otimes f_h = a \otimes f_l + (1 - a) \otimes f_h \quad (6)$$

式(6)中: \otimes 表示特征相乘操作。

如式(6), 通过训练 a_1, a_2 的权重, 注意力模块对不同尺度的输入特征进行自适应的特征融合, 从而获得理想的输出。

1.1.3 RDU Sub-Network 网络结构

如图 5 所示, RDU Sub-Network 延续 U 形网的整体设计: 保持 Encoder-Decoder 结构、包含 4 个比例尺、下采样阶段的编码器进行特征提取、上采样阶段的解码器进行图像重建。改进之处为: 将局部残差块(local residual block, LRB) 和扩张局部残差块(dilated local residual block, DLRB) 组成的 GRB 代替 U 形网中的特征提取模块; 通过引入图 3 的 DE-CAM 代替编码解码阶段中同尺度特征

间的跳跃连接, 增强模型对不同尺度特征的关注。

RDU Sub-Network 模块工作时, 使用步幅为 2、卷积核为 3×3 的卷积层对特征图进行下采样, 当编码网络到达第四尺度特征图即最小尺度特征图时, 则从下采样阶段转到上采样阶段, 上采样由比例系数为 2 的双线性插值和步幅为 1 的卷积完成。

4 个比例尺特征图均由 GRB 进行提取。在下采样阶段, 每个 GRB 包含一个 LRB 和一个 DLRB; 在上采样阶段, 每个 GRB 包含两个 LRB。

LRB 采用归一化策略^[19], 首先利用层归一化(layer normalization, LN)避免模型深度加深导致网络梯度不稳定、模型退化等问题; 然后通过两组卷积核大小为 3×3 的卷积层与 ReLU 激活函数以密集连接的方式进行连接, 这可以充分利用所有卷积层提取的局部特征; 最后采用 Concat 层与卷积核为 1×1 的卷积层进行局部特征融合、降维, 自适应地累积特征, 减少特征传播过程中的信息丢失。DLRB 则是将 LRB 中的卷积层改为卷积核大小为 3×3 , 扩

