Check for updates

An underwater image enhancement model for domain adaptation

Xiwen Deng[1,2], Tao Liu[1]*, Shuangyan He[1,2]*, Xinyao Xiao[1,2], Peiliang Li[1,2] and Yanzhen Gu[1,2]

[1]Ocean College, Zhejiang University, Zhoushan, China, [2]Hainan Institute, Zhejiang University, Sanya, China

Underwater imaging has been suffering from color imbalance, low contrast, and low-light environment due to strong spectral attenuation of light in the water. Owing to its complex physical imaging mechanism, enhancing the underwater imaging quality based on the deep learning method has been well-developed recently. However, individual studies use different underwater image datasets, leading to low generalization ability in other water conditions. To solve this domain adaptation problem, this paper proposes an underwater image enhancement scheme that combines individually degraded images and publicly available datasets for domain adaptation. Firstly, an underwater dataset fitting model (UDFM) is proposed to merge the individual localized and publicly available degraded datasets into a combined degraded one. Then an underwater image enhancement model (UIEM) is developed base on the combined degraded and open available clear image pairs dataset. The experiment proves that clear images can be recovered by only collecting the degraded images at some specific sea area. Thus, by use of the scheme in this study, the domain adaptation problem could be solved with the increase of underwater images collected at various sea areas. Also, the generalization ability of the underwater image enhancement model is supposed to become more robust. The code is available at https://github.com/fanren5599/UIEM.

KEYWORDS

underwater image, image enhancement, underwater dataset, domain adaptation, deep learning

1 Introduction

High-quality underwater photography and videography benefit ocean resource development and sustainable utilization (Mariani et al., 2018; Liu et al., 2020). However, underwater image degradation, including blurred edges and reduced visibility, is unavoidable due to the strong absorption and scattering of light in the water (Xie et al., 2022). The degree of degradation varies with different water constituents in various water areas (Fu et al., 2022). The enhanced visibility can make scenes and objects more

highlighted (Wang et al., 2023), although enhancing these poor-quality underwater images has been a challenging prerequisite for underwater object detection, monocular depth estimation, and underwater object tracking (Gao et al., 2019).

As shown in Figure 1, the underwater image quality mainly depends on two aspects (Rizzi et al., 2002; Foster, 2011; Ancuti et al., 2018). The first is brightness consistency. The reflected light of the underwater object is absorbed and scattered by the particles in the medium, and the camera usually captures low-contrast images. The second is color consistency. Light at long waves is more easily absorbed than short ones (Liu et al., 2020), thus underwater images generally appear bluish or greenish (Galdran et al., 2015). Accordingly, underwater image enhancement (UIE) algorithms basically aim at contrast enhancement and color cast correction (van de Weijer et al., 2007). In recent years, a variety of methods have been proposed for underwater image enhancement. Generally, the UIE methods can be divided into three categories: traditional methods without explicit model-building, physical prior-based methods, and data-driven deep learning methods.

Several traditional method scan better deal with noise, such as (Buchsbaum, 1980; Singh and Kapoor, 2014; Singh et al., 2015), lack of explicit modeling, which usually leads to problems of low contrast and color deviation. Several physical prior-based methods, such as (Kaiming He et al., 2009; Drews et al., 2016), can make the restored image conform to physical concepts, but the disadvantage is that there is often a lack of prior knowledge of experts and hand-crafted features (Lan et al., 2023). Several deep learning-based methods, such as (Li et al., 2017; Cao et al., 2018; Hou et al., 2018), have shown good enhancement and restoration performance. However, these studies are based on the data-driven approach, and thus their performance is limited by their specific dataset in two aspects. Firstly it is the domain adaptation problem (Huang and Belongie, 2017). The dataset used by different studies were collected in different water environments, and their optically active constituents vary with time and location (Devlin et al., 2009; Peterson et al., 2020). Besides, solar illumination also differs, and it also makes a significant difference in underwater objects' appearance and water styles. Limited training sample scan not

sufficiently cover changeably complex underwater environments (Chen et al., 2021). Therefore, even though one model recovers the underwater image of a specific sea area well, it is usually difficult to apply with good generalization abilities in other sea areas. Secondly, it is difficult for the marine engineering community to obtain a large dataset of degraded-clear pairwise underwater images (Wang et al., 2022a). The performance of the deep learning-based models based on pairwise datasets is supposed to be better than those using unpaired ones. In natural water such as rivers, lakes, and oceans, the degraded underwater images can be continuously collected, while it is usually impossible to obtain pairwise clear ones simultaneously.

As a result, in order to solve the above domain adaptation problem, we propose an underwater attenuation fitting network in this study. By combining utilizing individual degraded images and publicly available pairwise underwater data sets, it can merge the individual localized and publicly available degraded datasets into a combined degraded one. Thus, the combined degraded images and open available clear images compose a pairwise dataset. An underwater image enhancement method based on this dataset is also provided in this study. Using our method, even if only degraded images of a certain sea area are collected, an underwater enhancement model in the sea area can be built and applied well. In this study, we used Underwater Image Enhancement Benchmark (UIEB) (Li et al., 2020b), which contains 950 real Underwater images. Among them, 890 have corresponding reference images. We collected three datasets from different sea areas using the submarine online observation system for training and testing. We evaluated the quality of the generated images and the image quality of our dataset using the evaluation quality index of no reference images (NIQE) (Mittal et al., 2013). Then we propose a new image-to-image underwater image enhancement model to verify the validity of our dataset. The enhancement effect of the underwater image enhancement algorithms on our dataset is significantly higher than that on the public dataset.

This paper is organized as follows. Chapter two reviews the existing UIE algorithms and public datasets. In chapter three, the structure of our proposed UDFM and UIEM are described in detail. In the fourth chapter, the performance of experiments on different
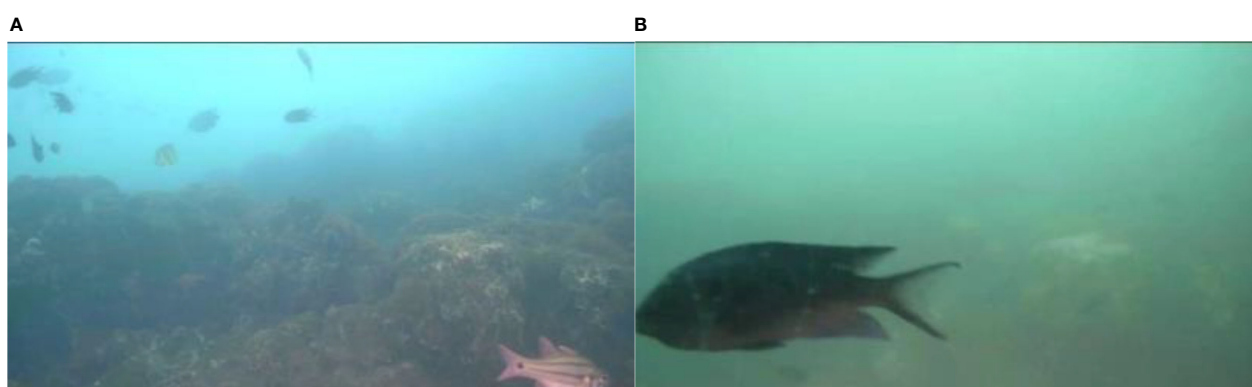

FIGURE 1
Challenges and problems in natural marine ranch environments: **(A)** a bluishimage with color distortion; **(B)** a low contrast and texture blurred image.

datasets is evaluated and compared with other image enhancement methods. The fifth part concludes and discusses this paper.

