A Cascade-Based Composite Neural Network for Underwater Image Enhancement

Qiuyue Huang[1], Chaoqun Yang[2*], Qi Yang[2], Linqiang Li[2]
1 Department of Public Basic Education, Liuzhou Institute of Technology, Liuzhou, 545616, Guangxi, China
2 College of Information Science and Engineering, Liuzhou Institute of Technology, Liuzhou, 545616, Guangxi, China
E-mail: huang3636528@163.com
*Corresponding author

Keywords: underwater image enhancement, composite neural network model, routing mechanism, evaluation metric

Received: Aprli 10, 2025

Underwater images are often affected by various degradation phenomena, such as low contrast, blurred details, color distortion, poor clarity, non-uniform illumination, and limited viewing distance. To address these issues, this paper proposes a cascaded composite neural network for underwater image enhancement, which incorporates a deep learning-based routing mechanism. Three individual neural networks, namely UWCNN (UW), Deep Wave-Net (DW), and PUIE-Net (PU), are employed as core components, and a method library is constructed using pairwise superimposed serial composite enhancement models. This framework is designed to enhance degraded underwater images and investigate the performance of the composite models. Experimental evaluations are conducted using metrics including PSNR, SSIM, UIQM, and UCIQE. The results indicate that the representative composite neural network model DW-PU achieves favorable performance with indicators of 20.495 (PSNR), 0.874 (SSIM), 3.270 (UIQM), and 0.897 (UCIQE), outperforming current mainstream underwater image enhancement models in certain aspects. Comparative analysis of images enhanced by multiple methods reveals that, in most underwater scenarios, the DW-PU model can effectively correct the color of degraded underwater images, making them more suitable for observing underwater conditions.

Povzetek: Članek predlaga kaskadni kompozitni nevronski model z učečim usmerjanjem, ki združuje UWCNN, Deep Wave-Net in PUIE-Net za izboljšavo podvodnih slik.

1 Introduction

Clear and high-quality underwater images are critical for deep-sea topographic surveys and seabed resource investigations, and they have been widely applied in fields such as underwater target identification and detection [1]. However, due to the influence of the special underwater environment, underwater images are often subject to degradation, which severely impairs their imaging quality and recognition performance. Consequently, there is an urgent need for underwater image enhancement technologies to restore degraded underwater images and obtain clearer ones.

In recent years, numerous scholars have conducted research on underwater degraded image enhancement technologies and proposed various methods, among which deep learning methods are the most prevalent, as exemplified in [2-18].

Chen et al. [15] proposed an underwater image enhancement framework based on self-attention and contrastive learning (UIESC) to address the issues of low contrast, color distortion, and blurred details. Local features and global dependencies are constructed through spatial and channel dual attention, while crisscross attention is utilized to mitigate the high computational complexity of self-attention. Finally, smoothed histogram equalization is employed for further optimization to adapt to complex and variable underwater scenes. Zhou et al. [16] put forward an efficient and fully guided information flow network (UGIF-Net) for underwater image enhancement. This network accurately approximates color information by integrating features from two color spaces within a unified framework. Subsequently, a dense attention block (DAB) is adopted to guide the network in thoroughly extracting color information from both color spaces while adaptively perceiving critical color information. Galdran et al. [17] leveraged the characteristic that the brightness of the dark channel decays rapidly in underwater images, taking the brightness value of the brightest pixel in the dark channel as the background light. Based on this, they derived the expression of transmittance and then estimated the undegraded underwater images through inverse transmittance transformation. Nevertheless, dehazing methods are ineffective for underwater image restoration due to color attenuation and blue (green) color tones caused by the selective absorption of water and light scattering. Yang et al. [18] proposed a new underwater image restoration method based on the idea of first removing color distortion and then eliminating background scattering, aiming to overcome the shortcomings of such methods. According to the attenuation characteristics of light in water, the transmittance of each channel is corrected using the relationship between the scattering coefficient and wavelength.

Convolutional neural networks (CNNs) are among the most commonly used deep learning structures. CNN methods combined with physical models aim to obtain more accurate transmission maps through CNN networks, thereby generating better underwater images. Fu et al. [2] decomposed underwater image enhancement (UIE) into distribution estimation and consensus processes and proposed a novel probabilistic network (PUIE) that combines conditional variational autoencoders with adaptive instance normalization to construct enhanced distributions. The consensus process is then used to predict deterministic outcomes from a set of samples in the distribution. By learning the enhanced distribution, this method can, to a certain extent, cope with the bias introduced in the labeling of reference images. Li et al. [3] proposed an underwater image enhancement convolutional neural network (CNN) model based on underwater scene priors, named UWCNN. By combining an underwater imaging physical model with the optical properties of underwater scenes, they first synthesized underwater image degradation datasets covering a diverse range of water types and degradation levels. Then, a lightweight CNN model was designed for enhancing each type of underwater scene, which was trained using the corresponding training data. Finally, this UWCNN model was directly extended to underwater video enhancement. Sharma et al. [4] incorporated an attentive skip mechanism to adaptively refine the learned multi-contextual features. The proposed framework, called Deep Wave-Net, is optimized using traditional pixel-wise and feature-based cost functions.

1.1 Research gaps

In the field of underwater image enhancement, composite models systematically integrate the advantages of multiple individual models through innovative architectures that combine enhancement models in series. Such models have achieved significant breakthroughs in key technical areas including the restoration of underwater image clarity and the correction of color distortion thereby addressing existing limitations in these domains. Specifically, deep learning-based composite models can cascade Retinex theory-based image enhancement modules with Generative Adversarial Network (GAN) texture restoration modules, effectively mitigating the insufficient enhancement performance of traditional single models in complex aquatic environments and advancing the practical application of underwater visual processing technologies. Below is a detailed analysis to elaborate on our research motivation:

- Single neural networks such as UWCNN and Deep Wave-Net predominantly adopt fixed architectures for feature extraction, limiting their ability to simultaneously capture the global color distribution and local details of underwater images. For example, while UWCNN integrates a physical model, its shallow network structure may fail to extract deep semantic features, leading to incomplete color correction in complex scenes. Similarly, although Deep Wave-Net's skip connection mechanism enables multi-scale feature fusion, a single network lacks sufficient adaptability to the unique light attenuation patterns inherent in underwater environments.
- Existing single models primarily rely on synthetic datasets, which struggle to fully simulate the complex degradation processes present in real underwater environments. For instance, the dataset used to train UWCNN may not include images of extremely turbid waters, resulting in limited generalization capabilities for the model in real-world scenarios. Although the composite model proposed in this paper combines multiple single models, it does not address the domain shift problem between training data and real-world scenes.
- Underwater image enhancement requires simultaneous handling of multiple tasks, such as contrast enhancement, color correction, and noise reduction. Single models often utilize a single loss function for optimization, making it challenging to balance the objectives of each task. For example, optimizing solely for PSNR may lead to excessive image smoothing, causing the loss of texture details; conversely, focusing exclusively on color correction may neglect noise suppression. While the PU-DW model demonstrates superior performance in quantitative metrics, single models generally lack designs for multi-task collaborative optimization, thereby limiting improvements in comprehensive performance.

