



# Article Adaptive Weighted Multi-Discriminator CycleGAN for Underwater Image Enhancement

# Jaihyun Park<sup>1</sup>, David K. Han<sup>2</sup> and Hanseok Ko<sup>1,\*</sup>

- <sup>1</sup> School of Electrical Engineering, Korea University, Seoul 02841, Korea
- <sup>2</sup> Information Science Division, ARL, Adelphi, MD 20783, USA
- \* Correspondence: hsko@korea.ac.kr; Tel.: +82-2-3290-3239

Received: 28 May 2019; Accepted: 26 June 2019; Published: 28 June 2019



**Abstract:** In this paper, we propose a novel underwater image enhancement method. Typical deep learning models for underwater image enhancement are trained by paired synthetic dataset. Therefore, these models are mostly effective for synthetic image enhancement but less so for real-world images. In contrast, cycle-consistent generative adversarial networks (CycleGAN) can be trained with unpaired dataset. However, performance of the CycleGAN is highly dependent upon the dataset, thus it may generate unrealistic images with less content information than original images. A novel solution we propose here is by starting with a CycleGAN, we add a pair of discriminators to preserve contents of input image while enhancing the image. As a part of the solution, we introduce an adaptive weighting method for limiting losses of the two types of discriminators to balance their influence and stabilize the training procedure. Extensive experiments demonstrate that the proposed method significantly outperforms the state-of-the-art methods on real-world underwater images.

Keywords: underwater; image enhancement; generative adversarial networks

# 1. Introduction

Technological advancement enables people to capture underwater images and videos in a wide variety of environments. As a result, underwater imaging has been an important issue of scientific research [1]. However, unlike in-air-captured images, visibility of underwater images is excessively compromised due to the attenuation of the propagated light, which results from the absorption and scattering effects. The absorption reduces the energy of light, whereas the scattering changes the direction of light propagation. They cause foggy appearance and contrast degradation, resulting in misty images at a distance. Furthermore, the colors of underwater images are faded as light absorption in the medium varies based on the wavelength. As a consequence, underwater images often show greenish or bluish tinge while red components are suppressed. It results in significant degradation of the underwater image quality and adversely affects performance of computer vision algorithms. To alleviate these problems, various underwater image enhancement methods have been presented.

Typical enhancement method is based on the optical process of underwater imaging, called underwater image formation model, which is formulated as:

$$I^{c}(x) = J^{c}(x)t^{c}(x) + B^{c}(1 - t^{c}(x)), \ c \in \{r, g, b\},$$
(1)

where  $I^{c}(x)$  is the observed intensity in color channel c,  $J^{c}(x)$  represents the true scene radiance,  $B^{c}$  means the global background light,  $t^{c}(x)$  is the medium transmission, and x indexes the pixel location in the

2 of 15

image. The medium transmission can be expressed as  $t^c(x) = e^{(-\beta^c d(x))}$ , where  $\beta^c$  is attenuation coefficient for channel *c* and d(x) means the scene depth. If the medium transmission and the global background light are properly estimated from the given turbid image *I*, the true scene radiance (clean image) can be recovered. Accurate estimations of global background light and medium transmission would result an effective underwater image enhancement. Meanwhile, approaches for the underwater image enhancement can be divided into two parts: handcrafted and learning-based methods. The former establishes prior knowledge to define the global background light and the medium transmission while the latter leverages dataset of image pairs to estimate these variables.

Handcrafted method: Bazeille et al. [2] proposed a filtering based underwater image enhancement method which includes a series of filters such as homomorphic filter and anisotropic filter. Ancuti et al. [3] proposed a method for enhancing underwater images using fusion technique. They derive the two images, which are color corrected image and contrast enhanced image, from original turbid image. Then, these two images are combined by four weight maps which are Laplacian contrast, local contrast, saliency, and exposedness. Lu et al. [4] proposed a method for enhancing underwater image using guided trigonometric bilateral filter and fast automatic color correction. They estimated the medium transmission using red color channel after applying  $\alpha$  ACE-based color correction. Then, medium transmission is refined by guided trigonometric bilateral filters [5]. Li [6] suggested different enhancement methods depending on color channel. Blue-green channels of turbid images are recovered via a dehazing algorithm based on modified dark channel prior (DCP) [7]. Then, red channel is corrected following the Gray-World (GW) assumption theory [8]. Chen et al. [9] proposed a region-specialized background light estimation method to deal with an inhomogeneous ambient light condition. Park et al. [10] leveraged the underwater image formation model then applied non-local means denoising method to the enhanced image to alleviate increased noise issue. To improve the underwater image formation model, Peng and Cosman [11] proposed a novel depth estimation method based on image blurriness and light absorption. Although the underwater image formation model is effective for many underwater scenes, there are claims that this model excludes some sources that affect the estimating of the variables. To accurately estimate the global background light and the medium transmission, some revised underwater image formation models are proposed [12,13], which supplement new variables in the model such as irradiance, sensor spectral response, and reflectance. However, as the model complexity is increased, it would overfit to the training data and may lead to poor generalization on new test data.