2 Relation work

2.1 Underwater image enhancement methods

Spatial domain methods such as histogram equalization (Singh and Kapoor, 2014)and gray world hypothesis algorithm (Buchsbaum, 1980), and frequency domain methods such as Fourier transform (Prabhakar and Praveen Kumar, 2011) and wavelet transform (Singh et al., 2015) are traditional methods. These methods can handle noises better but often bring problems of low contrast, loss of details, and color deviation.

Some researchers also built underwater imaging models to characterize the physical formation process of the underwater images and then estimated the depth-related transmission coefficients (Drews et al., 2016). Thus, they can reverse the degradation process to restore a clear underwater image, such as the Kai-ming DCP method (Kaiming He et al., 2009). The advantage of the method is that it can make the image enhancement process conform to physical concepts. The disadvantage is that experts' prior domain knowledge needs to be introduced into the model (Cheng et al., 2019), which does not always hold. These priors are often sensitive to the water environment where underwater images are collected, which causes the model may lack generalization. Besides, the imaging processing could be very complicated in various natural waters, and it is difficult to establish an appropriate general model.

With the improvement of the computing ability of hardware platforms, the introduction of excellent networks such as AlexNet (Krizhevsky et al., 2017) in 2012, and an ever-expanding database of digital images (Wang et al., 2022b), deep learning methods have been widely used in image processing (Li et al., 2017; Cao et al., 2018; Hou et al., 2018), natural language analysis (dos Santos and Gatti, 2014), and speech processing (Bollepalli et al., 2017). Data-driven deep learning-based methods are end-to-end methods, and they can solve the above problems easily by making the model learn basic parameters directly from the input. Deep convolutional network CNNs on large-scale datasets have shown excellent performance in many computer vision tasks (LeCun et al., 2015; He et al., 2016), which has motivated the development of data-driven UIE methods. At the same time, the proper availability of patterns for the training set affects the accuracy of test results (Naeem et al., 2022). However, real paired image datasets that meet the training objectives are often scarce. Thus, UIE methods usually use synthetic degraded images and high-quality counterpart images for training. Some UIE methods take a different approach and consider training with unpaired images.

As a pioneering work, Li et al. (Li et al., 2017) proposed the WaterGAN, which utilizes a generative adversarial network (GAN) and an image formation model to synthesize degraded/clear image pairs for unsupervised learning. To avoid the requirement of paired training data, Li et al. (Li et al., 2018) proposed a weakly supervised underwater color transformation model based on cycle-consistent adversarial networks, which alleviated the need for paired underwater images for training. However, the nature of the multiple possible outputs tends to produce unrealistic results in some cases. Li et al. (Li et al., 2020a) proposed to simulate real underwater images according to different water types and underwater imaging physical models. They first synthesized ten underwater images based on a revised underwater imaging model (Chiang and Chen, 2012). Then they use synthetic images to train the corresponding ten underwater image enhancement (UWCNN) models. This method can obtain a stable and good output. But facing the input underwater images, how to choose the appropriate UWCNN model is a challenge. Recently, Li et al. (Li et al., 2020b) collected a real-world paired underwater image dataset UIEBD to train a deep network. UIEBD can provide high-quality paired training data for depth models and a good evaluation of various underwater image enhancement methods.

2.2 Underwater image enhancement datasets

There are public underwater datasets, such as Fish4Knowledge (Boom et al., 2012), LifeCLEF2014 (Salman et al., 2016), LifeCLEF2015 (Salman et al., 2016), Sea-THRU (Salman et al., 2016), Haze-Line (Akkaynak and Treibitz, 2019), and UIEB(Devlin et al., 2009). These datasets are mainly used for object detection tasks. The Fish4Knowledge contains about 700,000 10-minute video clips of coral reefs spanned a time period of five years, including videos taken from sunrise to sunset. Note that these images were collected with varying water attenuation coefficients and solar illumination conditions and cannot be directly used in underwater image enhancement algorithms. The Sea-THRU dataset includes 1100 underwater images and range maps (Akkaynak and Treibitz, 2019), and the Haze-Line dataset has complete content, providing original images and camera calibration files (Berman et al., 2021). The UIEB (Li et al., 2020b) dataset contains 890 pairs of sharp and degraded pictures. In general, these existing datasets are individually collected in some specific sea areas in rather clear waters, so their representativeness is limited for real underwater images. Therefore, Li et al. (Li et al., 2020a) proposed to simulate real underwater images in different waters using underwater imaging physical models. They used the synthetic images to train ten underwater image enhancement (UWCNN) models. However, the underwater environment is complex, and ten models are not enough to characterize the underwater environment.

3 Method

The underwater imaging model (Chavez, 1988) can be expressed as:

$$I(x) = J(x) \star T(x) + A \star (1 - T(x)) \tag{1}$$

Where I(x) is the degraded underwater image; J(x) is the image to be recovered; A is background light; T(x) is the underwater medium transmission map depending on the water quality conditions, such as scattering and absorption coefficients of the water constituents. Therefore, the water quality conditions of other sea areas can be fused into the UIEB degraded images by extracting the medium transmission map.

3.1 Underwater dataset fitting model

Our underwater dataset fitting network is shown in Figure 2A. The input I are the degraded images of UIEB, and input II are the underwater images collected in other certain sea areas. In this study, the images were collected by an online seabed observation system in China Wei-zhou island. We first calculate the underwater medium transmission map of image II, and then splice the RGB image I and underwater medium transmission map of image II into the four-channel image III. Next, the output of a new degraded underwater image is generated by inputting Image III into the encoder and decoder. Note that the output image is generated from the degraded image I of UIEB, and it is further blurred with the underwater medium transmission map of image II by a neural network.

The neural network structure adopts the Encoder-Decoder structure. The structure of the Decoder is shown in Figure 2B. Encoder1 and Encoder2 use the first 31 layers of the VGG model. But the convolution kernel of the first layer is modified to 4×4 in Encoder1.

