

DALE : Dark Region-Aware Low-light Image Enhancement

Dokyeong Kwon*
dkdk6638@gmail.com

Guisik Kim*
specialre@naver.com

Junseok Kwon†
jskwon@cau.ac.kr

School of
Computer Science and Engineering
Chung-Ang University
Seoul, Korea

Abstract

In this paper, we present a novel low-light image enhancement method called dark region-aware low-light image enhancement (DALE), where dark regions are accurately recognized by the proposed visual attention module and their brightness are intensively enhanced. Our method can estimate the visual attention in an efficient manner using super-pixels without any complicated process. Thus, the method can preserve the color, tone, and brightness of original images and prevents normally illuminated areas of the images from being saturated and distorted. Experimental results show that our method accurately identifies dark regions via the proposed visual attention, and qualitatively and quantitatively outperforms state-of-the-art methods.

1 Introduction

Real-world images for outdoor scenes typically contain low-light areas, especially if the images are captured during nighttime or there exists backlit. However, using these low-light images, conventional computer vision algorithms (*e.g.*, object detection and tracking) cannot produce accurate results, because low-light regions cause images to lose local details and significantly reduce image quality. Therefore, low-light image enhancement is essential to prevent conventional computer vision algorithms from degrading their performance.

Low-light image enhancement has a long history. For example, Pizer *et al.* [28] enhanced the brightness and contrast of images based on histogram equalization [43]. The methods in [14, 34] introduced retinex theory [15] for low-light enhancement and used illumination information. High dynamic range (HDR)-based methods [6, 23] have been proposed to enhance the brightness of images, where HDR requires to combine multiple images with different exposures for the same scene.

Recently, several methods successfully implemented deep neural networks to solve low-level computer vision problems (*e.g.*, image super-resolution [19], dehazing [8], and deraining [31]). Low-light image enhancement problems have been also addressed in deep learning frameworks [13]. However, it is nontrivial to train deep neural networks for low-level vision problems, because the networks require be a large number of paired data (*i.e.*, image and

*Indicates equal contributions. † Corresponding author.

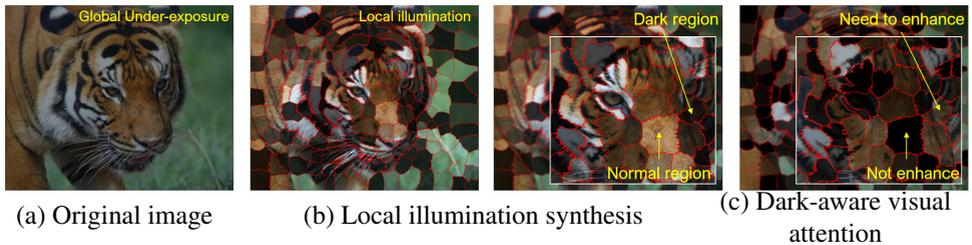


Figure 1: **Example of the proposed visual attention for low-light enhancement.** (b) Our method applies a different level of local illumination to each superpixel of a given image. (c) Our method can generate the attention map, where dark areas that need to be enhanced have large values and bright areas have small values.

ground truth) for the training. To collect a large number of paired data, in super-resolution problems, low-resolution images are synthetically generated by applying noise and blur to high-resolution images. Low-light image enhancement methods [18, 69] also synthetically generate paired data by applying several different levels of global illumination to the images. The methods using globally illuminated training images can improve the overall brightness and contrast of images. However, they cannot accurately enhance the brightness of locally illuminated regions, which are frequently contained in real-world images. Applying global illumination is more problematic, if both bright and dark areas exist simultaneously in an image; bright areas of original images can be over-saturated by global illumination. To solve this problem, methods in [11, 21, 22] proposed to synthesize realistic illumination. EnlightenGAN [11] performed low-light image enhancement in an unsupervised manner without paired data. Methods in [2, 11, 21] employed attention modules for low-level vision tasks to focus on the areas, which need to be improved. However, in low-light image enhancement, the attention mechanism has not been actively studied and was used only to estimate the illumination channel as an attention map [11, 69].

