



Local-CycleGAN: a general end-to-end network for visual enhancement in complex deep-water environment

Xianhui Zong¹ · Zhehan Chen¹ · Dadong Wang²

Accepted: 5 September 2020 / Published online: 21 October 2020
© Springer Science+Business Media, LLC, part of Springer Nature 2020

Abstract

Underwater image analysis is crucial for many applications such as seafloor survey, biological and environment monitoring, underwater vehicle navigation, inspection and maintenance of underwater infrastructure etc. However, due to light absorption and scattering, the images acquired underwater are always blurry and distorted in color. Most existing image enhancement algorithms typically focus on a few features of the imaging environments, and enhanced results depend on the characteristics of original images. In this study, a local cycle-consistent generative adversarial network is proposed to enhance images acquired in a complex deep-water environment. The proposed network uses a combination of a local discriminator and a global discriminator. Additionally, quality-monitor loss is adopted to evaluate the effect of the generated images. Experimental results show that the local cycle-consistent generative adversarial network is robust and can be generalized for many different image enhancement tasks in different types of complex deep-water environment with varied turbidity.

Keywords Local auxiliary discriminator · Quality-monitor loss function · Cycle-consistency generative adversarial network · Image enhancement · Turbid deep-water environment

1 Introduction

With the development of robotic technology, remotely operated underwater vehicles (ROVs), autonomous underwater vehicles (AUVs), and other underwater equipment are increasingly widely used in various deep-water engineering projects. Optical visual systems are one of the main components of underwater intelligent equipment and are important for object recognition and navigation in a complex deep-water environment. However, the underwater optical image processing is a challenging field and associated with complexity and uncertainty. Underwater images exhibit low brightness, blurs, and color distortions, these attributes significantly affect the visual effect. The attenuations of blue and green lights are weaker than that of red light in the underwater transmission process, and thus the underwater images generally appear green, yellow, cyan, and blue in color.

Water bodies are affected by the diameter and quantity of suspended particles and exhibit different turbidity levels.

At present, the dark channel prior (DCP) algorithm [1], the Retinex algorithm [2], the Automatic Color Enhancement (ACE) algorithm [3], and other traditional image enhancement algorithms are mostly used to solve the underwater image enhancement problem. The DCP algorithm restores a blurred image by estimating the atmospheric scattering model, while the Retinex algorithm estimates the reflection image in the object reflection model. Retinex and DCP have the same perspective that establishes the model for a certain situation, and the ACE algorithm is a combination of Retinex and white balance theory. However, underwater image processing cannot be treated as the same. These algorithms have ideal processing effects for the test samples used in the experiments reported in these papers. Once underwater images of other styles are selected, the image enhancement results are often not satisfactory. Therefore, different algorithms or their associated parameters have to be adopted for processing images acquired from different deep-water environments. Deep learning can generalize well and can be used to solve the problem of General Adaptability by learning the mapping between blurred images and clear images with a single model.

In this study, the local cycle-consistent generative adversarial network (Local-CycleGAN) is proposed for enhancing

✉ Zhehan Chen
chenzh_ustb@163.com

¹ School of Mechanical Engineering, University of Science & Technology Beijing, Beijing, China

² Quantitative Imaging Research Team, Data61, CSIRO, Sydney, Australia

deep-water images. Although the original Generated adversarial network (GAN) [4] can generate the style features of target images, it is difficult to generate the accurate content if the training epoch is not enough, and the training process is not very stable. Therefore, based on CycleGAN [5], the proposed network adopts a loop structure to ensure that the entire content of the generated image is not separated from that of the blurred image. A local auxiliary discriminator is added to identify whether the content and style of the generated image match the features of underwater clear images from both the global and local perspectives. And correspondingly local loss is used to constrain local details of the generated image. Besides, quality-monitor loss is designed for monitoring the quality of generated images during the training process. Through feedback from quality-monitor loss, the generator is continuously optimized and generates clear images. Experimental results show that our proposed network outperforms the others and exhibits good generalization ability, adaptability and stability.

The rest of the paper is organized as follows: First, we give a summary of prior work in Section 2, followed by the detailed description of the proposed method in Section 3. Section 4 presents our experimental results which are qualitatively and quantitatively assessed in Section 5. Finally, conclusions are drawn based on our experiments and analysis on the comparison among our proposed Local-CycleGAN and other existing algorithms.

2 Prior work

2.1 Underwater image enhancement

Yan [6] proposed an underwater grayscale image enhancement algorithm based on local complexity to sharpen the edge and improve the overall contrast of the image. Yang et al. [7] used the adaptive histogram equalization method to enhance the sea cucumber images with poor underwater contrast. However, their experiments were based on grayscale images only. Zhang et al. [8] modified the original Retinex algorithm to process the underwater images and got the competitive performance in color restoration aspect. However, the image clarity was not ideal. Javier et al. [9] proposed to estimate the depth of the underwater scene by using a neural network, then applied to a dehazing method, such as DCP, ACE and CLAHE. However, the enhancing effect was limited by the dehazing method chosen. Ho et al. [10] divided the underwater image enhancement task into two sections. A successive color correction method was presented to restore color, and then a superpixel-based DCP algorithm was used to dehaze. The proposed enhancement method outperformed the most existing approaches, but it was not a unified scheme.

Some scholars tried to restore underwater images based on underwater imaging models. This method is relatively complex, and in complex underwater situations, the general underwater imaging models and their associated parameters work for specific underwater environments only, and does not generalize. Li et al. [11] proposed an algorithm to demist based on minimum information loss and increase contrast based on histogram distribution prior. However, experiments showed that the colors of some enhanced images were excessively saturated. Zhang et al. [12] supposed the colors of light source was the global background light and proposed a new underwater image formation model. The method was limited by non-uniformly lighting and was inapplicable to other complex underwater scenes.

The imaging process of underwater images is similar to that of foggy images, and it seems that defogging algorithms exhibit some influence on the development of algorithms for enhancing underwater blurred images. DCP was first proposed by He et al. [1], and its effect on foggy image was evident. Therefore, based on the features of underwater images, some researchers optimized the DCP algorithm to adapt to underwater images. Extant studies (Ma et al. [13], Song et al. [14], Tang et al. [15], Yang et al. [16], Yu et al. [17]) discussed the dark channel defogging algorithm and proposed a corresponding improved algorithm. To improve feature clarity in underwater images with low quality and illumination, Ma et al. [13] employed an improved wavelet fusion algorithm to fuse the results of the DCP algorithm and the gray world algorithm. Given the fast attenuation of red light in water, Song et al. [14] estimated background light by calculating the difference between the red channel and blue-green channels, restored underwater images based on the dark channel prior principle, and post-processed the images using the color balance algorithm. Although the method improves image quality, it exhibits a few limitations such as an estimation error and color supersaturation. In Tang et al. [15], the color of the processed images was excessively supersaturated and some areas of the images were distorted; and the processed images were in the same single. Yang et al. [16] proposed a new background light estimation method on the basis of which DCP was used to enhance underwater images for the elimination of image blur and color deviation. Yu et al. [17] combined the wavelet fusion technology and the DCP algorithm to enhance image clarity and contrast.