Learning-based method: As deep learning outperformed the handcrafted methods in various fields, generative adversarial networks (GAN) frameworks are also exploited to the underwater image enhancement task. Li et al. [14] embedded the underwater image formation model in the GAN structure to generate turbid underwater images from clean in-air images. This paired dataset is then leveraged to train convolutional neural network based underwater image enhancement model. On the contrary, Chen et al. [15] produced clean underwater images from turbid underwater images via the underwater image formation model. Then, they use this paired dataset to train conditional GAN for image enhancement. The main issue of the learning-based methods is that, including GAN, they require a paired image dataset to train the models. Collecting a sufficient amount of paired turbid and clean underwater images is practically impossible. To mitigate this problem, synthesized dataset can be exploited. However, unlike the simulation settings, accurately estimating the variables in the optical model is difficult in real environments. Therefore, learning-based models trained by synthetic datasets are effective on simulated image enhancement but probably less so on real-world image. These algorithms, when applied to real underwater environments, typically result in hazy images or unrealistic color shift. Zhu et al. [16] proposed the CycleGAN for image-to-image translation. In this framework, a cycle-consistency loss enables the network to be trained without paired dataset. Therefore, CycleGAN can be trained without turbid-clean image pairs. By constructing unpaired dataset using turbid underwater images and clean in-air images, Li et al. [17] designed an underwater image enhancement network without leveraging the underwater image formation model. Lu et al. [18] found direct mapping between turbid and clean underwater images via CycleGAN while estimating the medium transmission via the DCP to improve underwater image quality. Fabbri et al. [19] just exploited the CycleGAN to transfer clean underwater images to turbid underwater images then these synthesized pair of images are used for supervised training of conditional

in enhanced images. In this paper, we propose a novel learning-based underwater image enhancement method by solving the issues of GAN, which generate artifacts, shift color, and lose details. We exploit key advantages of the CycleGAN that does not need a paired dataset for training. While CycleGAN is trained to find a mapping function between unpaired images, content information in the original image, like shape and color of each instance, are usually not well recovered. Furthermore, artifacts and unrealistic color shift could appear. To mitigate this problem, we add a pair of discriminators in the CycleGAN to preserve content information of the original image. In the training phase of CycleGAN framework, we introduce an adaptive weighting method for loss of each discriminator to stabilize the training process and improve the visual quality of generated image. Since the entire CycleGAN's model is jointly trained, this framework is an end-to-end learning method. In the proposed architecture, conventional discriminators remove haze and correct colors while novel discriminators preserve content information in the image without producing any artifacts and unrealistic color change. Our key contribution in this work is that, by adding discriminators and an adaptive weighting method, influence of discriminators with different roles can be balanced while suppressing their negative side effects. As a consequence, the drawbacks of CycleGAN mentioned above are considerably alleviated. Figure 1 shows an example of enhanced result via the proposed method wherein left image is a turbid input and right image is an output of our enhancing method. An excessive greenish tinge of the input image is removed in the output image while faint colors in the input are recovered in the output.

GAN. However, these models still have drawbacks such as artifacts, unrealistic color shift or blurred effect



**Figure 1.** An example of underwater image enhancement result using the proposed method. (**left**): turbid input. (**right**): enhanced result.

