Neural Processing Letters (2024) 56:93
https://doi.org/10.1007/s11063-024-11565-5

Rethinking Zero-DCE for Low-Light Image Enhancement

Aizhong Mi[1] · Wenhui Luo[1] · Yingxu Qiao[2] · Zhanqiang Huo[1]

Accepted: 11 February 2024 / Published online: 7 March 2024

Abstract

Zero-Reference Deep Curve Estimation (Zero-DCE) is currently one of the most popular low-light image enhancement methods. Through extensive experimentation, we observe that: (i) the excellent performance of Zero-DCE depends on the training data with multiple exposure levels, (ii) it cannot effectively handle uneven light, extremely low light, or overexposed images in natural environments. Therefore, we propose an improved zero-reference dual-illumination deep curve estimation method for low-light image enhancement named Zero-DiDCE, which can enhance, suppress, or maintain light levels for images. The adaptive light enhancement curve was designed to handle images with different exposure levels. An iterator and amplitude controller are designed to control the curve enhancement intensity by calculating the gap between the input image and the optimal light level. Furthermore, instead of the DCE-Net in Zero-DCE only taking the input image as network input, our DiDCE-Net in Zero-DiDCE takes the input image and the inverted input image simultaneously as network input to ensure that the training set contains samples with multiple exposure levels. A piecewise non-reference loss function is designed to guide the training of DiDCE-Net from the perspective of information loss. Qualitative and quantitative experiments show that our method can handle images with different levels of exposure well and outperforms state-of-the-art methods. In addition, the proposed curve and iterator can be integrated into other methods to improve their enhancement effects. The code is available at https://github.com/Wenhui-Luo/Zero-DiDCE.

Keywords Low-light image enhancement · Zero-reference · Dual-illumination · Curve estimation.

✉ Zhanqiang Huo
hzq@hpu.edu.cn

Aizhong Mi
miaizhong@hpu.edu.cn

Wenhui Luo
212109010014@home.hpu.edu.cn

Yingxu Qiao
qiaoyingxu@hpu.edu.cn

[1] School of Software, Henan Polytechnic University, Jiaozuo 454003, Henan, China

[2] School of Computer Science and Technology, Henan Polytechnic University, Jiaozuo 454003, Henan, China

1 Introduction

Many low-light images are degraded to varying degrees. Low-light image enhancement methods were proposed to solve such problems and applied to various fields [1].

Supervised learning methods mostly require a terrific amount of paired training data much-paired to achieve excellent performance. However, capturing many low-/normal-light images takes much work, limiting its application. Therefore, some scholars proposed to use of synthetic data. Nevertheless, synthetic data leads to poor generalization performance when processing images of natural scenes. To address these problems, some scholars have proposed unsupervised learning methods, such as GAN-based [2, 3] and zero-shot learning methods [4, 5], which can be trained without using paired data and have achieved good enhancement results. GAN-based methods [2, 3] use adversarial learning to achieve low-light enhancement, which requires careful training data selection to achieve better enhancement results. Zero-shot learning methods, such as Zero-DCE [4] and Zero-DCE++ [5] consider low-light enhancement as a image-specific curve estimation task with deep networks. These methods propose several loss functions to guide the model training and achieve image enhancement after a fixed curve iteration, as shown in Fig. 1.

The fixed iteration scheme results in limited dynamic adjustment ability and cannot handle different exposure images. Most of these methods require carefully selective images containing multiple exposure levels for training to achieve better enhancement results. This limitation on datasets reduces the robustness of the model. Many unsupervised methods are adept at enhancing ordinary low-light images but perform poorly on uneven and extremely light images. It is challenging to process extremely low light, overexposed, and uneven light images using one model, as shown in Fig. 2.

We propose a novel Zero-DiDCE algorithm to address the above problems. It not only has lower restrictions on training data but also can process images with different exposure levels, which significantly improves the model's generalization. The adaptive light enhancement curve evaluates and calculates the input image through amplitude controller, iterator, and DiDCE-Net and adjusts the curve according to the enhanced results to obtain the final image.

The adaptive light enhancement curve (ALE-curve) is designed to handle images with different exposure levels. ALE-curve performs different levels of enhance, suppress, or

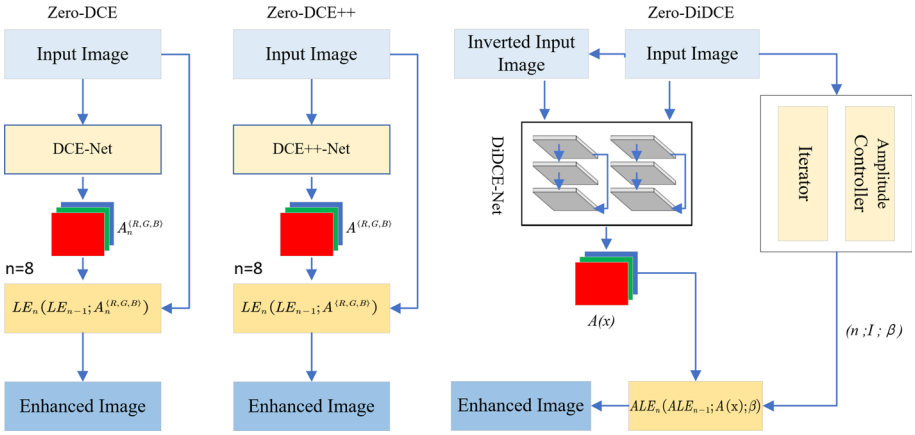

Fig. 1 Comparison of the overall framework of Zero-DCE, Zero-DCE++, and Zero-DiDCE

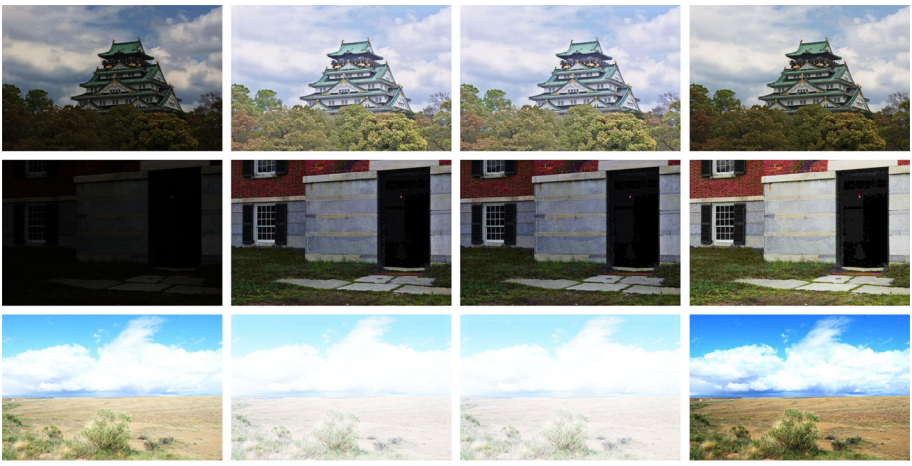

Fig. 2 Visual comparisons of Zero-DiDCE, Zero-DCE and Zero-DCE++. The Zero-DCE and Zero-DCE++ enhanced images appear overexposed and underexposed. Zero-DiDCE performs best in terms of image detail, color, and contrast, especially in terms of color, such as more realistic colors in house tiles, grass, and sky backgrounds

maintain operations by calculating the difference between the input image and the target light level. The iterator and amplitude controller can perform different numbers of iterations and enhancement amplitudes depending on the input image. The iterator and amplitude controller effectively avoid overexposure and underexposure problems in some images. The piecewise non-reference loss function and DiDCE-Net can enhance the learning ability of the model. The piecewise non-reference loss function guides the learning of the input image and inverse input image by DiDCE-Net based on the loss of input image lighting information. The improved learning capability further reduces the limitations of the training data. The iterator and ALE-curve can also be used for other algorithms.

