Journal of Contemporary Educational Research, 2025, Volume 9, Issue 11
http://ojs.bbwpublisher.com/index.php/JCER
ISSN Online: 2208-8474
ISSN Print: 2208-8466

Low-Light Image Enhancement Based on Wavelet Local and Global Feature Fusion Network

Shun Song[1]*, Xiangqian Jiang[1], Dawei Zhao[2]

[1]Henan Vocational University of Science and Technology, Zhoukou 453003, Henan, China
[2]State Grid Zhoukou Power Supply Company, Zhoukou 453003, Henan, China

**Author to whom correspondence should be addressed.*

Abstract: A wavelet-based local and global feature fusion network (LAGN) is proposed for low-light image enhancement, aiming to enhance image details and restore colors in dark areas. This study focuses on addressing three key issues in low-light image enhancement: Enhancing low-light images using LAGN to preserve image details and colors; extracting image edge information via wavelet transform to enhance image details; and extracting local and global features of images through convolutional neural networks and Transformer to improve image contrast. Comparisons with state-of-the-art methods on two datasets verify that LAGN achieves the best performance in terms of details, brightness, and contrast.

Keywords: Image enhancement; Feature fusion; Wavelet transform; Convolutional Neural Network (CNN); Transformer

Online publication: December 8, 2025

1. Introduction

Low-light image enhancement is one of the important research directions in the field of computer vision. Its goal is to improve the visual quality and information availability of images captured under low-light conditions through algorithms. In practical scenarios such as security monitoring, autonomous driving, medical image analysis, and remote sensing detection, insufficient illumination causes problems like low brightness, unbalanced contrast, severe noise interference, and loss of details in images. These issues directly affect the performance of downstream tasks, including object detection, semantic segmentation, and feature matching. For example, if an autonomous driving system fails to effectively enhance low-light images at night or when light changes suddenly in a tunnel environment, it may lead to misjudgments by the perception module. If medical endoscope images have blurred tissue textures due to uneven illumination, it will interfere with doctors' pathological diagnosis. Therefore, improving the visual quality and semantic information restoration ability of low-light images has significant practical significance.

2. Related work

Traditional enhancement methods mostly rely on physical illumination models or hand-crafted features. For instance, histogram equalization improves contrast by expanding the dynamic range [1–4], but it tends to cause local overexposure. Methods based on Retinex theory decompose an image into illumination components and reflection components, yet they depend on prior assumptions and are sensitive to noise [5–8]. Additionally, these methods usually struggle to adapt to complex illumination distributions and noise types. In recent years, deep learning technology has advanced data-driven research on low-light enhancement. End-to-end models based on Convolutional Neural Networks (CNN) can learn the mapping relationship from low-light to normal-light conditions [9–12], while Generative Adversarial Networks (GAN) further enhance the authenticity of detail generation through adversarial training. However, existing methods still face three challenges: first, low-light images cannot be used to effectively restore image details; second, insufficient utilization of image features leads to inadequate contrast enhancement; third, key features cannot be fully fused during the feature extraction process.

3. Methodology

To achieve detail enhancement of low-light images and color restoration of dark areas, this paper proposes a wavelet-based local and global feature fusion network.

3.1. Network architecture

The architecture of the wavelet local and global feature fusion network (LAGN) designed in this paper is shown in **Figure 1**. It includes three modules: the Multi-scale Feature Module (MFM) for local feature extraction, the Double-branch Transformer Module (DTM) for global feature extraction, and the Fusion Module (FM) for feature fusion.

For modeling convenience, the input image is denoted as $LR \in \mathbb{R}^{w \times h \times c}$, and the output image as $HR \in \mathbb{R}^{w \times h \times c}$. Here, w, h, and c represent the length, width, and number of bands of the image, respectively. To enhance image details, the low-light image is first subjected to wavelet transform, and then its number of channels is adjusted to match that of the input image, resulting in a wavelet image $LR\_W \in \mathbb{R}^{w \times h \times c}$ which contains rich detailed information. Subsequently, the low-light image and the wavelet image are stacked to form $I = C\left(LR, LR\_W\right)$, which serves as the input to LAGN.

In LAGN, the input features pass through MFM to obtain local features and through DTM to obtain global features. These local and global features are then adaptively fused in FM to acquire the most effective features for low-light image enhancement. This stage is progressive and can be divided into four steps, which are expressed as follows.

$$I_{i+1} = \begin{cases} F_{FM}\left(F_{MFM}\left(I_i\right), F_{DTM}\left(I_i, LR\_W\right)\right), i = 1 \\ F_{FM}\left(F_{MFM}\left(I_i + F_{MFM}\left(I_{i-1}\right)\right), F_{DTM}\left(I_i + F_{DTM}\left(I_{i-1}, LR\_W\right), LR\_W\right)\right), 1 < i \le 4 \end{cases} \tag{1.1}$$

Here, $F_{MFM}(\cdot)$, $F_{DTM}(\cdot)$, and $F_{FM}(\cdot)$ represent the Multi-scale Feature Module, Double-branch Transformer Module, and Fusion Module, respectively. The final enhanced image is obtained through the following reconstruction process:

$$HR = F_c(I_4) \tag{1.2}$$

where $F_c(\cdot)$ denotes the reconstruction module, which is implemented via a 3×3 convolution.

