

Article

An Underwater Image Enhancement Method for a Preprocessing Framework Based on Generative Adversarial Network

Xiao Jiang ¹, Haibin Yu ^{1,2} , Yaxin Zhang ¹, Mian Pan ¹, Zhu Li ¹ , Jingbiao Liu ³ and Shuaishuai Lv ^{1,*} 

¹ College of Electronics and Information, Hangzhou Dianzi University, Hangzhou 310018, China; 2143040030@hdu.edu.cn (X.J.); shoreyhb@hdu.edu.cn (H.Y.); bruce_zyx@hdu.edu.cn (Y.Z.); ai@hdu.edu.cn (M.P.); lz1126@hdu.edu.cn (Z.L.)

² Ningbo Institute of Oceanography, Ningbo 315832, China

³ Ocean Technology and Equipment Research Center, Hangzhou Dianzi University, Hangzhou 310018, China; ab@hdu.edu.cn

* Correspondence: lvshuai@hdu.edu.cn

Abstract: This paper presents an efficient underwater image enhancement method, named ECO-GAN, to address the challenges of color distortion, low contrast, and motion blur in underwater robot photography. The proposed method is built upon a preprocessing framework using a generative adversarial network. ECO-GAN incorporates a convolutional neural network that specifically targets three underwater issues: motion blur, low brightness, and color deviation. To optimize computation and inference speed, an encoder is employed to extract features, whereas different enhancement tasks are handled by dedicated decoders. Moreover, ECO-GAN employs cross-stage fusion modules between the decoders to strengthen the connection and enhance the quality of output images. The model is trained using supervised learning with paired datasets, enabling blind image enhancement without additional physical knowledge or prior information. Experimental results demonstrate that ECO-GAN effectively achieves denoising, deblurring, and color deviation removal simultaneously. Compared with methods relying on individual modules or simple combinations of multiple modules, our proposed method achieves superior underwater image enhancement and offers the flexibility for expansion into multiple underwater image enhancement functions.



Citation: Jiang, X.; Yu, H.; Zhang, Y.; Pan, M.; Li, Z.; Liu, J.; Lv, S. An Underwater Image Enhancement Method for a Preprocessing Framework Based on Generative Adversarial Network. *Sensors* **2023**, *23*, 5774. <https://doi.org/10.3390/s23135774>

Academic Editors: Dongsheng Zhang and Zhilong Su

Received: 7 May 2023

Revised: 16 June 2023

Accepted: 17 June 2023

Published: 21 June 2023



Copyright: © 2023 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 (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: underwater image enhancement; convolutional neural network (CNN); generative adversarial networks (GANs); feature extraction; cross-stage fusion

1. Introduction

Underwater vision has emerged as a vital tool for exploring the marine environment, offering non-invasive access and rich information content. High-quality underwater images play a crucial role in providing valuable information for underwater robots engaged in various missions, such as underwater exploration, marine archaeology, underwater rescue, and underwater imaging. However, underwater images are prone to color distortion, low contrast, and motion blur due to the absorption and scattering of light [1–5]. Consequently, enhancing the quality of underwater optical imaging has become a significant research focus in the field of ocean engineering and computer vision [6–10].

Over the past few decades, researchers and experts have devoted substantial efforts to underwater image enhancement. Various methods can be categorized as physical model-based [1–5], non-physical model-based [6–9], and data-driven approaches [10–22]. Physical model-based methods treat underwater image enhancement as an inverse problem and rely on estimating the parameters of the image formation model to obtain clear underwater images. One of the most widely used methods is the dark channel prior (DCP) model, including underwater dark channel prior (UDCP) [1] and generalized dark channel prior (GDCP) [2]. Another approach considers the optical characteristics of underwater images, such as forecasting scene transmission using the attenuation differences of RGB color

channels [3], underwater image enhancement based on minimum information loss and histogram prior distribution considering the relationship between background color and optical properties of the medium [4,5], and underwater image color correction methods [6]. However, physical model-based methods suffer from two main drawbacks. First, they often require prior knowledge due to the complex underwater environment, limiting their effectiveness. Second, these methods rely on simplified image formation models that assume uniform decay coefficients across color channels, resulting in unstable enhancements and suboptimal visual effects.

In contrast to physical model-based methods, non-physical model-based approaches directly manipulate pixel values to enhance underwater images. These methods include white balance [7], grayscale world theory [8], and histogram equalization [9]. For instance, a white balance method presented in [7] enhances underwater images through gamma correction and sharpening, whereas a variational Retinex method proposed in [8] involves color correction, layer decomposition, and late enhancement. Furthermore, a two-step approach combining color correction and contrast enhancement was introduced in [9]. However, these methods often overlook the unique optical properties of underwater images, leading to color biases and oversaturation or undersaturation in different regions. They also fail to fully consider the degradation mechanisms, resulting in red artifacts and potential introduction of pseudo-targets.

As deep learning methods are widely used in image processing and computer vision, enhanced methods for underwater images have been developed rapidly. Chang et al. [23] systematically discussed the application of generative adversarial networks (GANs) in scene graphs. Zhang et al. [24] highlighted the transferability aspect of the proposed method. In zero-shot learning [25] and unsupervised methods [26], GANs demonstrate their flexibility in application. By designing end-to-end networks, complex underwater image degradation models can be avoided. Typical underwater image enhancement methods include GAN-based and convolutional neural network-based (CNN-based) models. Fabbri et al. [10] employed conditional GAN (cGAN) to tackle underwater image enhancement as an image-to-image conversion problem. Building upon this work, Yu et al. [11] incorporated perceived loss into the cGAN framework for underwater image color correction. Subsequent studies explored multi-scale dense GANs for underwater image enhancement [12,13]. Liu et al. proposed a deep multiscale feature fusion network for underwater image color correction. However, GAN-based underwater image enhancement methods heavily rely on aligned underwater image pairs, which are often challenging to obtain. Additionally, most of these methods primarily focus on color correction and overlook overall detail enhancement, leading to suboptimal visual effects. To address these issues, unsupervised underwater image enhancement methods have been proposed [12,15,16]. Li et al. [15] introduced WaterGAN, a GAN-based algorithm that generates realistic underwater images from aerial images and their corresponding depth maps, providing unsupervised color correction. Islam et al. [16] proposed a rapid underwater image enhancement method for real-time preprocessing in an autonomous pipeline using a CycleGAN-based visual-guided underwater robot. Furthermore, Hong et al. [20] proposed a weakly-supervised underwater image enhancement (WSUIE) method to reduce reliance on raw/enhanced underwater image alignment. However, the practical application of these algorithms is significantly limited due to the difficulty in obtaining corresponding depth maps.