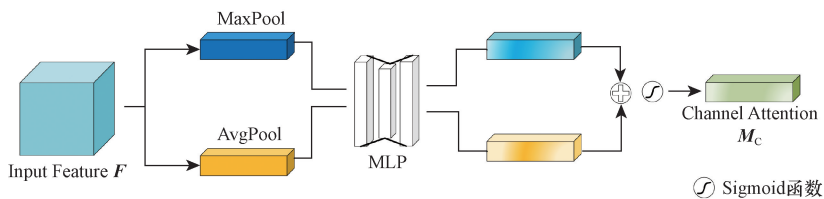


图 4 通道注意力模块

Fig. 4 Channel attention module

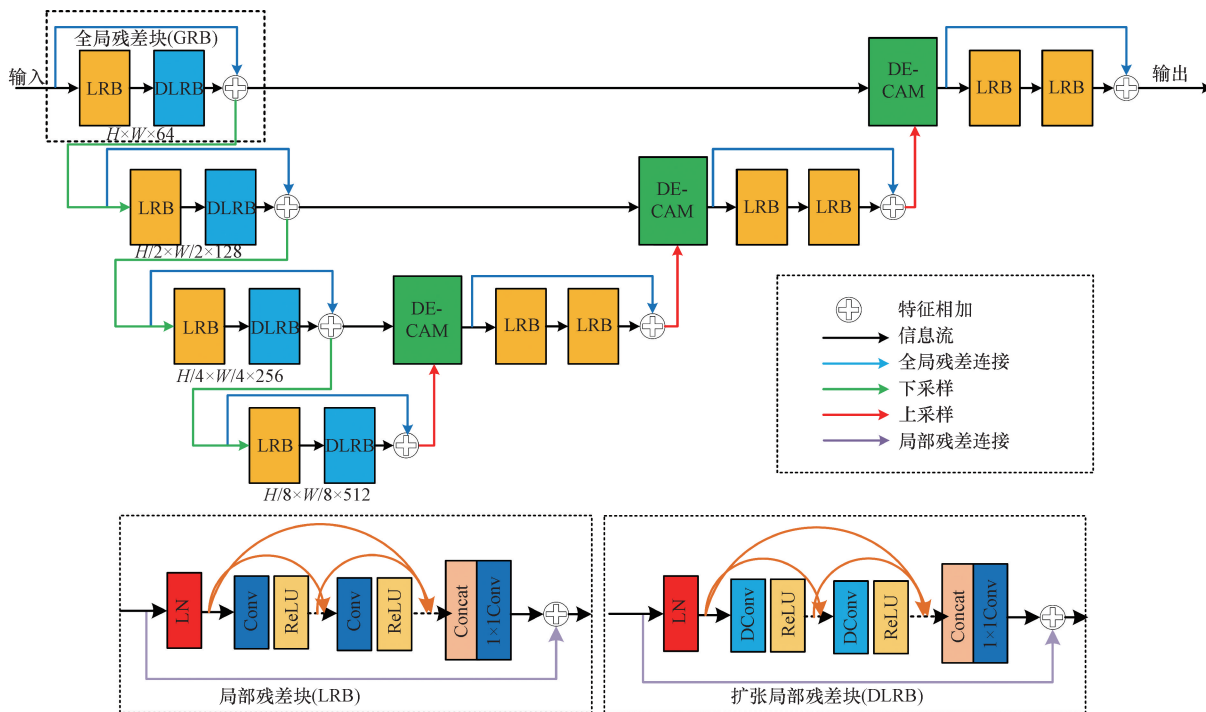


图 5 RDU Sub-Network 模块

Fig. 5 RDU Sub-Network module

张系数为2的空洞卷积(dilated convolution, DConv)层^[20],DLRB的设计是为了在下采样过程中GRB能拥有更多的感受野。

通过式(6),将上采样输出与上一尺度的GRB输出作为DE-CAM输入,实现相同尺度特征之间的连接,同时完成自适应的多尺度特征融合,使解码器在恢复图像特征时,能够关注不同分辨率的特征信息,改善特征尺度变化过程中的信息损失。

1.2 损失函数

模型训练过程使用峰值信噪比(peak signal to noise ratio, PSNR)作为损失函数的度量,即PSNR损失(PSNRLoss),其表达式为

$$L_{\text{psnr}} = 10 \lg \frac{\text{MaxValue}^2}{\text{MSE}} = 10 \lg \frac{2^{\text{bits}-1}}{\text{MSE}} \quad (7)$$

式(7)中:MaxValue为图像最大像素值;bits为像素位数;MSE为均方差误差(mean-square error),表达式为

$$\text{MSE}(\hat{y}, y) = \frac{1}{N} \sum_{i=1}^N \|\hat{y}_i - y_i\|^2 \quad (8)$$

式(8)中: \hat{y} 为图像预测值; y 为图像真实值; N 为图像数量。

2 实验结果与分析

2.1 运行环境

网络训练、测试和消融实验均基于Pytorch深度学习平台、NVIDIA GeForce RTX3090 GPU;初始学习率设为 1×10^{-4} ,每经过30个Epoch后学习率下降0.1倍,当学习率到达 1×10^{-6} 时停止下降;训练过程中Batch size设置为4,epoch设置为2000,采用Adam优化器^[21],使用默认参数。

2.2 实验数据集

采用伯克利分割数据集(Berkeley segmentation dataset, BSD)^[22]和Div2K^[23]作为训练集;采用Set12数据集^[24]和BSD68数据集中全部图像作为测试集。通过PSNR和结构相似性(structural similarity, SSIM)对降噪效果进行客观评价^[25]。

2.3 降噪效果分析

将提出的RDIDNet降噪模型和14种代表性降噪方法进行对比,这14种方法分别是BM3D^[26]、WNNM^[27]、DnCNN、FFDNet、DBSN^[28]、DeamNet^[29]、DRUNet、SwinIR、HDCNN、Blind2Unblind、RDDCNN^[30]、NHNet^[31]、WINNet、KBNet,其中BM3D、WNNM属于传统方法,其余是基于深度学习的方法。基于深度学习的方法中,SwinIR是基于ViT架构的方法,DBSN、Blind2Unblind是无监督学习的方法。分别选取Set12数据集中“Starfish”、BSD68数据集中“test044”进行降噪分析,因为HDCNN作者未公开代码,因此对其余13种方法进行主要分析。

2.3.1 主观评价

图6为Set12数据集中“Starfish”图添加标准差 $\sigma = 50$ 高斯噪声后的降噪效果对比,由整幅图及其兴趣区(region of interest, ROI)对比可见,经BM3D、WNNM、DnCNN、FFDNet、RDDCNN、WINNet、KBNet、DBSN降噪后图像变模糊,无法复原海星的条纹;DeamNet、DRUNet、Blind2Unblind、SwinIR、NHNet虽然可以复原出海星大部分纹理,但使更细小条纹变模糊,比如:SwinIR、DRUNet在两个ROI中出现错误的重建纹理,导致原纹路被遗漏、扭曲;DeamNet重建的海星纹理模糊,出现雾化;Blind2Unblind在绿框ROI中有严重雾化现象;而提出的网络RDIDNet在有效降噪的同时,更准确地恢复了海星的全部纹理。