Our contributions are summarized as follows:

- We propose a novel low-light enhancement method that is less restrictive for training data. It does not rely on paired or unpaired multi-exposure data, further avoiding the risk of overfitting and possessing stronger robustness.
- We propose a new adaptive light enhancement curve. The curve can perform different enhancement, suppression, or maintenance operations by determining the difference between the input image and the target light level.
- We propose an iterator, amplitude controller, and piecewise non-reference loss functions. The former two can develop appropriate enhancement or suppression amplitudes. The latter guides DiDCE-Net training based on the loss of light information in the input image.
- We have conducted numerous experiments, resulting in Zero-DiDCE outperforming other methods.

2 Related Works

The work in this paper focuses on extremely low light or overexposed images, which have been rarely covered in previous work. Related work is reviewed below.

2.1 Conventional Methods

Histogram Equalization (HE) performs nonlinear stretching of gray image values by adjusting the input image's histogram so that the gray level is evenly distributed over the entire gray range, achieving low-light enhancement by increasing the overall contrast of the image. However, conventional HE algorithms may result in overexposure of the image. Thus, some scholars have proposed brightness-maintaining [6] and contrast-limiting [7] schemes. Methods [8–13] based on Retinex theory have received increasing attention compared to methods with Histogram equalization. The Retinex model approach usually decomposes the low-illumination image into reflectance and illumination components by a priori or regularization methods, and the estimated reflectance component is considered the result of enhancement. Guo et al. [14] used the maximum intensity of pixels in the RGB channel to estimate the illumination map, which was then refined by a structural before achieving enhancement. Fu et al. [15] used a weighted change model to estimate the input image's reflectance and illumination. Retinex theory only sometimes holds, leading to color restoration.

2.2 Deep Learning Methods

Most deep learning-based methods have better enhancement effects than conventional methods. Lore et al. pioneered a deep learning-based LLIE method called LLNet [16]. After that, many methods have been proposed, including supervised learning methods [17–22], unsupervised learning methods [2], zero-shot learning methods [23–26], and semi-supervised learning methods [27, 28]. Zhu et al. [29] proposed an EEMEFN method to achieve enhancement through multiple exposure fusion and edge enhancement. Yang et al. [27] proposed a deep recursive band network (DRBN). DRBN obtained an improved representation by recombining the given bands. DRBN improved performance through long short-term memory (LSTM) networks [28]. Ma et al. [30] proposed a self-calibrating illumination framework for low-light image enhancement. Zhang et al. [23] proposed the ExCNet method, which used a network to estimate the S-curve of the input image and a bootstrap filter [31] to separate the base and detail layers. Weber contrast [32] blended the detail and adjusted the base layers for low-light enhancement. Zhang et al. [33] constructed a new lightweight architecture that works better on CPU devices. Peng et al. [34] presented a nonconvex method for denoising hyperspectral images that separates low-rank and sparse components more accurately. EnlightenGAN [2] used U-Net [35] as a generator and global–local discriminator to ensure the realism of the enhanced images. Since the training does not use paired training data, it has better robustness but requires careful training data selection. Chen et al. [36] proposed a new solution from the perspective of sparse representation by combining global metrics with traditional local metrics via deep fusion networks.

Zero-DiDCE is more capable of learning than Zero-DCE and Zero-DCE++. Firstly, DiDCE-Net improves the learning ability of Zero-DiDCE by learning images in dual illuminations. Secondly, the iterator and amplitude controller estimate the optimal curve, which improves the processing power of Zero-DiDCE while reducing the limitation of training data.

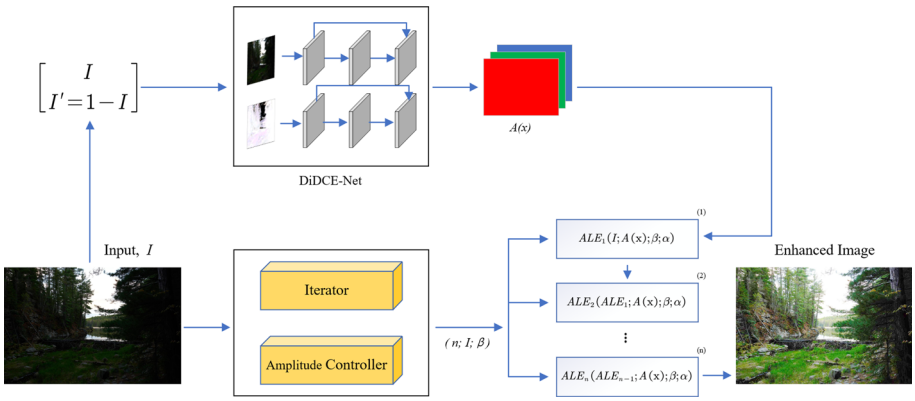

Fig. 3 Overall framework of Zero-DiDCE. DiDCE-Net, iterator, and amplitude controller estimate the best enhancement curve

3 Method

Zero-DiDCE has deeper zero-shot learning and can handle different exposure images. Zero-DiDCE is shown in Fig. 3.

Given the input image I, DiDCE-Net estimates the lighting information from the dual illuminations of the input image I and obtains the curve parameter maps for the adaptive light enhancement curve. The iterator and amplitude controller will be designed as a targeted enhancement scheme based on the light level of I and further adjust the adaptive light enhancement curve precisely to achieve the best enhancement or suppression effect. Iterate the ALE-curve according to the enhancement scheme and curve parameter map to obtain an enhanced image.