In this study, the underwater medium transmission map of T(x) is calculated based on the generalized dark channel prior algorithm (Peng et al., 2018), in which we improve the depth estimation method as shown in Equation (2).

$$D(x) = \min_{c,y \in \Theta(x)} (1 - w_c + w_c \times (|s_c - I^c(y)|)) \qquad (2)$$

In the formula, C represents the R, G, B three channels. s_c and w_c are determined by a_c. The a_c represents the proportional relation of C(R, G, or B channel) with the change of depth of field. Under different water quality conditions, the proportional relationship between light intensity and depth of field is different. A gradient map is first computed as the rough depth image on R, G, and B channel by using Sobel operators. The proportional relation can be fitted by the linear least square method under depth image. The a_c is the slope of the linear regression equation. If a_c is greater than 0, it means that the object is further away from the lens, the intensity of the corresponding channel is higher, and if a_c is less than 0, it means that the object is closer to the lens, and the intensity of the corresponding channel is higher. The relationship between a_c and s_c and w_c is shown as follows:

$$s_c = \begin{cases} 1 & if a_c > 0 \\ 0 & if a_c < 0 \end{cases} \qquad (3)$$
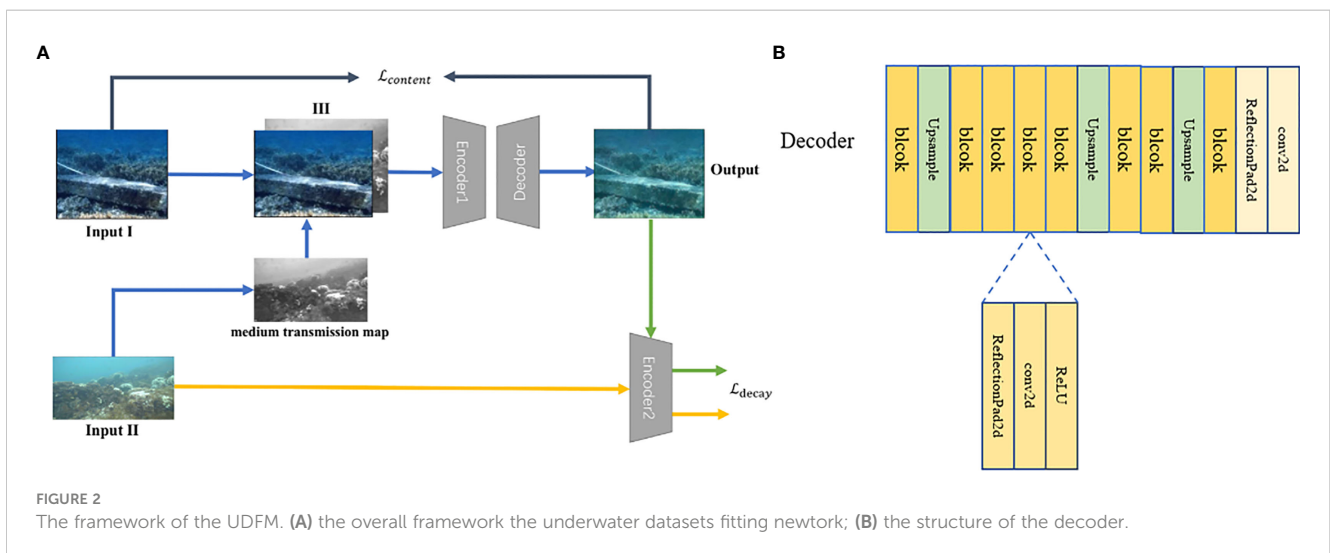
$$w_c = tanh(k|a_c| \qquad (4)$$

The w_c is the weight factor of each channel, and k is constant, which is generally taken as 5. The largest top 0.1% pixel values from the depth map are averaged as P. Then the Equation (5) can be used to estimate the media transmission map:

$$T(x) = \max_{c,y \in \Theta(x)} \left(\frac{|P - I^c(x)|}{max(P, \quad 1 - P)} \right) \qquad (5)$$

In the neural network structure, our loss function is divided into two parts: the content loss function and the ray attenuation loss function. The content loss function controls that the content of the generated image and the input I are the same object. The image attenuation of UIEB's degraded image I is controlled by the ray attenuation loss function. The ray attenuation is carried out according to the degraded degree of the underwater image II in a certain sea area. The content loss function is shown in Equation (6):

$$\mathcal{L}_{content} = \frac{\sum_{i=1}^{n}(Enc(x1)_i - Enc(x2)_i)^2}{n} \qquad (6)$$

Where $x1$ is the output image, $x2$ is the input image I, Enc is the first 31 layers of VGG model (VGG31), n is the pixel number of one image. The ray attenuation loss function uses the L1 loss function,

FIGURE 2
The framework of the UDFM. **(A)** the overall framework the underwater datasets fitting newtork; **(B)** the structure of the decoder.

and the specific equation is shown in Equation (7):

$$\mathcal{L}_{decay} = \frac{\sum_i^n |Mean(Relu_i(\text{Enc}(x1))) - Mean(Relu_i(\text{Enc}(x1)))|^2}{n} \quad (7)$$

Where x1 is the output image, x2 is the input image II, and $Relu_i$ refers to the RELU layer network of VGG31. In specific, it refers to the last four RELU layer networks. Then the output difference of each layer is averaged and calculated.

3.2 Underwater image enhancement model

In the underwater image enhancement task, more feature information about each pixel needs to be used than in other tasks, such as object detection and object classification. For example, each pixel uses global information to estimate the relative depth of the current data, to correct the color deviation. In addition, for every single pixel, the feature information of its nearby pixels is also needed to restore its texture information. To retain a large amount of existing information, we use the fifteen residual dense block (RDB) modules as the backbone network to further improve the enhanced image quality. The RDB module is composed of four dense blocks connected by residual structure, as shown in Figure 3. The bypass keeps the input information retained in the dense block (Huang et al., 2018), and the network layer learns and generates new information. Through the above design, our model is suitable for the image enhancement task. Moreover, we add residual structure between dense blocks to make training and convergence easier.

In the UIEM, our loss function is divided into the content loss function and the pix-to-pix loss function. The content loss function assures that the content of the generated image and the input are the same object. The U-net network (Ronneberger et al., 2015) implements the content loss function. The content loss function is shown in Equation (8):

$$\mathcal{L}_{content} = \frac{\sum_{i=1}^n (Unet(output)_i - Unet(gt)_i)^2}{n} \quad (8)$$

Where output is the output image, gt is the clear image in UIEB, Unet is the U-net model, n is the pixel number of an image. The pix-

to-pix loss function uses the L1 loss function, and the specific equation is shown in Equation (9):

$$\mathcal{L}_{pixtopix} = \frac{\sum_{i=1}^n (output_{i,j} - gt_{i,j})^2}{n} \quad (9)$$

4 Experiments and results

The underwater image enhancement has gradually formed its indicators, divided into two parts. One part is the comparison of result images, which we have put in Figure 4. The other part is the comparison of quantitative indicators. PSNR(peak signal-to-noise ratio) and SSIM(structural similarity) are most commonly used. However, it is found that the above indicators are not well consistent with human perception. And then, some scholars proposed an improved indicator named NIQE(no-reference image evaluation index). Therefore, we adopt NIQE in this paper. The results are compared in Tables 1, Table 2.

