IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

SPECIAL SECTION ON ADVANCED OPTICAL IMAGING FOR EXTREME ENVIRONMENTS

Received July 4, 2019, accepted July 12, 2019, date of publication July 15, 2019, date of current version July 31, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2928976

Underwater Image Enhancement With a Deep Residual Framework

PENG LIU[1,2], (Member, IEEE), GUOYU WANG[2], HAO QI[3], CHUFENG ZHANG[2], HAIYONG ZHENG[2], (Member, IEEE), AND ZHIBIN YU[2], (Member, IEEE)

[1]Computing Center, Ocean University of China, Qingdao 266100, China
[2]College of Information Science and Engineering, Ocean University of China, Qingdao 266100, China
[3]College of Management, Ocean University of China, Qingdao 266100, China

Corresponding author: Zhibin Yu (yuzhibin@ouc.edu.cn)

This work was supported in part by the MOE (Ministry of Education in China) Project of Humanities and Social Sciences under Grant 18YJCZH103, in part by the National Natural Science Foundation of China under Grant 61571407, in part by the Fundamental Research Funds for the Central Universities of China under Grant 201615003, and in part by the National Natural Science Foundation of China under Grant 61701463.

ABSTRACT Owing to refraction, absorption, and scattering of light by suspended particles in water, raw underwater images have low contrast, blurred details, and color distortion. These characteristics can significantly interfere with visual tasks, such as segmentation and tracking. This paper proposes an underwater image enhancement solution through a deep residual framework. First, the cycle-consistent adversarial networks (CycleGAN) is employed to generate synthetic underwater images as training data for convolution neural network models. Second, the very-deep super-resolution reconstruction model (VDSR) is introduced to underwater resolution applications; with it, the Underwater Resnet model is proposed, which is a residual learning model for underwater image enhancement tasks. Furthermore, the loss function and training mode are improved. A multi-term loss function is formed with mean squared error loss and a proposed edge difference loss. An asynchronous training mode is also proposed to improve the performance of the multi-term loss function. Finally, the impact of batch normalization is discussed. According to the underwater image enhancement experiments and a comparative analysis, the color correction and detail enhancement performance of the proposed methods are superior to that of previous deep learning models and traditional methods.

INDEX TERMS Asynchronous training, edge difference loss, residual learning, underwater image enhancement.

I. INTRODUCTION

Recently, vision-guided Autonomous Underwater Vehicles (AUVs) and Remote Operated Vehicles (ROVs) have integrally impacted the exploration of marine resources [41]–[44]. Clear underwater images are an important prerequisite for such vision-guided explorations. However, the raw underwater images currently obtained have low contrast, blurred details, and color distortion. This results from light refraction, absorption, and scattering of suspended particles.

The absorption of light by water is selective: the absorption rate of red light is higher, whereas the transmission rate of blue and green light is better. Accordingly, raw underwater images are mostly blue or green, unlike that of an in-air image. The scattering of light in water can be divided into two types: forward scattering and backward scattering [1]. Forward scattering light comes from the object and thus usually contributes to in the blurred texture features of underwater objects. Backward scattering constitutes the light that is reflected back before reaching the target object; this atomizes the underwater image and causes noise. These challenges bring difficulties on such tasks as segmentation, tracking, and vision-based navigation system. Therefore, underwater image restoration plays an extremely important role in underwater vision tasks, such as underwater monitoring, deep-sea exploration, underwater robots and so on. Underwater image enhancement can promote the reliability of underwater vision tasks by increasing the underwater image contrast and reducing the degradation caused by scattering and attenuation. To address these challenges, this paper proposes an underwater image enhancement solution through a deep residual

The associate editor coordinating the review of this manuscript and approving it for publication was Huimin Lu.

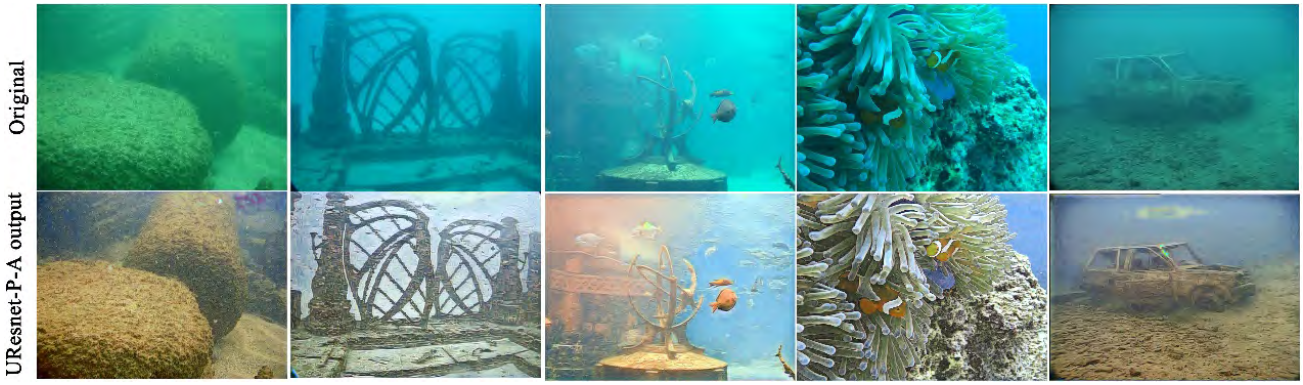


FIGURE 1. Comparison of underwater images and the images enhanced by the proposed UResnet-P-A.

framework. Unlike other deep learning based underwater enhancement approaches which focus on the cooperation between generative adversarial networks and weakly supervised learning [27]–[30], we provide a residual learning based framework which aims to build a deeper network and improve the performance for underwater image enhancement.

Firstly, the cycle-consistent adversarial networks (CycleGAN) [2] is employed to generate approximately 4000 pairs of synthetic underwater images and label images as a training set for underwater image enhancement. This addressed the training data insufficiency of using powerful supervised learning method to enhance underwater image.

Secondly, the Underwater Resnet (UResnet) is proposed for underwater image enhancement, according to the idea of residual learning. The VDSR model [3], which is used for super-resolution reconstruction tasks, was introduced into the field of underwater image enhancement.

Thirdly, the loss function and training mode were improved. Edge difference loss (EDL) is proposed to improve the detail enhancement ability of deep learning models. An asynchronous training mode is also proposed that propagates mean square error (MSE) loss and EDL asynchronously during training to improve the performance of the multi-term loss function. The structure of the proposed UResnet model with the EDL penalty term is shown in Fig. 5. The underwater image enhancement effects of the proposed UResnet with EDL using the asynchronous training mode are shown in Fig. 1.

II. RELATED WORKS

Research of underwater image restoration or enhancement is a frontier topic. Although some studies [16]–[22] on this topic have been conducted, the achievements are not satisfied. Because this issue is complicated, it is influenced by many factors, such as illumination, water quality. Meanwhile underwater image restoration or enhancement, which is usually used as a pre-processing method, is necessary for underwater robot navigation, underwater target detection, terrain detection, exploration and archaeology. Traditional underwater image restoration or enhancement methods either involve a non-physical model method or a physical model-based method. But traditional approaches always need to set many parameters manually (e.g. the scattering ratio of red, green and blue channel). Recently, the development of deep learning methods brings another solution for this research topic.

Underwater image enhancement based on a non-physical model refers to the method of improving visual effects by adjusting image pixel values rather than establishing a mathematical and physical model to simulate the image optical imaging characteristics. Histogram equalization and its improved algorithm [4]–[6] are commonly used non-physical contrast enhancement methods. By stretching the pixel value nonlinearly, the original random distribution histogram is transformed into a uniform distribution. Thus, the contrast of the image can be improved. Underwater images usually have significant color deviation; when light travels in water, different colors of light are attenuated at different rates. White balance [7], [8] is a classical image color correction algorithm. Images taken in different illumination environments will produce different extents of color deviation. Thus, the white balance algorithm corrects the color by adjusting the ratio of RGB three-channel pixels. There are also image enhancement methods based on human vision brightness and color perception, such as the classical Retinex image enhancement algorithm and its improved algorithms [9]–[12], which have great influence in the field of image enhancement.

The physical model-based methods model the degradation process of the image: by estimating the parameters of the model, the degradation process can be inverted to obtain the normal image. Because the scattering of particles in water is similar to that of aerosols in the atmosphere, many image dehazing algorithms have been applied to underwater image restoration, such as the dark channel algorithm [13], the non-local algorithm [14], the Fattal image dehazing algorithm [15], and other image restoration algorithms based on the atmospheric scattering model. However, the underwater image has more forward scattering and selective absorption of visible light than an in-air haze image. The underwater image also often has color deviation, and is more blurred than

an in-air image. Thus, applying image dehazing algorithms to underwater image restoration is usually insufficient.

There are also many studies on image restoration algorithms, particularly for underwater image optical characteristics [16]–[21]. For example, the red channel underwater image restoration method [22] is based on the physical fact that light decays exponentially in water and that different wavelengths of light have different attenuation rates. Considering that red light has the fastest attenuation, a red channel method is proposed, as a variant of the dark channel dehazing algorithm.

The non-physical model methods are sufficient for in-air image processing; however, because they ignore the optical properties specific to underwater images, they can easily lead to color deviation when applied to underwater images, wherein the enhanced image appears over-saturated and under-saturated in disparate regions. The physical model-based methods take into account the optical properties of underwater images, but often rely on environmental assumptions and prior knowledge. Mathematical models have great limitations and the estimation methods of model parameters are complex. They are difficult to generalize for different scenarios, and their robustness is poor.

