Supplementary Material for Paper: Learning to See in the Extremely Dark

Hai Jiang[1,*], Binhao Guan[2,*], Zhen Liu[2], Xiaohong Liu[3], Jian Yu[4], Zheng Liu[4],
Songchen Han[1], Shuaicheng Liu[2,†]
[1]School of Aeronautics and Astronautics, Sichuan University
[2]University of Electronic Science and Technology of China
[3]Shanghai Jiao Tong University [4]National Innovation Center for UHD Video Technology
jianghai@stu.scu.edu.cn, {guanbinhao@std., liushuaicheng@}uestc.edu.cn

This supplementary material is organized as follows:

- Sec. 1 provides more detailed information of our proposed See-in-the-Extremely-Dark (SIED) dataset.
- Sec. 2 provides a more comprehensive analysis of our method and the recently published diffusion-based low-light RAW2RAW image enhancement method ExposureDiff [19].
- Sec. 3 presents the limitation of our method.
- Sec. 4 conducts more experiments including real-world generalization analysis, cross-sensor generalization analysis, detailed ablation studies of our proposed AICM, and computational efficiency analysis.
- Sec. 5 provides more qualitative results on our proposed SIED dataset, the SID [2] dataset, and the results on the realistic extremely dark scenes.

1. More Details of the SIED Dataset

1.1. Noise Addition

In realistic scenes, the noise distribution can be divided into signal-dependent noise and signal-independent noise, where the signal-dependent noise follows the Poisson distribution and previous works typically treat the signal-independent noise as the Gaussian distribution [5, 6]. To this end, we first fit the Gaussian N_G and Poisson N_P noise distributions under various ISO values in the optical laboratory following [18], where the estimated noise parameters k and σ^2 of the Sony α7RIII and Canon EOS R cameras are illustrated in Fig. S1. With the ISO-k and ISO-σ^2 curves well fitted, the noise parameters under any ISO setting can be easily calculated, thus satisfying the requirements for synthesizing noisy data.

However, the noise model under extremely low-light conditions should not be considered a pure P+G model [21, 24]. Therefore, we additionally construct the dark-frame database N_D as shown in Fig. S2 (b) based on the dark frames used for Gaussian noise calibration to better satisfy the characteristics of realistic noise distribution following [24], where the 10 dark frames are captured for each ISO and the noise types that are hard to model explicitly, such as dark current noise, fixed pattern noise, and row noise, are all accumulated and saved in the shooting of the dark frames. Overall, the calibrated noise model can be formulated as $N_{all} = N_G + N_D + N_P$. Finally, we adopt the ISO-dependent noise addition strategy [18] to apply the calibrated noise model within various ISO values to construct realistic noisy images. As shown in Fig. S2, we illustrate the noise addition pipeline with an equivalent ISO of 6400 for Sony subset image generation.

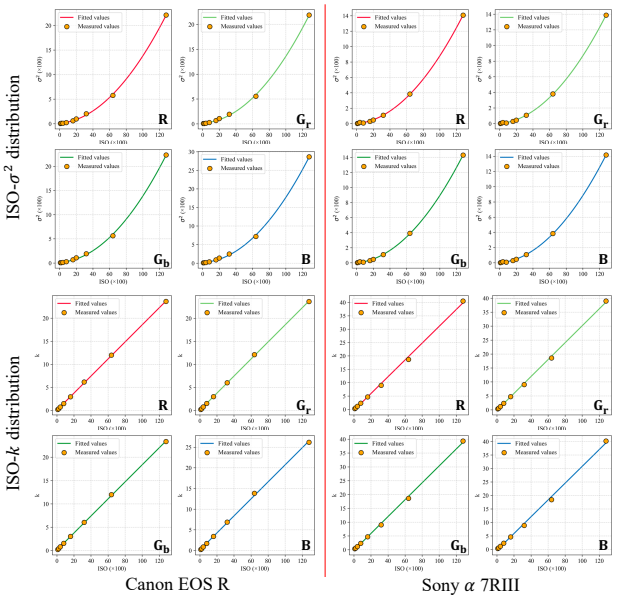


Figure S1. The estimated noise parameters ISO-k and ISO-σ^2 curves of the Canon EOS R and Sony α7RIII cameras.

*Equal contribution
†Corresponding author

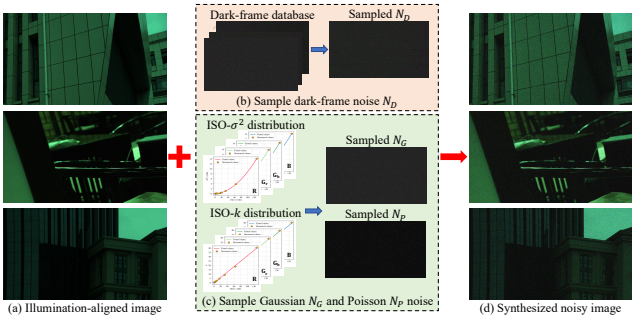


Figure S2. Illustration of the noise addition pipeline for Sony subset image synthesis with the illuminance of 0.01-0.1lux at ISO 6400, where (a) and (d) show the illumination-aligned images and the final noisy images which are visualized by a simple demosaicing process with a fixed amplification factor $\times 100$. (b) presents the dark-frame database and the sampled noise N_D, (c) presents the estimated ISO-k and ISO-σ^2 curves and the sampled Gaussian N_G and Poisson N_P noise.

Table S1. The KL divergence (↓) of illumination histogram distributions between the images synthesized by our strategy and previous strategies [1, 12, 20, 23] and laboratory images.

Method	0.01-0.1 lux		0.001-0.01 lux		0.0001-0.001 lux	
	Canon	Sony	Canon	Sony	Canon	Sony
Previous	0.165	0.166	0.227	0.147	0.216	0.339
Ours	0.017	0.011	0.018	0.009	0.011	0.059

1.2. Compare with Previous Synthesis Strategies

Previous dataset synthesis approaches [1, 12, 20, 23] typically generate low-light images by degrading the illumination of normal-light images, which inadequately captures the imaging characteristics of realistic dark scenes. To this end, we present a paired-to-paired data synthesis strategy to generate well-calibrated extremely dark images from collected real-world low-light data.

To validate the superiority of our data synthesis pipeline, we employ the illumination alignment and noise addition processes in Sec.3 of the main paper to transform the collected normal-light images into low-light counterparts following previous approaches, and endeavor to align the illumination distributions of the synthesized images consistent with standard images. However, the illumination distribution of low-light images synthesized from normal-light images is difficult to align with the standard image, as shown in Fig. S3. The KL divergence (↓) of illumination histogram distributions between the synthesized images obtained in the above approach and laboratory images is at least 10 times over our paired-to-paired synthesis strategy, as reported in Table S1.

