

# Learning a Simple Low-light Image Enhancer from Paired Low-light Instances

Zhenqi Fu<sup>1</sup>, Yan Yang<sup>2</sup>, Xiaotong Tu<sup>1\*</sup>, Yue Huang<sup>1</sup>, Xinghao Ding<sup>1</sup>, Kai-Kuang Ma<sup>3</sup>

<sup>1</sup>Key Laboratory of Multimedia Trusted Perception and Efficient Computing,

Ministry of Education of China, School of Informatics, Xiamen University, China

<sup>2</sup>School of Computer Science and Technology, Hangzhou Dianzi University, China

<sup>3</sup> School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore

fuzhenqi@stu.xmu.edu.cn, yangyanyy@hdu.edu.cn,

{xttu, yhuang2010, dxh}@xmu.edu.cn, ekkma@ntu.edu.sg

## Abstract

*Low-light Image Enhancement (LIE) aims at improving contrast and restoring details for images captured in low-light conditions. Most of the previous LIE algorithms adjust illumination using a single input image with several handcrafted priors. Those solutions, however, often fail in revealing image details due to the limited information in a single image and the poor adaptability of handcrafted priors. To this end, we propose PairLIE, an unsupervised approach that learns adaptive priors from low-light image pairs. First, the network is expected to generate the same clean images as the two inputs share the same image content. To achieve this, we impose the network with the Retinex theory and make the two reflectance components consistent. Second, to assist the Retinex decomposition, we propose to remove inappropriate features in the raw image with a simple self-supervised mechanism. Extensive experiments on public datasets show that the proposed PairLIE achieves comparable performance against the state-of-the-art approaches with a simpler network and fewer handcrafted priors. Code is available at: <https://github.com/zhenqifu/PairLIE>.*

## 1. Introduction

Images captured under low-light environments always suffer from multiple distortions, such as low contrast, poor visibility, and sensor noise. Those low-light images are unsatisfactory for information transmission because they incur challenges in human visualization and subsequent computer vision tasks [25]. To correct contrast, uncover textures, and remove sensor noise, great efforts have been made in developing Low-light Image Enhancement (LIE) algorithms in the past decades [1, 5, 6, 8, 28, 35].

\*Corresponding author.

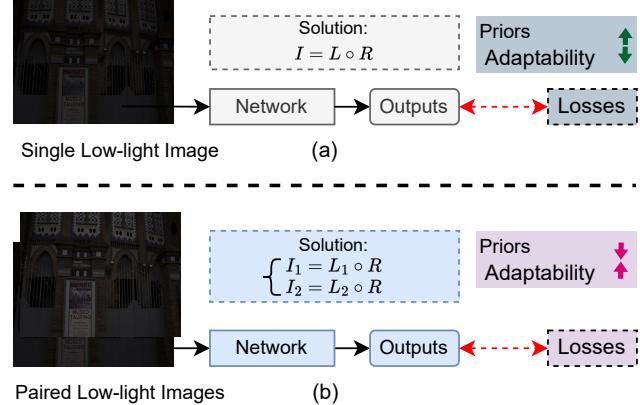


Figure 1. Comparison between the previous solution (a) and the proposed method (b) from the aspect of Retinex theory. The key idea of our method is to learn adaptive priors from low-light image pairs. As a result, our solution needs fewer handcrafted priors and the network is more robust. Note that image pairs are only used in the training phase.

Histogram-based and Retinex-based approaches are two well-known LIE techniques. The former enhances the contrast of an image by redistributing the luminous intensity on the histogram [3, 14]. The latter decomposes an observed image  $I$  into illumination  $L$  and reflectance  $R$  via  $I = L \circ R$ , where  $\circ$  denotes the element-wise multiplication [6, 13, 17]. Specifically, the reflectance component  $R$  is assumed to be consistent under different light conditions because  $R$  represents the physical properties of the objects. As the Retinex theory can well model color perceptions of human vision, Retinex-based methods have attracted relatively more attention in the LIE community.

Recent years have witnessed great success in developing learning-based LIE algorithms. Among these approaches, most solutions rely on low-light and normal-light image

pairs [33, 38]. However, collecting high-quality reference maps in real-world scenarios is time-consuming and expensive [32]. To eliminate the requirement for normal-light images, unsupervised and zero-shot LIE approaches are proposed. Concretely, the former trains a deep neural network using a set of collected low-light samples [7, 18], while the latter only employs the test image itself in the network optimization [40, 41]. Due to the absence of reference images, unsupervised and zero-shot LIE approaches depend on handcrafted priors to guide network training. Nevertheless, due to the complex natural scenes and the limited information in a single low-light image, it is difficult for those methods to attain a high-quality result.

To tackle the issues of limited information in a single low-light image and the poor adaptability of handcrafted priors, we propose to leverage paired low-light instances to train the LIE network. The main difference between our solution and previous approaches is illustrated in Fig. 1. Note that acquiring paired low-light images will complicate the imaging process since it needs to cope with the misalignment between the two images. Nevertheless, compared with collecting low-light and normal-light image pairs, our solution is more practical. Additionally, twice-exposure images provide useful information for solving the LIE task. As a result, our solution can reduce the demand for handcrafted priors and improve the adaptability of the network.

With paired low-light instances, we propose a novel learning-based LIE method, termed PairLIE. The core insight of our approach is to sufficiently exploit priors from paired low-light images. Therefore, we consider employing the Retinex theory and deep learning to decompose low-light images into illumination and reflectance components. First, since the two low-light inputs share the same content, the estimated reflectance components are expected to be consistent. Second, instead of directly imposing the Retinex decomposition on original low-light images, we adopt a simple self-supervised mechanism to remove inappropriate features and implement the Retinex decomposition on the optimized image. This can avoid sub-optimal estimations because the Retinex model has limitations in low-light modeling. As a result, with fewer prior constraints and a simpler network, the proposed PairLIE achieves competitive performance in public LIE datasets. In summary, the contributions of this paper are as follows:

- We propose a generic LIE solution using paired low-light images. The network is based on Retinex decomposition with several novel reference-free losses.
- To achieve an accurate decomposition, we first project the original image to remove inappropriate features.
- With fewer manually designed priors and a simpler network, the proposed solution achieves comparable performance against state-of-the-art methods.