Recently, deep learning methods have been applied to various fields [23], [24] such as computer vision and natural language processing. These methods can likewise be applied to underwater image restoration and enhancement [25], [26]. J. Li, *et al*. [27] propose WaterGAN, a generative adversarial network (GAN) for generating realistic underwater images from in-air image and depth pairings in an unsupervised pipeline used for color correction of monocular underwater images. C. Li, *et al*. [28] propose a weakly supervised color transfer method to correct underwater color distortion based on CycleGAN [2] and a multi-term loss function. C. Fabbri, *et al*. [29] use CycleGAN to generate paired training data for underwater image restoration tasks, and proposed underwater GAN (UGAN) based on pix2pix [30] to improve underwater image quality.

Currently, most of the deep learning methods for underwater image restoration and enhancement are weakly supervised learning methods based on generative adversarial networks. Most of these methods solely focus on color correction. Few methods can enhance overall image details effectively. Proposing a more comprehensive approach, this paper employs powerful supervised learning methods for underwater image enhancement tasks. According to the characteristics of an underwater image, the loss function and training mode are improved to promote the enhancement effect.

III. METHODOLOGY

A. PREPARING THE TRAINING DATA SET BY CYCLEGAN

Powerful supervised learning methods have achieved remarkable results in many fields of computer vision, but they have rarely been used in underwater image enhancement. Because the blurred underwater images have no corresponding clear images as the ground truth, powerful supervised learning methods such as CNN (Convolutional Neural Networks), which need a lot of ''paired'' training data, cannot be applied in this field.

To address this lack of training data, CycleGAN [2] is employed to generate training data. CycleGAN is a variant of a GAN [31], which can learn mapping from one data distribution to another data distribution without paired training data.

CycleGAN consists of discriminators D_X and D_Y and generators F and G. The discriminator D_X learns the features of the in-air images to judge whether the output result is the in-air image; the discriminator D_Y learns the features of the underwater images to judge whether the output result is the underwater image. The generator G learns the mapping from the in-air images to the underwater images; the generator F learns the mapping from the underwater images to the in-air images to complete the mutual conversion between the in-air images and the underwater images.

Approximately 4500 in-air images were collected from public data sets such as urban100 and bsd100 as domain X, and approximately 5000 underwater images were collected from Flickr as domain Y. CycleGAN is employed to learn the mapping from X to Y. Thus, CycleGAN is used to transform clear in-air images to underwater-style images, which generates paired training data for a powerful supervised learning model. CycleGAN can also learn the process from Y to X, that is, the underwater images can be restored to in-air image quality. In Section 4.4, the underwater image enhancement effects are experimentally compared between the proposed methods, CycleGAN, and other methods.

The image pairs consisting of the in-air images and the generated ''synthetic underwater images'' are used as the training set for the powerful supervised learning models. Approximately 4000 pairs were generated. Data set examples are shown in Fig. 2.

B. UNDERWATER IMAGE ENHANCEMENT WITH THE VDSR

The challenges of underwater image enhancement are similar to that of super-resolution reconstruction. Compared to an in-air image, a raw underwater image's color is distorted. Moreover, underwater images are blurred, and details, such as the edges of underwater objects, are muted or lost. Thus, underwater image enhancement requires color deviation correction and detail restoration. Similarly, super-resolution reconstruction aims to restore images' details.

After supplying training data for a powerful supervised learning model with CycleGAN, the very-deep super-resolution (VDSR) model [3] was introduced to the underwater image enhancement task.

The VDSR model has 20 convolution layers. Each convolution layer uses 3*3 size filters, with a stride of 1 and zero-padding with 1 pixel. Such parameter settings ensure that the resolution of the input image is identical to that of the output image. Except for the first and the last layers, each

FIGURE 2. Samples of in-air and synthetic underwater images pairs.

convolution layer has 64 channels. The first layer receives three-channel image data as input, generates 64-channel feature maps, and transmits them to the main body of the network. The last layer is the reconstruction layer. It receives 64-channel feature maps and outputs three-channel residual images. The residual images are added to the input images to generate the final restored images. When the VDSR model is used for super-resolution reconstruction, the input image is a high-resolution image generated by bicubic interpolation of a low-resolution image, such that the input image and the output image are the same size.

Therefore, when the VDSR model is applied to underwater image restoration, the size of input and output images does not need to be adjusted, and neither does the network structure. Only appropriate training data are needed for the network to learn the difference between underwater and in-air images.

C. EDGE DIFFERENCE LOSS

Most image-to-image translation models use per pixel difference loss functions such as the MSE or L1 loss function. The original VDSR model used MSE Loss, wherein it tries to make the VDSR model learn the pixel level difference between the two images. Using MSE Loss, the model can achieve a higher peak signal-to-noise ratio (PSNR) [32] score, but the generated images do not provide good visual effects; MSE Loss averages the differences at the pixel level and fails to take higher-level information, such as an overall structure, into account. Thus, the MSE Loss function tends to average the solution and make the image details too smooth, which is not conducive to the enhancement of high-frequency information.

Due to the significant detail loss of underwater images, especially with regard to edge information, a penalty term is proposed called edge difference loss (EDL). By penalizing the models with EDL, the details of generated images are promoted to a higher level.

To calculate the EDL, a Laplacian operator is adopted, serving as a sensitive edge detection operator. The Laplacian template

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & -8 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

(denoted as *lap*) is used as the convolution kernel to perform the convolution operation on the output results of the last layer of the model. Simultaneously, the convolution operation is performed on the label image using the Laplacian template. The MSE Loss of these two feature maps is then calculated as the EDL.

The output images from the deep learning models are denoted as I^g; these are the enhanced images from the underwater style images; the corresponding clear images (label images) are denoted as I^c. MSE Loss and EDL are thus calculated as follows:

$$MSE\ Loss = \mathbb{E}\left[\left\|I^c - I^g\right\|_2\right], \tag{1}$$

$$EDL = \mathbb{E}\left[\left\|I^c \otimes lap - I^g \otimes lap\right\|_2\right]. \tag{2}$$

Here, MSE Loss is the second-order norm of the difference between the enhanced image and the real image. EDL is the second-order norm of the feature map difference between the output image and the label image after the Laplacian convolution operation. Then, the loss function is composed of two parts: the MSE loss and the EDL:

$$Loss = MSE\ Loss + k * EDL. \tag{3}$$

The coefficient k is a super-parameter. It adjusts the proportion of the two parts of the loss. The value of parameter k in equation (3) is obtained by greedy search, which requires a lot of experiments.

The k value varies for different data sets in order to make the size of the two parts of the loss function at the same level. This is referred to as the synchronous training mode. The VDSR model with the EDL penalty term in synchronous

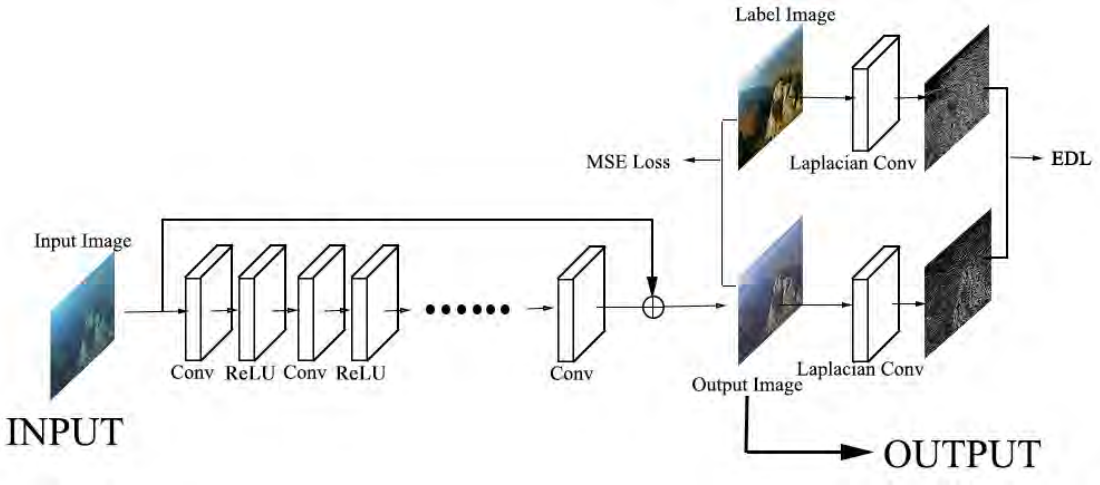


FIGURE 3. VDSR model with the proposed EDL penalty term.

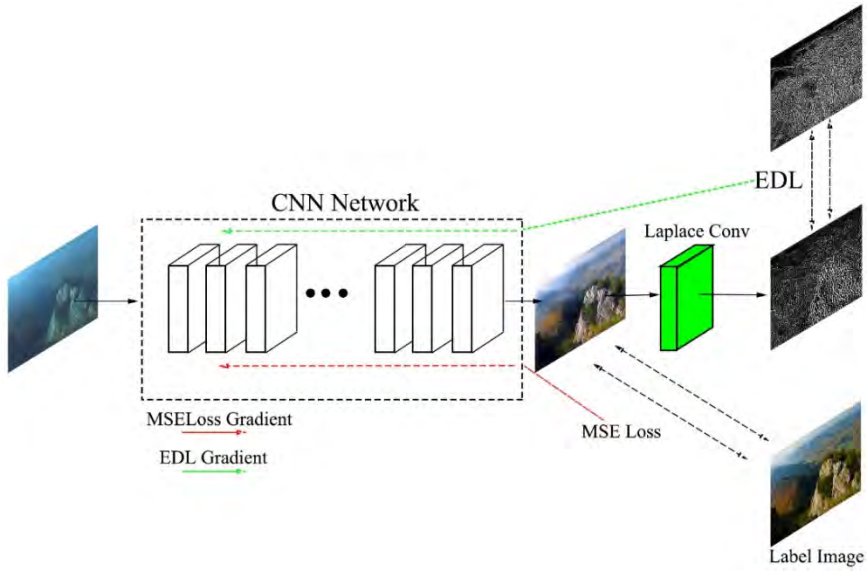


FIGURE 4. Proposed asynchronous training mode.

training mode is denoted as VDSR-P-S. This model structure is shown in Fig. 3.

