

Research Article

# Underwater Image Restoration and Enhancement via Residual Two-Fold Attention Networks

Bo Fu<sup>1,\*</sup>, Liyan Wang<sup>1</sup>, Ruizi Wang<sup>1</sup>, Shilin Fu<sup>1</sup>, Fangfei Liu<sup>1</sup>, Xin Liu<sup>2</sup>

<sup>1</sup>School of Computer and Information Technology, Liaoning Normal University, No. 1, Liu-shu-nan Street, Dalian, Liaoning, 116081, China

<sup>2</sup>School of Mathematics, Liaoning Normal University, No. 850, Huang-he Road, Dalian, Liaoning, 116029, China

## ARTICLE INFO

### Article History

Received 18 Jul 2020  
Accepted 26 Oct 2020

### Keywords

Deep residual network  
Underwater image restoration  
Nonlocal attention  
Channel attention  
Image de-noising  
Image color enhancement

## ABSTRACT

Underwater images or videos are common but essential information carrier for observation, fishery industry and intelligent analysis system in underwater vehicles. But underwater images are usually suffering from more complex imaging interfering impacts. This paper describes a novel residual two-fold attention networks for underwater image restoration and enhancement to eliminate the interference of color deviation and noise at the same time. In our network framework, nonlocal attention and channel attention mechanisms are respectively embedded to mine and enhance more features. Quantitative and qualitative experiment data demonstrates that our proposed approach generates more visually appealing images, and also provides higher objective evaluation index score.

© 2021 The Authors. Published by Atlantis Press B.V.

This is an open access article distributed under the CC BY-NC 4.0 license (<http://creativecommons.org/licenses/by-nc/4.0/>).

## 1. INTRODUCTION

In recent years, protecting and exploring marine resources and developing marine related technologies have gradually become an important strategy of major marine countries in the world. Underwater images or videos are common but essential information carrier for observation, fishery industry and intelligent analysis in underwater vehicles. In the practical operation, some tasks such as underwater exploration, underwater terrain browsing, fish density analysis, sea-hood health status analysis and even automatic recognition and detection rely on high-quality underwater images to ensure the smooth progress of visual analysis tasks. However, compared with the natural images, the real raw underwater images and videos usually suffer from more complex imaging interfering impacts, e.g., light scattering and absorbing by suspended particles, light darkening in deep water environment, camera blurring, and so on. No matter act alone or jointly, these interference factors result in underwater images significant visual degradation effects such as low contrast and brightness, color deviation, details blur and many kinds of noise. Figure 1 shows the visual contract effects between the degenerate underwater images affected by noise and color deviation. It can be seen that the noise and color deviation by the degradation of underwater images are the important factors affecting the quality of underwater images. Underwater image restoration can improve the visibility by eliminating the adverse effects and restoring destroyed content in the imaging process [1].



**Figure 1** | Three images in the top row are ground truth images, three images in the below row are degraded images.

In past few decades, underwater image enhancement and restoration methods range from two categories, i.e., physical priori model-based methods and data-driven techniques. Physical priori model-based methods tend to construct a model that can describe the degradation process of underwater images. And then, they usually adopt any optimizer to estimate parameters of model. For example, Trucco and Olmos-Antillon [2] designed a simplified Jaffe- MeGlamery underwater imaging model, the parameters used in the filter can be estimated by a contrast quality decision function. Fan *et al.* [3] proposed point spread function (PSF) and used

\*Corresponding author. Email: [fubo@lnnu.edu.cn](mailto:fubo@lnnu.edu.cn)

it in underwater image restoration. Many scholars have applied the dark channel model which proposed by He *et al.* [4] to underwater image restoration. Some statistics properties-based methods do not build the physical represent model, but construct subjective statistics criteria in some aspects (e.g., histogram), so as to achieve satisfactory visual effect. Ramesh [5] used histogram equalization, noise de-noising filter and repeated interpolation method to improve underwater images quality and enhance detail information. Although the physical priori model-based methods improve the quality of underwater images, these methods usually are self-learning rather data-learning. Less learning data often leads to inaccurate estimation of model's parameters, and disable to obtain satisfactory restoration effect.

With rapid development of deep learning technology, more and more deep learning tools, e.g., convolutional neural network (CNN), recurrent neural network (RNN) and generative adversative network (GAN) are respectively applied in the computer vision field and achieved significantly improved performances [6–9].

As a kind of typical methods of data-driven techniques, underwater image enhancement methods based on deep learning can fully learning parameters of model or target function through huge amount of data. Some researchers have attempted to utilize the deep learning techniques to the underwater image enhancement. For example, Li *et al.* [10] used GAN to estimate clear image through a large number of training image data sets and some underwater depth data as an end-to-end input, so as to restore image features and enhance visual effect. Zhu *et al.* [11] introduced a notion of image translation to generate two training sets, and peer sub-networks can learn self-style and transfer by adversarial loss each other. It got obvious visual color restoration effect but still owned some deficiencies in geometry restoration effect. Islam *et al.* [12] introduced many factors, i.e., image content, color and texture details to design target function of network. However, abovementioned methods usually only consider one degradation factor, we are unable to observe a deep learning method for underwater image enhancement which take into account many kinds of underwater degradation factors. Meanwhile, the phenomenon of gradient disappearing and exploding often appears in networks for underwater image enhancement, which limits the training performance.

