

文章编号: 1673-3193(2024)05-0608-06

极弱光环境下噪声图像的生成

秦嘉豪¹, 秦品乐¹, 柴锐¹, 陈作钧¹, 高艺鹏², 王宝³

(1. 中北大学 计算机科学与技术学院, 山西 太原 030051; 2. 山西和信基业科技股份有限公司, 山西 太原 030006;
3. 山西焦煤华晋焦煤有限责任公司沙曲二号煤矿, 山西 吕梁 033300)

摘要: 针对弱光环境中的噪声复杂, 建立噪声模型十分困难的问题, 本文提出一种基于扩散模型的噪声图像生成方法。首先, 通过扩散模型的前向过程中的公式向清晰图像中加入多种类型的噪声, 然后, 将其与清晰的条件图像一同输入到网络中, 最后, 通过冷扩散模型的采样算法循环迭代的生成噪声图像, 从而建立更加真实的噪声模型。在弱光数据集上的实验结果表明, 与其他算法相比, 本文算法生成的噪声图像在客观指标上的KL散度值为0.068, 比现有最好的方法降低了0.001, 主观上质量更高, 与弱光环境中的图像更接近, 建立的噪声模型更加准确。本文方法可以在弱光环境下建立质量较高的噪声模型, 为弱光环境下的噪声建模提供了新的思路, 且将扩散模型应用于弱光噪声建模领域。

关键词: 微光成像; 扩散模型; 噪声生成; 图像处理; 神经网络; 深度学习

中图分类号: TP391.41 **文献标识码:** A **doi:** 10.3969/j.issn.1673-3193.2024.05.007

引用格式: 秦嘉豪, 秦品乐, 柴锐, 等. 极弱光环境下噪声图像的生成[J]. 中北大学学报(自然科学版), 2024, 45(5): 608-613.

QIN Jiahao, QIN Pinle, CHAI Rui, et al. Generation of noisy images in extremely low-light environments[J]. Journal of North University of China(Natural Science Edition), 2024, 45(5): 608-613.

Generation of Noisy Images in Extremely Low-Light Environments

QIN Jiahao¹, QIN Pinle¹, CHAI Rui¹, CHEN Zuojun¹, GAO Yipeng², Wang Bao³

(1. School of Computer Science and Technology, North University of China, Taiyuan 030051, China;
2. Shanxi Hexin Jiye Technology Co., Ltd., Taiyuan 030006, China;
3. Shanxi Coking Coal Huajin Coking Coal Co., Ltd. Shaqu No. 2 Coal Mine, Lüliang 033300, China)

Abstract: Due to the complexity of noise in low light environments, it is very difficult to establish noise models. In this regard, this paper proposed a noise image generation method based on diffusion model. Firstly, this article introduced various types of noise into clean images using the forward process of diffusion models based on prior knowledge. Then, a conditional diffusion model was used to input it together with the clean image into the network. Finally, the reverse process of the cold diffusion model was used to iteratively generate noisy images. The experimental results on a low light dataset show that compared with other algorithms, the noise images generated by our algorithm have an objective KL divergence value of 0.068, which is 0.001 lower than the existing best methods. The subjective quality is higher, and it is closest to images in low light environments. The established noise model is the most accurate. This method successfully established a high-quality noise model in low light environments, which provided a

收稿日期: 2023-12-27

基金项目: 山西省科技重大专项计划(202101010101018)

作者简介: 秦嘉豪(2000-), 男, 硕士生, 主要从事低水平视觉计算的研究。

通信作者: 秦品乐(1978-), 男, 教授, 博士, 主要从事机器视觉、大数据的研究。E-mail: qpl@nuc.edu.cn.

new approach for noise modeling in low light environments, and the diffusion model was used to the field of low light noise modeling.

