



# Article LPS-Net: Lightweight Parallel Strategy Network for Underwater Image Enhancement

Jingxia Jiang<sup>1</sup>, Peiyun Huang<sup>1</sup>, Lihan Tong<sup>1</sup>, Junjie Yin<sup>1</sup> and Erkang Chen<sup>1,2,\*</sup>

- <sup>1</sup> School of Ocean Information Engineering, Jimei University, Xiamen 361021, China; 202021114006@jmu.edu.cn (J.J.); 202021114008@jmu.edu.cn (P.H.); 202121301035@jmu.edu.cn (L.T.); 202021112006@jmu.edu.cn (J.Y.)
- <sup>2</sup> Fujian Provincial Key Laboratory of Oceanic Information Perception and Intelligent Processing, Xiamen 361021, China
- \* Correspondence: ekchen@jmu.edu.cn

Abstract: Underwater images are frequently subject to color distortion and loss of details. However, previous enhancement methods did not tackle these mixed degradations by dividing them into subproblems that could be effectively addressed. Moreover, the parameters and computations required for these methods are usually costly for underwater equipment, which has limited power supply, processing capabilities, and memory capacity. To address these challenges, this work proposes a Lightweight Parallel Strategy Network (LPS-Net). Firstly, a Dual-Attention Enhancement Block and a Mirror Large Receptiveness Block are introduced to, respectively, enhance the color and restore details in degraded images. Secondly, we employed these blocks on parallel branches at each stage of LPS-Net, with the goal of achieving effective image color and detail rendering simultaneously. Thirdly, a Gated Fusion Unit is proposed to merge features from different branches at each stage. Finally, the network utilizes four stages of parallel enhancement, achieving a balanced trade-off between performance and parameters. Extensive experiments demonstrated that LPS-Net achieves optimal color enhancement and superior detail restoration in terms of visual quality. Furthermore, it attains state-of-the-art underwater image enhancement performance on the evaluation metrics, while using only 80.12 k parameters.

**Keywords:** Dual-Attention Enhancement Block; Mirror Large Receptiveness Block; underwater image enhancement

# 1. Introduction

Underwater images have broad application value in the marine field, playing an essential role in marine scientific research [1,2] and protection and environmental monitoring [3]. However, underwater image quality is typically limited by the absorption, scattering, and attenuation of light, resulting in the loss of contrast, blurring, and color distortion. With the rapid development of computer vision, many scholars have conducted extensive research on Underwater Image Enhancement (UIE) and have made significant advancements in the field [4–6].

According to the Jaffe–McGlamery imaging model [7,8], underwater imaging consists of a linear superposition of direct, back-scattered, and forward-scattered components. In general, the effects of forward scattering are negligible; thus, the imaging model can be simplified as:

$$\mathbf{I}_{\mathbf{c}}(\mathbf{x}) = \mathbf{J}_{\mathbf{c}}(\mathbf{x})t_{\mathbf{c}}(\mathbf{x}) + \mathbf{B}_{\mathbf{c}}(1 - t_{\mathbf{c}}(\mathbf{x})), \mathbf{c} \in \{R, G, B\}$$
(1)

where  $\mathbf{I}_{c}(\mathbf{x})$  is the observed intensity in the color channel *c* of the input image at the pixel **x**,  $\mathbf{J}_{c}(\mathbf{x})$  represents the radiation of the scene at **x**,  $\mathbf{B}_{c}$  represents the background light,  $t_{c}(\mathbf{x})$  is the transmission map, and *c* represents the red, green, and blue channels. UIE aims to recover the radiation  $\mathbf{J}(x)$  of the scene from the observed image  $\mathbf{I}(x)$ , which is an ill-posed



Citation: Jiang, J.; Huang, P.; Tong, L.; Yin, J.; Chen, E. LPS-Net:Lightweight Parallel Strategy Network for Underwater Image Enhancement. *Appl. Sci.* 2023, *13*, 9419. https:// doi.org/10.3390/app13169419

Academic Editor: Atsushi Mase

Received: 14 June 2023 Revised: 8 August 2023 Accepted: 17 August 2023 Published: 19 August 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/).

2 of 22

problem. Thus, in one line of traditional UIE methods, physical priors [9,10] were utilized to estimate the unknown transmission map and background light. However, these physical priors might not always hold true in complex underwater scenes, resulting in poor performance when physical priors are violated. Another line of traditional methods [11–13] directly modifies image pixel values to improve visual quality, regardless of the physical model. These methods usually rely on hand-crafted features, thus exhibiting poor generalization ability. In recent years, UIE methods based on Convolutional Neural Networks (CNNs) have achieved promising results [14–18]. For example, Li et al. [14] presented an underwater image enhancement network via medium transmission-guided multicolor space embedding, called Ucolor. Huo et al. [15] used wavelet-enhanced learning units to build a UIE framework. However, these methods usually require complex module designs in order to remove mixed degradations directly, resulting in a large number of parameters. Furthermore, these methods often have high computational complexity. For example, it takes the Ucolor [14] method 443.85 GFLOPs to process a 720P image. The considerable number of computational operations and parameters associated with these methods pose challenges for their deployment in underwater equipment with limited power supply, processing capabilities, and memory capacity. Therefore, designing a lightweight and effective underwater image enhancement model has become a challenge.

To address the above issues, we propose a Lightweight Parallel Strategy Network (LPS-Net) for underwater image enhancement. As shown in Figure 1, LPS-Net achieves state-of-the-art underwater image enhancement performance with very few parameters. Specifically, a Dual-Attention Enhancement Block (DAEB) and a Mirror Large Receptiveness Block (MLB) are proposed to, respectively, enhance color and restore detail in degraded images. Problem splitting can avoid the use of complex modules, as well as reduce the amount of computation and parameters. With the aim of achieving remarkable image color and detail rendering, we repeatedly employed these blocks on parallel branches at each stage of LPS-Net and designed a Gated Fusion Unit (GFU) to combine features from different branches.



**Figure 1. Left:** Compared with existing state-of-the-art methods, our method achieves the best performance on the T90 dataset using only 80.12 k parameters (1/35 of that of MFEF [19], 1/80 of the parameter number of PRW-Net [15], and 1/2000 of that of Ucolor [14]). The size of the stars represents the number of trainable parameters. **Right:** A sample image from the T90 dataset and its enhanced results produced by different methods including our LPS-Net.

In water, different wavelengths of light have varying rates of attenuation, with red light attenuating the fastest and blue–green light attenuating the slowest. As a result, there

are different degrees of degradation among the three channels of RGB images, but previous methods [20–22] have not adequately addressed the hidden visual hazards caused by such differences. To deal with such degradation differences in the R, G, and B channels, the network should capture the distribution of color features across different channels. To this end, the incoming features are divided into three bands in the proposed DAEB. Each band passes through an Instance Normalization (IN) module, which will reduce the differences among different bands. Equipped with subsequent Channel Attention (CA) [23] and Spatial Attention (SA) [24], DAEB is bestowed with the ability to enhance the color of different color channels adaptively and separately.

To restore image details, a large receptive field in the UIE network is desired, since increasing the receptive field allows for more contextual information to be obtained. Inspired by the long-range modeling ability of ViTs [25,26], we adopted large-kernel convolutions to expand the receptive field for improving model performance, as exemplified by the remarkable ConvNet [27], which employs  $7 \times 7$  depthwise convolutions. Specifically, the proposed MLB uses four large-kernel convolutions with different kernel sizes. We conduct a series of ablation studies in Section 5.2 and empirically set the kernel sizes in the MLB as  $7 \times 7$  and  $9 \times 9$ , respectively. Additionally, we constructed parallel branches using the DAEB and MLB as primary components, then applied the GFU to integrate features from different branches. Extensive experiments demonstrated that LPS-Net achieves state-of-the-art UIE performance while using only 80.12 k parameters.

