

# 小波域条件扩散与跨频率协同建模的水下图像增强方法

杨培晗

华北理工大学电气工程学院, 河北 唐山

收稿日期: 2026年6月2日; 录用日期: 2026年6月24日; 发布日期: 2026年7月7日

## 摘要

水下图像增强在海洋探测与水下视觉感知中具有重要应用价值, 但水体的光吸收与散射效应会导致严重的颜色失真与细节退化。现有方法多依赖空间域建模, 缺乏对结构与纹理的解耦建模; 扩散模型虽具备强生成能力, 但计算开销大且难以兼顾多频信息恢复。为此, 本文提出一种融合小波域条件扩散与跨频率协同建模的水下图像增强方法, 实现结构与细节的协同恢复。具体而言, 通过两级小波分解将图像解耦为低频结构与高频细节, 并在低频域引入条件扩散模型以恢复全局结构; 设计方向感知高频增强模块(HFEM), 通过方向建模与子带交互强化纹理表达; 提出跨频率校正模块(CFC), 实现高低频信息的双向融合与一致性约束, 从而提升整体重建质量。在UIEB数据集上的实验结果表明, 所提方法的PSNR达23.75 dB, 且在SSIM、LPIPS等指标上均优于现有方法; 同时在真实场景中表现出更自然的颜色恢复与更清晰的细节重建, 展现出良好的泛化能力。

## 关键词

水下图像增强, 扩散模型, 小波变换, 跨频率建模

# Wavelet-Domain Conditional Diffusion with Cross-Frequency Collaborative Modeling for Underwater Image Enhancement

Peihan Yang

School of Electrical Engineering, North China University of Science and Technology, Tangshan Hebei

Received: June 2, 2026; accepted: June 24, 2026; published: July 7, 2026

## Abstract

Underwater image enhancement is of significant importance for marine exploration and underwater visual perception. However, complex light absorption and scattering effects in water often lead to severe color distortion and detail degradation. Existing methods predominantly rely on spatial-domain modeling and lack explicit decoupling of structural and textural information. Although diffusion models have demonstrated strong generative capability, they typically incur high computational cost and struggle to jointly recover multi-frequency information. To address these challenges, we propose a wavelet-domain conditional diffusion framework with cross-frequency collaborative modeling for underwater image enhancement, enabling effective joint restoration of structure and fine details. Specifically, a two-level wavelet decomposition is first employed to separate the input image into low-frequency structural components and high-frequency detail components. A conditional diffusion model is then introduced in the low-frequency domain to recover global structures. Meanwhile, a direction-aware high-frequency enhancement module (HFEM) is designed to refine texture details via directional modeling and sub-band interactions. Furthermore, a cross-frequency correction module (CFC) is proposed to facilitate bidirectional fusion and consistency between low- and high-frequency representations, thereby improving overall reconstruction quality. Experimental results on the UIEB dataset demonstrate that the proposed method achieves a PSNR of 23.75 dB and consistently outperforms existing approaches in terms of SSIM and LPIPS. In addition, it produces more natural color restoration and clearer details in real-world scenarios, indicating strong generalization capability.

## Keywords

Underwater Image Enhancement, Diffusion Models, Wavelet Transform, Cross-Frequency Modeling

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

海洋环境在资源开发、生态监测以及水下工程等领域具有重要价值，而水下视觉技术则是实现海洋信息获取的重要手段。然而，由于水体对光传播具有显著的吸收与散射效应，水下成像往往表现出对比度降低、颜色失真以及细节模糊等问题。其中，不同波长光在水中的衰减速率差异会导致图像产生明显的色偏，而悬浮颗粒产生的前向散射与后向散射则进一步降低了图像的清晰度。这些退化因素严重影响了水下目标检测、环境监测以及自主水下航行器视觉感知等任务的可靠性，因此水下图像增强技术逐渐成为水下视觉研究的重要方向。

近年来，研究者提出了大量水下图像增强方法以改善图像视觉质量。这些方法在一定程度上缓解了水下图像退化问题，但仍然存在一些挑战。一方面，水下图像退化具有复杂的物理机制，传统方法难以准确建模不同环境条件下的退化过程。另一方面，虽然深度学习方法能够通过数据驱动学习复杂映射关系，但许多方法仍然主要在空间域进行特征建模，缺乏对图像结构信息与细节信息的显式分离与协同建模，从而限制了增强效果。

扩散模型[1]凭借其逐步去噪的生成机制，在图像恢复任务中展现出强大的建模能力。然而，直接将

扩散模型应用于水下图像增强仍面临困难：一方面，扩散模型通常需要在高维图像空间中进行迭代采样，计算开销较大；另一方面，水下图像中结构信息与纹理细节往往受到不同程度的退化影响，单一空间域建模难以同时恢复结构与细节。

针对上述问题，本文提出一种融合小波域条件扩散与方向感知高频增强的水下图像增强方法。该方法通过频域分解与生成模型的有机结合，实现水下图像中结构信息与细节信息的协同恢复。实验结果表明，本文方法在客观指标和主观视觉上均优于现有方法。本文的主要贡献如下：

1) 小波域条件扩散模型：在低频空间引入条件扩散模型，以退化图像的低频分量为条件引导恢复过程，在降低计算复杂度的同时提升结构恢复能力。

2) 方向感知高频增强模块：通过方向投影与自适应加权，针对性恢复水下图像纹理细节。

3) 跨频率校正模块：实现低频结构与高频细节的双向融合，确保重建结果的一致性与丰富性。

## 2. 相关工作

水下图像增强(Underwater Image Enhancement, UIE)旨在恢复由于水体吸收与散射造成的图像退化问题，例如颜色失真、对比度下降以及细节模糊等现象。现有方法大致可分为三类：基于物理模型的方法、基于图像处理的方法以及基于深度学习的方法。

基于物理模型的方法从水下光学成像机理出发，通过估计光传播参数实现图像复原。例如，Drews 等人[2]提出水下暗通道先验(UDCP)，利用水下图像中的暗通道统计特性估计透射率，从而改善图像对比度并恢复退化图像。随后，Peng 等[3]进一步结合图像模糊度与光吸收特性估计场景深度，以提升图像恢复效果。此外，Akkaynak 等人[4]提出 Sea-thru 模型，通过引入波长相关的光衰减模型对水下光传播进行更精确的建模，从而实现更加可靠的颜色恢复。然而，这类方法通常依赖较强的先验假设，并需要准确估计水体参数，在复杂水下环境中容易产生误差累积。

基于非物理模型的方法通常不依赖成像机理，而是通过图像处理技术直接增强图像视觉质量。例如，直方图均衡化[5]和 Retinex 理论[6]被广泛用于提升图像对比度与亮度。此外，多尺度融合方法[7]通过融合不同增强结果来改善图像视觉效果。然而，由于缺乏对水下成像机理的建模，这类方法往往难以同时恢复图像颜色与细节信息。

