Generalized Lightness Adaptation with Channel Selective Normalization

Mingde Yao[1]*, Jie Huang[1]*, Xin Jin[2], Ruikang Xu[1], Shenglong Zhou[1], Man Zhou[3], Zhiwei Xiong[1]†
[1]University of Science and Technology of China
[2]Eastern Institute of Technology
[3]Nanyang Technological University

Abstract

Lightness adaptation is vital to the success of image processing to avoid unexpected visual deterioration, which covers multiple aspects, e.g., low-light image enhancement, image retouching, and inverse tone mapping. Existing methods typically work well on their trained lightness conditions but perform poorly in unknown ones due to their limited generalization ability. To address this limitation, we propose a novel generalized lightness adaptation algorithm that extends conventional normalization techniques through a channel filtering design, dubbed Channel Selective Normalization (CSNorm). The proposed CSNorm purposely normalizes the statistics of lightness-relevant channels and keeps other channels unchanged, so as to improve feature generalization and discrimination. To optimize CSNorm, we propose an alternating training strategy that effectively identifies lightness-relevant channels. The model equipped with our CSNorm only needs to be trained on one lightness condition and can be well generalized to unknown lightness conditions. Experimental results on multiple benchmark datasets demonstrate the effectiveness of CSNorm in enhancing the generalization ability for the existing lightness adaptation methods. Code is available at https://github.com/mdyao/CSNorm.

1. Introduction

Lightness adaptation is a vital step in image processing, encompassing tasks such as low-light image enhancement [27, 31], image retouching [29], and inverse tone mapping [7]. These tasks have benefited significantly from the development of advanced neural network architectures. Although numerous powerful lightness adaptation methods have been proposed, the generalization problem [56, 46] for lightness adaptation still exists and is rarely explored.

In real-world applications, applying lightness adaptation models to unknown lightness conditions is quite challenging due to the brightness discrepancies between training and testing data [39, 52]. Existing lightness adaptation approaches [31, 65, 59, 32, 1, 49, 7, 24] primarily focus on addressing the challenge of accurate image reconstruction. However, they often underperform on wide-range scenes with other lightness conditions due to their over-fitting to the training lightness component, leading to unsatisfactory visual effects (Fig. 1) and inadequate generalization in complex real-world scenarios.

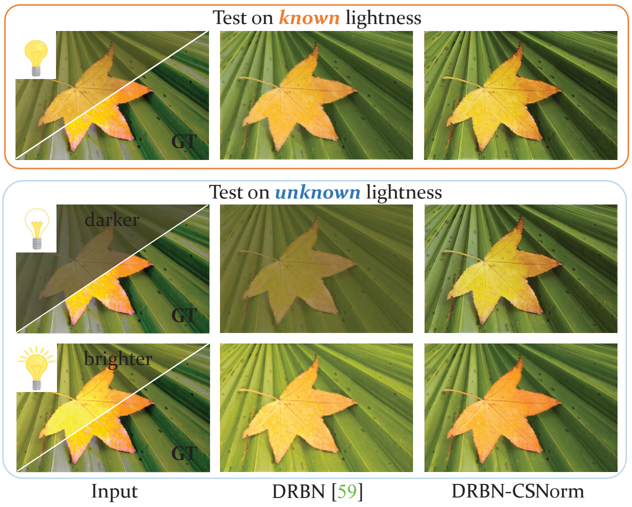


Figure 1: Visual comparisons on *known* and *unknown* lightness conditions. The model equipped with our CSNorm can generalize well to other *unknown* lightness while keeping the performance on the *known* lightness.

An alternative way is constructing a larger mixed-lightness dataset including more lightness conditions, but it is impractical for many complicated cases and too time-consuming for cumbersome acquisition from diverse domains [47, 67]. Besides, existing models suffer from the drawback of inadequately encapsulating generalization and discrimination abilities, where the former is responsible for the performance on unknown lightness conditions and the latter mainly corresponds to the reconstruction characteris-

*Co-first author.
†Corresponding author: zwxiong@ustc.edu.cn.

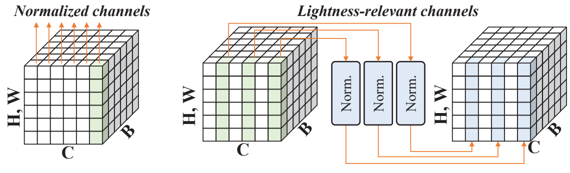


Figure 2: Comparisons of different normalization techniques. Our proposed CSNorm adaptively normalizes the lightness-relevant channels for generalization and keeps other channels unchanged for accurate reconstruction.

tics on the known lightness condition.

In this paper, we focus on designing a mechanism that empowers existing lightness adaptation methods with the generalization ability to wide-range lightness scenes. The key challenge lies in obtaining the above generalization ability while keeping the discrimination ability. To achieve this goal, we introduce the normalization technique, which has the good property of extracting invariant representations from the given features [22], especially for lightness components [40, 17]. However, normalization is a *double-edged sword* due to its inevitable loss of information, which might degrade the reconstruction accuracy [40, 68]. Therefore, we explore normalizing particular channels that are highly sensitive to lightness changes while keeping other channels unchanged (Fig. 2). Such a design enhances the generalization ability and keeps the discrimination of features.

To this end, we propose a concept of *Channel Selection Normalization* (CSNorm) to purposely select and normalize the lightness-relevant channels. CSNorm consists of two major parts: an instance-level lightness normalization module for eliminating lightness-relevant information and a differentiable gating module for adaptively selecting lightness-relevant channels. The gating module outputs a series of binary indicators to combine the normalized and original channels, which feasibly enhances the model's generalization capabilities and mitigates the information loss caused by normalization. The proposed CSNorm is simple, lightweight, and plug-and-play.

To identify lightness-relevant channels in CSNorm, we meticulously design an alternating training strategy. The network is alternately optimized with different inputs of two steps. Specifically, in the first step, the network inputs the images to learn an essential ability for lightness adaptation. In the second step, we slightly perturb the lightness of the above input image and solely optimize CSNorm with other parameters frozen. Since the only variable in the input images is the lightness condition, CSNorm can adaptively identify lightness-relevant channels and normalize them accordingly, thereby exhibiting superior performance in terms of generalization and discrimination.

In summary, we make the following contributions.

- To our best knowledge, this is the first work that improves the generalization ability of lightness adaptation methods in wide-range lightness scenarios.
- We propose CSNorm, which selectively normalizes the lightness-relevant channels according to their sensitivity to lightness changes. The model equipped with our CSNorm can generalize well to unknown lightness conditions while keeping the reconstruction ability on known lightness conditions.
- An alternate training strategy is meticulously designed to effectively optimize CSNorm for identifying lightness-relevant channels.
- We conduct extensive experiments to validate the advantage and versatility of CSNorm over existing lightness adaptation methods for improving their generalization in wide-range lightness scenarios.

