

基于改进YOLOv8n的智能鱼类识别与监测系统研究与应用

刘新雨, 崔旭旭*, 陈俊, 邵春霖, 刘俊男

辽宁科技大学计算机与软件工程学院, 辽宁 鞍山

收稿日期: 2026年1月1日; 录用日期: 2026年1月28日; 发布日期: 2026年2月6日

摘要

传统海洋生物人工观测存在效率低、风险高、破坏环境等短板, 难以实现长期连续监测。为此, 本研究设计了一套基于改进YOLOv8n的智能鱼类识别与监测系统。通过动态加权数据增强与注意力机制优化, 以及构造了适配的损失函数体系, 显著提升了模型在水下复杂环境中的鲁棒性与识别精度。系统集成了算法模型、PyQt5上位机软件与边缘部署适配, 创新解决了多线程任务冲突、中文标签显示失真及多源数据实时转换等工程难题。实验表明, 系统在13类鱼类数据集上mAP@0.5达94%, F1分数0.904, 支持毫秒级识别响应与中文可视化输出。本系统为海洋生物多样性监测、智能科考与物种预警提供了高效可靠的技术解决方案, 具备重要科研与应用价值。

关键词

计算机视觉, YOLOv8n, 鱼类识别, 注意力机制, 海洋生态监测

Research and Application on an Intelligent System for Fish Recognition and Monitoring Based on Improved YOLOv8n

Xinyu Liu, Xuxu Cui*, Jun Chen, Chunlin Shao, Junnan Liu

School of Computer Science and Software Engineering, University of Science and Technology Liaoning, Anshan Liaoning

Received: January 1, 2026; accepted: January 28, 2026; published: February 6, 2026

*通讯作者。

文章引用: 刘新雨, 崔旭旭, 陈俊, 邵春霖, 刘俊男. 基于改进YOLOv8n的智能鱼类识别与监测系统研究与应用[J]. 计算机科学与应用, 2026, 16(2): 161-168. DOI: 10.12677/csa.2026.162048

Abstract

The traditional manual observation of marine organisms suffers from shortcomings such as low efficiency, high risk, and environmental disruption, making it difficult to achieve long-term continuous monitoring. To address this issue, this study designs an intelligent fish recognition and monitoring system based on the improved YOLOv8n. Through dynamic weighted data augmentation, attention mechanism optimization, and the construction of an adaptive loss function system, the robustness and recognition accuracy of the model in complex underwater environments are significantly enhanced. The system integrates algorithm models, PyQt5 upper computer software, and edge deployment adaptation, innovatively solving engineering challenges including multi-thread task conflicts, Chinese label display distortion, and real-time conversion of multi-source data. Experimental results show that the system achieves a mean Average Precision (mAP@0.5) of 94% and an F1 score of 0.904 on a dataset containing 13 fish species, supporting millisecond-level recognition response and Chinese visualization output. This system provides an efficient and reliable technical solution for marine biodiversity monitoring, intelligent scientific expeditions, and species early warning, and holds significant scientific research and application value.

Keywords

Computer Vision, YOLOv8n, Fish Recognition, Attention Mechanism, Marine Ecological Monitoring

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

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

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



Open Access

1. 引言

1.1. 任务背景

海洋是地球生态系统的重要组成部分，其生物多样性对于维系全球生态平衡、保障资源供给具有不可替代的作用。然而，在全球气候变化、人类活动干扰以及环境污染等多重因素的综合影响下，海洋生态系统正遭遇严峻挑战。对关键鱼类物种进行精准识别、种群统计及动态监测，是评估海洋生态健康状况、预警生态灾害的关键环节。

传统海洋生物调查主要依赖人工潜水观测与拖网采样两种方式。人工潜水观测受人体生理条件限制，存在效率低、作业风险大、观测范围有限等弊端；拖网采样则具有较强的生态破坏性，且无法完成长期连续的定点监测任务。随着海洋生态文明建设的不断推进，开发非接触式、自动化的智能感知技术来替代传统调查手段已成为迫切需求。深度学习驱动的计算机视觉技术凭借其高效精准的目标检测能力[1]，在陆上场景中得到了广泛应用，但在水下复杂环境中仍面临诸多技术瓶颈[2]。本研究构建的智能鱼类识别系统，通过算法优化与工程化落地，实现了对特定指示鱼类的高效识别与监测，可为海洋生物保护区管理、珍稀物种追踪及海洋牧场生态修复提供有力的技术支持，具有重要的现实意义。