近年来，基于深度学习的方法逐渐成为水下图像增强研究的主流方向。这类方法通过数据驱动的方式学习退化图像与高质量图像之间的映射关系，在复杂场景下表现出更好的鲁棒性。早期研究多基于卷积神经网络构建端到端增强模型，其中 Anwar 等[8]提出的 UWCNN 颇具代表性，该方法结合残差学习与密集连接的轻量级 CNN 架构，针对多种水质条件下的退化特性进行建模，并通过联合优化 MSE 与 SSIM 损失函数，实现了水下图像的色偏校正与对比度提升。在此基础上，一些研究进一步将颜色空间信息或物理先验嵌入网络结构，以增强模型对水下退化特性的刻画能力。例如，Wang 等人[9]提出 UIEC<sup>2</sup>-Net，在统一 CNN 框架中联合引入 RGB 与 HSV 颜色空间，构建包含 RGB 像素级去噪与色偏去除、HSV 全局调节的双核心分支，其中 HSV 分支通过学习型分段线性曲线对亮度与饱和度进行全局优化，并通过注意力模块对两分支输出进行跨颜色空间的像素级自适应融合，从而提升颜色恢复能力；Li 等人[10]提出 Ucolor，引入 RGB、HSV 与 Lab 多颜色空间构建多分支编码器进行联合特征学习，利用不同颜色空间的表达互补性增强对复杂色偏的建模能力，并通过介质透射率引导的解码模块融合其相关的物理退化信息，以引导网络关注退化区域，实现更有效的图像增强。基于 CNN 的方法在增强效果和稳定性方面取得了显著进展，但其特征提取主要依赖局部卷积操作，对复杂水下退化条件下的长距离空间依赖建模能力仍然有限。为了进一步提升增强图像的视觉真实性，部分研究引入生成对抗网络框架，突破传统 CNN 仅依靠像素或结构损失优化的局限，通过对抗训练让模型学习真实的图像视觉特征，有效改善增强图像模糊、

纹理失真的问题。Fabbri 等[11]提出的 UGAN 是其中的典型代表,该方法依托 WGAN-GP 解决传统 GAN 的训练不稳定性问题,先通过 CycleGAN 生成无配对样本下的训练数据,再结合 U-Net 生成器与 PatchGAN 判别器完成对抗训练,还可引入梯度差异损失提升图像细节锐度,其增强结果在色彩还原与视觉质感上更贴近真实场景,同时能有效提升水下跟踪等下游视觉任务的性能。GAN 类方法在视觉效果方面具有一定优势,但由于对抗训练的不稳定性,其结果在颜色一致性和细节恢复方面仍可能存在波动。为增强网络对全局信息的建模能力,近年来一些研究开始引入注意力机制或 Transformer 结构,借助自注意力的长距离依赖捕捉特性,适配水下图像非均匀退化的特点。例如, Peng 等[12]提出 U-shape Transformer,将 Transformer 与 U 型编解码结构相结合,并引入注意力机制以增强退化区域的特征表达能力,从而提升图像的色彩恢复效果。然而,该方法主要在空间域进行建模,未显式对结构信息与细节信息进行解耦,在复杂退化场景下可能难以兼顾二者的协同恢复性能。进一步地, Chen 等[13]提出 UWFormer,通过引入小波变换对图像进行高低频分解,实现结构信息与纹理细节的解耦建模,并针对低频分量设计多尺度 Transformer 结构进行全局建模,同时结合频率感知注意力机制与半监督学习策略,有效缓解了水下配对数据稀缺问题,在多个主流基准上取得了优异性能。然而,该方法在高低频分支融合过程中缺乏显式的跨频率交互校正机制,在复杂退化场景下可能存在细节一致性不足的问题;此外,其低频分支所采用的 Transformer 架构未针对轻量化进行优化,计算开销相对较大,限制了其在水下端侧或实时应用中的部署。总体而言,基于 Transformer 的方法有效突破了传统 CNN 在局部建模方面的局限,但在多尺度协同建模与计算效率之间的平衡方面仍存在进一步提升空间。

针对 Transformer 方法在复杂退化建模方面的局限,扩散模型作为一种基于逐步噪声建模的生成方法,近年来被引入水下图像增强领域。该方法通过构建前向加噪与反向去噪过程,将图像恢复问题转化为从噪声分布逐步逼近目标分布的过程,从而能够更细致地刻画水下图像中的非线性退化特征,在结构恢复与细节重建方面表现出更强的稳定性与表达能力。Guan 等人[14]提出 DiffWater,将颜色补偿后的水下图像作为条件,引导去噪扩散概率模型(DDPM)进行图像增强。该方法在 RGB 颜色空间进行通道式颜色补偿,根据不同的水体条件和光照场景自适应调整补偿策略,以改善增强结果的色彩外观。Tang 等人[15]提出基于 Transformer 的扩散模型(TBDM),通过设计轻量化去噪网络并结合非均匀跳跃采样策略,在保证增强效果的同时提升了推理效率。Du 等人[16]提出 UIEDP,利用预训练的扩散模型所蕴含的自然图像先验,并结合任意现有增强方法生成伪标签以指导条件采样,从而缓解了配对数据质量不足对模型性能的影响。Shi 等人[17]提出 CPDM,通过构建多条件输入机制,将原始退化图像与噪声图像之间的差值作为辅助条件,同时引入内容补偿模块以融合低层结构信息,从而提升生成图像的结构保真度。现有模型虽在颜色恢复和细节重建方面取得了一定进展,但由于扩散模型通常依赖多步迭代过程,其计算开销仍然较高,不同方法在复杂水下退化场景下的表现也存在一定差异。

### 3. 本文方法

#### 3.1. 算法整体框架

为同时恢复水下图像的全局结构信息与细节纹理信息,本文提出一种基于频域分解与条件扩散模型的水下图像增强框架,其整体结构如图 1 所示。

首先,对输入水下图像与真实清晰图像分别进行两级离散小波变换(2-DWT),将其分解为低频子带与多个高频子带。低频子带主要包含图像的整体结构与颜色信息,因此在低频空间中引入条件扩散模型,对真实清晰图像的低频子带执行前向噪声扩散,逐步添加高斯噪声得到带噪低频特征;再通过逆向去噪过程,以输入退化图像的低频子带为条件约束,学习从带噪低频特征中恢复清晰结构的能力。高频子带主要描述图像的边缘与纹理细节,为此本文设计高频增强模块(HFEM)对高频特征进行方向感知增强与多

尺度细节恢复。为进一步缓解低频恢复与高频增强之间的结构不一致问题，本文提出跨频率校正模块(CFC)，实现低频结构信息与高频细节信息之间的特征交互。最后，通过逆小波变换(IDWT)对处理后的低频与高频子带进行融合重建，从而得到增强后的水下图像。

