Article

A Two-Stage Network for Zero-Shot Low-Illumination Image Restoration

Hao Tang [1], Linfeng Fei [1], Hongyu Zhu [1], Huanjie Tao [2] and Chao Xie [1,3,*]

1 College of Mechanical and Electronic Engineering, Nanjing Forestry University, Nanjing 210037, China
2 School of Computer Science, Northwestern Polytechnical University, Xi'an 710072, China
3 College of Landscape Architecture, Nanjing Forestry University, Nanjing 210037, China
* Correspondence: chaoxie@njfu.edu.cn

Abstract: Due to the influence of poor lighting conditions and the limitations of existing imaging equipment, captured low-illumination images produce noise, artifacts, darkening, and other unpleasant visual problems. Such problems will have an adverse impact on the following high-level image understanding tasks. To overcome this, a two-stage network is proposed in this paper for better restoring low-illumination images. Specifically, instead of manipulating the raw input directly, our network first decomposes the low-illumination image into three different maps (i.e., reflectance, illumination, and feature) via a Decom-Net. During the decomposition process, only reflectance and illumination are further denoised to suppress the effect of noise, while the feature is preserved to reduce the loss of image details. Subsequently, the illumination is deeply adjusted via another well-designed subnetwork called Enhance-Net. Finally, the three restored maps are fused together to generate the final enhanced output. The entire proposed network is optimized in a zero-shot fashion using a newly introduced loss function. Experimental results demonstrate that the proposed network achieves better performance in terms of both objective evaluation and visual quality.

Keywords: low-illumination image enhancement; zero-shot learning; Retinex theory; image feature

1. Introduction

In the field of image processing, low-illumination image restoration is one of the most important branches and can be used in a wide range of advanced vision tasks such as military, surveillance, and security [1,2]. However, due to the limitations in recording equipment and environmental factors, images and videos, especially when captured at night, are severely degraded and large amounts of information can be lost in high-level visual processing tasks [3,4]. Although a longer exposure time improves the image brightness to some extent, it has few practical applications [5]. The aim of low-light image enhancement is to highlight useful features of an image while weakening or eliminating noise and improving contrast to produce better visual perception for human eyes [6,7]. Researchers have proposed a large number of low-light image enhancement algorithms, and they can be roughly divided into conventional enhancement methods and deep-learning-based methods [8,9].

Conventional ones are mainly developed on the basis of image histograms [10] and Retinex theory [11]. Based on histogram equalization (HE), Zhu et al. [12] proposed local histogram equalization, which uses a sliding window to chunk the image, resulting in significant local detail enhancement compared to original HE. Since then, sliding window overlap algorithms, sliding window partial overlap algorithms, and so on have emerged successively. Retinex theory decomposes the image into a light map and a reflectance map. Earlier algorithms used the reflection map as an image enhancement effect, and the method would amplify the noise in the image, which would affect the final result. Researchers have made many improvements to address the problem of image noise, including the single

scale Retinex algorithm (SSR) [13], multi-scale Retinex algorithm (MSR) [14], multi-scale Retinex algorithm with image recovery (MSRCR) [15], multi-scale Retinex algorithm based on HSV space, etc.

Deep learning has achieved good performance in the field of low-illumination image enhancement because of its powerful feature representation and non-linear mapping capabilities. Lore et al. [16] were the first to implement low-light image enhancement (LLNet) using a deep learning approach, which proposes a deep autoencoder to identify signal features from low-light images, capable of adaptively enhancing image brightness without oversaturation. The increase in the image datasets in terms of quality as well as in quantity has given rise to many Retinex-based deep learning methods that enable low-light image enhancement by better estimating the reflectance component and enhancing the luminance of illuminated images. Some of the Retinex-based deep learning methods incorporate BM3D or channel attention mechanisms as a way of reducing image noise and extracting image features to achieve better visual results [17]. Subsequent methods have been improved in terms of targeting the limitations of the dataset, as well as overfitting and real-time problems, improving the application of deep learning methods in the field of low-illumination image enhancement.

Deep-learning-based approaches are the main trend in this research direction [18]. Although some of the low-light image enhancement algorithms based on deep learning have achieved remarkable results, there are still several problems to be solved. Firstly, existing methods mostly use fully supervised learning, although the illumination estimation under such a framework is inherently ill-posed. Secondly, the enhancement effect of fully supervised learning methods is heavily influenced by their training datasets, and thus there are some common problems such as loss of image details, noise amplification, and color distortion in the enhancement effect. Therefore, we propose a low-illumination enhancement algorithm that requires only a single low-illumination image input in this paper, and our contributions are summarized as follows:

- We propose a two-stage low-illumination image restoration network, in which a pre-decomposition submodule is incorporated to divide the original image into illumination, reflectance, and feature. Moreover, the whole network is optimized in a zero-shot way instead of using supervised learning.
- To guide the decomposition network focusing on the dark area of the image, a new loss function is proposed for our network. The loss function can also obtain relatively clearer texture features in the dark area, and avoid the problem of overexposure or underexposure in the others areas.
- Experiments show that our method achieves better performance on the benchmark datasets. Compared with recent methods based on decomposition theory, the proposed method can visually better retain the detailed features of images and avoid the problem of overexposure. There is a significant improvement in *PSNR* and *SSIM* (reference evaluation indices), and *NIQE* and *LOE* (no-reference evaluation indices).

The rest of this paper is organized as follows: In Section 2, we review representative Retinex-model-based and deep-learning-based methods for low-illumination image enhancement. In Section 3, the model proposed in this paper is presented. The fourth section introduces the loss function in detail. Section 5 gives the experimental results and evaluation. The sixth part summarizes the whole paper.