4.1 Underwater dataset fitting model experiments

Implementation details. The training and reference of the algorithm were carried out on a server configured with Intel(R) Xeon(R) Gold 6342 CPU ×2and NVIDIA A100 40 GB × 4 in this experiment. The PyTorch was used to build the backbone network and run on the Ubuntu 18.04 LTS operating system. We used Adam to optimize the overall objective. The learning rate was initialized as 1×10^{-4} and then decayed to 1×10^{-6} in the last 20 epochs. We trained with a batch size of 16 (on 4 GPUS) for 300K epochs.

Firstly, the performance of the data fitting model was verified. The loss function curve is shown in Figure 5 in the training stage, and it can be seen that the neural network began to converge at the 300K epoch, so we tested the model trained tillthe300K epoch.

In addition, the proportions of the content and ray loss functions were tested. The $\mathcal{L}_{content}/\mathcal{L}_{decay}$ was tested at 1:1, 5:1, 10:1 and 20:1, respectively. The convergence function curves are shown in Figure 6.

It has been proved by a lot of experiments that, as the data index of the experiment, higher peak signal-to-noise ratio (PSNR) or structural similarity (SSIM) does not always represent a better reconstruction effect, and the reconstructed texture often fails to meet the expectation of human eyes. Thus, the no-reference image evaluation index (NIQE) is used in this study. NIQE (Mittal et al., 2013) is based on 'quality perception' features. Usually, the features of simple and highly regular natural landscapes are extracted and used to fit a multivariate Gaussian model. The index quantifies the difference in the multivariate distribution of the image. In this study, the NIQE is used to quantify the difference between a set of generated images and an input set of images. The NIQE average is calculated in Equation (10).I are the UIEB degraded datasets, II are the other nature degraded datasets. III are the output images of the underwater dataset fitting model.
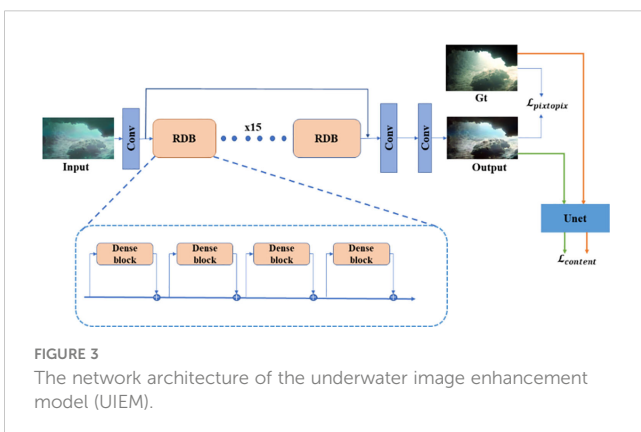

FIGURE 3
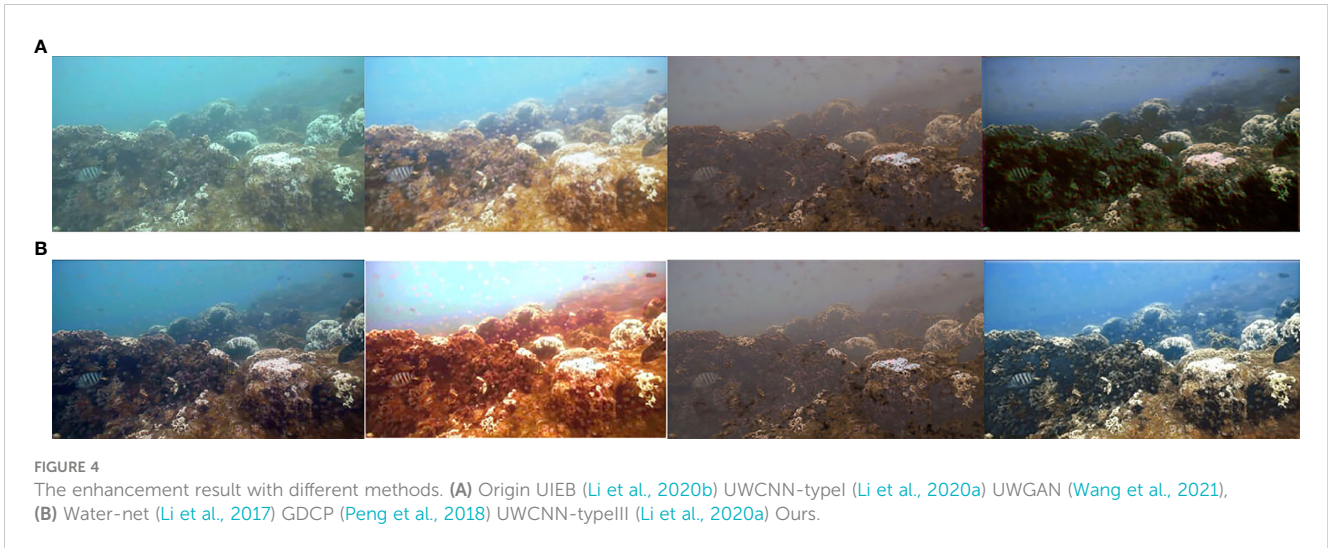The network architecture of the underwater image enhancement model (UIEM).

FIGURE 4
The enhancement result with different methods. **(A)** Origin UIEB (Li et al., 2020b) UWCNN-typeI (Li et al., 2020a) UWGAN (Wang et al., 2021),
(B) Water-net (Li et al., 2017) GDCP (Peng et al., 2018) UWCNN-typeIII (Li et al., 2020a) Ours.

TABLE 1 Comparison of results before and after using the underwater dataset fitting model.

