

## Deep parametric Retinex decomposition model for low-light image enhancement

Xiaofang Li <sup>a</sup>, Weiwei Wang <sup>a,b,\*</sup>, Xiangchu Feng <sup>a</sup>, Min Li <sup>b</sup>

<sup>a</sup> School of Mathematics and Statistics, Xidian University, Xi'an 710071, China

<sup>b</sup> College of Mathematics and Statistics, Shenzhen University, Shenzhen, 518060, China



### ARTICLE INFO

Communicated by Jinshan Pan

MSC:  
41A05  
41A10  
65D05  
65D17

**Keywords:**

Low-light image enhancement  
Parametric Retinex decomposition  
Deep network  
Denoising

### ABSTRACT

Images captured under low light conditions often suffer from various degradations. The Retinex models are highly effective in enhancing low-light images. The analytical optimization models are interpretable but inflexible to various scenes. The data-driven learning models are flexible to various scenes but less interpretable. To reconcile the advantages of both, we propose a parametric Retinex model with pixel-wise varying parameters. Then we unroll its iterative algorithm into an unfolding network so that the parameters can be learned. We call it Deep Parametric REtinex Decomposition (DPRED). Based on the Retinex decomposition, we present a novel network for low-light image enhancement, also called DPRED. The whole network comprises three modules: parametric Retinex decomposition, enhancement and refinement. The first two modules operate on the V channel in the HSV space, avoiding color deviation. The refinement module aims to remove noise in the enhanced RGB image. Extensive experiments demonstrate the proposed method is effective in low-light image enhancement and it significantly outperforms recent baselines.

### 1. Introduction

In computer vision and image analysis, high-quality images are needed in many tasks, such as object detection, tracking, segmentation and so on. However, images captured under low-light conditions often suffer from various degradations, including low visibility, low contrast, noise and color deviation, etc (Lv et al., 2022; Wang et al., 2023; Ren et al., 2019). Therefore, it is necessary to enhance low-light images before subsequent processing and analysis. In the past few decades, considerable efforts have been made in low-light image enhancement (LLIE). Existing methods roughly fall into two categories: analytical methods (Pizer et al., 1990; Jobson et al., 1997a,b; Guo et al., 2016; Cai et al., 2017; Lee et al., 2013) and recently developed deep learning methods (Lore et al., 2017; Lv et al., 2018; Wei et al., 2018; Zhang et al., 2019, 2021; Yang et al., 2021; Wu et al., 2022; Liu et al., 2022; Jiang et al., 2021a; Liu et al., 2021; Jiang et al., 2021b; Guo et al., 2020; Li et al., 2022).

Histogram Equalization (HE) (Pizer et al., 1990) and Retinex (Land, 1977) based methods (Jobson et al., 1997b,a; Guo et al., 2016; Cai et al., 2017; Xu et al., 2020) are representatives of the first category. HE (Pizer et al., 1990) methods stretch the dynamic range of images, which also amplify intensive noise hidden in the dark region. As for the Retinex (Land, 1977) based methods (Jobson et al., 1997b,a; Guo et al., 2016; Cai et al., 2017; Xu et al., 2020), the LLIE task is performed in

two phases: Retinex decomposition and enhancement. In the first phase, an image is factorized into an illumination component and a reflectance component by using regularities such as Total Variation (TV) (Rudin et al., 1992) and its variants (Guo et al., 2016; Cai et al., 2017; Xu et al., 2020). The illumination component describes the light intensity shot on objects, and the reflectance component characterizes the intrinsic property of objects. In the enhancement phase, some works (Jobson et al., 1997b,a) simply take the reflectance as the final enhanced image, yet it looks unrealistic and has color deviation. Other works (Guo et al., 2016; Cai et al., 2017; Xu et al., 2020) further enhance the estimated illumination by the Gamma Correction (GC) and then recombine it with the estimated reflectance to obtain the final enhanced image.

In recent years, inspired by the great success of deep neural networks in images processing, deep learning-based methods (Lore et al., 2017; Lv et al., 2018; Wei et al., 2018; Zhang et al., 2019, 2021; Yang et al., 2021; Wu et al., 2022; Liu et al., 2022; Jiang et al., 2021a; Liu et al., 2021; Jiang et al., 2021b; Guo et al., 2020; Li et al., 2022) have been proposed. Early supervised learning methods (Lore et al., 2017; Lv et al., 2018) train simple end-to-end networks to implement the whole task by using synthetic image pairs. Unfortunately, the performance is unsatisfactory for real-world images. Later, the Retinex decomposition framework inspired some complex networks. Wei et al. (2018) proposed the RetinexNet, which consists of two sub-networks:

\* Corresponding author at: School of Mathematics and Statistics, Xidian University, Xi'an 710071, China.

E-mail addresses: [xli\\_96@stu.xidian.edu.cn](mailto:xli_96@stu.xidian.edu.cn) (X. Li), [wwwang@mail.xidian.edu.cn](mailto:wwwang@mail.xidian.edu.cn) (W. Wang), [xcfeng@mail.xidian.edu.cn](mailto:xcfeng@mail.xidian.edu.cn) (X. Feng), [limin800@szu.edu.cn](mailto:limin800@szu.edu.cn) (M. Li).

the Retinex decomposition network and the illumination enhancement network. The former one conducts the Retinex decomposition, and the later one enhances the illumination component. The analytic method Block-Matching and 3D filtering (BM3D) (Dabov et al., 2007) is also used to refine the reflectance. To train the network, the authors (Wei et al., 2018) constructed the Low Light (LOL) dataset, in which real images were captured at various exposure times and ISO (International Organization for Standardization). Kindling the Darkness (KinD) (Zhang et al., 2019) is a modified RetinexNet (Wei et al., 2018), with the BM3D (Dabov et al., 2007) being replaced by a sub-network for reflectance refinement. KinD++ (Zhang et al., 2021) goes further by incorporating a multi-scale illumination attention module in the reflectance restoration network. Yang et al. (2021) also improved their RetinexNet (Wei et al., 2018) by using a reflectance restoration network to substitute the BM3D (Dabov et al., 2007), and introducing the sparse gradient regularization into the loss of the Retinex decomposition network. Different from these models, some networks (Wu et al., 2022; Liu et al., 2022) are constructed by unfolding optimization algorithms of the Retinex decomposition. Specifically, Wu et al. (2022) applied the Half Quadratic Splitting (HQS) (Geman and Yang, 1995) scheme to an implicitly regularized Retinex decomposition model and unfolded the iteration into the whole URetinex-Net, in which the sub-problems involving the illumination and the reflectance are replaced by two sub-networks. Note that the illuminance is further refined by an illumination enhancement network. Liu et al. (2022) formulated LLIE by a bi-level programming, with image enhancement as the upper level task and the Retinex decomposition as the lower level task. The gradient descent algorithm is used to solve the regularization problem of the illuminance and the gradient of the regularization term is computed by a network. The reflectance is also smoothed by a noise removal network.

Using unpaired images for training, EnlightenGAN (Jiang et al., 2021a) adopts an attention-guided U-Net as generator, and uses a global-local discriminator to encourage the enhanced result close to realistic image with normal-light. Without need of paired or unpaired images, and using non-reference loss in training, zero-shot learning methods (Liu et al., 2021; Jiang et al., 2021b; Guo et al., 2020; Li et al., 2022) have recently attracted much attention. Liu et al. (2021) proposed the Retinex-inspired Unrolling with Architecture Search (RUAS), which is trained by non-reference losses and the network architecture is searched automatically in the training phase. Jiang et al. (2021b) designed a self-regularized reflectance estimation network, which encourages reflectance consistency between the input and its randomly disturbed brightness, and the estimated reflectance is taken as the enhanced result. Note that the Retinex decomposition is performed on the V channel in HSV color space (Smith, 1978), and the enhanced result is recombined with the H and S channels to obtain the RGB enhanced result. Guo et al. (2020) regarded LLIE as image specific curve estimation so that the curve can map the low-light input to a normal light output. To achieve that, they designed the Zero-Reference Deep Curve Estimation (Zero-DCE) network to enhance the low-light image. Further, they reduced the number of parameters in Zero-DCE, and presented an accelerated and lightweight version Zero-DCE++ (Li et al., 2022).

Although unsupervised learning methods (Jiang et al., 2021a) and zero-shot learning methods (Liu et al., 2021; Jiang et al., 2021b; Guo et al., 2020; Li et al., 2022) have shown good performance, they are less effective than the supervised methods in recovering details lost under low light and removing noise. Yet, previous supervised learning methods also have limitations. For example, the methods (Wei et al., 2018; Zhang et al., 2019, 2021; Yang et al., 2021) conduct Retinex decomposition by pure networks, lacking interpretability. Second, performing Retinex decomposition and enhancement in the RGB space, the methods (Wei et al., 2018; Zhang et al., 2019, 2021; Yang et al., 2021; Liu et al., 2022, 2021) often suffer from color deviation. The methods (Cai et al., 2017; Xu et al., 2020; Jiang et al., 2021b) perform

Retinex decomposition and enhancement on the V channel in the HSV space, showing effective reduction in color deviation, yet they ignore the noise in the H and S channels.

In this work, we propose a new network for enhancing low-light images. Specifically, it consists of three modules: Retinex decomposition, enhancement and refinement. For Retinex decomposition, we first present an analytical Retinex decomposition model for V channel of low light color images. In the model we introduce a new parametric regularization with pixel-wise varying parameters. To learn the pixel-wise varying parameters, we present an iterative algorithm for solving the model and then unroll the iteration into an unfolding network with parameters embedded in. The parameters in the network are learned by supervised learning. Secondly, we use two networks to enhance the illuminance and reflectance respectively. Note that this operation is performed on the V channel in HSV color space to avoid color deviation. Finally, we recombine the enhanced V channel with unchanged H and S channels, and convert the HSV images to the RGB color space to get a coarsely enhanced result. As the channels H and S may be noisy, the coarsely enhanced result is also noisy. Then a refinement network is utilized to remove the noise in the coarsely enhanced result. In line with the previous methods (Wei et al., 2018; Zhang et al., 2019, 2021; Yang et al., 2021; Wu et al., 2022; Liu et al., 2022, 2021; Jiang et al., 2021b), our proposed method adopts the Retinex decomposition framework. However, our approach differs from the methods (Wei et al., 2018; Zhang et al., 2019, 2021; Yang et al., 2021; Jiang et al., 2021b) in that they implement the Retinex decomposition through pure networks, while we give a parametric regularization model and our network unfolds the iteration algorithm of our model. The Retinex decomposition network in methods (Wu et al., 2022; Liu et al., 2022, 2021) are also constructed by unrolling optimization iterations. However, in their networks, the regularity terms are completely implicit while our regularity is explicit, only the pixel-wise varying parameters need to be learned.

In summary, the contributions of this work are as follows:

- We present a parametric regularization model with pixel-wise varying parameters for Retinex decomposition of the V channel of low light color images. To learn the pixel-wise varying parameters, we present an iterative algorithm for solving the model and unroll the iteration into an unfolding network with parameters embedded in.
- We develop a novel network for LLIE, which consists of three modules: parametric Retinex decomposition, enhancement, and refinement.
- Experimental results show that our method can effectively enhance the low-light images and generate more visually pleasant and realistic enhanced results than the state-of-the-art methods.

The rest of this article is organized as follows. We first review some related works in Section 2. Then we present the proposed network for LLIE in Section 3. Experimental results and analysis are provided in Section 4. Conclusions are finally drawn in Section 5.

## 2. Related work

### 2.1. Retinex decomposition

In general, the Retinex model (Land, 1977) can be formulated in the following form:

$$\mathbf{O} = \mathbf{I} \odot \mathbf{R}, \quad (1)$$

where  $\mathbf{O}$ ,  $\mathbf{I}$  and  $\mathbf{R}$  respectively denote the input color image, illumination and reflectance,  $\odot$  denotes element-wise product. The general regularization model for solving this ill-posed problem is

$$\min_{\mathbf{I}, \mathbf{R}} \|\mathbf{O} - \mathbf{I} \odot \mathbf{R}\|_2^2 + \alpha \mathcal{R}_1(\mathbf{I}) + \beta \mathcal{R}_2(\mathbf{R}), \quad (2)$$

where  $\mathcal{R}_1(\mathbf{I})$  and  $\mathcal{R}_2(\mathbf{R})$  are the regularities for the illumination and reflectance components, respectively;  $\alpha$  and  $\beta$  are tuning parameters.

To obtain good estimations of illumination and reflectance, various regularization models (Guo et al., 2016; Cai et al., 2017; Xu et al., 2020) have been proposed. For example, Guo et al. (2016) adopted weighted TV (Rudin et al., 1992) to regularize the illumination. Cai et al. (2017) constrained the illumination by local variation deviation and the reflectance by TV (Rudin et al., 1992). Xu et al. (2020) employed the weighted TV- $L_2$  norm to regularize both the illumination and reflectance components, with the weights generated from the exponential local derivatives of the observed image. Although the hand-crafted regularities have shown empirical success, they are less flexible to adapt to various scenes.