图7为对BSD68数据集“test044”图添加标准差 $\sigma = 50$ 高斯噪声后的降噪效果对比。由全图及其ROI对比可见,图像经过BM3D、WNNM、DnCNN、FFDNet、RDDCNN、NHNet、KBNet处理后,只能复原出玻璃金字塔的整体轮廓,对较粗的钢结构的复原效果不够理想;DRUNet、SwinIR虽然能复原出金字塔较粗的结构特征,但对细节纹理、尤其边缘特征产生过度平滑,对玻璃金字塔内部复杂结构与边缘结构纹理丢失严重;Blind2Unblind、WINNet虽然能对边缘特征有效处理,仍使部分钢结构出现雾化;DeamNet能较完整地复原复杂的钢结构纹理,并且对边缘特征有良好的处理效果,但依旧出现错误的重建纹理,导致原来连续的钢结构纹理被错误地分成两截。而经过本文提出的网络RDIDNet降噪后,图像纹理更清晰、伪影更少,视觉效果明显优于其他方法。

2.3.2 客观评价

如表1所示,给出不同方法对Set12数据集全部12张图像添加 $\sigma = 50$ 高斯噪声的降噪评价参数PSNR、SSIM($[0, 1]$),两者值越大表明降噪效果越好。而BSD68数据集拥有68张测试图像,和使用Set12数据集分析结果类似,未在表1给出所有测试图像的具体分析。由表1中数据对比可知,RDIDNet在12张测试图中,有6张图像取得最佳PSNR、SSIM,并且对12张测试图的PSNR、SSIM平均值最高,其PSNR、SSIM指标比DnCNN高0.83 dB、0.008 1,比SwinIR高0.09 dB、0.002 5。

如表2所示,给出不同降噪方法通过Set12数据集、BSD68数据集对高斯噪声进行降噪分析得到的PSNR、SSIM。可以看出,RDIDNet降噪后得到的PSNR、SSIM值仅在BSD68数据集中 $\sigma = 15$ 时略低于SwinIR,其他均处于首位。而基于ViT的SwinIR始终优于本文降噪方法以外的其他方法。