In CNN-based underwater image enhancement, Anwar et al. [18] introduced UWCNN, an end-to-end model trained on a synthetic underwater image dataset. Sun et al. [19] proposed a deep model for underwater image enhancement using an encoder–decoder framework with skip connections to preserve low-level features and expedite training. Wu et al. [21] developed an underwater image CNN (UWCNN-SD) based on structure decomposition to address color distortion, blurred details, and low contrast in underwater images. Li et al. [22] proposed an adaptive algorithm utilizing random wired neural networks (RWNN) and co-evolution (SE) to perform color adjustment, contrast improvement, luminance enhancement, and detail enhancement with edge retention techniques.

However, despite the powerful feature extraction capabilities of CNNs, most CNN-based underwater image enhancement algorithms operate in the RGB color space. Although the RGB color space can address scattering issues and improve color deviation, it fails to directly capture critical parameters related to quality degradation, such as low contrast, saturation, and luminance.

Given the complex light propagation characteristics in the ocean, leading to severe color distortion, low contrast, and motion blur in underwater robot photography, this paper proposes an efficient image preprocessing framework called ECO-GAN based on a generative adversarial network. The framework enables blind enhancement of individual underwater images. ECO-GAN incorporates a CNN to tackle three underwater problems: dynamic blur, low brightness, and color bias. Firstly, we utilize the U-Net structure to sample the input underwater image and extract features of different scales through the encoder. Subsequently, by analyzing the multi-scale features and under the constraint of the objective function, we alleviate the quality issues while upsampling and recovering the image to its original size, thereby achieving blind image denoising (removing unknown noise from noisy images). The contributions of this article are as follows:

- (1) We propose a generative adversarial network capable of efficiently addressing various image enhancement tasks for underwater images with multiple defects.
- (2) The designed image enhancement model enables simultaneous enhancement of three distinct image enhancement tasks. By employing feature extraction through repeated encoders and dedicated decoders for different enhancement tasks, computational complexity is reduced and inference speed is improved.
- (3) ECO-GAN incorporates cross-stage fusion modules between decoders for multiple enhancement tasks, thereby enhancing the interconnection between decoders and improving the quality of output images.

2. Underwater Image Enhancement Network

Figure 1 presents the block diagram of the ECO-GAN structure for underwater image enhancement. ECO-GAN consists of a generator and three discriminators. During the training stage, original underwater images are input to the generator to learn the corresponding multiscale residual map, recovering three potentially clear images at different upsampling stages. These results are then passed to the corresponding discriminator, which independently assesses their authenticity at different scales. In the inference stage, the generator enables end-to-end underwater image enhancement for input underwater images. Now, we will delve into the details of the generator and the discriminator.

2.1. Generator

The generator structure of ECO-GAN, as depicted in Figure 2, comprises three parts: downsampled feature extraction, upsampled image recovery, and cross-stage fusion. The main functions of the generator include luminance enhancement, blur removal, and white balance.