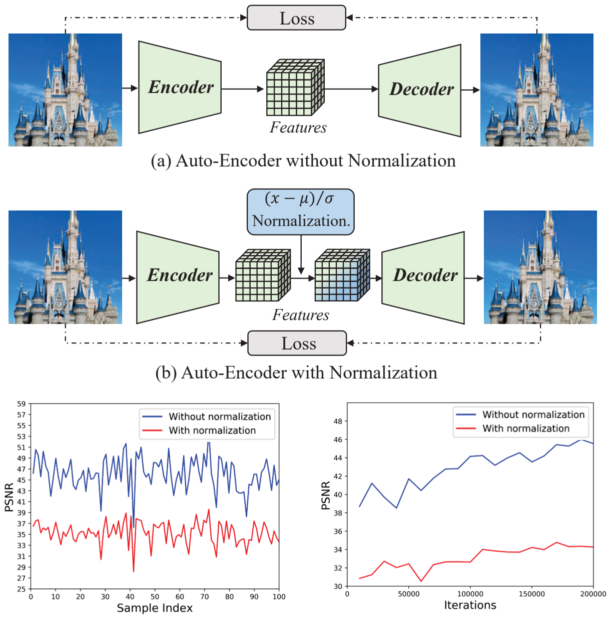
2. Related Work

2.1. Lightness Adaptation

As a key step in image restoration [66, 41, 12, 57, 33], lightness adaptation tasks [58, 23, 64, 19, 1], such as low-light image enhancement [27, 31, 64, 65, 59, 32, 34], image retouching [29, 63], and inverse tone mapping [35, 7, 61, 24, 61], aim to adjust lightness components (e.g., illumination, color, and dynamic range) from a degraded version to a normal version. Low-light image enhancement aims to improve the visibility and quality of images captured under low-light conditions. In recent years, various deep learning-based approaches [31, 65, 59, 32, 43] have shown promising results in this field. For image retouching, CSRNet [16] formulates pixel-independent operations by multi-layer perceptrons (MLPs), which learns implicit step-wise retouching operations. Inverse tone mapping aims to translate images from high dynamic range to low dynamic range. HDRTVNet [7] proposes a multi-stage method to adjust the global intensity and local contrast step-by-step. SR-ITM [24] proposes a dynamic filter to jointly learn the super-resolution and inverse tone mapping with a single network. Although prior works have made significant progress on lightness adaptation, they inherently tend to overfit the training data, resulting in poor generalization performance. Our proposed CSNorm enables the model to generalize to unknown lightness conditions, which only needs to be trained with limited lightness conditions and avoids time-consuming data collection. Besides, its lightweight and plug-and-play nature allows for easy integration into various networks.

2.2. Generalization

A well-generalized model exhibits the ability to infer meaningful patterns, relationships, and features from its training data and apply them effectively to new, unseen



(a) Auto-Encoder without Normalization

(b) Auto-Encoder with Normalization

(c) Analysis on Performance

Figure 3: The effect of applying normalization for image reconstruction. As can be seen, the introduction of instance normalization is harmful to image reconstruction.

instances. Various methods have been developed to address this problem, including domain generalization [46, 4], self-supervised learning [60, 28, 11], unsupervised learning [25], contrastive learning [30], and zero-shot learning [56]. In this paper, we focus on the generalization problem of lightness adaptation tasks. Domain generalization [46, 4, 20] aims to learn the domain-invariant representation from multiple source domains which could generalize well on unseen domains. Existing methods tried to address it mainly from the dataset synthesis aspect [47, 67] or optimization algorithm aspect [26, 8, 9]. Beyond domain generalization, single domain generalization [42, 10] has gained interest recently. This task aims to learn the model from one source domain to get a well generalization ability on other unseen domains. Following the development trajectory of domain generalization, methods based on adversarial domain augmentation have been proposed. However, our proposed method is distinct from existing approaches in that it is simple, lightweight, and specifically designed to meet the lightness adaptation requirements.

2.3. Normalization

Normalization plays an essential role in image processing, especially for lightness-relevant tasks. Formally, normalization subtracts the mean value and is derived by standard to scale the image in classical image processing. In deep learning-based methods, normalization serves as a basic layer [21, 54, 50, 2] with various varieties. Recently works [10, 48, 45, 38] point that normalization has a good property of generalization for neural networks. BN-Test [37] calibrats the model under covariate shift at the test stage. ASR-Norm [10] adapts to each individual input sample to avoid dependency on the testing samples. Though normalization has been investigated in high-level vision, it is rarely explored in the low-level lightness adaptation field.

3. Motivation

Since lightness differs and varies substantially in real-world captures, the processing of the lightness adaptation method is significantly variable. Consequently, it is challenging to directly deploy existing networks for real-world scenarios, particularly in lightness conditions absent from the training set. An alternative way is to increase the dataset's capacity by creating an enlarged mixed dataset, including extra lightness conditions for training. However, the exorbitant cost of the data collection makes it a challenging proposition to pursue. Furthermore, the mixed dataset has a greater propensity to cause ambiguity during training, which might bias the network toward particular lightness and result in an imbalanced training issue [55].

Normalization has good properties of eliminating lightness-relevant components [17] and reducing the discrepancy between images [36]. It can effectively lessen the impact of lightness and competently extract lightness-independent information, which enables the network to learn robust representations and improve the generalization ability. Based on this point, we aim to present a general normalization algorithm to address the generalization problem for lightness adaptation.

Despite these benefits, normalization is a *double-edged sword* for networks due to the inevitable loss of information (*e.g.*, statistical characteristics including mean and variance) [40, 68], resulting in inferior reconstruction performance. To comprehend the influence of normalization intuitively, we conduct a self-reconstruction task to illustrate the information loss induced by normalization. As shown in Fig. 3, we train two auto-encoder networks separately with and without inserting the IN [50] operation and we calculate the relative reconstruction accuracy (*i.e.,* PSNR) on different images. It is obvious that normalization ruins the network's reconstruction ability, and in fact, the harm of information loss caused by normalization outweighs its potential benefits in terms of generalization. This motivates us to design CSNorm to selectively normalize the channels, concurrently considering the generalization ability and reconstruction accuracy for lightness adaptation.

4. Method

4.1. Overview

Based on the above analysis, we propose a simple yet effective method as shown in Fig. 4. Particularly, we design

