*International Journal of Innovative Computing, Information and Control*
Volume **20**, Number **2**, April **2024**

ICIC International ©2024 ISSN 1349-4198
pp. **479–496**

# LOW-LIGHT IMAGE ENHANCEMENT USING ASYMMETRIC CONVOLUTIONAL NEURAL NETWORKS

Jiajia Liu* and Zhixiang Deng

Institute of Electronic and Electrical Engineering
Civil Aviation Flight University of China
No. 46, 4th Section, Nanchang Road, Guanghan 618307, P. R. China
dzxcafuc@163.com
*Corresponding author: cafucljj@cafuc.edu.cn

Received June 2023; revised October 2023

Abstract. *In response to the challenges posed by high noise, low brightness, and inadequate detail information in low-light images, this paper presents a novel approach for enhancing such images using a convolutional neural network (CNN) with asymmetric convolution kernels. Firstly, the low-light image is decomposed by a decomposition network, resulting in the extraction of the reflection component and the brightness component, which are then fed into the enhancement network. The enhancement network includes a denoising component and an illumination compensation component. In the denoising component, a novel asymmetric convolution kernel is introduced to replace the conventional square convolution kernel typically used for denoising the low-light reflection component. This innovative kernel design reduces the parameter operation load while yielding superior denoising performance. By employing this approach, the method achieves notable improvements in denoising effectiveness. In the illumination compensation component, a Convolutional Block Attention Module (CBAM) with an attention mechanism is introduced to suppress redundant information in the background of the low-light brightness component, while simultaneously enhancing high-weighted brightness and detail information. Finally, the denoised low-light reflection component and the enhanced low-light brightness component are fused to reconstruct the final enhanced image. Experimental evaluations validate the effectiveness of the proposed method. The results demonstrate improvements in evaluation parameters such as Peak Signal-to-Noise Ratio (PSNR), Structural Similarity Index (SSIM), and Mean Squared Error (MSE). These improvements substantiate the ability of the proposed method to enhance low-light images, offering superior image quality while preserving essential details and minimizing noise artifacts.*

**Keywords:** Image enhancement, Convolutional neural network, Asymmetric convolution kernel, Attention mechanism

1. **Introduction.** Digital images play a crucial role in everyday life as they serve as important carriers for acquiring, transmitting, and receiving information. A significant number of digital images suffer from blurred details in shadow regions due to various factors such as nighttime shooting, backlighting during dusk, and occlusions caused by obstacles. This not only diminishes the visual quality of the images but also affects subsequent computer vision tasks such as object detection and 3D modeling. Hence, it is of great importance to undertake research focused on enhancing images captured under low-light conditions [1].

Low-light image enhancement methods can be mainly categorized into two types: traditional algorithms and deep learning methods. However, the algorithms built based on

DOI: 10.24507/ijicic.20.02.479

traditional methods often process different channels of the image separately, which can lead to the presence of noise. Since the emergence of convolutional neural networks, low-light image enhancement algorithms based on neural networks have greatly improved color distortion and random noise in image processing [2-6]. However, they still face challenges in dealing with residual noise and preserving details adequately. To address these challenges, this paper presents a Asymmetric Convolution and Convolutional Block Attention Module based on Retinex Network, which is shorted as ACARNet to address the issues of noise reduction and image detail enhancement, ultimately leading to significant improvements in overall image quality. Firstly, by employing asymmetric convolutional kernels, the feature extraction capability of the low-light reflection component is enhanced, reducing the influence of image noise. Simultaneously, a color evaluation is conducted between the enhanced image result and the target image to avoid color distortion in the low-light reflection component during the denoising process, preserving color information. Moreover, this study incorporates a Convolutional Block Attention Module (CBAM) attention mechanism for enhancing brightness in low-light images by extracting luminance features more effectively. The CBAM attention mechanism plays a crucial role in improving the brightness of the low-light brightness component, thereby enhancing image details. By skillfully fusing and reconstructing the processed low-light image reflection component and brightness component, it becomes possible to significantly reduce noise, enhance image details, and achieve remarkable brightness enhancement.

1) An asymmetric convolution kernel is proposed to replace the traditional square convolution kernel for denoising the low-light reflection component.

2) The CBAM is introduced for brightness enhancement in low-light images to further extract luminance features.

3) We propose a new loss function in the domain which corrects in the low-light underexposed images effificiently and can be plugged into any deep framework easily without additional cost.

4) We verifified the improvement by conducting subjective experiments and observed that the improvement in scores is consistent with human judgement.

This article is organized in the following order. In Section 2, we review the classical approaches and deep learning-based approaches for low-light image enhancement tasks. In Section 3, we introduce the proposed frameworks and describe the loss function used for low-light image enhancement. In Section 4, we provide the details of our experiments by providing information about the datasets we have used in training and testing our methods. We present comparative performance evaluation results derived from the experiments on different datasets and also a comparison with other studies that used similar datasets. Lastly, the concluding statements are presented in Section 5.

## 2. **Related Work.**

2.1. **Classical approaches.** In traditional low-light image enhancement methods, the common approach is to process the luminance channel of the image in different color spaces. Jobson et al. [7] proposed Single-Scale Retinex (SSR), which estimates the illumination component using a low-pass filter. This method has low computational complexity and is relatively simple to implement. However, it tends to lose image details and introduce color distortion, making it difficult to accurately determine the scale. Jobson et al. [8] also introduced the Multi-Scale Retinex (MSR) to address the scale problem of SSR, but it does not effectively solve the issue of color distortion. Rahman et al. [9] proposed Color Restoration Retinex (MSRCR), which better restores image colors, enhances the image while suppressing noise amplification, and improves color fidelity. Guo et al. [10] focused

on estimating the structured illumination map from an initial one (LIME), which only processes the brightness component (LIME). It constructs a luminance map by finding the maximum value among the R, G, and B channels and then applies an algorithm based on augmented Lagrange multipliers to refining the luminance component and obtaining an enhanced image. LIME achieves an improvement in visual quality but is prone to color distortion. Traditional low-light image enhancement methods based on the luminance channel processing often encounter the following two problems: 1) ineffective extraction of illumination components from images with illumination variations, leading to loss of shadow details from low to high illuminations; 2) inaccurate estimation of illumination values, resulting in the presence of noise in the illumination compensation under low-light conditions.

2.2. **Deep learning-based approaches.** Fortunately, in recent times, learning-based methods utilizing Convolutional Neural Networks (CNNs) have achieved remarkable success in the field of low-light image enhancement [11]. In addition to [2-6], we have listed more algorithms to demonstrate the advantages of deep learning-based approaches. These methods have exhibited exceptionally high performance and have surpassed conventional prior-based enhancement approaches [12-17]. To more effectively address the aforementioned problems, researchers have recently attempted to combine the Retinex theory with deep learning methods for enhancing low-light images. Shen et al. [18] proposed the low-light net (MSR-Net), which is designed to simulate the traditional MSR algorithm. It primarily learns the mapping relationship between the dark channel and the bright channel and replaces the Gaussian surround function with convolutional layers. The low-light/normal-light images are used as the input and output of the network, respectively. This method avoids color distortion and excessive enhancement of brightness but does not yield significant overall brightness improvement and fails to preserve image details adequately. Wei et al. [19] introduced the Deep Retinex-Net (Retinex-Net), which consists of the Decom-Net for decomposition and the Enhance-Net for enhancement. This model is trained in a data-driven manner. Retinex-Net effectively improves image brightness but suffers from blurring of the reflection component at image edges due to the denoising operation using the denoising tool (BM3D [20]) module before reconstruction and fusion. Zhang et al. [21] constructed a convolutional neural network (KinD-Net), which performs denoising on the reflection component of low-light images and enhances the luminance component, resulting in excessive image sharpening. Jiang et al. [22] presented the low-light net (EnlightGan) for enhancing low-light images, using self-extracted information from input images to normalize training on unpaired images. However, this method tends to produce regional color artifacts, compromising visual quality. While the aforementioned methods have shown improvements in brightness, they still suffer from the presence of residual noise and inadequate preservation of details.