Recently, deep learning algorithms attracted considerable attention and are increasingly applied to image enhancement. Jin et al. [18] presented a model based on conditional generative adversarial networks to improve the performance of underwater image enhancement. Residual dense blocks were added to the generative model to mitigate the gradient disappearance problem. Liu et al. [19] also utilized the deep residual network to enhance underwater blurred images. When compared with the traditional algorithms, the network

exhibited optimal performance in terms of image detail restoration and color correction. Lu et al. [20] designed an adaptive SSIM loss in MCycle GAN system to improve the underwater image restoration performance. However, this model couldn't restore an image under an inhomogeneous illumination. Fu et al. [21] integrated data-driven deep learning and hand-crafted image enhancement for single underwater image enhancement. This lightweight model could compensate both the global distorted color and local reduced contrast. Li et al. [22] also presented a lightweight underwater image enhancement CNN model based on scene prior. These two methods were similar and had improvement space in network structures.

Traditional image enhancement algorithms process all pixels of an image in the same way or process different pixels by setting thresholds manually, which inevitably have limitations. Literature review reveals that the generalization of existing algorithms is limited and they cannot be easily adapted to diverse deep-water environments. In this study, we transform the image enhancement problem into an image generation problem. Local-CycleGAN is adopted to realize underwater image enhancement which learns and regenerates every pixel through the game between two networks. By adding an auxiliary discriminator and designing a feedback structure, the proposed method can learn high-level semantic information that is invisible and generate clear images. That is to say, the proposed network not only generates images with good quality, but also has general adaptability for most of underwater blurred images.

2.2 Generative adversarial network

GAN, proposed by Goodfellow et al. [4], was initially used for image generation task, to synthesize an image from noise or text, or translate images from original domain to target domain. When compared with the general neural network model, it is composed of two models, namely the generator module and discriminator module, which were used to generate more real samples through adversarial training. After GAN was proposed, it is increasingly used as one of the most popular models in deep learning.

Although GAN is becoming so popular, training a GAN can be challenging and the quality of generated images may be poor. Hence, some researchers have proposed improved models that achieve good results in various fields such as character generation, landscape image generation, image style transfer, image restoration, image fusion, and 3D model generation. Most improved strategies of generative adversarial networks were realized by optimizing the loss function and/or modifying the network structures.

Radford et al. [23] applied deep convolution network to the DCGAN model for image generation and proved that the representation level from object to scene was learned in the

adversarial training of generator and discriminator. Skinner et al. [24] proposed a network termed as WaterGAN. The internal structure of generator simulated an underwater imaging process. Underwater datasets were generated from the RGB-D dataset captured in air. Mirza and Osindero [25] proposed Conditional-GAN, which could construct images with simple input data and perform conditional conversion on the generator and discriminator. Zhu et al. [5] proposed the CycleGAN model to train unpaired datasets. The model comprises two generators and two discriminators that used inverse mapping to constrain target mapping and introduced cycle-consistency loss. It was applied to tasks such as style transfer and object translation. Chen et al. [26] introduced the InfoGAN model which combined the generative adversarial network and information theory to learn separation representation on the MNIST dataset and 3D rendered images. Li et al. [27] used a deep convolution network to build the BeautyGAN model to generate delicate facial makeup. Additionally, four loss items, namely adversarial loss, cycle-consistent loss, makeup loss, and perception loss, were designed in the study. Junho et al. [28] designed the U-GAT-IT model by combining GAN and a new attention module and employed a loss function with four parts. The model was used to generate anime images.

In our proposed network, a local monitoring process and a local discriminator are introduced to control image quality by learning the mapping from underwater blurred images to underwater clear images. More details of our proposed network can be found in the following section.

3 Proposed method

3.1 Overall framework

To realize the enhancement of blurred underwater images, we proposed a Local-CycleGAN to generate clear images from the blurred images. It is based on the CycleGAN model. The blurred underwater images and clear underwater images respectively obey two different distributions, namely B and C. As shown in Fig. 1, the proposed model is composed of a forward process and a backward process. In the forward process, the generator G_{B2C} in Local-CycleGAN is designed to translate the blurred underwater images belonging to B domain into C domain, and the global discriminator D_C is designed to determine whether an input image belongs to C domain. Initially, a blurred image b_t belonging to B domain is converted into an image c' by the generator G_{B2C} , where $c' = G_{B2C}(b_t)$. Subsequently, c' is inputted to D_C , which distinguishes the C domain images from the non-C domain images. If the generated image c' exhibits poor quality, then the discriminator D_C can easily distinguish the image $c' \notin C$ and determine c' as false, i.e., $D_C[G_{B2C}(b_t)] = 0$. If the generated

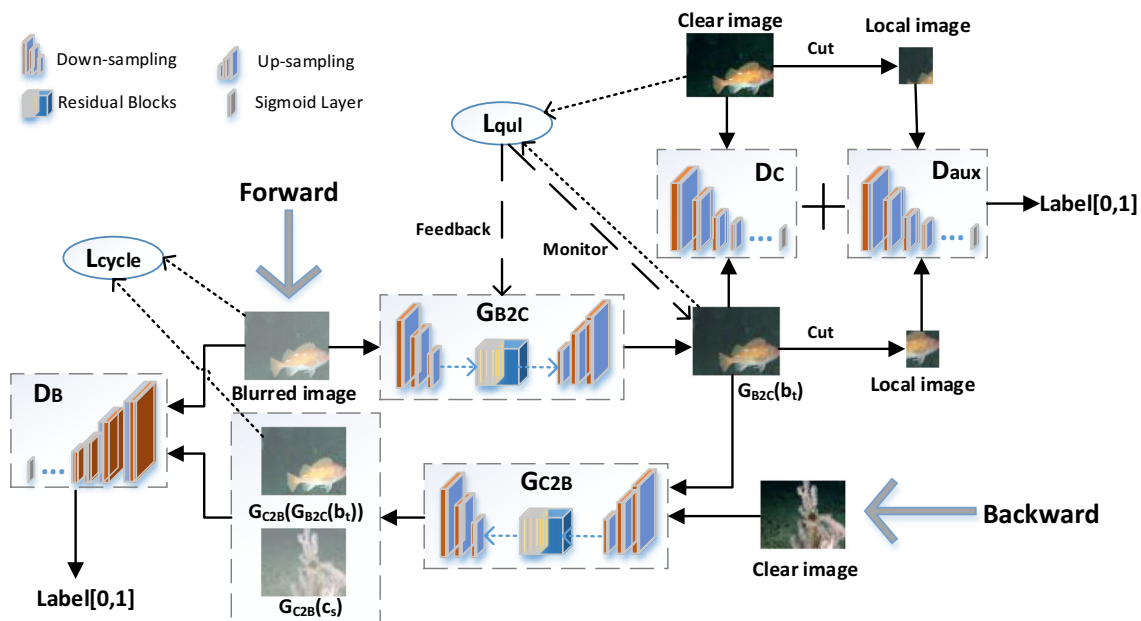


Fig. 1 Overall framework of our proposed Local-CycleGAN

image c' is very clear and is similar to the C domain images, then D_C can probably result in an inaccurate prediction. Finally, the predicted result is fed back to the generator G_{B2C} for fine-tuning.

