Journal of *Marine Science and Engineering*

MDPI

Article

Efficient Underwater Image Super-Resolution via Learnable Color Conversion and Dual-Branch Mamba Network

Yu-Yang Lin [1], Wan-Jen Huang [1], Chia-Hung Yeh [2,3,*], Yi-Shiuan Yang [2] and Chua-Chin Wang [3,4,5]

[1] Institute of Communications Engineering, National Sun Yat-sen University, Kaohsiung 80404, Taiwan
[2] Department of Electrical Engineering, National Taiwan Normal University, Taipei 10610, Taiwan
[3] Department of Electrical Engineering, National Sun Yat-sen University, Kaohsiung 80404, Taiwan
[4] Institute of IC Design, National Sun Yat-Sen University, Kaohsiung 80404, Taiwan
[5] Institute of Undersea Technology, National Sun Yat-sen University, Kaohsiung 80404, Taiwan
* Correspondence: chyeh@ntnu.edu.tw or yeh@mail.ee.nsysu.edu.tw

Abstract

Underwater image super-resolution plays an important role in marine exploration, as it aims to recover clearer and higher-resolution images from degraded low-resolution observations. However, in underwater environments, this task is particularly challenging due to non-uniform spectral attenuation and particle scattering. Underwater, red light fades quickly, leading to color distortion and loss of details. To address this while keeping the model lightweight, we propose a dynamic color-space network for single-image super-resolution. Traditional methods use fixed color conversion and cannot handle non-uniform attenuation well. We instead use a learnable module to adaptively obtain luminance (Y) and chrominance (CbCr) representations. Building upon this representation, we employ a highly efficient dual-branch architecture where the Y-channel features are processed using light Mamba blocks to capture spatial dependencies with linear complexity, while the CbCr-channel features are restored by a lightweight convolutional neural network. Finally, a lightweight residual model utilizing depth-wise separable convolutions (DSC), deformable convolutions, and pixel attention mechanisms is introduced to refine the features and suppress artifacts. Experimental results demonstrate that while achieving competitive restoration quality, the proposed method drastically reduces computational complexity and parameter count compared to other large-scale models. This balance between visual quality and computational efficiency makes the proposed method well-suited for real-time deployment on resource-constrained autonomous underwater vehicles (AUVs).

Keywords: underwater image restoration; deep learning; supervised learning; super resolution

Check for updates

Academic Editor: Weicheng Cui

Received: 24 May 2026
Revised: 25 June 2026
Accepted: 28 June 2026
Published: 30 June 2026

Copyright: © 2026 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.

1. Introduction

Because land resources become limited and the global economy grows, deep-sea exploration has become more important. In ocean engineering, visual data from AUVs is widely used for tasks such as seafloor mapping and navigation. Previous studies have introduced various image processing techniques to facilitate deep-sea scientific exploration [1,2]. However, underwater images are usually more blurred and lower in resolution than images taken in the air, mainly because of light scattering and absorption. Because of this, the captured images often lose fine details, which makes accurate target detection and scene understanding more difficult. While the emergence of deep learning has revolutionized

image restoration [3], the performance of these models heavily depends on the availability of extensive training datasets.

Underwater image restoration has branched into conventional techniques and deep learning-based approaches. Early works, such as those by Chiang et al. [4] and Li et al. [5], focused on physical models to address attenuation and dehazing. While fusion-based frameworks [6] have been effective for color compensation, they often struggle to recover high-frequency textures essential for high-resolution tasks. Following the widespread application of deep learning, several convolutional neural network (CNN) models have been introduced, including Deep SESR [7] and Lit-Net [8]. However, Generative Adversarial Network (GAN)-based frameworks like WaterGAN [9] and CycleGAN [10], while yielding visually pleasing results, are prone to model collapse and often fail to maintain structural integrity. To solve these problems, models like Dual-CycleGAN [11] have been developed. They use light-field information to make the restoration more stable.

Image restoration can improve degraded images, but they often miss fine details. To address this, underwater image super-resolution has been studied to further enhance resolution and recover structural information. Recently, more work has focused on reconstructing high-resolution images from degraded inputs. Early studies primarily extended conventional super-resolution techniques to underwater scenarios. For example, methods based on the SRCNN [12] were directly applied to underwater images to recover spatial details, demonstrating the feasibility of deep learning-based SR in aquatic environments. Subsequently, more specialized approaches were proposed to better handle underwater degradations. Islam et al. developed residual-based generative models, including SRDRM [13] and SRDRM-GAN [13], to enhance image quality and restore details in underwater scenes. They later proposed Deep SESR [7], which addresses both color distortion and resolution degradation together. More recently, various advanced architectures have been developed to further tackle complex underwater degradations. For instance, generative models such as SRGAN [14] were adapted to enhance perceptual realism. To capture multi-scale features and wavelength-dependent attenuation, attention mechanisms and wave-based networks, such as AMPCNet [15] and Deep WaveNet [16], have been utilized. Transformer-based models like URSCT [17] have been used to capture long-range information. Methods such as DSMN [18] focus on recovering structural details and fine textures.

These methods use techniques like attention and transformers to improve features, but most are still based on general image restoration models. These generic strategies implicitly assume that degradation is uniformly distributed across all color channels. However, an underwater image is fundamentally affected by wavelength-dependent light absorption, which creates a severe spectral imbalance. Because existing methods fail to explicitly decouple and model this non-uniform degradation, they often attempt to recover visual sharpness by stacking massive network layers. General color-space learning methods typically address uniform illumination changes or standard atmospheric scattering. In contrast, our module is explicitly designed to mitigate underwater non-uniform spectral attenuation, where red light degrades significantly faster than green and blue light. Furthermore, unlike existing methods that apply color transformations within heavy convolutional backbones, our conversion uniquely serves as a feature decoupling mechanism. It isolates the luminance component so that the Light Mamba blocks [19] can model spatial dependencies with linear complexity, while delegating the chrominance features to a lightweight CNN. Consequently, although these large-scale deep learning models can improve restoration quality, they require massive computational resources and parameter counts. This heavy computational burden severely limits their practical deployment on platforms with limited computational resources like AUVs.

To tackle these challenges while maintaining extreme computational efficiency, this paper presents a lightweight dynamic color-space learning framework integrated with a dual-branch Mamba network for single underwater image super-resolution. Instead of relying on standard fixed YCbCr matrices that cannot adapt to non-uniform underwater spectral attenuation, we introduce a learnable color conversion module to dynamically optimize the weights based on the input image content. This adaptive conversion effectively decouples the degradation process into optimal luminance and chrominance representations without requiring a computationally heavy backbone, facilitating more accurate restoration of structural details and color information. Based on this decomposition, our model then employs a highly efficient dual-branch upsampling architecture, where the long-range spatial dependencies of Y-channel features are modeled using light Mamba blocks [19] to maintain linear computational complexity, while the CbCr-channel features are restored by a lightweight convolutional neural network. Furthermore, a lightweight residual refinement model utilizing depth-wise separable convolutions, deformable convolutions, and pixel attention mechanisms is introduced to refine the fused features and suppress artifacts. Extensive experiments on the UFO-120 dataset [7] demonstrate that the proposed Mamba-enhanced method significantly reduces computational complexity as well as total parameters relative to prominent high-capacity models, while achieving competitive restoration quality in terms of PSNR, SSIM, and UIQM [20]. This optimal balance makes it highly suitable for real-world oceanic deployments.

