Exposure Normalization and Compensation for Multiple-Exposure Correction

Jie Huang*, Yajing Liu*, Xueyang Fu, Man Zhou, Yang Wang, Feng Zhao†, Zhiwei Xiong
University of Science and Technology of China
{hj0117,lyj123,manman}@mail.ustc.edu.cn, {xyfu,ywang120,fzhao956,zwxiong}@ustc.edu.cn

Abstract

Images captured with improper exposures usually bring unsatisfactory visual effects. Previous works mainly focus on either underexposure or overexposure correction, resulting in poor generalization to various exposures. An alternative solution is to mix the multiple exposure data for training a single network. However, the procedures of correcting underexposure and overexposure to normal exposures are much different from each other, leading to large discrepancies for the network in correcting multiple-exposures, thus resulting in poor performance. The key point to address this issue lies in bridging different exposure representations. To achieve this goal, we design a multiple exposure correction framework based on an Exposure Normalization and Compensation (ENC) module. Specifically, the ENC module consists of an exposure normalization part for mapping different exposure features to the exposure-invariant feature space, and a compensation part for integrating the initial features unprocessed by the exposure normalization part to ensure the completeness of information. Besides, to further alleviate the imbalanced performance caused by variations in the optimization process, we introduce a parameter regularization fine-tuning strategy to improve the performance of the worst-performed exposure without degrading other exposures. Our model empowered by ENC outperforms the existing methods by more than 2dB and is robust to multiple image enhancement tasks, demonstrating its effectiveness and generalization capability for real-world applications. Code: https://github.com/KevinJ-Huang/ExposureNorm-Compensation.

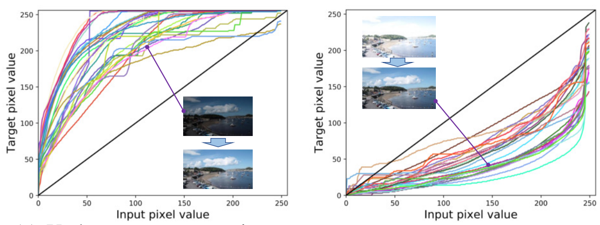
1. Introduction

In recent years, camera devices are used to capture photographs of a wide range of scenes at any time. As different scenes present various exposure conditions, images captured with underexposures or overexposures often suffer from unsatisfactory visual effects. For this reason, several exposure correction methods have been proposed, including model-driven-based [4, 9, 11, 21, 38] and deep learning-based approaches [10, 33, 34, 36, 41]. However, most of them focus on either underexposure or overexposed scenes, causing poor generalization to other exposures. This makes them incapable of being deployed in practical applications.

A naive way to solve this problem is to train specific networks corresponding to each exposure condition, leading to a significant increase in training time and parameter space. Alternatively, the network can be trained with a mixture of data from various exposure conditions to improve its capability of multiple-exposure correction. However, due to the variations of representations between underexposure and overexposure, the procedures of correcting them to normal exposures differ greatly from each other, as shown in Fig. 1. This introduces discrepancies for the network in correcting lightness and color across multiple-exposures, thus making it difficult to train a single network and resulting in poor performance. Besides, the variations in optimization processes also make the network incline to overlook disadvantaged data of the mixed datasets [35] and bring about imbalanced performance across exposures.

In this paper, we propose a framework for improving the multiple-exposure correction performances. The key point lies in narrowing the gap of different exposure representations. To this end, we design an Exposure Normalization and Compensation (ENC) module, as shown in Fig. 3. It consists of an exposure normalization part and a compensation part. Specifically, the exposure normalization part maps different exposure features to the exposure-invariant feature space. It is implemented by the Instance Normalization to coarsely align different exposure features, followed by a normalization distilling loss on the normalized features to further reduce the exposure effect. However, normalization will inevitably induce the loss of the image discriminative features [17, 27] for image reconstruction. Therefore, the compensation part is introduced for integrating the features unprocessed by the exposure normalization part in the spatial and channel dimension, which ensures the completeness information. Besides, to further enhance the correction effect of the worst-performed exposure caused by imbal-

*Equal contribution.
†Corresponding author.



(a) Underexposure correction (b) Overexposure correction

Figure 1. The static illustration of underexposure and overexposure correction curves on samples from SICE dataset, which are significantly different from each other.

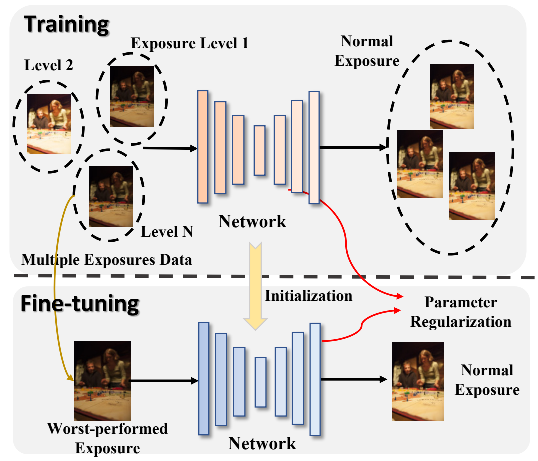


Figure 2. The overview of our proposed framework. During training, the network is trained for correcting multiple-exposures to normal exposures. During fine-tuning, we fine-tune the worst-performed exposure at training phase by parameter regularization.

anced performance, an intuitive way for fine-tuning this exposure will result in performance degradation for other exposures. Here, we fix the parameter of the ENC to keep its normalization ability to all the exposures, and fine-tune the other parts using a parameter regularization strategy. This strategy reduces the update rate of parameters that are important for other exposures, leading to a network with balanced improvement for all the exposures.

Moreover, our proposed ENC can be extended to tackle other image degradations, such as image quality degradation captured by different mobile phones, demonstrating the generalization capability of our method.

The main contributions of this work are summarized as:

- We propose a framework for multiple exposure correction by narrowing the gap of different exposure representations. Particularly, we develop an Exposure Normalization and Compensation (ENC) module, which is simple yet effective and can be used as a plug-and-play module for existing exposure correction architectures.
- Inside ENC, we design an exposure normalization part for mapping different exposure features to the exposure-invariant feature space, and a compensation part for integrating the features unprocessed by normalization to ensure the completeness of information.
- Aiming to improve the worst-performed exposure during training, we employ a parameter regularization strategy for fine-tuning it on the network except for the ENC module, resulting in a balanced improvement.
- We validate the effectiveness of our framework on several datasets. Furthermore, we extend it to various image-enhancement tasks and achieve remarkable performances, which demonstrate its generalization superiority for potential usage in real applications.

