

Zero-Reference Learning for Low-Light Image Enhancement (Supplementary Material)

Chunle Guo^{1,2*} Chongyi Li^{1,2*} Jichang Guo^{1†}
Chen Change Loy³ Junhui Hou² Sam Kwong² Runmin Cong⁴

¹ BIIT Lab, Tianjin University ² City University of Hong Kong ³ Nanyang Technological University ⁴ Beijing Jiaotong University

{guochunle, lichongyi, jcguo}@tju.edu.cn ccloy@ntu.edu.sg

{jh.hou, cssamk}@cityu.edu.hk rmcong@bjtu.edu.cn

https://li-chongyi.github.io/Proj_Zero-DCE.html/

Abstract

This supplementary material provides detailed network architecture and parameter settings of our Deep Curve Estimation Network (DCE-Net), an extra ablation study (the advantage of three-channel adjustment), more ablation study results, visual comparisons and face detection results.

1. Detailed Architecture of DCE-Net

In Fig. 1, we present the detailed network architecture and parameter settings of our Deep Curve Estimation Network (DCE-Net). As shown, the DCE-Net contains seven convolutional layers with symmetrical skip-connection. In the first six convolutional layers, each convolutional layer consists of 32 convolutional kernels of size 3×3 and stride 1 followed by the ReLU activation function. The last convolutional layer consists of 24 convolutional kernels of size 3×3 and stride 1 followed by the Tanh activation function, which produces 24 curve parameter maps for eight iterations, where each iteration requires three curve parameter maps for the three channels (*i.e.*, RGB channels).

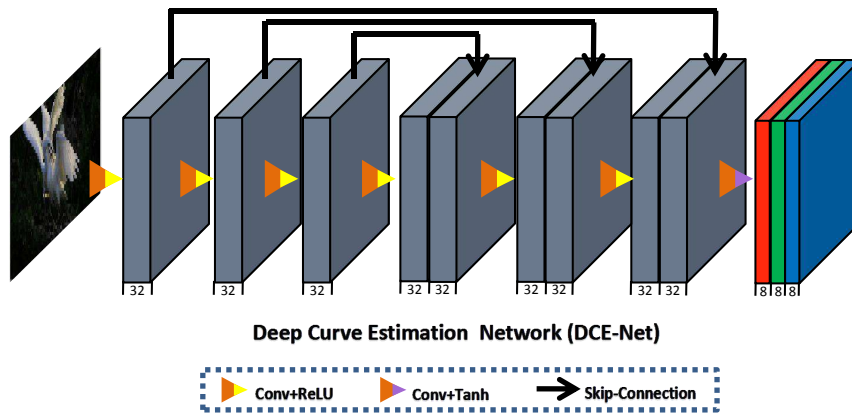


Figure 1: The architecture of Deep Curve Estimation Network (DCE-Net).

*The first two authors contribute equally to this work.

†Jichang Guo (jcguo@tju.edu.cn) is the corresponding author.

2. Ablation Study

In this section, we report an extra ablation study, *i.e.*, **the advantage of three-channel adjustment**. To investigate the advantage of RGB three-channel adjustment, we conduct experiments to only adjust the illumination related channels in CIE Lab and YCbCr color spaces by using the same configurations as the adjustment in RGB color space, except removing the color constancy loss that is only available for three-channel adjustment.

Specifically, we first transfer the input low-light image from RGB color space to CIE Lab (resp. YCbCr) color space, then feed the L (resp. Y) component to the DCE-Net for estimating a set of curve parameter maps. We compute each loss in L (resp. Y) channel during the training. At last, we adjust the L (resp. Y) component by pixel-wise curve mapping with the estimated curve parameter maps. After the adjustment of L (resp. Y) component, the corresponding ab (resp. CbCr) components are adjusted accordingly (equal proportion adjustment). In Fig. 3, we show an example to show the advantage of three-channel adjustment.