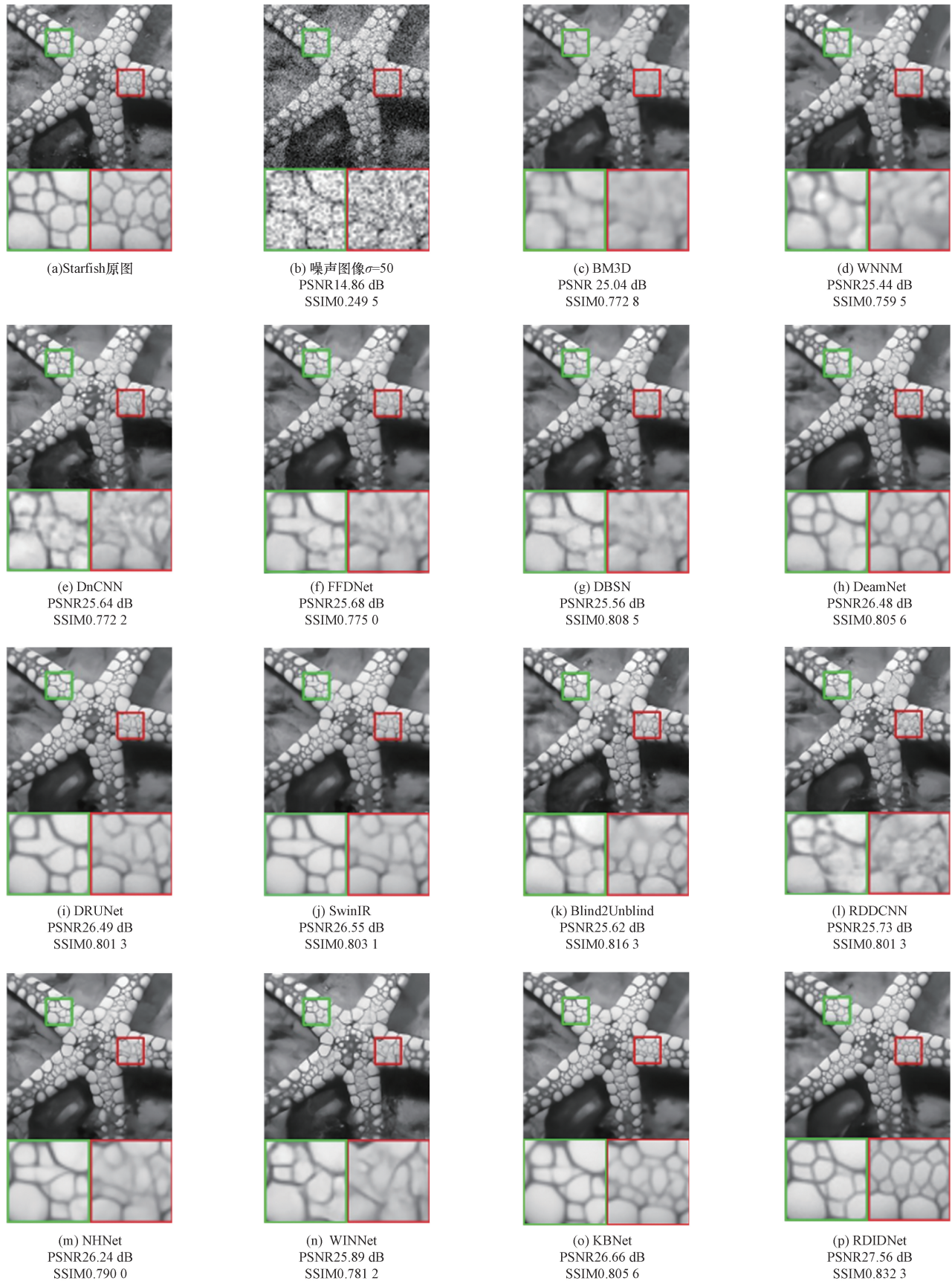


图6 不同方法对 Set12 中“Starfish”添加 $\sigma = 50$ 高斯噪声后的降噪效果对比
 Fig.6 Comparison of denoising effects using different methods on the “Starfish” image from Set12 after adding Gaussian noise with $\sigma = 50$

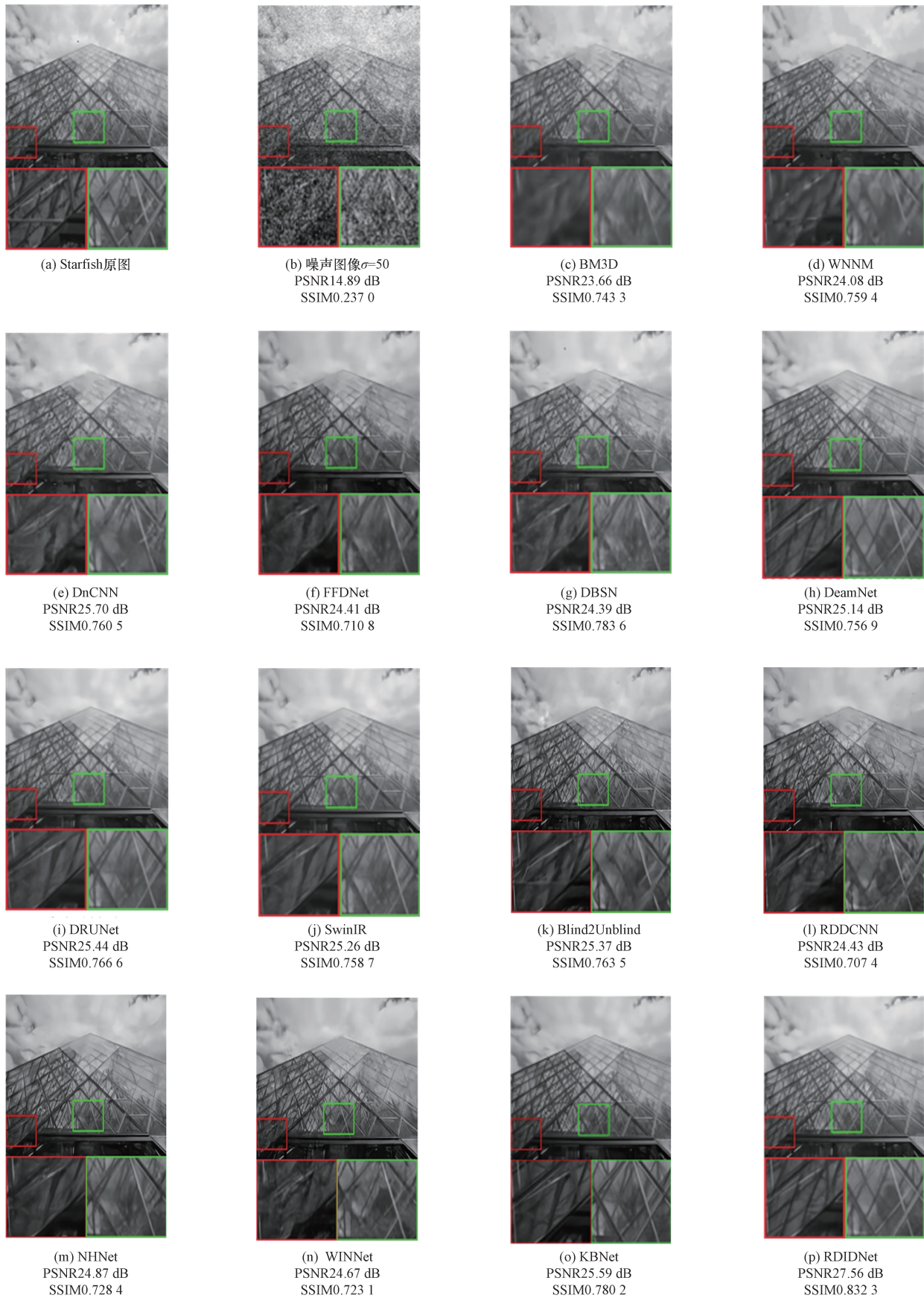
图7 不同方法对BSD68中“test044”添加 $\sigma=50$ 高斯噪声后降噪效果对比Fig. 7 Comparison of denoising effects using different methods on “test044” in BSD68 after adding gaussian noise with $\sigma=50$