Our main contributions in this paper are summarized as follows:

- We propose the DAEB to capture the distribution of color features across the R, G, and B channels. IN can reduce the difference between channels, and by employing CA and SA, the network is able to enhance the color of degraded images at different levels.
- We introduce a novel block named the MLB to focus on texture information in degraded images. With the large receptive field of the mirror-designed convolutional kernel, we were able to obtain more contextual information to restore fine details.
- Our LPS-Net needs only 80.12 k parameters, making it well-suited for underwater devices with limited memory capacity. Extensive experiments demonstrated that LPS-Net achieves state-of-the-art UIE performance in terms of visual quality and evaluative metrics.

The remaining sections of this paper are organized as follows: Section 2 reviews the traditional and deep-learning-based UIE methods. Subsequently, in Section 3, we introduce the proposed LPS-Net. Section 4 presents comprehensive qualitative and quantitative evaluations, aiming at comparing the efficacy of the proposed method against state-of-the-art approaches. Section 5 provides ablation studies of each module within the network, while Section 6 sheds light on the identified limitations and drawbacks. Finally, Section 7 encompasses a concise summary of our work as well as expounds upon future avenues of research and broader implications.

# 2. Related Works

In recent years, there has been a growing interest in underwater image enhancement [28–31]. In general, there are mainly two types of methods: traditional methods [9,11–13,19,32–36] and deep-learning-based methods [14–18,37–40].

#### 2.1. Traditional Methods

The traditional methods employed for enhancing underwater images, such as histogram equalization [11], wavelet transforms [12], and the Retinex algorithm [13], exhibit certain drawbacks. AbuNase et al. [32] utilized Particle Swarm Optimization (PSO) to mitigate the effects of light absorption and scattering, but this approach may be susceptible to noise and demands meticulous parameter tuning. Ghani et al. [33] proposed a Rayleigh distribution histogram stretching method that enhances contrast; however, it can lead to excessive enhancement and loss of image details. Fu et al. [34] presented a variational Retinex-based method for underwater image enhancement, which may result in oversmoothing and the loss of fine details. Zhuang et al. [35] developed a Bayesian Retinex method incorporating multi-order gradient priors, but its iterative optimization process can

method incorporating multi-order gradient priors, but its iterative optimization process can impose a significant computational burden. The UDCP algorithm [9], based on the dark channel prior, offers simplicity in implementation; however, it may yield inaccurate results in the presence of complex underwater color interference. Furthermore, Zhang et al. [36] introduced MLLE, an efficient and robust approach, but it may still exhibit limitations in handling certain challenging underwater scenarios. The depth estimation model proposed by Song et al. [10] based on the Underwater Light Attenuation Prior (ULAP) suffers from performance degradation in complex underwater environments due to the presence of color interference factors. These conventional methods often lack the flexibility required for implementation, necessitating accurate parameter estimation and valid underwater optical properties to achieve satisfactory results. Additionally, the computational complexity and multi-step processing of these algorithms can result in significant time overhead.

# 2.2. Deep-Learning-Based Methods

With the successful applications of deep learning in some high-level computer vision tasks (such as image classification [41] and recognition [42]), more and more researchers began to explore low-level computer vision tasks such as image enhancement, which has achieved outstanding results in terms of performance indicators and visual effects.

Li et al. [5] proposed the WaterNet model, which employed adaptive filters and deep convolutional neural networks for image enhancement and noise reduction. However, this method did not incorporate dedicated modules to address the color shift and texture loss commonly observed in degraded images. Jiang et al. [37] designed a domain adaptation framework based on transfer learning to transform aerial image deblurring into realistic underwater image enhancement. While it achieved notable performance indicators, it lacks dedicated modules to handle color shifts and texture loss. Naik et al. [38] introduced Shallownet, a lightweight underwater enhancement framework that sacrifices accuracy for reduced model size and faster processing. Although it offers fast processing, the reduced accuracy may limit its applicability. Islam [43] presented FUNIE-GAN, a conditional generative-adversarial-network-based model for real-time underwater image enhancement. Despite its excellent visual effects, it suffers from extensive network parameters and high computational requirements, which make it unsuitable for existing underwater devices. Huo et al. [15] utilized wavelet-enhanced learning units to decompose hierarchical features into high-frequency and low-frequency components, enhancing them through normalization and attention mechanisms. However, it does not explicitly address the color shift and texture loss associated with degraded images. Li et al. [14] proposed Ucolor, which incorporates medium transmission-guided multicolor space embedding to enhance the network's response towards quality-degraded regions. Despite its excellent visual effects, it exhibits extensive network parameters and high computational complexity, rendering it unsuitable for existing underwater devices. Wei et al. [18] introduced UHD-SFNet, an efficient two-path model that recovers color and texture in underwater blurred images through frequency and spatial domain processing. However, the specific limitations or drawbacks of this method were not explicitly mentioned. Fu et al. [16] addressed underwater image enhancement by estimating the distribution and employing a consensus process using a novel probabilistic network. While it provides a unique approach, specific limitations or drawbacks were not explicitly discussed. A Hierarchical Attention Aggregation with Multi-resolution feature learning for GAN-based underwater image enhancement (HAAM-GAN) [40] was proposed to tackle color bias, underexposure, and blurring in underwater images.

# 3. Proposed Method

The overall structure of the proposed LPS-Net is constituted by four recurring enhancement stages, as shown in Figure 2. At each stage, a pair of intricately designed components, i.e., the Dual-Attention Enhancement Block and the Mirror Large Receptiveness Block, is utilized to, respectively, augment color and restore image details. Specifically, the incoming features are fed into these two blocks on parallel branches. Then, at the end of each stage, the Gated Fusion Unit adaptively merges features from the DAEB and MLB, forming a unified representation. Throughout this procedure, the DAEB alleviates color distortion resulting from different degrees of light attenuation in different color channels. Concurrently, the MLB, with its large field of view, gradually reinstates image details at every stage, leveraging its ability to garner more contextual data to correct Gaussian blur caused by light scattering and other texture distortions. Feature maps from these two branches are then fused by the GFU in order to achieve effective image color and detail rendering simultaneously.



**Figure 2.** The overall structure of LPS-Net. The degraded image is fed into an end-to-end network that goes through a four-stage enhancement and correction process. The  $3 \times 3$  convolution on the branch in the top left corner is used for channel expansion.

## 3.1. Dual-Attention Enhancement Block

Since the absorption coefficients of different channels in RGB images are different, the distribution of pixels from the R, G, and B channels of images always appears unbalanced, resulting in serious color deviations. Previous methods did not pay attention to the hidden visual dangers caused by the differences among the three channels. To address this issue, we present the DAEB to better capture the distribution of color features in the R, G, and B channels. As shown in Figure 2b, the block has three small branches, and weights are not shared between branches. Specifically, in different branches, we used IN to reduce the differences among different levels. Then, the augmented results are summed. This block can accurately weigh color channels and extract color feature information. Since the attenuation of the red channel is larger than that of the blue and green channels, the DAEB is applied to assign different learning weights to different color channels, which compensates for the feature map of the red channel and improves the overall color features.

