

# 基于混合注意力的对偶残差去噪网络

尹海涛<sup>\*</sup>, 邓皓

南京邮电大学自动化学院, 人工智能学院, 江苏 南京 210023

**摘要** 提出了一种基于混合注意力机制和对偶残差学习的图像去噪网络。该网络采用了基于不同尺寸卷积核的对偶残差网络学习结构, 不仅能降低更深网络结构的拟合难度, 同时还能表示图像中的多尺度结构。所提去噪网络采用局部和非局部的混合注意力模块, 对卷积神经网络的特征通道进行自适应调整, 使得卷积神经网络不仅能注意图像的局部特征, 还能刻画图像中的长距离依赖关系。与几种常见深度去噪网络的对比实验表明, 本文算法能有效抑制不同强度的噪声, 并且针对高强度噪声的去除性能更优。

**关键词** 图像处理; 卷积神经网络; 图像去噪; 对偶残差网络; 混合注意力

中图分类号 TP391

文献标志码 A

doi: 10.3788/LOP202158.1410002

## Dual Residual Denoising Network Based on Hybrid Attention

Yin Haitao<sup>\*</sup>, Deng Hao

College of Automation and College of Artificial Intelligence, Nanjing University of Posts and Telecommunications, Nanjing, Jiangsu 210023, China

**Abstract** In this paper, an image denoising network based on a hybrid attention mechanism and dual residual learning is proposed. The network uses a dual residual network learning structure based on different sizes of convolution kernels, which can not only reduce the difficulty of fitting deeper network structures, but also represent the multi-scale structure in the image. In the proposed denoising network, the feature channels are adaptively adjusted through hybrid local and non-local attention modules. Such hybrid attention module ensures that convolutional neural network can not only pay attention to the local features, but also depict the long-range dependencies in image. By comparing with several common deep denoising networks, the experimental results show that the proposed method can effectively suppress noise at different levels, specifically for the high-level noise.

**Key words** image processing; convolutional neural network; image denoising; dual residual network; hybrid attention

**OCIS codes** 100.3020;100.3010;100.2000

## 1 引言

图像去噪是一种能有效去除图像中的噪声并能保持图像结构信息的图像处理技术, 有助于后续的图像分类、分割、识别等任务, 能进一步推广图像在监控、医学、遥感、交通等领域的应用。图像去噪技术一直是图像处理、计算机视觉以及相关应用领域

的研究热点。

传统图像去噪算法大致可以分为三类: 1) 空间域图像滤波去噪。此类算法主要是通过利用图像中相邻像素之间的结构关系, 设计相关的滤波器进行去噪, 例如高斯滤波<sup>[1]</sup>、维纳滤波<sup>[2]</sup>、双边滤波<sup>[3]</sup>、引导滤波<sup>[4]</sup>、非局部均值滤波<sup>[5]</sup>等。2) 变换域去噪。此类算法主要通过图像变换, 在变换域中实现图像

收稿日期: 2020-10-12; 修回日期: 2020-11-10; 录用日期: 2020-11-12

基金项目: 国家自然科学基金(61971237, 61501255)

通信作者: \*haitaoyin@njupt.edu.cn

信息与噪声的分离,并通过逆变换得到去噪图像,常见的变换包括傅里叶变换<sup>[6]</sup>、小波变换<sup>[7]</sup>等。3)稀疏表示去噪。主要是利用图像的稀疏特性对图像结构信息进行表示,经典算法包括基于 K-means 奇异值分解(K-SVD)字典学习算法<sup>[8]</sup>、非局部稀疏表示算法<sup>[9]</sup>、三维块匹配(BM3D)算法<sup>[10]</sup>以及低秩去噪算法<sup>[11]</sup>等。传统去噪算法在某些情况下也能获得较好的去噪效果,但算法仍依赖于较强的先验模型假设以及人工调参,算法的普适性和一般性还不够强。

近些年,卷积神经网络在图像处理、计算机视觉、模式识别和机器学习等领域取得了里程碑式的进展<sup>[12-14]</sup>,也成功应用于图像去噪。文献[15]利用卷积神经网络训练噪声图像到干净图像之间的非线性映射,并实现端到端的学习,取得了较好的去噪性能。目前,残差网络结构<sup>[16]</sup>被广泛应用于卷积神经网络的搭建,不仅解决深层卷积神经网络的训练问题,还提升了卷积神经网络的性能<sup>[17-18]</sup>。国内外研究人员也相继提出了基于残差网络的图像去噪模型并得到了广泛关注,例如 DnCNN 算法<sup>[19]</sup>、IRCNN 算法<sup>[20]</sup>以及 RDN 算法<sup>[21]</sup>等。此外,卷积神经网络也成功实现了对真实噪声的去除,比如文献[22]提出的 FFDNet 算法。

与传统图像去噪算法相比,基于卷积神经网络

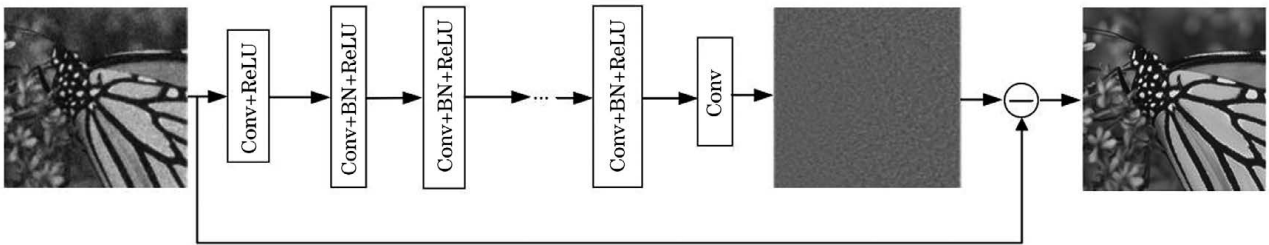


图 1 DnCNN 网络结构

Fig. 1 Network structure of DnCNN

DnCNN 网络共有 17 层,其中第 1 层由 64 个  $3 \times 3$  卷积核与 ReLU 激活函数组成;第 2~16 层由 64 个  $3 \times 3$  卷积核、批规范化操作处理(BN)与 ReLU 激活函数组成;第 17 层由 1 个  $3 \times 3$  的卷积核构成。DnCNN 网络通过下式残差学习策略进行训练,