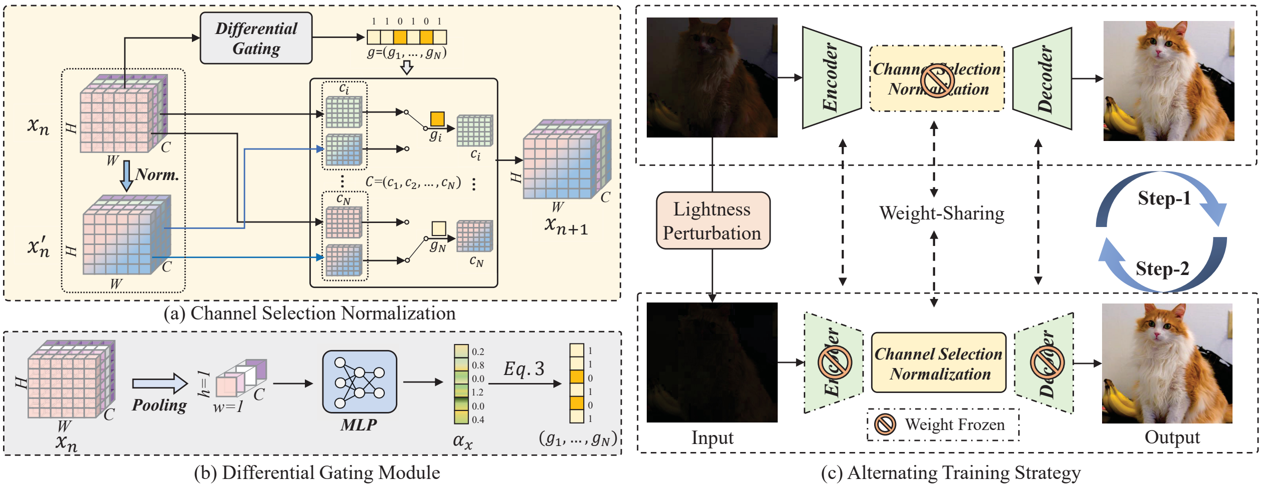


Figure 4: Overview of our proposed method. (a) Channel selective normalization (CSNorm), which consists of an instance-level normalization module and a differential gating module. (b) Differential gating module. It outputs a series of on-off switch gates for binarized channel selection in CSNorm. (c) Alternating training strategy. In the first step, we optimize the parameters outside CSNorm to keep an essential ability for lightness adaptation. In the second step, we only update the parameters inside CSNorm (see (a)&(b)) with lightness-perturbed images. The two steps drive CSNorm to select channels sensitive to lightness changes, which are normalized in x_{n+1}.

CSNorm (Fig. 4a) to improve the network's capacity for generalization, which can be used as a plug-and-play module for existing lightness adaptation networks. In CSNorm, a differentiable gating module (Fig. 4b) is introduced to efficiently select the original and normalized features along the channel dimension and then combine them to be passed to the next layer. Such a gating module operates as a de facto channel-selection function.

Further, we propose an alternating training strategy to force the gating module to select lightness-relevant channels, which is driven by performance stability under lightness perturbations (Fig. 4c). During the training stage, we only access one dataset with limited lightness conditions and train the model equipped with our CSNorm. Once trained, the model can work directly on other unknown lightness conditions.

4.2. Channel Selective Normalization

As depicted in Fig. 4a, CSNorm consists of two parts: an instance-level lightness normalization module for eliminating light-relevant information and a differentiable gating module for adaptively selecting light-relevant channels.

4.2.1 Instance-level Lightness Normalization

To facilitate the subsequent selection of channels, we normalize the channels and adopt IN as the implementation to operate precisely on individual instances and channels. Given a feature x with the shape of $H \times W \times C$, IN normalizes x by subtracting the mean value $\mu(x)$ followed by dividing the standard deviation $\sigma(x)$, which is expressed as

$$x' = \mathrm{IN}(x) = \gamma \frac{x - \mu(x)}{\sigma(x)} + \beta, \tag{1}$$

where $\mu(x)$ and $\sigma(x)$ are calculated independently across spatial dimensions for each channel and instance, and $\gamma, \beta \in \mathbb{R}^C$ are scalable parameters learned from data. Since IN can reduce the lightness discrepancy among instances [22], the normalized feature x' has a robust representation irrelevant to the lightness conditions, enabling the network to adapt to various lightness scenarios and improving its generalization capability.

4.2.2 Differentiable Gating Module

To achieve adaptive channel selection with minimum network modification costs, we introduce a differentiable gating module for channel selection, which feasibly enhances the model's generalization capabilities and mitigates the information loss caused by normalization. As depicted in Fig. 4a, the differentiable gating module outputs a series of binary indicators to combine the normalized and original channels, which can be expressed as

$$x_{n+1} = (\mathbf{1} - g) \odot x_n + g \odot x'_n. \tag{2}$$

where g represents the binary indicators across the channel dimension, and $\odot$ is the channel-wise multiplication. The gating operation activates or deactivates the channels by the binary indicators to normalize channels selectively. Consequently, the generated feature x_{n+1} eliminates the effects of lightness to obtain an invariant representation for generalization, and retains the essential information with unchanged channels for accurate reconstruction.

Specifically, the gating operation is expected to be differentiable and capable of biasing the output to zero or one for channel selection. Inspired by the pruning methods sampling the filters [62], we construct the gating module as

$$g = G(\alpha_x) = \frac{\alpha_x^2}{\alpha_x^2 + \epsilon}, \tag{3}$$

where $\alpha_x \in \mathbb{R}^C$ is an intermediate vector generated from the feature x, and ϵ is a small positive number. Specifically, to obtain α_x, we first employ adaptive pooling to shrink the spatial size of x to a single pixel, followed by several fully-connected layers with ReLU activations (Fig. 4b).

When $\alpha_x = 0$, it is obvious that $G(0) = 0$; when $\alpha_x \neq 0$, we can infer that $G(\alpha_x) \approx 1$ since ϵ is small enough. This function transforms α_x to a value close to one or zero, resulting in an on-off switch gate without requiring additional manual threshold design. Moreover, leveraging its differentiable character, we design an alternating optimization strategy (Sec. 4.3) to adaptively select the lightness-relevant channels.

It is worth noting that, the gating module can easily fall into a trivial solution that keeps all the channels unchanged to preserve the reconstruction accuracy, since the output of the function can easily be one (when $\alpha_x \neq 0$). Consequently, we make g to directly multiply with normalized channels x' in Eq. 2, pushing the network to prefer normalized channels to keep an elegant balance between the model's generalization ability and the discrimination of features for reconstruction.

4.3. Alternating Training Strategy

4.3.1 Training Strategy

We propose an alternating training strategy, as illustrated in Fig. 4c, to locate lightness-relevant channels in CSNorm. The rationale behind our strategy is that, by slightly perturbing the lightness condition of the input images, CSNorm is forced to locate and filter out lightness-relevant channels to achieve optimal performance on both original and perturbed images. Specifically, the strategy alternately optimizes the network on the original dataset to learn an essential ability for lightness adaptation, and optimizes CSNorm with slightly perturbed input images to identify lightness-relevant channels. This ensures that CSNorm can efficiently normalize lightness-relevant channels, exhibiting superior performance in both generalization and discrimination.

To optimize the network, we separate its parameters into two groups based on whether they belong to CSNorm, and use different loss functions to update them, as shown in Fig. 4c. In the first step, we input the original image and update parameters outside CSNorm by minimizing the loss function

$$\mathcal{L}_1 = |\hat{o}_1 - o_{gt}|_2, \tag{4}$$

where $\hat{o}_1$ is the output image of the network and o_{gt} is the ground-truth image. By doing so, the network's essential lightness-adaptation capability is maintained and all channels are preserved in their natural state.

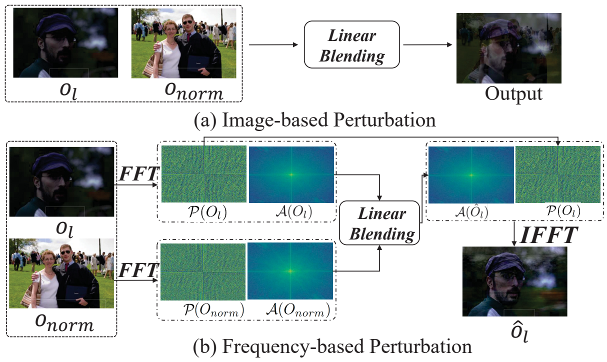


Figure 5: Comparisons of lightness perturbations. Our frequency-based perturbation captures the essential component of lightness, while avoiding interfering with other image components such as structural information.