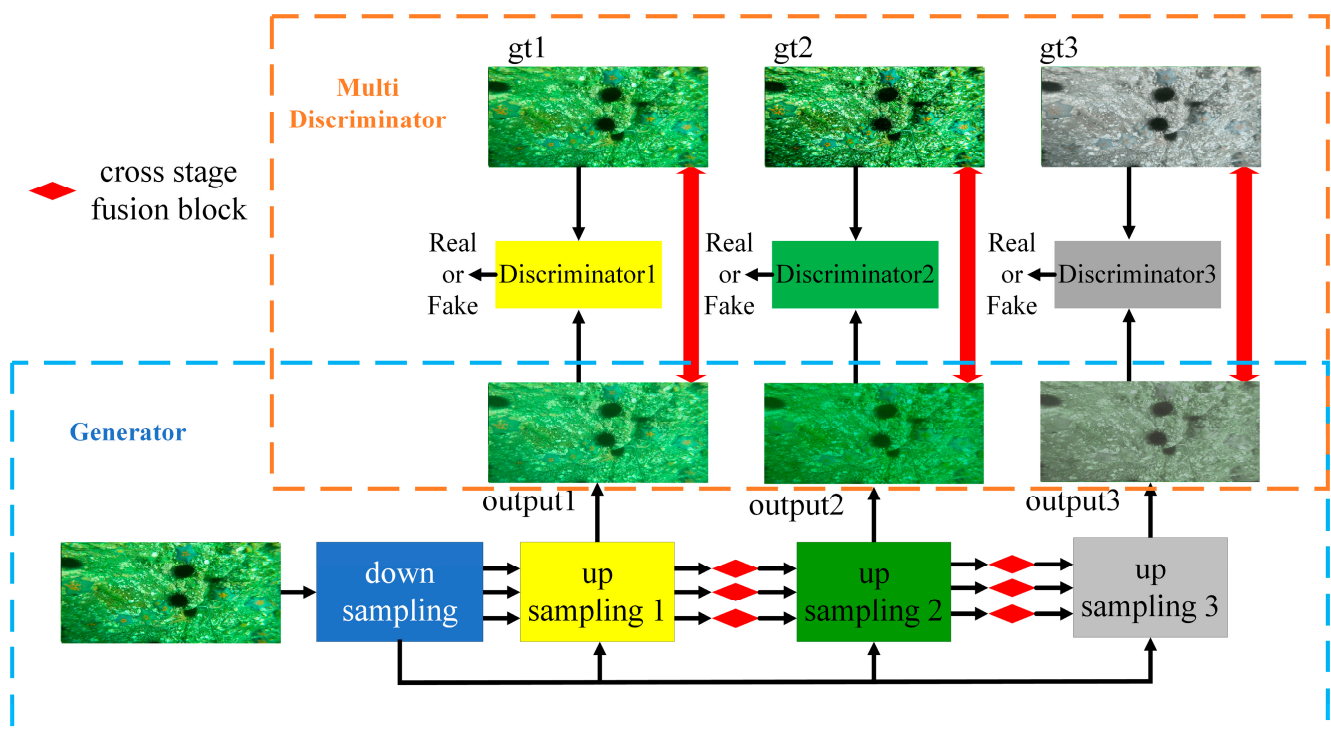


Figure 1. Overview of the proposed ECO-GAN for underwater image enhancement.

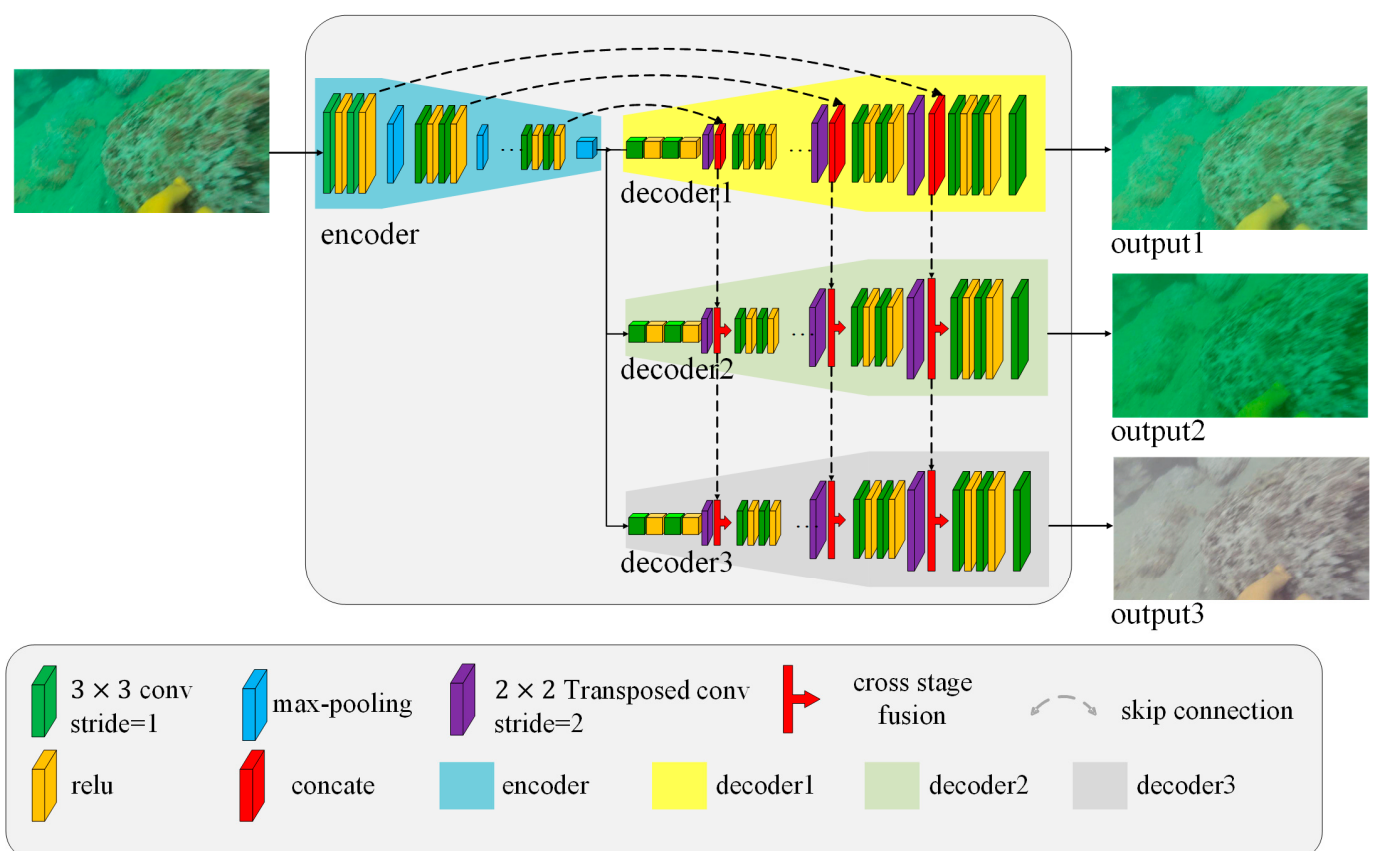


Figure 2. Structure of ECO-GAN generator.

To reduce computational complexity, we reused the backbone feature extraction network and extracted features only once. This design was motivated by Zhang et al. [27], who

implemented three separate functions and combined them in series. However, repeatedly extracting features three times can be computationally expensive. In our approach, we employed a classical U-Net structure and proposed a cross-stage fusion module to establish connections between the three branches. The U-Net structure resembles the letter “U” and enables efficient feature extraction and image recovery. The cross-stage fusion module is to combine the feature extracted by current-stage network and last-stage network.

As shown in Figure 2, our model is designed to realize three tasks and reuse backbone for reducing calculations. To recover the raw input image I_{in} to the final image I_f , we utilize a function $F(\cdot)$ to extract features at different scales and a function $G(\cdot)$ to recover them.

$$I_f = G(F(I_{in})) \quad (1)$$

For the three specific tasks, we introduce three different functions, namely, G_1 , G_2 , and G_3 , to recover extracted features. Our target tasks are specified by paired datasets.

$$output1 = G_1(F(I_{in})) \quad (2)$$

$$output2 = G_2(F(I_{in})) \quad (3)$$

$$output3 = G_3(F(I_{in})) \quad (4)$$

where $output1$, $output2$, and $output3$ are outputs of G_1 , G_2 , and G_3 , respectively.

