

Article

Not peer-reviewed version

Global and Local Enhancement of Low-light Endoscopic Images

En Mou , [Huiqian Wang](#) ^{*} , Meifang Yang , Enling Cao , [Yuanyuan Chen](#) , [Chunlan Ran](#) , [Yu Pang](#) ^{*}

Posted Date: 30 November 2023

doi: 10.20944/preprints202311.1954.v1

Keywords: endoscopic image enhancement; dual attention; higher-order curve function



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Article

Global and Local Enhancement of Low-Light Endoscopic Images

En Mou^{1,2}, Huiqian Wang^{3,*}, Meifang Yang^{4,5}, Enling Cao⁶, Yuanyuan Chen^{1,7}, Chunlan Ran³ and Yu Pang^{3,*}

¹ School of Communications and Information Engineering, Chongqing University of Posts and Telecommunications, Chongqing 400065, China; en_mou@swmu.edu.cn, chenyuanyuan@zknmu.edu.cn

² School of Medical Information and Engineering, Southwest Medical University, Luzhou 646000, China; en_mou@swmu.edu.cn

³ School of Optoelectronic Engineering, Chongqing University of Posts and Telecommunications, Chongqing 400065, China; wanghq@cqupt.edu.cn, 2022210993@stu.cqupt.edu.cn, pangyu@cqupt.edu.cn

⁴ School of Nursing, Southwest Medical University, Luzhou 646000, China; yangmeifang@swmu.edu.cn

⁵ School of Nursing, Silliman University, Dumaguete 6267, Philippines; yangmeifang@swmu.edu.cn

⁶ School of Software Engineering, Chongqing University of Posts and Telecommunications, Chongqing 400065, China; 2020215130@stu.cqupt.edu.cn

⁷ College of Physics and Telecommunication Engineering, Zhoukou Normal University, Zhoukou, 466001, China; chenyuanyuan@zknmu.edu.cn

* Correspondence: wanghq@cqupt.edu.cn(H.W.); pangyu@cqupt.edu.cn(Y.P.); Tel.: +86-13752970863(H.W.); +86-13372653576(Y.P.)

Abstract: Endoscopic medical images can suffer from uneven illumination, low contrast, and lack of texture information due to the use of point directional light sources and the presence of narrow tissue structures, posing diagnostic difficulties for physicians. In this paper, a deep learning-based supervised illumination enhancement network is designed for low-light endoscopic images, aiming to improve both global illumination and local details. Initially, a global illumination enhancement module is formulated utilizing a higher-order curve function to improve global illumination. Secondly, a local feature extraction module incorporating dual attention is designed to capture local detailed features. Considering the significance of color fidelity in biomedical scenarios, the designed loss function prioritizes introducing color loss to alleviate image color distortion. Compared with seven state-of-the-art enhancement algorithms on Endo4IE endoscopic datasets, experimental results show that the proposed method can better enhance low-light endoscopic images and avoid image color distortion. It provides an efficient method to enhance images captured by endoscopes which can effectively assist clinical diagnosis and treatment.

Keywords: endoscopic image enhancement; dual attention; higher-order curve function

MSC: 68U10

1. Introduction

Endoscopic technology plays an important role in the diagnosis and treatment of diseases. However, the practical implementation of medical endoscopy is frequently impeded by low-light environments, stemming from the intricate physiological structures of internal organs and the utilization of point directional light sources. This impedes the ability of physicians to accurately identify and localize lesions or areas of pathology. Low-light image enhancement has emerged as an effective method to address endoscopic image quality issues, aiming to enhance the visibility and interpretability of the images.

Image enhancement methods typically fall into two main categories: traditional algorithms and deep learning approaches. Traditional image enhancement models include histogram equalization

(HE) [1–6] and the Retinex model [7–14]. Histogram equalization enhances contrast by adjusting the dynamic pixel range of an image to approach a uniform distribution. Since histogram equalization does not take into account the pixel relationships between images, it leads to the problem of information loss and over-enhancement of illumination. To solve this problem, Ibrahim et al. [2] proposed a histogram regionalization method that assigns a new dynamic range to each partition. The Contrast Limited Adaptive Histogram Equalization (CLAHE) [1] algorithm enhances image contrast by adaptively adjusting the histogram of local regions, preventing the over-amplification of noise and improving overall visibility. The Retinex method aims at presenting the image in a human-perceptible way, which assumes that the image consists of two parts: reflections and illumination. Typically, only illumination is considered and reflectance is treated as a constant. The approach centers on the lighting component, improving the perceived image quality by emphasizing the role of lighting in our visual experience. Guo et al. proposed the LIME algorithm [11] to enhance low-light images by estimating the illumination map. Tan et al. [15] decomposed the image into two layers: the detail layer and the base layer. The vascular information is extended through the channels in the detail layer, while adaptive light correction is applied to the base layer. In EIEN [16], the image is decomposed into light and reflection components, which are then processed separately. Finally, the reconstructed image is obtained by multiplying the enhanced optical and reflected components. Tanaka et al. [17] proposed a gradient-based low illumination image enhancement algorithm emphasizing the gradient of the enhancement of the dark region. Wang et al. [13] proposed an initial illumination weighting method to improve the illumination uniformity of an image by incorporating the inverse square law of illumination while maintaining exposure, chromatic aberration, and noise control. This method effectively improves the illumination and uniformity of endoscopic images from both visual perception and objective evaluation. Fang et al. [3] proposed a conventional algorithm to enhance the illumination of endoscopic images, which is based on a modified unsharpened mask and the CLAHE algorithm. Acharya et al. [4] presented an adaptive histogram equalization technique based on a genetic algorithm. The framework incorporates a genetic algorithm, histogram segmentation, and a modified probability density function. LR3M [18] considers noise generated during low-light image or video enhancement and applies two stages to enhance the image and suppress noise, respectively. These traditional algorithms provide the benefits of high reliability and interpretability. Nevertheless, they often involve manual feature selection in their physical models, and the effectiveness of enhancement results depends on the accuracy of the selected features.