3.1 Adaptive Light Enhancement Curve

In order to reduce the model's dependence on training data and handle images with different exposure levels, we design an adaptive light enhancement curve. The ALE-curve is highly flexible and dynamically adjustable, allowing the light level of an image to be enhanced or suppressed to a specified interval by iteration. ALE-curve is expressed as:

$$\begin{aligned} ALE_n(x) &= ALE_{n-1}(x) \\ &\quad + A(x)\left(ALE_{n-1}(x)^2 - ALE_{n-1}(x)\right)\frac{\alpha - ALE_{n-1}(x)}{\beta - ALE_{n-1}(x)} \end{aligned} \tag{1}$$

where $ALE_n(x)$ is the result of the enhancement $ALE_{n-1}(x)$, $ALE_0(x)$ represents the initial input image. In this paper, n is dynamic, the iterator controls the size of n, and n controls the number of iterations. $A(x)$ is a curve parameter map, and β controls the enhancement amplitude. Naturally, we consider the appropriate exposure level for an image to be around 0.6. The algorithm uses $\alpha = 0.63$ as the target light level of the image.

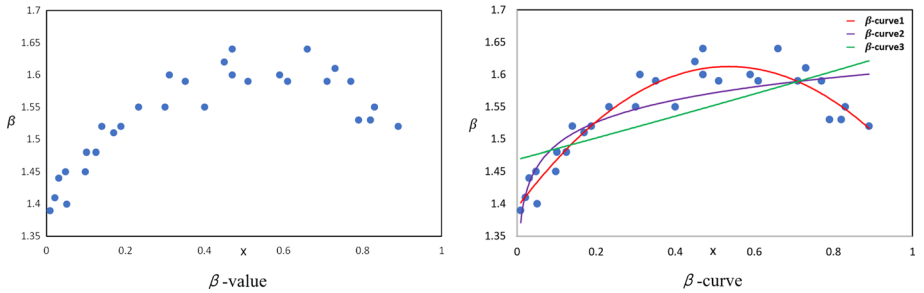

Fig. 4 The horizontal axis indicates the light level of the image, while the vertical axis indicates the appropriate enhancement or suppression amplitude. β-value is a representative amplitude distribution. Considering that it is impossible to visualize all amplitude values of 1000 images, representative amplitude values are selected to facilitate the observation of their distribution characteristics. β-curve indicates the amplitude curves

3.2 Amplitude Controller

The amplitude of enhancement or suppression required for images with different exposure levels is different, and using the same amplitude will affect the quality of the processed images. The optimal enhancement or suppression amplitude required for different exposure images was determined through extensive experiments. LOL [37] contains only a large number of standard and low-light images, and SICE [38] contains a wide range of images with different exposure levels, so they were chosen for the experiment. One thousand images from the SICE [38] and LOL [37] datasets with different exposure levels were selected for the experiment. The enhancement and suppression amplitude values of the images with suitable different exposure levels were found and recorded by many tests. Considering the unavoidable differences between image samples, the enhancement or suppression amplitude of images with the same exposure level has a specific pattern to follow. The amplitude values of the different distributions are fitted by a simple and effective amplitude curve, as shown in Fig. 4.

Three different curves were designed according to the distribution of amplitude values. The best-fit $\beta - curve1$ is selected as the fitting curve to calculate the best enhancement or suppression amplitude required for the input images with different exposure levels. Not only that, a reasonable enhancement amplitude can effectively avoid amplification noise. The amplitude curve is expressed as:

$$\beta = -0.79x^2 + 0.81x + 1.41 \tag{2}$$

3.3 Iterator

Zero-DCE and Zero-DCE++ use eight iterations to enhance images. They do not consider the processing of images with different exposures, so the generality of these methods is limited. The iterator is designed to select the appropriate number of iterations for the input images. The design of this module is similar to the amplitude controller, fitting the iterative scheme required for differently exposed images through an iterative curve, which is denoted by I_{iter}.

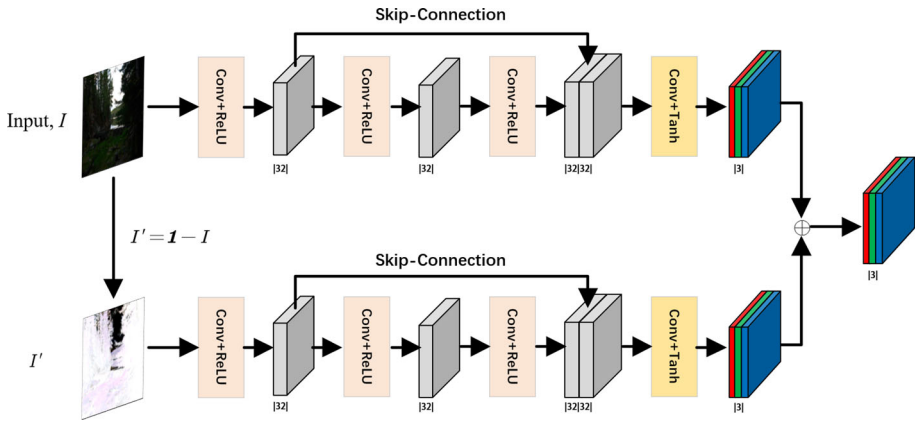

Fig. 5 The architecture of DiDCE-Net. The convolutional layer consists of 32 convolutional kernels. The size of the convolution kernel is 3×3, and the stride is 1. ⊕ represents element-wise addition

The I_{iter} is defined by the following equation:

$$I_{iter} = \begin{cases} -25x + 10, x < 0.1 \\ 17.14x^2 - 15.14x + 10, 0.1 \leq x < 0.45 \\ 5.66x^2 - 2.93x + 7.2, x \geq 0.45 \end{cases} \tag{3}$$

$$n = \lfloor I_{iter} \rfloor \tag{4}$$

where x denotes the normalized image pixel mean value, n is the number of iterations. Given the input image, the iterator extracts the lighting information and develops a suitable iteration scheme.

3.4 Dual-Illumination Deep Curve Estimation Network

Most current deep-learning methods tend to extract the lighting information of the input image. However, the reversed input image also contains much lighting information. In the input image after reversal, underexposed regions are represented as overexposed, and overexposed regions will be represented as underexposed. Therefore, this method uses the input image and the inverted input image together to represent the light estimation problem of the image. In this way, the model can acquire information on different light intensities in one input image, reducing the model's limitation on the training data and improving the model's generality. DiDCE-Net can extract the lighting information of the input image and the inverted input image separately, thus allowing the model to learn richer lighting information. The architecture of the DiDCE-Net is shown in Fig. 5.

Firstly, after image I input, the inverted image (I') is calculated by $I' = 1 - I$. Secondly, a 4-layer side-by-side ordered ordinary CNN convolution is executed on the input and inverted input images. The first 3 convolution layers consist of 32 convolution kernels of size 3 × 3 with a step size of 1, where a ReLU activation function follows each layer. The last layer is a Skip-Connection convolutional layer with a Tanh activation function. Finally, the two output three-channel (RGB) curve parameter maps are combined to obtain each iteration's final parameter map $A(x)$. The 4-layer convolutional structure is very lightweight, which is a significant advantage on devices with limited computational resources.

3.5 Piecewise Non-reference Loss Function

Low-light images have a certain degree of information loss, especially those with extremely light. The piecewise non-reference loss function can distinguish between extremely light images with more information loss and general low-light images with less information loss and perform different losses separately.

