


## Article

# DCTE-LLIE: A Dual Color-and-Texture-Enhancement-Based Method for Low-Light Image Enhancement

Hua Wang <sup>1,2</sup> , Jianzhong Cao <sup>1</sup>, Lei Yang <sup>2</sup> and Jijiang Huang <sup>1,\*</sup>

<sup>1</sup> Xi'an Institute of Optics and Precision Mechanics, Chinese Academy of Sciences, Xi'an 710119, China; wanghua@opt.ac.cn (H.W.); cjz@opt.ac.cn (J.C.)

<sup>2</sup> University of Chinese Academy of Sciences, Beijing 100049, China; yanglei@opt.ac.cn

\* Correspondence: huangjijiang@opt.ac.cn

**Abstract:** The enhancement of images captured under low-light conditions plays a vitally important role in the area of image processing and can significantly affect the performance of following operations. In recent years, deep learning techniques have been leveraged in the area of low-light image enhancement tasks, and deep-learning-based low-light image enhancement methods have been the mainstream for low-light image enhancement tasks. However, due to the inability of existing methods to effectively maintain the color distribution of the original input image and to effectively handle feature descriptions at different scales, the final enhanced image exhibits color distortion and local blurring phenomena. So, in this paper, a novel dual color-and-texture-enhancement-based low-light image enhancement method is proposed, which can effectively enhance low-light images. Firstly, a novel color enhancement block is leveraged to help maintain color distribution during the enhancement process, which can further eliminate the color distortion effect; after that, an attention-based multiscale texture enhancement block is proposed to help the network focus on multiscale local regions and extract more reliable texture representations automatically, and a fusion strategy is leveraged to fuse the multiscale feature representations automatically and finally generate the enhanced reflection component. The experimental results on public datasets and real-world low-light images established the effectiveness of the proposed method on low-light image enhancement tasks.

**Keywords:** low-light image enhancement; deep learning; color enhancement block; texture enhancement block; multiscale feature fusion



**Citation:** Wang, H.; Cao, J.; Yang, L.; Huang, J. DCTE-LLIE: A Dual Color-and-Texture-Enhancement-Based Method for Low-Light Image Enhancement. *Computers* **2024**, *13*, 134. <https://doi.org/10.3390/computers13060134>

Academic Editor: Lucia Maddalena

Received: 11 April 2024

Revised: 20 May 2024

Accepted: 24 May 2024

Published: 27 May 2024



**Copyright:** © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Images contain sufficient real-world information, and image processing techniques allow automatic information extraction from images, further enabling different applications of the aforementioned information. However, under certain extreme conditions, it may not be possible to acquire images under sufficient light conditions. There may be a lot of causes for insufficient light conditions, such as low-light environments, limited performance of photography equipment, and improper usage of the equipment. Low-light conditions may lead to multiple degradations, such as poor visibility, low contrast, and accidental noise, which means that compared with images captured under sufficient imaging conditions, it is a challenge to extract information from low-light images. So, based on the discussions above, low-light image enhancement (LLIE) has become an important branch in the image processing area.

In recent years, researchers have paid great attention to the LLIE task, and multiple methods have been proposed to improve the subjective and objective quality of low-light images [1]. Previous methods [2–5] mainly focus on the spatial low-light image enhancement tasks, such methods directly operating on the pixels of input images, making the enhancement process simple and efficient, and the speed can also be very fast. However, although such methods can improve the brightness of input images to a certain extent

and enhance the details, the overall characteristics of the image are ignored, so the final enhanced image has a poor sense of hierarchy. The histogram equalization-based methods are simple and have a good enhancement performance on most low-light images, but they cannot conform to the law of object imaging models, and the enhanced image may contain problems, such as dark brightness, detail loss, and color distortion, which cannot achieve satisfactory visual effects.

Another category is based on the Retinex theory. The Retinex theory was first proposed by Land in 1977 [6] and was a human visual characteristics-based image enhancement method. The Retinex theory assumes that the color of an object depends on the object's ability to reflect light, rather than the absolute value of the texture of reflected light, and the color is not affected by lighting inhomogeneity. So, the Retinex theory needs to decompose the observed image into a reflection component and an illumination component, and only the reflection component is used to achieve enhancement on observed images. The Retinex-theory-based methods can realize brightness unity, detail enhancement, and color fidelity, which has been favored by scholars around the world and has been widely used in the field of image enhancement.

In recent years, deep learning [7] has been established to be a very powerful tool for feature extraction, and it has been widely used in multiple research areas, such as image classification [8–11], object detection [12–16], image segmentation [17–20], etc. By optimizing the feature extractor and classifier in an end-to-end fashion, deep learning models can automatically extract high-level feature representations and achieve the global optimum during the training process. Deep learning has also been applied to the area of image enhancement, especially for low-light image enhancement tasks. Researchers also proposed RetinexNet [21], which combined Retinex theory with deep learning technology, and further proposed multiple low-light image enhancement methods. The general process of Retinex-theory-based methods is to extract the illumination component and reflection component through a decomposition model first, and then enhancement is applied to the reflection component, and the enhanced reflection component and illumination are composed again to generate the final enhanced image.

However, although deep-learning-based image enhancement methods [22–28] have achieved very outstanding performance, the low-light image enhancement task is still considered a huge challenge, and most of the existing methods still suffer from the problem of color distortion and local blurring. This is because most of the existing enhancement methods have difficulty retaining color distribution, while also failing to extract multiscale region texture representations during the enhancement process. So, in response to the above questions, a novel low-light image enhancement method, which is called the dual color-and-texture-enhancement-based LLIE (DCTE-LLIE) method is proposed in this paper. The contribution of the proposed DCTE-LLIE method can be summarized as follows:

1. A novel method called the DCTE-LLIE method is proposed in this paper. The proposed DCTE-LLIE method can extract more realistic color and texture feature representations of low-light images during the enhancement process, which can help eliminate the color distortion and local blurring effect on the final enhanced image more effectively.
2. A novel color enhancement block (CEB) is proposed to extract more realistic color representations by maintaining the color distribution of the low-light images during the enhancement process, help extract more reasonable color representations, and finally, eliminate the color distortion of the final enhanced image.
3. A multiscale attention-based texture enhancement block (ATEB) is proposed to help the network focus on the local regions and extract more effective and reliable texture feature representations during the training process; also, a multiscale feature fusion strategy is proposed to fuse multiscale features automatically and finally generate more reliable texture representations.

The rest of the paper is organized as follows: The related work of low-light image enhancement is introduced in Section 2, and the detailed description of the proposed

DCTE-LLIE method is described in Section 3. The experimental results on public datasets and real-world images are shown in Section 4. The paper is concluded in Section 5.

## 2. Related Work

The low-light image enhancement methods can be divided into three categories: (a) generic enhancement methods; (b) Retinex-theory-based enhancement methods; and (c) deep-learning-based enhancement methods.

