



**ARTICLE**

# Underwater Diver Image Enhancement via Dual-Guided Filtering

Jingchun Zhou<sup>1,\*</sup>, Taian Shi<sup>1</sup>, Weishi Zhang<sup>1,\*</sup> and Weishen Chu<sup>2</sup>

<sup>1</sup>College of Information Science and Technology, Dalian Maritime University, Dalian, 116026, China

<sup>2</sup>Department of Computer and Information Science, University of Pennsylvania, Philadelphia, PA 19104, USA

\*Corresponding Authors: Weishi Zhang. Email: teesiv@dlnu.edu.cn; Jingchun Zhou. Email: zhoujingchun@dlnu.edu.cn

Received: 25 September 2021 Accepted: 09 November 2021

## ABSTRACT

The scattering and absorption of light propagating underwater cause the underwater images to present low contrast, color deviation, and loss of details, which in turn make human posture recognition challenging. To address these issues, this study introduced the dual-guided filtering technique and developed an underwater diver image improvement method. First, the color distortion of the underwater diver image was solved using white balance technology to obtain a color-corrected image. Second, dual-guided filtering was applied to the white balanced image to correct the distorted color and enhance its details. Four feature weight maps of the two images were then calculated, and two normalized weight maps were constructed for multi-scale fusion using normalization. To better preserve the obtained image details, the fusion image was histogram-stretched to obtain the final enhanced result. The experimental results validated that this method has improved the accuracy of underwater human posture recognition.

## KEYWORDS

Multi-scale fusion; image enhancement; guided filter; underwater diver images

## 1 Introduction

Computer vision technology is developing rapidly and is continuously being applied in new fields, including human-centered applications. For example, computer vision technology can be used for human posture recognition, human body position detection, etc. At present, there are many human body recognition technologies, most of which are based on deep learning technology [1–3]. Although human body recognition in normal lighting has been well developed, recognition results in underwater environments are not ideal. This is because the absorption of light underwater causes a loss of detailed information about the target. Therefore, accurate human posture recognition in underwater environments still has many challenges.

One way to address issues with underwater human posture recognition is enhancing the underwater image to obtain an image that is close to one taken in a terrestrial environment. This allows the classical human body recognition method to work properly. Researchers have proposed various methods for improving underwater images. According to whether the physical principle of



underwater light propagation is applied, the processing methods fall mainly into two categories: image enhancement and image restoration.

Image restoration mainly depends on the physical principle of underwater light. The underwater imaging model is built through physical principles. A simplified underwater image model is calculated:

$$I_{\lambda}(x, y) = J_{\lambda}(x, y) t_{\lambda}(x, y) + (1 - t_{\lambda}(x, y)) B_{\lambda}(x, y) \quad (1)$$

where  $(x, y)$  refers to a pixel of the image;  $\lambda \in \{R, G, B\}$ ,  $I_{\lambda}(x, y)$  and  $J_{\lambda}(x, y)$  refer to the original underwater image and the restored image, respectively;  $B_{\lambda}(x, y)$  refers to the background light; and  $t_{\lambda}(x, y)$  represents the transmission map. By estimating the model parameters and inverting the degradation process, the original illumination intensity of the scene is restored to get the enhanced image.

He et al. [4] developed a dehazing algorithm—DCP—specifically for outdoor image scenes. Their method assumes that at least one channel of a sunny image contains very low pixel values (close to 0). Chiang et al. [5] combined this DCP algorithm with a wavelength-dependent compensation algorithm to make the algorithm suitable for underwater environments and remove fog and correct color in underwater images.

Galdran et al. [6] established a model to restore underwater images since red light attenuates faster underwater. However, this model does not apply to all images, as some will have varying degrees of red channel saturation. Berman et al. [7] considered spectral profiles of different water types and simplified image restoration process to single-image defogging by estimating the attenuation ratio of blue-red and blue-green channels. Fattal [8] deduced a model to restore the scene transfer map based on the colored lines in a hazy scene and described a model dedicated to generating a complete and regularized transfer map based on estimations of noise and dispersion. Yang et al. [9] built an improved model based on dark channel reflection illumination decomposition and local backscattered illumination estimation and obtained images with sharp edge details and improved colors. Peng et al. [10] first used depth-related color changes to estimate the ambient light and scene transmission map utilizing the difference between the observed intensity and the ambient light. And they introduced adaptive color correction based on this estimation. This method, which achieved good results, was used for images degraded by light scattering and absorption. Zhou et al. [11] introduced secondary guided transfer map and proposed a restoration method for underwater image. An improved guided filter was used to optimize the transfer map. However, despite these various underwater image restoration methods, accurately estimating various parameters of the optical image model is still difficult because of the complexity of the underwater environment and illumination conditions. Therefore, such restoration methods need to be further optimized.

Numerous research studies have attempted to enhance the underwater image quality. The image fusion has been found to be an effective way for underwater image enhancement. In 2012, Ancuti et al. [12] built a fusion-based model for image enhancement using the Laplacian contrast weight, local contrast weight, and saliency weight. This multi-scale fusion pyramid strategy obtained good results; the images had improved sharpness, contrast, and color distribution. In 2017, Ancuti et al. [13] improved the proposed fusion method, proposing an approach that utilized red channel compensation before fusion. Additionally, a new weight calculation method was used to obtain dark area display enhancement, improved overall contrast, and edge sharpening.

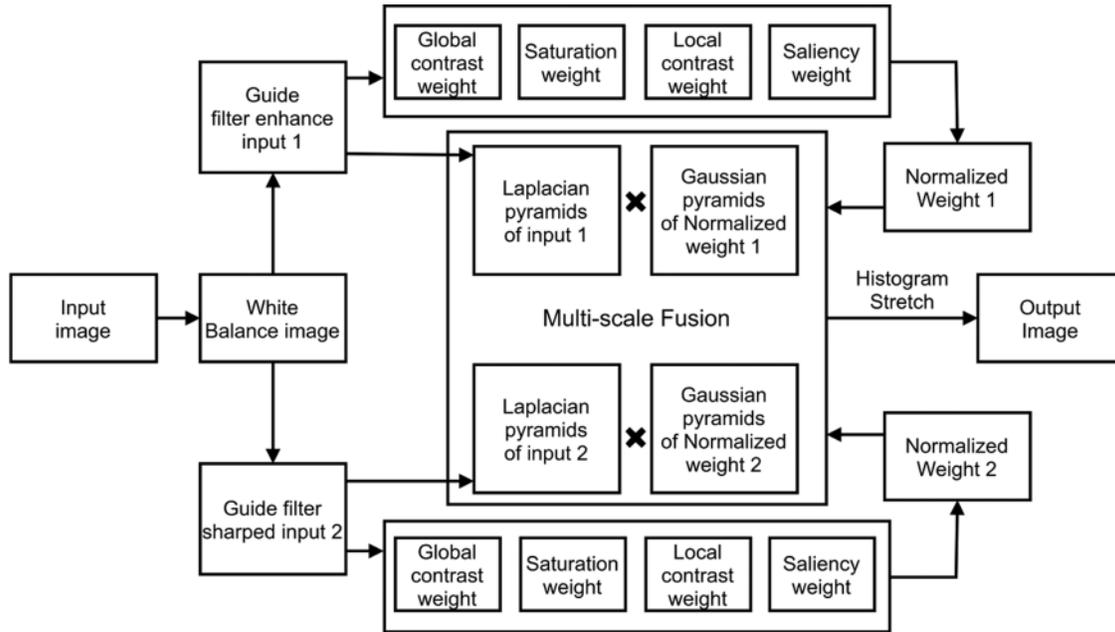
Zhou et al. [14] proposed the multi feature fusion method by fusing the image after color correction and guided filtering and then enhancing the edge to obtain balanced color and improved contrast. Fu et al. [15] established a two-stage underwater image improvement model for underwater images, seeking to correct colors and enhance contrast. Fu et al. [16] also applied the Retinex method to underwater image enhancement to obtain images with better visual effects and colors in another study. However, their method was relatively complex, requiring multiple iterations. Zhuang et al. [17] proposed a Bayesian Retinex algorithm with multi order gradient priors of reflectance and illumination. Most of the proposed enhancement methods enhanced the qualities of underwater images, however, they still did not fundamentally solve the degradation of underwater images.

