Article

RECAD: Retinex-Based Efficient Channel Attention with Dark Area Detection for Underwater Images Enhancement

Tianchi Zhang [1,*], Qiang Liu [1], Hongwei Qin [1] and Xing Liu [2]

[1] School of Information Science and Engineering, Chongqing Jiaotong University, Chongqing 400074, China
[2] School of Mechanical and Electrical Engineering, Harbin Engineering University, Harbin 150001, China
* Correspondence: zhangtianchi@cqjtu.edu.cn

Abstract

Focusing on visual target detection for Autonomous Underwater Vehicles (AUVs), this paper investigates enhancement methods for weakly illuminated underwater images, which typically suffer from blurring, color distortion, and non-uniform illumination. Although deep learning-based approaches have received considerable attention, existing methods still face limitations such as insufficient feature extraction, poor detail detection, and high computational costs. To address these issues, we propose RECAD—a lightweight and efficient underwater image enhancement method based on Retinex theory. The approach incorporates a dark region detection mechanism to significantly improve feature extraction from low-light areas, along with an efficient channel attention module to reduce computational complexity. A residual learning strategy is adopted in the image reconstruction stage to effectively preserve structural consistency, achieving an SSIM value of 0.91. Extensive experiments on the UIEB and LSUI benchmark datasets demonstrate that RECAD outperforms state-of-the-art models including FUnIEGAN and U-Transformer, achieving a high SSIM of 0.91 and competitive UIQM scores (up to 3.19), while improving PSNR by 3.77 dB and 0.69–1.09 dB, respectively, and attaining a leading inference speed of 97 FPS, all while using only 0.42 M parameters, which substantially reduces computational resource consumption.

Keywords: underwater image enhancement; retinex theory; efficient channel attention; dark area detection

1. Introduction

Autonomous Underwater Vehicles (AUVs) [1] are widely employed in the field of underwater target detection due to their strong autonomy and extensive operational range. Underwater imaging serves as the primary method by which AUVs acquire close-range target information. However, due to light scattering and attenuation in water, the captured underwater images often suffer from degradation such as blurring, loss of target and edge details, and low contrast, presenting a hazy appearance overall [2,3]. Consequently, image enhancement is crucial for providing AUVs with clear and accurate target information.

Underwater image enhancement must carefully account for lighting conditions. Images captured under near-surface solar light differ significantly from those under artificial deep-sea lighting. Under solar light, the foreground typically appears darker than the background, whereas the opposite occurs with artificial illumination [4]. Solar light provides relatively uniform illumination with predictable attenuation patterns, while artificial light is highly non-uniform and exhibits non-linear intensity decay, the characteristics of

which remain not fully understood [5]. Consequently, different enhancement methods are required for these two scenarios [6]. During the process of deep-sea exploration with AUVs, artificial light serves as the sole illumination source. However, its limited intensity decays significantly in water, often leaving distant areas of the image in a low-light state. Such conditions not only exacerbate the aforementioned degradation issues but also exhibit substantially reduced signal-to-noise ratio, increased noise, and severe loss of critical detail [7]. Hence, enhancing effective information extraction capability in low-light regions remains a central challenge in deep-sea image enhancement.

To address issues of weakly illuminated underwater images, researchers have developed various solutions based on Underwater Image Enhancement (UIE) techniques [8], which can broadly be categorized into traditional methods and deep learning-based methods. Traditional UIE methods are further divided into two categories: physics-based and non-physics-based techniques. Physics-based methods simulate the underwater optical imaging process and attempt to reconstruct clear images by estimating model parameters. However, their high computational cost limits their suitability for real-time AUV target detection. In contrast, non-physics-based methods focus on enhancing local or specific visual qualities. Representative methods include Underwater Dark Channel Prior (UDCP) [9], Image Blurriness and Light Absorption (IBLA) [10], and Contrast-Limited Adaptive Histogram Equalization (CLAHE) [11]. Despite their effectiveness in certain scenarios, most of these methods neglect the distinct characteristics of artificially illuminated and solar-illuminated underwater images, often resulting in regional over-enhancement, detail loss, and inadequate noise suppression.

With the rapid development of AI technology, deep learning-based underwater image processing has emerged as a major research direction and a hot topic in the field. Researchers have investigated a range of architectures, which can be broadly categorized into three categories: Generative Adversarial Networks (GANs), Convolutional Neural Networks (CNNs), and Transformers [12]. GAN-based methods utilize adversarial training to improve image quality but often suffer from training instability, leading to artifacts or unrealistic results—challenges that become particularly pronounced under artificial lighting conditions. For instance, FUnIEGAN [13] integrates physics-model guidance into adversarial training to produce visually natural and high-fidelity results. However, it still suffers from the dual drawbacks of unstable training and high computational cost. More recently, many researchers have explored hybrid architectures combining CNNs and Transformers for underwater image enhancement, such as U-Transformer [14]. These methods excel in global modeling and are particularly effective in handling overall degradation caused by artificial light scattering. Nonetheless, they demand significant computational resources and large training data, making it challenging to meet the real-time processing requirements of AUVs during missions. Therefore, designing lightweight model structures remains a core challenge for deep-sea image enhancement in AUV applications.

To address these challenges, this paper proposes RECAD (Retinex-based Efficient Channel Attention with Dark area Detection) to achieve high-quality and computationally efficient enhancement of weakly illuminated underwater images. Firstly, RECAD adopts the Retinex theory to decompose a low-light underwater image into an illumination map and a reflectance map. Unlike previous methods, it incorporates a dark area detection module to explicitly locate regions requiring enhancement. This allows the illumination map to be processed through multi-scale and independent pathways for targeted feature enhancement. Then, RECAD utilizes a custom U-shaped enhancer to refine the input image and recover finer details. The enhanced grayscale image is element-wise multiplied with the reflectance map, and the original image is added back to preserve structural similarity. Finally, a denoising step is applied to generate a more visually consistent result.

The main technical contributions of this paper are summarized as follows:

- A Retinex-based single-stage enhancement network: To tackle the challenges of computational efficiency and model complexity bottlenecks in enhancing weakly illuminated underwater images, we propose RECAD, a lightweight single-stage enhancement network. This network avoids complex multi-stage training pipelines and significantly reduces computational cost while maintaining enhancement performance comparable to state-of-the-art methods.
- Dark area detection mechanism: To enhance feature extraction in poorly illuminated regions (which often contain critical details), RECAD innovatively incorporates a dark area detection module. This module adaptively identifies and emphasizes low-light regions, enabling the network to perform more targeted enhancement.
- Efficient channel attention (ECA) mechanism: To balance performance and computational efficiency, RECAD incorporates an ECA module. This mechanism adaptively recalibrates channel-wise feature responses without significantly increasing parameters or computational load, thereby enhancing the network's discriminative feature extraction capability and overall enhancement quality.
- Residual learning for improved reconstruction quality: To ensure the enhanced image retains authentic structure with minimal distortion, RECAD introduces residual learning in the image reconstruction module. This strategy maintains geometric structural consistency between the output and the original low-light image, leading to a higher Structural Similarity Index (SSIM) performance.

This paper is organized into four subsequent sections. Commencing with the theoretical foundations in Section 2, it proceeds to elaborate on the detailed architecture of the proposed RECAD framework in Section 3. Section 4 then showcases a comparative analysis with established methods, and the work concludes with final remarks in Section 5.

2. Materials and Methods

This section establishes the theoretical foundation for proposed RECAD method. First, we analyze the categories, advantages, and key challenges of existing underwater image enhancement (UIE) methods, including issues such as model simplification, difficulty in parameter estimation, limited interpretability, and poor generalization. Next, we introduce Retinex theory, whose image decomposition provides insight into underwater degradation and guides the design of physical models. Finally, we briefly describe the Efficient Channel Attention Network (ECANet), whose lightweight and efficient characteristics help control model complexity. Collectively, these theories form the basis of the RECAD method.

2.1. UIE Methods

Research into underwater image degradation has yielded various enhancement (UIE) schemes, which are predominantly grouped as physics-based, non-model-based, and deep learning-based methods [8].

Methods based on physical models, such as the Jaffe-McGlamery underwater imaging model [15], restore degraded underwater images into clear ones by estimating model parameters (e.g., background light, transmission map, and attenuation coefficients), offering good physical interpretability. However, accurately estimating these parameters in complex underwater environments is challenging, and such methods often fail to meet the real-time requirements of AUV-based target image processing. In contrast, non-model-based methods enhance images by adjusting contrast or brightness distributions, offering advantages in algorithmic simplicity and computational efficiency. Nevertheless, they lack an understanding of the underwater imaging physical process and local scene characteristics,

making them susceptible to artifacts, noise amplification, and reduced adaptability in complex water conditions and low-light scenarios.

