IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 15 February 2024, accepted 17 March 2024, date of publication 25 March 2024, date of current version 2 April 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3381514

RESEARCH ARTICLE

A Low-Light Image Enhancement Algorithm Incorporating Cross-Mixed Attention and Receptive Field Expansion Mechanism

RONGFENG ZHOU, PING LI, MINGYANG ZHANG, QINGYAO LIN, YUANYUAN WANG, XUESHENG BIAN, FENG ZHOU, AND RUGANG WANG

School of Information Technology, Yancheng Institute of Technology, Yancheng 224051, China

Corresponding author: Ping Li (infoliping@ycit.edu.cn)

This work was supported in part by Jiangsu Graduate Practical Innovation Project under Grant SJCX22_1685, in part by the Natural Science Foundation of China under Grant 62301473, in part by the Major Project of Natural Science Research of Jiangsu Province Colleges and Universities under Grant 19KJA110002, and in part by the Natural Science Research Project of Jiangsu University under Grant 18KJD510010.

ABSTRACT Aiming at the conventional low-light enhancement algorithms for low-light image enhancement with problems of loss of details, low contrast and low color saturation, a detection algorithm called CE-Retinex(Cross Expansion Retinex) by incorporating the cross-mixed attention mechanism and the receptive field expansion mechanism is proposed and demonstrated for the first time. It mainly consists of three parts: initial module, optimization module, and detail restoration. Firstly, in the initial module, the image is decomposed into reflectance and illuminance using two different U-shaped networks with multiple layers of convolution. Secondly, in the optimization network module, the image brightness is enhanced using multi-scale lighting mechanism. Denoising is also performed. After that, a multi-layer convolutional fusion cross-mixed attention mechanism is used to filter the information from the four dimensions of channel, spatial, vertical and horizontal attention, so that it can be effectively reduce the negative effects of the low-light image enhancement process. In the detail restoration module, the receptive field expansion mechanism is utilized to enhance the receptive field and strengthen the detail information. Also, the color consistency loss function is used to recover colors in the loss function. The CE-Retinex algorithm was experimentally analyzed on the LOL dataset, and the PSNR of the CE-Retinex algorithm was 25.33, and the NIQE was 3.37, and the subjective feelings and objective evaluation indicators have effectively improved. So the proposed algorithm could be effectively solve the problems of loss of detail, low contrast, and low color saturation in low-light image enhancement.

INDEX TERMS U-shaped network, multi-scale illumination mechanism, cross-mixed attention mechanism, receptive field expansion mechanism.

I. INTRODUCTION

Low-light image enhancement is an important area in the field of computer vision, where the main objective is to enhance the visual quality of images in low light environments. So, the low-light image enhancement have received tremendous boost owing to many potential applications, such as nighttime aerial photography [1], [2], self-driving cars at night [3], [4], image segmentation [5], [6], [7], [8],etc. Low-light image enhancement algorithms are mainly classified into three categories: distribution mapping-based methods [9], [10], model optimization-based methods [11], [12], and deep learning-based methods [13], [14], [15]. The curve transformations, histogram equalization, and other means were used to adjust information of low-light images in the distribution mapping-based methods, however, in these

The associate editor coordinating the review of this manuscript and approving it for publication was Zhenbao Liu.

methods, there will be drawbacks such as overexposure and underexposure. In model optimization-based methods, many model parameters of Retinex theory must be manually adjusted to improve image brightness, so that the methods were not adaptable enough. The method based on deep learning can effectively overcome the problems of poor robustness of traditional algorithms in manually selecting thresholds, and the large influence of invalid background information on statistical features based on global information, and has good enhancement effects. So that the deep learning-based methods have better accuracy and speed compared to the first two methods.

Currently, deep learning-based methods have become the mainstream methods for low-light image enhancement. In 2017, the LLNet algorithm which employs a variant of the superimposed sparse denoising autoencoder for low-light image enhancement and denoising is proposed by Lore et al., and it is the first deep-learning based algorithm for low-light image enhancement [16]. In 2018, the end-to-end supervised learning multi-branch fusion MBLLEN algorithm is proposed by Lv et al., which improves the performance of the algorithm through the feature extraction module, enhancement module, and fusion module, but the algorithm has insufficient contrast and brightness [17]. In order to accurately estimate the illumination map, in 2019, through learning the image-to-illumination mapping relationship, the Deep UPE is proposed to extract global and local features by Wang et al., which presents pleasant contrast but suffers from underexposure and lack of details [18]. In 2020, the TBEFN algorithm is proposed and which uses a two-branch averaging estimation scheme to fuse the images end-to-end and refine the results, but it still needs to be optimized in terms of performance [19]. In 2021, the EnlightenGAN algorithm for self-supervised low-light image enhancement algorithm is proposed by Jiang et al., which trains the generator and discriminator and incorporates a self-attention mechanism and self-regularized perceptual loss for enhancement [20]. In 2022, in order to obtain better performance, the optimized URetinex-Net algorithm is proposed, and the new constraints and advanced network design was added to the URetinex-Net algorithm [21], [22]. In 2023, the PairLIE algorithm is proposed to reduce reliance on manual priors. It benefits from both Retinex theory and deep learning-based solutions [23]. From the existing research results, on the basis of low-light enhancement algorithms, low-light image has been effectively enhanced by learning adaptive prior knowledge to remove inappropriate features from the original image. However, the existing models still suffer from regional degradation, such as loss of details, low contrast and low color saturation, which prevents the whole image from achieving a satisfactory visual quality.

Aiming at the problems of detail loss, low contrast and low color saturation, a low-light image enhancement algorithm that incorporates the cross-mixed attention mechanism and the receptive field expansion mechanism(CE-Retinex) is proposed in this paper. This algorithm is divided into three modules, initial module, optimization module, and detail restoration. In the initial module, the image is decomposed into reflectance and illuminance using U-network and Retinex theory, the parameters are shared from the normal light image to the low light image. In the optimization module, a multi-scale illumination mechanism is used to adjust illumination and noise. A cross-hybrid attention mechanism is fused after convolution to remove negative interference from four aspects: channel, spatial, vertical and horizontal attention. In the detail restoration module, the receptive field expansion mechanism is a multi-branch structure consisting of convolutions of voids of different sizes. It is used for perceptual enhancement. The color is recovered using the color consistency loss function.

II. ANALYSIS OF THE CE-RETINEX ALGORITHM

Conventional low-light image enhancement algorithms mainly focus on brightness enhancement, which leads to problems such as loss of details and low color saturation in the enhanced image, and this results in poor image quality. To address this problem, the CE-Retinex algorithm is proposed in this paper, and the overall framework of this algorithm is shown in Figure 1. It could be seen that the framework of this algorithm consists of three main parts, which are the initial module, the optimization module, and the detail restoration module. Initial module, two different U-shaped networks are constructed using multilayer convolution and pooling. And the reflection and light images are normalized by sigmoid function. Normal light images and low light images are fed into the network at the same time and the network model learns the parameters of the normal light images. Secondly, in the optimization module, the illumination and denoising were optimized by using the convolutional fusion multi-scale illumination mechanism. After that the U-shaped network fuses the cross-mixed attention mechanism to filter the negative information of channel, spatial, vertical and horizontal attention. Finally, in the detail restoration module, feature maps of different sizes are adjusted using nearest neighbor interpolation and merged to be fed into the receptive field expansion mechanism. The receptive field expansion mechanism utilizes hollow convolution to enhance the perceptual field of the feature maps and strengthen the detail information. It makes the whole image structure more compatible with the comfort of human eyes.

