


## Article

# An Underwater Image Enhancement Algorithm Based on Generative Adversarial Network and Natural Image Quality Evaluation Index

Kai Hu <sup>1,2,\*</sup> , Yanwen Zhang <sup>1,2</sup>, Chenghang Weng <sup>1,2</sup>, Pengsheng Wang <sup>1,2</sup>, Zhiliang Deng <sup>1,2</sup> and Yunping Liu <sup>1,2</sup> 

<sup>1</sup> School of Automation, Nanjing University of Information Science & Technology, Nanjing 210044, China; 20181223105@nuist.edu.cn (Y.Z.); 20201249145@nuist.edu.cn (C.W.); 201983240002@nuist.edu.cn (P.W.); 002858@nuist.edu.cn (Z.D.); liuyunping@nuist.edu.cn (Y.L.)

<sup>2</sup> Jiangsu Collaborative Innovation Center of Atmospheric Environment and Equipment Technology (CICAET), Nanjing University of Information Science & Technology, Nanjing 210044, China

\* Correspondence: 001600@nuist.edu.cn; Tel.: +86-137-7056-9871

**Abstract:** When underwater vehicles work, underwater images are often absorbed by light and scattered and diffused by floating objects, which leads to the degradation of underwater images. The generative adversarial network (GAN) is widely used in underwater image enhancement tasks because it can complete image-style conversions with high efficiency and high quality. Although the GAN converts low-quality underwater images into high-quality underwater images (truth images), the dataset of truth images also affects high-quality underwater images. However, an underwater truth image lacks underwater image enhancement, which leads to a poor effect of the generated image. Thus, this paper proposes to add the natural image quality evaluation (NIQE) index to the GAN to provide generated images with higher contrast and make them more in line with the perception of the human eye, and at the same time, grant generated images a better effect than the truth images set by the existing dataset. In this paper, several groups of experiments are compared, and through the subjective evaluation and objective evaluation indicators, it is verified that the enhanced image of this algorithm is better than the truth image set by the existing dataset.

**Keywords:** underwater image enhancement; generative adversarial network; natural image quality evaluation index; the true image



**Citation:** Hu, K.; Zhang, Y.; Weng, C.; Wang, P.; Deng, Z.; Liu, Y. An Underwater Image Enhancement Algorithm Based on Generative Adversarial Network and Natural Image Quality Evaluation Index. *J. Mar. Sci. Eng.* **2021**, *9*, 691. <https://doi.org/10.3390/jmse9070691>

Academic Editors: Michael O'Byrne, Vikram Pakrashi and Bidisha Ghosh

Received: 18 May 2021  
Accepted: 20 June 2021  
Published: 24 June 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 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

Marine resources are rich and have not been fully exploited. Compared with rivers, lakes, and other waters, the underwater environment of the ocean is more complex and hazardous, and the risk coefficient of artificial exploration and development is too high. Autonomous underwater vehicles (AUV) have become essential tools for human beings to explore the ocean [1–3], and visual images play an important role in exploring and perceiving the surrounding environment of an AUV. Due to the absorption and scattering of the water's body, however, underwater images have some problems, such as low contrast, image blur and color deviation, which affect the follow-up vision task of the underwater vehicle. Therefore, the acquisition of high-quality underwater images is of great significance for human exploration and the development of our understanding of the ocean. Existing underwater image enhancement algorithms are mainly divided into the physical model-based algorithm [4,5], the non-physical model-based algorithm [6,7], and the neural network-based algorithm [8,9].

The physical model-based algorithm is mainly represented by dark channel prior algorithm. In 2009, He proposed the dark channel prior (DCP) algorithm [10]. The DCP algorithm builds the mathematical model of the degradation process of the foggy image, and inverts the degradation process of the image by estimating the unknown parameters and combining them with the known parameters, so as to enhance the image quality.

Because the attenuation process of the underwater image is similar to that of an outdoor image, a large number of underwater image enhancement algorithms based on DCP have been proposed. The non-physical model-based algorithm is mainly represented by Retinex algorithm. In 1963, Edwin H. Land proposed the Retinex theory [6]. The Retinex algorithm can balance the three aspects of edge enhancement, dynamic range compression and color constancy in the process of image processing, but the problem of exposure will occur when the illumination intensity is high. In order to solve the problems existing in the Retinex algorithm, many scholars have improved it based on the Retinex algorithm. The Multi-Scale Retinex algorithm (MSR) [11] and Multi-Scale Retinex with Color Restoration (MSRCR) [12] have been successively proposed. The neural network-based algorithm is mainly represented by generative adversarial network. In 2014, Goodfellow proposed the concept of generative adversarial network (GAN) [13]. GAN uses the network to directly learn the mapping relationship between degraded underwater images and clear images, trains the model, and then restores images. After GAN theory was proposed, many scholars improved the GAN and applied it to different fields, and the deep learning theory began to be applied in the field of underwater image enhancement. Among them, the GAN based on the neural network algorithm [14–18] is the most widely used in underwater image enhancement.

Many scholars apply the GAN model to the field of underwater image enhancement, among which the representative algorithms are UWGAN [19], WaterGAN [20], UGAN [21], FUnIE-GAN [22], etc. In 2017, Li [20] proposed an underwater image generative adversarial network (WaterGAN). WaterGAN first uses atmospheric images and depth maps to synthesize underwater images. It then takes the synthesized underwater images as datasets and constructs an end-to-end color correction network to realize the real-time color correction of underwater images. In 2018, Fabbri [21] proposed an underwater GAN (UGAN), UGAN, which first uses unpaired clear underwater images and degraded underwater images to train CycleGAN, and then inputs clear underwater images into CycleGAN to generate corresponding degraded underwater images. Then, it uses pairs of underwater images as datasets for subsequent network training. Finally, L1 loss and gradient loss are added to the original loss of Wasserstein GAN to restore a degraded underwater image. In 2019, Guo [19] proposed a new underwater GAN (UWGAN), UWGAN, which adds a residual multi-scale dense block (RMSDB) to the generator; this is used to correct image color and restore image details. In 2020, fast underwater image enhancement GAN (FUnIE-GAN) [22] was proposed, which has good robustness and efficiency and can be applied to underwater vehicles. Generally speaking, in the underwater image enhancement algorithm, the GAN learns the mapping relationship from underwater degraded images to true underwater images. GAN needs a lot of training data for underwater image enhancement, including underwater degraded images and their corresponding truth images. However, existing truth images are limited by their unique imaging environment, and it is difficult to obtain an underwater truth image in the learning sample. At present, a truth sample depends on a variety of traditional algorithms. In selecting the best image from the captured images, or synthesizing an image by estimating the random parameters of an underwater image as the true image, there is a certain gap between a true image and a real underwater image. As a result, the image quality generated by the GAN is not ideal.

In the FUnIE-GAN algorithm, the contrast effect of the generated image is not very good. We analyzed that the low contrast may be caused by the following reasons:

- (1) There are some low contrast underwater images in the truth samples of the training set, so the training effect of the model is not good.
- (2) The function that can improve the image contrast is not added in the FUnIE-GAN algorithm, so the contrast of the generated image is not high.

Considering the above problems, a fast underwater image enhancement algorithm based on the NIQE index (FUnIE-GAN-NIQE) is proposed in this paper. The main contributions of this paper are as follows:

- (1) To solve the problem of low contrast images in truth datasets, this paper filters the images into truth images based on EUVP datasets to screen out truth images that meet the requirements.
- (2) To solve the problem of the low contrast of the generated image, this paper takes the NIQE as a part of the loss function of the generator in FUnIE-GAN and becomes its enhancement index.
- (3) To make the discriminant factors more diversified, this paper adds NIQE as FUnIE-GAN to the structure of the discriminator as part of the discriminator, which makes the resulting image more uniform in the color histogram distribution and more consistent with the perception of the human eye; this makes the generated image exceed the effect of the truth image set in the existing dataset.
- (4) In FUnIE-GAN-NIQE, there are four loss items in the loss function of the generator, including the adversarial loss function of the standard conditional GAN, L1 loss, content loss, and image quality loss. The weight loss of each part will affect the training result of the generator in the whole network; thus, this paper proposes to train 10 generators and 10 discriminators, traverse the weights of three parts (L1 loss, content loss, and image quality loss), and select the best generator among the 10 generators to generate the image. This method not only enhances the underwater image, but it can be applied to the enhancement of non-underwater images.

Section 1 provides a brief introduction to the development of underwater image enhancement and the idea of this paper. Section 2 briefly introduces relevant background knowledge. Section 3 introduces the main work of this paper. Section 4 describes experiments on the algorithm proposed in this paper and objectively analyzes its performance. Section 5 is a summary and conclusion of the findings of this paper.

## 2. Fast Underwater Image Enhancement Algorithm

In 2020, Islam [22] proposed a fast underwater image enhancement algorithm (FUnIE-GAN) to improve visual perception. The algorithm trains the model by establishing a multi-modal objective function. It also combines global similarity, local texture and style information, making the generated image closer to the true image and improving computing speed by reducing the model's parameters.

### 2.1. Network Architecture of Fast Underwater Image Enhancement Algorithm

#### 2.1.1. Generator Network Architecture

In the FUnIE-GAN algorithm, a generator is designed to improve U-Net architecture. The generator is the encoder–decoder architecture ( $e_1 \sim e_5, d_1 \sim d_5$ ). The specific framework model is shown in Figure 1. The input of the generator is at the left end of the network and is an image with a size of  $256 \times 256$  and three channels. The output of the generator is at the right end of the network and is an image with a size of  $256 \times 256$  and three channels. The left half of the generator is an encoder. Five orange rectangles in the left half of Figure 1 constitute the encoder, where each orange rectangle represents a multi-channel feature map. These five-layer multi-channel feature images are mainly obtained through  $4 \times 4$  convolution kernel convolution, nonlinear, and batch normalization operations. The main function of the encoder is continuous downsampling, reducing image size and extracting the shallow information of the image.

The right half of the generator is a decoder. Four mixed rectangles and one green rectangle in the right half of Figure 1 constitute a decoder, in which each rectangle in the decoder represents a multi-channel feature map. The multi-channel feature map in the decoder is obtained through  $4 \times 4$  convolutional kernel deconvolution, drop-out, and batch normalization operations. The role of the decoder is to expand the size of the image and extract the deep information of the image through continuous upsampling.

