

DAF-YOLO:改进 YOLOv8s 的轻量化水下目标检测算法

刘儒一¹, 刘自超¹, 孙媛媛², 宋廷强¹

(1. 青岛科技大学 信息科学技术学院, 山东 青岛 266061;

2. 青岛科技大学 数据科学学院, 山东 青岛 266061)

摘要: 在复杂海洋探测场景中, 针对现有目标检测算法因光线散射、悬浮物遮挡等环境干扰导致的检测精度低和误检漏检问题, 以及模型参数量大与水下勘探设备有限算力间的部署矛盾, 提出基于改进 YOLOv8s 的轻量级水下目标检测算法; 该算法对 YOLOv8s 的主干网络进行了改进, 设计了一种改进的 C2f 模块 D2F, 增强了模型对形变目标的表征能力; 采用渐进特征金字塔网络 AFPN 替代原颈部网络, 通过跨层直接交互与自适应空间融合操作, 促进了多尺度特征有效融合, 解决了非邻近层特征融合时的语义差距和信息冲突问题; 为提升模型检测速度, 引入了 FasterNet 中的 PConv 新型卷积方式, 构建 C2f-Faster 模块降低模型复杂度; 将融合 D2F、AFPN 与 C2f-Faster 模块后的模型命名为 DAF-YOLO, 经实验测试表明, 该模型在 URPC2020 数据集上的准确率提高了 3.9%, 参数量和计算量分别减少了 10.52% 和 10.91%, 同时以每秒 115.6 帧的速度实现了高速检测, 进一步实现了算法在精度和速度间的平衡。

关键词: 水下目标检测; YOLOv8s; D2F; AFPN; FasterNet

DAF-YOLO: A Lightweight Underwater Target Detection Algorithm Based on Improved YOLOv8s

LIU Ruyi¹, LIU Zichao¹, SUN Yuanyuan², SONGTingqiang¹

(1. College of Information Science and Technology, Qingdao University of Science and Technology,

Qingdao 266061, China;

2. College of Data Science, Qingdao University of Science and Technology, Qingdao 266061, China)

Abstract: In complex ocean detection scenarios, due to environmental interference such as light scattering and suspended object occlusion, existing target detection algorithms have the characteristics of low detection accuracy, false monitoring and missed detection, as well as deployment conflicts between the large number of model parameters and the limited computing power of underwater exploration equipment. To address this issue, a lightweight underwater target detection algorithm based on improved YOLOv8s is proposed, which improves the backbone network of YOLOv8s. An improved C2f module D2F is designed, which enhances the ability of the model to characterize the deformed target; The asymptotic feature pyramid network AFPN is replaced as the original neck network, which effectively fuses multiscale features through the cross-layer direct interaction and adaptive spatial fusion, achieving the semantic differences and information conflicts in the feature fusion of non-neighboring layers; To enhance the detection speed of the model, the PConv novel convolution method in the FasterNet is introduced, and the C2f-Faster module is constructed to reduce the complexity of the model. The model that integrates the D2F, AFPN and C2f-Faster modules is named after DAF-YOLO, and Experimental tests show that the accuracy of this model on the URPC2020 dataset is improved by 3.9%, the number of parameters and the computation amount are

收稿日期:2025-01-20; 修回日期:2025-02-26。

基金项目:国家自然科学基金青年项目(32301702)。

作者简介:刘儒一(1999-),男,硕士研究生。

孙媛媛(1990-),女,博士,副教授。

通讯作者:宋廷强(1971-),男,博士,副教授。

引用格式:刘儒一,刘自超,孙媛媛,等. DAF-YOLO:改进 YOLOv8s 的轻量化水下目标检测算法[J]. 计算机测量与控制, 2026, 34(2):31-38.

reduced by 10.52% and 10.91%, respectively, with a detection speed of 115.6 frames per second, which further achieves a balance between the algorithm's accuracy and speed.