The specific research contributions are as follows:

- **A cascaded composite model was proposed**: To address the limitations of single neural networks in extracting features from underwater images, an innovative approach was developed to connect multiple individual models (including UWCNN, Deep Wave-Net, and PUIE-Net) into a composite neural network. By designing a cascaded architecture, each model is assigned specific tasks such as defogging, color correction, and contrast enhancement forming an orderly processing pipeline that enables multi-task collaborative optimization. This approach compensates for the inability of single models to simultaneously capture global color distribution and local details.
- **A learnable routing mechanism was implemented**: A learnable routing mechanism

was integrated into the composite model architecture. Through training, this mechanism can automatically analyze characteristic information of input underwater images, including degradation degree, water quality, and lighting conditions. Based on these analyses, it adaptively selects processing paths across different single models and dynamically adjusts parameter combinations for each model, thereby formulating optimal enhancement strategies for diverse underwater scenarios.

- **Efficient processing of degraded images was achieved**: By combining the cascaded composite model architecture with the learnable routing mechanism, the model's computational processes and parameter configurations were optimized. On one hand, the advantages of individual models are leveraged to enable parallel processing, reducing computational redundancy. On the other hand, the learnable routing mechanism intelligently allocates tasks to avoid unnecessary computations. This approach significantly improves the processing efficiency of degraded underwater images without compromising enhancement performance.

1.2 Objectives

This study proposes a series - connected composite neural network for underwater image enhancement, aiming to achieve superior image enhancement performance. Experimental results demonstrate that, in complex underwater image enhancement scenarios, the series - connected composite network for underwater image enhancement outperforms other mainstream underwater image enhancement models. Furthermore, the systematic structure of this network allows for better integration of other underwater image enhancement algorithms, enhancing its extensibility.

1.3 Contributions

Existing methods in the field all utilize deep learning techniques to learn the mapping relationship between low - quality input images and high - quality output images. Despite differences in their specific implementations and technical details, these methods share a core objective: to effectively improve the quality of underwater images through deep learning. Building on three neural network methods UWCNN, Deep Wave - Net, and PUIE Net this study explores the feasibility of a Cascade-Based composite neural network and proposes a composite model based on these three networks. Experiments validate that the proposed composite model achieves better performance than the original individual models.

The overall goal of this study is, based on the three existing underwater image enhancement neural network methods (UWCNN, Deep Wave - Net, and PUIE - Net), to explore the feasibility and effectiveness of composite neural networks in underwater image enhancement tasks. It intends to construct a model that can more efficiently address the problem of underwater image quality degradation, thereby improving the visual quality and application value of underwater images.

Specific contributions are as follows:

- A cascaded composite model is proposed. The underwater image enhancement models are concatenated and integrated to form a composite model. This architecture achieves the collaborative resolution of degradation problems through a staged processing mechanism.
- A learnable routing mechanism is proposed. To tackle the issues of feature conflicts and computational redundancy in traditional serial structures, a gated feature routing module is developed.
- Efficient processing of degraded images is realized. By integrating the composite model architecture with the learnable routing mechanism, the computing process and parameter configuration are optimized. This enables parallel processing to reduce redundancy, and tasks are intelligently allocated to avoid unnecessary computations, thus improving the efficiency of underwater image processing.

2 Related work

The PUIE-Net proposed in Reference [2] enhances the image's detail processing capability by optimizing edge detection and feature extraction through improved loss functions. The UWCNN introduced in Reference [3] adopts an architecture consisting of convolutional layers and pooling layers for underwater image processing. In Reference [4], the proposed Deep Wave-Net converts wavelength information into data features, aiming to perform processing from the perspective of fundamental intrinsic data characteristics. The UIESC presented in Reference [15] utilizes multi-scale convolution for feature extraction, enabling the acquisition of image information across different scales. The UGIF-Net proposed in Reference [16] generates and processes images based on the GAN framework, leveraging the adversarial mechanism to enhance image quality. Reference [19] introduces UColor, which employs deep learning algorithms to adjust RGB channels for color restoration. The UGAN proposed in Reference [20] generates enhanced images through a generator combined with prior knowledge, with the goal of optimizing images using such prior information. In Reference [21], SGUIE adopts a consistency loss and pseudo-label mechanism, aiming to effectively utilize unlabeled data. The CECF introduced in Reference [22] integrates a local contrast enhancement algorithm and a feature pyramid to fuse features of different scales, with the objective of improving the comprehensive processing effect of image features. Reference [23] presents Semi-UIR, which uses labeled data for training and combines pseudo-labels to jointly optimize the restoration of damaged images. The UIE-DM proposed in Reference [24] dynamically adjusts the network structure and parameters via an environmental perception module to adapt to different environments. In Reference [25], HCLR-Net processes images by combining high-contrast learning and low-rank representation. Finally, the UIE-WD introduced in Reference [26] integrates image features based on weighted decision-making, achieving advantage integration through weighting. Table 1 Analysis of related work.

3 Proposed work

3.1 System method

In our system, underwater degraded images are evaluated by the assessment module, which is designed to determine whether the input image is degraded. If classified as a normal image, it is directly output. Conversely, if the image is identified as degraded by the assessment module, it is forwarded to the routing module. Through the decision sub-module within the routing module, a sequential composite enhancement model from the method library is selected for the enhancement process. After enhancement, a processed image is generated, which then undergoes re-assessment by the image assessment module. If this image is judged to be normal, it is output; otherwise, it is redirected to the routing module for another round of image enhancement operations. The system flowchart is shown in Figure 1.

3.2 Component description

To achieve our cascaded composite neural network, we use multiple underwater image enhancement models as components. By connecting these components in series, we can form a composite neural network. Therefore, we need to analyze the relevant characteristics and functions of each component.