#### 3.2. Mirror Large Receptiveness Block

Inspired by ConvNet [27], we found that, when four deep convolutional layers in series are replaced by a parallel connection of deep convolutional layers with the same configuration in preliminary experiments, superior results can be achieved. Under certain constraints, the receptive field's size increases with the convolutional kernel's increase. A large receptive field can capture more contextual information. Therefore, we kept the parallel settings unchanged and selected the optimal configuration by changing the sizes

of the kernels. Our four parallel large-kernel convolutions are centrally symmetrical, so it is called the Mirror Large Receptiveness Block, shown in Figure 2c. Detailed ablation experiments showed that the network's performance is superior when the convolutional kernel sizes are set as  $7 \times 7$  and  $9 \times 9$ , respectively.

Given the feature  $F \in R^{C \times H \times W}$ , the key operations of the MLB can be expressed as follows: F' = Chunk(Control - 1) i = 1, 2, 3, 4

$$F_{i} = Chunk(Conv_{1\times1}(F), \dim = 1), i = 1, 2, 3, 4$$

$$F_{out1} = Conv_{7\times7}(F_{1}')$$

$$F_{out2} = Conv_{9\times9}(F_{2}')$$

$$F_{out3} = Conv_{9\times9}(F_{3}')$$

$$F_{out4} = Conv_{7\times7}(F_{4}')$$
(2)

where  $Chunk(\cdot)$  means the cutting operation in dimension one.  $Conv_{1\times 1}(\cdot)$ ,  $Conv_{7\times 7}(\cdot)$ ,  $Conv_{9\times 9}(\cdot)$  refer to the convolution with kernel sizes of  $1 \times 1, 7 \times 7$ , and  $9 \times 9$ , respectively.

#### 3.3. Gated Fusion Unit

In CNNs, the gating mechanism is usually implemented by some particular neurons. These neurons can decide whether information can pass through them. If gating neurons fire less, they can block the flow of information, preventing some information from being passed on to the next layer. Inspired by the Gated Linear Unit (GLU) [44] and self-attention [45], the GFU was designed for underwater enhancement, as shown in Figure 2d. The Gated Fusion Unit further extracts the input color information and texture features through a  $3 \times 3$  convolutional layer and depthwise separable convolution. Next, it fuses the weighted features with the original features and controls the weights of different feature through a gating mechanism. Finally, the Gated Fusion Unit takes the processed feature vector as the output. It is worth mentioning that the function of the Gated Fusion Unit is to enhance the quality and clarity of the feature maps at different stages and to filter the noise and irrelevant information so that the details of helpful information can be highlighted.

### 3.4. Loss Function

We introduce the Charbonnier Loss [46] as our basic reconstruction loss:

$$\mathcal{L}_{rec} = \mathcal{L}_c(\mathcal{I}(X), \mathcal{J}_{gt}) \tag{3}$$

where  $\mathcal{I}$  is our LPS-Net and X and  $\mathcal{J}_{gt}$  stand for the input and ground-truth.  $\mathcal{L}_c$  denotes the Charbonnier Loss, which can be express as:

$$\mathcal{L}_{c} = \frac{1}{N} \sum_{i=1}^{N} \sqrt{\|X^{i} - Y^{i}\|^{2} + \epsilon^{2}}$$
(4)

where constant  $\epsilon$  was empirically set to  $1 \times 10^{-3}$  for all experiments. In addition, the perceptual level of the restored image is also critical. We applied the perceptual loss to improve the restoration performance. The perceptual loss can be formulated as follows:

$$\mathcal{L}_{perceptual} = \sum_{j=1}^{2} \frac{1}{C_j H_j W_j} ||\phi_j(\mathcal{I}(x)) - \phi_j(\mathcal{Y})||_1$$
(5)

where  $\phi_j$  represents the 1-st and the 3-rd layers of VGG19 [47].  $C_j$ ,  $H_j$ , and  $W_j$  represent the channel number, height, and width of the feature map, respectively.

In order to better reflect the human visual system's perception of image quality, we adopted the negative SSIM loss to focus on luminance, contrast, and structure. The negative SSIM loss is:

$$\mathcal{L}_{ssim} = -SSIM(\mathcal{I}(X), \mathcal{J}_{gt})$$
(6)

where  $\mathcal{I}(\cdot)$  is LPS-Net and *X* and  $\mathcal{J}_{gt}$  stand for the input and the corresponding ground-truth, respectively.

The overall loss function can be expressed as:

$$\mathcal{L} = \lambda_1 \mathcal{L}_c + \lambda_2 \mathcal{L}_{percentual} + \lambda_3 \mathcal{L}_{ssim} \tag{7}$$

where the  $\lambda_1$ ,  $\lambda_2$  and  $\lambda_3$  are set to 1, 0.2 and 0.5, respectively.

#### 4. Experiments and Analysis

This section firstly introduces the datasets used in the experiments: Underwater Image Enhancement Benchmark (UIEB) dataset [5], non-reference underwater image dataset U45 [48], and Enhancement of Underwater Visual Perception (EUVP) dataset [43]. Secondly, we describe the experimental setup and implementation details. To demonstrate the feasibility and superiority of the proposed method, we compared LPS-Net with traditional methods [9,36,49] and state-of-the-art deep learning underwater methods [5,14–19,22,38,40,43,50]. Finally, we conducted an analysis of the objective data presented in the tables and provide detailed visualizations to demonstrate the advanced and effective nature of our method.

# 4.1. Datasets

The UIEB dataset contains 890 high-resolution raw underwater images, corresponding high-quality reference images, and 60 challenge images (C60) for which no corresponding reference images were obtained. UIEB covers a variety of underwater scenes, and the image content covers a wide range of areas, such as marine life, divers, submarine corals, and coral reefs. The quality of underwater images is significantly lower. The creators of the dataset devised a method for generating high-quality reference images. Firstly, they applied twelve popular image-enhancement methods to generate enhanced results. Then, fifty volunteers were recruited to evaluate the quality of the enhanced results. The reference image for each original underwater image was determined by majority voting based on pairwise comparisons.

The U45 dataset consists of 45 degraded small images without reference images. Considering that the underwater enhancement task does not have publicly available datasets for a test like the single-image super-resolution task, Li et al. carefully selected 45 authentic underwater images, named U45. It is divided into three subsets of green, blue, and haze, where the subsets correspond to the color cast of underwater degradation, low contrast, and blur effects.

The EUVP dataset contains separate sets of paired and unpaired image samples of poor and good perceptual quality. The creators employed seven distinct camera models, including multiple GoPros, uEye cameras integrated into the Aqua AUV, low-light USB cameras, and high-definition cameras mounted on the Trident ROV, to capture the image data. The data acquisition occurred at diverse locations and under various conditions during marine exploration and robot navigation. In our experiments, two paired sets were used: Underwater Dark (5550 pairs) and Underwater ImageNet (3700 pairs).

#### 4.2. Experimental Setup and Implementation Details

# 4.2.1. Experiment Details

All experiments were implemented under the Pytorch [51] framework and accelerated on the NVIDIA RTX A100GPU (40GB). During training, the training epochs were set to 400, and the batch size was 30. We used the ADAM optimizer as the optimization algorithm. The learning rate was set to  $4 \times 10^{-4}$  first, and the default values of  $\beta_1$  and  $\beta_2$  were 0.5 and 0.999, respectively. We used CyclicLR to adjust the learning rate, with an initial momentum of 0.9 and 0.999. Data augmentation included horizontal flipping, random cropping, and randomly rotating the image to 90, 180, and 270°.