The piecewise non-reference loss function (L) contains two parts of loss with different weights ($W_{1,2}$ for weights), determined by the regional light quality of the input image to perform different losses as a function. L is expressed as:

$$L = \begin{cases} W_1 L_1, Y_i \leq Q_1 or Y_i \geq Q_2 \\ W_2 L_2, Y_i \in (Q_1, Q_2) \end{cases} \tag{5}$$

In the first part, when the local area of the input image is extremely low light or overexposed, the difference between the light quality of its local area and good light is evaluated, i.e., L_1. (Y_i indicates the area light quality of the input image. $Q_{1,2}$ is the threshold value ($Q_{1,2} \in (0.2, 0.8)$) that distinguishes between extremely low light or overexposed images. When $Y_i \leq Q_1$ or $Y_i \geq Q_2$, it indicates an extremely low light or overexposure and performs L_1 loss. Otherwise, it indicates an area of general light and performs L_2). L_1 is expressed as:

$$L_1 = \frac{1}{K} \sum_{e=1}^{K} (Y_e - E)^2 \tag{6}$$

where K is the number of local areas of size 16×16, and Y_e denotes the regional light quality of the enhanced image. E is the well-lighted value, set to the gray level in RGB color space.

In the second part, when the input image is in the general light interval ($Y_i \in (Q_1, Q_2)$), the difference between the light quality of the local area of the image and the well-light quality is evaluated, i.e., L_2. m here is responsible for controlling L_2 the weights. L_2 is expressed as:

$$L_2 = \frac{1}{K} \sum_{i=1,e=1}^{K} (Y_e - E)^2 \frac{1}{1 + mY_i} \tag{7}$$

4 Experiments

4.1 Implementation Details

Firstly, to demonstrate that Zero-DiDCE is less limited to data. This experiment uses two datasets with different exposure levels, and the following conditions must be satisfied:

1. The number of identical scene images is the same between the two datasets.
2. There are no images other than the images of the same scene between the datasets.
3. There is only variation in the exposure level of the images.

Two datasets with the same scene but different exposures are designed on the LOL dataset [37] for the experiments, with an image size of 400×600 and jpg format. In the first one, the normal light image of the LOL dataset is selected twice for training to verify the learning ability of the model in the absence of low-light images, called Lack-Low. In the second one, the 485 low-light images are selected twice as the dataset to verify the learning ability of the model when using a dataset lacking exposure information, called Lack-Exposure.

Secondly, the LOL dataset [37] is selected for quantitative experiments. The algorithm is also visually compared on the SICE [38] and DICM datasets. SICE [38] and DICM contain

Fig. 6 Zero-DiDCE applies a visual comparison of different curves. Zero-DiDCE-LE indicates Zero-DiDCE with the LE-curve. Zero-DiDCE-ALE indicates Zero-DiDCE with the ALE-curve

images with multiple exposure levels and can visualize the enhancement effect of Zero-DiDCE on images with different exposures.

This experiment uses NVIDIA A100 GPU. The batch size is set to 8, and the network is optimized by training 100 cycles from zero, using an ADAM optimizer with a learning rate of $1e^{-4}$. We used three metrics Peak Signal-to-Noise Ratio (PSNR), Structural Similarity (SSIM) [39], and Neural Image Assessment (NIMA), for quantitative comparison. The higher the PSNR value, the less distortion in the image; the higher the NIMA score, the better the image. SSIM is a measure of the similarity between two images, and the higher the value, the smaller the degree of loss of the image.

4.2 Ablation Studies

We design ablation experiments to prove the validity of each part of the model by replacing or removing ALE-curve, iterator, and amplitude controller. Some of the methods are applied to other networks to see if they can improve the performance of similar networks.

Advantage of adaptive light enhancement curve. The LE-curve replaces the ALE-curve in Zero-DiDCE, and Fig. 6 shows the results of both visualizations. The LE-curve is expressed as:

$$LE_n(x) = LE_{n-1}(x) + A(x)\left(LE_{n-1}(x) - LE_{n-1}(x)^2\right) \tag{8}$$

Zero-DiDCE-ALE performs well, and its enhanced images are naturally exposed with clear details. Zero-DiDCE-LE image overexposure. ALE-curve is irreplaceable in the model.

ALE-curve can enhance the performance of other models. The ALE-curve is used instead of the LE-curve in Zero-DCE++ [5]. Since the remaining parameters of Zero-DCE++ cannot be changed, we plan to use the ALE-curve instead of the LE-curve of Zero-DCE++ when the light level of the enhanced image is higher than 0.7 ($\alpha = 0.65$, $\beta = 1.95$). The experiment is shown in Fig. 7.

Zero-DCE++-ALE has the best enhancement results with more natural lighting. ALE-curve greatly enhances Zero-DCE++'s ability to handle different exposure levels of images and improves the model's generality.

Effect of target light level (α). We evaluated the effect of the target light level. Three sets of significantly different α values (0.43, 0.63, and 0.83) were used in this experiment to facilitate the observation of the differences among them. Figure 8 illustrates a visual comparison of different α values.

The algorithm allows flexible adjustment of the image exposure level by setting different α. Thus, Zero-DiDCE can synthesize low-light datasets with different exposure levels using the ability of α to adjust the images. The potential of Zero-DiDCE for synthesizing low-light datasets will be explored in the future.

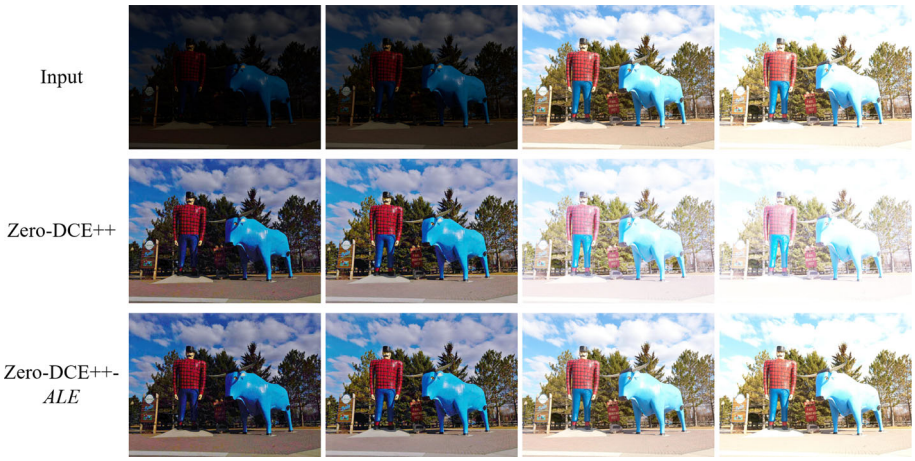

Fig. 7 Visual comparison of Zero-DCE++ [5] before and after using ALE-curve. Zero-DCE++-ALE indicates the visible result of Zero-DCE++ after using the ALE-curve

Fig. 8 Comparison results using different values of α. Zero-DiDCE-α indicates the results of the runs using different values of α

Fig. 9 Visual comparison of different α values

In addition, three sets of α values (0.53, 0.63, and 0.73) were designed for this experiment to illustrate the excellence of α selection values, as shown in Fig. 9.

