



## Structure aware transfer function network for low light image enhancement

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

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

<sup>b</sup> School of Mathematics Science, Shenzhen University, Shenzhen, 518060, China

### ARTICLE INFO

**Keywords:**

Low-light image enhancement  
Transfer function  
Retinex decomposition  
Structure  
Denoising

### ABSTRACT

Low-light images are prevalent in surveillance, photography, and underwater imaging, but suffer from poor visibility. It is demanding to develop effective techniques for enhancing low-light images. Compared with the methods based on Retinex decomposition, the transfer function provides a simpler pipeline for low light image enhancement. However, the previous method has limitations in estimating the transfer function and the final enhancement. In this work, we present a framework for low light image enhancement, wherein we give a novel supervised approach for learning the transfer function from the low-light image. We call the proposed model Structure Aware Transfer function Network (SATNet), comprising two sub-networks: the Structure-based Transfer network (*STnet*), and the Refinement network (*Rnet*). The *STnet* learns the transfer function from the low-light input image, then a coarsely enhanced image is obtained simply by multiplying the low-light input by the estimated transfer function. Then the *Rnet* eliminates the noise in the coarsely enhanced image. Extensive experiments on seven datasets demonstrate that the proposed method outperforms baselines in brightening low-light images and preserving details.

### 1. Introduction

Low-light images are prevalent in many applications, including surveillance, photography, underwater imaging, and computer vision (Chen et al., 2023a; Li et al., 2023; Cui et al., 2022; Fan et al., 2023; He et al., 2023). These images suffer from poor visibility due to various reasons such as insufficient light sources or inappropriate exposure time setting. In such images, the objects or scenes may appear dark, noisy, and color deviation, making it difficult to discern details or extract useful information. Therefore, it is demanding to develop effective techniques for enhancing low-light images.

Over the past decade, significant efforts have been made in low-light image enhancement (LLIE), ranging from analytical methods (Pizer et al., 1990; Jobson et al., 1997b,a; Guo et al., 2016; Cai et al., 2017; Xu et al., 2020a) to more advanced learning-based techniques (Lore et al., 2017; Lv et al., 2018, 2021; Xu et al., 2022b; Zhang et al., 2019b; Guo et al., 2020; Jiang et al., 2021; Yang et al., 2021; Zheng et al., 2021; Xu et al., 2022a; Wei et al., 2018; Zhang et al., 2019a, 2021; Wang and Zhang, 2023; Wu et al., 2022; Liu et al., 2023; Wang et al., 2019; Liu et al., 2021; Lu and Zhang, 2020). Analytical methods (Pizer et al., 1990; Jobson et al., 1997b,a; Guo et al., 2016; Cai et al., 2017; Xu et al., 2020a) involve histogram equalization (HE) (Pizer et al., 1990), Gamma correction (GC) and Retinex theory (Land, 1977) based methods (Jobson et al., 1997b,a; Guo et al., 2016; Cai et al.,

2017; Xu et al., 2020a). However, both HE and GC amplify noise while enhancing the brightness of images. The Retinex-based methods (Jobson et al., 1997b,a; Guo et al., 2016; Cai et al., 2017; Xu et al., 2020a) incorporate the physical mechanism of image formation and have received the majority attention. They perform LLIE in two phases: Retinex decomposition and enhancement. The first phase decomposes an image into illumination and reflectance. In the second phase, some methods (Jobson et al., 1997b,a) simply take the reflectance as the final enhanced result, yet having color distortion and looking unrealistic. Other methods (Guo et al., 2016; Cai et al., 2017; Xu et al., 2020a) modify the illumination by using GC and then combine it with the reflectance to generate the final enhanced image.

Benefiting from the powerful learning ability, various deep learning based methods (Lore et al., 2017; Lv et al., 2018, 2021; Xu et al., 2022b; Zhang et al., 2019b; Guo et al., 2020; Jiang et al., 2021; Yang et al., 2021; Zheng et al., 2021; Xu et al., 2022a; Wei et al., 2018; Zhang et al., 2019a, 2021; Wang and Zhang, 2023; Wu et al., 2022; Liu et al., 2023; Wang et al., 2019; Liu et al., 2021; Lu and Zhang, 2020) have been developed. Though showing effectiveness in low-light image enhancement, previous methods appear incorrect exposure, color distortion, or artifacts to degrade visual quality. Lu and Zhang (2020) proposed a simpler pipeline for LLIE, in which estimating a single

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

E-mail addresses: [xfl\\_96@stu.xidian.edu.cn](mailto:xfl_96@stu.xidian.edu.cn) (X. Li), [wwwang@mail.xidian.edu.cn](mailto:wwwang@mail.xidian.edu.cn) (W. Wang), [hany@szu.edu.cn](mailto:hany@szu.edu.cn) (Y. Han), [xcfeng@mail.xidian.edu.cn](mailto:xcfeng@mail.xidian.edu.cn) (X. Feng).

transfer function plays the key role. Considering noise in the low-light input, they introduced a two-branch exposure fusion network (TBEFN). One branch estimates the transfer function directly from the low-light input, while the other branch estimates the transfer function from a denoised version of the low-light input. The enhanced results from both branches are subsequently fused to produce the final enhanced image. Though TBEFN (Lu and Zhang, 2020) has achieved good performance, either branch of TBEFN (Lu and Zhang, 2020) for estimating the transfer function has limitations. On one hand, the low-light image contains noise, estimating the transfer function directly from it suffers from the noise wherein. On the other hand, though denoising the low-light image can reduce noise, it results in loss of details, which also influences the subsequent transfer function estimation. Moreover, the estimated transfer function heavily depends on the training data, and applying it to test data may cause over-exposure or under-exposure, if the test data and the training data are not obtained under the same exposure.

In this work, we present a new end-to-end neural network for low-light image enhancement. Our network adopts the pipeline in Lu and Zhang (2020) because it is simpler than the Retinex decomposition-enhancement framework (Wei et al., 2018; Zhang et al., 2019a, 2021; Wang and Zhang, 2023; Wu et al., 2022; Liu et al., 2023; Wang et al., 2019; Liu et al., 2021). Estimation of the transfer function is a key step in the pipeline. For that, we present a novel supervised approach for learning the transfer function from the structure of the low-light image. To implement the process, we design a subnetwork, called structure based transfer network (*STnet*). By using the estimated transfer function, we can easily obtain a normal-light image, simply by multiplying the estimated transfer function with the low-light input. However, this normal-light image contains noise since the low-light input usually contains noise. To remove the noise, we design a subnet, called *Rnet* to refine the noisy normal-light image. The whole network is called structure-aware transfer function network (*SATNet*). For training the whole network, we define a new loss to guarantee the color consistency of the estimated normal-light image and the ground-truth. Additionally, to enhance adaptability of the transfer function to test data captured at various exposures (may be different from that the training data captured), in test phase, the estimated transfer function is adjusted by using Gamma correction. In summary, the contributions of this work are three-fold:

- The main contribution is the proposed approach for estimating the transfer function from the structure of the low-light image. Base on this, we design a new network, called *SATNet* for LLIE, featuring two stages: transfer function estimation and refinement.
- To train the proposed network, we define a novel color consistency loss, effectively mitigating color deviation in the enhanced image. Moreover, we provide users the flexibility to fine-tune the transfer function using the Gamma correction, so that the proposed method can well adapt to enhance images at various exposure levels.
- Experimental results show that our method can effectively enhance the low-light images and generate more visually pleasant and realistic enhanced results than state-of-the-art methods.

The rest of this article is organized as follows. We first review some closely related works in Section 2. Then we describe the proposed network for LLIE in Section 3. Experimental results and analysis are provided in Sections 4 and 5. Conclusions are finally drawn in Section 6.

## 2. Related work

In this section, we briefly review learning-based LLIE methods, transfer function, structure–texture decomposition and color constancy loss for LLIE.

### 2.1. Learning-based LLIE methods

In Fig. 1, we categorize most existing LLIE methods into two groups: analytical methods (Pizer et al., 1990; Jobson et al., 1997b,a; Guo et al., 2016; Cai et al., 2017; Xu et al., 2020a) and deep learning methods (Lore et al., 2017; Lv et al., 2018, 2021; Xu et al., 2022b; Zhang et al., 2019b; Guo et al., 2020; Jiang et al., 2021; Yang et al., 2021; Zheng et al., 2021; Xu et al., 2022a; Wei et al., 2018; Zhang et al., 2019a, 2021; Wang and Zhang, 2023; Wu et al., 2022; Liu et al., 2023; Wang et al., 2019; Liu et al., 2021; Lu and Zhang, 2020). The former group includes histogram equalization (Pizer et al., 1990), Gamma correction, and Retinex based methods (Jobson et al., 1997b,a; Guo et al., 2016; Cai et al., 2017; Xu et al., 2020a). In this subsection, we mainly review the deep learning methods (Lore et al., 2017; Lv et al., 2018, 2021; Xu et al., 2022b; Zhang et al., 2019b; Guo et al., 2020; Jiang et al., 2021; Yang et al., 2021; Zheng et al., 2021; Xu et al., 2022a; Wei et al., 2018; Zhang et al., 2019a, 2021; Wang and Zhang, 2023; Wu et al., 2022; Liu et al., 2023; Wang et al., 2019; Liu et al., 2021; Lu and Zhang, 2020), which we further classify into single-stage methods (Lore et al., 2017; Lv et al., 2018, 2021; Xu et al., 2022b; Zhang et al., 2019b; Guo et al., 2020) and multi-stage methods (Jiang et al., 2021; Yang et al., 2021; Zheng et al., 2021; Xu et al., 2022a; Wei et al., 2018; Zhang et al., 2019a, 2021; Wang and Zhang, 2023; Wu et al., 2022; Liu et al., 2023; Wang et al., 2019; Liu et al., 2021).

The single-stage methods (Lore et al., 2017; Lv et al., 2018, 2021; Xu et al., 2022b; Zhang et al., 2019b; Guo et al., 2020) train various convolutional neural networks (CNNs) to map low-light images directly to well-lit images. For example, Lore et al. (2017) proposed a deep auto-encoder, called LLNet, to brighten low-light images. Lv et al. (2018) presented a multi-branch low light enhancement network (MBLLEN), which enhances features at different scales and fuses the enhanced features to get the enhanced images. Later, Lv et al. (2021) incorporated illumination and noise attention maps in MBLLEN to improve the feature enhancement module. Xu et al. (2022b) presented a signal-to-noise-ratio-aware transformer network, which takes signal-to-noise-ratio (SNR) prior to guide the feature fusion module. The above models are completely based on neural network, lacking interpretation and demanding huge computational resources (Lu and Zhang, 2020). In contrast, some researchers regard LLIE as image-specific curve estimation, which can map the low-light input to a normal light output. For example, Zhang et al. (2019b) proposed a zero-shot exposure correction network (ExCNet), which learns an S-curve to brighten the low-light images. Guo et al. (2020) designed the zero-reference deep curve estimation (Zero-DCE) network. These methods perform LLIE without using paired data, yet the enhancement quality is limited (Zhang et al., 2022).

