Article

AUIE–GAN: Adaptive Underwater Image Enhancement Based on Generative Adversarial Networks

Fengxu Guan *, Siqi Lu, Haitao Lai and Xue Du

College of Intelligent Systems Science and Engineering, Harbin Engineering University, Harbin 150001, China; lusiqi9803@163.com (S.L.)
* Correspondence: guanfengxu@hrbeu.edu.cn

Abstract: Underwater optical imaging devices are often affected by the complex underwater environment and the characteristics of the water column, which leads to serious degradation and distortion of the images they capture. Deep learning-based underwater image enhancement (UIE) methods reduce the reliance on physical parameters in traditional methods and have powerful fitting capabilities, becoming a new baseline method for UIE tasks. However, the results of these methods often suffer from color distortion and lack of realism because they tend to have poor generalization and self-adaptation capabilities. Generating adversarial networks (GANs) provides a better fit and shows powerful capabilities on UIE tasks. Therefore, we designed a new network structure for the UIE task based on GANs. In this work, we changed the learning of the self-attention mechanism by introducing a trainable weight to balance the effect of the mechanism, improving the self-adaptive capability of the model. In addition, we designed a feature extractor based on residuals with multi-level residuals for better feature recovery. To further improve the performance of the generator, we proposed a dual path discriminator and a loss function with multiple weighted fusions to help model fitting in the frequency domain, improving image quality. We evaluated our method on the UIE task using challenging real underwater image datasets and a synthetic image dataset and compared it to state-of-the-art models. The method ensures increased enhancement quality, and the enhancement effect of the model for different styles of images is also relatively stable.

Keywords: underwater image enhancement (UIE); underwater image recovery (UIR); generative adversarial networks (GANs); deep learning; adaptive enhancement

1. Introduction

The rapid development of underwater intelligence bodies, such as autonomous underwater vehicles (AUVs) and unmanned underwater vehicles (UUVs), has made the human exploration of marine resources possible. The development and conservation of aquatic resources have become increasingly important globally and the focal point of international strategic development due to their abundance. Underwater vision is the most direct and effective way for an underwater intelligent body to perceive its environment, as well as the key to the human observation of the underwater environment and the state of the robot. However, due to the complex underwater environment, the images taken by the underwater optical imaging equipment often suffer from severe color distortion, underexposure, and blur, leading to serious degradation and poor quality [1,2].

Research on underwater image enhancement (UIE) methods is very significant. On the one hand, compared with raw images, underwater robots can obtain more meaningful feature information from enhanced images, which is the key to accomplishing underwater advanced vision tasks (e.g., target tracking) with high quality and performance. On the other hand, UIE tasks have excellent implications for tasks such as resource exploration, as enhanced images and videos can help human beings to discover more marine resources. In recent years, more and more scholars have become interested in UIE tasks, proposing

many methods to enhance degraded images. However, these approaches are constrained by different conditions and still face many challenges. Some traditional methods, such as image processing-based methods and model-based methods, are constrained by complex environments and physical parameters and often show poor enhancement ability. Although deep learning-based methods solve these problems to some extent; they are still accompanied by severe color distortion in their enhancement results, and they have weak generalization ability and unstable enhancement results for different styles of images. Notably, these problems are particularly evident in images with significant feature loss.

Generative adversarial networks (GANs) have more powerful fitting capabilities and have achieved significant results in many computer vision tasks [3–5], and many scholars have proposed some satisfactory UIE models based on GANs. However, the above-mentioned problems have not been effectively solved in most of the works. The work of Guo et al. [2] showed that the capability of the feature extraction module has a significant impact on the effectiveness of the UIE task. Therefore, we propose a more capable and efficient feature extraction module based on self-attention and multi-level residual. In addition, we find that the model trained using time-domain loss functions still has some bias in the frequency domain despite the enhancement effect, which leads to a decrease in the quality of the enhanced results. Based on this finding, we propose a dual path discriminator model and a frequency domain loss function. Overall, we propose an adaptive UIE method based on GANs, which we refer to as AUIE–GAN.

Overall, the principal contributions of this paper are as follows:

(1) An efficient and powerful feature extraction module and a new generator structure. We have changed the learning of the self-attention mechanism by introducing a trainable weight to balance the effect of the mechanism, improving the self-adaptive capability of the model. Based on this, we have designed a feature extractor based on residuals with multi-level residuals for better feature recovery. Based on the module in addition, we have improved the standard U–NET structure by introducing global feature vectors and self-attention skip connection, which could improve the model's ability to grasp global information and improve the global features of enhanced images (see Section 3.1).

(2) A new discriminator structure and a frequency loss function. We propose a loss function in the frequency domain and use a dual path discriminator to consider enhanced images in both time and frequency domains (see Section 3.2).

(3) State-of-the-art enhancement. We show that the performance of AUIE–GAN on different test datasets produces better processing results than SOTA models (see Sections 4.3 and 4.4). Figure 1 shows a preliminary comparison of the enhancement results in terms of subjective.
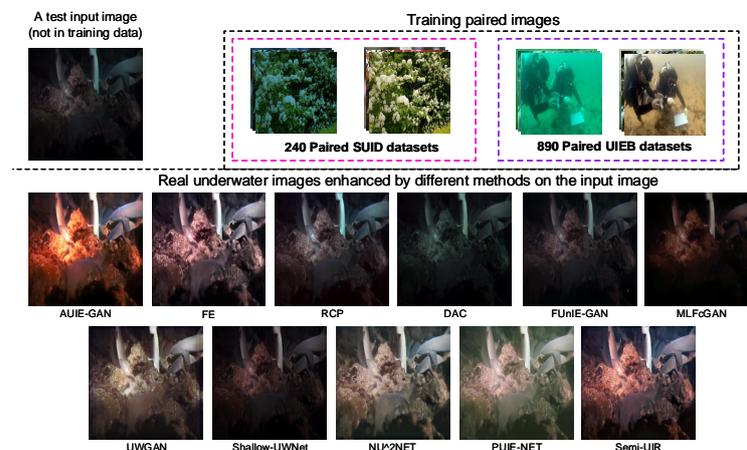

Figure 1. Enhancement comparison of different methods. In this paper, we trained AUIE–GAN on 1130 images and compared it with 10 alternative approaches.

2. Related Work

With the continuous development of underwater vision technology, UIE tasks have received widespread attention from related scholars in recent years. In general, the development of UIE tasks can be divided into the following parts.

2.1. Enhancement-Based Methods

These methods can obtain high-quality enhanced images by changing the pixel values in a single degraded underwater image [6]. After changing, the image has a satisfactory distribution. Some traditional works belong to this method including white balance (WB) [7], Retinex [8], and traditional contrast limited adaptive histogram equalization (CLAHE) [9]. For example, Li et al. [10] designed a segmented linear function for histogram transformation, adapting to the entire RGB values. The images processed by this method had more details on the RGB three channels. The CLAHE color model [11] applied a hybrid contrast adaptive histogram equalization method to the RGB and HSV models, effectively improving the underwater image quality, enhancing the contrast, and reducing noise. Tang et al. [12] proposed a new UIE algorithm based on adaptive feedback and the Retinex algorithm. The algorithm improved the color saturation, color richness, local contrast, and sharpness of underwater images through a series of operations such as adaptive feedback adjustment, guided filtering, and pixel fusion. Li et al. [13] proposed an underwater image dehazing algorithm based on the minimum information loss principle and a contrast enhancement algorithm based on a histogram distribution prior, which could be used to enhance degraded underwater images when we need to obtain more details from those. Zhang et al. [14] proposed an UIE algorithm based on extended multiscale Retinex. It could obtain enhanced images with better visual effects by combining bilateral and trilateral filtering of the three channels of the image in CIELAB color space according to the characteristics of each channel. FE [15] is a fusion-based strategy for enhancing underwater images and videos. It is a single−image method that effectively improves features such as contrast and enriches the detail information of an image just by deriving the inputs and weights from the degraded image.

Although these methods can enhance degraded images in most cases, they often under-enhance or over-enhance images with challenging scenes because they do not consider the degradation model.

2.2. Restoration-Based Methods

These methods estimate the parameters in the degradation model and obtain high-quality enhanced images by inverse-solving the degradation model. Red channel prior (RCP) [16], maximum intensity prior (MIP) [17], underwater light attenuation prior (ULAP) [18], Dark channel prior (DCP) [19], and underwater dark channel prior (UDCP) [20] belong to this method. Nicholas et al. [18] proposed a simple and effective a priori knowledge to estimate the scene's depth using the attenuation difference between the three image color channels in water. Yang et al. [21] proposed an efficient and low-complexity underwater image enhancement method based on dark channels prior. This method used median filtering to estimate the depth map of the image and a color correction method to enhance the contrast of the underwater image. Wang et al. [22] simplified the underwater light propagation model into a transmission model and proposed the maximum attenuation identification method based on it. They used this method to extract the depth map of underwater images for underwater image recovery. Paulo et al. [20] proposed a dark channel based on observing the R channel's absorption rate in many underwater images to recover high-quality images (UDCP). However, the variability of the underwater scene makes UDCP unstable, and the method requires many physical parameters and underwater optical properties, so it is usually difficult to implement.

