Supplemental Material for LIVENet: A novel network for real-world low-light image denoising and enhancement

Dhruv Makwana
Independent Researcher
dmakwana503@gmail.com

Gayatri Deshmukh
Independent Researcher
dgayatri9850@gmail.com

Onkar Susladkar
Independent Researcher
onkarsus13@gmail.com

Sparsh Mittal
IIT Roorkee
sparsh.mittal@mfs.iitr.ac.in

Sai Chandra Teja R
Green PMU Semi Pvt Ltd
saichandrateja@gmail.com

SUPPLEMENTARY CONTENT

In this supplementary material, we present the following:

- We discuss additional related work based on histogram equalization (HE)-based techniques and retinex-based techniques (S.1).
- We describe the processes for creating coarse map ground truth (S.2).
- We present the details of datasets used for training and evaluations (S.3).
- We discuss the details of experimental platform and evaluation metrics (S.4).
- We provide additional qualitative results on ablation study of our proposed method (S.5).
- We present additional qualitative results on various datasets (S.6).

S.1. Additional related works

HE-based methods: Lee et. al. [10] propose contrast enhancement algorithm based on layered difference representation. They enhance image contrast by amplifying the gray level differences between adjacent pixels. Celik et. al. [2] use interpixel contextual information to enhance the image. They use a 2D histogram of the input image constructed using a mutual relationship between each pixel and its neighboring pixels. Wang et. al [19] introduce the probability distribution function of an image, which is modified by weighting and thresholding before the histogram equalization to enhance the quality of the image. Wadud et. al. [1] partition the image histogram based on local minima and assign specific gray level ranges for each partition before equalizing them separately. Haidi et. al. [7] propose brightness-preserving dynamic histogram equalization (BPDHE). It can produce the output image with a mean intensity almost equal to the mean intensity of the input, thus fulfilling the requirement of maintaining the mean brightness of the image.

Retinex-based Methods: Early methods [8, 16] focused on color restoration at the cost of color consistency. Guo et. al. [6] propose a simple method for low-light image enhancement where the illumination of each pixel is first estimated individually by finding the maximum value in the R, G, and B channels. Li. et. al. [12] propose a Retinex model that additionally considers a noise map to deal with images with excessive noise. Ren et. al. [17] propose Low-Rank Regularized Retinex Model (LR3M), which injects the low-rank prior into a Retinex decomposition process to suppress noise in the reflectance map. Fu et. al. [3] propose a model which can preserve the reflectance layer with more refined details with the help of a bright channel prior.

S.2. coarse map

The steps for generating a transmission map, an atmospheric light map, and coarse map ground truth are as follows:

1. First, we use the min filter and the max filter to get the dark channel prior and the bright channel prior.
2. Next, we compute global atmospheric light using a bright channel prior.
3. Following that, we use the bright channel to get an initial transmission estimate.
4. The dark channel is then used as a supplement to adjust any incorrect transmission estimates derived from the bright channel prior earlier.
5. Finally, we obtain an enhanced image coarse map using the global atmospheric light and transmission map by feeding them into the atmospheric scattering model.

Step-1: Dark Channel Prior (DCP) represents the minimum intensity of pixels in the local patch. It can be expressed as

$$I_{dark}(x) = \min_{c\in[r,g,b]}\left(\min_{y\in\Omega(x)}(I^c(y))\right) \tag{1}$$

Where I^c is the color channel of image I, $\Omega(x)$ is a local patch centered at x, and y is the pixel in the local patch of $\Omega(x)$. Similarly, Bright Channel Prior (BCP) is the maximum intensity of pixels in the local patch. It can be expressed as

$$I_{bright}(x) = \max_{c\in[r,g,b]}\left(\max_{y\in\Omega(x)}(I^c(y))\right) \tag{2}$$

The DCP and BCP for given input low-light image are depicted in Figure S.1

(a) Input Image (b) DCP (c) BCP

Figure S.1. (a) Input low-light image, (b) dark channel prior (DCP) for low-light image, (c) bright channel prior (BCP) for the low-light image.

Step-2: For global atmospheric light, we use the top 10% of BCP. The BCP approximates the dark denseness of the atmosphere of the input image very well; therefore, the BCP is used to improve the atmospheric light estimation. This 10% of values is used to verify that a little abnormality does not significantly impact lighting.

Step-3: According to the BCP, the bright channel I_{bright} of a well-illuminated radiance (I) should be 255. The initial transmission map can be expressed using BCP based on the following equation.

$$t_{bright}(x) = \frac{I_{bright}(x) - \mathcal{A}^c}{255 - \mathcal{A}^c} \tag{3}$$

The BCP considers atmospheric light colorless and attempts to filter out this color by measuring it from the BCP. This transmission map fails with a patch containing bright objects like lamp lights. To address the concerns raised above, we use the dark channel in the next step as a supplement to rectify possibly incorrect transmission estimations derived from the BCP earlier.

