applied sciences

MDPI

Article

Underwater Image Super-Resolution via Dual-aware Integrated Network

Aiye Shi * and Haimin Ding

College of Computer and Information, Hohai University, Nanjing 211100, China; 221307020002@hhu.edu.cn
* Correspondence: ayshi@hhu.edu.cn

Abstract: Underwater scenes are often affected by issues such as blurred details, color distortion, and low contrast, which are primarily caused by wavelength-dependent light scattering; these factors significantly impact human visual perception. Convolutional neural networks (CNNs) have recently displayed very promising performance in underwater super-resolution (SR). However, the nature of CNN-based methods is local operations, making it difficult to reconstruct rich features. To solve these problems, we present an efficient and lightweight dual-aware integrated network (DAIN) comprising a series of dual-aware enhancement modules (DAEMs) for underwater SR tasks. In particular, DAEMs primarily consist of a multi-scale color correction block (MCCB) and a swin transformer layer (STL). These components work together to incorporate both local and global features, thereby enhancing the quality of image reconstruction. MCCBs can use multiple channels to process the different colors of underwater images to restore the uneven underwater light decay-affected real color and details of the images. The STL captures long-range dependencies and global contextual information, enabling the extraction of neglected features in underwater images. Experimental results demonstrate significant enhancements with a DAIN over conventional SR methods.

Keywords: underwater image; super-resolution; transformer; multi-scale

Citation: Shi, A.; Ding, H. Underwater Image Super-Resolution via Dual-aware Integrated Network. *Appl. Sci.* **2023**, *13*, 12985. https://doi.org/10.3390/app132412985

Academic Editors: Jie Tian and Atsushi Mase

Received: 1 November 2023
Revised: 1 December 2023
Accepted: 1 December 2023
Published: 5 December 2023

1. Introduction

As global pressure on land-based resources continues to grow, increasing attention is beginning to be focused on the oceans and seas as an important source of natural resources. The oceans cover the vast majority of the Earth's surface, which contains rich biodiversity, energy potential, and mineral resources. However, due to complex factors such as scattering, absorption, and color deviation of light in underwater environments, the quality of underwater scenes is significantly degraded. To address these challenging issues, many methods [1–6] have been developed to improve degraded image quality. In the beginning, people tended to make more elaborate filters to recreate as many images as possible captured in underwater environments, largely inspired by the theory of the human retina [7]. Later, physical models were often used to mimic the respective complex environments in real scenarios [8]. In addition to this, they also used a priori methods to globally enhance the picture [9]. Image super-resolution (SR) techniques are designed to restore high-resolution (HR) images from low-resolution (LR) images, which can restore the details and clarity of the images and improve the visualization and application of underwater images. Although the image improvement of these methods performs well in some specific contexts, it may suffer from poor performance when dealing with complex, dynamic, and variable real-world images.

Over the past few years, advances in deep learning have offered fresh solutions for underwater image SR, allowing researchers to better utilize models such as convolutional neural networks (CNNs) to enhance the visual quality and detail of degraded images. Zhang et al. [10] proposed a multipath crossing module that contains both residual and dilation blocks to boost the learning ability of the model and enhance the representation

of abstract features. Wang et al. [11] presented a lightweight multi-stage information distillation network, referred to as MSIDN. The MSIDN was designed to strike a better balance between performance and applicability by aggregating locally distilled features from different stages to capture more potent feature representations. Sharma et al. [12] utilized a convolutional block attention module (CBAM)) [13] to specify weights for channel features derived from convolutional neural networks (CNNs) with varying receptive field sizes. Wang et al. [14] proposed a progressive frequency interleaved network called PFIN to enhance and restore underwater images, performing effective color bias correction and detail enhancement in underwater images.

Despite a large amount of previous research on CNN-based methods, previous studies have encountered challenges related to reconstructing optical artificialities (e.g., color distortions) in underwater images. Wang et al. [15] were the first to integrate the HSV color space with deep learning techniques for color correction, noise reduction in the RGB space, and optimization of luminance and satellite images in the HSV space. Li et al. [16] proposed a transmission-guided framework that seeks to enhance feature representation by incorporating multi-color features. Liu et al. [17] proposed an enhancement method based on a super-resolution convolutional neural network (SRCNN) and perceptual fusion by combining learned and unlearned methods, which achieved exciting results in underwater image deblurring and color enhancement. However, existing methods have limitations that do not take into account the global information interaction.

Nowadays, the transformer is widely used in computer vision with its powerful global modeling capabilities. Similarly, various transformer structures are introduced in underwater image restoration tasks to achieve better restoration results. Liu et al. [18] proposed a structure with unique hierarchical segmentation and multiple levels based on a transformer, which overcame the challenge of processing large-size images and effectively captured information regarding long-range dependencies. Peng et al. [19] were the first to apply the U-shaped structure in combination with a transformer to UIE and achieved exciting results. A dual attention transformer-based method was introduced by Shen et al. [20] for underwater image enhancement that can better reconstruct the image. Liang et al. [21] offered the SwinIR method for image restoration using the swin transformer to achieve good image reconstruction. Huang et al. [22] proposed a new adaptive group attention method which was added to the transformer to mitigate the diffusion due to underwater magnesium. Guo et al. [23] used air images to guide underwater image clearness and leveraged the transformer to acquire overall information from the underwater images. Although all of the above methods have yielded some results, they all have certain drawbacks. Models based on CNN approaches may require a large number of parameters and computational resources, and deploying such models on resource-constrained underwater devices may be challenging, which happens to be one of the advantages of transformers. Similarly, transformer-based models are based on a self-attentive mechanism that pays global attention to the entire input sequence. In some underwater scenarios, images may contain a large number of local structures and details, and transformers may not be as effective as methods such as CNNs in dealing with these local structures.

To address these issues, we propose an active dual-aware integrated network (DAIN) for underwater image SR in this paper. The general architecture can be seen in Figure 1, which shows that the DAIN mainly comprises a sequence of dual-aware enhancement modules (DAEMs). A multi-scale color correction block (MCCB) and swin transformer layer (STL) constitute the DAEM, which can utilize the short-range ability of the CNN and the long-range ability of the STL to better integrate local and global information, thereby improving the temporal resolution of degraded underwater images. The MCCB utilizes multiple channels to process the different colors of underwater images to reduce color bias, while the STL captures global contextual information to improve image detail reconstruction. Experimental results display that our DAIN outperforms the most popular methods with low model capacity.

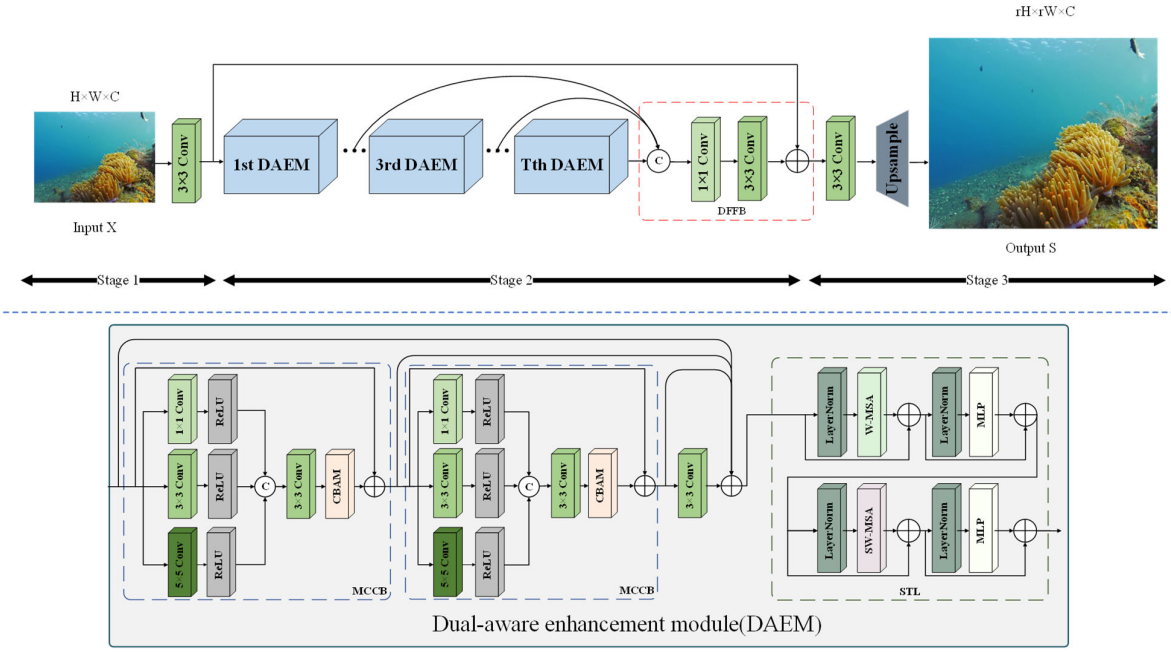

Figure 1. Network architecture of our proposed DAIN.

