Lighting Every Darkness in Two Pairs: A Calibration-Free Pipeline for RAW Denoising

Supplementary Materials

Xin Jin[1*] Jia-Wen Xiao[1*] Ling-Hao Han[1] Chunle Guo[1†]
Ruixun Zhang[2] Xialei Liu[1] Chongyi Li[1,3]
[1]VCIP, CS, Nankai University [2]Peking University [3]S-Lab, Nanyang Technological University
{xjin, xiaojw, lhhan}@mail.nankai.edu.cn, zhangruixun@pku.edu.cn,
{guochunle, xialei, lichongyi}@nankai.edu.cn
https://srameo.github.io/projects/led-iccv23

A. Network Architecture

We illustrate the detailed network architecture of the proposed LED in Fig. 1, a UNet-style [6] architecture with five stages. Both in encoder and decoder, each stage is consisted of two sequential RepNR blocks. It is worth noting that, except AINDNet [3], all other methods shared a same UNet architecture as SID [1]. Moreover, the LED would finally yield the same architecture for fair comparison after reparameterization (Sec. B).

B. Structural Reparameterization Process

In this section, we would detail the process of structural reparameterization. As stated in Sec. 3.4 in the main paper, the RepNR block consists of serial vs. parallel linear mapping, which can be fused to a single one. Specifically, the RepNR block can be transformed into a plain 3×3 convolution. Formally, this architecture contains two 3×3 convolutions with weights $\{W_0, W_1\}$ and bias $\{b_0, b_1\}$, and one of them follows a CSA layer. Let $CSA(x) = kx + b$, where k, b denote the weight and bias of it. Thus, the result for the input x can be represented as:

$$\begin{aligned}
\tilde{x} &= W_0(CSA(x)) + b_0 + W_1x + b_1 \\
&= W_0(kx + b) + b_0 + W_1x + b_1 \\
&= (W_0k + W_1)x + (W_0b + b_0 + b_1) \\
&= \tilde{W}x + \tilde{b},
\end{aligned} \tag{1}$$

where the whole deployment process is formulated. It demonstrates that our RepNR block can be transformed into a plain 3×3 convolution, and brings no extra costs during inference. It worth noting that we leveraged the online reparameterization strategy same as [2], thus there is no performance gap at all between training and testing.

C. More Visual Results

LED could better recover details compared with ELD [7] (calibration-based method) in Fig. 2. As shown in Fig. 3, LED outperforms other calibration-based methods [7, 8] in removing out-of-model noise. In Fig. 4 and Fig. 5-8, we provide more results on two benchmarks: ELD [7] and SID [1]. The restoration results of Kristina *et al.* [5], Noise2Noise [4], AINDNet [3], Zhang *et al.* [8] and ELD [7] are presented for comparison.

References

[1] Chen Chen, Qifeng Chen, Jia Xu, and Vladlen Koltun. Learning to see in the dark. In *CVPR*, 2018. 1, 2, 4, 5, 6, 7

[2] Mu Hu, Junyi Feng, Jiashen Hua, Baisheng Lai, Jianqiang Huang, Xiaojin Gong, and Xian-Sheng Hua. Online convolutional re-parameterization. In *CVPR*, 2022. 1

[3] Yoonsik Kim, Jae Woong Soh, Gu Yong Park, and Nam Ik Cho. Transfer learning from synthetic to real-noise denoising with adaptive instance normalization. In *CVPR*, 2020. 1, 4, 5, 6, 7

[4] Jaakko Lehtinen, Jacob Munkberg, Jon Hasselgren, Samuli Laine, Tero Karras, Miika Aittala, and Timo Aila. Noise2noise: Learning image restoration without clean data. *CVPR*, 2018. 1, 4, 5, 6, 7

[5] Kristina Monakhova, Stephan R Richter, Laura Waller, and Vladlen Koltun. Dancing under the stars: video denoising in starlight. In *CVPR*, 2022. 1, 4, 5, 6, 7

