Unsupervised Low-light Image Enhancement with Decoupled Networks

Wei Xiong∗, Ding Liu†, Xiaohui Shen†, Chen Fang† and Jiebo Luo∗
∗ Department of Computer Science, University of Rochester, USA
† ByteDance Inc., USA
‡ Tencent

Abstract—In this paper, we tackle the problem of enhancing real-world low-light images with significant noise in an unsupervised fashion. Conventional unsupervised approaches focus primarily on illumination or contrast enhancement but fail to suppress the noise in real-world low-light images. To address this issue, we decouple this task into two sub-tasks: illumination enhancement and noise suppression. We propose a two-stage, fully unsupervised model to handle these tasks separately. In the noise suppression stage, we propose an illumination-aware denoising model so that real noise at different locations is removed with the guidance of the illumination conditions. To facilitate the unsupervised training, we construct pseudo triplet samples and propose an adaptive content loss correspondingly to preserve contextual details. To thoroughly evaluate the performance of the enhancement models, we build a new unpaired real-world low-light enhancement dataset. Extensive experiments show that our proposed method outperforms the state-of-the-art unsupervised methods concerning both illumination enhancement and noise reduction.

I. INTRODUCTION

Real-world low-light image enhancement [1]–[6] is challenging since images captured under real low-light conditions usually exhibit low illumination, contain heavy noise and degraded contextual details. Enhancing these images requires adjusting illumination and color to normal conditions and, more importantly, simultaneously suppressing the real noise while preserving the details. Conventional methods for this task primarily focus on adjusting the contrast via a fixed tonemapping [1], [3], [7], resulting in limited performance. Data-driven methods learn illumination enhancement patterns from data [8]–[19]. However, many of them heavily rely on low-light and normal-light image pairs, which are expensive or even impossible to obtain in real-world scenarios.

Recently, unsupervised learning-based methods [11], [17], [19]–[22] have been developed for image enhancement to eliminate the reliance on paired training data. A severe limitation of these methods is that although they can improve the illumination of low-light images, they fail to reduce the noise that ubiquitously exists in images taken under real-world low-light conditions. When these methods are adopted to build a practical real-world image enhancement system, a denoising procedure is usually necessary as a separate post-processing step. Although noise removal approaches have been widely studied [23]–[34], they have not yet been perfectly integrated with the low-light enhancement approaches under the unsupervised learning setting. Therefore, the enhancement and denoising performance of the existing systems still cannot meet the real user requests.

To address these issues, we explicitly decouple the enhancement task into two sub-tasks: 1) illumination enhancement and 2) noise suppression. We propose a two-stage model to handle each sub-task separately. Specifically, in Stage I, a Retinex-based network is trained in an unsupervised fashion to enhance the illumination of low-light images while preserving the contextual details. In Stage II, an unsupervised illumination-aware denoising model is proposed. Our denoising model is explicitly guided by the illumination conditions of the original low-light image and the enhanced image from Stage I. With such a design, different image regions can be denoised with appropriate intensities to preserve texture details and prevent over-smoothness. Inspired by the success of pseudo labeling methods in semi-supervised learning [35], [36], in this work, we propose to construct pseudo triples, i.e., a pseudo low-light image, a pseudo image after illumination enhancement, and a real noise-free normal-light image with the same content, to facilitate the unsupervised training for image denoising. To use these pseudo samples more effectively, we especially design an adaptive content loss to enhance the contextual details based on the illumination conditions.

Existing low-light enhancement datasets usually consist of images without significant real noise. To thoroughly evaluate the models for enhancing real-world low-light images, we build an Unpaired Real-world Low-light Image enhancement dataset (URL). Our dataset comprises 1) diverse indoor and outdoor scenes captured under real-world low-light conditions with varying levels of noise and 2) normal-light images collected from existing data galleries. Extensive experiments on public datasets and our dataset show that our method outperforms the state-of-the-art unsupervised methods in terms of both illumination enhancement and noise suppression.

Our contributions are as follows. First, we propose a decoupled framework for unsupervised real-world low-light image enhancement. The primary novelty comes from the unsupervised illumination-aware denoising model. We construct pseudo triples and propose an adaptive content loss to denoise regions guided by illumination conditions. Second, we build an unpaired low-light image enhancement dataset containing scenes with various noise levels as an essential complement to the existing low-light enhancement datasets.
II. OUR APPROACH

As shown in Fig. 1 (a), our unsupervised model for real-world low-light image enhancement consists of two stages. In Stage I, we perform illumination enhancement on the low-light images while preserving contextual details. In Stage II, we propose an illumination-aware denoising model to suppress the noise existing in the output image of Stage I.