In the second step, we perturb the lightness of the input image (Sec. 4.3.2) and fix the parameters outside CSNorm. In other words, we only update the parameters in CSNorm, by minimizing the loss function

$$\mathcal{L}_2 = |\hat{o}_2 - o_{gt}|_2 + |\mathcal{A}(\hat{o}_2) - \mathcal{A}(o_{gt})|_2, \tag{5}$$

where $\hat{o}_2$ is the output, and $\mathcal{A}$ reprensents the amplitude information in frequency domain. This enables CSNorm to adaptively select the lightness-relevant channels to keep the performance on perturbed images. In particular, since lightness is related to magnitude in the frequency domain [18], we add the amplitude loss $|\mathcal{A}(\hat{o}_2) - \mathcal{A}(o_{gt})|_2$ in Eq. 5 to allow the network to focus more on lightness information and effectively select lightness-relevant channels.

The two steps are alternately optimized by the above two objectives, and the overall optimization function is given by

$$\mathcal{L} = \mathcal{L}_{ori} + \delta \mathcal{L}_{amp}, \tag{6}$$

where δ is a balance factor.

4.3.2 Lightness Perturbation

As previously discussed, in order to automatically identify lightness-relevant channels during training, we need to perturb the lightness component of input images. These perturbations should capture the essence of lightness adaptation, while avoiding interfering with other image components such as structural information. To achieve this, we propose a frequency-based perturbation scheme that linearly interpolates the amplitudes of two images, since amplitude information contains more lightness information [18] that can prevent augmentation artifacts (Fig. 5).

Taking the low light enhancement task as an example, we define the low light and normal light images as o_l and

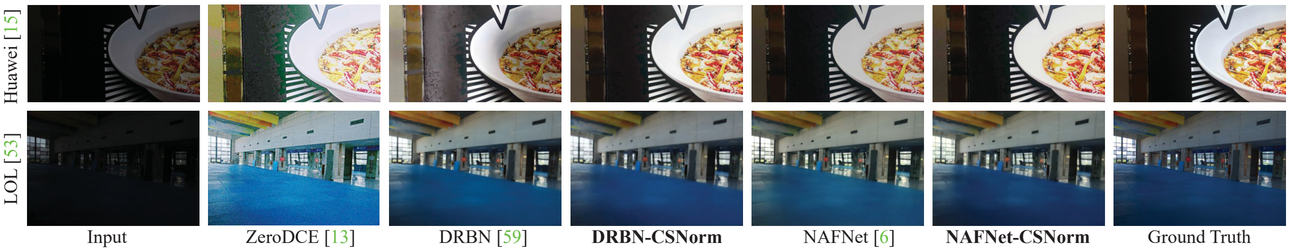


Figure 6: Visual results of generalized low-light image enhancement on Huawei [15] and LOL [53] datasets. The models are trained on one dataset (LOL [53] or Huawei [15] dataset) and tested on the other (Huawei [15] or LOL [53] dataset). Equipped with our CSNorm, the generalization abilities of base networks (DRBN [59] and NAFNet [6]) are significantly improved.

Table 1: Quantitative results of low-light image enhancement methods on synthetic lightnesses in terms of PSNR and SSIM.

Method	LOL [53]				Huawei [15]			
	original	interp	scale	average	original	interp	scale	average
CLANE [44]	12.77/0.5703	13.56/0.5852	13.39/0.5806	13.24/0.5787	13.29/0.4399	17.43/0.7220	14.51/0.4202	15.08/0.5274
LIME [14]	17.18/0.6130	17.49/0.9028	17.47/0.8188	17.38/0.7782	17.09/0.5932	15.03/0.7860	16.78/0.5778	16.30/0.6523
RetinexNet [53]	16.77/0.5393	15.02/0.7870	16.33/0.5043	16.04/0.6102	16.65/0.5010	13.12/0.5282	16.54/0.4693	15.44/0.5495
ZeroDCE [13]	15.47/0.6521	16.35/0.7432	15.96/0.6747	15.92/0.6900	12.53/0.4215	8.26/0.5539	11.33/0.4022	10.70/0.4592
LLFlow [51]	24.93/0.8922	18.32/0.9177	23.12/0.8901	22.12/0.9000	20.11/0.6645	12.97/0.7122	18.11/0.6381	17.06/0.6716
SID [5]	20.52/0.8382	15.12/0.7994	19.47/0.8305	18.37/0.8227	18.64/0.6251	12.15/0.6687	18.05/0.6011	16.28/0.6316
SID-CSNorm	**21.11/0.8312**	**16.63/0.8130**	**20.80/0.8341**	**19.52/0.8194**	**18.86/0.6128**	**12.77/0.6508**	**18.23/0.6114**	**16.62/0.6250**
DRBN [59]	20.75/0.8426	17.43/0.8745	20.15/0.8673	19.44/0.8614	18.80/0.6449	11.73/0.6797	17.58/0.6303	16.04/0.6516
DRBN-CSNorm	**21.05/0.8533**	**18.52/0.8822**	**20.64/0.8737**	**20.09/0.8697**	**18.81/0.6465**	**14.87/0.7285**	**18.74/0.6383**	**17.47/0.6711**
NAFNet [6]	23.02/0.8498	17.21/0.8733	20.22/0.8702	19.48/0.8846	19.07/0.6483	12.53/0.6822	17.75/0.6125	16.45/0.6476
NAFNet-CSNorm	**23.10/0.8544**	**18.75/0.8836**	**20.97/0.8767**	**20.27/0.6476**	**19.26/0.6492**	**15.02/0.7325**	**18.81/0.6410**	**17.69/0.6742**

o_{norm}, and their Fourier representations as O_l and O_{norm}. We linearly combine the amplitude components of O_l and O_{norm} as

$$\mathcal{A}(\hat{O}_l) = \lambda \mathcal{A}(O_l) + (1-\lambda)\mathcal{A}(O_{norm}), \quad (7)$$

where $\mathcal{A}$ represents amplitude information and $\lambda \in [0, 1]$ is randomly sampled. Then the perturbed image $\hat{o}$ is reconstructed through an inverse Fourier transformation $\mathcal{F}^{-1}$ as $\hat{o} = \mathcal{F}^{-1}(\mathcal{A}(\hat{O}_l), \mathcal{P}(O_l))$, where $\mathcal{P}$ is phase information.

As shown in Fig. 5, our frequency-based perturbation mitigates the influence of other factors in the image, such as structure and noise, and focuses more on the lightness itself. The perturbed and original images are used as inputs for different training steps to optimize CSNorm, enabling CSNorm to purposely select the lightness-relevant channels thereby enhancing the network's generalization ability.

5. Experiments

We do comprehensive evaluations on low-light image enhancement, inverse tone mapping, and image retouching to demonstrate the efficacy of our CSNorm.

5.1. Low-light Image Enhancement

Settings. We conducted experiments on the Huawei [15] and LOL [53] datasets. Representative methods such as CLANE [44], LIME [14], RetinexNet [53], LLFlow[51], and ZeroDCE [13] are used for comparison. We select SID [5], DRBN [59], and NAFNet [6] as base networks and integrate our CSNorm into them. The peak signal-to-noise ratio (PSNR) and structural similarity index measure (SSIM) are used as evaluation metrics.

Table 2: Quantitative results of low-light image enhancement methods across datasets in terms of PSNR and SSIM.

