

Article

Not peer-reviewed version

Underwater Image Enhancement Based on Dense Residual Blocks and Attention Mechanism

[Chuang Zhou](#), [Dexing Wang](#)^{*}, [Hongchun Yuan](#)

Posted Date: 17 March 2025

doi: 10.20944/preprints202503.1206.v1

Keywords: underwater image enhancement; convolutional neural networks; multi-scale feature extraction; image processing; color correction



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a Creative Commons CC BY 4.0 license, which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Article

Underwater Image Enhancement Based on Dense Residual Blocks and Attention Mechanism

Chuang Zhou, Dexing Wang * and Hongchun Yuan

College of Information Technology, Shanghai Ocean University, Shanghai 201306, China

* Correspondence: dxwang@shou.edu.cn

Featured Application: This work aims for recovering realistic colors of underwater images, which is very important for the detection and classification of underwater targets.

Abstract: Underwater images usually suffer from blurring, contrast degradation, color deviation, and loss of local details, which hinder the visual quality and subsequent processing of underwater images. In order to address these challenges, this study proposes an underwater image enhancement method based on dense residual blocks and attention mechanisms. The encoder introduces a contextual feature aggregation module to extract multi-scale features from the input image for global feature aggregation. In the decoder, the dense residual block enriches the detailed features of the image. At the same time, the bi-level routing attention mechanism captures critical local detail features, thus effectively reducing local detail loss. Experimental results on standard datasets show that the proposed method outperforms previous deep learning models in recovering blur, color bias, and local details in underwater images.

Keywords: underwater image enhancement; convolutional neural networks; multi-scale feature extraction; image processing; color correction

1. Introduction

With the rapid development of computer vision technology, underwater imaging technology has been widely used in underwater scenes such as marine biology and archaeology, underwater surveying and mapping [1]. However, due to factors such as light absorption, scattering, and water turbidity, underwater images may experience decreased clarity, insufficient contrast, and color distortion [2], seriously affecting the performance of downstream tasks based on visual guidance. Using underwater image enhancement technology to improve the quality of underwater images can provide more substantial support for related applications in fishery aquaculture, underwater biometric recognition, underwater exploration, and other fields.

Traditional underwater image enhancement techniques, such as the Retinex algorithm [3], histogram equalization (HE) [4], white balance [5], and underwater dark channel prior (UDCP) [6], do not perform well in highly complex underwater scenes such as low illumination and fogging. With the deep development of learning technology, models based on deep learning have achieved significant success in improving image quality, such as Convolutional Neural Networks (CNN) and Generative Adversarial Networks (GAN). The image enhancement method based on the convolutional neural network automatically learns the complex structure of the image from the training data through end-to-end training, thereby generating higher-quality images. Li et al. [7] proposed a gate-controlled fusion CNN model (Water Net), which generates three inputs by applying white balance, gamma correction, and histogram equalization and then combines the three input images into an enhanced result using a gate-controlled fusion network architecture. However, the complex CNN architecture of the network is affected by backscattering. Cao et al. [8] propose a multi-

scale global-local feature fusion network (MGDFN) consisting of two cascading subnets to improve degraded underwater images progressively from coarse to fine. The optimized encoder-decoder subnetwork (OEDNet) learns multi-scale context information through a feature attention module and multi-scale fusion mechanism, and the pyramid-based feature refinement subnetwork (PFRNet) preserves fine spatial details using a detail enhancement module and a self-supervised recovery module. Deploy tail enhancement and supervised recovery modules between subnetworks to improve image color and detail recovery. An et al. [9] proposed an underwater image enhancement model based on inverted residual networks, which uses double-layer inverted residual blocks to extract deep feature information from input images, significantly improving the quality of underwater images. Jiang et al. [10] proposed the Mutual Retinex model, which combines the self-attention mechanism and convolutional neural network and achieves complementary optimization between the reflection and illumination components through the mutual representation module. The results generated by the model show excellent color consistency and naturalness. Tun et al. [11] propose a color refinement network called ColorRefineNet, which incorporates a channel and spatial attention-based Convolutional Block Attention Module (CBAM) in each ConvBlock of Shallow-UWnet and integrates a Structural Similarity (SSIM) loss function as an additional loss to enhance the color brightness and contrast of the underwater images while preserving structural details, sharpness, and color vibrancy.

