

基于分频式生成对抗网络的非成对水下图像增强

牛玉贞, 张凌昕, 兰 杰, 许 瑞, 柯 道*

(福州大学计算机与大数据学院, 福建福州 350108)

摘 要: 增强水下图像质量对水下作业领域的发展具有重要意义。现有的水下图像增强方法通常基于成对的水下图像和参考图像进行训练, 然而实际获取与水下图像对应的参考图像比较困难, 相比之下获得非成对高质量水下图像或者陆上图像较为容易。此外, 现有的水下图像增强方法很难同时针对各种失真类型进行图像增强。为了避免对成对训练数据的依赖和进一步降低获得训练数据的难度, 并应对多样的水下图像失真类型, 本文提出了一种基于分频式生成对抗网络(Frequency-Decomposed Generative Adversarial Network, FD-GAN)的非成对水下图像增强方法, 并在此基础上设计了高低频双分支生成器用于重建高质量水下增强图像。具体来说, 本文引入特征级别的小波变换将特征分为低频和高频部分, 并基于循环一致性生成对抗网络对低频和高频部分区分处理。其中, 低频分支采用结合低频注意力机制的编码-解码器结构实现对图像颜色和亮度的增强, 高频分支则采用并行的高频注意力机制对各高频分量进行增强, 从而实现图像细节的恢复。在多个标准水下图像数据集上的实验结果表明, 本文提出的方法在使用非成对的高质量水下图像和引入部分陆上图像的情况下, 均能有效生成高质量的水下增强图像, 且有效性和泛化性均优于当前主流的水下图像增强方法。

关键词: 水下图像增强; 生成对抗网络; 小波变换; 注意力机制; 高低频双分支生成器

基金项目: 国家自然科学基金(No.U21A20472, No.61972097); 国家重点研发计划(No.2021YFB3600503); 福建省科技重大专项(No.2021HZ022007); 福建省自然科学基金(No.2023J01067, No.2020J01494); 福建省科技厅高校产学研合作项目(No.2021H6022)

中图分类号: TP391

文献标识码: A

文章编号: 0372-2112(2025)02-0527-18

电子学报 URL: <http://www.ejournal.org.cn>

DOI: 10.12263/DZXB.20240265

FD-GAN: Frequency-Decomposed Generative Adversarial Network for Unpaired Underwater Image Enhancement

NIU Yu-zhen, ZHANG Ling-xin, LAN Jie, XU Rui, KE Xiao*

(1. College of Computer and Data Science, Fuzhou University, Fuzhou, Fujian 350108, China)

Abstract: Enhancing the quality of underwater images is crucial for advancements in the fields of underwater exploration and underwater rescue. Existing underwater image enhancement methods typically rely on paired underwater images and reference images for training. However, obtaining corresponding reference images for underwater images is challenging in practice. In contrast, acquiring high-quality unpaired underwater images or images captured on land are relatively more straightforward. Furthermore, existing techniques for underwater image enhancement often struggle to address a variety of distortion types simultaneously. To avoid the reliance on paired training data, reduce the difficulty of acquiring training data, and effectively handle diverse types of underwater image distortions, in this paper, we propose a novel unpaired underwater image enhancement method based on the frequency-decomposed generative adversarial network (FD-GAN). We design a dual-branch generator based on high and low frequencies to reconstruct high-quality underwater images. Specifically, feature-level wavelet transform is introduced to separate the features into low-frequency and high-frequency parts. Then the separated features are processed by a cycle-consistent generative adversarial network, so as to simultaneously enhance the color and luminance in the low-frequency component and details in the high-frequency part. More specific, the low-frequency branch employs an encoder-decoder structure with a low-frequency attention mechanism to enhance the color and brightness of the image. The high-frequency branch utilizes parallel high-frequency attention mechanisms to enhance various

high-frequency components, thereby achieving the restoration of image details. Experimental results on multiple datasets show that the proposed method trained with unpaired high-quality underwater images or unpaired high-quality underwater images and on-land images, can effectively generate high-quality underwater enhanced images and the proposed method is superior to the state-of-the-art underwater image enhancement methods in terms of effectiveness and generalization.

Key words: underwater image enhancement; generative adversarial networks; wavelet transform; attention mechanism; dual-branch generator based on high and low frequencies

Foundation Item(s): National Natural Science Foundation of China (No.U21A20472, No.61972097); National Key Research and Development Program of China (No.2021YFB3600503); Major Science and Technology Project of Fujian Province (No.2021HZ022007); Natural Science Foundation of Fujian Province (No.2023J01067, No.2020J01494); Fujian Provincial Department of Science and Technology University Industry-Academy Cooperation Project (No.2021H6022)

1 引言

水下成像技术已被广泛应用于水下救援、海底勘探、水下目标检测等众多领域。由于不同光波在水下的能量衰减率有显著差异,如蓝光和绿光衰减最慢,而红光衰减最快,水下图像通常呈现蓝绿色,进而导致水下图像的颜色对比度和亮度降低。同时,水中的悬浮颗粒为水下图像引入噪声,导致水下图像模糊。这些失真的水下图像因无法准确反映水下环境的真实情况,给基于人眼视觉和计算机视觉的水下作业带来困难。因此,构建通用且有效的水下图像增强方法已成为当前水下成像领域中一个亟待解决的重要研究问题。

近年来,研究者提出了诸多水下图像增强方法,但这些方法通常需要成对的水下图像和参考图像进行训练(如图1(a)和图1(b)所示)。而在现实中获取与原始水下图像对应的参考图像难度较高,需要消耗大量的人力和物力成本。因此,为避免对成对训练数据的依赖,需要对非成对水下图像增强进行研究。如图1所示,相比于水下图像(a)成对的参考图像(b),获得非成对的高质量水下图像(c)和陆上图像(d)相对容易。尽管如此,现有的水下图像增强方法基本采用全监督方式,无法直接使用非成对水下图像进行训练,这使得基于非成对图像进行水下图像增强的研究变得尤为必要且具有挑战性。

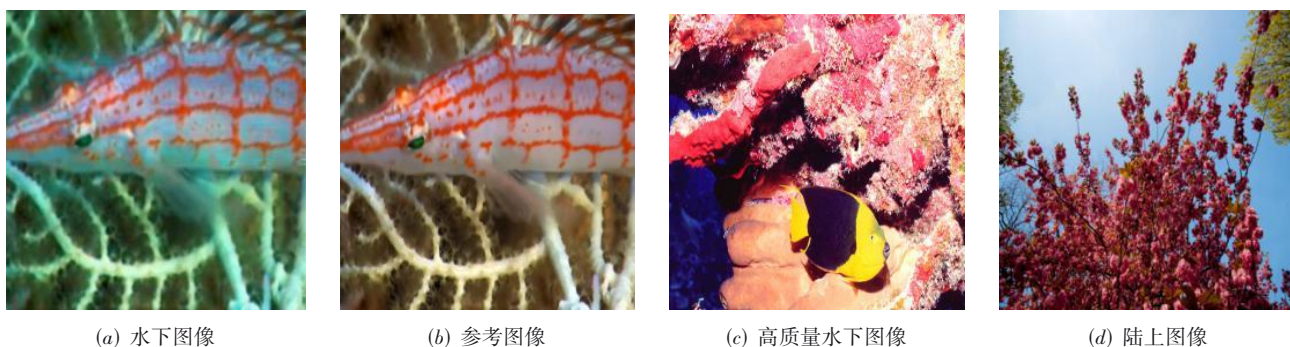


图1 水下图像及相应训练图像示例图

已有的水下图像增强方法主要包括传统的方法和基于深度学习的方法。其中,传统的水下图像增强方法大致分为两类:一类通过调整水下图像的像素值提高图像质量;另一类则先通过数学建模反演水下图像的退化过程,再估计物理模型参数以重构高质量的水下增强图像。如 Iqbal 等人^[1]设计了一种对比度校正模型,通过校正水下图像的颜色提高水下图像质量。刘柯等人^[2]则基于直方图均衡化法恢复水下图像颜色,并通过滤波处理去除图像中的雾霾,增强水下图像质量。尽管这些方法在某些特定的场景下可以取得良好的效果,但它们的泛化能力仅限于特定分布的数据集。而且,传统的基于物理模型的水下图像增强方法通常需

要对水体的大量光学参数、相机参数及拍摄距离等各种复杂的物理信息建立对应的模型。但是,这些额外的参数难以手动测量或推断。因此,建立能够精确匹配水下退化过程的物理模型对研究者来说仍是一项十分困难的任务,这直接影响到反演得到的图像的质量。

近年来,基于深度学习的方法在多种图像增强任务中表现良好^[3-5]。相比传统方法,基于深度学习的水下图像增强方法^[6-25]通过学习输入图像的特征表示,在特征层面去除原始水下图像的失真,并将其重构成高质量的水下增强图像。如 Islam 等人^[9]假设失真图像和增强图像之间存在非线性映射,将水下图像增强表示为图像到图像的转换问题,从而对水下图像进行增强。

Naik 等人^[10]提出了一种水下浅层神经网络架构,使用更少的模型参数生成质量较优的水下增强图像. Liu 等人^[14]通过结合不同空间信息提出了自适应学习注意网络,并在此基础上设计平行注意模块和浅层学习模块自适应学习重要水下特征信息. 重建高质量的水下增强图像要求模型对水下图像的特征进行充分学习,对不同类型的水下图像失真问题进行针对性处理. 然而,

现有的水下图像增强方法主要致力于改善图像的颜色,对特征进行统一学习,难以同时解决水下图像多种类型的失真问题. 这些失真问题不仅包括颜色退化问题,还有亮度不足、模糊和噪声等现象. 如图 2 所示,现有的水下图像增强方法生成的结果中仍然存在亮度不足[图 2(b)]、模糊[图 2(c)]以及伪影[图 2(d)红色箭头所示]等问题.

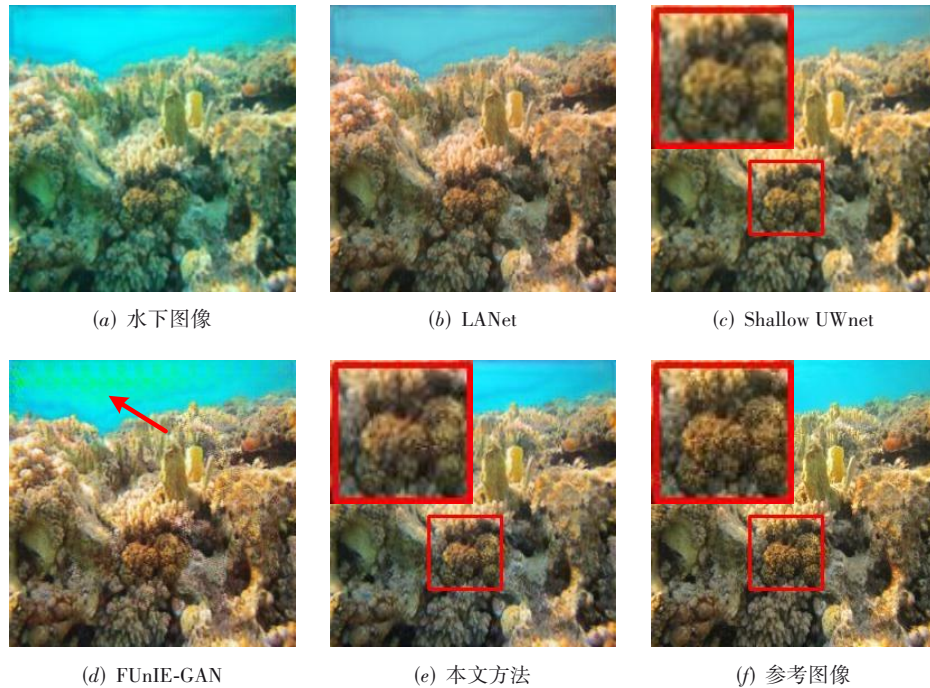


图 2 不同方法得到的增强图像失真问题示例图

为避免对成对训练数据的依赖并应对多样的水下图像失真类型,本文提出了一种基于分频式生成对抗网络(Frequency-Decomposed Generative Adversarial Network, FD-GAN)的非成对水下图像增强方法. 首先,为了进一步降低获得训练数据的难度,本文提出了一种基于循环一致性生成对抗网络^[16](Generative Adversarial Network, GAN)的水下图像增强方法,该方法无需对成对图像进行训练. 具体而言,受到循环生成一致性思想^[17]的启发,本文构建了一个以非成对的高质量水下图像和陆上图像为监督条件的水下图像增强模型框架,旨在解决由于数据分布不均、特征复杂等问题导致的生成对抗网络表现不稳定且训练困难的问题. 此外,为了处理水下图像中复杂而多样的失真类型,本文对各种失真类型与图像特征之间的关系进行了分析. 具体来说,水下图像的颜色和亮度等信息主要由特征图的低频部分表征,而水下图像的细节纹理等信息主要由特征图的高频部分表征. 因此,本文结合特征级别小波变换,将水下图像特征分解为低频和高频部分,并分

别对低频和高频部分进行处理,达到同时增强低频颜色、亮度和高频细节的目的,从而恢复出高质量的水下增强图像. 特别地,本文设计了高低频双分支生成器,采用双分支的生成对抗网络对小波变换后的低频和高频部分分别有针对性地进行处理.