A. INITIAL MODULE

The contrast and brightness of each part of the image is adjusted simultaneously in traditional low-light image enhancement algorithms, and this will lead to color distortion and overexposure problems in the enhanced image. The reflectance and light information are the key parameters to low-light image enhancement, among them, reflectance is

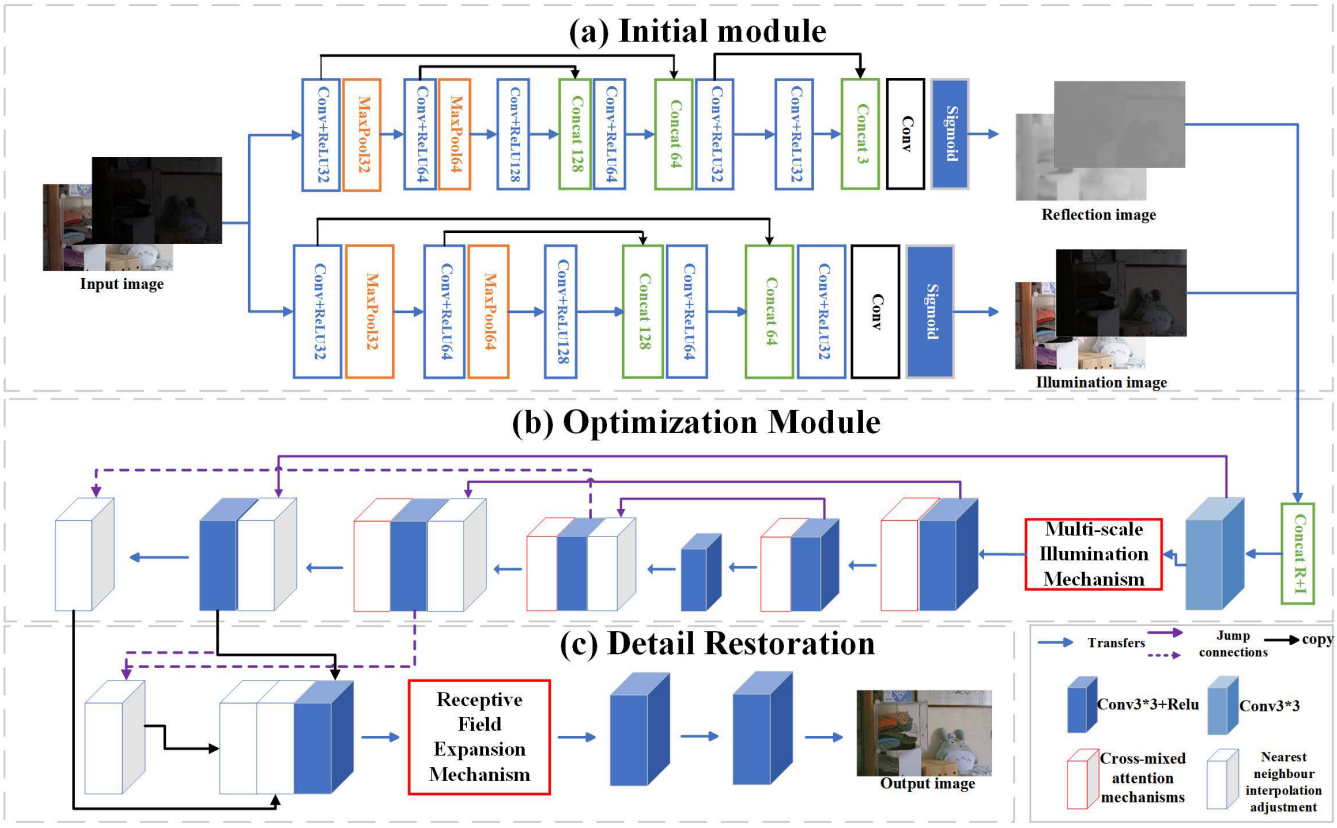


FIGURE 1. CE-Retinex overall framework diagram.

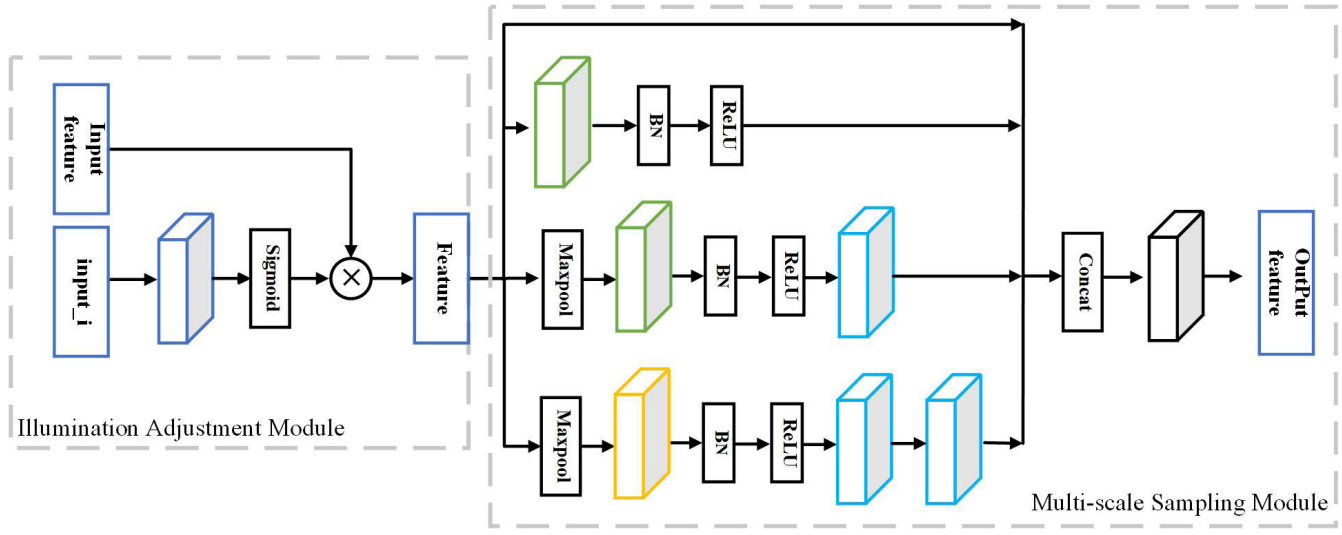


FIGURE 2. Structure of the multi-scale illumination mechanism(MIM).

determined by the surface properties of the object itself, and while the light information is the illumination from an external light source. Retinex theory simulates the way of perceiving the details and brightness of the picture in the human visual system, so the initial module of this paper simulates the Retinex theory by decomposing the image into illuminance and reflectance.Retinex theory is shown in equation (1). Where S is the input image, I is the illuminance, R is the reflectance and $\circ$ is the element level multiplication. The initial module structure is shown in Figure 1(a), it can be

seen from the figure the initial module obtains illuminance and reflectance information from the input normal light image. Illuminance is used to restore image brightness, and reflectance carries details and possible degradation.When a low-light image is input to the initial module, the low-light image learns the illuminance information and reflectance information of the normal-light image in the training dataset.

$$S = I \circ R \tag{1}$$

Both the U-network for reflection maps and the U-network for light maps consist of a 3×3 convolution, a ReLU function and maximum pooling. They extract feature maps with output channels of 32 and 64, respectively, and then they go through the bottleneck layer feature map with output channel of 128 to retain the important features. The feature map is upsampled using connect and inverse convolution to recover the image resolution. But in the reflection map U-network module, more detailed information is needed. So after the above convolution, a 3×3 convolution with a ReLU activation function is used for feature extraction. The final reflection map U-network is formed, and after that the sigmoid function is fed to constrain the reflection map and illumination map. Finally, the light and reflection maps are fused and fed into the optimization module. The initial module retains sufficient feature information and illumination information, which can help the optimization module and detail restoration module to remove negative interference and recover features from low-light images.