Generative Adversarial Network-based image enhancement methods enhance image quality through mutual game adversarial learning of generative and discriminative models. Ma et al. [12] proposed a conditional generative adversarial network consisting of a multi-scale generator and a dual discriminator, which introduces a parallel attention mechanism combining channel and pixel attention to improve model performance. Wang et al. [13] proposed a generative adversarial network (GAN) with a lightweight U-Net for underwater image enhancement and designed the GAN generator using a lightweight U-Net. They merged the convolutional block attention module into the jump connection, effectively solving underwater images' color bias and blurring problems. Wang et al. [14] proposed a new method called the Self-Adversarial Generative Adversarial Network (SA-GAN). This method introduces a self-adversarial model, thus providing the generator with stronger decision boundary constraints to improve the quality of image generation. Tandekar et al. [15] proposed a fast underwater image enhancement generative adversarial network (FUnIE-GAN) with an augmented encoder that integrates radial and angular convolution in the encoder module of the GAN architecture to enhance the visual quality and realism of underwater images. Xin et al. [16] proposed a cGAN-based approach to image enhancement that makes use of the cooperation between the generator and discriminator to accomplish accurate enhancement of both global and local details in images. Zhang et al. [17] proposed a multi-teacher knowledge distillation GAN for underwater image enhancement (MTUW-GAN), which guides the student network to achieve joint optimization of color and detail from different degradation perspectives through a multi-teacher knowledge distillation mechanism. The method also innovatively fuses intermediate-level channel distillation with an attentional feature screening strategy to migrate hierarchical feature representations of teacher models effectively.

However, the above deep learning-based methods often overlook the problem of information loss when performing color correction on images, resulting in the loss of local details in underwater images. Moreover, the model's generalization ability is insufficient, and the effect is often unstable when applied to different scenarios. In order to solve the above problems, a method for underwater image enhancement based on dense residual blocks and attention mechanisms is proposed. The U-Net network structure is adopted, and the Contextual Feature Aggregation Module is used in the encoder part to extract multi-scale features from the input image for global feature aggregation. The Dense Residual Block (DRB) and Bi-Level Routing Attention (BRA) are used locally to capture local detail features of the image, effectively improving the color deviation of underwater images while fully preserving the local details of the image.

In CFAB, a 4×4 convolution layer is first used to obtain input features and feed them into three branches to capture different levels of features. Each branch consists of three dilated convolution layers with dilation sizes of 1, 3, and 5, which are mainly used to capture more contextual information. The output of each branch is added to the input of the next branch, which is computed as:

$$Cout_k^i = \begin{cases} Conv_d(F_c^i) & k = 1 \\ Conv_d(F_c^i + Cout_{k-1}^i) & k = 2, 3 \end{cases} \quad (1)$$

where F_c^i denotes the feature map produced by the i th CFAB block, $Cout_k^i$ denotes the output of the k th branch of the i th CFAB block, and $Conv_d$ represents the stacked dilated convolution layers. Then, the outputs of the three branches are connected by a 1×1 convolution layer to reduce the channels and add to the input features F_c^i . Finally, a 4×4 convolution layer is employed to obtain the final contextual feature map $CFAB_i$, which can be expressed as:

$$CFAB_i = Conv_{4 \times 4} \left(F_c^i + Conv_{1 \times 1} \left(Cat_{k=1}^3 (Cout_k^i) \right) \right) \quad (2)$$

where $Cat_{k=1}^3$ denotes the connection of all three branches.

2.2. Dense Residual Block (DRB)

The DenseNet [19] network is based on ResNet [20], which directly passes the output of each layer to all subsequent layers through dense connections, enhancing feature reuse and reducing the number of parameters to a certain extent. This article proposes a Dense Residual Block (DRB) that enhances feature representation by promoting complex connections, effectively reducing information degradation during encoding, and decoding processes and improving image quality. The detailed structure of the DRB module is shown in Figure 3.

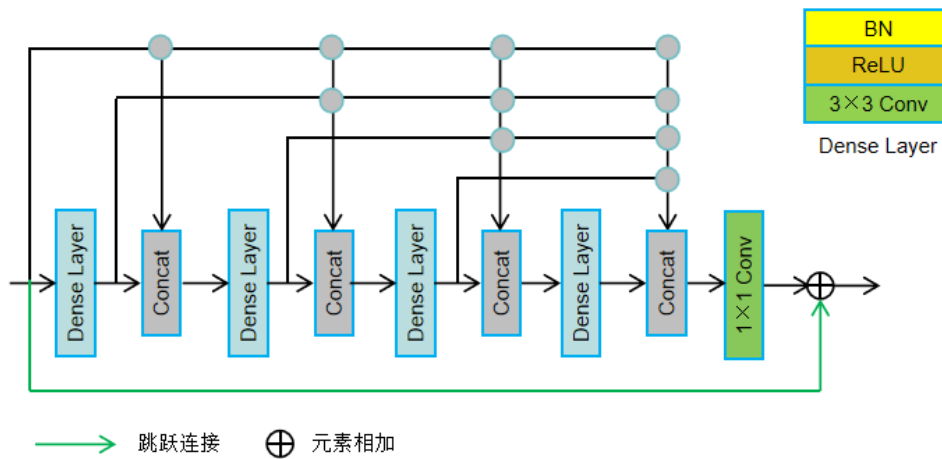


Figure 3. Dense residual block.