Each downsampling of the encoder halves the size of the feature map, and every time the decoder undergoes upsampling, the size of the feature map is doubled. The feature maps of the corresponding layers of the encoder and the decoder are stitched together, and

the corresponding relations between the multi-channel feature maps in the decoder and the multi-channel feature maps in the encoder are as follows:  $(e_1, d_5), (e_2, d_4), (e_3, d_2), (e_4, d_4)$ . The stitched feature map combines deep feature information with shallow feature information, and improves the performance of the generator. As shown in Figure 1, the whole black box represents the process of image processing, the horizontal lines between adjacent convolution layers represent convolution or deconvolution, among other operations, and the horizontal lines between non-adjacent convolution layers represent splicing operations. In the last five layers of the decoder, the orange rectangle is the downsampled multi-channel feature map in the encoder and the green rectangle is the multi-channel feature map obtained by the upsampling of the decoder.

The generator in the FUnIE-GAN algorithm is different from the traditional U-Net. The generator requires fewer parameters when building the model, and the encoder of the generator only needs to learn 256 feature maps with a size of  $8 \times 8$ .

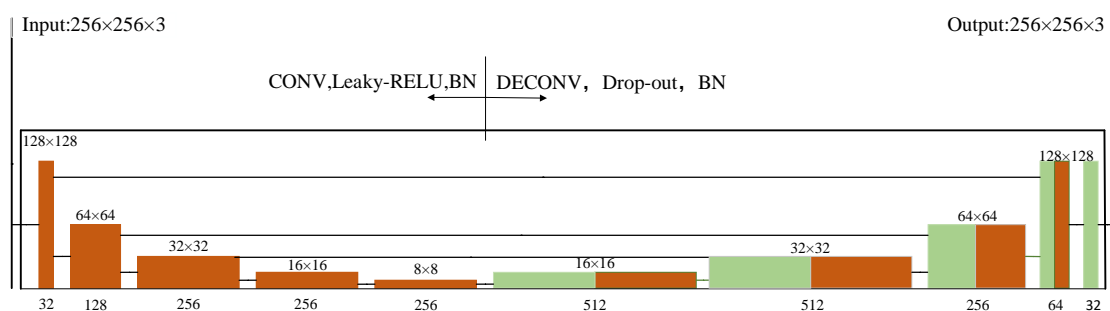


Figure 1. Generator network architecture.

### 2.1.2. Discriminator Network Architecture

The discriminator uses a Markovian PatchGAN. The Markovian PatchGAN is a full convolution network. The output of each layer in PatchGAN represents a receptive field in the original image, and can obtain high-frequency information related to local texture and style. The specific framework model is shown in Figure 2. The network model of the discriminator is the fully convolutional network, the input of the discriminator is at the left end of the whole network—it is an image with a size of  $256 \times 256$ —and the number of channels is three (truth image and generated image). The output of the discriminator is at the right end of the network; it is a  $16 \times 16$  vector. The yellow rectangle in Figure 2 is a multi-channel feature map. The image is input into the discriminator and after the convolution of nonlinear and batch normalization operations, with a convolution kernel size of  $3 \times 3$  and a convolution kernel stride of two for five times, a  $16 \times 16$  vector is obtained. The discriminator judges the authenticity of each patch. The results of all patches of an image are averaged as the final discriminator output to judge the probability of whether the input image is a true-value image. The entire large black box in Figure 2 represents the image processing process, and the horizontal lines between adjacent convolution layers represent convolution, non-linearity and batch normalization operations.

### 2.2. Fast Underwater Image Enhancement Loss Function

The FUnIE-GAN algorithm trains paired datasets, and the objective function is mainly composed of the following three parts:

- (1) In the FUnIE-GAN algorithm, the  $L_{cGAN}$  function is the adversarial loss function of the standard conditional GAN:

$$L_{cGAN}(G, D) = E_{X,Y}[\log D(X, Y)] + E_{X,Y}[\log(1 - D(X, G(X, Z)))] \quad (1)$$

In Equation (1),  $X$  and  $Y$  represent lossy image and truth image, respectively,  $Z$  is random noise,  $G$  is generator,  $D$  is discriminator and  $E$  is mathematical expectation. The lossy image  $X$  is input into the generator as a condition; together with random noise

$Z$ , the input of the generator is  $\{X, Z\}$  and the output of the generator is  $\{G(X, Z)\}$ . The discriminator needs to distinguish a pair of  $\{X, Y\}, \{X, G(X, Z)\}$  images so that it can not only generate real images, but also ensure that the generated images match the input images.

- (2) The  $L_1$  function is a global similarity loss that aims to enable the generator to sample from the global similarity space. The generated image is closer to the true image in the global appearance. The calculation formula is shown in Equation (2):

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

where the symbol  $\|\cdot\|_1$  denotes the one norm in functional mathematics.

- (3) The  $L_{con}$  function is a loss of content to encourage the generator to generate images with advanced features similar to true images. In the FUnIE-GAN algorithm, the image content function  $\varphi$  is defined as the extraction of high-level feature functions. The advanced features extracted in this paper are extracted from the block5-conv2 layer of the pre-training network VGG-19. The formula for calculating content loss is shown in Equation (3):

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

where the symbol  $\|\cdot\|_2$  denotes the L2 norm in functional mathematics. The objective function of the classical FUnIE-GAN algorithm for training paired datasets is shown in Equation (4):

$$G1^* = \underset{G}{\operatorname{argmin}}(\max_D L_{cGAN}(G, D) + \lambda_1 L_1(G) + \lambda_c L_{con}(G)) \quad (4)$$

where  $\lambda_1 = 0.7, \lambda_c = 0.3$ .

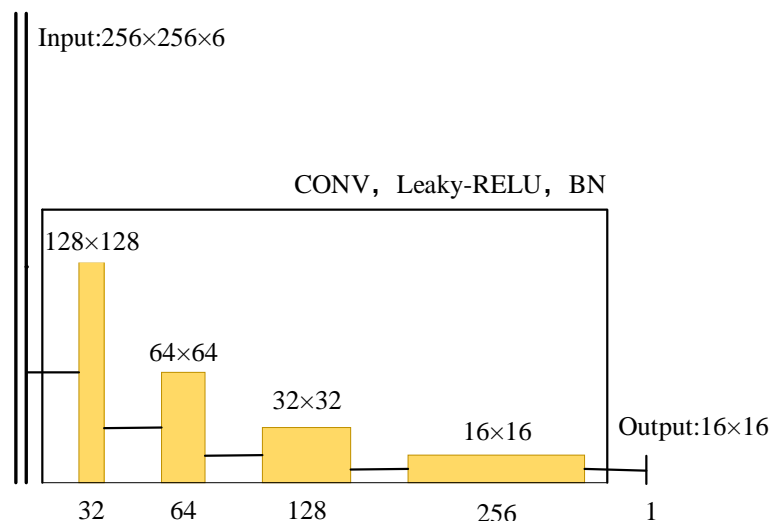


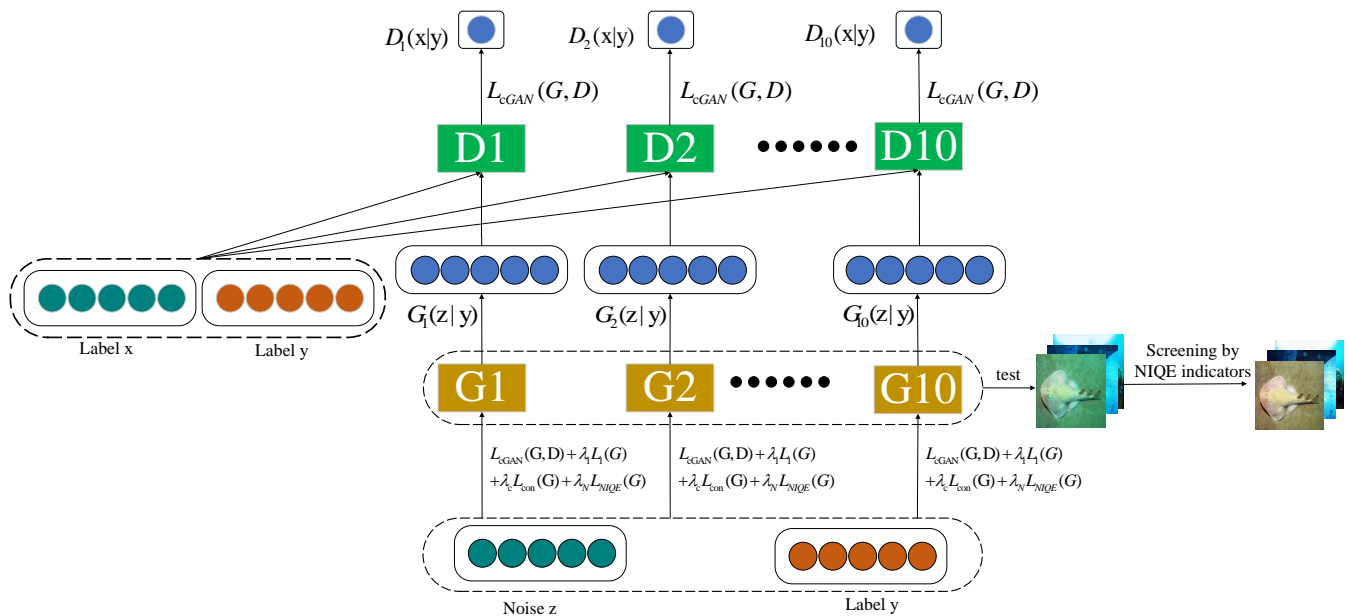
Figure 2. Discriminator network architecture.

### 3. Algorithm of Fast Underwater Image Enhancement Algorithm Based on NIQE Index

This paper proposes a fast underwater image enhancement algorithm (FUnIE-GAN-NIQE) based on the NIQE index. The framework structure of FUnIE-GAN-NIQE algorithm is shown in Figure 3. By building the generator and discriminator network framework, the FUnIE-GAN-NIQE algorithm learns the mapping relationship from the underwater lossy image to the underwater truth image, and trains the weight parameters of the generator network. By inputting the underwater damaged image into the trained generator model, we can output a clear underwater image, i.e., the enhanced underwater image. The pseudo-code of the algorithm flow is shown in Algorithm 1.