The multi-stage methods (Jiang et al., 2021; Yang et al., 2021; Zheng et al., 2021; Xu et al., 2022a; Wei et al., 2018; Zhang et al., 2019a, 2021; Wang and Zhang, 2023; Wu et al., 2022; Liu et al., 2023; Wang et al., 2019; Liu et al., 2021) can be divided into three categories: generative adversarial network (GAN)-based (Jiang et al., 2021; Yang et al., 2021), structure–texture decomposition (STD)-based (Zheng et al., 2021; Xu et al., 2022a), and Retinex-based (Wei et al., 2018; Zhang et al., 2019a, 2021; Wang and Zhang, 2023; Wu et al., 2022; Liu et al., 2023; Wang et al., 2019; Liu et al., 2021) methods. GAN-based methods (Jiang et al., 2021; Yang et al., 2021) are trained with unpaired data and consist of two stages: generating the enhanced image and discriminating whether the enhanced image is real or not. The EnlightenGAN (Jiang et al., 2021) uses an attention-guided U-Net as the generator, and the global-local discriminator encourages the enhanced result to be close to a realistic normal-light image. In Yang et al. (2021), domain adaptation and noise reduction mechanisms are introduced to guide the global U-net generator.

STD-based methods (Zheng et al., 2021; Xu et al., 2022a) follow two stages: STD and enhancement. The STD stage decomposes an image into two components: structure and texture. The former component mainly

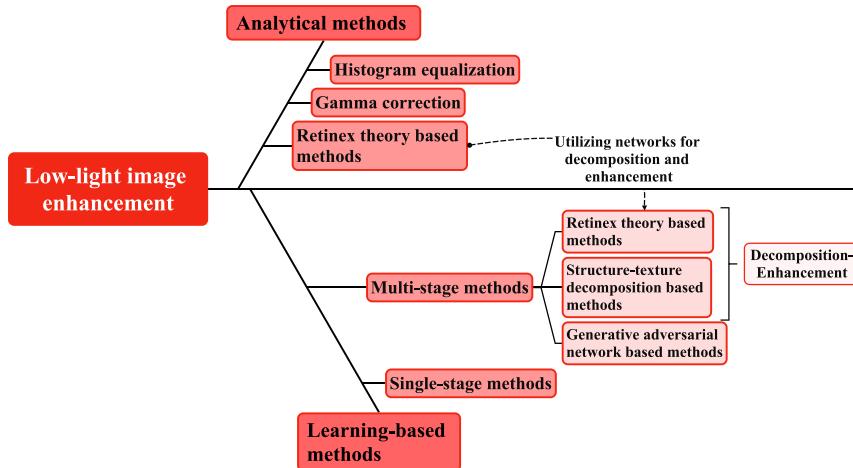


Fig. 1. Overview of LLIE methods.

contains the primary structure, comprising homogeneous regions, contours, and sharp edges. The texture component encompasses oscillating patterns such as fine details and locally repeated features. The enhancement stage first modifies the structure and the texture respectively, then adds them together to yield the final enhanced image. The unfolding total variation network (UTVNet) (Zheng et al., 2021) consists of three modules: structure estimation, structure brightening, and texture denoising. The structure–texture aware network (STANet) (Xu et al., 2022a) utilizes guided filtering (GF) (He et al., 2012) to obtain the structure and the texture, for which attention maps are generated to guide the enhancement process.

The Retinex learning methods (Wei et al., 2018; Zhang et al., 2019a, 2021; Wang and Zhang, 2023; Wu et al., 2022; Liu et al., 2023; Wang et al., 2019; Liu et al., 2021) also consist of two stages: Retinex decomposition and enhancement. Wei et al. (2018) proposed the RetinexNet, which utilizes two sub-networks to perform the Retinex decomposition and illumination enhancement respectively. The analytic method block-matching and 3D filtering (BM3D) (Dabov et al., 2007) is also used to refine the reflectance. Zhang et al. (2019a) modified the RetinexNet (Wei et al., 2018) by replacing the BM3D (Dabov et al., 2007) with a reflectance restoration network, called KinD. KinD++ (Zhang et al., 2021) further incorporates a multi-scale illumination attention module in KinD (Zhang et al., 2019a). Wang and Zhang (2023) introduced a global attention module in their reflectance restoration network. Differently, the Retinex decomposition networks in Wu et al. (2022), Liu et al. (2023) are constructed by unrolling some optimization algorithms. Specifically, the URetinexNet (Wu et al., 2022) is obtained by unrolling the half quadratic splitting (HQS) (Geman and Yang, 1995) scheme of an implicitly regularized Retinex decomposition model. The Retinex decomposition network in Liu et al. (2023) is constructed by unrolling the proximal gradient descent (PGD) iteration scheme of a Retinex model involving implicit and explicit regularizations; the illumination is subsequently improved by using a global brightness parameter, while the reflectance is refined through a local brightness attention map. The methods in literature (Wang et al., 2019; Liu et al., 2021) remove the illumination from the low-light input to extract a noisy reflectance, and a denoising network is applied to obtain the final enhanced image. To be specific, in Wang et al. (2019), the illumination is estimated by a network and the final enhanced image is the product of the low-light input and the inverse illumination. In Liu et al. (2021), the illuminance is estimated by using the gradient descent (GD) solution of a network regularized model, and a denoising network is used to refine the reflectance.

## 2.2. Transfer function

Let  $(\mathbf{O}_l, \mathbf{O}_n)$  be a pair of low-light and the corresponding normal-light images. Following the Retinex theory (Land, 1977),  $\mathbf{O}_l$  and  $\mathbf{O}_n$  have the same reflectance  $\mathbf{R}$ , or

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

$$\mathbf{O}_n = \mathbf{I}_n \cdot \mathbf{R}, \quad (2)$$

where  $\mathbf{I}_l$  and  $\mathbf{I}_n$  are low-light illumination and normal light illumination respectively.  $\cdot$  denotes element-wise product. TBEFN (Lu and Zhang, 2020) obtains the following observation:

$$\mathbf{O}_n = \mathbf{O}_l \cdot (\mathbf{I}_n / \mathbf{I}_l) = \mathbf{O}_l \cdot f, \quad (3)$$

where  $/$  denotes the element-wise division.  $f = \mathbf{I}_n / \mathbf{I}_l$  is called transfer function. Eq. (3) provides a new pipeline for LLIE, where estimation of the transfer function plays the key role. Considering presence of noise  $\mathbf{N}$  in the low-light input, the robust Retinex model is formulated as

$$\mathbf{O}_l = \mathbf{I}_l \cdot \mathbf{R} + \mathbf{N}. \quad (4)$$

Then Eq. (3) can be rewritten as

$$\mathbf{O}_n = (\mathbf{O}_l - \mathbf{N}) \cdot f. \quad (5)$$

To estimate the transfer function  $f$ , TBEFN (Lu and Zhang, 2020) employs two branches. One branch estimates a transfer function directly from the low-light input  $\mathbf{O}_l$ ; while the other branch estimates a transfer function from a denoised version of the low-light input. The low-light image contains noise, estimating the transfer function directly from it suffers from the noise wherein. Denoising the low-light image results in blurring, which also influence the subsequent transfer function estimation. Except this, TBEFN (Lu and Zhang, 2020) has other limitations discussed in the introduction.

## 2.3. Structure–texture decomposition

The structure–texture decomposition (STD) is a fundamental task in image processing and has wide applications in edge detection, image restoration and so on. In general, the STD can be formulated in the following form:

$$\mathbf{O} = \mathbf{S} + \mathbf{T}, \quad (6)$$

where  $\mathbf{O}$ ,  $\mathbf{S}$  and  $\mathbf{T}$  respectively denote the input image, structure and texture. Existing STD methods (He et al., 2012; Rudin et al., 1992; Xu et al., 2012; Tomasi and Manduchi, 1998; Ule, 1955) roughly fall into two categories. The first category (Rudin et al., 1992; Xu

et al., 2012) takes the structure and the texture as functions in certain functional spaces, and regularizes them by norms of the corresponding functional spaces. Among these methods, total variation (TV) (Rudin et al., 1992) and relative total variation (RTV) (Xu et al., 2012) are mostly used norms for structure. The second category designs different edge-preserving filters such as GF (He et al., 2012), bilateral filtering (BF) (Tomasi and Manduchi, 1998) and weighted least squares filtering (WLSF) (Ule, 1955) to obtain the structure. For LLIE task, researchers usually brighten the structure layer and remove the noise in the texture layer to obtain the final enhanced image.

#### 2.4. Color constancy loss for low-light image enhancement

The three color channels of RGB images are correlated with the illumination, and the Retinex-based learning methods (Wei et al., 2018; Zhang et al., 2019a, 2021; Wang and Zhang, 2023; Wu et al., 2022; Liu et al., 2023; Wang et al., 2019; Liu et al., 2021) show that performing Retinex decomposition on RGB images results in color deviation. To overcome this problem, the methods in Cai et al. (2017), Xu et al. (2020a) only enhance the V channel in HSV color space. Chen et al. (2023b) design a two-branch network to enhance images in YCbCr color space. One branch aims to enhance the luminance, while the other branch aims to preserve the chrominance information. Nevertheless, color space conversion may yield extra noise, requiring additional post-processing. Zhang et al. (2022) utilized two networks to generate enhanced gray image and its color histograms, which are then fused to generate the enhanced color image. However, the back-propagation of the histogram loss is challenging. The methods in Guo et al. (2020), Xu et al. (2022a), Wang et al. (2019) designed various color constancy losses for LLIE. For example, following the gray-world color constancy hypothesis, Guo et al. (2020) assumes the three color channels have uniform means. Nonetheless, this hypothesis is sensitive to the color and lighting in the scene. Xu et al. (2022a) enforced similarity of the color saturation between the enhanced image and the corresponding normal light image, which ignores the influence of image hue. Wang et al. (2019) constrained the angle similarity between two images by regarding the RGB images as 3D vectors, which quantifies the overall color discrepancy and neglects the color consistency in individual channels. To train our proposed network on RGB images, we define a new loss to guarantee the color consistency of the estimated normal-light image and the ground-truth. We will explain it in Section 3.2.2.