Key words: low light imaging; diffusion model; noisy image generation; image process; neural network; deep learning

0 引言

弱光成像技术在国防军事、天文学等多个领域有广泛的应用需求^[1]。然而,由于非常有限的光子数和不可避免的噪声,使得夜间和弱光源成像变得非常困难^[2],一般的解决方案是通过扩大光圈来收集更多的光子数,但是这种做法只是对噪声进行了削减,并不能对噪声进行去除。

弱光环境下所成图像的噪声包括散粒噪声、读出噪声、暗电流噪声、量化噪声等。深度生成模型在微光去噪方面具有优良的性能,如SID^[3]、SIDD^[4]等,还有很多经典去噪算法,如BM3D^[5]、VBM4D^[6]等。另外,ELD方法^[7]从传感器的物理成像角度出发,对微光成像过程中每一步所产生的噪声进行了估计,同时,ELD方法将成像过程分为光子、电子、电压和数字信号四个阶段,较为合理地将微光环境下产生的噪声表示出来。Monakhova等^[8]在ELD的基础上,重新设计了噪声种类,并通过Wasserstain GAN^[9]网络学习了每一种类型的噪声参数,从而得到了质量较高的噪声图。

本文在Monakhova等^[8]工作的基础上提出了一种新的噪声建模的方法LIGDM(Low-light Image Generation Diffusion Model)。LIGDM通过学习一系列细化步骤将复杂的噪声分布转化为经验数据分布,类似Langevin动力学,它是用于训练噪声模型,迭代的输出噪声,本文参考SR3^[10]对U-Net网络进行了简单的修改,使其能够适应有条件的图像生成,然后使用冷扩散模型的采样方法,迭代输出得到最终的噪声图。

1 相关工作

1.1 图像去噪与噪声生成

经典的去噪方法通常依赖于图像的先验知识,如自相似度^[11]、平滑度^[12]等。同时,有EMCCD的视频降噪算法^[13]的性能较为优异。与预先设置图像的经典方法不同,基于深度学习的方法^[14]是通过学习图像数据的分布得到图像的先

验,其在图像质量方面比经典方法有显著的提升,但是由于缺少成对的数据集,基于深度学习的方法的可学习性往往较差。

为了解决复杂的数据分布带来的瓶颈问题,已经有学者进行了提高弱光下和夜间噪声建模真实感的研究。这类研究从传感器的自身特性出发,考虑了传感器本身带来的噪声(包括散粒噪声和读出噪声等),对弱光下可能会产生的噪声进行精准了建模。例如:泊松-高斯噪声^[15]、ELD^[7]。弱光摄影中的一个常用方法是将多幅图像进行合并降噪,即Brust Denoising^[16],但这类方法通常需要大量的成对数据集进行训练,这在实际场景中是十分困难的。

1.2 扩散模型

目前,扩散模型^[17]已经成为建模的强大工具。扩散模型有很多种类,但是基本都是围绕随机噪声去除的概念建立的。Ho等^[18]提出了条件扩散模型,在训练过程中引入了条件 y 来引导模型的生成。Bansal等^[19]认为,扩散模型的前向过程不局限于高斯噪声,它可以使用任意形式的噪声替换,甚至无噪声反演任意图像,即冷扩散。

2 噪声图像生成方法

本文的噪声图像生成算法整体结构如图1所示。

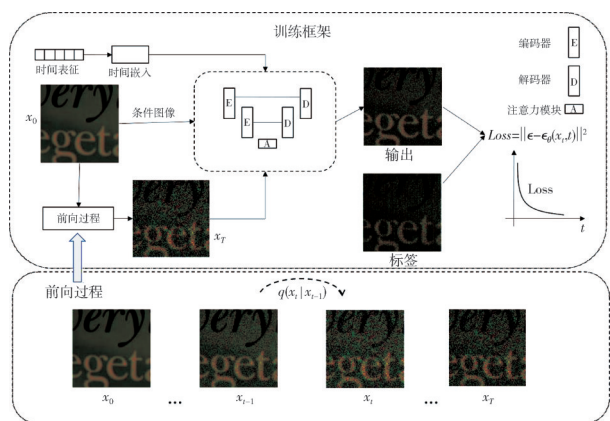


图1 算法整体结构

Fig. 1 The overall structure of the algorithm

2.1 噪声特性

数字图像传感器所产生的噪声图 D 可以用简单的线性模型表示为^[7]

