

Supplementary Material

Color-wise Attention Network for Low-light Image Enhancement

Yousef Atoum
Yarmouk University

Mao Ye
BOSCH Research North America

Liu Ren
BOSCH Research North America

Ying Tai
Tencent YouTu

Xiaoming Liu
Michigan State University

1. Contents of the supplementary material

In this document, we provide additional materials to supplement our main paper submission. We mainly focus on qualitative results of our proposed method in the provided materials, since the main submission has sufficient quantitative experiments. Note that all images used in this supplementary document are down-sampled by a factor of 0.25 to satisfy the size limitation when submitted.

The contents of this document is as follows:

- In Sec. 2, we evaluate CWAN on both synthetic and real low-light videos. The results of the low-light video enhancement are provided in separate supplementary video files.
- In Sec. 3, we provide visual examples of the proposed $CWAN_L$ model. We show the enhanced lightness \mathbf{X}_L of the five test images used in the LLNet paper [10].
- In Sec. 4, we provide additional qualitative results on SID_{Sony} , SID_{Fuji} and $PASCAL_{1000}$.
- In Sec. 5, we provide additional qualitative results on HDRDB.
- In Sec. 6, we have a discussion regarding generating very colorful images, sometimes even better than the ground truth images.

2. Evaluating CWAN on low-light videos

The ability to enhance low-light images is the main objective of our submission, where CWAN is proven to achieve state-of-the-art results compared to many baselines. We further conduct two additional experiments on low-light video enhancement, comparing our proposed CWAN method with the top three baselines which achieved the best results based on our main paper submission experiments, namely LIME [6], LightenNet [8], and SID [1].

In the first experiment, we use a small subset of seven videos obtained from the e-VDS35 database [2]. We randomly select the videos from six different object categories, which are chair, cup, lawn, computer, dog, and toys (two videos from toys). These specific categories were selected because they contained most color variations compared to other categories, i.e., ceiling, floor, or hand. The original videos of e-VDS35 have good lighting conditions, and therefore, we synthetically generate low-light videos by processing every frame using the Retinex-based approach as performed in [8], reducing the illumination map by 85%. The results are illustrated in a supplementary video named "*Video_1_Paper_ID_12.mp4*" demonstrating the enhanced video results along with plotting the PSNR values at each frame.

In the second experiment, we collect three videos of extreme low-light conditions in a real-world environment during night time. The first video was collected indoors with the only light source present in the room being from the laptop screen, while the remaining two videos were captured outdoors with very far light sources, i.e., moon and street lights. The videos are captured using a Canon EOS Rebel T6 camera, and saved as AVI format. Other than the extreme low light conditions in these videos, severe noise is also present. The results are illustrated in a supplementary video named "*Video_2_Paper_ID_12.mp4*" comparing the performance of CWAN with the three baselines, along with plotting the LOE values at each frame. From this video, we can observe that LIME usually over enhances low-light, producing high intensities at object surfaces reflecting light. LIME also tends to amplify the noise caused by extreme light conditions producing unpleasant scenes. LightenNet and SID produce very similar results in enhancing low-light videos, but both methods seem to produce poor colors in the enhanced image. On the other hand, with the aid of CWAN_{AB}, our enhanced low-light image contains rich colors closer to the true object colors, and tends to suppress noise very well.

In summary, all three baselines including CWAN perform very well on enhancing low-light video, with a clear advantage by CWAN in synthesizing good color features and reducing the effect of noise. In Table 1, we report the average PSNR/LOE results. One important outcome of enhancing low-light videos, is the ability to enhance the lighting in a consistent manner throughout the video frames. We can observe from Table 1, that CWAN in most videos has the lowest standard deviation, demonstrating the stable performance compared to other methods.

Database	Videos	LIME [6]	LightenNet [8]	SID [1]	CWAN ^{Sony}
Synthetic video e-VDS35 [2] (PSNR)	Chair	19.87 \pm 0.97	18.33 \pm 1.20	21.11 \pm 0.25	24.58 \pm 0.57
	Cup	23.40 \pm 1.13	21.56 \pm 0.49	22.92 \pm 0.38	25.10 \pm 0.19
	Dog	18.02 \pm 4.09	21.76 \pm 1.33	21.20 \pm 1.42	26.25 \pm 1.06
	Lawn	21.76 \pm 1.76	21.34 \pm 0.95	21.18 \pm 0.33	23.87 \pm 0.27
	Computer	22.47 \pm 2.02	14.33 \pm 1.57	23.59 \pm 1.07	27.62 \pm 1.45
	Toy-1	18.36 \pm 1.24	10.97 \pm 1.00	21.80 \pm 0.35	27.20 \pm 0.30
	Toy-2	14.82 \pm 0.32	19.53 \pm 0.75	19.41 \pm 0.17	23.10 \pm 0.13
Real low-light videos (LOE)	Indoor toys	109.64 \pm 2.34	91.06 \pm 5.02	56.89 \pm 1.19	50.59 \pm 1.36
	Outdoor cactus	127.63 \pm 31.94	101.53 \pm 32.53	96.50 \pm 32.22	83.49 \pm 28.25
	Outdoor garden	108.88 \pm 24.82	75.70 \pm 18.03	77.43 \pm 18.03	71.19 \pm 14.67

Table 1. Average results on low-light video enhancement. All synthetic videos from e-VDS35 [2] are evaluated using the PSNR measure, while all collected real-world videos are evaluated using the LOE measure. Red/blue font means the best/second best results.