值得注意的是,特征图低频部分和高频部分具有不同特性,对这两部分进行统一处理无法同时达到期望的增强效果. 因此,为了更充分地利用这两部分的特性,本文分别对特征的低频部分与高频部分进行处理. 具体来说,本文在高低频双分支生成器中针对特征的低频部分和高频部分的不同信息,分别设计了不同的网络结构和对应的注意力机制,更具针对性地解决了水下图像中不同的失真问题. 本文进一步将传统的小波池化层^[18]推广到特征层面,将空间域特征根据水平和垂直方向上的低频和高频,分解为低频部分和高频部分. 一方面,本文针对特征的低频部分设计了低频处理分支,提出结合低频注意力机制的编码-解码结构实现对低频特征的整体增强,进而重点增强水下图像的

颜色与亮度. 另一方面, 由于高频部分包含多种高频分量, 本文在设计高频处理分支时, 采用并行的高频注意力机制对各高频分量进行残差增强, 以恢复图像的细节纹理.

为了验证本文提出方法的有效性, 本文在标准水下图像数据集上进行了系统性的实验, 包括 EUVP^[9]、UFO120^[11]、UIEB^[8]、Challenge60^[8]和 RUIE^[19]. 此外, 我们还引入陆上图像数据集 MIT-Adobe FiveK^[20]作为高质量水下图像的替换数据, 以验证本文方法的创新性. 实验结果证明, 本文方法相较现有的先进方法具有更优的性能, 且具有更好的泛化能力. 同时, 详尽的消融实验也验证了本文方法各部分的有效性.

综上所述, 本文的主要贡献可以概括如下:

(1) 本文提出了一种基于分频式生成对抗网络的非成对水下图像增强方法, 并结合循环一致性思想进行框架搭建. 该方法能够在给定非成对的高质量水下图像和陆上图像作为监督的条件下, 有效应对多种类型的水下图像失真问题.

(2) 本文设计了一种高低频双分支生成器结构, 低频分支采用结合低频注意力机制的编码-解码器结构实现对图像颜色和亮度的增强, 高频分支则采用并行的高频注意力机制对各高频分量进行增强, 从而实现对图像细节的恢复.

(3) 大量且系统性的实验结果表明, 本文提出的方法在有效性和泛化性上均优于目前先进的基于成对数据训练的水下图像增强方法.

2 相关工作

2.1 传统水下图像增强方法

传统水下图像增强方法主要分为基于非物理模型的方法和基于物理模型的方法. 基于非物理模型的方法主要通过调整水下图像的像素值提高图像质量, 无需数学模型. Iqbal 等人^[1]设计了一种对比度校正模型, 通过校正水下图像中的蓝色和红色还原图像原始的颜色. 为了恢复水下图像的细节, Ghani 等人^[21]通过在一定范围内遵循 Rayleigh 分布去除水下图像的伪影. 为了解决因水下介质散射导致的图像质量下降问题, Ancuti 等人^[22]提出去除水下的雾霾效果, 并通过改变水下图像的颜色信息提高视觉效果. Ancuti 等人^[23]提出了一种颜色通道补偿方法, 通过对相关的颜色通道进行补偿, 在颜色外观上有效增强了水下图像. 由于水下环境复杂且观测条件有限, 这些基于非物理模型的方法生成的增强图像经常存在颜色较暗和细节丢失等问题.

基于物理模型的方法通过模拟物理成像过程, 并根据观测数据和各种先验假设估计成像参数, 实现水

下图像增强. 其中, 基于 Retinex 算法的方法^[2,24,25]主要依据高分辨率图像和水下图像的像素值分布规律, 通过映射函数扩大水下图像像素值的动态范围, 从而提高图像对比度. 此外, 许多基于物理模型的水下图像增强方法使用了暗通道先验(Dark Channel Prior, DCP)^[26]原理, 通过在图像中寻找颜色通道数值较低的像素点, 使用滤波来增强水下图像. Chiang 等人^[27]提出将暗通道先验与波长相关的补偿算法相结合实现水下图像增强. 与结合波长的算法不同, Drews 等人^[28]针对在水下图像中红通道信息不可靠的问题, 提出了一种专用于水下图像的 DCP 算法.

为了进一步恢复图像亮度并提高颜色对比度, Peng 等人^[29]提出了一种广义 DCP 算法, 使用与深度相关的颜色变化估计环境光, 并通过场景深度的估计提高背景光的精度, 进而改善传统 DCP 方法的效果. 与暗通道方法不同, Berman 等人^[30]考虑不同类型水下环境不同的光谱剖面, 通过估计蓝红和蓝绿通道的衰减比, 恢复退化的水下图像. 这些基于物理模型的方法在水下图像增强任务中取得了较好的效果, 但是这些方法通常计算耗时较长, 且对水下图像的失真类型敏感, 方法的泛化能力较低.

2.2 基于深度学习的水下图像增强方法

近年来, 深度卷积神经网络(Convolutional Neural Networks, CNN)促进了图像增强技术的迅速发展^[3-5,31-34]. 在水下图像增强任务中, 基于深度学习的方法也快速发展并取得了先进的性能. 基于深度学习的水下图像增强方法主要分为3类, 即基于融合网络的方法、基于生成对抗网络的方法和基于轻量级网络的方法, 这3类方法侧重点各不相同.

基于融合网络的方法通过神经网络模型, 将各个基于物理模型的方法的初步增强图像进行加权融合得到最终的水下增强图像. Ancuti 等人^[35]通过特征金字塔将白平衡法和通道均衡法得到的增强结果依照不同权重进行融合. 为适应水下图像退化的特点, Li 等人^[8]提出 WaterNet 方法, 通过神经网络融合3种基于物理模型的图像增强方法的结果, 包括白平衡法、伽马校正法和直方图均衡法.

生成对抗网络(Generative Adversarial Network, GAN)^[16]被广泛应用于图像增强, 在水下图像增强任务中也有较好的表现. Li 等人^[6]基于生成对抗网络建立了 WaterGAN 模型, 在结合水下图像形成过程的基础上, 构建高分辨率的水下增强图像. 为了更好地学习水下图像特征, Guo 等人^[36]提出了一种多尺度致密生成对抗网络(Dense-GAN), 通过学习图像在特征域中的分布, 应对水下图像的失真现象. 为了更好地提取与重建特征, Fabbri 等人^[7]提出了 UGAN 网络结构, 运用了 U-

Net^[37]结构作为生成器,通过降采样操作提取水下图像特征,并通过上采样操作重构图像;网络中卷积层之间的跳跃连接则有助于水下特征的学习.为了解决水下图像的细节模糊问题,Islam等人^[9]提出FUnIE-GAN(Fast Underwater Image Enhancement GAN),假设失真图像和增强图像之间存在非线性映射,将水下图像增强表示为图像到图像的转换问题,从而消除水下图像中的模糊.

近年来,轻量级的水下图像增强方法得到快速发展.这些方法使用较少的参数且耗时较短,能实现良好的增强效果.Islam等人^[11]提出了一种可以实现水下超分辨率的水下图像增强方法(Deep Simultaneous Enhancement and Super Resolution, Deep SESR),运用端到端的网络结构,设计辅助注意力网络及密集残差子网络,促进网络模型学习局部的特征信息,进而恢复水下图像细节.Li等人^[12]提出了一种基于水下场景先验的轻量级水下图像增强卷积神经网络模型,根据不同的水域成像类型和水下成像物理模型来模拟真实的水下图像.

为了提高水下图像增强方法的泛化能力,Naik等人^[10]提出了一种水下浅层神经网络架构Shallow UWnet,它由多个卷积结合跳跃连接残差模块组成,使用更少的模型参数生成质量较优的水下增强图像.Wang等人^[13]提出了域自适应的水下图像增强网络(Two-phase Underwater Domain Adaptation network, TUDA),通过最小化域间和域内的域差距,更准确地评估感知质量,进而增强水下图像.Jiang等人^[38]提出的LCNet模型运用特征金字塔结构提取水下图像特征,并通过多尺度的网络结构对水下图像进行增强.Zhuang等人^[39]利用超拉普拉斯先验,增强水下图像细节,并恢复图像颜色.Zhou等人^[40]通过多色空间导向色彩估计模块和密集注意力模块引导网络提取颜色信息,从而对水下图像进行增强.为更好地应对不同的水下环境问题,Fu等人^[15]将水下图像增强分解为分布估计和共识过程,提出了一种新的水下图像增强概率网络学习退化水下图像的分布.为了更好地解决水下浑浊环境造成的失真问题,Liu等人^[14]通过结合不同空间信息提出了自适应学习注意力网络LANet,并在此基础上设计平行注意模块和浅层学习模块自适应学习重要水下特征信息,实现水下图像增强.

上述基于深度学习的水下图像增强方法虽然可以在一定程度上生成较高质量的增强图像,但是无法全面地恢复水下图像在细节和颜色上的失真.此外,这些方法需要对成对的水下图像进行训练,而收集成对的水下图像数据集具有较大的难度.与上述方法不同,本文方法不仅不需要使用成对数据进行监督,而且对特征的低频部分和低频部分分别进行处理,可以应对复

杂多样的水下图像失真类型,进而更好地恢复水下图像的细节与颜色信息,最终获得高质量的水下增强图像.

3 本文方法

本章首先在3.1节介绍提出的分频式生成对抗网络框架;然后,在3.2节中阐明该框架核心包含的高低频双分支生成器;进一步地,在3.3节和3.4节中介绍低频处理分支和低频注意力机制、高频处理分支和高频注意力机制;最后,在3.5节中阐述框架训练所涉及的损失函数.

3.1 分频式生成对抗网络框架

为避免模型对成对水下图像的依赖,本文提出的分频式生成对抗网络基于循环一致性思想^[17]进行框架搭建,缓解了由于数据分布不均、特征复杂等导致的GAN表现不稳定且训练困难的问题.此外,本文设计了高低频双分支生成器,以处理水下图像存在的多种失真类型.通过两组高低频双分支生成器和对应的图像领域判别器之间的循环生成对抗,实现了非成对的水下图像质量增强.本文提出的分频式生成对抗网络FD-GAN的网络框架如图3所示.为方便表述,本文用C表示清晰且高质量的水下图像领域分布,用W表示原始水下图像领域分布.

如图3所示,该网络包括两个高低频双分支生成器(高质量图像生成器 G_{w2c} 和 underwater 图像生成器 G_{c2w})以及两个对应的图像领域判别器(水下图像领域判别器 D_w 和高质量图像领域判别器 D_c).其中,生成器 G_{w2c} 将输入的低质量水下图像增强为高质量水下图像.反之, G_{c2w} 则将输入的高质量水下图像转换为低质量水下图像.生成器 G_{w2c} 和 G_{c2w} 构成循环一致的关系.高质量图像领域判别器 D_c 用于区分生成器 G_{w2c} 生成的高质量水下图像和原始的高质量水下图像.而水下图像领域判别器 D_w 则用于区分生成器 G_{c2w} 生成的水下图像和原始的水下图像.本文采用马尔科夫性质判别器^[17],基于全卷积的网络结构对图像的各局部区域进行判断.通过引入该判别器,可以确保增强图像呈现高对比度和丰富的纹理细节,并防止伪影及噪声的出现.两组生成器与判别器之间通过循环一致性地生成对抗,最终实现了在无需成对水下图像和参考图像的前提下生成高质量的水下增强图像.