Input I	Input II	NIQE average for the original dataset	NIQE average of the dataset after fitting
UIEB degraded datasets	North Sea dataset	3.88	0.73
UIEB degraded datasets	East Pole Island dataset	5.32	0.92

$$NIQE_{avg} = \frac{\sum_{i=1}^{n}(NIQE(I,\ II) - NIQE(III,\ II))}{n} \qquad (10)$$
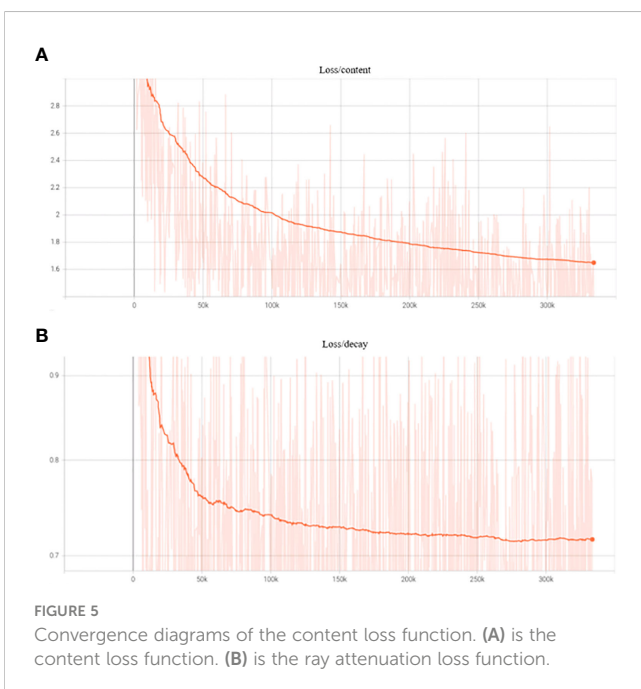

FIGURE 5
Convergence diagrams of the content loss function. **(A)** is the content loss function. **(B)** is the ray attenuation loss function.

TABLE 2 The Comparison of results from different encoders.

Model	$\mathcal{L}_{content}/\mathcal{L}_{decay}$	NIQE average
VGG31	1.61/0.73	0.73
DenseNet121	1.9/1.6	0.95
ResNet50	1.21/0.32	0.70
ResNet101	0.96/0.28	0.65

We tested two nature datasets, and the results are shown in Table 1. The *NIQE* average has all dropped below 1.

Besides that, the output results using different encoders are also given in Table 2. The neural network can encode more apparent information with the complex network structure, thus improving performance. As can be seen from the table, Resnet101 has the best effect as an encoder. The NIQE average is 0.65.

From the visual perspective, we demonstrate some images in Figure 7. Class (a) is the images of the publicly available dataset, class (b) is the underwater images of a sea area, and class (c) is the generated images.

4.2 Underwater image enhancement experiments

To evaluate the image restoration performance in a real underwater dataset, we compared UIEM with six deep learning methods, including UWCNN, UWGAN, Water-net, GDCP (Chavez, 1988), and UIEB. The network's input is the same underwater image without additional input, and the reasoning parameters of each network are the default parameters provided in the open-source network. It can be seen that when the input image quality is poor, the enhancement effects of UWCNN and UWGAN are not obvious and even appear more blurred. Water-net, GDCP, and UIEB can enhance the image to different degrees. However, Water-Net loses the picture's original color and enhances the contrast, making the whole image dark. UIEB also appears to strengthen excessively brightness,

and the details and texture of the picture after GDCP enhancement are unclear. The UIEM enhancement is the best. All the above image enhancement models use the generated dataset as the training set. The closer part of the results obtained by the proposed method is more precise than the UWCNN-typeI and UWCNN-typeIII, but the farther part is blurrier. In this study, we need to calculate the underwater medium transmission map, which represents the attenuation coefficient of underwater images. The distance of the contents in the image limits its accuracy. Its accuracy for the results of the near part is much greater than that of the far part, so it can be seen that the recovery of the near part is better than that of the far part. However, UWCNN-typeI and UWCNN-typeIII perform better in the restoration of the farther part of the image because the UWCNN method designs a variety of water types and then selects the algorithm model according to the water type. Therefore, The algorithm tends to restore the far water rather than the near target.
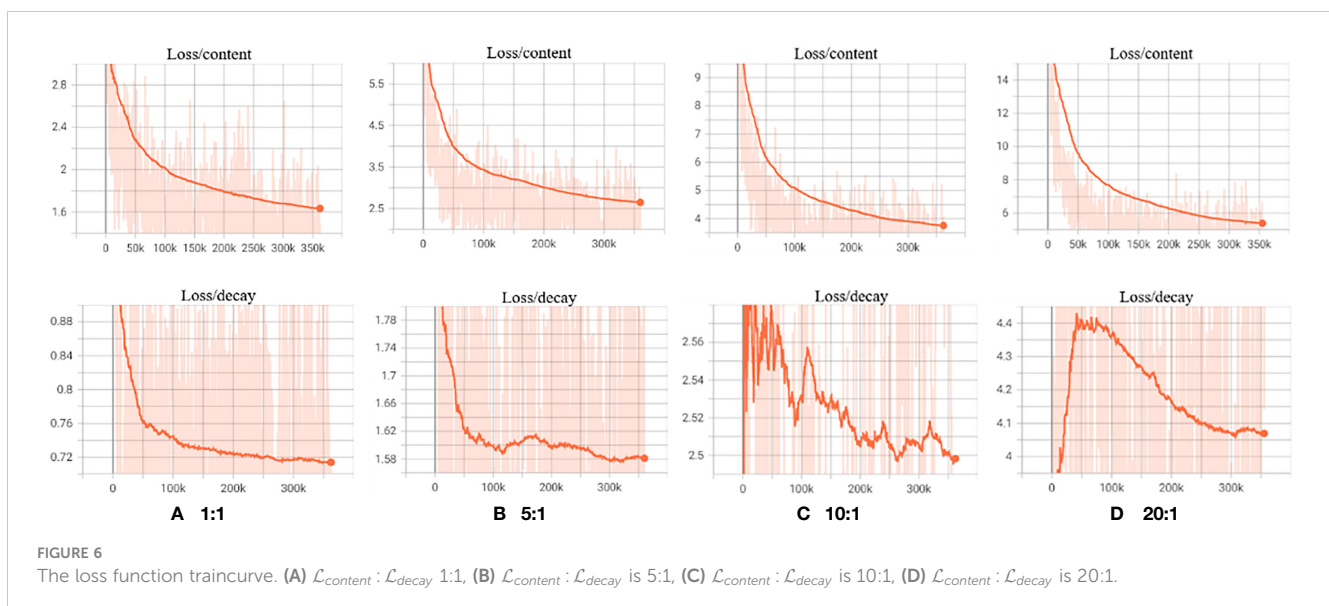
5 Discussion and conclusion

In this paper, we proposed a dataset fitting model UDFM to solve a challenging problem in underwater image enhancement. There are few high-quality images, not to mention corresponding clear-enough reference images for a specific sea area. Although the UIEB dataset has contributed clear paired datasets of underwater images for training, the results are unsatisfactory to test in a specific sea area. Our model has good generalization. The features of underwater images in specific sea areas are extracted and given to real images in UIEB to obtain synthetic images. Specific sea areas characterize these synthetic images. We use NIQE of the reference images (Mittal et al., 2013) to ensure that the quality of the generated images is close to our sea areas. Finally, image-to-image enhancement training can be performed using synthetic images and UIEB reference images. In the testing phase, we adopted the weights obtained after the training to the real blurred pictures. Experiments verify that our dataset augmentation method can be applied to underwater images under different sea area conditions. Just put the image of any sea area into the encoder to simulate the underwater environment. Experiments also verify that these image-to-image underwater image enhancement models achieve better results on our dataset than on the UIEB dataset. Our method can provide a baseline for the dataset synthesis method so that the training of image-to-image underwater image enhancement algorithms is no longer limited to the dataset. Our method provides new ideas for the study of domain adaptation.