D. ASYNCHRONOUS TRAINING MODE

In practice, it is difficult to find the best k value in equation (3) to optimize the output. Because the Laplacian operator is sensitive to edge information, it is susceptible to noise. If the value of k is not appropriate, it will lead to noise amplification and thus reduce the quality of the output image. Therefore, the asynchronous training method is proposed for the VDSR-P model. This training mode is also applicable to other deep learning models which use a multi-term loss function. The specific steps of the asynchronous training method are as follows:

The network has to train each batch twice. In the first round, EDL is used to calculate the gradients and execute back propagation to update the weights of the network; in the second round, the gradient is calculated using MSE loss, and propagated back to update the weights of the network. Thus, each batch is trained twice, and the weights of the network are updated twice for each batch. The flow chart of the asynchronous training mode is shown in Fig. 4.

The asynchronous training mode is adopted and EDL used in the first training round. This method can take advantage of EDL's ability to provide edge information and thus assists the network in restoring edge information and details. However, EDL's influence on the network is restricted by the second training, which limits the network to focus on the difference of the pixel level between the output and label images, thus suppressing the amplification effect of the Laplacian operator on noise.

Additionally, if the two parts of loss function are trained with different weights by constructing a multi-term loss function, as the traditional multi-loss training model does,

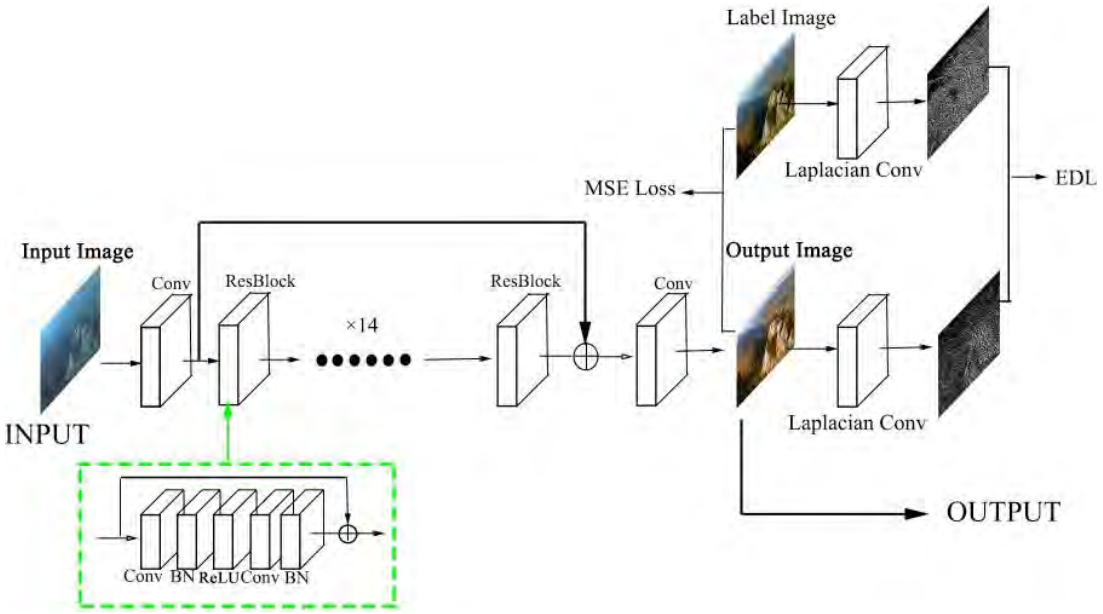


FIGURE 5. The proposed UResnet model with the EDL penalty term.

the appropriate loss weights allocation needs to be identified via a large number of experiments. This presents a challenge in determining the optimal weights. However, the weights' determination is often unalterable and can sacrifice the robustness of the model.

Alternatively, asynchronous training is equivalent to learning the k value in equation (3) by the network, wherein the proportion of the two parts in the loss function is automatically adjusted to achieve the optimal solution. VDSR-P in asynchronous training mode is denoted as VDSR-P-A.

E. UNDERWATER RESNET (URESNET)

The VDSR model sufficiently achieves underwater images enhancement; however, the VDSR model is a relatively shallow model, with 20 convolution layers and only one skip connection. It is well known that the simplest way to enhance the performance of the CNN model is stacking more layers. Generally, the deeper CNNs have more parameters and better potential to handle complex tasks. However, the deep CNNs are difficult to train. ResBlocks [35] and skip-connections [45] can ease the training of deep CNNs. The VDSR model has one skip connection. However it does not use ResBlocks, which limits the depth of VDSR model. To further improve underwater image enhancement, a deeper model is proposed, denoted Underwater Resnet (UResnet).

The proposed UResnet is a residual learning [35] model. It is composed of ResBlocks, which add the input of one convolution layer to the output of the next convolution layer. Utilization of ResBlocks ensure that the information from the previous layer can be fully transmitted to the ensuing layers. Stacking ResBlocks allows deeper networks to be trained. UResnet is inspired by the super-resolution reconstruction models EDSR [33] and SRResnet [34].

The proposed UResnet model is comprised of three main sections: a head, body, and tail. Inspired by VDSR and EDSR, a long distance skip connection is included from the head section outputs to the body section outputs. The long distance skip connection adds the feature information of the input layer to the output layer of body, which constrains ResBlock modules to learn the difference between label images and input images. The head contains one convolution layer. Considering the time consuming of training, the body part stacks 16 ResBlocks arranged in the following order: [*Conv-BN-ReLU-Conv-BN*]. The tail contains one convolution layer. In sum, there are 34 convolution layers. In the network, a 3×3 convolution is used with a stride of 1 pixel and a zero-padding of 1 pixel to maintain the shape of feature maps, which allows UResnet to obtain inputs with arbitrary shapes.

In UResnet, the proposed EDL could be included, and the asynchronous training mode could be introduced. Because of the resounding effect of the BN layers on the underwater image enhancement task (Section 4.3), UResnet was designed with BN layers.

The detailed configuration of the proposed UResnet model is shown in Table 1, and the structure shown in Fig. 5.

IV. EXPERIMENTS

A. TESTING DATA SETS

The testing data sets are used to test the enhancement effect of the proposed models. The evaluation includes both subjective and objective evaluations. The testing data sets include two parts: (1) Testing data set A includes 221 underwater images for subjective and objective evaluation. Sample images from data set A are shown in Fig. 6. (2) Testing data set B includes 60 pairs of ''synthetic underwater images'' generated by CycleGAN and their corresponding in-air images, which

TABLE 1. Configuration of convolution layers in URESNET.

Layer Name	Number of kernels	Kernel size	Stride	Padding	Output size
Conv	64	3×3×3	1	1	w×h×64
ResBlock_1:Conv_1	64	3×3×64	1	1	w×h×64
ResBlock_1:Conv_2	64	3×3×64	1	1	w×h×64
…	…	…	…	…	…
ResBlock_16:Conv_1	64	3×3×64	1	1	w×h×64
ResBlock_16:Conv_2	64	3×3×64	1	1	w×h×64
Conv	3	3×3×64	1	1	w×h×3

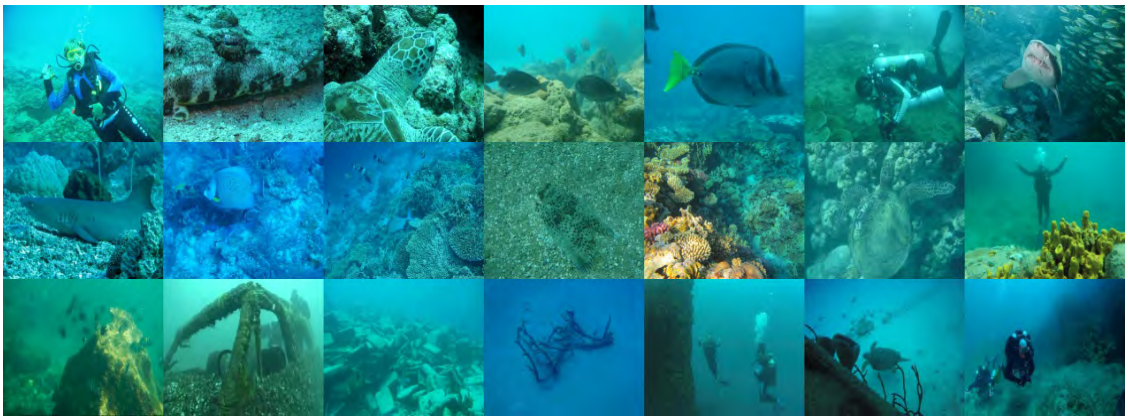


FIGURE 6. Sample images from data set A.

were not used for powerful supervised learning models. The images in testing data set B are similar to those in Fig. 2. Because they are pairs, they can be used for both objective evaluation, such as by PSNR [32] and structural similarity (SSIM) [36] evaluation, and subjective evaluation.

B. EXPERIMENTAL COMPARISONS OF THE ORIGINAL VDSR AND IMPROVED VDSR MODELS WITH THE PROPOSED EDL AND ASYNCHRONOUS TRAINING MODEL

To test the effects of the proposed EDL and asynchronous training mode, three models were trained: the original VDSR model, the VDSR model with the proposed EDL penalty term in the synchronous training mode (VDSR-P-S), and the one in the asynchronous training mode (VDSR-P-A).

The experiments were conducted on a workstation with an Intel Xeon E5-1630 processor and 3 NVIDIA 1080 Ti graphics cards. The experimental software environment is Ubuntu 16.04 with PyTorch 0.4.1.

