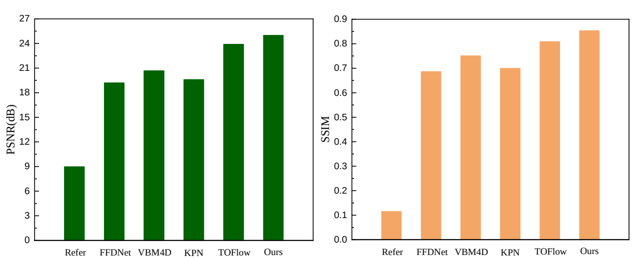
Supplementary Material

1. Experiments on Grasshopper3 camera

Additional analysis about Grasshopper3 camera model and more results of Canon camera can be found in this supplementary material. The noise parameters of the Grasshopper3 camera model are shown in the body of our paper.

Effect of network individually. To demonstrate the superiorities of the proposed network individually, we first compare our network with FFDNet [8], VBM4D [5], KPN [6] and TOFlow [7] upon the proposed noise model. Note that for fair comparison, the parameters of VBM4D [5] are choosen at the best performance and the learning based method, i.e. FFDNet [8], KPN [6], TOFlow [7] and our network, are trained on the same dataset and tested on the same noisy videos. As shown in Fig. 1, at least 1.1dB/0.05 improvement in PSNR/SSIM are introduced compared with the other methods. As shown in Fig. 2, the proposed network outperforms other methods in color fidelity and contrast. Note that both the FFDNet and the KPN do not work well in our scenario where the noise level is extremely higher than normal cases, although with the same training data of TOFlow and our method,

As a whole, the performance improvement introduced by the proposed network is demonstrated both quantitatively and qualitatively.



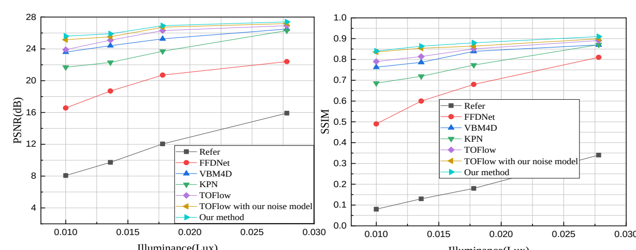
(a) (b)

Figure 1. Quantitative comparisons on the effect of network individually.

Effect of noise model. Then, to further verify the effectiveness of the proposed noise model individually, we train our network model upon different noise models, i.e. AWGN model, mixture of Gaussian and Poisson noise model, and the proposed noise model. The real captured dark noisy video is denoised by the trained network models as shown in the 1-3 th columns in Fig. 3. Note that limited by the body paper length, we also show a set additional results of Cannon camera video here. With the same network, the results of proposed noise model are of the best quality in terms of much less chrominance artifacts, more structual details and higher contrast, validating the superiorities of the proposed noise model for enhancing videos in low light condition.

In all, we demonstrate that the superiorities of the proposed method are attributed to both the proposed network and the proposed noise model.

Performance Analysis on Synthetic Data. We test our method on the synthetic test dataset generated with our practical noise model, and compare it with the other state-of-the-art algorithms. The comparisons are conducted on six test videos generated by simulating the environment illuminance from 0.01 to 0.03 Lux. As shown in Fig. 4, the proposed method achieves the highest scores in both PSNR and SSIM, over all the test videos and different luminance levels.



(a) (b)

Figure 4. Quantitative comparisons with the other methods under different illuminance intensities.