To sum up, our main contributions can be summarized as being three-fold:

- We introduce a lightweight and efficient DAIN tailored for the underwater SR domain. Leveraging the DAME, which consists of an MCCB and STL, our approach demonstrates superior reconstruction performance, outperforming the majority of existing underwater SR methods.
- Our proposed DAEM adeptly integrates local and global features, enhancing the model's capacity to capture intricate details.
- The MCCB employs multi-channel and multi-scale strategies to address the challenge of uneven color attenuation in underwater images. Notably, the incorporation of the attention mechanism further elevates the representational prowess of the network.

While transformer-based methods excel in global modeling, they exhibit limitations in local modeling. Furthermore, to streamline computations and enhance model efficiency, transformers often entail a significant reduction in computational parameters. Consequently, recent efforts have explored hybrid models that synergize global features extracted by the transformer with local features extracted by a CNN.

2. Related Work

2.1. CNN-Based Underwater Image Super-Resolution

Developing deep learning for various computing tasks shows great benefits [24–28]. Image SR has been a major focus of research in the area of computer vision. It is focused on increasing the quality of photos and videography. SRCNNs [29] first used CNNs in the SR task, which can learn the relationships between paired images, achieving significant improvements over conventional approaches. In order to push the boundaries of reconstruction accuracy, numerous studies have been carried out in this area [30–36]. In contrast, less research has been performed on underwater images. Owing to the paucity of extensive underwater image datasets, the execution of the aforementioned underwater SR model has some deficiencies. Thankfully, Islam et al. [37] proposed a dataset termed USR-248 that contains paired LR–HR images and devised an underwater SR generative model named SRDRM. To allow for deeper exploration of the underwater SR domain, they also developed an adversarial version of SRDRM, called SRDRM-GAN, which adopts Markov PatchGAN [38] as the discriminator to reconstruct more texture details. Subsequent to this, an increasing number of CNN-based and GAN-based models have been utilized in underwater tasks. Cherian et al. [39] proposed an

approach, called the alpha super-resolution generative adversarial network (AlphaSRGAN), to improve restore accuracy. Improving underwater imagery quality results in higher resolution and more concise details. Li et al. [40] proposed WaterGAN for synthesizing underwater images based on indoor images and depth maps. Huo et al. [41] proposed an underwater residual convolutional neural network (URCNN) with great depth based on VGG [42]. Li et al. [43] developed an algorithm for underwater image processing to reduce image blurriness, and they then utilized an end-to-end network to create natural and color-enhanced images. Although these approaches are effective in addressing the undesirable effects of underwater scenes, they lack modeling of global information and they are not conducive to generating more natural and realistic textures.

2.2. Transformer-Based Underwater Images

In contrast to CNN-based methods, Alexey et al. [44] suggested that the direct application of a pure transformer to a sequence of patches in a visual transformer (ViT) could perform the task of image classification well. Liu et al. [18] put forth a shifted windows transformer (swin transformer) to address the difficulty of matching transformers from the verbal domain to the visual domain due to the differences between the two domains. Wang et al. [45] put forth an architecture that includes a novel block of locally enhanced window transformers for stripping, denoising, and deblurring. Recently, transformers have gradually been increasingly utilized in underwater imaging with notable success. For example, Peng et al. [19] developed a U-shaped transformer, integrating a channel-level multi-scale feature fusion transformer block and a spatial-level multiscale feature fusion transformer block to model global features so as to enhance underwater images. Ren et al. [46] presented a dual transformer structure based on a U-shaped structure for super-resolution reconstruction of underwater images and achieved exciting results. Zhang et al. [47] developed the WaterFormer, a two-stage network that combines deep learning and an underwater physical model to tackle the numerous distortions found in underwater images as a result of water's absorption and scattering properties. Sun et al. [48] developed a model comprising a strengthened LeWin transformer block-based encoder and decoder to enhance color accuracy. The authors of [49] constructed a model consisting of a grey-scale attention and phase transformer block for underwater enhancement. Qi et al. [50] constructed an underwater image enhancement network, known as SGUIENet, utilizing semantic information as high-level guidance to improve the acquisition of locally enhanced features. Although transformers perform well in terms of capturing global dependencies, they are relatively weak in terms of local details. This is a challenge for image SR tasks because local features are critical for understanding details in an image. In this paper, we propose a module, called the DAEM, which efficiently models both global and local information and facilitates high-quality underwater image reconstruction.

3. Methods

3.1. Overall Network Architecture

The proposed DAIN, as illustrated in Figure 1, is composed of three phases. Stage 1 strives to obtain shallow feature information; Stage 2 strives to obtain and merge deeper features; and Stage 3 strives to merge dense features and reconstruct underwater images. The input of our network is an LR image ($X \in R^{H \times W \times 3}$) and its output is an HR image ($S \in R^{rH \times rW \times 3}$). The height and width of the image are denoted by H and W, respectively. The scale factor is represented by r.

In Stage 1, we first use a 3×3 convolution to extract shallow information:

$$F_0 = H_{SFE}(X) \tag{1}$$

where $H_{SFE}(\cdot)$ is the 3×3 convolution operation and the deviation term of the convolutional layer is omitted for simplicity. F_0 is then delivered to Stage 2 as an input to the DAEM, which uses two MCCBs and an STL to model the multi-scale and long-range dependencies

of features. More detailed information will be described in Sections 3.2 and 3.3. Assuming the amount of DAEMs is D, the result of the dth DAEM $F_d (1 \leq d \leq D)$ can be expressed as:

$$F_d = H^d_{DAEM}(H^{d-1}_{DAEM} \cdots ((H^1_{DAEM}(F_0)))) \tag{2}$$

where $H^d_{DAEM}(\cdot)$ is the operation of the dth DAEM and F_d denotes the result of the dth DAEM. More importantly, all the results of these DAEMs are combined and transmitted to the dense feature fusion block (DFFB), which aggregates all the hierarchical features to generate more expressive feature representations. To ease the learning difficulties, we add a global residual learning strategy. This specific process is defined as follows:

$$F_{DFFB} = H_{DFFB}[F_0, F_1, F_2, \cdots, F_D] \tag{3}$$

where $H_{DFFB}(\cdot)$ denotes the operation of the DFFB, which contains a 3×3 convolution and a 1×1 convolution. $[F_1, F_2, \cdots, F_d]$ is the combination of all the characters produced by the DAEM. Finally, Stage 3 functions as the element for recuperation. An upsampling operation is utilized to upgrade the depth detail to the HR image required. We use 3×3 convolutional layers and sub-pixel convolution to reconstruct the SR image S as follows:

$$S = H_{UP}(F_{DFFB}) \tag{4}$$

where $H_{UP}(\cdot)$ is the function of upsampling.