2. Related Work

2.1. Retinex-Model-Based Approach

The Retinex-model-based approach enhances low-illumination images by the idea of decomposition. The principle of Retinex theory, first proposed by Land and widely used in image processing, can be expressed as $S = R * I$, where $*$ denotes elemental multiplication. S is the original image, R is the reflection map, and I is the illumination map. The brightness and contrast of the image are improved by retaining the reflectance properties while adjusting the intensity of the light map. Jobson et al. [11] proposed the

single scale Retinex algorithm (SSR), which uses the construction of a Gaussian surround function to filter each of the three-color channels of the image as the estimated light component, and obtains the output after subtracting the logarithmically processed original image from the light map. The multi-scale Retinex algorithm (MSR) [19] adds the number of Gaussian surround functions to the SSR and can be seen as a weighted summation of multiple SSRs that do not pass the scale. The adaptive multi-scale Retinex algorithm based on the HSV color channel first transforms the image spatially, estimating the image light component in luminance space before transferring back to RGB space [20].

2.2. Learning-Based Approach

In recent years, low-light image enhancement methods based on deep learning have become popular and have made good progress. Starting in 2017, a deep-encoder-based method was first proposed, which uses a variant of the stacked sparse denoising self-encoder to identify features from shimmering images, adaptively enhance and denoise them, and simultaneously highlight and denoise the images [16]. Shen et al. [21] proposed a multi-scale Retinex combined with a CNN model that learns the end-to-end mapping between low-light and bright images directly, with the network structure containing a residual structure that makes full use of the information in each convolutional layer and the relevant parameters set by backpropagation. Wang et al. [22] proposed the GLADNet network structure, which first performs global illumination estimation of low-illumination images and then reorganizes the details using the global part of the generated illumination in connection with the original input, and the connection to the input image complements the details. The aforementioned fully supervised learning methods have achieved good results in self-illuminating image enhancement, but such methods require image pairs that often need to be manually adjusted for parameters. There is still a certain gap between the manually adjusted images and the real images, which limits the generalizability of fully supervised learning methods in real environments.

Compared to fully supervised learning methods, unsupervised and zero-shot learning methods have become more popular in recent years as they effectively avoid reliance on datasets by constructing loss functions and constraining the images themselves to achieve enhancement. Jiang et al. [23] first proposed the unsupervised learning method called EnlightenGAN. EnlightenGAN is the first method to successfully introduce unpaired training into low-light image enhancement. The method creates a mapping between unpaired low-light and normal-light image domains to enable unsupervised training. Fu et al. [24] proposed the low-illumination enhancement network (LE-GAN) using an identity invariant loss and attention module, which enhances the feature extraction of images using an illumination-aware attention module to improve visual quality while achieving noise reduction and detail enhancement. In addition, identity-invariant loss can solve the overexposure problem. Zhang et al. [25] trained a small CNN network (ExCNet), where the network does not need to be trained in advance and then tested. It can estimate the best-fit s-curve directly for a given backlit image. With its s-curve, the backlit image can be recovered accordingly. These unsupervised learning methods have improved considerably in terms of generalization ability, but there is still much to be achieved in terms of image brightness, contrast, and color bias.

Compared with the above methods, the proposed method integrates image smoothing into the decomposition network to extract the global features of the image and avoid the loss of image details caused by the smoothing operation. This method can accurately estimate the reflectance map, illumination map, and feature map of the image. Moreover, this method does not need pairwise image pairs for training, which ensures its ability to generalize to the illumination environment.

3. Methodology

In order to better preserve image details and improve image brightness at the same time, as shown in Figure 1, the method in this paper is divided into two parts: decompo-

sition and enhancement. In the decomposition part, the image is first decomposed into illumination map, reflection map, and feature map. Then the illumination map is separately enhanced, and finally the final effect is obtained by adaptive fusion. The remainder of this section describes both modules in detail.
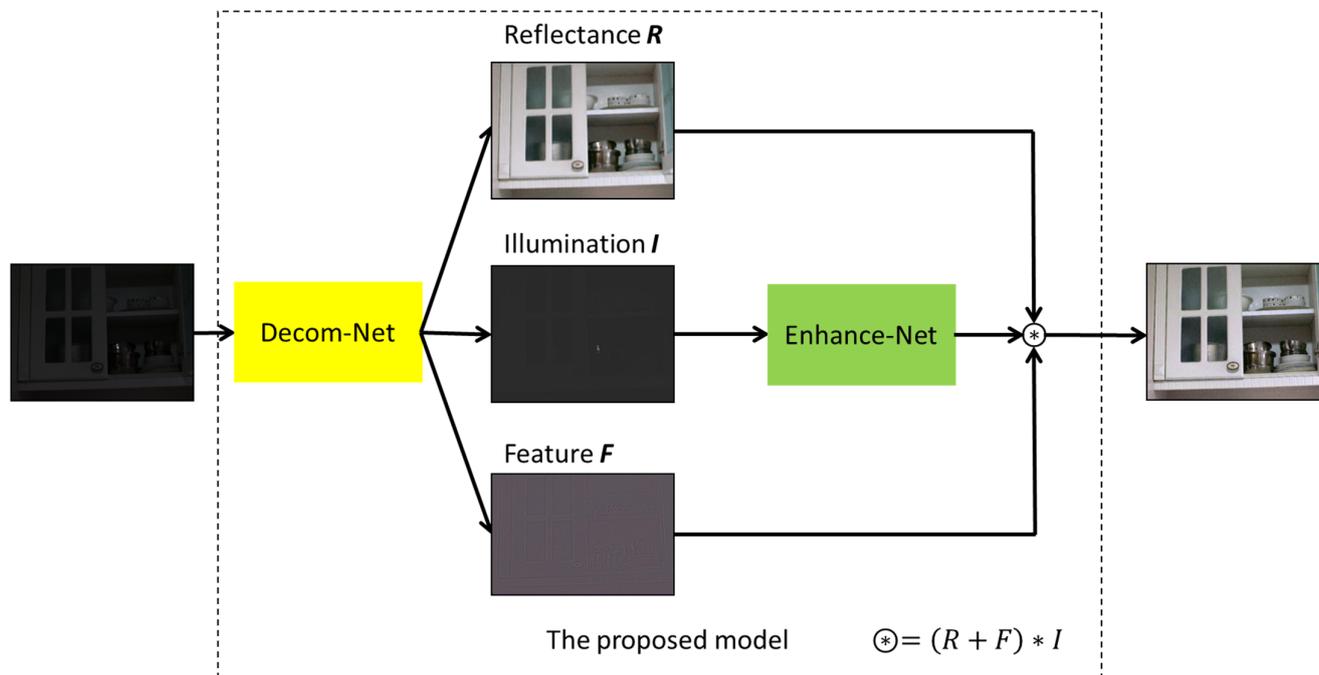

Figure 1. The framework of the proposed model.