Throughout the training process, we standardized the resolution of both the input and output to  $256 \times 256$ . From the UIEB dataset, we randomly selected 800 pairs of original images and their corresponding clear images to compose the training set for our model.

During the training validation phase, we employed the remaining 90 images from the UIEB dataset (referred to as T90) to evaluate our method's performance on degraded images. For the formal testing stage, we continued to use the T90 dataset; however, in contrast to the training phase, we refrained from making any alterations to the format and size of the images, aiming to restore the model's capability to enhance real underwater images. Furthermore, to assess the generalization capacity, we conducted evaluations using the C60 and U45 datasets with various methods. To assess the generalization capability of our model, some comparative experiments were conducted on the EUVP dataset. We used the ImageNet sub-dataset under the EUVP dataset to contain 3700 pairs of images for training and randomly sampled 1000 pairs of images from the Dark sub-dataset for testing.

#### 4.2.2. Evaluation Metrics

For quantitative evaluations, several objective evaluation metrics were employed to assess the performance of the proposed underwater image enhancement method. These metrics served as references for measuring image quality in a comprehensive manner.

The first metric used was the Peak-Signal-to-Noise Ratio (PSNR) [52], which is a wellestablished full-reference image quality evaluation metric. It quantifies the errors between corresponding pixels in the enhanced image and the reference image. A higher PSNR score indicates better image quality in terms of minimizing pixelwise errors.

To evaluate the visual quality of the enhanced images, the Structural Similarity Index (SSIM) [53] was employed. The SSIM measures the similarity between the enhanced image and the reference image in terms of three key features: brightness, contrast, and structure. A higher SSIM value implies a higher perceptual similarity between the enhanced and reference images.

The Mean-Squared Error (MSE) [54] was another metric used, which calculates the average squared difference between the pixels of the enhanced image and the reference image. A lower MSE value indicates better image quality, as it quantifies the overall distortion in the enhanced image compared to the reference image.

For a comprehensive evaluation of underwater image quality, two additional metrics were employed. The Underwater Color Image Quality Evaluation (UCIQE) [55] metric primarily measures the degree of detail and color recovery in distorted images. It is considered one of the most-comprehensive image-evaluation standards for underwater images. The Underwater Image Quality Metric (UIQM) [56] is specifically designed to assess color, sharpness, and contrast in underwater images. By considering both local and global image features, the UIQM offers a holistic evaluation of underwater image quality.

By utilizing these objective evaluation metrics, the proposed method aims to provide a quantitative assessment of its performance in terms of fidelity, perceptual similarity, color reproduction, detail recovery, and overall image quality. These metrics serve as valuable references to measure and compare the effectiveness of the proposed underwater image enhancement technique.

#### 4.3. Compared Methods

We compared LPS-Net with state-of-the-art methods, including traditional methods and deep learning methods. Traditional methods included UDCP [9], IBLA [49], and MLLE [36], and deep learning methods included UWCNN [22], Water-Net [5], PRW-Net [15], Shallownet [38], Ucolor [14], UIEC^2-Net [50], PUIE-Net [16], FUNIE-GAN [43], UHD-SFNet [18], HAAM-GAN [40], MFEF [19], and the latest NU2Net [17] for underwater image enhancement.

#### 4.4. Quantitative Comparisons

We firstly performed a quantitative comparison on 90 images from T90, 60 images from C60, and 45 images from U45. Table 1 gives the quantitative results of different methods. The PSNR, SSIM, and MSE were used as full-reference evaluation metrics and

provided only the T90 test set, since T90 has ground-truth images. UCIQE and UIQM were used as non-reference evaluation metrics and provided for all test sets.

**Table 1.** Experimental results on the T90 [5], C60 [5], and U45 [48] datasets. The training set of UIEB was used for training. The best results are <u>underlined</u>. The second-best results are in **bold**.  $\uparrow$  represents that higher is better, and  $\downarrow$  represents that lower is better.

	<b>T90</b> [5]					C60 [5]		U45 [48]	
Method	PSNR↑	<b>SSIM</b> ↑	MSE↓	UCIQE↑	UIQM↑	UCIQE↑	UIQM↑	UCIQE↑	UIQM↑
UDCP(ICCVW'13) [9]	13.415	0.749	0.228	0.572	2.755	0.560	1.859	0.574	2.275
IBLA(TIP'17) [49]	18.054	0.808	0.142	0.582	2.557	0.584	1.662	0.565	2.387
WaterNet(TIP'19) [5]	16.305	0.797	0.161	0.564	2.916	0.550	2.113	0.576	2.957
UWCNN(PR'20) [22]	17.949	0.847	0.221	0.517	3.011	0.492	2.222	0.527	3.063
FUNIE-GAN(RAL'20) [43]	17.114	0.701	0.216	0.564	3.092	0.556	2.867	0.545	2.495
PRW-Net(ICCVW'21) [15]	20.787	0.823	0.099	0.603	3.062	0.572	2.717	0.621	3.026
Shallownet(AAAI'21) [38]	18.278	0.855	0.131	0.544	2.942	0.521	2.212	0.545	3.109
Ucolor(TIP'21) [14]	21.093	0.872	0.096	0.555	3.049	0.530	2.167	0.554	3.148
UIEC^2-Net(SPIC'21) [50]	22.958	0.907	0.078	0.599	2.999	0.580	2.228	0.604	3.125
MLLE(TIP'22) [36]	19.561	0.845	0.115	0.592	2.624	0.581	1.977	0.597	2.454
PUIE-Net(ECCV'22) [16]	21.382	0.882	0.093	0.566	3.021	0.543	2.155	0.563	<u>3.192</u>
UHD-SFNet(ACCV'22) [18]	18.877	0.810	0.144	0.559	2.551	0.528	1.741	0.585	2.826
NU2Net(AAAI'23,Oral) [17]	22.419	<u>0.923</u>	0.086	0.587	2.936	0.555	2.222	0.593	3.185
MFEF(EAAI'23) [19]	23.352	0.910	0.076	0.602	1.333	-	-	-	-
HAAM-GAN(EAAI'23) [40]	22.947	0.889	0.081	0.602	<u>3.164</u>	0.569	<u>2.867</u>	0.606	3.161
Ours	23.905	0.915	0.071	0.620	2.833	0.585	2.242	0.625	3.142

From Table 1, it can be seen that our method achieved the best results in terms of the PSNR and MSE metrics on the T90 test set and the best results in terms of the UCIQE metric on all three test sets. Specifically, on the T90 test set with reference images, the proposed LPS-Net scored the best PSNR and MSE metrics and obtained the second-best SSIM metric, proving that the proposed network has achieved state-of-the-art UIE performance, and its network architecture is able to restore color and recover image details. Compared with the recently developed MFEF method, which roughly ranked second-best, LPS-Net led by 0.553 dB, 0.005, and 0.005 in terms of the PSNR, SSIM, and MSE.

On the other hand, higher UCIQE and UIQM values indicate better performance in terms of sharpness, color correction ability, and contrast in images. Among the evaluated methods, LPS-Net achieved the highest UCIQE values, demonstrating its efficacy in mitigating color shifts and enhancing sharpness. Although FUNIE-GAN and HAAM-GAN slightly outperformed our method in terms of the UIQM, a qualitative comparison revealed that GAN-based approaches suffered from localized exposure issues, leading to significant image distortions. Conversely, the enhanced results obtained by LPS-Net effectively addressed color biases without introducing additional degradation problems.