Train / Test	Huawei / LOL	LOL / Huawei
CLANE [44]	10.25/0.5602	11.11/0.4152
LIME [14]	15.25/0.7994	15.20/0.5321
RetinexNet [53]	15.35/0.5102	15.32/0.4855
ZeroDCE [13]	15.01/0.5974	12.25/0.4194
SID [5]	17.93/0.7159	16.10/0.5689
SID-CSNorm	**18.33/0.7725**	**17.31/0.6105**
DRBN [59]	18.10/0.8033	15.21/0.5477
DRBN-CSNorm	**18.65/0.8105**	**17.42/0.6122**
NAFNet [6]	19.05/0.7901	17.02/0.6002
NAFNet-CSNorm	**19.63/0.8322**	**17.53/0.6257**

We conduct experiments in synthetic and realistic settings. For the synthetic setting, we simulate two input lightness conditions, i.e., interp: the input image is interpolated by the original low-light and ground-truth images (weight is 0.5), and scale: the input image x is manipulated as $x' = \Lambda x^{\eta}$ ($\eta = 1.1, \Lambda = 1.2$). For the real-world setting, we use LOL [53] and Huawei [15] for cross-validation. The model is trained on one dataset and tested on the other.

Results. Table 1 compares the performance of our proposed method (name with **-CSNorm**) and baseline meth-

Figure 7: Visual results of the generalized inverse tone mapping on Kim *et al.* [24] and HDRTVNet [7] datasets. The models are trained on one dataset (Kim *et al.* [24] or HDRTVNet [7] dataset) and tested on the other (HDRTVNet [7] or Kim *et al.* [24] dataset). The colors of results seem light-colored since they are visualized in the standard Rec.2020 color space.

Table 3: Quantitative results of inverse tone mapping methods on synthetic lightnesses in terms of PSNR and SSIM.

Method	Kim *et al.* [24]				HDRTVNet [7]			
	original	interp	scale	average	original	interp	scale	average
CSRNet [16]	32.22/0.9472	26.74/0.9545	27.43/0.9282	28.79/0.9433	36.01/0.9717	29.22/0.9764	27.93/0.9417	31.05/0.9632
CSRNet-CSNorm	**32.53/0.9496**	**27.14/0.9562**	**27.83/0.9301**	**29.16/0.9453**	**36.15/0.9730**	**29.46/0.9766**	**28.15/0.9428**	**31.25/0.9641**
AGCM [7]	32.44/0.9482	26.87/0.9551	27.51/0.9297	28.94/0.9443	36.25/0.9733	29.36/0.9767	27.98/0.9418	31.19/0.9639
AGCM-CSNorm	**32.61/0.9501**	**27.26/0.9573**	**27.91/0.9322**	**29.26/0.9465**	**36.31/0.9734**	**29.57/0.9781**	**28.26/0.9433**	**31.38/0.9649**

ods on Huawei and LOL datasets. While previous methods achieve good results on original low-light conditions, they have a poor generalization ability on unknown lightness conditions. In contrast, our methods exhibit superior generalization ability, outperforming corresponding backbones by over 0.6dB on the two datasets. Our CSNorm also keeps the performance on original low-light images. Note that our objective is to improve the initial network rather than achieve state-of-the-art performance.

Table 2 shows generalization performance across different datasets. It can be seen that, previous methods tend to overfit the training dataset and have poor abilities for generalization. Our CSNorm improves the performance of all base networks, which greatly enhances their generalization abilities on unknown lightness conditions. We also show qualitative results in Fig. 6. Even though there is a large discrepancy between training and testing images (two datasets are separately captured with different lightnesses), the base networks equipped with CSNorm produce visually pleasing results on unknown lightness conditions.

5.2. Inverse Tone Mapping

Settings. We perform inverse tone mapping using CSRNet [16] and AGCM [7] as base networks, where we deepen the AGCM We use HDRTVNet [7] and Kim *et al.* [24] for training and testing. We conduct experiments on the original, the interpolated, and the scaled lightness conditions same as Sec. 5.1.

Results. Table 3 shows the quantitative results on the synthetic lightness conditions. CSRNet and AGCM have

Table 4: Quantitative results of inverse tone mapping methods using realistic datasets in terms of PSNR and SSIM.

Train / Test	Kim *et al.* / HDRTVNet	HDRTVNet / Kim *et al.*
CSRNet [16]	33.47/0.9642	31.06/0.9501
CSRNet-CSNorm	**34.02/0.9677**	**31.35/0.9523**
AGCM [7]	33.52/0.9651	31.22/0.9512
AGCM-CSNorm	**34.11/0.9687**	**31.67/0.9539**

good performances on original SDR frames but perform poorly when the lightness condition changes. In contrast, CSRNet-CSNorm and AGCM-CSNorm, which only add our CSNorm to base networks, achieve well performances on unknown lightness conditions. We also show real-world results in Table 4 and Fig. 7. The results demonstrate that our CSNorm has a strong generalization ability across different lightness conditions. It is worth noting that, CSNorm does not affect the performance of the original SDR frames, which confirms that the selected channels only affect the lightness-relevant information without altering the overall data distribution.

5.3. Image Retouching

Settings. We adopt MIT-Adobe FiveK [3] for training and testing. The experiments are conducted on lightness conditions as Sec. 5.1. We select CSRNet [16], DRBN [59], and NAFNet [6] as base networks to plug our CSNorm.

Results. We show the quantitative results in Table 5. It can be seen that, the base network (CSRNet) has a good performance on the original image but suffers from poor generalization ability. Its performance drops about 5dB when lightness changes from the original one to the interpolation

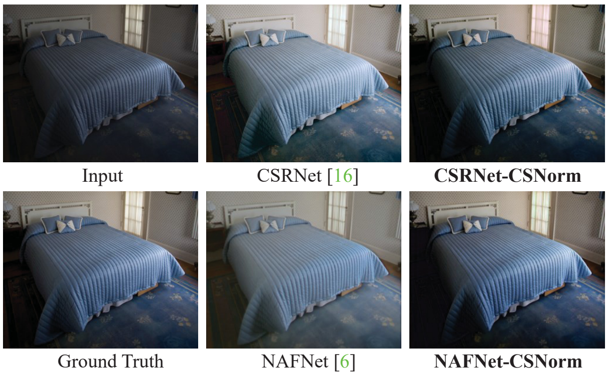


Figure 8: Visual comparisons of the generalized image retouching on the MIT-Adobe FiveK [3] dataset. The models are trained on the original dataset and tested on the scaled lightness condition.

Table 5: Quantitative results of image retouching methods on synthetic lightnesses in terms of PSNR and SSIM.

Method	original	interp	scale	average
DRBN [59]	22.11/0.8622	23.54/0.8757	20.51/0.8544	22.05/0.8641
DRBN-CSNorm	**22.43/0.8679**	**23.97/0.8792**	**20.68/0.8561**	**22.36/0.8677**
CSRNet [16]	**23.52**/0.8865	24.69/0.8943	22.76/0.8651	23.65/0.8823
CSRNet-CSNorm	23.41/**0.8838**	**25.34/0.9026**	**22.85/0.8676**	**23.86/0.8846**
NAFNet [6]	**23.71/0.8901**	24.77/0.8966	22.85/0.8685	23.82/0.8850
NAFNet-CSNorm	23.54/0.8922	**25.40/0.9033**	**22.97/0.8690**	**24.03/0.8881**

Table 6: Quantitative results of normalization methods, where the models are trained on the LOL [53] dataset and tested on the Huawei [15] dataset.