Keywords: underwater target detection; YOLOv8s; D2F; AFPN; FasterNet

0 引言

随着时间的推移,水下探测技术的不断进步为海洋科学和资源管理等核心领域带来了显著的推动力,其重要性不言而喻。这项技术精确地获取物体的位置、形态、数量和行为等数据,为海洋研究和资源的持续利用提供了加速的动力。同时,水下探测在海洋安全、生态监测和文化遗产保护等方面扮演着至关重要的角色,为海洋生态的平衡和资源的长期开发提供了坚实的技术支持^[1]。

人工智能领域的快速发展使得深度学习技术迅速扩展到多个领域,特别是在海洋生物识别方面取得了重大突破。目前,基于深度学习的检测技术主要分为两大类:双阶段和单阶段检测算法。在双阶段检测算法中,RCNN^[2]及其衍生算法因其卓越的精确度而备受关注。Faster R-CNN^[3]在物体检测方面展现了卓越的识别能力,而 Mask R-CNN^[4]在实例分割任务中同样表现出色,两者均对该领域的发展做出了重要贡献。在单阶段检测算法领域,YOLO^[5]以其快速且高效的识别性能脱颖而出,成为业界的焦点,在实际应用中显示出其独特的优势。随着技术进步,单阶段算法不仅保持了速度上的优势,其准确度也在不断提升,这为海洋生物目标检测带来了新的发展机遇和可能性。

为了将基于深度学习的目标检测算法应用于水下场景中,叶赵兵等人^[6]在图像预处理时引用水下图像增强来减轻图像模糊带来的负面影响,为后续目标检测提供支持。

Lei 等人^[7]采用 Swin Transformer 模型作为 YOLOv5 的主网络,使该模型在水下环境中,能够有效提取目标有用特征,提高模型的泛化能力。Liang 等人^[8]在 YOLOv5 主干网络中基于 Ghost 卷积改进 C3 模块以及使用 Rep-ECA 模块进行结构重参数化,有效降低了模型的数量,提升了推理速度。辛世澳等人^[9]在 YOLOv7 的骨干网络 backbone 上,引入 ShuffleNetv2 轻量级网络和 BiFPN 特征聚合网络,来提高水下目标检测精度。

尽管这些研究在提升检测速度和精度方面有所进展,但未能兼顾模型的可移植性和实时性,无法同时满足速度和精度的要求。为此,提出改进 YOLOv8s^[10] 的轻量级目标检测算法。它提供了一种轻量级解决方案,既提高了检测速度,又提高了准确率,主要贡献如下:首先,提出了 D2F 特征提取模块,它结合可变形卷积、标准卷积和 EMA 多尺度注意力机制,增强了对复杂细

节的捕捉和边缘形变特征的恢复处理。其次,为防止特征丢失和退化,并解决水下目标漏检和误检问题,改进颈部融合网络,引入 AFPN 渐进式特征金字塔网络,用于多尺度特征融合,这种方法能有效地处理不同分辨率的图像,以较低的成本增强非相邻层特征的融合,并减轻因参数降低而造成的精度损失,提高模型检测效果。最后,融合 FasterNet 中的 PConv 卷积,形成 C2f-Faster 结构,替换 YOLOv8 颈部的 C2f 模块,从而有效减少模型参数数量和计算量,提升检测速度。

1 YOLOv8 目标检测算法

YOLOv8 模型由 4 部分组成:输入阶段、特征提取骨干、颈部网络和输出层。在输入阶段,应用了 Mosaic 数据增强技术,并根据模型的规模灵活调整超参数设置。对于较大模型,进一步引入了 MixUp 和 CopyPaste 等数据增强方法,以增加数据集的多样性。这些策略有助于提升模型的泛化能力和鲁棒性,使其在不同场景下都能保持较高的检测精度和稳定性。骨干网络使用 Conv、C2f 和 SPPF 模块,从图像中提取特征。C2f 结构基于 C3 模块和 YOLOv7^[11] 的 ELAN 的残差设计,在丰富梯度流的同时减小了体积,同时它还能自适应调整通道数,优化模型性能。此外,SPPF 则负责从不同尺度特征上捕获上下文信息,进行多尺度特征融合。颈部部分使用 FPN^[12]+PAN^[13] 双金字塔结构实现层与层之间的特征流动,增强不同尺度的语义表现和定位能力,从而实现更全面的特征融合。输出层则根据前面提取的特征信息,获取目标的类别和位置,来实现目标检测。该结构通过使用解耦头部设计,解决了分类与检测各自关注点不匹配的问题。

