Learning Sample Relationship for Exposure Correction

Jie Huang[*], Feng Zhao[†], Man Zhou[*†], Jie Xiao, Naishan Zheng, Kaiwen Zheng, Zhiwei Xiong
University of Science and Technology of China
{hj0117,manman,ustchbxj,nszheng,kezh}@mail.ustc.edu.cn, {fzhao956,zwxiong}@ustc.edu.cn

Abstract

Exposure correction task aims to correct the underexposure and its adverse overexposure images to the normal exposure in a single network. As well recognized, the optimization flow is the opposite. Despite great advancement, existing exposure correction methods are usually trained with a mini-batch of both underexposure and overexposure mixed samples and have not explored the relationship between them to solve the optimization inconsistency.

In this paper, we introduce a new perspective to conjunct their optimization processes by correlating and constraining the relationship of correction procedure in a mini-batch. The core designs of our framework consist of two steps: 1) formulating the exposure relationship of samples across the batch dimension via a context-irrelevant pretext task. 2) delivering the above sample relationship design as the regularization term within the loss function to promote optimization consistency. The proposed sample relationship design as a general term can be easily integrated into existing exposure correction methods without any computational burden in inference time. Extensive experiments over multiple representative exposure correction benchmarks demonstrate consistent performance gains by introducing our sample relationship design.

1. Introduction

The images captured under non-ideal illumination conditions, *i.e.,* underexposure or overexposure scenes, usually suffer from unpleasant visual effects and thus count against the down-streaming vision tasks. To this end, exposure correction techniques have been developed, which aim to correct both underexposure and overexposure images to the normal exposure automatically. It is recognized that a single algorithm is challenging to take for exposure correction since the mapping flows of correcting underexposure and overexposure are quite different [12].

Recent years have witnessed explosive advancement only on the single underexposure correction, including conventional methods that rely on manually designed strategies [2, 8, 10, 20, 28, 31], and deep-learning-driven methods that account for the powerful learning capability of complicated neural networks [32–34, 39], where deep-learning methods have achieved improvement in restoring corruptions [25, 35–37]. Seldom efforts have been devoted to both underexposure and overexposure scenes within a single algorithm for meeting the practical application. Very recently, some promising works [1, 12, 14, 24] attempt to solve the above issue. Both of them follow the common principle of alleviating the optimization process inconsistency by conjugating their exposure representations in spatial domain [12, 14, 24, 29] or in frequency transform domain [13].

In fact, most above exposure correction approaches are trained with a mini-batch that contains both underexposure and overexposure mixed samples (see Fig. 1). Within the mini-batch, the optimization process of a single network is the opposite. On the other hand, correlating the relationship of samples across the mini-batch could conjunct their optimization processes [11]. Therefore, by constraining the relationship across the batch dimension, the adverse effect of opposite optimization in the mini-batch could be relieved.

To this end, in this paper, we introduce a new perspective that conjuncts the optimization processes across the batch dimension via sample relationship learning and further improves the optimization processes of exposure correction. To achieve this, we construct an exposure-relationship learning (ERL) framework consisting of two steps (see Fig. 2). In the first step (see Fig. 3), we devise a batch-correlation module (BCM) that captures the relationship of samples across the batch dimension. To enable such a relationship focusing on the exposure-related representations, we train BCM via a pretext task that excludes context information correlation. Then, in the second step (see Fig. 4), we deliver the above sample relationship as the additional training regularization term within the loss function of exposure correction algorithms, where the relationship of corrected results is optimized on the trained BCM. In this way, the optimization processes within a mini-batch

[*]Equal contributions.
[†]Corresponding authors.

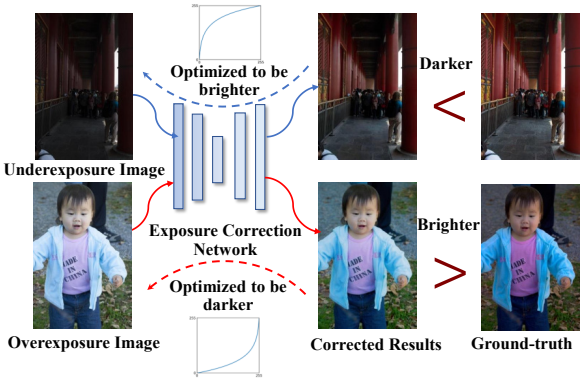


Figure 1. The illustration of optimization flow of underexposure and overexposure correction. As can be seen, the corrected results of underexposure samples are still obviously darker than normal exposure, while the corrected overexposure samples behave oppositely. This demonstrates the corrected results of underexposure and overexposure are optimized to approach their corresponding ground truth in the opposite direction.

are conjunct and the adverse effect of inconsistency optimization can be reduced.

Our proposed ERL framework is general and could be integrated with existing exposure correction approaches. The above sample relationship regularization is only adopted during the training procedure and does not introduce any computation burden in inference time. Extensive experiments on exposure correction datasets demonstrate consistent gains by applying our ERL framework. Moreover, it can also be extended to the mixed image enhancement task, demonstrating the extensive capability of our method.

We summarize the main contributions of this work as:

- This work is the first time to solve inconsistency optimization of exposure correction from a new perspective of batch dimension. By exploring the relationship of samples within a mini-batch, their optimization processes are conjunct to relieve the adverse inconsistency optimization effect.
- We propose an exposure relationship learning (ERL) framework to correlate and constrain the relationship of corrected samples across the mini-batch. The learned sample relationship acts as an additional regularization term within the loss function to assist the model optimization.
- Our ERL framework is general and can be integrated into the existing exposure correction methods without introducing any computation burden during inference.
- Extensive experiments over multiple exposure correction datasets demonstrate consistent performance gains by introducing our sample relationship learning mechanism.

2. Related Work

2.1. Conventional Methods

In the past decades, various methods have been proposed for correcting the exposure of images with poor illumination. Conventional works mainly rely on manually designed models and strategies. A line of work is based on histogram-based techniques [2, 3, 15, 16], which improve the lightness and contrast of an improper exposure image. While another line of work is based on Retinex theory [18], which decomposes an image into the reflectance component and the illumination component. By enhancing the illumination component that improves the lightness, while regularizing the reflectance component that suppresses the noises, these methods have achieved prominent effect for exposure correction [8, 10, 20, 30].