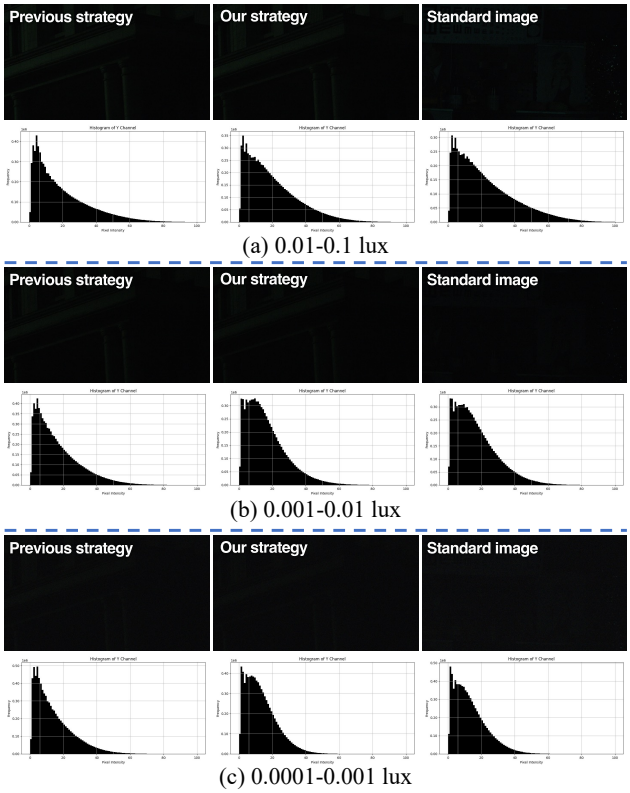


Figure S3. Comparison of images generated by previous data synthesis strategies [1, 12, 20, 23] and our strategy.

1.3. Align with Standard Data

As we described in Sec. 3 in the main paper, we adopt a fixed ISP pipeline to transform the generated RAW image and standard images into YUV space, and manually finetune η in Eq.(1) to match the illumination histograms of the two images in the Y channel. As shown in Fig. S4, we present several samples of illumination histograms for our generated images and standard laboratory images. Moreover, we employ fixed amplification factors, i.e., $\times 100$, $\times 200$, $\times 300$, and $\times 500$, to improve the contrast of the low-light RAW images followed by a simple demosaicing process for visualization. As we can see, our generated images are quite approaching the standard data in terms of both illumination distribution and subjective visualization, proving that our paired-to-paired data synthesis strategy is capable of generating realistic extremely low-light images.

2. More Detailed Analysis

In this section, we provide an analysis of our method and the recently published diffusion-based RAW2RAW image enhancement method ExposureDiff [19].

- ExposedDiff targets RAW2RAW image enhancement that only takes exposure correction and denoising in the

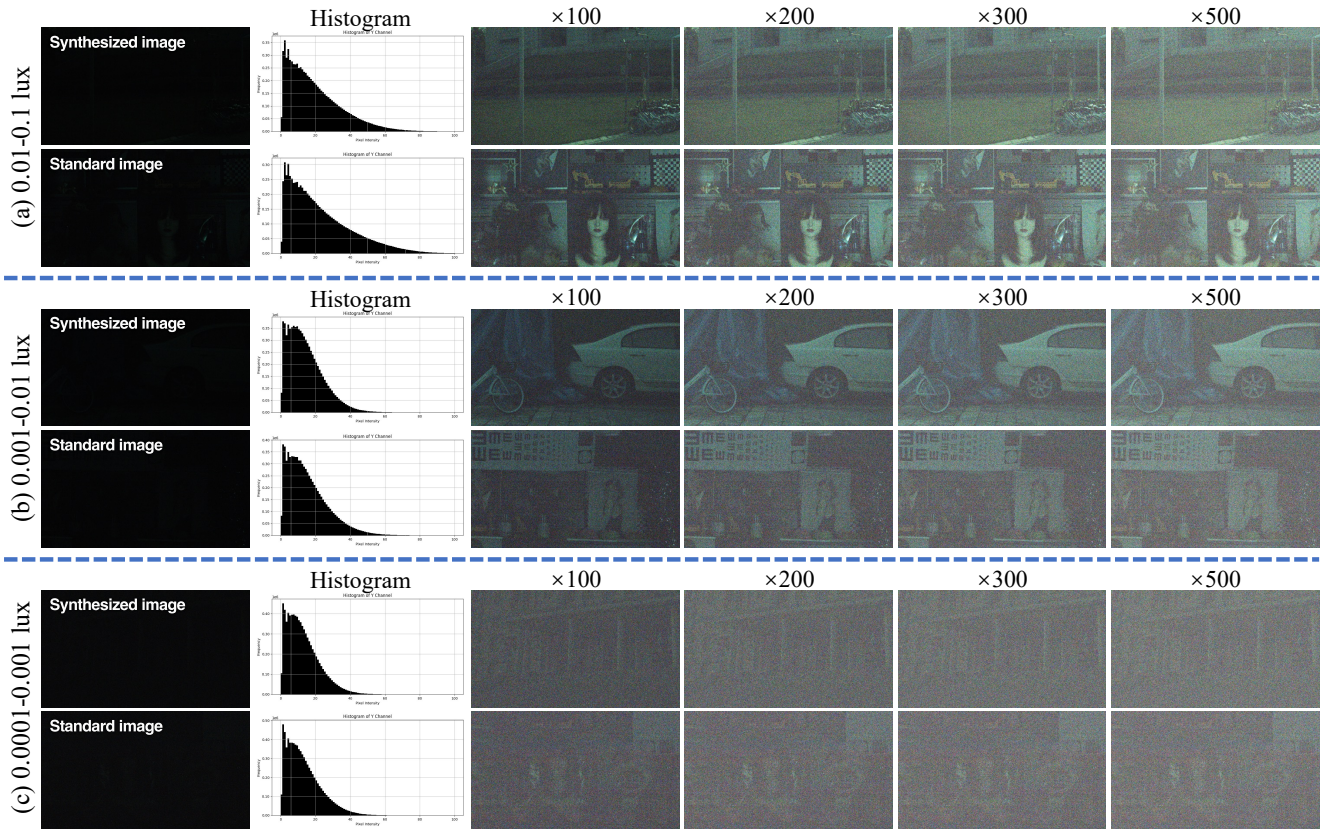


Figure S4. Illustration of illumination histograms for our generated images and standard images, as well as the visualization results with various fixed amplification factors, i.e., ×100, ×200, ×300, and ×500.

Figure S5. Capturing images in the realistic extremely dark scene using various camera parameters. These images are in sRGB format produced by camera default ISP in a 0.001lux-0.01lux scene.

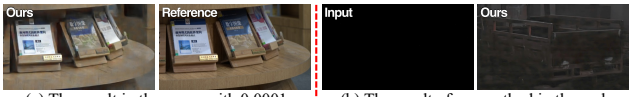
RAW domain into account. In contrast, our method aims to transform low-light noisy RAW images into well-exposed high-quality sRGB images, which not only takes into account exposure correction and denoising but also the complex ISP processes modeling, thus being more challenging than the RAW2RAW image enhancement.