表1 在 $\sigma=50$ 高斯噪声下不同方法通过Set12数据集降噪得到的PSNR,SSIM比较
Table 1 Comparison of PSNR and SSIM obtained by different methods on the Set12 dataset for denoising after adding Gaussian noise with $\sigma=50$

模型	指标	Man	House	Peppers	Starfish	Monarch	Airplane	Parrot
BM3D	PSNR/dB	26.13	29.69	26.68	25.04	25.82	25.10	25.90
	SSIM	0.794 7	0.852 2	0.819 3	0.772 8	0.828 3	0.797 5	0.812 3
WNNM	PSNR/dB	26.45	30.33	26.95	25.44	26.32	25.42	26.14
	SSIM	0.784 4	0.822 2	0.799 9	0.759 5	0.834 3	0.784 0	0.784 3
DnCNN	PSNR/dB	27.26	29.96	27.35	25.64	26.83	25.83	26.42
	SSIM	0.807 7	0.818 5	0.809 0	0.772 2	0.851 3	0.797 8	0.795 2
FFDNet	PSNR/dB	27.24	30.36	27.41	25.68	26.92	25.79	26.57
	SSIM	0.813 8	0.827 3	0.816 4	0.775 0	0.858 5	0.799 7	0.800 4
DBSN	PSNR/dB	27.17	30.09	27.41	25.56	26.77	25.73	26.48
	SSIM	0.800 5	0.817 4	0.824 7	0.808 5	0.846 8	0.802 8	0.798 5
DeamNet	PSNR/dB	27.43	31.19	27.78	26.48	27.19	26.08	26.72
	SSIM	0.814 3	0.840 6	0.832 6	0.805 6	0.869 1	0.808 5	0.803 3
DRUNet	PSNR/dB	27.80	31.26	27.87	26.49	27.31	26.08	26.92
	SSIM	0.826 7	0.841 2	0.827 1	0.801 3	0.872 4	0.809 3	0.809 8
Blind2Unblind	PSNR/dB	26.32	29.72	26.84	25.62	25.96	25.21	26.19
	SSIM	0.796 2	0.811 2	0.800 1	0.816 3	0.834 7	0.784 7	0.782 4
RDDCNN	PSNR/dB	27.16	30.21	27.38	25.73	26.84	25.88	26.53
	SSIM	0.805 9	0.824 6	0.820 3	0.780 3	0.853 9	0.800 1	0.794 6
NHNet	PSNR/dB	27.54	30.85	27.84	26.24	27.10	26.00	26.76
	SSIM	0.811 3	0.834 2	0.824 4	0.790 0	0.862 3	0.803 7	0.802 3
WINNet	PSNR/dB	27.21	30.57	27.47	25.89	26.99	25.95	26.63
	SSIM	0.810 8	0.829 6	0.818 7	0.781 2	0.863 7	0.805 2	0.798 0
SwinIR*	PSNR/dB	27.79	31.11	27.91	26.55	27.31	26.14	26.91
	SSIM	0.827 8	0.838 2	0.827 6	0.803 1	0.872 3	0.808 7	0.807 8
RDIDNet	PSNR/dB	27.63	31.51	27.96	27.56	27.39	26.20	26.78
	SSIM	0.820 0	0.842 8	0.835 3	0.832 3	0.875 4	0.811 0	0.804 9
模型	指标	Lena	Barbara	Boat	Man	Couple	平均值	
BM3D	PSNR/dB	29.05	27.22	26.78	26.81	26.46	26.72	
	SSIM	0.847 3	0.818 7	0.771 1	0.769 0	0.765 7	0.804 0	
WNNM	PSNR/dB	29.25	27.79	26.97	26.94	26.64	27.05	
	SSIM	0.805 3	0.819 9	0.708 3	0.709 1	0.713 7	0.771 8	
DnCNN	PSNR/dB	29.34	26.15	27.19	27.19	26.86	27.17	
	SSIM	0.810 4	0.768 2	0.717 5	0.721 5	0.723 8	0.782 9	
FFDNet	PSNR/dB	29.63	26.41	27.30	27.26	27.04	27.30	
	SSIM	0.821 8	0.779 2	0.725 3	0.727 7	0.734 3	0.789 9	
DBSN	PSNR/dB	29.46	26.51	27.16	27.11	26.86	27.19	
	SSIM	0.809 2	0.786 2	0.712 8	0.707 2	0.726 3	0.785 3	
DeamNet	PSNR/dB	29.96	27.62	27.61	27.38	27.45	27.74	
	SSIM	0.832 6	0.828 3	0.741 7	0.733 3	0.758 3	0.805 7	
DRUNet	PSNR/dB	30.15	28.16	27.66	27.46	27.59	27.90	
	SSIM	0.839 2	0.842 3	0.744 0	0.737 2	0.765 1	0.809 6	
Blind2Unblind	PSNR/dB	29.14	27.36	26.84	26.96	26.67	26.90	
	SSIM	0.801 7	0.821 0	0.707 1	0.706 3	0.720 1	0.781 8	
RDDCNN	PSNR/dB	29.32	26.36	27.23	27.22	26.88	27.23	
	SSIM	0.812 7	0.776 1	0.719 3	0.720 8	0.725 3	0.786 2	
NHNet	PSNR/dB	29.83	27.19	27.46	27.32	27.28	27.62	
	SSIM	0.825 8	0.801 3	0.728 0	0.727 5	0.745 9	0.796 4	
WINNet	PSNR/dB	29.67	26.69	27.37	27.31	27.13	27.41	
	SSIM	0.822 4	0.793 1	0.728 9	0.729 7	0.740 4	0.793 5	
SwinIR*	PSNR/dB	30.11	28.41	27.70	27.45	27.53	27.91	
	SSIM	0.837 4	0.848 0	0.743 7	0.737 5	0.762 5	0.809 6	
RDIDNet	PSNR/dB	30.07	28.11	27.73	27.44	27.57	28.00	
	SSIM	0.836 8	0.841 6	0.745 7	0.735 3	0.764 3	0.812 1	

