SCIPEDIA

# MTNet: Multi-Task Underwater Image Enhancement Method Based on Retinex

Yang Song, Xing Deng*, Haijian Shao and Fei Wang

School of Computer, Jiangsu University of Science and Technology, Zhenjiang, 212003, China


**INFORMATION**

**Keywords:**
Image enhancement
Retinex
convolution
discrete wavelet transform (DWT)
deep curve

**DOI:** 10.23967/j.rimni.2025.10.60509

Revista Internacional de Métodos Numéricos para Cálculo y Diseño en Ingeniería

# MTNet: Multi-Task Underwater Image Enhancement Method Based on Retinex

**Yang Song, Xing Deng*, Haijian Shao and Fei Wang**

School of Computer, Jiangsu University of Science and Technology, Zhenjiang, 212003, China

**ABSTRACT**

Underwater images play a critical role in underwater exploration and related tasks. However, due to light attenuation and other underwater factors, underwater images often suffer from color distortion and low contrast, which to some extent limit the efficiency and safety of underwater exploration. To meticulously address these issues and enhance the accuracy and reliability of underwater exploration, this paper proposes a multi-task underwater image enhancement method based on Retinex theory. This method divides the underwater image enhancement task into several sub-tasks, including image decomposition, color correction, detail reconstruction, and illumination adjustment. Specialized sub-networks—DecomNet, DecolorNet, and DelightNet—are designed to specifically address these problems, thereby alleviating color distortion, enhancing image details, and improving contrast. Experiments conducted on several publicly underwater image datasets indicate that the quality of underwater images is significantly improved after enhancement with the proposed method, compared to other representative underwater image processing techniques. For example, on the real-world dataset Underwater Image Enhancement Benchmark, the MSE, Structural Similarity Index Measure, and Peak signal-to-noise ratio scores achieved were 453.480, 0.901, and 25.145, respectively. This study holds significant implications for underwater exploration, with potential applications in the fields of marine research and underwater archaeology.

OPEN ACCESS

**Received:** 03/11/2024

**Accepted:** 27/12/2024

**Published:** 20/04/2025

**DOI**
10.23967/j.rimni.2025.10.60509

**Keywords:**
Image enhancement
Retinex
convolution
discrete wavelet transform (DWT)
deep curve

## 1 Introduction

Underwater images serve as a pivotal conduit for the acquisition of subaquatic information, with its quality being paramount to the success of underwater exploration and related operational tasks [1]. However, the complex of the underwater environment poses significant challenges to the collection of high-quality underwater images. For example, the selective attenuation of light by water often results in images captured by imaging devices exhibiting a blue or green hue [2]. Furthermore, the presence of suspended particulate matter in water induces forward and backward scattering of

*Correspondence: Xing Deng (xdeng@just.edu.cn). This is an article distributed under the terms of the Creative Commons BY-NC-SA license

light before it reaches the imaging equipment, leading to a reduction in image contrast and increased blurriness [3]. These phenomena severely impede the efficiency and accuracy of advanced visual processing tasks such as image segmentation, object classification, and detection. Consequently, the effective preprocessing of acquired underwater images to enhance their quality has become an urgent requirement in the research of underwater visual systems.

In the early stages of research, image enhancement relied on the direct manipulation of image pixels [4–6] or the utilization of environmental prior knowledge to achieve image restoration based on imaging models [7–9]. However, these methods are faced with the issues of insufficient generalization ability and excessive dependence on prior knowledge. With the rapid advancement of Deep Neural Networks (DNNs), a suite of learning-based methods has been applied to Underwater Image Enhancement (UIE) [10–13]. The core of these methods is to use DNNs to learn the mapping function between degraded underwater images and clear images in a brute-force manner, in order to directly reconstruct undistorted underwater images. Although these learning-based methods are effective to some extent, they often struggle to finely address issues such as color bias and low contrast caused by underwater light attenuation and scattering.

To address these challenges, this paper presents a multi-task underwater image enhancement method based on Retinex theory [14], referred to as MTNet, whose core lies in decomposition and reconstruction. In the decomposition stage, MTNet decomposes the underwater image into an illumination map and a reflectance map. In the reconstruction stage, MTNet uses the illumination and reflectance maps to correct color distortion, reconstruct true colors, and enhance image details and adjust lighting using the illumination map. Specifically, the UIE task is decoupled by MTNet into three sub-tasks: image decomposition, color reconstruction, and illumination detail reconstruction. Three specialized sub-networks are designed for these tasks: DecomNet, DecolorNet, and DelightNet. To begin with, based on Retinex theory, the degraded underwater image is decomposed into an illumination map and a reflectance map using DecomNet. Subsequently, DecolorNet takes the illumination and reflectance maps as inputs, aiming to adjust the color information in the reflectance map while suppressing noise. Due to the blue-green bias in underwater images, we design a Multi-Branch Color Correction (MCC) module to capture the color feature distributions of each channel in the underwater image separately. Meanwhile, the illumination map serves as a constraint to further suppress noise in the reflectance map, thereby achieving better visual quality. Ultimately, to avoid the degradation of image texture and loss of detail information by DNNs, DelightNet employs continuous wavelet convolutions to construct a network for illumination map reconstruction. It also utilizes deep curve estimation for progressive pixel-level adjustments of the reconstructed image, effectively preserving the texture features of the illumination map. By modularizing the specific issues related to underwater image enhancement, we improve the enhancement results, making the reconstructed images in terms of color, contrast, and texture more closely aligned with reality. This approach enhances the model's adaptability and generalization capabilities under varying underwater environmental conditions.

The main contributions of this paper are summarized as follows:

1. This paper proposes a decomposition-reconstruction approach for underwater image enhancement. In the decomposition phase, the image is separated into an illumination map and a reflectance map, while the reconstruction phase seeks to adjust the color and illumination levels of the image to reconstruct a clear underwater image.
2. This paper introduces a multi-branch color correction network. It captures the color feature distributions of the R, G, and B channels in underwater images separately and performs intensity adjustments for each channel's information.

3. A dual-branch progressive illumination adjustment network is proposed. This network utilizes continuous wavelet convolutions to reconstruct the illumination map and employs iterative functions to progressively adjust the reconstructed image.

The remainder of this paper is organized as follows: Section 2 introduces related work, Section 3 describes the details of the proposed method, Section 4 presents the experimental results, Section 5 presents the results of ablation study, and Section 6 concludes the paper.

## 2 Related Work

The extant methodologies for enhancing underwater imagery can be categorized into three principal categories: model-free approaches, physical model-based approaches, and deep learning-based approaches.

### *2.1 Model-Free Approaches*

Model-free approaches directly manipulate the pixel-level information of images to enhance their visual quality. Reference [15] proposed a hybrid approach that combines color casts correction algorithms with the Retinex to rectify image colors. Reference [16] proposed a fusion-based technique that leverages white balance and histogram equalization to generate preliminary images, which are subsequently synthesized through a multiscale fusion strategy. Reference [17] presented a multiscale fusion strategy predicated on two color spaces, RGB and CIELAB, aimed at augmenting the textural attributes of underwater images. Reference [18] introduced a Retinex-based underwater image texture enhancement method, which employs bilateral filtering and trilateral filtering within the CIELAB color space to further refine the enhancement process. Despite these advancements, the application of traditional image processing techniques often leads to the introduction of artifacts and noise.

### *2.2 Physical Model-Based Approaches*

In contrast to model-free approaches, physical model-based approaches leverage the a priori knowledge of underwater environments to estimate the parameters of physical models [19] for simulating the underwater imaging process. Subsequently, the original image content is restored through the inverse solution of the model [20], on the basis of defogging algorithms [21], considering that the blue and green channels are the sources of underwater visual information, proposed the UDCP. Reference [22] have achieved the restoration of underwater images by leveraging the red channel prior. Reference [23] proposed a new model to explain the mechanism of light attenuation underwater. Reference [24] achieved image restoration by estimating the attenuation ratios. The quality of methods predicated on physical models hinges on prior knowledge and the models themselves, which can restrict the methods' ability to generalize.