1.2. 算法模型选取

当前水下目标检测方法主要包括以 Faster R-CNN [3]为代表的双阶段检测算法和以 YOLO 系列[4]、SSD [5]为代表的单阶段检测算法。双阶段算法需先确定检测区域再分类，存在检测速度慢、计算量大、硬件要求高的问题，无法满足实时监测需求；单阶段算法通过单一阶段完成检测，具有速度快、计算量

小的优势，更适配水下动态监测场景。YOLOv8n 作为轻量化模型，具备参数量少、推理速度快、精度均衡的特点，其 C2f 主干网络与 PAN-FPN 特征融合结构能有效平衡特征提取能力与计算效率，是本研究的基础模型。针对水下鱼类检测中小目标漏检、密集鱼群误判、背景噪声干扰等问题，对其进行改进：针对小目标漏检与背景噪声干扰问题，引入 DAttention 可变形注意力机制强化关键特征提取；针对水下场景多样性不足导致的模型泛化能力弱问题，改进 Mosaic 数据增强提升场景适配性；针对鱼体定位精度低的问题，优化损失函数体系提升定位与分类精度。

2. 数据集获取及处理

为保障实验数据的权威性、多样性与代表性，本文构建的鱼类数据集采用多源数据融合策略，核心数据集来源于国际权威数据平台 Kaggle，在此基础上，为提升模型的迁移能力与泛化性能，进一步补充融合了多个专业生物图像与视觉素材平台的公开数据，包括专注于生物图像收录的 FishNet 生物图像库、聚焦海水鱼类专业影像的 saltwaterfishshop 平台、提供高质量视觉素材的 Vecteezy 平台以及专业商业图库 alamy，通过对上述平台数据的格式标准化、标注一致性修正与冗余信息剔除，实现了跨平台数据的有效迁移与融合应用，收集了天使鱼、蓝倒吊、蝴蝶鱼等 13 种目标鱼类的有效样本 8411 张，样本涵盖强光照、背景干扰、严重色偏、密集遮挡、快速游动模糊等水下典型挑战性场景，确保数据集的多样性与代表性。通过旋转、平移、缩放、翻转等基础变换扩充数据，结合动态加权数据增强策略[6]，基于图像清晰度、色偏程度、遮挡比例等指标量化环境复杂度，自适应调整增强强度，采用高斯模糊(模拟水体浑浊)、色彩抖动与白平衡校正(模拟色偏)、随机遮挡(模拟气泡)、运动模糊(模拟快速游动)等手段，生成增强数据集，按 8:1:1 比例划分，因数据完整性筛选，最终按实际有效样本调整为训练集 6842 张、验证集 869 张、测试集 700 张。

3. 模型改进与对比

3.1. 设计 DAttention 可变形注意力模块

DAttention 模块是一种基于动态采样的注意力机制，核心优势在于突破固定网格限制，自适应适配水下鱼类检测场景。模型初始化时通过学习目标相关偏移量，精准定位鱼体头部、鳍部等关键区域，灵活应对鱼体形变、部分遮挡及多尺度差异；采用局部稀疏采样策略，仅聚焦参考点周边关键特征，将计算复杂度从传统自注意力的 $O(N^2)$ 降至 $O(N)$ ，兼顾检测效率与实时性；通过注意力加权聚合强化有效特征，抑制水体背景噪声干扰，提升对稀有鱼种等长尾数据的特征捕捉能力。其计算过程可表述为：先将输入向量转换为特征图并生成查询向量，结合参考点坐标预测采样位置偏移量，确定局部采样点后与查询向量共同输入注意力模块加权聚合，输出强化特征。