## 2. Related Work

Over the decades, extensive LIE methods have been presented, which can be roughly categorized into conventional approaches and learning-based techniques.

### 2.1. Conventional Methods

Histogram-based techniques perform light enhancement by expanding the dynamic range of an image. For example, Park *et al.* [23] separated the dynamic range of the histogram into several parts and resized the gray-scale range based on the area ratio. Arici *et al.* [1] introduced penalty terms to avoid the unnatural look and visual artifacts of the enhanced image. Lee *et al.* [14] applied the layered difference representation of 2D histograms to amplify the gray-level differences between adjacent pixels.

Retinex-based methods first decompose the low-light image into reflectance and illumination components. Subsequently, those approaches either consider the reflectance as the enhanced image or adjust the illumination and then recompose it with reflectance to generate the enhanced result. Wang *et al.* [29] proposed a LIE algorithm to promote naturalness and enhance details for non-uniform illumination images. Fu *et al.* [6] used a weighted variational model to preserve the reflectance with more details. Guo *et al.* [8] first estimated the illumination by calculating the maximum value in R, G, and B channels. Then, they refined the illumination map by imposing a structure prior. Li *et al.* [17] improved the performance of LIE by introducing a noise map in the Retinex model. Xu *et al.* [34] proposed a texture-aware Retinex model solved by an alternating optimization algorithm. Hao *et al.* [9] presented a novel Retinex-based LIE method performed in a semi-decoupled way.

### 2.2. Learning-based Methods

Commonly, learning-based LIE methods rely on paired low-light and normal-light images. Lore *et al.* [19] designed a stacked sparse denoising auto-encoder to enhance low-light images. The proposed model was trained on synthetic image pairs. Wei *et al.* [32] first built a real-world low-light image enhancement dataset including low-light and normal-light image pairs. Then, they trained an end-to-end network in a fully-supervised way. Wang *et al.* [28] introduced an intermediate illumination map in the network to associate the low-light input with the expected enhancement result. Chen *et al.* [4] collected a dataset of short-exposure low-light images, with corresponding long-exposure reference maps. Based on the dataset, the authors developed a fully-convolutional network for enhancing low-light images. Zhang *et al.* [38] proposed a supervised approach to decompose the low-light image into illumination and reflectance components. Wu *et al.* [33] proposed a Retinex-based deep unfolding network to promote adaptability and

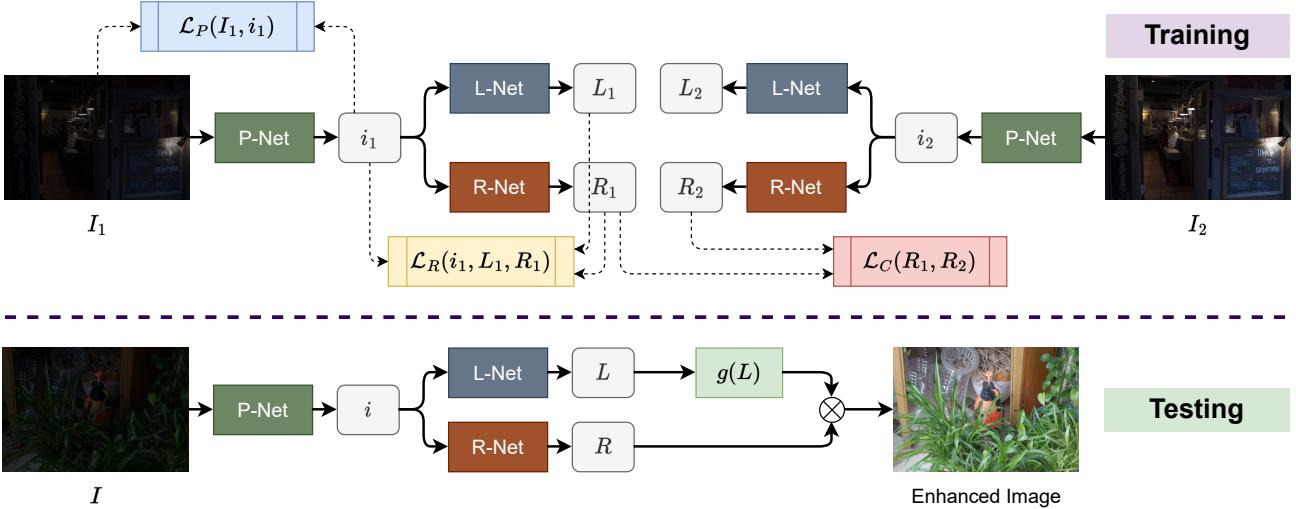


Figure 2. The architecture of PairLIE. In the training phase, P-Net is first adopted to remove inappropriate features of the original image. L-Net and R-Net are then employed to estimate the illumination and reflectance components. Three loss functions are used to guide the network optimization, including a self-supervised projection loss  $\mathcal{L}_P$ , a reflectance consistency loss  $\mathcal{L}_C$ , and a Retinex loss  $\mathcal{L}_R$ . In the testing stage, given a low-light image, P-Net, L-Net, and R-Net are used to decompose the input into reflectance and illumination. Subsequently, PairLIE adjusts the illumination and recomposes it with the reflectance to capture the enhanced image.

efficiency. Xu *et al.* [35] combined signal-to-noise-ratio-aware transformers and convolutional networks to enhance low-light images. Zhang *et al.* [39] proposed a color-consistent network to mitigate the color difference between the enhanced image and ground-truth.

Recently, unsupervised (also known as self-supervised) networks are developed to eliminate the requirement for reference images. For example, Zhu *et al.* [41] proposed a zero-shot LIE approach that trained a deep network using the input image itself. Guo *et al.* [7] proposed a reference-free LIE approach based on curve estimation. Their network was optimized with a set of non-reference loss functions. Liu *et al.* [18] proposed a lightweight LIE network by integrating the unrolling techniques and prior architecture search strategies. Jiang *et al.* [12] proposed a LIE method based on the generative adversarial network and unpaired training data. Zhao *et al.* [40] designed a unified zero-reference network for enhancing low-light images based on Deep image Prior (DIP) [27]. Jiang *et al.* [11] presented a Retinex-based unsupervised decomposition and correction network for LIE. Ma *et al.* [22] established a cascaded illumination estimation process to achieve fast, flexible, and robust LIE in complex scenarios.