Table 1: Analysis of related work

Reference number	Method	Proposed method	Limitation
[2]	PUIE-Net	Improve the loss function to optimize edge detection and feature extraction	It has high requirements for the dataset, and small samples are prone to overfitting
[3]	UWCNN	Adopt a convolutional layer and pooling layer architecture	It has poor adaptability to complex underwater environments, low algorithm efficiency and insufficient real-time performance
[4]	DeepWave-Net	Convert wavelength information into data features	The ability to handle unstructured data is limited and its universality is poor
[15]	UIESC	Features are extracted using multi-scale convolution	The semantic segmentation boundaries are ambiguous, and the recognition rate of small targets is low
[16]	UGIF-Net	Dense attention block (DAB)	The training is unstable, prone to mode collapse, and it is difficult to balance the realistic and diverse images
[19]	UColor	Adjust the RGB channels using deep learning a lgorithms	The color recovery varies greatly with different water qualities and lacks an adaptive mechanism
[20]	UGAN	The generator combines prior knowledge to generate enhanced images	The training is unstable, prone to mode collapse, and it is difficult to balance the realistic and diverse images
[21]	SGUIE	Consistency loss and pseudo-labeling mechanism	Unlabeled data has low utilization and the model converges slowly
[22]	CECF	The local contrast enhancement algorithm fuses features of different scales with the feature pyramid	Fusion is prone to losing details, which affects image quality
[23]	Semi-UIR	Repair damaged images by training with labeled data and jointly optimizing with pseudo-labels	The repaired area does not match the surrounding area well and the effect is unnatural
[24]	UIE-DM	The environmental perception module dynamically adjusts the network structure and parameters	The calculation is complex, the hardware consumption is high and the processing is slow
[25]	HCLR-Net	Combine high-contrast learning with low-rank representation	The feature extraction from complex backgrounds is poor, and the low-rank assumption is difficult to satisfy
[26]	UIE-WD	Make weighted decisions based on image features to determine the comprehensive advantages	Weights rely on experience and lack adaptability, making it difficult to adapt to diverse scenarios

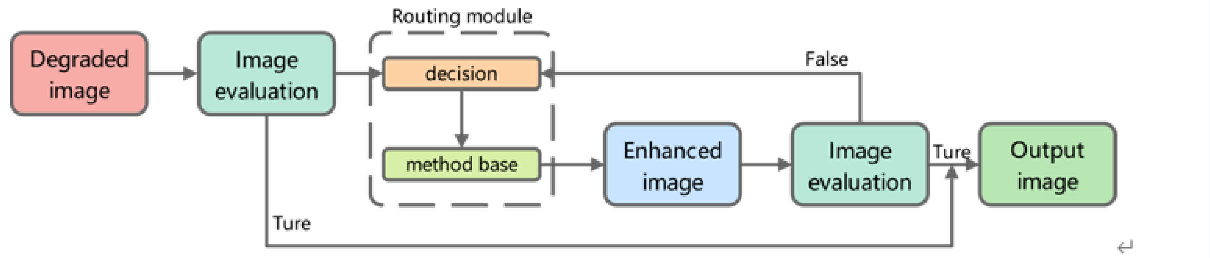


Figure 1: System flowchart

3.2.1 Component description

Therefore, we conducted numerical comparisons, calculating the Peak Signal-to-Noise Ratio (PSNR), Structure Similarity Index Measure (SSIM), Underwater Colour Image Quality Evaluation (UCIQE), and Underwater Image Quality Measure (UIQM) of the degraded images after applying different neural network enhancement techniques. These metrics were used to assess the effectiveness and quality of the image enhancement. Thus, we were able to evaluate the performance of the image enhancement model.

SSIM: This is an index used to evaluate the similarity between two images, often employed to measure the similarity between an image before and after distortion. The calculation of SSIM is based on the sliding window method, that is, each calculation takes a window of size N×N from the image, calculates the SSIM index based on the window, traverses the entire image, and then takes the average of all window values as the SSIM index of the entire image. Let x represent the data in the window of the first image, and y represent the data in the window of the second image. The similarity of the images consists of three parts: $l(x,y)$ for brightness, $c(x,y)$ for contrast, and $s(x,y)$ for structure.

$$l(x,y)=\frac{2\mu_x\mu_y+c_1}{\mu_x^2+\mu_y^2+c_1} \quad (1)$$

$$c(x,y)=\frac{2\sigma_x\sigma_y+c_2}{\sigma_x^2+\sigma_y^2+c_2} \quad (2)$$

$$s(x,y)=\frac{\sigma_{xy}+c_3}{\sigma_x\sigma_y+c_3} \quad (3)$$

$$SSIM(x,y)=\frac{(2\mu_x\mu_y+c_1)(2\sigma_x\sigma_y+c_2)}{(\mu_x^2+\mu_y^2+c_1)(\sigma_x^2+\sigma_y^2+c_2)} \quad (4)$$

Here, μ_x and μ_y respectively represent the mean values of *x* and *y*, σ_x and σ_y respectively represent the variances of *x* and *y*, σ_{xy} represents the covariance between *x* and *y*, $c_1=(k_1L)^2$, $c_2=(k_2L)^2$ and $c_3=c_2/2$ represent three constants, avoiding division by zero, k_1 and k_2 respectively default to 0.01 and 0.03, L represents the range of image pixel values, and $L=2^B-1$, B represent the number of pixel bits.

PSNR: It is commonly used to evaluate the degree of distortion of compressed images or videos compared to the original images or videos. The higher the PSNR, the higher the similarity between the compressed image and video and the original image and video, and the better the quality. MAX represents the maximum value of the pixel values after 8-bit image normalization, which is 255. MSE represents the mean square error. $I(i,j)$ represents the value of the image at pixel position (i,j), and $R(i,j)$ represents the value of the reference image at pixel position (i,j).

$$PSNR=10\cdot\log_{10}\left(\frac{MAX^2}{MSE}\right) \quad (5)$$

$$MES=\frac{1}{M\cdot N}\sum_{i=1}^{M}\sum_{j=1}^{N}\left(I(i,j)-R(i,j)\right)^2 \quad (6)$$

UCIQE: UCIQE refers to the evaluation of the color quality of underwater images to determine the visual performance and effectiveness of the images. The closer the value is to 1, the richer the image's color. σ_c represents the standard deviation of chromaticity, con_l represents the brightness contrast, μ_s represents the color saturation, c_1,c_2,c_3 represent weighting coefficients, and they are usually set as $c_1=0.4680$, $c_2=0.2745$, $c_3=0.2576$.

UIQM is an indicator used to evaluate the quality of underwater images. It combines the color measurement indicator UICM, the clarity measurement indicator UISM, and the contrast measurement indicator UIConM, reflecting the color, clarity and contrast of underwater images. Here, w_1,w_2,w_3 represents the weighting coefficient, and it is generally set to $w_1=0.5$, $w_2=0.3$, $w_3=0.2$.

Evaluation of input image: The input image $I_{in}\in\mathbb{R}^{H\times W\times 3}$, where *H* and *W* represent the height and width of the image respectively, and 3 represents the number of RGB channels. By calculating the unified evaluation index of image $q_{quality}=UIQM(I_{in})$ quality, $q_{color}=UCIQE(I_{in})$ color quality index, and $s=[H,W]$ pixel size of the image. When UIQM < 3, it indicates that the image has color distortion. When UCIQE < 0.2, it indicates that the image has insufficient color saturation. When $s=[H,W]>400\times 400$, it indicates that the image is too large, and it will consume a certain amount of time during image enhancement.

