Local Color Distributions Prior for Image Enhancement Supplemental Material

Haoyuan Wang, Ke Xu*, and Rynson W.H. Lau*

Department of Computer Science, City University of Hong Kong

In this supplemental, we provide more visual comparisons between our results and those of existing methods and more details about our loss function. .[1]

1 Visual Comparisons

We first show the comparisons on our proposed dataset in Fig. 1, Fig. 2, Fig. 3, Fig. 4, and Fig. 5. The input images cover diverse scenes (*e.g.*, outdoor and indoor, daytime and nighttime), and our method can reconstruct the details in both over-exposed and under-exposed regions. We then show comparisons of over-exposure correction on the MSEC [1] dataset in Fig. 6, Fig. 7 and Fig. 8. Comparing to the state-of-the-art MSEC [1], our method reconstructs better textures and colors. Lastly, we show comparisons of images with both over-/under-exposure from the Internet in Fig. 9 and Fig. 10.

2 Loss Function Details

The overall function is:

$$\mathcal{L}_{\text{total}} = \lambda_1 \mathcal{L}_{\text{mse}} + \lambda_2 \mathcal{L}_{\text{cos}} + \lambda_3 \mathcal{L}_{\text{tv1}} + \lambda_4 \mathcal{L}_{\text{tv2}}, \tag{1}$$

where λ_1, λ_2, λ_3 and λ_4 are the balancing hyper-parameters, $\mathcal{L}_{\text{mse}}, \mathcal{L}_{\text{cos}}, \mathcal{L}_{\text{tv1}}, \mathcal{L}_{\text{tv2}}$ are MSE loss, cos similarity loss, under-illumination smoothness loss and over-illumination smoothness loss, respectively.

MSE Loss. We use the MSE loss to measure the pixel-wise difference:

between the output image and the expert-retouched reference image to help image content reconstruction:

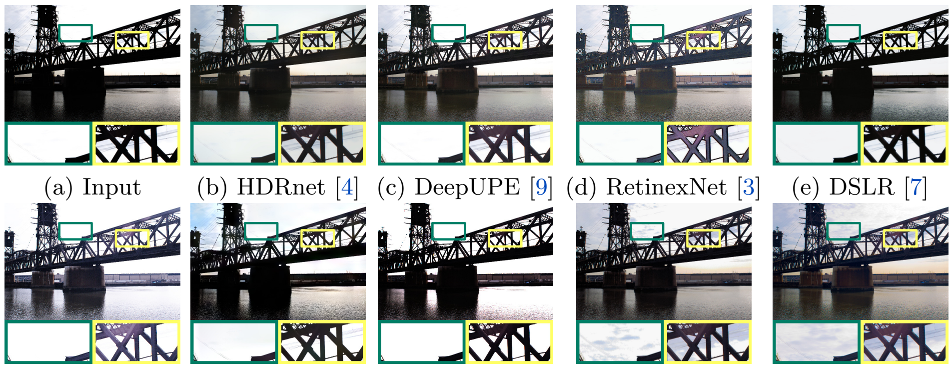
$$\mathcal{L}_{\text{mse}} = \frac{1}{hw} ||I_y - I'_y||_2^2, \tag{2}$$

where I_y, I'_y are output image and ground truth. in a mini-batch with size N.

Cosine Similarity Loss. Since uneven illumination causes color distortion, we use the cosine similarity loss to measure the color difference:

it is necessary to reconstruct the colors. We use the cosine similarity loss to measure the color difference between the output image and ground truth. Maximizing the cosine similarity of the pixels in the RGB space is equivalent

[1] * Joint corresponding authors.



(a) Input (b) HDRnet [4] (c) DeepUPE [9] (d) RetinexNet [3] (e) DSLR [7]

(f) ZeroDCE [5] (g) MSEC [1] (h) RUAS [8] (i) Ours (j) GT

Fig. 1: Visual comparison of an outdoor scene with over-/under-exposure from our dataset.

to minimizing the angle between two colors vectors, which helps restore the color-damaged regions. The function is expressed as:

$$\mathcal{L}_{\cos} = \frac{1}{hw} \sum_{p} s\left(1 - \frac{I_{y_p} \cdot I'_{y_p}}{||I_{y_p}||_2 \cdot ||I'_{y_p}||_2}\right), \tag{3}$$

where p is the spatial position of pixels across the image and $s(\cdot)$ is the sigmoid function.