$$D = KI + N, \quad (1)$$

式中: I 为光电子数; K 为模拟增益和数字增益的总和; N 为所有噪声的总和。相机的成像过程可以分为光子、电子、电压和数字信号四个阶段。在整个过程中, 每一阶段都会产生不同的噪声, 从而影响整个图像的质量。在光照良好的环境中, 由于传感器收集的光子数较多, 噪声对图像的影响较小, 主要是散粒噪声, 因此可以将其建模为异方差高斯噪声。然而, 在光线较差的环境中, 传感器所能接收到的光子数较少, 噪声在整个图像中所占的比例较大, 之前的噪声模型不能够很好地描述弱光环境中复杂的噪声特性, ELD^[7] 将弱光环境中的噪声分解为散粒噪声、读出噪声、行噪声和固定模式噪声。Monakhova 等^[8] 在 ELD 的基础上将噪声细分为散粒噪声、读出噪声、行噪声、时间行噪声、量化噪声、固定模式噪声和周期噪声, 并将这些噪声参数输入到 GAN 网络中进行学习, 不需要手动校准噪声参数就可以得到质量较高的噪声模型。本文在此基础上提出了一种基于条件扩散模型和冷扩散模型的噪声生成器 LIGDM, 它先重复地往清晰的图像中加入噪声, 然后再反向得到噪声图像。LIGDM 不需要对抗训练, 只需要通过扩散过程即可得到噪声图像。本文的方法使用 Monakhova 等^[8] 开源的数据集进行训练, 并成功合成真实的噪声。

本文的 LIGDM 噪声模型考虑了散粒噪声、读出噪声、量化噪声和固定模式噪声。图像从光子到电子的转化过程中, 由于光子的量子性质, 通

常被建模为泊松分布, 而读出噪声可以被近似为高斯分布, 两者可以用一个异方差高斯噪声来表示, 从电压到数字信号的转化过程中, 会产生量化噪声, 其可以被看作输入电压与输出数字之间的舍入误差, 服从均匀分布。对于传感器捕获的原始图像, 本文将噪声的形成过程建模为

$$N = N_s + N_r + N_q + N_p, \quad (2)$$

式中: N_s 为散粒噪声; N_r 为读出噪声; N_q 为量化噪声; N_p 为固定模式噪声。

在确定噪声模型后, 本文通过模拟条件扩散模型的前向过程给清晰的图像添加噪声, 使图像变得模糊, 然后将模糊图像和清晰图像同时输入 Unet 网络进行噪声特性学习, 进一步通过反向推理合成噪声图像。

2.2 前向扩散模块

给定一个由配对图像组成的数据集, 记为 $D = \{x_i, y_i\}_{i=1}^N$, 表示从未知的条件分布中抽取的样本, 这是一个一对多的映射, 其中多个图像 y 可能与单个图像 x 对应, 而本文主要研究图像 y 与 x 相同的映射, 并学习它的随机迭代优化过程。本文将扩散概率模型应用于条件图像生成来解决这个问题。

条件扩散模型可以通过条件来控制输出结果, 对此, LIGDM 修改了 DDPM 的模型, 将模型的输入修改为两幅图像, 即引入一幅参考图像作为条件, 以此来引导扩散模型的输出。模型输入的另一幅图像是经过前向过程的模糊图像, 该过程通过逐渐地向清晰图像中加入噪声而得到过度加噪的图像, 该图像与参考图像一起输入到网络中, 最后输出噪声图像。

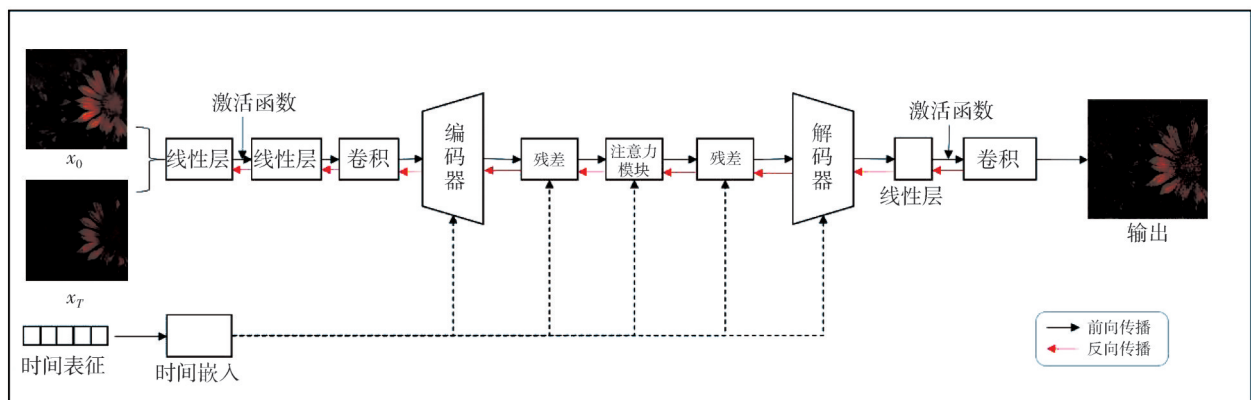


图2 网络图

Fig. 2 Network diagram

LIGDM的架构类似于DDPM中的Unet结构,如图2所示,输入的图像会先经过两个线性层和一个卷积层,然后会经过四层编码器,中间层由两个残差块和一个注意力模块组成,然后会经过四层解码器以及一个GroupNorm,最终输出结果。本文的损失函数使用均方误差损失,计算公式为

