

Article

MTUW-GAN: A Multi-Teacher Knowledge Distillation Generative Adversarial Network for Underwater Image Enhancement

Tianchi Zhang and Yuxuan Liu *

School of Information Science and Engineering, Chongqing Jiaotong University, Chongqing 400074, China; zhangtianchi@cqjtu.edu.cn

* Correspondence: liuyx0086@gmail.com

Abstract: Underwater imagery is plagued by issues such as image blurring and color distortion, which significantly impede the detection and operational capabilities of underwater robots, specifically Autonomous Underwater Vehicles (AUVs). Previous approaches to image fusion or multi-scale feature fusion based on deep learning necessitated multi-branch image preprocessing prior to merging through fusion modules. However, these methods have intricate network structures and a high demand for computational resources, rendering them unsuitable for deployment on AUVs, which have limited resources at their disposal. To tackle these challenges, we propose a multi-teacher knowledge distillation GAN for underwater image enhancement (MTUW-GAN). Our approach entails multiple teacher networks instructing student networks simultaneously, enabling them to enhance color and detail in degraded images from various perspectives, thus achieving an image-fusion-level performance. Additionally, we employ middle layer channel distillation in conjunction with the attention mechanism to extract and transfer rich middle layer feature information from the teacher model to the student model. By eliminating multiplexed branching and fusion modules, our lightweight student model can directly generate enhanced underwater images through model compression. Furthermore, we introduce a multimodal objective enhancement function to refine the overall framework training, striking a balance between a low computational effort and high-quality image enhancement. Experimental results, obtained by comparing our method with existing approaches, demonstrate the clear advantages of our proposed method in terms of visual quality, model parameters, and real-time performance. Consequently, our method serves as an effective solution for real-time underwater image enhancement, specifically tailored for deployment on AUVs.

Keywords: underwater image enhancement; image fusion; multi-teacher knowledge distillation; channel distillation



Citation: Zhang, T.; Liu, Y. MTUW-GAN: A Multi-Teacher Knowledge Distillation Generative Adversarial Network for Underwater Image Enhancement. *Appl. Sci.* **2024**, *14*, 529. <https://doi.org/10.3390/app14020529>

Academic Editor: Atsushi Mase

Received: 20 November 2023

Revised: 3 January 2024

Accepted: 3 January 2024

Published: 8 January 2024



Copyright: © 2024 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/>).

1. Introduction

Autonomous Underwater Vehicles (AUVs) possess distinctive features such as unmanned and cable-free operation, as well as a wide range of mobility, making them extensively utilized in ocean observation, mine search, rescue missions, and marine military applications [1,2]. During underwater exploration and operations, AUVs rely primarily on visual images as a means to gather scene information at short distances [3]. However, the quality of underwater visual images is considerably lower than that of land-based images, exhibiting issues such as blurring, color distortion, low contrast, and high noise [4]. Consequently, it has become imperative to enhance underwater images to enable effective target identification and localization by AUVs.

Traditional methods for improving the clarity of degraded underwater images can be broadly categorized into two approaches: image restoration methods based on physical models and image enhancement methods based on non-physical models [5]. In the realm

of underwater image restoration, researchers have drawn inspiration from the DCP algorithm [6]. This algorithm aims to clarify underwater images by estimating the degradation model and parameters, thereby inverting the imaging process [7,8]. However, the diverse range of degradation types encountered in underwater images poses a challenge in determining a single imaging model with fixed parameters based on a priori assumptions. On the other hand, underwater image enhancement methods bypass the need to consider the imaging process directly. For instance, histogram equalization [9], a fundamental linear mapping method, enhances the image contrast by adjusting the tonal distribution through a constructed function. Methods rooted in Retinex theory [10] are employed to correct color deviations in underwater images, thus enhancing the overall color quality. Nevertheless, these techniques are prone to issues such as local over-enhancement and color inhomogeneity, leading to the introduction of new noise and color biases in the image. To address these limitations, image fusion algorithms [11] have been developed. These algorithms combine multiple enhanced underwater images with specific weights, effectively reducing noise and local oversaturation resulting from a single enhancement method. However, the requirement of preprocessing multiple enhanced images prior to fusion restricts the applicability of such methods in high real-time processing scenarios involving AUVs.

In recent years, the rapid advancement of AI technology has propelled the rapid development of deep learning, which boasts powerful feature learning capabilities. Image processing methods based on deep learning have exhibited remarkable performance [12–14]. Consequently, numerous enhanced techniques have emerged within the domain of underwater image enhancement research [15–20]. Li et al. [17] conducted experiments that demonstrate the commendable performance of the fusion base [11] on a diverse range of underwater images. Furthermore, they propose a gated fusion network called Water-Net, which enhances underwater images by leveraging white balance, histogram equalization, and gamma correction algorithms as inputs to the network model. Subsequently, the network model predicts corresponding confidence maps, which are then fused with multiple inputs to obtain the enhanced image. While this method effectively enhances the performance of image fusion, its network structure is intricate and it requires different algorithms to generate three types of augmented images as input data for the network model. Consequently, its practicality is significantly impacted, rendering it unsuitable for real-time augmentation of underwater images by AUVs.

Several researchers have proposed the use of multi-scale feature fusion techniques to enhance network models for underwater image enhancement [21–23]. Unlike image fusion methods, these approaches do not require the fusion of multiple inputs for enhancement. Instead, they leverage the feature information from the intermediate layer of the neural network and fuse multi-scale feature maps generated through gradual convolution of underwater images within the network. This fusion process aims to improve the overall performance of underwater image enhancement. Liu et al. [21] introduced MLFcGAN, a multiscale feature fusion framework. This framework involves extracting feature maps at different scales by gradually downsampling the input image using a convolutional layer. The fused features are then fed into the corresponding decoder layer, which incorporates a feature fusion module and skip-connect operation. The resulting enhanced image is obtained through multilevel feature fusion. Tian et al. [23] proposed an end-to-end framework called heterogeneous feature fusion and dynamic feature enhancement. This framework integrates a feature fusion module with an attention mechanism, enhancing the encoder structure by introducing an improved feature attention module. Features are gradually extracted at different scales, ranging from low to high levels, and the decoder employs an upsampling operation to reconstruct feature vectors and progressively restore them to clear images. Furthermore, Liu et al. [24] presented MGF-cGAN, a multiscale gated fusion framework for underwater image enhancement. The generator in this framework utilizes a multiscale feature extraction module to extract feature information at different scales from three parallel sub-networks. A gated fusion module (GFM) is introduced to gradually fuse the three feature images collected from the multiscale feature extraction module using a

recursive strategy, thereby improving contrast and color saturation in underwater images, while the multiscale feature fusion methods eliminate the need for preprocessing multiple enhanced images. The operations involved, such as multiscale feature extraction, fusion modules, and dense connectivity in the network, increase the complexity and the number of parameters. Consequently, these algorithms struggle to strike a balance between low computation requirements and high-quality image enhancement in hardware-constrained practical applications.

Among the available underwater image enhancement methods, image fusion methods and multi-scale feature fusion methods exhibit superior performance. However, their network models employ complex multi-branch structures with a substantial number of parameters. This poses a challenge for AUVs equipped with small-capacity batteries, as they often have a limited computational power and storage capacity. Consequently, these advanced algorithms are impractical for local real-time applications. In underwater environments, effective real-time image processing algorithms are essential. A typical AUV camera captures and transmits between 10 and 30 frames per second. To ensure a seamless underwater observation experience for operators, the processing time for each image should ideally be less than 0.1 s. Therefore, there is a pressing need for a comprehensive conceptual and structural enhancement of fusion-based algorithms that can deliver both an improved performance and real-time capabilities.

