

中图法分类号: TP391 文献标识码: A 文章编号: 1006-8961(2023)07-2026-11

论文引用格式: Ding Y H, Wu H, Kong F L, Xu D and Yuan G W. 2023. A dual of real image noise-related blind denoising technique. Journal of Image and Graphics, 28(07):2026-2036(丁岳皓, 吴昊, 孔凤玲, 徐丹, 袁国武. 2023. 面向真实图像噪声的两阶段盲去噪. 中国图象图形学报, 28(07):2026-2036)[DOI:10.11834/jig.211020]

面向真实图像噪声的两阶段盲去噪

丁岳皓, 吴昊*, 孔凤玲, 徐丹, 袁国武

云南大学信息学院, 昆明 650500

摘要: 目的 现有的深度图像去噪算法在去除加性高斯噪声上效果显著,但在去除任意分布的真实图像噪声时表现不佳;去噪模型的深度在不断增加,但去噪效果上却并未能显著提高。对此,设计了一种简单有效的两阶段深度图像去噪算法。**方法** 首先基于注意力机制估计真实图像上的噪声分布水平,然后使用一个混合膨胀卷积和普通卷积的多尺度去噪模块进行非盲降噪。**结果** 在DND(darmstadt noise dataset)、SIDD(smartphone image denoising dataset)、Nam和PolyU(the Hong Kong Polytechnic University)等4个图像去噪领域常用数据集上进行去噪实验,选择峰值信噪比(peak signal-to-noise ratio, PSNR)和结构相似性(structural similarity, SSIM)作为去噪效果的评价指标,得到的平均PSNR值分别为39.23 dB, 38.54 dB, 40.45 dB, 37.34 dB,并与几种传统去噪方法和基于深度学习的去噪方法进行比较。实验结果表明,本文的去噪算法在去噪效果和视觉质量上有明显提升。同时,在SIDD数据集上进行消融实验以验证算法中模块的有效性。**结论** 本文算法使用的跳跃连接、噪声水平估计以及多尺度模块均可以有效提升真实图像去噪效果。与现有方法相比,本文算法不仅能有效去除真实图像噪声,而且能通过简单的模块参数设置控制去噪网络的计算效率。

关键词: 深度学习;真实图像去噪;注意力机制;噪声水平估计;多尺度模块

A dual of real image noise-related blind denoising technique

Ding Yuehao, Wu Hao*, Kong Fengling, Xu Dan, Yuan Guowu

School of Information Science and Engineering, Yunnan University, Kunming 650500, China

Abstract: Objective Image denoising technique can be as one of essential aspects in relevance to low-level vision. For image processing tasks, preprocessing is required for such contexts of high image quality, image segmentation, fine-grained detection and flaw detection. For real life scenario, image noise is inevitable because the image is collected and defected from the digital device itself, which will affect people's visual perception of the image. Image denoising-related research is originated from removing additive Gaussian white noise. There is a great difference between additive Gaussian white noise and real noise. The causes of real noise are diverse, and its distribution is different from that of Gaussian noise. The distribution of real noise is complex and may not only manipulated for one distribution, in which poor denoising results are obtained in terms of the removal of real noise. Moreover, traditional image denoising methods are required to set parameters manually, and deep learning based denoising algorithm has a poor effect in removing real noise. To get better results of denoising, the network model of these denoising methods is complex and time-consuming, but actual denoising

收稿日期:2021-11-04;修回日期:2022-07-01;预印本日期:2022-07-08

*通信作者:吴昊 haowu1982@vip.163.com

基金项目:国家自然科学基金项目(62061049);云南省应用基础研究计划重点项目(202001BB050032);云南省应用基础研究计划项目(2018FB100);云南省科技厅—云南大学“双一流”建设联合基金重点项目(2019FY003012)

Supported by: National Natural Science Foundation of China (62061049)

effect is not significantly improved. It is difficult to use an end-to-end algorithm to remove real image noise well beyond non-blind noise reduction. Therefore, we develop a dual of real image denoising model based on deep learning, which uses a simplified network structure and has ideal denoising results and visual performance on the image denoising dataset.

Method A blind denoising network model of two-stage modular denoising network (TMNet) is designed and illustrated, which is divided into two stages, the first stage is concerned of noise level estimation, and the second stage is concerned of non-blind denoising of the image. For the noise level estimation, a 4-layer fully convolutional neural network is used to estimate the noise level of the input image. The pooling layer is not used, but a convolutional layer is linked with a convolution kernel size of 3 only, and a padding operation is performed after the convolutional layer to keep the obtained feature size consistent from each layer. At the same time, due to the attention mechanism can effectively extract network features and improve network performance, before the last convolutional layer, the squeeze and excitation (SE) module can be used to process the features obtained via the noise level estimation, and such weights is adaptively assigned to the features of different channels. Finally, the estimated noise level and the input image are used as the output of this stage. The non-blind denoising stage can be focused on taking the estimated noise level and the image together as the input of this stage, and real image denoising becomes the problem of image denoising, which can remove the known noise distribution level. At this stage, we design a multi-scale structure, which can melt image features into two branches for processing. One branch is as an expanded convolutional layer, and the other branch is as an ordinary convolutional layer as well. Dilation convolution can be applied to get a larger receptive field by injecting holes into features of non-additional new parameters. In this way, image characteristics can be obtained under different receptive fields. The algorithm is designed for those skip connections are used in the multi-scale structure because effective skip connections are appeared in image denoising, and the loss of image information will be reduced in the process of parameter transmission when the number of network layers increases. For the convolution operation used in the non-blind denoising stage, the size of the convolution kernel is 3, and the corresponding padding operation is also guaranteed that the size of the image feature will not change in the entire network. Due to the high resolution of the original image of the smartphone image denoising dataset (SIDD) dataset used in the training model, our experiment is focused on dividing the image into 256 patches and performs 4 000 epochs of training randomly. The loss function used is L1 loss. All experiments are done on NVIDIA GTX 2080ti GPU. **Result** Four sort of commonly used datasets in image denoising fields including darmstadt noise dataset (DND), SIDD, Nam, and the Hong Kong Polytechnic University (PolyU) are used to verify the effectiveness of the algorithm, and the peak signal-to-noise ratio (PSNR) and structural similarity (SSIM) are used to evaluate the denoising results. Denoising experiments are carried out on DND, SIDD, Nam, and PolyU datasets, and the average PSNR values obtained are 39.23 dB, 38.54 dB, 40.45 dB, and 37.34 dB, respectively. Compared to the traditional algorithm, the denoising effect and visual sense are significantly improved, and the denoising effect can be similar to that of the deep learning denoising algorithm in related to fewer network layers. Furthermore, ablation experiments are designed to verify the effectiveness of the methods proposed, such as skip connection, channel attention, and noise level estimation. Experiments show that our methods proposed can effectively improve the effect of real image denoising to a certain extent. **Conclusion** We design a convolutional neural network (CNN) model for removing noise from real images. Our model has a simple network structure, in which most of the convolution operations with a convolution kernel size of 3 are carried out, and the network training speed is fast. Our network can get richer residual structure in image denoising as well. Additionally, we use multiple convolution to obtain more effective feature information. Experimental results show that our network is potential to removing real image noise further to some extent.