3.2 高低频双分支生成器

水下图像的失真类型主要包括颜色失真、亮度失真、细节纹理失真等,如果仅针对单一类型失真进行处理,可能会导致生成的水下增强图像存在遗留失真问题.在此基础上,本文观察到特征的低频部分主要包含颜色和亮度信息,而高频部分主要包含细节纹理信息,

同时高频部分的信息数值通常较小. 鉴于高频部分和低频部分呈现不同的特点, 本文提出对其进行分频式学习. 具体而言, 本文提出了高低频双分支生成器(如图4所示)对水下图像特征的低频和

失真类型. 网络包括小波池化层、低频处理分支(上)、高频处理分支(下)、小波反池化层及若干卷积层. 其中, 低频处理分支包含降采样模块、上采样模块和3个不同尺度上的低频注意力机制, 高频处理分支包括3层并行的高频注意力机制.

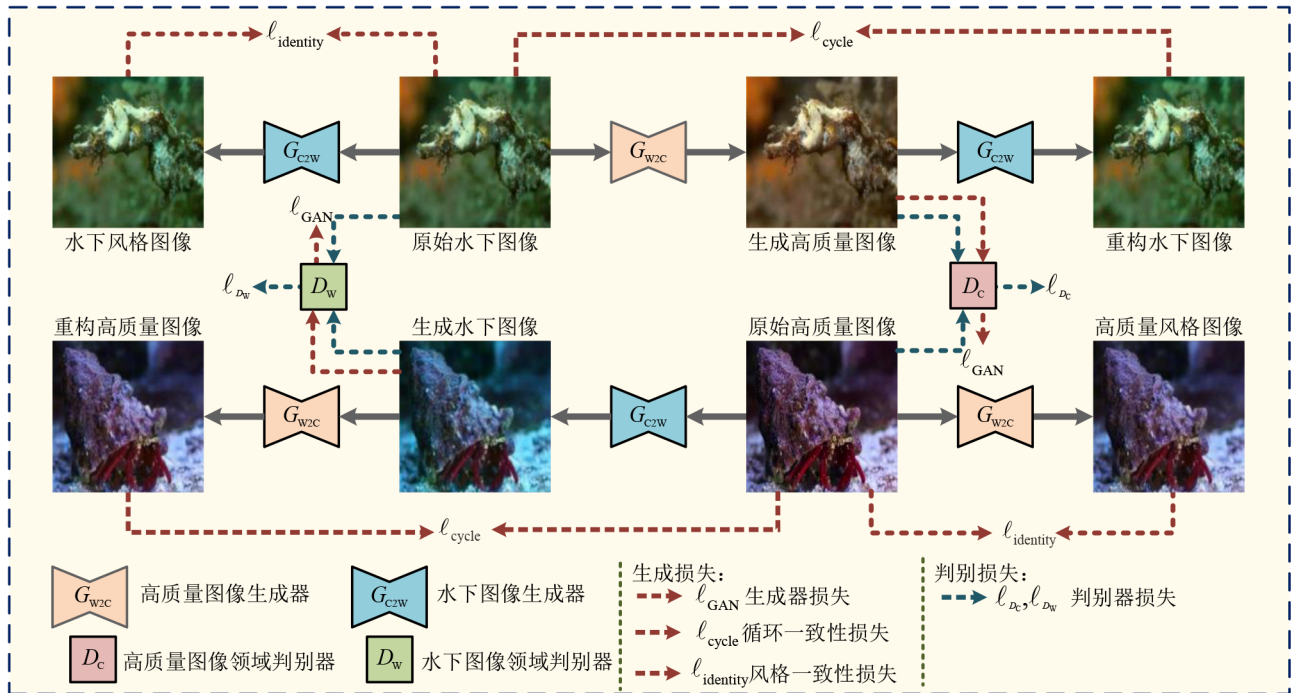


图3 分频式生成对抗网络框架图

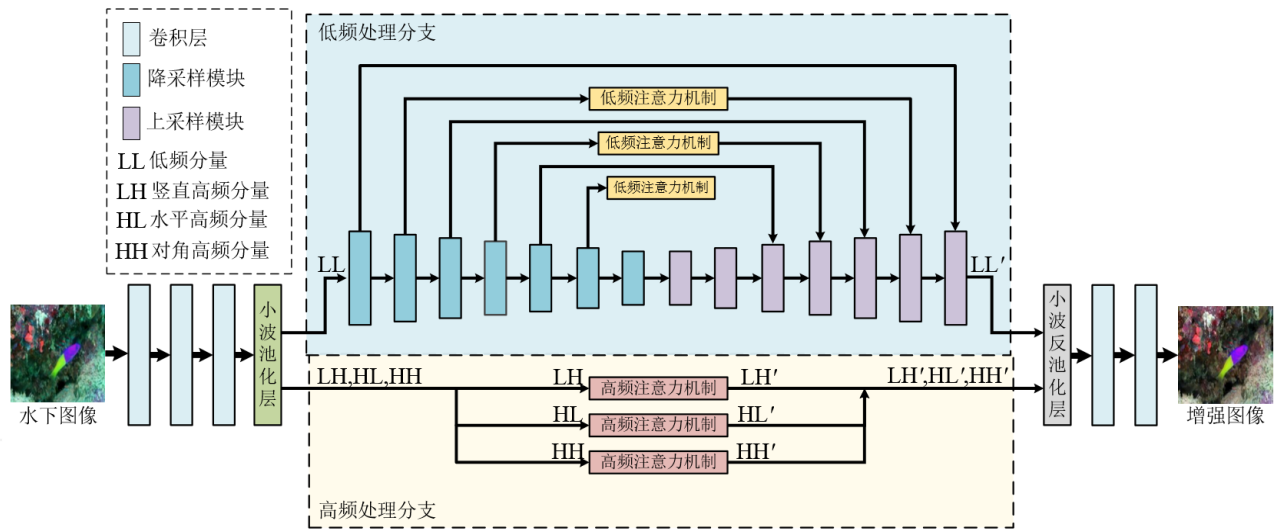


图4 高低频双分支生成器结构图

鉴于小波变换^[41]在计算机视觉领域有较广泛的应用, 可以将图像分解为不同频域的子带, 从而提取出图像的颜色、亮度和纹理等特征. 因此, 本文在小波变换理论的基础上引入小波池化层^[18], 根据水平和垂直方向上的低频与高频分量, 通过4个分解核将特征分解为

水平和垂直方向上的低频分量(LL)、水平方向低频和垂直方向高频的分量(垂直高频分量, LH)、水平方向高频和垂直方向低频的分量(水平高频分量, HL)以及水平和垂直方向均为高频的分量(对角高频分量, HH). 其中, 特征的低频部分(LL)包含颜色和亮度信息; 高频

部分(LH、HL和HH)包含细节纹理信息. 由于小波变换可以将特征的低频域和高频域进行分离, 高频部分中还包含细化的3类高频分量, 使得高频部分更易于被充分且针对性地学习. 于是, 本文使用的4个分解核中的低频滤波核和低频滤波核可以分别表示为

$$F_L^T = \frac{1}{\sqrt{2}}[1, 1], F_H^T = \frac{1}{\sqrt{2}}[-1, 1] \quad (1)$$

其中, F_L^T 和 F_H^T 分别为低频和高频滤波核.

另外, 小波池化层中的4个分量LL、LH、HL及HH所对应的分解核 K_{LL} 、 K_{LH} 、 K_{HL} 和 K_{HH} 分别表示为

$$\begin{aligned} K_{LL} &= \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, & K_{LH} &= \frac{1}{2} \begin{bmatrix} -1 & 1 \\ -1 & 1 \end{bmatrix}, \\ K_{HL} &= \frac{1}{2} \begin{bmatrix} -1 & -1 \\ 1 & -1 \end{bmatrix}, & K_{HH} &= \frac{1}{2} \begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix} \end{aligned} \quad (2)$$

如图5所示, 输入的水下图像(a)的原始特征被低频和高频处理分支学习并处理后, 得到增强后的低频部分(i)和高频部分(j)~(l), 网络通过小波反池化层聚合并生成融合后的增强特征(h), 进一步经过卷积神经网络对特征进行整合, 最后输出高质量的水下图像(g). 其中, 小波反池化层使用了和小波池化层相同的分解核, 对低频与高频4个特征分量进行反池化操作,

再对其进行融合, 进而得到融合特征.

本文提出的基于小波变换的高低频双分支生成器可以表示为

$$LL, LH, HL, HH = \text{WavePool}(\text{Conv}_{3 \times 3}^3(I_W)) \quad (3)$$

$$LL' = \text{LFPB}(LL) \quad (4)$$

$$LH', HL', HH' = \text{HFPB}(LH, HL, HH) \quad (5)$$

$$I_E = \text{Conv}_{3 \times 3}^2(\text{WaveUnpool}(LL', LH', HL', HH')) \quad (6)$$

其中, I_W 和 I_E 分别表示输入高低频双分支生成器的低质量水下图像和输出的增强图像; LL表示小波特征分解得到的特征的低频分量, LH、HL和HH则表示小波特征分解得到特征的3个方向的高频分量; LL'表示经过低频处理分支处理的低频增强分量, LH'、HL'和HH'则表示经过高频处理分支增强的3种高频增强分量; WavePool和WaveUnpool分别对应小波池化层和小波反池化层; LFPB表示低频处理分支, HFPB表示高频处理分支. $\text{Conv}_{3 \times 3}^n$ 表示经过 n 次 3×3 卷积操作. 式(3)表示小波变化前使用的3层卷积串连的网络结构, 用来对输入图像进行初步特征提取; 式(6)则表示小波变换后使用的是2层卷积串连的网络结构, 由小波反池化层的输出特征得到结果图像.

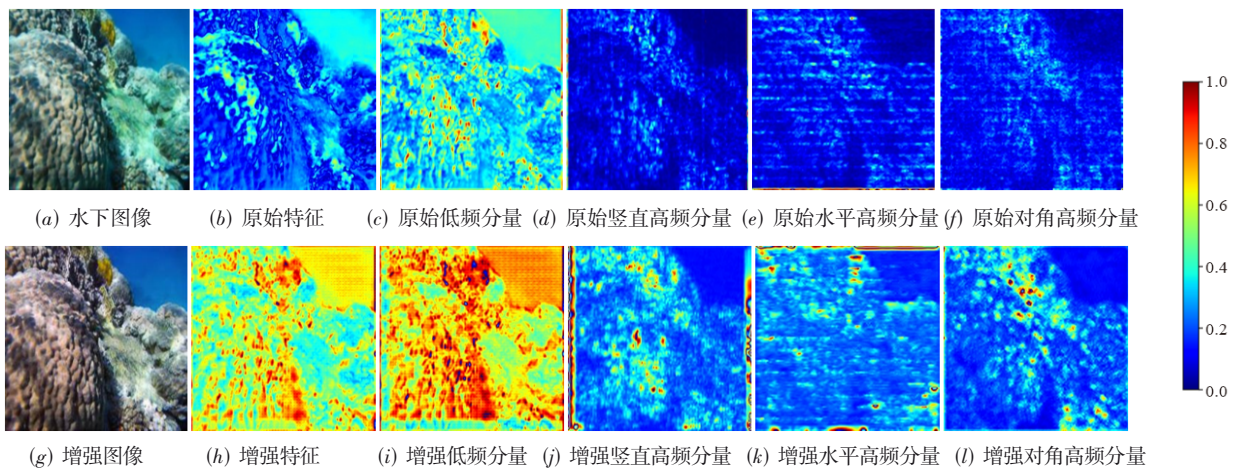


图5 特征级别小波变换得到的特征的低频和高频分量及其增强结果的可视化

3.3 低频处理分支

特征图的低频部分反映了特征值变换相对平缓的区域, 主要对应图像的低频颜色和亮度信息. 为此, 本文构建了低频处理分支以增强水下图像的颜色和亮度. 进一步地, 为了实现对特征的充分学习和重建, 本文将低频处理分支设计为一个结合低频注意力机制的7层编码-解码结构, 以增强特征的低频部分. 在低频处理分支中, 通过对特征图进行7次降采样和上采样操作, 可以在多个尺度上有效提取并调整颜色和亮度信息. 同时, 通过跳跃连接操作, 在采样过程中丢失的信

息得到了有效的补充. 此外, 为了加强对低频部分颜色信息的处理, 本文在第2层、第4层和第6层的跳跃连接处结合低频注意力机制, 在多个尺度上对低频分量进行增强. 其中, 降采样模块由卷积层、归一化层和激活函数层依次串联组成, 而上采样层由反卷积层、归一化层和激活函数层依次串联组成.