In this paper, we construct a new dataset, which can be used to learn a visual attention map. Then, the proposed method enhances the brightness of dark areas, which can be recognized by the aforementioned attention map. We synthesize differently illuminated superpixels and generate a locally illuminated image dataset, as shown in Fig.1(b). As a result, our method can produce more accurate low-light enhanced images than existing deep-learning based methods. Fig.1(c) shows an estimated visual attention map, where dark areas are accurately recognized and described using large values.

The contributions of our method is as follows:

- We present a new attention module to recognize dark areas. For this purpose, we synthesize images to train the attention map, where each superpixel of the images has a different local illumination. Experimental results demonstrate the effectiveness of the proposed dark-aware visual attention.
- We propose a novel low-light enhancement method using the proposed dark-aware visual attention. We call this method dark region-aware low-light image enhancement (DALE). Our method can intensively enhance the brightness of dark areas, while preserving the brightness of other areas.
- We exhaustively conduct the experiments to demonstrate the effectiveness of the proposed method and provide a locally illuminated image dataset, which is used in all experiments. This dataset will be publicly available to re-train conventional low-light enhancement methods and improve their accuracy.

2 Related Work

2.1 Low-light image enhancement

Non-deep learning-based techniques such as histogram equalization and its variations [28] have been commonly adopted for low-light enhancement, in which contrast-limited adaptive histogram equalization [43] showed the best performance. Retinex theory was employed for low-light enhancement in the single-scale retinex-based method [15]. Multi-scale retinex with color restoration [17] enhanced the single-scale retinex method by adding the color restoration process. The haze-model was employed to solve low-light enhancement problems [5], while low-light images are similar to reversed hazy images. Lime [8] used a structure prior to enhance the brightness of low-light regions and outperformed naturalness preserved enhancement [67], multi-deviation fusion [6], and simultaneous reflection and illumination estimation [0]. LECARM [62] presented a new camera response model based on the exposure ratio estimation for each pixel. Inspired by the human visual system, BIMEF [41] proposed a multi-exposure fusion technique for low-light enhancement. Deep learning methods, MSRNet [65] and LLNet [20], proposed stacked sparse denoising autoencoders to solve low-light enhancement problems. LightenNet [18] estimated the illumination map using deep neural networks to improve the brightness of images, while RetinexNet [39] estimated both reflection and illumination maps, inducing further improvement. DeepFuse [29] solved low-light enhancement problems by combining HDR with an unsupervised learning method. MBLEN [27] improves performance using multi-branch fusion methods. Lv and Lu [21] improved the accuracy of low-light enhancement using attention mechanisms and produced the illumination map. In contrast, our method directly recognizes dark areas using visual attention modules. Thus, our method more intensively enhances the brightness of dark areas.

Most deep-learning-based methods have a difficult in gathering sufficient training images. To address this problem, methods in [18, 65] synthesized training data by applying global illumination to the entire image. However, these methods are not suitable for real-world images, which contain both low-light and normal illumination areas concurrently. Recently, low-lightGAN [13] proposed a generative adversarial network (GAN) in low-light enhancement. EnlighthGAN [40] present a low-light enhancement method, which can be trained without paired dataset in an unsupervised manner. In contrast to these methods, our method synthesizes training data using superpixels efficiently.

2.2 Attention Mechanism

Attention mechanisms have been widely used in many of recent computer vision tasks. This attention mechanism stem from the human perception system, where the human brain typically focus on important areas. In deep learning, the attention mechanism has been implemented using convolution layers, which results in channel and spatial attention networks. There exists a different type of attention mechanisms, which focuses on an image itself instead of convolution layers. For example, in deraining problems, the method in [60] recognized raindrop areas using the attention mechanism.

In low-light image enhancement problems, attention-guided method in [21] used the attention map and EnlighthGAN [40] proposed a self-regularized attention map to enable unsupervised learning. In contrast to the aforementioned methods that estimate the attention maps to obtain exposure and illumination information, our method directly uses the attention map to enhance the brightness of dark areas.

