

Article

An Improved Retinex-Based Approach Based on Attention Mechanisms for Low-Light Image Enhancement

Shan Jiang ^{1,2}, Yingshan Shi ^{1,2}, Yingchun Zhang ^{1,2} and Yulin Zhang ^{1,2,*}

¹ Key Laboratory of Ethnic Language Intelligent Analysis and Security Governance of MOE, Minzu University of China, Beijing 100081, China; jshan@muc.edu.cn (S.J.); ysshi@muc.edu.cn (Y.S.); 21302036@muc.edu.cn (Y.Z.)

² School of Information Engineering, Minzu University of China, Beijing 100081, China

* Correspondence: yzhang@muc.edu.cn

Abstract: Captured images often suffer from issues like color distortion, detail loss, and significant noise. Therefore, it is necessary to improve image quality for reliable threat detection. Balancing brightness enhancement with the preservation of natural colors and details is particularly challenging in low-light image enhancement. To address these issues, this paper proposes an unsupervised low-light image enhancement approach using a U-net neural network with Retinex theory and a Convolutional Block Attention Module (CBAM). This method leverages Retinex-based decomposition to separate and enhance the reflectance map, ensuring visibility and contrast without introducing artifacts. A local adaptive enhancement function improves the brightness of the reflection map, while the designed loss function addresses illumination smoothness, brightness enhancement, color restoration, and denoising. Experiments validate the effectiveness of our method, revealing improved image brightness, reduced color deviation, and superior color restoration compared to leading approaches.

Keywords: low-light image enhancement; Retinex theory; attention mechanism; unsupervised learning



Citation: Jiang, S.; Shi, Y.; Zhang, Y.; Zhang, Y. An Improved Retinex-Based Approach Based on Attention Mechanisms for Low-Light Image Enhancement. *Electronics* **2024**, *13*, 3645. <https://doi.org/10.3390/electronics13183645>

Academic Editor: Stefano Scanzio

Received: 31 July 2024

Revised: 8 September 2024

Accepted: 10 September 2024

Published: 13 September 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

Low-light conditions often result in poor image quality, hindering accurate threat detection and compromising security measures. Traditional supervised methods require extensive labeled datasets, which are impractical to obtain in diverse and dynamic environments. Unsupervised learning approaches, on the other hand, can leverage abundant low-light images without the need for labeled counterparts, enabling the development of robust enhancement algorithms. These algorithms improve the visibility and detail of images captured under challenging lighting conditions, thus enhancing the reliability and effectiveness of security systems across heterogeneous networks. By ensuring better image quality, unsupervised low-light image enhancement contributes to more accurate monitoring and analysis, ultimately strengthening overall network security. Therefore, improving low-light images is essential for security and surveillance.

Algorithms for enhancing low-light images can be broadly divided into two categories: conventional techniques and deep learning-based methods. Among conventional techniques, histogram equalization and its related methods [1] are the most commonly used. These methods enhance image brightness by calculating the gray levels of the picture and redistributing them evenly across the full range of gray levels. The main advantage of this method is its fast computation speed, which effectively enhances the brightness of general grayscale images. However, it is less effective for RGB image enhancement, often resulting in significant noise, color deviation, and over-enhancement issues.

Researchers have developed various low-light image enhancement techniques [2,3] rooted in Retinex theory [4]. With the success of deep learning in image reconstruction and restoration, neural network-based algorithms like CNN [5] have gained considerable

attention. Among them, supervised learning methods rely on large datasets of paired low-light and well-lit images for training. These methods can achieve impressive results but are limited by the availability of high-quality paired datasets. In contrast, unsupervised learning methods do not require paired datasets. Instead, they use loss functions that incorporate illumination smoothness, brightness consistency, and color restoration to guide the network. The unsupervised approaches are more flexible and practical in scenarios where paired datasets are scarce, enabling effective enhancement by training solely on low-light images. However, unsupervised learning-based algorithms still encounter challenges, including significant color distortion and the blurring of details.

Significant progress has been made in low-light image enhancement in recent years. However, in complex low-light scenarios, image enhancement methods still struggle with preserving natural colors and details while dealing with significant noise. To address these issues, this paper constructs a U-net network based on the Convolutional Block Attention Module (CBAM) to decompose images and enhance the brightness of the reflectance map. The attention mechanism allows the network to concentrate on key features and regions in the image. Retinex theory explains that image brightness is determined by both illumination and object reflection. By separating these components, we can effectively enhance image details. In our approach, we divide the initial image into a low-light RGB image and its corresponding brightness values. Since the illumination conditions in different regions of the image may be different, in order to avoid a one-size-fits-all processing method and enhance local details more finely, we propose a local adaptive enhancement function when enhancing the low-light image. Meanwhile, an unsupervised learning loss function is designed for illumination smoothness, brightness consistency, and color restoration, guiding the network to effectively enhance low-light images. The main contributions of this paper are as follows:

- (1) Combining the U-net network with the CBAM grounded in Retinex theory to achieve the decomposition of the images.
- (2) Establishing a local adaptive enhancement function that calculates the local gray mean of the image through a block operation and adjusts the enhancement effect according to the specific values of each gray block. The parameters within the function allow for the flexible adjustment of the enhancement degree, avoiding over-enhancement.
- (3) Designing an unsupervised learning loss function that introduces a color restoration loss term, further optimizing color restoration, effectively improving image brightness and preserving image details.