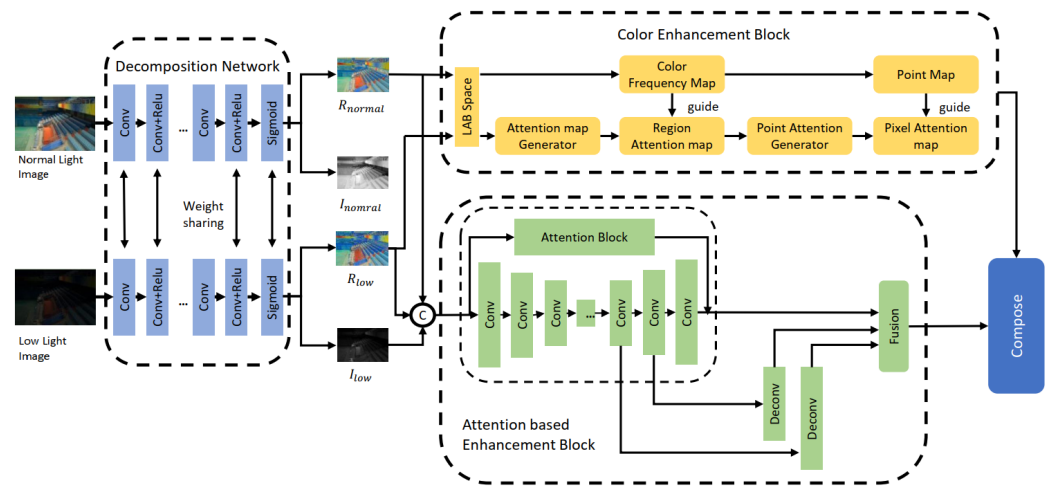
The generic enhancement methods: Histogram equalization (HE) [29] and its variations are first applied to the enhancement process to make sure the histogram of the output image can meet the constraints. The HE-based methods can be categorized into global histogram-equalization-based methods [29] and local histogram-equalization-based methods [30–32]. The global HE-based method aims to adjust the overall gray level of the low-light image, but the performance will be worse if the input image is darker, as the object information of the image cannot be highlighted and the details cannot be well preserved. The local HE-based method is to split the input image into multiple subblocks, equalize the histograms for each subblock separately, and finally superimpose the subblocks to achieve the enhanced image. As each subblock is a neighborhood of a certain pixel, the high-frequency gray level can be improved, and the low-frequency gray level is suppressed. The dehazing-based method [33] takes advantage of the inverse connection between low-light images and images under hazy environments. However, the above methods may lead to amplified intensive noise during the enhancement process.

Another category is the Retinex-theory-based method. Retinex theory can decompose the input image into an illumination component and reflection component, and the enhancement process can be applied to the illumination component and reflection component, respectively. Finally, the enhanced illumination component and reflection component are composed again to obtain the final enhanced image. Multiple methods have been proposed for the decomposition process [34–36]. The single-scale Retinex [34] method first leverages the Gaussian filter to smooth the illumination component and further enhance the low-light image. The multiscale Retinex method [35] extends the above methods by multiscale Gaussian filters. The LIME [37] method takes advantage of structure priors to estimate the illumination components and finally uses the reflection component as the enhanced result. Also, multiple methods [38,39] are proposed to enhance low-light images by simultaneously achieving image enhancement and noise removal.

Although existing methods have shown outstanding performance on the low-light image enhancement task and are established to obtain pleasant results, the performance of the above methods may be limited due to the models' capacity. However, in recent years, with the rapid development of deep neural networks, convolution neural networks (CNNs) have been widely used in the compute vision area. Also, as CNNs have power in the low-level computer vision area [22–28], CNN-based LLIE methods have drawn great attention. Wang et al. [22] proposed novel lightening back-projection blocks to learn the residual for normal-light estimations. Wang et al. [24] proposed a novel stream regularization model LLFlow to achieve outstanding performance on multiple tasks. Guo et al. [27] transfer an input image into luminance–chrominance color space and eliminate noise in the brightened luminance. Many existing methods also take advantage of the Retinex theory to help improve the performance of LLIE methods. Zhao et al. [23] proposed a novel generative strategy for Retinex decomposition, further estimated the latent components, and performed low-light image enhancement. Hai et al. [25] took advantage of frequency information to preserve image details and achieved more robust results. Yang et al. [26] proposed an enhanced method that takes advantage of ViT to acquire high-level global fine-tuned features and achieved further improvement on enhancement. Cai et al. [28] proposed an illumination-guided transformer that utilizes illumination representations to direct the modeling of nonlocal interactions of regions with different light conditions.

### 3. DCTE-LLIE Method

Although researchers have proposed multiple low-light image enhancement methods, and the enhancement performance is established, existing methods still suffer from color distortion and local blurring effects. So, considering the above problems, a novel method called the DCTE-LLIE method is proposed in this paper to deal with low-light image enhancement tasks. The proposed method contains three main blocks: (a) the decomposition subnetwork; (b) the color enhancement block (CEB), and (c) the attention-based texture enhancement block (ATEB). The whole structure of the proposed DCTE-LLIE method is shown in Figure 1. As the DCTE-LLIE method is designed based on the Retinex theory, the input image is first decomposed into the reflection component and illumination component, and the enhancement method is then applied to the reflection and illumination components, respectively. The illumination component is enhanced by a denoising operation that is the same as RetinexNet, while the proposed CEB and ATEB are applied to the reflection component. Finally, the enhanced illumination component and reflection component are composed to obtain the final enhanced image.



**Figure 1.** The structure of the proposed reflection enhancement branch of the DCTE-LLIE method.

#### 3.1. Decomposition Subnetwork

As the DCTE-LLIE method is designed based on the Retinex theory, the input image should be first decomposed into a reflection component and an illumination component. In this paper, a decomposition subnetwork is used to achieve image decomposition. Based on the Retinex theory, the image  $I$  can be decomposed as

$$I(x, y) = R(x, y) \times L(x, y) \quad (1)$$

where  $L(x, y)$  is the reflection component,  $R(x, y)$  is the illumination component,  $\times$  represents the multiplication operation. According to the Retinex theory, the reflection component is a constant part that is determined by the nature of the object itself, while the illumination component is affected by the external light. So, the enhancement process can be achieved on the reflection component by removing the influence of lighting while correcting the illumination component.

The proposed decomposition network is constructed by 5 convolution layers followed by ReLU, and a two-channel output that corresponds to the reflection and illumination component is generated. The output goes through a sigmoid function to generate the final output which has the same size as the input. For each input image pair (including the low-light image and corresponding normal-light image), the low-light image is decomposed into reflection component  $R_{low}$  and illumination component  $L_{low}$ , while the normal-light image is decomposed into reflection component  $R_{normal}$  and illumination component  $L_{normal}$ . A constrained loss function is leveraged to optimize the parameter of

the proposed decomposition subnetwork, and the final loss function is composed of reconstruction loss  $L_{recon}$ , reflection component consistency loss  $L_{ir}$ , and illumination smoothing loss  $L_{is}$ . The decomposition loss  $L_D$  can be formed as

$$L_D = L_{recon} + \lambda_{ir}L_{ir} + \lambda_{is}L_{is} \quad (2)$$

where the  $L_{recon}$  loss is defined as

$$L_{recon} = \sum_{i=low,normal} \sum_{j=low,normal} \lambda_{ij} \|R_i * I_j - S_j\|_1 \quad (3)$$

where  $*$  represents the inner product operation.  $L_{recon}$  is designed to enable the reflection and illumination components to reconstruct the origin input as much as possible. The reflection component consistency loss  $L_{ir}$  is defined as