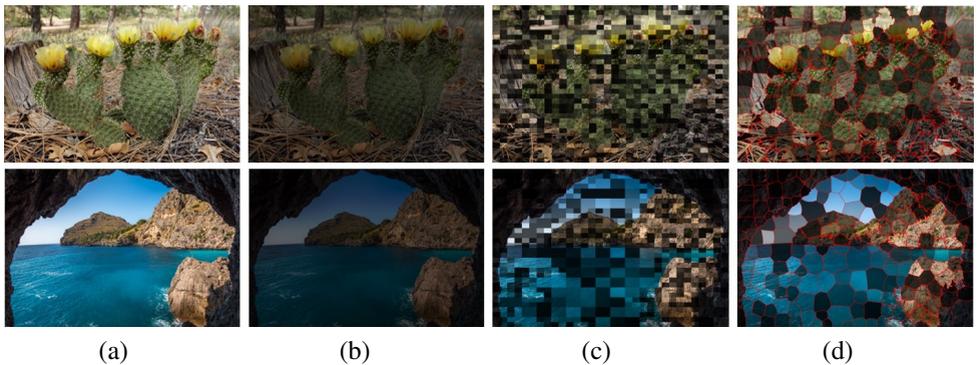


Figure 2: **Proposed local illumination method for training data synthesis.** (a) Input image, (b) Global illumination method, (c) Quad-tree-based method [13], and (d) Our superpixel-based method.

3 The Proposed Method

We construct a new low-light driven training dataset (Section 3.1). Then, we present a novel attention network to recognize dark areas (Section 3.2). In Section 3.3, we describe the proposed low-light enhancement method with detailed network architectures and loss functions.

3.1 Low-light Driven Training Dataset

To construct a large set of paired data, the deraining method in [40] synthesized raindrops. The dehazing methods in [9, 17] synthesized haze images using depth information through single image depth estimation. For low-light enhancement, methods in [18, 65] synthesized low-light images. However, it is nontrivial for these methods to accurately describe illumination in real-world. Most conventional methods synthesized global illumination as shown in Fig. 2(b). However, these methods are only suitable for under exposure images and cannot handle low-light images properly if dark areas are caused by natural illumination. Under the real-world environment, global and local illumination can exist simultaneously in the same scene. If we attempt to apply global illumination to real-world images, normally illuminated areas in the images can be easily saturated. To solve this problem, low-lightGAN [13] presented a quad-tree local illumination synthesis method, as shown in Fig. 2(c).

We propose a local illumination synthesis method based on superpixels, as shown in Fig. 2(d). We apply a randomly different level of illumination to each superpixel and synthesize both the low-light and normal illumination areas. In addition, the proposed superpixel-based method can synthesize local illumination according to object boundaries, because superpixels describe object shapes. The local illumination synthesis using superpixel is formulated as follows:

$$I_{local} = SLIC(I) \times L, \text{ for } L = \{0.1, 0.2, \dots, 0.9, 1.0\}, \quad (1)$$

where the SLIC function [11] outputs superpixels of image I . In (1), L denotes the illumination weight. If $L = 1.0$, the original brightness is maintained. If $L = 0.1$, the corresponding superpixel is considerably darkened. This synthetic data is used to train the dark-aware attention network, which is explained in the next section.

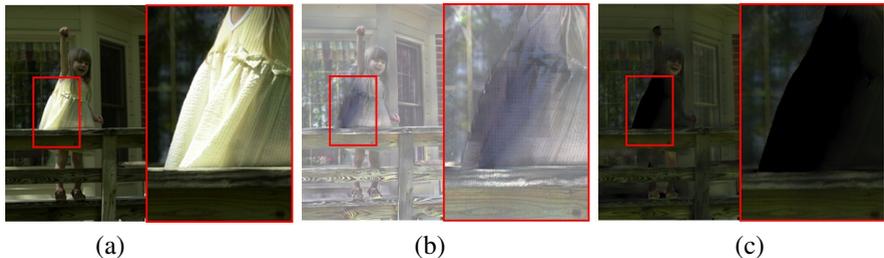


Figure 3: **Proposed visual attention map.** (a) Input image, (b) Quad-tree-based method [13], and (c) Our dark-aware attention method.

3.2 Dark-aware Visual Attention

The methods in [7, 8, 34] used existing attention mechanisms and produced illumination maps based on retinex theory. Lv and Lu [21] estimated attention maps based on max channels in original and low-light images. CWAN [2] created binary mask maps for the foreground color and used the maps for training. EnlightenGAN [10] used illumination channels to estimate a self-regularized attention map.