**Algorithm 1** FUnIE-GAN-NIQE algorithm flow.

- 1: Dataset filtering
- 2: **for**  $i = 10$  **do**
- 3:   Initialize a discriminator  $D$  and generator  $G$ ;
- 4:   Cyclic iteration training process:
- 5:   The process of training discriminant  $D$ :
- 6:     Input lossy images and noises into the generator to generate truth images;
- 7:     The image generated by the generator and the truth-value image are put into the discriminator. The loss function of the discriminator is used to train the discriminator. The discriminator can distinguish the image generated by the generator and the truth-value image correctly, and the parameters of the discriminator are updated;
- 8:   The process of training discriminant  $G$ :
- 9:     By training the generator with its loss function, the generated image can make the discriminator indistinguishable, and the generator parameters can be updated;
- 10: **end for**

**Figure 3.** Frame structure diagram of FUnIE-GAN-NIQE algorithm.**3.1. Screening of Datasets**

The training of the FUnIE-GAN requires pairs of datasets;  $I^C$  is the underwater truth image and  $I^D$  is the corresponding underwater lossy image, as it is difficult to obtain  $I^C$  and  $I^D$  at the same time. FUnIE-GAN adopts the method in UGAN, which uses CycleGAN [23] to generate  $I^C$  into  $I^D$  to generate paired datasets. By limiting the depth, lighting conditions, camera model and physical position in an underwater environment, an image with little or no distortion can be obtained. FUnIE-GAN selects the above image (the image with little or no distortion) as the true image by subjective vision. Because different people have different perceptions of color, the true image selected by subjective vision is not objective. The artificial selection of a true image that does not meet the requirements will affect the training of the network model.

To reduce the influence of artificial subjective factors and make truth images not only meet the requirements of artificial subjective vision but also meet the requirements of objective indicators, this paper selected the sub-dataset, “Underwater Dark”, of the EUVP dataset. There are 5550 training pairs in “Underwater ImageNet”, and there are 570 images in the verification set. In this paper, the NIQE index is used as the optimization index of underwater image quality. Hu [24] and others have proven that the NIQE index has the characteristics of simple calculation, high accuracy, and that it is in line with the visual

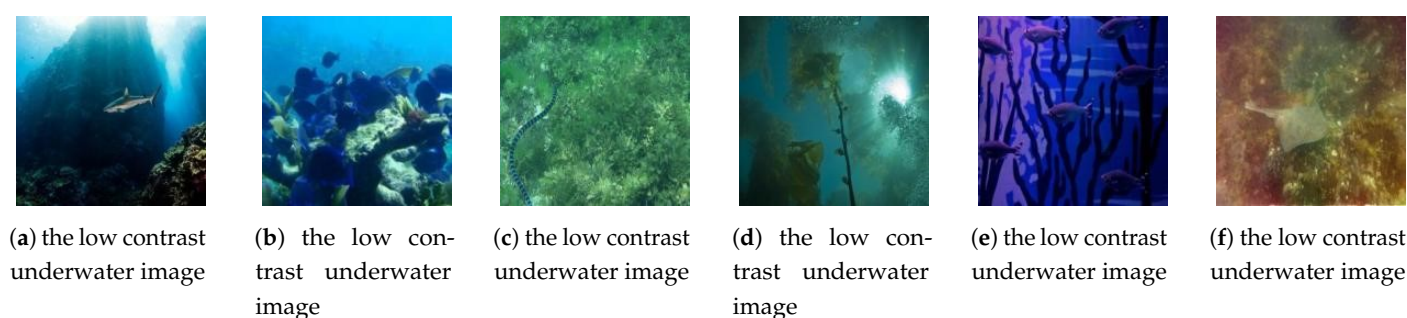


characteristics of the human eye. The higher the NIQE value, the worse the image quality; on the contrary, the better the image quality, the lower the NIQE value. In this study, the true image in the “Underwater Dark” dataset was screened by the NIQE index. The lossy image dataset was restored by a variety of mainstream underwater image enhancement methods. The NIQE value of the enhanced image was calculated, and the NIQE value of the corresponding image was processed. This paper takes the processed NIQE value as the standard NIQE value of the truth image, and the standard NIQE value will be compared with the NIQE value of the original truth image. In this method, the NIQE values of two corresponding truth images were selected and compared by the same measurement standards to filter out the truth images in the dataset that do not meet the standard. The specific filtering process is as follows:

- (1) Four mainstream underwater image enhancement methods (CycleGAN [23], FUnIE-GAN [22], FUnIE-GAN-Up [22], UGAN [21]) were used to enhance the lossy images in the “Underwater Dark” dataset, and four enhanced image sets were obtained. The four image sets were named  $I^{CycleGAN}$ ,  $I^{FUnIE-GAN}$ ,  $I^{FUnIE-GAN-Up}$  and  $I^{UGAN}$ , respectively. Then, each image in four datasets was evaluated by the NIQE index. Because the existing four mainstream methods could not effectively restore all images in specific applications to remove random errors and improve the accuracy of datasets, this paper used the method of error analysis to choose the four values of each image. The main implementation step was to remove the maximum value and the minimum value, and then take the average of the two values to obtain the NIQE value of the damaged image after effective enhancement.
- (2) The average value of NIQE obtained by the mainstream method in 4 was compared with the NIQE value of the corresponding truth image, and the truth image was screened by the error analysis method. The specific calculation formula is as follows:

$$\frac{|NIQE_{ave} - NIQE_{true}|}{NIQE_{ave}} \leq 20\% \quad (5)$$

Among them,  $NIQE_{ave}$  is the average value of NIQE obtained by four mainstream methods and  $NIQE_{true}$  is the NIQE value of the corresponding truth image. The truth images whose error fluctuation range was greater than 20% were eliminated. Through the screening of truth images, a total of 153 truth images, which did not meet the requirements of the NIQE index, were screened out. Figure 4 shows the truth images selected in “Underwater Dark” that did not meet the requirements.



**Figure 4.** Failure truth sample.

### 3.2. Loss Function of the Generator

In FUnIE-GAN, the discriminator discriminates that the image generated by the generator and the truth sample is of a true probability, which also requires the truth sample. There is still a certain gap between the current truth image and the real underwater lossless image, so the image generated in FUnIE-GAN cannot break through the shackles of the truth image. This paper innovates and adapts the loss function of the generator to solve this problem. Hu [24] has proven that the natural image quality evaluation method [25]

(NIQE) can effectively reflect the changing trend of image quality. This paper uses the NIQE index as a part of the loss function of FUnIE-GAN to enhance the image, which makes the histogram distribution of the color more uniform and more consistent to perception by the human eye, and makes the generated image exceed the effect of the truth image set by the existing dataset.

The color enhancement Equation (6) in this paper is as follows:

$$L_{NIQE}(G) = NIQE(G(X, Z)) \quad (6)$$

### 3.3. Structure of Discriminator

In FUnIE-GAN, the discriminator discriminates the probability that the image generated by the generator and the truth sample is true, which also requires the truth sample and the discriminator. The training of the discriminator cannot break through the shackles of the truth image. This paper posits a method for adding the NIQE index to the discriminator, which diversifies the factors of the discriminator in FUnIE-GAN. In this study, the structure of the original discriminator was improved, the image generated by the generator was divided into 16 batches, the size of each batch was  $16 \times 16$ , each batch was evaluated by the NIQE index, and the output of the discriminator improves the sum of the probability of generating the image to be true in relation to the NIQE index. The output of the original discriminant matrix is  $16 \times 16$  matrix  $P$ , and the output of the calculated NIQE for the image is  $16 \times 16$  matrix  $N$ . Then, the output of the improved discriminator is  $D = \lambda_P * P + \lambda_N * \frac{1}{N}$ . Where  $\lambda_P = 0.8$ ,  $\lambda_N = 0.2$ . The frame structure of the improved discriminator is shown in Figure 5.

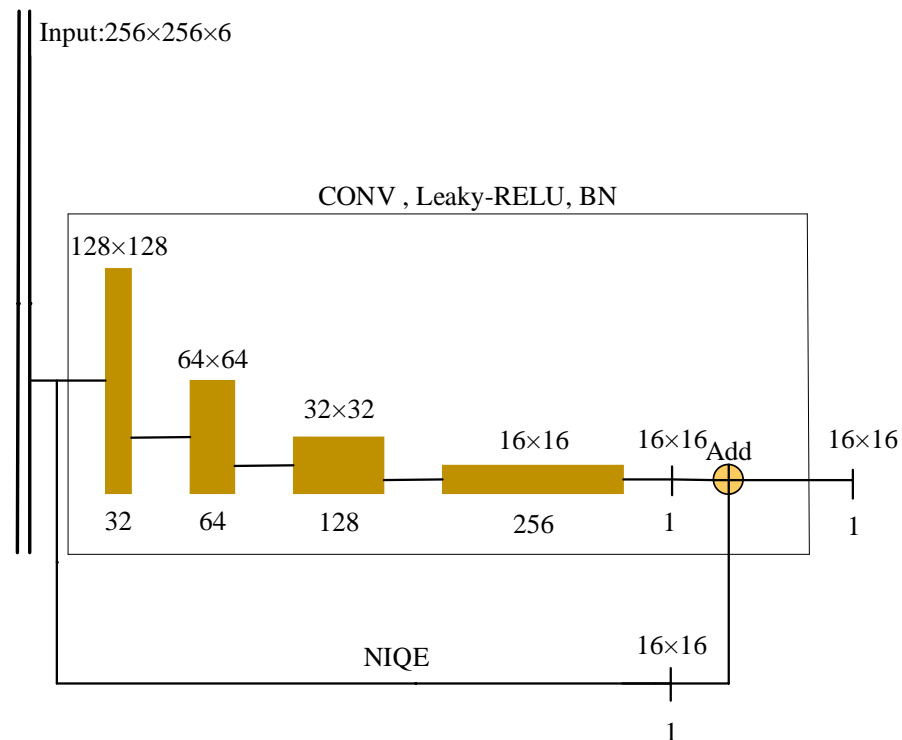


Figure 5. Frame structure diagram of the improved generator.