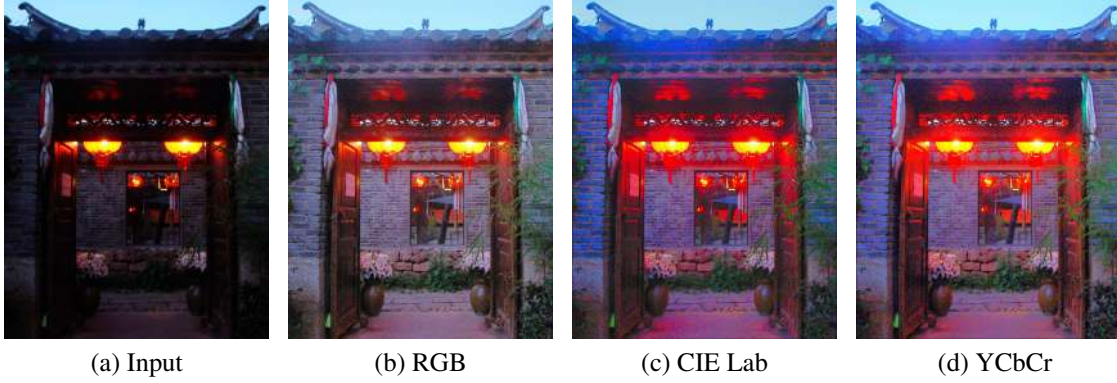


Figure 2: Ablation study of the advantage of three-channel adjustment (RGB, CIE Lab, YCbCr color spaces).

As observed, all the results have improved brightness and contrast; however the results adjusted in CIE Lab and YCbCr color spaces as shown in Fig. 3(c) and (d) have obvious color deviations (*e.g.*, the color of wall) and over-saturation (*e.g.*, the regions of lantern). The visual comparisons imply that three-channel adjustment can preserve the inherent color and reduce the risk of over-saturation as shown in Fig. 3(b).

3. Ablation Study Results

In this section, we provide more results of ablation study, including the advantage of three-channel adjustment, the impact of training data, the contribution of each loss, and the effect of parameter settings.

Results of the ablation study: the advantage of three-channel adjustment. As shown in Fig. 3, both the single channel adjustment (CIE Lab and YCbCr color spaces) and the three-channel adjustment (RGB color space) can improve the brightness of the given low-light image. However, it is difficult to control the color of the enhanced result when only adjusting the illumination related channels (*i.e.*, L channel in CIE Lab color space and Y channel in YCbCr color space).

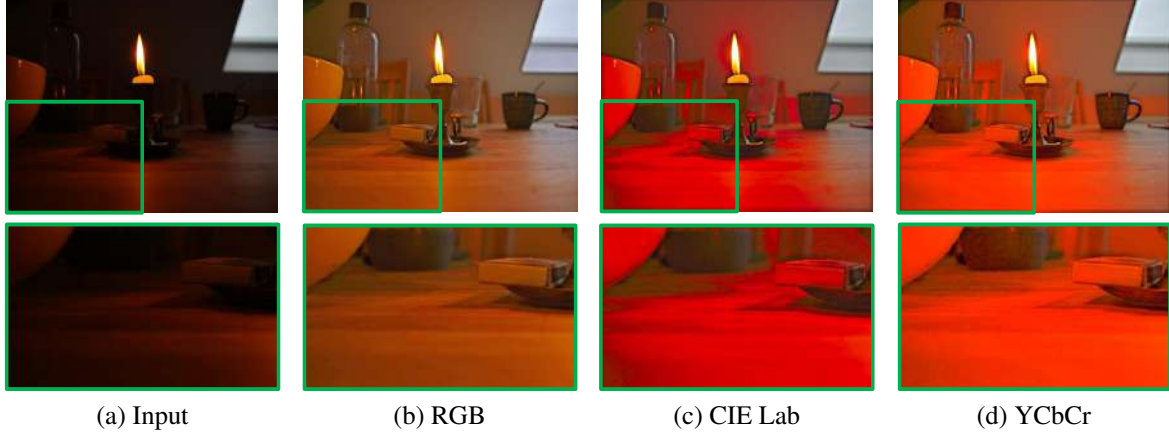


Figure 3: Ablation study of the advantage of three-channel adjustment (RGB, CIE Lab, and YCbCr color spaces). **Green boxes indicate the obvious differences and amplified details.**

Results of the ablation study: the impact of training data. As shown in Fig. 4(c) and (d), only using the low-light images as the training data (*i.e.*, Zero-DCE_{Low} and Zero-DCE_{LargeL}), the Zero-DCE tends to over-enhance the well-lit regions (*e.g.*, the cup). In contrast, our Zero-DCE as shown in Fig. 4(b) has a good balance between over-enhancement and under-enhancement. In Fig. 4(e), the Zero-DCE can better recover the dark regions (*e.g.* the roses) when more multi-exposure data are used as training data (*i.e.*, Zero-DCE_{LargeLH}).