2 DAF-YOLO 水下目标检测算法原理

尽管 YOLOv8 在目标检测方面表现出色,但在水下探测应用中仍面临挑战。复杂的水下环境和干扰因素会导致目标漏检或误检现象。此外,为实现勘探识别任务的高精度检测需求,所需的模型大都参数多、体积大,部署困难,难以应用在水下勘探设备上。为应对这些问题,对原始 YOLOv8 检测模型进行了改进,将融合了 D2F、AFPIN 和 C2f-Faster 模块后的网络命名为 DAF-YOLO,其网络结构如图 1 所示。

2.1 设计 D2F 模块改进主干网络

2.1.1 可变形卷积替换标准卷积

水下环境充满独特的复杂性,在进行特征提取时,目标边界细节繁杂且易发生形变,单纯依赖标准卷积进

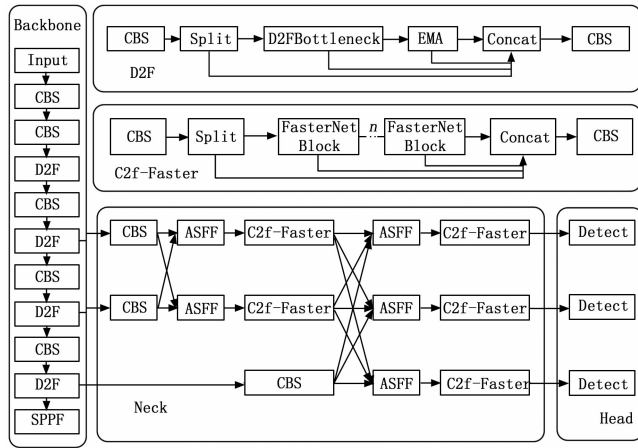


图 1 DAF-YOLO 网络结构

行特征提取往往难以应对, 可能导致一系列识别问题。首先, 标准卷积 (SC, standard convolution)^[14] 难以有效捕捉细粒度细节, 尤其在目标存在遮挡或重叠时, 其感受野调整能力有限, 导致细节丢失或模糊。为此, 在特征提取阶段引入了可变形卷积第二版 (DCNv2^[15]), 该卷积通过可学习的形变参数增强感受野调整能力, 更好地捕捉细节并适应形变, 弥补了标准卷积的不足。此优化提高了特征表示与目标定位的准确性, 增强了水下检测系统的性能和鲁棒性, 有助于在复杂场景中更准确地检测和识别水下目标。

可变形卷积网络 (DCN) 增强了传统的卷积能力, 通过引入可学习的偏移量, 使卷积核能够动态调整其采样位置, 从而更有效地提取特征。DCNv2 在此基础上进一步优化, 加入了偏移量和加权策略, 提升了模型对

目标的捕捉能力。其操作流程分为两步: 第一步, 生成偏移量, 卷积层计算出卷积核在特征图上 x 和 y 方向的采样点偏移; 第二步, 通过双线性插值确定采样位置并执行卷积操作。该过程使 DCNv2 能够更精准提取目标特征, 提升检测精度。

尽管 DCNv2 的应用对模型的参数数量和计算复杂度的影响相对较小, 不恰当的使用仍可能导致计算时间显著增加。为了优化效率和性能, 在模型改进方案中, 特别在骨干 C2f 模块的瓶颈部分用 DCNv2 可变形卷积取代了标准卷积, 这种策略性替换形成了 D2FBottleneck 模块。

2.1.2 EMA 模块嵌入 C2f 网络

在水下目标检测任务中, 光线折射导致目标边界模糊, 同时目标周身背景等干扰因素也增加了目标细节特征的提取难度。为应对这一问题, 将 EMA^[16] 注意力机制应用在模型中。注意力机制在图像特征处理中至关重要, 它能够捕捉局部关键信息, 使模型更专注于目标的相关特征, 从而提高检测的准确性。高效多尺度 EMA 机制通过将部分通道转换为批量维度来处理降维问题, 又能在减少计算负荷的同时保留独立信息, 提高了模型特征信息获取能力。EMA 结构如图 2 所示。

EMA 按通道数将输入特征 $X \in R^{C \times H \times W}$, 划分为 G 个子特征, 获取其中隐藏的潜在特征语义信息。特征 $X = [X_0, X_1, \dots, X_{G-1}]$, $X_i \in R^{C // G \times H \times W}$ 中, 假设 G 远远小于 C , 且在这一前提下, 通过学习得到的权重描述符将有效地增强每个子特征中目标区域的特征表达能力。

