

OPEN ACCESS

EDITED BY
Yakun Ju,
University of Leicester, United Kingdom

REVIEWED BY
Hao Wang,
China University of Petroleum, China
Yunyi Yan,
Xidian University, China

*CORRESPONDENCE
Xiuping Liu
✉ liuxiuping8@126.com

RECEIVED 18 February 2025
ACCEPTED 02 July 2025
PUBLISHED 28 July 2025

CITATION
Han L, Liu X, Xu T and Wang Y (2025)
Hybrid Mamba for amphibious Limulidae
low-light image enhancement.
Front. Mar. Sci. 12:1578735.
doi: 10.3389/fmars.2025.1578735

COPYRIGHT
© 2025 Han, Liu, Xu and Wang. This is an
open-access article distributed under the terms
of the [Creative Commons Attribution License
\(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction
in other forums is permitted, provided the
original author(s) and the copyright owner(s)
are credited and that the original publication
in this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

Hybrid Mamba for amphibious Limulidae low-light image enhancement

Lili Han, Xiuping Liu*, Tao Xu and Yuangan Wang

College of Electronics and Information Engineering, Beibu Gulf University, Qinzhou, Guangxi, China

Obtaining high-quality images of Limulidae in amphibious environments is a challenging task due to insufficient light and the complex optical properties of water, such as light absorption and scattering, which often result in low contrast, color distortion, and blurring. These issues severely impact applications like nocturnal biological monitoring, underwater archaeology, and resource exploration. Traditional image enhancement methods struggle with the complex degradation of such images, but recent advancements in deep learning have shown promise. This paper proposes a novel method for amphibious low-light image enhancement based on hybrid Mamba, which integrates wavelet transform, Discrete Cosine Transform (DCT), and Fast Fourier Transform (FFT) within the Mamba framework. Wavelet transform effectively decomposes images at multiple scales, capturing feature information at different frequencies and excelling in noise removal and detail preservation, whereas DCT concentrates and compresses image energy, aiding in the restoration of high-frequency components and improving clarity. FFT provides efficient frequency domain analysis, accurately locating key information in the image spectrum and enhancing image quality. Mamba, as an emerging optimization strategy, offers unique computational characteristics and optimization capabilities, making it well suited for this task. The main contributions include the construction of the amphibious low-light image dataset (ALID) in collaboration with the Beibu Gulf Key Laboratory of Marine Biodiversity Conservation and the introduction of the hybrid Mamba method. Extensive experiments on the ALID dataset demonstrate that our method outperforms state-of-the-art approaches in both subjective visual assessment and quantitative analysis, achieving superior results in brightness enhancement and detail reconstruction, thus paving new paths for amphibious low-light image processing and promoting further development in related industries and research.

KEYWORDS

amphibious Limulidae, SSM, Mamba, low-light image enhancement, UIE

1 Introduction

Obtaining high-quality Limulidae images in amphibious environments has always been an extremely challenging task. Due to insufficient light in low-light conditions and the complex optical properties of water, such as absorption and scattering of light, amphibious low-light images often suffer from issues like low contrast, color distortion, and blurring. These problems severely impact subsequent analysis and applications, such as nocturnal biological monitoring, underwater archaeology, and resource exploration in amphibious environments.

Traditional image enhancement methods (Han et al., 2023; Wang W. et al., 2024; Zheng et al., 2024) have limitations when dealing with the complex degradation of amphibious low-light Limulidae images. Zhang et al. (Zhuang et al., 2024) proposed a color correction strategy guided by an attenuation map, which employs a global contrast enhancement strategy optimized by maximum information entropy to enhance the global contrast of color-corrected images. Additionally, a local contrast enhancement strategy optimized by fast integration is applied to obtain locally enhanced images. To leverage the complementary nature of globally and locally enhanced images, a weighted wavelet visual perception fusion strategy is introduced. This strategy fuses high-frequency and low-frequency components of images at different scales to produce high-quality underwater images. Wang Y. et al. (2024) integrated multi-level wavelet transforms and color compensation priors into a multi-stage enhancement framework. Each stage includes a multi-level wavelet enhancement module, a color compensation prior extraction module, and a color filter with prior-aware weights.

In recent years, the technology for enhancing amphibious low-light images has continuously evolved, achieving significant progress from traditional methods to deep learning-based approaches (Wang et al., 2023; Wang et al., 2024a; Wang et al., 2025; Zhuang et al., 2025). Wang et al. (2024b) developed a novel reinforcement learning framework that selects a series of image enhancement methods and configures their parameters in a self-organizing manner to achieve underwater image enhancement (UIE). Peng L. et al. (2023) constructed a large-scale underwater image dataset and proposed a U-shape transformer that integrates a channel multi-scale feature fusion transformer module and a spatial global feature modeling transformer module, specifically designed for UIE tasks. This enhances the network's attention to parts of the image that suffer more severe attenuation in color channels and spatial regions. Zhou et al. (2023) introduced a novel multi-feature UIE method on the basis of an embedded fusion mechanism. By preprocessing to obtain high-quality images, they incorporated a white balance algorithm and contrast-limited adaptive histogram equalization algorithm and adopted multi-path inputs to extract rich features from multiple perspectives.