3.1. Decom-Net

Our Decom-Net is inspired by the Retinex theory. However, the original RetinexNet proposed by Chen et al. [17] erases many detailed features of the image when noise reduction is applied to the image. In response, we propose a three-branch convolutional decomposition network, shown in Figure 2, with three branches for estimating the illumination map, the reflectance map, and the feature map, respectively. It features a pooling layer in the decomposition process, which allows the illumination and reflectance maps to be denoised during the decomposition process. The Sigmoid function is chosen for the activation functions of the illumination and reflectance maps to ensure that the output is between 0 and 1. The activation function for the feature map is chosen as the tanh function. The output of tanh is between -1 and 1, which has faster model convergence. The decomposition of the image is carried out when the number of iterations has been reached and the loss function has reached its minimum value.

3.2. Enhance-Net

The main purpose of the Enhance-Net is to adjust the brightness of the illumination, as shown in Figure 3. The input of the module is the illumination map output by the Decom-Net. The module is composed of eight convolutional layers, which can effectively obtain the illumination information of the illumination map. In order to make up for the possible loss of the effective information of the illumination map during the process, the input layer is finally spliced to the last layer, and the output is the adjusted illumination map.
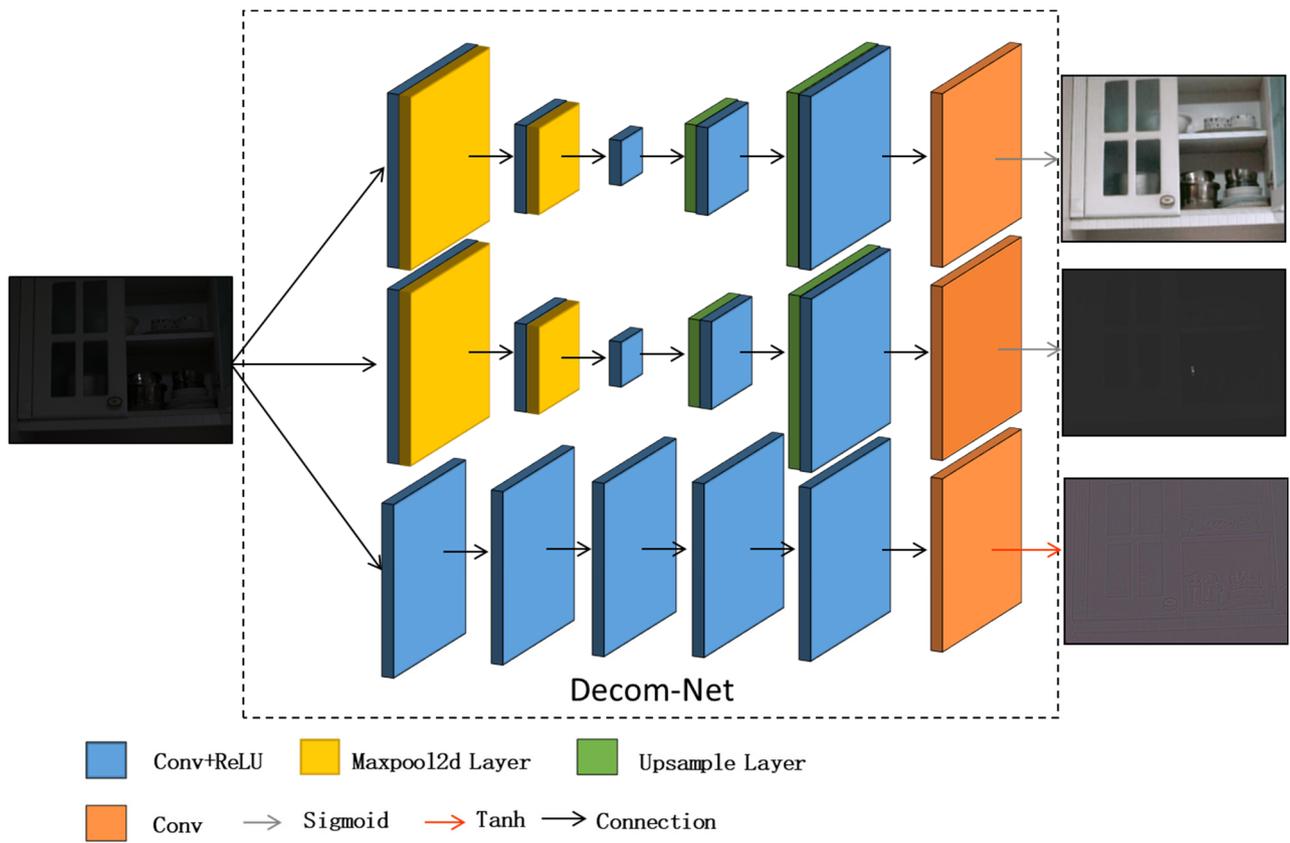
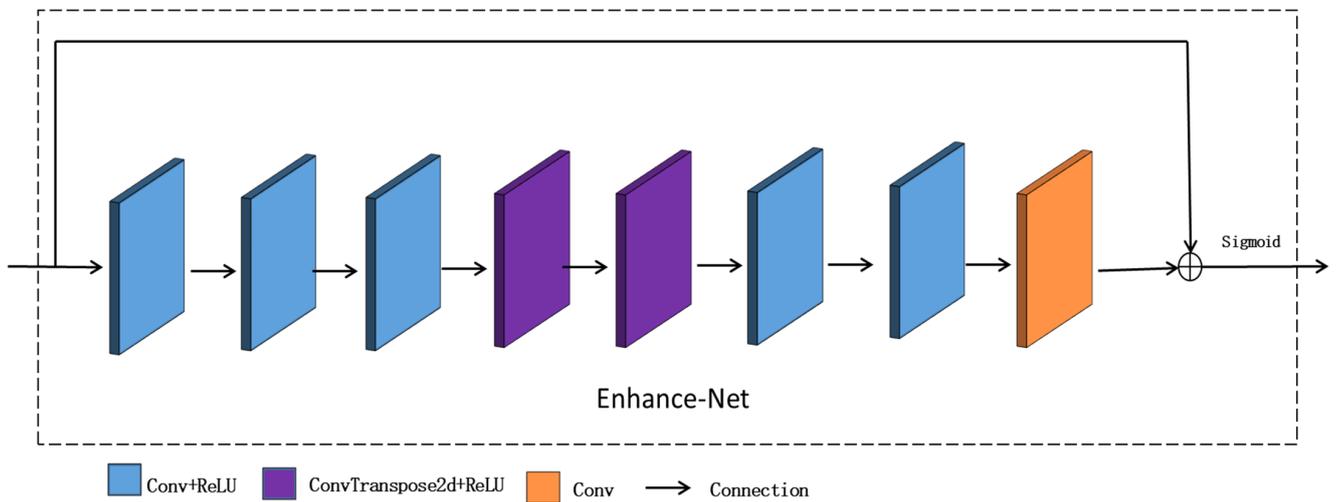**Figure 2.** The model structure diagram of the Decom-Net.