However, using atmospheric degradation models to estimate underwater images can lead to deterioration in enhancement and poor quality. In addition, the complex

underwater environment and uncertainty of physical parameters also make the model modeling difficult.

2.3. Polarization Imaging-Based Methods

These methods are based on optical theory and they use the polarization properties of light to model and enhance underwater images. They are also a significant branch of the UIE task due to their low cost and high-quality enhancement [23]. Hu et al. [24] noticed the non-uniform light situation caused by active illumination. They regarded the backscattered light polarization as a spatial variable and derived the backscattered light polarization and intensity by a three-dimensional fitting method, thus obtaining a clear underwater image. In addition, to solve the degradation of the effect in the case of high turbidity, Hu et al. [25] utilized the combination of circular polarization and linear polarization and obtained an excellent enhanced image. Li et al. [26] combined polarization techniques with digital image processing techniques to preprocess polarized images through histogram equalization operations, and obtained enhanced images with rich details. Treibitz et al. [27] considered the polarization characteristics of the target signal together with the classical polarization equation, which can enhance underwater images efficiently. However, the method suffered from a serious noise amplification problem. Han et al. [28] further considered forward-scattered light and improved the enhancement quality by estimating the point diffusion function for forward scattering, inverting the imaging model.

Although polarization imaging-based methods are characterized by high quality and low cost, such methods suffer from high subjectivity, noise amplification, and poor real-time performance.

2.4. Deep Learning-Based Methods

In recent years, deep convolution-based methods have become the preferred base methods for UIE tasks because they can obtain high-quality enhanced images without complex modeling. In addition, these methods reduce the dependence on parameters in traditional methods and have higher generalization capability. Shallow–UWNet [29] can reduce the number of model parameters while maintaining enhancement quality to achieve lightweight enhancement. Guo et al. [30] proposed a normalization method for the tail of UIE methods, which can effectively enhance the enhancement of UIE methods. Fu et al. [31] proposed that UIE is a one-to-many task, where one input corresponds to multiple outputs. Therefore, they combined conditional variational autoencoder (cVAE) with adaptive instance normalization methods to propose PUIE–NET with a high diversity of enhanced results. However, these results are not bright enough in color and there is some fogging effect. UIEC^2–Net [32] efficiently and effectively integrates both RGB Color Space and HSV Color Space in one single CNN and obtains high-quality images. Successively captured underwater images are usually accompanied by the same type of degradation, and the images can integrate with each other. Qi et al. [33] proposed a cooperative processing and joint learning strategy in the same scene, which can process multiple images at once and obtain better enhancement results.

Generative adversarial networks (GANs) have been widely used in various tasks of computer vision with great success because of their own powerful fitting ability [3–5]. In recent years, many UIE works have used generative adversarial networks as their baseline model, obtaining higher-quality images. Inspired by the cycle consistency loss of Cycle–GAN, Li et al. [34] designed a multinomial loss function, and proposed a weakly supervised color transfer method to correct the color. UWGAN [2] proposed a multi-scale feature extraction module and obtained enhanced images based on GANs. However, the images obtained by UWGAN have color distortion and checkerboard artifacts, and the quality needs to be improved. FUnIE–GAN [35] is a real-time enhancement network, but the quality of the enhancement is lower due to the fewer parameters. The MLFcGAN [36] used global features to enhance local features at each scale to perform color correction and image detail preservation. Wu et al. [6] proposed a multi-scale fusion enhancement method based

on GANs to improve the enhancement and generalization ability of the model by encoding and combining three different priors.

We propose AUIE–GAN, a UIE method based on GANs. Compared with enhanced-based methods, AUIE–GAN has better enhancement quality and adaptivity, and its processing performance is robust to different scenes; unlike restoration-based methods and polarization imaging-based methods, AUIE–GAN does not require complex modeling steps and has faster processing speed. In conclusion, the proposed method achieves better enhanced qualitative and quantitative results than existing methods.

3. Methods

3.1. Generator Architecture

Ref. [2] stated that the feature extraction ability of the generator directly affects the number and quality of features that can be learned by the model from degraded images, which directly affects the enhancement quality. In our AUIE–GAN, we designed an Encoder–Decoder structure generator based on the U–Net [37]. The inputs and outputs of U–Net have the same size, and it is the most commonly used network structure for end-to-end models and is widely used for various types of vision tasks [35,37]. The difference is we propose to use an efficient and powerful feature extraction module in the backbone of the U–Net. Moreover, we combine self-attention skip connecting with global feature embedding in the up-sampling stage to improve the global ability of the model.

Self-attention mechanism has been successfully applied to computer vision tasks because it can focus on the correlation between pixel points at different locations in an image. Although the degradation of the underwater images is uniform, with the entire image exhibiting almost the same degradation style, the degree of degradation of different images is again diverse and uncontrollable, which leads to the standard self-attention mechanism not being able to learn the complex degradation style and reduces the self-adaptive capability of the model. Through numerous experiments, we found that this problem can be solved using a simple method. In this paper, we propose to use a trainable weight to balance the output of the self-attention block, shown in the bottom right corner of Figure 2.
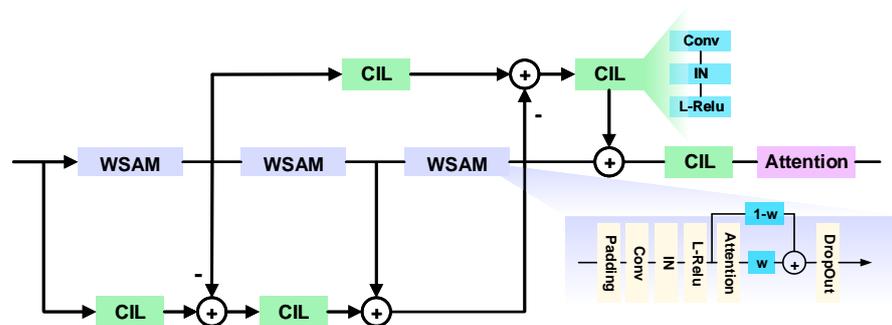

Figure 2. Self-Attention Residual Module. WSAM is the weight self-attention module, and the topology is shown in the bottom right corner. CIL blocks are used to change the number of channels of the image to facilitate feature computation. In the CIL block, we use a 3×3 convolution, an instance normalization, and a Leaky-Relu activation function.

Several end-to-end UIE models have used complex multi-scale convolutional modules to improve the feature extraction capability of the model [2]. The U–Net-based end-to-end UIE method uses an encoder to convert the input image into coded information in the latent space and achieves the enhancement of the image by processing the latent vectors accordingly. These complex multi-scale convolutional modules indeed improve the handling of hidden vectors and enhance the quality of enhancement. Based on this, in this paper, we use multi-level residuals, which allows the model to learn more useful information and improve enhancement by simply computing different fusion features at different stages; we refer to this module as Self-Attention Residual Module (SARM).

Based on SARM, we propose a new generator topology, shown in Figure 3. The structure is based on an improved U–Net as the main structure. In the structure, we have consider color priori information using global feature vectors, which are considered by [38] to eliminate artifacts that appear in enhanced images in image enhancement tasks. Because there are often uniform color projections in underwater images, inspired by [38], we have embedded these global vectors with color projection information at the start of the up-sampling operation. Moreover, we have added the self-attention mechanism in the middle of the skip connection, which can help the network recover the features better. In details, we have replaced all Batch Normalization (BN) layers with Instance Normalization (IN) layers because BN layers show a significant weakening effect when trained in a small batch.
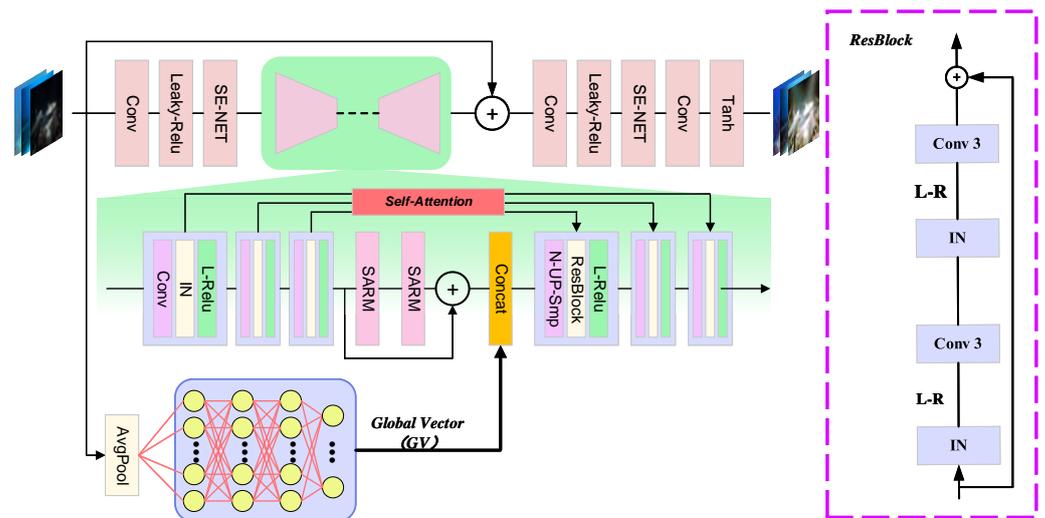

Figure 3. Generator structure in the AUIE–GAN. In this paper, both input and output image sizes are 256^2. The right side of the figure shows the residual block structure used in this paper, which is used three times in our AUIE–GAN. In the figure, IN is an instance normalization operation, L–Relu is a Leaky–Relu function, and N–UP–Smp is the nearest up-sample operation.