3. Lightness component enhancement \hat{X}_L in $CWAN_L$

In Fig. 1, we illustrate the enhanced $CWAN_L$ images of the ablation study "Analysis of $CWAN_L$ ".



Figure 1. Visual analysis of $CWAN_L$. First row, the ground truth images of the five test set images used in [10]. Second row, the synthesized low-light images using the same protocol in [10]. Forth row, the synthesized low-light with Gaussian noise images using the same protocol in [10]. Third and fifth row, are the estimated enhanced lightness images from $CWAN_L$ given the corresponding test image.

4. Visual results of SID_{Sony} , SID_{Fuji} and PASCAL₁₀₀₀

We provide several examples on low-light enhancement illustrated in Fig. 2, 3, 4, 5, 6 and 7, which were obtained from SID_{Sony} , SID_{Fuji} and PASCAL₁₀₀₀. The PSNR and SSIM values for every image in these examples are reported in Table 2.

Method	Fig. 2	Fig. 3	Fig. 4	Fig. 5	Fig. 6	Fig. 7
HE	20.21/0.511	21.27/0.765	23.03/0.844	13.89/0.577	26.49/0.823	28.55/0.828
Dong [3]	22.39/0.770	20.15/0.837	20.31/0.909	26.66/0.931	18.62/0.917	23.51/0.807
BIMEF [13]	21.68/0.803	11.17/0.704	14.37/0.683	15.46/0.671	11.06/0.687	21.34/0.821
Ying [14]	22.23/0.800	15.67/0.783	19.09/0.785	21.01/0.790	16.17/0.805	20.72/0.770
JED [11]	20.24/0.716	10.68/0.723	13.81/0.766	15.41/0.801	10.91/0.667	19.97/0.731
AMSR [7]	17.18/0.632	11.88/0.716	15.21/0.818	18.91/0.890	11.24/0.745	17.43/0.634
LIME [6]	19.70/0.745	16.90/0.863	22.00/0.932	22.56/0.920	15.98/0.935	18.59/0.694
Li [9]	20.38/0.717	10.67/0.717	13.83/0.775	15.41/0.799	10.90/0.670	19.85/0.715
LECARM [12]	22.23/0.851	13.89/0.821	16.67/0.829	17.53/0.820	14.38/0.855	23.78/0.944
MBLLEN [5]	13.29/0.668	13.06/0.746	15.46/0.744	15.90/0.780	21.22/0.814	13.45/0.553
LightenNet [8]	25.70/0.858	17.69/0.818	15.21/0.852	21.34/0.922	14.25/0.845	21.84/0.895
SID [1]	26.63/0.848	26.36/0.868	25.81/0.905	26.71/0.939	26.22/0.858	29.19/0.900
CWAN ^{Sony}	28.72/0.896	29.25/0.913	27.34/0.947	27.53/0.947	31.85/0.957	32.60/0.915
CWAN ^{Fuji}	28.09/0.892	28.70/0.921	25.93/0.950	25.64/0.946	30.69/0.958	30.19/0.871

Table 2. PSNR and SSIM for the examples in the listed figures.