A. Stage I: Illumination Enhancement

Given a noisy low-light image \( I_l \), our goal in this stage is to learn a model \( G_e \) to generate an enhanced image \( I_e \) with proper illumination, contrast, color, and realistic content details. Recent work [15] on low-light image enhancement suggests that illumination maps for natural images usually have relatively simple forms. Thus using Retinex-based illumination modeling can facilitate the learning process, leading to an enhancer with a better generalization ability. Inspired by this, we adopt a Retinex-based network as our generator to enhance the low-light images. An image \( I \) can be modeled as \( I = S \odot R \), where \( S \) is the illumination, \( \odot \) denotes element-wise multiplication, and \( R \) is the reflectance. Similar to [13], [15], we regard the reflectance as a well-exposed image \( I_e \), then we have \( I_l = S \odot I_e \). In reverse, the enhanced image \( I_e \) can be recovered from the low-light image \( I_l \) given the predicted illumination map \( S \). As shown in Fig. 1, we use the generator \( G_e \) to estimate an illumination map (with RGB channels) \( S = G_e(I_l) \) from low-light image \( I_l \). Then we obtain the enhanced image \( I_e = I_l/S \), where / is element-wise division.

Model Architecture. As shown in Fig. 1, our generator in Stage I consists of several down sampling convolution blocks, followed by a pyramid module and several up sampling convolution blocks. The pyramid module, inspired by PSPNet [37], is customized to enlarge the receptive field of our network. This module performs pooling on the feature maps at multiple resolutions and fuses the channels into a condensed feature map.

Loss Functions. We adopt global and local discriminators to encourage the generator \( G_e \) to produce realistic normal-light images. We train \( G_e \) using adversarial losses and perceptual losses with unpaired data. The details of the pyramid module and the loss functions are present in the supplementary materials.

B. Stage II: Illumination-aware Noise Suppression

As shown in Fig. 1, our noise suppression model \( G_n \) adopts the original low-light image \( I_l \), the enhanced image (with noise) \( I_e \) from Stage I, and an illumination mask \( M \) as inputs to generate the final clean image \( I_c \) with reduced noise as well as a good illumination condition, i.e., \( I_c = G_n(I_l, I_e, M) \). \( M \) indicates how much illumination is increased from low-light image \( I_l \) to enhanced image \( I_e \). It is defined as \( M = \max(\text{illu}(I_c) - \text{illu}(I_l), 0) \), where \( \text{illu}(\cdot) \) means extracting the illumination of an image. In this work, we directly use the grayscale version of an image as its illumination map. With \( M, I_l, \) and \( I_e \) as conditions, our denoising model is explicitly guided by the illumination conditions.

Model Architecture. Our denoising generator \( G_n \) in Stage II adopts an encoder-decoder architecture, with several convolutional blocks followed by several Resnet blocks then decoded back to an image. We adopt a multi-scale discriminator [38] to predict the realness of images at multiple resolutions. The detailed architecture of the networks can be found in the supplementary materials.

Loss Functions. Since there is no ground truth for the low-light image during training, we adopt an LSGAN-based adversarial loss to encourage the generated image to be as clean as the real-world clean normal-light images. The discriminator produces probabilities of a synthesized to be real against a randomly sampled clean image. Note that the discriminator needs not only to judge whether the illumination and color
of the generated image are realistic enough but also whether the generated image is clean without much noise. Training a single discriminator with clean or synthesized images to do both tasks simultaneously is challenging. As our goal in this stage is noise suppression, when feeding the discriminator, we first perform an instance normalization on both the synthesized image and the normal-light clean image to reduce the influence of image illumination, color, and contrast.

During training, we randomly match the denoised image \( I_c \) to a normal-light clean image \( J_c \). Our adversarial loss for training the discriminator \( D_n \) is:

\[
L_D = \mathbb{E}_{I_c \in \mathbb{P}}[(D_n(I_c) - 1)^2] + \mathbb{E}_{I_c \in \mathbb{Q}}[(D_n(I_c))^2].
\] (1)

The corresponding loss for updating the generator is \( L_G = \mathbb{E}_{I_c \in \mathbb{P}}[(D_n(I_c) - 1)^2] \), where \( \mathbb{P} \) and \( \mathbb{Q} \) are the normal-light clean image distribution and the distribution of generated images in Stage II, respectively.