B. OPTIMIZATION MODULE

Conventional low-light enhancement algorithms have problems such as low contrast and negative information interference after enhancement of low-light images. To address these problems, the multi-scale illumination mechanism and the cross-mixed attention mechanism are fused in the optimization module. The structure of the optimization module is shown in Figure 1(b). First, A multi-scale illumination mechanism is designed with two parts, namely the illumination adjustment module and the multi-scale sampling module. In order to balance image lighting and smooth image noise. Then, cross-mixed attention mechanism is fused with U-shaped network. The 5-layer convolution of the U-shaped network is used to obtain feature maps. The 4-layer cross mixed attention mechanism is used to remove negative information from spatial, channel, horizontal, and vertical attention in feature maps. Finally, through nearest neighbor interpolation adjustment and convolution, the detailed feature map is input into the detail restoration module.

1) MULTI-SCALE ILLUMINATION MECHANISM(MIM)

Low-light image enhancement algorithms may have the problem of overexposure, halo and noise after enhancement, for this problem, the multi-scale illumination mechanism is proposed in this paper. And the structure of the multi-scale illumination mechanism is shown in Figure 2. It can be seen from the Figure 2 the multi-scale illumination mechanism consists of a illumination adjustment module and a multi-scale sampling module. First, the light map is fed to the network, in which the light information is adjusted and obtained through a 3×3 convolution and sigmoid function. The input feature maps are multiplied with the light information, thus directing the network to focus on the areas with the most severe light degradation. The multi-scale sampling module obtains feature maps from the light adjustment module, and the different feature information was obtained by using four branches. The first layer feature map is obtained from the illumination adjustment module. The second layer feature map is obtained by using 3×3 channel convolution, normalization and ReLU functions. The third layer feature map is obtained from the maximum pooling, 3×3 channel convolution, normalization, the ReLU function, and 2×2 channel inverse convolution. The fourth layer feature map is obtained from the maximum pooling, 1×1 channel convolution, normalization, ReLU function, and two 2×2 channel inverse convolution. The detailed channel information is obtained at the channel convolution. Finally, the four layers of feature maps are stacked. The output feature map is obtained by a 1×1 convolution. The multi-scale sampling module extracts rich features. The image illumination can be adjusted in the illumination adjustment module. Thereby, the multi-scale illumination mechanism can balance the illumination and remove uneven speckle and noise.

2) CROSS-MIXED ATTENTION MECHANISM(CAM)

In response to the effects of negative information generated by the feature fusion process, a cross-mixing attention mechanism is utilised to remove the negative interference from four aspects: channel, spatial, vertical and horizontal attention. The structure of the cross-mixed attention mechanism is shown in Figure 3. First, the channel attention mechanism calculates the size of the adaptive convolution kernel based on the number of channels in the input feature map. After pooling by global averaging, a one-dimensional convolution is utilized. The channel attention mechanism has no large number of convolutional structures. It achieves cross-channel information interaction with fewer number of parameters. The spatial attention mechanism deals with the spatial domain of the feature map, which is stacked by maximum pooling and average pooling. The spatial feature information is fused with 2D convolution, and it normalised using the sigmoid function. The channel attention mechanism multiply the input feature maps and the associated weights, in order to prevent feature details from being overlooked. The spatial attention mechanism operates similarly. The channel attention mechanism and the spatial attention mechanism merge feature maps to the row-column attention mechanism. The row-column attention mechanism focuses on feature information in both the vertical and horizontal directions.

First, the image mean of the feature map rows is computed by the reduce_mean function. The row attention is then computed using the 2D convolution and softmax function, and the row attention is applied to the columns. The column attention is the same operation. When the row and column attention maps are obtained separately, the row and column attention information is shared and stacked. The final output is a cross-mixed attention feature map.

C. DETAIL RESTORATION MODULE

For the problem of low detail and color saturation, the receptive field is enhanced using the receptive field expansion mechanism to strengthen the detail information. The color consistency loss function is used for color restoration. The structure of the detail restoration module is shown in Figure 1(c). First, rich feature information is obtained through stacking of nearest neighbor interpolation. Feature information of different sizes is acquired after the receptive field expansion mechanism, thus enhancing the feature details. Finally, after passing through two layers of 3×3 convolution and ReLU function, output the final enhanced image.

1) RECEPTIVE FIELD EXPANSION MECHANISM(RFEM)

The receptive field expansion mechanism is capable of receiving information from a wide range of regions. The structural diagram of the receptive field expansion mechanism is shown in Figure 4. The receptive field expansion mechanism consists of five branches. Four of branches utilize 1×1 convolution, 1×3 convolution and 3×1 convolution to obtain feature maps of different sizes, respectively. The expansion convolution with expansion rates of 1, 3 and 5 is utilized to expand the receptive field. Than, the four feature maps with different expansion rates are stacked, fused by 1×1 convolution, and finally summed with the front layer feature map after 1×1 convolution to output the detail feature map. Features at different scales and locations in the image are extracted by using different sizes and types of convolution operations as well as extended convolution. This helps the network to learn more feature representations at different scales and locations, which improves the performance of the model.

D. LOSS FUNCTION

Aiming at the problem that real data is needed to guide the decomposition of an image into reflectance and illuminance, and that low-light images need to learn normal light images, a modular loss function is designed in this paper. This modular loss function is divided into two main parts, the initial loss function and the optimization loss function. In this, the initial loss function is used to adjust the two components of reflectance and illuminance, in order to obtain the best reflectance and illuminance. In the optimization loss function, the low light image is optimized by learning the normal light image, and the colour details are recovered using the colour consistency loss function.

Equation (2) is the initial total loss function(L^d).From equation, it can be seen that the initial loss function consists of error loss $L_{\mathrm{ss}}^d = \|I_l - R_l \circ L_l\|_1 + \|I_h - R_h \circ L_h\|_1$,reflectance loss $L_{\mathrm{fs}}^d = \|R_l - R_h\|$, illumination loss $L_{\mathrm{gz}}^d = \left\|\frac{\nabla L_l}{\max(\nabla I_l, \mathcal{E})}\right\|_1 + \left\|\frac{\nabla L_h}{\max(\nabla I_h, \mathcal{E})}\right\|_1$ and consistency loss $L_{\mathrm{yz}}^d = \left\|(|\nabla L_l| + |\nabla L_h|) \circ \exp(-10 \cdot (|\nabla L_l| + |\nabla L_h|)\right\|_1$. Among them, $\|\|_1$ is the paradigm L1. R_l is the reflectivity of low-light image. R_h is the reflectivity of normal light image. I_l is the illumination of low-light image. I_h is the illumination of normal light image. $L_l$is the low-light image. L_h is the normal light image. $\overline{\nabla}$ is the first-order derivative operator in the horizontal and the vertical direction. $\mathcal{E}$ is 0.01, in order to avoid a zero denominator. | | find the absolute value.

$$L^d = L_{ss}^d + 0.01L_{fs}^d + 0.15L_{gz}^d + 0.2L_{yz}^d \tag{2}$$

where the error loss L_{ss}^d is used to constrain the reconstruction error that occurs during the decomposition of normal light images and low light images. Reflectance loss L_{fs}^d is used to regularize the similarity of reflectance. The illumination loss L_{gz}^d is used to constrain the smoothness of the illumination. Consistency loss L_{yz}^d is used to ensure the completeness of the image information.