[6] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *MICCAI*, 2015. 1

[7] Kaixuan Wei, Ying Fu, Yinqiang Zheng, and Jiaolong Yang. Physics-based noise modeling for extreme low-light photography. 2021. 1, 2, 3, 4, 5, 6, 7

[8] Yi Zhang, Hongwei Qin, Xiaogang Wang, and Hongsheng Li. Rethinking noise synthesis and modeling in raw denoising. In *ICCV*, 2021. 1, 2, 4, 5, 6, 7

*Equal contribution.
†C. L. Guo is the corresponding author.

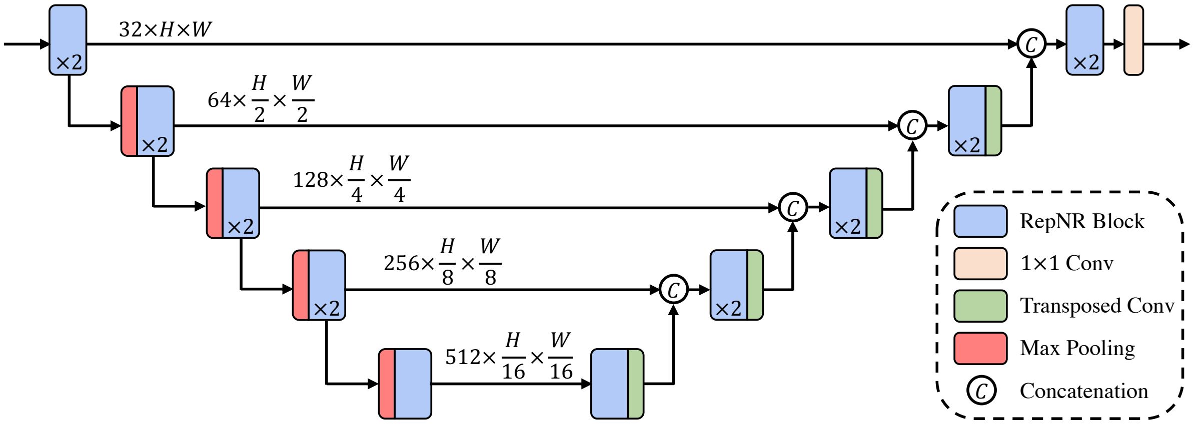
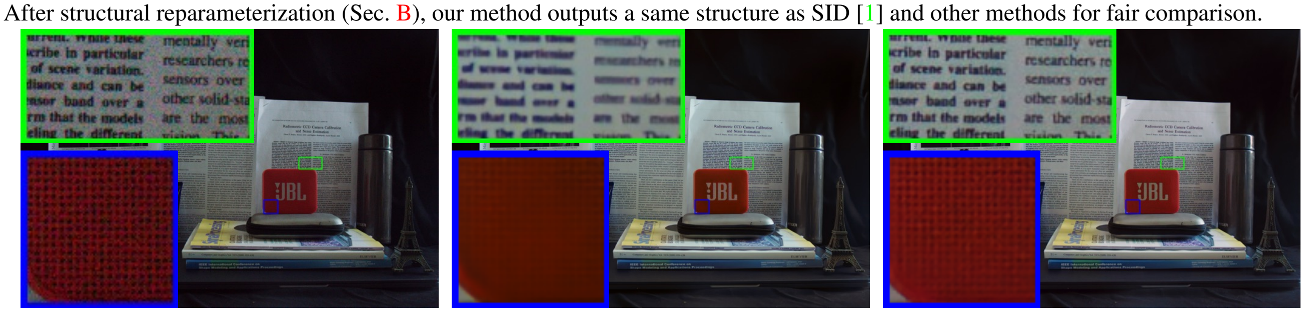


Figure 1. Detailed network architecture for our proposed LED. The $\hat{C} \times \hat{H} \times \hat{W}$ formatted expression on the arrow indicates the feature size for the corresponding stage. $H \times W$ is the input resolution. RepNR block with $\times 2$ denotes two RepNR blocks in a sequential way. After structural reparameterization (Sec. B), our method outputs a same structure as SID [1] and other methods for fair comparison.