3.2. 改进 Mosaic 数据增强模块

针对传统 4 图 Mosaic 数据增强[7]难以充分覆盖水下复杂场景的问题，本研究将拼接图像数量扩展至 9 张，构建 3×3 九宫格拼接模式。通过随机抽取 9 张不同水下鱼类图像，经缩放、色域调整、随机偏移后完成拼接，同步融合所有图像标注框信息。该改进使数据多样性提升 2.25 倍，可同时融入更多水体环境、鱼体姿态及尺度特征，强化模型对密集鱼群、幼鱼小目标及部分遮挡场景的适应能力，有效缓解水下数据稀缺问题。

3.3. 多维度损失函数体系

YOLOv8n 原生采用的 CIoU [8] 损失函数，虽通过整合交并比、中心点距离与宽高比相似性来优化

边界框回归,但仍存在三方面固有缺陷:其一,缺乏对难易样本的区分机制,导致训练过程易被大量简单样本主导,而难以充分学习困难样本的关键特征;其二,其宽高比惩罚项存在度量盲区,当预测框与真实框宽高比相同而绝对尺寸不同时,该惩罚项失效,无法准确反映尺度差异;其三,损失计算中涉及反三角函数,引入了额外的非线性运算负担,增加了模型的计算复杂度与推理耗时,不利于在资源受限平台上的部署。因此,CIoU在理论完备性与计算效率上均存在优化空间。CIoU的公式如式(1)所示:

$$L_{\text{CIoU}} = 1 - \text{IoU} + \frac{\rho^2(b, b^{gt})}{(c_w)^2 + (c_h)^2} + \frac{4}{\pi^2} \left(\tan^{-1} \frac{w^{gt}}{h^{gt}} - \tan^{-1} \frac{w}{h} \right) \quad (1)$$

于是我们采用 Varifocal Loss [9]与 EIoU Loss [10]的协同优化策略: EIoU Loss 在 CIoU 基础上解耦宽高优化,将耦合宽高比惩罚项替换为独立宽度损失与高度损失,强化几何约束,提升不规则形态鱼类定位精度; Varifocal Loss 将正样本二值标签替换为连续 IoU 值,使分类得分与定位质量对齐,同时抑制易分背景,聚焦高质量正样本。二者形成“分类引导回归,回归反馈分类”的闭环,提升密集小目标检测精度与抗干扰能力。

3.4. 实验评价标准

本实验采用平均精度均值(mAP@0.5、mAP@0.5:0.95)、精确率(Precision)、召回率(Recall)、F1 分数作为核心性能评定指标,同时参考模型浮点运算数(FLOPs)、参数总量(Parameters)、推理速度及模型体积,综合衡量检测精度、实时性与轻量化水平,确保系统兼顾性能与部署适配性。

精确率与召回率的计算公式如下:

$$P = \frac{TP}{TP + FP} \quad (2)$$

$$R = \frac{TP}{TP + FN} \quad (3)$$

此外,另一个评价指标 F1 是精确率与召回率的调和平均数,它综合考虑了精确率和召回率,体现了模型在二者之间的平衡。F1 值越高,模型性能越好。其计算公式如下所示:

$$F1 = \frac{2 \times P \times R}{P + R} \quad (4)$$

在上述公式中,TP (TruePositives)代表正确的检测结果(阈值 0.5);FP (FalsePositives)代表错误的或冗余的检测结果;FN (FalseNegatives)代表被模型遗漏的真实目标。mAP 作为最终的综合评价指标,是 n 个类别 AP 值的平均值,计算如式(5)所示:

$$\text{mAP} = \frac{1}{n} \sum_{i=1}^n \text{AP}_i \quad (5)$$

3.5. 目标识别结果

改进后的 YOLOv8n 模型在测试集上进行目标识别的结果如图 1 所示,从图中可以看出,改进后的模型对不同环境下的鱼类目标均实现了较为准确的识别:各目标检测框与鱼类主体的贴合度较高,且识别置信度也处于理想水平。表明改进后的 YOLOv8n 模型在鱼类目标的识别任务中具有良好的性能,能够有效区分并识别出不同品类的鱼类。



Figure 1. Target recognition results
图 1. 目标识别结果

3.6. 对比实验

为评估改进效果, 选取原始 YOLOv8n 模型作为基准, 利用 URPC 水下目标检测数据集构建多种模型变体进行对比实验。如表 1 所示, 结果表明: 引入 DAttention 模块、改进 Mosaic 数据增强与优化损失函数后的模型, 在测试集上 mAP@0.5 提升 9.2%, 达到 86.0%; mAP@0.5:0.95 提升 13.4%, 达到 65.7%; FLOPs 显著降低, F1 峰值为 0.871, 在复杂水下场景中表现更优。