The optimization loss function $\mathcal{L}^e$,as shown in equation (3). Consists of a mean square error loss function *MSE*, a structural similarity loss function *SSIM*, a gradient loss function *Grad* and a color consistency loss function *color*. R_l is the reflectance of a low-light image. R_h is the reflectance of a normal-light image.

$$\begin{aligned} L^e = {} & MSE(R_l, R_h) + 0.5SSIM(R_l, R_h) \\ & + 0.5Grad(R_l, R_h) + color \end{aligned} \tag{3}$$

From equation (3) it can be seen that the low light image learns the reflectance extracted from the normal light image, thus optimizes the low light image. Where the color consistency loss function *color*, as shown in equation (4). c is the number of color channels, I_{hc} is the illumination of the normal light image, R_{hc} is the reflectance of the normal light image.

$$color = \frac{1}{c}((\sum_{c=1}^{c} I_{hc} - R_{hc})^2 + \sum_{c=1}^{c}(\sqrt{I_{hc}} - \sqrt{R_{hc}})^2) \tag{4}$$

The color consistency loss function, which reprocesses the color information of a normal light image on the basis of the color information, achieves the recovery of color details.

III. EXPERIMENTS AND ANALYSIS OF RESULTS

A. DATA SETS AND EVALUATION INDICATORS

In order to verify whether the proposed algorithm is effective or not, the LOL dataset is selected for training in the experimental process, and the LOL, SCIE, LIME, MEF, and DICM datasets are used for testing, respectively. The LOL dataset consists of low-light and normal-light images taken in real-world environments. The SCIE dataset contains images

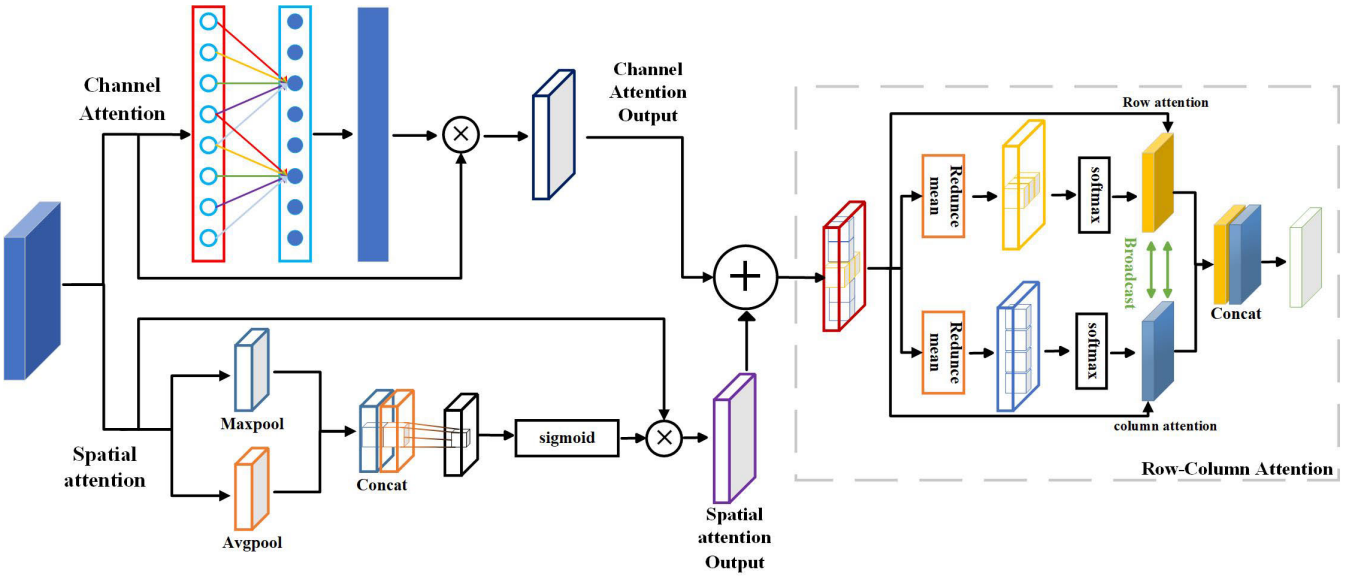


FIGURE 3. Structure of the Cross-mixed Attention Mechanism(CAM).

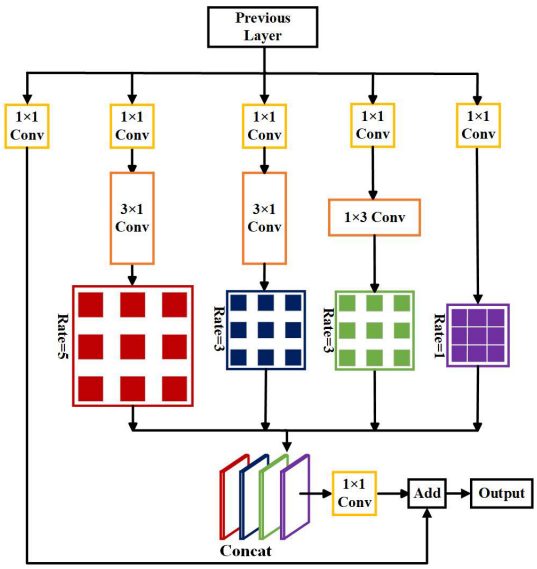


FIGURE 4. Structural diagram of the Receptive Field Expansion Mechanism(RFEM).

with different levels of exposure. The LIME dataset was taken in a nighttime, evening environment. The MEF dataset contains images taken indoors in low-light conditions and in overcast weather conditions. The DICM dataset contains images taken in the evening, at dawn, and under a variety of back-lighting conditions.

In order to verify the objective validity of the proposed algorithm, they are evaluated using referenced evaluation metrics: Peak Signal-to-Noise Ratio (PSNR), Structural Similarity Index (SSIM), Learned Perceptual Image Patch Similarity (LPIPS), and Root Mean Square Error (RMSE); and non-referenced evaluation metrics: Blind/Reference-less Image Spatial Quality Evaluator (BRISQUE) and Natural Image Quality Evaluator (NIQE).

PSNR is a commonly used image quality evaluation metric for comparing image noise. SSIM compares the quality of images in terms of brightness, contrast and structure. The LPIPS metric measures the similarity of images. RMSE is an objective metric based on pixel error. BRISQUE is a reference-free image quality assessment model that uses locally normalized luminance coefficients to obtain the corresponding parametric features. NIQE measures the difference between natural images over a multivariate distribution. Among them, the larger the value of PSNR and SSIM metrics and the smaller the value of LPIPS, RMSE, BRISQUE and NIQE metrics, the better the quality of the image is indicated.

B. EXPERIMENTAL SETTING

The environment used in the training experiments is NVIDIA GeForce RTX 3060, based on the TensorFlow-GPU2.4 framework, Python version 3.7, and Adam is used as the optimizer during model training. The specific parameters are shown in Table 1.

TABLE 1. Experimental parameter settings.

