

文章编号 1004-924X(2026)08-1330-17

## 融合动态感知与渐进细化的水下图像增强网络

孙士明, 张成雯\*, 李广峰, 高绍姝, 蒋铭皓

(中国石油大学(华东)青岛软件学院、计算机科学与技术学院,  
山东省智能油气工业软件重点实验室, 山东 青岛 266580)

**摘要:** 针对水体吸收与散射导致的图像颜色失真、对比度不足和尺度依赖性模糊问题, 本文提出了一种融合动态感知与渐进细化的 U 型水下图像增强网络。首先, 在编码器中设计动态调优 Transformer 模块, 通过提取各向异性特征来动态感知水下空间的非均匀退化, 实现空间上下文细节增强和全局色彩恢复。然后, 在解码器中构建由空洞卷积聚合模块和渐进式语义引导上采样模块组成的渐进细化架构。其中, 空洞卷积聚合模块利用阶梯式膨胀策略捕获多尺度上下文特征, 并整合通道混洗与残差注意力机制优化模糊区域重建, 提升图像的整体对比度和色彩保真度; 渐进式语义引导上采样模块采用级联式梯度保持策略补偿高频细节损失, 实现深层语义重构与浅层细节优化间的动态平衡。在 5 个公开数据集上与 9 种先进算法的对比结果表明, 所提方法在 UIEB 和 LSUI 数据集上的 PSNR、SSIM 及 MSE 指标均达到最优。相较于次优方法, 其在 UIEB 上的 PSNR 和 SSIM 分别提升了 1.47% 和 1.26%, MSE 降低了 8.33%; 在 LSUI 上 PSNR 和 SSIM 分别提升了 14.51% 和 6.92%, MSE 降低了 45.00%。此外, 在 UCCE, U45 和 C60 无参考数据集上, 本文方法在 UCIQE 指标上亦取得了最优结果。上述结果证明了所提方法在色偏纠正、对比度提升与细节恢复方面的优越性, 可为高级视觉任务提供高质量的图像输入。

**关键词:** 水下图像增强; 注意力机制; 动态感知; 渐进细化; 色彩校正; 特征融合

**中图分类号:** TP394.1; TH691.9 **文献标识码:** A

**doi:** 10.37188/OPE.20263408.1330 **CSTR:** 32169.14.OPE.20263408.1330

## Underwater image enhancement network integrating dynamic perception and progressive refinement

SUN Shiming, ZHANG Chengwen\*, LI Guangfeng, GAO Shaoshu, JIANG Minghao

(Qingdao Institute of Software, College of Computer Science and Technology, China University of Petroleum (East China), Shandong Key Laboratory of Intelligent Oil & Gas Industrial Software, Qingdao 266580, China)

\* Corresponding author, E-mail: z24070011@s.upc.edu.cn

**Abstract:** To address the color distortion, low contrast, and scale-dependent blurring caused by absorption and scattering in the underwater medium, a U-shaped underwater image enhancement network integrating dynamic perception and progressive refinement is proposed. Firstly, a dynamic optimization Transformer module is designed in the encoder, which dynamically perceives non-uniform degradation in underwater space by extracting anisotropic features, thereby enabling spatial contextual detail enhancement and

收稿日期: 2026-02-02; 修订日期: 2026-03-03.

基金项目: 国家自然科学基金(No. 61801517)

global color restoration. Subsequently, a progressive refinement architecture is constructed in the decoder, comprising a dilated convolution aggregation module and a progressive semantic-guided upsampling module. Specifically, the dilated convolution aggregation module utilizes a staircase dilation strategy to capture multi-scale contextual features, and integrates a channel shuffle operation and a residual attention mechanism to optimize blurred region reconstruction, thereby improving the overall contrast and color fidelity of the images. The progressive semantic-guided upsampling module adopts a cascaded gradient preservation strategy to compensate for high-frequency detail loss, achieving a dynamic balance between deep semantic reconstruction and shallow detail optimization. Compared to nine state-of-the-art algorithms across five public datasets, the proposed method achieves the best PSNR, SSIM, and MSE metrics on the UIEB and LSUI datasets. Relative to the second-best approaches, it represents improvements of 1.47% in PSNR and 1.26% in SSIM, alongside an 8.33% reduction in MSE on the UIEB dataset. On the LSUI dataset, the PSNR and SSIM are improved by 14.51% and 6.92%, respectively, while the MSE is reduced by 45.00%. Furthermore, on the UCCS, U45, and C60 no-reference datasets, our method also yields the best results in terms of the UCIQE metric. These results validate the superiority of the proposed method in correcting color deviation, improving contrast, and recovering details, indicating that the generated images can serve as high-quality inputs for advanced visual tasks.

**Key words:** underwater image enhancement; attention mechanism; dynamic perception; progressive refinement; color correction; feature fusion

## 1 引言

高质量的水下图像在水下导航、海洋生态监测和水下考古等任务中发挥着至关重要的作用<sup>[1-2]</sup>。然而,受复杂成像环境的影响,获取的水下图像常伴有颜色失真、对比度低和细节模糊等非均匀退化现象<sup>[3-4]</sup>。水下图像增强旨在恢复受损的视觉信息并提升图像清晰度,以满足广泛的实际应用需求<sup>[5-6]</sup>。

现有的水下图像增强方法主要分为传统方法和基于深度学习的方法<sup>[7-8]</sup>。其中,传统方法又可细分为基于像素空间调整的方法和基于物理模型的方法。基于像素空间调整的方法侧重于直接调整像素值来改善图像质量,常用技术包括对比度拉伸<sup>[9]</sup>、自动白平衡<sup>[10]</sup>等。例如,Subramani等人<sup>[11]</sup>利用高斯金字塔变换和上下文对比度优化策略提升图像质量。然而,此类方法通常忽略了水下光选择性吸收特性所导致的空间非均匀退化,易导致增强结果出现色彩失真或视觉伪影<sup>[12]</sup>。基于物理模型的方法则通过反演水下成像模型,结合现有先验知识和假设估计关键参数,进而通过逆运算推断出增强图像<sup>[13]</sup>。例如,Kong等人<sup>[14]</sup>提出一种浊度分析策略,通过融合

局部对比度、饱和度和边缘密度特征来实现透射率的估计。然而,受限于真实水下光照环境与浑浊度的动态多变性,以及待估参数的欠定性问题,此类方法的泛化能力较为有限<sup>[15]</sup>。

相比之下,基于数据驱动的深度学习方法凭借其强大的非线性表征能力,能够自适应学习关键特征,从而在复杂多变的水下环境中展现出更优的鲁棒性<sup>[16]</sup>。然而,现有主流网络架构仍存在一定的局限:卷积神经网络虽在局部特征提取方面能力突出,但其固定大小的感受野难以充分捕捉全局语义信息,使其缺乏对全局非均匀光照分布的感知能力;而Transformer虽具备长距离建模优势,但其自注意力机制在视觉重构任务中易削弱或丢失结构信息<sup>[17]</sup>。特别是在复杂水下环境中,悬浮颗粒引起的前向散射已导致图像的高频细节丢失与边缘模糊,若仅使用Transformer网络,其对结构特征的弱化会进一步加剧该现象,导致增强的图像缺乏锐度。

针对上述挑战,构建CNN和Transformer的混合框架已成为当下提升水下图像增强性能的有效方案之一<sup>[18]</sup>。Chen等人<sup>[19]</sup>提出半监督框架UWFormer,利用非线性频率感知注意力模块提升水下图像的复原质量。Shang等人<sup>[20]</sup>引入亮度

引导的 Transformer, 以实现多尺度特征的动态加权融合。Lee 等人<sup>[21]</sup>构建了 Sobel 引导的稀疏注意力和交叉嵌套 Transformer 结构, 通过区分颜色偏移与物体散射来缓解颜色通道中的不规则损失。Khan 等人<sup>[22]</sup>利用相位对退化不敏感的特性, 设计了一种轻量级相位注意力机制, 能够有效传输结构信息并抑制冗余特征。然而, 由于光线随景深的非线性衰减, 近景与远景的模糊程度及噪声分布存在显著差异, 导致图像呈现出强烈的尺度依赖性模糊。而现有方法大多采用单一的特征提取范式, 难以在特征层面对此类多尺度退化进行有效解耦与自适应感知, 因此在恢复多尺度结构细节时仍存在局限性。

针对上述问题, 本文提出一种融合动态感知与渐进细化的 U 型水下图像增强网络。首先, 在编码器中设计动态调优 Transformer 模块, 该模块内嵌自适应动态特征校准机制和特征细化模块。前者融合风车形卷积、动态响应调制单元和卷积注意力机制, 恢复图像的高频纹理和全局色彩。后者整合残差连接与动态激活函数, 聚合局部上下文信息并滤除冗余特征。然后, 针对水下图像的尺度依赖性模糊问题设计空洞卷积聚合模块。该模块通过构建多尺度感受野来解耦不同尺度的退化特征, 并引入通道混洗和残差注意力机制自适应地恢复结构细节。最后, 针对网络深层特征重建过程中的纹理丢失问题提出渐进

式语义引导上采样模块。该模块利用级联式梯度保持策略增强局部感知能力, 实现对空间解码特征的结构约束与纹理细化。

## 2 本文方法

### 2.1 总体网络结构

本文提出的水下图像增强网络结构如图 1 所示。输入图像首先经由风车形卷积 (Pinwheel-shaped convolution, PConv)<sup>[23]</sup> 处理, 提取为初始浅层特征。然后, 该特征进入编码器, 通过四个级联的动态调优 Transformer 模块 (Dynamic optimization Transformer module, DOTM) 逐步映射至潜在特征空间。为弥补下采样引起的高频细节损失, 各层级的编码特征通过基于 PConv 的跳跃连接传输至解码器。在解码阶段, 网络采用了由空洞卷积聚合模块 (Dilated convolution aggregation module, DCAM) 与渐进式语义引导上采样模块 (Progressive semantic-guided upsampling module, PSUM) 构成的渐进细化架构。其中, 前一层的深层特征与相应编码器特征首先在 DCAM 中进行多尺度融合, 旨在恢复结构细节并抑制噪声。随后, 融合特征由 PSUM 恢复空间分辨率并增强结构与纹理细节。最后, 解码器的输出特征经由末端卷积层映射回图像空间, 生成高质量的增强图像。