$$Loss = \|\epsilon - \epsilon_{\theta}(x_t, t)\|^2 \quad (3)$$

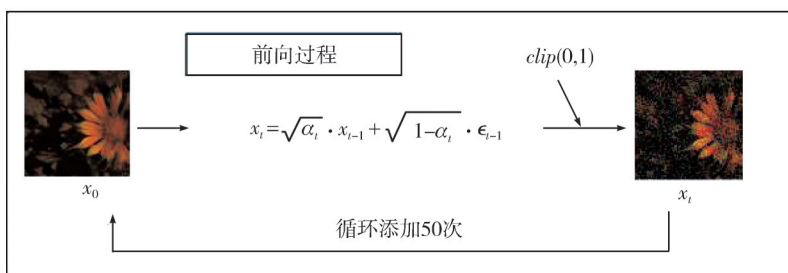


图3 前向过程示意图

Fig. 3 Forward process diagram

2.3 反向扩散模块

选择模糊算子S并训练模型R来进行恢复后,这些算子可以串联使用,通过使用扩散模型文献中的方法来逆转较为模糊的结果。对于t值较小的模糊来说,进行一次R就可以获得理想的结果,然而,当t较大时,会产生模糊的结果,因此需要迭代多次R。在冷扩散的光滑/可微情况下(非高斯随机噪声),标准扩散模型由于加入非高

之前的扩散模型在推理过程中需要1~2 000个扩散步骤,导致生成噪声图的效率较低。本文提出的LIGDM采用冷扩散模型中的方法来实现更加有效的推理,这允许其使用更少的推理步骤。如图3所示,本文参考冷扩散中的雪花算子推理步骤,将推理步骤设置为50步,并将各类噪声加入到这50步之内。

斯噪声导致推理结果很差,所以本文采用冷扩散中的算法2来进行推理。本文的迭代方法,每次迭代都需计算前一个状态的 $S(x_0, t)$,即前一时刻的过程图像。其中,迭代次数通过实验确定。

通过归纳发现,使用冷扩散模型中的采样算法2,生成的值能够达到本文方法的预期,该方法不通过对抗训练就可以获得质量较高的噪声图像。推理过程如图4所示。

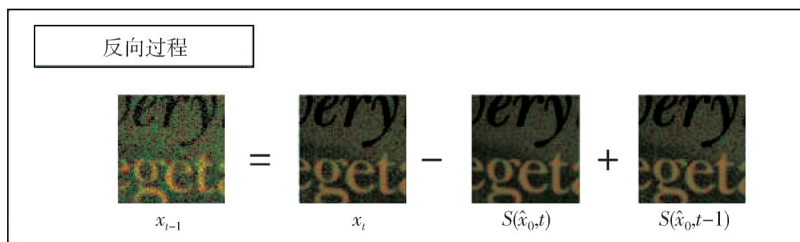


图4 反向过程示意图

Fig. 4 Reverse process diagram

3 实验验证

3.1 数据集及实验环境

本文使用Monakhova等^[8]公开的数据集进行训练和验证。其中,将成对的干净/嘈杂的静态场景中的灰度图像数据集用于训练本文LIGDM的噪声生成器,将自然场景下配对的干净/嘈杂的图像数据集用于验证,该数据集由67个成对的图像对组成,每个干净图像包含16个噪声脉冲。所有的图像都是在最低光环境下拍摄的,能够满足本

文进行噪声建模的任务要求。

本文的实验开发环境为基于python 3.7.13的深度学习框架pytorch 1.12.1,训练模型使用的是系统为Ubuntu 18.04 64位的NVIDIA Station服务器,内存为128 G,4个Tesla V100显卡,每张显卡的显存为32 G。本文的模型采用Adam优化器进行训练,初始学习率设置为0.000 01, batchsize设置为1。所有的图像都是以raw数据的形式储存,使用均方误差损失进行了700个epoch的训练,训练过程中学习率会降低,最低的学习率为 10^{-6} 。在训练到第616个epoch时,损失最低。