Parameter Name	Epoch	Batch size	Patch size	Learning rate
Parameter Value	300	16	48	0.001

C. ANALYSIS OF EXPERIMENTAL RESULTS

1) PERFORMANCE COMPARISON

In order to test the quantitative metrics of the algorithms in this paper, Table 2 lists the comparison of six metrics of ten different algorithms on the LOL dataset. Bolded are the best

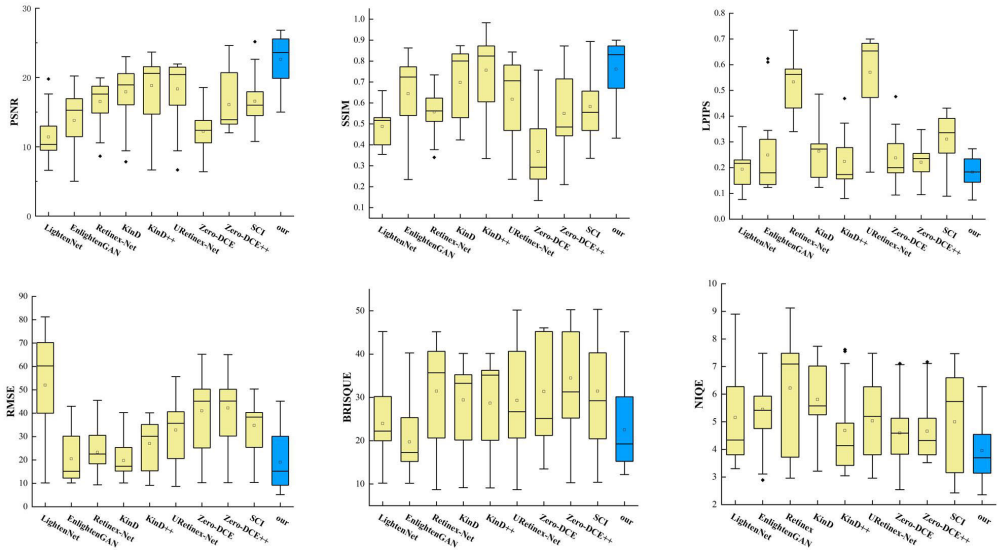


FIGURE 5. Comparison of average performance of different algorithms on LOL dataset.

results. In this paper, the PSNR value of the algorithm is 25.33, SSIM value is 0.83, LPIPS value is 0.128, RMSE value is 10.5731, BRISQUE value is 19.4487 and NIQE value is 3.37. The proposed algorithm achieves sub-optimal results in BRISQUE metrics but best results in PSNR, SSIM, LPIPS, RMSE and NIQE metrics. The table shows that the proposed algorithm achieves better results.

In order to comprehensively analyze the proposed algorithm on each evaluation metric, the box plots were utilized for average performance evaluation, and the comparison of average performance of different algorithms on LOL dataset was shown in Figure 5. It can be seen from the figure the first row shows the comparison of different algorithms on PSNR, SSIM and LPIPS evaluation metrics, and the second row shows the comparison of different algorithms on RMSE, BRISQUE and NIQE evaluation metrics. From the figure, it can be seen that the proposed algorithm is relatively centralized in terms of values and achieves better results.

In order to test the performance effect of the proposed algorithm on different datasets, this paper evaluates the performance of different algorithms on the public datasets DICM, LIME, SCIE, and MEF using the no-reference metrics BRISQUE and NIQE. As shown in Table 3, the bolded ones are the best results. The proposed algorithm, has a BRISQUE value of 10.1794 and NIQE value of 2.7692 on the DICM dataset; BRISQUE value of 19.6649 and NIQE value of 3.4159 on the LIME dataset; BRISQUE value of 10.4074 and NIQE value of 2.8309 on the SCIE dataset; BRISQUE value of 21.5537 and NIQE value of 2.7617 on the MEF dataset. It can be seen that the proposed algorithm BRISQUE metric achieves the best results on the DICM, SCIE dataset. The NIQE indicator achieved the best results on the DICM, SCIE, and MEF datasets. In summary, the proposed algorithm achieves better experimental results on several different datasets.

To test the effectiveness of each mechanism module of the proposed algorithm, ablation experiments are performed on the LOL dataset. As shown in Table 4, comparisons were made using the reference metric PSNR and the non-reference metric NIQE, with bolded being the best result. Add Receptive Field Expansion Mechanism(RFEM), Cross-mixed Attention Mechanism (CAM) and Multi-scale Illumination Mechanism(MIM) to the basic model of the algorithm one by one. In several experiments, we found that the various convolutional structures used in this paper perform best with the parameter combinations listed in Table 1. When the proposed algorithm was tested in the base module, the PSNR value was 18.81 and the NIQE value was 7.49. With the addition of the receptive field expansion mechanism, the algorithm in this paper expands the sensory field range, the PSNR value rises to 19.17, and the NIQE value decreases significantly to 3.85. Afterwards, the cross-mixed attention mechanism was added to remove negative information from the four dimensions of channel, spatial, vertical, and horizontal attention, and the PSNR rose to 22.91 and the NIQE fell to 3.64. Finally, by adding the multi-scale illumination mechanism to the above, the illumination of the proposed algorithm is corrected and the PSNR rises to 25.33 and the NIQE falls to 3.37. The results of the ablation experiments show that the individual mechanism modules of the CE-Retinex algorithm are effective.

2) QUALITATIVE COMPARISON

In order to verify the effectiveness and convergence of the algorithms in this paper, we conducted tests time and convergence on training. With the parameters in Table 1, the

TABLE 2. Metric values for different algorithms on the LOL dataset.

Methods	PSNR↑	SSIM↑	LPIPS↓	RMSE↓	BRISQUE↓	NIQE↓
LightenNet	9.88	0.51	0.2172	70.0184	22.907	3.86
EnlightenGAN	15.74	0.72	0.1603	12.054	**16.538**	5.39
Retinex-Net	18.81	0.56	0.6423	22.0625	37.7640	7.49
KinD	20.86	0.80	0.2879	17.6113	33.6675	5.14
KinD++	21.30	0.82	0.1584	27.7209	36.3942	3.88
URetinex-Net	19.72	0.77	0.6663	39.8628	26.4183	5.32
Zero-DCE	12.31	0.30	0.1806	43.6384	21.706	4.01
Zero-DCE++	13.97	0.48	0.1898	45.2415	31.6029	3.48
SCI	14.99	0.55	0.339	38.9246	29.5999	7.23
Our	**25.33**	**0.83**	**0.128**	**10.5731**	19.4487	**3.37**

TABLE 3. Comparison of BRISQUE and NIQE values for different algorithms and datasets.