### 3.4. Network Structure

In FUnIE-GAN-NIQE, the loss of the generator is mainly composed of four parts, including the adversarial loss function of standard conditional GAN,  $L_1$  loss, content loss and image quality loss. The calculation formula is as follows:

$$G^* = \underset{G}{\operatorname{argmin}} (\max_D L_{cGAN}(G, D) + \lambda_1 L_1(G) + \lambda_c L_{con}(G) + \lambda_N L_{NIQE}(G)) \quad (7)$$



The weight loss of each part will affect the training result of the generator in the whole network. As such, this paper proposes training 10 generators, traverse the weights of three parts ( $L_1$  loss, content loss, image quality loss) and select the best generator among the 10 generators to generate the image. In this paper, the ratio relationship between  $\lambda_1$  and  $\lambda_c$  in the original FUnIE-GaN is preserved; i.e.,  $\lambda_1 : \lambda_c = 7 : 3$ . Ten generators are trained by controlling  $\lambda_1 + \lambda_c + \lambda_N = 10$  in the FUnIE-GAN-NIQE algorithm in this paper, and the numerical settings of  $\lambda_1$ ,  $\lambda_c$  and  $\lambda_N$  of the 10 generators are shown in Table 1. A large number of experiments show that the image generated by the FUnIE-GAN-NIQE algorithm is relatively good when  $\lambda_1 : \lambda_c : \lambda_N = 1.4 : 0.6 : 8$ .

**Table 1.** Parameter setting of  $\lambda_1$ ,  $\lambda_c$ ,  $\lambda_N$ .

Generator Number	$\lambda_1$	$\lambda_c$	$\lambda_N$
Generator 1	6.3	2.7	1
Generator 2	5.6	2.4	2
Generator 3	4.9	2.1	3
Generator 4	4.2	1.8	4
Generator 5	3.5	1.5	5
Generator 6	2.8	1.2	6
Generator 7	2.1	0.9	7
Generator 8	1.4	0.6	8
Generator 9	0.7	0.3	9
Generator 10	0.07	0.03	9.9

## 4. Analysis of Experimental Results

### 4.1. The Experimental Configuration

The algorithm is based on the Linux system. The CPU of the experimental environment was an Intel i7-9700 K with 64 G memory. The graphics card was an Nvidia Geforce GTX2080Ti and the software environment was Python 3.5. This algorithm used the Pytorch open source framework. In this paper, experiments were carried out on the open dataset EUVP. The batch size was 1, using the Adam optimizer,  $\beta_1$  is 0.5,  $\beta_2$  is 0.99. The learning rate of the generator and discriminator was 0.0003.

### 4.2. Dataset Setting

The experimental data in this paper are from the EUVP dataset. Islam proposed the EVUP dataset in 2020, which contains the paired and unpaired image samples of lossful images and truth images to train underwater image enhancement models. The EUVP dataset consists of three paired datasets and one unpaired dataset. Islam used seven different cameras to capture images in different locations and indifferent visibility conditions. Additionally, some of the images were taken from publicly available YouTube videos. The dataset contains images of different scenes of water environments with different turbidities and illumination degrees, making it diverse.

The method for obtaining EUVP unpaired datasets was as follows: the acquired images were divided into damaged images and truth images from a visual point of view (color, contrast, sharpness) by six people. Figure 6 shows some EUVP unpaired datasets. In Figure 6, the upper row represents the damaged images and the lower row represents the truth images.

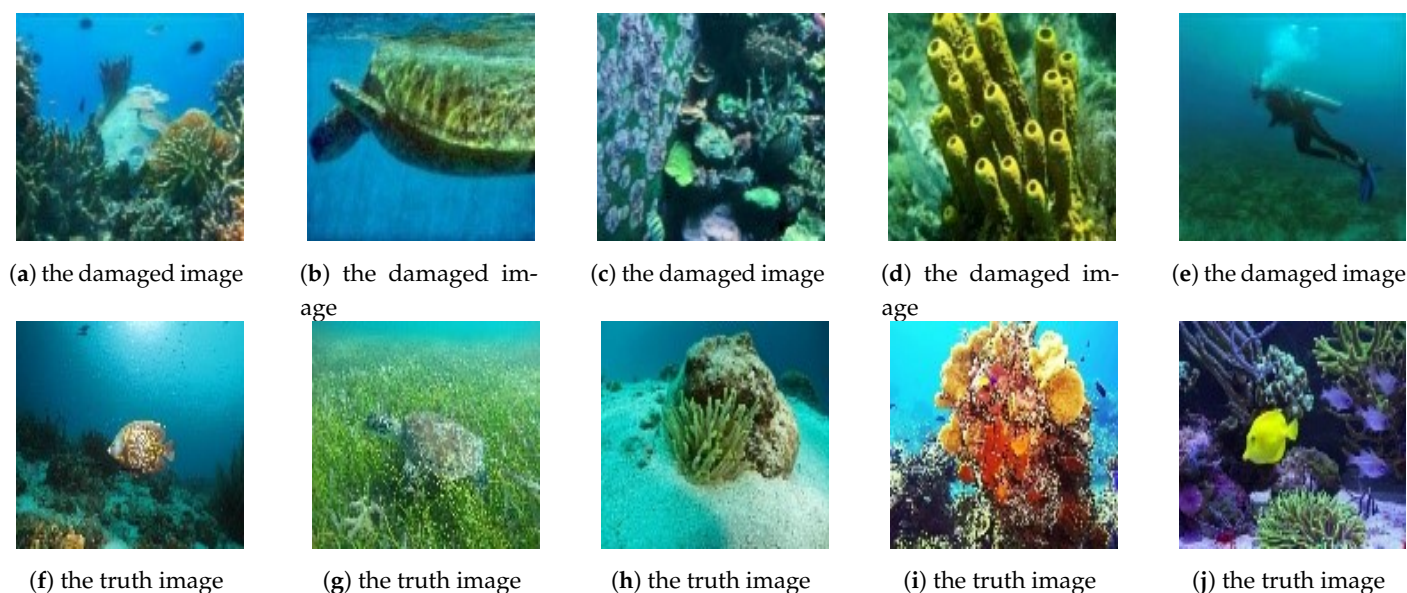


Figure 6. EUVP unpaired dataset.

EUVP paired datasets were obtained through  $X$  and  $Y$  sets of data, where  $I^C \in X$ ,  $I^D \in Y$ ,  $I^D$  is the truth image, and  $I^C$  is the lossy image; the network model is obtained by learning the mapping relationship between the truth image and the lossy image through CycleGAN. The artificially selected truth image is inputted into the model to generate a lossy image, thus generating a pair of datasets. Figure 7 shows a portion of the EUVP paired dataset. In Figure 7, the upper row represents the damaged images and the lower row are the truth images. The image sample data of EVUP paired datasets are shown in Table 2. The sample data of the unpaired dataset are shown in Table 3.

In this paper, the “Underwater Dark” sub-dataset was selected for network training. In Section 3.1, the summary of the “Underwater Dark” dataset was processed. This paper presents 5397 pairs of training pairs in the training set and 570 images in the verification set. In this paper, 100 lossy underwater images were selected from the “Underwater ImageNet” sub-data for subsequent testing.

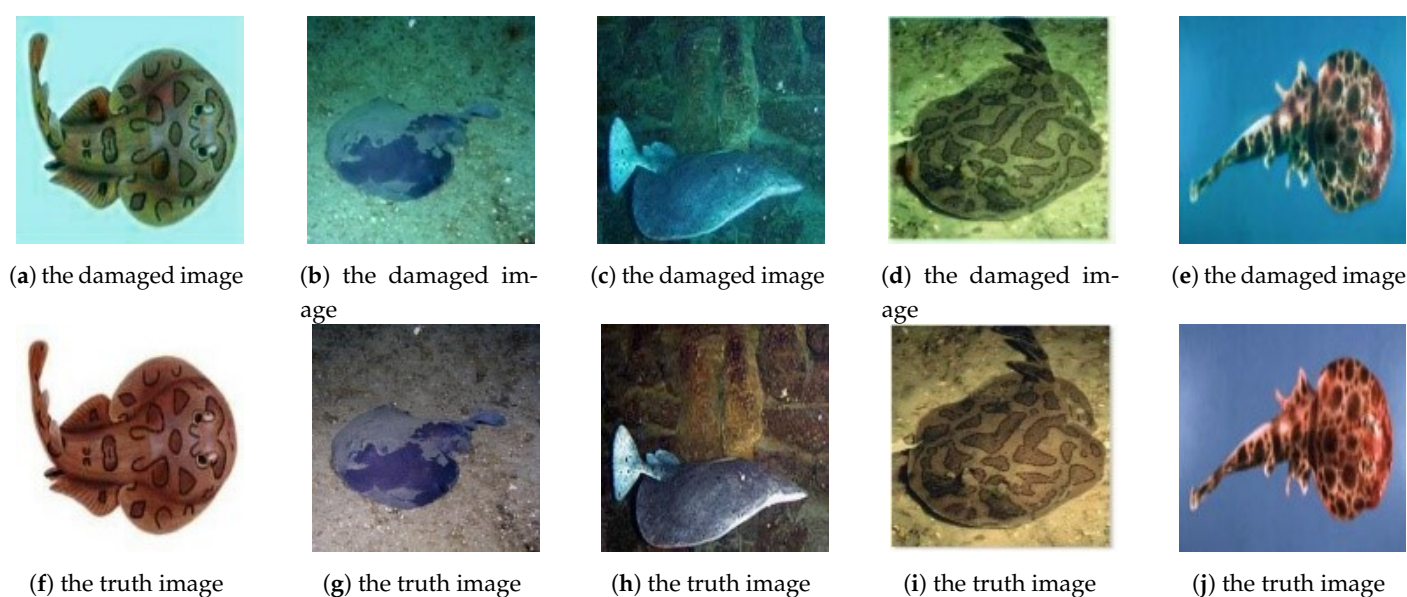


Figure 7. EUVP paired dataset.