The paper proposes a method based on multiple transformation domain Mamba for amphibious low-light image enhancement. This method integrates wavelet transform, DCT, and FFT within the Mamba framework to enhance amphibious low-light images. Multiple transformation domain techniques are introduced into

the research of amphibious low-light image enhancement. Among them, wavelet transform (Wavelet) can effectively decompose images at multiple scales, capturing feature information at different frequencies, and excels in noise removal and detail preservation; DCT has unique advantages in concentrating and compressing image energy, aiding in the restoration of high-frequency components and improving clarity; FFT provides an efficient means for frequency domain analysis, accurately locating key information in the image spectrum and assisting in enhancing image quality. Mamba, as an emerging optimization strategy or algorithmic component, possesses unique computational characteristics and optimization capabilities. Combining it with Wavelet, DCT, and FFT is expected to break through the current bottlenecks in amphibious low-light image enhancement technology. This study aims to explore the application potential of the innovative combination method based on multiple transformation domain Mamba in amphibious low-light image enhancement. Through rigorous experimental design and data analysis, the practical effects of this method on improving the quality of amphibious low-light images are evaluated, paving new paths for the field of amphibious low-light image processing and promoting further development in related industries and research.

The main contributions of this paper are as follows:

1. Construction of a dataset: In partnership with the Beibu Gulf Key Laboratory of Marine Biodiversity Conservation, we jointly established amphibious low-light image dataset (ALID), and extensive experiments are conducted on this dataset to validate the effectiveness of the proposed method.
2. Proposal of the multiple transformation domain Mamba method: A novel combination of wavelet transform, DCT, and FFT is introduced for amphibious low-light image enhancement.

2 Related works

2.1 Mamba-based UIE

The Mamba-based amphibious low-light image enhancement method demonstrates strong adaptability to complex amphibious low-light environments while maintaining lightweight and high efficiency. It can address various degradation scenarios such as color distortion, blurring, and insufficient illumination and can be flexibly extended to different tasks and models, providing an efficient and reliable solution for amphibious low-light vision tasks.

Dong et al. (2024) proposed O-mamba, which adopts an O-shaped dual-branch network to model spatial information and cross-channel information separately, leveraging the efficient global receptive field of a state-space model optimized for underwater images. Zhang et al. (2024) introduced Mamba-UIE, which divides the input image into four components: underwater scene radiance, direct transmission map, backward scattering transmission map, and global background light. These

components are recombined according to a revised underwater image formation model, and a reconstruction consistency constraint is applied between the reconstructed image and the original image, thereby achieving effective physical constraints during UIE. Chang et al. (2025) proposed a Mamba-enhanced spectral attention wavelet network for effective underwater image restoration. This network includes three modules: a spatial detail enhancement encoder, a state-space model for local information compensation, and a spectral cross-attention residual module.

2.2 Hybrid Mamba-based UIE

Hybrid Mamba methods have demonstrated significant advantages in the field of amphibious low-light image processing. By integrating multiple frequency analysis techniques and leveraging the efficient state-space model of the Mamba architecture, these methods not only enhance the quality of amphibious low-light image enhancement but also provide more efficient solutions for complex visual tasks, showcasing their immense potential in practical applications.

Tan et al. (2024) proposed the WalMaFa model on the basis of wavelet and Fourier adjustments, which consists of a wavelet-based Mamba module and a fast Fourier adjustment module. The model employs an encoder-latent layer-decoder structure to achieve end-to-end transformation. Bai et al. (2024) introduced the RetinexMamba architecture, which retains the physical intuitiveness of traditional Retinex methods while integrating the deep learning framework of Retinexformer and utilizing the computational efficiency of state-space models (SSMs) to improve processing speed. This architecture incorporates innovative illumination estimators and damage-repair mechanisms to maintain image quality during enhancement. Zou et al. (2024) proposed the Wave-Mamba method on the basis of the wavelet domain, which improves SSMs with a low-frequency state-space block (LFSSBlock) focused on restoring information in low-frequency sub-bands, and a high-frequency enhancement block (HFEBlock) to process high-frequency sub-band information. By using enhanced low-frequency information to correct high-frequency information, this method effectively restores accurate high-frequency details.

3 Preliminaries

3.1 2D-DWT-DCT

2D-DWT is used on the input amphibious low-light image I , obtaining four sub-bands: LL , LH , HL , and HH . LL contains the low-frequency information of the image (brightness and overall structure). LH , HL , and HH contain high-frequency detail information.