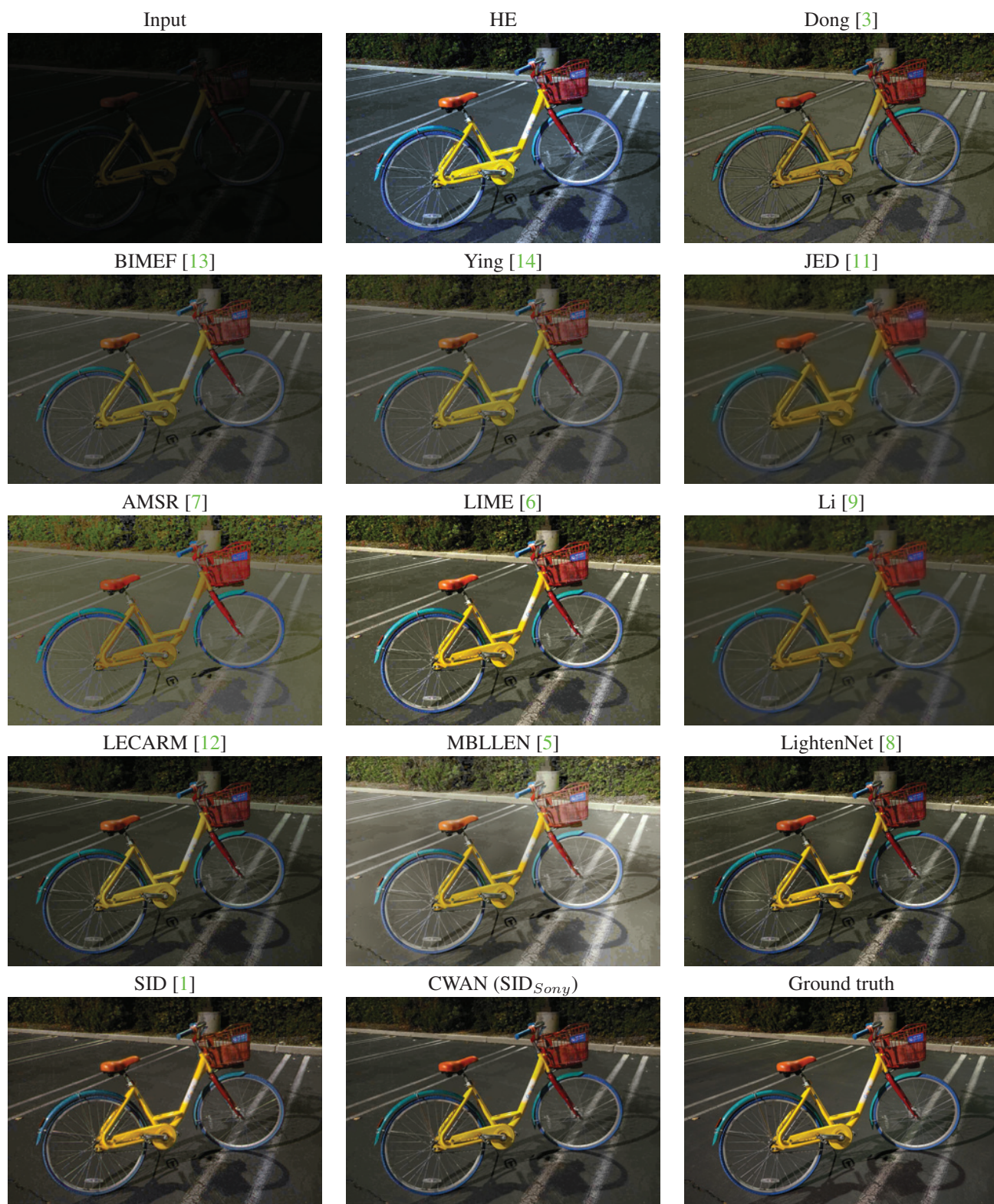


Figure 2. Example from SID_{Sony}.

