

文章编号: 1674-5566(2026)01-0254-16

DOI: 10.12024/jso.20250304795

DUIE-YOLO:一种基于图像增强的水下鱿鱼目标检测算法

曹莉凌, 胡浩宇, 曹守启

(上海海洋大学 工程学院, 上海 201306)

摘要:为了解决水下图像因模糊和色偏导致的目标检测精度下降问题,提升复杂水下环境中鱿鱼检测的准确性和鲁棒性,本研究提出一种基于图像增强的水下鱿鱼检测算法DUIE-YOLO,采用“先增强后检测”的级联框架,由DUIE-Net增强模块和YOLOv8-HD检测模块组成。DUIE-Net模块通过颜色校正、多尺度特征融合、特征恢复与增强及去雾优化,显著提升图像质量;YOLOv8-HD检测模块结合FasterNet网络、小目标检测头、CoordAttention注意力机制及ShapeIoU损失函数,优化特征提取能力与小目标检测精度。实验结果表明,DUIE-YOLO相比原始YOLOv8n在Precision、Recall、F1-score和mAP等4个关键指标上分别提升4.2%、6.8%、5.7%和5.5%。联合实验结果显示,DUIE-Net与YOLOv8-HD的组合相比基线(Raw+YOLOv8n),mAP提升40.3%,Precision提升10.5%,Recall提升53%,F1-score提升31%,证明该算法具有显著的级联优化效果。研究表明,DUIE-YOLO通过图像增强与检测模块的协同优化,有效解决了水下图像质量差导致的检测性能下降问题。本研究为复杂水下环境中的目标识别提供了高精度的解决方案,对海洋生物监测与资源开发具有重要应用价值。

关键词:水下鱿鱼检测; 目标检测; 图像增强; 多尺度特征融合; YOLOv8

中图分类号: S 951.2; TP 183 文献标志码: A

随着全球渔业资源枯竭和海洋环境恶化,水下目标检测技术在海洋科学、渔业管理和环保等领域日益重要。鱿鱼作为高价值水产资源,其捕捞和监测不仅关系到渔业经济的可持续发展,还对生态保护起着重要作用,但水中悬浮颗粒的散射、水体的吸收和衰减等因素会导致水下图像出现模糊、色偏、对比度低等质量退化问题,严重影响图像的后续利用^[1]。这些因素严重影响了水下目标检测的精度,尤其是在检测和识别鱿鱼等目标时,现有的目标检测算法面临着巨大挑战。

为应对水下图像模糊、色偏等问题,图像增强算法应运而生,该算法可以有效提高水下图像的视觉质量。水下图像增强的发展历程可以追溯到数十年前,早期的直方图均衡化^[2]等传统方法主要通过调整图像的灰度分布来改善对比度,但未能针对水下环境特有的光散射和色彩偏移问题进行优化;变换域方法^[3]将图像转换到频域

进行处理,但由于需要手动调整参数,难以适应复杂多变的水下光照条件;而基于Retinex方法^[4]虽然能够通过分离光照分量来增强图像,但其简化的光照模型在面对水下环境中的散射光和光谱失真时表现欠佳,不仅难以准确还原真实色彩,还容易引入噪声干扰。

近年来,随着深度学习技术的飞速发展,水下图像增强迎来了革命性的变化。针对水下图像模糊和色偏问题,DehazeNet^[5]、GridDehazeNet^[6]、4kDehazing^[7]等模型通过学习水下图像的物理特性,实现去雾和清晰化。对于水下图像普遍分辨率低的问题,SDFlow^[8]模型能够在缺乏配对数据的情况下学习低分辨率和高分辨率图像之间的双向映射关系,LapSRN^[9]模型则是基于拉普拉斯金字塔的超分辨率网络,通过多尺度特征提取逐步重建高分辨率图像,使模型在速度和精度之间取得了良好平衡,这些技术的

收稿日期: 2025-03-06 修回日期: 2025-05-14

基金项目: 国家重点研发计划(2023YFD2401302)

作者简介: 曹莉凌(1982—),女,高级工程师,博士,研究方向为渔业工程装备。E-mail: llcao@shou.edu.cn

通信作者: 曹守启, E-mail: sqcao@shou.edu.cn

版权所有 ©《上海海洋大学学报》编辑部(CC BY-NC-ND 4.0)

Copyright © Editorial Office of Journal of Shanghai Ocean University (CC BY-NC-ND 4.0)

<http://www.shhydxxb.com>

结合为水下图像增强提供了全面的解决方案,推动了水下目标检测、环境监测等应用的发展。

在基于深度学习的目标检测中,深度神经网络以其强大的特征提取能力和端到端的学习方式,显著提升了检测精度和效率,其通过构建深层卷积网络,能够从复杂的图像数据中提取多层次特征,并结合注意力机制、多尺度特征融合等技术,实现对目标的精确定位和分类。R-CNN^[10]首先提出了将CNN应用于目标检测的思路,通过选择性搜索生成候选区域,再对每个区域进行分类和边界框回归。

R-CNN属于两阶段检测算法,识别精度高但速度较慢,基于此诞生出了单阶段检测算法,单阶段检测算法通过损失一定的识别精度,提高检测速度,代表算法有YOLO^[11]和SSD^[12]。陈欣等^[13]提出增强型多尺度特征融合SSD模型,采用ResNet101作为主干网络,显著提升了特征提取能力。WANG等^[14]提出轻量化L-SSD系统,结合深度可分离卷积、MobileNetv2及优化结构(如UFFM、BiFPN),在保持精度的同时降低计算量。赵晨等^[15]提出YOLOv7-RFCW,通过引入P-ELAN、CBAM注意力机制和WIOU-V3损失函数,优化了复杂环境下的检测性能,但计算量仍较大,需进一步轻量化以适应边缘设备部署。



图1 DUIE-YOLO算法框架结构
Fig. 1 Framework structure of DUIE-YOLO algorithm

1.2 数据来源

本研究采用多源数据集,分别针对图像增强模块(DUIE-Net)和目标检测模块(YOLOv8-HD)进行训练与评估,以确保模型的泛化能力和鲁棒性。对于DUIE-Net模型,本研究在公开数据集上进行实验,包括UIEB^[16]和C60数据集。其中,UIEB数据集中包含890张模糊水下图像及其对应的清晰对照图像。本实验随机选取其中的800对图像作为训练集,以学习从模糊图像到清晰图像的映射关系,其余90对图像作为测试集,用于有参考图像的增强效果评估。C60数据集则包含60张不同环境下的水下模糊图像,无清

1 材料与方法

1.1 DUIE-YOLO算法框架概述

传统的目标检测算法往往无法有效应对水下图像的退化特性,导致识别精度较低。为了克服上述问题,本研究提出了DUIE-YOLO算法框架,这是一种级联式深度学习架构,专门用于水下鱿鱼目标检测任务,该算法通过图像增强与目标检测的级联优化,实现了从退化图像输入到精准目标检测的端到端解决方案。DUIE-YOLO算法框架包含DUIE-Net和YOLOv8-HD两个核心子模块,DUIE-Net作为前置图像增强模块,专门处理水下图像特有的颜色失真、模糊和低对比度等问题;YOLOv8-HD作为后端检测模块,在增强后的图像上执行高精度目标检测,形成完整的处理流水线。

本算法采用“先增强后检测”的基础架构设计,算法框架结构如图1所示。原始水下图像首先经过DUIE-Net的多阶段增强处理,包括颜色校正、多尺度特征融合和去雾优化等操作,这些处理显著改善了图像的视觉质量和特征可辨识度。增强后的图像输入到YOLOv8-HD模块,该模块通过改进的骨干网络和注意力机制,实现对鱿鱼目标的精确定位。

晰对照图像,本研究将其用作无参考图像的增强实验数据集,以评估模型的通用性和适应性。

对于YOLOv8-HD模型,本研究采用Lignjev.2 Computer Vision Project中的鱿鱼数据集^[17],从该数据集中挑选1 000张高质量鱿鱼图像,所有图像均经过严格的筛选与专业标注。为确保模型的泛化能力与评估效果,数据集被划分为训练集、验证集和测试集,其中900张图像用于训练,80张用于验证,剩余20张用于测试。

2 图像增强模块DUIE-Net

图像增强模块DUIE-Net专为提升水下图像

质量而设计。DUIE-Net采用端到端的深度学习架构,主要由生成器网络、VGG19特征提取器、Adam优化器以及多任务损失函数4个关键组件构成,本文的研究重点聚焦于生成器网络的结构改进与多任务损失函数的优化设计。

DUIE-Net生成器网络模型主要包括颜色校正模块、金字塔多尺度融合模块、特征恢复与增强、去雾优化等4个部分(图2)。其中,颜色正

模块负责纠正由于水下光传播导致的颜色失真问题;金字塔多尺度融合模块采用分层结构设计,提取多层次的图像特征信息,增强整体视觉效果;特征恢复与增强进一步优化图像特征,恢复图像的清晰度;去雾优化模块去除水下图像中的模糊和低对比度问题,提升图像质量,生成清晰、自然且视觉效果显著增强的水下图像。