Apart from enhancement performance, efficiency is also very important for UIE methods. Table 2 provides comparison results in terms of various efficiency metrics, including the number of computational operations measured in GFLOPs, the number of parameters, and runtime speed. It can be seen that LPS-Net needed only 0.08 M (80.12 k) parameters, substantially fewer than other methods, which is beneficial for deployment in equipment with limited memory capacity. In terms of GFLOPs, LPS-Net ranked the second-best, after UHD-SFNet, whose UIE performance was much worse compared to LPS-Net, as shown in Table 1. Compared with Shallownet [38], we reduced the parameters by more than 50% and reduced the time complexity by more than 71%, but the PSNR and SSIM were 5.627 dB and 0.06 higher, which fully proved the superiority and feasibility of our method. Although LPS-Net did not have the best runtime metric, taking into consideration its performance and efficiency overall, it can be concluded that LPS-Net achieved state-of-the-art underwater image enhancement performance while enjoying superior model efficiency.

**Table 2.** The efficiency evaluation was conducted using  $1280 \times 720$  images as the input on an RTX 3090 graphics card. The results are presented below, with the best-performing outcomes indicated with <u>underlining</u> and the second-best results highlighted in **bold**. The symbol  $\uparrow$  denotes that higher values indicate better performance, while  $\downarrow$  signifies that lower values are desirable. This evaluation aimed to assess the computational efficiency of the methods under consideration and provide insights into their relative performance based on the given criteria.

	Efficiency						
Method	GFLOPs (G)↓	# of Params (M) $\downarrow$	Runtime (s) $\downarrow$	FPS (f/s)↑			
WaterNet(TIP'19) [5]	193.70	24.81	0.680	-			
FUNIE GAN(RAL'20) [43]	-	7.02	-	-			
PRW-Net(ICCVW'21) [15]	223.40	6.30	0.216	4.624			
Shallow-uwnet(AAAI'21) [38]	304.75	0.22	0.031	31.836			
Ucolor(TIP'21) [14]	443.85	157.42	2.758	-			
UIEC^2-Net(SPIC'21) [50]	367.53	0.53	0.174	5.742			
PUIE-Net(ECCV'22) [16]	423.05	1.40	0.071	14.194			
UHD-SFNet(ACCV'22) [18]	15.24	37.31	0.059	16.769			
NU2Net(AAAI'23,Oral) [17]	146.64	3.15	0.024	42.345			
MFEF(EAAI'23) [19]	-	2.79	-	-			
Ours	90.02	<u>0.08</u>	0.070	14.552			

Additionally, for the EUVP dataset, we employed the PSNR, SSIM, UCIQE, and UIQM as quantitative metrics to assess the recovery outcomes for more-extensive research. Notably, the results in Table 3 produced by LPS-Net demonstrate excellent performance in terms of the PSNR, SSIM, and UICIQE, yielding the highest scores. This observation highlights the robustness and generalizability of LPS-Net for underwater image enhancement.

**Table 3.** Experimental results on the EUVP-Dark [43] dataset. Best results are <u>underlined</u>. The second-best results are in **bold**.  $\uparrow$  represents that higher is better.

	EUVP-Dark (1000) [43]						
Method	<b>PSNR</b> ↑	<b>SSIM</b> ↑	UCIQE↑	UIQM↑			
Shallownet(AAAI'21) [38] NU2Net(AAAI'23,Oral) [17]	19.6058 <b>20.0908</b>	0.8167 <b>0.8631</b>	<b>0.5537</b> 0.5536	2.9411 <u>3.0775</u>			
Ours	<u>20.4129</u>	0.8802	0.5610	3.0466			

## 4.5. Visual Comparisons

In the case of the T90 dataset, a subset of images from the UIEB dataset was carefully chosen for conducting a visual comparison against seven state-of-the-art methods. These images were meticulously classified into various categories, encompassing yellow-toned, green-toned, blue-toned, low light, shallow-water areas, and deep-water areas. Figures 3 and 4 showcase the performance of each method. Shallownet succeeded in enhancing contrast, but fell short in improving the quality of yellow-toned and green-toned underwater images. UWGAN exhibited notable improvements in yellow-toned images; however, oversaturation issues arose in many instances. PRW-Net's applicability was limited across a wide range of scenarios, as it introduced additional chromatic aberrations or artifacts in deep-water areas and low-light conditions. UIEC^2-Net employed a multicolor spatial encoder for natural color processing, but inaccurate transmission map estimation resulted in contrast and saturation degradation. PUIE-Net and NU2Net encountered challenges with overall darkness and low contrast in a majority of the processed images. In contrast, our method effectively combines the strengths of multiple features, resulting in visually pleasing outcomes in terms of contrast, saturation, and detail processing.



Figure 3. Visual comparison of enhancement results on the T90 testing set.

Additionally, Figures 5–7 verify the generalization performance of LPS-Net, which can be finely enhanced for underwater degraded images of different qualities. Our method exhibited quite compatible color and detail recovery, enhancing the entire degraded image and making its contrast and texture details meet the sensory requirements of the human eyes. That is beneficial from our well-designed DAEB, MLB, and GFU.



Figure 4. Visual comparison of enhancement results on the T90 testing set.

Further, we also give an intuitive comparison with two previous SOTA methods on the EUVP dataset. As seen in Figures 8 and 9, Shallownet was incapable of sufficiently restoring underwater images due to its straightforward network structure. On the other hand, NU2Net lacked precise color control, hence leading to visible chromatic deviations that can be noticed by the human eye. In contrast, our method was better at removing color casts and handling details.



**Figure 5.** Visual comparison of enhancement results on the U45 testing set for generalization ability evaluation.



**Figure 6.** Visual comparison of enhancement results on the U45 testing set for generalization ability evaluation.



**Figure 7.** Visual comparison of enhancement results on the C60 testing set for generalization ability evaluation.



Figure 8. Visual comparison of enhancement results on EUVP-Dark.



Figure 9. Visual comparison of enhancement results on EUVP-Dark.

# 5. Ablation Study

In this section, a series of ablation studies are conducted to validate the role of the DAEB, MLB, and GFU in the network. For the ablation study, we followed the basic setup presented above and conducted experiments to demonstrate the effectiveness of the individual components of our proposed comprehensive approach. We performed different ablation studies to analyze different components. We first constructed the primary network as the baseline of the enhanced network, which mainly consisted of six convolutional layers, two IN layers, and ReLU activation functions.

# 5.1. Effectiveness of Key Components in the DAEB

We aimed to showcase the effectiveness of the proposed DAEB structure and present the experimental results in Table 4. Notably, the PSNR gain achieved by the IN layer was higher than that of the Batch Normalization (BN) layer, which can be attributed to their distinct normalization methods. While the BN layer does not account for channel differences during normalization, the IN layer utilizes unique means and variances for different layers. Additionally, the exclusion of the SA and CA mechanisms reduced the performance of the DAEB by restricting its focus to specific layers. To provide a more-intuitive understanding of the utility of the DAEB, we visualize the channel histogram before and after enhancement. As illustrated in Figure 10, the DAEB effectively reduced the differences between the three channels, as anticipated.

**Table 4.** Ablation study on the Dual-Attention Color Enhancement block. Among them, w/o SA and w/o CA refer to removing the Spatial Attention and Channel Attention, respectively.  $\uparrow$  represents that higher is better, and  $\downarrow$  represents that lower is better. <u>Underlines</u> indicate the best results.