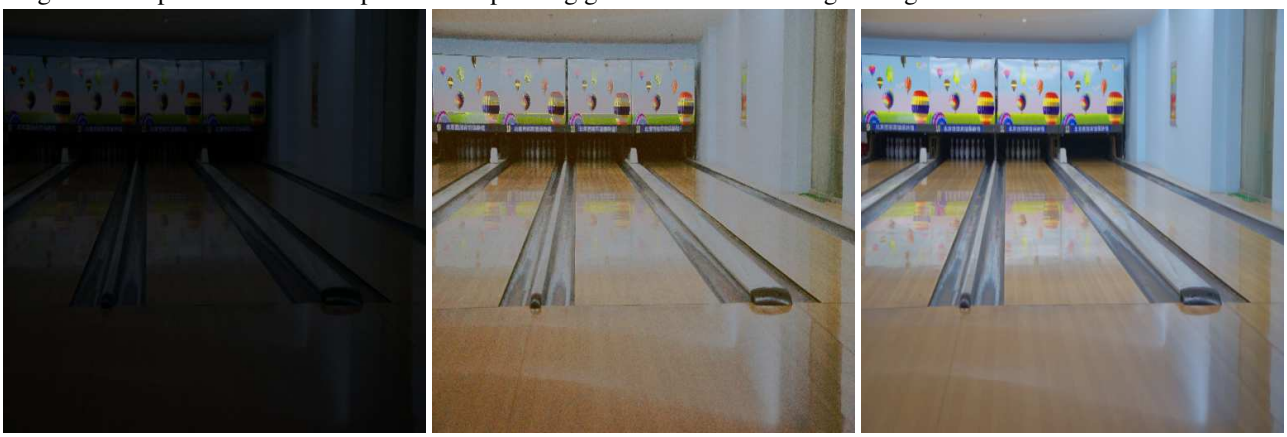
Step-4: The DCP is also used to generate a transmission map, and the difference between the DCP and BCP is computed. This computation is performed to adjust possibly incorrect transmission estimations derived from the BCP in the previous step. Any pixel x with a difference smaller than the specified value of α (0.6, established empirically) is in a dark object, making its depth untrustworthy. This renders pixel x transmission unstable. As a result, by using the transmission maps' product, the faulty transmission may be remedied.

Step-5: The derived transmission map $t(x)$ and atmospheric light $\mathcal{A}$ can be used to recover the scene radiance using the atmospheric scattering model using the following equation:

$$J(x) = \frac{I(x) - \mathcal{A}}{max(t(x), t_0)} + \mathcal{A} \tag{4}$$

When the transmission $t(x)$ is near zero, the term $\frac{I(x)-\mathcal{A}}{t(x)}$ has a very high value, and the immediately recovered scene radiance J will contain a high amount of noise. To overcome this issue, we limit the transmission $t(x)$ value to a lower constraint value t_0. We use the value of t_0 as 0.15, which was determined by experimental analysis.

Figure S.2 depicts the coarse map and corresponding ground-truth normal-light image.



(a) Input Image (b) coarse map (c) normal-light GT

Figure S.2. (a) Input low-light image, (b) Coarse map generated from low-light and normal-light image, (c) Ground truth normal-light image

S.3. Dataset

We use LOL-v1 [23] and LOL-v2-real [25] paired dataset for training. Both the v1 and v2 versions of LOL include noticeable noise. The LOL-v1 contains 485 low/normal image pairs for training and 15 pairs for testing. Each pair has a low-light input image and a well-exposed reference image. LOL-v2 is split into two categories: real and synthetic. LOL-v2-real includes 689 pairs of low/normal-light image pairs for training and 100 pairs for testing. Most low-light images were captured by varying the exposure duration and ISO while keeping the other camera settings constant. We use DICM [10], LIME [5], MEF [13] NPE [20], and VV [18] datasets for testing efficacy of LLIE techniques.