Key words: deep learning; real image denoising; attention mechanism; noise level estimation; multiscale module

0 引言

图像去噪是底层视觉研究的重要领域,也是许多图像处理任务的预处理步骤。去噪算法的好坏直

接关系到后续图像处理的效果,如图像分割、目标识别和边缘提取等(于虹和甄彤,2020)。图像去噪的实质是对数据本身的恢复和重建(管美静等,2020),经典的加性噪声模型一般可以表示为 $I' = I + n$,图像去噪任务就是将原始图像 I' 中的噪声 n 去除,得

到不含噪的图像 I 。在真实场景中,通过数字设备采集的图像都含有噪声,这些噪声通常来源广泛。在采集图像时,由于采集对象的纹理过于复杂,采集到的图像会有较多噪声。此外,受拍摄场景的影响,如亮度不均匀(过明/过暗)、物体运动和拍摄抖动等也会使图像产生噪声。另外,图像在传输过程中受传输介质影响也会产生多种真实噪声。这些噪声还会经过图像采集设备的机内系统处理,如去马赛克、图像灰度校正和压缩等(谢舒闽等,2020)。因此,受各种内部和外部因素的影响,在采集图像时容易产生特别复杂的真实噪声,且可能不是仅服从一种分布水平,这与去除特定分布噪声有很大不同。

近年许多基于深度学习的图像去噪算法都仅用于去除加性高斯白噪声。这种噪声模型的表示简单,去除也相对容易。在训练这种图像去噪算法时,只需在清晰图像上加上特定分布水平的噪声,即可得到用于训练的清晰图像与噪声图像对。这种去噪算法仅在去除特定分布水平噪声上有较好表现(王瀚深,2021),而真实场景下的噪声可能与加性高斯白噪声有很大不同。真实图像的噪声分布水平往往是未知的,而且在去噪时使用过高或过低的噪声水平都会导致最终的去噪效果不佳,若使用的噪声水平低于真实水平,真实图像上的噪声就会去除不干净,若高于真实水平就会有丢失边缘信息和图像细节信息的风险。因此,传统的卷积神经网络去噪器通常不能很好地处理真实场景下的噪声图像(谢舒闽等,2020)。而现有的一些去除真实图像噪声的方法在去噪效果上还有很大不足(刘金山,2021),且为了获得更好的去噪效果,这些网络的结构大多很复杂,需要花费大量的时间来训练网络,但实际的去噪效果却没有明显提高。同时,设计层次非常深的神经网络结构容易出现过拟合和非凸优化问题(徐少平等,2019)。所以,本文提出一种去除图像噪声的神经网络去噪模型,使用两阶段的网络结构对真实场景中的图像进行去噪,并在常用的图像去噪数据集上有较好的评价指标数值与视觉表现。本文的贡献如下:1)设计了一个两阶段的面向真实场景图像的盲去噪网络模型 TMNet (two-stage modular denoising net)。第1阶段用于噪声水平的估计,第2阶段对图像进行非盲去噪。2)设计了一个可以提取多尺度特征的模块结构。根据不同的真实场景,通过改变模块的个数可以保证去噪效果的同时减少网络参数。

1 相关工作

1.1 传统去噪算法

传统的用于图像降噪的方法主要分为基于空间和基于频率域两种。

基于空间域的算法常见的有双边滤波(Tomasi和Manduchi,1998)、非局部均值去噪(Buades等,2005)和引导滤波(He等,2013)等。双边滤波算法同时考虑图像的距离因素和像素值差异,在去噪时可以很好地保存图像边缘细节,但效率不是太高,花费的时间也比较长。引导滤波(导向滤波)通过一幅引导图对图像进行去噪处理,引导图可以是图像本身或其他单独的图像。非局部均值在去噪时使用图像包含的全部信息,虽然可以有效去除图像中的高斯噪声,但计算复杂度很高、耗时长,在实际应用中不适用。

基于频域的算法主要有傅里叶变换和小波变换(Chang等,2000)等。这类算法在有用信号和干扰信号的频谱没有重叠的前提下,才能将有用信号和干扰信号完全分开,但实际上很难分开。

除此之外,还有一些其他常见的传统去噪算法。利用稀疏编码去噪的KSVD(k-singular value decomposition)(Aharon等,2006)和TWSC(trilateral weighted sparse coding)算法(Xu等,2018b),利用滤波去噪的BM3D(block-matching and 3D filtering)算法(Dabov等,2008)以及利用外部先验去噪的EPLL算法(Zoran和Weiss,2011)和利用低秩去噪的WNNM(weighted nuclear norm minimization)算法(Gu等,2014)等。尽管传统方法在图像去噪的某些方面,如去除特定噪声水平的加性高斯白噪声,已经取得不错的性能,但大多数都需要手动设置参数以及单个降噪任务的特定模型(刘利平等,2021)。

1.2 基于深度学习的去噪算法

由于传统的去噪方法存在许多不足,同时深度学习已经在图像处理领域取得较好的效果,因而基于深度学习的图像去噪方法得到了迅猛发展。

在去除特定噪声方面,Jain和Seung(2008)首先提出用卷积神经网络(convolutional neural network, CNN)处理自然图像的去噪问题,说明特定形式的CNN可以被视为图像去噪的马尔可夫模型推断的结果的一种近似,但是神经网络模型可以避免马尔