Merely using the adversarial loss can cause color shifting problems, i.e., the color of the generated images can be easily distorted since we only constrain the images after instance normalization to be similar to the normal-light images. As we have already obtained an image \( I_c \) with good contrast and color from Stage I, in Stage II, we only need to preserve the contrast and color of \( I_c \). Therefore, we use a color loss to constrain the generated image \( I_c \) to have the same color as \( I_c \).

Specifically, we first down-sample the images with average pooling to \( I_c^l \) and \( I_c^l \) to suppress the noise in \( I_c \), then perform the color matching. We have:

\[
L_{color} = \sum_p \angle((I_c^l)_p, (I_c^l)_p)/N_n,
\]

where \( p \) is the location of a pixel in the down-sampled image, \( \angle(x, y) \) calculates the inner product between two 3-D vectors which are composed of RGB channels of a pixel location, \( N_n \) is the number of pixels in the downsized image.

**Pseudo Triples Construction.** We propose a Pseudo Labeling technique to facilitate the unsupervised training of the denoising model. As shown in Fig. 1 (b), we first estimate the noise in image \( I_c \) as \( I_n = I_c - I_c \). Then given the randomly matched normal-light clean real image \( J_c \), we can simulate a pseudo noisy image \( J_\gamma \) by adding the estimated noise to the clean image, i.e., \( J_\gamma = J_c + I_n \). We also use Gamma Correction [39] to decrease the brightness of \( J_\gamma \), to obtain a corresponding pseudo low-light image \( J_l = (J_c)^\lambda \), where \( \lambda \) is estimated as \( \lambda = \log \hat{T}_c/\log \hat{T}_l \). \( \hat{T}_c \) and \( \hat{T}_l \) are the average pixel values over all pixel locations of image \( I_c \) and \( I_l \), respectively. After these steps, we obtain a pseudo triple \( J = \{J_l, J_c, J_\gamma\} \), where \( J_l \) is the constructed pseudo low-light image, \( J_c \) is the pseudo enhanced image (with noise) and \( J_\gamma \) is the real clean image from our training set. Similarly, we construct the illumination mask for the pseudo triple as \( M_f = \max(iill(J_c) - iill(J_\gamma), 0) \). We can then predict a denoised image \( J_\gamma \) by \( G_n(J_l, J_c, M_f) \) from the constructed fake images and use \( J_\gamma \) as the supervision to train \( G_n \).

**Adaptive Content Loss for Pseudo Triples.** To train \( G_n \), we propose an adaptive content loss to constrain the generated image \( J_\gamma \) to be perceptually close to the clean normal-light image \( J_c \). Specifically, we use both perceptual loss and L1 reconstruction loss between \( J_\gamma \) and \( J_c \). As different regions may have different lighting conditions, regions with a significant brightness increase after the first stage may contain heavy noise, and regions without a significant brightness increase may contain less noise. When imposing the reconstruction constraint, we encourage the network to focus more on dark regions where noise is usually heavier. We then formulate the adaptive content loss for the pseudo triples as:

\[
L_{con}^{adapt} = \sum_l \|M_{jl}^l \circ (\Phi_l(J_\gamma) - \Phi_l(J_c))\|^2_2/N_l + \gamma_p\|J_\gamma - J_c\|_1/N, \] (2)

where \( M_{jl}^l \) is the downsized version of \( M_f \) that matches the spatial size of the VGG features at the \( l \)-th VGG layer. \( N_l \) serves as the weight for each pixel in the image. \( N_l \) is the number of elements in image \( J_c \). \( N_l \) is the number of elements in the feature maps of the \( l \)-th layer in VGG network. \( \Phi_l(I) \) is the feature in the \( l \)-th layer of VGG given the input image \( I \). \( \gamma_p \) is the weight to balance the losses from the RGB image domain and VGG feature domain. In our work, we choose the layers of “relu1_2”, “relu2_2”, “relu3_2”, “relu4_4”, “relu5_4” to perform both low-level and high-level feature matching, and \( \gamma_p \) is set as 10. Since we need to preserve the color and contrast, we do not use instance normalization on the VGG feature maps of the pseudo triple.

**Interpretation of Pseudo Labeling.** Note that at the early training stage, the estimated noise may contain apparent structures of the objects in the input noisy image \( I_c \). As a result, the constructed pseudo image \( J_\gamma \) may also contain object structures from \( I_c \). Since our network is trained to remove noise, it will be difficult for the network to remove the high-frequency object structure patterns. Then the generated image \( J_\gamma \) may also contain object structures from \( I_c \). By minimizing our proposed adaptive content loss, we are essentially encouraging the estimated noise to contain fewer object structures, i.e., encouraging the network to estimate the noise pattern more accurately.