- For exposure correction, ExposureDiff follows the previous methods [2, 3, 8] that utilizes the exposure values of GT images as a prior for amplification. However, it is impractical to obtain high-quality reference images in realistic extremely dark scenes. As shown in Fig. S5, the high-quality normal-light images are unable to be obtained in the realistic extremely dark scene even with exposure times as long as 30 seconds. To this end, we propose an Adaptive Illumination Correction Module (AICM) to estimate the amplification factor from the low-light RAW feature without relying on the reference images.
- Our method designs a color consistency loss to facilitate the diffusion model to generate reconstructed sRGB features with vivid color, while ExposureDiff focuses on image enhancement in the RAW domain without considering the color mapping in the RAW2RGB process.

To validate the advantages of our method over ExposureDiff, we have modified our framework to RAW2RAW image enhancement and conducted experiments on the Sony subset of the SID [2] dataset following the settings of ExposureDiff. As shown in Table S2, our method also achieves performance superior to ExposureDiff, further proving the effectiveness of our method.

Table S2. Quantitative results of our method and ExposureDiff [19] for RAW2RAW image enhancement on the Sony subset of the SID [2] dataset. The best results are highlighted in **bold**.

Ratio	Metric	ExposureDiff [19]	Ours
×100	PSNR ↑	38.89	**39.44**
	SSIM ↑	0.902	**0.921**
×250	PSNR ↑	36.02	**36.37**
	SSIM ↑	0.832	**0.843**
×300	PSNR ↑	35.00	**35.19**
	SSIM ↑	0.808	**0.814**



(a) The result in the scene with 0.0001-0.001 lux and ISO=40000

(b) The result of our method in the real-world scene below 0.0001 lux

Figure S6. The results of our method in the scene with 0.0001-0.001 lux and ISO=4000 (a) and the scene below 0.0001 lux (b).

3. Limitations

Even our method works well on most scenes, while detail blurring may occur under extremely low-light and high-noise conditions, as shown in Fig. S6(a). Moreover, as shown in Fig. S6(b), our method still retains its capability in real-world dark scenes below 0.0001 lux. However, the result remains suboptimal due to the extremely limited content information and more severe degradation. Solving such ultra-dark scenes will be our future work. Another limitation of our method lies in the inference efficiency since diffusion-based methods iteratively denoise the randomly sampled Gaussian noise to obtain reconstructed results in the inference phase, leading to the inference speed being affected by the sampling step. Therefore, our method shows inferior efficiency compared to some lightweight methods, as discussed in Sec. 4.4. In the future, we will explore more effective sampling strategies, such as DPM-Solver [13] and the consistency model [16], to improve inference efficiency and investigate the effectiveness of our method for extremely low-light RAW video enhancement.

4. More Experiments

4.1. Real-world Generalization

As we mentioned in the main manuscript, we collect several images in realistic extremely dark scenes using the Sony camera and determine their illuminance level according to the matching of the illumination histograms in the Y-channel to standard laboratory images, containing 11 images with the range of 0.01-0.1 lux, 13 images with the range of 0.001-0.01 lux, and 6 images with the range of 0.0001-0.001 lux. To validate the generalization ability of our method, we adopt three non-reference perceptual metrics NIQE [15], MUSIQ [9], and CLIPIQA [17] to measure the visual quality of the restored results of our method and comparison methods. As reported in Table S3, our method achieves the best MUSIQ and CLIPIQA scores in all three illuminance ranges, proving the effectiveness of the proposed method and our synthesized dataset is capable of supporting methods to generalize to real-world scenes.

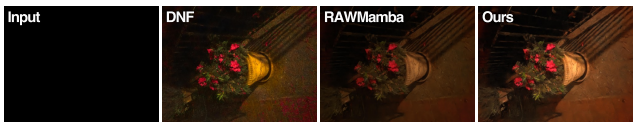


Figure S7. The Cross-sensor generalization results on the iPhone image of DNF [8], RAWMamba [3], and our method that are trained on the Sony subset of the SIED dataset.

Figure S8. Qualitative comparison of RAWMamba [3] with various amplification factors and our method on the realistic scene.

4.2. Cross-sensor Generalization

To validate the generalization ability of the methods trained on our SIED dataset to images captured by other sensors, we apply the model trained on the Sony subset of the SIED dataset to restore the realistic dark images captured by iPhone 6s [2]. As shown in Fig. S7, the restored results of our method and previous methods DNF [8] and RAWMamba [3] present property contrast, low noise, and vivid color, demonstrating the methods trained on our dataset perform well in cross-sensor generalization.

4.3. More Detailed Results of the Ablation Study

In this section, we conduct more detailed experiments to evaluate the effectiveness of our proposed Adaptive Illumination Correction Module (AICM). We use the implementation details described in Sec. 5.1 of the main paper for training, and quantitative results on the Sony subset of the SID [2] dataset are illustrated in Table S4. Detailed experiment settings are discussed below.

We first remove the proposed AICM from the overall framework, as shown in row 1 of Table S4, which would cause overall performance degradation. Then, we adopt fixed amplification factors (Amp.) of 100, 200, and 300 to improve the contrast of the input RAW images, as shown in row 2-4 of Table S4, while such constant amplification factors are unable to achieve satisfactory exposure improvement in realistic scenes where the illumination degradation is diverse and unknown. Meanwhile, we follow previous methods [2, 3, 7, 8, 11, 14] to employ the exposure infor-

Table S3. Non-reference perceptual metrics of our method and comparison methods on the realistic extremely dark scenes. The best results are highlighted in **bold** and the second-best results are in underlined.

	Method	0.01-0.1 lux			0.001-0.01 lux			0.0001-0.001 lux		
		NIQE ↓	MUSIQ ↑	CLIPIQA ↑	NIQE ↓	MUSIQ ↑	CLIPIQA ↑	NIQE ↓	MUSIQ ↑	CLIPIQA ↑
Single-stage	SID [2]	7.359	23.122	0.284	7.788	22.484	0.286	8.654	20.458	0.253
	DID [14]	7.642	24.423	0.251	8.365	22.392	0.238	8.750	21.429	0.227
	SGN [7]	7.386	24.380	0.302	8.486	22.464	0.288	8.983	21.461	0.268
	LLPackNet [11]	8.636	22.684	0.259	9.260	22.163	0.261	10.096	20.380	0.260
	RRT [10]	8.082	21.199	0.263	8.492	20.739	0.277	9.875	19.642	0.259
Multi-stage	LDC [22]	9.051	21.374	0.275	10.504	19.866	0.239	12.438	19.657	0.273
	MCR [4]	7.913	23.703	0.271	8.897	22.958	0.274	9.440	20.816	0.256
	DNF [8]	7.712	23.855	0.288	8.912	22.769	0.276	**8.451**	21.339	0.256
	RAWMamba [3]	**7.354**	23.318	0.276	8.044	22.075	0.283	8.489	21.826	0.266
	Ours	7.654	**24.694**	**0.321**	**7.403**	**23.031**	**0.314**	8.480	**21.913**	**0.312**

Table S4. Quantitative results of ablation studies on the Sony subset of the SID [2] dataset. 'w/o' denotes without.