Later, the Retinex-based learning methods (Wei et al., 2018; Zhang et al., 2019, 2021; Yang et al., 2021; Wu et al., 2022; Liu et al., 2022, 2021; Jiang et al., 2021b) have been developed. Early methods (Wei et al., 2018; Zhang et al., 2019, 2021; Yang et al., 2021; Jiang et al., 2021b) simply perform Retinex decomposition by pure networks, which lack interpretability. Integrating the advantages of both analytical and learning based methods, the unfolding networks (Wu et al., 2022; Liu et al., 2022, 2021) are more interpretable, which utilize implicit regularized Retinex model. Wu et al. (2022) unfolded the iterative optimization process of Eq. (2) by replacing the sub-problems corresponding to the illuminance and reflectance with two sub-networks. Liu et al. (2022, 2021) utilized the gradient descent algorithm to solve the sub-problem containing the illumination and estimated the gradient of the regularity by a network.

To reconcile the advantages of both analytical and learning based methods, we propose an explicit parametric regularization Retinex model with pixel-wise varying parameters and unroll the iterative algorithm for solving the model into an unfolding network with the parameters embedded in.

## 2.2. Retinex decomposition in HSV color space

Most existing methods (Wei et al., 2018; Zhang et al., 2019, 2021; Yang et al., 2021; Liu et al., 2022, 2021) perform Retinex decomposition in RGB space, but the enhanced images show severe color deviation. The methods (Cai et al., 2017; Xu et al., 2020; Jiang et al., 2021b) perform Retinex decomposition on the V channel in HSV space to enhance the brightness of the low-light image without changing the hue and saturation. The enhanced images show little color deviation. However, the hue and saturation of low-light images are often noisy compared to those of normal-light images (as shown in Figs. 4(e) and 4(f)), which leads to the presence of noise in the enhanced RGB images (as shown in Fig. 4(c)).

This motivates our proposed network to perform Retinex decomposition and enhancement on the V channel in HSV space, and refine the enhanced image by removing the noise.

## 3. Proposed method

### 3.1. Proposed network

Our proposed network DPRED is illustrated in Fig. 1. It consists of three modules: Retinex decomposition, enhancement and refinement. For Retinex decomposition, different from existing unfolding methods (Wu et al., 2022; Liu et al., 2021, 2022), which use implicit regularized Retinex model, we propose an explicit parametric regularized Retinex model with pixel-wise varying parameters. And we construct an unfolding network which mimics the iteration solution of our parametric regularization decomposition model with the parameters embedded in. The second module conducts enhancement on the illuminance and reflectance of V channel by two networks. To reduce color deviation, we perform the former two modules to the V channel in HSV color space, which decouples the color and the illuminance. Then

the illuminance and reflectance of V channel and the original H and S channels are used to obtain a coarse enhanced RGB image. Considering the noise hidden in the H and S channels leads to noisy coarse enhanced result, we add a refinement module to remove the noise in the coarse enhanced RGB image and output the final result.

Denote the low light image and the corresponding normal light image pair by  $(\mathbf{O}_l, \mathbf{O}_n)$ , their HSV channels by  $(\mathbf{H}_l, \mathbf{S}_l, \mathbf{V}_l)$  and  $(\mathbf{H}_n, \mathbf{S}_n, \mathbf{V}_n)$ . To avoid color deviation, we perform the Retinex decomposition and enhancement for the V channel. Let  $(\mathbf{I}_l, \mathbf{R}_l)$  and  $(\mathbf{I}_n, \mathbf{R}_n)$  denote the illuminance and reflectance of the low light image and the normal light image, we can obtain them by  $RDnet$  (the specific structure will be illustrated in Section 3.2). Then  $\mathbf{I}_l$  and  $\mathbf{R}_l$  are input to the corresponding enhancement network:  $INet$  and  $RENnet$  (illustrated in Section 3.3) respectively to obtain the enhanced illuminance  $\hat{\mathbf{I}}$  and reflectance  $\hat{\mathbf{R}}$ . Note that  $INet$  brightens the illuminance under the guidance of  $\mathbf{I}_r = \mathbf{I}_l \oslash (\mathbf{I}_n + \epsilon)$  ( $\oslash$  denotes element-wise division,  $\epsilon$  is a small positive number to avoid division by zero), and  $RENnet$  takes  $\mathbf{I}_l$  for reference since the noise in the reflectance is related to the illuminance. After that, the enhanced V channel, denoted by  $\hat{\mathbf{V}}$ , can be obtained by  $\hat{\mathbf{I}} \odot \hat{\mathbf{R}}$ . Then  $(\mathbf{H}_l, \mathbf{S}_l, \hat{\mathbf{V}})$  are converted to obtain the enhanced RGB image  $\hat{\mathbf{E}}$ .

Previous methods (Cai et al., 2017; Xu et al., 2020; Jiang et al., 2021b) simply take  $\hat{\mathbf{E}}$  as the final enhanced result, which is satisfactory for images captured under slightly low light conditions. However, for images captured under extremely low light conditions, perceivable noise exists in the  $\mathbf{H}_l$ ,  $\mathbf{S}_l$  and  $\hat{\mathbf{E}}$ . Therefore, we further add a refinement network  $Fnet$  to remove the noise in  $\hat{\mathbf{E}}$  and obtain the final enhanced result

$$\mathbf{E} = Fnet(\hat{\mathbf{E}}, \mathbf{M}), \quad (3)$$

where  $\mathbf{M}$  denotes noise map. The experimental results in Fig. 4 show this module is effective in refining the enhanced images.

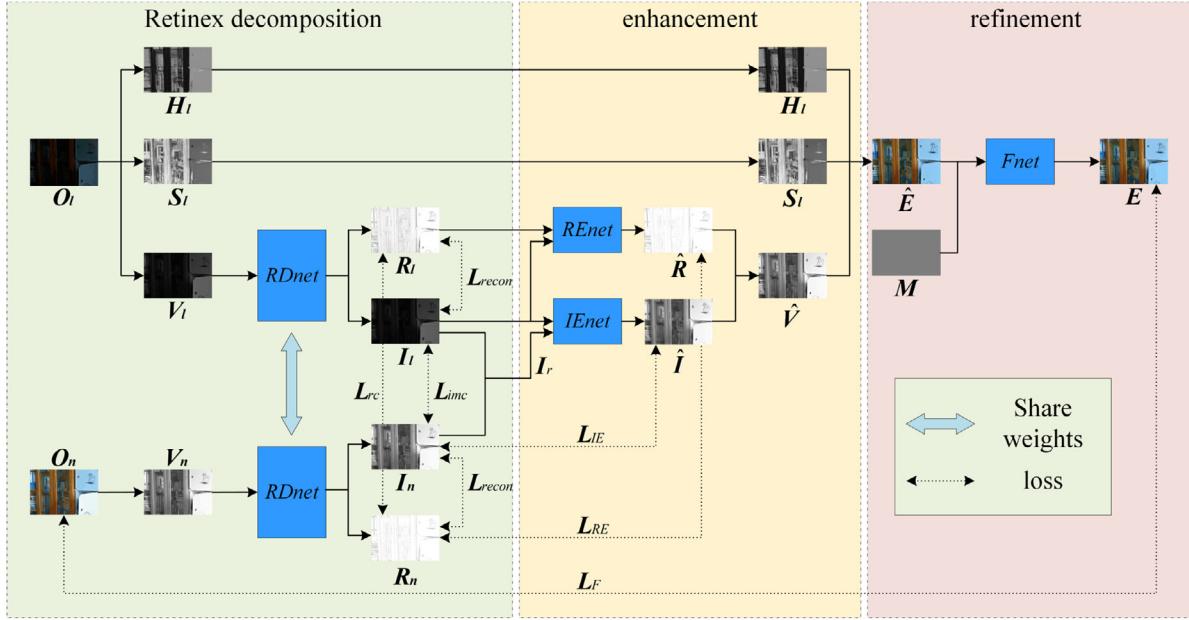
### 3.2. Parametric Retinex decomposition model

The analytical Retinex decomposition models (Guo et al., 2016; Cai et al., 2017; Xu et al., 2020), especially the regularizations lack data adaptivity, and parameters involved need manually selection. In contrast, learning methods (Wei et al., 2018; Zhang et al., 2019, 2021; Yang et al., 2021; Jiang et al., 2021b) lack interpretability. In this work, we propose a model to reconcile the advantages of analytical model and learning methods. Specifically, we present an explicit parametric regularization model with pixel-wise varying parameters for Retinex decomposition of V channel of low light color images, and these pixel-wise varying parameters are learned by a network constructed by unrolling the iterative solution of our model.

We present the following explicit parametric Retinex decomposition model:

$$\min_{\mathbf{I}} \|\mathbf{V} - \mathbf{I}\|_2^2 + \sum_{i=1}^N \lambda_i (|(\nabla_h \mathbf{I})_i|^{p_i} + |(\nabla_v \mathbf{I})_i|^{q_i}), \quad (4)$$

where  $\mathbf{V} \in \mathbb{R}^N$  is the V channel of the input image  $\mathbf{O}$ ,  $\mathbf{I} \in \mathbb{R}^N$  is the illuminance of  $\mathbf{V}$ .  $\|\cdot\|_2$  denotes the Euclidean norm of vectors and  $|\cdot|$  denotes the absolute value operator.  $|\cdot|$  also means element-wise absolute value when applied to vectors.  $\nabla_h \mathbf{I} = ((\nabla_h \mathbf{I})_1, \dots, (\nabla_h \mathbf{I})_N)^T \in \mathbb{R}^N$  and  $\nabla_v \mathbf{I} = ((\nabla_v \mathbf{I})_1, \dots, (\nabla_v \mathbf{I})_N)^T \in \mathbb{R}^N$  are the horizontal and vertical partial derivatives of  $\mathbf{I}$ .  $|(\nabla_h \mathbf{I})_i|^{p_i}$  and  $|(\nabla_v \mathbf{I})_i|^{q_i}$  are the  $p_i$ -th and  $q_i$ -th power of  $|(\nabla_h \mathbf{I})_i|$  and  $|(\nabla_v \mathbf{I})_i|$ , respectively. The first term is the data fitting term and the second term is the regularity. Our regularity term basically extends the TV (Rudin et al., 1992) by incorporating pixel-wise varying parameters  $\lambda = (\lambda_1, \dots, \lambda_N)^T \in (0, 1]^N$ ,  $p = (p_1, \dots, p_N)^T \in [0, 1]^N$  and  $q = (q_1, \dots, q_N)^T \in [0, 1]^N$ , so that it can more flexibly characterize the local regularity of an image. For better data adaption, we propose an unfolding network to learn these parameters. For simplicity, the reflectance  $\mathbf{R} \in \mathbb{R}^N$  of  $\mathbf{V}$  is computed easily by  $\mathbf{R} = \mathbf{V} \oslash \mathbf{I}$ .



**Fig. 1.** Proposed network DPRED for LLIE, consisting of three modules: Retinex decomposition, enhancement and refinement. The solid line indicates the data flow in the network. First, the low-light RGB image  $O_l$  is converted into HSV color space, then the V channel  $V_l$  is decomposed into illumination and reflectance ( $I_l, R_l$ ) by the  $RDnet$ . Then, the low-light illumination and reflectance ( $I_l, R_l$ ) are enhanced by the  $IEnet$  and  $REnet$  respectively to gain the enhanced illuminance  $\hat{I}$ , enhanced reflectance  $\hat{R}$  and enhanced V channel  $\hat{V}$ . Next, ( $H_l, S_l, \hat{V}$ ) are converted to the enhanced RGB image  $\hat{E}$ . Finally,  $\hat{E}$  is refined by the refinement network  $Fnet$ . Note that the V channel  $V_n$  of the normal light image  $O_n$  is also decomposed by  $RDnet$  to obtain ( $I_n, R_n$ ) to supervise the training of  $RDnet$ ,  $IEnet$  and  $REnet$ .