3.2.2 Module design

Routing Module: The routing module is a module composed of a decision-making module and a method library module. When dealing with the selection of multi-model cascading schemes, we innovatively proposed a learnable routing mechanism. The core of this mechanism lies in its ability to make intelligent routing optimization decisions between subnets. Specifically, this mechanism selects the combination model processing scheme by training a dedicated routing network, based on the underwater image features input and the image quality evaluation indicators processed by each sub-model.

Decision Module: The training process of the decision module. By designing a loss function L that minimizes, we can learn the parameters W_1,b_1,W_2,b_2 of the decision module. The training process can be expressed as *D* representing the training dataset, which includes the degraded images I_{in} and the corresponding enhanced images I_{target} along with their related parameters.

The decision-making process of the decision module. By obtaining the feature quantities $x=[f_{img}, q_{color}, q_{quality}, s]$ of the input image and through the decision function R(x), the probability distribution $p \in \mathbb{R}^{M}$ is output. Here, p_i represents the probability of the selection method M_i. This achieves the selection of an appropriate composite model method for the image to be reinforced by calculating the distribution probability $p=\mathrm{soft}\max(W_1 x+b_1)$ of method M_i in the method library, where $W_1 \in \mathbb{R}^{M\times d}$ and $b_1 \in \mathbb{R}^{M}$ are learnable parameters, and d is the input feature x. The decision-making mechanism of the routing module is shown in Figure 2.

The underwater image enhancement methods in the method library:

Deep Wave-Net (DW): This model is distinguished by its incorporation of an attentive skip mechanism and wavelength-aware feature transformation, converting wavelength information into discriminative data features [4]. This unique characteristic allows it to capture multi-scale contextual details, especially in scenarios with non-uniform illumination and blurred textures. Its primary function is to refine local details and strengthen edge information, complementing the global feature processing of preceding components in the cascade.

UWCNN (UW): This model is built on a convolutional neural network architecture that integrates underwater scene priors and physical imaging models [3]. Its key characteristic is its lightweight structure, enabling efficient extraction of low-to-medium level features from underwater images. Functionally, it excels in preliminary processing tasks such as mitigating mild color distortion and enhancing basic contrast, making it suitable as an initial component in the cascaded framework to lay a foundation for subsequent enhancement steps.

PUIE-Net (PU): Leveraging a probabilistic network design that combines conditional variational autoencoders with adaptive instance normalization, this model focuses on learning enhanced feature distributions

$$\min_{W_1,b_1,W_2,b_2} \mathrm{E}_{(I_{in},I_{target})\sim \mathrm{D}}\left[\mathrm{L}(I_{in},I_{target};W_1,b_1,W_2,b_2)\right] \quad (7)$$

[2]. A notable characteristic is its robustness in handling complex color degradation, which is attributed to its optimized loss function that prioritizes edge preservation and fine-grained feature extraction. In the cascaded structure, its core function is to perform advanced color correction and suppress residual noise, thereby enhancing the overall visual quality of images processed by upstream components.

3.3 Model building

To optimize the selection of multi-model cascading schemes, a learnable routing mechanism is employed to refine model selection decisions. This mechanism involves training a routing network that selects enhancement schemes from the method library based on the features of input underwater images and the image quality evaluation metrics derived from processing by each sub-model. The process is as follows: first, feature extraction is performed on the input underwater images. Subsequently, metrics such as PSNR, SSIM, UIQM, and UCIQE are computed for both the extracted features and the images processed by each sub-model. These features and metrics are then fed into the routing network, enabling it to determine the optimal combination of model processing schemes.

Individual enhancement models from the method library serve as middleware for the cascaded composite model. During the data processing phase, after each neural network model enhances the image data, its output is processed and used as the input for the next cascaded neural network component. By exploring various combination strategies of individual neural network components, a cascaded neural network composite model is constructed. The specific workflow includes dataset acquisition, image normalization, sub-model enhancement, another round of image normalization and sub-model enhancement, and final multi-index evaluation of the generated images. Taking the PUIE-Net-Deep Wave-Net composite model as an example, the relevant steps are illustrated in Figure 3.

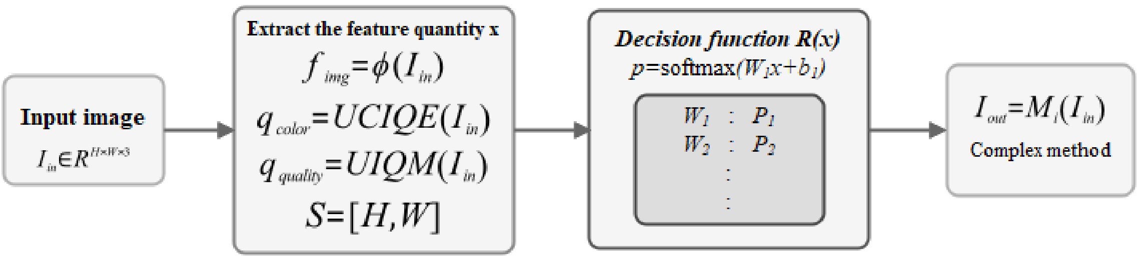


Figure 2: Routing module decision-making mechanism

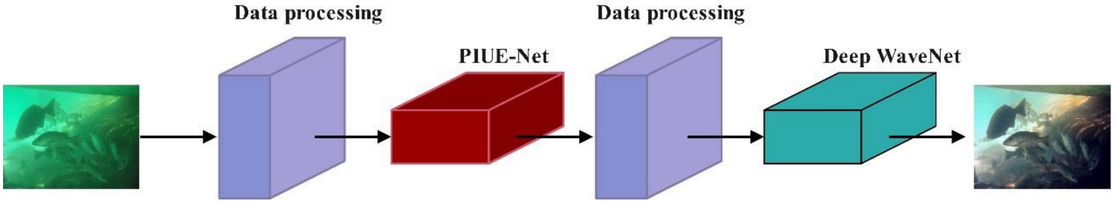


Figure 3: Processing flow of composite model

To systematically investigate the performance improvements of composite models relative to single enhancement models, this paper constructs various composite model architectures and acquires relevant data through comparative experiments. The specific results are presented in Table 2.