The rapid progress in artificial intelligence has made deep learning a prominent approach in underwater image enhancement. These methods leverage deep neural networks (DNNs) to learn the mapping from degraded to clear images, enabling the restoration of degraded underwater imagery. Early representative works, such as WaterNet [16] and UWGAN [17], utilized convolutional neural networks (CNNs) or generative adversarial networks (GANs), demonstrating the potential of deep learning in UIE tasks. Furthermore, Retinex-based approaches have been integrated with CNNs (e.g., RetinexNet) for low-light enhancement [18], while subsequent work by Wang et al. [19] demonstrates continued improvements in this line of research. These methods highlight the potential of combining physical models with deep learning. Subsequent studies incorporated strategies like dense connections (e.g., shallow networks with dense connections) to mitigate overfitting. The field has recently extended to incorporate Transformer-based methods (e.g., U-Transformer [14]) for superior feature fusion and global modeling, as well as diffusion models (e.g., WF-Diff, Reti-Diff) that achieve state-of-the-art generation quality, despite their higher computational intensity. Furthermore, comprehensive frameworks like Metalantis [20] have been proposed to address a wide range of underwater degradations through unified deep learning architectures. Despite these advances, data-driven methods still face core challenges, including insufficient theoretical guidance, weak interpretability of the enhancement process, and limited generalization to unseen degradation patterns (cross-degradation).

To overcome the limitations of existing methods—such as oversimplification, parameter estimation challenges, and poor interpretability—this paper proposes an integrated framework merging Retinex theory with convolutional networks and attention mechanisms. This integration is designed to balance high expressive capability with model interpretability, and its lightweight nature ensures enhanced practical usability.

2.2. Retinex Theory

Retinex theory [21], a cornerstone model of color vision, offers a computational explanation for how the human visual system separates the influence of illumination from an object's intrinsic color to achieve color constancy. This theory has furnished a valuable mathematical framework for simulating perceptual mechanisms, leading to its broad adoption in low-light image enhancement. The core principle of Retinex is mathematically represented by the following equation:

$$S = I \cdot R \tag{1}$$

where S is the captured image, I signifies the illumination map, R corresponds to the reflectance map, and the $\cdot$ operator indicates element-wise multiplication over the spatial dimensions (H $\times$ W) for each RGB channel.

Based on this formulation, the image S is factorized into two multiplicative components: the illumination and the reflectance. The illumination map I characterizes the external lighting environment, accounting for factors such as ambient light, direct sources, and shadows. The reflectance map R, in contrast, embodies the inherent surface properties and color of objects, which remain constant under different lighting.

Owing to its effective decomposition, methodologies grounded in Retinex theory have found utility in numerous imaging tasks, including low-light and underwater image enhancement, as well as high dynamic range (HDR) imaging [22]. By isolating and independently processing the illumination component, these approaches adeptly mitigate challenges posed by non-uniform lighting and significantly improve visual quality.

2.3. Efficient Channel Attention Network

The Efficient Channel Attention Network (ECANet) [23] introduces a lightweight and efficient channel attention mechanism designed to optimize the computational efficiency and performance of channel dependency modeling. Its core innovation lies in abandoning the dimensionality-reducing fully connected layers used in traditional channel attention mechanisms. Instead, ECANet employs one-dimensional convolution (1D convolution) without dimensionality reduction to directly capture local cross-channel interactions.

ECANet operates through three key steps. First, channel statistical descriptors $(1 \times 1 \times C)$ are obtained via global average pooling. Next, a 1D convolutional layer (with kernel size k adaptively determined based on the number of channels C) is applied to this descriptor to efficiently learn dependencies between adjacent channels and compute channel attention weights. Finally, the attention weights are normalized using a Sigmoid activation function and used to recalibrate the original input feature map channel-wise, thereby enhancing discriminative features.

The primary advantage of ECANet is its high efficiency. By avoiding dimensionality reduction and complex fully connected layers, and requiring only a lightweight 1D convolutional layer (with significantly reduced parameters), it effectively models interdependencies between channels. This enables ECANet to maintain excellent performance on benchmark tasks such as image classification while substantially reducing model complexity and computational overhead. Consequently, this mechanism is widely used in mainstream computer vision tasks such as image classification and object detection and is particularly suitable for resource-constrained lightweight network architectures. Compared to typical channel attention methods, ECANet achieves equal or superior feature selection capability at a lower computational cost.

3. RECAD

To address the complex degradation inherent in weakly illuminated underwater images, this paper proposes RECAD, a Retinex-based lightweight and efficient method for underwater image enhancement method. Unlike traditional singular enhancement approaches, RECAD adopts a multi-stage, well-defined, and collaboratively optimized strategy. By introducing a dark region detection mechanism and an integrated post-processing unit, RECAD achieves joint and precise control over image illumination, contrast, and noise levels. This section provides a detailed explanation of the overall design philosophy of the RECAD network and the architectural details of its core modules.

3.1. Architecture of RECAD

As illustrated in Figure 1, the proposed method consists of five modules: the Decom Module, the Dark Area Detection Module, the Illumination Enhancement Module, the Refinement Module, and the Denoising Module.

The working principle and process of RECAD are described as follows:

(1) Based on Retinex theory, the decomposition module separates the weakly illuminated underwater image into a reflectance map and an illumination map.
(2) The dark region detection module is responsible for accurately locating dark areas that require priority enhancement.
(3) The enhancement module performs brightness improvement on the illumination map.
(4) The refinement module adjusts contrast and optimizes detail representation.
(5) The denoising module applies noise suppression to produce an enhanced result consistent with human visual perception.
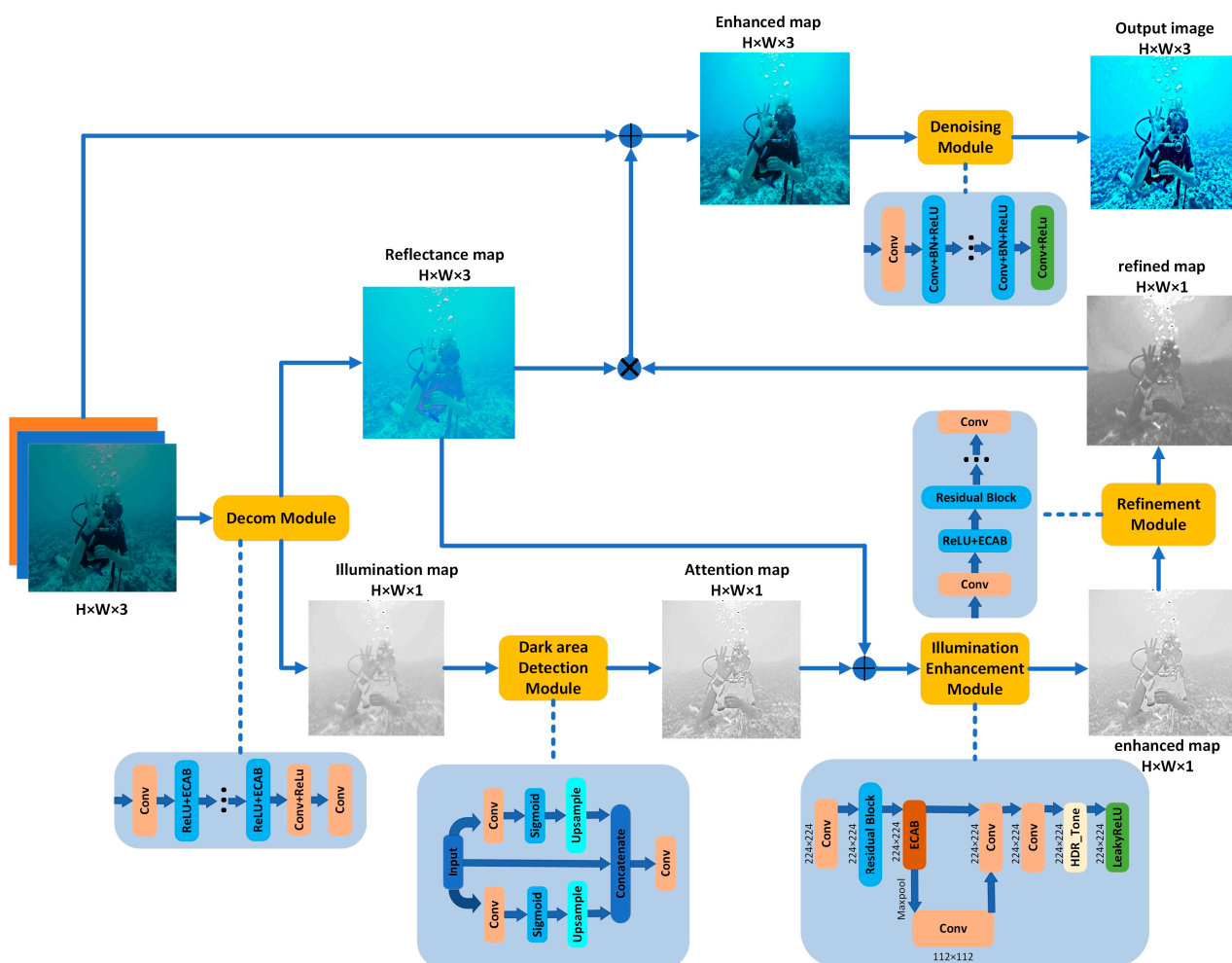
Figure 1. RECAD framework diagram.