### 3. Proposed Method

We first give the problem definition of using low-light image pairs for LIE. Then, we detail the pipeline and loss functions. Finally, we introduce the dataset for training the proposed network.

#### 3.1. Retinex Model with Paired Low-light Images

According to the Retinex theory, low-light image  $I$  can be decomposed into illumination  $L$  and reflectance  $R$  as:

$$I = L \circ R, \quad (1)$$

where  $\circ$  denotes the element-wise multiplication. Illumination  $L$  describes the light intensity of objects.  $L$  should be piece-wise continuous and textureless. Reflectance  $R$  represents the physical properties of the objects.  $R$  should contain textures and details in the observed image. The Retinex decomposition is highly ill-posed. Various methods have been proposed to handle this problem [6, 17, 29, 34]. A generic solution of the Retinex decomposition is to minimize the following energy function:

$$\underset{L,R}{\operatorname{argmin}} \|L \circ R - I\|_2 + \lambda_L f_L(L) + \lambda_R f_R(R), \quad (2)$$

where  $f_L$  and  $f_R$  are prior constraints for  $L$  and  $R$ , respectively.  $\lambda_L$  and  $\lambda_R$  denote the weights.  $\|L \circ R - I\|_2$  is the data-fidelity term between the input and reconstructed image. To achieve a reasonable decomposition, most of the LIE approaches focus on introducing assorted prior constraints enforced on the two components, such as priors of structure [17], smooth [32], and bright channel [15, 26].

However, handcrafted priors are commonly not adaptive enough due to the diverse natural scenes and light conditions. In this paper, instead of exploiting handcrafted priors for  $L$  and  $R$  from a single image, we apply paired low-light images to automatically learn adaptive priors in a data-driven fashion. Those low-light image pairs share the same

scene content but different illumination. Mathematically, Retinex decomposition with low-light image pairs can be expressed as:

$$\begin{cases} I_1 = L_1 \circ R \\ I_2 = L_2 \circ R \end{cases}, \quad (3)$$

where  $I_1$  and  $I_2$  are low-light image pairs, which share the reflectance component  $R$ . Intuitively, Eq. 3 can address the decomposition better than Eq. 1 because more information and constraints are introduced.

### 3.2. Network Structure

The whole pipeline of our method is illustrated in Fig. 2. We use L-Net and R-Net to estimate the illumination and reflectance components, respectively. L-Net and R-Net are very similar and simple, both of which contain five convolutional layers. The activation function of the first four convolutional layers is ReLU. L-Net and R-Net end with a sigmoid layer to normalize the output into  $[0, 1]$ . According to the Retinex theory, the three color channels are assumed to have the same illumination. Therefore, the output channel of L-Net is set as 1 while that of R-Net is set as 3. Note that, this paper does not focus on designing modernistic network structures. In contrast, we aim at providing a generic solution for LIE under paired low-light instances. In our experiments, we found those simple network already achieves comparable performance. Apart from L-Net and R-Net, we introduce P-Net to remove inappropriate features from the original image. Specifically, the structure of the P-Net is identical to the R-Net.

In the training phase, the original low-light image pairs  $I_1$  and  $I_2$  are first taken into the P-Net, yielding two optimized versions  $i_1$  and  $i_2$ . Then, L-Net and R-Net are applied to estimate the latent illumination ( $L_1$  and  $L_2$ ) and reflectance ( $R_1$  and  $R_2$ ). To optimize the network, three loss functions are designed in PairLIE. The first is the projection loss  $\mathcal{L}_P$  that measures the difference between  $I$  and  $i$ . The second is the reflectance consistency loss  $\mathcal{L}_C$  calculated based on  $R_1$  and  $R_2$ . The third is the Retinex loss  $\mathcal{L}_R$  that restricts the decomposed components to satisfy the Retinex theory. In the testing period, given a low-light image, P-Net, R-Net, and L-Net are applied to calculate the final enhanced image:

$$I_{en} = g(L) \circ R = L^\lambda \circ R, \quad (4)$$

where  $\lambda$  is the illumination correction factor,  $I_{en}$  denotes the enhanced image.

### 3.3. Projection Loss

Instead of performing the deep Retinex decomposition on the original low-light image, we propose to first remove inappropriate features to ensure the input can be accurately

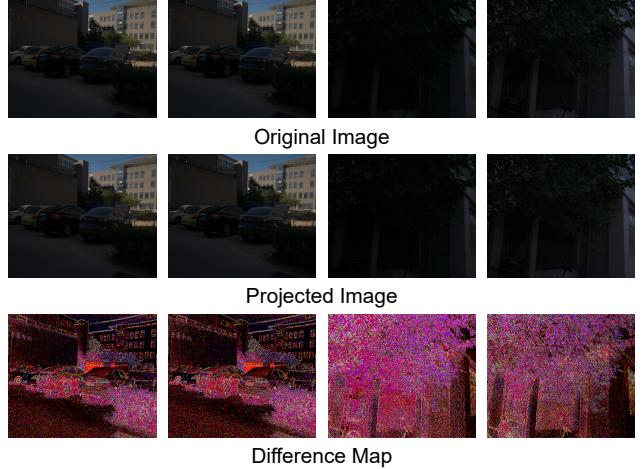


Figure 3. Example of projected images and corresponding difference maps. A smaller value indicates a higher similarity between the projected and original images.