## 2. Materials and Methods

## 2.1. Image Acquisition and Analysis

To train the proposed model, an unpaired underwater image dataset is established from Places dataset [20]. In the Places dataset, 'underwater-ocean deep' category includes 5000 different underwater images. We sort these images as turbid or clean according to their subjective quality. Specifically, five researchers classified these images as turbid or clean according to their perceived grades for clarity and color balance. Then, these images were re-categorized by a majority vote. Finally, 937 and 1014 images were sorted as clean and turbid in the training dataset respectively, while images that only receive 3 or less votes or contain very similar content with other images are excluded. For the quantitative evaluation, 890 underwater images in the benchmark dataset [21] are exploited with the underwater color image quality evaluation (UCIQE) [22] metric. UCIQE metric quantifies the nonuniform color cast, blurring, and low-contrast that characterize underwater images. For the qualitative evaluation, samples in the previous works [23] are used. Considering computational load, all the images are resized to  $256 \times 256$  pixels via bicubic downsampling. As images of  $256 \times 256$  size sufficiently represent detailed features of any object in the real-world, this image size is widely used for the deep learning based image to image translation task. Samples from our training dataset are illustrated in Figure 2.



**Figure 2.** Samples from the training dataset. (**First row**): images in the clean domain. (**Second row**): images in the turbid domain.

For the turbid domain *I* and the clean domain *J* in the training dataset, intensity histograms for the RGB channels are illustrated in Figure 3 where the intensity means pixel-wise average value for each domain. Clean underwater images in the dataset have a wide range of red colors while the red color histogram of turbid images is relatively narrow due to the severe attenuation of the propagated light. Therefore, translating an image from the clean domain to the turbid domain is relatively easy compared to translating an image from the turbid domain to the clean domain. In other words, estimating accurate colors of clean domain from images in the turbid domain is difficult without any ground truth during

the model training. Therefore, when CycleGANs are applied to the underwater image enhancement task, these models are prone to generate clean images from turbid images with unrealistic color change. Moreover, GAN's inherent drawbacks, producing a blurred image with artifacts, make enhancing the underwater image more difficult.

It can be assumed that an image is encoded as features in a content space and a style space according to the previous works [24,25]. If we define the content as underling features, such as original color, shape, and texture of each instance, and the style as rendering of the content, they can be disentangled from images by different types of networks. Following this assumption, images in the turbid domain and the clean domain would share the same content space as shown in Figure 4 because images in both domains are captured in underwater. For example, their images may include some objects like marine life, diver, or submarine. On the other hand, it is difficult for images in different domains to share the same style space as there is a considerable difference between their underwater environments. For instance, some underwater images show greenish tinge while some images have poor visibility due to the haze effect. Based on this assumption, we add a pair of discriminators in the CycleGAN's architecture as shown in Figure 5. Unlike conventional discriminators in the CycleGAN, these discriminators do not distinguish images as real or fake. They distinguish features in the content space as real or fake, instead. To fool these discriminators, generators try to produce clean underwater images while preserving content in turbid input images.



Figure 3. Average color histograms for the training dataset. (left): clean domain. (right): turbid domain.



Figure 4. Diagram for the relationship between the turbid domain and the clean domain in feature space.



**Figure 5.** Framework of the proposed method. (a) Overall framework. (b) Unfolded framework for  $I \rightarrow G(I) \rightarrow F(G(I))$  where dotted lines indicate losses.

## 2.2. Architecture of Multi-Discriminator CycleGAN

In this section, the architecture of our proposed CycleGAN is explained. As shown in Figure 5, our model consists of two generators and four discriminators. Generator *G* translates images in the turbid domain *I* into the clean domain *J*. Generator *F* transfers clean images *J* in the opposite direction. We refer to the conventional discriminators as style discriminators  $D_{stl_J}$  and  $D_{stl_I}$  because the main role of this pair is to distinguish real image from the generated images. The novel discriminators are termed content discriminators  $D_{cnt_J}$  and  $D_{cnt_I}$  as they receive features in the content space as inputs to distinguish real image's features. To ensure that content features are preserved in the generated images, a pre-trained VGG-16 [26] model is added in the content discriminator. Thus, the VGG-16 model is excluded from the training of model parameters. The type of loss function, which is a least-squares loss, and the network structures of generators and discriminators are the same as those of the basic CycleGAN [16].

