

RESEARCH

Open Access



Retinex theory-based nonlinear luminance enhancement and denoising for low-light endoscopic images

En Mou^{1,2}, Huiqian Wang^{3*}, Xiaodong Chen², Zhangyong Li³, Enling Cao⁴, Yuanyuan Chen^{1,5}, Zhiwei Huang² and Yu Pang^{3*}

Abstract

Background The quality of low-light endoscopic images involves applications in medical disciplines such as physiology and anatomy for the identification and judgement of tissue structures. Due to the use of point light sources and the constraints of narrow physiological structures, medical endoscopic images display uneven brightness, low contrast, and a lack of texture information, presenting diagnostic challenges for physicians.

Methods In this paper, a nonlinear brightness enhancement and denoising network based on Retinex theory is designed to improve the brightness and details of low-light endoscopic images. The nonlinear luminance enhancement module uses higher-order curvilinear functions to improve overall brightness; the dual-attention denoising module captures detailed features of anatomical structures; and the color loss function mitigates color distortion.

Results Experimental results on the Endo4IE 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). The PSNR is 27.2202, SSIM is 0.8342, and the LPIPS is 0.1492. It provides a method to enhance image quality in clinical diagnosis and treatment.

Conclusions It offers an efficient method to enhance images captured by endoscopes and offers valuable insights into intricate human physiological structures, which can effectively assist clinical diagnosis and treatment.

Keywords Retinex theory, Nonlinear luminance enhancement, Low-light endoscopic images, Denoising

*Correspondence:

Huiqian Wang
wanghq@cqupt.edu.cn
Yu Pang
pangyu@cqupt.edu.cn

¹School of Communications and Information Engineering, Chongqing University of Posts and Telecommunications, Chongqing 400065, China

²School of Medical Information and Engineering, Southwest Medical University, Luzhou 646000, China

³School of Optoelectronic Engineering, Chongqing University of Posts and Telecommunications, Chongqing 400065, China

⁴School of Software Engineering, Chongqing University of Posts and Telecommunications, Chongqing 400065, China

⁵College of Physics and Telecommunication Engineering, Zhoukou Normal University, Zhoukou 466001, China



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

Background

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 models include spatial domain-based methods [1, 2], frequency domain-based methods [3, 4], Retinex model-based methods [5, 6], fusion-based methods [7, 8], deep learning methods [9, 10], etc. The principle of the spatial or frequency domain approach is to transform a low-light image into a specific feature space by adjusting its pixels' distribution and dynamic range. The fusion-based model is designed to extract complementary details from a set of low-light images of the same scene. This method does not consider the correlation between neighboring pixels and different channels, resulting in images with noticeable artifacts and blurred details. The Retinex theory is a computational model for human color vision that can be adapted to be used as an image enhancer [11–14]. Retinex has inspired some illuminant and reflectance image decomposition methods [15, 16]. Deep learning approaches have the potential to learn complex nonlinear end-to-end mappings and have shown impressive performance in the field of image restoration [17]. Tan et al. [15] decomposed the endoscopic 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 endoscopic image is decomposed into light and reflection components, which are then processed separately. Finally, the reconstructed image is obtained by multiplying the enhanced illuminance and reflected components. Wang et al. [6] proposed an initial illumination weighting method to improve the illumination uniformity of endoscopic images 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. 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 is used in a wide range of fields, such as computer vision [10, 19], biomedical signal processing [20], and drug discovery [21, 22]. Low-light image enhancement, a subset of computer vision, has made significant progress and achieved excellent results. The pioneering LLNet [9] 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 [10, 23]. ExCNet [24] estimates the S-curve for the whole image, the curve can be simply parameterized by shadow amount and high-light amount. Many researchers employed adversarial generative networks to generate synthetic datasets. ZeroDCE [25] 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 [26] designed a lightweight enhancement network with global and local feature extraction adjustments, proving effective for enhancing low-light images. SMNet [10] designed a multi-scale network considering both local details and global brightness for low-light image enhancement. MBPNet [23] designed a multi-branch network with different scales, which gradually enhanced low-light images from coarse to fine. While these algorithms have yielded satisfactory outcomes in enhancing natural images, their efficacy is constrained when applied to medical endoscopic images. The special physiological 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 luminance in cavity depressions, and they fail to consider overall image luminance uniformity and overexposure, both critical for expanding the surgeon's field of view and executing surgical maneuvers. Existing algorithms enhance the dark parts of an image and increase the brightness, improving the image quality. However, there are deficiencies in overall brightness, texture details, and colors, such as inadequate representation of texture details, insufficient brightness, and color shifts. These aspects are very important for doctors to diagnose and treat. Current networks often cause detailed information about physiological tissue structures in endoscopic images to be smoothed during the brightness enhancement process. However, anatomical details in weakly textured images are important for diagnosis and treatment and require special attention. Unlike natural images, the physiological environment of endoscopic applications requires strict color fidelity, but existing methods typically exhibit

significant color deviations in this environment and are therefore not suitable for direct application to scene brightness enhancement. Noise in the input data is a common problem in medical imaging [27, 28]. The presence of noise can significantly degrade the quality and reliability of the data, leading to several challenges such as reduced signal clarity, increased error rates, compromised data integrity, difficulties in feature extraction, and so on. While noise is generally seen as detrimental, the concept of stochastic resonance (SR) illustrates that noise can sometimes be harnessed to improve signal processing. In medical imaging, stochastic resonance can be used to enhance the visibility of faint structures in noisy images. By adding a controlled amount of noise, the contrast of these structures can be improved, aiding in better diagnosis. In communication systems and signal processing, stochastic resonance can improve the detection of weak signals in noisy environments, leading to more robust data transmission. By understanding and leveraging this phenomenon, we can address some of the challenges posed by noise in input data and improve the performance of various systems in medical imaging [29, 30], and beyond. Significant advances have been made in medical image processing [31–39], providing a solid foundation for medical diagnosis and treatment. We propose a new method for image enhancement that exploits an image decomposition approach inspired by Retinex theory. The network comprises a nonlinear luminance enhancement module and a denoising module. The loss function accounts for color difference, structure, and luminance aspects. Experimental results on the Endo4IE [40] 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) [41], and Learned Perceptual Image Patch Similarity (LPIPS) [42].

