

Article

Underwater Image Enhancement via Triple-Branch Dense Block and Generative Adversarial Network

Peng Yang ^{1,2,†}, Chunhua He ^{1,2,†} , Shaojuan Luo ^{3,†}, Tao Wang ^{1,2,*} and Heng Wu ^{1,2,*} 

¹ Guangdong Provincial Key Laboratory of Cyber-Physical System, School of Automation, Guangdong University of Technology, Guangzhou 510006, China

² School of Computer, Guangdong University of Technology, Guangzhou 510006, China

³ School of Chemical Engineering and Light Industry, Guangdong University of Technology, Guangzhou 510006, China

* Correspondence: wangtao_cps@gdut.edu.cn (T.W.); heng.wu@foxmail.com (H.W.)

† These authors contributed equally to this work.

Abstract: The complex underwater environment and light scattering effect lead to severe degradation problems in underwater images, such as color distortion, noise interference, and loss of details. However, the degradation problems of underwater images bring a significant challenge to underwater applications. To address the color distortion, noise interference, and loss of detail problems in underwater images, we propose a triple-branch dense block-based generative adversarial network (TDGAN) for the quality enhancement of underwater images. A residual triple-branch dense block is designed in the generator, which improves performance and feature extraction efficiency and retains more image details. A dual-branch discriminator network is also developed, which helps to capture more high-frequency information and guides the generator to use more global content and detailed features. Experimental results show that TDGAN is more competitive than many advanced methods from the perspective of visual perception and quantitative metrics. Many application tests illustrate that TDGAN can significantly improve the accuracy of underwater target detection, and it is also applicable in image segmentation and saliency detection.

Keywords: generative adversarial network (GAN); underwater image enhancement; multiscale dense; residual learning



Citation: Yang, P.; He, C.; Luo, S.; Wang, T.; Wu, H. Underwater Image Enhancement via Triple-Branch Dense Block and Generative Adversarial Network. *J. Mar. Sci. Eng.* **2023**, *11*, 1124. <https://doi.org/10.3390/jmse11061124>

Academic Editor: Fabio Bruno

Received: 8 May 2023

Revised: 19 May 2023

Accepted: 24 May 2023

Published: 26 May 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Underwater imaging technology is widely used in deep-sea resource exploration, marine rescue, biodiversity monitoring, and submarine cable laying. However, the images captured by the underwater cameras encounter several degradation problems, such as color distortion, low contrast, and blurring [1]. The main reasons for the above problems are threefold. Firstly, the underwater propagation of light is exponentially attenuated, and this attenuation acts on different wavelengths. In particular, the red channel has the most obvious attenuation. Therefore, raw underwater images are always bluish or greenish compared to images taken in the air. Secondly, stray light is added to the sensor due to the scattering effect [2], resulting in a haze effect on the entire scene and reducing the image resolution and quality. Finally, the colors of underwater images are often distorted, owing to the influence of external environments such as water depth and lighting conditions [3]. The above three aspects cause the underwater image degradation, which cannot meet the practical application requirements. Moreover, the limitations of camera equipment are also a factor in the degradation of underwater images. The reason is that cameras cannot, as the human eye can, capture the full range of brightness [4]. Note that underwater images refer to photographs captured by cameras in the underwater environment.

Many image enhancement and restoration methods have been developed to improve the quality of underwater images in the past few years. Underwater image enhancement methods

are mainly based on the redistribution of pixel intensities to improve the color and contrast of images; for example, the image enhancement method [5] enhances underwater images by adjusting image pixel values. Unlike underwater image enhancement, underwater image restoration usually requires the establishment of an effective degradation model and consideration of the underwater imaging mechanism and the physical properties of underwater light propagating. The key parameters of the constructed physical model are derived through prior knowledge, and the underwater image is restored through a retention compensation process [6]. For example, Refs. [7–9] used an image formation model to restore underwater images. However, considering the complex underwater physics and optics in the image enhancement and restoration process, the limitations of traditional methods are apparent. Limited by the problems of insufficient training data and parameter selection, traditional methods encounter the poor generalization performance problem, which is manifested in the cases wherein the enhanced images of some scenes are over-enhanced or under-enhanced. With the development of artificial intelligence technology, deep learning methods have been applied to underwater image processing [10–13]. Under the premise of abundant training data, deep learning dramatically improves the generalization ability of the traditional methods and can enhance the image quality in different underwater scenes. For instance, Li et al. [10] proposed WaterGAN, which uses synthetic real-world images, raw underwater images, and depth data to train a deep learning network for correcting color underwater images. Similar to waterGAN, Fabbri et al. also used a GAN method to enhance underwater images. Firstly, the distorted image is reconstructed based on the CycleGAN [14], and then the underwater GAN (UGAN) [15] is used to train the reconstructed underwater image pairs. Finally, a clear underwater image is obtained based on the pix2pix [16] model. Furthermore, based on Ref. [16], Islam et al. [17] designed a fast underwater image enhancement model (FunieGAN) using the U-net network to generate underwater images with rich visual perception. Although deep learning-based methods can obtain high-quality underwater images, the training time and image quality (e.g., noise and image details) depend on the structure of the neural network. Moreover, color consistency and training stability are the main problems that restrict the performance of the existing deep learning-based methods.

To solve the problems of color distortion, noise interference, and detail loss in the underwater image, we propose a generative adversarial network with a triple-branch dense block (TDGAN). Firstly, a triple-branch dense block (TBDB) is developed without building an underwater degradation model and prior image knowledge. The TBDB, which combines dense stitching, multi-scale techniques, and residual learning, can fully utilize feature information and recover image details. Secondly, a dual-branch discriminative network that can extract high-frequency information is designed to extract high-dimensional features and obtain low-dimensional discriminative information. The discriminator can guide the generator to pay attention to the global semantics and local details of the image and output images with prominent local details. In addition, a multinomial loss function is constructed to enrich the visual appearance of the images and obtain high-quality underwater images that align with human visual perception. The non-reference and full-reference metrics are utilized for quantitative comparison, and many experiments prove that TDGAN has higher evaluation metrics on underwater images. Moreover, two ablation studies are used to demonstrate the function of each component in the module. Finally, application tests are used to verify the effectiveness of TDGAN.

The main contributions of this paper are summarized as follows:

- (1) We propose a TDGAN for underwater image enhancement. Extensive experiments demonstrate that TDGAN can improve the quality of underwater images and has potential applications in the fields of image denoising, object detection, image segmentation, and so on;
- (2) We design a dual-branch discriminator to reconstruct underwater images. The discriminator can guide the generator to exploit global semantics and local details fully;
- (3) We develop a TBDB that can significantly improve the feature mining ability of the network and make full use of the underlying semantic information. Compared with

the dual-scale channel MSDB in UWGAN [18], the TBDB adopts three channels with different scales, which can obtain different levels of detailed information and is more sensitive to detail changes.

2. Related Work

Underwater image enhancement methods can be divided into traditional and deep learning-based methods. Generally, underwater image enhancement methods revolve around one or more image degradation problems, such as image contrast enhancement, color restoration, detail reproduction, brightness enhancement, and noise removal.

2.1. Traditional Underwater Image Enhancement Methods

Recently, many traditional methods have been proposed to enhance or restore underwater images, such as dark channel prior (DCP) [19], histogram equalization (HE) [20], contrast limited adaptive histogram equalization (CLAHE) [21], unsupervised colour correction method (UCM) [22], and underwater light attenuation prior (ULAP) [23], etc. On the other hand, with the purpose of improving underwater image contrast, Deng et al. presented a removing light source color and dehazing (RLSCD) method [24], which considered the scene depth correlation with decay. The results show that the RLSCD method has the advantage of improving the contrast and brightness of underwater images. Tao et al. [25] reconstructed high-quality underwater images by improving white balance and image fusion strategies. In [25], a multi-scale fusion scheme is designed to adjust the contrast, saturation, and brightness of the image.

Moreover, to solve the problem of color cast and low visibility of underwater images, Ref. [26] proposed an underwater image enhancement method called MLLE. In this method, the color and details of the image are locally adjusted through a fusion strategy, and the contrast of the image is adaptively adjusted by the mean and variance of the local image blocks. The underwater images of MLLE are characterized by high contrast and clarity. To solve the problem of color distortion in underwater images, Ke et al. [27] designed a framework for underwater image restoration. Firstly, a color correction is performed in the Lab color space to remove the color cast. Then, the transmission map of each channel is corrected using the relationship between the scattering coefficient and wavelength. Experiments show that this method significantly improves underwater image detail and color saturation. For the problem of underwater image color correction, Zhang et al. [28] proposed a color correction method, where a dual-histogram-based iterative thresholding method was developed to obtain global contrast-enhanced images and a finite histogram method with Rayleigh distribution was designed to obtain local contrast-enhanced images.

2.2. Underwater Image Enhancement Method Based on Deep Learning

Over the past few years, deep learning-based underwater image enhancement methods have made remarkable achievements. However, many underwater image enhancement techniques based on deep learning often produce artifacts and color distortion. To solve these problems, Wang et al. [29] proposed a two-phase underwater domain adaptation network (TUDA) to generate underwater images competitive in both visual quality and quantitative metrics. Sun et al. [30] developed an underwater multi-scene generative adversarial network (UMGAN) to enhance underwater images. This method uses a feedback mechanism and a denoising network to address noise and artifacts in generated images. To study the inherent degradation factors of underwater images and improve the network's generalization ability, Xue et al. [31] designed a multi-branch aggregation network (MBANet). The MBANet analyzes underwater degradation factors from the perspective of color distortion and the veil effect and can significantly improve the performance of underwater object detection.

Furthermore, Cai et al. [32] proposed a CURE-Net to enhance the details of underwater images. The CURE-Net is composed of three cascaded subnetworks, a detail enhancement block, and a supervisory restoration block. The results indicate that CURE-Net achieves a

gradual improvement of degraded underwater images. Ref. [33] developed a priori-guided adaptive underwater compressed sensing framework (UCSNet) to reproduce underwater image details. UCSNet uses the principle of multiple networks, where the sampling matrix generation network (SMGNet) is used to capture structural information and highlight image details.