$$L_{ir} = \|R_{low} - R_{normal}\|_1 \quad (4)$$

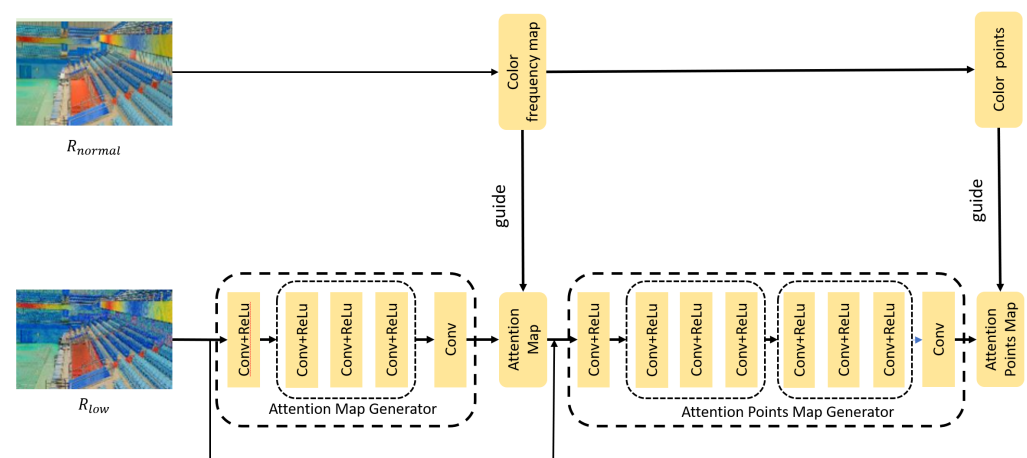
where  $\| \cdot \|_1$  means the L1 loss.  $L_{ir}$  aims to make sure the reflection components of low-light image  $R_{low}$  and normal-light image  $R_{normal}$  should be as similar as possible. The smoothing loss  $L_{is}$  aims to represent that an ideal illumination component should be as smooth as possible. The illumination smooth loss  $L_{is}$  is then defined as

$$L_{is} = \sum_{i=low,normal} \|\nabla I_i \circ \exp(-\lambda_g \nabla R_i)\| \quad (5)$$

where  $\nabla I_i$  means calculating the gradient of  $I_i$ , and  $\nabla R_i$  represents the gradient of  $R_i$ . The proposed smooth loss assigns weights to the gradient map of the lighting components by finding a gradient for the reflection components so that the areas that are smoother on the reflection components should also be as smooth as possible on the lighting components.

### 3.2. Color Enhancement Block

Retinex-theory-based LLIE methods can decompose an image into a reflection component and illumination component, and then apply enhancement on the reflection component. However, the final enhanced image may still suffer from the color distortion effect because existing Retinex-theory-based methods cannot maintain the color distributions during the enhancement process, which means they cannot extract realistic color representations and finally lead to color distortion on the final enhanced image. So, in order to extract more realistic color representations, a novel color enhancement block (CEB) is proposed in this paper. The structure of the proposed CEB is shown in Figure 2.



**Figure 2.** The structure of proposed color enhancement block.

In order to extract more realistic color representations, a novel CEB is proposed in this paper to help the network extract more realistic color representations by maintaining color distribution during the enhancement process. As shown in Figure 2, the reflection component  $R_{low}$  is first transformed into the LAB color space. LAB space is designed based on the human perception of color, so compared with other color spaces such as RGB and CMYK, LAB is more in line with human vision and easier to adjust. L channel can be used to adjust image brightness, and the AB channel can be used to adjust image color representation.

So, in this paper, the AB channel is considered as the input of the color enhancement block (CEB). The AB channel first goes through the attention map generator and generates a two-channel attention map  $\hat{M}$  corresponding to the AB channel. The generated attention map  $\hat{M}$  is concated with the AB channel, and the concated feature further goes through the attention point generator to generate an attention point map  $\hat{P}$ . The generated attention map  $\hat{M}$  and attention point map  $\hat{P}$  are used to help the proposed network maintain the color distribution of the input image and extract more realistic color representations. The generated attention map  $\hat{M}$  and attention point map  $\hat{P}$  were supervised by the color distribution map  $M$  and  $P$ , which were generated from the reflection component  $R_{normal}$  of the normal-light image, as the images captured under normal-light conditions contain more realistic and precious color information. For a given image, the color distribution map  $M$  was calculated by counting the color appearance for each pixel. The color distribution map  $M$  is calculated as

$$M = R_{normal}[AB] \odot \bar{F} \quad (6)$$

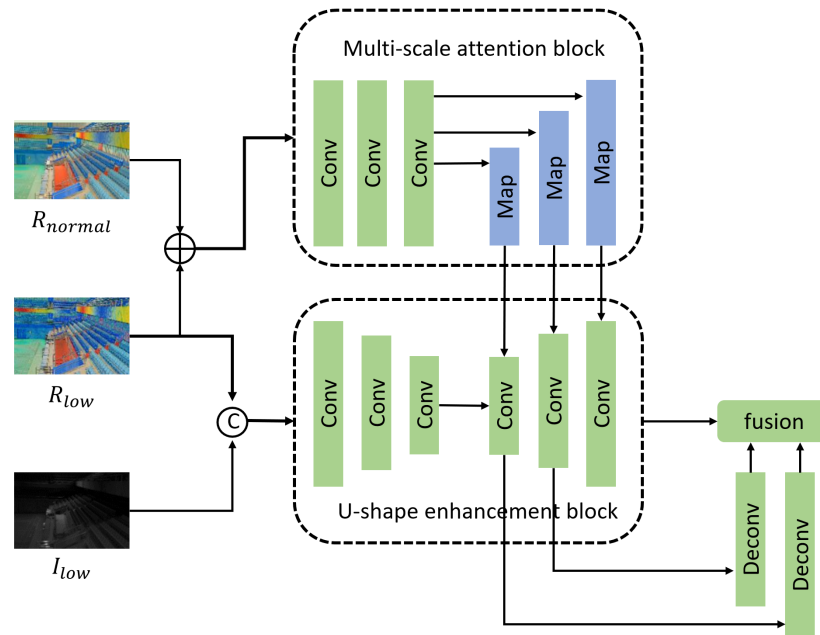
where  $\odot$  means Hadamard product operation,  $R_{normal}[AB]$  represents the AB channel of  $R_{normal}$ , so  $i \in 1, 2$  as the  $R_{normal}[AB]$  has only two channels, and  $\bar{F}$  is the color frequency map. Each component  $\bar{F}(x, y)$  equals the number of occurrences of the color of the input image. So,  $\bar{F}$  can be used to represent the color distribution of the input image. After that, in order to eliminate the effect of noise and background, a threshold is used to eliminate the dominant color distribution and useless noise frequency to finally focus on the foreground colors, as the color background of an input image should be very similar, which is represented as very high values in  $\bar{F}$ . After  $M$  is calculated, multiple foreground points are randomly selected from  $M$  to generate attention point map  $P$ . In this part,  $M$  is used to supervise the attention map  $\hat{M}$ , and  $P$  is used to supervise the attention point map  $\hat{P}$ . The first part can help the network extract local color representations, while the second part guides the network in extracting pixel-level color representations. It is easy to understand that map  $M$  has much more complicated and duplicated colors than  $P$ , so the supervision of  $P$  can cover most of the color distribution while using the least constraints.

It is worth noticing that the generated color attention map  $\hat{M}$  and attention point map  $\hat{P}$  are supervised by the color distribution map generated by the illumination component  $R_{normal}$  from the normal-light image. This is because the reflection component of the normal-light image contains more realistic color information, so the proposed CEB can help the network extract more realistic color representations.