In summary, this approach contributes the following key elements:

- (1) A novel network architecture is proposed for the luminance and detailed feature enhancement of low-light images in endoscopic environments. The network addresses the global luminance imbalance and the weak organizational texture commonly found in endoscopic images by integrating global luminance, and local and spatial detail enhancement for noise reduction, thereby achieving a balanced enhancement of luminance in endoscopic images;
- (2) The nonlinear luminance enhancement module mitigates the illuminance inhomogeneity in endoscopic images resulting from the use of point light sources and the physiological tissue structure environment. This is achieved by enhancing the overall image luminance perspective. Inspired by the

Retinex methodology, the module extracts the overall image luminance through the decomposition of the model and optimizes the higher-order curve function using histogram information to automatically supplement the image luminance;

- (3) Addressing the weak texture characteristics of endoscopic images, the denoising module incorporates a feature enhancement with a dual-attention mechanism. This mechanism enhances the 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, Sect. 2 details the proposed image enhancement method, Sect. 3 covers the related experiments, Sect. 4 provides the discussions, and Sect. 5 offers conclusions.

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.

Overall pipeline

Figure 1 shows the comprehensive framework of the low-light image enhancement network. The nonlinear luminance enhancement module automatically improves light luminance by employing illumination separation and nonlinear enhancement, aiming at global luminance equalization and luminance enhancement of dark areas. The denoising module enhances physiological structure details by amplifying local and spatial 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 nonlinear luminance enhancement module and the denoising module. Given a low-light image, an illumination map and a reflection map can be obtained using an image decomposition approach. The decomposition process includes three convolutions with LeakyRelu functions and one convolution with the Relu function. Then, the illumination map is enhanced by using a nonlinear luminance enhancement module with high order curve function. Afterward, the denoising module extracts detailed features with 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 luminance of dark areas and maintain color fidelity. It outperforms existing methods in terms of visual effects and performance metrics.

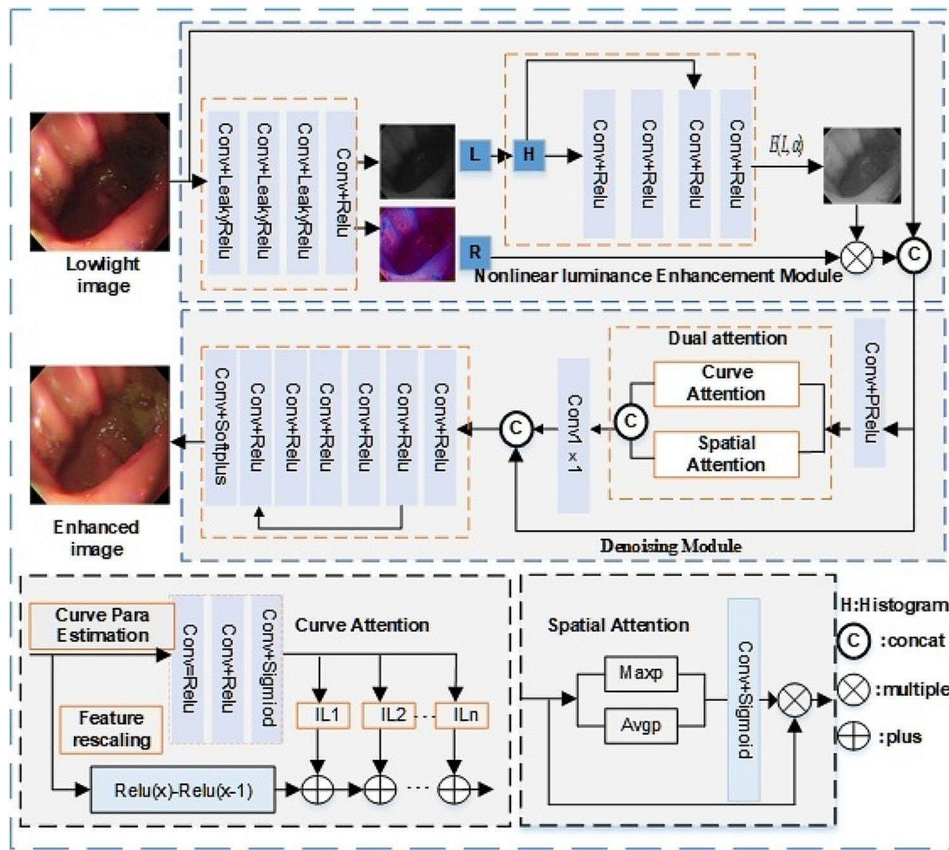


Fig. 1 Framework flowchart

Nonlinear luminance enhancement module

To enhance the global luminance, a nonlinear luminance enhancement module is designed in this paper to enhance the illumination map by the higher-order curve function [25]. The method proposed in this paper operates at the pixel level, assuming each pixel follows a higher-order curve, i.e., the pixel-wise curve. With the fitting maps, the enhanced version image can be directly obtained by pixel-wise curve mapping. We design a decomposition method to obtain the illumination map by the following equation:

$$I = L \cdot R \tag{1}$$

Where I is the low-light image, L is an illumination map, and R is a reflection map. Inspired by the ZeroDCE principle, a higher-order curve function operation is performed on the illumination map L to achieve global luminance enhancement. The parameters of the higher-order curve function are computed from the histogram of the illumination map and consist of a four-layer convolution and a LeakyReLU activation function. According to the literature [25], the equation can be applied by the following equation:

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

where the parameter n is the number of iterations and controls the curvature of the curve; alpha is the trainable curve parameter, the value ranges from -1 to 1 ; E denotes the illuminance value at a coordinate x . Here, we assume that pixels in a local region have the same intensity (also the same adjustment curves), and thus the neighboring pixels in the output result still preserve the monotonous relations. We formulate alpha as a pixel-wise parameter, i.e., each pixel of the given input image has a corresponding curve with the best-fitting alpha to adjust its dynamic range. So, both dark and bright pixels are adjusted according to this curve. In our method, we set the value of n to 7. 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. Then the enhanced illumination map L_{en} and the reflection map R are multiplied to obtain a global luminance enhanced image I_g . The equation is as follows:

$$I_g = concat([L_{en}, L_{en}, L_{en}]) \times R \tag{3}$$

Denoising module

The image is adjusted with local attention and spatial attention. To better extract the physiological structural detail 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 can better capture features near image pixel points, especially in complex geometric structures or textures. This is particularly important for fine anatomical structures in medical images, such as blood vessels and nerve fibers. Spatial attention enhances the model's ability to represent key features by focusing on the most important regions in the image while suppressing the interference of irrelevant information. Thus, a denoising effect is achieved. Curve attention proposed in the literature [43] can effectively extract the detailed features, the equation is as follows:

$$\frac{IL_{n(c)}}{IL_{n-1(c)}} = Curve_{n-1}(1 - IL_{n-1(c)}) \quad (4)$$

Where $IL_{n(c)}$ is the curve function, and c represents the feature location coordinates. Here, this curvilinear attention is used to improve the six-channel image obtained earlier, resulting in a curvilinear attention map of the overall image. It is estimated by three Conv2D layers and Sigmoid activation. The equation is as follows:

$$CA(I_D) = IL_{n(c)}(I_D) \quad (5)$$

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

Spatial attention [44] adopted global average pooling and max pooling on input features respectively to get the interspatial dependencies of convolutional features, which are then concatenated for convolution and Sigmoid operations. The equation is as follows:

$$\begin{aligned} SPA(I_D) \\ = I_D \otimes Sig(Conv(concat(Max(I_D), Avg(I_D)))) \end{aligned} \quad (6)$$

$SPA(I_D)$ 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:

$$I_{dual} = Conv_{1 \times 1}\{concat(CA(I_D), SPA(I_D))\}, \quad (7)$$

Then, there are seven layers of convolution and activation functions to extract detailed features.

Total loss

The loss functions in FLW mainly consider the relative loss to keep the structure, saturation, and brightness of

the image consistent with the reference image. These loss functions are designed to ensure that the enhanced image retains as much structural information, i.e., luminance information, as possible. We add a color loss function to the loss function designed in FLW to recover more color information. 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:

$$l_{total} = l_1 + l_{SSIM} + l_{hs} + l_b + l_{str} + l_{cd} \quad (8)$$

An explanation of the symbols in the formula will be provided later. CIELAB (also known as the CIE Lab* color space) is considered to be closer to human color perception because it was specifically designed to be perceptually uniform. This means that a given numerical change in CIELAB values corresponds to a roughly equal change in perceived color. The CIELAB color space models human vision by incorporating knowledge of the nonlinearities of human color perception, making it more accurate in representing how humans perceive color differences compared to other color spaces, such as RGB [45–48]. Given its alignment with human visual perception, the CIELAB color space enables more effective comparison and measurement of color differences. So, 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 in this paper. The color difference equation l_{cd} is as follows:

$$l_{cd} = dl + dc / \left((1 + 0.045c_1)^2 \right) + dh / \left((1 + 0.015c_1)^2 \right), \quad (9)$$

Where dl , dc , dh , c_1 denote the square of the L channel difference, the square of the c channel difference, the color channel difference, and the mean value of the color 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:

$$l_1 = \frac{1}{mn} \sum_{i=1, j=1}^{m, n} |I_{en}(i, j) - I_{ref}(i, j)|, \quad (10)$$

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

$$l_{SSIM} = 1 - SSIM(I_{en}, I_{ref}), \quad (11)$$

The structural similarity function is as follows:

$$SSIM(I_{en}, I_{ref}) = \frac{2\mu_1\mu_2 + c_1}{\mu_1^2 + \mu_2^2 + c_1} \cdot \frac{2\sigma_1\sigma_2 + c_2}{\sigma_1^2 + \sigma_2^2 + c_2}, \quad (12)$$

l_{hs} [26] measures the hue and saturation difference between two pixels, the equation is as follows:

$$l_{hs} = 1 - \sum_{i=1, j=1}^{m, n} \langle I_{en}(i, j), I_{ref}(i, j) \rangle, \quad (13)$$

Where $\langle \bullet, \bullet \rangle$ represents the cosine similarity of two vectors.

l_b [26] express the luminance relation, the equation is as follows:

$$l_b = 1 - \sum_{c \in \{R, G, B\}} \sum_{i=1, j=1}^{m, n} \langle I_{en}^c(i, j) - \min I_{en}^c(i, j), I_{ref}^c(i, j) - \min I_{ref}^c(i, j) \rangle \quad (14)$$

l_{str} [26] is the gradient loss. l_{str} includes the sum of the horizontal and vertical losses. The function calculates the gradient from the horizontal and vertical directions. To compute the gradient consistency loss, it first subtracts the smallest value in the image to ensure that no negative values interfere with the calculation. It then calculates the gradient consistency loss for each channel, both independently between color channels and collectively. Specifically, it uses cosine similarity to measure the gradient consistency between the reference and enhanced images. The gradient consistency between the reference and enhanced images across different color channels and spatial locations is ensured in greater detail. This approach is especially useful for processing color images because it captures and maintains the details and color information more accurately. 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:

$$l_{str} = \sum_{c \in \{R, G, B\}} \sum_{i=1, j=1}^{m, n} \langle \nabla I_{en}^c(i, j) - \min \nabla I_{en}^c(i, j), \nabla I_{ref}^c(i, j) - \min \nabla I_{ref}^c(i, j) \rangle \quad (15)$$

Implementation details

The Endo4IE dataset is employed for training and testing. This dataset is developed from the EAD2020 Challenge dataset [49], EAD 2.0 from the EndoCV2022 Challenge dataset [50], and the HyperKvasir dataset [51], 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 images for training, 29 pairs for validation, and 266 pairs for testing. The proposed method is

quantitatively compared with several benchmark experiments: PSNR, SSIM, and LPIPS.

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. Throughout the training process, a batch size of 100 was employed, and the model utilized the Adam optimizer with a learning rate of 0.0001. Visual comparisons with existing methods show that while these methods enhance the luminance of the image, they tend to fall short of preserving image details. The method presented in this paper demonstrates superior performance compared to state-of-the-art methods, as indicated by quantitative evaluations using PSNR, SSIM, and LPIPS metrics.

Results

The visual comparisons of results using the Endo4IE dataset are illustrated in Figs. 2 and 3, and 4. The various indicators of quantitative comparison are displayed in Table 1. The results validate the effectiveness of our method, which can effectively brighten the dark area in endoscopic images, highlight structural details, and keep the images underexposed with natural colors. The other seven contrast methods can also enhance the brightness of medical endoscopy images. However, the brightness enhancement of the recessed area in the center of the cavity is insufficient, and the texture details of the recessed area are not fully displayed. As for the LIME and Zero-DCE methods, there is obvious color distortion in the enhanced images.