Based on the existing image SR, the L1 loss can be integrated into the training of the DAIN, which helps maintain rich texture and local structure [51].

$$L = L_1(G) = E_{X,S}[\|S - D(X)\|_1] \tag{5}$$

Here, $\|\cdot\|_1$ represents the L1-norm function. The mapping relationship between LR and HR is implemented by $D : \{X\} \rightarrow S$.

3.2. Dual-Aware Enhancement Module (DAEM)

In the realm of underwater image processing, the correction of color distortions and the recovery of texture details in degraded images necessitate meticulous decomposition. Many existing methods tend to overlook the integration of optical and visual perception, relying solely on simple convolution for high-frequency feature extraction. However, the unique characteristics of underwater images demand separate operations for distinct colors due to uneven color recession. Additionally, the entire image is affected by uniform blurring caused by light scattering and absorption issues, which extend beyond localized regions. Consequently, incorporating long-range dependencies becomes crucial to further enhance reconstruction results.

To address the challenges posed by the diverse complexities of the underwater environment and to elevate image reconstruction performance, we propose the dual-aware enhancement module (DAEM). This module captures dual-aware information, integrating both local and global features to progressively generate natural and realistic textures.

As shown in Figure 1, the DAEM consists of two MCCBs, a 3×3 convolutional layer, and an STL. The MCCB mitigates color shifts and distortions in underwater images caused by uneven light attenuation. More details will be shown in Section 3.3. Given the input features F_{d-1}, it first undergoes processing in MCCBs to achieve color bias correction and high-frequency detail recovery, and then fuses the different features using a 3×3 convolution. Finally, the STL models long-range dependencies to generate powerful feature representation.

3.3. Multi-Scale Color Correction Block (MCCB)

In underwater images, uneven decay of light propagation leads to low contrast and high color distortion leading to color asymmetry, with the degree of decay varying with wavelength. Conventional underwater processing methods usually adopt a global color

compensation strategy and ignore the problem that different color channels are subject to different degrees of attenuation in underwater environments. Consequently, we propose an MCCB that differentially handles three color channels and effectively compensates for different attenuation conditions, avoiding the shortcomings of global processing.

As illustrated in Figure 1, the input features were initially divided into three scales through the use of various convolution sizes dependent on the level of underwater light reduction. Since the attenuation of red light is the most significant, we use a 1×1 convolution, while green light and blue light utilize 3×3 and 5×5 convolution kernels, respectively. This is because different sizes of convolution kernels have different performance in terms of image details and a wide range of features. Let us take the first MCCB as an example, where the procedure could be written as follows:

$$F_{R,1} = \sigma(f_{1\times1}(F_{d-1})) \tag{6}$$

$$F_{G,3} = \sigma(f_{3\times3}(F_{d-1})) \tag{7}$$

$$F_{B,5} = \sigma(f_{5\times5}(F_{d-1})) \tag{8}$$

where $f(\cdot)$ represents the convolution operation, subscripts represent convolution kernels, and $\sigma(\cdot)$ represents the Relu activation function.

Subsequently, we splice and fuse the feature information extracted from each of the three colors:

$$F_{cat} = F_{R,1} \odot F_{G,3} \odot F_{B,5} \tag{9}$$

where $\odot$ denotes the channel connection and $F_{cat}(\cdot)$ denotes the output after the channel connection. Also, to combine the spatial and channel attentional weights, we use a convolutional block attention module (CBAM) [13]. Finally, we fuse and extract more advanced feature information through the CBAM, which can be defined as follows:

$$F_{MCCB} = H_{CBAM}(F_{cat}) \tag{10}$$

where $H_{CBAM}(\cdot)$ denotes the operation of the CBAM. F_{MCCB} is the output of the MCCB.

The CBAM is an effective and efficient attention module that is widely used in many computer vision tasks. It integrates channel and spatial attention mechanisms to improve the model's ability to perceive different feature channels and spatial locations. Overall, the use of a CBAM module in underwater image processing can assist the neural network in concentrating on significant features, intensifying the details and contrast of the image and mitigating the lighting problem, thereby enhancing the quality and visualization of underwater images.

3.4. Swin Transformer Layer (STL)

The swin transformer originates from the design of the initial transformer layer, which relies on regular multi-head self-attention to model long-range dependence and thus enhance network representation.

As illustrated in Figure 1, with an initial input of $F_{MCCB} \in \mathbb{R}^{H\times W\times C}$, the swin transformer reshapes the input $\frac{HW}{M^2} \times M^2 \times C$ feature first, where $\frac{HW}{M^2}$ represents the total number of windows. Then it performs self-attention calculations simultaneously h times, where h represents the quantity of self-attention heads. For a local window feature $F_{in}^{swt} \in R^{M^2\times C}$, the query, key, and value matrices Q, K, and $V \in R^{M^2\times C}$ are expressed as

$$Q = F_{in}^{swt} W_Q, K = F_{in}^{swt} W_K, V = F_{in}^{swt} W_V \tag{11}$$

where $d = \frac{C}{H}$,W_Q, W_K, and W_V are shared learnable projections over different windows. The attention matrix $Attn(Q, K, V)$ is then computed via the self-attention in the local window.

$$Attn(Q, K, V) = SoftMax(\frac{QK^T}{\sqrt{d}} + b)V \tag{12}$$

Here, b is the trainable relative positional coding. Multi-head self-attention (MSA) is combined to maintain consistent embedding dimensions. After the attention operation, there are two Multilayer Perceptrons (MLPs) with GELU activation in them. The Layer-Norms (LNs) are then added ahead of MSA and the MLP and use residual connections. The transformer's overall functionality can be summarized as follows:

$$\begin{aligned} F_{\text{inter}}^{swt} &= H_{MSA}(H_{LN}(F_{in}^{swt})) + F_{in}^{swt} \\ F_{out}^{swt} &= H_{MLP}(H_{LN}(F_{\text{inter}})) + F_{\text{inter}}^{swt} \end{aligned} \tag{13}$$

where $H_{LN}(\cdot)$ denotes the LN operation, $H_{MSA}(\cdot)$ represents MSA operation, and $H_{MLP}(\cdot)$ denotes the operation function. On the one hand, the STL exploits regular and shift window partition alternately, realizing efficient information transmission and interaction between different windows. On the other hand, it uses local attention and a shifted window mechanism to greatly reduce computational effort while maintaining strong modeling power. Therefore, we introduce the STL to better catch local features and details, thus enhancing underwater image processing.

4. Experiments

4.1. Experimental Setup

Two open datasets were adopted to train our proposed work, namely USR-248 [52] and UFO-120 [53], as well as the EUVP dataset [52]. USR-248 was the first dataset made for the SR reconstruction task of underwater optical images, which uses 1060 pairs of underwater images for training and 248 pairs of samples for testing. UFO-120 contains over 1500 pairs of training samples and 120 pairs of testing samples. It is important to note that the LR samples in the USR-248 and UFO-120 datasets are acquired by artificial distortion. Both of them follow the standard procedures [54,55] for optical/spatial image degradation and use manually labeled saliency mappings to create pairs of data.

Adam optimization is utilized to minimize the objective function, with the parameters configured as follows: $\beta_1 = 0.9$,$\beta_2 = 0.999$, and $\varepsilon = 10^{-8}$. The learning rate is initially set to 10^{-4} and then decreased by 50% after 300 iterations. In the STL, the number of multi-heads is set to h = 5 and the number of channels is set to C = 50. The performance of image super-resolution (SR) was thoroughly assessed using three criteria: peak signal-to-noise ratio (PSNR), the structural similarity index (SSIM), and the underwater image quality measure (UIQM). Our model is implemented using the PyTorch framework and was executed on an NVIDIA RTX 3080 GPU.