**Content Preserving Loss.** To ensure that the denoised image \( I_c \) preserves contextual details of the enhanced image \( I_c \), we also impose a perceptual loss and a reconstruction loss between images \( I_c \) and \( I_c \). To reduce the influence of color and contrast, we perform instance normalization to the images before imposing the perceptual loss and reconstruction loss. The content loss is formulated as:

\[
L_{con} = \sum_l \|\Phi_l(I_c) - \Phi_l(I_c)\|^2_2/N_l + \gamma_c\|I_c - I_c\|_1/N, \] (3)

where the layers and operations used are the same as \( L_{con}^{adapt} \). \( \gamma_c \) is a weight balance term similar to \( \gamma_p \) and set as 10 in our
work. The total loss $L$ for training $G_n$ is a combination of all the losses.

$$L = LG + \lambda_cL_{color} + \lambda_pL_{con}^{adapt} + \lambda_cL_{con}, \quad (4)$$

and we empirically find that setting $\lambda_c, \lambda_p, \lambda_c^{adapt}, \lambda_c$ as 10, 1, 1, respectively, yields the best result.

### III. Experiments

In this section, we first introduce the datasets used in our experiments. Then we report the performance of each model concerning only illumination enhancement on low-light images without much noise. Following that, we conduct experiments for both illumination enhancement and noise suppression on low-light images containing significant noise. We put all the implementation details of our model in the supplementary.

#### A. Datasets

**Unpaired Enhancement Dataset.** Jiang et al. [22] collect an unpaired dataset for training contrast enhancement models. The training set comprises 914 low-light images, which are dark yet containing no significant noise, and 1916 normal-light images from public datasets.

**LOW-LIGHT (LOL) Dataset** [13]. LOL comprises 500 low-light and normal-light image pairs and is split into 485 training pairs and 15 testing pairs. To adapt the dataset to our unsupervised setting, we adopt the training images as our low-light train set, and the normal-light images in the Unpaired Enhancement Dataset [22] as the normal-light train set. The testing images remain the same as the LOL dataset.

**URL Dataset.** There are a quite limited number of real-world low-light datasets publicly available. Some public low-light datasets are composed of synthetic images, while many other datasets, such as ExDark [40] or Adobe FiveK [41] contain dark images without significant noise. These datasets do not meet the objective of our study. Therefore, we collect an Unsupervised Real-world Low-light dataset (URL dataset) composed of 414 high-resolution real-world low-light images taken by iPhone-6s and 3,837 normal-light images selected from Adobe FiveK. Our URL dataset is quite diverse, containing various indoor and outdoor scenes under different light conditions. Consequently, the noise level in each image or even different regions of the same image varies considerably across the dataset. We divide the low-light images into 328 training images and 86 testing images. This dataset thus complements the existing datasets in those two regards. Note that there is no corresponding ground-truth image for each testing image. Each low-light image is resized to 1008 × 756. More dataset details are in the supplementary materials.

#### B. Experiments for Illumination Enhancement Only

We compare our Stage I model which does only illumination enhancement on the Unpaired Enhancement Dataset [22] with state-of-the-art unsupervised models, including CycleGAN [42], UNIT [38] and EnlightenGAN [22]. As shown in Fig. 2, our model can generate normal-light images with reasonable contrast and color in both global and local regions. EnlightenGAN improves the illumination of the low-light images. However, the resulting images still suffer from color distortion and inconsistency in regions indicated by the red boxes. The other unsupervised methods produce images with imperfect color, contrast, and illumination. We report the PSNR and SSIM of the generated images only as a complement to the visual results on the Unpaired Enhancement Dataset. Table I shows that our model performs significantly better than the existing models, which are consistent with the visual results.

#### C. Experiments for Both Illumination Enhancement and Noise Suppression

**Experiment Settings.** We train the two stages of our model separately, as we observe that jointly training the two stages leads to an unstable training procedure under the unsupervised training setting. We evaluate the models on the real-world low-light image enhancement datasets LOL and our URL datasets and compare our model with the state-of-the-art unsupervised enhancement models, including CycleGAN [42], UNIT [38], and EnlightenGAN [22]. We also compare the performance of the combination of the illumination enhancement model and denoising model. Since our Stage I outperforms the existing models concerning illumination enhancement as demonstrated in the previous section, in this section, we only compare the combination of our Stage I model and other denoising methods. Specifically, we compare our two-stage model with our Stage I + BM3D [29], our Stage I + ADN [43], [44] and our Stage I + GCBD [45]. BM3D [29] is a robust image denoising method. However, it requires a known noise level as input. To use BM3D in our task, we first estimate a rough noise level for each testing image, then apply BM3D to the testing images.