Datasets	DICM		LIME		SCIE		MEF	
Methods	BRISQUE↓	NIQE↓	BRISQUE↓	NIQE↓	BRISQUE↓	NIQE↓	BRISQUE↓	NIQE↓
LightenNet	28.7627	3.6591	29.2530	3.2085	17.9913	3.2612	41.6791	5.0974
EnlightenGAN	18.848	3.118	19.634	5.649	21.1175	3.2905	20.304	3.421
Retinex-Net	37.0398	4.3257	43.3844	3.5095	17.9523	3.3379	28.3467	3.6571
KinD	30.1685	2.7771	18.9679	3.3857	18.7446	3.7426	45.8744	3.1761
KinD++	29.68	4.2636	**16.0635**	**2.9172**	17.2617	3.2914	29.2544	3.6379
URetinex-Net	47.6793	3.7959	26.3036	5.3161	24.7510	3.2785	37.154	5.4279
Zero-DCE	25.9088	4.0047	20.969	3.5482	16.6224	3.6004	47.0696	3.4585
Zero-DCE++	21.1623	3.8641	16.4807	5.3325	23.7083	4.8132	**12.3999**	3.5161
SCI	47.5223	3.1676	33.4984	3.1617	19.0887	3.7360	25.3836	3.5884
Our	**10.1794**	**2.7692**	19.6649	3.4159	**10.4074**	**2.8309**	21.5537	**2.7617**

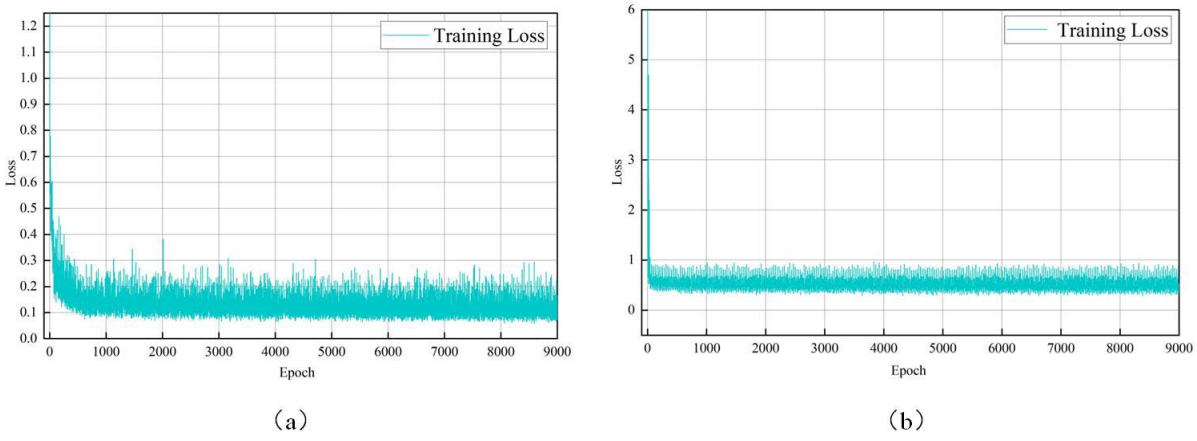


FIGURE 6. The convergence graph of the loss function.

TABLE 4. Analysis of ablation experiments.

Basic Module	RFEM	CAM	MIM	PSNR↑	NIQE↓
√				18.81	7.49
√	√			19.17	3.85
√	√	√		22.91	3.64
√	√	√	√	**25.33**	**3.37**

training time is about twenty minutes, which changes with parameter or other changes. Under many experiments, the relevant parameters in Table 1 make the algorithm of this paper work best. The convergence speed of the loss function is also a kind of important index to test the efficiency of the algorithm, such as Figure 6 shows the convergence graph of the loss function during the training of the CE-Retinex algorithm in this paper. The convergence curve of the loss function during the training of the initial module is shown in Figure 6(a), when the value of the loss function decreases from 1.2 to 0.3, a significant improvement can be observed, and the performance of the model continues to improve. The loss function convergence curves for the optimization and detail restoration modules are shown in Figure 6(b), where the loss function curves fluctuate within the range of 0.9 to 0.4, and the model maintains smooth operation. This gradually decreasing and smoothing loss function value reflects the fact that the model is constantly approaching the optimal solution during the optimization process, thus indicating the effectiveness and convergence of the algorithm.

FIGURE 7. Visual comparison chart.

In order to verify how the proposed algorithm is visually effective in practical applications, Figure 7 shows the visual comparison chart between the CE-Retinex algorithm of this paper and different low-light enhancement algorithms. The first row is the input low-light image. The second row is the LightenNet algorithm. LightenNet algorithm [24] takes the low light image as input and then estimates its light map. The third row is the EnlightenGAN algorithm. The EnlightenGAN algorithm [20] is based on generative adversarial networks, which mainly train generator and discriminator networks, where the generator network generates realistic low-light images and the discriminator network evaluates the realism of the generated images. The fourth row is the Retinex-Net algorithm [22]. The Retinex-Net algorithm makes use of Retinex theory to adjust low-light images. The fifth row is the KinD algorithm [25]. The KinD algorithm also utilises Retinex theory and is designed with three components: layer decomposition, reflectivity recovery and light adjustment. The sixth row is the KinD++ algorithm [26]. The KinD++ algorithm adds a multi-scale illumination attention mechanism to the KinD algorithm to mitigate some of the visual defects. The seventh row is the URetinex-Net algorithm [21]. The URetinex-Net algorithm adds new constraints and advanced network design to the Retinex-Net algorithm to obtain better performance. The eighth row is the Zero-DCE algorithm [27]. The Zero-DCE algorithm designs a pixel-by-pixel higher-order curve that can efficiently perform luminance mapping over a wide dynamic range through multiple iterations. The ninth row is the Zero-DCE++ algorithm [28]. The Zero-DCE++ algorithm is an update to the Zero-DCE algorithm, which is lighter and replaces the normal convolution with a depth-separable convolution. The tenth row is the SCI algorithm [29]. The SCI algorithm constructs self-calibrating light learning and utilises an unsupervised loss function to make the network structure simpler. The last row is the proposed algorithm in this paper. Among them, it can be seen from the locally enlarged portion of Figure 7 that the LightenNet algorithm suffers from low brightness and loss of detail compared to the algorithm in this paper. EnlightenGAN, Retinex-Net, KinD, KinD++, and URetinex-Net algorithms suffer from blurring and color distortion compared to the algorithms in this paper. Zero-DCE, Zero-DCE++, and SCI algorithms suffer from low brightness compared to the algorithm in this paper. In summary, the CE-Retinex algorithm in this paper obtains better visualization results.

To test the grayscale distribution of the CE-Retinex algorithm in this paper. At the same time, it is verified that the luminance enhancement effect of this paper's algorithm's modules such as the multi-scale illuminate mechanism is effective. A grayscale histogram is used to visualize it, as shown in Figure 8. The first column is the gray scale distribution of the low-light image. The second column is the gray scale distribution of the enhanced low-light image. The gray scale distribution of the low-light image in the first column is mainly concentrated in the darker parts, and the enhanced image in the second column has a more uniform gray scale distribution. As a result, the comparison of the grayscale histograms shows that the grayscale distribution of the proposed algorithm is more uniform and the brightness is well improved.

To test the color recovery of the CE-Retinex algorithm in this paper, the RGB color histogram is visualized using the RGB color histogram as shown in Figure 9. The first column is the RGB color distribution of the low-light image, the second column is the RGB color distribution of the enhanced low-light image. From the figure, it can be seen that the color distribution of the enhanced low-light image is more uniform and color recovery is achieved. Therefore, the algorithm in this paper restores the colour of low light images significantly with the help of modules such as colour loss function.