Setting	<b>PSNR</b> ↑	<b>SSIM</b> ↑	MSE↓
(a) Base	19.412	0.863	0.127
(b) Base+DAEB(IN->BN)	22.110	0.894	0.087
(c) Base+DAEB(w/o SA)	22.338	0.898	0.085
(c) Base+DAEB(w/o CA)	22.558	0.903	0.082
Ours	22.651	<u>0.904</u>	0.072



Figure 10. Color correction results of degenerated underwater image by DAEB.

#### 5.2. Effectiveness with Different Kernel Sizes of the MLB

We performed experiments with different sizes of convolutional kernels in the MLB and evaluated the performance on the same datasets. As shown in Table 5, the PSNR and SSIM showed an upward trend for our design as the convolutional kernel became larger. Such observations are consistent with our assumption that the receptive field increases as the convolutional kernel increases, capturing more contextual information. As the number of network parameters increases, the generalization ability also increases. The performance begins to saturate and decay when the convolutional kernel is further expanded. Therefore, we considered using kernels of sizes  $7 \times 7$  and  $9 \times 9$  as the backbone of the MLB. We also visualize the convergence speed and performance of different sizes of convolutional kernels during training in Figure 11. We can observe that: compared with other configurations, when the convolutional kernels were  $7 \times 7$  and  $9 \times 9$ , the convergence speed was faster, and the performance increase was stable, which shows that the MLB under this configuration can extract rich attention context features with cross-dimensional interactions to capture good detail information and restore degraded image details.

Madal	Setting_kernel_size		Metrics			# . ( <b>D</b>	CELOD				
Model	1	3	5	7	9	11	<b>PSNR</b> ↑	<b>SSIM</b> ↑	MSE↓	# of Params	GFLOPS
(a)	$\checkmark$	$\checkmark$					23.535	0.909	0.0742	78.46 k	92.91 G
(b)		$\checkmark$	$\checkmark$				23.622	0.910	0.0736	78.53 k	93.18 G
(c)			$\checkmark$	$\checkmark$			23.771	0.913	0.0725	79.18 k	95.57 G
(d)					$\checkmark$	$\checkmark$	23.423	0.910	0.0744	81.34 k	103.53 G
Ours				$\checkmark$	$\checkmark$		23.905	0.915	0.0717	80.12 k	99.02 G

**Table 5.** The ablation study of the MLB components.  $\uparrow$  represents that higher is better, and  $\downarrow$  represents that lower is better. <u>Underlines</u> indicate the best results.



Figure 11. Metric visualization of the ablution study of the MLB during training.

# 5.3. Effectiveness of the GFU

To showcase the efficacy of the proposed GFU, we conducted an experiment by removing the GFU, which revealed a decline in both the PSNR and SSIM indicators when simply adding the output of the two branches. This can be attributed to the fact that the utility and focus of different modules are distinct, often resulting in significant disparities between their respective outputs. Directly adding these outputs can exacerbate such differences, leading to a loss of performance. Subsequently, we substituted the well-designed GFU with the Attention Feature Fusion (AFF) module proposed by Dai et al. [57]. The results presented in Table 6 demonstrate that, while AFF also led to considerable gains, its performance fell short when compared to the GFU. These findings reinforce the superiority of the GFU in effectively addressing feature differences across different branches.

**Table 6.** Ablation study on the Gated Fusion Unit.  $S_1$  means simply adding the results of the two branches.  $\uparrow$  represents that higher is better, and  $\downarrow$  represents that lower is better. <u>Underlines</u> indicate the best results.

Setting	Model	PSNR↑	<b>SSIM</b> ↑	MSE↓
(a)	S1	23.624	0.910	0.0732
(b)	AFF	23.891	0.914	0.0721
(c)	GFU	<u>23.905</u>	<u>0.915</u>	0.0717

#### 6. Limitations and Error Cases

Despite demonstrating effectiveness and exceptional performance in underwater image enhancement tasks through experiments on multiple datasets, LPS-Net is still constrained compared to other tasks due to the challenges in collecting underwater datasets, resulting in insufficient dataset sizes and suboptimal model optimization. Specifically, LPS-Net may require further model design and optimization to overcome the overfitting issue caused by small datasets and improve its performance in handling complex underwater image details. For instance, the utilization of unsupervised training methods could eliminate the need for collecting a large number of expensive paired data. Additionally, although the proposed model is expected to exhibit efficiency and flexibility in underwater robotics, further experiments are needed to validate its performance and reliability in practical applications.

Besides, in some uncommon cases, where the degraded images appear to be pink or orange, the DAEB, which was designed to process individual color channels, cannot effectively handle these particular colors, leading to unsatisfactory enhancement results. In the future, we will consider incorporating more color-processing branches to address the challenges posed by complex and diverse color environments.

#### 7. Conclusions and Future Work

This work proposes a Lightweight Parallel Strategy Network for underwater image enhancement. By introducing a Dual-Attention Enhancement Block and a Mirror Large Receptiveness Block, we addressed the mixed degeneracies of color distortion and the loss of detail in underwater raw images. LPS-Net utilizes these components as the building blocks of parallel branches and employs Gated Fusion Units to fuse features from different branches. Extensive experiments demonstrated that LPS-Net achieved state-of-the-art underwater image enhancement performance in terms of visual quality and evaluative metrics, while using only 80.12 k parameters. The success of LPS-Net proves that tackling mixed degradations by dividing them into sub-problems has the potential of avoiding the use of complex modules and reducing the amount of computation and parameters.

Despite the superiority of LPS-Net, there is still room for improvement. Looking ahead, we plan to incorporate frequency domain operations into LPS-Net since processing information in the Fourier space is capable of capturing the global representation in the frequency domain, while normal convolution focuses on learning local representations in the spatial domain. Another direction is to take advantage of both visual transformers and CNNs. It is worth investigating how to bring the merits of transformers while achieving a good balance between UIE performance and model complexity.

The proposed LPS-Net demonstrated impressive performance in real-world underwater image enhancement, making it beneficial for other downstream vision applications, such as underwater photography, underwater archeology, and marine biological survey. We also plan to investigate whether LPS-Net can be adapted to other image-restoration tasks [58], such as image dehazing and image deraining. Of course, efforts should be made to tackle mixed degradations in these tasks by dividing them into sub-problems that could be effectively addressed.

**Author Contributions:** Conceptualization, J.J.; methodology, J.J. and L.T.; software, P.H. and J.Y.; validation, J.J. and J.Y.; formal analysis, E.C. and P.H.; investigation, E.C., J.Y. and J.J.; resources, E.C.; data curation, J.J.; writing—original draft preparation, J.J., L.T. and E.C.; project administration, E.C.; funding acquisition, E.C. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the Natural Science Foundation of Fujian Province of China under Grant (2021J01867), Xiamen Ocean and Fisheries Development Special Funds (22CZB013HJ04), and the Youth Science and Technology Innovation Program of Xiamen Ocean and Fisheries Development Special Funds (23ZHZB039QCB24).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Publicly available datasets were analyzed in this study. These data can be found here: https://drive.google.com/file/d/12W\_kkblc2Vryb9zHQ6BfGQ\_NKUfXYk13/view (accessed on 20 October 2022), https://drive.google.com/file/d/1cA-8CzajnVEL4feBRKdBxjEe6hw ql6Z7/view (accessed on 20 October 2022), https://drive.google.com/file/d/1Ew\_r83nXzVk0hlkfu omWqsAIxuq6kaN4/view (accessed on 20 October 2022), and https://github.com/IPNUISTlegal/underwater-test-dataset-U45- (accessed on 20 October 2022). The divided UIEB dataset has 800 pairs of quantity images and 90 pairs of test images: https://pan.baidu.com/s/1dylR8g2lMK5Cd2eIVm\_XLQ?pwd=r3i2 (accessed on 20 October 2022).