	Method	PSNR ↑	SSIM ↑	LPIPS ↓
1)	w/o AICM	30.02 (−1.18)	0.785 (−0.016)	0.369 (+0.030)
2)	Amp. = 100	30.91 (−0.29)	0.794 (−0.007)	0.351 (+0.012)
3)	Amp. = 200	30.77 (−0.43)	0.798 (−0.003)	0.349 (+0.010)
4)	Amp. = 300	30.63 (−0.57)	0.793 (−0.008)	0.355 (+0.016)
5)	GT exposure	31.29 (+0.09)	0.806 (+0.005)	0.333 (−0.006)
6)	default	31.20 (+0.00)	0.801 (+0.000)	0.339 (+0.000)

Table S5. Run time (Time) and memory consumption (Mem.) of our method and comparison methods on NVIDIA A100 when performing inference on an image with 3840×2160 resolution.

	SID [2]	DID [14]	SGN [7]	LLPackNet [11]	RRT [10]
Time (s)	0.053	0.393	0.066	0.026	0.033
Mem. (GB)	4.166	9.119	3.090	1.676	1.838
	LDC [22]	MCR [4]	DNF [8]	RAWMamba [3]	Ours
Time (s)	0.275	0.094	1.143	1.935	0.951
Mem. (GB)	14.76	7.117	7.395	38.164	7.736

mation of GT images as priors to amplify the input low-light RAW images. As shown in rows 5-6 of Table S4, our AICM presents comparable performance compared to adopting GT exposure for pre-amplification, which is unlikely to be available in practical applications, demonstrating the effectiveness of our proposed AICM.

Furthermore, for previous methods [2, 3, 8, 11, 14] rely on the ratio between the exposure time of normal/low-light images as a prior for amplification during the training stage, we use the ratio of mean values between paired images as the factor in [10.6, 565.8] for retraining on our SIED dataset and apply fixed amplification factor for evaluation in the realistic dark scenes as in SID [2]. As shown in Fig. S8, we also apply various factors to RAWMamba [3] for inference in the realistic scenes, where different factors would bring different exposure levels, while it is impractical to adapt proper factors to variable scenes, while our well-designed AICM can achieve adaptive exposure correction, further proving the effectiveness of our AICM.

4.4. Computational Efficiency

In this section, we report the inference time and memory consumption of our method and comparison methods on NVIDIA A100 GPU when performing inference on the test set of our proposed SIED dataset, where the input RAW images are with the resolution of 3,840×2,160. As reported in Table S5, the computational efficiency and memory consumption of our method are 0.951s and 7.736GB, respectively, which is inferior to lightweight methods, such as LLPackNet [11] and RRT [10], while our method achieves considerable performance improvements over them. Moreover, compared with recently published state-of-the-art competitors DNF [8] and RAWMamba [3], our approach achieves advantages in computational resource consumption and inference speed, especially compared to RAWMamba with a significant reduction in memory consumption, while achieving performance superior.

5. More Qualitative Results

We present more visual comparisons of our method and comparison methods on each illuminance level of the Canon and Sony subsets of our proposed SIED dataset in Fig. S9, Fig. S10, Fig. S11, Fig. S12, Fig. S13, and Fig. S14. As we can see, previous methods [2–4, 7, 8, 10, 11, 14, 22] encounter unexpected artifacts, color distortion, or noise amplification. In contrast, our method properly improves global and local contrast, reconstructs sharper details, and presents vivid color, resulting in visual pleasant results.

We also provide the visual comparisons of our method and comparison methods on the Sony subset of the SID dataset, as illustrated in Fig. S15. Our method does not use the GT exposure as a prior for pre-amplification while gen-

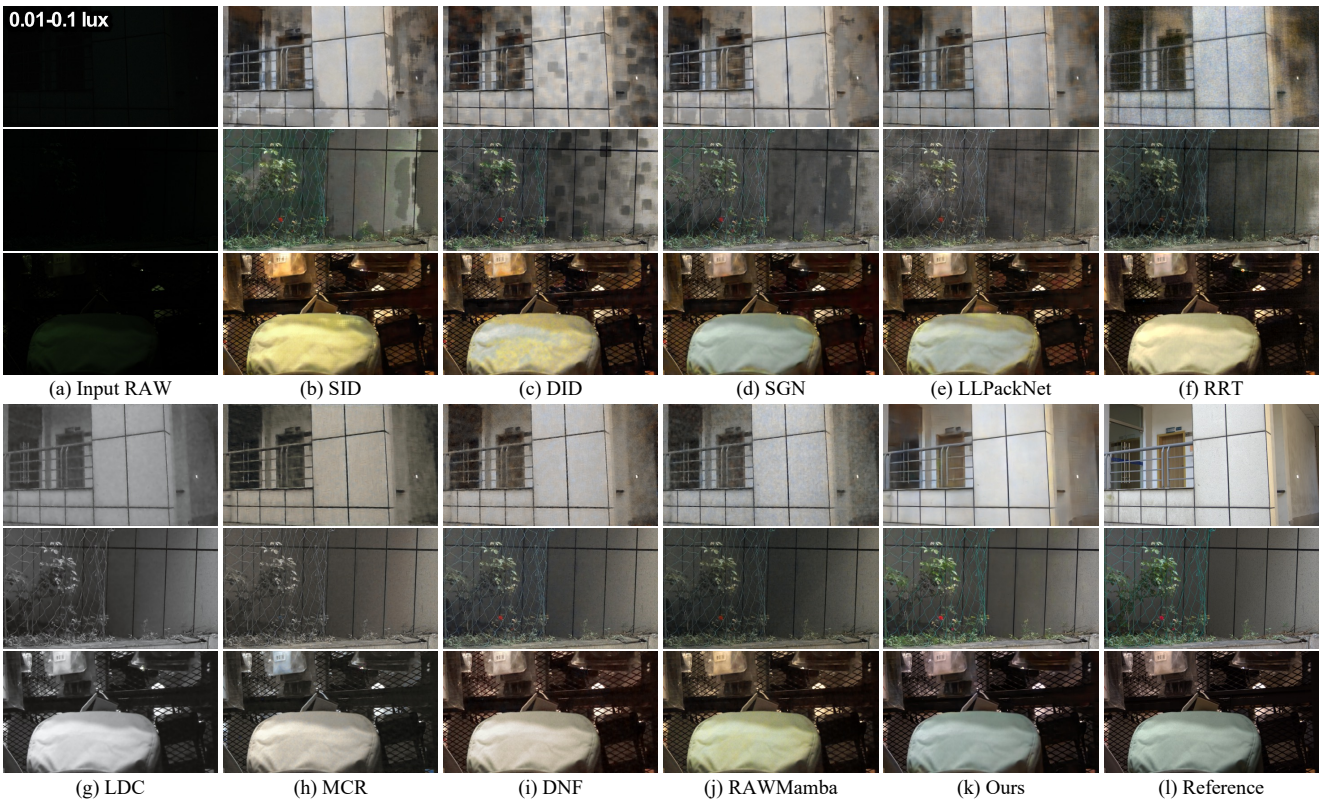


Figure S9. Qualitative comparison of our method and comparison methods [2–4, 7, 8, 10, 11, 14, 22] on the Canon subset of our SIED dataset at the illuminance range of 0.01-0.1 lux. The input RAW images are visualized by a simple demosaicing process.

erating results closer to reference sRGB images, proving the effectiveness of our approach. Moreover, we provide the restored results of our method and comparison method on the realistic extremely dark scenes in Fig S16, where all methods are capable of transforming realistic extremely low-light inputs into normal-light sRGB images while our method demonstrates superior performance.