**Table 2.** EVUP paired dataset image sample data.

Dataset Name	Paired Datasets	Validation	Total Number of Image Samples
Underwater Dark	5550	570	11,670
Underwater ImageNet	3700	1270	8670
Underwater Scenes	2185	130	4500

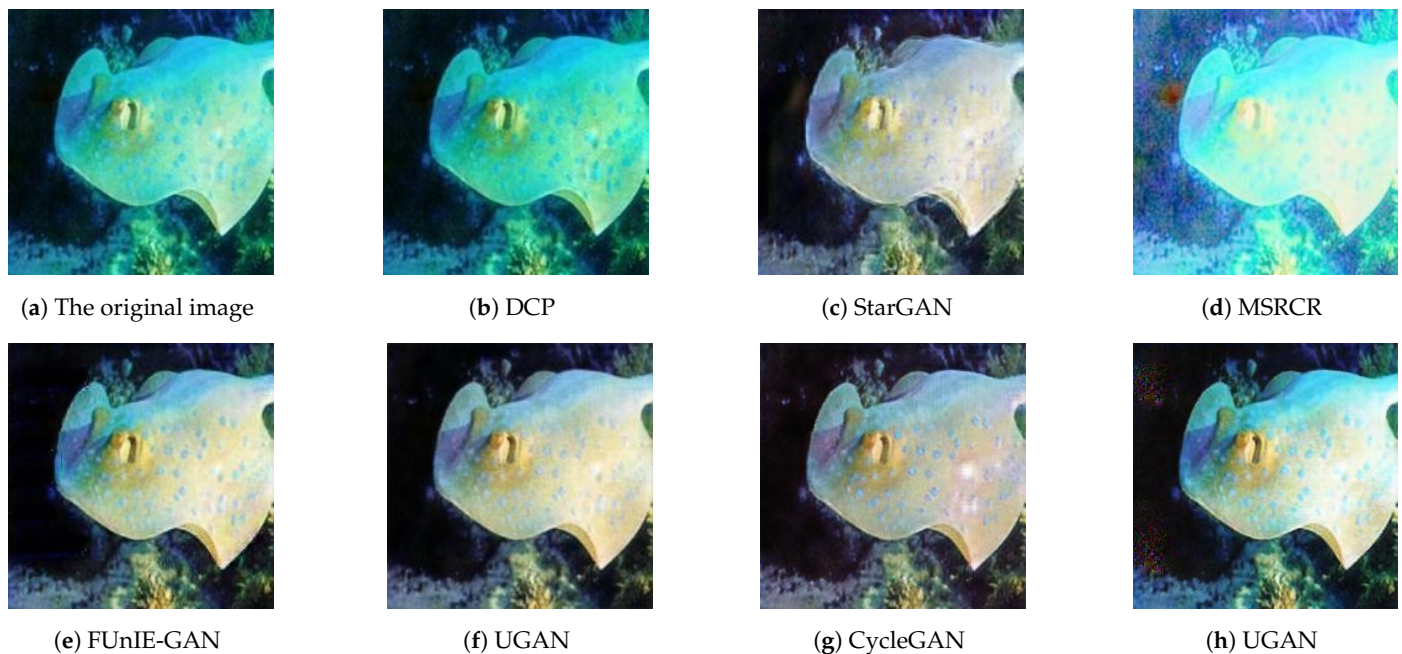
**Table 3.** EVUP paired dataset image sample data.

Damaged Images	Truth Images	Validation	Total Number of Image Samples
3195	3140	330	6665

### 4.3. Experiment

#### 4.3.1. Subjective Assessment

In this paper, five blue and green underwater images were selected, including different fish and seaweed scenes. This paper used the DCP [10] algorithm, the StarGAN [9] algorithm, the MSRCR [12] algorithm, the FUnIE-GAN [22] algorithm, the UGAN [21] algorithm and the CycleGAN [23] algorithm as a comparison to the FUnIE-GAN-NIQE algorithm proposed in this section (see Figures 8–12).

**Figure 8.** Scene 1 algorithm effect diagram.

By analyzing the effect images of the seven image enhancement algorithms used in the simulation experiments in Figures 8–12, we can see that the contrast of the images processed by the seven algorithms improves in varying degrees. The contrast of the enhanced image of the DCP algorithm and MSRCR algorithm is not high, and there is a serious color deviation difficulty. Although the color of the images enhanced by StarGAN and CycleGAN are more pronounced, the effects of image contrast and sharpness are poor. By comparing the effect images of the CycleGAN algorithm and the FUnIE-GAN-NIQE algorithm, we can find that although the image generated by the CycleGAN algorithm is similar to the FUnIE-GAN-NIQE algorithm in contrast, the effect of the FUnIE-GAN-NIQE algorithm is obviously better than that of the CycleGAN algorithm in the texture details. The loss of the FUnIE-GAN-NIQE algorithm consists of four parts. In this paper, the weight of each part is set reasonably to improve the texture and the contrast of the generated image.



However, the contrast of the image generated by the CycleGAN algorithm seems to be improved, but the exposed area appears in the image shown in Figure 8g, and the boundary information of the object shown in Figures 9g and 10g is not clear, and the color contrast between objects is low. The image quality loss added in the FUnIE-GAN-NIQE algorithm can reduce the exposure and low contrast of the generated underwater image. At the same time, the CycleGAN algorithm is not always effective in improving image contrast, and the function of improving image contrast is not added into CycleGAN algorithm. The image enhanced by FUnIE-GAN and UGAN is close to the trained truth image in color and contrast, but there is still a clear gap between the definition of the image and the truth image. Because the NIQE index based on human vision is added as a loss in this algorithm, the color and contrast of the generated image are slightly different to those of the true image. Compared with FUnIE-GAN and UGAN, the color of the image generated by this algorithm is more vivid and consistent with the conventional feeling of the human eye, and the contrast is also somewhat improved.

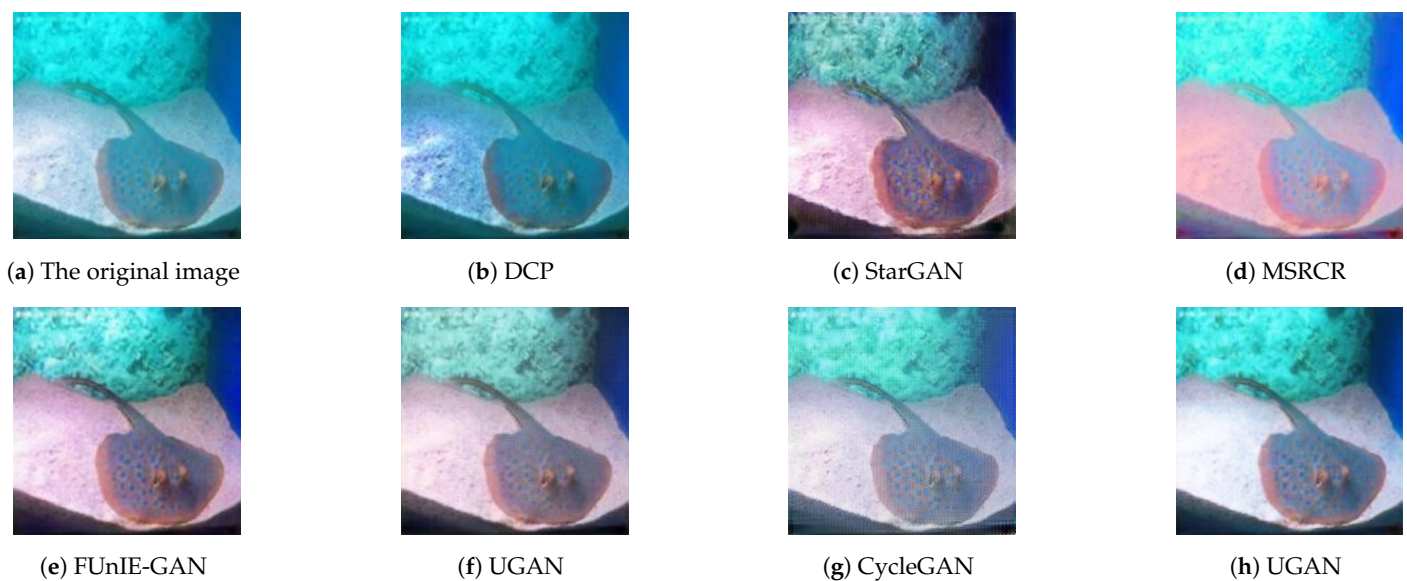


Figure 9. Scene 2 algorithm effect diagram.