In the domain of image dehazing, Lan et al. [12] proposed the Online Knowledge Distillation Network (OKDNet) as a solution. This approach involves the construction of a multiscale feature extraction network utilizing residual dense blocks guided by an attention mechanism. The network generates rich features, which are then sent to different branches for further processing. Supervised training is employed to generate two styles of dehazed images, which are subsequently fused using parallel convolution. Additionally, the model is optimized through online knowledge distillation, resulting in a reduction in the model parameters to 2.58 M. This optimization significantly enhances the algorithm's real-time applicability. Hence, employing knowledge distillation for model compression in real-time underwater image enhancement is a viable approach. Furthermore, Wang et al. [25] proposed that a student can learn better from multiple teachers, who are more informative and instructive than a single teacher. Motivated by their work, we present a novel framework called multi-teacher knowledge distillation GAN (MTUW-GAN) for underwater image enhancement. Through extensive experiments, we demonstrate that MTUW-GAN surpasses existing methods in terms of visual quality, model parameters, and complexity. Moreover, it effectively balances an enhanced performance with real-time capabilities.

In conclusion, fusion-based methods for underwater image enhancement exhibit a superior performance and have been extensively studied. While these methods contribute to enhancing the overall image quality, there are several important considerations to address in underwater image enhancement scenarios:

- The requirement to preprocess multiple enhanced images in the image fusion algorithm hampers practical usability and necessitates innovative approaches to the overall concept.
- The inclusion of feature extraction branches and fusion modules in the multi-scale feature fusion method elevates computational demands and compromises the real-time performance of the algorithm.
- The image translation loss function exhibits limited capability in rectifying color bias and facilitating detail recovery when employed for underwater image enhancement.
- Previous fusion-based underwater image enhancement algorithms encounter challenges in striking a balance between enhancement effectiveness and real-time performance, with the need to process each image within a time frame of 0.1 s.

To tackle these challenges, we aim to develop AUV underwater image enhancement algorithms that meet the demand for a high-quality output while ensuring real-time performance. In this regard, we introduce MTUW-GAN as our proposed solution. The key innovations of our approach are as follows:

- We propose a framework called MTUW-GAN, which changes the idea of “multi-image fusion” to “multi-teacher teaching”, and guides the student network to perform underwater image enhancement through multiple teacher networks.
- We incorporate the attention mechanism to extract the channel attention information from the teacher network and transfer it to the middle layer of the student network in the form of a loss function, which improves the performance without increasing the number of parameters and computations of the student network.
- We formulate a multimodal objective function to fully restore the visual quality of underwater images by removing noise and color distortion caused by underwater light scattering based on the global content, color, and texture information of the images.
- We comprehensively verify that the method in this paper can achieve image fusion level performance by designing comparison experiments and ablation experiments, taking into account the performance enhancement and real-time performance.

To provide a more intuitive illustration of the innovative ideas presented in this paper, particularly in comparison to previous image fusion and multi-scale feature fusion methods, Figure 1 showcases the overall process enhancement achieved by the proposed method.

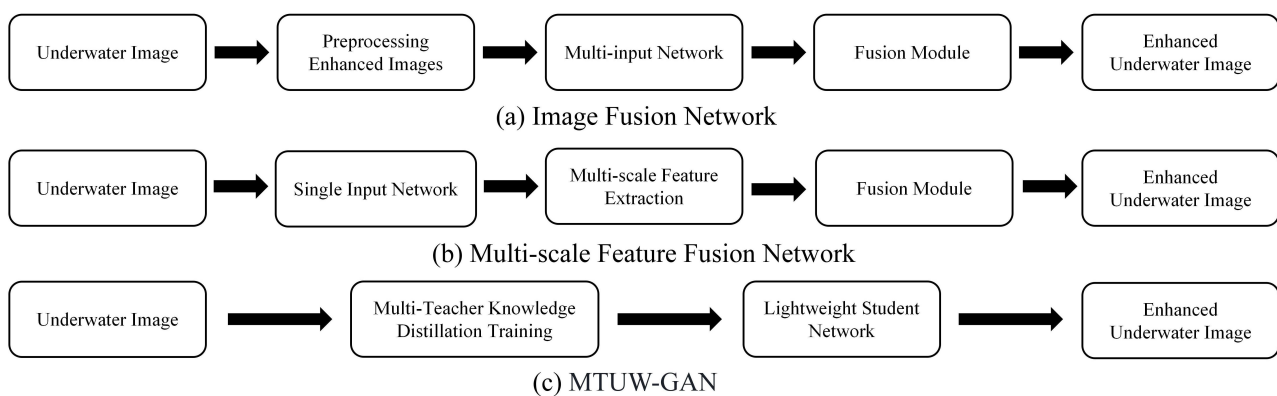


Figure 1. Comparison between MTUW-GAN and fusion-based method processes. MTUW-GAN does not need a preprocessing enhanced images and fusion module, and only needs to output the enhanced underwater images directly via a lightweight student network after framework training is completed.

The remainder of this paper is structured as follows: In Section 2, we present an introduction to MTUW-GAN. Section 3 conducts quantitative and qualitative analyses, along with a comprehensive ablation experiment. Lastly, Section 4 provides a summary of our work.

2. Materials and Methods

In this paper, we propose MTUW-GAN, an underwater image enhancement method based on multi-teacher knowledge distillation. The overall framework is depicted in Figure 2, consisting of three main components: the Main Teacher Generator (G_T^M), responsible for the primary instructional contribution; the Assistance Teacher Generator (G_T^A), responsible for the secondary instructional contribution; and the Lightweight Student Generator (G_S). During the training process, multiple teacher networks concurrently learn the end-to-end transformations from degraded underwater images to real underwater images. The student networks, guided by the teachers, enhance the underwater images and merge them to achieve image fusion. Moreover, G_S are trained to distill the knowledge from the intermediate layers of the G_T^M , ensuring the preservation of crucial details. To address the issue of over-fitting in offline distillation [26], we employ online knowledge distillation throughout the training process, allowing the multi-teacher network to gradually guide the students in learning the feature information, as depicted in Figure 2. During training, all mentioned tasks are executed simultaneously, while in the testing phase, each teacher

and student model can be independently run. Only the component represented by the blue solid line, corresponding to the Student Generator (G_S) after model compression, is utilized, as it aligns with our task objective. The specific details and loss functions of each component are presented below.