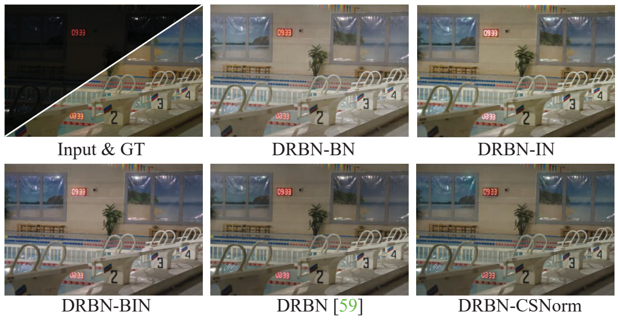
Method	DRBN-BN	DRBN-IN	DRBN-BIN	DRBN-CSNorm
PSNR/SSIM	16.05/0.5537	16.98/0.5821	17.11/0.5817	**17.42/0.6122**

one. Based on CSRNet network, our method (CSRNet-CSNorm) remarkably improves the original model's average capability by over 1 dB, which demonstrates CSNorm's powerful generalization ability. We also show the visual results in Fig. 8.

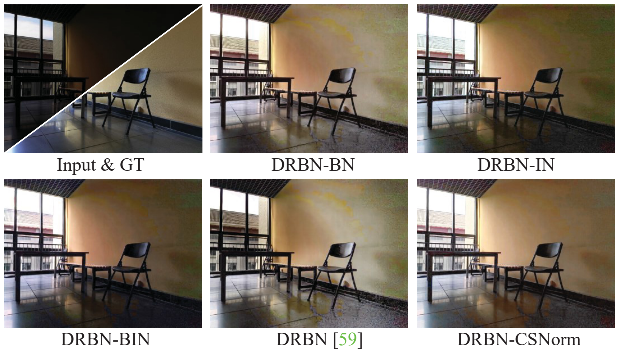
6. Analysis

Feature visualization. We visualize the selected channels to demonstrate our CSNorm can effectively enhance the model's ability to generalize to different lightnesses. We show the selected channels (*i.e.,* lightness-relevant channels) in Fig. 10. It can be seen that, the channels extracted from different lightness conditions have different characteristics, which may lead to poor generalization ability. Our CSNorm selects this channel and normalizes it, which substantially produces lightness-independent information for generalization.

Comparisons to other normalizations. We compare our CSNorm with conventional normalization techniques, including Batch Normalization (BN) [21], Instance Normalization (IN) [50], and Batch Instance Normalization



(a) The models are trained and tested on the LOL [53] dataset.

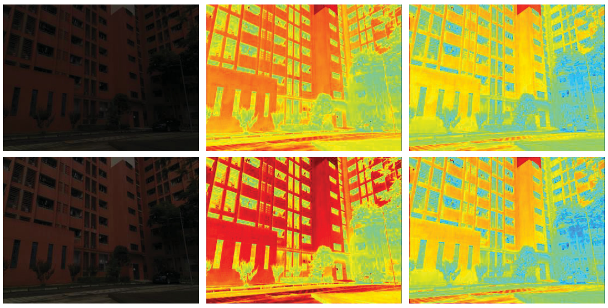


(b) The models are trained on the LOL [53] dataset and tested on the Huawei [15] dataset.

Figure 9: Ablation study on normalization techniques. (a) CSNorm keeps the base network's ability. (b) CSNorm can generalize to other unknown lightness conditions.

(BIN) [38], by plugging them into DRBN [59]. Since the alternating strategy is specially designed for our CSNorm and may be harmful to conventional normalization techniques, we train BN, IN, and BIN only with the data perturbation. We show the quantitative and qualitative results in Table 6 and Fig. 9, respectively. Compared with conventional normalization techniques, our CSNorm effectively keeps the performance on the known lightness condition and has a well generalization ability to unknown lightness conditions, which avoids unsatisfied artifacts.

Training strategy and data perturbation. We take ablation studies on training strategy and data perturbation. For the training strategy, we replace the alternating training strategy with the mixed training strategy, where the network is trained by mixing original and perturbed data. For the data perturbation, we replace the frequency-based perturbation with linearly blending the input images and the ground truth images. These aforementioned ablation experiments are conducted on the LOL [53] dataset and tested on the Huawei [15] dataset. As shown in Table 7, the alternating training strategy gets a better performance compared with mixed training, while our frequency-based data perturbation improves the model's generalization, demonstrating



Input Before Normalization After Normalization

Figure 10: Visualization of lightness-relevant channels. The channels are adaptively selected by our CSNorm.

Table 7: Ablation study on training strategy and data perturbation.

Training strategy	Mixed		Alternating	
Data perturbation	Linear	Amp	Linear	Amp (ours)
DRBN [59]	14.35/0.5526	14.10/0.5382	-	-
DRBN-CSNorm	16.79/0.5950	16.81/0.5972	17.15/0.6097	**17.42/0.6122**

Table 8: Experiment results of exposure correction in terms of PSNR (dB) and SSIM.

Method	ME [1]	Exp0	Exp6
DRBN [59]	19.65/0.8292	18.10/0.7843	13.43/0.7339
DRBN w/o ATS	21.31/0.8345	18.59/0.7321	15.47/0.7917
DRBN w/ ATS	**21.63/0.8396**	**18.98/0.7673**	**16.31/0.7928**

Table 9: Experiments on lightness-only datasets.

Train / Test	LOL / LOL	LOL / Huawei	Huawei / Huawei	Huawei / LOL
DRBN [59]	20.98/0.8922	15.13/0.5778	18.94/0.6517	18.76/0.8190
DRBN-CSNorm	**24.33/0.8998**	**17.40/0.5883**	**19.10/0.6505**	**19.15/0.8255**

the effectiveness of our design. Simply training CSNorm with a diverse range of lighting conditions images cannot effectively identify the lightness-relevant channels due to the influence of content, which is detrimental to the lightness generalization. We conduct experiments on the ME dataset [1] (Table 8). It can be seen that, without the alternating training strategy (ATS), the network cannot generalize well to unknown lightness conditions (Exp0 and Exp6 are unknown lower and higher lightness).

Parameters. Our proposed CSNorm is lightweight that can be plugged into existing networks with nearly no parameter increase, which avoids the heavy storage cost. For example, given a feature with 64 channels, our CSNorm just requires 16.5k parameters to identify and normalize lightness-relevant channels, owing to the gating module and the affine transformation in normalization. The CSNorm's number of parameters grows linearly with the number of channels.

Evaluation on known lightness-only datasets. To further demonstrate the effectiveness of our CSNorm, we conduct experiments on lightness-only datasets. We transform the color image into Ycbcr color space and use the Y channel since it represents the luminance or lightness information. Experiment results in Table 9 show that our method can effectively enhance generalization ability without sacrificing the discrimination of the features.

Amplitude-related information. Amplitude-related information has been proven related to the lightness components in previous works [18]. We implicitly utilize the amplitude-related information as a detailed lightness perturbation manner in the alternating training strategy, enabling CSNorm to identify the lightness-related channels. Thus, this amplitude-based lightness perturbation is orthogonal to CSNorm and its format is not introduced into CSNorm. Note that other lightness perturbation manner can also drive CSNorm's training, *e.g.*, linear interpolation in Table 7, while our adopted amplitude-based lightness perturbation experimentally achieves higher performance.