### 3. Proposed method

In this section, we introduce the motivation and the details of the proposed method.

#### 3.1. Motivation

As defined in Eq. (2), the Retinex decomposition-enhancement methods (Wei et al., 2018; Zhang et al., 2019a, 2021; Wang and Zhang, 2023; Wu et al., 2022; Liu et al., 2023; Wang et al., 2019; Liu et al., 2021) need to estimate both the illuminance and the reflectance of the normal-light image. The reflectance can be estimated or learned from the low-light image. However, in test phase, the illuminance of the normal-light image is unknown, and its estimation is quite challenging. In contrast, Eq. (3) provides a simpler pipeline for LLIE, which only needs estimating the transfer function  $f$ . More significantly, estimating  $f$  does not need Retinex decomposition. To be specific, in case of  $\mathbf{O}_l$  noise-free, we have

$$\mathbf{f} = \mathbf{O}_n / \mathbf{O}_l, \quad (7)$$

meaning that  $f$  can be estimated directly by  $\mathbf{O}_n / \mathbf{O}_l$ . In case of noisy  $\mathbf{O}_l$ ,  $\mathbf{O}_n$  can be estimated by the Eq. (5), and the transfer function  $f$  can be estimated by

$$\mathbf{f} = \mathbf{O}_n / (\mathbf{O}_l - N) = \mathbf{O}_n / \hat{\mathbf{O}}_l, \quad (8)$$

**Table 1**

The PSNR and SSIM metrics of the enhanced images obtained by different transfer functions.

Methods	PSNR	SSIM
$\mathbf{O}_l \cdot f$ , $f$ estimated from $\mathbf{O}_l$ by TBEFN (Lu and Zhang, 2020)	16.2956	0.3472
$\hat{\mathbf{O}}_l \cdot f$ , $f$ estimated from $\hat{\mathbf{O}}_l$ by TBEFN (Lu and Zhang, 2020)	14.7493	0.6856
TBEFN (Lu and Zhang, 2020)	18.8596	0.7304
$\mathbf{O}_l \cdot (\mathbf{I}_n / \mathbf{I}_l)$	22.0453	0.4850
$\mathbf{O}_l \cdot (\mathbf{O}_n / \hat{\mathbf{O}}_l)$	25.9552	0.5318
$\mathbf{O}_l \cdot (\mathbf{S}_n / \mathbf{S}_l)$	25.9255	0.5627
$\mathbf{O}_l \cdot (\mathbf{T}_n / \mathbf{T}_l)$	11.5528	0.2053

where  $\hat{\mathbf{O}}_l = \mathbf{O}_l - N$ , denotes the latent noise-free low-light image. Again,  $f$  can be estimated from  $\mathbf{O}_l$  and  $\mathbf{O}_n$ , without using the Retinex decomposition.

TBEFN (Lu and Zhang, 2020) firstly estimates  $\hat{\mathbf{O}}_l$  by denoising the low-light input  $\mathbf{O}_l$ , and then estimates the transfer function  $f$  from  $\hat{\mathbf{O}}_l$ . However, estimation of  $\hat{\mathbf{O}}_l$  by denoising has bias, which adversely affects the transfer function estimation. Additionally, TBEFN (Lu and Zhang, 2020) also estimates a transfer function directly from the noisy low-light input, which is influenced by the noise. In Figs. 2(c)-2(e), we show the enhanced images  $\mathbf{O}_l \cdot f$ ,  $\hat{\mathbf{O}}_l \cdot f$ , and the final enhanced image. The metrics Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity (SSIM) (Wang et al., 2004) are provided in Table 1 for objective evaluation. The transfer functions are estimated by TBEFN (Lu and Zhang, 2020) directly from  $\mathbf{O}_l$  and  $\hat{\mathbf{O}}_l$ , respectively.  $\hat{\mathbf{O}}_l$  is obtained by the denoising network in TBEFN (Lu and Zhang, 2020). We can note that, the transfer function  $f$  estimated directly from  $\mathbf{O}_l$  leads to noisy enhanced image  $\mathbf{O}_l \cdot f$ ; the transfer function  $f$  estimated from  $\hat{\mathbf{O}}_l$  results in severe color deviation in the enhanced image  $\hat{\mathbf{O}}_l \cdot f$ , which also spreads to the final fused image.

In this work, we present a new method for estimation of the transfer function. Our method is inspired by the STD of images. Let

$$\mathbf{O}_n = \mathbf{S}_n + \mathbf{T}_n, \quad (9)$$

and

$$\mathbf{O}_l = \mathbf{S}_l + \mathbf{T}_l + \mathbf{N}, \quad (10)$$

where  $\mathbf{S}_n$  and  $\mathbf{T}_n$  denote the structure and the texture of the normal-light image, respectively;  $\mathbf{S}_l$  and  $\mathbf{T}_l$  denote the noise-free structure and texture of the low-light counterpart, respectively. We get the following observation:

$$\mathbf{O}_n = (\mathbf{S}_l + \mathbf{T}_l) \cdot f = \mathbf{S}_l \cdot f + \mathbf{T}_l \cdot f = \mathbf{S}_n + \mathbf{T}_n. \quad (11)$$

Intuitively,

$$\mathbf{S}_n = \mathbf{S}_l \cdot f, \quad (12)$$

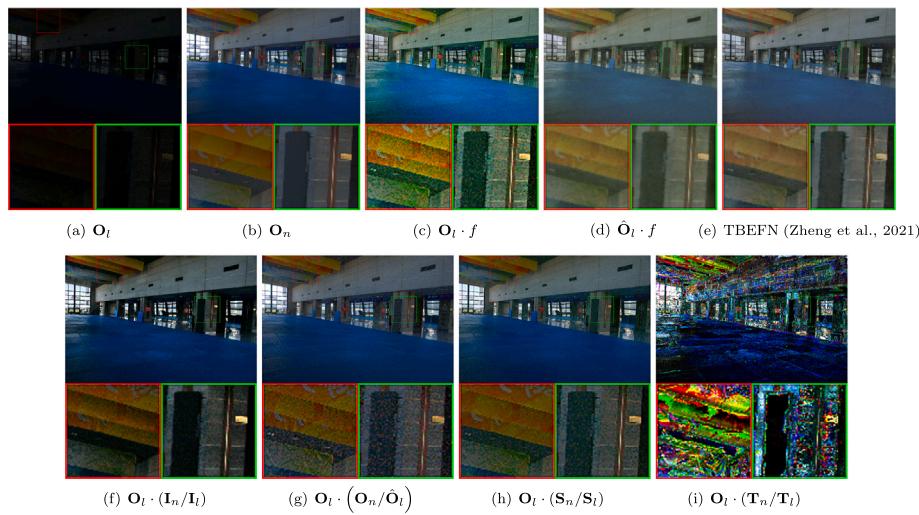
and

$$\mathbf{T}_n = \mathbf{T}_l \cdot f, \quad (13)$$

The above equations provide new perspectives for estimating the transfer function.

In general, the structure contains the mainframe of the image, free of noise. For computation convenience, the texture  $\mathbf{T}_l$  is simply obtained by  $\mathbf{O}_l - \mathbf{S}_l$ , which contains noise if  $\mathbf{O}_l$  is noisy. Therefore, we choose Eq. (12) to estimate the transfer function  $f$ , irrespective of  $\mathbf{O}_l$  noise-free or not.

To further verify the rationality of our strategy for estimating the transfer function  $f$ , we compare the enhanced images obtained by  $\mathbf{O}_l \cdot f$  with  $f = \mathbf{I}_n / \mathbf{I}_l$ ,  $\mathbf{O}_n / \hat{\mathbf{O}}_l$ ,  $\mathbf{S}_n / \mathbf{S}_l$  and  $\mathbf{T}_n / \mathbf{T}_l$ , respectively. In this experiment, we mainly compare the effects of different estimations of  $f$ . For convenience, we simply use  $\mathbf{O}_l$  to obtain the enhanced image. Note that  $\mathbf{I}_n$  and  $\mathbf{I}_l$  are computed by using the method in Xu



**Fig. 2.** Comparison of the enhanced images obtained by different transfer functions. (c)  $f$  estimated from  $\mathbf{O}_l$  by TBEFN (Lu and Zhang, 2020); (d)  $f$  estimated from  $\hat{\mathbf{O}}_l$  by TBEFN (Lu and Zhang, 2020). The objects in color boxes are zoomed in for better perception.

et al. (2020a),  $\hat{\mathbf{O}}_l$  is obtained by applying the denoising algorithm BM3D (Dabov et al., 2007) to  $\mathbf{O}_l$ .  $\mathbf{S}_n$  and  $\mathbf{S}_l$  are computed by using the RTV (Xu et al., 2012), and the textures are obtained simply by  $\mathbf{T}_n = \mathbf{O}_n - \mathbf{S}_n$  and  $\mathbf{T}_l = \hat{\mathbf{O}}_l - \mathbf{S}_l$ , respectively. The original normal/low light image pairs and the enhanced images are shown in Fig. 2. The metrics PSNR and SSIM (Wang et al., 2004) are provided in Table 1 for objective evaluation. We can note that, the enhanced image obtained with  $f = \mathbf{T}_n/\mathbf{T}_l$  exhibits severe distortion, suggesting its unsuitability for the transfer function estimation. The enhanced images obtained with  $f = \mathbf{O}_n/\hat{\mathbf{O}}_l$  and  $f = \mathbf{S}_n/\mathbf{S}_l$  look much better than that obtained with  $f = \mathbf{I}_n/\mathbf{I}_l$  in brightness. Comparing the enhanced images by using  $f = \mathbf{S}_n/\mathbf{S}_l$ , and  $f = \mathbf{O}_n/\hat{\mathbf{O}}_l$ , we can note that the later one suffers from more noticeable noise. This empirical analysis verifies the advantages of our strategy (Eq. (12)) for estimating the transfer function.

Once the transfer function  $f$  is estimated, the enhanced image  $\mathbf{O}_n$  can be restored from the low-light counterpart by using Eq. (5), or  $\mathbf{O}_n = (\mathbf{O}_l - \mathbf{N}) \cdot f = \hat{\mathbf{O}}_l \cdot f$ . This means denoising  $\mathbf{O}_l$  to obtain  $\hat{\mathbf{O}}_l$  before lighting by using  $f$ . However, the possible residual noise (even very weak) in the denoised low light image  $\hat{\mathbf{O}}_l$  can be enlarged by  $f$ -fold, and  $f = \mathbf{I}_n/\mathbf{I}_l > 1$  could be relatively large. So the enhanced image  $\mathbf{O}_n$  may have strong noise, demanding a post-denoising operation.