### 3.3. Attention-Based Texture Enhancement Block

To further eliminate the local blurring effect in the enhanced image, a novel attention-based texture enhancement block (ATEB) is also proposed to help the network focus on the local regions and further extract more reliable local region representations. As existing methods cannot extract multiscale region representations and further lead to a local blur effect, it is essential to extract more effective multiscale local representations. Attention mechanisms [40] are established to show excellent ability to help networks focus on the region of interest and extract more reliable feature representations. So, in this paper, an attention mechanism is also applied to the texture enhancement block, and an attention-based multiscale texture enhancement block (ATEB) is proposed.

The structure of the proposed ATEB is shown in Figure 3. The input image is first decomposed into the illumination component and reflection component by the decomposition subnetwork, and the enhancement process is then applied to the reflection component. The proposed ATEB contains three parts: (a) the multiscale attention block, which used to help network focus on the multiscale region of interest of input image by extracting multiscale attention maps; (b) a U-shaped enhancement block, where a U-shape network structure is constructed to extract multiscale feature presentations; and (c) a multiscale feature fusion strategy, where the multiscale attention map and multiscale feature representations are fused to generate feature representations with sufficient local region texture information.



**Figure 3.** The structure of the proposed attention-based texture enhancement block.

The reflection component  $R_{low}$  and  $R_{normal}$  contains sufficient but various texture information, so  $R_{low}$  and  $R_{normal}$  are first mixed to form a mixed reflection representation  $R_{mix}$ . The mixed feature representation  $R_{mix} = (R_{low} + R_{normal})/2$ . After that,  $R_{mix}$  is sent into the multiscale attention block to generate multiscale spatial attention maps. The proposed multiple-scale attention block is constructed by multiple  $3 \times 3$  convolution layers and one  $1 \times 1$  convolution layer. Firstly, two  $3 \times 3$  convolution layers are applied on the mixed reflection feature to extract multiscale regions of interest and remove background information. After that, a  $1 \times 1$  convolution layer is used to reduce the channel number and computation cost. Finally, two  $3 \times 3$  convolution layers are leveraged to extract the final multiscale attention maps. The generated multiscale attention maps further correspond to the multiscale feature representations generated from the U-shape network.

As shown in Figure 3, the reflection component  $R_{low}$  and illumination component  $I_{low}$  from the low-light image are concatenated to form a concatenated feature representation as the input of the U-shape feature enhancement block. In this paper, a U-shape enhancement network is leveraged to extract multiscale local region feature representations and achieve enhancement on the reflection component. The left part of the U-shape network can capture feature representations from low level to high level, while the right part can recover the details of each layer and continuously pass high-level semantic information to the bottom. So, the proposed U-shape enhancement network can effectively extract multiscale feature representations.

After the multiscale feature is extracted from the U-shape enhancement network, the multiscale attention maps are applied to the multiscale feature representations to generate the final feature representation which contains sufficient local region texture information.

In order to fuse multiscale feature representations effectively, a novel fusion strategy is also proposed, where multiple deconvolution layers are used to transform feature maps with different sizes into the same size and finally fuse multiscale feature maps into an enhanced reflection component. The final reflection component is then combined with the enhanced illumination component to generate the final enhanced image.

### 3.4. Training Strategy

The proposed method cannot optimize the model in an end-to-end fashion, especially the decomposition subnetwork and color enhancement block. This is because the decomposition subnetwork will seriously affect the performance of the following operation. For the color enhancement block, two color distribution maps  $M$  and  $P$  are leveraged to supervise the realistic color representation extraction process. So, in this paper, a three-step training strategy is proposed to achieve parameter optimization for the proposed DCTE-LLIE method. The training process for the proposed method can be divided into three steps: (a) decomposition subnetwork optimization to optimize the Retinex decomposition subnetwork by minimizing the loss function  $L_D$ ; (b) color enhancement block optimization, where the color distribution map and color point map is generated to guide network maintaining color distribution and help extract more realistic color representations, and this step is achieved by minimizing the color maintaining loss function  $L_C$ ; and (c) the whole network optimization, where the whole network is optimized in an end-to-end fashion, while the parameter of the decomposition subnetwork and the CEB are also fine-tuned during this step.

The second step was to optimize the color attention generator, where the color attention map  $\hat{M}$  was supervised by the color distribution map  $M$  through the following function:

$$L_M = \|\hat{M} - M\|_1 \quad (7)$$

The color point attention map  $\hat{P}$  was supervised by the color points map  $P$  by the following function:

$$L_P = \|\hat{P} - P\|_1 \quad (8)$$

the final color maintaining loss function  $L_C$  for color enhancement block optimization is defined as

$$L_C = L_M + \alpha L_P \quad (9)$$

where  $\alpha$  is the balance parameter to balance the color attention map loss  $L_M$  and color point attention map loss  $L_P$ .

After that, the whole DCTE-LLIE method can be optimized in an end-to-end fashion, while the model trained in the first two steps can also be fine-tuned during the training process. The final loss function is formed as

$$L_F = \sum_{i=low} \sum_{j=low} \|\hat{R}_i \odot \hat{I}_j - S_j\|^2 \quad (10)$$

where  $\hat{R}_i$  is the enhanced reflection component, while  $\hat{I}_j$  is the enhanced illumination component,  $\odot$  represent for the Hardward. The final loss function  $L_F$  aims to minimize the difference between the enhanced image which is composed of the enhanced reflection component and illumination component and the ground truth image.

## 4. Experiment

### 4.1. Experiment Details

To establish the effectiveness of the proposed method, the proposed method is applied to three public datasets: the LOL, SID, and MIT-Adobe FiveK datasets, and the proposed DCTE-LLIE method is also applied to the low-light images captured in real-world scenarios.



The LOL dataset is the first dataset that captures low-light images and corresponding normal-light images in a real-world environment. Also, the LOL dataset consists of two categories, 500 real-world image pairs and 1000 composite image pairs. The real-world images are captured under real-world environments, while the composite images are generated from the normal-light image from the RAISE dataset. In this paper, only real-world images are used in this experiment, in which 485 image pairs are used for training and 15 image pairs are used for testing.

The SID dataset includes 5094 raw images captured under low-light conditions, and each low-light image has a corresponding normal-light high-quality image. Also, the SID dataset contains indoor and outdoor images, and the outdoor images are usually captured under moonlight or street lighting conditions. Images in the SID dataset are captured by two different cameras: Sony  $\alpha 7S$  and Fujifilm X-T2, so the SID contains two subsets:  $ID_{Sony}$  and  $SID_{Fuji}$ .

The MIT-Adobe FiveK dataset was proposed in 2011, including 5000 raw images captured by DSLR camera. Each image was postadjust by 5 experienced photographers using Adobe Lightroom for color tone. As this dataset contains pairwise data of the original image and 5 postimages, and there are multiple photos from the same photographer, so it can be used for a later style of learning.

#### 4.2. Evaluation Index

In order to evaluate the effectiveness of the proposed method, multiple evaluation indexes are used. In this paper, PSNR, SSIM [41], and LPIPS [42] are used.

Peak signal-to-noise ratio (PSNR) indicates the ratio between the maximum power a signal can reach and the noise power that can affect it. The calculation method can be formed as follows:

$$PSNR(I, R) = 10 \log_{10} \left[ \frac{(2^L - 1)^2}{MSE(I, R)} \right] \quad (11)$$

where  $MSE$  is the mean squared error. The  $MSE$  can be calculated as

$$MSE(I, R) = \frac{1}{mn} \sum_{i=1}^m \sum_{j=1}^n ||I(i, j) - R(i, j)||^2 \quad (12)$$

where  $I$  means the images captured under normal-light conditions,  $R$  means the enhanced image, and  $L$  means the largest gray level, and its value is  $2^n - 1$ .  $n$  is the number of pixel bits. When the PSNR is bigger, the enhancement affection is better.