7. Conclusion and Discussion

In this work, we propose CSNorm, a novel normalization technique customized for generalized lightness adaptation. It purposely normalizes lightness-relevant channels while keeping other channels unchanged, which empowers existing lightness adaptation methods with the generalization ability to wide-range lightness scenes. Except for the sufficient generalization ability on the unknown lightnesses, CSNorm keeps the reconstruction accuracy on the known lightness. The proposed CSNorm is architecture-agnostic which we validate, making it simple, lightweight, and plug-and-play. Extensive experiments on multiple tasks and benchmark datasets verify the effectiveness of our proposed CSNorm to enhance the generalization of existing lightness adaptation methods. We believe our method would inspire more valuable generalization methods for lightness adaptation, and holds potential for application in other tasks.

Despite the promising preliminary results, there are still some real-world conditions that have not been considered in this paper. For images with no discernible details or high-level noise levels, it is indeed challenging for our CSNorm due to the loss of necessary information. However, our lightweight and plug-and-play design may enable it to handily collaborate with other methods, either by inserting or by following inpainting or denoising networks. Moreover, we believe there is great potential to explore more robust data perturbation methods, e.g., Retinex model-based lightness decomposition, to further uncover the capabilities of our method.

Acknowledgments

This work was supported in part by the National Natural Science Foundation of China under Grants 62131003 and 62021001.

References

[1] Mahmoud Afifi, Konstantinos G Derpanis, Bjorn Ommer, and Michael S Brown. Learning multi-scale photo exposure correction. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 9157–9167, 2021. 1, 2, 9

[2] Jimmy Lei Ba, Jamie Ryan Kiros, and Geoffrey E Hinton. Layer normalization. *arXiv preprint arXiv:1607.06450*, 2016. 3

[3] 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 Conference on Computer Vision and Pattern Recognition*, pages 97–104, 2011. 7, 8

[4] Prithvijit Chattopadhyay, Yogesh Balaji, and Judy Hoffman. Learning to balance specificity and invariance for in and out of domain generalization. In *European Conference on Computer Vision*, pages 301–318. Springer, 2020. 3

[5] Chen Chen, Qifeng Chen, Jia Xu, and Vladlen Koltun. Learning to see in the dark. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 3291–3300, 2018. 6

[6] Liangyu Chen, Xiaojie Chu, Xiangyu Zhang, and Jian Sun. Simple baselines for image restoration. In *European Conference on Computer Vision*, pages 17–33. Springer, 2022. 6, 7, 8

[7] Xiangyu Chen, Zhengwen Zhang, Jimmy S Ren, Lynhoo Tian, Yu Qiao, and Chao Dong. A new journey from sdrtv to hdrtv. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 4500–4509, 2021. 1, 2, 7

[8] Qi Dou, Daniel Coelho de Castro, Konstantinos Kamnitsas, and Ben Glocker. Domain generalization via model-agnostic learning of semantic features. *Advances in Neural Information Processing Systems*, 32, 2019. 3

[9] Yingjun Du, Jun Xu, Huan Xiong, Qiang Qiu, Xiantong Zhen, Cees GM Snoek, and Ling Shao. Learning to learn with variational information bottleneck for domain generalization. In *European Conference on Computer Vision*, pages 200–216. Springer, 2020. 3

[10] Xinjie Fan, Qifei Wang, Junjie Ke, Feng Yang, Boqing Gong, and Mingyuan Zhou. Adversarially adaptive normalization for single domain generalization. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 8208–8217, 2021. 3

[11] Hao Feng, Jiajun Deng, Wendi Wang, Wengang Zhou, Li Li, and Houqiang Li. Simfir: A simple framework for fisheye image rectification with self-supervised representation learning. In *Proceedings of the International Conference on Computer Vision*, 2023. 3

[12] Xueyang Fu, Menglu Wang, Xiangyong Cao, Xinghao Ding, and Zheng-Jun Zha. A model-driven deep unfolding method for jpeg artifacts removal. *IEEE Transactions on Neural Networks and Learning Systems*, 33(11):6802–6816, 2021. 2

[13] 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/CVF Conference on Computer Vision and Pattern Recognition*, pages 1780–1789, 2020. 6

[14] 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. 6

[15] Jiang Hai, Zhu Xuan, Ren Yang, Yutong Hao, Fengzhu Zou, Fang Lin, and Songchen Han. R2rnet: Low-light image enhancement via real-low to real-normal network. *arXiv preprint arXiv:2106.14501*, 2021. 6, 8

[16] Jingwen He, Yihao Liu, Yu Qiao, and Chao Dong. Conditional sequential modulation for efficient global image retouching. In *European Conference on Computer Vision*, pages 679–695. Springer, 2020. 2, 7, 8

[17] 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. 2, 3

[18] 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 *European Conference on Computer Vision*, pages 163–180. Springer, 2022. 5, 9

[19] 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*, pages 6309–6317, 2022. 2

[20] Zeyi Huang, Haohan Wang, Eric P Xing, and Dong Huang. Self-challenging improves cross-domain generalization. In *European Conference on Computer Vision*, pages 124–140. Springer, 2020. 3

[21] Sergey Ioffe and Christian Szegedy. Batch normalization: Accelerating deep network training by reducing internal covariate shift. In *International conference on machine learning*, pages 448–456. PMLR, 2015. 3, 8

[22] Xin Jin, Cuiling Lan, Wenjun Zeng, Zhibo Chen, and Li Zhang. Style normalization and restitution for generalizable person re-identification. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 3143–3152, 2020. 2, 4

[23] Yeying Jin, Wenhan Yang, and Robby T Tan. Unsupervised night image enhancement: When layer decomposition meets light-effects suppression. In *European Conference on Computer Vision*, pages 404–421. Springer, 2022. 2

[24] Soo Ye Kim, Jihyong Oh, and Munchurl Kim. Deep sr-itm: Joint learning of super-resolution and inverse tone-mapping for 4k uhd hdr applications. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 3116–3125, 2019. 1, 2, 7

[25] Boyun Li, Yuanbiao Gou, Shuhang Gu, Jerry Zitao Liu, Joey Tianyi Zhou, and Xi Peng. You only look yourself: Unsupervised and untrained single image dehazing neural network. *International Journal of Computer Vision*, 129:1754–1767, 2021. 3

[26] Da Li, Yongxin Yang, Yi-Zhe Song, and Timothy M Hospedales. Learning to generalize: Meta-learning for do-

main generalization. In *Thirty-Second AAAI Conference on Artificial Intelligence*, 2018. 3

[27] 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

[28] Rui Li and Dong Liu. Spatial-then-temporal self-supervised learning for video correspondence. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 2279–2288, 2023. 3

[29] Jie Liang, Hui Zeng, Miaomiao Cui, Xuansong Xie, and Lei Zhang. PPR10K: A large-scale portrait photo retouching dataset with human-region mask and group-level consistency. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 653–661, 2021. 1, 2

[30] Yijie Lin, Yuanbiao Gou, Zitao Liu, Boyun Li, Jiancheng Lv, and Xi Peng. Completer: Incomplete multi-view clustering via contrastive prediction. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 11174–11183, 2021. 3

[31] 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/CVF Conference on Computer Vision and Pattern Recognition*, pages 10561–10570, 2021. 1, 2