The structure of this paper is as follows: Section 2 reviews related research. Section 3 introduces the unsupervised low-light enhancement algorithm developed in this study, detailing the design of the local adaptive function and the loss function. The experimental results are presented in Section 4, and Section 5 offers several conclusions.

2. Related Works

2.1. Unsupervised Low-Light Image Enhancement Algorithms

Many researchers have proposed unsupervised learning-based algorithms to enhance the generalization and robustness of image enhancement models. Jiang et al. [6] proposed an unsupervised decomposition and correction network inspired by the Retinex model. Hu et al. [7] introduced a method that first used the traditional retinol-based method to pre-enhance images, and then used the thinning network for additional quality improvement. Shi et al. [8] proposed a structure-aware unsupervised network comprising four modules. Ma et al. [9] proposed a region-based, unsupervised, low-light image enhancement algorithm that utilizes explicit domain supervision to convert unsupervised segmentation into a supervised process, developing several region-based loss functions to establish semantic consistency between regions and daytime. Guo et al. [10] introduced a new hybrid loss function that combines quality, task, and perception to tackle problems such as blurring and unnatural colors. Wang et al. [11] presented a mixed-attention-guided Generative Adversarial Network (GAN) in a fully unsupervised fashion. Fu et al. [12] introduced an

unsupervised learning network featuring an illumination-aware attention module and a novel identity-invariant loss.

Unsupervised learning methods leverage loss functions to guide the network without the need for paired examples. By focusing on intrinsic properties such as illumination smoothness, brightness consistency, and color restoration, unsupervised approaches offer greater flexibility and robustness.

2.2. Retinex Theory

An image is decomposed into a reflectance component that represents the inherent color of objects and an illumination component representing the varying light conditions. This decomposition allows for the enhancement of images by adjusting the illumination component without destroying the natural colors and details. Assuming the input image is $S(x, y)$, it can be expressed as follows:

$$S(x, y) = I(x, y) \odot R(x, y) \quad (1)$$

where $I(x, y)$ represents the image illumination component, $R(x, y)$ represents the reflectance component of the image, and \odot denotes pixel-wise multiplication.

Retinex theory is commonly applied to enhance visibility and contrast without introducing significant artifacts. Wu et al. [2] proposed a Retinex-based deep unfolding network to obtain noise suppression and detail preservation. Zhao et al. [13] presented a generative strategy for Retinex decomposition, casting the decomposition as a generative problem. Jiang et al. [14] proposed a self-regularized method that preserves all colors while integrating Retinex theory solely for brightness adjustments. Liu et al. [15] constructed a deep learning framework, comprising a decomposition network and adjustment networks that address both global and local brightness. Ma et al. [16] introduced a Retinex-based variational model that effectively produces noise-free images and shows generalizability across various lighting conditions. Yang et al. [17] proposed an image enhancement algorithm that integrates a fast and robust fuzzy C-means clustering technique with Retinex theory, producing enhanced images characterized by their rich detail and texture.

Numerous studies have demonstrated the effectiveness of Retinex in this field. Retinex-based methods can effectively improve visibility and contrast in low-light conditions by accurately decomposing an image into its illumination and reflectance components.

2.3. Attention Mechanisms

To further enhance the feature-learning capability of networks and improve image enhancement results, many scholars have introduced attention mechanisms to boost performance. Chen et al. [18] proposed an attention-based network that incorporates Retinex theory, featuring an attention mechanism module integrated into the convolutional layer. Ai and Kwon [19] developed a convolutional network that integrates an attention gate with a U-net network. Lv et al. [20] separated the tasks of brightness enhancement and noise reduction and completed them separately with two attention maps. Atoum et al. [21] introduced a color-wise attention network that identifies useful color cues to assist with color enhancement. Zhang and Wang [22] proposed an illumination attention map to identify areas of different illumination levels, and a multi-scale attention Retinex network. Zhang et al. [23] developed a neural network that incorporates channel attention and spatial attention modules and achieved a positive effect.

3. Methodology

3.1. Neural Network Structure

This paper constructs a U-net architecture in which the CBAM is introduced [24]. The U-net decomposes low-light images and regards the reflectance map as the enhanced image. In Figure 1, the network receives low-light RGB images and the corresponding brightness values, while the output of the network includes the decomposed reflectance and illumination maps. The encoder of the network gradually reduces the size of the

input tensor and increases the number of layers through convolution and max pooling. The decoder consists of upsampling and convolution layers, which restore the tensor size through convolution and upsampling. During the upsampling process, the CBAM is employed to assign weights to the extracted feature maps to enhance feature refinement and focus during the reconstruction phase, optimizing the network performance. By applying channel attention mechanisms during decoding, we can selectively enhance important features that are identified during encoding, improving detail restoration. Spatial attention enables focus on relevant areas of the image, which is crucial for accurately reconstructing details under low-light conditions. Meanwhile, placing these mechanisms in the decoder reduces the computational load during encoding, allowing for efficient feature extraction.