为了更好地校正低频部分的颜色和亮度信息, 本文设计了低频注意力机制(Low Frequency Attention Mechanism, LFAM). 低频注意力机制对输入特征的低频分量进行了多尺度的增强, 同时通过引入通道注意

力机制对增强后的特征进行了通道维度上的重加权,以获得增强后的颜色和亮度特征.

如图6所示,低频注意力机制主要包括扩展尺度学习模块和通道注意力层.其中,扩展尺度学习模块从多个尺度学习低频分量,由三支多尺度卷积块并联组成,每个分支中采用1个 1×1 卷积和不同数量的 3×3 卷积,生成不同尺度的特征,然后对多个尺度的信息进行特征融合.之后再通过通道注意力层(Squeeze-and-Excitation Layer, SELayer)^[42]对扩展尺度学习模块输出的特征进行通道维度上的增强.低频注意力机制计算公式为

$$L_1 = \text{ReLU_Conv}_{1 \times 1}(L_{in}) \quad (7)$$

$$L_2 = \text{ReLU_Conv}_{3 \times 3}(\text{ReLU_Conv}_{1 \times 1}(L_{in})) \quad (8)$$

$$L_3 = \text{ReLU_Conv}_{3 \times 3}^2(\text{ReLU_Conv}_{1 \times 1}(L_{in})) \quad (9)$$

$$L_{out} = \text{SELayer}(\text{Conv}_{3 \times 3}(\text{Cat}(L_1, L_2, L_3))) \quad (10)$$

其中, L_{in} 表示输入低频注意力机制的特征, L_{out} 表示对应的输出特征,ReLU表示修正线性单元的激活函数, $\text{Conv}_{k \times k}$ 表示 $k \times k$ 卷积,ReLU_Conv表示卷积层后使用激活函数,ReLU_Convⁿ表示对ReLU_Conv结构进行n次堆叠,Cat表示特征按通道拼接操作,SELayer表示通道注意力层.

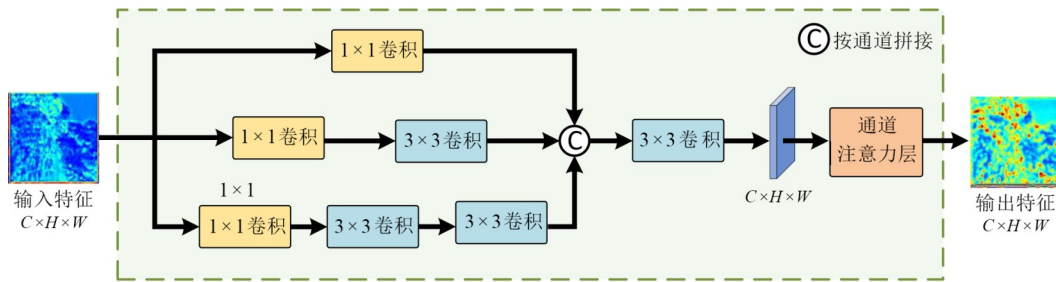


图6 低频注意力机制结构图

3.4 高频处理分支

特征的高频部分主要包含特征图的边缘细节信息,其数值范围通常较小,因此需要与低频部分进行区别学习.针对特征的高频部分,我们设计了基于并行高频注意力机制的高频处理分支,以增强特征的高频部分表示,进而提升水下图像的边缘和轮廓的可视质量,同时去除原始图像中的模糊和噪声现象.高频处理分支如图4下半部分所示,即对小波池化层细分出的高频分量LH、HL和HH分别进行自适应处理.由于3个高频分量所包含的特征信息互不相同(LH代表垂直高频分量,HL代表水平高频分量,HH代表对角高频分量),本文设计了三支并行的高频注意力机制,用于分别增强小波池化层提取的LH、HL和HH这3个高频分量.最后,该模块输出增强后的3个高频分量,进而与已增强的低频分量进行特征级别的融合.

本文设计了高频注意力机制(High Frequency Attention Mechanism, HFAM)以进一步增强细节信息.通过对高频信息进行多尺度的学习和融合,再结合空间注意力与通道注意力对特征进行空间维度和通道维度上的重加权,从而获得边缘和细节增强后的特征表示.

如图7所示,高频注意力机制主要包括扩展尺度学习模块、空间注意力层和通道注意力层.高频处理分支中没有采用降采样模块,因此为了获得更大的感受野,高频注意力机制中的扩展尺度学习模块全部由 3×3 卷积组成,从而对已提取的高频分量进行多尺度的加强学习.之后,对多尺度特征进行特征级别融合操作,再

经由空间注意力层(Spatial Attention Layer, SALayer)^[43]和通道注意力层对特征进行空间和通道双维度的图像特征强度重加权.与低频注意力机制不同,高频注意力机制中增加了空间注意力层,用于进一步增强空间中有用的细节纹理特征,以获得更有效的多尺度融合的高频分量表示.提出的高频注意力机制可表示为

$$H_1 = \text{ReLU_Conv}_{3 \times 3}(H_{in}) \quad (11)$$

$$H_2 = \text{ReLU_Conv}_{3 \times 3}^2(H_{in}) \quad (12)$$

$$H_3 = \text{ReLU_Conv}_{3 \times 3}^3(H_{in}) \quad (13)$$

$$H_4 = \text{ReLU_Conv}_{3 \times 3}(H_{in}) \quad (14)$$

$$H_{out} = \text{SELayer}(\text{SALayer}(\text{Conv}_{3 \times 3}(\text{Cat}[H_1, H_2, H_3]))) \quad (15)$$

其中, H_{in} 表示高频注意力机制的输入特征, H_{out} 表示对应的输出特征,SALayer表示空间注意力层,SELayer表示通道注意力层.

3.5 损失函数

本文综合使用生成损失和判别损失对生成器模型和判别器模型进行监督.采用的生成损失函数定义为

$$\ell_{total} = \lambda_1 \cdot \ell_{GAN} + \lambda_2 \cdot \ell_{cycle} + \lambda_3 \cdot \ell_{identity} \quad (16)$$

其中, ℓ_{GAN} 、 ℓ_{cycle} 和 $\ell_{identity}$ 分别表示生成器损失、循环一致性损失和风格一致性损失,而 λ_1 、 λ_2 和 λ_3 分别表示对应各项损失函数的加权系数.

具体地,式(16)中生成器损失函数的计算公式为

$$\ell_{GAN} = L_{MSE}(1, D_c(G_{w2c}(I_w))) + L_{MSE}(1, D_w(G_{c2w}(I_c))) \quad (17)$$

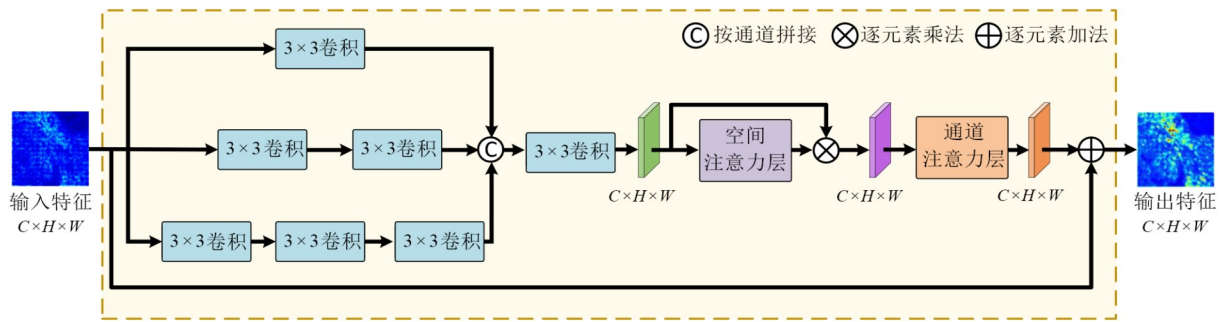


图7 高频注意力机制结构图

其中, L_{MSE} 表示均方误差损失函数, I_W 为输入的原始水下图像, I_C 为输入的高质量图像, 判别器 $D_C(\cdot)$ 用于对生成器 $G_{W2C}(\cdot)$ 生成的增强图像 $G_{W2C}(I_W)$ 进行领域判别, 判别器 $D_W(\cdot)$ 则用于对生成器 $G_{C2W}(\cdot)$ 生成的水下图像 $G_{C2W}(I_C)$ 进行领域判别。

为了保证生成器模型训练的稳定性, 同时为图像领域的迁移提供监督信息, 本文进一步引入循环一致性损失, 具体为

$$\begin{aligned} \ell_{\text{cycle}} = & L_1(G_{C2W}(G_{W2C}(I_W)), I_W) \\ & + L_1(G_{W2C}(G_{C2W}(I_C)), I_C) \end{aligned} \quad (18)$$

其中, L_1 表示 L1 损失, I_W 为输入的原始水下图像, 而 I_C 为输入的高质量图像。值得注意的是, 这两个图像无对应关系, 是不成对图像。本文使用原始水下图像的重构图像 $G_{C2W}(G_{W2C}(I_W))$ 和原始高质量图像的重构图像 $G_{W2C}(G_{C2W}(I_C))$ 分别作为原始水下图像和原始高质量图像的参考, 实现循环一致性监督。

接着, 为保证生成器学习到目标领域的分布, 本文引入风格一致性损失, 具体为

$$\ell_{\text{identity}} = L_1(G_{W2C}(I_C), I_C) + L_1(G_{C2W}(I_W), I_W) \quad (19)$$

上述的风格一致性损失用于确保当输入的图片为目标领域的图像时, 生成器不改变图像的风格。比如将高质量图像 I_C 输入到高质量图像生成器 $G_{W2C}(\cdot)$ 时, 生成器的输出要与 I_C 接近。

本文设计的两个判别器主要用于判断输入的生成图像是否属于各自的数据分布领域, 本文定义判别损失计算公式为

$$\ell_D = \ell_{D_C} + \ell_{D_W} \quad (20)$$

其中, ℓ_{D_C} 表示高质量图像领域判别器 $D_C(\cdot)$ 的判别器损失, 而 ℓ_{D_W} 表示水下图像领域判别器 $D_W(\cdot)$ 的判别器损失, 具体的损失函数计算公式为

$$\ell_{D_C} = L_{MSE}(0, D_C(G_{W2C}(I_W))) + L_{MSE}(1, D_C(I_C)) \quad (21)$$

$$\ell_{D_W} = L_{MSE}(0, D_W(G_{C2W}(I_C))) + L_{MSE}(1, D_W(I_W)) \quad (22)$$

本文对生成器和判别器进行交替对抗训练, 生成

器致力于生成更加真实的数据分布, 而判别器致力于判断输入图像是否服从该领域分布。生成器和判别器交替对抗直至训练收敛且达到稳定的纳什平衡。最后, 将 $G_{W2C}(\cdot)$ 作为最终的水下图像增强模型。

4 实验

4.1 数据集和实现细节

4.1.1 数据集

在对数据集及对比方法进行选择时, 本文尽可能涵盖当前水下图像增强领域主流新近数据集和方法, 从而保证实验的可重复性和公平性。首先, 采用水下图像增强数据集 EUVP^[9] 构建非成对训练集。使用训练好的模型直接在 UFO120、UIEB、RUIE 和 Challenge60 数据集上进行测试, 以验证本文方法的有效性和泛化性。其中, RUIE 和 Challenge60 是真实数据集, 且没有提供参考图像。EUVP 是水下图像增强任务常用的数据集, 包含图像的失真类型较多, 且数据量大, 共包含 11 435 对水下图像和参考图像训练集、非成对的 3 195 幅水下图像和 3 140 幅高质量图像训练集以及 515 对测试集。其次, 为了更全面地验证本文方法的创新性, 本文还引入陆上数据集 MIT-Adobe FiveK^[20] 作为可替换数据。该数据集是图像增强领域常用的陆上数据集。本文将该数据集中的图像按照不同比例替换了 EUVP 中的高质量水下图像, 以进行相关实验。

为模拟非成对水下图像的训练条件, 本文构建的训练数据集包含两部分, 即水下图像集合和高质量图像集合。为确保成对的水下图像和参考图像不会同时出现在两个集合中, 本文将 EUVP 数据集中成对训练集随机分为两组, 前一组中的水下图像加入水下图像集合, 后一组中的高质量图像加入高质量图像集合。然后, 本文将 EUVP 数据集中非成对训练集中的水下图像加入水下图像集合, 高质量图像加入高质量图像集合, 最终构建了一个包含 8 939 幅水下图像和 8 827 幅高质量图像的 EUVP 非成对水下图像训练集。

4.1.2 实验细节

本文在 NVIDIA GeForce RTX 2080 SUPER GPU 和

开源深度学习框架 PyTorch^[44]进行模型训练和性能测试. 本文在上述划分的 EUVP 非成对水下图像训练集上进行模型训练, 损失函数加权系数 λ_1 , λ_2 和 λ_3 分别设置为 1、10 和 5, 初始学习率设置为 2×10^{-4} . 本文采用 Adam^[45]算法优化损失函数, 共训练 200 轮数据迭代, 并将训练数据批次大小设置为 8, 在 100 轮数据迭代之后将学习率降至 1×10^{-6} , 以便网络参数收敛. 训练时, 本文将输入图像尺寸缩放至 256×256 , 并通过随机裁剪和随机翻转对数据进行增强.