The rest of this paper is organized as follows. Section 2 introduces super-resolution backbones and recent work on state–space models as the foundation of our method. Section 3 presents the proposed multi-stage residual refinement framework, including its formulation and the Mamba-based residual design. Section 4 provides experimental results along with comparative analysis. Concluding remarks are made in Section 5.

2. Related Work

2.1. Lightweight Deep Learning-Based Super-Resolution

In contrast to open-air photography, restoring subsea environments presents fundamentally different reconstruction challenges because of complex degradation, such as light scattering and absorption, making it difficult to accurately recover high-frequency details with traditional interpolation methods. To resolve these issues, deep learning methods for super-resolution approaches, starting from early frameworks like SRCNN [12], have been introduced to capture and reconstruct detailed spatial information. To further enhance efficiency, subsequent research introduced techniques such as sub-pixel convolution, commonly known as pixel shuffle [21], which significantly reduces computational complexity by performing feature extraction in the low-resolution space and shifting the upsampling operation to the end of the network. Furthermore, information distillation mechanisms, such as those proposed in IMDN [22], have been adopted to progressively extract useful features while discarding redundant information, thereby maintaining a compact model size. However, many recent State-of-the-Art super-resolution models rely on very deep and complex networks, resulting in massive computational overhead and huge parameter counts. These large models can recover fine details, but they are not suitable for real-time use on resource-limited platforms. To reduce computation, many lightweight methods separate image processing into different parts, such as treating luminance and chrominance differently. A common strategy used in models like SRCNN [12] and FSRCNN [23] is to convert images to YCbCr and focus most of the computation on the luminance channel, while keeping the chrominance processing simple. By utilizing such color-space conversions and lightweight convolutional designs, networks can strategically allocate their limited computational capacity to the most perceptually critical components. In the proposed

method, we do not use heavy backbones. Instead, we design a dual-branch upsampling network that works on dynamically learned Y and CbCr channels. With pixel shuffle, the model can recover structural details while remaining efficient.

Recent advancements have further introduced frameworks such as DeepAquaEnhance, which utilizes hierarchical multi-head attention to address complex underwater illumination, and self-attention-guided approaches that integrate depth-aware priors. Additionally, multi-conditional diffusion probabilistic models like MCLL-Diff have shown strong capabilities in handling severe visual degradation and illumination imbalance. Other notable recent advances include dynamic lightweight super-resolution frameworks [24], scale-aware attention mechanisms [25], frequency-aware refinement strategies [26], and deformable multi-scale feature fusion frameworks [27]. Furthermore, advanced image restoration models leveraging channel-attention [28,29] and physics-aware color and illumination modeling approaches [30] have provided critical insights into addressing underwater spectral attenuation.

2.2. State–Space Models and Mamba

Mamba, extended from State–Space Models (SSM), which was originally proposed by Gu and Dao [19], aims at enhancing long-range dependency modeling with linear time complexity. Recently, Mamba-based architectures have been applied to several image restoration tasks, achieving good results in capturing global context [31]. Unlike traditional transformers that suffer from quadratic complexity, Mamba learns a selective mechanism that can infer complex spatial mappings efficiently, showing good performance in high-resolution image processing. Recent Mamba-based visual representation frameworks [32] have further established the efficacy of state-space models in complex visual tasks. Although deep learning has worked well for different underwater applications, capturing global dependencies remains a significant challenge due to the non-uniform blurring in underwater environments. This aligns with the ongoing demand for lightweight deep neural networks in underwater exploration, where joint learning of color conversion and task-specific goals (e.g., object detection) has proven effective for resource-constrained platforms [33]. Mamba-based methods are particularly suitable for underwater super-resolution as they provide an optimal trade-off between receptive field size and computational cost. This perfectly aligns with the stringent requirements for lightweight deep neural networks in real-time underwater exploration. In our framework, we specifically deploy Light Mamba blocks [19] within the luminance Y-channel upsampling branch. This integration empowers the model to extract long-range spatial structural dependencies under linear time complexity, contributing to the extremely low floating-point operations (FLOPs) of our system while maintaining high visual quality.

3. Proposed Framework

Underwater images are often affected by color distortion and loss of details, mainly due to light absorption and scattering. Existing deep learning models can restore these images, but they often depend on RGB processing and large networks, making them hard to use in real-time applications. To address this, we propose a lightweight dynamic color-space learning framework that decouples the restoration of luminance and chrominance. Given a low-resolution underwater input image, $RGB_{LR} \in \mathbb{R}^{3\times H\times W}$, the overall architecture is specifically designed to minimize parameters and FLOPs while preserving high-frequency details. As illustrated in Figure 1, the system progressively reconstructs the final target high-resolution image $RGB_{final} \in \mathbb{R}^{3\times 2H\times 2W}$ through three primary stages: dynamic color-space conversion, a dual-branch upsampling network, and lightweight residual refinement.

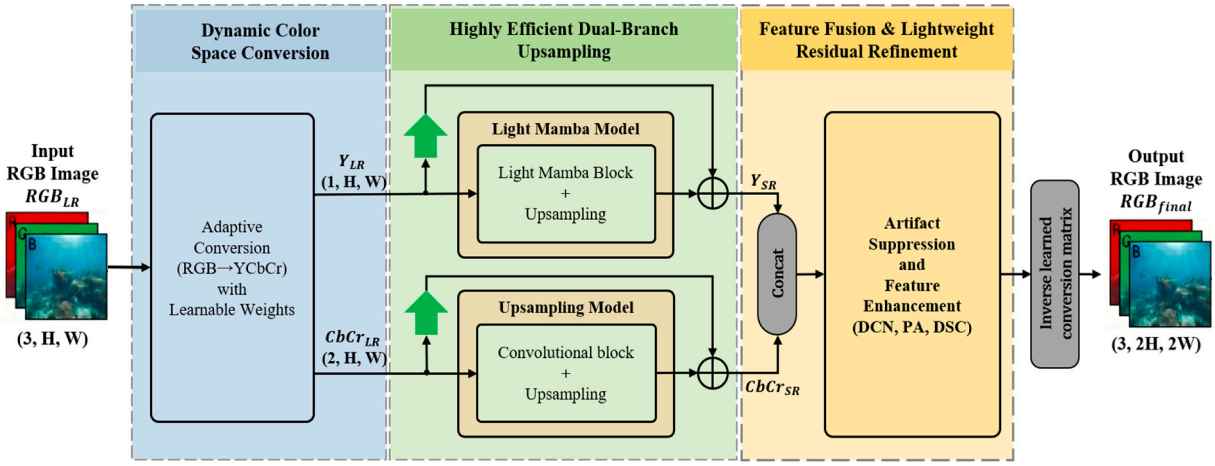


Figure 1. The overall architecture of the proposed framework.