Figure 3. The model structure diagram of the Enhance-Net.

4. Loss Function

We decompose low-illumination image *S* into three parts: illumination map *I*, reflectance map *R*, and feature map *F*.

$$S = (R + F) * I \tag{1}$$

To better configure the network weights, we set up a loss function to guide the current network to generate a more accurate branching section. This loss function is as follows:

$$L = L_{recon} + \lambda_1 L_s + \lambda_2 L_t \tag{2}$$

where L_{recon} is the reconstruction loss function, L_s is the smoothness loss, and L_t is the feature estimation loss. λ_1 and λ_2 are the weight factors.

4.1. Reconstruction Loss

In Retinex theory, the channel maxima of *R*, *G*, and *B* are usually used as an initial estimate of illumination, and the reflectance image is obtained by pixel segmentation between the original image and the illumination map. Moreover, we assume that the three-color channels have the same illumination. Here, we follow the idea of Retinex theory as a constraint on reflectance and illumination. The reconstruction losses in this paper are as follows:

$$L_{recon} = \|S - \widetilde{S}\|_1 + \|I - \widetilde{I}\|_1 + \|R - \widetilde{R}\|_1 \tag{3}$$

where S represents the input image, $\widetilde{S}$ is the reconstructed image, I is the decomposed illumination map, and $\widetilde{I}$ is the color channel maximum. R is the decomposed reflectance map, $\widetilde{R} = S/I$, and L_1-norm is used to bootstrap all loss functions in this paper.

4.2. Smoothness Loss

In this paper, in terms of noise, the reflectance and illumination maps should be properly denoised to avoid the amplified noise on the enhancement effect. The loss function is as follows:

$$L_s = \frac{1}{HWC} \sum_{i=1}^{H} \sum_{w=1}^{W} \sum_{c=1}^{C} \left(\|\Delta_x \widetilde{S}\| + \|\Delta_y \widetilde{S}\| \right) \tag{4}$$

where *H*, *W*, and *C*, respectively, represent the height, width, and channel of the image.

4.3. Feature Estimation Loss

In a low-illumination image enhancement task, it is inevitable to lose some detail features. In the RetinexNet and other methods, smoothing reduces the sharpness of the image itself, and the image becomes blurred. Therefore, it is necessary to extract the image features separately in the network using the smoothing operation. In this paper, weighted guide image features are extracted according to the estimated illumination map. The loss function is as follows:

$$L_t = \|S \cdot F\|_F + \frac{\left[\|\beta \cdot (\alpha_x R)^2\|_1 + \|\beta \cdot (\alpha_y R)^2\|_1 \right]}{\lambda_2} \tag{5}$$

where $\| \ \|_F$ represents the Frobenius norm of the matrix, β is the illumination guidance weight, and its expression is as follows:

$$\beta = normalize \left[I \cdot (\alpha_x R)^2 \cdot (\alpha_y R)^2 \right]^{-1} \tag{6}$$

where *normalize* denotes min-max normalization.

5. Experimental Results and Analysis

In all experiments, we empirically set $\lambda_1 = 0.5$ and $\lambda_2 = 5000$. All experiments in this article were conducted in the same configuration environment, the training environment configuration: Intel I7-8700 CPU, 32 GB RAM, and NVIDIA GeForce RTX2080 Ti GPU. PyTorch framework, PyCharm software in 32 GB environment, and Anaconda Python 3.7 interpreter built the network framework. For the sake of fairness, two low-light image datasets, LOL [17] and 5 K [26], were selected for comparison. The methods selected

for comparison include HE [10], Retinex [19], ExCNet [25], RRDNet [27], LightenNet [28], Zero-DCE [29], DSLR [30], and LLNet [16], and the results of all comparison methods are reproduced from their official code.