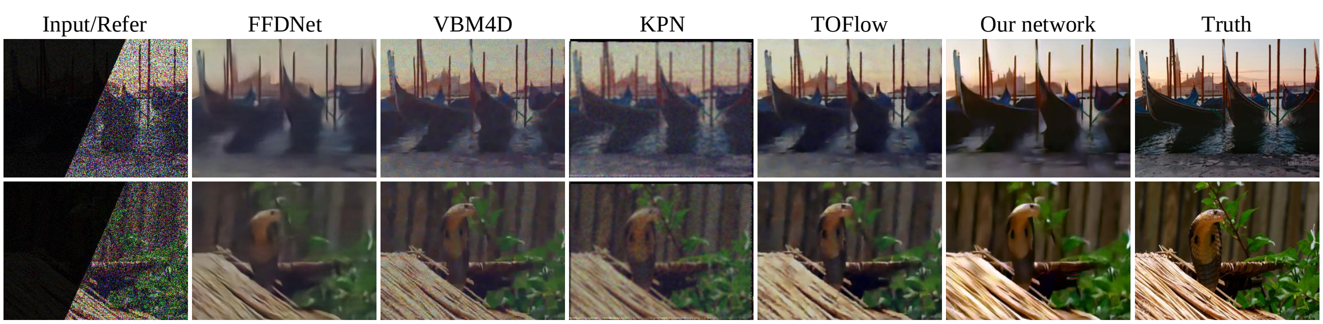


Figure 2. Qualitative comparisons on the effect of network individually. To facilitate visualization of the original input frames, we show in the first column the left halves of dark noisy input frames and the right halves of the brightness-scaled 'refer' frame.

Figure 3. Comparisons on the effect of the proposed noise model. Results show our network trained with different noise models enhancing real videos captured by Canon 5D MarkIII and Grasshopper3 GS3-U3-32S4C. 'Real data with light' denotes the same scene with the light turned on and the same camera setting parameters.

2. Experiments on Sony α7S 2 camera

The camera model of the Sony α7S 2 camera is added and the noise parameters of the camera model are calibrated shown in Tab. 1.

Comparisons with other methods on the real videos. We compare our method with FFDNet [8], VBM4D [5] and TOFlow [7] upon the real Captured Videos by Sony α7S 2 camera. As shown in Fig. 6, the results of the other methods lose many details on the extremely noisy region and cannot deal with the stripe noise well in the flat region, while our method could remove the noise effect clearly while preserving the details of images well.

Compare with Learning to See in the Dark [3]. Chen et al. [3] trained a map from the short exposed noisy images to the long exposed clean ones to enhance the noisy images captured in dark environment. However, due to the requirement of long exposure for capturing clean images, it is difficult to extend this method to dynamic video denoising. Besides, two specific camera models (Sony α7S 2 and Fujifilm X-T2) are provided with only the raw image as input, while those cameras are unable to obtain raw data when taking videos. However, we are still interested in comparing with this method at the same camera settings in the same low light scenario. We have calibrated the noise model parameters and trained the network for Sony α7S 2. A thorough comparison can be found at Fig. 7. Through comparing with the image captured during the day, our method is demonstrated to recover higher fidelity both in color and spatial details.

Table 1. Parameters of our practical noise model calibrated of Sony α7S 2 camera.

Sony α7S 2			
Parameters	R	G	B
$\sigma^2_{\beta^r_c}$	0.0082	0.0054	0.0058
K_c	4.51	3.26	3.49
N_d	35.16/s		
σ^2_R	9.81		

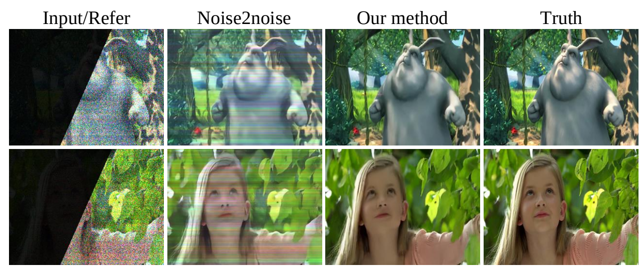


Figure 5. Comparisons on the synthetic videos with the noise-to-noise method.

3. Compare with noise2noise [4]

Lehtinen et al. [4] proposed to learn to turn bad images into good images by only looking at bad images, which is quite innovative and inspiring, so that we also compare the proposed method with noise2noise here. We train the noise-to-noise network with our noise model, following the official guide in [2]. The results are shown in Fig. 5. As compared, both method could remove a large extent of noise and recover the spatial details, while the proposed method could further remove the streak noise which could not be handled with [4].