In this paper, focus on underwater image degradation problem with multiple factors, i.e., color deviation and noise, we propose a novel residual two-fold attention network named RTFAN. In RTFAN, nonlocal attention and channel attention (CA) mechanisms are introduced into a whole residual network framework to suppress noise and enhance details respectively. Quantitative and qualitative experiment data demonstrates that our proposed approach generates more visually appealing image, and also provides higher objective evaluation index score.

## 2. RELATED WORK

Nowadays, a number of successful deep learning approaches toward different underwater image degradation have been proposed, e.g., de-noising, color restoring and de-blurring problems, etc. [13]. Different architectures of deep neural networks often show different characteristics of data processing. From an architecture perspective, the methods mentioned above can be basically divided into

(1) encoder to decoder, (2) generative adversative, (3) multi-branch and (4) modular architectures.

In general, methods based on encoder to decoder architecture are adopted in restoring details of images, so as to often be applied in de-noising and de-blurring tasks. For example, Sun *et al.* [14] used pixel to pixel (P2P) network which an encoder to decoder model to enhance underwater images. Compare to encoder to decoder model, GAN model owns an additional adversative discriminator, so it can consider more constraint for image color style. Li *et al.* [10] used GAN to restore image features and enhance visual effect. Uplavikar *et al.* [15] used domain adversarial learning to learn agnostic model where it can enhance multi-types underwater images. Fabbri *et al.* [16] used gradient penalty as the soft constraint on the underwater output which named underwater generative adversarial network (UGAN).

Although these two categories of methods have achieved good results, they are not good at dealing with underwater degraded images with multiple degradation factors. There are two reasons leading to the shortcomings of the above algorithms. Firstly, underwater image features extracting becomes more difficulty with multiple degradation factors interfering. It leads to that the feature extracted by the convolution layer is not accurate enough, which is not conducive to the recovery of features. Secondly, due to the inaccuracy of feature extraction, it is often necessary to increase the network depth. But simply increasing the depth will produce gradient disappearance or explosion. Multi-branch architecture is usually adopted to overcome the above problems. For example, Wang *et al.* [17] presented a deep CNN method for enhancement of underwater images, namely, UIE-Net, which is composed of three sub-networks, and each branch can focus on extracting some specific features. Another modular architecture such as residual network and dense net can also increase the number of network layers by skip connection. Due residual architectures can partly prevent the phenomenon of gradient disappearing, thus increasing the number of layers of the network and improving the effect of feature extraction. Guo *et al.* [18] introduced a multi-scale dense block, named DenseGAN which employs the use of dense connections, residual learning and multi-scales network for underwater image enhancement.

Attention mechanism is another effective way to enhance network features. The core ideal of attention mechanism is enhancing some particular encoder features. At present, there are two types of attention mechanisms, i.e., CA and spatial attention (SA). The CA attention acts on the channel scale, weighting different channel features and SA attention is weighted on the spatial scale, weighting different spatial regions. Zhang *et al.* [19] introduced CA attention to improve the performance of the network and propose a residual CA networks named RCAN. Some scholars try to introduce manifold constraints into deep networks to mine local features correlation and suppress noise. Liu *et al.* [9] proposed a recurrent network for image reconstruction. Zhang *et al.* [20] proposed a residual nonlocal attention network for high-quality image restoration. It also has good performance in image de-noising, mosaic removal, compression, artifact reduction and super-resolution.

Inspired by the advantages of the attention mechanism, such as the enhancement of specific network features, and deeper network architecture by skip connection. We propose a novel two-fold

attention residual network for underwater image enhancement and restoration with noise and color degradation factors simultaneously interfere. In our solution, residual architecture is adopted to ensure network deep enough, and a nonlocal attention is conducted in shallow layer to suppress noise and extract color features, and CAIs conducted in subsequent layers to enhance detail features.

### 3. METHODOLOGY

The real raw underwater images usually suffer from complex imaging interfering impacts, e.g., light scattering and absorbing by suspended particles, light darkening, which leads to color deviation, noise effects of underwater images. In this paper, focus on underwater images degradation problem with multiple factors, i.e., color deviation and noise, we propose a novel residual two-fold attention network named RTFAN.

#### 3.1. Underwater Image Modelling

According to Jaffe-MeGlamery underwater imaging model, underwater imaging is mainly composed of three parts, i.e., direct component, forward scattering component and back scattering component. Figure 2 shows the basic composition of the underwater model [21]. Where,  $x$  denotes the target position and  $D$  denotes the depth of the water.  $D(x)$  represents the distance from the target to the sea level and  $d(x)$  represents the distance between the target and the acquisition equipment.

The underwater image model can be formally expressed as follows:

$$I(x) = J(x)t(x) + A(1 - t(x)) \quad (1)$$

Here,  $x$  denotes the  $x$ -th pixel in image,  $I(x)$  denotes degraded underwater image,  $J(x)$  denotes the source clear image,  $t(x)$  denotes transmittance of scene light,  $A(1 - t(x))$  is the background scattering component. And underwater image is also degraded by noise  $n$ , and  $n$  is assumed to be i.i.d Gaussian noise with zero mean, variance  $\sigma^2$ .