Each DRB module contains four dense layers, which are densely connected to facilitate the propagation of information at different network depths. Each dense layer is first regularized using batch normalization (BN), then activated using the ReLU activation function, introducing non-linear changes, and finally using a 3×3 convolution kernel to generate feature maps. The output calculation method of the dense layer is shown in formula (3):

$$O_i = Conv_{3 \times 3} \left(Cat_{k=1}^i (X, O_k) \right) \quad (3)$$

Among them, X represents the original input of the module, O_i represents the output of the i -th dense layer, and the input is the concatenation of X and the outputs of all previous dense layers.

At the end of the DRB module, a 1×1 convolution is used to reduce the number of channels in the feature map, ensuring that the number of feature channels in the input and output of the DRB module is the same, and then added to the input as the final output of the module. As shown in formula (4):

$$O = X + \text{Conv}_{1 \times 1} \left(\text{Cat}_{k=1}^4 (X, O_k) \right) \quad (4)$$

Two DRB modules were added to the encoder, which were multiplied by different decoder depths to enrich the features during the decoding process. In addition, a DRB module is also placed at the end of the decoder to enhance the refinement of the final feature map.

2.3. Bi-Level Routing Attention (BRA)

The traditional Multi-Head Self-Attention (MHSA) [21] mechanism requires each query to compute all key-value pairs, resulting in high computational complexity. Therefore, the bi-level routing attention module (BRA) is introduced. This is a query-aware sparse attention mechanism that can filter out irrelevant key-value pairs at a rough region level, leaving only a small portion of the routing region. This achieves efficient and accurate feature extraction while effectively reducing computational complexity. The structure of the BRA is shown in Figure 4.

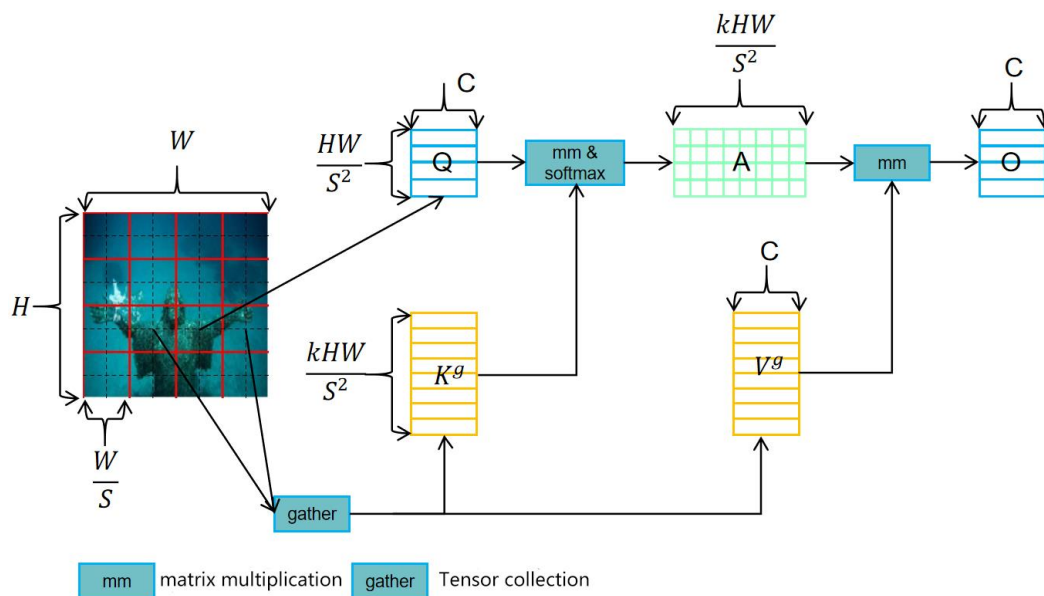


Figure 4. Bi-Level routing attention.

The implementation process of BRA is as follows:

(1) Regional division and input projection. For a 2D input feature map $X \in R^{H \times W \times C}$, divide X into $S \times S$ non-overlapping regions, each region containing $\frac{HW}{S^2}$ feature vectors. Linear projection is applied to the feature vectors in each region to obtain the query (Q) tensor, key (K) tensor, and value (V) tensor, as shown in formula (5):

$$Q = X^T W^q, K = X^T W^k, V = X^T W^v \quad (5)$$

Among $W^q, W^k, W^v \in R^{C \times C}$ are the projection weights of Q, K, and V.

(2) Region-to-region routing with a graph. Determine the correlation between regions by constructing a directed graph. Firstly, the feature vectors of Q and K are averaged within each region

to generate region-level queries Q^r and key K^r . Then, the adjacency matrix between regions is derived through matrix multiplication, as shown in formula (6):

$$A^r = Q^r (K^r)^T \quad (6)$$

Among them, the entries in the adjacency matrix A^r are used to measure the semantic correlation between regions. Next, the Top-k algorithm is used to retain each region's k most relevant regions, generating a routing index matrix I^r . Its i-th row contains the k indices of the most relevant region for the i-th region, as shown in formula (7):