Deep learning methods have made impressive and significant advances in enhancement results in recent years, attributed to their capacity to automatically extract features from a large dataset of images for low-light image enhancement. The pioneering LLNet [19] was the first deep learning network designed to enhance images captured in low-light natural scenes. Subsequently, numerous deep learning methods for enhancing image illumination have emerged [20–26]. Many researchers employed adversarial generative networks to generate synthetic datasets. Adversarial generative networks play an important role in synthesizing datasets to overcome data pairing issues, with commonly employed synthetic dataset generation methods like Pix2Pix [27] and CycleGan [28]. Zero-DCE [29] estimated the light intensity as a specific curve and designed a non-referenced loss function for deep training within a given image dynamic range, which is in line with the requirements of lightweight networks. FLW [30] designed a lightweight enhancement network with global and local feature extraction adjustments, proving effective for enhancing low-light images. While these algorithms have yielded satisfactory outcomes in enhancing natural images, their efficacy is constrained when applied to medical endoscopic images. The internal cavity environment of endoscopes exhibits weak texture characteristics due to non-Lambertian reflections from tissues, and the structural configuration of the internal cavity, coupled with the use of point light sources, leads to images displaying uneven light and darkness. Utilizing existing algorithms directly in such environments proves ineffective in enhancing image brightness in cavity depressions, and they fail to consider overall image brightness uniformity and overexposure, both critical for expanding the surgeon's field of view and executing surgical maneuvers. The existing network model introduces a

certain degree of smoothing effect on the detailed information of tissue structure in endoscopic images during the brightness enhancement process. However, the detailed information in weak texture images serves as a crucial basis for diagnosis and treatment, necessitating emphasis. Unlike natural images, endoscopic application environments demand strict color fidelity maintenance, and prevailing methods typically exhibit substantial color bias in such settings, rendering them unsuitable for direct application to scene brightness enhancement.

This paper specifically designs a low-light endoscopic image illumination enhancement network. The network comprises a decomposition module, a global illumination module, a local feature extraction module with dual attention, and a denoising module. The loss function accounts for color difference, structure, and illumination aspects. Experimental results on the Endo4IE [31] dataset demonstrate that the proposed method outperforms existing state-of-the-art methods in terms of Peak Signal-to-Noise Ratio (PSNR), Structural Similarity (SSIM), and Learned Perceptual Image Patch Similarity (LPIPS).

In summary, this approach contributes the following key elements:

- A novel network architecture is proposed for the global and local enhancement of low-light images in endoscopic environments. The network addresses the global brightness imbalance and the weak organizational texture commonly found in endoscopic images by integrating global illumination, local detail enhancement, and noise reduction, thereby achieving a balanced enhancement of brightness in endoscopic images;
- The global illumination enhancement Module mitigates the luminance inhomogeneity in endoscopic images resulting from the use of point light sources and the tissue structure environment. This is achieved by enhancing the overall image illumination perspective. Inspired by the Retinex methodology, the module extracts the overall image illumination through the decomposition of the model and optimizes the higher-order curve function using histogram information to automatically supplement the image luminance;
- Addressing the weak texture characteristics of endoscopic images, the local enhancement module incorporates a feature enhancement with a dual-attention mechanism. This mechanism enhances the local detailed feature expression of images by integrating curvilinear attention and spatial attention, effectively improving the detailed expression of the image organizational structure.

In this paper, Section 2 details the proposed image enhancement method, Section 3 covers the related experiments, Section 4 provides the conclusions, and Section 5 offers discussions.

2. Materials and Methods

We first present the general framework of the methodology, and then follow the details of the method and the composition of the corresponding modules.

2.1. Overall Pipeline

Figure 1 shows the comprehensive framework of the low-light image enhancement network. The global illumination enhancement module automatically improves light brightness by employing illumination separation and nonlinear enhancement, aiming at global brightness equalization and brightness enhancement of dark areas. The local detail enhancement module enhances organizational structure details by amplifying local dimensional interaction feature information through a dual attention mechanism. To ensure color realism, the loss function design prioritizes chromaticity loss. The proposed framework in this paper comprises two key components: the global illumination enhancement module and the local enhancement module. Given a low-light image $I_{low} \in R^{H \times W \times 3}$, the decomposition module is employed to decompose the low-light image to an illumination map $I \in R^{H \times W \times 1}$ and a reflectance map $R \in R^{H \times W \times 3}$. The decomposition process includes three convolutions with LeakyRelu functions and one convolution with the Relu function. Then, the illumination map obtained from decomposition is enhanced by using a global illumination

enhancement module with high order curve function. Afterward, the local enhancement module extracts detailed features by dual attention and removes the noise generated in the previous processes. The denoising process includes seven Conv2D layers with Relu activation functions and one convolution with Sigmoid activation function. The network enables low-light images to be enhanced to improve the illumination of dark areas and maintain color fidelity. It outperforms existing methods in terms of visual effects and performance metrics.