### *2.3 Deep Learning-Based Approaches*

In recent years, deep learning-based approaches have achieved promising results in the field of UIE. UWCNN [2] learn mappings from degraded to high-definition images, reconstructing clear underwater imagery directly. Reference [25] presented a model that adeptly amalgamates the RGB and HSV color spaces within a unified network architecture. Reference [26] proposed WaterNet and concurrently established the Underwater Image Enhancement Benchmark (UIEB) dataset. Reference [27] introduced UGAN, leveraging the CycleGAN [28] framework to generate underwater images, which were then subjected to training with corresponding pairs of images. Reference [29] engineered a

lightweight network architecture characterized by its simplicity and parsimony in parameters. Reference [30] employed a two-stage enhancement architecture to accomplish the UIE task, successfully compressing the model parameters to 9 k. However, due to the lack of illumination in deep-sea environments, the majority of deep-sea images exhibit low-light characteristics, making low-light image enhancement [31] an important aspect of underwater image enhancement. Reference [32] combined traditional Retinex theory with CycleGAN to enhance low-light images without paired data. Reference [33] designed a noise data training strategy that introduces noise priors during training to generate better image structures. Reference [34] combined enhancement networks with classifiers to facilitate the integration of enhanced results with downstream tasks.

Deep learning-based methods, propelled by extensive repositories of underwater imagery data, typically outperform model-based and non-model-based approaches.

## 3 Proposed Method

This section elaborates a multi-task underwater image enhancement model based on Retinex theory, termed MTNet, which comprises three sub-networks: DecomNet, DecolorNet, and DelightNet. The architecture of MTNet is illustrated in Fig. 1. Specifically, DecomNet decomposes the degraded underwater image into illumination and reflection components. DecolorNet takes the illumination and reflectance maps as inputs, aiming to adjust the color information in the reflectance map while suppressing noise. DelightNet is responsible for adjusting the illumination map to enhance the image contrast. The final enhanced image is obtained by performing element-wise multiplication of the results from DecolorNet and DelightNet.

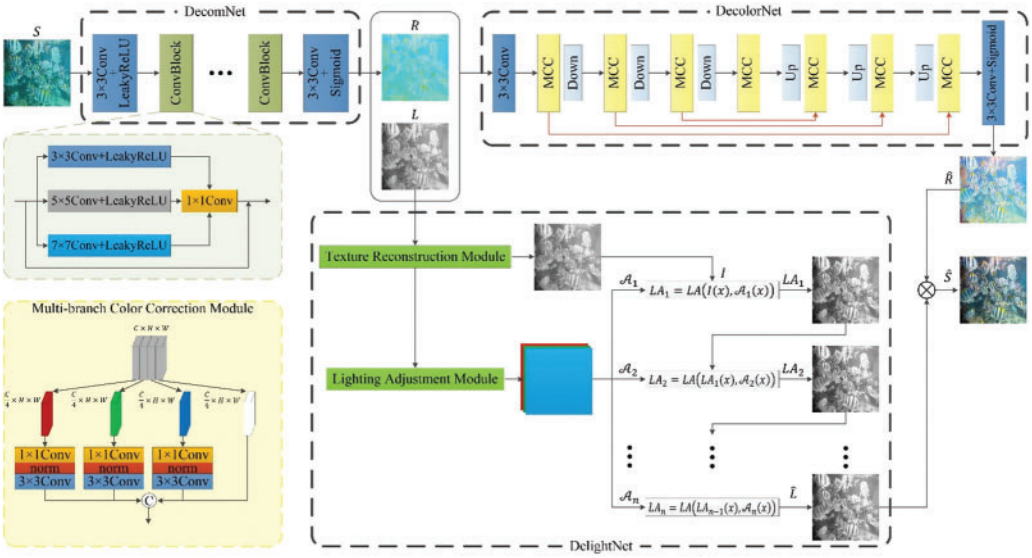


**Figure 1:** The construction of MTNet. MTNet consists of three sub-networks: DecomNet, DecolorNet, and DelightNet, which respectively perform tasks of image decomposition, color correction, texture reconstruction and lighting adjustment. DecomNet decomposes underwater images into illumination maps and reflectance maps. DecolorNet then performs color correction on the decomposed reflectance maps. DelightNet reconstructs the texture details of the illumination maps and adjusts the lighting

### *3.1 DecomNet*

According to the Retinex theory, underwater images can be represented as the product of the reflectance map and illumination map of the scene, expressed as:

$$S(x) = R(x) * L(x), \tag{1}$$

where $R(x)$ denotes the reflectance component, and $L(x)$ signifies the illumination component. Therefore, the key to reconstructing clear underwater images lies in obtaining high-quality illumination and reflectance maps. Since downsampling may lead to image texture distortion and information loss, especially in pixel-level computer vision tasks. This paper designs the DecomNet network as a continuous convolutional network without any downsampling. Additionally, considering the limitations of single-scale convolutional kernels in feature extraction, we introduce multi-scale convolutional kernels to provide multi-scale features for the decomposition task, enhancing the ability to capture both detailed and global information. Residual networks [35,36] help optimize deep neural networks during the training process, avoiding the vanishing or exploding gradient problem. Based on this, residual connections are used multiple times within the convolutional blocks to achieve better decomposition results. The specific design details of the DecomNet network are shown in Fig. 1, which consists of two single-scale $3 \times 3$ convolutional layers and five multi-scale convolutional layers with residual connections. Except for the last layer and $1 \times 1$ convolution, LeakyReLU is used as the activation function after each convolution, and the purpose of using a Sigmoid activation function in the last layer is to restrict the output values within the range of [0, 1].

### *3.2 DecolorNet*

Due to the absorptive properties of water with respect to light, light experiences energy attenuation during underwater transmission. Generally, red light attenuates the fastest in water, while blue and green light attenuate more slowly. This phenomenon results in underwater images typically exhibiting a bluish-green tint. Therefore, for underwater images, the decomposed reflectance map exhibits a significant blue-green bias, which can severely impact the final enhancement results. DecolorNet, which is designed to correct color bias, is a U-Net architecture composed of a series of MCC and convolutional layers. MCC adopts a multi-branch structure to capture and adjust the color feature distribution of the R, G, and B channels. The input is divided into four branches, with no parameter sharing between the branches. As shown in Fig. 1, three of the branches are processed using $1 \times 1$ convolutional layers, while the fourth branch is directly passed to the output layer. MCC utilizes $1 \times 1$ convolutions to capture color information at the pixel level and implement cross-channel feature interaction. Subsequently, normalization layers are applied to adjust the color feature distribution, and a $3 \times 3$ convolution is employed for local feature extraction. To reduce the model's parameter count and improve computational efficiency, MCC does not alter the number of feature maps in the feature branches during processing. The decomposed illumination map, serving as a constraint, is inputted into DecolorNet along with the reflectance map to ensure that important edge information is not lost during the color correction process.

### *3.3 DelightNet*

DelightNet is tasked with adjusting illumination to enhance the contrast of enhanced images and alleviate the blurring effects in underwater images. It consists of two modules: the Texture Reconstruction Module (TRM) and the Lighting Adjustment Module (LAM). The decomposed illumination map contains a wealth of textural details of the underwater image. However, conventional downsampling methods inevitably lead to information loss. Discrete Wavelet Transform (DWT) has

been widely applied in low-level vision tasks [37,38]. It can downsample the input to half of its original size, but due to the biorthogonal property of DWT, it does not result in information loss. Therefore, in the TRM, DWT is used in place of conventional downsampling methods. The specific structure of the TRM is shown in Fig. 2a, which is also a contraction-expansion network that utilizes a skip structure to further reduce the loss of feature information.