5.1. Subjective Evaluation

We show the enhancement effects of various algorithms in Figures 4–7, and enlarge some of the details to better carry out subjective visual evaluation. Figures 4 and 5 belong to the LOL dataset. In Figure 4, HE significantly improves the image brightness by pulling up the contrast, but the overall image distortion is serious. Retinex has the best visual brightness improvement effect in Figure 4, but the clothing color is distorted and obvious noise can be seen in the enlarged image. Compared with the previous two methods, ExCNet avoids the impact of color distortion, but the detail is lost severely and the image is white as a whole. RRDNet performed well on the original brighter images but performed relatively poorly on the darker images, where the brightness boost was not significant, making it difficult to achieve good visual effects. The LightenNet method showed average brightness improvement in the comparison method, but the image showed a white blocky phenomenon. Although Zero-DCE can retain the detail features of the image well, the brightness enhancement is not obvious, and the color contrast of the image is significantly reduced. The enhancement effect of DSLR produces obvious stacked block phenomenon, the whole wardrobe part appears as a block effect, and there are artifacts. For LLNet, it can be seen from the hanger and the enlarged part that details are seriously lost, and the image enhancement effect is blurred as a whole. Compared with other methods, the brightness improvement effect of the proposed method may not be the most ideal, but it effectively avoids other problems, such as color distortion, detail loss and artifact phenomenon, etc., and the brightness of the enhanced effect is improved as a whole without overexposure or underexposure phenomena. Figures 6 and 7 are from the 5 K dataset. In Figure 6, the Retinex method has the highest visual brightness, but the overall image is overexposed and the visible details are seriously lost. The enhancement effect of the HE method is overall white. RRDNet, as with DSLR, has a low brightness boost, which has a big impact on visual effects.
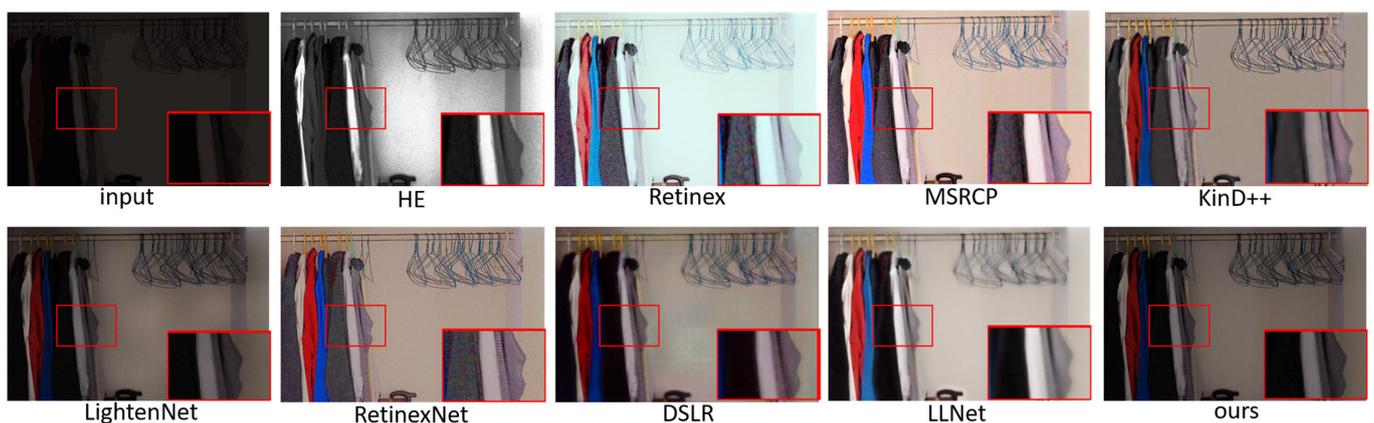

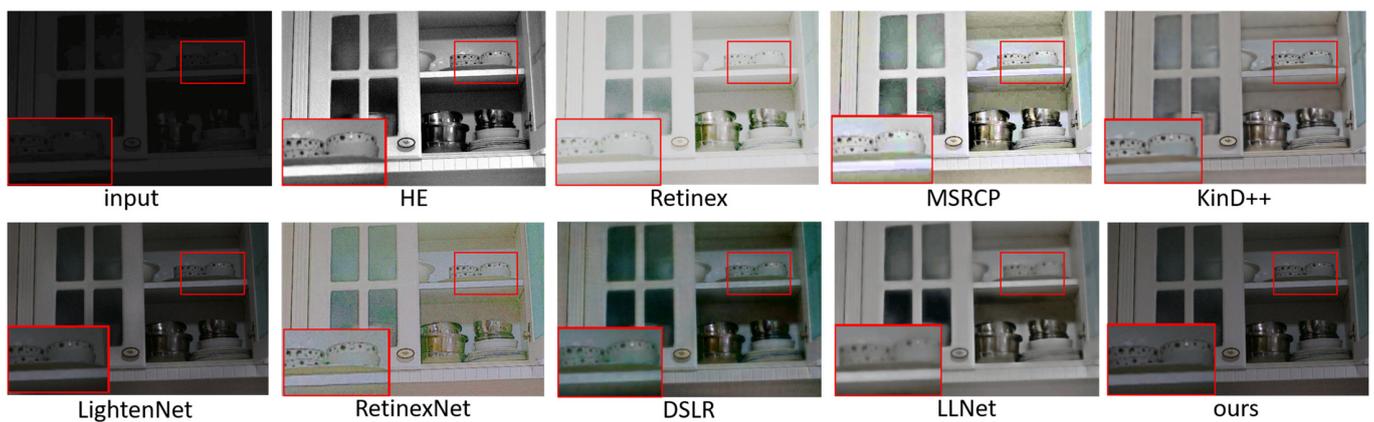**Figure 4.** Visual comparison of enhanced renderings of low-illumination images.

Figure 5. Visual comparison of enhanced renderings of low-illumination images.