In addition, we propose a deep learning-based image-to-image underwater image enhancement model UIEM. Our algorithm is based on dense blocks that can obtain more structural texture information, and we add residual structure. It is beneficial to the image restoration stage. This method can eliminate the effects of degradation and scatter on underwater images. Finally, we apply the MSE loss and introduce the Content loss (Wang et al., 2021) to train the network for further control over the content information. In addition, the results of our model are compared with existing algorithms. The effects of other algorithms are reddish, dark, and blurred. The UIEM has a leading position in underwater datasets.

The shortcoming of this model is that the model needs to be retrained when the data sets of other sea areas are collected, and the fitting ability of the model can only become more robust with the expansion of the data sets. Secondly, for some pictures, there will be color bands. As shown in Figure 8, blue color bands exist in the upper part of the processed picture, which is caused by the fact that the data fitting model is not optimal. Our future research direction is to combine these two models and solve the phenomenon of color bands in the model.

In future work, we will try to collect data from different sea areas to improve the generalization ability of our underwater image enhancement model. And we will further optimize our data fitting model to solve the phenomenon of color bands. We will also try to combine our method with algorithms such as object detection for better progress in more advanced vision tasks.
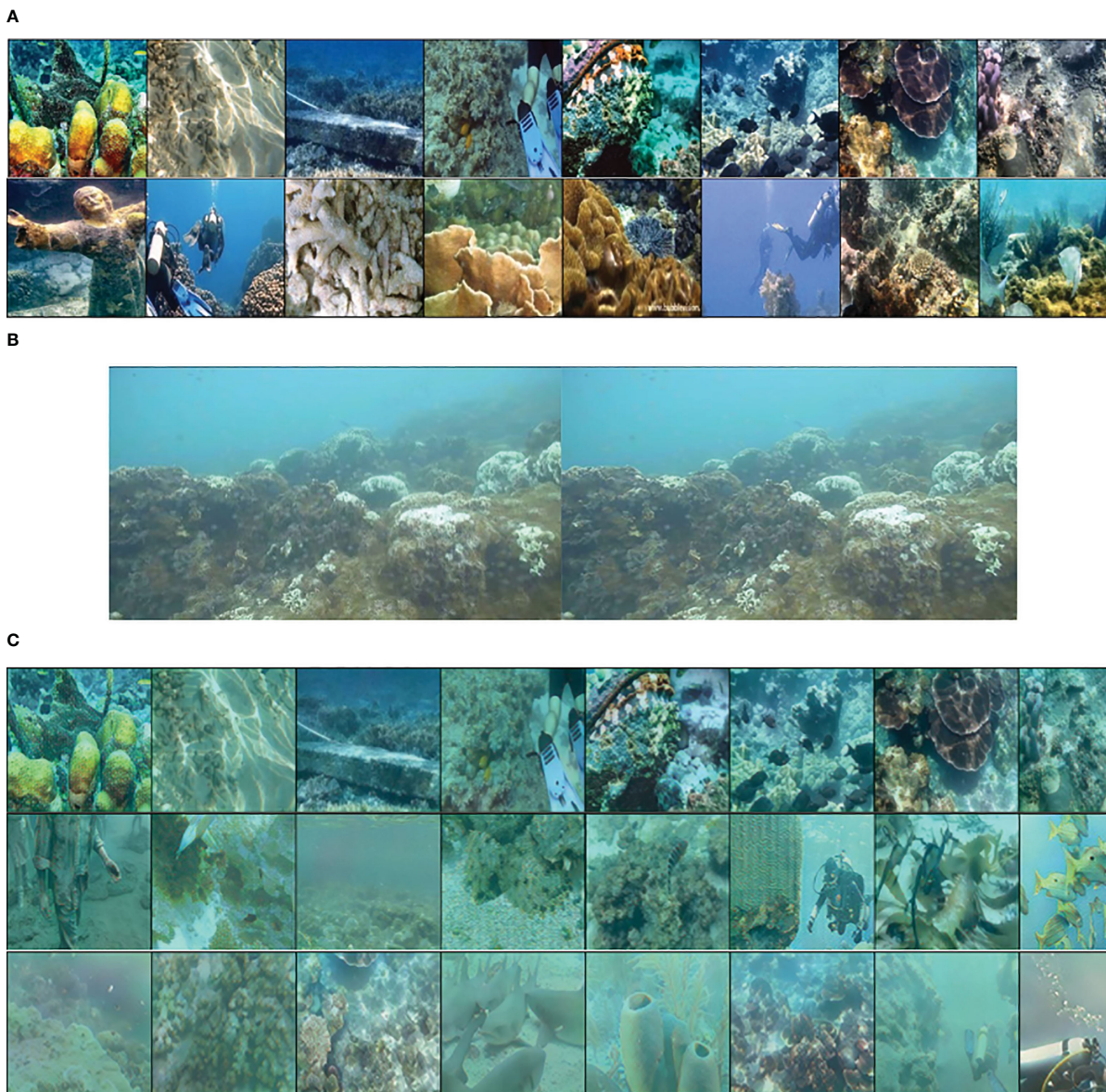

FIGURE 6
The loss function traincurve. (A) $\mathcal{L}_{content} : \mathcal{L}_{decay}$ 1:1, (B) $\mathcal{L}_{content} : \mathcal{L}_{decay}$ is 5:1, (C) $\mathcal{L}_{content} : \mathcal{L}_{decay}$ is 10:1, (D) $\mathcal{L}_{content} : \mathcal{L}_{decay}$ is 20:1.