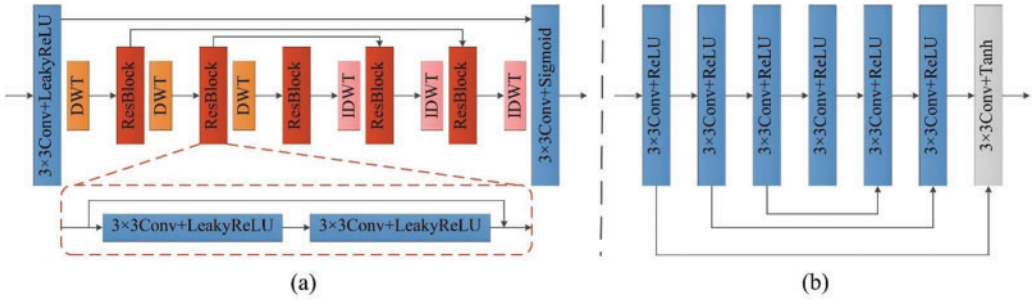


**Figure 2:** The detailed architecture of (a) Texture Reconstruction Module and (b) Lighting Adjustment Module

The LAM provides a progressive lighting adjustment curve for the output of the TRM. Inspired by [39], this paper applies depth curves to the adjustment of the illumination map. The LAM consists of seven convolutional blocks, with the specific structure shown in Fig. 2b. The first six convolutional blocks adopt the same structure, which includes only 3 × 3 convolutions and LeakyReLU activation functions. The last convolutional block is followed by a Tanh activation function to obtain pixel-level depth curve parameters. The lighting adjustment curve can be represented as:

$$LA_n(x) = LA_{n-1}(x) + \mathcal{A}(x)\, LA_{n-1}(x)\,(1 - LA_{n-1}(x)), \tag{2}$$

where $x$ denotes the pixel coordinates, $LA_n(x)$ represents the adjusted illumination map, and $\mathcal{A}(x)$ is a parameter map of the same size as the input image.

### *3.4 Loss Function*

During the training phase, the three sub-networks are trained separately. For DecomNet, it computes the L1 loss and perceptual loss between the reflectance map $R_o$ and illumination map $\mathrm{L}_o$ egmented from the original underwater image $S_o$ and the reflectance map $R_{gt}$ and illumination map $L_{gt}$ of the reference image $S_{gt}$. The L1 loss can be represented as:

$$L_1 = \mathrm{E}[||S_o - R_o \mathrm{L}_o||_1] + \mathrm{E}\left[\left|\left|S_{gt} - R_{gt}\mathrm{L}_{gt}\right|\right|_1\right] + \mathrm{E}\left[\left|\left|S_{gt} - R_o\mathrm{L}_{gt}\right|\right|_1\right] + E\left[||S_o - R_{gt}L_o||_1\right]. \tag{3}$$

The content loss function is expressed as follows:

$$L_{con} = \mathrm{E}[||\phi(S_o) - \phi(R_o\mathrm{L}_o)||_2] + \mathrm{E}\left[||\phi(S_{gt}) - \phi(R_{gt}\mathrm{L}_{gt})||_2\right], \tag{4}$$

where $\phi(\cdot)$ is defined as the feature map from the block4_conv3 layer of VGG-16. The total loss of DecomNet can be represented as:

$$L_{com} = L_1 + \lambda_{con} L_{con}, \tag{5}$$

where $\lambda_{con} = 0.1$ is the weight that balance the contribution of the L1 loss and perceptual loss.

For DecolorNet, it calculates the L1 loss, perceptual loss, and Structural Similarity Index Measure (SSIM) loss between the corrected reflectance map $\hat{R}_o$ and the reference image's reflectance map $R_{gt}$.

The L1 loss can be represented as:

$$L_1 = \mathrm{E}\left[||R_{gt} - \hat{R}_o||_1\right]. \tag{6}$$

The content loss function is expressed as follows:

$$L_{con} = \mathrm{E}\left[||\phi(R_{gt}) - \phi(\hat{R}_o)||_2\right], \tag{7}$$

where $\phi()$ is defined as the feature map from the *block4_conv3* layer of VGG-16. In general, the SSIM [40] score is used to measure the similarity of image textures. The formula is as follows:

$$SSIM\,(x, y) = \mathrm{l}\,(x, y)^{\alpha} * \mathrm{c}\,(x, y)^{\beta} * \mathrm{s}\,(x, y)^{\gamma}, \tag{8}$$

where $x$ and $y$ represent images, and $\mathrm{l}\,(x, y)$, $\mathrm{c}\,(x, y)$, and $\mathrm{s}\,(x, y)$ respectively compare the Luminance, contrast, and structure between the images. The SSIM loss is expressed as follows:

$$L_{SSIM} = \mathrm{E}\left[1 - SSIM(R_{gt}, \hat{R}_o)\right], \tag{9}$$

The total loss of DecolorNet can be represented as:

$$L_{color} = L_1 + \lambda_{con} L_{con} + \lambda_{SSIM} L_{SSIM}, \tag{10}$$

where $\lambda_{con} = 0.1$ and $\lambda_{con} = 0.4$.

For DelightNet, it calculates the L1 loss, perceptual loss, and SSIM loss between the corrected enhanced image $\hat{S}_o$ and the reference image $S_{gt}$. The L1 loss can be represented as:

$$L_1 = \mathrm{E}\left[||S_{gt} - \hat{S}_o||_1\right], \tag{11}$$

where $\hat{S}_o$ is the product of the corrected reflectance map and the adjusted illumination map. The content loss function is expressed as follows:

$$L_{con} = \mathrm{E}\left[||\phi(S_{gt}) - \phi(\hat{S}_o)||_2\right], \tag{12}$$

where $\phi(\cdot)$ is defined as the feature map from the block4_conv3 layer of VGG-16. The SSIM loss is expressed as follows:

$$L_{SSIM} = \mathrm{E}\left[1 - SSIM(S_{gt}, \hat{S}_o)\right]. \tag{13}$$

The total loss of DelightNet can be represented as:

$$L_{light} = L_1 + \lambda_{con} L_{con} + \lambda_{SSIM} L_{SSIM}, \tag{14}$$

where $\lambda_{con} = 0.1$ and $\lambda_{SSIM} = 0.5$

## 4 Experiments

To assess the performance of MTNet, this section conducts both qualitative and quantitative comparisons of MTNet with several representative methods on the synthetic dataset EUVP [29] and the real-world underwater image datasets UIEB [26] and LSUI [41]. For synthetic underwater images, we randomly selected 800 pairs of synthetically paired images from EUVP to serve as the test set, denoted as T800. For real-world underwater images, we selected 90 pairs of paired images and 60 unpaired images from the dataset UIEB to be used as test images, denoted as T90 and T60, respectively. We also selected 879 pairs of images from LSUI to form a rich test set, denoted

as T879. The comparison methods included two model-free approaches (UDCP [20], ULAP [42]), two physical model-based approaches (Retinex-based [15], UCM [43]), and six deep learning-based approaches (MWCNN [44], UGAN-P [27], Shallow-UWnet [45], FUnIE-GAN [29], UWCNN [2], UIEC^2-Net [25]). To ensure optimal performance, the implementations followed the parameter settings recommended by their respective authors.

### *4.1 Experiment Settings*

The MTNet is implemented within the PyTorch framework and trained on a GeForce RTX 3090 (24 G) platform. During the training process, both the training and test images are cropped to a size of 256 × 256 and are randomly adjusted brightness, saturation, contrast, and hue. The $\beta_1$ and $\beta_2$ are set to 0.9 and 0.999. Adam was used as the optimizer. Each sub-network has a fixed learning rate of 0.0001 and a batch size of 8. Each sub-network was trained for 200 epochs to obtain the final model parameters.

### *4.2 Experiment on Synthetic Datasets*

Underwater environments pose significant challenges for acquiring paired image data. Therefore, synthesizing paired underwater images to facilitate model training has become an effective approach. This section presents a qualitative and quantitative comparison of MTNet with other methods on the synthetic dataset T800. Visual results are detailed in Fig. 3.