2. Related Work

Various approaches have been developed for exposure correction. Some traditional ones propose to employ the histogram-based technique to enhance contrast and lightness [1, 30, 31, 38], while another line of works is based on the Retinex theory [19], which improve the lightness of images through enhancing the illumination components and suppressing noises by the regularization of the reflectance components [4, 11, 21, 29, 39].

Recently, with the emergence of deep learning schemes, exposure correction task has benefited from the deep learning models [7, 13, 22, 26, 37]. Based on the Retinex theory, RetinexNet [34] proposes to restore the illumination in a data-driven form and KIND [41] further introduces a sub-network for recovering the reflectance component. As another form of component decomposition, Ren *et al.* [28] uses two distinct streams to learn global content and salient structures simultaneously, and DRBN [36] decomposes the features into different band representations for band recursive learning. In addition, self-supervised methods [10, 20, 40] were proposed for the adaptive illumination adjustment. However, most of these methods focus on either underexposure or overexposure correction, limiting their applications for various exposure conditions. Although MSEC [2] corrects varieties of exposures in a coarse-to-fine manner, it fails to achieve consistent correction across exposures, thus generating results with lightness shifts. Compared with these methods, our algorithm aims to narrow the gap of different exposure features for effectively improving the training performance.

3. Method

3.1. Motivation and Overview

Images captured under different scenes often suffer from underexposure or overexposure problems. For multiple-exposure correction, we aim at designing a unified framework training on the mixed multiple-exposure datasets for correcting various exposures to normal ones.

As shown in Fig. 1, since underexposure and overexposure obviously present different exposure representations

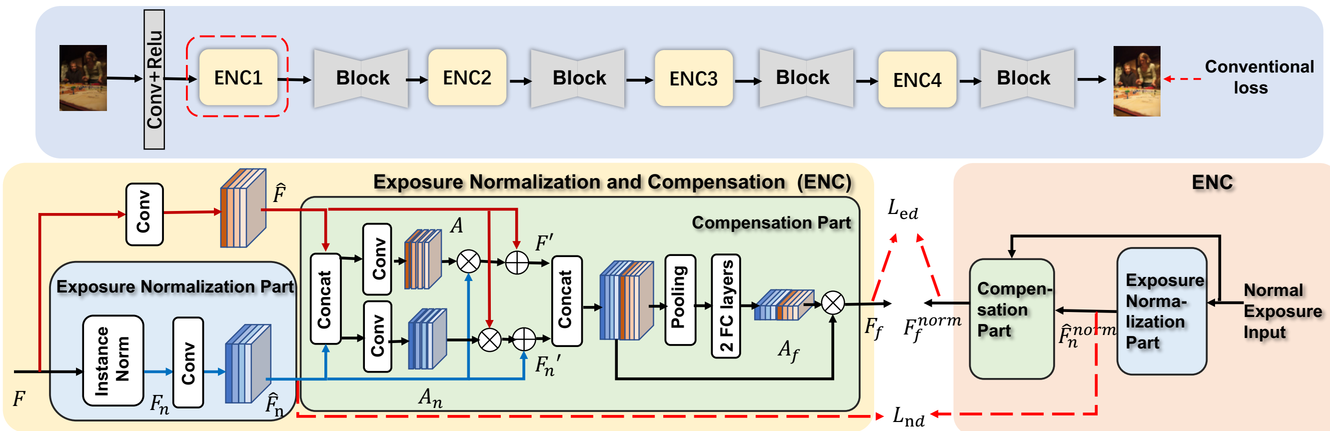


Figure 3. The illustration of our proposed ENC module in DRBN [36] network, consisting of an exposure normalization part and a compensation part. The exposure normalization part maps different exposure features F to an exposure-invariant feature $\hat{F}_n$, while the compensation part compensates for the information lost caused by normalization, which integrates the features that unprocessed by the exposure normalization $\hat{F}$ in the spatial and channel dimensions. The normalization distilling loss L_{nd} and the exposure distilling loss L_{ed} are further introduced for reducing the effect of the exposure on features.

such as lightness, the procedures of correcting underexposure and overexposure differ significantly. Therefore, there exist large discrepancies for a network to correct lightness and color across multiple-exposures, which introduces difficulty in the training process and results in poor performance, bringing color and lightness distortion problems. Furthermore, training on mixed datasets may cause variations in the optimization processes, leading to the network incline to overlook the disadvantaged data in the mixed datasets [35]. Therefore, the network has poorer performance in some exposures than other exposures, thus resulting in imbalanced performance across exposures.

To address the above issues, we design a framework (see Fig. 2) consisting of training and fine-tuning phases. During the training, the network is trained on mixed data of various exposures with its conventional loss. Specifically, to solve the poor performance problem, we propose an ENC module to narrow the gap of multiple exposure representations, leading to consistent lightness and color correction across exposures. The ENC module can be used as a plug-and-play module for existing exposure correction networks. Taking the DRBN [36] network as an example (see Fig. 3), ENC can be added after each block of DRBN. To further alleviate the imbalanced performance, we name the exposure that performs poorer than other exposures as the worst-performed exposure, and fine-tune it with a parameter regularization strategy for improving its correction performance, while keeping the performance of other exposures.

3.2. Exposure Normalization and Compensation

As shown in Fig. 3, we implement our ENC module with two parts: the exposure normalization part that is designed for mapping various exposure features to exposure invariant feature space, and the compensation part which is proposed to integrate the features unprocessed by normalization for compensating the removed image discriminative information caused by normalization.

Exposure Normalization Part. In the exposure normalization part, we first employ Instance Normalization to coarsely align features. Assuming the input features as F, we perform Instance Normalization by:

$$F_n = IN(F) = \gamma \frac{F - \mu(F)}{\sigma(F)} + \beta, \tag{1}$$

where $\mu(\cdot)$ and $\sigma(\cdot)$ denote the mean and standard deviation computed across spatial dimensions for each channel and each sample, γ and β are parameters learned from data. With Instance Normalization equipped in the feature space, it can normalize feature statistics for style normalization [14]. Since each exposure can be viewed as a kind of style, different exposures are aligned with Instance Normalization that reduces their representation discrepancies.