To integrate these tasks, we need G_t , which utilizes the features extracted by F and recovered by G_1 , G_2 , and G_3 at different scales. We can express the estimate of the final enhanced image as follows:

$$\begin{aligned} \hat{I}_f &= G_t(F(I_{in}), G_1(F(I_{in})), G_2(F(I_{in})), G_3(F(I_{in}))) \\ &= G_t'(F(I_{in})) \end{aligned} \quad (5)$$

where $\hat{I}_f$ is the estimate of I_f ; G_t is a function that can utilize features extracted by F and recovered by G_1 , G_2 , and G_3 at different scales; and G_t' is an alternate form of G_t .

To optimize the model and reduce computational costs, we leverage a backbone feature extraction network to extract features at different scales in the downsampling stage. The backbone feature extraction network adopts a modular design, incorporating convolutional layers, batch normalization layers, and activation layers. The outputs of each downsampling module are connected to the corresponding module in the upsampling image recovery stage through skip connections. This approach effectively addresses the issue of gradient disappearance. In the upsampling stage of image recovery, we employ three upsampled image recovery branches that correspond to the three image enhancement tasks: luminance enhancement, blur removal, and white balance. The image recovery branch also adopts a modular design, where each module consists of deconvolution, batch normalization layers, and activation layers.

Both the backbone feature extraction network and the image recovery branches follow a U-Net structure. The downsampling module utilizes skip connections to propagate extracted features to the subsequent upsampling recovery module, mitigating gradient disappearance and guiding the direction of image repair. This is crucial because the upsampling module outputs variations in the image rather than the image itself.

In the upsampling stage, an input image undergoes three separate image repair branches to achieve blur removal, low-light enhancement, and color bias correction. To improve the quality of image recovery, a cross-stage fusion module is incorporated to integrate the recovery features from different enhancement tasks. The cross-stage fusion module concatenates input features from different stages, adding two channels of ‘ch’ along the channel dimension to aid the recovery task at the current stage. This results in a channel number of $2 \times \text{ch}$ while maintaining the original width and height of the

image. The number of channels is subsequently reduced to ‘ch’ through convolutional layers, generating features that retain the original input feature size and contain multistage information. Additionally, the cross-stage fusion module stacks the features of the previous stage and the current stage. By introducing the features of the previous stage, it assists with the recovery task of the current stage and further integrates the features through convolution, maintaining the feature dimension. As a single model, the entire image enhancement model can accomplish various enhancement tasks, striking a balance between computational efficiency and performance.

2.2. Discriminator

The ECO-GAN discriminator distinguishes image blocks of size 70×70 to obtain local-scale discrimination results and combines them with a standard global discriminator to assess the entire image. This approach yields a two-scale least squares discriminator. In this study, we adopt a two-scale least squares discriminator to determine the authenticity of the image. The two-scale representation assesses the authenticity of the input image from both a global semantic level and local detail level. This guides the generator to generate real images at both scales, facilitating the production of realistic image details that improve visual results. Since the ECO-GAN handles three image enhancement tasks, we utilize three multilayer perceptrons as discriminant networks to evaluate the output of each generator separately. The structure of the discriminator is illustrated in Figure 3, where ‘output i ’ ($i = 1, 2, 3$) refers to the three images generated by the generator.

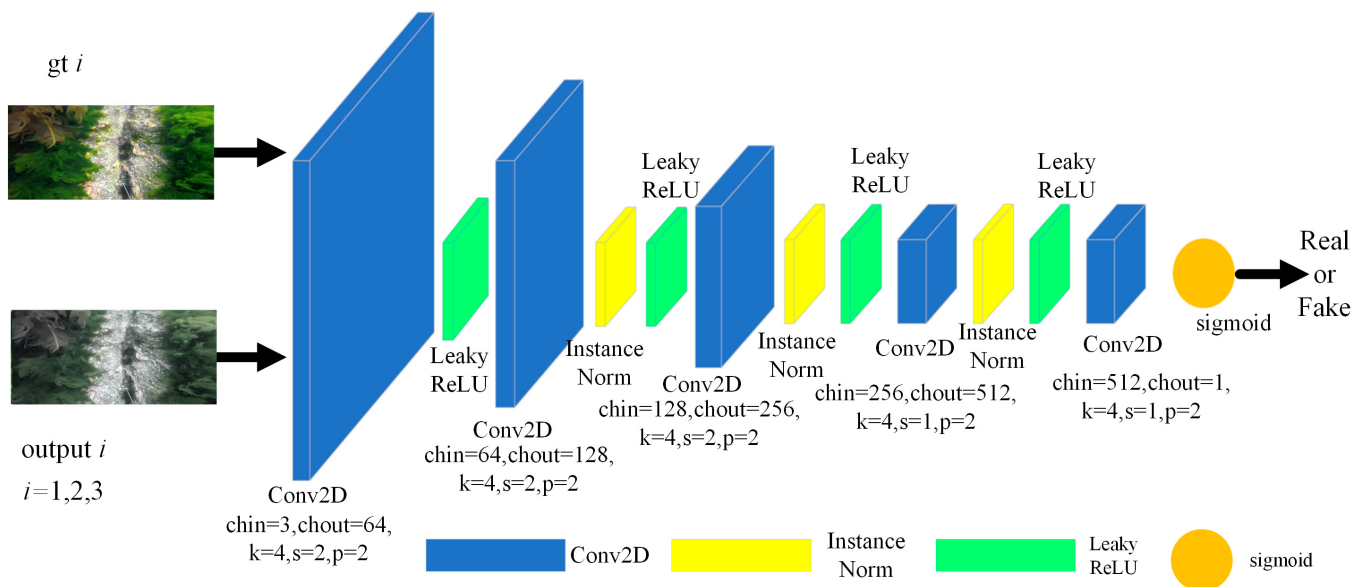


Figure 3. ECO-GAN discriminator.