2.2. Learning-based Methods

In recent years, deep-learning-based methods have attracted more attention [41–44] which is also adopted for adjusting images with poor illumination, which are learned to correct exposures in a data-driven manner. Some of these approaches are also based on the Retinex theory [30, 32, 33, 38, 39]. As a representative, RetinexNet [32] and KIND [39] decompose an image into illumination and reflectance components, which enhances the illumination and the reflectance components to approach corresponding components of ground truth. Alternatively, another kind of components decomposed method is designed to divide the representation into different frequency bands and then enhance them progressively [21, 27, 34], such as DRBN [34]. Additionally, a few methods are also developed based on a self-supervised manner [9, 22, 23, 40]. For instance, Zero-DCE [9] formulates exposure adjustment as a task of image-specific curve estimation with a deep network. However, most of these methods are not dedicated to correcting various exposures, limiting their wide-range applications under various light conditions.

More recently, a few works have been developed for correcting both underexposure and overexposure images. This is challenging since the optimization flows of underexposure and overexposure correction are opposite. As a pioneer work, MSEC [1] proposes to correct varieties of exposures with a Laplacian pyramid architecture, which restores lightness and details in a coarse-to-fine manner. To ease the correction of wide-range exposures with inconsistency optimization flows, Huang *et al.* proposed to narrow the distribution of exposure representation with exposure normalization [12]. Similarly, CMEC [24] proposes to map different exposures to an exposure-invariant space with the assistance of a transformer for exposure correction, while ECLNet [14] derives exposure-consistency representations with the bilateral activation mechanism to as-

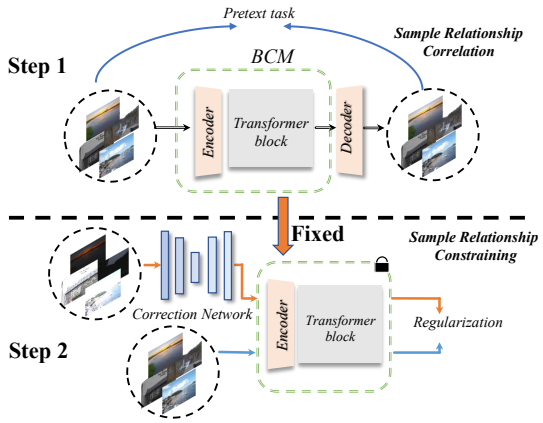


Figure 2. The overview of our proposed ERL framework, consisting of two steps. In the first step (Sec. 3.2), we formulate the exposure relationship of samples across the batch dimension on the BCM via a pretext task. In the second step (Sec. 3.3), the above-trained BCM serves as a regularization function to constrain the sample relationship of exposure correction.

sist exposure correction. To solve this problem in other viewpoints, Wang *et al.* proposed a local color distribution operator for exposure correction to meet the problem of non-uniform illumination problem [29], while FECNet [13] corrects lightness and structures progressively in a Fourier-based perspective. Although these algorithms have definitely achieved remarkable progress, there remains room for improving the optimization process across the batch dimension that contains both underexposure and overexposure corrections. In this work, we introduce a new perspective that explores the sample relationships within a mini-batch to conjunct their optimization processes, which reduces the adverse effect of opposite optimization flows.

3. Method

In this section, we first present an overview of the exposure relationship learning framework. We then introduce the proposed BCM in detail via its pretext task learning that captures sample relationships. Finally, we detail the above sample relationship as the regularization term to train exposure correction approaches.

3.1. Motivation and Overview

Given images with incorrect exposure settings, the exposure correction task aims to correct underexposure and its adverse overexposure images to normal exposure in a single network. Previous works are usually trained with a mini-batch containing both underexposure and overexposure samples, which enables the network to learn to correct various exposures simultaneously. However, the large discrepancies between underexposure and overexposure corrections would inevitably lead to opposite optimization flows across the batch dimension. As depicted in Fig. 1, the underexposure and overexposure corrected results across the mini-batch are optimized oppositely. Such a large inconsistency optimization leads to a sub-optimal optimization process, which has rarely been explored for exposure correction.

To this end, we start a new perspective for exposure correction that conjuncts the optimization of samples within a mini-batch to reduce the adverse effect of inconsistency optimization, which is built upon learning the sample relationship [11]. To achieve this, we correlate and constrain the relationship of correction procedures on the corrected results and design an ERL framework with a two-step mechanism (see Fig. 2). In the first step, it correlates the exposure relationship of samples within a mini-batch by employing the BCM. To exclude the influence of context information on the relationship, the BCM is trained with a context-irrelevant pretext task. While in the second step, the above sample relationship built upon BCM acts as an additional regularization term within the loss function of exposure correction methods, which constrains the relationship of samples. In this way, their optimization processes across the batch dimension are conjunct during training, and it would not introduce any computation burden in inference time.

3.2. Sample Relationship Correlation

As the first step of the ERL framework, we propose to correlate the sample relationship across the batch dimension, which is implemented by BCM via a context-irrelevant task.

Batch-correlation module (BCM). We design the BCM to capture the sample relationship in feature space. Its architecture is shown in Fig. 3, which consists of a feature encoder f_e with several CNN-based layers, and a sample relationship part with a transformer encoder t_e. In addition, the decoder f_d consists of several CNN-based layers are set to follow the transformer encoder that reconstructs the input image, which assists the training of BCM (see Sec. 3.3).

