

基于非对称卷积神经网络的图像去噪

甘建旺¹, 沙芸^{1*}, 张国英²

¹北京石油化工学院信息工程学院, 北京 102617;

²中国矿业大学(北京)机电与信息工程学院, 北京 100083

摘要 由于图像的像素越来越小,数字成像传感器输出的信号对光子噪声的敏感性越来越强,使光子噪声成为数字图像传感器噪声的主要来源。鉴于此,提出一种基于非对称卷积神经网络的图像去噪算法。为了提高模型的泛化能力,将网络框架分为噪声评估网络和去噪网络两部分。为了减少编码器与解码器中网络特征映射之间的语义差距,对去噪网络中的跳跃连接进行改进,使特征在语义上更相似,以便于任务的优化处理。从定性和定量方面进行对比实验,实验结果表明,改进后的网络模型的去噪性能更佳。

关键词 图像处理; 非对称卷积神经网络; 去噪; 跳跃连接; 光子噪声

中图分类号 TP391.4

文献标志码 A

doi: 10.3788/LOP57.221018

Image Denoising Based on Asymmetric Convolutional Neural Networks

Gan Jianwang¹, Sha Yun^{1*}, Zhang Guoying²

¹School of Information Engineering, Beijing Institute of Petrochemical Technology, Beijing 102617, China;

²School of Mechanical Electronic & Information Engineering, China University of Mining and Technology (Beijing), Beijing 100083, China

Abstract Owing to the continuing decrement in the pixels of the images, the signal output of the digital imaging sensor is increasingly sensitive to photon noise, making the photon noise the main source of noise in the digital image sensor. To address this issue, an image denoising algorithm based on asymmetric convolutional neural networks is proposed herein. To enhance the generalization ability of the model, the network framework is divided into two parts: noise evaluation network and denoising network. To reduce the semantic gap between the network feature mapping in the encoder and the decoder, the skip connection in the denoising network is improved to make the features more similar in semantics to facilitate task optimization. From the qualitative and quantitative aspects of comparative experiments, the experimental results show that the proposed network model exhibits better denoising performance.

Key words image processing; asymmetric convolutional neural network; denoising; skip connection; photon noise

OCIS codes 100.2000; 150.1135; 150.4065

1 引言

图像去噪是计算机视觉领域中非常经典的问题之一,其是从包含噪声的图像中还原潜在的且相对干净的图像。虽然传感器的电和热特性已有显著改善^[1],但目前每个像素的尺寸越来越小。数字图像传感器输出的信号对光子噪声的敏感性越来越强,因此光子噪声是数字图像传感器中噪声的主要来源^[2]。图像中多余噪声的去除可以带来更佳的

视觉效果,提高图像质量,处理后的图像会影响后续图像的分类、分割和识别等^[3]。

图像去噪的方法分为基于传统算法的图像去噪和基于深度学习的图像去噪两类。基于传统算法的图像去噪的处理方法:一类是专注于噪声图像的局部平滑,在去噪的基础上尽可能地保留图像边缘细节,使用的滤波方法有双边滤波^[4]和高斯滤波^[5]等;另一类是利用图像先验模型对图像进行去噪,使用的算法有非局部算法^[6]和 BM3D(Block-

收稿日期: 2020-04-02; 修回日期: 2020-04-20; 录用日期: 2020-04-21

* E-mail: shayun@bipt.edu.cn

Matching and 3D filtering)算法^[7]等。其中 BM3D 算法是传统算法中经典的去噪算法之一,采用该算法将图像分为一定大小的块,将具有相似特征的块合并成三维数组,采用联合滤波的方法对其进行处理,之后通过逆变换获得去噪后的图像。基于深度学习的图像去噪方法主要分为基于多层感知器(MLP)^[8]和全卷积网络(FCN)^[9]。基于 MLP 的图像去噪方法的处理过程:首先将原始图像拆分为大小相同的图像块,然后对每个图像块进行去噪,最后对去噪后的图像块进行拼接,对重叠部分的图像块进行聚合,从而得到去噪后的图像^[8]。采用该方法处理图像耗时较长且容易造成网络模型的过拟合。基于 FCN 的去噪方法有去噪卷积神经网络(DnCNN)^[10]和卷积盲去噪网络(CBDNet)^[11]。DnCNN 为 FCN^[10],其噪声类型为加性高斯白噪声,而数据集中的噪声具有单一性,则训练的模型对于数字成像传感器产生的噪声,去噪效果较差。CBDNet 模型^[11]分为噪声评估模型和去噪模型两部分,其训练数据集使用泊松高斯(Poissonian-Gaussian)模型^[2]的模拟数据集和真实噪声数据集,该模型对相机内部噪声的去噪效果较好。

