

文章编号 1004-924X(2024)10-1582-13

综合多尺度信息和注意力机制的水下图像增强

夏晓华^{1*}, 钟预全¹, 胡 鹏¹, 姚运仕^{1,2}, 耿继光², 张良奇²

(1. 长安大学 道路施工技术装备教育部重点实验室, 陕西 西安 710064;

2. 河南万里交通科技集团股份有限公司, 河南 许昌 461000)

摘要:针对水下图像由于水的散射和吸收而存在颜色失真和细节丢失等问题,提出了一种综合多尺度信息和注意力机制的生成对抗网络模型来增强水下图像。首先,为了充分利用和增强图像的局部信息和全局信息,使用局部编码器和全局编码器分别提取图像的局部特征和全局特征,并互相融合以实现互补性。接着,设计多尺度混合卷积来捕捉多尺度信息,增加网络对不同尺度特征的适应性。然后,利用注意力机制增加特征提取的准确性,加强网络对高价值特征的关注度。最后,重复使用多尺度混合卷积和注意力机制进一步细化特征后,逐步上采样得到增强图像。与六种经典和最新的方法相比,提出的模型不仅在主观评价中取得了最好的视觉感受,而且在整个测试集上,峰值信噪比(PSNR)、结构相似指数(SSIM)、水下图像质量指标(UIQM)和自然图像质量(NIQE)四种客观评价指标分别取得了22.499,0.789,2.911和4.175的平均分数,均优于六种对比方法,较对比方法中的最优值分别提升0.353,0.002,0.025和0.307,证明提出的模型不仅能够矫正图像颜色失真,而且在恢复图像细节、增加图像对比度和清晰度等方面均有较好的表现,具有良好的应用前景。

关键词:水下图像增强;生成对抗网络;编码器;多尺度混合卷积;注意力机制

中图分类号:TP391 **文献标识码:**A **doi:**10.37188/OPE.20243210.1582

Underwater image enhancement synthesizing multi-scale information and attention mechanisms

XIA Xiaohua^{1*}, ZHONG Yuquan¹, HU Peng¹, YAO Yunshi^{1,2}, GENG Jiguang², ZHANG Liangqi²

(1. Key Laboratory of Road Construction Technology and Equipment, Ministry of Education, Chang'an University, Xi'an 710064, China;

2. Henan Wanli Transportation Technology Group Co. Ltd., Xuchang 461000, China)

* Corresponding author, E-mail: xhxia@chd.edu.cn

Abstract: Aiming at the problems of color distortion and detail loss in underwater images due to water scattering and absorption, a generative adversarial network model integrating multi-scale information and attention mechanism was proposed to enhance underwater images. Firstly, to fully exploit and enhance both local and global information of the image, local encoders and global encoders were employed to extract local and global features respectively, which were then fused to achieve complementarity. Next, a multi-scale hybrid convolution was designed to capture multi-scale information, increasing the network's adaptability to features at different scales. Subsequently, attention mechanisms were utilized to enhance

收稿日期:2023-12-14;修订日期:2024-02-20.

基金项目:国家自然科学基金(No. 61901056);秦创原引用高层次创新创业人才项目(No. QCYRCXM-2022-352)

the accuracy of feature extraction, emphasizing the focus on high-value features. Finally, by iteratively applying multi-scale hybrid convolution and attention mechanisms to refine features, the enhanced image was gradually up-sampled. Compared with the six classical and state-of-the-art methods, the proposed model not only achieved the best visual perception in subjective evaluations but also outperformed the six comparative methods on the entire test set in terms of four objective evaluation metrics peak signal-to-noise ratio (PSNR), structural similarity (SSIM), underwater image quality measurement (UIQM), and natural image quality evaluation (NIQE) with average scores of 22.499, 0.789, 2.911, and 4.175, respectively. The improvements over the best scores among the comparative methods are 0.353, 0.002, 0.025, and 0.307, respectively. These results indicate that the proposed model not only corrects image color distortion but also performs well in restoring image details, increasing image contrast, and enhancing clarity. Therefore, it shows promising prospects for practical applications in underwater image enhancement.

Key words: underwater image enhancement; generative adversarial network; encoder; multi-scale hybrid convolution; attention mechanism

1 引 言

21 世纪以来,海洋战略地位不断提升,海底资源探测、海洋生物多样化、海洋军事等领域的发展都离不开海洋这块基石^[1-3]。在此背景下,王明军等人^[4]通过研究水下三维重建,分析了温度差和盐度差对三维重建效果的影响,为水下探测提供了重要的参考价值。另外,图像作为获取信息快速而又高效的媒介之一,海洋信息的获取离不开海洋图像的支撑。然而,在水下复杂的环境中,光会与水中的颗粒产生反射和散射现象,在水中不同波长的光因为吸收作用产生不同程度的衰减,造成水下图像出现色偏、低对比度、细节模糊等现象,这严重影响了分类、识别、分割等后续高级计算机视觉任务的准确性。因此,有必要通过图像处理技术实现水下图像质量提升。

水下图像质量提升方法可分为基于传统方法的水下图像质量提升和基于深度学习方法的水下图像质量提升。在传统方法方面,GHANI 等人^[5]在直方图均衡化算法的基础上提出了综合全局和局部矫正的水下图像增强方法,取得了一定的效果。胡振宇^[6]等人将水下图像依次进行颜色矫正和对比度增强,较好地去除了色偏和色调偏暗等现象。张楠楠^[7]等人通过改进传统的大气散射模型,在此物理模型基础上结合灰度世界算法修复图像,避免了修复过程中色彩丢失的问题。基于传统方法的水下图像质量提升近年来虽然取得了一定的进展,但是一般只适用于特定环境,并且在

图像清晰度、色彩过渡等方面提升欠佳^[8]。

近年来,深度学习因其具有强大的特征学习能力,被广泛应用于各类视觉任务,并取得了极具竞争力的效果,相关研究者将其运用至水下图像质量提升任务并取得了一定的研究成果。Fabbri 等人^[9]利用 CycleGAN^[10]将高质量无失真的水下图像映射到失真的水下图像,从而得到成对的水下图像数据集,进而提出像素到像素的水下图像增强网络(UGAN),但该网络的泛化性较差。Islam 等人^[11]提出的 FUnIE-GAN 网络,通过对损失函数的多角度调整修改,在图像的色彩、细节轮廓、主体内容上复原水下图像的真实面貌。但是该模型的深度较浅,不能学习到图像的全部特征导致泛化性较差。Li 等人^[12]结合水下成像模型设计了一种由多颜色空间的编码器和透射率引导的解码器组成的结构网络,该网络增强的水下图像在颜色过渡上较为自然,但是面对低照度水下图像时效果较差。Ma 等人^[13]提出了一种基于小波的双流网络(UIE-WD),该网络使用离散的小波变换将图像分解,然后输入到双流网络中进行色彩矫正和细节增强。Fu 等人^[14]引入了一种概率网络学习退化水下图像的增强分布,将水下图像增强分解为分布估计和共识过程。米泽田等人^[15]级联多尺度原始图像和对应的特征图像,有效地利用不同尺度特征的稀疏性,增强细节保持,并通过构建损失函数解决细节恢复不均的问题。李云等人^[16]将传统的水下图像增强方法结合到深度学习中,融合来自三种