Unlike conventional attention-based low-light enhancement methods, our method can recognize visually dark areas rather than illumination areas based on retinex theory. Thus, our method is a visual attention method (*i.e.*, dark-aware attention network), which focuses on dark areas for low-light image enhancement. To train the proposed dark-aware attention network in a supervised manner, we synthesize the ground-truth attention map I_{VA} , as follows:

$$I_{VA} = I - I_{local}, \quad (2)$$

where I and I_{local} in (2) denote the original image and locally illuminated image, respectively. Local illumination and superpixels with various forms enable the proposed attention network to learn various types of bright and dark areas during the training. Fig.3(c) shows the estimated attention map of the proposed method. Our attention map accurately differentiates between dark areas and bright regions, where bright regions are represented using small values (*i.e.*, red box) and dark areas are represented using large values. The conventional method [13] cannot accurately estimate visual attention maps with blocking artifacts, as shown in Fig.3(b).

3.3 Dark Region-aware Low-light Enhancement Network

The proposed low-light enhancement network consists of visual attention and enhancement networks, as shown in Fig.4. The attention network produces the attention map that can recognize dark areas, whereas the enhancement network outputs low-light enhanced images.

Visual Attention Network (VAN). The proposed VAN adopts the U-Net [33] structure (*i.e.*, encoder and decoder) as a backbone network. The first convolution layer has the kernel of 3×3 size with stride 1. The Encoder has three convolution layers, residual blocks, and two down-sample layers, as shown in Fig.4. Residual Blocks consist of convolution layers with the kernels of 1×1 size, ReLUs, and squeeze-and-excitation blocks [9], which have different dilation factors (*i.e.*, 3, 2, and 1). The decoder has three convolution layers, residual blocks, and two up-sample layers. Residual Blocks use different dilation factors (*i.e.*, 1, 2, and 3).

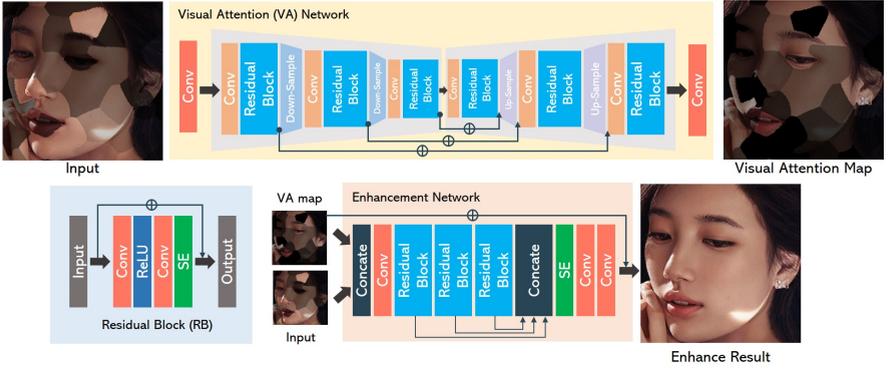


Figure 4: **The proposed network architecture** consisting of visual attention network (VAN), enhancement network (EN), and residual block (RB).

We compute l_2 loss between the estimated attention map $VA(I_{local})$ and ground truth I_{VAGT} at the pixel level:

$$\mathcal{L}_a = \|VA(I_{local}) - I_{VAGT}\|_2, \quad (3)$$

where I_{local} in (1) denotes the low-light image with locally illuminated areas and the function VA estimates visual attention maps. We also calculate perceptual loss, which measures the similarity between $VA(I_{local})$ and I_{VAGT} at the feature level:

$$\mathcal{L}_p = \|\phi(VA(I_{local}) + I_{local}) - \phi(I_{GT})\|_1, \quad (4)$$

where ϕ is 16-th feature map obtained by the pre-trained VGG-16 network [4] and I_{GT} denotes ground truth for the low-light enhanced image. Then, the total loss for VAN is designed as follows.

$$\mathcal{L}_{VAN} = \lambda_1 \mathcal{L}_a + \lambda_2 \mathcal{L}_p, \quad (5)$$

where λ_1 and λ_2 are weighting hyper-parameters.