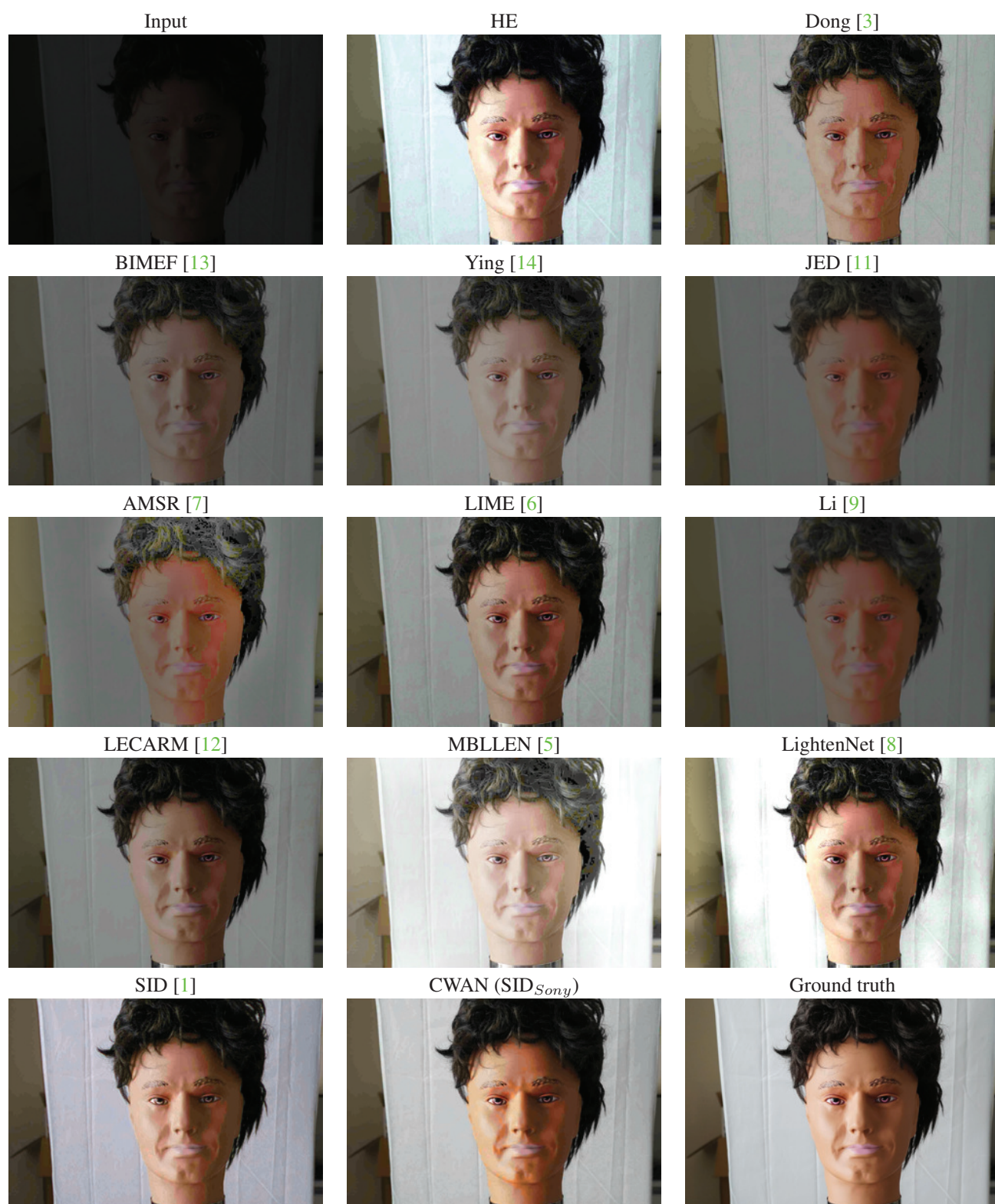


Figure 3. Example from SID_{Sony}.

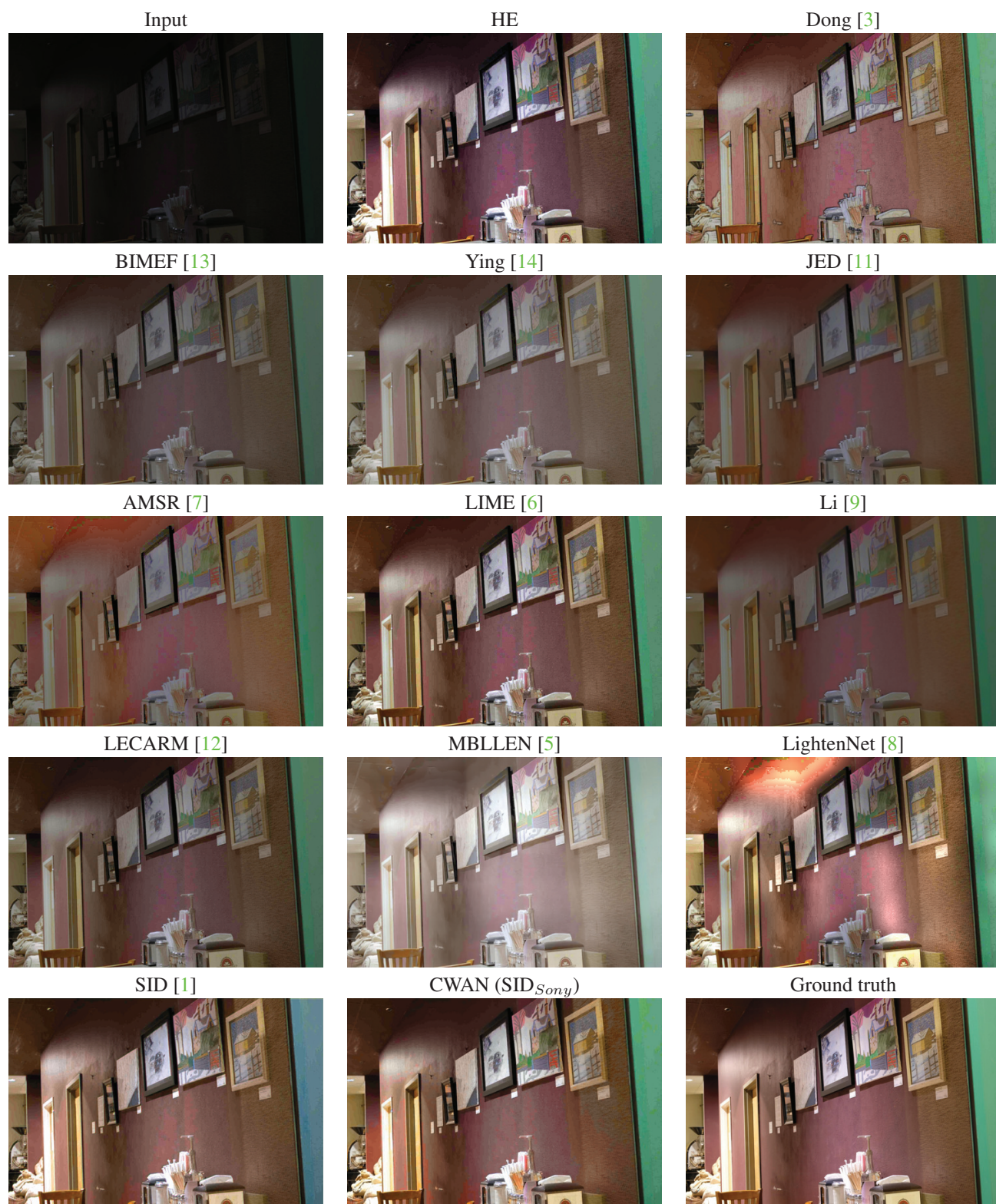
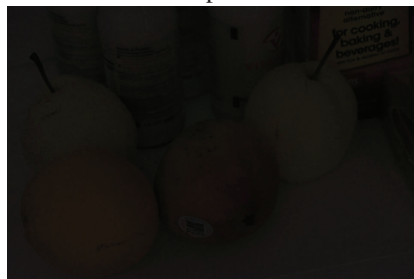