Artificial intelligence (AI) is advancing rapidly and has been applied to a wide range of fields. Researchers have begun to incorporate AI-assisted analysis to underwater imaging improvement models as well. To ameliorate the color shift and increase the contrast of underwater images, Li et al. [18] presented a multi-media transmission guide that used an underwater image enhancement network embedded in color space. Guo et al. [19] built a multi-scale dense generation confrontation network to solve underwater image color distortion, underexposure, and blur. To do so, they proposed a residual multi-scale dense block in the generator. Li et al. [20] used a weakly supervised color transfer algorithm to ameliorate the color deviation. To address the inability of real-time and adaptive underwater enhancement methods, Chen et al. [21] built a recovery scheme with a generative confrontation network (GAN-RS). In this scheme, a multi-branch discriminator removes underwater noise while preserving the image content. Furthermore, a loss function based on an underwater index trains the evaluation branch of underwater noise suppression. Li et al. [22] proposed Zero-DCE, which formulated light enhancement as a task of image-specific curve estimation with a deep network. Compared to image enhancement and restoration methods, deep learning methods require many high-quality image datasets. However, since lossless underwater images are hard to obtain, deep learning methods usually cannot be fully trained and validated. Furthermore, Anwar et al. [23] comprehensively investigated underwater image enhancement methods with assistance of deep learning and found that methods with deep learning lag the most advanced traditional methods in most cases.

Therefore, this study proposed an underwater diver image improvement method with dual-guided filtering. This method ensured correction of the diver image's color and enhanced the image's details and contrast. Additionally, it was suitable for diver images in various underwater environments. The main innovations of this study are:

- (1) The method managed to improve the single underwater diver image quality without the need to consider complex image degradation models.
- (2) The images obtained through different technologies had different characteristics, and the advantages of these results were combined through fusion methods to obtain the final enhanced images.
- (3) Experiments were performed and compared with other advanced techniques. Both qualitative and quantitative methods were analyzed. The results showed that the approach introduced could accurately correct the color of underwater diver images, enhance detailed information, and improve the effectiveness of human posture recognition.

The flow of the method introduced in this paper is shown in Fig. 1. Each stage will be described in detail below.



**Figure 1:** Summary of the introduced method

## 2 Color Correction

The attenuation rate of light changes for different wavelengths. Generally, the shorter the wavelength, the stronger the penetration ability. Among red, green, and blue channels, the red light disappears first as it has the longest wavelength. However, in turbid or plankton-rich waters, blue light will be attenuated due to the absorption of organic matter, resulting in color distortion in the diver image. Therefore, the color deviation must be corrected before image enhancement. Because the average value of the red channel is minimal, applying conventional color correction methods like the gray world may lead to overcompensation for the red channel, resulting in over-saturation of that channel. Since the wavelength of green light is relatively short, the attenuation of green light in the underwater environment is small, and the preservation is relatively complete [13]. Thus, some green channels can be used to make up for the attenuation of the red and green channels. According to the gray world hypothesis theory [24], the mean values of the three channels should be the same. Therefore, the compensation of the red channel should be proportional to the difference between the mean values of the red and green channels.

First, the three channels were normalized according to the dynamic distribution and limited to [0,1]. The compensation coefficient of the red channel was as follows:

$$r_{rc}(x, y) = I_g(x, y) \times (1 - I_r(x, y)) \quad (2)$$

where  $I_r(x, y)$  and  $I_g(x, y)$  are the intensity values of the red and green channels, respectively. The final pixel value of the red channel was expressed in (3):

$$I_{rc}(x, y) = I_r(x, y) + r_{rc}(x, y) \times (\bar{I}_g - \bar{I}_r) \times \lambda \quad (3)$$

When the blue channel was in turbid or plankton-rich water, it was greatly attenuated because of the absorption of marine organisms. In these instances, the blue channel needed to be

compensated. For the blue channel, there were also the following:

$$r_{bc}(x, y) = I_g(x, y) \times (1 - I_b(x, y)) \quad (4)$$

$$I_{bc}(x, y) = I_b(x, y) + r_{bc}(x, y) \times (\bar{I}_g - \bar{I}_b) \times \lambda \quad (5)$$

where,  $I_r(x)$ ,  $I_g(x)$ , and  $I_b(x)$  represent the intensity values of the red, blue, and green channels, respectively, and  $\bar{I}_g$ ,  $\bar{I}_r$ , and  $\bar{I}_b$  refer to the average values of the intensity of each channel.  $\lambda$  is a dynamic parameter used to prevent overcompensation. It is usually set between 0 and 1 to adapt to different types of underwater environment, and in this case was set to 0.5.

After color attenuation compensation, the gray world white balance method was used for color correction. First, the gain coefficients of each channel were obtained as follows:

$$K_C = \frac{\text{Max}\left(\sum_{i=1}^m \sum_{j=1}^n R(i, j), \sum_{i=1}^m \sum_{j=1}^n G(i, j), \sum_{i=1}^m \sum_{j=1}^n B(i, j)\right)}{\sum_{i=1}^m \sum_{j=1}^n C(i, j)} \cdot \lambda \quad (6)$$

where  $C \in \{R, G, B\}$  represents different channels,  $K_C$  refers to the gain of the specific channel,  $\lambda$  represents a dynamic parameter ( $\lambda = 0.005$ ),  $m$  and  $n$  refer to the length the width of the image, respectively. The pixel value interval of each channel was obtained  $([L_S, H_S], S \in \{R, G, B\})$  according to the corresponding gain value.

Underwater diver images become blurred due to the influence of light scattering. To ensure a more natural visual effect in the output image, an adaptive contrast stretching was applied to stretch elements in range  $[L_s, H_s]$  between  $[0, 255]$ . The specific method was as follows:

$$T_C(i, j) = \frac{255 \cdot [I_C(i, j) - L_C]}{H_C - L_C} \quad (7)$$

where  $C$  represents different channels, and  $T_C(i, j)$  represents the value after each channel undergoes white balancing.

However, the white balance method can only recover color information; it cannot enhance the edge and detail information. Thus, a fusion-based method was applied to recover the lost details and edge information.

### 3 Detail Enhancement

Underwater diver images are influenced by light attenuation and thus have low contrast. The white balance method can largely alleviate the color deviation, but recovering the details of the degraded image is difficult under some complex conditions. Therefore, the image must be enhanced to improve the visual effect. He et al. [25] found that guided filtering has an outstanding edge-preserving characteristics and could improve computational efficiency.

To prevent the color information of the image from being affected while the contrast of the image was enhanced, the image was transformed from RGB channel to Lab channel when the guided filtering was applied, and only the L channel was processed [14]. Guided filtering can be calculated as (8):