4.2. Experimental Evaluation on the USR-248 Dataset

The assessment of the USR-248 dataset is presented in Table 1. The proposed DAIN is compared to some popular SR networks, namely SRCNN [29], VDSR [31], DSRCNN [56], EDSRGAN [57], SRGAN [58], ESRGAN [59], SRDRM [37], SRDRM-GAN [37], LatticeNet [60], Deep WaveNet [12], ESRGCNN [61], AMPCNet [10], and RDLN [62]. We can observe that our proposed DAIN attains comparable results when compared to popular methods. In comparison to mainstream SR methods, our method obtains up to 0.08 dB and 0.1 increases in PSNR and SSIM. Although the DAIN fails to obtain optimal values in terms of UIQM, it still achieves positive results. For example, in the case of a scale factor of $\times 8$, compared to AMPCNet, our DAIN lags behind by 0.05 in terms of UIQM but boosts PSNR and SSIM by 0.14 dB and 0.02, respectively. Visual comparisons of the USR-248 dataset are shown in Figure 2. It is evident that our DAIN produces

clearer and more natural image reconstruction. SRDRM and SRDRM-GAN generate visible reconstruction artifacts.

Table 1. Quantitative results of different methods in the USR-248 dataset with scale factors of ×2, ×4, and ×8. Bold is the best performance.

Scale	Method	FLOPs (G)	Params (M)	PSNR (dB)	SSIM	UIQM
×2	SRCNN [29]	21.30	0.06	26.81	0.76	2.74
	VDSR [31]	205.28	0.67	28.98	0.79	2.57
	DSRCNN [56]	54.22	1.11	27.14	0.77	2.71
	EDSRGAN [57]	273.34	1.38	27.12	0.77	2.67
	SRGAN [58]	377.76	5.95	28.08	0.78	2.74
	ESRGAN [59]	4274.68	16.70	26.66	0.75	2.70
	SRDRM [37]	203.91	0.83	28.36	0.80	**2.78**
	SRDRM-GAN [37]	289.38	11.31	28.55	0.81	2.77
	LatticeNet [60]	56.84	0.76	29.47	0.80	2.65
	Deep WaveNet [12]	21.47	0.28	29.09	0.80	2.73
	AMPCNet [10]	—	1.15	29.54	0.80	2.77
	RDLN [62]	74.86	0.84	29.96	0.83	2.68
	DAIN(ours)	85.55	1.16	**29.98**	**0.84**	2.77
×4	SRCNN [29]	21.30	0.06	23.38	0.67	2.38
	VDSR [31]	205.28	0.67	25.70	0.68	2.44
	DSRCNN [56]	15.77	1.11	23.61	0.67	2.36
	EDSRGAN [57]	206.42	1.97	21.65	0.65	2.40
	SRGAN [58]	529.86	5.95	24.76	0.69	2.42
	ESRGAN [59]	1504.09	16.70	23.79	0.65	2.38
	SRDRM [37]	291.73	1.90	24.64	0.68	2.46
	SRDRM-GAN [37]	377.20	12.38	24.62	0.69	2.48
	LatticeNet [60]	14.61	0.78	26.06	0.65	2.43
	Deep WaveNet [12]	5.59	0.29	25.20	0.68	2.54
	AMPCNet [10]	—	1.17	25.90	0.68	**2.58**
	RDLN [62]	29.56	0.84	26.16	0.66	2.38
	DAIN (ours)	21.78	1.18	**26.23**	**0.70**	2.56
×8	SRCNN [29]	21.30	0.06	19.97	0.57	2.01
	VDSR [31]	205.28	0.67	23.58	0.63	2.17
	DSRCNN [56]	6.15	1.11	20.14	0.56	2.04
	EDSRGAN [57]	189.69	2.56	19.87	0.58	2.12
	SRGAN [58]	567.88	5.95	20.14	0.60	2.10
	ESRGAN [59]	811.44	16.70	19.75	0.58	2.05
	SRDRM [37]	313.68	2.97	21.20	0.60	2.18
	SRDRM-GAN [37]	399.15	13.45	20.25	0.61	2.17
	LatticeNet [60]	4.05	0.86	23.88	0.54	2.21
	Deep WaveNet [12]	1.62	0.34	23.25	0.62	2.21
	AMPCNet [10]	—	1.25	23.83	0.62	**2.25**
	RDLN [62]	18.23	0.84	23.91	0.54	2.18
	DAIN (ours)	5.99	1.26	**23.97**	**0.64**	2.20

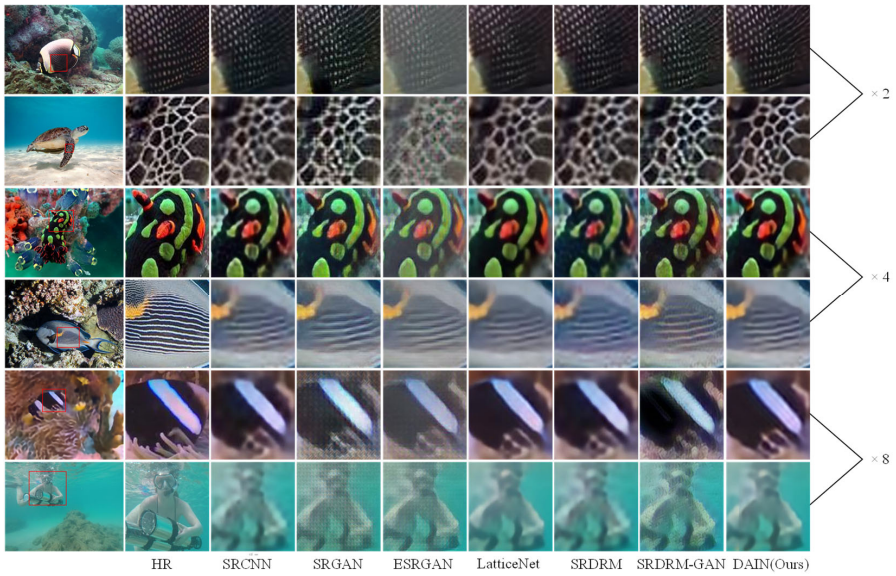

Figure 2. Visual comparisons between our DAIN and popular methods were conducted utilizing the USR-248 dataset for scale factors of ×2, ×4, and ×8. Left to right: the original HR image, SRCNN, SRGAN, ESRGAN, LatticeNet, SRDRM, SRDRM-GAN, and the proposed DAIN.

4.3. Experimental Evaluation on the UFO-120 Dataset

In the case of Table 2, our proposed DAIN demonstrates competitive performance, particularly achieving the highest values in terms of PSNR. Our DAIN is slightly inferior to Deep WaveNet in terms of SSIM and UIQM, but the gap is at most 0.05. The quantitative results of AMPCNet lag far behind our method. Compared with transformer-based methods such as URSCT and RDLN, our method yields a 0.42 dB improvement in PSNR at the scale factor of ×2. In the case of ×4, our DAIN has better performance, obtaining a 0.06 increase in SSIM compared to URSCT. Additionally, the DAIN does not surpass the RDLN in terms of UIQM, but the margin is 0.02 at a scale factor of ×4.

Table 2. Quantitative results of different methods in the UFO-120 dataset with scale factors of ×2, ×3, and ×4. Bold is the best performance.