Table 2: Method library composite enhancement model table

	UWCNN (UW)	PUIE-Net (PU)	Deep Wave-Net (DW)
UWCNN (UW)	—	UW-PU	UW-DW
PUIE-Net (PU)	PU-UW	—	PU-DW
Deep Wave-Net (DW)	DW-UW	DW-PU	—

4 Results and discussion

4.1 Experimental environment

The experiment was conducted on a system equipped with an Intel Core i7 processor, 16GB of RAM, and an NVIDIA RTX 4060 graphics card, which provided robust computational support for model training and testing. The experimental environment was developed based on Python 3.10. Specifically, the PyTorch deep learning framework was adopted for model construction, while the OpenCV library was utilized for image preprocessing and postprocessing operations. Additionally, the NumPy and Pandas libraries were employed for data processing and analysis, and the Matplotlib library was used for data visualization. The experimental environment is shown in Table 3.

Table 3: Experimental Environment List

Configuration	Experimental environment
Intel Core i7	CPU
16GB	Memory size
NVIDIARTX060	GPU
Python 3.10	Programing language
OpenCV	Image processing and analysis

4.2 Data preparation

The datasets utilized in this study are summarized in Table 4. The U45 dataset [27] focuses on underwater multi-task research and includes image samples exhibiting underwater distortion characteristics. The EUVP dataset [28] is specifically designed for underwater image enhancement tasks, with its data structure comprising paired and unpaired data: the former consists of three subsets (Underwater Dark, Image Net, and Scenes) totaling 24,840 images, which are suitable for supervised learning scenarios; the latter contains 6,665 low/high-quality image pairs, supporting unsupervised or semi-supervised training. As the first benchmark for real underwater scene enhancement, the UIEB dataset [29] is divided into a supervised training subset (890 pairs of original images and artificially enhanced reference images) and a challenge test subset (60 reference-free images for evaluating algorithm robustness). The Underwater_ImageNet (UWIN) dataset [20] is an open-source underwater vision dataset extended from the traditional ImageNet, integrating the degradation characteristics of the underwater environment with the semantic diversity of natural images.

To construct a dataset for training the routing module and decision-making module, 500 images were randomly selected from each of the EUVP, UIEB, and Underwater_ImageNet datasets. Corresponding enhanced images were generated using the method library, and the degraded and enhanced images were paired. This dataset was then split into training, test, and validation sets at a ratio of 8:2:2. Additionally, the UIEB dataset was used as a benchmark to compare performance with the recently proposed outstanding underwater image enhancement algorithm UGIF-Net.

Table 4: The dataset used in the experiment

Dataset name	Reference materials
U45-[28]	https://github.com/IPNUISTlegal/underwater-test-dataset-U45-
EUVP[29]	https://irvlab.cs.umn.edu/resources/euvp-dataset
UIEB[22]	https://li-chongyi.github.io/proj_benchmark.html
UWIN[20]	https://github.com/xinzhichao/underwater_datasets

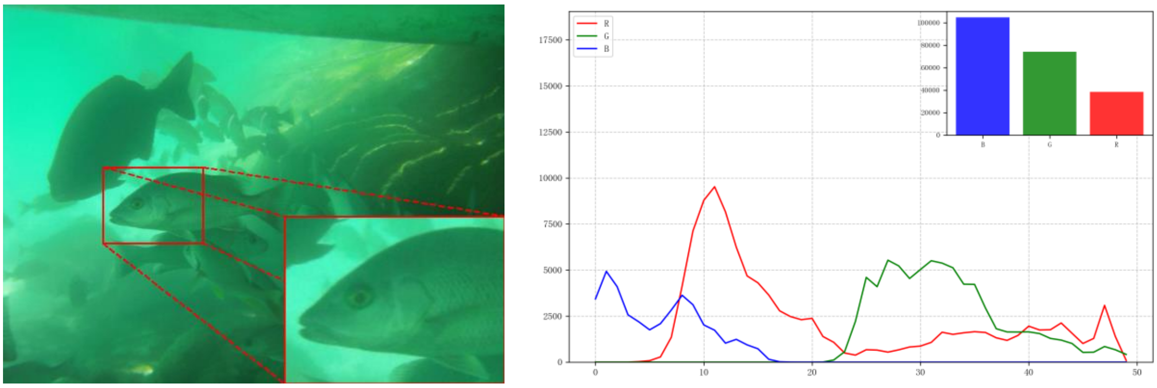
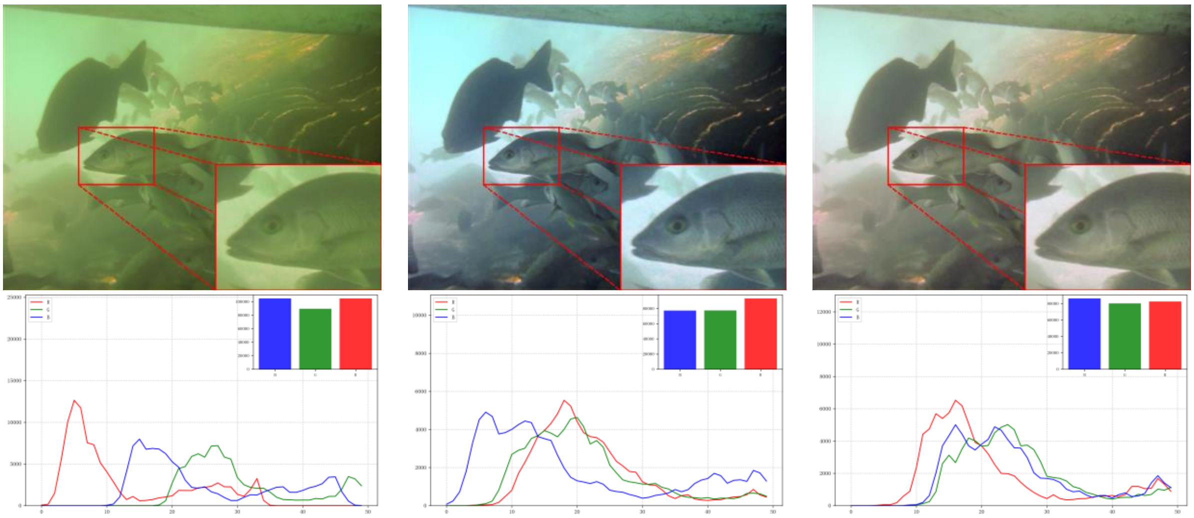
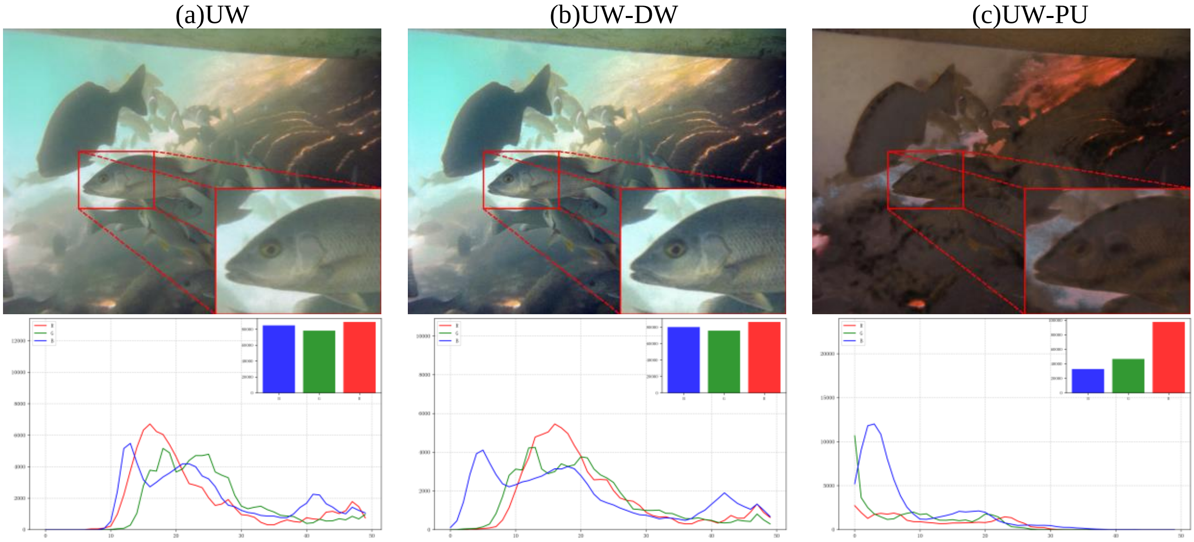