decomposed by an ideal Retinex model (i.e., Eq. 1). Specifically, we design the following loss to guide projection:

$$\mathcal{L}_p = \|I_1 - i_1\|_2^2, \quad (5)$$

where  $i_1$  refers to the projected image. The loss function  $\mathcal{L}_p$  transforms the original image to a specific image that is more suitable for the Retinex decomposition. Concretely, this image is expected to be noise free since we do not consider the noise component in the Retinex model. Besides, some useless features are also discarded in this stage. Our projection loss can be interpreted from the view of error re-allocation, which can be expressed as:

$$\begin{aligned} & \underset{L, R}{\operatorname{argmin}} \|I - L \circ R - \varepsilon\| \\ &= \underset{L, R}{\operatorname{argmin}} \|I - i + i - L \circ R - \varepsilon\| \\ &\leq \|I - i - \varepsilon\| + \underset{L, R}{\operatorname{argmin}} \|i - L \circ R\|, \end{aligned} \quad (6)$$

where  $\varepsilon$  denotes the error. According to Eq. 6, the goal of  $\mathcal{L}_p$  is to move the error from the Retinex decomposition to the projection stage. Fig. 3 gives an example of the calculated projected image and corresponding difference map. As can be seen, the projected image is highly similar to the original one, and sensor noise is removed since it cannot be formulated by the Retinex model. Note that the projection loss needs to cooperate with the other constraints to avoid a trivial solution, i.e.,  $i_1 = I_1$ .

### 3.4. Reflectance Consistency Loss

Reflectance consistency loss  $\mathcal{L}_C$  is calculated based on low-light image pairs and the Retinex theory. Compared

with handcrafted priors,  $\mathcal{L}_C$  is more accurate and adaptive because it reveals the physical properties of the objects. Mathematically,  $\mathcal{L}_C$  is formulated as:

$$\mathcal{L}_C = \|R_1 - R_2\|_2^2, \quad (7)$$

where  $R_1$  and  $R_2$  refer to the reflectance components of paired low-light images.  $\mathcal{L}_C$  enforces the network to predict the same reflectance components because low-light image pairs share the same objects.

Since sensor noise hidden in dark regions will be amplified when the contrast is improved. To cope with the noise issue, existing approaches either add a smoothness term on the estimated reflectance or perform a denoise operation after the enhancement. In our method, the sensor noise can be implicitly removed by Eq. 7. This is because the two low-light images contain independent noise of the same scene. As discussed in [16], paired noisy images can be leveraged to train denoising models. This is because noise is random and different in two images, the deep network cannot fit the noise in one image to another. In our case, the two low-light images can help each other to remove noise during the Retinex decomposition. Therefore, PairLIE does not require additional handcrafted constraints of noise.

### 3.5. Retinex Loss

Since PairLIE employs the Retinex theory to decompose low-light images into illumination and reflectance components, some basic constraints of Retinex decomposition are introduced, which can be formulated as:

$$\begin{aligned} \mathcal{L}_R = & \|R \circ L - i\|_2^2 + \|R - i/\text{stopgrad}(L)\|_2^2 \\ & + \|L - L_0\|_2^2 + \|\nabla L\|_1, \end{aligned} \quad (8)$$

where  $i$  denotes the projected image,  $L_0$  refers to the initial estimation of illumination,  $\nabla$  represents the horizontal and vertical gradients. First, the decomposed components should satisfy the requirement for reconstructing the input image. Therefore, a reconstruction term  $\|R \circ L - i\|_2^2$  is applied to ensure a reasonable decomposition. Once the illumination is estimated, the reflectance can be calculated through the pixel-wise division between the low-light image and its illumination map. Therefore, we additionally add a term  $\|R - i/\text{stopgrad}(L)\|_2^2$  to guide the decomposition. Note that we stop the gradient of the illumination to make the training stable. Different from most existing methods that use a lot of handcrafted priors, PairLIE only imposes a smooth term and an initialization term on  $L$ . Specifically, the initialized illumination  $L_0$  is calculated via the maximum of the R, G, and B channels:

$$L_0 = \max_{c \in \{R, G, B\}} I^c(x). \quad (9)$$

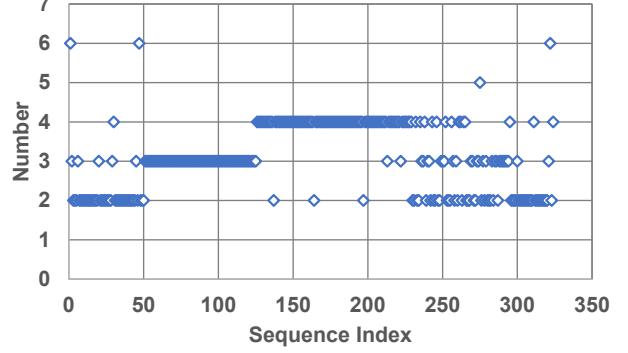


Figure 4. Statistics concerning sequence length of the collected data. Most sequences have 2 to 4 low-light images.

### 3.6. Overall Loss

The overall loss function for training our PairLIE is a liner combination of each loss:

$$\mathcal{L}_{All} = w_0 \mathcal{L}_P + w_1 \mathcal{L}_C + w_2 \mathcal{L}_R, \quad (10)$$

where  $w_0$ ,  $w_1$ , and  $w_2$  denote the weights.

### 3.7. Training Data Collection

We collect low-light image pairs from SICE (part2) [2] and LOL (training set) [32], which contain multi-exposure images. These datasets adopt some specific operators to deal with the misalignment caused by camera shaking or object moving. Note that the SICE dataset consists of both under and over-exposed images. We only select under-exposure and well-aligned cases for constructing the training set. As a result, we collect 324 sequences (a total of 1000 low-light images). As shown in Fig. 4, each sequence includes 2 to 6 samples. In the training phase, we randomly select two images from each sequence to constitute a pair.

## 4. Experiments

We first describe implementation details, evaluation datasets, and performance criteria. Then, we present the quantitative and qualitative comparisons with state-of-the-art methods. Finally, we conduct ablative experiments to validate each component.

### 4.1. Implementation Details

We implement PairLIE with PyTorch. In the training phase, we randomly crop images to the size of  $128 \times 128$ . A batch size of 1 is applied. We use ADAM with the initial learning rate of  $1 \times 10^{-4}$  to optimize the network. The number of training epochs is set to 400. The learning rate is half-decayed per 100 epochs. The default correction factor  $\lambda$  is 0.2. In extremely dark cases such as the LOL dataset, we set  $\lambda = 0.14$ . As for the hyper-parameter  $w_0$ ,  $w_1$ , and  $w_2$  in Eq. 10, we set  $w_0 = 500$ ,  $w_1 = w_2 = 1$ , empirically.