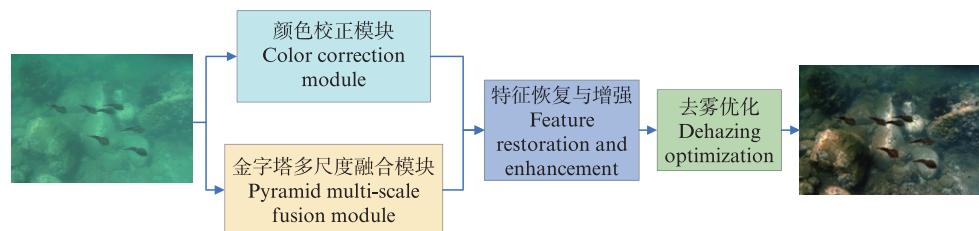


图2 DUIE-Net生成器网络模型
Fig. 2 DUIE-Net generator network model

2.1 颜色校正模块

颜色校正模块是专门针对水下图像颜色失真问题设计的网络模块,通过分层特征提取、注意力权重优化以及通道特征融合等一系列操作,能够有效地修复和增强水下图像的颜色信息。

颜色校正模块网络结构如图3所示,本模块参考Deep-WaveNet^[18]算法的Stage1~Stage3,采用分通道处理策略对输入RGB图像进行特征提取。由于水下图像中常出现红色信息缺失现象,本模块在红色通道上使用尺寸较小的 3×3 卷积核,以提取更加细致的局部特征;对于绿色和蓝色通道,由于它们在水下环境中保留的特征相对较多,本模块分别采用 5×5 和 7×7 较大的卷积核,提取更广泛的上下文信息,增强图像整体的色彩平衡性和细节还原能力。各通道特征均经过CBAM注意力模块处理,通过通道注意力和空间注意力的双重机制,有效增强关键区域的表征能力,最后通过特征融合将优化后的多尺度特征进行整合,在保持色彩自然平衡的同时显著提升图像细节的还原质量。

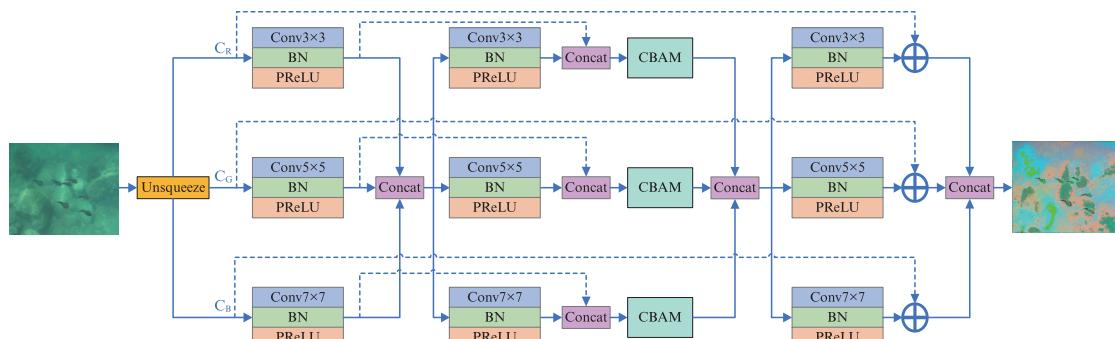


图3 颜色校正模块网络结构
Fig. 3 Color correction module network structure

2.2 金字塔多尺度融合模块

金字塔多尺度融合模块是一种通过结合多尺度金字塔池化和注意力机制来增强特征图表达能力的模块,旨在提升深度学习模型在处理复杂输入时的性能。本模块设计思路是基于图像中多尺度特征的提取和关注重要区域的信

息,提高模型对多样化输入的理解和处理能力,捕捉并融合不同尺度的信息,同时抑制无关或冗余的特征,从而在视觉任务中增强特征的表示能力。

金字塔多尺度融合模块网络结构如图4所示,多尺度融合操作计算过程如下:

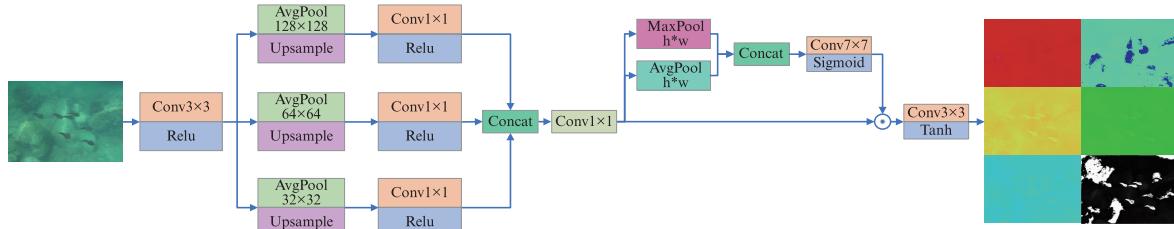


图4 金字塔多尺度融合模块网络结构

Fig. 4 Pyramid multi-scale fusion module network structure

$$X' = \text{ReLU}(\text{Conv3} \times 3(X)) \quad (1)$$

$$P_f(X', s) = \text{ReLU}(\text{Conv1} \times 1 \\ (\text{Upsample}(\text{AvgPool}(X', s)))) \quad (2)$$

$$X_p = \text{Conv1} \times 1 \\ (\text{Concat}(P_f(X', s_{128}), P_f(X', s_{64}), P_f(X', s_{32}))) \quad (3)$$

式中: X 为输入图像;Conv3 \times 3为尺寸为3 \times 3的卷积;ReLU为ReLU激活函数; X' 为预处理后的特征图;AvgPool为平均池化操作;Upsample为上采样操作;Conv1 \times 1为尺寸为1 \times 1的卷积; s 为不同尺度,取128 \times 128、64 \times 64、32 \times 32; $P_f(X', s)$ 为对 X' 进行不同尺度的操作;Concat为拼接操作; X_p 为多尺度融合结果。

首先对输入图像进行标准化特征处理,使用3 \times 3卷积和ReLU激活函数进行预处理,这个过程不仅有助于压缩输入通道,减少计算量,还能有效保持模型的表达能力。在完成输入特征的标准化处理后,特征图被送入多尺度处理模块。金字塔池化模型通过3个不同尺度的池化层实现渐进式的特征提取:128 \times 128的大尺度池化层主要捕获全局性的宏观特征;64 \times 64的中尺度池化层作为过渡层级,负责整合中等范围的特征信息;32 \times 32的小尺度池化层则专注于提取精细的局部特征,这种分级设计使模型能够有效应对具有多尺度特性的复杂输入数据。在特征处理阶段,模型采用双线性插值算法将各尺度的特征图统一上采样至输入特征图的空间分辨率,确保不同尺度特征在空间维度上保持对齐。为进一步增强特征表达能力,金字塔多尺度融合模块为每个尺度的池化操作配置了独立的1 \times 1卷积核和ReLU激活函数,避免跨尺度特征干扰。随后将3个尺度的特征图通过通道拼接(Concat)操作进行融合,并使用1 \times 1卷积对拼接后的特征进行跨通道信息整合和维度调整,这种架构的设计能够使模型实现更全面的特征表示。

为进一步优化特征表达,模块引入全局最大池化和平均池化,生成注意力图以捕捉局部区域的依赖关系,并通过Sigmoid函数将输出值限制在0到1之间,确保注意力图的稳定性。具体来说,模块应用全局最大池化和平均池化提取显著特征,其中最大池化捕捉水下图像中目标边缘(如珊瑚、岩石)的局部极值,避免因散射导致的细节丢失,而平均池化平滑水体背景的均匀区域(如悬浮颗粒)。两者通过Concat拼接后使用7 \times 7卷积和Sigmoid进行融合,融合后的注意力权重图A能自适应地强化高频细节(A \rightarrow 1)并抑制平滑背景(A \rightarrow 0),而7 \times 7卷积的大感受野能覆盖多尺度上下文,适应水下目标的复杂分布。生成的注意力图随后对原始特征图进行逐元素加权,结合3 \times 3卷积和Tanh激活函数,进一步强化高权重区域的关键特征表达,同时抑制低权重区域的冗余信息干扰。通过这种精细加权,模块显著提升模型对重要特征的关注与理解能力,实现更精准的特征提取与决策优化:

$$GM(X_p) = \text{Sigmoid}(\text{Conv7} \times 7 \\ (\text{Concat}(\text{MaxPool}(X_p), \text{AvgPool}(X_p)))) \quad (4)$$

$$X_{att} = \text{Tanh}(\text{Conv3} \times 3(GM(X_p) \odot X_p)) \quad (5)$$

式中:MaxPool为最大池化操作;Conv7 \times 7为尺寸为7 \times 7的卷积;Sigmoid为Sigmoid激活函数;GM(X_p)为对 X_p 进行操作;Tanh为Tanh激活函数; X_{att} 为生成的注意力图。

