Signal Processing: Image Communication 81 (2020) 115723

Contents lists available at ScienceDirect

Signal Processing: Image Communication

journal homepage: www.elsevier.com/locate/image

Underwater image enhancement based on conditional generative adversarial network☆

Miao Yang [a,b,c,d], Ke Hu [a,*], Yixiang Du [a], Zhiqiang Wei [c], Zhibin Sheng [a], Jintong Hu [a]

[a] School of Electronic Engineering, Jiangsu Ocean University, Lianyungang 222005, China
[b] Marine Equipment and Technology Institute, Jiangsu University of Science and Technology, Zhenjiang 212000, China
[c] Qingdao National Laboratory of Marine Science and Technology, Qingdao 266100, China
[d] Department of Biomedical Engineering, Department of Electrical and Computer Engineering, University of Virginia, USA

ARTICLE INFO

Keywords:
Underwater image enhancement
Conditional generative adversarial networks
Adversarial learning
Deep learning

ABSTRACT

Underwater images play an essential role in acquiring and understanding underwater information. High-quality underwater images can guarantee the reliability of underwater intelligent systems. Unfortunately, underwater images are characterized by low contrast, color casts, blurring, low light, and uneven illumination, which severely affects the perception and processing of underwater information. To improve the quality of acquired underwater images, numerous methods have been proposed, particularly with the emergence of deep learning technologies. However, the performance of underwater image enhancement methods is still unsatisfactory due to lacking sufficient training data and effective network structures. In this paper, we solve this problem based on a conditional generative adversarial network (cGAN), where the clear underwater image is achieved by a multi-scale generator. Besides, we employ a dual discriminator to grab local and global semantic information, which enforces the generated results by the multi-scale generator realistic and natural. Experiments on real-world and synthetic underwater images demonstrate that the proposed method performs favorable against the state-of-the-art underwater image enhancement methods.

1. Introduction

Underwater images carry significant information of the underwater environment and are widely-used for exploring, recognizing, and monitoring underwater world. High-quality underwater images are desired in practical applications, which can guarantee the reliability of analysis and processing. However, suffering from the effects of light absorption and scattering, underwater images have obvious quality degradation issues, such as color distortion, low contrast, blurring, low light, uneven illumination, etc [1]. Concretely, the wavelength related attenuation causes color casts while the scattering brings a distance-dependent component into a clear image, which causes low-contrast of underwater images. Besides, underwater images also suffer from the noise and blurring induced by the cameras and motion. Several typical quality degraded underwater images are presented in Fig. 1. Such degraded underwater images limit their further applications in underwater object detection, underwater scene understanding, underwater robot inspection, underwater 3D reconstruction, to name a few [2,3].

To obtain high-quality underwater images, the existing methods range from polarization filters, reversing physical models to deep learning-based technologies [4]. Due to the diversity of underwater image degradation [5], the performance of current methods has room to be improved. In recent years, much attention has been paid to deep learning-based visual tasks. Moreover, the deep learning-based methods have achieved encouraging results in many visual tasks [6–12]. Nevertheless, the deep learning-based underwater image enhancement methods usually fall behind the traditional underwater image enhancement methods. The main reasons focus on the limited and monotonous training data as well as ineffective network structures [1]. More recently, the success of cGANs [13] inspires us to explore the performance of cGANs in underwater image enhancement. Additionally, we also propose to solve the issues of limited underwater image training data based on a kind of random sampling strategy and the blending of synthetic and real data. The promising performance of recent low-level visual tasks [14–17] also encourages our method. Experimental results on synthetic and real-world underwater image datasets demonstrate the robust and decent performance of our proposed method.

The contributions of this paper are summarized as follows.

- We propose a conditional generative adversarial network (cGAN) for underwater image enhancement. To the best of our

☆ No author associated with this paper has disclosed any potential or pertinent conflicts which may be perceived to have impending conflict with this work. For full disclosure statements refer to https://doi.org/10.1016/j.image.2019.115723.

* Corresponding author.
E-mail addresses: kexisibest@outlook.com, 2009000013@jou.edu.cn (K. Hu).

https://doi.org/10.1016/j.image.2019.115723
Received 4 October 2019; Received in revised form 8 November 2019; Accepted 25 November 2019
Available online 2 December 2019

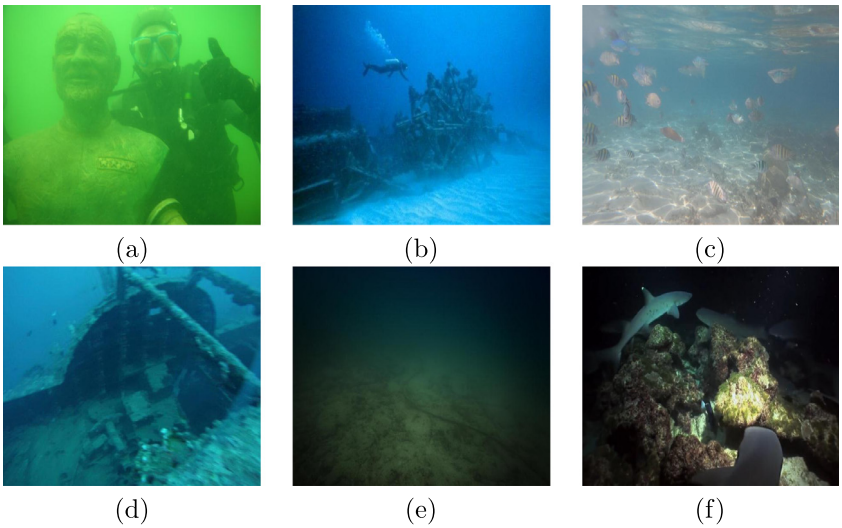

Fig. 1. Typical underwater images captured in diverse underwater scenes. (a) Greenish color deviation. (b) Bluish color deviation. (c) Low contrast. (d) Blurring. (e) Low light. (f) Uneven illumination.

knowledge, this is the first cGAN with dual discriminator for underwater image enhancement problem.
- To produce more realistic and natural results, we develop a dual discriminator which determines the reliability of generated results from different views.
- Our proposed method can improve the contrast and remove color casts of underwater images, and achieves state-of-the-art performance on diverse scenes.