# 2.2.1. Style Discriminator

We employ the style discriminators for turbid underwater images to remove haze and correct colors. As these discriminators receive an image as input, they focus on how realistic the input image is. To fool the style discriminators, generators are trained to mimic the style of the images in target domain while less attention is given to retaining the contents in the input image. For this reason we refer to the conventional discriminators in CycleGAN as the style discriminators. For the mapping function  $G : I(turbid) \rightarrow J(clean)$  and its style discriminator  $D_{stl}$ , the style loss,  $L_{stl}$ , is defined as follows:

$$L_{stl}(G, D_{stl\_J}, I, J) = \mathbb{E}_{j \sim p_{data(j)}}[(D_{stl\_J}(j) - 1)^2] + \mathbb{E}_{i \sim p_{data(i)}}[D_{stl\_J}(G(i))^2],$$
(2)

where  $j \sim p_{data(j)}$  and  $i \sim p_{data(i)}$  indicate image sample j and i from the clean domain and the turbid domain respectively and  $\mathbb{E}$  means expectation. The adversarial loss for the mapping function  $F: J(clean) \rightarrow I(turbid)$  and its style discriminator  $D_{stl \ I}$  is defined in the same way.

#### 2.2.2. Content Discriminator

To preserve the content information, low-level features are extracted from the input image and the corresponding generated image via the pre-trained VGG-16. Specifically, the content feature is extracted via the first convolution and activation function, which is rectified linear unit (ReLU). We call this feature extraction stage ReLU1-1. The deeper the layer we choose, the more network concentrates on global features and patterns. As a result, detailed local contents in the output image starts to degrade with deeper layers, because the network begins to capture only broader patterns. Since the cycle-consistency loss and the style loss are reasonably effective in preserving larger patterns, we focus on low-level features as content information to be recovered by the content loss. The extracted features from original images are considered as real and features from generated images are defined as fake. Therefore, generators *G* and *F* try to fool the content discriminators  $D_{cnt_J}$  and  $D_{cnt_J}$  by producing fake images containing original contents of the inputs. For the mapping function  $G : I(turbid) \rightarrow J(clean)$  and its content discriminator  $D_{cnt_J}$ , the content loss,  $L_{cnt}$ , is formulated as follows:

$$L_{cnt}(G, D_{cnt_{J}}, I, J) = \mathbb{E}_{\phi(j) \sim p_{data(\phi(j))}}[(D_{cnt_{J}}(\phi(j)) - 1)^{2}] + \mathbb{E}_{\phi(i) \sim p_{data(\phi(i))}}[D_{cnt_{J}}(\phi(G(i)))^{2}],$$
(3)

where  $\phi(\cdot)$  means the VGG-16 feature extractor.  $\phi(j) \sim p_{data(\phi(j))}$  and  $\phi(i) \sim p_{data(\phi(i))}$  indicate content feature sample  $\phi(j)$  and  $\phi(i)$  from the clean domain and the turbid domain respectively. The adversarial loss for the mapping function  $F : J(clean) \rightarrow I(turbid)$  and its content discriminator  $D_{cnt_I}$  is defined in the same way.

## 2.3. Overall Loss Function

The structures of the proposed CycleGAN is trained by the loss function as follows:

$$L = \lambda_{cnt} L_{cnt} + \lambda_{stl} L_{stl} + \lambda_{cyc} L_{cyc}, \tag{4}$$

where  $L_{cyc}$  indicates the cycle-consistency loss, which guarantees the individual mapping between samples of different domains. Without this loss, generators are prone to transfer a certain type of images in the input domain into random images in the target domain. By forcing an image produced by generators to recover an original image, this loss can reduce the space of possible mapping via generators. As networks are trained to reduce the overall loss, if a certain loss forms a big part of the overall loss, the networks are trained to satisfy a certain aspect emphasized by that loss while the others are ignored. Therefore, to balance the effects of two types of discriminators during the training phase, adaptive weights  $\lambda_{cnt}$  and  $\lambda_{stl}$  are leveraged. The weight for the cycle-consistency loss,  $\lambda_{cyc}$ , is fixed.

# Adaptive Weighted Adversarial Loss

We assume that CycleGANs usually generate blurry images because they concentrate on transferring the style between the different image domains. Thus, generated turbid or clean images lose original content information and have fewer edges than real input images because the content information in images, such as texture and shape of any object, can be represented by the edge features. If contents of the input image are well preserved in the generated image, this tendency would be alleviated. This assumption can be demonstrated by histograms for the amount of detected edges in an image. To detect edges, RGB color images are converted to grayscale. Once the RGB color images in the test dataset are converted to grayscale, we then make edge maps via Sobel edge detector. Specifically, this detector calculates gradient of each pixel and assigns a value 0 or 1 to each pixel according to the pre-defined threshold where we choose default threshold of the TensorFlow built-in function. As these edge maps are binary images, we

can easily count the number of the detected edge pixels in the image by summing up total pixel intensities in image. As illustrated in Figure 6, images in the test dataset have the largest number of edges on average while outputs of the proposed CycleGAN are second and outputs of the CycleGAN are placed third. Based on this experimental result, we apply an adaptive weighting method according to the amount of detected edges in the image and balance the adversarial losses of style discriminators and content discriminators. To calculate each weight of adversarial loss, edge maps of input images j, i and output images F(j), G(i)are extracted via Sobel edge detector.