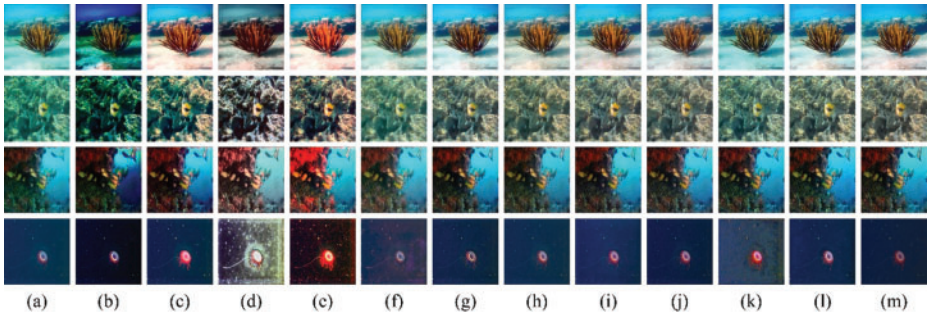


**Figure 3:** Subjective visual comparison of enhanced results sampled from the T800. From left to right are (a) original underwater images, the results of (b) UDCP [20], (c) ULAP [42], (d) Retinex-based [15], (e) UCM [43], (f) MWCNN [44], (g) UGAN-P [27], (h) Shallow-UWnet [45], (i) FUnIE-GAN [29], (j) UWCNN [2], (k) UIEC^2-Net [25], (l) MTNet and (m) reference images

As shown in Fig. 3, UDCP [20] fails to correct color bias. ULAP [42] relies on physical models, making it difficult to accurately infer the imaging process, which sometimes leads to image distortion. Retinex-based [15] and UCM [43] lack generalization, and as shown in Fig. 3d,e, their enhanced results exhibit significant white and red noise. MWCNN [44] successfully enhances image details due to its unique wavelet transformation, but it shows insufficient enhancement in color saturation. UGAN-P [27] generates checkerboard artifacts in the image. Shallow-UWnet [45] and UWCNN [2] only use simple convolutional structures, with their enhancement results showing a large gap from the reference image. UIEC^2-Net [25] fails to improve the blue color bias in the image and generates artifacts in dark areas, as shown in Fig. 3k. Compared to these methods, MTNet effectively restores the image color and alleviates the low contrast issue through detail reconstruction and illumination adjustment, making it visually closer to the reference image.

For a quantitative assessment of various methods, this section employs the MSE, SSIM [40] and Peak signal-to-noise ratio (PSNR) [46] to measure the pixel-level discrepancies, content differences, and structural consistency between the restored or enhanced images and the reference images, respectively. Typically, the higher the quality of image enhancement, the lower the MSE becomes, while the PSNR and SSIM indices increase correspondingly. Table 1 presents the quantitative evaluation results of various methods on the T800 dataset. Compared to other methods, MTNet achieved the best performance in all three metrics, indicating that the enhancement results of MTNet are structurally and content-wise closer to the reference image.

**Table 1:** Quantitative results of different methods on T800

| Methods | MSE | SSIM | PSNR (dB) |
|---|---|---|---|
| UDCP [20] | 3044.062 | 0.562 | 16.429 |
| ULAP [42] | 1061.054 | 0.689 | 21.383 |
| Retinex-based [15] | 1860.767 | 0.646 | 18.163 |
| UCM [43] | 1180.458 | 0.748 | 21.721 |
| MWCNN [44] | 675.782 | 0.742 | 23.210 |
| UGAN-P [27] | 350.542 | 0.794 | 25.586 |
| Shallow-UWnet [45] | 401.982 | 0.787 | 25.601 |
| FUnIE-GAN [29] | 375.343 | 0.764 | 25.099 |
| UWCNN [2] | 399.718 | 0.792 | 25.517 |
| UIEC^2-Net [25] | 415.869 | 0.814 | 26.225 |
| Ours | **239.186** | **0.833** | **27.049** |

Note: The best results are highlighted in bold.

### *4.3 Evaluation on Real-World Underwater Images*

In order to further validate the generalization capability of MTNet, we conducted tests on MTNet and other comparison methods using the T879 and T90. In contrast to the synthetic underwater datasets, T879 contains a richer variety of underwater scenes (Fig. 4), including different water types, lighting conditions, and target categories. T90 encompasses a diverse array of real-world underwater scenes, including the blurring effects of shallow water areas (Fig. 5) and common colors such as green (Fig. 6), blue (Fig. 7) and yellow (Fig. 8). These characteristics are pervasive within the realm of subaquatic imagery, exerting a profound influence on the fidelity and color of the images.

As shown in Fig. 4, UDCP [20] and ULAP [42] are model-based methods that heavily rely on physical models and prior knowledge. As a result, they struggle to accurately simulate imaging in complex underwater environments and cannot recover the image accurately. The ULAP [42] method weakens the green channel in the image but fails to properly compensate for the red channel, resulting in a grayish-white tone. UCM [43] excessively enhances the red channel, leading to an overemphasis of the red color bias in the image. MWCNN [44] preserves sufficient image structural information, but it fails to correctly recover the image color. UGAN-P [27] tends to produce checkerboard artifacts in the enhanced images, which distorts some of the image structure. Shallow-UWnet [45] and UWCNN [2] use simple network architectures, which lack expressive ability, resulting in generated images with some degree of haze. FUnIE-GAN [29] introduces color noise into the image, as shown in Fig. 5i. UIEC^2-Net [25] can enhance image details, but it incorrectly amplifies the blue tone of the image,

giving the overall image a blue tint. In contrast, as shown in Fig. 4l, MTNet is able to accurately restore the colors of underwater scenes with a bluish-green tint, indicating that the proposed method demonstrates a significant advantage in color restoration.

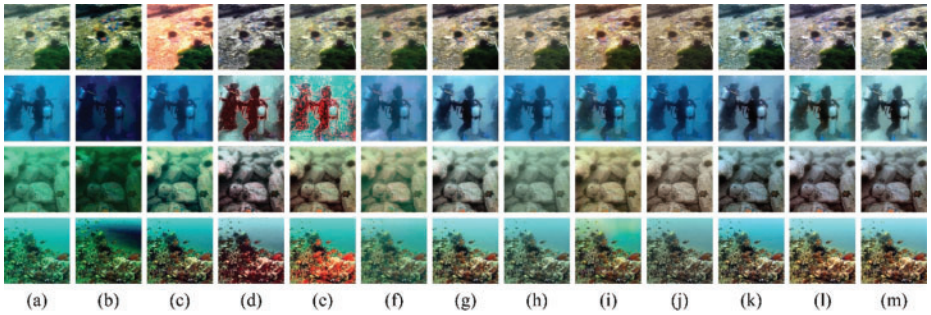


**Figure 4:** The subjective visual comparison of the enhancement results sampled from the T879. From left to right are (a) original underwater images, the results of (b) UDCP [20], (c) ULAP [42], (d) Retinex-based [15], (e) UCM [43], (f) MWCNN [44], (g) UGAN-P [27], (h) Shallow-UWnet [45], (i) FUnIE-GAN [29], (j) UWCNN [2], (k) UIEC^2-Net [25], (l) MTNet and (m) reference images

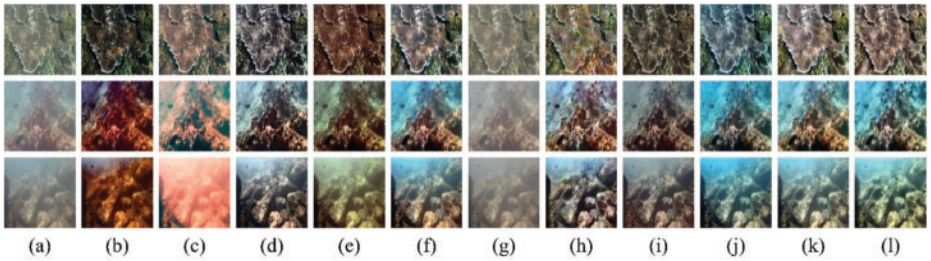


**Figure 5:** The subjective visual comparison of the enhancement results for the blurred underwater images from shallow water areas sampled from the T90. From left to right are (a) original underwater images, the results of (b) UDCP [20], (c) ULAP [42], (d) Retinex-based [15], (e) UCM [43], (f) MWCNN [44], (g) UGAN-P [27], (h) Shallow-UWnet [45], (i) FUnIE-GAN [29], (j) UWCNN [2], (k) UIEC^2-Net [25], (l) MTNet