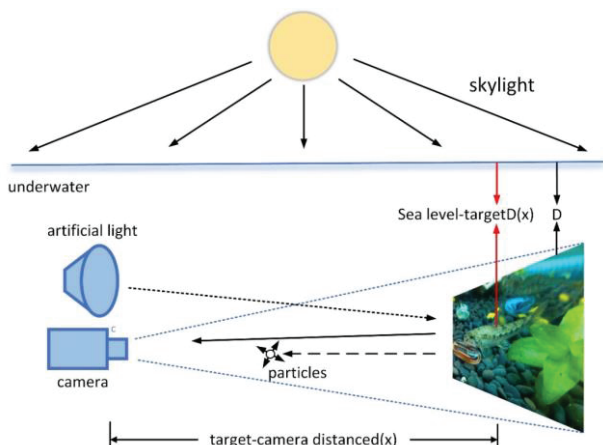


Figure 2 | Figure of underwater imaging model.

The degraded image  $F(x)$  taking into account noise and color deviation can be expressed as follows:

$$F(x) = J(x)t(x) + A(1 - t(x)) + n \quad (2)$$

So, we can express source image  $J(x)$  as follows:

$$J(x) = \frac{F(x) - A(1 - t(x)) - n}{t(x)} \quad (3)$$

#### 3.2. Objective Function Minimization

Given a source domain  $X$  (degraded image  $F(x)$ ) and an expected domain  $E$ , we adopt any neural network can get the mapping function as follows:  $G : X \{F(x)\} \rightarrow E$ . Our goal is to apply our neural network to learn the mapping function to achieve automatic underwater image restoration and enhancement. The minimization form of the objective function is follows:

$$\arg \min_{J(x)} \frac{1}{2} \|E - J(x)\|_2^2 \quad (4)$$

Here,  $E$  denotes the expected image,  $J(x)$  denotes the source clear image. The goal of our algorithm is to find a mapping that minimizes the function by training the network.

#### 3.3. Architecture of Residual Two-Fold Attention Networks (RTFAN)

Figure 3 shows a basic architecture of our RTFAN, includes three parts, i.e., shallow feature de-noising module, deep feature enhancement module and reconstruction module. In shallow feature de-noising module, there are one convolution layer and one nonlocal de-noising module named Conv\_nl, respectively. The nonlocal de-noising module can suppress noise and extract feature. The deep feature enhancement module is to further mine the image details, and it consists of four residual CA blocks (RCAB), one convolution layer and one long skip connection, detailed RCAB structure is shown in Figure 4 [19]. In RCAB, there are two Conv layers, a Relu activation layer and a CA module, which is introduced in section 3.4.2. Following the structure of RCAB, there is a deconvolutional layer to reconstruct the image.

Based on the RTFAN, we further formalize the objective function of formula into the following minimum loss function.

$$\min L(I_{RTFAN}(F(x)), J(x)) \quad (5)$$

Here,  $I_{RTFAN}$  denotes mapping function training by RTFAN model,  $F(x)$  denotes degraded underwater image,  $J(x)$  denotes clear image.

#### 3.4. Two-Fold Attention Mechanism

In RTFAN framework, nonlocal attention and CA are embedding to extract and enhance features. Nonlocal attention modular is set in shallow layers, and it focus on suppressing noise and extracting color style features. CA modular is set in deeper layers, and it focus on enhancing detail features.

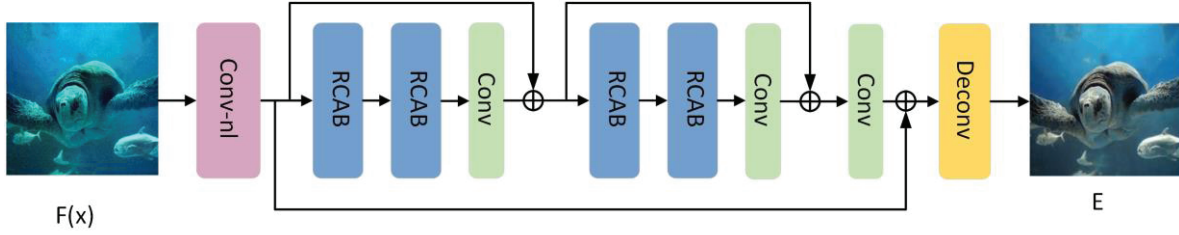


Figure 3 | Architecture of residual two-fold attention networks (RTFANs).

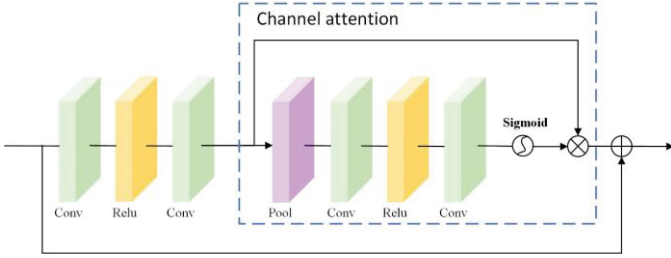


Figure 4 | Residual channel attention block (RCAB).