Figure 4: Ablation study of the impact of training data. **Red and green boxes indicate the obvious differences and amplified details.**

Results of the ablation study: the contribution of each loss. As shown, without the spatial consistency loss L_{spa} , Fig. 5 (c) has relatively low contrast, such as the amplified region. The exposure control loss L_{exp} significantly affects the exposure level as shown in Fig. 5(d). When removing the color constancy loss L_{col} , the severe color deviation appears as shown in Fig. 5(e). As presented in Fig. 5(f), removing the illumination smoothness loss L_{tv_A} leads to obvious artifacts.

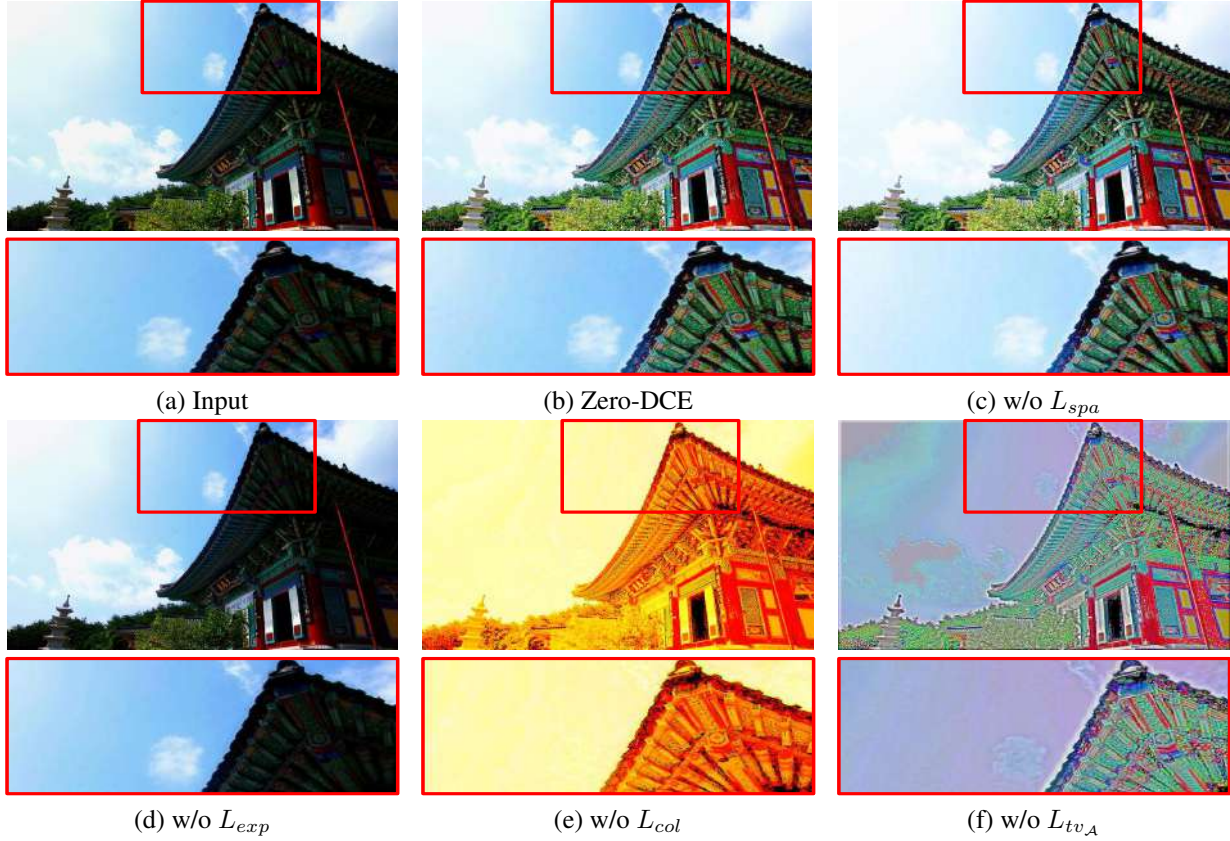


Figure 5: Ablation study of the contribution of each loss (spatial consistency loss L_{spa} , exposure control loss L_{exp} , color constancy loss L_{col} , illumination smoothness loss L_{tv_A}). **Red boxes indicate the obvious differences and amplified details.**

Results of the ablation study: the effect of parameter settings. In Table 1, we report the quantitative comparisons of different parameter settings of our Zero-DCE. Comparing the input with the Zero-DCE₃₋₃₂₋₈, it indicates that the Zero-DCE₃₋₃₂₋₈ can improve the quantitative performance in spite of only using three convolutional layers. The Zero-DCE₇₋₃₂₋₁ achieves the worst quantitative performance due to the limited adjustment capability of only one-time iteration (*i.e.*, $n=1$), which implies the importance of the iterations. When increasing the numbers of feature maps of each layer from 16 to 32, the quantitative performance is improved (*i.e.*, Zero-DCE₇₋₁₆₋₈ and Zero-DCE₇₋₃₂₋₈). When increasing the numbers of iterations from 8 to 16, the average PSNR value is slightly improved (*i.e.*, Zero-DCE₇₋₃₂₋₈ and Zero-DCE₇₋₃₂₋₁₆).