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 SMNet trained on the Endo4IE dataset. As indicated in Table 1, the PSNR, SSIM, and LPIPS values for the method presented in this paper are 27.2202 dB, 0.8342, and 0.1492, respectively. In the quantitative experiments that have been conducted, the index values of our proposed method exceed those of seven excellent algorithms including FLW, SMNet, and MBPNet. 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.

In Fig. 2, the image generated by the Zero-DCE method has an overall grey appearance with insufficient enhancement of the central black areas, and the images tested with the LIME and LR3M methods tend to be reddish. The luminance of LIME and LR3M methods was improved, but there were some noise spots in the edge portion of the image, as in Fig. 3. From Fig. 3, the image enhancement results of Ying's method were better than those of LR3M and LIME, but there was still a

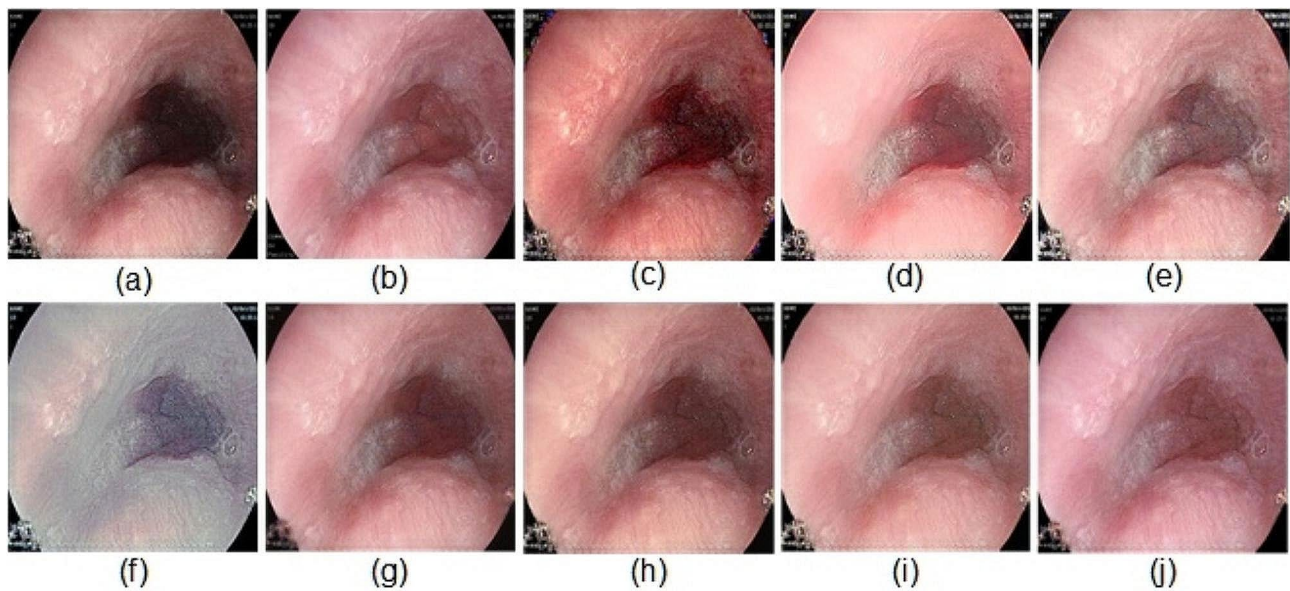


Fig. 2 Visual comparison results on Endo4IE. **a** The lowlight image. **b** The reference image. **c** LIME. **d** LR3M. **e** Ying's. **f** Zero-DCE. **g** SMNet. **h** FLW. **i** MBPNet. **j** Ours

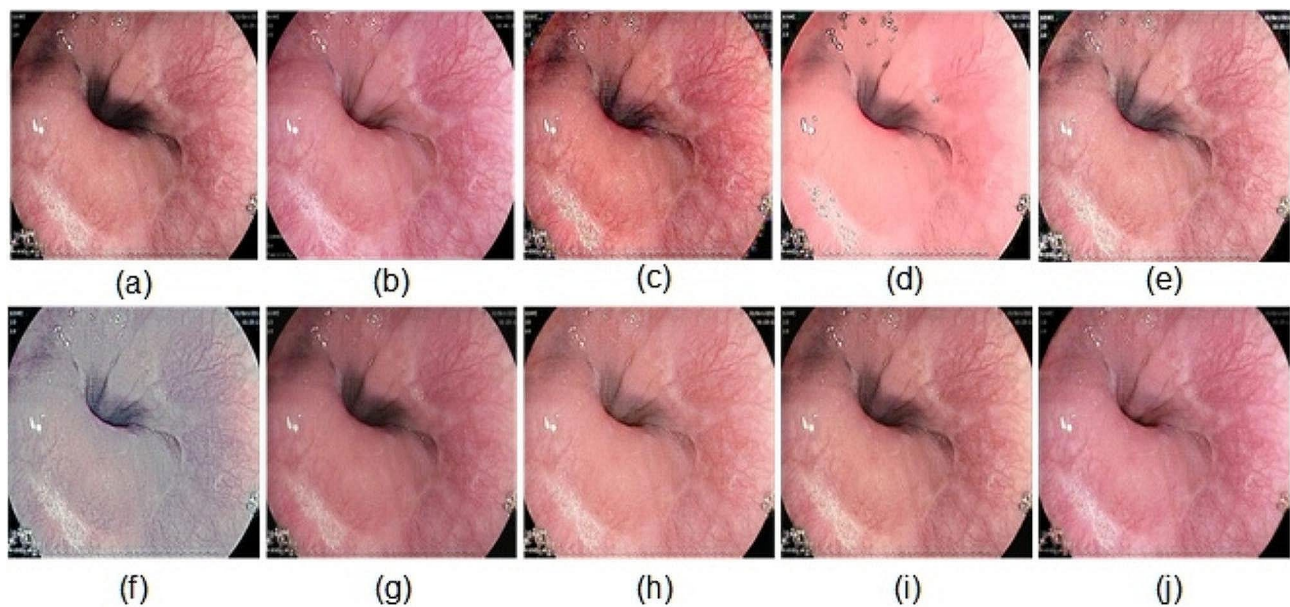


Fig. 3 Visual comparison results on Endo4IE. **a** The lowlight image. **b** The reference image. **c** LIME. **d** LR3M. **e** Ying's. **f** Zero-DCE. **g** SMNet. **h** FLW. **i** MBPNet. **j** Ours

slight deficiency in terms of color maintenance when compared to the reference image. In Fig. 4, the luminance of the image enhanced by FLW was improved, but the color fidelity was not well controlled; the image tested by SMNet showed the same phenomenon as the two algorithms just described in the central dark area. MBPNet gives better results than the previous methods, with detail in darker areas needing to be enhanced. From Fig. 4, it can be seen that the luminance of the

edge part can still be improved to get a more anatomically detailed texture. Compared with these methods, the method proposed in this paper outperforms the same training dataset. It ensures uniform luminance enhancement in the central recessed region, improves detailed features and contrast, maintains color consistency, and enhances physicians' visibility of complex physiological structures. This aids doctors in diagnosing and treating physiological abnormalities. Our training and testing