注:BM3D、WNNM 基于传统方法; * 基于 ViT;其他基于 CNN。

表2 不同程度高斯噪声下各方法降噪后的 PSNR、SSIM 平均值比较
Table 2 Comparison of the average PSNR and SSIM of different methods for denoising under various levels of Gaussian noise

模型	指标	$\sigma = 15$		$\sigma = 25$		$\sigma = 50$	
		Set12	BSD68	Set12	BSD68	Set12	BSD68
BM3D	PSNR/dB	32.37	31.07	29.97	28.57	26.72	25.62
	SSIM	0.895 2	0.871 7	0.850 4	0.801 3	0.804 0	0.686 4
WNNM	PSNR/dB	32.70	31.37	30.28	28.83	27.05	25.87
	SSIM	0.898 2	0.876 6	0.857 7	0.808 7	0.771 8	0.698 2
DnCNN	PSNR/dB	32.86	31.73	30.44	29.23	27.17	26.23
	SSIM	0.903 1	0.890 7	0.862 2	0.827 8	0.782 9	0.718 9
FFDNET	PSNR/dB	32.75	31.63	30.43	29.19	27.30	26.29
	SSIM	0.902 7	0.890 2	0.863 4	0.828 9	0.789 9	0.724 5
DBSN	PSNR/dB	32.76	31.63	30.32	29.12	27.19	26.19
	SSIM	0.906 2	0.891 0	0.860 1	0.820 0	0.785 3	0.720 3
DeamNet	PSNR/dB	33.19	31.91	30.81	29.44	27.74	26.54
	SSIM	0.909 7	0.895 7	0.871 7	0.837 3	0.805 7	0.736 8
DRUNet	PSNR/dB	33.25	31.91	30.94	29.48	27.90	26.59
	SSIM	0.909 8	0.895 2	0.873 2	0.837 1	0.809 6	0.737 8
HDCNN	PSNR/dB	32.86	31.74	30.44	29.25	27.20	26.23
	SSIM	—	—	—	—	—	—
Blind2Unblind	PSNR/dB	32.46	31.44	30.09	28.99	26.90	26.09
	SSIM	0.897 1	0.884 7	0.854 4	0.820 3	0.781 8	0.715 2
RDDCNN	PSNR/dB	32.88	31.76	30.47	29.25	27.23	26.30
	SSIM	0.903 3	0.891 4	0.862 6	0.828 6	0.786 2	0.721 0
NHNet	PSNR/dB	33.15	31.85	30.77	29.37	27.62	26.43
	SSIM	0.906 4	0.893 1	0.867 5	0.831 4	0.796 4	0.726 1
WINNet	PSNR/dB	32.85	31.70	30.54	29.27	27.41	26.36
	SSIM	0.902 8	0.890 2	0.865 2	0.830 0	0.793 5	0.727 5
SwinIR *	PSNR/dB	33.36	31.97	31.01	29.50	27.91	26.58
	SSIM	0.911 0	0.896 0	0.874 1	0.837 6	0.809 6	0.737 7
RDIDNet	PSNR/dB	33.41	31.92	31.06	29.51	28.00	26.60
	SSIM	0.912 6	0.895 4	0.875 9	0.839 1	0.812 1	0.739 1