The training data set consists of approximately 4000 pairs of images with a resolution of 256*256.

In order to determine the training configuration, a lot of experiments were conducted. Take the VDSR model for example, the training effect under different learning rates was

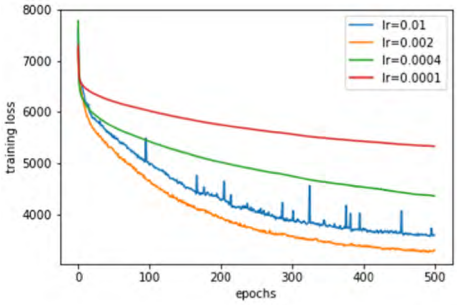


FIGURE 7. Loss curves of VDSR model at different learning rate.

tested as shown in Fig. 7. According to Fig. 7, the learning rate was determined to be 0.002. The learning rate curves of VDSR-P-S and VDSR-P-A are similar to those of VDSR. On all these models, 0.002 is the best choice for learning rate. Similarly, other experiments were conducted, such as the enhancement effects in different training epochs, and so on.

Based on the experimental results, each model uses the same training configuration. The size of each batch is 1 and the learning rate is 0.002. A stochastic gradient descent (SGD) optimizer was used. The learning rate decreases by 0.7 per-400 epochs. All three models were trained for 2000 epochs. For the VDSR-P-S model,

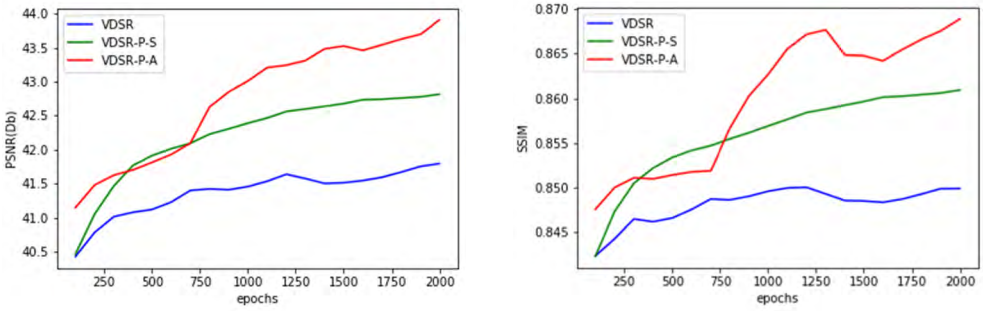


FIGURE 8. Curves of PSNR and SSIM scores for testing data set B during training of the VDSR, VDSR-P-S, and VDSR-P-A models.

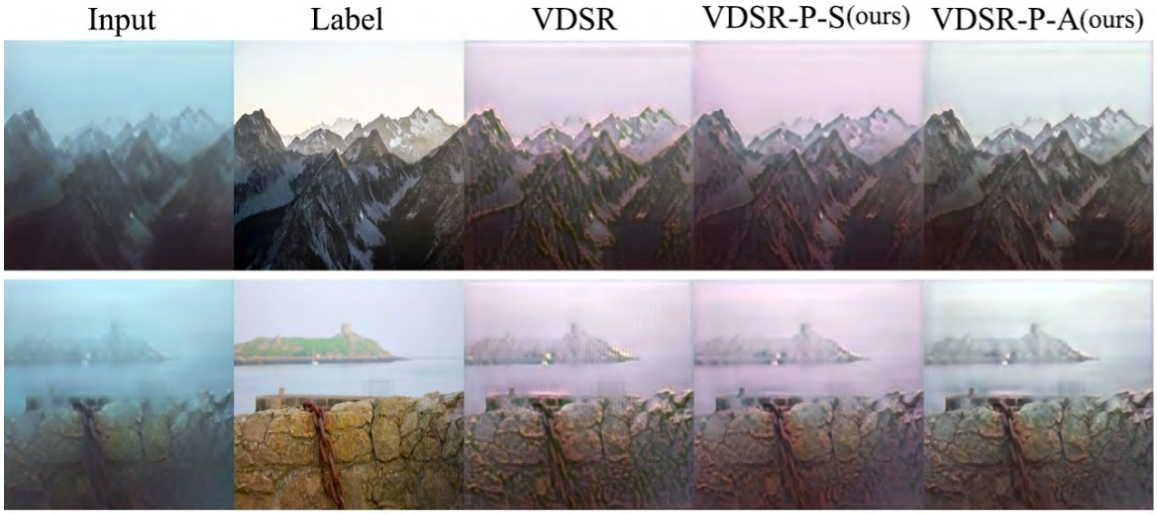


FIGURE 9. Comparison of the restoration effects of sample images in testing data set B by VDSR, VDSR-P-S, and VDSR-P-A.

10 different EDL weight coefficients were attempted between 0.0001 and 0.5. According to the experimental results, the selected weight coefficient of EDL was 0.07.

The performance of the three models was compared with regard to the PSNR and the SSIM in testing data set B under different training epochs. As shown in Fig. 8, the *VDSR* models with the proposed EDL achieve higher PSNR and SSIM scores than the original VDSR model under the same number of training epochs, and the VDSR-P-A achieves the highest scores overall. This suggests that EDL and asynchronous training can accelerate convergence and improve the training effect. The "synthetic underwater image" restoration effect of testing data set B is shown in Fig. 9.

In addition to the image restoration effect of testing data set B, the PSNR and SSIM scores represent the training level of the models. As seen in the PSNR scores, SSIM scores, and restoration effect, including EDL is beneficial to training; as shown in Table 2 the effect of the VDSR-P-A model is the best. As shown in Fig. 9, the color correction accuracy of the VDSR-P-A model is higher. As seen in the Loss curve in Fig. 10, the model converges faster and better when the proposed EDL is added in asynchronous training mode.

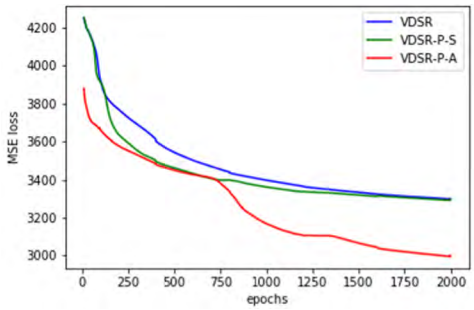


FIGURE 10. Loss curves of VDSR, VDSR-P-S, and VDSR-P-A.

TABLE 2. PSNR and SSIM scores of VDSR, VDSR-P-S, and VDSR-P-A on testing data set B (2000 EPOCHS).

Method	PSNR	SSIM
VDSR	41.7954	0.8498
VDSR-P-S	42.8171	0.8609
VDSR-P-A	**43.9115**	**0.8689**

The performance of the VDSR model was compared with our improved models on testing data set A, the real underwater images (Fig. 11). Fig. 12 shows the comparison of detail

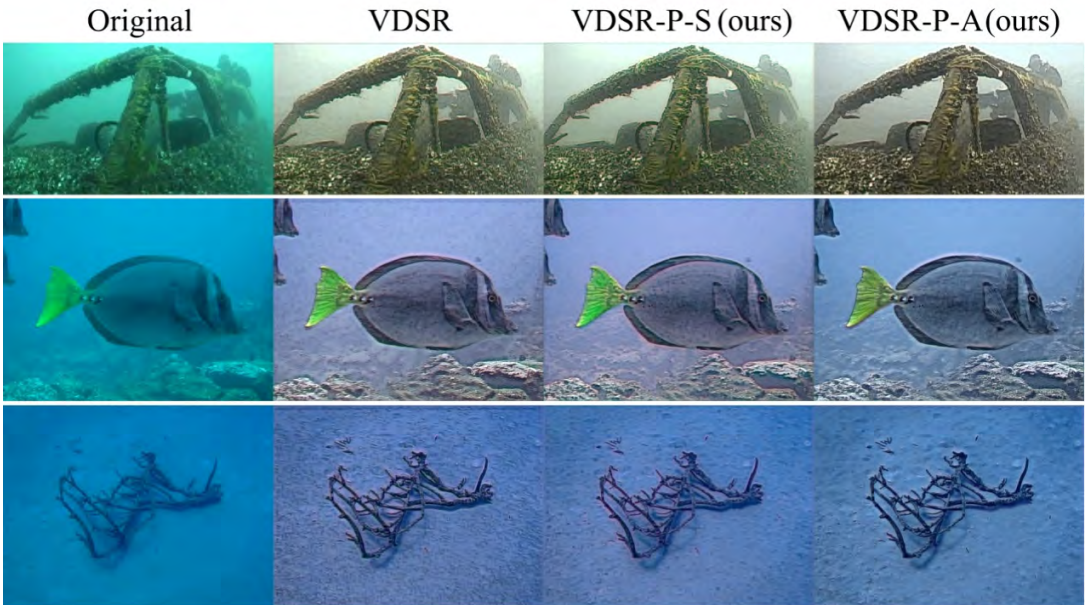


FIGURE 11. Comparison of enhancement effects of sample images in testing data set A by VDSR, VDSR-P-S, and VDSR-P-A.