可夫模型在概率学习和推断过程中的计算困难,从而降低计算的复杂度。Burger等人(2012)提出用多层感知机的方法实现神经网络去噪,并说明如果训练集足够,MLP(multi-layer perceptron)模型可以达到图像去噪的当前最优水平。MLP模型对噪声类型不敏感,对于不同类型的噪声,只要对带有该类噪声的图像进行处理,即可得到较好的结果。Chen和Pock(2017)提出一种可训练的非线性反应扩散模型TNRD(trainable nonlinear reaction diffusion),但在捕捉图像结构的全局特征方面有限。Zhang等人(2017)提出较深层的CNN网络DnCNN(denoising convolutional neural network)。为了解决网络层数加深导致的梯度弥散效应,DnCNN并不对图像进行学习,而是以输出与噪声的范数为损失函数来训练网络。Kim等人(2016)提出使用非常深的网络进行模型预测,使用残差学习的方法直接预测超分的残差,极大地加速了模型的训练过程,也使得最终输出中保留的细节较好。Zhang等人(2018)又提出了FFDNet(fast and flexible denoising convolutional neural network),将噪声强度与含噪图像一起输入网络,得到了较好的去噪效果,但真实图像噪声的分布情况是未知的。Tai等人(2017)提出一种长期记忆网络MemNet(memory network),采用密集链接的方法,最大限度地使信息在不同记忆单元之间流动,同时,网络中的记忆模块使用递归单元和门控单元,对靠前记忆单元传来的信息和本记忆单元内的信息进行权重分配,结合为记忆块最终输出。吴从中等人(2018)提出结合残差学习和感知损失的图像去噪方法,有效解决了图像去噪后边缘过光滑,易产生伪影的问题。

这些去噪算法的去噪对象都为特定分布的加性高斯白噪声,使用这些方法去除真实噪声时效果并不理想。于是,人们研究的重点又转移到设计去除真实图像噪声上来。Guo等人(2019)提出一个去除真实噪声的去噪算法CBDNet(convolutional blind denoising network),使用两阶段结构,分步对真实图像进行去噪。RIDNet(real image denoising network)(Anwar和Barnes,2019)利用残差结构上的残差来缓解低频信息的流动,并利用特征注意来挖掘信道的相关性,在合成图像和真实图像的定量和定性上都达到了当时的最高水平。Zhao等人(2019)提出一种金字塔真实图像去噪网络PRIDNet(pyramid real image denoising network),使用通道注意力机制对图

像不同通道上的特征进行提取,并自适应地调整通道重要性。Tian等人(2020)使用批量重归一化对图像进行去噪,通过增加网络宽度来提高网络的学习能力。Cheng等人(2021)提出一种基于子空间注意力模块的图像降噪算法,取得了较好的去噪效果。此外,基于深度学习的真实图像去噪方法还有利用相机管道模型去噪的CycleISP算法(Zamir等,2020a)、并行卷积流多尺度模型MIRNet(Zamir等,2020b)以及使用强化学习进行去噪的Path-Restore(Yu等,2022)模型等。

以上方法都使用了较多的网络层数,导致训练速度缓慢。另一些网络虽然使用较少的网络层数,但在去噪效果上不太理想。对此,设计了一个两阶段的真实图像去噪器,在噪声估计阶段使用通道注意力提高噪声水平估计性能,在非盲降噪阶段使用多尺度模块,可以通过模块的增减来控制网络层数,同时在去噪效果上也具有较好的表现。

2 算法模型

用于去除特定分布水平噪声的去噪算法如DnCNN,FFDNet,MemNet等,都是用一次端到端的算法直接去除图像噪声,但真实图像噪声分布水平是未知的,使用一次端到端的算法难以对真实图像有很好的去噪效果。所以需要首先对真实图像上的噪声分布水平进行估计,再将估计出来的噪声分布水平与真实图像一起输入算法的下一步骤,这样真实图像去噪问题就简化为去除特定分布水平的高斯噪声。另外,现有去噪算法在去除特定分布水平噪声上效果显著,对本文的去噪算法有先验作用。由此本文的去噪网络模型分为两个阶段:第1阶段对噪声图像的噪声水平估计;第2阶段对图像进行非盲去噪处理。算法模型如图1所示。

2.1 噪声水平估计阶段

这个阶段的任务是对真实图像上的噪声进行水平估计。真实噪声图像模型可表示为 $y = x + n(x)$,式中, y 表示真实噪声图像, x 表示干净图像, $n(x)$ 表示真实噪声,且 $n(x) \sim N(0, \sigma^2(x))$ 。由于网络中非盲降噪阶段的任务是去除真实噪声图像上的噪声 $n(x)$,但其分布水平参数 σ 未知,降噪模型无法准确

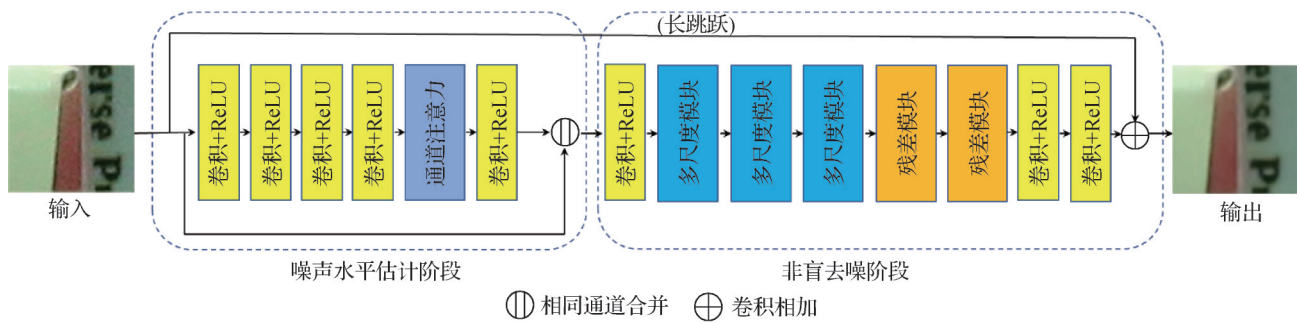


图1 算法模型图

Fig. 1 Algorithm model diagram

地学习到噪声图像到干净图像的映射关系,所以将其噪声水平估计阶段的结果作为分布水平参数 σ 输入非盲降噪阶段。

对此,本文设计了一个噪声水平估计子网。这一阶段使用4层的全卷积神经网络对输入图像的噪声水平进行估计,这一部分不使用池化层,也不进行BN(batch normalization)操作,而是仅使用卷积核大小为 3×3 的卷积层,并在每层后都进行相应的padding操作,使每一层得出来的特征大小都相同。由于注意力机制可以有效提取网络特征,提高网络性能,所以在最后一个卷积层前使用通道注意力机制(squeeze and excitation, SEblock)(Hu等,2019)对噪声水平估计得到的特征进行通道加权处理,最后将估计出的噪声水平和输入图像一起作为本阶段的输出。通道注意力结构如图2所示,其中“ \otimes ”表示逐通道乘法。