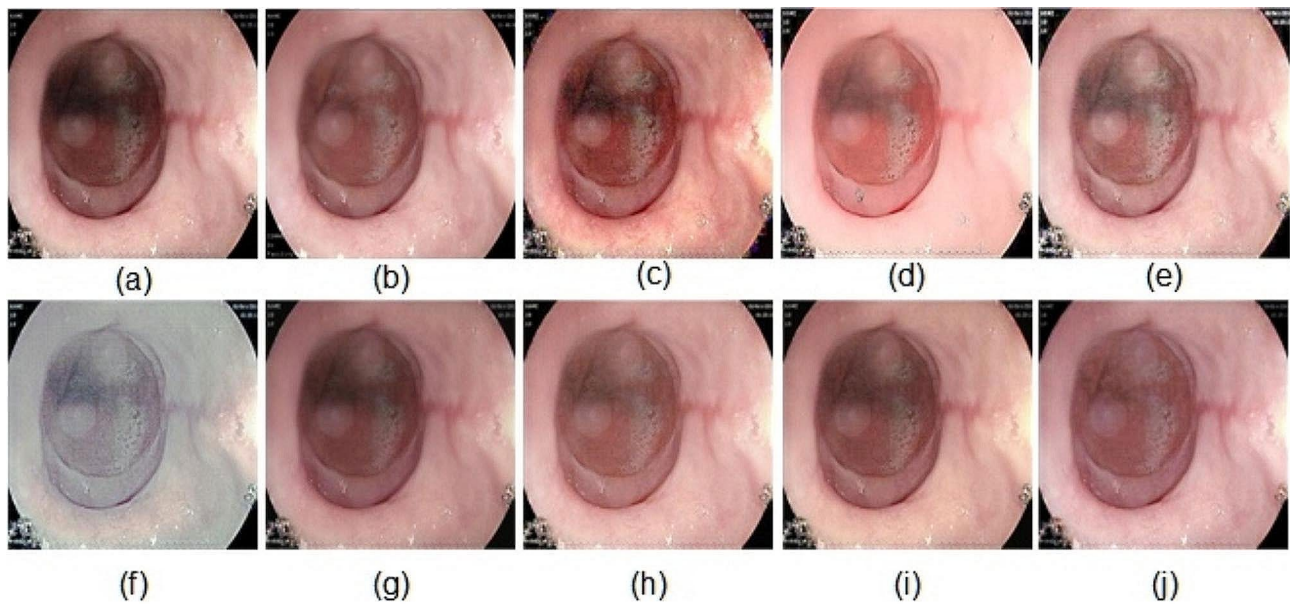


Fig. 4 Visual comparison results on Endo4IE. **a** The lowlight image. **b** The reference image. **c** LIME. **d** LR3M. **e** Ying's. **f** Zero-DCE. **g** SMNet. **h** FLW. **i** MBPNet. **j** Ours

Table 1 Quantitative comparison results on Endo4IE

Method	Published	PSNR (dB)↑	SSIM↑	LPIPS↓
LR3M [18]	TIP	14.7685	0.6998	0.2630
LIME [5]	TIP	19.8880	0.7039	0.2294
Ying's [52]	ICCV	18.0095	0.7128	0.1839
Zero-DCE [25]	CVPR	15.0306	0.6751	0.3016
SMNet [10]	TMM	22.3031	0.7902	0.1763
FLW [26]	ArXiv	25.1273	0.8238	0.1513
MBPNet [23]	TIP	25.4923	0.7950	0.2005
ours		27.2202	0.8342	0.1492

Table 2 Results of quantitative comparisons about different components of the model on training

Sequence	Nonlinear luminance Enhancement Module	Denoising Module	PSNR (dB)↑	SSIM↑	LPIPS↓
1		√	26.5592	0.8324	0.1499
2	√		26.8640	0.8274	0.1557
3	√	√	27.2202	0.8342	0.1492

are done on the GPU. The proposed method has 0.13 M trainable parameters and 1.15G Flops for an input image of size $96 \times 96 \times 3$. Currently, it can process 500 images per second. In the next step, if the platform is ported to a system-on-chip [53, 54], its performance will be further improved, enhancing the possibility of its application in clinical diagnostic therapy. Both qualitative and quantitative experiments validate the efficacy of the proposed method in enhancing low-light endoscopic images.

Table 3 Results of quantitative comparisons about color difference loss function on training

Sequence	$l_1 + l_{SSIM} + l_{hs} + l_b + l_{str}$	l_{cd}	PSNR (dB)↑	SSIM↑
1	√		25.6055	0.8333
2		√	27.2202	0.8342

Ablation study

The ablation study demonstrated the individual contribution of each component and color difference loss in the network. We designed a new network including a nonlinear luminance enhancement 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 and color difference loss function using the Endo4IE dataset to demonstrate the effectiveness of each module, the quantitative comparisons are shown in Tables 2 and 3.