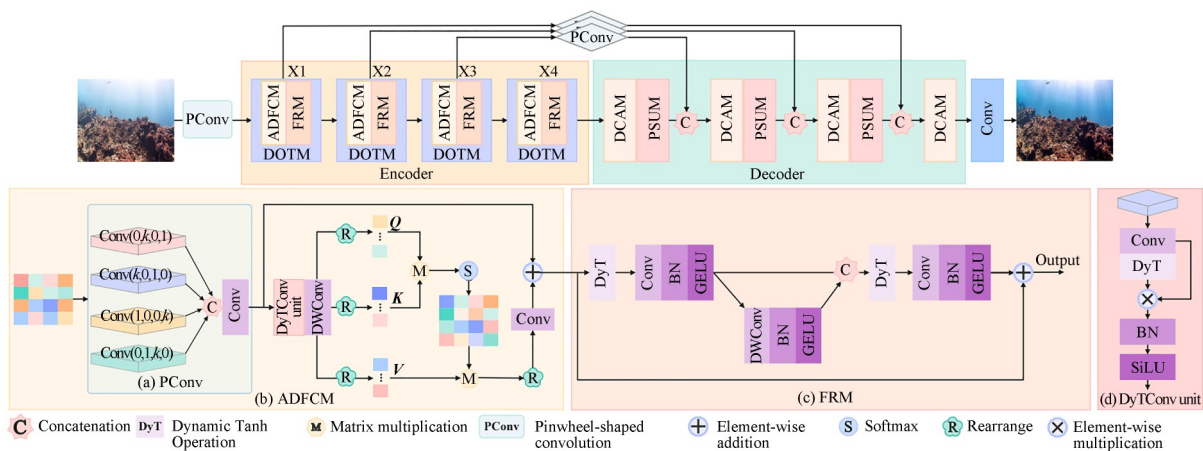


图 1 网络整体结构图 (其中, (a), (b), (c) 和 (d) 分别为 DOTM 模块中 PConv, ADFCM, FRM 和 DyTConv unit 的结构示意图。X1, X2, X3 和 X4 分别对应编码器四个阶段的输出特征)

Fig. 1 Overall architecture of the proposed network (Specifically, (a), (b), (c), and (d) are illustrations of the PConv, ADFCM, FRM and DyTConv unit structures in the DOTM. X1, X2, X3, and X4 correspond to the output features of the four stages of the encoder, respectively)

## 2.2 风车形卷积(PCConv)

受水下定向光线散射和不规则光照波动的影响,水下图像的纹理退化具有强烈的各向异性。传统标准卷积受限于固定的方形感受野,难以有效捕捉图像中的各向异性特征。相比之下,PCConv利用多方向条形卷积核,能够更敏锐地感知特定方向上的结构变化,减少水下微小细节在浅层提取阶段的特征损失。因此,本文引入PCConv作为浅层特征提取和跳跃连接的基本单元,增强网络对各向异性特征的提取能力。

如图1(a)所示,PCConv采用四组并行非对称填充的条形卷积,旨在有效捕捉多方向的局部结构与纹理特征,实现在显著扩大感受野的同时有效控制参数量的增长。随后,将提取的四个方向特征在通道维度拼接,并通过无填充的 $2 \times 2$ 卷积进行融合以生成最终输出。这种设计打破了传统卷积的空间对称性束缚,实现了对水下场景中复杂空间结构信息的各向异性感知。

## 2.3 动态调优Transformer模块(DOTM)

为应对水下环境中的空间非均匀退化,本文在编码器中设计了DOTM。该模块由自适应动态特征校准机制(Adaptive Dynamic Feature Calibration Mechanism, ADFCM)和特征细化模块(Feature Refinement Module, FRM)级联构成。

### 2.3.1 自适应动态特征校准机制(ADFCM)

如图1(b)所示,ADFCM通过融合风车形卷积、动态响应调制单元和卷积注意力机制,实现退化水下图像的细节增强和全局色彩恢复。输入特征首先经由PCConv处理以提取各向异性特征,生成中间特征 $Y$ 。随后,特征 $Y$ 输入至动态响应调制单元(DyTConv unit)进行特征自适应调制。针对水下场景特有的非均匀光照与散射退化,DyTConv unit引入可学习的非线性激活函数 $DyT(x)$ <sup>[24]</sup>来动态校准特征,实现对低频噪声的抑制和局部细节的增强。相较于传统的层归一化,DyT(x)利用非线性映射有效抑制极端激活值的干扰,在提升网络对特征尺度变化适应性的同时,减轻了模型对均值与方差的依赖。DyTConv unit的核心计算如下:

$$DyT(x) = \gamma * \tanh(\alpha x) + \beta, \quad (1)$$

$$M_d = DyT(Conv_1(Y)) \odot Conv_1(Y), \quad (2)$$

$$M_m = SiLU(BN(M_d))$$

其中: $\odot$ 表示逐元素相乘, $Conv_1$ 表示 $1 \times 1$ 卷积, $\alpha, \beta$ 和 $\gamma$ 为可学习参数, $BN$ 表示批归一化。

在完成局部特征的动态校准后,为进一步纠正由水体吸收导致的全局色彩失真,本文引入卷积注意力机制<sup>[25]</sup>来建立所有位置之间的长程依赖关系。具体而言,增强后的特征 $M_m$ 经过深度卷积层为每个注意力头 $h$ 映射相应的 $Q_h, K_h$ 和 $V_h$ 。为稳定注意力分布并防止数值爆炸,模型对 $Q_h$ 和 $K_h$ 采用L2归一化,并设置可学习的参数 $T_h$ 替代标准缩放因子,使模块能灵活聚焦于不同退化程度的关键区域。参数 $T_h$ 应用于Softmax函数前,用于动态调节 $K$ 与 $Q$ 点积的大小。随后,利用Rearrange操作恢复原始的空间维度排列,并通过 $1 \times 1$ 卷积整合通道信息,获得特征 $Y_m$ 。卷积注意力机制的核心计算公式如下:

$$Q'_h = Norm(Q_h), K'_h = Norm(K_h), \quad (3)$$

$$Y_m = Conv_1 \left( Rearrange \left( Sm \left( \frac{Q'_h \cdot K'_h}{T_h} \right) \cdot V_h \right) \right), \quad (4)$$

其中: $Sm$ 表示Softmax激活函数, $\cdot$ 表示矩阵乘法。最后,为防止深层特征退化,将特征 $Y_m$ 与PCConv提取的局部空间特征 $Y$ 进行逐元素相加,实现局部细节与全局语义特征的自适应融合。

### 2.3.2 特征细化模块(FRM)

考虑到单独使用ADFCM易导致特征同质化和局部细节丢失问题,为在保持全局色彩的前提下恢复局部高频细节,本文提出FRM模块。该模块通过整合动态激活函数(DyT(x))和残差连接,实现聚合局部上下文信息并滤除冗余特征。FRM结构如图1(c)所示,输入特征 $Y$ 首先经过DyT(x)抑制极端噪声并稳定特征分布。然后,特征输入至CBG单元(包含 $1 \times 1$ 卷积、批归一化和GELU激活函数)进行跨通道交互,生成特征 $M$ 。为缓解深层特征提取过程中产生的细节模糊,将特征 $M$ 与经由DCBG单元(包含深度卷积、批归一化和GELU激活函数)处理后的特征在通道维度拼接,生成特征 $Y'_d$ 。其中,DCBG单元利用深度卷积对局部空间邻域进行像素感知,通过增强相邻像素间的结构关联来有效聚合上下文信息,同时将特征通道扩展至原来的四倍。随后, $Y'_d$ 经由DyT(x)和CBG单元进行特征压缩与重建。在此过程中,DyT(x)对 $Y'_d$

进行自适应的非线性校准,旨在抑制拼接过程中引入的散射噪声。CBG 单元则利用  $1 \times 1$  卷积进行降维和跨通道的特征提取。最后,FRM 利用残差连接将重建的精细特征与初始输入  $Y_o$  进行逐元素相加,生成最终输出  $Y_{out}$ 。残差连接将  $Y_o$  提供的全局色彩信息和基础结构信息作为恒等映射,与主干路径提取的高频特征融合,在进一步细化局部结构的同时,减轻网络层次加深导致的特征退化问题。FRM 的核心计算公式如下:

$$M = CBG(DyT(Y_o)), \quad (5)$$

$$Y'_o = Concat(M, DCBG(M)), \quad (6)$$

$$Y_{out} = CBG(DyT(Y'_o)) + Y_o, \quad (7)$$

其中,  $Concat(\cdot)$  代表通道拼接操作。

#### 2.4 空洞卷积聚合模块(DCAM)

非线性光散射导致水下图像呈现尺度依赖性模糊:近景区域纹理保留相对完整但易受噪声干扰,而中远景区域的结构细节因长距离散射丢失严重。传统固定感受野的卷积难以自适应这种随深度动态变化的退化特征。针对上述问题,本文设计了 DCAM 模块。该模块利用空洞卷积