Table 1: Quantitative comparisons in terms of full-reference image quality assessment metrics ((PSNR, dB), SSIM, MAE)). This comparisons are carried out on the Part2 subset of SICE dataset [1]. l-f-n represents the proposed Zero-DCE with l convolutional layers, f feature maps of each layer (except the last layer), and n iterations.

Method	PSNR\uparrow	SSIM\uparrow	MAE\downarrow
Input	10.71	0.33	209.65
Zero-DCE ₃₋₃₂₋₈	14.50	0.56	119.40
Zero-DCE ₇₋₁₆₋₈	15.67	0.58	111.01
Zero-DCE ₇₋₃₂₋₁	12.21	0.42	172.89
Zero-DCE ₇₋₃₂₋₈	16.21	0.59	98.78
Zero-DCE ₇₋₃₂₋₁₆	16.79	0.57	98.70

4. Visual Comparison Results

In this section, we present more visual comparison results on different datasets (DICM, LIME, MEF, NPE, VV, Part2 subset) as shown in Figs. 6, 7, 8, 9, 10, 11.



Figure 6: Visual comparisons on DICM set [5]. Our Zero-DCE enhances the low-light regions and well preserves the inherent color and structure details. **Red and green boxes indicate the obvious differences and amplified details.**

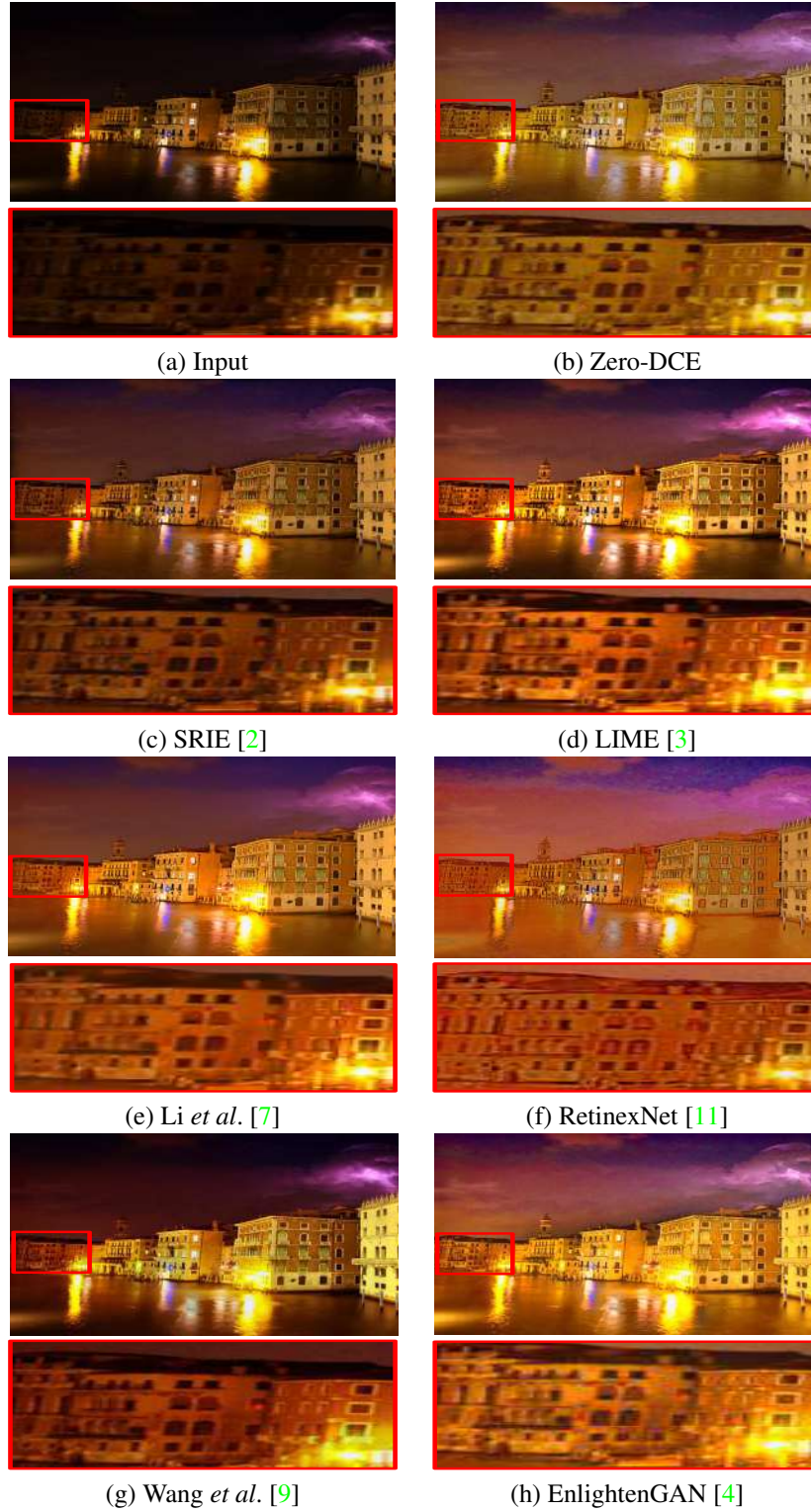


Figure 7: Visual comparisons on LIME set [3]. Our Zero-DCE produces more clear details (*e.g.*, the region in the red box). **Red boxes indicate the obvious differences and amplified details.**