注:BM3D、WNNM 基于传统方法; * 基于 ViT;其他基于 CNN。

2.4 消融实验

RDIDNet 是以无注意力机制、采用 MSE 损失函数的 4 层 U-Net 为基准模型。基准模型的具体工作过程如下:将含噪声图像通过特征域模块扩展通道后,送入 U-Net 进行特征提取,将此操作重复 4 次后,将结果经重建模块进行图像恢复。在模型训练过程中以 MSE 为损失函数。

对 RDIDNet 进行消融实验,验证在基准模型上使用 RDU Sub-Network、DE-CAM、PSNRLoss 进行网络优化的有效性。选用污染较严重的 $\sigma = 50$ 高斯噪声,使用数据集 Set12 进行测试,对评价指标 PSNR、SSIM 取均值,分析结果如表 3 所示。对比基准模型可见,RDU Sub-Network 将 PSNR、SSIM 分别提升 0.31 dB、0.007 8;添加 DE-CAM 注意力模块将 Baseline 的 PSNR、SSIM 分别提升 0.25 dB、0.006 6;同时采用 RDU Sub-Network 和 DE-CAM 注意力模块将基准模型的 PSNR、SSIM 分别提升 0.46 dB、0.011 9;同时采用 RDU Sub-Network、DE-CAM 注意力模块和替换

表3 消融实验中不同模块对 PSNR、SSIM 的影响

Table 3 The impact of different modules on PSNR and SSIM in ablation studies

模块	指标平均值	
Baseline	PSNR/dB	27.52
	SSIM	0.799 6
Baseline + DE-CAM	PSNR/dB	27.77
	SSIM	0.806 2
Baseline + RDU Sub-Network	PSNR/dB	27.83
	SSIM	0.807 4
Baseline + DE-CAM + RDU Sub-Network	PSNR/dB	27.98
	SSIM	0.811 5
Baseline + DE-CAM + PSNRLoss + RDU Sub-Network	PSNR/dB	28.00
	SSIM	0.812 1

MSELoss 为 PSNRLoss 得到模型最佳效果,将基准模型的 PSNR、SSIM 分别提升 0.48 dB、0.012 5。

3 结论

基于 CNN 中 U-Net 为基准模型,提出一种改进的降噪网络 RDIDNet,在降噪性能上优于当前基于

CNN 和 ViT 的主流降噪网络,所体现的创新如下。

(1) 利用所设计的局部残差块组成全局残差模块替代编码器-解码器中简单的单层卷积,不仅加强了模型的非线性映射,而且全局残差模块中局部残差连接和全局残差连接的设计也加强了模型的稳定性。

(2) 通过双元素卷积注意力机制替代编码解码阶段中同尺度特征间的跳跃连接、不同尺度特征间的简单拼接,实现自适应特征融合。

(3) 采用 PSNR 评价函数作为损失函数,获得了比 L1、L2 损失函数更高的性能表现。

在对噪声污染较严重的 $\sigma = 50$ 高斯噪声的降噪性能分析中,通过 Set12 测试集得到其降噪后图像 PSNR、SSIM 均值分别为 28.00 dB、0.812 1;通过 BSD68 测试集得到其降噪后图像 PSNR、SSIM 均值分别为 26.60 dB、0.739 1。与传统方法 BM3D、基于 ViT 架构的 SwinIR 方法、最新的基于 CNN 的 NHNet 方法,以及其他几种主流方法相比,经 RDIDNet 降噪后图像的视觉效果和评价参数均具有优势,这说明 RDIDNet 性能稳定,在有效消除图像噪声时能够更好地保留图像细节。

