symmetry

MDPI

Article

DAE-GAN: Underwater Image Super-Resolution Based on Symmetric Degradation Attention Enhanced Generative Adversarial Network

Miaowei Gao [1], Zhongguo Li [1,*], Qi Wang [1,2] and Wenbin Fan [3]

[1] School of Mechanical Engineering, Jiangsu University of Science and Technology, Zhenjiang 212100, China; gaomw745@stu.just.edu.cn (M.G.); wqi003@126.com (Q.W.)
[2] School of Automotive Engineering, Nantong Institute of Technology, Nantong 226001, China
[3] Jiangsu JBPV Intelligent Equipment Co., Ltd., Zhangjiagang 215634, China; fanwenbin@jbpv.com
* Correspondence: lzg1975@163.com

Abstract: Underwater images often exhibit detail blurring and color distortion due to light scattering, impurities, and other influences, obscuring essential textures and details. This presents a challenge for existing super-resolution techniques in identifying and extracting effective features, making high-quality reconstruction difficult. This research aims to innovate underwater image super-resolution technology to tackle this challenge. Initially, an underwater image degradation model was created by integrating random subsampling, Gaussian blur, mixed noise, and suspended particle simulation to generate a highly realistic synthetic dataset, thereby training the network to adapt to various degradation factors. Subsequently, to enhance the network's capability to extract key features, improvements were made based on the symmetrically structured blind super-resolution generative adversarial network (BSRGAN) model architecture. An attention mechanism based on energy functions was introduced within the generator to assess the importance of each pixel, and a weighted fusion strategy of adversarial loss, reconstruction loss, and perceptual loss was utilized to improve the quality of image reconstruction. Experimental results demonstrated that the proposed method achieved significant improvements in peak signal-to-noise ratio (PSNR) and underwater image quality measure (UIQM) by 0.85 dB and 0.19, respectively, significantly enhancing the visual perception quality and indicating its feasibility in super-resolution applications.

Keywords: underwater image super-resolution; degradation model; generative adversarial network; attention mechanism

Citation: Gao, M.; Li, Z.; Wang, Q.; Fan, W. DAE-GAN: Underwater Image Super-Resolution Based on Symmetric Degradation Attention Enhanced Generative Adversarial Network. *Symmetry* **2024**, *16*, 588. https://doi.org/10.3390/sym16050588

Academic Editor: Changxin Gao

Received: 2 April 2024
Revised: 30 April 2024
Accepted: 1 May 2024
Published: 9 May 2024

1. Introduction

Underwater imaging plays a pivotal role in marine science and engineering applications, offering significant value to oceanographic research, ecological monitoring, exploration of marine resources, and maintenance of underwater equipment [1]. It not only enhances the monitoring capabilities of marine life and coral reefs but also plays a central role in the precise localization, detection, and identification of underwater targets. However, the complexity of the underwater environment leads to loss of detail, reduced contrast, color distortion, blurred images, and increased noise in underwater images. Consequently, super-resolution technology for underwater images becomes critical. This technology compensates for various quality deficiencies in low-resolution images by reconstructing high-resolution images, thereby significantly enhancing image quality.

Single image super-resolution (SISR) is a well-established challenge within the realm of computer vision and image processing. The goal is to reconstruct a high-resolution image from a provided low-resolution input. The application of deep learning techniques has markedly improved the capabilities in this super-resolution (SR) task [2]; several methods based on the convolutional neural network (CNN) have been suggested [3–7] and have

almost dominated this field in the recent years. Subsequently, super-resolution methods based on generative adversarial networks (GANs) [8–11] have garnered attention. These methods enhance the quality of generated images through adversarial training, particularly in terms of restoring image details and textures. Techniques based on GANs have emerged as a significant branch within the field of image super-resolution, demonstrating their broad potential across multiple application domains. Recently, due to the advancements in natural language processing, Transformer [12] has captured the interest of the computer vision community. After making rapid progress on high-level vision application [13], methods based on the Transformer architecture are now being applied to low-level vision tasks, including super-resolution [14–17].

Underwater image super-resolution technology faces unique challenges. Firstly, the degradation process from high-resolution (HR) to low-resolution (LR) images underwater is often unknown, deviating from the specific methods typically used (such as bilinear interpolation, nearest-neighbor interpolation, etc.) to generate paired training data from HR to LR. Consequently, when models generate high-resolution images from low-resolution images, their performance often falls short of expectations. Secondly, in the complex underwater environment, the abundance of impurities, suspended particles, and the optical properties of water significantly affect image quality, leading to distortion and a considerable amount of noise in underwater images. These factors interfere with the basic structure of images and, to a certain extent, obscure important texture and detail information [18,19]. This makes it difficult for networks to extract and recognize useful features from underwater images, thereby impacting the quality and accuracy of super-resolution reconstruction.

Building upon an in-depth analysis and leveraging insights from the super-resolution domain, specifically ESRGAN [20], this study has developed and optimized a specialized underwater image super-resolution network tailored to address the unique challenges of underwater imaging. Firstly, a novel approach was designed to simulate the degradation process of underwater images, enabling the network to better learn the mapping relationship between HR and LR images, thereby enhancing the quality of underwater image reconstruction. Secondly, a series of innovative adjustments and optimizations were made to the model. A significant improvement involves the integration of an adaptive residual attention module within the dense residual blocks of the model, aimed at bolstering the network's ability to recognize and extract key features in underwater images [21]. Furthermore, a suite of targeted design optimizations was implemented, involving adjustments to the network's loss function and improvements to the configuration of convolutional layers, along with the introduction of spectral normalization (SN) layers to enhance the model's stability and generalization capacity. These comprehensive improvement strategies work in synergy to elevate the model's performance in processing underwater images.

The key contributions of this paper are outlined as follows:

- We propose a method to simulate the actual degradation process for underwater images, enabling the network to better learn the mapping between high-resolution and low-resolution images, thereby enhancing the quality of reconstructed images.
- The adaptive residual attention module designed for underwater images automatically assesses image importance using an energy function and, when integrated into dense residual blocks, enhances the precision of key feature extraction and the effectiveness of super-resolution reconstruction.
- Experimental results demonstrate that our approach achieves high PSNR values while maintaining low LPIPS scores, traditionally seen as opposing outcomes.

2. Related Work

2.1. Deep Networks for Image Super-Resolution

Since SRCNN first applied deep convolution neural networks to the image SR task and obtains superior performance over conventional SR methods, a variety of deep learning models [22–24] have been developed to further elevate image reconstruction quality. For instance, many methods apply more elaborate convolution module designs, such as residual

block [25–27] and dense block [28], to boost the representational power of the models. Some studies also investigate alternative architectures like recursive neural networks and graph neural networks [29]. To enhance perceptual quality, [30] employs adversarial learning to generate more realistic results. By integrating attention mechanism, [31–33] achieve further improvement in terms of reconstruction fidelity. Recently, a new wave of Transformer-based networks [34] have been introduced, consistently setting new benchmarks in the SR field and demonstrating the robust representational capabilities of the Transformer.

2.2. Degradation Models

In current super-resolution research, many networks still rely on simple interpolation methods [35] or traditional degradation models [36], which often struggle to accurately simulate the complex degradation phenomena present in the real world. Underwater images are typically affected by a variety of complex factors, including the scattering and absorption of light, the suspension of particulate matter, and dynamic blurring caused by water flow, all of which contribute to a decline in image quality. To effectively address this issue, we have designed a degradation process specifically tailored for the underwater environment. By simulating the unique degradation characteristics of underwater settings, our method ensures that the processed high-resolution images more closely resemble the properties of real underwater images.

2.3. Attention-Based Image Super-Resolution

Attention mechanisms enhance image reconstruction quality and model adaptability to the diversity and complexity of underwater images by highlighting critical features and extracting detailed information. Thus, they have become essential for improving underwater image super-resolution. For instance, RCAN [37] enhances network performance through channel attention mechanisms; SAN [38] leverages second-order channel attention to strengthen feature correlation learning; NLSN demonstrates the potential of attention mechanisms in addressing non-local dependencies; and SwinIR employs self-attention mechanisms from transformers. Moreover, CAL-GAN [39] effectively improves the super-resolution quality of photorealistic images by adopting a content-aware local generative adversarial network strategy, while DAT achieves efficient feature aggregation by merging features across spatial and channel dimensions.