[32] Kin Gwn Lore, Adedotun Akintayo, and Soumik Sarkar. Llnet: A deep autoencoder approach to natural low-light image enhancement. *Pattern Recognition*, 61:650–662, 2017. 1, 2

[33] Xin Luo, Yunan Zhu, Shunxin Xu, and Dong Liu. On the effectiveness of spectral discriminators for perceptual quality improvement. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2023. 2

[34] 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*, pages 5637–5646, 2022. 2

[35] Rafał Mantiuk, Scott Daly, and Louis Kerofsky. Display adaptive tone mapping. In *ACM SIGGRAPH 2008 papers*. ACM, 2020. 2

[36] Krikamol Muandet, David Balduzzi, and Bernhard Schölkopf. Domain generalization via invariant feature representation. In *International Conference on Machine Learning*, pages 10–18. PMLR, 2013. 3

[37] Zachary Nado, Shreyas Padhy, D Sculley, Alexander D'Amour, Balaji Lakshminarayanan, and Jasper Snoek. Evaluating prediction-time batch normalization for robustness under covariate shift. *arXiv preprint arXiv:2006.10963*, 2020. 3

[38] Hyeonseob Nam and Hyo-Eun Kim. Batch-instance normalization for adaptively style-invariant neural networks. *Advances in Neural Information Processing Systems*, 31, 2018. 3, 8

[39] Zhangkai Ni, Wenhan Yang, Shiqi Wang, Lin Ma, and Sam Kwong. Towards unsupervised deep image enhancement with generative adversarial network. *IEEE Transactions on Image Processing*, 29:9140–9151, 2020. 1

[40] Xingang Pan, Ping Luo, Jianping Shi, and Xiaoou Tang. Two at once: Enhancing learning and generalization capacities via ibn-net. In *Proceedings of the European Conference on Computer Vision*, pages 464–479, 2018. 2, 3

[41] 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. 2

[42] Fengchun Qiao, Long Zhao, and Xi Peng. Learning to learn single domain generalization. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 12556–12565, 2020. 3

[43] 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

[44] Ali M Reza. Realization of the contrast limited adaptive histogram equalization (clahe) for real-time image enhancement. *Journal of VLSI signal processing systems for signal, image and video technology*, 38(1):35–44, 2004. 6

[45] Mattia Segu, Alessio Tonioni, and Federico Tombari. Batch normalization embeddings for deep domain generalization. *arXiv preprint arXiv:2011.12672*, 2020. 3

[46] Seonguk Seo, Yumin Suh, Dongwan Kim, Geeho Kim, Jongwoo Han, and Bohyung Han. Learning to optimize domain specific normalization for domain generalization. In *European Conference on Computer Vision*, pages 68–83, 2020. 1, 3

[47] Shiv Shankar, Vihari Piratla, Soumen Chakrabarti, Siddhartha Chaudhuri, Preethi Jyothi, and Sunita Sarawagi. Generalizing across domains via cross-gradient training. *arXiv preprint arXiv:1804.10745*, 2018. 1, 3

[48] Ivan Skorokhodov and Mohamed Elhoseiny. Class normalization for (continual)? generalized zero-shot learning. *arXiv preprint arXiv:2006.11328*, 2020. 3

[49] Yuda Song, Hui Qian, and Xin Du. Starenhancer: Learning real-time and style-aware image enhancement. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 4126–4135, 2021. 1

[50] Dmitry Ulyanov, Andrea Vedaldi, and Victor Lempitsky. Instance normalization: The missing ingredient for fast stylization. *arXiv preprint arXiv:1607.08022*, 2016. 3, 8

[51] Yufei Wang, Renjie Wan, Wenhan Yang, Haoliang Li, Lap-Pui Chau, and Alex Kot. Low-light image enhancement with normalizing flow. In *Proceedings of the AAAI conference on artificial intelligence*, 2022. 6

[52] Yongzhen Wang, Xuefeng Yan, Fu Lee Wang, Haoran Xie, Wenhan Yang, Mingqiang Wei, and Jing Qin. Ucl-dehaze: Towards real-world image dehazing via unsupervised contrastive learning. *arXiv preprint arXiv:2205.01871*, 2022. 1

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

[54] Yuxin Wu and Kaiming He. Group normalization. In *Proceedings of the European conference on computer vision*, pages 3–19, 2018. 3

[55] Jie Xiao, Man Zhou, Xueyang Fu, Aiping Liu, and Zheng-Jun Zha. Improving de-raining generalization via neural reorganization. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 4987–4996, 2021. 3

[56] Ruikang Xu, Mingde Yao, and Zhiwei Xiong. Zero-shot dual-lens super-resolution. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 9130–9139, 2023. 1, 3

[57] Gang Yang, Man Zhou, Keyu Yan, Aiping Liu, Xueyang Fu, and Fan Wang. Memory-augmented deep conditional unfolding network for pan-sharpening. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 1788–1797, 2022. 2

[58] Kai-Fu Yang, Cheng Cheng, Shi-Xuan Zhao, Hong-Mei Yan, Xian-Shi Zhang, and Yong-Jie Li. Learning to adapt to light. *International Journal of Computer Vision*, 131(4):1022–1041, 2023. 2

[59] 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, 7, 8, 9

[60] Mingde Yao, Dongliang He, Xin Li, Fu Li, and Zhiwei Xiong. Towards interactive self-supervised denoising. *IEEE Transactions on Circuits and Systems for Video Technology*, 2023. 3

[61] 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. 2

[62] Zhonghui You, Kun Yan, Jinmian Ye, Meng Ma, and Ping Wang. Gate decorator: Global filter pruning method for accelerating deep convolutional neural networks. *Advances in neural information processing systems*, 32, 2019. 5

[63] Huimin Zeng, Jie Huang, Jiacheng Li, and Zhiwei Xiong. Region-aware portrait retouching with sparse interactive guidance. *IEEE Transactions on Multimedia*, 2023. 2

[64] Jinghao Zhang, Jie Huang, Mingde Yao, Man Zhou, and Feng Zhao. Structure-and texture-aware learning for low-light image enhancement. In *Proceedings of the 30th ACM International Conference on Multimedia*, pages 6483–6492, 2022. 2

[65] Yonghua Zhang, Jiawan Zhang, and Xiaojie Guo. Kindling the darkness: A practical low-light image enhancer. In *Proceedings of the 27th ACM international conference on multimedia*, pages 1632–1640, 2019. 1, 2

[66] Zixiang Zhao, Haowen Bai, Yuanzhi Zhu, Jiangshe Zhang, Shuang Xu, Yulun Zhang, Kai Zhang, Deyu Meng, Radu Timofte, and Luc Van Gool. Ddfm: denoising diffusion model for multi-modality image fusion. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2023. 2

[67] Kaiyang Zhou, Yongxin Yang, Timothy Hospedales, and Tao Xiang. Learning to generate novel domains for domain generalization. In *European Conference on Computer Vision*, pages 561–578. Springer, 2020. 1, 3

[68] Man Zhou, Jie Huang, Keyu Yan, Gang Yang, Aiping Liu, Chongyi Li, and Feng Zhao. Normalization-based feature selection and restitution for pan-sharpening. In *Proceedings of the 30th ACM International Conference on Multimedia*, pages 3365–3374, 2022. 2, 3