$$G_L(x, y) = \text{guide}(I, r, \varepsilon) * T_{L,i}(x, y) \quad (8)$$

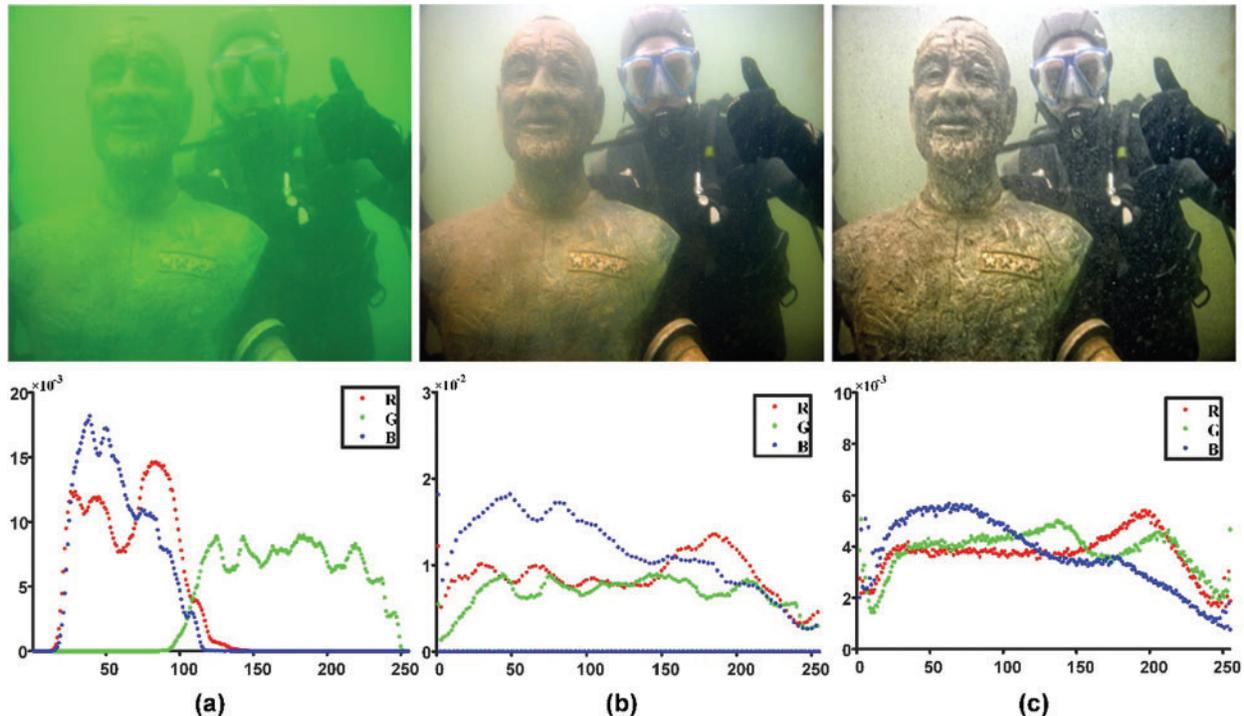
where  $G_L(x, y)$  refers to the filtered image,  $T_{L,i}(x, y)$  refers to the pixel value of input image  $L$  channel,  $I$  refers to guide image,  $r$  refers to the radius of the filtering window,  $\varepsilon$  refers to the regularization coefficient related to the smoothing effect, and  $*$  is the filtering operation.

The enhanced L channel was determined in (9):

$$L(x, y) = \lambda \times (I(x, y) - G_L(x, y)) + G_L(x, y) \quad (9)$$

where  $\lambda$  is a dynamic parameter. To make the image fully enhanced, set  $\lambda$  to 5.

As shown in Fig. 2, the image enhanced by guided filtering retained the effect of white balance color correction. It had better contrast and detail information, and the distribution of the RGB histogram was more balanced.



**Figure 2:** Results of our image improvement approach: (a) raw image, (b) white balance image, and (c) guided filter image. In the corresponding RGB histogram distribution plot below, the abscissa refers to the pixel values, and the ordinate refers to the normalized frequency

#### 4 Image Sharpening

Underwater diver images are usually fuzzy and lack detail. Guided filtering-enhanced images have greater brightness and contrast, but some details are still lost. Therefore, this study introduced a sharpened version underwater image as another input to compensate for the loss of detail in the underwater diver image and reduce the degradation caused by light scattering. In the present study, the unsharpened mask method [13,26] was utilized to sharpen the image.

In the traditional sharpening method, a Gaussian kernel is first used to smooth the original image, which is then subtracted from the original image. Finally, the difference is added to the original image. In this study, however, a guided filter was used instead of a Gaussian kernel to smooth the image, since it had good edge-preserving characteristics [25]. Inspired by MSR [27], a guided filter with low, medium, and high scales was selected to obtain three different scale versions of the image, and the average value was calculated to obtain a more accurate smooth image.

In this study, multiple smooth images were obtained by selecting guided filtering with different regularization coefficients, and a weight value was assigned to each smooth image. These images were combined into a final smooth image, as shown below:

$$T(x, y) = \sum_{n=1}^N \omega_n [\text{guide}(q, r, \varepsilon_n) * I(x, y)] \quad (10)$$

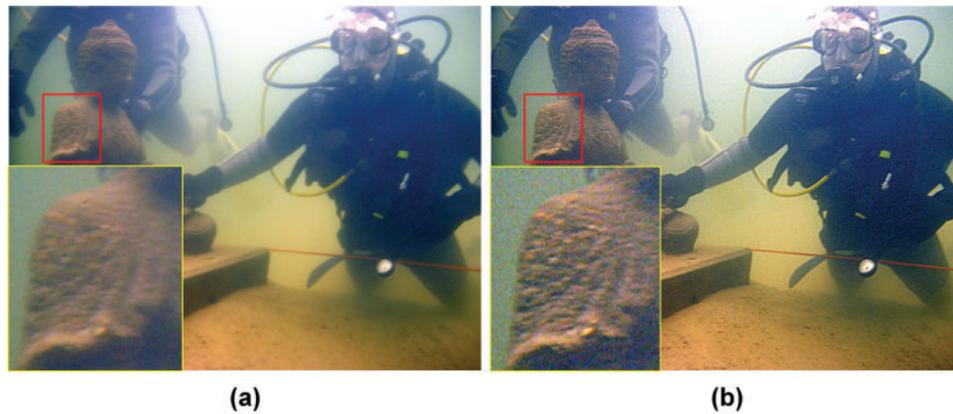
where  $I(x, y)$  represents the input image,  $q$  represents to the guide image (the original image was used as the guide image),  $N$  represents the number of images at different scales,  $\omega$  refers to the weight of each scale (set to  $\frac{1}{N}$ ), and  $T(x, y)$  represents the smoothed image.

To obtain the edge detail image, the smoothed image was subtracted from the input image. Finally, the edge detail image was linearly fused with the input image to obtain the sharpened image. The final sharpened result was formulated as follows:

$$S(x, y) = s \times [I(x, y) - T(x, y)] + I(x, y) \quad (11)$$

where  $s$  is the control factor of sharpening degree. After many experiments, the value was 1.5.

As shown in Fig. 3, the approach introduced obtained more obvious texture information than the raw image.



**Figure 3:** Image sharpened result. From left to right: (a) raw image, (b) sharpened image

## 5 Computation Fusion Weight

After image sharpening and detail enhancement, the white balance image was used as the two fused input images. The image fusion effect was enhanced by extracting the special weight map of the image. Considering the low contrast and fuzzy details of underwater images, four weights were selected for the input image: global contrast weight, local contrast weight, saturation weight, and significance weight. The weight map was used to increase the proportion of pixels with higher weight values in the resulting image [13].

### 5.1 Global Contrast Weight

Laplace filter could enhance the image edge and texture to a certain extent. In this study, we denoted  $W_{La,k}(x, y)$  as the global contrast weight, The Laplace filter was applied to the L channel in a lab space to estimate the global contrast weight, which highlighted edge details, but the

gradient and flat areas could not be distinguished. Therefore, other weights had to be introduced to supplement.

### 5.2 Saturation Weight

Saturation weight enables the fusion process to use the highly saturated region to adapt to the color information, but the saturation weight will reduce the contrast to some extent. Therefore, global contrast and local contrast weight must be introduced to solve the low contrast.

The saturation weight was first used to convert the image to Lab space, then the standard deviation between the RGB and L channels was calculated. The specific calculation method was applied:

$$W_{Sat,k}(x,y) = \sqrt{\frac{1}{3} \left\{ \sum_C [C_k(x,y) - L_k(x,y)]^2 \right\}} \quad C \in \{R, G, B\} \quad (12)$$

where  $W_{Sat,k}(x,y)$  represents the saturation weight;  $R_k(x,y)$ ,  $G_k(x,y)$ , and  $B_k(x,y)$  refer to the red, green, and blue color channels of the input image  $k$  respectively; and  $L_k(x,y)$  is the L channel of the input image lab space.

### 5.3 Saliency Weight

Saliency weight was used to highlight some objects in the underwater environment that are more easily perceived by human vision. Fu et al. [28] developed a collaborative saliency detection algorithm with clustering to obtain a more accurate saliency map by integrating contrast and spatial clues.

Contrast cues mainly represent the uniqueness of visual features of a single image, which are calculated as follows:

$$\omega^C(k) = \sum_{i=i, i \neq k}^K \left( \frac{n^i}{N} \|u^k - u^i\|_2 \right) \quad (13)$$

where  $n^i$  represents the pixel number of  $C^i$ ,  $N$  represents the total pixel number, and  $u^k$  refers to the cluster center associated with the cluster  $C^k$ . L2 norm is used for distance calculation on the feature space.

