

基于改进 YOLOv7 的轻量化海洋生物检测算法

郑凯东¹ 蒋瑞娜¹

ZHENG Kaidong JIANG Ruina

摘要

水下设备在内存和计算能力方面的限制, 以及其复杂的运行环境, 导致检测过程中频繁出现误检和漏检现象。为解决这些难题, 文章提出了一种耗资少且检测效率与精度高的轻量化海洋生物检测算法 YOLOv7-RDBC。具体而言, 在 YOLOv7 骨干网络中融入 RepViT 轻量级网络, 设计 D-ELAN 和 D-MPConv 模块, 进一步实现网络的轻量化, 加快检测速度。另外, 借助 BiFPN 实施多尺度特征融合, 再结合 CBAM 卷积注意力机制, 增强关键特征的表达能力, 提高检测精度。实验结果显示, 相较于传统的 YOLOv7 模型, YOLOv7-RDBC 在多个性能指标上都有明显提升, 在海洋生物检测任务中表现出色, 具备广阔的应用前景。

关键词

RepViT 轻量级网络; YOLOv7; 多尺度特征融合; 注意力机制; 目标检测; 水下环境

doi: 10.3969/j.issn.1672-9528.2025.07.005

0 引言

海洋作为地球上重要的资源储备库, 随着陆地资源日益枯竭, 人类对海洋资源的需求愈发迫切, 海洋开发成为新热点^[1]。近年来, 水下探测技术取得突破性进展, 为海洋勘探活动提供了有力技术支持, 推动其快速发展。然而水下目标检测技术作为海洋开发关键, 面临着严峻挑战。一方面是复杂的水下环境使光的衰减与散射效应增强, 导致目标特征退化, 极大增加检测难度。另一方面是水下设备的计算和存储能力有限, 要求检测算法必须轻量化^[2]。所以研发适应复杂水下环境、满足设备性能要求的高效检测算法, 是推动海洋开发进程的关键所在。

在目标检测技术的发展过程中, 基于深度学习的算法已经形成了两大技术路线。两阶段检测器^[3]通过区域建议机制实现了高精度的目标检测, 但其计算复杂度比较高, 难以满足实时性的需求。相比之下, 单阶段检测器^[4]采用端到端的架构, 在速度性能上具有显著优势, 更适合水下设备的应用场景, 但是当前的研究大多通过架构的扩展来提升检测性能, 这在一定程度上导致了模型复杂度与检测效率之间的不平衡问题。尽管如此, 仍然有一些研究者的研究方法在模型的轻量化与实时性协同优化方面取得了阶段性进展。例如, Zhang 等人^[5]通过引入 MobileNetV3 优化 YOLOv4, 实现了甲壳类动物的高效检测, Chen 等人^[6]在 YOLOv7 中引入 Conv2Former, 强化了特征融合能力, Li 等人^[7]通过三重注意机制改进 YOLOv5 的小目标检测能力, 另外有研究基于

YOLOv3-SPP 改进, 有效解决了误检问题^[8]。这些方法虽然在一定程度上取得了阶段性进展, 但是在模型轻量化与实时性协同优化方面仍存在明显不足。

为了进一步提升目标检测的性能, 一些研究引入了先进的特征融合网络, 如双向特征金字塔网络 (BiFPN)。BiFPN 通过双向跨尺度连接和加权特征融合机制, 解决了传统特征金字塔网络在处理多尺度目标时的局限性。通过在其自上而下与自下而上的双向路径中引入加权机制, 动态地调整各输入特征在输出结果中的权重占比, 以此实现更具效率的特征融合过程。

基于上述内容, 本研究提出了一种基于 YOLOv7 的创新架构 YOLOv7-RDBC。该架构以 YOLOv7^[9] 为基线模型, 通过多尺度特征优化和模型压缩技术, 构建了具有检测精度高与计算效率高的水下目标检测解决方案。研究着重解决了现有算法在海洋特殊环境下的适应性不足问题, 为水下设备的实时目标检测提供了新的技术范式。YOLOv7-RDBC 通过引入高效的特征提取模块和经过优化的网络结构, 显著降低了模型复杂度的同时提升了模型的目标检测精度。该架构还借鉴了剪枝和量化等模型压缩技术, 进一步优化了模型的计算效率。通过这些改进, YOLOv7-RDBC 模型不仅在检测精度上优于现有方法, 还在实时性方面表现出色, 为水下目标检测领域提供了新的研究方向。