为验证注意力图通过加权特征增强局部区域的空间依赖(如像素间强度相关性)和语义依赖(如前景与背景的关联),开展消融实验,比较含全局池化的PAM模型与无全局池化的PAM_N模型的性能。实验结果如表1所示,PAM在图像增强任务中显著优于PAM_N:均方误差(MSE)从0.0060降至0.0047,峰值信噪比(PSNR)从23.775 dB升至24.129 dB,结构相似性(SSIM)从0.879升至0.895。这些提升表明全局池化通过捕

提局部区域的空间依赖和语义依赖,改善了图像的重建精度和结构保留能力。其中,SSIM的提升反映了模块对图像结构信息(如边缘和纹理)的更好建模,而PSNR的提高表明其增强了细节

表1 全局池化对模型性能的影响
Tab. 1 Impact of global pooling on model performance

| 模型 Models | 均方误差 MSE | 峰值信噪比 PSNR/dB | 水下图像质量评价 UCIQE | 结构相似性 SSIM |
|------------------|-------------|------------------|-------------------|---------------|
| Basic+Loss+PAM_N | 0.006 0 | 23.775 | 5.590 | 0.879 |
| Basic+Loss+PAM | 0.004 7 | 24.129 | 5.489 | 0.895 |

2.3 特征恢复与增强模块

特征恢复与增强模块是专门针对图像处理过程中细节恢复和特征增强问题设计的一种网络模块,该模块结合了反卷积层、特征融合技术以及注意力机制,旨在恢复图像的真实细节,提升图像质量,并加强图像中关键特征的表达能力。

特征恢复与增强网络结构如图5所示,本模块参考Deep-WaveNet^[18]算法的Stage4,通过拼接颜色校正模块和金字塔多尺度融合模块的特征

保留和噪声抑制能力,但PAM的水下图像质量评价(UCIQE)从5.590降至5.489,可能因全局池化过于聚焦显著区域,影响了全局色度平衡。

图,整合多尺度的上下文信息,而反卷积层的上采样操作则恢复了图像的空间结构,增加了分辨率和细节层次。结合CBAM注意力机制,模块能够为不同通道分配权重,优化色彩和纹理,最终生成的图像在细节、色彩和纹理等方面都得到了显著提升。这一设计不仅有效解决了水下图像处理中的颜色失真、细节丢失和分辨率低等问题,还增强了图像的视觉效果和信息提取能力,为水下目标检测和环境监测等应用提供了强有力的支持。

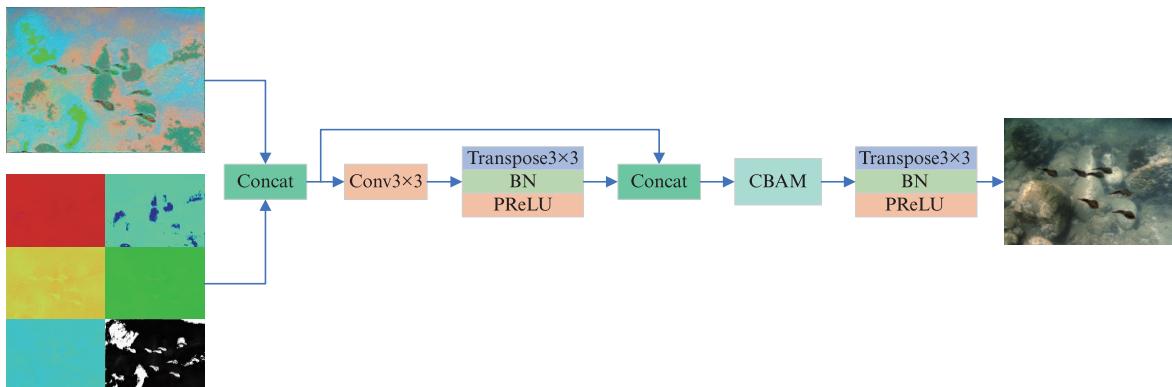


图5 特征恢复与增强网络结构
Fig. 5 Feature restoration and enhancement network structure

2.4 去雾优化

在本研究中,去雾优化模块基于UDCP(Underwater dark channel prior)^[19]算法来实现水下图像的去雾与优化。UDCP算法通过对蓝绿色通道的暗通道先验计算,准确估计透射率和背景光,结合多尺度优化与后处理手段,实现了水下图像的高效去雾和色彩修复,进而恢复图像的清晰度和真实色彩,模块结构如图6所示。

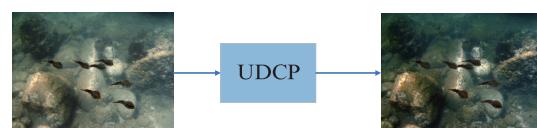


图6 去雾优化模块
Fig. 6 Dehazing optimization module

2.5 损失函数

本研究在图像增强模型的训练中,综合使用了均方误差(MSE)损失、VGG感知损失、结构相

似性(SSIM)损失和颜色损失,多角度联合优化提升生成图像质量。其中,MSE保证像素级重建精度,VGG感知损失维持高级语义特征一致性,SSIM损失保护图像结构完整性,颜色损失校正色彩失真。通过多种损失函数的协同作用,网络能够在像素级精确重建的基础上,进一步优化图像的语义信息、结构相似性、色彩一致性和局部平滑性,不仅提高了图像的视觉效果,还增强了网络对复杂水下场景的适应性和泛化能力。

3 目标检测模型 YOLOv8-HD

目标检测模块 YOLOv8-HD 在 YOLOv8 框架基础上进行了多项关键性优化,具体网络架构

如图 7 所示。对于 Backbone 网络,采用了 FasterNet 作为 YOLOv8 中使用的基础网络,FasterNet 以其轻量化设计和较高的精度优势,能够在保证计算效率的同时,提取深层次的图像特征,使得更加准确地识别目标;为了进一步提升模型精度,引入了 CoordAttention 注意力机制,该机制结合了自适应池化和通道注意力,通过在空间和通道维度上自适应地调整特征的权重,能够有效提高目标的定位精度;对于损失函数,采用 ShapeIoU 损失函数,与传统的 IoU 度量不同,ShapeIoU 不仅关注矩形框的重叠程度,还考虑了框的形状(宽高比)和位置,能够提供更准确的框间相似性度量。