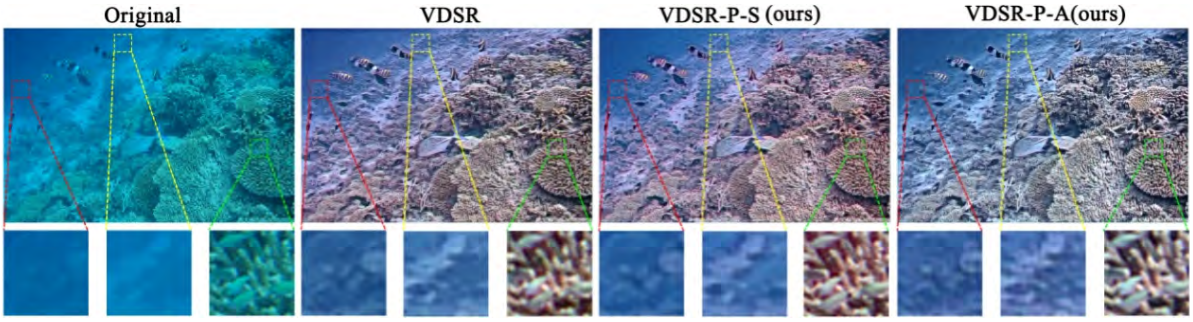


FIGURE 12. Comparison of detail enhancement effects by VDSR, VDSR-P-S, and VDSR-P-A.

TABLE 3. Visible edge scores of VDSR, VDSR-P-S, and VDSR-P-A on testing data set A (2000 EPOCHS).

Method	e	$\bar{r}$	δ
VDSR	2.228949	1.950276	0.005028
VDSR-P-S	2.432708	2.15557	0.006564
VDSR-P-A	**2.833926**	**2.667856**	**0.003343**

enhancement effects of the VDSR model and our improved models. The original trained VDSR model corrects the color deviation and enhances the details. After our improvement, the VDSR model's performance improves even further: the VDSR-P-A model has the best enhancement effect and the details are clearest.

Table 3 shows the visible edge [37] scores of the VDSR model and our improved models on testing data set A. The visible edge measurement does not need a true value, but can be calculated with the original image and the enhanced image. The visible edge measurement can evaluate edge details and contrast enhancement by parameters **e** and $\bar{r}$, respectively. Parameter δ represents the proportion of saturated pixels in the restored image. Higher values of e and $\bar{r}$, and a lower value of δ denotes better images quality. These parameters are widely used in the field of image dehazing. In Table 3, it can be seen that VDSR-P-A is superior to VDSR-P-S, and VDSR-P-S is superior to the original VDSR; the proposed VDSR-P-A model restores the most details. This is attributed to the better training supplied by EDL and the asynchronous training mode. Accordingly, with the same training data and number of training epochs, the VDSR model with EDL in asynchronous training mode learns more than the original VDSR model.

C. BN LAYERS AND THE PROPOSED URESNET MODEL

Batch Normalization (BN) [38] can reduce the difficulty of training deep neural networks. The effects of BN layers

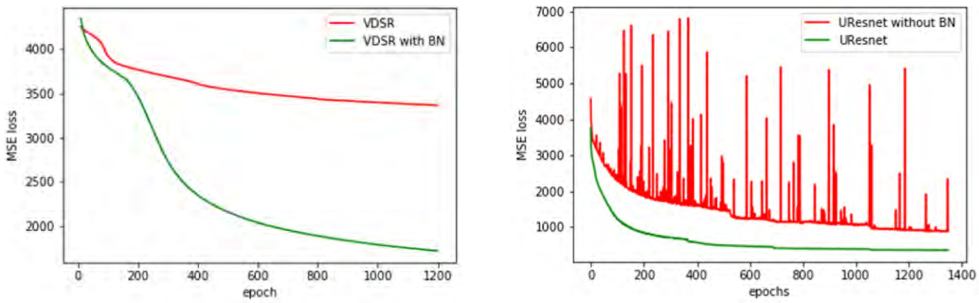


FIGURE 13. Loss curves of VDSR and UResnet with and without BN layers.

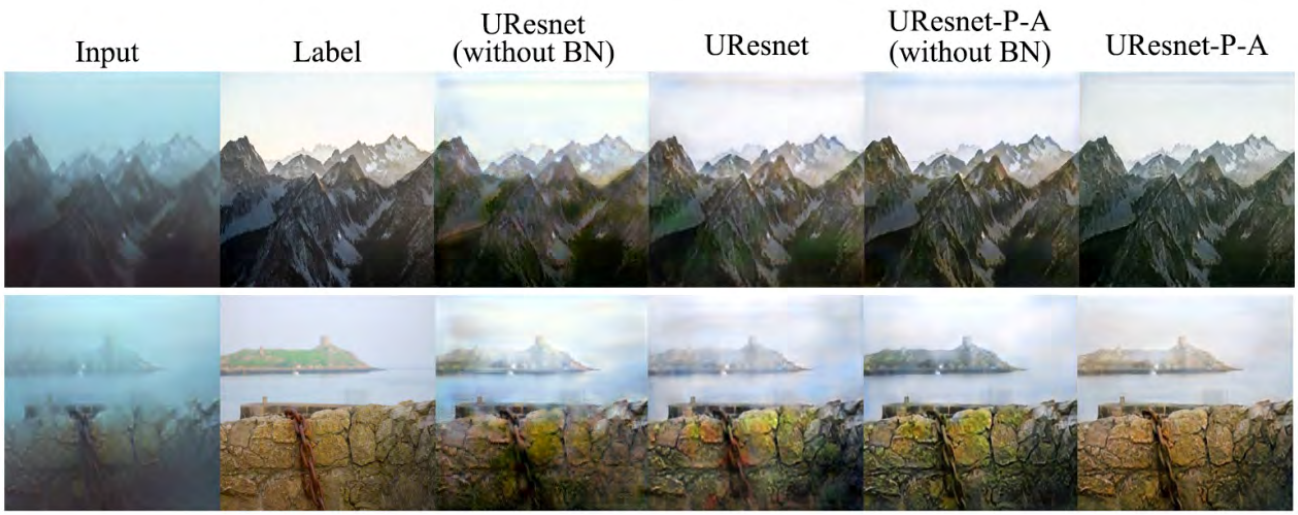


FIGURE 14. Comparison of the restoration effect of the proposed UResnet and UResnet-P-A, each with and without BN on testing data set B.

include accelerating convergence and regularization, which can prevent the trained model from over-fitting. Since it was proposed, BN has been widely applied to many CNN-based classification tasks. BN normalizes feature maps on spatial scales and batch dimensions in the network; when a batch size is 1, BN can be called IN (Instance Normalization) [39].

The original VDSR model and some other super-resolution reconstruction models, such as EDSR [33], do not use BN layers. This is because the BN layer normalizes the features, which diminishes the network's range flexibility [33]. For the image super-resolution reconstruction task, the output image needs to be consistent with the input in color, contrast, and brightness; the only adjustment should be to improve the resolution and details.

After the image passes through the BN layer, its color distribution is normalized, which is similar to contrast stretching. This affects the original contrast information of the image, so the BN layers reduce the quality of the output image in the image super-resolution reconstruction task.

Underwater image enhancement is different from the super-resolution reconstruction task. An underwater image has color deviation, blurring, and low contrast. Enhancing an underwater image thus requires color deviation correction and contrast improvement. Therefore, the addition of BN layer does not hinder the training, but rather improves the training effect and thus accelerate convergence. As shown in Fig. 13, it was found that the BN layer accelerates convergence. Additionally, the output has a better visual perceptual effect (Figs. 14 and 15).

For the VDSR and UResnet models on the underwater image enhancement task, BN layers are helpful for model training, as it can accelerate the convergence of the model and reduce the effect of oscillation. This informed the decision to include BN layers in the design of UResnet.

By including BN layers, the training process is more stable, and oscillation is reduced. Moreover, compared with the original model, the loss is reduced to a significantly lower level. From Table 4, it can be seen that with the inclusion of BN layers, the performance of the UResnet model in the training process is superior to that without BN layers. Accordingly, BN layers are conducive to training.

From Table 5, it can be seen that the training effect with the inclusion of BN layers is better than that without BN layers. Fig. 14 shows the image restoration effect of the proposed

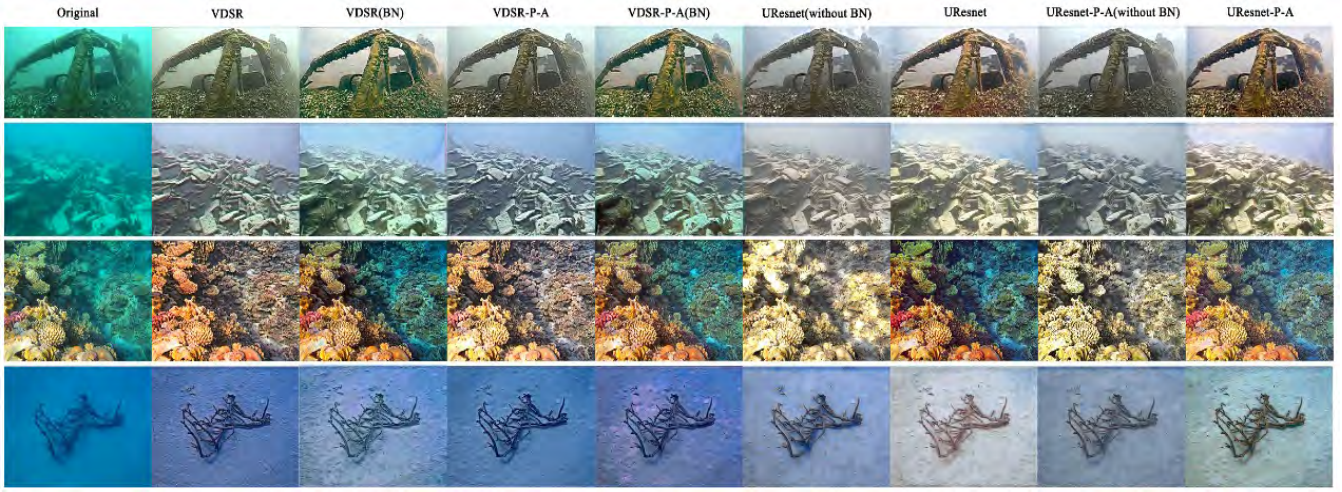


FIGURE 15. Comparison of enhancement effects of VDSR and UResnet models with and without BN layers on sample images on testing data set A.