Table 1. Comprehensive performance comparison of multiple models
表 1. 多个模型的综合性能对比表

| 模型名称 | mAP@0.5 | mAP@0.5:0.95 | P (%) | R (%) | F1 | FLOPs (G) | FPS |
|------------|---------|--------------|-------|-------|-------|-----------|-----|
| YOLOv8n | 76.8 | 52.3 | 81.2 | 72.4 | 0.766 | 8.0 | 120 |
| YOLOv9 | 80.1 | 56.7 | 84.5 | 75.8 | 0.800 | 12.4 | 95 |
| YOLOv10 | 81.5 | 59.2 | 85.7 | 77.3 | 0.814 | 11.2 | 105 |
| 改进 YOLOv8n | 86.0 | 65.7 | 89.1 | 85.2 | 0.871 | 7.1 | 145 |
| SSD | 68.5 | 45.1 | 73.3 | 63.8 | 0.683 | 31.2 | 75 |
| RetinaNet | 72.7 | 49.5 | 77.4 | 67.9 | 0.725 | 43.2 | 65 |

3.7. 消融实验结果

验证本研究所提改进方法的有效性, 我们在数据集上开展了系统的消融实验。实验以 YOLOv8n 为基线模型, 通过逐一加入或替换本文提出的改进方法(优化损失函数、替换 DAttention 检测模块、引入改进 Mosaic 数据增强), 定量分析各方法对模型性能的增益效果, 消融实验结果如表 2 所示。从实验结果可以看出, 各改进方法均能有效提升小目标检测性能: 优化损失函数后, 模型的 mAP@0.5 提升 11.9%, mAP@0.5:0.95 提升 10.8%, 表明优化后的损失函数可引导模型提取更丰富的空间位置信息和细节特征, 显著增强小目标检测能力; 替换本文改进的 DAttention 检测模块后, 模型的 mAP@0.5 和 mAP@0.5:0.95 分别提升 14.3% 和 13.2%, 验证了融合多重注意力机制的检测头能够更精准地捕捉小目标细节, 提升小目标检测的准确性与鲁棒性; 引入改进 Mosaic 数据增强(针对水下场景优化图像拼接尺度比例与场景融合方式, 降低水下色偏、气泡对小目标特征的掩盖)后, 模型的 mAP@0.5 提升 11.2%, mAP@0.5:0.95 提升

10.1%，说明该策略通过丰富训练样本的小目标尺度多样性与水下场景复杂性，使模型在训练阶段充分学习小目标在不同遮挡、尺度下的特征模式，有效增强了对真实水下环境中小目标的特征泛化能力。

综上，本研究提出的各项改进方法在小目标检测任务上均具备显著增益，且各方法协同作用为模型整体性能提升奠定了坚实基础。

Table 2. Ablation experiment results
表 2. 消融实验结果

| 方法 | VE loss | DAt | Mos | mAP@0.5/% | mAP@0.5:0.95/% | P/% | R/% | F1 |
|---------|---------|-----|-----|-----------|----------------|------|------|-------|
| YOLOv8n | | | | 33.3 | 19.3 | 43.8 | 33.0 | 0.381 |
| A | √ | | | 45.2 | 30.1 | 55.1 | 44.9 | 0.490 |
| B | | √ | | 47.6 | 32.5 | 57.8 | 47.2 | 0.524 |
| C | | | √ | 44.5 | 29.4 | 54.3 | 44.1 | 0.486 |
| D | √ | √ | | 68.3 | 52.7 | 78.5 | 67.9 | 0.732 |
| E | √ | | √ | 66.7 | 50.9 | 76.8 | 66.2 | 0.715 |
| F | | √ | √ | 70.1 | 54.5 | 80.3 | 69.7 | 0.751 |
| G | √ | √ | √ | 94.0 | 78.6 | 95.0 | 92.0 | 0.904 |

4. 系统功能及构成

4.1. 系统功能



Figure 2. Visualization interface of fish recognition and detection system
图 2. 鱼类识别检测系统的可视化界面