Additionally, a local discriminator D_{aux} is introduced into the forward process to discriminate among local areas of the input image and aid G_{B2C} in training and generating local details. The global discriminator D_C and local discriminator D_{aux} are combined, and thus the global performance and the local appearance of the generated images are improved simultaneously. Furthermore, an inverse mapping G_{C2B} is added to the backward process to form a closed cycle network with G_{B2C} . Specifically, G_{C2B} is used to generate an image b' where the input of an image c_s belongs to the C domain, i.e. $b' = G_{C2B}(c_s)$. Furthermore, the discriminator D_B is used to discriminate whether the input image belongs to the B domain. Therefore, the proposed Local-CycleGAN includes two generator modules and three discriminator modules.

3.2 Generator

The generator module adopts an end-to-end full convolution network including the residual learning unit that can receive the input of different image sizes. Among them, the kernel size of all the convolution layers used in the generator is 3×3 . The residual learning unit can alleviate the problem of gradient disappearance while retaining the low-level features extracted from the previous layer. Generally, the up-sampling layer adopts either linear interpolation methods or transposed convolution methods. As shown in Fig. 2, the first row

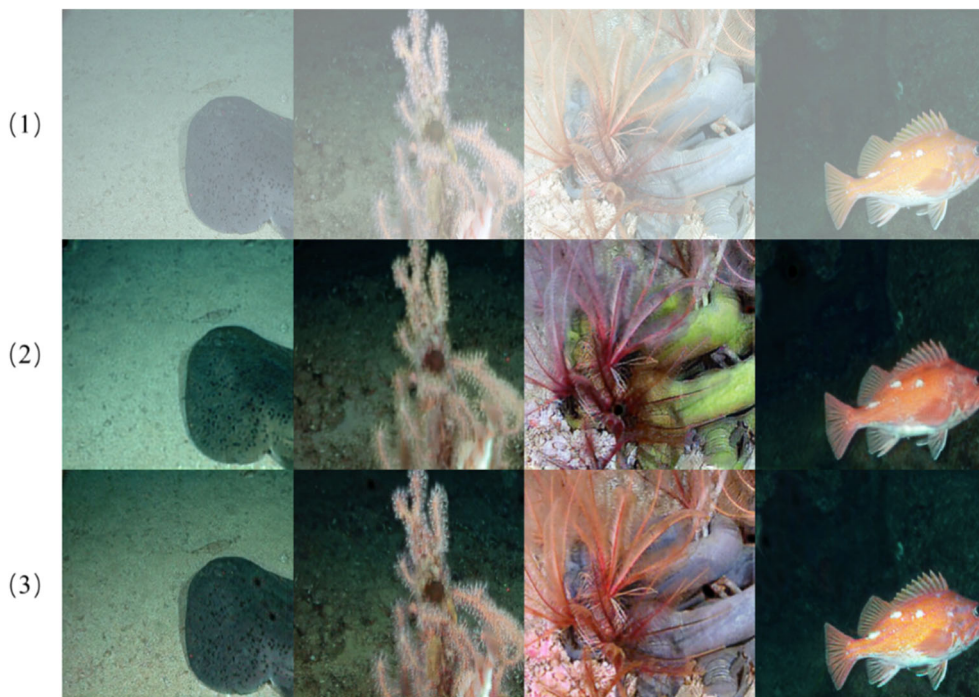
corresponds to the blurred underwater images, the second row corresponds to the results obtained by using the nearest neighbor interpolation which is a linear interpolation method, and the third row corresponds to the results obtained by the transposed convolution method. The former method exhibits color distortion in the object area of the third image, and the generated images are not as clear as the ones produced by the transposed convolution. After comparing the performance of the two methods, we select the transposed convolution method to realize the up-sampling operation.

3.3 Local auxiliary discriminator

Underwater visual enhancement is an important research area with broad significance. The enhanced images are beneficial to the subsequent target recognition or navigation tasks. We focus on the object area which may only occupy a small proportion of an underwater image while the global discriminator focuses on the overall effect. Hence, the clarity of the local areas cannot be judged accurately by the global discrimination. The basic structure of the global discriminator D_C corresponds to a simple binary classification network with only 8 convolutional layers, and it finally connects with a sigmoid layer to output the prediction result. In order to improve the capability of the global discriminator, a local auxiliary discriminator D_{aux} is designed to identify the object area.

Only one local image is randomly cropped for each input image of D_C . We suppose that the local image u_{ij} is fed to D_{aux} for training, which is cropped from the image u , where i and j denote the fixed height and fixed width of the cropped local image, $i = 64, j = 64, u \in \{c_s, G_{B2C}(b_t)\}$. The prediction result

Fig. 2 Results obtained from two different up-sampling methods. (1) the blurred underwater images, (2) the result images obtained from the nearest neighbor interpolation method, and (3) the result images produced by the transposed convolution method



$D_{aux}(u_{ij})$ denotes the output of D_{aux} , the prediction result of the global discriminator D_C is $D_C(u)$, and the final discrimination result is:

$$Pre(u) = \alpha D_c(u) + \beta D_{aux}(u_{ij}) \tag{1}$$

Specifically, α and β denote the weights that control the effect of each item of the prediction result, and $\alpha + \beta = 1$. With respect to the visual enhancement task, the backward process of generating blurred underwater images from the clear underwater images does not exhibit research significance, which is only trained as an auxiliary process. In order to simplify the proposed model, the local auxiliary discriminator is not introduced in the backward process.

3.4 Feedback mechanism

The purpose of this study is to improve the quality of the generated images while the overall contents of the images are not changed, that is to say, generate more useful information. Therefore, we introduce a feedback mechanism in the forward process to control the quality of the generated images. Specifically, we design a quality loss function to evaluate the quality of each generated image and feed back to the generation module to regulate image details. During the training process, G_{B2C} continuously receives the feedbacks from the quality loss function and the generated images get clearer and clearer.

We assume that X denotes a generated image and Y denotes its corresponding real image. Similarity between the generated

image and the corresponding real image can be calculated to measure the quality of the image X .

$$F_{qual}(X) = M(X, Y) \tag{2}$$

Where $M(\cdot)$ denotes a function to compute the pixel difference between the real image and the generated image. The quality loss between the image X and the image Y involves calculating the difference between each corresponding pixel of the image X and the image Y and then averaging the sum of the differences. This can be expressed as follows:

$$M(x, y) = \frac{1}{n} \sum_{pix=1}^n \|x_{pix} - y_{pix}\|^2 \tag{3}$$

Where pix denotes each pixel of an image, and n denotes the total number of pixels in the image.