Method	Params (M)			PSNR (dB)			SSIM			UIQM		
	×2	×3	×4	×2	×3	×4	×2	×3	×4	×2	×3	×4
SRCNN [29]	0.06	0.06	0.06	24.75	22.22	19.05	0.72	0.65	0.56	2.39	2.24	2.02
SRGAN [58]	5.95	5.95	5.95	26.11	23.87	21.08	0.75	0.70	0.58	2.44	2.39	2.56
SRDRM [37]	0.83	-	1.90	24.62	-	23.15	0.72	-	0.67	2.59	-	2.57
SRDRM-GAN [37]	11.31	-	12.38	24.61	-	23.26	0.72	-	0.67	2.59	-	2.55
Deep WaveNet [12]	0.28	0.28	0.29	25.71	25.23	25.08	0.77	**0.76**	**0.74**	**2.99**	2.96	**2.97**
AMPCNet [10]	1.15	1.16	1.17	25.24	25.73	24.70	0.71	0.70	0.70	2.93	2.85	2.88
ESRGCNN [59]	1.53	1.53	1.53	25.82	26.19	25.20	0.72	0.71	0.70	2.98	2.96	2.85
LatticeNet [60]	0.76	0.77	0.78	25.86	26.13	25.10	0.71	0.71	0.70	2.97	2.94	2.94
HNCT [63]	0.36	0.36	0.36	25.73	25.86	24.91	0.71	0.71	0.69	2.96	2.88	2.84
URSCT [46]	11.37	-	16.07	25.96	-	23.59	**0.80**	-	0.66	-	-	-
RDLN [62]	0.84	0.84	0.84	25.96	26.55	25.37	0.76	0.74	0.73	2.98	**2.98**	2.94
DAIN(ours)	**1.16**	**1.17**	**1.18**	**26.38**	**26.62**	**25.56**	0.76	**0.76**	0.72	2.97	2.94	2.92

The visual comparisons of the UFO-120 dataset are presented in Figure 3. Our proposed DAIN consistently delivers superior results and effectively recovers finer texture features, bringing it closest to the HR image. Notably, LatticeNet exhibits over-saturation, while AMPCNET and Deep WaveNet result in significant blurring artifacts and distortions. In contrast, our proposed method is better equipped to mitigate these adverse effects.

This is attributed to the fact that the DAIN comprises a series of chained DAEMs, which effectively capture both local and global image information, thereby enhancing the quality of image reconstruction.

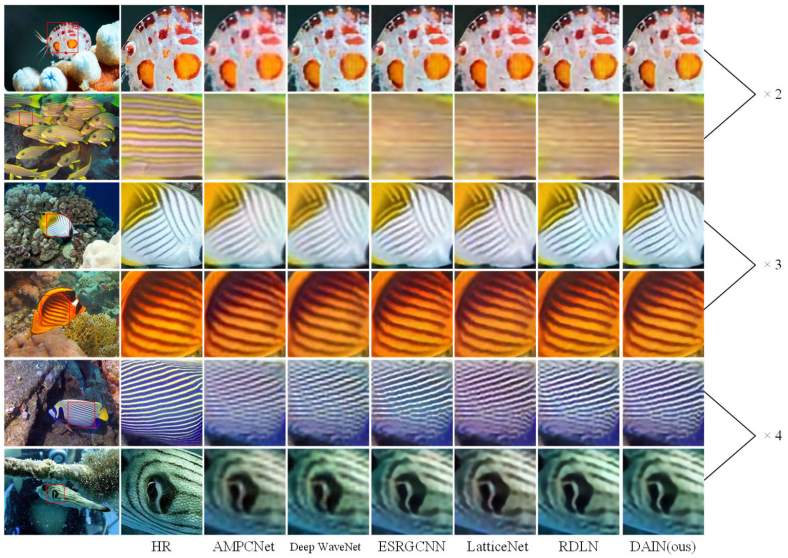

Figure 3. Visual comparisons between our DAIN and popular methods were conducted utilizing the UFO-120 dataset for scale factors of ×2, ×3, and ×4. Left to right: the original HR image, AMPCNet, Deep WaveNet, ESRGCNN, LatticeNet, RDLN, and the proposed DAIN.

4.4. Experimental Evaluation on the EUVP Dataset

To further show the robustness and effectiveness of the proposed DAIN, we performed enhancement experiments on the EUVP dataset and compared our model with models including UGAN [54], UGAN-P [54], Funie-GAN [52], Funie-GAN-UP [52], Deep SESR [53], and Deep WaveNet [12]. EUVP consists of 11,435 pairs of underwater images for training and 515 pairs of samples for testing. They follow the standard procedures [54,55] for optical/spatial image degradation and use manually labeled saliency mappings to create pairs of data.

Table 3 displays the competitive performance of the DAIN proposed by us, attaining the highest values in terms of both PSNR and SSIM. Specifically, PSNR and SSIM are improved by at least 0.57 dB and 0.02, respectively. The DAIN is found to be slightly lower than Deep SESR in terms of UIQM, but the difference is only 0.03. The quantitative results of Funie-GAN-UP are far below our method. A visual comparison of the EUVP dataset is shown in Figure 4. Our DAIN produces much sharper image reconstruction. Most algorithms, including UGAN, Funie-GAN, and Deep SESR, ignore global dependency modeling, resulting in poor performance. By contrast, our DAIN adequately incorporates global dependency modeling whilst also considering local information, enabling our network to acquire more valuable information and enhance image quality.

Table 3. Quantitative results of different methods in the EVUP dataset. Bold is the best performance.

Numbers	PSNR(dB)	SSIM	UIQM
UGAN [54]	26.45	0.79	2.87
UGAN-P [54]	26.44	0.79	2.91
Funie-GAN [52]	26.16	0.78	2.95
Funie-GAN-UP [52]	25.16	0.78	2.91
Deep SESR [53]	27.03	0.80	**3.06**
Deep WaveNet [12]	28.56	0.83	3.02
DAIN(ours)	**29.13**	**0.85**	3.03

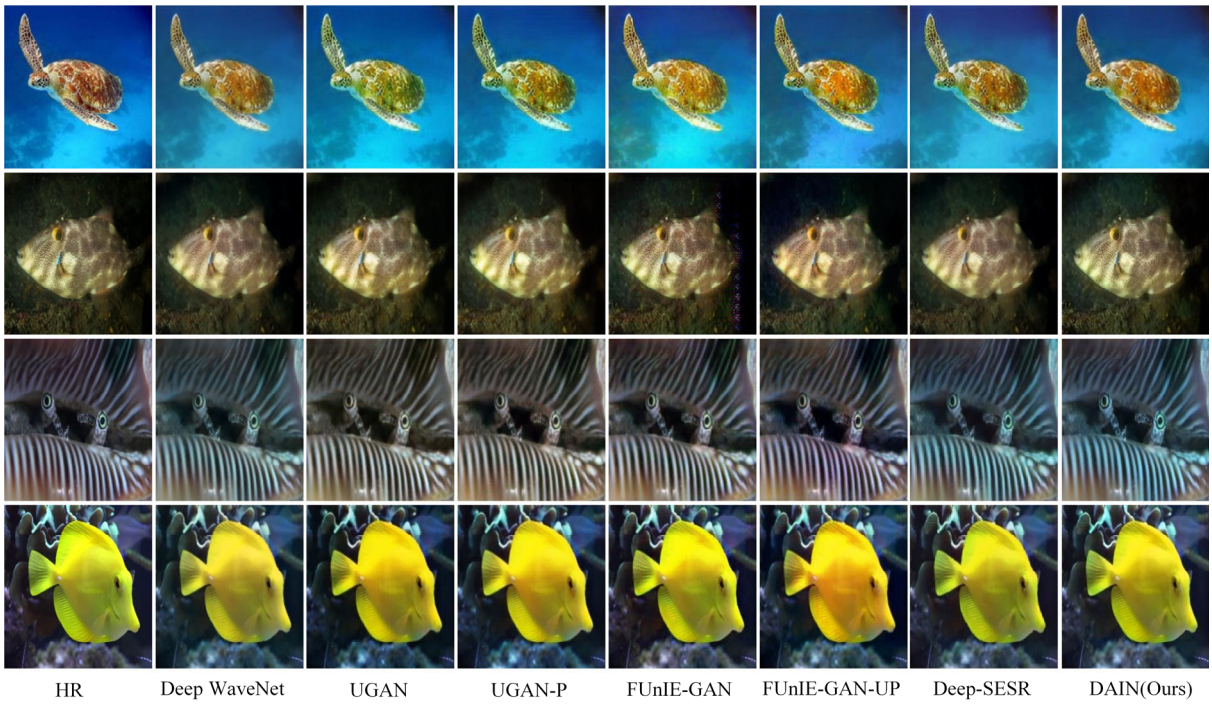

Figure 4. Visual comparisons between our DAIN and popular methods were conducted utilizing the EUVP dataset. Left to right: the original HR image, LR image, UGAN, UGAN-P, Funie-GAN, Funie-GAN-UP, Deep-SESR, and the proposed DAIN.