4.2 实验结果

本文采用峰值信噪比 (Peak Signal to Noise Ratio, PSNR)、结构相似度指数 (Structural SIMilarity index, SSIM)、无参考评价指标 UCIQE (Underwater Color Image Quality Evaluation) 和 UIQM (Underwater Image Quality Measure) 作为水下图像增强结果的图像质量评价指标. 由于目前没有可以对比的其他非成对水下图像增强方法, 本文方法与当前先进的成对水下图像增强方法进行对比, 包括 WaterNet^[8]、FUInE-GAN^[9]、Deep SESR^[11]、Shallow UWnet^[10]、LANet^[14]、PUIE-Net^[15] 和 TUDA^[13].

4.2.1 水下图像增强性能对比实验结果

表 1 列出了本文方法 (包括 FD-GAN-0% 和 FD-GAN-50%) 与现有先进的成对水下图像增强方法在 EUVP^[9] 数据集上的对比结果, 其中 FD-GAN-0% 表示完全使用高质量水下图像训练, FD-GAN-50% 则表示将训练数据中 50% 的水下图像替换为陆上图像. 从表 1 中可以看出, 本文方法 FD-GAN-0% 在 SSIM 指标上取得了最优的结果, 与性能第 2 的 PUIE-Net 方法相比提升了 0.02. 在 PSNR 指标上, FD-GAN-0% 取得了第 2 的结果, 仅与性能最优的 Shallow UWnet 方法低 0.07 dB. 此外, 在 UCIQE 指标上, FD-GAN-0% 取得了最优的结果, 与性能第 2 的 PUIE-Net 方法相比取得了 0.29 的提升. 在 UIQM 指标上, FD-GAN-0% 也排名第 1. 值得关注的是, FD-GAN-50% 虽然用陆上图像替换了 50% 的高质量水下图像, 其性能依旧与 FD-GAN-0% 相当. 综上所述, 本文提出的方法在不依赖成对数据集的前提下, 可以在颜色、亮度和细节上较好地增强水下图像, 有效解决了原始水下图像中存在的失真问题.

表 1 不同方法在 EUVP、UFO120 和 UIEB 测试集上的指标结果

	方法	EUVP				UFO120		UIEB	
		PSNR \uparrow /dB	SSIM \uparrow	UCIQE \uparrow	UIQM \uparrow	PSNR \uparrow /dB	SSIM \uparrow	PSNR \uparrow /dB	SSIM \uparrow
P	WaterNet ^[8]	20.14 \pm 2.31	0.68 \pm 0.18	0.550	2.992	23.12 \pm 3.31	0.73 \pm 0.07	19.11 \pm 3.68	0.79 \pm 0.09
	FUInE-GAN ^[9]	26.22 \pm 3.07	0.79 \pm 0.06	0.556	3.079	24.32 \pm 2.19	0.74 \pm 0.05	20.36 \pm 3.17	0.69 \pm 0.09
	Deep SESR ^[11]	27.08 \pm 2.74	0.80 \pm 0.07	0.558	3.102	26.46 \pm 3.13	0.78 \pm 0.07	19.26 \pm 3.56	0.73 \pm 0.11
	Shallow UWnet ^[10]	27.46\pm2.34	0.80 \pm 0.04	0.576	3.124	27.80 \pm 2.34	0.79 \pm 0.05	21.58\pm4.62	0.80\pm0.10
	LANet ^[14]	27.27 \pm 2.31	0.81 \pm 0.04	0.582	3.175	27.82 \pm 3.21	0.81 \pm 0.05	21.25 \pm 3.56	0.81\pm0.11
	PUIE-Net ^[15]	27.35\pm2.94	0.82\pm0.04	0.598	3.199	27.93\pm2.34	0.82\pm0.06	21.32 \pm 4.21	0.81\pm0.09
	TUDA ^[13]	—	—	0.588	3.287	—	—	—	—
U	FD-GAN-0%	27.39\pm2.32	0.84\pm0.05	0.627	3.293	28.56\pm2.18	0.84\pm0.05	21.38\pm4.23	0.82\pm0.10
	FD-GAN-50%	27.18 \pm 2.21	0.83\pm0.06	0.619	3.148	28.25 \pm 2.07	0.84\pm0.04	21.67\pm3.95	0.81\pm0.04

注: “P” 标记成对水下图像增强方法, “U” 标记非成对水下图像增强方法, **加粗下划线/加粗/下划线** 分别表示性能第 1/第 2/第 3 的数据.

图 8 给出了不同水下图像增强方法的增强结果对比, 两个示例分别来自 EUVP 和 UFO120 数据集. 从第 1 个示例展示的细节区域可以看出, 各水下图像增强方法都提高了原始水下图像的对比度和亮度, 但是这些方法的增强图像仍存在细节模糊、颜色不均匀、亮度较低等问题. 具体而言, WaterNet 方法增强图像的整体亮度不足, 且图像色调偏灰; FUInE-GAN 方法的增强图像存在颜色分布不均匀的现象; Deep SESR 和 Shallow UWnet 方法的增强图像存在细节模糊的现象, 且图像颜色对比度较低; LANet 方法的增强图像存在细节失真现象; PUIE-Net 方法增强图像的整体色调偏蓝绿色. 同其他方法相比, 本文方法 FD-GAN-0% 所得到的水下图像增强结果不仅在颜色、亮度和对比度上与参考图像更为接近, 而且从原始水下图像中恢复了更多细节.

4.2.2 水下图像增强泛化性能对比实验结果

为验证本文提出的水下图像增强方法的跨域泛化能力, 本文将在 EUVP 非成对水下图像训练集上训练得到的水下图像增强模型直接应用到 UFO120 数据集、UIEB 数据集以及真实水下图像数据集 Challenge60 和 RUIE 上进行泛化性验证. 本文方法 (包括 FD-GAN-0% 和 FD-GAN-50%) 在 UFO120 数据集和 UIEB 数据集上与当前先进水下图像增强方法的对比结果如表 1 所示. 在 UFO120 数据集上, 本文方法 FD-GAN-0% 得到的水下增强图像在 PSNR 和 SSIM 指标上均取得了最优的效果, 较性能排第 2 位的方法分别取得了 0.63 dB 和 0.02 的较大幅度提升. 在 UIEB 数据集上, FD-GAN-0% 得到的增强图像在 SSIM 指标上取得了最优的性能, 在 PSNR 指标上取得了次优的性能. 特别地, FD-GAN-50% 不仅

在两个数据集上取得具有竞争力的性能,在 UIEB 数据集上还体现了超越 FD-GAN-0% 的 PSNR 指标性能. 在两个数据集上的综合表现,体现了本文方法在跨分布数据上优越的泛化性能.

图 8 第 2 个示例展示了不同方法在 UFO120 测试集上的增强结果以及对应的细节区域. 如图 8 所示, WaterNet 和 FUnIE-GAN 方法的增强结果整体偏黄色,与参考图像色差较大; Deep SESR, Shallow UWnet 和 LANet 方法的增强结果存在细节模糊现象; PUIE-Net 方法的增强结果与参考图像的颜色差异也较为明显. 相比之

下,本文方法 FD-GAN-0% 生成的增强图像的颜色和细节与参考图像更为相近.

表 2 列出了本文方法在真实数据集 Challenge60 和 RUIE 上与先进方法的对比结果. 如表 2 所示,本文方法 FD-GAN-0% 在两个真实图像数据集上的 UCIQE 指标均排名第 1,在 UIQM 指标上分别排名第 1 和第 3. 此外,尽管 FD-GAN-50% 将一半比例的高质量水下图像训练数据替换为陆上图像,在两个数据集上也体现出更好的 UCIQE 指标水平,而 UIQM 指标性能也具有竞争力.

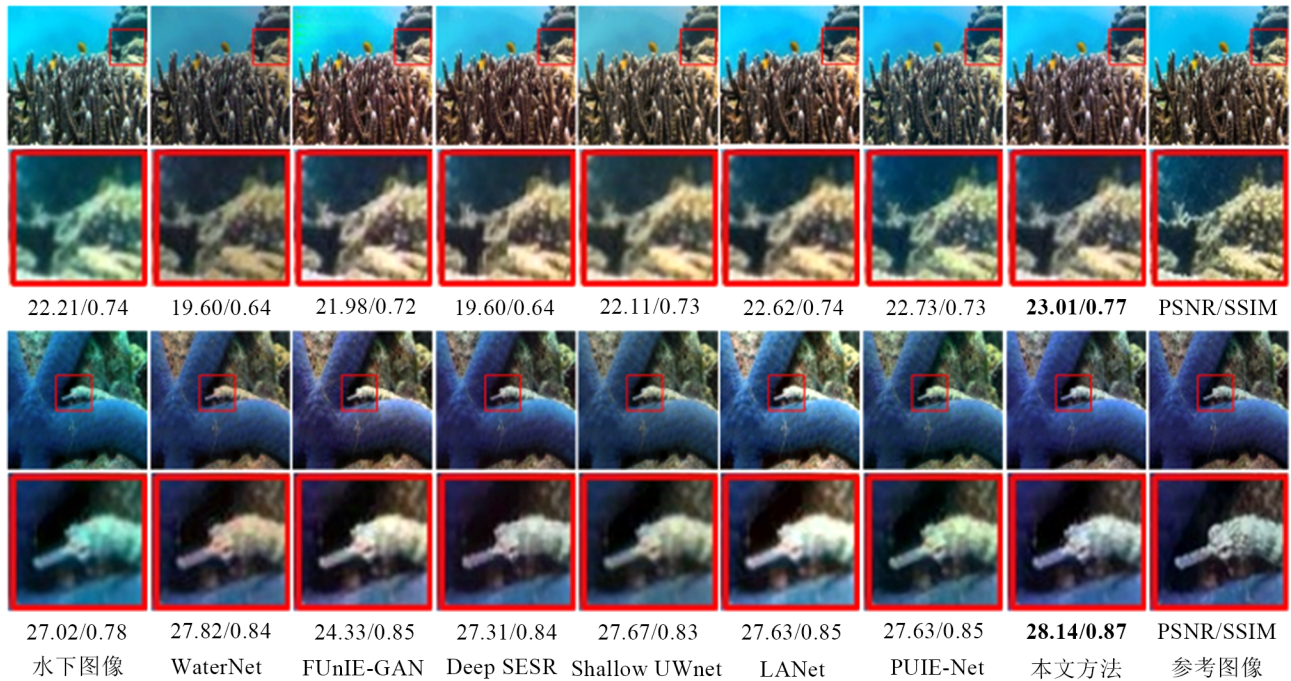


图 8 不同水下图像增强方法的增强结果及细节区域对比

表 2 不同方法在 Challenge60 和 RUIE 数据集上的无参考评价指标结果

	方法	Challenge60		RUIE	
		UCIQE ↑	UIQM ↑	UCIQE ↑	UIQM ↑
P	WaterNet ^[8]	0.549	2.288	0.538	2.877
	FUnIE-GAN ^[9]	0.556	2.335	0.562	3.032
	Deep SESR ^[11]	0.577	2.485	0.551	3.124
	Shallow UWnet ^[10]	0.601	2.561	0.569	3.298
	LANet ^[14]	0.599	2.594	0.576	3.394
	PUIE-Net ^[15]	<u>0.604</u>	<u>2.756</u>	<u>0.583</u>	<u>3.451</u>
U	FD-GAN-0%	<u>0.613</u>	<u>2.782</u>	<u>0.601</u>	<u>3.346</u>
	FD-GAN-50%	0.607	<u>2.729</u>	0.594	3.334

注:“P”标记成对水下图像增强方法,“U”标记非成对水下图像增强方法,加粗下划线/加粗/下划线分别表示性能第1/第2/第3的数据.