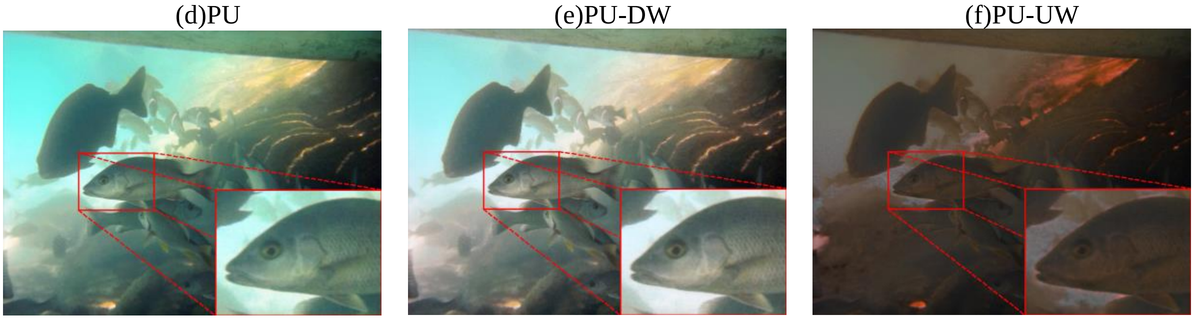
Figure 4: Original image and RGB channel color histogram

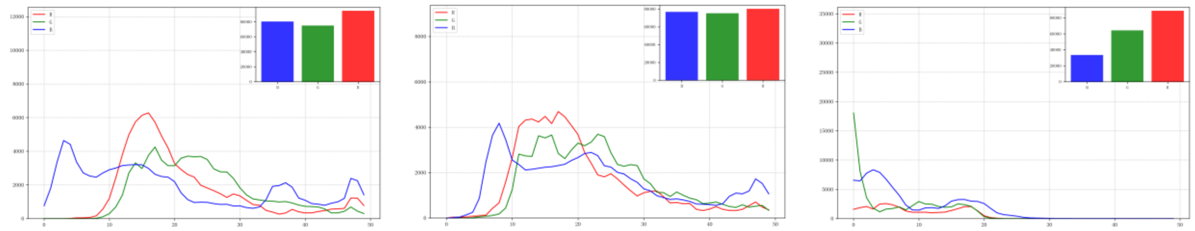


(a)UW (b)UW-DW (c)UW-PU



(d)PU (e)PU-DW (f)PU-UW





(g)DW (h)DW-PU (i)DW-UW

Figure 5: Enhanced image and RGB channel color histogram

4.3 Numerical experiment

On the UIEB dataset, numerical experiments were conducted using the cascaded underwater image enhancement composite neural network model, with an in-depth analysis performed on the color histograms of the RGB channels for the output images of each enhancement model. The specific experimental results are presented in Figure 4 and Figure 5.

A comparative analysis of the color histogram of the enhanced image in Figure 5, the original image in Figure 4, and their respective RGB channel color histograms reveals that the RGB channel values in the original image's color histogram exhibit significant fluctuations, with the red channel showing the most pronounced amplitude. In contrast, for the image enhanced by the proposed model, the distribution curves of the RGB channels in the color histogram are more balanced, and the differences in values between channels are significantly reduced. This result intuitively validates the effectiveness of the image enhancement process. From a visual perception perspective, the composite neural network model demonstrates better visual adaptability in enhancement effects compared to the single neural network model, aligning more closely with human visual preferences.

However, not all models achieve ideal enhancement effects. Taking the DW-UW and PU-UW models as examples, the processed images exhibit obvious color discrepancy issues, accompanied by the persistence of low-light phenomena. Additionally, the color histograms of the enhanced images still show significant fluctuations, reflecting the limitations of these two model types in image enhancement. To achieve an objective and precise evaluation of model performance, this study introduces image quality assessment metrics such as PSNR, SSIM, UCIQE, and UIQM for quantitative analysis of the test images. The specific evaluation results are presented in Table 5 below.

This study systematically compares the objective performance of nine underwater image enhancement methods across three standard datasets: EUVP, LSUI, and UWIN. The experimental results indicate that the DW-PU method exhibits significant advantages: on the LSUI dataset, it ranks first with a PSNR of 22.232±0.321 and an SSIM of 0.870±0.007; on the EUVP dataset, its PSNR and SSIM metrics reach 21.247±0.293 and 0.818±0.009, respectively; on the UWIN dataset, it also maintains a leading position, achieving excellent performance with a PSNR of 20.619±0.343 and SSIM values of 0.818±0.009 and 0.803 ±0.013.