3.2. Retinex-Based Decomposition Module

RECAD first applies Retinex theory to the input weakly illuminated underwater image. The decomposition module separates the underwater low-light image S into a reflectance map R and an illumination map I, as shown in Figure 2.
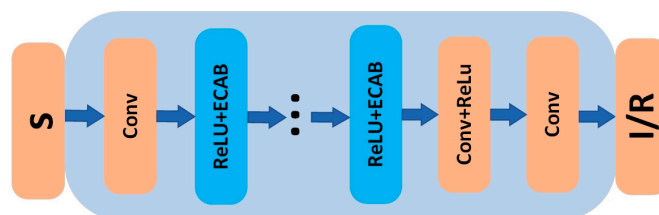

Figure 2. Schematic Diagram of the Decom Module.

The input image is first passed through a preliminary convolutional layer for general feature extraction. This is followed by deep feature processing using two consecutive Efficient Channel Attention Blocks (ECAB) combined with ReLU activation, which represents the key innovation of this module. The ECAB modules capture global contextual information between channels through global average pooling and utilize an attention mechanism to dynamically compute the weight of each channel, thereby implicitly learning and separating the reflectance features (which represent the intrinsic properties and rich details of objects) from the illumination features (which represent ambient lighting and smooth variations). The attention-weighted features are further processed through a Con eLU

combination for feature fusion and non-linear mapping. Finally, a convolutional output layer maps the high-dimensional features to the final decomposition result, simultaneously outputting both the reflectance map and the illumination map.

Unlike traditional methods such as RetinexNet [18], which adopt a fixed channel assignment strategy, our method employs a dual-path heterogeneous output structure after feature extraction: a three-channel convolutional layer and a single-channel convolutional layer are used in parallel to output R and I, respectively, avoiding the limitations of manual channel specification. Moreover, ECAB modules are embedded in the encoding path to enhance the representational capacity of key features through feature recalibration, thereby improving the decomposition quality of the reflectance and illumination maps.

3.3. Multi-Attention-Based Dark Region Detection Module

To improve the effective information extraction capability in low-light regions, inspired by the RSEND method [24], we innovatively introduce an attention-based dark region detection module, the details of which are shown in Figure 3.
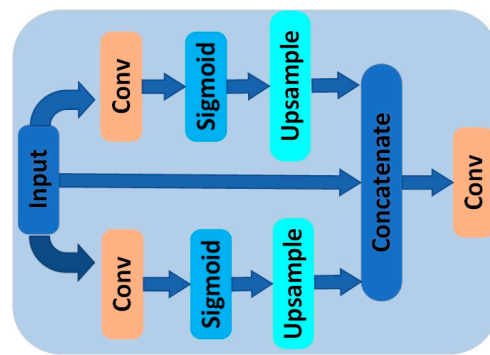

Figure 3. Schematic Diagram of the Dark area Detection Module.

First, the module adopts a multi-scale feature extraction strategy on the illumination map: a 3×3 convolution kernel is applied to the original image, while a 5×5 kernel with a stride of 2 is used for downsampling to capture features at different scales. The extracted features are then passed through a Sigmoid activation function to generate a spatial attention map, which adaptively weights the degree of enhancement required for each region.

The features at different scales are then upsampled to the original resolution and concatenated with the original feature map, forming a feature map with a depth of number_of_channels $\times$ 3 (integrating features from the original and two attention-enhanced paths).

Finally, an additional convolutional layer (typically 1×1 convolution) fuses and refines the concatenated features and outputs a single-channel attention weight map. When this weight map is multiplied element-wise with the features of the main network, it adaptively enhances the response intensity of features in dark regions while suppressing those in non-dark regions. This enables the entire network to adaptively focus on darker areas requiring priority enhancement, significantly improving the targeting and overall effectiveness of underwater image enhancement.

Our method departs from the traditional approach of directly enhancing the illumination map. Instead, it innovatively uses a dark region detection module to achieve region-adaptive feature weighting, emphasizing the priority processing of low-light regions. This effectively enhances the clarity of details in dark areas while maintaining natural overall image brightness.

3.4. U-Net-Based Illumination Enhancement Module

Building on the dark region detection from the previous module, this module enhances the initial illumination map. The specific principle is illustrated in Figure 4.
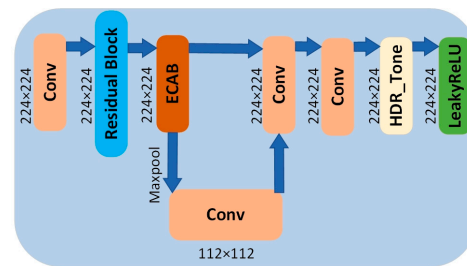

Figure 4. Schematic Diagram of the Illumination Enhancement Module.

The illumination enhancer of this module is based on an improved U-Net architecture [25], designed to deeply optimize and map the initially estimated illumination component. The module centers on an encoder–decoder structure: the encoder path uses Residual Blocks to learn residual features of the illumination and downsamples via Max-Pooling to a low-resolution space, where an Efficient Channel Attention Block (ECAB) is innovatively embedded to adaptively calibrate feature channel weights, strengthening the modeling of global context in illumination distribution. The decoder path gradually restores spatial resolution through upsampling and utilizes skip connections to fuse multi-scale features from the encoder, effectively preserving detail information.

Unlike previous methods, the module introduces a learnable HDR tone mapping unit (HDR_Tone) before output. This unit non-linearly maps the low dynamic range illumination map to a more optimal contrast distribution, significantly recovering details in both bright and dark regions. The HDR_Tone unit employs dual-branch processing: a global branch with learnable scaling and log transformation, and a local branch with two convolutional layers. The combined output is normalized through LayerNorm and Sigmoid to generate the enhanced illumination map. Finally, the enhanced illumination map is output via a LeakyReLU activation function, improving visual quality while ensuring numerical stability. This design enables the module to collaboratively achieve illumination balancing, detail enhancement, and dynamic range expansion, laying a foundation for subsequent image reconstruction.

3.5. Refinement Module

To address potential noise points that may arise after illumination map enhancement, a refinement module is introduced into the model. This module performs detail adjustment, contrast modification, and overall quality improvement on the enhanced illumination map. The specific principle is shown in Figure 5.
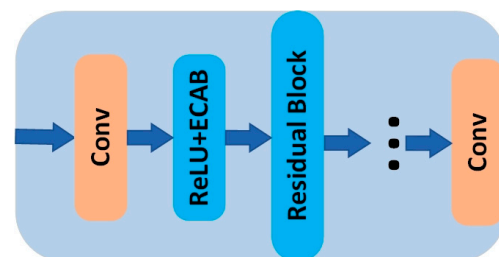

Figure 5. Schematic Diagram of the Refinement Module.

The refinement module has a simple and efficient structure, built around multiple convolutional layers with a kernel size of 3 and padding of 1. Consistent with the earlier decomposition module, this module also incorporates Residual Blocks and ECABlocks to further strengthen its feature extraction and reconstruction capabilities. The residual structure effectively alleviates the vanishing gradient problem and facilitates deep network training, while the ECABlock adaptively calibrates feature weights through a channel attention mechanism, suppressing irrelevant noise and highlighting important details.