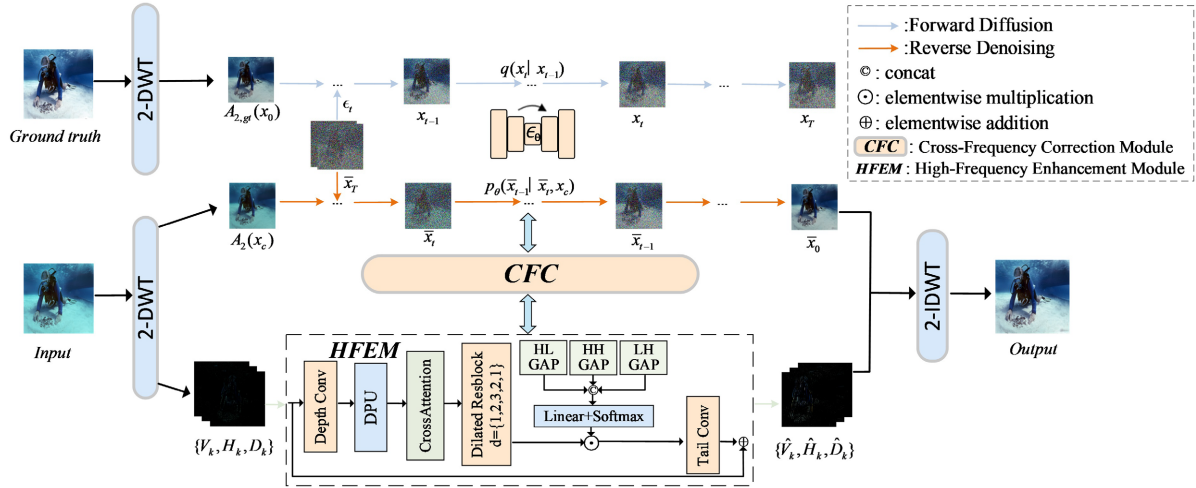


Figure 1. Overall flowchart  
图 1. 整体流程图

### 3.2. 小波域频率分解

通过离散小波变换(Discrete Wavelet Transform, DWT)可以将图像从空间域映射到小波域，从而实现不同频率信息的分解。在实现中，考虑到分解层数与计算复杂度之间的权衡，本文采用两级小波分解作为默认设置。

设输入退化水下图像为  $I \in \mathbb{R}^{H \times W \times c}$ ，其二维小波分解表示为：

$$\{A_1, V_0, H_0, D_0\} = \text{DWT}(I) \quad (1)$$

其中， $A_1 \in \mathbb{R}^{\frac{H}{2} \times \frac{W}{2} \times c}$  表示第一层低频子带，主要包含图像的整体结构与颜色信息； $V_0$ 、 $H_0$  和  $D_0$  分别表示垂直、水平以及对角方向的高频子带，用于描述图像中的边缘与纹理细节。

对低频子带  $A_1$  再次进行小波分解，得到第二层低频子带  $A_2$  与高频子带  $V_1, H_1, D_1$ ：

$$\{A_2, V_1, H_1, D_1\} = \text{DWT}(A_1) \quad (2)$$

为简化表示，定义第  $k$  层高频集合为  $F_k = \{V_k, H_k, D_k\}, k \in \{0, 1\}$ 。通过两级小波分解，图像在保留主要信息的同时实现空间维度压缩，从而降低后续处理的计算复杂度。

### 3.3. 低频条件扩散模型

为实现水下图像低频结构信息的有效恢复，本文在小波分解域中构建条件扩散模型，对二级小波分解后的低频子带进行建模。通过引入退化图像的低频结构作为条件约束，引导扩散过程在结构一致的前提下完成重建，从而提升恢复稳定性与精度。

给定目标图像经小波分解得到的低频子带  $x_0 = A_{2,gr}$ ，通过逐步向其添加高斯噪声构建前向扩散过程。具体而言，对于时间步  $t \in \{1, 2, \dots, T\}$ ，扩散过程由方差调度  $\beta_t$  控制：

$$q(x_t | x_{t-1}) = \mathcal{N}(x_t; \sqrt{1 - \beta_t} x_{t-1}, \beta_t \mathbf{I}) \quad (3)$$

通过重参数化技巧，任意时间步的噪声图像可表示为：

$$x_t = \sqrt{\bar{\alpha}_t} x_0 + \sqrt{1 - \bar{\alpha}_t} \epsilon, \epsilon \sim \mathcal{N}(0, I) \quad (4)$$

其中  $\bar{\alpha}_t = \prod_{s=1}^t (1 - \beta_s)$ 。在训练过程中，通过随机采样时间步  $t$ ，将真实低频子带逐步扰动至噪声空间，从而构建从数据分布到高斯分布的映射关系，为后续反向去噪提供学习目标。

反向去噪过程旨在从纯噪声分布逐步恢复出目标图像，其核心是学习条件概率分布  $p_\theta(\bar{x}_{t-1} | \bar{x}_t, x_c)$ ，其中  $x_c$  为退化图像。根据贝叶斯公式，该分布可近似为高斯形式：

$$p_\theta(\bar{x}_{t-1} | \bar{x}_t, x_c) = \mathcal{N}(\bar{x}_{t-1}; \mu_\theta(\bar{x}_t, x_c, t), \sigma_t^2 \mathbf{I}) \quad (5)$$

其中均值  $\mu_\theta$  由神经网络预测，方差  $\sigma_t^2$  通常固定为与前向过程相关的常数。

为实现条件引导的去噪过程，本文采用噪声预测网络  $\epsilon_\theta$  对目标进行建模。根据扩散模型的参数化策略，均值可表示为：

$$\mu_\theta(\bar{x}_t, x_c, t) = \frac{1}{\alpha_t} \left( \bar{x}_t - \frac{\beta_t}{1 - \bar{\alpha}_t} \epsilon_\theta(\bar{x}_t, x_c, t) \right) \quad (6)$$

其中  $\alpha_t = 1 - \beta_t$ 。网络输入为当前噪声图像与条件图像的拼接形式  $[x_c, \bar{x}_t]$ ，输出为预测的噪声分量。

在推理阶段，本文采用 DDIM 采样策略对反向过程进行加速，在保证重建质量的前提下显著减少采样步数，提高推理效率。

### 3.4. 方向感知高频增强模块(HFEM)

高频子带承载图像的边缘与纹理信息，但在水下成像中易受散射和吸收影响而衰减。为此，本文设计高频增强模块，对多尺度高频分量进行针对性恢复。

设输入特征为  $X \in \mathbb{R}^{H \times W \times c}$ ，首先通过深度可分离卷积进行通道对齐，随后对三个子带分别进行方向感知增强、子带交互与自适应加权。

#### 3.4.1. 方向感知投影(DPU)

水下图像中的纹理通常具有明显的方向性。为了增强不同方向上的纹理响应，本文设计方向感知投影单元(Direction-Aware Projection Unit, DPU)，分别沿高度和宽度方向进行特征建模：

$$g_h(j) = \frac{1}{H} \sum_{i=1}^H X_{i,j}, \quad g_w(i) = \frac{1}{W} \sum_{j=1}^W X_{i,j} \quad (7)$$

其中， $g_h$  和  $g_w$  分别表示沿高度方向和宽度方向压缩后的统计描述。之后，分别通过两个  $1 \times 1$  卷积映射和非线性激活生成方向注意权重：

$$a_h = \sigma(f_h(g_h)), \quad a_w = \sigma(f_w(g_w)) \quad (8)$$

其中  $f_h$  和  $f_w$  表示  $1 \times 1$  卷积映射函数， $\sigma$  为 Sigmoid 激活函数。

通过广播操作将两个方向权重扩展至空间尺寸，并计算方向门控掩码：

$$M_{i,j} = a_h(j) \cdot a_w(i) \quad (9)$$

最终得到增强后的特征表示：

$$X' = X \odot (1 + M) \quad (10)$$

其中  $\odot$  表示逐元素乘法。该残差门控机制能够在保持原始特征信息的同时，自适应增强具有显著方向性

的纹理结构。将 DPU 分别应用于  $V_k, H_k, D_k$ ，得到初步增强后的高频特征。

### 3.4.2. 子带交互与自适应融合

三个高频子带分别描述水平、垂直和对角方向的纹理信息，它们之间既存在互补性，又因水下场景差异而重要性不同。为此，我们设计一个联合的子带交互与自适应加权模块，对 DPU 处理后的特征(简记为  $V, H, D$ )进行联合建模。该模块包含三个连续步骤：