Spatial cues are calculated assuming the central area of the image can attract people's attention more than other areas:

$$\omega^S(k) = \frac{1}{n^k} \sum_{j=i}^M \sum_{i=1}^{N_j} \left[ G \left( \|z_i^j - c^j\|^2, 0, \sigma^2 \right) \cdot \delta \left[ b(p_i^j) - C^k \right] \right] \quad (14)$$

where  $\delta(\cdot)$  denotes the Kronecker delta function,  $c^j$  is image center,  $G(\cdot)$  is used to calculate the Euclidean distance between pixels  $z_i^j$  and  $c^j$ ,  $\sigma^2$  is the normalized radius of images, and  $n^k$  is the pixel number of  $C^k$ . The final saliency map was obtained by combining these two clues.

### 5.4 Local Contrast Weight

This study also introduced local contrast weight [13] to address the local detail loss issue. The local contrast weight highlighted areas in the input image where there were more local pixel value changes and increased the transition between the lighter and darker areas. Letting  $W_{Lc,k}(x,y)$  be

the local contrast weight, it can be obtained by calculating the standard deviation of the mean value of each pixel and its surrounding pixels, as follows:

$$W_{Lc,k}(x,y) = \|L_k(x,y) - L_{\omega_{hc,k}}(x,y)\| \quad (15)$$

where  $L_k(x,y)$  is the brightness value of the input image, and  $L_{\omega_{hc,k}}(x,y)$  is calculated from the  $5 \times 5$  separable binomial kernel  $1/16[1,4,6,4,1]$ .

### 5.5 Multifeatured Fusion

To take full advantage of the above four weight maps, a multi-scale fusion method was introduced; the saturation weight map could balance the color difference, the global contrast weight map could improve the overall contrast, and the saliency weight graph could highlight prominent objects that lost saliency in the underwater images. Finally, the local contrast weight was employed to polish the image details.

The four weight maps were merged into a weight map  $W$  with normalization, and the fused image was calculated:

$$W_K(x,y) = \sum W_{i,k}(x,y) \quad i \in \{Sat, La, Sa, Lc\} \quad (16)$$

where  $W_k(x,y)$  refers to the  $k$ th normalized weight map, and  $\{Sat, La, Sa, Lc\}$  represents the saliency, global contrast, saturation, and the local contrast weight respectively.

To avoid the artifacts that may appear in simple linear weighted fusion, the Laplacian pyramid fusion method was introduced, and the normalized weight map was deconstructed into multi-scale Laplacian pyramids. Additionally, the Gaussian pyramid was applied to decompose the input image. Finally, the obtained Gaussian pyramid and Laplacian pyramid were fused on multiple scales.

First, the normalized weight map was obtained as follows:

$$\overline{W}_k(x,y) = \frac{W_k(x,y) + \delta}{\sum_{k=1}^2 (W_k(x,y) + \delta)} \quad (17)$$

where  $\overline{W}_k(x,y)$  refers to the normalized weight map of the  $k$ th input map, and  $\delta$  is a constant set to 0.001 to ensure that each weight map contributed to the results.

Laplacian pyramid decomposition and Gaussian pyramid decomposition were performed on the fused input image and its normalized weight map, respectively, and the fusion was performed on multiple scales, as follows:

$$R_{l,k}(x,y) = \sum_k G_l[\overline{W}_k(x,y)] L_l[I_k(x,y)] \quad (18)$$

where  $R_{l,k}(x,y)$  represents the  $l$ -th layer of the output image pyramid for the  $k$ -th input image,  $G_l[\overline{W}_k(x,y)]$  refers to the normalized weight map after the Gaussian filtering, and  $L_l[I_k(x,y)]$  refers to the input Laplacian filtering pyramid of the image.

To obtain the final output image, the output pyramid was sampled, as expressed in (19):

$$E_{res}(x,y) = \sum_l U[R_{l,k}(x,y)] \quad (19)$$

where  $E_{res}(x,y)$  is the output image, and  $U[R_{l,k}(x,y)]$  refers to the sampled output image pyramid.

### 5.6 Histogram Stretching

The color and contrast of the fused image were greatly improved after these steps, but some areas were still too dark or too bright. To get a high-quality image, these parts had to be removed to generate more reasonable image brightness. The over-bright and over-dark regions were removed by stretching the pixels whose size was between the pixel values at quantile 0.1% and 99.9% in each channel to [0,255].

$$R_E(x,y) = \frac{255 \cdot [E_{res}(x,y) - M(T \cdot C)]}{M(C - T \cdot C) - M(T \cdot C)} \quad (20)$$

where  $M$  represents an array of pixels in a certain channel sorted in ascending order,  $C$  refers to the number of image pixels,  $T$  refers to a threshold ( $T = 0.01$ ), and  $R_E(x,y)$  refers to the stretched image.

## 6 Experiment Results and Discussion

This section compares the method introduced in this paper with the existing restoration approaches for underwater images. Each method was analyzed qualitatively and quantitatively. The selected comparison methods were UDCP [29], RBE [16], TSP [15], IBLA [10], ULAP [30], GDCP [31], RGHS [26], UWCNN [32], and WaterNet [33]. All experiments were run on Win10 PC, AMD ryzen 5 4600u, and MATLAB r2020b.

A total of 890 images from the UIEB [33] dataset was utilized to verify the introduced method. The UIEB dataset is widely used. The dataset was divided into two categories: Data A, about underwater divers, and Data B, about images of other underwater scenes.

First, the experimental results were analyzed qualitatively, and some objective metrics were then used for quantitative analysis. Next, the average value was calculated according to the quantitative metric results of all images in the dataset, and a box plot was created to demonstrate the superiority of the approach introduced in this paper.

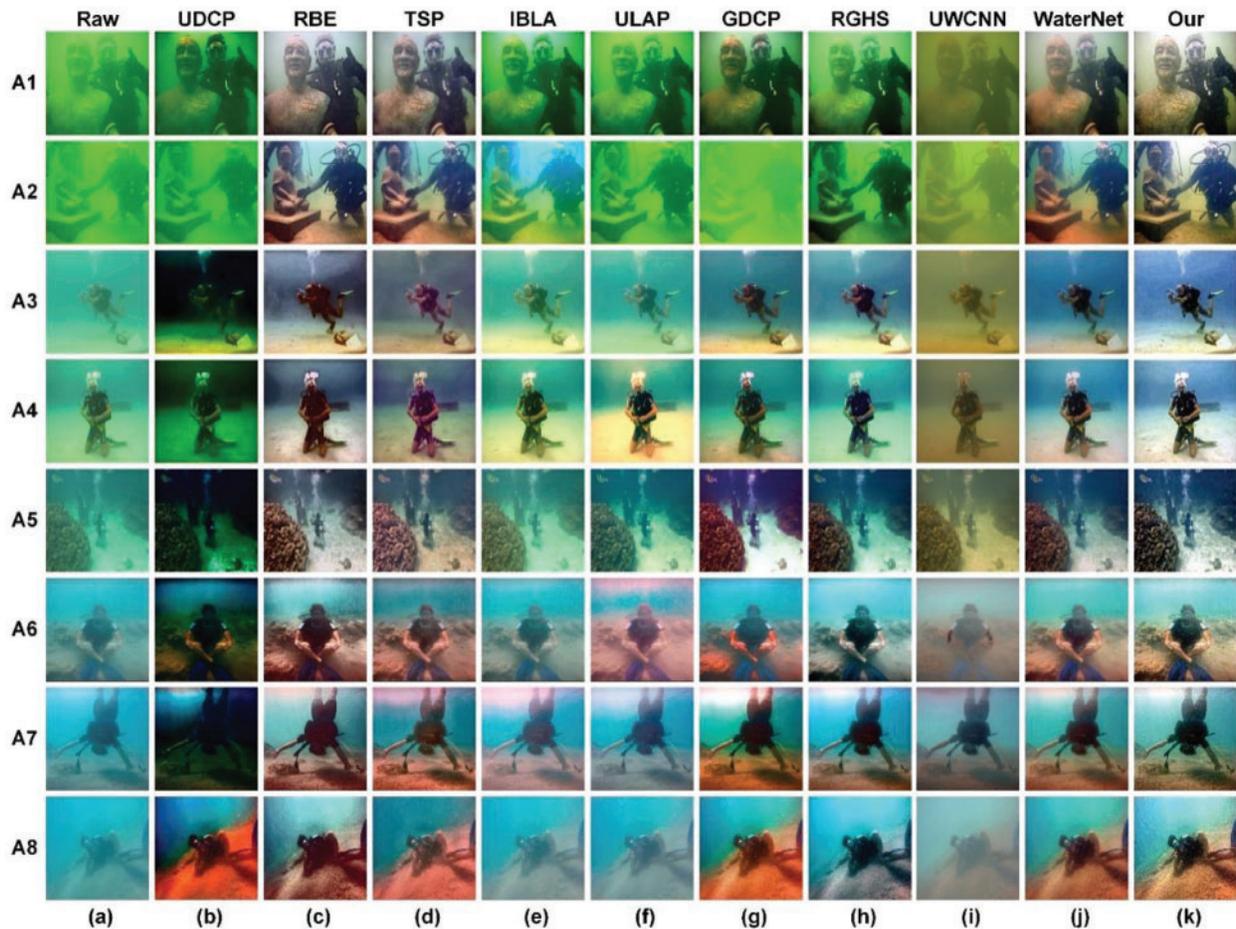
### 6.1 Qualitative Analysis

For qualitative analysis, eight images were selected from the two categories in the dataset.