3. Methods

3.1. A Practical Degradation Model

The SISR degradation model [40] can be mathematically formulated as

$$x = (y \otimes k) \downarrow_s \tag{1}$$

where y denotes the HR image; x denotes the LR image; k is the blur kernel; $\otimes$ denotes convolution operator; and $\downarrow_s$ denotes sub-sampling operator with stride of s.

Underwater imaging is subject to unique degradation factors distinct from those affecting conventional images, rendering traditional models inadequate for underwater image restoration. To address this, we have devised a degradation model specifically for underwater scenes, concentrating on unique aquatic factors to minimize computational overhead and enhance processing efficiency. The degradation dynamics of underwater images are captured by the following formula:

$$x = ((y \otimes k) + n + p) \downarrow_s \tag{2}$$

where n denotes the added noise, p denotes the suspended particles in the underwater environment.

As shown in Figure 1, the degradation model employs a first-order degradation process, and the detailed choices included in each degradation process are listed. In our previous experiments, we tested different sequences of degradation steps and found that their impact on the final results was negligible.

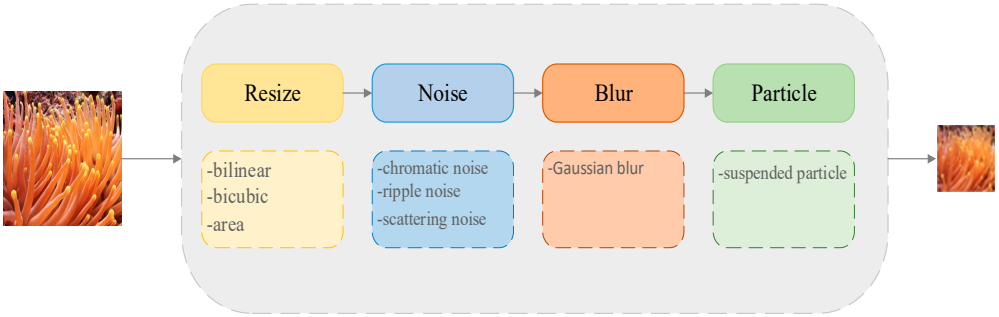

Figure 1. Overview of the degradation model, where each degradation process employs the classical degradation model.

3.1.1. Resize

Downsampling is a fundamental operation for generating low-resolution images in the realm of super-resolution [41]. Broadening our scope, we evaluate both downsampling and upsampling, that is, the resizing procedure. Various algorithms for resizing exist, including nearest-neighbor interpolation, area resizing, bilinear interpolation, and bicubic interpolation. Each method introduces its own distinctive effect, with some leading to blurriness, while others may produce overly sharp images accompanied by overshoot artifacts [42].

To encompass a richer array of complex resizing effects, we incorporate a stochastic selection of resizing techniques from the methods mentioned. Due to the misalignment complications presented by nearest-neighbor interpolation, we discount this method in favor of area, bilinear, and bicubic techniques.

3.1.2. Noise

The refractive index variation of air particles is mainly attributed to scattering, indicated as n_{scatter}, with chromatic dispersion, $n_{\text{chromatic}}$, being secondary and often overlooked. The wave-based index, n_{wave}, is essential for accurate long-distance light transmission. To accurately simulate these influences, a comprehensive stochastic noise model has been constructed:

$$n = \alpha n_{\text{scatter}} + \beta n_{\text{chromatic}} + \gamma n_{\text{wave}} \tag{3}$$

In the proposed model, weighting coefficients α, β and γ quantify the relative contributions of distinct noise sources, adhering to the normalization condition $\alpha + \beta + \gamma = 1$. This ensures precise modulation of each noise component within the comprehensive noise framework. Additionally, γ is restricted to $0 \leq \gamma \leq 0.1$, permitting nuanced adjustment of noise influence in a primarily linear domain, which is crucial for accurate noise behavior analysis. Through weighted fusion, we can better control the impact of noise on the system, improving system performance and stability.

3.1.3. Blur

To simulate the uniform blurring effects caused by less-than-ideal lighting conditions or environmental particulates, an isotropic Gaussian kernel is utilized. This kernel is founded on a two-dimensional normal distribution, with the standard deviation being equal in all directions, thus ensuring the uniformity of the blur effect. The isotropic Gaussian kernel k can be represented as follows:

$$k(i,j) = \frac{1}{N} \exp\left(-\frac{\frac{1}{2}(i^2 + j^2)}{\sigma^2} \right) \tag{4}$$

Within the matrix, (i,j) denotes the spatial coordinates, N is the normalization factor ensuring the sum of all weights equals 1, and σ represents the standard deviation [43]. During experimentation, kernels of various dimensions—3×3, 5×5, 7×7, and 9×9—

were implemented to replicate blurring effects across different area widths. The standard deviation was modulated from 1 to 3 to span blur intensities ranging from slight to severe.

3.1.4. Suspended Particles

In underwater imaging, image quality is notably impacted by suspended particulates that scatter incident light, producing distinct light spots. This research utilizes random field theory [44] to quantify scattering in this heterogeneous medium. The method allows simulation of the stochastic interactions between light and particles, generating statistically characterized scatter patterns. The function of the kernel based on spatial coordinates $I(x, y)$ is defined as follows:

$$I(x, y) = A \exp\left(-\frac{(x - x_0)^2 + (y - y_0)^2}{2\sigma^2}\right) \tag{5}$$

The coordinates (x_0, y_0) denote the centroid of the osculating circle located within the central segment of the elliptical distribution, where σ symbolizes the standard deviation thereof. The parameter A signifies the amplitude of said distribution, providing an index of its density, whereas the standard deviation σ conveys the extent of its dispersion.

3.1.5. Validation of the Degradation Model Efficacy

To accurately evaluate the effectiveness of the designed degradation model, this study quantified the degree of degradation by calculating the standard deviation (STD) of texture and noise in image samples. This approach considers the common noise and texture distortion characteristics of underwater images. By analyzing the standard deviations, it is possible to quantitatively assess how different models simulate underwater environments, thereby identifying the most suitable model for super-resolution reconstruction. The analysis results, as illustrated in Figure 2, revealed that the degraded image had a texture STD of 19.83 and a noise STD of 15.05, figures that align more closely with the characteristic high noise levels and lower texture clarity found in underwater imaging environments. In contrast, the texture STD of the image subjected to direct downsampling significantly increased to 78.95, while the noise STD decreased to 5.51. In summary, the texture and noise characterizations of the degraded image further affirm the suitability and superiority of our model for processing underwater images.

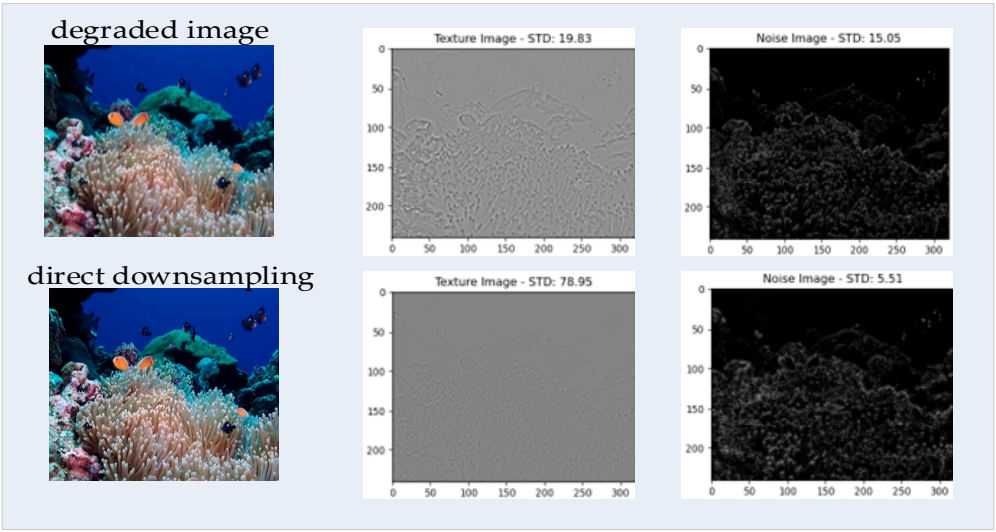

Figure 2. Comparative analysis of texture and noise in degraded versus directly downsampled underwater images.