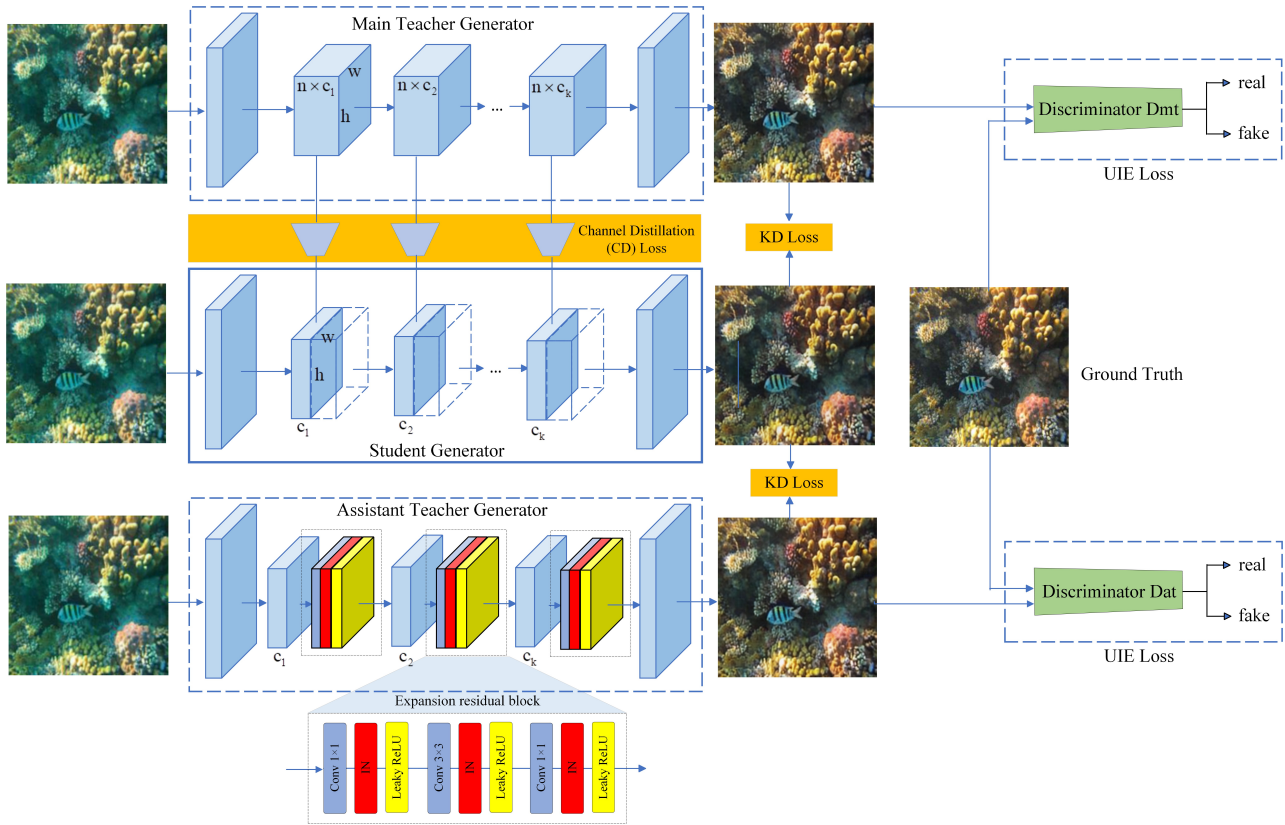


Figure 2. The overall structure of MTUW-GAN. The different structural teacher generators provide complementary knowledge, and the Lightweight Student Generator (G_S) in the blue solid box learns to enhance underwater images under the guidance of multiple teachers (G_T^M and G_T^A). n in G_T^M denotes each channel of the convolutional layer divided by the channel compression factor.

2.1. Multi-Teacher Teaching

To meet the requirements of real-time enhancement of underwater images by AUVs, achieving a balance between a high-quality output and a low complexity is crucial. Previous image fusion methods have relied on the merging of multiple enhanced images to synthesize clear underwater images, exploiting the benefits of information integration. However, such approaches often involve multiple pre-processing steps and image fusion network modules, resulting in increased computational demands. In order to address this challenge, we propose a novel approach that shifts the paradigm from “multi-image fusion” to “multi-teacher teaching”. In this approach, the student model learns from multiple teacher models, enabling it to incorporate the complementary advantages of multiple enhanced underwater images. Consequently, our method eliminates the need for image preprocessing and image fusion network modules. Furthermore, the adoption of model compression in the lightweight student model enables a harmonious balance between performance enhancements and real-time capabilities.

The concept of multi-teacher knowledge distillation resembles the cognitive process of human learning, where students deepen their understanding by receiving guidance from multiple teachers. The utilization of multiple teachers effectively enhances the performance of the student model. However, when there is a significant capacity gap between the teacher and student models, the transfer of knowledge from the teacher to the student model may

be insufficient, resulting in a degradation of the student model's performance [27]. To address this issue, our framework adopts a consistent MobileNet-Style [28] architecture for both the teacher and student models. This architectural consistency reduces the number of model parameters while ensuring effective knowledge transfer. Li et al. [29] demonstrated that the ResBlock layer in the ResNet-based generator [30] accounts for most of the model parameters and computational cost, remaining unaffected by decomposition. Consequently, to reduce model redundancy, we employ channel pruning techniques [31,32] to decrease the channel width of the ResBlock layer. By applying channel pruning to the robust G_T^M , we obtain a more streamlined G_S . The specific settings are to set the channel compression factor n to 4 and to reduce the number of convolution filters of G_T^M from 64 to 16 for G_S . This model compression approach not only improves computational efficiency but also provides a better foundation for subsequent model optimization and deployment.

To achieve image fusion-level performance for G_S , we incorporated the Assistance Teacher Generator (G_T^A) into the framework. Inspired by the residual network [33], we incorporate expansion residual blocks to facilitate additional nonlinear transformations that extract more expressive information features from underwater images. This aids in accurate recovery of details, colors, and contrasts in the underwater images. The robust G_T^A exhibits a faster convergence rate compared to the shallow G_S , aligning with the cognitive process of teachers guiding students in the learning process.

By engaging in collaborative training between the G_T^M and the G_T^A , teacher generators with diverse structures can effectively capture the intricate relationships between input image features and augmented images from different perspectives. Consequently, the knowledge gained from these diverse perspectives can be transferred to the G_S . This approach offers several advantages. Firstly, the G_S is relieved from the need to preprocess the underwater image into multiple augmented images. Additionally, there is no requirement for an additional image fusion module to merge these images. Furthermore, during the training process, the G_S is not involved in adversarial learning with the discriminator. This reduction in complexity and training cost significantly alleviates the hardware requirements of the entire framework.

2.2. Channel Distillation Module

Current methods for enhancing underwater images employ multi-scale feature fusion techniques, which involve extracting and merging feature information from intermediate layers of neural networks. These methods typically utilize feature extraction modules based on residual blocks [33] and dense blocks [34] to extract features from convolutionally generated multiscale feature maps. These features are then fused together using operations like channel splicing and dense joining to generate enhanced underwater images [21–23]. However, these feature extraction branches and fusion modules significantly increase computational demands and impact the real-time performance of the algorithm. To address this issue, we leverage the knowledge distillation framework to transfer middle layer feature information from the teacher model to the student model through a loss function. This approach effectively improves the model's performance without increasing the number of parameters or computational requirements of the student network.

In this study, we employ the channel distillation [35] to effectively enhance the texture details of underwater images by extracting and transferring middle layer feature information from the teacher model. Initially, a 1×1 convolution is utilized to address the discrepancy in the number of channels between the G_T^M and the G_S feature maps. This expansion of channels in the student generator ensures compatibility. Subsequently, we utilize global average pooling (GAP) to compute the importance of each channel's feature map, which reflects the attention information associated with that channel. We regard the attention information of each channel's feature map as knowledge. The G_T^M and the G_S independently calculate the attention information of each channel from their respective feature maps. The G_T^M then guides the G_S to learn the attention information of each channel and transfers this knowledge to the G_S . Through this process, the G_S can acquire

the teacher’s attention information for each channel, resulting in an enhancement of its performance. Consequently, the texture information of the underwater image is further improved. Notably, the Channel Distillation (CD) module only utilizes computational resources during the training process of the framework, and there is no need to invoke this module when utilizing the student model. Figure 3 provides a detailed depiction of the CD module architecture.

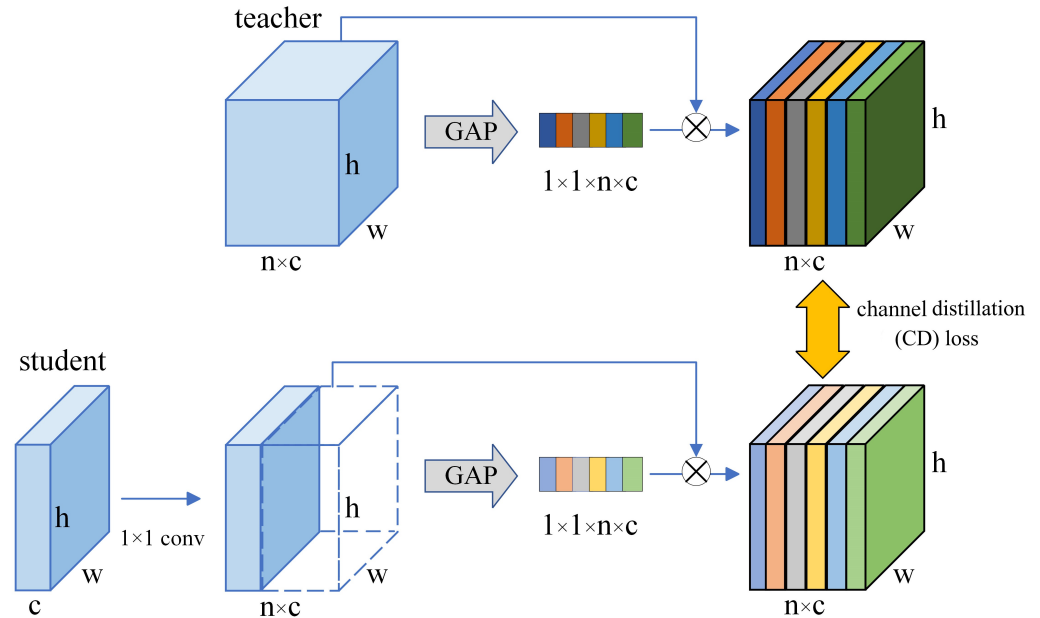


Figure 3. The Channel Distillation module.