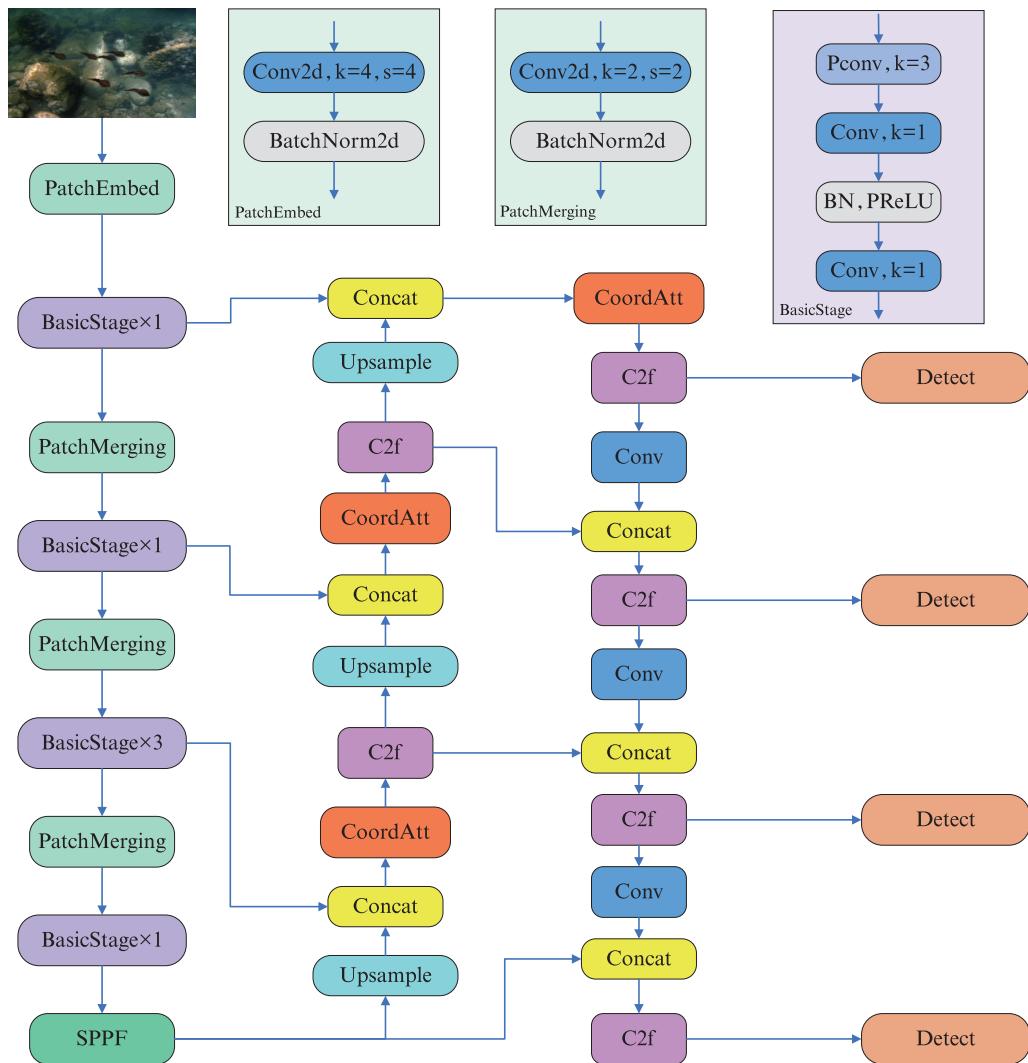


图 7 YOLOv8-HD 网络结构
Fig. 7 YOLOv8-HD network structure

3.1 YOLOv8 网络结构改进

在目标检测任务中,特征提取网络通过提取

高质量、多尺度的特征,并结合上下文信息,能够有效区分目标与背景,提升检测性能。所以,如

何选择和优化特征提取网络是提升目标检测效果的关键。

ShuffleNet^[20]是一种轻量级网络,专为移动设备等计算资源受限的场景设计,它通过组卷积和通道混洗技术显著减少了计算量和参数量,但当组数较多时,该网络的组间信息交互受限,组卷积可能导致特征表达能力的下降。EfficientNet^[21]创新性地采用了复合缩放机制,运用神经架构搜索(NAS)技术进行自动化设计,但某些深度和宽度扩展导致特征图内存占用增加,且复合缩放需要大量超参数调试和神经架构搜索。SqueezeNet^[22]提出了一种基于Fire模块的结构,通过将传统的 3×3 卷积操作分解为 1×1 和 3×3 卷积的组合,显著降低了模型的参数规模和计算复杂度,但是网络结构单一,无法很好地适配复杂任务或大规模数据集。

针对上述网络的缺点,本研究选用FasterNet作为YOLOv8的骨干网络,以提取输入图像的特征。FasterNet是一种轻量化且高效的卷积神经网络(CNN)架构,其设计核心在于将卷积神经网络和多层次感知机(MLP)相结合,充分发挥二者的优势,并引入部分卷积技术,解决了ShuffleNet和SqueezeNet在特征表达能力上的不足。而部分卷积^[23]通过将卷积处理后的通道与未处理通道进行混合,使得网络在保留输入特征的同时能够高效提取局部空间信息,不仅提高了特征表达能力,还避免了组卷积在信息交互上的局限性。

3.2 CoordAttention 注意力机制

在本研究中,引用CoordAttention^[24]模块来增强特征提取能力,与传统的注意力机制如CBAM^[25]和SE^[26]不同,CoordAttention是一种基于空间和通道信息的注意力机制,旨在通过对输入特征图的空间位置进行编码,增强网络对不同空间位置特征的关注能力,从而提高特征表达的准确性和多样性。CoordAttention模块的创新之处在于它将空间维度的特征分开处理,并通过自适应池化、卷积和激活操作有效捕获不同方向(高度和宽度)上的重要特征信息,从而能够在处理复杂的目标检测任务时,识别出更加细致的特征,并避免了其他注意力机制中可能存在的忽视位置关系的缺陷。

3.3 边界框回归损失优化

本研究引入ShapeIoU边界框损失函数,替代

传统IoU损失。传统IoU损失仅基于预测框与真实框的交并比来计算优化目标,当预测框与真实框没有重叠时,IoU损失的梯度为零,难以引导模型对框位置的有效回归。虽然IoU的改进方法GIoU^[27]和DIOU^[28]试图解决这一问题,但它们在实际应用中仍然存在局限性。为弥补不足,ShapeIoU损失函数在IoU的基础上进一步引入了对边界框几何形状的敏感性,ShapeIoU边界框损失函数的独特之处在于结合了框的几何形状特性,不仅关注框的重叠区域,还进一步评估了框的长宽比、形状比例等关键特征。与GIoU和DIOU相比,ShapeIoU在低重叠或无重叠情况下表现出更高的稳定性,并且能够对目标框的形状特性进行精细优化,ShapeIoU不仅提升了边界框位置的收敛效率,还提高了预测框与真实框在形状和位置上的匹配精度。

4 结果与分析

4.1 实验条件与评价指标

模型训练实验在GPU服务器上完成,服务器配置:CPU为Intel Xeon Silver 4214R,显卡型号为RTX 3080 Ti。软件环境:Python版本为3.8,torch版本为1.8.1,CUDA版本为11.1。DUIE-Net模型使用Adam优化器,Batch Size为6,Epochs为400;YOLOv8-HD使用Adam优化器,Batch Size为8,Epochs为300。

在图像增强模型的评估中,本研究采用多种指标衡量增强效果,包括图像熵(Entropy)、均方误差(MSE)、峰值信噪比(PSNR)、水下图像质量指数(UCIQE)和结构相似性指数(SSIM),这些指标能够有效评估图像的细节保留、失真程度、视觉质量以及与参考图像的相似性,提供了多维度的评价体系。

在目标检测模型的评估中,本研究采用精确率(Precision)、召回率(Recall)、F1 Score、平均精度均值(Mean average precision, mAP)、模型参数量(Model parameters)和计算量(FLOPs)衡量检测性能和模型效率,这些指标能够有效评估模型的检测准确性、目标定位能力、综合性能以及计算效率,为目标检测提供了多维度的评价体系。

4.2 对比实验

4.2.1 DUIE-Net模型实验结果及分析

为了验证本模型的泛化能力,将DUIE-Net

模型与UDCP、Fusion^[29]、FUnIE-GAN^[30]、LiteEnhanceNet^[31]、PUGAN^[32]、Shallow-water^[33]、TCTL-Net^[34]进行对比试验。

表2数据表明,DUIE-Net在UCIQE和SSIM性能指标上表现优异,显示出其在图像质量和结构保真度上的显著优势,证明其在水下图像增强和目标检测等任务中具有较强的竞争力。在重建精度方面,DUIE-Net能够有效地提升图像的

视觉质量,保证细节的丰富性和清晰度,其均方误差(MSE)和峰值信噪比(PSNR)值表现出色,表明该模型生成的图像与原图差异较小,且图像质量非常高。模型的UCIQE和SSIM值也反映出其在颜色增强、去雾和结构保真度方面的优越性,能够在改善图像质量的同时,保持图像的真实感和自然感。

表2 各模型参数对比
Tab. 2 Comparison of parameters for each model

| 模型 Models | 均方误差 MSE | 峰值信噪比 PSNR/dB | 水下图像质量评价 UCIQE | 结构相似性 SSIM |
|----------------|-------------|------------------|-------------------|---------------|
| UDCP | 0.027 5 | 17.300 | 5.230 | 0.668 |
| Fusion | 0.006 4 | 22.789 | 5.222 | 0.872 |
| FUnIE-GAN | 0.020 3 | 17.740 | 4.691 | 0.549 |
| LiteEnhanceNet | 0.007 1 | 22.996 | 5.027 | 0.834 |
| PUGAN | 0.017 7 | 18.165 | 5.941 | 0.562 |
| Shallow-water | 0.030 5 | 18.866 | 5.394 | 0.658 |
| TCTL-Net | 0.005 1 | 24.242 | 5.154 | 0.847 |
| DUIE-Net | 0.006 3 | 22.276 | 5.941 | 0.888 |

由图8可以看出,不同模型在实际图像增强中的表现存在显著差异,UDCP、FUnIE-GAN、PUGAN在图像增强方面效果不尽如人意,整体颜色偏灰暗,缺乏足够的图像细节;TCTL-Net和LiteEnhanceNet的增强效果稍好一些,能够在一定程度上改善图像质量,但这些模型在去雾方面的效果并不理想,增强后的图像中仍然保留了较多的雾霾效果,导致图像中的内容特征不够明显,影响了图像的可读性和可用性;Fusion和Shallow-water虽然能够在部分图像中取得较好的效果,成功地还原了较为清晰的水下图像,但在某些情况下,图像的红色通道增强过于强烈,导致部分图像出现颜色失真,失去了原有的自然色调,影响图像的真实感和自然感;相比之下,DUIE-Net在各项指标上均表现适中,能够为平衡地增强图像的亮度和对比度,避免过度强化某一通道的色彩,从而还原出更自然、真实的水下图像。模型在去雾和细节保留方面的表现也较为出色,能够有效地提取出图像中的重要特征,增强图像的清晰度和视觉效果,减少了