$$I^r = \text{topkIndex}(A^r) \quad (7)$$

(3) Token to token attention. Calculate token-to-token attention in the union of k-selected routing regions using the routing index matrix I^r . Due to the possibility of routing regions being scattered across the entire feature map, direct computation efficiency is relatively low. Therefore, first collect the key value tensor of the relevant region, as shown in formula (8):

$$K^g = \text{gather}(K, I^r), V^g = \text{gather}(V, I^r) \quad (8)$$

Among $K^g, V^g \in \mathbb{R}^{S^2 \times \frac{kHW}{S^2} \times C}$ is the key value tensor. Then, perform attention operations on the collected key-value pairs, as shown in formula (9):

$$O = \text{Attention}(Q, K^g, V^g) \quad (9)$$

2.4. Loss Function

To optimize the performance of the model, mean square error (MSE) loss and perceptual loss inspired by the VGG 19 network [22] were used to train the model. Integrating perceptual loss can enhance the consistency between model output and perceptual quality, reflecting human perceptual characteristics. The overall loss function is shown in formula (10):

$$L = L_{MSE} + \lambda L_{VGG} \quad (10)$$

Among them, L_{MSE} represents the MSE loss, L_{VGG} represents the perceptual loss obtained from the intermediate feature map of the VGG19 network. The balance between hyperparameter λ controls losses.

The mean square error loss L_{MSE} calculates the difference between the pixel values of the generated image and the reference image to optimize model training, as shown in formula (11):

$$L_{MSE} = \frac{1}{N} \sum_{i=1}^N (\hat{I}_i - I_i)^2 \quad (11)$$

Among them, N represents the batch size, \hat{I}_i represents the output of the model, and I_i represents the true value.

The perception loss L_{VGG} can correct the color distribution of an image, making the enhanced image more detailed, as shown in formula (12):

$$L_{VGG} = \frac{1}{N} \sum_{i=1}^N (VGG(\hat{I}) - VGG(I_i))^2 \quad (9)$$

Among them, N represents the batch size, I_i represents the i-th input image, $VGG(\hat{I})$ and $VGG(I_i)$ respectively represent the feature maps extracted from the real image (I_i) corresponding to the output of the model (\hat{I}_i) in this paper.

3. Results

3.1. Datasets

The dataset used for training in this study is a mixed dataset of the EUVP dataset [23] and the UIEBD dataset [24]. Over 10000 pairs of paired underwater image instances and 7000 unpaired underwater image instances are in the EUVP dataset, collected using seven different cameras in various scenarios. After manual selection, it can adapt to various natural changes in the data, but there is a problem of poor perceptual quality. The UIEBD dataset contains 890 pairs of paired underwater image instances and 60 unpaired underwater image instances, covering a variety of complex underwater environments. Different images exhibit varying degrees of degradation patterns, and the reference images are manually adjusted by experts as high-quality visual enhancement targets. However, the dataset size is small and insufficient to meet the large-scale training needs of deep learning models. Therefore, this study used a mixed dataset consisting of the EUVP and UIEBD datasets. The training set was composed of 2000 pairs of underwater images selected from the EUVP dataset and 800 pairs of high-quality paired images selected from the UIEBD dataset, mainly to improve the robustness and generalization ability of the model. The test set is divided into two parts: test set A is the remaining 90 paired images of the UIEBD dataset, and test set B is the 60 reference-free images of the UIEBD dataset.

3.2. Experimental Setup

The implementation of this experiment is based on the PyTorch framework, using Python version 3.8 and NVIDIA's GeForce RTX3090 GPU. During the training process, the input image will be adjusted to 256×256 , the batch size will be set to 16, and the Adam [25] optimizer will be used to optimize the performance of the model. The initial learning rate will be set to 1×10^{-4} and gradually reduced to 1×10^{-6} using the cosine annealing algorithm [26], with a total of 100 iterations.

3.3. Compared Methods

In order to evaluate the effectiveness of the proposed method, we compare it with classical underwater image enhancement methods, including three traditional methods (CLAHE [27], UDCP, ICM [30]) and two deep learning-based methods (Water-Net [24], FUnIE-GAN [23]).

3.4. Quantitative Comparisons

The quantitative comparison results on test set A are shown in Table 1. This study used three fully referenced image evaluation metrics: Structural Similarity Index (SSIM), Peak Signal Noise Ratio (PSNR) [29], and Learning Perceived Image Block Similarity (LPIPS) [30]. Among them, PSNR measures the pixel similarity between the enhanced and reference images. In contrast, SSIM measures the similarity in structure and texture between the enhanced image and the reference image. PSNR and SSIM values have higher scores, indicating better surface image quality. LPIPS reflects the visual perception difference between the enhanced and reference images, with a lower LPIPS score indicating better performance. As can be seen from Table 1, our method outperforms the comparison algorithm in both PSNR and LPIPS metrics. Although it is not optimal in SSIM metrics, it is only lower than the Water Net and FUnIE-GAN methods and significantly higher than the other three comparison methods. The experimental results show that the proposed method can significantly enhance the details and clarity of images, and the enhanced images are closer to human visual perception, demonstrating the effectiveness of the method.