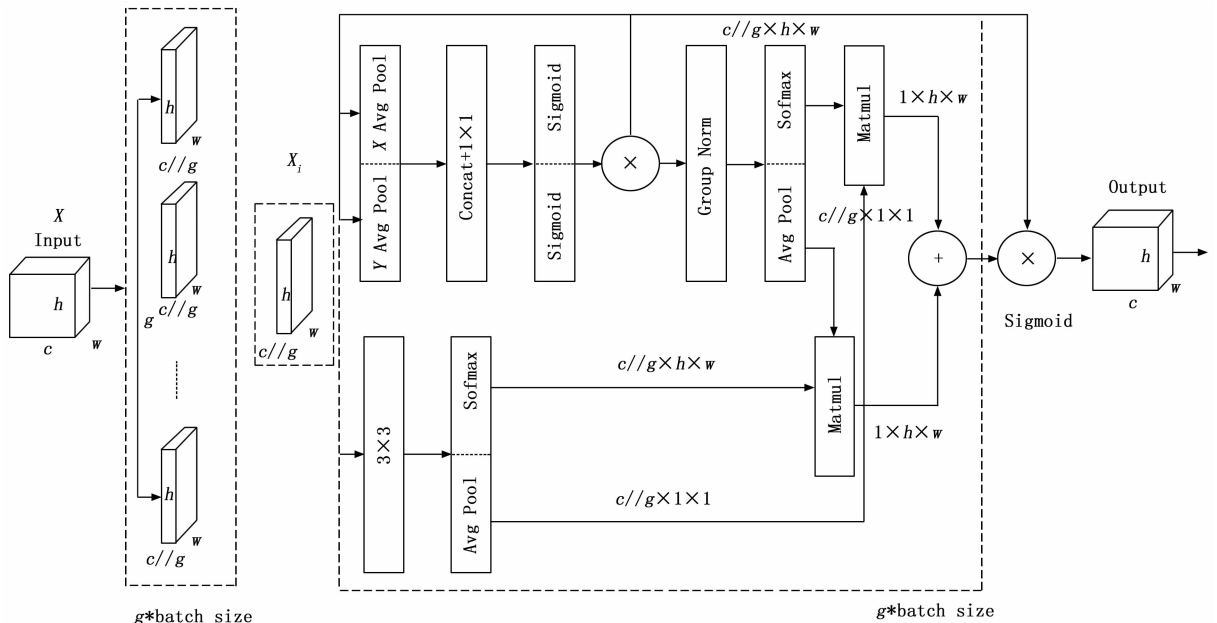


图 2 EMA 结构

EMA 采用双路径策略从分组特征图中提取关键权重描述符, 这两条路径分别是 1×1 卷积路径和 3×3 卷积路径。在 1×1 路径中, 通过两次一维全局平均池化操作, 能够在不减少特征维度的情况下, 有效地编码和整合跨空间维度的通道信息, 从而确保通道中关键信息的完整性得以保持。随后, 1×1 卷积的输出被分成两个向量。这些矢量会经过两个 Sigmoid 函数来拟合线性卷积上的 2D 二进制分布, 帮助建立数据分布模型, 最后这一过程通过将向量相乘来实现跨信道连接交互。另一方面, 在 3×3 路径中, 单个 3×3 卷积用于捕捉不同尺度的特征, 使用二维全局平均池化对两条路径的输出进行全局空间信息编码操作。然后将这些输出调整到适当的维度, 由 $R_1^{1 \times C // G} \times R_2^{C // \times HW}$ 表示。之后, 应用 softmax 函数来拟合数据的线性变换, 并将两条路径的输出合并并转换为 $R^{1 \times H \times W}$ 这种格式。接着, 通过矩阵点积将并行输出合并, 形成收集多尺度的信息空间注意力图。EMA 的最终结果与输入 (X) 大小相同, 因此可以轻松集成到 YOLOv8 网络中。为了充分发挥 EMA 的优势, 将其嵌入到骨干网络的 C2f 模块中。