图 9 展示了不同方法在 Challenge60 和 RUIE 数据集上的增强结果. 如图 9 所示,其他对比方法的增强结果偏蓝绿色,且增强图像的亮度和对比度较低. 本文方法 FD-GAN-0% 的增强结果的亮度和对比度较高,具有

更好的视觉效果.

综上,在不依赖成对数据集的前提下,本文方法的跨域泛化性能优异,且在应对不同分布的水下图像失真类型时均有较好的表现.

4.2.3 陆上图像训练对比实验结果

为验证本文提出方法的实用性,本文采用陆上图像作为一部分原始高质量图像进行实验. 具体而言,本文使用陆上图像数据集 MIT-Adobe FiveK^[20]中的图像,替换了部分高质量水下图像进行模型的训练和测试,替换比例包括 25%、50%、75% 和 100%. 实验结果如

表 3 所示,其中陆上数据 0%、25%、50%、75% 和 100% 分别表示采用 0% 的陆上图像和 100% 的高质量水下图像、25% 的陆上图像和 75% 的高质量水下图像、50% 的陆上图像和 50% 的高质量水下图像、75% 的陆上图像和 25% 的高质量水下图像,以及 100% 的陆上图像和 0% 的高质量水下图像.

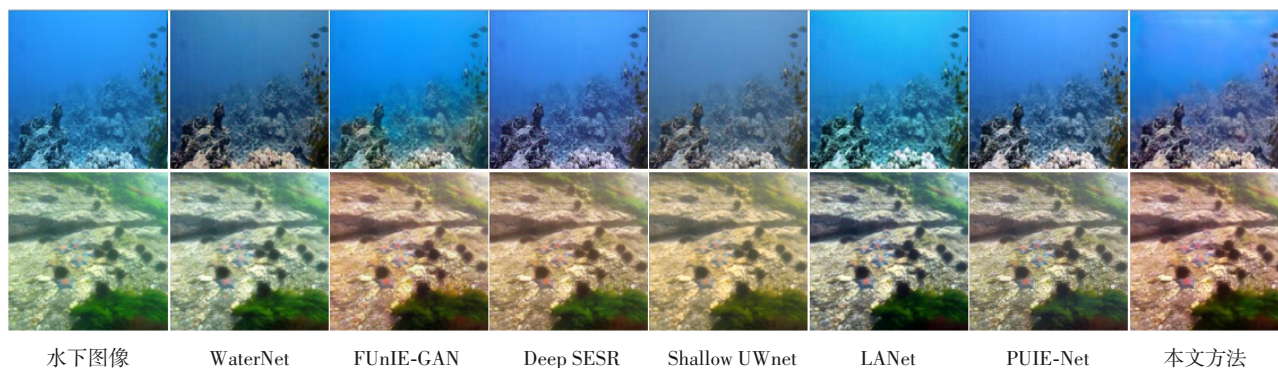


图 9 不同方法在 Challenge60 和 UIEB 数据集上的水下图像增强结果对比(第 1 行的示例来自 Challenge60 数据集,第 2 行的示例来自 UIEB 数据集)

表 3 不同比例陆上数据训练模型在 EUVP、UFO120 和 UIEB 测试集上的指标结果

	EUVP				UFO120		UIEB	
	PSNR ↑ /dB	SSIM ↑	UCIQE ↑	UIQM ↑	PSNR ↑ /dB	SSIM ↑	PSNR ↑ /dB	SSIM ↑
FD-GAN-0%	27.39±2.32	0.84±0.05	0.627	3.293	28.56±2.18	0.84±0.05	21.38±4.23	0.82±0.10
FD-GAN-25%	27.25±2.34	0.84±0.05	0.621	3.251	<u>28.27±2.02</u>	0.84±0.05	21.79±4.13	0.82±0.05
FD-GAN-50%	27.18±2.21	<u>0.83±0.06</u>	0.619	3.148	28.25±2.07	0.84±0.04	<u>21.67±3.95</u>	<u>0.81±0.04</u>
FD-GAN-75%	26.63±2.33	0.82±0.06	0.578	2.869	27.92±2.03	0.82±0.05	21.12±3.99	0.80±0.07
FD-GAN-100%	24.42±2.28	0.72±0.07	0.557	2.584	25.59±2.14	0.70±0.06	17.83±4.05	0.61±0.06

注:加粗/下划线分别表示性能第 1/第 2 的数据.

如表 3 所示,当用 25% 和 50% 的陆上图像替换高质量水下图像进行训练时,模型的性能改变较小. 特别是在 UIEB 测试集上,PSNR 指标分别提升了 0.41 dB 和 0.29 dB. 这说明当使用陆上图像替换部分高质量水下图像时,模型既能从高质量水下图像学习水下图像的特征,又能通过陆上图像从更多样的数据中进行学习,从而提升了模型的泛化能力. 然而,当使用 75% 和 100% 的陆上图像替换高质量水下图像进行训练时,由于陆上和两个域之间存在较大的差异,模型无法从足够多的水下图像中进行学习,从而导致性能出现显著下降. 此外,如图 10 所示,当使用 25% 和 50% 的陆上图像训练模型时,其增强结果与仅使用高质量水下图像训练模型的增强结果较为接近. 然而,当使用 75% 和 100% 陆上图像训练模型时,由于受到陆上图像的影响,增强结果中出现了颜色偏差和伪影(如图 10 中红色箭头所示). 另外,图 10(B)的图像从左往右分别是使用 0%、25%、50%、75% 和 100% 陆上图像进行模型训练的增强结果.

4.3 消融实验

在本节中,对所提出的方法及各模块的有效性进行消融验证. 为了确保公平性和可重复性,本文采用完整的高质量水下图像训练所提出的方法,即用 FD-GAN-0% 进行相应实验. 为了表示方便,我们将该方法称为 FD-GAN.

4.3.1 高低频双分支生成器模块消融实验

本小节对提出的高低频双分支生成器中的小波池化层和小波反池化层、低频注意力机制和高频注意力机制进行模块有效性的验证. 具体来说,本文采用了相同的训练策略在 EUVP^[9]数据集上进行训练,并对模型各模块的不同组合方式进行性能上的比较. 消融实验包括如下 6 组不同的模型:作为基准的 U-Net^[37]网络、去除小波池化层和反池化层的高低频双分支生成器(w/o WavePool)、去除高频注意力机制的高低频双分支生成器(w/o HFAM)、去除低频注意力机制的高低频双分支生成器(w/o LFAM)以及完整的高低频双分支生成器. 其中,去除小波池化层和反池化层的高低频双分支生成器先将水下图像特征分别输入低频处理模块和高

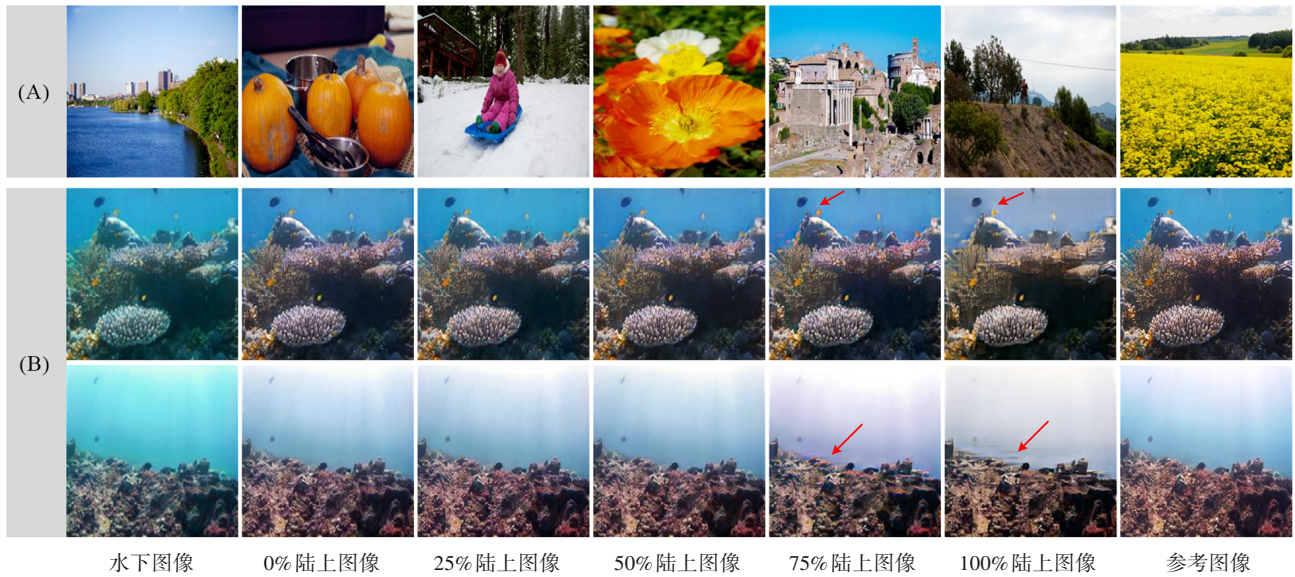


图 10 陆上图像示例(A)和不同比例陆上图像训练模型在 EUVP 数据集上的水下图像(B)增强结果对比

频处理模块,再将两个模块分别增强后的水下图像特征按通道拼接。

如表 4 第 2 行结果所示,去除小波池化层和反池化层后,方法性能出现显著下降,PSNR 降低了 0.96 dB,这说明对水下图像特征的高低频部分进行区分处理具有必要性。在第 3 行,本文讨论了使用高频注意力机制的必要性,可见在去除了高频注意力机制后,方法性能下降明显,PSNR 降低了 1.06 dB,这说明本文提出的高频注意力机制能够对图像特征的高频部分进行有效的增强,进而更充分地恢复水下图像细节。再如第 4 行所示,在去除低频注意力机制后,方法性能明显下降,PSNR 降低了 0.92 dB,这进一步验证了本文所提的低频注意力机制对水下图像的颜色和亮度进行了有效的修正。

表 4 高低频双分支生成器模块消融实验指标结果

生成器	PSNR ↑ /dB	SSIM ↑
U-Net	26.03±2.52	0.79±0.05
w/o WavePool	26.43±2.48	0.81±0.07
w/o HFAM	26.33±2.48	0.81±0.07
w/o LFAM	<u>26.47±2.43</u>	<u>0.82±0.06</u>
FD-GAN	27.39±2.32	0.84±0.05

注:加粗/下划线分别表示性能第 1/第 2 的数据。

为探究所提出的低频注意力机制在低频处理分支的编码-解码结构中所处位置的合理性,本文对低频注意力机制的位置进行了消融实验。具体来说,本文对低频处理分支中低频注意力机制的位置进行了不同设置,并基于此训练了两个对比模型:一个在低频处理分支中的所有跳跃连接处均设置低频注意力机制,另一个仅在低频处理分支中的第 1、3、5 层跳跃连接处设置

低频注意力机制。两个模型均训练收敛,得到的最终模型在 EUVP 测试集上的增强结果如表 5 所示。从第 1 行数据可以看出,在所有层均设置低频注意力机制时,模型性能低于仅在 1、3、5 层或仅在 2、4、6 层设置低频注意力机制,说明过多的低频注意力模块会导致原始低频信息的损失。再如第 3 行数据,在 2、4、6 层设置低频注意力机制模型性能最高,说明本文在低频处理分支配置低频注意力机制位置的合理性。

表 5 低频处理分支中低频注意力机制位置消融实验指标结果

位置	PSNR ↑ /dB	SSIM ↑
所有层	26.93±2.51	0.82±0.05
1,3,5 层	<u>27.13±2.35</u>	<u>0.83±0.06</u>
FD-GAN(2,4,6 层)	27.39±2.32	0.84±0.05

注:加粗/下划线分别表示性能第 1/第 2 的数据。

高频注意力机制内部各模块的消融实验。为了验证本文所提出的高频注意力机制内部各模块的有效性及其整体结构的合理性,本文对高频注意力机制的结构和各组成部分进行消融实验。具体来说,本文对高频注意力机制中的组成部分进行了不同结构上的组合,构建了如下 4 组模型:去除高频注意力机制的通道注意力层(w/o SELayer)、去除空间注意力层的高频注意力机制(w/o SALayer)、去除扩展尺度学习模块的高频注意力机制(w/o Multiscale)和将高频注意力机制中的空间和通道注意力替换为 CBAM^[46]机制(→CBAM)。4 个模型均训练至收敛,得到的最终模型在 EUVP 测试集上的增强结果如表 6 所示。

从表 6 第 1~3 行结果可见,去除通道注意力机制、空间注意力机制及扩展尺度学习模块后,模型性能指标均有所下降,其中 PSNR 分别下降 0.61 dB、0.93 dB 和

0.85 dB,这说明上述3个模块对增强特征的高频部分均有促进作用.