(a) Input (b) ELD [7] (c) **LED (Ours)**

Figure 2. Proposed LED outperforms current state-of-the-art method in detail recovery significantly.

Input | Zhang *et al.* [8] | ELD [7] | **LED (Ours)** | GT

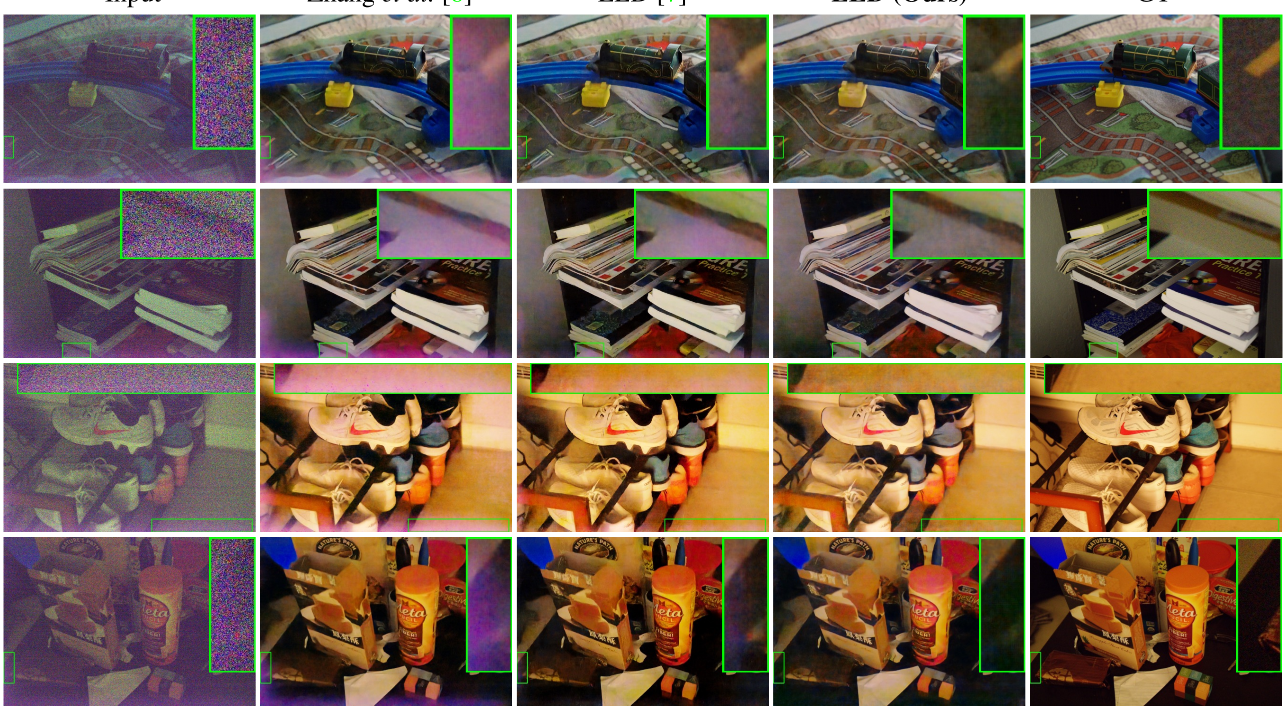


Figure 3. Compared with state-of-the-art calibration-based methods: ELD [7] and Zhang *et al.* [8], proposed LED is able to remove the out-of-model noise (*Zoom-in for best view*).