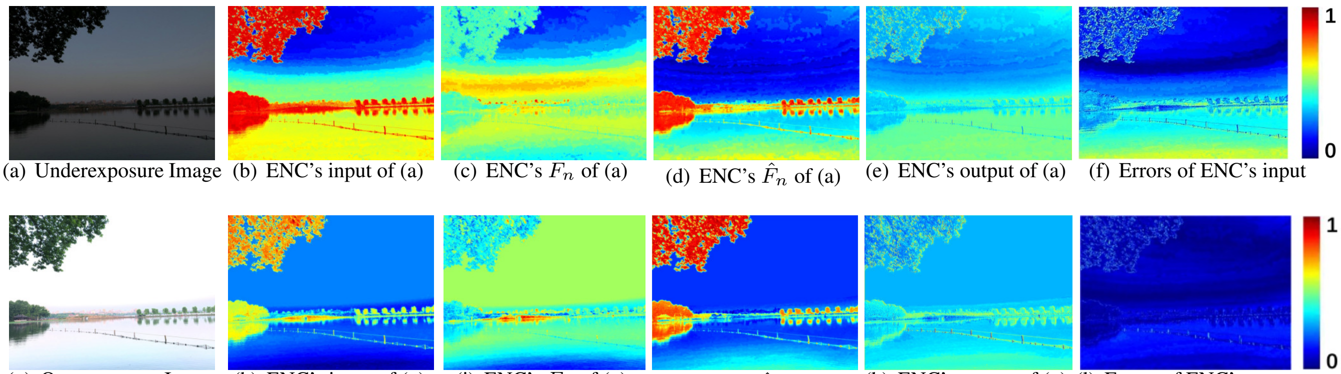
Following that, we introduce a normalization distilling loss to further reduce the exposure's effect on the normalized features. Particularly, followed by a convolution layer, we implement this loss between the normalized features of different exposures ($\hat{F}_n$) and those of the normal exposures ($\hat{F}_n^{norm}$), which is defined as:

$$L_{nd} = ||\hat{F}_n - \hat{F}_n^{norm}||_1, \tag{2}$$

where $||.||_1$ represents the L_1 distance between two terms. It efficiently forces the normalized features of different exposures to be similar to those of the normal exposures, thus reducing their discrepancy.

Fig. 4 presents the feature visualization of different components in our ENC, where the underexposure and overexposure features processed by the Instance Normalization are more similar, and L_{nd} further reduces their discrepancy.

Compensation Part. Normalization inevitably removes discriminative information [17, 27], thus resulting in inadequate information for image reconstruction. To tackle this



(a) Underexposure Image (b) ENC's input of (a) (c) ENC's F_n of (a) (d) ENC's $\hat{F}_n$ of (a) (e) ENC's output of (a) (f) Errors of ENC's input

(g) Overexposure Image (h) ENC's input of (g) (i) ENC's F_n of (g) (j) ENC's $\hat{F}_n$ of (g) (k) ENC's output of (g) (l) Errors of ENC's output

Figure 4. Feature visualization of different components in ENC on samples from SICE dataset. As can be seen, the ENC's input features F from underexposure and overexposure differ greatly shown in (b) and (h). As shown in (c) and (i) as well as (d) and (j), after being processed by the exposure normalization part, input features are mapped to the exposure invariant space and their discrepancies are progressively reduced. With ENC, the gap of representations between underexposure and overexposure is obviously narrowed as illustrated in (e) and (k) as well as (f) and (l).

shortcoming, as shown in Fig. 3, we propose a compensation part for integrating the initial features unprocessed by the exposure normalization part to ensure the completeness information [23]. Specifically, we implement the compensation part in both spatial and channel dimensions, which can comprehensively obtain correlations between the initial and normalized features. These correlations reflect their information relationships, thus helping guide the integration of the lost information from the initial features.

In the spatial dimension, the normalized features $\hat{F}_n$ and the initial features unprocessed by normalization $\hat{F}$ are integrated with the attention maps A and A_n. Here, A and A_n represent the correlations between features $\hat{F}$ and $\hat{F}_n$, and A is derived by the spatial attention as:

$$A = sigmoid(W_0 * [\hat{F}, \hat{F_n}]), \tag{3}$$

where W_0 is the kernels' weight matrix, $*$ stands for the convolution operation, and $[\cdot]$ means the concatenate operation. Similarly, A_n is generated as well. The spatially-interacted features F' and F'_n can be obtained by:

$$\begin{aligned} F' &= \hat{F_n} \cdot A + F, \\ F'_n &= \hat{F} \cdot A_n + F_n, \end{aligned} \tag{4}$$

where $\cdot$ denotes the element-wise multiplication.

Then, we further integrate the two features in the channel dimension. Specifically, we re-weight the concatenated features of F' and F'_n by applying the attention weight A_f to adaptively integrate them, and the A_f is derived by the SE-like [12] channel attention. In particular, A_f is obtained by a pooling layer and two FC layers parameterized by W_1 and W_2, which can be denoted as:

$$A_f = sigmoid(W_2 \cdot relu(W_1 * pool([F', F'_n]))). \tag{5}$$

Notably, we implement the pooling operation in Eq. 5 with global contrast average pooling for capturing global

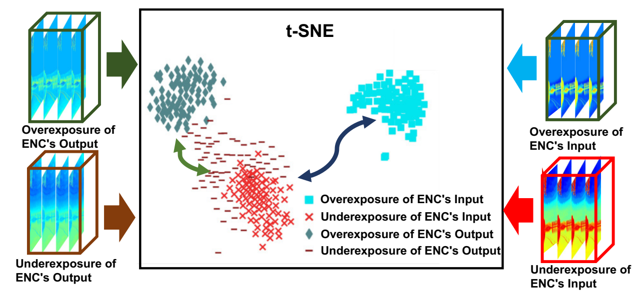


Figure 5. t-SNE [32] visualization of the features in ENC. As can be seen, after being processed by ENC, the underexposure and overexposure representations tend to be intersected together.