Table 1. Quantitative comparison results of test set A.

Method	SSIM	PSNR	LPIPS
CLAHE	0.723	19.378	0.354
UDCP	0.582	15.453	0.369

ICM	0.738	21.293	0.315
Water-Net	0.796	23.879	0.261
FUnIE-GAN	0.817	24.382	0.226
Ours	0.784	25.235	0.182

Three reference-free measurement methods were used for quantitative comparison on test set B: Underwater Color Image Quality Assessment (UCIQE) [31], Underwater Image Quality Measurement (UIQM) [32], and Natural Image Quality (NIQE) [33]. UCIQE is an indicator used to evaluate the color quality of underwater images. The higher the UCIQE value, the more natural and contrasting the color of the image. UIQM combines multiple image quality assessment standards to evaluate various aspects of an image comprehensively. The higher the UIQM value, the better the image quality. NIQE is based on the statistical characteristics of natural scenes and measures whether an image conforms to the distribution of natural scenes. The smaller the value, the better the effect. The quantitative comparison results on test set B are shown in Table 2. Our method is slightly lower than the Water Net method in the UCIQE index but higher than FUnIE-GAN and other comparison methods. The NIQE index is lower than the FUnIE-GAN method and higher than other comparisons, such as Water-Net. The UIQM metric outperforms all comparison algorithms. Overall, the method proposed in this paper is superior to the Water-Net and FUnIE-GAN methods in enhancing image quality and significantly outperforms the three traditional methods of CLAHE, UDCP, and ICM.

Table 2. Quantitative comparison results of test set B.

Method	UCIQE	UIQM	NIQE
CLAHE	0.492	2.434	7.625
UDCP	0.519	2.326	7.779
ICM	0.488	2.351	7.572
Water-Net	0.564	2.615	7.853
FUnIE-GAN	0.535	2.636	8.126
Ours	0.557	2.715	7.833

3.5. Visual Comparisons

Figure 5 shows the visual comparison results on test set A. According to Figure 5, it can be seen that the CLAHE method improves the brightness and contrast of the image, but there is a problem of excessive color deviation in some areas. The image enhanced by the UDCP method tends to be darker overall, especially in deep water areas where details are almost lost, and color distortion is severe, resulting in unnatural blue-green tones. The brightness of the image enhanced by the ICM method is higher than that of UDCP. However, it is still darker, with improved blue-green tones, insufficient image contrast, and a weak ability to represent details. The Water Net method significantly improves the overall brightness and contrast of the image, but the color correction of the overall image is insufficient, and blue-green deviation still exists. The overall details and contrast of the image enhanced by the FUNIE GAN method show good performance, with significant improvement in color correction, but occasional slight color distortion may occur. In contrast, the method proposed in this article performs well in contrast, brightness, and color correction. The details in deep water areas are also well preserved, and it is obvious that due to the contrast method, it is very close to the reference image.

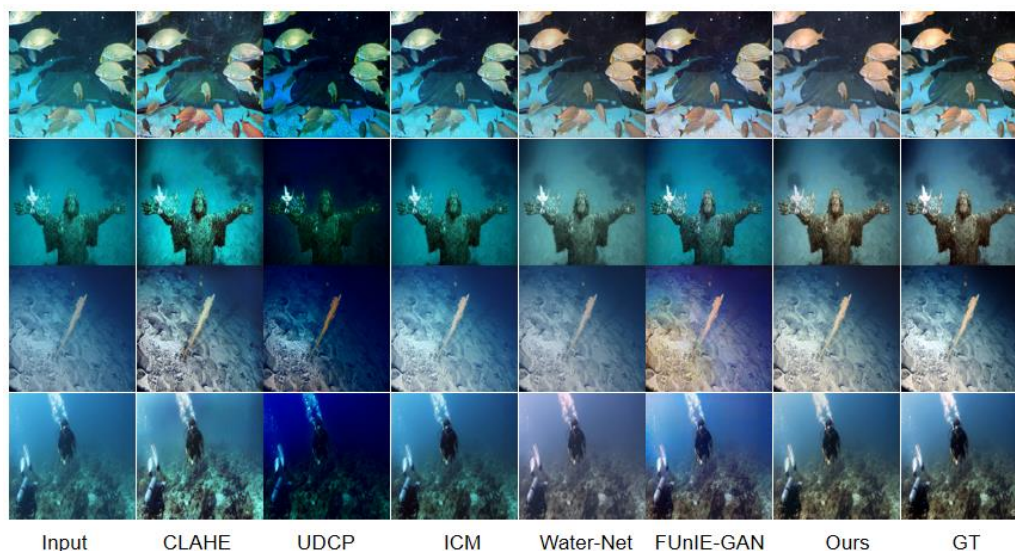


Figure 5. Visual comparison results of test set A.