3.2. Discriminator Architecture

Ref. [39] proposed that the Fourier transform changes the spatial coherence of the image, and it is not possible to use convolutional neural networks to extract the features of the spectrogram. Therefore, in this paper, we propose a dual path discriminator to consider enhanced images in both time and frequency domains. The discriminator can be divided into a time domain path and a frequency domain path, as shown in Figure 4. It is worth noting that the time domain here is relative to the frequency domain and is not the real existing time axis. On the time domain path, the input is a combination of a degraded image and a reference or enhanced image. We used a Patch–GAN structure to extract feature information from the input images. In addition, we used a spectral normalization operation to stabilize the training of GANs, which is shown to have significant effects on the generation task [40]. We input spectrograms corresponding to the inputs on the time domain path to the frequency domain path, which was composed of fully connected layers. At the end of the frequency domain path, we used a single neuron to output the probability. Finally, this probability was mapped to the same scale as the probability matrix of the time domain path by replication and weighted superposition to obtain the final probability matrix. It is well known that the discriminator is used to determine the source of the input image and to guide the generator in updating the weights during the backpropagation process. The output of the dual-path discriminator considers the authenticity of the input image from both time and frequency domains, which can further improve its guiding ability. In addition, different from the standard GAN's discriminator, we used tanh as the activation function of the output layer to accommodate the hinge loss. The structure of the

discriminator of the AUIE–GAN is shown in Figure 4. Therefore, the output probability matrix is:

$$D(\cdot) = \alpha_t D_t + (1 - \alpha_t) D_f \tag{1}$$

where α_t is the weight of the two path outputs of the dual-path discriminator, D_t is the probability matrix of the time-domain path outputs, and D_f is the probability matrix of the frequency-domain path outputs.
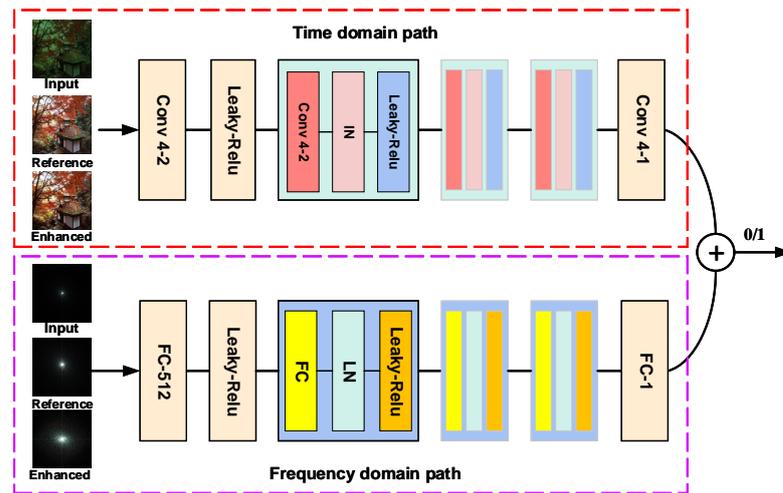

Figure 4. Discriminator structure of the AUIE–GAN. In the time domain path, we set the stride of the convolution layer to 2 to complete the down-sampling instead of the pooling layer. In the frequency domain path, all fully connected layers except the last one have 512 neurons. The LN in the figure is the layer normalization operation.

3.3. Loss Functions

We found that the performance of GANs-based models trained with single or fewer losses tended to be limited, and these models did not better represent their advantages in UIE tasks. Although several image translation works have proposed many effective loss functions in recent years, there are no UIE works that integrate the rich loss functions and train their GANs-based models. We found that the weighted fusion of loss functions with different capabilities can lead to further enhancement of the models by choosing reasonable weights. In this paper, we trained our AUIE–GAN using a loss function with multiple weighted fusions.

In this section, the following definitions are available: x is the ground-truth image (GT), y is the degenerate image, and z is the enhanced image output by the generator of the AUIE–GAN.

We used a hinge loss of GANs to improve the stability of the training, which is mathematically expressed as:

$$L_H(G, D) = L_D + L_G \tag{2}$$

$$L_D = E[\max(0, 1 - D(x))] + E[\max(0, 1 + D(G(y)))] \tag{3}$$

$$L_G = -E[D(G(y)] \tag{4}$$

where $D(x)$ is the probability that the image input to the discriminator is the GT image, and $D(G(y))$ is the probability that the image input to the discriminator is the enhanced image.

We trained L1 and L2 losses alternately to avoid fitting problems when training with a single loss, and we referred to this alternate training as the L loss. The loss can be expressed as:

$$
loss_L = \begin{cases} \frac{1}{N} \sum_{i=1}^{N} |x_i - z_i| & epochs\%200 < 100 \\ \frac{1}{N} \sum_{i=1}^{N} (x_i - z_i)^2 & epochs\%200 > 100 \end{cases} \tag{5}
$$

where, in this paper, we used an alternation frequency of 100 epochs.

SSIM loss is a common objective function in image translation tasks, which can effectively enhance the quality of enhanced images. It can be expressed as:

$$
loss_S = 1 - SSIM(x,z) = 1 - \frac{(2\mu_x \mu_z + C_1)(2\sigma_{xz} + C_2)}{(\mu_x^2 + \mu_z^2 + C_1)(\sigma_x^2 + \sigma_z^2 + C_2)} \tag{6}
$$

where μ_x and μ_z denote the mean value of each image pixel, σ_x and σ_z are the unbiased estimate of the standard deviation of the image, σ_{xz} is the correlation coefficient of the two images, and C is constant.

Enhanced images often suffer from feature-level color loss because L loss neglects phase angle information between pixels [41]. Ref. [42] proposed image processing using Gaussian blurring, preserving only color information and fitting at the color level. This operation can be expressed as:

$$
loss_C = \|Z_b - X_b\|_2^2 \tag{7}
$$

where Z_b and X_b are the blurred images of z and x.

The perceptual loss [43] can be fitted to the features of the image in the latent space, which improves the quality of the image. It can be expressed as:

$$
loss_V = \frac{1}{N} \sum_{i=1}^{N} (VGG(x_i) - VGG(z_i))^2 \tag{8}
$$

where $VGG(\cdot)$ is the output of the Conv4–3 layer in the pre-trained VGG16 network.

Actually, the difference between the enhanced image and the ground-truth image is also evident in the spectrum, and this bias causes different degrees of loss of certain feature information (e.g., texture, color, etc.) in the enhanced image. Unfortunately, the loss function in the time domain does not solve this problem well [2,35,36,39]. Therefore, we propose to fit distributions on the frequency domain.

In this section, we assumed that $F(a,b)$ is the frequency domain information after the FFT transform of the image. The information can be expressed as Equation (9).

$$
F_{qm}(a,b) = a_{qm} + b_{qm}i \tag{9}
$$

where qm denotes the image's pixel with coordinates (q,m). In this paper, we use $F(a,b)$ to denote the enhanced image and $F(a',b')$ to denote the ground-truth image.

In contrast to existing work, we discussed amplitude and phase separately to ensure reasonableness. The loss of amplitude and phase in this paper can be expressed as:

$$
|F(a,b) - F(a',b')|^2 = (a - a')^2 + (b - b')^2 \tag{10}
$$

$$
|\angle F(a,b) - \angle F(a',b')| = \left| \arctan \frac{b}{a} - \arctan \frac{b'}{a'} \right| \tag{11}
$$

Therefore, we can obtain the loss function of amplitude and phase:

$$loss_{fa} = \frac{1}{HWC} \sum_{q=0}^{H} \sum_{m=0}^{W} \sqrt{\left[(a_{qm} - a'_{qm})^2 + (b_{qm} - b'_{qm})^2\right]^2 + \varepsilon^2} \tag{12}$$

$$loss_{f\angle} = \frac{1}{HWC} \sum_{q=0}^{H} \sum_{m=0}^{W} \sqrt{\left[a'_{qm}b_{qm} - a_{qm}b'_{qm}\right]^2 + \varepsilon^2} \tag{13}$$

where ε takes 1×10^{-3}, W, H and C denote the width, height, and channel of the image. Finally, the frequency domain loss can be expressed as:

$$loss_F = 0.5 \times loss_{fa} + 0.5 \times loss_{f\angle} \tag{14}$$

Considering all the above loss functions comprehensively, the weights of each loss were determined by repeated experiments and combined with experience. The total loss function is:

$$Loss = L_H(G, D) + \lambda_L \times L_L(G) + \lambda_S \times L_S(G) + \lambda_V \times L_V(G) + \lambda_C \times L_C(G) + \lambda_F \times L_F(G) \tag{15}$$

We set $\lambda_L = 2$, $\lambda_S = 15$, $\lambda_V = 4$, $\lambda_C = 20$, and $\lambda_F = 10$.