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

#### 20 of 22

## Abbreviations

The following abbreviations are used in this manuscript:

AFF	Attention Feature Fusion
BN	Batch Normalization
CA	Channel Attention
CNNs	Convolutional Neural Networks
DAEB	Dual-Attention Enhancement Block
EUVP	Enhancement of Underwater Visual Perception
GLU	Gated Linear Unit
GAN	Generative Adversarial Network
GFU	Gated Fusion Unit
IBLA	Image Blurriness and Light Absorption
IN	Instance Normalization
LPS-Net	Lightweight Parallel Strategy Network
MLB	Mirror Large Receptiveness Block
MSE	Mean-Squared Error
PSO	Particle Swarm Optimization
PSNR	Peak-Signal-to-Noise Ratio
SOTA	State-Of-The-Art
SSIM	Structural Similarity index
SA	Spatial Attention
UIE	Underwater Image Enhancement
ULAP	Underwater Light Attenuation Prior
UDCP	Underwater Dark Channel Prior
UCIQE	Underwater Color Image Quality Evaluation
UIEB	Underwater Image Enhancement Benchmark
UIQM	Underwater Image Quality Metric

# References

- 1. Hubert, A.M. Marine scientific research and the protection of the seas and oceans. In *Research Handbook on International Marine Environmental Law;* Edward Elgar Publishing: Cheltenham, UK , 2023; pp. 385–408.
- Hubert, A.M. The new paradox in marine scientific research: Regulating the potential environmental impacts of conducting ocean science. Ocean. Dev. Int. Law 2011, 42, 329–355. [CrossRef]
- 3. Schander, C.; Willassen, E. What can biological barcoding do for marine biology? Mar. Biol. Res. 2005, 1, 79–83. [CrossRef]
- Schettini, R.; Corchs, S. Underwater image processing: State of the art of restoration and image enhancement methods. EURASIP J. Adv. Signal Process. 2010, 2010, 746052. [CrossRef]
- 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]
- Jiang, J.; Bai, J.; Liu, Y.; Yin, J.; Chen, S.; Ye, T.; Chen, E. RSFDM-Net: Real-time Spatial and Frequency Domains Modulation Network for Underwater Image Enhancement. *arXiv* 2023, arXiv:2302.12186.
- McGlamery, B. A computer model for underwater camera systems. In *Proceedings of the Ocean Optics VI, Monterey, CA, USA,* 23–25 October 1979; SPIE: Bellingham, WA, USA, 1980; Volume 208, pp. 221–231.
- Jaffe, J.S. Computer modeling and the design of optimal underwater imaging systems. *IEEE J. Ocean. Eng.* 1990, 15, 101–111. [CrossRef]
- 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.
- Song, W.; Wang, Y.; Huang, D.; Tjondronegoro, D. A rapid scene depth estimation model based on underwater light attenuation prior for underwater image restoration. In *Proceedings, Part I 19, Proceedings of the Advances in Multimedia Information Processing– PCM 2018: 19th Pacific-Rim Conference on Multimedia, Hefei, China, 21–22 September 2018; Springer: New York, NY, USA, 2018;* pp. 678–688.
- 11. Abdullah-Al-Wadud, M.; Kabir, M.H.; Dewan, M.A.A.; Chae, O. A dynamic histogram equalization for image contrast enhancement. *IEEE Trans. Consum. Electron.* 2007, 53, 593–600. [CrossRef]
- Singh, G.; Jaggi, N.; Vasamsetti, S.; Sardana, H.; Kumar, S.; Mittal, N. Underwater image/video enhancement using wavelet based color correction (WBCC) method. In *Proceedings of the 2015 IEEE Underwater Technology (UT), Chennai, India, 23–25 February* 2015; IEEE: New York, NY, USA, 2015; pp. 1–5.
- Zhang, S.; Wang, T.; Dong, J.; Yu, H. Underwater image enhancement via extended multi-scale Retinex. *Neurocomputing* 2017, 245, 1–9. [CrossRef]