We use the iterative re-weighting strategy in [Chartrand and Yin \(2008\)](#) to solve Eq. (4) approximately. Let  $\mathbf{w}_h = ((\mathbf{w}_h)_1, \dots, (\mathbf{w}_h)_N)^T \in \mathbb{R}^N$  and  $\mathbf{w}_v = ((\mathbf{w}_v)_1, \dots, (\mathbf{w}_v)_N)^T \in \mathbb{R}^N$ , with the  $i$ th elements  $(\mathbf{w}_h)_i = \sqrt{\lambda_i / (|\nabla_h \mathbf{I}|_i^{2-p_i} + \epsilon)}$  and  $(\mathbf{w}_v)_i = \sqrt{\lambda_i / (|\nabla_v \mathbf{I}|_i^{2-q_i} + \epsilon)}$  respectively,  $\epsilon$  being a small positive number to avoid dividing by zero. Eq. (4) can be approximated by

$$\min_{\mathbf{I}} \|\mathbf{V} - \mathbf{I}\|_2^2 + \sum_{i=1}^N (|\mathbf{w}_h|_i |\nabla_h \mathbf{I}|_i^2 + |\mathbf{w}_v|_i |\nabla_v \mathbf{I}|_i^2), \quad (5)$$

or

$$\min_{\mathbf{I}} \|\mathbf{V} - \mathbf{I}\|_2^2 + \|\mathbf{w}_h \odot \nabla_h \mathbf{I}\|_2^2 + \|\mathbf{w}_v \odot \nabla_v \mathbf{I}\|_2^2. \quad (6)$$

Then we optimize Eq. (6) iteratively as follow:

$$(\mathbf{w}_h^{k+1})_i = \sqrt{\lambda_i / (|\nabla_h \mathbf{I}^k|_i^{2-p_i} + \epsilon)}, \quad (7)$$

$$(\mathbf{w}_v^{k+1})_i = \sqrt{\lambda_i / (|\nabla_v \mathbf{I}^k|_i^{2-q_i} + \epsilon)}, \quad (8)$$

$$\mathbf{I}^{k+1} = \arg \min_{\mathbf{I}} \|\mathbf{V} - \mathbf{I}\|_2^2 + \|\mathbf{w}_h^{k+1} \odot \nabla_h \mathbf{I}\|_2^2 + \|\mathbf{w}_v^{k+1} \odot \nabla_v \mathbf{I}\|_2^2. \quad (9)$$

To solve sub-problem (9), we use the Alternating Direction Method of Multipliers (ADMM) ([Boyd et al., 2011](#)) framework. Introducing two auxiliary vectors  $\mathbf{d}_h = \nabla_h \mathbf{I}$  and  $\mathbf{d}_v = \nabla_v \mathbf{I}$ , we define the corresponding augmented Lagrangian function as

$$\begin{aligned} L(\mathbf{I}, \mathbf{d}_h, \mathbf{d}_v, \mathbf{b}_h, \mathbf{b}_v) &= \|\mathbf{V} - \mathbf{I}\|_2^2 + \|\mathbf{w}_h^{k+1} \odot \mathbf{d}_h\|_2^2 + \|\mathbf{w}_v^{k+1} \odot \mathbf{d}_v\|_2^2 \\ &\quad + \mu \|\mathbf{d}_h - \nabla_h \mathbf{I} + \mathbf{b}_h\|_2^2 + \mu \|\mathbf{d}_v - \nabla_v \mathbf{I} + \mathbf{b}_v\|_2^2, \end{aligned} \quad (10)$$

where  $\mathbf{b}_h, \mathbf{b}_v$  are Lagrangian multipliers, and  $\mu$  is a penalty parameter. We update one variable by fixing others at each iteration step, and let  $\mu$  vary in each iteration. The corresponding sub-problems for  $\mathbf{I}, \mathbf{d}_h$  and  $\mathbf{d}_v$  are as follows:

$$\begin{aligned} \mathbf{I}^{t+1} &= \arg \min_{\mathbf{I}} \|\mathbf{V} - \mathbf{I}\|_2^2 \\ &\quad + \mu^{t+1} \|\mathbf{d}_h^t - \nabla_h \mathbf{I} + \mathbf{b}_h^t\|_2^2 + \mu^{t+1} \|\mathbf{d}_v^t - \nabla_v \mathbf{I} + \mathbf{b}_v^t\|_2^2, \end{aligned} \quad (11)$$

$$\mathbf{d}_h^{t+1} = \arg \min_{\mathbf{d}_h} \|\mathbf{w}_h^{k+1} \odot \mathbf{d}_h\|_2^2 + \mu^{t+1} \|\mathbf{d}_h - \nabla_h \mathbf{I}^{t+1} + \mathbf{b}_h^t\|_2^2, \quad (12)$$

$$\mathbf{d}_v^{t+1} = \arg \min_{\mathbf{d}_v} \|\mathbf{w}_v^{k+1} \odot \mathbf{d}_v\|_2^2 + \mu^{t+1} \|\mathbf{d}_v - \nabla_v \mathbf{I}^{t+1} + \mathbf{b}_v^t\|_2^2. \quad (13)$$

#### Algorithm 1 The iterative algorithm for approximate solution of Eq. (4)

**Input:** The V channel  $\mathbf{V}$ .

1: Initialize:  $\mathbf{I}^0 = \mathbf{V}, \mathbf{d}_h^0 = 0, \mathbf{d}_v^0 = 0, \mathbf{b}_h^0 = 0, \mathbf{b}_v^0 = 0, \lambda = f_\lambda(\mathbf{V}), p = f_p(\mathbf{V}), q = f_q(\mathbf{V}), [\mu^1, \dots, \mu^K] = f_\mu(\mathbf{V}), K$  (the number of iterations).

2: Output:  $\mathbf{I}^K, \mathbf{R}^K$ .

3: **for**  $k = 0 : K - 1$  **do**

4: compute  $\mathbf{w}_h^{k+1}$  and  $\mathbf{w}_v^{k+1}$  by Eq. (7) and Eq. (8) respectively;

5: update  $\mathbf{I}^{k+1}$  by Eq. (16);

6: update  $\mathbf{d}_h^{k+1}$  and  $\mathbf{d}_v^{k+1}$  by Eq. (17) and Eq. (18) respectively;

7: update  $\mathbf{b}_h^{k+1}$  and  $\mathbf{b}_v^{k+1}$  by Eq. (14) and Eq. (15) respectively;

8: **end for**

9:  $\mathbf{R}^K = \mathbf{V} \oslash \mathbf{I}^K$

The multipliers can be updated by the gradient ascending algorithm:

$$\mathbf{b}_h^{t+1} = \mathbf{b}_h^t + \mathbf{d}_h^{t+1} - \nabla_h \mathbf{I}^{t+1}, \quad (14)$$

$$\mathbf{b}_v^{t+1} = \mathbf{b}_v^t + \mathbf{d}_v^{t+1} - \nabla_v \mathbf{I}^{t+1}. \quad (15)$$

The solution of problem (11) can be obtained by

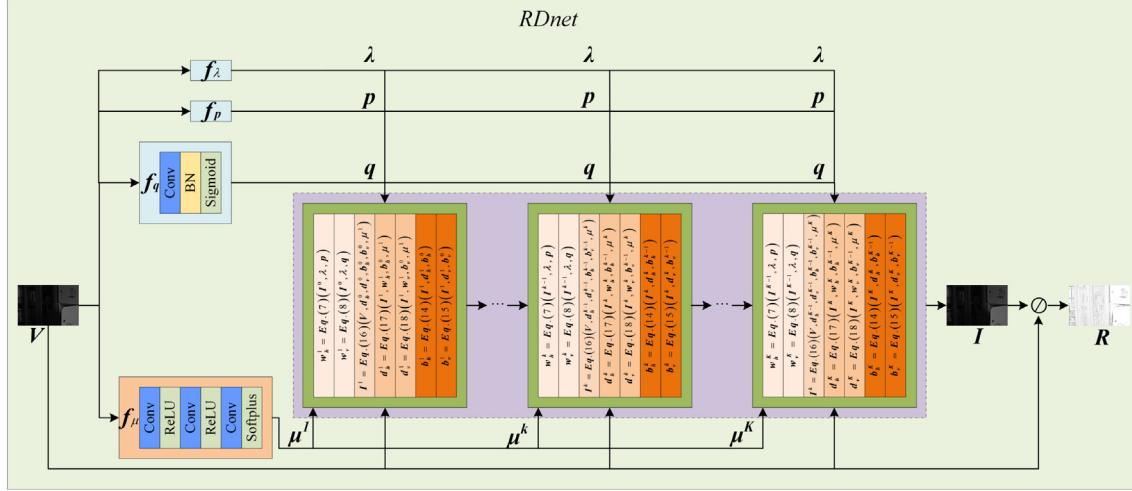
$$\mathbf{I} = \mathcal{F}^{-1} \left( \frac{\mathcal{F}(\mathbf{V}) + \mu^{t+1} \left( \overline{\mathcal{F}(\nabla_h)} (\mathbf{d}_h^t + \mathbf{b}_h^t) + \overline{\mathcal{F}(\nabla_v)} (\mathbf{d}_v^t + \mathbf{b}_v^t) \right)}{1 + \mu^{t+1} \left( \overline{\mathcal{F}(\nabla_h)} \mathcal{F}(\nabla_h) + \overline{\mathcal{F}(\nabla_v)} \mathcal{F}(\nabla_v) \right)} \right), \quad (16)$$

where  $\mathcal{F}(\cdot)$  and  $\mathcal{F}^{-1}(\cdot)$  denote the Fast Fourier Transform (FFT) and inverse FFT,  $\overline{\mathcal{F}(\cdot)}$  denotes the complex conjugate of  $\mathcal{F}(\cdot)$ . Problems (12) and (13) have the following closed-form solution, respectively:

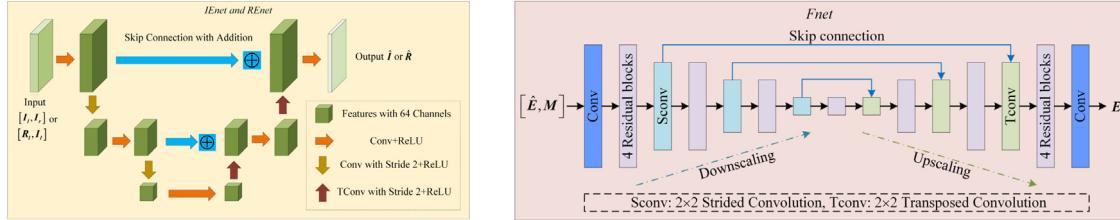
$$\mathbf{d}_h^{t+1} = (\mu^{t+1} (\nabla_h \mathbf{I}^{t+1} - \mathbf{b}_h^t)) \oslash (\mathbf{w}_h^{k+1} \odot \mathbf{w}_h^{k+1} + \mu^{t+1}), \quad (17)$$

$$\mathbf{d}_v^{t+1} = (\mu^{t+1} (\nabla_v \mathbf{I}^{t+1} - \mathbf{b}_v^t)) \oslash (\mathbf{w}_v^{k+1} \odot \mathbf{w}_v^{k+1} + \mu^{t+1}). \quad (18)$$

For simplicity, the iteration in the inner loop is performed only once. The whole procedure is summarized in Algorithm 1.



(a) The unfolding Retinex decomposition network: *RDnet*. The  $V$  channel is input to the parameter estimation modules  $f_\lambda$ ,  $f_p$ ,  $f_q$  and  $f_\mu$  to learn the parameters  $\lambda$ ,  $p$ ,  $q$  and  $\mu$ . Then these learned parameters are substituted into the iterations to gain the illumination  $I$ . The reflectance is obtained by  $R = V \oslash I$ .



(b) The Enhancement networks: *IEnet* and *REnet*, which enhance the low-light illumination and reflectance  $I_l$  and  $R_l$  respectively to obtain  $(\hat{I}, \hat{R})$ . *IEnet* brightens the illuminance under the guidance of  $I_r$ . *REnet* takes  $I_l$  for reference since the noise in the reflectance is related to the illuminance.

(c) The refinement network: *Fnet*, which removes the noise in the coarsely enhanced image  $\hat{E}$  to obtain the final enhanced image  $E$ ;  $M$  denotes noise map.

Fig. 2. The network architecture of *RDnet*, *IEnet*, *REnet* and *Fnet*.

The model parameters  $\lambda, p, q$  and the penalty parameters  $\mu = (\mu^1, \dots, \mu^K)^T$  in the algorithm will be predicted by using their corresponding modules, denoted by

$$\lambda = f_\lambda(V), p = f_p(V), q = f_q(V), \mu = f_\mu(V). \quad (19)$$

We unfold the iterations in the algorithm above into a network, with the parameter prediction modules embedded in. For  $\lambda, p$  and  $q$ , we use the same simple architecture: Conv+BN+Sigmoid to predict their values. For  $\mu$ , we use two Conv+ReLU layers and a Conv+Softplus layer. The whole Retinex decomposition network, called *RDnet*, is shown in Fig. 2(a).

### 3.3. Enhancement sub-networks and refinement sub-network

For the enhancement networks *IEnet* and *REnet*, we design a U-Net (Ronneberger et al., 2015)-like architecture, as shown in Fig. 2(b). Note that the number of feature channels is fixed as 64. And the concatenation of features in the vanilla U-Net (Ronneberger et al., 2015) is replaced by the addition operator.

The DRUNet (Zhang et al., 2022) is a state-of-the-art image denoising network, so we use it as our *Fnet* to remove the noises in the coarsely enhanced RGB image. The details are illustrated in Fig. 2(c).