Our work is motivated from and close to the work done by [21]. In this work, the authors created the KinD-Net and decomposed a Layer Decomposition and Reflectance Restoration architecture for low-image enhancement. However, we contribute a new framework-independent loss function which works for images and performs signifificantly better than their solution.

## 3. **Proposed Approach.**

3.1. **Retinex theory.** The fundamental idea of the Retinex theory is to eliminate the influence of the incident component on the reflectance component, thereby achieving image enhancement [23]. According to the Retinex theory, the observed image by the human eye

can be divided into the intrinsic reflectance component of the object and the brightness component. This can be represented by Formula (1):

$$L_{(x,y)} = R_{(x,y)} \cdot I_{(x,y)} \tag{1}$$

where $L$ represents the imaging of the human eye, $R$ represents the reflectance component, $I$ represents the brightness component, and $(x, y)$ represents the pixel at any given point.

3.2. **Framework.** This paper is based on the consistency theory in Retinex theory, where the reflectance component is considered as the intrinsic attribute of an image. The proposed neural network, ACARNet combines asymmetric convolutional kernels with attention mechanisms. The model structure is illustrated in Figure 1 and consists of two main parts: the Decom-net for decomposition and the Enhance-net for enhancement.

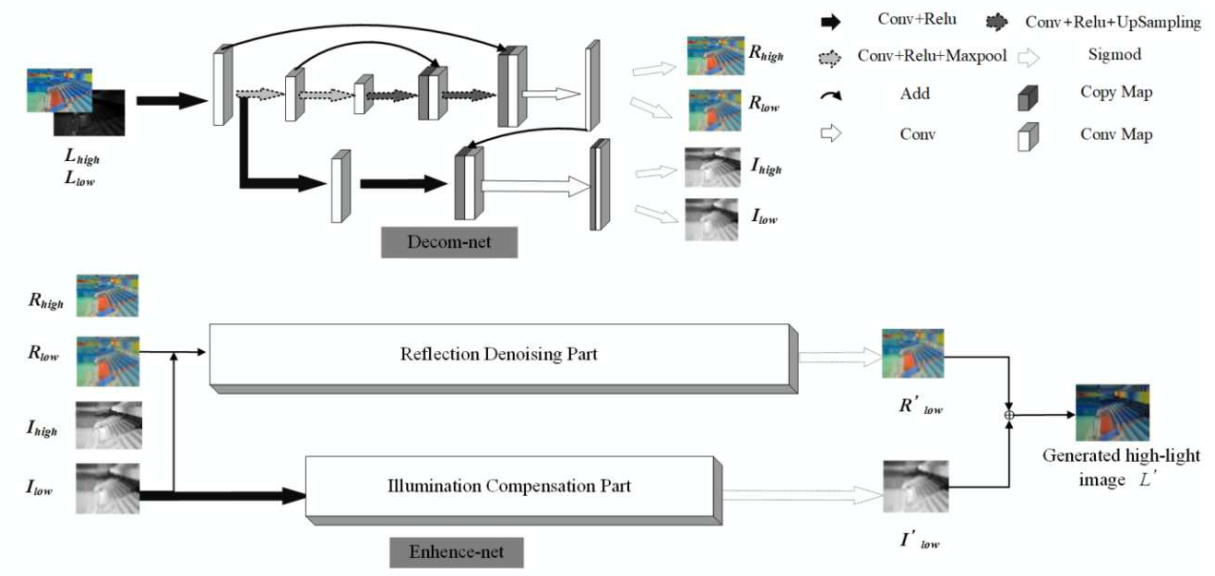



FIGURE 1. The proposed ACARNet architecture

The Decom-net handles paired normal and low-illumination images and decomposes them into the reflectance component $R_{high}$, $R_{low}$ and the brightness component $I_{high}$, $I_{low}$. Firstly, the paired images undergo $3 \times 3$ convolution operations and ReLU activation function. The convolution extracts image features, while the activation function adjusts and updates the linear relationship of the model's parameter weights, mapping the features to the reflectance component $R_{high}$, $R_{low}$ and the brightness component $I_{high}$, $I_{low}$, allowing the neural network to learn optimal results. The presence of max pooling reduces the dimensionality of the image features and further extracts features. Then, the image dimensions are enlarged through upsampling, followed by a convolution operation with a replicated image of the same dimension, resulting in a new feature map with a dimension of 3. The triple concatenation helps to preserve image details in the brightness component. Another convolution is performed to generate a new feature map. Finally, the Sigmoid activation function is used to constrain $R_{high}$, $R_{low}$ and $I_{high}$, $I_{low}$ within the range of $(0, 1)$.

The Enhance-net consists of a denoising component and an illumination compensation component. Firstly, the low-illumination reflectance component $R_{low}$ is denoised, and the illumination information of the low-illumination brightness component $I_{low}$ is enhanced. Then, the final enhanced result image is obtained by fusing them with certain weights.

1) Denoising component

Due to the interference of factors such as illumination and exposure during the acquisition of low-illumination images, a significant amount of noise is often present, which can deteriorate the visual quality and blur the image details. Therefore, it is necessary to perform denoising on $R_{low}$ to preserve the fine details of the image. In this paper, an improved network structure called New RED-Net (NRED-Net), based on the RED-Net [24], is utilized for denoising $R_{low}$. The RED-Net network architecture consists of convolutional layers and deconvolutional layers. RED-Net exhibits remarkable Gaussian denoising capabilities. However, although symmetric network structures have certain advantages, they are relatively complex with a large number of parameters. They also have redundant training data, leading to increased training difficulty and longer training time. The use of RED-Net can result in performance degradation and extended training time. To address these issues, this paper proposes a method to reduce the computational complexity and improve algorithm performance by replacing the core convolutional layer's kernel. By fusing and replacing the original standard square $3 \times 3$ kernel with a combination of a $3 \times 1$ kernel and a $1 \times 3$ kernel, better performance and enhanced representational capability can be achieved through training [25]. To preserve image information, only the convolutional layer's kernel in RED-Net is replaced, while the deconvolutional layer remains unchanged. The advantage of this approach is that it significantly improves algorithm performance and reduces training complexity. This kernel replacement method can effectively enhance the reconstruction quality of low-illumination images and preserve more detailed information, achieving desirable results. Figure 2 illustrates the architecture of the denoising subnetwork, including the kernel replacement structure.

