Generalized Lightness Adaptation with Channel Selective Normalization

Mingde Yao	extsuperscript{1}, Jie Huang	extsuperscript{1}, Xin Jin	extsuperscript{2}, Ruikang Xu	extsuperscript{1}, Shenglong Zhou	extsuperscript{1}, Man Zhou	extsuperscript{3}, Zhiwei Xiong	extsuperscript{1†}

	extsuperscript{1}University of Science and Technology of China
	extsuperscript{2}Eastern Institute of Technology
	extsuperscript{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.

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-
deteriorate the reconstruction accuracy \([40, 68]\). Therefore, sword due to its inevitable loss of information, which might components \([40, 17]\). However, normalization is a from the given features \([22]\), especially for lightness com-

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 ad-
aptation. 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 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 stepwise 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
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] calibrates the model under covariate shift at the test stage. ASR-Norm [10] adapts 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
Given a feature \( x \), we normalize the channels and adopt IN as the implementation 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

\[
x' = \text{IN}(x) = \frac{x - \mu(x)}{\sigma(x)} + \beta,
\]

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. 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

\[
x' = \text{IN}(x) = \frac{x - \mu(x)}{\sigma(x)} + \beta,
\]

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

\[
x_{n+1} = (1 - g) \odot x_n + g \odot x'_n,
\]

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 \( \epsilon \) is a small positive number. Specifically, to obtain \( \alpha_x \), 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 \( \epsilon \) is small enough. This function transforms \( \alpha_x \) 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 lightness-adaptation capability and all channels are preserved in their natural state.

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

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

where \( \hat{o}_2 \) is the output, and \( A \) represents 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 \( |A(\hat{o}_2) - A(o_{gt})|_2 \) 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 \( \delta \) 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
\[ a_{\text{norm}}, \quad \text{and their Fourier representations as } O_l \text{ and } O_{\text{norm}}. \]

We linearly combine the amplitude components of \( O_l \) and \( O_{\text{norm}} \) as

\[
A(\hat{O}_l) = \lambda A(O_l) + (1 - \lambda) A(O_{\text{norm}}),
\]

where \( \lambda \) 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

\[
\hat{O} = \mathcal{F}^{-1}(A(\hat{O}_l), P(O_l)), \quad \text{where } P \text{ is phase information.}
\]

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

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

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>original</td>
<td>interp</td>
</tr>
<tr>
<td>CLANE [44]</td>
<td>12.77/0.5703</td>
<td>13.56/0.5852</td>
</tr>
<tr>
<td>LIME [14]</td>
<td>17.18/0.6130</td>
<td>17.49/0.6188</td>
</tr>
<tr>
<td>RetinexNet [53]</td>
<td>16.77/0.5393</td>
<td>15.02/0.6043</td>
</tr>
<tr>
<td>ZeroDCE [13]</td>
<td>15.47/0.6521</td>
<td>16.35/0.6747</td>
</tr>
<tr>
<td>LLFlow [51]</td>
<td>24.93/0.8922</td>
<td>18.32/0.9177</td>
</tr>
<tr>
<td>SID [5]</td>
<td>20.52/0.8382</td>
<td>15.12/0.7999</td>
</tr>
<tr>
<td>SID-CSNorm</td>
<td>21.11/0.8312</td>
<td>16.63/0.8130</td>
</tr>
<tr>
<td>DRBN [59]</td>
<td>20.75/0.8426</td>
<td>17.43/0.8745</td>
</tr>
<tr>
<td>DRBN-CSNorm</td>
<td>21.05/0.8533</td>
<td>18.52/0.8822</td>
</tr>
<tr>
<td>RetinexNet [53]</td>
<td>15.35/0.5102</td>
<td>15.32/0.4855</td>
</tr>
<tr>
<td>ZeroDCE [13]</td>
<td>15.21/0.6822</td>
<td>17.55/0.7122</td>
</tr>
<tr>
<td>NAFNet [6]</td>
<td>23.02/0.8498</td>
<td>17.21/0.8733</td>
</tr>
<tr>
<td>NAFNet-CSNorm</td>
<td>23.10/0.8544</td>
<td>18.75/0.8836</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Method</th>
<th>Huawei / LOL</th>
<th>LOL / Huawei</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>original</td>
<td>interp</td>
</tr>
<tr>
<td>CLANE [44]</td>
<td>10.25/0.5602</td>
<td>11.11/0.4113</td>
</tr>
<tr>
<td>LIME [14]</td>
<td>15.25/0.7994</td>
<td>15.20/0.5321</td>
</tr>
<tr>
<td>RetinexNet [53]</td>
<td>15.35/0.5102</td>
<td>15.32/0.4855</td>
</tr>
<tr>
<td>ZeroDCE [13]</td>
<td>15.01/0.5074</td>
<td>12.25/0.4194</td>
</tr>
<tr>
<td>SID [5]</td>
<td>17.93/0.7159</td>
<td>16.10/0.5689</td>
</tr>
<tr>
<td>SID-CSNorm</td>
<td>18.33/0.7725</td>
<td>17.31/0.6105</td>
</tr>
<tr>
<td>DRBN [59]</td>
<td>18.73/0.8103</td>
<td>15.21/0.5477</td>
</tr>
<tr>
<td>DRBN-CSNorm</td>
<td>18.65/0.8105</td>
<td>17.42/0.6122</td>
</tr>
<tr>
<td>NAFNet [6]</td>
<td>19.05/0.7901</td>
<td>17.02/0.6002</td>
</tr>
<tr>
<td>NAFNet-CSNorm</td>
<td>19.63/0.8322</td>
<td>17.53/0.6257</td>
</tr>
</tbody>
</table>
Table 3: Quantitative results of inverse tone mapping methods on synthetic lightnesses in terms of PSNR and SSIM.