We then describe the workflow of BCM that captures sample relationships as follows. Let $I \in R^{B\times 3\times H\times W}$ denote a sequence of input images in a mini-batch, where B indicates the length of the sequence, and H and W are the spatial dimensions. The CNN-based encoder f_e first extracts feature from images as $F \in R^{B\times C\times H\times W}$, where C is the channel dimension. Then these features are resized to the resolution of 16×16, which can filter out most content-relevant information since low-resolution contains less content. After that, these features are reshaped to the dimension of $R^{B\times C'}$ that transfers the dimension of the feature as C'. Finally, we correlate the relationship of these features in the batch dimension with the transformer encoder t_e, and obtain the features $F' \in R^{B\times C'}$ that are correlated across the batch dimension. The whole above process is expressed as:

$$F' = t_e(f_e(I)), \tag{1}$$

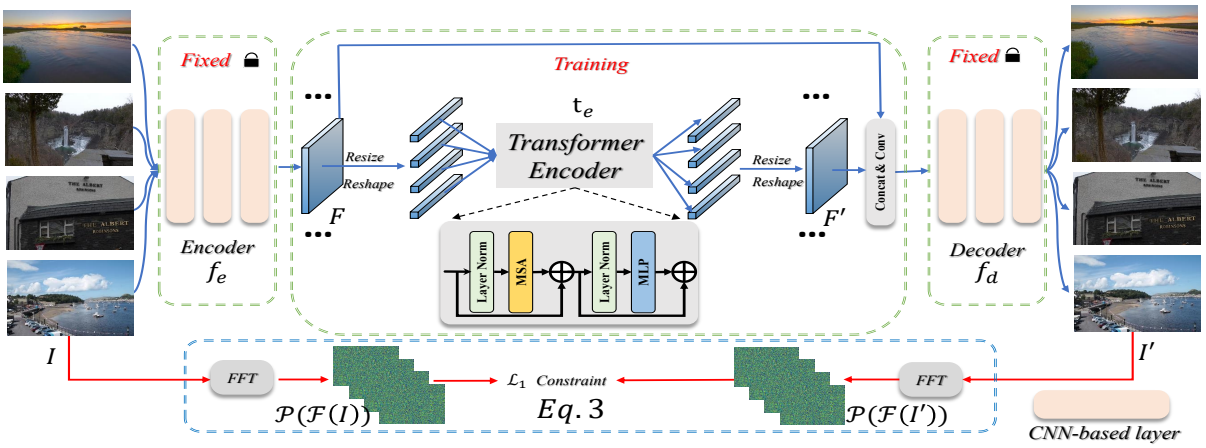


Figure 3. The illustration of our proposed BCM, which correlates the exposure relationship of samples across the batch dimension. Specifically, it consists of an encoder part f_e that extracts features and a decoder part f_d that reconstructs the image, both of which have been trained in a self-reconstruction manner. We then apply a transformer encoder t_e between the encoder and decoder parts to explore the sample relationship. Furthermore, we train the BCM with a pretext task that reconstructs the content-related phase information, leading to it focusing more on correlating exposure-related information.

where the resize and reshape operations are omitted here.

We further detail how the transformer encoder t_e explores the sample relationship across the batch dimension. Transformer usually contains a multihead self-attention layer that models the relationship of spatial and channel dimensions [7]. By transposing the feature into the dimension that the transformer layer can operate on the batch dimension, the self-attention mechanism of the multi-head self-attention layer could cross attention samples in a mini-batch, thus correlating their relationships.

Train the BCM via a context-irrelevant task. In preliminary, the encoder and decoder parts of the BCM are first trained in a self-reconstruction manner that can reconstruct input images of themselves, then these parts are fixed and a transformer encoder t_e is incorporated between them to construct BCM. Thus, when training BCM, only the transformer encoder t_e is trained to learn the sample relationship.

Since the optimization processes of the correction are mainly affected by the exposure condition, we hope the sample relationship learning correlates more with exposure-related components. To this end, the context relationship across the batch dimension should not be focused, and we now need to construct a pretext task that trains BCM to exclude the influences of context relationships.

As shown in Fig. 3, we train a context-irrelevant pretext task on the BCM via a self-reconstruction manner. For a sample correlating with other samples, it is hard to reconstruct itself since the information of other samples is introduced. Therefore, we construct the pretext task to reconstruct content-relevant information, which suppresses the correlation of the content component, thus driving the content-irrelevant information to be more correlated across the mini-batch. Specifically, a mini-batch containing normal exposure samples $I \in R^{B\times 3\times H\times W}$ are sent to the BCM, which are substantially derived correlated feature F' as depicted in Eq. 1. Then F' are reshaped and resized to the dimension of $F' \in R^{B\times C\times H\times W}$, which are further fused with F. Finally, the decoder f_d reconstructs the integrated features to the image space as I_{rec}. This procedure is expressed as:

$$I' = f_d(W(cat([F, F'])))), \tag{2}$$

where W denotes the 1×1 convolution operation, and $[\cdot]$ means the concatenate operation.

We then introduce a phase component constraint as the learning objective of the BCM. Based on FECNet [13], the amplitude component reflects the exposure-related information, while the phase component corresponds more to content information. Therefore, we construct the learning objective $\mathcal{L}_{pre}$ to constrain the phase component of the reconstructed images I' to be closed with that of the input images I:

$$\mathcal{L}_{pre} = ||\mathcal{P}(\mathcal{F}(I)) - \mathcal{P}(\mathcal{F}(I'))||_1, \tag{3}$$

where $||\cdot||_1$ denotes the mean absolutely error, $\mathcal{F}$ is the Fourier transform operation, and $\mathcal{P}(\cdot)$ represents getting the phase component. Based on this regularization, the BCM reduces the content information correlation across the mini-batch which would put a negative effect on content reconstruction, thus leading to BCM focusing more on correlating exposure-related information.