Illumination Smoothness Loss. According to [2,6,9], natural illumination has local smoothness. To preserve such characteristics, in the estimated under-illumination map, we define the illumination smoothness loss according to [9,10], by accumulating the variation of the gradient map. For the forward illumination estimation path, we have:

$$\begin{aligned}\mathcal{L}_{\text{tv1}} =& \mathcal{L}_{\text{tv}}(L, I_x) \\ =& \sum_{p} \left(\frac{|\nabla_h L_p|^2}{|\nabla_h \log(I_{x_p})|^\alpha} + \frac{|\nabla_v L_p|^2}{|\nabla_h \log(I_{x_p})|^\alpha} \right),\end{aligned} \tag{4}$$

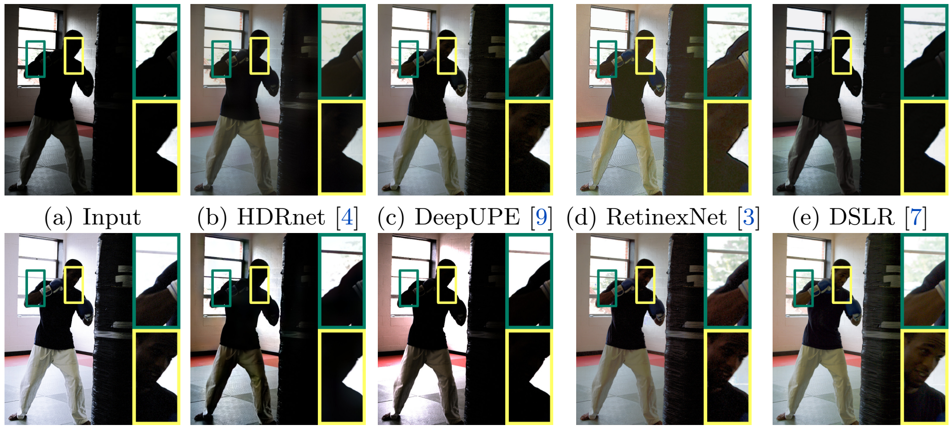
where p is the spatial position. ∇_h and ∇_v represent the horizontal and vertical gradient operators. $\alpha = 1.2$ controls the sensitivity to the image gradient map.

The loss function only penalizes the unsmoothness where the gradient value is small in the illumination map, which can be regarded as an edge-preserving prior to make the illumination map more natural.

Similarly, the smoothness loss function of the over-illumination estimation path could be formulated as $L_{\text{tv2}} = L_{\text{tv}}(L', I'_x)$.

References

1. Afifi, M., Derpanis, K.G., Ommer, B., Brown, M.S.: Learning multi-scale photo exposure correction. In: CVPR (2021)



(a) Input (b) HDRnet [4] (c) DeepUPE [9] (d) RetinexNet [3] (e) DSLR [7]

(f) ZeroDCE [5] (g) MSEC [1] (h) RUAS [8] (i) Ours (j) GT

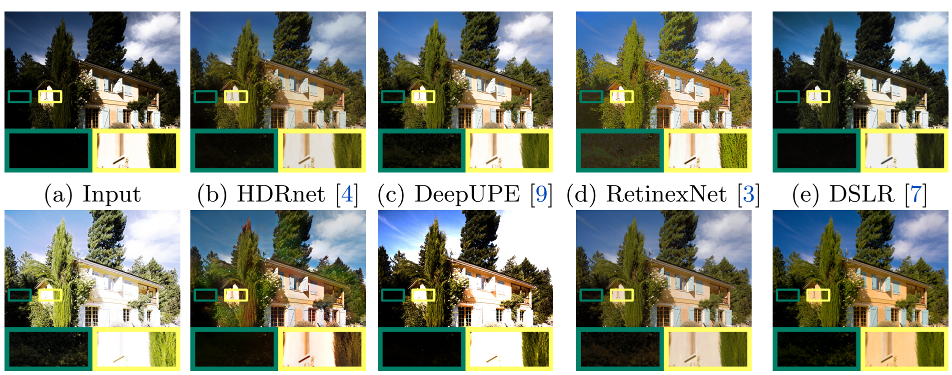
Fig. 2: Visual comparison of an indoor scene with over-/under-exposure from our dataset.