In Figure 3, we use GAP to compute the importance of each channel’s feature map, which represents the attention information of each channel [35,36]. The weight of each channel is defined as follows:

$$w_c = \frac{1}{H \times W} \sum_{i=1}^H \sum_{h=1}^W \mu_c(i, j) \tag{1}$$

where w_c denotes the weight of the c th channel. H, W are the spatial dimensions of the feature map and $\mu_c(i, j)$ is the activation work. When calculating the channel distillation (CD) loss, due to the inconsistency of the number of G_S and G_T^M channels, it is necessary to upgrade the dimensionality by using 1×1 convolution first, and then perform channel distillation. The CD loss is defined as follows:

$$\mathcal{L}_{CD}(G_T^M, G_S) = \frac{\sum_{i=1}^f \sum_{j=1}^c (w_{t_M}^{ij} - w_s^{ij})^2}{f \times c} \tag{2}$$

where w^{ij} denotes the attention weight of the j th channel of the i th feature map of the model, f denotes the number of feature maps, and c represents the number of channels.

2.3. Loss Function

The loss function of the proposed MTUW-GAN in this paper comprises three primary components: the underwater image enhancement (UIE) loss, the knowledge distillation (KD) loss, and the channel distillation (CD) loss.

When image translation losses [37] are applied to underwater image enhancement, they often encounter challenges, resulting in deficiencies such as blurred details, a low contrast, and overall darkness in the generated images. To address these issues and generate high-quality enhanced underwater images, we propose the underwater image enhancement

(UIE) loss. This loss formulation effectively eliminates noise and color distortion caused by underwater light scattering by restoring the visual perception quality of underwater images. The restoration process is achieved by utilizing the global content, color, and local texture information of the images. Among them, for the teacher model, the GAN objective is formalized as:

$$\min_{G_T} \max_D V(D, G_T) = \mathbb{E}_{y \sim p_{\text{train}}(y)} [\log D(y)] + \mathbb{E}_{x \sim p_{\text{gen}}(x)} [\log(1 - D(G_T(x)))] \quad (3)$$

In the context of the image dehazing task, the \mathcal{L}_1 loss demonstrates superior capability in aligning the feature distribution between pixels in hazy and clear images compared to the \mathcal{L}_2 loss [38]. Moreover, the \mathcal{L}_1 loss mitigates the risk of introducing blurred artifacts. Therefore, we employ the \mathcal{L}_1 loss to enhance the dehazing performance of the network. The calculation formula for the \mathcal{L}_1 loss is expressed as follows:

$$\mathcal{L}_1(G_T) = \mathbb{E}_{x,y} [\|y - G_T(x)\|_1] \quad (4)$$

To enhance the generation of images with advanced features and details, we calculate the content loss [19] by utilizing the output of the middle layer of VGG-19 [39] as a feature representation. This approach assists the generator in producing high-quality images. The content loss is defined as the Euclidean distance between the generated underwater image from the teacher model and the corresponding ground truth. It can be expressed as follows:

$$\mathcal{L}_{\text{content}}(G_T) = \mathbb{E}_{(x,y)} [\|\varphi_j(y) - \varphi_j(G_T(x))\|_2] \quad (5)$$

The UIE loss is represented as follows:

$$\mathcal{L}_{\text{UIE}}(G_T^*) = \min_{G_T} \max_D V(D, G_T) + \lambda_1 \mathcal{L}_1(G_T) + \lambda_{\text{content}} \mathcal{L}_{\text{content}}(G_T) \quad (6)$$

To optimize the G_S , it is necessary to simulate the outputs of multiple G_T models. To accomplish this, we introduce the knowledge distillation (KD) loss. For the sake of clarity, we denote the outputs of the G_T and G_S models as t and s , respectively. The optimization process of G_S involves simulating the outputs of multiple G_T models.

The SSIM index [40] quantifies the similarity between two images by considering their brightness, contrast, and structure. It serves as a valuable metric for measuring image similarity. In our study, we utilize the SSIM index to enhance the local structure and detail information of the generated image. The SSIM loss for t and s is defined as follows:

$$\mathcal{L}_{\text{SSIM}}(t, s) = 1 - \text{SSIM}(t, s) \quad (7)$$

Relying solely on the SSIM loss during training may lead to significant geometric distortions in the generated images. Furthermore, since the SSIM loss is primarily designed for grayscale images, it can fall short in accurately assessing the quality of color images, resulting in color bias within the generated outputs. To mitigate these issues, we integrate perceptual loss [41] into the training process. By leveraging a pre-trained VGG network, we extract high-level feature representations from t and s to approximate the perceptual information between them. The expression for perceptual loss is as follows:

$$\mathcal{L}_{\text{Per}}(t, s) = \frac{1}{C_j H_j W_j} \|\varphi_j(t) - \varphi_j(s)\|_2^2 \quad (8)$$

To enhance the spatial smoothness of the generated images produced by G_S , we incorporate the total variation (TV) loss [42] into the training process. This additional loss term serves to constrain network learning by reducing the discrepancies between neighboring pixel points. By doing so, it effectively preserves the fine details of the generated images while removing excessive noise. The TV loss is represented as follows:

$$\mathcal{L}_{TV} = \text{Mean} \left\| \sum_{h,w} (s_{h+1,w} - s_{h,w}) + (s_{h,w+1} - s_{h,w}) \right\|_2^2 \quad (9)$$

The overall KD loss is obtained by combining \mathcal{L}_{SSIM} , \mathcal{L}_{Per} , and \mathcal{L}_{TV} , and can be expressed as follows:

$$\mathcal{L}_{KD}(t^*, s) = \lambda_{SSIM} \mathcal{L}_{SSIM} + \lambda_{Per} \mathcal{L}_{Per} + \lambda_{TV} \mathcal{L}_{TV} \quad (10)$$

To sum up, after adding the CD loss, the total target loss form of the whole distillation is:

$$\mathcal{L}(G_T^M, G_T^A, G_S) = \mathcal{L}_{UIE}^{multi}(G_T^M, G_T^A) + \mathcal{L}_{KD}^{multi}(t^M, t^A, s) + \lambda_{CD} \mathcal{L}_{CD}(G_T^M, G_S) \quad (11)$$

3. Experiments and Analysis

Previous fusion-based methods for enhancing underwater images face challenges in balancing the quality of enhancement and real-time usage. This is primarily due to the complexity of the process and network structure involved. To validate the effectiveness of the proposed method in this paper, a comprehensive evaluation is conducted in this section through comparative experiments and ablation experiments. Firstly, we discuss the detailed setup of the proposed framework training. Subsequently, we demonstrate the performance of our method by comparing it with other non-deep and deep learning methods on both synthetic and real underwater images. This comparison is carried out using subjective and objective evaluation methods, focusing on parametric quantities, GFLOPs, and real-time performance of different algorithms. Finally, ablation experiments are performed to showcase the effectiveness of each module within the Channel Distillation (CD) module, loss function, and multi-teacher model employed in the framework of this paper.