### 3.4.1. Nonlocal attention modular

Motivated by literature [20,22], nonlocal attention operator is modularized into a modular block, it is shown in Figure 5. In nonlocal attention block, input data size of  $B \times 3 \times H \times W$  passes a convolution layer with 64 filters to obtain a set of shallow features size of  $B \times 64 \times H \times W$ . And then compute nonlocal correlation and enhance features. In this step, we adopt  $1 \times 1$  convolutional operators and softmax function to learn three weight matrix  $W_\theta$ ,  $W_\phi$ ,  $W_g$ . Input data passes this three matrix and then pass a  $1 \times 1$  convolutional operators again to get output nonlocal operator result with size of  $B \times 64 \times H \times W$ . The nonlocal attention block can be formally expressed as follows:

$$Y = \text{soft max} \left( [I_{Conv}(X)]^T W_\theta^T W_\phi^T X \right) W_g [I_{Conv}(X)] \quad (6)$$

Here,  $I_{Conv}$  denotes a convolutional operation,  $X$  is input of network, and  $Y$  is output. And then  $Y$  is added to  $X$ . Formula express is as follows:

$$Z = I_{Conv}(X) + Y * W_Z \quad (7)$$

Here,  $W_Z$  is initialized to 0. In order to simplify the expression, we summarize Formulas (6) and (7) to one expression as follows:

$$Z = I_{Conv\_nl}(X) \quad (8)$$

### 3.4.2. CA modular

The core idea of CA is cooperative work of channels and gating mechanism (sigmoid). Figure 6 shows basic structure of CA. Firstly, the spatial information is transformed into channel feature with  $1 \times 1 \times C$  size by a global average pooling. Secondly, convolutional operation is carried out with the weight of  $W_D$ , which is activated by ReLU activation function, and the channel feature with  $1 \times 1 \times C/r$  size is down-sampled with the ratio  $r$ . Then, another convolution operation is carried out with the weight of  $W_U$  to upscale

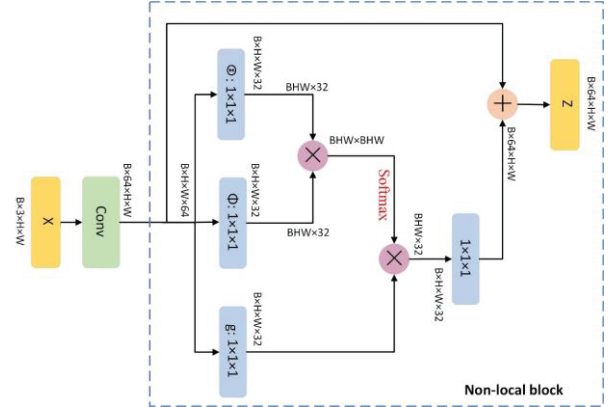


Figure 5 | Nonlocal attention.

with the ratio  $r$ ,  $1 \times 1 \times C$  channel feature is obtained. Finally, after being activated by sigmoid function (denoted as  $f$ ), the final channel feature with  $1 \times 1 \times C$  is obtained, and is further multiplied with input to obtain output.

It can be seen from Figure 6, CA mechanism can better enhance deeper features. In view of this characteristic of CA mechanism, our RTFAN apply CA modular in deep layers and the deep feature enhancement module named as  $I_{Res\_ca}$ , its express is as follows:

$$W = I_{Res\_ca}(Z) \quad (9)$$

Here,  $Z$  and  $W$  denote the input and output of the deep feature enhancement module, respectively.

In the last layer of RTFAN, we apply a deconvolution to reconstruct features, marked as  $I_{Dec}$ . Thus, the expectation of the final result estimated by RTFAN network can be expressed as follows:

$$E = I_{Dec} (I_{Res\_ca} (I_{Conv\_nl}(X))) \quad (10)$$

## 4. EXPERIMENTS

### 4.1. Experimental Setting

**Dataset** We use 5550 paired underwater images from EUVP [12] as the data set, where 5022 pairs are used for training and the rest are verified and tested. For the noisy image training set, the simulated Gaussian noise is added to the original data set.

**Baseline algorithm** We use four existing underwater image enhancement algorithms as baselines, including cycleGan [11],



funieGAN [12], RCAN [19]. We also compare with our previous work, named Color-diff Net [22]. For fair comparison, all methods' code implementations are their publicly available versions and their parameters are set following the guidelines in original articles. In addition, our algorithm and all comparison algorithms adopt the above consistent data set.

**Evaluation Metric** The peak signal-to-noise ratio (PSNR) and the structural similarity (SSIM) index are adopted to measure the objective performance of our algorithm.

## 4.2. Network Architecture and Training Details

We choose the network architecture of Zhang *et al.* [19] as our basic network framework. Our network consists of a shallow feature de-noising module, a deep feature enhancement module and a reconstruction module. Among them, the shallow feature de-noising module is composed of a convolutional layer and a nonlocal module, while the deep feature module mainly contains four residual channel attention blocks (RCABs) and skip links. In each RCAB, there are two convolutional layers with size of  $3 \times 3$  filters and a CA block module. In the CA block module, input data is squeezed into size of  $1 \times 1 \times C$  data by average pooling, then this  $1 \times 1 \times C$  descriptors are put through the convolutional layers and activated by sigmoid function. During training step, the training images are cut into the patches with size of  $48 \times 48$ . The batch size is set as 16, the learning rate is set as  $1e-4$  and the ADAM is adopted as optimizer of our model.

In our method, we only train 300 epoch for our model. We choose same training dataset, iteration times and network parameter for comparison methods. Our network framework is based on the PyTorch platform, and all experiments are implemented on a workstation with NVIDIA Titan X GPU and Core (TM) i7-7700K CPU.