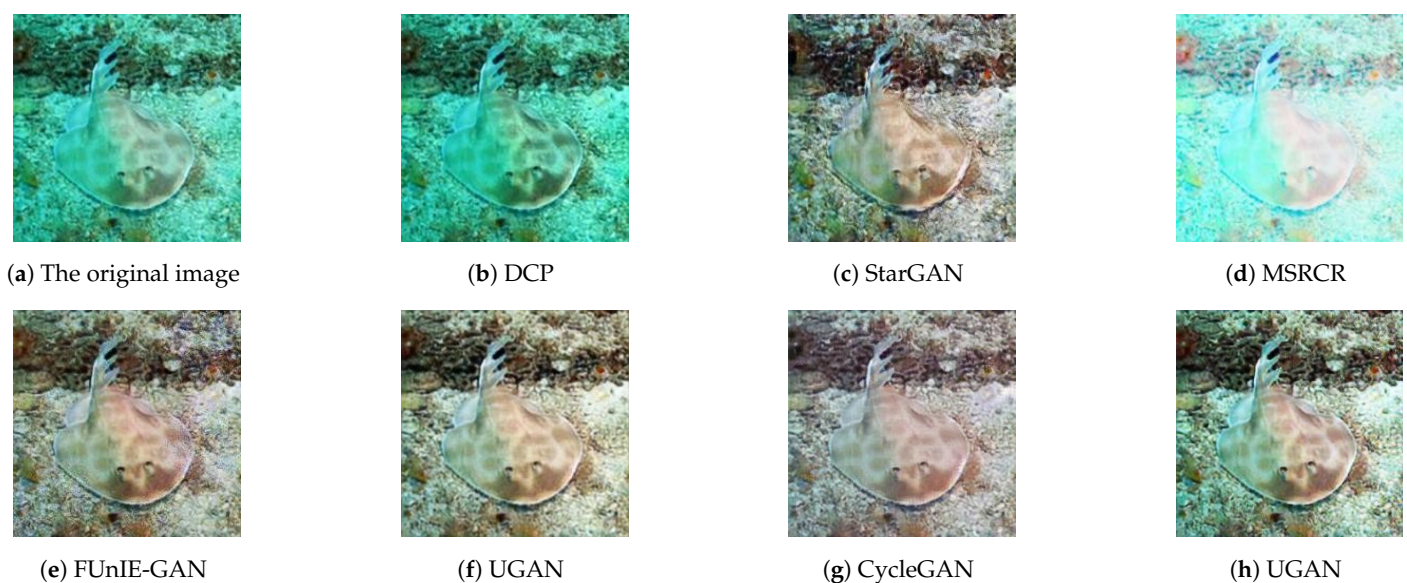
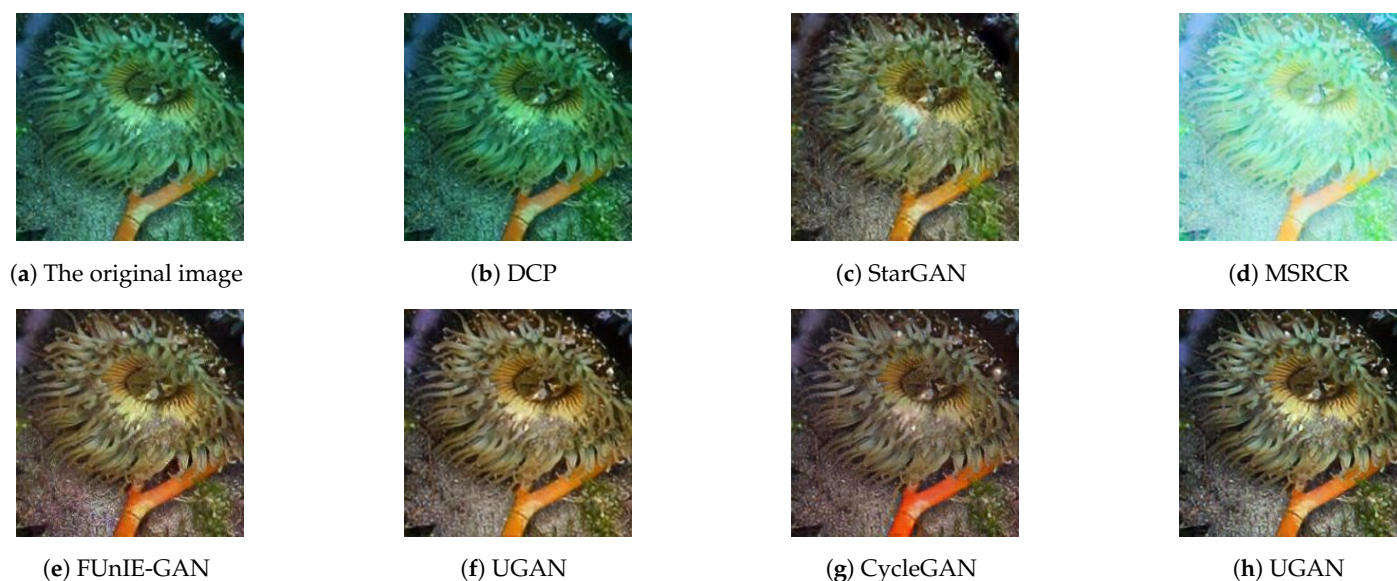
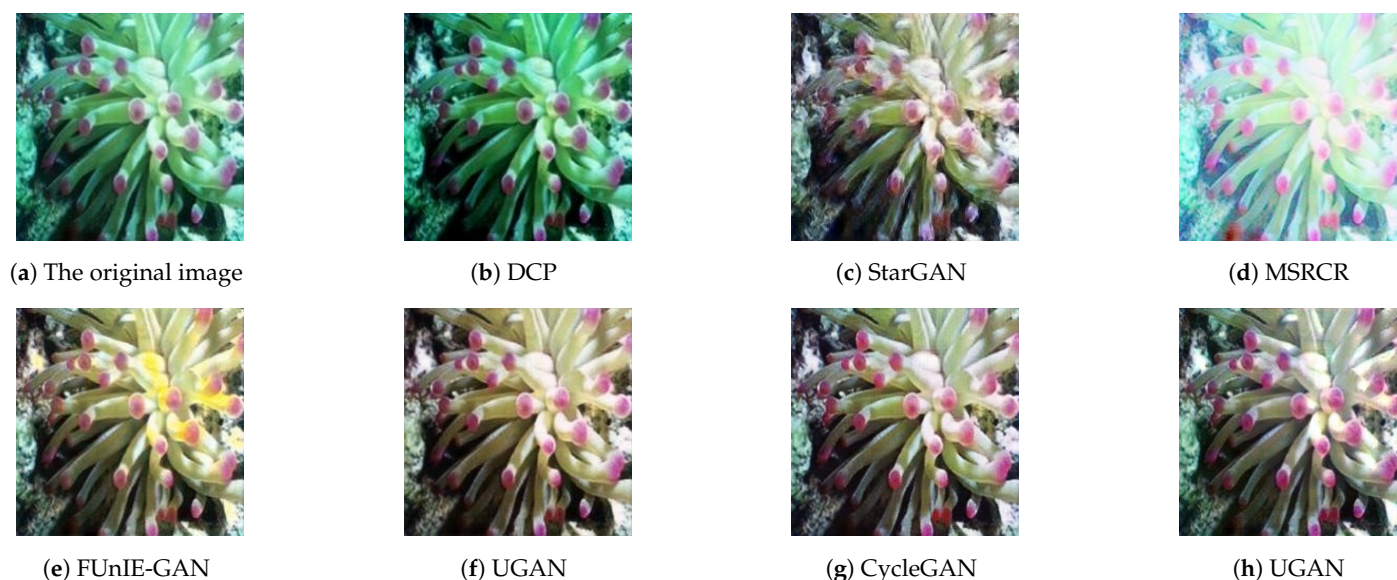


Figure 10. Scene 3 algorithm effect diagram.



**Figure 11.** Scene 4 algorithm effect diagram.



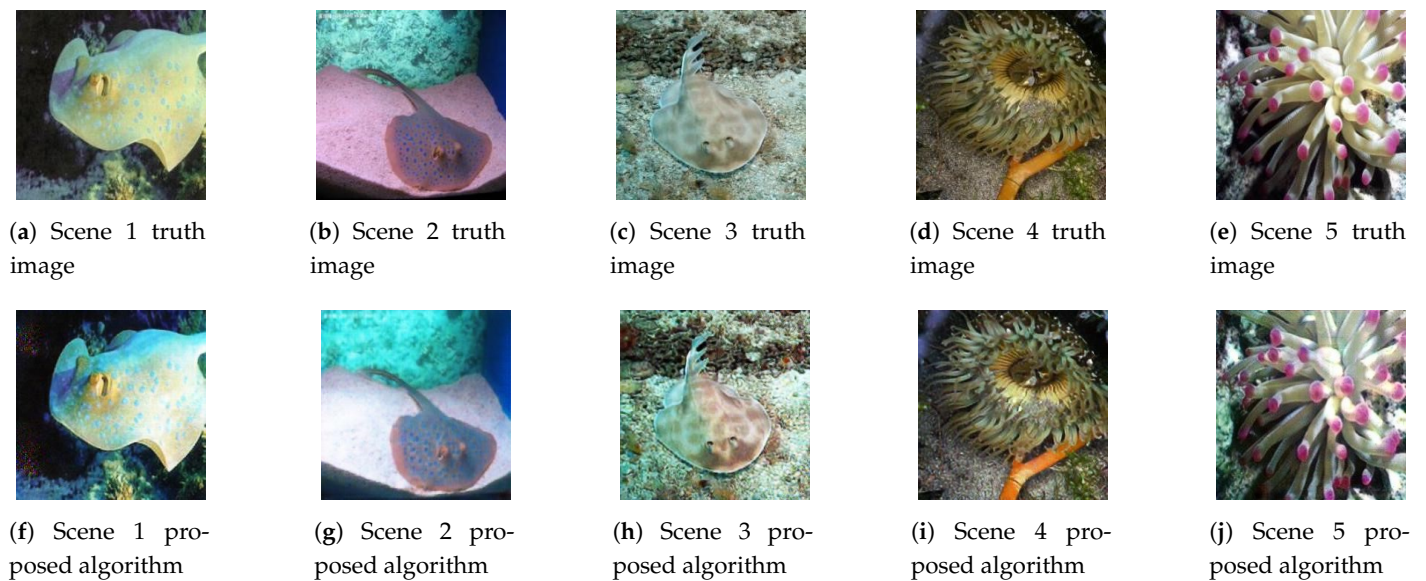
**Figure 12.** Scene 5 algorithm effect diagram.

Figure 13 shows the truth map of five scenes and the algorithm effect diagram of this section. By analyzing Figure 13, we find that the contrast of the algorithm in this section is clearer than that of the true image. However, by comparing the enhanced effect images of the seven algorithms in Figures 8–12, as well as the truth images in Figure 13, we find that the sharpness of the seven enhanced effect images is not as pronounced as in the truth images. This is also the most common problem for underwater image enhancement in specific engineering applications: the relationship between speed and clarity.