在不增加参数量及计算开销的前提下能显著扩大感受野的固有优势,通过构建阶梯式多尺度感知域来解耦不同深度的退化特征。

DCAM 结构如图 2 所示,输入特征  $D$  首先经由四个并行的 CBR 分支处理。受 DeepLabv3<sup>[26]</sup> 启发,这四个分支中的卷积层分别采用了  $d \in 1, 6, 12, 18$  的膨胀率,旨在构建多层次的空间感知域。具体而言,  $d=1$  的 CBR 分支专注于恢复局部细粒度纹理,以防止在特征聚合时高频细节的丢失。 $d=6, 12, 18$  的 CBR 分支通过构建由近及远的层次化感受野,实现对中远景区域内不同散射效应的有效分离与特征重建。这种从局部到全局的阶梯式感受野设计能有效避免大空洞率易引发的局部信息丢失问题,实现多尺度特征提取时的平滑过渡。其中, CBR 分支的定义如公式(8)所示:

$$F_m = ReLU(BN(DConv_d(D))) \quad (8)$$

$$d = 1, 6, 12, 18,$$

其中:  $DConv_d$  表示膨胀率为  $d$  的卷积。随后,各分支输出特征  $F_m$  沿通道维度拼接,生成融合特征  $X_{out}$ 。

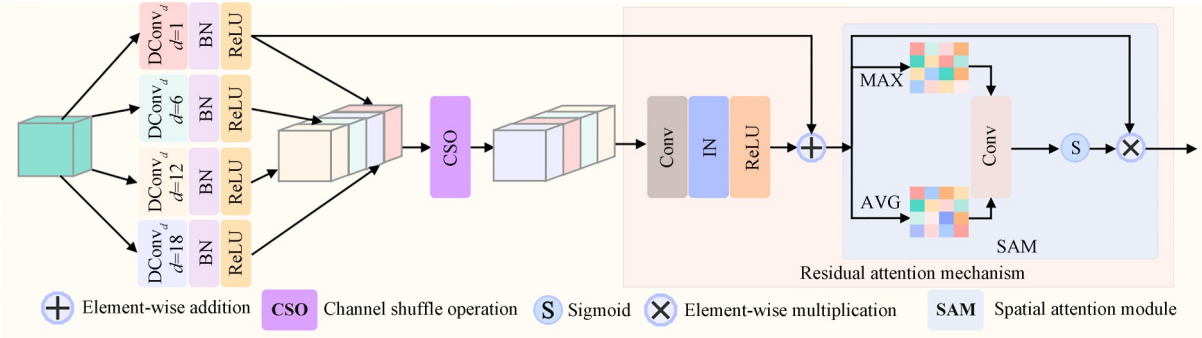


图 2 空洞卷积聚合模块

Fig. 2 Dilated convolution aggregation module

为促进跨尺度特征的交互与颜色互补,在完成多尺度特征的提取与拼接后,引入通道混洗操作进行通道重排。此外,针对大感受野可能引入背景噪声的问题,并确保亮度与色彩的自然过渡,本文进一步设计了融合 CIR 单元、残差连接和空间注意力机制的残差注意力机制。其中,由卷积、实例归一化和 ReLU 激活函数组成的 CIR 单元完成特征整合后,通过残差连接将其与  $d=1$  CBR 分支的输出跨层相加。该设计旨在防止网

络过深引起的色彩退化,同时增强多尺度特征的结构细节。接着,空间注意力机制利用池化操作对特征频率的差异化响应,自适应地学习空间权重分布,在实现局部细节增强和抑制噪声的同时,确保亮度与色彩的自然过渡。空间注意力机制的核心计算公式为:

$$M_s = \sigma\{Conv[AAP(F_s); MP(F_s)]\} \odot F_s, \quad (9)$$

其中:  $MP$  表示沿通道维度的最大池化操作,

AAP为自适应平均池化操作, $\sigma$ 表示Sigmoid激活函数。

## 2.5 渐进式语义引导下采样模块(PSUM)

为解决水下图像在特征重建阶段由于细节退化与结构异常导致的累积误差,本文提出PSUM模块。该模块结合双线性插值操作和级联式梯度保持策略,恢复空间分辨率并增强结构与纹理细节。

PSUM结构如图3所示,输入特征 $P_m$ 首先经过双线性插值恢复空间分辨率,随后通过PConv

进一步细化各向异性空间结构,得到特征 $E_y$ 。最后, $E_y$ 进入级联式高频建模机制进行处理,旨在通过梯度保持与特征重建的联合优化,实现空间解码特征的结构约束与纹理细化。具体而言,特征 $E_y$ 在各分支中首先经过 $3 \times 3$ 的最大池化操作提取显著的结构响应特征。接着依次通过深度卷积、实例归一化和Sigmoid激活函数生成特征 $E_l$ 。此过程旨在防止强纹理信号在传递过程中被稀释,并缓解反向传播过程中的梯度衰减问题。

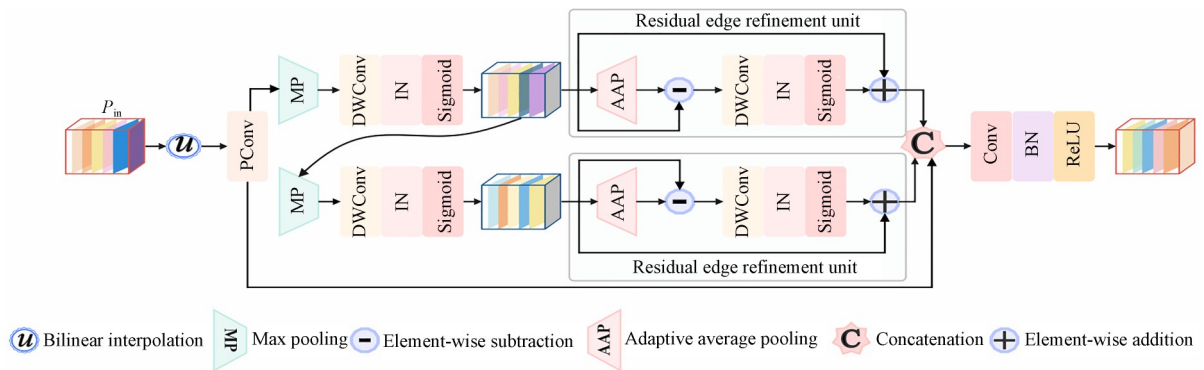


图3 渐进式语义引导下采样模块

Fig. 3 Progressive semantic-guided upsampling module

为补偿上采样过程中不可避免的高频细节损失,受迭代反向投影更新机制的启发,本文在级联链路的每一级中设计了残差边缘重建单元(Residual Edge Refinement Unit, RER)实施隐式高频建模。具体而言,RER首先计算特征 $E_l$ 与其自适应平均池化特征的差值。通过从原始特征中减去平滑后的低频特征,以实现高频边缘残差的精确解耦。随后,利用深度卷积、实例归一化和Sigmoid激活函数对提取的残差特征进行非线性校准,并将校准后的高频信息与输入特征 $E_l$ 相加,旨在确保微小的纹理细节在多级聚合过程中不会被淹没,进而实现纹理锐度与结构一致性的双重提升。RER的核心计算过程如公式(10)所示:

$$D_l = \sigma\left(\text{IN}\left(\text{DWConv}_l\left(E_l - \text{AAP}\left(E_l\right)\right)\right)\right) + E_l, \quad (10)$$

其中: $\text{DWConv}_l$ 代表深度卷积, $l$ 为分支索引,IN代表实例归一化。最后,将各分支的输出沿通道维度级联,并与特征 $E_y$ 融合,经由 $1 \times 1$ 卷积、批

归一化和ReLU激活函数处理,输出最终的高分辨率重建特征 $D_{\text{out}}$ :

$$D_{\text{out}} = \text{ReLU}\left(\text{BN}\left(\text{Conv}\left(\text{Concat}\left(E_y, D_1, D_2\right)\right)\right)\right). \quad (11)$$

## 2.6 损失函数

为实现水下图像中色彩与纹理结构的一致性,本文采用分阶段反向传播策略<sup>[27]</sup>构建联合损失函数 $L_{\text{joint}}$ :

$$L_{\text{joint}} = \{L_{\text{lap}}, L_{\text{R}}\}. \quad (12)$$

$L_{\text{joint}}$ 通过在单次迭代中解耦优化目标,先后利用拉普拉斯损失 $L_{\text{lap}}$ 与图像感知重建损失 $L_{\text{R}}$ 独立更新网络参数,旨在避免多损失干扰的同时,锐化图像的结构边缘,提升整体视觉感知质量。下面对各阶段损失函数进行具体说明。

(1) 拉普拉斯损失:为减少图像边缘纹理信息的损失,引入拉普拉斯算子作为边缘检测算子,增强模型的鲁棒性与泛化能力。拉普拉斯模

板定义如式(13)所示:

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & -8 & 1 \\ 1 & 1 & 1 \end{pmatrix}. \quad (13)$$

设增强后的图像为  $I_t$ , 真实标签图像为  $I_g$ 。拉普拉斯损失定义为:

$$L_{\text{lap}} = \frac{1}{N} \sum_{i=1}^N (\text{lap}(I_g^i) - \text{lap}(I_t^i))^2, \quad (14)$$

其中:  $N$  代表像素总数,  $i$  表示像素位置,  $\text{lap}(\cdot)$  表示拉普拉斯卷积操作。

(2) 图像感知重建损失: 图像感知重建损失  $L_R$  结合了平滑  $L_1$  损失和 VGG 感知损失。其中, 平滑  $L_1$  损失为小误差提供了平滑的梯度, 并降低了对大偏差的敏感度, 有助于实现像素级一致性和训练稳定性。平滑  $L_1$  损失的定义如公式(15)所示:

$$\text{Smooth}L_1(x) = \begin{cases} 0.5x^2, & \text{if } |x| < 1 \\ |x| - 0.5, & \text{otherwise} \end{cases}, x = I_g^i - I_t^i, \quad (15)$$