后续处理的难度,提高了目标识别的效率和准确性。

4.2.2 YOLOv8-HD模型实验结果及分析

为了验证本模型的泛化能力,将YOLOv8-HD模型与Faster R-CNN^[35]、SSD、YOLOv5、YOLOv10、YOLOv11进行对比试验。

图9展示了Faster R-CNN、SSD、YOLOv10s、YOLOv11s、YOLOv8-HD在不同图像上的预测结果,其中仅显示置信度高于0.5的预测框。从图中可以看出,YOLOv8-HD模型在单目标检测和低光照环境等场景中表现优异,检测结果稳定可靠,但在多目标和目标重叠场景中,该模型仍存在一定的局限性:当目标密集分布或相互遮挡时,会出现个别目标漏检现象,检测性能明显下降。分析认为,这一现象可能与模型对复杂场景的特征表征能力不足有关,后续研究可从以下两方面进行改进:(1)引入高效的注意力机制,增强模型对遮挡目标的判别能力;(2)通过数据增强策略丰富训练样本的多样性。通过这些改进措施,有望提升模型在多目标场景中的检测性能。

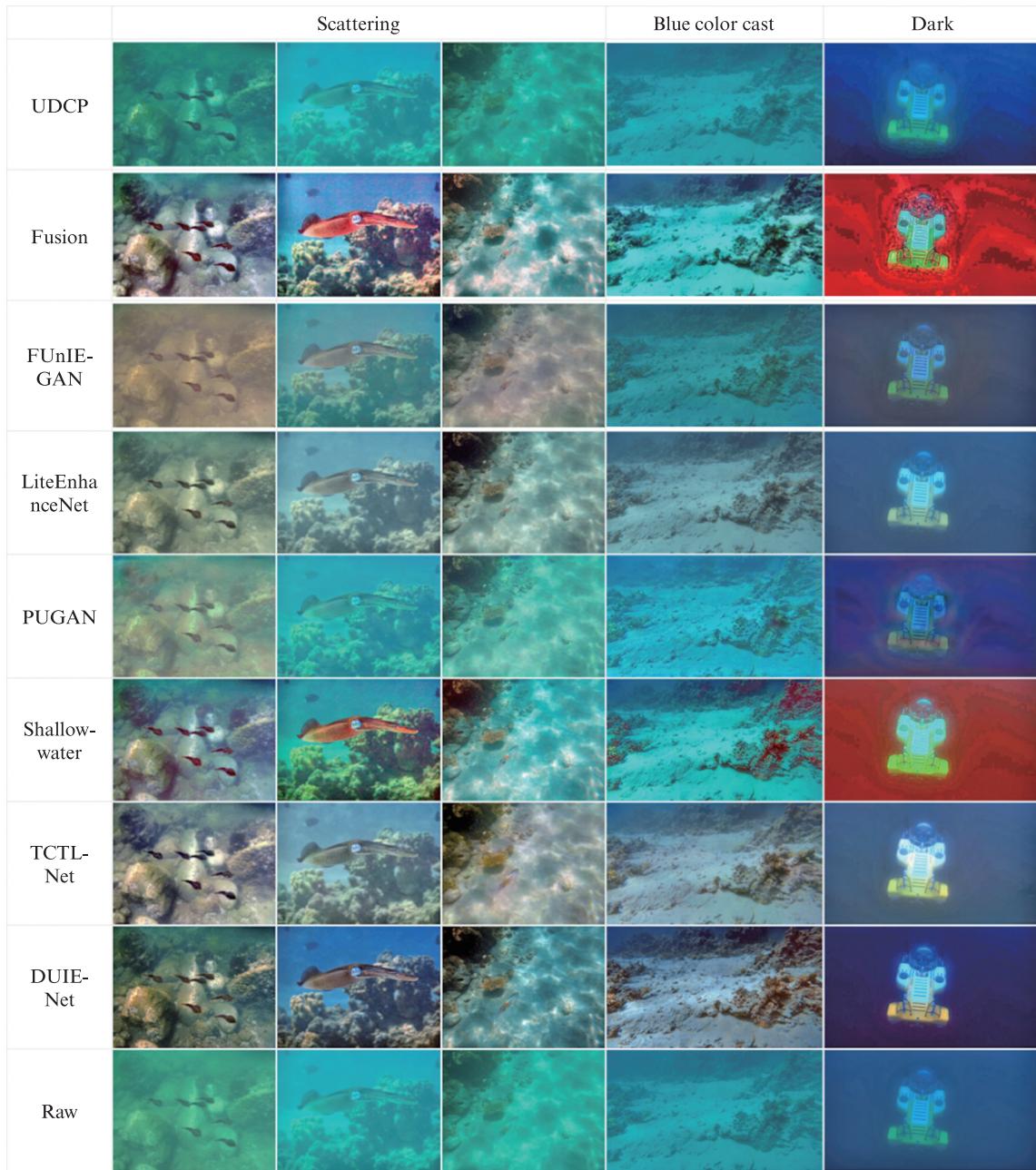


图8 各模型增强效果对比图
Fig. 8 Enhanced effect comparison of different models

表3数据表明, YOLOv8-HD的Recall值为0.838,较其他模型表现更加均衡,表明该模型在检测到目标时更具鲁棒性,尤其在处理目标多样性较大或小物体检测时,具有显著优势。F1值为0.894,处于所有模型的较高水平,表明YOLOv8-HD在Precision与Recall之间找到了较好的平衡,能够在保证准确度的同时,有效减少漏检和误检的情况。尽管在Precision指标上, YOLOv8-HD略逊于YOLOv5m,但其在Recall和mAP上的优势弥补了这一差距,而且

YOLOv8-HD相对适中的较低计算复杂度和模型参数数量也使其在保持高性能的同时,更具实用性和部署优势。

4.3 消融实验

4.3.1 DUIE-Net模型消融实验

在DUIE-Net模型消融实验中,选用颜色校正模块和特征恢复与增强模块结合作为基准模型(Basic),逐步增加损失函数(Loss)、无全局池化金字塔模块(PAM_N)、含全局池化金字塔模块(PAM)、去雾优化模块(UDCP)。