The remainder of this paper is scheduled as follows. In Section 2, we briefly review the related works, including single image dehazing methods, underwater image enhancement methods, and cGANs in low-level visual tasks. In Section 3, the proposed cGAN for underwater image enhancement is presented and analyzed. In Section 4, extensive experiments are conducted to demonstrate the effectiveness of the proposed method and the advantages over current methods. In Section 5, we discuss and conclude this paper.

2. Related work

Similar to underwater images, the images captured under hazy day also suffer from the effect of light scattering. Some underwater image enhancement methods borrow some techniques or ideas from the research area of single image dehazing. Thus, single image dehazing is related to our underwater image enhancement. Besides, as the key network skeleton, the cGANs are also related to our work. In this section, we will briefly review the related works including single image dehazing, underwater image enhancement, and cGANs in low-level visual tasks.

2.1. Single image dehazing

As a pioneer work, He et al. [18] proposed the classical dark channel prior (DCP) for single image dehazing. The assumption of DCP is that haze-free outdoor images that at least one channel has some pixels with very low intensities (approximating zero). In [19], an optimization-based method which uses the contextual regularization to constrain the estimation of transmission was proposed. In [20], Zhu et al. proposed a color attenuation prior for the estimation of the scene depth from the camera to the object. Based on the assumption that colors of a haze-free image are well approximated by a few hundred distinct colors that form tight clusters in RGB space, a non-local prior was proposed in [21]. Wang et al. [22] proposed a single image dehazing method based on a physical model and the brightness components of the image by using a multi-scale retinex with color restoration algorithm. Based on the powerful modeling capabilities of Convolutional Neural Networks (CNNs), CNNs have been widely used for estimating the model parameters of haze formation models or the latent results. In [23], the CNN is used to estimate the transmission map which is a significant component of a haze formation model. Moreover, another important component, global atmospheric light, was estimated by traditional algorithms. In [24], a cascaded CNN was proposed to separately estimate the transmission and global atmospheric light. In [25], a unified framework was proposed for reconstructing the latent result in an end-to-end manner, where the transmission and atmospheric light are jointly estimated. For better perceptual visual quality, Li et al. proposed a perception-inspired single image dehazing network [26], which includes a subnetwork that reconstructs the latent dehazed image and a refinement network that further enhances the contrast and color of the dehazed result. With the emergence of GAN, Li et al. [27] introduced the cGAN into the image dehazing, which generates realistic clear images. Galdran et al. [28] provided the theoretical proof that Retinex on inverted intensities is a solution to the image dehazing task. Similarly, Wang et al. [29] proposed a single image dehazing method based on a physical model and the brightness components of the image by using a multi-scale Retinex.

2.2. Underwater image enhancement

In [30], a fusion-based underwater image enhancement method was proposed, where four weights are used to determine which pixel appears in the final image. In [31], the image dehazing algorithm was combined with the characteristics of underwater imaging for underwater color correction and scattering removal. Besides, this method considered the effect of artificial lighting in water. In [32], an underwater image restoration method was proposed, which considers the characteristics of the selective attenuation of different channels. Thus, the authors [32] separately processes different channels of an underwater image. In [33], the traditional dark channel prior originally proposed for single image dehazing was extended to underwater image enhancement. A hybrid method based on color correction and underwater image dehazing was proposed in [34]. This method corrects the color casts of the underwater image using image color prior and improves the visibility by a modified image dehazing algorithm. In [35], the contrast of underwater images was improved by recovering the colors associated with short wavelengths. In [36,37], an algorithm which not only enhances the contrast of underwater image but also restores its color was proposed. Based on the blurriness as a prior,

an underwater image enhancement method was proposed in [38]. Inspired by the morphology and function of the teleost fish retina, Gao et al. [39] proposed an underwater image enhancement method which solves the problems of underwater image degradation raised by the blurring and nonuniform color bias. With the rapid development of deep learning, the CNNs have been applied in underwater image enhancement in recent years. For example, Li et al. [40] proposed a CNN based underwater image color correction model based on synthetic underwater images generated in a weakly supervised learning manner. In [41], an underwater image color correction method based on weakly supervised color transfer was proposed. Such a method learns a cross-domain mapping function between underwater images and air images. In [42], a GAN-based underwater image restoration network was proposed, where an underwater image dataset by simulating underwater imaging model was provided. In [1], a fusion-based CNN for underwater image enhancement was proposed. This model was trained based on self-collected paired underwater images and the corresponding reference images. Recently, Li et al. [43] proposed an underwater image and video enhancement network based on synthetic underwater image datasets which simulate ten types of underwater scenes. More recently, a revised underwater image formation model was proposed [5] and then a physical model-based underwater image color correction method [44] was proposed according to this revised model.

2.3. Conditional Generative Adversarial Networks

Owing to the conditional information, the cGANs [13] have better stability and more powerful representation capability than the original GANs [45], especially for image enhancement and restoration tasks, where the realistic-looking images are generated. For example, Li et al. [27] proposed an end-to-end trainable cGAN for single image dehazing, where the generator is an encoder–decoder architecture. Furthermore, the authors developed a new loss function based on the pre-trained VGG features and an L_1 regularized gradient prior to generate realistic clear images and remove artifacts. In [46], a cGAN was proposed for image super-resolution task, where the adversarial loss enforces the results to the natural image manifold using a discriminator network. Zhang et al. [14] addressed the problem of image de-raining by a cGAN. This network constraints the generated results must be indistinguishable from the corresponding clear images, and thus, obtains superior results than the CNNs-based methods. In [47], Wang et al. proposed a method for synthesizing high-resolution photo-realistic images from semantic label maps based on cGANs. This method can generate high-resolution and high-quality results with a novel adversarial loss and the multi-scale generator and discriminator architectures.

3. Our approach

In this section, we introduce the proposed approach, including the network structure and parameter settings (generator and discriminator), loss function, and implementation details.