传统方法增强后的图像特征,并设计注意力学习模块来增强水下图像。Zhang 等人^[17]提出了一种具有多分辨率特征学习的分层注意力聚合网络来增强水下图像。上述基于深度学习的水下图像增强方法,虽然取得了一定程度的效果,提高了原始图像的质量,但由于水下环境的高度复杂性,增强后的图像在清晰度、对比度、细节显现等方面仍有待进一步提高。此外,单一尺度信息的输入可能会导致局部信息和全局信息的混淆。因此,有必要开展相关研究以更好地区分和处理图像中的局部信息和全局信息,以达到更优的增强效果。

针对水下图像存在色偏、低对比度、模糊等质量问题,本文在生成对抗网络的基础上,改进生成器和判别器的网络结构,使其更准确地学习水下图像到理想图像的映射关系。首先受 U-Net^[18]强大的特征学习能力启发,生成器采用类似 U-Net 的编码器-解码器对称结构。其次考虑到 U-Net 无法提取图像的全局特征^[19],在原有局部编码器的基础上增加了全局编码器,构

成了双分支编码结构,使网络在学习图像局部特征的同时,兼顾图像全局特征,达到统筹兼顾的目的;然后在编码器之后使用多尺度混合卷积和注意力机制,增加网络对不同尺度特征的适应性和重要特征提取的准确性,使高价值特征得到更好地利用,低价值特征及时摒弃,减少低价值特征对解码过程的干扰;最后引入了残差连接,其能够减少编码器和解码器之间内容的损失,提取更多的图像细节,并在一定程度上避免梯度消失的问题。实验结果表明本文模型在主观评价和客观评价上均取得了较好的结果。

2 本文模型

2.1 生成器网络结构

如图 1 所示,本文模型使用的生成器包括局部编码器、全局编码器、多尺度混合卷积模块、注意力模块和解码器。其中局部编码器和解码器呈对称结构,类似于文献[18]。由于图像在下采

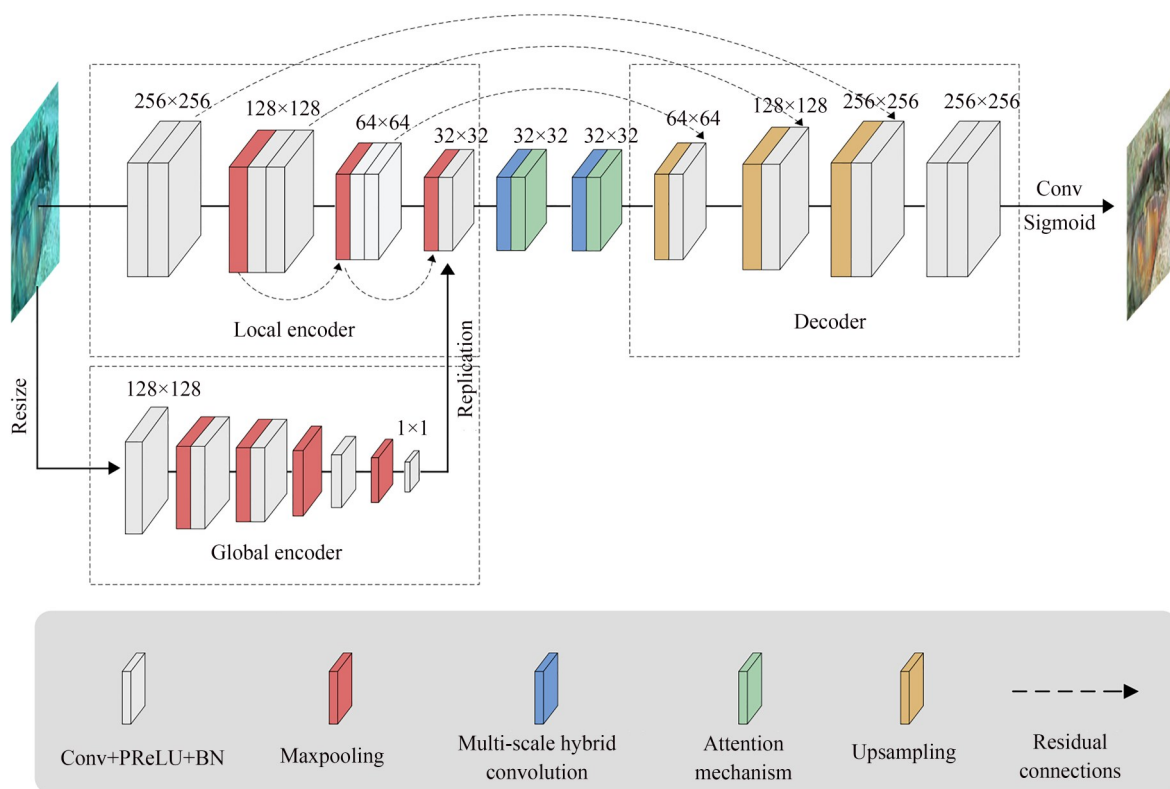


图 1 生成器网络结构

Fig. 1 Generator network structure

样过程中不可避免地存在特征损失,所以在局部编码器中以及局部编码器和解码器间添加残差连接来保证内容特征的完整性,减少图像细节损失,并在一定程度上避免梯度消失问题。局部编码器的输入图像为 256×256 大小的 RGB 图像,首先经过两层卷积单元提取低级语义信息,接着将输出进行下采样,图像大小缩小一半,再重复上述过程两次得到局部编码器的输出特征图。然后与全局编码器的特征图进行融合,经过多尺度混合卷积模块和注意力模块后输入解码器,依次进行上采样和卷积单元恢复图像原始大小和细节,提升图像内容表达。经过两层卷积单元进一步整合特征,加强网络表达能力,最后再进行一次卷积核大小为 4×4 ,步长为 1 的卷积操作,输出通道数为 3 的特征图,经 Sigmoid 操作后得到最终的增强图像。

2.1.1 全局编码器

由于局部编码器着眼于图像局部特征的提取,没有兼顾图像的全局特征,所以本文引入全局编码器,对图像的全局特征进行提取增强。全局编码器输入图像为 128×128 大小的 RGB 图像,依次通过卷积单元和下采样降低图像大小,浓缩全局信息。其中卷积单元卷积核的大小均为 3×3 ,前三层的步长均为 1,第四层和第五层的步长分别为 2 和 4,最终得到大小为 1×1 的特征图,与局部编码器的结果融合后作为下一模块的输入。

2.1.2 多尺度混合卷积模块

如图 2 所示,多尺度混合卷积模块的第一部分由四组不同卷积核大小的卷积层构成。由于较小的卷积核能够捕捉到局部细节信息,而较大的卷积核则可以捕捉到更广泛的上下文信息,并且该模块处理的特征图大小为 32×32 ,因此,本文选取卷积核大小分别为 1×1 、 3×3 、 5×5 、 7×7 ,步长均为 1 的卷积层。另外,不同尺度卷积核的组合形式对多尺度特征更加敏感,同时还能消除不同尺度的噪声,减少图像伪影和细节模糊现象。接着将卷积层输出的特征图拼接后采用参数修正线性单元(PReLU)进行修正,相比常规的修正线性单元(ReLU),PReLU 使网络的拟合能力更强。然后依次经过卷积