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

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

<table>
<thead>
<tr>
<th>Train / Test</th>
<th>Kim et al. / HDRTVNet</th>
<th>HDRTVNet / Kim et al.</th>
</tr>
</thead>
<tbody>
<tr>
<td>CSRNet [16]</td>
<td>33.47/0.9642</td>
<td>31.06/0.9501</td>
</tr>
<tr>
<td>CSRNet-CSNorm</td>
<td>34.02/0.9677</td>
<td>31.35/0.9523</td>
</tr>
<tr>
<td>AGCM [7]</td>
<td>33.52/0.9651</td>
<td>31.22/0.9512</td>
</tr>
<tr>
<td>AGCM-CSNorm</td>
<td>34.11/0.9687</td>
<td>31.67/0.9539</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Method</th>
<th>original interp</th>
<th>scale</th>
<th>average</th>
</tr>
</thead>
<tbody>
<tr>
<td>DRBN [59]</td>
<td>22.11/0.8679</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CSRNet [16]</td>
<td>23.52/0.8885</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NAFNet [6]</td>
<td>23.71/0.8901</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CSRNet-CSNorm</td>
<td>23.54/0.8922</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NAFNet-CSNorm</td>
<td>23.14/0.8838</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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 (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...
Before Normalization After Normalization

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

<table>
<thead>
<tr>
<th>Training strategy</th>
<th>Data perturbation</th>
<th>Mixed</th>
<th>Alternating</th>
</tr>
</thead>
<tbody>
<tr>
<td>DRBN [59]</td>
<td>Linear</td>
<td>14.35/0.5526</td>
<td>14.10/0.5382</td>
</tr>
<tr>
<td>DRBN-CSNorm</td>
<td>Linear</td>
<td>16.79/0.5950</td>
<td>16.81/0.5972</td>
</tr>
<tr>
<td></td>
<td>Amp</td>
<td>17.15/0.6097</td>
<td>17.42/0.6122</td>
</tr>
<tr>
<td></td>
<td>Linear Amp</td>
<td>17.15/0.6097</td>
<td>17.42/0.6122</td>
</tr>
<tr>
<td></td>
<td>Linear Amp (ours)</td>
<td>17.15/0.6097</td>
<td>17.42/0.6122</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Method</th>
<th>ME [1]</th>
<th>Exp0</th>
<th>Exp6</th>
</tr>
</thead>
<tbody>
<tr>
<td>DRBN [59]</td>
<td>19.65/0.8292</td>
<td>18.10/0.7843</td>
<td>13.43/0.7339</td>
</tr>
<tr>
<td>DRBN w/o ATS</td>
<td>21.31/0.8345</td>
<td>18.90/0.7321</td>
<td>15.47/0.7917</td>
</tr>
<tr>
<td>DRBN w/ ATS</td>
<td>21.63/0.8396</td>
<td>18.98/0.7673</td>
<td>16.31/0.7928</td>
</tr>
</tbody>
</table>

Table 9: Experiments on lightness-only datasets.

<table>
<thead>
<tr>
<th>Train / Test</th>
<th>LOL / LOL</th>
<th>LOL / Huawei</th>
<th>Huawei / Huawei</th>
<th>Huawei / LOL</th>
</tr>
</thead>
<tbody>
<tr>
<td>DRBN [59]</td>
<td>20.98/0.8922</td>
<td>15.13/0.5778</td>
<td>18.94/0.6517</td>
<td>18.76/0.8190</td>
</tr>
<tr>
<td>DRBN-CSNorm</td>
<td>24.33/0.8998</td>
<td>17.46/0.583</td>
<td>19.16/0.6585</td>
<td>19.15/0.6255</td>
</tr>
</tbody>
</table>

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


[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


[41] Zhihong Pan, Baoqi Li, Dongliang He, Mingde Yao, Wen-hao 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


[46] Seonguk Seo, Yumin Suh, Dongwan Kim, Geelho 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


10679