4. Experiment

4.1. Datasets and Metrics

In this work, we trained our AUIE–GAN using paired UIEB datasets and part of SUID dataset. UIEB is an underwater image benchmark dataset open-sourced by Li et al. [44] in 2020. It consists of 890 pairs of real underwater images, with the reference images selected from the best of various enhancement methods. The SUID datasets [45] acquire 30 ground truth images and algorithmically simulated degraded images with different degradation levels, including fogging and blue-green tint. To demonstrate the performance of the AUIE–GAN, on the one hand, we tested our AUIE–GAN on the real image datasets which are Challenge60, U45, and RUIE datasets. On the other hand, we tested the model on the SUID dataset because it is a synthetic dataset that can test the color realism of the enhanced images. The Challenge60 and U45 datasets are common test sets of real underwater images used in UIE tasks, and they include a variety of styles of degraded images. The RUIE dataset [46] is the first large database of realistic underwater images specifically set up for multi-angle algorithm evaluation, proposed by Liu et al. in 2020. The RUIE dataset was sampled at Roe Deer Island near the Yellow Sea, with over 250 h of video taken under natural marine ecosystems, covering a wide range of illumination, depth of field, and color shift.

We quantified the effect of AUIE–GAN testing on real underwater image datasets using the UIQM and UCIQE metrics, and the test results on synthetic datasets using the SSIM and PSNR metrics. PSNR was used for image evaluation by mean square error (MSE). It estimates the similarity between the target image and the enhanced image by comparing their MSE values and is a standard evaluation metric in image enhancement. SSIM more closely resembles human visual perception and evaluates the similarity of two images by focusing mainly on edge and texture similarity. UCIQE evaluates enhanced images using a linear combination of chromaticity, saturation, and contrast to quantify uneven color cast, blur, and low contrast. UIQM is obtained by weighted summation of underwater image color metric (UICM), underwater image sharpness metric (UISM), and underwater image contrast metric (UIConM).

4.2. Implementations and Baselines

To train AUIE–GAN, we used 890 pairs of images in UIEB and 240 pairs of images in SUID as the training set, which contains a total of 1130 pairs of images. In addition, we trained AUIE–GAN for 400 epochs using the Adam optimizer with default parameter

values ($\beta1 = 0.9$, $\beta2 = 0.999$). In addition to the loss function weights set in Section 3.3, we set the learning rate = 1×10^{-4} and set the output weights of the dual path discriminator $\alpha_t = 0.8$. We converted input images to 256^2 for model training. All experiments were implemented using Tensorflow on an NVIDIA GeForce RTX 3090 GPU.

To demonstrate that AUIE–GAN promotes the development of the UIE task, we compared it with the mainstream UIE methods in recent years, which include FE (2012) [15], RCP (2015) [16], DAC (2022) [47] among traditional methods, FUnIE–GAN (2020) [35], MLFcGAN (2020) [36], UWGAN (2020) [2], Shallow–UWNet (2021) [29], NU^2NET (2022/SOTA) [30], PUIE–NET(2022/SOTA) [31] and Semi–UIR (2023/SOTA) [48] among deep learning-based methods.

4.3. Results on Full-Reference Datasets

We tested the AUIE–GAN on the SUID dataset to demonstrate the fitting ability of the model. The qualitative results are shown in Figure 5.

Figure 5. Comparison of the enhancement effects on synthetic images. The four images on the left that possessed different degrees and different styles of degradation were selected from the SUID dataset.

In Figure 5, compared to the degraded image, different methods showed diverse fitting effects. Compared with other methods, FE, UWGAN, and our AUIE–GAN had significantly excellent visualization results. Furthermore, we plotted the frequency distribution curves of the degraded image, the enhanced image and the reference image, which can visualize our enhancement results and further illustrate the contribution of AUIE–GAN to the UIE task, as in Figure 6.

In Figure 6, the graphs on the left are frequency distribution of pixels on RGB space for each degraded image. The faulty color projection in the image is caused by the high frequency component of the corresponding channel. The graphs on the right are frequency distribution of the corresponding enhanced image and the reference image. The enhanced image has the same curve shape as the reference image, with the difference that the position of the curve is shifted, which is expressed in the image as a difference in brightness and contrast.

We sampled 71 images from the SUID dataset that had different degradation styles for testing. The PSNR and SSIM metrics were used to quantify the test results of different methods on the SUID dataset, as in Table 1.

On the SUID dataset, AUIE–GAN achieved the best PSNR and SSIM metrics, with a 1.52% improvement in PSNR metrics and a 0.02% improvement in SSIM metrics compared to the second-best metrics. Referring to the qualitative experimental results, the results of AUIE–GAN are more similar to the reference ground-truth image. In general, AUIE–GAN achieved excellent enhancement results on the synthesized dataset, and the enhanced images not only removed the color projection but also had clearer features and more realistic colorimetry in the images.
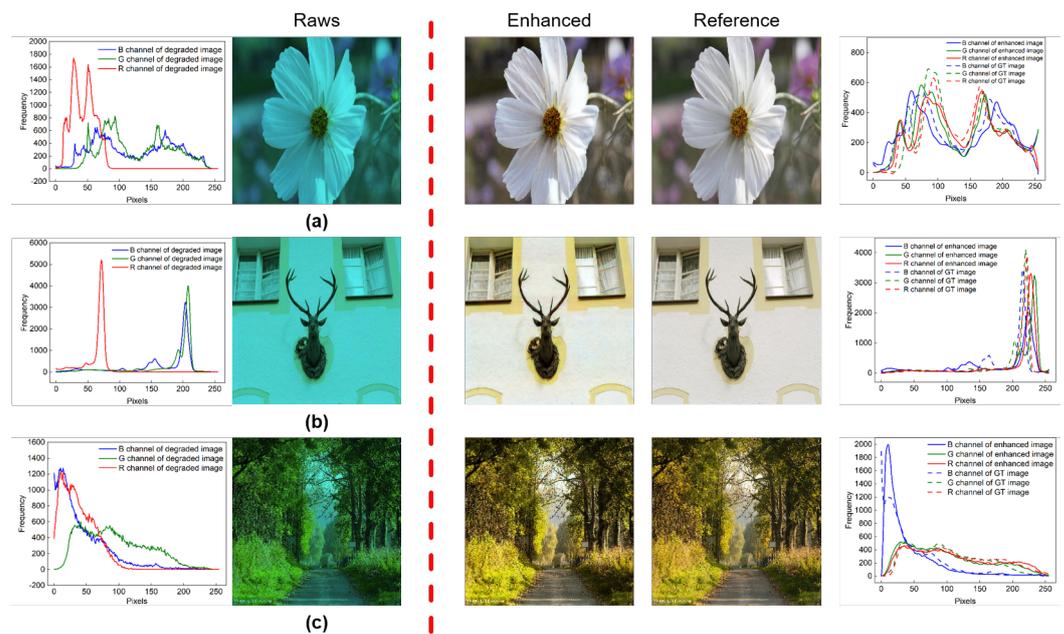
Figure 6. Comparison of the frequency distribution curves.

Table 1. Quantitative comparison results of different methods.

Baselines	PSNR	SSIM
Raw	12.2289	0.7598
FE (2012)	20.3589	0.8007
RCP (2015)	20.2441	0.8780
DAC (2022)	13.8407	0.7531
FUnIE–GAN (2020)	16.2501	0.8058
MLFcGAN (2020)	13.7275	0.5767
UWGAN (2020)	28.7628	0.9138
Shallow–UWNet (2021)	14.4293	0.6985
NUˆ2NET (2022)	23.2156	0.9300
PUIE–NET (2022)	19.6275	0.9028
Semi–UIR (2023)	22.3050	0.9192
AUIE–GAN (ours)	29.1992	0.9302

Red indicates the best metrics in each column, and blue indicates the second–best metrics in each column.

4.4. Results on Non-Reference Datasets

In this section, we tested AUIE–GAN on real underwater image datasets, which are Challenge60, U45, and RUIE. In fact, the enhancement results on real underwater images are more representative of the quality of the model because the realism of color bias present in these images cannot be synthesised by algorithm. Therefore, we sampled images with different degraded tones from above datasets for demonstration. Figures 7–9 show the qualitative results on different datasets.