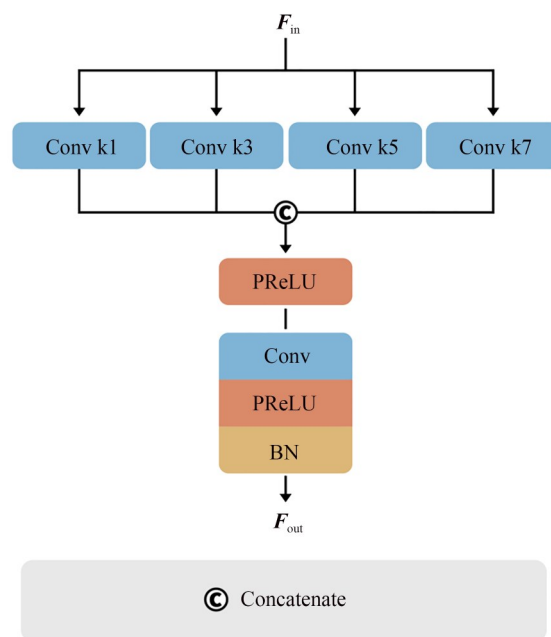


图 2 多尺度混合卷积结构

Fig. 2 Multi-scale hybrid convolution structure

层、PReLU 层和批归一化层(BN)得到内容更丰富的特征图。

2.1.3 注意力机制

注意力机制的核心目标是保留特征图中的高价值信息,并且剔除特征图中低价值或无用的信息。近几年来,基于该模块的通用性,众多学者将其应用在各个领域,并取得了令人满意的效果。本文使用 Convolutional Block Attention Module(CBAM)由通道注意力(CA)和空间注意力(SA)串联组成^[20],如图 3(a)所示。

通道注意力主要关注特征图哪个通道的信息是高价值的,其主要结构如图 3(b)所示,首先输入特征图通过全局平均池化和全局最大池化操作,接着送入 Dense1 压缩通道数,Dense2 扩充通道数,然后将得到的两幅权值图相加后经 Sigmoid 层操作与输入特征图相乘,得到通道注意力的结果。空间注意力主要关注特征图哪个位置的信息是高价值的,其主要结构如图 3(c)所示,首先将输入特征图在通道层面通过全局平均池化和全局最大池化操作,接着将两个分支进行拼接,然后通过卷积层压缩通道值至 1,经 Sigmoid 层操作后与输入特征图相乘得到空间注意力的结果。输入特征图通过 CBAM 得到输出特征

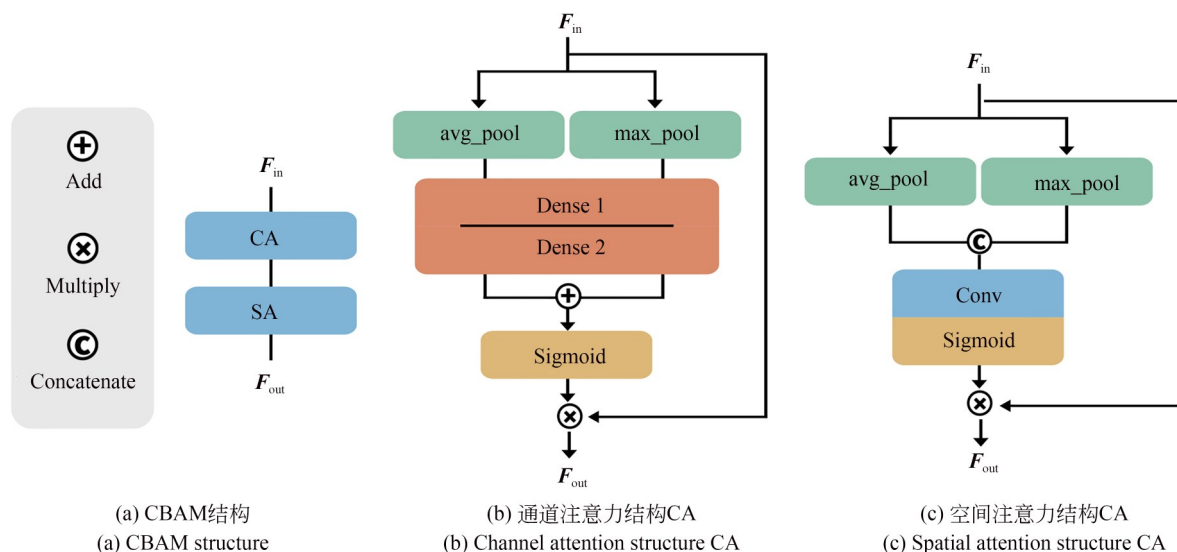


图3 注意力机制结构

Fig. 3 Attention mechanism structure

图,输出特征图经过空间层面和通道层面的提炼,更多的关键信息被捕捉增强,特征图得到细化。

2.2 判别器网络结构

判别器在生成对抗网络中也占据着非常重要的地位,其表现的优劣很大程度上影响到生成器网络的训练结果,本文采用马尔可夫判别器^[21]。该判别器假设每个像素块都是独立的,各像素块之间相互不受影响,每个像素块只根据自身来判断真假,这一假设对于有效捕获高频特征如局部纹理和样式非常重要^[22]。如图4所示,该网络首先使用四层卷积操作扩充通道数和降低图像大小,其中每层卷积核大小均为 3×3 ,步长均为2,除第一层外,其他三层后均连接BN层。最后再通过一层卷积核大小为4,步长为2的卷积层压缩通道,得到 $16 \times 16 \times 1$ 的概率图。

2.3 损失函数

考虑到使用多个损失函数的线性组合能保持图像细节特征,提升图像增强效果^[23],本文采用多种损失函数的线性组合对模型进行训练。为减少输出图像和目标图像之间的差异,采用输出图像和目标图像的 L_1 距离作为损失函数之一优化网络,其定义如式(1)所示:

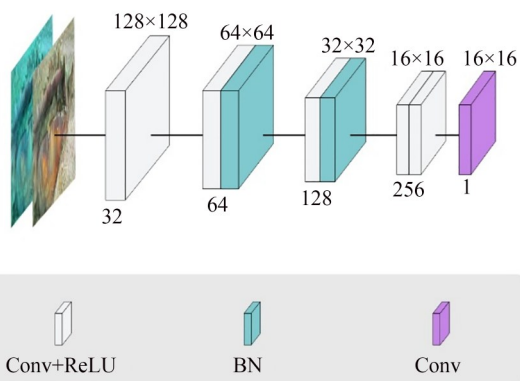


图4 判别器网络结构

Fig. 4 Discriminator network structure

$$L_{L_1} = \frac{1}{n} \sum_{i=1}^n |Y_i - G(X_i)|, \quad (1)$$

式中: n 代表训练样本集的总数; G 代表生成器; X_i 代表模型输入的第 i 张图像; Y_i 代表与 X_i 相对应的目标图像。

另外,将本文网络的输出图像和目标图像输入VGG19预训练网络提取网络第四次最大池化后的第2个卷积层的输出特征,然后计算高级特征之间的差异得到感知损失。感知损失可以减少高频特征丢失造成的图像降质,增强图像的细节使生成的图像更加真实,符合人类视觉感受,其定义如式(2)所示:

$$L_{VGG} = \frac{1}{n} \sum_{i=1}^n \|\phi(Y_i) - \phi(G(X_i))\|^2, \quad (2)$$

式中: ϕ 表示提取VGG19网络中高级特征的操作。生成对抗网络的对抗损失定义如式(3)所示:

$$L_{GAN}(G, D) = E_{X, Y} [\lg D(Y, X)] + E_{X, Y} [\lg (1 - D(X, G(X)))], \quad (3)$$

式中: D 代表判别器; X, Y 分别代表输入图像和目标图像。当 $D(Y, X)$ 的值越大时,证明判别器判定输入为真实图像对的概率越大。 L_{GAN} 越大,证明判别器分辨真假图像对的能力越强;而 L_{GAN} 越小,证明判别器分辨真假图像对的能力越弱,生成器的生成能力越强,其生成的数据逼近真实数据,干扰了判别器的判定,两者互相博弈优化。最终本文模型的损失函数如式(4)所示:

$$L_{loss} = \lambda_1 L_{L_1} + \lambda_2 L_{VGG} + L_{GAN}, \quad (4)$$

式中, λ_1, λ_2 分别为 L_{L_1} 和 L_{VGG} 的权重值,本文使用FUnIE-GAN^[11]中的权重: $\lambda_1 = 0.7; \lambda_2 = 0.3$ 。通过上述定义的网络和损失函数,对整个网络进行反复迭代训练,直至损失函数收敛为止,即得到了水下图像到增强图像的映射关系。

3 实验结果与分析

3.1 实验设置与模型训练

为验证本文模型的有效性,本文选取性能优异的六种水下图像增强方法进行对比实验,分别为水下暗通道先验(UDCP),RGHS^[24],IBLA^[25]三种经典的方法和UGAN^[9],FUnIE-GAN^[11],UIE-WD^[13]三种最新的深度学习模型,为保证实验的可比性,相关实验均在Linux操作系统下完成,CPU为Intel Xeon E5-2680 V4, GPU为NVIDIA Tesla P40 T1,显存大小22 G;实验过程中选择Adam优化器,学习率设置为0.000 3,每一批次大小设置为8;另外本文选择具有丰富水下退化环境的EUVP数据集^[11]进行训练测试,具体为合并EUVP数据集下的underwater_dark和underwater_imageNet两个子集,并随机挑选出8 608对图像作为训练集,余下做验证集使用;测试集本文选择EUVP数据集下的

测试集,共515对图像。其中训练过程的损失函数收敛曲线如图5所示,随着迭代次数的增加,生成器和判别器的损失函数值稳定下降,直至收敛后结束训练。

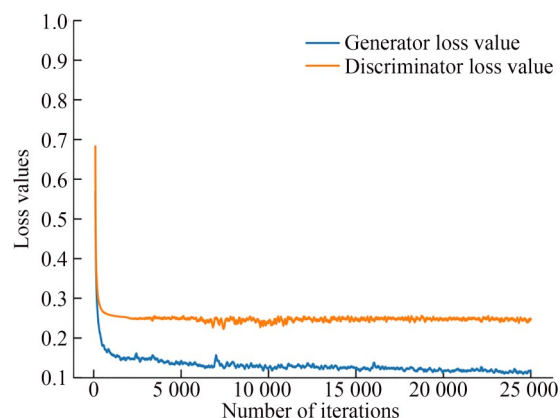


图5 损失函数值随迭代次数变化图

Fig. 5 Loss function variation chart with iterations

3.2 主观评价

由于版面篇幅有限,本文选取具有代表性的五幅图像进行主观分析。各方法的实验结果图如图6所示(彩图见期刊电子版),UDCP算法对图像的颜色矫正效果较差,图像整体色调偏暗,并会加剧蓝绿色偏的现象,不能处理亮度较低的水下图像(图6中圆圈所示)。RGHS算法整体上能够去除色偏,但会给部分图像带来伪影,如Image4和Image5。IBLA算法对图像的亮度有一定的补偿,但会造成一定的阴影,在颜色过渡上显得不自然,如Image3。基于深度学习的方法相比传统方法在恢复图像细节,提高图像清晰度和色彩过渡的自然性上更具优势。UGAN模型在图像清晰度上表现较好,但会过度补偿使图像整体颜色偏黄,如Image1,Image2和Image3。FUnIE-GAN模型为轻量级网络,在检测速度上优势较大,但增强图像同样存在色调偏黄的现象,并会带入略微伪影,如Image4和Image5。UIE-WD模型色调偏暗并且存在色偏,如Image3和Image5。从主观上来看,本文模型不仅提高了图像的清晰度、对比度和亮度,而且没有造成色偏、伪影等现象,增强了细节表达和色彩均衡性,使得图像与参考图像更为接近,呈现出更加令人满意的视觉效果。