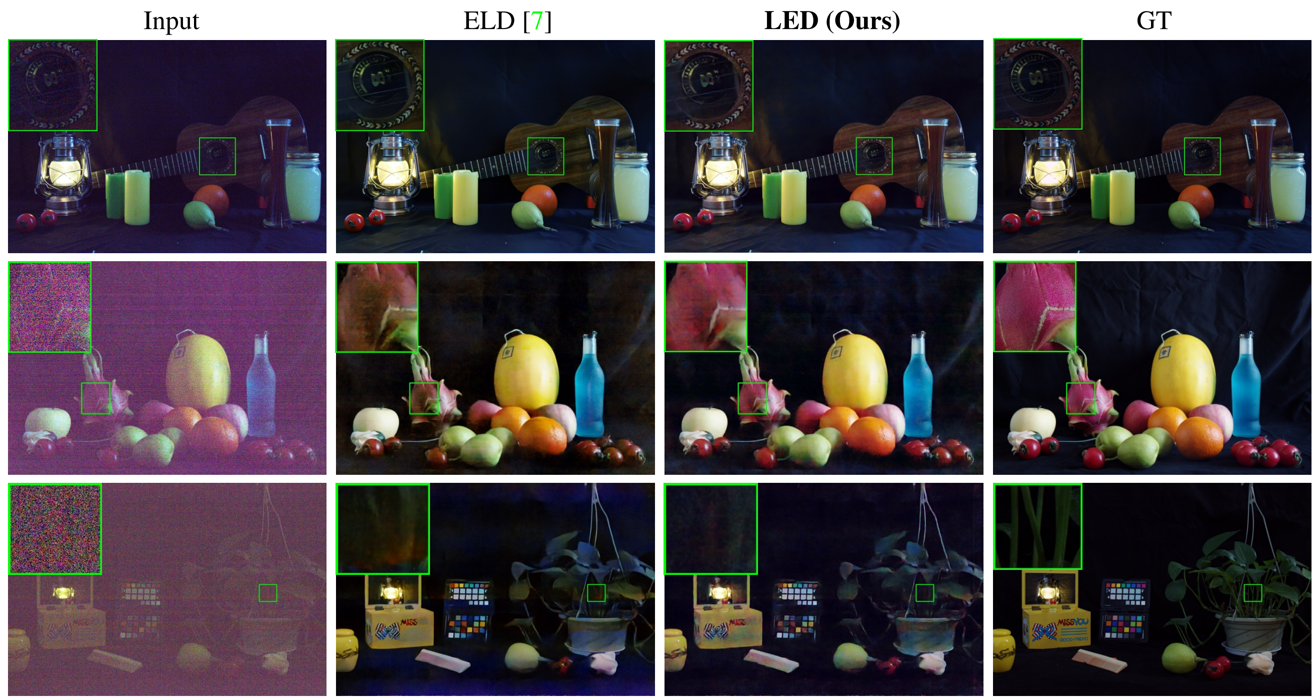


Figure 4. Visual comparison between our LED and other state-of-the-art methods on the ELD [7] dataset (*Zoom-in for best view*). We amplified and post-processed the input images with the same ISP as ELD [7].