2.3. Underwater Image Evaluation Metrics

Since image quality is affected by many factors, the assessment of the image quality is usually divided into two types: qualitative and quantitative assessment. Underwater image quality evaluation indicators commonly used by researchers are Underwater Color Image Quality Evaluation (UCIQE) [34] and Underwater Image Quality Metric (UIQM) [35]. In 2015, Yang et al. [34] found a correlation among the image sharpness, color, and subjective perception of the image quality and proposed an image quality evaluation method (UCIQE) for underwater images. UCIQE is a linear model involving contrast, hue, and saturation. Like UCIQE, UIQM constructs a linear combination of Underwater Image Color Metric (UICM), Underwater Image Sharpness Metric (UISM), and Underwater Image Contrast Metric (UIConM). Therefore, the larger the UCIQE and UIQM, the better the underwater image quality.

Additionally, full-reference image quality assessment metrics Peak Signal to Noise Ratio (PSNR) [36] and Structural Similarity Index Measurement (SSIM) [37] are also often used to evaluate the quality between the generated image and the reference image. The larger the PSNR and SSIM values, the better quality of the generated images.

3. Method

A TDGAN is proposed to restore clear underwater images (photographs) from the original ones. The architecture of TDGAN is shown in Figure 1. TDGAN consists of two parts, the generator and discriminator networks. Specifically, a TBDB is developed for the generator, and a dual-branch discriminator is designed to guide the generator network to generate underwater images with more salient details. Specifically, as shown in Figure 1, the data flow of the proposed method is expressed as $\{[X_{GT} + X_G]; [X_{1/2GT} + X_{1/2G}]\} \rightarrow [B_1; B_2]_D \rightarrow [c, s, c] \rightarrow L_{TDGAN}$, where X_{GT} and X_G denote the reference image (ground truth) in the dataset and the image generated by the generator, respectively; $X_{1/2GT}$ and $X_{1/2G}$ represent the images after reducing the width of the X_{GT} and X_G images by half, respectively; B_1 and B_2 mean the two branches of the discriminator network; c , s , and L_{TDGAN} are the convolution, sigmoid activation function, and loss function, respectively.

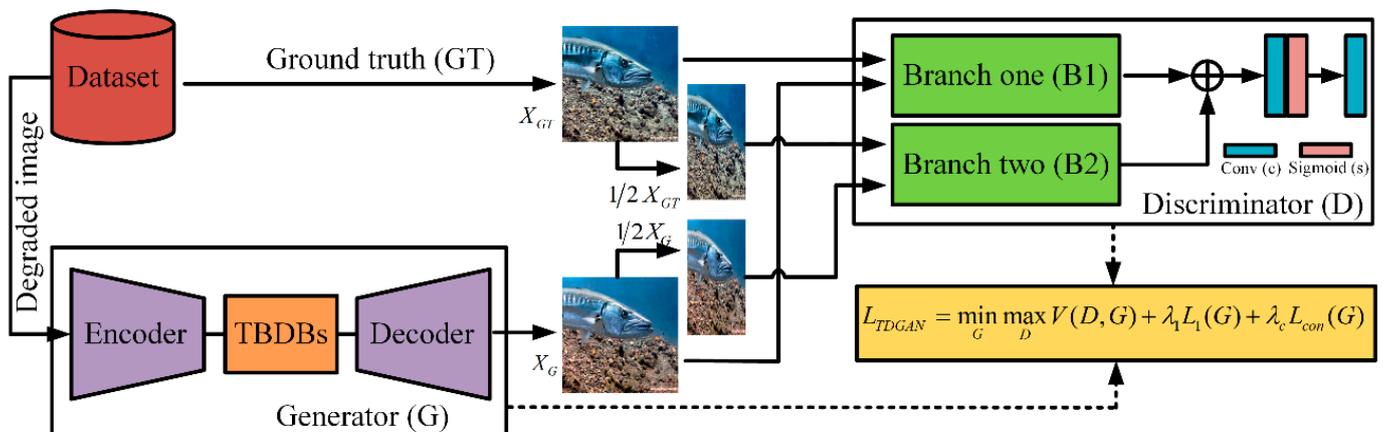


Figure 1. Model overview of TDGAN.

3.1. Generator Network

As shown in Figure 2, the proposed generator includes six convolutional blocks and two triple-branch dense blocks (TBDBs). The three convolution blocks on the left of TBDBs

perform the down-sampling and encoding operations. The three convolutional blocks on the right side of TBDBs perform the up-sampling and decoding operations. The network structure and parameters of TDGAN are shown in Table 1, where BN stands for batch normalization [38], and Leaky_ReLU stands for Leaky Rectified Linear Unit. The slope of all Leaky_ReLU activation functions is 0.2. Three convolutional layers are placed in the first three layers of the generator. The first, second, and third kernel sizes are 7×7 , 5×5 , and 3×3 , respectively. The corresponding feature maps of the first three layers are 64, 128, and 256, respectively. Note that each layer of the first three is processed by Leaky_ReLU and BN, respectively. The convolutions of the first three layers can reduce the size of the feature map and extract preliminary features. Therefore, to capture contextual image information at multiple scales, we concatenate multi-scale dense blocks (TBDBs) at the outputs of the first three layers. The following three deconvolution layers are used to reconstruct the image. The last deconvolution layer ensures that the numbers of output channels and input channels are the same, and it uses the Tanh function to match the distribution of input channels $[-1, 1]$. Both the input and output image sizes of the generator network are $256 \times 256 \times 3$.

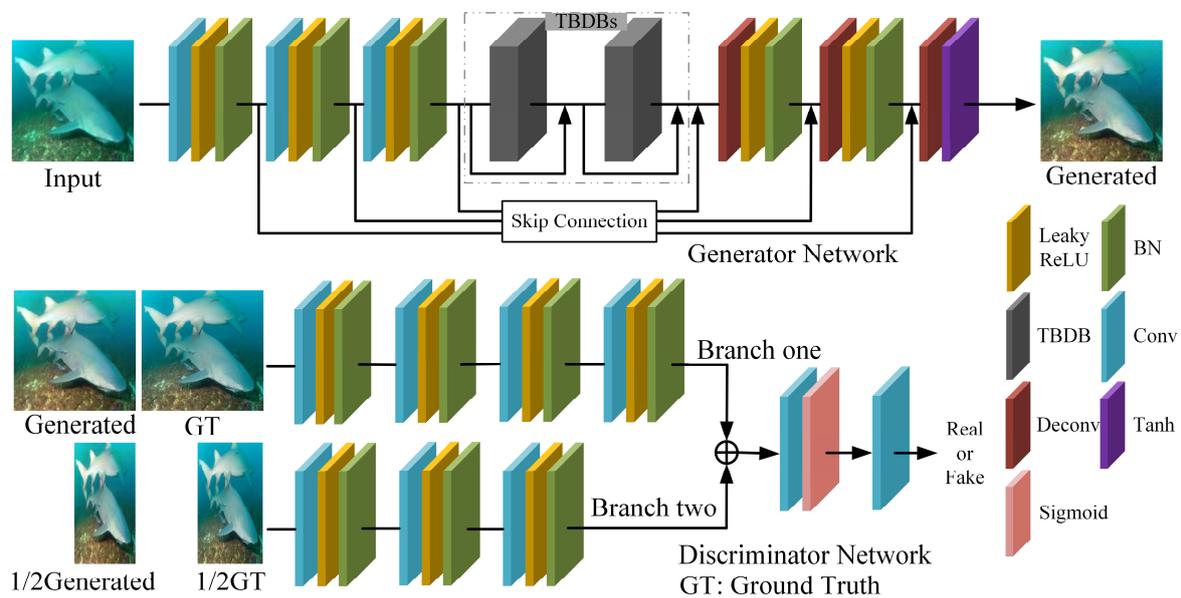


Figure 2. The network structure of TDGAN, including generator and discriminator networks. TBDBs, triple-branch dense blocks.

Table 1. The network structure and parameters.

Generator Network			Discriminator Network					
Layer	Kernel Size	Output Shape	Branch One			Branch Two		
			Layer	Kernel size	Output shape	Layer	Kernel size	Output shape
Conv, Leaky_ReLU, BN	[7,7,2]	$h/2 \times w/2 \times 64$	Conv, Leaky_ReLU, BN	[3,3,2]	$h/2 \times w/2 \times 32$	Conv, Leaky_ReLU, BN	[3,3,2]	$h/4 \times w/4 \times 64$
Conv, Leaky_ReLU, BN	[5,5,2]	$h/4 \times w/4 \times 128$	Conv, Leaky_ReLU, BN	[3,3,2]	$h/4 \times w/4 \times 64$	Conv, Leaky_ReLU, BN	[3,3,2]	$h/8 \times w/8 \times 128$
Conv, Leaky_ReLU, BN	[3,3,2]	$h/8 \times w/8 \times 256$	Conv, Leaky_ReLU, BN	[3,3,2]	$h/8 \times w/8 \times 128$	Conv, Leaky_ReLU, BN	[3,3,2]	$h/16 \times w/16 \times 256$
TBDBs	—	$h/8 \times w/8 \times 256$	Conv, Leaky_ReLU, BN	[3,3,2]	$h/16 \times w/16 \times 256$			
Deconv, Leaky_ReLU, BN	[3,3,2]	$h/4 \times w/4 \times 128$	Conv, Leaky_ReLU, BN	[3,3,2]	$h/16 \times w/16 \times 256$			
Deconv, Leaky_ReLU, BN	[5,5,2]	$h/2 \times w/2 \times 64$	Conv, Sigmoid	[3,3,2]	$h/16 \times w/16 \times 1$			

Inspired by the feature extraction modules [18,29–33], we develop a TBDB, as shown in Figure 3. The features of different scales cannot be fully utilized by simply connecting them together at the end of a block [29]. Adding TBDB to the network increases the depth of the network and improves the utilization of features. At the same time, the multi-scale structure design makes TBDB more sensitive to changes in feature map details. Three intermediate paths with different kernel sizes are designed to extract feature maps at different scales. A, B, and C, respectively, are the three different paths with different convolution (Conv) kernels (scales). The final 1×1 Conv is used to facilitate feature fusion and improve computational efficiency and keep the same number of input and output feature maps. X_{n-1} and X_n are the input and output image information, respectively. The variables T , P , F and O_1 are the network parameters.