参 考 文 献

- [1] 白茹鑫, 梁尚敏. 图像处理中应用图像降噪算法的研究综述[J]. 现代信息科技, 2024, 8(10): 21-25, 31.
Bai Ruxin, Luan Shangmin. Research review of the application of image denoising algorithms in image processing[J]. Modern Information Technology, 2024, 8(10): 21-25, 31.
- [2] 季维勋. 一种新的基于 SVD 图像降噪方法的探究[J]. 科学技术与工程, 2010, 10(20): 4949-4954.
Ji Weixun. A new research for the method of based on SVD image denoising[J]. Science Technology and Engineering, 2010, 10(20): 4949-4954.
- [3] Zhang K, Zuo W, Chen Y, 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.
- [4] Zhang K, Zuo W, 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] Guo S, Yan Z, Zhang K, et al. Toward convolutional blind denoising of real photographs[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach; IEEE, 2019: 1712-1722.
- [6] 白瑞君, 李众, 张启尧, 等. 一种改进的深度卷积神经网络图像去噪算法[J]. 科学技术与工程, 2019, 19(36): 247-252.
Bai Ruijun, Li Zhong, Zhang Qiyao, et al. Research on an improved deep convolution neural network image denoising algorithm[J]. Science Technology and Engineering, 2019, 19(36): 247-252.
- [7] Zhang K, Li Y, Zuo W, et al. Plug-and-play image restoration with deep denoiser prior[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021, 44(10): 6360-6376.
- [8] Liang J, Cao J, Sun G, et al. SwinIR: image restoration using Swin Transformer[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. Montreal; IEEE, 2021: 1833-1844.
- [9] Wang Z, Liu J, Li G, et al. Blind2Unblind: self-supervised image denoising with visible blind spots[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. New Orleans; IEEE, 2022: 2027-2036.
- [10] 姜旭, 赵荣彩, 刘勇杰, 等. 基于改进 U-Net 神经网络的图像去噪算法[J]. 科学技术与工程, 2022, 22(9): 3629-3635.
Jiang Xu, Zhao Rongcai, Liu Yongjie, et al. Image denoising algorithm based on improved U-Net neural network[J]. Science Technology and Engineering, 2022, 22(9): 3629-3635.
- [11] Zheng M, Zhi K, Zeng J, et al. A hybrid CNN for image denoising[J]. Journal of Artificial Intelligence and Technology, 2022, 2(3): 93-99.
- [12] Huang J J, Dragotti P L. WINNet: wavelet-inspired invertible network for image denoising[J]. IEEE Transactions on Image Processing, 2022, 31: 4377-4392.
- [13] 周联敏, 周冬明, 杨浩. 结合迁移学习的真实图像去噪算法[J]. 科学技术与工程, 2022, 22(34): 15237-15244.
Zhou Lianmin, Zhou Dongmin, Yang Hao. Real-world images denoising algorithm combined with transfer learning[J]. Science Technology and Engineering, 2022, 22(34): 15237-15244.
- [14] Lin H, Zhuang Y, Ding X, et al. Self-supervised image denoising using implicit deep denoiser prior[C]//Proceedings of the AAAI Conference on Artificial Intelligence. Washington DC; AAAI, 2023, 37(2): 1586-1594.
- [15] Zhang Y, Li D, Shi X, et al. Kbnnet: kernel basis network for image restoration[J]. ArXiv Preprint ArXiv, 2023: 2303.02881.
- [16] Xenarios I, Rice D W, Salwinski L, et al. DIP: the database of interacting proteins[J]. Nucleic Acids Research, 2000, 28(1): 289-291.
- [17] 徐少平, 李芬, 陈孝国, 等. 一种利用改进深度图像先验构建的图像降噪模型[J]. 电子学报, 2022, 50(7): 1573-1578.
Xu Shaoping, Li Fen, Chen Xiaoguo, et al. An image denoising model using the improved deep image prior[J]. Acta Electronica Sinica, 2022, 50(7): 1573-1578.
- [18] Woo S, Park J, Lee J Y, et al. CBAM: convolutional block attention module[J]. ArXiv, 2018: 1807.06521.
- [19] Ioffe S, Szegedy C. Batch normalization: accelerating deep network training by reducing internal covariate shift[J]. ArXiv, 2015: 1502.03167.
- [20] Yu F, Koltun V. Multi-scale context aggregation by dilated convolutions[J]. ArXiv, 2016: 1511.07122.
- [21] Kingma D P, Ba J. Adam: a method for stochastic optimization[J]. ArXiv, 2015: 1412.6980.
- [22] 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. Vancouver; IEEE, 2001(2): 416-423.
- [23] Timofte R, Agustsson E, Van Gool L, et al. Ntire 2017 challenge on single image super-resolution: methods and results[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops. Honolulu; IEEE, 2017: 114-125.

- [24] Zhang K, Zuo W, Chen Y, 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.
- [25] 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.
- [26] 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.
- [27] Gu S, Zhang L, Zuo W, et al. Weighted nuclear norm minimization with application to image denoising [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Columbus; IEEE, 2014: 2862-2869.
- [28] Wu X, Liu M, Cao Y, et al. Unpaired learning of deep image denoising [C]//European Conference on Computer Vision. Abu Dhabi; IEEE, 2020: 352-368.
- [29] Ren C, He X, Wang C, et al. Adaptive consistency prior based deep network for image denoising [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville; IEEE, 2021: 8596-8606.
- [30] Zhang Q, Xiao J, Tian C, et al. A robust deformed convolutional neural network (CNN) for image denoising [J]. CAAI Transactions on Intelligence Technology, 2023, 8(2): 331-342.
- [31] Zhang J, Cao L, Wang T, et al. NHNet: a non-local hierarchical network for image denoising [J]. IET Image Processing, 2022, 16(9): 2446-2456.