The grass is slightly overexposed when $\alpha = 0.73$ and the grass color is slightly whitish. When $\alpha = 0.53$, the grass is underexposed. When $\alpha = 0.63$, the visual effect is the best. Grass exposure is most natural.

Advantage of the iterator. The algorithm uses a fixed number (3, 5, 7, and 20) of iterations instead of an iterator. Figure 10 shows the visualization of the different iteration schemes.

Zero-DCE++ increases the image's brightness as the number of iterations increases. Zero-DiDCE does not increase the light level of the image infinitely with the number of iterations, and even using 20 iterations cannot cause overexposure of the enhanced image. As a result, the ALE-curve has a strong dynamic adjustment capability. The ALE-curve can gradually converge the input image illumination to α by successive iterations. However, this does not mean that the more iterations, the better the enhancement effect. Firstly, the optimal

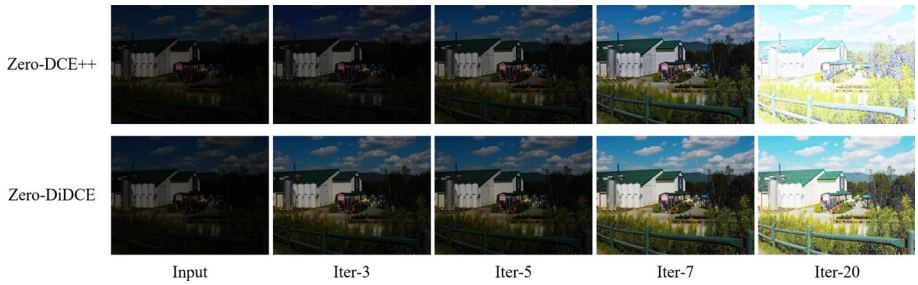

Fig. 10 Visual comparison of the two methods using the same number of iterations. $Iter - i$ denotes the result of running with a fixed number of i iterations

Table 1 Comparison of different iterative schemes

Method	PSNR↑	SSIM ↑	NIMA↑
Zero-DiDCE-3	10.15	0.45	3.87
Zero-DiDCE-5	13.36	0.63	4.03
Zero-DiDCE-7	17.15	0.71	4.34
Zero-DiDCE-20	15.56	0.61	**4.57**
Zero-DiDCE-iter	**18.23**	**0.72**	4.52

Zero-DiDCE-i means use i iterations, and iter means iterator. The best results are highlighted in bold

Fig. 11 Visual comparison of Zero-DCE++ [5] before and after using the iterator. Zero-DCE++-iter indicates the running result of Zero-DCE++ after using the iterator

iteration scheme varies for each input image, and using a fixed iteration scheme will affect the network's computational complexity and reduce the model's efficiency. Secondly, too many iterations do not make the enhanced image brightness infinitely close to α but will make it stay near α and cannot continue to improve. Therefore, Zero-DiDCE's iterator is designed with targeted iteration schemes for different images to reduce the computational complexity of the network while improving computational efficiency. The results of quantitative comparison of different iterative schemes are shown in Table 1.

As shown in Table 1, the scheme using the iterator performs the best. The iterator plays an important role.

Iterator can be applied to Zero-DCE++ [5] to improve its performance. Using an iterator instead of the fixed number of 8 iterations scheme in Zero-DCE++, keeping the remaining parameters unchanged, the run results are shown in Fig. 11. Since the two methods are entirely different regarding network structures, enhancement curves, and loss functions, their enhancement strengths are not the same. So the dynamic range of the iterator needs to be controlled between 2 and 10 iterations.

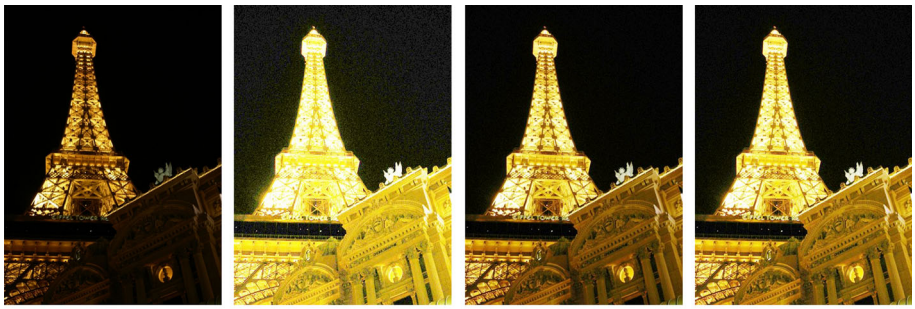

Fig. 12 Visual comparison of the model using fixed amplitude values with the amplitude controller. β=1/β=2 indicates the results of the run using fixed amplitude values

Table 2 The algorithm compares the results separately after using the amplitude controller and fixed amplitude values. The best results are highlighted in bold

Method	PSNR↑	SSIM ↑	NIMA↑
β=1	17.56	0.64	**4.61**
β=2	14.98	0.69	4.17
Zero-DiDCE	**18.23**	**0.72**	4.52

Table 3 Comparison of different networks. The best results are highlighted in bold

Method	PSNR↑	SSIM ↑	NIMA↑
CNN-3	18.00	**0.72**	4.50
CNN-4	**18.23**	**0.72**	**4.52**
CNN-5	**18.23**	**0.72**	**4.52**

Zero-DCE++-iter has the best enhancement, with sharper image details and more natural lighting. Zero-DCE++ images show underexposure and overexposure.

Advantage of amplitude controller. The experiment compares the effect of the amplitude controller and the fixed amplitude value on the algorithm, as shown in Fig. 12.

When $\beta = 1$, the enhanced image appears overexposed. When $\beta = 2$, local areas of the image appeared underexposed, such as dark walls. Zero-DiDCE had the best enhancement with an image. The walls did not appear overexposed or underexposed.

Additionally, the amplitude controller and the fixed amplitude values were compared quantitatively, as shown in Table 2.

Zero-DiDCE performs optimally in evaluation metrics. It is necessary to use the amplitude controller to design various degrees of enhancement or suppression for different images.

Effect of DiDCE-Net. The input image and the inverted input image perform side-by-side ordered ordinary CNN convolution of different layers (3, 4, and 5) denoted as CNN-3, CNN-4, and CNN-5. Skip connection in a network connects only the first and last layers of the network.

Observation of Table 3 shows that CNN-3 performs lower than CNN-4 and CNN-5. CNN-4 has the same enhancement effect as CNN-5, but the network complexity of CNN-5 is higher than that of CNN-4. The algorithm chooses CNN-4 as the network structure.