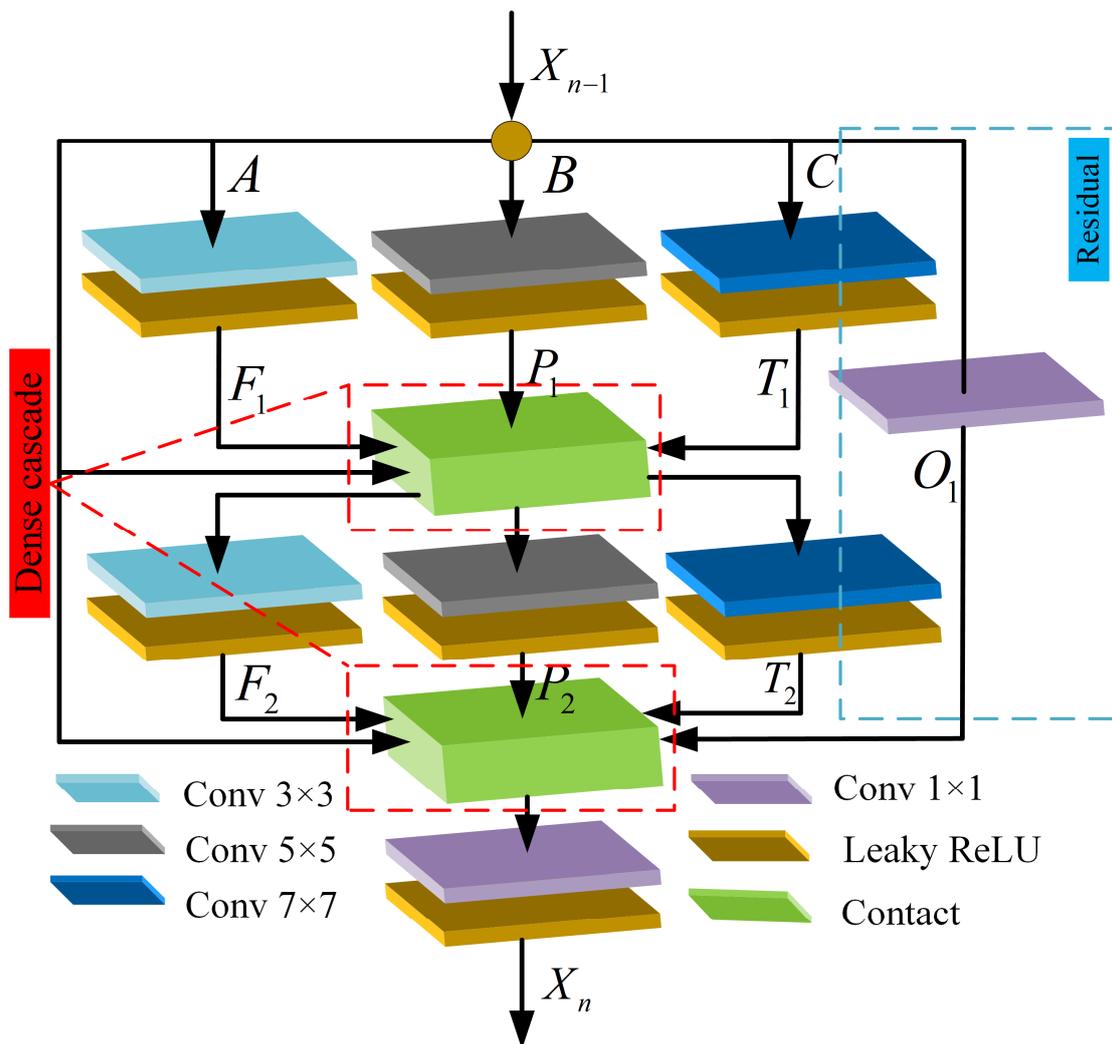


Figure 3. Structure of the TBDB. “Contact” denotes the dense concatenation.

In Figure 3, each layer in TBDB uses a Conv kernel with stride 1. Skip connections are added to TBDB, which widens the flow channels of image information and gradients. As shown in Figure 4, more than two TBDBs can improve the network performance in the experiments. However, this approach introduces too many parameters and increases training time. Therefore, after careful consideration, TDGAN uses two TBDBs.

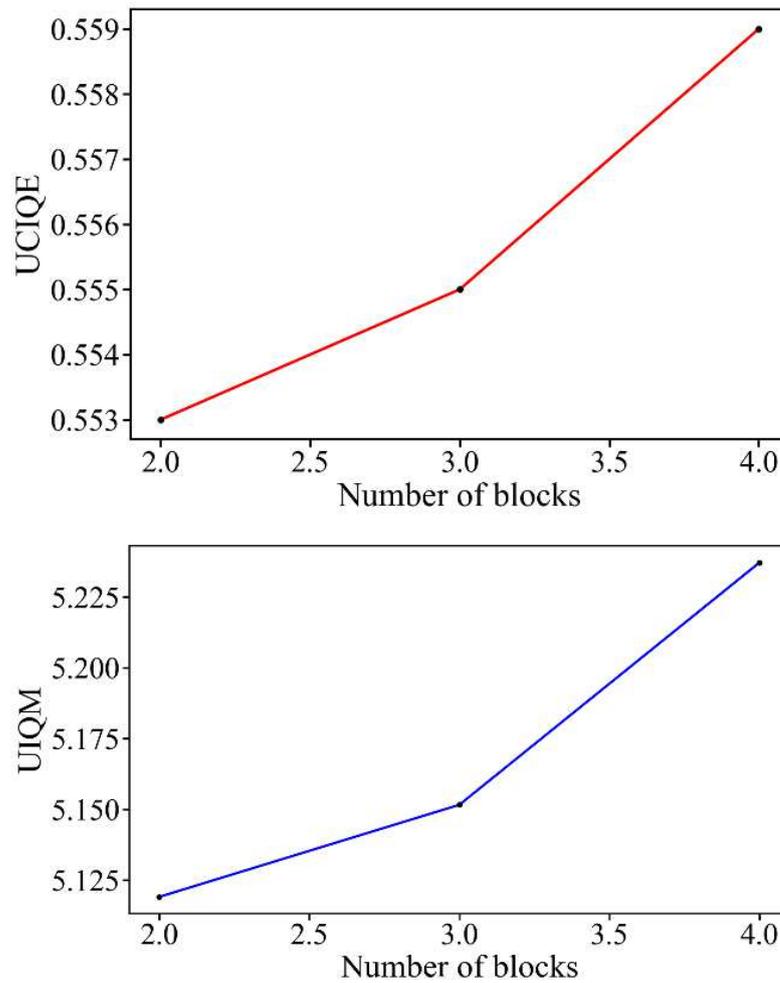


Figure 4. The results for different numbers of blocks. The results are obtained by averaging two hundred images.

3.2. Discriminator Network

Inspired by Markov Patch GAN architecture [16], a discriminator network is designed to distinguish patch-level information and capture high-frequency features, such as local texture and style [39]. The discriminator is designed as a dual-branch structure to guide the generator to pay more attention to global content and local details. The input image of the discriminator contains the images generated by the generator and the GT images. In addition, the input image size of the dual-branch discriminator is the original size and half of the original size, which can provide multi-resolution input. Here, all the Conv has a stride of 2. The discriminator network transforms a $256 \times 256 \times 6$ input image (GT images and generated images) into a $16 \times 16 \times 1$ output.

3.3. Loss Function

We design a loss function composed of conditional adversarial loss, L1 loss, and content loss to guide the generator to generate perceptually strong images. Generally, the conditional adversarial loss [37] is expressed as

$$L_{GAN} = \min_G \max_D V(D, G) = E_{X,Y}[\log D(Y)] + E_{X,Y}[\log(1 - D(X)G(X))] \tag{1}$$

where $D(Y)$ and Y denote the probability and real underwater image (original noisy image), respectively; $G(X)$ is the probability of the generator output; X stands for the image gen-

erated by the generator; $E[\cdot]$ denotes the mathematical expectation function; and $V(D, G)$ denotes cross entropy loss.

The L1 loss helps the generator G learn the global similarity of the image, avoid color distortion, and ensure that the gradient is stable [16]. The L1 loss is defined as

$$L_1(G) = E_{X,Y}[\|Y - G(X)\|_1]. \quad (2)$$

The content loss function is given by [40]

$$L_{con}(G) = E_{X,Y}[\|\Phi(Y) - \Phi(G(X))\|_2], \quad (3)$$

where $\Phi(\cdot)$ is the function extracted after the block5_conv2 layer of the pretrained VGG-19 network [40].

Consequently, the complete loss function of TDGAN is defined as

$$L_{TDGAN} = \min_G \max_D V(D, G) + \lambda_1 L_1(G) + \lambda_c L_{con}(G), \quad (4)$$

where λ_1 and λ_{con} signifies the weights for L_1 and content losses, respectively.

4. Experiments

4.1. Experiment Settings

(1) *Datasets.* The datasets used in the experiments include RUIE [41], UIEB [42], U45 [43], and EUVP [17]. The model training datasets include 14,266 images from EUVP and UIEB. The model testing data is composed of 1200 images from EUVP, UIEB, and U45. The data for application tests contains 460 images from RUIE and U45. The composition of the above training and testing datasets is randomly selected.