Considering

$$\mathbf{O}_n = (\mathbf{O}_l - \mathbf{N}) \cdot f = \mathbf{O}_l \cdot f - \mathbf{N} \cdot f. \quad (14)$$

We prefer to recover the enhanced image  $\mathbf{O}_n$  by first computing  $\mathbf{O}_l \cdot f$  and then removing the noise  $\mathbf{N} \cdot f$  from  $\mathbf{O}_l \cdot f$ . Note that such strategy needs only one denoising operation.

In summary, our method first estimates the transfer function by using Eq. (12). Once  $f$  is obtained, we compute a noisy normal-light image  $\mathbf{O}_l \cdot f$ , then obtain  $\mathbf{O}_n = \text{Denoiser}(\mathbf{O}_l \cdot f)$ , where  $\text{Denoiser}$  denotes a denoising operator.

### 3.2. Proposed network

To learn the transfer function and the denoiser, we propose a structure-aware transfer function network (SATNet) for LLIE. Note that the SATNet is trained and tested on RGB images. As illustrated in Fig. 3, the network comprises two modules:  $STnet$  and  $Rnet$ . The  $STnet$  estimates the transfer function  $f$  by using image structures, while the  $Rnet$  removes noise in  $\mathbf{O}_l \cdot f$ . The final normal-light image can be formulated as

$$f = STnet(\mathbf{O}_l; \theta_{ST}), \quad (15)$$

$$\mathbf{O}_n = Rnet(\mathbf{O}_l \cdot f; \theta_R), \quad (16)$$

where  $STnet(\mathbf{O}_l; \theta_{ST})$  ( $\theta_{ST}$  denotes its parameters) learns  $f$  from a given set of low-light images.  $Rnet(\theta_R)$  denotes the parameters) refines  $\mathbf{O}_l \cdot f$  to generate a normal light image  $\mathbf{O}_n$ . We will discuss the details in the following subsections.

#### 3.2.1. Transfer function estimation module

As discussed in the motivation, we utilize the structures of the low light image  $\mathbf{O}_l$  and the ground-truth normal light image  $\mathbf{O}_{gt}$ , denoted by  $\mathbf{S}_l$  and  $\mathbf{S}_{gt}$ , respectively to train  $STnet$  for estimating the transfer function  $f$ . The problem can be formulated as

$$\min_{\theta_{ST}} \mathcal{L}_{ST}(\mathbf{S}_l \cdot f, \mathbf{S}_{gt}), \quad (17)$$

where  $f = STnet(\mathbf{O}_l; \theta_{ST})$  is the estimated transfer function, the loss function  $\mathcal{L}_{ST}$  enforces  $\mathbf{S}_l \cdot f$  to be consistent with the counterpart structure  $\mathbf{S}_{gt}$ . We define  $\mathcal{L}_{ST}$  by

$$\mathcal{L}_{ST} = \lambda_1 \mathcal{L}_{mse} + \mathcal{L}_{exp}, \quad (18)$$

where  $\mathcal{L}_{mse}$  denotes the mean square error (MSE) between  $\mathbf{S}_l \cdot f$  and  $\mathbf{S}_{gt}$ ,  $\mathcal{L}_{exp}$  is the exposure loss, originated in Zero-DCE (Zhang et al., 2019b), and  $\lambda_1$  is a tuning parameter. Specifically,

$$\mathcal{L}_{mse} = \frac{1}{M} \|\mathbf{S}_l \cdot f - \mathbf{S}_{gt}\|_2^2, \quad (19)$$

where  $M$  is the total number of pixels, and

$$\mathcal{L}_{exp} = \frac{1}{W} \sum_{k=1}^W \|m_{\mathbf{S}_l \cdot f}(P_k) - m_{\mathbf{S}_{gt}}(P_k)\|_2^2, \quad (20)$$

where  $P_k$  denotes a local patch, of size  $16 \times 16$ ,  $W$  denotes the total number of such non-overlapping patches.  $m$  denotes the mean intensity value of an image in a patch. Note that, we use  $\mathcal{L}_{exp}$  to enhance local exposure consistency between  $\mathbf{S}_l \cdot f$  and  $\mathbf{S}_{gt}$ , so that the transfer function  $f$  can be estimated more precisely.

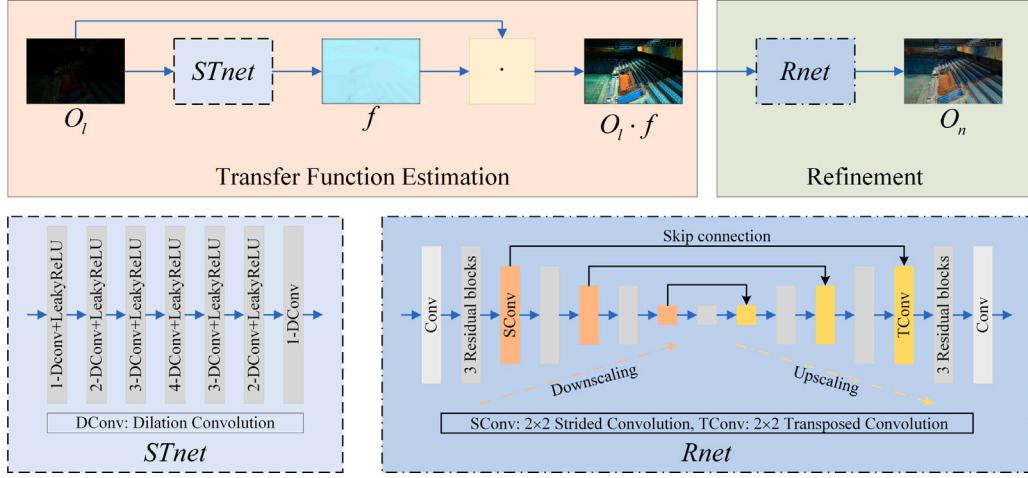
#### 3.2.2. Refinement module

The refinement module mainly aims to remove the noise in the coarsely enhanced image  $\mathbf{O}_l \cdot f$  and gain the final enhanced image  $\mathbf{O}_n = Rnet(\mathbf{O}_l \cdot f; \theta_R)$ . The training can be formulated as

$$\min_{\theta_R} \mathcal{L}_R(\mathbf{O}_n, \mathbf{O}_{gt}), \quad (21)$$

where  $\mathcal{L}_R$  is the loss function, defined as

$$\mathcal{L}_R = \lambda_2 \mathcal{L}_{ssim} + \lambda_3 \mathcal{L}_{color}. \quad (22)$$



**Fig. 3.** Proposed network SATNet for LLIE, consisting of two modules: *STnet* for transfer function estimation, and *Rnet* for refinement. The solid line indicates the data flow in the network.

The  $\mathcal{L}_{ssim} = 1 - SSIM(\mathbf{O}_{gt}, \mathbf{O}_n)$  aims to maximize the SSIM (Wang et al., 2004) between  $\mathbf{O}_{gt}$  and  $\mathbf{O}_n$ , which essentially enforces the consistency of the two images in luminance, contrast, and structure. To ensure color consistency between  $\mathbf{O}_{gt}$  and  $\mathbf{O}_n$ , we define the following color consistency loss,

$$\mathcal{L}_{color} = \frac{1}{W} \sum_{k=1}^W \sum_{c \in \{r,g,b\}} \|m_{\mathbf{O}_{gt}(c)}(P_k) - m_{\mathbf{O}_n(c)}(P_k)\|_2^2, \quad (23)$$

where  $P_k$  denotes a patch of size  $4 \times 4$ ,  $W$  denotes the total number of such non-overlapping patches,  $m$  denotes the mean value of the  $c \in \{r,g,b\}$  channel of an image in a patch. It is easy to see that,  $\mathcal{L}_{color}$  enforces the local consistency of each color channel between the estimated normal light image and ground-truth.

### 3.3. Network architecture and implementation details

The architecture of *STnet* is depicted in Fig. 3, which comprises 7 convolutional layers and the LeakyReLU activation function. Each convolutional layer employs a  $3 \times 3$  kernel size and is followed by a LeakyReLU activation with a negative slope of 0.2. *STnet* utilizes dilation convolutions with increasing dilation rates (1, 2, 3, 4, 3, 2) in successive layers to capture both local and global context information effectively.

For the refinement network *Rnet*, we adopt a U-Net (Ronneberger et al., 2015)-like architecture with channel and spatial attention mechanisms proposed by Woo et al. (2018). The details are illustrated in Fig. 3. It begins with a single convolutional layer, followed by 3 down-sampling operations, each consisting of 3 residual blocks and a strided convolution. Then, 3 residual blocks follow. Next, there are 3 up-sampling operations, each containing a transposed convolution and 3 residual blocks. Finally, there is a single convolutional layer. The residual blocks consist of Conv+ReLU+Conv. The channel configuration changes from 3 to 64, 128, 256, 512, 256, 128, 64, and 3. Between the down-sampling and up-sampling operations, there are skip connections. Note that the first convolutional layer and the residual blocks are equipped with channel and spatial attention mechanisms respectively.

We implement our method with PyTorch on a NVIDIA GTX 1660 GPU. We use RTV model (Xu et al., 2012) to get  $S_l$  and  $S_{gt}$  with parameters set to 0.0005 and 0.005 respectively. For every network in the proposed method, the batch size is set to 10, the patch size is  $112 \times 112$ . As shown in Fig. 4, we train our proposed network with a two-stage strategy: first, *STnet* is pre-trained by minimizing the loss  $\mathcal{L}_{ST}$  in Eq. (17); then *STnet* and *Rnet* are trained jointly by minimizing the loss  $\mathcal{L}_R$  in Eq. (22). The learning rate of *STnet* is set to  $5e-5$  in

the first stage. The learning rates of *STnet* and *Rnet* are set to  $5e-5$  and  $1e-4$  in the second stage. We train our method for 200 epochs and evaluate it every 20 epochs. We use Adam optimizer (Kingma and Ba, 2015) with default parameters. The tuning parameters  $\lambda_1, \lambda_2, \lambda_3$  are set to 5, 60, 100 empirically.