References

[1] Vladimir Bychkovsky, Sylvain Paris, Eric Chan, and Frédo Durand. Learning photographic global tonal adjustment with a database of input/output image pairs. In *CVPR*, pages 97–104, 2011. 2

[2] Chen Chen, Qifeng Chen, Jia Xu, and Vladlen Koltun. Learning to see in the dark. In *CVPR*, pages 3291–3300, 2018. 1, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12

[3] Xianmin Chen, Peiliang Huang, Xiaoxu Feng, Dingwen Zhang, Longfei Han, and Junwei Han. Retinex-rawmamba: Bridging demosaicing and denoising for low-light raw image enhancement. *arXiv preprint arXiv:2409.07040*, 2024. 3, 4, 5

[4] Xingbo Dong, Wanyan Xu, Zhihui Miao, Lan Ma, Chao Zhang, Jiewen Yang, Zhe Jin, Andrew Beng Jin Teoh, and Jiajun Shen. Abandoning the bayer-filter to see in the dark. In *CVPR*, pages 17431–17440, 2022. 5, 6, 7, 8, 9, 10, 11, 12

[5] Alessandro Foi. Clipped noisy images: Heteroskedastic modeling and practical denoising. *Signal Processing*, 89 (12):2609–2629, 2009. 1

[6] Alessandro Foi, Mejdi Trimeche, Vladimir Katkovnik, and Karen Egiazarian. Practical poissonian-gaussian noise modeling and fitting for single-image raw-data. *IEEE TIP*, 17 (10):1737–1754, 2008. 1

[7] Shuhang Gu, Yawei Li, Luc Van Gool, and Radu Timofte. Self-guided network for fast image denoising. In *ICCV*, pages 2511–2520, 2019. 4, 5, 6, 7, 8, 9, 10, 11, 12

[8] Xin Jin, Ling-Hao Han, Zhen Li, Chun-Le Guo, Zhi Chai, and Chongyi Li. Dnf: Decouple and feedback network for seeing in the dark. In *CVPR*, pages 18135–18144, 2023. 3, 4, 5, 6, 7, 8, 9, 10, 11, 12

[9] Junjie Ke, Qifei Wang, Yilin Wang, Peyman Milanfar, and Feng Yang. Musiq: Multi-scale image quality transformer. In *ICCV*, pages 5148–5157, 2021. 4

[10] Mohit Lamba and Kaushik Mitra. Restoring extremely dark images in real time. In *CVPR*, pages 3487–3497, 2021. 5, 6, 7, 8, 9, 10, 11, 12

[11] Mohit Lamba, Atul Balaji, and Kaushik Mitra. Towards fast and light-weight restoration of dark images. *arXiv preprint arXiv:2011.14133*, 2020. 4, 5, 6, 7, 8, 9, 10, 11, 12

[12] Kin Gwn Lore, Adedotun Akintayo, and Soumik Sarkar. Llnet: A deep autoencoder approach to natural low-light image enhancement. *PR*, 61:650–662, 2017. 2

(a) Input RAW (b) SID (c) DID (d) SGN (e) LLPackNet (f) RRT

(g) LDC (h) MCR (i) DNF (j) RAWMamba (k) Ours (l) Reference

Figure S10. Qualitative comparison of our method and comparison methods [2–4, 7, 8, 10, 11, 14, 22] on the Canon subset of our SIED dataset at the illuminance range of 0.001-0.01 lux. The input RAW images are visualized by a simple demosaicing process.

[13] Cheng Lu, Yuhao Zhou, Fan Bao, Jianfei Chen, Chongxuan Li, and Jun Zhu. Dpm-solver: A fast ode solver for diffusion probabilistic model sampling in around 10 steps. In *NeurIPS*, pages 5775–5787, 2022. 4

[14] Paras Maharjan, Li Li, Zhu Li, Ning Xu, Chongyang Ma, and Yue Li. Improving extreme low-light image denoising via residual learning. In *ICME*, pages 916–921, 2019. 4, 5, 6, 7, 8, 9, 10, 11, 12

[15] Anish Mittal, Rajiv Soundararajan, and Alan C Bovik. Making a "completely blind" image quality analyzer. *IEEE Sign. Process. Letters*, 20(3):209–212, 2012. 4

[16] Yang Song, Prafulla Dhariwal, Mark Chen, and Ilya Sutskever. Consistency models. In *ICML*, pages 32211–32252, 2023. 4

[17] Jianyi Wang, Kelvin CK Chan, and Chen Change Loy. Exploring clip for assessing the look and feel of images. In *AAAI*, pages 2555–2563, 2023. 4

[18] Yuzhi Wang, Haibin Huang, Qin Xu, Jiaming Liu, Yiqun Liu, and Jue Wang. Practical deep raw image denoising on mobile devices. In *ECCV*, pages 1–16, 2020. 1

[19] Yufei Wang, Yi Yu, Wenhan Yang, Lanqing Guo, Lap-Pui Chau, Alex C Kot, and Bihan Wen. Exposurediffusion: Learning to expose for low-light image enhancement. In *ICCV*, pages 12438–12448, 2023. 1, 2, 4

[20] Chen Wei, Wenjing Wang, Wenhan Yang, and Jiaying Liu. Deep retinex decomposition for low-light enhancement. In *BMVC*, 2018. 2

[21] Kaixuan Wei, Ying Fu, Yinqiang Zheng, and Jiaolong Yang. Physics-based noise modeling for extreme low-light photography. *IEEE TPAMI*, 44(11):8520–8537, 2021. 1

[22] Ke Xu, Xin Yang, Baocai Yin, and Rynson WH Lau. Learning to restore low-light images via decomposition-and-enhancement. In *CVPR*, pages 2281–2290, 2020. 5, 6, 7, 8, 9, 10, 11

[23] Wenhan Yang, Wenjing Wang, Haofeng Huang, Shiqi Wang, and Jiaying Liu. Sparse gradient regularized deep retinex network for robust low-light image enhancement. *IEEE TIP*, 30:2072–2086, 2021. 2

[24] Yi Zhang, Hongwei Qin, Xiaogang Wang, and Hongsheng Li. Rethinking noise synthesis and modeling in raw denoising. In *ICCV*, pages 4593–4601, 2021. 1