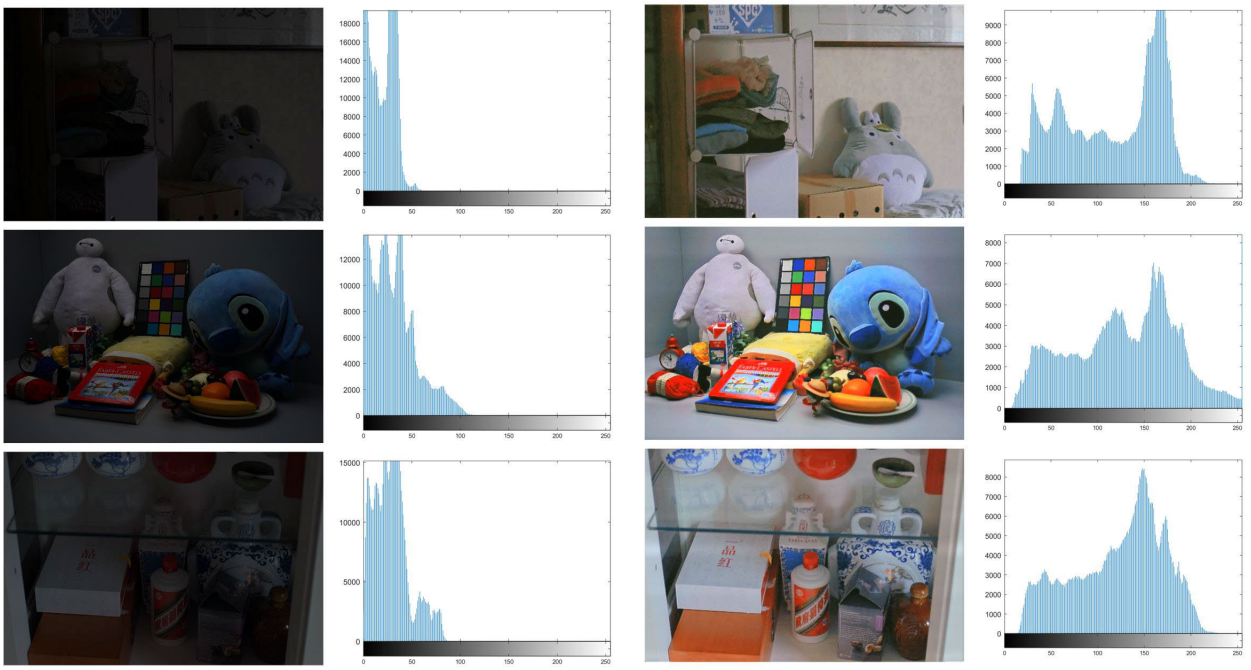


FIGURE 8. Comparison of grayscale histograms.

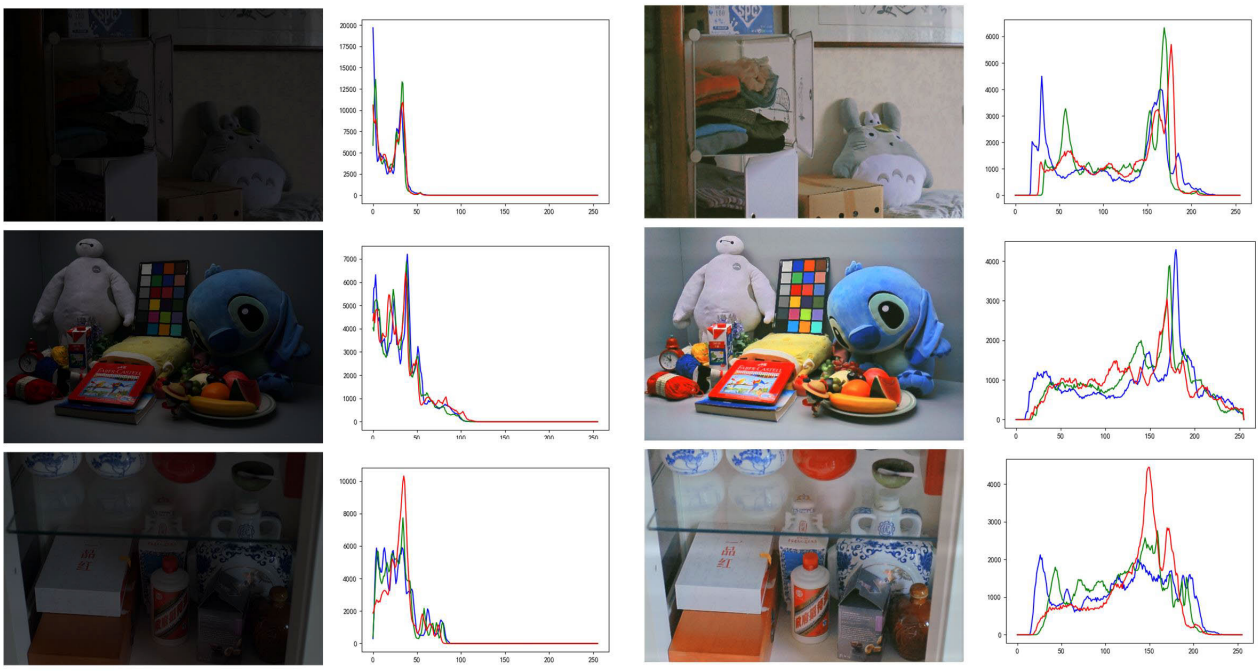


FIGURE 9. Comparison of RGB color histograms.

To verify the visual effectiveness of the CE-Retinex algorithm of this paper for noise removal. Meanwhile, it is verified that the detail enhancement effect of the cross-mixed attention mechanism and receptive field expansion mechanism of the algorithm in this paper is effective. Figure 10 shows a local zoom comparison. The first column is the local zoom effect after the enhancement of the base module, and the second column is the local zoom effect after the enhancement of the proposed algorithm. As can be seen from the localized zoom section, the image noise is significantly

FIGURE 10. Local zoom comparison.

removed after enhancement by the proposed algorithm, and the image is clearer and smoother.

IV. CONCLUSION

In order to improve the performance of the algorithm in low-light image enhancement and to address the problems of loss of detail, low contrast and low color saturation in low-light images, a low-light image enhancement algorithm that incorporates a cross-mixed attention mechanism and a receptive field expansion mechanism is proposed in this paper. The multi-scale illumination mechanism in the initial module solves the problem of low light and low contrast, in order to adapt to more application environments. A cross-mixed attention mechanism is utilized to remove negative information generated during feature fusion in four dimensions: channel, spatial, horizontal and vertical attention. To address the problem of low detail and colour saturation, the perceptual field is enhanced using the receptive field expansion mechanism to strengthen the detail information, and the colour loss function is used for colour recovery. The CE-Retinex algorithm was measured on the LOL dataset for objective evaluation metrics, with the PSNR of 25.33 and the NIQE of 3.37. The results show that the CE-Retinex algorithm has good results for low-light image enhancement. Although the algorithm in this paper has been improved in subjective vision and objective evaluation, how to make realize network self-supervision is still the main direction of future research.

DATA AVAILABILITY

The data used to support the findings of this study are available from the corresponding author upon request.

CONFLICTS OF INTEREST

There is no conflict of interest regarding the publication of this paper.

REFERENCES

[1] Y. Gonglin, H. Jing, and Y. Kuiying, "Night-time aerial image vehicle recognition technology based on transfer learning and image enhancement," *J. Comput.-Aided Des. Comput. Graph.*, vol. 31, no. 3, pp. 467–473, 2019.

[2] S. X. Zong, C. Q. Wang, and Y. J. Zhou, "Low-light aerial image enhancement method based on retinex and multi-attention mechanism," *Electron. Opt. Control*, vol. 30, no. 5, pp. 23–28, 2023.

[3] X. U. Fenglin, M. Yubin, and Z. Ming, "Navigation image enhancement based on color weighted guided image filtering-Retinex algorithm," *J. Shanghai Jiaotong Univ.*, vol. 53, no. 8, pp. 921–927, 2019.

[4] Z. H. Miao and L. J. Zhao, "Image enhancement of low-light parking space based on retinex," *Automot. Eng.*, vol. 45, no. 6, pp. 989–996, 2023.