3.1. Overview of network architecture

As shown in Fig. 2, we present an overview of our cGAN for underwater image enhancement. Our cGAN consists of a multi-scale generator and a dual discriminator. In the phase of training, given a batch of underwater images as input, we first feed it to the multi-scale generator to learn the corresponding residual maps. After adding the residual maps to the input underwater images, the latent clear images are obtained. The results are sent to the dual discriminator, and then the discriminator determines the generated results are real or fake under different scales. In the inference phase, the input underwater image is enhanced through the multi-scale generator in an end-to-end fashion. In Fig. 3, we show several results generated by our multi-scale generator in the phase of inference. Next, we will illustrate the details of multi-scale generator and dual discriminator.

Multi-Scale Generator. The multi-scale generator of our cGAN consists of a multi-scale features extraction unit, features refine unit, and residual map estimation unit. In detail, the multi-scale feature extraction unit is built by three sets of multi-scale convolutions (with kernels of sizes $7 \times 7, 5 \times 5$, and 3×3, respectively). Each set of convolution includes five convolutional layers (numbers of filters are increased from 16 to 256 by a factor of 2) and each convolutional layer is followed by the non-linear activation ReLU. The purpose of such a multi-scale feature extractor is to grab different receptive fields which help to obtain the statistic information from inputs on various scales. After that, we down-sample the multi-scale features to half of the original feature size, concatenate them, and feed them to features refine unit. The down-sampling operation is to capture global features and reduce computational costs. After processing by three convolutional layers, the refined features are down-sampled by $2\times$ down-sampling and followed by four successive convolutional layers. The filter number of the convolutional layers in feature refine unit is 64. Then, we up-sample the feature maps to the original size and feed them to three successive convolutional layers (each layer has 64 filters). Finally, the residual map is estimated by a convolutional layer without non-linear activity, which can be used to achieve the final enhanced result by an element-wise addition operation. Please note that zero padding is applied to each convolutional layer to ensure the same input and output sizes. Moreover, except for the convolutional layers in the multi-scale features extractor, all convolutional layers have 3×3 kernel sizes.

Different from the commonly-used encoder–decoder and cGAN network structures [48], our generator contains a multi-scale features extraction unit which is designed for enriching the capability of network and also adapting fickle underwater scenes. Besides, our generator does not adopt the skip-connections owing to our light-weight and shallow structure.

Dual Discriminator. The dual discriminator consists of two sub-discriminators which share the same network structure but different weights. Besides, the sizes of inputs to these two sub-discriminators are different (original size and half of the original size). The purpose of the dual discriminator is to enable the discriminator to guide the generator to generate realistic images in both the global semantic level and local detail level. The reason behind such a design is that the existing discriminator is limited in guiding the generator to produce realistic details. Thus, we feed multi-resolution inputs to different discriminators to improve the visual quality of results. The network structure and parameter settings of these two sub-discriminators are the same as the discriminator used in [46] which is widely-used based on its effectiveness and efficiency. To be specific, the sub-discriminator contains eight convolutional layers with an increasing number of 3×3 filters, increasing from 64 to 512 by a factor of 2. To reduce the image resolutions, the strided convolutions are used. At last, the 512 feature maps are fed to two fully connected layers for predicting the probability of the inputs to be real or fake. Different from our multi-scale generator, the first convolution in the sub-discriminator followed by Leaky ReLU non-linear activation and the other convolutions are followed by batch normalization and Leaky ReLU. The last fully connected layer is followed by Sigmoid non-linear activation to predict the probability, which is commonly-used in image classification tasks. These two sub-discriminators are employed to guide our multi-scale generator.

3.2. Loss function

In the optimization process of our multi-scale generator network, we use four losses: an MSE loss L_m, a perceptual loss L_p, and two adversarial losses (L_{a1} and L_{a2}). To obtain better quantitative values, we introduce the MSE loss while the perceptual loss leads to the better visual quality of the enhanced results. These two adversarial losses