FCN 去噪的成功依赖于训练数据集与真实噪声之间是否匹配,因此本文采用 Poissonian-Gaussian 模型模拟真实噪声的形成过程,同时使用公共数据集 RENOIR^[12]作为模型的训练数据集,此外为了提高模型的泛化能力,将模型分为噪声评估网络模型和 U 型去噪网络模型两部分。

2 所提方法

2.1 图像噪声模型

实验使用的训练数据集可分为模拟数据集和公共数据集 RENOIR^[12]两部分。其中模拟数据集

是使用 Poissonian-Gaussian 模型^[2]来模拟图像噪声以及光子传感器产生的噪声,剩余的稳态扰动使用高斯模型来模拟,噪声模拟的具体方法如下。

与信号相关的噪声模型可以表示为

$$F(x) = y(x) \sigma[y(x)] \xi(x), \quad (1)$$

式中: x 表示像素的位置; $F(x)$ 表示待观测信号; $y(x)$ 表示原始信号; $\xi(x)$ 表示期望 $E[\xi(x)] = 0$, 方差 $\text{Var}[\xi(x)] = 1$ 的独立随机噪声; $\sigma[y(x)]$ 表示总体分量的标准差。

假设 $\sigma[y(x)]\xi(x)$ 由泊松分布 η_p 和高斯分布 η_g 组成,可以表示为

$$\sigma[y(x)]\xi(x) = \eta_p[y(x)] + \eta_g(x). \quad (2)$$

由分布函数可以得到

$$\eta_g \sim N(0, b), \quad (3)$$

$$\lambda\{y(x) + \eta_p[y(x)]\} \sim \eta_p[\lambda y(x)], \quad (4)$$

式中: b, λ 表示参数,此时 $b \geq 0, \lambda > 0$; N 表示正态分布。根据泊松分布的性质,可以得到 $\text{Var}\{\eta_p[y(x)]\} = ay(x)$, 其中 $a = \lambda^{-1}$, 因此(1)式的 $\sigma[y(x)]$ 可以表示为

$$\sigma^2[y(x)] = ay(x) + b. \quad (5)$$

泊松高斯模型服从异方差高斯分布,可以表示为

$$\sigma[y(x)]\xi(x) \sim N\{0, \sigma^2[y(x)]\}. \quad (6)$$

图像信号传感器输出的信号经过图像处理器(ISP)处理后,使得噪声在空间与色彩上相关,因此对噪声模型进行改进,可以表示为

$$Z(x) = f\{D\{y(x) + n[y(x)]\}\}, \quad (7)$$

式中: $f(\cdot)$ 表示相机的响应函数(CRF)^[13]; $D(\cdot)$ 表示去噪函数^[14],其插值会涉及不同通道和空间位置的像素; $Z(x)$ 表示合成的噪声图像,如图 1 所示。其中每组图像中左侧为原始图像,右侧为合成的噪声图像。

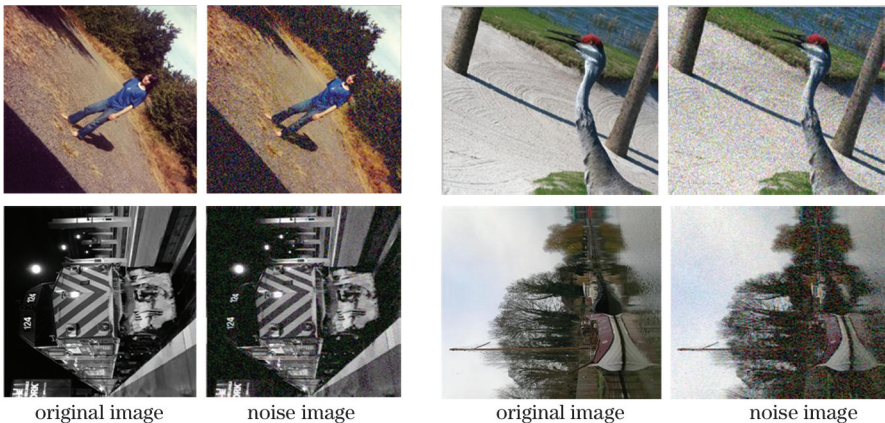


图 1 部分合成的噪声图像。

Fig. 1 Partially synthesized noise images.