[5] G. Cheng, C. Lang, and J. Han, "Holistic prototype activation for few-shot segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 45, no. 4, pp. 4650–4666, Apr. 2023.

[6] C. Lang, G. Cheng, B. Tu, C. Li, and J. Han, "Base and meta: A new perspective on few-shot segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 45, no. 9, pp. 10669–106866, Jul. 2023.

[7] C. Lang, J. Wang, G. Cheng, B. Tu, and J. Han, "Progressive parsing and commonality distillation for few-shot remote sensing segmentation," *IEEE Trans. Geosci. Remote Sens.*, vol. 61, 2023, Art. no. 5613610.

[8] C. Lang, G. Cheng, B. Tu, and J. Han, "Global rectification and decoupled registration for few-shot segmentation in remote sensing imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 61, 2023, Art. no. 5617211.

[9] Y. Wang, Q. Chen, and B. Zhang, "Image enhancement based on equal area dualistic sub-image histogram equalization method," *IEEE Trans. Consum. Electron.*, vol. 45, no. 1, pp. 68–75, Feb. 1999.

[10] E. P. Bennett and L. Mcmillan, "Video enhancement using per-pixel virtual exposures," *ACM Trans. Graph.*, vol. 24, no. 3, pp. 845–852, Jul. 2005.

[11] X. Ren, W. Yang, W.-H. Cheng, and J. Liu, "LR3M: Robust low-light enhancement via low-rank regularized retinex model," *IEEE Trans. Image Process.*, vol. 29, pp. 5862–5876, 2020.

[12] S. Hao, X. Han, Y. Guo, X. Xu, and M. Wang, "Low-light image enhancement with semi-decoupled decomposition," *IEEE Trans. Multimedia*, vol. 22, no. 12, pp. 3025–3038, Dec. 2020.

[13] R. Zhou, R. Wang, Y. Wang, F. Zhou, and N. Guo, "Research on low-light image enhancement based on MER-retinex algorithm," *Signal, Image Video Process.*, vol. 18, no. 1, pp. 803–811, Feb. 2024, doi: 10.1007/s11760-023-02801-x.

[14] L.-W. Wang, Z.-S. Liu, W.-C. Siu, and D. P. K. Lun, "Lightening network for low-light image enhancement," *IEEE Trans. Image Process.*, vol. 29, pp. 7984–7996, 2020.

[15] M. He, R. Wang, Y. Wang, F. Zhou, and N. Guo, "DMPH-net: A deep multi-scale pyramid hybrid network for low-light image enhancement with attention mechanism and noise reduction," *Signal, Image Video Process.*, vol. 17, no. 8, pp. 4533–4542, Nov. 2023, doi: 10.1007/s11760-023-02687-9.

[16] K. G. Lore, A. Akintayo, and S. Sarkar, "LLNet: A deep autoencoder approach to natural low-light image enhancement," *Pattern Recognit.*, vol. 61, pp. 650–662, Jan. 2017.

[17] F. Lv, F. Lu, and J. Wu, "MBLLEN: Low-light image/video enhancement using CNNs," in *Proc. Brit. Mach. Vis. Conf.*, vol. 220, 2018, p. 4.

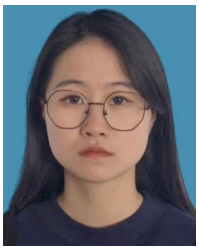
[18] R. Wang, Q. Zhang, C.-W. Fu, X. Shen, W.-S. Zheng, and J. Jia, "Underexposed photo enhancement using deep illumination estimation," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 6842–6850.

[19] K. Lu and L. Zhang, "TBEFN: A two-branch exposure-fusion network for low-light image enhancement," *IEEE Trans. Multimedia*, vol. 23, pp. 4093–4105, 2021.

[20] Y. Jiang, X. Gong, D. Liu, Y. Cheng, C. Fang, X. Shen, J. Yang, P. Zhou, and Z. Wang, "EnlightenGAN: Deep light enhancement without paired supervision," *IEEE Trans. Image Process.*, vol. 30, pp. 2340–2349, 2021.

[21] W. Wu, J. Weng, P. Zhang, X. Wang, W. Yang, and J. Jiang, "URetinex-net: Retinex-based deep unfolding network for low-light image enhancement," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2022, pp. 5891–5900.

[22] C. Wei, W. Wang, W. Yang, and J. Liu, "Deep retinex decomposition for low-light enhancement," 2018, *arXiv:1808.04560*.
[23] Z. Fu, Y. Yang, X. Tu, Y. Huang, X. Ding, and K.-K. Ma, "Learning a simple low-light image enhancer from paired low-light instances," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2023, pp. 22252–22261.
[24] C. Li, J. Guo, F. Porikli, and Y. Pang, "LightenNet: A convolutional neural network for weakly illuminated image enhancement," *Pattern Recognit. Lett.*, vol. 104, pp. 15–22, Mar. 2018.
[25] Y. Zhang, J. Zhang, and X. Guo, "Kindling the darkness: A practical low-light image enhancer," in *Proc. 27th ACM Int. Conf. Multimedia*, Oct. 2019, pp. 1632–1640.
[26] Y. Zhang, X. Guo, J. Ma, W. Liu, and J. Zhang, "Beyond brightening low-light images," *Int. J. Comput. Vis.*, vol. 129, no. 4, pp. 1013–1037, Apr. 2021.
[27] 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," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 1777–1786.
[28] C. Li, C. Guo, and C. C. Loy, "Learning to enhance low-light image via zero-reference deep curve estimation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 44, no. 8, pp. 4225–4238, Aug. 2022.
[29] L. Ma, T. Ma, R. Liu, X. Fan, and Z. Luo, "Toward fast, flexible, and robust low-light image enhancement," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2022, pp. 5627–5636.



RONGFENG ZHOU received the B.S. degree from Yancheng Institute of Technology, Yancheng, China, in 2022, where she is currently pursuing the M.Eng. degree. Her research interests include computer vision technology and intelligent control systems and signal detection.

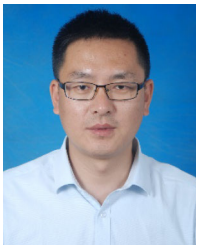
PING LI received the B.S. degree from Suzhou University, Suzhou, China, in 1991, and the M.S. degree from Nanjing University of Aeronautics and Astronautics, Nanjing, China, in 2002. She is currently an Associate Professor with the College of Information Engineering, Yancheng Institute of Technology, Yancheng, China. Her research interest includes web database application.

MINGYANG ZHANG, photograph and biography not available at the time of publication.

QINGYAO LIN, photograph and biography not available at the time of publication.

YUANYUAN WANG, photograph and biography not available at the time of publication.

XUESHENG BIAN, photograph and biography not available at the time of publication.



FENG ZHOU received the B.S. and M.S. degrees from Southeast University, Nanjing, China, in 2004 and 2012, respectively. He is currently pursuing the Ph.D. degree with the Army Engineering University of PLA. He is also a Professor with the College of Information Engineering, Yancheng Institute of Technology, Yancheng, China. His research interests include cooperative communication, computer vision technology, and image processing technology.

RUGANG WANG received the B.S. degree from Wuhan University of Technology, Wuhan, China, in 1999, the M.S. degree from Jinan University, Guangzhou, China, in 2007, and the Ph.D. degree from Nanjing University, Nanjing, China, in 2012. He is currently a Professor with the College of Information Engineering, Yancheng Institute of Technology, Yancheng, China. His research interests include image processing technology and optical communication networks.