### 3.4. Loss functions

The loss function  $\mathcal{L}_{RD}$  for *RDnet* consists of three terms: the reconstruction loss  $\mathcal{L}_{recon}$ , the reflectance consistency loss  $\mathcal{L}_{rc}$  and the illumination mutual consistency loss  $\mathcal{L}_{imc}$ :

$$\mathcal{L}_{RD} = \mathcal{L}_{recon} + \lambda_{rc} \mathcal{L}_{rc} + \lambda_{imc} \mathcal{L}_{imc}, \quad (20)$$

where

$$\mathcal{L}_{recon} = \|V_l - I_l \odot R_l\|_1 + \|V_n - I_n \odot R_n\|_1, \quad (21)$$

$$\mathcal{L}_{rc} = \|R_l - R_n\|_1, \text{ and} \quad (22)$$

$$\mathcal{L}_{imc} = \|\mathbf{M}_I \odot \exp(-10\mathbf{M}_I)\|_1, \mathbf{M}_I = |\nabla I_l| + |\nabla I_n|. \quad (23)$$

$\lambda_{rc}$  and  $\lambda_{imc}$  are tuning parameters.  $\|\cdot\|_1$  denotes the  $l_1$  norm of vectors,  $\exp(\cdot)$  is the element-wise exponential operator for vectors and  $\nabla : \mathbb{R}^N \rightarrow \mathbb{R}^N \times \mathbb{R}^N$  stands for the gradient operator.  $\mathcal{L}_{recon}$  is used to constrain the reconstruction error of the Retinex decomposition model. Based on the Retinex theory (Land, 1977),  $(V_l, V_n)$  should have the same reflectance, thus  $\mathcal{L}_{rc}$  aims to minimize the discrepancy between  $R_l$  and  $R_n$ .  $\mathcal{L}_{imc}$  maintains the mutual consistency of the structure of the low/normal light illuminance, which is introduced by KinD (Zhang et al., 2019).

The loss  $\mathcal{L}_{IE}$  for  $IEnet$  is

$$\mathcal{L}_{IE} = \|\hat{\mathbf{I}} - \mathbf{I}_n\|_1 + \|\mathbf{V}_n - \hat{\mathbf{I}} \odot \mathbf{R}_n\|_1 + \||\nabla \hat{\mathbf{I}}| - |\nabla \mathbf{I}_n|\|_2^2. \quad (24)$$

The former two terms are the fidelity terms between the enhanced illumination  $\hat{\mathbf{I}}$  and that of the normal-light one  $\mathbf{I}_n$ . The third term measures the distance between the absolute value of the gradient of  $\hat{\mathbf{I}}$  and  $\mathbf{I}_n$ .

Similarly, the loss  $\mathcal{L}_{RE}$  for  $REnet$  is

$$\mathcal{L}_{RE} = \|\hat{\mathbf{R}} - \mathbf{R}_n\|_1 + \|\mathbf{V}_n - \mathbf{I}_n \odot \hat{\mathbf{R}}\|_1 + \||\nabla \hat{\mathbf{R}}| - |\nabla \mathbf{R}_n|\|_2^2. \quad (25)$$

The loss  $\mathcal{L}_F$  of  $Fnet$  is

$$\mathcal{L}_F = \|\mathbf{E} - \mathbf{O}_n\|_2^2 + 1 - SSIM(\mathbf{E}, \mathbf{O}_n). \quad (26)$$

The first term minimizes the Mean Squared Error (MSE) between the enhanced result  $\mathbf{E}$  and normal light image  $\mathbf{O}_n$ . The last two terms maximize the Structural Similarity (SSIM) index distance (Wang et al., 2004) between  $\mathbf{E}$  and  $\mathbf{O}_n$ .

#### 4. Experiments and discussion

In this section, we analyze our method empirically and verify its performance on datasets with/without ground-truth. In Section 4.1, we describe the implementation details of the proposed method. In Section 4.2, we discuss some notable aspects of the proposed method. First, we analyze the learned parameters in  $RDnet$ . Then we conduct four ablation studies to analyze the effect of our model. In Section 4.3, we evaluate the performance of our method on two datasets with ground-truth: LOL dataset (Wei et al., 2018) and MIT-Adobe 5K dataset (Bychkovsky et al., 2011), and four datasets without ground-truth: LIME (Guo et al., 2016), NPE (Wang et al., 2013), MEF (Ma et al., 2015) and DICM (Lee et al., 2013). The LOL dataset (Wei et al., 2018) contains 500 low/normal light image pairs, with 485 paired images for training and 15 paired images for testing. The MIT-Adobe 5K dataset (Bychkovsky et al., 2011) contains 5000 low-light/reference image pairs, and we randomly select 500 image pairs for evaluation. The non-reference datasets (Guo et al., 2016; Wang et al., 2013; Ma et al., 2015; Lee et al., 2013) contain 10, 84, 17, 64 images respectively. Note that the proposed method is trained only on LOL dataset (Wei et al., 2018), and tested on other datasets without retraining or fine-tuning to validate its generalization performance. In Section 4.4, we provide the limitations of our model.

We compare the proposed method with state-of-the-art learning-based methods, including RetinexNet (Wei et al., 2018), KinD (Zhang et al., 2019), KinD++ (Zhang et al., 2021), URRetinex-Net (Wu et al., 2022), LLFlow (Wang et al., 2022), Retinexformer (Cai et al., 2023), RQ-LLIE (Liu et al., 2023), EnlightenGAN (Jiang et al., 2021a), RUAS (Liu et al., 2021; Jiang et al., 2021b), Zero-DCE (Guo et al., 2020) and Zero-DCE++ (Li et al., 2022). The former seven supervised learning methods have been trained on LOL dataset (Wei et al., 2018), such that we utilize their provided models for evaluation. The latter five methods are unsupervised, we also use their provided models for evaluation.

We adopt the well-known Peak Signal-to-Noise Ratio (PSNR), SSIM (Wang et al., 2004), Learned Perceptual Image Patch Similarity (LPIPS) (Zhang et al., 2018), DeltaE (McLaren, 1976) and Lightness-Order-Error (LOE) (Wang et al., 2013) for quantitative comparison. PSNR, SSIM and LPIPS measure similarity between the enhanced image and ground-truth, while DeltaE measures the color difference between the enhanced image and ground-truth. LOE (Wang et al., 2013) is a metric specifically designed to evaluate the performance of LLIE methods. It is a general image quality metric without using the ground-truth. However, in its original definition, the reference image used to compute this metric is chosen as the input low-light image. When the ground truth is available, a more appropriate modification (referred to as  $LOE_{ref}$ ) can be made to LOE (Wang et al., 2013) by replacing the low-light image with its corresponding normal-light image as the Zhang et al. (2019). Higher values of PSNR and SSIM indicate better quality while lower values of LPIPS, DelatE, LOE and  $LOE_{ref}$  indicate better quality.

#### 4.1. Implementation details

We implement our method with PyTorch on a NVIDIA GTX 1660 GPU. For every network in the proposed method, the batch size is set to 10, the patch size is  $48 \times 48$ , and the learning rate is set to  $10^{-4}$ , except for the refinement network where the learning rate is specifically set to  $10^{-5}$ . We train our method for 100 epochs and evaluate it every 10 epochs. We use Adam optimizer (Kingma and Ba, 2015) with default parameters. For the decomposition network, the learning rate is adjusted every 10 epochs. For other networks, the learning rates are fixed at the training stage. The number of iterations  $K$  in  $RDnet$  is empirically set to 10.

Due to the complexity of the network architecture, training the entire network with an end-to-end strategy is challenging. Therefore, we adopt a strategy of training each sub-network individually with others being fixed. The Retinex decomposition network  $RDnet$  is trained by optimizing the loss  $\mathcal{L}_{RD}$  in Eq. (20). The parameters  $\lambda_{rc}$  and  $\lambda_{imc}$  in Eq. (20) are all empirically set to 0.1. The enhancement networks  $IEnet$  and  $REnet$  are trained by minimizing the losses  $\mathcal{L}_{IE}$  in Eq. (24) and  $\mathcal{L}_{RE}$  in Eq. (25), respectively. Note that we use  $\mathbf{I}_n^{1/\gamma}$  instead of  $\mathbf{I}_n$  as input to  $IEnet$  at the test stage for the non-reference datasets (Guo et al., 2016; Wang et al., 2013; Ma et al., 2015; Lee et al., 2013),  $\gamma$  is empirically set to 3 in our experiments. Since the noise distribution is unknown, we initialize the refinement network with the pre-trained parameters of DRUNet (Zhang et al., 2022), which is capable of handling various types of noise. Subsequently, we fine-tune the parameters of  $Fnet$  using the loss  $\mathcal{L}_F$  in Eq. (26) to better adapt to the dataset. Note that the input of  $Fnet$  is the coarsely enhanced RGB image  $\hat{\mathbf{E}}$  and a manually set noise level map  $\mathbf{M}$ . At the training stage, we set the noise level as 15. At the test stage, it is empirically set to 15 for the LOL dataset (Wei et al., 2018), 3 for the MIT-Adobe 5K dataset (Bychkovsky et al., 2011), and 5 for the non-reference datasets (Guo et al., 2016; Wang et al., 2013; Ma et al., 2015; Lee et al., 2013). Besides, we use the corresponding normal light image as the ground-truth to train  $Fnet$ .

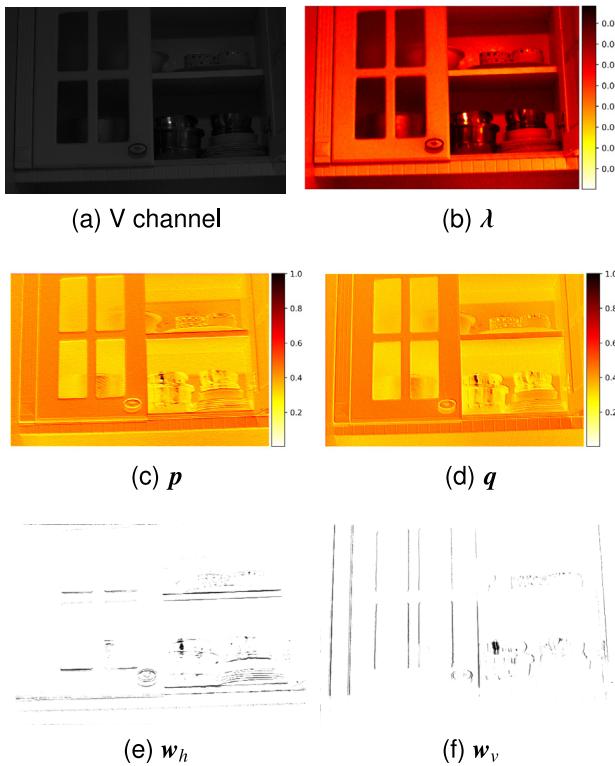
#### 4.2. Model analysis and ablation study

##### 4.2.1. The learned parameters by $RDnet$

One of our main contribution is the parametric regularization model for Retinex decomposition, with parameters learned through the sub-network  $RDnet$ . Our experiment shows that the learned parameters are data adaptive and explainable. In Fig. 3, we show the V channel  $\mathbf{V}$ , the learned parameters  $\lambda, p, q$ , and the weights  $\mathbf{w}_h, \mathbf{w}_v$  of a sample image from LOL dataset (Wei et al., 2018). It can be observed that, the learned parameters are pixel-wise adaptive to the input image, and they characterize local features of the image.

First, we note that the value of  $\lambda$  is small in lighted areas while large in dark areas. As expected, such a  $\lambda$  favors estimation of illuminance by our model Eq. (4). In lighted areas,  $\lambda$  is small, the data fidelity term dominates the objective function so that the estimated illuminance tends to approach the light strength of V channel; while in dark areas,  $\lambda$  is large, which amplifies the role of the regularity term in the objective function so that the estimated illuminance tends to retreat from the light strength of V channel, but preserves local geometric better.

Surprisingly, the values of  $p$  and  $q$  seem dependent on the light strength of  $\mathbf{V}$  more heavily than the gradients of  $\mathbf{V}$ . In general, the values of the exponents  $p$  and  $q$  are large in lighted areas while small in dark areas. Such  $p$  and  $q$  also facilitate estimation of illuminance, but taking a contrary position, compared with  $\lambda$ . In particular, in lighted areas,  $p$  and  $q$  take large values, which amplify the role of the regularity, so that the estimated illuminance can preserve local geometric structure well; while in dark areas,  $p$  and  $q$  take small values, which attenuate the role of the regularity, so that the estimated illuminance tends to approach the light strength of V channel. Moreover, we also observe differences between  $p$  and  $q$ . Specifically,  $p$  characterizes more information about the horizontal edges, while  $q$  characterizes more information



**Fig. 3.** Visualization of the V channel, the learned pixel-varying parameters  $\lambda, p, q$  in  $RDnet$ , the weights  $w_h, w_v$  of a sample image in LOL dataset (Wei et al., 2018). The learned parameters  $\lambda, p$  and  $q$  are obtained by  $f_\lambda(V), f_p(V)$  and  $f_q(V)$  respectively. The weights  $w_h, w_v$  are computed by Eqs. (7) and (8) respectively.

**Table 1**

Average metrics over 15 test images from LOL dataset (Wei et al., 2018). The best results are in bold.