Figure 6 shows the visual comparison results on test set B without reference. It can be seen that the CLAHE method does not fully restore the color of the image, and there is still a significant blue-green color cast in the image. In addition, some areas in the image are excessively enhanced, resulting in high brightness or loss of details. The image enhanced by the UDCP method tends to be darker overall, resulting in the loss of some details and insufficient color restoration. The ICM method has a reasonable contrast enhancement effect, with some details being more precise, but there is still a certain degree of blue-green residue in the overall image. The Water Net method eliminates most of the blue-green color cast in the image, providing a good color reproduction effect. However, there are slight over-enhancement phenomena in some areas of the image, and the details are slightly unnatural. The FUNIE GAN method performs well in color correction and contrast. However, there is a slight red color shift in some areas of the image, and some images' degree of detail restoration is still insufficient. In contrast, the method proposed in this article performs well in detail preservation and color correction, capable of restoring more details of underwater images, and the enhanced image colors are closer to the actual scene. In addition, the contrast and brightness adjustments of the images are more natural and in line with human visual perception, which proves that this model can effectively improve the quality of generated images.

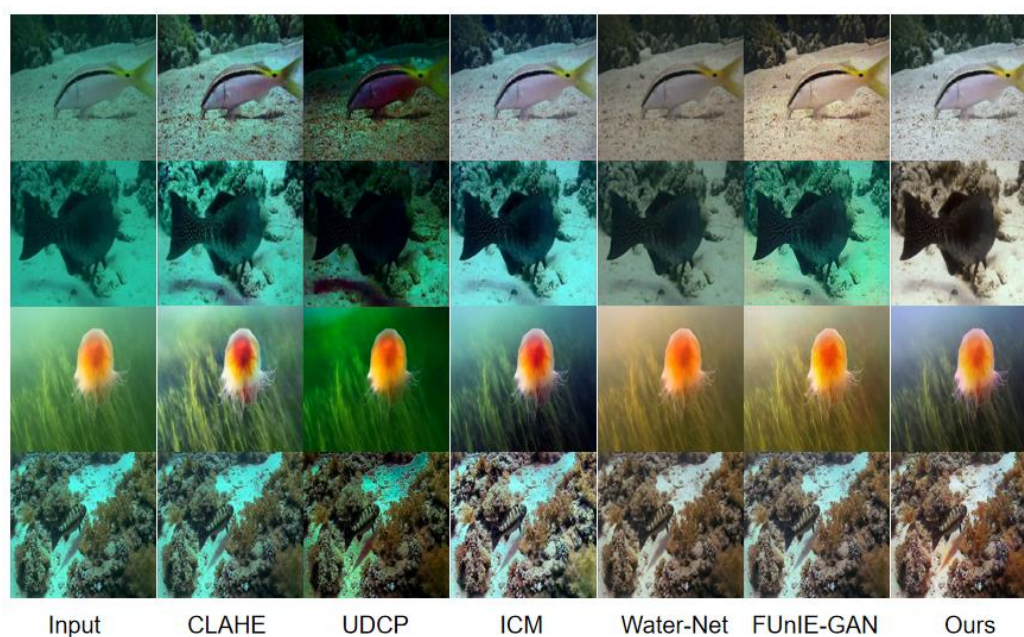


Figure 6. Visual comparison results of test set B.

3.6. Ablation Study

In order to verify the effectiveness and necessity of each module in the model, ablation experiments were conducted on each module. The experimental parameter settings should be consistent with the comparative experiments mentioned above.

The results of the ablation experiments on test set A are shown in Table 3. The results in the table show that the CFAM module has a significant impact on the SSIM and LPIPS indicators. The reason for this is that the CFAM module uses multi-scale dilated convolution blocks for feature extraction, effectively preserving the image's high-level and structural features. The DRB module and BRA module have a significant impact on the PSNR metric, with a focus on feature preservation at the pixel level and color correction. The combination of the CFAM module, DRB module, and BRA module results in the best performance of the model, which can comprehensively improve the visual effect of the image.

Table 3. Ablation experimental results of test set A.

CFAM	DRB	BRA	SSIM	PSNR	LPIPS
			0.651	16.037	0.582
√			0.743	19.324	0.328
	√		0.726	20.691	0.436
		√	0.717	21.552	0.453
	√	√	0.762	23.368	0.287
√	√	√	0.784	25.235	0.182

The results of the ablation experiments on test set B are shown in Table 4. After analyzing the data in the table, we see that the CFAM module significantly impacts the NIOE index, and the image is more in line with the natural scene distribution. The DRB module and BRA module significantly improve the color distribution and contrast of the image, thus having a significant impact on the UCIQE and UIQM metrics. The combination of the CFAM module, DRB module, and BRA module results in the best performance of the model, which can comprehensively improve the visual effect of the image.

Table 4. Ablation experimental results of test set B.