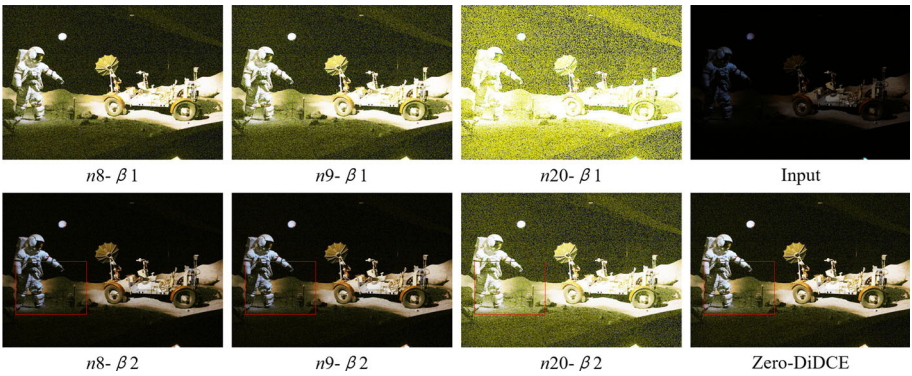

Fig. 13 Effect of different enhancement or suppression schemes on image noise. $n - \beta$ indicates the scheme of Zero-DiDCE using n iterations and β amplitude values

4.3 Benchmark Evaluations

The performance of Zero-DiDCE is tested through comparative experiments with other algorithms.

Analysis and effects of image noise. Low-light images are usually accompanied by some degree of noise. Most current low-light enhancement methods choose to deal with their noise problem after image enhancement, which poses a problem. This model requires an additional denoising module, which increases the complexity of the model and reduces computational efficiency. This is unacceptable, especially on devices with limited computing resources, such as mobile devices. Zhang et al. [40] proposed a self-supervised method to reduce noise while improving image contrast, avoiding image blur caused by pre- or post-enhancement denoising. Inspired by this, we have conducted many experiments and found that many noise problems of enhanced images are due to unreasonable enhancement methods that amplify the noise. The noise of the images themselves is not severe. Zero-DiDCE precisely controls the enhancement or suppression intensity of each input image through ALE-curve, iterator, and amplitude controller to avoid amplifying image noise, resulting in noise attenuation, as shown in Fig. 13.

As shown in Fig. 13, the amplitude controller and iterator play a massive role in avoiding amplified image noise. When the enhancement intensity of the input image is low, the image has an underexposed problem, such as $n8/9 - \beta 2$. When the enhancement intensity is too large, it inevitably amplifies the noise of the image, such as $n8/9/20 - \beta 1$ with $n20 - \beta 2$. Zero-DiDCE achieves good exposure and low noise, which shows that a well-designed enhancement intensity for each image can significantly enhance low-light images and avoid noise problems.

Deeper zero-shot learning. Compared to existing zero-shot learning methods [4, 5], Zero-DiDCE has lower limitations on training data and does not require a dataset containing multiple exposure images. The experimental result is shown in Fig. 14.

When using Lack-Exposure, the training dataset lacks exposure information. Zero-DCE [4] and Zero-DCE++ [5] are slightly overexposed, resulting in unrealistic colors. When using the dataset Lack-Low, which lacks low-lighting information, the images of Zero-DCE [4] and Zero-DCE++ [5] show significant underexposure. The mainstream zero-shot learning methods still have some limitations in the training dataset selection. They cannot learn the correct exposure information when the training dataset lacks exposure information, resulting

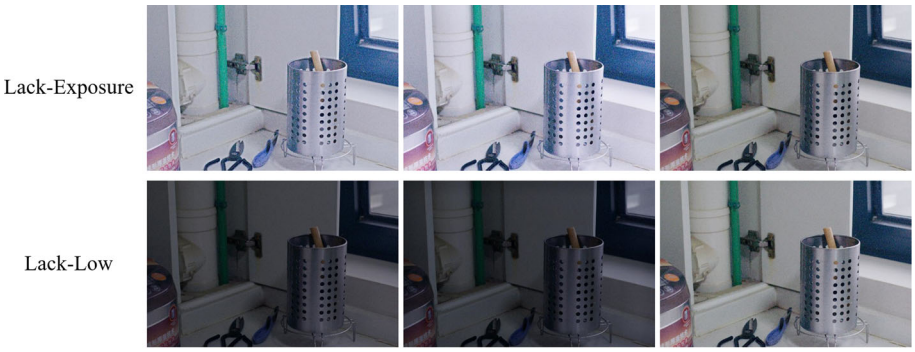

Fig. 14 Visual comparison on the Lack-Exposure and Lack-Low datasets

in the overexposure of their enhanced images. If low-lighting information is lacking in the dataset, the model cannot learn the accurate low-lighting information, resulting in underexposed images after restoration. Thanks to the specially designed piecewise non-reference loss function, DiDCE-Net, and ALE-curve, Zero-DiDCE has a more robust learning capability. The model learns critical information despite the lack of lighting information in the training dataset. Zero-DiDCE images performed the best. Compared with existing zero-shot learning methods, Zero-DiDCE has fewer restrictions on the dataset and more learning capability.

Visual comparisons. Compared with other methods, Zero-DiDCE can process images with different exposure levels better. The visual comparison of Zero-DiDCE with other methods is shown in Fig. 15.

Zero-DiDCE performs best when processing input images with different exposure levels. The model increases exposure levels in underexposed areas, maintains well-exposed areas, and suppresses overexposed areas.

Zero-DiDCE images perform best when the input images are low-light ones (Fig. 15a/b/c). MBLLEN [41] contrast is too high, and shadow areas cannot be corrected. RetinexNet's [37] images suffer from color bias. EnlightenGAN [2] and Zero-DCE++ [5] suffer from a slight underexposure (Fig. 15a). Zero-DCE++ [5] and Zero-DCE [4] are gradually overexposed as the exposure level increases.

When the input image is a normal light image (Fig. 15d), RetinexNet [37], EnlightenGAN [2], Zero-DCE++ [5], and Zero-DCE [4] images are overexposed. MBLLEN [41] contrast is too high, and the exposure of shadowed areas cannot be corrected.

The overexposure phenomenon of RetinexNet [37], EnlightenGAN [2], Zero-DCE++ [5], and Zero-DCE [4] images is more severe when the input image is overexposed (Fig. 15e). The Zero-DiDCE image is closest to the Reference image.

Quantitative comparisons. Experiments were performed on the LOL [37] for quantitative comparison with other methods. As shown in Table 4, Zero-DiDCE performs best even without using paired or unpaired multiple exposure data as the training dataset.