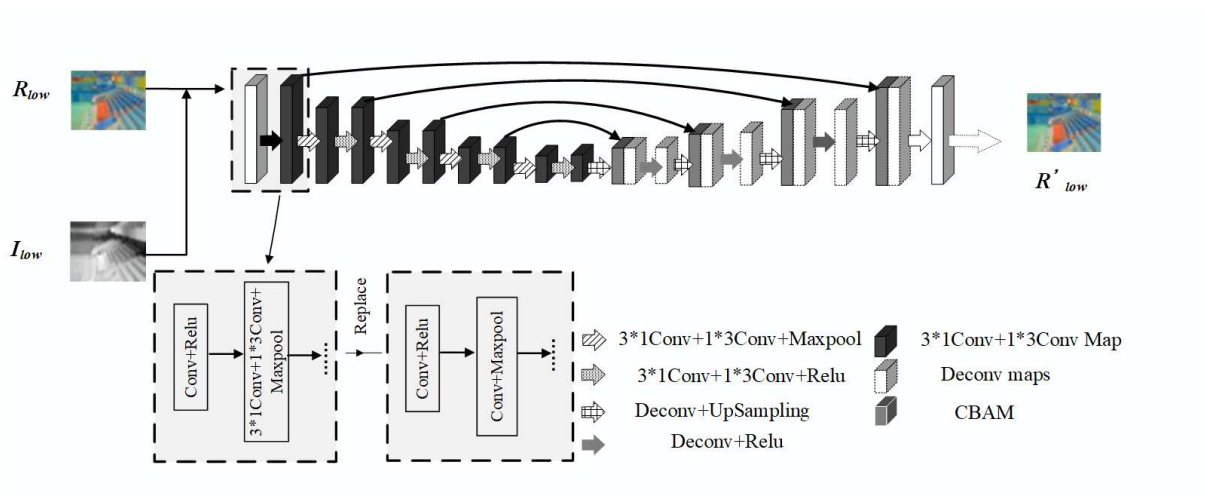



FIGURE 2. The architecture of the reflection denoising component

After the replacement, the number of parameters is reduced, and the modified model does not require additional computational cost. However, the model can be trained to achieve higher performance. Additionally, NRED-Net performs corresponding max pooling operations and upsampling, followed by convolution to generate a feature map with a dimension of 3. Finally, the Sigmoid activation function is applied to activating the denoising part of the low-illumination reflection component image.

2) Illumination compensation component

To address the issues of information blur and difficult-to-discern details in the brightness component $I_{low}$ of the Decom-net for low-illumination images, this paper introduces

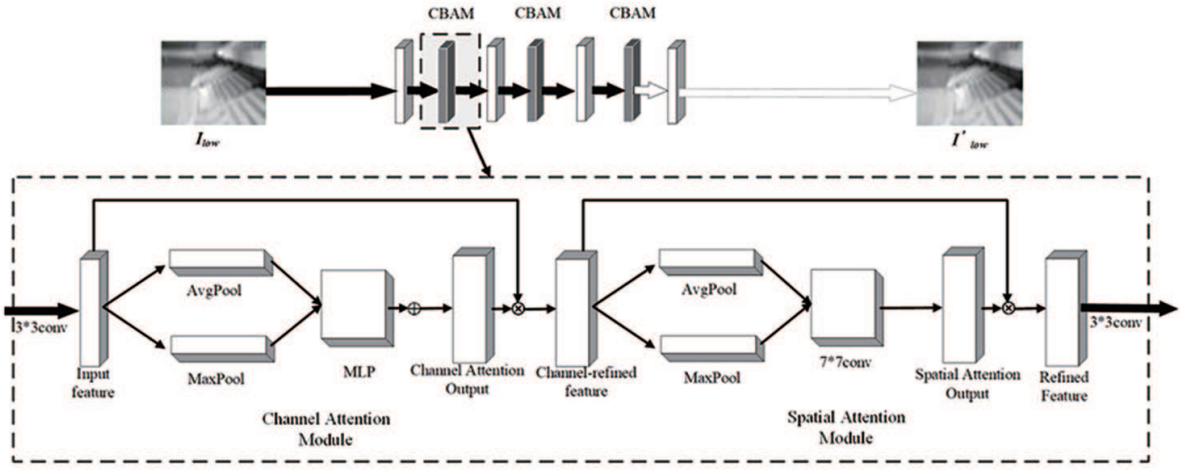



FIGURE 3. The architecture of the illumination compensation component

the Convolutional Block Attention Module (CBAM) [26] attention mechanism. The utilization of CBAM can effectively enhance the processing capability for low-illumination images and improve the clarity and recognition accuracy of the brightness component $I_{low}$ in the Decom-net. CBAM is a lightweight feed-forward convolutional neural network that is highly suitable for computer vision tasks. The architecture of the light compensation sub-network, including the CBAM structure, is depicted in Figure 3.

CBAM, as a lightweight and versatile module, can be integrated into any CNNs framework. As shown in Figure 3, the brightness component $I_{low}$ of the low-illumination image undergoes convolution and max pooling operations before being fed into the CBAM module. There are three CBAM modules in total. After passing through a regular convolution layer, a 3-dimensional feature map is generated, which is then constrained and activated by the Sigmoid activation function. CBAM redistributes the weights of the input feature map to create a new feature map. Channel attention enhances the weights of effective channels, focusing on areas with higher weights. Spatial attention exploits the non-local correlation of the feature map to capture the distribution of noise in different regions, enabling the mapping of the brightness component to remove corresponding noise. Global max pooling and global average pooling operations are used to extract rich high-level features.

The Enhance-net generates the denoised low-illumination reflection component $R'_{low}$ and the light-compensated low-illumination brightness component $I'_{low}$. These components are multiplied element-wise to produce an image with normal brightness.

3.3. **Loss function.** According to the Retinex theory, the reflection component is an inherent property of objects. The loss function $L_{Decom}$ of the decomposition network consists of three parts. The first part, $L^1_{Decom}$, includes the reconstruction loss $L_{re}$ and the reflection component loss $L_r$, aiming to minimize the discrepancy between the image and the product of the decomposed reflection component and the brightness component. It also encourages the low-light image's reflection component to be close to that of the normal-light image. The mathematical expressions are as follows, given by Equations (2), (3), and (4):

$$L^1_{Decom} = L_{re} + \lambda_r L_r \tag{2}$$

$$L_{re} = \|L_{low} - R_{low} \cdot I_{low}\|_1 + \|L_{high} - R_{high} \cdot I_{high}\|_1 \tag{3}$$

$$L_r = \|R_{low} - R_{high}\|_2^2 \tag{4}$$

In the equation, $R_{low}$ and $I_{low}$ represent the low-light decomposed reflection component and brightness component generated by the decomposition network. $R_{high}$ and $I_{high}$ represent the normal-light decomposed reflection component and brightness component generated by the decomposition network, and their product represents the directly fused mapping image. $\lambda_r$ denotes the weight parameter, $\| \|_1$ represents the L1 norm operation, and $\| \|_2$ represents the L2 norm operation.

The second part of the loss function, $L^2_{Decom}$, incorporates the smoothness consistency loss from the KinD-Net, which exhibits strong gradient preservation capabilities. It is formulated as Equation (5):

$$L^2_{Decom} = \lambda_i \left[ \left\| \frac{\nabla I_{low}}{\max\left(|\nabla L_{low}|, \varepsilon_1\right)} \right\|_1 + \left\| \frac{\nabla I_{high}}{\max\left(|\nabla L_{high}|, \varepsilon_1\right)} \right\|_1 \right] \tag{5}$$

In Equation (5), $\nabla$ represents the first-order derivative operator with respect to the horizontal direction $X$ and the vertical direction $Y$. $L_{low}$ and $L_{high}$ represent the input low-light and normal-light images, respectively. The value $\varepsilon_1$ denotes a small positive constant, such as 0.01, to prevent division by zero in the equation. $||$ represents the absolute value operation, and $\lambda_i$ represents a weight parameter.

In the third part, the decomposition network of the low-light image is guided by rich features from real images. Here, the Charbonnier loss [27] is employed, as shown in Equation (6):

$$L_{Charbonnier} = \sqrt{\|R_i I_i - L_i^*\| + \varepsilon_2^2}, \quad i \in (low, high) \tag{6}$$

In Equation (6), $L^*$ represents the input ground truth image, and $\varepsilon_2$ is a constant factor set to $10^{-3}$ in the experiments. Therefore, the overall loss function of the decomposition network is the sum of the three individual loss functions. This combined loss function guides the training process of the network to simultaneously minimize the reconstruction loss, ensure consistency between the low-light and high-light components, and leverage the rich features from real images for accurate decomposition.