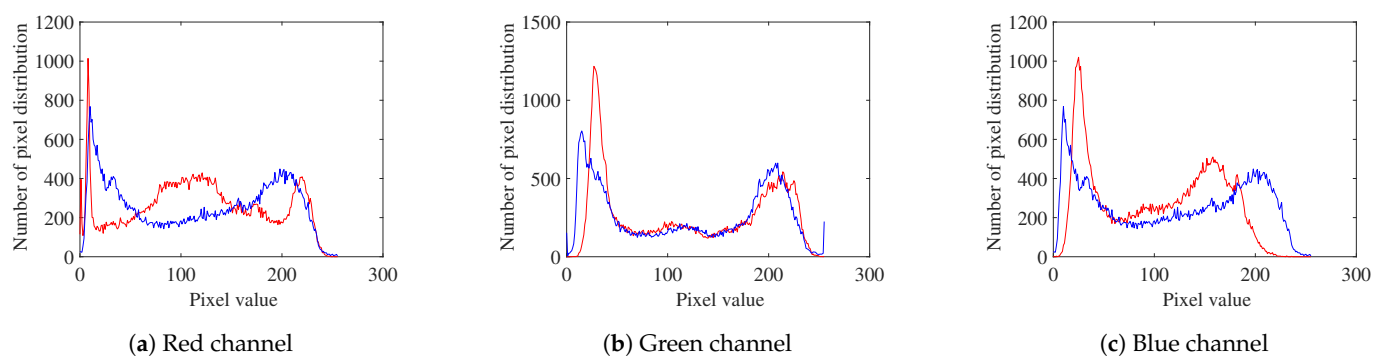
Figures 14–18 are a comparison of the truth map of the five scenes in Figure 13 with the histograms of R, G, and B of the algorithm effect map in this section, where the  $x$  axis is the pixel value and the  $y$  axis is the number of pixels distributed. According to the analysis of Figure 15, the number of 0 pixel in the red channel of the truth map in Figure 15a is too many, while the number of 0 pixel in the red channel of the proposed algorithm is 0. The number of other pixels is neither too high nor too low, which indicates that the histogram-stretching effect of the algorithm is more efficient. The contrast of the image generated by this algorithm is improved compared to the truth map. Similarly, the



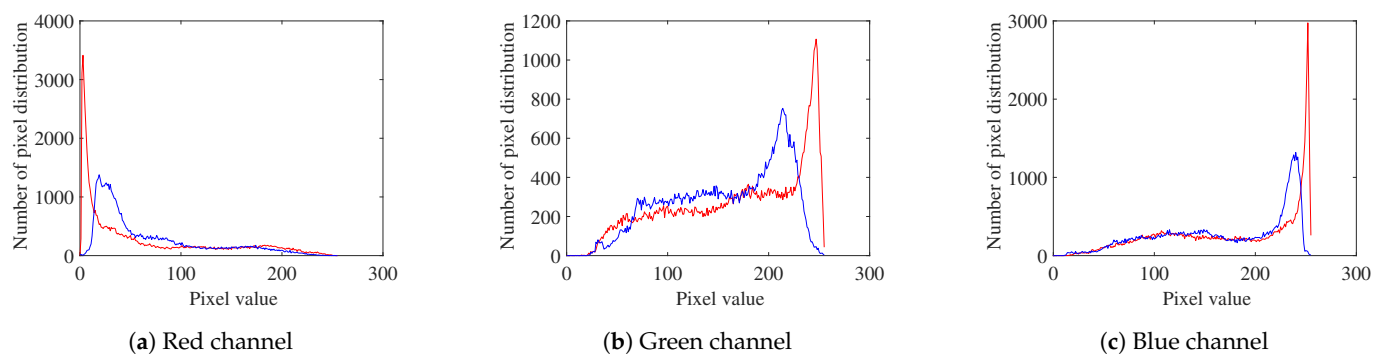
histogram of the algorithm effect diagram in Figures 17c and 18a has an obvious stretching effect, which can also objectively explain the results of naked-eye observation in Figure 13.



**Figure 13.** Comparison of the effect of the truth map with that of the algorithm in this paper.



**Figure 14.** Histogram comparison between the truth graph of Scene 1 and the proposed algorithm.



**Figure 15.** Histogram comparison between the truth graph of Scene 2 and the proposed algorithm.



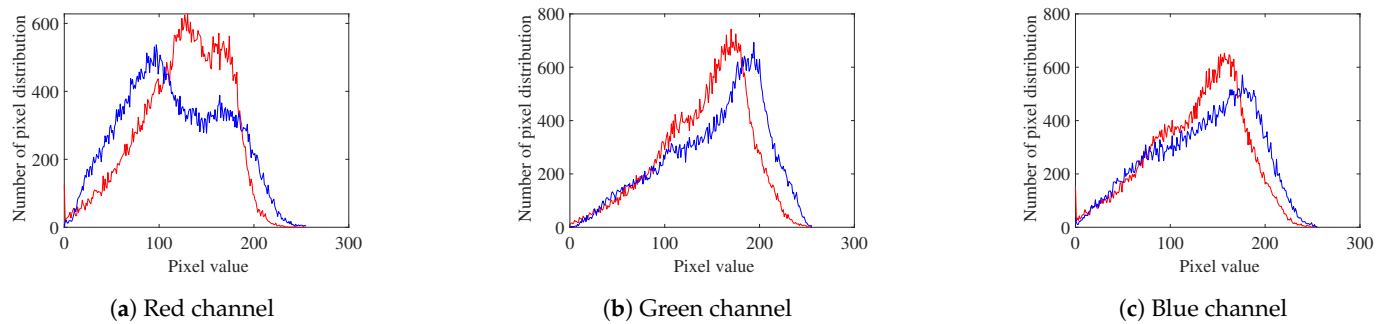


Figure 16. Histogram comparison between the truth graph of Scene 3 and the proposed algorithm.

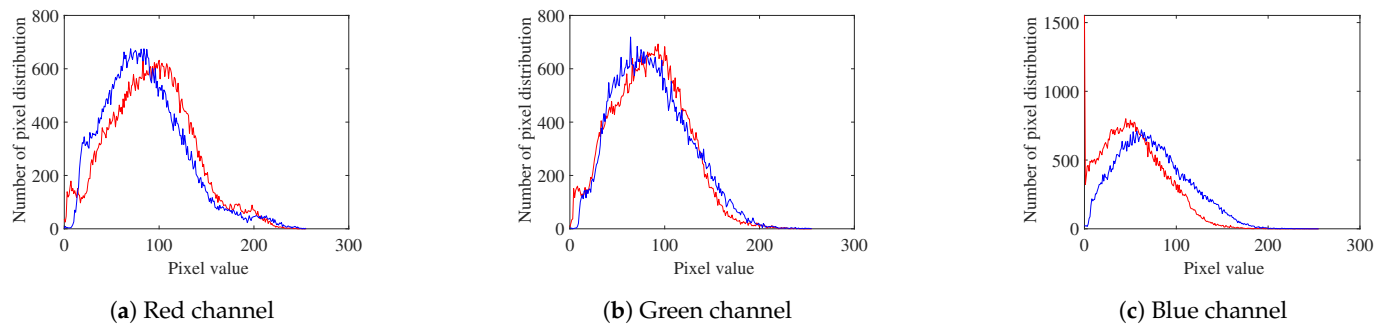


Figure 17. Histogram comparison between the truth graph of Scene 4 and the proposed algorithm.

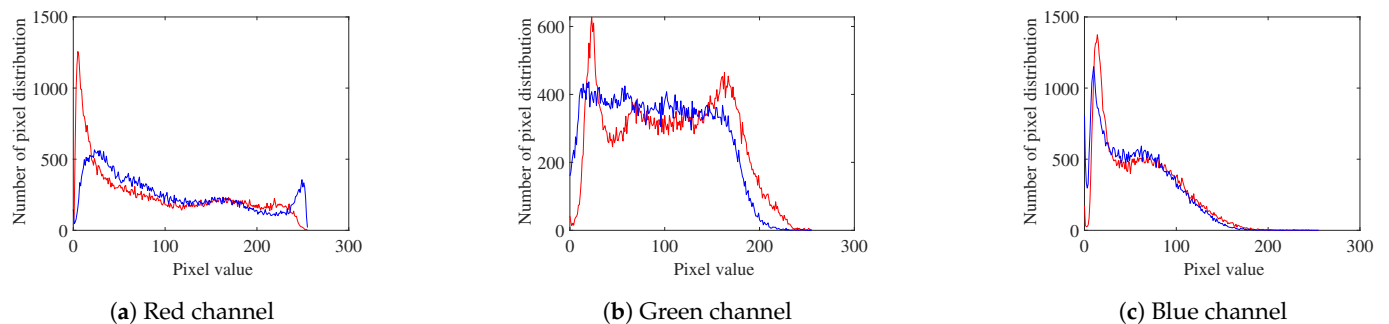


Figure 18. Histogram comparison between the truth graph of Scene 5 and the proposed algorithm.

#### 4.3.2. Objective Assessment

The image generated by the FUNIE-GAN-NIQE algorithm proposed in this paper exceeds the truth image set by the existing dataset. The full reference image quality evaluation index cannot be used. This paper uses no-reference image quality evaluation indexes to evaluate the image generated by the FUNIE-GAN-NIQE algorithm. Tables 4 and 5 correspond to the NIQE and UCIQE [26] indicators of Figures 8–12. Bold fonts in the Tables 4 and 5 represent the optimal index.

Table 4. NIQE index.

Image \ Scene	Scene 1	Scene 2	Scene 3	Scene 4	Scene 5
The original image	3.832	4.432	5.856	6.023	4.579
DCP	3.683	4.367	5.856	6.114	4.681
StarGAN	3.465	<b>3.700</b>	5.011	5.705	4.196
MSRCR	3.740	4.383	5.490	5.728	4.452
FUnIE-GAN	3.306	4.929	5.228	6.208	4.655
UGAN	3.727	4.054	5.626	6.963	5.314
CycleGAN	3.403	4.323	5.704	5.842	5.124
The proposed algorithm	<b>3.262</b>	4.400	<b>4.990</b>	<b>5.611</b>	<b>4.146</b>

**Table 5.** UCIQE index.

Image \ Scence	Scence 1	Scence 2	Scence 3	Scence 4	Scence 5
The original image	0.603	0.524	0.517	0.535	0.561
DCP	0.603	0.498	0.522	0.534	0.601
StarGAN	0.594	0.623	0.568	0.567	0.591
MSRCR	0.523	0.560	0.411	0.426	0.453
FUnIE-GAN	0.627	<b>0.627</b>	0.541	0.565	0.616
UGAN	0.611	0.593	0.565	0.582	0.600
CycleGAN	0.576	0.545	0.497	0.531	0.594
The proposed algorithm	<b>0.631</b>	0.593	<b>0.570</b>	<b>0.583</b>	<b>0.624</b>

By analyzing Tables 4 and 5, we find that the values of NIQE and UCIQE of the enhanced images of the FUnIE-GAN-NIQE algorithm are better than those of the other algorithms in the five scenes. The above seven algorithms are used to enhance the 100 images in the test set of this paper, and the enhanced results are compared with NIQE and UCIQE indicators. The score statistics are shown in Table 6, which contains the enhanced average and standard deviation of 100 underwater images. Bold font is the best index in Table 6. By analyzing Table 6, it can be observed that the FUnIE-GAN-NIQE algorithm proposed in this paper is relatively effective in using the NIQE index and UCIQE index to evaluate the dataset composed of 100 underwater images.