CFAM	DRB	BRA	UCIQE	UIQM	NIQE
			0.421	2.246	8.314
√			0.449	2.326	7.987
	√		0.462	2.527	8.129
		√	0.485	2.468	8.035
	√	√	0.526	2.623	7.952
√	√	√	0.557	2.715	7.833

Based on the results of the ablation experiments on the two test sets above, the CFAM module effectively improves the model's retention of overall image features. In contrast, the DRB module and BRA module enhance the model's local feature extraction for images. When used in combination, the model can improve image quality in multiple dimensions, thereby comprehensively enhancing the visual effect of the image.

4. Conclusions

This study proposes an underwater image enhancement method based on dense residual blocks and attention mechanisms. The overall U-Net network structure is adopted, and multiple multi-scale dilated convolution blocks are used in the CFAM module to continuously enlarge the receptive field, which can effectively aggregate the contextual features of the image and help the network capture the global features of the image. The DRB module enhances feature representation through dense residual connections, which can enrich features during decoding. The BRA module uses a sparse attention mechanism to focus on key regions in the features during the decoding process, achieving efficient and accurate feature extraction. Through the synergistic effect of the three modules, the model-enhanced image is closer to the reference image. It is better than the comparative image enhancement methods in terms of color, clarity, and contrast, as well as more prosperous and more natural detail performance and better visual effect. In the future, the effectiveness of the model for image enhancement will be further improved while reducing the number of parameters and floating-point calculations, making the model applicable to low-resource real-time enhancement scenarios to meet a wider range of application requirements.

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

Funding: This research was funded by National Natural Science Foundation of China, grant number 41776142.

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

References

1. Wang, Y.; Song, W.; Fortino, G.; Qi, L.-Z.; Zhang, W.; Liotta, A. An Experimental-Based Review of Image Enhancement and Image Restoration Methods for Underwater Imaging. *IEEE Access* 2019, 7, 140233–140251.
2. Hu, K.; Weng, C.; Zhang, Y.; Jin, J.; Xia, Q. An Overview of Underwater Vision Enhancement: From Traditional Methods to Recent Deep Learning. *J. Mar. Sci. Eng.* 2022, 10, 241.
3. Fu, X.; Zhuang, P.; Huang, Y.; Liao, Y.; Zhang, X.-P.; Ding, X. A Retinex-Based Enhancing Approach for Single Underwater Image. In *Proceedings of the 2014 IEEE International Conference on Image Processing (ICIP)*, Paris, France, 2014, pp. 4572-4576.
4. Muniyappan, S.; Allirani, A.; Saraswathi, S. A novel approach for image enhancement by using contrast limited adaptive histogram equalization method. In *Proceedings of the 2013 Fourth International Conference on Computing, Communications and Networking Technologies (ICCCNT)*, Tiruchengode, India, 2013, pp. 1-6.
5. Zhang, H.; Li, D.; Sun, L.; Li, Y. An Underwater Image Enhancement Method Based on Local White Balance. In *Proceedings of the 2020 5th International Conference on Mechanical, Control and Computer Engineering (ICMCCE)*, Harbin, China, 2020, pp. 2055-2060.
6. Gur Emre Guraksin; Kose, U.; Deperlioglu, O. Underwater image enhancement based on contrast adjustment via differential evolution algorithm. In *Proceedings of the 2016 International Symposium on INnovations in Intelligent SysTems and Applications (INISTA)*, Sinaia, Romania, 2016, pp. 1-5.
7. Li, C.; Guo, C.; Ren, W.; Cong, R.; Hou, J.; Kwong, S.; Tao, D. An Underwater Image Enhancement Benchmark Dataset and Beyond. *IEEE Trans. Image Process.* 2020, 29, 4376-4389.
8. Cao, W.; Zhang, W. Multi Scale Perceptual Underwater Image Enhancement Based on Feature Fusion Network. In *Proceedings of the 2024 IEEE 6th Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC)*, Chongqing, China, 2024, pp. 1128-1136.