$$\hat{x} = y - F(x), \quad (1)$$

式中: $\hat{x}$  为去噪图像; $y$  为噪声图像; $F(x)$  为卷积神经网络预测的噪声。

## 2.2 注意力模块

注意力机制能使卷积神经网络自适应地注意图

的图像去噪模型取得了较大的性能提升,但是现有去噪网络在高噪声情况下的效果仍有很大改进空间。首先,现有去噪网络主要是通过局部卷积来实现,感受视野较小,对图像的整体结构刻画还不足(比如非局部相似结构)。此外,针对图像中的纹理部分,由于没有重点“注意”这些显著特征,使得现有去噪网络仍会丢失一些边缘信息,从而产生边缘平滑现象。基于上述的分析,本文提出了基于混合注意力机制和对偶残差学习的图像去噪网络。本文设计的混合注意力机制由局部注意模块与非局部注意模块混合构成,每个模块包括主干分支与掩模分支。对于主干分支,我们采用对偶残差网络结构;对于掩模分支,通过大步长的卷积与反卷积操作实现特征降维和升维操作,从而扩大了感受视野范围,提升网络对图像全局结构的刻画能力。最后,采用整体残差学习思想对设计的网络进行参数训练。在不同噪声强度和不同类型噪声的实验结果表明,本文提出的网络能较好地保留边缘信息与纹理细节。

## 2 相关工作

### 2.1 DnCNN 去噪网络

DnCNN 去噪网络是一种具有代表性的基于残差学习的卷积神经网络去噪模型,网络整体架构示意图如图 1 所示。

像中显著区域,提升卷积神经网络的表示性能,广泛应用于自然语言处理、图像识别以及语音识别等领域。Squeeze-Excitation 网络(SE-Net)<sup>[23]</sup>和非局部注意神经网络(Non-Local-Net)<sup>[24]</sup>是两种常见的注意力网络。

SE-Net 主要包括 Squeeze 操作与 Excitation 操作两部分。假设  $x \in \mathbf{R}^{W \times H \times C}$  为 SE-Net 的输入特征,其中  $C$  为特征的通道数。Squeeze 操作对  $x$  进行全局平均池化,从而得到全局特征,表达式为

$$z_c = F_{sq}(x_c) = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W x_c(i, j),$$

$$c = 1, 2, \dots, C, \quad (2)$$

式中:  $x_c$  表示第  $c$  个特征通道。接下来, Excitation 操作由下式实现,

$$y = \sigma [W_2 \delta (W_1 z)], \quad (3)$$

式中:  $\delta$  为 ReLU 激活函数;  $\sigma$  代表 Sigmoid 激活函数;  $W_1 \in \mathbf{R}^{\frac{C}{r} \times C}$  和  $W_2 \in \mathbf{R}^{C \times \frac{C}{r}}$  为全连接变换分别实现特征的降维和升维;  $r$  为维度变化比例。SE-Net 可移植性较强, 可以嵌入到其他网络结构中。图 2 为 SE-Net 与 ResNet 结合的网络结构示意图。

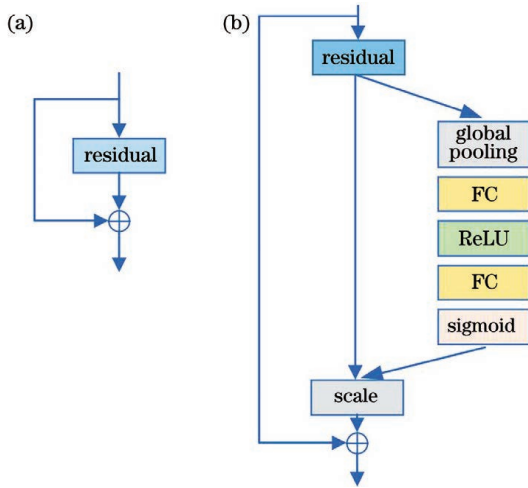


图 2 ResNet 结构与 SE-ResNet 结构。(a)传统 ResNet 模块; (b) SE-ResNet 模块

Fig. 2 Structure of ResNet block and SE-ResNet block.

(a) Traditional ResNet block; (b) SE-ResNet block

在非局部均值滤波和 SE-Net 基础上, 文献[24]提出了 Non-Local-Net, 能刻画图像中远距离间的关联性, 利用两个位置点之间的相似关系对特征进行加权, 从而增强网络对全局特征的感知能力。图 3 为 Non-Local-Net 网络结构示意图。

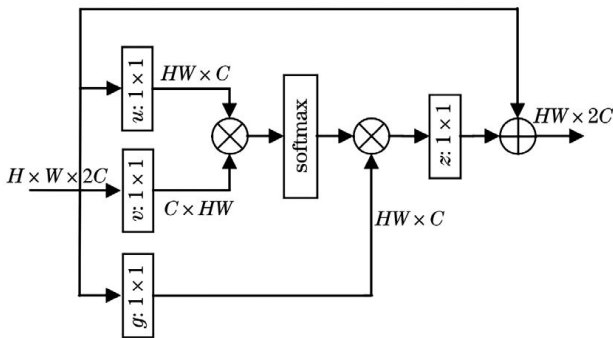


图 3 非局部模块结构

Fig. 3 Structure of non-local block

非局部操作借鉴了非局部均值滤波思想, 能表

示图像中不同位置间的依赖关系。具体定义为

$$y_i = \left[ \sum_{\forall j} k(x_i, x_j) g(x_j) \right] / \sum_{\forall j} k(x_i, x_j), \quad (4)$$

式中:  $k(x_i, x_j)$  计算输入  $x$  的第  $i$  与第  $j$  个位置之间的相似关系;  $g(x_j)$  对  $x_j$  进行  $g$  变换。其中  $k(x_i, x_j)$  可以定义为

$$k(x_i, x_j) = \exp \{ \langle u(x_i), v(x_j) \rangle \} = \exp \{ \langle W_u x_i, W_v x_j \rangle \}, \quad (5)$$

式中:  $W_u$  与  $W_v$  是权重矩阵。  $g$  函数可以定义为

$$g(x_j) = W_g x_j, \quad (6)$$

式中:  $W_g$  为权重矩阵。从而, 非局部注意模块输出可以表示为

$$z_i = W_z y_i + x_i =$$

$$W_z \text{softmax} \{ \langle u(x_i), v(x_j) \rangle \} g(x_j) + x_i, \quad (7)$$