5 Conclusion and Future Work

We propose a zero-reference dual-illumination deep curve estimation method for low-light image enhancement. Zero-DiDCE is less restrictive on training data than current zero-shot

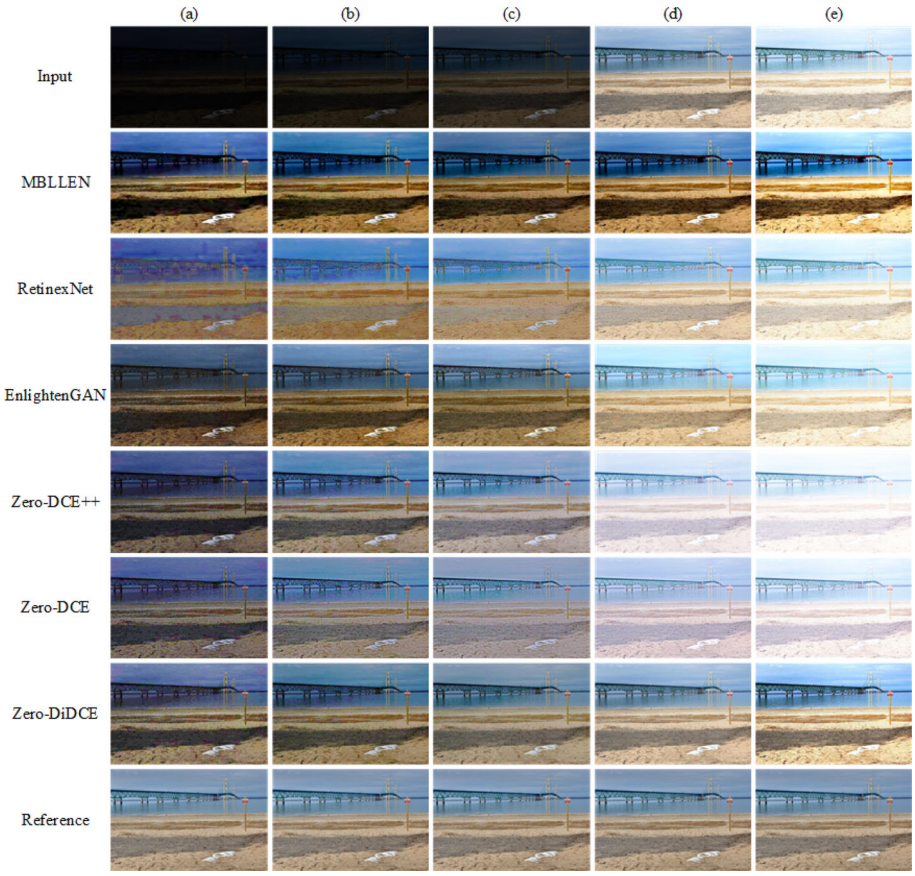

Fig. 15 Visual comparison with other methods on differently exposed images

Table 4 Quantitative comparison of results. The best results are highlighted in bold

Method	PSNR↑	SSIM ↑	NIMA↑
MBLLEN [41]	17.90	0.71	4.37
RetinexNet [37]	16.77	0.42	3.37
EnlightenGAN [2]	17.48	0.71	4.36
Zero-DCE++ [5]	14.11	0.50	**4.60**
Zero-DCE [4]	14.86	0.58	4.09
Zero-DiDCE	**18.23**	**0.72**	4.52

learning methods. It also handles images with different exposure levels well and outperforms various advanced algorithms in subjective and objective metrics. These are achieved through the powerful adjustment capabilities of the adaptive light enhancement curve. The curve ensures accurate and targeted adjustment through DiDCE-Net, piecewise non-reference loss function, amplitude controller, and iterator. The algorithm still has room for improvement in suppressing image noise. In the future, we will further explore the image noise problem to improve the image enhancement effect.

Acknowledgements This work was supported by the National Natural Science Foundation of China (62273292).

Declarations

Conflict of interest All authors declare no conflict of interest.

References

1. Zheng S, Ma Y, Pan J, Lu C, Gupta G (2022) Low-light image and video enhancement: a comprehensive survey and beyond. arXiv preprint arXiv:2212.10772
2. 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
3. Zhu J-Y, Park T, Isola P, Efros AA (2017) Unpaired image-to-image translation using cycle-consistent adversarial networks. In: Proceedings of the IEEE international conference on computer vision, pp 2223–2232
4. Guo C, Li C, Guo J, Loy CC, Hou J, Kwong S, Cong R (2020) Zero-reference deep curve estimation for low-light image enhancement. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp 1780–1789
5. Li C, Guo C, Loy CC (2021) Learning to enhance low-light image via zero-reference deep curve estimation. IEEE Trans Pattern Anal Mach Intell 44(8):4225–4238
6. Wang C, Ye Z (2005) Brightness preserving histogram equalization with maximum entropy: a variational perspective. IEEE Trans Consum Electron 51(4):1326–1334
7. Reza AM (2004) Realization of the contrast limited adaptive histogram equalization (clahe) for real-time image enhancement. J VLSI Signal Process Syst Signal Image Video Technol 38:35–44
8. Land EH (1986) An alternative technique for the computation of the designator in the retinex theory of color vision. Proc Natl Acad Sci 83(10):3078–3080
9. 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
10. Jobson DJ, Rahman Z-U, Woodell GA (1997) Properties and performance of a center/surround retinex. IEEE Trans Image Process 6(3):451–462
11. Li M, Liu J, Yang W, Sun X, Guo Z (2018) Structure-revealing low-light image enhancement via robust retinex model. IEEE Trans Image Process 27(6):2828–2841
12. Gu Z, Li F, Fang F, Zhang G (2019) A novel retinex-based fractional-order variational model for images with severely low light. IEEE Trans Image Process 29:3239–3253
13. Ren X, Yang W, Cheng W-H, Liu J (2020) Lr3m: robust low-light enhancement via low-rank regularized retinex model. IEEE Trans Image Process 29:5862–5876
14. Guo X, Li Y, Ling H (2016) Lime: low-light image enhancement via illumination map estimation. IEEE Trans Image Process 26(2):982–993
15. Fu X, Zeng D, Huang Y, Zhang X-P, Ding X (2016) A weighted variational model for simultaneous reflectance and illumination estimation. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 2782–2790
16. Lore KG, Akintayo A, Sarkar S (2017) Llnet: a deep autoencoder approach to natural low-light image enhancement. Pattern Recogn 61:650–662
17. 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