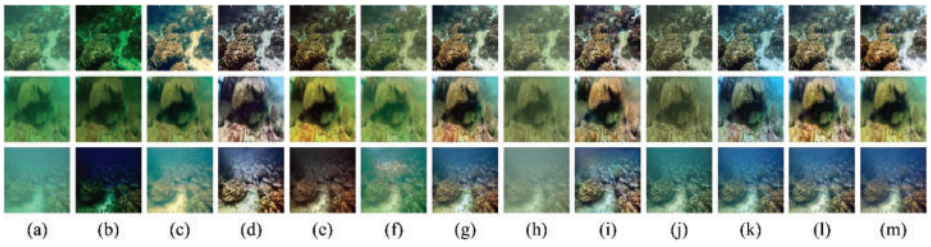


**Figure 6:** The subjective visual comparison of the enhancement results for the green underwater images sampled from the T90. From left to right are (a) original underwater images, the results of (b) UDCP [20], (c) ULAP [42], (d) Retinex-based [15], (e) UCM [43], (f) MWCNN [44], (g) UGAN-P [27], (h) Shallow-UWnet [45], (i) FUnIE-GAN [29], (j) UWCNN [2], (k) UIEC^2-Net [25], (l) MTNet and (m) reference images

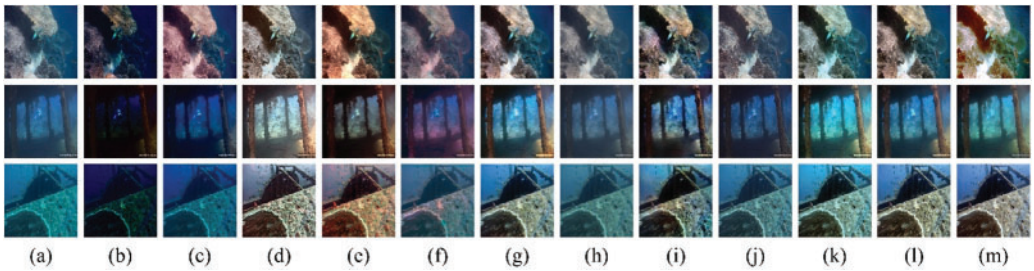


**Figure 7:** The subjective visual comparison of the enhancement results for the blue underwater images sampled from the T90. From left to right are (a) original underwater images, the results of (b) UDCP [20], (c) ULAP [42], (d) Retinex-based [15], (e) UCM [43], (f) MWCNN [44], (g) UGAN-P [27], (h) Shallow-UWnet [45], (i) FUnIE-GAN [29], (j) UWCNN [2], (k) UIEC^2-Net [25], (l) MTNet and (m) reference images

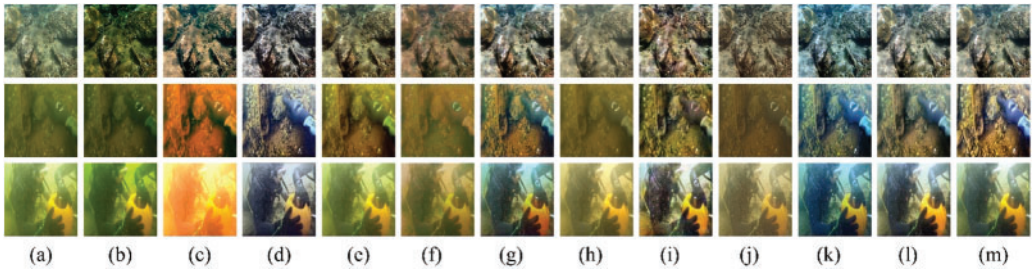


**Figure 8:** The subjective visual comparison of the enhancement results for the yellow underwater images sampled from the T90. From left to right are (a) original underwater images, the results of (b) UDCP [20], (c) ULAP [42], (d) Retinex-based [15], (e) UCM [43], (f) MWCNN [44], (g) UGAN-P [27], (h) Shallow-UWnet [45], (i) FUnIE-GAN [29], (j) UWCNN [2], (k) UIEC^2-Net [25], (l) MTNet and (m) reference images

Similar to Section 4.2, we choose MSE, PSNR, and SSIM to evaluate the enhancement results of each method on real-world underwater images. The quantitative results of different methods on T879 and T90 are presented in Tables 2 and 3. Among them, the best score is indicated in bold. It is evident that MTNet outperforms its peers on all three metrics across both test sets, meaning that MTNet is more competitive than other methods in preserving image structural information and content restoration.

**Table 2:** Quantitative results of different methods on T879

| Methods | MSE | SSIM | PSNR (dB) |
|---|---|---|---|
| UDCP [20] | 4216.245 | 0.546 | 14.453 |
| ULAP [42] | 1547.289 | 0.723 | 20.906 |
| Retinex-based [15] | 1547.941 | 0.682 | 18.262 |
| UCM [43] | 1044.551 | 0.796 | 22.939 |
| MWCNN [44] | 790.683 | 0.782 | 22.900 |
| UGAN-P [27] | 345.371 | 0.819 | 25.412 |

(Continued)

**Table 2 (continued)**

| Methods | MSE | SSIM | PSNR (dB) |
|---|---|---|---|
| Shallow-UWnet [45] | 649.632 | 0.813 | 23.910 |
| FUnIE-GAN [29] | 842.241 | 0.767 | 23.791 |
| UWCNN [2] | 548.053 | 0.842 | 24.267 |
| UIEC^2-Net [25] | 449.672 | 0.869 | 26.259 |
| Ours | **192.164** | **0.885** | **28.768** |

Note: The best results are highlighted in bold.

**Table 3:** Quantitative results of different methods on T90

| Methods | MSE | SSIM | PSNR (dB) |
|---|---|---|---|
| UDCP [20] | 5916.636 | 0.498 | 12.062 |
| ULAP [42] | 2258.868 | 0.744 | 19.155 |
| Retinex-based [15] | 1238.529 | 0.676 | 20.136 |
| UCM [43] | 1818.908 | 0.795 | 20.889 |
| MWCNN [44] | 1043.352 | 0.834 | 23.223 |
| UGAN-P [27] | 627.567 | 0.768 | 23.210 |
| Shallow-UWnet [45] | 1493.890 | 0.756 | 19.145 |
| FUnIE-GAN [29] | 1010.414 | 0.783 | 20.436 |
| UWCNN [2] | 1154.478 | 0.838 | 20.647 |
| UIEC^2-Net [25] | 763.383 | 0.882 | 23.459 |
| Ours | **453.480** | **0.901** | **25.145** |

Note: The best results are highlighted in bold.

We continue to demonstrate the performance of MTNet based on challenging underwater images in the T60, and present enhancement results of various methods in Fig. 9. Suffering from severely light absorption and scattering, challenging underwater images can hardly be recovered well.

As depicted in Fig. 9, the majority of the methods struggle to illuminate severely dark underwater scenes. The Retinex-based [15] is capable of enhancing dark scenes but introduces significant white noise. UGAN-P [27], while illuminating the underwater environment, imparts a distinct checkerboard artifact to the image. Concurrently, UIEC^2-Net [25] continues to exhibit a tendency to amplify blue tones in the background, thereby adversely affecting the overall visual appeal of the image. In contrast, MTNet not only successfully improves the illumination of severely dark underwater scenes and enhances color contrast but also effectively preserves the detailed textures of the image.

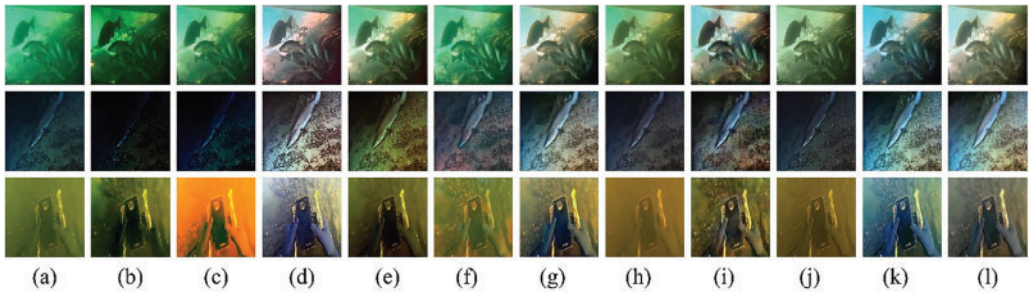


**Figure 9:** Subjective visual comparison of enhanced results sampled from the T60. From left to right are (a) original underwater images, the results of (b) UDCP [20], (c) ULAP [42], (d) Retinex-based [15], (e) UCM [43], (f) MWCNN [44], (g) UGAN-P [27], (h) Shallow-UWnet [45], (i) FUnIE-GAN [29], (j) UWCNN [2], (k) UIEC^2-Net [25], and (l) MTNet