1 基于改进 YOLOv7 的轻量化目标检测算法

改进的 YOLOv7 算法通过三大关键优化实现了高效性和适应性的提升。第一是骨干网络引入 RepViT 模块, 融合

1. 西安石油大学 陕西西安 710065

了视觉 Transformer 和轻量级 CNN 的优势, 利用轻量级自注意力机制和结构重参数化技术, 大幅降低了参数量和计算复杂度, 从而提升了训练与推理效率; 第二是特征融合模块采用 D-ELAN、D-ELAN-H 和 D-MPConv 模块, 以深度可分离卷积替代传统卷积, 减少参数量并提升推理速度; 第三是特征金字塔网络引入 BiFPN, 结合卷积注意力模块 (CBAM) 的优化策略, 增强了多尺度目标检测能力。图 1 展示了 YOLOv7-RDBC 模型架构。这些改进在保持检测精度的同时, 显著降低了计算资源消耗, 提升了轻量化程度和实时性表现, 为水下目标检测提供了更高效、更适应的解决方案。

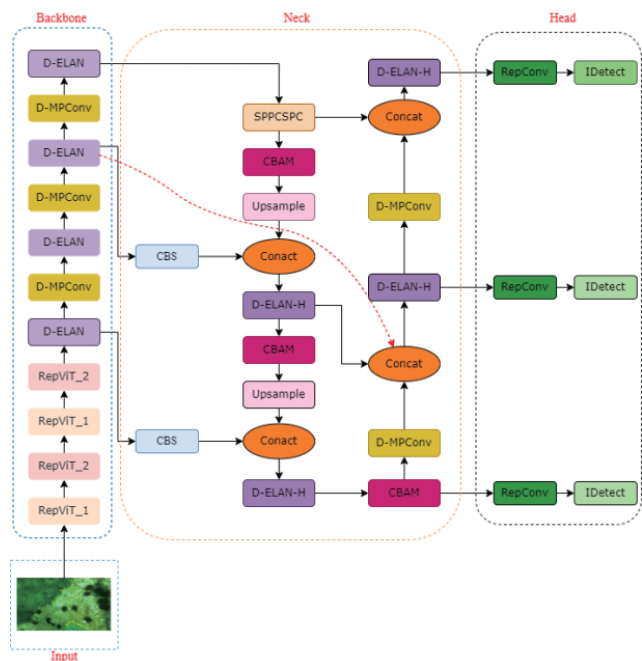


图 1 YOLOv7-RDBC 模型架构图

1.1 RepViT 的骨干网络设计

在 YOLOv7 中, 其骨干网络基于 CSPDarknetV3 架构, 虽具备强大的特征提取能力, 但因参数量较大, 导致检测速度受限^[10]。所以本文引入 RepViT 模块替换传统 CBS 模块, 以优化骨干网络。RepViT 融合了视觉 Transformer (ViT) 的高效架构设计与轻量级卷积神经网络 (CNN) 的优势, 通过轻量级自注意力机制、优化的激活与归一化层以及高效的下采样策略, 显著降低了模型的参数量与计算量^[11]。而且 RepViT 采用结构重参数化技术, 提升训练效率的同时减少推理阶段的计算与内存成本。通过调整网络宽度、扩展比以及跨块放置 SE 层, RepViT 在保持轻量化的基础上, 可以实现性能与延迟的最佳平衡, 如图 2 所示。

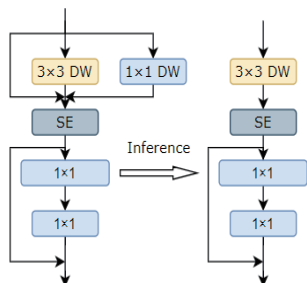


图 2 RepViT 模块

1.2 ELAN 和 MPConv 改进模块

为有效削减模型参数量, 并显著提升检测效率, 设计了 D-ELAN、D-ELAN-H 以及 D-MPConv 模块。在设计过程中, 运用深度可分离卷积, 对 ELAN 和 MPConv 模块中的传统标准卷积进行替换, 减少计算复杂度和参数量。这在保留信息传递的同时, 显著减少了浮点运算量和模型参数, 如图 3 所示。