(2) *Training details.* During the training process, the training and test images are of a size $256 \times 256 \times 3$ and normalized into the range $[-1, 1]$. In TDGAN, the parameters are set as follows: $\lambda_1 = 0.7$ and $\lambda_{con} = 0.3$. The LReLU slope and learning rate are set as 0.2 and 0.0003, respectively. The batch size is set to 32. The optimizer is Adam [44]. When the generator is updated once, the discriminator is updated five times. TDGAN is trained on a GeForce RTX 3090 GPU for 200 epochs using the PyTorch framework.

(3) *Comparing methods.* We compare TDGAN with other methods on real underwater images. These comparison methods include Dark Channel Prior (UDCP) [45], Image Blur and Light Absorption (UIBLA) [46], Weakly Supervised Color Transfer (WSCT) [47], CycleGAN [14], Water-net [42] and FGAN [43], HLRP [48], and TACL [49]. The model parameters of the training processes of all comparison methods are set according to the reference.

(4) *Quantitative measurement index of underwater images.* Including non-reference image quality evaluation metrics UCIQE and UIQM and the full-reference image quality evaluation metrics PSNR and SSIM.

4.2. Qualitative Analysis

We conduct comparative experiments to verify the performance of TDGAN in different underwater backgrounds. The experimental results are shown in Figures 5–7. Note that the public datasets U45 and UIEB do not provide “ground truth” as a reference. Consequently, the results are not compared with the reference image. The EUVP dataset provides “ground truth” as a reference, so we compare the results with the reference image. Furthermore, the green underwater image in Figure 5 contains dark green and teal images. It should be noted that the real environment means the natural environment (natural scene without human intervention). Thus, the real image in this paper is the image collected in the natural environment.

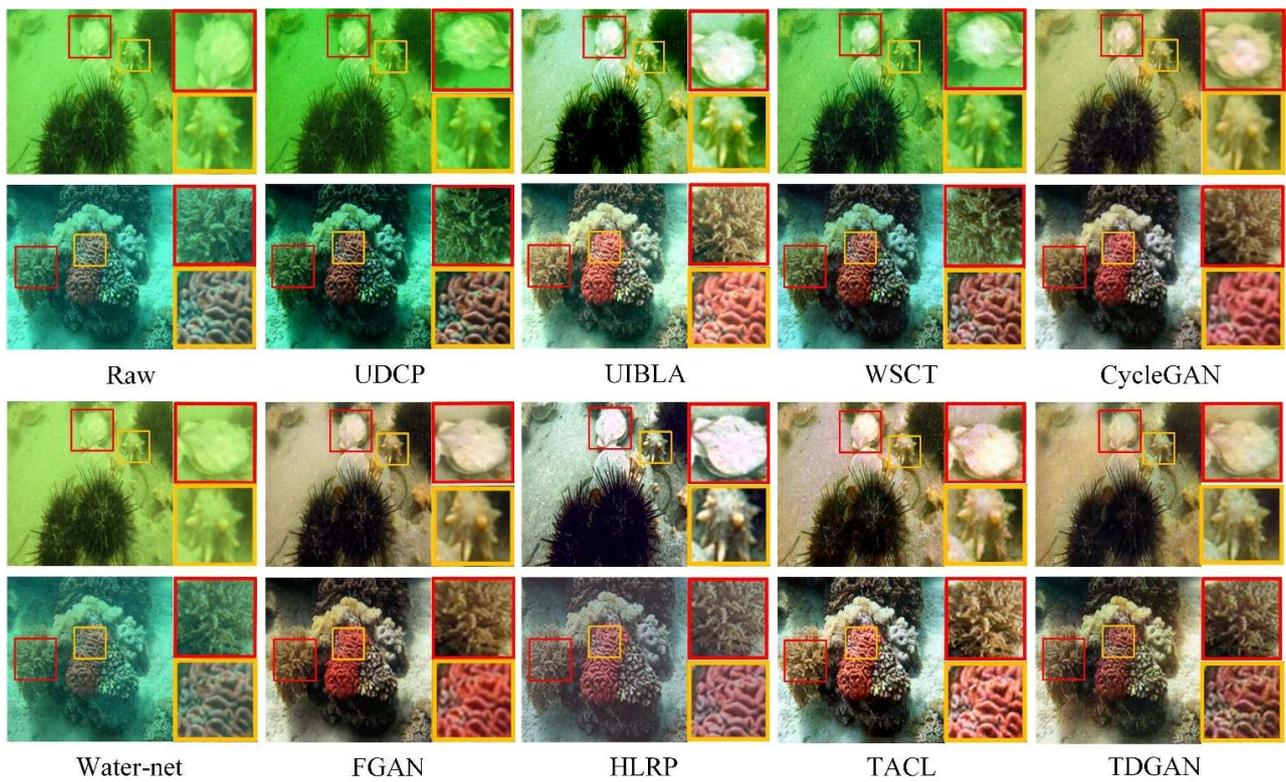


Figure 5. Visual comparison for samples from green underwater images. The Raw is from the dataset U45 [43].

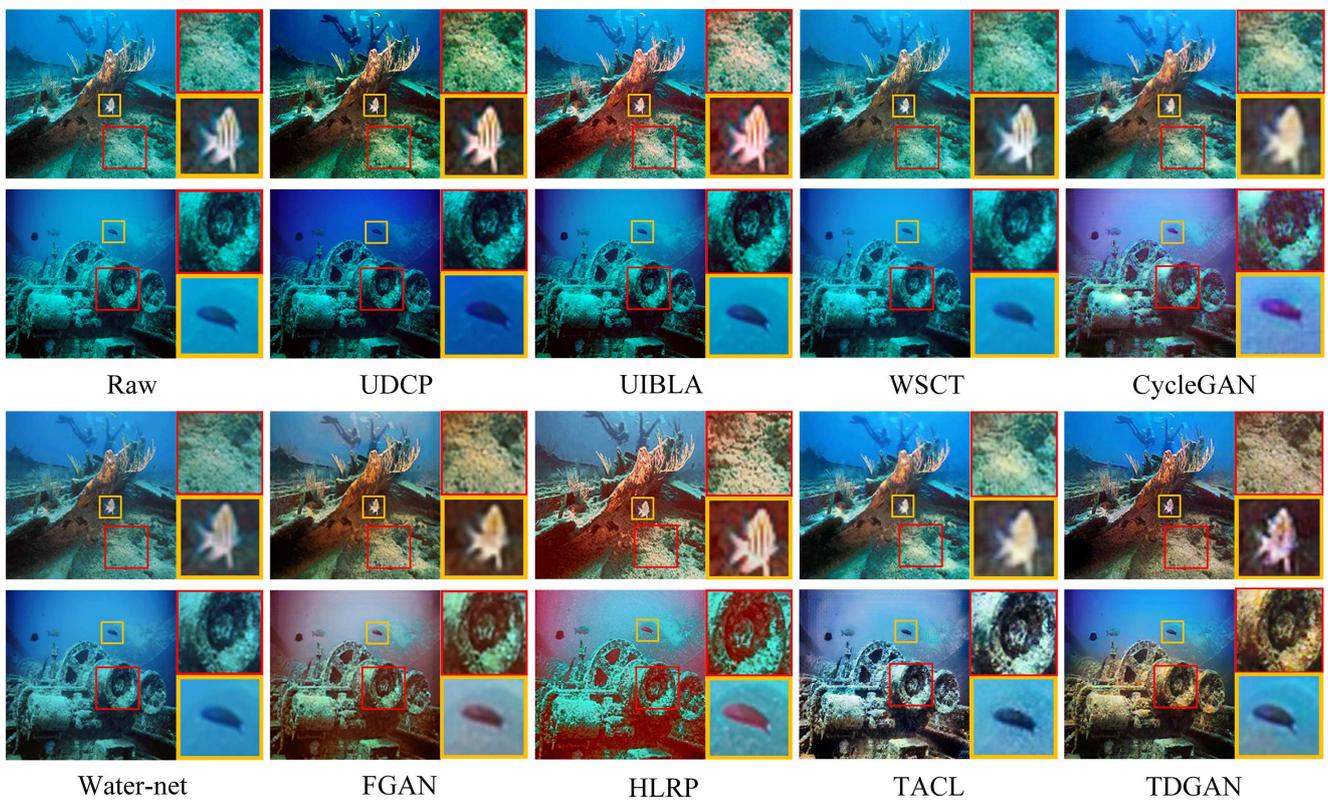


Figure 6. Visual comparison for samples from blue underwater images. The Raw is from the dataset UIEB [42].