and local information [15], which is propitious for image processing. This operation is defined as:

$$\begin{aligned} F_o =& \frac{1}{HW} \sum_{(x,y)\epsilon F_i} F_i^{x,y} + \\ & \sqrt{\frac{1}{HW} \sum_{(x,y)\epsilon F_i} (F_i^{x,y} - \frac{1}{HW} \sum_{(x,y)\epsilon F_i} F_i^{x,y})^2}, \end{aligned} \tag{6}$$

where F_i and F_o represent the input and output features of the global pooling operation, x and y are the position coordinates, H and W denote the spatial size. Finally, the output features of ENC (denoted as F_f) are derived by weighting the concatenated features $[F', F'_n]$ with A_f, which can be denoted by:

$$F_f = A_f \cdot [F', F'_n]. \tag{7}$$

To further maintain the exposure invariant property of the integrated features F_f, we need to reduce the exposure effect introduced from the ENC's input features F on F_f. Therefore, we apply the exposure distilling loss between F_f and the integrated normal exposure features F_f^{norm} as:

$$L_{ed} = ||F_f - F_f^{norm}||_1. \tag{8}$$

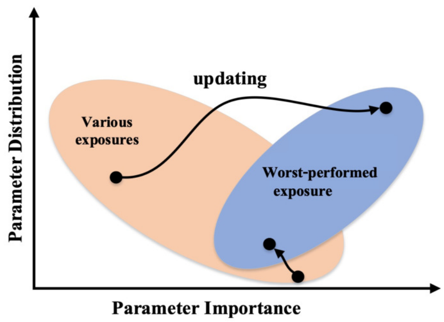


Figure 6. The schematics of the parameter regularization strategy. At the bottom, the parameters that are more important to various exposures are less updated; while at the top, the parameters that are less important to various exposures but more important to the worst-performed exposure are more updated.

After being processed by ENC, the gap between underexposure and overexposure representations is narrowed. As shown in the last column of Fig. 4, the errors between underexposure and overexposure features are smaller after being processed by ENC. Fig. 5 further presents the statistical visualization in the feature space, demonstrating the effectiveness of our method for bridging their representations.

Plugging into exposure correction networks. As a plug-and-play module, the proposed ENC module can be incorporated into most existing exposure correction networks. Two representative baseline networks, SID [6] and DRBN [36], are chosen as the backbones. For SID, it is a U-Net-based architecture with an encoder and decoder. We replace its first layer with the ENC module and denote it as SID-ENC, providing the features with bridged representations for the subsequent processing. As for DRBN, it is a framework consisting of multiple U-shape-based blocks. As shown in Fig. 3, the ENC module can also be set as the first layer of DRBN in its first block (DRBN-ENC) or in all the four blocks (DRBN-ENC-4), both resulting in better performance on exposure correction.

3.3. Fine-tuning with Parameter Regularization

To further improve the worst-performed exposure that has poorer performance than other exposures in the training phase, an intuitive way is to fine-tune it. However, this may lead to performance degradation of other exposures. Inspired by the EWC method of continuous learning [18], we employ a parameter regularization strategy with a first- and second-order parameter importance modification scheme. Specifically, since ENC is important for all the exposures and helps narrow the gap of their representations, we fix its parameters and update the other parameters according to the parameter importance. In doing so, we can improve the results of the worst-performed exposure while retaining the performance of the other exposures.

As shown in Fig. 6, we present the schematics of the parameter regularization strategy. In particular, we first calculate the parameter importance of the network obtained from the multiple-exposure training phase, then fix the parameters of ENC, and update the other parameters based on the parameter importance. Specially, by denoting the training on various exposures as task 0 and fine-tuning on the worst-performed exposure as task 1, the parameter importance weight Ω_{θ_k} is computed by accumulating the gradients over various exposure data points:

$$\Omega_{\theta_k} = f(x;\theta_k^1) - f(x;\theta_k^0), \tag{9}$$

where $f(.)$ represents the mapping function of our network, θ_k denotes any parameter of the network, and $\theta_k^1 = \theta_k^0 + \delta\theta_k$, δ denotes the parameter change magnitude, and x indicates the input various exposure data. In particular, the above equation can be written as:

$$\Omega_{\theta_k} = \nabla_{\theta_k}\mathcal{L}\,|\delta\theta_k| + \frac{1}{2}\cdot\nabla^2_{\theta_k}\mathcal{L}\,|\delta\theta_k|^2 + O(|\delta\theta_k|^3), \tag{10}$$

where $\mathcal{L}$ is the conventional loss of baseline method. Here, we adopt the first two terms for approximation *.

To improve the worst-performed exposure while maintaining the performance of other exposures, we add a regularization term based on the baseline's conventional loss to keep the knowledge of training on all the exposures. In summary, the total loss $\mathcal{L}'$ for fine-tuning is formulated as:

$$\begin{aligned} \mathcal{L}' &= \mathcal{L} + \lambda\sum_{k=1}^{m}\Omega_{\theta_k} \\ &= \mathcal{L} + \lambda\sum_{k=1}^{m}\left[\nabla_{\theta_k}\mathcal{L}\,|\delta\theta_k| + \frac{1}{2}\cdot\nabla^2_{\theta_k}\mathcal{L}\,|\delta\theta_k|^2\right]. \end{aligned} \tag{11}$$

In this way, the parameters of the network that are important to other exposures are less updated, thus maintaining their performance.

4. Experiments

4.1. Settings

Datasets. The experiments are evaluated on two datasets of multiple-exposures, including the multiple-exposures (ME) dataset collected by MSEC [2] and the SICE dataset [5]. The ME dataset contains exposure images of five levels. To demonstrate the effectiveness of our method, we conduct experiments on two settings for the ME dataset. Following [3], the retouched version of the middle exposure subset is selected as the ground truth in the standard ME dataset [2], which includes 17,675 training sample pairs, 750 validation sample pairs, and 5,905 testing sample pairs. Furthermore, we also conduct experiments on revised ME-v2 dataset, which selects the middle exposure subset in ME dataset as the ground truth and preserves the other exposure subsets as the multiple exposure input. Note that ME-v2

*The details of the derivation process can be found in the supplementary material.