To thoroughly validate the degradation model's capability to mimic authentic underwater image characteristics, we expanded our sample size and extracted a total of 100 images from the USR-248 dataset in increments of 5 for a comprehensive evaluation. The study involved a comparative analysis of images synthesized by the model against those produced by standard downsampling, focusing on the standard deviations of noise and texture. The analysis organized texture deviation in descending order and noise deviation in ascending order to establish trends. As shown in Figure 3, images processed by the degradation model demonstrate a higher standard deviation in noise and a lower one in texture, aligning more closely with the inherent properties of underwater images.

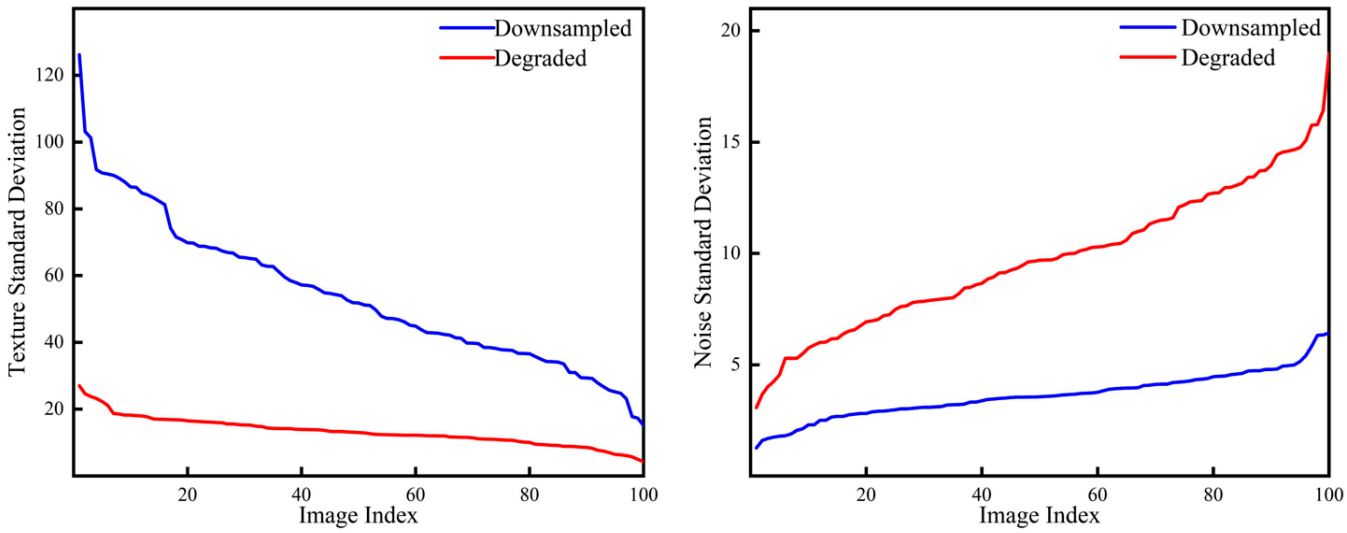

Figure 3. Comparative analysis of noise and texture standard deviations between the degradation model and direct downsampling, with the red line representing the degradation model and the blue line denoting direct downsampling.

3.2. Network Architecture

3.2.1. Overall Structure

As shown in Figure 4, our proposed symmetric degradation-aware and attention enhanced generative adversarial network (DAE-GAN) is structured into three main components: the degradation model, the generator, and the discriminator. The degradation model converts high-resolution images to low-resolution ones to mimic underwater conditions. The generator comprises a tripartite architecture, encompassing modules for shallow feature extraction, deep feature extraction, and image reconstruction. Specifically, for a given low-resolution input $I_{LR} \in \mathbb{R}^{H \times W \times C_{in}}$, we first exploit one convolution layer to extract the shallow feature $\mathrm{F}_0 \in \mathbb{R}^{\mathrm{H} \times \mathrm{W} \times \mathrm{C}}$, where C_{in} and C denote the channel number of the input and the intermediate feature. Then, a series of attention enhanced residual dense blocks (AERDB) and one 3×3 convolution layer are utilized to perform the deep feature extraction, with a total of 7 blocks ultimately employed in the final experiment. The final feature reconstruction layer ultimately generates high-resolution images. The discriminator is composed of feature extraction and activation layers followed by a classifier. It processes generated and real high-resolution images, uses convolutional layers and spectral normalization to stabilize training, and outputs a scalar authenticity score via a sigmoid function after a fully connected layer.

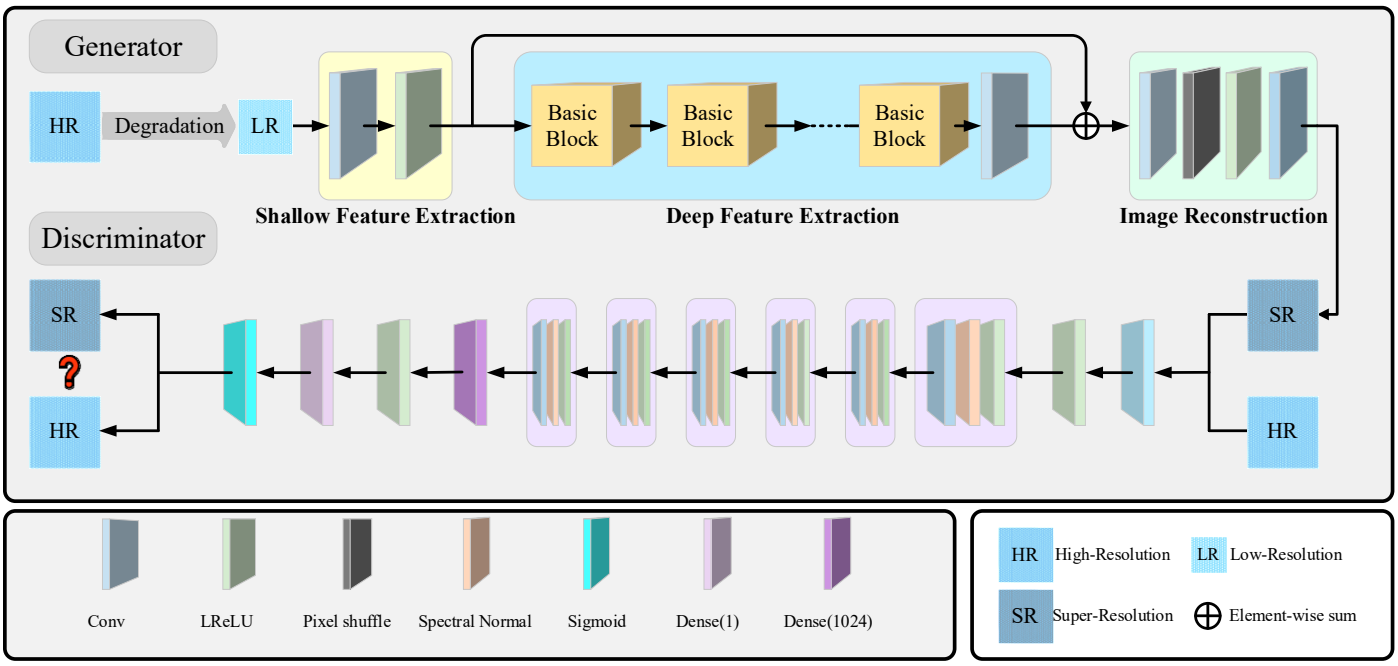

Figure 4. The overall architecture of DAE-GAN, with a symmetric structure in both the generator and discriminator components.

3.2.2. Attention Enhanced Residual Dense Block

As shown in Figure 5, each AERDB fuses the residual in residual dense block (RRDB) with the adaptive residual attention module (ARAM), the latter focusing on key features through its unique energy function, thereby enhancing the overall performance of the module. It is important to highlight that this structure exhibits symmetry, with both the main architecture and the branching structure being symmetric. This symmetry ensures balanced processing and optimization across the entire network, contributing to its effectiveness in image super-resolution.