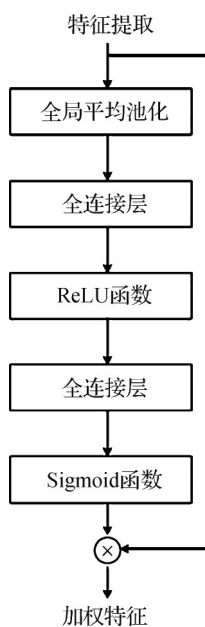


图2 SEblock结构

Fig. 2 The structure of SEblock

对于一个输入图像 x ,本文算法在噪声水平估计阶段的数学模型可表示为

$$x_1 = conv(SE(conv^4(x))) \quad (1)$$

式中, $conv$ 表示卷积加ReLU(rectified linear unit), SE (squeeze and excitation)表示通道注意力模块, x_1 表示噪声估计水平阶段的输出。

2.2 非盲降噪阶段

2.2.1 多尺度去噪模块

在这一阶段构造了一个多尺度结构,其中一个分支是两个卷积核大小为 3×3 的卷积层,另一个分支也同样是两个卷积核大小为 3×3 的卷积层,但是与第1个分支不同的是,对它使用了膨胀卷积的策略。膨胀卷积通过对特征注入空洞来获得更大的感受野,两个卷积层的膨胀率都为2,并使用相应的填充策略,使图像特征的大小保持相同。最后将两个分支获得的特征结合起来,并结合残差连接一起作为下一网络层的输入。这样,得到的多尺度结构一方面使用膨胀卷积获得较为庞大的特征信息,另一方面又保留了局部的特征信息,再将两种信息结合,这样就可以获得更加全面的特征。设计的多尺度模块结构如图3所示,图中颜色较浅的层是使用膨胀卷积的部分,“ \oplus ”表示对相同通道进行相加。

2.2.2 非盲降噪阶段

这个阶段将第1阶段的输出作为输入,即其输

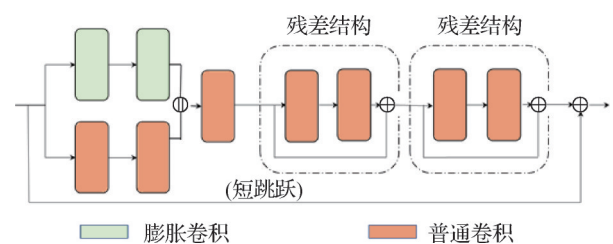


图3 多尺度模块结构

Fig. 3 The structure of multiscale module

入包括输入图像以及在输入图像上估计出来的噪声水平,因而在去噪模型的第2个阶段仅需去除输入图像上特定分布的噪声。由于残差网络在图像去噪上的不错表现,在这一阶段,残差结构扮演着重要的角色。

先使用一个卷积核大小为 3×3 的卷积层,后面跟着3个多尺度结构,并在网络最后使用两个残差结构以及两个普通卷积层调整最后输出的维度。每个残差结构由两个卷积层构成,并在第1个卷积层后使用ReLU操作,在第2个卷积后的输出和残差块的输入进行相加后再次使用ReLU。这样,一个残差块就部署完成。与第1阶段一样,本文算法在第2个阶段未使用BN和池化,并且同样使用了相应的padding操作,使特征大小在整个网络中都不改变。

在非盲降噪阶段使用模块化结构,可以根据真实场景的复杂情况,对网络深度进行调整。在拍摄条件较差的情况下,如光源不足、设备性能不高以及拍摄对象有大量的纹理细节,都会导致采集到的真实图像中的噪声增多,此时可以通过增加网络中的模块数来获得更好的去噪效果。本文算法在非盲降噪阶段数学模型为

$$\mathbf{x}_2 = \text{conv}^2(\text{Res}^2(\text{MD}^3(\text{conv}(\mathbf{x}_1)))) \quad (2)$$

式中, \mathbf{x}_1 表示非盲降噪阶段的输入, conv 表示卷积加ReLU操作, MD 表示本文设计的多尺度模块, Res 表示残差模块, \mathbf{x}_2 表示本阶段的输出。

2.2.3 损失函数

图像去噪领域常用的损失函数主要有L1损失函数和L2损失函数。L1损失函数是图像中每个像素真实值与预测值的绝对值之和,L2损失函数是

图像中每个像素真实值与预测值之间插值的平方和。由于L2损失函数对误差取平方,会放大真实值与预测值的误差,而且对相对较大的误差放大的更加明显,惩罚力度更大,因而使用L2损失函数可能会导致图像边缘模糊和细节信息丢失等。因此本文算法使用L1损失函数对模型进行优化,具体为

$$L(w) = \frac{1}{N} \sum_{i=1}^N \|TMNet(\mathbf{x}_i) - \mathbf{y}_i\| \quad (3)$$

式中, $TMNet(\cdot)$ 表示本文算法, w 表示网络所有要学习的权重参数, \mathbf{x} 表示原始图像, \mathbf{y} 表示图像对应的ground-truth。

3 实验结果

3.1 去噪效果比较

为了验证本文算法的去噪效果,与常见的传统去噪算法KSVD、TWSC和BM3D以及基于深度学习的图像去噪算法DnCNN、CBDNet和DRNet(Li等,2017)进行比较。

3.1.1 SIDD数据集

本文使用SIDD(smartphone image denoising dataset)数据集(Abdelhamed等,2018)作为训练数据集。该数据集的图像为png文件,是使用5个不同相机,在4种相机参数下,分别拍摄200个场景,每个场景连续拍摄150幅图像并进行一定处理后制作的图像去噪数据集。实验时将200个场景中的160个场景作为训练集,剩下的40个场景作为测试集。对SIDD Benchmark上的40个图像对进行测试,去噪结果如图4所示。

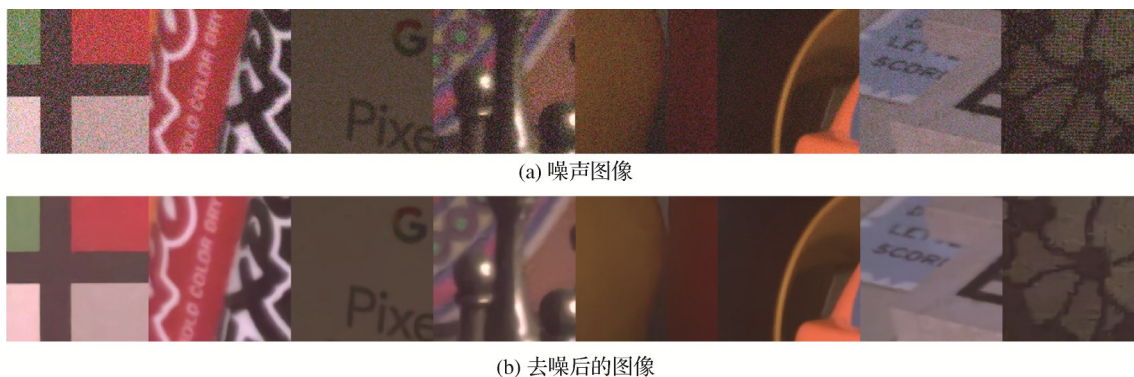


图4 SIDD数据集去噪结果

Fig. 4 Denoising results on the SIDD dataset ((a) noisy images; (b) denoised images)