LL contains the primary brightness information of the image, enhancing low-light image primarily involves increasing the brightness and contrast of LL . To more effectively process the LL

sub-band, it can be further divided into smaller blocks (e.g., 8×8 blocks), and DCT is used for each block:

$$DCT(\mathbf{LL}_{block}) = \mathbf{C}(u, v) \quad (1)$$

In Equation 1, the low-frequency coefficients (e.g., $C(0, 0)$, $C(0, 1)$, $C(1, 0)$) carry the energy and brightness information of the image. By adjusting these low-frequency coefficients, image brightness enhancement and contrast improvement can be achieved (Equation 2).

$$\mathbf{C}'(u, v) = \gamma \mathbf{C}(u, v) \text{ for low frequencies} \quad (2)$$

where $\gamma > 1$ is the amplification factor.

The adjusted coefficient \mathbf{C}' is transformed back to the spatial domain using the inverse DCT, resulting \mathbf{LL}' :

$$\mathbf{LL}'_{block} = IDCT(\mathbf{C}'(u, v)) \quad (3)$$

In Equation 3, \mathbf{LL}' is combined with the original high-frequency sub-bands LH , HL , and HH . Using the inverse wavelet transform (IWT), the enhanced image I' is reconstructed.

3.2 FFT

In amphibious Limulidae low-light image enhancement, FFT can effectively improve the image's contrast, clarity, and illumination distribution by processing its frequency components. Suppose 2D image $I(x, y)$ with scale $M \times N$, Discrete Fourier Transform (DFT) is defined as Equation 4.

$$F(u, v) = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} I(x, y) \cdot e^{-j2\pi(\frac{ux}{M} + \frac{vy}{N})} \quad (4)$$

where $I(x, y)$ is the pixel value in spatial domain, $F(u, v)$ is the transformed result in the frequency domain, u, v are the frequency coordinates in the frequency domain, and $e^{-j2\pi(\frac{ux}{M} + \frac{vy}{N})}$ is the complex exponential basis function.

FFT decomposes the image into two parts: the magnitude spectrum and the phase spectrum. The magnitude spectrum represents the intensity of each frequency component in the image, whereas the phase spectrum represents the phase information of the frequency components. The result is a complex matrix containing both the magnitude spectrum $|F(u, v)|$ and the phase spectrum $\arg(F(u, v))$ in Equation 5.

$$\begin{aligned} |F(u, v)| &= \sqrt{\text{Re}(F(u, v))^2 + \text{Im}(F(u, v))^2} \\ \arg(F(u, v)) &= \arctan\left(\frac{\text{Im}(F(u, v))}{\text{Re}(F(u, v))}\right) \end{aligned} \quad (5)$$

Using the inverse FFT (IFFT), the enhanced frequency domain result $F'(u, v)$ is converted back to the spatial domain Equations 6:

$$I'_f(x, y) = IFFT(F'(u, v)) \quad (6)$$

The result of the IFFT contains minor imaginary parts. Because the input image is real-valued, the absolute value or the real part can be directly taken. The enhanced image is $I'_f(x, y)$.

Essentially, both 2D-DWT-DCT and FFT decompose the image into two components: one is more sensitive to color brightness, and the other is more sensitive to texture details. The low-frequency LL component in 2D-DWT-DCT is clearly superior to FFT amplitude component in terms of color brightness, whereas FFT phase component is evidently better at capturing texture details than the high-frequency LH , HL , and HH components in 2D-DWT-DCT. This paper proposes using the low-frequency LL component of 2D-DWT-DCT to enhance global color brightness while employing the phase component of FFT to restore local texture details and smoothness.

4 Methodology

4.1 State-space model

The structured state-space model (SSM) (S4) is based on the theoretical foundation of continuous systems, explaining the dynamic relationship between the input signal $x(t)$ and the output signal $y(t)$ within a time-invariant linear framework. Essentially, this model maps one-dimensional function or sequence $x(t) \in \mathbb{R}^L$ to the corresponding output $y(t) \in \mathbb{R}^L$. The implicit latent state $h(t) \in \mathbb{R}^N$. The system can be succinctly represented in the form of ordinary differential equations (ODEs) as follows (Equations 7, 8):

$$h'(t) = \mathbf{A}h(t) + \mathbf{B}x(t) \tag{7}$$

$$y(t) = \mathbf{C}h(t) + \mathbf{D}x(t) \tag{8}$$

where $\mathbf{A} \in \mathbb{R}^{N \times N}$, $\mathbf{B} \in \mathbb{R}^{N \times 1}$, $\mathbf{C} \in \mathbb{R}^{1 \times N}$ are parameters of the state size N , and $\mathbf{D} \in \mathbb{R}^1$ represents a jump connection.