4.5. Model Analysis

Analysis of varying numbers of DAEMs. To explore the influence of model depth on reconstruction accuracy, we set the number of DAEMs to D = 10, 12, and 14. From Table 4, it can be seen that there is no significant improvement in the reconstruction results as depth increases. This is because the deeper the depth of the network, the more prone network weights are to deactivation, which makes it difficult to improve performance. Considering the size of the network and restoration accuracy, we opted for D = 12 as the number of DAEMs.

Table 4. Effect of the number of DAEMs on the USR-248 dataset at a scale factor of ×4.

Numbers	FLOPs (G)	PSNR (dB)	SSIM	UIQM
10	18.53	29.83	0.8177	2.7585
12	21.79	**29.86**	**0.8196**	2.7724
14	25.04	29.81	0.8177	**2.7803**

Analysis of edge detection. To further investigate different methods for preserving edge features in the restored images, we employed the Canny algorithm [64] to illustrate the benefits of our proposed method. As depicted in Figures 5 and 6, our DAIN exhibits more prominent edge detection features compared to the majority of methods. These outcomes demonstrate the efficiency of our DAIN in extracting feature information and aiding in the recovery of additional texture details, resulting in visually pleasing effects.

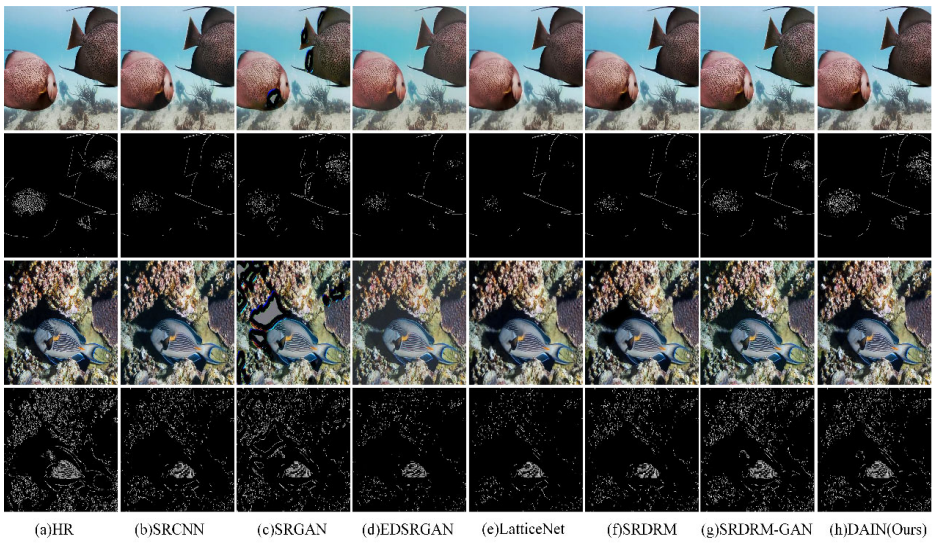

Figure 5. Canny edge detection on the USR-248 dataset.

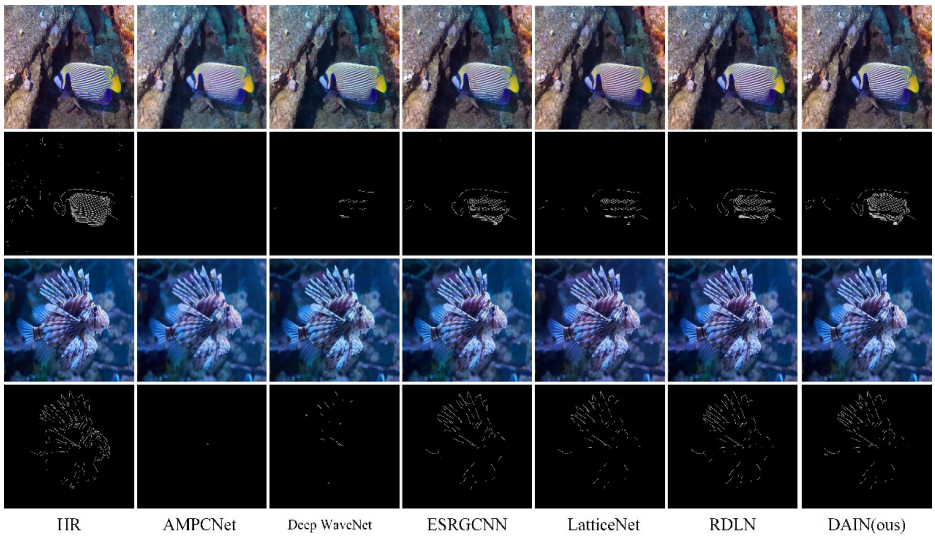

Figure 6. Canny edge detection on the UFO-120 dataset.

4.6. Ablation Study

To clearly show how our proposed components can improve reconstruction performance, we conducted various experiments, as documented in Table 4. We conducted a stepwise removal of the MCCB, the STL, and the cascades within the MCCB to illustrate the effectiveness of each component. Notably, we modified the MCCB, where we changed the different convolutional sizes to a uniform use of 3×3 convolution for feature extraction, and named it as a DAIN with a PCB. We retrained the network and obtained another four models.

As is evident in Table 5, when the MCCB is erased, reconstruction performance drops drastically, with PSNR and SSIM dropping by at least 0.06 dB and 0.0029. It can be observed that both the DAIN without an MCCB and the DAIN with a PCB have worse performance than the DAIN alone. Similarly, although the number of parameters decreases after removing the STL and concatenation, model performance decreases substantially, with PSNR and SSIM decreasing by at least 0.15 dB and 0.0187, respectively. In summary, our proposed MCCB, STL, and concatenation components have a beneficial effect on both the recovery performance and computational efficiency of the network.

Table 5. Ablation studies of different components on the UFO-120 dataset at a scale factor of ×4.

Models	Params (M)	PSNR (dB)	SSIM	UIQM
DAIN w/o MCCB	0.93	25.35	0.7132	2.91
DAIN with PCB	1.54	25.34	0.7143	2.91
DAIN w/o STL	0.67	25.26	0.6974	2.89
DAIN w/o DFFB	1.02	24.79	0.6961	2.89
DAIN	1.16	25.41	0.7161	2.92

5. Conclusions

Underwater images display unique characteristics such as light absorption, scattering, and color attenuation, unlike terrestrial images. To bridge this gap, our study introduces an effective approach called the DAIN for SR reconstruction of underwater imagery. Specifically, the DAIN encompasses a series of DAEMs that tackle image distortion and blurriness by exploring both local and global features. MCCBs examine various colors in an underwater image using multiple channels to reduce color bias and restore texture details. Meanwhile, the STL is capable of modeling global features to enhance high-quality images even further. Thanks to the incorporation of an MCCB and an STL in the DAEM, the network can considerably improve the accuracy of underwater SR. It is proven that our model has higher robustness and efficiency. Through experimentation on benchmark datasets, findings demonstrate that the proposed DAIN achieves competitive performance while requiring fewer parameters compared to other popular methods.

Author Contributions: Conceptualization, A.S. and H.D.; methodology, A.S. and H.D.; software, H.D.; validation, H.D.; formal analysis, H.D.; investigation, H.D.; resources, H.D.; data curation, H.D.; writing—original draft preparation, H.D.; writing—review and editing, A.S. and H.D. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data provided in this study are publicly available at: https://irvlab.cs.umn.edu/resources/ufo-120-dataset (accessed on 1 November 2023) and https://irvlab.cs.umn.edu/resources/usr-248-dataset (accessed on 1 November 2023) and https://irvlab.cs.umn.edu/resources/euvp-dataset (accessed on 1 November 2023).