本文方法在SIDD验证集上的峰值信噪比(peak signal-to-noise ratio, PSNR)为51.63 dB,结构相似性(structural similarity, SSIM)(Wang等,2004)为0.9958。表1是本文方法获得的去噪结果与FoE(fields of experts)(Roth和Black,2005)、CBDNet和KSVD等方法在SIDD Benchmark上去噪结果的比较。可以看出,本文方法在峰值信噪比和结构相似性上都有不错的表现。

表1 不同方法在SIDD数据集上去噪结果对比
Table 1 Comparison of denoising results of different methods on SIDD dataset

方法	sRGB		RAW	
	PSNR/dB	SSIM	PSNR/dB	SSIM
DnCNN	23.66	0.583	43.30	0.965
MLP	24.71	0.641	43.17	0.965
TNRD	24.73	0.643	42.77	0.945
FoE	25.58	0.792	43.13	0.969
BM3D	25.65	0.685	45.52	0.980
WNNM	25.78	0.809	44.85	0.975
KSVD	26.88	0.842	43.26	0.969
EPLL	27.11	0.870	40.73	0.935
CBDNet	33.28	0.868	-	-
本文	38.54	0.946	50.99	0.99

注:加粗字体表示各列最优结果,“-”表示无数据。

3.1.2 Nam数据集

Nam数据集(Nam等,2016)由11个静态场景和相应的无噪声图像组成,这些图像是通过500个噪声图像在同一场景中的平均值获得的,图像为png文件。表2是本文方法与CDnCNN-B、BM3D和CBDNet在Nam数据集上去噪效果比较。图5是使用本文去噪方法在Nam数据集上去噪结果。

表2 不同方法在Nam数据集上去噪结果对比
Table 2 Comparison of denoising results of different methods on Nam dataset

方法	PSNR/dB	SSIM
CDnCNN-B	37.49	0.927 2
BM3D	39.84	0.965 7
CBDNet	40.02	0.968 7
本文	40.45	0.989 8

注:加粗字体表示各列最优结果。

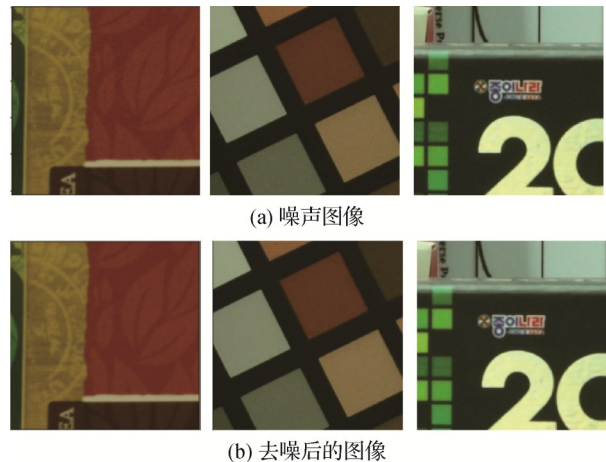


图5 本文算法在Nam数据集上去噪结果

Fig. 5 Denoising results of our algorithm on the Nam dataset ((a) noisy images; (b) denoised images)

3.1.3 PolyU数据集

PolyU(the Hong Kong Polytechnic University)数据集(Xu等,2018a)包含100个噪声图像和平均图像对,图像为JPG文件。表3是本文方法与CBDNet、DIDN(deep iterative down-up convolutional neural network)(Yu等,2019)和AINDNet(Kim等,2020)在PolyU数据集上的对比。图6是本文方法在PolyU数据集上去噪结果。可以看出本文去噪方法在PolyU数据集上表现良好。

表3 不同方法在PolyU数据集上去噪结果对比
Table 3 Comparison of denoising results of different methods on PolyU dataset

方法	PSNR/dB	SSIM
CBDNet	36.93	0.945 7
AINDNet	37.22	0.946 0
DIDN	37.32	0.952 3
本文	37.34	0.980 1

注:加粗字体表示各列最优结果。

3.1.4 DND数据集

DND(darmstadt noise dataset)数据集(Plötz和Roth,2017)是使用第2种方法制作而成,也是目前图像去噪领域比较常用的数据集。表4是本文方法与CDnCNN-B、BM3D、NI(neat image)、NC(noise clinic)(Lebrun等,2015)、DRNet(Li等,2017)和Path-Restore等方法在DND数据集上的比较结果。可以看出,本文方法的图像去噪评价指标都有明显

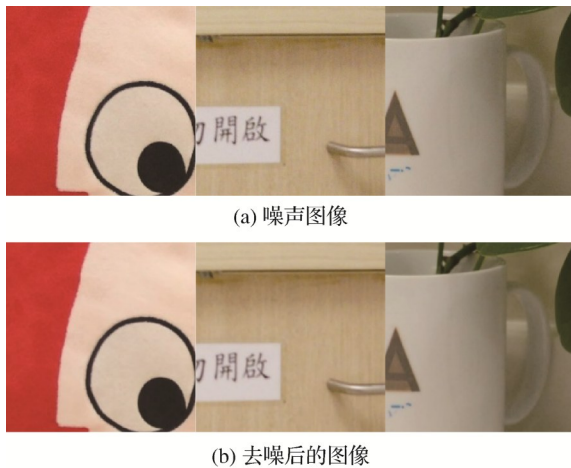


图6 本文算法在PolyU数据集上去噪结果

Fig. 6 Denoising results of our algorithm on the PolyU dataset
((a) noisy images; (b) denoised images)

表4 不同方法在DND数据集上去噪结果对比

Table 4 Comparison of denoising results of different methods on DND dataset

方法	sRGB	
	PSNR/dB	SSIM
CDnCNN-B	32.43	0.790 0
EPLL	33.51	0.824 4
TNRD	33.65	0.830 6
MLP	34.23	0.833 1
BM3D	34.51	0.850 7
FoE	34.62	0.884 5
WNNM	34.67	0.864 6
NI	35.11	0.877 8
NC	35.43	0.884 1
DRNet	35.89	0.935 1
KSVD	36.49	0.897 8
FFDNet+	37.61	0.941 5
TWSC	37.96	0.941 6
CBDNet	38.06	0.942 1
Path-Restore	39.00	0.954 2
本文	39.23	0.949 4
PRIDNet	39.42	0.952 8
NBNet	39.89	0.955 0