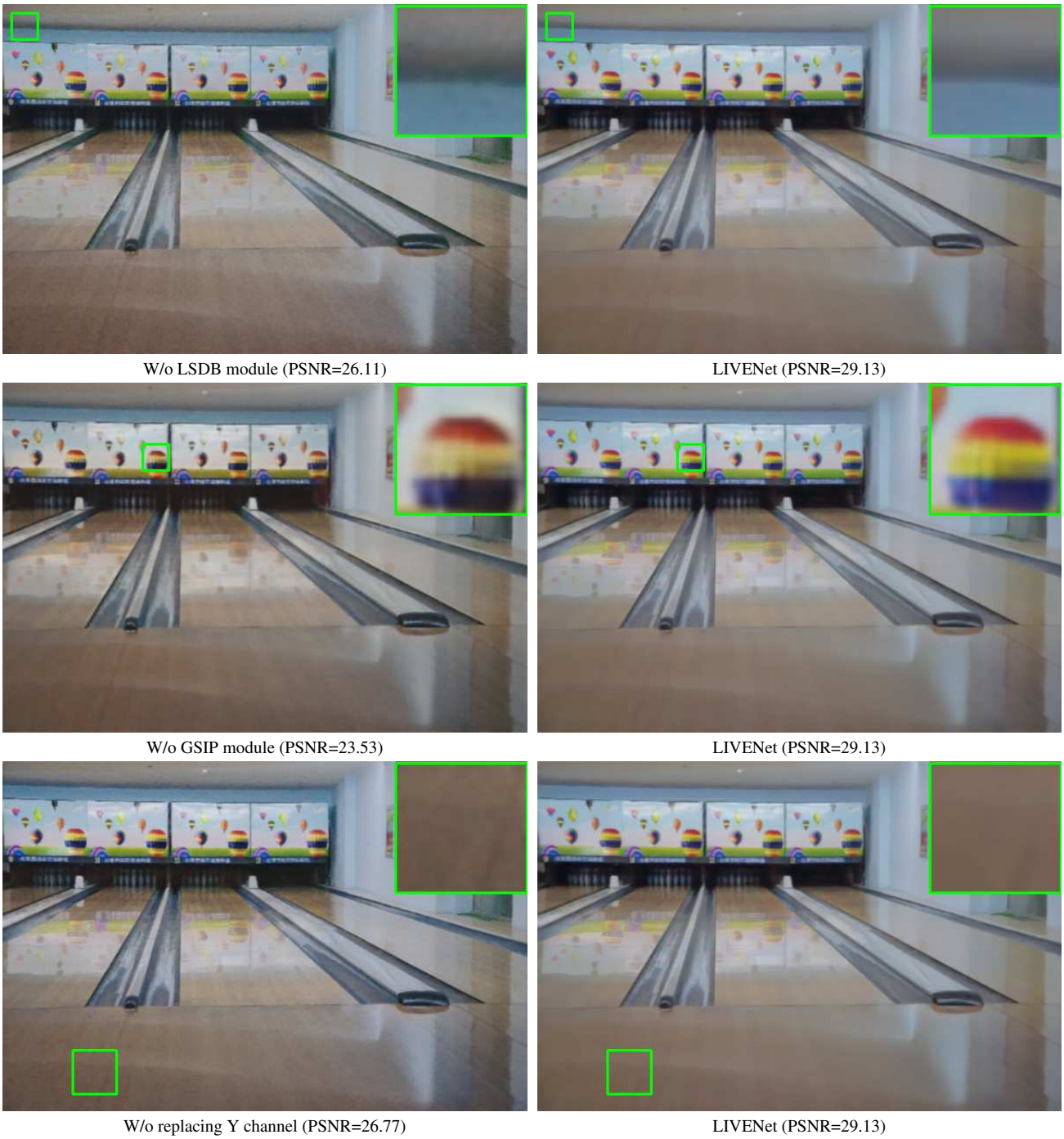
S.4. Implementation platform and metrics

All the experiments are carried out using Pytorch [15] framework with an NVIDIA A100 GPU. All the training and testing images are resized to 512×512 pixels. We use the Adam optimizer [9] with a batch size of 16. We train LIVENet for 3.1×10^4 number of iterations with a learning rate of 3×10^{-4}. We employ both full-reference and non-reference image quality assessment measures to compare the performance of various LLIE algorithms. Peak signal-to-noise ratio (PSNR), structural similarity (SSIM) [22], mean absolute error (MAE), and Learned Perceptual Image Patch Similarity (LPIPS) [26] are used for LOL-v1, and LOL-v2 paired testing data. NIQE [20] is used for DICM, LIME, MEF, NPE, and VV unpaired datasets. Note that the higher the PSNR and SSIM, the better the enhancement. The lower the MAE, LPIPS, and NIQE, the more realistic the refined images. We compare against the following nine techniques: ZeroDCE [4], ZeroDCE++ [11], SCI [14], RetinexNet [23], KinD [28], KinD++ [27], SNR [24], DCC-Net [29], and LLFlow [21]. We assess all networks on the same train and test data set, with the same loss and hyperparameters.