In these comparison methods, various methods had different enhancement effects on each style of image. The FE method had high robustness, but it showed significant red distortion when processing blue images. RCP only faded the degradation when processing these images and did not do an excellent job (especially the blue images and green images). The result of DAC achieved poor visual results because it only stretched in RGB space and did not consider real underwater degradation models. FUnIE–GAN and Shallow–UWNet improved the real-time performance by reducing the parameters of the model, so they showed weaker performance in processing images with higher degradation. MLFcGAN was designed to solve the color bias problem, but it still suffers from poor results when processing these highly degraded images. UWGAN uses a complex feature extraction

module to improve the enhancement quality, but some artifacts can appear in the enhanced images. Moreover, the enhancement stability of the model was different when processing different styles of images. Although NU^2NET and PUIE–NET achieved more pleasant results, they also suffered from some problems, such as the low brightness and contrast of the results of NU^2NET and the fogging effect of the results of PUIE–NET. Semi–UIR was the latest deep learning-based UIE method, and it achieved more pleasant enhancement results in our experiment. However, some enhanced images would appear blurred (the bottom image in Figure 7). AUIE–GAN had better visual results for all styles of images, they did not blindly increase brightness, contrast, and color, but targeted enhancement of different features, which cannot be achieved without the role of our WSAM module (Section 4.5 proved this point). However, qualitative analysis is highly subjective due to the different preferences of users. Therefore, the performance evaluation of UIE methods requires a combination of qualitative and quantitative results [48].
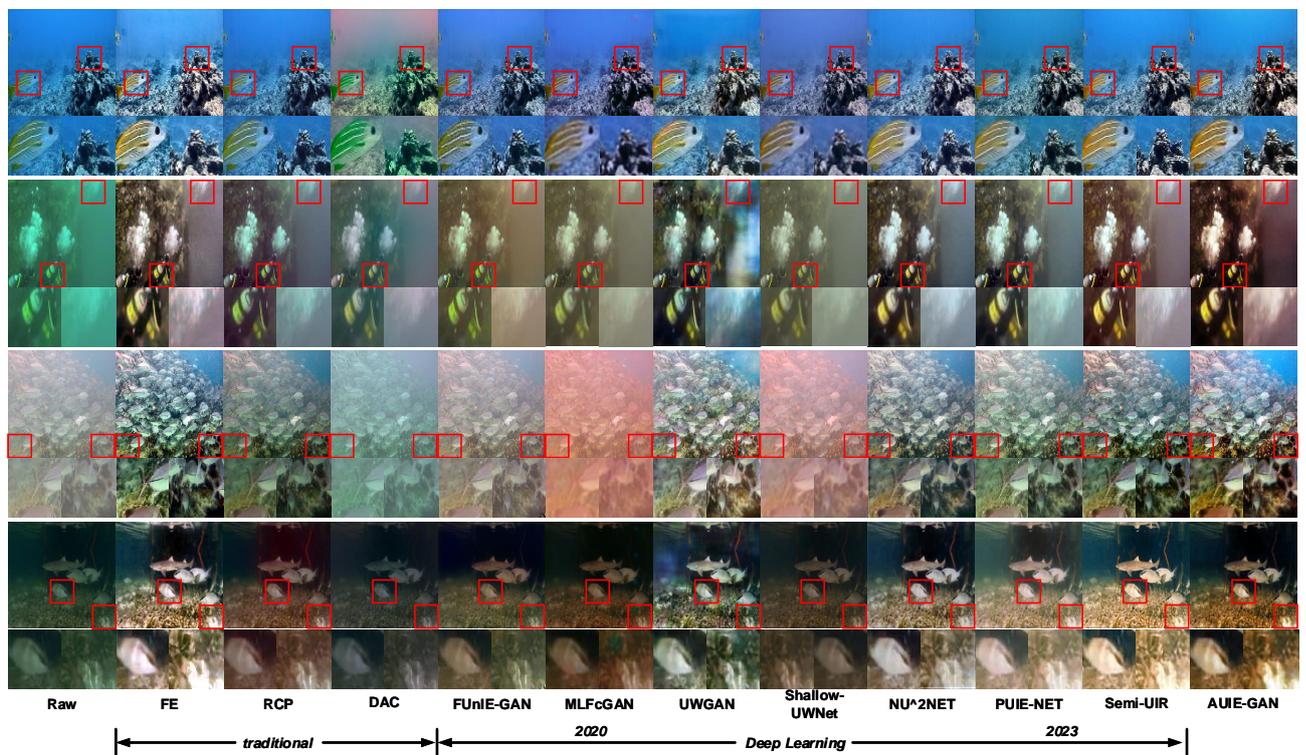

Figure 7. Comparison of results on the Challenge60 dataset. These images had different styles: blue, green, haze, and low light. The boxes are examples, zoom in on the image to see better results.

We quantified the enhancement results using UIQM and UCIQE and presented them in Table 2. Our AUIE–GAN achieves the best test results in most cases on three datasets. Specifically, for the UIQM metric, AUIE–GAN improves over the second-best result by 4.99%, 3.86%, and 0.11% on the three datasets, respectively. For the UCIQE metric, our AUIE–GAN achieved the best results on the Challenge60 and U45 datasets, with 0.17% and 2.70% improvement compared to the second-best results, respectively.

Considering the qualitative and quantitative experimental results together, we believe that AUIE–GAN achieved pleasant enhancement results and obtained better processing results compared to SOTA models.

Complex underwater environments cause uncontrollable degradation of the images acquired by underwater imaging equipment. Present UIE methods are weakly adaptive, making them less effective in enhancing some specific image styles (e.g., the red noise that appears in the FE method). To further illustrate the contribution of AUIE–GAN, we divided U45 and Challenge60 into four categories according to degradation styles, which are blue, green, haze, and low-light. Based on the classification, we performed independent quantitative calculations for each class of images to evaluate the enhancement of different models for different styles of images. The quantitative results are shown in Table 3 and Figure 10.
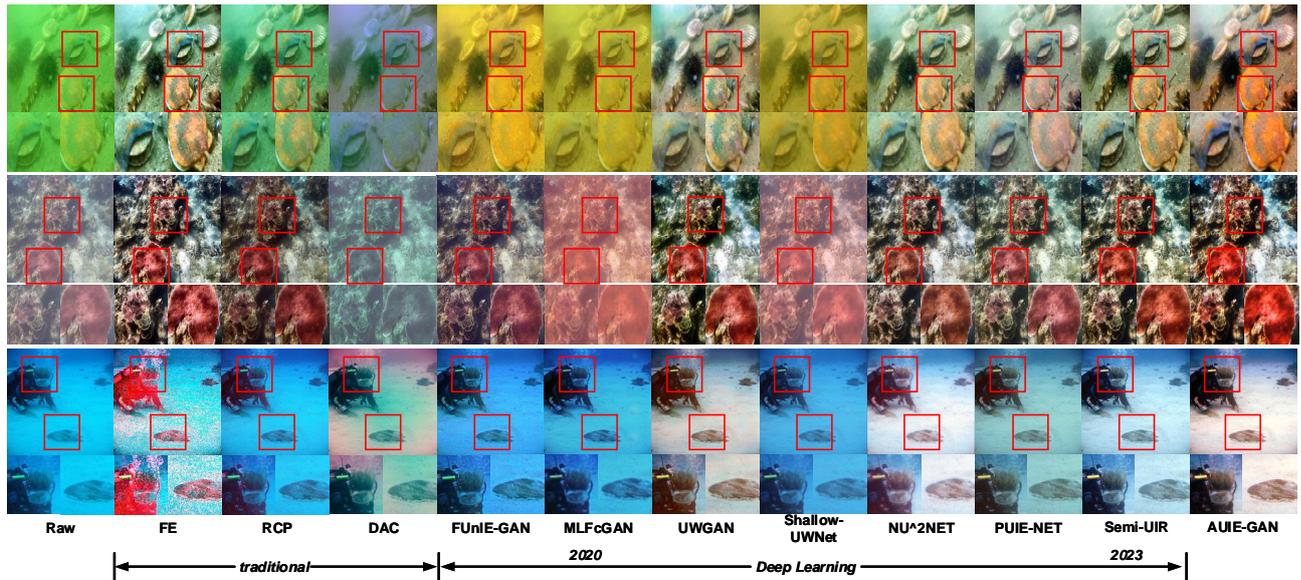

Figure 8. Comparison of results on the U45 dataset. These images had different styles: green, haze, and blue. Different from the images sampled in the Challenge60 dataset, these images suffered from more severe degradation. The boxes are examples, zoom in on the image to see better results.
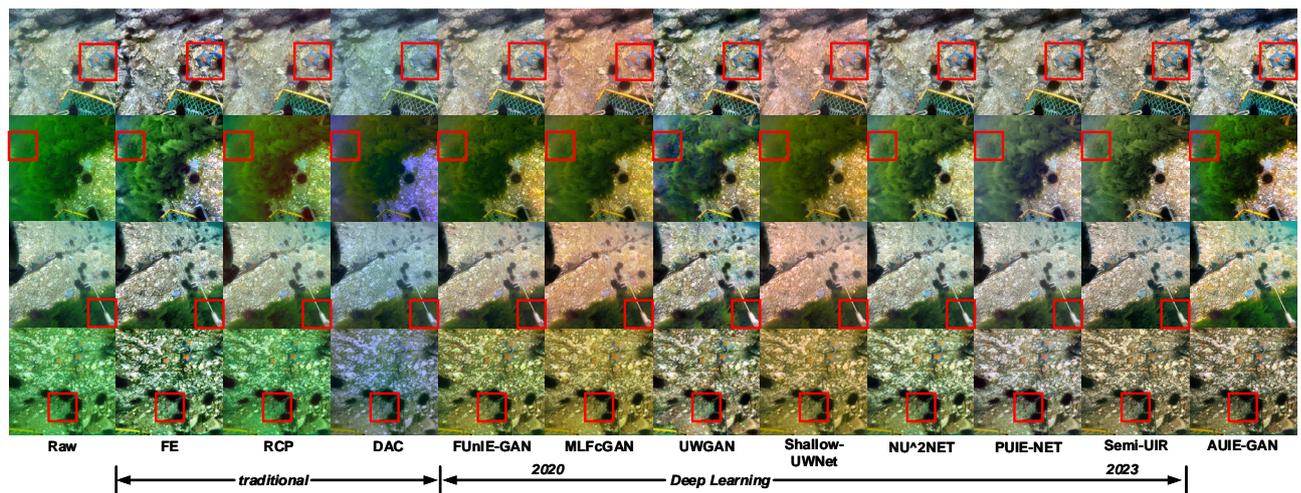

Figure 9. Comparison of results on the RUIE dataset. More detailed information on the comparison graph can be viewed by enlarging the image.