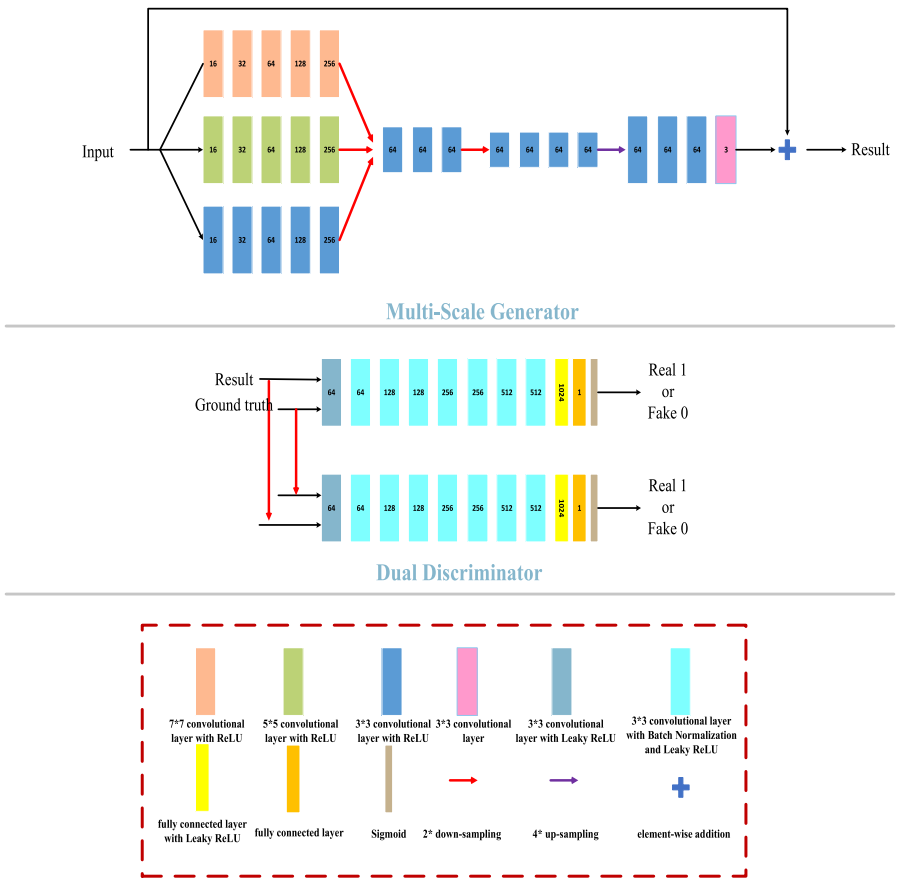

Fig. 2. An overview of the proposed cGAN for underwater image enhancement. It consists of two subnetworks: a multi-scale generator and a dual discriminator. **Top (multi-scale generator)**: Firstly, the multi-scale features are extracted by a multi-scale feature extractor which includes a set of $7 \times 7, 5 \times 5$, and 3×3 stacked convolutional layers. By 2× down-sampling operation (average pooling), the multi-scale features are integrated and the feature dimensions are reduced. After successively refining features, the features are resized to the same size as the input underwater image by 4× up-sampling operation (neighborhood linear interpolation). At last, the residual between the input underwater image and the corresponding clear underwater image is estimated. The latent result is achieved by simple element-wise addition between input and the corresponding residual map. **Middle (dual discriminator)**: The dual discriminator includes two sub-discriminators, where the scaled results and ground truth images are fed to them at the same time. The inputs with half-size are generated by 2× down-sampling operation on the generated result and corresponding ground truth image. **Bottom (network details)**: The detailed network components are listed. The numbers on the convolutional layers indicate the number of the filters. **Best viewed with zoom-in on a digital display**.

Fig. 3. Sample results by our approach. The input underwater images are captured in diverse underwater scenes.

make the enhanced results realistic and natural. The final loss function is the combination of above-mentioned four losses:

$$L_{final} = \alpha L_m + \beta L_p + \gamma(L_{a1} + L_{a2}), \quad (1)$$

where α, β, and γ are the weight of L_m, L_p, and $(L_{a1} + L_{a2})$, respectively. These weights are set according to preliminary experiments on the training data empirically. As follows, we will introduce these losses in details.

Adversarial Loss. Given the input underwater images $\{I_i, i = 1, 2, \ldots, N\}$ and the corresponding ground truth images $\{J_i, i = 1, 2, \ldots,$

$N\}$, the adversarial losses L_{a1} and L_{a2} can be expressed as:

$$L_{a1} = \frac{1}{N}\sum_{i=1}^{N} log[1 - D1(G(I))], \tag{2}$$

$$L_{a2} = \frac{1}{N}\sum_{i=1}^{N} log[1 - D2(G(I)_{half})], \tag{3}$$

where G is our multi-scale generator which tries to generate fake images to fool the dual discriminator, $D1$ and $D2$ are the sub-discriminators which try to distinguish fake images from the real images (ground truth images). $G(I)_{half}$ means the 2× down-sampled results.

In the phase of the alternate optimization, we solve the following optimization problem:

$$\min_G \max_{D1} \mathbb{E}_I[log(1 - D1(G(I)))] + \mathbb{E}_{I,J}[log D1(J)], \tag{4}$$

$$\min_G \max_{D2} \mathbb{E}_I[log(1 - D2(G(I)_{half}))] + \mathbb{E}_{I,J}[log D2(J_{half})], \tag{5}$$

where J_{half} means the 2× down-sampled ground truth images.

MSE Loss. We compute the difference between the generated results and the corresponding ground truth images, and the MSE loss is expressed as:

$$L_m = \frac{1}{N}\sum_{i=1}^{N}(I_i - J_i)^2, \tag{6}$$

Perceptual Loss. We measure the distance between the feature representations of the generated results and the feature representations of the corresponding ground truth images. The perceptual loss can be expressed as:

$$L_{con} = \frac{1}{N}\sum_{i=1}^{N} \|\phi_j(I_i) - \phi_j(J_i)\|, \tag{7}$$

where $\phi_j()$ is the jth convolution layer of the VGG19 network ϕ pretrained on the ImageNet [49].

3.3. Implementations

To the best of our knowledge, there is only one real-world underwater image enhancement dataset with corresponding reference images (i.e., UIEBD [1]) available. The UIEBD dataset includes 890 real-world underwater images captured under diverse scenes and the corresponding reference images are collected by subjective selection from the results of the state-of-the-art image enhancement algorithms. Specifically, we use 800 underwater images from UIEBD [1] as the training data and the rest images are treated as the testing data. However, we found that the UIEBD dataset [1] is too small to train our cGAN. To overcome this issue, we employ two strategies. First, in the process of training, we randomly sample the patches with fixed sizes of 64 × 64 from input underwater image and the corresponding ground truth image simultaneously. Second, we randomly sample 500 synthesized underwater images from the recent synthetic underwater image dataset [43] to supply our training data. This synthetic underwater image dataset includes ten subsets which simulate ten types of underwater scenes. We randomly sample 50 images from each subset. Thus, our training data contain real-world underwater images and synthetic underwater images. In this way, we augment our training data. In Fig. 4, we show several samples of our train data.

Our cGAN for underwater image enhancement is implemented on a computer with an Nvidia GTX 1080Ti GPU, Intel I7 6700 CPU 4.0 GHz and 32 GB RAM using the TensorFlow framework. We alternatively train the generator and discriminator using ADAM with the learning rate 0.0002, β_1 0.9 and β_2 0.999. We fix the learning rate in the entire training procedure. The batch size is set to 16. It takes around two days to optimize our cGAN. In addition, we measure the perceptual loss at layer relu5_4 of the VGG19 network [50]. The weights α, β, and γ are set to 0.001, 1, and 10 based on our training data, respectively.

Table 1
Quantitative evaluations on synthesized underwater image testing set in terms of PSNR and SSIM. The values in bold represent the best result while the underline means the second best result.

Metrics	RED [35]	UDCP [33]	UIBLA [38]	UWCNN [43]	Ours
PSNR	14.125	13.250	14.052	**19.512**	<u>17.715</u>
SSIM	0.6125	0.5375	0.5250	**0.7250**	<u>0.6552</u>

4. Experiments

We compare the proposed cGN for underwater image enhancement with several state-of-the-art underwater image enhancement methods on both synthetic and real-world underwater images. The competing methods are UDCP [33], RED [35], UIBLA [38], and UWCNN [43]. The results of these methods are generated by the source codes provided by the authors with the recommended parameter settings. In [43], ten UWCNN models trained for different types of water are provided. We run these ten models and produce the best result for real-world underwater images. For the synthetic underwater images, we randomly select ten underwater images from each type of scene [43] for testing. For real underwater images, we employ the rest 90 underwater images from [1] for testing. There is no overlap between training data and testing data.