The proposed loss function for denoising the reflection component in the Enhance-net is expressed as Equation (7):

$$L_{reflect} = \left\| R'_{low} - R_{high} \right\|_2 + \left\| \nabla R'_{low} - \nabla R_{high} \right\|_2 + \lambda_c L_{color} \tag{7}$$

In Equation (7), $R'_{low}$ represents the denoised low-light reflection component, and $\lambda_c$ denotes the weight parameter. To prevent image distortion and evaluate the color fidelity between the enhanced image result and the target image, a color loss function $L_{color}$ based on the Grey-World algorithm [28] is introduced. This $L_{color}$ function constrains the relationship between the three channels to correct color errors within each channel. It is expressed as Equations (8) and (9):

$$\overline{gray} = \frac{\bar{R} + \bar{G} + \bar{B}}{3} \tag{8}$$

$$L_{color} = \sqrt{\sum\nolimits_{k=1}^{N} W_K \left(C_i - \overline{gray}\right)^2}, \quad C_i \in \left(\bar{R}, \bar{G}, \bar{B}\right) \tag{9}$$

In Equations (8) and (9), $\bar{R}$, $\bar{G}$, $\bar{B}$ represent the average values of the three color channels in the enhanced image. $\overline{gray}$ represents the average grayscale value, $W_K$ is a weight parameter with a value of 1/3, and $N$ is a constant equal to 3. The loss function imposes constraints on the three-channel components, aiming for their mean values to be consistent.

The loss function proposed for enhancing the brightness compensation component of low-light images in the enhancement network is expressed as Equation (10):

$$L_{enhance} = \|I'_{low} - I_{high}\|_2^2 + \|\nabla I'_{low} - \nabla I_{high}\|_2^2 \tag{10}$$

4. **Experiments.** The experiments were conducted using a Windows 10 operating system with a computer configuration of 4GB RAM, i5 2.50GHz processor, 16GB of running memory, and an NVIDIA GeForce RTX 3050 GPU with CUDA version 11.2. The software platform used was PyCharm. The experimental image dataset used was the LOL dataset, which consists of 500 low-light and normal-light images, with 485 images in the training set and 15 images in the testing set. Additionally, the NPE dataset [29], ExDark dataset [30] and VV dataset [31] were also used, with 84 low-light images selected from the NPE dataset, 242 randomly selected images from the ExDark dataset and 24 images from the VV dataset.

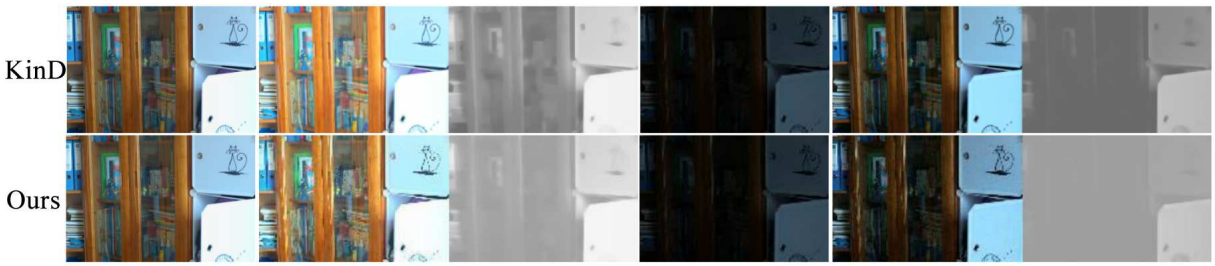
The experimental parameters were set as follows: $\lambda_r = 0.01$, $\lambda_i = 0.15$, $\lambda_c = 0.004$. We referred to the parameter settings in [19] to determine the approximate parameter range, and then conducted multiple experiments to finally determine the parameter values. In order to verify the rationality of the selected value parameters, we conducted the following experiments: in the LOL dataset, the model was trained with $\lambda_r$, $\lambda_i$, and $\lambda_c$ values of half, original, and twice, respectively. 15 images in the testing set with enhanced brightness were obtained. Peak Signal-to-Noise Ratio (PSNR) [32], Structural Similarity Index (SSIM) [33], and Mean Squared Error (MSE) are used as evaluation criteria. PSNR measures the quality of the image by calculating the pixel-wise error, and a higher PSNR value indicates better enhancement. SSIM is used to evaluate the similarity between the original and enhanced images, and a higher SSIM value indicates better enhancement. MSE represents the mean squared error between the original and enhanced images. For the first two metrics, higher values indicate better performance (represented as '↑' in this study). As for the MSE metric, lower values indicate better performance (represented as '↓' in this study). The average values of these metrics for the LOL datasets are summarized in Table 1. The number of iterations for the Decom-net and the illumination compensation component of the Enhance-net was set to 2000, while for the denoising component of the Enhance-net, it was set to 1000. The learning rate was set to 0.0001, and the optimizer used was Adam.

Table 1. Experimental parameters setting verification

| Image group | Target | Values of half | Values of twice | Values of original image |
|---|---|---|---|---|
| $\lambda_r$ | PSNR↑ | 19.633 | 19.050 | **19.860** |
| | SSIM↑ | 0.762 | 0.765 | **0.783** |
| | MSE↓ | 0.043 | 0.041 | **0.038** |
| $\lambda_i$ | PSNR↑ | 18.903 | 18.962 | **19.860** |
| | SSIM↑ | 0.735 | 0.766 | **0.783** |
| | MSE↓ | 0.048 | 0.040 | **0.038** |
| $\lambda_c$ | PSNR↑ | 19.256 | 19.213 | **19.860** |
| | SSIM↑ | 0.750 | 0.757 | **0.783** |
| | MSE↓ | 0.046 | 0.045 | **0.038** |

In Table 1, taking $\lambda_r$ as an example, whether the parameter value is reduced or expanded, the number of PSNR, SSIM, and MSE will decrease. $\lambda_i$ and $\lambda_c$ are also the same. So we get $\lambda_r = 0.01$, $\lambda_i = 0.15$, $\lambda_c = 0.004$.

4.1. **Validation of the effectiveness of the Decom-net.** This section primarily validates the impact of adding Charbonnier loss in the decomposition network on the quality of image decomposition. Additionally, a comparison was made between the proposed algorithm and the KinD-Net [21], as they share similarities in the decomposition process. The comparison results are shown in Figure 4.




(a) $L$ (b) The reflection (c) The brightness (d) $L$ (e) The reflection (f) The brightness

FIGURE 4. Comparison of Decom-net result

As for the mutual consistency, the method used by KinD-Net was to preserve strong mutual edges and depress weak ones. However, it is difficult to distinguish between strong mutual edges and weak ones. To solve this problem, the proposed algorithm incorporates Charbonnier loss in the loss function to ensure the consistency of the reflectance components between low-light and normal-light images. Directly use the plentiful features of real images for pixel level mapping of low illumination images to avoid distinguishing between strong and weak mutual edges. Besides, the brightness component contains a significant amount of low-frequency information, and this loss function helps to achieve maximum smoothness and uniformity in the brightness component while preserving fine details. From Figure 4, it can be observed that the proposed method yields smoother brightness components for both low-light and normal-light images compared to the KinD-Net. The presence of fewer high-frequency details in the brightness component allows for more comprehensive feature extraction from the images. The smoother nature of the brightness components in the proposed algorithm contributes to the preservation of image details while promoting a more uniform representation. This leads to improved feature extraction capabilities and enhances the overall performance of the algorithm.