3.1. Experiment Details

To assess the effectiveness of the algorithms, all experiments were conducted within the PyTorch framework and executed on NVIDIA GeForce GTX 3060 GPUs (12G). Data augmentation techniques, including random rotations and horizontal flips, were applied during the training of MTUW-GAN. Additionally, all input images in the training set were resized to dimensions of 256×256 pixels. The adversarial training involved simultaneous and progressive optimization of G_T^M , G_T^A , and G_S , as the entire optimization process follows an online knowledge distillation framework. The Adam optimizer was utilized for optimization, with an initial learning rate of 2×10^{-4} . The parameters β_1 and β_2 were set to 0.5 and 0.999, respectively. The batch size was set to 1, and a total of 200 epochs were trained.

To demonstrate the robustness of our proposed method, we conducted training and testing using two publicly available underwater datasets: EUVP [19] and UIEBD [17]. The training set comprises real underwater images paired with their corresponding clear images. Specifically, we employed 11,435 images from the EUVP dataset to train our model, encompassing both degraded and high-quality underwater images. For testing, we randomly selected 300 pairs of images from the EUVP test set, representing various underwater scenes.

The UIEBD dataset offers a diverse collection of underwater scenes and encompasses a wide range of degradation types. The images within this dataset cover various domains such as marine ecology, divers, submarine corals, and coral reefs. We utilized a random subset of 800 images from the UIEBD dataset for training purposes, while the remaining 90 images were reserved for testing.

3.2. Compared Methods

To address the problem that traditional enhancement methods have an insufficient generalization ability and that deep-learning-based methods, especially fusion-based methods, are unable to balance performance enhancements and real-time performance, we compared the method proposed in this paper with the following seven methods: UDCP [7], Fusion-based [11], UGAN [18], Water-Net [17], CWR [20], and two feature-based algo-

rithms, MSBDN-DFF [43] and FFA-Net [44], for image dehazing. Among them, UDCP and Fusion-based algorithms are traditional methods for underwater image processing, while the others are deep-learning-based methods that require the measurement of model parameters and the number of GFLOPs. Water-Net, MSBDN-DFF, and FFA-Net are based on image fusion or multi-scale feature fusion, while the other methods are non-fusion methods.

3.3. Evaluation Metrics

To validate the image enhancement quality achieved by the method proposed in this paper, we employed several objective evaluation metrics [45] to assess the performance of our underwater image enhancement technique. Firstly, we utilize the mean square error (MSE) to calculate the average squared difference between the pixels of the enhanced underwater image and the reference image. The MSE provides a quantitative measure of the overall distortion between the enhanced image and the reference image, with a lower MSE value indicating a better image quality. The second metric employed is the peak signal-to-noise ratio (PSNR), commonly used to evaluate the signal reconstruction quality in domains such as image compression. The PSNR measures the ratio between the peak signal and the noise level and serves as an indicator of the extent of distortion in the processed underwater image compared to the reference image. Higher PSNR values indicate lower levels of distortion. To assess the visual quality of the enhanced image, we utilized the structural similarity index (SSIM). The SSIM evaluates the similarity between the enhanced image and the reference image based on key features, including luminance, contrast, and structural information. Higher SSIM values indicate a higher degree of similarity between the enhanced image and the reference image, reflecting an improved visual quality.

Furthermore, we calculated the average underwater image quality measure [46] (UIQM) for the quantitatively enhanced images. The UIQM is a metric that takes into account the degradation mechanism and optical imaging properties specific to underwater images. This metric comprises three distinct components: the underwater image color metric (UICM), the underwater image sharpness metric (UISM), and the underwater image contrast metric (UIConM). These metrics, respectively, evaluate the color fidelity, sharpness, and contrast of the underwater images. Each attribute metric can be utilized independently to assess a specific aspect of underwater image degradation. Higher scores in the UIQM indicate a more perceptually consistent outcome in line with human visual perception. The expression for the UIQM is as follows:

$$\text{UIQM} = c_1 \times \text{UICM} + c_2 \times \text{UISM} + c_3 \times \text{UIConM} \quad (12)$$

where the color, sharpness, and contrast of the image are represented as linear combinations, with respective weighting coefficients of $c_1 = 0.0282$, $c_2 = 0.2953$, and $c_3 = 3.5753$. The weight values utilized in this study are derived from [46].

In order to validate the superiority of our proposed method in terms of computational requirements and real-time performance, we evaluate the efficiency of the model based on the number of model parameters, GFLOPs, and the single image processing time. These metrics provide insights into the computational demands of the algorithm. Notably, the number of parameters serves as a key indicator of model complexity, whereby a higher parameter count indicates greater computational resource and data requirements for model training and execution. GFLOPs, on the other hand, quantify the number of floating-point operations necessary to execute a network model once, thus serving as a measure of computational efficiency and speed. Furthermore, the single image processing time directly signifies the algorithm's efficiency in handling individual images.

3.4. Objective Evaluation

Table 1 presents the average MSE, PSNR, and SSIM values of the enhancement results on the EUVP and UIEBD datasets, showcasing the competitiveness of our method. Our approach, MTUW-GAN, achieves a remarkable performance in underwater image enhance-

ment, as evidenced by achieving the highest scores in MSE, PSNR, and SSIM on the EUVP test set. This outcome underscores the network architecture's capability to restore colors and revive image details in underwater images. Compared to the second-best method, our approach exhibits improvements of 31.40 in the MSE, 1.04 in the PSNR, and 0.06 in the SSIM. Additionally, on the UIEBD test set, our method attains the best PSNR and SSIM scores, and the second-best MSE score. These findings demonstrate that our method excels in handling image details and local contrast in degraded images, while also displaying robust generalization capabilities across a wide range of underwater scene images. In contrast to prevalent methods based on image fusion and multi-scale feature fusion, our proposed approach exhibits a superior performance.

Table 1. Average MSE, PSNR, and SSIM values of the enhancement results on the EUVP and UIEBD datasets. ↑ represents that higher is better, and ↓ represents that lower is better. Red represents the best, while blue represents the second best.

Method	EUVP			UIEBD		
	MSE↓	PSNR↑	SSIM↑	MSE↓	PSNR↑	SSIM↑
UDCP	2319.73	15.90	0.55	4341.56	12.45	0.52
Fusion-based	1068.93	18.58	0.68	583.66	21.63	0.75
UGAN	817.39	20.29	0.71	970.40	21.36	0.73
Water-Net	385.44	23.89	0.78	594.83	22.14	0.75
CWR	263.76	24.99	0.76	782.46	20.71	0.72
MSBDN-DFE	187.84	26.45	0.73	170.40	24.60	0.79
FFA-Net	154.96	27.16	0.75	240.25	24.51	0.80
Ours	123.56	28.20	0.84	205.71	25.88	0.83

Table 2 presents the average UICM, UISM, UIConM, and UIQM results for the enhanced images. Notably, our proposed method consistently achieves the highest UIQM score across the board. Among the three component metrics, UISM and UIConM obtain the highest scores, indicating superior performance in terms of image sharpness and contrast enhancement. Although the UICM score ranks third among the compared methods, it is worth noting that the UICM, being a color metric component, tends to favor images with a higher color saturation. Consequently, certain underwater images may receive inflated color metric scores when processed by fusion-based methods, leading to discrepancies between score judgments and the visual perception of human observers. Nonetheless, our algorithm remains highly competitive in enhancing the clarity and contrast of underwater images, delivering an exceptional performance akin to image-fusion-based approaches.

Table 2. Average UICM, UISM, UIConM, and UIQM results for enhanced images. ↑ represents that higher is better. Red represents the best, while blue represents the second best.