## 4. Experiments

In this section, we analyze our method empirically and verify its performance on datasets with/without ground-truth. Note that, the transfer function estimated by our SATNet relies on the training data. The test data may be obtained in different exposures. To enhance the adaptability of the estimated transfer function to different test data, we employ Gamma correction to the estimated transfer function. So in the test phase, the final enhanced image is obtained by

$$\mathbf{O}_n = Rnet(\mathbf{O}_l \cdot f^\gamma; \theta_R), \quad (24)$$

where  $\gamma$  is a user tuning parameter.

In Section 4.1, we provide the experiment settings, include the datasets, the baselines and the evaluation metrics. In Section 4.2, we analyze the transfer function estimation strategy of *STnet* and conduct ablation studies on the refinement module *Rnet*, the SSIM loss  $\mathcal{L}_{ssim}$  and the proposed color consistency loss  $\mathcal{L}_{color}$ .

### 4.1. Experiment settings

**(1) Datasets.** we evaluate the performance of our method on four datasets with ground-truth: LOL (Wei et al., 2018), MIT-Adobe 5K (Bychkovsky et al., 2011), RLMP (Chen et al., 2022) and UIEB (Li et al., 2019), and three datasets without ground-truth: LIME (Guo et al., 2016), NPE (Wang et al., 2013) and DICM (Lee et al., 2013). LOL (Wei et al., 2018) contains 500 low/normal light image pairs, with 485 paired images for training and 15 paired images for testing. MIT-Adobe 5K dataset (Bychkovsky et al., 2011) contains 5000 low-light/reference image pairs, and we randomly select 500 image pairs for evaluation. RLMP (Chen et al., 2022) were captured by three mobile phone models: Huawei Mate 10, Huawei P30, and Vivo NEX. There are 30 image pairs with both high-resolution ( $3952 \times 2960$ ) and low-resolution ( $800 \times 600$ ) versions for testing. We evaluate our proposed model on the low-resolution ( $800 \times 600$ ) version. Note the exposure is uneven, resulting in noise with non-uniform distributions. UIEB (Li et al., 2019) consists of 950 real-world underwater images, among which 890 images have corresponding reference images. We use the 890 images for evaluation. The non-reference datasets LIME (Guo et al., 2016), NPE (Wang et al., 2013) and DICM (Lee et al., 2013)

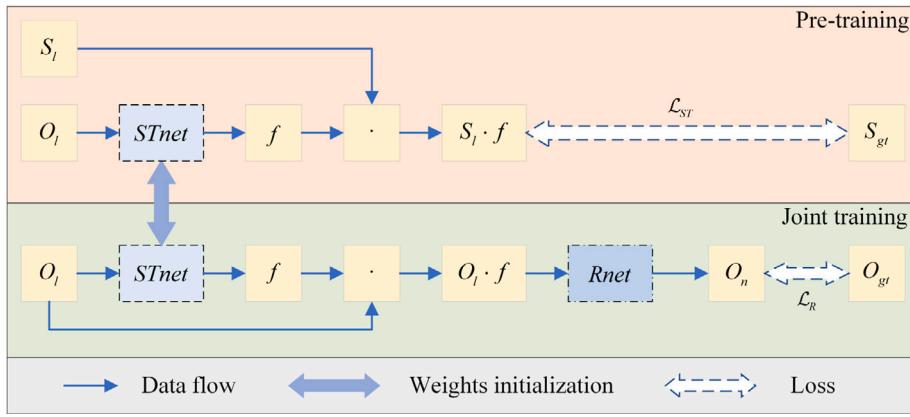


Fig. 4. The training strategy for the proposed network SATNet.

contain 10, 84 and 64 images respectively. Note that the proposed method is trained only on the LOL dataset (Wei et al., 2018), and tested on other datasets without retraining to validate its generalization performance. We provide the results with  $\gamma = 1$  and 1.2 for MIT-Adobe 5K (Bychkovsky et al., 2011),  $\gamma = 0.4$  for RLMP (Chen et al., 2022) and  $\gamma = 1$  for UIEB (Li et al., 2019). For the non-reference datasets LIME (Guo et al., 2016), NPE (Wang et al., 2013) and DICM (Lee et al., 2013), we provide the results with  $\gamma = 0.8, 1$  and 1 respectively.

**(2) Baselines.** We compare the proposed method with the state-of-the-art learning-based methods, including UTVNet (Zheng et al., 2021), STANet (Xu et al., 2022a), RetinexNet (Wei et al., 2018), KinD (Zhang et al., 2019a), KinD++ (Zhang et al., 2021), URetinex-Net (Wu et al., 2022), DeepUPE (Wang et al., 2019), and TBEFN (Lu and Zhang, 2020). Since RetinexNet (Wei et al., 2018), KinD (Zhang et al., 2019a), KinD++ (Zhang et al., 2021), URetinex-Net (Wu et al., 2022) and TBEFN (Lu and Zhang, 2020) were trained on LOL (Wei et al., 2018), we utilize their provided models for evaluation. DeepUPE (Wang et al., 2019) was originally trained on MIT-Adobe 5K (Bychkovsky et al., 2011). For fairness, we retrain it on LOL (Wei et al., 2018), then use it for evaluation. UTVNet (Zheng et al., 2021) was originally trained on sRGB-SID (Xu et al., 2020b), which contains 4,198 image pairs for training and 1,196 image pairs for testing. Since the training code is unavailable, we utilize their provided model directly for evaluation.

**(3) Evaluation Metrics.** We employ several well-established metrics for quantitative comparison, including 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). The PSNR, SSIM, and LPIPS assess similarity between the enhanced image and the ground-truth, while DeltaE measures color differences. LOE<sub>ref</sub> and LOE (Wang et al., 2013) are metrics specially designed for evaluating LLIE methods, by using the ground truth, and without using the ground-truth, respectively. Higher values of PSNR and SSIM indicate better quality, whereas lower values of LPIPS, DeltaE, LOE<sub>ref</sub>, and LOE indicate better quality.

#### 4.2. Model analysis and ablation study

In this section, we analyze the transfer function estimation strategy of *STnet* and conduct ablation studies on the refinement module *Rnet* and loss functions.

##### 4.2.1. Transfer function estimation

Our major contribution is learning the transfer function by *STnet* with  $(S_l, S_n)$ . To verify its effectiveness and advantage, we train *STnet* with  $(O_l, O_n)$  and  $(S_l, S_n)$ , respectively, on LOL (Wei et al., 2018). Then we fix the parameters of *STnet* and train *Rnet* by optimizing the loss  $L_R$  in Eq. (22). The average PSNR and SSIM of the enhanced results for test images in LOL (Wei et al., 2018) are shown in the second and

Table 2

Average metrics over 15 test images from LOL (Wei et al., 2018).

Methods	PSNR	SSIM
Training <i>STnet</i> with $(O_l, O_n)$	21.4258	0.8381
Training <i>STnet</i> with $(S_l, S_n)$	21.5304	0.8442
SATNet without <i>Rnet</i> ( $O_l \cdot f$ )	16.6382	0.4328
$\hat{O}_l \cdot f$	17.8091	0.7390
TBEFN (Lu and Zhang, 2020)	17.3510	0.7823
SATNet without $L_{ssim}$	11.7871	0.0784
SATNet without $L_{color}$	20.5594	0.8301
SATNet with $L_{ssim} + L_{mse}$	21.5059	0.8413
SATNet with $L_{ssim} + L_{color}$	<b>21.6981</b>	<b>0.8437</b>

the third rows of Table 2. It can be observed that the proposed strategy gains higher performance, demonstrating its advantage in estimating the transfer function.

##### 4.2.2. Ablation study

**(1) Effect of the refinement module *Rnet*.** To validate the effectiveness of the refinement module *Rnet*, we conduct test of our SATNet over the 15 test images. Table 2 and Fig. 5 respectively show the quantitative and the visual results obtained  $O_l \cdot f$  (SATNet without *Rnet*),  $\hat{O}_l \cdot f$ , TBEFN (Lu and Zhang, 2020) and *Rnet*( $O_l \cdot f$ ) (SATNet with *Rnet*).  $\hat{O}_l$  is obtained by applying the denoising algorithm BM3D (Dabov et al., 2007) to  $O_l$  and  $f$  is estimated by *STnet* with  $(S_l, S_n)$ . It can be observed that, the enhanced image  $O_l \cdot f$  has good lightness, yet it contains strong noise. This is consistent with the equation  $O_n = O_l \cdot f - N \cdot f$ , where the nearly invisible noise hidden in the low-light image is amplified by the transfer function. The enhanced image has comparable metrics with TBEFN (Lu and Zhang, 2020), suggesting that the transfer function estimated by the *STnet* is good and the noise mainly comes from the original low light image. Thus the refinement module *Rnet* plays a critical role in suppressing the noise.

**(2) Effect of the loss functions.** This section verifies the effectiveness of the SSIM loss  $L_{ssim}$  and the proposed color constancy loss  $L_{color}$ . The average PSNR and SSIM of the enhanced images from LOL dataset (Wei et al., 2018) are reported in the sixth and the seventh rows of Table 2. The visual results are shown in Fig. 6. We can note that,  $L_{ssim}$  plays dominating role in recovering the geometric structure and colors from the low light image, while  $L_{color}$  plays a supplementary role in that. We also provide comparison between the proposed color constancy loss and the MSE loss. It can be seen that the proposed color constancy loss obtains better metrics than the MSE loss. This improvement can be attributed to the fact that the proposed color constancy loss considers patch-level distance, whereas the MSE loss penalizes pixel-level distance.



Fig. 5. Visual comparison of the ablation study on the refinement module.



Fig. 6. Visual comparison of enhanced images obtained by our SATNet trained by using either  $L_{ssim}$ ,  $L_{color}$ ,  $L_{ssim} + L_{mse}$  and  $L_{ssim} + L_{color}$ .

## 5. Discussion

In Section 5.1, we evaluate the performance of our method on four datasets with ground-truth: LOL (Wei et al., 2018), MIT-Adobe 5K (Bychkovsky et al., 2011), RLMP (Chen et al., 2022) and UIEB (Li

et al., 2019), and three datasets without ground-truth: LIME (Guo et al., 2016), NPE (Wang et al., 2013) and DICM (Lee et al., 2013). In Section 5.2, application of the proposed method in X-ray image enhancement and lung segmentation is provided. In Section 5.3, the limitation of the proposed method is analyzed.