Figure 6. Visual comparison of enhanced renderings of low-illumination images.

Figure 7. Visual comparison of enhanced renderings of low-illumination images.
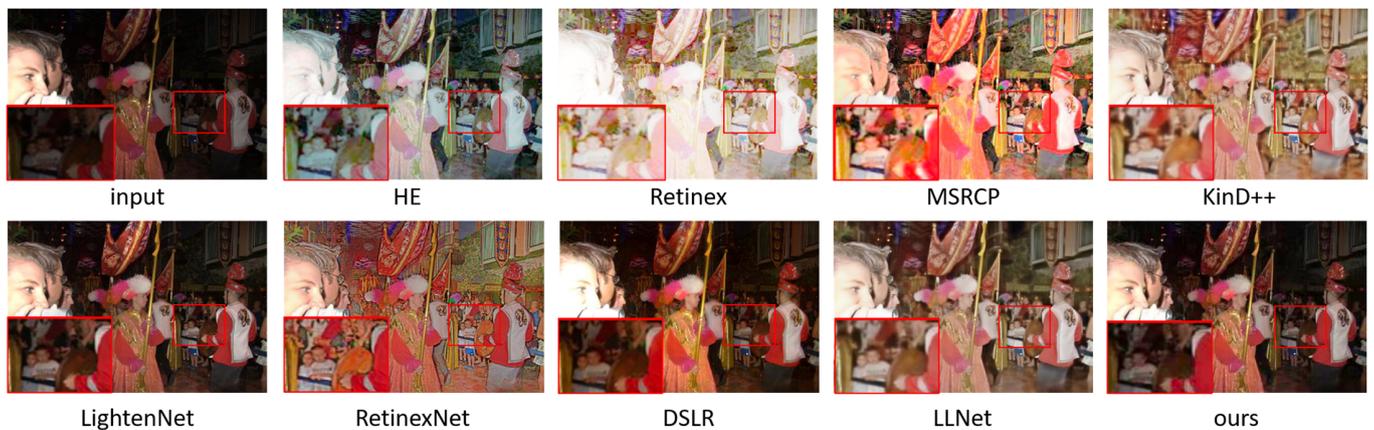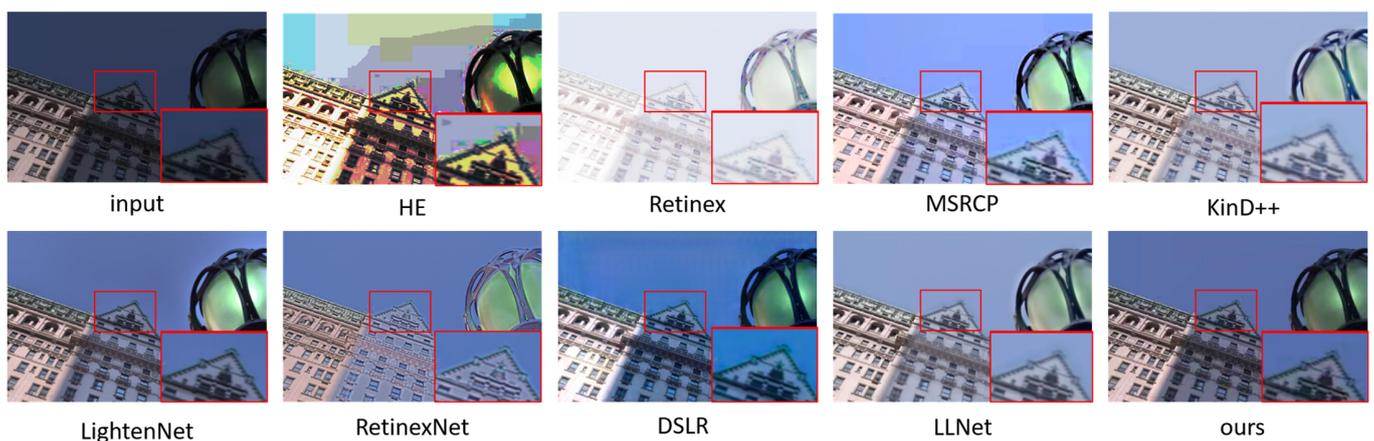
5.2. Objective Evaluation

In addition to subjective visual evaluation, recognized image quality metrics for quantitative comparisons are used to illustrate the effectiveness of the algorithms in this paper. Many image quality evaluation metrics have been proposed in various image processing fields, of which *PSNR* [31] is the most widely used objective evaluation metric. *PSNR* stands for Peak Signal to Noise Ratio and is measured in dB, with larger values indicating less distortion. *SSIM* [32] is a measure of the similarity between two images. *SSIM* uses two images, one processed and the other real, to measure the similarity of the

two images in terms of brightness, contrast, and structure, respectively. The value of *SSIM* is between 0 and 1, and the closer to 1 the higher the similarity. Both of these metrics have references, and for non-paired datasets, there are also non-reference metrics, such as *NIQE* [33], which tests the test images by extracting features from the natural landscape to more closely match the human eye's visual perception. The *LOE* [34] reflects the natural holding power of the image, with smaller values indicating a better brightness order and a more natural look.

PSNR is used to evaluate the differences between images. It is widely used in image quality evaluation of low-level image processing tasks such as image de-fogging, image noise reduction, and image enhancement. The *PSNR* formula can be expressed as:

$$PSNR = 10 \times \log \frac{MaxValue^2}{MSE} \tag{7}$$

where *MSE* is the mean square error between images, and *MaxValue* is the maximum pixel value of two images. The formula of *MSE* can be expressed as:

$$MSE = \frac{1}{M \times N} \sum_{i=1}^{M} \sum_{j=1}^{N} [g(x,y) - \hat{g}(x,y)]^2 \tag{8}$$

where *H* stands for image height and *W* stands for image width, and $g(x,y)$ and $\hat{g}(x,y)$ stand for the original image and enhanced image, respectively.