2.1.3 融合 DCNv2、SC 和 EMA

将融合 DCNv2、SC、与 EMA 协同架构命名为 D2F 模块, 其通过“基础特征约束—动态形变适配—判别性特征筛选”的协同架构实现了精密特征提取。首先通过 1×1 标准卷积 (SC) 构建稳定的基础特征表达, 避免形变建模过程中的特征空间失稳; 随后串联 DCNv2 动态卷积, 通过可行变采样点自适应捕捉目标边界形变; 二者输出后经过批量归一化 (BN, batch normalization)^[17] 和 SiLU 激活函数 (SiLU, sigmoid-weighted linear unit)^[18] 处理后通过跳跃连接融合, 形成兼具结构鲁棒性与形变适应性的 D2FBottleneck 模块。为进一步提升特征提取能力, 在 Bottleneck 末端嵌入 EMA 注意力机制, 通过双路径空间通道交互能力抑制水下光线折射和背景信息的干扰, 强化目标关键区域的关注能力。通过三者的协同工作优化架构, 实现了水下目标形变、模糊以及多尺度干扰的联合建模, 显著提升了场景下的检测鲁棒性。

通过把 D2F 模块融合至骨干网络中, 可使模型对目标特征更为敏感, 能自适应地选取目标关键特征并予以增强。这种融合方式在不增加模型复杂度的情况下, 提升了目标特征的提取能力, 进而提高了模型的目标检测性能。

2.2 改进颈部特征融合网络

2.2.1 基于 AFPN 的特征融合网络

在优化主干网络的特征提取能力后, 促进颈部网络的多尺度特征有效融合成为制约模型检测精度提高的关

键因素。YOLOv8 的特征融合网络由 FPN+PAN 的特征金字塔结构组成, 其通过经典的自上而下和自下而上的特征金字塔网络进行多尺度特征融合, 但这种方式在一定程度上会造成特征信息丢失以及多尺度目标特征信息冲突问题, 削弱非相邻层级的特征融合效果, 进而降低检测精度。鉴于此, 采用渐进特征金字塔网络 (AFPN^[19]) 让非相邻层特征直接交互, 并利用自适应空间融合操作 (ASFF^[20]) 缓解多目标信息冲突, 实现各层级不同尺度特征的充分融合, 减少漏检和误检情况, 提高模型检测精度。AFPN 渐进特征金字塔结构如图 3 所示, 其中平行箭头和交叉箭头, 分别表示卷积步骤和上下采样步骤。

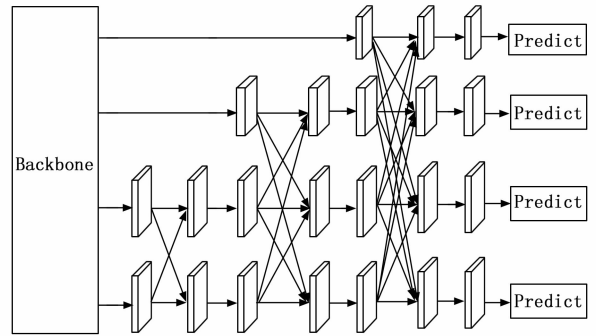


图 3 渐进特征金字塔网络

Backbone 网络在自下而上提取特征时, 先组合两个不同尺度的浅层特征开启特征融合, 再逐步把深层特征加入融合过程, 最后融合 Backbone 的顶级特征。这种渐进融合方式使浅层、深层和顶级特征相互融合, 让非相邻层级特征直接交互, 避免了多级特征传输中的信息丢失与退化, 消除了非相邻层级特征融合产生的语义差距, 实现不同尺度特征的充分融合。在特征融合过程中, 不同层级特征的空间位置可能出现多目标信息冲突, 元素求和融合效果不佳的问题。因此, AFPN 采用自适应空间融合操作, 为不同层级特征分配不同空间权重, 增强关键层级特征的重要性, 充分融合不同尺度特征, 提高模型检测精度。自适应空间融合操作结构如图 4 所示。