TABLE 4. PSNR and SSIM scores of URESNET (without BN) and URESNET (with BN) on testing data set B for different training epochs.

Method	100epoch		300epoch		500epoch		700epoch	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
UResnet (without BN)	49.11	0.9056	50.83	0.9118	50.83	0.9142	51.10	0.9181
UResnet (with BN)	50.51	0.9156	52.22	0.9136	53.46	0.9289	53.71	0.9293

TABLE 5. PSNR and SSIM scores of VDSR and URESNET models with and without BN layers on testing data set B.

Method	PSNR	SSIM
VDSR	41.7954	0.8498
VDSR(BN)	45.7284	0.8818
VDSR-P-A	43.9115	0.8689
VDSR-P-A(BN)	**45.8858**	**0.8867**
UResnet(without BN)	51.0678	0.9162
UResnet	53.5879	0.9268
UResnet-P-A(without BN)	51.5362	0.9251
UResnet-P-A	**54.4502**	**0.9361**

TABLE 6. Visible edge scores of VDSR models and URESNET models with and without BN layers on testing data set A.

Method	e	$\bar{r}$	δ
VDSR	2.228949	1.950276	0.005028
VDSR(BN)	2.928499	2.252044	0.00068
VDSR-P-A	2.833926	2.667856	0.003343
VDSR-P-A(BN)	3.345048	2.514672	0.000692
UResnet(withoutBN)	2.188772	2.248601	0.003712
UResnet	2.953616	2.541302	0.001628
UResnet-P-A(withoutBN)	2.322385	2.44466	0.004049
UResnet-P-A	**3.507927**	**2.678004**	0.004495

UResnet models on testing data set B. The models with BN layers are better than those without BN layers. UResnet-P-A performs the best, with the restored images quality very close to the label images. Comparing the visual image quality of Fig. 14 with that of Fig. 9, the proposed UResnet models are better than the VDSR models. The performance of the VDSR models was compared with the proposed UResnet models with and without BN layers on testing data set A, the real underwater images (Fig. 15). From Fig. 15, through visual observation, the proposed UResnet models are better than VDSR models, and the BN layers are helpful for underwater image enhancement.

On the testing data set A (real underwater images), the returned visible edge evaluation scores are as expected, that is, the proposed UResnet models achieved higher scores than the VDSR models. The models with BN layers are better than those without BN layers, as shown in Table 6. It can be seen that the BN layers, the proposed EDL, and the asynchronous training mode can enhance image details. Subjectively, from Fig. 15, following the inclusion of the BN, the colors of the enhanced images are more pronounced, and the contrast is higher.

The inclusion of BN layers is beneficial to the task of underwater image enhancement, which is embodied by the following three aspects:

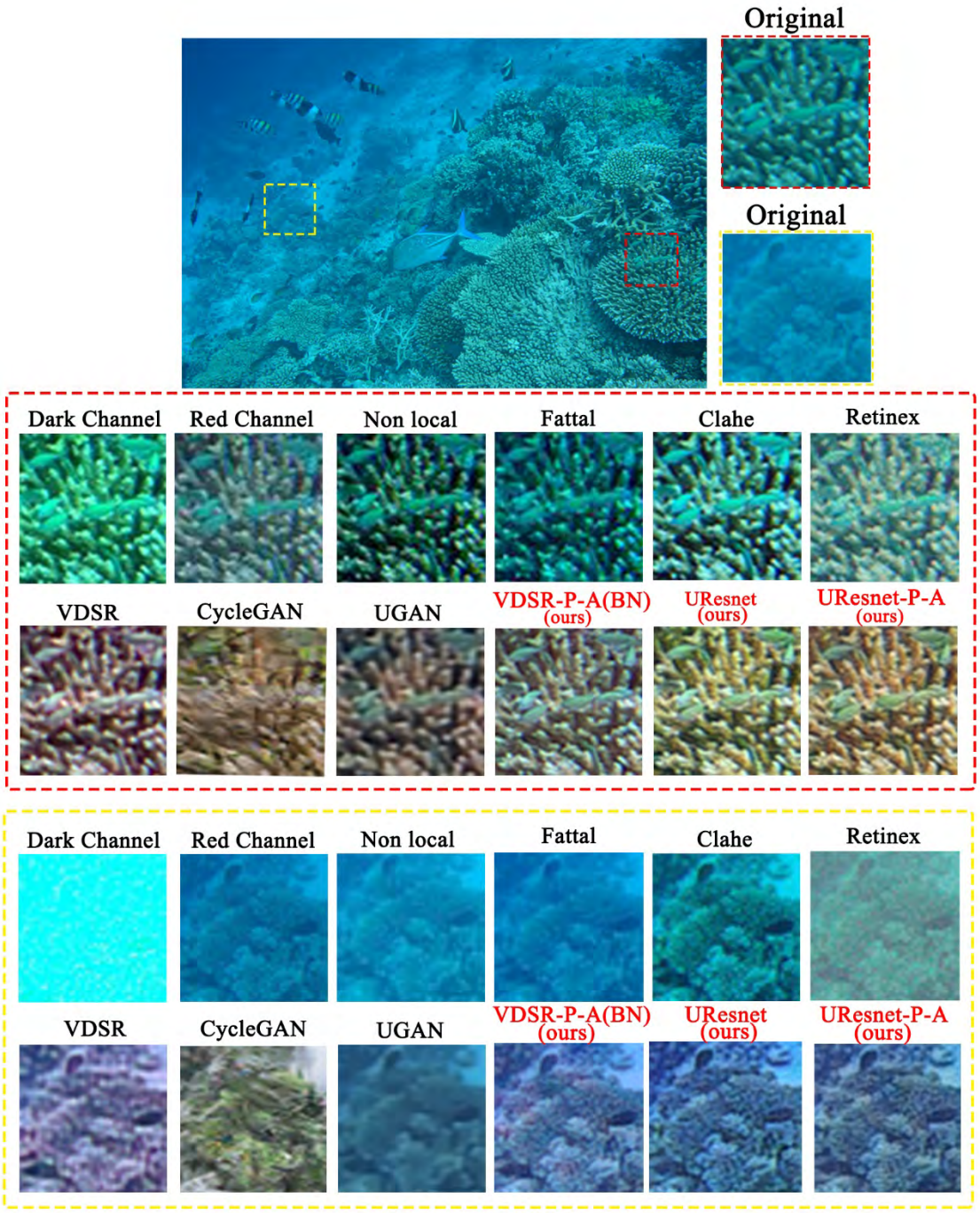


FIGURE 16. Detailed comparison of underwater image enhancement effects between the proposed methods and the other methods.

(1) BN layers are advantageous to training, and the models with BN layers converge faster.

(2) The models with BN layers restore more details, as seen in the higher the visible edge scores.

(3) BN layers enhance the contrast of the image and improve the color saturation.

In summary, the proposed UResnet models outperform the VDSR models in image enhancement, but the UResnet models are deeper than the VDSR model. Training and testing UResnets requires more time and more computing resources, as shown in Table 7. The inclusion of BN layers also increases time consumption.

D. COMPARATIVE EXPERIMENTS WITH OTHER METHODS

To further evaluate the enhancement effects, an underwater image quality measure (UIQM) [40] was employed to provide