本文将数据集裁剪为像素 128×128 的图像块,在减去干净的图像后,计算了合成噪声和真实噪声之间的KL散度,以此来判断本文方法生成的噪声图像的真实性。

3.2 结果对比

在训练完LIGDM之后,利用星光下跳舞中的自然场景配对的数据集来验证本文所提方法LIGDM的可行性与先进性。本文比较了基于GAN网络的starlight模型^[8](baseline)、非深度低光噪声模型ELD^[7]、以及两种基于深度学习的噪声模型CA-GAN^[20]和noise Flow^[21]。如表1所示,Noise Flow和CA-CAN都忽略了真实噪声中

存在的条带噪声,ELD丢失了量化噪声,Starlight需要对抗训练,而本文方法LIGDM在该数据集中的表现较好,在定量指标方面,LIGDM所计算的KL散度与baseline接近。如图5所示,在定性效果方面,LIGDM能够较好地学习到最低光照条件下的噪声,与baseline效果接近,同时,本文方法不需要对抗训练,具有更好的可解释性。

表1 各类算法所计算的KL散度值

Tab.1 KL divergence values calculated by various algorithms

模型	KL散度值
ELD ^[7]	1.360
Noise Flow ^[21]	0.385
CA-GAN model ^[20]	0.513
Starlight ^[8]	0.069
本文算法	0.068

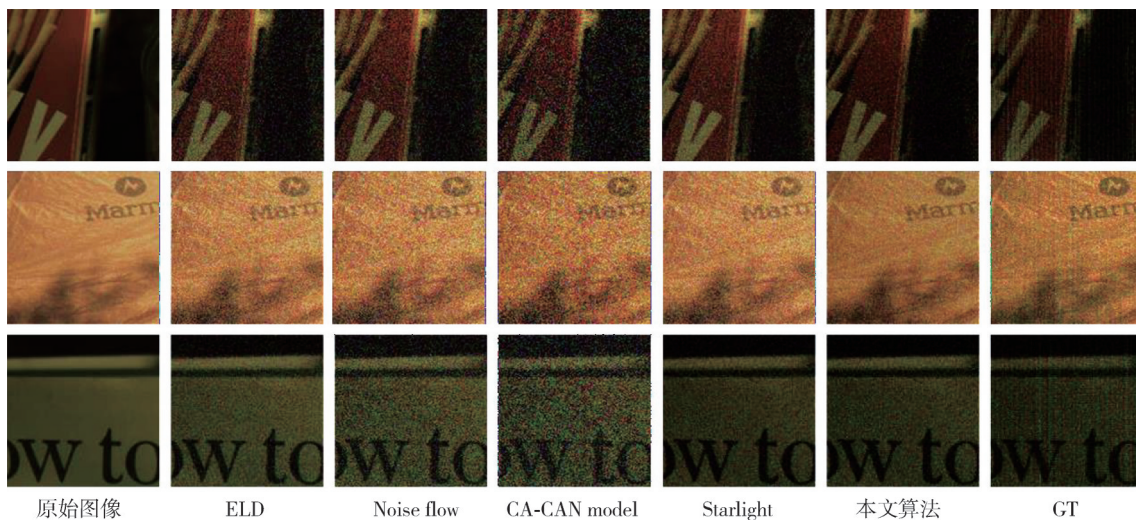


图5 不同算法生成的结果对比

Fig.5 Comparison of results generated by different algorithms

3.3 消融实验

本文比较了单独使用条件扩散模型和冷扩散模型的效果,并计算了合成噪声与真实噪声之间的KL散度值,结果如表2所示。

表2 单模块时的KL散度值

Tab.2 KL divergence value of single module

噪声模型类型	KL散度值
条件扩散模型	0.473
冷扩散模型	0.106
本文算法模型	0.068

进一步分析可知,当只使用条件扩散模型时,由于本文添加的噪声服从非高斯分布,因此不能生成非高斯分布的噪声图像;当只使用冷扩散模型时,由于模型训练过程中丢失了原始图像