3.3. Sample Relationship Constraining

Deliver the sample relationship as regularization. After correlating the sample relationship across batch dimensions, we further deliver the sample relationship as an additional regularization term that conjuncts the optimization within a mini-batch. For implementation, when training the exposure correction network, we constrain the relationship

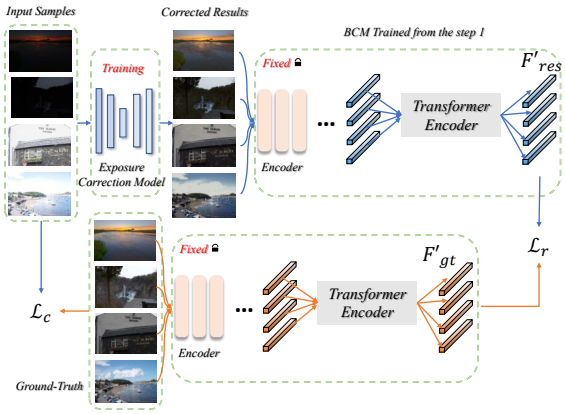


Figure 4. The illustration of delivering the sample relationship as regularization for exposure correction. We employ the encoder and transformer parts of the trained BCM as the regularization function, and integrate this regularization term $\mathcal{L}_r$ with the conventional loss functions $\mathcal{L}_c$ of exposure correction methods to optimize the exposure correction model.

of the corrected samples to be closed to that of the ground truth.

Specifically, we employ the trained BCM to construct the sample relationship within a mini-batch, where we discard its decoder part and deploy the correlated feature F' as the sample relationships representation. To this end, we apply the sample relationship regularization on the F' during the training of exposure correction, and its learning objective $\mathcal{L}_r$ is expressed as:

$$\mathcal{L}_r = ||F'_{res} - F'_{gt}||_1 \tag{4}$$

where F'_{res} and F'_{gt} denote the sample relationships representation of corrected results and ground truth, respectively.

Integrate the above regularization term to existing methods. The above sample relationship regularization term is general and could be plugged into the training paradigm of existing exposure correction algorithms as an additional optimization objective. As shown in Fig. 4, the total learning objective for training existing exposure correction method is the combination of the conventional loss $\mathcal{L}_c$ and the relationship regularization term $\mathcal{L}_r$:

$$\mathcal{L} = \mathcal{L}_c + \alpha\mathcal{L}_r \tag{5}$$

where α is the weight factor and we set it to 0.1 empirically, and the selection of α will be discussed in the ablation study. The conventional loss denotes the original loss function employed in these exposure correction methods.

3.4. Discussion

To better understand how the proposed ERL framework conjunct the optimization processes by modeling sample relationships across the batch dimension, we provide an intuitive explanation as demonstrated in Fig. 5. Given the

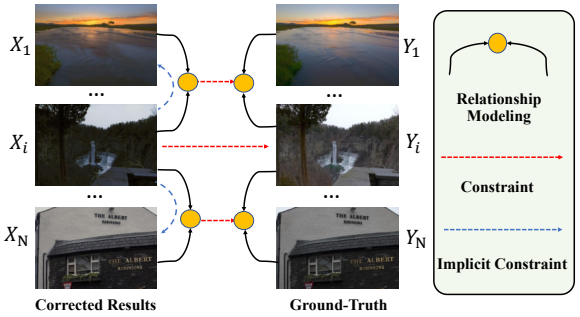


Figure 5. With the regularization of sample relationship, besides being optimized to the ground truth Y_i, the corrected results X_i are also implicitly constrained with other samples to keep their relationships to be closed with those of the ground truth.

corrected results $X(X = X_1, .., X_i, ..X_N)$ and their corresponding ground truth $Y(Y = Y_1, .., Y_i, ..Y_N)$ in a mini-batch, when the network is only optimized with the conventional loss, the sample X_i is only optimized to approach the corresponding ground truth Y_i, and their optimization flows are separated and often opposite in exposure correction.

While with the employ of sample relationship constraining, the corrected result X_i are further constrained to other corrected results X_j (where $j \neq i$) in a mini-batch, which drives their relationships to be closed to those of Y_i and Y_j (where $j \neq i$). To this end, the optimization flow of X_i is also affected by other samples, which could lead to a more consistent exposure correction effect across the mini-batch. In CMEC [24], it models consistency learning of different exposures that regularize the underexposure and overexposure representations to be similar in the network. Similarly, we consider our sample relationship regularization is another form of consistency learning, and it implicitly regularizes the results' exposure representation to be correlated and constrained in a mini-batch.

4. Experiments

4.1. Experimental Settings

Datasets. To evaluate the performance of our proposed method, we conduct experiments on two representative datasets, including the MSEC dataset proposed in MSEC [1] and SICE dataset [5] proposed in SICE. The MSEC dataset is rendered from the MIT-Adobe FiveK dataset [4], consisting of five exposure levels, where its retouched version by experts of middle exposure is set as ground truth. Finally, it consists of 17675 images for training, 750 images for validation, and 5905 images for testing. While for the SICE dataset, the middle exposure subset of this dataset is selected as the ground truth, and the second and last second exposure subsets are set as the underexposed and overexposed images, respectively. Finally, the SICE dataset contains 512 images for training and 30 images for testing.