4.2. **Validation of the effectiveness of the Enhance-net.**

1) Denoising component

To validate the effectiveness of NRED-Net on low-light reflectance components, a comparative experiment was designed by removing certain components from the loss function. Specifically, the color loss was excluded from the loss function of NRED-Net for denoising the reflectance component $R_{low}$. The resulting denoised reflectance component is denoted as $R'_{low\_COLOR}$, and the corresponding restored brightness map is denoted as $L'_{COLOR}$. Additionally, a separate experiment was conducted without utilizing the NRED-Net network structure, where the complete loss function was used to train a denoising model. The outcome of this experiment is represented by $R'_{low\_NRED}$, and the corresponding restored brightness map is denoted as $L'_{NRED}$. The results obtained by the proposed algorithm are denoted as $R'_{low}$, with the corresponding restored brightness map denoted as $L'$. The original low-light image is represented as $L$. The comparative effects are illustrated in Figure 5 and Figure 6.

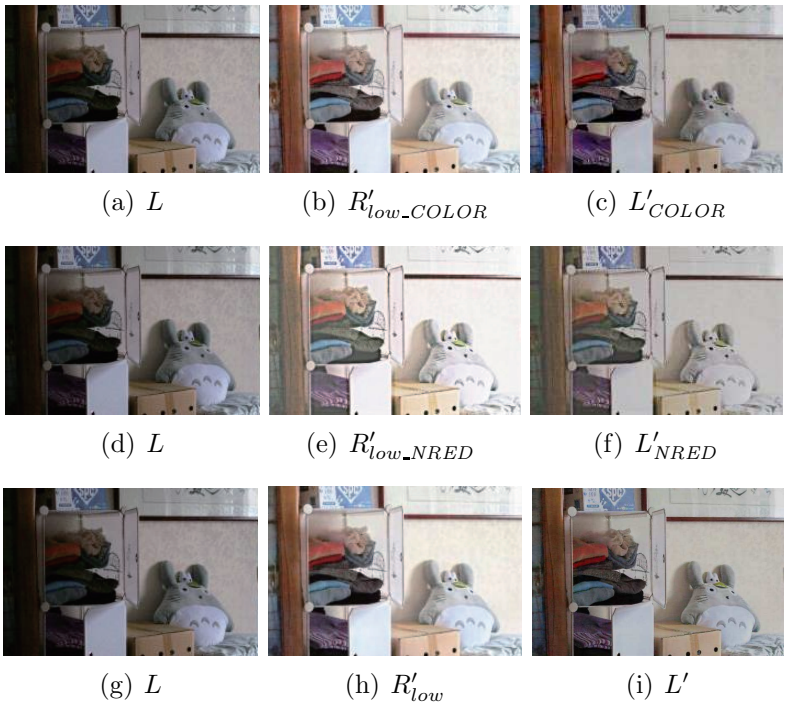
From Figures 5 and 6, it can be observed that the denoised reflectance image $R'_{low\_COLOR}$ obtained without using the color loss exhibits color distortion, as seen in Figure 5(b) where

(a) $L$ (b) $R'_{low\_COLOR}$ (c) $L'_{COLOR}$

(d) $L$ (e) $R'_{low\_NRED}$ (f) $L'_{NRED}$

(g) $L$ (h) $R'_{low}$ (i) $L'$

FIGURE 5. Module removal comparison (1)



(a) $L$ (b) $R'_{low\_COLOR}$ (c) $L'_{COLOR}$

(d) $L$ (e) $R'_{low\_NRED}$ (f) $L'_{NRED}$

(g) $L$ (h) $R'_{low}$ (i) $L'$

FIGURE 6. Module removal comparison (2)

there are black redundant regions in the background, and the overall brightness enhancement in the generated restored brightness map, as shown in Figure 5(c), is not significant. The denoised reflectance image $R'_{low\_NRED}$ obtained without utilizing the NRED-Net network structure contains some noise artifacts, such as the edges of the storage cabinet on the left side in Figure 6(e), and the generated restored brightness map also suffers from

noise contamination. In contrast, the denoised reflectance image $R'_{low}$ obtained using the NRED-Net approach shows natural enhancement in brightness, well-preserved color, and clear contours.

Furthermore, to objectively evaluate the denoising performance, several quality metrics were utilized, including Information Entropy (IE), Standard Deviation (SD), and Natural Image Quality Evaluator (NIQE) [34]. For the first two metrics, higher values indicate better performance, indicating richer image information and details. As for the NIQE metric, lower values indicate better performance, indicating closer resemblance to natural images. These metrics measure image information content, image variability, and naturalness, respectively. The average values of these metrics for the image sets in Figures 5 and 6 are summarized in Table 2.

TABLE 2. Comparison of indicators of the removal module of the enhanced network denoising component

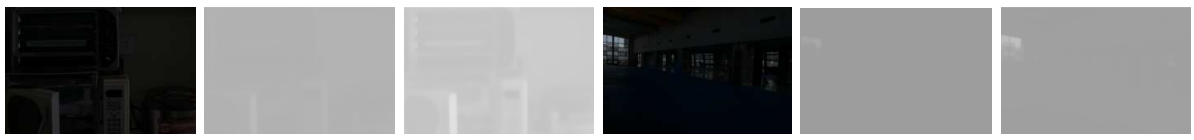
| Image group | $\overline{IE}\uparrow$ | $\overline{SD}\uparrow$ | $\overline{NIQE}\downarrow$ |
|---|---|---|---|
| $L$ | 6.732 | 41.752 | 11.200 |
| $R'_{low\_COLOR}$ | 6.964 | 54.043 | 11.432 |
| $R'_{low\_NRED}$ | 6.781 | 46.781 | 13.611 |
| $R'_{low}$ | **7.222** | **57.911** | **10.314** |
| $L'_{COLOR}$ | 6.843 | 40.792 | 10.972 |
| $L'_{NRED}$ | 6.494 | 35.442 | 13.650 |
| $L'$ | 6.950 | 43.200 | 10.531 |

From Table 2, it can be observed that the proposed method, which utilizes the NRED-Net network structure and incorporates the color loss, achieves higher values in most of the evaluated parameters compared to the algorithms that solely utilize the color loss or the NRED-Net network structure alone. This validates the effectiveness of the denoising component in the enhancement network of the proposed algorithm.

2) Illumination compensation component

To verify the brightness enhancement effect of CBAM on the low-light brightness component, this section compares the algorithm with CBAM to the case where CBAM is removed. Specifically, one set of images is processed without CBAM, denoted as $I'_{CBAM}$, and the corresponding restored luminance image is denoted as $L'_{CBAM}$. The other set of images includes the introduction of CBAM in the proposed algorithm, denoted as $I'_{low}$, and the corresponding restored luminance image is denoted as $L'$, with the low-light image denoted as $L$. The comparison results are shown in Figures 7 and 8.

From Figure 7, it can be observed that CBAM performs information filtering on the global low-light image and suppresses the background redundant information, resulting in enhanced brightness of the low-light brightness component. For example, in Figure 7, the edges of the oven door and the windows of the stadium are more distinct. Figure



(a) $L$ (b) $I'_{CBAM}$ (c) $I'_{low}$ (d) $L$ (e) $I'_{CBAM}$ (f) $I'_{low}$

FIGURE 7. Comparison of whether CBAM is used

(a) $L'_{CBAM}$ (b) $L'$

FIGURE 8. Comparison of normal illumination map generated by CBAM or not

TABLE 3. Comparison of indicators of the removal module of the CBAM

| Image | Image group | *IE* ↑ | *SD* ↑ | *NIQE* ↓ |
|---|---|---|---|---|
| Oven | $L$ | 4.763 | 9.213 | 12.243 |
| | $I'_{CBAM}$ | 2.592 | 10.211 | 22.722 |
| | $I'_{low}$ | 3.481 | 22.013 | 17.861 |
| | $L'_{CBAM}$ | 6.990 | 41.074 | 14.742 |
| | $L'$ | **7.033** | **41.240** | **12.983** |
| Gymnasium | $L$ | 4.452 | 14.031 | 14.020 |
| | $I'_{CBAM}$ | 1.381 | 4.082 | 20.750 |
| | $I'_{low}$ | 1.983 | 4.404 | 19.752 |
| | $L'_{CBAM}$ | 6.344 | 30.474 | 13.721 |
| | $L'$ | **6.372** | **30.741** | **12.282** |

8 demonstrates that the visually appealing high-brightness images are generated when CBAM is utilized, with richer details in the floor of the stadium, as shown in Figure 8. Furthermore, to objectively evaluate the brightness enhancement effect, IE, SD, and NIQE are used to compare the brightness component images restored with and without CBAM, as well as the corresponding generated high-brightness images. The results are presented in Table 3.