其中,  $x = I_g^i - I_t^i$  表示真实标签图像与增强图像之间的像素差值。

为提升高层语义一致性, 采用基于预训练 VGG-16 网络的感知损失。该损失通过评估增强图像与参考图像在特征空间上的差异来优化视觉质量, 感知损失  $L_v$  定义如下:

$$L_v = \frac{1}{CWH} \sum_{c=1}^C \sum_{w=1}^W \sum_{h=1}^H \left( V(I_g)_{c,w,h} - V(I_t)_{c,w,h} \right)^2, \quad (16)$$

其中:  $C, W, H$  分别表示通道数、宽度和高度,  $V(\cdot)$  表示由 VGG 网络提取的特征图。

图像感知重建损失定义为:

$$L_R = \text{Smooth}L_1 + \lambda \cdot L_v, \lambda = 0.5. \quad (17)$$

### 3 实验设计与结果分析

#### 3.1 实验环境

所有实验均在 Ubuntu 操作系统下进行, 使用硬件平台配置为 NVIDIA Tesla V100 GPU 进行训练。训练时所有输入图像设置为  $256 \times 256$  大小, 优化器选用 AdamW, Batch Size 设置为 4, 初始学习率设置为 0.000 2。

#### 3.2 实验设置

本文选用五个公开的水下图像数据集进行

模型的训练与性能评估, 具体如下:

(1) LSUI<sup>[28]</sup>: 包含 4 279 对真实水下图像, 涵盖了不同的水域类型和光照条件。本文从中随机划分 3 423 对构建训练集 (Train-LSUI), 剩余全部图像构成测试集 (Test-LS856)。

(2) UIEB<sup>[29]</sup>: 该数据集包含 890 对配对的水下图像。实验时随机选取 800 对用于训练 (Train-UIEB), 剩余 90 对用作测试 (Test-UI90)。

(3) C60: 作为 UIEB 数据集的挑战子集, 该数据集包含 60 张极端退化场景的水下图像, 其全部图像用于构建测试集 (Test-C60), 以评估模型在恶劣水下环境中的泛化能力。

(4) UCCS<sup>[30]</sup>: 该数据集为 RUIE 数据集的子集, 其中包括不同级别的模糊、色偏及光照变化的图像。UCCS 的所有图像作为测试集 (Test-UCCS) 来重点评估模型的色彩恢复性能。

(5) U45<sup>[31]</sup>: 该数据集包含 45 幅极端水下场景下的图像, 具有显著的色偏、低对比度及雾化特征。其所有图像均用于构建测试集 (Test-U45), 旨在检验算法处理多重极端退化特征时的鲁棒性。

为了定量评估不同方法的增强性能, 本文选取了三项全参考图像评价指标: 峰值信噪比 (PSNR)<sup>[32]</sup>、结构相似性 (SSIM)<sup>[33]</sup> 和均方误差 (MSE)<sup>[34]</sup>。其中, PSNR 用于衡量增强图像与真值图像在像素值上的差异程度, 其计算公式如下:

$$\text{PSNR} = 10 \log_{10} \left( \frac{255^2}{\text{MSE}} \right), \quad (18)$$

其中: 255 为 8-bit 图像中像素的最大值, MSE 为增强图像与真值图像的均方误差。

SSIM 从亮度、对比度和结构三个方面衡量两个图像之间的相似度, 其计算公式如下:

$$\text{SSIM}(x, y) = \frac{(2\mu_x\mu_y + \beta_1)(2\sigma_{xy} + \beta_2)}{(\mu_x^2 + \mu_y^2 + \beta_1)(\sigma_x^2 + \sigma_y^2 + \beta_2)}, \quad (19)$$

其中:  $x, y$  分别为增强图像  $I_t$  与真值图像  $I_g$ ;  $\mu_x, \mu_y$  为图像  $x, y$  的平均值;  $\sigma_x^2, \sigma_y^2$  为图像  $x, y$  的方差;  $\sigma_{xy}$  为图像  $x, y$  的协方差;  $\beta_1, \beta_2$  为两个常数, 用于避免分母为零。

MSE 用于量化像素级误差的强度。对于大

小为  $m \times n$  的图像, MSE 的计算公式为:

$$MSE = \frac{1}{mn} \sum_{i=1}^m \sum_{j=1}^n (I_g(i, j) - I_l(i, j))^2, \quad (20)$$

其中:  $I_g(i, j)$  为真值图像在像素点  $(i, j)$  的像素值,  $I_l(i, j)$  为增强图像在像素点  $(i, j)$  的像素值。

此外, 本文还引入了四项无参考图像质量评价指标, 旨在从视觉感知角度全面评估增强图像的视觉效果, 包括水下图像质量度量 (UIQM)<sup>[35]</sup>、水下彩色图像质量评价 (UCIQE)<sup>[36]</sup>、自然图像质量评价 (NIQE)<sup>[37]</sup> 和感知图像质量评价 (PIQE)<sup>[38]</sup>。在上述指标中, 除 MSE、NIQE 和 PIQE 数值越低代表图像质量越好外, 其余指标数值越高表示图像质量越好。

### 3.3 对比实验

本文将所提方法与 9 种主流方法进行了对比, 具体包括 UDCP<sup>[39]</sup>, DICAM<sup>[40]</sup>, UDNet<sup>[41]</sup>, SuirSIR<sup>[42]</sup>, Ushape<sup>[28]</sup>, SMDR<sup>[43]</sup>, FGUIR<sup>[44]</sup>, UIEAnythingv2<sup>[45]</sup> 和 PyUIE<sup>[46]</sup>。

#### 3.3.1 定性分析

图 4 和图 5 分别展示了 10 种方法在 Test-UI90 和 Test-LS856 数据集上的可视化增强结果。具体对比分析如下: UDCP 处理后的图像整体对比度偏低。DICAM 虽然改善了图像的整体色彩, 但在暗光场景中易引入伪影。UDNet 和

Ushape 未能有效去除水下雾化效应, 导致结果呈现不同程度的蓝色色偏。SuirSIR, FGUIR 和 UIEAnythingv2 对红色通道进行了过度补偿, 使图像呈现不自然的红色基调 (彩图见期刊电子版)。SMDR 有效改善了细节模糊现象, 但导致生成图像的整体亮度过高。PyUIE 增强后的图像存在细节丢失和伪影。相比之下, 本文方法不仅实现了更高的色彩保真度, 还精准恢复了光影纹理 (如水波纹) 及细微结构 (如鱼鳍、礁石表面), 生成结果在视觉上更接近真实的水下自然场景。

为了进一步评估各方法的泛化能力, 本文从 Test-C60, Test-UCCS 和 Test-U45 三个无参考数据集中分别选择两幅代表性图像进行详细分析。如图 6 所示, UDCP 算法压低了背景亮度, 加剧了图像的蓝绿色偏。DICAM 生成的图像亮度偏高, 在处理偏绿色场景时存在色彩恢复不足的问题。UDNet 整体对比度欠佳, 无法有效去除雾状干扰。SuirSIR 和 Ushape 的噪声抑制能力较弱, 容易在增强过程中引入伪影。SMDR、FGUIR 和 PyUIE 算法虽然在纹理细节恢复方面表现尚可, 但在处理复杂场景 (如 Image4、Image5 和 Image6) 时, 容易引发色偏及局部过曝现象。UIEAnythingv2 算法的色彩恢复效果欠佳, 且生成图像的细节模糊。相比之下, 本文算法能够去