Figs. 4 and 5 show the resulting images obtained by different imaging improvement methods.

The UDCP approach applies DCP to estimate the transmission map according to the blue-green channel, but it does not take the attenuation of the red channel into consideration, which is prone to color deviation. Additionally, the overall brightness of the image is low. For example, a green color cast appeared in images A3, A4, and B1, while a red color cast appeared in images A8, B4, and B7. RBE first performs color correction and then solves the underexposure and blur based on Retinex. However, because of the oversaturation of the red channel, the images generally show red, such as images B1, and B2. TSP adopts a two-step approach to improve the image. It mainly solves color deviation and low contrast through pixel value transformation. However, due to its inability to select an accurate target image when adjusting the contrast, this method does not improve the contrast much. This results in red channel oversaturation, such as that seen in images A6, A7, A8, B1, B2, and B3. IBLA restores the image by estimating the transmission image and background light. However, these estimates are not accurate, resulting in color correction and contrast enhancement that are not obvious. As a result, the results have different degrees of color

deviation, and the contrast is not largely enhanced. ULAP estimates the background light based on the maximum intensity difference between the red and blue-green channels. However, because of the limited improvement of the red and blue channels, the color degradation of its output image is not completely removed. For example, A1, A2, and B1 showed green color cast; A7, and A8 had a blue color shift; A6, B2, and B8 images showed a red cast. Additionally, the contrast of these images was low. This was because ULAP has difficulty accurately estimating the background light in some images with serious green color cast, resulting in a poor restoration effect.

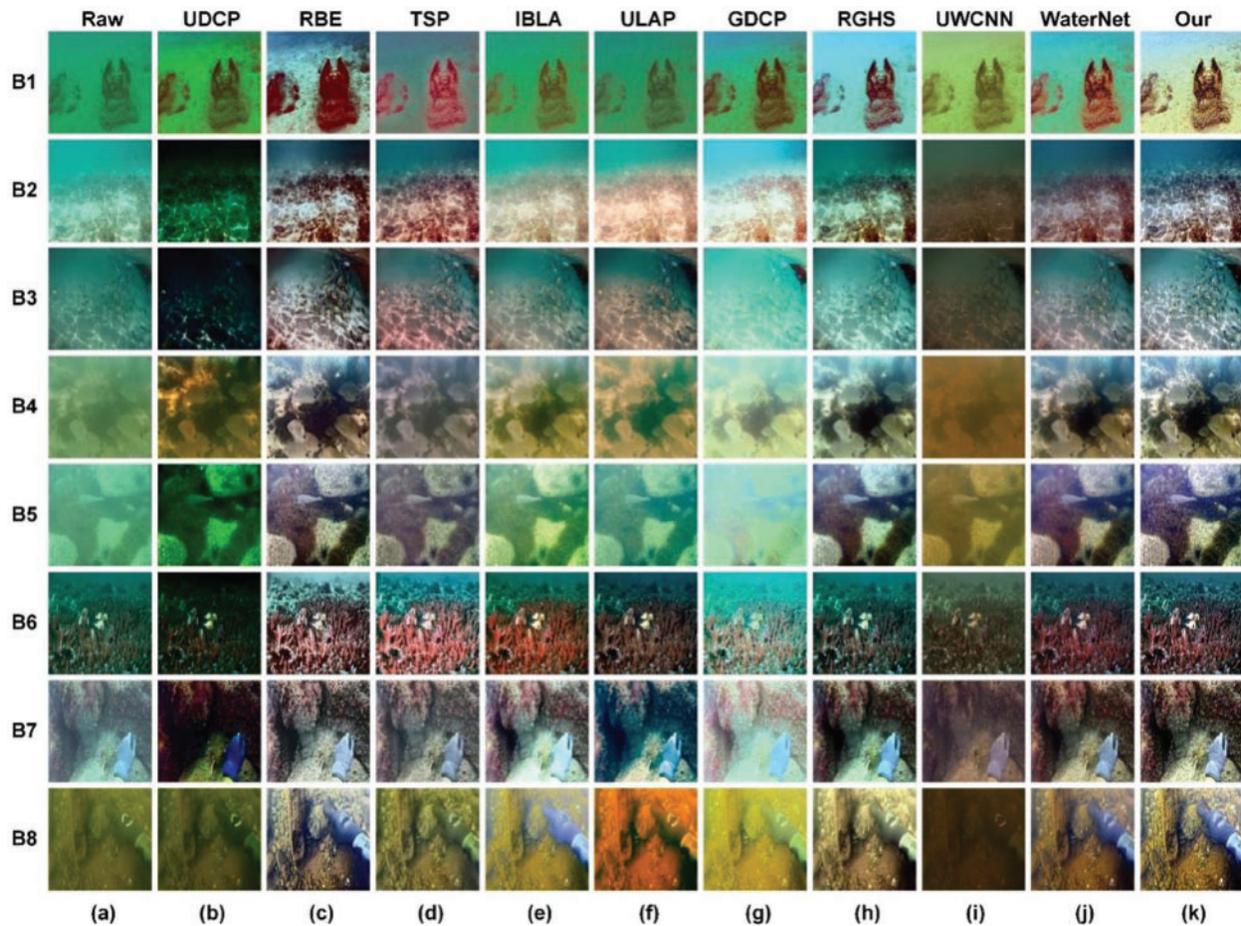


**Figure 4:** The qualitative evaluation result of eight selected images from Data A: (a) Raw images, (b) UDCP, (c) RBE, (d) TSP, (e) IBLA, (f) ULAP, (g) GDCP, (h) RGHS, (i) UWCNN, (j) WaterNet, (k) Our method

Like UDCP, GDCP estimates the ambient light through the color change related to the water depth and then obtains the scene transmission map. However, this method is based on IFM and does not apply to all underwater environments. Furthermore, although it can improve contrast and brightness to a certain extent, the method can introduce color deviations when enhancing images, such as those found in images A2, B2, and B5. RGHS first equalizes the blue-green channel and then redistributes the histogram using the dynamic parameters related to the original image and underwater wavelength attenuation. Although the contrast of the image can be enhanced, the

detailed information is not recovered well as the approach ignores the relationship between image degradation and scene depth, resulting in serious color deviation, such as that seen in images A1, A2, and B1. UWCNN applies the deep learning method to underwater diver image enhancement, but the combination of UWCNN with the physical model did not obtain ideal results. Severe color deviation remained in the obtained images, and the low contrast and low brightness were not solved. WaterNet constructs a special dataset to train the CNN network. Still, due to the lack of real images in the underwater environment, backscattering obtained by WaterNet is difficult to eliminate. Although the color deviation of the image was resolved, the details of some images were not restored, such as images A1, A2, B7, etc.

In contrast, the present method did not introduce color cast, over-enhancement, or under-enhancement. As a result, it obtained satisfactory images with color, contrast, brightness, and clearer texture details.