注:加粗字体表示各列最优结果。

提高。虽然 PRIDNet 和 NBNet 算法的评价指标更好,但与本文算法的差异并不明显。本文算法模型

较为简单,且由于多尺度模块的设计,本文模型的部署也更为灵活。图7显示了各方法在DND数据集上已去噪的图像,以及它们的PSNR数值(单位dB)。

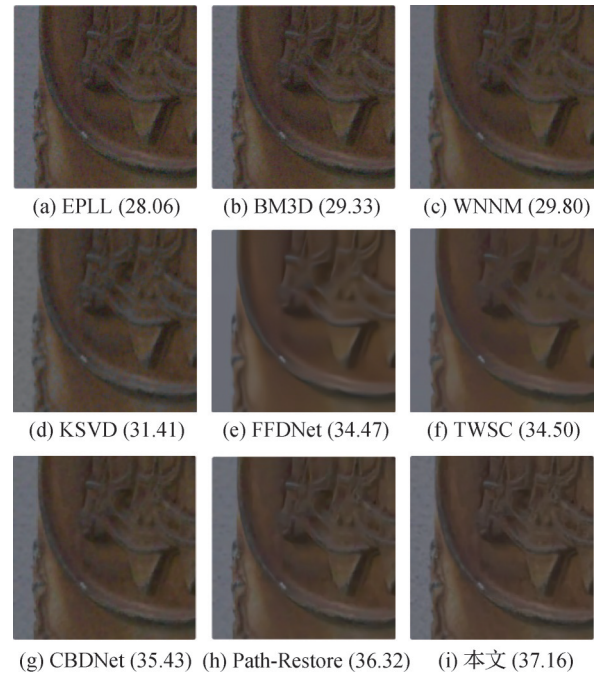


图7 不同方法在DND数据集上去噪结果(dB)

Fig. 7 Denoising results of different methods on the DND dataset (dB) ((a) EPLL (28.06); (b) BM3D (29.33); (c) WNNM (29.80); (d) KSVD (31.41); (e) FFDNet (34.47); (f) TWSC (34.50); (g) CBDNet (35.43); (h) Path-Restore (36.32); (i) ours (37.16))

3.1.5 补充实验

本文算法是一个两阶段的真实图像去噪算法,需要对真实图像进行噪声水平估计,为了验证噪声水平估计对真实图像去噪效果的影响,通过是否使用噪声水平估计,分别对图像进行去噪处理。具体做法是在BSD68(Berkeley segmentation data)数据集上添加不同分布水平的高斯噪声(σ 分别为15, 25, 50),并对其进行降噪,实验结果如表5所示。实验表明,使用噪声水平估计的算法在去除噪声时效果更加明显,说明在去除真实图像噪声时对噪声进行水平估计是必要的。

为了证明加深网络即增加网络参数并不能明显提升所有图像的去噪效果,尤其对于纹理细节相对较少的图像,去噪效果的提升并不明显。在Nam数据集中挑选4幅纹理细节明显不同的图像,通过使用不同数量的多尺度模块设置网络深度,再对挑选

表5 噪声水平估计对去噪结果的影响
Table 5 Influence of noise level estimation on denoising results

噪声水平估计	PSNR/dB		
	$\sigma = 15$	$\sigma = 25$	$\sigma = 50$
无	30.21	29.08	26.26
有	30.26	29.11	26.31

注:加粗字体表示各列最优结果。

出的图像进行去噪处理,去噪结果如图8所示。其中,结果1使用了2个多尺度模块,结果2使用了3个多尺度模块。

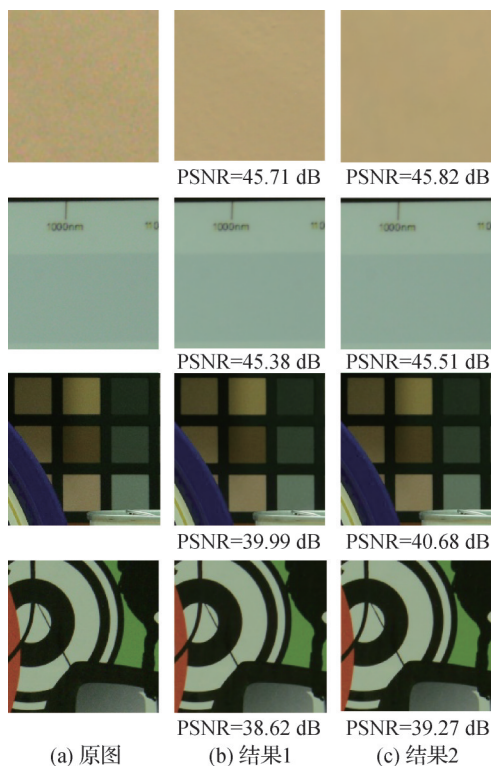


图8 网络参数对去噪效果的影响

Fig. 8 The influence of network parameters on denoising results ((a) original pictures; (b) result 1; (c) result 2)

3.1.6 消融实验

本文去噪网络模型大量使用了长、短跳跃连接,一方面因为残差结构在去除图像噪声上有良好表现;另一方面,使用跳跃连接可以将模型中的深层信息和浅层信息融合起来,减少由于模型层数的增加而导致的信息丢失。本文模型在输入图像与最后的输出之间使用了跳跃连接,这是比较长的跳跃连接,表示为长连接。在多尺度结构中的模块输入和输出部分也使用了跳跃连接,这是比较短的跳跃连接,表

示为短连接。此外,在噪声水平估计阶段使用了通道注意力机制,也会影响实验的结果。为了证明跳跃连接和通道注意力机制的有效性,在SIDD验证集上进行消融实验,结果如表6所示。

表6 消融实验
Table 6 Ablation experiments

长跳跃	短跳跃	SEBlock	水平估计	PSNR/dB
-	-	-	-	51.30
-	-	√	√	51.36
-	√	√	√	51.46
√	√	-	-	51.61
√	√	-	√	51.64
√	√	√	√	51.70

注:加粗字体表示最优结果。“-”表示不使用,“√”表示使用。

3.2 实验细节

由于SIDD数据集原始图像分辨率过高,本文将图像随机切分为 256×256 像素的patch,再对数据进行4 000轮训练,在前1 500轮学习率为 10^{-4} ,之后每隔1 000轮,学习率降低1/10。所有的实验均在NVIDIA GTX 2080ti GPU上使用pytorch框架完成。