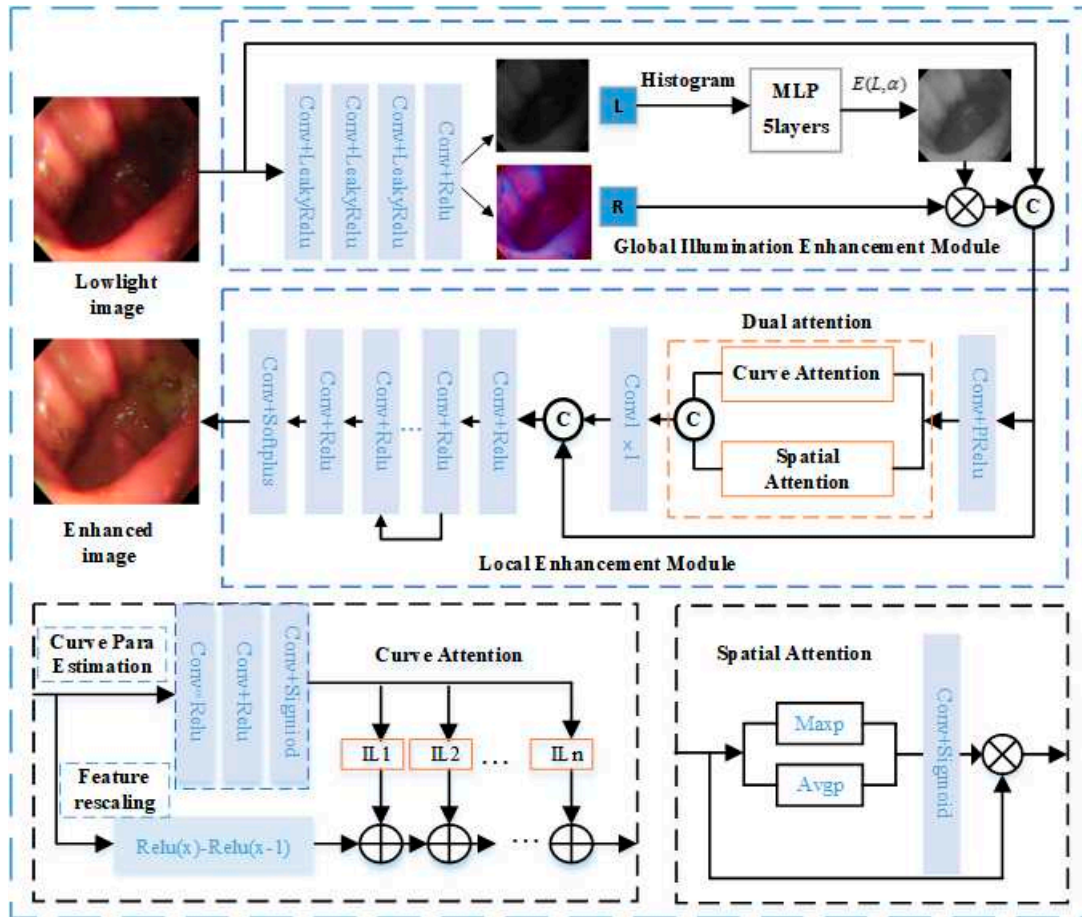


Figure 1. Framework Flowchart.

2.2. Global Illumination Enhancement

To enhance the global illumination, a global illumination enhancement module is designed in this paper to enhance the illumination map by the higher-order curve function [29]. According to the Retinex theory, a low-light image can be decomposed into an illumination map and a reflection map, the decomposition equation is as follows:

$$I = L \cdot R, \quad (1)$$

Where I is the low-light image, L is an illumination map, and R is a reflection map. Inspired by the Zero-DCE principle [29], a higher-order curve function operation is performed on the illumination map L to achieve global illumination enhancement. The parameters of the higher-order curve function are computed from the histogram of the illumination map and consist of a five-layer convolution and a LeakyReLU activation function. According to the literature [29], the equation representing this curve is as follows:

$$E(L(x), \alpha) = L(x) + \alpha L(x) - \alpha L^2(x), \quad (2)$$

where x represents the illumination of a pixel point, and $E(L(x), \alpha)$ signifies the enhanced illumination of $L(x)$. The training parameter α adjusts the curve's amplitude and controls the exposure level, with its value range limited to $[-1, 1]$.

The light-enhancement curve [29], as defined in Equation (2), can be applied by the following equation:

$$E_n(x) = E_{n-1}(x) + \alpha E_{n-1}(x) - \alpha E_{n-1}^2(x), \quad (3)$$

where the parameter n represents the number of iterations and controls the curvature of the curve. In our method, we set the value of n to 7. Equation (3) is equivalent to Equation (2) when n is equal to 1.

Equation (3) is applied globally to all pixels, representing a universal adjustment. However, this global mapping can sometimes overpower or underpower specific local regions. To solve this problem, parameters are introduced for each pixel. The modified expression of Equation (3) is given below:

$$E_n(x) = E_{n-1}(x) + \alpha_n(x)E_{n-1}(x) - \alpha_n(x)E_{n-1}^2(x), \quad (4)$$

The dimensionality of α_n matches that of the input image, ensuring adherence to the necessary conditions. α_n is obtained from the illumination map of the low-light input image by calculating its histogram, which is computed from the 5-layer MLP. The higher-order curves are normalized, monotonically increasing, and differentiable. Normalization ensures that each adjusted pixel remains in the range $[0, 1]$, thus preventing information loss due to potential overflow. Monotonicity preserves the distinction between neighboring pixels and ensures consistent contrast enhancement. Distinguish ability in turn facilitates efficient backpropagation during training. Then the enhanced illumination map L_{en} and the reflection map R are multiplied to obtain a global illumination enhanced image I_g . The equation is as follows:

$$I_g = \text{concat}([L_{en}, L_{en}, L_{en}]) \times R, \quad (5)$$

2.3. Dual Attention

The image is adjusted with global and local spatial attention. To better extract the detailed contrast of the image, a dual attention mechanism is adopted for local and spatial feature extraction. Dual attention includes curve attention and spatial attention, which operate parallel. Curve attention proposed in the literature [32] can effectively extract the local feature, the equation is as follows:

$$\frac{IL_{n(c)}}{IL_{n-1(c)}} = \text{Curve}_{n-1}(1 - IL_{n-1(c)}), \quad (6)$$

Where $IL_{n(c)}$ is the curve function, and c represents the feature location coordinates. It is estimated by three Conv2D layers and Sigmoid activation. The equation is as follows:

$$CA(I_D) = IL_{n(c)}(I_D), \quad (7)$$

$CA(I_D)$ denotes that curve attention is applied to image I_D to extract local features to extract local features.