Method	UICM↑	UISM↑	UIConM↑	UIQM↑
UDCP	6.251	5.422	0.046	1.942
Fusion-based	7.853	6.829	0.141	2.745
UGAN	6.522	6.623	0.202	2.863
Water-Net	5.408	6.850	0.254	3.083
CWR	4.911	6.401	0.233	2.862
MSBDN-DFE	4.371	5.715	0.181	2.457
FFA-Net	4.014	5.670	0.214	2.551
Ours	6.314	7.258	0.268	3.278

In addition to evaluating performance, our focus also extends to comparing the efficiency of each algorithm. Table 3 presents these data, indicating that our model exhibits a superior efficiency across multiple metrics. Specifically, MTUW-GAN demonstrates remarkable advantages in terms of the number of parameters, GFLOPs, and the running speed. MTUW-GAN requires only 0.137 M parameters, which is more than 87% less than

the image fusion network Water-Net. In terms of GFLOPs, our method significantly outperforms the second-ranked multiscale feature fusion network, MSBDN-DFF. This observation highlights the considerably lower computational resource requirements of our approach, making it more suitable for deployment on AUVs with a constrained GPU performance and battery capacity. To further assess efficiency, we compared the time required to enhance a single image among the seven methods. The results were averaged using the test data. Our method achieves an impressive average time of 0.069 s for enhancing each underwater image, providing a significant advantage over image fusion and multi-scale feature fusion algorithms. Collectively, the multiple experiments conducted in this section prove that the method proposed in this paper effectively balances performance and real-time processing capabilities.

Table 3. Comparison of the model parameter, GFLOPs, runtime and FPS for different methods. \uparrow represents that higher is better, and \downarrow represents that lower is better. The red value represents the best, while the blue value represents the second best.

Method	# Model Param (M) \downarrow	GFLOPs (G) \downarrow	Times (s) \downarrow	FPS (f/s) \uparrow
UDCP	-	-	1.693	0.591
Fusion-based	-	-	0.442	2.262
UGAN	38.7	18.14	0.104	9.615
Water-Net	1.1	142.9	0.161	6.211
CWR	6.1	42.37	0.136	7.353
MSBDN-DFF	31.4	16.14	0.143	6.993
FFA-Net	4.7	302.7	0.209	4.785
Ours	0.137	0.82	0.069	14.493

3.5. Subjective Evaluation

Figure 4 showcases the visual results obtained from our proposed method and seven comparison methods. To ensure a comprehensive evaluation across diverse underwater scenes, we deliberately selected different types of images for the experimental comparison. The findings reveal distinct characteristics for each method. UDCP demonstrates a dehazing effect but exhibits an overall greenish hue, failing to effectively eliminate significant color biases in underwater images. Fusion-based approaches effectively enhance the image contrast but exhibit tendencies towards over-enhancement and introduce substantial amounts of red noise in certain images. UGAN suffers from localized over-enhancement and inadequate recovery of image details. Water-Net and CWR exhibit varying degrees of color bias when handling low-light underwater scenes, resulting in an uneven color distribution and insufficient preservation of fine textures. Feature-fusion-based methods, such as MSBDN-DFF and FFA-Net, succeed in reducing image blurring caused by underwater scattered light, but lack sufficient color correction capabilities, leading to a subpar restoration of the original colors with an overall blue-green bias. In contrast, our method excels in enhancing underwater images by providing visually clearer texture details of marine plants and coral reefs, more vivid fish colors, and accurate reproduction of divers' skin tones. It consistently demonstrates an excellent performance in restoring underwater image details, mitigating the effects of scattered light and effectively correcting color biases across multiple scenes.

3.6. Ablation Study

To comprehensively validate the roles of each module proposed in this paper within the framework, namely the Channel Distillation (CD) module for feature extraction from the teacher's network middle layer, the UIE loss for perception-based quality recovery of underwater images, and the Assistance Teacher Generator G_T^A for achieving fusion-level performance, ablation study were conducted. Multiple models were trained on the EUVP dataset to facilitate detailed analysis and validation, and the results are presented in Table 4. In index-A, where all three modules were removed, and the proposed UIE loss was replaced

by the image translation loss in pix2pix, the performance was unsatisfactory, with only a PSNR score of 24.40 and an SSIM score of 0.77. In index-B, by incorporating the CD module for knowledge distillation learning, the scores for G_S -enhanced images improved to 24.56 and 0.78, respectively. In index-C, substituting the objective loss function of the teacher network with the UIE loss resulted in a significant boost in PSNR scores by 3.45 and SSIM scores by 0.04. Lastly, in index-3, with the addition of G_T^A for multi-teacher knowledge distillation, there was further improvement, with an increase of 0.19 in PSNR scores and 0.02 in SSIM scores. These ablation study effectively demonstrate the efficacy of the individual modules proposed in this paper.

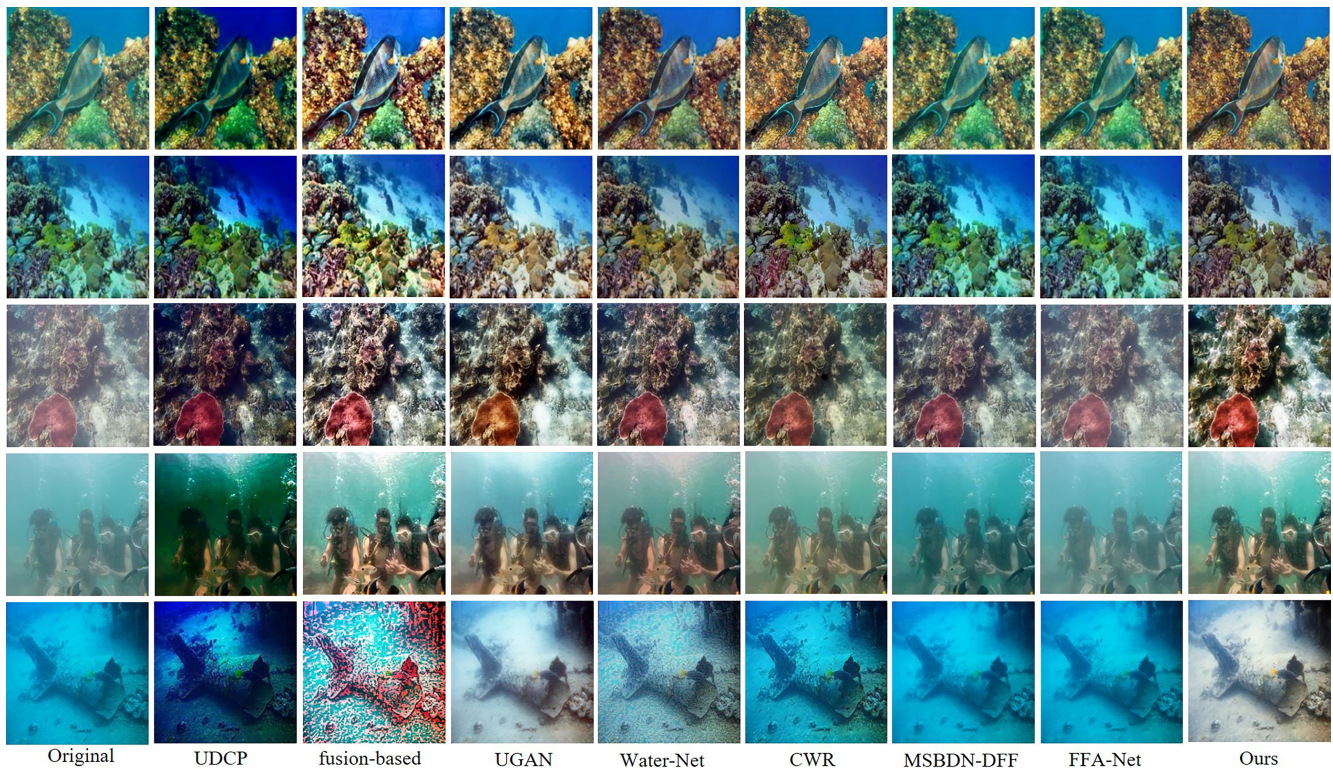


Figure 4. Visual comparison of different types of images. The first and second rows are from the test set of EUVP, and the third, fourth, and fifth rows are from UIEDB.