9. An, Y.; Feng, Y.; Yuan, N.; Ji, Z.; Ganchev, I. DIRBW-Net: An Improved Inverted Residual Network Model for Underwater Image Enhancement. *IEEE Access*. 2024, 12, 75474–75482.
10. Jiang, K.; Wang, Q.; An, Z.; Wang, Z.; Zhang, C.; Lin, C.-W. Mutual Retinex: Combining Transformer and CNN for Image Enhancement. *IEEE Trans. Emerging Top. Comput.* 2024, 8, 2240–2252.
11. Tun, M. T.; Shimamura, T. ColorRefineNet: Color Refining Underwater Image Enhancement Based on Attention Network. In *Proceedings of the 2024 5th International Conference on Advanced Information Technologies (ICAIT)*, Yangon, Myanmar, 2024, pp. 1-6.
12. Ma, C.; Yang, L.; Hu, H.; Chen, Y.; Bu, A. Underwater Image Enhancement Method Based on CGAN and Parallel Attention Mechanism. In *Proceedings of the 2023 3rd International Conference on Neural Networks, Information and Communication Engineering (NNICE)*, Guangzhou, China, 2023, pp. 171-174.
13. Wang, X.; Zhou, H.; Le, X.; Ding, J.; Tao, H.; Cheng, L.; Chen, W.; Wang, R.; Yang, Q.; Chen, C.; Kong, M. Generative Adversarial Network with Lightweight U-Net for Underwater Optical Image Enhancement. In *Proceedings of the 2024 12th International Conference on Intelligent Computing and Wireless Optical Communications (ICWOC)*, Chongqing, China, 2024, pp. 29-34.
14. Wang, H.; Yang, M.; Yin, G.; Dong, J. Self-Adversarial Generative Adversarial Network for Underwater Image Enhancement. *IEEE J. Oceanic Eng.* 2024, 49, 237–248.
15. Minal Tandekar; Anil Singh Parihar. Underwater Image Enhancement through Deep Learning and Advanced Convolutional Encoders. In *Proceedings of the 2022 13th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, Kamand, India, 2024, pp. 1-6.
16. Xin, D.; Wang, X.; Wang T. Image Enhancement Method Based on Conditional Generative Adversarial Network. In *Proceedings of the 2024 IEEE 7th International Conference on Information Systems and Computer Aided Education (ICISCAE)*, Dalian, China, 2024, pp. 850-857.
17. Zhang, T.; Liu, Y. MTUW-GAN: A Multi-Teacher Knowledge Distillation Generative Adversarial Network for Underwater Image Enhancement. *Appl. Sci.* 2024, 14, 529.
18. Chong, F.; Dong, Z.; Yang, X.; Zeng, Q. SAR and Multispectral Image Fusion Based on Dual-Channel Hybrid Attention Block and Dilated Convolution. In *Proceedings of the 2023 3rd International Conference on Neural Networks, Information and Communication Engineering (NNICE)*, Guangzhou, China, 2023, pp. 602-607.
19. Huang, G.; Liu, Z.; Van Der Maaten, L.; Weinberger, K. Q. Densely Connected Convolutional Networks. In *Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI, USA, 2017, pp. 2261-2269.
20. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. In *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA, 2016, pp. 770-778.
21. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; N. Gomez, A.; Kaiser, Ł.; Polosukhin, I. Attention is all you need[J]. In *Proceedings of the 31st International Conference on Neural Information Processing Systems (NIPS'17)*. Curran Associates Inc., Red Hook, NY, USA, 2017, pp. 6000–6010.
22. Johnson, J.; Alahi, A.; Fei-Fei, L. Perceptual Losses for Real-Time Style Transfer and Super-Resolution. In *Proceedings of the Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, 2016*, pp:694-711.
23. Islam, J.; Xia, Y.; Sattar, J. Fast Underwater Image Enhancement for Improved Visual Perception. *IEEE Rob. Autom. Lett.* 2020, 5, 3227–3234.
24. Li, C.; Guo, C.; Ren, W.; Cong, R.; Hou, J.; Kwong, S.; Tao, D. An Underwater Image Enhancement Benchmark Dataset and Beyond. *IEEE Trans. Image Process.* 2020, 29, 4376-4389.
25. ZHANG, Z. Improved adam optimizer for deep neural networks. In *Proceedings of the 2018 IEEE/ ACM 26th international symposium on quality of service (IWQoS)*, Banff, AB, Canada, 2018, pp. 1-2.
26. Loshchilov, I.; Hutter, F. SGDR: Stochastic Gradient Descent with Warm Restarts. In *Proceedings of the 5th International Conference on Learning Representations*, Washington DC, USA, 2017, pp. 1-16.
27. Reza, A. M. Realization of the Contrast Limited Adaptive Histogram Equalization (CLAHE) for Real-Time Image Enhancement. *J SIGNAL PROCESS SYS.* 2004, 38, 35-44.

28. Iqbal, K.; Salam, R.A.; Osman, A.M.; Talib, A.Z. Underwater image enhancement using an integrated colour model. *IAENG Int. J. Comput. Sci.* 2007, 34, 239-244.
29. Horé, A.; Ziou, D. Image quality metrics: PSNR vs. SSIM. In *Proceedings of the 2010 20th International Conference on Pattern Recognition, Istanbul, Turkey, 2010*, pp. 2366-2369.
30. Zhang, R.; Isola, P.; Efros, A. A.; Shechtman, E.; Wang, O. The unreasonable effectiveness of deep features as a perceptual metric. In *Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 2018*, pp. 586-595.
31. Yang, M.; Sowmya, A. An Underwater Color Image Quality Evaluation Metric. *IEEE Trans. Image Process.* 2015, 24, 6062-6071.
32. Panetta, K.; Gao, C.; Aghaian, S. S. Human-Visual-System-Inspired Underwater Image Quality Measures. *IEEE J. Oceanic Eng.* 2016, 41, 541-551.
33. Mittal, A.; Soundararajan, R.; Bovik, A. C. Making a “Completely Blind” Image Quality Analyzer. *IEEE Signal Process Lett.* 2013, 20, 209–212.

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.