$$\lambda_{cnt} = \alpha * \frac{\sum_{x=1}^{X} (\pi(j(x)) + \pi(i(x)))}{\sum_{x=1}^{X} (\pi(j(x)) + \pi(i(x))) + \sum_{x=1}^{X} (\pi(F(j(x))) + \pi(G(i(x))))},$$
  

$$\lambda_{stl} = \beta * \frac{\sum_{x=1}^{X} (\pi(F(j(x))) + \pi(G(i(x))))}{\sum_{x=1}^{X} (\pi(f(j(x))) + \pi(i(x))) + \sum_{x=1}^{X} (\pi(F(j(x))) + \pi(G(i(x))))},$$
(5)

where  $\pi(\cdot)$  is the Sobel edge detector that produces a binary edge map.  $\alpha$  and  $\beta$  are constants for scaling and X means the number of pixels in the image. As the fractions of  $\lambda_{cnt}$  and  $\lambda_{stl}$  are the complementary events, when one weight increases at the same time the other decreases. By the effect of adaptive weighting, if generated images F(j), G(i) severely lose information of the original contents including edges, weight for the content discriminators,  $\lambda_{cnt}$ , is increased while weight for the style discriminators,  $\lambda_{stl}$ , is decreased. On the other hand, if generated images retain sufficient content information and it is better to pay attention to the style transfer, the adaptive weighting works in the opposite way.



**Figure 6.** Histograms for the average proportion of detected edge pixels in grayscale image. (a) Edge histogram for original images in the test dataset. (b) Edge histogram for enhanced images by the CycleGAN. (c) Edge histogram for enhanced images by the proposed model.

Although the adaptive weighting method contributes to improving the performance of underwater image enhancement and mitigating the problem caused by unstable learning, it may have negative effects in the early stage of learning because adaptive weights could intensify fluctuations of the loss. Thus, constants  $\alpha$  and  $\beta$  should be carefully selected.

#### 3. Results

In this section, we demonstrate the effectiveness of the proposed method against the state-of-the-art approaches for the real-world underwater image enhancement. Namely, we show the validity of VGG-16 features as a representative content information indicator. Additionally, effects of the content discriminators and the adaptive weighting method in our CycleGAN framework are to be shown.

#### 3.1. Experimental Settings

For the CycleGAN training, we select  $\lambda_{cyc}$  as 10 and exploit the Adam solver of the TensorFlow where the initial learning rate and the momentum term for this solver are 0.0002 and 0.5, respectively. Batch size is 1 and epoch is set to 200. Both the training procedure and the performance of image enhancement are susceptible to the adaptive weights  $\lambda_{cnt}$  and  $\lambda_{stl}$ . In general, training procedure of GANs is very sensitive even for small changes in some constants. Therefore, we changed  $\alpha$  and  $\beta$  at 0.1 intervals respectively. Then,  $\alpha = 0.5$  and  $\beta = 1.5$  have shown to be the best for these weights. All the parameters are selected via cross-validation.

# 3.2. Ablation Study

To verify the effectiveness of the content discriminators and the adaptive weighting method, we compare the image enhancement results of the following models. (a) **CycleGAN**: basic CycleGAN model which does not include the content discriminators. (b) **CycleGAN+CD+W0.5**: CycleGAN model including the content discriminators where the weights for the content discriminators and the style discriminators are fixed to 0.5 and 1, respectively. (c) **CycleGAN+CD+W1**: CycleGAN model containing the content discriminators where the weights for the two types of discriminators are equal to 1. (d) **CycleGAN+CD+AW**: the proposed model that includes the adaptive weighted adversarial losses for the content discriminators and the style discriminators.