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

References

1. Anwar, S.; Khan, S.; Barnes, N. A Deep Journey into Super-Resolution: A Survey. *ACM Comput. Surv.* **2020**, *53*, 60. [CrossRef]
2. Yang, W.; Zhang, X.; Tian, Y.; Wang, W.; Xue, J.-H.; Liao, Q. Deep Learning for Single Image Super-Resolution: A Brief Review. *IEEE Trans. Multimed.* **2019**, *21*, 3106–3121. [CrossRef]
3. Li, C.-Y.; Guo, J.-C.; Cong, R.-M.; Pang, Y.-W.; Wang, B. Underwater Image Enhancement by Dehazing with Minimum Information Loss and Histogram Distribution Prior. *IEEE Trans. Image Process.* **2016**, *25*, 5664–5677. [CrossRef] [PubMed]
4. Ancuti, C.; Ancuti, C.O.; Haber, T.; Bekaert, P. Enhancing Underwater Images and Videos by Fusion. In Proceedings of the 2012 IEEE Conference on Computer Vision and Pattern Recognition, Providence, RI, USA, 16–21 June 2012; pp. 81–88.
5. Abdul Ghani, A.S.; Mat Isa, N.A. Underwater Image Quality Enhancement through Composition of Dual-Intensity Images and Rayleigh-Stretching. *SpringerPlus* **2014**, *3*, 757. [CrossRef] [PubMed]
6. Wang, L.; Li, K.; Tang, J.; Liang, Y. Image Super-Resolution via Lightweight Attention-Directed Feature Aggregation Network. *ACM Trans. Multimed. Comput. Commun. Appl.* **2023**, *19*, 60. [CrossRef]
7. Jobson, D.J.; Rahman, Z.; Woodell, G.A. A Multiscale Retinex for Bridging the Gap between Color Images and the Human Observation of Scenes. *IEEE Trans. Image Process.* **1997**, *6*, 965–976. [CrossRef]
8. Cho, Y.; Jeong, J.; Kim, A. Model-Assisted Multiband Fusion for Single Image Enhancement and Applications to Robot Vision. *IEEE Robot. Autom. Lett.* **2018**, *3*, 2822–2829.
9. 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]

10. Zhang, Y.; Yang, S.; Sun, Y.; Liu, S.; Li, X. Attention-Guided Multi-Path Cross-CNN for Underwater Image Super-Resolution. *Signal Image Video Process.* **2022**, *16*, 155–163. [CrossRef]
11. Wang, H.; Wu, H.; Hu, Q.; Chi, J.; Yu, X.; Wu, C. Underwater Image Super-Resolution Using Multi-Stage Information Distillation Networks. *J. Vis. Commun. Image Represent.* **2021**, *77*, 103136. [CrossRef]
12. Sharma, P.; Bisht, I.; Sur, A. Wavelength-Based Attributed Deep Neural Network for Underwater Image Restoration. *ACM Trans. Multimed. Comput. Commun. Appl.* **2023**, *19*, 2. [CrossRef]
13. Woo, S.; Park, J.; Lee, J.-Y.; Kweon, I.S. CBAM: Convolutional Block Attention Module. In Proceedings of the European Conference on Computer Vision (ECCV), Munich, Germany, 8–14 September 2018; Ferrari, V., Hebert, M., Sminchisescu, C., Weiss, Y., Eds.; Lecture Notes in Computer Science. Springer: Cham, Switzerland, 2018; Volume 11211, pp. 3–19, ISBN 978-3-030-01233-5.
14. Wang, L.; Xu, L.; Tian, W.; Zhang, Y.; Feng, H.; Chen, Z. Underwater Image Super-Resolution and Enhancement via Progressive Frequency-Interleaved Network. *J. Vis. Commun. Image Represent.* **2022**, *86*, 103545. [CrossRef]
15. Wang, Y.; Guo, J.; Gao, H.; Yue, H. UIEC^2-Net: CNN-Based Underwater Image Enhancement Using Two Color Space. *Signal Process. Image Commun.* **2021**, *96*, 116250. [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. Liu, K.; Liang, Y. Underwater Optical Image Enhancement Based on Super-Resolution Convolutional Neural Network and Perceptual Fusion. *Opt. Express* **2023**, *31*, 9688. [CrossRef]
18. 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 2021 IEEE/CVF International Conference on Computer Vision (ICCV), Montreal, QC, Canada, 10–17 October 2021; pp. 9992–10002.
19. Peng, L.; Zhu, C.; Bian, L. U-Shape Transformer for Underwater Image Enhancement. *IEEE Trans. Image Process.* **2023**, *32*, 3066–3079. [CrossRef] [PubMed]
20. Shen, Z.; Xu, H.; Luo, T.; Song, Y.; He, Z. UDAformer: Underwater Image Enhancement Based on Dual Attention Transformer. *Comput. Graph.* **2023**, *111*, 77–88. [CrossRef]
21. Liang, J.; Cao, J.; Sun, G.; Zhang, K.; Van Gool, L.; Timofte, R. SwinIR: Image Restoration Using Swin Transformer. In Proceedings of the 2021 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW), Montreal, BC, Canada, 10–17 October 2021; pp. 1833–1844.
22. Huang, Z.; Li, J.; Hua, Z.; Fan, L. Underwater Image Enhancement via Adaptive Group Attention-Based Multiscale Cascade Transformer. *IEEE Trans. Instrum. Meas.* **2022**, *71*, 5015618. [CrossRef]
23. Guo, Z.; Guo, D.; Gu, Z.; Zheng, H.; Zheng, B.; Wang, G. Unsupervised Underwater Image Clearness via Transformer. In Proceedings of the OCEANS 2022—Chennai, Chennai, India, 21–24 February 2022; pp. 1–4.
24. Lu, T.; Wang, Y.; Zhang, Y.; Wang, Y.; Wei, L.; Wang, Z.; Jiang, J. Face Hallucination via Split-Attention in Split-Attention Network. In Proceedings of the 29th ACM International Conference on Multimedia, Virtual Event, 20–24 October 2021; pp. 5501–5509.
25. Zhang, D.; Shao, J.; Liang, Z.; Gao, L.; Shen, H.T. Large Factor Image Super-Resolution with Cascaded Convolutional Neural Networks. *IEEE Trans. Multimed.* **2020**, *23*, 2172–2184. [CrossRef]
26. Wang, J.; Shao, Z.; Huang, X.; Lu, T.; Zhang, R.; Ma, J. Enhanced Image Prior for Unsupervised Remoting Sensing Super-Resolution. *Neural Netw.* **2021**, *143*, 400–412. [CrossRef]
27. Ouyang, D.; Shao, J.; Jiang, H.; Nguang, S.K.; Shen, H.T. Impulsive Synchronization of Coupled Delayed Neural Networks with Actuator Saturation and Its Application to Image Encryption. *Neural Netw.* **2020**, *128*, 158–171. [CrossRef] [PubMed]
28. Ouyang, D.; Zhang, Y.; Shao, J. Video-Based Person Re-Identification via Spatio-Temporal Attentional and Two-Stream Fusion Convolutional Networks. *Pattern Recognit. Lett.* **2019**, *117*, 153–160. [CrossRef]
29. Dong, C.; Loy, C.C.; He, K.; Tang, X. Learning a Deep Convolutional Network for Image Super-Resolution. In Proceedings of the Computer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland, 6–12 September 2014; Part IV 13. Springer: Berlin/Heidelberg, Germany, 2014; pp. 184–199.
30. Dong, C.; Loy, C.C.; Tang, X. Accelerating the Super-Resolution Convolutional Neural Network. In Proceedings of the Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, 11–14 October 2016; Part II 14. Springer: Berlin/Heidelberg, Germany, 2016; pp. 391–407.
31. Kim, J.; Lee, J.K.; Lee, K.M. Accurate Image Super-Resolution Using Very Deep Convolutional Networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 1646–1654.
32. Lai, W.-S.; Huang, J.-B.; Ahuja, N.; Yang, M.-H. Deep Laplacian Pyramid Networks for Fast and Accurate Super-Resolution. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 624–632.
33. Kim, J.; Lee, J.K.; Lee, K.M. Deeply-Recursive Convolutional Network for Image Super-Resolution. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 1637–1645.
34. Tong, T.; Li, G.; Liu, X.; Gao, Q. Image Super-Resolution Using Dense Skip Connections. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017; pp. 4799–4807.
35. Tai, Y.; Yang, J.; Liu, X. Image Super-Resolution via Deep Recursive Residual Network. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 3147–3155.