S.5. Additional Ablation Results

Figure S.3 depicts the qualitative outcomes of several LIVENet components. We also compare the PSNR measure between the entire LIVENet network and LIVENet without the corresponding module to highlight the impact on noise removal. The green rectangle magnifies the small area of the image to highlight finer details. In Figure S.3, the first figure on the left shows

that on removing the LSDB module, the PSNR reduces to 26.11. This shows that LSDB plays an important role in noise reduction from the low-light image. In Figure S.3, the second figure on the left shows that on removing the GSIP module, the output image contains more noise and poor texture; this is confirmed by the fact that the PSNR score without GSIP is just 23.53. In Figure S.3, the third figure on the left shows that when the Y channel is not replaced with a noise-free grayscale image, there is a small amount of noise on the floor.



W/o LSDB module (PSNR=26.11) LIVENet (PSNR=29.13)

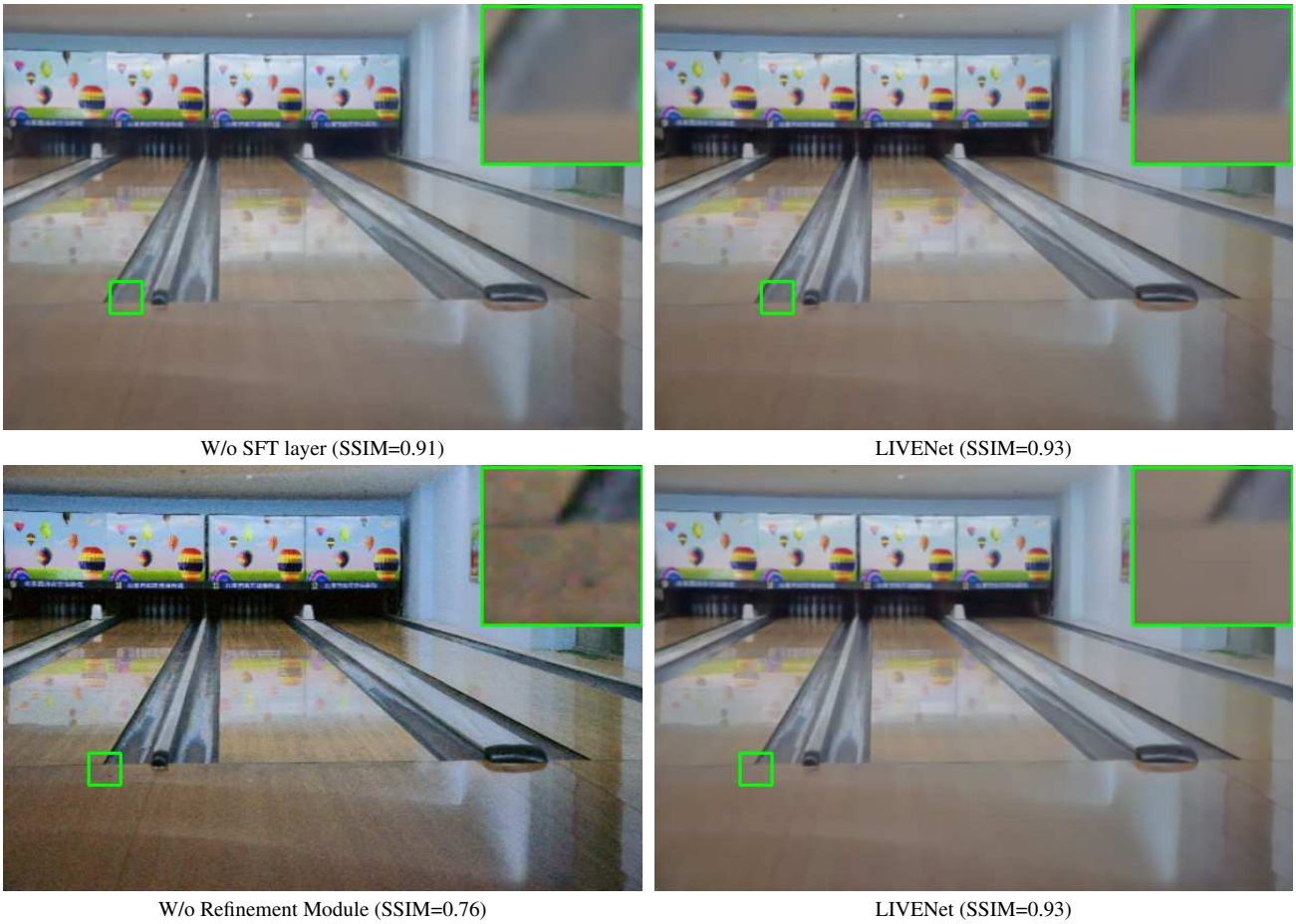
W/o GSIP module (PSNR=23.53) LIVENet (PSNR=29.13)

W/o replacing Y channel (PSNR=26.77) LIVENet (PSNR=29.13)

Figure S.3. Qualitative results of ablation study: The noise removal effect of different components in LIVENet. Note that the three images on the right-hand side are all the same images, but the magnified areas are different.