3.1. Dynamic Color-Space Conversion

Instead of employing a fixed transformation matrix, we introduce a learnable color conversion module to adaptively extract the optimal luminance (Y_{LR}) and chrominance ($CbCr_{LR}$) representations based on the input underwater image RGB_{LR}. In traditional image processing, RGB is often converted to YCbCr using a fixed matrix. According to the ITU-R BT.601 [34] standard, the luminance and chrominance components are defined as:

$$Y = 0.299 \cdot R + 0.587 \cdot G + 0.114 \cdot B, \tag{1}$$

$$Cb = \frac{B - Y}{1.772} + 0.5, \tag{2}$$

$$Cr = \frac{R - Y}{1.402} + 0.5. \tag{3}$$

These fixed coefficients are determined based on human visual perception under standard atmospheric lighting conditions, assuming uniform light propagation. However, underwater imaging environments exhibit non-uniform spectral attenuation, where red light degrades significantly faster than green and blue light. Because of this, a fixed matrix cannot deal with the different color distortions in underwater images. We therefore use a learnable color conversion module that adapts the weights based on the input image. This helps the model better extract useful features and handle wavelength-related degradation. The structure of this module is shown in Figure 2.

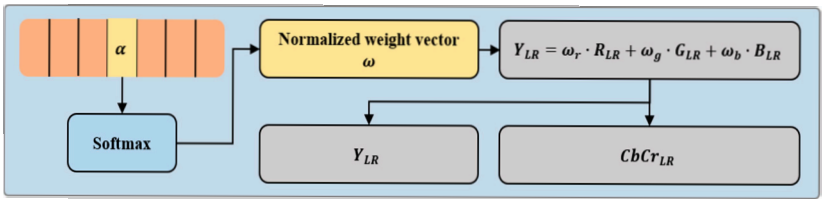


Figure 2. Dynamic color-space conversion.

The adaptive conversion process is driven by a learnable parameter vector α. To ensure proper scaling and stability, this vector is passed through a Softmax activation layer to generate a normalized weight vector ω. This dynamic weight vector, denoted as $\omega = [\omega_r, \omega_g, \omega_b]$, is then utilized to perform a channel-wise weighted summation over the input RGB channels (R_{LR}, G_{LR}, B_{LR}) of RGB_{LR} This process adaptively synthesizes the luminance component through the formulation $Y_{LR} = \omega_r \cdot R_{LR} + \omega_g \cdot G_{LR} + \omega_b \cdot B_{LR}$, which carries the primary structural information of the underwater scene. Concurrently,

the remaining transformation paths generate the corresponding chrominance components $CbCr_{LR}$. By decoupling the feature space through this dynamic conversion scheme, the framework can allocate computational resources more strategically, thereby avoiding the need for high-complexity RGB backbones while effectively mitigating the non-uniform wavelength attenuation inherent in aquatic environments.

3.2. Highly Efficient Dual-Branch Upsampling

Recognizing that human visual perception is highly sensitive to structural details in the luminance channel but more tolerant to variations in the chrominance channels, we propose a specialized dual-branch architecture to allocate computational resources efficiently. As illustrated in Figure 3, the proposed framework explicitly decouples the low-resolution input into a luminance branch ($Y_{LR} \in \mathbb{R}^{1\times H\times W}$) and a chrominance branch ($CbCr_{LR} \in \mathbb{R}^{2\times H\times W}$), processing them through two parallel streams to ensure optimal structural and chromatic reconstruction.

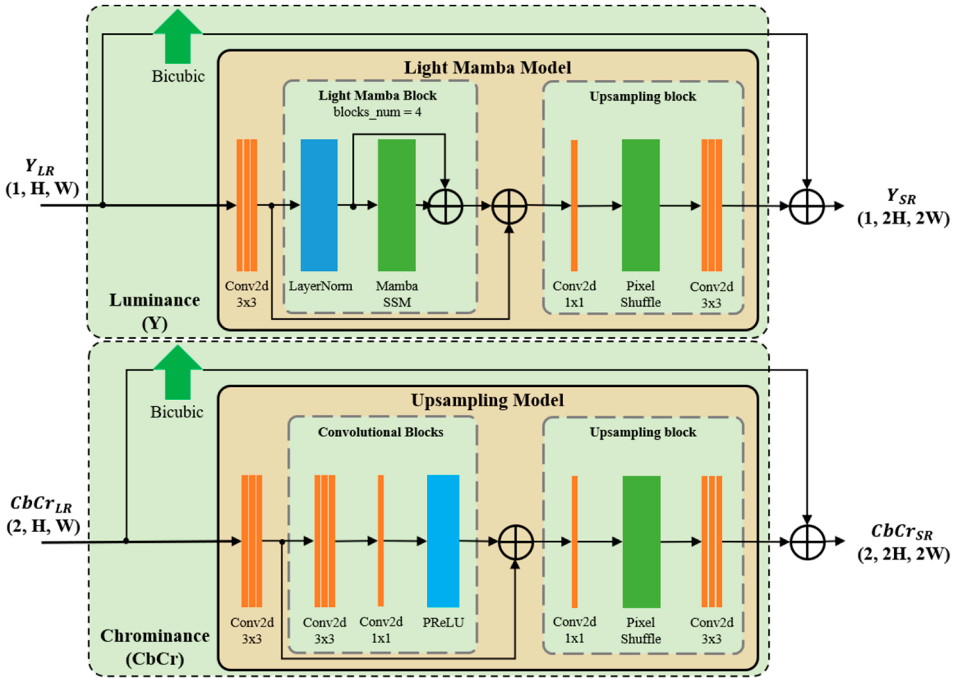


Figure 3. Highly Efficient Dual-Branch Upsampling.

For the luminance branch, the input Y_{LR} enters the light Mamba model and is first passed through a shallow feature-extraction layer consisting of a 3×3 convolution layer to generate the initial features. These features are then split into a block-level shortcut path and a main processing path. In the main path, the features are fed into the light Mamba block component configured with four cascaded blocks. The selection of four Mamba blocks [19] was determined through empirical optimization rather than a random choice. Extensive preliminary experiments demonstrated that configuring the module with four blocks yields the optimal trade-off between restoration performance and computational complexity, preserving the lightweight characteristics of the framework. Within this component, to maintain a clean and concise visual schematic, the core structure is represented as "Mamba SSM," whose underlying mathematical formulation is governed by the selective state–space model. Specifically, for a processed sequence x_t, the continuous-time evolution of the hidden state h_t within this module is defined as:

$$h'(t) = A \times h(t) + B \times x(t), \tag{4}$$

$$y(t) = C \times h(t), \tag{5}$$

where $h'(t)$ denotes the time derivative of the hidden state representing its rate of change, y_t signifies the continuous output sequence, and A, B, and C are the system matrices. To integrate this continuous-time formulation into our discrete deep learning network, these continuous parameters must be transformed into corresponding discrete forms $(\overline{A},\ \overline{B})$ via a discretization step with a step size parameter Δ, expressed as:

$$\overline{A} = \exp(\Delta A), \tag{6}$$

$$\overline{B} = (\Delta A)^{-1}(\overline{A} - I) \cdot \Delta B. \tag{7}$$

This discretization mechanism enables the "Mamba SSM" component to model extended spatial structural interactions within a linear computational bound, which is essential for minimizing the total FLOPs of our framework. Within each block, the features are normalized and processed by the selective SSM with a local shortcut. An internal residual connection helps combine the features. Then, a block-level residual connection adds the shallow features to the output of the light Mamba block. The aggregated feature maps are subsequently forwarded to the upsampling block, which sequentially comprises a 1×1 convolution, a pixel shuffle layer for $\times 2$ and $\times 4$ spatial scaling, and a final 3×3 convolution layer. To facilitate residual learning and accelerate training convergence, a global skip connection utilizing bicubic interpolation directly upsamples the original input Y_{LR} and aggregates it with the output of the upsampling block, yielding the final super-resolved luminance feature map $Y_{SR} \in \mathbb{R}^{1\times 2H\times 2W}$.