Method	PSNR	SSIM
V1	25.5147	0.8169
V2	25.8684	0.8424
V3	18.7066	0.7663
DPRED	<b>26.0414</b>	<b>0.8506</b>
DPRED (without $IEnet$ )	7.9046	0.2050
DPRED (without $REnet$ )	24.6108	0.8073
DPRED (without $Fnet$ )	22.5753	0.5390
JieP (Cai et al., 2017)	17.2813	0.5172
STAR (Xu et al., 2020)	17.4133	0.4946
Jiang et al. (2021b)	16.8629	0.4827
DPRED-RGB	25.2860	0.7256

about the vertical edges. This confirms that using individual exponent for the horizontal derivative and the vertical derivative in the regularity is meaningful. As expected, the weight  $w_h$  represents horizontal edges while the weight  $w_v$  represents vertical edges.

#### 4.2.2. Ablation study

In this section, we conduct ablation studies to analyze the effect of our model. First, we show the advantages of the  $RDnet$  by comparing it with related Retinex decomposition networks. Second, the effects of the modules  $IEnet$  and  $REnet$  are empirically analyzed. Third, we study the effect of the refinement module  $Fnet$ . Finally, we compare our DPRED with its variant DPRED-RGB. Experiments are conducted on LOL dataset (Wei et al., 2018) and the results are reported in Table 1.

**Advantages of the parametric Retinex decomposition module  $RDnet$ .** To show the advantages of our parametric Retinex decomposition module  $RDnet$ , we compare our network with its three variants, all based on our network DPRED shown in Fig. 1, different only in

the Retinex decomposition module. The first variant, called V1, is a special case of our method with  $p = q = 1$ . This special case aims to mimic the TV regularization (Rudin et al., 1992) of the illuminance, with the parameters  $\lambda$  and  $\mu$  learned by using our network. The second variant V2 replaces our  $RDnet$  by the architecture in URetinex-Net (Wu et al., 2022), which unfolds the iterative optimization process of Eq. (2) by replacing the sub-problems corresponding to the illuminance and reflectance with two sub-networks. The third variant V3 replaces our  $RDnet$  by the architecture in Liu et al. (2022, 2021), which utilizes gradient descent algorithm to solve the sub-problem involving the illuminance and estimates the gradient of the regularity by a network. Note that our  $RDnet$  and V1 employ analytical regularities with pixel-varying parameters, while V2 and V3 utilize implicit regularities. We adopt the same training settings (described in Section 4.1) for the variants as for the proposed method. The average PSNRs and SSIMs over 15 test images are reported in Table 1. With the proposed  $RDnet$ , our whole network (the fourth row in Table 1) obtains the best metrics. This demonstrates that, with learned data adaptive parameters, our analytical regularity parameters favor in the final results.

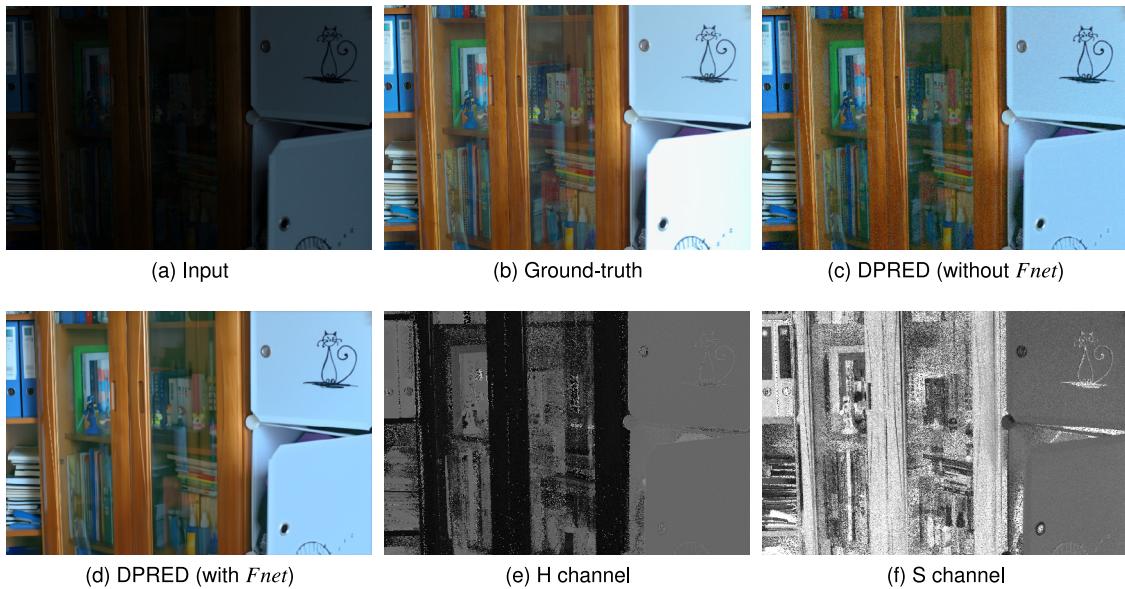
**Effects of the enhancement modules  $IEnet$  and  $REnet$ .** The fifth and the sixth rows of Table 1 respectively present the average PSNRs and SSIMs over 15 test images obtained by using our network in two cases: one without  $IEnet$  and the other without  $REnet$ . Without  $IEnet$ , the metrics of the final enhanced images degenerate severely (compared with the metrics obtained by our whole network), verifying the essential role of this module in our network. On the other hand, without  $REnet$ , the metrics also degenerate, but not so quickly. This indicates that  $REnet$  is also important in facilitating the final result.

**Effect of the refinement module  $Fnet$ .** To study the effect of the refinement module  $Fnet$ , we conduct test of our DPRED over the 15 test images by dropping  $Fnet$ . The average metrics are presented in the seventh row of Table 1. The final enhanced color image of a sample test image is shown in Fig. 4. It can be observed that, without the refinement module, the final enhanced image contains perceptible noise, as depicted in Fig. 4(c). The noise mainly comes from H and S channels, as shown in Fig. 4(e) and 4(f). By using the refinement module  $Fnet$ , our proposed network can successfully remove noise and obtains higher performance. This indicates the necessity of the refinement module in our network. We also report the results obtained by JieP (Cai et al., 2017), STAR (Xu et al., 2020), and Jiang et al. (2021b) in rows 8, 9, and 10 for comparison. Without  $Fnet$ , our method achieves superior metrics.

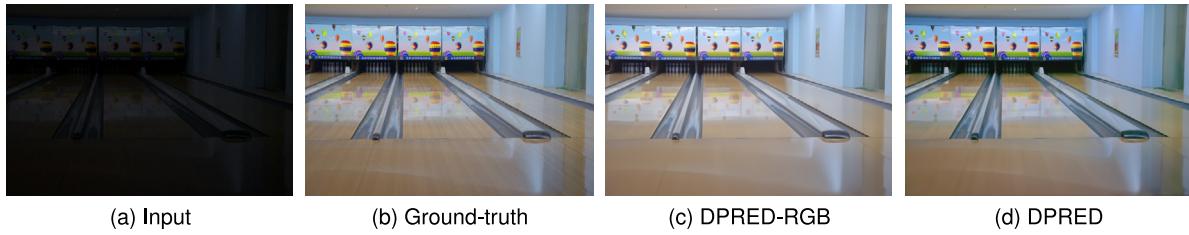
**DPRED vs. DPRED-RGB.** This experiment compares our method with a variant DPRED-RGB, which is obtained by modifying the input of  $RDnet$ ,  $IEnet$  and  $REnet$  from one channel into three channels so that it can be applied directly to RGB images. DPRED-RGB is trained under the same training settings as we train DPRED. The metrics are reported in the last row of Table 1. Visual comparison on a test image from LOL dataset (Wei et al., 2018) is shown in Fig. 5. The enhanced images look almost the same, except that the image obtained by DPRED-RGB is a little brighter. However, the metrics obtained by our DPRED are significantly better than that obtained by DPRED-RGB.

#### 4.3. Evaluation of the proposed method

In this subsection, we evaluate the performance of our method on the dataset LOL (Wei et al., 2018). Additionally, we assess the generalization performance of our method on other datasets. Furthermore, we compare the proposed method with some state-of-the-art methods. Moreover, we compare the computational complexity between the proposed method and the state-of-the-art methods.



**Fig. 4.** Effect of the refinement module *Fnet* on the final enhanced result of a test image from LOL dataset (Wei et al., 2018). (a) The low-light input (7.22 dB); (b) Ground-truth (c) Enhanced RGB image by the proposed network without *Fnet* (19.80 dB); (d) Enhanced RGB image by the proposed network with *Fnet* (21.64 dB); (e) and (f) H and S channels of the low-light input.



**Fig. 5.** Visual comparison of our DPRED and DPRED-RGB. (a) The low-light input (PSNR=6.75 dB, SSIM=0.1848); (b) Ground-truth (c) Enhanced RGB image by DPRED-RGB (PSNR=21.93 dB, SSIM=0.8941); (d) Enhanced RGB image by DPRED (PSNR=28.38 dB, SSIM=0.9082).

**Table 2**

Average metrics over 15 test images from LOL dataset (Wei et al., 2018). The best results are in bold and the second best results are underlined.

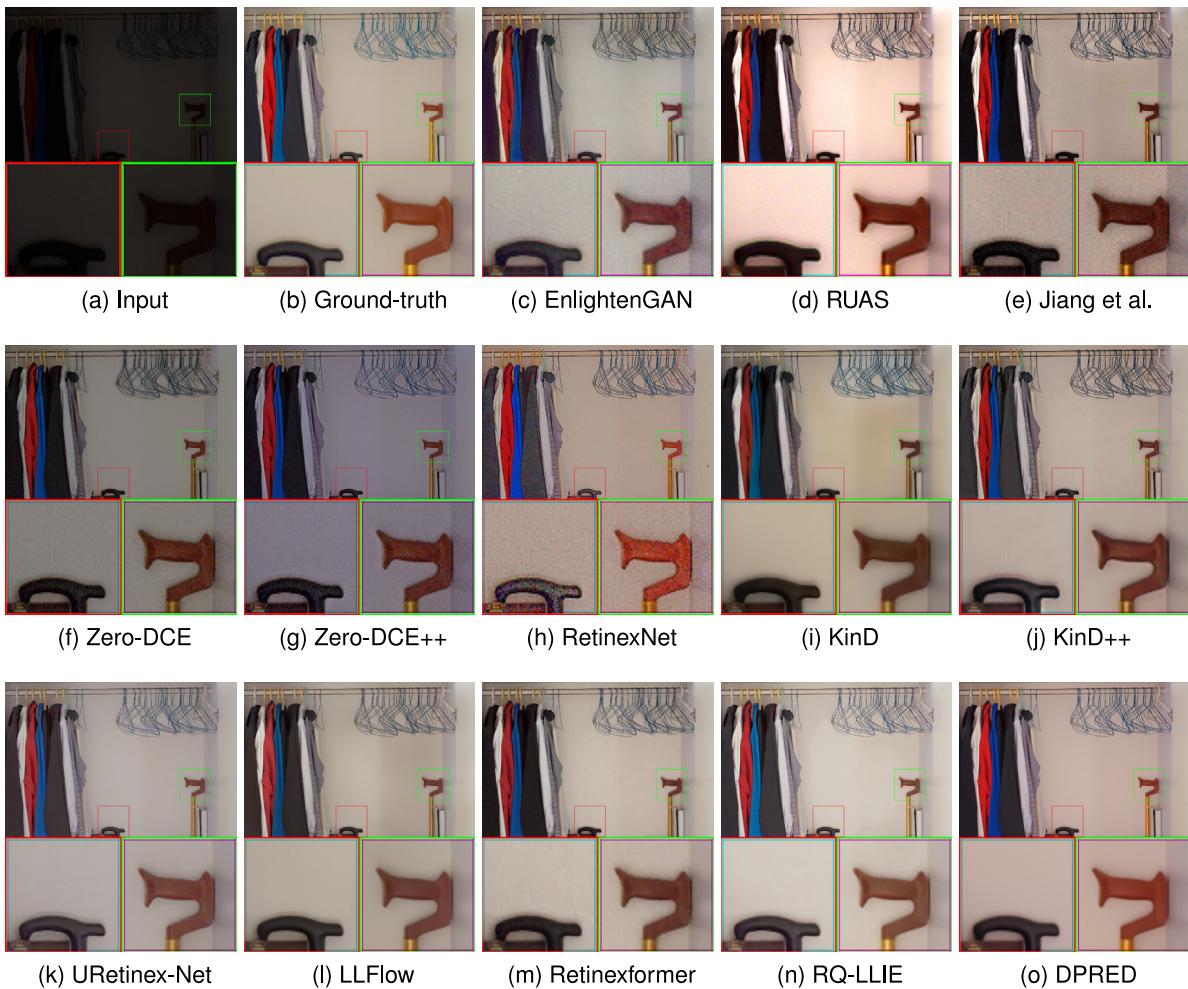
Method	PSNR	SSIM	LPIPS	DeltaE	LOE	LOE <sub>ref</sub>
EnlightenGAN (Jiang et al., 2021a)	17.4829	0.6542	0.3903	10.3047	355.3146	395.5841
RUAS (Liu et al., 2021)	16.4462	0.5048	0.3818	10.4141	<b>93.2858</b>	218.8997
Jiang et al. (2021b)	16.8629	0.4827	0.3997	10.3984	328.1878	405.2548
Zero-DCE (Guo et al., 2020)	14.8607	0.5615	0.3852	10.4453	190.7118	272.5778
Zero-DCE++ (Li et al., 2022)	15.1416	0.5647	0.3870	10.4609	168.3146	256.5977
RetinexNet (Wei et al., 2018)	16.7740	0.4285	0.4667	10.5000	475.0293	460.1811
KinD (Zhang et al., 2019)	20.3792	0.8302	0.2711	9.7109	208.7519	276.4276
KinD++ (Zhang et al., 2021)	21.8035	0.8357	0.2889	9.5625	228.4131	280.4782
URetinex-Net (Wu et al., 2022)	21.3282	0.8355	0.2384	10.0391	<u>110.9729</u>	197.0188
LLFlow (Wang et al., 2022)	24.9987	<b>0.8731</b>	<b>0.2244</b>	8.8594	277.5963	309.2483
Retinexformer (Cai et al., 2023)	<b>27.1825</b>	0.8500	0.2530	<b>7.8555</b>	122.1203	<u>162.8535</u>
RQ-LLIE (Liu et al., 2023)	23.1596	<u>0.8510</u>	<u>0.2327</u>	9.6953	148.1233	177.8530
DPRED	<u>26.0414</u>	0.8506	0.3680	<u>8.1719</u>	214.8746	<b>88.1509</b>

#### 4.3.1. Evaluation on LOL (Wei et al., 2018)

The dataset LOL (Wei et al., 2018) contains 500 low/normal light image pairs. We use 485 paired images for training and 15 paired images for testing. The metrics are reported in Table 2. We also provide a visual comparison of an image in Fig. 6. From the results, we can draw the following analysis.