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 nonlinear luminance 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 nonlinear luminance enhancement module. Removing the nonlinear luminance enhancement module results in the LPIPS metric that is 0.0017 higher than the LPIPS of the proposed algorithm. Similarly, removing the denoising module results in an LPIPS that is 0.0094 higher than that of the entire network framework. Applying either the

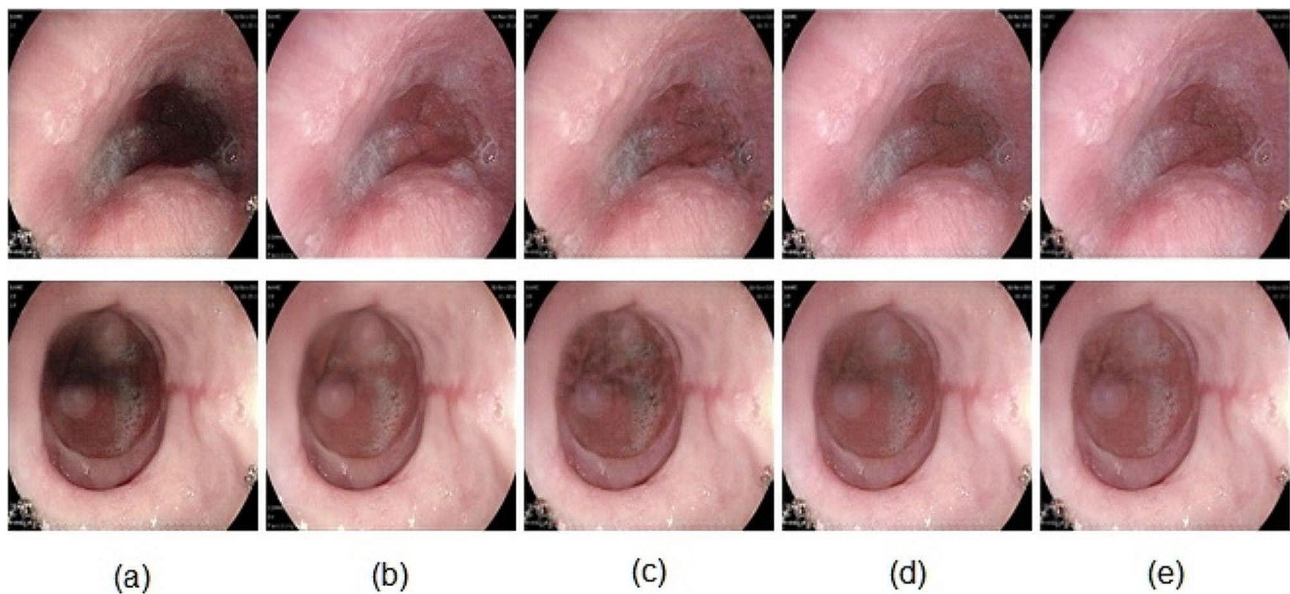


Fig. 5 Visual comparison results for the ablation experiment on Endo4IE. **a** The lowlight image. **b** The reference image. **c** Absence of the nonlinear luminance enhancement module. **d** Absence of the denoising module. **e** Our method

nonlinear luminance enhancement module or the denoising module for detailed feature 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.

The contribution of individual modules to image enhancement is shown in Fig. 5. As shown in Fig. 5(c), the lack of a local enhancement module leads to poor overall anatomical image detail enhancement, which can be improved in both luminance uniformity and contrast enhancement. Correspondingly, Fig. 5(d) shows the appearance of black dots in the image, which shows the inconsistency of the enhancement of the dark region and proves the contribution of the global enhancement module.

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.

Discussion

We propose an efficient luminance enhancement and denoising method specifically optimized for low-light endoscopic images. Experimental results (shown in Table 1; Figs. 2 and 3, and 4) indicate that our model has

significant advantages in image enhancement. Although the method proposed in this paper can enhance endoscopic low-light images well, further efforts are needed to lighten the parameters and computation so that it can meet the requirements of clinical diagnosis and treatment. Overall, the proposed method demonstrates strong potential in the field of endoscopic image enhancement, not only in terms of brightness and noise processing but also in maintaining color fidelity and detail sharpness. This provides an enhancement method for endoscopic low-light images.

Conclusions

This paper introduces a novel model for enhancing low-light endoscopic images through Retinex theory-based nonlinear luminance and denoising for detailed features. Inspired by the Retinex traditional method, the model separated low-light images to obtain the illuminance map and iteratively optimized higher-order curve functions using histogram information for balanced global luminance enhancement. To address the importance of detail organization structure, a denoising module with a dual attention mechanism is integrated to amplify interaction features in local and spatial 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 hardware the network, enhancing real-time feasibility through hardware

implementation. Algorithm optimization will involve integrating endoscopic luminance enhancement as a pre-processing step for subsequent tasks such as depth prediction and target recognition in endoscopy.

Acknowledgements

Not applicable.

Author contributions

E.M., H.W., X.C., Y.P., Z.H. and Z. L. designed the network and wrote the original manuscript. E.M. and H.W. contributed to the review and editing. H.W. and Y.P. were mainly responsible for this project. E.C. and Y. C. completed the typesetting of the paper. All authors reviewed the manuscripts.

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, Chongqing Natural Science Foundation Innovation and Development Joint Fund Project under Grant No.CSTB2023NSCQ-LZX0064, Sichuan Province Science and Technology Support Program under Grant No.2022NSFSC0508 and No.2022YFS0616, Research Project of Southwest Medical University under Grant No.2021ZKZD019. The work was funded by Central Nervous System Drug Key Laboratory of Sichuan Province (230005-01SZ), Medical Engineering & Medical Informatics Integration and Transformational Medicine of Luzhou Key Laboratory.

Data availability

The datasets for this study can be found: <https://data.mendeley.com/datasets/3j3tmghw33/1>.

Declarations

Ethics approval and consent to participate

This article does not contain any studies with human participants or animals performed by any of the authors.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Received: 25 February 2024 / Accepted: 1 August 2024