**Table 3**

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

Method	PSNR	SSIM	LPIPS	DeltaE	LOE	$LOE_{ref}$
UTVNet (Zheng et al., 2021)	15.6155	0.7267	0.3275	10.4062	195.3877	261.5160
STANet (Xu et al., 2022a)	<b>22.0524</b>	0.8195	0.2885	<b>9.5078</b>	192.8302	224.6847
RetinexNet (Wei et al., 2018)	16.7740	0.4285	0.4667	10.5000	475.0293	460.1811
KinD (Zhang et al., 2019a)	17.6476	0.7759	0.2752	10.1641	335.2898	399.5347
KinD++(Zhang et al., 2021)	17.7518	0.7690	0.3079	10.1719	683.0142	700.7935
URetinex-Net (Wu et al., 2022)	20.1096	<u>0.8255</u>	<u>0.2407</u>	10.0391	<u>130.4353</u>	<u>202.6324</u>
DeepUPE (Wang et al., 2019)	17.5489	0.5357	0.4025	10.4453	339.2532	379.3071
TBEFN (Lu and Zhang, 2020)	17.3510	0.7823	0.3243	10.2891	331.0364	324.9512
SATNet	21.6981	<b>0.8437</b>	<u>0.2663</u>	<u>9.6875</u>	<b>125.4189</b>	<b>171.9630</b>

**Table 4**

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

Method	PSNR	SSIM	LPIPS	DeltaE	LOE	$LOE_{ref}$
UTVNet (Zheng et al., 2021)	<u>16.7205</u>	0.7149	0.3358	10.4453	235.3516	253.6157
STANet (Xu et al., 2022a)	12.6228	<u>0.6627</u>	0.2982	10.3906	204.7837	208.0105
RetinexNet (Wei et al., 2018)	12.3053	0.6318	0.3027	10.1953	1377.3805	1352.7196
KinD (Zhang et al., 2019a)	15.5472	0.7303	<b>0.2003</b>	<u>10.1328</u>	446.1454	439.8536
KinD++(Zhang et al., 2021)	14.6387	0.7052	0.2141	<b>10.1172</b>	643.2417	632.6330
URetinex-Net (Wu et al., 2022)	13.8512	0.6817	0.2327	10.2266	322.9547	316.3901
DeepUPE (Wang et al., 2019)	15.4859	0.6680	0.3299	10.5703	425.8036	436.9561
TBEFN (Lu and Zhang, 2020)	14.3285	0.7017	<u>0.2050</u>	10.2109	519.9448	501.5762
SATNet( $\gamma = 1$ )	16.0894	<u>0.7421</u>	0.2354	10.1953	<b>95.2634</b>	<b>112.9223</b>
SATNet( $\gamma = 1.2$ )	<b>16.8396</b>	<u>0.7482</u>	0.2554	10.2188	<u>120.3595</u>	<u>144.8358</u>

**Table 5**

Average metrics over 30 test images from RLMP (Chen et al., 2022). The best results are in bold and the second best results are underlined.

Method	PSNR	SSIM	LPIPS	DeltaE	LOE	$LOE_{ref}$
UTVNet (Zheng et al., 2021)	<u>20.1185</u>	<u>0.7332</u>	<b>0.3846</b>	<b>9.9766</b>	117.4364	141.0401
STANet (Xu et al., 2022a)	19.3108	0.5977	0.4418	10.0781	<u>99.6759</u>	<u>125.1350</u>
RetinexNet (Wei et al., 2018)	18.6546	0.5250	0.4672	10.1953	474.5346	440.7726
KinD (Zhang et al., 2019a)	19.4481	0.6271	0.4270	10.0234	292.2018	298.1908
KinD++(Zhang et al., 2021)	19.3679	0.6220	0.4317	<u>9.9922</u>	405.4397	401.3843
URetinex-Net (Wu et al., 2022)	19.9453	0.5794	0.4469	10.0469	<b>95.7269</b>	<b>113.1207</b>
DeepUPE (Wang et al., 2019)	19.7753	0.5687	0.4482	10.0938	213.9652	221.8969
TBEFN (Lu and Zhang, 2020)	19.3602	0.6308	0.4280	<b>9.9766</b>	324.1894	302.5406
SATNet( $\gamma = 0.4$ )	<b>22.9520</b>	<u>0.8365</u>	0.4080	10.1484	129.6620	131.8010

### 5.1. Evaluation of the proposed method

Recall that the proposed model is trained only on LOL (Wei et al., 2018). In testing on LOL (Wei et al., 2018), we set  $\gamma = 1$ . The metrics of the enhanced images are reported in Table 3. We can note that, the proposed method outperforms the baselines in terms of SSIM[38], LOE (Wang et al., 2013), and  $LOE_{ref}$ , while obtains the second-best metrics of PSNR, LPIPS (Zhang et al., 2018) and DeltaE (McLaren, 1976). Fig. 7 provides visual comparison of a test image. It can be noted that, the enhanced images obtained by UTVNet (Zheng et al., 2021), STANet (Xu et al., 2022a), RetinexNet (Wei et al., 2018), UReticex-Net (Wu et al., 2022), DeepUPE (Wang et al., 2019) and TBEFN (Lu and Zhang, 2020) exhibit color deviation; the ones obtained by RetinexNet (Wei et al., 2018) and DeepUPE (Wang et al., 2019) exist noticeable noise; the ones obtained by KinD (Zhang et al., 2019a) and KinD++ (Zhang et al., 2021) look blurry; the ones obtained by UTVNet (Zheng et al., 2021) and TBEFN (Lu and Zhang, 2020) are not sufficiently brightened. In contrast, the proposed method performs effectively in enhancing the brightness, preserving colors and geometric structures, as well as suppressing noise.

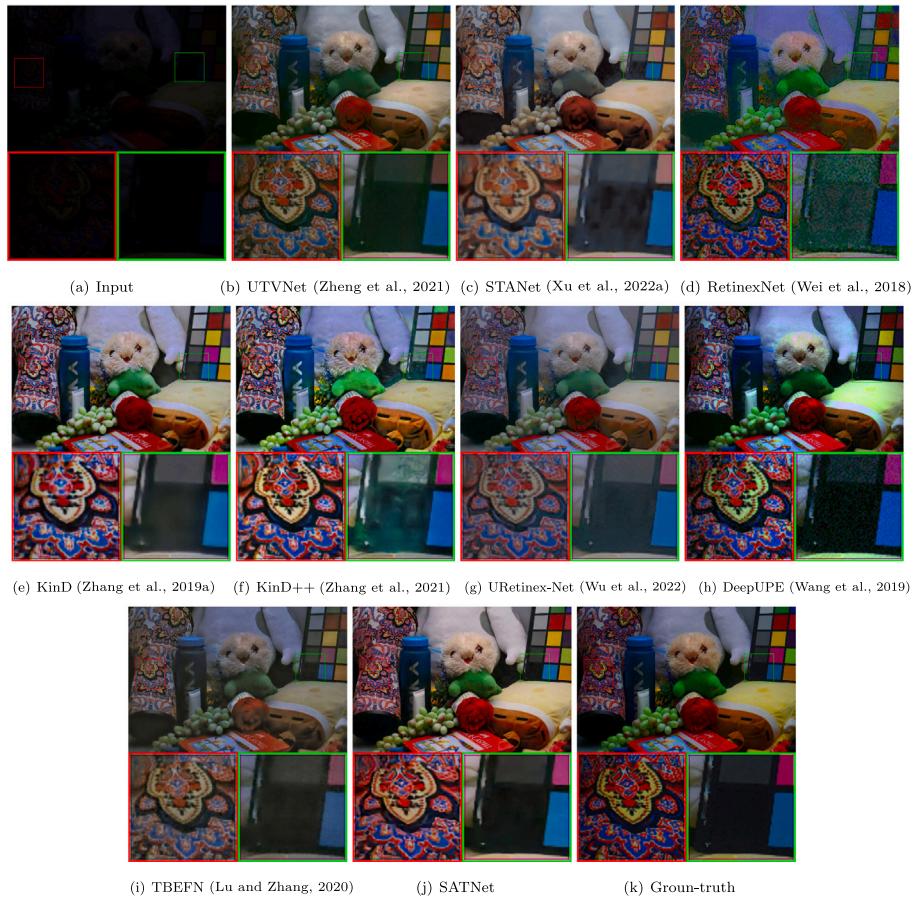
Table 4 presents the average metrics obtained by all methods on MIT-Adobe 5K (Bychkovsky et al., 2011) (without retraining). With  $\gamma = 1$ , the proposed method outperforms the baselines in terms of SSIM (Wang et al., 2004), LOE (Wang et al., 2013) and  $LOE_{ref}$ . To better adapt to the test data in MIT-Adobe 5K (Bychkovsky et al., 2011), we can fine-tune the estimated transfer function by adjusting  $\gamma$ . For example, setting  $\gamma = 1.2$ , the proposed method obtains higher

PSNR and SSIM (Wang et al., 2004). Fig. 8 shows visual comparison on a test image. It can be seen that, the enhanced images obtained by UTVNet (Zheng et al., 2021) and DeepUPE (Wang et al., 2019) have severe color deviation; the ones obtained by STANet (Xu et al., 2022a), RetinexNet (Wei et al., 2018), KinD (Zhang et al., 2019a), KinD++ (Zhang et al., 2021), UReticex-Net (Wu et al., 2022) and TBEFN (Lu and Zhang, 2020) cannot correctly brighten dark objects. In contrast, the proposed method obtains satisfying enhanced image.

Table 5 presents the average metrics obtained by all methods on the low light noisy dataset RLMP (Chen et al., 2022) (without retraining). With  $\gamma = 0.4$ , the proposed method outperforms the baselines in terms of PSNR and SSIM. Fig. 9 shows visual comparison on a test image. We can observe that, previous methods introduce local color deviation or noise, while the proposed method produces visually satisfying result.

Table 6 presents the average metrics obtained by all methods on the underwater dataset UIEB (Li et al., 2019) (without retraining). With  $\gamma = 1$ , the proposed method outperforms the baselines in terms of PSNR, DeltaE and  $LOE_{ref}$ . Fig. 10 shows visual comparison on a test image. It can be seen that, the enhanced images obtained by the baselines exhibit more or less color deviation in some local regions; while the enhanced image obtained by our method is closer to the ground-truth normal light image.

We also test the trained network on the three non-reference datasets (Guo et al., 2016; Wang et al., 2013; Lee et al., 2013). Figs. 11–13 show the enhanced images of sample images from the datasets LIME (Guo et al., 2016), NPE (Wang et al., 2013) and DICM (Lee et al., 2013) respectively. We can draw similar observations: the proposed



**Fig. 7.** Visual comparison on a test image from LOL (Wei et al., 2018). The objects in color boxes are zoomed in for better perception.