Enhancement Network (EN). The proposed EN enhances the brightness of low-light images using the estimated visual attention maps. The EN takes the concatenation of low-light image and visual attention map as an input. Similar to the VAN, all convolution layers have the kernel of 3×3 size with stride 1. Three residual blocks use different dilation factors from 3 to 1, while we concatenate all residual blocks to fuse the information.

We compute l_2 loss between the low-light enhanced image $EN(I_{EN})$ and ground truth I_{GT} at the pixel level:

$$\mathcal{L}_e = \|EN(I_{EN}) - I_{GT}\|_2, \quad (6)$$

where $I_{EN} = VAN(I_{local}) + I_{local}$. We also calculate perceptual loss, which measures the similarity between $EN(I_{EN})$ and I_{GT} at the feature level:

$$\mathcal{L}_{ep} = \|\phi(EN(I_{EN})) - \phi(I_{GT})\|_1. \quad (7)$$

The total variation loss aims to make output images spatially smooth:

$$\mathcal{L}_{tv} = \frac{1}{CHW} \|\nabla_x EN(I_{EN}) + \nabla_y EN(I_{EN})\|^2, \quad (8)$$



where ∇_x and ∇_y differentiate the images via the x and y directions. C , H , and W are the channel, height, and width of the enhanced image, respectively. Then, the total loss for EN is designed as follows.

$$\mathcal{L}_{EN} = \lambda_1 \mathcal{L}_e + \lambda_2 \mathcal{L}_{ep} + \lambda_3 \mathcal{L}_{lv}. \quad (9)$$

4 Experiments

For training, we synthesized locally illuminated images using the DIV2K [36] and Flickr2K [36] datasets, as explained Section 3.1. The proposed visual attention network was trained with learning rate of $1e-5$ for 130 epochs, whereas the proposed enhancement network was trained with learning rate of $1e-5$ for 40 epochs. We trained the whole network for 24 hours on NVIDIA GeForce GTX TITAN Xp GPUs. Hyper-parameters λ_1 and λ_2 in (5) were set to 0.5 and 1, respectively. Hyper-parameters λ_1 , λ_2 , and λ_3 in (9) were set to 1, 5, and 1, respectively. We randomly cropped approximately 3500 images with 2K resolutions into 240×240 patches. We compared our method (DALE) with state-of-the-art methods including NPE [33], LIME [8], MEF [24], and DICM [16].

4.1 Ablation Study

Refinement via adversarial learning. The proposed visual attention intensively enhances the brightness of dark areas, while it preserves the brightness of other regions. Thus, if the visual attention map is accurately estimated, the following enhancement network has little effect on the performance of low-light enhancement and parameters of the network are

Table 1: **Quantitative comparison with state-of-the-art low-light enhancement methods using NIQE and BRISQUE.** Red and blue numbers denote the best and second best results, respectively.

NIQE / BRISQUE	DICM	LIME	MEF	NPE
LIME [8]	3.63 / 26.8	4.35 / 22.3	3.83 / 24.1	3.84 / 26.1
BIMEF [14]	3.38 / 26.8	3.55 / 23.2	3.13 / 19.3	3.5 / 24.5
RetinexNet [19]	4.31 / 26.7	4.91 / 26.1	4.90 / 26.0	4.07 / 26.9
EnlightenGAN [14]	3.05 / 26.3	3.37 / 20.6	2.89 / 23.6	3.34 / 27.3
DALE	3.78 / 21.4	4.33 / 22.2	4.01 / 22.8	3.38 / 22.1
DALEGAN	3.61 / 22.2	4.16 / 23.5	3.80 / 23.9	3.31 / 19.6

Table 2: **Quantitative comparison with state-of-the-art low-light enhancement methods using LOE.** Red and blue numbers denote the best and second best results, respectively.