As shown in Figure 7a, the CycleGAN generates images with unrealistic colors and artifacts. This tendency results from the loss function and the composition of training dataset. Specifically, as many images in the clean domain contain transparent color like gray or white, the generator *G* tries to color the instances in turbid images as these colors. By doing this, the generator may fool the discriminator that the loss can be minimized regardless of the true colors of the instances in turbid images. The resultant images typically shift their color spectrum while their content information such as edges may get lost as shown in Figure 7a. To mitigate these problems, the content discriminators we added in the CycleGAN framework preserve the fine details as shown in Figure 7b–d. As the influence of the content discriminators increases from (b) to (c) in Figure 7, the generator excessively focuses on maintaining original contents of the input images to reduce the loss for the content discriminators. As a result, influence of the style discriminators is considerably weakened and haze effect and blueish or greenish tinges persist in the output images as indicated by Figure 7c. As shown in Figure 7d, the proposed method produces relatively hazeless and color corrected images without unrealistic color change by balancing the influences of the content discriminator.

To verify the validity of VGG-16's low-level feature as a content feature extractor, we compare the underwater image enhancement results of following models. Because our final goal is to design the CycleGAN model in which the contents of the input image are well preserved. Therefore, we consider that the comparison between the enhancement results of each model is more meaningful than the evaluation of the extracted feature itself. (a) **Conv1-1**: layer after ReLU1-1, which is our choice. (b) **Conv2-1**: layer after ReLU2-1. (c) **Conv3-1**: layer after ReLU3-1. (d) **Conv4-1**: layer after ReLU4-1. As the feature extraction layer becomes deeper, the features lose a specific content information. As a consequence, it is difficult for generated images to retain original color of each instance in the input images as shown in Figure 8. These experimental results demonstrate that low-level features of the VGG-16, extracted from the layer after ReLU1-1, are appropriate for mapping an image into the content space.



Figure 7. Underwater image enhancement results according to the content discriminators and the adaptive weighting method. (a) CycleGAN. (b) CycleGAN+CD+W0.5. (c) CycleGAN+CD+W1. (d) CycleGAN+CD+AW.

# 3.3. Quantitative Evaluation

For the UCIQE metric, we compare the proposed method with four state-of-the-art methods: GW [8], DCP [7], CAP [27], and CycleGAN [16]. As shown in Table 1, GW shows the lowest performance while the proposed method achieves the best performance for the test dataset compared to the other methods considered here.

**Table 1.** Underwater image enhancement evaluation based on UCIQE [22] metric. Larger metric means better result.

	Input	GW [8]	DCP [7]	CAP [27]	CycleGAN [16]	Ours
UCIQE	26.08	26.83	30.44	29.71	31.73	32.00



**Figure 8.** Underwater image enhancement results according to the feature extraction layer in the VGG-16. (a) Conv1-1. (b) Conv2-1. (c) Conv3-1. (d) Conv4-1.

# 3.4. Qualitative Evaluation

In this section, performance of the proposed method is qualitatively verified. We compare the proposed method with five state-of-the-art methods: Ancuti and Ancuti [28], Galdran et al. [29], Emberton et al. [30], Ancuti et al. [23], and Zhu et al. [16]. Eight samples and the corresponding enhanced results are illustrated in Figure 9. For the results of Ancuti and Ancuti [28], artifacts or color shift are rarely observed. However, haze effect and bluish or greenish tinges still remain. For the bluish images, specifically the first image and the second image in Figure 9, results of Galdran et al. [29] seem better than those of Ancuti and Ancuti [28] with respect to the color correction. However, their method is less effective for the greenish images like the sixth, seventh, and eighth. In general, Emberton et al. [30]'s results lose color balance as red components are over-emphasized both on objects and backgrounds. Results of Ancuti et al. [23] seem effective in terms of dehazing and correcting colors. Nevertheless, their method also produces reddish outputs. The method of Zhu et al. [16], which is the CycleGAN, generates images with unnatural colors and this tendency is particularly serious when turbidity of an image is strong. On the other hand, our method produces haze-free images with minimal observable artifacts and unrealistic color change. In addition, it performs better in delivering contrast-enhanced images than the other state-of-the-art methods while it better suppresses the tendency of over-emphasizing red colors.





Figure 9. Underwater image enhancement results. (a) Ancuti and Ancuti [28]. (b) Galdran et al. [29]. (c) Emberton et al. [30]. (d) Ancuti et al. [23]. (e) Zhu et al. [16]. (f) Our results.