Spatial attention adopted global average pooling and max pooling on input features respectively to get the interspatial dependencies of convolutional features [33], which are then concatenated for convolution and sigmoid operations. The equation is as follows:

$$SPA(I_D) = I_D \otimes \text{Sig}(\text{Conv}(\text{concat}(\text{Max}(I_D), \text{Avg}(I_D))))), \quad (8)$$

$SPA(I_D)$ denotes the computation of spatial attention on I_D to extract the global features of the image. $CA(I_D)$ is concatenated with $SPA(I_D)$ followed by a 1×1 Conv2D layer, the equation is as follows:

$$I_{dual} = Conv_{1 \times 1} \{concat(CA(I_D), SPA(I_D))\}, \quad (9)$$

2.4. Total Loss

Apart from the commonly employed pixel absolute and structural losses, a hue and saturation loss function, an illumination loss function, and a structural loss basis for FLW [30] are applied, and a color difference loss is employed to maintain the color fidelity of the image. The total loss is calculated as a combination of different loss functions for endoscopic low-light image enhancement, the total loss function is as follows:

$$l_{total} = l_1 + l_{SSIM} + l_{hs} + l_b + l_{str} + l_{cd}, \quad (10)$$

Given its alignment with human visual perception, the CIELAB color space enables more effective comparison and measurement of color differences. So, in this paper, we convert the sRGB space to the CIELAB color space to more accurately assess the color difference between the enhanced image and the reference image. The color difference equation l_{cd} is as follows:

$$l_{cd} = dl + dc / ((1 + 0.045c_1)^2) + dh / ((1 + 0.015c_1)^2), \quad (11)$$

Where dl , dc , dh , c_1 denote the square of the L channel difference, the square of the C channel difference, the chrominance channel difference, and the mean value of the chrominance channel about the enhanced image respectively.

l_1 represents the pixel-level loss between the enhanced image and the reference image, the equation is as follows:

$$l_1 = \frac{1}{mn} \sum_{i=1, j=1}^{m, n} |I_{en}(i, j) - I_{ref}(i, j)|, \quad (12)$$

l_{SSIM} is the SSIM loss, the equation is as follows:

$$l_{SSIM} = 1 - SSIM(I_{en}, I_{ref}), \quad (13)$$

The structural similarity [34] function is as follows:

$$SSIM(I_{en}, I_{ref}) = \frac{2\mu_1\mu_2 + c_1}{\mu_1^2 + \mu_2^2 + c_1} \cdot \frac{2\sigma_1\sigma_2 + c_2}{\sigma_1^2 + \sigma_2^2 + c_2}, \quad (14)$$

l_{hs} measures the hue and saturation difference between two pixels, the equation is as follows:

$$l_{hs} = 1 - \sum_{i=1, j=1}^{m, n} \langle I_{en}(i, j), I_{ref}(i, j) \rangle, \quad (15)$$

where $\langle \cdot, \cdot \rangle$ represents the cosine similarity of two vectors.

l_b express the illumination relation, the equation is as follows:

$$l_b = 1 - \sum_{c \in \{R, G, B\}} \sum_{i=1, j=1}^{m, n} \langle I_{en}^c(i, j) - \min I_{en}^c(i, j), I_{ref}^c(i, j) - \min I_{ref}^c(i, j) \rangle, \quad (16)$$

l_{str} is the gradient loss, which expresses structure information. l_{str} includes the sum of the horizontal and vertical losses. The horizontal and vertical gradients of the enhanced image and the reference image in the R , G , and B channels are calculated by subtracting the cosine similarity by 1 to calculate the gradient consistency loss. The equation is as follows:

$$l_{str} = 1 - \sum_{c \in \{R, G, B\}} \sum_{i=1}^{m,n} \left\langle \nabla I_{en}^c(i, j) - \min \nabla I_{en}^c(i, j), \nabla I_{ref}^c - \min \nabla I_{ref}^c(i, j) \right\rangle, \quad (17)$$

3. Results

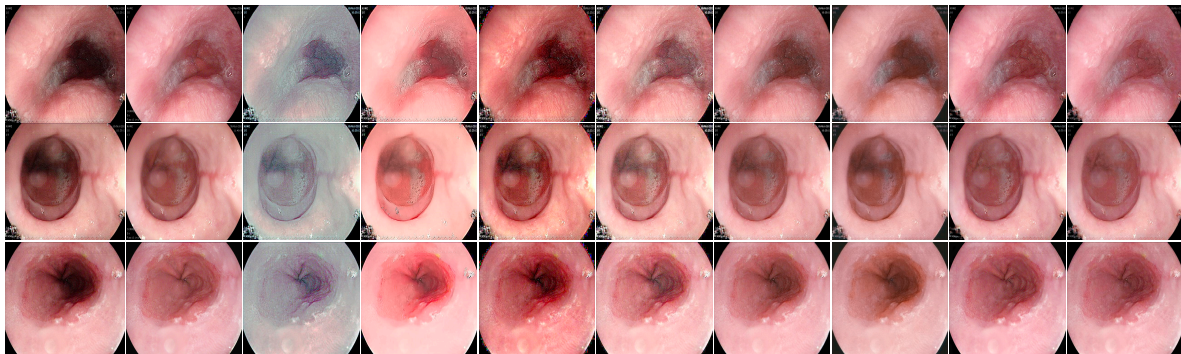
3.1. Implementation Details

The Endo4IE [31] dataset is employed for training and testing. This dataset is developed from the EAD2020 Challenge dataset [26], EAD 2.0 from the EndoCV2022 Challenge dataset [35], and the HyperKvasir dataset [36], where the exposure images are synthesized by CycleGAN after the removal of non-informative frames to obtain the paired data. Each image in the dataset has dimensions of 512×512 . This dataset consists of 690 pairs of low-light and reference-light images for training, 29 pairs for validation, and 266 pairs for testing. The proposed method is quantitatively compared with several benchmark experiments: PSNR and SSIM.