From Table 3, it can be observed that the information entropy, standard deviation, and natural image quality evaluation of the low-light brightness component image and the corresponding restored brightness image with CBAM are higher than those without CBAM. This verifies the effectiveness of introducing CBAM.

4.3. **Generalization ability.** To validate the universality of the proposed algorithm and ensure its applicability in various scenarios, a universality validation is conducted using multiple datasets, including the LOL dataset, NPE dataset, ExDark dataset, and VV dataset. Specifically, 15 low-light images from the LOL dataset, 84 images from the NPE dataset, 24 images from the VV dataset, and 242 randomly selected images from the ExDark dataset are chosen for the universality validation. The comparison results are shown in Figure 9. Figure 9 illustrates the comparative performance of the algorithm on different datasets, showcasing its ability to handle diverse low-light scenarios. The results demonstrate that the proposed algorithm consistently achieves improved image quality and effective enhancement across various datasets, reinforcing its universality and robustness.

In Figure 9, the images are arranged in pairs, showing the low-light images from the datasets and the corresponding enhanced images using the proposed algorithm. Each dataset consists of three selected images to demonstrate the effectiveness of the image

(a) LOL dataset

(b) VV

(c) NPE

(d) ExDark

FIGURE 9. Comparison of effects of different data sets before and after algorithm enhancement in our method

enhancement. From Figure 9, it is evident that the proposed algorithm significantly improves the brightness of the images. Furthermore, to objectively evaluate the improvement in image quality achieved by the algorithm, three quality evaluation metrics are used: IE, SD, and NIQE. The average values of these metrics are computed for each dataset, with the low-light images denoted as $L$ and the enhanced images denoted as $L'$. The results are presented in Table 4. Table 4 shows the average values of the three evaluation metrics for the datasets used in this study. The results demonstrate that the proposed algorithm effectively enhances image quality, as indicated by higher IE and SD values and lower NIQE values. This confirms the improved visual quality and naturalness of the enhanced images produced by the algorithm.

From Table 4, it can be observed that the proposed algorithm improves and enhances various metrics across all datasets. Particularly, the SD value for the LOLdataset is increased fourfold, indicating the most significant improvement in image enhancement. These results validate that the proposed algorithm effectively enhances the details and brightness of low-light images, demonstrating its strong universality and robustness.

4.4. **Evaluation.** To compare the proposed algorithm with MSR [8], MSRCR [9], Retinex-Net [19], LIME [10], and KinD-Net [21], representative images from the LOL dataset are selected for comparison. The results are shown in Figure 10.

In Figure 10, MSR, MSRCR, Retinex-Net, LIME, and KinD-Net are compared in terms of their impact on image brightness. These algorithms show varying degrees of improvement in image brightness. However, MSR and MSRCR algorithms exhibit more overall noise, and in G1, there is color distortion around the wooden vertical bar on the left, as well as color distortion in the wardrobe in G2. The Retinex-Net algorithm also

TABLE 4. Comparison of evaluation indicators of different data sets before and after algorithm enhancement in our method

| Dataset | Image group | $\overline{IE}\uparrow$ | $\overline{SD}\uparrow$ | $\overline{NIQE}\downarrow$ |
|---|---|---|---|---|
| (a) LOL dataset | $L$ | 4.700 | 10.070 | 11.233 |
| | $L'$ | 6.930 | 43.420 | 10.012 |
| (b) VV | $L$ | 6.630 | 69.880 | 10.741 |
| | $L'$ | 7.360 | 75.270 | 10.293 |
| (c) NPE | $L$ | 6.890 | 56.982 | 12.603 |
| | $L'$ | 7.330 | 66.200 | 12.042 |
| (d) ExDark | $L$ | 5.840 | 40.614 | 12.343 |
| | $L'$ | 6.930 | 56.873 | 11.550 |

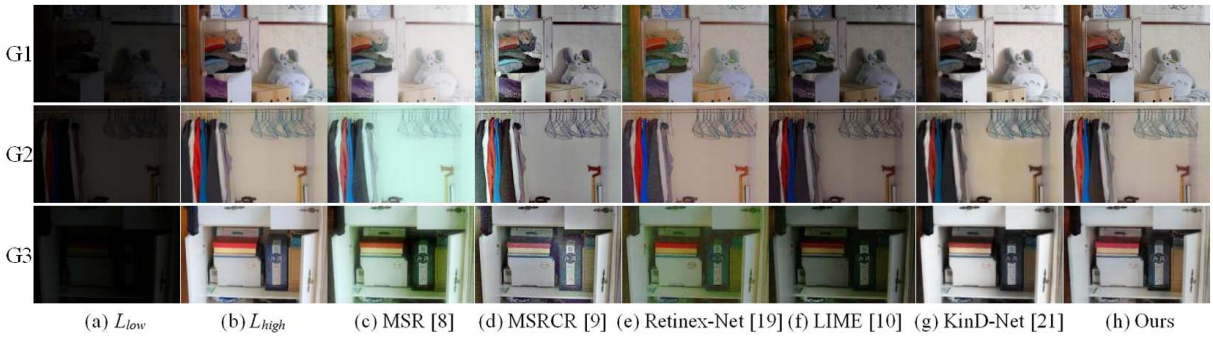



FIGURE 10. Comparison of enhancement effects of different algorithms in LOL dataset

suffers from color distortion, with green color distortion appearing in the bookshelf in G3. The results of the LIME algorithm do not adequately preserve the details, resulting in less clarity. For example, the objects around the doll in G1 appear darker overall. KinD-Net tends to oversharpen image details, leading to the loss of details in brighter areas. For instance, the details of the wooden vertical bar on the left in G1 are not easily observable. In contrast, the proposed algorithm in this paper achieves a natural color restoration and rich hierarchical details. To objectively evaluate the algorithm's performance, PSNR, SSIM, and MSE are used as evaluation criteria. The comparative results are shown in Table 5.

In Table 5, the proposed algorithm performs the best in several evaluation metrics, particularly in terms of MSE and SSIM, indicating that it enhances image quality with minimal distortion. Additionally, it demonstrates good performance in contrast enhancement and exhibits a certain level of robustness. In summary, the proposed algorithm exhibits good universality and robustness. Furthermore, to ensure the algorithm's generalizability, a comparison is conducted between the proposed algorithm and other different algorithms using the NPE dataset. The results are shown in Figure 11.