2.3. Loss Function

2.3.1. Loss Function of the Generator

The loss function for the ECO-GAN generator is complex, encompassing multiple components. It includes blur removal loss, low-light enhancement loss, and color bias correction loss. From the neural network perspective, it comprises the generator loss L_G and the discriminator loss L_D . The generator loss further consists of pixel detail loss, abstract content loss, and adversarial loss.

$$L_G = L_{db-G} + L_{lt-G} + L_{wb-G} \quad (6)$$

$$L_D = E(D(img^g)) - E(D(img^{gt})) + gp \quad (7)$$

$$gp = \lambda E \left[\left(\left| \nabla_{img^{gt}} D(img^{gt}) \right|_2 - 1 \right)^2 \right] \quad (8)$$

where the corresponding discriminator of each branch adopts the same loss function, and L_D , img , img' , and img^{gt} represent the input original image, the generated image of the generator output, and the label truth-value image, respectively. A gradient penalty term proposed by Gulrajani et al., namely, Formula (3), is referred to here, and it is applicable to multiple generator structures and requires little adjustment of hyperparameters, i.e., $\lambda = 10$ in this experiment.

The loss function of the generator for the blur removal branch is as follows:

$$L_{db-G} = \sum_{pix}^{pixels} |img' - img^{gt}|^2 + L_{X-db} + Loss_{db}^{RaLSGAN}(img, img^{gt}) \quad (9)$$

The loss function of the generator for the low-light enhancement branch is as follows:

$$L_{lt-G} = \sum_{pix}^{pixels} |img' - img^{gt}|^2 + L_{X-lt} + Loss_{lt}^{RaLSGAN}(img, img^{gt}) \quad (10)$$

The loss function of the generator for the color bias correction branch is as follows:

$$L_{wb-G} = \sum_{pix}^{pixels} |img' - img^{gt}|^1 + L_{X-wb} + Loss_{wb}^{RaLSGAN}(img, img^{gt}) \quad (11)$$

$$L_X = \frac{1}{W_{i,j}H_{i,j}} \sum_{x=1}^{W_{i,j}} \sum_{y=1}^{H_{i,j}} (M_{i,j}(img') - M_{i,j}(img^{gt}))^2 \quad (12)$$

where $M_{i,j}$ represents the feature map between the convolutional layer i and convolutional layer j of the convolutional neural network VGG19, and $W_{i,j}H_{i,j}$ is the height and width between the convolutional layer i and convolutional layer j .

2.3.2. The Loss Function of the Discriminator

To make the discriminator training faster and the network more stable, we use the RaGAN-LS loss function proposed by Tetiana Martyniuk et al. in [28]:

$$Loss^{RaLSGAN}(x, z) = E_{x \sim P_{data}(x)} \left[\left(D(x) - E_{z \sim P_z(z)} [D(G(z))] - 1 \right)^2 \right] + E_{z \sim P_z(z)} \left[\left(D(G(z)) - E_{x \sim P_{data}(x)} [D(x)] + 1 \right)^2 \right] \quad (13)$$

where x represents the input data, which follows the distribution $P_{data}(x)$ and represents the data generated by the generator, $\tilde{x} = G(z)$; z represents a sample from the noise distribution, which follows the distribution $P_z(z)$; G and D represent the generator model and the discriminator model, respectively; and E indicates the expectation.

3. Experiment

3.1. Evaluation Indicator

In this paper, we utilized PSNR as an evaluation index to measure the enhancement effect of ECO-GAN. The width and height of a given image are h and w , respectively. The enhanced image is recorded as I_c , whereas the original noise image is recorded as I_n . The mean square error (MSE) of the enhanced and original images is defined as $MSE = \frac{1}{hw} \sum_{i=0}^{h-1} \sum_{j=0}^{w-1} [I_c(i, j) - I_n(i, j)]^2$, and the PSNR (dB) is defined as:

$$PSNR = 10 \cdot \log_{10} \left(\frac{MAX_I}{MSE} \right) \quad (14)$$

where MAX_I is the maximum pixel value of the image. If each pixel is represented by a B-bit binary number, then $MAX_I = 2^B - 1$. In this paper, if each pixel is represented by an 8-bit binary number, then the MAX_I is 255. In this paper, we accumulated PSNR of the three channels of RGB.

In addition, we used structural similarity SSIM to measure the luminance, contrast, and structure (structure) between samples x and y .

$$l(x, y) = \frac{2\mu_x\mu_y + c_1}{\mu_x^2 + \mu_y^2 + c_1} \quad (15)$$

$$c(x, y) = \frac{2\sigma_x\sigma_y + c_2}{\sigma_x^2 + \sigma_y^2 + c_2} \quad (16)$$

$$s(x, y) = \frac{\sigma_{xy} + c_3}{\sigma_x\sigma_y + c_3} \quad (17)$$

where μ_x and μ_y are the mean of x and y , respectively; σ_x^2 and σ_y^2 are the variance of x and y , respectively; σ_{xy} is the covariance of x and y ; and $c_1 = (k_1 MAX_I)^2$ and $c_2 = (k_2 MAX_I)^2$ are the two constants. Take $c_3 = c_2/2$ and avoid dividing by zero. MAX_I represents the maximum value of pixels in the B-bit image, which is 255 in this paper.

Take $k_1 = 0.01$, $k_2 = 0.03$ as the default value, then

$$SSIM(x, y) = [l(x, y)^\alpha \cdot c(x, y)^\beta \cdot s(x, y)^\gamma] \quad (18)$$

Set α , β , γ to 1, then

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \quad (19)$$

3.2. Experimental Results

The data used in this experiment were provided by the Underwater Robot Picking Competition (URPC) organized by the National Natural Science Foundation of China. The dataset used in this paper was URPC2019, which consists of images captured by an underwater robot with a camera and primarily exhibits quality issues such as blur, low light, and color distortion. We used Adobe's LightRoom Classic software (version 10.0) for batch processing of the images. The dataset included 5543 images, which were divided into a training set and a test set in a 7:3 proportion. The training set comprised 3880 deblurred truth-value images, 3880 enhanced truth-value images, and 3880 dedistorted truth-value images. The test set consisted of 1663 deblurred truth-value images, 1663 enhanced truth-value images, and 1663 dedistorted truth-value images. The proposed ECO-GAN and the comparison model were executed on a single NVIDIA GeForce RTX 2080Ti video card.