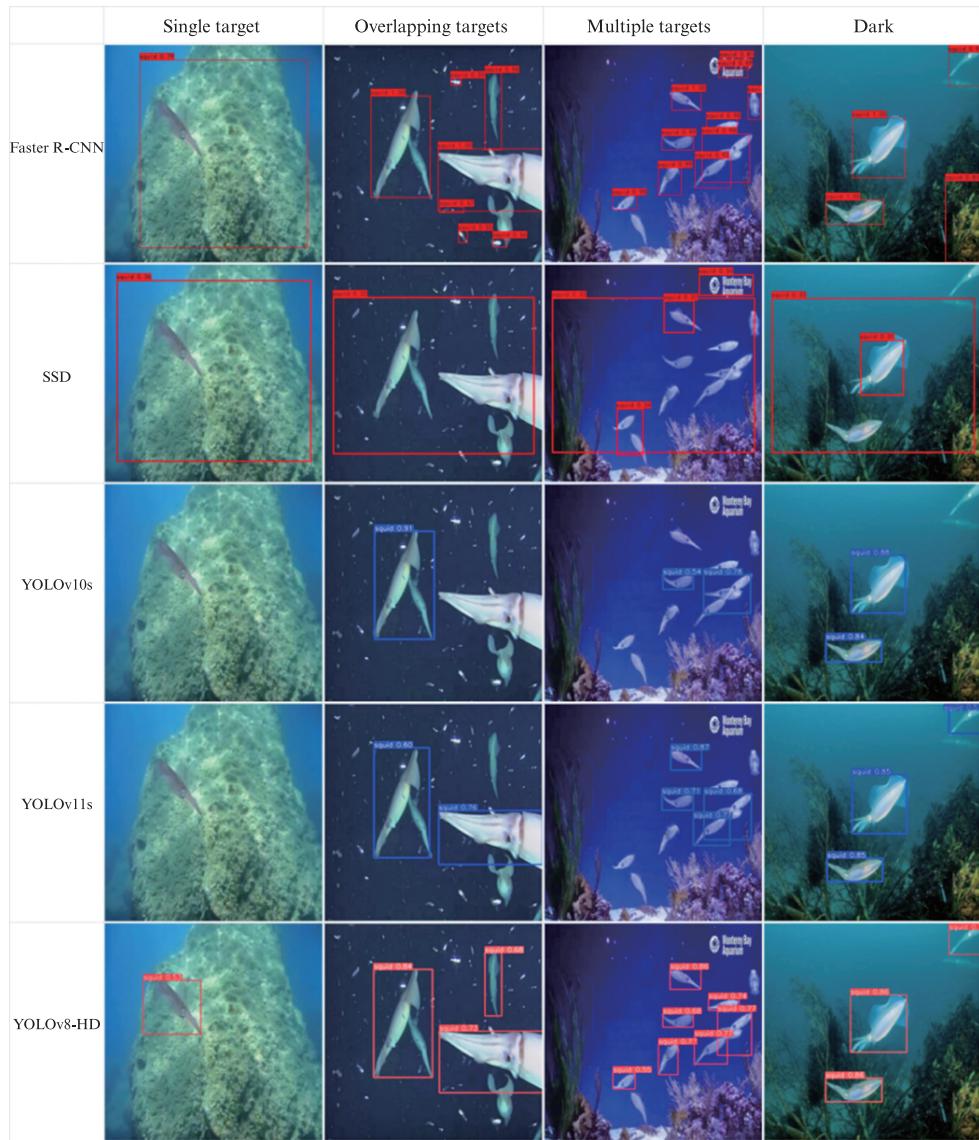


图9 各模型预测结果
Fig. 9 Prediction results of each model

表3 各模型参数对比
Tab. 3 Comparison of parameters for each model

| 模型 Models | 精确率 Precision | 召回率 Recall | F1值 F1-Score | 平均精度 mAP/% | 参数量 Parameters | 浮点运算量 FLOPs/G |
|--------------|---------------|------------|--------------|------------|----------------|---------------|
| Faster R-CNN | 0.434 | 0.867 | 0.580 | 0.827 | 41 755 286 | 355.16 |
| SSD | 0.820 | 0.300 | 0.430 | 0.490 | 14 812 195 | 24.60 |
| YOLOv5n | 0.900 | 0.832 | 0.865 | 0.913 | 2 503 139 | 7.10 |
| YOLOv5s | 0.863 | 0.761 | 0.809 | 0.876 | 9 111 923 | 23.80 |
| YOLOv5m | 0.965 | 0.779 | 0.862 | 0.886 | 25 045 795 | 64.00 |
| YOLOv10n | 0.898 | 0.698 | 0.785 | 0.829 | 2 694 806 | 8.20 |
| YOLOv10s | 0.944 | 0.751 | 0.837 | 0.876 | 8 035 734 | 24.40 |
| YOLOv10m | 0.788 | 0.758 | 0.773 | 0.820 | 16 451 542 | 63.40 |
| YOLOv11n | 0.938 | 0.810 | 0.869 | 0.904 | 2 582 347 | 6.30 |
| YOLOv11s | 0.940 | 0.828 | 0.880 | 0.900 | 9 413 187 | 21.30 |
| YOLOv11m | 0.912 | 0.805 | 0.855 | 0.902 | 20 030 803 | 67.60 |
| YOLOv8-HD | 0.959 | 0.838 | 0.894 | 0.924 | 11 103 060 | 27.90 |

由图10可以看出, Basic+Loss+PAM+UDCP组合方案能够有效降低水体悬浮颗粒造成的光散射效应, 准确还原水下图像的真实色彩, 并显著提升黑暗环境下的图像可视性。表4数据表明 Basic+Loss+PAM 的组合在公开数据集UIEB上的性能指标表现最佳, 但生成的增强图像在视觉效果上仍存在一定不足, 尤其是在去雾效果和颜

色校正的真实性上有所欠缺。当加入UDCP模块后, 虽然MSE、PSNR、SSIM指标有所下降, 但是该组合生成的增强图像视觉效果更加突出, 尤其在去雾、颜色还原以及细节恢复方面表现更优, 使得增强后的图像特征更加明显, 这对后续的目标检测任务尤其重要。

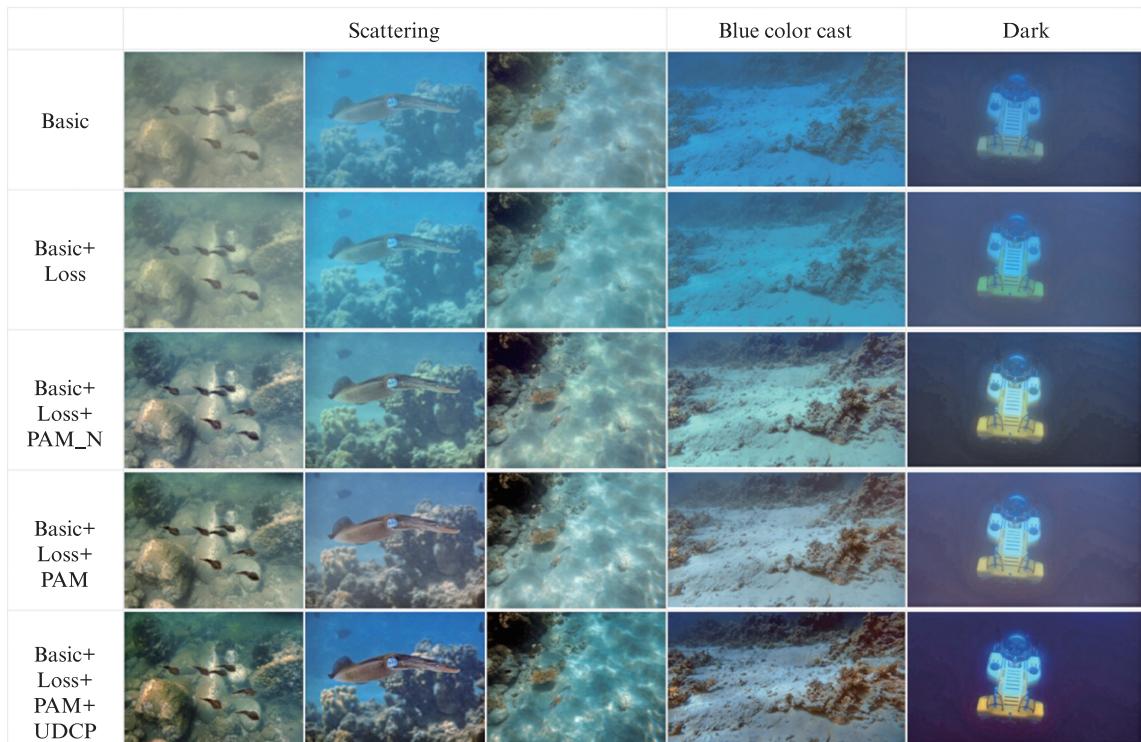


图10 不同模块的增强效果图
Fig. 10 Enhancement effects of different modules

表4 不同模块对模型性能的影响
Tab. 4 Impact of different modules on model performance

| 模型 Models | 均方误差 MSE | 峰值信噪比 PSNR/dB | 水下图像质量评价 UCIQE | 结构相似性 SSIM |
|---------------------|-------------|------------------|-------------------|---------------|
| Basic | 0.017 1 | 19.183 | 5.553 | 0.616 |
| Basic+Loss | 0.017 1 | 19.250 | 5.565 | 0.634 |
| Basic+Loss+PAM_N | 0.006 0 | 23.775 | 5.590 | 0.879 |
| Basic+Loss+PAM | 0.004 7 | 24.129 | 5.489 | 0.895 |
| Basic+Loss+PAM+UDCP | 0.006 3 | 22.276 | 5.941 | 0.888 |

4.3.2 YOLOv8-HD模型消融实验

在YOLOv8-HD模型消融实验中, 使用YOLOv8n作为基准模型, 依次添加FasterNet模块、CoordAttention模块、小目标检测头(Small Target Detection Head)和ShapeIoU损失函数。

表5数据表明, YOLOv8n逐步引入FasterNet(FN)、CoordAttention(CA)、小目标检测头(Small

Target Detection Head, STDH)和ShapeIoU(SIoU)损失函数能够显著提升模型的整体性能, 相比于YOLOv8n, YOLOv8-HD在Precision、Recall、F1和mAP 4个关键指标上分别提升了4.2%、6.8%、5.7%、5.5%, 达到最佳水平, 表明该模型能够在保持较高准确率的同时, 提供优异的整体检测性能。

表5 不同改进对模型性能的影响
Tab. 5 Impact of different improvements on model performance

| 模型 Models | 精确率 Precision | 召回率 Recall | F1值 F1-Score | 平均精度 mAP/% |
|------------------------|------------------|---------------|-----------------|---------------|
| YOLOv8 | 0.917 | 0.770 | 0.837 | 0.869 |
| YOLOv8+FN | 0.960 | 0.814 | 0.881 | 0.920 |
| YOLOv8+FN+CA | 0.974 | 0.788 | 0.871 | 0.892 |
| YOLOv8+FN+STDH | 0.902 | 0.817 | 0.857 | 0.878 |
| YOLOv8+FN+CA+STDH | 0.954 | 0.814 | 0.878 | 0.912 |
| YOLOv8+FN+CA+STDH+SIoU | 0.959 | 0.838 | 0.894 | 0.924 |

4.4 DUIE-YOLO联合验证实验

为了深入探究水下原始图像的质量对目标检测模型性能的影响,本研究设计了DUIE-YOLO联合验证实验,具体实验方案如下:从数据集中筛选出69张具有典型水下成像问题的原始图像,包括颜色失真、模糊、低对比度等常见退化现象,为确保实验的公正性,这些图像均不参与模型的训练过程。将模糊的原始图像输入到预先训练好的DUIE-Net图像增强模型中,通过该模型的优化处理,获得对应的清晰增强图像数据集,并对原始模糊数据集和增强后的清晰数据集分别进行了精细的人工标注,确保标注质量的一致性。完成数据准备后,将这2个数据集输入YOLOv8-HD改进模型和YOLOv8n基准模型进行推理测试,实验过程分为原始数据集和YOLOv8n基准模型(Raw+YOLOv8n)、原始数据集和YOLOv8-HD改进模型(Raw+YOLOv8-HD)、DUIE-Net增强数据集和YOLOv8n基准模型(DUIE-Net+YOLOv8n)、DUIE-Net增强数据集和YOLOv8-HD改进模型(DUIE-Net+YOLOv8-HD)4种组合,以对比分析不同模块组合下的检测性能差异。