## 4.3. Evaluation

### 4.3.1. Qualitative evaluations

Firstly, we qualitatively analyze the visual quality of degraded images with color deviation and noise factors restored by RCAN and RTFAN, respectively. As shown in Figure 7, most of the color information can be recovered from the RCAN-generated result when only the color deviation is considered. As the gaussian noise with noise intensity of 10 is added, simple RCAN cannot achieve a good effect of restoring color and removing noise at the same time. The RCAN network with only CA can correct the color properly, but it's de-noising ability is not strong. While, RTFAN network with

only nonlocal attention can achieve simple de-noising tasks, but it cannot recover the real color information and detailed features. However, our RTFAN model can remove noise, perform color correction and get closer to the real image in visual effect. Therefore, we further prove the advantages in our model.

Next, we use four underwater image enhancement models based on deep learning to make a qualitative comparison with our RTFAN model: (1) cycleGan [11], (2) funieGAN [12], (3)Color-diff Net [22] (our previously completed work), (4) RCAN [19]. We use the same noise data set for training on these models, and the test results are amplified, as shown in Figures 8-10.

Figures 8-10 shows the contrast of enhancement algorithms for underwater images with 5, 10 and 15 intensity of Gaussian noise, respectively. It can be seen that the cycleGan model and the funieGAN model cannot correct the color bias in the image, while Color-diff Net can correct the color, but the de-noising effect is not well. Relative to results above the performance of RCAN is better, but

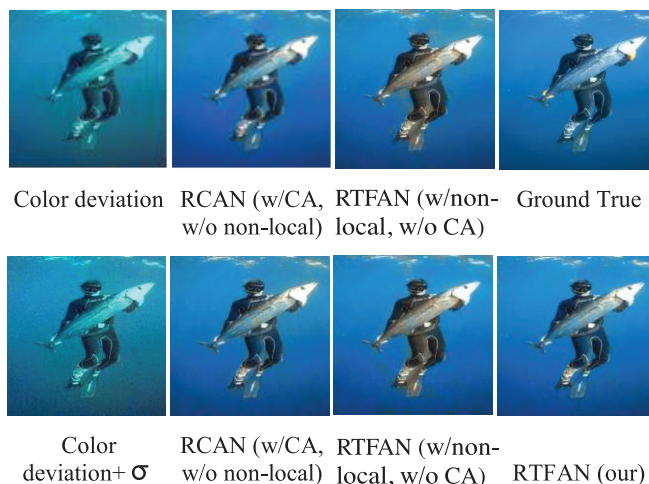


Figure 7 | Visual contrast between RCAN model and our residual two-fold attention network (RTFAN) model with multiple factors.

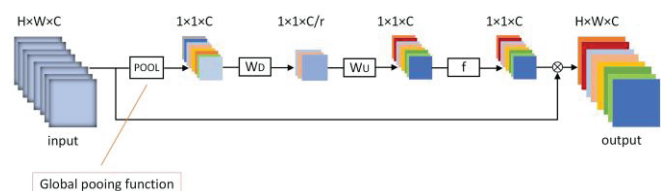


Figure 6 | Channel attention.

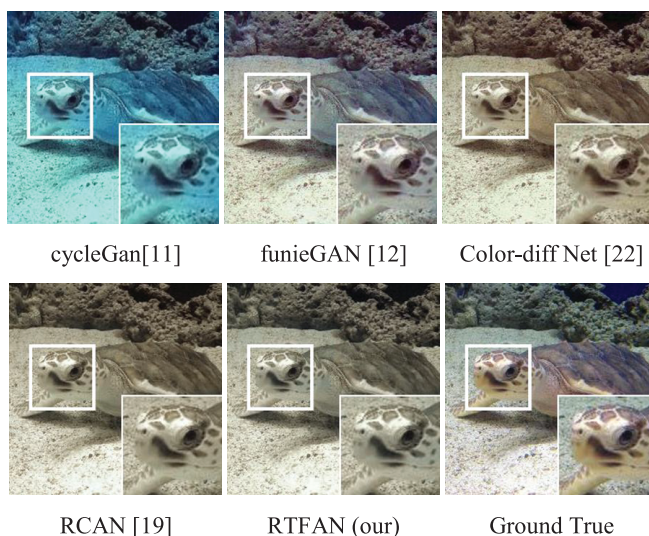


Figure 8 | Comparison of enlarged view with noise intensity of 5.

slightly inferior to our method. As a result, our approach is generally effective, which can not only remove noise and perform color correction, but also get closer to the real image in visual effect.

### 4.3.2. Quantitative evaluation

After qualitative evaluation, we use PSNR and SSIM to quantitatively compare the evaluation values of our method and other methods. As an objective standard for image evaluation, the PSNR mainly calculates the mean square error (MSE) between the real image and the processed image. The formula (11) is expressed as follows:

$$PSNR(x, y) = 10 * \log_{10} \left[ \frac{255^2}{MSE(x, y)} \right] \quad (11)$$