Method	MSEC						SICE						#Param
	Under		Over		Average		Under		Over		Average		
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	
HE [26]	16.52	0.6918	16.53	0.6991	16.53	0.6959	14.69	0.5651	12.87	0.4991	13.78	0.5376	-
CLAHE [28]	16.77	0.6211	14.45	0.5842	15.38	0.5990	12.69	0.5037	10.21	0.4847	11.45	0.4942	-
RetinexNet [32]	12.13	0.6209	10.47	0.5953	11.14	0.6048	12.94	0.5171	12.87	0.5252	12.90	0.5212	0.84M
URetinexNet [33]	13.85	0.7371	9.81	0.6733	11.42	0.6988	17.39	0.6448	7.40	0.4543	12.40	0.5496	1.32M
Zero-DCE [9]	14.55	0.5887	10.40	0.5142	12.06	0.5441	16.92	0.6330	7.11	0.4292	12.02	0.5311	0.079M
Zero-DCE++ [19]	13.82	0.5887	9.74	0.5142	11.37	0.5583	11.93	0.4755	6.88	0.4088	9.41	0.4422	0.010M
DPED [17]	20.06	0.6826	13.14	0.5812	15.91	0.6219	16.83	0.6133	7.99	0.4300	12.41	0.5217	0.39M
KIND [39]	15.51	0.7115	11.66	0.7300	13.20	0.7424	13.43	0.4837	7.85	0.4779	10.64	0.4808	8.54M
SID [6]	19.37	0.8103	18.83	0.8055	19.04	0.8074	19.51	0.6635	16.79	0.6444	18.15	0.6540	7.40M
RUAS [22]	13.43	0.6807	6.39	0.4655	9.20	0.5515	16.63	0.5589	4.54	0.3196	10.59	0.4393	0.0014M
SCI [23]	9.965	0.6681	5.837	0.5190	7.49	0.5786	17.86	0.6401	4.45	0.3629	12.49	0.5051	0.0003M
MSEC [1]	20.52	0.8129	19.79	0.8156	20.08	0.8210	19.62	0.6512	17.59	0.6560	18.58	0.6536	7.04M
CMEC [24]	22.23	0.8140	22.75	0.8336	22.54	0.8257	17.68	0.6592	18.17	0.6811	17.93	0.6702	5.40M
SID-ENC [12]	22.59	0.8423	22.36	0.8519	22.45	0.8481	21.36	0.6652	19.38	0.6843	20.37	0.6748	7.45M
LCDPNet [29]	22.35	0.8650	22.17	0.8476	22.30	0.8552	17.45	0.5622	17.04	0.6463	17.25	0.6043	0.96M
DRBN [34]	19.74	0.8290	19.37	0.8321	19.52	0.8309	17.96	0.6767	17.33	0.6828	17.65	0.6798	0.53M
DRBN+ERL	**19.91**	**0.8305**	**19.60**	**0.8384**	**19.73**	**0.8355**	**18.09**	0.6735	**17.93**	**0.6866**	**18.01**	0.6796	0.53M
DRBN-ENC [12]	22.72	0.8544	22.11	0.8521	22.35	0.8530	21.89	0.7071	19.09	0.7229	20.49	0.7150	0.58M
DRBN-ENC+ERL	**22.92**	**0.8704**	**22.45**	**0.8724**	**22.67**	**0.8716**	**22.06**	0.7053	**19.50**	0.7205	**20.78**	0.7129	0.58M
ECLNet [14]	22.37	0.8566	22.70	0.8673	22.57	0.8631	22.05	0.6893	19.25	0.6872	20.65	0.6861	0.018M
ECLNet+ERL	**22.90**	**0.8624**	22.58	**0.8676**	**22.70**	**0.8655**	**22.14**	**0.6908**	**19.47**	**0.6982**	**20.81**	**0.6945**	0.018M
FECNet [13]	22.96	0.8598	23.22	0.8748	23.12	0.8688	22.01	0.6737	19.91	0.6961	20.96	0.6849	0.15M
FECNet+ERL	**23.10**	**0.8639**	23.18	**0.8759**	**23.15**	**0.8711**	**22.35**	0.6671	**20.10**	0.6891	**21.22**	0.6781	0.15M

Table 1. Quantitative results of different methods on the MSEC and the SICE datasets in terms of PSNR and SSIM. **Bold** denotes the results that are improved by incorporating ERL. The underline denotes the best result.

Implementation details. We implement our proposed model with PyTorch on a single NVIDIA GTX 3090 GPU. Firstly, we train the BCM on the normal exposure images for 100 epochs that captures the exposure sample relationship. Then, to couple the above sample relationship as additional regularization of existing methods, we select several methods as equipped methods, including DRBN [34], DRBN-ENC-4 [12], ECLNet [14] and FECNet [13]. After the integration, we add + between their names and ERL to denote the improved method (such as DRBN+ERL).

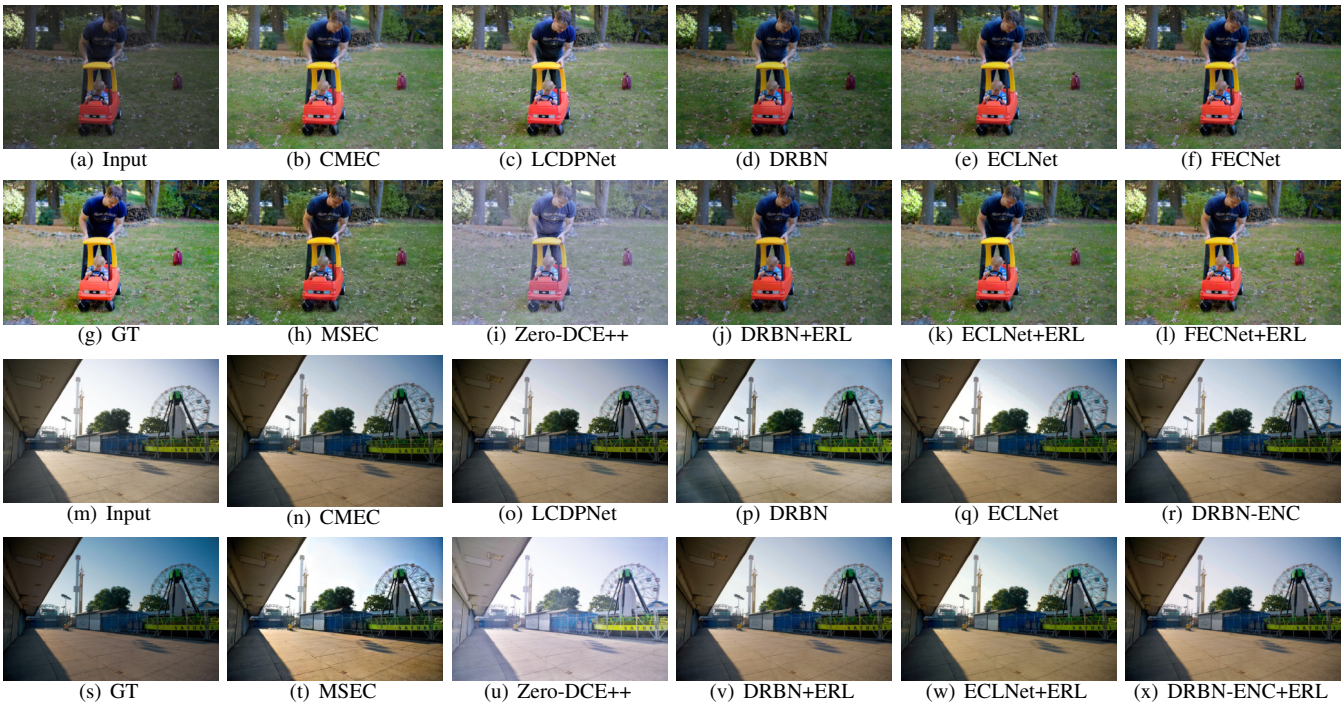
During training, we employ an Adam optimizer with $\beta_1 = 0.9$, $\beta_2 = 0.99$ to update the model. We set the training settings following the original settings of the above-equipped methods, including the update of the learning rate, the number of training epochs, and the patch size. All of the loss functions for exposure correction are optimized in an end-to-end manner. During the evaluation, we implement Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index (SSIM) for numerical evaluation.