In the process of integrating SSMs into deep learning models, the ODE process is discretized to match the sampling frequency of the latent signals in the input data. This discretization process typically uses the zero-order hold (ZOH) method, combined with a time scale parameter Δ , to convert the continuous parameters A and B into discrete forms $\bar{\mathbf{A}}$ and $\bar{\mathbf{B}}$. The definitions are (Equations 9–12).

$$h'_t = \bar{\mathbf{A}}h_{t-1} + \bar{\mathbf{B}}x_t \tag{9}$$

$$y_t = \mathbf{C}h_t + \mathbf{D}x_t \tag{10}$$

$$\bar{\mathbf{A}} = e^{\Delta \mathbf{A}} \tag{11}$$

$$\bar{\mathbf{B}} = (\Delta \mathbf{A})^{-1}(e^{\Delta \mathbf{A}} - \mathbf{I}) \cdot \Delta \mathbf{B} \tag{12}$$

where $\Delta \in \mathbb{R}^D$, $\mathbf{B}, \mathbf{C} \in \mathbb{R}^{D \times N}$.

Mamba continues to evolve, allowing parameters B and C to be adjusted based on input, achieving dynamic feature representation. Essentially, Mamba employs a recursive structure that can process and retain information from long sequences, ensuring that more pixels can participate in the recovery process. Additionally, Mamba utilizes parallel scanning algorithms, reflecting the advantages of

parallel processing, and improving the efficiency of training and inference processes. Based on the Mamba architecture, combined with hybrid transformation methods, the channel-wise Mamba approach further explores changes in color and brightness within the channel dimension.

4.2 Hybrid Mamba

2D-DWT-DCT can effectively extract low-frequency components, which is of vital importance for capturing global color and brightness information. The channel-wise Mamba module compensates for missing spatial information in the channel dimension. By leveraging its linear analysis capability over long-distance sequences, it can effectively capture global contextual information while maintaining a low computational cost. Compared with transformers, the Mamba module exhibits lower computational complexity when processing long sequences (see Figure 1).

WDMB follows the efficient token mixer of transformer, which can be represented as (Equations 13):

$$\begin{aligned} I' &= \text{WDM}(\text{LN}(x)) + x, \\ I'' &= \text{FFTB}(\text{LN}(I')) + I' \end{aligned} \tag{13}$$

where WDM denotes the Mamba operation based on Wavelet and DCT, and LN represents the LayerNorm layer normalization operation.

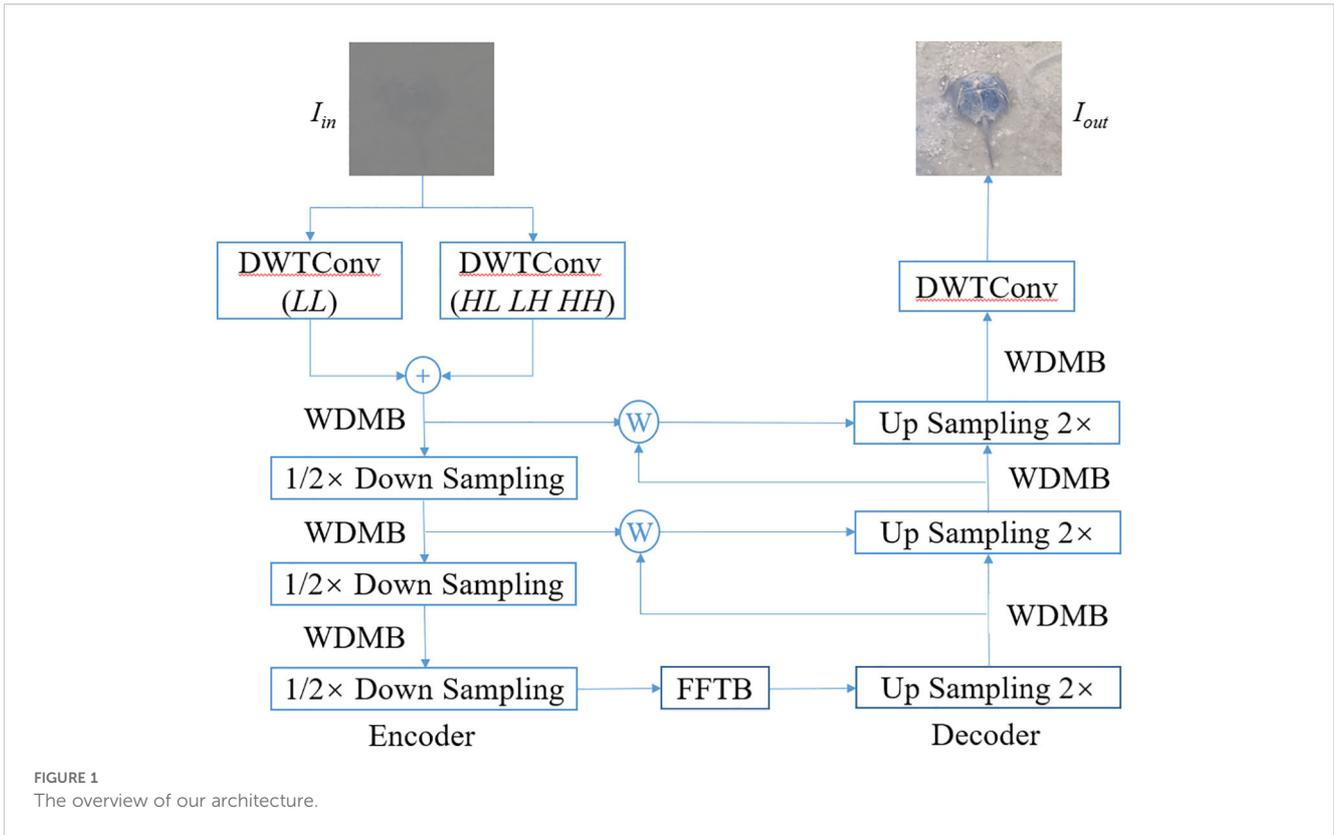
Given the input feature map $F_{in} \in \mathbb{R}^{H \times W \times C}$, it is decomposed into four parts: low-frequency sub-bands $F_{LL} \in \mathbb{R}^{\frac{H}{2} \times \frac{W}{2} \times C}$ and high-frequency sub-bands $\{F_{LH}, F_{HL}, F_{HH}\} \in \mathbb{R}^{3 \times \frac{H}{2} \times \frac{W}{2} \times C}$. For the low-frequency sub-bands, F_{LL} is fed into a 3×3 convolutional activation block, and then, the height and width dimensions are merged to generate $F_{LL}^C \in \mathbb{R}^{B,C,L}$, where $L = H \times W$, which integrates spatial features while retaining the channels for subsequent modeling on the channel dimension. After Mamba and 3×3 convolution, we obtain the low-frequency enhanced output $F'_{LL} \in \mathbb{R}^{\frac{H}{2} \times \frac{W}{2} \times C}$ through a reconstruction operation. For the processing of high-frequency sub-bands, F_{LH}, F_{HL}, F_{HH} are directly fed into a 3×3 convolution to generate $\{F'_{LH}, F'_{HL}, F'_{HH}\} \in \mathbb{R}^{3 \times \frac{H}{2} \times \frac{W}{2} \times C}$. F'_{LL} and $\{F'_{LH}, F'_{HL}, F'_{HH}\}$ are reconstructed into the image space through the IWT operation, resulting in $F_{out} \in \mathbb{R}^{H \times W \times C}$ (see Figure 2).