Structural similarity (SSIM) can be used to evaluate the similarity between two images and the value range from  $-1$  to  $1$ , and a large value indicates that the enhanced image is more structurally similar to the image captured under normal-light conditions. It is measured from three aspects: brightness, contrast, and structure. The SSIM between image  $I$  and image  $R$  can be calculated as

$$SSIM(I, R) = \frac{(2\mu_I\mu_R + c_1)(2\delta_{IR} + C_2)}{(\mu_I^2 + \mu_R^2 + C_1)(\delta_I^2 + \delta_R^2 + C_2)} \quad (13)$$

where  $\mu_I$  means the luminance, and it can be calculated as the mean of all pixel values. The formula can be represented as

$$\mu_I = \frac{1}{N} \sum_{i=1}^N I_i \quad (14)$$

Learning perceptual image patch similarity (LPIPS) is also used as an evaluation index, which uses deep features to measure the perceptual similarity of images. This indicator is learned through deep learning methods. Compared with PSNR and SSIM, LPIPS can more truly reflect the human eye's perception of image quality.

#### 4.3. Comparison with State-of-the-Art Methods

The public LOL dataset is used to evaluate the effectiveness of the proposed DCTE-LLIE method, and multiple deep-learning-based methods are also compared with the proposed method. The comparison results are shown in Table 1. It is easy to understand that compared with generic methods, deep-learning-based methods can achieve more outstanding performance, which establishes the feature extraction ability that deep learning technology showed in the low-light image enhancement task. So, in this paper, only deep-learning-based enhancement methods are used for comparison with the proposed DCTE-LLIE method.

**Table 1.** The performance comparison between multiple low-light image enhancement methods and the proposed method on the LOL dataset.

Method	PSNR	SSIM	LPIPS
LLNet [43]	15.96	0.611	0.529
LightenNet [44]	15.35	0.604	0.541
RetinexNet [21]	16.77	0.635	0.436
MBLLEN [45]	17.91	0.729	0.356
KinD [46]	20.87	0.804	0.207
KinD++ [47]	21.30	0.822	0.175
TBEFN [48]	19.35	0.671	0.237
DSLRL [49]	19.05	0.723	0.244
ElighenGAN [50]	17.48	0.683	0.314
DRBN [51]	19.86	0.748	0.261
ExCNet [52]	18.78	0.716	0.282
Zero-DCE [53]	18.86	0.734	0.311
RRDNet [54]	21.39	0.791	0.157
LLFlow [24]	24.13	0.872	0.117
DCTE-LLIE	24.21	0.826	0.131

As shown in Table 1, one can easily find that the proposed method can achieve outstanding performance on the low-light image enhancement task. This is mainly because the proposed DCTE-LLIE method can extract more realistic color representations and more reliable region representations during the enhancement process. The proposed color enhancement block can help the proposed network maintain color distribution during the enhancement process, while the multiscale attention-based texture enhancement block can help extract more reliable multiscale local texture representations. So the proposed DCTE-LLIE method can achieve better performance on three evaluation indexes than most of the deep-learning-based methods.

#### 4.4. The Effectiveness of Color Enhancement Block

As shown in Table 2, it is easy to find that the performance of the DCTE-LLIE method without a color enhancement block is worse than that of the method with a color enhancement block. This comparison can establish the effectiveness of the proposed color enhancement block on the performance of a low-light image enhancement block. This is because the proposed color enhancement block leveraged a steady color maintenance strategy to extract a more reliable color representation. By calculating the color distribution map and color point map, the color distribution of the input image can be well preserved and the extracted color representation can be more reliable. So, the proposed CEB can make sure the color distribution can be well preserved in the final enhanced image, and the enhancement performance can be effectively improved by the performance comparison.

**Table 2.** The performance of ablation study on the DCTE-LLIE method.

Method	PSNR	SSIM	LPIPS
DCTE-LLIE without CEB	23.05	0.814	0.174
DCTE-LLIE without ATEB	23.68	0.808	0.157
DCTE-LLIE without multiscale feature fusion strategy	23.14	0.796	0.163
DCTE-LLIE	24.21	0.826	0.131

#### 4.5. The Effectiveness of Attention-Based Texture Enhancement Block

As shown in Table 2, the effectiveness of the proposed texture block is also established. By comparing the proposed DCTE-LLIE method with DCTE-LLIE without ATEB, one can easily find that the proposed ATEB block can significantly improve the final enhancement performance. This is because the attention module can help the module focus on multiscale local region texture details and further maintain local details during the enhancement process. Also, by comparing the performance between the DCTE-LLIE method and the DCTE-LLIE method without a multiscale feature fusion strategy, the effectiveness of the proposed multiscale feature fusion strategy can be established. The proposed ATEB can make the network focus on the region of interest from different perspectives, and the U-shape feature extractor can extract multiscale feature representations, which makes the final fused feature representation more reliable and can significantly improve the final enhancement result.

#### 4.6. Further Experimental Results on MIT-Adobe FiveK Dataset

To further evaluate the effectiveness and genetic ability of the proposed method, the proposed DCTE-LLIE method is also applied to the MIT-Adobe FiveK dataset, and its performance is also compared with multiple existing enhancement methods. The experimental results are shown in Table 3. One can also find that the proposed DCTE-LLIE method can still achieve satisfactory enhancement performance, and the performance of the DCTE-LLIE method on PSNR and SSIM is only a little bit worse than that of LLFlow [24], while it is better than other deep-learning-based enhancement methods. This means that the proposed DCTE-LLIE method can still achieve outstanding performance on different enhancement tasks, and the general ability of the proposed method can be promised.

**Table 3.** The performance comparison between multiple low-light image enhancement methods with the proposed method on the MIT-Adobe FiveK dataset.

Method	PSNR	SSIM	LPIPS
LLNet [43]	15.84	0.624	0.537
LightenNet [44]	15.57	0.618	0.528
RetinexNet [21]	16.83	0.674	0.462
MBLEN [45]	18.15	0.740	0.374
KinD [46]	19.22	0.783	0.241
KinD++ [47]	20.41	0.801	0.188
TBEFN [48]	19.06	0.781	0.237
DSLR [49]	18.71	0.739	0.244
ElightenGAN [50]	17.07	0.655	0.339
DRBN [51]	18.53	0.718	0.320
ExCNet [52]	18.04	0.724	0.401
Zero-DCE [53]	17.33	0.736	0.397
RRDNet [54]	20.14	0.785	0.196
LLFlow [24]	22.64	0.824	0.149
DCTE-LLIE	22.39	0.813	0.137

#### 4.7. Further Experimental Results on the SID Dataset

The proposed DCTE-LLIE method is also applied to the SID dataset. The SID dataset has two subsets,  $SID_{Sony}$  and  $SID_{Fuji}$ , corresponding to images captured under a Sony

camera and a Fuji camera, respectively. The experiments on both subsets are shown in Table 4.

**Table 4.** The performance comparison between multiple low-light image enhancement methods with the proposed method on the SID dataset.