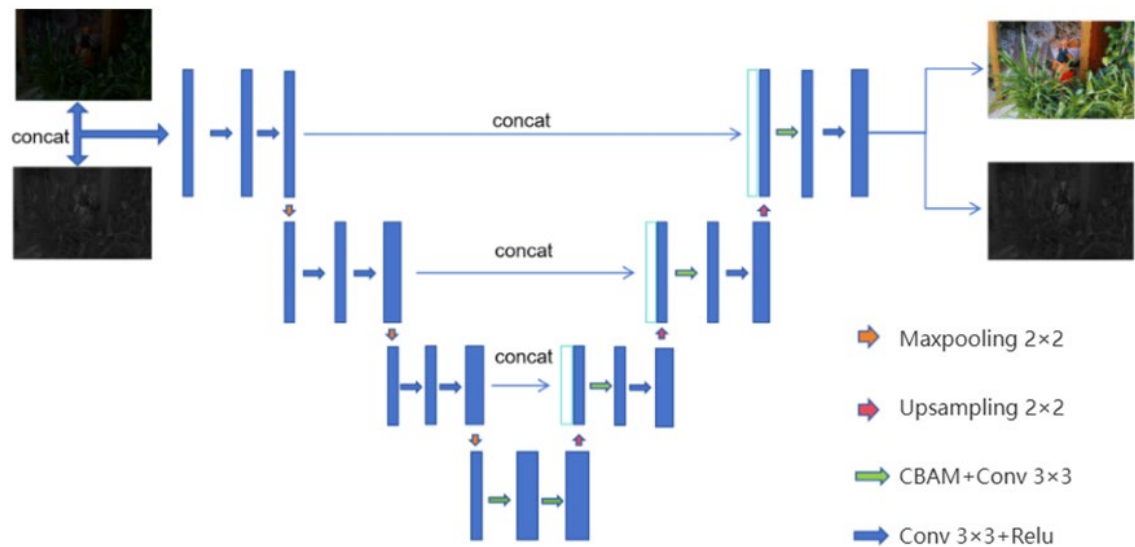


Figure 1. U-net network with attention mechanisms.

The CBAM is a type of Channel and Spatial Mixed Attention (CSMA) mechanism, as shown in Figure 2. The network initially utilizes channel attention, which emphasizes the relationships and dependencies among various feature channels, enabling the network to weigh the importance of each channel differently. The channel attention mechanism can be expressed as follows:

$$M_c(F) = \sigma(MLP(AvgPool(F)) + MLP(MaxPool(F))) \quad (2)$$

where $\sigma(\cdot)$ is the S-shaped function, $MLP(\cdot)$ denotes the multilayer perceptron, $AvgPool(\cdot)$ represents the average pooling, and $MaxPool(\cdot)$ is the maximum pooling. It then applies spatial attention to emphasize the significance of various spatial locations, allowing the network to focus on the relevant regions of the image. The attention module can be written as follows:

$$M_s(F) = \delta(Con\upsilon^{7 \times 7}([AvgPool(F); MaxPool(F)])) \quad (3)$$

where $\delta(\cdot)$ denotes the sigmoid function and $Con\upsilon^{7 \times 7}(\cdot)$ refers to the convolution operation through a convolution kernel, sized 7×7 . By combining these two types of attention, the network can capture complex patterns and dependencies in the data, leading to improved performance in image enhancement where both channel and spatial information are crucial for achieving high-quality results.

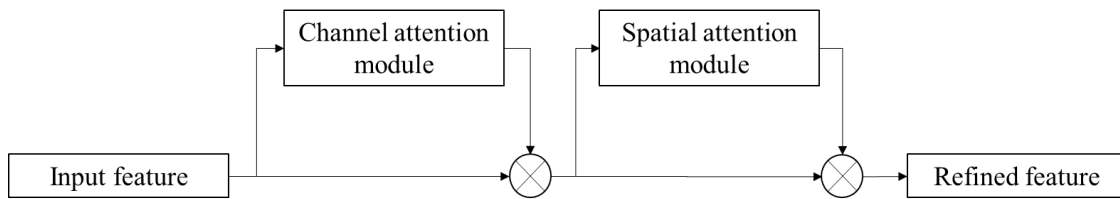


Figure 2. CBAM module.

3.2. Adaptive Enhancement Function

This paper proposes a local adaptive enhancement function that divides a low-light image into multiple blocks and enhances them separately. The mathematical expression is as follows:

$$L_g(x, y) = \min\left(\frac{\alpha + L_{w,max}}{\alpha + \bar{L}_w}, \frac{L_w(x, y)}{L_{w,max}} + \frac{L_w(x, y)}{\bar{L}_w}\right) * \frac{\log(L_w(x, y)/\bar{L}_w + 1)}{\log(L_{w,max}/\bar{L}_w + 1)} \quad (4)$$

where $L_g(x, y)$ represents the corresponding pixel value of the globally adaptive output, $L_w(x, y)$ represents the corresponding input pixel value, $L_{w,max}$ represents the highest brightness value of the input image, α is a constant parameter that regulates the overall enhancement level, and \bar{L}_w represents the logarithmic mean brightness, which can be written as:

$$\bar{L}_w = \exp\left\{\frac{1}{N} \sum \log[\delta + L_w(x, y)]\right\} \quad (5)$$

Equation (4) uses the minimum function to flexibly adjust the enhancement level based on the grayscale values of each block, preventing the excessive enhancement of brightness values.