Considering that there is no corresponding reference image in T60, we evaluated the performance of the different methods by UIQM [47], UCIQE [48], Twice Mixing [49], BRISQUE [50], and NIQE [51] and listed the scores in Table 4. For Twice Mixing [49], we used the pre-training model provided by the author to predict the score. Higher scores on the first three metrics are indicative of superior quality in the enhanced images. In the BRISQUE [50] and NIQE [51] algorithms, a lower score indicates better image quality. It is noteworthy that MTNet only achieved a lower ranking in the UCIQE and UIQM metrics. However, for the UCIQE metric, although the UCM [43] obtained the highest score, it over-enhanced the red channel, resulting in a red cast. Regarding the UIQM, the checkerboard artifacts introduced by UGAN-P [27] also adversely affected the visual quality of the images to some extent. Therefore, it is hypothesized in this study that these two metrics may not fully represent human visual perception, and the lower ranking of MTNet is considered acceptable. Furthermore, MTNet ranked first in Twice Mixing, which also serves to illustrate its outstanding visual performance. The enhanced images from MTNet achieved third and fifth place in BRISQUE and NIQE, respectively. Although MWCNN [44], achieved the top rank in BRISQUE, its visual effect is far from the reference image. The Retinex-based [15], despite its good NIQE score, tends to render all images in a grayish hue, lacking generalization ability. The proposed approach, while not ranking high, demonstrates significant advantages in terms of visual effect.

**Table 4:** Quantitative results of different methods on T60

| Methods | UCIQE | UIQM | Twice Mixing | BRISQUE | NIQE |
|---|---|---|---|---|---|
| UDCP [20] | 0.520 | 1.215 | 0.072 | 26.058 | 39.143 |
| ULAP [42] | 0.560 | 1.594 | 0.275 | 26.736 | 37.847 |
| Retinex-based [15] | 0.602 | 2.877 | 0.372 | 22.702 | **36.785** |
| UCM [43] | **0.608** | 2.367 | 0.334 | 20.961 | 41.672 |
| MWCNN [44] | 0.538 | 2.975 | 1.295 | **10.334** | 43.093 |
| UGAN-P [27] | 0.597 | **3.336** | 0.852 | 22.226 | 43.593 |
| Shallow-UWnet [45] | 0.518 | 2.414 | 0.906 | 33.440 | 43.992 |
| FUnIE-GAN [29] | 0.580 | 3.263 | 1.476 | 15.357 | 39.899 |
| UWCNN [2] | 0.519 | 2.578 | 0.998 | 22.307 | 43.059 |
| UIEC^2-Net [25] | 0.589 | 2.910 | 1.575 | 17.024 | 41.486 |

(Continued)

**Table 4 (continued)**

| Methods | UCIQE | UIQM | Twice Mixing | BRISQUE | NIQE |
|---|---|---|---|---|---|
| Ours | 0.586 | 3.022 | **1.621** | 15.771 | 40.956 |

Note: The best results are highlighted in bold.

### *4.4 The Practical Application*

To validate the positive impact of MTNet on vision-based tasks, we selected SIFT [52] to test the effect of different underwater image enhancement methods on feature matching tasks. The feature matching results of the enhanced underwater images are shown in Fig. 10. The numbers in the top-left corner indicate the number of keypoints successfully matched under the same scene and matching conditions. As shown in Fig. 10, the color distortion and blurring present in the original underwater images severely limit the performance of underwater feature matching algorithms. Moreover, compared to other methods, MTNet effectively increases the number of keypoints matched in underwater feature matching.

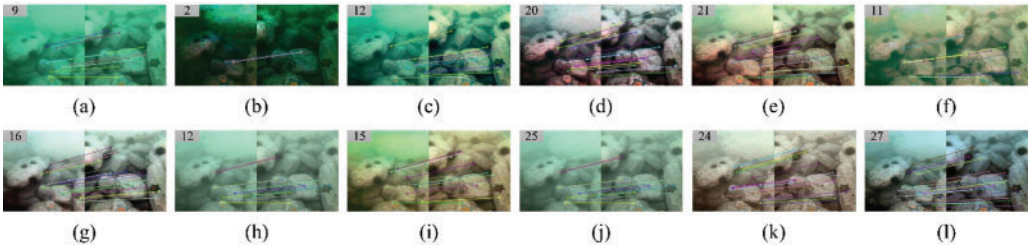


**Figure 10:** The feature matching results of SIFT. From left to right are the matching results of (a) original underwater images, the results of (b) UDCP [20], (c) ULAP [42], (d) Retinex-based [15], (e) UCM [43], (f) MWCNN [44], (g) UGAN-P [27], (h) Shallow-UWnet [45], (i) FUnIE-GAN [29], (j) UWCNN [2], (k) UIEC^2-Net [25], and (l) MTNet

In addition, the practical application requirements of various comparison methods are documented in Table 5. Since traditional machine learning methods typically do not rely on GPU acceleration, we only computed the running speed, computational load, and parameter volume of the learning-based methods on the same GPU. Compared to other learning-based methods, MTNet ranks 5th in terms of running speed and computational load, and 4th in terms of parameter volume. This result indicates that although MTNet is not the most optimal in terms of speed, computational load, and parameter volume, it is still sufficient to meet the requirements for real-time applications. While its running speed is not the fastest, we believe that the slight trade-off in speed is reasonable given the outstanding performance of MTNet in image enhancement. In future work, we will explore optimization strategies, including model pruning, quantization, and knowledge distillation, to reduce the parameter volume and computational demands while striving to maintain the enhancement of visual quality.

**Table 5:** The application requirements of different methods on the GPU

| Methods | MWCNN [44] | UGAN-P [27] | Shallow-UWnet [45] | FUnIE-GAN [29] | UWCNN [2] | UIEC^2-Net [25] | Ours |
|---|---|---|---|---|---|---|---|
| FPS(f/s) | 44.036 | 155.736 | 123.377 | **265.592** | 142.858 | 36.859 | 57.127 |
| Params(M) | 13.819 | 54.404 | **0.214** | 7.020 | 0.039 | 0.522 | 6.838 |
| GFLOPs(G) | 55.952 | 18.155 | 21.630 | 10.239 | **2.610** | 10.239 | 21.873 |

Note: The best results are highlighted in bold.

### *4.5 Limitations and Failure Cases*

Although MTNet has made certain progress in addressing the color bias and low contrast issues in most underwater images, it still struggles to restore the colors of underwater images with severe greenish tones to match the reference image. As shown in Fig. 11, the underwater image exhibits significant green color bias, and in this case, MTNet fails to remove the greenish hue from the image. The likely cause of this unsatisfactory result is the limited number of extreme color bias samples in the training dataset. To improve the model's ability to handle such images, we plan to add more underwater image samples with extreme color biases to the training set.

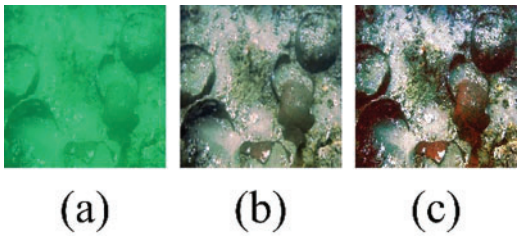


**Figure 11:** The failure cases. From left to right are the matching results of (a) original image, (b) the result of MTNet, and (c) reference image

## 5 Ablation Study

This section presents ablation studies conducted to ascertain the contributions of various components within the MTNet network architecture. To evaluate the role of the MCC module within the DecolorNet framework, this study replaces the MCC module with a traditional $3 \times 3$ convolution, denoted as "w/o MCC". Furthermore, to verify the effects of the TRM and LAM in DelightNet, this research examines the performance changes of the model after removing each module individually, labeled as "w/o TRM" and "w/o LAM", respectively. The results of the ablation study, as shown in Fig. 12, clearly demonstrate the impact of each component on the model's performance.

The results of the ablation study, as depicted in Fig. 12, distinctly elucidate the impact of each component on the model's performance. Notably, the DecolorNet devoid of the MCC module fails to accurately correct image color, exhibiting a significant color bias compared to the enhancement outcomes of MTNet and the reference images.