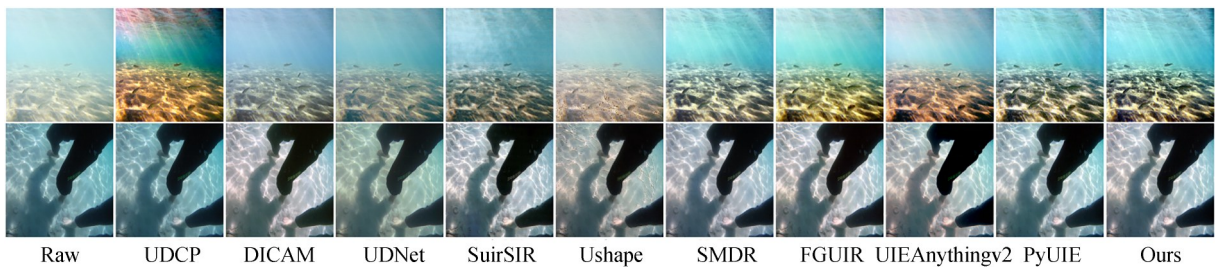


图 4 在 Test-UI90 数据集上的对比结果

Fig. 4 Comparison results on the Test-UI90 dataset

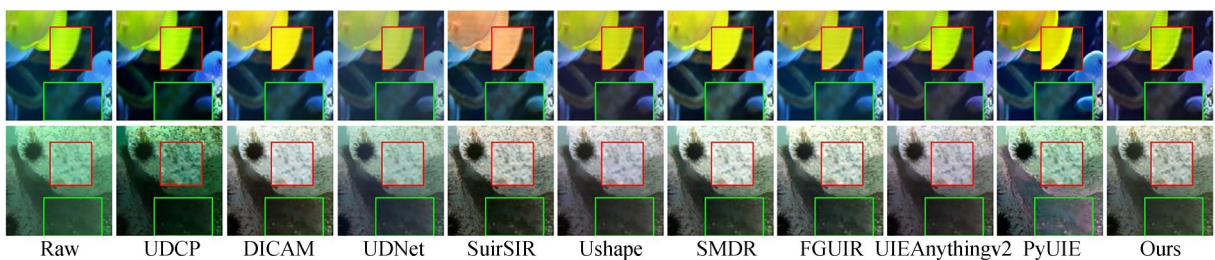


图 5 在 Test-LS856 数据集上的对比结果

Fig. 5 Comparison results on the Test-LS856 dataset

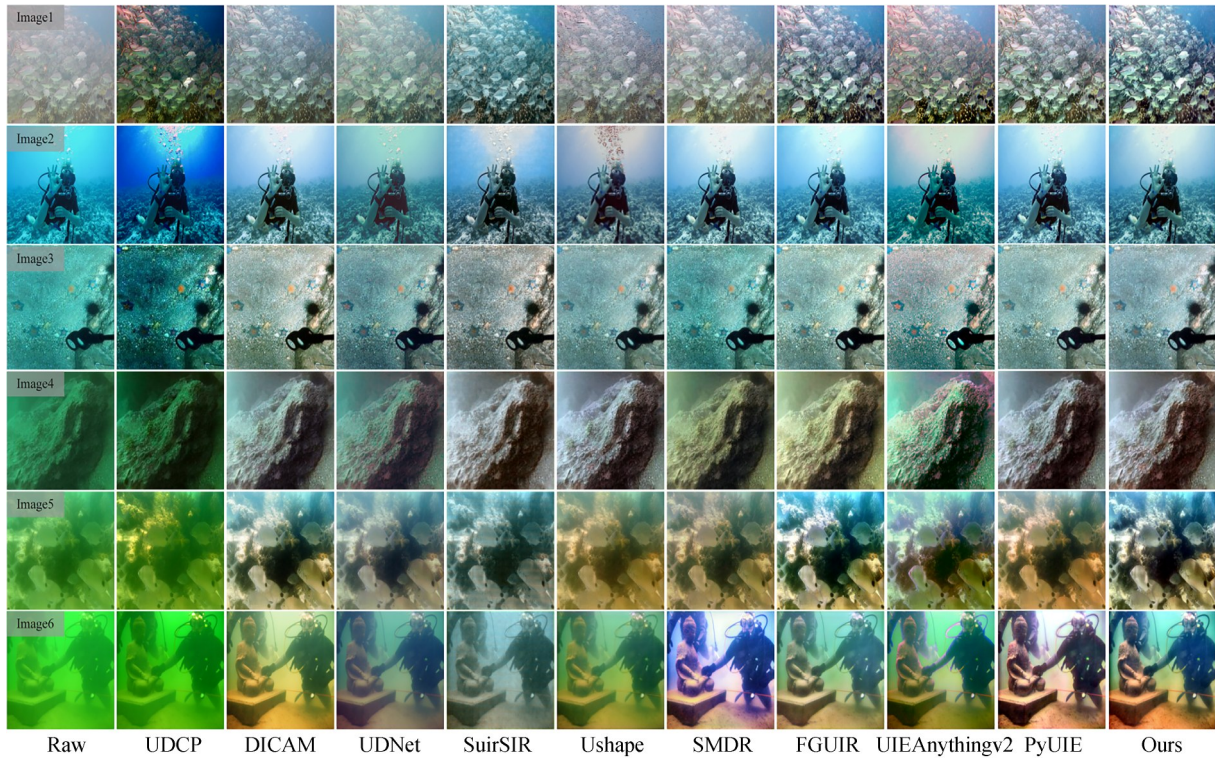


图6 在三个无参考数据集上的对比结果

Fig. 6 Comparison results on the three no-reference datasets

除类雾效应,在恢复丰富纹理细节的同时实现了自然平滑的色彩过渡(彩图见期刊电子版)。

### 3.3.2 定量分析

本节采用粗体和下划线分别标注各指标的最优值与次优值,以直观展示不同方法的性能。

表1汇总了10种方法在 Test-UI90 和 Test-

LS856数据集上的全参考指标(PSNR、SSIM和MSE)评估结果。数据显示,本文模型在上述三种指标上均取得了最高成绩。表2展示了各方法在 Test-C60, Test-UCCS 和 Test-U45数据集上的无参考指标(UIQM, UCIQE, NIQE 和 PIQE)评估结果。由表可知,本文方法在三个测试集上均获得了最高的 UCIQE 分数,表明本文算法对

表1 在 Test-UI90 和 Test-LS856数据集上的实验结果

Tab. 1 Experimental results on Test-UI90 and Test-LS856 datasets

Methods	Params/ M	FLOPs/ G	FPS/ ( $\text{f}\cdot\text{s}^{-1}$ )	PGM/ MB	Test-UI90			Test-LS856		
					PSNR	SSIM	MSE	PSNR	SSIM	MSE
UDCP	—	—	3.8	<b>12.89</b>	13.38	0.646	0.0549	14.80	0.643	0.0454
DICAM	<b>0.94</b>	53.14	27.5	1189.11	24.34	0.918	0.0058	20.84	0.818	0.0121
UDNet	<u>1.40</u>	30.15	<b>78.9</b>	203.55	19.89	0.832	0.0133	19.71	0.789	0.0132
SuirSIR	9.21	219.06	4.5	1816.05	22.31	0.878	0.0077	20.82	0.801	0.0107
Ushape	65.60	66.20	37.2	211.99	20.83	0.804	0.0115	24.52	0.827	0.0049
SMDR	12.25	48.32	39.5	177.07	25.69	0.939	0.0041	20.43	0.813	0.0123
FGUR	1.91	<u>17.84</u>	21.4	124.48	27.11	0.945	0.0027	20.95	0.809	0.0121
UIEAnythingv2	24.18	160.66	6.3	<u>119.49</u>	18.99	0.823	0.0167	19.00	0.754	0.0146
PyUIE	4.45	<b>12.76</b>	<u>73.2</u>	209.25	<u>27.91</u>	<u>0.952</u>	<u>0.0024</u>	<u>25.23</u>	<u>0.853</u>	<u>0.0040</u>
Ours	9.34	32.69	27.6	262.26	<b>28.32</b>	<b>0.964</b>	<b>0.0022</b>	<b>28.89</b>	<b>0.912</b>	<b>0.0022</b>

表 2 在 Test-C60, Test-UCCS 和 Test-U45 数据集上的实验结果

Tab. 2 Experimental results on Test-C60, Test-UCCS, and Test-U45 datasets

Methods	Test-C60				Test-UCCS				Test-U45			
	UIQM	UCIQE	NIQE	PIQE	UIQM	UCIQE	NIQE	PIQE	UIQM	UCIQE	NIQE	PIQE
UDCP	1.475	0.428 3	39.16	17.79	2.144	0.396 2	44.30	20.80	2.197	0.477 8	46.67	5.35
DICAM	2.480	0.533 8	37.98	17.53	2.905	0.514 0	44.36	19.36	<b>3.210</b>	0.554 7	47.50	4.41
UDNet	2.494	0.485 8	38.98	16.12	2.954	0.496 7	46.35	19.23	<u>3.154</u>	0.501 0	47.34	5.11
SuirSIR	<b>2.991</b>	0.518 7	45.05	<b>9.19</b>	<b>3.037</b>	0.503 4	52.93	17.55	3.069	0.543 0	50.08	<u>3.76</u>
Ushape	2.563	0.483 7	40.79	14.78	<u>2.990</u>	0.507 5	51.54	<u>11.04</u>	3.008	0.508 7	43.53	8.64
SMDR	2.539	0.506 5	37.36	16.63	2.914	0.496 6	<u>44.04</u>	20.75	3.087	0.549 1	43.86	4.29
FGUIR	2.687	0.535 5	38.16	<u>12.19</u>	2.943	<u>0.536 5</u>	45.34	17.78	3.012	<u>0.567 0</u>	43.46	<b>3.70</b>
UIEAnythingv2	1.857	<u>0.536 2</u>	<b>35.97</b>	15.22	2.392	0.532 6	45.86	<b>10.67</b>	2.356	0.543 6	<b>41.63</b>	5.92
PyUIE	2.552	0.521 6	38.79	15.33	2.891	0.514 9	44.40	19.27	2.968	0.558 0	43.06	4.16
Ours	<u>2.965</u>	<b>0.541 1</b>	<u>36.61</u>	13.45	2.978	<b>0.537 6</b>	<b>44.01</b>	12.83	3.027	<b>0.568 2</b>	<u>42.69</u>	4.10

图像色彩和细节恢复得更为自然。尽管在 UIQM, NIQE 和 PIQE 指标上略低于其最佳对手,但结合图 4~图 6 的可视化结果可知,本文方法在细节保真度和色彩均衡性上的综合表现更为稳健。

上述定量与定性实验结果均表明,本文方法在多种数据集和复杂场景下具有强泛化能力,在

应对低对比度、细节模糊及色偏等退化问题时展现出优异的鲁棒性。

### 3.4 消融实验

为验证网络中各模块及其内部组件的有效性 with 合理性,本文采用 Test-UI90 测试集进行消融实验。其中 Full model 为本文所提的完整模型,定量测评结果如表 3 所示,定性结果如图 7 所示。

表 3 消融实验

Tab. 3 Ablation experiment

Methods	PSNR	SSIM	MSE	Methods	PSNR	SSIM	MSE
w/o PConv	27.14	0.953	0.003 1	w/o CRA	27.84	0.957	<u>0.002 4</u>
w/o DyTConv	27.89	0.954	0.002 9	w/o PSUM	27.22	<u>0.958</u>	0.003 0
w/o Attention	25.40	0.941	0.004 1	w/o RER	27.48	0.948	0.003 3
w/o ADFCM	25.29	0.946	0.004 4	BPSUM	24.20	0.930	0.005 6
w/o FRM	27.14	0.953	0.003 0	CPSUM	26.08	0.948	0.003 7
w/o DOTM	24.60	0.929	0.006 1	w/o LP	<u>28.17</u>	0.957	0.002 5
w/o DCAM	27.03	<u>0.958</u>	0.002 8	Full model	<b>28.32</b>	<b>0.964</b>	<b>0.002 2</b>
w/o DConv	24.84	0.938	0.004 7				