Through a combination of multi-level convolution and non-linearity, this module can effectively smooth abnormal noise, enhance edge clarity, and optimize global contrast distribution while maintaining image structural consistency, thereby significantly improving the visual quality and stability of the enhanced image.

3.6. Reconstruction and Denoising Module

After obtaining the enhanced and refined illumination map, it must be multiplied element-wise with the reflectance map to reconstruct the final image, as shown in Figure 1. However, unlike traditional Retinex methods that directly use the product result, and to avoid detail loss and structural distortion, this paper draws on the idea of residual learning [26] and innovatively adds the original low-light image S to the product in the form of a residual. This design not only helps maintain structural consistency between the enhanced image and the original input but also effectively preserves the details of the original scene. Thereby, while improving illumination levels, it ensures the image content remains authentic and natural, leading to a significant optimization of the overall visual effect.

Although Retinex-based enhancement methods perform well in enhancing weakly illuminated underwater images, the enhancement process may introduce or amplify noise. The reason is that when the brightness of dark areas is increased, the inherent noise in these areas also becomes visible. This noise originates not only from the increased visibility of inherent sensor noise in brightened dark areas but also, more critically, from a fundamental model mismatch. The multiplicative Retinex decomposition model does not explicitly account for the additive backscatter component prevalent in underwater environments (as described by physical models like Jaffe-McGlamery). Consequently, during decomposition, high-frequency backscatter patterns risk being misinterpreted by the network as part of the reflectance map. When this 'contaminated' reflectance is subsequently used for reconstruction, these artifacts are preserved and amplified. Therefore, after the reconstruction stage, our method incorporates a dedicated denoising module. This module aims to ensure that the final output image not only has appropriate illumination but also effectively suppresses newly introduced noise, thereby yielding a result with superior visual quality. It is important to emphasize that for images captured in real underwater environments, especially under low-light conditions, the presence of noise is a common phenomenon. Thus, denoising is an indispensable key step in the entire image enhancement pipeline. The specific principle is shown in Figure 6.
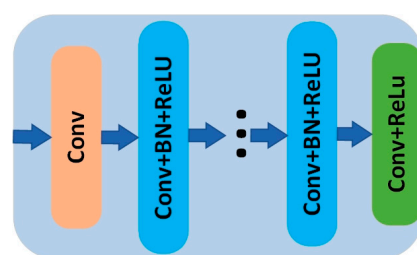

Figure 6. Schematic Diagram of the Denoising Module.

The module is composed of stacked units of multiple convolutional (Conv) layers, batch normalization (BN), and ReLU activation functions: the convolutional layers are responsible for progressively extracting features and integrating contextual information to distinguish between noise and image structure; batch normalization accelerates training convergence and improves generalization; the ReLU function introduces non-linearity and enhances feature sparsity. Together, they achieve efficient noise modeling. Finally, a noise residual is output through a final convolutional layer, and a residual learning mechanism is used to subtract the predicted noise from the input. This effectively suppresses noise while preserving image detail integrity, significantly enhancing the visual quality of the output image.

Compared to previous methods, our approach does not treat denoising as an independent post-processing step or rely on external general-purpose denoisers. Instead, it is embedded as an end-to-end learnable internal module within the enhancement pipeline. This makes the denoising process more tailored to the image enhancement task and is co-optimized specifically for the noise characteristics of underwater low-light images.

3.7. Loss Functions

To effectively improve the quality of underwater image enhancement, this paper designs a composite loss function that combines perceptual loss and underwater image characteristic loss, consisting of the following two parts:

First, the perceptual loss (VGGLoss) based on the VGG19 network is adopted [27]. This loss uses a pre-trained VGG19 model to extract multi-scale features and constrains the predicted and ground-truth images in the feature space to ensure the enhanced image remains consistent with the real image in structure and texture. The VGG loss is defined as:

$$L_{\text{VGG}} = \sum_{i=1}^{5} w_i \cdot \parallel \phi_i(I_{\text{pred}}) - \phi_i(I_{\text{gt}}) \parallel_1 \tag{2}$$

where ϕ_i represents the feature maps from the i-th layer of VGG19, and w_i are the layer weights. It is computed as a weighted sum of the L1 distances between features at different layers, with the weights [1.0/32, 1.0/16, 1.0/8, 1.0/4, 1.0] increasing with network depth. To ensure training stability, feature values are clipped to the range $[-1 \times 10^6, 1 \times 10^6]$ to prevent gradient explosion.

Second, to address the common issues of color distortion and low contrast in underwater images, an underwater-specific loss (UnderwaterLoss) is introduced. The underwater loss is formulated as:

$$L_{\text{UW}} = \lambda_{\text{color}} \cdot L_{\text{color}} + \lambda_{\text{contrast}} \cdot L_{\text{contrast}} \tag{3}$$

where $\lambda_{\text{color}} = \lambda_{\text{contrast}} = 0.5$. This loss is further decomposed into a color loss (ColorLoss) and a contrast loss (ContrastLoss): the color loss computes the difference between the predicted and target images in the Lab color space to better maintain color consistency; the contrast loss extracts image edge information using a Laplacian operator and enhances local contrast through a negative correlation constraint.

Finally, the two losses are adaptively fused via a learnable dynamic weighting module (Dynamic Weighted Loss) to form the overall loss function:

$$L_{total} = \alpha \cdot L_{VGG} + (1 - \alpha) \cdot L_{UW} \tag{4}$$

where α is a weight parameter learned dynamically during training. The weight α is a trainable parameter that is initialized to 0.5 and updated alongside other network parameters. This mechanism enables the model to automatically find the optimal balance between

the perceptual loss L_{VGG} and the underwater image loss L_{UW}. This loss function both preserves the structural authenticity of the image and optimizes the color and contrast characteristics of the underwater image during training, thereby effectively improving the overall quality of the enhanced image.

4. Experiments

To thoroughly validate the efficacy of the proposed RECAD model in enhancing underwater images and to benchmark it against current state-of-the-art methods, this section outlines a detailed experimental plan. The plan covers dataset descriptions, experimental setup, evaluation metrics, presentation of results, subjective and objective assessments, analysis of computational efficiency, and a series of ablation experiments.

4.1. Experimental Details

(1) Datasets

Similarly to most existing studies [28–30], Our evaluation relies on two publicly accessible datasets containing paired underwater images: UIEB [31] and LSUI [32], for our experiments. These datasets encompass a wide range of underwater degradation scenarios. Specifically, the training set was constructed by randomly selecting 700 paired images from UIEB (which provides 890 pairs in total) and 3429 pairs from LSUI (which contains a total of 4279 pairs).

To ensure a comprehensive performance evaluation, a multi-dimensional test set was used in the objective evaluation phase. These test datasets include: the remaining samples from the training sets (190 pairs from UIEB and 850 pairs from LSUI), and three no-reference datasets: Challenging-60 (60 challenging cases from UIEB), the Seathru [33] dataset (an underwater image dataset with high-quality details), and the HURLA [34] dataset (real deep-sea images captured under artificial lighting).

(2) Experimental Setting

The proposed deep learning framework was implemented based on the PyTorch platform (version 2.6.0), adopting a modular component training paradigm. All input images underwent standardized preprocessing and were resized to a resolution of 224×224 pixels. The Adam optimizer was chosen for model optimization, with initial hyperparameters set to $\beta_1 = 0.9$, $\beta_2 = 0.999$, and an initial learning rate of 1.0×10^{-4}. Training was performed on hardware comprising an Intel Core i5-12400 processor (Intel Corporation, Santa Clara, CA, USA), 32 GB of RAM, and an NVIDIA GeForce RTX 4060 GPU (NVIDIA Corporation, Santa Clara, CA, USA), with a batch size of 4 for 200 epochs.

(3) Comparative Methods

In this experiment, the proposed RECAD was compared against a range of techniques spanning multiple methodological categories in underwater image enhancement (UIE). Among traditional methods, we selected IBLA, CLAHE, and UDCP, representing classical approaches to UIE. For deep learning-based methods, we evaluated two advanced models: FUnIEGAN (an end-to-end GAN-based model focused on realistic enhancement and color correction of underwater images) and U-Transformer (a network utilizing the Transformer architecture to capture long-range dependencies for improving global consistency and detail recovery in underwater scenes). Comparative analysis with these representative methods was conducted to comprehensively measure the performance of RECAD.

(4) Evaluation Metrics

This study utilized both full-reference and no-reference image quality assessments. The full-reference evaluation incorporated SSIM [35] and PSNR [36]. The SSIM metric

gauges the perceptual structural similarity between the output and reference images, with a higher score denoting better preservation. PSNR quantifies the ratio of the maximum possible signal power to the corrupting noise power, wherein an increase signifies enhanced signal fidelity and reconstruction quality. The formulas are as follows.