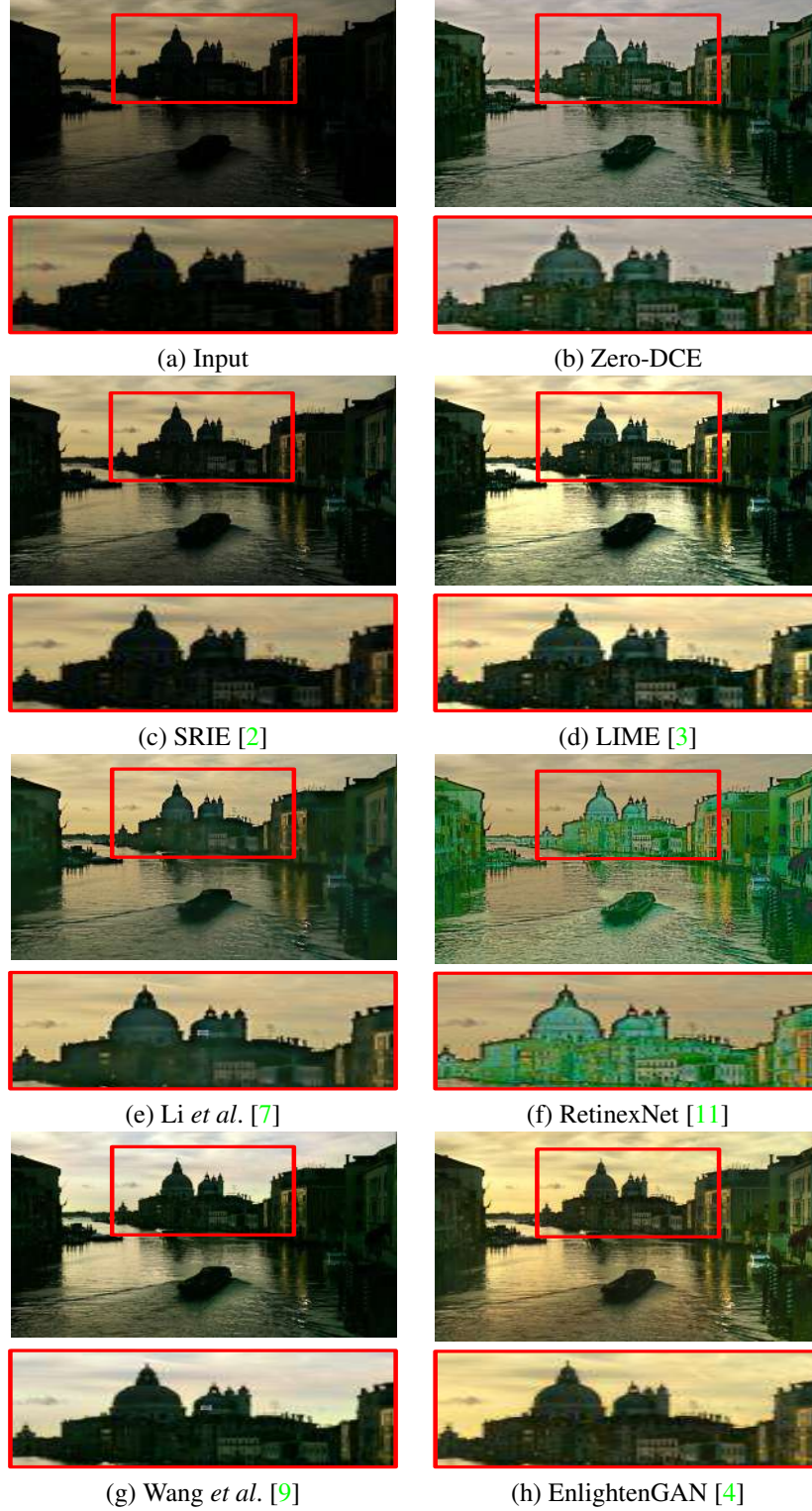


Figure 8: Visual comparisons on MEF set [8]. Our Zero-DCE well uncovers the structure details in the dark regions and does not introduce extra color artifacts. **Red boxes indicate the obvious differences and amplified details.**