式中:  $W_z$  为权重矩阵。

### 2.3 对偶残差网络结构

Liu 等<sup>[25]</sup> 提出了一种对偶残差网络模型并应用于图像修复。与残差网络相比, 对偶残差网络引入“上采样与下采样”、“大尺寸卷积核与小尺寸卷积核”等成对操作, 通过该对偶操作可以有效地提取图像多尺度特征。除此之外, 与传统的单一残差网络相比, 对偶残差网络引入了跨模块连接的操作, 即上一个模块内部与下一个模块内部相连接, 从而进一步提高网络性能。图 4 和图 5 分别为传统残差网络和对偶残差网络结构示意图。在传统的残差网络结构中(图 4),  $F_1, F_2, F_3$  等为常规卷积核; 在对偶残差网络中(图 5), 通过加入  $G_1, G_2, G_3$  等大小不同的卷积核, 构成了多组成对的操作。

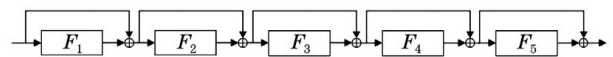


图 4 传统残差网络结构

Fig. 4 Structure of traditional residual network

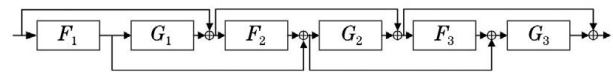


图 5 对偶残差网络结构

Fig. 5 Structure of dual residual network

## 3 本文算法

### 3.1 HDDNet 网络结构

本文通过结合局部注意力模块、非局部注意力模型以及对偶残差学习, 提出一种基于混合注意力的对偶残差去噪网络, 简称 HDDNet。图 6 为 HDDNet 网络架构示意图, 主要由三种模块构成: 特征提取模块、注意力机制模块(D-block)、非局部

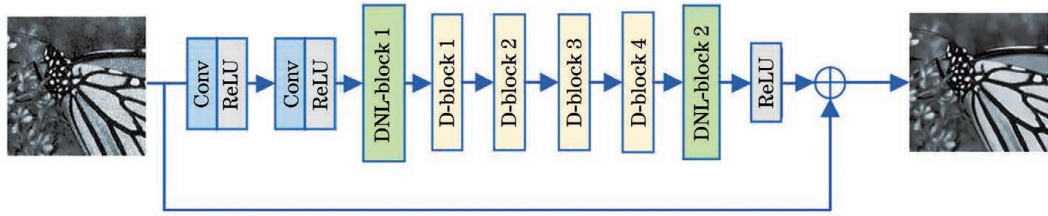


图 6 HDDNet 网络模型结构

Fig. 6 Structure of HDDNet network model

注意模块 (DNL-block)。其中, 特征提取模块由  $3 \times 3$  卷积和 ReLU 激活函数构成。

### 3.2 D-block 模块

图 7 为 D-block 模块的结构图, 主要构成部分包括: 前端和后端的残差单元、主干分支与掩模分支。

与原始残差模块不同, D-block 前端和后端采

用了一种简化的残差单元, 即去掉 BN 层, 由 2 层卷积层与 1 层 ReLU 激活函数层组成, 如图 8 所示。文献[26]指出简化残差单元能提高计算速度、减少显存, 便于搭建更深的神经网络; 文献[27]也表明简化残差单元具有较好的图像恢复性能。因此, 本文采用了简化形式的残差单元。

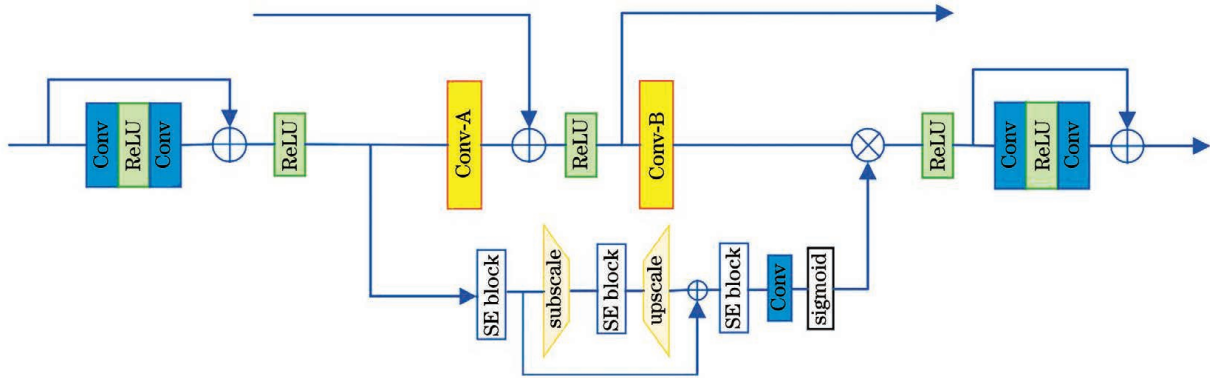


图 7 D-block 模块结构图

Fig. 7 Structure of D-block module

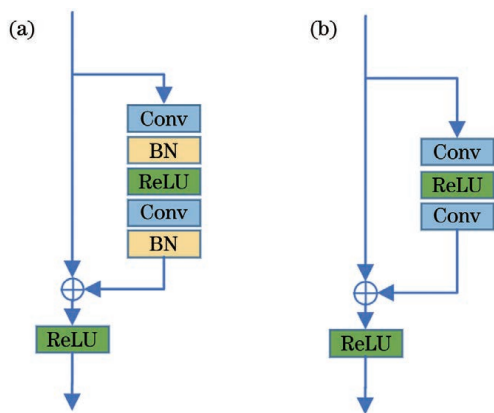


图 8 ResNet 模块的结构图。(a) 传统 ResNet 模块;  
(b) 简化残差模块

Fig. 8 Structure of ResNet block. (a) Traditional ResNet block; (b) simplified ResNet block

D-block 模块的主干分支由一对不同尺寸的卷积核 Conv-A 与 Conv-B 组成, 通过不同尺寸的卷积核可以有效地提取多尺度的特征<sup>[28-29]</sup>。此外, 为了进一步提高网络学习效率与增加网络模型深度, 在两个卷积核 (Conv-A 与 Conv-B) 之间分别添加了一