Dataset	SID <sub>Sony</sub>			SID <sub>Fuji</sub>		
	Method	PSNR	SSIM	LPIPS	PSNR	SSIM
LLNet [43]	15.21	0.602	0.527	15.49	0.611	0.547
LightenNet [44]	15.77	0.611	0.533	16.04	0.602	0.516
RetinexNet [21]	16.81	0.647	0.419	16.75	0.688	0.497
MBLLEN [45]	17.69	0.686	0.385	17.53	0.670	0.402
KinD [46]	20.03	0.788	0.181	20.36	0.796	0.193
KinD++ [47]	20.97	0.801	0.149	21.13	0.814	0.161
TBEFN [48]	19.78	0.783	0.274	19.62	19.83	0.261
DSLR [49]	19.41	0.774	0.258	19.73	19.54	0.245
ElightenGAN [50]	18.50	0.758	0.284	18.37	0.703	0.278
DRBN [51]	19.11	0.773	0.262	19.02	0.779	0.283
ExCNet [52]	18.47	0.713	0.302	18.61	0.714	0.297
Zero-DCE [53]	18.59	0.748	0.281	18.28	0.743	0.272
RRDNet [54]	21.33	0.792	0.154	20.69	0.782	0.188
LLFlow [24]	23.15	0.824	0.129	22.83	0.820	0.120
DCTE-LLIE	23.04	0.829	0.125	23.32	0.818	0.127

It is easy to find out that the proposed DCTE-LLIE method can still achieve outstanding performance on the SID dataset, no matter whether on a Fuji camera or a Sony camera. Compared with existing low-light enhancement methods, the proposed method can achieve the best performance. This is because the proposed method can extract realistic color representation and reliable local region texture representation, which can further improve the final enhancement performance. At the same time, the experimental results can prove that the proposed method can be applied to different equipment, which further confirms the robustness of the proposed method.

#### 4.8. Visualization Performance on Real-World Scenarios

Some visual comparisons between the proposed DCTE-LLIE method and other enhancement methods are also shown in this paper. Some real-world low-light images captured from real-world scenarios at noon are used to evaluate the performance of the proposed method. As shown in Figures 4–7, one can easily find that most of the existing methods struggle with extracting realistic color and texture presentations. For example, LIME overenhanced the moderately brightness area, which makes the final result particularly unnatural and exhibits varying degrees of color distortion; RetinexNet can effectively improve the brightness of input image, but there still exists color distortion and local region blurring effect; and the EnlightenGAN method is based on the GAN model, and this method can enhance the illumination better while producing less noise. However, compared with the ground truth image, there are still weak light areas in the local area, and the contrast is not high. Among all the enhanced images, the result of the proposed DCTE-LLIE method is much closer to the ground truth image.

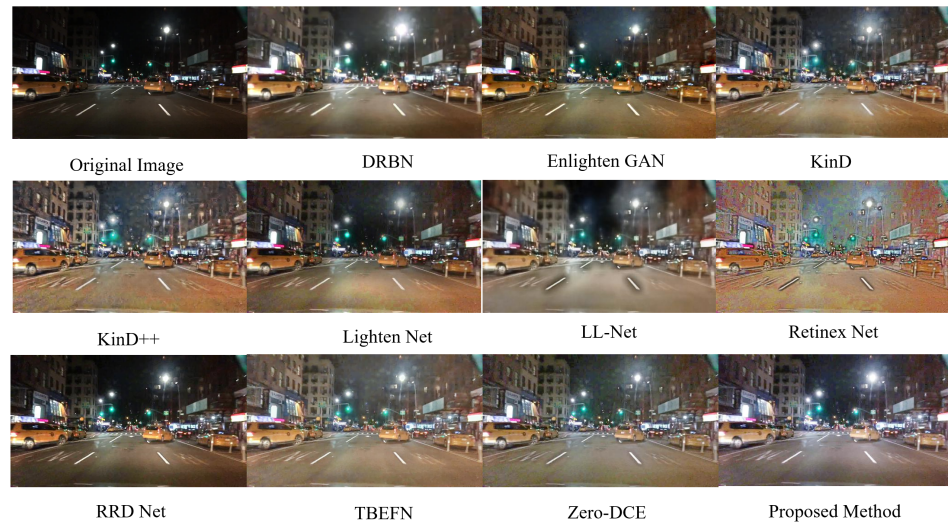


Figure 4. Performance comparison A.

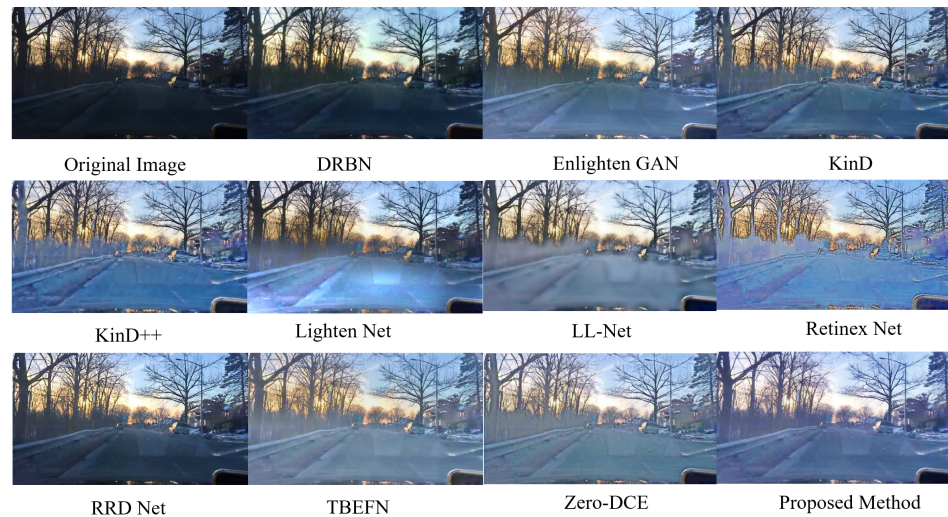


Figure 5. Performance comparison B.



Figure 6. Performance comparison C.

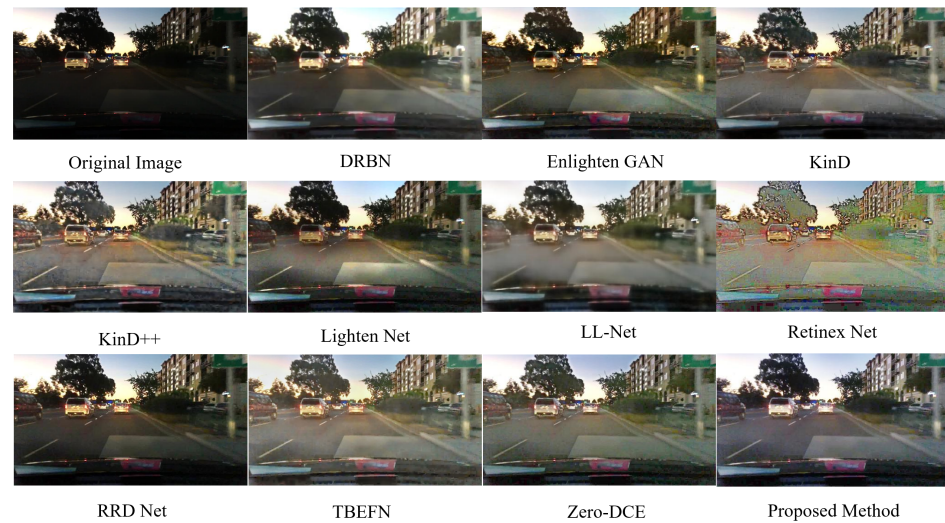


Figure 7. Performance comparison D.