The metrics in Table 2 indicate that, most of the supervised methods perform better than the unsupervised and zero-shot methods on the dataset LOL (Wei et al., 2018). Retinexformer (Cai et al., 2023) achieves the best PSNR and DeltaE, closely followed by our DPRED; LLFlow (Wang et al., 2022) achieves the best SSIM and LPIPS, also closely followed by our DPRED; however, the LOE<sub>ref</sub> index of our method is far beyond that of other methods. As shown in Fig. 6, there

exists perceivable noise in the enhanced results of the unsupervised and zero-shot methods, thus the metrics of these methods fall behind. Among the supervised methods, RetinexNet (Wei et al., 2018) retains noise in the final enhanced images; while KinD (Zhang et al., 2019), KinD++ (Zhang et al., 2021), URetinex-Net (Wu et al., 2022), LLFlow (Wang et al., 2022), Retinexformer (Cai et al., 2023), RQ-LLIE (Liu et al., 2023) and our method suppress noise well. Comparing the results obtained by KinD (Zhang et al., 2019), KinD++ (Zhang et al., 2021), URetinex-Net (Wu et al., 2022), LLFlow (Wang et al., 2022), Retinexformer (Cai et al., 2023), RQ-LLIE (Liu et al., 2023) and our method, it is easy to see that, our result has better fidelity in brightness and colors. In all, with the parametric Retinex decomposition module and the enhancement modules, the proposed network can well recover



**Fig. 6.** Visual comparison of an image from the LOL dataset (Wei et al., 2018). The objects in color boxes are zoomed in for better perception. (a) Input (6.63 dB); (b) Ground-truth; (c) EnlightenGAN (Jiang et al., 2021a) (24.62 dB); (d) RUAS (Liu et al., 2021) (16.24 dB); (e) (Jiang et al., 2021b) (21.16 dB); (f) Zero-DCE (Guo et al., 2020) (15.48 dB); (g) Zero-DCE++ (Li et al., 2022) (15.57 dB); (h) RetinexNet (Wei et al., 2018) (21.78 dB); (i) KinD (Zhang et al., 2019) (23.30 dB); (j) KinD++ (Zhang et al., 2021) (25.02 dB); (k) URetinex-Net (Wu et al., 2022) (26.94 dB); (l) LLFlow (Wang et al., 2022) (26.41 dB); (m) Retinexformer (Cai et al., 2023) (25.53 dB); (n) RQ-LLIE (Liu et al., 2023) (21.05 dB); (o) DPRED (28.57 dB).

**Table 3**

Average metrics over 500 test images from MIT-Adobe 5K dataset (Bychkovsky et al., 2011). The best results are in bold and the second best results are underlined.

Method	PSNR	SSIM	LPIPS	DeltaE	LOE	LOE <sub>ref</sub>
EnlightenGAN (Jiang et al., 2021a)	15.6214	0.7403	0.2064	10.2969	552.1799	549.2273
RUAS (Liu et al., 2021)	7.7616	0.4675	0.4999	10.3984	825.8669	816.9359
Jiang et al. (2021b)	15.4296	0.7238	0.2246	10.0625	688.6929	689.0131
Zero-DCE (Guo et al., 2020)	15.4630	0.7011	0.2805	10.3984	365.0499	355.1534
Zero-DCE++ (Li et al., 2022)	14.9995	0.6952	0.3107	10.5391	328.3433	333.8250
RetinexNet (Wei et al., 2018)	12.3053	0.6318	0.3027	10.1953	1377.3805	1352.7196
KinD (Zhang et al., 2019)	21.2163	0.8042	0.2292	9.8203	212.8142	225.9808
KinD+ (Zhang et al., 2021)	<u>21.6492</u>	<u>0.8188</u>	<u>0.1994</u>	9.8828	<u>140.2021</u>	<u>154.8203</u>
URetinex-Net (Wu et al., 2022)	13.6056	0.6876	0.2316	10.2344	355.6524	345.4250
LLFlow (Wang et al., 2022)	19.7882	0.7806	<u>0.2016</u>	10.0000	303.9430	3180.595
Retinexformer (Cai et al., 2023)	18.7697	0.7361	0.2404	10.1172	210.2383	223.9037
RQ-LLIE (Liu et al., 2023)	13.9310	0.7013	0.2290	10.2812	152.9560	161.1115
DPRED	<b>24.4446</b>	<u>0.8336</u>	0.2317	<u>8.0859</u>	<u>118.7620</u>	<u>61.3503</u>

the brightness and colors of low light images; with the additional refinement module, the proposed network can well suppress noise in low light images.

#### 4.3.2. Generalization performance

To validate the generalization performance of the proposed method, we first test our method on 500 images from MIT-Adobe 5K dataset (Bychkovsky et al., 2011) without retraining or fine-tuning. Table 3 presents the average metrics obtained by all methods. Fig. 7 shows

visual comparison of a test image. The metrics in Table 3 show that the proposed method significantly outperforms previous methods again in terms of PSNR, SSIM, DeltaE, LOE and LOE<sub>ref</sub>. Compared with the ground-truth normal light image, the enhanced images by the unsupervised and zero-shot methods have obvious color deviation or distortions. Among the supervised methods, RetinexNet (Wei et al., 2018) and URetinex-Net (Wu et al., 2022) cause color deviation and halo artifacts around the letters; KinD (Zhang et al., 2019), KinD++ (Zhang et al., 2021) and RQ-LLIE (Liu et al., 2023) change the



**Fig. 7.** Visual comparison of an image from the MIT-Adobe 5K dataset (Bychkovsky et al., 2011). The objects in color boxes are zoomed in for better perception. (a) Input(10.04 dB); (b) Ground-truth; (c) EnlightenGAN (Jiang et al., 2021a) (13.96 dB); (d) RUAS (Liu et al., 2021) (8.99 dB); (e) (Jiang et al., 2021b) (14.39 dB); (f) Zero-DCE (Guo et al., 2020) (13.36 dB); (g) Zero-DCE++ (Li et al., 2022) (13.32 dB); (h) RetinexNet (Wei et al., 2018) (10.27 dB); (i) KinD (Zhang et al., 2019) (18.80 dB); (j) KinD++ (Zhang et al., 2021) (18.79 dB); (k) URetinex-Net (Wu et al., 2022) (11.24 dB); (l) LLFlow (Wang et al., 2022) (15.43 dB); (m) Retinexformer (Cai et al., 2023) (15.85 dB); (n) RQ-LLIE (Liu et al., 2023) (13.25 dB); (o)DPRED (21.39 dB).

**Table 4**

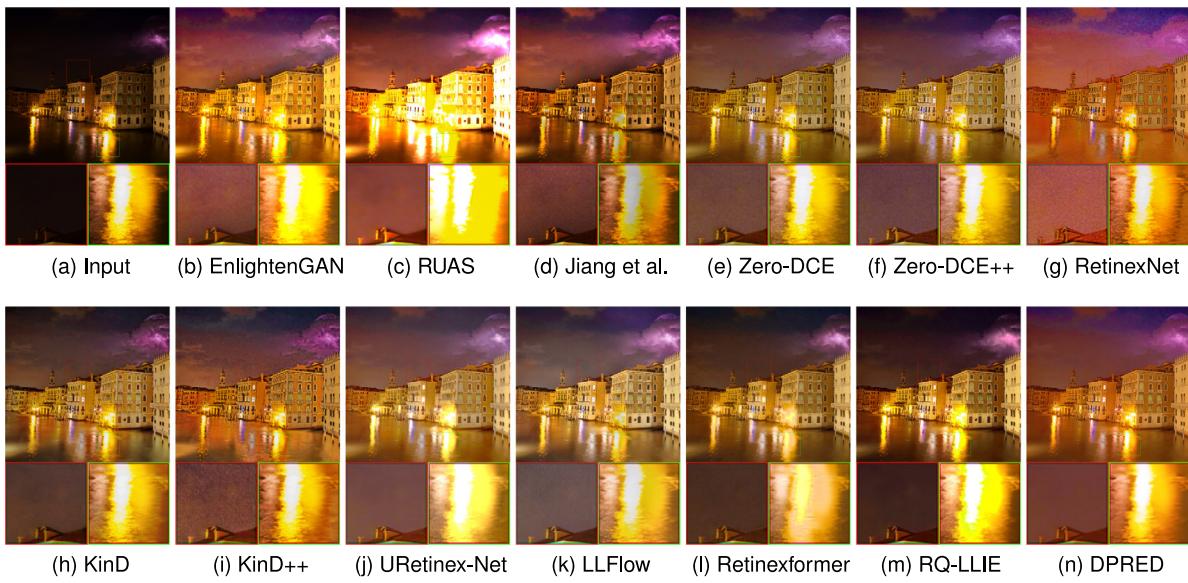
Average LOE (Guo et al., 2016) metrics of non-reference datasets (Guo et al., 2016; Wang et al., 2013; Ma et al., 2015; Lee et al., 2013). The best results are in bold and the second best results are underlined.

Method	LIME (Guo et al., 2016)	NPE (Wang et al., 2013)	MEF (Ma et al., 2015)	DICM (Lee et al., 2013)
EnlightenGAN (Jiang et al., 2021a)	396.4410	494.7369	416.8930	483.6156
RUAS (Liu et al., 2021)	267.8673	702.4092	292.4628	845.5868
Jiang et al. (2021b)	512.4983	707.5117	479.7571	644.1845
Zero-DCE (Guo et al., 2020)	175.7179	216.7553	167.0044	<u>177.1926</u>
Zero-DCE++ (Li et al., 2022)	126.0155	212.6347	<u>125.4071</u>	195.1223
RetinexNet (Wei et al., 2018)	542.6483	700.7733	615.5681	447.6674
KinD (Zhang et al., 2019)	167.0566	198.7513	158.5977	189.7126
KinD++ (Zhang et al., 2021)	353.7506	457.7775	358.5340	337.2845
URetinex-Net (Wu et al., 2022)	143.0156	<u>182.1935</u>	146.3790	178.2939
LLFlow (Wang et al., 2022)	193.1637	239.4261	253.5817	300.5299
Retinexformer (Cai et al., 2023)	<u>112.5222</u>	242.1262	145.0835	524.1901
RQ-LLIE (Liu et al., 2023)	125.4934	195.8325	136.0383	238.7689
DPRED	<b>103.9885</b>	<b>129.2544</b>	<b>111.2308</b>	<b>164.4262</b>

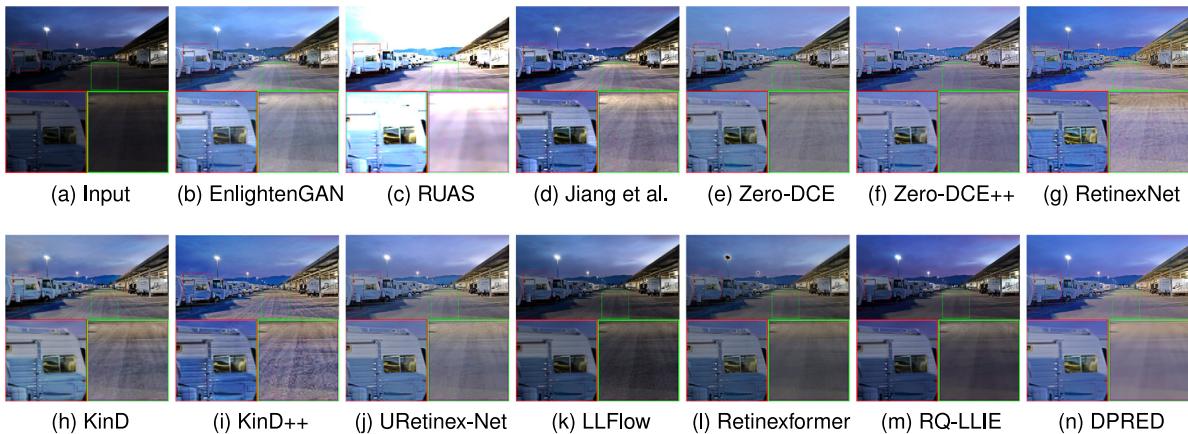
colors of the letters; LLFlow (Wang et al., 2022) and Retinexformer (Cai et al., 2023) do not brighten the image well. Benefiting from the *RDnet*, *IEnet* and *REnet*, the proposed method can better light up the low light input and maintain colors; with the additional *Fnet*, our proposed method can eliminate noise and artifacts in the enhanced images.