4 结论

本文设计了一个用于去除真实图像噪声的深度神经网络模型,网络中的卷积操作大多都是 3×3 卷积,网络训练速度较快。由于残差结构在图像去噪中的优良表现,本文网络也使用丰富的残差结构。此外,本文将膨胀卷积与普通卷积相结合,从不同感受野中得到更为丰富的特征信息。在4个真实图像去噪数据集上的实验结果表明,本文设计的去噪网络相较于其他更深更复杂的网络在去除真实图像噪声上效果明显。

本文设计的两阶段盲去噪网络使用简单的网络结构可以实现真实图像的有效降噪,但与其他较深的去噪网络模型相比,在去噪效果上仍有差距,更深更复杂的网络相较于本文的网络可以获得更为丰富的图像信息。因此,如何在尽量减少图像信息丢失的同时,保证网络结构的简单和灵活性是今后的主要研究方向。

参考文献 (References)

- Abdelhamed A, Lin S and Brown M S. 2018. A high-quality denoising dataset for smartphone cameras//Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA: IEEE: 1692-1700 [DOI: 10.1109/CVPR.2018.00182]
- Aharon M, Elad M and Bruckstein A. 2006. K-SVD: an algorithm for designing overcomplete dictionaries for sparse representation. *IEEE Transactions on Signal Processing*, 54(11): 4311-4322 [DOI: 10.1109/TSP.2006.881199]
- Anwar S and Barnes N. 2019. Real image denoising with feature attention//Proceedings of 2019 IEEE/CVF International Conference on Computer Vision. Seoul, Korea (South): IEEE: 3155-3164 [DOI: 10.1109/ICCV.2019.00325]
- Buades A, Coll B and Morel J M. 2005. A non-local algorithm for image denoising//Proceedings of 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. San Diego, USA: IEEE: 60-65 [DOI: 10.1109/CVPR.2005.38]
- Burger H C, Schuler C J and Harmeling S. 2012. Image denoising: can plain neural networks compete with BM3D?//Proceedings of 2012 IEEE Conference on Computer Vision and Pattern Recognition. Providence, USA: IEEE: 2392-2399 [DOI: 10.1109/CVPR.2012.6247952]
- Chang S G, Yu B and Vetterli M. 2000. Adaptive wavelet thresholding for image denoising and compression. *IEEE Transactions on Image Processing*, 9(9): 1532-1546 [DOI: 10.1109/83.862633]
- Chen Y J and Pock T. 2017. Trainable nonlinear reaction diffusion: a flexible framework for fast and effective image restoration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(6): 1256-1272 [DOI: 10.1109/TPAMI.2016.2596743]
- Cheng S, Wang Y Z, Huang H B, Liu D H, Fan H Q and Liu S C. 2021. NBNet: noise basis learning for image denoising with subspace projection//Proceedings of 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville, USA: IEEE: 4896-4906 [DOI: 10.1109/CVPR46437.2021.00486]
- Dabov K, Foi A, Katkovnik V and Egiazarian K. 2008. Image restoration by sparse 3D transform-domain collaborative filtering//Proceedings of SPIE 6812, Image Processing: Algorithms and Systems VI. San Jose, USA: SPIE: #681207 [DOI: 10.1117/12.766355]
- Gu S H, Zhang L, Zuo W M and Feng X C. 2014. Weighted nuclear norm minimization with application to image denoising//Proceedings of 2014 IEEE Conference on Computer Vision and Pattern Recognition. Columbus, USA: IEEE: 2862-2869 [DOI: 10.1109/CVPR.2014.366]
- Guan M J, Zhu J and Wu Y R. 2020. Research progress of image denoising methods. *Computer Era*, 2: 29-32, 35 (管美静, 珠杰, 吴燕如. 2020. 图像去噪方法研究进展. *计算机时代*, (2): 29-32, 35) [DOI: 10.16644/j.cnki.cn33-1094/tp.2020.02.008]
- Guo S, Yan Z F, Zhang K, Zuo W M and Zhang L. 2019. Toward convolutional blind denoising of real photographs//Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach, USA: IEEE: 1712-1722 [DOI: 10.1109/CVPR.2019.00181]
- He K M, Sun J and Tang X O. 2013. Guided image filtering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(6): 1397-1409 [DOI: 10.1109/TPAMI.2012.213]
- Hu J, Shen L, Albanie S, Sun G and Wu E. 2019. Squeeze-and-excitation networks//Proceedings of 2019 IEEE Conference on Computer Vision and Pattern Recognition. New York, USA: IEEE: 183-192 [DOI: 10.1109/TPAMI.2019.2913372]
- Jain V and Seung H S. 2008. Natural image denoising with convolutional networks//Proceedings of the 21st International Conference on Neural Information Processing Systems. Vancouver British, Canada: Curran Associates Inc.: 769-776
- Kim J, Lee J K and Lee K M. 2016. Accurate image super-resolution using very deep convolutional networks//Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, USA: IEEE: 1646-1654 [DOI: 10.1109/CVPR.2016.182]
- Kim Y, Soh J W, Park G Y and Cho N I. 2020. Transfer learning from synthetic to real-noise denoising with adaptive instance normalization//Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle, USA: IEEE: 3479-3489 [DOI: 10.1109/CVPR42600.2020.00354]
- Lebrun M, Colom M and Morel J M. 2015. The noise clinic: a blind image denoising algorithm. *Image Processing on Line*, 5: 1-54 [DOI: 10.5201/ipol.2015.125]
- Li C Y, Guo J C, Porikli F, Guo C, Fu H Z and Li X. 2017. DR-Net: transmission steered single image dehazing network with weakly supervised refinement [EB/OL]. [2021-12-02]. <https://arxiv.org/pdf/1712.00621.pdf>
- Liu J S. 2021. Research on Real-World Image Denoising Based on Deep Neural Network. Xi'an: Xidian University (刘金山. 2021. 基于深度神经网络的真实图像去噪的研究. 西安: 西安电子科技大学) [DOI: 10.27389/d.cnki.gxadu.2021.002156]
- Liu L P, Qiao L L and Jiang L C. 2021. Overview of image denoising methods. *Journal of Frontiers of Computer Science and Technology*, 15(8): 1418-1431 (刘利平, 乔乐乐, 蒋柳成. 2021. 图像去噪方法概述. *计算机科学与探索*, 15(8): 1418-1431) [DOI: 10.3778/j.issn.1673-9418.2101035]
- Nam S, Hwang Y, Matsushita Y and Kim S J. 2016. A holistic approach to cross-channel image noise modeling and its application to image denoising//Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, USA: IEEE: 1683-1691 [DOI: 10.1109/CVPR.2016.186]
- Plötz T and Roth S. 2017. Benchmarking denoising algorithms with real photographs//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, USA: IEEE: 2750-2759 [DOI: 10.1109/CVPR.2017.294]
- Roth S and Black M J. 2005. Fields of experts: a framework for learning image priors//Proceedings of 2005 IEEE Computer Society Confer-