Method	ME						ME-v2						#Param
	Under		Over		Average		Under		Over		Average		
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	
CLAHE [30]	16.77	0.6211	14.45	0.5842	15.38	0.5990	19.06	0.8558	15.76	0.7643	17.41	0.8100	-
RetinexNet [8]	12.13	0.6209	10.47	0.5953	11.14	0.6048	14.03	0.7354	16.32	0.7781	15.18	0.7568	1.70M
Zero-DCE [10]	14.55	0.5887	10.40	0.5142	12.06	0.5441	15.95	0.7912	9.89	0.6837	12.92	0.7374	0.33M
MSEC [2]	20.52	0.8129	19.79	0.8156	20.08	0.8145	24.48	0.9055	24.30	0.9153	24.29	0.9104	7.04M
DRBN [36]	19.74	0.8290	19.37	0.8321	19.52	0.8309	24.10	0.9156	23.15	0.9055	23.63	0.9105	0.53M
DRBN-L	19.84	0.8319	19.50	0.8355	19.64	0.8340	24.35	0.9168	23.34	0.9083	23.84	0.9126	0.67M
I-DRBN (Ours)	22.05	0.8476	**22.17**	**0.8542**	22.12	0.8516	25.32	0.9100	26.73	**0.9337**	26.02	0.9218	0.54M
I-DRBN-4 (Ours)	**22.72**	**0.8544**	22.11	0.8521	**22.35**	**0.8530**	**27.29**	**0.9264**	**26.75**	0.9275	**27.02**	**0.9270**	0.58M
SID [6]	19.37	0.8103	18.83	0.8055	19.04	0.8074	24.04	0.9072	21.1	0.8669	22.57	0.8871	7.40M
SID-L	19.32	0.8099	18.95	0.8073	19.10	0.8083	23.07	0.8966	22.25	0.8773	22.66	0.8870	11.56M
I-SID (Ours)	**22.59**	**0.8423**	**22.36**	**0.8519**	**22.45**	**0.8481**	**26.09**	**0.9133**	**26.62**	**0.9166**	**26.35**	**0.9149**	7.45M

Table 1. Quantitative results of different methods on ME and ME-v2 Datasets in terms of PSNR and SSIM.

Method	Under	Over	Average
CLAHE [30]	12.69/0.5037	10.21/0.4847	11.45/0.4942
RetinexNet	12.94/0.5171	12.87/0.5252	12.90/0.5212
Zero-DCE [10]	16.92/0.6330	7.11/0.4292	12.02/0.5311
MSEC [2]	19.62/0.6512	17.59/0.6560	18.58/0.6536
DRBN [36]	17.96/0.6767	17.33/0.6828	17.65/0.6798
DRBN-L	18.04/0.6746	17.61/0.6923	17.83/0.6835
I-DRBN (Ours)	20.47/0.7050	19.22/0.7222	19.85/0.7136
I-DRBN-4 (Ours)	**21.77/0.7052**	**19.57/0.7267**	**20.67/0.7160**
SID [6]	19.51/0.6635	16.79/0.6444	18.15/0.6540
SID-L	19.43/0.6644	17.00/**0.6495**	18.22/0.6570
I-SID(Ours)	**21.30/0.6645**	**19.63**/0.6941	**20.47/0.6793**

Table 2. Quantitative results of various methods on SICE dataset in terms of PSNR and SSIM.

dataset includes 14,144 pairs for training, 600 pairs for validation, and 4,724 pairs for testing. For the SICE dataset, we adopt the middle exposure subset as the ground truth, while the second and last second exposure subsets are set as the underexposed and overexposed images, respectively. The number of the training, validation, and testing pairs in the SICE dataset is individually set to 1,000, 24, and 60.

Comparison of Methods. For performance comparison, we compare our method with MSEC and the baseline networks. Besides, CLAHE [30], RetinexNet [8], and Zero-DCE [10] are chosen for comparison. More comparative results with other methods are provided in the supplementary material. Due to the introducing of more parameters in ENC, we expand our baseline networks by increasing the number of channels for a fair comparison, which are denoted as DRBN-L and SID-L. Additionally, the SID-ENC, DRBN-ENC, and DRBN-ENC-4 models mentioned in Sec. 3.2 with our fine-tuning strategy are separately denoted as I-SID (Improved-SID), I-DRBN (Improved-DRBN), and I-DRBN-4 (Improved-DRBN-4).

Implementation Details. We conduct all our experiments on an NVIDIA 2080Ti GPU, which are based on the released code of the baseline networks with the same training settings. Specifically, our SID is trained with the batch size of 1 and patch size of 384×384, while DRBN is trained with the batch size of 4 and patch size of 256×256. During training, we optimize the networks by the Adam optimizer with a learning rate of 1×10^{-4} for 80 epochs. During fine-tuning, we set the λ in Eq. 11 to 0.7, and the network is trained for 40 epochs with a learning rate of 4×10^{-5}. All the methods are evaluated in terms of PSNR and SSIM.

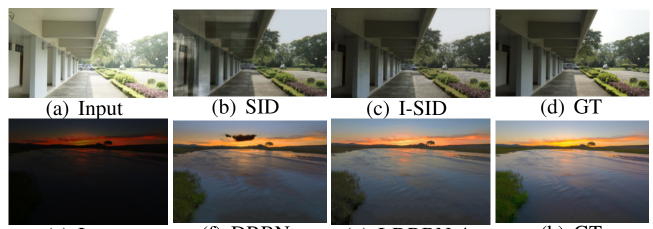
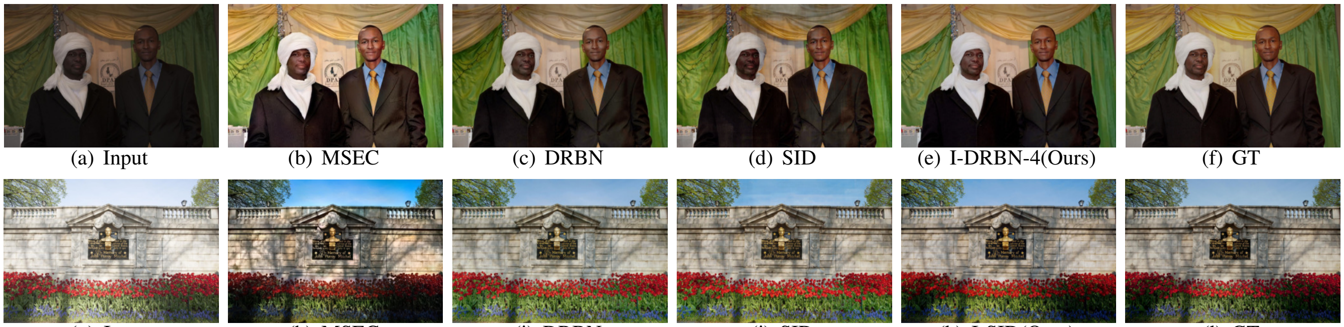

(a) Input (b) SID (c) I-SID (d) GT
(e) Input (f) DRBN (g) I-DRBN-4 (h) GT

Figure 7. Visualization results on SICE dataset of (top) overexposure correction and (bottom) underexposure correction.