Concurrently, the chrominance branch restores the $CbCr_{LR}$ features via a lightweight upsampling model to maintain color fidelity with a minimal computational footprint. The input $CbCr_{LR}$ is initially transformed by a shallow 3×3 convolution layer. Similar to the luminance stream, the extracted features are split into a shortcut path and a main path that flows through the convolutional block component, which sequentially consists of a 3×3 convolution, a 1×1 convolution, and a PReLU activation function. A block-level residual connection then sums the initial features from the shallow layer with the output of the convolutional block component. The fused chrominance features are projected to the high-resolution space using an identical upsampling block structure containing a 1×1 convolution, a pixel shuffle layer, and a 3×3 convolution layer. Finally, a global bicubic skip connection up-samples the original $CbCr_{LR}$ input and adds it to the output of the upsampling block, producing the reconstructed chrominance map $CbCr_{SR} \in \mathbb{R}^{2\times 2H\times 2W}$.

3.3. Lightweight Residual Refinement

To achieve high-resolution restoration while maintaining a minimal computational footprint, the super-resolved luminance feature Y_{SR} and the chrominance feature $CbCr_{SR}$ obtained from the dual-branch network are concatenated along the channel dimension via a concat block to form a unified feature map $YCbCr_{SR} \in \mathbb{R}^{3\times 2H\times 2W}$. This concatenated feature map is then processed by a lightweight residual model to refine structural boundaries and suppress reconstruction artifacts. The prediction of the final refined high-resolution image can be mathematically expressed as:

$$YCbCr_{final} = G_{ref}(Concat(Y_{SR},\ CbCr_{SR})) + YCbCr_{SR}, \tag{8}$$

where G_{ref} denotes the refinement function mapping performed by the sequential layers prior to the global residual addition.

The internal architecture of this refinement model is sequentially organized into four main stages to progressively restore high-frequency details. In the first stage, a shallow

feature extraction layer, comprising a 3 × 3 convolution layer and a PReLU activation function, projects the 3-channel concatenated input into the high-dimensional feature space. In the second stage, these features flow through a deep body network containing four cascaded light residual blocks (LightResBlock). As illustrated in Figure 4, within each LightResBlock, the feature maps are sequentially processed by a 5 × 5 DSC, a PReLU activation layer, a second 5 × 5 DSC, and a Pixel Attention block (PABlock) [35]. A local residual connection aggregates the block's input with the output of the PABlock to stabilize the gradient flow and promote feature reuse. In the third stage, to effectively compensate for geometric distortions and non-uniform blurring caused by underwater light scattering, the features from the body network are forwarded to a cascaded deformable convolution stage consisting of two sequential deformable convolutional networks (DCN) [36] blocks. Within each DCN block, the input features are parallelly fed into parameter generators to dynamically estimate the coordinate offsets Δp and the modulation masks Δm based on the local data context. The learned offsets and masks are then passed into the core DCN layer to dynamically adjust the receptive field according to the structural topology of underwater objects, sharpening object boundaries and eliminating remaining turbid artifacts.

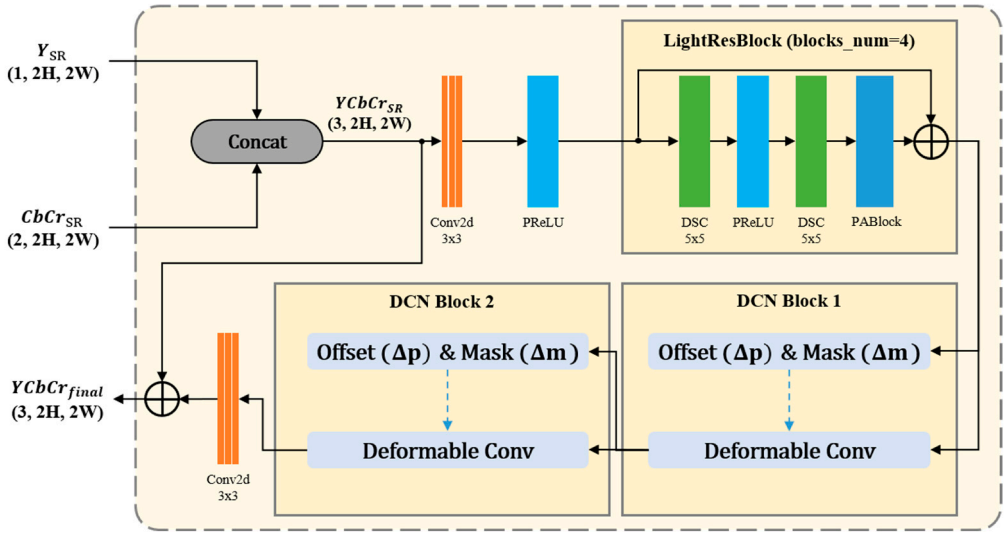


Figure 4. Feature Fusion and Lightweight Residual Refinement.

Finally, in the fourth stage, a tail reconstruction layer utilizes a 3 × 3 convolution layer to adjust the feature channels back to the output dimensions. To preserve the fundamental low-frequency content and prevent over-smoothing, a global residual connection performs an element-wise addition between the concatenated feature map $YCbCr_{SR}$ and the output of the tail layer, generating the final refined high-resolution representation $YCbCr_{final} = \left[Y_{final},\ Cb_{final},\ Cr_{final}\right]^T \in \mathbb{R}^{3\times 2H\times 2W}$.

Following this refinement stage, as illustrated in the final block of Figure 1, this unified representation is transformed back into the primary RGB color space to produce the ultimate super-resolved output image $RGB_{final} = \left[R_{final},\ G_{final},\ B_{final}\right]^T \in \mathbb{R}^{3\times 2H\times 2W}$ via an inverse learned conversion matrix. To achieve this, the normalized color weights shared from the dynamic color-space conversion module are utilized. Defining the chrominance differences as $Cb_{diff} = Cb_{final} - 0.5$ and $Cr_{diff} = Cr_{final} - 0.5$, the closed-form formulation inverse transformation for each reconstructed color channel is expressed as:

$$R_{final} = Y_{final} + 2(1 - \omega_r)\cdot Cr_{diff}, \tag{9}$$

$$B_{final} = Y_{final} + 2(1 - \omega_b)\cdot Cb_{diff}, \tag{10}$$

$$G_{final} = Y_{final} - \frac{2\omega_b(1-\omega_b)}{\omega_g+\epsilon}\cdot Cb_{diff} - \frac{2\omega_r(1-\omega_r)}{\omega_g+\epsilon}\cdot Cr_{diff}, \tag{11}$$

where $\epsilon = 10^{-8}$ denotes a minute offset integrated for numerical robustness against zero-division faults. This process guarantees both theoretical invertibility and numerical stability under all learning parameter circumstances. Since the dynamic weight vector ω is generated via a Softmax activation layer, the weights are strictly constrained to (0, 1). Because none of the RGB channel weights can collapse to exactly zero, the linear independence of the color channels is maintained, ensuring that the transformation matrix is non-singular and perfectly invertible. Furthermore, because the Softmax layer prevents ω_g from being negative, the denominator $(\omega_g + \epsilon)$ is strictly bounded away from zero, completely eliminating the risk of numerical overflow or division-by-zero faults. The proposed symmetric weight-sharing inverse formulation enforces a strict closed-loop constraint over the color learning framework, ensuring that the back-projection to the high-resolution RGB space accurately and deterministically compensates for the wavelength-dependent underwater light attenuation modeled at the beginning of the pipeline.