#### 3.4.1 动态调优 Transformer 模块有效性分析

为验证动态调优 Transformer 模块(DOTM)的有效性,本文对其内部的自适应动态特征校准机制(ADFCM)、特征细化模块(FRM)等关键组件进行了消融分析。

本文针对 ADFCM 设计了四组消融实验,分别为移除风车形卷积的 w/o PConv、移除动态响应调制单元的 w/o DyTConv、移除卷积注意力机制的 w/o Attention 和移除 ADFCM 的 w/o ADF-

CM。定量结果如表 3 所示,与完整模型相比,移除上述任何一个组件都会导致网络的性能下降。上述实验的定性结果如图 7 所示,w/o PConv 和 w/o DyTConv 的生成结果均出现了各向异性细节丢失和对比度下降的问题。w/o Attention 和 w/o ADFCM 的生成图像出现了明显的黄绿色偏,且海葵触手等边缘区域细节丢失严重(彩图见期刊电子版)。上述实验结果证明 ADFCM 通过融合风车形卷积、动态响应调制单元和卷积注意

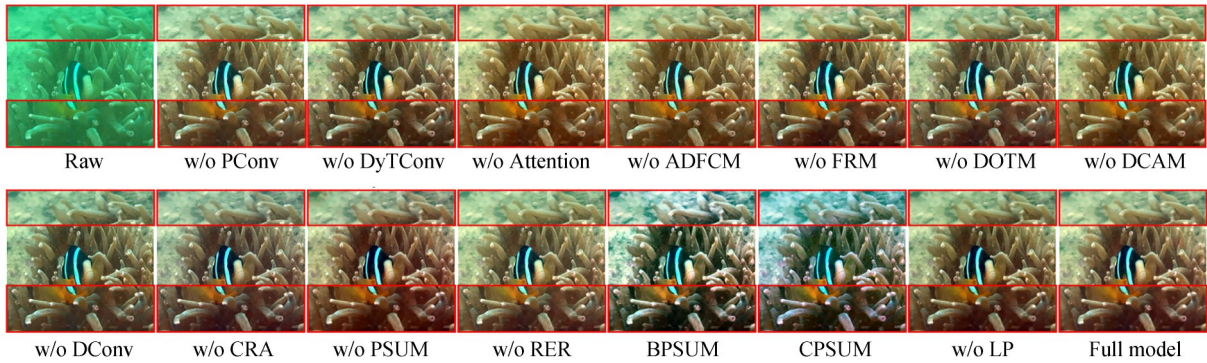


图 7 消融实验的可视化结果

Fig. 7 Visual demonstration of the ablation experiment

力机制,有效实现了上下文细节增强与色彩恢复。

移除特征细化模块后,w/o FRM对图像明暗交界处的纹理恢复能力和去噪能力下降,并出现局部细节丢失现象。上述结果证明FRM能够有效地聚合局部上下文信息并滤除冗余特征。

最后,移除整个DOTM模块后,w/o DOTM的PSNR降至24.60,且生成图像存在严重的细节丢失和绿色色偏,证明DOTM模块能够实现图像空间上下文细节的增强和全局色彩的恢复。

#### 3.4.2 空洞卷积聚合模块有效性分析

DCAM模块旨在解决水下图像的尺度依赖性模糊问题。为验证其有效性与合理性,本文针对DCAM设计了三组消融实验,分别为移除多尺度空洞卷积的w/o DConv(只保留 $d=1$ 的CBR分支)、移除通道混洗与残差注意力机制的w/o CRA和移除整个DCAM模块的w/o DCAM。

如表3所示,w/o DConv的PSNR降至24.84,MSE增至0.0047。结合如图7的可视化结果可见,w/o DConv的生成图像出现了明显的局部细节模糊与边缘过度平滑现象,验证了单一感受野无法处理水下复杂的散射效应,导致整体视觉效果严重缺乏层次感。当仅保留多尺度空洞卷积而移除通道混洗与残差注意力机制时,w/o CRA生成图像出现结构细节丢失、蓝色色偏以及色彩过渡不自然等问题。

此外,在移除整个DCAM模块后,w/o DCAM生成的图像存在整体色彩偏黄、小丑鱼体表的色彩过渡不自然等问题,导致图像的语义表

达能力下降。上述三组实验结果充分证明了DCAM模块及其内部组件设计的合理性与必要性。

#### 3.4.3 渐进式语义引导上采样模块有效性分析

为验证所提PSUM的级联式梯度保持策略能够实现空间解码特征的结构约束与纹理细化,设计了四组消融实验:采用并联架构的PSUM模型(BPSUM)、采用串联架构的PSUM模型(CPSUM)、移除残差边缘重建单元的w/o RER和移除整个PSUM模块的w/o PSUM。其中,w/o PSUM中采用双线性插值法替换整个PSUM模块;BPSUM和CPSUM的具体结构如图8所示。

由表3和图7可知,BPSUM和CPSUM的生成图像均出现了严重的色偏与高频细节丢失问题。而采用本文设计的Full model在各项评价指标上均取得最优表现,其生成图像在视觉上更接近真实的水下场景,证明级联式梯度补偿策略能够实现深层语义重构与浅层细节优化的动态平衡。

如图7所示,w/o RER生成的图像中,礁石与海葵触手等区域存在边缘锐度下降问题,证明RER能够有效保留微小的纹理细节,有效补偿了U型网络在分辨率恢复过程中的高频信息损失。

此外,为验证在联合损失函数中引入拉普拉斯算子的合理性,设计了w/o LP消融实验。其可视化结果如图7所示,w/o LP的生成图像存在明显的细节丢失现象。相比之下,完整模型生成的图像明暗过渡更自然平滑,有效还原了图像的纹理细节与边缘锐度。

上述实验结果表明,网络各模块的协同作用在颜色校正、对比度提升和细节增强方面展现出

优越性能,进一步验证了本文所提网络整体架构的有效性。

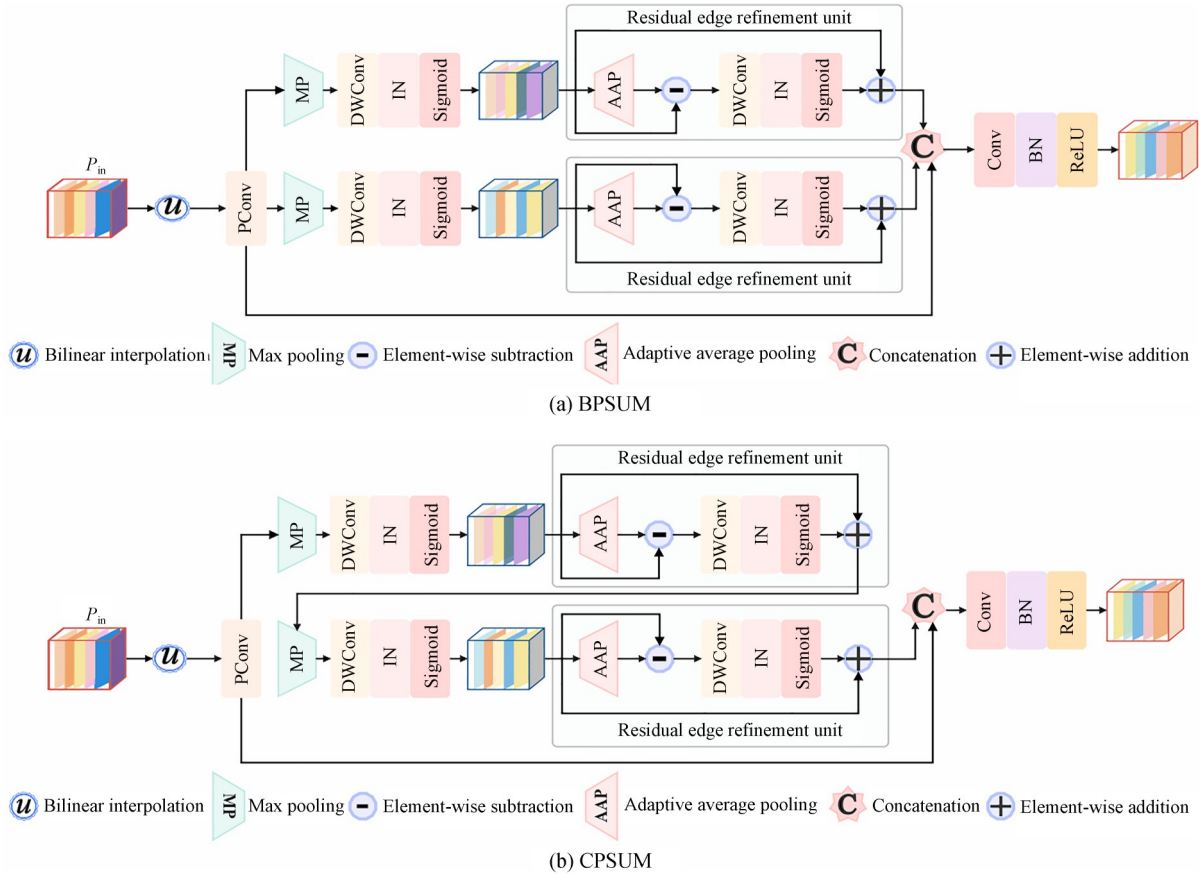


图 8 BPSUM 模块和 CPSUM 模块  
Fig. 8 BPSUM and CPSUM modules

### 3.5 计算复杂度与算力资源需求分析

本节从模型参数量(Params)、浮点运算数(FLOPs)、每秒传输帧数(FPS)和峰值显存占用(Peak GPU Memory, PGM)四个维度,对 10 种方法的参数规模、计算复杂度和算力资源需求进行了详细分析。上述四种指标均在第 3.1 节所述的实验环境(NVIDIA Tesla V100 32G GPU,输入分辨率为  $256 \times 256$ )下测得,确保分析结果与实际部署场景具有一致的可参照性。实验结果如表 1 所示,在参数规模与计算复杂度方面,本文模型的 Params 为 9.34 M, FLOPs 为 32.69 G。相较于 Ushape 和 UIEAnythingv2 等大参数量模型,本文方法在保持更优特征表达能力的同时,实现了计算复杂度与增强性能之间良好平衡。在