As illustrated in Fig. 12c, the TRM is capable of meticulously reconstructing the illumination map. Due to the low contrast issue inherent in underwater images, the details within the decomposed

illumination map are far less rich than those in the reference image's illumination map, a fact that this experiment effectively demonstrates the necessity of TRM in the task of detail reconstruction. Moreover, the structure of LAM is relatively simple, and it cannot learn all the information on its own. The data in Table 6 supports this assertion.

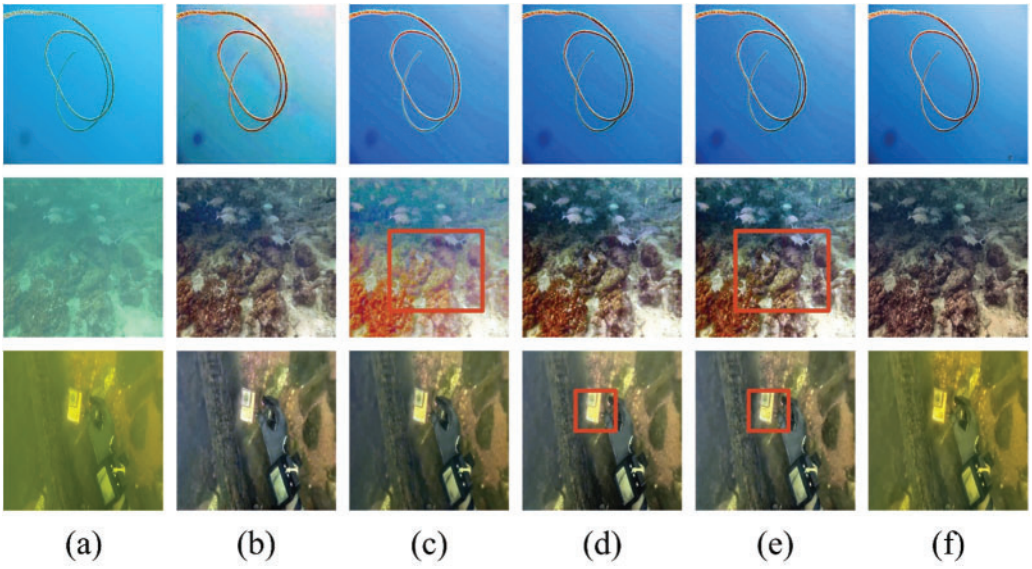


**Figure 12:** Enhancement results of ablation study. From left to right are (a) original underwater images and the results of (b) w/o MCC, (c) w/o TRM, (d) w/o LAM, (e) results of MTNet, and (f) reference images

**Table 6:** Quantitative results of ablation study

| Methods | MSE | SSIM | PSNR (dB) |
|---|---|---|---|
| w/o MCC | 498.871 | 0.892 | 24.718 |
| w/o TRM | 910.010 | 0.861 | 21.378 |
| w/o LAM | 459.224 | 0.898 | 25.089 |
| Ours | 453.480 | 0.901 | 25.145 |

The role of LAM is to further adjust the illumination map to preserve image details and enhance lighting effects. As shown in Fig. 12, the removal of LAM leads to localized overexposure in the model's enhancement results, which detracts from the visual quality of the images.

## 6 Conclusion

This paper presents a multi-task underwater image enhancement method based on Retinex, referred to as MTNet. The network decomposes the underwater image enhancement task into multiple sub-tasks and designs three dedicated sub-networks: DecomNet, DecolorNet and DelightNet, to specifically address the tasks of color correction and contrast enhancement for underwater images. To tackle the issue of color bias in underwater images, a multi-branch color correction strategy is employed to accurately restore the image colors. To address the problem of low contrast, texture

reconstruction is performed on the illumination map to enhance its texture information, followed by a progressive pixel-level adjustment of the enhanced illumination map to improve the visual quality of the enhanced image. Qualitative and quantitative evaluations on both synthetic and real-world underwater datasets demonstrate that the proposed method outperforms several representative underwater image processing techniques. Furthermore, various ablation studies clearly illustrate the contributions of each component within MTNet.

**Acknowledgement:** Not applicable.

**Funding Statement:** This study was supported by National Natural Science Foundation of China (No. 62202210).

**Author Contributions:** The authors confirm contribution to the paper as follows: study conception and design: Yang Song, Xing Deng; data collection: Yang Song, Haijian Shao; analysis and interpretation of results: Yang Song, Xing Deng, Haijian Shao; draft manuscript preparation: Yang Song, Haijian Shao, Fei Wang. All authors reviewed the results and approved the final version of the manuscript.

**Availability of Data and Materials:** The data that support the findings of this study are available from the corresponding author upon reasonable request.

**Ethics Approval:** Not applicable.

**Conflicts of Interest:** The authors declare no conflicts of interest to report regarding the present study.

## References

1. Zhang Z, Zhang L, Wang L, Wu H. Scene 3-D reconstruction system in scattering medium. Comput Mater Contin. 2024;80(2):3405–20. doi:10.32604/cmc.2024.052144.
2. Li C, Anwar S, Porikli F. Underwater scene prior inspired deep underwater image and video enhancement. Pattern Recognit. 2020;98(1):107038. doi:10.1016/j.patcog.2019.107038.
3. Zhou J, Yang T, Zhang W. Underwater vision enhancement technologies: a comprehensive review, challenges, and recent trends. Appl Intell. 2022;53(3):3594–621. doi:10.1007/s10489-022-03767-y.
4. Abdul Ghani AS, Mat Isa NA. Enhancement of low quality underwater image through integrated global and local contrast correction. Appl Soft Comput. 2015;37(1):332–44. doi:10.1016/j.asoc.2015.08.033.
5. Fu X, Fan Z, Ling M, Huang Y, Ding X. Two-step approach for single underwater image enhancement. In: 2017 International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS); 2017; Piscataway, NJ: IEEE. p. 789–94.
6. Abdul Ghani AS, Mat Isa NA. Underwater image quality enhancement through integrated color model with rayleigh distribution. Appl Soft Comput. 2015;27:219–30. doi:10.1016/j.asoc.2014.11.020.
7. Drews PL, Nascimento ER, Botelho SS, Montenegro Campos MF. Underwater depth estimation and image restoration based on single images. IEEE Comput Graph Appl. 2016;36(2):24–35. doi:10.1109/MCG.2016.26.
8. Alhajlah M. Underwater image enhancement using customized CLAHE and adaptive color correction. Comput Mater Contin. 2023;74(3):5157–72. doi:10.32604/cmc.2023.033339.
9. Peng Y-T, Cao K, Cosman PC. Generalization of the dark channel prior for single image restoration. IEEE Trans Image Process. 2018;27(6):2856–68. doi:10.1109/TIP.2018.2813092.
10. Aswathy K, Cherian EP. A novel AlphaSRGAN for underwater image super resolution. Comput Mater Contin. 2021;69(2):1537–52. doi:10.32604/cmc.2021.018213.