ASFF 的核心思想就是在逐元素相加特征融合方式的基础上, 添加一个可学习的空间权重, 以实现自适应高效特征融合。假设 $P_{mn}^{x \rightarrow l}$ 表示特征图 x 到特征图 l 在位置 (m, n) 处的特征融合分量, Y_{mn}^l 表示多级特征经自适应空间融合操作后的输出特征图, 由特征分量 $P_{mn}^{1 \rightarrow l}$ 、 $P_{mn}^{2 \rightarrow l}$ 和 $P_{mn}^{3 \rightarrow l}$ 的组合定义公式如下所示:

$$Y_{mn}^l = \alpha_{mn}^l P_{mn}^{1 \rightarrow l} + \beta_{mn}^l P_{mn}^{2 \rightarrow l} + \gamma_{mn}^l P_{mn}^{3 \rightarrow l} \quad (1)$$

$$\alpha_{mn}^l + \beta_{mn}^l + \gamma_{mn}^l = 1 \quad (2)$$

其中: α_{mn}^l 、 β_{mn}^l 和 γ_{mn}^l 分别为各个特征分量在融合输

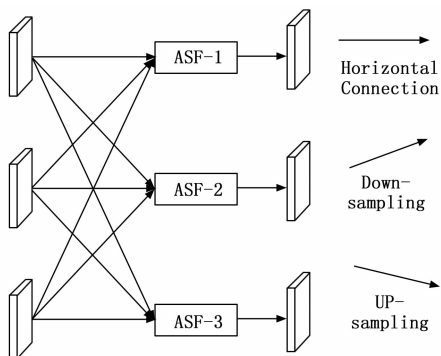


图 4 自适应空间融合操作

出特征图 Y_{mm}^i 处的可学习空间权重, 且 $\alpha_{mm}^i, \beta_{mm}^i, \gamma_{mm}^i \in [0, 1]$ 。

2.2.2 融合新型卷积 PConv

引入 AFPN 思想对特征融合网络进行重构后, 实现了不同尺度特征的充分融合, 丰富了特征信息且提高了模型检测精度, 但代价是计算复杂度、内存开销增大, 权重系数增多, 导致模型计算推理速度降低, 尤其在内存受限的水下移动设备中不利于水下目标实时检测。

为在保证模型精度的同时降低参数量和计算量、提高检测速度, 在特征融合网络中引入了 FasterNet^[21] 中的新型卷积 PConv 来解决该问题。FasterNet 提出的新卷积方式 PConv 能够降低计算冗余和内存访问, 有效利用计算资源提取空间特征, 其结构如图 5 所示。

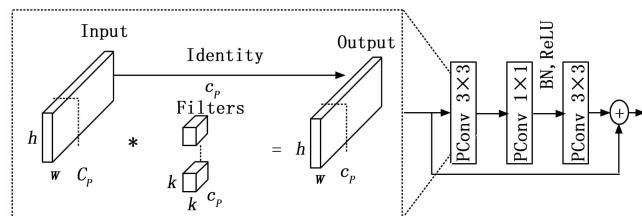


图 5 PConv 和 FasterNet Block 结构

从图 5 中可知 PConv 的工作原理。它只对部分输入和输出通道应用常规卷积 (Conv), 而其余通道保持不变。为了进行连续有序的处理, 会使用第一组或最后一组连续通道来表示整个特征图。这样可以确保输入和输出特征图中的通道数量保持不变, 而不会丢失任何重要信息。PConv 的计算如公式 (3) 所示:

$$h \times w \times k^2 \times c_p^2 \quad (3)$$

对于典型的部分比率 $r = c_p/c = 1/4$, PConv 的计算量仅为常规 Conv 的 1/16 倍。

此外, PConv 的内存访问量较小, 计算量如公式 (4) 所示:

$$h \times w \times 2c_p + k^2 \times c_p^2 \approx h \times w \times 2c \quad (4)$$

其中: h 为通道的高, w 为通道的宽, c_p 为连续的

网络通道, k 为滤波器。

为了解决模型改进后因参数量和计算量较大而导致检测速度变慢的问题, 基于 PConv 和 Bottleneck 设计了 FasterNet Block 模块, 并用其替换 C2f 结构中的 Bottleneck 模块, 形成 C2f-Faster 模块。然后, 将特征融合网络中的所有 C2f 模块替换为 C2f-Faster。