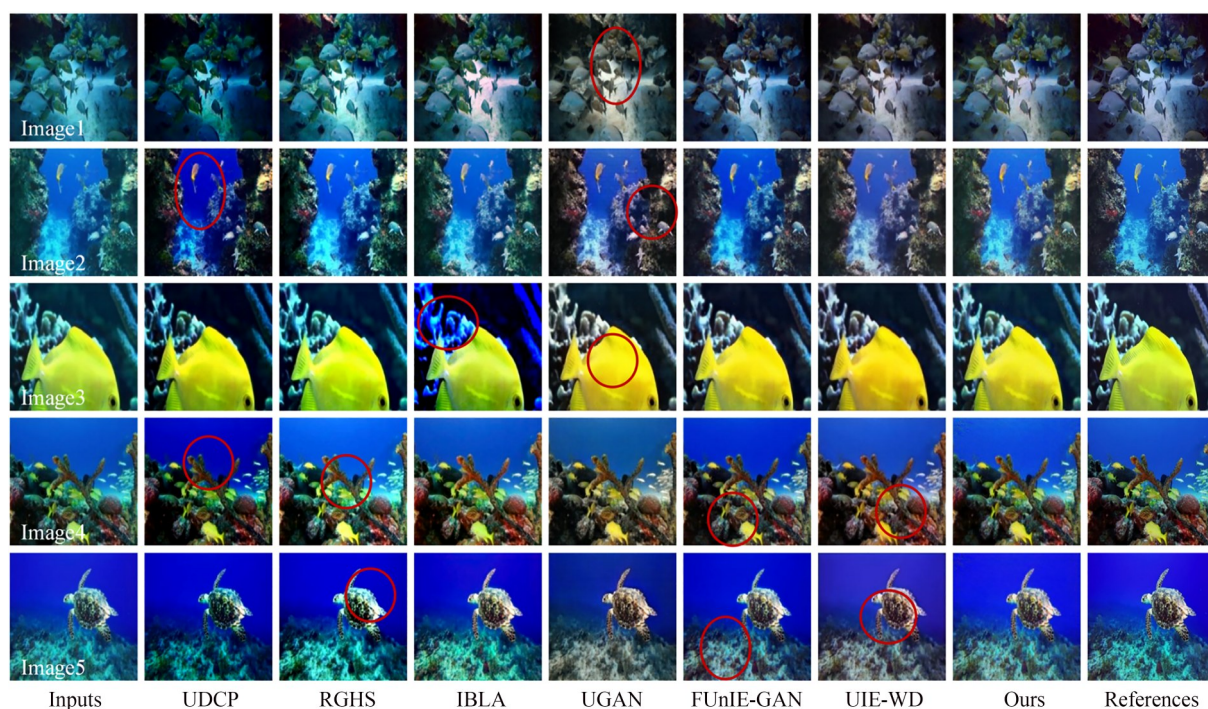


图6 7种方法处理结果对比图

Fig. 6 Comparison of the processing results of the seven methods

3.3 客观评价

因为水下图像质量退化因素复杂多样,个别图像使用单一评价指标可能与主观感受存在一定误差,所以本文使用 PSNR, SSIM, UIQM 和 NIQE 四种评价指标^[26-28]来对本文模型和其他方法的客观评价价值进行综合比较。PSNR 和 SSIM 为全参考质量评价指标,能够评价测试图像与参考图像的相似性,得分越高证明与参考图像的差异越小,细节保留越多,图像增强效果越好。UIQM 和 NIQE 为无参考质量评价指标,其中 UIQM 是一种基于视觉系统的水下图像质量评价指标。针对水下图像退化特点,该指标由色彩测量指标(UICM)、清晰度测量指标(UISM)和对比度测量指标(UIconM)线性组合而成,该值越大,证明图像的色彩、清晰度和对比度越好,UIQM 的定义如式(5)所示:

$$\text{UIQM} = c_1 \times \text{UICM} + c_2 \times \text{UISM} + c_3 \times \text{UIconM}, \quad (5)$$

式中: c_1, c_2, c_3 代表各分量的权重因子,本文中权重因子 $c_1 = 0.0282, c_2 = 0.2953, c_3 = 3.5753$ 。NIQE 为自然图像评价指标,通过计算测试图像与自然清晰图像多元高斯模型的距离来衡量图像的质量,值越小则证明图像质量越好。图 7~

10 详细给出了主观评价中五幅图像不同指标的得分,并给出平均分数,表 1 给出了在整个测试集上各方法的平均得分,其中最佳得分用粗体标记。

由图 7~图 8 可知,本文提出的模型所得结果在全参考质量评价指标上几乎全部取得了最优得分。尽管 UIE-WD 模型的 Image1 在 SSIM 指标上稍高于本文结果,但结合具体图像(图 6 第 1 行)可以清楚地观察到无论从色彩、清晰度、图像细节上本文模型所得出的结果都优于 UIE-WD 模型所得出的结果,这说明本文模型优于 UIE-WD 模型。

由图 9 和图 10 可知,本文提出的模型所得结果在无参考质量评价指标上大部分取得了最优得分。尽管 UGAN 模型的 Image3 在 UIQM 和 NIQE 指标上稍优于本文结果;同样结合具体图像(图 6)可以看出 UGAN 模型所得的 Image3 相对于参考图像存在黄色色偏,并且背景发白。本文模型所得的图像色彩更加自然,颜色变化更加协调(彩图见期刊电子版)。

从上述分析可知,单一的评价指标并不能全方面评估单一图像的质量,存在一定的误差,图像所得质量分数的优劣可能与人类的视

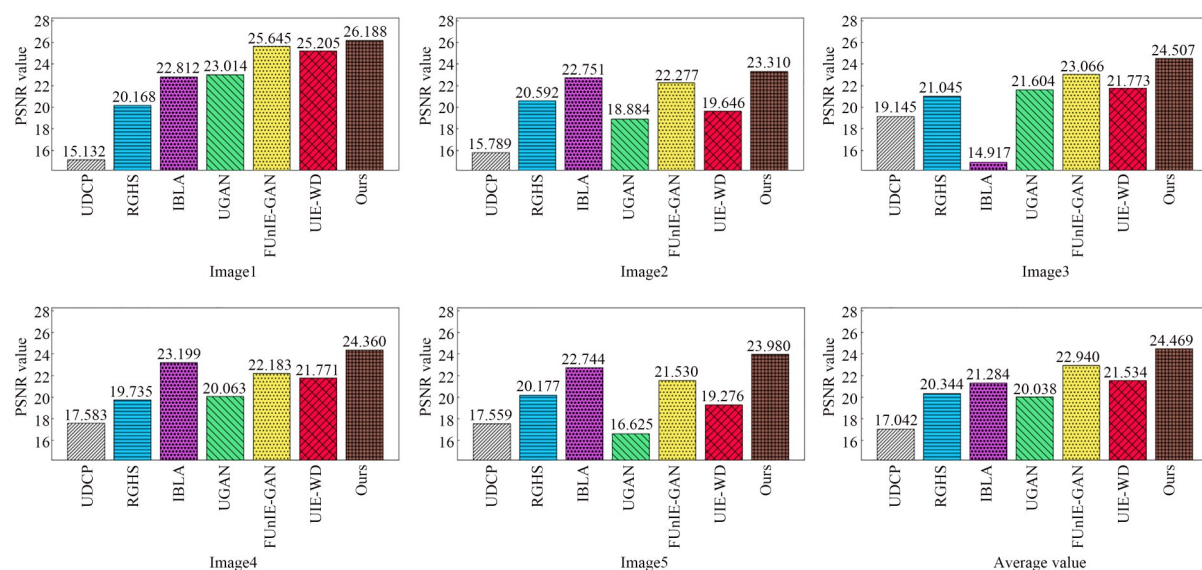


图7 本文提出的模型和对比方法所得结果的PSNR

Fig. 7 PSNR of the result obtained by the model proposed in this paper and comparison methods