SSIM compares the structural similarity between the reference and enhanced images, considering luminance and contrast:

$$SSIM(F, E) = \frac{(2\mu_F\mu_E + C_1)(2\sigma_{FE} + C_2)}{(\mu_F{}^2 + \mu_E{}^2 + C_1)(\sigma_F{}^2 + \sigma_E{}^2 + C_2)} \tag{5}$$

where μ_F and μ_E are the mean intensities, σ_F and σ_E are the variances, $\mu_F \mu_E$ is the covariance of F and E. C_1 and C_2 are constants to stabilize the division.

MSE quantifies the mean squared error between the enhanced image $E(i, j)$ and the reference image $F(i, j)$:

$$MSE = \frac{1}{MN} \sum_{i=1}^{M} \sum_{j=1}^{N} [F(i, j) - E(i, j)]^2 \tag{6}$$

where M and N are the image dimensions.

PSNR is derived from MSE and measures the peak error:

$$PSNR = 20 log_{10} \left(\frac{MAX_F}{\sqrt{MSE}} \right) \tag{7}$$

where MAX_F is the maximum possible pixel value of the image.

For no-reference evaluation, the Underwater Image Quality Measure (UIQM) [37], designed specifically for assessing underwater image quality, was employed. The development of such dedicated metrics, along with benchmark databases [38], has been crucial for advancing objective evaluation in underwater vision research. UIQM combines multiple quality factors via this weighted formula:

$$UIQM = \alpha \cdot UICM + \beta \cdot UISM + \gamma \cdot UIConM \tag{8}$$

where $UICM$, $UISM$, and $UIConM$ denote the measures for image colorfulness, sharpness, and contrast, respectively, with α, β, and γ being their respective weighting factors. In this experiment, following common practice (e.g., [22]), the weights were set to α = 0.0282, β = 0.2953, γ = 3.5753.

4.2. Experimental Results

Based on the experimental process described above, the results for the proposed method and comparative methods are shown in Figures 7 and 8. Evaluation of these results is presented below.

Samples A–H are specific examples from multiple typical underwater image datasets. Among them, A, B, D, E, and F were selected from the Challenging-60 and LSUI datasets. These images cover various typical underwater visual environments, such as scenes at different depths, varying water turbidity levels, and under natural lighting changes, offering strong representativeness and diversity. C is from the Seathru dataset, presenting a seabed scene under uniform low-light conditions; G and H are sourced from the HURLA dataset, representing real deep-sea images captured under artificial lighting.

Figure 7. Processing Effects of Different Methods in Various Underwater Environments. (**A**) Challenging scene; (**B**) Natural light scene; (**C**) Low-light seabed; (**D**) Reef structures; (**E**) Marine life close-up; (**F**) Greenish color cast; (**G**) Deep-sea artificial light; (**H**) Deep-sea with particles.
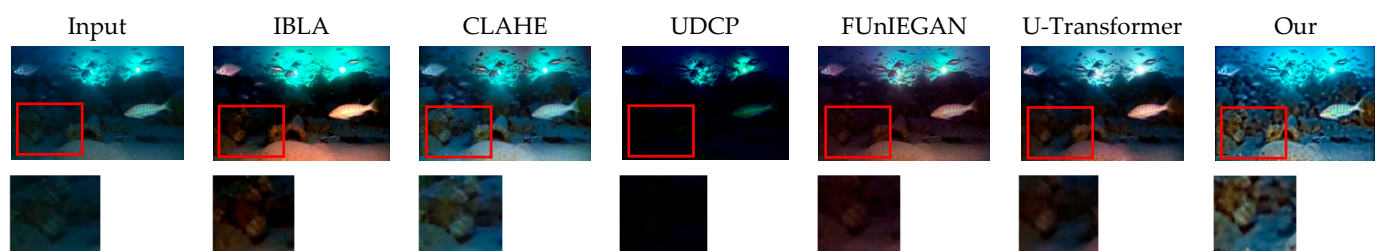

Figure 8. Partial Enlargement Effects.

4.3. Subjective Evaluation

To comprehensively evaluate the performance of underwater image enhancement algorithms, this paper systematically compares the proposed method with other mainstream enhancement and restoration methods in terms of color cast correction, brightness compensation, and sharpness enhancement based on visual comparison experiments.

Regarding color cast removal: As can be seen from Figure 7, compared to other methods, the proposed method effectively suppresses the blue-green cast in the images, offers higher overall visibility, and does not exhibit the reddish tint seen in the U-Transformer results in Figure 7C–E. Simultaneously, our method better restores the true colors of objects. The FUnIEGAN method suffers from severe color cast issues, showing obvious redness accompanied by color artifacts in Figure 7D–F. Clearly, our method outperforms FUnIEGAN in both color restoration and detail preservation. The results of the CLAHE method are closest to ours, but residual green cast can still be observed in Figure 7G; in contrast, our method is visually superior.

Regarding brightness compensation: As shown in Figure 7, our method demonstrates the best brightness adjustment both globally and locally. In Figure 7C, under overall low brightness, compared with FUnIEGAN and UDCP, the IBLA method partially brightens the target object, yet the seabed sand remains dark. The U-Transformer improves the brightness of both the target and surroundings, but the seabed sand is still dim, and details remain obscured. By contrast, our method applies moderate brightness compensation to both the target and its surroundings, presenting clearer details. To further illustrate local brightness compensation, we provide a magnified view of a local area in Figure 7D (Figure 8). Our method performs best in terms of brightness in the reef area, with no artifacts in the seawater and sand background. While FUnIEGAN and U-Transformer still exhibit some color cast in the seawater area, our method not only avoids this defect but also surpasses other methods in presenting sand textures and coral reef details, demonstrating the effectiveness of our method in brightness compensation and the innovative advantage of the introduced dark region detection module.

Regarding sharpness enhancement: As shown in Figure 7, our method also significantly outperforms other comparative methods. From Figure 7E, it can be seen that our method more clearly and prominently restores the black spots on the fish's head and the surrounding coral details. The results in Figure 7H show that our method effectively suppresses image noise while preserving the color and textural details of the seabed sand. Thus, our method demonstrates strong performance in both sharpness improvement and detail preservation, further validating the superiority of the residual learning strategy.

Despite the overall satisfactory performance, we observed that RECAD encounters challenges in scenes with extreme backscatter, as shown in Figure 7H. In such environments with dense suspended particles causing intense forward scattering, the model may occasionally insufficiently remove the haze-like effect, leading to residual fogging in the output. We attribute this limitation to the inherent constraint of the multiplicative Retinex decomposition framework in handling strong additive components like backscatter, which is a known challenge in Retinex-based models. Future work will focus on explicitly modeling the scattering component to address this challenge.

In summary, compared to mainstream methods such as U-Transformer, FUnIEGAN, CLAHE, IBLA, and UDCP, our method performs superiorly in color cast removal, brightness compensation, and sharpness enhancement. It not only effectively corrects color deviations and achieves natural and realistic brightness recovery but also significantly improves the sharpness and structural integrity of image details. The overall visual effect is significantly better than that of comparative algorithms, fully demonstrating the comprehensive performance and practical value of the proposed method.

4.4. Objective Evaluation

Building upon the subjective visual comparison, to further quantitatively assess the comprehensive performance of the proposed method, this section employs cross-dataset full-reference image quality metrics (PSNR, SSIM) and a no-reference metric (UIQM) for

systematic evaluation, thereby providing more convincing data support for the effectiveness and superiority of our method.

(1)   Full-Reference Image Quality Metrics Comparison

This experiment randomly selected 100 underwater images covering various scenes and distortion types from the UIEB and LSUI datasets, processed them using the aforementioned six methods, and conducted quantitative evaluation using the two full-reference image quality metrics, PSNR and SSIM. The final scores are the averages of the results from the 100 images, with specific results shown in Table 1.

Table 1. Quantitative Comparison Results of Different UIE Methods on Full-Reference Metrics.

Methods	UIEB		LSUI	
	PSNR	**SSIM**	**PSNR**	**SSIM**
Input	17.23	0.77	18.05	0.75
LBIA	15.86	0.59	11.44	0.13
CLAHE	16.86	0.76	14.24	0.18
UDCP	9.04	0.22	11.32	0.14
FUnIEGAN	19.64	0.81	23.62	0.87
U-Transformer	22.32	0.88	21.48	0.85
Our	23.41	0.91	22.17	0.84

Note: For each metric, the optimal and suboptimal values are highlighted in red and blue, respectively.