2.2 网络模型

实验将 UNet++^[15] 中的跳跃连接应用到 CBDNet^[11] 的 U 型去噪子网中, 加入的跳跃连接旨在减少编码器和解码器中子网特征映射之间的语义差距。当特征映射在语义上相似时, 优化任务将更容易被处理^[15]。

CBDNet^[11] 模型如图 2 所示, 该网络模型由噪

声评估网络模型和去噪网络模型两部分组成, 目的是为了提高模型的泛化能力。噪声评估网络架构中使用 5 个卷积层, 每个卷积层中的卷积核大小为 3×3 , 步长为 1, 输出通道数为 32, 激活函数为线性整流函数; 去噪网络架构为 16 层 U 型架构, 其中跳跃连接是将编码器和解码器中的特征映射进行连接。

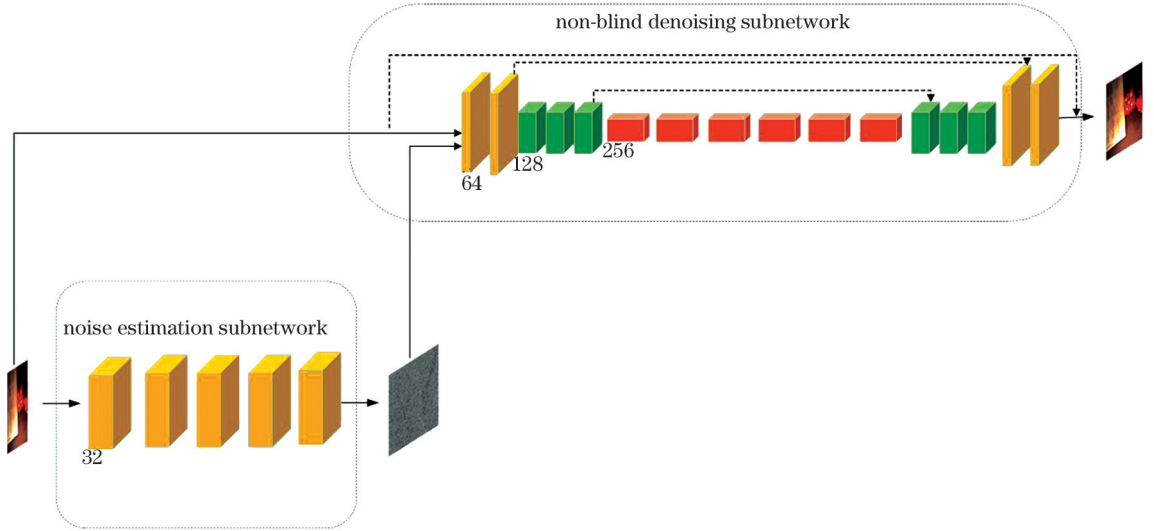


图 2 CBDNet 模型

Fig. 2 CBDNet model

对图 2 中的 U 型网络架构进行简化, 结果如图 3 所示。其中 $X_{i,j}$ 表示 CBDNet 去噪子网中具有相同输出通道数的特征映射层, i 表示沿着编码器的索引, j 表示沿着跳跃连接的索引, 如 $X_{1,0}$ 表示 CBDNet 去噪子网输出通道大小为 64 的卷积层, 对 $X_{1,0}$ 进行下采样后得到输出通道数为 128 的 $X_{2,0}$ 。

设计跳跃连接, 减少编码器与解码器中子网特征映射之间的语义差距, 使得特征映射在语义上相似, 优化任务变得简单^[15]。 $X_{i,j}$ 的计算表达式为

$$X_{i,j} = \begin{cases} H(X_{i-1,j}), & j = 0 \\ H[X_{i,k}, U(X_{i+1,j-1})], & j > 0 \end{cases} \quad (8)$$