All experiments are implemented using PyTorch in Python on a system with an Intel i9-12900KF 3.20 GHz CPU, NVIDIA RTX 3090Ti GPU, and 32 GB RAM. During training, the batch size was set to 100, the Adam optimizer was utilized in the model, employing a learning rate of 0.0001. Visual comparisons with existing methods show that while these methods enhance the illumination of the image, they tend to fall short of preserving image details. Quantitative evaluation, utilizing PSNR, SSIM, and LPIPS [37] metrics, demonstrates superior performance of the proposed method compared to the state-of-the-art methods.

3.2. Enhancement Results

Figure 2 shows visual comparisons of results using the Endo4IE dataset. The methods compared all enhance the illumination of the medical endoscopic image, but in the recessed area in the center of the cavity, the illumination enhancement is insufficient, texture details are not demonstrated, and color is distorted. The image generated by the Zero-DCE method has an overall grey appearance with insufficient enhancement of the central black areas. The images tested with the LIME and LR3M methods tend to be reddish, with improved brightness, but with noise spots in the edge portion of the image, as in the fourth row of images in Figure 2. The image enhancement results of Ying's and FLW methods are better than those of LR3M and LIME, but there is still a slight deficiency in terms of color maintenance when compared to the reference image. The images tested by EFINet show the same phenomenon as the two algorithms just described, with darker illumination in the central area. The method proposed in this paper obtains better results on the same training dataset, achieving consistency in illumination enhancement in the central recessed region, better resolution, and higher color retention. MBPNet gives better results than the previous methods, with detail in darker areas needing to be enhanced. From the fourth row of Figure 2, it can be seen that the edge part of the illumination can still be improved to get a more detailed texture. A comparison of these methods with the proposed in this paper shows that the proposed method not only enhances the illumination effect well but also enhances the detail features and maintains color consistency.



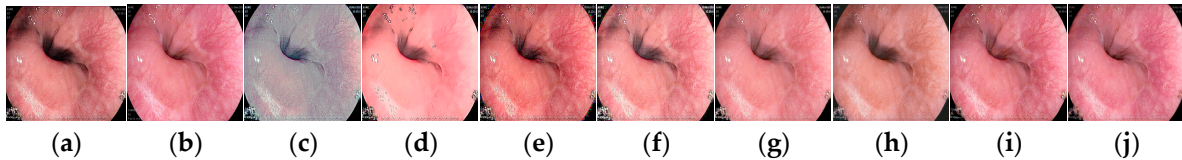


Figure 2. Visual comparison results on Endo4IE [31]. (a) The lowlight image; (b) The reference image; (c) Zero-DCE; (d) LR3M; (e) LIME; (f) Ying's; (g) FLW; (h) EFINet; (i) MBPNet; (j) ours.

Table 1 provides the quantitative comparison results, where the method of this paper shows better results in PSNR, SSIM, and LPIPS values with seven methods including MBPNet and EFINet trained on the Endo4IE [31] dataset. As indicated in Table 1, the PSNR, SSIM, and LPIPS values for the method presented in this paper are 27.3245 dB, 0.8349, and 0.1472, respectively. In the conducted quantitative experiments, the PSNR values of our proposed method surpass those of Zero-DCE, FLW, EFINet, and MBPNet by 12.8459 dB, 1.717 dB, 4.1963 dB, and 0.0915 dB, respectively. Similarly, SSIM metrics demonstrate improvements of 0.1579, 0.0153, 0.0612, and 0.028 over the corresponding values for Zero-DCE, FLW, EFINet, and MBPNet, respectively. Additionally, LPIPS scores are 0.154, 0.0014, 0.038, and 0.0211 lower than those of Zero-DCE, FLW, EFINet, and MBPNet methods, respectively. The results strongly indicate that the enhanced images produced through the methodology elucidated in this paper exhibit reduced disparities and superior structural resemblance when compared with the reference images. Both qualitative and quantitative experiments validate the efficacy of the proposed method in enhancing low-light endoscopic images.

Table 1. Quantitative comparison results on Endo4IE [31]. The best results are shown in bold.

Method	Published	PSNR(dB)↑	SSIM↑	LPIPS↓
LR3M [18]	TIP(2020)	14.8124	0.7007	0.2634
LIME [11]	TIP(2016)	19.9308	0.7046	0.2279
Ying's [38]	ICCV(2017)	18.0976	0.7149	0.1821
Zero-DCE [29]	CVPR (2020)	15.0663	0.6796	0.3012
EFINet [39]	TCSVT(2022)	23.1282	0.7737	0.1852
FLW [30]	ArXiv(2023)	25.6075	0.8196	0.1486
MBPNet [40]	TIP(2023)	27.2330	0.8069	0.1683
ours		27.3245	0.8349	0.1472

3.3. Ablation Study

The ablation study demonstrated the individual contribution of each component in the network. We designed a new network including an image decomposition module, a global illumination enhancement module, a local feature extraction module, and a denoising module. A color difference loss function is employed for the total loss function. Therefore, we performed ablation experiments on each module using the dataset [31] to demonstrate the effectiveness of each module, the quantitative comparisons are shown in Table 2.

Table 2. Results of quantitative comparisons about different components of the model on training. The best results are shown in bold.