On the UIEB dataset, our method achieved the optimal values in both PSNR and SSIM (PSNR: 23.41 dB, SSIM: 0.91). Compared to traditional non-model methods, our method shows significant advantages: compared to IBLA (PSNR: 15.86 dB, SSIM: 0.59), PSNR increased by 7.55 dB and SSIM by 0.32; compared to CLAHE (PSNR: 16.86 dB, SSIM: 0.76), PSNR increased by 6.55 dB and SSIM by 0.15; compared to UDCP (PSNR: 9.04 dB, SSIM: 0.22), PSNR lead reached 14.37 dB and SSIM improved by 0.69, indicating our method is far superior to traditional strategies in image fidelity and structure recovery. In comparison with advanced deep learning models, our method also performs outstandingly: compared to U-Transformer (PSNR: 22.32 dB, SSIM: 0.88), PSNR increased by 1.09 dB and SSIM by 0.03; compared to FUnIEGAN (PSNR: 19.64 dB, SSIM: 0.81), PSNR lead was 3.77 dB and SSIM improved by 0.10, indicating our method has advantages in distortion suppression while maintaining high perceptual quality.

On the LSUI dataset, FUnIEGAN ranked first in PSNR with 23.62 dB, while our method ranked second with 22.17 dB, still showing good competitiveness. Compared to traditional methods, our method still leads significantly: PSNR increased by 10.73 dB compared to IBLA (PSNR: 11.44 dB), SSIM (0.84) improved by 0.71; increased by 7.93 dB compared to CLAHE (PSNR: 14.24 dB), SSIM improved by 0.66; increased by 10.85 dB compared to UDCP (PSNR: 11.32 dB), SSIM improved by 0.70. In terms of SSIM, our method achieved 0.84, ranking third, only 0.03 lower than the best method FUnIEGAN (0.87) and 0.01 lower than U-Transformer (0.85), indicating comparable performance in image structure preservation to current advanced methods and high stability.

The slightly superior performance of FUnIEGAN on the LSUI dataset in terms of PSNR (23.62 dB vs. 22.17 dB) may be attributed to the dataset's specific characteristics. LSUI contains a substantial number of images with complex color variations and organic matter, where the GAN-based approach of FUnIEGAN might better capture the intricate color distributions through its adversarial training. However, our method maintains advantages in structural preservation (SSIM) and computational efficiency, making it more suitable for resource-constrained deployments.

In summary, the full-reference image quality evaluation results objectively demonstrate that our method not only shows absolute advantages over traditional non-model methods but also exhibits highly competitive performance against existing advanced deep learning models, validating the good generalization ability and stability of our method on datasets with different characteristics, enabling effective handling of various underwater degradation scenarios.

(2) No-Reference Image Quality Metrics Comparison

This experiment further randomly selected 60 underwater images covering various scenes and distortion types from each of the four datasets: Challenging-60, LSUI, Seathru, and HURLA, processed them using the same six methods, and measured them using the UIQM no-reference evaluation metric. The final result is the average of the 60 images, as shown in Table 2.

Table 2. Quantitative Comparison Results of Different UIE Methods on No-Reference Metrics.

Methods	Challenging-60 UIQM	LSUI UIQM	Seathru UIQM	HURLA UIQM
Input	2.01	2.13	2.75	2.30
LBIA	2.04	2.46	2.54	2.35
CLAHE	2.03	2.45	2.97	2.48
UDCP	1.64	1.80	2.11	1.85
FUnIEGAN	2.84	2.95	3.21	3.04
U-Transformer	2.76	3.07	3.14	3.08
Our	2.85	2.99	3.19	3.11

Note: For each metric, the optimal and suboptimal values are highlighted in red and blue, respectively.

Across all four datasets, our method performed stably, consistently ranking in the top two, demonstrating good generalization ability and robustness. Specifically, on the Challenging-60 dataset, our method ranked first with a UIQM value of 2.85, outperforming all comparative methods; on the HURLA dataset, our method also achieved optimal performance with a score of 3.11. On the LSUI and Seathru datasets, our method achieved UIQM values of 2.99 and 3.19, respectively. Although slightly lower than the best methods on these datasets (by 0.08 and 0.02, respectively), it still significantly leads most traditional methods and shows performance comparable to current optimal deep learning models.

Compared to traditional image enhancement methods, the UIQM values of our method on all four datasets are much higher than those of non-model methods such as IBLA, CLAHE, and UDCP. For example, on LSUI, it improved by 1.19 compared to UDCP (1.80), indicating a fundamental advantage in restoring the visual quality of underwater images. When compared with deep learning–based methods, including FUnIEGAN and U-Transformer, our approach remains competitive in most cases, particularly on challenging datasets such as Challenging-60 and HURLA that feature complex degradation patterns.

These results indicate that our method consistently performs excellently under the no-reference evaluation system, can effectively adapt to underwater image data from different sources and with different degradation types, further confirming the effectiveness and robustness of the method in real complex underwater environments.

4.5. Computational Efficiency Analysis

In practical applications, underwater image enhancement algorithms need not only excellent performance but also to meet computational efficiency requirements, especially on platforms with limited computational resources or in real-time processing scenarios such as AUV online target detection. These applications impose strict limits on the algorithm's

real-time processing capability and resource usage. Therefore, a systematic computational efficiency analysis of the method proposed in this study is necessary and of significant practical importance.

This section analyzes three aspects: Floating Point Operations (FLOPs), number of parameters (Params), and inference speed (Frames Per Second, FPS). The pursuit of high efficiency in underwater image enhancement aligns with the broader trend in computer vision of designing lightweight and efficient network architectures, such as IremulbNet [39]. FLOPs measure the computational complexity of the model, with lower values indicating higher computational efficiency [40]. The number of parameters (Params) reflects the model size and memory footprint, with smaller values indicating a lighter model [41]. FPS represents the number of image frames processed per second, with higher values indicating better real-time performance; this metric directly determines whether the algorithm can meet the real-time requirements of AUV online target detection. A comprehensive comparison was made between our method and current mainstream models, including the GAN-based FUnIEGAN, the self-attention-based U-Transformer, the Retinex-GAN fusion-based UIEVUS, and the classic lightweight method WaterNet. The specific performance comparison of each method on the above metrics is presented in Table 3.

Table 3. Computational Efficiency Comparison Results of Various Deep Learning Methods.

Methods	FLOPs (G)	Params (M)	FPS
FUnIEGAN	7.35	7.06	294
U-Transformer	15.82	32.52	28
UIEVUS	35.57	1.82	58
WaterNet	35.72	1.09	78
Our	19.33	0.42	97

Note: For each metric, the optimal and suboptimal values are highlighted in red and blue, respectively.

In terms of the number of parameters, our method requires only 0.42 M parameters, significantly lower than all comparative models: this is about 5.9% of FUnIEGAN (7.06 M), about 1.3% of U-Transformer (32.52 M), and about 61.5% less than WaterNet (1.09 M), demonstrating excellent model lightweight characteristics and significantly reducing storage and memory requirements.

In terms of computational complexity, the FLOPs of our method are 19.33 G, placing it at a medium level among the comparative models. Although higher than FUnIEGAN (7.35 G) and U-Transformer (15.82 G), it is significantly lower than UIEVUS (35.57 G) and WaterNet (35.72 G), a reduction of about 45.7%, showing good computational efficiency, especially suitable for deployment on medium-to-low complexity devices.

Particularly in real-time performance, our method achieves an inference speed of 97 FPS, greatly exceeding the common benchmark for real-time processing (30 FPS). Specifically, it is 3.46 times faster than U-Transformer (28 FPS). Although there is still a gap compared to FUnIEGAN (294 FPS), it is significantly better than UIEVUS (58 FPS) and WaterNet (78 FPS), with improvement rates of 67.2% and 24.4%, respectively. This excellent real-time processing capability can fully meet the high frame rate and low latency requirements for image processing in AUV online target detection, ensuring real-time environmental perception and response for underwater autonomous platforms.

It is noteworthy that the reported FPS was measured with an input size of 224 × 224. While the computational cost increases for higher resolutions, RECAD's fully convolutional architecture allows it to process arbitrary image sizes. For deployment on AUVs, a sliding window strategy can be employed to efficiently handle high-resolution inputs.