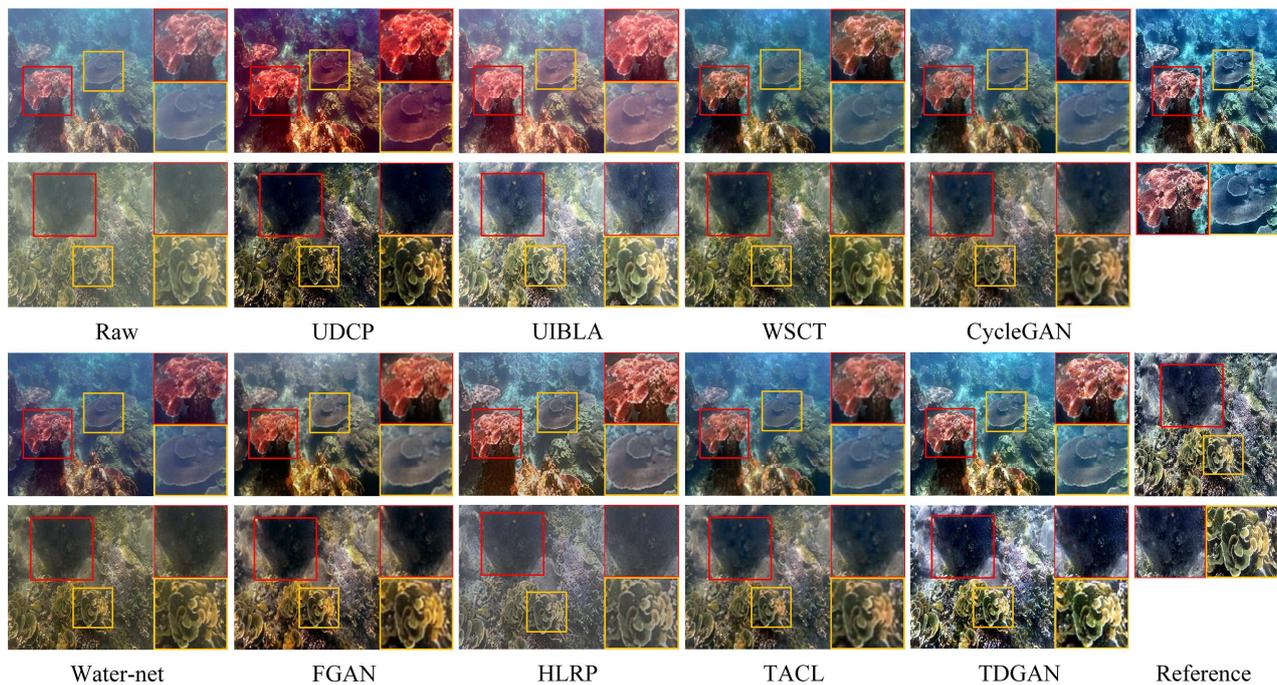


Figure 7. Visual comparison for samples from haze underwater images. The Raw is from the dataset EUVP [17].

Specifically, Figure 5 shows the comparison of experimental results of green underwater images. From the perspective of color enhancement, UDCP, UIBLA, WSCT, and Water-net cannot filter out the green background, while CycleGAN, FGAN, HLRP, TACL, and TDGAN can significantly remove the green environment background. Regarding the intuitive feeling of the restored image, the color of TDGAN is softer, and the brightness of the restored image is relatively higher. The underwater sand recovered by CycleGAN, HLRP, and TACL has a slightly dark background, which is inconsistent with the real environment. From the perspective of texture details, TACA, HLRP, and TDGAN also have excellent restoration capabilities. For example, in terms of the shape of sea cucumbers, these three methods can reconstruct the antennae outline of sea cucumbers. However, as shown in the reconstruction results of shells, HLRP and TACL have different degrees of artifacts, affecting the visual appearance of the images. Additionally, HLRP generates white artifacts that make the restored images look brighter. TACL produces a small number of red artifacts and some color distortions.

Figure 6 shows the enhancement results of underwater images with a blue background. The results show that UDCP, UIBLA, WSCT, and Water-net have no apparent effect on the images with the dark blue environment. FGAN and HLPR produce many red artifacts. TACL and TDGAN perform well, but the visual experience of TDGAN is full. Additionally, TDGAN has restored the tan fuselage of the sunken ship well, which is not achieved by other methods. Moreover, the fish body lines in TDGAN are clearer and more layered than in other comparison methods. Compared with other methods, TDGAN can effectively restore the color of underwater images, providing images with higher recognition, higher definition, and more prominent textures.

Figure 7 shows the comparison results of various methods in a light-colored underwater environment. In Figure 6, as the surface of the light-colored underwater image has a layer of blur-like water mist; thus, the performance of removing the water mist layer can intuitively reflect the enhancement performance of a method. From this point of view, UDCP, FGAN, and TDGAN have more advantages. However, UDCP is not clear enough, and the overall color is red. FGAN has clear layers, but the image details are not sufficient. For flower-like plants, TDGAN shows a stronger image detail restoration capability.

As the results shown in Figures 5–7, TDGAN is more effective than the other comparison methods for enhancing underwater images in green, blue, and light-colored environments. Specifically, TDGAN can effectively restore the color and texture of underwater images, recover more detailed features, and generate more visual images.

4.3. Quantitative Analysis

During the experiments, we chose different image quality evaluation indicators for different datasets. For the U45 and UIEB datasets, we adopt the no-reference image quality evaluation metrics UIQM and UCIQE to evaluate the performance of various methods quantitatively. Here, UIQM contains three components: UICM, UIConM, and UISM. For the EUVP dataset, we employ full-reference image quality evaluation metrics, including PSNR and SSIM. Table 2 lists the mean values of the metrics for 500 randomly selected images. As shown in Table 2, the assessment indicators of UDCP are relatively poor, and UIQM of UDCP is the worst. WSCT achieves the lowest UCIQE, probably because WSCT is unsuitable for underwater image enhancement with blue backgrounds, affecting its overall performance. TDGAN achieves the best UIConM, UIQM, UCIQE, PSNR, SSIM, and the second best UICM and UISM, proving the superiority of TDGAN.

Table 2. Quantitative evaluation of different underwater image datasets. The best and second best results are, respectively, highlighted and underlined.

Method	UIEB and U45 (500 Images) (Non-Reference)			EUVP (500 Images) (Full-Reference)			
	UICM	UIConM	UISM	UIQM	UCIQE	PSNR	SSIM
Raws	−14.913	0.629	7.058	4.017	0.501	19.017	0.704
UDCP	−64.298	0.827	7.027	3.220	0.543	19.401	<u>0.891</u>
UIBLA	−28.646	0.871	<u>7.289</u>	4.459	0.522	19.825	0.874
WSCT	−33.678	<u>0.896</u>	7.231	4.389	0.474	21.613	0.806
CycleGAN	−2.011	0.893	7.051	5.219	0.554	21.654	0.776
Water-net	−55.209	0.894	7.177	3.759	0.507	20.107	0.725
FGAN	5.770	0.895	7.098	<u>5.458</u>	0.566	<u>22.258</u>	0.832
HLRP	−2.265	0.867	7.443	<u>5.225</u>	<u>0.585</u>	21.563	0.796
TACL	0.259	0.916	6.962	5.337	0.528	21.737	0.837
TDGAN	<u>5.501</u>	0.925	7.203	5.588	0.571	25.434	0.911

4.4. Ablation Study

(1) *Ablation Study 1.* Two hundred different underwater images are tested in the ablation study. The ablation study methods are variants of TDGAN. The variants of TDGAN are introduced in Table 3. The residual and dense cascade blocks are shown in Figure 3. The experimental results of the ablation study are presented in Table 4. Here, the ✓ and ✗ denote the retention and removal, respectively. All the residual learning, dense cascade, and multi-scale operations can improve the evaluation metrics UICM, UIQM, and UCIQE. The original TDGAN performs better than its three variants for UICM, UIQM, and UCIQE. Moreover, an underwater image is used to show the variant differences of TDGAN visually. The results are shown in Figure 8, which are consistent with the results in Table 4. TDGAN obtains the best visual effect, and the image is more colorful than the others. It can also be seen that the visual effect turns bad when removing any of the operations mentioned above.

Table 3. The specific operation instructions.

Models	Residual	Dense Cascade	TDBBs
−RL	✗	✓	✓
−DC	✓	✗	✓
−Ms	✓	✓	✗
TDGAN	✓	✓	✓

Table 4. The comparison results of different modules in TDGAN.

Method	UICM	UIConM	UISM	UIQM	UCIQE
–RL	–20.112	0.890	7.002	4.683	0.562
–DC	–20.225	0.896	6.859	4.660	0.576
–Ms	–11.857	0.870	6.753	4.770	0.566
TDGAN	5.484	0.887	6.823	5.341	0.580

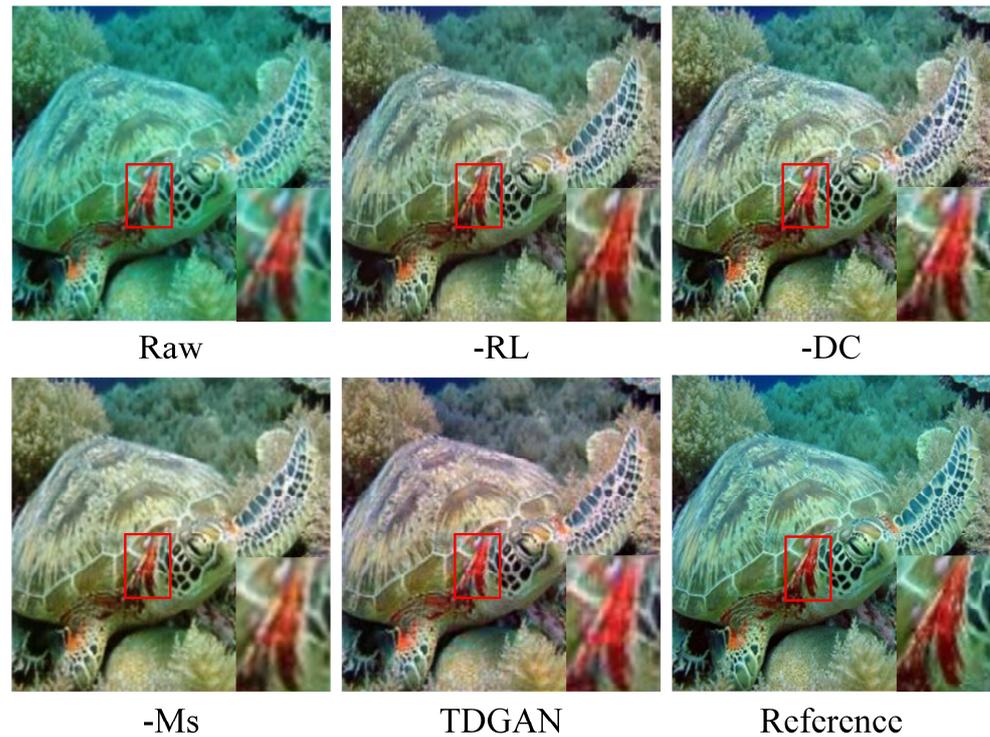


Figure 8. Visual comparison of different modules in TDGAN. The Raw and Reference are from EUVP [17].