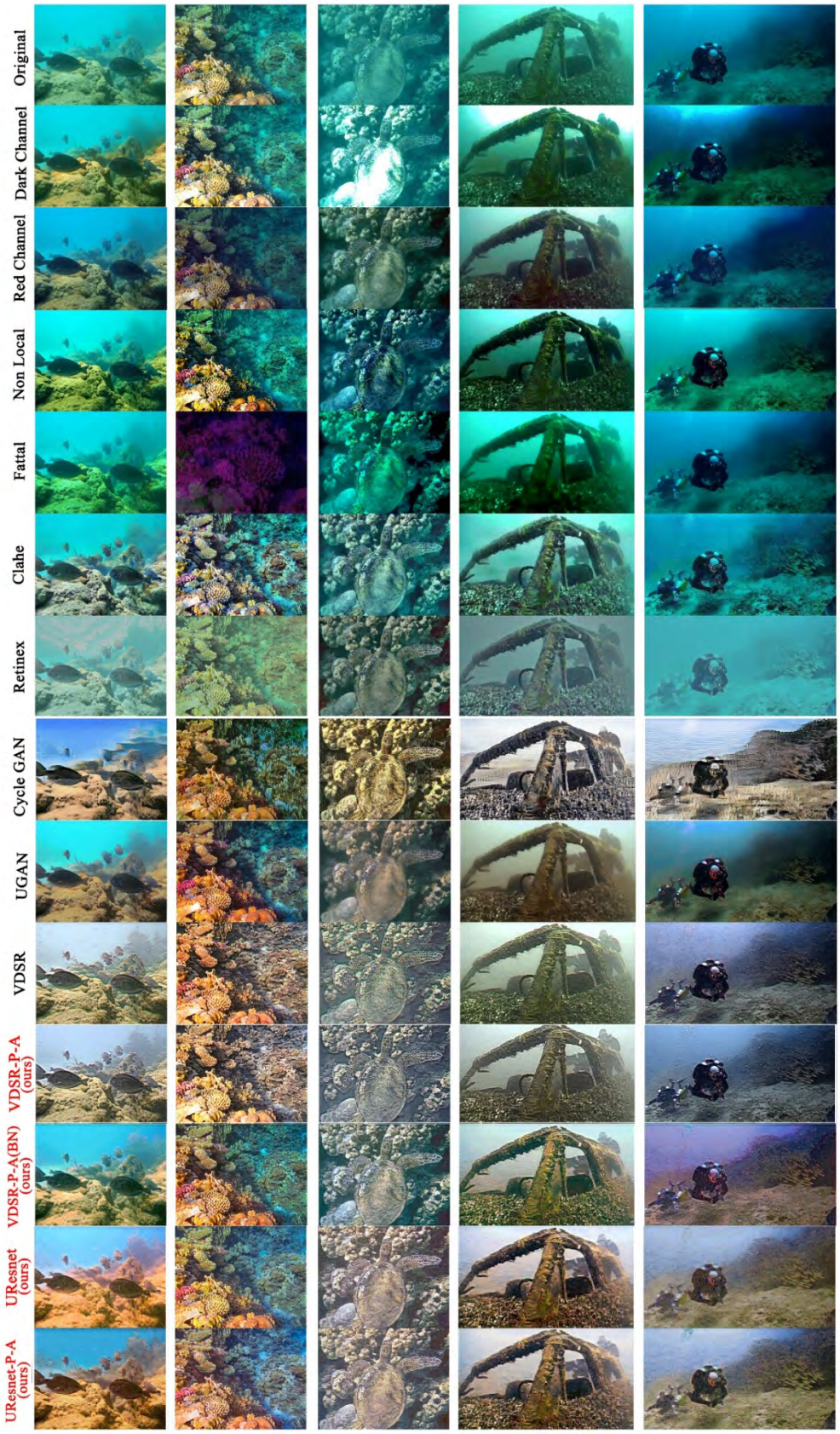


FIGURE 17. Comparison of enhancement effects of the proposed methods and other methods on sample images in testing data set A.

TABLE 7. Time consumption comparison of training and testing with different models.

Method	Training time consumption(s/epoch)	Testing time consumption(s/image)
VDSR	167.7253	0.001025041
VDSR(BN)	172.7353	0.002232181
VDSR-P-A	321.9669	0.001140436
VDSR-P-A(BN)	326.3544	0.002294064
UResnet(withoutBN)	294.3012	0.00236294
UResnet	354.3772	0.004727628
UResnet-P-A(withoutBN)	554.3045	0.002615558
UResnet-P-A	702.7952	0.004466772

TABLE 8. Visible edge scores on testing data set a from the proposed methods and other methods.

Method	e	$\bar{r}$	δ
Dark channel[13]	0.096252	1.064544	0.008245
Red channel [22]	0.634426	1.119801	0.181192
Non local [14]	0.692538	2.039882	0
Fattal [15]	-0.07745	3.355831	0.066663
Clahe [6]	1.2145430	1.298229	0.00000623
Retinex [9]	0.599079	1.195826	0.000187
CycleGAN [2]	2.732066	2.501146	1.467265
UGAN [29]	1.2133209	1.68746650	0.61078355
VDSR [3]	2.228949	1.950276	0.005028
VDSR-P-A	2.833926	2.667856	0.003343
VDSR-P-A(BN)	3.345048	2.514672	0.000692
UResnet	2.953616	2.541302	0.001628
UResnet-P-A	**3.507927**	**2.678004**	0.004495

no-reference metrics. The UIQM is comprised of three properties: the underwater image colorfulness measure (UICM), the underwater image sharpness measure (UISM), and the underwater image contrast measure (UIConM). Higher values of the UIQMs denote better image quality.

The performance of each model on testing set A can be seen in Table 8, Table 9, Fig. 16, and Fig. 17. The proposed UResnet-P-A model achieves the highest scores with the visible edge metric, showing excellent detail recovery. With the UIQM metrics, the proposed UResnet-P-A model also achieves the highest score.

Fig. 16 show the differences in detail restoration and contrast level of the enhanced image by the proposed methods and the others. There is a group of fish on the coral reef in the lower right corner (outlined with the red dotted line frame) of the image. In the original image, the color of the fish and the coral reef is similar and blurred. On the left side of the image (outlined with the yellow dotted line frame), the coral reefs are also blurred.

TABLE 9. UIQM scores on testing data set a with the proposed methods and other methods.

Method	UICM	UISM	UIconM	UIQM
Original	3.8791	4.7655	0.0402	3.1339
Dark channel[13]	5.2405	4.5348	0.0482	3.7098
Red channel [22]	7.2381	5.1463	0.0327	4.8085
Non local [14]	7.0958	4.8980	0.0137	4.6689
Fattal [15]	7.3229	3.4026	0.0025	4.3618
Clahe [6]	5.3820	6.0513	0.0505	4.1929
Retinex [9]	1.0908	6.2486	0.2664	2.2944
CycleGAN [2]	4.3100	8.3418	0.2258	4.3651
UGAN [29]	5.4489	7.0735	0.2572	4.5580
VDSR [3]	3.5367	7.5994	0.2301	3.8005
VDSR-P-A	3.6881	7.3863	0.2253	3.8116
VDSR-P-A(BN)	6.1184	7.4961	0.1943	4.9711
UResnet	6.0953	7.4800	0.2487	4.9699
UResnet-P-A	6.6041	7.5703	0.1971	**5.2196**

Seen from the detail enhancement images in the red dotted line frame, most of the deep learning methods perform better than traditional methods with regard to reducing color deviation. The best enhancement effect is the proposed UResnet-P-A method, followed by the proposed UResnet, and then the VDSR-P-A (BN) methods.

The enhanced image from the UResnet-P-A method has a natural color and a moderate contrast. Although the enhanced image from UGAN has color corrected to some extent, the enhanced image from UGAN is blurred and its effect is inferior to that of UResnet-P-A. With regard to detail restoration, the UResnet-P-A, UResnet, and VDSR-P-A (BN) images are visibly better than UGAN. The enhanced image of CycleGAN corrects the color deviation, but the details are blurred and even worse than the original image.

From the detailed enhancement of coral reefs in the yellow dotted line frame in Fig. 16, the proposed VDSR-P-A (BN), UResnet, and UResnet-P-A models enhanced detail the most; among the traditional methods, Clahe has the best effect, whereas Retinex generated visible noise; The enhancement effect of CycleGAN is fuzzy; except for some minor color changes, UGAN does not contain significant detail enhancement.

Thus, compared with other methods, the proposed UResnet-P-A model performs the best with regard to color correction and detail enhancement in underwater image enhancement. This is followed by the proposed UResnet and VDSR-P-A (BN) methods.

V. CONCLUSION

This paper proposes an underwater image enhancement solution by a deep residual framework. Firstly, CycleGAN was employed to generate synthetic underwater images as training

data for the CNN models. Secondly, the super-resolution reconstruction model VDSR was introduced into the field of underwater image enhancement, and the residual learning model, Underwater Resnet (UResnet) was proposed. Furthermore, the loss function and training mode were improved; a multi-term loss function was formed with the proposed edge difference loss (EDL) and MSE loss indices. An asynchronous training mode was also proposed to improve the performance of the multi-term loss function. The experimental results show the effectiveness of the proposed methods for underwater image restoration. EDL and the asynchronous training mode can improve the performances of CNN models on the underwater image enhancement task. The proposed UResnet-P-A model achieved the best performance with regard to both color correction and detail enhancement than the other methods we compared, followed by the proposed UResnet and VDSR-P-A (BN) models. It has also been shown that BN layers, though harmful to super-resolution reconstruction, are helpful in the underwater image enhancement task. BN layers can accelerate convergence in training. Furthermore, the inclusion of BN layers can assist in further restoring details and enhancing image contrast.

The proposed methods can significantly improve the visual effects of underwater images, which are helpful to the implementation of vision-based underwater tasks, such as segmentation and tracking. Furthermore, we consider applying the proposed methods to the similar domains, such as image dehazing and super-resolution reconstruction to test the generality of the proposed methods. We leave these to our future work.