4.1. Results on synthetic dataset

We first carry experiments on synthesized underwater images. In Fig. 5, we show several sample results by different methods. The synthesized underwater images have diverse color deviations and even invisible details. The RED [35] introduces pinkish color costs for the first image because it over-enhances the red channel of underwater image. The UDCP [33] shows limitation for the underwater image with invisible details and induces obvious artifacts. The visual results of UDCP [33] also indicates its brittle robustness. The UIBLA [38] leaves the haze on the result for the fourth image. This is because the blurring-prior of UIBLA [38] does not always hold in some cases. In comparison, the UWCNN [43] and our method can effectively remove the color deviations and recover the textures. For the second image, the UWCNN [43] tends to introduce artifacts on the wall. Please note that the training data of UWCNN [43] are similar to the testing data and we pick up the best results from the candidate results from ten models of UWCNN [43]. In contrast, our method only consists of one model and obtains comparable performance with the UWCNN [43].

We further compare the quantitative scores of different methods on synthesized underwater images. When the ground truth images are available, we employ two full-reference image quality assessment metrics to evaluate the performance of different methods. The commonly-used PSNR and SSIM [51] are used in this part. The higher PSNR and SSIM scores indicate better performance and also represent the results are more similar to the ground truth images. The average scores on the synthesized underwater image testing data under different metrics are reported in Table 1.

Observing Table 1, the results of UWCNN [43] and the proposed method achieve the best and second best performance in terms of PSNR and SSIM metrics. By contrast, the methods of RED [35], UDCP [33], and UIBLA [38] show limitations when they are used to process diverse underwater images. Additionally, it is reasonable that our method is a little lower than the UWCNN [43] on synthetic data. However, our method has better generalization capability for real-world underwater images. We will demonstrate the advantages of our method on real data in the next section.

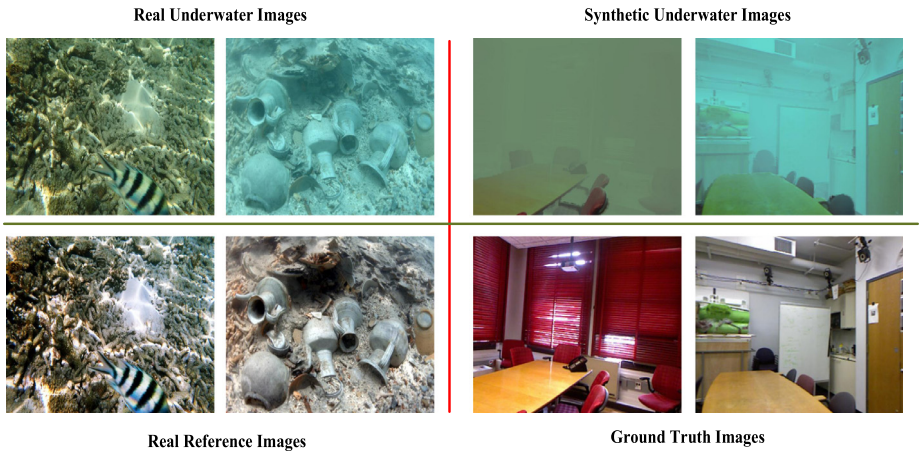

Fig. 4. Samples of training data.

Fig. 5. Results on synthetic underwater images. From top to bottom are underwater images, results of RED [35], results of UDCP [33], results of UIBLA [38], results of UWCNN [43], our results, and ground truth images.