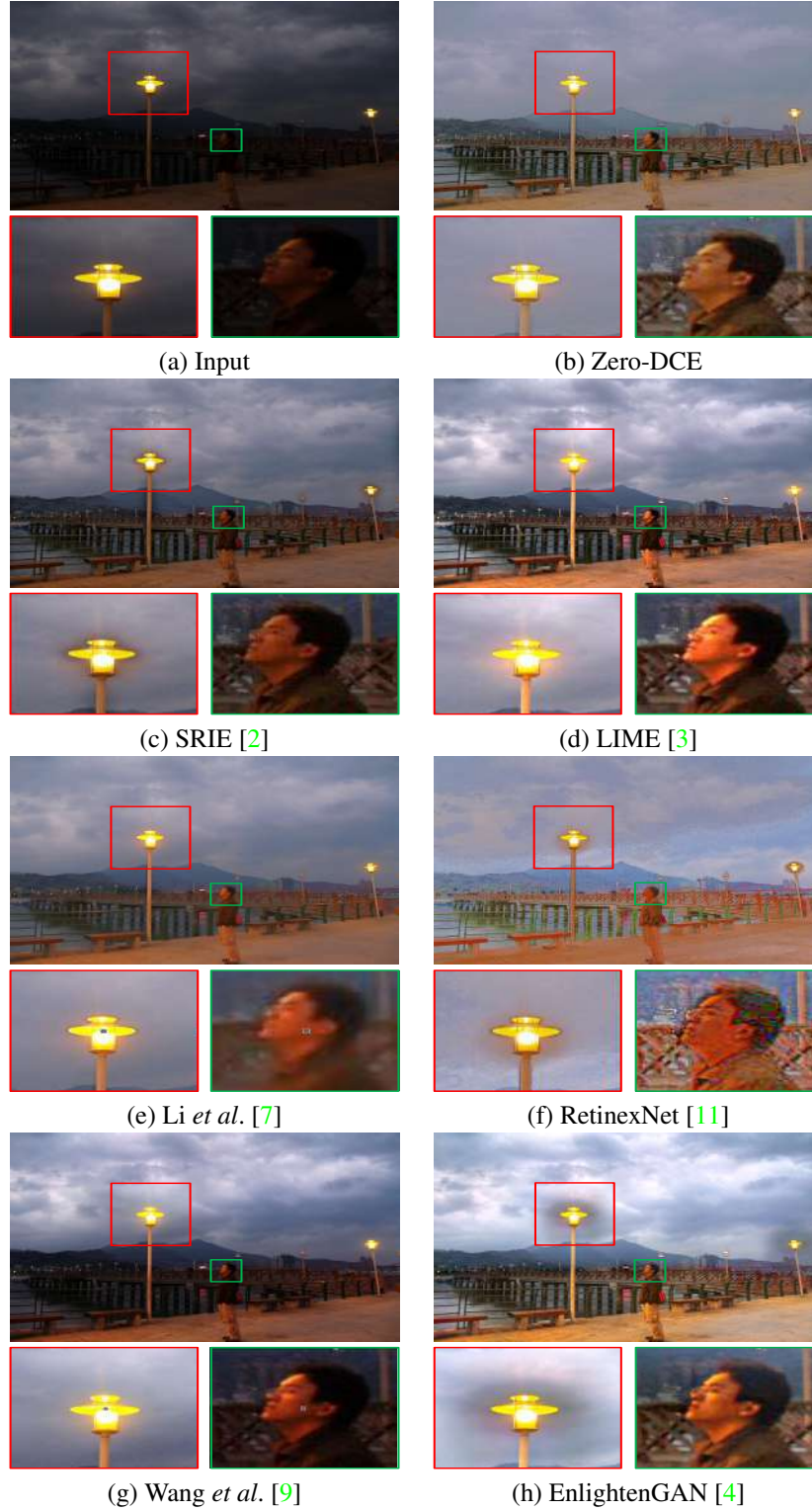


Figure 9: Visual comparisons on NPE set [10]. Our Zero-DCE introduces less halo artifacts and over-exposure than the compared methods. **Red and green boxes indicate the obvious differences and amplified details.**



Figure 10: Visual comparisons on VV set. Our Zero-DCE does not over-/under-enhance the low-light regions (*e.g.*, the face) and preserves the good contrast of the original image (*e.g.*, the mountain). **Red and green boxes indicate the obvious differences and amplified details.**