**Figure 5:** The qualitative evaluation result of eight selected images from Data B: (a) Raw images, (b) UDCP, (c) RBE, (d) TSP, (e) IBLA, (f) ULAP, (g) GDCP, (h) RGHS, (i) UWCNN, (j) WaterNet, (k) Our method

### 6.2 Quantitative Analysis

Evaluation of underwater images often focuses on whether the color, contrast, and overall visual quality of the image have been improved. Such evaluation requires objective methods. This section presents this study’s approach and other methods using four objective metrics: average gradient (AG), underwater image quality measurement (UIQM), underwater image quality measurement (UCIQE), and patch-based contrast quality (PCQI). UIQM and UCIQE indicators are used particularly for underwater image evaluation, mainly to evaluate the image’s color richness, clarity, and contrast. PCQI represents the contrast of the image, while AG mainly indicates the sharpness of the image. Higher PCQI and AG values indicate better contrast and more detailed texture information of an image. The objective evaluation results of the images shown in Figs. 4 and 5 were presented using a broken line diagram, as shown in Figs. 6 and 7. The abscissa represents the image sequence number, and the ordinate represents the result value from different methods.

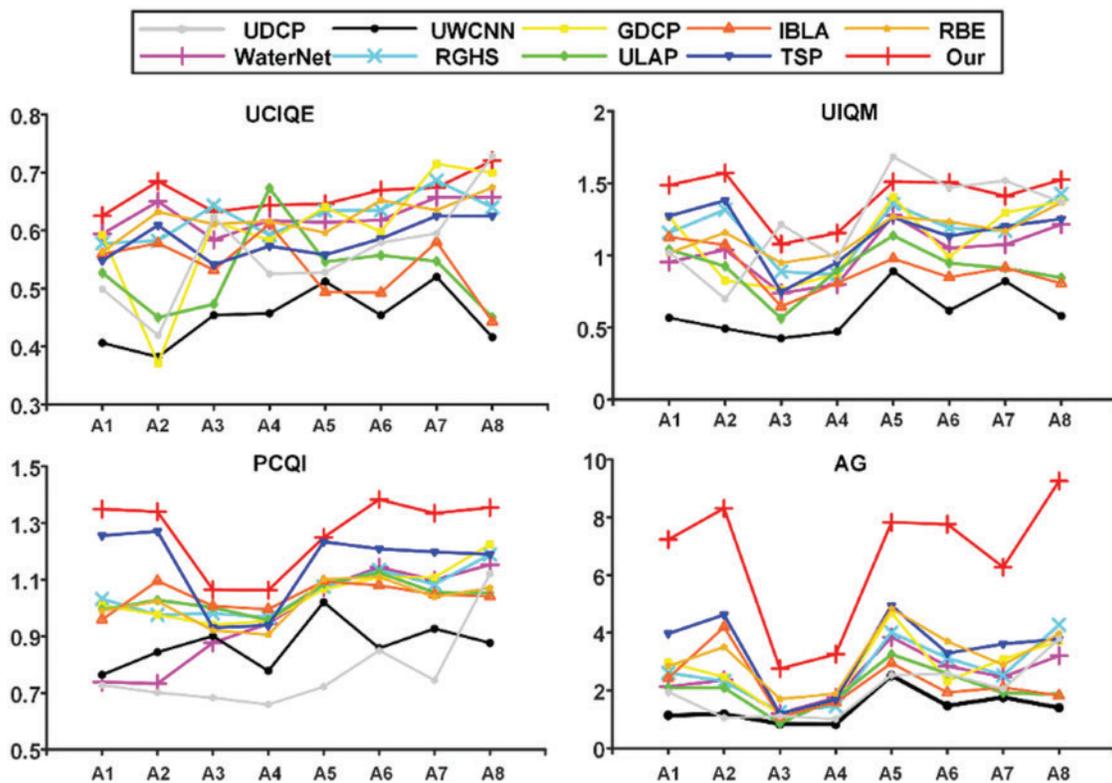


Figure 6: Results of objective evaluation of Fig. 4

As seen in the broken line diagram, the UCIQE, UIQM, PCQI, and AG values obtained by the present approach were higher than those acquired by other approaches in most cases. Among the UCIQE values, A3 was lower than RGHS, A4 was lower than ULAP, and A7 was lower than GDCP because UCIQE is a linear addition of color concentration, saturation, and contrast. Thus, an excessively high factor will lead to a high UCIQE value. As seen in Fig. 4, the RGHS, ULAP, and GDCP methods resulted in supersaturated images with color deviation. Since UCIQE does not fully consider color deviation and artifact [33], the UCIQE value was high. For

A3, A5, A7, B2, and B3, the UIQM value of the present approach was slightly lower than that of the UDCP. As shown in Figs. 4 and 5, the UDCP approach resulted in low brightness and contrast, as well as serious color cast. For B1 images, the PCQI value of the present method was slightly lower than TSP and UDCP, while other images had higher PCQI values than other methods. Additionally, the AG values of all experimental images of the present approach were higher than other approaches. Thus, the contrast and detail enhancement method found in this study effectively enhanced valuable information and improved contrast.

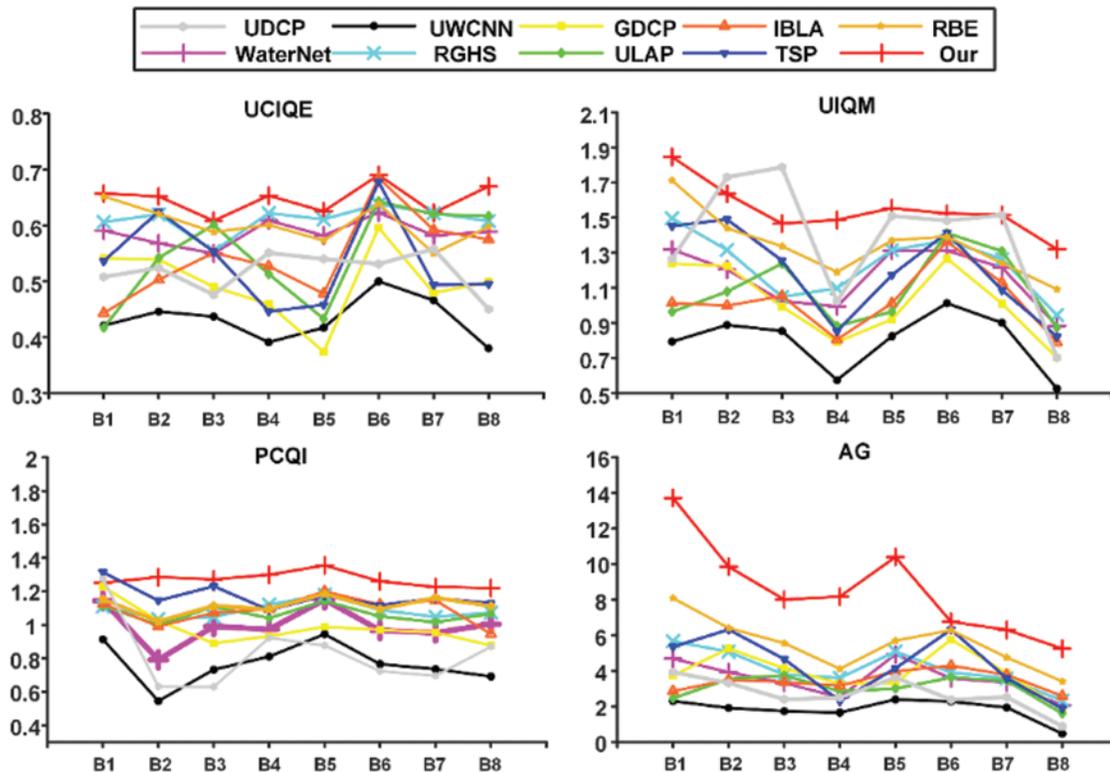
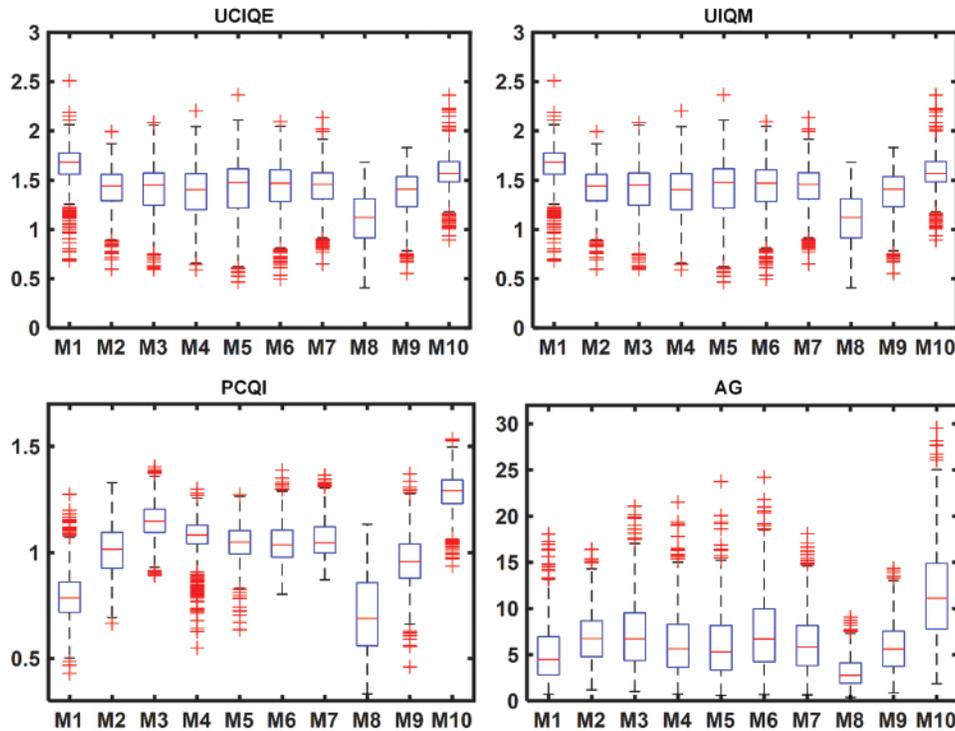