Through the above analysis, it can be seen that the proposed method (19.33 G FLOPs/0.42 M Params/97 FPS) achieves an optimal balance in the three dimensions of computational complexity, parameter quantity and inference speed. While maintaining competitive enhancement performance, it has ultra-lightweight parameter configuration and excellent real-time processing capability, showing significant practical application advantages for deployment on resource-constrained platforms.

4.6. Ablation Study

To verify the effectiveness of each component in the proposed framework and systematically evaluate its necessity, ablation studies were conducted on the widely used UIEB underwater image dataset. This dataset contains a large number of real underwater scene images. 700 image pairs were randomly selected as the training set, and 190 pairs were used as the test set. Model training continued for over 100 epochs to ensure sufficient convergence and stable performance.

The ablation study involved sequentially removing four key modules from the framework—ECABlock, Denoising Module, Dark Area Detection Module, and Refinement Module—to investigate the contribution of each module to the overall performance. For each experiment, the rest of the structure and hyperparameters remained consistent. Performance differences were evaluated both quantitatively and qualitatively, with specific quantitative results shown in Table 4 and representative visual comparisons provided in Figure 9.

Table 4. Ablation Study Results of Different Modules.

ECABlock	Dark Area Detection Module	Refinement Module	Denoising Module	PSNR	SSIM
√	√	√	√	23.41	0.91
/	√	√	√	20.18	0.87
√	/	√	√	21.54	0.86
√	√	/	√	22.14	0.89
√	√	√	/	20.87	0.84

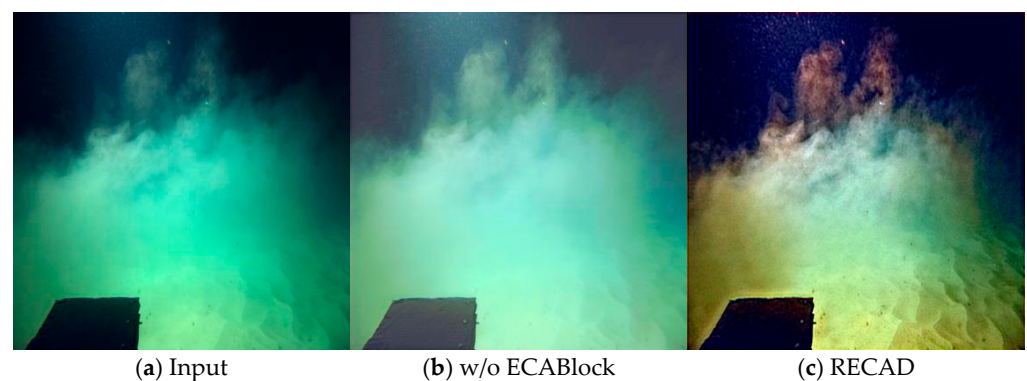(**a**) Input   (**b**) w/o ECABlock   (**c**) RECAD

Figure 9. Visual comparison of the ablation study on the ECABlock.

As shown in Table 4, the removal of the ECABlock, a component embedding the channel attention mechanism, caused a significant drop in PSNR by 3.23 dB and SSIM by 0.04, indicating its core role in global feature calibration and key information extraction. This performance degradation is visually corroborated in Figure 9b,c, where the output without ECABlock exhibits noticeable deficiencies in both color fidelity and detail clarity. Specifically, the image processed without ECABlock appears less vibrant with inferior color restoration, and fine structural details are less distinct compared to the result from the

complete model with ECABlock, which demonstrates more natural color distribution and enhanced textualural details. The absence of the Denoising Module resulted in a PSNR decrease of 2.27 dB and an SSIM decrease of 0.07, so this module contributes significantly to suppressing underwater noise and artifacts, directly affecting the image's signal-to-noise ratio and structural integrity. In comparison, the Dark Area Detection Module, focused on enhancing details in local dark regions, caused a PSNR drop of 1.87 dB and an SSIM drop of 0.05 upon removal, indicating its important auxiliary function in improving visual visibility. The Refinement Module, as a post-processing optimization unit, had the relatively smallest impact; its removal still led to a PSNR decrease of 1.27 dB and an SSIM decrease of 0.02, reflecting its supplementary role in detail recovery and edge optimization.

From the magnitude of performance degradation, it can be inferred that the ECABlock and Denoising Module form the core pillars of the model, ensuring image quality from the key aspects of feature enhancement and noise purification, respectively; the Dark Area Detection Module and Refinement Module undertake more functions of local enhancement and detail refinement, further improving the overall consistency and naturalness of the visual effect. In summary, the ablation study not only quantitatively verifies the effectiveness of each module through metrics but also provides qualitative evidence of their contributions through visual comparisons. The results clearly reveal the hierarchical nature of their contributions, fully demonstrating the indispensability and functional complementarity of each component in the proposed framework.

4.7. Experimental Summary

In summary, through comprehensive experimental validation from multiple dimensions including subjective visual comparison, objective metric evaluation, computational efficiency analysis, and ablation studies, this section comprehensively evaluated the performance of the RECAD method in underwater image enhancement tasks. The experimental results show that the proposed method significantly outperforms existing mainstream methods in color correction, detail enhancement, noise suppression, and brightness recovery, while also demonstrating excellent generalization ability and robustness across multiple datasets. These performance improvements are particularly beneficial for enhancing the target detection and recognition capabilities of AUVs in complex underwater environments. Furthermore, RECAD excels in computational efficiency, requiring a small number of parameters and fast inference speed, making it especially suitable for real-time processing needs in resource-constrained environments. The ablation study further confirms the effectiveness and necessity of each core module, clarifying their functional and synergistic mechanisms within the overall architecture. These results provide evidence for the effectiveness and practicality of the RECAD method.

5. Conclusions

To address the challenge of low-light image degradation faced by Autonomous Underwater Vehicles (AUVs) in deep-sea exploration, this paper proposes RECAD, a lightweight and efficient underwater image enhancement method based on Retinex theory. The method introduces a dark region detection mechanism to enhance feature extraction capability in low-light areas; integrates an efficient channel attention module to improve model representational ability while effectively balancing computational complexity; and employs a residual reconstruction strategy to maintain image geometric structure consistency, thereby enhancing visual fidelity.

Extensive experiments on multiple benchmark datasets show that the proposed method exhibits competitive enhancement performance and maintains good robustness across diverse image types (UIQM metric stably in the range of 2.85–3.19). Furthermore,

RECAD attains high computational efficiency, requiring only 0.42 M parameters while reaching a PSNR of 23.41 dB and a UIQM of 3.19. These results highlight RECAD as an effective enhancement solution for underwater visual perception systems in resource-constrained environments, with promising practical application potential.

The proposed RECAD method demonstrates competitive performance but has certain limitations. In highly turbid waters with intense light scattering, it may occasionally produce residual haze or color inaccuracies. Furthermore, while its dark area detection module improves low-light enhancement, its effectiveness is constrained by the input signal-to-noise ratio, potentially leading to amplified artifacts in extremely low-light conditions. Future work will focus on integrating turbidity-aware mechanisms and advanced noise modeling to better handle these challenging scenarios.

Author Contributions: Q.L. proposed the original idea and wrote the manuscript. H.Q. and X.L. collected materials and wrote the manuscript. T.Z. supervised the manuscript. All authors have read and agreed to the published version of the manuscript.

Data Availability Statement: The raw data supporting the conclusions of this article will be made available by the corresponding author on request.