子带间交叉注意力：以对角子带  $D$  作为信息桥梁，增强水平子带  $H$  和垂直子带  $V$  的表示：

$$F_{V \leftarrow D} = \text{CrossAttention}(V, D), F_{H \leftarrow D} = \text{CrossAttention}(H, D) \quad (11)$$

其中  $\text{CrossAttention}$  以第一个输入为查询，第二个为键值，实现跨方向的信息流动。

多尺度空洞卷积增强：交互后的特征与原子带拼接，送入空洞卷积残差块，在扩大感受野的同时保持分辨率：

$$\tilde{V} = f_{\text{dilated}}(\text{Concat}(V, F_{V \leftarrow D})), \tilde{H} = f_{\text{dilated}}(\text{Concat}(H, F_{H \leftarrow D})) \quad (12)$$

其中  $f_{\text{dilated}}$  为空洞卷积残差块，采用空洞率逐步增大再减小的设计(1→2→3→2→1)，在扩大感受野的同时避免网格效应。

值得注意的是，增强后的水平子带  $\tilde{H}$  和垂直子带  $\tilde{V}$  通过融合操作进一步增强对角子带：

$$\tilde{D}_{\text{fused}} = f_{\text{conv}}(\text{Concat}(F_{V \leftarrow D}, F_{H \leftarrow D})) \quad (13)$$

$$\tilde{D}_{\text{final}} = f_{\text{dilated}}(\tilde{D}_{\text{fused}}) \quad (14)$$

由此形成从对角子带流向水平、垂直子带，再回流至对角子带的双向信息交互，使三个子带能够相互补充，提升细节恢复的一致性。

自适应子带加权：考虑到不同场景下三个子带的重要性差异，我们引入可学习加权机制：

$$[\omega_1, \omega_2, \omega_3] = \text{Softmax}\left(\left[\text{GAP}(\tilde{V}), \text{GAP}(\tilde{H}), \text{GAP}(\tilde{D}_{\text{final}})\right]^T\right) \quad (15)$$

其中  $\text{GAP}$  为全局平均池化。最终输出为加权后的子带特征：

$$\hat{V} = \omega_1 \cdot \tilde{V}, \hat{H} = \omega_2 \cdot \tilde{H}, \hat{D} = \omega_3 \cdot \tilde{D}_{\text{final}} \quad (16)$$

将上述处理分别应用于  $k=0$  和  $k=1$  两层高频子带，即可得到增强后的多尺度高频特征  $\hat{V}_0, \hat{H}_0, \hat{D}_0$  和  $\hat{V}_1, \hat{H}_1, \hat{D}_1$ 。该设计使网络能够根据图像内容动态调整各子带的贡献度，提升对不同水下场景的适应能力。

### 3.5. 跨频率校正模块(CFC)

低频结构和高频细节若独立恢复，易导致重建结果不一致。为此，我们设计跨频率校正模块，在两个重建阶段分别实现高低频信息融合，具体如图 2 所示。

给定低频特征  $A$  和高频特征  $F = \{V, H, D\}$ ，首先通过深度可分离卷积将低频和高频特征映射到高维特征空间，得到  $A_f, F_f$ ，以增强特征表达能力

随后将三个高频子带平均，得到与低频维度对齐的表示：

$$H_{\text{avg}} = \frac{V_f + H_f + D_f}{3} \in \mathbb{R}^{B \times C \times H \times W} \quad (17)$$

以低频为查询、平均高频为键值，计算交叉注意力：

$$Q = W_q(A_f), K = W_k(H_{\text{avg}}), V_p = W_v(H_{\text{avg}}) \quad (18)$$

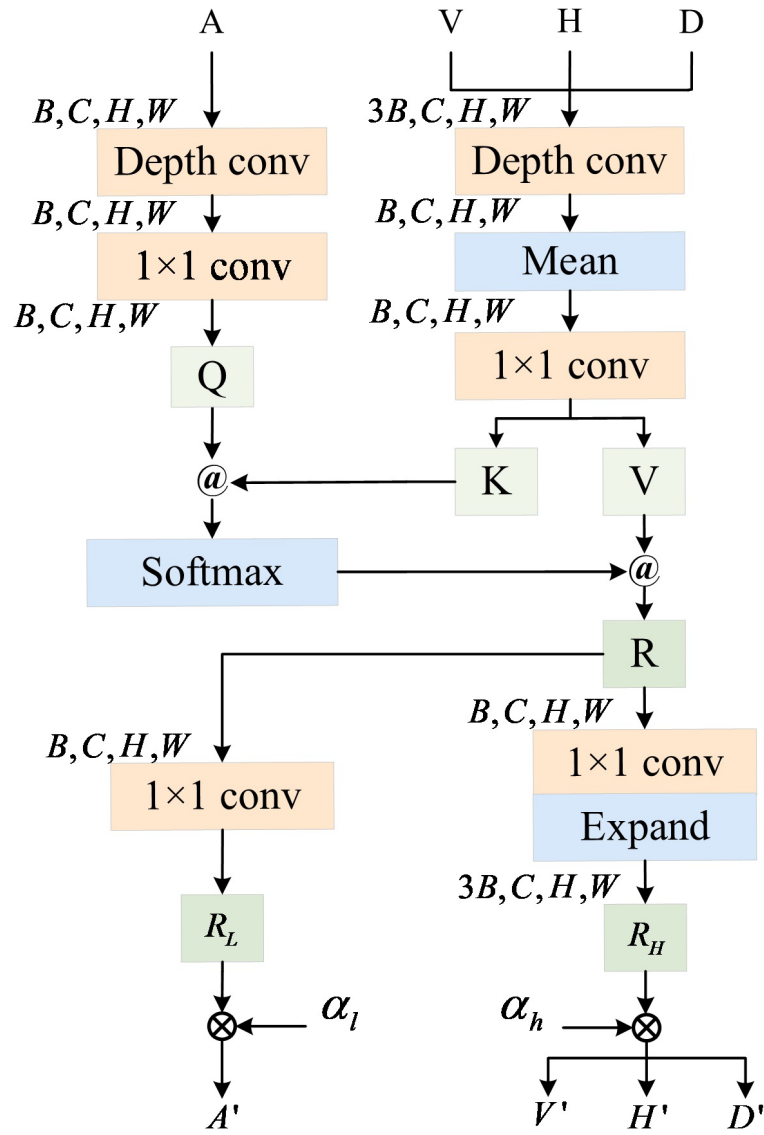


Figure 2. Cross-frequency calibration module  
图 2. 跨频率校正模块

$$R = \text{Softmax} \left( \frac{QK^T}{\sqrt{C_f}} \right) V_p \tag{19}$$

注意力结果  $R$  分别用于更新低频和低频分支：

$$A' = A_f + \lambda_a \cdot W_a(R) \tag{20}$$

$$F' = F_f + \lambda_h \cdot \text{Expand}(W_h(R)) \tag{21}$$

其中  $\lambda_a, \lambda_h$  为可学习参数(初始值 0.1)，Expand 将单份特征复制扩展为三份，恢复高频的三子带结构。

该模块在两级重建中各使用一次：先融合扩散恢复后的  $\hat{A}_2$  与增强后的二级高频  $V_1, H_1, D_1$ ，经逆小波变换得到一级低频；再融合  $\hat{A}_1$  与增强后的一级高频  $V_0, H_0, D_0$ ，最终重建增强图像。