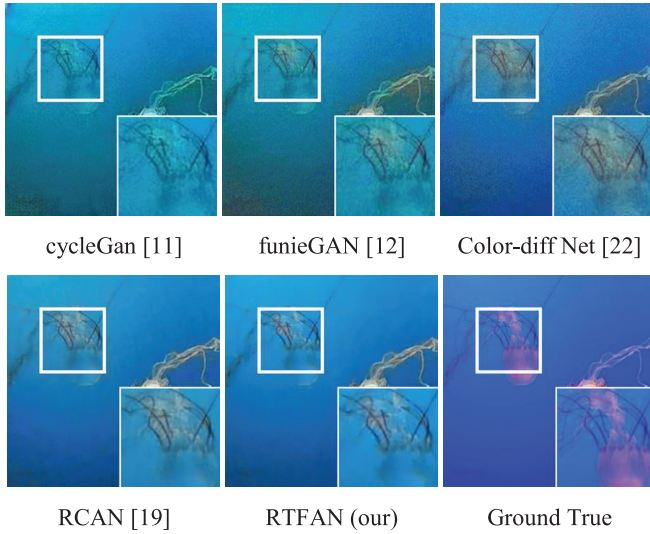


Figure 9 | Comparison of enlarged view with noise intensity of 10.

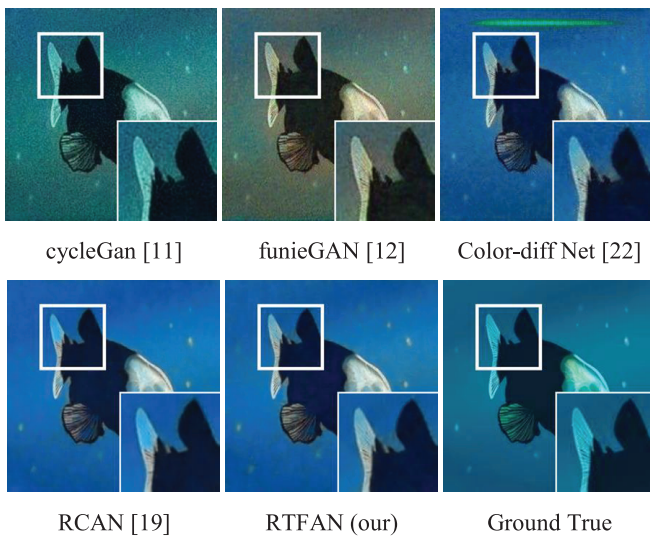


Figure 10 | Comparison of enlarged view with noise intensity of 15.

The SSIM is an index to measure the similarity of two images, which mainly considers the brightness, contrast and structure. The formula (12) is as follows:

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + c_1) (2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1) (\sigma_x^2 + \sigma_y^2 + c_2)} \quad (12)$$

where,  $\mu_x$  and  $\mu_y$  are the mean,  $\sigma_x^2$  and  $\sigma_y^2$  are the variance,  $\sigma_{xy}$  denotes the covariance,  $c_1$  and  $c_2$  are the constant that's used to maintain stability. Table 1 compares the PSNR score and SSIM score of our method with some of the underwater image enhancement methods.

It can be seen that in Table 1, our algorithm has obtained obvious advantages compared with other algorithms on PSNR, but the SSIM is a little deficient, the reason is that our RTFAN both considering de-noising and color deviation, so de-noising operation inevitably smooths the target image, which results in changing tiny detail construct. However, due to the underwater image is often polluted by noise, it is necessary to consider de-noising item. Figure 11 shows our algorithms more superior with increasing intensity of 10  $\sigma$  noise.

### 4.4. Ablation Experiments

In order to further prove the effectiveness of our proposed algorithm and highlight the superiority of the two-fold attention mechanism, we conduct three different tests respectively. For example, (1) in the case of without noise, RCAN network with only CA mechanism; (2) RCAN network with only CA mechanism after adding 10 Gaussian noises; (3) in the case of without noise, RTFAN network with only nonlocal mechanism; (4) RTFAN network with only nonlocal mechanism after adding 10 Gaussian noises; (5) RTFAN network with the two-fold attention mechanism after adding 10 Gaussian noises. The visual effect is shown in Figure 7. The actual results were quantified by PSNR and SSIM, as shown in Table 2.