组残差连接, 分别连接前后两个模块, 从而形成对偶残差结构, 能有效地提高网络训练的效率并解决网络加深所带来的退化问题。

D-block 模块的掩模分支由 2 个 SE-Block、1 个 SE-Block 残差单元、1 个卷积层以及 1 个 Sigmoid 激活层构成。其中, SE-Block 通过 (2) 式和 (3) 式实现。SE-Block 残差单元包括下采样层、SE-Block、上采样层以及残差连接等部分, 能进一步扩大 SE-Block 的感受区域, 提升网络对纹理边缘信息的感知能力, 使得网络模型在处理复杂纹理图像时, 能够更加有效地关注边缘与纹理部分, 从而有效地保持图像的纹理与边缘细节。下采样层和上采样层分别通过大步长 (步长  $\geq 2$ ) 的卷积和反卷积实现。最后, 通过 Sigmoid 函数产生权重并对主干分支的特征进行加权, 实现 D-block 模块的“注意”功能。

### 3.3 DNL-block 模块

图 9 为 DNL-block 模块的结构图, 也是由前端和后端的残差单元、主干分支与掩模分支等部分组成。其中, 前端和后端的残差单元、主干分支与

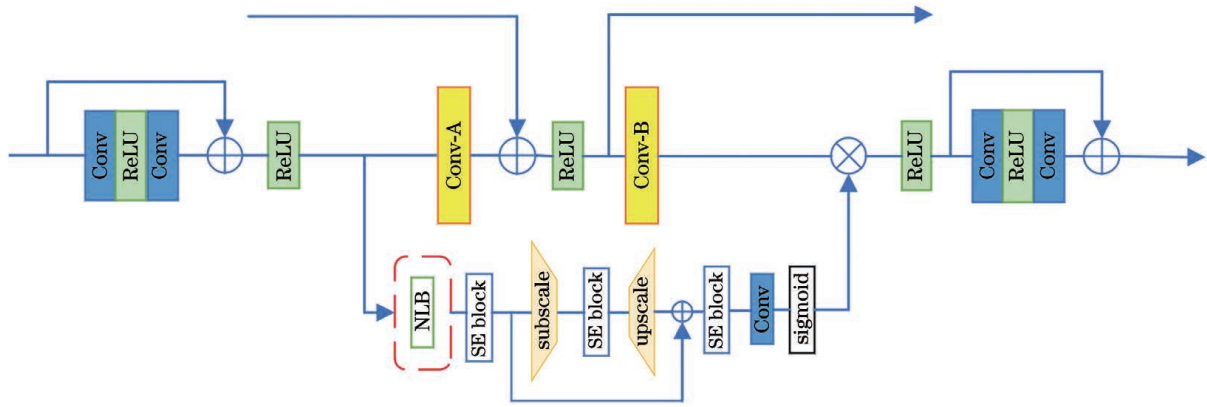


图 9 DNL-block 模块的结构

Fig. 9 Structure of DNL-block module

D-block 中相应单元的结构相同。不同之处在于：DNL-block 模块在掩模分支中添加了非局部注意模块(NLB),其中 NLB 结构如图 3 所示。传统卷积神经网络通常采用局部卷积操作,感受野较小,不利于网络捕获大范围信息以及不同位置之间的联系。从而导致网络在对图像进行表示时,不能有效地刻画图像的全局和非局部特征。因此,本文算法在 DNL-block 模块的分支中加入非局部注意模块,通过 NLB 模块进一步提升网络对图像长距离非局部特征的表示能力,同时提高网络模型的全局信息感知能力。

### 3.4 网络参数

本文提出的 HDDNet 网络整体结构包括 2 个特征提取模块、4 个 D-block 模块和 2 个 DNL-block 模块。D-block 模块与 DNL-block 模块中对偶卷积(Conv-A 与 Conv-B)的具体参数设置见表 1。

表 1 DNL-block 与 D-block 模块参数信息

Table 1 Parameters of DNL-block and D-block modules

Module	Conv-A kernel size	Conv-B kernel size	Dilation
DNL-block1	5	3	1
D-block1	7	5	1
D-block2	7	5	2
D-block3	11	7	2
D-block4	11	5	1
DNL-block2	11	7	3

采用均方误差作为损失函数对网络参数进行训练,其数学表达式为

$$L(\Theta) = \frac{1}{N} \sum_{i=1}^N \|R(y_i; \Theta) - (y_i - x_i)\|_F^2, \quad (8)$$

式中: $\Theta$  为 HDDNet 网络参数; $R(y_i; \Theta)$  为网络学习到的残差图像; $y_i$  为噪声图像; $x_i$  代表干净图像; $N$  为训练样本数。

采用 Train400 数据集作为训练数据集,训练过程的主要设置为:每个 epoch 随机提取 40 个尺寸为  $128 \times 128$  的图像块作为样本,并采用 Adam 优化器进行训练,初始学习率设为  $1 \times 10^{-4}$ ,每 10 个 epoch 学习率下降 0.2 倍。

## 4 实验结果与分析

### 4.1 实验设置

为了验证本文提出网络的去噪性能,与 DnCNN<sup>[19]</sup>、FDnCNN<sup>[19]</sup>、FFDNet<sup>[22]</sup>、IRCNN<sup>[20]</sup>、DuRN<sup>[25]</sup> 几种常见卷积神经网络去噪模型进行了实验对比,其中 DnCNN、FDnCNN、FFDNet、IRCNN 等算法的代码下载地址为 <https://github.com/cszn/KAIR>;DuRN 算法的代码下载地址为 <https://github.com/liu-vis/DualResidualNetworks>。DnCNN、FDnCNN、FFDNet、IRCNN 以及 DuRN 网络均采用网站提供的网络参数进行测试。测试图像由 Set12 数据集中 7 幅常用的图像组成,如图 10 所示。