4.3 Loss function

The loss function L is composed of three parts: 2D-DWT-DCT loss (L_{wd}), FFT loss (L_f), and the Charbonnier loss (L_c) in (Equations 14–17).

$$L_{wd} = \|T_{LL'} - G_{LL}\|_2 \tag{14}$$

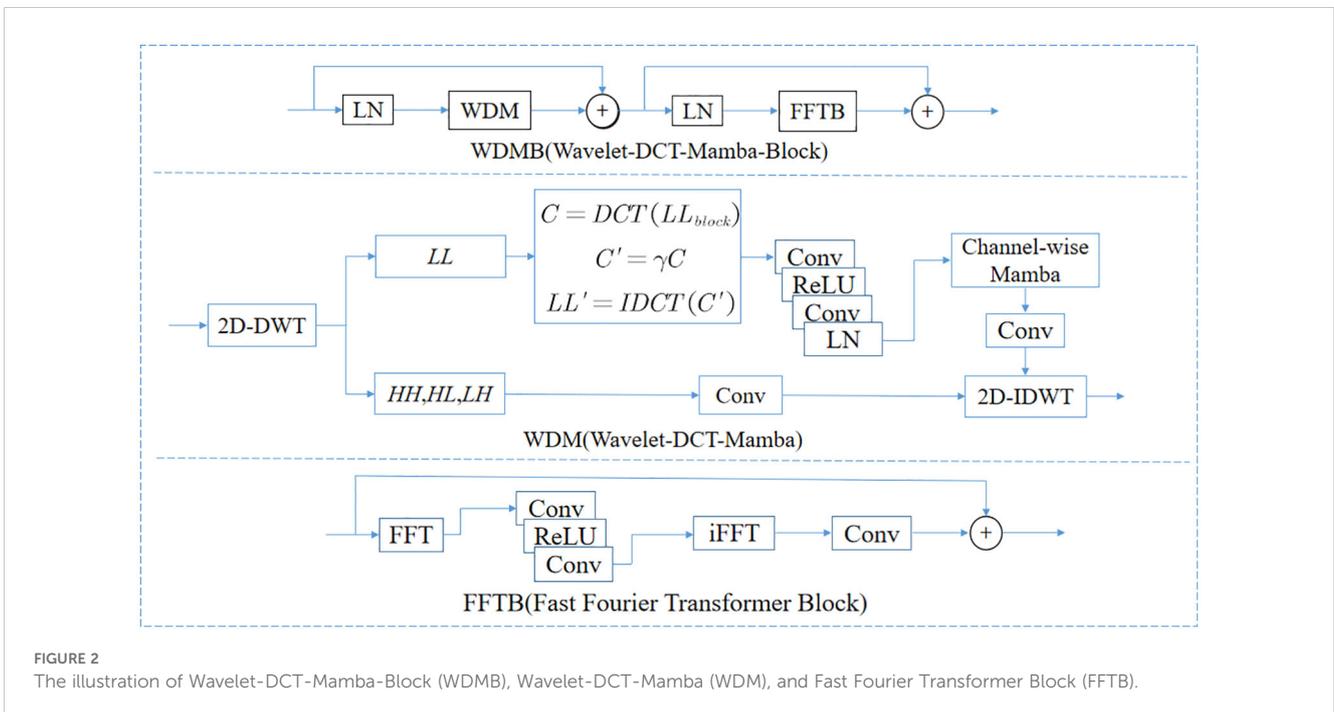
$$L_f = \left\| P(I'_f) - P(G) \right\|_2 \tag{15}$$