The influence of LIVENet components on texture enhancement is shown in Figure S.4. We show the results without the second stage (Refinement Module) and the SFT layer. The SSIM of LIVENet is reduced from 0.93 to 0.91 by not utilizing the

SFT layer. Additionally, the figure illustrates that a two-stage LIVENet network can recover texture information efficiently since the SSIM score without the second stage (W/o Refinement Module) is only 0.76.



W/o SFT layer (SSIM=0.91) LIVENet (SSIM=0.93)

W/o Refinement Module (SSIM=0.76) LIVENet (SSIM=0.93)

Figure S.4. The texture improvement effect of different components for the same network. Note that the two images on the right-hand side are both the same images, but the magnified areas are different.

S.6. Additional Visual Results

Our main paper had adequate quantitative (Table 1-2) and qualitative results (Figure. 6-8). Here, we provide more qualitative analysis by comparing against all SOTA methods, including ZeroDCE [4], ZeroDCE++ [11], SCI [14], RetinexNet [23], KinD [28], KinD++ [27], SNR [24], DCC-Net [29], and LLFlow [21]. Figure S.5 shows the qualitative comparison on the LOL-v1 [23] dataset. Figure S.6 shows the qualitative comparison on the LOL-v2-real [25] dataset. Both the LOL-v1 and LOL-v2-real datasets are paired, and GT symbolizes the Ground Truth.

Figure S.7 shows the visual results of the DICM [10] dataset. Figure S.8 presents the LIME [5] dataset results. Figure S.9 shows the results on the MEF [13] dataset. Figure S.10 and Figure S.11 presents the outcomes of NPE [20], and VV [18] datasets, respectively. It can be observed that, on several benchmarks, LIVENet outperforms the state-of-the-art techniques.

References

[1] M. Abdullah-Al-Wadud, Md. Hasanul Kabir, M. Ali Akber Dewan, and Oksam Chae. A dynamic histogram equalization for image contrast enhancement. *IEEE Transactions on Consumer Electronics*, 53(2):593–600, 2007. 1

[2] Turgay Celik and Tardi Tjahjadi. Contextual and variational contrast enhancement. *IEEE Transactions on Image Processing*, 20(12):3431–3441, 2011. 1

[3] Gang Fu, Lian Duan, and Chunxia Xiao. A hybrid l2 -lp variational model for single low-light image enhancement with bright channel prior. In *2019 IEEE International Conference on Image Processing (ICIP)*, pages 1925–1929, 2019. 1

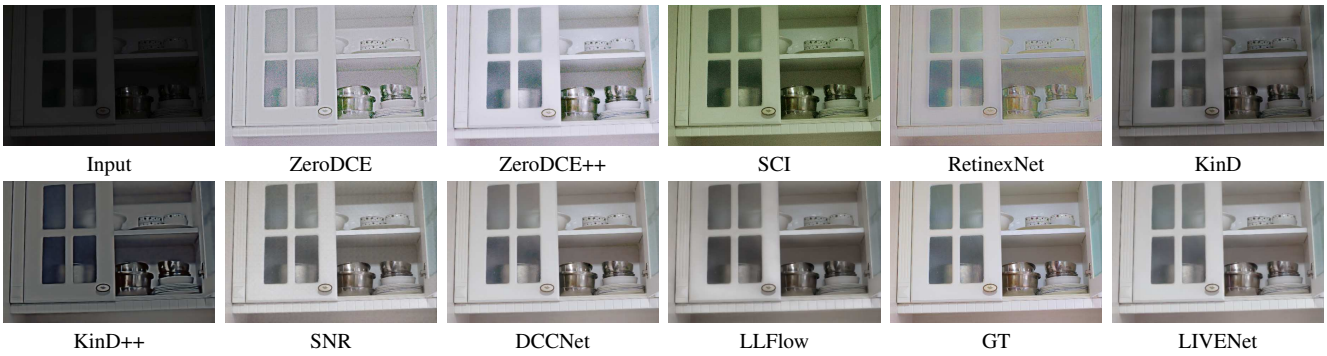


Figure S.5. Results on LOL-v1 dataset.