To further verify the generalization performance of the proposed method, we test the trained network on some non-reference datasets

(Guo et al., 2016; Wang et al., 2013; Ma et al., 2015; Lee et al., 2013). Figs. 8–11 show the enhanced images of sample images from the datasets LIME (Guo et al., 2016), NPE (Wang et al., 2013), MEF (Ma et al., 2015) and DICM (Lee et al., 2013), respectively. We can draw similar observations: the proposed method generates visually pleasant enhanced results, while previous methods cause color deviation, noise or distortions. Table 4 displays the average LOE (Wang et al., 2013) metrics for non-reference datasets (Guo et al., 2016; Wang et al., 2013;



**Fig. 8.** Visual comparison of an image from LIME (Guo et al., 2016). The objects in color boxes are zoomed in for better perception. (a) Input; (b) EnlightenGAN (Jiang et al., 2021a); (c) RUAS (Liu et al., 2021); (d) (Jiang et al., 2021b); (e) Zero-DCE (Guo et al., 2020); (f) Zero-DCE++ (Li et al., 2022); (g) RetinexNet (Wei et al., 2018); (h) KinD (Zhang et al., 2019); (i) KinD++ (Zhang et al., 2021); (j) URetinex-Net (Wu et al., 2022); (k) LLFlow (Wang et al., 2022); (l) Retinexformer (Cai et al., 2023); (m) RQ-LLIE (Liu et al., 2023); (n) DPRED.



**Fig. 9.** Visual comparison of an image from NPE (Wang et al., 2013). The objects in color boxes are zoomed in for better perception. (a) Input; (b) EnlightenGAN (Jiang et al., 2021a); (c) RUAS (Liu et al., 2021); (d) (Jiang et al., 2021b); (e) Zero-DCE (Guo et al., 2020); (f) Zero-DCE++ (Li et al., 2022); (g) RetinexNet (Wei et al., 2018); (h) KinD (Zhang et al., 2019); (i) KinD++ (Zhang et al., 2021); (j) URetinex-Net (Wu et al., 2022); (k) LLFlow (Wang et al., 2022); (l) Retinexformer (Cai et al., 2023); (m) RQ-LLIE (Liu et al., 2023); (n) DPRED.

Ma et al., 2015; Lee et al., 2013). The results indicate that our proposed method achieves the highest performance, demonstrating its promising ability in the field of LLIE.

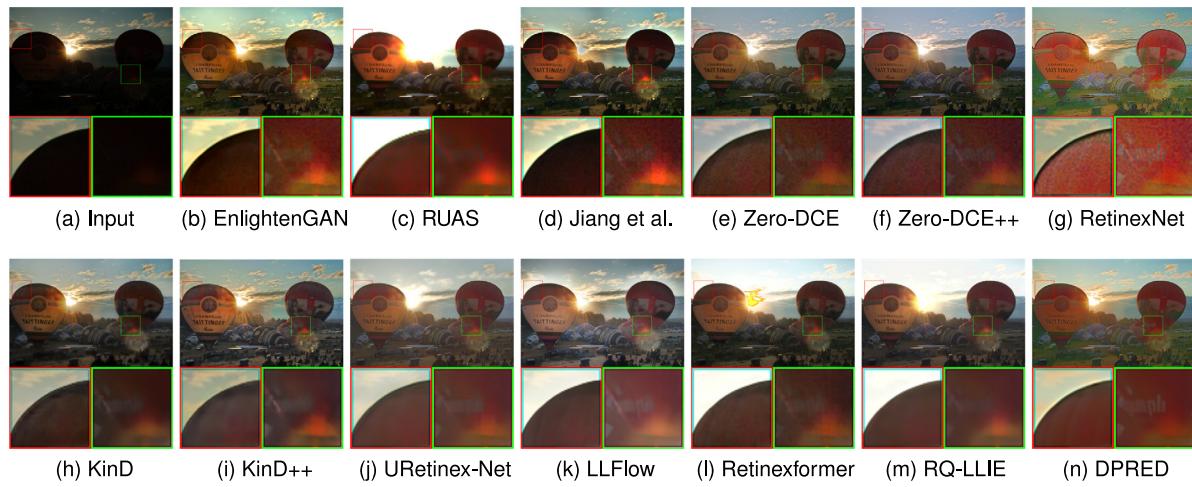
#### 4.3.3. Computational complexity

To compare the computational complexity between the proposed method and the baselines, we present in Table 5 the running time, the number of trainable parameters, and the floating point operations (FLOPs) over images of size  $600 \times 400 \times 3$ . The running time and FLOPs of our method are nearly at the median level of all methods. Compared with the supervised learning methods KinD++ (Zhang et al., 2021) and LLFlow (Wang et al., 2022), our method has much less FLOPs and running times; Though our method has more parameters and FLOPs than the supervised learning methods RetinexNet (Wei et al., 2018), KinD (Zhang et al., 2019), URetinex-Net (Wu et al., 2022), Retinexformer (Cai et al., 2023) and RQ-LLIE (Liu et al., 2023), the

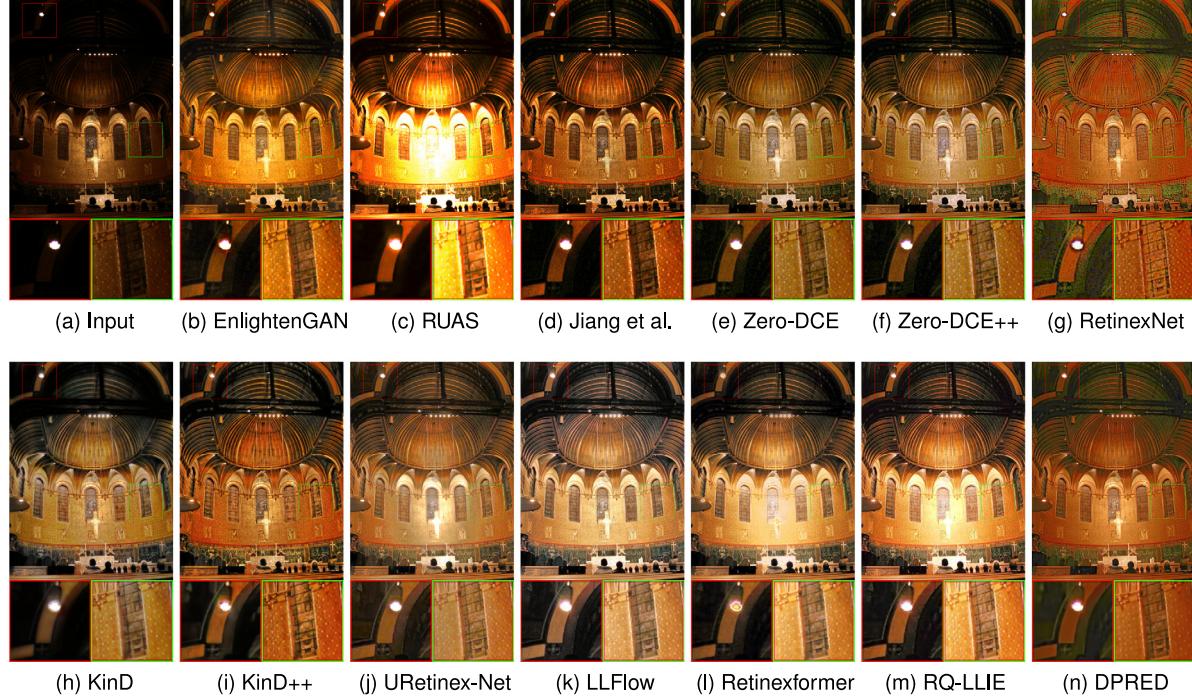
running time of our method is comparable with that of these five methods. In all, at the cost of mild parameters, FLOPs and running time, our method achieves competing enhanced results.

#### 4.4. Limitations of the proposed method

Compared with pure network-based Retinex model, the proposed parametric Retinex model may have limited capability in characterizing the prior of underlying normal-light images since it assumes the form of the regularity. However, the explicit form of the regularity has good explanation, and the learned pixel-wise varying parameters can explain the structure of the image (as shown in Fig. 3). The major limitation of the proposed method is the large size of the trainable parameters and FLOPs (as presented in Table 5). If dropping the refinement module  $F_{net}$  from the proposed model, the numbers of the trainable parameters and FLOPs reduce to 0.714M and 73.338G, respectively.



**Fig. 10.** Visual comparison of an image from MEF (Ma et al., 2015). The objects in color boxes are zoomed in for better perception. (a) Input; (b) EnlightenGAN (Jiang et al., 2021a); (c) RUAS (Liu et al., 2021); (d) (Jiang et al., 2021b); (e) Zero-DCE (Guo et al., 2020); (f) Zero-DCE++ (Li et al., 2022); (g) RetinexNet (Wei et al., 2018); (h) KinD (Zhang et al., 2019); (i) KinD++ (Zhang et al., 2021); (j) URetinex-Net (Wu et al., 2022); (k) LLFlow (Wang et al., 2022); (l) Retinexformer (Cai et al., 2023); (m) RQ-LLIE (Liu et al., 2023); (n) DPRED.



**Fig. 11.** Visual comparison of an image from DICM (Lee et al., 2013). The objects in color boxes are zoomed in for better perception. (a) Input; (b) EnlightenGAN (Jiang et al., 2021a); (c) RUAS (Liu et al., 2021); (d) (Jiang et al., 2021b); (e) Zero-DCE (Guo et al., 2020); (f) Zero-DCE++ (Li et al., 2022); (g) RetinexNet (Wei et al., 2018); (h) KinD (Zhang et al., 2019); (i) KinD++ (Zhang et al., 2021); (j) URetinex-Net (Wu et al., 2022); (k) LLFlow (Wang et al., 2022); (l) Retinexformer (Cai et al., 2023); (m) RQ-LLIE (Liu et al., 2023); (n) DPRED.

In the following Fig. 12, we show one example for which the results of our method may be unsatisfying. The enhanced images obtained by RetinexNet (Wei et al., 2018), KinD++ (Zhang et al., 2021) and URetinex-Net (Wu et al., 2022) are brighter than the ground-truth image, and the details in the dark objects are more definite. However, there exists color deviation between the enhanced image and the ground-truth. In comparison, the enhanced images obtained by Jiang et al. (2021b), RUAS (Liu et al., 2021), KinD (Zhang et al., 2019), Retinexformer (Cai et al., 2023), RQ-LLIE (Liu et al., 2023) and our method are closer to the ground-truth, yet the details in the originally dark objects are less invisible.

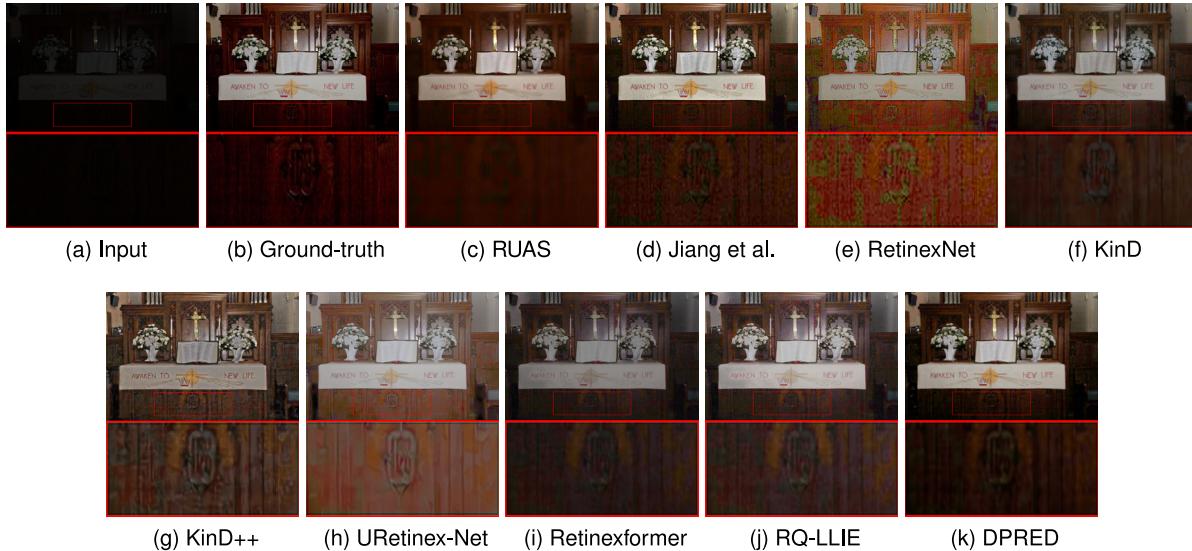
## 5. Conclusion

In this paper, we propose a novel network DPRED for low-light image enhancement. The proposed network consists of three modules: parametric Retinex decomposition, enhancement and refinement. The proposed Retinex decomposition module reconciles the advantages of analytical models and learning methods, and our ablation study shows that it greatly favors in illuminance and reflectance estimation. The enhancement networks *IEnet* and *REnet* enhance the estimated illuminance and reflectance respectively. Ablation study shows both, especially the *IEnet*, play key role in enhancing the lightness of the

**Table 5**

Quantitative comparisons of computational complexity in terms of runtime (in seconds), number of trainable parameters (#Parameters) (in M), and FLOPs (in G or T) (Note that the platform affects the running time).