3.3. Loss Function

By applying the ternary Bayesian theorem, the illumination and reflectance components can be derived from Equation (1):

$$P(R, I|S) \propto P(S|R, I)P(R)P(I) \quad (6)$$

The designed loss function includes an illumination smoothness term, a brightness enhancement term, a color restoration term, and a smoothness term. The reconstruction loss recon can be written as follows:

$$L_{rc} = \left\| S - R \odot I \right\|_1 \quad (7)$$

where S denotes the low-light picture, R represents the reflectance component, I represents the illumination component, and \odot denotes pixel-wise multiplication. The reconstruction loss primarily constrains the decomposition effect of the network, making the decomposed image closer to the real image. The RGB image is transformed into an HSV image and the brightness value V channel can be calculated. And then, the unsupervised learning reflectance loss L_R can be obtained as follows:

$$L_R = \left\| \max_{C \in R, G, B} R_C - F\left(\max_{C \in R, G, B} s_C\right) \right\|_1 \quad (8)$$

where $\max_{C \in R, G, B} R_C$ refers to the maximum value channel of the reflectance component, and $F(\cdot)$ represents the local adaptive enhancement function which can be obtained by Equation (4). The illumination loss L_I is calculated as follows:

$$L_I = \left\| \nabla I \odot \exp(-\lambda \nabla R) \right\|_1 \quad (9)$$

where ∇I represents the gradient variation of the illumination component, controlled by the gradient change of the reflectance, and λ is a constant parameter. The color restoration loss L_C is established as follows:

$$L_C = \sum_{i,j \in (R,G,B)} |R_I - R_J|^2 \quad (10)$$

Equation (10) expresses a relationship between the RGB channels of the reflectance map, further enhancing the color restoration effect. In addition, the noise issue after enhancement is also considered. Thus, the loss function includes the reconstruction loss L_{rc} , the reflectance loss L_R , the color restoration loss L_C , and the denoising loss, expressed as follows:

$$L = \lambda_1 \|S - R \odot I\|_1 + \lambda_2 \left\| \max_{C \in R,G,B} R_C - F \left(\max_{C \in R,G,B} s_C \right) \right\|_1 + \lambda_3 \sum_{i,j \in (R,G,B)} |R_I - R_J|^2 + \lambda_4 \|\nabla I \odot \exp(-\lambda_5 \nabla R)\|_1 + \lambda_6 \|\nabla R\|_1 \quad (11)$$

where ∇R represents the gradient of the reflectance component, mainly used for image denoising. Through the illumination enhancement function, the brightness value of the reflectance map is increased, ultimately achieving low-light image enhancement.

4. Experimental Results and Analysis

4.1. Experimental Setup

The experiments are implemented with an AMD 7945HX processor (Advanced Micro Devices, Santa Clara, CA, USA), an NVIDIA GTX 4060 GPU (Colorful, Shenzhen, China), and 32 GB of memory (Crucial, Meridian, MS, USA). The experimental environment includes Python version 3.9 and TensorFlow version 2.10.0. The training dataset used is LOL, consisting of 500 pairs of low-light and normal-light images captured under various illumination conditions and scenes. In this experiment, LOL contains 485 low/normal-light image pairs for training and 15 pairs for testing. This dataset comprises images from both indoor and outdoor scenes, each with a resolution of 600×400 . We meticulously examine various hyperparameters, including batch size, learning rate, the number of epochs, and weight decay. The search range for each hyperparameter is specified, emphasizing the optimal configuration that yields the best performance. Ultimately, we find that setting the batch size to 8, learning rate to 0.001, number of epochs to 160, and weight decay to 0.0005 result in the best performance of our model on the LOL dataset. The selection of this optimal configuration undergoes thorough experimental validation, ensuring the robustness and effectiveness of our proposed low-light image enhancement method in various aspects.

The average training duration is approximately 60 min, while the processing time for each image in the test set is 0.465 s. Through many experiments, the values of the six parameters $\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5, \lambda_6$ in Equation (11) are set to 20, 1, 5, 0.1, 10, and 0.01, respectively. This parameter combination has been validated to achieve the best results on the test set. In this combination, $\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5, \lambda_6$ control the weights of reconstruction loss, reflection loss, color recovery loss, illumination loss, local reflection component gradient, and global reflection component gradient, respectively. Fine-tuning in different datasets and application scenarios according to actual needs is recommended.

4.2. Experimental Results

To assess the effectiveness of the proposed algorithm, several comparative experiments are carried out with methods like Retines [25], LIME [26], SCI [27], Zero-DCE [28], EnlightenGAN [29], and GLADNet [30]. We also choose the methods of class unsupervised learning, based on deep learning techniques such as RUAS [31] and SSIE [32].

4.2.1. Subjective Evaluation Results

The enhanced results are shown in Figure 3. The SCI algorithm results in less noise and better color restoration, but its brightness enhancement is still somewhat lacking compared to normal light images. Compared to the enhancement results of the presented algorithm, the brightness of the SCI method is relatively dim. The Zero-DCE and LIME algorithms exhibit less color distortion, but their brightness enhancement is inadequate, leaving the dark areas of the image insufficiently illuminated. The Retinex-Net algorithm achieves positive brightness enhancement, but it suffers from a loss of image detail and texture, and the variation in brightness across different areas is not natural enough. The EnlightenGAN algorithm provides positive brightness enhancement, but the dark areas of the image show noticeable noise, and there is some color distortion present. The RUAS algorithm provides limited enhancement in extremely dark regions, while the brighter local areas are excessively enhanced, leading to overexposure. The SSIE algorithm clearly improves brightness, but the background color of the enhanced image shows a significant deviation.