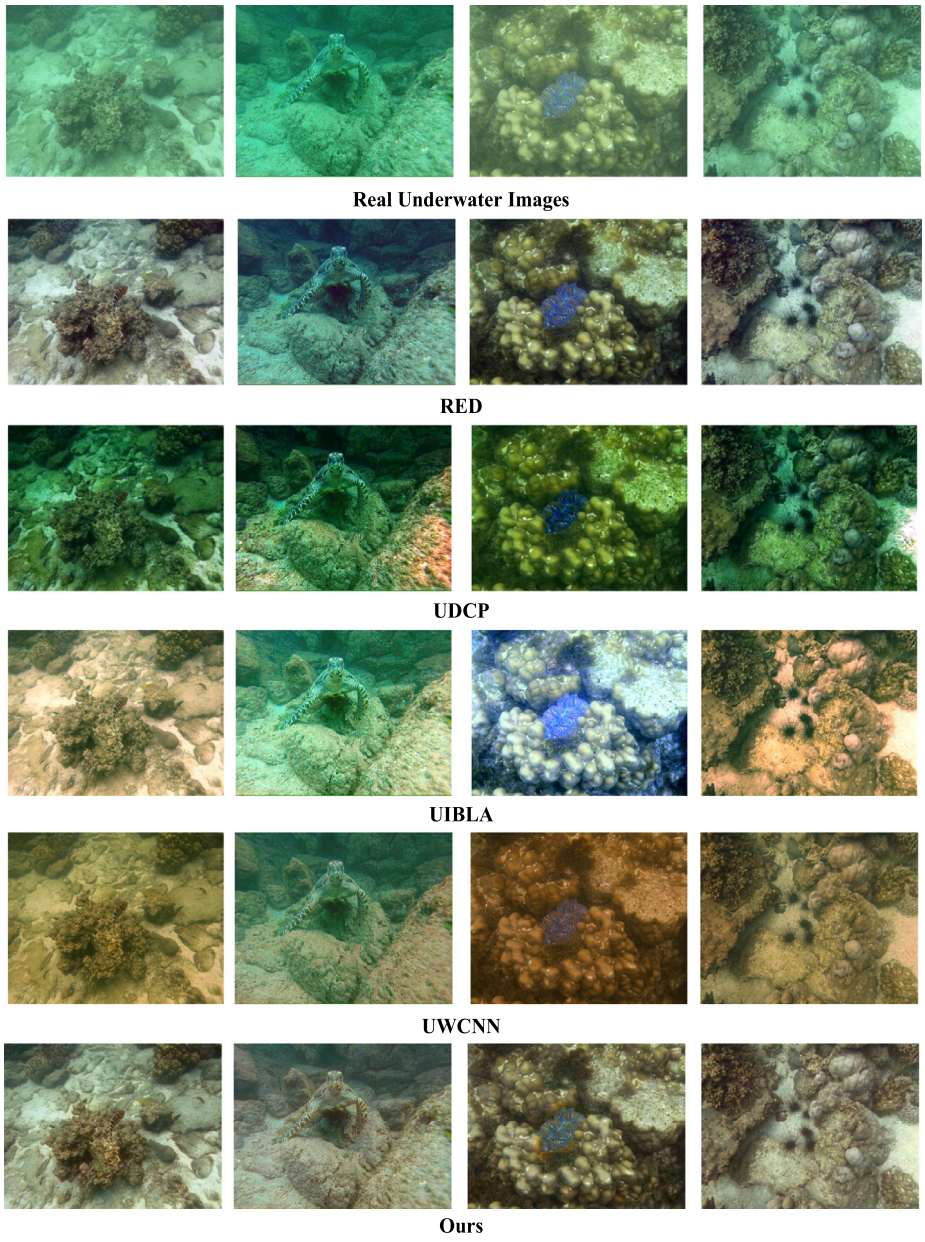

Fig. 6. Results on real-world underwater images. From top to bottom are underwater images with color casts, results of RED [35], results of UDCP [33], results of UIBLA [38], results of UWCNN [43], and our results.

4.2. Results on real-world dataset

We carry out the visual comparisons on real-world underwater images. To further analyze the advantages and disadvantages of different methods. We select two sets of typical underwater images from [1], including underwater images with color casts and underwater images with low contrast. The corresponding results are presented in Figs. 6 and 7.

As shown in Fig. 6, the underwater images have obvious color casts which might challenge the performance of color-based underwater intelligent systems. Compared with the original underwater images, our method can totally remove different color casts and preserve the details of original images. The RED [35] are only effective for part of color casts while the UDCP [33] makes the results greenish. In addition, the UWCNN [43] does not achieve the same performance as it is used in synthetic data. All in all, our method achieves visually pleasing results.

When coming to the underwater images with low contrast as shown in Fig. 7, the performance of the UWCNN [43] drops rapidly due to the limited generalization of synthetic training data used in [43]. The RED [35] and UIBLA [38] improve the contrast to some extent, but leave the haze on some regions. The UDCP [33] over-enhances the underwater images and even introduces color casts. In contrast, our method significantly improves the contrast and does not introduce over-enhancement and color casts.

5. Conclusion

In this paper, we proposed a cGAN for underwater image enhancement. We use a multi-scale generator for extracting larger receptive fields while the proposed dual discriminator well pushes the generator towards better visual quality. Also, we propose two strategies to augment the training data, which is useful for solving the issues of limited training data in low-level visual tasks. Experimental results show that our method achieves visually pleasing and realistic results and surpasses the state-of-the-art underwater image enhancement methods.

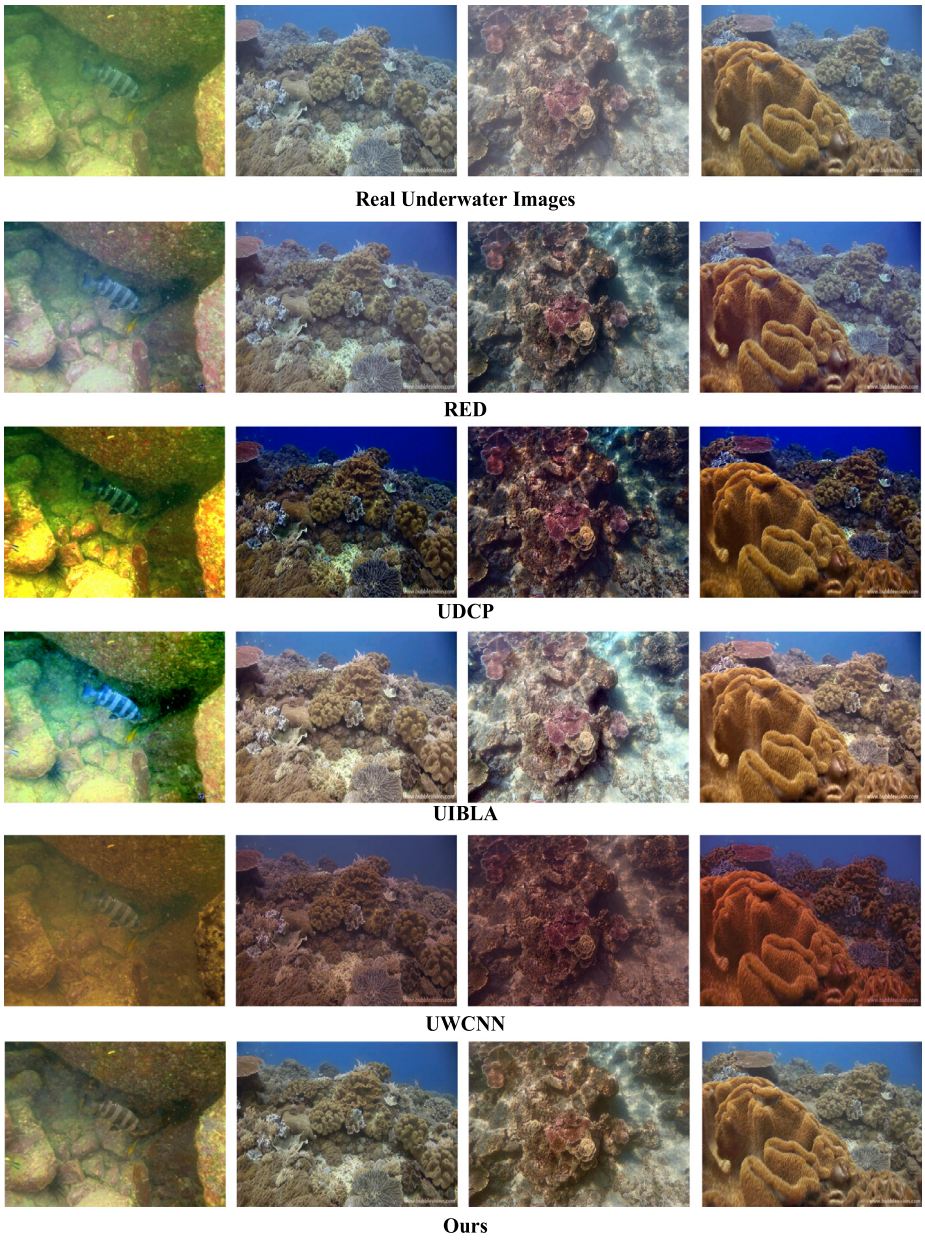

Fig. 7. Results on real-world underwater images. From top to bottom are underwater images with low contrast, results of RED [35], results of UDCP [33], results of UIBLA [38], results of UWCNN [43], and our results.