3.5 Loss function

3.5.1 Adversarial loss

Adversarial loss is a loss function in the generative adversarial network. A deviation exists between the predicted result and the actual label of an input image, thereby resulting in a loss value. The generator is trained to generate images that are closer to clear underwater images while the discriminator is trained to accurately identify images coming from the generator as real or fake. The aim of the training of the generator and discriminator is to make the generated images look as

real as possible. The adversarial losses $L_A^{B2C}(G_{B2C}, D_C)$ and $L_A^{C2B}(G_{C2B}, D_B)$ are expressed as follows:

$$L_A^{B2C}(G_{B2C}, D_C) = E_{c_s \sim P_{data}(C)} [\log D_C(c_s)] + E_{b_t \sim P_{data}(B)} \left[\log \left(1 - D_C \left(G_{B2C}(b_t) \right) \right) \right] \tag{4}$$

$$L_A^{C2B}(G_{C2B}, D_B) = E_{b_t \sim P_{data}(B)} [\log D_B(b_t)] + E_{c_s \sim P_{data}(C)} \left[\log \left(1 - D_B \left(G_{C2B}(c_s) \right) \right) \right] \tag{5}$$

3.5.2 Local loss

A local auxiliary discriminator is introduced into Local-CycleGAN, and thus a local loss is added to the proposed network, which has a similar formulation as the adversarial loss.

$$L_{Local}^{B2C}(G_{B2C}, D_L) = E_{c_s \sim P_{data}(C)}^{u_{ij} \in c_s} [\log D_L(u_{ij})] + E_{b_t \sim P_{data}(B)}^{u_{ij} \in G_{B2C}(b_t)} \left[\log \left(1 - D_L \left(G_{B2C}(b_t) \right) \right) \right] \tag{6}$$

3.5.3 Cycle loss

With our proposed network, an image of B domain can be translated into itself after the mapping of G_{B2C} and the mapping of G_{C2B} successively. Similarly, an image of C domain can be translated into the original image after the mapping of G_{C2B} and mapping of G_{B2C} successively. Cycle loss $L_{cycle}(G_{B2C}, G_{C2B})$ and cycle loss $L_{cycle}(G_{C2B}, G_{B2C})$ are expressed as follows:

$$L_{cycle}(G_{B2C}, G_{C2B}) = E_{b_t \sim P_{data}(B)} \left[\|G_{C2B}(G_{B2C}(b_t)) - b_t\|_1 \right] \tag{7}$$

$$L_{cycle}(G_{C2B}, G_{B2C}) = E_{c_s \sim P_{data}(C)} \left[\|G_{B2C}(G_{C2B}(c_s)) - c_s\|_1 \right] \tag{8}$$

3.5.4 Domain-style loss

An image originally belonging to the C domain does not change its content after the mapping G_{B2C} , and thus it is still a C domain image. Under ideal conditions, $G_{B2C}(c_s) = c_s$. However, a deviation exists between the generated image $G_{B2C}(c_s)$ and the input image c_s . The deviation denotes the domain style loss which can be calculated as shown in Fig. 3.

Domain-style loss $L_{Dom}(G_{B2C})$ is expressed as follows:

$$L_{Dom}(G_{B2C}) = E_{c_s \sim P_{data}(C)} \left[\|G_{B2C}(c_s) - c_s\|_1 \right] \tag{9}$$

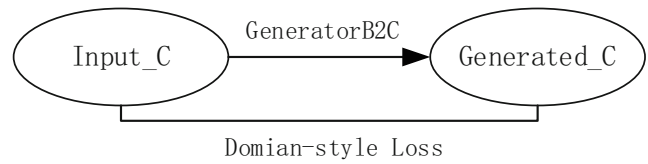


Fig. 3 Domain-style Loss

3.5.5 Quality-monitor loss

We define clear underwater images as standard images, and similarity between a generated image and its corresponding standard image can be measured. Therefore, we introduce the quality-monitor loss to monitor the quality of the generated image and feed it back to the generator.

Referring to Eq. 2, the quality-monitor loss is expressed as follows:

$$L_{qul}(G_{B2C}) = M(G_{B2C}(b_t), c_s) \tag{10}$$

From Eqs. 3 and 10, we can obtain the following expression:

$$L_{qul}(G_{B2C}) = \frac{1}{n} \sum_{pix=1}^n \left\| G_{B2C}(b_t)_{pix} - c_s_{pix} \right\|^2 \tag{11}$$

Given the quality monitoring in the forward process, the generated images become increasingly real and clear.

3.5.6 Objective function

The objective function denotes the sum of the adversarial loss, the local loss, the cycle loss, the domain-style loss, and the quality-monitor loss, and this is expressed as follows:

$$L = \min_{G_{B2C}, G_{C2B}} \max_{D_B, D_C} \left(\gamma_1 (L_A^{B2C}(G_{B2C}, D_C) + L_A^{C2B}(G_{C2B}, D_B)) + \gamma_2 L_{Aux}^{B2C}(G_{B2C}, D_{aux}) + \gamma_3 (L_{cycle}(G_{B2C}, G_{C2B}) + L_{cycle}(G_{C2B}, G_{B2C})) + \gamma_4 L_{Dom}(G_{B2C}) + \gamma_5 L_q(G_{B2C}) \right) \tag{12}$$

Specifically, $\gamma_1, \gamma_2, \gamma_3, \gamma_4,$ and γ_5 denote hyper-parameters that control the weight of each loss value. Based on the training experience, our proposed model yields better performance when $\gamma_1 = 1, \gamma_2 = 1, \gamma_3 = 10, \gamma_4 = 1,$ and $\gamma_5 = 0.5$. In the process of continuously training to determine the optimal solution of the objective function, the generated images become closer to the real images.

4 Experiments

4.1 Dataset and parameters

Training a network requires a large number of samples to ensure the performance of the model and avoid over-

fitting. This holds for training Local-CycleGAN as well. In the study, 694 pairs of blurred and clear underwater images are curated using Fish Dataset and the images extracted from the underwater videos downloaded from the NATURE FOOTAGE website. Given that the number of underwater samples is not sufficient, we adopt transfer learning to train our model. Specifically, 1449 pairs of in-air images from the NYU dataset are selected to pre-train the model. We then fine-tune the model with the paired underwater images. For an end-to-end full convolution network, our network can accept input images of any size to facilitate comparison with other algorithms, and the input images are uniformly resized to 320×320 . Our network uses the Adam method with a learning rate of 2×10^{-5} to train 50 epochs with the in-air dataset and then with a learning rate of 2×10^{-5} to train another 50 epochs with the underwater dataset. The β value of Adam is set to 0.5.

4.2 Results

In Fig. 4, the first row corresponds to five blurred images, the second row corresponds to the enhancement results obtained from our proposed method, and the third row corresponds to the clear images. The first column corresponds to in-air images while the last four columns correspond to the underwater images of green style, white style, dark style, and blue style. The proposed model is pre-trained using in-air datasets, and thus it is evident that Local-CycleGAN is effective in deblurring both the in-air foggy images and the blurred underwater images. As shown in Fig. 4, the enhanced images are slightly darker, and this is because most of the underwater images used for training correspond to the green style.