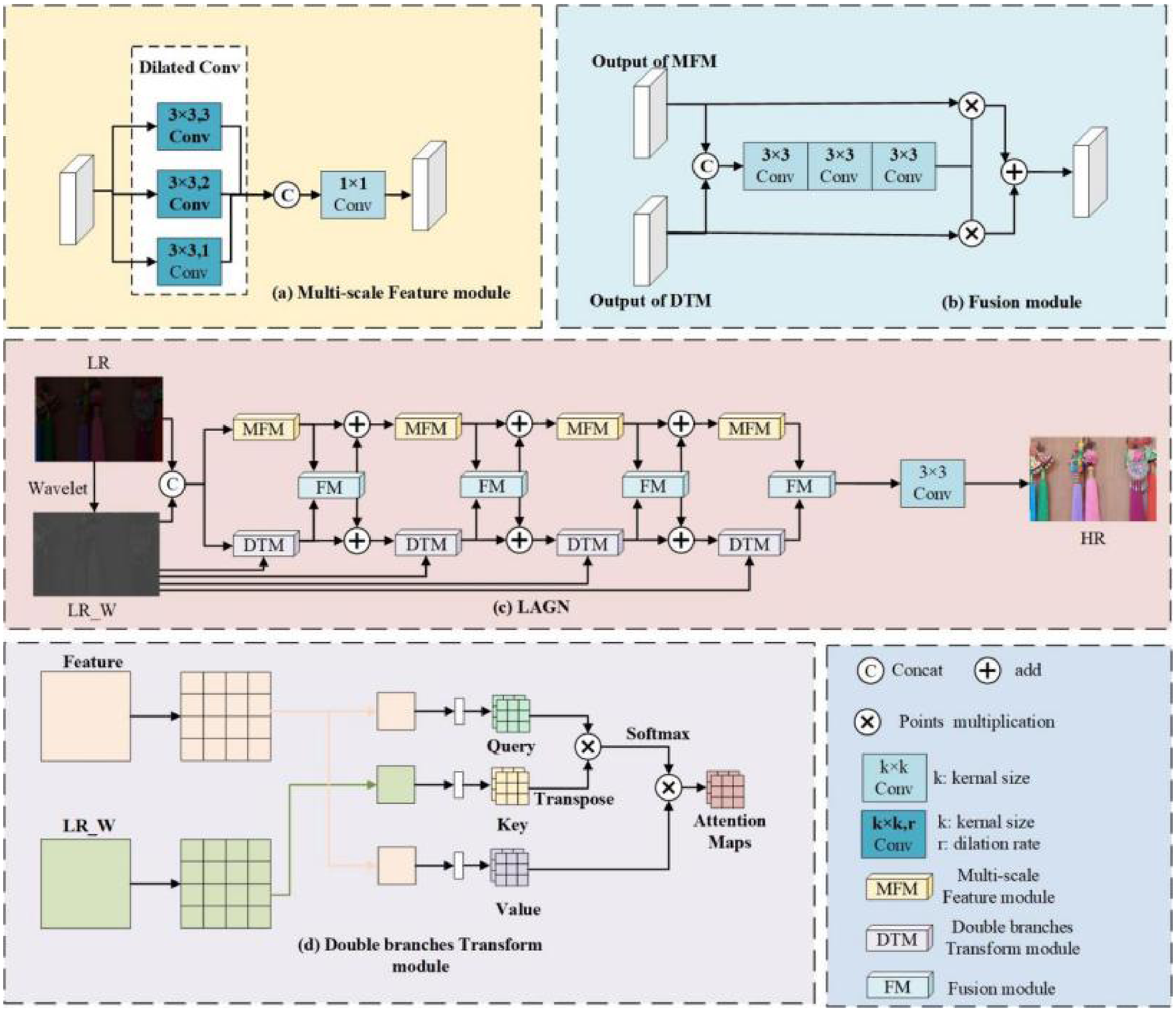

Figure 1. Overall framework diagram. (a) Multi-scale Feature Module (MFM); (b) Double-branch Transformer Module (DTM); (c) Wavelet Local and Global Feature Fusion Network (LAGN); (d) Fusion Module (FM)

3.2. Multi-scale Feature Module (MFM)

To extract local features of the image, a multi-scale feature module is designed in this paper, as shown in **Figure 1(a)**. It uses dilated convolutions with different receptive fields to extract multi-scale features while minimizing computational complexity as much as possible and optimizing learning efficiency. Specifically, this module consists of three parallel branches of dilated convolutions with dilation rates of 1, 2, and 3, respectively, for multi-scale feature extraction. Features from different receptive fields are then stacked, followed by a 1×1 convolution to obtain integrated features.