Method	Type	LOL				SICE			
		PSNR↑	SSIM↑	LPIPS↓	DeltaE ↓	PSNR↑	SSIM↑	LPIPS↓	DeltaE ↓
SDD [9]	T	13.34	0.637	0.743	21.83	15.35	0.741	0.232	16.08
STAR [34]	T	12.91	0.518	0.366	23.46	15.17	0.727	0.246	16.35
MBLLEN [20]	S	17.86	0.727	<b>0.225</b>	13.68	13.64	0.632	0.297	18.60
RetinexNet [32]	S	17.55	0.648	0.379	12.69	19.89	0.783	0.276	8.715
GLADNet [30]	S	<b>19.72</b>	0.680	0.321	<b>12.28</b>	19.98	0.837	<b>0.203</b>	8.947
KinD [38]	S	17.65	<b>0.775</b>	<b>0.171</b>	12.49	<b>21.10</b>	<b>0.838</b>	<b>0.195</b>	<b>8.009</b>
DRBN [36]	S	16.29	0.551	0.260	13.44	15.58	0.522	0.289	13.78
URetinexNet [33]	S	<b>19.84</b>	<b>0.826</b>	<b>0.128</b>	<b>10.65</b>	<b>21.64</b>	<b>0.843</b>	<b>0.192</b>	<b>7.728</b>
ZeroDCE [7]	U	14.86	0.559	0.335	18.81	18.69	0.810	0.207	11.93
RRDNet [41]	U	11.40	0.457	0.362	26.43	13.28	0.678	0.221	19.64
RUAS [18]	U	16.40	0.500	0.270	16.85	13.18	0.734	0.363	16.81
SCI [22]	U	14.78	0.522	0.339	19.52	15.95	0.787	0.235	13.71
EnlightenGAN [12]	U	17.48	0.651	0.322	14.50	18.73	0.822	0.216	10.42
PairLIE (Ours)	U	<b>19.51</b>	<b>0.736</b>	0.248	<b>10.80</b>	<b>21.32</b>	<b>0.840</b>	0.216	<b>7.835</b>

Table 1. Quantitative comparisons with state-of-the-art methods on LOL and SICE datasets. “T”, “S”, and “U” represent “Traditional”, “Supervised”, and “Unsupervised” methods, respectively. The top three results are marked in bold.

## 4.2. Datasets and Criteria

As described before, low-light image pairs collected from SICE and LOL are applied to train PairLIE. We select another 50 sequences (150 images) from SICE and use the official evaluation set (15 images) of LOL to measure the model performance. Since SICE and LOL contains reference images, we employ PSNR, SSIM [31], LPIPS [37], and DeltaE with CIE2000 standard [24] to objectively evaluate the performance of each method. A higher PSNR/SSIM score indicates the result is closer to the reference. A lower LPIPS/DeltaE value denotes better enhancement performance. Furthermore, we adopt the MEF dataset [21] for visual comparisons.

## 4.3. Compared Methods

PairLIE is compared with 13 state-of-the-art LIE methods, which can be divided into the following three categories: traditional methods (SDD [9], STAR [34]), supervised approaches (MBLLEN [20], RetinexNet [32], GALDNet [30], KinD [38], DRBN [36], URetinexNet [33]), and unsupervised methods (Zero-DCE [7], RRDNet [41], RUAS [18], SCI [22], and EnlightenGAN [12]). Note that the results of all those methods are reproduced by using the official codes with recommended parameters.

## 4.4. Quantitative Comparisons

Tab. 1 lists the quantitative performance of the LOL and SICE datasets. As can be observed, traditional and unsupervised methods obtain relatively poor results. This is reasonable because it is difficult to learn an accurate enhancement model without the reference image. Besides, the performance of those solutions is highly dependent on the used

priors. However, hand-crafted features may not be adaptive enough for assorted light conditions. In Tab. 1, PairLIE achieves the best performance among the five unsupervised methods and gets competitive results compared with supervised approaches. Since paired low-light images provide sufficient information for solving the LIE task, PairLIE can reduce the demand for handcrafted priors. As a result, PairLIE achieves significant performance improvement.

## 4.5. Visual Comparisons

Fig. 5 illustrates the visual comparisons of different LIE methods. Our observations are as follows: 1) The proposed method achieves visually pleasing results in terms of brightness, color, contrast, and naturalness. While other methods fail to cope with the extreme black light in this case. 2) Although supervised approaches GLADNet, KinD, and URetinexNet show good performance on LOL and SICE datasets according to Tab. 1, their generalization ability may be limited as supervised learning is sensitive to the data distribution. We further show an example of noise suppression in Fig. 6. As can be seen, although PairLIE does not introduce any handcrafted priors about the noise, our method can successfully suppress sensor noise in dark regions and the result is clear and natural. In contrast, the competitors either amplify noise or fail in correcting color and contrast, leading to poor visual quality.