5 Analysis

5.1 Quality assessment

5.1.1 Qualitative evaluation

In our study, CycleGAN and MUNIT (Huang et al.) [29] are used as the baseline models and trained with the same datasets as used in training our proposed Local-CycleGAN. All models are trained by 100 epochs. We have also compared with other existing algorithms including the Retinex algorithm, the CLAHE algorithm, the DCP algorithm, and the ACE algorithm. The results are illustrated in Fig. 5 where the first row shows the blurred underwater images with different characteristics, and the rows 2 to 8 demonstrate the enhancement images obtained from the Retinex algorithm, the CLAHE algorithm, the DCP algorithm, the ACE algorithm, the CycleGAN, the MUNIT and our proposed Local-CycleGAN. The Retinex algorithm exhibits a high performance on the third, fourth, and fifth images. Other enhanced images are significantly distorted and over-processed, and this will increase the difficulty for dark object recognition. The CLAHE algorithm exhibits defogging effects on the original blurred images while generating problems like color distortion and significant graying. The DCP algorithm shows optimal enhancement performance on the second and fourth images without significant color bias although it does not evidently affect images of blue or green style. This is related to the dark channel prior principle. Dark channel prior implies that each local area in a sharp image is likely to exhibit shadows, or things of bright color or black, and the fog makes the original image fade. The DCP algorithm is only used to eliminate that foggy layer and does not focus on the color correction problem. The ACE algorithm exhibits the optimal color correction effect on the first five images albeit with relatively high

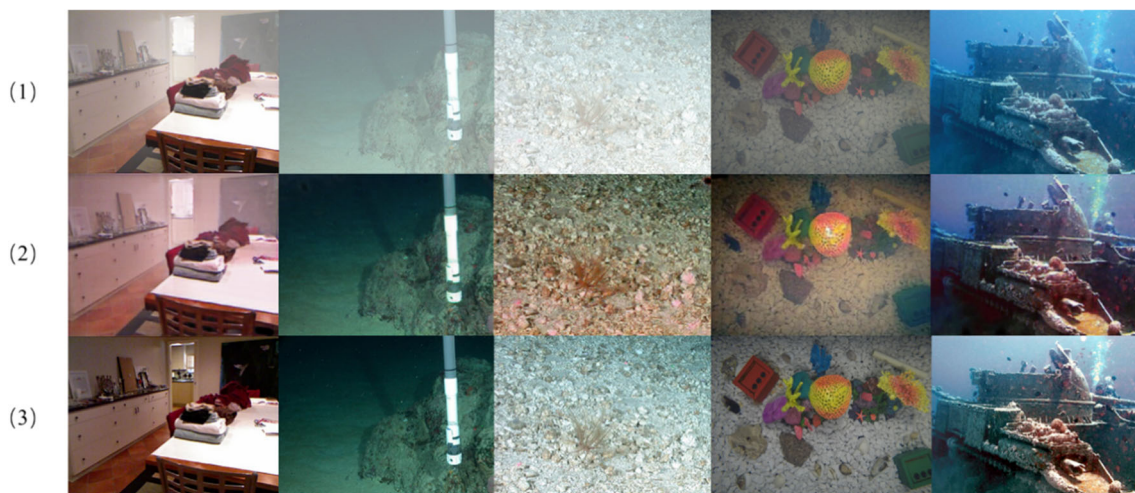


Fig. 4 Image enhancement results obtained from our proposed Local-CycleGAN. (1) the In-air and blurred underwater images, (2) the enhanced images generated by our proposed method, and (3) the in-air and clear underwater images

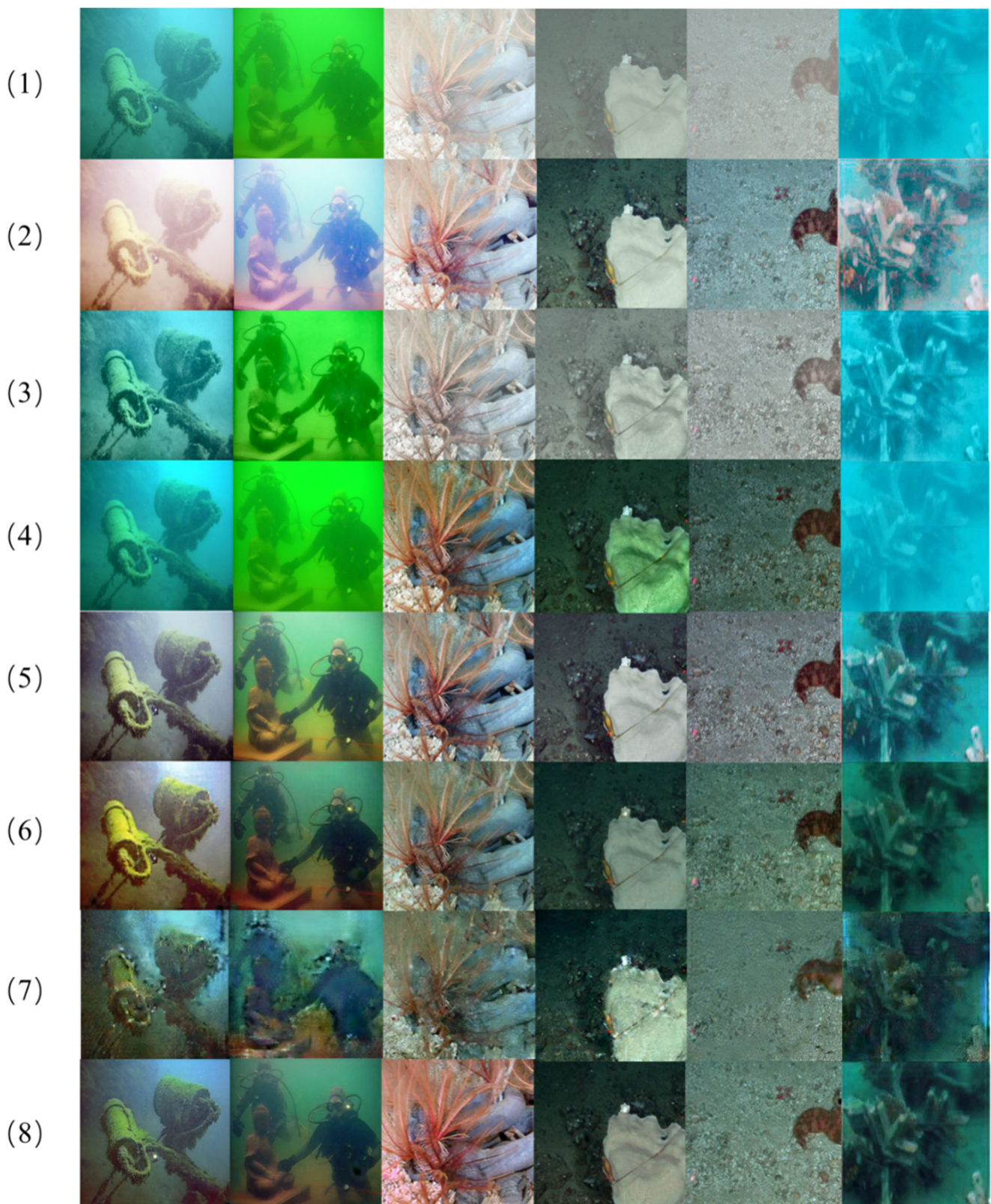


Fig. 5 Comparison of the performance of different algorithms for image enhancement. (1) The blurred underwater images, (2) the results produced using the Retinex algorithm, (3) the results generated by the CLAHE algorithm, (4) the results obtained from the DCP algorithm, (5)

the results produced by the ACE algorithm, (6) the results generated by the CycleGAN, (7) the results obtained from the MUNIT, and (8) the results from our proposed Local-CycleGAN

contrast. The enhancement of the sixth image is not evident, and additional noise and distortion are observed. The artifacts and noise of images generated by the CycleGAN model are significant, and the overall tone is dark. The results of MUNIT are very poor and many details are lost, which may be due to the relatively complex model and slow fitting process. However, our Local-CycleGAN demonstrates good enhancement performance on all blurred underwater images, and it also outperforms the ACE algorithm in enhancing the sixth image. Generally, the ACE and Local-CycleGAN exhibit the optimal performance among the aforementioned methods in terms of visual effect.