3.3. Double-branch Transformer Module (DTM)

To extract global features for enhancing image details, during the fusion process, this paper takes the wavelet image and fused features as inputs and designs a double-branch Transformer module. As shown in **Figure 1(b)**, by modeling the relationship between these two types of inputs, the module fully learns the global features of the image and enhances the global details of the image.

3.4. Feature Fusion Module (FM)

To fully fuse the local and global features obtained from the image, a fusion module is designed in this paper, as shown in **Figure 1(d)**. This module is composed of three layers of 3×3 convolutions. Weighting factors for the fused features are obtained through the convolutional layers, and these weighting factors are then applied to the two inputs. Finally, the weighted fused features are element-wise multiplied with the input features and summed to serve as the fused features.

3.5. Loss function

To supervise the training process, a hybrid loss function combining SSIM (Structural Similarity Index) loss and VGG (Visual Geometry Group) loss is adopted in this paper, which is expressed as follows:

$$L_{total} = \lambda_1 L_{ssim} + \lambda_2 L_{vgg} \tag{1.3}$$

where λ_1 and λ_2 are both set to 1.

4. Experiments and analysis

To evaluate the performance of the proposed method, experiments were conducted on two datasets (**Table 1**). For the fairness of the experiments, all experiments were performed on a computer equipped with an RTX 2070 GPU. The Adam optimizer was used with its default parameters. The training batch size was set to 16, and the size of the input image patches was 96×96. The learning rate was set to 2×10^{-4}.

Table 1. Numerical results on LOL and MIT5K datasets

Method	LOL		MIT5K	
	PSNR	SSIM	PSNR	SSIM
MF	18.23	0.569	17.48	0.780
NPE	17.20	0.529	17.20	0.771
SRIE	14.25	0.537	19.48	0.799
LIME	17.37	0.545	14.54	0.750
KinD	20.38	0.804	21.84	0.794
LPNet	21.70	0.775	24.55	0.902
Zero-DCE++	17.42	0.781	20.21	0.805
LADN	23.31	0.815	24.95	0.901

LAGN was compared with traditional methods (MF [5], NPE [6], SRIE [7], LIME [8]) and deep learning-based methods (KinD [10], LPNet [11], Zero-DCE++[12]) in the following aspects.

As shown in **Table 1**, by comparing these methods, it can be observed that LADN achieves the best performance on both the MIT5K and LOL datasets. The results presented in **Figures 2** and **3** also demonstrate that LADN delivers the optimal subjective effects, and it has considerable advantages in both enhancing image details and restoring colors in dark areas.