FIGURE 7
(A) The images of the public data set, **(B)** the underwater images of a sea area, and **(C)** the generated images.
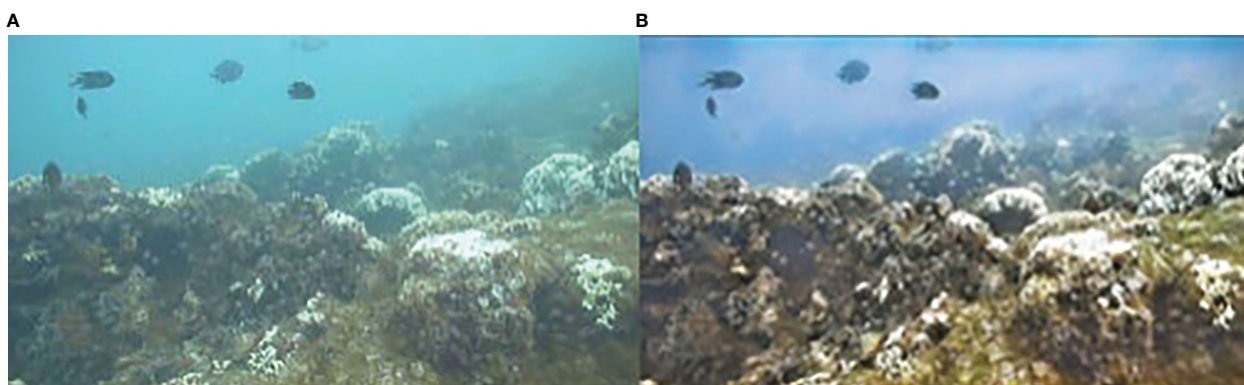

FIGURE 8
The enhancement result from the phenomenon of color bands. **(A)** The origin picture. **(B)** The enhancement picture with color bands.

Data availability statement

The data is provided by Hainan Observation and Research Station of Ecological Environment and Fishery Resource in Yazhou Bay. The original contributions presented in the study are included in the article/supplementary material. Further inquiries can be directed to the corresponding authors.

Author contributions

The algorithm studied by XD and TL, and the experimental results organized by SH and XX. Paper by PL and YG. All authors contributed to the article and approved the submitted version.

Funding

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

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

References

Akkaynak, D., and Treibitz, T. (2019). "Sea-Thru: A method for removing water from underwater images," in *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Long Beach, CA, USA. 1682–1691 (IEEE). doi: 10.1109/CVPR.2019.00178

Ancuti, C. O., Ancuti, C., De Vleeschouwer, C., and Bekaert, P. (2018). Color balance and fusion for underwater image enhancement. *IEEE Trans. Image Process.* 27, 379–393. doi: 10.1109/TIP.2017.2759252

Berman, D., Levy, D., Avidan, S., and Treibitz, T. (2021). Underwater single image color restoration using haze-lines and a new quantitative dataset. *IEEE Trans. Pattern Anal. Mach. Intell.* 43, 2822–2837. doi: 10.1109/TPAMI.2020.2977624

Bollepalli, B., Juvela, L., and Alku, P. (2017). "Generative adversarial network-based glottal waveform model for statistical parametric speech synthesis," in *Interspeech 2017*. 3394–3398. doi: 10.21437/Interspeech.2017-1288

Boom, B. J., Huang, P. X., He, J., and Fisher, R. B. (2012). "Supporting ground-truth annotation of image datasets using clustering," in *Proceedings of the 21st International Conference on Pattern Recognition (ICPR2012)*. 1542–1545.

Buchsbaum, G. (1980). A spatial processor model for object colour perception. *J. Franklin Institute* 310, 1–26. doi: 10.1016/0016-0032(80)90058-7

Cao, K., Peng, Y.-T., and Cosman, P. C. (2018). "Underwater image restoration using deep networks to estimate background light and scene depth," in *2018 IEEE Southwest Symposium on Image Analysis and Interpretation (SSIAI)*, Las Vegas, NV. 1–4 (IEEE). doi: 10.1109/SSIAI.2018.8470347

Chavez, P. S. (1988). An improved dark-object subtraction technique for atmospheric scattering correction of multispectral data. *Remote Sens. Environ.* 24, 459–479. doi: 10.1016/0034-4257(88)90119-3

Chen, T., Wang, N., Wang, R., Zhao, H., and Zhang, G. (2021). One-stage CNN detector-based benthonic organisms detection with limited training dataset. *Neural Networks* 144, 247–259. doi: 10.1016/j.neunet.2021.08.014

Cheng, S., Zhao, K., and Zhang, D. (2019). Abnormal water quality monitoring based on visual sensing of three-dimensional motion behavior of fish. *Symmetry* 11, 1179. doi: 10.3390/sym11091179

Chiang, J. Y., and Chen, Y.-C. (2012). Underwater image enhancement by wavelength compensation and dehazing. *IEEE Trans. Image Process.* 21, 1756–1769. doi: 10.1109/TIP.2011.2179666

Devlin, M. J., Barry, J., Mills, D. K., Gowen, R. J., Foden, J., Sivyer, D., et al. (2009). Estimating the diffuse attenuation coefficient from optically active constituents in UK marine waters. *Estuarine Coast. Shelf Sci.* 82, 73–83. doi: 10.1016/j.ecss.2008.12.015

dos Santos, C. N., and Gatti, M. (2014). Deep convolutional neural networks for sentiment analysis. *COLING* 10, 69–78. Available at: https://aclanthology.org/C14-1008

Drews, P. L. J., Nascimento, E. R., Botelho, S. S. C., and Montenegro Campos, M. F. (2016). Underwater depth estimation and image restoration based on single images. *IEEE Comput. Grap. Appl.* 36, 24–35. doi: 10.1109/MCG.2016.26

Foster, D. H. (2011). Color constancy. *Vision Res.* 51, 674–700. doi: 10.1016/j.visres.2010.09.006

Fu, Z., Lin, X., Wang, W., Huang, Y., and Ding, X. (2022)Underwater image enhancement via learning water type desensitized representations. In: *arXiv:2102.00676*. Available at: http://arxiv.org/abs/2102.00676 (Accessed April 8, 2022).

Galdran, A., Pardo, D., Picón, A., and Alvarez-Gila, A. (2015). Automatic red-channel underwater image restoration. *J. Visual Communication Image Representation* 26, 132–145. doi: 10.1016/j.jvcir.2014.11.006

Gao, S.-B., Zhang, M., Zhao, Q., Zhang, X.-S., and Li, Y.-J. (2019). Underwater image enhancement using adaptive retinal mechanisms. *IEEE Trans. Image Process.* 28, 5580–5595. doi: 10.1109/TIP.2019.2919947

He, K., Sun, J., and Tang, X. (2009). "Single image haze removal using dark channel prior," in *2009 IEEE Conference on Computer Vision and Pattern Recognition*, Miami, FL. 1956–1963 (IEEE). doi: 10.1109/CVPR.2009.5206515

He, K., Zhang, X., Ren, S., and Sun, J. (2016). "Deep residual learning for image recognition," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA. 770–778 (IEEE). doi: 10.1109/CVPR.2016.90

Hou, M., Liu, R., Fan, X., and Luo, Z. (2018). "Joint residual learning for underwater image enhancement," in *2018 25th IEEE International Conference on Image Processing (ICIP)*, Athens. 4043–4047 (IEEE). doi: 10.1109/ICIP.2018.8451209

Huang, X., and Belongie, S. (2017)Arbitrary style transfer in real-time with adaptive instance normalization. In: *arXiv:1703.06868*. Available at: http://arxiv.org/abs/1703.06868 (Accessed April 8, 2022).