From Figure 11, it can be observed that the selected comparative algorithms have considerable effects on enhancing low-light images. Overall, the MSR algorithm, MSRCR algorithm, and Retinex-Net algorithm exhibit varying degrees of color distortion. The LIME algorithm lacks sufficient improvement in brightness, as seen in Figure 11 where the brightness enhancement in the vegetation of G3 is not very pronounced. The KinD-Net algorithm suffers from oversharpening, as evident in Figure 11 where the feathers of

TABLE 5. Comparison of low illumination image enhancement effects of different algorithms in LOL dataset

| Image group | Target | MSR [8] | MSRCR [9] | Retinex-Net [19] | LIME [10] | KinD-Net [21] | Ours |
|---|---|---|---|---|---|---|---|
| G1 | PSNR↑ | 15.140 | 13.510 | 17.900 | 12.370 | 19.500 | **23.550** |
| | SSIM↑ | 0.800 | 0.680 | 0.670 | 0.760 | 0.930 | **0.940** |
| | MSE↓ | 0.300 | 0.044 | 0.016 | 0.057 | 0.011 | **0.004** |
| G2 | PSNR↑ | **18.080** | 12.600 | 8.750 | 6.340 | 8.880 | 10.050 |
| | SSIM↑ | **0.850** | 0.780 | 0.360 | 0.350 | 0.430 | 0.450 |
| | MSE↓ | **0.020** | 0.054 | 0.133 | 0.232 | 0.129 | 0.098 |
| G3 | PSNR↑ | **28.890** | 21.780 | 25.070 | 14.080 | 25.550 | 28.270 |
| | SSIM↑ | 0.790 | 0.640 | 0.570 | 0.500 | 0.880 | **0.900** |
| | MSE↓ | 0.010 | 0.026 | 0.012 | 0.156 | 0.011 | **0.005** |

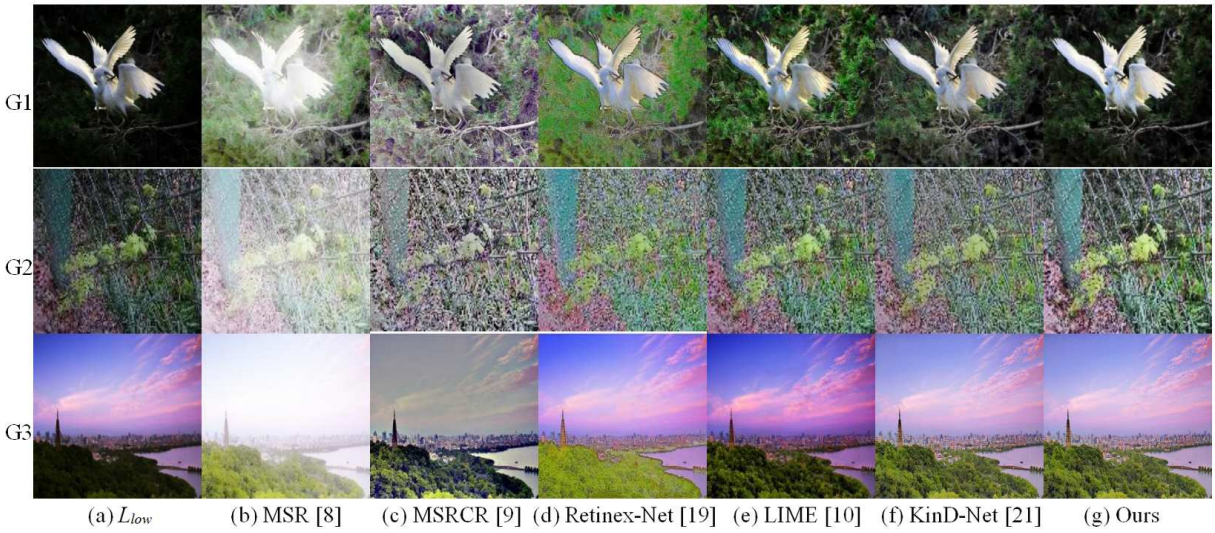



FIGURE 11. Comparison of enhancement effects of different algorithms in NPE dataset

the bird in G1 appear overly sharp, and there is a color deviation between the wire mesh and the middle vegetation in G2. In contrast, the proposed algorithm maintains natural colors, clear contours, and natural brightness enhancement. The comparative results are shown in Table 6.

From Table 6, it can be observed that the proposed algorithm performs well across all three image groups, particularly excelling in the MSE metric. This indicates that the algorithm can enhance image quality with minimal distortion. Furthermore, the algorithm demonstrates the best performance in terms of PSNR and SSIM metrics in G1 and G3. These results collectively highlight the advantages of the proposed algorithm across multiple evaluation metrics.

5. **Conclusion.** This paper proposes a method for enhancing low-light images using a neural network with asymmetric convolutional kernels combined with an attention mechanism. The neural network framework consists of a decomposition network and an enhancement network, with the enhancement network comprising a denoising component and an illumination compensation component. The main contributions of this paper are

TABLE 6. Comparison of low illumination image enhancement effects of different algorithms in NPE

| Image group | Target | MSR [8] | MSRCR [9] | Retinex-Net [19] | LIME [10] | KinD-Net [21] | Ours |
|---|---|---|---|---|---|---|---|
| G1 | PSNR↑ | 6.390 | 7.330 | 9.040 | 13.390 | 14.370 | **19.350** |
| | SSIM↑ | 0.210 | 0.240 | 0.250 | 0.340 | 0.420 | **0.730** |
| | MSE↓ | 0.229 | 0.184 | 0.124 | 0.045 | 0.036 | **0.011** |
| G2 | PSNR↑ | 5.860 | 10.950 | 11.220 | 13.360 | 12.820 | **13.640** |
| | SSIM↑ | 0.520 | 0.610 | 0.630 | 0.720 | 0.730 | **0.780** |
| | MSE↓ | 0.259 | 0.080 | 0.075 | 0.046 | 0.052 | **0.043** |
| G3 | PSNR↑ | 7.970 | 14.840 | 13.140 | 17.330 | 15.650 | **19.260** |
| | SSIM↑ | 0.690 | 0.790 | 0.770 | 0.870 | 0.810 | **0.920** |
| | MSE↓ | 0.159 | 0.099 | 0.048 | 0.018 | 0.027 | **0.011** |

as follows: in the decomposition network, the Charbonnier loss is introduced in the loss function to enforce smoothness and uniformity in the low-light luminance component. In the denoising component of the enhancement network, the RED-Net model is improved by using asymmetric convolutional kernels to denoise the low-light reflectance component. Additionally, a color loss term is added to the loss function to mitigate color distortion during the denoising process. In the illumination compensation component of the enhancement network, the attention mechanism CBAM is introduced to enhance the image enhancement results. Experimental results demonstrate the effectiveness and generalization capability of the proposed algorithm. It can be observed that the algorithm improves both image quality and visual effects. However, it should be noted that currently, there is a scarcity of low-light/normal-light image datasets, and the process of creating such datasets is laborious. Further research is needed to explore no-reference methods for enhancing low-light images in the future.

**Acknowledgments.** This work was partially supported by "the Fundamental Research Funds for the Central Universities (J2023-024)".

## REFERENCES

[1] A. N. N. Afifah, Indrabayu, A. Suyuti and Syafaruddin, Improving the image quality of grayscale thermal images taking from photovoltaic panel with contrast enhancement method, *International Journal of Innovative Computing, Information and Control*, vol.19, no.1, pp.197-212, DOI: 10.24507/ijicic.19.01.197, 2023.

[2] H. Tang, H. Zhu, L. Fei et al., Low-illumination image enhancement based on deep learning techniques: A brief review, *J. Photonics.*, vol.10, no.2, 198, DOI: 10.3390/photonics10020198, 2023.

[3] Y. Wu, C. Pan, G. Wang et al., Learning semantic-aware knowledge guidance for low-light image enhancement, *Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp.1662-1671, DOI: 10.3390/photonics10020198, 2023.

[4] Y. Huang, X. Tu, G. Fu et al., Low-light image enhancement by learning contrastive representations in spatial and frequency domains, *arXiv Preprint*, arXiv: 2303.13412, 2023.

[5] J. Chen, Q. Lian, X. Zhang et al., Hcsam-Net: Multistage network with a hybrid of convolution and self-attention mechanism for low-light image enhancement, *SSRN Electronic Journal*, DOI: 10.2139/ssrn.4237486, 2022.

[6] Y. Li, T. Liu, J. Fan et al., LDNet: Low-light image enhancement with joint lighting and denoising, *J. Machine Vision and Applications*, vol.34, no.1, 13, DOI: 10.1007/s00138-022-01365-z, 2023.