## 5. Conclusions

In this paper, a novel deep learning and Retinex-theory-based method is proposed to deal with low-light image enhancement. Considering that existing enhancement methods can lead to color distortion and blur effects during the enhancement process, the proposed methods have the following improvements. Firstly, a color enhancement block is proposed to help the network extract more realistic color representations. Secondly, a multiscale attention mechanism is applied to the texture enhancement block to help the network focus on the local region areas and further help extract multiscale feature representations, and the extracted multiscale feature representations are finally fused to achieve a final enhanced reflection component through multiscale feature fusion strategy. The experimental results on public datasets also established the effectiveness of the proposed method on low-light image enhancement tasks.

**Author Contributions:** Conceptualization, H.W., J.C., L.Y., and J.H.; methodology, H.W., J.C., L.Y., and J.H.; software, H.W.; validation, H.W., J.C., L.Y., and J.H.; formal analysis, H.W.; investigation, H.W. and J.C.; resources, H.W. and J.C.; data curation, H.W.; writing—original draft preparation, H.W.; writing—review and editing, J.C., L.Y. and J.H.; visualization, H.W.; supervision, J.C., L.Y., and J.H.; project administration, J.C.; funding acquisition, J.C. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work is supported by the National Science Basic Research Plan in Shannxi Province of China (program: No. 2023-JQ-QC-0714).

**Data Availability Statement:** The original data presented in the study are openly available at <https://daooshee.github.io/BMVC2018website/> (LOL Dataset, accessed on 10 April 2024); <https://cchen156.github.io/SID.html> (SID Dataset, accessed on 10 April 2024); <https://data.csail.mit.edu/graphics/fivek/> (MIT-Adobe FiveK Dataset, accessed on 10 April 2024).

**Conflicts of Interest:** The authors declare no conflicts of interest.

## References

1. Liu, J.; Xu, D.; Yang, W.; Fan, M.; Huang, H. Benchmarking low-light image enhancement and beyond. *Int. J. Comput. Vis.* **2021**, *129*, 1153–1184. [CrossRef]
2. Ibrahim, H.; Kong, N.S.P. Brightness preserving dynamic histogram equalization for image contrast enhancement. *IEEE Trans. Consum. Electron.* **2007**, *53*, 1752–1758. [CrossRef]
3. Nakai, K.; Hoshi, Y.; Taguchi, A. Color image contrast enhancement method based on differential intensity/saturation gray-levels histograms. In Proceedings of the 2013 International Symposium on Intelligent Signal Processing and Communication Systems, Naha, Japan, 12–15 November 2013; pp. 445–449.
4. Celik, T.; Tjahjadi, T. Contextual and variational contrast enhancement. *IEEE Trans. Image Process.* **2011**, *20*, 3431–3441. [CrossRef] [PubMed]

5. Lee, C.; Lee, C.; Kim, C.S. Contrast enhancement based on layered difference representation of 2D histograms. *IEEE Trans. Image Process.* **2013**, *22*, 5372–5384. [[CrossRef](#)]
6. Land, E.H. The retinex theory of color vision. *Sci. Am.* **1974**, *237*, 108–128. [[CrossRef](#)] [[PubMed](#)]
7. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436–444. [[CrossRef](#)] [[PubMed](#)]
8. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. ImageNet classification with deep convolutional neural networks. *Commun. ACM* **2017**, *60*, 84–90. [[CrossRef](#)]
9. Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. *arXiv* **2014**, arXiv:1409.1556.
10. Zhang, Y.; Li, W.; Sun, W.; Tao, R.; Du, Q. Single-source domain expansion network for cross-scene hyperspectral image classification. *IEEE Trans. Image Process.* **2023**, *32*, 1498–1512. [[CrossRef](#)]
11. Bhatti, U.A.; Huang, M.; Neira-Molina, H.; Marjan, S.; Baryalai, M.; Tang, H.; Wu, G.; Bazai, S.U. MFFCG–Multi feature fusion for hyperspectral image classification using graph attention network. *Expert Syst. Appl.* **2023**, *229*, 120496. [[CrossRef](#)]
12. Liu, W.; Anguelov, D.; Erhan, D.; Szegedy, C.; Reed, S.; Fu, C.Y.; Berg, A.C. Ssd: Single shot multibox detector. In Proceedings of the Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, 11–14 October 2016; Springer: Cham, Switzerland, 2016; pp. 21–37.
13. Ren, S.; He, K.; Girshick, R.; Sun, J. Faster r-cnn: Towards real-time object detection with region proposal networks. In Proceedings of the Advances in Neural Information Processing Systems 28, Montreal, QC, Canada, 7–12 December 2015.
14. Zou, Z.; Chen, K.; Shi, Z.; Guo, Y.; Ye, J. Object detection in 20 years: A survey. *Proc. IEEE* **2023**, *111*, 257–276. [[CrossRef](#)]
15. Chen, S.; Sun, P.; Song, Y.; Luo, P. Diffusiondet: Diffusion model for object detection. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Paris, France, 2–6 October 2023; pp. 19830–19843.
16. Wang, Z.; Li, Y.; Chen, X.; Lim, S.N.; Torralba, A.; Zhao, H.; Wang, S. Detecting everything in the open world: Towards universal object detection. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Vancouver, BC, Canada, 17–24 June 2023; pp. 11433–11443.
17. Chen, L.C.; Papandreou, G.; Kokkinos, I.; Murphy, K.; Yuille, A.L. Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *40*, 834–848. [[CrossRef](#)] [[PubMed](#)]
18. He, K.; Gkioxari, G.; Dollár, P.; Girshick, R. Mask r-cnn. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017; pp. 2961–2969.
19. Jain, J.; Li, J.; Chiu, M.T.; Hassani, A.; Orlov, N.; Shi, H. Oneformer: One transformer to rule universal image segmentation. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Vancouver, BC, Canada, 17–24 June 2023; pp. 2989–2998.
20. Wu, J.; Fu, R.; Fang, H.; Zhang, Y.; Yang, Y.; Xiong, H.; Liu, H.; Xu, Y. Medsegdiff: Medical image segmentation with diffusion probabilistic model. In Proceedings of the Medical Imaging with Deep Learning, Nashville, NJ, USA, 10–12 July 2023; pp. 1623–1639.
21. Wei, C.; Wang, W.; Yang, W.; Liu, J. Deep retinex decomposition for low-light enhancement. *arXiv* **2018**, arXiv:1808.04560.
22. Wang, L.W.; Liu, Z.S.; Siu, W.C.; Lun, D.P. Lightning network for low-light image enhancement. *IEEE Trans. Image Process.* **2020**, *29*, 7984–7996. [[CrossRef](#)]
23. Zhao, Z.; Xiong, B.; Wang, L.; Ou, Q.; Yu, L.; Kuang, F. RetinexDIP: A unified deep framework for low-light image enhancement. *IEEE Trans. Circuits Syst. Video Technol.* **2021**, *32*, 1076–1088. [[CrossRef](#)]
24. Wang, Y.; Wan, R.; Yang, W.; Li, H.; Chau, L.P.; Kot, A. Low-light image enhancement with normalizing flow. In Proceedings of the AAAI Conference on Artificial Intelligence, Vancouver, BC, Canada, 22 February–1 March 2022; Volume 36, pp. 2604–2612.
25. Hai, J.; Xuan, Z.; Yang, R.; Hao, Y.; Zou, F.; Lin, F.; Han, S. R2rnet: Low-light image enhancement via real-low to real-normal network. *J. Vis. Commun. Image Represent.* **2023**, *90*, 103712. [[CrossRef](#)]
26. Yang, S.; Zhou, D.; Cao, J.; Guo, Y. LightingNet: An integrated learning method for low-light image enhancement. *IEEE Trans. Comput. Imaging* **2023**, *9*, 29–42. [[CrossRef](#)]
27. Guo, X.; Hu, Q. Low-light image enhancement via breaking down the darkness. *Int. J. Comput. Vis.* **2023**, *131*, 48–66. [[CrossRef](#)]
28. Cai, Y.; Bian, H.; Lin, J.; Wang, H.; Timofte, R.; Zhang, Y. Retinexformer: One-stage retinex-based transformer for low-light image enhancement. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Paris, France, 2–6 October 2023; pp. 12504–12513.
29. Pizer, S.M.; Amburn, E.P.; Austin, J.D.; Cromartie, R.; Geselowitz, A.; Greer, T.; ter Haar Romeny, B.; Zimmerman, J.B.; Zuiderveld, K. Adaptive histogram equalization and its variations. *Comput. Vis. Graph. Image Process.* **1987**, *39*, 355–368. [[CrossRef](#)]
30. Hussain, K.; Rahman, S.; Khaled, S.M.; Abdullah-Al-Wadud, M.; Shoyaib, M. Dark image enhancement by locally transformed histogram. In Proceedings of the 8th International Conference on Software, Knowledge, Information Management and Applications (SKIMA 2014), Dhaka, Bangladesh, 18–20 December 2014; pp. 1–7.
31. Liu, B.; Jin, W.; Chen, Y.; Liu, C.; Li, L. Contrast enhancement using non-overlapped sub-blocks and local histogram projection. *IEEE Trans. Consum. Electron.* **2011**, *57*, 583–588. [[CrossRef](#)]
32. Kim, J.Y.; Kim, L.S.; Hwang, S.H. An advanced contrast enhancement using partially overlapped sub-block histogram equalization. *IEEE Trans. Circuits Syst. Video Technol.* **2001**, *11*, 475–484.
33. Dong, X.; Pang, Y.; Wen, J. Fast efficient algorithm for enhancement of low lighting video. In Proceedings of the ACM SIGGRAPH 2010 Posters, Los Angeles, CA, USA, 26–30 July 2010; p. 1.