To validate the effectiveness of the luminance enhancement, blur removal, and color bias correction modules in the generator, separate experiments were conducted on each module, and the three modules were also tested in combination to assess their individual and combined efficacy. The experimental results are presented in Table 1, which demonstrates that the PSNR and SSIM indices for image enhancement using the proposed ECO-GAN method outperformed both individual and combined enhancement modules. In addition, we used UIQM and UCIQE, which are non-reference evaluation metrics. In Table 1, GT serves as the ground truth for calculating reference evaluation indicators. Light, Sharp, and WB represent modules for low-light enhancement, blur removal, and color bias correction, respectively. "Light + Sharp" represents a combination of a low-light enhancement module and a blur removal module. "Light + WB" represents a combination of a low-light enhancement module and a bias correction module. "Sharp + WB" represents a combination of a blur removal module and a bias correction module. "Light + Sharp +

WB” represents a combination of a low-light enhancement module, a blur removal module, and a color bias correction module. The combination method among three modules is in series. In order to verify the superiority of the proposed ECO-GAN, we compared ECO-GAN with the method proposed by Zhang et al. in [27] using a similar three-module serial structure. Due to the enhanced intercorrelation between multiple modules achieved through repeated feature extraction in the U-Net, the three enhancement modules collectively performed better than their independent counterparts, thereby facilitating flexible expansion of image enhancement capabilities. Moreover, Zhang et al. [24] highlighted the transferability aspect, which aids in accelerating model stability. We observed that ECO-GAN loaded with a pretrained model achieved faster stability compared with ECO-GAN without pretrained weights.

Table 1. Comparison of experimental results.

Method			PSNR	SSIM	UIQM	UCIQE
GT			/	1.0	0.2337	0.2524
Light			16.19	0.9075	0.1404	0.2166
	Sharp		11.48	0.7224	0.1467	0.2094
		WB	31.40	0.9613	0.1744	0.2191
Light	Sharp		11.95	0.7609	0.1523	0.2125
Light		WB	24.56	0.9587	0.1867	0.2214
	Sharp	WB	21.79	0.9487	0.1899	0.2170
Light	Sharp	WB	24.12	0.9521	0.1985	0.2263
Method proposed by Zhang et al. in [27]			17.29	0.7648	0.1556	0.1678
ECO-GAN			32.71	0.9783	0.1970	0.2016

The underwater image enhancement results obtained by different methods are presented in Figure 4. The raw image (Figure 4a) represents the original underwater image and serves as the input of the algorithm. GT (Figure 4b) serves as the ground truth for calculating reference evaluation indicators. Figure 4c–e show the enhanced results of GT after being processed by three modules, namely, “Light,” “Sharp,” and “WB,” respectively. Meanwhile, they were used to calculate the loss between corresponding image enhancement results. Figure 4f–k show the results of the single module and pairwise combinations of three modules on the raw image. Figure 4l–n show the results of the serial combination of three modules, the method proposed in [27], and the proposed ECO-GAN applied to the raw image, respectively. Due to the low quality of images in harsh underwater scenes, the non-reference values of both truth and result images are generally low. For reference indicators, the results demonstrate that ECO-GAN exhibited excellent performance in underwater image enhancement tasks.