**Fig. 8.** Visual comparison on a test image from MIT-Adobe 5K (Bychkovsky et al., 2011). (a) Input; (b) UTVNet (Zheng et al., 2021); (c) STANet (Xu et al., 2022a); (d) RetinexNet (Wei et al., 2018); (e) KinD (Zhang et al., 2019a); (f) KinD++ (Zhang et al., 2021); (g) URetinex-Net (Wu et al., 2022); (h) DeepUPE (Wang et al., 2019); (i) TBEFN (Lu and Zhang, 2020); (j) SATNet( $\gamma = 1$ ); (k) SATNet( $\gamma = 1.2$ ); (l) Ground-truth. The objects in color boxes are zoomed in for better perception.



**Fig. 9.** Visual comparison on a test image from RLMP (Chen et al., 2022). The objects in color boxes are zoomed in for better perception.

**Table 6**

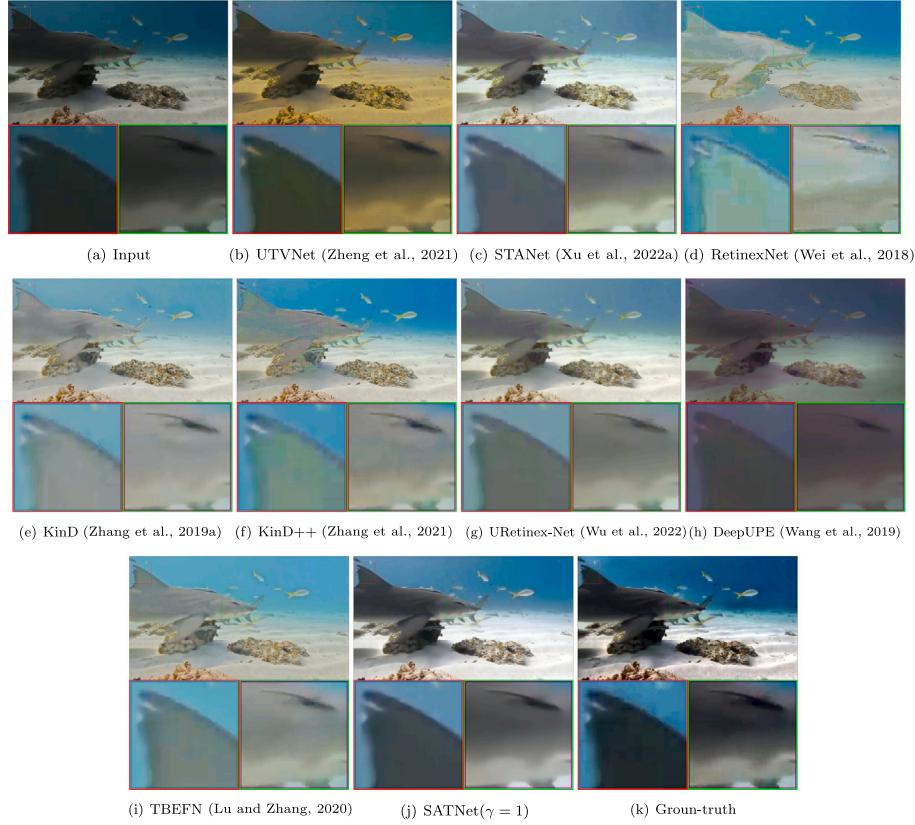
Average metrics over 890 test images from UIEB (Li et al., 2019). The best results are in bold and the second best results are underlined.

Method	PSNR	SSIM	LPIPS	DeltaE	LOE	LOE <sub>ref</sub>
UTVNet (Zheng et al., 2021)	13.0691	0.6556	0.4568	10.5938	652.2712	722.8677
STANet (Xu et al., 2022a)	11.7838	0.6441	0.4274	10.5781	451.1774	564.2830
RetinexNet (Wei et al., 2018)	11.9464	0.6570	0.3987	10.5703	472.4489	629.7689
KinD (Zhang et al., 2019a)	12.3384	0.6764	0.3817	10.5625	318.6282	450.8058
KinD++(Zhang et al., 2021)	12.5633	<u>0.6771</u>	0.3732	10.5547	412.5450	556.8265
URRetinex-Net (Wu et al., 2022)	12.5910	<b>0.6801</b>	<u>0.3429</u>	10.5547	<b>140.9932</b>	<u>335.1794</u>
DeepUPE (Wang et al., 2019)	12.6477	0.6513	0.3719	<u>10.5469</u>	656.2274	<u>739.4472</u>
TBEFN (Lu and Zhang, 2020)	12.3844	0.6546	<u>0.3701</u>	10.5547	548.6010	653.0820
SATNet( $\gamma = 1$ )	<b>13.6359</b>	0.6669	0.3940	<b>10.5234</b>	<u>153.7206</u>	<b>304.7554</b>

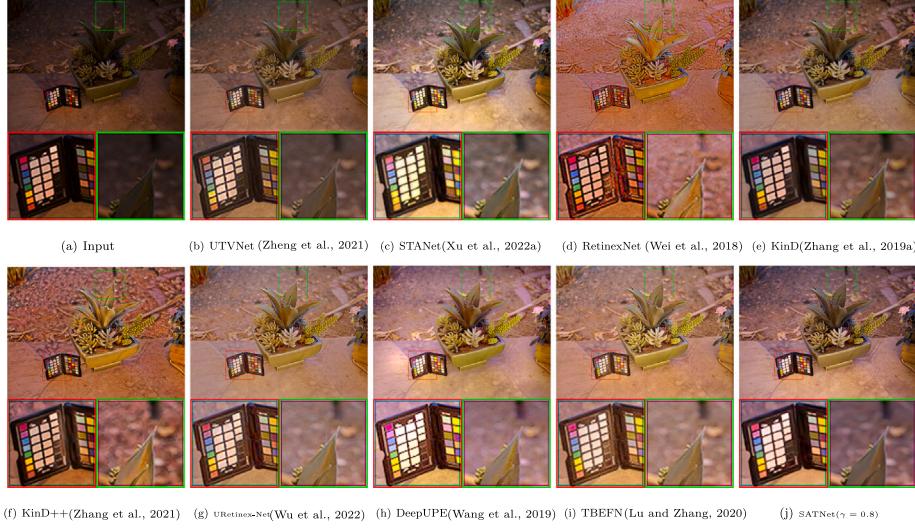
**Table 7**

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

Method	LIME (Guo et al., 2016)( $\gamma = 0.8$ )	NPE (Wang et al., 2013)( $\gamma = 1$ )	DICM (Lee et al., 2013)( $\gamma = 1$ )
UTVNet (Zheng et al., 2021)	168.9427	219.6094	313.4116
STANet (Xu et al., 2022a)	<u>111.9475</u>	243.1970	313.1666
RetinexNet (Wei et al., 2018)	542.6483	700.7733	447.6674
KinD (Zhang et al., 2019a)	167.0566	198.7513	189.7126
KinD++(Zhang et al., 2021)	353.7506	457.7775	337.2845
URRetinex-Net (Wu et al., 2022)	143.0156	<u>182.1935</u>	<u>178.2939</u>
DeepUPE (Wang et al., 2019)	269.4831	386.3850	485.8956
TBEFN (Lu and Zhang, 2020)	291.3698	380.8496	432.5450
SATNet	<b>107.7923</b>	<b>108.7870</b>	<b>145.7047</b>



**Fig. 10.** Visual comparison on a test image from UIEB (Li et al., 2019). The objects in color boxes are zoomed in for better perception.



**Fig. 11.** Visual comparison on a test image from LIME (Guo et al., 2016). The objects in color boxes are zoomed in for better perception.

method generates visually pleasant enhanced results, while previous methods cause color deviation, noise or artifacts. Table 7 displays the average LOE (Wang et al., 2013) metrics. It can be seen that, the metrics obtained by our method are significantly better than that obtained by the baselines.

The advantageous results of our method on these datasets indicate that our model has good generalization capability.

## 5.2. Application of the proposed method in X-ray image enhancement

We also provide an application of the proposed method in X-ray image enhancement. The chest X-ray images usually suffer from low contrast and noise, which bring challenges to lung segmentation. We first apply one baseline or the proposed method to enhance the chest X-ray image, then use a trained network to conduct lung



**Fig. 12.** Visual comparison on a test image from NPE (Wang et al., 2013). The objects in color boxes are zoomed in for better perception.



**Fig. 13.** Visual comparison on a test image from DICM (Lee et al., 2013). The objects in color boxes are zoomed in for better perception.

**Table 8**

Average IOU (Jaccard, 1912) and Dice (Dice, 1945) metrics on 20 chest X-ray images for segmenting lungs. The best results are in bold and the second best results are underlined.

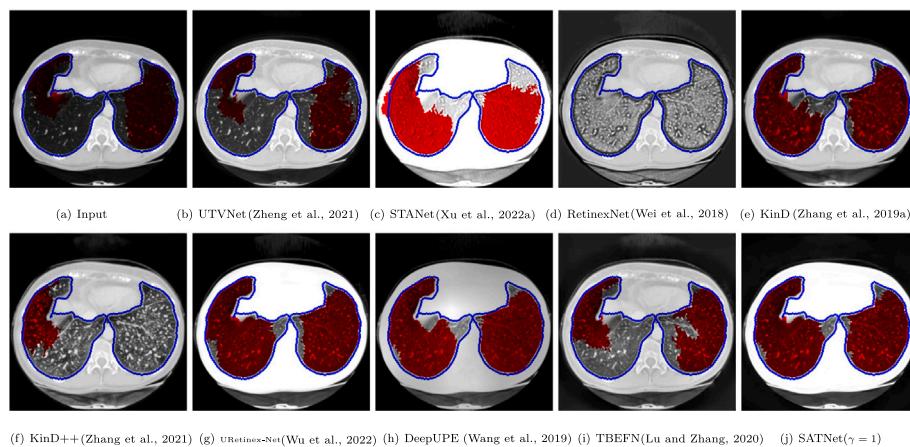
Methods	Original	UTVNet	STANet	RetinexNet	KinD
IOU (Jaccard, 1912)	0.1955	0.2167	0.3089	0.0200	0.4330
Dice (Dice, 1945)	0.2957	0.3386	0.3904	0.0330	0.5230
Methods	KinD++	URetinex-Net	DeepUPE	TBEFN	SATNet ( $\gamma = 0.5$ )
IOU (Jaccard, 1912)	0.1208	0.5805	<u>0.5960</u>	0.3256	<b>0.6032</b>
Dice (Dice, 1945)	0.2020	0.7158	<u>0.7227</u>	0.4717	<b>0.7362</b>

segmentation from the enhanced image. The dataset and the segmentation code can be accessed from <https://github.com/imlab-uiip/lung-segmentation-2d>. The quantitative and the visual results are presented in Table 8 and Fig. 14 respectively. Note that all LLIE methods are used without retraining, and we set  $\gamma = 0.5$  for our proposed method. As shown in Table 8, the proposed method leads to the best IOU (Jaccard, 1912) and Dice (Dice, 1945) metrics over 20 test images. Fig. 14 shows the normalized input and enhanced X-ray images, in which the ground-truth lung fields are outlined in blue and the predicted lung fields are filled in red. As we can see, the enhanced image by our

proposed method yields the best segmentation. This shows the potential application of our method in medical image enhancement.