3.4. Loss Design

The proposed framework is optimized using a composite loss function that prioritizes both pixel-wise accuracy and structural perception. To explicitly guide the network toward high-fidelity restoration, the loss functions are computed directly between the ultimate super-resolved output image RGB_{final} and its corresponding ground-truth high-resolution image, denoted as $RGB_{GT} \in \mathbb{R}^{3\times 2H\times 2W}$. We employ the Charbonnier loss as a robust alternative to the standard L1 loss, which is defined as:

$$\mathcal{L}_{cb} = \frac{1}{N}\sum_{i=1}^{N}\sqrt{\left(RGB_{final}(i) - RGB_{GT}(i)\right)^2 + \epsilon^2}, \tag{12}$$

where i represents the pixel index across the joint spatial and spectral dimensions, N denotes the total number of pixels, and ϵ is set to 10^{-6} to keep the computation stable. To further preserve high-frequency structural details, the SSIM loss is incorporated as:

$$\mathcal{L}_{SSIM} = 1 - SSIM\left(RGB_{final}, RGB_{GT}\right) \tag{13}$$

The total objective function is expressed as:

$$\mathcal{L}_{total} = \lambda_1 \times \mathcal{L}_{cb} + \lambda_2 \times \mathcal{L}_{SSIM}, \tag{14}$$

where λ_1 and λ_2 are the weighting coefficients used to balance the contribution of pixel-wise reconstruction and structural consistency. In our implementation, these hyperparameters are empirically set to 2.0 and 5.0, respectively, to ensure optimal restoration performance.

4. Experimental Results

To train the proposed extremely lightweight dynamic color-space learning framework, the UFO-120 [7] dataset was selected for the underwater image super-resolution task, which provides paired low-resolution and high-resolution underwater images. The proposed framework was developed on an Ubuntu 24.04 LTS operating system using PyTorch version 2.9.0 with CUDA 12.8 in Python, and model optimization relied on the Adam optimizer [37] with a baseline learning rate of 10^{-4}, coupled with a Cosine Annealing learning rate scheduler. We used a training input patch size of 256 × 256 with a batch size of 8, and the model was trained for 2000 epochs to ensure network convergence. This

section outlines the experimental results through different subsections, focusing particularly on the trade-off between restoration quality and computational efficiency.

4.1. Quantitative Comparisons

The quantitative performance of the proposed method is validated using three recognized visual metrics: PSNR, SSIM, and UIQM [20]. Unlike traditional indices, UIQM is specifically tailored for the underwater environment by encompassing three critical visual attributes: color (UICM), sharpness (UISM), and contrast (UIConM). Generally, superior restoration performance is demonstrated by higher UIQM values. Furthermore, to explicitly evaluate the computational efficiency—which is the primary objective of our lightweight design—we incorporate the number of FLOPs and total trainable parameters (Params) as core comparative metrics. To evaluate the overall performance, we compare our model with several recent deep learning algorithms for underwater image enhancement and super-resolution: SRCNN [12], SRGAN [14], SRDRM [13], SRDRM-GAN [13], Deep SESR [7], AMPCNet [15], Deep WaveNet [16], DSMN [18], and MambaUSR [38]. Table 1 presents the quantitative performance and computational cost assessments on the UFO-120 [7] dataset for the scaling factor of $\times 2$ and $\times 4$.

Table 1. Quantitative evaluation results on the UFO-120 [7] dataset. Bold indicates the best performance, and underline denotes the second-best performance.

Method	PSNR (dB)		SSIM		UIQM		FLOPs (G)	Params (M)	Inference Time (ms)	FPS
	$\times 2$	$\times 4$	$\times 2$	$\times 4$	$\times 2$	$\times 4$				
SRCNN	24.75	19.05	0.72	0.56	2.39	2.02	**21.30**	**0.06**	28.44	35.16
SRGAN	26.11	21.08	0.75	0.58	2.44	2.56	377.76	5.95	504.42	1.98
SRDRM	24.62	23.15	0.72	0.67	2.59	2.57	203.91	0.83	272.28	3.67
SRDRM-GAN	24.61	23.26	0.72	0.67	2.59	2.55	289.38	11.31	386.41	2.59
Deep SESR	<u>27.15</u>	24.75	**0.84**	0.66	<u>3.13</u>	2.55	146.10	2.46	195.09	5.13
AMPCNet	25.24	24.70	0.71	0.70	2.93	2.88	-	1.15	-	-
Deep WaveNet	25.71	25.08	0.77	<u>0.74</u>	2.99	<u>2.97</u>	<u>21.47</u>	0.28	28.67	34.88
DSMN	26.02	<u>25.59</u>	<u>0.80</u>	0.72	3.00	2.89	47.98	0.55	64.07	15.61
MambaUSR	25.76	25.07	0.74	0.69	2.93	2.83	93.8	2.73	125.25	7.98
Proposed	**27.23**	**26.45**	<u>0.80</u>	**0.76**	**3.34**	**3.35**	26.96	<u>0.16</u>	36	27.78

To evaluate the framework on higher scaling factors, we conducted additional experiments for the $\times 2$ and $\times 4$ spatial scaling factor. The proposed method consistently outperforms contemporary models, including DSMN and recently proposed Mamba-based baseline MambaUSR, maintaining its superior performance in PSNR, SSIM, and UIQM under more challenging magnification.

While the raw prediction scores of baseline models are not publicly available for paired statistical testing, our framework consistently outperforms or matches State-of-the-Art baselines across multiple independent evaluation metrics and scaling factors. This systematic consistency strongly indicates robust architectural superiority rather than random statistical artifacts. Furthermore, an inference-time analysis was conducted to substantiate practical applicability. Tested on a ZOTAC Gaming Geforce NVIDIA RTX 5090 AMP Extreme Infinity (Made in Indonesia) GPU with a 320 $\times$ 240 input, our method requires only 36 ms for inference, achieving approximately 27.78 FPS, which is the fastest among the compared high-capacity models and verifies its viability for real-time processing.

As revealed by Table 1, the proposed framework delivers competitive numerical performance while maintaining a smaller computational footprint. Our proposed method reaches a PSNR of 27.23 dB and the highest UIQM score of 3.34 compared to other methods. This shows that it can better recover fine details and improve color quality in underwater images. At the same time, the model uses only 0.16M parameters and 26.96G FLOPs. Compared to other advanced architectures, the efficiency of our dynamic color-space and dual-branch design becomes evident. For instance, while Deep SESR achieves a slightly higher SSIM (0.84), it requires 146.10G FLOPs and 2.46M parameters, meaning our proposed framework consumes only about 18.4% of its computational operations and 6.5% of its parameters. Compared to GAN-based models such as SRGAN (377.76G FLOPs) and SRDRM-GAN (289.38G FLOPs), our method requires much less computation. While simpler models like SRCNN and Deep WaveNet have lower or similar FLOPs, their performance (e.g., PSNR of 24.75 dB and 25.71 dB) is noticeably lower. Overall, these results show that the proposed method can achieve a good trade-off between visual quality and computational efficiency, making it suitable for resource-constrained platforms.