(2) *Ablation Study 2.* The effect of kernel size in TBDB is also investigated experimentally. The architectures are shown in Table 5.

Table 5. The specific operation instruction 2.

Models	Kernel 7×7	Kernel 5×5	Kernel 3×3
–C	✗	✓	✓
–B	✓	✗	✓
–A	✓	✓	✗
TDGAN	✓	✓	✓

Three different architectures were tested on one hundred and sixteen different images. The results are shown in Table 6, where the best metric is marked in bold. The images tested under different scales (kernel size) have certain differences in UIQM and UCIQE values. –C and –B obtain a higher UIQM than –A; however, –C obtains a slightly lower UIQM than –A. Considering UIQM and UCIQE, TDGAN has more advantages. The intuitive results are displayed in Figure 9, where TDGAN has richer detailed features and colors. Specifically, in the absence of A or B, TDGAN cannot effectively remove the green background. In the absence of C, TDGAN is less expressive for red.

Table 6. The comparison results of different kernel sizes in TDGAN.

Method	UICM	UIConM	UISM	UIQM	UCIQE
−C	−95.251	0.635	6.466	1.495	0.551
−B	−79.637	0.533	6.459	1.566	0.544
−A	−97.816	0.639	6.674	1.500	0.527
TDGAN	−91.796	0.636	6.605	1.634	0.551

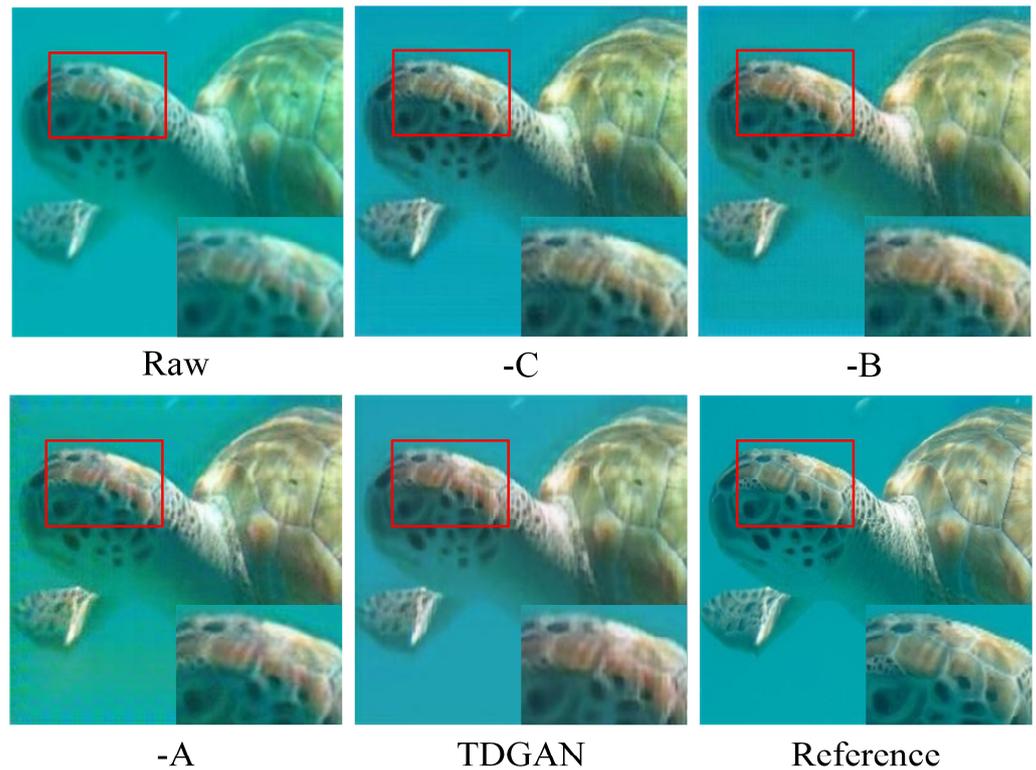


Figure 9. Visual comparison of different kernel sizes in TDGAN. The Raw and Reference are from EUVP [17].

The above two qualitative ablation experiments demonstrate that TDGAN is the optimal design.

4.5. Application Test

(1) *Application test 1.* To verify that TDGAN can be applied to the object detection fields, we use yolov5 [50] as a detection algorithm to detect underwater objects. The 3000 annotated images from the RUIE are adopted to train yolov5. The parameters of yolov5 are set as default, which is the same as the source code. Two hundred images from RUIE are used for the testing. The detection object is a starfish. Underwater images of different backgrounds are randomly selected for comparison of detection results, and the backgrounds include green, blue, and haze. The test results are shown in Figures 10–12. It can be seen that TDGAN removes the interference of the background color so that the detection object can be clearly displayed. Images enhanced by TDGAN can improve the probability and accuracy of object detection. This shows that TDGAN can be effectively applied in underwater target detection.

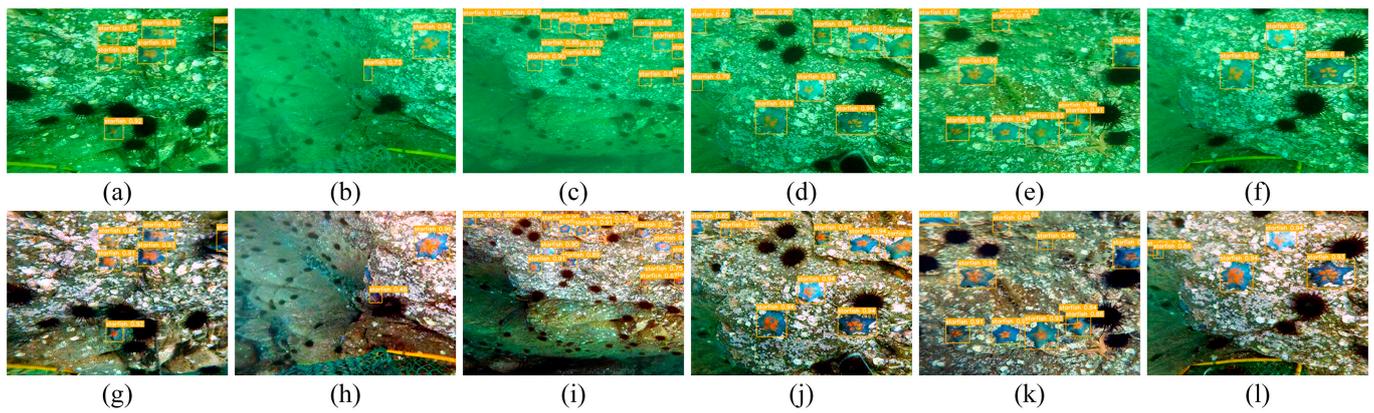


Figure 10. Comparison results of target detection in green environments. (a–f) are from RUIE [40], and (g–l) are from TDGAN.

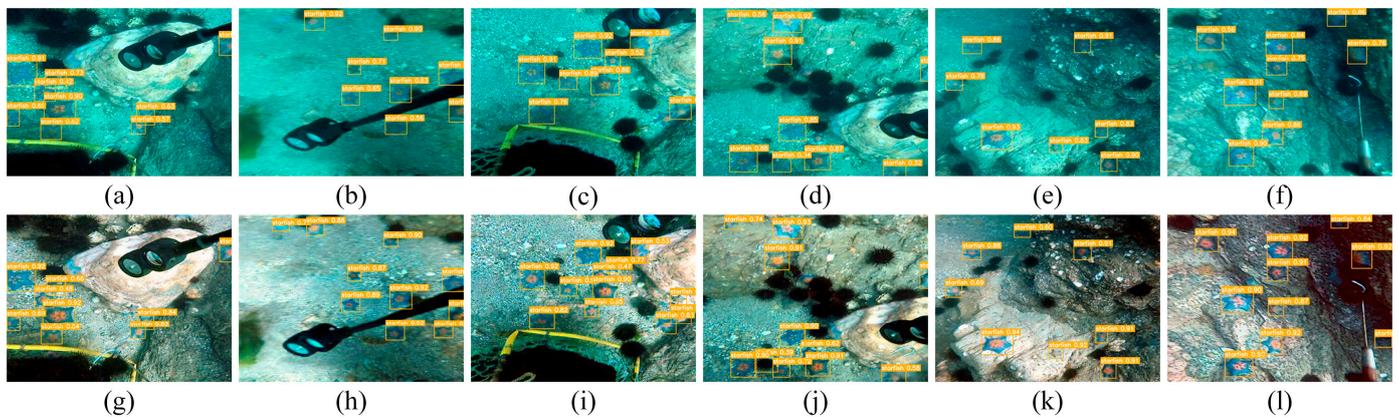


Figure 11. Comparison results of target detection in blue environments. (a–f) are from RUIE [40], and (g–l) are from TDGAN.

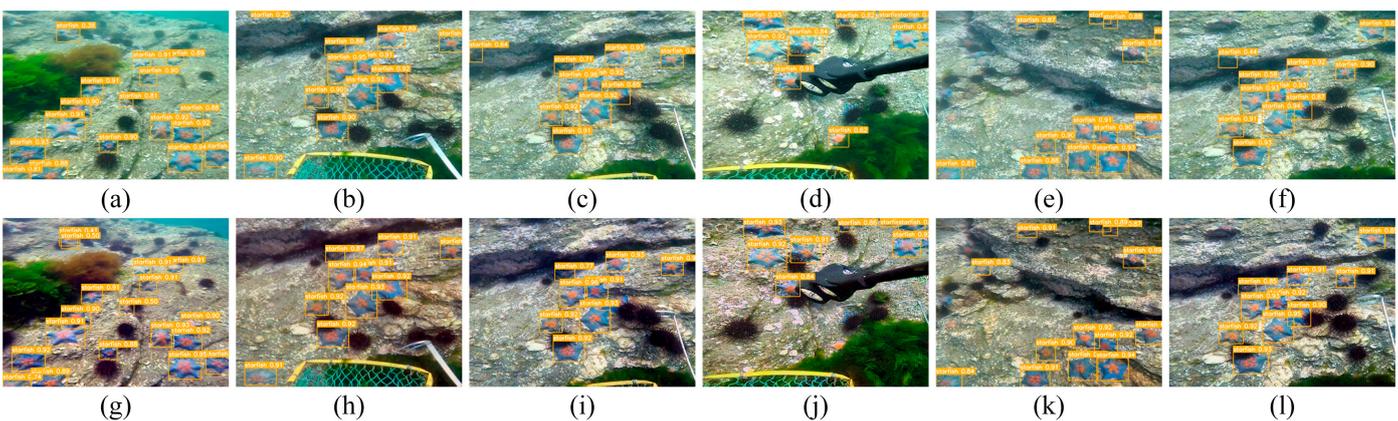


Figure 12. Comparison results of target detection in haze environments. (a–f) are from RUIE [40], and (g–l) are from TDGAN.