## 4. Discussion

In our experiments, the baseline CycleGAN's outputs look somewhat worse than samples presented in the previous works [17,18]. This may be because of the small-quantity and biased training dataset. So far, creating a dataset that includes richness and a variety of turbid and clean underwater images is difficult and time consuming. Nevertheless, our proposed method overcomes the data limitation and significantly improves the CycleGAN's performance. Therefore, if a rich dataset is given, we expect that the proposed method will achieve even better results. To stabilize the initial training phase, we heuristically chose the scaling constants for each adaptive weight. However, as shown in Figure 7d, tiny black artifacts are present in some enhanced images. Specifically, artifacts are observed at the bottom left of the first image, at the bottom right of the second image, and at the top left of the third image. The training issue could have produced these artifacts. Thus, there is still some room for improvement with regard to the training method. Furthermore, our results in Figure 9f are slightly blurred compared to those of Ancuti et al. [23]. Although we address the blurring issue by introducing the content discriminators, there are still some improvements to be made.

## 5. Conclusions

We propose a new underwater image enhancement method. Specifically, we improve architecture of the CycleGAN to generate a style-transferred image while preserving content information in the original image. The style discriminators force the generators to make color corrected and dehazed underwater images. Novel discriminators, called content discriminators, contribute to maintaining content in turbid input images like original color, shape, and texture of each instance. Additionally, our novelty resides with the introduction of the adaptive weighted loss for balancing the influence of two types of discriminators. The adaptive weighting exploits strengths of each discriminator while suppressing their negative effects. Experimental results demonstrate that the proposed method significantly outperforms the state-of-the-art methods on real-world underwater images, both quantitatively and qualitatively.

Author Contributions: Conceptualization, J.P. and D.K.H.; Methodology, J.P. and D.K.H.; Software, J.P.; Validation, J.P. and D.K.H.; Writing-original draft, J.P., D.K.H., and H.K.; Writing-review & editing, J.P., D.K.H., and H.K.; Visualization, J.P. and H.K.; Funding acquisition, D.K.H. and H.K.

**Funding:** This research was funded by the Air Force Office of Scientific Research under award number FA2386-19-1-4001.

Acknowledgments: This material is based upon work supported by the Air Force Office of Scientific Research under award number FA2386-19-1-4001.

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

## References

- 1. Kocak, D.M.; Dalgleish, F.R.; Caimi, F.M.; Schechner, Y.Y. A focus on recent developments and trends in underwater imaging. *Mar. Technol. Soc. J.* **2008**, *42*, 52–67. [CrossRef]
- Bazeille, S.; Quidu, I.; Jaulin, L.; Malkasse, J.P. Automatic underwater image pre-processing. In Proceedings of the Characterisation du Milieu Marin (CMM'06), Brest, France, 16–19 October 2006.
- Ancuti, C.; Ancuti, C.O.; Haber, T.; Bekaert, P. Enhancing underwater images and videos by fusion. In Proceedings of the 2012 IEEE Conference on Computer Vision and Pattern Recognition, Providence, RI, USA, 16–21 June 2012.
- Lu, H.; Li, Y.; Serikawa, S. Underwater image enhancement using guided trigonometric bilateral filter and fast automatic color correction. In Proceedings of the International Conference Image Processing, Melbourne, VIC, Australia, 15–18 September 2013.