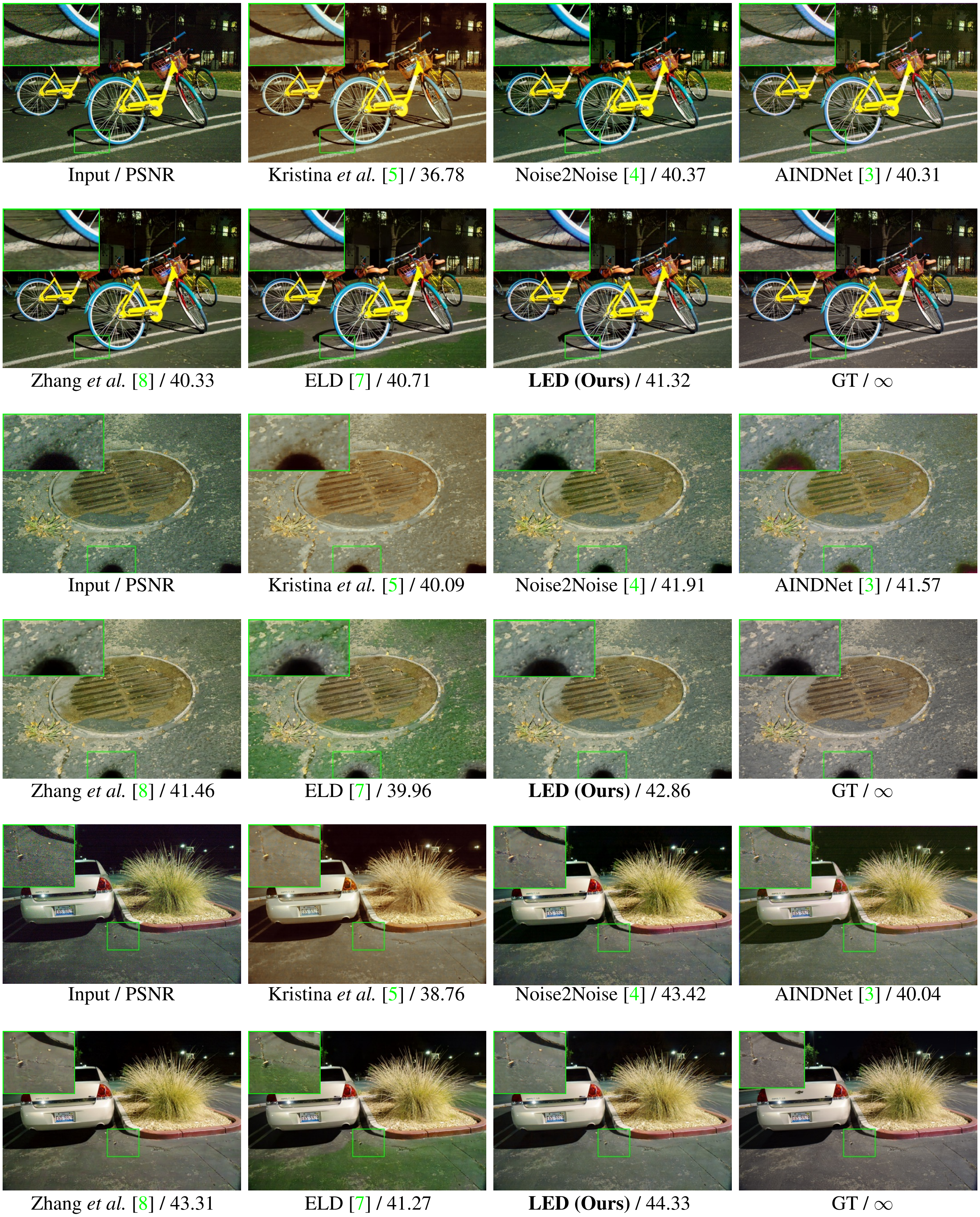


Figure 5. Visual comparison between our LED and other state-of-the-art methods on the SID [1] dataset (*Zoom-in for best view*). We amplified and post-processed the input images with the same ISP as ELD [7].