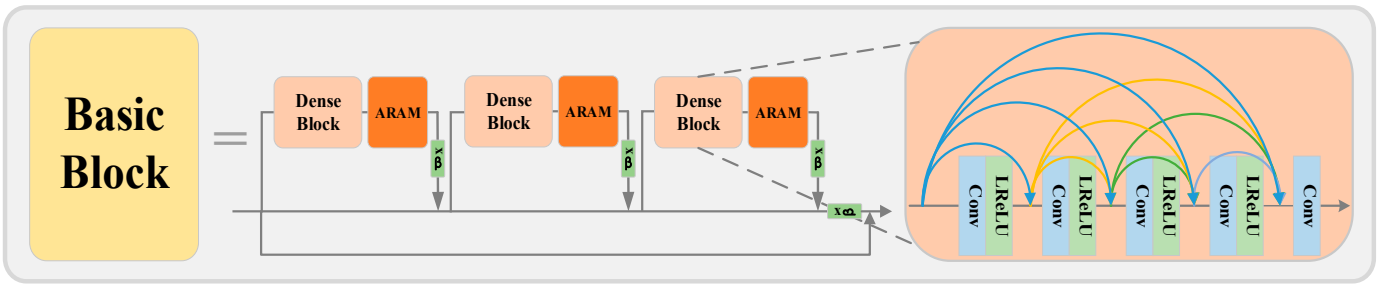

Figure 5. Schematic representation of the basic block, accentuating the positional relationship between the dense blocks and ARAMs.

Inspired by SimAM [45], the ARAM implements a unique energy function informed by neuroscience to modulate neuronal activity within feature maps, thereby boosting the model's ability to capture details. This function accounts for the spatial arrangement and activity patterns of neurons, enabling the network to process visual information with greater precision. The energy function is

$$e_t(w_t, b_t, y, x_i) = (y_t - \hat{t})^2 + \frac{1}{M-1}\sum_{i=1}^{M-1}(y_o - \hat{x}_i)^2 \quad (6)$$

In our approach, $\hat{t} = w_t t + b_t$ and $\hat{x}_i = w_t x_i + b_t$ represent linear transformations of t and x_i, where t is the target neuron and x_i signifies other neurons within a single channel of the input feature map X that belongs to $\mathbb{R}^{H \times W \times C}$. i is index over spatial dimension and $M = H \times W$ is the number of neurons on the channel. w_t and b_t are weight and bias the transform. All variables are scalars, achieving minimum values when t aligns with y_t and all x_i align with y_o, where y_t and y_o represent distinct scalar values. Minimizing this function is tantamount to ascertaining the linear separability of the target neuron t from its peers in the channel. To streamline this process, binary labels (1 and -1) are assigned to y_t and y_o Incorporating a regularizer into the equation, the final energy function is formulated as

$$e_t(w_t, b_t, y, x_i) = \frac{1}{M-1} \sum_{i=1}^{M-1} (-1 - (w_t \cdot x_i + b_t))^2 + (1 - (w_t \cdot t + b_t))^2 + \lambda w_t^2 \quad (7)$$

The energy function for each channel, in theory, can be computationally intensive to solve with iterative algorithms such as SGD. Fortunately, a rapid resolution method allows for the efficient calculation of the overall solution:

$$w_t = -\frac{(t - \mu_t)^2 + 2(t \cdot \mu_t)}{2\sigma_t^2 + 2\lambda} \quad (8)$$

$$b_t = -\frac{1}{2}(t^2 + t \cdot \mu_t) w_t \quad (9)$$

In the model, μ_t and σ_t^2, representing the mean and variance across all neurons, can be challenging to compute channel-wise. Therefore, we utilize global mean and variance as proxies. Under this assumption, these statistics are computed over all neurons and applied to adjust the aforementioned neurons, thereby alleviating the model's computational load. Consequently, the minimized energy function can be succinctly expressed by the following formula:

$$e_t^* = \frac{4(\sigma^2 + \lambda)}{(t - \mu)^2 + 2\sigma^2 + 2\lambda} \quad (10)$$

The improved e_t^* can help in differentiating neurons with significant characteristic differences, which is beneficial for feature extraction. Therefore, the importance of each neuron can be obtained by $1/e_t^*$.

In this approach, the model not only learns the intensity of each pixel but also the inter-pixel relationships. As shown in Figure 5, to incorporate SimAM into the RRDB module, a SimAM unit is placed right after the output of each dense block. Specifically, the feature map output from RRDB's dense block is fed into SimAM, which then calculates attention weights for every neuron within it, with the weighted output serving as the input for the following layer. Through this process, SimAM adaptively emphasizes features with higher variability or greater importance to the reconstruction task by minimizing its energy function, while suppressing less contributory information for the current task. This adaptive adjustment strategy not only improves RRDB's ability to discriminate features but also enhances the network's generalization capabilities in complex scenarios.

Grad-CAM [46] is applied for visualizing feature activation to evaluate the performance differences between the RRDB module used alone and the RRDB module integrated with an attention mechanism. As shown in Figure 6, the RRDB module with integrated attention mechanism enhances the network's focus on key areas during image processing. This focused attention leads to more pronounced activation of critical features, rather than a uniform distribution of attention across the entire image. Such improvements enhance the model's ability to recognize important information in images, significantly benefiting the accuracy and efficiency of deep learning models in image processing tasks.

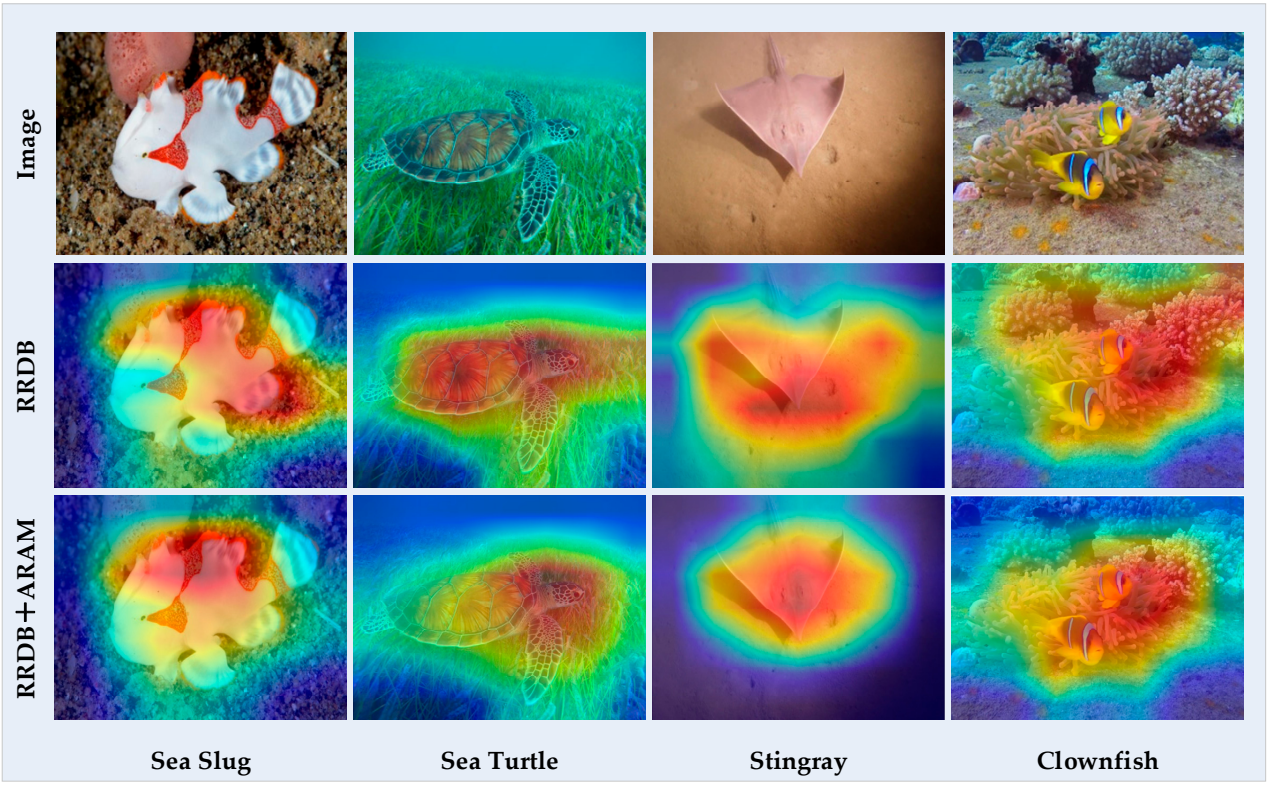

Figure 6. Feature activation comparison using Grad-CAM, illustrating the SimAM's impact on emphasizing key areas in underwater images for various marine species.

On the heatmap, it may be difficult to clearly observe significant differences between different models. Therefore, to assess the model's performance more accurately, we introduced attention scores as an additional metric. Attention scores quantify the degree of focus of deep neural networks on different regions of the image, serving as a measure of model performance. Higher scores indicate a greater focus on key areas, improving the recognition and understanding of critical information within the image. Thus, by utilizing attention scores, we can comprehensively evaluate the model's performance. The specific scores are shown in Table 1.