Although our method achieves superior performance than several state-of-the-art underwater image enhancement methods, it still exists several limitations for the challenging tasks. First, our method mainly focuses on removing the color casts and improving the contrast; however, real-world underwater images suffer from diverse degradations, such as noise and uneven light. Second, the real gap between the enhanced results and real-world clean underwater images is not clear due to the limited testing data and reliable evaluation metrics.

Acknowledgments

This work was supported in part by the National Natural Science Foundation under Grant 61601194, 11573011. Natural Science Foundation of Jiangsu province under Grant BK20191469, the "Double creation talents" science and technology deputy general manager project in Jiangsu Province (2017), and in part by Lianyungang "521" project, Lianyungang "Haiyan" funded project in China, and the project of Collaborative Innovation Center of Advanced Ship and Marine Equipment/Marine Equipment and Technology Institute, Jiangsu University of Science and Technology (HZ20190005).

References

[1] C. Li, J. Guo, W. Ren, R. Cong, J. Hou, S. Kwong, An underwater image enhancement benchmark dataset and beyond, IEEE Trans. Image Process. (2019).
[2] S. Anwar, C. Li, Diving deeper into underwater image enhancement: A survey, 2019, arXiv preprint arXiv:1907.07863.
[3] M. Yang, J. Hu, C. Li, G. Rohde, Y. Du, K. Hu, An in-depth survey of underwater image enhancement and restoration, IEEE Access 7 (2019) 123638–123657.
[4] M. Sheinin, Y. Schechner, The next best underwater view, in: Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit., CVPR, IEEE, 2017, pp. 1436–1443.
[5] D. Akkaynak, T. Treibitz, A revised underwater image formation, in: Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit., CVPR, IEEE, 2018, pp. 6723–6732.
[6] S. Anwar, C. Huynh, F. Porikli, Class-specific image deblurring, in: Proc. IEEE Int. Conf. Comput. Vis., ICCV, IEEE, 2015, pp. 495–503.
[7] C. Guo, C. Li, J. Guo, R. Cong, H. Fu, P. Han, Hierarchical features driven residual learning for depth map super-resolution, IEEE Trans. Image Process. 28 (5) (2019) 2545–2557.

[8] W. Wang, Y. Xu, J. Shen, S. Zhu, Attentive fashion grammar network for fashion landmark detection and clothing category classification, in: Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit., CVPR, IEEE, 2018, pp. 4271–4280.

[9] S. Anwar, C. Huynh, F. Porikli, Category-specific object image denoising, IEEE Trans. Image Process. 26 (11) (2017) 5506–5518.

[10] W. Wang, J.S.F. Guo, M. Cheng, A. Borji, Revisiting video saliency: A large-scale benchmark and a new model, in: Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit., CVPR, IEEE, 2018, pp. 4894–4903.

[11] C. Li, R. Cong, J. Hou, S. Zhang, Y. Qian, S. Kwong, Nested network with two-stream pyramid for salient object detection in optical remote sensing images, IEEE Trans. Geosci. Remote Sens. 57 (11) (2019) 9156–9166.

[12] R. Cong, J. Lei, H. Fu, W. Lin, Q. Huang, X. Cao, An iterative co-saliency framework for RGBD images, IEEE Trans. Cybernet. 49 (1) (2019) 233–246.

[13] M. Mirza, S. Osindero, Conditional generative adversarial nets, 2014, arXiv preprint arXiv:1411.1784.

[14] H. Zhang, V. Sindagi, V. Patel, Image de-raining using a conditional generative adversarial network, IEEE Trans. Circuits Syst. Video Technol. (2019).

[15] R. Cong, J. Lei, H. Fu, Q. Huang, X. Cao, C. Hou, Co-saliency detection for RGBD images based on multi-constraint feature matching and cross label propagation, IEEE Trans. Image Process. 27 (2) (2018) 568–579.

[16] C. Li, J. Guo, F. Porikli, Y. Pang, LightenNet: A convolutional neural network for weakly illuminated image enhancement, Pattern Recognit. Lett. 104 (2018) 15–22.

[17] R. Cong, J. Lei, H. Fu, Q. Huang, X. Cao, N. Ling, HSCS: Hierarchical sparsity based co-saliency detection for RGBD images, IEEE Trans. Multimedia 21 (7) (2019) 1660–1671.

[18] K. He, J. Sun, X. Tang, Single image haze removal using dark channel prior, IEEE Trans. Pattern Anal. Mach. Intell. 33 (12) (2011) 2341–2353.

[19] G. Meng, Y. Wang, J. Duan, S. Xiang, C. Pan, Efficient image dehazing with boundary constraint and contextual regularization, in: Proc. IEEE Int. Conf. Comput. Vis., ICCV, IEEE, 2013, pp. 617–624.

[20] Q. Zhu, J. Mai, L. Shao, A fast single image haze removal algorithm using color attenuation prior, IEEE Trans. Image Process. 24 (11) (2015) 3522–3533.

[21] D. Berman, S. Avidan, Non-local image dehazing, in: Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit., CVPR, IEEE, 2016, pp. 1674–1682.

[22] J. Wang, K. Lu, J. Xue, N. He, L. Shao, Single image dehazing based on the physical model and MSRCR algorithm, IEEE Trans. Circuits Syst. Video Technol. 28 (9) (2018) 2190–2199.