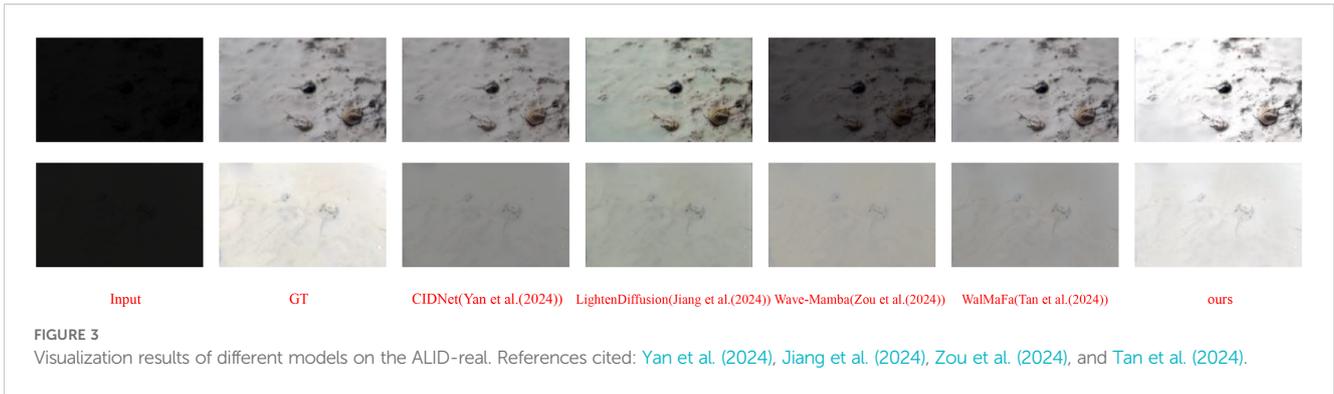


$$L_c = \sqrt{\|R_e - G\|_2 + \epsilon^2} \tag{16}$$

$$L = L_{wd} + L_f + L_c \tag{17}$$

Because the low-frequency component plays a crucial role in enhancing global image brightness, L_{wd} is formed by comparing the low-frequency branch output $T_{LL'}$ and the ground-truth low-frequency branch G_{LL} . Due to the significant texture-enhancing effect of the phase





component in Fourier transforms, L_f is constructed using the phase of the transformed image $P(I'_f)$ and the phase of the ground-truth image $P(G)$. L_c is composed of the global image enhancement result R_e and the ground-truth image G , and the default value of ϵ is 10^{-3} .

5 Experiments

5.1 Datasets and environment

In collaboration with the Beibu Gulf Key Laboratory of Marine Biodiversity Conservation, we have jointly constructed the ALID. This dataset comprises 485 pairs of low-light and highlight images of outdoor scenes at varying scales. ALID can be categorized into two types: ALID-real and ALID-synthetic. The ALID-real dataset was captured in real-world scenarios by adjusting ISO and exposure times. The ALID-synthetic dataset, on the other hand, was generated by synthesizing low-light images from original images through the analysis of illumination distribution in low-light conditions.

We trained the model on RTX 4090 with batch size of 4. The images are resized to 128×128 . We employ the AdamW optimizer (momentum of 0.9) and conducted mixed-precision training for 5,000 epochs. The initial learning rate is set to 8×10^{-4} , gradually reduced to 1×10^{-6} by the cosine annealing schedule.

5.2 Comparisons with state-of-the-art methods

We performed experiments to evaluate the image enhancement performance of two distinct architectures, transformer and Mamba, with comparisons conducted from two perspectives: subjective visual assessment and theoretical quantitative analysis.

The visualization results in Figure 3 indicate that the amphibious Limulidae low-light image enhancement outcomes produced by CIDNet exhibit a limited overall brightness improvement and suboptimal reconstruction of certain details within the images. LightenDiffusion demonstrates superior performance, with some details preserved clearly. WalMaFa excels in retaining image details, yet it offers only a modest enhancement in overall brightness. Wave-Mamba provides the least improvement in overall brightness and also retains the fewest details. Ours achieved the most significant enhancement in terms of overall brightness improvement and image detail reconstruction, and it is superior to other methods in noise suppression. However, ours performed inadequately in preserving texture details.