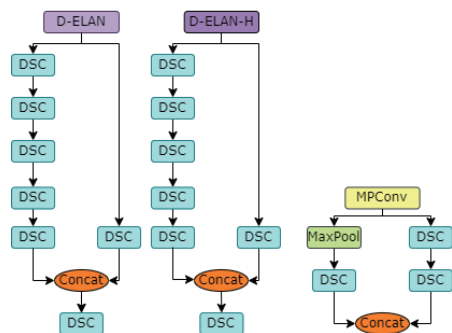


图 3 D-ELAN 模块、D-ELAN-H 模块和 D-MPConv 模块

假设输入特征张量的空间分辨率为 $D_K \times D_K$, 卷积滤波器的空间覆盖范围为 $D_F \times D_F$, 当输入特征的通道数量设定为 M , 输出特征的通道数量设定为 N 时, 深度可分离卷积 (DSC) 模块的总参数计算量计算公式为:

$$C = D_K \times D_K \times M \times N + D_K \times D_K \times D_F \times M \quad (1)$$

用于计算标准卷积参数量的公式为:

$$C = \frac{D_K \times D_K \times M \times N + D_K \times D_F \times D_F \times M}{D_K \times D_K \times M \times N \times D_K \times D_F} \quad (2)$$

两者之间的比值公式为:

$$\frac{D_K \times D_K \times M \times N + D_K \times D_F \times D_F \times M}{D_K \times D_K \times M \times N \times D_K \times D_F} = \frac{1}{N} + \frac{1}{D_F^2} \quad (3)$$

根据式 (3) 的分析, 前者计算量低于后者。因此, 在 ELAN 和 MPConv 模块的重构过程中, 采用深度可分离卷积替代传统标准卷积, 能够减少模型的参数量。模型参数量的减少, 不仅降低了计算复杂度, 还使得模型在推理过程中所需处理的数据量降低, 进而提升了模型的推理速度。

1.3 BiFPN 与 CBAM 结合的特征增强网络

为解决水下目标检测中多尺度特征融合效率低下以及遮挡目标表征能力不足的问题, 提出了一种结合加权双向特征金字塔网络 (BiFPN) 与卷积注意力模块 (CBAM) 的优化策略。BiFPN 替代了传统的 PANet 结构^[12], 在双向特征金字塔中嵌入可学习的权重参数, 动态调节各层级特征的贡献, 优化跨尺度特征融合能力。通过重复利用同一层级的输入与输出节点, 并增加跨层跳跃连接, 提升了浅层细节特征与深层语义信息之间的交互效率。PANet 结构图和 BiFPN 结构如图 4 所示。

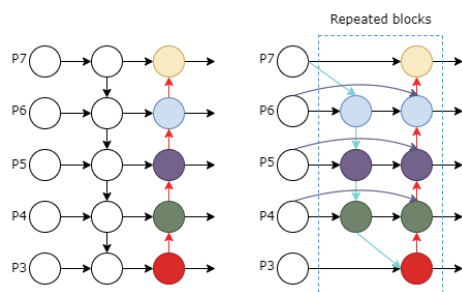


图4 PANet 结构和 BiFPN 结构

在此基础上, 针对复杂水下环境中光照不均匀和目标遮挡导致的特征细节丢失问题, 进一步引入了 CBAM 卷积注意力机制。CBAM 通过通道和空间双重注意力机制对特征图进行建模, 从而增强关键特征的表达能力。首先利用通道注意力模块对特征图的通道进行加权, 突出重要通道的特征信息, 接着通过空间注意力模块对特征图的空间维度进行加权, 聚焦于目标区域, 从而提升模型对关键特征的感知能力。BiFPN 与 CBAM 的协同作用形成了全局与局部特征的互补优化框架, 显著提升了模型对模糊目标和遮挡目标的检测能力。

2 实验结果分析

2.1 所用数据集

本研究采用 2020 年全国水下机器人专业比赛 (URPC) 数据集, 该数据集涵盖扇贝、海星、海胆和海参这四种不同海洋生物的水下光学图像, 保留 6 450 张有效图像。为保证数据分布的均衡性, 采用随机划分的方式将数据集分为训练集和测试集, 划分比例为 7:3。其中, 训练集包含 4 515 张图像, 测试集包含 1 935 张图像。数据集如图 5 所示。