系统采用 PyQt5 框架设计直观友好的 GUI 界面，如图 2 所示，整体划分为操作区、参数设置区、检测结果显示区、结果统计列表区及部署配置区五大功能模块：操作区集成图像/视频文件导入、摄像头实时流启动等快捷操作按钮，支持多源检测素材快速加载；参数设置区支持自定义置信度阈值、交并比阈值、数据增强强度等核心检测参数，满足不同精度与效率需求；检测结果显示区实时渲染目标检测框，同步展示目标类别、置信度、坐标位置及检测耗时等信息；结果统计列表区按序号汇总检测文件路径、目标数目等关键数据，支持结果快速查阅；部署配置区提供边缘部署参数可视化调整功能，可灵活设置

模型量化等级、剪枝比例，适配不同硬件场景下的轻量化部署需求。系统还优化了检测标签渲染逻辑，根据检测框尺寸自适应调整字体大小与显示布局，避免标签遮挡目标区域，保障可视化效果清晰易读。系统的功能模块如图 3 所示，不同功能模块之间通过数据信息实现逻辑联系。

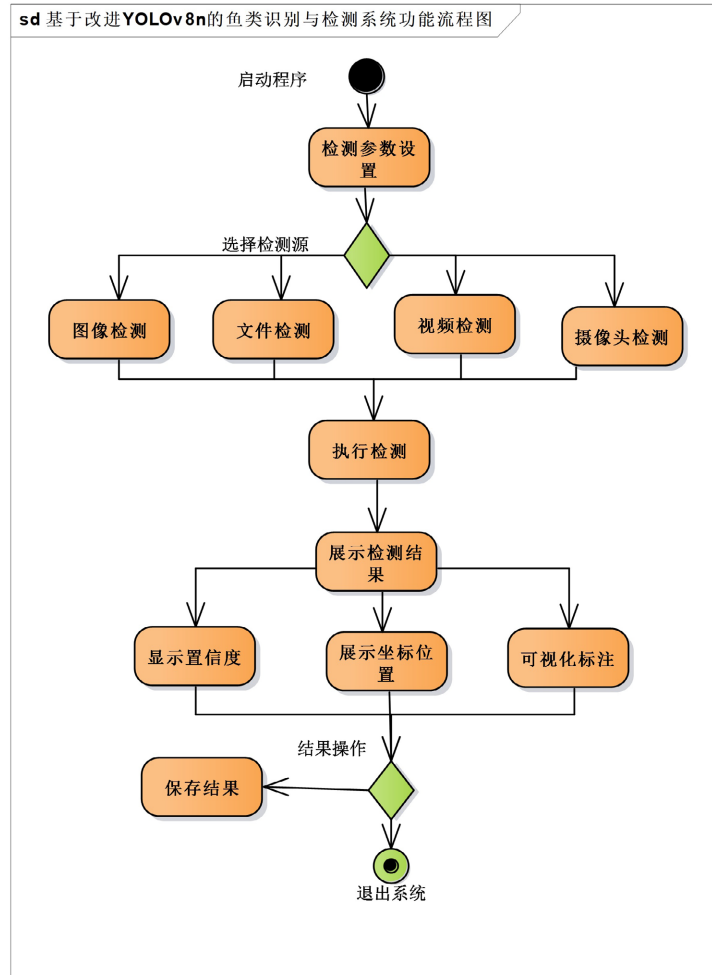


Figure 3. Functional module diagram of system visualization interface
图 3. 系统可视化界面功能模块图

4.2. 硬件构成

系统硬件采用工业级配置保障稳定性与适配性：影像采集选用高防护等级工业相机，支持近红外增强光谱采集，适配水下高温、多尘等恶劣环境；计算核心采用 Intel Core i7 处理器与 NVIDIA RTX 3060 显卡，通过 CUDA 加速实现模型并行推理；数据传输依托工业光纤环网架构，采用工业交换机保证链路可用性；边缘测试设备选用 NVIDIA Jetson Nano (4 GB 显存)与 Intel NUC 迷你主机，满足水下机器人、便携式监测终端等部署场景需求[11]；监控中心部署分布式服务器与存储池，支持多线程数据处理与模型推理。