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

Abbreviations

The following abbreviations are used in this manuscript:

RECAD	Retinex-based Efficient Channel Attention with Dark area Detection
AUV	Autonomous Underwater Vehicle
UIE	Underwater Image Enhanced
ECA	Efficient Channel Attention

References

1. Li, S.; Liu, Y.; Wang, Z.; Zhang, H. Application and prospect of unmanned underwater vehicle. *Bull. Chin. Acad. Sci. (Chin. Version)* **2022**, *37*, 910–920. [CrossRef]
2. Xu, Y.L.; Du, J.H.; Lei, Z.Y. Review: Applications status and key technologies of underwater robots in fishery. *Robot* **2023**, *45*, 110–128. [CrossRef]
3. Feng, X.S.; Cui, W.C.; Zhang, Z.Y.; Li, J.P. Development of deep-sea autonomous underwater vehicle and its applications in resource survey. *Chin. J. Nonferrous Met.* **2021**, *31*, 2746–2756. [CrossRef]
4. Zhao, X.; Tao, J.; Song, Q. Deriving inherent optical properties from background color and underwater image enhancement. *Ocean Eng.* **2015**, *94*, 163–172. [CrossRef]
5. Han, P.; Liu, F.; Yang, K.; Ma, J.; Li, J.; Shao, X. Active underwater descattering and image recovery. *Appl. Opt.* **2017**, *56*, 6631–6638. [CrossRef] [PubMed]
6. Hao, J.; Yang, H.; Hou, X.; Zhang, Y. Two-Stage Underwater Image Restoration Algorithm Based on Physical Model and Causal Intervention. *IEEE Signal Process. Lett.* **2023**, *30*, 120–124. [CrossRef]
7. Hu, K.; Weng, C.; Zhang, Y.; Jin, S. An overview of underwater vision enhancement: From traditional methods to recent deep learning. *J. Mar. Sci. Eng.* **2022**, *10*, 241. [CrossRef]
8. Raveendran, S.; Patil, M.D.; Birajdar, G.K. Underwater image enhancement: A comprehensive review, recent trends, challenges and applications. *Artif. Intell. Rev.* **2021**, *54*, 5413–5467. [CrossRef]
9. Drews, P.; Nascimento, E.; Moraes, F.; Botelho, S.; Campos, M. Transmission estimation in underwater single images. In Proceedings of the IEEE International Conference on Computer Vision Workshops, Sydney, Australia, 1–8 December 2013.
10. Ghani, A.S.A.; Isa, N.A.M. Underwater image quality enhancement through integrated color model with Rayleigh distribution. *Appl. Soft Comput.* **2015**, *27*, 219–230. [CrossRef]
11. Khan, S.A.; Hussain, S.; Yang, S. Contrast enhancement of low-contrast medical images using modified contrast limited adaptive histogram equalization. *J. Med. Imaging Health Inform.* **2020**, *10*, 1795–1803. [CrossRef]

12. Cong, X.; Zhao, Y.; Gui, J.; Liu, X.; Liu, Z. A comprehensive survey on underwater image enhancement based on deep learning. *arXiv* **2024**, arXiv:2405.19684. [CrossRef]

13. Islam, M.J.; Xia, Y.; Sattar, J. Fast underwater image enhancement for improved visual perception. *IEEE Robot. Autom. Lett.* **2020**, *5*, 3227–3234. [CrossRef]

14. Zhang, Y.; Chandler, D.M.; Leszczuk, M. Retinex-based underwater image enhancement via adaptive color correction and hierarchical U-shape transformer. *Opt. Express* **2024**, *32*, 24018–24040. [CrossRef] [PubMed]

15. Wu, X.; Li, H. A simple and comprehensive model for underwater image restoration. In Proceedings of the 2013 IEEE International Conference on Information and Automation (ICIA), Yingchuan, China, 26–28 August 2013.

16. Syariz, M.A.; Jaelani, L.M.; Subehi, L.; Pamungkas, A.; Koenhardono, E.S.; Sulisetyono, A. WaterNet: A convolutional neural network for chlorophyll-a concentration retrieval. *Remote Sens.* **2020**, *12*, 1966. [CrossRef]

17. Wang, N.; Zhou, Y.; Han, F.; Zhu, H.; Zheng, Y. UWGAN: Underwater GAN for real-world underwater color restoration and dehazing. *arXiv* **2019**, arXiv:1912.10269. [CrossRef]

18. Wei, C.; Wang, W.; Yang, W.; Liu, J. Deep retinex decomposition for low-light enhancement. *arXiv* **2018**, arXiv:1808.04560. [CrossRef]

19. Wang, J.; Wang, H.; Sun, Y.; Yang, J. Improved retinex-theory-based low-light image enhancement algorithm. *Appl. Sci.* **2023**, *13*, 8148.

20. Wang, H.; Zhang, W.; Bai, L.; Ren, P. Metalantis: A comprehensive underwater image enhancement framework. *IEEE Trans. Geosci. Remote Sens.* **2024**, *62*, 1–19.

21. Land, E.H.; McCann, J.J. Lightness and retinex theory. *J. Opt. Soc. Am.* **1971**, *61*, 1–11. [CrossRef] [PubMed]

22. Ren, S.; Zhou, W.; Qi, D.; Li, C. UIEVUS: An underwater image enhancement method for various underwater scenes. *Signal Process. Image Commun.* **2025**, *135*, 117264. [CrossRef]

23. Wang, Q.; Wu, B.; Zhu, P.; Li, P.; Zuo, W.; Hu, Q. ECA-Net: Efficient channel attention for deep convolutional neural networks. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 14–19 June 2020.

24. Li, J.; Zhang, K.; Zhang, Y.; Fan, J.; Liu, Y. RSEND: Retinex-based Squeeze and Excitation Network with Dark Region Detection for Efficient Low Light Image Enhancement. *arXiv* **2024**, arXiv:2406.09656. [CrossRef]

25. Du, G.; Cao, X.; Liang, J.; Chen, X.; Zhan, Y. Medical image segmentation based on U-net: A review. *J. Imaging Sci. Technol.* **2020**, *64*, 020508. [CrossRef]

26. Shafiq, M.; Gu, Z. Deep residual learning for image recognition: A survey. *Appl. Sci.* **2022**, *12*, 8972. [CrossRef]

27. 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, 11–14 October 2016; Volume 9906, pp. 694–711.

28. Wang, Y.; Zhang, J.; Cao, Y.; Wang, Z. UIEC^2-Net: CNN-based underwater image enhancement using two color space. *Signal Process. Image Commun.* **2021**, *96*, 116250. [CrossRef]

29. Li, C.; Guo, J.; Cong, R.; Pang, Y.; Wang, B. Underwater image enhancement via medium transmission-guided multi-color space embedding. *IEEE Trans. Image Process.* **2021**, *30*, 4985–5000. [CrossRef] [PubMed]

30. Chen, Y.; Wang, L.; Li, H.; Zhang, P. Underwater image enhancement based on improved water-net. In Proceedings of the 2022 IEEE International Conference on Cyborg and Bionic Systems (CBS), Wuhan, China, 24–26 March 2023.

31. Li, C.; Guo, C.; Ren, W.; Cong, R.; Hou, J.; Kwong, S.; Tao, D. An underwater image enhancement benchmark dataset and beyond. *IEEE Trans. Image Process.* **2019**, *29*, 4376–4389. [CrossRef]

32. Peng, L.; Zhu, C.; Bian, L. U-shape transformer for underwater image enhancement. *IEEE Trans. Image Process.* **2023**, *32*, 3066–3079. [CrossRef] [PubMed]

33. Akkaynak, D.; Treibitz, T. Sea-thru: A method for removing water from underwater images. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 15–20 June 2019.

34. Guo, W.; Zhu, Y.; Li, X.; Zhang, Q. Rapid deep-sea image restoration algorithm applied to unmanned underwater vehicles. *Acta Opt. Sin.* **2022**, *42*, 0410002. [CrossRef]

35. Chen, T.; Wang, R.; Liu, J. Underwater image quality assessment method based on color space multi-feature fusion. *Sci. Rep.* **2023**, *13*, 16838. [CrossRef]

36. Wang, Z.; Bovik, A.C.; Sheikh, H.R.; Simoncelli, E.P. Image quality assessment: From error visibility to structural similarity. *IEEE Trans. Image Process.* **2004**, *13*, 600–612. [CrossRef]

37. Panetta, K.; Gao, C.; Agaian, S. Human-visual-system-inspired underwater image quality measures. *IEEE J. Ocean. Eng.* **2015**, *41*, 541–551. [CrossRef]

38. Liu, Y.; Zhang, B.; Hu, R.; Gu, K.; Zhai, G.; Dong, J. Underwater image quality assessment: Benchmark database and objective method. *IEEE Trans. Multimed.* **2024**, *26*, 7734–7747. [CrossRef]

39. Su, T.; Liu, A.; Shi, Y.; Zhang, X. IremulbNet: Rethinking the inverted residual architecture for image recognition. *Neural Netw.* **2024**, *172*, 106140. [CrossRef] [PubMed]

40. Chen, J.; Kao, S.; He, S.; Zhuo, W.; Wen, S.; Lee, C.; Chan, S. Run, don't walk: Chasing higher FLOPS for faster neural networks. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Vancouver, BC, Canada, 17–24 June 2023.

41. Molchanov, P.; Tyree, S.; Karras, T.; Aila, T.; Lehtinen, J. Pruning convolutional neural networks for resource efficient inference. *arXiv* **2016**, arXiv:1611.06440. [CrossRef]