### 3.6. 损失函数设计

总损失由三部分组成:

$$\mathcal{L}_{total} = \mathcal{L}_{diff} + \mathcal{L}_{photo} + \mathcal{L}_{freq} \quad (22)$$

去噪损失: 扩散模型的核心目标是使预测噪声与真实噪声尽可能接近, 采用均方误差(MSE)作为去噪损失:

$$L_{diff} = \mathbb{E}_{x_0, t, \epsilon} \left[ \left\| \epsilon - \epsilon_{\theta}(x_t, x_c, t) \right\|^2 \right] \quad (23)$$

其中  $\epsilon \sim \mathcal{N}(0, \mathbf{I})$ , 为前向过程中添加的真实噪声。通过最小化该损失, 模型能够准确估计出当前时间步所添加的噪声, 从而在条件信息的引导下实现从噪声状态到清晰图像的逆映射。

图像重建损失: 定义真实清晰图像为  $J$ , 模型预测图像为  $\hat{J}$ 。为提升恢复图像的主观质量, 结合 L1 损失与结构相似性(SSIM)损失:

$$\mathcal{L}_{photo} = \left\| \hat{J} - J \right\|_1 + \left( 1 - \text{SSIM}(\hat{J}, J) \right) \quad (24)$$

频域一致性损失: 为了约束小波域中低频结构与高频细节的一致性, 定义:

$$\mathcal{L}_{freq} = \lambda_1 \left( \sum_{k=0}^1 \left\| F'_k - F_{k,gr} \right\|^2 + \left\| A_2 - A_{2,gr} \right\|^2 \right) + \lambda_2 \left( \sum_{k=0}^1 TV(F'_k) + TV(A'_2) \right) \quad (25)$$

其中  $\lambda_1 = 0.1, \lambda_2 = 0.01$ ,  $TV$  为全变分正则项, 抑制伪影。

## 4. 实验结果与分析

### 4.1. 数据集和评价指标

为系统评估所提方法的有效性与泛化能力, 本文在典型水下图像数据集 UIEB [18]和 U45 [19]上开展对比实验。其中, UIEB 数据集提供参考图像, 用于有参考定量评价; U45 数据集不包含参考图像, 用于无参考质量评估。两类数据集的结合使得方法能够从像素重建精度与视觉感知质量两个层面对其性能进行综合评估。

对比方法选取了当前具有代表性的水下图像增强算法, 涵盖传统方法(如 UDCP [2]、UNTV [20]、WWPF [21])、基于生成对抗网络的方法(如 FUnIE-GAN [22])、以及近年来表现较优的深度学习方法(如 PUIE-Net [23]、U-transformer [12]、Diffwater [14]), 以保证实验结果的客观性与充分性。

在评价指标方面, 对于 UIEB 数据集, 采用 PSNR、MSE、SSIM、LPIPS、PCQI 和 FSIM 指标, 从像素误差、结构相似性及感知质量等多个角度进行综合评估; 对于 U45 数据集, 采用 UCIQE、UIQM 和 NIQE 无参考指标, 以衡量增强图像在真实水下场景中的视觉质量表现。

### 4.2. 对比实验

#### 4.2.1. 定量评估

在 UIEB 数据集上的定量结果如表 1 所示:

**Table 1.** Performance comparison on the UIEB dataset

**表 1.** 在 UIEB 数据集上的性能比较

算法	PSNR $\uparrow$	MSE $\downarrow$	SSIM $\uparrow$	LPIPS $\downarrow$	PCQI $\uparrow$	FSIM $\uparrow$
UDCP	13.72	0.051797	0.7091	0.1010	0.5748	0.8875

续表

FUnIE-GAN	17.92	0.022141	0.8377	0.0603	0.6351	0.9010
UNTV	14.64	0.037691	0.5214	0.1225	0.6471	0.6855
PUIE-Net	21.20	0.008861	0.9160	0.0413	0.7805	0.9505
WWPF	17.88	0.020753	0.8177	0.0760	<b>1.0387</b>	0.8710
U-transformer	21.34	0.010685	0.8171	0.0552	0.6706	0.8967
Diffwater	19.86	0.015427	0.9147	0.0382	0.9526	0.9526
our	<b>23.75</b>	<b>0.006539</b>	<b>0.9390</b>	<b>0.0358</b>	0.9159	<b>0.9596</b>

从整体结果来看,本文方法在重建精度与结构保持方面均表现出明显优势。相比现有方法,其在 PSNR 上取得最高值 23.75 dB, SSIM 亦达到最优,说明所提方法能够在恢复图像细节的同时有效保持结构一致性;同时, LPIPS 降至 0.0358,表明增强结果在视觉感知上更接近真实图像分布。结构与感知指标的同步提升进一步表明,本文方法在结构信息恢复与细节重建之间实现了有效协同,而非依赖单一的对比度增强机制。值得注意的是,尽管部分方法(如 WWPF)在 PCQI 上表现较高,但其在结构相关指标上的表现相对较弱,可能存在过度增强现象。相比之下,本文方法在多项指标之间取得了更为均衡且稳定的综合性能,体现出更优的整体恢复能力。

为了进一步验证本文方法的泛化能力,我们在 U45 数据集上进行了对比实验,结果如表 2 所示:

**Table 2.** Performance comparison on the U45 dataset

**表 2.** 在 U45 数据集上的性能比较

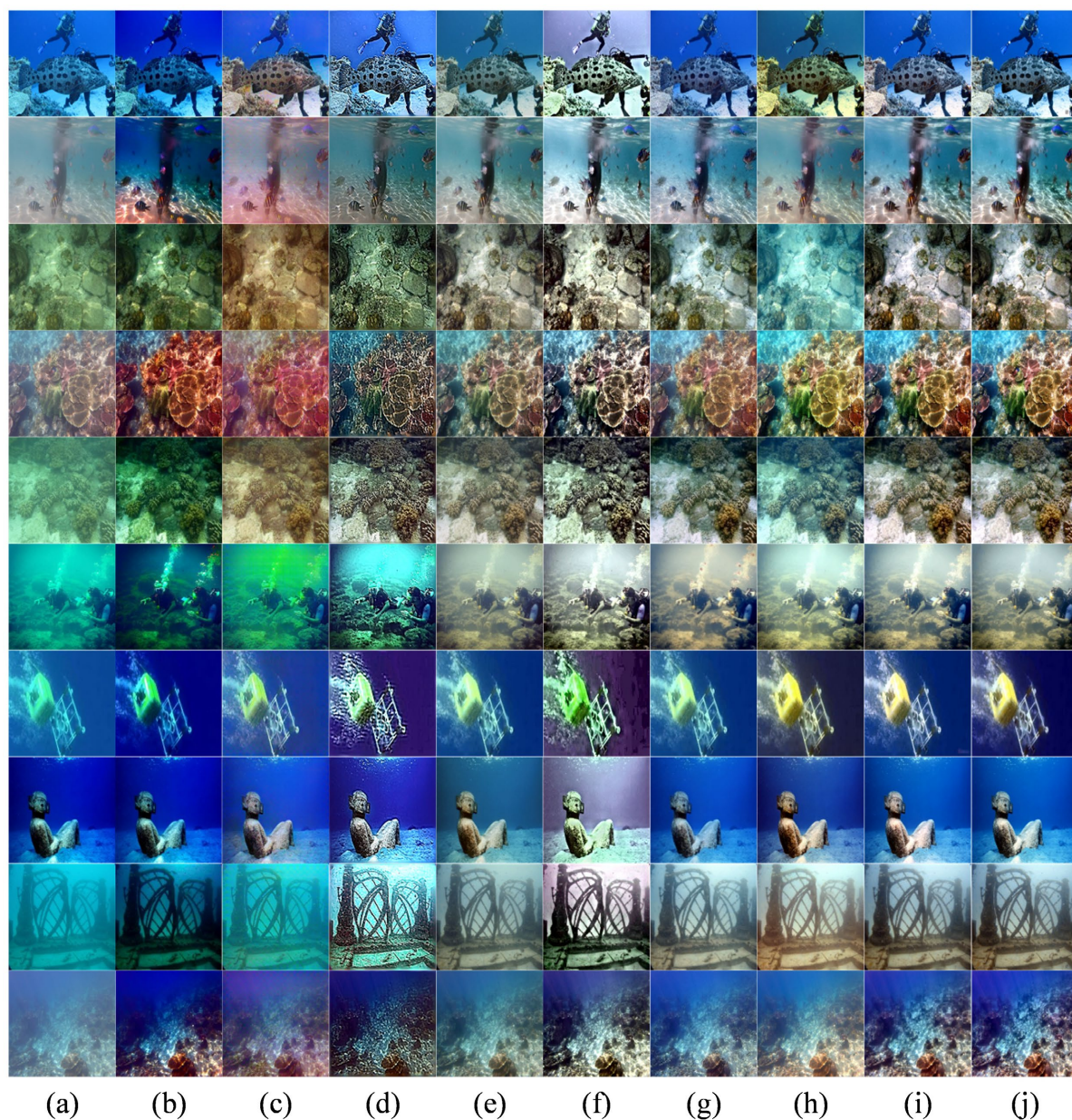
算法	UCIQE ↑	UIQM ↑	NIQE ↓
UDCP	0.571289	4.002527	6.30373
FUnIE-GAN	0.517628	4.338472	7.001771
UNTV	<b>0.606872</b>	3.976086	11.55812
PUIE-Net	0.527051	4.145846	6.061562
WWPF	0.594027	4.269372	6.745383
U-transformer	0.533973	4.209621	6.246205
Diffwater	0.578710	4.282377	<b>5.886452</b>
our	0.596174	<b>4.394435</b>	5.89939

实验结果表明,本文方法在无参考指标上同样表现出良好的综合性能。其中, UIQM 取得最优结果,说明增强图像在色彩表现与整体视觉质量方面具有明显优势;相较之下, UCIQE 略低于 UNTV 算法,但整体仍处于较高水平,表明该方法在对比度与综合色彩增强方面表现稳定。同时, NIQE 维持在较低水平 (5.899), 接近 Diffwater 的最优结果,说明在提升图像质量的同时未引入明显的伪影或不自然现象。

总体而言,本文方法在不同评价指标之间取得了较为均衡的表现,在提升视觉质量的同时避免了明显的失真问题,体现出良好的稳定性与泛化能力。

#### 4.2.2. 定性评估

为进一步验证各方法的视觉增强效果,图 3 和图 4 分别展示了 UIEB 与 U45 数据集上的对比结果。整体来看,不同方法在颜色恢复、细节保留以及视觉自然性方面存在明显差异。



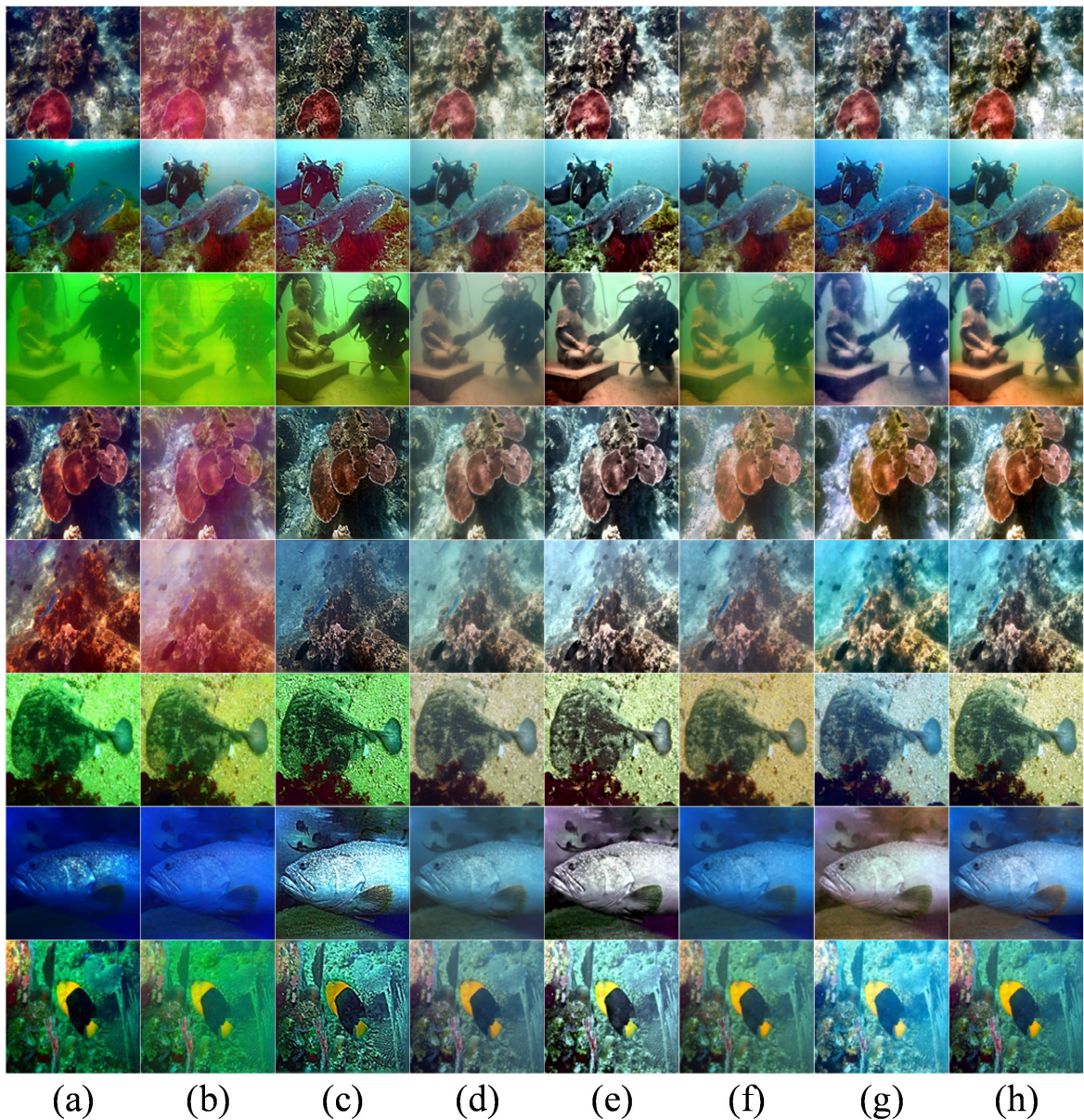
**Figure 3.** Visualization results of different methods on the UIEB dataset. From left to right, the images are: (a) original image, (b) UDCP, (c) FUnIE-GAN, (d) UNTV, (e) PUIE-Net, (f) WWPF, (g) U-transformer, (h) Diffwater, (i) Our, and (j) target image