The experimental results in Figure 4 demonstrate that LightenDiffusion performs optimally in terms of chromaticity preservation, yet it loses textural details during the image smoothing process. Wave-Mamba exhibits the poorest performance in brightness enhancement. WalMaFa alters the



TABLE 1 Comparison of quantitative results on ALID.

Methods	ALID-real	ALID-synthetic	AVG	Speed (s)	FLOPS (G)	Params (M)
	PSNR/SSIM	PSNR/SSIM	PSNR/SSIM	—	—	—
CIDNet	32.98/0.946	32.47/0.937	32.73/0.942	—	7.57	1.88
FlowIE (Zhu et al., 2024)	18.39/0.663	17.14/0.712	17.77/0.688	—	—	—
LightenDiffusion	33.17/0.957	33.21/0.942	33.19/0.950	0.314	—	—
RAUNE (Peng W. et al., 2023)	20.16/0.812	19.28/0.849	19.72/0.831	—	—	—
Wave-Mamba	25.29/0.854	24.56/0.861	24.93/0.858	—	—	2.3
WalMaFa	33.26/0.951	32.39/0.943	32.83/0.947	0.068	14.41	11.09
Ours	37.32/0.985	35.29/0.971	36.31/0.978	0.072	14.62	11.23

The bold values are the Optimal values.

TABLE 2 Ablation studies of different components.

Experiment	2D-DWTBlock	2D-DWT-DCTBlock	FFTBlock	PSNR	SSIM
Setting 1	✓			21.32	0.749
Setting 2		✓		27.51	0.921
Setting 3	✓		✓	33.45	0.963
ours		✓	✓	36.31	0.978

image color and shows limited effectiveness in brightness improvement. CIDNet suffers from color distortion during image reconstruction and displays poor fine-grained detail representation. Ours achieves the best performance in brightness enhancement but fails to adequately preserve the original image colors. In noise suppression, it slightly lags behind LightenDiffusion.

Table 1 presents a quantitative analysis of low-light image enhancement results for horseshoe crabs using methods based on Mamba and transformer architectures. The results indicate that FlowIE has the poorest overall performance 17.77/0.688, followed by RAUNE at 19.72/0.831. The Wave-Mamba model shows limited image enhancement effectiveness 24.93/0.858. In contrast, CIDNet, LightenDiffusion, and WalMaFa models demonstrate better enhancement results, achieving 32.73/0.942, 33.19/0.950, and 32.83/0.947, respectively. Our method outperforms the others, delivering the best image enhancement performance across different types of datasets and the entire dataset 36.31/0.978.

5.3 Ablation study

The ablation experiment results of the models are presented in Table 2. Setting 1 yields the poorest performance with a PSNR/SSIM of 21.32/0.749, whereas setting 2 shows relatively inferior results at 27.51/0.921. Setting 3 demonstrates better performance with a PSNR/SSIM of 33.45/0.963. Our hybrid-domain model achieves the optimal performance, attaining a PSNR/SSIM of 36.31/0.978.

6 Conclusion

In this paper, we proposed a novel hybrid mamba method for amphibious low-light image enhancement, integrating wavelet transform, DCT, and FFT within the Mamba framework. Our approach effectively addresses the challenges of low contrast, color distortion, and blurring in amphibious environments by leveraging the strengths of multiple transformation domains. The experimental results on the ALID demonstrate that our method significantly outperforms state-of-the-art techniques in both subjective visual quality and quantitative metrics, achieving superior brightness enhancement and detail reconstruction. The approach highlights the potential of combining traditional techniques with advanced optimization strategies like Mamba for complex image enhancement tasks. Future work will focus on further optimizing the computational efficiency of the model and exploring its applicability to other low-light vision tasks in diverse environments. This study not only advances the field of amphibious low-light image processing but also provides a robust framework for future research and practical applications in related domains.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material. Further inquiries can be directed to Lili Han, hanlili_can@126.com.

Author contributions

LH: Conceptualization, Data curation, Formal Analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. XL: Visualization, Writing – review & editing. TX: Data curation, Formal Analysis, Writing – review & editing. YW: Project administration, Resources, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This work was supported by the Natural Science Foundation of Guangxi Zhuang Autonomous Region (Grant No. 2025GXNSFH069149), the Major Guangxi Science and Technology (Grant No. Guike AA24206003).