Published online: 09 August 2024

References

- Pizer SM, Johnston RE, Ericksen JP, Yankaskas BC, Muller KE. Contrast-Limited Adaptive Histogram Equalization: Speed and Effectiveness. In: Proceedings of the First Conference on Visualization in Biomedical Computing. 1990; 337:2.
- 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). 2021: 234–239. <https://doi.org/10.1109/ICSIP52628.2021.9688796>
- Huang SC, Cheng FC, Chiu YS. Efficient contrast enhancement using adaptive gamma correction with weighting distribution. *IEEE Trans Image Process.* 2012;22(3):1032–41. <https://doi.org/10.1109/TIP.2012.2226047>.
- Ju M, Ding C, Zhang D, Guo YJ. Gamma-correction-based visibility restoration for single hazy images. *IEEE Signal Process Lett.* 2018;25(7):1084–8. <https://doi.org/10.1109/LSP.2018.2839580>.
- Guo X, Li Y, Ling H, LIME. Low-light image enhancement via illumination map estimation. *IEEE Trans Image Process.* 2016;26(2):982–93. <https://doi.org/10.1109/TIP.2016.2639450>.
- 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.* 2018;18(4):e2396. <https://doi.org/10.1002/rcs.2396>.
- Celebi AT, Duvar R, Urhan O. Fuzzy fusion based high dynamic range imaging using adaptive histogram separation. *IEEE Trans Consum Electron.* 2015;61(1):119–27. <https://doi.org/10.1109/TCE.2015.7064119>.
- Sugimura D, Mikami T, Yamashita H, et al. Enhancing color images of extremely low light scenes based on RGB/NIR images acquisition with different exposure times. *IEEE Trans Image Process.* 2015;24(11):3586–97. <https://doi.org/10.1109/TIP.2015.2448356>.
- Lore KG, Akintayo A, Sarkar S, LLNet: A deep autoencoder approach to natural low-light image enhancement. *Pattern Recogn.* 2017;61:650–62. <https://doi.org/10.1016/j.patcog.2016.06.008>.
- Lin SD, Tang F, Dong WM, Pan XJ, Xu CS, SMNet. Synchronous multi-scale low light Enhancement Network with local and global concern. *IEEE Trans Multimedia.* 2023;25:9506–17. <https://doi.org/10.1109/TMM.2023.3254141>.
- McCann J. Lessons learned from mondrians applied to real images and color gamuts//Color and imaging conference. Society of Imaging Science and Technology. 1999; 7: 1–8. <https://doi.org/10.2352/CIC.1999.7.1art00001>
- McCann JJ, Parraman C, Rizzi A. Reflectance, illumination, and appearance in color constancy. *Front Psychol.* 2014;5:57285. <https://doi.org/10.3389/fpsyg.2014.00005>.
- Rizzi A, Bonanomi C. Milano Retinex family. *J Electron Imaging.* 2017;26(3):031207–031207. <https://doi.org/10.1117/1.JEI.26.3.031207>.
- Lecca M. Generalized equation for real-world image enhancement by Milano Retinex family. *JOSA A.* 2020;37(5):849–58. <https://doi.org/10.1364/JOSAA.384197>.
- Tan W, Xu C, Lei F, Fang Q, An Z, Wang D, et al. An endoscope image Enhancement Algorithm based on image decomposition. *Electronics.* 2022;11(12):1909. <https://doi.org/10.3390/electronics11121909>.
- An Z, Xu C, Qian K, Han J, Tan W, Wang D, et al. EIEN: endoscopic image enhancement network based on Retinex theory. *Sensors.* 2022;22(14):5464. <https://doi.org/10.3390/s22145464>.
- Li C, Guo C, Han L, Jiang J, Cheng MM, Gu J, Loy CC. Low-light image and video enhancement using deep learning: a survey. *IEEE Trans Pattern Anal Mach Intell.* 2022;44(12):9396–416. <https://doi.org/10.1109/TPAMI.2021.3126387>.
- Ren X, Yang W, Cheng WH, Liu J. LR3M: robust low-light enhancement via low-rank regularized Retinex model. *IEEE Trans Image Process.* 2020;29:5862–76. <https://doi.org/10.1109/TIP.2020.2984098>.
- Ansari MY, Qaraqe M, Mefood. A large-scale representative benchmark of quotidian foods for the middle east. *IEEE Access.* 2023;11:4589–601. <https://doi.org/10.1109/ACCESS.2023.3234519>.
- Ansari MY, Qaraqe M, Charafeddine F, Serpedin E, Righetti R, Qaraqe K. Estimating age and gender from electrocardiogram signals: a comprehensive review of the past decade. *Artif Intell Med.* 2023;102690. <https://doi.org/10.1016/j.artmed.2023.102690>.
- Chandrasekar V, Ansari MY, Singh AV, Uddin S, Prabhu KS, Dash S, et al. Investigating the use of machine learning models to understand the drugs permeability across placenta. *IEEE Access.* 2023. <https://doi.org/10.1109/ACCESS.2023.3272987>.
- Ansari MY, Chandrasekar V, Singh AV, Dakua SP. Re-routing drugs to blood brain barrier: a comprehensive analysis of machine learning approaches with fingerprint amalgamation and data balancing. *IEEE Access.* 2022;11:9890–906. <https://doi.org/10.1109/ACCESS.2022.3233110>.
- Zhang KB, Yuan C, Li J, Gao XB, Li MQ. Multi-branch and Progressive Network for Low-Light Image Enhancement. *IEEE Trans Image Process.* 2023;32:2295–308. <https://doi.org/10.1109/TIP.2023.3266171>.
- Zhang L, Zhang L, Liu X, Shen Y, Zhang S, Zhao S. Zero-shot restoration of back-lit images using deep internal learning. In: Proceedings of the 27th ACM international conference on multimedia. 2019; 1623–1631. <https://doi.org/10.1145/3343031.3351069>
- Guo C, Li C, Guo J, Loy CC, Hou J, Kwong S et al. Zero-reference deep curve estimation for low-light image enhancement. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2020;1780–1789. <https://doi.org/10.48550/arXiv.2001.06826>
- Zhang Y, Di XG, Wu JD, FU R, Li Y, Wang Y et al. A Fast and Lightweight Network for Low-Light Image Enhancement. *arXiv:2304.02978.* 2023.
- Mohanty S, Dakua SP. Toward computing cross-modality symmetric non-rigid medical image registration. *IEEE Access.* 2022;10:24528–39. <https://doi.org/10.1109/ACCESS.2022.3154771>.
- Dakua SP, Abinahed J, Al-Ansari AA. Pathological liver segmentation using stochastic resonance and cellular automata. *J Vis Commun Image Represent.* 2016;34:89–102. <https://doi.org/10.1016/j.jvcir.2015.10.016>.