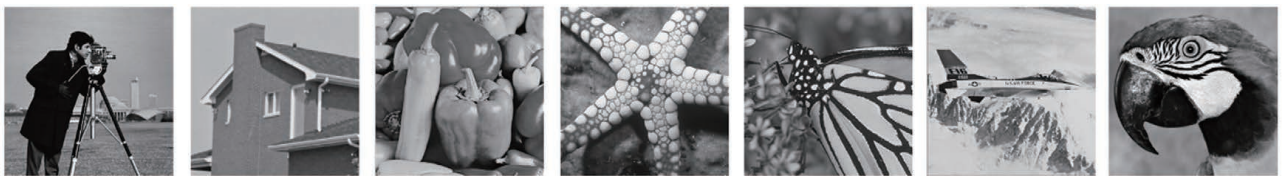


图 10 测试图像

Fig. 10 Test images

分别测试了高斯噪声和椒盐噪声两种类型噪声,其中噪声图像由人工合成,原图作为参考图像用于性能比较与分析。采用峰值信噪比(PSNR)、结构相似性(SSIM)<sup>[30]</sup>以及均方根误差(RMSE)作为评价指标对去噪结果质量进行客观评价。PSNR 与 SSIM 值越大表示去噪效果越好, RMSE 值越小表示去噪效果越好。

#### 4.2 高斯噪声去噪结果

表 2~4 分别给出了各种算法在不同高斯噪声

强度( $\sigma=15, \sigma=30, \sigma=50$ )下的 PSNR 值、SSIM 值和 RMSE 值。由于图像结构之间存在差异,使得每种去噪算法在不同测试图像上的去噪性能存在一定的差异。从表 2~4 中可以看出,本文提出的 HDDNet 在高强度高斯噪声情况下( $\sigma=30$  和  $\sigma=50$ )取得最好的平均指标值,表明 HDDNet 在高强度高斯噪声情况下去噪性能整体最优。对于低噪声情况( $\sigma=15$ ),虽然 HDDNet 获得的平均 PSNR 和 SSIM 指标不是最好的,但也都获得了第 2 名的

表 2 不同算法的 PSNR 值

Table 2 PSNR values of different methods

Noise level	Method	Cameraman	House	Peppers	Starfish	Butterfly	Airplane	Parrot	Average
$\sigma=15$	DnCNN	<b>32.59</b>	34.99	<b>33.24</b>	32.13	<b>33.25</b>	31.67	<b>31.88</b>	<b>32.82</b>
	FDnCNN	<b>32.59</b>	<b>35.13</b>	33.17	32.07	33.14	31.67	31.85	32.80
	FFDNet	32.37	35.05	33.02	31.95	32.92	31.55	31.79	32.66
	IRCNN	32.53	34.88	33.21	31.96	32.98	31.66	31.88	32.73
	DuRN	32.29	34.73	33.02	32.01	32.95	31.62	31.80	32.63
	HDDNet	<b>32.59</b>	34.96	33.21	<b>32.15</b>	33.22	<b>31.69</b>	31.86	32.81
$\sigma=30$	DnCNN	29.28	32.29	29.85	28.28	29.42	28.19	28.59	29.42
	FDnCNN	<b>29.43</b>	<b>32.59</b>	29.92	28.37	29.43	28.21	<b>28.70</b>	29.52
	FFDNet	29.28	32.57	29.87	28.34	29.39	28.13	28.65	29.46
	IRCNN	29.28	32.19	29.79	28.14	29.22	28.14	28.62	29.34
	DuRN	28.99	32.26	29.53	28.35	29.24	28.08	28.52	29.28
	HDDNet	<b>29.43</b>	32.43	<b>29.98</b>	<b>28.56</b>	<b>29.52</b>	<b>28.23</b>	28.65	<b>29.53</b>
$\sigma=50$	DnCNN	27.26	29.96	27.35	25.6	26.83	25.83	26.42	27.04
	FDnCNN	<b>27.35</b>	30.25	<b>27.41</b>	25.66	26.84	25.81	<b>26.59</b>	27.13
	FFDNet	27.34	<b>30.36</b>	<b>27.41</b>	25.68	26.92	25.79	26.57	27.14
	IRCNN	27.16	29.90	27.33	25.48	26.66	25.78	26.48	26.97
	DuRN	26.91	30.14	27.13	25.71	26.73	25.87	26.35	26.98
	HDDNet	27.33	30.35	27.37	<b>25.89</b>	<b>26.97</b>	<b>25.90</b>	26.47	<b>27.18</b>

表 3 不同方法的 SSIM 值

Table 3 SSIM values of different methods

Noise level	Method	Cameraman	House	Peppers	Starfish	Butterfly	Airplane	Parrot	Average
$\sigma=15$	DnCNN	0.9131	0.8855	0.9121	0.9146	0.9501	0.9077	0.9049	0.9126
	FDnCNN	0.9132	0.8870	0.9119	0.9135	0.9503	0.9080	0.9047	0.9127
	FFDNet	0.9118	<b>0.8877</b>	0.9112	0.9126	0.9491	0.9074	0.9045	0.9120
	IRCNN	0.9113	0.8831	0.9107	0.9123	0.9477	0.9064	0.9039	0.9108
	DuRN	0.9089	0.8834	<b>0.9153</b>	<b>0.9205</b>	0.9487	0.9071	<b>0.9078</b>	<b>0.9131</b>
	HDDNet	<b>0.9137</b>	0.8858	0.9117	0.9147	<b>0.9505</b>	<b>0.9081</b>	0.9050	0.9128

续表

Noise level	Method	Cameraman	House	Peppers	Starfish	Butterfly	Airplane	Parrot	Average
$\sigma=30$	DnCNN	0.8500	0.8518	0.8609	0.8453	0.9025	0.8511	0.8425	0.8577
	FDnCNN	0.8593	<b>0.8542</b>	0.8648	0.8463	0.9071	0.8537	0.8463	0.8617
	FFDNet	0.8599	<b>0.8542</b>	0.8652	0.8457	0.9073	0.8537	0.8467	0.8618
	IRCNN	0.8530	0.8474	0.8561	0.8412	0.8990	0.8475	0.8427	0.8553
	DuRN	0.8482	0.8507	0.8634	<b>0.8541</b>	0.9020	0.8494	<b>0.8474</b>	0.8593
	HDDNet	<b>0.8611</b>	0.8538	<b>0.8661</b>	0.8502	<b>0.9082</b>	<b>0.8545</b>	0.8455	<b>0.8628</b>
$\sigma=50$	DnCNN	0.8077	0.8185	0.8090	0.7722	0.8513	0.7978	0.7952	0.8074
	FDnCNN	0.8107	0.8245	0.8137	0.7747	0.8553	0.7986	0.7995	0.8110
	FFDNet	0.8138	0.8273	<b>0.8164</b>	0.7750	<b>0.8585</b>	0.7997	<b>0.8004</b>	0.8130
	IRCNN	0.8028	0.8159	0.8044	0.7675	0.8454	0.7953	0.7953	0.8038
	DuRN	0.7954	0.8211	0.8082	0.7800	0.8451	0.7954	0.7921	0.8053
	HDDNet	<b>0.8142</b>	<b>0.8293</b>	0.8144	<b>0.7812</b>	0.8575	<b>0.8012</b>	0.7971	<b>0.8136</b>