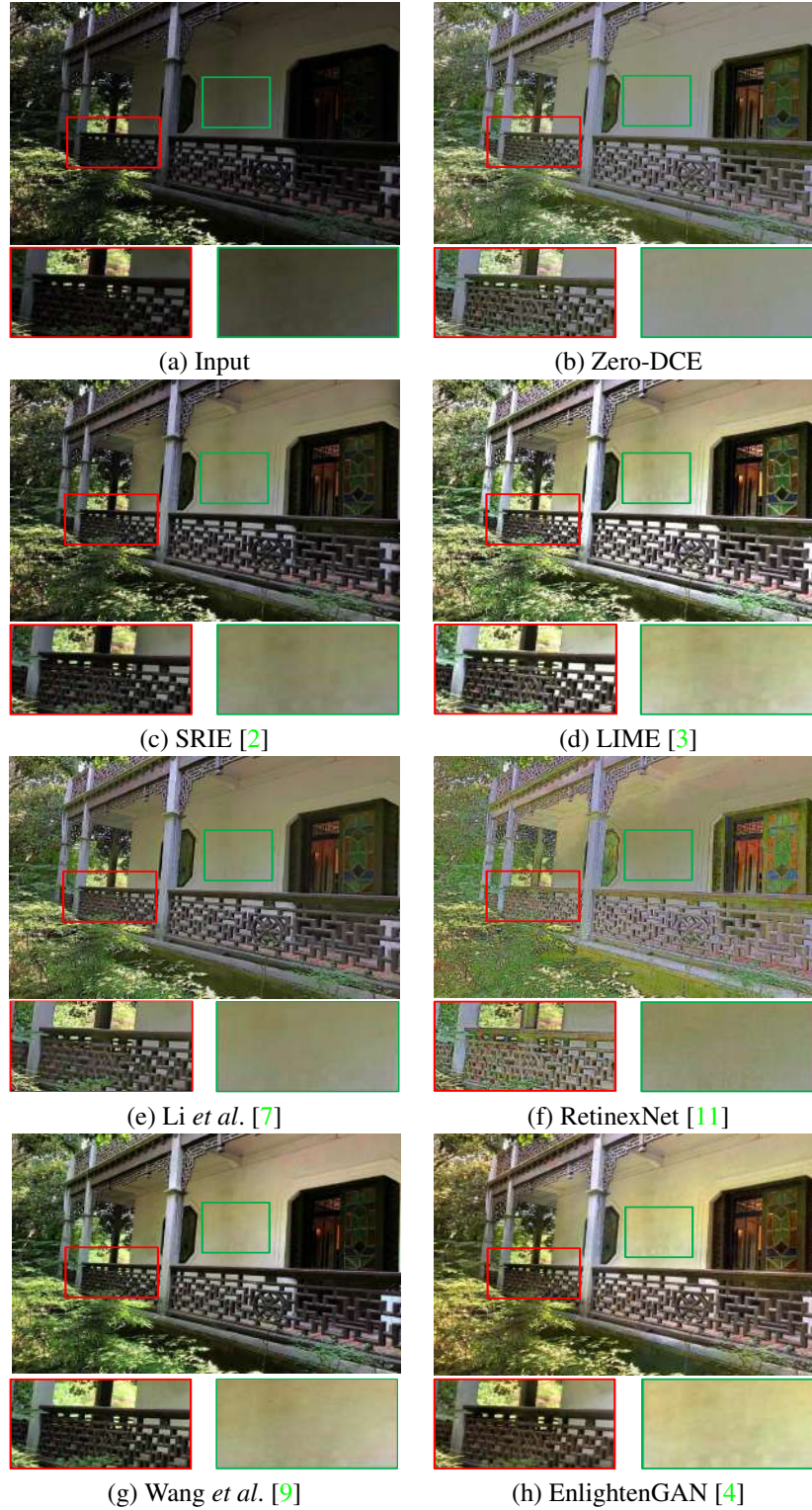


Figure 11: Visual comparisons on Part2 subset [1]. Our Zero-DCE produces more clear details and does not introduce blocking artifacts as shown in the green boxes of the compared methods. **Red and green boxes indicate the obvious differences and amplified details.**

5. Visual Face Detection Results

In this section, we present more visual face detection results for the low-light images from the DARK FACE dataset [12] in Fig. 12 and also provide more face detection results for the input images before and after enhanced by our Zero-DCE in Fig. 13.

Observing Fig. 12(a) and (b), after enhanced by our Zero-DCE, the face detector, Dual Shot Face Detector (DSFD) [6], can detect more faces than the raw low-light image as the input. By contrast, the results in Fig. 12(c), (e), (g) and (f) loss several faces that can be detected in our result. In addition, the result of Fig. 12(d) mistakenly detects the bicycle handlebar as a face as shown in the bottom right corner. Similarly, the result of Fig. 12(f) mistakenly detects a head (the face is invisible) as a face as shown in the leftmost side. Besides, in Fig. 13, it is obvious that our Zero-DCE effectively improves the performance of the face detector.



(a) Input



(b) Zero-DCE



(c) SRIE [2]



(d) LIME [3]



(e) Li *et al.* [7]



(f) RetinexNet [11]



(g) Wang *et al.* [9]



(h) EnlightenGAN [4]

Figure 12: Visual face detection comparisons. **Best viewed with zoom-in on a digital display.**



Figure 13: Visual face detection results before and after enhanced by our Zero-DCE.

References

- [1] Jianrui Cai, Shuhang Gu, and Lei Zhang. Learning a deep single image contrast enhancer from multi-exposure image. *IEEE Transactions on Image Processing*, 27(4):2049–2026, 2018. 5, 11