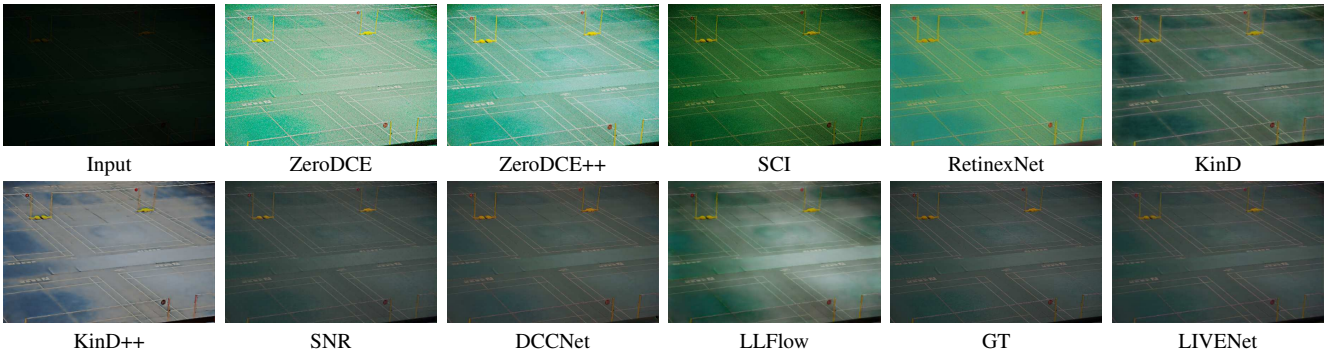


Figure S.6. Results on LOL-v2-real dataset.

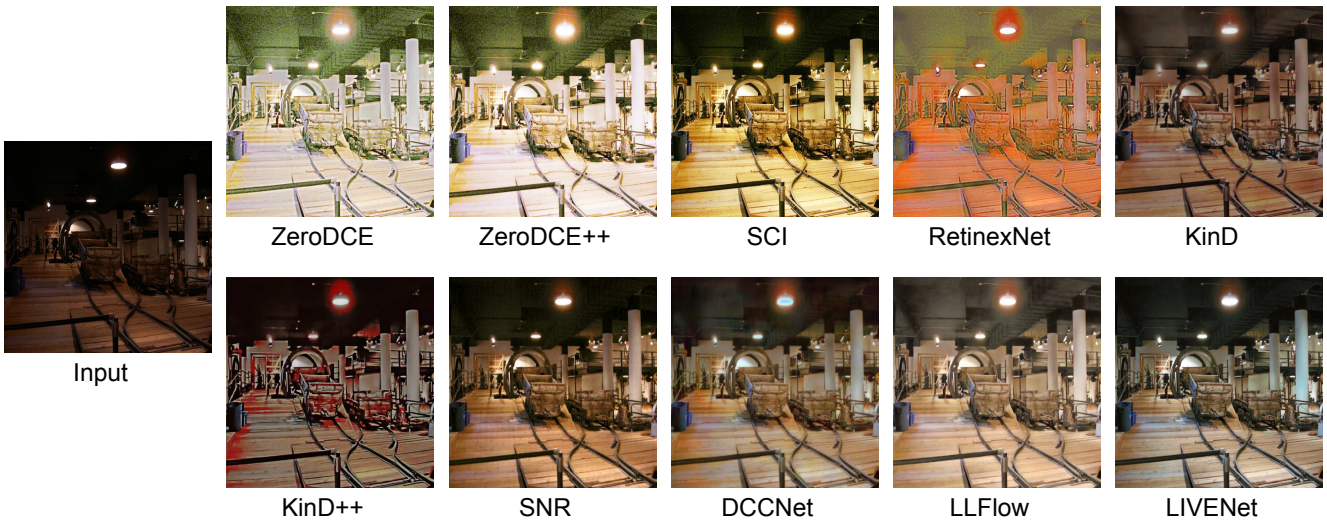


Figure S.7. Results on DICM dataset.

[4] Chunle Guo, Chongyi Li, Jichang Guo, Chen Change Loy, Junhui Hou, Sam Kwong, and Runmin Cong. Zero-reference deep curve estimation for low-light image enhancement. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 1780–1789, 2020. 3, 5

[5] Xiaojie Guo, Yu Li, and Haibin Ling. Lime: Low-light image enhancement via illumination map estimation. *IEEE Transactions on image processing*, 26(2):982–993, 2016. 3, 5

[6] Xiaojie Guo, Yu Li, and Haibin Ling. Lime: Low-light image enhancement via illumination map estimation. *IEEE Transactions on Image Processing*, 26(2):982–993, 2017. 1

[7] Haidi Ibrahim and Nicholas Sia Pik Kong. Brightness preserving dynamic histogram equalization for image contrast enhancement. *IEEE Transactions on Consumer Electronics*, 53(4):1752–1758, 2007. 1

Figure S.8. Results on LIME dataset.

Figure S.9. Results on MEF dataset.