- Li, C.; Anwar, S.; Hou, J.; Cong, R.; Guo, C.; Ren, W. Underwater image enhancement via medium transmission-guided multicolor space embedding. *IEEE Trans. Image Process.* 2021, 30, 4985–5000. [CrossRef]
- Huo, F.; Li, B.; Zhu, X. Efficient wavelet boost learning-based multi-stage progressive refinement network for underwater image enhancement. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Barcelona, Spain, 10–17 October 2021; pp. 1944–1952.
- Fu, Z.; Wang, W.; Huang, Y.; Ding, X.; Ma, K.K. Uncertainty Inspired Underwater Image Enhancement. In *Proceedings, Part XVIII, Proceedings of the Computer Vision–ECCV 2022: 17th European Conference, Tel Aviv, Israel, 23–27 October 2022; Springer: New York, NY, USA, 2022; pp. 465–482.*
- 17. Guo, C.; Wu, R.; Jin, X.; Han, L.; Chai, Z.; Zhang, W.; Li, C. Underwater Ranker: Learn Which Is Better and How to Be Better. In Proceedings of the AAAI Conference on Artificial Intelligence, Washington, DC, USA, 13–14 February 2023.
- Wei, Y.; Zheng, Z.; Jia, X. UHD Underwater Image Enhancement via Frequency-Spatial Domain Aware Network. In Proceedings of the Asian Conference on Computer Vision, Virtual, 22 February–1 March 2022; pp. 299–314.
- Zhou, J.; Sun, J.; Zhang, W.; Lin, Z. Multi-view underwater image enhancement method via embedded fusion mechanism. *Eng. Appl. Artif. Intell.* 2023, 121, 105946. [CrossRef]
- Wang, J.; Li, P.; Deng, J.; Du, Y.; Zhuang, J.; Liang, P.; Liu, P. CA-GAN: Class-condition attention GAN for underwater image enhancement. *IEEE Access* 2020, *8*, 130719–130728. [CrossRef]
- 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]
- Li, C.; Anwar, S.; Porikli, F. Underwater scene prior inspired deep underwater image and video enhancement. *Pattern Recognit.* 2020, 98, 107038. [CrossRef]
- Wang, Q.; Wu, B.; Zhu, P.; Li, P.; Zuo, W.; Hu, Q. ECA-Net: Efficient channel attention for deep convolutional neural networks. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 13–19 June 2020; pp. 11534–11542.
- Xu, H.; Saenko, K. Ask, attend and answer: Exploring question-guided spatial attention for visual question answering. In Proceedings, Part VII 14, Proceedings of the Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, 11–14 October 2016; Springer: New York, NY, USA, 2016; pp. 451–466.
- Liu, Z.; Lin, Y.; Cao, Y.; Hu, H.; Wei, Y.; Zhang, Z.; Lin, S.; Guo, B. Swin transformer: Hierarchical vision transformer using shifted windows. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Nashville, TN, USA, 20–25 June 2021; pp. 10012–10022.
- Khan, S.; Naseer, M.; Hayat, M.; Zamir, S.W.; Khan, F.S.; Shah, M. Transformers in vision: A survey. ACM Comput. Surv. (CSUR) 2022, 54, 1–41. [CrossRef]
- Liu, Z.; Mao, H.; Wu, C.Y.; Feichtenhofer, C.; Darrell, T.; Xie, S. A ConvNet for the 2020s. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), New Orleans, LA, USA, 18–24 June 2022; pp. 11976–11986.
- Raveendran, S.; Patil, M.D.; Birajdar, G.K. Underwater image enhancement: A comprehensive review, recent trends, challenges and applications. *Artif. Intell. Rev.* 2021, 54, 5413–5467. [CrossRef]
- 29. Sahu, P.; Gupta, N.; Sharma, N. A survey on underwater image enhancement techniques. Int. J. Comput. Appl. 2014, 87. [CrossRef]
- Yang, M.; Hu, J.; Li, C.; Rohde, G.; Du, Y.; Hu, K. An in-depth survey of underwater image enhancement and restoration. *IEEE Access* 2019, 7, 123638–123657. [CrossRef]
- Jiang, Q.; Gu, Y.; Li, C.; Cong, R.; Shao, F. Underwater image enhancement quality evaluation: Benchmark dataset and objective metric. *IEEE Trans. Circuits Syst. Video Technol.* 2022, 32, 5959–5974. [CrossRef]
- AbuNaser, A.; Doush, I.A.; Mansour, N.; Alshattnawi, S. Underwater image enhancement using particle swarm optimization. J. Intell. Syst. 2015, 24, 99–115. [CrossRef]
- Iqbal, K.; Odetayo, M.; James, A.; Salam, R.A.; Talib, A.Z.H. Enhancing the low quality images using unsupervised color correction method. In *Proceedings of the 2010 IEEE International Conference on Systems, Man and Cybernetics, Istanbul, Turkey, 10–13 October 2010; IEEE: New York, NY, USA, 2010; pp. 1703–1709.*
- 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; IEEE: New York, NY, USA, 2014; pp. 4572–4576.
- 35. Zhuang, P.; Li, C.; Wu, J. Bayesian Retinex underwater image enhancement. Eng. Appl. Artif. Intell. 2021, 101, 104171. [CrossRef]
- Zhang, W.; Zhuang, P.; Sun, H.H.; Li, G.; Kwong, S.; Li, C. Underwater image enhancement via minimal color loss and locally adaptive contrast enhancement. *IEEE Trans. Image Process.* 2022, 31, 3997–4010. [CrossRef] [PubMed]
- 37. Jiang, Q.; Zhang, Y.; Bao, F.; Zhao, X.; Zhang, C.; Liu, P. Two-step domain adaptation for underwater image enhancement. *Pattern Recognit.* **2022**, 122, 108324. [CrossRef]
- Naik, A.; Swarnakar, A.; Mittal, K. Shallow-uwnet: Compressed model for underwater image enhancement (student abstract). In Proceedings of the AAAI Conference on Artificial Intelligence, Vancouver, BC, Canada, 2–9 February 2021; Volume 35, pp. 15853–15854.
- Chen, X.; Yu, J.; Kong, S.; Wu, Z.; Fang, X.; Wen, L. Towards real-time advancement of underwater visual quality with GAN. IEEE Trans. Ind. Electron. 2019, 66, 9350–9359. [CrossRef]

- 40. Zhang, D.; Wu, C.; Zhou, J.; Zhang, W.; Li, C.; Lin, Z. Hierarchical attention aggregation with multi-resolution feature learning for GAN-based underwater image enhancement. *Eng. Appl. Artif. Intell.* **2023**, *125*, 106743. [CrossRef]
- Wang, F.; Jiang, M.; Qian, C.; Yang, S.; Li, C.; Zhang, H.; Wang, X.; Tang, X. Residual attention network for image classification. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 22–25 July 2017; pp. 3156–3164.
- Fujiyoshi, H.; Hirakawa, T.; Yamashita, T. Deep learning-based image recognition for autonomous driving. *IATSS Res.* 2019, 43, 244–252. [CrossRef]
- Islam, M.J.; Xia, Y.; Sattar, J. Fast Underwater Image Enhancement for Improved Visual Perception. IEEE Robot. Autom. Lett. 2020, 5, 3227–3234. [CrossRef]
- 44. Dauphin, Y.N.; Fan, A.; Auli, M.; Grangier, D. Language modeling with gated convolutional networks. In Proceedings of the International Conference on Machine Learning, Sydney, Australia, 6–11 August 2017; pp. 933–941.
- Zhao, H.; Jia, J.; Koltun, V. Exploring self-attention for image recognition. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 13–19 June 2020; pp. 10076–10085.
- Charbonnier, P.; Blanc-Feraud, L.; Aubert, G.; Barlaud, M. Two deterministic half-quadratic regularization algorithms for computed imaging. In *Proceedings of the 1st International Conference on Image Processing, Austin, TX, USA, 13–16 November 1994;* IEEE: New York, NY, USA, 1994; Volume 2, pp. 168–172.
- 47. Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. arXiv 2014, arXiv:1409.1556.
- 48. Li, H.; Li, J.; Wang, W. A fusion adversarial underwater image enhancement network with a public test dataset. *arXiv* 2019, arXiv:1906.06819.
- Peng, Y.T.; Cosman, P.C. Underwater image restoration based on image blurriness and light absorption. *IEEE Trans. Image Process.* 2017, 26, 1579–1594. [CrossRef]
- Wang, Y.; Guo, J.; Gao, H.; Yue, H. UIEC<sup>2</sup>-Net: CNN-based underwater image enhancement using two color space. *Signal Process. Image Commun.* 2021, 96, 116250. [CrossRef]
- 51. Paszke, A.; Gross, S.; Massa, F.; Lerer, A.; Bradbury, J.; Chanan, G.; Killeen, T.; Lin, Z.; Gimelshein, N.; Antiga, L.; et al. Pytorch: An imperative style, high-performance deep learning library. *Adv. Neural Inf. Process. Syst.* **2019**, *32*.
- Korhonen, J.; You, J. Peak signal-to-noise ratio revisited: Is simple beautiful? In Proceedings of the 2012 Fourth International Workshop on Quality of Multimedia Experience, Melbourne, VIC, Australia, 5–7 July 2012; IEEE: New York, NY, USA, 2012; pp. 37–38.
- 53. 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]
- 54. Marmolin, H. Subjective MSE measures. IEEE Trans. Syst. Man, Cybern. 1986, 16, 486–489. [CrossRef]
- 55. Yang, M.; Sowmya, A. An underwater color image quality evaluation metric. *IEEE Trans. Image Process.* **2015**, *24*, 6062–6071. [CrossRef] [PubMed]
- Panetta, K.; Gao, C.; Agaian, S. Human-visual-system-inspired underwater image quality measures. *IEEE J. Ocean. Eng.* 2015, 41, 541–551. [CrossRef]
- Dai, Y.; Gieseke, F.; Oehmcke, S.; Wu, Y.; Barnard, K. Attentional feature fusion. In Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, Waikoloa, HI, USA, 3–8 January 2021; pp. 3560–3569.
- Liu, Y.; Yan, Z.; Tan, J.; Li, Y. Multi-Purpose Oriented Single Nighttime Image Haze Removal Based on Unified Variational Retinex Model. *IEEE Trans. Circuits Syst. Video Technol.* 2023, 33, 1643–1657. [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.