Figure 7: Results of objective evaluation of Fig. 5

To objectively evaluate the effect of the method further, the average value was calculated, and a box plot, as shown in Fig. 8, was drawn according to each objective metric for all underwater images in the UIEB dataset to represent the discrete distribution of objective metrics in different methods. The average value is displayed in Table 1, in which the best measurement results are bold and underlined. Compared with most other advanced methods, our method had the first UCIQE, PCQI, and AG indicators and the second UIQM. However, compared with the first UDCP method, the difference was only 0.07.

As shown in the box plot in Fig. 7, On each box, the central mark represents the median, and the bottom and top edges of the box illustrate the upper and lower quartile, respectively. The whiskers above and below the box show the minimum and maximum value, respectively. red marks represent outliers. It can be seen from the box plot that for UCIQE and UIQM, the median and maximum values of our method are only less than those of UDCP method, but higher than those

of other methods, and the results are relatively stable. On the other hand, for PCQI and AG, our approach has better performance than other approaches.



**Figure 8:** Box plot of four objective metrics in UIEB dataset. M1: UDCP, M2: RBE, M3: TSP, M4: IBLA, M5: ULAP, M6: GDCP, M7: RGHS, M8: UWCNN, M9: WaterNet, M10: Our

**Table 1:** Evaluation results of average indexes of all images in UIEB dataset

Method	UDCP	RBE	TSP	IBLA	ULAP	GDCP	RGHS	UWCNN	WaterNet	Our
UCIQE	0.582	0.602	0.577	0.599	0.605	0.598	0.628	0.477	0.584	<b>0.641</b>
UIQM	<b>1.650</b>	1.407	1.397	1.373	1.410	1.427	1.429	1.105	1.365	1.572
PCQI	0.795	1.009	1.148	1.074	1.046	1.045	1.064	0.708	0.959	<b>1.280</b>
AG	5.184	6.870	7.296	6.237	6.037	7.463	6.270	3.084	5.833	<b>11.420</b>

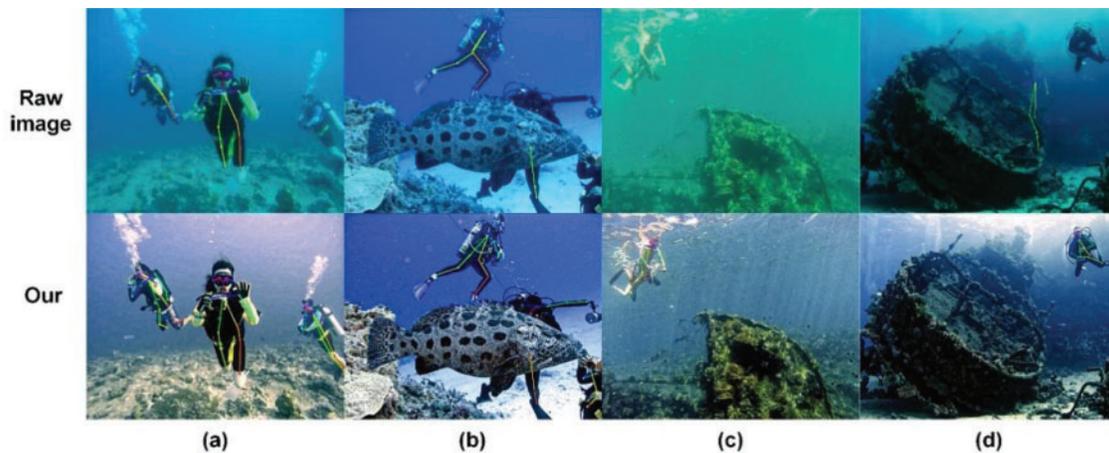
Thus, comparison showed that the four objective metrics obtained by the present method achieved good results. Our method first compensated for the red channel, then performed gray world white balance to alleviate the color distortion of the image, and then obtained the detail-enhanced and sharpened images based on guided filtering. Finally, the obtained images were multi-scale fused and histogram-stretched to obtain the enhanced result. The brightness, saturation, and contrast of the obtained images were balanced, with no obvious color cast.

Thus, our method could obtain images with significantly improved clarity, contrast, chroma, and brightness. Furthermore, all objective evaluation indices were high and had certain robustness. Therefore, our approach is suitable for underwater diver image enhancement in most environments.

### 6.3 Application Test

Human posture recognition is the basis of most human-centered computer vision applications. This section will illuminate the influence of the enhanced underwater diver image on human posture recognition. A method proposed by Cao et al. [1] was used to effectively detect the two-dimensional posture of multiple people in the image. This method is based on deep learning and uses Partial Affinity Field (PAFs) learning to associate body parts with individuals in the image, to achieve real-time performance, maintain high accuracy and can accurately recognize human posture in an environment with normal lighting.

As shown in Figs. 9a and 9b, only some divers could be identified in the raw images, and the legs and arms of some could not be accurately identified. In contrast, all divers could be completely identified in the enhanced image, and the recognition results were more accurate. As seen Figs. 9c and 9d, the divers could not be identified from the raw image, and other targets were identified as divers in the image (d). Because the image was affected by the underwater environment, it had color deviation, and the edge information was lost. In contrast, the enhanced image was closer to the daily environment. Thus, the results showed that the present approach could remarkably enhance the posture recognition of underwater divers.



**Figure 9:** Experimental results of underwater human posture recognition: (a) Scene 1, (b) Scene 2, (c) Scene 3, (d) Scene 4.

### 7 Conclusion

This paper presents an underwater diver image enhancement approach based on double-guided filtering. First, color correction was performed to restore the color cast. Second, a guided filter was used to enhance details and sharpen the image, and multi-scale fusion was performed to enhance the edge and contrast to avoid artifact. Finally, the areas where the image was too bright or dark were removed by histogram stretching. Our method significantly alleviated the color distortion of the underwater diver image, enhanced the contrast and detail information, obtained an image like that in a normal lighting environment, highlighted the edge information of the underwater diver, made the diver's image more identifiable, and improved underwater human posture recognition. Many experiments showed that the method had certain advantages over others. By comparing with the most advanced approaches, our approach has relatively large

advantages in various objective indicators. It is robust for underwater diver image enhancement and can effectively improve underwater human posture recognition.

**Acknowledgement:** Thanks to the data set provided by the joint laboratory of the Dalian University of Technology and Zhangzidao Group. We are also extremely grateful to the anonymous reviewers for their critical comments on the manuscript.

**Funding Statement:** National Natural Science Foundation of China (No. 61702074); the Liaoning Provincial Natural Science Foundation of China (No. 20170520196); the Fundamental Research Funds for the Central Universities (Nos. 3132019205 and 3132019354).