36. Shi, W.; Caballero, J.; Huszár, F.; Totz, J.; Aitken, A.P.; Bishop, R.; Rueckert, D.; Wang, Z. Real-Time Single Image and Video Super-Resolution Using an Efficient Sub-Pixel Convolutional Neural Network. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 1874–1883.
37. Islam, M.J.; Enan, S.S.; Luo, P.; Sattar, J. Underwater Image Super-Resolution Using Deep Residual Multipliers. In Proceedings of the 2020 IEEE International Conference on Robotics and Automation (ICRA), Paris, France, 31 May–31 August 2020; pp. 900–906.
38. 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.
39. Cherian, A.K.; Poovammal, E. A Novel AlphaSRGAN for Underwater Image Super Resolution. *Comput. Mater. Contin.* **2021**, *69*, 1537–1552. [CrossRef]
40. Li, J.; Skinner, K.A.; Eustice, R.M.; Johnson-Roberson, M. WaterGAN: Unsupervised Generative Network to Enable Real-Time Color Correction of Monocular Underwater Images. *IEEE Robot. Autom. Lett.* **2017**, *3*, 387–394. [CrossRef]
41. Hou, M.; Liu, R.; Fan, X.; Luo, Z. Joint Residual Learning for Underwater Image Enhancement. In Proceedings of the 2018 25th IEEE International Conference on Image Processing (ICIP), Athens, Greece, 7–10 October 2018; pp. 4043–4047.
42. Simonyan, K.; Zisserman, A. Very Deep Convolutional Networks for Large-Scale Image Recognition. *arXiv* **2014**, arXiv:1409.1556.
43. Li, H.; Zhang, C.; Wan, N.; Chen, Q.; Wang, D.; Song, D. An Improved Method for Underwater Image Super-Resolution and Enhancement. In Proceedings of the 2021 IEEE 4th International Conference on Electronics Technology (ICET), Chengdu, China, 7–10 May 2021; pp. 1295–1299.
44. Dosovitskiy, A.; Beyer, L.; Kolesnikov, A.; Weissenborn, D.; Zhai, X.; Unterthiner, T.; Dehghani, M.; Minderer, M.; Heigold, G.; Gelly, S.; et al. An Image Is Worth 16x16 Words: Transformers for Image Recognition at Scale. *arXiv* **2020**, arXiv:2010.11929.
45. Wang, Z.; Cun, X.; Bao, J.; Zhou, W.; Liu, J.; Li, H. Uformer: A General u-Shaped Transformer for Image Restoration. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans, LA, USA, 18–24 June 2022; pp. 17683–17693.
46. Ren, T.; Xu, H.; Jiang, G.; Yu, M.; Zhang, X.; Wang, B.; Luo, T. Reinforced Swin-Convs Transformer for Simultaneous Underwater Sensing Scene Image Enhancement and Super-Resolution. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 4209616. [CrossRef]
47. Zhang, Y.; Chen, D.; Zhang, Y.; Shen, M.; Zhao, W. A Two-Stage Network Based on Transformer and Physical Model for Single Underwater Image Enhancement. *J. Mar. Sci. Eng.* **2023**, *11*, 787. [CrossRef]
48. Sun, K.; Meng, F.; Tian, Y. Underwater Image Enhancement Based on Noise Residual and Color Correction Aggregation Network. *Digit. Signal Process.* **2022**, *129*, 103684. [CrossRef]
49. Khan, M.R.; Kulkarni, A.; Phutke, S.S.; Murala, S. Underwater Image Enhancement with Phase Transfer and Attention. In Proceedings of the 2023 International Joint Conference on Neural Networks (IJCNN), Gold Coast, Australia, 18–23 June 2023; pp. 1–8.
50. Qi, Q.; Li, K.; Zheng, H.; Gao, X.; Hou, G.; Sun, K. SGUIE-Net: Semantic Attention Guided Underwater Image Enhancement with Multi-Scale Perception. *IEEE Trans. Image Process.* **2022**, *31*, 6816–6830. [CrossRef]
51. Zhao, H.; Gallo, O.; Frosio, I.; Kautz, J. Loss Functions for Image Restoration with Neural Networks. *IEEE Trans. Comput. Imaging* **2016**, *3*, 47–57. [CrossRef]
52. Islam, M.J.; Xia, Y.; Sattar, J. Fast Underwater Image Enhancement for Improved Visual Perception. *IEEE Robot. Autom. Lett.* **2020**, *5*, 3227–3234. [CrossRef]
53. Islam, M.J.; Luo, P.; Sattar, J. Simultaneous Enhancement and Super-Resolution of Underwater Imagery for Improved Visual Perception. *arXiv* **2020**, arXiv:2002.01155.
54. 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, Australia, 21–25 May 2018; pp. 7159–7165.
55. Li, Z.; Yang, J.; Liu, Z.; Yang, X.; Jeon, G.; Wu, W. Feedback Network for Image Super-Resolution. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 15–20 June 2019; pp. 3867–3876.
56. Mao, X.-J.; Shen, C.; Yang, Y.-B. Image Restoration Using Convolutional Auto-Encoders with Symmetric Skip Connections. *arXiv* **2016**, arXiv:1606.08921.
57. Lim, B.; Son, S.; Kim, H.; Nah, S.; Mu Lee, K. Enhanced Deep Residual Networks for Single Image Super-Resolution. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, Honolulu, HI, USA, 21–26 July 2017; pp. 136–144.
58. Ledig, C.; Theis, L.; Huszár, F.; Caballero, J.; Cunningham, A.; Acosta, A.; Aitken, A.; Tejani, A.; Totz, J.; Wang, Z.; et al. Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 4681–4690.
59. Wang, X.; Yu, K.; Wu, S.; Gu, J.; Liu, Y.; Dong, C.; Qiao, Y.; Change Loy, C. Esrgan: Enhanced Super-Resolution Generative Adversarial Networks. In Proceedings of the European conference on computer vision (ECCV) workshops, Munich, Germany, 8–14 September 2018.
60. Luo, X.; Xie, Y.; Zhang, Y.; Qu, Y.; Li, C.; Fu, Y. Latticenet: Towards Lightweight Image Super-Resolution with Lattice Block. In Proceedings of the Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, 23–28 August 2020; Part XXII 16. Springer: Berlin/Heidelberg, Germany, 2020; pp. 272–289.
61. Tian, C.; Yuan, Y.; Zhang, S.; Lin, C.-W.; Zuo, W.; Zhang, D. Image Super-Resolution with an Enhanced Group Convolutional Neural Network. *Neural Netw.* **2022**, *153*, 373–385. [CrossRef]

62. Chen, Z.; Liu, C.; Zhang, K.; Chen, Y.; Wang, R.; Shi, X. Underwater-Image Super-Resolution via Range-Dependency Learning of Multiscale Features. *Comput. Electr. Eng.* **2023**, *110*, 108756. [CrossRef]
63. Fang, J.; Lin, H.; Chen, X.; Zeng, K. A Hybrid Network of Cnn and Transformer for Lightweight Image Super-Resolution. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans, LA, USA, 18–24 June 2022; pp. 1103–1112.
64. Canny, J. A Computational Approach to Edge Detection. *IEEE Trans. Pattern Anal. Mach. Intell.* **1986**, *PAMI-8*, 679–698. [CrossRef]