4.2. Qualitative Comparisons

The qualitative performance for the $\times 2$ scaling factor is illustrated in Figure 5, where our proposed method is compared against several State-of-the-Art approaches. The visual results clearly demonstrate that our framework achieves superior detail preservation and structural sharpness while maintaining a natural color balance. From the enlarged regions in Figure 5, early models like SRCNN [12] and SRDRM [13] tend to produce blurry images and miss fine details. GAN-based methods such as SRGAN [14] and SRDRM-GAN [13] improve sharpness, but can introduce artifacts or distort structures. Newer models like Deep WaveNet [16] and DSMN [18] perform better, but still show some over-smoothing in detailed areas, such as fish scales or shipwreck surfaces.

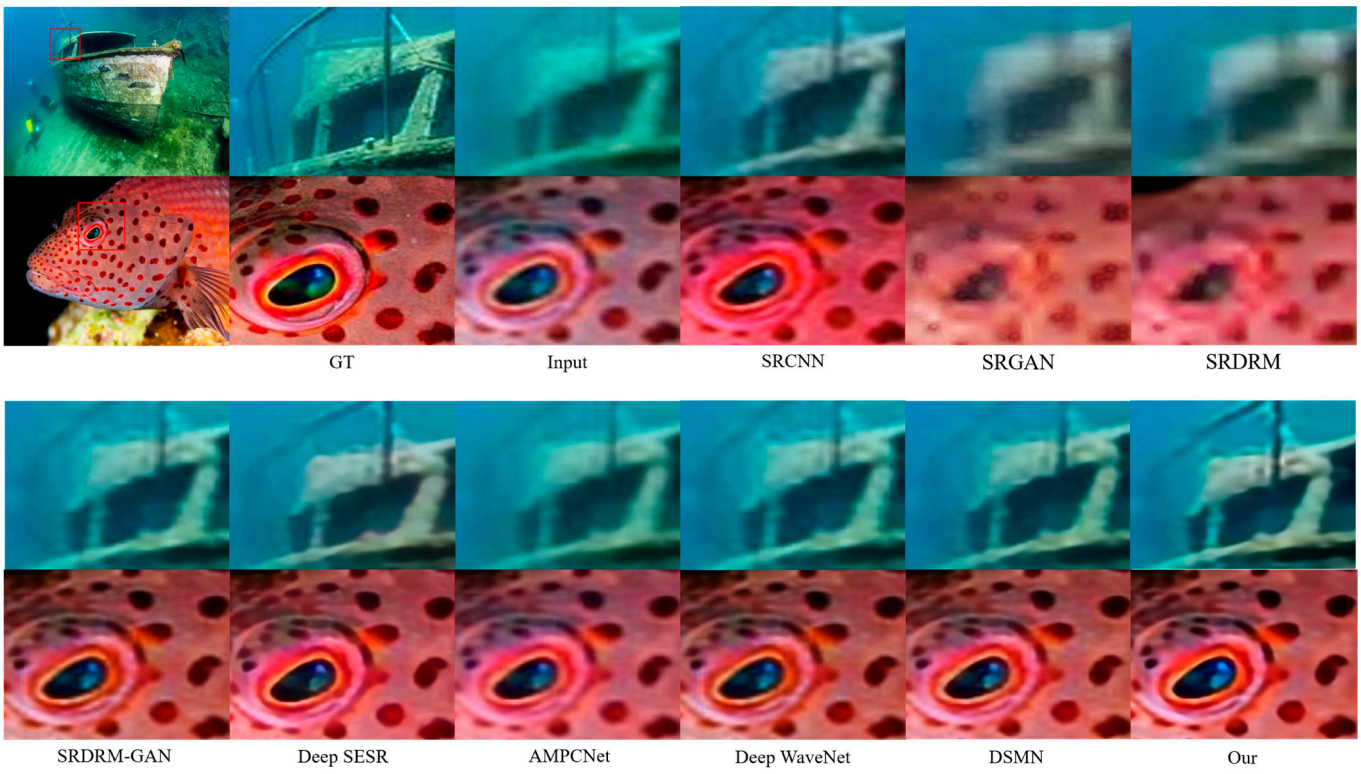


Figure 5. Qualitative evaluation results on the UFO-120 [7] dataset for $\times 2$ scaling.

In contrast, our proposed lightweight framework reconstructs intricate high-frequency details. This improvement comes from the dual-branch design. The Mamba blocks capture long-range information in the luminance channel, and the refinement stage sharpens edges using DCN and pixel attention. Also, our method learns luminance and chrominance information instead of treating RGB channels the same, which helps reduce color distortion. The results look more natural, making the method suitable for real-time use on resource-limited platforms.

To further demonstrate the robustness of our framework across diverse underwater degradation factors, we provide additional visual comparisons in Figure 6. These examples highlight specific challenges: Brain Coral (High-frequency details): This specimen contains dense pores and grooves that are heavily washed out by scattering in the blurred input. The model manages to recover these complex surface textures while pulling the shifted purplish-pink hue back toward the ground truth. Emperor Angelfish (High-contrast boundaries): Underwater blurring typically causes severe color bleeding across adjacent vivid regions, such as these blue and yellow stripes. By decoupling luminance and chrominance features, our dynamic color conversion module sharpens the dividing borders without introducing smearing artifacts. Starfish (Red wavelength attenuation and mosaic patterns): Because red light attenuates most rapidly underwater, the red patches in the input appear exceptionally dull. The network selectively compensates for this channel loss, retrieving the proper red tones alongside the grid-like skin pattern.



Figure 6. Additional visual comparisons of the proposed method on diverse underwater scenes.

4.3. Ablation Study

To validate the effectiveness of the core components in our lightweight framework, we conducted ablation studies focusing on the dynamic color-space conversion and the dual-branch upsampling architecture. The quantitative results at the $\times 2$ scaling factor are presented in Table 2.

Table 2. Quantitative results of ablation studies for UFO-120 [7].

	×2		
	PSNR	**SSIM**	**UIQM**
Complete	27.23	0.80	3.34
(w/o) Learnable Y	27.04	0.78	3.32
(w/o) Upsampling Model	26.04	0.78	3.39
(w/o) Mamba module	24.48	0.72	3.10
(w/o) Deformable Convolution	27.06	0.77	3.26
(w/o) Pixel Attention	27.24	0.78	3.42
(w/o) Residual Refinement	26.68	0.76	3.42

First, we evaluated the impact of the learnable color conversion module by replacing it with a standard, fixed RGB-to-YCbCr transformation matrix, denoted as (w/o) learnable Y. As shown in Table 2, the removal of this dynamic module leads to a drop in PSNR from 27.23 dB to 27.04 dB, and SSIM decreases from 0.80 to 0.78. This result suggests that a fixed color conversion does not work well under complex underwater conditions. By learning the weights dynamically using Softmax, the model can better extract luminance information and improve both detail and color quality.