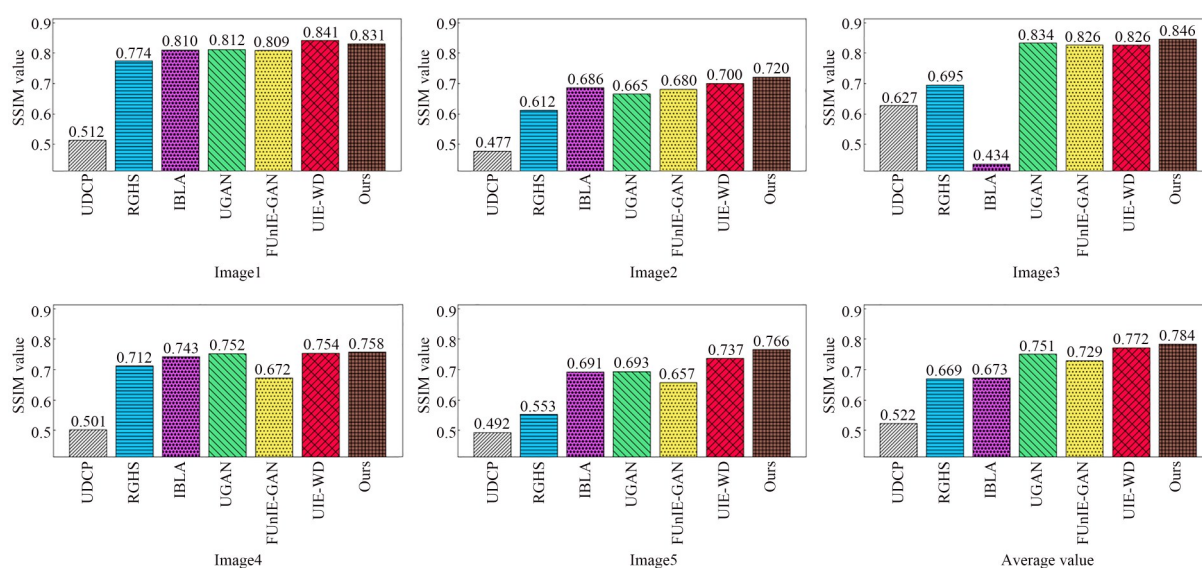
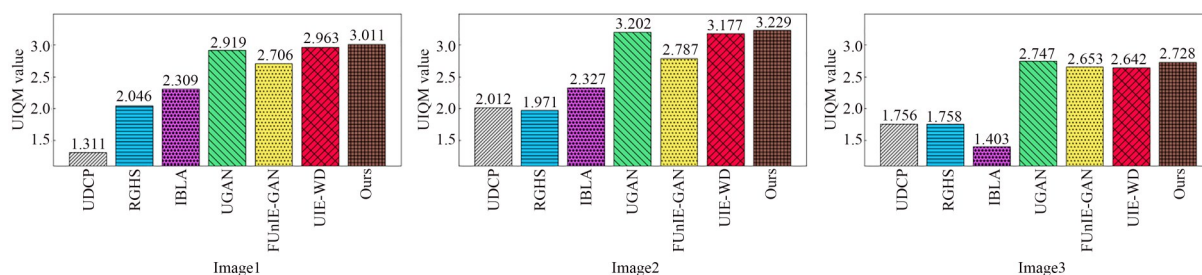


图8 本文提出的模型和对比方法所得结果的SSIM

Fig. 8 SSIM of the result obtained by the model proposed in this paper and comparison methods



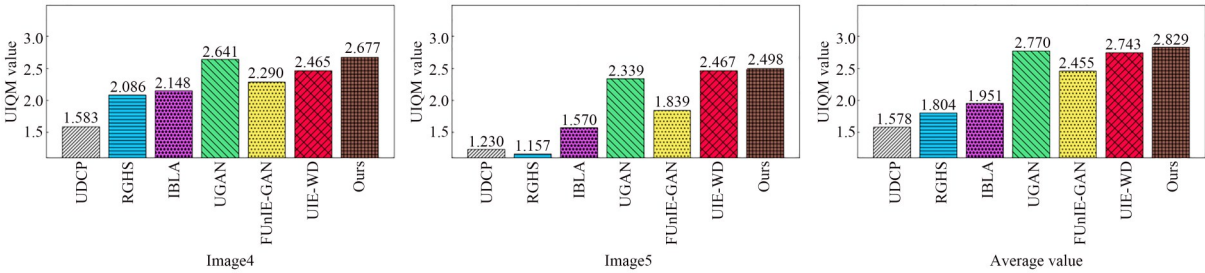


图9 本文提出的模型和对比方法所得结果的 UIQM

Fig. 9 UIQM of the result obtained by the model proposed in this paper and comparison methods

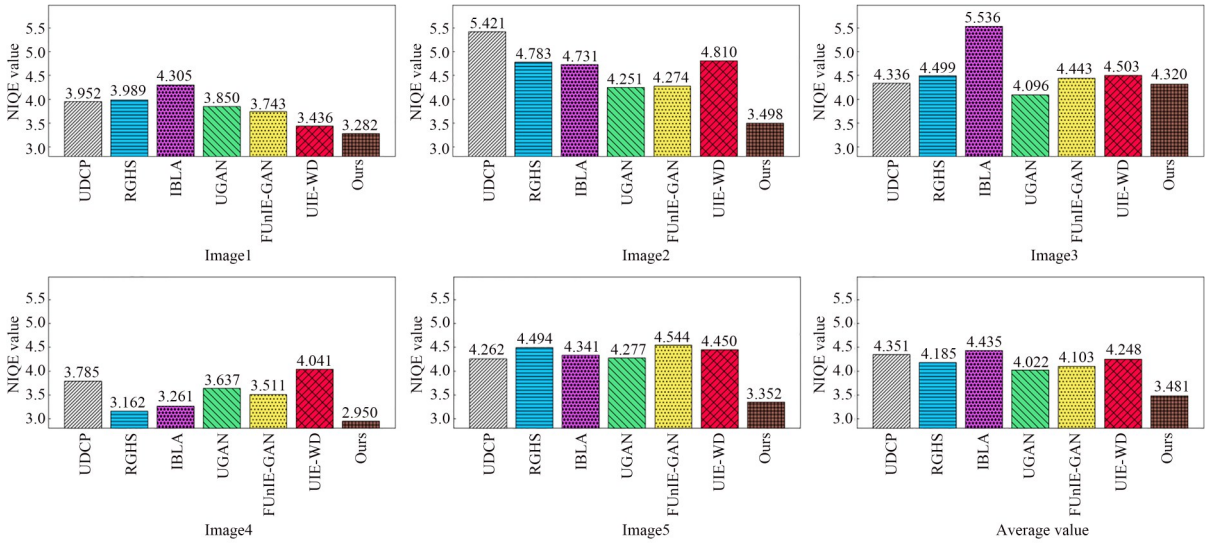


图10 本文提出的模型和对比方法所得结果的 NIQE

Fig. 10 NIQE of the result obtained by the model proposed in this paper and comparison methods

觉感受出现相矛盾的情况,只有当评价指标多样化或者测试样本足够多时,得到的结果才稳定可靠。例如,文献[29-30]通过两种评价指标在大量样本的情况下验证了方法的有效性,文献[31-32]通过四种评价指标综合验证了方法的先进性。因此本文选择四种质量评价指标综合考量图像质量,并在表1给出整个测试集在不同指标上的得分,使客观评价结果更加可信。

由表1可知,在整个测试集上,本文模型在PSNR和SSIM两种评价指标上获得最优得分,证明本文模型相对其他几种方法在提升图像质量的同时保留了原图像更多的内容细节,甚至有增强图像内容的效果。本文模型在UIQM和NIQE两种评价指标上获得最优得分,证明本文模型相对其他几种方法在提升清晰度、对比度和色彩均衡性等方面表现更好,视觉效果更符合人类视觉的预期。