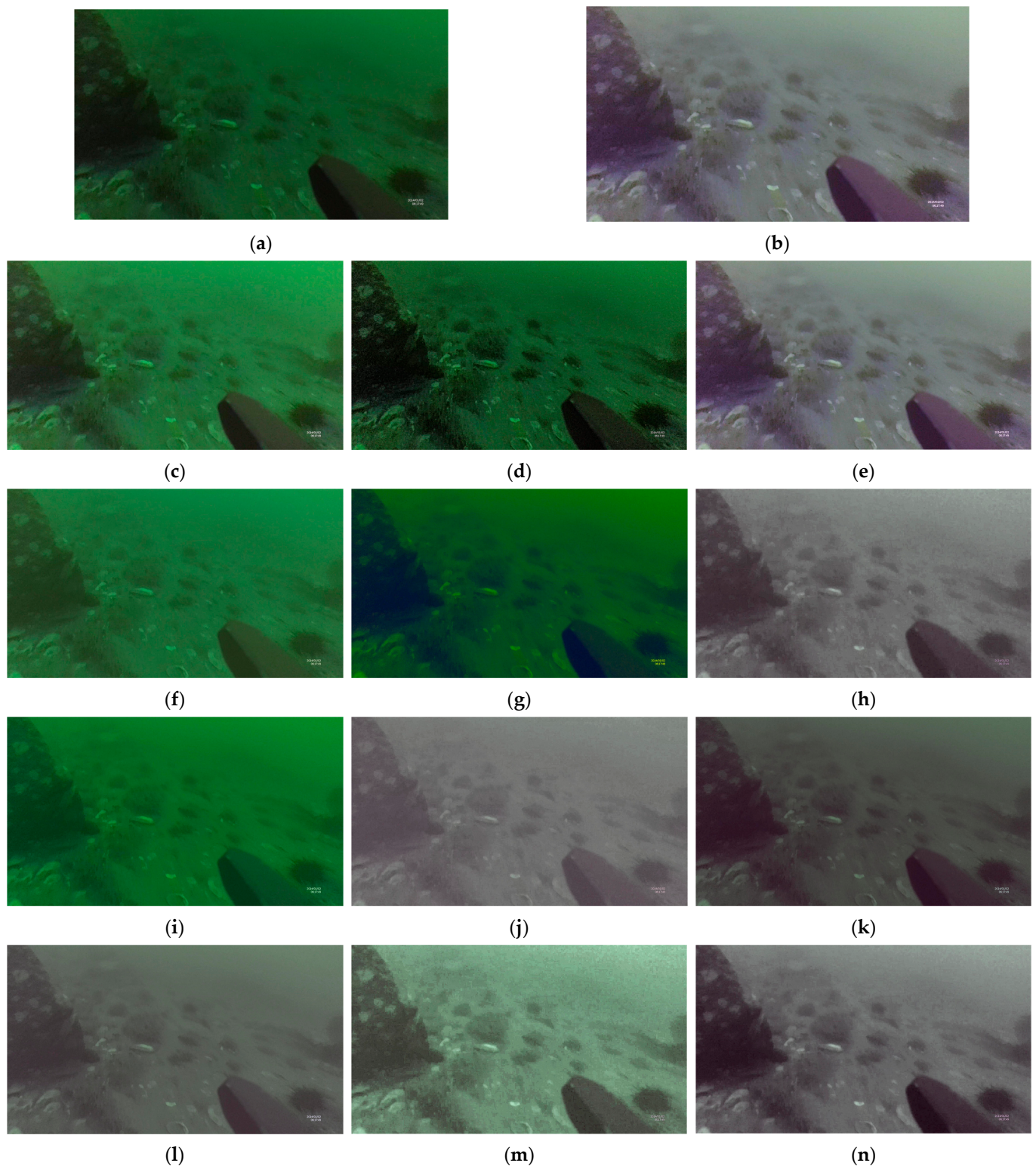


Figure 4. Underwater image enhancement results using different methods. (a) Raw; (b) GT; (c) Light on GT; (d) Sharp on GT; (e) WB on GT; (f) Light on raw; (g) Sharp on raw; (h) WB on raw; (i) “Light + Sharp” on raw; (j) “Light + WB” on raw; (k) “Sharp + WB” on raw; (l) “Light + Sharp + WB” on raw; (m) Method proposed in [27] on raw; (n) Proposed ECO-GAN on raw.

4. Discussion

In this paper, we propose an underwater image enhancement method, ECO-GAN, based on an image preprocessing framework for GANs. Our method addresses three key

challenges in underwater images: dynamic blur, low-light conditions, and color bias. The goal is to achieve blind enhancement of underwater images with multiple defects. To reduce computational effort and enhance inference speed, we employ the U-Net network structure and design decoders for different enhancement tasks by reusing the encoded features. By analyzing multi-scale features, we aim to alleviate image-quality issues. Additionally, a cross-stage fusion module is employed between the decoders for multiple enhancement tasks, strengthening their connection and enabling the recovery of details and improved image contrast. During the training process, the images are blindly enhanced using supervised training on paired datasets, eliminating the need for additional physical knowledge or prior information.

The proposed method significantly enhances the visualization of underwater images and facilitates the implementation of vision-based underwater tasks such as segmentation and tracking. In future work, we plan to explore the application of our method in related areas such as image defogging and super-resolution reconstruction to evaluate its generality. Additionally, we aim to apply the efficient model to unsupervised learning. These avenues remain as potential directions for future research.

Author Contributions: Conceptualization, H.Y. and M.P.; methodology, X.J. and M.P.; software, X.J., Z.L. and Y.Z.; validation, X.J., S.L. and Y.Z.; formal analysis, J.L., X.J. and S.L.; investigation, J.L. and X.J.; resources, M.P. and S.L.; data curation, Y.Z. and X.J.; writing—original draft preparation, X.J., Y.Z. and S.L.; writing—review and editing, S.L. and X.J.; visualization, X.J. and H.Y.; supervision, H.Y. and J.L.; project administration, H.Y. and J.L.; funding acquisition, M.P. and J.L. All authors have read and agreed to the published version of the manuscript.

Funding: This work was funded in part by the National Key Research and Development Project of China (grant No. 2022YFC2803600), the Key Research and Development Program of Zhejiang Province (grant Nos. 2022C03027, 2022C01144), the Public Welfare Technology Research Project of Zhejiang Province (grant Nos. LGF21E090004, LGF22E090006), the Hangzhou Science and Technology Development Plan Project (grant No. 202203B35), and the Zhejiang Provincial Key Lab of Equipment Electronics.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

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

References

1. Drews, P.; Nascimento, E.R.; Botelho, S.S.C.; Campos, M.F.M. Underwater Depth Estimation and Image Restoration Based on Single Images. *IEEE Comput. Graph. Appl.* **2016**, *36*, 24–35. [\[CrossRef\]](#) [\[PubMed\]](#)
2. Peng, Y.-T.; Cao, K.; Cosman, P.C. Generalization of the Dark Channel Prior for Single Image Restoration. *IEEE Trans. Image Process.* **2018**, *27*, 2856–2868. [\[CrossRef\]](#) [\[PubMed\]](#)
3. Carlevaris-Bianco, N.; Mohan, A.; Eustice, R.M. Initial results in underwater single image dehazing. In Proceedings of the Oceans 2010 Mts/IEEE Seattle, Seattle, DC, USA, 20–23 September 2010; pp. 1–8. [\[CrossRef\]](#)
4. Li, C.-Y.; Guo, J.-C.; Cong, R.-M.; Pang, Y.-W.; Wang, B. Underwater Image Enhancement by Dehazing With Minimum Information Loss and Histogram Distribution Prior. *IEEE Trans. Image Process.* **2016**, *25*, 5664–5677. [\[CrossRef\]](#) [\[PubMed\]](#)
5. Li, C.; Guo, J.; Chen, S.; Tang, Y.; Pang, Y.; Wang, J. Underwater image restoration based on minimum information loss principle and optical properties of underwater imaging. In Proceedings of the 2016 IEEE International Conference on Image Processing (ICIP), Phoenix, AZ, USA, 25–28 September 2016; pp. 1993–1997. [\[CrossRef\]](#)
6. Akkaynak, D.; Treibitz, T. A Revised Underwater Image Formation Model. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2018, Salt Lake City, UT, USA, 18–22 June 2018; pp. 6723–6732. [\[CrossRef\]](#)
7. 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. [\[CrossRef\]](#) [\[PubMed\]](#)
8. Fu, X.; Zhuang, P.; Huang, Y.; Liao, Y.; Zhang, X.-P.; Ding, X. A retinex-based enhancing approach for single underwater image. In Proceedings of the 2014 IEEE International Conference on Image Processing (ICIP), Paris, France, 27–30 October 2014; pp. 4572–4576. [\[CrossRef\]](#)