**Results.** Table II shows PSNR, SSIM, and the perceptual scores (LPIPS) of the enhanced images on the LOL dataset (as these testing data have ground truth), demonstrating a consistent superiority of our model over the existing methods. Fig. 3 shows the visual results on both the LOL dataset and our URL dataset. Most existing methods suffer from heavy color distortion on the URL dataset and cannot preserve fine
TABLE III
CONFIGURATION OF DIFFERENT VERSIONS OF OUR MODEL.

| Model      | $L_G$ | $L_{\text{color}}$ | $L_{\text{mg}}$ | $L_{\text{com}}$ |
|------------|-------|---------------------|------------------|------------------|
| Version 1  | ✓     | ✓                   | ✓                | ✓                |
| Version 2  | ✓     | ✓                   | ✓                | ✓                |
| Version 3  | ✓     | ✓                   | ✓                | ✓                |
| Full Model | ✓     | ✓                   | ✓                | ✓                |
| Plain      | ✓     | ✓                   | ✓                | vanilla          |

TABLE IV
QUANTITATIVE RESULTS OF ABLATION STUDY.

|                      | V3 | Plain | Full |
|----------------------|----|-------|------|
| BRISQUE on URL       | 33.66 | 26.95 | 22.13 |
| NIQE on URL          | 4.12  | 3.44  | 2.53  |

User Study. We randomly select 20 low-light testing images from the URL dataset. For each input image, we show each user the final enhanced result of each method and ask the user to select the most visually pleasing result among all methods, considering illumination condition, color, texture realness, and noise level. In total, we obtain 200 preference opinions. Table II (last column) shows that our method consistently outperforms the state-of-the-art methods in terms of the preference opinion score (POS).

D. Ablation Study

We analyze the importance of each component in Stage II, which is the core contribution of this work. As shown in Table III, we compare the versions of our Stage II model with different losses imposed. We also compare our model to a non-adaptive denoising model, which we call the Plain Model, i.e., the generator only takes the enhanced image as input without illumination guidance. The four losses used in the Plain Model remain the same as our Full Model, except that we use the vanilla content loss instead of adaptive content loss.

Fig. 4 shows the results of each variant on the URL dataset. Enhanced images from Stage I still contain heavy noise, showing evidence that it is challenging to simultaneously model illumination and noise patterns with a single network. Merely using adversarial loss (V1) or additional color loss (V2) leads to distorted images as the content is not well-bounded. Imposing the learning with pseudo triples (V3), the network can produce realistic contents and perform noise suppression. The results compared to V1 and V2 demonstrates that our pseudo triple construction plays a key role in stabilizing the learning process. However, several regions are over-smoothed with this version. Please pay attention to the trees in these results in Fig. 4. The results of the Plain Model contain noticeable noise in the sky and other regions. The illumination of the image explicitly guides our full model. Therefore it can capture the noise pattern under various illumination conditions more effectively and produce better results.

Since the URL dataset does not contain ground truth images, we use non-reference image quality assessment methods BRISQUE [46] and NIQE [47] to quantify the visual quality of the results, as shown in Table IV. We do not include the scores of V1 or V2, as these two versions can only produce heavily distorted images, as shown in Fig. 4. These quantitative results further demonstrate the importance of using illumination as guidance for real-world noise modeling.
IV. CONCLUSION AND FUTURE WORK

In this paper, we have presented a decoupled model to address the real-world low-light image enhancement problem in an unsupervised fashion. Our model first enhances the illumination and color of a low-light image, then removes the real noise in the enhanced image and preserves contextual details. Experiments on real-world datasets show that our model outperforms the state-of-the-art models concerning both illumination enhancement and noise removal. Although our model achieves satisfactory performance, there are still challenging cases where our model is unable to process perfectly. One reason can be the choice of the normal-light image dataset, which will affect the stability of the unsupervised training. Another possible reason can be the pseudo triple construction which generates pseudo noisy images using only additive noise, while the real noise can be any form. A more realistic pseudo noisy image construction method may further improve the generalization of our model.