LOE	DICM	LIME	MEF	NPE
LIME [8]	1260.8	1323.8	1079.4	1119.6
BIMEF [14]	351.82	478.57	325.86	308.12
RetinexNet [19]	1565.8	1882.5	1777.4	1224.5
EnlightenGAN [14]	1318.8	1361.5	1141.9	1346.2
DALE	888.7	810	829.2	678.7
DALEGAN	920.91	849.6	892.4	714.6

hardly updated. To solve this problem, we trained the network in an adversarial manner (*i.e.*, DALEGAN). We used the enhancement network as a generative network and employed PatchGAN [14, 27] as a discriminator network. Fig.5 shows that the enhancement network with adversarial learning further enhanced the brightness of low-light region.

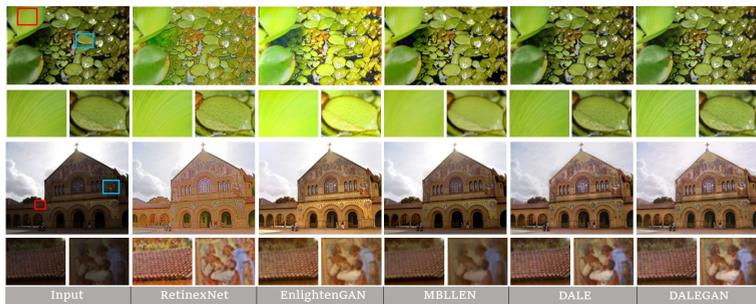
Dark-aware visual attention. Fig.6 shows examples of estimated visual attention maps. The proposed visual attention method accurately recognized dark areas in low-light images. The method did not focus on bright areas such as sky regions, because the regions appear clearly and there exist sufficient brightness. In the last column of Fig.6, the visual attention method produced very small values for a tree or left person, because they received direct sunlight. In contrast, the method produced large values for a right person, because he was in dark areas.

4.2 Comparison with Other Methods

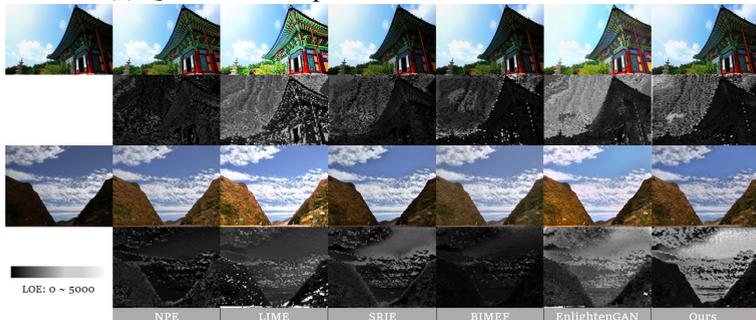
We quantitatively evaluated several methods in terms of lightness order error (LOE) [58], naturalness image quality evaluator (NIQE) and blind/referenceless image spatial quality evaluator (BRISQUE) [25] *i.e.*, no-reference image quality evaluation metric in [26]. Tables 1 and 2 show the low-light enhancement results for five standard benchmark datasets, which contain real-world low-light images. The proposed DALE and EnlightenGAN showed state-of-the-art performance in terms of NIQE and BRISQUE, where overall image quality has been improved. In terms LOE, BIMEF produced the best results. Please note that LOE evaluates the degree of light distortion. However, it produces good scores (*i.e.*, small values), even though the brightness of images is not much improved.

Thus, BIMEF still has visually low-light regions, as shown in the second row of Fig.7(b). In contrast, LIME, EnlightenGAN, and the proposed DALE produced qualitatively better low-light enhance results than BIMEF, although they have very high values of LOE.

Fig.7(a) qualitatively compared saturation and details of low-light enhanced images. Fig.7(b) qualitatively compared lightness distortion using LOE maps in low-light enhanced images. As shown in the third row of Fig.7(c), LIME over-improved the brightness of images, which results in saturation problems. BIMEF and MBLLEN produced clear images but



(a) Qualitative comparison on saturation and details



(b) Qualitative comparison on lightness distortion



(c) Qualitative comparison with other methods

Figure 7: **Qualitative comparison with state-of-the-art-methods.** (b) first and third rows: input images, second and fourth rows: estimated LOE maps.

lacked the brightness. EnlightenGAN produced good image enhancement results. However, some colors distorted. In contrast these methods, the proposed DALE qualitatively outperforms other state-of-the-art methods.