5.1.2 Quantitative evaluation

Metrics of image quality assessment include mean square error (MSE), peak signal-to-noise ratio (PSNR), structural similarity (SSIM), and feature similarity (FSIM). We select 30 underwater images of different styles to evaluate the aforementioned methods. As shown in Table 1 more useful details are omitted in the enhanced underwater images produced from the the Retinex and CLAHE algorithms. The DCP algorithm yields the highest PSNR score, the highest FSIM score and the lowest MSE among all methods. It is evident that the DCP still exhibits specific advantages in handling underwater images. With the CycleGAN and MUNIT, the lowest FSIM score is observed. This indicates that some features are lost in the image generation process. Our Local-CycleGAN shows the highest SSIM score and relatively good scores for other metrics. Table 1 demonstrates that the proposed method has good overall performance when assessed using the metrics. The metrics from CycleGAN look good too, however, it generates significant artifacts and noise in the enhanced images as explained in the Qualitative Evaluation section.

5.2 Ablation study

An ablation study has been done to prove the effectiveness of our proposed model equipped with the auxiliary discriminator and the quality feedback mechanism. We have conducted different experiments for the ablation study as shown in Table 2. The experimental results are shown in Fig. 6 and the results of quantitative evaluation on ablation study are shown in Table 3.

Table 1 Results of quantitative evaluation on several algorithms

Indicator	Retinex	CLAHE	DCP	ACE	CycleGAN	MUNIT	Local-CycleGAN
MSE	3589.49	5912.07	1055.49	1862.32	1156.34	3340.41	1181.05
PSNR	14.82	12.35	22.69	16.87	21.85	14.53	22.13
SSIM	0.85	0.76	0.93	0.88	0.93	0.82	0.96
FSIM	4.64	4.64	4.91	4.56	4.45	3.94	4.66

Table 2 The experiments for the ablation study

Experiment	L_{cycle}	L_{Local}^{B2C}	L_{qui}
A	✓		
B	✓	✓	
C	✓	✓	✓

In Experiment A, we directly remove L_{Local}^{B2C} and L_{qui} from Eq. 12. Then we add L_{Local}^{B2C} in Experiment B and add L_{qui} in Experiment C step by step. With each experiment, we trained the machine learning model by 100 epochs using the same dataset. Experiments B and C are designed respectively to explore the contribution of local auxiliary discriminator and the quality feedback mechanism. In Fig. 6, the first column shows the blurred underwater images, the columns 2 to 4 illustrate the results of the experiments A, B, and C, which are listed in Table 2. Compared with the results obtained from the experiment A, the experiment B can generate fewer artifacts in darker local areas. Compared with the results from the experiment B, the details of objects in the images generated from the experiment C are clearer and richer, such as the crab in the first row and the patterns of starfishes in the second row. As can be seen from the Table 3, the quantitative results of experiment C get the best scores for all metrics. This proved that our proposed model is effective in the enhancement of blurred underwater images.

5.3 General adaptability analysis

5.3.1 Underwater images with different colors

Underwater images taken in different water areas or depths show different color styles. We select several typical underwater images of yellow, blue, cyan, and green styles, and use the ACE algorithm and Local-CycleGAN to enhance them. The enhancement results are shown in Fig. 7. The ACE algorithm performs well in terms of details. However, the color of the first image is excessively saturated, a few areas of the second image show an over-exposure phenomenon, and the third image exhibits an evident red tone. Color bias is observed in the target area of the fourth image. Although the ACE algorithm demonstrates promising performs on the

Table 3 Quantitative evaluation on ablation study

Indicator	MSE	PSNR	SSIM	FSIM
A	506.82	22.11	0.97	0.58
B	259.25	24.39	0.99	0.60
C	200.50	25.60	0.99	0.62

whole image enhancement, its performance is limited in the color correction of these images. Comparing to the ACE algorithm, less details are observed in the enhanced images generated from our proposed Local-CycleGAN. This is mainly because traditional algorithms like the ACE enhance the original pixels of an image while Local-CycleGAN regenerates every pixel. The latter is more difficult. However, the targets in the Local-CycleGAN enhanced images can be easily identified, and the image color is more realistic and adapts to human visual perception. In the third image, a red area is also observed, which is smaller than that in the image enhanced by the ACE algorithm. Generally, when there are sufficient and diverse training sample, the proposed method can be trained to suit most image styles and produces balanced enhancement results in terms of image details and color.

5.3.2 Underwater images with different turbidity

In the operation of an underwater machine, water can be stirred up, and this may lead to increased turbidity. When the machine stops working, the suspended particles in the water will slowly settle down, and the turbidity will gradually decrease. Thus, the turbidity can vary at different times, thereby affect the captured underwater images. Experiments have been conducted to test our proposed Local-CycleGAN on the underwater images with different turbidity, and the experimental results are shown in Fig. 8. The first and third rows correspond to the original blurred underwater images selected from TURBID 3D dataset, and the second and fourth rows correspond to the images generated by the Local-CycleGAN model. Although the visual enhancement effect is directly relevant to the original images, the proposed method exhibits good robustness when enhancing the blurred underwater images with different turbidity. Furthermore, when compared with the original blurred underwater images, the clarity, brightness, and contrast of the enhanced images are significantly improved.