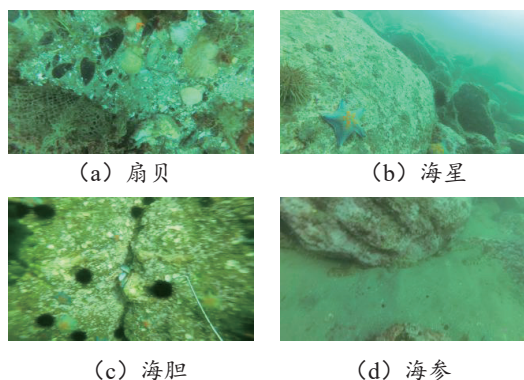


图5 URPC2020 数据集

2.2 评价标准

为全面评估改进后的 YOLOv7 算法性能, 本文从检测精度和速度两个维度分析。检测精度上, 选用精确率 (P)、召回率 (R)、平均精度 (AP) 和平均精度均值 (mAP), 衡量算法对不同目标的识别准确性。检测速度方面, 用帧率 (FPS) 体现实时处理能力, 借助浮点运算次数 (GFLOPs) 评估计算量, 通过参数量 (params) 衡量模型复杂度。平均

精度均值 mAP 越高, 意味着模型在各类目标识别中的整体性能越优异, 对目标的识别精度也越高, 其计算公式如式 (4) 所示。AP 是对精确率-召回率 (PR) 曲线下面积的量化体现, 曲线下面积越大, 表明模型在对应类别上的识别精度越高, 具体计算参考式 (5)。另外, 精确率和召回率同样是评估模型性能的重要指标, 分别从不同角度反映模型的表现, 其计算分别参考式 (6) 和式 (7)。通过综合分析, 能准确验证改进算法在实际应用中的优势。

$$mAP = \frac{1}{c} \sum_{i=1}^c AP(i) \quad (4)$$

$$AP = \int_0^1 P(r) dr \quad (5)$$

$$P = \frac{TP}{TP + FP} \quad (6)$$

$$R = \frac{TP}{TP + FN} \quad (7)$$

2.3 实验过程

为验证本文提出的改进算法的有效性, 分别对原始 YOLOv7 算法和 YOLOv7-RDBC 算法在 URPC2020 数据集上进行了训练。在实验过程中, 确保所有参数设置和条件保持一致, 将迭代次数设置为 100 次, 权重衰减系数设置为 0.000 5, 初始学习率设置 0.001, 批量大小设置 16, 模型的训练结果如表 1 所示。

表1 改进前后模型训练结果对比

类别	YOLOv7 训练结果	YOLOv7-RDBC 训练结果
扇贝	0.792	0.839
海星	0.893	0.916
海胆	0.898	0.902
海参	0.736	0.764
所有类别	0.841	0.865
(mAP@0.5)		

通过表 1 可以看出, 以平均精度均值 (mAP) 衡量, 该算法的性能从 84.1% 提升至 86.5%, 涨幅达 2.4%。在具体目标类别中, 扇贝的平均精度 (AP) 提升幅度为 4.7%; 海星的 AP 值增长了 1.9%; 海胆的 AP 值提升了 0.4%; 海参的 AP 值增高了 3.2%。数据表明, 改进后的 YOLOv7 算法在不同海洋生物目标检测中, 均展现出更优的性能表现。

2.4 实验结果和分析

消融实验的目的是对每个模型的改进版本进行测评其效果的变化。以 YOLOv7 为基准模型, 在 URPC2020 数据集上添加不同模块评估性能, 环境及参数统一。如表 2 所示。实验结果表明, 引入 RepViT 轻量级网络模块和深度可分离卷积 (DSC) 使模型参数量显著减少, 帧率提升 34.1%, 但导致平均精度 (mAP) 下降 1.9%; 通过整合加权双向特征金字塔网络 (BiFPN), 达成多尺度特征的动态融合效果, mAP

回升至 85.3%，最终结合 CBAM 卷积注意力机制，模型在参数量和计算量仅微量增加的情况下，mAP 提升至 86.5%，较基准模型实现 2.4% 的精度增益。改进后的 YOLOv7 模型在保证实时性的同时，达成参数量降低、计算量减少且检测精度显著提升的均衡优化效果，为复杂水下场景的目标检测提供了有效的轻量化解决方案。

表 2 消融实验中每个模型的评估指标

Model	RepViT	DSC	BiFPN	CBAM	FPS/(帧 · s ⁻¹)	Params/10 ⁶	GFLOPs	mAp/%
YOLOv7	×	×	×	×	68.4	36.5	105.1	84.1
YOLOv7	√	×	×	×	77.6	24.5	87.5	83.3
YOLOv7	√	√	×	×	102.5	13.6	58.9	82.2
YOLOv7	√	√	√	×	92.3	15.2	62.4	85.3
YOLOv7	√	√	√	√	101.4	16.2	63.4	86.5