3 实验设置与结果分析

3.1 实验环境

实验环境搭建依托于 PyTorch 1.13.0 深度学习框架, 用于训练和测试水下检测模型。开发环境是 PyCharm 2022.1.2 和 Python 3.8, CPU Intel (R) Core (TM) i9-12700KF@3.61 GHz, GPU 为 NVIDIA GeForce RTX 3090, 24 GB 显存, 运行内存 64 GB。在模型训练中, 输入图像分辨率设为 640×640 像素, 使用 SGD 优化器, 并设置随机种子为 1 以确保结果可复现。数据增强系数为 1, 动量参数为 0.937, 批量大小为 32, 迭代次数为 150, 初始学习率为 0.001, 衰减系数为 0.05, 以提高泛化能力、加速收敛、平衡效率与效果, 并实现平稳的优化过程。

在本实验中, 采用了 URPC 的数据集, 该数据集来自全国水下机器人大赛。其包括海参、扇贝、海胆和海星四类海洋生物的水下图像。虽然该数据集涵盖了典型目标, 但没有涵盖珊瑚、鱼类等生物以及缺乏海藻等背景干扰, 存在类别单一和背景复杂度不足的限制性, 会影响模型在复杂水域的泛化能力。为提升模型鲁棒性和泛化能力, 采用 Mosaic 数据增强方式, 随机选取 4 张图像进行裁剪缩放, 拼接为单一样本, 进一步扩充数据集。随后将数据集以 7:3 的比例随机分成训练集和测试集, 以便于模型的准确训练和评估。

3.2 评价指标

为了清晰有效地展示对 YOLOv8 所做的改进表现, 重点介绍了该模型在两个关键方面的优势: 检测精度和检测速度。为了评估检测精度, 使用了标准的目标检测指标, 包括精确度 (P)、召回率 (R)、平均精度 (AP) 和平均精度均值 (mAP)。这些指标的计算公式如下:

$$P = \frac{TP}{TP + FP} \quad (5)$$

$$R = \frac{TP}{TP + FN} \quad (6)$$

TP (真阳性) 是指模型正确预测的阳性样本数量; FP (假阳性) 是指被错误预测为阳性的阴性样本数; FN (假阴性) 是指被错误预测为阴性的阳性样本数。一个类别的召回率用 “ r ” 表示, “ N ” 代表类别总数。 mAP 衡量所有类别的平均 AP , 数值越高, 表示识别准确率越高。 AP 是精确度-召回 (PR) 曲线下的面积, 也表示模型的识别精确度, AP 越大, 精确度越高。

AP 和 mAP 的计算详见公式 (7) 和 (8):

$$AP = \int_0^1 P(r) dr \quad (7)$$

$$mAP = \frac{1}{N} \sum_{i=1}^N AP_i \quad (8)$$

为了评估检测速度,我们使用了 3 个关键指标:每秒帧数(FPS)、浮点运算次数(GFLOPs)和模型参数(Params)。FPS 反映模型单位时间内的画面处理能力,数值越高代表实时流畅性越好;GFLOPs 量化模型的计算复杂度,体现单次推理所需的算力;Params 表征模型的总参数量,三者共同构成评估模型的重要指标。

3.3 消融实验

为了验证 DAF-YOLO 网络各结构的有效性,以 YOLOv8s 作为基线网络,对 D2F、AFPN 和 C2f-Faster 进行消融实验,并在 URPC2020 数据集上进行验证。为确保公平比较,对环境参数进行了标准化处理,实验结果见表 1。“√”表示,使用了该方法。

表 1 消融实验结果对比

Model	D2F	AFPN	C2f-Faster	FPS	Params /M	FLOPs /G	mAP /%
YOLOv8s				82	11.12	28.4	82.2
YOLOv8s	√			60.6	11.16	28.7	83.9
YOLOv8s	√	√		82.5	11.23	28.3	85.3
YOLOv8s	√	√	√	115.6	9.95	25.3	86.1

从表 1 能够看出,每一项改进都在不同程度上使网络的检测性能得到了提高,成功平衡了模型的精度与轻量化,为水下终端设备上的部署提供了一种实用且可行的解决方案。在本实验中,我们把基线模型 YOLOv8s 的实验结果作为后面 3 组实验的对比基线,该基线模型的检测精度为 82.2%,FPS 为 82,参数量为 11.12 M,GFLOPs 为 28.4 GFLOPs。当把 D2F 模块融合到骨干网络中并替换掉 C2f 层时,模型参数和 GFLOPs 分别仅增加了 0.04 M 和 0.3 GFLOPs,而 mAP 却提升了 1.7%,这说明 D2F 模块的引入在基本不增加参数量和计算量的情况下,增强了模型的特征表达能力,让模型能更好地提取特征信息,从而提升了检测精度。