11. Ren Q, Xiang Y, Wang G, Gao J, Wu Y, Chen R-P. The underwater polarization dehazing imaging with a lightweight convolutional neural network. Optik. 2022;251:168381. doi:10.1016/j.ijleo.2021.168381.
12. Lin P, Wang Y, Wang G, Yan X, Jiang G, Fu X. Conditional generative adversarial network with dual-branch progressive generator for underwater image enhancement. Signal Process: Image Commun. 2022;108:116805. doi:10.1016/j.image.2022.116805.
13. Wu J, Liu X, Lu Q, Lin Z, Qin N, Shi Q. FW-GAN: underwater image enhancement using generative adversarial network with multi-scale fusion. Signal Process: Image Commun. 2022;109:116855. doi:10.1016/j.image.2022.116855.
14. Land EH. Recent advances in Retinex theory. Vision Res. 1986;26(1):7–21. doi:10.1016/0042-6989(86)90067-2.
15. Fu X, Zhuang P, Huang Y, Liao Y, Zhang X-P, Ding X. A Retinex-based enhancing approach for single underwater image. In: 2014 IEEE International Conference on Image Processing (ICIP); 2014; Piscataway, NJ: IEEE. p. 4572–6.
16. Ancuti C, Ancuti CO, Haber T, Bekaert P. Enhancing underwater images and videos by fusion. In: 2012 IEEE Conference on Computer Vision and Pattern Recognition; 2012; Piscataway, NJ: IEEE. p. 81–8.
17. Yuan J, Cai Z, Cao W. TEBCF: real-world underwater image texture enhancement model based on blurriness and color fusion. IEEE Trans Geosci Remote Sens. 2022;60:1–15.
18. Zhang S, Wang T, Dong J, Yu H. Underwater image enhancement via extended multi-scale Retinex. Neurocomputing. 2017;245:1–9. doi:10.1016/j.neucom.2017.03.029.
19. McGlamery BL. A Computer model for underwater camera systems. In: Duntley SQ, editor. Ocean optics VI. Washington, DC, USA: SPIE; 1980. vol. 0208, p. 221–31.
20. Drews PJr, Moraes F, Botelho S, Campos M. Transmission estimation in underwater single images. In: 2013 IEEE International Conference on Computer Vision Workshops; 2013; Piscataway, NJ: IEEE. p. 825–30.
21. He K, Sun J, Tang X. Single image haze removal using dark channel prior. IEEE Trans Pattern Anal Mach Intell. 2011;33(12):2341–53. doi:10.1109/TPAMI.2010.168.
22. Galdran A, Pardo D, Picón A, Alvarez-Gila A. Automatic red-channel underwater image restoration. J Vis Commun Image Represent. 2015;26:132–45. doi:10.1016/j.jvcir.2014.11.006.
23. Akkaynak D, Treibitz T. Sea-Thru: a method for removing water from underwater images. In: 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR); 2019; Piscataway, NJ: IEEE. p. 1682–91.
24. Berman D, Levy D, Avidan S, Treibitz T. Underwater single image color restoration using haze-lines and a new quantitative dataset. IEEE Trans Pattern Anal Mach Intell. 2021;43(8):2822–37. doi:10.1109/TPAMI.2020.2977624.
25. Wang Y, Guo J, Gao H, Yue H. UIEC^2-Net: CNN-based underwater image enhancement using two color space. Signal Process: Image Commun. 2021;96:116250. doi:10.1016/j.image.2021.116250.
26. Li C, Guo C, Ren W, Cong R, Hou J, Kwong S, et al. An underwater image enhancement benchmark dataset and beyond. IEEE Trans Image Process. 2020;29:4376–89. doi:10.1109/TIP.2019.2955241.
27. Fabbri C, Islam MJ, Sattar J. Enhancing underwater imagery using generative adversarial networks. In: 2018 IEEE International Conference on Robotics and Automation (ICRA); 2018; Piscataway, NJ: IEEE. p. 7159–65.
28. Zhu J-Y, Park T, Isola P, Efros AA. Unpaired image-to-image translation using cycle-consistent adversarial networks. In: 2017 IEEE International Conference on Computer Vision (ICCV); 2017; Piscataway, NJ: IEEE. p. 2242–51.
29. Islam MJ, Xia Y, Sattar J. Fast underwater image enhancement for improved visual perception. IEEE Robot Autom Lett. 2020;5(2):3227–34. doi:10.1109/LRA.2020.2974710.
30. Jiang J, Ye T, Bai J, Chen S, Chai W, Jun S, et al. Five $A^+$ Network: you only need 9K parameters for underwater image enhancement. arXiv:2305.08824. 2023.

31. Al Sobbahi R, Tekli J. Comparing deep learning models for low-light natural scene image enhancement and their impact on object detection and classification: overview, empirical evaluation, and challenges. Signal Process: Image Commun. 2022;109:116848. doi:10.1016/j.image.2022.116848.
32. Wu K, Huang J, Ma Y, Fan F, Ma J. Cycle-Retinex: Unpaired low-light image enhancement via Retinex-inline CycleGAN. IEEE Trans Multimedia. 2024;26:1213–28. doi:10.1109/TMM.2023.3278385.
33. Ang K, Lim WT, Loh YP, Ong S. Noise-aware zero-reference low-light image enhancement for object detection. In: 2022 International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS); 2022; Piscataway, NJ: IEEE. p. 1–4.
34. Al Sobbahi R, Tekli J. Low-light homomorphic filtering network for integrating image enhancement and classification. Signal Process: Image Commun. 2022;100:116527. doi:10.1016/j.image.2021.116527.
35. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR); 2016; Piscataway, NJ: IEEE. p. 770–8.
36. Lei S, Deng X, Shao H, Xia W. Improved res2net based on multi-scale attention mechanism for breast cancer image classification. In: Liu Q, Liu X, Cheng J, Shen T, Tian Y, editors. Proceedings of the 12th International Conference on Computer Engineering and Networks; 2022; Singapore: Springer Nature Singapore. p. 597–611.
37. Huang H, He R, Sun Z, Tan T. Wavelet-SRNet: a wavelet-based CNN for multi-scale face super resolution. In: 2017 IEEE International Conference on Computer Vision (ICCV); 2017; Piscataway, NJ: IEEE. p. 1698–706.
38. Kang E, Chang W, Yoo J, Ye JC. Deep convolutional framelet denosing for low-dose CT via wavelet residual network. IEEE Trans Med Imaging. 2018;37(6):1358–69. doi:10.1109/TMI.2018.2823756.
39. Guo C, Li C, Guo J, Loy CC, Hou J, Kwong S, et al. Zero-reference deep curve estimation for low-light image enhancement. In: 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR); 2020; Piscataway, NJ: IEEE. p. 1777–86.
40. Wang Z, Bovik A, Sheikh H, Simoncelli E. Image quality assessment: from error visibility to structural similarity. IEEE Trans Image Process. 2004;13(4):600–12. doi:10.1109/TIP.2003.819861.
41. Peng L, Zhu C, Bian L. U-shape transformer for underwater image enhancement. IEEE Trans Image Process. 2023;32:3066–79. doi:10.1109/TIP.2023.3276332.
42. Song W, Wang Y, Huang D, Tjondronegoro D. A rapid scene depth estimation model based on underwater light attenuation prior for underwater image restoration. In: Hong R, Cheng W-H, Yamasaki T, Wang M, Ngo C-W, editors. Advances in multimedia information processing-PCM 2018. Cham: Springer International Publishing; 2018. p. 678–88.
43. Iqbal K, Odetayo M, James A, Salam RA, Talib AZH. Enhancing the low quality images using unsupervised colour correction method. In: 2010 IEEE International Conference on Systems, Man and Cybernetics; 2010; Piscataway, NJ: IEEE. p. 1703–9.
44. Liu P, Zhang H, Zhang K, Lin L, Zuo W. Multi-level wavelet-CNN for image restoration. In: 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW); 2018; Piscataway, NJ: IEEE. p. 886–95.
45. Naik A, Swarnakar A, Mittal K. Shallow-UWnet: compressed model for underwater image enhancement; arXiv:2101.02073. 2021.
46. Horé A, Ziou D. Image quality metrics: PSNR vs. SSIM. In: 2010 20th International Conference on Pattern Recognition; 2010; Piscataway, NJ: IEEE. p. 2366–9.
47. Yang M, Sowmya A. An underwater color image quality evaluation metric. IEEE Trans Image Process. 2015;24(12):6062–71. doi:10.1109/TIP.2015.2491020.
48. Panetta K, Gao C, Agaian S. Human-visual-system-inspired underwater image quality measures. IEEE J Oceanic Eng. 2016;41(3):541–51. doi:10.1109/JOE.2015.2469915.

49. Fu Z, Fu X, Huang Y, Ding X. Twice mixing: a rank learning based quality assessment approach for underwater image enhancement. Signal Process: Image Commun. 2022;102:116622. doi:10.1016/j.image.2021.116622.
50. Mittal A, Moorthy AK, Bovik AC. No-reference image quality assessment in the spatial domain. IEEE Trans Image Process. 2012;21(12):4695–708. doi:10.1109/TIP.2012.2214050.
51. Mittal A, Soundararajan R, Bovik AC. Making a "Completely Blind" image quality analyzer. IEEE Signal Process Lett. 2013;20(3):209–12. doi:10.1109/LSP.2012.2227726.
52. Burger W, Burge MJ. Scale-invariant feature transform (SIFT). London: Springer London; 2016. p. 609–64.