9. Fu, X.; Fan, Z.; Ling, M.; Huang, Y.; Ding, X. Two-step approach for single underwater image enhancement. In Proceedings of the 2017 International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS), Xiamen, China, 6–9 November 2017; pp. 789–794. [\[CrossRef\]](#)
10. Fabbri, C.; Islam, J.; Sattar, J. Enhancing Underwater Imagery Using Generative Adversarial Networks. *arXiv* **2018**, arXiv:1801.04011.
11. Yu, X.; Qu, Y.; Hong, M. Underwater-GAN: Underwater Image Restoration via Conditional Generative Adversarial Network. In Proceedings of the Pattern Recognition and Information Forensics: ICPR 2018 International Workshops, CVAUI, IWCF, and MIPPSNA, Beijing, China, 20–24 August 2018; Springer: Cham, Switzerland, 2018; pp. 66–75. [\[CrossRef\]](#)
12. 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. [\[CrossRef\]](#)
13. Lu, J.; Li, N.; Zhang, S.; Yu, Z.; Zheng, H.; Zheng, B. Multi-scale adversarial network for underwater image restoration. *Opt. Laser Technol.* **2018**, *110*, 105–113. [\[CrossRef\]](#)
14. Liu, X.; Gao, Z.; Chen, B.M. MLFcGAN: Multilevel Feature Fusion-Based Conditional GAN for Underwater Image Color Correction. *IEEE Geosci. Remote Sens. Lett.* **2019**, *17*, 1488–1492. [\[CrossRef\]](#)
15. 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.* **2017**, *3*, 387–394. [\[CrossRef\]](#)
16. Islam, J.; Xia, Y.; Sattar, J. Fast Underwater Image Enhancement for Improved Visual Perception. *IEEE Robot. Autom. Lett.* **2020**, *5*, 3227–3234. [\[CrossRef\]](#)
17. 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 on Computer Vision, Venice, Italy, 22–29 October 2017; pp. 2223–2232.
18. Anwar, S.; Li, C.; Porikli, F. Underwater scene prior inspired deep underwater image and video enhancement. *Pattern Recognit.* **2020**, *98*, 107038.
19. Sun, X.; Liu, L.; Li, Q.; Dong, J.; Lima, E.; Yin, R. Deep pixel-to-pixel network for underwater image enhancement and restoration. *IET Image Process.* **2019**, *13*, 469–474. [\[CrossRef\]](#)
20. Hong, L.; Wang, X.; Xiao, Z.; Zhang, G.; Liu, J. WSUIE: Weakly Supervised Underwater Image Enhancement for Improved Visual Perception. *IEEE Robot. Autom. Lett.* **2021**, *6*, 8237–8244. [\[CrossRef\]](#)
21. Wu, S.; Luo, T.; Jiang, G.; Yu, M.; Xu, H.; Zhu, Z.; Song, Y. A Two-Stage Underwater Enhancement Network Based on Structure Decomposition and Characteristics of Underwater Imaging. *IEEE J. Ocean. Eng.* **2021**, *46*, 1213–1227. [\[CrossRef\]](#)
22. Li, Y.; Chen, R. SE-RWNN: An synergistic evolution and randomly wired neural network-based model for adaptive underwater image enhancement. *IET Image Process.* **2020**, *14*, 4349–4358. [\[CrossRef\]](#)
23. Chang, X.; Ren, P.; Xu, P.; Li, Z.; Chen, X.; Hauptmann, A. A Comprehensive Survey of Scene Graphs: Generation and Application. *IEEE Trans. Pattern Anal. Mach. Intell.* **2021**, *45*, 1–26. [\[CrossRef\]](#) [\[PubMed\]](#)
24. Zhang, L.; Chang, X.; Liu, J.; Luo, M.; Li, Z.; Yao, L.; Hauptmann, A. TN-ZSTAD: Transferable Network for Zero-Shot Temporal Activity Detection. *IEEE Trans. Pattern Anal. Mach. Intell.* **2022**, *45*, 3848–3861. [\[CrossRef\]](#) [\[PubMed\]](#)
25. Yan, C.; Chang, X.; Li, Z.; Guan, W.; Ge, Z.; Zhu, L.; Zheng, Q. ZeroNAS: Differentiable Generative Adversarial Networks Search for Zero-Shot Learning. *IEEE Trans. Pattern Anal. Mach. Intell.* **2021**, *44*, 9733–9740. [\[CrossRef\]](#) [\[PubMed\]](#)
26. Li, M.; Huang, P.Y.; Chang, X.; Hu, J.; Yang, Y.; Hauptmann, A. Video Pivoting Unsupervised Multi-Modal Machine Translation. *IEEE Trans. Pattern Anal. Mach. Intell.* **2023**, *45*, 3918–3932. [\[CrossRef\]](#) [\[PubMed\]](#)
27. Zhang, X.; Fang, X.; Pan, M.; Yuan, L.; Zhang, Y.; Yuan, M.; Lv, S.; Yu, H. A Marine Organism Detection Framework Based on the Joint Optimization of Image Enhancement and Object Detection. *Sensors* **2021**, *21*, 7205. [\[CrossRef\]](#) [\[PubMed\]](#)
28. Kupyn, O.; Martyniuk, T.; Wu, J.; Wang, Z. DeblurGAN-v2: Deblurring (Orders-of-Magnitude) Faster and Better. In Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision (ICCV), Seoul, Republic of Korea, 27 October–2 November 2019; pp. 8877–8886.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.