Figure 4. Example from SID_{Fuji}.

Input



HE



Dong [3]



BIMEF [13]



Ying [14]



JED [11]



AMSR [7]



LIME [6]



Li [9]



LECARM [12]



MBLLEN [5]



LightenNet [8]



SID [1]



CWAN (SID_{Sony})



Ground truth



Figure 5. Example from SID_{Fuji}.

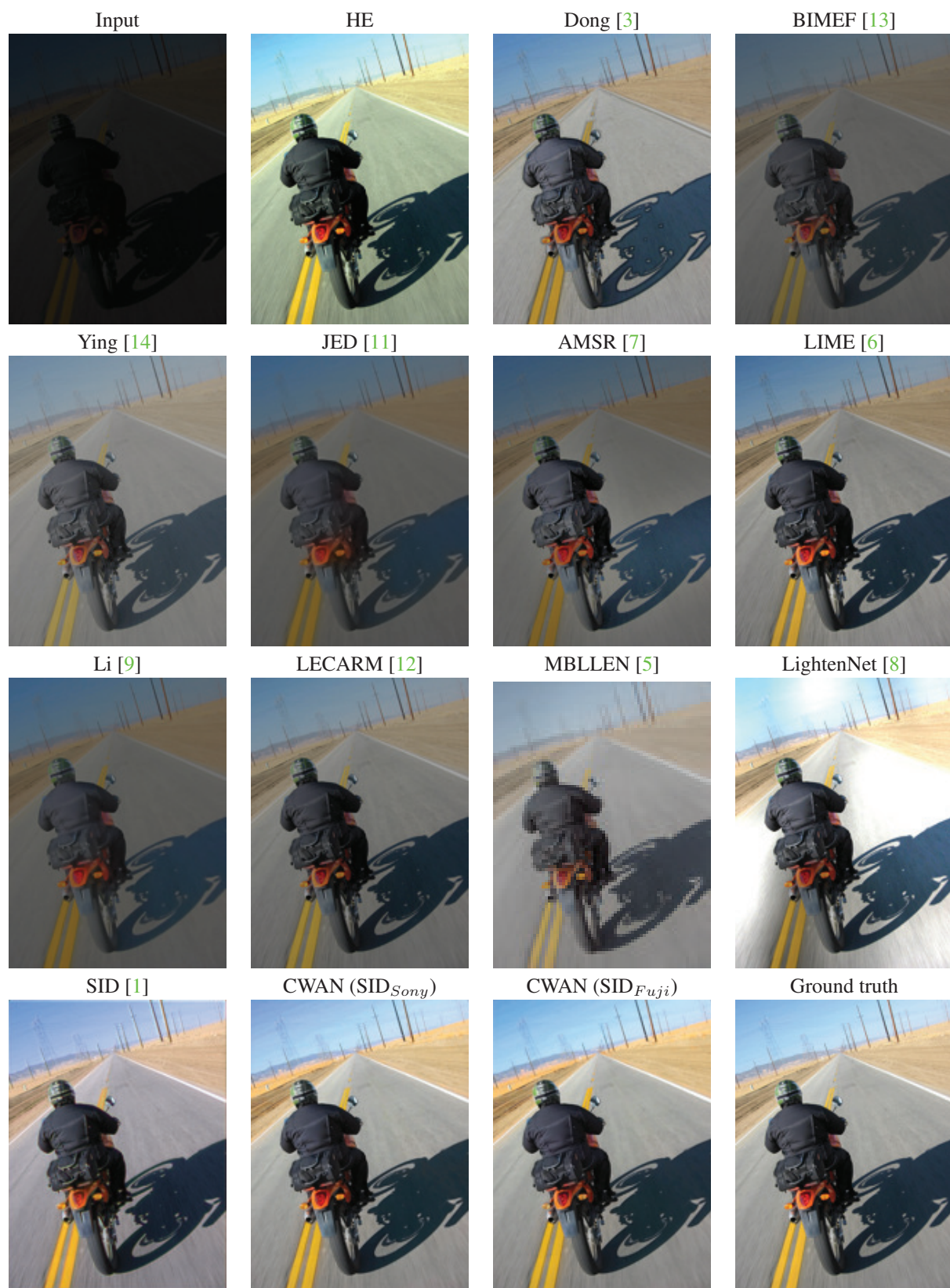


Figure 6. Example from PASCAL₁₀₀₀.