SSIM is used to highlight the brightness, contrast, and structural similarity between two images, and the value range is 0–1; the closer the value is to 1, the more similar the two images are. Assuming that x and y are two input images, the formula is:

$$SSIM = [l(x,y)]^{\alpha} [C(x,y)]^{\beta} [S(x,y)]^{\gamma} \tag{9}$$

where $l(x,y)$ is brightness comparison, $C(x,y)$ is contrast comparison, and $S(x,y)$ is structural comparison. α, β, γ are greater than 0, and are used to adjust the three-part specific gravity. $l(x,y)$, $C(x,y)$, and $S(x,y)$ have the following formulas, respectively:

$$l(x,y) = \frac{2\mu_x\mu_y + c_1}{\mu_x^2 + \mu_y^2 + c_1} \, , \, C(x,y) = \frac{2\sigma_{xy} + c_2}{\sigma_x^2 + \sigma_y^2 + c_2} \, , S(x,y) = \frac{\sigma_{xy} + c_3}{\sigma_x\sigma_y + c_3} \tag{10}$$

where μ_x and μ_y, respectively, represent the average values of the two images, σ_x and σ_y represent the standard deviations of the two images. σ_{xy} represents the covariance of the two images. The function of c_1, c_2, and c_3 is to avoid the denominator being 0.

LOE is the sequential difference of brightness of an image, and the illumination change of an image is evaluated by evaluating the sequential change process of the brightness of the image in the neighborhood. *LOE* reflects the natural retention ability of the image. A smaller value indicates that the image has a better luminance order and looks more natural. The formula is:

$$LOE = \frac{1}{M \times N} \sum_{i=1}^{M} \sum_{j=1}^{N} RD(i,j) \tag{11}$$

where $RD(i, j)$ is the difference in the relative brightness order between the original image and the enhanced image.

NIQE is based on a set of "quality-aware" features and fits them into the MVG model. The quality perception features are derived from a simple but highly regularized NSS model. Then, the *NIQE* index of a given test image is expressed as the distance between the MVG model of NSS features extracted from the test image and the MVG model of quality perception features extracted from the natural image corpus. The *NIQE* formula is:

$$NIQE = D(v_1, v_2, m_1, m_2) = \sqrt{\left((v_1 - v_2)^T \left(\frac{m_1 + m_2}{2} \right)^{-1} (v_1 - v_2) \right)} \tag{12}$$

where v_1, v_2, m_1, and m_2 represent the mean vector and covariance matrix of the natural MVG model and distorted image MVG model, respectively.

We evaluated the results of the proposed method and eight other representative methods on the *PSNR*, *SSIM*, *NIQE*, and *LOE* indicators on the LOL and 5 K datasets. It can be seen from Tables 1 and 2 that no method can obtain the optimal value among all the image quality detection indicators. However, in the LOL dataset test, our method has the best performance on the *PSNR* index, and the second *SSIM* index is also better than most methods. In the test of the 5 K dataset, the optimal value of the *LOE* index was obtained, followed by *NIQE*, which also obtained the second place. Tables 1 and 2 more strongly illustrate the effectiveness and applicability of the proposed approach.

Table 1. The LOL dataset (paired dataset) was quantitatively evaluated according to *PSNR* and *SSIM*. The best results are shown in bold, the second is italic, and the third is underlined.

Method	PSNR↑	SSIM↑
Input	5.10	0.19
HE [10]	*15.23*	0.59
Retinex [19]	10.92	0.37
ExCNet [25]	<u>14.32</u>	<u>0.75</u>
RRDNet [27]	8.72	0.60
LightenNet [28]	7.87	0.52
Zero-DCE [29]	11.98	*0.76*
DSLR [30]	10.78	0.67
LLNet [16]	14.00	**0.78**
Ours	**17.84**	0.74

Table 2. The 5 K dataset (unpaired dataset) was quantitatively evaluated according to *NIQE* and *LOE*. The best results are shown in bold, the second is italic, and the third is underlined.

Method	NIQE↓	LOE↓
Input	28.12	0
HE [10]	30.76	<u>254.87</u>
Retinex [19]	23.33	291.14
ExCNet [25]	**17.96**	316.85
RRDNet [27]	18.47	*251.37*
LightenNet [28]	20.97	305.50
Zero-DCE [29]	21.50	351.37
DSLR [30]	<u>18.40</u>	272.58
LLNet [16]	26.35	302.76
Ours	*18.02*	**249.25**

6. Conclusions

In this paper, we propose a two-stage zero-shot low-illumination image enhancement network. Considering the mixed noise in low-illumination images and the loss of stylistic features by general methods, the model adopts a smoothing operation in the decomposition network to reduce image noise and obtain image texture feature images, which can effectively avoid the above two problems. Comparative experiments show that the proposed method is more consistent with human perception in the subjective visual angle, and performs well in the comparison of objective evaluation indicators.