34. Jobson, D.J.; Rahman, Z.-u.; Woodell, G.A. Properties and performance of a center/surround retinex. *IEEE Trans. Image Process.* **1997**, *6*, 451–462. [[CrossRef](#)] [[PubMed](#)]
35. Jobson, D.J.; Rahman, Z.-u.; Woodell, G.A. A multiscale retinex for bridging the gap between color images and the human observation of scenes. *IEEE Trans. Image Process.* **1997**, *6*, 965–976. [[CrossRef](#)] [[PubMed](#)]
36. Wang, S.; Zheng, J.; Hu, H.M.; Li, B. Naturalness preserved enhancement algorithm for non-uniform illumination images. *IEEE Trans. Image Process.* **2013**, *22*, 3538–3548. [[CrossRef](#)] [[PubMed](#)]
37. Guo, X.; Li, Y.; Ling, H. LIME: Low-light image enhancement via illumination map estimation. *IEEE Trans. Image Process.* **2016**, *26*, 982–993. [[CrossRef](#)] [[PubMed](#)]
38. Li, M.; Liu, J.; Yang, W.; Guo, Z. Joint denoising and enhancement for low-light images via retinex model. In Proceedings of the International Forum on Digital TV and Wireless Multimedia Communications, Shanghai, China, 8–9 November 2017; Springer: Singapore, 2017; pp. 91–99.
39. Li, M.; Liu, J.; Yang, W.; Sun, X.; Guo, Z. Structure-revealing low-light image enhancement via robust retinex model. *IEEE Trans. Image Process.* **2018**, *27*, 2828–2841. [[CrossRef](#)] [[PubMed](#)]
40. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, Ł.; Polosukhin, I. Attention is all you need. In Proceedings of the Advances in Neural Information Processing Systems 30, Long Beach, CA, USA, 4–9 December 2017.
41. Wang, Z.; Bovik, A.C.; Sheikh, H.R.; Simoncelli, E.P. Image quality assessment: From error visibility to structural similarity. *IEEE Trans. Image Process.* **2004**, *13*, 600–612. [[CrossRef](#)]
42. Zhang, R.; Isola, P.; Efros, A.A.; Shechtman, E.; Wang, O. The unreasonable effectiveness of deep features as a perceptual metric. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–22 June 2018; pp. 586–595.
43. Lore, K.G.; Akintayo, A.; Sarkar, S. LLNet: A deep autoencoder approach to natural low-light image enhancement. *Pattern Recognit.* **2017**, *61*, 650–662. [[CrossRef](#)]
44. Li, C.; Guo, J.; Porikli, F.; Pang, Y. LightenNet: A convolutional neural network for weakly illuminated image enhancement. *Pattern Recognit. Lett.* **2018**, *104*, 15–22. [[CrossRef](#)]
45. Lv, F.; Lu, F.; Wu, J.; Lim, C. MBLLN: Low-Light Image/Video Enhancement Using CNNs. In Proceedings of the BMVC Newcastle, Newcastle, UK, 3–6 September 2018; Volume 220, p. 4.
46. Zhang, Y.; Zhang, J.; Guo, X. Kindling the darkness: A practical low-light image enhancer. In Proceedings of the 27th ACM International Conference on Multimedia, Nice, France, 21–25 October 2019; pp. 1632–1640.
47. Zhang, Y.; Guo, X.; Ma, J.; Liu, W.; Zhang, J. Beyond brightening low-light images. *Int. J. Comput. Vis.* **2021**, *129*, 1013–1037. [[CrossRef](#)]
48. Lu, K.; Zhang, L. TBFEFN: A two-branch exposure-fusion network for low-light image enhancement. *IEEE Trans. Multimed.* **2020**, *23*, 4093–4105. [[CrossRef](#)]
49. Lim, S.; Kim, W. DSLR: Deep stacked Laplacian restorer for low-light image enhancement. *IEEE Trans. Multimed.* **2020**, *23*, 4272–4284. [[CrossRef](#)]
50. Jiang, Y.; Gong, X.; Liu, D.; Cheng, Y.; Fang, C.; Shen, X.; Yang, J.; Zhou, P.; Wang, Z. Enlightengan: Deep light enhancement without paired supervision. *IEEE Trans. Image Process.* **2021**, *30*, 2340–2349. [[CrossRef](#)] [[PubMed](#)]
51. Yang, W.; Wang, S.; Fang, Y.; Wang, Y.; Liu, J. 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, Seattle, WA, USA, 14–19 June 2020; pp. 3063–3072.
52. Zhang, L.; Zhang, L.; Liu, X.; Shen, Y.; Zhang, S.; Zhao, S. Zero-shot restoration of back-lit images using deep internal learning. In Proceedings of the 27th ACM International Conference on Multimedia, Nice, France, 21–25 October 2019; pp. 1623–1631.
53. Guo, C.; Li, C.; Guo, J.; Loy, C.C.; Hou, J.; Kwong, S.; Cong, R. Zero-reference deep curve estimation for low-light image enhancement. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 14–19 June 2020; pp. 1780–1789.
54. Zhu, A.; Zhang, L.; Shen, Y.; Ma, Y.; Zhao, S.; Zhou, Y. Zero-shot restoration of underexposed images via robust retinex decomposition. In Proceedings of the 2020 IEEE International Conference on Multimedia and Expo (ICME), London, UK, 6–10 July 2020; pp. 1–6.

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.