5 Conclusion

In this paper, we present a novel low-light enhancement method (DALE) based on a new visual attention network that can recognize dark regions. We synthesize different local illumination for each super-pixel and accurately estimate the brightness of low-light images using synthetic training images. Experimental results demonstrate that our method outperforms state-of-the-art methods.

Acknowledgments: This research was supported by the the Seoul R&BD Program (CY190032) and the ITRC program (IITP-2020-2018-0-01799).

References

- [1] Radhakrishna Achanta, Appu Shaji, Kevin Smith, Aurelien Lucchi, Pascal Fua, and Sabine Süsstrunk. Slic superpixels compared to state-of-the-art superpixel methods. *IEEE TPAMI*, 34(11):2274–2282, 2012.
- [2] Yousef Atoum, Mao Ye, Liu Ren, Ying Tai, and Xiaoming Liu. Color-wise attention network for low-light image enhancement. *arXiv preprint arXiv:1911.08681*, 2019.
- [3] Bolun Cai, Xiangmin Xu, Kui Jia, Chunmei Qing, and Dacheng Tao. Dehazenet: An end-to-end system for single image haze removal. *TIP*, 25(11):5187–5198, 2016.
- [4] K. Chatfield, K. Simonyan, A. Vedaldi, and A. Zisserman. Return of the devil in the details: Delving deep into convolutional nets. In *BMVC*, 2014.
- [5] 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 *ICME*, 2011.
- [6] Xueyang Fu, Delu Zeng, Yue Huang, Yinghao Liao, Xinghao Ding, and John Paisley. A fusion-based enhancing method for weakly illuminated images. *Signal Processing*, 129:82–96, 2016.
- [7] 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.
- [8] Xiaojie Guo, Yu Li, and Haibin Ling. LIME: Low-light image enhancement via illumination map estimation. *TIP*, 26(2):982–993, 2017.
- [9] Jie Hu, Li Shen, and Gang Sun. Squeeze-and-excitation networks. In *CVPR*, 2018.
- [10] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and Alexei A Efros. Image-to-image translation with conditional adversarial networks. In *CVPR*, 2017.
- [11] 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. *arXiv preprint arXiv:1906.06972*, 2019.
- [12] 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. *TIP*, 6(7):965–976, 1997.
- [13] Guisik Kim, Dokyeong Kwon, and Junseok Kwon. Low-LightGAN: Low-light enhancement via advanced generative adversarial network with task-driven training. In *ICIP*, 2019.
- [14] Ron Kimmel, Michael Elad, Doron Shaked, Renato Keshet, and Irwin Sobel. A variational framework for retinex. *IJCV*, 52(1):7–23, 2003.
- [15] Edwin H Land and John J McCann. Lightness and retinex theory. *Josa*, 61(1):1–11, 1971.
- [16] Chulwoo Lee, Chul Lee, and Chang-Su Kim. Contrast enhancement based on layered difference representation. In *ICIP*, 2012.