Author Contributions: Conceptualization, H.T. (Hao Tang) and C.X.; methodology, H.T. (Hao Tang); software, H.T. (Hao Tang); validation, H.T. (Huanjie Tao); formal analysis, H.Z.; investigation, L.F.; resources, C.X.; data curation, L.F.; writing—original draft preparation, H.T. (Hao Tang); writing—review and editing, C.X.; visualization, H.T. (Huanjie Tao); supervision, C.X.; project administration, C.X.; funding acquisition, C.X. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are openly available at https://arxiv.org/abs/1808.04560 and https://data.csail.mit.edu/graphics/fivek/, accessed on 10 August 2018 and 2019, respectively.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Fang, M.; Li, H.; Lei, L. A review on low light video image enhancement algorithms. *J. Changchun Univ. Sci. Technol.* **2016**, *39*, 56–64.
2. Ackar, H.; Abd Almisreb, A.; Saleh, M.A. A review on image enhancement techniques. *Southeast Eur. J. Soft Comput.* **2019**, *8*, 42–48. [CrossRef]
3. Wang, Y.-F.; Liu, H.-M.; Fu, Z.-W. Low-light image enhancement via the absorption light scattering model. *IEEE Trans. Image Process.* **2019**, *28*, 5679–5690. [CrossRef]
4. Yan, X.; Liu, T.; Fu, M.; Ye, M.; Jia, M. Bearing Fault Feature Extraction Method Based on Enhanced Differential Product Weighted Morphological Filtering. *Sensors* **2022**, *22*, 6184. [CrossRef]
5. Han, Q.; Liu, Y.; Li, Z.; Zhang, Y.; Chen, Y. Investigation of energy accommodation coefficient at gas-solid interface of a hypersonic flying vehicle. *Aerosp. Sci. Technol.* **2022**, *126*, 107585. [CrossRef]
6. Abdullah-Al-Wadud, M.; Kabir, M.H.; Dewan, M.A.A.; Chae, O. A dynamic histogram equalization for image contrast enhancement. *IEEE Trans. Consum. Electron.* **2007**, *53*, 593–600. [CrossRef]
7. Zhu, Y.; Feng, K.; Hua, C.; Wang, X.; Hu, Z.; Wang, H.; Su, H. Model Analysis and Experimental Investigation of Soft Pneumatic Manipulator for Fruit Grasping. *Sensors* **2022**, *22*, 4532. [CrossRef]
8. Li, C.; Guo, C.; Han, L.; Jiang, J.; Cheng, M.-M.; Gu, J.; Loy, C.C. Lighting the darkness in the deep learning era. *arXiv* **2021**, arXiv:2104.10729.
9. Jin, X.; Bagavathiannan, M.; McCullough, P.E.; Chen, Y.; Yu, J. A deep learning-based method for classification, detection, and localization of weeds in turfgrass. *Pest Manag. Sci.* **2022**, *78*, 4809–4821. [CrossRef] [PubMed]
10. Stark, J.A. Adaptive image contrast enhancement using generalizations of histogram equalization. *IEEE Trans. Image Process.* **2000**, *9*, 889–896. [CrossRef] [PubMed]
11. 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]
12. Zhu, H.; Chan, F.H.; Lam, F.K. Image contrast enhancement by constrained local histogram equalization. *Comput. Vis. Image Underst.* **1999**, *73*, 281–290. [CrossRef]
13. Rahman, Z.-u.; Jobson, D.J.; Woodell, G.A. Multi-scale retinex for color image enhancement. In Proceedings of the 3rd IEEE International Conference on Image Processing, Lausanne, Switzerland, 19 September 1996; pp. 1003–1006.
14. Barnard, K.; Funt, B. Investigations into Multi-Scale Retinex. In *Colour Imaging: Vision and Technology*; Wiley: Hoboken, NJ, USA, 1998.
15. Parthasarathy, S.; Sankaran, P. An automated multi scale retinex with color restoration for image enhancement. In Proceedings of the 2012 National Conference on Communications (NCC), Kharagpur, India, 3–5 February 2012; pp. 1–5.
16. 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]
17. Wei, C.; Wang, W.; Yang, W.; Liu, J. Deep retinex decomposition for low-light enhancement. *arXiv* **2018**, arXiv:1808.04560.
18. Wang, W.; Wu, X.; Yuan, X.; Gao, Z. An experiment-based review of low-light image enhancement methods. *IEEE Access* **2020**, *8*, 87884–87917. [CrossRef]
19. 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]
20. Zotin, A. Fast algorithm of image enhancement based on multi-scale retinex. *Procedia Comput. Sci.* **2018**, *131*, 6–14. [CrossRef]
21. Shen, L.; Yue, Z.; Feng, F.; Chen, Q.; Liu, S.; Ma, J. Msr-net: Low-light image enhancement using deep convolutional network. *arXiv* **2017**, arXiv:1711.02488.
22. Wang, W.; Wei, C.; Yang, W.; Liu, J. Gladnet: Low-light enhancement network with global awareness. In Proceedings of the 2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018), Xi'an, China, 15–19 May 2018; pp. 751–755.
23. 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]

24. Fu, Y.; Hong, Y.; Chen, L.; You, S. LE-GAN: Unsupervised low-light image enhancement network using attention module and identity invariant loss. *Knowl. Based Syst.* **2022**, *240*, 108010. [CrossRef]

25. 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.

26. Wang, R.; Zhang, Q.; Fu, C.-W.; Shen, X.; Zheng, W.-S.; Jia, J. Underexposed photo enhancement using deep illumination estimation. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 15–20 June 2019; pp. 6849–6857.

27. 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.

28. 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]

29. 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.

30. Lim, S.; Kim, W. DSLR: Deep stacked Laplacian restorer for low-light image enhancement. *IEEE Trans. Multimed.* **2020**, *23*, 4272–4284. [CrossRef]

31. Yu, C.; Xu, X.; Lin, H.; Xinyan, Y. Low-illumination image enhancement method based on a fog-degraded model. *J. Image Graph.* **2017**, *22*, 1194–1205.

32. Ignatov, A.; Kobyshev, N.; Timofte, R.; Vanhoey, K.; Van Gool, L. Wespe: Weakly supervised photo enhancer for digital cameras. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, Salt Lake City, UT, USA, 18–22 June 2018; pp. 691–700.

33. Mittal, A.; Soundararajan, R.; Bovik, A.C. Making a "completely blind" image quality analyzer. *IEEE Signal Process. Lett.* **2012**, *20*, 209–212. [CrossRef]

34. Fu, X.; Zeng, D.; Huang, Y.; Zhang, X.-P.; Ding, X. A weighted variational model for simultaneous reflectance and illumination estimation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 2782–2790.