4.2. Quantitative Evaluation

The evaluation results on the ME and ME-v2 datasets are reported in Table 1. To simplify, we average the results of the first two levels' exposures and the rest levels' exposures as the underexposure and overexposure results, respectively. As can be seen, the MSEC method performs better than our baseline networks with the well-designed networks, and the introduced channels in SID-L and DRBN-L cannot improve the performance significantly. With the assistance of our method, the I-SID and I-DRBN networks both achieve better performance and obtain superior results than the MSEC method. As the model size only increases by 3%, I-SID and I-DRBN-4 remarkably improve the PSNR, which proves the effectiveness of our ENC module.

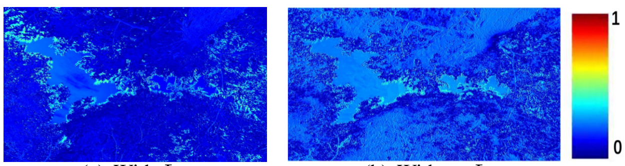
To further demonstrate the capability of our model, we also perform experiments on the SICE dataset. As shown in Table 2, with the introducing of our method, the PSNR and SSIM of SID and DRBN are improved greatly on both underexposure and overexposure subsets, which outperform other methods by a large margin.



(a) Input (b) MSEC (c) DRBN (d) SID (e) I-DRBN-4(Ours) (f) GT

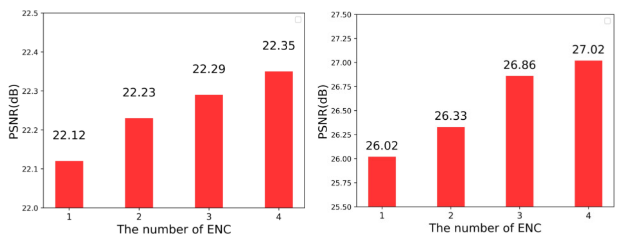
(g) Input (h) MSEC (i) DRBN (j) SID (k) I-SID(Ours) (l) GT

Figure 8. Visualization results on ME dataset of (top) underexposure correction and (bottom) overexposure correction. As can be seen, for both the underexposed and overexposed images, there exist color and lightness shift problems in DRBN and MSEC, while SID tends to generate artifacts. On the contrary, our method can simultaneously achieve color and lightness recovery while preserving the structures.



(a) With L_{ed} (b) Without L_{ed}

Figure 9. The visualization of the ENC's output errors between underexposure and overexposure. With the employing of exposure distilling loss, the errors are further reduced.



(a) ME Dataset (b) ME-v2 Dataset

Figure 10. Ablation study for the number of ENC modules based on the DRBN network. The increasing number of ENC modules leads to a better performance.

Method	Under	Over	Average
DRBN-ENC	**21.89/0.7071**	19.09/0.7229	20.49/0.7150
DRBN-ENC-4-SEQ	7.92/0.1346	**19.96/0.7315**	13.94/0.4331
I-DRBN-4 (Ours)	21.77/0.7052	19.57/0.7267	**20.67/0.7160**
SID-ENC	**21.36/0.6652**	19.38/0.6843	20.37/0.6748
SID-ENC-SEQ	7.93/0.1199	**19.95/0.7137**	13.94/0.4168
I-SID (Ours)	21.30/0.6645	19.63/0.6941	**20.47/0.6793**

Table 3. Ablation study for parameter regularization on SICE dataset, the overexposure subset is set as the worst-performed subset due to its lower performance.

4.3. Qualitative Evaluation

Fig. 8 exhibits some visualization results on the ME dataset. It can be seen that our method achieves better color and lightness recovery effects. We further present visual comparisons on the SICE dataset in Fig. 7. With the employing of our method, the artifacts can be reduced remarkably. More visualization results are provided in the supplementary material.

4.4. Ablation Studies

We perform ablation studies to prove the effectiveness of the proposed ENC module and parameter regularization strategy. More results of ablation studies are provided in the supplementary material.

ENC Module. Based on the SID network, we conduct experiments on the ME and SICE datasets to investigate the effectiveness of different components in ENC module. As shown in Table 6, the network performance drops significantly without Instance Normalization, demonstrating the effectiveness of mapping different exposures to the exposure-invariant space, and the introducing of normalization distilling loss further strengths this effect. The compensation part can also improve the performance since it integrates the initial features for ensuring the completeness of information. Specifically, both of the integration processes in the spatial and channel dimensions obtain improvements. Additionally, the introducing of exposure distilling loss helps to reduce the exposure effect on the integrated features (see Fig. 9), thus leading to enhancement. Note that only the employment of exposure loss can also contribute to the improvement due to the introducing of feature constraints. We further investigate the influence of the number of ENC modules for the DRBN baseline network. As shown in Fig. 10, it validates the effectiveness of our ENC module for improving exposure correction. Additionally, we provide results of comparing our ENC module with other plug-and-play modules in the supplementary material.

Parameter regularization strategy. To further demonstrate the effectiveness of our parameter regularization strategy, we carry out ablation studies on the SICE dataset. As depicted in Table 3, the simply fine-tuned DRBN and SID on the worst-performed overexposure dataset are denoted as DRBN-ENC-4-SEQ and SID-ENC-4-SEQ, which lead to the notable performance drop on the underexposure dataset. With the employing of parameter regularization, the overexposure performance can be enhanced with little performance drop on other-level exposures.

Method	LOL		SICE_OVER		FIVE5K		Average		SICE_UNDER		BRIGHTEN	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
DRBN	19.39	0.8165	19.46	0.7286	21.77	0.8585	20.21	0.8012	9.27	0.2644	15.19	0.6235
DRBN-L	18.97	0.8147	19.04	0.7272	21.93	0.8600	19.98	0.8006	9.62	0.3197	15.97	0.6328
I-DRBN-4 (Ours)	**22.31**	**0.8366**	**20.83**	**0.7385**	**23.71**	**0.8703**	**22.28**	**0.8151**	**13.19**	**0.4005**	**19.80**	**0.6969**
SID	19.78	0.7617	18.91	0.6885	21.39	0.8390	20.03	0.7631	11.13	**0.4300**	15.58	0.5949
SID-L	20.21	0.7696	19.57	0.7042	21.40	0.8442	20.39	0.7727	9.90	0.3477	16.28	0.6154
I-SID (Ours)	**21.27**	**0.7823**	**20.54**	**0.7107**	**23.48**	**0.8624**	**21.76**	**0.7851**	**12.00**	0.3681	**17.13**	**0.6646**

Table 4. Quantitative results of various methods for multiple enhancement tasks and generalizability evaluation.