Sequence	Global Illumination Enhancement Module	Local Enhancement Module	PSNR(dB)↑	SSIM↑	LPIPS↓
1		√	26.6236	0.8311	0.1489
2	√		26.8364	0.8282	0.1566
3	√	√	27.3245	0.8349	0.1472

Ablation experiments were conducted to assess the individual contributions of various components of the proposed method. The quantitative metric values PSNR, SSIM, and LPIPS reflect the contribution of each module in Table 2. Changes in PSNR and LPIPS metrics in particular are

evident in Table 2. Sequence 1 removes the global illumination enhancement module, and the PSNR value obtained is 26.6236 dB, which is 0.7009 dB less than the whole algorithm, verifying the role of the global illumination enhancement module. Removing the global illumination enhancement module results in the LPIPS metric that is 0.0017 higher than the LPIPS of the proposed algorithm. Similarly, removing the local feature enhancement network results in an LPIPS that is 0.0094 higher than that of the entire network framework. Applying either the global illumination enhancement module or the local enhancement module individually resulted in a decrease in the PSNR of the test image. This decline indicates that a larger disparity between the processed image and the reference image corresponds to a more significant deviation in image quality from the reference standard. Furthermore, higher LPIPS values were observed when testing each module individually compared to the proposed network. This suggests an increased perceived difference between the processed image and the reference image, accompanied by a corresponding decrease in overall image quality. These ablation studies highlight the essential role played by individual modules within the network structure, underscoring their significance and interpretability in enhancing low-light endoscopic images.

4. Discussion

We provide an efficient and effective global and local enhancement method for low-light endoscopic images. The proposed method models endoscopic low-light images to effectively enhance the illumination and details of endoscopic images. As illustrated in Table 1 and Figure 2, our model achieves more efficient image enhancement compared to existing methods. Ablation experiments indicate that the superior performance is attributed to the global illumination enhancement module with the higher-order curve and the local feature extraction module with dual attention. In the local enhancement module, the dual attention mechanism is followed by an effective denoising process, which eliminates and suppresses noise from the previous step. To reduce the color bias, a color difference loss function is introduced during the training process to prevent undesired color shifts.

5. Conclusions

This paper introduces a novel model for enhancing low-light endoscopic images through global illumination and local detail enhancement. Inspired by the Retinex traditional method, the model is designed to separate illuminance maps and iteratively optimize higher-order curve functions using histogram information for balanced global illuminance enhancement. To address the importance of detail organization structure, a local detail enhancement module with a dual attention mechanism is integrated to amplify interaction features in local dimensions. The training process includes a chromatic aberration loss function to preserve the color realism of the medical scene. Comprehensive experiments demonstrate the quantitative and qualitative superiority of the proposed method over existing approaches, providing an effective solution for low-light image enhancement in endoscopic scenes. Future work is planned to hardwareize the network, enhancing real-time feasibility through hardware implementation. Algorithm optimization will involve integrating endoscopic luminance enhancement as a preprocessing step for subsequent tasks such as depth prediction and target recognition in endoscopy.

Author Contributions: Conceptualization, En Mou. and Huiqian Wang.; methodology, En Mou.; software, En Mou. and Enling Cao.; validation, En Mou., Huiqian Wang.; formal analysis, En Mou.; investigation, En Mou. and Meifang Yang.; resources, En Mou. and Huiqian Wang.; data curation, En Mou.; writing—original draft preparation, En Mou., Meifang Yang., and Yuanyuan Chen.; writing—review and editing, En Mou. and Chunlan Ran; visualization, En Mou. and Yu Pang.; supervision, Huiqian Wang.; project administration, Huiqian Wang.; funding acquisition, Huiqian Wang. and Yu Pang. All authors have read and agreed to the published version of the manuscript.

Funding: The project was funded by the National Natural Science Foundation of China under Grant No. U21A20447, the Science and Technology Research Program of Chongqing Municipal Education Commission under Grant No.KJQN202100602, Project funded by China Postdoctoral Science Foundation under Grant No.2022MD713702, Special Postdoctoral Support from Chongqing Municipal People's Social Security Bureau

under Grant No.2021XM3010, Chongqing Technical Innovation and Application Development Special Project under Grant No.CSTB2022TIAD-KPX0062.

Data Availability Statement: Endo4IE was employed for the research described in the article. Data available at: <https://data.mendeley.com/datasets/3j3tmghw33/1>.

Acknowledgments: We thank the School of Communications and Information Engineering and the School of Optoelectronic Engineering of Chongqing University of Posts and Telecommunications for their assistance in the research.