Second, we investigated the necessity of the chrominance upsampling branch by completely removing the upsampling model for the CbCr channels, denoted as (w/o) upsampling model. In this configuration, the network only learns the super-resolution of the Y channel, while the CbCr channels are merely upsampled using basic bicubic interpolation. The results show a significant decrease in PSNR to 26.04 dB. The slight increase in the UIQM score to 3.39 is attributed to the unnatural contrast or over-saturation introduced by the unrefined, bicubic-upsampled chrominance combined with a highly sharpened luminance channel. This effect raises the metric score, but reduces PSNR and visual quality. This also explains why we use a dual-branch design, where a lightweight CNN processes the CbCr channels to keep the colors natural and the overall result stable.

Furthermore, to comprehensively justify the inclusion of key architectural components, we evaluated the removal of the Mamba module, deformable convolution blocks, pixel attention mechanism, and residual refinement network. The empirical results demonstrate that removing the Mamba module leads to the most significant performance drop (PSNR decreasing to 24.48 dB), confirming its critical role in capturing long-range dependencies. Deformable convolution and pixel attention are essential for handling spatially varying distortions and preserving fine-grained features, respectively.

To validate that the learnable color conversion module genuinely captures underwater spectral attenuation rather than acting as a simple projection layer, we analyzed the dynamically generated weight vector ω during the inference phase (Table 3). While the standard ITU-R BT.601 [34] conversion fixes the red channel weight at 0.299, our statistical analysis across the test set reveals that the network adaptively elevates the red channel weight to approximately 0.36–0.37. This consistent shift directly correlates with the physical reality of underwater optics, where red light undergoes the most severe attenuation. Visualizations of the intermediate learned luminance (Y_{LR}) and chrominance ($CbCr_{LR}$) feature maps further demonstrate that, compared to the standard fixed Y-channel, our dynamically extracted Y_{LR} representation preserves significantly sharper structural edges and fine textures under various underwater conditions.

Table 3. Dynamically learned RGB conversion weights.

Image Index	ω_r	ω_g	ω_b
1	0.3704	0.5401	0.0895
2	0.3700	0.5424	0.0877
3	0.3672	0.5590	0.0738
4	0.3711	0.5488	0.0801
5	0.3723	0.5311	0.0965
6	0.3661	0.5515	0.0824
7	0.3731	0.5432	0.0837
8	0.3652	0.5481	0.0867
9	0.3627	0.5472	0.0901
10	0.3685	0.5432	0.0883

4.4. Discussion

The experimental results and ablation studies show that the proposed lightweight framework works well for underwater image super-resolution. Table 1 indicates that high-quality restoration can be achieved without relying on large backbones or computationally expensive Transformer models. Although the PSNR improvement over models like Deep SESR (27.23 dB vs. 27.15 dB) is marginal, the real-world utility of our proposed method lies in achieving competitive baseline performance while offering extreme computational efficiency. By consuming approximately 18.4% of the computational operations and 6.5% of the parameters of Deep SESR, our model is significantly more viable for deployment on resource-constrained hardware like AUVs. This extreme efficiency is further confirmed by its actual GPU memory footprint; during inference on an NVIDIA RTX 5090, the maximum GPU memory allocated by our framework is merely 2684.13 MB. This is mainly due to the way the method handles non-uniform spectral attenuation. By learning channel weights adaptively, the model can compensate for the loss of certain wavelengths, such as red light, before feature extraction. This reduces the need for additional convolution operations to correct color distortion. Table 2 also suggests that allocating more modeling capacity to perceptually important components is effective. Applying the light Mamba blocks [19] mainly to the luminance channel helps preserve structural details and fine textures. At the same time, simplifying the processing of chrominance channels reduces computation while maintaining overall color quality. The ablation results in Table 2 also suggest that allocating more modeling capacity to perceptually important components is effective. Applying the light Mamba blocks [19] mainly to the luminance channel helps preserve structural details and fine textures. At the same time, simplifying the processing of chrominance channels reduces computation while maintaining overall color quality.

Despite these advantages, we acknowledge several limitations in the current study. First, the reliance on a single dataset (UFO-120 [7]) for quantitative evaluation limits the demonstration of broad generalization, although UFO-120 [7] encompasses a diverse array of subsea conditions and varying levels of water turbidity. Second, the framework may struggle in highly turbid environments with dense suspended particles, where heavy backscattering masks high-frequency details, potentially resulting in overly smoothed textures rather than sharp edges. Under extreme low-light deep-sea conditions lacking sufficient artificial illumination, the network might inadvertently amplify underlying sensor noise alongside structural features. Furthermore, when deployed in unseen oceanic regions with optical properties drastically different from the training distribution, the dynamically learned weights may shift sub-optimally, leading to slight color deviations.

To address these constraints, our future research will focus on collecting broader datasets, utilizing unpaired real-world data, and incorporating advanced no-reference underwater image quality assessment frameworks, such as PUIQA [39], for comprehensive perceptual restoration quality evaluation.

5. Conclusions

This paper has presented a lightweight dynamic color-space learning framework integrated with a dual-branch Mamba network for single underwater image super-resolution. By adaptively decoupling the input into optimal luminance and chrominance representations, the framework explicitly addresses non-uniform spectral attenuation while maintaining low computational demands. Experiments on the UFO-120 dataset [7] show that our method achieves a PSNR of 27.23 dB and a UIQM score of 3.34, with competitive performance compared to recent models. It uses only 0.16M parameters and 26.96G FLOPs, which are much lower than those of many existing approaches. The model maintains good image quality while keeping the computational cost low, making it suitable for real-time applications on limited-resource systems. In future work, we will improve generalization under different underwater conditions and explore multi-frame methods for low-power devices.

Author Contributions: Conceptualization, C.-H.Y.; Methodology, C.-H.Y., Y.-Y.L. and Y.-S.Y.; Software, Y.-Y.L.; Validation, Y.-Y.L. and Y.-S.Y.; Formal analysis, C.-H.Y. and Y.-S.Y.; Investigation, Y.-Y.L.; Data curation, Y.-Y.L.; Writing—original draft, Y.-Y.L.; Writing—review and editing, C.-H.Y. and W.-J.H.; Visualization, Y.-Y.L.; Supervision, C.-H.Y., W.-J.H. and C.-C.W.; Project administration, C.-H.Y. and C.-C.W.; Funding acquisition, W.-J.H. and C.-C.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Science and Technology Council (grant numbers: NSTC 113-2221-E-003-018-MY3, NSTC 114-2218-E-110-006, NSTC 114-2622-B-110-003, and NSTC 115-2640-E-110-004).

Data Availability Statement: The data presented in this study are available in [7].