推理效率与显存消耗方面,本文方法在 Tesla V100(32 G)GPU 上达到了 27.6 FPS 的推理速度,PGM 仅为 262.26 MB,其推理速度明显优于 DICAM 和 SuirSIR 等高显存消耗模型。综合第 3.3 节的定量与定性分析可知,本文方法在有效控制模型规模与显存占用的前提下,在推理效率和增强图像质量之间取得了较好的平衡,充分验证了所提网络的高效性。

### 3.6 多层次视觉特征评估

为验证本文方法在不同层级视觉任务中的特征恢复能力,本文从测试集中随机选取 4 幅代表性图像,依次进行 SIFT 关键点匹配、Canny 边缘检测和显著目标检测(采用加载官方预训练权重的 U2-Net<sup>[47]</sup>)对比测试。上述实验旨在评估本

文模型的局部特征提取能力、结构边界感知能力与全局语义理解能力,并验证底层特征的有效恢

复对高层语义理解的促进作用,可视化结果如图 9 和图 10 所示。

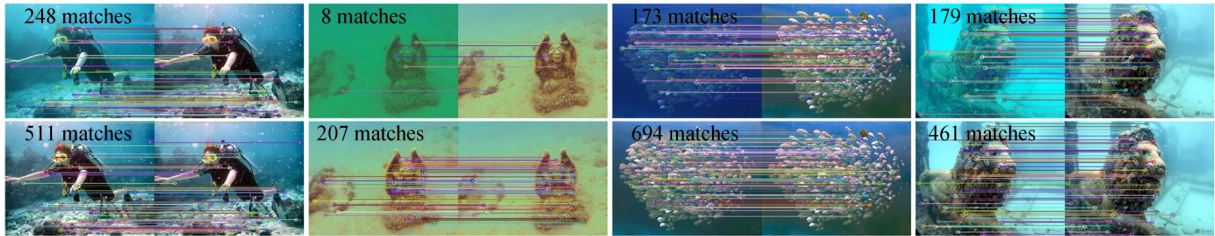
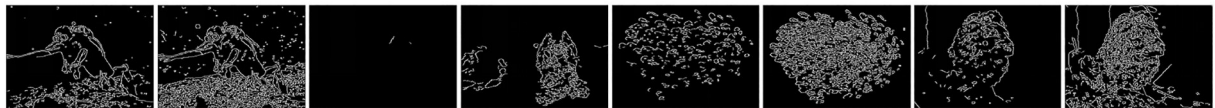


图 9 关键点匹配任务的可视化结果

Fig. 9 Visual demonstration of the key-point matching task



(a) 原始图像与增强后的图像  
(a) Original and enhanced images



(b) 边缘检测结果  
(b) Canny edge detection results



(c) 显著目标检测结果  
(c) Salient object detection results

图 10 原始图像与增强图像在边缘检测及显著目标检测任务上的可视化结果

Fig. 10 Visual demonstration of original and enhanced images on edge detection and salient object detection tasks

(1) 局部特征提取:结果如图 9 所示,增强后图像与真值图像间的关键点匹配数量分别增加了 263, 199, 521 和 282,直观证明了本文方法能够有效消除水下退化干扰,重建受损的高频纹理特征。

(2) 结构边界感知:在准确恢复局部特征的基础上,结构边界的清晰度直接决定了视觉模型对目标轮廓的感知质量。如图 10(b)所示,增强后图像的边缘结构更加清晰和连贯,有效抑制了水下模糊带来的边缘缺失现象,为后续目标定位提供了可靠的结构基础。

(3) 全局语义理解:如图 10(c)所示,原始图像受色偏与散射影响,前景与背景融合严重,导

致显著图边界模糊甚至出现目标漏检。得益于局部细节与边界信息的恢复,增强后的图像提高了前景与背景的分度,有助于更精准地定位显著目标,并生成边界清晰、轮廓完整的预测结果。

上述实验结果表明,本文方法实现了局部细节、结构轮廓到全局语义三个层级上的优化,在提升高级视觉任务性能方面具有实用价值。

## 4 结 论

本文提出了一种融合动态感知与渐进细化的水下图像增强网络,旨在应对水下图像中存在的色彩失真、对比度低和尺度依赖性模糊等问

题。首先,在编码器中提出了动态调优 Transformer 模块。该模块通过整合自适应动态特征校准机制与特征细化模块,动态感知水下空间的非均匀退化,实现空间上下文细节的增强和全局色彩的恢复。然后,针对尺度依赖性模糊问题设计了空洞卷积聚合模块,利用并行多分支空洞卷积提取多尺度上下文特征,并引入通道混洗与残差注意力机制重建模糊区域的结构细节,确保图像中亮度与色彩的自然过渡。最后,在解码器中构建了渐进式语义引导下采样模块,采用双线性插值和级联式梯度保持策略,实现对空间解码特征的结构约束与纹理细化。实验结果表明,本文方法在 UIEB 和 LSUI 数据集上的 PSNR、SSIM 及 MSE 指标均达到最优。相较于次优方法,本文方法在 UIEB 上的 PSNR 和 SSIM 分别提升了

1.47% 和 1.26%,MSE 降低了 8.33%;在 LSUI 上的 PSNR 和 SSIM 分别提升了 14.51% 和 6.92%,MSE 降低了 45.00%。此外,在 UCCS, U45 和 C60 无参考数据集上,所提方法在 UCIQE 指标上亦取得了最优结果。上述结果证明本文方法在校正色偏、提升对比度和恢复细节方面的优越性,可为高级视觉任务提供高质量的图像输入。

#### 作者贡献声明:

孙士明:项目管理,提供资源,写作指导;  
张成雯:模型设计,初稿写作,实验验证;  
李广峰:数据分析,实验验证;  
高绍姝:获取资助,写作指导;  
蒋铭皓:调查研究,数据分析。

#### 参考文献:

- [1] SHANG J S, LI Y, YUAN J Y, *et al.* Collaboratively optimized multi-level adaptive image enhancement for underwater ship inspection[J]. *Expert Systems with Applications*, 2026, 309: 131073.
- [2] DONG N, LIU S F, IP A W H, *et al.* End-to-end autonomous underwater vehicle path following control method based on improved soft actor - critic for deep space exploration[J]. *Journal of Industrial Information Integration*, 2025, 45: 100792.
- [3] 陶洋, 龚霖霆, 周立群. 基于 Retinex 的可变注意力低照度水下图像增强[J]. *液晶与显示*, 2025, 40(3): 481-492.  
TAO Y, GONG J T, ZHOU L Q. Variable attention low illumination underwater image enhancement based on Retinex [J]. *Chinese Journal of Liquid Crystals and Displays*, 2025, 40(3): 481-492. (in Chinese)
- [4] 郭梨, 黄智奇, 何润丰, 等. 一种基于改进 DCGAN 的水下图像增强方法[J]. *光学学报*, 2026, 46(2): 0210005.  
GUO L, HUANG Z Q, HE R F, *et al.* Underwater image enhancement method based on improved DCGAN [J]. *Acta Optica Sinica*, 2026, 46(2): 0210005. (in Chinese)
- [5] 彭子芊, 程志远, 胡佳豪. 基于色彩校正与注意力机制的水下图像增强[J]. *激光与光电子学进展*, 2025, 62(14): 63-71.  
PENG Z Q, CHENG Z Y, HU J H. Underwater image enhancement algorithm using color correction and attention mechanism [J]. *Laser & Optoelectronics Progress*, 2025, 62(14): 63-71. (in Chinese)
- [6] 商家硕, 李颖, 袁靖懿, 等. 均值强度差先验驱动的水下图像复原[J]. *光学精密工程*, 2025, 33(18): 2962-2979.  
SHANG J S, LI Y, YUAN J Y, *et al.* Underwater image restoration driven by the mean intensity difference prior [J]. *Opt. Precision Eng.*, 2025, 33(18): 2962-2979. (in Chinese)
- [7] CAO T Y, YU Z B, ZHENG B. DNnet: a lightweight network for real-time 4K underwater image enhancement using dynamic range and average normalization [J]. *Expert Systems with Applications*, 2025, 270: 126561.
- [8] 陶洋, 谭浩, 周立群. ULCF-Net: 跨尺度结构与色彩融合的水下低照度图像增强算法[J]. *激光与光电子学进展*, 2025, 62(2): 420-430.  
TAO Y, TAN H, ZHOU L Q. ULCF-net: an underwater low-illumination image enhancement algorithm based on a cross-scale structure and color fusion [J]. *Laser & Optoelectronics Progress*, 2025, 62(2): 420-430. (in Chinese)