Figure 7. Example from PASCAL₁₀₀₀.

5. Visual results of HDRDB

Our qualitative and case study experiment in the main paper was carried out using HDRDB [4]. We provide more examples using this set on low-light enhancement illustrated in Fig. 8 and Fig. 9.



Figure 8. Example from HDRDB.

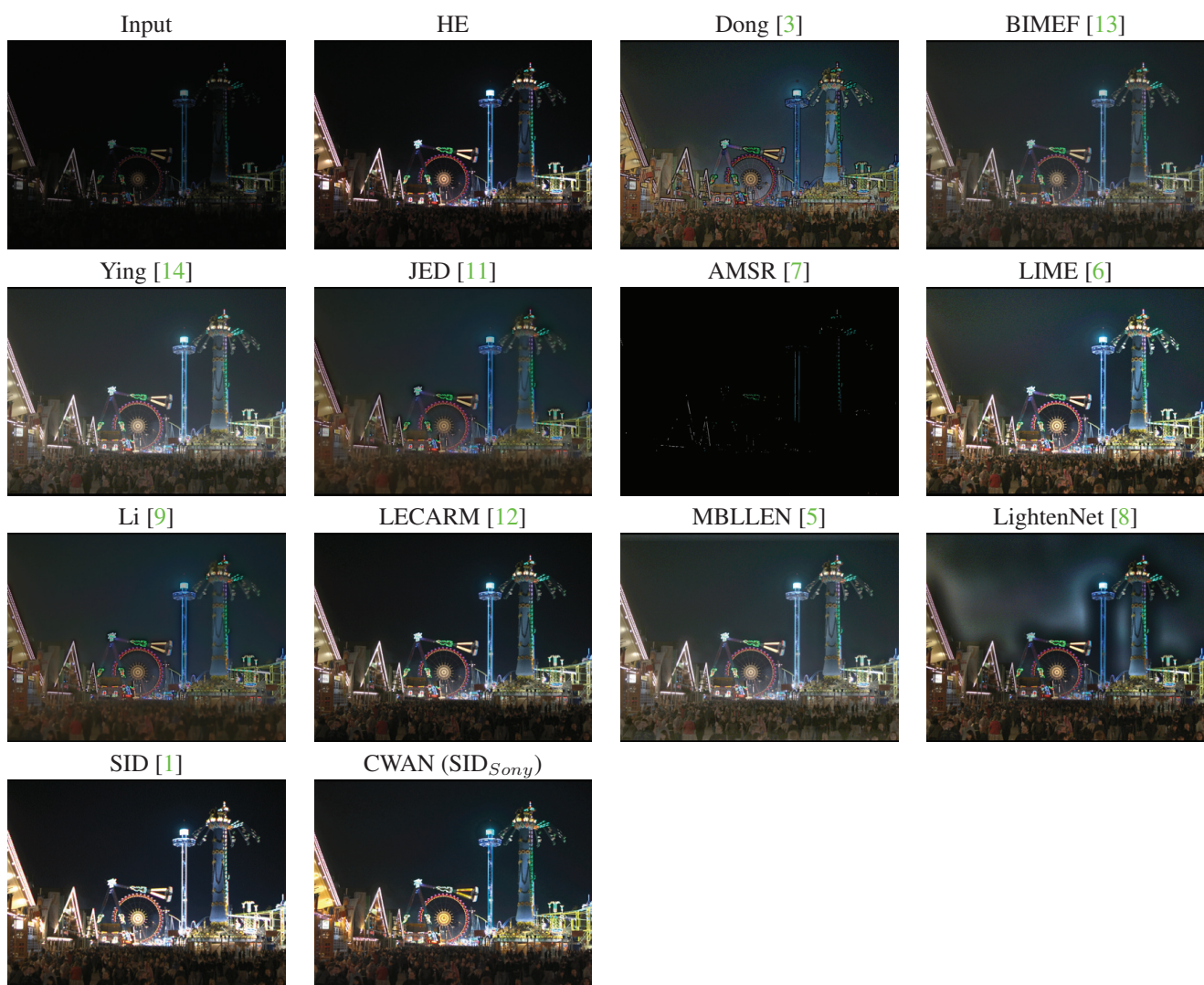


Figure 9. Example from HDRDB.

6. Generating colorful images

CWAN_{AB} is supervised by key color cues generated by the attention maps $\hat{\mathbf{M}}$, to synthesize vibrant and colorful enhanced low-light images. We observe in many test cases, the enhanced image seems to have better colors than the ground truth image as in Fig. 10. Note that CWAN_{AB} is learned to map from low-light images to enhanced colorful images using the SID database. Therefore, it will attempt to convert any low-light image into its best colorful version, which in turn might be better than the ground truth. The colorfulness metrics are also correlated with our observation. For example, the ground truth images of PASCAL₁₀₀₀ have $C = 32.51$, where our estimated images have $C = 37.21$.