Method	RunTime/s	#Parameters/M	FLOPs	Platform
EnlightenGAN (Jiang et al., 2021a)	0.1781	8.637	61.010G	PyTorch
RUAS (Liu et al., 2021)	0.1250	0.003	0.784G	PyTorch
Jiang et al. (2021b)	0.1567	4.317	36.717G	PyTorch
Zero-DCE (Guo et al., 2020)	0.1272	0.079	19.008G	PyTorch
Zero-DCE++ (Li et al., 2022)	0.1255	0.011	2.418G	PyTorch
RetinexNet (Wei et al., 2018)	0.5355	0.555	135.997G	TensorFlow
KinD (Zhang et al., 2019)	0.5442	8.160	127.768G	TensorFlow
KinD++ (Zhang et al., 2021)	6.1809	8.275	2.532T	TensorFlow
URetinex-Net (Wu et al., 2022)	0.2661	0.340	208.497G	PyTorch
LLFlow (Wang et al., 2022)	5.6695	17.420	1.050T	PyTorch
Retinexformer (Cai et al., 2023)	0.1820	1.606	62.323G	PyTorch
RQ-LLIE (Liu et al., 2023)	0.3789	11.383	593.499G	PyTorch
DPRED	0.3948	33.355	599.249G	PyTorch



**Fig. 12.** An example of unsatisfying case. (a) Input; (b) Ground-truth; (c) RUAS (Liu et al., 2021); (d) (Jiang et al., 2021b); (e) RetinexNet (Wei et al., 2018); (f) KinD (Zhang et al., 2019); (g) KinD++ (Zhang et al., 2021); (h) URetinex-Net (Wu et al., 2022); (i) Retinexformer (Cai et al., 2023); (j) RQ-LLIE (Liu et al., 2023); (k) DPRED.

image. Our experimental results show that performing Retinex decomposition and enhancement on the V channel in HSV color space can effectively avoid color deviation. Ablation study shows the refinement network is effective in suppressing noise in the enhanced RGB image. Extensive experiments show that the proposed method outperforms state-of-the-art methods and have good generalization capability. In our future research, we will focus on addressing more complex degradations that occur in low-light images.

#### CRediT authorship contribution statement

**Xiaofang Li:** Methodology, Software, Writing – original draft. **Wei-wei Wang:** Methodology, Supervision, Writing – review & editing. **Xiangchu Feng:** Conceptualization, Writing – review & editing. **Min Li:** Methodology, Writing – review & editing.

#### Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

#### Data availability

Data will be made available on request.

#### Acknowledgments

The authors would like to thank the editors and the anonymous reviewers for their constructive comments and suggestions. This paper is supported by the National Natural Science Foundation of China (Grant Nos. 61972264, 62072312, 62372302) and Natural Science Basic Research Program of Shaanxi (Program No.2024JC-YBMS-527).

#### References

- Boyd, S., Parikh, N., Chu, E., Peleato, B., Eckstein, J., et al., 2011. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Found. Trends Mach. Learn.* 3 (1), 1–122.
- Bychkovsky, V., Paris, S., Chan, E., Durand, F., 2011. Learning photographic global tonal adjustment with a database of input/output image pairs. In: Proc. IEEE Conf. Comput. Vis. Pattern Recog. pp. 97–104.
- Cai, Y., Bian, H., Lin, J., Wang, H., Timofte, R., Zhang, Y., 2023. Retinexformer: One-stage Retinex-based transformer for low-light image enhancement. In: Proc. IEEE Int. Conf. Comput. Vis. pp. 12504–12513.
- Cai, B., Xu, X., Guo, K., Jia, K., Hu, B., Tao, D., 2017. A joint intrinsic-extrinsic prior model for Retinex. In: Proc. IEEE Int. Conf. Comput. Vis. pp. 4000–4009.
- Chartrand, R., Yin, W., 2008. Iteratively reweighted algorithms for compressive sensing. In: IEEE Int. Conf. Acoust. Speech Signal Process. pp. 3869–3872.
- Dabov, K., Foi, A., Katkovnik, V., Egiazarian, K., 2007. Image denoising by sparse 3-D transform-domain collaborative filtering. *IEEE Trans. Image Process.* 16 (8), 2080–2095.
- Geman, D., Yang, C., 1995. Nonlinear image recovery with half-quadratic regularization. *IEEE Trans. Image Process.* 4 (7), 932–946.
- Guo, C., Li, C., Guo, J., Loy, C.C., Hou, J., Kwong, S., Cong, R., 2020. Zero-reference deep curve estimation for low-light image enhancement. In: Proc. IEEE Conf. Comput. Vis. Pattern Recog. pp. 1780–1789.

- Guo, X., Li, Y., Ling, H., 2016. LIME: Low-light image enhancement via illumination map estimation. *IEEE Trans. Image Process.* 26 (2), 982–993.
- Jiang, Y., Gong, X., Liu, D., Cheng, Y., Fang, C., Shen, X., Yang, J., Zhou, P., Wang, Z., 2021a. EnlightenGAN: Deep light enhancement without paired supervision. *IEEE Trans. Image Process.* 30, 2340–2349.
- Jiang, Z., Li, H., Liu, L., Men, A., Wang, H., 2021b. A switched view of retinex: deep self-regularized low-light image enhancement. *Neurocomputing* 454, 361–372.
- Jobson, D.J., Rahman, Z.u., Woodell, G.A., 1997a. A multiscale Retinex for bridging the gap between color images and the human observation of scenes. *IEEE Trans. Image Process.* 6 (7), 965–976.
- Jobson, D.J., Rahman, Z.u., Woodell, G.A., 1997b. Properties and performance of a center/surround Retinex. *IEEE Trans. Image Process.* 6 (3), 451–462.
- Kingma, D.P., Ba, J., 2015. Adam: A method for stochastic optimization. In: Bengio, Y., LeCun, Y. (Eds.), *Int. Conf. Learn. Represent.* pp. 1–15.
- Land, E.H., 1977. The Retinex theory of color vision. *Sci. Am.* 237 (6), 108–129.
- Lee, C., Lee, C., Kim, C.-S., 2013. Contrast enhancement based on layered difference representation of 2D histograms. *IEEE Trans. Image Process.* 22 (12), 5372–5384.
- Li, C., Guo, C., Loy, C.C., 2022. Learning to enhance low-light image via zero-reference deep curve estimation. *IEEE Trans. Pattern Anal. Mach. Intell.* 44 (8), 4225–4238.
- Liu, Y., Huang, T., Dong, W., Wu, F., Li, X., Shi, G., 2023. Low-light image enhancement with multi-stage residue quantization and brightness-aware attention. In: Proc. IEEE Int. Conf. Comput. Vis. pp. 12140–12149.
- Liu, R., Ma, L., Ma, T., Fan, X., Luo, Z., 2022. Learning with nested scene modeling and cooperative architecture search for low-light vision. *IEEE Trans. Pattern Anal. Mach. Intell.* (01), 1–17.
- Liu, R., Ma, L., Zhang, J., Fan, X., Luo, Z., 2021. Retinex-inspired unrolling with cooperative prior architecture search for low-light image enhancement. In: Proc. IEEE Conf. Comput. Vis. Pattern Recog. pp. 10561–10570.
- Lore, K.G., Akintayo, A., Sarkar, S., 2017. LLNet: A deep autoencoder approach to natural low-light image enhancement. *Pattern Recognit.* 61, 650–662.
- Lv, F., Lu, F., Wu, J., Lim, C., 2018. MBLLEN: Low-light image/video enhancement using CNNs. In: Proc. Brit. Mach. Vis. Conf. pp. 1–13.
- Lv, X., Zhang, S., Liu, Q., Xie, H., Zhong, B., Zhou, H., 2022. BacklitNet: A dataset and network for backlit image enhancement. *Comput. Vis. Image Underst.* 218, 103403.
- Ma, K., Zeng, K., Wang, Z., 2015. Perceptual quality assessment for multi-exposure image fusion. *IEEE Trans. Image Process.* 24 (11), 3345–3356.
- McLaren, K., 1976. XIII – the development of the CIE 1976 ( $L^*$ ,  $a^*$ ,  $b^*$ ) uniform colour space and colour-difference formula. *J. Soc. Dyers Colourists* 92 (9), 338–341.
- Pizer, S., Johnston, R., Erickson, J., Yankaskas, B., Muller, K., 1990. Contrast-limited adaptive histogram equalization: speed and effectiveness. In: Proc. 1st Conf. Visualiz. Biomed. Comput. pp. 337–345.
- Ren, W., Liu, S., Ma, L., Xu, Q., Xu, X., Cao, X., Du, J., Yang, M.-H., 2019. Low-light image enhancement via a deep hybrid network. *IEEE Trans. Image Process.* 28 (9), 4364–4375.
- Ronneberger, O., Fischer, P., Brox, T., 2015. U-net: Convolutional networks for biomedical image segmentation. In: Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Interv. pp. 234–241.
- Rudin, L.I., Osher, S., Fatemi, E., 1992. Nonlinear total variation based noise removal algorithms. *Physica D* 60 (1–4), 259–268.
- Smith, A.R., 1978. Color gamut transform pairs. *ACM Siggraph Comput. Graph.* 12 (3), 12–19.
- Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P., 2004. Image quality assessment: from error visibility to structural similarity. *IEEE Trans. Image Process.* 13 (4), 600–612.
- Wang, M., Li, J., Zhang, C., 2023. Low-light image enhancement by deep learning network for improved illumination map. *Comput. Vis. Image Underst.* 232, 103681.
- Wang, Y., Wan, R., Yang, W., Li, H., Chau, L.-P., Kot, A., 2022. Low-light image enhancement with normalizing flow. In: AAAI Conf. Artif. Intell., Vol. 36, No. 3. pp. 2604–2612.
- Wang, S., Zheng, J., Hu, H.-M., Li, B., 2013. Naturalness preserved enhancement algorithm for non-uniform illumination images. *IEEE Trans. Image Process.* 22 (9), 3538–3548.
- Wei, C., Wang, W., Yang, W., Liu, J., 2018. Deep retinex decomposition for low-light enhancement. In: Proc. Brit. Mach. Vis. Conf. pp. 1–12.
- Wu, W., Weng, J., Zhang, P., Wang, X., Yang, W., Jiang, J., 2022. URRetinex-Net: Retinex-based deep unfolding network for low-light image enhancement. In: Proc. IEEE Conf. Comput. Vis. Pattern Recog. pp. 5901–5910.
- Xu, J., Hou, Y., Ren, D., Liu, L., Zhu, F., Yu, M., Wang, H., Shao, L., 2020. STAR: A structure and texture aware Retinex model. *IEEE Trans. Image Process.* 29, 5022–5037.
- Yang, W., Wang, W., Huang, H., Wang, S., Liu, J., 2021. Sparse gradient regularized deep Retinex network for robust low-light image enhancement. *IEEE Trans. Image Process.* 30, 2072–2086.
- Zhang, Y., Guo, X., Ma, J., Liu, W., Zhang, J., 2021. Beyond brightening low-light images. *Int. J. Comput. Vis.* 129 (4), 1013–1037.
- Zhang, R., Isola, P., Efros, A.A., Shechtman, E., Wang, O., 2018. The unreasonable effectiveness of deep features as a perceptual metric. In: Proc. IEEE Conf. Comput. Vis. Pattern Recog.. pp. 586–595.
- Zhang, K., Li, Y., Zuo, W., Zhang, L., Van Gool, L., Timofte, R., 2022. Plug-and-play image restoration with deep denoiser prior. *IEEE Trans. Pattern Anal. Mach. Intell.* 44 (10), 6360–6376.
- Zhang, Y., Zhang, J., Guo, X., 2019. Kindling the darkness: A practical low-light image enhancer. In: Proc. 27th ACM Int. Conf. Multimedia. pp. 1632–1640.