## 4.6. Decomposition Visualization

We visualize the reflectance and illumination components to validate the effectiveness of our method. As shown in Fig. 7, the reflectance contains rich textures and details while the illumination is piece-wise continuous and textureless, demonstrating PairLIE can accurately decompose the

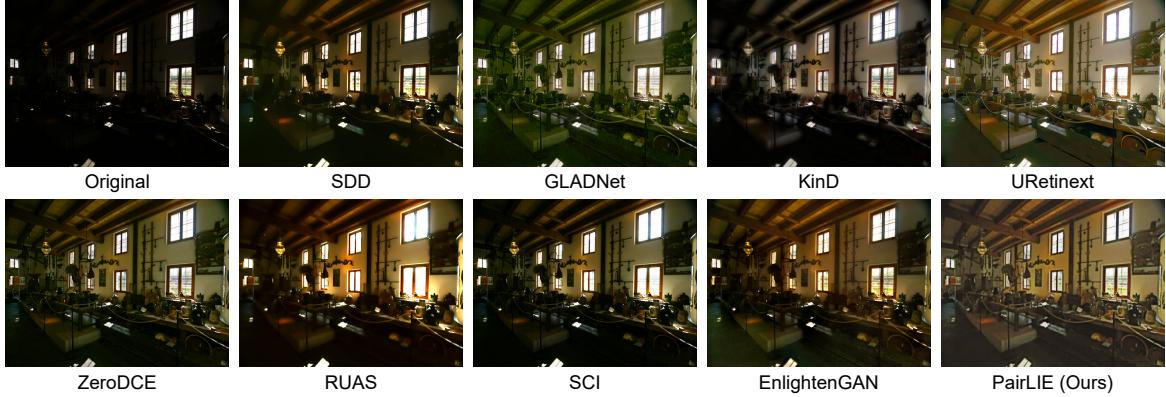


Figure 5. Visual comparisons of different LIE methods. SDD [9], RetinexNet [32], GaldNet [30], KinD [38], URetinext [33], Zero-DCE [7], RRDNet [41], RUAS [18], SCI [22], EnlightenGAN [12], and PairLIE (Ours).

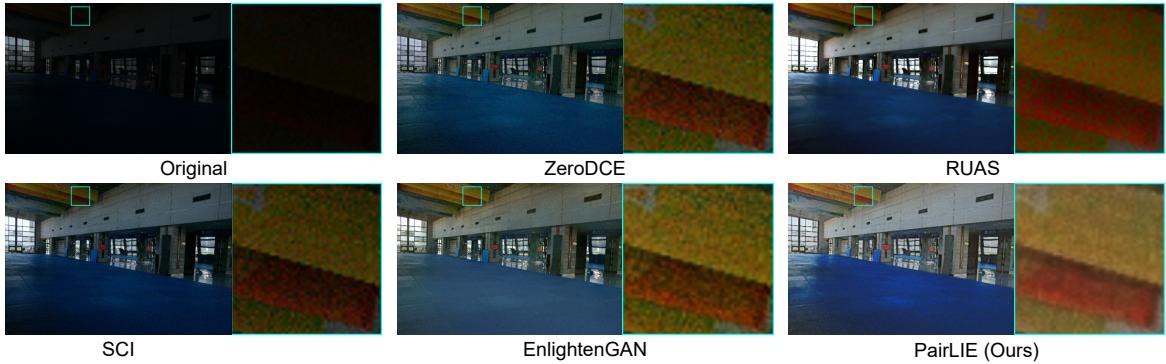


Figure 6. Performance comparisons in noise removal. Zoom in for the best view. Zero-DCE [7], RUAS [18], SCI [22], EnlightenGAN [12], and PairLIE (Ours). Our result is visually satisfactory without obvious noise.

low-light image. We further show the enhancement results with different correction factors. With the increase of  $\lambda$ , the image brightness decreases gradually. When  $\lambda$  is larger than 0.4 or smaller than 0.2, the enhanced images show obvious under/over-enhancement effects. In this work, the default  $\lambda$  is set as 0.2. Note that users can adjust  $\lambda$  in the testing phase according to their personal preferences.

#### 4.7. Ablation Studies

We conduct ablation studies under different settings to understand how each component affects performance. We have tried the following three variations over the original PairLIE. 1) Setting A: without  $\mathcal{L}_C$ . 2) Setting B: without  $\mathcal{L}_P$ . 3) Setting C: without prior terms, i.e., removing  $\|L - L_0\|_2^2$  and  $\|\nabla L\|_1$  from  $\mathcal{L}_R$ .

Tab. 2 and Fig. 8 report the results of our ablation studies on the LOL dataset. We can observe that our approach significantly outperforms setting A, which demonstrates the effectiveness of learning adaptive priors using low-light image pairs. Compared with setting B, PairLIE achieves slight performance improvement. The results of setting B suggest

that P-Net can assist the Retinex decomposition, but it is not necessary for stably training the network. To further understand the role of P-Net, we show the reconstruction error  $\kappa$  in Fig. 9. The blue line denotes the decomposition using original low-light images, i.e.,  $\kappa = |L \circ R - I|$ , while the orange line refers to the decomposition using projected images, i.e.,  $\kappa = |L \circ R - i|$ . As can be seen, the proposed method obtains a smaller reconstruction error than the baseline. Therefore, P-Net can remove inappropriate features and the output is more suitable for the Retinex decomposition. The results of Setting C show that when priors terms are removed, the performance drops significantly. This is because the Retinex decomposition requires some basic priors to initialize and restrain the illumination.

#### 4.8. Discussions

**Acquiring low-light image pairs.** Although PairLIE achieves promising performance with fewer handcrafted priors and a simpler network, collecting low-light image pairs is relatively expensive compared with recording a single low-light image. This problem can be partly solved by

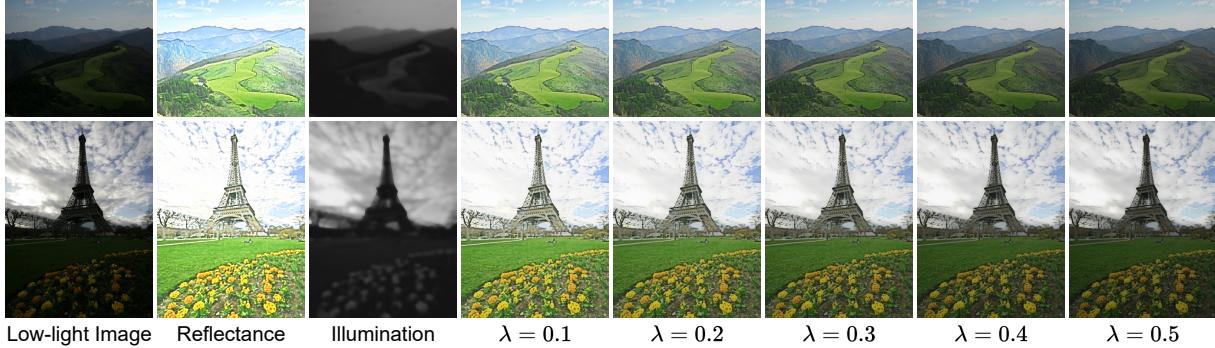


Figure 7. Visualization of the Retinex decomposition. The enhanced results are presented under different correction factors. We set the default  $\lambda$  as 0.2, but users can adjust the  $\lambda$  according to their preferences.