表 4 不同算法的 RMSE 值

Table 4 RMSE values of different methods

Noise level	Method	Cameraman	House	Peppers	Starfish	Butterfly	Airplane	Parrot	Average
$\sigma=15$	DnCNN	5.99	4.54	<b>5.55</b>	6.31	<b>5.55</b>	6.66	<b>6.49</b>	<b>5.87</b>
	FDnCNN	5.99	<b>4.47</b>	5.60	6.36	5.62	6.65	6.52	5.89
	FFDNet	6.14	4.51	5.70	6.44	5.76	6.74	6.56	5.98
	IRCNN	6.03	4.60	5.57	6.44	5.72	6.66	<b>6.49</b>	5.93
	DuRN	6.20	4.68	5.70	6.40	5.74	6.69	6.55	5.99
	HDDNet	<b>5.98</b>	4.56	5.57	<b>6.30</b>	5.57	<b>6.64</b>	6.51	<b>5.87</b>
$\sigma=30$	DnCNN	8.76	6.19	8.20	9.82	8.62	9.93	9.48	8.72
	FDnCNN	<b>8.61</b>	<b>5.98</b>	<b>8.14</b>	9.73	8.61	9.91	<b>9.36</b>	8.62
	FFDNet	8.76	6.00	8.19	9.76	8.65	10.00	9.42	8.68
	IRCNN	8.76	6.27	8.26	9.99	8.82	9.99	9.46	8.74
	DuRN	9.06	6.22	8.51	9.75	8.80	10.05	9.57	8.85
	HDDNet	<b>8.61</b>	6.09	<b>8.14</b>	<b>9.52</b>	<b>8.52</b>	<b>9.89</b>	9.41	<b>8.60</b>
$\sigma=50$	DnCNN	11.05	8.10	10.94	13.32	11.62	13.04	12.18	11.46
	FDnCNN	<b>10.94</b>	7.83	<b>10.87</b>	13.29	11.60	13.07	<b>11.95</b>	11.36
	FFDNet	11.08	<b>7.74</b>	<b>10.87</b>	13.26	11.50	13.09	11.97	11.36
	IRCNN	11.18	8.15	10.96	13.57	11.85	13.11	12.10	11.56
	DuRN	11.51	7.94	11.22	13.21	11.75	12.98	12.28	11.55
	HDDNet	10.97	<b>7.74</b>	10.91	<b>12.95</b>	<b>11.44</b>	<b>12.93</b>	12.11	<b>11.29</b>

结果,与最优结果的差别很小,比如平均 PSNR 值与最优的 DnCNN 仅差 0.01;平均 SSIM 值与最优的 DuRN 仅差 0.0003。此外,表 4 中的 RMSE 指

标结果表明 HDDNet 在  $\sigma=15$  噪声强度下的平均 RMSE 值是最优的。这也验证了 HDDNet 在低噪声情况下也具有一定的竞争力。

图 11 和图 12 分别给出了“海星”( $\sigma=30$ )和“蝴蝶”( $\sigma=50$ )去噪图像的视觉效果比较。“海星”图像中含有丰富的纹理结构,在去噪结果中很容易产生伪影。从图 11 中局部区域放大结果对比可以看出,DnCNN 算法、FFDNet 算法和 IRCNN 算法去结果图中仍存在一些噪声点,而 HDDNet 结果图中的纹理结构整体效果较好。类似地,图 12 的视觉效果对比表明,随着噪声强

度加大,DnCNN 算法、IRCNN 算法、DuRN 算法结果中仍存在一些噪声的伪影。此外,“蝴蝶”的翅膀区域具有显著的边缘特征,从图 12 中局部区域放大结果对比可以看出,HDDNet 算法在高噪声强度情况下仍能有效地保持“蝴蝶”的翅膀边缘特征。图 11 和图 12 的视觉对比进一步验证了 HDDNet 算法在纹理与边缘结构保持上具有一定的优势。

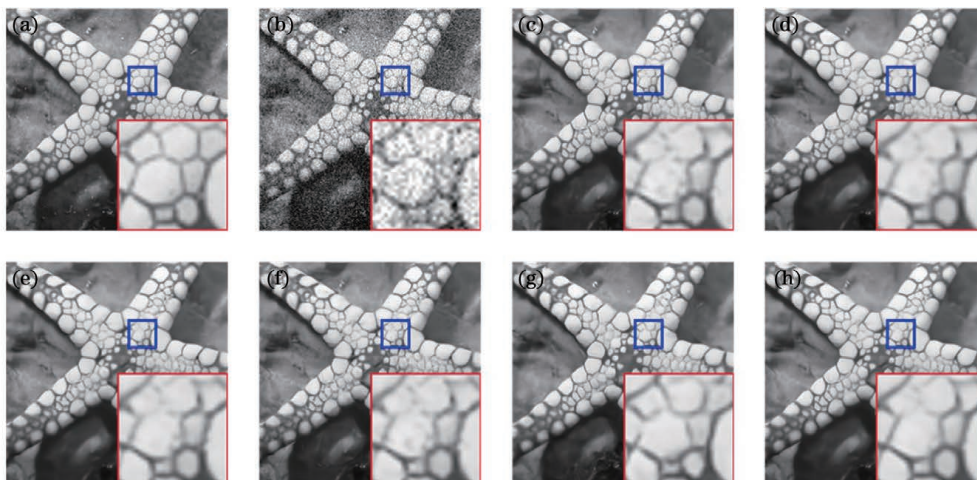


图 11 不同算法针对“海星”的去噪结果 ( $\sigma=30$ )。 (a)原图; (b)噪声图; (c) DnCNN; (d) FDnCNN; (e) FFDNet; (f) IRCNN; (g) DuRN; (h) HDDNet