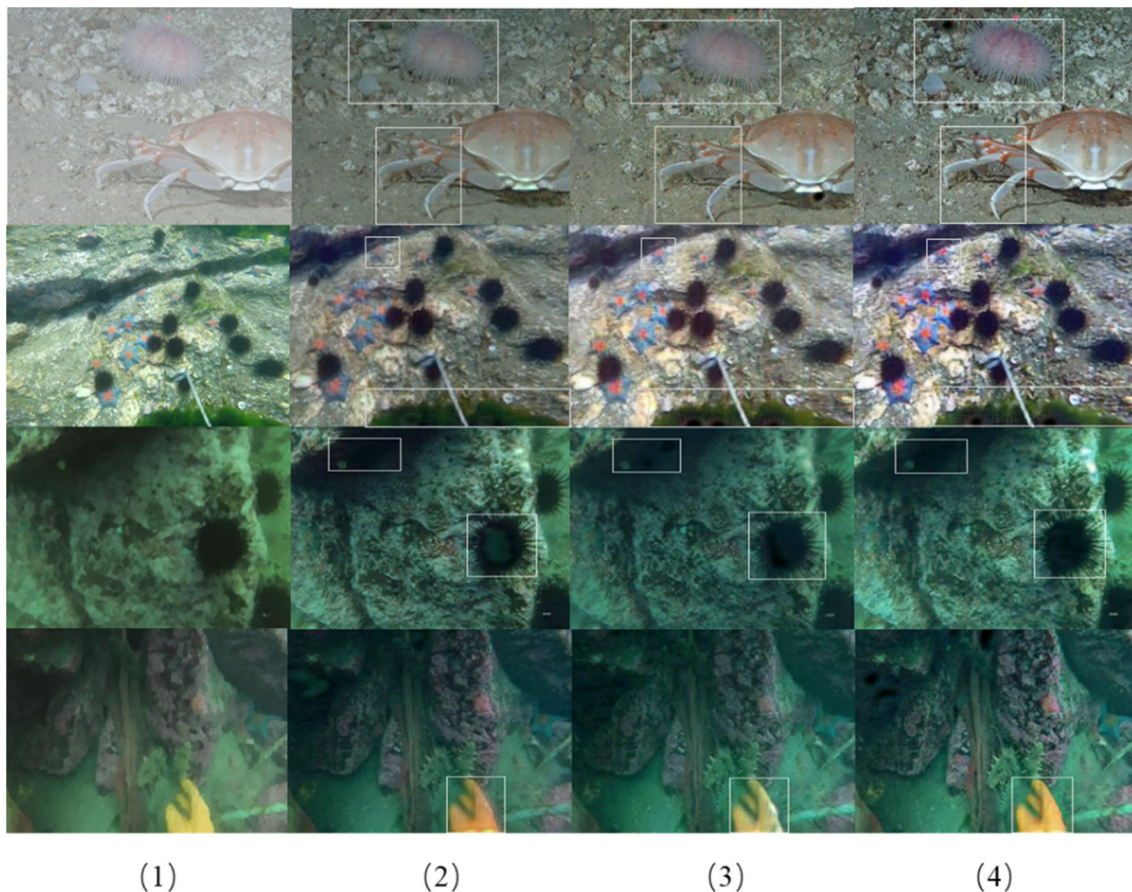


Fig. 6 Results of ablation study. (1) the blurred underwater images, (2)(3)(4) the results of the experiments A, B, and C, which are listed in Table 2

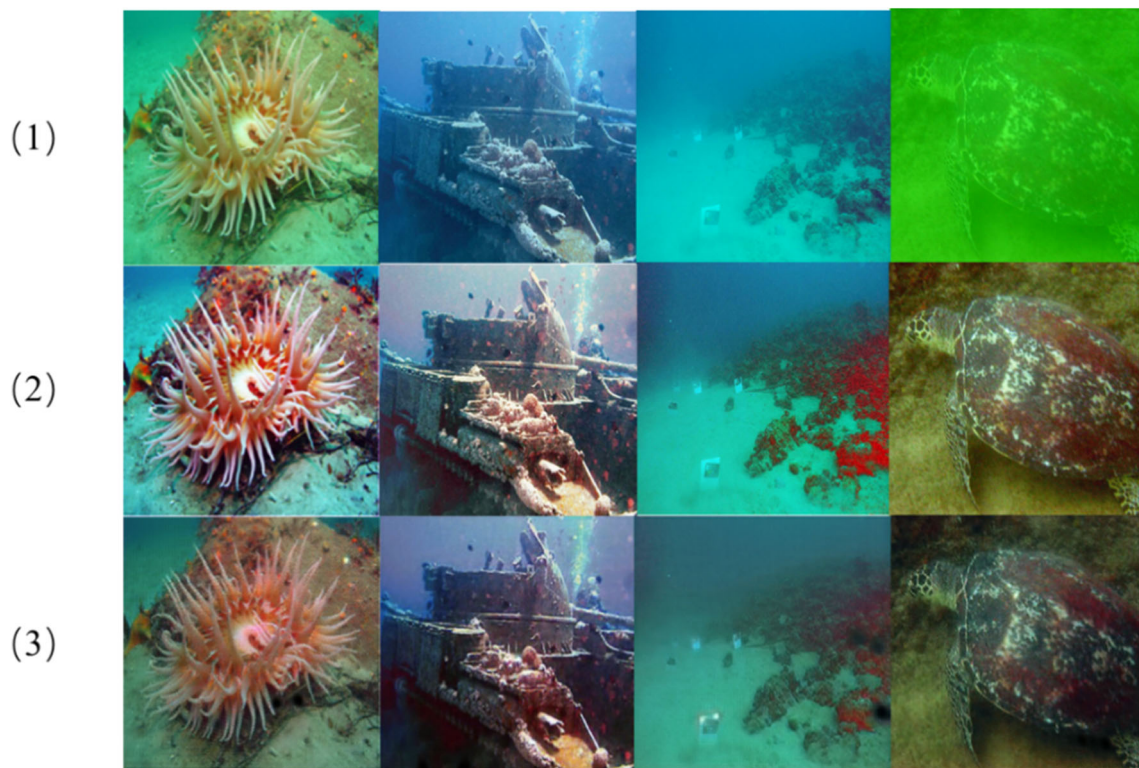


Fig. 7 Comparison of the performance of the ACE algorithm and the proposed method in enhancing images of different styles. (1) The blurred underwater images, (2) the enhanced images produced from the ACE

algorithm, and (3) the enhanced images generated from our proposed Local-CycleGAN