2. Barron, J.T., Malik, J.: Shape, illumination, and reflectance from shading. IEEE TPAMI (2014)
3. Chen, W., Wenjing, W., Wenhan, Y., Jiaying, L.: Deep retinex decomposition for low-light enhancement. In: BMVC (2018)
4. Gharbi, M., Chen, J., Barron, J.T., Hasinoff, S.W., Durand, F.: Deep bilateral learning for real-time image enhancement. ACM TOG (2017)
5. Guo, C., Li, C., Guo, J., Loy, C.C., Hou, J., Kwong, S., Cong, R.: Zero-reference deep curve estimation for low-light image enhancement. In: CVPR (2020)
6. Guo, X., Li, Y., Ling, H.: Lime: Low-light image enhancement via illumination map estimation. IEEE TIP (2017)
7. Ignatov, A., Kobyshev, N., Timofte, R., Vanhoey, K., Van Gool, L.: Dslr-quality photos on mobile devices with deep convolutional networks. In: ICCV (2017)
8. Risheng, L., Long, M., Jiaao, Z., Xin, F., Zhongxuan, L.: Retinex-inspired unrolling with cooperative prior architecture search for low-light image enhancement. In: CVPR (2021)
9. Wang, R., Zhang, Q., Fu, C., Shen, X., Zheng, W., Jia, J.: Underexposed photo enhancement using deep illumination estimation. In: CVPR (2019)
10. Xu, L., Yan, Q., Xia, Y., Jia, J.: Structure extraction from texture via natural variation measure. ACM TOG (2012)

(a) Input (b) HDRnet [4] (c) DeepUPE [9] (d) RetinexNet [3] (e) DSLR [7]

(f) ZeroDCE [5] (g) MSEC [1] (h) RUAS [8] (i) Ours (j) GT

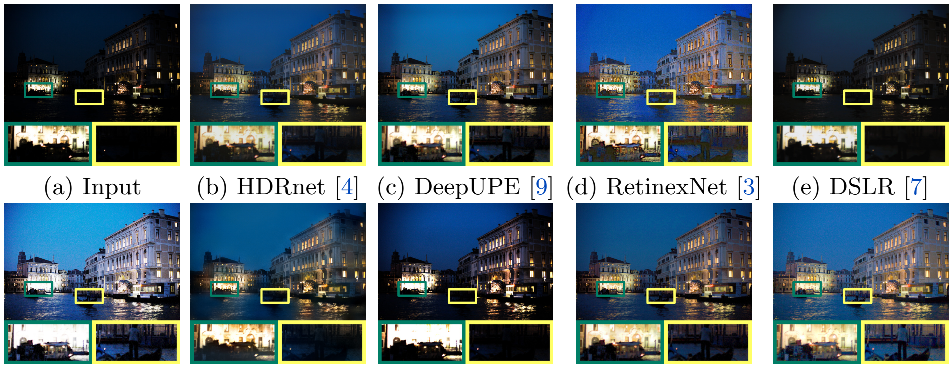
Fig. 3: Visual comparison of an outdoor scene with over-/under-exposure from our dataset.



(a) Input (b) HDRnet [4] (c) DeepUPE [9] (d) RetinexNet [3] (e) DSLR [7]

(f) ZeroDCE [5] (g) MSEC [1] (h) RUAS [8] (i) Ours (j) GT

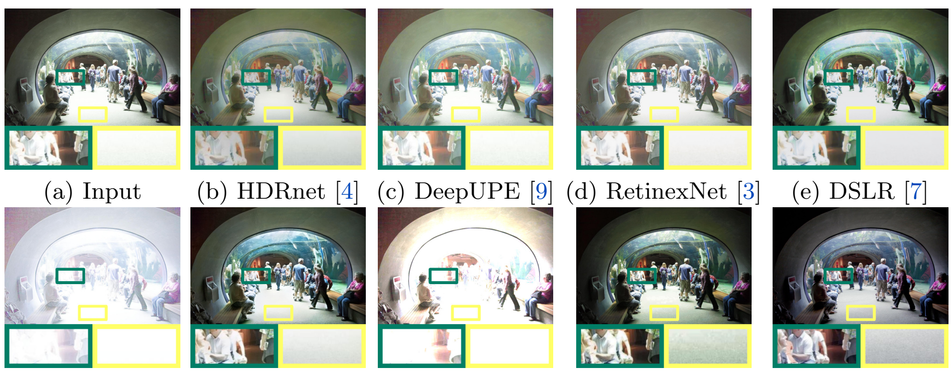
Fig. 4: Visual comparison of an outdoor scene with over-/under-exposure from our dataset.