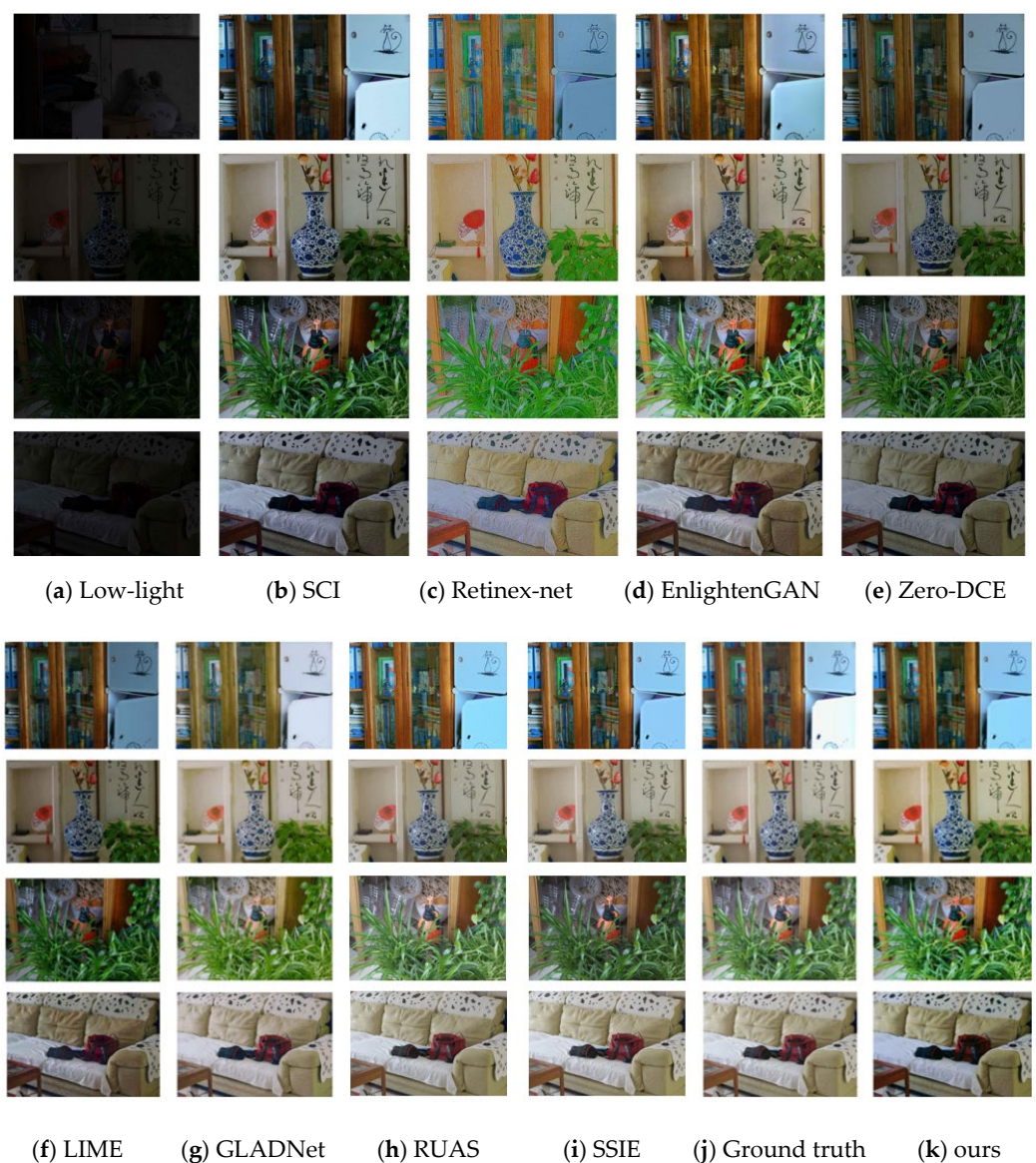


Figure 3. Comparison of different methods.

Figure 4 illustrates the enhancement effects of various algorithms when applied to a single image. Compared with SCI, the proposed model performs better in enhancing

light-colored areas. The EnlightenGAN enhancement algorithm faces certain difficulties in distinguishing between the black areas of the image, failing to differentiate between inherent color and low-light conditions. The enhanced results of Retinex-net indicate significant color deviation. The GLADNet-enhanced image has less noise but exhibits low color saturation. The enhanced results of RUAS algorithm are not obvious in the dark. The SSIE algorithm faces image color distortion. It is noted that the proposed approach results in less color distortion compared to other methods.