Table 4. Ablation study involving CD Modules, UIE loss, and G_T^A validity.

Index	CD Module	UIE Loss	G_T^A	PSNR	SSIM
A	-	-	-	24.40	0.77
B	✓	-	-	24.56	0.78
C	✓	✓	-	28.01	0.82
Ours	✓	✓	✓	28.20	0.84

Figure 5 presents a partial example of the ablation study, illustrating the functions of each module. The channel distillation module plays a vital role in restoring detailed image textures and reducing blurring. The UIE loss contributes to color correction of the overall image and effectively mitigates the impact of scattered light. Additionally, the incorporation of G_T^A integrates diverse styles of enhanced images into the student model, enabling it to achieve a performance comparable to image fusion algorithms. This integration further enhances the quality of detailed image features. The ablation study's quantitative and qualitative results are summarized in Table 4 and Figure 5, revealing that the removal of any module diminishes the enhancement of underwater images.

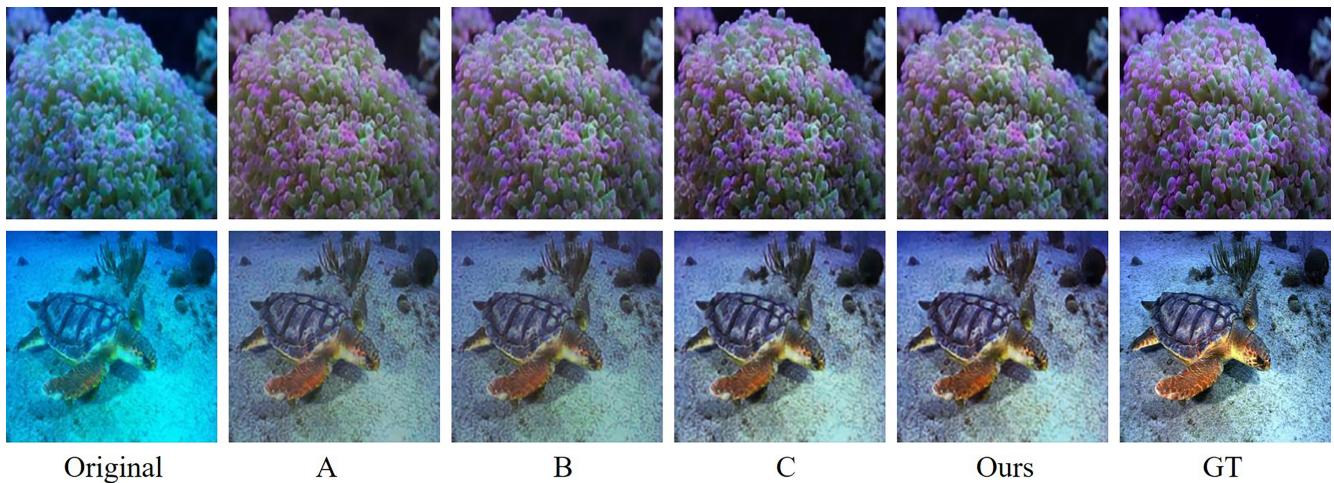


Figure 5. Example of control group results from an ablation study.

4. Conclusions

In this paper, we propose a multi-teacher knowledge distillation GAN for underwater image enhancement (MTUW-GAN). By employing a multi-teacher network to simultaneously instruct a student network, we achieve a comparable performance to image fusion methods while addressing color distortion and detail loss issues in underwater images. The student model of MTUW-GAN acquires rich intermediate-layer feature information from the teacher model through channel distillation, enhancing image color and feature details without increasing the parameters or computational requirements. Additionally, our proposed underwater image enhancement function effectively removes noise and color distortion caused by underwater light scattering, restoring image quality. Compared to typical fusion networks, MTUW-GAN eliminates multiplexed branching and fusion modules, using online knowledge distillation for model compression, significantly reducing the computational requirements for the lightweight student model. Numerous experiments demonstrate that MTUW-GAN achieves a state-of-the-art performance in terms of evaluation metrics and visual quality in underwater image enhancement. It has low computational requirements and real-time performance, making it viable for future applications in AUVs for operational tasks like underwater mineral resource exploration and biological resource investigation. In addition, we plan to file a patent application based on the method presented in this paper.

While MTUW-GAN offers notable advantages, there is still room for improvement. Firstly, the setup of the multi-structured teacher network results in a large number of overall parameters within the framework, potentially leading to slower training. Secondly, the student model in our algorithm struggles to effectively recover certain specialized underwater images. To address these issues, we intend to conduct further research in our future work. Our plan involves exploring several more streamlined and effective teacher networks to optimize the framework's training process. Additionally, we aim to identify the root cause of the problem by utilizing statistics such as image histograms, with the goal of narrowing the gap between the outputs of the student model and the teacher model.

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

Funding: This work was supported by the National Natural Science Foundation of China under grant 52001039.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: In this study, publicly available datasets were analyzed and can be accessed at the following: <https://irvlab.cs.umn.edu/resources/euvp-dataset> (accessed on 27 December 2023), https://li-chongyi.github.io/proj_benchmark.html (accessed on 27 December 2023)