Conflicts of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Pizer, S.M.; Johnston, R.E.; Ericksen, J.P.; Yankaskas, B.C.; Muller, K.E. Contrast-Limited Adaptive Histogram Equalization: Speed and Effectiveness. In Proceedings of the First Conference on Visualization in Biomedical Computing, Atlanta, Georgia, USA, 22-25 May 1990.
2. Ibrahim, H.; Kong, N.S.P. Brightness Preserving Dynamic Histogram Equalization for Image Contrast Enhancement. *IEEE Trans. Consum Electr*, **2007**, *53*, 1752-1758. <https://doi.org/10.1109/TCE.2007.4429280>.
3. Fang, S.; Xu, C.; Feng, B.; Zhu, Y. Color Endoscopic Image Enhancement Technology Based on Nonlinear Unsharp Mask and CLAHE. In Proceedings of the 2021 IEEE 6th International Conference on Signal and Image Processing (ICSIP), Beijing, China, 9-11 July 2021. <https://doi.org/10.1109/ICSIP52628.2021.9688796>.
4. Acharya, U.K.; Kumar, S. Genetic algorithm based adaptive histogram equalization (GAAHE) technique for medical image enhancement. *Optik*, **2021**, *230*, 166273. <https://doi.org/10.1016/j.ijleo.2021.166273>.
5. Rundo, L.; Tangherloni, A.; Nobile, M.S.; Militello, C. MedGA: a novel evolutionary method for image enhancement in medical imaging systems. *Expert Syst Appl*, **2019**, *119*, 387-399. <https://doi.org/10.1016/j.eswa.2018.11.013>.
6. Lu, L.; Zhou, Y.; Panetta, K.; Agaian, S. Comparative study of histogram equalization algorithms for image enhancement. *Mobile Multimedia/Image Processing, Security, and Applications*, **2010**, *7708*, 337-347. <https://doi.org/10.1117/12.853502>.
7. Rahman, Z.; Jobson, D.J.; Woodell, G.A. Multi-scale retinex for color image enhancement. In Proceedings of 3rd IEEE International Conference on Image Processing, Lausanne, Switzerland, 19-19 September 1996. <https://doi.org/10.1109/ICIP.1996.560995>.
8. Wang, S.; Zheng, J.; Hu, H.M.; Li, B. Naturalness preserved enhancement algorithm for non-uniform illumination images. *IEEE Trans. Image Process*, **2013**, *22*, 3538-3548. <https://doi.org/10.1109/TIP.2013.2261309>.
9. Fu, X.; Zeng, D.; Huang, Y.; Zhang, X.P.; Ding, X. A weighted variational model for simultaneous reflectance and illumination estimation. In Proceedings of the IEEE conference on computer vision and pattern recognition, Las Vegas, USA, 5 July 2016.
10. Fu, X.; Zeng, D.; Huang, Y.; Liao, Y.; Ding, X.; Paisley, J. A fusion-based enhancing method for weakly illuminated images. *Signal Process*, **2016**, *129*, 82-96. <https://doi.org/10.1016/j.sigpro.2016.05.031>
11. Guo, X.; Li, Y.; Ling, H. LIME: Low-light image enhancement via illumination map estimation. *IEEE Trans. Image Process*, **2016**, *26*, 982-993. <https://doi.org/10.1109/TIP.2016.2639450>.
12. Wei, C.; Wang, W.; Yang, W.; Liu, J. Deep retinex decomposition for low-light enhancement. *arXiv* **2018**, arXiv:1808.04560. <https://doi.org/10.48550/arXiv.1808.04560>.
13. Wang, L.; Wu, B.; Wang, X.; Zhu, Q.; Xu, K. Endoscopic image illumination enhancement based on the inverse square law for illumination and retinex. *Int J Med Robot*, **2022**, *18*, e2396. <https://doi.org/10.1002/rcs.2396>.
14. Xia, W.; Chen, E.C.S.; Peters, T. Endoscopic image enhancement with noise suppression. *Healthc Technol Lett*, **2018**, *5*, 154-157. <https://doi.org/10.1049/htl.2018.5067>.
15. Tan, W.; Xu, C.; Lei, F.; Fang, Q.; An, Z.; Wang, D.; Han, J.; Qian, K.; Feng, B. An Endoscope Image Enhancement Algorithm Based on Image Decomposition. *Electronics*, **2022**, *11*, 1909. <https://doi.org/10.3390/electronics11121909>.
16. An, Z.; Xu, C.; Qian, K.; Han, J.; Tan, W.; Wang, D.; Fang, Q. EIEN: endoscopic image enhancement network based on retinex theory. *Sensors*, **2022**, *22*, 5464. <https://doi.org/10.3390/s22145464>.
17. Tanaka, M.; Shibata, T.; Okutomi, M. Gradient-based low-light image enhancement. In Proceedings of the 2019 IEEE International Conference on Consumer Electronics (ICCE), Las Vegas, NV, USA, 11-13 January 2019. <https://doi.org/10.1109/ICCE.2019.8662059>.
18. Ren, X.; Yang, W.; Cheng, W.H.; Liu, J. LR3M: Robust low-light enhancement via low-rank regularized retinex model. *IEEE Trans. Image Process*, **2020**, *29*, 5862-5876. <https://doi.org/10.1109/TIP.2020.2984098>.
19. Lore, K.G.; Akintayo, A.; Sarkar, S. LLNet: A deep autoencoder approach to natural low-light image enhancement, *Pattern Recognit.*, **2017**, *61*, 650-662. <https://doi.org/10.1016/j.patcog.2016.06.008>.