表6 高频注意力机制内部各模块消融实验指标结果

模型	PSNR ↑ /dB	SSIM ↑
w/o SELayer	26.78±2.44	<u>0.83±0.05</u>
w/o SALayer	26.46±2.45	0.81±0.07
w/o Multiscale	26.54±2.41	0.82±0.09
→CBAM	<u>26.98±2.42</u>	<u>0.83±0.06</u>
FD-GAN	27.39±2.32	0.84±0.05

注:加粗/下划线分别表示性能第1/第2的数据.

4.3.2 频域注意力机制消融实验

本节对所提的高低频注意力机制的设计以及高频处理模块中高频注意力机制的组合进行消融实验,以验证高频注意力机制和低频注意力机制不同结构的有效性.本文使用了相同的训练策略在EUVP^[9]数据集上进行训练,并对高频处理分支和低频处理分支的注意力机制搭配进行了不同的组合.为此,本文训练了如下4个组合对应的模型:将高频和低频处理模块中的注意力机制统一改为高频注意力机制(HFAM),将高频和低频处理模块中的注意力机制统一改为低频注意力机制(LFAM),将高频注意力机制和低频注意力机制统一改为CBAM机制(CBAM),将高频处理模块中的3个高频注意力机制并联的结构改为由统一的高频注意力机制处理(LFAM+1个HFAM).4个模型均训练至收敛,得到的最终模型在EUVP测试集上进行性能测试,增强结果的性能对比如表7所示.

从表7可以看出,当将高频和低频处理模块中的注意力机制统一改为高频注意力机制或低频注意力机制后,模型性能指标均有所下降,PSNR分别下降0.22 dB

表7 频域注意力机制消融实验指标结果

模型	PSNR ↑ /dB	SSIM ↑
HFAM	<u>27.17±2.87</u>	<u>0.83±0.06</u>
LFAM	26.98±3.01	<u>0.83±0.06</u>
CBAM	26.63±2.93	0.82±0.05
LFAM +1个HFAM	27.04±2.82	<u>0.83±0.06</u>
FD-GAN	27.39±2.32	0.84±0.05

注:加粗/下划线分别表示性能第1/第2的数据.

和0.41 dB,SSIM均下降0.01,这验证了本文针对高频和低频特征分量分别设计注意力机制的有效性.同样,当将高频注意力机制和低频注意力机制统一改为CBAM机制时,模型性能指标均下降,PSNR下降0.76 dB,SSIM指标下降0.02,这证明了本文提出的注意力机制的有效性.

4.3.3 频域消融实验

本节对特征级别小波变换所划分的各特征频域进行消融实验,以验证各频域分量对水下图像增强工作的有效性.具体来说,本文使用了相同的训练策略在EUVP^[9]数据集上进行训练,并对特征的低频部分和高频部分进行了各种组合上的探究.为此,本文训练了如下7个组合对应的模型:低频分量(LL)、低频分量及垂直高频分量(LL+LH)、低频分量及水平高频分量(LL+HL)、低频分量及对角线高频分量(LL+HH)、去除垂直高频分量(LL+HL+HH)、去除水平高频分量(LL+LH+HH)和去除对角线高频分量(LL+LH+HL).7个模型均训练至收敛,得到的最终模型在EUVP测试集上进行性能测试,增强结果的性能对比如表8所示,图像示例结果及对应细节区域如图11所示,其中包含增强图像的子区域与参考图像的子区域差值的可视化结果.

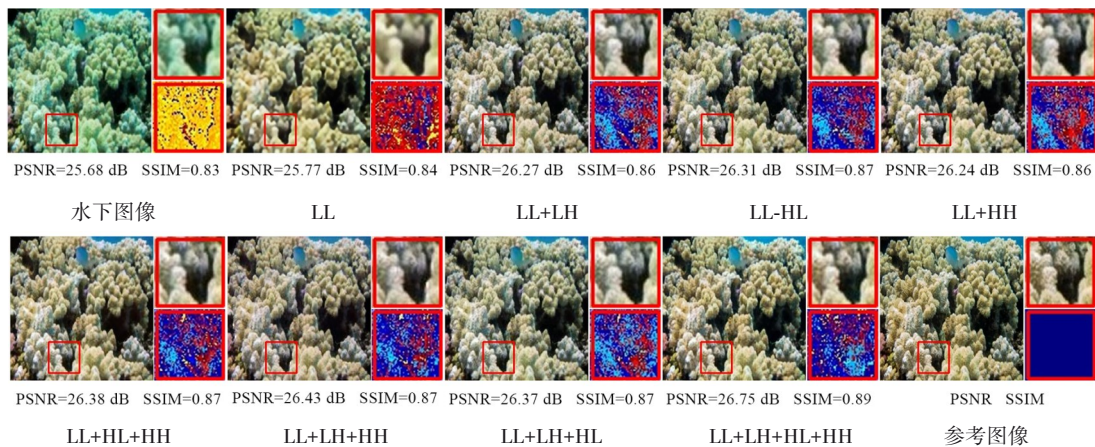


图11 不同频域组合消融实验结果对比

从表8可以看出,随着不同频域分量的增加,模型性能整体呈逐渐提高的趋势.其中,LL模型的性能最低,LL+LH、LL+HL和LL+HH模型的性能较LL模型有

所提升,LL+HL+HH、LL+LH+HH和LL+LH+HL模型的性能又有进一步提升,而本文使用LL+LH+HL+HH的完整模型性能最优.此外,如图11所示,对于LL模

型的增强结果,由于其只对特征的低频部分进行增强,虽然增强图像在颜色和亮度上有所改善,但与参考图像相比整体较为模糊.在图11中,除了增强结果和细节区域外,本文还给出了增强图像的细节区域与参考图像对应的细节区域的差值可视化结果.可以看到,LL模型在图像细节处与参考图像相比存在明显失真.这说明增强特征的低频部分对水下图像的颜色和亮度增强具有一定的促进作用,但生成高质量的水下图像还需要对细节边缘信息进行增强.

表8 频域消融实验指标结果

频域组合	PSNR ↑ /dB	SSIM ↑
LL	25.68±2.44	0.81±0.05
LL+LH	26.53±2.98	0.82±0.06
LL+HL	26.49±2.61	<u>0.83±0.09</u>
LL+HH	26.36±2.42	0.82±0.04
LL+HL+HH	26.78±2.81	<u>0.83±0.05</u>
LL+LH+HH	<u>26.86±2.84</u>	<u>0.83±0.09</u>
LL+LH+HL	26.63±2.98	<u>0.83±0.06</u>
FD-GAN	27.39±2.43	0.84±0.05

注:加粗/下划线分别表示性能第1/第2的数据.

表9 不同方法参数量和运行时间对比

模型	UGAN	FUnIE-GAN	TUDA	FD-GAN
参数量/M	58.17	7.020	<u>31.36</u>	69.4
运行时间/s	0.337	0.173	—	<u>0.329</u>

注:加粗/下划线分别表示性能第1/第2的数据.

对于LL和任意方向上的高频分量相融合模型(LL+LH、LL+HL、LL+HH),从对应的增强图像可以看出,在加入特征的高频部分后,增强图像的清晰度显著提升,细节上和参考图像也更加接近.这表明融合高频部分的特征可以有效增强水下图像的细节,并改善水下图像的清晰程度.对于融合LL和两个方向的高频分量的模型(LL+HL+HH、LL+LH+HH、LL+LH+HL),从对应的增强图像可以看出,在加入两个方向的高频分量后,增强图像更加接近参考图像,这体现了特征的各

个方向高频分量的叠加效果.此外,融合两个方向的高频分量相比于融合单一方向的高频分量得到的增强图像的效果更好,这也说明了每个方向的高频分量对整体图像增强有着各异的增强作用.对于融合所有特征频域的模型,可以看出其增强效果最好,且增强图像的细节部分最接近参考图像.这说明特征的每个频域对整体图像增强有着各异的增强作用,通过叠加所有频域分量会获得最好的水下图像增强效果.

为了验证消融实验中各模块消融结果的显著性意义,本文采用P值分析,比较了本文提出的FD-GAN的结果与各个消融结果之间的差异.在表4至表8所示的各模块消融结果中,所有的P值都 ≤ 0.01 ,说明这些消融结果的差异都具有显著性意义.具体来说,8个P值 < 0.001 ,32个P值在0.01~0.001之间,2个P值=0.01.这些P值进一步表明,对于各个消融模块,其结果与FD-GAN的结果之间的差异是显著的.

4.4 讨论

本节首先基于图像分割算法,采用任务驱动的评价方式,对水下图像增强算法的性能进行验证,然后对所提方法的局限性进行讨论.

如图12所示,原始水下图像直接使用图像分割算法^[47]不能较好地分割出红色矩形框中的物体,而采用本文方法增强后的水下图像可以较好地分割出红色矩形框中的物体和背景分离.与其他水下图像增强方法如Deep SESR^[11]、Shallow UWnet^[10]、LANet^[14]以及PUIE-Net^[15]的增强图像相比,本文方法增强图像的分割结果与参考图像的分割结果更为接近,显示了本文方法在后续的计算机视觉应用中的潜能.

此外,本文提出的方法还具有一定的局限性.首先,在应用域的差异方面,虽然本文方法使用陆上图像进一步降低了对高质量水下图像的需求,但由于陆上和两个域之间存在较大的差异,本文方法无法利用大规模陆上数据进行训练.其次,在模型参数量方面,如表9所示,本文提出模型的参数量和运行速度与UGAN^[7]相当,但是与FUnIE-GAN^[9]和TUDA^[13]方法相

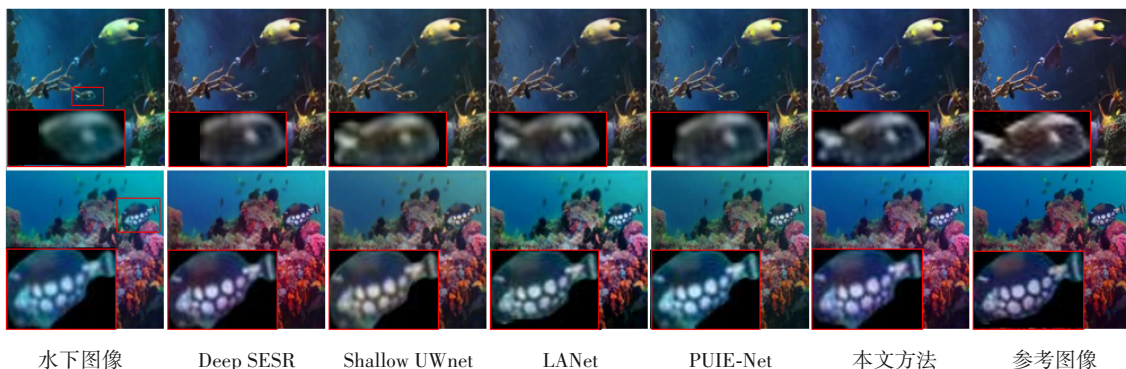


图12 不同水下图像增强结果的分割效果展示

比,本文提出的FD-GAN参数量偏大.我们将在未来的工作中进一步优化模型,以改进在参数量和运行时间方面的局限性.