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

References

1. Peng, L.; Zhu, C.; Bian, L. U-Shape transformer for underwater image enhancement. *IEEE Trans. Image Process.* **2023**, *32*, 3066–3079. [CrossRef]
2. Liu, X.; Chen, Z.; Xu, Z.; Zheng, Z.; Ma, F.; Wang, Y. Enhancement of underwater images through parallel fusion of transformer and CNN. *J. Mar. Sci. Eng.* **2024**, *12*, 1467. [CrossRef]
3. Yeh, C.-H.; Lai, Y.W.; Lin, Y.Y.; Chen, M.J.; Wang, C.C. Underwater image enhancement based on light field guided rendering network. *J. Mar. Sci. Eng.* **2024**, *12*, 1217. [CrossRef]
4. Chiang, J.Y.; Chen, Y.-C. Underwater image enhancement by wavelength compensation and dehazing. *IEEE Trans. Image Process.* **2012**, *21*, 1756–1769. [PubMed]
5. Li, C.; Guo, J.; Pang, Y.; Chen, S.; Wang, J. Single underwater image restoration by blue-green channels dehazing and red channel correction. In Proceedings of the IEEE International Conference on Acoustics, Shanghai, China, 20–25 March 2016; pp. 1731–1735.
6. Ancuti, C.O.; Ancuti, C.; Vleeschouwer, C.D.; Bekaert, P. Color balance and fusion for underwater image enhancement. *IEEE Trans. Image Process.* **2018**, *27*, 379–393. [PubMed]
7. Islam, M.J.; Luo, P.; Sattar, J. Simultaneous enhancement and super-resolution of underwater imagery for improved visual perception. *arXiv* **2020**, arXiv:2002.01155.
8. Pramanick, A.; Sur, A.; Saradhi, V.V. Harnessing multi-resolution and multi-scale attention for underwater image restoration. *Vis. Comput.* **2025**, *41*, 8235–8254.
9. 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.* **2018**, *3*, 387–394.
10. Zhu, J.-Y.; Park, T.; Isola, P.; Efros, A.A. Unpaired image-to-image translation using cycle-consistent adversarial networks. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017; pp. 2223–2232.

11. Lin, Y.-Y.; Huang, W.-J.; Yeh, C.-H. Dual-CycleGANs with dynamic guidance for robust underwater image restoration. *J. Mar. Sci. Eng.* **2025**, *13*, 231.
12. Dong, C.; Loy, C.C.; He, K.; Tang, X. Learning a deep convolutional network for image super-resolution. In Proceedings of the European Conference on Computer Vision, Zurich, Switzerland, 6–12 September 2014; pp. 184–199.
13. 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, Paris, France, 31 May 2020–31 August 2020; pp. 900–906.
14. Ledig, C.; Theis, L.; Huszar, 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.
15. 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.
16. Sharma, P.; Bisht, I.; Sur, A. Wavelength-based attributed deep neural network for underwater image restoration. *ACM Trans. Multimed. Comput. Commun. Appl.* **2023**, *19*, 1–23. [CrossRef]
17. 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*, 1–16. [CrossRef]
18. Wang, L.; Li, K.; Dong, C.; Shen, K.; Mu, Y. Dynamic structure-aware modulation network for underwater image super-resolution. *Biomimetics* **2024**, *9*, 774. [CrossRef] [PubMed]
19. Gu, A.; Dao, T. Mamba: Linear-Time Sequence Modeling with Selective State Spaces. In Proceedings of the First Conference on Language Modeling, Philadelphia, PA, USA, 7 October 2024.
20. Panetta, K.; Gao, C.; Agaian, S. Human-visual-system-inspired underwater image quality measures. *IEEE J. Ocean. Eng.* **2016**, *41*, 541–551.
21. 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.
22. Hui, Z.; Gao, X.; Yang, Y.; Wang, X. Lightweight image super-resolution with information multi-distillation network. In Proceedings of the 27th ACM International Conference on Multimedia, Nice, France, 21–25 October 2019; pp. 2024–2032.
23. Dong, C.; Loy, C.C.; Tang, X. Accelerating the super-resolution convolutional neural network. In *Proceedings of European Conference on Computer Vision*; Springer: Cham, Switzerland, 2016; pp. 391–407.
24. Gong, Y.; Wu, C.; Zheng, W.; Lu, S.; Xu, G.; Zhang, L.; Yin, L. DDNet: A novel dynamic lightweight super-resolution algorithm for arbitrary scales. *Comput. Model. Eng. Sci.* **2025**, *145*, 2223. [CrossRef]
25. Liu, X.; Li, H.; Sheng, V.; Ma, Y.; Liang, X.; Wang, G. Scale-aware attention and multi-modal prompt learning with fusion adapter for RGBT tracking. *IEEE Trans. Multimed.* **2025**, *28*, 344–359.
26. Zhang, H.; Zhong, Q.; Liu, J.; Chen, R.; Wang, Z. SAFARI-net: Scale-adaptive frequency-aware refinement infrastructure for infrared small target detection. *Opt. Laser Technol.* **2026**, *199*, 115049.
27. Yin, Y.; Deng, B.; He, Y.; Cheng, L.; Zhang, Y.; Wang, Y. DAMNet: RGB–liDAR fusion for traffic and water-surface detection using dynamic direction multiscale scanning priors and deformable window cross-modal. *IEEE Sens. J.* **2026**, *26*, 17062–17073.
28. Zhu, M.; Xu, Z.; Zhang, Q.; Liu, Y.; Gu, D.; Xu, S.S.D. GCSTormer: Gated swin transformer with channel weights for image denoising. *Expert Syst. Appl.* **2025**, *284*, 127924. [CrossRef]
29. Zhu, M.L.; Yuan, J.H.; Kong, E.; Zhao, L.L.; Xiao, L.; Gu, D.B. Generative adversarial networks with noise optimization and pyramid coordinate attention for robust image denoising. *Int. J. Intell. Syst.* **2025**, *2015*, 1546016.
30. Xue, X.; Hu, H.M.; He, Z.; Zheng, H. Towards multi-source illumination color constancy through physics-based rendering and spectral power distribution embedding. *IEEE Trans. Comput. Imaging* **2025**, *11*, 1349–1360.
31. Zhu, L.; Liao, B.; Zhang, Q.; Wang, X.; Liu, W.; Wang, X. Vision mamba: Efficient visual representation learning with bidirectional state space model. *arXiv* **2024**, arXiv:2401.09417.
32. Liang, Z.; Bao, S.; Su, H.; Zhang, W.; Wang, H.; Yan, H. Daily subsurface salinity reconstruction from multisource satellite observations using wavelet-enhanced 3-D Mamba. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2025**, *119*, 3371–3388. [CrossRef]
33. Yeh, C.-H.; Lin, C.-H.; Kang, L.-W.; Lin, M.-H.; Chang, C.-Y.; Wang, C.-C. Lightweight deep neural network for joint learning of underwater object detection and color conversion. *IEEE Trans. Neural Netw. Learn. Syst.* **2021**, *33*, 6129–6143. [CrossRef]
34. *ITU-R BT.601*; Studio Encoding Parameters of Digital Television for Standard 4:3 and Wide-Screen 16:9 Aspect Ratios. International Telecommunication Union: Geneva, Switzerland, 2011.
35. Zhao, H.; Kong, X.; He, J.; Qiao, Y.; Dong, C. Efficient image super-resolution using pixel attention. In *Proceedings of European Conference on Computer Vision*; Springer: Cham, Switzerland, 2020; pp. 56–72.
36. Dai, J.; Qi, H.; Xiong, Y.; Li, Y.; Zhang, G.; Hu, H.; Wei, Y. Deformable convolutional networks. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017; pp. 764–773.

37. Kinga, D.; Adam, J.B. A method for stochastic optimization. In Proceedings of the International Conference on Learning Representations, San Diego, CA, USA, 7–9 May 2015.
38. Shen, G.; Zhang, J.; Chen, Z. MambaUSR: Mamba and frequency interaction network for underwater image super-resolution. *Appl. Sci.* **2025**, *15*, 11263. [CrossRef]
39. Rehman, M.U.; Abbas, Z.; Nasir, M.F.; Hussain, I. A multiscale physics-informed framework for robust no-reference underwater image quality evaluation. *Alex. Eng. J.* **2026**, *135*, 114–125.

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.