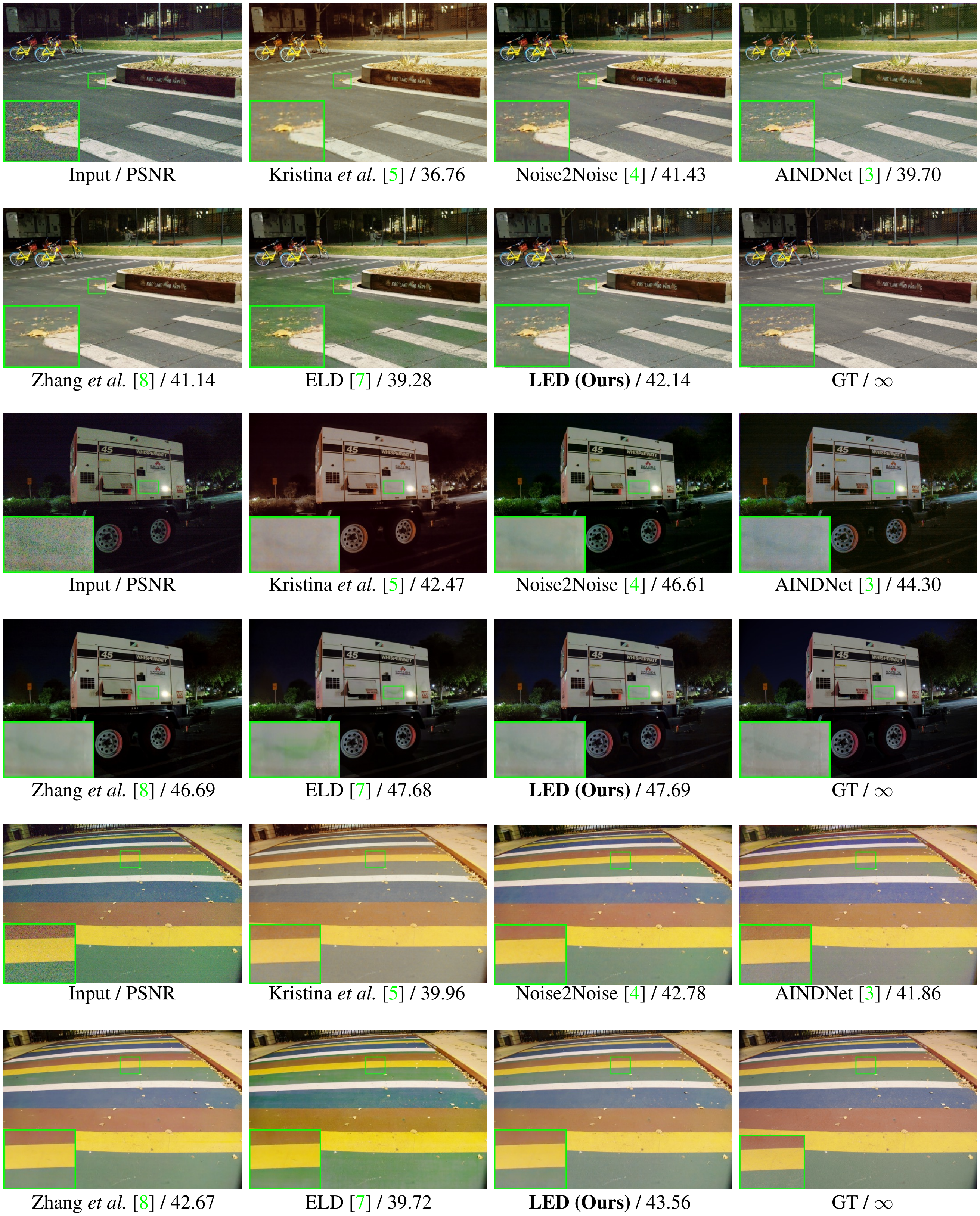


Figure 6. Visual comparison between our LED and other state-of-the-art methods on the SID [1] dataset (*Zoom-in for best view*). We amplified and post-processed the input images with the same ISP as ELD [7].