Table 1. The attention scores when using the RRDB module alone and when integrating the RRDB module with an attention mechanism, where higher scores indicate a greater focus on key areas by the model.

Image Sample	RRDB Attention Score	RRDB + ARAM Attention Score
Sea Slug	0.72	0.78
Sea Turtle	0.53	0.75
Stingray	0.58	0.68
Clownfish	0.49	0.56

3.3. Networks and Training

3.3.1. Generator

This study has innovatively enhanced the generator architecture of ESRGAN by integrating an attention mechanism within its RRDB modules. Moreover, the architecture, initially designed for 4× upscaling, has been expanded to support super-resolution at higher scale factors. These improvements not only bolster the network's capability for detail processing but also enhance its versatility across different magnification rates.

3.3.2. Discriminator with Spectral Normalization

To ensure overall model stability, this study has incorporated spectral normalization into the discriminator architecture to enhance its resistance to interference. The integration of a spectral normalization layer within the discriminator helps to maintain the spectral norm of the weights during training, preventing the weights from growing indefinitely. This approach effectively ensures the stability of the model and significantly improves its ability to resist interference.

3.3.3. Loss Function

The loss function used by the network in this study is defined as

$$L_{\mathrm{GAN}} = \lambda_1 L_{\mathrm{recons}} + \lambda_2 L_{\mathrm{percep}} + \lambda_3 L_{\mathrm{adv}} \tag{11}$$

where L_{recons} is the pixel-wise reconstruction loss, L_{percep} is the perceptual loss measuring the feature distance in VGG feature space, and L_{adv} denotes the adversarial loss. The coefficients λ_1, λ_2, and λ_3 are the balancing parameters, with values typically set to 0.1, 1, and 0.005, respectively.

4. Experiments

4.1. Datasets and Experiments Settings

4.1.1. Datasets

To train our model, two open-source datasets were selected, namely USR-248 [47] and UFO120 [48]. The USR-248 dataset is the first dataset designed for the super-resolution reconstruction of underwater images, containing 1060 pairs of underwater images for training and 248 pairs for testing. The UFO120 dataset consists of 1500 training samples and 120 testing samples. The low-resolution images in both datasets are created through artificial simulation and deformation. All samples were processed according to standard procedures for optical and spatial image degradation and combined with manually labeled saliency mappings to generate data pairs. Additionally, to validate the effectiveness of the model, two more datasets, EUVP and SQUID, were used as test datasets. We trained the model separately on the USR-248 dataset and the UFO120 dataset, and then used the model trained on the USR-248 dataset to test the EUVP dataset and the SQUID dataset. The specific information of the datasets used in our experiments is shown in Table 2.

Table 2. Detailed information on training and validation datasets.

Name	Type	Train Samps	Val/Test Samps	HR Image Size
USR-248	Train	1060	248	640 × 480
UFO120	Train	1500	120	640 × 480
EUVP	Val	-	12,000	256 × 256
SQUID	Val	-	57	256 × 256

4.1.2. Implementation Details

The model in this study was developed within the PyTorch framework and trained using the Adam optimizer, with the hyperparameters β1 and β2 for the optimizer set at 0.9 and 0.99, respectively. The initial learning rate was set to 2×10^{-4} and was halved after 200k iterations, with the entire training process spanning 400k iterations. Training on the dataset utilized input image patches of 64 × 64 pixels, with a batch size set to 32. To improve training effectiveness, symmetric image flipping was used as an augmentation technique. This exposed the model to a wider range of image variations, enhancing generalization performance and adaptability to various underwater conditions. We used NVIDIA RTX 3090 GPUs with CUDA acceleration for all training processes in the experiments. The computer had 32GB of RAM, and the CPU was an i7 13700k.

4.1.3. Evaluation Metrics

This study employs PSNR, structural similarity index measure (SSIM), UIQM, and learned perceptual image patch similarity (LPIPS) to evaluate underwater image super-resolution. PSNR and SSIM gauge signal-to-noise ratio and visual similarity of reconstructed images. UIQM addresses quality degradation from underwater scattering and absorption. LPIPS, using deep learning, assesses perceptual image quality, aligning evaluation with human visual observation. Specifically, PSNR and SSIM are computed on the Y channel in the YCbCr space.

4.2. Comparisons of Super-Resolution Results

4.2.1. Quantitative Results

Table 3 shows the quantitative comparison of our method against other methods on the USR-248 dataset—Deep WaveNet [49], RDLN [50], etc.—while Table 4 presents the quantitative comparison of our method against other methods on the UFO120 dataset: SRDRM [51], AMPCNet [52], HNCT [53], URSCT [54], etc. These outcomes are derived from the average performance metrics across all test samples. Notably, DAE-GAN achieved significant improvements in both SSIM and PSNR metrics and also performed admirably with respect to UIQM and LPIPS metrics. It is imperative to underscore that a lower LPIPS score indicates superior image quality.

Table 3. Experimental evaluation on the USR-248 dataset, offering a quantitative comparison at magnification factors ×2, ×4, and ×8 with other methods, utilizing four metrics: PSNR (dB)↑, SSIM↑, UIQM↑, and LPIPS↓. The best results are highlighted in red and the second-best in blue.

Method	PSNR(dB)↑			SSIM↑			UIQM↑			LPIPS↓		
	×2	×4	×8	×2	×4	×8	×2	×4	×8	×2	×4	×8
SRCNN	26.81	23.68	19.07	0.76	0.65	0.57	2.59	2.38	2.01	0.56	0.71	0.86
VDSR	27.98	24.70	20.15	0.79	0.69	0.61	2.61	2.44	2.09	0.53	0.67	0.83
SRGAN	26.68	23.46	19.83	0.73	0.63	0.54	2.55	2.38	1.98	0.30	0.48	0.61
ESRGAN	28.08	24.50	20.08	0.76	0.67	0.57	2.59	2.45	2.02	0.24	0.40	0.54
BSRGAN	28.15	25.05	20.33	0.79	0.69	0.59	2.63	2.47	2.07	0.20	0.32	0.42
Real-ESRGAN	28.86	25.11	20.45	0.80	0.71	0.62	2.68	2.50	2.10	0.19	0.33	0.44
Deep WaveNet	29.09	25.40	21.70	0.83	0.73	0.63	2.72	2.53	2.13	0.44	0.61	0.72
RDLN	29.76	25.59	22.40	0.82	0.71	0.62	2.74	2.58	2.19	0.29	0.50	0.66
DAIN	29.97	26.16	22.86	0.84	0.73	0.63	2.77	2.64	2.17	0.43	0.63	0.69
DAE-GAN(ours)	29.95	26.23	22.83	0.85	0.75	0.64	2.80	2.68	2.20	0.19	0.31	0.40

Table 4. Experimental evaluation on the UFO120 dataset, offering a quantitative comparison at magnification factors ×2, ×3, and ×4 with other methods, utilizing four metrics: PSNR (dB)↑, SSIM↑, UIQM↑, and LPIPS↓. The best results are highlighted in red and the second-best in blue.

Method	PSNR(dB)↑			SSIM↑			UIQM↑			LPIPS↓		
	×2	×3	×4	×2	×3	×4	×2	×3	×4	×2	×3	×4
SRCNN	24.75	22.22	19.05	0.72	0.65	0.56	2.39	2.24	2.12	0.56	0.65	0.71
SRGAN	25.11	23.01	19.93	0.75	0.70	0.58	2.44	2.39	2.35	0.24	0.33	0.37
Deep WaveNet	25.71	25.23	23.26	0.77	0.76	0.73	2.89	2.86	2.85	0.40	-	0.53
AMPCNet	25.24	25.43	25.08	0.71	0.70	0.70	2.76	2.65	2.68	0.31	-	0.47
ESRGCNN	25.82	25.98	24.70	0.73	0.71	0.71	2.88	2.86	2.75	0.34	0.46	0.51
HNCT	25.73	25.86	24.91	0.72	0.73	0.70	2.76	2.78	2.64	0.27	0.40	0.47
URSCT	25.96	-	25.37	0.81	-	0.69	-	-	-	0.37	0.49	0.50
RDLN	26.20	26.13	25.56	0.78	0.74	0.73	2.87	2.84	2.83	0.29	0.37	0.39
DAE-GAN(ours)	26.26	26.19	25.89	0.80	0.76	0.74	2.88	2.87	2.85	0.19	0.25	0.30

Upon evaluating the DAE-GAN approach against other state-of-the-art methods, it emerges as the clear leader in the USR-248 dataset, particularly excelling at a 2× scale with a PSNR of 29.95dB and an SSIM of 0.85. Its prowess extends to superior image quality

metrics, outshining competitors with the lowest LPIPS score, indicative of higher image fidelity at a 4× scale. As magnification increases to 8×, DAE-GAN consistently upholds its exceptional performance, achieving a PSNR of 23.83dB and an SSIM of 0.64, reinforcing its robustness in enhancing image resolution and quality across scales. Further analysis on the UFO120 dataset corroborates DAE-GAN's superior capabilities. It leads with a notable margin, particularly at 2× magnification, where it achieves a PSNR of 26.26dB and a remarkable SSIM of 0.80, surpassing other methodologies. Even at higher magnifications of 3× and 4×, DAE-GAN maintains its supremacy, reflected through its consistent scores, notably, a top-tier LPIPS of 0.25 at 3× and 0.30 at 4× magnification.