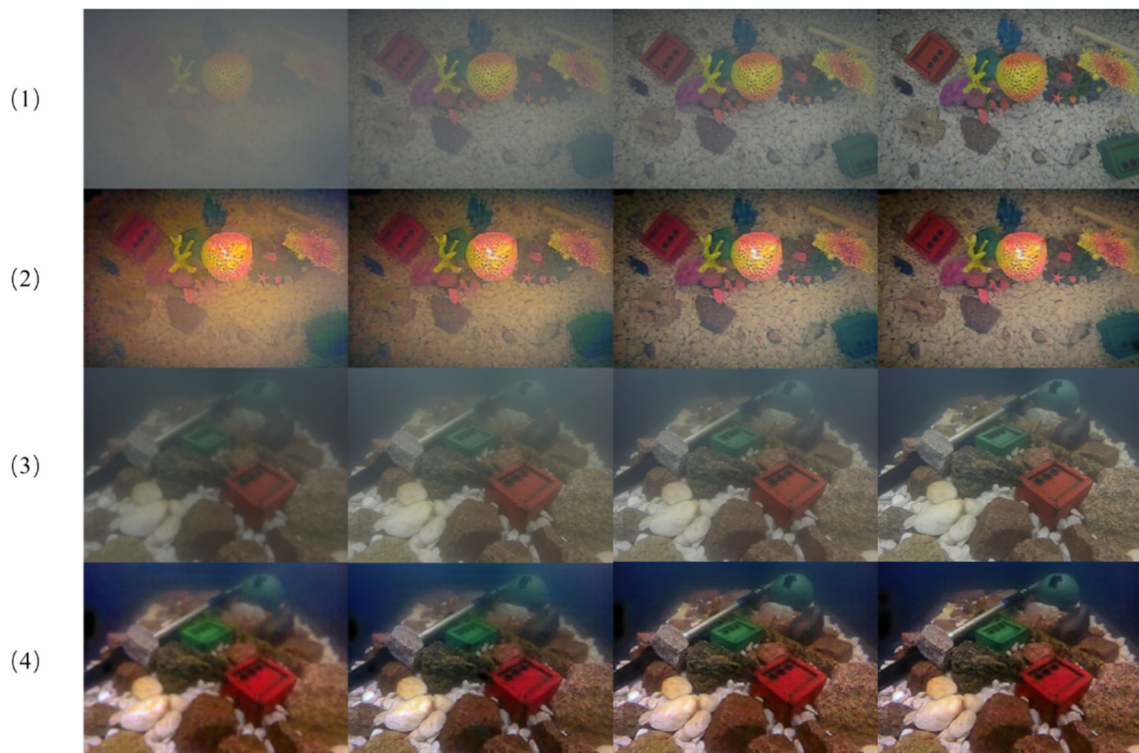


Fig. 8 Comparison of the blurred underwater images with different turbidity and their corresponding images enhanced by the proposed Local-CycleGAN. (1) the blurred underwater images, (2) the enhanced

images by our proposed method, (3) the blurred underwater images, and (4) the enhanced images generated by the proposed method

6 Conclusions

We present a Local-CycleGAN network to enhance underwater images, which combines a local discriminator and a global discriminator to constrain the global and local visual effects of enhanced images. To generate better quality images, we use a feedback mechanism to control the training process of the proposed network, and formulate the total loss function with the quality-monitor loss, local loss, and domain-style loss. The ablation study has demonstrated the performance of our proposed method. Extensive experiments have been carried out and the experimental results show that the proposed network outperforms other existing algorithms, not only in the image quality, but also in adaptability. That said, our experimental results also reveal that the images generated by our proposed Local-CycleGAN may result in slight loss of image details in some cases. Future work will focus on further improvement of the performance of the proposed network, especially focusing on the preservation of image details.

Compliance with ethical standards

Declarations This research was funded by the National Key R&D Program of China (Grant No. 2018YFC0810500) and the Scientific, the Fundamental Research Funds for the Central Universities (Grant No. RF-TP-20-009A3) and Technological Innovation Foundation of Shunde Graduate School, USTB (Grant No. BK19BE003).

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

- He KM, Sun J, Tang X (2009) Single image haze removal using Dark Channel prior. *IEEE Trans Pattern Anal Mach Intell* 33(12): 1956–1963
- Land EH (1977) The Retinex theory of color vision. *Sci Am* 237: 108–128
- Getreuer P (2012) Automatic color enhancement (ACE) and its fast implementation. *Image Process Line* 2:266–277
- Lan G, Jean P-A, Mehdi M, et al, 2014. Generative adversarial networks[J]. arXiv:1406.2661
- Zhu JY, Park T, Isola P et al (2017) Unpaired image-to-image translation using cycle-consistent adversarial networks. arXiv: 1703.10593
- Yan XC (2009) A new method for underwater image enhancement based on local complexity. *Modern Manuf Eng* (12):101–103
- Yang WZ, Xu YL, Qiao X et al (2016) Method for image intensification of Underwater Sea cucumber based on contrast-limited adaptive histogram equalization. *Trans Chin Soc Agric Eng* 32(06):197–203
- Zhang S, Wang T, Dong JY et al (2017) Underwater image enhancement via extended multi-scale Retinex. *Neurocomputing*. 245:1–9
- Javier P, Mitch B, Stefan BW et al (2020) Recovering depth from still images for underwater Dehazing using deep learning. *Sensors* 20(16):4580
- Ho SL, Sang WM et al (2020) Underwater image enhancement using successive color correction and Superpixel Dark Channel prior. *Symmetry* 12(8):1220
- Li CY, Guo JC, Cong RM, Pang YW, Wang B (2016) Underwater image enhancement by Dehazing with minimum information loss and histogram distribution prior. *IEEE Trans Image Process* 25(12): 5664–5677
- Zhang MH, Peng JH (2018) Underwater image restoration based on a new underwater image formation model *IEEE access* 6:58634–58644
- Ma XM, Chen ZH, Feng ZP 2019 Underwater image restoration through a combination of improved Dark Channel prior and gray world algorithms. *J Electron Imaging* 28(5)
- Song W, Wang Y, Huang DM et al (2018) Combining background light fusion and underwater Dark Channel prior with color balancing for underwater. *Pattern Recogn Artificial Intell* 31(09): 856–868
- Tang ZQ, Zhou B, Dai XZ (2018) Underwater robot visual enhancements based on the improved DCP algorithm. *Robot* 40(2): 222–230
- Yang SD, Chen ZH, Feng ZP 2019 Underwater Image Enhancement Using Scene Depth-Based Adaptive Background Light Estimation and Dark Channel Prior Algorithms *IEEE Access*
- Yu H, Li X, Lou Q, Lei C, Liu Z (2020) Underwater image enhancement based on DCP and depth transmission map. *Multimed Tools Appl* 79:20373–20390
- Jin WP, Guo JC, Qi Q (2019) Underwater image enhancement based on conditional generative adversarial network. *Signal Processing: Image Communication* 81
- Liu P, Wang GY, Qi H et al (2019) Underwater image enhancement with a deep residual framework. *IEEE Access* 7:94614–94629
- Lu JY, Li N, Zhang AY et al (2019) Multi-scale adversarial network for underwater image restoration. *Opt Laser Technol* 110:105–113
- Fu XY, Cao XY 2020 Underwater image enhancement with global–local networks and compressed-histogram equalization. *Signal Process: Image Commun* 86
- Li CY, Anwar S, Porikli F 2020 Underwater Scene Prior Inspired Deep Underwater Image and Video Enhancement. *Pattern Recogn* 98
- Radford A Metz L, Chintala S (2016) Unsupervised representation learning with deep convolutional generative adversarial networks. *The International Conference on Learning Representations* 10667: 97–108
- Li J, Skinner K, Eustice R et al (2018) WaterGAN: unsupervised generative network to enable real-time color correction of monocular underwater images. *IEEE Robotics Automation Letters* 3(1): 387–394
- Mirza M, Osindero S, 2014 Conditional generative adversarial nets. arXiv:1411.1784
- Chen X, Duan Y, Houthoofd R et al (2016) InfoGAN: interpretable representation learning by information maximizing generative adversarial nets. arXiv:1606.03657
- Li T, Qian RH, Chao D et al. 2018 BeautyGAN: instance-level facial makeup transfer with deep generative adversarial network. *ACM Multimedia Conference*
- Junho K, Minjae K, Hyeonwoo K et al 2019 U-GAT-IT: unsupervised generative Attentional networks with adaptive layer-instance normalization for image-to-image translation. arXiv:1907.10830
- Huang X., Liu M.Y., Belongie S., et al., 2018. Multimodal unsupervised image-to-image translation. *European Conference on Computer Vision, Multimodal Unsupervised Image-to-Image Translation*

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.