ACKNOWLEDGEMENT

This work is supported by our corporate sponsors.
REFERENCES

[1] E. H. Land, “The retinex theory of color vision,” Scientific american, vol. 237, no. 6, pp. 108–129, 1977.
[2] S. Lee, “An efficient content-based image enhancement in the compressed-domain using retinex theory,” IEEE transactions on circuits and systems for video technology, vol. 17, no. 2, pp. 199–213, 2007.
[3] D. J. Jobson, Z.-u. Rahman, and G. A. Woodell, “A multiscale retinex for bridging the gap between color images and the human observation of scenes,” IEEE Transactions on Image processing, vol. 6, no. 7, pp. 965–976, 1997.
[4] 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, vol. 29, no. 4, pp. 968–981, 2018.
[5] Z. Zhao, B. Xiong, L. Wang, Q. Ou, L. Yu, and F. Kuang, “Retinexdp: A unified deep framework for low-light image enhancement,” IEEE Transactions on Circuits and Systems for Video Technology, 2021.
[6] J. Li, X. Feng, and Z. Hua, “Low-light image enhancement via progressive-recursive network,” IEEE Transactions on Circuits and Systems for Video Technology, 2021.
[7] R. Mantris, S. Daly, and L. Kerofsky, “Display adaptive tone mapping,” in ACM SIGGRAPH 2008 papers, 2008, pp. 1–10.
[8] X. Fu, D. Zeng, Y. Huang, X.-P. Zhang, and X. Ding, “A weighted variational model for simultaneous reflectance and illumination estimation,” in CVPR, 2016, pp. 2782–2790.
[9] X. Guo, Y. Li, and H. Ling, “Lime: Low-light image enhancement via illumination map estimation,” IEEE Transactions on Image Processing, vol. 26, no. 2, pp. 982–993, 2016.
[10] 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, vol. 27, no. 6, pp. 2828–2841, 2018.
[11] K. G. Lore, A. Akintayo, and S. Sarkar, “Linnet: A deep autoencoder approach to natural low-light image enhancement,” Pattern Recognition, vol. 61, pp. 650–662, 2017.
[12] M. Gharbi, J. Chen, J. T. Barron, S. W. Hasinoff, and F. Durand, “Deep bilateral learning for real-time image enhancement,” ACM Transactions on Graphics (TOG), vol. 36, no. 4, pp. 1–12, 2017.
[13] C. Wei, W. Wang, W. Yang, and J. Liu, “Deep retinex decomposition for low-light enhancement,” in BMVC, 2018.
[14] C. Chen, Q. Chen, J. Xu, and V. Kolm, “Learning to see in the dark,” in CVPR, 2018, pp. 3291–3300.
[15] R. Wang, Q. Zhang, C.-W. Fu, X. Shen, W.-S. Zheng, and J. Jia, “Underexposed photo enhancement using deep illumination estimation,” in CVPR, 2019, pp. 6849–6857.
[16] K. Xu, X. Yang, B. Yin, and R. W. Lau, “Learning to restore low-light images via decomposition-and-enhancement,” in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020, pp. 2281–2290.
[17] C. Guo, C. Li, J. Guo, C. C. Loy, J. Hou, S. Kwong, and R. Cong, “Zero-reference deep curve estimation for low-light image enhancement,” in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020, pp. 1780–1789.
[18] S. Moran, P. Marza, S. McDonagh, S. Parisot, and G. Slabaugh, “Deepplif: Deep local parametric filters for image enhancement,” in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020, pp. 12 826–12 835.
[19] W. Yang, S. Wang, Y. Fang, Y. Wang, and J. Liu, “From fidelity to perceptual quality: A semi-supervised approach for low-light image enhancement,” in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020, pp. 3063–3072.
[20] S. M. Pizer, E. P. Amburn, J. D. Austin, R. Cromptie, A. Geselowitz, T. Greer, B. Ier Haar Romeny, J. B. Zimmerman, and K. Zuiderveld, “Adaptive histogram equalization and its variations,” Computer vision, graphics, and image processing, vol. 39, no. 3, pp. 355–368, 1987.
[21] Y.-S. Chen, Y.-C. Wang, M.-H. Kao, and Y.-Y. Chuang, “Deep photo enhancer: Unpaired learning for image enhancement from photographs with gaps,” in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018, pp. 6306–6314.
[22] Y. Jiang, X. Gong, D. Liu, Y. Cheng, C. Fang, X. Shen, J. Yang, P. Zhou, and Z. Wang, “Enlightengan: Deep light enhancement without paired supervision,” IEEE Transactions on Image Processing, vol. 30, pp. 2340–2349, 2021.