5 结论

为了降低对成对训练数据的依赖,同时考虑到水下图像增强任务中复杂且多样的水下图像失真类型,本文提出了一种基于分频式生成对抗网络的非成对水下图像增强方法,并采用循环一致性思想构建了框架.本文通过对水下图像特征进行分频域处理,提出了高低频双分支生成器,并进一步针对特征的低频和高频部分构建对应的低频处理分支和高频处理分支,增强了水下图像在不同视觉特性上的表现.本文方法不仅可以在无需成对水下图像和参考图像的条件下实现模型训练,还可以实现对水下图像的多种类型的失真进行有效的针对性增强,以生成高质量的水下增强图像.大量的实验结果表明,本文方法不仅能获得比当前先进的成对水下图像增强方法更好的性能,还具有更好的泛化能力.此外,本文提出的方法在使用非成对的高质量水下图像和引入部分陆上图像的情况下,均能有效生成高质量的水下增强图像.消融实验也进一步验证了本文方法的各个模块的有效性和必要性.在今后的工作中,我们将探讨如何利用更大规模的陆上数据显著降低对高质量水下图像的需求.

参考文献

- [1] IQBAL K, ODETAYO M, JAMES A, et al. Enhancing the low quality images using unsupervised colour correction method[C]//2010 IEEE International Conference on Systems, Man and Cybernetics. Piscataway: IEEE, 2010: 1703-1709.
- [2] 刘柯, 李旭健. 水下和微光图像的去雾及增强方法[J]. 光学学报, 2020, 40(19): 73-85.
LIU K, LI X J. De-hazing and enhancement methods for underwater and low-light images[J]. Acta Optica Sinica, 2020, 40(19): 73-85. (in Chinese)
- [3] 牛玉贞, 陈铭铭, 李悦洲, 等. 基于任务解耦的低照度图像增强方法[J]. 电子学报, 2024, 52(1): 34-45.
NIU Y Z, CHEN M M, LI Y Z, et al. Task decoupling guided low-light image enhancement[J]. Acta Electronica Sinica, 2024, 52(1): 34-45. (in Chinese)
- [4] 肖进胜, 申梦瑶, 雷俊锋, 等. 基于生成对抗网络的雾霾场景图像转换算法[J]. 计算机学报, 2020, 43(1): 165-176.
XIAO J S, SHEN M Y, LEI J F, et al. Image conversion algorithm for haze scene based on generative adversarial networks[J]. Chinese Journal of Computers, 2020, 43(1): 165-176. (in Chinese)
- [5] 徐少平, 陈孝国, 李芬, 等. 采用两阶段混合策略实现的低照度图像增强算法[J]. 电子学报, 2021, 49(11): 2166-2170.
XU S P, CHEN X G, LI F, et al. A low-light image enhancement algorithm using two-stage hybrid strategy[J]. Acta Electronica Sinica, 2021, 49(11): 2166-2170. (in Chinese)
- [6] LI J, SKINNER K A, EUSTICE R M, et al. WaterGAN: Unsupervised generative network to enable real-time color correction of monocular underwater images[J]. IEEE Robotics and Automation Letters, 2018, 3(1): 387-394.
- [7] FABBRI C, ISLAM M J, SATTAR J. Enhancing underwater imagery using generative adversarial networks[C]//2018 IEEE International Conference on Robotics and Automation (ICRA). Piscataway: IEEE, 2018: 7159-7165.
- [8] 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.
- [9] 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.
- [10] NAIK A, SWARNAKAR A, MITTAL K. Shallow-UWnet: Compressed model for underwater image enhancement (student abstract)[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2021, 35(18): 15853-15854.
- [11] ISLAM M J, LUO P G, SATTAR J. Simultaneous enhancement and super-resolution of underwater imagery for improved visual perception[EB/OL]. (2020-02-04)[2024-03-24]. <https://arxiv.org/abs/2002.01155v1>.
- [12] LI C Y, ANWAR S, PORIKLI F. Underwater scene prior inspired deep underwater image and video enhancement[J]. Pattern Recognition, 2020, 98: 107038.
- [13] WANG Z Y, SHEN L Q, XU M, et al. Domain adaptation for underwater image enhancement[J]. IEEE Transactions on Image Processing, 2023, 32: 1442-1457.
- [14] 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.
- [15] FU Z Q, WANG W, HUANG Y, et al. Uncertainty inspired underwater image enhancement[M]//Lecture Notes in Computer Science. Cham: Springer Nature Switzerland, 2022: 465-482.
- [16] GOODFELLOW I, POUGET-ABADIE J, MIRZA M, et al. Generative adversarial networks[J]. Communications of the ACM, 2020, 63(11): 139-144.
- [17] 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). Piscataway: IEEE, 2017: 2242-2251.
- [18] LI B, CHEN X F. Wavelet-based numerical analysis: A review and classification[J]. *Finite Elements in Analysis and Design*, 2014, 81: 14-31.
- [19] 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.
- [20] BYCHKOVSKY V, PARIS S, CHAN E, et al. Learning photographic global tonal adjustment with a database of input/output image pairs[C]//2011 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2011: 97-104.
- [21] ABDUL GHANI A S, MAT ISA N A. Underwater image quality enhancement through integrated color model with rayleigh distribution[J]. *Applied Soft Computing*, 2015, 27: 219-230.
- [22] ANCUTI C O, ANCUTI C, DE VLEESCHOUWER C, et al. Color balance and fusion for underwater image enhancement[J]. *IEEE Transactions on Image Processing*, 2018, 27(1): 379-393.
- [23] ANCUTI C O, ANCUTI C, DE VLEESCHOUWER C, et al. Color channel compensation (3C): A fundamental preprocessing step for image enhancement[J]. *IEEE Transactions on Image Processing*, 2020, 29: 2653-2665.
- [24] FU X Y, ZHUANG P X, HUANG Y, et al. A retinex-based enhancing approach for single underwater image[C]//2014 IEEE International Conference on Image Processing (ICIP). Piscataway: IEEE, 2014: 4572-4576.
- [25] ZHANG S, WANG T, DONG J Y, et al. Underwater image enhancement via extended multi-scale Retinex[J]. *Neurocomputing*, 2017, 245: 1-9.
- [26] XIE B, GUO F, CAI Z X. Improved single image dehazing using dark channel prior and multi-scale retinex[C]//2010 International Conference on Intelligent System Design and Engineering Application. Piscataway: IEEE, 2010: 848-851.
- [27] CHIANG J Y, CHEN Y C. Underwater image enhancement by wavelength compensation and dehazing[J]. *IEEE Transactions on Image Processing*, 2012, 21(4): 1756-1769.
- [28] DREWS P L J JR, NASCIMENTO E R, BOTELHO S S C, et al. Underwater depth estimation and image restoration based on single images[J]. *IEEE Computer Graphics and Applications*, 2016, 36(2): 24-35.
- [29] PENG Y T, CAO K M, COSMAN P C. Generalization of the dark channel prior for single image restoration[J]. *IEEE Transactions on Image Processing*, 2018, 27(6): 2856-2868.
- [30] BERMAN D, LEVY D, AVIDAN S, et al. Underwater single image color restoration using haze-lines and a new quantitative dataset[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2021, 43(8): 2822-2837.
- [31] JIANG Y F, ZHANG H, ZHANG J M, et al. SSH: A self-supervised framework for image harmonization[C]//2021 IEEE/CVF International Conference on Computer Vision (ICCV). Piscataway: IEEE, 2021: 4812-4821.
- [32] 陈勇, 陈东, 刘焕淋, 等. 基于深度卷积神经网络的无参考低照度图像增强[J]. *电子与信息学报*, 2022, 44(6): 2166-2174.
- CHEN Y, CHEN D, LIU H L, et al. Unreferenced low-lighting image enhancement based on deep convolutional neural network[J]. *Journal of Electronics & Information Technology*, 2022, 44(6): 2166-2174. (in Chinese)
- [33] DAN R T, MOTIIAN S, FAIETA B, et al. ALADIN: All layer adaptive instance normalization for fine-grained style similarity[C]//2021 IEEE/CVF International Conference on Computer Vision (ICCV). Piscataway: IEEE, 2021: 11906-11915.
- [34] NESHATAVAR R, YAVARTANOO M, SON S, et al. CVF-SID: Cyclic multi-variate function for self-supervised image denoising by disentangling noise from image[C]//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2022: 17562-17570.
- [35] ANCUTI C, ANCUTI C O, HABER T, et al. Enhancing underwater images and videos by fusion[C]//2012 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2012: 81-88.
- [36] GUO Y C, LI H Y, ZHUANG P X. Underwater image enhancement using a multiscale dense generative adversarial network[J]. *IEEE Journal of Oceanic Engineering*, 2020, 45(3): 862-870.
- [37] RONNEBERGER O, FISCHER P, BROX T. U-Net: Convolutional networks for biomedical image segmentation[M]// *Lecture Notes in Computer Science*. Cham: Springer International Publishing, 2015: 234-241.
- [38] JIANG N F, CHEN W L, LIN Y T, et al. Underwater image enhancement with lightweight cascaded network[J].

- IEEE Transactions on Multimedia, 2022, 24: 4301-4313.
- [39] ZHUANG P X, WU J M, PORIKLI F, et al. Underwater image enhancement with hyper-Laplacian reflectance priors[J]. IEEE Transactions on Image Processing, 2022, 31: 5442-5455.
- [40] ZHOU J C, LI B S, ZHANG D H, et al. UGIF-net: An efficient fully guided information flow network for underwater image enhancement[J]. IEEE Transactions on Geoscience and Remote Sensing, 2023, 61: 4206117.
- [41] RHIF M, ABBES A BEN, FARAH I R, et al. Wavelet transform application for/in non-stationary time-series analysis: A review[J]. Applied Sciences, 2019, 9(7): 1345.
- [42] HU J, SHEN L, SUN G. Squeeze-and-excitation networks[C]// 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2018: 7132-7141.
- [43] JADERBERG M, SIMONYAN K, ZISSERMAN A. Spatial transformer networks[C]//Proceedings of the Advances in Neural Information Processing Systems. Quebec: NIPS Foundation, 2015: 8024-8035.
- [44] PASZKE A, GROSS S, MASSA F, et al. Pytorch: An imperative style, high-performance deep learning library[C]// Proceedings of the Advances in Neural Information Processing Systems. British Columbia: NIPS Foundation, 2019: 2017-2025.
- [45] KINGMA D P, BA J, HAMMAD M M. Adam: A method for stochastic optimization[EB/OL]. (2021-12-22) [2024-03-24]. <https://arxiv.org/abs/1412.6980v9>.
- [46] WOO S, PARK J, LEE J Y, et al. CBAM: Convolutional block attention module[M]//Lecture Notes in Computer Science. Cham: Springer International Publishing, 2018: 3-19.
- [47] KIRILLOV A, MINTUN E, RAVI N, et al. Segment anything[C]//2023 IEEE/CVF International Conference on Computer Vision (ICCV). Piscataway: IEEE, 2023: 3992-4003.

作者简介



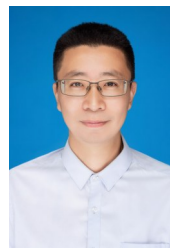
牛玉贞 女,1982年出生于山东省济南市,博士。现为福州大学计算机与大数据学院教授、博士生导师。主要研究方向为计算机视觉。
E-mail: yuzhenniu@gmail.com



许瑞 男,1995年出生于江苏省扬州市。现为福州大学计算机与大数据学院博士研究生。主要研究方向为图像增强与复原、视觉目标跟踪。
E-mail: xurui.ryan.chn@gmail.com



张凌昕 女,1999年出生于福建省福州市。现为福州大学计算机与大数据学院硕士研究生。主要研究方向为计算机视觉。
E-mail: fzuzhanglingxin@163.com



柯道 男,1982年出生福建省福州市,博士。现为福州大学教授、博士生导师。主要研究方向为计算机视觉、模式识别。
E-mail: kex@fzu.edu.cn



兰杰 男,2000年出生于福建省武平县。现为福州大学计算机与大数据学院硕士研究生。主要研究方向为计算机视觉。
E-mail: jielan.blue@gmail.com