4.2.2. Qualitative Results

To conduct a comprehensive evaluation of DAE-GAN's performance, this study visually compares the effectiveness of various methods. Figures 7 and 8 illustrate the reconstruction results of our method at a 4× scale using a single network for arbitrary-scale SR, clearly showing from the comparisons that DAE-GAN excels in restoring image clarity and texture details, particularly at the edges. These visual results underscore the significant advantages of DAE-GAN in enhancing image quality, demonstrating its efficacy in the task of refined image restoration. This paper employs symmetrical images for visualizing results, clearly demonstrating the DAE-GAN model's precision in symmetry reconstruction. These visualizations are both appealing and assist in the detailed evaluation of the reconstruction process.

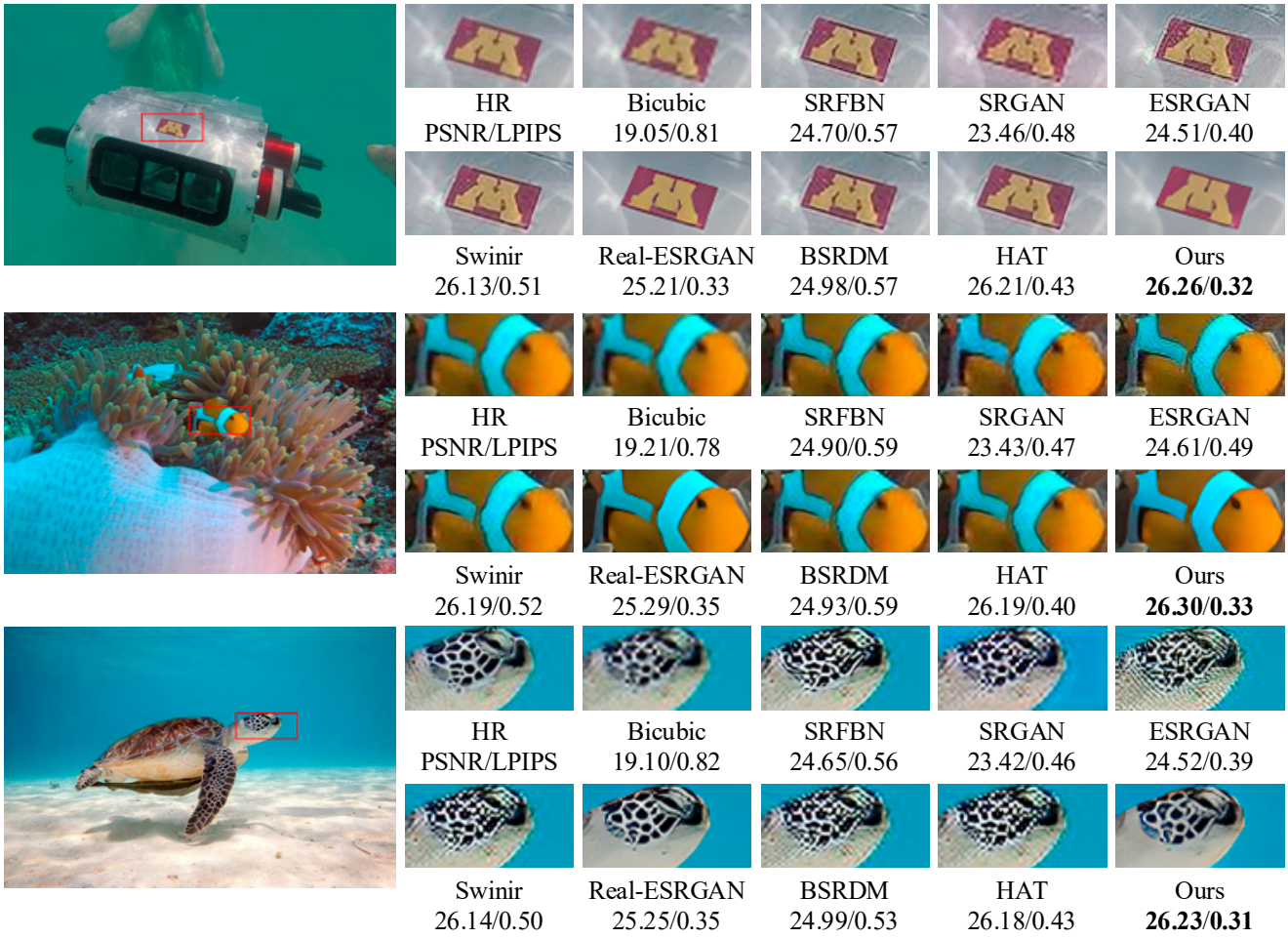

Figure 7. Qualitative comparison of different methods on ×4 super-resolution for the USR-248 dataset. Regions for comparison are highlighted with red boxes in the original high-resolution images. Zoom in for the best view. Higher PSNR, better quality; lower LPIPS, closer to original.

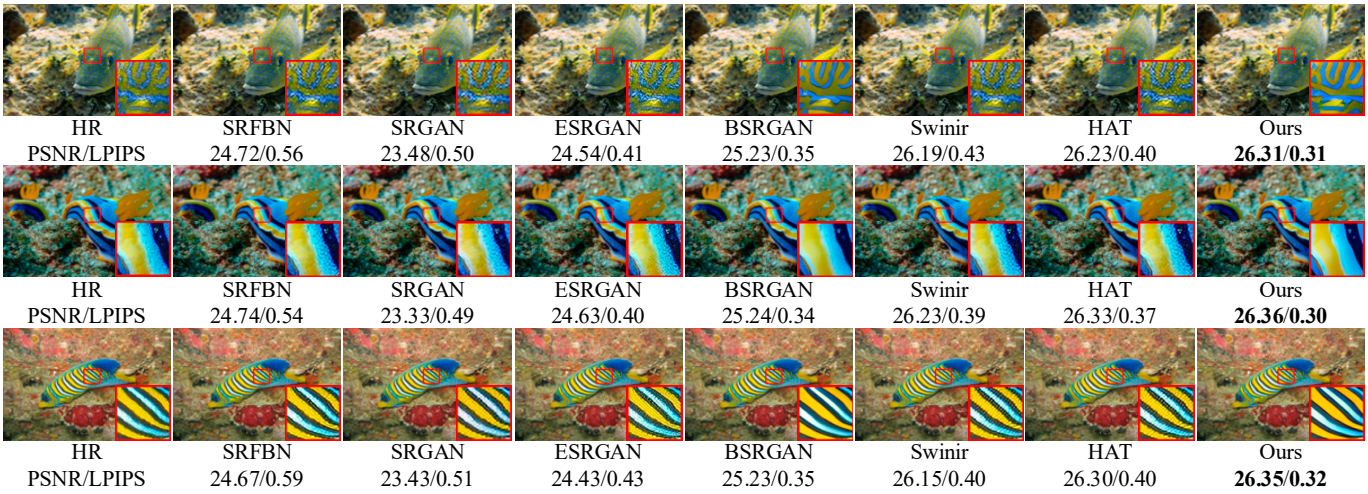

Figure 8. Qualitative comparison of different methods on ×4 super-resolution for the UFO120 dataset. Regions for comparison are highlighted with red boxes in the original high-resolution images. Zoom in for the best view. Higher PSNR, better quality; lower LPIPS, closer to original.