**Table 6.** Objective evaluation index of underwater image quality.

Algorithm \ Index	DCP	StarGAN	MSRCR	FUnIE-GAN	UGAN	CycleGAN	The Proposed Algorithm
NIQE	4.95 ± 1.80	4.80 ± 1.83	5.12 ± 1.78	4.70 ± 1.92	4.68 ± 1.60	4.70 ± 1.72	<b>4.65 ± 1.68</b>
UCIQE	0.49 ± 0.06	0.54 ± 0.07	0.50 ± 0.08	0.56 ± 0.05	0.56 ± 0.07	0.54 ± 0.06	<b>0.58 ± 0.06</b>

#### 4.4. Engineering Analysis

Underwater image enhancement algorithms have high requirements for engineering applications. FUnIE-GAN has a memory requirement of 17 MB and runs at 25.4 frames per second on an embedded platform (Nvidia Jetson TX2), and 148.5 frames per second on a graphics card (Nvidia GTX 1080). The running speed on the robot CPU (Intel Core-I3 6100U) is 7.9 frames. The algorithm FUnIE-GAN-NIQE proposed in this paper retains the advantage of a fast calculation speed of FUnIE-GAN, and can be fully applied to the real-time underwater image enhancement process of an underwater robot.

## 5. Conclusions

This paper aimed to solve the difficulty of obtaining underwater truth images when a supervised generative adversarial network, which works in underwater image enhancement, leads to low contrast for generated images. First of all, this paper filtered the dataset of EUVP truth images. Secondly, this paper proposed to add the NIQE index to the loss of the generator to provide the generated image with higher contrast and make it more in line with the perception of the human eye. At the same time, we attempted to make a generated image exceed the effects of the truth image set by the existing dataset. Then, this paper proposed to add the NIQE index to the discriminator to diversify its discriminant factors in FUnIE-GAN. Finally, this paper proposed a new structure of GAN, and selected the generated image suitable for this paper by training 10 generators. In this paper, several groups of experiments were compared. Through subjective evaluation and objective evaluation indicators, it was verified that the enhanced image contrast of this algorithm was better than the truth image contrast set by the existing dataset theory. At the end of this paper, the real-time performance of the algorithm was analyzed to verify that the algorithm could be used in engineering.

In this paper, it was verified that the FUnIE-GAN-NIQE algorithm proposed in this paper can improve the contrast of the enhanced image through subjective evaluation and objective index evaluation. This paper does not prove the superiority of this algorithm through specific engineering examples, such as underwater target detection, underwater image segmentation, etc. In future work, we will continue to study the task in this direction, and verify the superiority of the underwater image enhancement algorithm through more abundant engineering examples.

The contrast of the image generated by the FUnIE-GAN-NIQE algorithm proposed in this paper is better than that of the truth image, but the clarity of the image is far lower than that of the truth image. This is because the underwater image enhancement algorithm has high engineering requirements, and the running speed can only be improved by reducing the amount of computation, which is a common problem in the existing underwater image enhancement algorithms. Therefore, it is hoped that the future work can achieve a balance between clarity and speed, and ensure the running speed of the algorithm in the case of greatly reducing the clarity.

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

**Funding:** The research in this paper is supported by the National Natural Science Foundation of China (61773219, 61701244), the key special project of the National Key R&D Program (2018YFC1405703), and the NUIST Students' Platform for Innovation and Entrepreneurship Training Program (202010300300). Heartfelt thanks are expressed to the reviewers who submitted valuable revisions to this article.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The code used to support the findings of this study are available from the corresponding author upon request (nuistpanda@163.com).

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

## References

1. Li, Y.; Lu, H.; Zhang, L.; Li, J.; Serikawa, S. Real-time visualization system for deep-sea surveying. *Math. Probl. Eng.* **2014**, *2014*, 437071. [\[CrossRef\]](#)
2. Ahn, J.; Yasukawa, S.; Sonoda, T.; Nishida, Y.; Ishii, K.; Ura, T. An optical image transmission system for deep sea creature sampling missions using autonomous underwater vehicle. *IEEE J. Ocean. Eng.* **2018**, *45*, 350–361. [\[CrossRef\]](#)
3. Sai, S.; Sai, I. Artificial Object Images Synthesis in Underwater Robot Vision System. In Proceedings of the 2020 International Conference on Industrial Engineering, Applications and Manufacturing, Sochi, Russia, 18–22 May 2020; pp. 1–6.
4. 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 Proceedings of the Pacific Rim Conference on Multimedia, Hefei, China, 21–22 September 2018; pp. 678–688.
5. Iwamoto, Y.; Hashimoto, N.; Chen, Y. Fast Dark Channel Prior Based Haze Removal from a Single Image. In Proceedings of the 2018 14th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery, Huangshan, China, 28–30 July 2018; pp. 458–461.
6. Jobson, D.; Rahman, Z.; Woodell, G. Properties and performance of a center/surround retinex. *IEEE Trans. Image Process.* **1997**, *6*, 451–462. [\[CrossRef\]](#) [\[PubMed\]](#)
7. Singhai, J.; Rawat, P. Image enhancement method for underwater, ground and satellite images using brightness preserving histogram equalization with maximum entropy. In Proceedings of the International Conference on Computational Intelligence and Multimedia Applications, Sivakasi, India, 13–15 December 2007; pp. 507–512.
8. Isola, P.; Zhu, J.; Zhou, T.; Efros, A. Image-to-image translation with conditional adversarial networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 1125–1134.
9. Choi, Y.; Choi, M.; Kim, M.; Ha, J.; Kim, S.; Choo, J. Stargan: Unified generative adversarial networks for multi-domain image-to-image translation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 8789–8797.

10. He, K.; Sun, J.; Tang, X. Single Image Haze Removal Using Dark Channel Prior. *IEEE Trans. Pattern Anal. Mach. Intell.* **2011**, *33*, 2341–2353. [[PubMed](#)]
11. Rahman, Z.; Jobson, D.; Woodell, G. Multi-Scale Retinex for Color Image Enhancement. In Proceedings of the 3rd IEEE International Conference on Image Processing, Lausanne, Switzerland, 16–19 September 1996; pp. 1003–1006.
12. Jobson, D.; Rahman, Z.; Woodell, G. A multiscale retinex for bridging the gap between color images and the human observation of scenes. *IEEE Trans. Image Process.* **1997**, *6*, 965–976. [[CrossRef](#)] [[PubMed](#)]
13. Goodfellow, I.; Pouget-Abadie, J.; Mirza, M.; Bing, X.; Bengio, Y. Generative Adversarial Nets. In Proceedings of the 27th International Conference on Neural Information Processing Systems, Montreal, QC, Canada, 8 December 2014; pp. 2672–2680.
14. Chen, B.; Xia, M.; Huang, J. MFANet: A Multi-Level Feature Aggregation Network for Semantic Segmentation of Land Cover. *Remote Sens.* **2021**, *13*, 731. [[CrossRef](#)]
15. Xia, M.; Wang, T.; Zhang, Y.; Liu, J.; Xu, Y. Cloud/shadow segmentation based on global attention feature fusion residual network for remote sensing imagery. *Int. J. Remote Sens.* **2021**, *42*, 2022–2045. [[CrossRef](#)]
16. Xia, M.; Cui, Y.; Zhang, Y.; Xu, Y.; Xu, Y. DAU-Net: A novel water areas segmentation structure for remote sensing image. *Int. J. Remote Sens.* **2021**, *42*, 2594–2621. [[CrossRef](#)]
17. Xia, M.; Liu, W. Non-intrusive load disaggregation based on composite deep long short-term memory network. *Expert Syst. Appl.* **2020**, *160*, 113669. [[CrossRef](#)]
18. Xia, M.; Zhang, X.; Liu, W.; Weng, L.; Xu, Y. Multi-Stage Feature Constraints Learning for Age Estimation. *IEEE Trans. Inf. Forensics Secur.* **2020**, *15*, 2417–2428. [[CrossRef](#)]
19. Arjovsky, M.; Bottou, L. Towards principled methods for training generative adversarial networks. *arXiv* **2017**, arXiv:1701.04862.
20. Li, J.; Skinner, K.; Eustice, R.; 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](#)]
21. Fabbri, C.; Islam, M.; Sattar, J. Enhancing underwater imagery using generative adversarial networks. In Proceedings of the 2018 IEEE International Conference on Robotics and Automation, Brisbane, QLD, Australia, 21–25 May 2018; pp. 7159–7165.
22. Islam, M.; Xia, Y.; Sattar, J. Fast underwater image enhancement for improved visual perception. *IEEE Robot. Autom. Lett.* **2020**, *5*, 3227–3234. [[CrossRef](#)]
23. Zhu, J.; Park, T.; Isola, P.; Efros, A. Unpaired image-to-image translation using cycle-consistent adversarial networks. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 25 December 2017; pp. 2223–2232.
24. Hu, K.; Zhang, Y. An Underwater Image Enhancement Algorithm Based on MSR Parameter Optimization. *J. Mar. Sci. Eng.* **2020**, *8*, 741. [[CrossRef](#)]
25. Mittal, A.; Soundararajan, R.; Bovik, A. Making a “completely blind” image quality analyzer. *IEEE Signal Process. Lett.* **2012**, *20*, 209–212. [[CrossRef](#)]
26. Yang, M.; Sowmya, A. An underwater color image quality evaluation metric. *IEEE Trans. Image Process.* **2015**, *24*, 6062–6071. [[CrossRef](#)] [[PubMed](#)]