Figure 8. Visual comparisons of the ablation studies. Setting A: without  $\mathcal{L}_C$ . Setting B: without  $\mathcal{L}_P$ . Setting C: without prior terms. Our result is visually closer to the reference.

Method	PSNR↑	SSIM↑
Setting A	17.70	0.567
Setting B	19.50	0.724
Setting C	14.69	0.684
<b>PairLIE</b>	<b>19.51</b>	<b>0.736</b>

Table 2. Quantitative results of ablation studies on the LOL dataset. The best results are marked in bold.

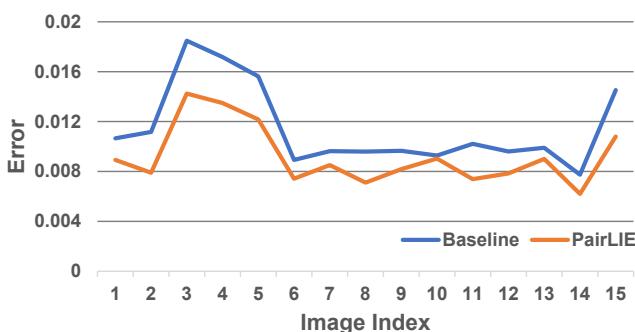


Figure 9. Visualization of the reconstruction error. Baseline refers to the decomposition with original low-light images.

generating training pairs from a single low-light image. For example, one can use neighbor masks [10] to generate two similar low-light images from a degraded observation. Additionally, specific recorruption techniques can be leveraged to capture paired training instances.

**Illumination components.** PairLIE solves the Retinex decomposition with paired low-light images. The basic idea is to generate the same reflectance components since the reflectance is independent of the light. However, the relationship between illumination components is not sufficiently exploited in this paper. Illumination plays a critical role in the Retinex model. Therefore, we believe that PairLIE can be further improved by introducing new constraints between two illumination maps.

## 5. Conclusions

This paper proposes PairLIE, a reference-free approach for low-light image enhancement that benefits from both Retinex-based and learning-based solutions. By learning adaptive constraints from low-light image pairs, PairLIE reduces the dependence on handcrafted priors and thus generalizes well on various scenes. To assist the decomposition, PairLIE first removes inappropriate features in the original image and then implements the decomposition on the optimized image. Extensive experiments on public benchmarks show that PairLIE outperforms the state-of-the-art unsupervised methods significantly. In future works, we will concentrate on exploiting effective priors of illumination in a data-driven fashion.

**Acknowledgements:** The work was supported partly by the National Natural Science Foundation of China under Grants 82172033, U19B2031, 61971369, 52105126, 82272071, and 62271430.