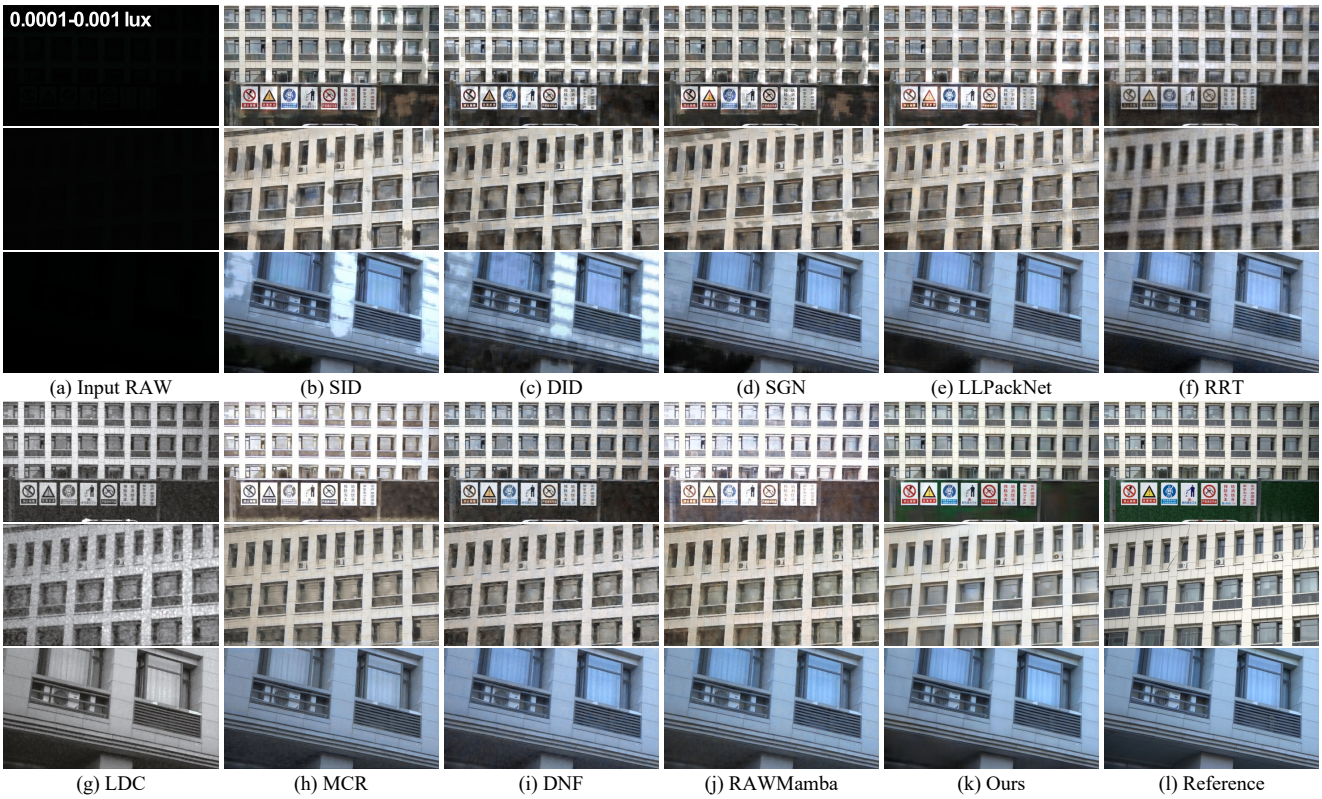
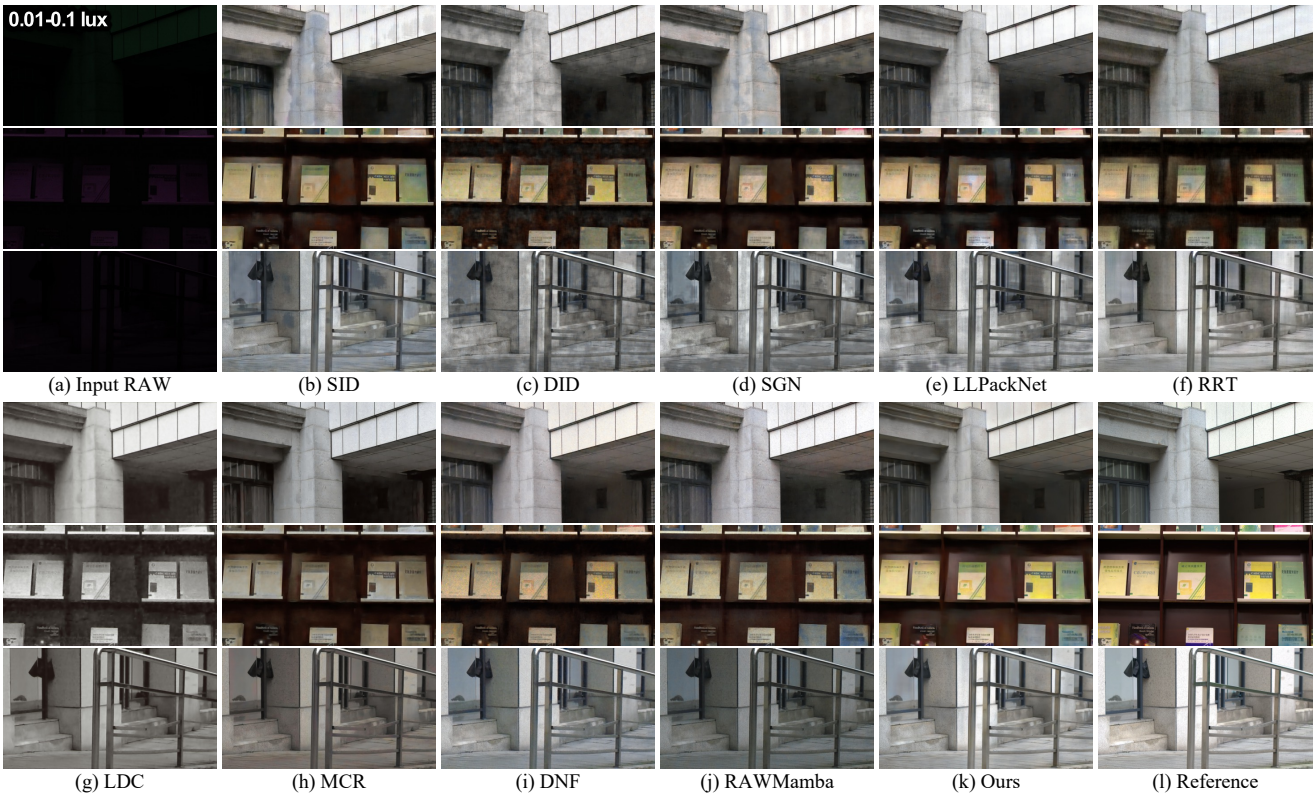


Figure S11. Qualitative comparison of our method and comparison methods [2–4, 7, 8, 10, 11, 14, 22] on the Canon subset of our SIED dataset at the illuminance range of 0.0001-0.001 lux. The input RAW images are visualized by a simple demosaicing process.



(a) Input RAW (b) SID (c) DID (d) SGN (e) LLPackNet (f) RRT

(g) LDC (h) MCR (i) DNF (j) RAWMamba (k) Ours (l) Reference

Figure S12. Qualitative comparison of our method and comparison methods [2–4, 7, 8, 10, 11, 14, 22] on the Sony subset of our SIED dataset at the illuminance range of 0.01-0.1 lux. The input RAW images are visualized by a simple demosaicing process.