式中: $H(\cdot)$ 表示卷积运算; $U(\cdot)$ 表示上采样。

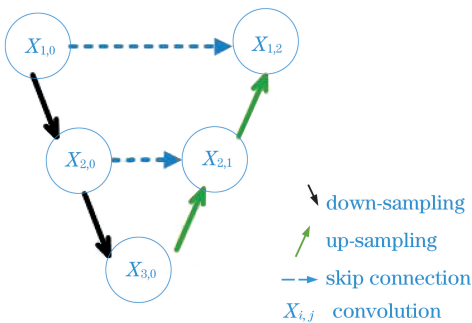


图 3 CBDNet 去噪子网架构

Fig. 3 CBDNet denoising subnet architecture

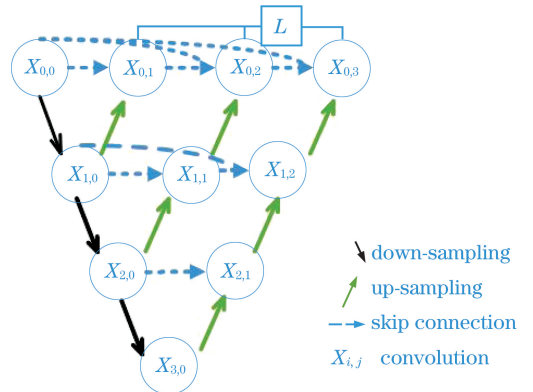


图 4 UNet++ 架构

Fig. 4 UNet++ architecture

UNet++^[15] 模型如图 4 所示, 该网络模型由编码器和解码器嵌套组成, 本质上为自编码器网络, 其中 L 表示总的损失函数。UNet++^[15] 模型中, 重新

将 UNet++ 模型中重新设计的跳跃连接应用到 CBDNet 的去噪子网中, 即可得到实验使用的模型, 如图 5 所示。

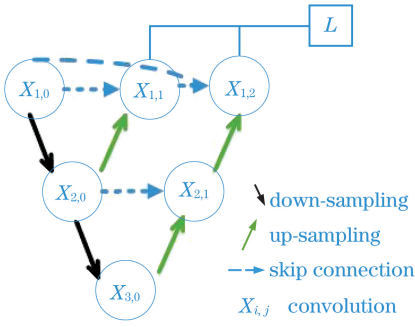


图5 改进的网络模型

Fig. 5 Improved network model

2.3 损失函数

损失函数采用 CBDNet^[11] 模型中的损失函数, 其中损失函数分为噪声评估模型损失函数和去噪模型损失函数两部分。噪声评估模型损失函数可以表示为

$$L_1 = \sum_x |\alpha - I_e| \times [\hat{v}(o_x) - v(o_x)]^2, \quad (9)$$

式中: α 表示参数, 此时 $0 < \alpha < 0.5$; $v(o_x)$ 和 $\hat{v}(o_x)$ 分别表示在位置 x 处像素 o_x 的真实噪声和估计噪声; $\hat{v}(o_x) - v(o_x) I_e$ 表示当 $e < 0$ 时 $I_e = 1$, 其他条件为零, 则更大的惩罚项将作用于参数的调节。此外为了约束 $\hat{v}(o_x)$ 过度平滑, 将在 L_1 后加上正则项, 可表示为

$$L_r = \|\nabla_h \hat{v}(o_x)\|_2^2 + \|\nabla_v \hat{v}(o_x)\|_2^2, \quad (10)$$

式中: ∇_h 和 ∇_v 分别表示沿着水平和垂直方向上的梯度算子。

去噪模型的损失函数可以表示为

$$L_2 = \|\hat{\mathbf{x}} - \mathbf{x}\|_2^2, \quad (11)$$

式中: $\hat{\mathbf{x}}$ 表示预测图像; \mathbf{x} 表示真实无噪图像。

总的损失函数可表示为

$$L = L_2 + \lambda_1 L_1 + \lambda_r L_r, \quad (12)$$

式中: λ_1 和 λ_r 表示噪声评估模型的权衡参数。

3 实验结果与分析

3.1 数据集和评价指标

3.1.1 数据集

在经典的 RENOIR 数据集^[12] 和 VOC 数据集中, 使用 Poissonian-Gaussian 模型生成的数据作为训练数据集, 将公共数据集 PloyU^[16] 作为测试数据集, 实验从测试集中任选 10 张图像进行测试。

3.1.2 评价标准

评价标准有客观评价和主观评价。客观评价标准是采用峰值信噪比 (PSNR)^[17] 和结构相似性

(SSIM)^[17] 作为定量的评价标准, 表达式为

$$P_{\text{PSNR}} = 10 \cdot \lg \left[\frac{(2^n - 1)^2}{E_{\text{MSE}}} \right], \quad (13)$$

$$S_{\text{SSIM}}(x, y) = \frac{(2\epsilon_1\epsilon_2 + d_1)(2\gamma_{12} + d_2)}{(\epsilon_1^2 + \epsilon_2^2 + d_1)(\gamma_1^2 + \gamma_2^2 + d_2)}, \quad (14)$$