Acknowledgments

The Beibu Gulf Key Laboratory of Marine Biodiversity Conservation. We sincerely thank the researchers and technicians at the Beibu Gulf Key Laboratory for their assistance in data collection. Finally, we extend our heartfelt gratitude to the

References

- Bai, J., Yin, Y., He, Q., et al (2024). Retinexmamba: Retinex-based mamba for low-light image enhancement. *International Conference on Neural Information Processing*. Singapore: Springer Nature Singapore, 427–442. doi: 10.1007/978-981-96-6596-9_30
- Chang, B., Yuan, G., and Li, J. (2025). Mamba-enhanced spectral-attentive wavelet network for underwater image restoration. *Engineering Applications of Artificial Intelligence* 143, 109999.
- Dong, C., Zhao, C., Cai, W., and Yang, B. (2024). O-Mamba: O-shape State-Space Model for Underwater Image Enhancement. *arXiv preprint arXiv 2408.12816*. doi: 10.48550/arXiv.2408.12816
- Han, L., Wang, Y., Chen, M., Huo, J., and Dang, H. (2023). Non-local self-similarity recurrent neural network: dataset and study. *Applied Intelligence* 53, 3963–3973.
- Jiang, H., Luo, A., Liu, X., Han, S., and Liu, S. (2024). Lightdiffusion: Unsupervised low-light image enhancement with latent-retinex diffusion models. *European Conference on Computer Vision*, Springer, Cham, 161–179.
- Peng, W., Zhou, C., Hu, R., Cao, J., and Liu, Y. (2023). Raune-Net: a residual and attention-driven underwater image enhancement method. *International Forum on Digital TV and Wireless Multimedia Communications*, Singapore, Springer Nature Singapore, 15–27.
- Peng, L., Zhu, C., and Bian, L. (2023). U-shape transformer for underwater image enhancement. *IEEE Transactions on Image Processing* 32, 3066–3079.
- Tan, J., Pei, S., Qin, W., Fu, B., Li, X., and Huang, L. (2024). Wavelet-based Mamba with Fourier Adjustment for Low-light Image Enhancement. *Proceedings of the Asian Conference on Computer Vision*, 3449–3464.
- Wang, Y., Hu, S., Yin, S., Zeng, D., and Yang, Y. H. (2024). A multi-level wavelet-based underwater image enhancement network with color compensation prior. *Expert Systems with Applications* 242, 122710.
- Wang, H., Köser, K., and Ren, P. (2025). Large foundation model empowered discriminative underwater image enhancement. *IEEE Transactions on Geoscience and Remote Sensing* 63, 5609317.
- Wang, H., Sun, S., Chang, L., Li, H., Zhang, W., Frery, A. C., et al (2024). INSPIRATION: A reinforcement learning-based human visual perception-driven image enhancement paradigm for underwater scenes. *Engineering Applications of Artificial Intelligence* 133, 108411.
- Wang, W., Sun, Y. F., Gao, W., Xu, W. K., Zhang, Y. X., and Huang, D. X. (2024). Quantitative detection algorithm for deep-sea megabenthic organisms based on improved YOLOv5. *Frontiers in Marine Science* 11, 1301024.
- Wang, H., Sun, S., and Ren, P. (2023). Underwater color disparities: Cues for enhancing underwater images toward natural color consistencies. *IEEE Transactions on Circuits and Systems for Video Technology* 34, 738–753.
- Wang, H., Zhang, W., and Ren, P. (2024). Self-organized underwater image enhancement. *ISPRS Journal of Photogrammetry and Remote Sensing* 215, 1–14.
- Yan, Q., Feng, Y., Zhang, C., Wang, P., Wu, P., and Dong, W. (2024). You only need one color space: An efficient network for low-light image enhancement. *arXiv preprint arXiv 2402.05809*. doi: 10.48550/arXiv.2402.05809
- Zhang, S., Duan, Y., Li, D., and Zhao, R. (2024). Mamba-UIE: Enhancing Underwater Images with Physical Model Constraint. *arXiv preprint arXiv 2407.19248*. doi: 10.48550/arXiv.2407.19248
- Zheng, H., Bi, H., Cheng, X., and MC, Benfield (2024). Deep learning for marine science. *Frontiers in Marine Science* 11, 1407053. doi: 10.3389/fmars.2024.1407053
- Zhou, J., Sun, J., Zhang, W., and Lin, Z. (2023). Multi-view underwater image enhancement method via embedded fusion mechanism. *Engineering Applications of Artificial Intelligence* 121, 105946.
- Zhu, Y., Zhao, W., Li, A., Tang, Y., Zhou, J., and Lu, J. (2024). FlowIE: Efficient Image Enhancement via Rectified Flow. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 13–22.
- Zhuang, J., Chen, W., Huang, X., and Yan, Y. (2025). Band Selection Algorithm Based on Multi-Feature and Affinity Propagation Clustering. *Remote Sensing* 17, 193.
- Zhuang, J., Zheng, Y., Guo, B., and Yan, Y. (2024). Globally Deformable Information Selection Transformer for Underwater Image Enhancement. *IEEE Transactions on Circuits and Systems for Video Technology* 35, 19–32. doi: 10.1109/TCSVT.2024.3451553
- Zou, W., Gao, H., Yang, W., and Liu, T. (2024). Wave-mamba: Wavelet state space model for ultra-high-definition low-light image enhancement. *Proceedings of the 32nd ACM International Conference on Multimedia*, 1534–1543.

reviewers for their constructive feedback, which significantly enhanced the quality of this manuscript.

Conflict 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.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.