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

References

1. Thompson, F.; Guihen, D. Review of mission planning for autonomous marine vehicle fleets. *J. Field Robot.* **2019**, *36*, 333–354. [[CrossRef](#)]
2. Li, D.; Du, L. Auv trajectory tracking models and control strategies: A review. *J. Mar. Sci. Eng.* **2021**, *9*, 1020. [[CrossRef](#)]
3. Mungekar, R.P.; Jayagowri, R. Color tone determination prior algorithm for depth variant underwater images from AUV's to improve processing time and image quality. *Multimed. Tools Appl.* **2023**, *82*, 31211–31231. [[CrossRef](#)]
4. Hu, K.; Weng, C.; Zhang, Y.; Jin, J.; Xia, Q. An overview of underwater vision enhancement: From traditional methods to recent deep learning. *J. Mar. Sci. Eng.* **2022**, *10*, 241. [[CrossRef](#)]
5. Zhou, J.; Yang, T.; Zhang, W. Underwater vision enhancement technologies: A comprehensive review, challenges, and recent trends. *Appl. Intell.* **2023**, *53*, 3594–3621. [[CrossRef](#)]
6. He, K.; Sun, J.; Tang, X. Single image haze removal using dark channel prior. *IEEE Trans. Pattern Anal. Mach. Intell.* **2010**, *33*, 2341–2353. [[CrossRef](#)] [[PubMed](#)]
7. Drews, P.; Nascimento, E.; Moraes, F.; Botelho, S.; Campos, M. Transmission estimation in underwater single images. In Proceedings of the IEEE International Conference on Computer Vision Workshops, Sydney, NSW, Australia, 2–8 December 2013; pp. 825–830. [[CrossRef](#)]
8. Berman, D.; Levy, D.; Avidan, S.; Treibitz, T. Underwater single image color restoration using haze-lines and a new quantitative dataset. *IEEE Trans. Pattern Anal. Mach. Intell.* **2020**, *43*, 2822–2837. [[CrossRef](#)] [[PubMed](#)]
9. Hummel, R. Image enhancement by histogram transformation. *Comput. Graph. Image Process.* **1977**, *6*, 184–195. [[CrossRef](#)]
10. 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](#)]
11. Ancuti, C.; Ancuti, C.O.; Haber, T.; Bekaert, P. Enhancing underwater images and videos by fusion. In Proceedings of the 2012 IEEE Conference on Computer Vision and Pattern Recognition, Washington, DC, USA, 16–21 June 2012; pp. 81–88. [[CrossRef](#)]
12. Lan, Y.; Cui, Z.; Su, Y.; Wang, N.; Li, A.; Zhang, W.; Li, Q.; Zhong, X. Online knowledge distillation network for single image dehazing. *Sci. Rep.* **2022**, *12*, 14927. [[CrossRef](#)]
13. Ashtiani, F.; Geers, A.J.; Aflatouni, F. An on-chip photonic deep neural network for image classification. *Nature* **2022**, *606*, 501–506. [[CrossRef](#)]
14. Wang, T.; Zhao, L.; Huang, P.; Zhang, X.; Xu, J. Haze concentration adaptive network for image dehazing. *Neurocomputing* **2021**, *439*, 75–85. [[CrossRef](#)]
15. Wang, K.; Hu, Y.; Chen, J.; Wu, X.; Zhao, X.; Li, Y. Underwater image restoration based on a parallel convolutional neural network. *Remote Sens.* **2019**, *11*, 1591. [[CrossRef](#)]
16. Li, C.; Anwar, S.; Hou, J.; Cong, R.; Guo, C.; Ren, W. Underwater image enhancement via medium transmission-guided multi-color space embedding. *IEEE Trans. Image Process.* **2021**, *30*, 4985–5000. [[CrossRef](#)] [[PubMed](#)]
17. Li, C.; Guo, C.; Ren, W.; Cong, R.; Hou, J.; Kwong, S.; Tao, D. An underwater image enhancement benchmark dataset and beyond. *IEEE Trans. Image Process.* **2019**, *29*, 4376–4389. [[CrossRef](#)] [[PubMed](#)]
18. Fabbri, C.; Islam, M.J.; Sattar, J. Enhancing underwater imagery using generative adversarial networks. In Proceedings of the 2018 IEEE International Conference on Robotics and Automation (ICRA), Brisbane, QLD, Australia, 21–25 May 2018; pp. 7159–7165. [[CrossRef](#)]
19. Islam, M.J.; Xia, Y.; Sattar, J. Fast underwater image enhancement for improved visual perception. *IEEE Robot. Autom. Lett.* **2020**, *5*, 3227–3234. [[CrossRef](#)]
20. Han, J.; Shoeiby, M.; Malthus, T.; Botha, E.; Anstee, J.; Anwar, S.; Wei, R.; Armin, M.A.; Li, H.; Petersson, L. Underwater image restoration via contrastive learning and a real-world dataset. *Remote Sens.* **2022**, *14*, 4297. [[CrossRef](#)]
21. 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](#)]
22. Guo, Y.; Li, H.; Zhuang, P. Underwater image enhancement using a multiscale dense generative adversarial network. *IEEE J. Ocean. Eng.* **2019**, *45*, 862–870. [[CrossRef](#)]
23. Tian, Y.; Xu, Y.; Zhou, J. Underwater Image Enhancement Method Based on Feature Fusion Neural Network. *IEEE Access* **2022**, *10*, 107536–107548. [[CrossRef](#)]
24. Liu, X.; Lin, S.; Tao, Z. Learning multiscale pipeline gated fusion for underwater image enhancement. *Multimed. Tools Appl.* **2023**, *82*, 32281–32304. [[CrossRef](#)]
25. Wang, L.; Yoon, K.J. Knowledge distillation and student-teacher learning for visual intelligence: A review and new outlooks. *IEEE Trans. Pattern Anal. Mach. Intell.* **2021**, *44*, 3048–3068. [[CrossRef](#)] [[PubMed](#)]

26. Guo, Q.; Wang, X.; Wu, Y.; Yu, Z.; Liang, D.; Hu, X.; Luo, P. Online knowledge distillation via collaborative learning. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 13–19 June 2020; pp. 11020–11029.
27. Kim, J.; Bhalgat, Y.; Lee, J.; Patel, C.; Kwak, N. Qkd: Quantization-aware knowledge distillation. *arXiv* **2019**, arXiv:1911.12491. [[CrossRef](#)]
28. Howard, A.G.; Zhu, M.; Chen, B.; Kalenichenko, D.; Wang, W.; Weyand, T.; Andreetto, M.; Adam, H. Mobilenets: Efficient convolutional neural networks for mobile vision applications. *arXiv* **2017**, arXiv:1704.04861. [[CrossRef](#)]
29. Li, M.; Lin, J.; Ding, Y.; Liu, Z.; Zhu, J.Y.; Han, S. Gan compression: Efficient architectures for interactive conditional gans. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 13–19 June 2020; pp. 5284–5294. [[CrossRef](#)]
30. 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. [[CrossRef](#)]
31. He, Y.; Zhang, X.; Sun, J. Channel pruning for accelerating very deep neural networks. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017; pp. 1389–1397. [[CrossRef](#)]
32. Guo, J.; Zhang, W.; Ouyang, W.; Xu, D. Model compression using progressive channel pruning. *IEEE Trans. Circuits Syst. Video Technol.* **2020**, *31*, 1114–1124. [[CrossRef](#)]
33. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 26 June–1 July 2016; pp. 770–778. [[CrossRef](#)]
34. Huang, G.; Liu, Z.; Van Der Maaten, L.; Weinberger, K.Q. Densely connected convolutional networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 4700–4708. [[CrossRef](#)]
35. Zhou, Z.; Zhuge, C.; Guan, X.; Liu, W. Channel distillation: Channel-wise attention for knowledge distillation. *arXiv* **2020**, arXiv:2006.01683. [[CrossRef](#)]
36. Hu, J.; Shen, L.; Sun, G. Squeeze-and-excitation networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 7132–7141. [[CrossRef](#)]
37. Isola, P.; Zhu, J.Y.; Zhou, T.; Efros, A.A. Image-to-image translation with conditional adversarial networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 1125–1134. [[CrossRef](#)]
38. 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: Berlin/Heidelberg, Germany, 2019; pp. 66–75. [[CrossRef](#)]
39. Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. *arXiv* **2014**, arXiv:1409.1556. [[CrossRef](#)]
40. Wang, Z.; Bovik, A.C.; Sheikh, H.R.; Simoncelli, E.P. Image quality assessment: From error visibility to structural similarity. *IEEE Trans. Image Process.* **2004**, *13*, 600–612. [[CrossRef](#)]
41. Johnson, J.; Alahi, A.; Li, F.-F. Perceptual losses for real-time style transfer and super-resolution. In Proceedings of the Computer Vision—ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, 11–14 October 2016; Springer: Berlin/Heidelberg, Germany, 2016; pp. 694–711. [[CrossRef](#)]
42. Rudin, L.I.; Osher, S.; Fatemi, E. Nonlinear total variation based noise removal algorithms. *Phys. D* **1992**, *60*, 259–268. [[CrossRef](#)]
43. Dong, H.; Pan, J.; Xiang, L.; Hu, Z.; Zhang, X.; Wang, F.; Yang, M.H. Multi-scale boosted dehazing network with dense feature fusion. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans, LA, USA, 18–24 June 2020; pp. 2157–2167. [[CrossRef](#)]
44. Qin, X.; Wang, Z.; Bai, Y.; Xie, X.; Jia, H. FFA-Net: Feature fusion attention network for single image dehazing. In Proceedings of the AAAI Conference on Artificial Intelligence, New York, NY, USA, 7–12 February 2020; Volume 34, pp. 11908–11915. [[CrossRef](#)]
45. Sara, U.; Akter, M.; Uddin, M.S. Image quality assessment through FSIM, SSIM, MSE and PSNR—A comparative study. *J. Comput. Commun.* **2019**, *7*, 8–18. [[CrossRef](#)]
46. Panetta, K.; Gao, C.; Agaian, S. Human-visual-system-inspired underwater image quality measures. *IEEE J. Ocean. Eng.* **2015**, *41*, 541–551. [[CrossRef](#)]

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.