Method	iPhone	Sony	Blackberry	Average
DRBN	22.84/0.824	25.15/0.879	24.14/0.850	24.04/0.852
DRBN-L	22.84/0.823	**26.10/0.883**	24.53/**0.853**	24.49/0.850
I-DRBN-4	**23.54/0.825**	26.04/0.882	**24.67**/0.851	**24.75/0.852**

Table 5. Quantitative results of different DRBN methods for various kinds of phone image enhancement on the DPED dataset.

IS	IC	IN	L_{nd}	L_{ed}	SICE	ME
					18.15/0.6540	19.04/0.8074
		✓			19.81/0.6666	21.36/0.8373
	✓	✓			20.01/0.6678	21.78/0.8392
✓		✓			19.98/0.6721	21.53/0.8369
✓	✓				18.82/0.6654	21.27/0.8378
✓	✓	✓			20.07/0.6726	22.10/0.8452
				✓	19.45/0.6671	20.67/0.8391
✓	✓	✓	✓		20.16/0.6734	22.21/0.8458
✓	✓	✓		✓	20.24/0.6742	22.29/0.8467
✓	✓	✓	✓	✓	**20.37/0.6748**	**22.37/0.8472**

Table 6. Ablation study for investigating the components of ENC module. IS and IC denote integration in the spatial and channel, respectively. IN represents Instance Normalization. L_{nd} is normalization distilling loss, and L_{ed} is the exposure distilling loss.

(a) Input (b) DRBN (c) I-DRBN-4 (d) GT

(e) Input (f) SID (g) I-SID (h) GT

Figure 11. Visualization results on the Brighten dataset.

4.5. Extension and Discussion

To demonstrate the potentials of our method for real-world applications, we extend it to different kinds of enhancement tasks. First, we blend several datasets and simultaneously address three image enhancement tasks, including image retouching, low-light enhancement, and overexposure correction. Second, to prove the generalizability of our method, we evaluate the trained model on unknown enhancement datasets without fine-tuning. Third, we conduct enhancement for image quality degradation caused by different phone capturing, which is a real-world problem.

Multiple enhancement tasks. We blend the LOL dataset [8] designed for low-light enhancement, MIT-FiveK dataset [3] collected for image retouching, and the overexposure subset from the SICE dataset to build a Task-mix dataset. The results are shown in Table 4. With the introducing of our method, the performance of each task on the multi-task dataset can be improved significantly.

Generalizability evaluation. To evaluate the generalizability of the proposed framework [24, 25], we evaluate the trained model on Brighten dataset [8] and the underexposure subset from the SICE dataset. As shown in Table 4 and Fig. 11, the introducing of instance normalization improves the robustness of our module. The generalization results of our method can also be improved, which demonstrate the potential usage of our work for real-world applications.

Various kinds of phone image enhancement. We adopt the DPED dataset [16] for experiments of various kinds of phone image enhancement, which contains images captured by three types of phones. We select 2,048 images and 380 images from each phone type as the training and testing sets. As described in Table 5, with the introducing of our method, the performance of DRBN can be improved, displaying the effectiveness of our algorithm for more applications.

5. Conclusion and Limitation

In this paper, we develop a framework for multiple exposure correction. An Exposure Normalization and Compensation (ENC) module is proposed to narrow the gap of multiple exposure representations, leading to consistency correction across exposures. Then, we employ the parameter regularization fine-tuning strategy to obtain a network with a balanced improvement for all the exposures. The experimental results show that our method achieves superior performance for multiple-exposure corrections. However, our method fails to incorporate a specific design for handling severe noise corruption that often appears in extremely dark conditions, which can be investigated in the future.

Acknowledgments. This work was supported by the Anhui Provincial Natural Science Foundation under Grant 2108085UD12, the JKW Research Funds under Grant 20-163-14-LZ-001-004-01, and the National Natural Science Foundation of China under Grants 62131003 and 62021001.

References

[1] M. Abdullah-Al-Wadud, M. H. Kabir, M. A. Akber Dewan, and O. Chae. A dynamic histogram equalization for image contrast enhancement. *IEEE Transactions on Consumer Electronics*, 53(2):593–600, 2007. 2

[2] Mahmoud Afifi, Konstantinos G Derpanis, Bjorn Ommer, and Michael S Brown. Learning multi-scale photo exposure correction. In *CVPR*, 2021. 2, 5, 6

[3] Vladimir Bychkovsky, Sylvain Paris, Eric Chan, and Frédo Durand. Learning photographic global tonal adjustment with a database of input output image pairs. In *CVPR*, 2011. 5, 8

[4] Bolun Cai, Xianming Xu, Kailing Guo, Kui Jia, Bin Hu, and Dacheng Tao. A joint intrinsic-extrinsic prior model for retinex. In *ICCV*, pages 4000–4009, 2017. 1, 2

[5] Jianrui Cai, Shuhang Gu, and Lei Zhang. Learning a deep single image contrast enhancer from multi-exposure images. *IEEE Transactions on Image Processing*, 27(4):2049–2062, 2018. 5

[6] Chen Chen, Qifeng Chen, Jia Xu, and Vladlen Koltun. Learning to see in the dark. *arXiv preprint arXiv:1805.01934*, 2018. 5, 6

[7] Yu-Sheng Chen, Yu-Ching Wang, Man-Hsin Kao, and Yung-Yu Chuang. Deep photo enhancer: Unpaired learning for image enhancement from photographs with gans. In *CVPR*, pages 6306–6314, 2018. 2

[8] Wenhan Yang Jiaying Liu Chen Wei, Wenjing Wang. Deep retinex decomposition for low-light enhancement. In *BMVC*, 2018. 6, 8

[9] Xuan Dong, Xiaoyan Hu, Weixin Li, Xiaojie Wang, and Yunhong Wang. Miehdr cnn: Main image enhancement based ghost-free high dynamic range imaging using dual-lens systems. In *AAAI*, volume 35, pages 1264–1272, 2021. 1

[10] Chunle Guo Guo, Chongyi Li, Jichang Guo, Chen Change Loy, Junhui Hou, Sam Kwong, and Runmin Cong. Zero-reference deep curve estimation for low-light image enhancement. In *CVPR*, pages 1780–1789, 2020. 1, 2, 6

[11] 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, 2016. 1, 2

[12] Jie Hu, Li Shen, and Gang Sun. Squeeze-and-excitation networks. In *CVPR*, pages 7132–7141, 2018. 4

[13] Jie Huang, Zhiwei Xiong, Xueyang Fu, Dong Liu, and Zheng-Jun Zha. Hybrid image enhancement with progressive laplacian enhancing unit. In *ACM MM*, pages 1614–1622, 2019. 2