(2) *Application test 2.* Recently, underwater image enhancement methods have been widely applied in image segmentation and saliency detection, such as [51]. Here, a group of experiments was conducted to verify the effectiveness of TDGAN in image segmentation and saliency detection areas. A superpixel-based clustering algorithm [52] was first adopted for image segmentation. Then, an SR model [53] was used for image saliency detection. The image segmentation and saliency detection results are shown in Figure 13. It can be found from Figure 12 that the image segmentation boundary of TDGAN is more significant than that of other methods. Additionally, compared with other methods, TDGAN can

distinguish image layers more clearly, and the saliency map of TDGAN contains more complete contours and better boundaries. The above experiments show that TDGAN performs better in underwater image segmentation and saliency detection applications than the other methods (e.g., FGAN, CycleGAN, HLRP, TACL, etc.).

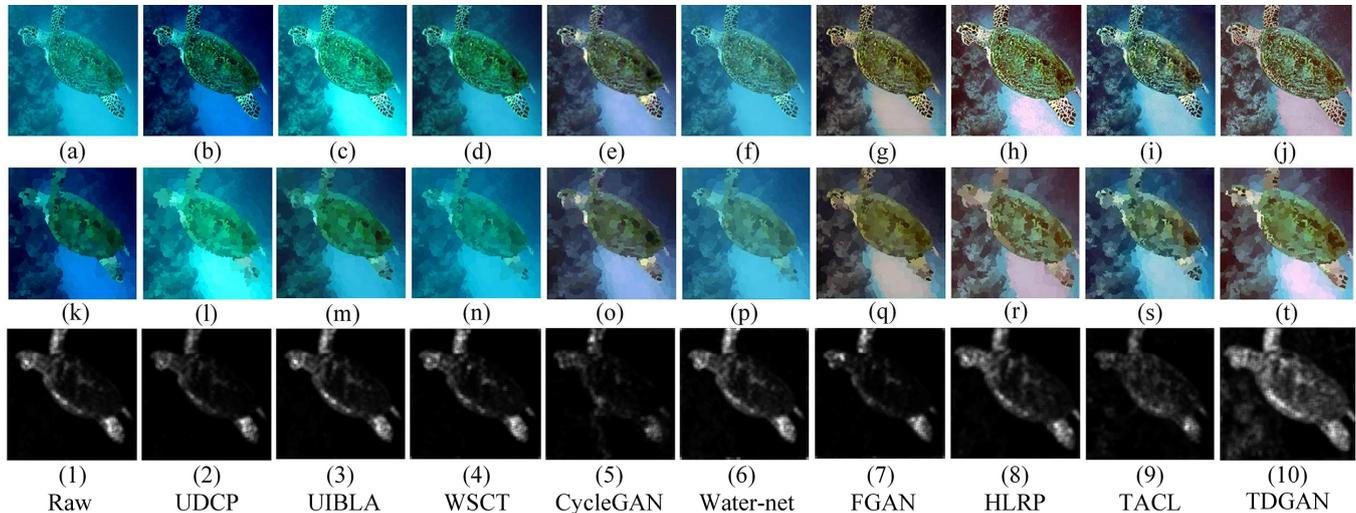


Figure 13. Application tests of underwater image segmentation and saliency detection with TDGAN. (b–j) are the images generated by the corresponding methods (in the bottom of the figure), (1–t) are the image segmentation results of the corresponding methods, and (2–10) are the image saliency detection results of the corresponding methods. The raw image (a) is from the dataset U45 [43]. (k,l) are the results of [52,53], respectively.

5. Discussion

The experimental results in Figures 5–13 and Table 2 demonstrate that TDGAN achieves competitive performance and has many advantages compared with the other methods. For instance, TDGAN can effectively restore the color and texture of underwater images, remove the water mist layer, generate more detailed features, and produce more visual images than the previous methods. These are attributed to the dual-branch discriminative network, which improves the network’s ability to perceive changes in image detail features, and the TBDB, which improves the information mining capability of the network. TDGAN has potential applications in underwater target detection and image segmentation. Experiments show that the proposed method can improve the accuracy of underwater target detection and obtain a complete boundary contour. However, TDGAN is restricted by the numerous training and large dataset, which limits its applicability. In the future, we will focus on optimizing TDGAN and small sample network design for underwater image and video enhancement.

6. Conclusions

We have presented a GAN-based underwater image enhancement network TDGAN to achieve high-quality underwater image restoration. We developed a TBDB to improve the feature-reusing ability of the generator network. TBDB helps the generator combine multi-scale features to fully use image information for training, thus generating high-quality images. A dual-branch discriminator network is designed to guide the generator to create images from global semantics to local details and enhances the local details of images. A loss function is developed to regulate the direction of network training. Qualitative and quantitative experiments are conducted to verify the effectiveness and superiority of TDGAN. Ablation studies are implemented to show the contribution of each component in TDGAN. The practicability of TDGAN is examined through application tests. Experimental results show that TDGAN can effectively enhance the underwater image quality, provide richer salient features, and improve the accuracy of underwater target detection. In future

work, the computational complexity of TDGAN will be optimized, and the performance of TDGAN will be further improved.

Author Contributions: Conceptualization, C.H. and S.L.; methodology, H.W.; software, P.Y.; validation, P.Y., H.W. and T.W.; formal analysis, S.L.; investigation, P.Y.; resources, P.Y.; data curation, P.Y.; writing—original draft preparation, P.Y.; writing—review and editing, H.W.; visualization, S.L.; supervision, T.W.; project administration, C.H.; funding acquisition, H.W. All authors have read and agreed to the published version of the manuscript.

Funding: National Natural Science Foundation of China (62173098, 62104047, U20A6003, U2001201), Guangdong Provincial Key Laboratory of Cyber-Physical System (2020B1212060069).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data underlying the results presented in this paper are not publicly available at this time but may be obtained from the authors upon reasonable request.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Kocak, D.M.; Dalgleish, F.R.; Caimi, F.M.; Schechner, Y.Y. A focus on recent developments and trends in underwater imaging. *Mar. Technol. Soc. J.* **2008**, *42*, 52. [[CrossRef](#)]
- Ghani, A.S.A.; Isa, N.A.M. Underwater image quality enhancement through integrated color model with Rayleigh distribution. *Appl. Soft Comput.* **2015**, *27*, 219–230. [[CrossRef](#)]
- Ronneberger, O.; Fischer, P.; Brox, T. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*; Springer: Berlin/Heidelberg, Germany, 2015; pp. 234–241.
- Abd-Alhamid, F.; Kent, M.; Bennett, C.; Calautit, J.; Wu, Y. Developing an innovative method for visual perception evaluation in a physical-based virtual environment. *Build. Environ.* **2019**, *162*, 106278. [[CrossRef](#)]
- Li, C.; Guo, J.; Cong, R.; Pang, Y.; Wang, B. Underwater image enhancement by dehazing with minimum information loss and histogram distribution prior. *IEEE Trans. Image Process.* **2016**, *25*, 5664–5677. [[CrossRef](#)]
- Chang, H.; Cheng, C.; Sung, C. Single underwater image restoration based on depth estimation and transmission compensation. *IEEE J. Oceanic Eng.* **2018**, *44*, 1130–1149. [[CrossRef](#)]
- Kar, A.; Dhara, S.K.; Sen, D.; Biswas, P.K. Zero-Shot Single Image Restoration Through Controlled Perturbation of Koschmieder’s Model. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition 2021, Nashville, TN, USA, 20–25 June 2021*; pp. 16205–16215.
- Marques, T.P.; Albu, A.B. L2uwe: A framework for the efficient enhancement of low-light underwater images using local contrast and multi-scale fusion. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops 2020, Seattle, WA, USA, 14–19 June 2020*; pp. 538–539.
- Galdran, A.; Pardo, D.; Picón, A.; Alvarez-Gila, A. Automatic red-channel underwater image restoration. *J. Vis. Commun. Image Represent.* **2015**, *26*, 132–145. [[CrossRef](#)]
- Li, J.; Skinner, K.A.; Eustice, R.M.; Johnson-Roberson, M. WaterGAN: Unsupervised generative network to enable real-time color correction of monocular underwater images. *IEEE Robot. Autom. Lett.* **2017**, *3*, 387–394. [[CrossRef](#)]
- Naik, A.; Swarnakar, A.; Mittal, K. Shallow-UWnet: Compressed model for underwater image enhancement. *arXiv* **2021**, arXiv:2101.02073.
- Ye, X.; Li, Z.; Sun, B.; Wang, Z.; Xu, R.; Li, H.; Fan, X. Deep joint depth estimation and color correction from monocular underwater images based on unsupervised adaptation networks. *IEEE Trans. Circ. Syst. Vid.* **2019**, *30*, 3995–4008. [[CrossRef](#)]
- Yang, H.; Huang, K.; Chen, W. Laffnet: A lightweight adaptive feature fusion network for underwater image enhancement. In *Proceedings of the 2021 IEEE International Conference on Robotics and Automation (ICRA), Xi’an, China, 30 May–5 June 2021*; pp. 685–692.
- Zhu, J.; Park, T.; Isola, P.; Efros, A.A. Unpaired image-to-image translation using cycle-consistent adversarial networks. In *Proceedings of the IEEE International Conference on Computer Vision 2017, Venice, Italy, 22–29 October 2017*; pp. 2223–2232.
- Fabbri, C.; Islam, M.J.; Sattar, J. Enhancing underwater imagery using generative adversarial networks. In *Proceedings of the 2018 IEEE International Conference on Robotics and Automation (ICRA), Brisbane, QLD, Australia, 21–25 May 2018*; pp. 7159–7165.
- Isola, P.; Zhu, J.; Zhou, T.; Efros, A.A. Image-to-image translation with conditional adversarial networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition 2017, Honolulu, HI, USA, 21–26 July 2017*; pp. 1125–1134.
- Islam, M.J.; Xia, Y.; Sattar, J. Fast Underwater Image Enhancement for Improved Visual Perception. *IEEE Robot. Autom. Lett.* **2020**, *5*, 3227–3234. [[CrossRef](#)]
- Guo, Y.; Li, H.; Zhuang, P. Underwater image enhancement using a multiscale dense generative adversarial network. *IEEE J. Ocean. Eng.* **2019**, *45*, 862–870. [[CrossRef](#)]