式中: E_{MSE} 表示原图像与去噪图像的均方误差 (MSE); ϵ_1 和 ϵ_2 表示两幅图像的均值; γ_1^2 和 γ_2^2 表示两幅图像的方差; γ_{12} 表示图像 1 和 2 的协方差; $d_1 = (k_1 o_x)^2$, $d_2 = (k_2 o_x)^2$; $k_1 = 0.01$, $k_2 = 0.03$; n 为常数。计算得到, PSNR 值为 0~100, SSIM 值为 0~1, 两幅图像越相似, PSNR 和 SSIM 值越大, 即去噪效果越好。

主观评价标准是在人的主观视觉中, 根据自然图像的先验知识, 将图像与已有先验知识的场景进行对比。

3.2 实验环境与参数选择

网络的训练和测试都是在 Pytorch 工具上进行的。实验硬件设备为 Intel 2.20 GHz CPU, NVIDIA GeForce GTX 1080 Ti。采用 Adam 梯度下降算法对回归目标进行优化训练, 初始学习率为 0.0001, 训练 201 次迭代后终止训练。

3.3 结果分析

为了验证网络模型的有效性, 对网络模型进行实验对比。首先从公共数据集 PloyU 中选取 10 张图像, 将改进的 CBDNet 与标准 CBDNet 得到的结果进行对比, PSNR 和 SSIM 的结果如表 1 和表 2 所示。然后将公共数据集 PolyU^[16] 中的 40 张图像作为测试数据集, 分别使用 BM3D、CBDNet 及改进 CBDNet 方法得到的结果进行对比, 实验结果如表 3 所示, 从整体的测试结果来看, 改进 CBDNet 的 PSNR 值和 SSIM 值大于 BM3D^[7] 和 CBDNet^[11]。

表 1 10 张图像的 PSNR

Table 1 PSNR of 10 images

No.	CBDNet	Improved CBDNet
1	35.769720	35.880955
2	31.461794	30.128640
3	36.202120	36.103760
4	33.205257	33.806810
5	35.034992	34.901047
6	38.669327	40.513270
7	37.770805	37.781620
8	34.800518	34.688103
9	34.610195	34.964275
10	34.232903	34.225975
Mean value	35.175762	35.299446

表2 10张图像的结构相似度
Table 2 SSIM of 10 images

No.	CBDNet	Improved CBDNet
1	0.94174070	0.94231653
2	0.96375257	0.96220344
3	0.97295123	0.97360283
4	0.97548705	0.97699260
5	0.97754630	0.97723440
6	0.97962856	0.97987250
7	0.98251104	0.98210240
8	0.98582670	0.98583820
9	0.96994860	0.97106440
10	0.97935563	0.97972226
Mean value	0.97287480	0.97309494

表3 PSNR与SSIM的均值
Table 3 Mean of PSNR and SSIM

Type	BM3D	CBDNet	Improved CBDNet
PSNR	35.00	35.200	35.400
SSIM	0.86	0.972	0.973

从PolyU数据集^[16]中随机选取三张样本图像,去噪结果如图6~8所示,其中图中数据为算法处理后的PSNR值。从图6~8可以看到,改进的CBDNet方法的视觉效果更好。

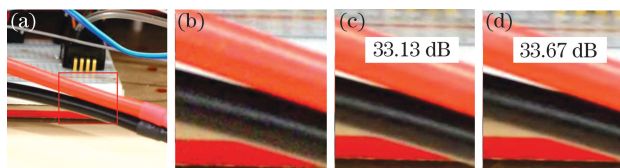


图6 图像1的去噪结果。(a)原始图像;(b)噪声图像;
(c) CBDNet;(d)改进的CBDNet

Fig. 6 Denoising results of image 1. (a) Original image;
(b) noise image; (c) CBDNet; (d) improved
CBDNet

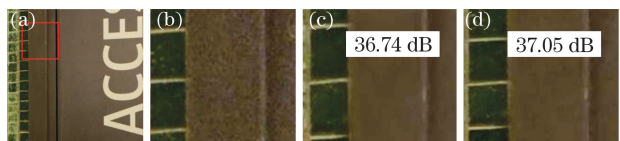


图7 图像2的去噪结果。(a)原始图像;(b)噪声图像;
(c) CBDNet;(d)改进的CBDNet

Fig. 7 Denoising results of image 2. (a) Original image;
(b) noise image; (c) CBDNet; (d) improved
CBDNet