From Table 3 and Figure 10, AUIE–GAN achieved better results in most of the experiments, and the metric curves were smoother compared with other methods. The experimental results indicate that AUIE–GAN can perform stable enhancement for underwater images with different styles.

Table 2. Comparison of objective evaluation of different methods.

Baselines	UIQM (↑)			UCIQE (↑)		
	U45	Challenge60	RUIE	U45	Challenge60	RUIE
RAW	2.3821	2.9388	3.6634	0.3754	0.3590	0.3571
FE	4.1226	4.3528	4.8685	0.4791	0.4782	0.4464
RCP	3.5829	4.1845	4.7307	0.4365	0.4243	0.4047
DAC	3.6855	3.7572	4.8680	0.3274	0.3049	0.3354
FUnIE–GAN	4.3503	4.2946	5.3997	0.4000	0.3592	0.3633
MLFcGAN	4.2029	4.2923	5.3902	0.3909	0.3681	0.3740
UWGAN	3.9840	4.2963	5.4074	0.4836	0.5043	0.3887
Shallow–UWNet	4.0523	4.1125	5.3939	0.3704	0.3348	0.3719
NU^2NET	4.3921	4.4623	5.3521	0.4186	0.4125	0.4088
PUIE–NET	4.2863	4.0633	5.1813	0.3874	0.3772	0.3820
Semi–UIR	4.5154	4.3051	5.0206	0.4283	0.4357	0.4041
AUIE–GAN	4.7409	4.6346	5.4136	0.4844	0.5179	0.4326

Red indicates the best metrics in each column, and blue indicates the second-best metrics in each column.

Table 3. Comparison of enhancement results by styles.

Baselines	Blue		Green		Hazy		Low-Light	
	UIQM	UCIQE	UIQM	UCIQE	UIQM	UCIQE	UIQM	UCIQE
RAW	2.1982	0.4229	1.8274	0.3857	3.7035	0.2788	4.2557	0.3322
FE	3.7578	0.4943	4.1200	0.4702	4.7638	0.4621	5.0450	0.4719
RCP	3.4218	0.4579	3.3204	0.4190	4.7891	0.4134	4.9475	0.3943
DAC	3.6714	0.3600	3.6919	0.3414	3.4419	0.2312	4.5746	0.2912
FUnIE–GAN	3.6850	0.3992	4.5069	0.3848	4.8411	0.3452	4.8492	0.3449
MLFcGAN	3.6804	0.3935	3.9696	0.3408	5.1186	0.3865	4.8290	0.3884
UWGAN	3.4863	0.4247	4.1250	0.3808	4.6873	0.4020	5.4309	0.4413
Shallow–UWNet	3.5399	0.3616	4.0139	0.3249	4.6907	0.3536	4.7070	0.3445
NU^2NET	4.0057	0.4295	4.4596	0.4052	4.7934	0.4124	4.9543	0.4028
PUIE–NET	3.7370	0.3888	4.2769	0.3818	4.4543	0.3783	4.6661	0.3713
Semi–UIR	3.8503	0.4312	4.6147	0.4113	4.6627	0.4422	5.1683	0.4611
AUIE–GAN	4.1562	0.4904	4.6068	0.4699	4.7696	0.4963	5.9332	0.5360

Red indicates the best metrics in each column, and blue indicates the second-best metrics in each column.
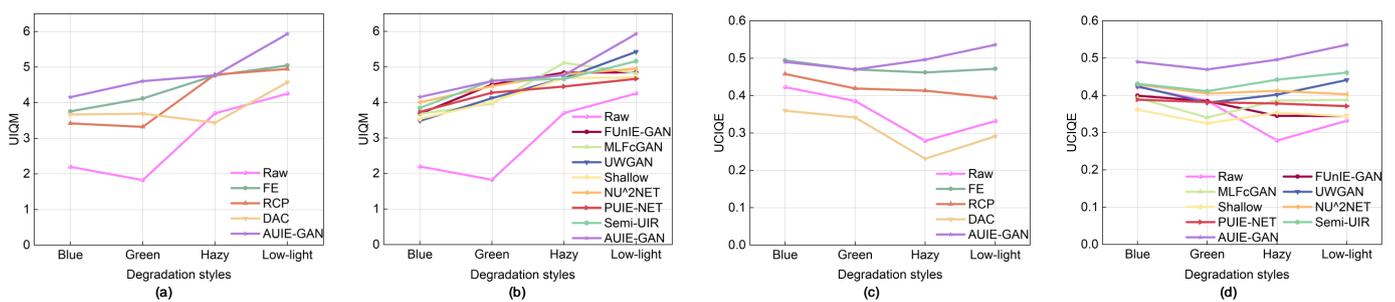

Figure 10. Comparison of enhancement results by styles. We show the traditional enhancement methods and deep learning-based enhancement methods separately. Graphs (**a**) and (**c**) represent the UIQM and UCIQE metrics of the traditional approaches, respectively, while graphs (**b**) and (**d**) represent the UIQM and UCIQE metrics of the deep learning-based approaches, respectively.

4.5. Ablation Study

We performed a series of ablation experiments to demonstrate the effectiveness of our proposed method. To improve the model's adaptive enhanced ability, we proposed to use the weight self-attention module (WSAM). In this section, firstly, we investigated the contribution of the WSAM to the model.

On the one hand, we reduced the number of WASM from the full model (Model Pruning) and used regular convolution layers to match the number of parameters of the full model; on the other hand, we used WASM without trainable weights (WASM w/o tw) in the full model.

The model pruning operation is shown in the Figure 11. We cut the complete model in turn and filled it with different numbers of convolution layers.

The results were shown in Figure 12 and Table 4. We sampled four images with different styles in the Challenge60 and processed them using the model with different connections (see Figure 11).
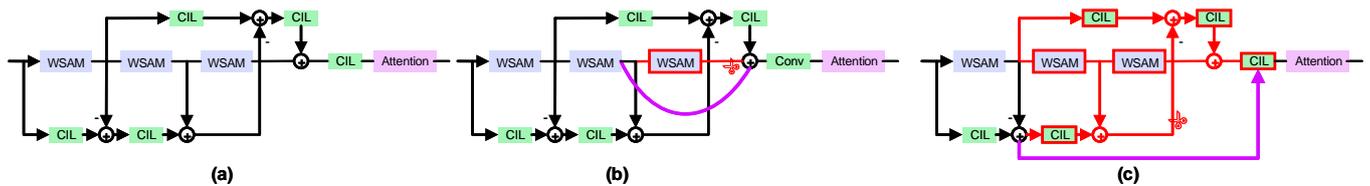
Figure 11. Pruning of SARM at different stages. (**a**) is the complete SARM, and (**b**,**c**) reduce one WSAM in turn. Significantly, model pruning would break the flow of the feature maps, so we jumped the connection to these locations where the corruption occurred. The role of the convolution layer that appears in the figure was to balance the number of parameters of the model, and the size of the perceptual field is 3.
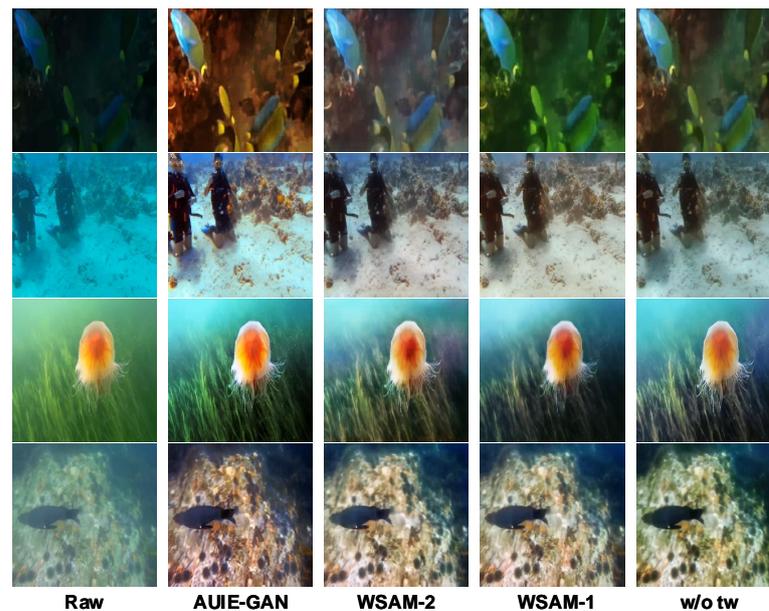

Figure 12. Visual comparison of different pruning situations. WSAM-n represents the number of WSAM in each SARM as n. tw represents the trainable weights w in the WSAM.

Table 4. Comparison of objective evaluation metrics in different pruning situations.