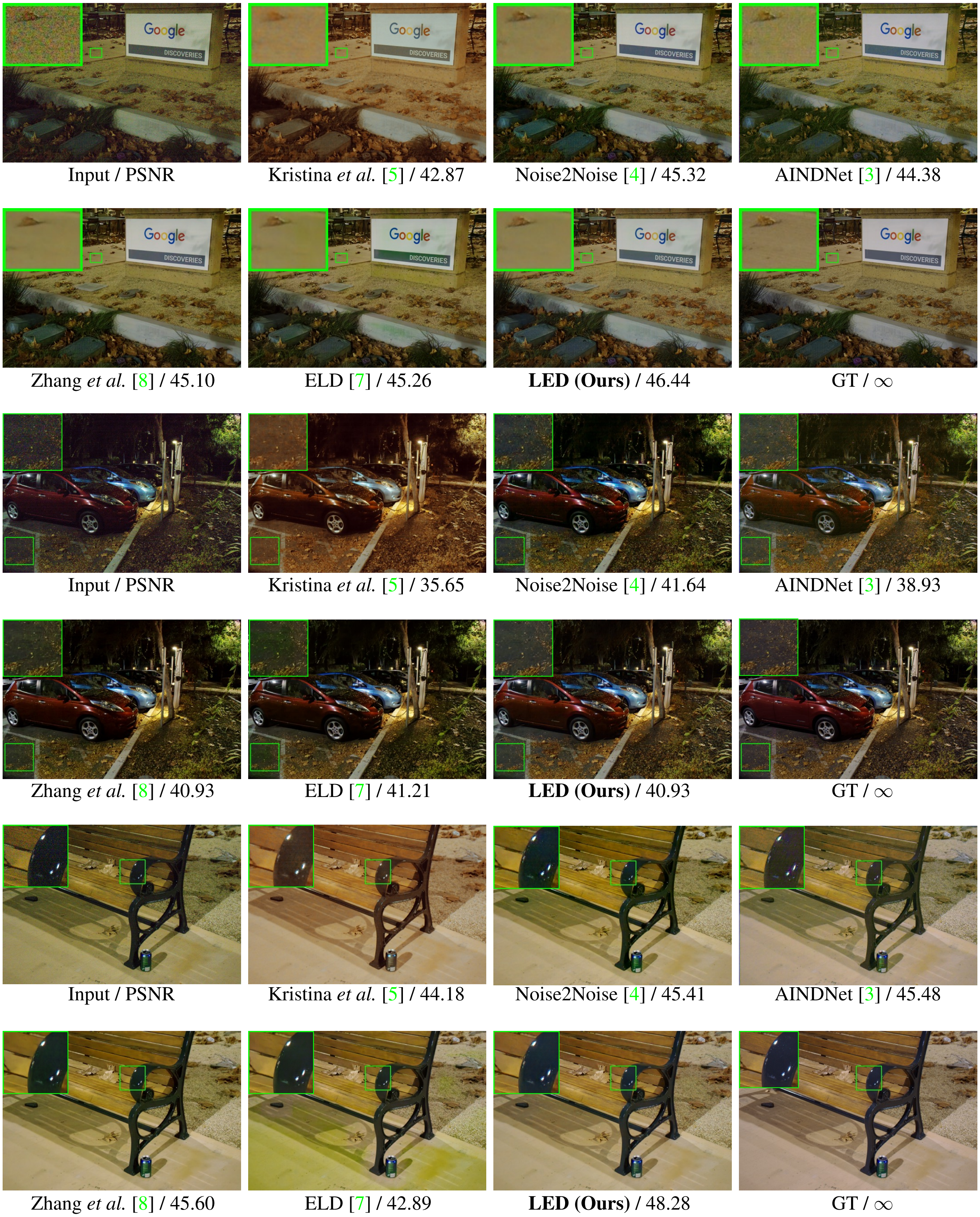


Figure 7. Visual comparison between our LED and other state-of-the-art methods on the SID [1] dataset (*Zoom-in for best view*). We amplified and post-processed the input images with the same ISP as ELD [7].