- 5. Tomasi, C.; Manduchi, R. Bilateral filtering for gray and color images. In Proceedings of the International Conference Image Processing, Bombay, India, 7 January 1998.
- 6. Li, C.; Quo, J.; Pang, Y.; Chen, S.; Wang, J. Single underwater image restoration by blue-green channels dehazing and red channel correction. In Proceedings of the IEEE International Conference Acoustics, Speech and Signal Processing, Shanghai, China, 20–25 March 2016.
- He, K.; Sun, J.; Tang, X. Single image haze removal using dark channel prior. *IEEE Trans. Pattern. Anal. Mach. Intell.* 2011, 33, 2341–2353. [PubMed]
- Provenzi, E.; Gatta, C.; Fierro, M.; Rizzi, A. A spatially variant white-patch and gray-world method for color image enhancement driven by local contrast. *IEEE Trans. Pattern. Anal. Mach. Intell.* 2008, 30, 1757–1770. [CrossRef] [PubMed]
- 9. Chen, Z.; Wang, H.; Shen, J.; Li, X.; Xu, L. Region-specialized underwater image restoration in inhomogeneous optical environments. *Optik* 2014, 125, 2090–2098. [CrossRef]
- 10. Park, D.; Han, D.K.; Ko, H. Enhancing underwater color images via optical imaging model and non-local means denoising. *IEICE Trans. Inf. Syst.* 2017, 100, 1475–1483. [CrossRef]
- 11. Peng, Y.T.; Cosman, P.C. Underwater image restoration based on image blurriness and light absorption. *IEEE Trans. Image Process.* **2017**, *26*, 1579–1594. [CrossRef] [PubMed]
- 12. Akkaynak, D.; Treibitz, T. A revised underwater image formation model. In Proceedings of the IEEE Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018.
- 13. Zhang, M.; Peng, J. Underwater image restoration based on a new underwater image formation model. *IEEE Access* **2018**, *6*, 58634–58644. [CrossRef]
- 14. Li, J.; Skinner, K.A.; Eustice, R.M.; Johnson-Roberson, M. WaterGAN: Unsupervised generative network to enable real-time color correction of monocular underwater images. *IEEE Robot. Autom. Lett.* **2018**, *3*, 387–394. [CrossRef]
- 15. Chen, X.; Yu, J.; Kong, S.; Wu, Z.; Fang, X.; Wen, L. Towards real-time advancement of underwater visual quality with GAN. *IEEE Trans. Ind. Electron.* **2019**. [CrossRef]
- Zhu, J.Y.; Park, T.; Isola, P.; Efros, A.A. Unpaired image-to-image translation using cycle-consistent adversarial networks. In Proceedings of the IEEE International Conference Computer Vision, Venice, Italy, 22–29 October 2017.
- 17. Li, C.; Guo, J.; Guo, C. Emerging from water: Underwater image color correction based on weakly supervised color transfer. *IEEE Signal. Process. Lett.* **2018**, *25*, 323–327. [CrossRef]
- 18. Lu, J.; Li, N.; Zhang, S.; Yu, Z.; Zheng, H.; Zheng, B. Multi-scale adversarial network for underwater image restoration. *Opt. Laser. Technol.* **2019**, *110*, 105–113. [CrossRef]
- Fabbri, C.; Islam, M.J.; Sattar, J. Enhancing underwater imagery using generative adversarial networks. In Proceedings of the IEEE International Conference on Robotics and Automation, Brisbane, Australia, 21–25 May 2018.
- 20. Zhou, B.; Lapedriza, A.; Khosla, A.; Oliva, A.; Torralba, A. Places: A 10 million image database for scene recognition. *IEEE Trans. Pattern. Anal. Mach. Intell.* **2018**, 40, 1452–1464. [CrossRef] [PubMed]
- 21. Li, C.; Guo, C.; Ren, W.; Cong, R.; Hou, J.; Kwong, S.; Tao, D. An underwater image enhancement benchmark dataset and beyond. *arXiv* **2019**, arXiv:1901.05495.
- 22. Yang, M.; Sowmya, A. An underwater color image quality evaluation metric. *IEEE Trans. Image Process.* 2015, 24, 6062–6071. [CrossRef] [PubMed]
- 23. Ancuti, C.O.; Ancuti, C.; De Vleeschouwer, C.; Bekaert, P. Color balance and fusion for underwater image enhancement. *IEEE Trans. Image Process.* **2018**, *27*, 379–393. [CrossRef] [PubMed]
- 24. Huang, X.; Liu, M.Y.; Belongie, S.; Kautz, J. Multimodal unsupervised image-to-image translation. In Proceedings of the European Conference on Computer Vision, Beijing, China, 20–24 August 2018.
- 25. Gatys, L.A.; Ecker, A.S.; Bethge, M. Image style transfer using convolutional neural networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016.
- 26. Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. *arXiv* **2014**, arXiv:1409.1556.

- 27. Zhu, Q.; Mai, J.; Shao, L. A fast single image haze removal algorithm using color attenuation prior. *IEEE Trans. Image Process.* **2015**, *24*, 3522–3533. [PubMed]
- 28. Ancuti, C.O.; Ancuti, C. Single image dehazing by multi-scale fusion. *IEEE Trans. Image Process.* **2013**, *22*, 3271–3282. [CrossRef] [PubMed]
- 29. Galdran, A.; Pardo, D.; Picón, A.; Alvarez-Gila, A. Automatic red-channel underwater image restoration. *J. Vis. Commun. Image Represent.* 2015, *26*, 132–145. [CrossRef]
- Emberton, S.; Chittka, L.; Cavallaro, A. Hierarchical rank-based veiling light estimation for underwater dehazing. In Proceedings of the BMVC, Swansea, UK, 8–10 September 2015.



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).