More results of our paper can be found: `https://github.com/xiaotiantianweiweiwang/EnhancingLowLightVideos`.

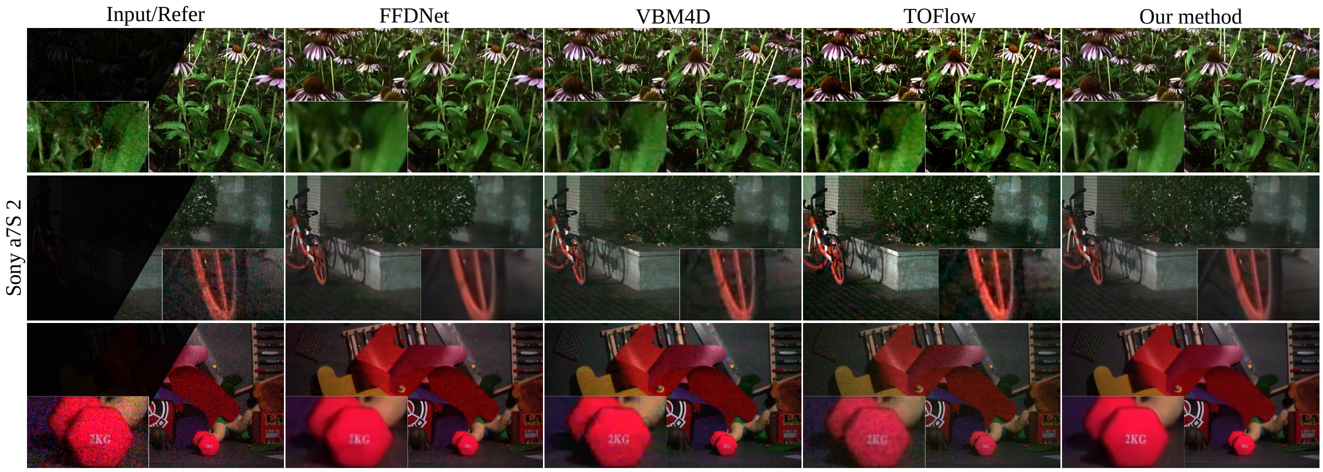


Figure 6. Results on real videos captured by Sony α7S 2.

Figure 7. Comparisons with see-in-dark at the same camera settings(ISO:128000,Exposure time:1/30) in the same low light scenario.

References

[1] Learning-to-see-in-the-dark. https://github.com/cchen156/Learning-to-See-in-the-Dark.

[2] Noise2noise. https://github.com/NVlabs/noise2noise. 2

[3] Chen Chen, Qifeng Chen, Jia Xu, and Vladlen Koltun. Learning to see in the dark. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 3291–3300, 2018. 2

[4] Jaakko Lehtinen, Jacob Munkberg, Jon Hasselgren, Samuli Laine, Tero Karras, Miika Aittala, and Timo Aila. Noise2noise: Learning image restoration without clean data. *arXiv preprint arXiv:1803.04189*, 2018. 2

[5] Matteo Maggioni, Giacomo Boracchi, Alessandro Foi, and Karen Egiazarian. Video denoising, deblocking, and enhancement through separable 4-d nonlocal spatiotemporal transforms. *IEEE Transactions on Image Processing*, 21(9):3952–3966, 2012. 1, 2

[6] Ben Mildenhall, Jonathan T Barron, Jiawen Chen, Dillon Sharlet, Ren Ng, and Robert Carroll. Burst denoising with kernel prediction networks. In *International Conference on Computer Vision and Pattern Recognition*, 2018. 1

[7] Tianfan Xue, Baian Chen, Jiajun Wu, Donglai Wei, and William T Freeman. Video enhancement with task-oriented flow. *International Journal of Computer Vision*, Feb 2019. 1, 2

[8] Kai Zhang, Wangmeng Zuo, and Lei Zhang. Ffdnet: Toward a fast and flexible solution for cnn-based image denoising. *IEEE Transactions on Image Processing*, 27(9):4608–4622, 2018. 1, 2