表1 各方法在整个测试集上的平均得分

Tab. 1 Average score of each method on the entire test set

Methods	PSNR	SSIM	UIQM	NIQE
UDCP	14.496	0.559	1.843	4.748
RGHS	18.016	0.746	2.350	4.574
IBLA	18.870	0.700	2.158	4.728
UGAN	21.069	0.786	2.886	4.496
FUnIE-GAN	22.146	0.787	2.819	4.482
UIE-WD	21.605	0.787	2.845	4.615
Ours	22.499	0.789	2.911	4.175

3.4 实际验证分析

为进一步验证提出模型在实际拍摄水下图像的增强效果,本文还进行了标准色卡的水下测试,实验结果如图11所示(彩图见期刊电子版),每幅图像的左上角为标准色卡。其中UD-

CP, RGHS 和 IBLA 算法色调偏暗, UGAN, FUnIE-GAN 和 UIE-WD 模型整体色彩不鲜明, 本文模型色彩更加鲜丽, 色块与标准色卡更为接近。

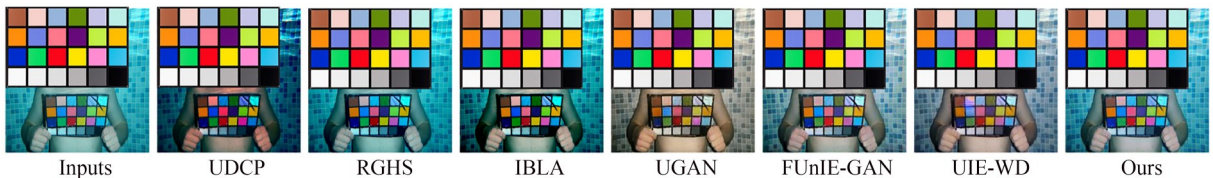


图 11 标准色卡测试对比图

Fig. 11 Comparison of standard color card tests

3.5 运行时间

针对本文模型以及六种对比方法的运行时间进行了统计分析,记录各方法处理整个测试集所需的时间,进一步计算获取每种方法处理单张图像的耗时,具体结果如表 2 所示。尽管本文方法在运行时间上不属于最优方法,但与另外三种深度学习方法相比,耗时相当,并且远远优于经典方法。

表 2 运行时间对比结果

Tab. 2 Comparison results of operating time

Methods	UDCP	RGHS	IBLA	UGAN	FUnIE-GAN	UIE-WD	Ours
Time/s	2.652	0.729	5.678	0.014	0.013	0.010	0.014

3.6 消融实验

为了验证本文引入全局编码器模块、多尺度混合卷积模块和注意力机制模块的有效性,针对这三个模块在整个测试集上分别进行消融实验,实验结果如表 3 所示。由表 3 可知,使用三个模块的网络模型在 PSNR, SSIM 和 UIQM 三个评价指标上取得了最优得分,在 NIQE 评价指标上取得了次优评分,综合四个指标的对比结果,证明了三个模块对图像增强的有效性。

表 3 消融实验对比结果

Tab. 3 Comparison results of ablation experiments

Attention mechanism	Hybrid convolution	Global encoder	PSNR	SSIM	UIQM	NIQE
✓	✓	✓	22.499	0.789	2.911	4.175
×	✓	✓	20.758	0.696	2.803	5.026
✓	×	✓	21.638	0.743	2.804	4.061
✓	✓	×	21.973	0.758	2.770	4.289

4 结 论

针对水下环境对光的吸收和散射造成水下图像颜色失真和细节丢失等问题,本文提出了一种基于生成对抗网络的水下图像增强方法。生成器中的编码器使用局部编码器和全局编码器双分支结构,兼顾水下图像的局部信息和全局信息。使用多尺度混合卷积和注意力机制加强网络的拟合能力。实验结果表明,对比其他几种方法,在相同的训练集和测试集上训练测试,本文模型不仅在主观视觉感受上表现最优,而且在整个测试集上,PSNR, SSIM, UIQM, NIQE 四种质

量评价指标分别取得 22.499, 0.789, 2.911 和 4.175 的平均分数, 均优于其他方法, 并且消融实验证明了各个模块的有效性。

尽管本文模型在提升图像质量方面取得了优异的效果, 但网络较为复杂, 后续将在降低网络复杂度的情况下, 使其依然保持同等性能方面进行探索。

参考文献:

- [1] 郭继昌, 李重仪, 郭春乐, 等. 水下图像增强和复原方法研究进展[J]. 中国图象图形学报, 2017, 22(3): 273-287.
GUO J C, LI C Y, GUO C L, *et al.* Research progress of underwater image enhancement and restoration methods[J]. *Journal of Image and Graphics*, 2017, 22(3): 273-287. (in Chinese)
- [2] 袁国铭, 杨光, 王金峰, 等. 由粗到细的多级小波变换水下图像增强[J]. 光学精密工程, 2022, 30(22): 2939-2951.
YUAN G M, YANG G, WANG J F, *et al.* Coarse-to-fine underwater image enhancement based on multi-level wavelet transform[J]. *Opt. Precision Eng.*, 2022, 30(22): 2939-2951. (in Chinese)
- [3] 林森, 刘旭. 门控融合对抗网络的水下图像增强[J]. 图学报, 2021, 42(6): 948-956.
LIN S, LIU X. Underwater image enhancement algorithm using gated fusion generative adversarial network[J]. *Journal of Graphics*, 2021, 42(6): 948-956. (in Chinese)
- [4] 王明军, 彭月, 刘燕荣, 等. 模拟水体湍流环境下目标激光点云数据的三维重建与分析[J]. 光电工程, 2023, 50(6): 230004.
WANG M J, PENG Y, LIU Y R, *et al.* Three-dimensional reconstruction and analysis of target laser point cloud data in simulated turbulent water environment[J]. *Opto-Electronic Engineering*, 2023, 50(6): 230004. (in Chinese)
- [5] GHANI A S A, ISA N A M. Enhancement of low quality underwater image through integrated global and local contrast correction[J]. *Applied Soft Computing*, 2015, 37(C): 332-344.
- [6] 胡振宇, 陈琦, 朱大奇. 基于颜色平衡和多尺度融合的水下图像增强[J]. 光学精密工程, 2022, 30(17): 2133-2146.
HU Z Y, CHEN Q, ZHU D Q. Underwater image enhancement based on color balance and multi-scale fusion[J]. *Opt. Precision Eng.*, 2022, 30(17): 2133-2146. (in Chinese)
- [7] 张楠楠, 李志伟, 郭新军, 等. 使用改进型大气散射模型的双阶段图像修复[J]. 光学精密工程, 2022, 30(18): 2267-2279.
ZHANG N N, LI Z W, GUO X J, *et al.* Two-stage image restoration using improved atmospheric scattering model[J]. *Opt. Precision Eng.*, 2022, 30(18): 2267-2279. (in Chinese)
- [8] 丛润民, 张禹墨, 张晨, 等. 深度学习驱动的水下图像增强与复原研究进展[J]. 信号处理, 2020, 36(9): 1377-1389.
CONG R M, ZHANG Y M, ZHANG C, *et al.* Research progress of deep learning driven underwater image enhancement and restoration[J]. *Journal of Signal Processing*, 2020, 36(9): 1377-1389. (in Chinese)
- [9] FABBRI C, ISLAM M J, SATTAR J. Enhancing underwater imagery using generative adversarial networks[C]. 2018 *IEEE International Conference on Robotics and Automation (ICRA)*. Brisbane, QLD, Australia. IEEE, 2018: 7159-7165.
- [10] ZHU J Y, PARK T, ISOLA P, *et al.* Unpaired image-to-image translation using cycle-consistent adversarial networks[C]. 2017 *IEEE International Conference on Computer Vision (ICCV)*. Venice, Italy. IEEE, 2017: 2242-2251.
- [11] ISLAM M J, XIA Y Y, SATTAR J. Fast underwater image enhancement for improved visual perception[J]. *IEEE Robotics and Automation Letters*, 2020, 5(2): 3227-3234.
- [12] LI C Y, ANWAR S, HOU J H, *et al.* Underwater image enhancement via medium transmission-guided multi-color space embedding [J]. *IEEE Transactions on Image Processing*, 2021, 30: 4985-5000.
- [13] MA Z Y, OH C. A wavelet-based dual-stream network for underwater image enhancement [C]. *ICASSP 2022 - 2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. Singapore, Singapore. IEEE, 2022: 2769-2773.
- [14] FU Z Q, WANG W, HUANG Y, *et al.* Uncertainty inspired underwater image enhancement[C]. *AVIDAN S, BROSTOW G, Cissé M, et al.*