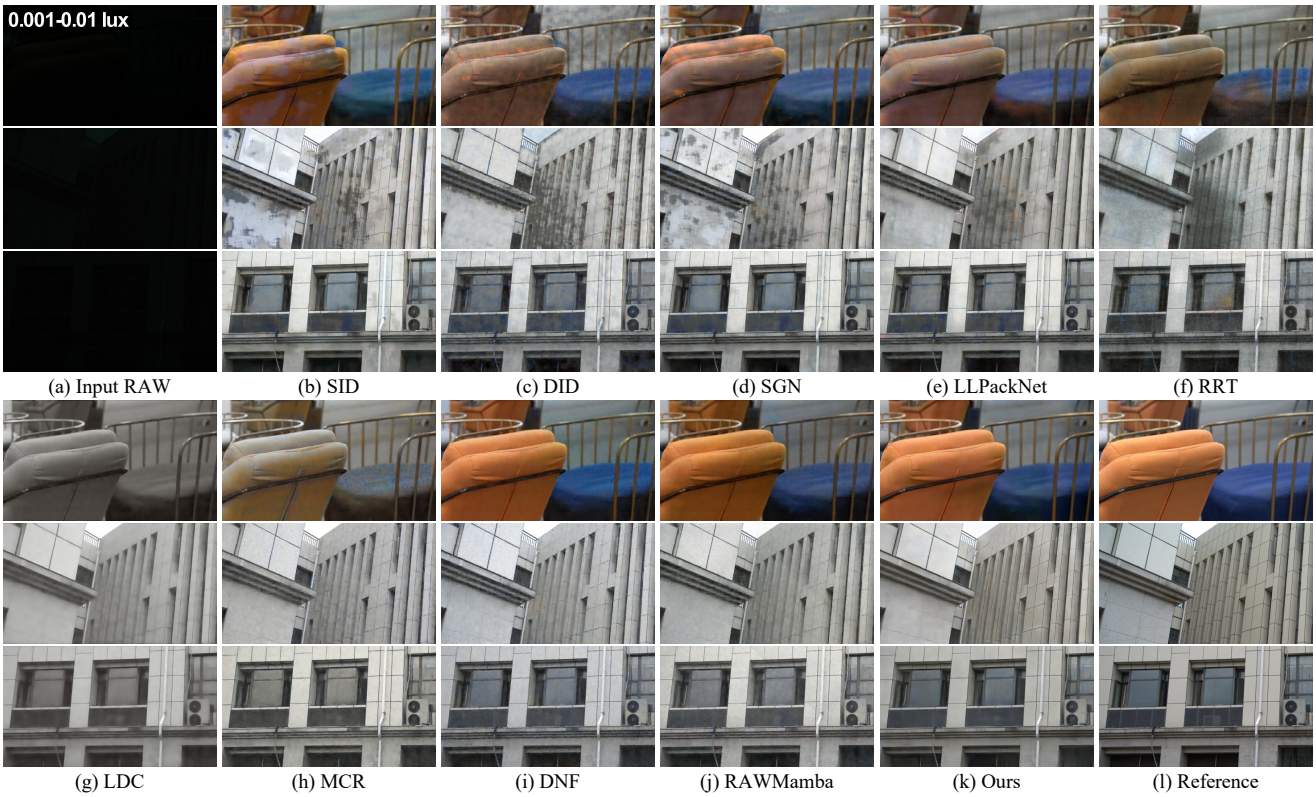


Figure S13. Qualitative comparison of our method and comparison methods [2–4, 7, 8, 10, 11, 14, 22] on the Sony subset of our SIED dataset at the illuminance range of 0.001-0.01 lux. The input RAW images are visualized by a simple demosaicing process.

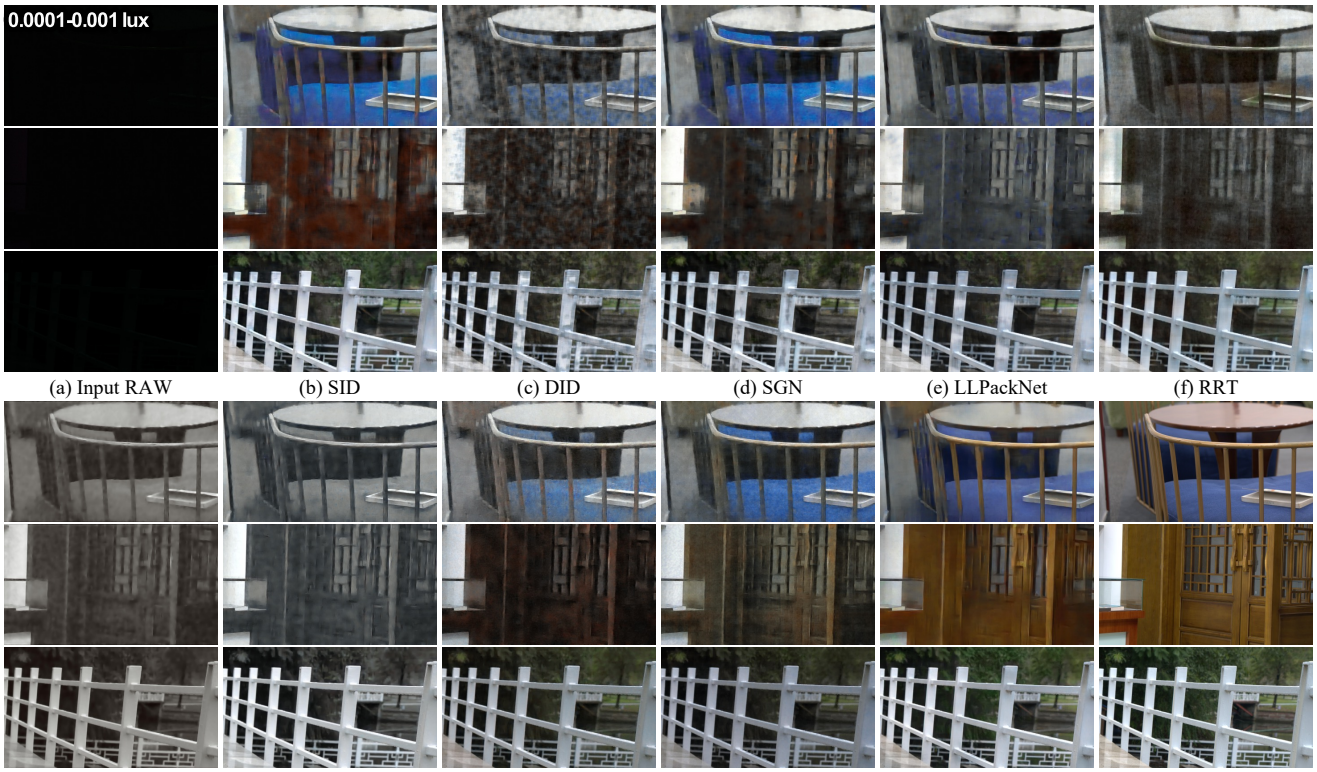


Figure S14. Qualitative comparison of our method and comparison methods [2–4, 7, 8, 10, 11, 14, 22] on the Sony subset of our SIED dataset at the illuminance range of 0.0001-0.001 lux. The input RAW images are visualized by a simple demosaicing process.