19. He, K.; Sun, J.; Tang, X. Single image haze removal using dark channel prior. *IEEE Trans. Pattern Anal. Mach. Intell.* **2010**, *33*, 2341–2353. [[PubMed](#)]
20. Pizer, S.M.; Amburn, E.P.; Austin, J.D.; Cromartie, R.; Geselowitz, A.; Greer, T.; Romeny, B.T.H.; Zimmerman, J.B.; Zuiderveld, K. Adaptive histogram equalization and its variations. *Comput. Vis. Graph. Image Process.* **1987**, *39*, 355–368. [[CrossRef](#)]
21. Pisano, E.D.; Zong, S.; Hemminger, B.M.; DeLuca, M.; Johnston, R.E.; Muller, K.; Braeuning, M.P.; Pizer, S.M. Contrast limited adaptive histogram equalization image processing to improve the detection of simulated spiculations in dense mammograms. *J. Digit. Imaging* **1998**, *11*, 193–200. [[CrossRef](#)] [[PubMed](#)]
22. Iqbal, K.; Odetayo, M.; James, A.; Salam, R.A.; Talib, A.Z.H. Enhancing the low quality images using unsupervised colour correction method. In Proceedings of the 2010 IEEE International Conference on Systems, Man and Cybernetics, Istanbul, Turkey, 10–13 October 2010; pp. 1703–1709.
23. Song, W.; Wang, Y.; Huang, D.; Tjondronegoro, D. A rapid scene depth estimation model based on underwater light attenuation prior for underwater image restoration. In *Advances in Multimedia Information Processing–PCM 2018: Proceedings of the 19th Pacific-Rim Conference on Multimedia, Hefei, China, 21–22 September 2018*; Springer: Berlin/Heidelberg, Germany, 2018; pp. 678–688.
24. Deng, X.; Wang, H.; Liu, X. Underwater image enhancement based on removing light source color and dehazing. *IEEE Access* **2019**, *7*, 114297–114309. [[CrossRef](#)]
25. Tao, Y.; Dong, L.; Xu, W. A novel two-step strategy based on white-balancing and fusion for underwater image enhancement. *IEEE Access* **2020**, *8*, 217651–217670. [[CrossRef](#)]
26. Zhang, W.; Zhuang, P.; Sun, H.; Li, G.; Kwong, S.; Li, C. Underwater image enhancement via minimal color loss and locally adaptive contrast enhancement. *IEEE Trans. Image Process.* **2022**, *31*, 3997–4010. [[CrossRef](#)]
27. Ke, K.; Zhang, C.; Wang, Y.; Zhang, Y.; Yao, B. Single underwater image restoration based on color correction and optimized transmission map estimation. *Meas. Sci. Technol.* **2023**, *34*, 55408. [[CrossRef](#)]
28. Zhang, W.; Wang, Y.; Li, C. Underwater image enhancement by attenuated color channel correction and detail preserved contrast enhancement. *IEEE J. Ocean. Eng.* **2022**, *47*, 718–735. [[CrossRef](#)]
29. Li, J.; Fang, F.; Mei, K.; Zhang, G. Multi-scale residual network for image super-resolution. In Proceedings of the European Conference on Computer Vision (ECCV), Munich, Germany, 8–14 September 2018; pp. 517–532.
30. Sun, B.; Mei, Y.; Yan, N.; Chen, Y. UMGAN: Underwater Image Enhancement Network for Unpaired Image-to-Image Translation. *J. Mar. Sci. Eng.* **2023**, *11*, 447. [[CrossRef](#)]
31. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition 2016, Las Vegas, NV, USA, 26 June–1 July 2016.
32. Cai, X.; Jiang, N.; Chen, W.; Hu, J.; Zhao, T. CURE-Net: A Cascaded Deep Network for Underwater Image Enhancement. *IEEE J. Ocean. Eng.* **2023**. [[CrossRef](#)]
33. Huang, G.; Liu, Z.; Laurens, V.; Weinberger, K.Q. Densely connected convolutional networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 4700–4708.
34. Yang, M.; Sowmya, A. An underwater color image quality evaluation metric. *IEEE Trans. Image Process.* **2015**, *24*, 6062–6071. [[CrossRef](#)]
35. Panetta, K.; Gao, C.; Agaian, S. Human-Visual-System-Inspired Underwater Image Quality Measures. *IEEE J. Ocean. Eng.* **2016**, *41*, 541–551. [[CrossRef](#)]
36. Hore, A.; Ziou, D. Image quality metrics: PSNR vs. SSIM. In Proceedings of the 2010 20th International Conference on Pattern Recognition, Istanbul, Turkey, 23–26 August 2010; pp. 2366–2369.
37. Wang, Z.; Bovik, A.C.; Sheikh, H.R.; Simoncelli, E.P. Image quality assessment: From error visibility to structural similarity. *IEEE Trans. Image Process.* **2004**, *13*, 600–612. [[CrossRef](#)]
38. Ioffe, S.; Normalization, C.S.B. Accelerating deep network training by reducing internal covariate shift. *arXiv* **2014**, arXiv:1502.03167.
39. Yi, Z.; Zhang, H.; Tan, P.; Gong, M. Dualgan: Unsupervised dual learning for image-to-image translation. In Proceedings of the IEEE International Conference on Computer Vision 2017, Venice, Italy, 22–29 October 2017; pp. 2849–2857.
40. Johnson, J.; Alahi, A.; Fei-Fei, L. Perceptual losses for real-time style transfer and super-resolution. In *Computer Vision–ECCV 2016: Proceedings of the 14th European Conference, Amsterdam, The Netherlands, 11–14 October 2016*; Springer: Berlin/Heidelberg, Germany, 2016; pp. 694–711.
41. Liu, R.; Fan, X.; Zhu, M.; Hou, M.; Luo, Z. Real-world underwater enhancement: Challenges, benchmarks, and solutions under natural light. *IEEE Trans. Circ. Syst. Vid.* **2020**, *30*, 4861–4875. [[CrossRef](#)]
42. Li, C.; Guo, C.; Ren, W.; Cong, R.; Hou, J.; Kwong, S.; Tao, D. An underwater image enhancement benchmark dataset and beyond. *IEEE Trans. Image Process.* **2019**, *29*, 4376–4389. [[CrossRef](#)]
43. Li, H.; Li, J.; Wang, W. A fusion adversarial underwater image enhancement network with a public test dataset. *arXiv* **2019**, arXiv:1906.06819.
44. Kingma, D.P.; Ba, J. Adam: A method for stochastic optimization. *arXiv* **2014**, arXiv:1412.6980.
45. Drews, P.L.; Nascimento, E.R.; Botelho, S.S.; Campos, M.F.M. Underwater depth estimation and image restoration based on single images. *IEEE Comput. Graph.* **2016**, *36*, 24–35. [[CrossRef](#)]
46. Peng, Y.; Cosman, P.C. Underwater image restoration based on image blurriness and light absorption. *IEEE Trans. Image Process.* **2017**, *26*, 1579–1594. [[CrossRef](#)]

47. Li, C.; Guo, J.; Guo, C. Emerging from water: Underwater image color correction based on weakly supervised color transfer. *IEEE Signal Process. Lett.* **2018**, *25*, 323–327. [[CrossRef](#)]
48. Zhuang, P.; Wu, J.; Porikli, F.; Li, C. Underwater image enhancement with hyper-laplacian reflectance priors. *IEEE Trans. Image Process.* **2022**, *31*, 5442–5455. [[CrossRef](#)]
49. Liu, R.; Jiang, Z.; Yang, S.; Fan, X. Twin adversarial contrastive learning for underwater image enhancement and beyond. *IEEE Trans. Image Process.* **2022**, *31*, 4922–4936. [[CrossRef](#)]
50. Bochkovskiy, A.; Wang, C.; Liao, H.M. Yolov4: Optimal speed and accuracy of object detection. *arXiv* **2020**, arXiv:2004.10934.
51. Zhuang, P.; Li, C.; Wu, J. Bayesian retinex underwater image enhancement. *Eng. Appl. Artif. Intell.* **2021**, *101*, 104171. [[CrossRef](#)]
52. Lei, T.; Jia, X.; Zhang, Y.; Liu, S.; Meng, H.; Nandi, A.K. Superpixel-based fast fuzzy C-means clustering for color image segmentation. *IEEE Trans. Fuzzy Syst.* **2018**, *27*, 1753–1766. [[CrossRef](#)]
53. Hou, X.; Zhang, L. Saliency detection: A spectral residual approach. In Proceedings of the 2007 IEEE Conference on Computer Vision and Pattern Recognition, Minneapolis, MN, USA, 17–22 June 2007; pp. 1–8.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.