Fig. 11 Denoised images of different algorithms on Starfish ( $\sigma=30$ ). (a) Original image; (b) noisy image; (c) DnCNN; (d) FDnCNN; (e) FFDNet; (f) IRCNN; (g) DuRN; (h) HDDNet

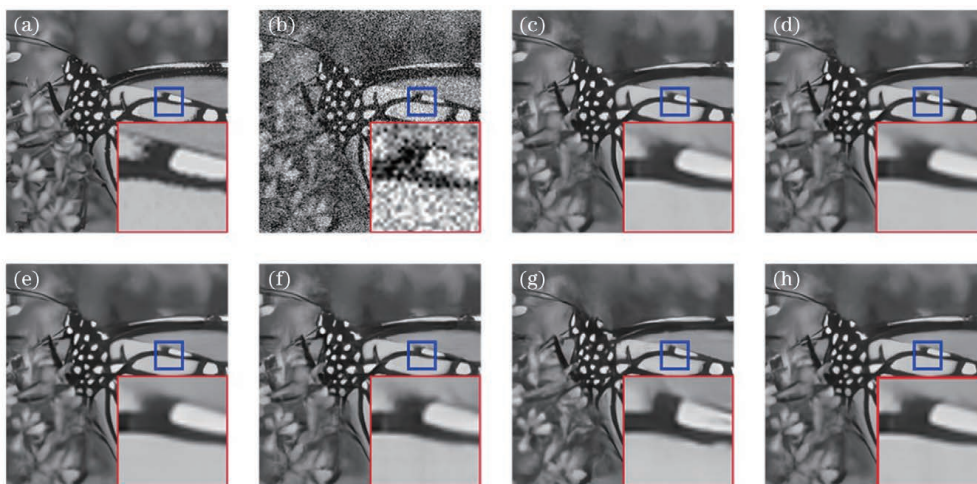


图 12 不同算法针对“蝴蝶”的去噪结果( $\sigma=50$ )。 (a)原图; (b)噪声图; (c) DnCNN; (d) FDnCNN; (e) FFDNet; (f) IRCNN; (g) DuRN; (h) HDDNet

Fig. 12 Denoised images of different algorithms on Butterfly ( $\sigma=50$ ). (a) Original image; (b) noisy image; (c) DnCNN; (d) FDnCNN; (e) FFDNet; (f) IRCNN; (g) DuRN; (h) HDDNet

### 4.3 椒盐噪声去噪结果

在本节测试了 HDDNet 模型针对椒盐噪声的去除效果。首先,针对椒盐噪声,在 Train400 数据集上重新训练了的 HDDNet 模型。然后,在图 10

给出的测试集上进行测试。图 13 为 HDDNet 模型在 10%椒盐噪声情况下的去噪效果。与干净的原图(图 10)对比可以得出,HDDNet 模型也能有效地去除椒盐噪声。



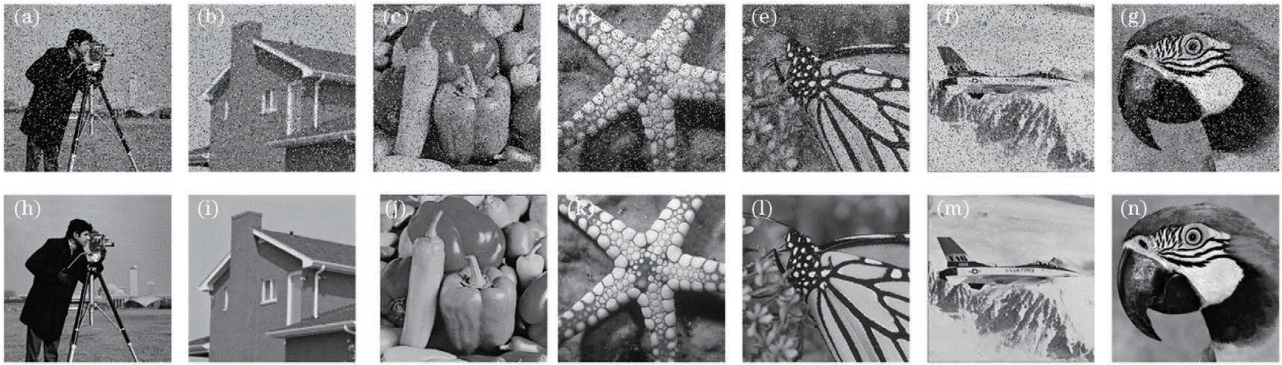


图 13 椒盐噪声去噪结果图(噪声密度为 10%)。(a)~(g)椒盐噪声图;(h)~(n)去噪后图像

Fig. 13 Denoised images for pepper & salt noise (noise density is 10%). (a)–(g) Pepper & salt noise images; (h)–(n) denoised images

## 5 结 论

本文提出了一种 HDDNet 去噪网络,该网络采用了局部注意力与非局部注意力的混合注意机制,能有效刻画图像中的局部和非局部特征,使网络更加有效地处理纹理边缘等显著区域。同时, HDDNet 采用了对偶残差结构,提高了信息利用效率和学习效率。实验结果验证了本文算法能取得较好的去噪效果,特别在高强度噪声情况下,本文算法仍能更加有效地减少边缘细节的损失、边界伪影和清晰度变差等问题。接下来,我们将继续研究卷积神经网络的自适应训练,提高卷积神经网络对不同图像特征的代表性能,进一步提升去噪性能。