(a) Input (b) HDRnet [4] (c) DeepUPE [9] (d) RetinexNet [3] (e) DSLR [7]

(f) ZeroDCE [5] (g) MSEC [1] (h) RUAS [8] (i) Ours (j) GT

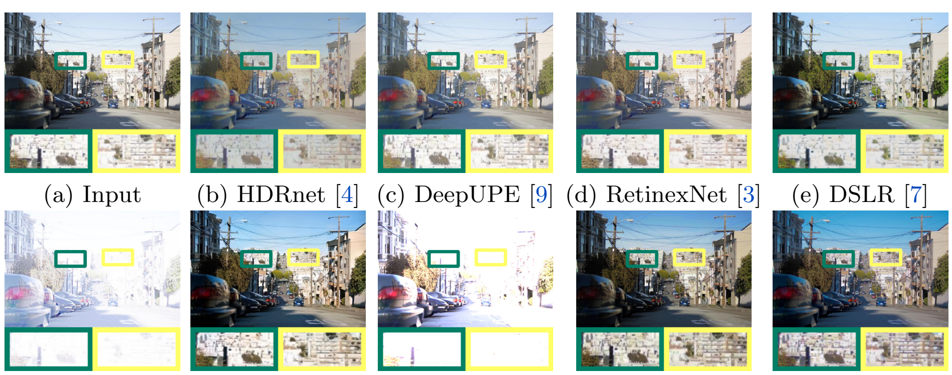
Fig. 5: Visual comparison of a nighttime scene with over-/under-exposure from our dataset.



(a) Input (b) HDRnet [4] (c) DeepUPE [9] (d) RetinexNet [3] (e) DSLR [7]

(f) ZeroDCE [5] (g) MSEC [1] (h) RUAS [8] (i) Ours (j) GT

Fig. 6: Visual comparison of an over-exposed image from the MSEC [1] dataset.

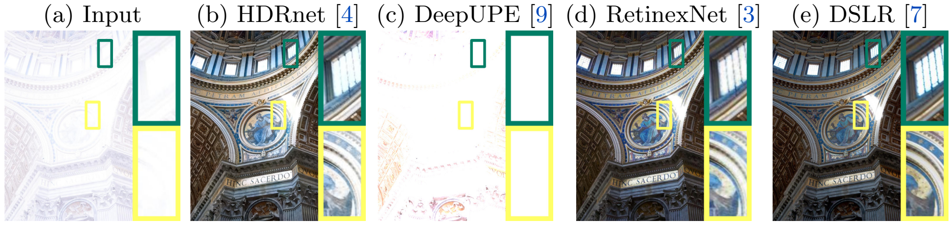


(a) Input (b) HDRnet [4] (c) DeepUPE [9] (d) RetinexNet [3] (e) DSLR [7]

(f) ZeroDCE [5] (g) MSEC [1] (h) RUAS [8] (i) Ours (j) GT

Fig. 7: Visual comparison of an over-exposed image from the MSEC [1] dataset.

(a) Input (b) HDRnet [4] (c) DeepUPE [9] (d) RetinexNet [3] (e) DSLR [7]



(f) ZeroDCE [5] (g) MSEC [1] (h) RUAS [8] (i) Ours (j) GT

Fig. 8: Visual comparison of an over-exposed image from the MSEC [1] dataset.

(a) Input (b) RetinexNet [3] (c) ZeroDCE [5] (d) DSLR [7]

(e) DeepUPE [9] (f) RUAS [8] (g) MSEC [1] (h) Ours

Fig. 9: Visual comparison of an outdoor image with over-/under-exposure from the Internet.

(a) Input (b) RetinexNet [3] (c) ZeroDCE [5] (d) DSLR [7]

(e) DeepUPE [9] (f) RUAS [8] (g) MSEC [1] (h) Ours

Fig. 10: Visual comparison of an outdoor image with over-/under-exposure from the Internet.