- [9] ZHANG W D, ZHUANG P X, SUN H H, *et al.* Underwater image enhancement *via* minimal color loss and locally adaptive contrast enhancement [J]. *IEEE Transactions on Image Processing*, 2022, 31: 3997-4010.
- [10] GÜRAKSIN G E, KÖSE U, DEPERLİOĞLU Ö. Underwater image enhancement based on contrast adjustment *via* differential evolution algorithm [C]. *2016 International Symposium on INnovations in Intelligent SysTems and Applications (INISTA)*. August 2-5, 2016, Sinaia, Romania. IEEE, 2016: 1-5.
- [11] SUBRAMANI B, VELUCHAMY M. Pixel intensity optimization and detail-preserving contextual contrast enhancement for underwater images[J]. *Optics & Laser Technology*, 2025, 180: 111464.
- [12] 麦仁贵, 王骥. 基于关键特征传递的KFT-GAN水下图像增强模型[J/OL]. *电子测量与仪器学报*, 2026: 1-11.  
MAI R G, WANG J. An Underwater Image Enhancement Model Based on Generative Adversarial Network with Key Feature Transfer [J/OL]. *Journal of Electronic Measurement and Instrumentation*, 2026: 1-11. (in Chinese)
- [13] 李文彪, 陶洋, 董源, 等. 基于颜色先验引导和注意力机制的水下图像增强[J]. *液晶与显示*, 2025, 40(8): 1163-1176.  
LI W B, TAO Y, DONG Y, *et al.* Underwater image enhancement based on color prior guidance and attention mechanism [J]. *Chinese Journal of Liquid Crystals and Displays*, 2025, 40(8): 1163-1176. (in Chinese)
- [14] KONG H, YUAN A M, CHEN Q, *et al.* Underwater engineering image enhancement *via* turbidity prior and multi-scale adaptive[J]. *Engineering Applications of Artificial Intelligence*, 2026, 166: 113685.
- [15] 程竹明, 李佳轩, 黄三傲, 等. 基于多分支残差注意力网络的水下图像增强[J]. *光学精密工程*, 2025, 33(7): 1141-1151.  
CHENG Z M, LI J X, HUANG S A, *et al.* Underwater image enhancement based on multi-branch residual attention network[J]. *Opt. Precision Eng.*, 2025, 33(7): 1141-1151. (in Chinese)
- [16] 郑浩君, 王振, 张佳鹏, 等. 基于混合注意力模块改进StarGAN的水下图像增强[J]. *南方水产科学*, 2025, 21(1): 185-196.  
ZHENG H J, WANG Z, ZHANG J P, *et al.* Enhancement of underwater images on improved StarGAN by mixed attention module[J]. *South China Fisheries Science*, 2025, 21(1): 185-196. (in Chinese)
- [17] QIN N W, WU J J, LU Q H, *et al.* MFF-Trans: Multi-feature fusion transformer for underwater image enhancement [J]. *Digital Signal Processing*, 2025, 165: 105341.
- [18] 姚婷婷, 冯子豪, 肇恒鑫. 融合域迁移和注意力机制的水下图像增强[J]. *光学精密工程*, 2025, 33(2): 298-310.  
YAO T T, FENG Z H, ZHAO H X. Underwater image enhancement by integrating domain transfer and attention mechanisms [J]. *Opt. Precision Eng.*, 2025, 33(2): 298-310. (in Chinese)
- [19] CHEN W W, LEI Y T, LUO S H, *et al.* UWFormer: underwater image enhancement *via* a semi-supervised multi-scale transformer[C]. *2024 International Joint Conference on Neural Networks (IJCNN)*. June 30 - July 5, 2024, Yokohama, Japan. IEEE, 2024: 1-8.
- [20] SHANG J S, LI Y, XING H, *et al.* LGT: Luminance-guided transformer-based multi-feature fusion network for underwater image enhancement [J]. *Information Fusion*, 2025, 118: 102977.
- [21] LEE H, KANG S. Two-stage underwater image enhancement using domain adaptation and interlacing transformer[J]. *Displays*, 2025, 87: 102980.
- [22] KHAN M R, NEGI A, KULKARNI A, *et al.* Phaseformer: phase-based attention mechanism for underwater image restoration and beyond [C]. *2025 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*. February 26 - March 6, 2025, Tucson, AZ, USA. IEEE, 2025: 9618-9629.
- [23] YANG J N, LIU S L, WU J J, *et al.* Pinwheel-shaped convolution and scale-based dynamic loss for infrared small target detection[J]. *Proceedings of the AAAI Conference on Artificial Intelligence*, 2025, 39(9): 9202-9210.

- [24] ZHU J C, CHEN X L, HE K M, *et al.* Transformers without normalization [C]. 2025 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. June 10-17, 2025, Nashville, TN, USA. IEEE, 2025: 14901-14911.
- [25] ZAMIR S W, ARORA A, KHAN S, *et al.* Restormer: efficient transformer for high-resolution image restoration [C]. 2022 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. June 18-24, 2022, New Orleans, LA, USA. IEEE, 2022: 5718-5729.
- [26] CHEN L C, PAPANDREOU G, SCHROFF F, *et al.* Rethinking atrous convolution for semantic image segmentation [EB/OL]. 2017: *arXiv*: 1706.05587. <https://arxiv.org/abs/1706.05587>
- [27] LIU S B, FAN H J, LIN S, *et al.* Adaptive learning attention network for underwater image enhancement [J]. *IEEE Robotics and Automation Letters*, 2022, 7(2): 5326-5333.
- [28] PENG L T, ZHU C L, BIAN L H. U-shape transformer for underwater image enhancement [J]. *IEEE Transactions on Image Processing*, 2023, 32: 3066-3079.
- [29] LI C Y, GUO C L, REN W Q, *et al.* An underwater image enhancement benchmark dataset and beyond [J]. *IEEE Transactions on Image Processing*, 2020, 29: 4376-4389.
- [30] LIU R S, FAN X, ZHU M, *et al.* Real-world underwater enhancement: challenges, benchmarks, and solutions under natural light [J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2020, 30(12): 4861-4875.
- [31] LI H Y, LI J J, WANG W. A fusion adversarial underwater image enhancement network with a public test dataset [J]. *ArXiv: Image and Video Processing*. 2019
- [32] KORHONEN J, YOU J Y. Peak Signal-to-noise ratio revisited: is simple beautiful? [C]. 2012 *Fourth International Workshop on Quality of Multimedia Experience*. July 5-7, 2012, Melbourne, VIC, Australia. IEEE, 2012: 37-38.
- [33] WANG Z, BOVIK A C, SHEIKH H R, *et al.* Image quality assessment: from error visibility to structural similarity [J]. *IEEE Transactions on Image Processing*, 2004, 13(4): 600-612.
- [34] MARMOLIN H. Subjective MSE measures [J]. *IEEE Transactions on Systems, Man, and Cybernetics*, 1986, 16(3): 486-489.
- [35] PANETTA K, GAO C, AGAIAN S. Human-visual-system-inspired underwater image quality measures [J]. *IEEE Journal of Oceanic Engineering*, 2016, 41(3): 541-551.
- [36] YANG M, SOWMYA A. An underwater color image quality evaluation metric [J]. *IEEE Transactions on Image Processing*, 2015, 24(12): 6062-6071.
- [37] MITTAL A, SOUNDARARAJAN R, BOVIK A C. Making a "completely blind" image quality analyzer [J]. *IEEE Signal Processing Letters*, 2013, 20(3): 209-212.
- [38] VENKATANATH N, PRANEETH D, BH M C, *et al.* Blind image quality evaluation using perception based features [C]. 2015 *Twenty First National Conference on Communications (NCC)*. February 27-March 1, 2015, Mumbai, India. IEEE, 2015: 1-6.
- [39] DREWS P JR, NASCIMENTO EDO, MORAES F, *et al.* Transmission estimation in underwater single images [C]. 2013 *IEEE International Conference on Computer Vision Workshops*. December 2-8, 2013, Sydney, NSW, Australia. IEEE, 2014: 825-830.
- [40] TOLIE H F, REN J C, ELYAN E. DICAM: Deep inception and channel-wise attention modules for underwater image enhancement [J]. *Neurocomputing*, 2024, 584: 127585.
- [41] SALEH A, SHEAVES M, JERRY D, *et al.* Adaptive deep learning framework for robust unsupervised underwater image enhancement [J]. *Expert Systems with Applications*, 2025, 268: 126314.
- [42] LIN S, ZHANG R H, LI Y P. Sinking into water without seeing water: Semi-supervised underwater image enhancement based on simulated image re-degradation process [J]. *Optics & Laser Technology*, 2025, 184: 112545.
- [43] ZHANG D H, ZHOU J C, GUO C L, *et al.* Synergistic multiscale detail refinement *via* intrinsic su-

- pervision for underwater image enhancement [C]. *Proceedings of the AAAI Conference on Artificial Intelligence*, 2024, 38(7): 7033-7041.
- [44] ZHANG L, ZHANG X D, CHU M H, *et al.* A frequency-guided underwater image restoration network via progressive contrastive learning[J]. *Displays*, 2025, 90: 103102.
- [45] JIANG W C, TAN Y Q, QIU Z X, *et al.* PyUIE: A coarse-to-fine deep pyramid network for underwater image enhancement [J]. *IEEE Transactions on Multimedia*, 2026: 1-14.
- [46] SHAO J X, ZHANG H S, MIAO J M. Robust-SeaThru: a physics-guided underwater image restoration model with percentile-based backscatter estimation[J]. *Applied Soft Computing*, 2026, 192: 114763.
- [47] QIN X B, ZHANG Z C, HUANG C Y, *et al.* U2-Net: Going deeper with nested U-structure for salient object detection [J]. *Pattern Recognition*, 2020, 106: 107404.

## 作者简介:



孙士明(1970-),男,山东聊城人,副教授,硕士生导师,2003年于中国石油大学(华东)获得硕士学位,现就职于中国石油大学(华东)计算机与通信工程学院,主要从事图像处理、嵌入式系统方向的教学和研究工作。E-mail: sunshim@upc.edu.cn

## 通讯作者:



张成雯(1999-),女,河北平泉人,硕士研究生,2024年于沈阳师范大学获得网络工程专业学士学位,现为中国石油大学(华东)青岛软件学院、计算机科学与技术学院硕士研究生,主要从事图像增强方面的研究。E-mail: z24070011@s.upc.edu.cn