[7] D. J. Jobson, Z. Rahman and G. A. Woodell, Properties and performance of a center/surround retinex, *IEEE Transactions on Image Processing*, vol.6, no.3, pp.451-462, DOI: 10.1109/83.557356, 1997.

[8] D. J. Jobson, Z. Rahman and G. A. Woodell, A multiscale retinex for bridging the gap between color images and the human observation of scenes, *IEEE Transactions on Image Processing*, vol.6, no.7, pp.965-976, DOI: 10.1109/83.597272, 1997.

[9] Z. Rahman, D. J. Jobson and G. A. Woodell, Retinex processing for automatic image enhancement, *Journal of Electronic Imaging*, vol.13, no.1, DOI: 10.1117/1.1636183, 2004.

[10] X. Guo, Y. Li and H. Ling, LIME: Low-light image enhancement via illumination map estimation, *IEEE Transactions on Image Processing*, vol.26, no.2, pp.982-993, DOI: 10.1109/TIP.2016.2639450, 2017.

[11] A. Ignatov and R. Timofte, NTIRE 2019 challenge on image enhancement: Methods and results, *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, Long Beach, CA, USA, pp.2224-2232, DOI: 10.1109/CVPRW.2019.00275, 2019.

[12] Z. Cui, K. Li, L. Gu et al., Illumination adaptive transformer, *arXiv Preprint*, arXiv: 2205.14871, 2022.

[13] Y. Zhang, X. Guo, J. Ma et al., Beyond brightening low-light images, *International Journal of Computer Vision*, vol.129, pp.1013-1037, DOI: 10.1007/s11263-020-01407-x, 2021.

[14] W. Wu, J. Weng, P. Zhang et al., URetinex-Net: Retinex-based deep unfolding network for low-light image enhancement, *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, New Orleans, LA, USA, pp.5891-5900, DOI: 10.1109/CVPR52688.2022.00581, 2022.

[15] S. Surya and A. Muthukumaravel, Significance of deep learning in artificial intelligence systems, *2023 9th International Conference on Advanced Computing and Communication Systems (ICACCS)*, pp.2220-2223, 2023.

[16] Y. Wang, R. Wan, W. Yang, H. Li, L.-P. Chau and A. C. Kot, Low-light image enhancement with normalizing flow, *arXiv Preprint*, arXiv: 2109.05923, 2021.

[17] K. Hu, C. Weng, C. Shen et al., A multi-stage underwater image aesthetic enhancement algorithm based on a generative adversarial network, *Engineering Applications of Artificial Intelligence*, vol.123, 106196, DOI: 10.1016/j.engappai.2023.106196, 2023.

[18] L. Shen et al., MSR-net: Low-light image enhancement using deep convolutional network, *arXiv Preprint*, arXiv: 1711.02488, 2017.

[19] C. Wei et al., Deep Retinex decomposition for low-light enhancement, *arXiv Preprint*, arXiv: 1808.04560, 2018.

[20] K. Dabov, A. Foi, V. Katkovnik and K. Egiazarian, Image denoising by sparse 3-D transform-domain collaborative filtering, *IEEE Transactions on Image Processing*, vol.16, no.8, pp.2080-2095, DOI: 10.1109/TIP.2007.901238, 2007.

[21] Y. Zhang, J. Zhang and X. Guo, Kindling the darkness: A practical low-light image enhancer, *Proc. of the 27th ACM International Conference on Multimedia (MM'19)*, Nice, France, 2019.

[22] Y. Jiang, X. Gong, D. Liu et al., EnlightenGAN: Deep light enhancement without paired supervision, *IEEE Transactions on Image Processing*, vol.30, pp.2340-2349, DOI: 10.1109/TIP.2021.3051462, 2021.

[23] Z. Li et al., Low illumination video image enhancement, *IEEE Photonics Journal*, vol.12, no.4, pp.1-13, Art No. 3900613, DOI: 10.1109/JPHOT.2020.3010966, 2020.

[24] X. J. Mao, C. Shen and Y. B. Yang, Image restoration using convolutional auto-encoders with symmetric skip connections, *arXiv Preprint*, arXiv: 1606.08921, 2016.

[25] X. Ding, Y. Guo, G. Ding et al., ACNet: Strengthening the kernel skeletons for powerful CNN via asymmetric convolution blocks, *Proc. of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pp.1911-1920, 2019.

[26] S. Woo, J. Park, J. Y. Lee and I. S. Kweon, CBAM: Convolutional block attention module, in *Computer Vision – ECCV 2018. ECCV 2018. Lecture Notes in Computer Science*, V. Ferrari, M. Hebert, C. Sminchisescu and Y. Weiss (eds.), Cham, Springer, 2018.

[27] S. W. Zamir, A. Arora, S. Khan et al., Learning enriched features for real image restoration and enhancement, in *Computer Vision – ECCV 2020. ECCV 2020. Lecture Notes in Computer Science*, A. Vedaldi, H. Bischof, T. Brox and J. M. Frahm (eds.), Cham, Springer, 2020.

[28] G. Buchsbaum, A spatial processor model for object colour perception, *Journal of the Franklin Institute*, vol.310, no.1, pp.1-26, 1980.

[29] S. Wang, J. Zheng, H. M. Hu et al., Naturalness preserved enhancement algorithm for non-uniform illumination images, *IEEE Transactions on Image Processing*, vol.22, no.9, pp.3538-3548, DOI: 10.1109/TIP.2013.2261309, 2013.

[30] C. Li, C. Guo, L. Han et al., Low-light image and video enhancement using deep learning: A survey, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.44, no.12, pp.9396-9416, DOI: 10.1109/TPAMI.2021.3126387, 2021.
[31] C. Guo, C. Li, J. Guo, C. C. Loy, J. Hou, S. Kwong and R. Cong, Zero-reference deep curve estimation for low-light image enhancement, *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Seattle, WA, USA, pp.1777-1786, DOI: 10.1109/CVPR42600.2020.00185, 2020.
[32] Q. Huynh-Thu and M. Ghanbari, Scope of validity of PSNR in image/video quality assessment, *Electronics Letters*, vol.44, no.13, pp.800-801, DOI: 10.1049/el:20080522, 2008.
[33] Z. Wang, A. C. Bovik and H. R. Sheikh, Image quality assessment: From error visibility to structural similarity, *IEEE Transactions on Image Processing*, vol.13, no.4, pp.600-612, DOI: 10.1109/TIP.2003.819861, 2004.
[34] A. Mittal, Making a completely blind image quality analyzer, *IEEE Signal Processing Letters*, vol.20, no.3, pp.209-212, DOI: 10.1109/LSP.2012.2227726, 2013.

## Author Biography

**Jiajia Liu** is an associate professor at the Institute of Electronic and Electrical Engineering, Civil Aviation Flight University of China. She received her B.S. and M.S. degrees in Communication and Information Systems from the Sichuan University in 2008 and 2011, respectively. Her current research interests include image processing, and cloud computing. She is member of China Aviation Society and Sichuan Electronics Society. She is also holding a fixed wing (Class IV) pilot license for beyond visual range vertical takeoff and landing, a civil aircraft maintenance personnel license, a Class II AV professional license for PA-44-180 (LO-360) aircraft, and concurrently serving as a police unmanned aerial vehicle flight instructor in Jining, Shandong Province.

**Zhixiang Deng** received his B.S. degree in Computer Science and Technology from Yunnan Agricultural University, Kunming, China, in 2021. He is currently pursuing his M.S. degree in Machine at Institute of Electronic and Electrical Engineering, Civil Aviation Flight University of China, Guanghan, China. His research interest is image processing and computer vision, and he currently encourages his competence in UAVs projects.