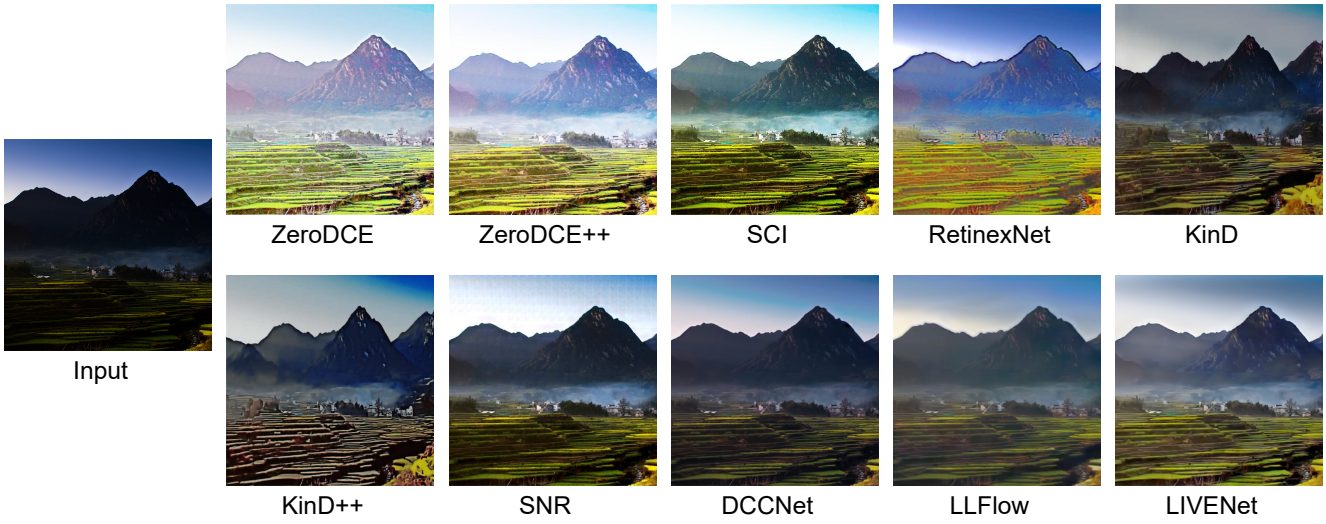


Figure S.10. Results on NPE dataset.

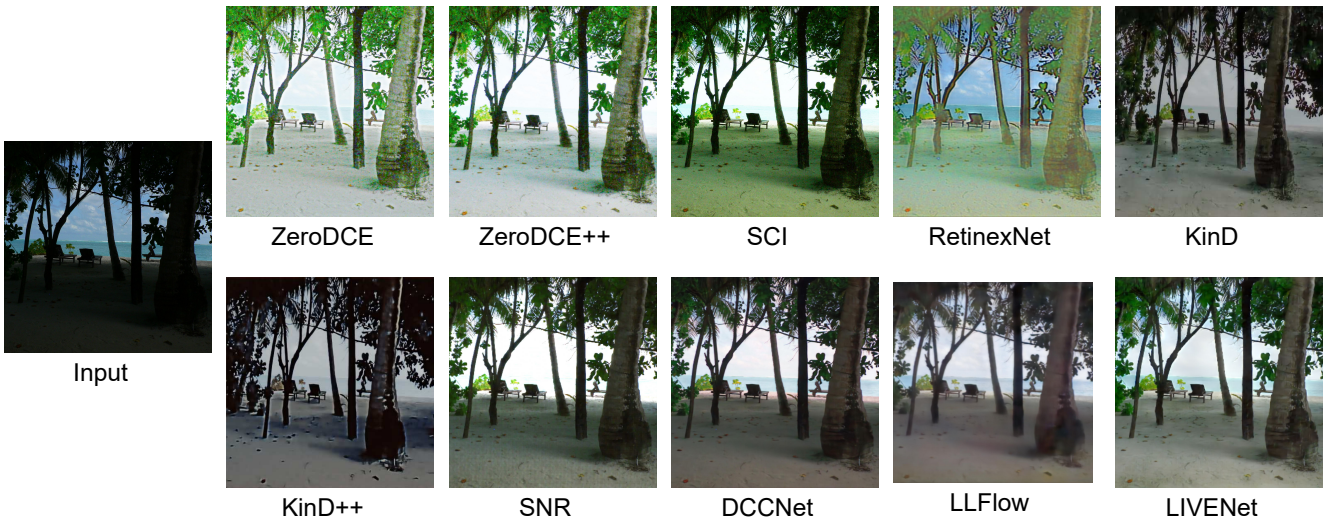


Figure S.11. Results on VV dataset.