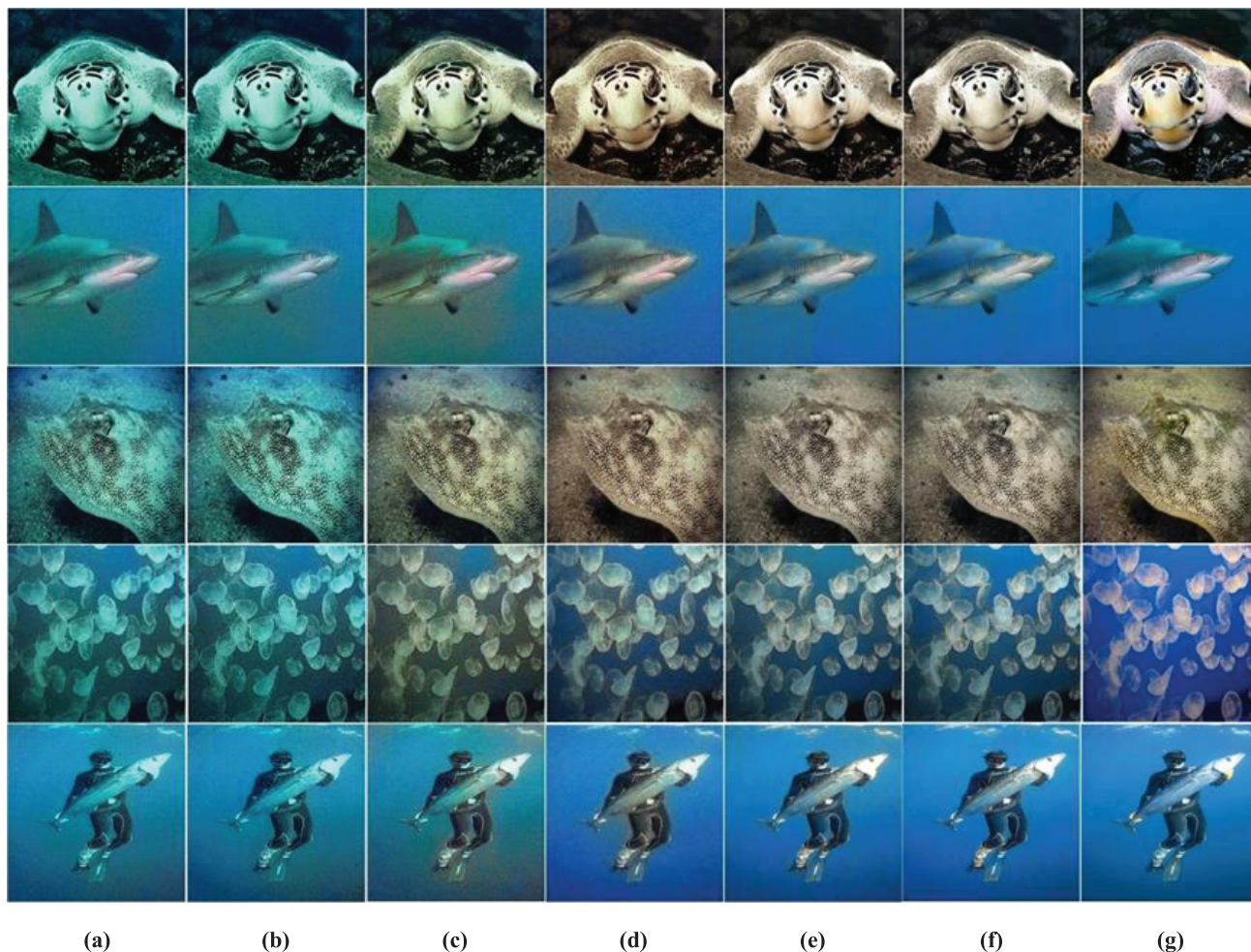
Whether from the visual effect or quantitative results, it can be concluded that our algorithm has certain advantages in the restoration of noisy underwater images compared with other algorithms.

Table 1 | The PSNR and SSIM scores of our method and some underwater image enhancement methods (best results are highlighted, our results are underlined).

Algorithm	Evaluation Index	EUVP + Noise( $\sigma$ )		
		5	10	15
cycleGan	PSNR	16.35	16.27	16.26
	SSIM	0.7773	0.7295	0.6858
funieGAN	PSNR	20.52	20.10	20.21
	SSIM	0.8797	0.8560	0.8381
Color-diff Net	PSNR	22.06	21.80	21.48
	SSIM	<b>0.9017</b>	<b>0.8842</b>	<b>0.8574</b>
RCAN	PSNR	22.36	22.15	21.96
	SSIM	0.8936	0.8729	0.8485
RTFAN	PSNR	<u>22.36</u>	<u>22.19</u>	<u>21.97</u>
	SSIM	<u>0.8936</u>	<u>0.8708</u>	<u>0.8498</u>

RTFAN, residual two-fold attention network; PSNR, peak signal-to-noise ratio; SSIM, structural similarity index.





**Figure 11** | The enhanced result images: (a) input, (b) cycleGan, (c) funieGAN, (d) Color-diff Net, (e) RCAN, (f) residual two-fold attention network (RTFAN) and (g) ground truth.

**Table 2** | Compared within and without noise (best results are highlighted, our results are underlined).

Dataset	Noise( $\sigma$ )	Methods	PSNR	SSIM
EUVP	0	RCAN(w/CA)	22.430	0.90224
EUVP	0	RTFAN(w/oCA)	22.139	0.89840
EUVP	10	RCAN(w/CA)	22.151	<b>0.87290</b>
EUVP	10	RTFAN(w/oCA)	21.798	0.86094
EUVP	10	RTFAN	<u>22.188</u>	<u>0.87075</u>

PSNR, peak signal-to-noise ratio; SSIM, structural similarity index.

## 5. CONCLUSION

In this paper, we show a novel residual two-fold attention network for underwater image restoration and enhancement to eliminate the interference of color deviation and noise at the same time. In our network framework, nonlocal attention and CA mechanisms are respectively embedded to mine and enhance features. Quantitative and qualitative experiment data demonstrates that our proposed approach generates more visually appealing images, and also provides higher objective evaluation index score.

## AUTHORS' CONTRIBUTIONS

Bo Fu designed model and wrote paper. Liyan Wang implemented the whole algorithm. Ruizi Wang and Fangfei Liu responsible for statistical algorithm related experimental data. And Xin Liu checked the innovation and data of proposed algorithm.

## ACKNOWLEDGMENTS

This work is supported by the National Natural Science Foundation of China (NSFC) Grant No. 61702246, 61772250, China Postdoctoral Science Foundation No. 2019M651123 and Science and Technology Innovation Found (Youth Science and Technology Star) of Dalian, China No. 2018RQ65.