29. Regaya Y, Amira A, Dakua SP. Development of a cerebral aneurysm segmentation method to prevent sentinel hemorrhage. *Netw Model Anal Health Inf Bioinf.* 2023;12(1):18. <https://doi.org/10.1007/s13721-023-00412-7>.
30. Dakua SP. LV segmentation using stochastic resonance and evolutionary cellular automata. *Int J Pattern Recognit Artif Intell.* 2015;29(03):1557002. <https://doi.org/10.1142/S0218001415570025>.
31. Han Z, Jian M, Wang GG, ConvUNeXt. An efficient convolution neural network for medical image segmentation. *Knowl Based Syst.* 2022;253:109512. <https://doi.org/10.1016/j.knosys.2022.109512>.
32. Ansari MY, Yang Y, Balakrishnan S, Abinahed J, AlAnsari A, Warfa M, et al. A lightweight neural network with multiscale feature enhancement for liver CT segmentation. *Sci Rep.* 2022;12(1):14153. <https://doi.org/10.1038/s41598-022-16828-6>.
33. Jafari M, Auer D, Francis S, Garibaldi J, Chen X. DRU-Net: an efficient deep convolutional neural network for medical image segmentation. In: 2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI). IEEE. 2020; 1144–1148. <https://doi.org/10.1109/ISBI45749.2020.9098391>
34. Ansari MY, Yang Y, Meher PK, Dakua SP. Dense-PSP-UNet: a neural network for fast inference liver ultrasound segmentation. *Comput Biol Med.* 2023;153:106478. <https://doi.org/10.1016/j.combiomed.2022.106478>.
35. Xie Y, Zhang J, Shen C, Xia Y, Cotr. Efficiently bridging cnn and transformer for 3d medical image segmentation. In *Medical Image Computing and Computer Assisted Intervention–MICCAI 2021: 24th International Conference, Strasbourg, France, September 27–October 1, 2021, Proceedings, Part III 24*. Springer International Publishing, 2021: 171–180. https://doi.org/10.1007/978-3-030-87199-4_16
36. Ansari MY, Abdalla A, Ansari MY, Ansari MI, Malluhi B, Mohanty S, et al. Practical utility of liver segmentation methods in clinical surgeries and interventions. *BMC Med Imaging.* 2022;22(1):97. <https://doi.org/10.1186/s12880-022-00825-2>.
37. Akhtar Y, Dakua SP, Abdalla A, Aboumarzouk OM, Ansari MY, Abinahed J, et al. Risk assessment of computer-aided diagnostic software for hepatic resection. *IEEE Trans Radiation Plasma Med Sci.* 2021;6(6):667–77. <https://doi.org/10.1109/TRPMS.2021.3071148>.
38. Rai P, Ansari MY, Warfa M, Al-Hamar H, Abinahed J, Barah A, et al. Efficacy of fusion imaging for immediate post-ablation assessment of malignant liver neoplasms: a systematic review. *Cancer Med.* 2023;12(13):14225–51. <https://doi.org/10.1002/cam4.6089>.
39. Ansari MY, Mangalote IAC, Meher PK, Aboumarzouk O, Al-Ansari A, Halabi O, et al. Advancements in Deep Learning for B-Mode Ultrasound Segmentation: a Comprehensive Review. *IEEE Trans Emerg Top Comput Intell.* 2024. <https://doi.org/10.1109/TETCI.2024.3377676>.
40. García-Vega A, Espinosa R, Ochoa-Ruiz G, Bazin T, Falcón-Morales L, Lamarque D et al. 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.* 2022; 267–281. <https://doi.org/10.1007/978-3-031>
41. Wang Z, Bovik AC, Sheikh HR, Simoncelli EP. Image quality assessment: from error visibility to structural similarity. *IEEE Trans Image Process.* 2004;13(4):600–12. <https://doi.org/10.1109/TIP.2003.819861>.
42. Zhang R, Isola P, Efros AA, 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.* 2018;586–595. <https://doi.org/10.48550/arXiv.1801.03924>
43. 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.* 2023;34–44. <https://doi.org/10.1007/978-3-031-43999-5>
44. Zamir SW, Arora A, Khan S, Hayat M, Khan FS, Yang MH et al. Learning enriched features for real image restoration and enhancement. In: *Proceedings of the Computer Vision–ECCV 2020: 16th European Conference.* 2020;492–511. https://doi.org/10.1007/978-3-030-58595-2_30
45. Connolly C, Fleiss T. A study of efficiency and accuracy in the transformation from RGB to CIELAB color space. *IEEE Trans Image Process.* 1997;6(7):1046–8. <https://doi.org/10.1109/83.597279>.
46. Kuehni RG. Hue uniformity and the CIELAB space and color difference formula. *Color Research & Application: endorsed by Inter-society Color Council, the Colour Group (Great Britain), Canadian Society for Color, Color Science Association of Japan, Dutch Society for the study of Color, the Swedish Colour Centre Foundation, Colour Society of Australia, Centre Français De La Couleur.* 1998;23(5):314–22. [https://doi.org/10.1002/\(SICI\)1520-6378\(199810\)23:5%3C314::AID-COL7%3E3.0.CO;2-Z](https://doi.org/10.1002/(SICI)1520-6378(199810)23:5%3C314::AID-COL7%3E3.0.CO;2-Z).
47. Lajevardi SM, Wu HR. Facial expression recognition in perceptual color space. *IEEE Trans Image Process.* 2012;21(8):3721–33. <https://doi.org/10.1109/TIP.2012.2197628>.
48. Hill B, Roger T, Vorhagen FW. Comparative analysis of the quantization of color spaces on the basis of the CIELAB color-difference formula. *ACM Trans Graphics (TOG).* 1997;16(2):109–54. <https://doi.org/10.1145/248210.248212>.
49. Ali S, Dmitrieva M, Ghatwary N, Bano S, Polat G, Temizel A, et al. A translational pathway of deep learning methods in gastrointestinal endoscopy. *arXiv: 2010.06034.* 2020;5(6):7.
50. Ali S, Ghatwary N. (2022) Endoscopic computer vision challenges 2.0. <https://endocv2022.grand-challenge.org/> 5, 6, 7.
51. Borgli H, Thambawita V, Smedsrud PH, Hicks S, Jha D, Eskeland SL, et al. Sci Data. 2020;7(1):283. <https://doi.org/10.1038/s41597-020-00622-y>. HyperKvasir, a comprehensive multi-class image and video dataset for gastrointestinal endoscopy.
52. 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.* 2017; 3015–3022.
53. Zhai X, Amira A, Bensaali F, Al-Shibani A, Al-Nassr A, El-Sayed A, et al. Zynq SoC based acceleration of the lattice boltzmann method. *Concurrency Comput Pract Experience.* 2019;31(17):e5184. <https://doi.org/10.1002/cpe.5184>.
54. Esfahani SS, Zhai X, Chen M, Amira A, Bensaali F, AbiNahed J, et al. Lattice-boltzmann interactive blood flow simulation pipeline. *Int J Comput Assist Radiol Surg.* 2020;15:629–39. <https://doi.org/10.1007/s11548-020-02120-3>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.