Huang, G., Liu, Z., van der Maaten, L., and Weinberger, K. Q. (2018) *Densely connected convolutional networks*. Available at: http://arxiv.org/abs/1608.06993 (Accessed September 6, 2022).

Krizhevsky, A., Sutskever, I., and Hinton, G. E. (2017). ImageNet classification with deep convolutional neural networks. *Commun. of the ACM* 60, 84-90. doi: 10.1145/3065386.

Lan, Z., Zhou, B., Zhao, W., and Wang, S. (2023). An optimized GAN method based on the que-attn and contrastive learning for underwater image enhancement. *PloS One* 18, e0279945. doi: 10.1371/journal.pone.0279945

LeCun, Y., Bengio, Y., and Hinton, G. (2015). Deep learning. *Nature* 521, 436–444. doi: 10.1038/nature14539

Li, C., Anwar, S., and Porikli, F. (2020a). Underwater scene prior inspired deep underwater image and video enhancement. *Pattern Recognition* 98, 107038. doi: 10.1016/j.patcog.2019.107038

Li, C., Guo, J., and Guo, C. (2018). Emerging from water: Underwater image color correction based on weakly supervised color transfer. *IEEE Signal Process. Lett.* 25, 323–327. doi: 10.1109/LSP.2018.2792050

Li, C., Guo, C., Ren, W., Cong, R., Hou, J., Kwong, S., et al. (2020b). An underwater image enhancement benchmark dataset and beyond. *IEEE Trans. Image Process.* 29, 4376–4389. doi: 10.1109/TIP.2019.2955241

Li, J., Skinner, K. A., Eustice, R. M., and Johnson-Roberson, M. (2017). WaterGAN: Unsupervised generative network to enable real-time color correction of monocular underwater images. *IEEE Robot. Autom. Lett.*, 1. doi: 10.1109/LRA.2017.2730363

Liu, R., Fan, X., Zhu, M., Hou, M., and Luo, Z. (2020). Real-world underwater enhancement: Challenges, benchmarks, and solutions under natural light. *IEEE Trans. Circuits Syst. Video Technol.* 30, 4861–4875. doi: 10.1109/TCSVT.2019.2963772

Mariani, P., Quincoces, I., Haugholt, K., Chardard, Y., Visser, A., Yates, C., et al. (2018). Range-gated imaging system for underwater monitoring in ocean environment. *Sustainability* 11, 162. doi: 10.3390/su11010162

Mittal, A., Soundararajan, R., and Bovik, A. C. (2013). Making a "Completely blind" image quality analyzer. *IEEE Signal Process. Lett.* 20, 209–212. doi: 10.1109/LSP.2012.2227726

Naeem, A., Anees, T., Ahmed, K. T., Naqvi, R. A., Ahmad, S., and Whangbo, T. (2022). Deep learned vectors' formation using auto-correlation, scaling, and derivations with CNN for complex and huge image retrieval. *Complex Intell. Syst.* doi: 10.1007/s40747-022-00866-8

Peng, Y.-T., Cao, K., and Cosman, P. C. (2018). Generalization of the dark channel prior for single image restoration. *IEEE Trans. Image Process.* 27, 2856–2868. doi: 10.1109/TIP.2018.2813092

Peterson, K. T., Sagan, V., and Sloan, J. J. (2020). Deep learning-based water quality estimation and anomaly detection using landsat-8/Sentinel-2 virtual constellation and cloud computing. *GIScience Remote Sens.* 57, 510–525. doi: 10.1080/15481603.2020.1738061

Prabhakar, C. J., and Praveen Kumar, P. U. (2011). An image based technique for enhancement of underwater images. *Int. J. Mach. Intell.* 3, 217–224. doi: 10.48550/arXiv.1212.0291

Rizzi, A., Gatta, C., and Marini, D. (2002). Color correction between gray world and white patch. *Human Vision and Electronic Imaging VII.* 4662, 367–375. doi: 10.1117/12.469534

Ronneberger, O., Fischer, P., and Brox, T. (2015). U-Net: Convolutional networks for biomedical image segmentation. *Medical image computing and computer-assisted intervention, PT III.* 9351, 234-241. doi: 10.48550/arXiv.1505.04597

Salman, A., Jalal, A., Shafait, F., Mian, A., Shortis, M., Seager, J., et al. (2016). Fish species classification in unconstrained underwater environments based on deep learning: Fish classification based on deep learning. *Limnol. Oceanogr. Methods* 14, 570–585. doi: 10.1002/lom3.10113

Singh, G., Jaggi, N., Vasamsetti, S., Sardana, H. K., Kumar, S., and Mittal, N. (2015). "Underwater image/video enhancement using wavelet based color correction (WBCC) method," in *2015 IEEE Underwater Technology (UT)*, Chennai, India. 1–5 (IEEE). doi: 10.1109/UT.2015.7108303

Singh, K., and Kapoor, R. (2014). Image enhancement using exposure based Sub image histogram equalization. *Pattern Recognition Lett.* 36, 10–14. doi: 10.1016/j.patrec.2013.08.024

van de Weijer, J., Gevers, T., and Gijsenij, A. (2007). Edge-based color constancy. *IEEE Trans. Image Process.* 16, 2207–2214. doi: 10.1109/TIP.2007.901808

Wang, W., Jiao, P., Liu, H., Ma, X., and Shang, Z. (2022b). Two-stage content based image retrieval using sparse representation and feature fusion. *Multimed Tools Appl.* 81, 16621–16644. doi: 10.1007/s11042-022-12348-7

Wang, Z., Shen, L., Xu, M., Yu, M., Wang, K., and Lin, Y. (2023). Domain adaptation for underwater image enhancement. *IEEE Trans. Image Process.* 32, 1442–1457. doi: 10.1109/TIP.2023.3244647

Wang, N., Wang, Y., and Er, M. J. (2022a). Review on deep learning techniques for marine object recognition: Architectures and algorithms. *Control Eng. Pract.* 118, 104458. doi: 10.1016/j.conengprac.2020.104458

Wang, N., Zhou, Y., Han, F., Zhu, H., and Yao, J. (2021). UWGAN: Underwater GAN for real-world underwater color restoration and dehazing. *CoRR* doi: 10.48550/arXiv.1912.10269

Xie, Y., Yu, Z., Yu, X., and Zheng, B. (2022). Lighting the darkness in the sea: A deep learning model for underwater image enhancement. *Front. Mar. Sci.* 9. doi: 10.3389/fmars.2022.921492