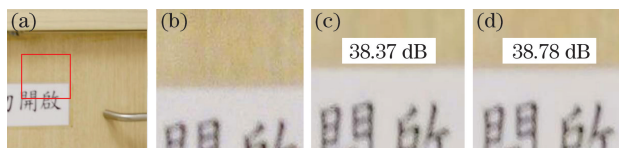


图8 图像3的去噪结果。(a)原始图像;(b)噪声图像;
(c) CBDNet;(d)改进的CBDNet

Fig. 8 Denoising results of image 3. (a) Original image;
(b) noise image; (c) CBDNet; (d) improved
CBDNet

4 结 论

基于CBDNet架构^[11]对其去噪部分的U型网络架构进行改进,将UNet++^[15]中的跳跃连接应用到CBDNet中,旨在减少编码器和解码器中子网特征映射之间的语义差距;去噪后的图像可以解决边缘细节的损失、边界伪影和清晰度变差的问题。实验结果表明,对CBDNet模型的改进能够提高PSNR和SSIM,这为后续的目标检测、识别和跟踪等应用提供很好的图像样本数据集,并具有很好的实用性。

参 考 文 献

- [1] Rhodes H, Agranov G, Hong C, et al. CMOS imager technology shrinks and image performance [C]//2004 IEEE Workshop on Microelectronics and Electron Devices, April 16-16, 2004, Boise, ID, USA. New York: IEEE, 2004: 7-18.
- [2] Foi A, Trimeche M, Katkovnik V, et al. Practical Poissonian-Gaussian noise modeling and fitting for single-image raw-data [J]. IEEE Transactions on Image Processing, 2008, 17(10): 1737-1754.
- [3] Chen X. Image denoising based on deep convolutional neural networks [D]. Hefei: Hefei University of Technology, 2019.
陈曦. 基于深度卷积神经网络的图像去噪[D]. 合肥: 合肥工业大学, 2019.
- [4] Zhang B Y, Allebach J P. Adaptive bilateral filter for sharpness enhancement and noise removal[J]. IEEE Transactions on Image Processing, 2008, 17(5): 664-678.
- [5] Weiss B. Fast median and bilateral filtering [J]. ACM Transactions on Graphics, 2006, 25(3): 519-526.
- [6] 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.

- [7] 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.
- [8] Burger H C, Schuler C J, Harmeling S. Image denoising with multi-layer perceptrons, part 1: comparison with existing algorithms and with bounds [EB/OL]. (2012-11-09) [2020-04-01]. <https://arxiv.org/abs/1211.1544>.
- [9] Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation[C]//2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 7-12, 2015, Boston, MA, USA. New York: IEEE, 2015: 3431-3440.
- [10] 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.
- [11] Guo S, Yan Z F, Zhang K, et al. Toward convolutional blind denoising of real photographs [C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 15-20, 2019, Long Beach, CA, USA. New York: IEEE, 2019: 1712-1722.
- [12] Anaya J, Barbu A. RENOIR - A dataset for real low-light image noise reduction[J]. Journal of Visual Communication and Image Representation, 2018, 51: 144-154.
- [13] Grossberg M D, Nayar S K. Modeling the space of camera response functions[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2004, 26(10): 1272-1282.
- [14] Malvar H S, He L W, Cutler R. High-quality linear interpolation for demosaicing of Bayer-patterned color images[C]//2004 IEEE International Conference on Acoustics, Speech, and Signal Processing, May 17-21, 2004, Montreal, Que., Canada. New York: IEEE, 2004: 8038960.
- [15] Zhou Z W, Rahman Siddiquee M M, Tajbakhsh N, et al. UNet++: a nested U-Net architecture for medical image segmentation [M]//Stoyanov D, Taylor Z, Carneiro G, et al. Deep learning in medical image analysis and multimodal learning for clinical decision support. Lecture notes in computer science. Cham: Springer, 2018, 11045: 3-11.
- [16] Xu J, Li H, Liang Z T, et al. Real-world noisy image denoising: a new benchmark[EB/OL]. (2018-04-07) [2020-04-01]. <https://arxiv.org/abs/1804.02603>.
- [17] 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.