[23] W. Ren, S. Liu, H. Zhang, J. Pan, X. Cao, Single image dehazing via multi-scale convolutional neural networks, in: Proc. Eur. Conf. Comput. Vis., ECCV, Springer, 2016, pp. 154–169.

[24] C. Li, J. Guo, F. Porikli, H. Fu, Y. Pang, A cascased convolutional neural network for single image dehazing, IEEE Access 6 (2018) 24877–24887.

[25] B. Li, X. Peng, Z. Wang, J. Xu, D. Feng, Aod-net: All-in-one dehazing network, in: Proc. IEEE Int. Conf. Comput. Vis., ICCV, IEEE, 2017, pp. 4780–4788.

[26] C. Li, C. Guo, J. Guo, P. Han, H. Fu, R. Cong, PDR-Net: Perception-inspired single image dehazing network with refinement, IEEE Trans. Multimedia (2019) 1.

[27] R. Li, J. Pan, Z. Li, J. Tang, Single image dehazing via conditional generative adversarial network, in: Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit., CVPR, IEEE, 2018, pp. 8202–8211.

[28] A. Galdran, A. Gila, A. Bria, J. Corral, M. Bertalmio, On the duality between retinex and image dehazing, in: Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit., CVPR, IEEE, 2018, pp. 8212–8221.

[29] J. Wang, K. Lu, J. Xue, N. He, L. Shao, Single image dehazing based on the physical model and MSRCR algorithm, IEEE Trans. Circuits Syst. Video Technol. 28 (9) (2018) 2190–2199.

[30] C. Ancuti, C.O. Ancuti, Enhancing underwater images and videos by fusion, in: Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit., CVPR, IEEE, 2012, pp. 81–88.

[31] J. Chiang, Y. Chen, Underwater image enhancement by wavelength compensation and dehazing, IEEE Trans. Image Process. 21 (4) (2012) 1756–1769.

[32] C. Li, J. Guo, Y. Pang, S. Chen, J. Wang, Single underwater image restoration by blue-green channels dehazing and red channel correction, in: Proc. IEEE Int. Conf. Acoustics, Speech and Signal Process., ICASSP, IEEE, 2016, pp. 1731–1735.

[33] P. Drews, E. Nascimento, S. Botelho, M. Campos, Underwater depth estimation and image restoration based on single images, IEEE Comput. Graph. Appl. 36 (2) (2016) 24–35.

[34] C. Li, J. Guo, C. Guo, R. Cong, J. Gong, A hybrid method for underwater image correction, Pattern Recognit. Lett. 94 (2017) 62–67.

[35] A. Galdran, D. Pardo, A. Picn, A. Alvarez-Gila, Automatic red-channel underwater image restoration, Vis. Commun. Image Represent. 26 (2015) 132–145.

[36] C. Li, J. Guo, S. Chen, Y. Tang, Y. Pang, J. Wang, Underwater image restoration based on minimum information loss principle and optical properties of underwater imaging, in: Proc. IEEE Int. Conf. Image Process., ICIP, IEEE, 2016, pp. 1993–1997.

[37] C. Li, J. Guo, R. Cong, Y. Pang, B. Wang, Underwater image enhancement by dehazing with minimum information loss and histogram distribution prior, IEEE Trans. Image Process. 25 (12) (2016) 5664–5677.

[38] Y. Peng, P. Cosman, Underwater image restoration based on image blurriness and light absorption, IEEE Trans. Image Process. 26 (4) (2017) 1579–1594.

[39] S. Gao, M. Zhang, Q. Zhao, X. Zhang, Y. Li, Underwater image enhancement using adaptive retinal mechanisms, IEEE Trans. Image Process. 28 (11) (2019) 5580–5595.

[40] J. Li, K. Skinner, R. Eustice, M. Roberson, WaterGAN: unsupervised generative network to enable real-time color correction of monocular underwater images, IEEE Robot. Autom. Lett. 3 (1) (2017) 387–394.

[41] C. Li, J. Guo, C. Guo, Emerging from water: underwater image color correction based on weakly supervised color transfer, IEEE Signal Process. Lett. 25 (3) (2018) 323–327.

[42] X. Yu, Y. Qu, M. Hong, Underwater-GAN: Underwater image restoraiton via conditional generative adversarial network, in: Proc. Int. Conf. Pattern Recognit., ICPR, Springer, 2018, pp. 66–75.

[43] C. Li, S. Anwar, F. Porikli, Underwater scene prior inspired deep underwater image and video enhancement, Pattern Recognit. 98 (2019).

[44] D. Akkaynak, T. Treibitz, Sea-thru: A method for removing water from underwater images, in: Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit., CVPR, IEEE, 2019, pp. 1682–1691.

[45] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, Y. Bengio, Generative adversarial nets, in: Proc. Neural Infor. Process. Syst., NeurIPS, 2014, pp. 2672–2680.

[46] C. Ledig, L. Theis, F. Huszar, J. Callero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang, W. Shi, Photo-realistic single image super-resolution using a generative adversarial network, in: Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit., CVPR, IEEE, 2017, pp. 105–114.

[47] T. Wang, M. Liu, J. Zhu, A. Tao, J. Kautz, B. Ctanzaro, High-resolution image synthesis and semantic manipulation with conditional GANs, in: Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit., CVPR, IEEE, 2018, pp. 8798–8807.

[48] O. Ronneberger, P. Fischer, T. Brox, U-net: Convolutional networks for biomedical image segmentagion, in: Proc. Int. Conf. Med. Image Comput. Computer Assisted Intervention, MICCAI, Springer, 2015, pp. 234–241.

[49] J. Deng, W. Dong, R. Socher, L. Li, K. Li, L. Fei, Imagenet: a large-scale hierarchical image database, in: Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit., CVPR, IEEE, 2009, pp. 248–255.

[50] K. Simonyan, A. Zisserman, Very deep convolutional networks for large-scale image recognition, 2014, arXiv preprint arXiv:1409.1556.

[51] Z. Wang, A. Bovik, H. Sherikh, E. Simoncelli, Image quality assessment: from error visibility to structual similarity, IEEE Trans. Image Process. 13 (4) (2004) 600–612.