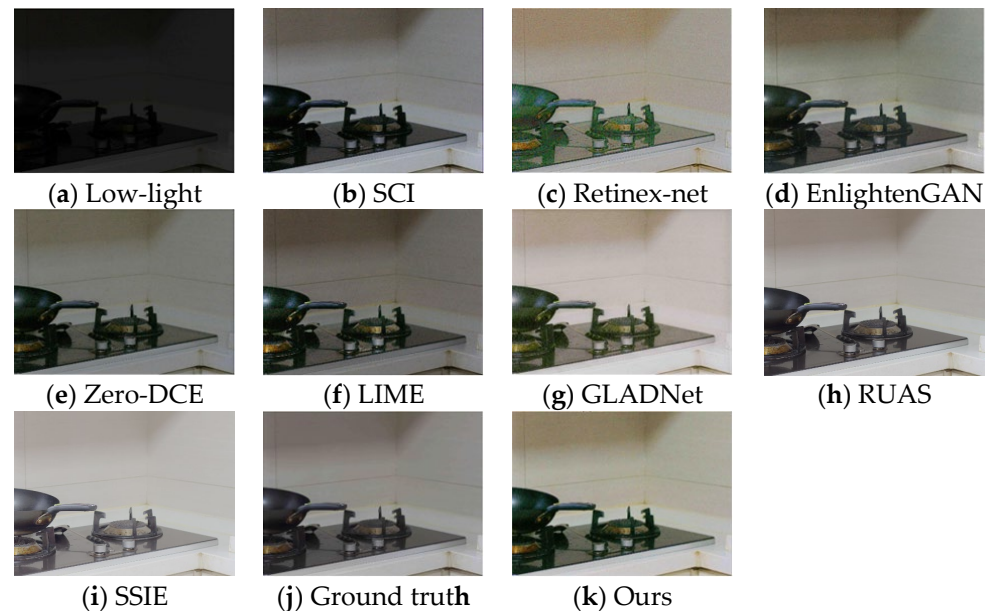


Figure 4. Comparison of details of different methods.

4.2.2. Objective Evaluation Results

The objective evaluation methods used include PSNR, SSIM [33], and NIQE [34]. In Table 1, an upward arrow (\uparrow) indicates that the metric is positively correlated with image quality, while a downward arrow (\downarrow) indicates a negative correlation.

Table 1. Objective evaluation indicators of different methods on the LOL dataset.

| | SCI | EnlightenGAN | LIME | Zero-DCE | Retinex-Net | GLADNet | RUAS | SSIE | Ours |
|-------------------|--------|--------------|--------|----------|-------------|---------|--------|--------|--------|
| SSIM \uparrow | 0.635 | 0.752 | 0.590 | 0.664 | 0.502 | 0.778 | 0.441 | 0.723 | 0.826 |
| PSNR \uparrow | 17.210 | 18.849 | 13.244 | 15.215 | 17.839 | 19.705 | 10.714 | 16.800 | 20.200 |
| NIQE \downarrow | 8.878 | 7.174 | 8.640 | 8.497 | 11.250 | 7.084 | 7.833 | 3.882 | 5.05 |

The SCI, EnlightenGAN, and Zero-DCE algorithms exhibit poor NIQE and PSNR metrics due to insufficient illumination enhancement and noise in dark areas. The GALDNet algorithm performs well across these three metrics due to its effective brightness enhancement and reduced image noise. However, subjective observation reveals significant color distortion in the results of this algorithm. LIME and Retinex-Net show poor performance in objective evaluation metrics. Their enhancement results contain substantial noise, and the Retinex-Net algorithm suffers from severe detail loss. The RUAS algorithm underperforms when using the proposed approach on all three indicators, and its dark region enhancement is insufficient. The SSIE algorithm performs best on the NIQE indicator, but not as well as the proposed approach on SSIM and PSNR, and it also has color distortion problems in terms of subjective performance.

4.3. Ablation Experiment

To demonstrate the effectiveness of the CBAM attention mechanism and the color restoration term in the loss function of the proposed algorithm, ablation studies are conducted on each module.

4.3.1. CBAM Attention Mechanism

This section quantitatively assesses the impact of the CBAM module. Training and testing are performed on the same dataset, and various evaluation metrics are compared, as shown in Table 2. The addition of the attention mechanism module lowers the NIQE index, while both SSIM and PSNR show improvements, indicating that the module improves brightness. Notably, the performance is best when two attention mechanism modules are added, with the approach of applying channel attention first, followed by spatial attention (CBAM), yielding the most effective results.

Table 2. Index data with attention mechanisms.

| Model | Attention Mechanism | SSIM | PSNR | NIQE |
|----------|--------------------------------------|-------|--------|-------|
| 1 | none | 0.782 | 20.144 | 7.030 |
| 2 | channel attention | 0.797 | 20.169 | 5.924 |
| 3 | spatial attention | 0.804 | 20.166 | 5.600 |
| 4 | spatial and channel attention | 0.822 | 20.188 | 5.105 |
| 5 (Ours) | channel and spatial attention (CBAM) | 0.826 | 20.200 | 5.057 |

4.3.2. Color Restoration Term

To validate the contribution of the color restoration term in boosting model performance, several ablation experiments are conducted, as shown in Table 3. In these experiments, the color restoration term is removed to observe changes in performance, providing a clearer understanding of its significance in enhancing overall model effectiveness. Initially, we modify a model that already uses our optimized loss function by removing the color restoration term. We then retrain the modified model using the same datasets and training parameters, documenting its performance. It is observed that the configuration using CBAM and our optimized loss function produces the highest performance. Even when CBAM is not used, the configuration using the loss function with the added color restoration loss item performs better than the configuration using the regular loss function. The color restoration loss term reduces the NIQE metric, making the enhanced images appear more natural by reducing color deviation.