	Raw	WSAM-1	WSAM-2	w/o tw	AUIE–GAN
UIQM	2.9388	4.1141	4.3672	4.2390	**4.6346**
UCIQE	0.3590	0.4133	0.4104	0.4173	**0.5179**

Qualitatively, the number of WSAM corresponded to the quality of the enhanced image. As the number increased, the enhanced image was not only significantly optimized in terms of global features such as brightness and contrast, but also the details of the features were gradually enriched and the image was more detailed. When the trainable weights were removed from the model (the last column in Figure 12), there was a single enhancement effect in the processed image, although there was no collapsing drop in the

enhancement effect. Specifically, these results simply removed the color projection from the input image and were not targeted to the specific image.

Quantitatively, the complete model achieved the highest values of evaluation metrics. Incomplete models showed varying degrees of indicator jitter, which may be caused by changes in the signal flow of the model during the pruning process. Compared to the model with a single WSAM module, the complete model improved by 12.65% in the UIQM metric and 25.31% in the UCIQE metric; the trainable weight also had a significant improvement on the model, with 9.33% and 24.11% improvement in the UIQM and UCIQE metrics, respectively. However, the need for calmness is that quantitative analysis had a certain result bias and needed to be considered in combination with qualitative analysis [48].

We proposed to use some improved structures and techniques, which could effectively improve the enhancement of the model. Therefore, we performed ablation experiments on these improvement points in the second part of the ablation study. In general, they include:

(1) AUIE–GAN removes Self-Attention Residual Module (w/o SARM);

(2) AUIE–GAN removes Self-Attention Skip Connections and Global Feature vectors of U–Net (w/o SAKC&GF);

(3) AUIE–GAN removes Frequency Loss function and dual path Discriminator structure (w/o FL&dpD).

We evaluated the results of ablation experiments on the Challenge60 and SUID datasets. Qualitative experimental results were shown in Figure 13. On the one hand, there was a severe grid effect occurred and the image was not smooth enough in the result of w/o SAKC&GF. On the other hand, severe color imbalance appears when AUIE–GAN removed the SARM or FL&dpD modules. The complete model is able to balance color and semantic continuity.
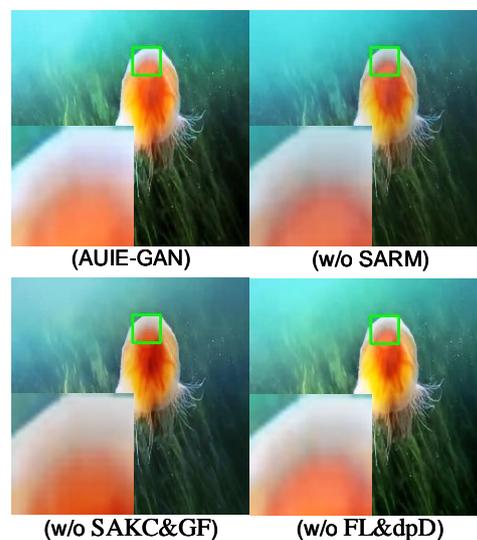

Figure 13. Results of the ablation experiment. The local zoom in the lower left corner was used to show more detailed texture information.

Quantitative experimental results were shown in Figure 14. Compared with the complete model, ablation models showed some decreases in all evaluation metrics. Furthermore, for different evaluation metrics, the lowest point of the curve appears in different models. In general, the Self-Attention Skip Connections and Global Feature vectors of U–Net had the greatest enhancement to the model.
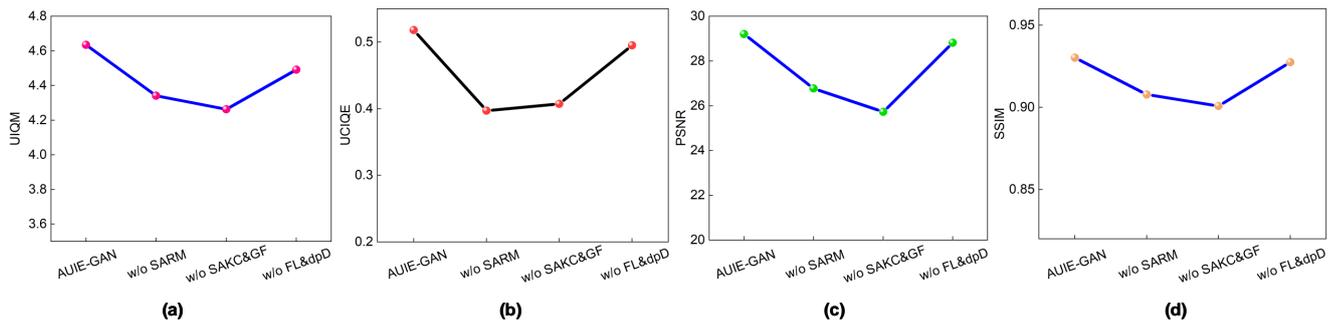
Figure 14. Comparison of quantitative evaluation of ablation experiments. We tested the UIQM and UCIQE of different ablation models on the Challenge60, and the PSNR and SSIM metrics on the SUID dataset. In the figure, (**a**,**b**) are results quantified using no-reference metrics, and (**c**,**d**) are results quantified using full-reference metrics.

5. Conclusions

In this paper, we present a new GANs-based model to resolve the color cast present in underwater images, which we refer to as AUIE–GAN. We have improved the self-attentive mechanism with a trainable weight and proposed WSAM; based on this module, we propose SARM for the generator backbone layer to improve the model's adaptability for different styles of underwater degraded images. Furthermore, we propose to use a discriminator network with a dual path structure and a frequency domain loss function to improve the enhancement quality. In general, AUIE–GAN not only achieved excellent visual results but also outperformed the best existing UIE methods in several evaluation metrics.

However, the frequency domain paths in the AUIE–GAN's discriminator were designed using fully-connected layers, which resulted in the model taking up a large amount of GPU resources during the training process. In addition, we found that AUIE–GAN shows a bias in fitting the R channel in RGB space, which led to some results is red. We will publish these results in our future works.

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]
2. Guo, Y.C.; Li, H.Y.; Zhuang, P.X. Underwater Image Enhancement Using a Multiscale Dense Generative Adversarial Network. *IEEE J. Ocean. Eng.* **2020**, *45*, 862–870. [CrossRef]
3. Goodfellow, I.; Pouget-Abadie, J.; Mirza, M.; Xu, B. Generative adversarial nets. *Commun. ACM* **2020**, *63*, 139–144. [CrossRef]
4. Li, R.; Wu, C.H.; Liu, S.C.; Wang, J. SDP-GAN: Saliency Detail Preservation Generative Adversarial Networks for High Perceptual Quality Style Transfer. *IEEE Trans. Image Process.* **2021**, *30*, 374–385. [CrossRef] [PubMed]
5. Xu, W.H.; Xie, X.H.; Lai, J.H. RelightGAN: Instance-level Generative Adversarial Network for Face Illumination Transfer. *IEEE Trans. Image Process.* **2021**, *30*, 3450–3460. [CrossRef]
6. Wu, J.J.; Liu, X.L.; Lu, Q.H.; Zin, Z. FW-GAN: Underwater image enhancement using generative adversarial network with multi-scale fusion. *Signal Process. Image Commun.* **2022**, *109*, 116855. [CrossRef]
7. Liu, Y.C.; Chan, W.H.; Chen, Y.Q. Automatic white balance for digital still camera. *IEEE Trans. Consum. Electron.* **1995**, *41*, 460–466.