[14] Xun Huang and Serge Belongie. Arbitrary style transfer in real-time with adaptive instance normalization. In *ICCV*. 3

[15] Zheng Hui, Xinbo Gao, Yunchu Yang, and Xiumei Wang. Lightweight image super-resolution with information multi-distillation network. In *ACM MM*, pages 2024–2032, 2019. 4

[16] Andrey Ignatov, Nikolay Kobyshev, Radu Timofte, Kenneth Vanhoey, and Luc Van Gool. Dslr-quality photos on mobile devices with deep convolutional networks. In *ICCV*, pages 3277–3285, 2017. 8

[17] Xin Jin, Cuiling Lan, Wenjun Zeng, Zhibo Chen, and Li Zhang. Style normalization and restitution for generalizable person re-identification. In *CVPR*, pages 3140–3149, June 2020. 1, 3

[18] James Kirkpatrick, Razvan Pascanu, Neil Rabinowitz, Joel Veness, Guillaume Desjardins, Andrei A. Rusu, Kieran Milan, John Quan, Tiago Ramalho, Agnieszka Grabska-Barwinska, Demis Hassabis, Claudia Clopath, Dharshan Kumaran, and Raia Hadsell. Overcoming catastrophic forgetting in neural networks. *Proceedings of the National Academy of Sciences*, 114(13):3521–3526, 2017. 5

[19] Edwin H Land. The retinex theory of color vision. *Scientific american*, 237(6):108–129, 1977. 2

[20] Chongyi Li, Chunle Guo, and Chen Change Loy. Learning to enhance low-light image via zero-reference deep curve estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pages 1–14, 2021. 2

[21] 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. 1, 2

[22] Jiaying Liu, Dejia Xu, Wenhan Yang, Minhao Fan, and Haofeng Huang. Benchmarking low-light image enhancement and beyond. *International Journal of Computer Vision*, 129:1153–1184, 04 2021. 2

[23] Yajing Liu, Xinmei Tian, Ya Li, Zhiwei Xiong, and Feng Wu. Compact feature learning for multi-domain image classification. In *CVPR*, pages 7193–7201, 2019. 4

[24] Yajing Liu, Zhiwei Xiong, Ya Li, Yuning Lu, Xinmei Tian, and Zheng-Jun Zha. Category-stitch learning for union domain generalization. *ACM Transactions on Multimedia Computing, Communications, and Applications*, 2022. 8

[25] Yajing Liu, Zhiwei Xiong, Ya Li, Xinmei Tian, and Zheng-Jun Zha. Domain generalization via encoding and resampling in a unified latent space. *IEEE Transactions on Multimedia*, 2021. 8

[26] Feifan Lv, Yu Li, and Feng Lu. Attention guided low-light image enhancement with a large scale low-light simulation dataset. *International Journal of Computer Vision*, 129:2175–2193, 07 2021. 2

[27] Xingang Pan, Ping Luo, Jianping Shi, and Xiaoou Tang. Two at once: Enhancing learning and generalization capacities via ibn-net. In *ECCV*, pages 464–479, 2018. 1, 3

[28] Wenqi Ren, Sifei Liu, Lin Ma, Qianqian Xu, Xiangyu Xu, Xiaochun Cao, Junping Du, and Ming-Hsuan Yang. Low-light image enhancement via a deep hybrid network. *IEEE Transactions on Image Processing*, 28(9):4364–4375, 2019. 2

[29] Xutong Ren, Wenhan Yang, Wen-Huang Cheng, and Jiaying Liu. Lr3m: Robust low-light enhancement via low-rank regularized retinex model. *IEEE Transactions on Image Processing*, 29:5862–5876, 2020. 2

[30] Ali M Reza. Realization of the contrast limited adaptive histogram equalization (clahe) for real-time image enhancement. *IEEE Transactions on Multimedia*, pages 35–44, 2004. 2, 6

[31] Qi-Chong Tian and Laurent D. Cohen. Global and local contrast adaptive enhancement for non-uniform illumination

color images. In *ICCV Workshops*, pages 3023–3030, Oct 2017. 2

[32] Laurens Van der Maaten and Geoffrey Hinton. Visualizing data using t-sne. *Journal of Machine Learning Research*, 9(11), 2008. 4

[33] Ruixing Wang, Qing Zhang, Chi-Wing Fu, Xiaoyong Shen, Wei-Shi Zheng, and Jiaya Jia. Underexposed photo enhancement using deep illumination estimation. In *CVPR*, pages 6849–6857, 2019. 1

[34] Chen Wei, Wenjing Wang, Wenhan Yang, and Jiaying Liu. Deep retinex decomposition for low-light enhancement. In *BMVC*, pages 155–165, 2018. 1, 2

[35] Jie Xiao, Man Zhou, Xueyang Fu, Aiping Liu, and Zheng-Jun Zha. Improving de-raining generalization via neural reorganization. In *ICCV*, pages 4987–4996, October 2021. 1, 3

[36] Wenhan Yang, Shiqi Wang, Yuming Fang, Yue Wang, and Jiaying Liu. From fidelity to perceptual quality: A semi-supervised approach for low-light image enhancement. In *CVPR*, pages 3063–3072, 2020. 1, 2, 3, 5, 6

[37] Wenhan Yang, Wenjing Wang, Haofeng Huang, Shiqi Wang, and Jiaying Liu. Sparse gradient regularized deep retinex network for robust low-light image enhancement. *IEEE Transactions on Image Processing*, 30:2072–2086, 2021. 2

[38] Zhenqiang Ying, Ge Li, Yurui Ren, Ronggang Wang, and Wenmin Wang. A new image contrast enhancement algorithm using exposure fusion framework. In *ICCP*, pages 36–46. Springer, 2017. 1, 2

[39] Qing Zhang, Ganzhao Yuan, Chunxia Xiao, Lei Zhu, and Wei-Shi Zheng. High-quality exposure correction of underexposed photos. In *ACM MM*, page 582–590, 2018. 2

[40] Yu Zhang, Xiaoguang Di, Bin Zhang, and Chunhui Wang. Self-supervised image enhancement network: Training with low light images only. *arXiv Preprint: 2002.11300*, 2020. 2

[41] Yonghua Zhang, Jiawan Zhang, and Xiaojie Guo. Kindling the darkness: A practical low-light image enhancer. In *ACM MM*, pages 1632–1640, 2019. 1, 2