Table 3. Index data with color restoration terms.

| Model | CBAM | Color Restoration Term | SSIM | PSNR | NIQE |
|----------|------|------------------------|-------|--------|-------|
| 1 | × | × | 0.793 | 19.822 | 6.142 |
| 2 | ✓ | × | 0.819 | 20.050 | 5.520 |
| 3 | × | ✓ | 0.782 | 20.144 | 7.030 |
| 4 (Ours) | ✓ | ✓ | 0.826 | 20.200 | 5.057 |

5. Conclusions

In this paper, a U-net network based on attention mechanisms is built, decomposing the illumination and reflectance components of low-light images based on Retinex theory. The local, adaptive enhancement function adjusts the enhancement level for different regions of the image, ensuring that both dark and bright areas are optimally enhanced for improved overall image quality. In addition, an unsupervised learning loss function is introduced. The proposed algorithm exhibits minimal dependence on the quality of the training dataset, low computational complexity, and rapid training speed. Experimental results show that the algorithm produces enhanced outcomes that closely mimic normal

lighting conditions, featuring more natural lighting enhancement, improved color restoration, minimal color deviation, and preserved image details. It also shows good model generalization and robustness, with the peak signal-to-noise ratio and structural similarity index on real images outperforming other state-of-the-art methods. However, there are still some shortcomings in this study. The proposed model is suitable for unsupervised low-light image enhancement, and can only process still images, not videos. The local adaptive enhancement function adopts the equal division strategy when it divides the image into blocks as well. In the future, we plan to extend the proposed model to handle video data, enabling real-time, low-light video enhancement. Additionally, we will explore more sophisticated image division strategies, such as adaptive or content-aware partitioning, to further improve the local adaptive enhancement function for better performance and finer detail preservation.

Author Contributions: S.J.: conceptualization, methodology, supervision, writing—original draft preparation, and funding acquisition; Y.S.: methodology and writing—review and editing; Y.Z. (Yingchun Zhang): methodology and validation; Y.Z. (Yulin Zhang): conceptualization, data curation, formal analysis, and investigation. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the National Natural Science Foundation of China (52105167).

Data Availability Statement: The data presented in this study are available in Low Light paired dataset (LOL) at <https://daooshee.github.io/BMVC2018website/> (accessed on 14 May 2024). These data were derived from the following resources available in the public domain: Wei, C., Wang, W., Yang, W. and Liu, J. (2018). Deep retinex decomposition for low-light enhancement. arXiv preprint arXiv:1808.04560 [27].

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

References

1. Zhang, F.; Shao, Y.; Sun, Y.; Gao, C.; Sang, N. Self-supervised Low-Light Image Enhancement via Histogram Equalization Prior. In Proceedings of the Chinese Conference on Pattern Recognition and Computer Vision (PRCV), Xiamen, China, 13–15 October 2023; Springer Nature: Singapore, 2023; pp. 63–75.
2. Wu, W.; Weng, J.; Zhang, P.; Wang, X.; Yang, W.; Jiang, J. Uretinex-net: Retinex-based deep unfolding network for low-light image enhancement. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans, LA, USA, 18–24 June 2022; pp. 5901–5910.
3. Yi, X.; Xu, H.; Zhang, H.; Tang, L.; Ma, J. Diff-retinex: Rethinking low-light image enhancement with a generative diffusion model. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Paris, France, 1–6 October 2023; pp. 12302–12311.
4. Land, E.H. The retinex. In *Ciba Foundation Symposium—Colour Vision: Physiology and Experimental Psychology*; John Wiley & Sons, Ltd.: Chichester, UK, 1965; pp. 217–227.
5. Tao, L.; Zhu, C.; Xiang, G.; Li, Y.; Jia, H.; Xie, X. LLCNN: A convolutional neural network for low-light image enhancement. In Proceedings of the 2017 IEEE Visual Communications and Image Processing (VCIP), St. Petersburg, FL, USA, 10–13 December 2017; IEEE: Piscataway, NJ, USA, 2017; pp. 1–4.
6. Jiang, Q.; Mao, Y.; Cong, R.; Ren, W.; Huang, C.; Shao, F. Unsupervised decomposition and correction network for low-light image enhancement. *IEEE Trans. Intell. Transp. Syst.* **2022**, *23*, 19440–19455. [[CrossRef](#)]
7. Hu, J.; Guo, X.; Chen, J.; Liang, G.; Deng, F.; Lam, T.L. A two-stage unsupervised approach for low light image enhancement. *IEEE Robot. Autom. Lett.* **2021**, *6*, 8363–8370. [[CrossRef](#)]
8. Shi, Y.; Wang, B.; Wu, X.; Zhu, M. Unsupervised low-light image enhancement by extracting structural similarity and color consistency. *IEEE Signal Process. Lett.* **2022**, *29*, 997–1001. [[CrossRef](#)]
9. Ma, Y.; Xie, S.; Xu, W.; Chen, X.; Huang, X.; Sun, Y.; Liu, W. Region-Based Unsupervised Low-Light Image Enhancement in the Wild with Explicit Domain Supervision. *IEEE Trans. Instrum. Meas.* **2024**, *73*, 5024511. [[CrossRef](#)]
10. Guo, H.; Xu, W.; Qiu, S. Unsupervised low-light image enhancement with quality-task-perception loss. In Proceedings of the 2021 International Joint Conference on Neural Networks (IJCNN), Shenzhen, China, 18–22 July 2021; IEEE: Piscataway, NJ, USA, 2021; pp. 1–8.
11. Wang, R.; Jiang, B.; Yang, C.; Li, Q.; Zhang, B. MAGAN: Unsupervised low-light image enhancement guided by mixed-attention. *Big Data Min. Anal.* **2022**, *5*, 110–119. [[CrossRef](#)]
12. 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](#)]