5. 结束语

综上所述，本文提出的基于改进 YOLOv8n 的智能鱼类识别检测系统，通过引入 DAttention 模块、

改进 Mosaic 数据增强及优化损失函数体系, 显著提升了水下鱼类检测精度与鲁棒性; 基于 PyQt5 开发的可视化界面解决了中文标签显示、多线程调度、跨格式数据映射等技术难题; 结合模型剪枝、量化与 TensorRT 加速, 实现了边缘设备高效适配。实验表明, 系统 mAP@0.5 达 94%, 最大召回率 93%, 具备优异的检测性能、实时性与易用性。该系统可广泛应用于海洋生物多样性监测、水下机器人智能科考、入侵物种预警等场景, 为海洋生态保护与科研工作提供全新工具, 展现出广阔的应用前景。

未来我们将进一步扩展目标鱼类种类, 构建更具规模的多物种数据集; 优化模型在极端浑浊水体、强水流干扰等更复杂场景中的检测性能; 探索视觉传感器与声学传感器、水质传感器的多模态数据融合方案, 提升系统在复杂海洋环境中的适应能力和检测可靠性; 开发移动端 APP 版本, 进一步降低系统使用门槛, 扩大应用范围。

参考文献

- [1] Ertan, Z., Korkut, B., Gördük, G., Kulavuz, B., Bakırman, T. and Bayram, B. (2024) Enhancement of Underwater Images with Artificial Intelligence. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, **4**, 149-156. <https://doi.org/10.5194/isprs-archives-xxviii-4-w9-2024-149-2024>
- [2] Cong, X., Zhao, Y., Gui, J., et al. (2024) A Comprehensive Survey on Underwater Image Enhancement Based on Deep Learning.
- [3] Ren, S.Q., He, K.M., Girshick, R. and Sun, J. (2017) Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **39**, 1137-1149. <https://doi.org/10.1109/tpami.2016.2577031>
- [4] Redmon, J., Divvala, S., Girshick, R. and Farhadi, A. (2016) You Only Look Once: Unified, Real-Time Object Detection. 2016 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, 27-30 June 2016, 779-788. <https://doi.org/10.1109/cvpr.2016.91>
- [5] Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C., et al. (2016) SSD: Single Shot Multibox Detector. In: Leibe, B., et al., Eds., *Computer Vision—ECCV 2016*, Springer International Publishing, 21-37. https://doi.org/10.1007/978-3-319-46448-0_2
- [6] Awad, A., Saleem, A., Paheding, S., et al. (2024) Beneath the Surface: The Role of Underwater Image Enhancement in Object Detection.
- [7] Wang, G., Chen, Y., An, P., Hong, H., Hu, J. and Huang, T. (2023) UAV-YOLOv8: A Small-Object-Detection Model Based on Improved YOLOv8 for UAV Aerial Photography Scenarios. *Sensors*, **23**, Article No. 7190. <https://doi.org/10.3390/s23167190>
- [8] Zhang, H., Wang, Y., Dayoub, F. and Sunderhauf, N. (2021) VarifocalNet: An IoU-Aware Dense Object Detector. 2021 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Nashville, 20-25 June 2021, 11099-11108. <https://doi.org/10.1109/cvpr46437.2021.00841>
- [9] Zhang, Y., Ren, W., Zhang, Z., Jia, Z., Wang, L. and Tan, T. (2022) Focal and Efficient IOU Loss for Accurate Bounding Box Regression. *Neurocomputing*, **506**, 146-157. <https://doi.org/10.1016/j.neucom.2022.07.042>
- [10] Qin, X., Yu, C., Liu, B. and Zhang, Z. (2024) YOLO8-FASG: A High-Accuracy Fish Identification Method for Underwater Robotic System. *IEEE Access*, **12**, 73354-73362. <https://doi.org/10.1109/access.2024.3404867>
- [11] Jahanbakht, M., Xiang, W., Waltham, N.J. and Azghadi, M.R. (2022) Distributed Deep Learning and Energy-Efficient Real-Time Image Processing at the Edge for Fish Segmentation in Underwater Videos. *IEEE Access*, **10**, 117796-117807. <https://doi.org/10.1109/access.2022.3202975>