## REFERENCES

[1] R.Q. Wang, Y.F. Wang, J. Zhang, Review on underwater image restoration and enhancement algorithms, in Proceedings of the 7th International Conference on Internet Multimedia Computing and Service, Zhangjiajie, China, 2015, pp. 1–6.

- [2] E. Trucco, A.T. Olmos-Antillon, Self-tuning underwater image restoration, *IEEE J. Ocean. Eng.* 31 (2006), 511–519.
- [3] F. Fan, K.C. Yang, M. Xia, Underwater image restoration by means of blind deconvolution approach, *Front. Optoelectron. China.* 3 (2010), 169–178.
- [4] K.M. He, J. Sun, X.O. Tang, Single image haze removal using dark channel prior, *IEEE Trans. Pattern Anal. Mach. Intell.* 33 (2011), 2341–2353.
- [5] R. Ramesh, Haze-Free Underwater Image Enhancement, *Inter. J. of Innovative Research and Development* 2 (2013), 533–543.
- [6] C. Dong, C.C. Loy, K.M. He, Learning a deep convolutional network for image super-resolution, in: D. Fleet, T. Pajdla, B. Schiele, T. Tuytelaars (Eds.), *Computer Vision – ECCV*, Springer International Publishing, Cham, Switzerland, 2014.
- [7] J. Kim, J.K. Lee, K.M. Lee, Accurate image super-resolution using very deep convolutional networks, in 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 2016, pp. 1646–1654.
- [8] K. Zhang, W.M. Zuo, Y.J. Chen, Beyond a Gaussian Denoiser: residual learning of deep CNN for image Denoising, *IEEE Trans. Image Process.* 26 (2016), 3142–3155.
- [9] D. Liu, B.H. Wen, Y.C. Fan, Non-local recurrent network for image restoration, in *Proceedings of the 32nd International Conference on Neural Information Processing Systems December (NIPS)*, Red Hook, NY, USA, 2018, pp. 1680–1689.
- [10] J. Li, K.A. Skinner, R.M. Eustice, WaterGAN: unsupervised generative network to enable real-time color correction of monocular underwater images, *IEEE Robot. Autom. Lett.* 3 (2018), 387–394.
- [11] J.Y. Zhu, T. Park, P. Isola, Unpaired image-to-image translation using cycle-consistent adversarial networks, in *IEEE International Conference on Computer Vision (ICCV)*, Venice, Italy, 2017, pp. 2223–2232.
- [12] M. J. Islam, Y. Xia, J. Sattar, Fast underwater image enhancement for improved visual perception, *IEEE Robot. Autom. Lett.* 5 (2020), 3227–3234.
- [13] S. Anwar, C. Li, Diving deeper into underwater image enhancement: a survey, *Signal Process. Image Commun.* 89 (2020), 115978.
- [14] X. Sun, L.P. Liu, Q. Li, Deep pixel to pixel network for underwater image enhancement and restoration, *IET Image Process.* 13 (2019), 469–474.
- [15] P. Uplavikar, Z. Wu, Z. Wang, All-in-one underwater image enhancement using domain-adversarial learning, arXiv preprint arXiv: 1905.13342, 2019. [https://openaccess.thecvf.com/content\\_CVPRW\\_2019/html/UG2\\_Prize\\_Challenge/Uplavikar\\_All-in-One\\_Underwater\\_Image\\_Enhancement\\_Using\\_Domain-Adversarial\\_Learning\\_CVPRW\\_2019\\_paper.html](https://openaccess.thecvf.com/content_CVPRW_2019/html/UG2_Prize_Challenge/Uplavikar_All-in-One_Underwater_Image_Enhancement_Using_Domain-Adversarial_Learning_CVPRW_2019_paper.html)
- [16] C. Fabbri, M.J. Islam, J. Sattar, Enhancing underwater imagery using generative adversarial networks, in 2018 IEEE International Conference on Robotics and Automation (ICRA), Brisbane, Australia, 2018.
- [17] Y. Wang, J. Zhang, Y. Cao, A deep CNN method for underwater image enhancement, in 2017 IEEE International Conference on Image Processing (ICIP), IEEE, Beijing, China, 2017.
- [18] Y.C. Guo, H.Y. Li, P.X. Zhuang, Underwater image enhancement using a multiscale dense generative adversarial network, *IEEE J. Ocean. Eng.* 45 (2020), 862–870.
- [19] Y.L. Zhang, K.P. Li, K. Li, Image super-resolution using very deep residual channel attention networks, in: V. Ferrari, M. Hebert, C. Sminchisescu, Y. Weiss (Eds.), *European Conference on Computer Vision (ECCV 2018)*, Springer, Cham, Switzerland, 2018.
- [20] Y.L. Zhang, K.P. Li, K. Li, Residual non-local attention networks for image restoration, 7th International Conference on Learning Representations (ICLR 2019), New Orleans, Louisiana, US, 2019.
- [21] X.Y. Ding, Underwater Image Enhancement Algorithm Based on Convolutional Neural Networks, Dalian Maritime University, Dalian, China, 2018.
- [22] B. Fu, R.Z. Wang, L.Y. Wang, Enhancement method of underwater color image based on deep convolution neural network, *J. Jilin Univ. Sci. Ed.* (2020).