4.2. Comparison with State-of-the-Art Methods

To evaluate the performance with existing exposure correction methods, several methods are adopted as a comparison, including two traditional methods HE [26] and CLAHE [28]. Additionally, several deep-learning-based methods are involved as well, including RetinexNet [32], DPED [17], SID [6], KinD [39], DRBN [34], URetinexNet [33], Zero-DCE [9], Zero-DCE++ [19], RUAS [22] and SCI [23]. We further include several specific approaches for exposure correction: MSEC [1], CMEC [24], SID-ENC [12], DRBN-ENC [12], ECLNet [14], LCDPNet [29] and FECNet [13].

Quantitative results. The quantitative results are shown in Table 1. For the MSEC dataset, following MSEC, we average the results of the exposures of the first two levels and the remaining levels of exposure as underexposure and overexposure results, respectively. As can be observed, recently proposed exposure correction approaches (such as FECNet and ECLNet) achieve excellent performance on the two datasets. With the incorporation of the ERL framework, their overall quantitative results are further improved on most subsets and metrics. It substantiates the effectiveness and flexibility of our proposed sample relationship regularization. Note that the introducing of our sample relationship regularization would not increase parameters or computation costs. In addition, the effectiveness of our framework can also be verified to improve the performance during the training procedure as shown in Fig. 7. As can be seen, our ERL framework improves performance during most training time with higher PSNR, demonstrating it improves optimization processes and thus lead to better performance.

Qualitative results. We provide visual comparisons on the MSEC dataset in Fig. 6, and visual comparisons on the SICE dataset in Fig. 8, respectively. As can be seen, with the employ of the sample relationship regularization term, the global lightness or local structures and details of the en-



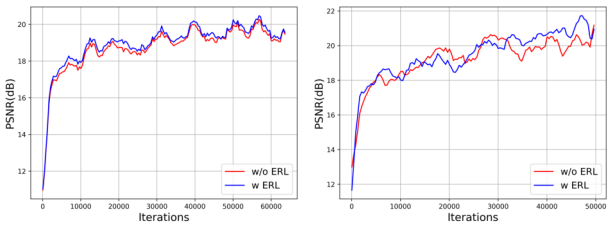
(a) Input (b) CMEC (c) LCDPNet (d) DRBN (e) ECLNet (f) FECNet

(g) GT (h) MSEC (i) Zero-DCE++ (j) DRBN+ERL (k) ECLNet+ERL (l) FECNet+ERL

(m) Input (n) CMEC (o) LCDPNet (p) DRBN (q) ECLNet (r) DRBN-ENC

(s) GT (t) MSEC (u) Zero-DCE++ (v) DRBN+ERL (w) ECLNet+ERL (x) DRBN-ENC+ERL

Figure 6. Visualization results on the MSEC dataset of (top) underexposure correction and (bottom) overexposure correction.



(a) Training of ECLNet (b) Training of DRBN-ENC

Figure 7. Graph of PSNR with/without ERL framework during the training process of the ECLNet and DRBN-ENC.

hanced results are further improved. Therefore, these results prove our method could improve existing methods to generate better visual results. We provide more visual results in the supplementary materials.

4.3. Ablation Studies

We conduct ablation studies on the SICE dataset to further investigate the effectiveness of our proposed sample relationship modeling under different configurations, and we set the baseline as training the exposure correction method ECLNet on its original settings.

Investigation of the BCM. To validate the effectiveness of the BCM design, we perform ablations with different configurations of the BCM and present the results in Table 2. As can be seen, without the employ of the pretext task (which means the t_e in the BCM is randomly initialized) or training the pretext task using pixel L_1 loss, the performance can also be improved compared with the baseline. This demonstrates that the sample relationship regularization mechanism itself is effective for improving performance, while the designing of our pretext task can further elevate the exposure correction effect.

Option	Baseline	w/o pretext task	with $\mathcal{L}_1$	with $\mathcal{L}_{pre}$
PSNR	20.65	20.76	20.72	20.81
SSIM	0.6861	0.6912	0.6897	0.6945

Table 2. Ablation study of investigating the configuration of the BCM based on the ECLNet in SICE dataset.

Investigation of weight coefficient. To investigate the influences of the weight coefficient α on sample relationship regularization term, we perform experiment experiments with setting different coefficients α. As shown in Fig. 9 (a), larger α could lead to a slight performance decrease. Overall, the performance remains stable and is robust to the weight coefficient changes.

Investigation of batch size. We perform ablation studies to investigate the batch size of sample relationship regularization for optimizing exposure correction in Fig. 9 (b). The results show that setting the batch size of 4 improves more performance than other settings, while a larger batch size can not improve performance significantly. This may be caused by the sample relationship of large batch size being hard to optimize due to the high variance.