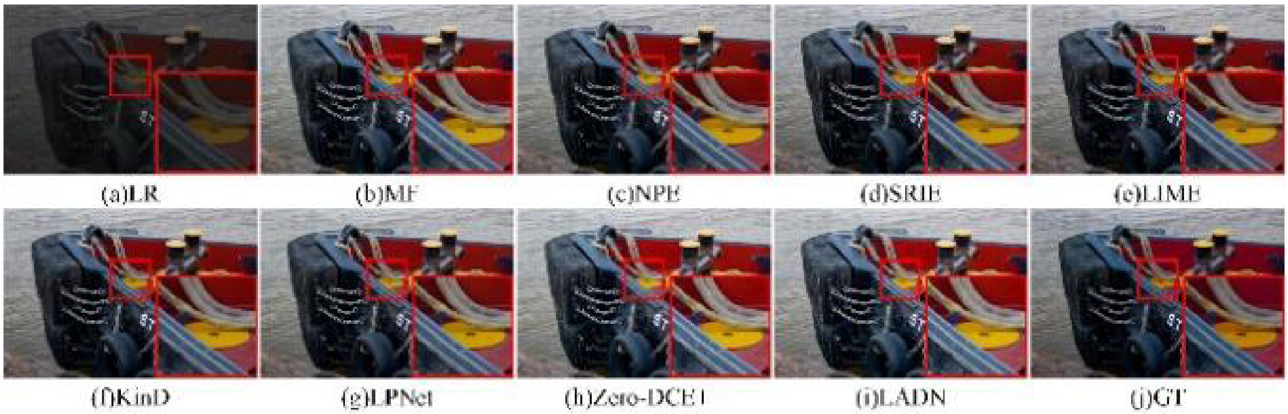

Figure 2. Results of comparative experiments on the MIT5K dataset

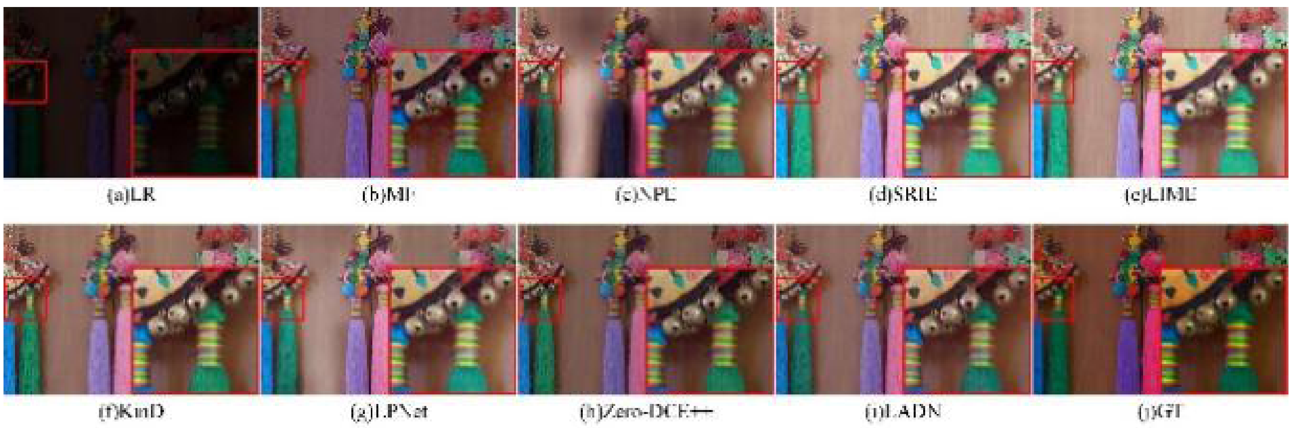

Figure 3. Results of comparative experiments on the LOL dataset

5. Conclusion

To address issues such as insufficient details and contrast in low-light images, this paper proposes a wavelet-based local and global feature fusion network. Through operations including wavelet transform and local-global feature fusion, the network realizes image detail enhancement and color restoration in dark areas.

Future research can be expanded in the following aspects:

(1) Conduct research on the generalization of low-light image enhancement and extend it to underwater image enhancement.

(2) Explore unsupervised image enhancement methods to reduce dependence on datasets.

Disclosure statement

The authors declare no conflict of interest.

References

[1] Pizer SM, Amburn EP, Austin JD, et al., 1987, Adaptive Histogram Equalization and Its Variations. Computer Vision, Graphics, and Image Processing, 39(3): 355–368.

[2] Narendra PM, Fitch RC, 1978, Real-Time Adaptive Contrast Enhancement For Imaging Sensors, Proceedings of SPIE - The International Society for Optical Engineering, 137.

[3] Reza AM, 2004, Realization of the Contrast Limited Adaptive Histogram Equalization (CLAHE) for Real-Time Image Enhancement. Journal of VLSI Signal Processing Systems for Signal, Image and Video Technology, 38: 35–44.

[4] Kim YT, 1997, Contrast Enhancement Using Brightness Preserving Bi-Histogram Equalization. IEEE Transactions on Consumer Electronics, 43(1): 1–8.

[5] Fu X, Zeng D, Huang Y, et al., 2016, A Fusion-Based Enhancing Method for Weakly Illuminated Images. Signal Processing, 129: 82–96.

[6] Wang S, Zheng J, Hu HM, et al., 2013, Naturalness Preserved Enhancement Algorithm for Non-Uniform Illumination Images. IEEE Transactions on Image Processing, 22(9): 3538–3548.

[7] Fu X, Zeng D, Huang Y, et al., 2016, A Weighted Variational Model for Simultaneous Reflectance and Illumination Estimation, 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE.

[8] Guo X, Li Y, Ling H, 2016, LIME: Low-Light Image Enhancement via Illumination Map Estimation. IEEE Transactions on Image Processing, 26(2): 982–993.

[9] Lore KG, Akintayo A, Sarkar S, 2017, LLNet: A Deep Autoencoder Approach to Natural Low-Light Image Enhancement. Pattern Recognition, 61: 650–662.

[10] Zhang Y, Zhang J, Guo X, 2019, Kindling the Darkness: A Practical Low-Light Image Enhancer, Proceedings of the 27th ACM International Conference on Multimedia, 1632–1640.

[11] Li J, Li J, Fang F, et al., 2020, Luminance-Aware Pyramid Network for Low-Light Image Enhancement. IEEE Transactions on Multimedia, 23: 3153–3165.

[12] Li C, Guo C, Chen CL, 2021, Learning to Enhance Low-Light Image via Zero-Reference Deep Curve Estimation. IEEE Transactions on Software Engineering, (99).