[8] Daniel J Jobson, Zia-ur Rahman, and Glenn A Woodell. A multiscale retinex for bridging the gap between color images and the human observation of scenes. *IEEE Transactions on Image processing*, 6(7):965–976, 1997. 1

[9] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014. 3

[10] Chulwoo Lee, Chul Lee, and Chang-Su Kim. Contrast enhancement based on layered difference representation. In *2012 19th IEEE international conference on image processing*, pages 965–968. IEEE, 2012. 1, 3, 5

[11] Chongyi Li, Chunle Guo, and Chen Change Loy. Learning to enhance low-light image via zero-reference deep curve estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(8):4225–4238, 2021. 3, 5

[12] Mading Li, Jiaying Liu, Wenhan Yang, Xiaoyan Sun, and Zongming Guo. Structure-revealing low-light image enhancement via robust retinex model. *IEEE Transactions on Image Processing*, 27(6):2828–2841, 2018. 1

[13] Kede Ma, Kai Zeng, and Zhou Wang. Perceptual quality assessment for multi-exposure image fusion. *IEEE Transactions on Image Processing*, 24(11):3345–3356, 2015. 3, 5

[14] Long Ma, Tengyu Ma, Risheng Liu, Xin Fan, and Zhongxuan Luo. Toward fast, flexible, and robust low-light image enhancement. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 5637–5646, 2022. 3, 5

[15] Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, et al. Pytorch: An imperative style, high-performance deep learning library. *Advances in neural information processing systems*, 32, 2019. 3

[16] Zia-ur Rahman, Daniel J Jobson, and Glenn A Woodell. Multi-scale retinex for color image enhancement. In *Proceedings of 3rd IEEE international conference on image processing*, volume 3, pages 1003–1006. IEEE, 1996. 1

[17] Xutong Ren, Wenhan Yang, Wen-Huang Cheng, and Jiaying Liu. Lr3m: Robust low-light enhancement via low-rank regularized retinex model. *IEEE Transactions on Image Processing*, 29:5862–5876, 2020. 1

[18] Vassilios Vonikakis, Dimitris Chrysostomou, Rigas Kouskouridas, and Antonios Gasteratos. Improving the robustness in feature detection by local contrast enhancement. In *2012 IEEE International Conference on Imaging Systems and Techniques Proceedings*, pages 158–163. IEEE, 2012. 3, 5

[19] Qing Wang and Rabab K. Ward. Fast image/video contrast enhancement based on weighted thresholded histogram equalization. *IEEE Transactions on Consumer Electronics*, 53(2):757–764, 2007. 1

[20] Shuhang Wang, Jin Zheng, Hai-Miao Hu, and Bo Li. Naturalness preserved enhancement algorithm for non-uniform illumination images. *IEEE transactions on image processing*, 22(9):3538–3548, 2013. 3, 5

[21] Yufei Wang, Renjie Wan, Wenhan Yang, Haoliang Li, Lap-Pui Chau, and Alex Kot. Low-light image enhancement with normalizing flow. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 36, pages 2604–2612, 2022. 3, 5

[22] Zhou Wang, Alan C Bovik, Hamid R Sheikh, and Eero P Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE transactions on image processing*, 13(4):600–612, 2004. 3

[23] Chen Wei, Wenjing Wang, Wenhan Yang, and Jiaying Liu. Deep retinex decomposition for low-light enhancement. *British Machine Vision Conference*, 2018. 3, 5

[24] Xiaogang Xu, Ruixing Wang, Chi-Wing Fu, and Jiaya Jia. Snr-aware low-light image enhancement. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 17714–17724, 2022. 3, 5

[25] Wenhan Yang, Wenjing Wang, Haofeng Huang, Shiqi Wang, and Jiaying Liu. Sparse gradient regularized deep retinex network for robust low-light image enhancement. *IEEE Transactions on Image Processing*, 30:2072–2086, 2021. 3, 5

[26] Richard Zhang, Phillip Isola, Alexei A Efros, Eli Shechtman, and Oliver Wang. The unreasonable effectiveness of deep features as a perceptual metric. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 586–595, 2018. 3

[27] Yonghua Zhang, Xiaojie Guo, Jiayi Ma, Wei Liu, and Jiawan Zhang. Beyond brightening low-light images. *International Journal of Computer Vision*, 129:1013–1037, 2021. 3, 5

[28] Yonghua Zhang, Jiawan Zhang, and Xiaojie Guo. Kindling the darkness: A practical low-light image enhancer. In *Proceedings of the 27th ACM international conference on multimedia*, pages 1632–1640, 2019. 3, 5

[29] Zhao Zhang, Huan Zheng, Richang Hong, Mingliang Xu, Shuicheng Yan, and Meng Wang. Deep color consistent network for low-light image enhancement. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 1899–1908, 2022. 3, 5