4.4. Extension on the Mixed Enhancement Task

To further demonstrate the extensive ability of our method, we further train and test the method on an image

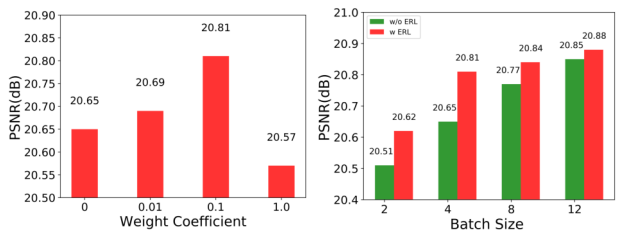
(a) Input (b) CMEC (c) LCDPNet (d) DRBN (e) ECLNet (f) FECNet

(g) GT (h) MSEC (i) Zero-DCE++ (j) DRBN+ERL (k) ECLNet+ERL (l) FECNet+ERL

(m) Input (n) CMEC (o) LCDPNet (p) DRBN (q) ECLNet (r) DRBN-ENC

(s) GT (t) MSEC (u) Zero-DCE++ (v) DRBN+ERL (w) ECLNet+ERL (x) DRBN-ENC+ERL

Figure 8. Visualization results on the SICE dataset of (top) underexposure correction and (bottom) overexposure correction.



(a) Ablation of weight coefficient α (b) Ablation of batch size

Figure 9. Ablation study for investigating the factors of (a) the weight coefficient α and (b) the batch size of our method based on the ECLNet network.

Method	LOL	SICE (over)	FiveK	Average
DRBN	19.39/0.817	19.46/0.729	21.77/0.859	20.21/0.801
DRBN+ERL	**19.84/0.824**	19.34/**0.730**	**22.14/0.873**	**20.44/0.809**
ECLNet	21.54/0.806	19.98/0.728	23.54/0.851	21.69/0.795
ECLNet+ERL	**22.01/0.827**	19.73/0.714	**23.71/0.853**	**21.82/0.798**
FECNet	22.03/0.836	19.78/0.718	23.82/0.849	21.88/0.800
FECNet+ERL	21.08/0.829	**20.57/0.722**	**24.18/0.864**	**21.94/0.805**

Table 3. Quantitative results of applying our framework on exposure correction methods for multiple enhancement tasks. **Bold** denotes the results that are improved by incorporating the ERL.

mixed enhancement task containing different enhancement tasks, where the optimization flows of them are quite different. Specifically, we blend the LOL dataset [32] designed for low-light enhancement, MIT-FiveK dataset [4] collected for image retouching, and the overexposure subset from the SICE dataset to build a Task-mix dataset. As shown in Table 3, with the equipped with our sample relationship regularization, the overall performance is improved. We provide visual results in the supplementary materials.

5. Conclusion

In this paper, for the first time, we explore the sample relationship within a mini-batch to meet the optimization inconsistency problem of exposure correction. Thus, we introduce a framework that conjunct the optimization process of exposure correction with two steps. The first step correlates the exposure relationship of samples via a context-irrelevant pretext task, while the second step constrains the relationship of corrected results by delivering the above sample relationship as a regularization term within the loss function. Our sample relationship design is compatible with existing exposure correction approaches that serve as a general regularization term. Extensive experiments demonstrate our framework consistently boosts representative methods' performance on exposure correction datasets. Although the performance improvement is not very significant in some cases, we hope our approach would shed light on investigating the sample relationship for improving the performance of low-level vision problems in the future.

Acknowledgments. This work was supported by the JKW Research Funds under Grant 20-163-14-LZ-001-004-01, and the Anhui Provincial Natural Science Foundation under Grant 2108085UD12. We acknowledge the support of GPU cluster built by MCC Lab of Information Science and Technology Institution, USTC.

References

[1] Mahmoud Afifi, Konstantinos G Derpanis, Bjorn Ommer, and Michael S Brown. Learning multi-scale photo exposure correction. In *Proceedings of the IEEE International Conference on Computer Vision*, 2021. 1, 2, 5, 6

[2] Tarik Arici, Salih Dikbas, and Yucel Altunbasak. A histogram modification framework and its application for image contrast enhancement. *IEEE Transactions on Image Processing*, 18(9):1921–1935, 2009. 1, 2

[3] Tarik Arici, Salih Dikbas, and Yucel Altunbasak. A histogram modification framework and its application for image contrast enhancement. *IEEE Transactions on image processing*, 18(9):1921–1935, 2009. 2

[4] Vladimir Bychkovsky, Sylvain Paris, Eric Chan, and Frédo Durand. Learning photographic global tonal adjustment with a database of input/output image pairs. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 97–104, 2011. 5, 8

[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. 6

[7] Alexey Dosovitskiy, Lucas Beyer, Alexander Kolesnikov, Dirk Weissenborn, Xiaohua Zhai, Thomas Unterthiner, Mostafa Dehghani, Matthias Minderer, Georg Heigold, Sylvain Gelly, et al. An image is worth 16x16 words: Transformers for image recognition at scale. *arXiv preprint arXiv:2010.11929*, 2020. 4

[8] Xueyang Fu, Delu Zeng, Yue Huang, Xiao-Ping Zhang, and Xinghao Ding. A weighted variational model for simultaneous reflectance and illumination estimation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 2782–2790, 2016. 1, 2

[9] Chunle 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 *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1780–1789, 2020. 2, 6

[10] 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

[11] Zhi Hou, Baosheng Yu, and Dacheng Tao. Batchformer: Learning to explore sample relationships for robust representation learning. In *CVPR*, 2022. 1, 3

[12] Jie Huang, Yajing Liu, Xueyang Fu, Man Zhou, Yang Wang, Feng Zhao, and Zhiwei Xiong. Exposure normalization and compensation for multiple-exposure correction. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 6043–6052, 2022. 1, 2, 6

[13] Jie Huang, Yajing Liu, Feng Zhao, Keyu Yan, Jinghao Zhang, Yukun Huang, Man Zhou, and Zhiwei Xiong. Deep fourier-based exposure correction network with spatial frequency interaction. In *Proceedings of the European Conference on Computer Vision (ECCV)*, 2022. 1, 3, 4, 6

[14] Jie Huang, Man Zhou, Yajing Liu, Mingde Yao, Feng Zhao, and Zhiwei Xiong. Exposure-consistency representation learning for exposure correction. In *Proceedings of the 30th ACM International Conference on Multimedia*, page 6309–6317, 2022. 1, 2, 6

[15] Haidi Ibrahim and Nicholas Sia Pik Kong. Brightness preserving dynamic histogram equalization for image contrast enhancement. *IEEE Transactions on Consumer Electronics*, 53(4):1752–1758, 2007. 2