- European Conference on Computer Vision*. Cham: Springer, 2022: 465-482.
- [15] 米泽田, 晋洁, 李圆圆, 等. 基于多尺度级联网络的水下图像增强方法[J]. 电子与信息学报, 2022, 44(10): 3353-3362.
MI Z T, JIN J, LI Y Y, *et al.* Underwater image enhancement method based on multi-scale cascade network[J]. *Journal of Electronics & Information Technology*, 2022, 44(10): 3353-3362. (in Chinese)
- [16] 李云, 孙山林, 黄晴, 等. 基于多路混合注意力机制的水下图像增强网络[J]. 电子与信息学报, 2024, 46(1): 118-128.
LI Y, SUN S L, HUANG Q, *et al.* Underwater image enhancement network based on multi-channel hybrid attention mechanism [J]. *Journal of Electronics & Information Technology*, 2024, 46(1): 118-128. (in Chinese)
- [17] ZHANG D H, WU C Y, ZHOU J C, *et al.* Hierarchical attention aggregation with multi-resolution feature learning for GAN-based underwater image enhancement[J]. *Engineering Applications of Artificial Intelligence*, 2023, 125: 106743.
- [18] RONNEBERGER O, FISCHER P, BROX T. U-Net: convolutional networks for biomedical image segmentation[C]. *Medical Image Computing and Computer-Assisted Intervention-MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18*. Springer International Publishing, 2015: 234-241.
- [19] KINOSHITA Y, KIYA H. Convolutional neural networks considering local and global features for image enhancement[C]. 2019 *IEEE International Conference on Image Processing (ICIP)*. Taipei, China. IEEE, 2019: 2110-2114.
- [20] WOO S, PARK J, LEE J Y, *et al.* CBAM: convolutional block attention module [C]. *European Conference on Computer Vision*. Cham: Springer, 2018: 3-19.
- [21] ISOLA P, ZHU J Y, ZHOU T H, *et al.* Image-to-image translation with conditional adversarial networks[C]. 2017 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Honolulu, HI, USA. IEEE, 2017: 5967-5976.
- [22] YI Z L, ZHANG H, TAN P, *et al.* DualGAN: unsupervised dual learning for image-to-image translation [C]. 2017 *IEEE International Conference on Computer Vision (ICCV)*. Venice, Italy. IEEE, 2017: 2868-2876.
- [23] 范新南, 杨鑫, 史朋飞, 等. 特征融合生成对抗网络的水下图像增强[J]. 计算机辅助设计与图形学学报, 2022, 34(2): 264-272.
FAN X N, YANG X, SHI P F, *et al.* Underwater image enhancement based on feature fusion generative adversarial networks[J]. *Journal of Computer-Aided Design & Computer Graphics*, 2022, 34(2): 264-272. (in Chinese)
- [24] WANG Y, SONG W, FORTINO G, *et al.* An experimental-based review of image enhancement and image restoration methods for underwater imaging[J]. *IEEE Access*, 2019, 7: 140233-140251.
- [25] PENG Y T, COSMAN P C. Underwater image restoration based on image blurriness and light absorption[J]. *IEEE Transactions on Image Processing*, 2017, 26(4): 1579-1594.
- [26] 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.
- [27] MITTAL A, SOUNDARARAJAN R, BOVIK A C. Making a “completely blind” image quality analyzer [J]. *IEEE Signal Processing Letters*, 2013, 20(3): 209-212.
- [28] HU K, WENG C H, ZHANG Y W, *et al.* An overview of underwater vision enhancement: from traditional methods to recent deep learning [J]. *Journal of Marine Science and Engineering*, 2022, 10(2): 241.
- [29] 王欣, 石慧. DRHA-UIE: 基于双重残差混合注意力模块的水下图像增强方法[J]. 电子学报, 2023, 51(9): 2398-2407.
WANG X, SHI H. DRHA-UIE: an underwater image enhancement method based on dual residual hybrid attention block [J]. *Acta Electronica Sinica*, 2023, 51(9): 2398-2407. (in Chinese)
- [30] ZHOU W H, ZHU D M, SHI M, *et al.* Deep images enhancement for turbid underwater images based on unsupervised learning[J]. *Computers and Electronics in Agriculture*, 2022, 202: 107372.
- [31] WANG T H, WANG L L, ZHANG E, *et al.* Underwater image enhancement based on optimal contrast and attenuation difference [J]. *IEEE Access*, 2023, 11: 68538-68549.

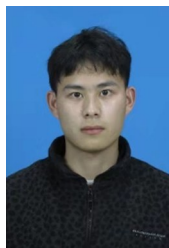
- [32] 巩文哲, 褚金奎, 成昊远, 等. 基于无监督学习和注意力机制的水下偏振图像融合[J]. 光学精密工程, 2023, 31(21): 3212-3220.
- GONG W Z, CHU J K, CHENG H Y, *et al.* Un-

derwater polarization image fusion based on unsupervised learning and attention mechanisms [J]. *Opt. Precision Eng.*, 2023, 31(21): 3212-3220. (in Chinese)

作者简介:



夏晓华(1987—),男,山东临沂人,博士,副教授,2009年于中国石油大学(华东)获得学士学位,2015年于西安交通大学获得博士学位,主要从事机器视觉与光机电一体化方面的研究。
E-mail: xhxia@chd.edu.cn



钟预全(2000—),男,湖南怀化人,硕士研究生,2022年于长安大学获得学士学位,主要从事机器视觉的研究。
E-mail: 1163515359@qq.com