### 5.3. Limitations of the proposed method

Table 9 presents the average running time, the number of trainable parameters, and the floating point operations (FLOPs) over 15 images of size  $600 \times 400 \times 3$ . The numbers of the trainable parameters and FLOPs of the transfer function estimation module are small, 12.483K and 2.972G, respectively. The large number of the trainable parameters and FLOPs of the whole network are caused by the high complexity of



**Fig. 14.** Visual comparison on a test image for lung segmentation. The input image and the enhanced images are normalized. The ground-truth lungs are outlined with the blue contour and the segmented lung regions are filled in red.

**Table 9**

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(G or T)	Platform
UTVNet (Zheng et al., 2021)	0.1787	7.745	83.083G	PyTorch
STANet (Xu et al., 2022a)	0.4861	0.472	103.228G	TensorFlow
RetinexNet (Wei et al., 2018)	0.5355	0.555	135.997G	TensorFlow
KinD (Zhang et al., 2019a)	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
DeepUPE (Wang et al., 2019)	0.1675	0.447	41.285G	PyTorch
TBEFN (Lu and Zhang, 2020)	0.1632	0.145	24.112G	PyTorch
SATNet	2.8808	25.004	405.313G	PyTorch

the refinement module. This is the major limitation of the proposed method and we are considering to design a lightweight refinement module in the future research.

## 6. Conclusion

Low-light image enhancement is demanding in many applications, including surveillance, photography and underwater imaging. Compared with Retinex-based methods, the transfer function provides a simpler pipeline for low-light image enhancement. However, the previous method has limitations in estimating the transfer function and the final enhancement. In this paper, we introduce a novel network SATNet for enhancing low-light images. The proposed network comprises two stages: transfer function estimation and refinement. In the first stage, we introduce a new strategy for estimating the transfer function from the structure of the low-light image. Experiments show its efficacy in estimating the transfer function. Then we train the refinement network and the transfer function estimation network jointly with the proposed color consistency loss. Ablation study confirms the effectiveness of the proposed color constancy loss in reducing color deviation. Experimental results validate the effectiveness of the trained refinement network in exhibiting strong capabilities in noise suppression and the preservation of both colors and details. Extensive experiments highlight that the proposed method outperforms state-of-the-art methods and exhibits robust generalization capability. The major limitation of the proposed network is that it contains more parameters and FLOPs, which will be the focus of our future research.

## CRediT authorship contribution statement

**Xiaofang Li:** Writing – original draft, Validation, Software, Methodology, Data curation. **Weiwei Wang:** Writing – review & editing, Supervision, Methodology, Funding acquisition, Conceptualization. **Yu Han:**

Writing – review & editing, Methodology. **Xiangchu Feng:** Writing – review & editing, Conceptualization.

## 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.

## 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).

## Data availability

Data will be made available on request.

## References

- 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, 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.
- Chen, Z., Jiang, Y., Liu, D., Wang, Z., 2022. CERL: A unified optimization framework for light enhancement with realistic noise. IEEE Trans. Image Process. 31, 4162–4172.
- Chen, J., Wang, Y., Han, Y., 2023a. A semi-supervised network framework for low-light image enhancement. Eng. Appl. Artif. Intell. 126, 107003.

- Chen, Y., Wen, C., Liu, W., He, W., 2023b. DBENet: Dual-branch brightness enhancement fusion network for low-light image enhancement. *Electronics* 12 (18), 3907.
- Cui, H., Li, J., Hua, Z., Fan, L., 2022. TPET: Two-stage perceptual enhancement transformer network for low-light image enhancement. *Eng. Appl. Artif. Intell.* 116, 105411.
- 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.
- Dice, L.R., 1945. Measures of the amount of ecologic association between species. *Ecology* 26 (3), 297–302.
- Fan, S., Liang, W., Ding, D., Yu, H., 2023. LACN: A lightweight attention-guided ConvNeXt network for low-light image enhancement. *Eng. Appl. Artif. Intell.* 117, 105632.
- 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.
- He, X., Chen, Z., Dai, L., Liang, L., Wu, J., Sheng, B., 2023. Global-and-local aware network for low-light image enhancement. *Eng. Appl. Artif. Intell.* 126, 106969.
- He, K., Sun, J., Tang, X., 2012. Guided image filtering. *IEEE Trans. Pattern Anal. Mach. Intell.* 35 (6), 1397–1409.
- Jaccard, P., 1912. The distribution of the flora in the alpine zone. *New Phytol.* 11 (2), 37–50.
- Jiang, Y., Gong, X., Liu, D., Cheng, Y., Fang, C., Shen, X., Yang, J., Zhou, P., Wang, Z., 2021. EnlightenGAN: Deep light enhancement without paired supervision. *IEEE Trans. Image Process.* 30, 2340–2349.
- 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., Ren, W., Cong, R., Hou, J., Kwong, S., Tao, D., 2019. An underwater image enhancement benchmark dataset and beyond. *IEEE Trans. Image Process.* 29, 4376–4389.
- Li, Y., Niu, Y., Xu, R., Chen, Y., 2023. Zero-referenced low-light image enhancement with adaptive filter network. *Eng. Appl. Artif. Intell.* 124, 106611.
- 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.
- Liu, X., Xie, Q., Zhao, Q., Wang, H., Meng, D., 2023. Low-light image enhancement by retinex-based algorithm unrolling and adjustment. *IEEE Trans. Neural Netw. Learn. Syst.*
- Lore, K.G., Akintayo, A., Sarkar, S., 2017. LLNet: A deep autoencoder approach to natural low-light image enhancement. *Pattern Recognit.* 61, 650–662.
- Lu, K., Zhang, L., 2020. TBEFN: A two-branch exposure-fusion network for low-light image enhancement. *IEEE Trans. Multimed.* 23, 4093–4105.
- Lv, F., Li, Y., Lu, F., 2021. Attention guided low-light image enhancement with a large scale low-light simulation dataset. *Int. J. Comput. Vis.* 129 (7), 2175–2193.
- Lv, F., Lu, F., Wu, J., Lim, C., 2018. MBLLEN: Low-light image/video enhancement using CNNs. In: Proc. Brit. Mach. Vis. Conf. p. 13.
- 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., Ericksen, J., Yankaskas, B., Muller, K., 1990. Contrast-limited adaptive histogram equalization: speed and effectiveness. In: Proc. 1st Conf. Visualiz. Biomed. Comput. pp. 337–345.
- 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.
- Tomasi, C., Manduchi, R., 1998. Bilateral filtering for gray and color images. In: Proc. IEEE Int. Conf. Comput. Vis. IEEE, pp. 839–846.
- Ule, L., 1955. Weighted least-squares smoothing filters. *IRE Trans. Circ. Theory* 2 (2), 197–203.
- 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, Y., Zhang, Z., 2023. Global attention retinex network for low light image enhancement. *J. Vis. Commun. Image Represent.* 92, 103795.
- Wang, R., Zhang, Q., Fu, C.-W., Shen, X., Zheng, W.-S., Jia, J., 2019. Underexposed photo enhancement using deep illumination estimation. In: Proc. IEEE Conf. Comput. Vis. Pattern Recog.. pp. 6849–6857.
- 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. p. 12.
- Woo, S., Park, J., Lee, J.-Y., Kweon, I.S., 2018. CBAM: Convolutional block attention module. In: Proceedings of the European Conference on Computer Vision. pp. 3–19.
- Wu, W., Weng, J., Zhang, P., Wang, X., Yang, W., Jiang, J., 2022. URetinex-Net: Retinex-based deep unfolding network for low-light image enhancement. In: Proc. IEEE Conf. Comput. Vis. Pattern Recog.. pp. 5901–5910.
- Xu, K., Chen, H., Xu, C., Jin, Y., Zhu, C., 2022a. Structure-texture aware network for low-light image enhancement. *IEEE Trans. Circuits Syst. Video Technol.* 32 (8), 4983–4996.
- Xu, J., Hou, Y., Ren, D., Liu, L., Zhu, F., Yu, M., Wang, H., Shao, L., 2020a. STAR: A structure and texture aware Retinex model. *IEEE Trans. Image Process.* 29, 5022–5037.
- Xu, X., Wang, R., Fu, C.-W., Jia, J., 2022b. SNR-aware low-light image enhancement. In: Proc. IEEE Conf. Comput. Vis. Pattern Recog.. pp. 17714–17724.
- Xu, L., Yan, Q., Xia, Y., Jia, J., 2012. Structure extraction from texture via relative total variation. *ACM Trans. Graph.* 31 (6), 1–10.
- Xu, K., Yang, X., Yin, B., Lau, R.W., 2020b. Learning to restore low-light images via decomposition-and-enhancement. In: Proc. IEEE Conf. Comput. Vis. Pattern Recog.. pp. 2281–2290.
- Yang, Q., Wu, Y., Cao, D., Luo, M., Wei, T., 2021. A lowlight image enhancement method learning from both paired and unpaired data by adversarial training. *Neurocomputing* 433, 83–95.
- 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, Y., Zhang, J., Guo, X., 2019a. Kindling the darkness: A practical low-light image enhancer. In: Proc. 27th ACM Int. Conf. Multimedia. pp. 1632–1640.
- Zhang, L., Zhang, L., Liu, X., Shen, Y., Zhang, S., Zhao, S., 2019b. Zero-shot restoration of back-lit images using deep internal learning. In: Proc. 27th ACM Int. Conf. Multimedia. pp. 1623–1631.
- Zhang, Z., Zheng, H., Hong, R., Xu, M., Yan, S., Wang, M., 2022. Deep color consistent network for low-light image enhancement. In: Proc. IEEE Conf. Comput. Vis. Pattern Recog.. pp. 1899–1908.
- Zheng, C., Shi, D., Shi, W., 2021. Adaptive unfolding total variation network for low-light image enhancement. In: Proc. IEEE Int. Conf. Comput. Vis. pp. 4439–4448.