[16] Haidi Ibrahim and Nicholas Sia Pik Kong. Brightness preserving dynamic histogram equalization for image contrast enhancement. *IEEE Transactions on Consumer Electronics*, 53(4):1752–1758, 2007. 2

[17] 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. 6

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

[19] Chongyi Li, Chunle Guo, and Chen Change Loy. Learning to enhance low-light image via zero-reference deep curve estimation. *arXiv preprint arXiv:2103.00860*, 2021. 6

[20] 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

[21] Seokjae Lim and Wonjun Kim. DSLR: Deep stacked laplacian restorer for low-light image enhancement. *IEEE Transactions on Multimedia*, 23:4272–4284, 2020. 2

[22] Risheng Liu, Long Ma, Jiaao Zhang, Xin Fan, and Zhongxuan Luo. Retinex-inspired unrolling with cooperative prior architecture search for low-light image enhancement. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 10561–10570, 2021. 2, 6

[23] Long Ma, Tengyu Ma, Risheng Liu, Xin Fan, and Zhongxuan Luo. Toward fast, flexible, and robust low-light image enhancement. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 5637–5646, June 2022. 2, 6

[24] Ntumba Elie Nsamp, Zhongyun Hu, and Qing Wang. Learning exposure correction via consistency modeling. In *Proceedings of the British Machine Vision Conference (BMVC)*, pages 1–12, 2018. 1, 2, 5, 6

[25] Zhihong Pan, Baopu Li, Dongliang He, Mingde Yao, Wenhao Wu, Tianwei Lin, Xin Li, and Errui Ding. Towards bidirectional arbitrary image rescaling: Joint optimization and cycle idempotence. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 17389–17398, 2022. 1

[26] Ioannis Pitas. *Digital image processing algorithms and applications*. John Wiley & Sons, 2000. 6

[27] Xutong Ren, Mading Li, Wen-Huang Cheng, and Jiaying Liu. Joint enhancement and denoising method via sequential decomposition. In *Proceedings of the IEEE International Symposium on Circuits and Systems*, pages 1–5. IEEE, 2018. 2

[28] 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. 1, 6

[29] Haoyuan Wang, Ke Xu, and Rynson W.H. Lau. Local color distributions prior for image enhancement. In *Proceedings of the European Conference on Computer Vision (ECCV)*, 2022. 1, 3, 6

[30] Ruixing Wang, Qing Zhang, Chi-Wing Fu, Xiaoyong Shen, Wei-Shi Zheng, and Jiaya Jia. Underexposed photo enhancement using deep illumination estimation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 6849–6857, 2019. 2

[31] Shuhang Wang, Jin Zheng, Hai-Miao Hu, and Bo Li. Naturalness preserved enhancement algorithm for non-uniform illumination images. *IEEE Transactions on Image Processing*, 22(9):3538–3548, 2013. 1

[32] Chen Wei, Wenjing Wang, Wenhan Yang, and Jiaying Liu. Deep Retinex decomposition for low-light enhancement. *arXiv preprint arXiv:1808.04560*, 2018. 1, 2, 6, 8

[33] Wenhui Wu, Jian Weng, Pingping Zhang, Xu Wang, Wenhan Yang, and Jianmin Jiang. Uretinex-net: Retinex-based deep unfolding network for low-light image enhancement. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 5901–5910, June 2022. 1, 2, 6

[34] 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 *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 3063–3072, 2020. 1, 2, 6

[35] Mingde Yao, Dongliang He, Xin Li, Zhihong Pan, and Zhiwei Xiong. Bidirectional translation between uhd-hdr and hd-sdr videos. *IEEE Transactions on Multimedia*, 2023. 1

[36] Mingde Yao, Zhiwei Xiong, Lizhi Wang, Dong Liu, and Xuejin Chen. Spectral-depth imaging with deep learning based reconstruction. *Optics express*, 27(26):38312–38325, 2019. 1

[37] Wenlong Zhang, Yihao Liu, Chao Dong, and Yu Qiao. RankSRGAN: Generative adversarial networks with ranker for image super-resolution. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 3096–3105, 2019. 1

[38] Yonghua Zhang, Xiaojie Guo, Jiayi Ma, Wei Liu, and Jiawan Zhang. Beyond brightening low-light images. *International Journal of Computer Vision*, 129(4):1013–1037, 2021. 2

[39] Yonghua Zhang, Jiawan Zhang, and Xiaojie Guo. Kindling the darkness: A practical low-light image enhancer. In *Proceedings of the ACM International Conference on Multimedia*, pages 1632–1640, 2019. 1, 2, 6

[40] Anqi Zhu, Lin Zhang, Ying Shen, Yong Ma, Shengjie Zhao, and Yicong Zhou. Zero-shot restoration of underexposed images via robust Retinex decomposition. In *Proceedings of the IEEE International Conference on Multimedia and Expo*, pages 1–6, 2020. 2

[41] Yurui Zhu, Xueyang Fu, Chengzhi Cao, Xi Wang, Qibin Sun, and Zheng-Jun Zha. Single image shadow detection via complementary mechanism. In *Proceedings of the 30th ACM International Conference on Multimedia*, pages 6717–6726, 2022. 2

[42] Yurui Zhu, Xueyang Fu, and Aiping Liu. Learning dual transformation networks for image contrast enhancement. *IEEE Signal Processing Letters*, 27:1999–2003, 2020. 2

[43] Yurui Zhu, Xi Wang, Xueyang Fu, and Xiaowei Hu. Enhanced coarse-to-fine network for image restoration from under-display cameras. In *Computer Vision–ECCV 2022 Workshops: Tel Aviv, Israel, October 23–27, 2022, Proceedings, Part V*, pages 130–146. Springer, 2023. 2

[44] Yurui Zhu, Zeyu Xiao, Yanchi Fang, Xueyang Fu, Zhiwei Xiong, and Zheng-Jun Zha. Efficient model-driven network for shadow removal. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 36, pages 3635–3643, 2022. 2