**图 3.** 不同方法在 UIEB 数据集上的可视化结果。从左到右，图像分别是(a) 原始图像，(b) UDCP，(c) FUnIE-GAN，(d) UNTV，(e) PUIE-Net，(f) WWPF，(g) U-transformer，(h) Diffwater，(i) Our，(j) 目标图像

观察可知，传统方法(如 UDCP)普遍存在颜色偏移和对比度不足的问题，图像整体偏蓝或偏绿，细节表现较弱；FUnIE-GAN 虽然在一定程度上提升了亮度，但部分区域出现颜色过饱和或不自然现象。PUIE-Net、U-transformer 及 Diffwater 等方法在结构恢复方面有所改善，但在复杂场景中仍存在细节模糊、局部颜色失真或对比度不稳定等问题，例如在珊瑚纹理、海底区域以及远距离目标处表现不够稳定。

相比之下，本文方法在不同数据集与多种场景下均展现出更稳定的视觉效果。一方面，能够有效校正水下常见的颜色偏移，使图像色彩更自然；另一方面，纹理细节与边缘区域的恢复更清晰，有效避免

了过度平滑或伪影问题。此外，在不同光照与水体条件下，本文方法能够保持较为均衡的对比度表现，既未出现增强不足，也无明显的过度增强现象。



**Figure 4.** Visualization results of different methods on the U45 dataset. From left to right, the images are (a) UDCP, (b) FUnIE-GAN, (c) UNTV, (d) PUIE-Net, (e) WWPF, (f) U-transformer, (g) Diffwater, and (h) Our

**图 4.** 不同方法在 U45 数据集上的可视化结果。从左到右，图像分别是(a) UDCP, (b) FUnIE-GAN, (c) UNTV, (d) PUIE-Net, (e) WWPF, (f) U-transformer, (g) Diffwater, (h) Our

综合而言，本文方法在颜色一致性、细节恢复及整体视觉自然性方面均表现出良好的稳定性，能够在不同数据集上获得自然且清晰的增强结果。

### 4.3. 消融实验

为验证本文所提各模块的有效性，我们在 UIEB 数据集上进行了消融实验。以仅包含基础扩散模型

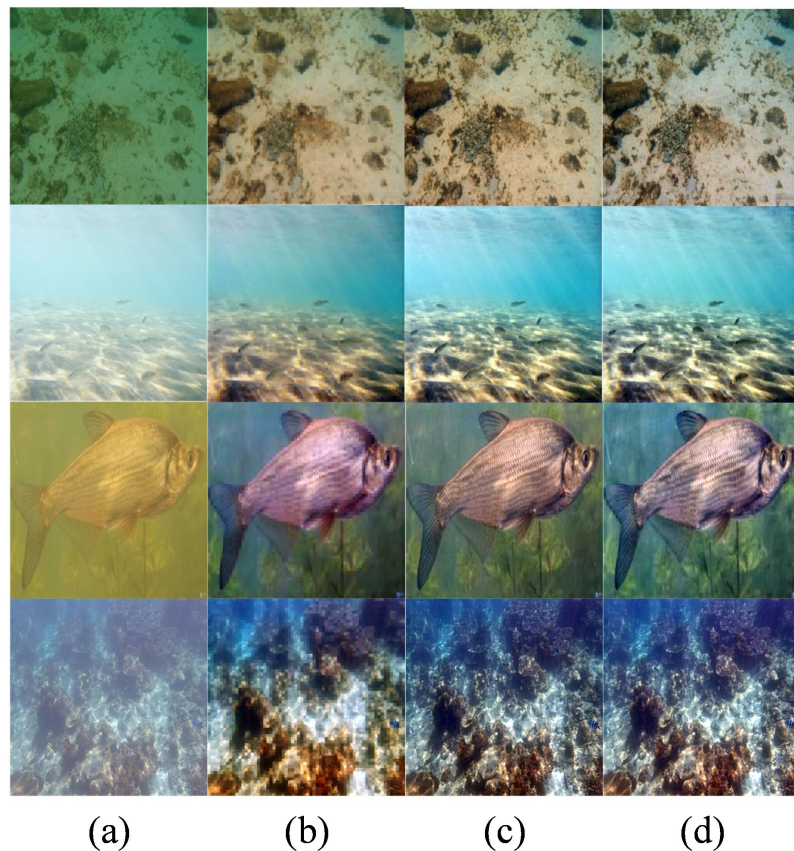
的结构为基线(记为 Base), 逐步引入高频增强模块 HFEM 和跨频率校正模块 CFC, 实验结果如表 3 所示:

**Table 3.** Ablation experiment results  
**表 3.** 消融实验结果

方法	PSNR $\uparrow$	MSE $\downarrow$	SSIM $\uparrow$	LPIPS $\downarrow$	PCQI $\uparrow$	FSIM $\uparrow$
Base	22.89	0.007413	0.8880	0.0414	0.7152	0.9112
Base + HFEM	23.59	0.006394	0.9384	0.0369	0.9151	0.9588
Base + HFEM + CFC	23.75	0.006539	0.9358	0.0358	0.9159	0.9596

从定量结果来看, 引入 HFEM 后, PSNR 由 22.89 dB 提升至 23.59 dB, SSIM 与 LPIPS 同步改善, 说明该模块能够有效增强纹理细节并提升结构表达能力。在此基础上进一步加入 CFC 后, PSNR 提升至 23.75 dB, LPIPS 降至 0.0358, FSIM 达到最优。尽管 SSIM 略有波动, 但整体性能仍保持最优水平, 说明该模块能够在细节恢复与结构保持之间实现更合理的平衡, 有效改善不同频率信息之间的一致性, 从而进一步提升整体重建质量。

在定量结果的基础上, 为进一步直观分析各模块对结构恢复与细节增强的影响, 本文给出了对应的可视化消融结果, 如图 5 所示。



**Figure 5.** Ablation results for each module of our method on the UIEB dataset. From left to right, the images are: (a) original image, (b) Base, (c) Base + HFEM, and (d) Base + HFEM + CFC

**图 5.** 在 UIEB 数据集上测试本方法各模块的消融结果。从左到右, 图像分别是(a) 原始图像, (b) Base, (c) Base + HFEM, (d) Base + HFEM + CFC

从视觉效果来看, 基线模型存在细节模糊和对比度不足的问题; 加入 HFEM 后, 纹理细节得到显著增强, 例如鱼体纹理和海底结构更加清晰; 进一步引入 CFC 后, 图像整体对比度与颜色过渡更加自然, 同时减少了局部过增强现象。

消融实验充分验证了本文提出的高频增强模块(HFEM)和跨频率校正模块(CFC)对水下图像增强任务的有效性。两者在细节增强与全局一致性方面形成互补, 共同实现了模型性能的进一步提升。

## 5. 结论

本文提出了一种融合小波频域分解与扩散模型的水下图像增强方法。该方法通过在小波域中解耦低频结构与高频细节信息, 引入条件扩散模型进行结构恢复, 并设计方向感知高频增强模块(HFEM)与跨频率校正模块(CFC), 实现结构与纹理的协同重建, 从而有效缓解水下图像中的颜色失真与细节退化问题。

在标准数据集上的实验结果表明, 本文方法具有较优的综合性能。相较现有最优方法, 在 UIEB 数据集上 PSNR 提升约 11.3%, 同时 LPIPS 降至 0.0358, 这说明在提升结构保真的同时显著改善了视觉感知质量。此外, 消融实验进一步验证了 HFEM 与 CFC 在细节增强与结构一致性方面的有效性。