Table 5: Quantitative results of supervised training on the PSNR and SSIM metrics, using the full-reference benchmark. The best results are indicated in red, and the second-best results are indicated in blue

Method	EUVP		LSUI		UWIN	
	PSNR↑	SSIM↑	PSNR↑	SSIM↑	PSNR↑	SSIM↑
DW	21.208±0.217	0.884±0.004	15.515±0.227	0.759±0.006	16.986±0.312	0.798±0.014
DW-PU	21.247±0.293	0.818±0.009	22.232±0.321	0.870±0.007	20.619±0.343	0.803±0.013
DW-UW	18.275±0.213	0.762±0.008	14.766±0.210	0.696±0.006	15.948±0.250	0.723±0.012
PU	13.721±0.252	0.764±0.008	14.031±0.233	0.799±0.007	13.501±0.297	0.733±0.015
PU-DW	12.389±0.233	0.729±0.008	12.660±0.238	0.778±0.007	11.699±0.260	0.695±0.014
PU-UW	13.069±0.233	0.693±0.008	13.583±0.217	0.740±0.007	13.087±0.273	0.686±0.012
UW	13.978±0.255	0.745±0.008	12.956±0.206	0.709±0.007	13.927±0.290	0.719±0.014
UW-DW	18.371±0.252	0.802±0.008	17.190±0.270	0.840±0.007	16.156±0.269	0.768±0.012
UW-PU	13.688±0.257	0.730±0.009	14.158±0.236	0.775±0.007	13.529±0.273	0.717±0.012

Table 6: Quantitative results based on no-reference benchmarks with UIQM and UCIQE as indicators. The best results are shown in red and the second-best results in blue

Method	EUVP		LSUI		UWIN	
	UIQM↑	UCIQE↑	UIQM↑	UCIQE↑	UIQM↑	UCIQE
DW	2.904±0.037	0.805±0.009	2.947±0.026	1.359±0.060	2.826±0.040	1.265±0.159
DW-PU	3.086±0.027	1.246±0.142	3.127±0.020	1.016±0.019	2.982±0.031	0.810±0.008
DW-UW	2.473±0.052	1.192±0.102	2.760±0.033	1.015±0.061	2.511±0.054	1.125±0.134
PU	3.044±0.028	0.802±0.010	3.092±0.030	0.753±0.014	3.012±0.031	0.772±0.010
PU-DW	2.977±0.031	0.901±0.012	3.034±0.030	0.817±0.013	2.909±0.030	0.886±0.013
PU-UW	2.739±0.034	0.882±0.055	2.932±0.029	0.795±0.037	2.765±0.035	0.875±0.065
UW	2.327±0.050	0.983±0.086	2.429±0.055	0.645±0.041	2.320±0.051	0.921±0.146

UW-DW	2.757±0.044	1.153±0.101	2.844±0.042	0.768±0.057	2.727±0.047	1.236±0.204
UW-PU	3.021±0.029	0.762±0.013	3.110±0.029	0.707±0.013	3.014±0.032	0.729±0.011

Table 7: Comparison Experimental Results under the UIEB Dataset. The best results are shown in red and the second-best results in blue

Method	PSNR	SSIM	UIQM↑	UCIQE↑
DW	20.091	0.866	2.888	0.866
DW-PU	20.495	0.874	3.270	0.897
DW-UW	24.156	0.746	3.223	0.783
PU	13.875	0.745	2.790	0.760
PU-DW	12.835	0.735	3.143	0.830
PU-UW	13.298	0.689	3.101	0.891
UW	14.206	0.797	2.721	0.765
UW-DW	18.767	0.827	2.808	0.868
UW-PU	14.073	0.727	3.265	0.703
UGIF-Net	24.466	0.915	3.129	0.622

Table 6 provides a detailed comparison of the performance of cascaded underwater image enhancement methods in terms of no-reference evaluation metrics across three major datasets: EUVP, LSUI, and UWIN. The experimental results reveal that the DW-PU model combination exhibits a significant advantage in the UIQM metric. It achieves the optimal results of 3.086±0.027 and 3.127±0.020 on the EUVP and LSUI datasets, respectively, and ranks first across all three datasets.

Table 7 conducts a comprehensive performance evaluation of the recent underwater image enhancement method UGIF-Net on the UIEB dataset. The experimental results indicate that the tandem composite model lags slightly behind UGIF-Net in terms of PSNR and SSIM metrics. The tandem composite model represented by DW-PU achieves a UIQM value of 3.270 and a UCIQE value of 0.897, while its PSNR value of 20.495 is slightly lower than that of the mainstream method UIESC (24.466). Notably, the tandem composite model demonstrates unique advantages in color restoration and overall image quality improvement, confirming its effectiveness in underwater image enhancement tasks.

To compare the performance differences between single and composite neural network models, three single models (UWCNN, PUIE-Net, and Deep Wave-Net) and six composite models (UW-PU, UW-DW, PU-UW, PU-DW, DW-UW, and DW-PU) were selected. Experiments were conducted using the U45 dataset [27], and the specific comparison results are presented in Figure 6.

(a) (b) (c) (d) (e) (f) (g) (h) (i) (j)

(a)Original (b)UWCNN (c)Deep Wave-Net (d)PUIE-Net (e)UW-DW (f)UW-PU (g)DW-UW (h)DW-PU (i)PU-DW (j)PU-UW

Figure 6: Comparison of enhancement effects of multiple composite models

After enhancement processing, the test images exhibit significant changes: compared with the original images, the enhanced images in column (b) still suffer from color distortion; the original image in column (g) was captured under low-light conditions, but its details become clearly distinguishable after enhancement; the original image in column (j) had blurring issues, which are effectively alleviated. As intuitively observed from the sample images, the cascaded enhancement model demonstrates certain advantages.

5 Conclusion

This study focuses on analyzing types of image degradation and evaluating metrics for enhanced images based on a cascaded underwater image enhancement composite neural network model. A cascaded underwater image enhancement model is constructed, which first establishes a routing mechanism to enable model allocation for serial enhancement schemes, followed by an analysis of the advantages of three underwater image enhancement models: UWCNN, Deep Wave-Net, and PUIE-Net. To explore the enhancement performance of composite models, the method library incorporates various serial enhancement model allocation methods, including DW-PU, DW-UW, PU-DW, PU-UW, UW-DW, and UW-PU. Using the UIEB dataset, experiments evaluate images processed by 3 single models, 6 composite models, and the underwater image enhancement algorithm UGIF-Net. Parametric evaluation metrics (PSNR and SSIM) and non-parametric evaluation metrics (UCIQE and UIQM) are calculated, with comparative analysis conducted against the mainstream method UGIF-Net. The results indicate that the cascaded underwater image enhancement composite neural network, exemplified by the DW-PU composite model (PSNR 20.495, SSIM 0.874, UIQM 3.270, UCIQE 0.897), exhibits certain advantages in most scenarios where visual quality is a key concern.

Funding

This work was supported by the Research Basic Ability Improvement Project of Middle and Young Teachers in Colleges and Universities of Guangxi (No. 2024KY1800), and the National College Students' Innovation and Entrepreneurship Training Program (No. 202413639007).

References

[1] Guo J, Li C, Guo C, et al. Research progress of underwater image enhancement and restoration methods. Journal of Image and Graphics, 22(3): 273-287, 2017. https://doi.org/10.11834/jig.20170301

[2] Fu Z, Wang W, Huang Y, et al. Uncertainty inspired underwater image enhancement. arXiv e-prints, 2022. https://doi.org/10.48550/arXiv.2207.09689.