3 结语

为解决当前水下目标检测算法普遍存在的问题，本文提出了一种轻量级海洋生物检测算法 YOLOv7-RDBC。该算法在维持高检测精度的前提下，借助创新性的结构设计与优化策略，有效减少了模型的参数量，大幅降低了计算复杂度，模型的实时性与适应性也得到显著提升。在算法设计上选用轻量级的 RepViT 对 YOLOv7 骨干网络进行改进，结合深度可分离卷积的特性，设计了 D-ELAN 和 D-MPConv 模块，同时引入 CBAM 卷积注意力机制和 BiFPN 改进特征融合部分，有效解决了复杂水下环境中遮挡目标和小目标检测的难题。实验结果表明，YOLOv7-RDBC 的 mAP 达到 86.5%，处理速度为 101.4 帧 /s，在精度和速度上都有出色的表现，为水下生物探测提供了新思路。尽管如此，该网络结构仍有改进空间，在未来的研究中，将围绕轻量化和加速方法在水下目标检测领域的应用展开深入探索，以进一步提升模型的实时性和适应性。

参考文献:

[1] ZHOU X, DING W Q, JIN W B. Microwave-assisted extraction of lipids, carotenoids, and other compounds from marine resources[J].Innovative and emerging technologies in the bio-marine food sector,2022: 375-394.

[2] CHRISTENSEN L, FERNANDEZ J D G, HILDEBRANDT M, et al. Recent advances in AI for navigation and control of underwater robots[J]. Current robotics reports, 2022, 3: 165-175.

[3] XU X Y, ZHAO M, SHI P X, et al. Crack detection and comparison study based on Faster R-CNN and Mask R-CNN[J]. Sensors, 2022,22(3):1215.

[4] KUMAR A, ZHANG Z P, LYU H B. Object detection in real time based on improved single shot multi-box detector algorithm[J/OL]. EURASIP journal on wireless communications and networking, 2020[2024-01-12].https://

jwcn-urasipjournals.springeropen.com/articles/10.1186/s13638-020-01826-x.

[5] ZHANG M H, XU S B, SONG W, et al. Lightweight underwater object detection based on YOLOv4 and multi-scale attentional feature fusion[J]. Remote sensing,2021,13(22): 4706.

[6] CHEN X, YUAN M J H, YANG Q, et al. Underwater-YCC: underwater target detection optimization algorithm based on YOLOv7[J]. Journal of marine science and engineering, 2023,11(5): 995.

[7] LI Y, BAI X Y, XIA C L. An improved YOLOv5 based on triplet attention and prediction head optimization for marine organism detection on underwater mobile platforms[J].Journal of marine science and engineering, 2022, 10(9): 1230.

[8] LIU T X, LI J X, CAI M Y, et al. An improved YOLOv3-SPP algorithm for image-based pothole detection[C]//Advances in Neural Networks – ISNN 2024:18th International Symposium on Neural Networks.NewYork:ACM,2024:328-335.

[9] WANG C Y, BOCHKOVSKIY A, LIAO H Y M. YOLOv7: trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[C]//2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway:IEEE,2023:7464-7475.

[10] GOUIDER C, SEDDIK H. YOLOv4 enhancement with efficient channel recalibration approach in CSPdarknet53[C/OL]//2022 IEEE Information Technologies & Smart Industrial Systems (ITSIS). Piscataway: IEEE, 2022[2024-12-02]. https://ieeexplore.ieee.org/document/10118431.DOI:10.1109/ITSIS56166.2022.10118431.

[11] WANG A, CHEN H, LIN Z J, et al. RepVIT: revisiting mobile cnn from VIT perspective[C]//2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway:IEEE,2024:15909-15920.

[12] LI X, ZHANG T, ZHANG Z, et al. Survey of transformer research in computer vision[J].Computer engineering and applications, 2023, 59(1): 1-14.

【作者简介】

郑凯东（1964—），男，广东汕头人，硕士，副教授、硕士生导师，研究方向：图形学与虚拟现实、深度学习与计算机视觉、程序设计。

蒋瑞娜（2002—），女，河南周口人，硕士研究生，研究方向：计算机视觉、深度学习。

（收稿日期：2025-02-24 修回日期：2025-07-02）