8. Rahman, Z.; Jobson, D.J.; Woodell, G.A. Multi-scale retinex for color image enhancement. In Proceedings of the 3rd IEEE International Conference on Image Processing, Lausanne, Switzerland, 19 September 1996; pp. 1003–1006.
9. Zuiderveld, K. Contrast limited adaptive histogram equalization. In *Graphics Gems IV*; Heckbert, P.S., Ed.; Academic Press Professional, Inc.: San Diego, CA, USA, 1994; pp. 474–485.
10. Li, C.L.; Tang, S.Q.; Kwan, H.K.; Yan, J. Color Correction Based on CFA and Enhancement Based on Retinex With Dense Pixels for Underwater Images. *IEEE Access* **2020**, *8*, 155732–155741. [CrossRef]
11. Hitam, M.S.; Yussof, W.N.J.H.W.; Awalludin, E.A.; Bachok, Z. Mixture Contrast Limited Adaptive Histogram Equalization for Underwater Image Enhancement. In Proceedings of the 2013 International Conference on Computer Applications Technology (ICCAT), Sousse, Tunisia, 20–22 January 2013; IEEE: New York, NY, USA, 2013.
12. Tang, Z.J.; Jiang, L.Z.; Luo, Z.H. A new underwater image enhancement algorithm based on adaptive feedback and Retinex algorithm. *Multimed. Tools Appl.* **2021**, *80*, 28487–28499. [CrossRef]
13. Li, C.Y.; Guo, J.C.; Cong, R.M.; Wang, B. Underwater Image Enhancement by Dehazing With Minimum Information Loss and Histogram Distribution Prior. *IEEE Trans. Image Process.* **2016**, *26*, 5664–5677. [CrossRef]
14. Zhang, S.; Wang, T.; Dong, J.Y.; Yu, H. Underwater image enhancement via extended multi-scale Retinex. *Neurocomputing* **2017**, *245*, 100622. [CrossRef]
15. Ancuti, C.; Ancuti, C.O.; Haber, T.; Bekaert, P. Enhancing Underwater Images and Videos by Fusion. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), New York, NY, USA, 16–21 June 2012; pp. 81–88.
16. Galdran, A.; Pardo, D.; Picon, A.; Alvarez-Gila, A. Automatic Red-Channel underwater image restoration. *J. Vis. Commun. Image Represent.* **2015**, *26*, 132–145. [CrossRef]
17. Carlevaris-Bianco , N.; Mohan, A.; Eustice, R.M. Initial Results in Underwater Single Image Dehazing. In Proceedings of the Washington State Conference and Trade Center (WSCTC), Seattle, WA, USA, 20–23 September 2010 ; IEEE: New York, NY, USA, 2010.
18. Song, W.; Wang, Y.; Huang, D.M.; Tjondronegoro, D. *A Rapid Scene Depth Estimation Model Based on Underwater Light Attenuation Prior for Underwater Image Restoration*. In Proceedings of the 19th Pacific-Rim Conference on Multimedia (PCM), Hefei, China, 21–22 September 2018 ; Springer: Berlin, Germany, 2018; pp. 678–688.
19. He, K.M.; Sun, J.; Tang, X.O. Single Image Haze Removal Using Dark Channel Prior. *IEEE Trans. Pattern Anal. Mach. Intell.* **2011**, *33*, 2341–2353.
20. Drews, P.; do Nascimento, E.; Moraes, F.; Campos, M. Transmission Estimation in Underwater Single Images. In Proceedings of the IEEE International Conference on Computer Vision (ICCV), Sydney, NSW, Australia, 1–8 December 2013; IEEE: New York, NY, USA, 2013; pp. 825–830.
21. Yang, H.Y.; Chen, P.Y.; Huang, C.C.; Shiau, Y.C. Low Complexity Underwater Image Enhancement Based on Dark Channel Prior. In Proceedings of the 2011 Second International Conference on Innovations in Bioinspired Computing and Applications, Washington, DC, USA, 16–18 December 2011; IEEE: Shenzhen, China, 2011; pp. 17–20.
22. Wang, N.; Zheng, H.Y.; Zheng, B. Underwater Image Restoration via Maximum Attenuation Identification. *IEEE Access* **2017**, *5*, 18941–18952. [CrossRef]
23. Schechner, Y.Y.; Karpel, N. Clear underwater vision. In Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), Washington, DC, USA, 27 June–2 July 2004. [CrossRef]
24. Hu, H.; Zhao, L.; Li, X.; Wang, H.; Liu, T. Underwater Image Recovery Under the Nonuniform Optical Field Based on Polarimetric Imaging. *IEEE Photonics J.* **2018**, *10*, 1–9. [CrossRef]
25. Hu, H.; Zhao, L.; Li, X.; Wang, H. Polarimetric image recovery in turbid media employing circularly polarized light. *Opt. Express.* **2018**, *26*, 25047–25059 . [CrossRef] [PubMed]
26. Li, X.; Hu, H.; Zhao, L.; Wang, H. Polarimetric image recovery method combining histogram stretching for underwater imaging. *Sci. Rep.* **2018**, *8*, 12430. [CrossRef] [PubMed]
27. Treibitz, T.; Schechner, Y.Y. Active Polarization Descattering. *IEEE Trans. Pattern Anal. Mach. Intell.* **2009**, *31*, 385–399. [CrossRef]
28. Han, P.; Liu, F.; Yang, K.; Ma, J. Active underwater descattering and image recovery. *Appl. Opt.* **2017**, *56*, 6631–6638. [CrossRef]
29. Naik, A.; Swarnakar, A.; Mittal, K. Shallow-UWnet: Compressed Model for Underwater Image Enhancement (Student Abstract). In Proceedings of the 35th AAAI Conference, San Francisco, CA, USA, 9–10 August 2011; Association Advancement Artificial Intelligence: Palo Alto, CA, USA, 2021; pp.15853–15854.
30. Guo, C.L.; Wu, R.Q.; Jin, X.; Han, L. Underwater Ranker: Learn Which Is Better and How to Be Better. Available online: https://arxiv.org/abs/2208.06857 (accessed on 18 June 2023).
31. Fu, Z.Q.; Wang, W.; Huang, Y.; Ding, X. Uncertainty Inspired Underwater Image Enhancement. In Proceedings of the 17th European Conference on Computer Vision (ECCV), Tel Aviv, Israel, 23–27 October 2022; Springer: Berlin, Germany, 2022; pp. 465–482.
32. Wang, Y.D.; Guo, J.C.; Gao, H.; Yue, H. UIEC2-Net: CNN-Based Underw. Image Enhanc. Using Two Color Space. *Signal Process. Image Commun.* **2021**, *96*, 116250. [CrossRef]
33. Qi, Q.; Zhang, Y.C.; Tian, F.; Wu, J. Underwater Image Co-Enhancement With Correlation Feature Matching and Joint Learning. *IEEE Trans. Circuits Syst. Video Technol.* **2022**, *32*, 1133–1147. [CrossRef]
34. Li, C.Y.; Guo, J.C.; Guo, C.L. Emerging From Water: Underwater Image Color Correction Based on Weakly Supervised Color Transfer. *IEEE Signal Process. Lett.* **2018**, *25*, 323–327. [CrossRef]

35. Islam, M.J.; Xia, Y.Y.; Sattar, J. Fast Underwater Image Enhancement for Improved Visual Perception. *IEEE Robot. Autom. Lett.* **2020**, *5*, 3227–3234. [CrossRef]

36. Liu, X.D.; Gao, Z.; Chen, B.M. MLFcGAN: Multilevel Feature Fusion-Based Conditional GAN for Underwater Image Color Correction. *IEEE Geosci. Remote Sens. Lett.* **2020**, *17*, 1488-1492. [CrossRef]

37. Ronneberger, O.; Fischer, P.; Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. In *Lecture Notes in Computer Science, Proceedings of the Medical Image Computing and Computer-Assisted Intervention (MICCAI), Munich, Germany, 5–9 October 2015*; Navab, N., Hornegger, J., Wells, W., Frangi, A., Eds.; Springer: Berlin/Heidelberg, Germany, 2015; Volume 9351.

38. Huang, J.; Zhu, P.F.; Geng, M.R.; Zhou, X. Range Scaling Global U-Net for Perceptual Image Enhancement on Mobile Devices. In Proceedings of the 15th European Conference on Computer Vision (ECCV), Munich, Germany, 8–14 September 2019; Springer: Berlin, Germany, 2019; pp. 230–242.

39. Fuoli, D.; Van Gool, L.; Timofte, R. Fourier Space Losses for Efficient Perceptual Image Super-Resolution. In Proceedings of the 18th IEEE/CVF International Conference on Computer Vision (ICCV), Montreal, QC, Canada, 11–17 October 2021; IEEE: New York, NY, USA, 2021; pp. 2340–2349.

40. Miyato, T.; Kataoka, T.; Koyama, M.; Yoshida, Y. Spectral Normalization for Generative Adversarial Networks. Available online: https://arxiv.org/abs/1802.05957 (accessed on 18 June 2023).

41. Wang, Z.; Bovik, A.C.; Sheikh, H.R.; Simoncelli, E.P. Image quality assessment: From error visibility to structural similarity. *IEEE Trans. Image Process.* **2020**, *13*, 600–612. [CrossRef]

42. Ignatov, A.; Kobyshev, N.; Timofte, R.; Vanhoey, K. DSLR-Quality Photos on Mobile Devices with Deep Convolutional Networks. In Proceedings of the 16th IEEE International Conference on Computer Vision (ICCV), Vancouver, BC, Canada, 22–23 September 2017; IEEE: New York, NY, USA, 2017; pp. 3297–3305.

43. Johnson, J.; Alahi, A.; Li, F.F. Perceptual Losses for Real-Time Style Transfer and Super-Resolution. In Proceedings of the 14th European Conference on Computer Vision (ECCV), Amsterdam, The Netherlands, 8–16 October 2016; Springer: Berlin, Germany, 2016; pp. 694–711.

44. Li, C.Y.; Guo, C.L.; Ren, W.Q.; Tao, D. An Underwater Image Enhancement Benchmark Dataset and Beyond. *IEEE Trans. Image Process.* **2020**, *29*, 4376–4389. [CrossRef] [PubMed]

45. Hou, G.J.; Zhao, X.; Pan, Z.K.; Li, J. Benchmarking Underwater Image Enhancement and Restoration, and Beyond. *IEEE Access* **2020**, *8*, 122078–122091. [CrossRef]

46. Liu, R.S.; Fan, X.; Zhu, M.; Hou, M. Real-World Underwater Enhancement: Challenges, Benchmarks, and Solutions Under Natural Light. *IEEE Trans. Circuits Syst. Video Technol.* **2020**, *30*, 4861–4875. [CrossRef]

47. Lai, Y.H.; Zhou, Z.; Su, B.H.; Chen, J. Single underwater image enhancement based on differential attenuation compensation. *Front. Mar. Sci.* **2022**, *9*, 1047053. [CrossRef]

48. Huang, S.R.; Wang, K.Y.; Liu, H.; Li, Y. Contrastive Semi-Supervised Learning for Underwater Image Restoration via Reliable Bank. Available online: https://arxiv.org/abs/2303.09101 (accessed on 18 June 2023).