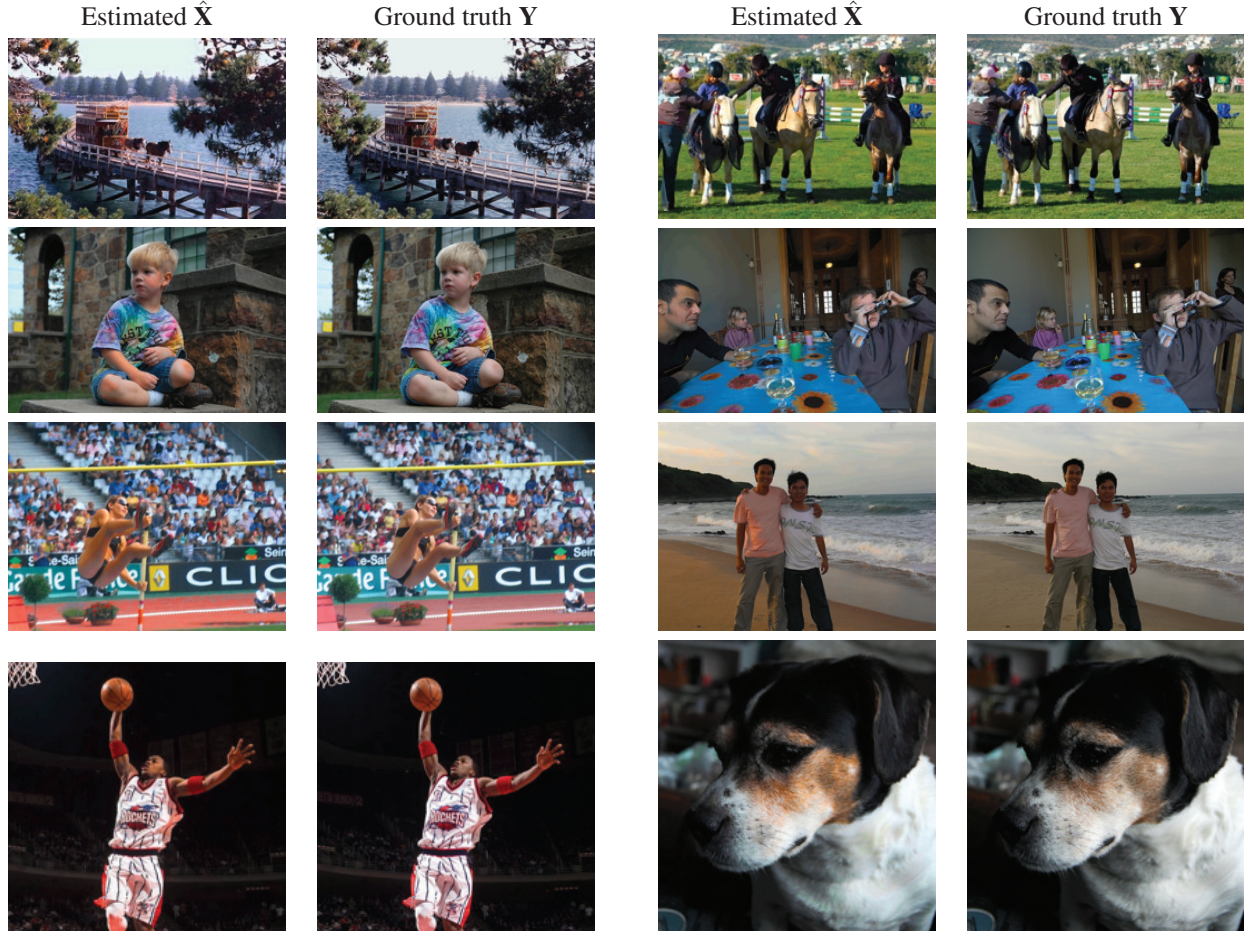


Figure 10. $\hat{\mathbf{X}}$ vs. \mathbf{Y} . Comparing estimated image from CWAN $\hat{\mathbf{X}}$ (Left), with the ground truth image \mathbf{Y} (Right).

Given that CWAN can generate more colorful, vibrant and lively images than the original image, this feature can be used to enhance pale, dull and faded colors as seen from the examples in Fig. 11. Note that CWAN is not a colorization model, and therefore it cannot colorize gray scale images. CWAN still requires degraded colors to be present in the scene in order to enhance the image. For example, the face in the third row in Fig. 11, barely contains any skin color-tone in the original image, allowing limited improvement by CWAN in enhancing the colors.