- [2] Xueyang Fu, Delu Zeng, Yue Huang, Xiao-Ping Zhang, and Xinghao Ding. A weighted variational model for simultaneous reflectance and illumination estimation. In *CVPR*, 2016. 6, 7, 8, 9, 10, 11, 13
- [3] 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. 6, 7, 8, 9, 10, 11, 13
- [4] Yifan Jiang, Xinyu Gong, Ding Liu, Yu Cheng, Chen Fang, Xiaohui Shen, Jianchao Yang, Pan Zhou, and Zhangyang Wang. EnlightenGAN: Deep light enhancement without paired supervision. In *CVPR*, 2019. 6, 7, 8, 9, 10, 11, 13
- [5] Chulwoo Lee, Chul Lee, and Chang-Su Kim. Contrast enhancement based on layered difference representation. In *ICIP*, 2012. 6
- [6] Jian Li, Yabiao Wang, Changan Wang, Ying Tai, Jianjun Qian, Jian Yang, Chengjie Wang, Jilin Li, and Feiyuen Huang. Dsfd: Dual shot face detector. In *CVPR*, 2019. 12
- [7] 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. 6, 7, 8, 9, 10, 11, 13
- [8] 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. 8
- [9] Ruixing Wang, Qing Zhang, Chi-Wing Fu, Xiaoyong Shen, Wei-Shi Zheng, and Jiaya Jia. Underexposed photo enhancement using deep illumination estimation. In *CVPR*, 2019. 6, 7, 8, 9, 10, 11, 13
- [10] 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. 9
- [11] Chen Wei, Wenjing Wang, Wenhan Yang, and Jiaying Liu. Deep retinex decomposition for low-light enhancement. In *BMVC*, 2018. 6, 7, 8, 9, 10, 11, 13
- [12] Ye Yuan, Wenhan Yang, Wenqi Ren, Jiaying Liu, Walter J Scheirer, and Wang Zhangyang. Ug+ track 2: A collective benchmark effort for evaluating and advancing image understanding in poor visibility environments, 2019. arXiv arXiv:1904.04474. 12