### 参 考 文 献

- [1] Buades A, Coll B, Morel J M. A review of image denoising algorithms, with a new one[J]. *Multiscale Modeling & Simulation*, 2005, 4(2): 490-530.
- [2] Benesty J, Chen J D, Huang Y T. Study of the widely linear Wiener filter for noise reduction[C]//2010 IEEE International Conference on Acoustics, Speech and Signal Processing, March 14-19, 2010, Dallas, TX, USA. New York: IEEE Press, 2010: 205-208.
- [3] Gavaskar R G, Chaudhury K N. Fast adaptive bilateral filtering [J]. *IEEE Transactions on Image Processing*, 2019, 28(2): 779-790.
- [4] He K, Sun J, Tang X. Guided image filtering [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2013, 35(6): 1397-1409.
- [5] Buades A, Coll B, Morel J M. A non-local algorithm for image denoising [C] // 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '05), June 20-25, 2005, San Diego, CA, USA. New York: IEEE Press, 2005: 60-65.
- [6] Komatsu T, Tyon K, Saito T. 3-D mean-separation-type short-time DFT with its application to moving-image denoising[C]//2017 IEEE International Conference on Image Processing (ICIP), September 17-20, 2017, Beijing. New York: IEEE Press, 2017: 2961-2965.
- [7] Shen Y, Liu Q, Lou S Q, et al. Wavelet-based total variation and nonlocal similarity model for image denoising [J]. *IEEE Signal Processing Letters*, 2017, 24(6): 877-881.
- [8] Elad M, Aharon M. Image denoising via sparse and redundant representations over learned dictionaries [J]. *IEEE Transactions on Image Processing*, 2006, 15(12): 3736-3745.
- [9] Mairal J, Bach F, Ponce J, et al. Non-local sparse models for image restoration [C] // 2009 IEEE 12th International Conference on Computer Vision, September 29-October 2, 2009, Kyoto, Japan. New York: IEEE Press, 2009: 2272-2279.
- [10] 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.
- [11] Yu L Q, Qin Y L, Zhang X S. Denoising of strong noisy image via gradient reweighted non-local averaging over learned dictionaries [J]. *Laser & Optoelectronics Progress*, 2016, 53(11): 111002. 余临倩, 覃亚丽, 张晓帅. 基于字典学习的梯度加权非局部平均的强噪声图像去噪 [J]. *激光与光电子学进展*, 2016, 53(11): 111002.
- [12] Lu Y S, Li Y X, Liu B, et al. Hyperspectral data haze monitoring based on deep residual network [J]. *Acta Optica Sinica*, 2017, 37(11): 1128001. 陆永帅, 李元祥, 刘波, 等. 基于深度残差网络的高光谱遥感数据霾监测 [J]. *光学学报*, 2017, 37(11):

- 1128001.
- [13] Li S M, Lei G Q, Fan R. Depth map super-resolution reconstruction based on convolutional neural networks [J]. *Acta Optica Sinica*, 2017, 37(12): 1210002.  
李素梅, 雷国庆, 范如. 基于卷积神经网络的深度图超分辨率重建 [J]. *光学学报*, 2017, 37(12): 1210002.
- [14] Chen Q J, Li Y, Chai Y Z. A multi-focus image fusion algorithm based on depth learning [J]. *Laser & Optoelectronics Progress*, 2018, 55(7): 071015.  
陈清江, 李毅, 柴昱洲. 一种基于深度学习的多聚焦图像融合算法 [J]. *激光与光电子学进展*, 2018, 55(7): 071015.
- [15] Jain V, Seung S. Natural image denoising with convolutional networks [C] // *Advances in Neural Information Processing Systems*, December 8, 2008, Red Hook, NY, United States. New York: Curran Associates Inc, 2008: 769-776.
- [16] He K M, Zhang X Y, Ren S Q, et al. Deep residual learning for image recognition [C] // *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 770-778.
- [17] Yuan K P, Xi Z H. Image super resolution based on depth jumping cascade [J]. *Acta Optica Sinica*, 2019, 39(7): 0715003.  
袁昆鹏, 席志红. 基于深度跳跃级联的图像超分辨率重建 [J]. *光学学报*, 2019, 39(7): 0715003.
- [18] Xi Z H, Hou C Y, Yuan K P, et al. Super-resolution reconstruction of accelerated image based on deep residual network [J]. *Acta Optica Sinica*, 2019, 39(2): 0210003.  
席志红, 侯彩燕, 袁昆鹏, 等. 基于深层残差网络的加速图像超分辨率重建 [J]. *光学学报*, 2019, 39(2): 0210003.
- [19] 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.
- [20] 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)*, July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 2808-2817.
- [21] Zhang Y L, Tian Y P, Kong Y, et al. Residual dense network for image super-resolution [C] // *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, June 18-23, 2018, Salt Lake City, UT, USA. New York: IEEE Press, 2018: 2472-2481.
- [22] 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.
- [23] Hu J, Shen L, Albanie S, et al. Squeeze-and-excitation networks [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2020, 42(8): 2011-2023.
- [24] Liu D, Wen B, Fan Y, et al. Non-local recurrent network for image restoration [C] // *NIPS'18: Proceedings of the 32nd International Conference on Neural Information Processing Systems*, December 3, 2018, Red Hook, NY, United States. New York: Curran Associates Inc, 2018: 1673-1682.
- [25] Liu X, Suganuma M, Sun Z, et al. Dual residual networks leveraging the potential of paired operations for image restoration [C] // *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 15-20, 2019, Long Beach, CA, USA. New York: IEEE Press, 2019: 7000-7009.
- [26] Lim B, Son S, Kim H, et al. Enhanced deep residual networks for single image super-resolution [C] // *2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 1132-1140.
- [27] Zhang Y L, Li K P, Li K, et al. Residual non-local attention networks for image restoration [EB/OL]. (2019-03-24) [2020-10-10]. <http://arxiv.org/abs/1903.10082>.
- [28] Suganuma M, Ozay M, Okatani T. Exploiting the potential of standard convolutional autoencoders for image restoration by evolutionary search [EB/OL]. (2018-03-01) [2020-10-10]. <https://arxiv.org/abs/1803.00370>.
- [29] Wu L, Lü G Q, Xue Z T, et al. Super-resolution reconstruction of images based on multi-scale recursive network [J]. *Acta Optica Sinica*, 2019, 39(6): 0610001.  
吴磊, 吕国强, 薛治天, 等. 基于多尺度递归网络的图像超分辨率重建 [J]. *光学学报*, 2019, 39(6): 0610001.
- [30] Tong Y B, Zhang Q S, Qi Y P. Image quality assessing by combining PSNR with SSIM [J]. *Journal of Image and Graphics*, 2006, 11(12): 1758-1763.  
佟雨兵, 张其善, 祁云平. 基于 PSNR 与 SSIM 联合的图像质量评价模型 [J]. *中国图象图形学报*, 2006, 11(12): 1758-1763.