尽管本文方法取得了较好的增强效果, 但在极端水下环境下仍存在一定挑战。未来工作将进一步优化扩散推理效率, 并结合物理先验信息, 以提升模型在复杂场景中的鲁棒性。

## 基金项目

2025 年度校级大学生创新创业训练项目(编号 X2025176)。

## 参考文献

- [1] Ho, J., Jain, A. and Abbeel, P. (2020) Denoising Diffusion Probabilistic Models. *Advances in Neural Information Processing Systems*, **33**, 6840-6851.
- [2] Drews Jr, P., do Nascimento, E., Moraes, F., Botelho, S. and Campos, M. (2013) Transmission Estimation in Underwater Single Images. 2013 *IEEE International Conference on Computer Vision Workshops*, Sydney, 2-8 December 2013, 825-830. <https://doi.org/10.1109/iccvw.2013.113>
- [3] Peng, Y. and Cosman, P.C. (2017) Underwater Image Restoration Based on Image Blurriness and Light Absorption. *IEEE Transactions on Image Processing*, **26**, 1579-1594. <https://doi.org/10.1109/tip.2017.2663846>
- [4] Akkaynak, D. and Treibitz, T. (2019) Sea-Thru: A Method for Removing Water from Underwater Images. 2019 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Long Beach, 15-20 June 2019, 1682-1691. <https://doi.org/10.1109/cvpr.2019.00178>
- [5] Iqbal, K., Odetayo, M., James, A., et al. (2010) Enhancing the Low-Quality Images Using Unsupervised Colour Correction Method. 2010 *IEEE International Conference on Systems, Man and Cybernetics*, Istanbul, 10-13 October 2010, 1703-1709. <https://doi.org/10.1109/icsmc.2010.5642311>
- [6] Fu, X., Zhuang, P., Huang, Y., Liao, Y., Zhang, X. and Ding, X. (2014) A Retinex-Based Enhancing Approach for Single Underwater Image. 2014 *IEEE International Conference on Image Processing (ICIP)*, Paris, 27-30 October 2014, 4572-4576. <https://doi.org/10.1109/icip.2014.7025927>
- [7] Ancuti, C., Ancuti, C.O., Haber, T. and Bekaert, P. (2012) Enhancing Underwater Images and Videos by Fusion. 2012 *IEEE Conference on Computer Vision and Pattern Recognition*, Providence, 16-21 June 2012, 81-88. <https://doi.org/10.1109/cvpr.2012.6247661>
- [8] Anwar, S., Li, C. and Porikli, F. (2018) Deep Underwater Image Enhancement.
- [9] Wang, Y., Guo, J., Gao, H., et al. (2021) UIEC<sup>2</sup>-Net: CNN-Based Underwater Image Enhancement Using Two Color Space. *Signal Processing: Image Communication*, **96**, Article 116250. <https://doi.org/10.1016/j.image.2021.116250>
- [10] Li, C., Anwar, S., Hou, J., Cong, R., Guo, C. and Ren, W. (2021) Underwater Image Enhancement via Medium Transmission-Guided Multi-Color Space Embedding. *IEEE Transactions on Image Processing*, **30**, 4985-5000. <https://doi.org/10.1109/tip.2021.3076367>
- [11] Fabbri, C., Islam, M.J. and Sattar, J. (2018) Enhancing Underwater Imagery Using Generative Adversarial Networks. 2018 *IEEE International Conference on Robotics and Automation (ICRA)*, Brisbane, 21-25 May 2018, 7159-7165.

- <https://doi.org/10.1109/icra.2018.8460552>
- [12] Peng, L., Zhu, C. and Bian, L. (2023) U-Shape Transformer for Underwater Image Enhancement. *IEEE Transactions on Image Processing*, **32**, 3066-3079. <https://doi.org/10.1109/tip.2023.3276332>
- [13] Chen, W., Lei, Y., Luo, S., Zhou, Z., Li, M. and Pun, C. (2024) UWFormer: Underwater Image Enhancement via a Semi-Supervised Multi-Scale Transformer. 2024 *International Joint Conference on Neural Networks (IJCNN)*, Yokohama, 30 June-5 July 2024, 1-8. <https://doi.org/10.1109/ijcnn60899.2024.10651366>
- [14] Guan, M., Xu, H., Jiang, G., Yu, M., Chen, Y., Luo, T., *et al.* (2024) DiffWater: Underwater Image Enhancement Based on Conditional Denoising Diffusion Probabilistic Model. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, **17**, 2319-2335. <https://doi.org/10.1109/jstars.2023.3344453>
- [15] Tang, Y., Kawasaki, H. and Iwaguchi, T. (2023) Underwater Image Enhancement by Transformer-Based Diffusion Model with Non-Uniform Sampling for Skip Strategy. *Proceedings of the 31st ACM International Conference on Multimedia*, Ottawa, 29 October-3 November 2023, 5419-5427. <https://doi.org/10.1145/3581783.3612378>
- [16] Du, D., Li, E., Si, L., Zhai, W., Xu, F., Niu, J., *et al.* (2025) UIEDP: Boosting Underwater Image Enhancement with Diffusion Prior. *Expert Systems with Applications*, **259**, Article 125271. <https://doi.org/10.1016/j.eswa.2024.125271>
- [17] Shi, X. and Wang, Y. (2024) CPDM: Content-Preserving Diffusion Model for Underwater Image Enhancement. *Scientific Reports*, **14**, Article No. 31309. <https://doi.org/10.1038/s41598-024-82803-y>
- [18] Li, C., Guo, C., Ren, W., Cong, R., Hou, J., Kwong, S., *et al.* (2019) An Underwater Image Enhancement Benchmark Dataset and beyond. *IEEE Transactions on Image Processing*, **29**, 4376-4389. <https://doi.org/10.1109/tip.2019.2955241>
- [19] Li, H., Li, J. and Wang, W. (2019) A Fusion Adversarial Underwater Image Enhancement Network with a Public Test Dataset.
- [20] Xie, J., Hou, G., Wang, G. and Pan, Z. (2021) A Variational Framework for Underwater Image Dehazing and Deblurring. *IEEE Transactions on Circuits and Systems for Video Technology*, **32**, 3514-3526. <https://doi.org/10.1109/tcsvt.2021.3115791>
- [21] Zhang, W., Zhou, L., Zhuang, P., Li, G., Pan, X., Zhao, W., *et al.* (2023) Underwater Image Enhancement via Weighted Wavelet Visual Perception Fusion. *IEEE Transactions on Circuits and Systems for Video Technology*, **34**, 2469-2483. <https://doi.org/10.1109/tcsvt.2023.3299314>
- [22] Islam, M.J., Xia, Y. and Sattar, J. (2020) Fast Underwater Image Enhancement for Improved Visual Perception. *IEEE Robotics and Automation Letters*, **5**, 3227-3234. <https://doi.org/10.1109/lra.2020.2974710>
- [23] Fu, Z., Wang, W., Huang, Y., Ding, X. and Ma, K. (2022) Uncertainty Inspired Underwater Image Enhancement. In: *Lecture Notes in Computer Science*, Springer, 465-482. [https://doi.org/10.1007/978-3-031-19797-0\\_27](https://doi.org/10.1007/978-3-031-19797-0_27)