20. Shen, L.; Yue, Z.; Feng, F.; Chen, Q.; Liu, S.; Ma, J. Msr-net: Low-light image enhancement using deep convolutional network. *arXiv* **2017**, arXiv: 1711.02488. <https://doi.org/10.48550/arXiv.1711.02488>.
21. Jiang, Y.; Gong, X.; Liu, D.; Cheng, Y.; Fang, C.; Shen, X.; Yang, J.; Zhou, P.; Wang, Z. Enlightengan: Deep light enhancement without paired supervision[J]. *IEEE Trans. Image Process*, **2021**, *30*, 2340-2349. <https://doi.org/10.1109/TIP.2021.3051462>.
22. Ren, W.; Liu, S.; Ma, L.; Xu, Q.; Xu, X.; Cao, X.; Du, J.; Yang, M.H. Low-light image enhancement via a deep hybrid network. *IEEE Trans. Image Process*, **2019**, *28*, 4364-4375. <https://doi.org/10.1109/TIP.2019.2910412>.
23. Huang, D.; Liu, J.; Zhou, S.; Tang, W. Deep unsupervised endoscopic image enhancement based on multi-image fusion. *Computer Methods and Programs in Biomedicine, Comput Methods Programs Biomed*, **2022**, *221*, 106800. <https://doi.org/10.1016/j.cmpb.2022.106800>.
24. Ma, Y.; Liu, Y.; Cheng, J.; Zheng, Y.L.; Ghahremani, M.; Chen, H.L.; Liu, J.; Zhao, Y.T. Cycle structure and illumination constrained GAN for medical image enhancement. In Proceedings of the Medical Image Computing and Computer Assisted Intervention–MICCAI 2020: 23rd International Conference, Lima, Peru, 4–8 October 2020. https://doi.org/10.1007/978-3-030-59713-9_64.
25. Ma, Y.; Liu, J.; Liu, Y.; Fu, H.; Hu, Y.; Cheng, J.; Qi, H.; Wu, Y.; Zhang, J.; Zhao, Y. Structure and illumination constrained GAN for medical image enhancement. *IEEE Trans. Med Imaging*, **2021**, *40*: 3955-3967. <https://doi.org/10.1109/TMI.2021.3101937>.
26. Ali, S.; Dmitrieva, M.; Ghatwary, N.; Bano, S.; Polat, G.; Temizel, A.; Krenzer, A.; Hekalo, A.; Guo, Y.B.; Matuszewski, B.; Rittscher, J. A translational pathway of deep learning methods in gastrointestinal endoscopy. *arXiv* **2020**, arXiv: 2010.06034.
27. Isola, P.; Zhu, J.Y.; Zhou, T.; Efros, A.A. Image-to-image translation with conditional adversarial networks. In Proceedings of the IEEE conference on computer vision and pattern recognition, Honolulu, HI, USA, 4 August 2017. <https://doi.org/10.48550/arXiv.1611.07004>.
28. Zhu, J.Y.; Park, T.; Isola, P.; Efros, A.A. Unpaired image-to-image translation using cycle-consistent adversarial networks. In Proceedings of the IEEE international conference on computer vision, Venice, Italy, 22-29 October 2017. <https://doi.org/10.48550/arXiv.1703.10593>.
29. Guo, C.; Li, C.; Guo, J.; Loy, C.C.; Hou, J.; Kwong, S.; Cong, R. Zero-reference deep curve estimation for low-light image enhancement[C]. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, Seattle, USA, 13-19 June 2020. <https://doi.org/10.48550/arXiv.2001.06826>.
30. Zhang, Y.; Di, X.G.; Wu, J.D.; FU, R.; Li, Y.; Wang, Y.; Xu, Y.W.; YANG, G.H.; Wang, C.H. A Fast and Lightweight Network for Low-Light Image Enhancement. *arXiv* **2023**, arXiv:2304.02978.
31. García-Vega, A.; Espinosa, R.; Ochoa-Ruiz, G.; Bazin, T.; Falcón-Morales, L.; Lamarque, D.; Daul, C. A novel hybrid endoscopic dataset for evaluating machine learning-based photometric image enhancement models. In Proceedings of the 21th Mexican International Conference on Artificial Intelligence. "Salón de Congresos", Tecnológico de Monterrey (ITESM), Monterrey, Mexico, 24-29 October 2022. https://doi.org/10.1007/978-3-031-19493-1_22.
32. Bai, L.; Chen, T.; Wu, Y.; Wang, A.; Islam, M.; Ren, H. LLCaps: Learning to illuminate low-light capsule endoscopy with curved wavelet attention and reverse diffusion. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Vancouver, Canada, 8-12 October 2023. https://doi.org/10.1007/978-3-031-43999-5_4.
33. Zamir, S.W.; Arora, A.; Khan, S.; Hayat, M.; Khan, F.S.; Yang, M.H.; Shao, L. Learning enriched features for real image restoration and enhancement. In Proceedings of the Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, 23–28 August 2020. https://doi.org/10.1007/978-3-030-58595-2_30.
34. Wang, Z.; Bovik, A.C.; Sheikh, H.R.; Simoncelli, E.P. Image quality assessment: from error visibility to structural similarity. *IEEE Trans. Image Process*, **2004**, *13*, 600-612. <https://doi.org/10.1109/TIP.2003.819861>.
35. Ali, S.; Ghatwary, N. Endoscopic computer vision challenges 2.0, **2022**. <https://endocv2022.grand-challenge.org/5,6,7>.
36. Borgli, H.; Thambawita, V.; Smedsrud, P.H.; Hicks, S.; Jha, D.; Eskeland, S.L.; Randel, K.R.; Pogorelov, K.; Lux, M.; Nguyen, D.T.D.; Johansen, D.; Griwodz, C.; Stensland, H.K.; Garcia-Ceja, E.; Schmidt, P.T.; Hammer, H.L.; Riegler, M.A.; Halvorsen, P.; Lange, T.D. HyperKvasir, a comprehensive multi-class image and video dataset for gastrointestinal endoscopy. *Sci. Data*, **2020**, *7*, 283. <https://doi.org/10.1038/s41597-020-00622-y>.
37. Zhang, R.; Isola, P.; Efros, A.A.; Shechtman, E.; Wang, O. The unreasonable effectiveness of deep features as a perceptual metric. In Proceedings of the IEEE conference on computer vision and pattern recognition, Salt Lake City, Utah, USA, 18-22 Jun 2018. <https://doi.org/10.48550/arXiv.1801.03924>.
38. Ying, Z.; Li, G.; Ren, Y.; Wang, R.; Wang, W. A new low-light image enhancement algorithm using camera response model. In Proceedings of the IEEE international conference on computer vision workshops, Hawaii Convention Center, Hawaii, USA, 21-26 July 2017.
39. Liu, C.; Wu, F.; Wang, X. EFINet: Restoration for low-light images via enhancement-fusion iterative network[J]. *IEEE Trans. Circ Syst Vid*, **2022**, *32*, 8486-8499. <https://doi.org/10.1109/TCSVT.2022.3195996>.

40. Zhang, K.B.; Yuan, C.; Li, J.; Gao, X.B., Li, M.Q. Multi-Branch and Progressive Network for Low-Light Image Enhancement. *IEEE Trans. Image Process.* **2023**, *32*, 2295-2308. <https://doi.org/10.1109/TIP.2023.3266171>.

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