的信息导致不能恢复出原始图像,所以不能生成噪声图像。当两者结合时,可以实现优势互补,最终生成较好的噪声图像。

4 结论

本文提出了一种弱光环境下的噪声图像生成方法,将扩散模型应用于弱光噪声建模领域,并取得较好的效果。首先,将不同噪声特性的噪声加入弱光干净图像中,提供足够的噪声分布,使其能够满足弱光环境中的噪声分布,然后通过反向迭代,细化每一种噪声参数,从而得到符合弱光环境下噪声分布的噪声图像,建立了更加真实的噪声模型。后续研究可用本文建立的噪声模型来去噪,以更加准确地去除噪声。

参考文献:

- [1] MCLEAN I S, WALKER G. Electronic imaging in astronomy: Detectors and instrumentation [J]. *Physics Today*, 1998, 51(6): 66.
- [2] 陈钱. 先进夜视成像技术发展探讨[J]. *红外与激光工程*, 2022, 51(2): 1-8.
CHEN Qian. Discussions on the development of advanced night vision imaging technology [J]. *Infrared and Laser Engineering*, 2022, 51(2): 1-8. (in Chinese)
- [3] CHEN C, CHEN Q F, XU J, et al. Learning to see in the dark[C]//IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2018: 3291-3300.
- [4] CHEN C, CHEN Q F, DO M, et al. Seeing motion in the dark[C]//IEEE/CVF International Conference on Computer Vision (ICCV), 2019: 3184-3193.
- [5] 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.
- [6] MAGGIONI M, BORACCHI G, FOI A, et al. Video denoising, deblocking, and enhancement through separable 4-d nonlocal spatiotemporal transforms [J]. *IEEE Transactions on Image Processing*, 2012, 21(9): 3952-3966.
- [7] WEI K X, FU Y, ZHENG Y Q, et al. Physics-based noise modeling for extreme low-light photography[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2022, 44(11): 8520-8537.
- [8] MONAKHOVA K, RICHTER S R, WALLER L, et al. Dancing under the stars: video denoising in starlight [C]//IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2022: 16220-16230.
- [9] GULRAJANI I, AHMED F, ARJOVSKY M, et al. Improved training of wasserstein gans [C]//Proceedings of the 31st International Conference on Neural Information Processing Systems, 2017: 5769-5779.
- [10] SAHARIA C, HO J, CHAN W, et al. Image super-resolution via iterative refinement [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2022, 45(4): 4713-4726.
- [11] OSHER S, BURGER M, GOLDFARB D, et al. An iterative regularization method for total variation-based image restoration [J]. *Multiscale Modeling & Simulation*, 2005, 4(2): 460-489.
- [12] RUDIN L I, OSHER S, FATEMI E. Nonlinear total variation based noise removal algorithms [J]. *Physica D: Nonlinear Phenomena*, 1992, 60(1/2/3/4): 259-268.
- [13] 陈作钧, 秦品乐, 柴锐, 等. 微光场景下的时空域 EMCCD 视频降噪算法 [J]. *中北大学学报(自然科学版)*, 2023, 44(2): 168-175.
CHEN Zuojun, QIN Pinle, CHAI Rui, et al. Spatio-temporal domain EMCCD video noise reduction algorithm in low light scene [J]. *Journal of North University of China(Natural Science Edition)*, 2023, 44(2): 168-175. (in Chinese)
- [14] 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.
- [15] 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.
- [16] GODARD C, MATZEN K, UYTENDAELE M. Deep burst denoising [C]//European Conference on Computer vision (ECCV), 2018: 538-554.
- [17] HO J, JAIN A, ABBEEL P. Denoising diffusion probabilistic models [J]. *Advances in Neural Information Processing Systems*, 2020, 33: 6840-6851.
- [18] HO J, SALIMANS T. Classifier-free diffusion guidance [DB/OL]. (2022-07-26)[2023-12-27]. <http://arxiv.org/abs/2207.12598>.
- [19] BANSAL A, BORGNIA E, CHU H M, et al. Cold diffusion: Inverting arbitrary image transforms without noise [DB/OL]. (2022-08-19)[2023-12-27]. <http://arxiv.org/abs/2208.09392>.
- [20] CHANG K C, WANG R, LIN H J, et al. Learning camera-aware noise models [C]//Computer Vision ECCV, 2020: 343-358.
- [21] ABDELHAMED A, BRUBAKER M A, BROWN M S. Noise flow: Noise modeling with conditional normalizing flows [C]//IEEE/CVF International Conference on Computer Vision, 2019: 3165-3173.