- ence on Computer Vision and Pattern Recognition. San Diego, USA: IEEE; 860-867 [DOI: 10.1109/CVPR.2005.160]
- Tai Y, Yang J, Liu X M and Xu C Y. 2017. MemNet: a persistent memory network for image restoration//Proceedings of 2017 IEEE International Conference on Computer Vision. Venice, Italy: IEEE; 4549-4557 [DOI: 10.1109/ICCV.2017.486]
- Tian C W, Xu Y and Zuo W M. 2020. Image denoising using deep CNN with batch renormalization. *Neural Networks*, 121: 461-473 [DOI: 10.1016/j.neunet.2019.08.022]
- Tomasi C and Manduchi R. 1998. Bilateral filtering for gray and color images//Proceedings of the 6th International Conference on Computer Vision. Bombay, India: IEEE; 839-846 [DOI: 10.1109/ICCV.1998.710815]
- Wang H C. 2021. Research on Real Image Denoising Algorithm Based on Deep Learning. Chengdu: University of Electronic Science and Technology of China (王瀚琛. 2021. 基于深度学习的真实图像去噪算法研究. 成都: 电子科技大学) [DOI: 10.27005/d.cnki.gdzku.2021.004590]
- Wang Z, Bovik A C, Sheikh H R and Simoncelli E P. 2004. Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13 (4): 600-612 [DOI: 10.1109/TIP.2003.819861]
- Wu C Z, Chen X, Ji D and Zhan S. 2018. Image denoising via residual network based on perceptual loss. *Journal of Image and Graphics*, 23(10): 1483-1491 (吴从中, 陈曦, 季栋, 詹曙. 2018. 结合深度残差学习和感知损失的图像去噪. 中国图象图形学报, 23(10): 1483-1491) [DOI: 10.11834/jig.180069]
- Xie S M, Deng L, Xie Z M and Li H P. 2020. Denoising algorithm based on convolutional neural network in real scene. *Computer and Telecommunication*, (11): 39-43 (谢舒闽, 邓磊, 谢中梅, 李海平. 2020. 真实场景中基于卷积神经网络的去噪算法. 电脑与电信, (11): 39-43) [DOI: 10.15966/j.cnki.dnydx.2020.11.012]
- Xu J, Li H, Liang Z T, Zhang D and Zhang L. 2018a. Real-world noisy image denoising: a new benchmark [EB/OL]. [2021-11-04]. <https://arxiv.org/pdf/1804.02603.pdf>
- Xu J, Zhang L and Zhang D. 2018b. A trilateral weighted sparse coding scheme for real-world image denoising//Proceedings of the 15th European Conference on Computer Vision. Munich, Germany: Springer [DOI: 10.1007/978-3-030-01237-3_2]
- Xu S P, Liu T Y, Lin Z Y, Zhang G Z and Li C X. 2019. Main bottlenecks and research prospects of the deep convolutional neural network-based denoising model. *Journal of Image and Graphics*, 24(8): 1207-1214 (徐少平, 刘婷云, 林珍玉, 张贵珍, 李崇禧. 2019. 深度卷积神经网络降噪模型的技术瓶颈与研究展望. 中国图象图形学报, 24(8): 1207-1214) [DOI: 10.11834/jig.190165]
- Yu H and Zhen T. 2020. Research on classical image denoising algorithms. *China Computer and Communication*, 32(14): 66-67 (于虹, 甄彤. 2020. 图像去噪经典算法研究. 信息与电脑(理论版), 32(14): 66-67) [DOI: 10.3969/j.issn.1003-9767.2020.14.021]
- Yu K, Wang X T, Dong C, Tang X O and Loy C C. 2022. Path-restore: learning network path selection for image restoration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44 (10): 7078-7092 [DOI: 10.1109/TPAMI.2021.3096255]
- Yu S, Park B and Jeong J. 2019. Deep iterative down-up CNN for image denoising//Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. Long Beach, USA: IEEE; 2095-2103 [DOI: 10.1109/CVPRW.2019.00262]
- Zamir S W, Arora A, Khan S, Hayat M, Khan S, Yang M H and Shao L. 2020a. CycleISP: real image restoration via improved data synthesis//Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle, USA: IEEE [DOI: 10.1109/CVPR42600.2020.00277]
- Zamir S W, Arora A, Khan S, Hayat M, Khan F S, Yang M H and Shao L. 2020b. Learning enriched features for real image restoration and enhancement//Proceedings of the 16th European Conference on Computer Vision. Glasgow, UK: Springer: 492-511 [DOI: 10.1007/978-3-030-58595-2_30]
- Zhang K, Zuo W M, Chen Y J, Meng D Y and Zhang L. 2017. Beyond a gaussian denoiser: residual learning of deep CNN for image denoising. *IEEE Transactions on Image Processing*, 26(7): 3142-3155 [DOI: 10.1109/TIP.2017.2662206]
- Zhang K, Zuo W M and Zhang L. 2018. FFDNet: toward a fast and flexible solution for CNN-based image denoising. *IEEE Transactions on Image Processing*, 27 (9): 4608-4622 [DOI: 10.1109/TIP.2018.2839891]
- Zhao Y Y, Jiang Z Q, Men A D and Ju G D. 2019. Pyramid real image denoising network//Proceedings of 2019 IEEE Visual Communications and Image Processing. Sydney, Australia: IEEE; 1-4 [DOI: 10.1109/VICIP47243.2019.8965754]
- Zoran D and Weiss Y. 2011. From learning models of natural image patches to whole image restoration//Proceedings of 2011 International Conference on Computer Vision. Barcelona, Spain: IEEE; 479-486 [DOI: 10.1109/ICCV.2011.6126278]

作者简介

丁岳皓,男,硕士研究生,主要研究方向为深度学习图像处理。E-mail: d1yh152003@163.com

吴昊,通信作者,男,讲师,主要研究方向为图像处理、计算机视觉。Email: haowu1982@vip.163.com

孔凤玲,女,硕士研究生,主要研究方向为图像处理。

E-mail: 2378015890@qq.com

徐丹,女,教授,主要研究方向为模式识别、计算机视觉。

E-mail: danxu@ynu.edu.cn

袁国武,男,副教授,主要研究方向为数字图像处理。

E-mail: gwuyuan@ynu.edu.cn