[3] Li C, Anwar S, Porikli F. Underwater scene prior inspired deep underwater image and video enhancement. Pattern Recognition, 98: 107038, 2020. https://doi.org/10.1016/j.patcog.2019.107038

[4] Sharma P, Bisht I, Sur A. Wavelength-based Attributed Deep Neural Network for Underwater Image Restoration. ACM Transactions on Multimedia Computing Communications and Applications, 2023. https://doi.org/10.1145/3511021

[5] Dong She. Retinex-based Visual Image Enhancement Algorithm for Coal Mine Exploration Robots. Informatica, vol. 48, no. 11, pp. 133-146, 2024. https://doi.org/10.31449/inf.v48i11.6003

[6] Pratima Sarkar, Sourav De, Sandeep Gurung. U-YOLOv3: A Model Focused on Underwater Object Detection. Informatica, vol. 49, no. 6, pp. 87-102, 2025. https://doi.org/10.31449/inf.v49i6.6642

[7] Wala'a Nsaif Jasim, Zainab Najem Nemer, Esra'a Jasem Harfash. Implementation of Multiple CNN Architectures to Classify the Sea Coral Images. Informatica, vol. 47, no. 1, pp. 43-50, 2023. https://doi.org/10.31449/inf.v47i1.4429

[8] Deng Z, Zhu L, Hu X, et al. Deep Multi-Model Fusion for Single-Image Dehazing. IEEE. https://doi.org/10.1109/ICCV.2019.00254

[9] Cong X, Zhao Y, Gui J, et al. A Comprehensive Survey on Underwater Image Enhancement Based on Deep Learning. 2024. https://doi.org/10.48550/arXiv.2405.19684

[10] Liu L, Wiberg A O J, Myslivets E, et al. Comparison of One- and Three-Mode Phase-Sensitive Wavelength Multicasting. Journal of Lightwave Technology, 34(10): 2491-2499, 2016. https://doi.org/10.1109/JLT.2016.2529680

[11] Ailong Tang, Ling Wei, Zhiping Ni, Qiuyong Huang. Multi-Modal Modified U-Net for Text-Image Restoration: A Diffusion-Based Multimodal Information Fusion Approach. Informatica, vol. 49, no. 2, pp. 319-332, 2024. https://doi.org/10.31449/inf.v49i2.8245

[12] Lei Y, Yu J, Dong Y, et al. UIE-UnFold: Deep Unfolding Network with Color Priors and Vision Transformer for Underwater Image Enhancement. 2024. https://doi.org/10.1109/DSAA61799.2024.10722842

[13] Wang J, Yu L, Tian S, et al. AMFNet: An attention-guided generative adversarial network for multi-model image fusion. Biomedical signal processing and control, 2022. https://doi.org/10.1016/j.bspc.2022.103990

[14] Lei C, Zhang H, Wang Z, Miao Q. Multi-Model Fusion Demand Forecasting Framework Based on Attention Mechanism. Processes 12: 2612, 2024. https://doi.org/10.3390/pr12112612

[15] Chen R, Cai Z, Yuan J. UIESC: An underwater image enhancement framework via self-attention and contrastive learning. IEEE Transactions on Industrial Informatics, 19(12): 11701-11711, 2023. https://doi.org/10.1109/TII.2023.3249794

[16] Zhou J, Li B, Zhang D, et al. UGIF-Net: An Efficient Fully Guided Information Flow Network for Underwater Image Enhancement. IEEE Transactions on Geoscience and Remoter Sensing, 61: 1-17, 2023. https://doi.org/10.1109/TGRS.2023.3293912.
[17] Galdran A, Pardo D, Picón A, et al. Automatic red-channel underwater image restoration. Journal of Visual Communication and Image Representation, 2015, 26: 132-145.
[18] Yang Aiping, Zheng Jia, Wang Jian, et al. Underwater image restoration based on color cast removal and dark channel prior. Journal of Electronics and Information Technology, 2015, 37(11): 2541-2547.
[19] Li C, Anwar S, Hou J, et al. Underwater Image Enhancement via Medium Transmission-Guided Multi-Color Space Embedding. IEEE Transactions on Image Processing, PP(99): 1-1, 2021. https://doi.org/10.1109/TIP.2021.3076367
[20] Fabbri C, Islam M J, Sattar J. Enhancing underwater imagery using generative adversarial networks. 2018 IEEE international conference on robotics and automation (ICRA). IEEE, 2018: 7159-7165. https://doi.org/10.1109/ICRA.2018.8460552
[21] Qi Q, Li K, Zheng H, et al. SGUIE-Net: Semantic Attention Guided Underwater Image Enhancement with Multi-Scale Perception. arXiv e-prints, 31: 6816-6830, 2022. https://doi.org/10.48550/arXiv.2201.02832.
[22] Cong X, Gui J, Hou J. Underwater organism color fine-tuning via decomposition and guidance. Proceedings of the AAAI conference on artificial intelligence, 38(2): 1389-1398, 2024. https://doi.org/10.1609/aaai.v38i2.2790
[23] Huang S, Wang K, Liu H, et al. Contrastive semi-supervised learning for underwater image restoration via reliable bank. Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2023: 18145-18155. https://doi.org/10.48550/arXiv.2303.09101
[24] Tang Y, Kawasaki H, Iwaguchi T. Underwater image enhancement by transformer-based diffusion model with non-uniform sampling for skip strategy. Proceedings of the 31st ACM international conference on multimedia. 2023: 5419-5427. https://arxiv.org/abs/2309.03445
[25] Zhou J, Sun J, Li C, et al. HCLR-Net: hybrid contrastive learning regularization with locally randomized perturbation for underwater image enhancement. International Journal of Computer Vision, 132(10): 4132-4156, 2024. https://doi.org/10.1007/s11263-024-01987-y
[26] Ma Z, Oh C. A wavelet-based dual-stream network for underwater image enhancement. ICASSP 2022-2022 IEEE international conference on acoustics, speech and signal processing (ICASSP). IEEE, 769-2773, 2022. https://doi.org/10.48550/arXiv.2202.08758.
[27] Ancuti C, Ancuti C O, Haber T, et al. Enhancing underwater images by fusion. 2011. https://doi.org/10.1145/2037715.2037753.
[28] Islam J, Xia Y, Sattar J. Fast Underwater Image Enhancement for Improved Visual Perception. International Conference on Robotics and Automation. IEEE, 2020. https://doi.org/10.1109/LRA.2020.2974710.
[29] Li C, Guo C, Ren W, et al. An Underwater Image Enhancement Benchmark Dataset and Beyond. IEEE, 2020. https://doi.org/10.1109/TIP.2019.295524