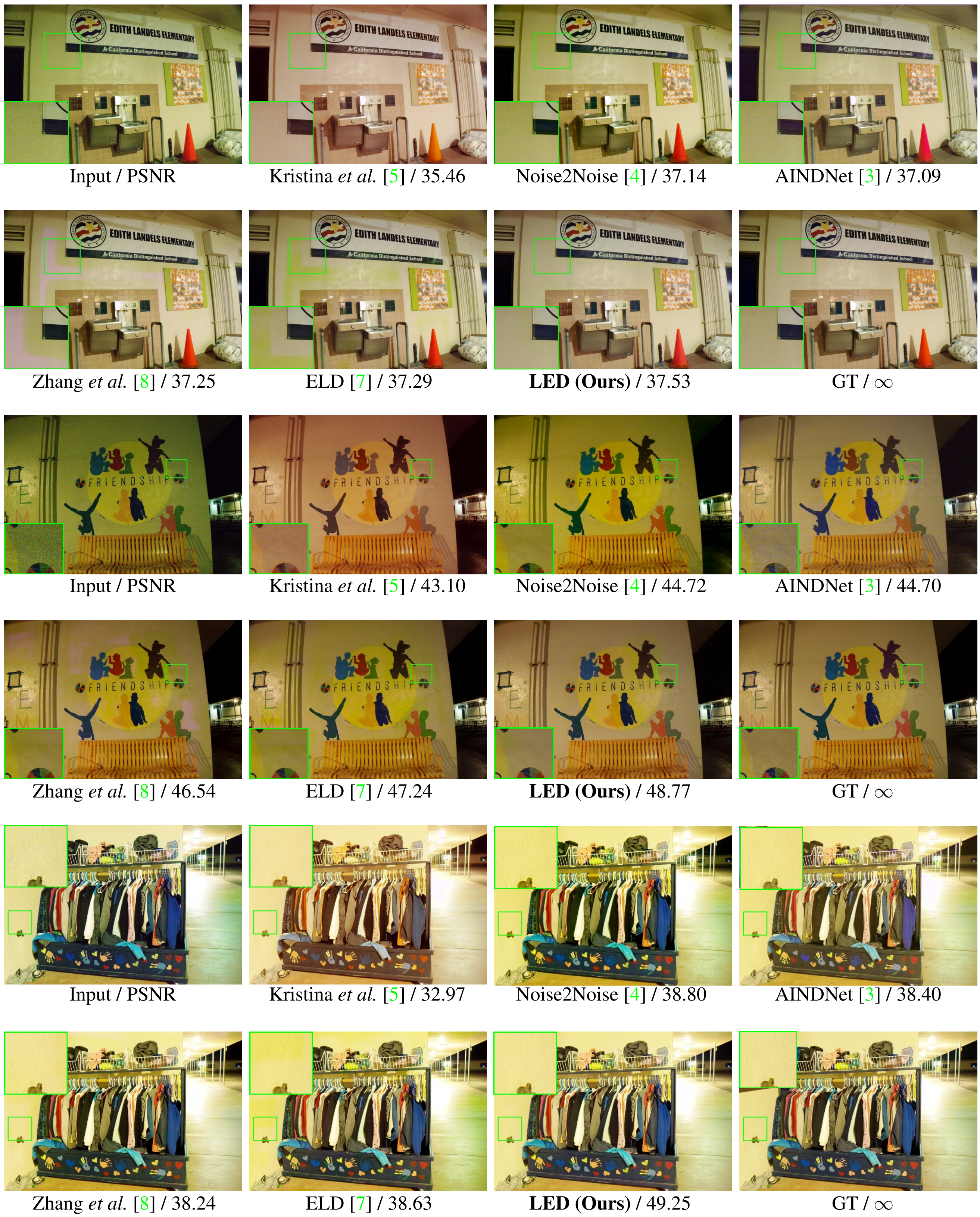


Figure 8. Visual comparison between our LED and other state-of-the-art methods on the SID [1] dataset (*Zoom-in for best view*). We amplified and post-processed the input images with the same ISP as ELD [7].