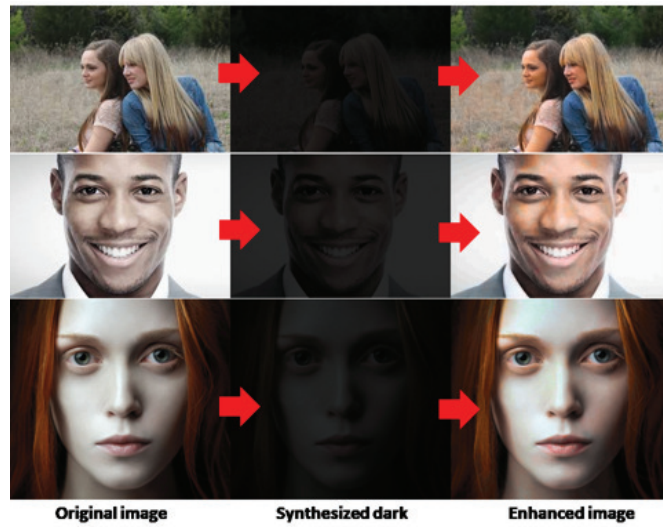


Figure 11. Enhancing pale color-tone faces randomly selected from the internet. First column is the original face. The second column is synthesized to low-light image using the method in [8]. The third column is the enhanced image, representing more colorful skin tones.

References

- [1] C. Chen, Q. Chen, J. Xu, and V. Koltun. Learning to see in the dark. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018. 2, 4, 5, 6, 7, 8, 9, 10, 11, 12
- [2] E. Culurciello and A. Canziani. e-Lab video data set. <https://engineering.purdue.edu/elab/eVDS/>, 2017. 2
- [3] X. Dong, G. Wang, Y. Pang, W. Li, J. Wen, W. Meng, and Y. Lu. Fast efficient algorithm for enhancement of low lighting video. In *2011 IEEE International Conference on Multimedia and Expo*, pages 1–6, July 2011. 4, 5, 6, 7, 8, 9, 10, 11, 12
- [4] G. Eilertsen, J. Kronander, G. Denes, R. Mantiuk, and J. Unger. HDR image reconstruction from a single exposure using deep CNNs. *ACM Transactions on Graphics*, 36(6), 2017. 11
- [5] J. W. Feifan Lv, Feng Lu and C. Lim. MBLLEN: Low-light image/video enhancement using CNNs. *British Machine Vision Conference*, 2018. 4, 5, 6, 7, 8, 9, 10, 11, 12
- [6] X. Guo, Y. Li, and H. Ling. LIME: Low-light image enhancement via illumination map estimation. *IEEE Transactions on Image Processing*, 26(2):982–993, 2017. 2, 4, 5, 6, 7, 8, 9, 10, 11, 12
- [7] C.-H. Lee, J.-L. Shih, C.-C. Lien, and C.-C. Han. Adaptive multiscale retinex for image contrast enhancement. In *International Conference on Signal-Image Technology & Internet-Based Systems*, pages 43–50. IEEE, 2013. 4, 5, 6, 7, 8, 9, 10, 11, 12
- [8] C. Li, J. Guo, F. Porikli, and Y. Pang. LightenNet: A convolutional neural network for weakly illuminated image enhancement. *Pattern Recognition Letters*, 104:15–22, 2018. 2, 4, 5, 6, 7, 8, 9, 10, 11, 12, 14
- [9] M. Li, J. Liu, W. Yang, X. Sun, and Z. Guo. Structure-revealing low-light image enhancement via robust retinex model. *IEEE Transactions on Image Processing*, 27(6):2828–2841, 2018. 4, 5, 6, 7, 8, 9, 10, 11, 12
- [10] K. G. Lore, A. Akintayo, and S. Sarkar. LLNet: A deep autoencoder approach to natural low-light image enhancement. *Pattern Recognition*, 61:650–662, 2017. 1, 3
- [11] X. Ren, M. Li, W.-H. Cheng, and J. Liu. Joint enhancement and denoising method via sequential decomposition. In *IEEE International Symposium on Circuits and Systems*, pages 1–5. IEEE, 2018. 4, 5, 6, 7, 8, 9, 10, 11, 12
- [12] Y. Ren, Z. Ying, T. H. Li, and G. Li. Lecarm: Low-light image enhancement using the camera response model. *IEEE Transactions on Circuits and Systems for Video Technology*, 29(4):968–981, 2018. 4, 5, 6, 7, 8, 9, 10, 11, 12
- [13] Z. Ying, G. Li, and W. Gao. A bio-inspired multi-exposure fusion framework for low-light image enhancement. *arXiv preprint arXiv:1711.00591*, 2017. 4, 5, 6, 7, 8, 9, 10, 11, 12
- [14] Z. Ying, G. Li, Y. Ren, R. Wang, and W. Wang. A new low-light image enhancement algorithm using camera response model. In *IEEE International Conference on Computer Vision Workshops*, pages 3015–3022, Oct 2017. 4, 5, 6, 7, 8, 9, 10, 11, 12