图11展示了不同组合在四种典型场景下的预测对比结果,包括黑暗环境(Dark)、模糊目标(Blurred target)、绿色色偏(Green color cast)和蓝色偏(Cyan color cast)。从图中可以看出,在低光、运动模糊及色偏干扰等复杂条件下,DUIE-

Net能够有效增强目标特征,使鱿鱼的形态和纹理更加清晰,YOLOv8-HD在此基础上进一步提升了鱿鱼目标的检测精度。表6数据进一步表明,DUIE-Net模块与YOLOv8-HD模块的组合目标检测性能最佳:Raw+YOLOv8n为基准组合,Raw+YOLOv8-HD和DUIE-Net+YOLOv8n则是在基准组合的基础上更换了不同的改进模块,各方面数据都有所上升,其中Raw+YOLOv8-HD的组合取得了0.783的较高精确率,但召回率仅为0.575,说明该组合模型对模糊图像的预测较为保守,虽然误检较少,但漏检严重;DUIE-Net+YOLOv8-HD组合在整体性能上表现最优,相比于基准组合,其Precision提升至0.724,Recall提升至0.782,mAP从0.586提高到0.822,F1分数也从0.574提升至0.752,但DUIE-Net+YOLOv8-HD组合的精确率比Raw+YOLOv8-HD组合低,可能是因为增强后的图像虽然提高了目标可见性,但也可能使背景特征更接近目标(如增强后的海藻、珊瑚等被误判为鱿鱼),导致误检增加。

总体而言,DUIE-Net与YOLOv8-HD的结合充分发挥了图像增强和目标检测模块的级联优化作用,在复杂水下环境中实现了更稳定、更全面的检测性能,这一结果不仅验证了多模块联合优化的有效性,也为未来水下目标检测算法的改进提供了重要参考。

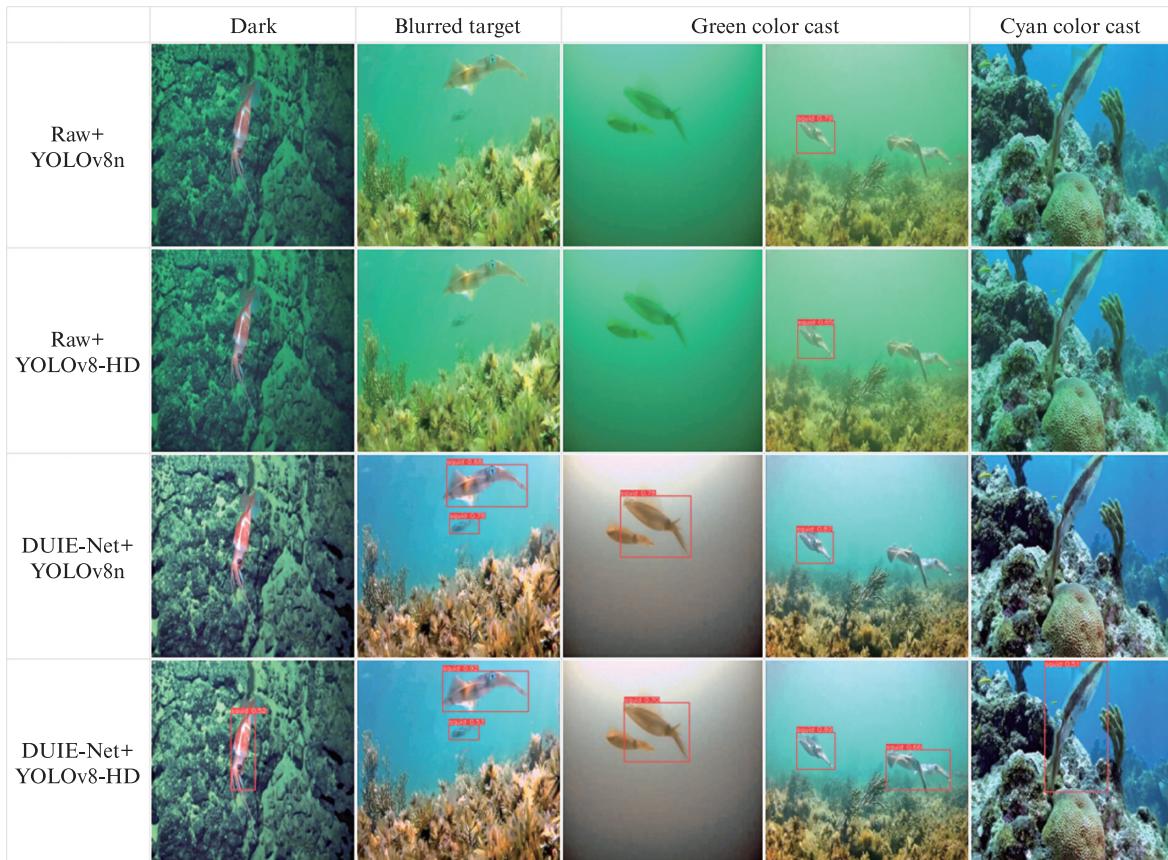


图 11 不同模块组合下的检测性能对比图

Fig. 11 Detection performance comparison with different module combinations

表 6 不同模块组合下的检测性能对比

Tab. 6 Detection performance comparison with different module combinations

| 模块组合 Module combination | 精确率 Precision | 召回率 Recall | F1 值 F1-Score | 平均精度 mAP/% |
|----------------------------|------------------|---------------|------------------|---------------|
| Raw + YOLOv8n | 0.655 | 0.511 | 0.574 | 0.586 |
| Raw + YOLOv8-HD | 0.783 | 0.575 | 0.663 | 0.700 |
| DUIE-Net + YOLOv8n | 0.722 | 0.648 | 0.683 | 0.705 |
| DUIE-Net + YOLOv8-HD | 0.724 | 0.782 | 0.752 | 0.822 |

5 结论

本研究提出了一种基于图像增强的水下鱿鱼目标检测算法 DUIE-YOLO, 旨在解决水下环境中常见的图像质量问题并提高目标检测性能。在图像增强方面, DUIE-Net 模型相比于 TCTL-Net 模型在 UCIQE 和 SSIM 方面分别高出 0.787 和 0.041, 但在 MSE 和 PSNR 方面稍差, 这是由于 DUIE-Net 引入了 UDCP 物理去雾方法, 使得生成的图像与参考图像有一定的差异, 在消融实验中去除 UDCP 模块得出的 MSE 和 PSNR 值与 TCTL-Net 模型相近, 但在主观方面, 增加 UDCP

模块能够进一步提升图像的清晰度, 故本文保留了 UDCP 模块。在目标检测方面, 改进后的 YOLOv8-HD 相比于 YOLOv5m 模型在 Recall、F1、mAP 方面高出 5.9%、3.2%、3.8%, 在 Precision 方面 YOLOv5m 模型比 YOLOv8-HD 高 0.6%, 但 YOLOv8-HD 模型的 Parameters 和 FLOPs 是 YOLOv5m 模型的一半, 降低了模型的复杂度。相比于 YOLOv11n, YOLOv8-HD 模型在精度方面确实有所提高, 但是 Parameters 和 FLOPs 达不到轻量化的要求, 需要进一步优化。

在 DUIE-YOLO 联合验证实验中, 实验结果表明 DUIE-Net 与 YOLOv8-HD 的组合方案展现

出最优异的性能表现,其mAP达0.822,F1分数据升至0.752,且召回率(0.782)显著高于单模块改进方案。这一系列数据充分证明,通过图像增强模块与目标检测模型的级联优化策略能够有效提升模型在复杂水下环境中的检测鲁棒性,为解决水下目标检测难题提供了可靠的技术方案。

尽管DUIE-Net和YOLOv8-HD模型在大多数水下环境中表现出色,能够保持较高的检测精度和稳定性,但在极端水下环境下,如极低光照、高浑浊度或强背景干扰等极端条件下,模型可能面临目标漏检或误检的风险。为了应对这一挑战,未来的研究将进一步优化DUIE-Net和YOLOv8-HD模型,通过引入新的图像增强技术、改进目标检测模块以及针对极端水下环境的训练策略,开发更具鲁棒性的算法。此外,本研究还将探索将该算法部署在硬件平台上的可行性,特别是针对智能渔业的应用需求,以便为渔业生产提供实时、高效的智能检测支持,推动智能化渔业的发展。

作者声明本文无利益冲突。

参考文献:

- [1] 霍冠英,盛兴琳,施淑娴. 基于全局背景光强度估计和对比度调整的无监督水下图像增强[J]. 东南大学学报(自然科学版),2025, 55(1): 297-305.
- HUO G Y, SHENG X L, SHI S X. Unsupervised underwater image enhancement with global background light intensity estimation and contrast adjustment[J]. Journal of Southeast University (Natural Science Edition), 2025, 55(1): 297-305.
- [2] ZUIDERVELD K. Contrast limited adaptive histogram equalization [M]//HECKBERT P S. Graphics GemsIV. Pittsburgh: Academic Press Professional, Inc., 1994: 474-485.
- [3] KAUR G, KAUR M. A study of transform domain based image enhancement techniques [J]. International Journal of Computer Applications, 2016, 152(9): 25-29.
- [4] JOSHI K R, KAMATHE R S. Quantification of retinex in enhancement of weather degraded images [C]//2008 International Conference on Audio, Language and Image Processing. Shanghai: IEEE, 2008: 1229-1233.
- [5] CAI B L, XU X M, JIA K, et al. DehazeNet: An end-to-end system for single image haze removal [J]. IEEE Transactions on Image Processing, 2016, 25(11): 5187-5198.
- [6] LIU X H, MA Y R, SHI Z H, et al. Griddehazenet: Attention-based multi-scale network for image dehazing [C]//2019 IEEE/CVF International Conference on Computer Vision. Seoul: IEEE, 2019: 7313-7322.
- [7] ZHENG Z R, REN W Q, CAO X C, et al. Ultra-high-definition image dehazing via multi-guided bilateral learning [C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Nashville: IEEE, 2021: 16180-16189.
- [8] SUN W J, CHEN Z Z. Learning many-to-many mapping for unpaired real-world image super-resolution and downscaling [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2024, 46(12): 9874-9889.
- [9] LAI W S, HUANG J B, AHUJA N, et al. Deep Laplacian Pyramid Networks for fast and accurate super-resolution [C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu: IEEE, 2017: 5835-5843.
- [10] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation [C]//2014 IEEE Conference on Computer Vision and Pattern Recognition. Columbus: IEEE, 2014: 580-587.
- [11] 曹宇,李佳阳,王芳. 基于改进YOLOv8n的水下鱼类目标识别轻量化模型[J]. 上海海洋大学学报, 2025, 34(1):188-200.
- CAO Y, LI J Y, WANG F. An improved YOLOv8n based on lightweight model for underwater fish target detection [J]. Journal of Shanghai Ocean University, 2025, 34(1):188-200.
- [12] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single shot MultiBoxdetector [C]//Proceedings of the 14th European Conference on Computer Vision - ECCV 2016. Amsterdam: Springer, 2016: 21-37.
- [13] 陈欣,万敏杰,马超,等. 采用多尺度特征融合SSD的遥感图像小目标检测[J]. 光学精密工程, 2021, 29(11): 2672-2682.
- CHEN X, WAN M J, MA C, et al. Recognition of small targets in remote sensing image using multi-scale feature fusion-based shot multi-box detector [J]. Optics and Precision Engineering, 2021, 29(11): 2672-2682.
- [14] WANG H L, QIAN H M, FENG S, et al. L-SSD: lightweight SSD target detection based on depth-separable convolution [J]. Journal of Real-Time Image Processing, 2024, 21(2): 33.
- [15] 赵晨,陈明. 水下底栖生物轻量化目标检测算法YOLOv7-RFPCW[J]. 农业工程学报, 2024, 40(11): 168-177.
- ZHAO C, CHEN M. A YOLOv7-RFPCW of a lightweight target detection algorithm for benthic organisms underwater [J]. Transactions of the Chinese Society of Agricultural Engineering, 2024, 40(11): 168-

- 177.
- [16] 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.
- [17] KIPOSAYER. Lignjev. 2 Vision Project [DB/OL]. Roboflow Universe, (2023-04-03) [2025-05-14]. <https://universe.roboflow.com/kiposlayer/lignjev.2>.
- [18] SHARMA P, BISHT I, SUR A. Wavelength-based attributed deep neural network for underwater image restoration [J]. *ACM Transactions on Multimedia Computing, Communications, and Applications*, 2023, 19(1): 1-23.
- [19] CHIANG J Y, CHEN Y C. Underwater image enhancement by wavelength compensation and dehazing [J]. *IEEE Transactions on Image Processing*, 2011, 21 (4): 1756-1769.
- [20] ZHANG X Y, ZHOU X Y, LIN M X, et al. ShuffleNet: An extremely efficient convolutional neural network for mobile devices [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018: 6848-6856.
- [21] TAN M X, LE Q V. EfficientNet: Rethinking model scaling for convolutional neural networks [C]//Proceedings of the 36th International Conference on Machine Learning. Mountain View, CA: PMLR, 2019: 6105-6114.
- [22] IANDOLA F N, HAN S, MOSKEWICZ M W, et al. SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and < 0.5 MB model size [J]. *arXiv preprint arXiv: 1602.07360*, 2016.
- [23] LIU G, REDA F A, SHIH K J, et al. Image inpainting for irregular holes using partial convolutions [C]//Proceedings of the 15th European Conference on Computer Vision -ECCV 2018. Munich: Springer, 2018: 89-105.
- [24] HOU Q B, ZHOU D Q, FENG J S. Coordinate attention for efficient mobile network design [C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Nashville: IEEE, 2021: 13708-13717.
- [25] WOO S, PARK J, LEE J Y, et al. CBAM: Convolutional block attention module [C]//Proceedings of the 15th European Conference on Computer Vision -ECCV2018. Munich: Springer, 2018: 3-19.
- [26] HU J, SHEN L, SUN G. Squeeze-and-excitation networks [C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, UT: IEEE, 2018: 7132-7141.
- [27] REZATOIGHI H, TSOI N, GWAK J, et al. Generalized intersection over union: A metric and a loss for bounding box regression [C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Long Beach, CA: IEEE, 2019: 658-666.
- [28] ZHENG Z H, WANG P, LIU W, et al. Distance-IoU loss: Faster and better learning for bounding box regression [C]//Proceedings of the 34th AAAI Conference on Artificial Intelligence. New York: AAAI, 2020: 12993-13000.
- [29] 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. Providence, RI: IEEE, 2012: 81-88.
- [30] 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.
- [31] ZHANG S, ZHAO S L, AN D, et al. LiteEnhanceNet: A lightweight network for real-time single underwater image enhancement [J]. *Expert Systems with Applications*, 2024, 240: 122546.
- [32] CONG R M, YANG W Y, ZHANG W, et al. PUGAN: Physical model-guided underwater image enhancement using GAN with dual-discriminators [J]. *IEEE Transactions on Image Processing*, 2023, 32: 4472-4485.
- [33] HUANG D M, WANG Y, SONG W, et al. Shallow-water image enhancement using relative global histogram stretching based on adaptive parameter acquisition [C]//Proceedings of the 24th International Conference on Multi Media Modeling. Bangkok: Springer, 2018: 453-465.
- [34] LI K Q, FAN H T, QI Q, et al. TCTL-Net: Template-free color transfer learning for self-attention driven underwater image enhancement [J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2023, 34 (6): 4682-4697.
- [35] REN S Q, HE K M, GIRSHICK R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks [J]. *arXiv preprint arXiv: 1506.01497*, 2016.

DUIE-YOLO: An image enhancement-based underwater squid target detection algorithm

CAO Liling, HU Haoyu, CAO Shouqi

(College of Engineering Science and Technology, Shanghai Ocean University, Shanghai 201306, China)

Abstract: To address the decline in target detection accuracy caused by blur and color deviation in underwater images and to improve the accuracy and robustness of squid detection in complex underwater environments, this study proposes an underwater squid detection algorithm named DUIE-YOLO based on image enhancement. The algorithm adopts a cascaded framework of “enhance first, detect later” consisting of the DUIE-Net enhancement module and the YOLOv8-HD detection module. The DUIE-Net module significantly improves image quality through color correction, multi-scale feature fusion, feature restoration and enhancement, and dehazing optimization. The YOLOv8-HD detection module combines the FasterNet network, a small-object detection head, the CoordAttention mechanism, and the ShapeIoU loss function to optimize feature extraction and small-object detection accuracy. Experimental results show that DUIE-YOLO outperforms the original YOLOv8n in four key metrics: Precision, Recall, F1-score, and mAP, with improvements of 4.2%, 6.8%, 5.7% and 5.5%, respectively. Joint experiments demonstrate that the combination of DUIE-Net and YOLOv8-HD achieves a 40.3% increase in mAP, a 10.5% increase in Precision, a 53% increase in Recall, and a 31% increase in F1-score compared to the baseline (Raw+YOLOv8n), proving the algorithm's significant cascaded optimization effect. The study indicates that DUIE-YOLO effectively mitigates the performance degradation caused by poor underwater image quality through the synergistic optimization of image enhancement and detection modules. This research provides a high-precision solution for target recognition in complex underwater environments, offering significant application value for marine biological monitoring and resource development.

Key words: underwater squid detection; object detection; image enhancement; multi-scale feature fusion; YOLOv8