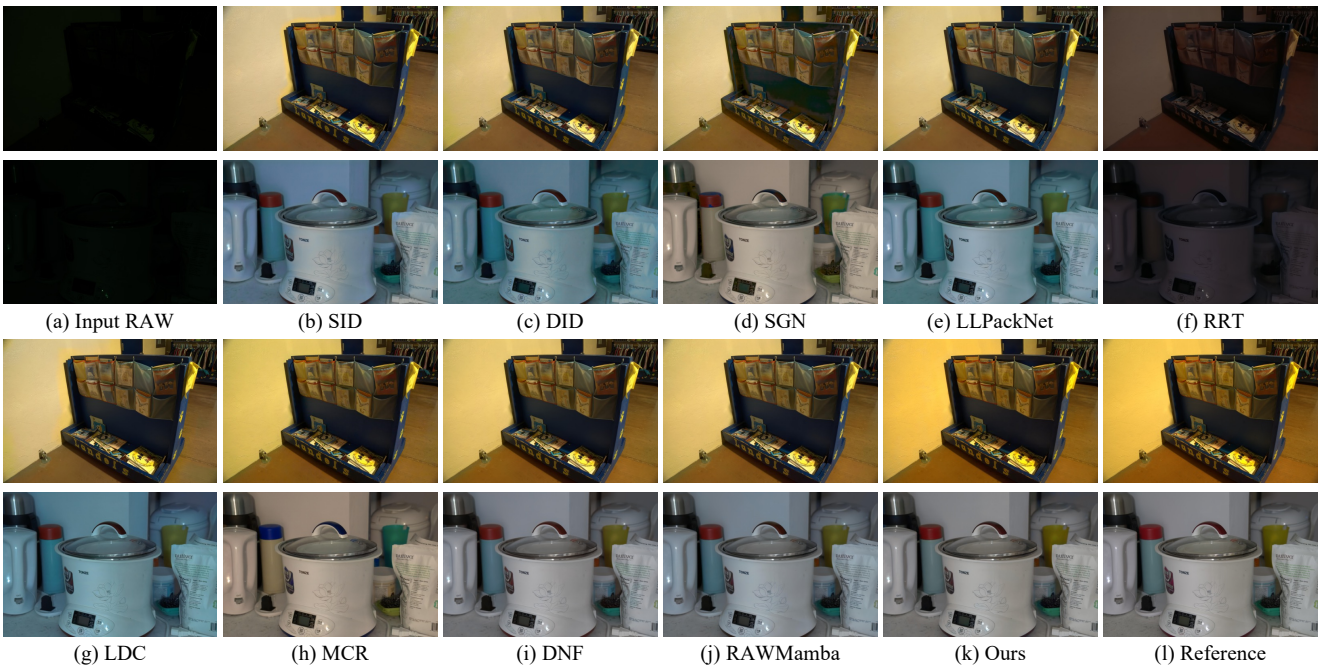
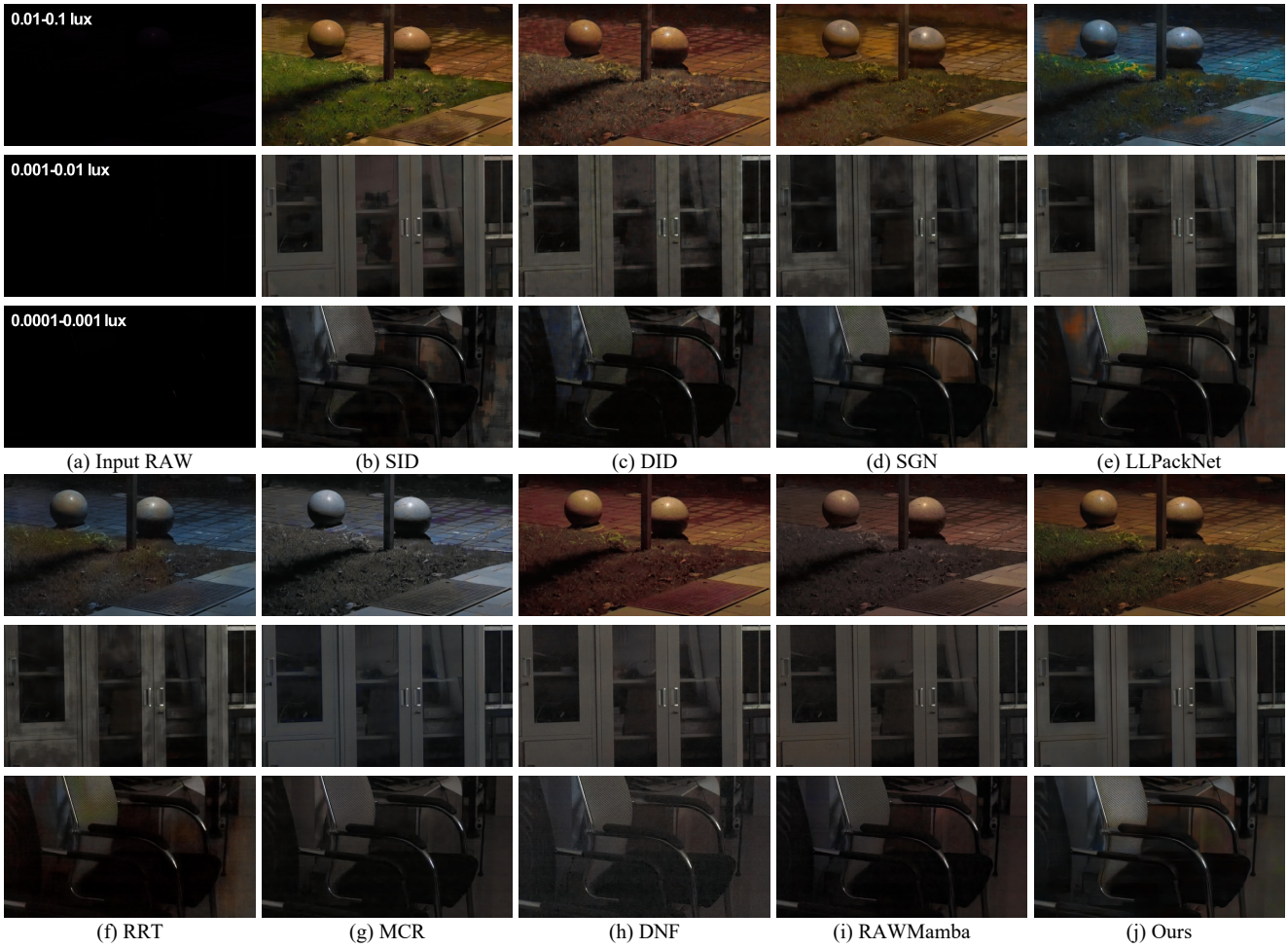


Figure S15. Qualitative comparison of our method and comparison methods [2–4, 7, 8, 10, 11, 14, 22] on the Sony subset of the SID [2] dataset. The input RAW images are visualized by a simple demosaicing process.



(a) Input RAW (b) SID (c) DID (d) SGN (e) LLPackNet

(f) RRT (g) MCR (h) DNF (i) RAWMamba (j) Ours

Figure S16. Qualitative comparison of our method and competitive methods [2–4, 7, 8, 10, 11, 14] on real-world extremely dark scenes.