4.3. Model Performance Evaluation on Test Datasets

To rigorously ascertain the efficacy and robustness of the DAE-GAN model propounded in this research, an extensive battery of tests was executed on the EUVP and SQUID datasets. As shown in Table 5, the DAE-GAN model proposed in this study demonstrates solid overall performance in the quantitative evaluation of 4× super-resolution, underscoring its effectiveness in tackling the challenges of underwater image super-resolution from various dimensions. Upon closer examination, the model displays the best performance across all assessment metrics on the SQUID dataset. On the EUVP dataset, it scores slightly lower in the LPIPS index compared to CAL-GAN and is slightly outperformed in the PSNR index by BSRDM [55]; however, it secures the top results in all other relevant evaluation metrics.

Table 5. Experimental evaluation on the EUVP and SQUID test datasets, offering a quantitative comparison exclusively at a magnification factor of ×4 against other methods, utilizing four metrics: PSNR (dB)↑, SSIM↑, UIQM↑, and LPIPS↓. The best results are highlighted in red and the second-best in blue.

Method	Scale	EUVP				SQUID			
		PSNR↑	SSIM↑	UIQM↑	LPIPS↓	PSNR↑	SSIM↑	UIQM↑	LPIPS↓
SRCNN	×4	23.64	0.63	2.37	0.70	23.66	0.65	2.38	0.70
SRGAN		23.31	0.59	2.38	0.48	23.37	0.63	2.40	0.45
ESRGAN		24.40	0.66	2.44	0.38	23.62	0.68	2.48	0.38
BSRGAN		24.89	0.70	2.47	0.32	25.11	0.72	2.51	0.31
Real-ESRGAN		25.01	0.73	2.48	0.33	25.23	0.75	2.53	0.33
PDM-SRGAN		25.89	0.74	-	0.29	26.04	0.74	-	0.28
BSRDM		26.35	0.76	2.40	0.38	26.40	0.73	2.43	0.36
CAL-GAN		26.09	0.71	-	0.31	26.11	0.69	-	0.29
DAE-GAN (ours)		26.33	0.78	2.63	0.30	26.41	0.77	2.68	0.29

4.4. Ablation Study

In this section, we examine the importance of every fundamental element within our suggested approach. Through a series of exhaustive ablation experiments on the USR-248 dataset, this study comprehensively evaluates and confirms the performance and effectiveness of the proposed degradation model (DM) and ARAM when applied independently and in conjunction, as shown in Table 6.

Table 6. Ablation study on the impact of degradation models and ARAM for ×4 super-resolution in the USR-248 dataset. The best results are highlighted in red and the second-best in blue.

DM	ARAM	PSNR↑	SSIM↑	UIQM↑	LPIPS↓
		24.50	0.66	2.45	0.40
√		25.02	0.68	2.64	0.34
	√	25.89	0.72	2.47	0.37
√	√	26.23	0.75	2.68	0.31

We have employed an innovative training strategy by horizontally flipping every other image and rotating them 180 degrees to augment the dataset, increasing it by 1.5 times. This approach helps the model better grasp symmetry features, enhancing image super-resolution performance. Introducing symmetry enables more accurate capture and reconstruction of symmetric structures, particularly in imagery rich in symmetry like underwater images. This strategy enriches dataset diversity and strengthens the model's ability to recognize and reconstruct various types of symmetric structures underwater, ultimately improving generalization and robustness. The other experimental conditions, consistent with previous experiments, are reflected in the experimental results shown in Table 7.

Table 7. The ablation study on the impact of image flipping for ×4 super-resolution in the USR-248 dataset aims to investigate the effect of image flipping on the symmetry features in the super-resolution task.

Image Flipping Applied	PSNR↑	SSIM↑	UIQM↑	LPIPS↓
Not Applied	26.23	0.75	2.68	0.31
Applied	26.33	0.77	2.69	0.30

5. Conclusions

In this work, an innovative generative adversarial network architecture, termed DAE-GAN, is introduced with the aim of enhancing the super-resolution processing of underwater images. To more accurately reflect the inherently complex and irregular degradation phenomena present in underwater environments, a specialized degradation model was specifically designed and placed at the forefront of the super-resolution network. This model not only simulates the unique degradation process of underwater images but also provides more realistic input conditions for subsequent super-resolution reconstruction. To effectively capture the delicate features in underwater images, an adaptive residual attention module and dense residual blocks were integrated, boosting the network's sensitivity to details and its feature extraction capability. Extensive experiments conducted on multiple datasets, and evaluations at different magnification scales, have demonstrated not only a significant improvement in visual effects but also outstanding performance across multiple objective evaluation metrics. These achievements indicate the potential and practical value of DAE-GAN in the field of underwater image super-resolution. Furthermore, this approach not only provides a fresh avenue for the enhancement of underwater visual technology but also carries substantial implications for the progression of underwater image processing methodologies. Future research directions may involve exploring super-resolution processing in more complex underwater environments, such as deep-sea or multimodal underwater imagery. Potential enhancements could entail the design of novel neural network architectures tailored to deep-sea characteristics and the development of multimodal fusion algorithms. Nevertheless, these avenues also present challenges, including the computational resource requirements for handling large-scale datasets and the high costs associated with data acquisition.

6. Patents

The work reported in this manuscript has led to the filing of a patent, currently in the acceptance stage. The patent is entitled "A Method for Super-Resolution Reconstruction of Underwater Images Based on Generative Adversarial Networks", with the application number 202410044547.5. This patent encompasses an innovative approach utilizing generative adversarial network technologies for the super-resolution reconstruction of underwater image, aiming to address some of the limitations in current image processing techniques and to enhance the clarity and quality of underwater imaging. Through this patent application, this research seeks to advance the field of image processing, particularly in the area of image reconstruction in underwater environments.

Author Contributions: Conceptualization, M.G. and Z.L.; methodology, M.G.; visualization, Q.W. and W.F.; investigation, M.G. and Z.L.; experiment, M.G.; writing—original draft preparation, Q.W.; validation, Z.L. and Q.W.; project administration, M.G., Z.L. and W.F.; funding acquisition, Z.L. and W.F. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the following: 1. The Key Research and Development Program of Jiangsu Province, grant number BE2022062; 2. The Zhangjiagang Science and Technology Planning Project, grant number ZKYY2314; 3. The Doctoral Scientific Research Start-up Fund Project of Nantong Institute of Technology, grant number 2023XK(B)02.

Data Availability Statement: This manuscript encompasses all data that were produced or examined throughout the course of this study. Accompanying scripts and computational methods integral to the creation of the data will be made available in due course.

Conflicts of Interest: Wenbin Fan is employed by the company Jiangsu JBPV Intelligent Equipment Co., Ltd. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Shi, A.; Ding, H. Underwater image super-resolution via dual-aware integrated network. *Appl. Sci.* **2023**, *13*, 12985. [CrossRef]
2. 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; pp. 184–199.
3. 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; pp. 391–407.
4. Kong, X.; Zhao, H.; Qiao, Y.; Dong, C. Classsr: A general framework to accelerate super-resolution networks by data characteristic. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Nashville, TN, USA, 20–25 June 2021; pp. 12016–12025.
5. Li, Z.; Liu, Y.; Chen, X.; Cai, H.; Gu, J.; Qiao, Y.; Dong, C. Blueprint separable residual network for efficient image super-resolution. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans, LA, USA, 19–20 June 2022; pp. 833–843.
6. Shi, S.; Xiangli, B.; Yin, Z. Multiframe super-resolution of color images based on cross channel prior. *Symmetry* **2021**, *13*, 901. [CrossRef]
7. Liang, J.; Sun, G.; Zhang, K.; Van Gool, L.; Timofte, R. Mutual affine network for spatially variant kernel estimation in blind image super-resolution. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Montreal, QC, Canada, 10–17 October 2021; pp. 4096–4105.
8. Park, S.H.; Moon, Y.S.; Cho, N.I. Perception-oriented single image super-resolution using optimal objective estimation. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Vancouver, BC, Canada, 17–24 June 2023; pp. 1725–1735.
9. Ledig, C.; Theis, L.; Huszár, F.; Caballero, J.; Cunningham, A.; Acosta, A.; Aitken, A.; Tejani, A.; Totz, J.; Wang, Z. 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.
10. Zhang, K.; Liang, J.; Van Gool, L.; Timofte, R. Designing a practical degradation model for deep blind image super-resolution. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Montreal, QC, Canada, 10–17 October 2021; pp. 4791–4800.
11. Wang, X.; Xie, L.; Dong, C.; Shan, Y. Real-esrgan: Training real-world blind super-resolution with pure synthetic data. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Montreal, BC, Canada, 11–17 October 2021; pp. 1905–1914.

12. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, Ł.; Polosukhin, I. Attention is all you need. In Proceedings of the Advances in Neural Information Processing Systems, Long Beach, CA, USA, 4–9 December 2017; Volume 30.
13. Liu, Z.; Lin, Y.; Cao, Y.; Hu, H.; Wei, Y.; Zhang, Z.; Lin, S.; Guo, B. Swin transformer: Hierarchical vision transformer using shifted windows. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Montreal, QC, Canada, 10–17 October 2021; pp. 10012–10022.
14. Wu, C.; Wang, D.; Bai, Y.; Mao, H.; Li, Y.; Shen, Q. Hsr-diff: Hyperspectral image super-resolution via conditional diffusion models. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Paris, France, 1–6 October 2023; pp. 7083–7093.
15. Zhang, T.; Yang, J. Transformer with Hybrid Attention Mechanism for Stereo Endoscopic Video Super Resolution. *Symmetry* **2023**, *15*, 1947. [CrossRef]
16. Zhao, Z.; Zhang, J.; Gu, X.; Tan, C.; Xu, S.; Zhang, Y.; Timofte, R.; Van Gool, L. Spherical space feature decomposition for guided depth map super-resolution. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Paris, France, 1–6 October 2023; pp. 12547–12558.
17. Choi, H.; Lee, J.; Yang, J. N-gram in swin transformers for efficient lightweight image super-resolution. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Vancouver, BC, Canada, 17–24 June 2023; pp. 2071–2081.
18. Aghelan, A.; Rouhani, M. Underwater image super-resolution using generative adversarial network-based model. In Proceedings of the 2023 13th International Conference on Computer and Knowledge Engineering (ICCKE), Mashhad, Iran, 1–2 November 2023; pp. 480–484.
19. Umer, R.M.; Micheloni, C. Real Image Super-Resolution using GAN through modeling of LR and HR process. In Proceedings of the 2022 18th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS), Madrid, Spain, 29 November 2022–2 December 2022; pp. 1–8.
20. 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, Dermany, 8–14 September 2018.
21. Chen, Z.; Zhang, Y.; Gu, J.; Kong, L.; Yang, X.; Yu, F. Dual aggregation transformer for image super-resolution. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Paris, France, 1–6 October 2023; pp. 12312–12321.
22. Liu, Y.; Chu, Z. A Dynamic Fusion of Local and Non-Local Features-Based Feedback Network on Super-Resolution. *Symmetry* **2023**, *15*, 885. [CrossRef]
23. Gao, Y.; Liu, J.; Li, W.; Hou, M.; Li, Y.; Zhao, H. Augmented Grad-CAM++: Super-Resolution Saliency Maps for Visual Interpretation of Deep Neural Network. *Electronics* **2023**, *12*, 4846. [CrossRef]
24. Mei, Y.; Fan, Y.; Zhou, Y. Image super-resolution with non-local sparse attention. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Nashville, TN, USA, 20–25 June 2021; pp. 3517–3526.
25. Niu, B.; Wen, W.; Ren, W.; Zhang, X.; Yang, L.; Wang, S.; Zhang, K.; Cao, X.; Shen, H. Single image super-resolution via a holistic attention network. In Proceedings of the Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, 23–28 August 2020; pp. 191–207.
26. Karras, T.; Laine, S.; Aittala, M.; Hellsten, J.; Lehtinen, J.; Aila, T. Analyzing and improving the image quality of stylegan. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 13–19 June 2020; pp. 8110–8119.
27. 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.
28. Zhang, Y.; Tian, Y.; Kong, Y.; Zhong, B.; Fu, Y. Residual dense network for image super-resolution. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 2472–2481.
29. Zhou, S.; Zhang, J.; Zuo, W.; Loy, C.C. Cross-scale internal graph neural network for image super-resolution. *Adv. Neural Inf. Process. Syst.* **2020**, *33*, 3499–3509.
30. Wang, Z.; Cun, X.; Bao, J.; Zhou, W.; Liu, J.; Li, H.U. A general u-shaped transformer for image restoration. *arXiv* **2021**, arXiv:2106.03106.
31. Li, G.; Zhao, L.; Sun, J.; Lan, Z.; Zhang, Z.; Chen, J.; Lin, Z.; Lin, H.; Xing, W. Rethinking Multi-Contrast MRI Super-Resolution: Rectangle-Window Cross-Attention Transformer and Arbitrary-Scale Upsampling. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Paris, France, 1–6 October 2023; pp. 21230–21240.
32. Li, A.; Zhang, L.; Liu, Y.; Zhu, C. Feature modulation transformer: Cross-refinement of global representation via high-frequency prior for image super-resolution. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Paris, France, 1–6 October 2023; pp. 12514–12524.
33. Zhou, Y.; Li, Z.; Guo, C.-L.; Bai, S.; Cheng, M.-M.; Hou, Q. Srformer: Permuted self-attention for single image super-resolution. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Paris, France, 1–6 October 2023; pp. 12780–12791.
34. Liang, J.; Cao, J.; Sun, G.; Zhang, K.; Van Gool, L.; Timofte, R. Swinir: Image restoration using swin transformer. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Montreal, BC, Canada, 11–17 October 2021; pp. 1833–1844.

35. 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.
36. Zhang, K.; Zuo, W.; Zhang, L. Deep plug-and-play super-resolution for arbitrary blur kernels. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 15–20 June 2019; pp. 1671–1681.
37. Zhang, Y.; Li, K.; Li, K.; Wang, L.; Zhong, B.; Fu, Y. Image super-resolution using very deep residual channel attention networks. In Proceedings of the European Conference on Computer Vision (ECCV), Munich, Dermany, 8–14 September 2018; pp. 286–301.
38. Dai, T.; Cai, J.; Zhang, Y.; Xia, S.-T.; Zhang, L. Second-order attention network for single 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. 11065–11074.
39. Park, J.; Son, S.; Lee, K.M. Content-aware local gan for photo-realistic super-resolution. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Paris, France, 1–6 October 2023; pp. 10585–10594.
40. Liu, C.; Sun, D. On Bayesian adaptive video super resolution. *IEEE Trans. Pattern Anal. Mach. Intell.* **2013**, *36*, 346–360. [CrossRef] [PubMed]
41. 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]
42. Keys, R. Cubic convolution interpolation for digital image processing. *IEEE Trans. Acoust. Speech Signal Process.* **1981**, *29*, 1153–1160. [CrossRef]
43. Luo, Z.; Huang, H.; Yu, L.; Li, Y.; Fan, H.; Liu, S. Deep constrained least squares for blind image super-resolution. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans, LA, USA, 19–20 June 2022; pp. 17642–17652.
44. Zhang, W.; Li, X.; Xu, S.; Li, X.; Yang, Y.; Xu, D.; Liu, T.; Hu, H. Underwater Image Restoration via Adaptive Color Correction and Contrast Enhancement Fusion. *Remote Sens.* **2023**, *15*, 4699. [CrossRef]
45. Yang, L.; Zhang, R.-Y.; Li, L.; Xie, X. Simam: A simple, parameter-free attention module for convolutional neural networks. In Proceedings of the International Conference on Machine Learning, Vienna, Austria, 18–24 July 2021; pp. 11863–11874.
46. Selvaraju, R.R.; Cogswell, M.; Das, A.; Vedantam, R.; Parikh, D.; Batra, D. Grad-cam: Visual explanations from deep networks via gradient-based localization. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017; pp. 618–626.
47. Islam, M.J.; Xia, Y.; Sattar, J. Fast underwater image enhancement for improved visual perception. *IEEE Robot. Autom. Lett.* **2020**, *5*, 3227–3234. [CrossRef]
48. Islam, M.J.; Luo, P.; Sattar, J. Simultaneous enhancement and super-resolution of underwater imagery for improved visual perception. *arXiv* **2020**, arXiv:2002.01155.
49. 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]
50. 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]
51. 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, 4–8 June 2020; pp. 900–906.
52. 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]
53. 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, 19–20 June 2022; pp. 1103–1112.
54. 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]
55. Yue, Z.; Zhao, Q.; Xie, J.; Zhang, L.; Meng, D.; Wong, K.-Y.K. Blind image super-resolution with elaborate degradation modeling on noise and kernel. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans, LA, USA, 19–20 June 2022; pp. 2128–2138.