13. 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](#)]
14. Jiang, Z.; Li, H.; Liu, L.; Men, A.; Wang, H. A switched view of Retinex: Deep self-regularized low-light image enhancement. *Neurocomputing* **2021**, *454*, 361–372. [[CrossRef](#)]
15. Liu, X.; Xie, Q.; Zhao, Q.; Wang, H.; Meng, D. Low-light image enhancement by retinex-based algorithm unrolling and adjustment. *IEEE Trans. Neural Netw. Learn. Syst.* **2023**, 1–14. [[CrossRef](#)] [[PubMed](#)]
16. Ma, Q.; Wang, Y.; Zeng, T. Retinex-based variational framework for low-light image enhancement and denoising. *IEEE Trans. Multimed.* **2022**, *25*, 5580–5588. [[CrossRef](#)]
17. Yang, J.; Wang, J.; Dong, L.; Chen, S.; Wu, H.; Zhong, Y. Optimization algorithm for low-light image enhancement based on Retinex theory. *IET Image Process.* **2023**, *17*, 505–517. [[CrossRef](#)]
18. Chen, X.; Li, J.; Hua, Z. Retinex low-light image enhancement network based on attention mechanism. *Multimed. Tools Appl.* **2023**, *82*, 4235–4255. [[CrossRef](#)]
19. Ai, S.; Kwon, J. Extreme low-light image enhancement for surveillance cameras using attention U-Net. *Sensors* **2020**, *20*, 495. [[CrossRef](#)] [[PubMed](#)]
20. Lv, F.; Li, Y.; Lu, F. Attention guided low-light image enhancement with a large scale low-light simulation dataset. *Int. J. Comput. Vis.* **2021**, *129*, 2175–2193. [[CrossRef](#)]
21. Atoum, Y.; Ye, M.; Ren, L.; Tai, Y.; Liu, X. Color-wise attention network for low-light image enhancement. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, Seattle, WA, USA, 14–19 June 2020; pp. 506–507.
22. Zhang, X.; Wang, X. Marn: Multi-scale attention retinex network for low-light image enhancement. *IEEE Access* **2021**, *9*, 50939–50948. [[CrossRef](#)]
23. Zhang, C.; Yan, Q.; Zhu, Y.; Li, X.; Sun, J.; Zhang, Y. Attention-based network for low-light image enhancement. In Proceedings of the 2020 IEEE International Conference on Multimedia and Expo (ICME), London, UK, 6–10 July 2020; IEEE: Piscataway, NJ, USA, 2020; pp. 1–6.
24. Woo, S.; Park, J.; Lee, J.Y.; Kweon, I.S. Cbam: Convolutional block attention module. In Proceedings of the European Conference on Computer Vision (ECCV), Munich, Germany, 8–14 September 2018; pp. 3–19.
25. Wei, C.; Wang, W.; Yang, W.; Liu, J. Deep retinex decomposition for low-light enhancement. *arXiv* **2018**, arXiv:1808.04560.
26. 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](#)]
27. Ma, L.; Ma, T.; Liu, R.; Fan, X.; Luo, Z. Toward fast, flexible, and robust low-light image enhancement. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans, LA, USA, 18–24 June 2022; pp. 5637–5646.
28. 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, 13–19 June 2020; pp. 1780–1789.
29. 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](#)]
30. 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; IEEE: Piscataway, NJ, USA, 2018; pp. 751–755.
31. Liu, R.; Ma, L.; Zhang, J.; Fan, X.; Luo, Z. Retinex-Inspired Unrolling With Cooperative Prior Architecture Search for Low-Light Image Enhancement. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Nashville, TN, USA, 20–25 June 2021; IEEE: Piscataway, NJ, USA, 2021; pp. 10561–10570.
32. Zhang, Y.; Di, X.; Zhang, B.; Wang, C. Self-supervised Image Enhancement Network: Training with Low Light Images Only. *arXiv* **2020**, arXiv:2002.11300.
33. Hore, A.; Ziou, D. Image quality metrics: PSNR vs. SSIM. In Proceedings of the 2010 20th International Conference on Pattern Recognition, Istanbul, Turkey, 23–26 August 2010; IEEE: Piscataway, NJ, USA, 2010; pp. 2366–2369.
34. Mittal, A.; Soundararajan, R.; Bovik, A.C. Making a “completely blind” image quality analyzer. *IEEE Signal Process. Lett.* **2012**, *20*, 209–212. [[CrossRef](#)]

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.