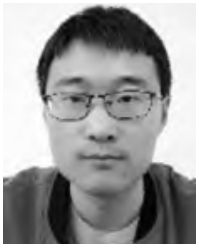
REFERENCES

[1] R. Schettini and S. Corchs, "Underwater image processing: State of the art of restoration and image enhancement methods," *EURASIP J. Adv. Signal Process.*, vol. 2010, Apr. 2010, Art. no. 746052.
[2] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," 2017, *arXiv:1703.10593*. [Online]. Available: https://arxiv.org/abs/1703.10593
[3] J. Kim, J. K. Lee, and K. M. Lee, "Accurate image super-resolution using very deep convolutional networks," in *Proc. Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Las Vegas, NV, USA, Jun. 2016, pp. 1646–1654.
[4] R. Hummel, "Image enhancement by histogram transformation," *Comput. Graph. Image Process.*, vol. 6, no. 2, pp. 184–195, 1977.
[5] S. M. Pizer, E. P. Amburn, J. D. Austin, R. Cromartie, A. Geselowitz, T. Greer, B. ter Haar Romeny, J. B. Zimmerman, and K. Zuiderveld, "Adaptive histogram equalization and its variations," *Comput. Vis., Graph., Image Process.*, vol. 39, no. 3, pp. 355–368, 1987.
[6] K. Zuiderveld, "Contrast limited adaptive histogram equalization," in *Graphics Gems IV*. San Diego, CA, USA: Academic, 1994, pp. 474–485.
[7] Y.-C. Liu, W.-H. Chan, and Y.-Q. Chen, "Automatic white balance for digital still camera," *IEEE Trans. Consum. Electron.*, vol. 41, no. 3, pp. 460–466, Aug. 1995.
[8] M. Ebner, *Color Constancy*. New York, NY, USA: Wiley, 2007.
[9] E. H. Land, "The Retinex theory of color vision," *Sci. Amer.*, vol. 237, no. 6, pp. 108–128, Dec. 1977.
[10] D. J. Jobson, Z.-U. Rahman, and G. A. Woodell, "A multiscale retinex for bridging the gap between color images and the human observation of scenes," *IEEE Trans. Image Process.*, vol. 6, no. 7, pp. 965–976, Jul. 1997.
[11] A. B. Petro, C. Sbert, and J.-M. Morel, "Multiscale retinex," *Image Process. Line*, vol. 4, pp. 71–88, Apr. 2014.
[12] S. Parthasarathy and P. Sankaran, "An automated multi scale retinex with color restoration for image enhancement," in *Proc. Nat. Conf. Commun. (NCC)*, Kharagpur, India, Feb. 2012, pp. 1–5.
[13] K. He, J. Sun, and X. Tang, "Single image haze removal using dark channel prior," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 12, pp. 2341–2353, Dec. 2011.
[14] D. Berman, T. Treibitz, and S. Avidan, "Non-local image dehazing," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Las Vegas, NV, USA, Jun. 2016, pp. 1674–1682.
[15] R. Fattal, "Single image dehazing," *ACM Trans. Graph.*, vol. 27, no. 3, pp. 1–9, 2008.
[16] M. Chambah, D. Semani, A. Renouf, P. Courtellemont, and A. Rizzi, "Underwater color constancy: Enhancement of automatic live fish recognition," *Proc. SPIE*, vol. 5293, pp. 157–168, Dec. 2003.
[17] B. Henke, M. Vahl, and Z. L. Zhou, "Removing color cast of underwater images through non-constant color constancy hypothesis," in *Proc. 8th Int. Symp. Image Signal Process. Anal. (ISPA)*, Trieste, Italy, Sep. 2013, pp. 20–24.
[18] K. Iqbal, R. A. Salam, A. Osman, and A. Z. Talib, "Underwater image enhancement using an integrated colour model," *IAENG Int. J. Comput. Sci.*, vol. 34, no. 2, pp. 239–244, 2007.
[19] A. Celebi and S. Ertürk, "Visual enhancement of underwater images using empirical mode decomposition," *Expert Syst. Appl.*, vol. 39, no. 1, pp. 800–805, 2012.
[20] C. Ancuti, C. O. Ancuti, T. Haber, and P. Bekaert, "Enhancing underwater images and videos by fusion," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Providence, RI, USA, Jun. 2012, pp. 81–88.
[21] S. Serikawa and H. Lu, "Underwater image dehazing using joint trilateral filter," *Comput. Elect. Eng.*, vol. 40, no. 1, pp. 41–50, 2014.
[22] A. Galdran, D. Pardo, A. Picón, and A. Alvarez-Gila, "Automatic red-channel underwater image restoration," *J. Vis. Commun. Image Represent.*, vol. 26, pp. 132–145, Jan. 2015.
[23] H. Lu, Y. Li, S. Mu, D. Wang, H. Kim, and S. Serikawa, "Motor anomaly detection for unmanned aerial vehicles using reinforcement learning," *IEEE Internet Things J.*, vol. 5, no. 4, pp. 2315–2322, Aug. 2018.
[24] H. Lu, Y. Li, M. Chen, H. Kim, and S. Serikawa, "Brain intelligence: Go beyond artificial intelligence," *Mobile Netw. Appl.*, vol. 23, no. 2, pp. 368–375, Apr. 2018.
[25] H. Lu, D. Wang, Y. Li, J. Li, X. Li, H. Kim, S. Serikawa, and I. Humar, "CONet: A cognitive ocean network," *IEEE Wireless Commun.*, to be published.
[26] H. Lu, Y. Li, T. Uemura, H. Kim, and S. Serikawa, "Low illumination underwater light field images reconstruction using deep convolutional neural networks," *Future Gener. Comput. Syst.*, vol. 82, pp. 142–148, May 2018.
[27] J. Li, K. A. Skinner, R. M. Eustice, and M. Johnson-Roberson, "WaterGAN: Unsupervised generative network to enable real-time color correction of monocular underwater images," *IEEE Robot. Autom. Lett.*, vol. 3, no. 1, pp. 387–394, Jan. 2018.
[28] C. Li, J. Guo, and C. Guo, "Emerging from water: Underwater image color correction based on weakly supervised color transfer," *IEEE Signal Process. Lett.*, vol. 25, no. 3, pp. 323–327, Mar. 2018.
[29] C. Fabbri, M. J. Islam, and J. Sattar, "Enhancing underwater imagery using generative adversarial networks," 2018, *arXiv:1801.04011*. [Online]. Available: https://arxiv.org/abs/1801.04011
[30] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-image translation with conditional adversarial networks," 2017, *arXiv:1611.07004*. [Online]. Available: https://arxiv.org/abs/1611.07004
[31] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial networks," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 3, 2014, pp. 2672–2680.
[32] I. Avcıbas, B. Sankur, and K. Sayood, "Statistical evaluation of image quality measures," *J. Electron. Imag.*, vol. 11, no. 2, pp. 206–223, Apr. 2002.
[33] B. Lim, S. Son, H. Kim, S. Nah, and K. M. Lee, "Enhanced deep residual networks for single image super-resolution," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Honolulu, HI, USA, Jul. 2017, pp. 1132–1140.
[34] C. Ledig, L. Theis, and F. Huszar, "Photo-realistic single image super-resolution using a generative adversarial network," in *Proc. Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Honolulu, HI, USA, Jul. 2017, pp. 105–114.
[35] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Las Vegas, NV, USA, Jun. 2016, pp. 770–778.
[36] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.

[37] N. Hautière, J.-P. Tarel, D. Aubert, and É. Dumont, "Blind contrast enhancement assessment by gradient ratioing at visible edges," *Image Anal. Stereol. J.*, vol. 27, no. 2, pp. 87–95, Jun. 2008.
[38] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," 2015, *arXiv:1502.03167*. [Online]. Available: https://arxiv.org/abs/1502.03167
[39] D. Ulyanov, A. Vedaldi, and V. Lempitsky, "Instance normalization: The missing ingredient for fast stylization," 2016, *arXiv:1607.08022*. [Online]. Available: https://arxiv.org/abs/1607.08022
[40] K. Panetta, C. Gao, and S. Agaian, "Human-visual-system-inspired underwater image quality measures," *IEEE J. Ocean. Eng.*, vol. 41, no. 3, pp. 541–551, Jul. 2015.
[41] L. Whitcomb, D. Yoerger, H. Singh, and J. Howland, "Advances in underwater robot vehicles for deep ocean exploration: Navigation, control, and survey operations," in *Robotics Research*. Springer, 2000, pp. 439–448. [Online]. Available: https://link.springer.com/chapter/10.1007%2F978-1-4471-0765-1_53
[42] T. Hyakudome, S. Tsukioka, H. Yoshida, T. Sawa, S. Ishibashi, A. Ishikawa, J. Ishiwata, K. Watanabe, M. Nakamura, and T. Aoki, "Autonomous underwater vehicle for surveying deep ocean," in *Proc. IEEE Int. Conf. Ind. Technol.*, Gippsland, VIC, Australia, Feb. 2009, pp. 1–6.
[43] M. Grasmueck, G. P. Eberli, D. A. Viggiano, T. Correa, G. Rathwell, and J. Luo, "Autonomous underwater vehicle (AUV) mapping reveals coral mound distribution, morphology, and oceanography in deep water of the Straits of Florida," *Geophys. Res. Lett.*, vol. 33, no. 23, 2006, Art. no. L23616.
[44] Y. Takemura, T. Himeno, and D. Ando, "Development of autonomous underwater vehicle 'OCTA' for coastal survey: Introduction of the 'OCTA' system and control experiment," in *Proc. Joint 7th Int. Conf. Soft Comput. Intell. Syst. (SCIS) 15th Int. Symp. Adv. Intell. Syst. (ISIS)*, Kitakyushu, Japan, Dec. 2014, pp. 236–241.
[45] K. He, X. Zhang, S. Ren, and J. Sun, "Identity mappings in deep residual networks," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*. Springer, 2016, pp. 630–645. [Online]. Available: https://link.springer.com/chapter/10.1007/978-3-319-46493-0_38

PENG LIU (M'19) received the B.S. degree in electronic information engineering and the master's degree in telecommunication and information system from the Ocean University of China, Qingdao, China, in 2004 and 2007, respectively. In 2007, he joined the Computing Center, Ocean University of China, where he is currently a Lecturer. He is pursuing the Ph.D. degree with the School of Information Science and Engineering, Ocean University of China. His current research interests include image processing, computer vision, and machine learning.

GUOYU WANG received the B.S. and M.S. degrees in physics from the Ocean College of Shandong, Qingdao, China, in 1984 and 1987, respectively. He received the Ph.D. degree from Twente University, the Netherlands, in 2000. He is currently a Professor with the Ocean University of China. His major research interests include image processing and pattern recognition.



HAO QI is currently pursuing the bachelor's degree with the Ocean University of China. His research interests include computer vision and deep learning.

CHUFENG ZHANG is currently pursuing the bachelor's degree with the Ocean University of China. His research interests include image processing, computer vision, and data mining.

HAIYONG ZHENG (M'12) received the B.S. degree in electronic information engineering and the Ph.D. degree in ocean information sensing and processing from the Ocean University of China, Qingdao, China, in 2004 and 2009, respectively. In 2009, he joined the Department of Electronic Engineering, Ocean University of China, where he is currently an Associate Professor. His research interests include image processing, computer vision, and machine learning.

ZHIBIN YU (M'16) received the master's degree in computer engineering and the Ph.D. degree in electrical engineering from Kyungpook National University, Daegu, Korea, in 2009 and 2016, respectively. In 2016, he joined the Department of Electronic Engineering, Ocean University of China, where, he is currently a Lecturer. His research interests include underwater image processing, artificial neural network, and underwater 3D reconstruction.