## References

- [1] Tarik Arici, Salih Dikbas, and Yucel Altunbasak. A histogram modification framework and its application for image contrast enhancement. *IEEE TIP*, 18(9):1921–1935, 2009. 1, 2
- [2] Jianrui Cai, Shuhang Gu, and Lei Zhang. Learning a deep single image contrast enhancer from multi-exposure images. *IEEE TIP*, 27(4):2049–2062, 2018. 5
- [3] Turgay Celik and Tardi Tjahjadi. Contextual and variational contrast enhancement. *IEEE TIP*, 20(12):3431–3441, 2011. 1
- [4] Chen Chen, Qifeng Chen, Jia Xu, and Vladlen Koltun. Learning to see in the dark. In *CVPR*, pages 3291–3300, 2018. 2
- [5] Xueyang Fu, Delu Zeng, Yue Huang, Yinghao Liao, Xinghao Ding, and John Paisley. A fusion-based enhancing method for weakly illuminated images. *Signal Processing*, 129:82–96, 2016. 1
- [6] Xueyang Fu, Delu Zeng, Yue Huang, Xiao-Ping Zhang, and Xinghao Ding. A weighted variational model for simultaneous reflectance and illumination estimation. In *CVPR*, pages 2782–2790, 2016. 1, 2, 3
- [7] 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 *CVPR*, pages 1777–1786, 2020. 2, 3, 6, 7
- [8] Xiaojie Guo, Yu Li, and Haibin Ling. Lime: Low-light image enhancement via illumination map estimation. *IEEE TIP*, 26(2):982–993, 2017. 1, 2
- [9] Shijie Hao, Xu Han, Yanrong Guo, Xin Xu, and Meng Wang. Low-light image enhancement with semi-decoupled decomposition. *IEEE TMM*, 22(12):3025–3038, 2020. 2, 6, 7
- [10] Tao Huang, Songjiang Li, Xu Jia, Huchuan Lu, and Jianzhuang Liu. Neighbor2neighbor: Self-supervised denoising from single noisy images. In *CVPR*, pages 14781–14790, 2021. 8
- [11] Qiuping Jiang, Yudong Mao, Runmin Cong, Wenqi Ren, Chao Huang, and Feng Shao. Unsupervised decomposition and correction network for low-light image enhancement. *IEEE TITS*, pages 1–16, 2022. 3
- [12] Yifan Jiang, Xinyu Gong, Ding Liu, Yu Cheng, Chen Fang, Xiaohui Shen, Jianchao Yang, Pan Zhou, and Zhangyang Wang. Enlightengan: Deep light enhancement without paired supervision. *IEEE TIP*, 30:2340–2349, 2021. 3, 6, 7
- [13] Edwin H Land. The retinex theory of color vision. *Scientific American*, 237(6):108–129, 1977. 1
- [14] Chulwoo Lee, Chul Lee, and Chang-Su Kim. Contrast enhancement based on layered difference representation of 2d histograms. *IEEE TIP*, 22(12):5372–5384, 2013. 1, 2
- [15] Hunsang Lee, Kwanghoon Sohn, and Dongbo Min. Unsupervised low-light image enhancement using bright channel prior. *IEEE SPL*, 27:251–255, 2020. 3
- [16] Jaakko Lehtinen, Jacob Munkberg, Jon Hasselgren, Samuli Laine, Tero Karras, Miika Aittala, and Timo Aila. Noise2Noise: Learning image restoration without clean data. In *ICML*, pages 2965–2974, 2018. 5
- [17] Mading Li, Jiaying Liu, Wenhan Yang, Xiaoyan Sun, and Zongming Guo. Structure-revealing low-light image enhancement via robust retinex model. *IEEE TIP*, 27(6):2828–2841, 2018. 1, 2, 3
- [18] 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 *CVPR*, pages 10561–10570, 2021. 2, 3, 6, 7
- [19] 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. 2
- [20] Feifan Lv, Feng Lu, Jianhua Wu, and Chongsoon Lim. Mbllen: Low-light image/video enhancement using cnns. In *BMVC*, pages 1–13, 2018. 6
- [21] Kede Ma, Kai Zeng, and Zhou Wang. Perceptual quality assessment for multi-exposure image fusion. *IEEE TIP*, 24(11):3345–3356, 2015. 6
- [22] Long Ma, Tengyu Ma, Risheng Liu, Xin Fan, and Zhongxuan Luo. Toward fast, flexible, and robust low-light image enhancement. In *CVPR*, pages 5637–5646, 2022. 3, 6, 7
- [23] Gyu-Hee Park, Hwa-Hyun Cho, and Myung-Ryul Choi. A contrast enhancement method using dynamic range separate histogram equalization. *IEEE TCE*, 54(4):1981–1987, 2008. 2
- [24] Gaurav Sharma, Wencheng Wu, and Edul N Dalal. The ciede2000 color-difference formula: Implementation notes, supplementary test data, and mathematical observations. *Color Research and Application*, 30(1):21–30, 2005. 6
- [25] Weitao Song, Masanori Suganuma, X. Liu, Norimasa Shimobayashi, D. Maruta, and Takayuki Okatani. Matching in the dark: A dataset for matching image pairs of low-light scenes. In *ICCV*, pages 6009–6018, 2021. 1
- [26] Li Tao, Chuang Zhu, Jiawen Song, Tao Lu, Huizhu Jia, and Xiaodong Xie. Low-light image enhancement using cnn and bright channel prior. In *ICIP*, pages 3215–3219, 2017. 3
- [27] Dmitry Ulyanov, Andrea Vedaldi, and Victor Lempitsky. Deep image prior. In *CVPR*, pages 9446–9454, 2018. 3
- [28] Ruixing Wang, Qing Zhang, Chi-Wing Fu, Xiaoyong Shen, Wei-Shi Zheng, and Jiaya Jia. Underexposed photo enhancement using deep illumination estimation. In *CVPR*, pages 6842–6850, 2019. 1, 2
- [29] Shuhang Wang, Jin Zheng, Hai-Miao Hu, and Bo Li. Naturalness preserved enhancement algorithm for non-uniform illumination images. *IEEE TIP*, 22(9):3538–3548, 2013. 2, 3
- [30] Wenjing Wang, Chen Wei, Wenhan Yang, and Jiaying Liu. Gladnet: Low-light enhancement network with global awareness. In *IEEE International Conference on Automatic Face and Gesture Recognition*, pages 751–755, 2018. 6, 7
- [31] Zhou Wang, Alan C Bovik, Hamid R Sheikh, and Eero P Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE TIP*, 13(4):600–612, 2004. 6
- [32] Chen Wei, Wenjing Wang, Wenhan Yang, and Jiaying Liu. Deep retinex decomposition for low-light enhancement. In *BMVC*, pages 1–12, 2018. 2, 3, 5, 6, 7
- [33] Wenhui Wu, Jian Weng, Pingping Zhang, Xu Wang, Wenhan Yang, and Jianmin Jiang. Uretinex-net: Retinex-based

- deep unfolding network for low-light image enhancement. In *CVPR*, pages 5901–5910, 2022. 2, 6, 7
- [34] Jun Xu, Yingkun Hou, Dongwei Ren, Li Liu, Fan Zhu, Mengyang Yu, Haoqian Wang, and Ling Shao. Star: A structure and texture aware retinex model. *IEEE TIP*, 29:5022–5037, 2020. 2, 3, 6
- [35] Xiaogang Xu, Ruixing Wang, Chi-Wing Fu, and Jiaya Jia. Snr-aware low-light image enhancement. In *CVPR*, pages 17714–17724, 2022. 1, 3
- [36] Wenhan Yang, Shiqi Wang, Yuming Fang, Yue Wang, and Jiaying Liu. From fidelity to perceptual quality: A semi-supervised approach for low-light image enhancement. In *CVPR*, pages 3060–3069, 2020. 6
- [37] Richard Zhang, Phillip Isola, Alexei A Efros, Eli Shechtman, and Oliver Wang. The unreasonable effectiveness of deep features as a perceptual metric. In *CVPR*, pages 586–595, 2018. 6
- [38] Yonghua Zhang, Jiawan Zhang, and Xiaojie Guo. Kindling the darkness: A practical low-light image enhancer. In *ACM MM*, pages 1632–1640, 2019. 2, 6, 7
- [39] Zhao Zhang, Huan Zheng, Richang Hong, Mingliang Xu, Shuicheng Yan, and Meng Wang. Deep color consistent network for low-light image enhancement. In *CVPR*, pages 1899–1908, 2022. 3
- [40] Zunjin Zhao, Bangshu Xiong, Lei Wang, Qiaofeng Ou, Lei Yu, and Fa Kuang. Retinexdip: A unified deep framework for low-light image enhancement. *IEEE TCSVT*, 32(3):1076–1088, 2022. 2, 3
- [41] Anqi Zhu, Lin Zhang, Ying Shen, Yong Ma, Shengjie Zhao, and Yicong Zhou. Zero-shot restoration of underexposed images via robust retinex decomposition. In *ICME*, pages 1–6, 2020. 2, 3, 6, 7