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

## References

1. Cao, Z., Simon, T., Wei, S. E., Sheikh, Y. (2017). Realtime multi-person 2D pose estimation using part affinity fields. arXiv: 1611.08050. <http://arxiv.org/abs/1611.08050>.
2. Chou, C. J., Chien, J. T., Chen, H. T. (2018). Self adversarial training for human pose estimation. *2018 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)*, Honolulu, HI, USA, pp. 17–30. DOI 10.23919/APSIPA.2018.8659538.
3. Chu, X., Yang, W., Ouyang, W., Ma, C., Yuille, A. L. et al. (2017). Multi-context attention for human pose estimation. arXiv: 1702.07432. <http://arxiv.org/abs/1702.07432>.
4. He, K., Sun, J., Tang, X. (2011). Single image haze removal using dark channel prior. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(12), 2341–2353. DOI 10.1109/TPAMI.2010.168.
5. Chiang, J. Y., Chen, Y. C. (2012). Underwater image enhancement by wavelength compensation and dehazing. *IEEE Transactions on Image Processing: A Publication of the IEEE Signal Processing Society*, 21(4), 1756–1769. DOI 10.1109/TIP.2011.2179666.
6. Galdran, A., Pardo, D., Picón, A., Alvarez-Gila, A. (2015). Automatic red-channel underwater image restoration. *Journal of Visual Communication and Image Representation*, 26(4), 132–145. DOI 10.1016/j.jvcir.2014.11.006.
7. Berman, D., Levy, D., Avidan, S., Treibitz, T. (2020). Underwater single image color restoration using haze-lines and a new quantitative dataset. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 43(8), 2822–2837. DOI 10.1109/TPAMI.2020.2977624.
8. Fattal, R. (2014). Dehazing using color-lines. *ACM Transactions on Graphics*, 34(1), 1–14. DOI 10.1145/2651362.
9. Yang, M., Sowmya, A., Wei, Z., Zheng, B. (2020). Offshore underwater image restoration using reflection-decomposition-based transmission map estimation. *IEEE Journal of Oceanic Engineering*, 45(2), 521–533. DOI 10.1109/JOE.2018.2886093.
10. Peng, Y. T., Cosman, P. C. (2017). Underwater image restoration based on image blurriness and light absorption. *IEEE Transactions on Image Processing: A Publication of the IEEE Signal Processing Society*, 26(4), 1579–1594. DOI 10.1109/TIP.2017.2663846.
11. Zhou, J., Liu, Z., Zhang, W., Zhang, D., Zhang, W. (2021). Underwater image restoration based on secondary guided transmission map. *Multimedia Tools and Applications*, 80(5), 7771–7788. DOI 10.1007/s11042-020-10049-7.
12. Ancuti, C., Ancuti, C. O., Haber, T., Bekaert, P. 2012. Enhancing underwater images and videos by fusion. *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 81–88. Providence, RI, USA. DOI 10.1109/CVPR.2012.6247661.
13. Ancuti, C. O., Ancuti, C., de Vleeschouwer, C., Bekaert, P. (2018). Color balance and fusion for underwater image enhancement. *IEEE Transactions on Image Processing: A Publication of the IEEE Signal Processing Society*, 27(1), 379–393. DOI 10.1109/TIP.2017.2759252.

14. Zhou, J., Zhang, D., Zhang, W. (2021). A multifeature fusion method for the color distortion and low contrast of underwater images. *Multimedia Tools and Applications*, 80(12), 17515–17541. DOI 10.1007/s11042-020-10273-1.
15. Fu, X., Fan, Z., Ling, M., Huang, Y., Ding, X. (2017). Two-step approach for single underwater image enhancement. *International Symposium on Intelligent Signal Processing and Communication Systems*, pp. 789–794, Xiamen, China. DOI 10.1109/ISPACS.2017.8266583.
16. Fu, X., Zhuang, P., Huang, Y., Liao, Y., Zhang, X. P. et al. (2014). A retinex-based enhancing approach for single underwater image. *IEEE International Conference on Image Processing*, pp. 4572–4576. Paris, France. DOI 10.1109/ICIP.2014.7025927.
17. Zhuang, P., Li, C., Wu, J. (2021). Bayesian retinex underwater image enhancement. *Engineering Applications of Artificial Intelligence*, 101(1), 104171. DOI 10.1016/j.engappai.2021.104171.
18. Li, C., Anwar, S., Hou, J., Cong, R., Guo, C. et al. (2021). Underwater image enhancement via medium transmission-guided multi-color space embedding. *IEEE Transactions on Image Processing: A Publication of the IEEE Signal Processing Society*, 30, 4985–5000. DOI 10.1109/TIP.2021.3076367.
19. Guo, Y., Li, H., Zhuang, P. (2020). Underwater image enhancement using a multiscale dense generative adversarial network. *IEEE Journal of Oceanic Engineering*, 45(3), 862–870. DOI 10.1109/JOE.2019.2911447.
20. Li, C., Guo, J., Guo, C. (2018). Emerging from water: Underwater image color correction based on weakly supervised color transfer. *IEEE Signal Processing Letters*, 25(3), 323–327. DOI 10.1109/LSP.2018.2792050.
21. Chen, X., Yu, J., Kong, S., Wu, Z., Fang, X. et al. (2019). Towards real-time advancement of underwater visual quality with GAN. *IEEE Transactions on Industrial Electronics*, 66(12), 9350–9359. DOI 10.1109/TIE.2019.2893840.
22. Li, C., Guo, C., Chen, C. L. (2021). Learning to enhance low-light image via zero-reference deep curve estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence (Early Access)*. DOI 10.1109/TPAMI.2021.3063604.
23. Anwar, S., Li, C. (2020). Diving deeper into underwater image enhancement: A survey. *Signal Processing: Image Communication*, 89(6), 115978. DOI 10.1016/j.image.2020.115978.
24. Buchsbaum, G. (1980). A spatial processor model for object colour perception. *Journal of the Franklin Institute*, 310(1), 1–26. DOI 10.1016/0016-0032(80)90058-7.
25. He, K., Sun, J., Tang, X. (2013). Guided image filtering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(6), 1397–1409. DOI 10.1109/TPAMI.2012.213.
26. Huang, D., Wang, Y., Song, W., Sequeira, J., Mavromatis, S. (2018). Shallow-water image enhancement using relative global histogram stretching based on adaptive parameter acquisition. In: Schoeffmann, K., Chalidabhongse, T. H., Ngo, C. W., Aramvith, S., O'Connor, N. E. et al. (Eds.), *MultiMedia modeling*, vol. 10704, pp. 453–465. Springer International Publishing, Springer, Cham Publishing. DOI 10.1007/978-3-319-73603-7\_37.
27. Jobson, D. J., Rahman, Z., Woodell, G. A. (1997). A multiscale retinex for bridging the gap between color images and the human observation of scenes. *IEEE Transactions on Image Processing: A Publication of the IEEE Signal Processing Society*, 6(7), 965–976. DOI 10.1109/83.597272.
28. Fu, H., Cao, X., Tu, Z. (2013). Cluster-based co-saliency detection. *IEEE Transactions on Image Processing: A Publication of the IEEE Signal Processing Society*, 22(10), 3766–3778. DOI 10.1109/TIP.2013.2260166.
29. Drews Jr, P., do Nascimento, E., Moraes, F., Botelho, S., Campos, M. et al. (2013). Transmission estimation in underwater single images. *IEEE International Conference on Computer Vision Workshops*, pp. 825–830. Sydney, NSW, Australia.
30. Song, W., Wang, Y., Huang, D., Tjondronegoro, D. (2018). 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. (Eds.), *Advances in multimedia information processing—PCM 2018*, vol. 11164, pp. 678–688. Springer International Publishing, Springer, Cham Publishing. DOI 10.1007/978-3-030-00776-8\_62.

31. Peng, Y. T., Cao, K., Cosman, P. C. (2018). Generalization of the dark channel prior for single image restoration. *IEEE Transactions on Image Processing*, 27(6), 2856–2868. DOI 10.1109/TIP.2018.2813092.
32. Li, C., Anwar, S., Porikli, F. (2020). Underwater scene prior inspired deep underwater image and video enhancement. *Pattern Recognition*, 98(1), 107038. DOI 10.1016/j.patcog.2019.107038.
33. Li, C., Guo, C., Ren, W., Cong, R., Hou, J. et al. (2019). An underwater image enhancement benchmark dataset and beyond. *IEEE Transactions on Image Processing: A Publication of the IEEE Signal Processing Society*, vol. 29, pp. 4376–4389. DOI 10.1109/TIP.2019.2955241.