18. Li C, Guo J, Porikli F, Pang Y (2018) Lightennet: a convolutional neural network for weakly illuminated image enhancement. Pattern Recogn Lett 104:15–22
19. Zhang Y, Zhang J, Guo X (2019) Kindling the darkness: a practical low-light image enhancer. In: Proceedings of the 27th ACM international conference on multimedia, pp 1632–1640
20. 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
21. Lim S, Kim W (2020) Dslr: deep stacked laplacian restorer for low-light image enhancement. IEEE Trans Multimed 23:4272–4284
22. Li J, Feng X, Hua Z (2021) Low-light image enhancement via progressive-recursive network. IEEE Trans Circuits Syst Video Technol 31(11):4227–4240
23. Zhang L, Zhang L, Liu X, Shen Y, Zhang S, Zhao S (2019) Zero-shot restoration of back-lit images using deep internal learning. In: Proceedings of the 27th ACM international conference on multimedia, pp 1623–1631
24. Zhu A, Zhang L, Shen Y, Ma Y, Zhao S, Zhou Y (2020) Zero-shot restoration of underexposed images via robust retinex decomposition. In: 2020 IEEE international conference on multimedia and expo (ICME), pp 1–6. IEEE
25. Zhao Z, Xiong B, Wang L, Ou Q, Yu L, Kuang F (2021) Retinexdip: a unified deep framework for low-light image enhancement. IEEE Trans Circuits Syst Video Technol 32(3):1076–1088
26. 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: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp 10561–10570
27. Yang W, Wang S, Fang Y, Wang Y, Liu J (2020) From fidelity to perceptual quality: a semi-supervised approach for low-light image enhancement. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp 3063–3072
28. Yang W, Wang S, Fang Y, Wang Y, Liu J (2021) Band representation-based semi-supervised low-light image enhancement: bridging the gap between signal fidelity and perceptual quality. IEEE Trans Image Process 30:3461–3473
29. Zhu M, Pan P, Chen W, Yang Y (2020) Eemefn: low-light image enhancement via edge-enhanced multi-exposure fusion network. In: Proceedings of the AAAI conference on artificial intelligence, vol 34, pp 13106–13113
30. Ma L, Ma T, Liu R, Fan X, Luo Z (2022) Toward fast, flexible, and robust low-light image enhancement. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp 5637–5646
31. He K, Sun J, Tang X (2012) Guided image filtering. IEEE Trans Pattern Anal Mach Intell 35(6):1397–1409
32. Whittle P (1994) The psychophysics of contrast brightness. In: Lightness, brightness and transparency. lawrence erlbaum associates, New Jersey, pp 35–110
33. Zhang X, Wang G, Yang L, Chen C (2023) Lightweight portrait segmentation via edge-optimized attention. In: ICASSP 2023-2023 IEEE international conference on acoustics, speech and signal processing (ICASSP), pp 1–5. IEEE
34. Peng C, Liu Y, Kang K, Chen Y, Wu X, Cheng A, Kang Z, Chen C, Cheng Q (2022) Hyperspectral image denoising using nonconvex local low-rank and sparse separation with spatial-spectral total variation regularization. IEEE Trans Geosci Remote Sens 60:1–17
35. Ronneberger O, Fischer P, Brox T (2015) U-net: convolutional networks for biomedical image segmentation. In: Medical image computing and computer-assisted intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18, Springer, pp 234–241
36. Chen C, Zhao H, Yang H, Yu T, Peng C, Qin H (2021) Full-reference screen content image quality assessment by fusing multilevel structure similarity. ACM Trans Multimed Comput Commun Appl (TOMM) 17(3):1–21
37. Wei C, Wang W, Yang W, Liu J (2018) Deep retinex decomposition for low-light enhancement. arXiv preprint arXiv:1808.04560
38. Cai J, Gu S, Zhang L (2018) Learning a deep single image contrast enhancer from multi-exposure images. IEEE Trans Image Process 27(4):2049–2062
39. Wang Z, Bovik AC, Sheikh HR, Simoncelli EP (2004) Image quality assessment: from error visibility to structural similarity. IEEE Trans Image Process 13(4):600–612
40. Zhang Y, Di X, Zhang B, Li Q, Yan S, Wang C (2021) Self-supervised low light image enhancement and denoising. arXiv preprint arXiv:2103.00832
41. Lv F, Lu F, Wu J, Lim C (2018) Mbllen: low-light image/video enhancement using CNNs. In: BMVC, vol 220, p 4

Terms and Conditions

Springer Nature journal content, brought to you courtesy of Springer Nature Customer Service Center GmbH ("Springer Nature").

Springer Nature supports a reasonable amount of sharing of research papers by authors, subscribers and authorised users ("Users"), for small-scale personal, non-commercial use provided that all copyright, trade and service marks and other proprietary notices are maintained. By accessing, sharing, receiving or otherwise using the Springer Nature journal content you agree to these terms of use ("Terms"). For these purposes, Springer Nature considers academic use (by researchers and students) to be non-commercial.

These Terms are supplementary and will apply in addition to any applicable website terms and conditions, a relevant site licence or a personal subscription. These Terms will prevail over any conflict or ambiguity with regards to the relevant terms, a site licence or a personal subscription (to the extent of the conflict or ambiguity only). For Creative Commons-licensed articles, the terms of the Creative Commons license used will apply.

We collect and use personal data to provide access to the Springer Nature journal content. We may also use these personal data internally within ResearchGate and Springer Nature and as agreed share it, in an anonymised way, for purposes of tracking, analysis and reporting. We will not otherwise disclose your personal data outside the ResearchGate or the Springer Nature group of companies unless we have your permission as detailed in the Privacy Policy.

While Users may use the Springer Nature journal content for small scale, personal non-commercial use, it is important to note that Users may not:

1. use such content for the purpose of providing other users with access on a regular or large scale basis or as a means to circumvent access control;
2. use such content where to do so would be considered a criminal or statutory offence in any jurisdiction, or gives rise to civil liability, or is otherwise unlawful;
3. falsely or misleadingly imply or suggest endorsement, approval , sponsorship, or association unless explicitly agreed to by Springer Nature in writing;
4. use bots or other automated methods to access the content or redirect messages
5. override any security feature or exclusionary protocol; or
6. share the content in order to create substitute for Springer Nature products or services or a systematic database of Springer Nature journal content.

In line with the restriction against commercial use, Springer Nature does not permit the creation of a product or service that creates revenue, royalties, rent or income from our content or its inclusion as part of a paid for service or for other commercial gain. Springer Nature journal content cannot be used for inter-library loans and librarians may not upload Springer Nature journal content on a large scale into their, or any other, institutional repository.

These terms of use are reviewed regularly and may be amended at any time. Springer Nature is not obligated to publish any information or content on this website and may remove it or features or functionality at our sole discretion, at any time with or without notice. Springer Nature may revoke this licence to you at any time and remove access to any copies of the Springer Nature journal content which have been saved.

To the fullest extent permitted by law, Springer Nature makes no warranties, representations or guarantees to Users, either express or implied with respect to the Springer nature journal content and all parties disclaim and waive any implied warranties or warranties imposed by law, including merchantability or fitness for any particular purpose.

Please note that these rights do not automatically extend to content, data or other material published by Springer Nature that may be licensed from third parties.

If you would like to use or distribute our Springer Nature journal content to a wider audience or on a regular basis or in any other manner not expressly permitted by these Terms, please contact Springer Nature at

onlineservice@springernature.com