- [17] Boyi Li, Wenqi Ren, Dengpan Fu, Dacheng Tao, Dan Feng, Wenjun Zeng, and Zhangyang Wang. Reside: A benchmark for single image dehazing. *arXiv preprint arXiv:1712.04143*, 1, 2017.
- [18] Chongyi Li, Jichang Guo, Fatih Porikli, and Yanwei Pang. LightenNet: A convolutional neural network for weakly illuminated image enhancement. *PRL*, 104(C):15–22, 2018.
- [19] Bee Lim, Sanghyun Son, Heewon Kim, Seungjun Nah, and Kyoung Mu Lee. Enhanced deep residual networks for single image super-resolution. In *CVPR*, 2017.
- [20] Kin Gwn Lore, Adedotun Akintayo, and Soumik Sarkar. LLNet: A deep autoencoder approach to natural low-light image enhancement. *PR*, 61:650–662, 2017.
- [21] Feifan Lv and Feng Lu. Attention-guided low-light image enhancement. *arXiv preprint arXiv:1908.00682*, 2019.
- [22] Feifan Lv, Feng Lu, Jianhua Wu, and Chongsoon Lim. Mblen: Low-light image/video enhancement using cnns. In *BMVC*, page 220, 2018.
- [23] Kede Ma and Zhou Wang. Multi-exposure image fusion: A patch-wise approach. In *ICIP*, 2015.
- [24] Kede Ma, Kai Zeng, and Zhou Wang. Perceptual quality assessment for multi-exposure image fusion. *TIP*, 24(11):3345–3356, 2015.
- [25] Anish Mittal, Anush Krishna Moorthy, and Alan Conrad Bovik. No-reference image quality assessment in the spatial domain. *TIP*, 21(12):4695–4708, 2012.
- [26] Anish Mittal, Rajiv Soundararajan, and Alan C Bovik. Making a “completely blind” image quality analyzer. *SPL*, 20(3):209–212, 2013.
- [27] Takeru Miyato, Toshiki Kataoka, Masanori Koyama, and Yuichi Yoshida. Spectral normalization for generative adversarial networks. In *ICLR*, 2018.
- [28] Stephen M Pizer, E Philip Amburn, John D Austin, Robert Cromartie, Ari Geselowitz, Trey Greer, Bart ter Haar Romeny, John B Zimmerman, and Karel Zuiderveld. Adaptive histogram equalization and its variations. *Comput. Gr. Image Process.*, 39(3):355–368, 1987.
- [29] K Ram Prabhakar, V Sai Srikar, and R Venkatesh Babu. Deepfuse: A deep unsupervised approach for exposure fusion with extreme exposure image pairs. In *ICCV*, 2017.
- [30] Rui Qian, Robby T Tan, Wenhan Yang, Jiajun Su, and Jiaying Liu. Attentive generative adversarial network for raindrop removal from a single image. In *CVPR*, 2018.
- [31] Dongwei Ren, Wangmeng Zuo, Qinghua Hu, Pengfei Zhu, and Deyu Meng. Progressive image deraining networks: a better and simpler baseline. In *CVPR*, 2019.
- [32] Yurui Ren, Zhenqiang Ying, Thomas H. Li, and Ge Li. LECARM: Low-light image enhancement using the camera response model. *TCSVT*, 29(4):968–981, 2019.
- [33] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, pages 234–241. Springer, 2015.

- [34] Chih-Tsung Shen and Wen-Liang Hwang. Color image enhancement using retinex with robust envelope. In *ICIP*, 2009.
- [35] Liang Shen, Zihan Yue, Fan Feng, Quan Chen, Shihao Liu, and Jie Ma. MSR-net: Low-light image enhancement using deep convolutional network. *arXiv preprint arXiv:1711.02488*, 2017.
- [36] Radu Timofte, Eirikur Agustsson, Luc Van Gool, Ming-Hsuan Yang, and Lei Zhang. Ntire 2017 challenge on single image super-resolution: Methods and results. In *CVPRW*, 2017.
- [37] Shuhang Wang, Jin Zheng, Hai-Miao Hu, and Bo Li. Naturalness preserved enhancement algorithm for non-uniform illumination images. *TIP*, 22(9):3538–3548, 2013.
- [38] Shuhang Wang, Jin Zheng, Hai-Miao Hu, and Bo Li. Naturalness preserved enhancement algorithm for non-uniform illumination images. *TIP*, 22(9):3538–3548, 2013.
- [39] Chen Wei, Wenjing Wang, Wenhan Yang, and Jiaying Liu. Deep retinex decomposition for low-light enhancement. In *BMVC*, 2018.
- [40] Wenhan Yang, Robby T Tan, Jiashi Feng, Jiaying Liu, Zongming Guo, and Shuicheng Yan. Deep joint rain detection and removal from a single image. In *CVPR*, 2017.
- [41] Zhenqiang Ying, Ge Li, and Wen Gao. A bio-inspired multi-exposure fusion framework for low-light image enhancement. *arXiv preprint arXiv:1711.00591*, 2017.
- [42] Zhenqiang Ying, Ge Li, and Wen Gao. A bio-inspired multi-exposure fusion framework for low-light image enhancement. *arXiv preprint arXiv:1711.00591*, 2017.
- [43] Karel Zuiderveld. Graphics gems iv. chapter Contrast Limited Adaptive Histogram Equalization, pages 474–485. Academic Press Professional, Inc., 1994.