在此基础上,我们引入 AFPN 渐进特征金字塔网络对模型的特征融合层进行了优化,尽管模型参数增加了 0.07 M,但 GFLOPs 减少了 0.4 GFLOPs,检测精度提升了 1.4%。这表明,AFPN 渐进特征金字塔网络能够获取不同尺度特征间的潜在语义联系,促使它们逐步融合,有效解决了非邻近层特征信息冲突问题,减少了错误识别的可能性。本文的改进方案结合了 D2F、AFPN 和 C2f-Faster 模块,在显著减少模型参数量和计算量的同时,检测精度提升至 86.1%,并将帧率(FPS)提高到 115.6,DAF-YOLO 相较于原始模型表

现出更高的效率和速度。尽管 C2f-Faster 网络的引入可能导致精度下降,但通过 AFPN 和 D2F 模块的改进,有效恢复并进一步提升了精度。

3.4 与其他目标检测算法对比实验

在复杂的水下环境中,DAF-YOLO 的检测精度和速度都非常出色。为了进一步评估其性能,我们将其与几种主流水下目标检测算法进行了比较,包括 Faster-RCNN、YOLOv5m^[22]、YOLOv7-tiny、YOLOv8s、YOLOv9m^[23]、YOLOv10m^[24] 和 EnYOLO^[25]。比较结果如表 2 所示,Faster-RCNN 在检测准确性方面存在局限性,由于其候选帧的大小有限,在识别不同比例和长宽比的物体时准确性较差,这可能会导致漏检现象,增加错误识别概率。YOLOv7-tiny 与 DAF-YOLO 相比速度更快,但在检测精度方面,DAF-YOLO 相较于其 mAP 提升了 7%。此外,DAF-YOLO 在检测精度和速度方面都优于 YOLOv5m 和 YOLOv9m,mAP 相较于它们分别提升 7.6% 和 3.7%。虽然 DAF-YOLO 的参数更多,需要的计算量也更大,但其 115.6 FPS 的检测速度仍能满足实时检测的要求,其准确性也明显超过了上述模型,在速度和准确性之间取得了良好的平衡。此外,改进后的算法在检测速度和精度上均优于原始 YOLOv8s,检测速度提高了 40.97%,mAP 提升了 3.9%,参数量减少了 10.52%,GFLOPs 减少了 10.91%。尽管 YOLOv10m 和 EnYOLO 算法在水下目标检测中具有较高精度,但其检测速度较慢,实时性差,不适合水下设备部署。相比之下,本文提出的算法 DAF-YOLO 在检测速度和准确性上都明显优于前述算法,展现出其在水下目标检测中的优势和优越性能。

表 2 不同目标检测算法结果

Model	FPS	Params/M	FLOPs/G	mAP/%
Faster-RCNN	36.52	136.77	368.25	57.9
YOLOv5m	78	21.2	49.0	78.5
YOLOv7-tiny	120	6.01	13.25	79.1
YOLOv8s	82	11.12	28.4	82.2
YOLOv9m	72	20.0	76.3	82.4
YOLOv10m	79	15.4	59.1	83.5
EnYOLO	97.2	20.88	20.0	82.6
DAF-YOLO	115.6	9.95	25.3	86.1

为了更清晰地评估 DAF-YOLO 的检测效果,我们在测试集中选取了若干代表性图像进行实验,并与当前主流的目标检测算法进行了对比分析,实验结果如图 6 所示。从检测效果来看,在图像模糊和目标聚集的情况下,YOLOv10m 和 EnYOLO 的表现优于 YOLOv7-tiny 算法,但仍然存在误检和漏检的问题。例如,一些目标由于特征不明显且相互遮挡,被错误识别为背景。然

而, 在此类场景中, 本文提出的算法 DAF-YOLO 表现出更好的目标识别能力, 尤其是在检测被遮挡物体时具有更强的优势。与 YOLOv10m 和 EnYOLO 相比, DAF-YOLO 不仅检测精度更高, 而且误检和漏检现象有所减少, 能够有效识别相互遮挡的目标, 提升了模型的检测能力。

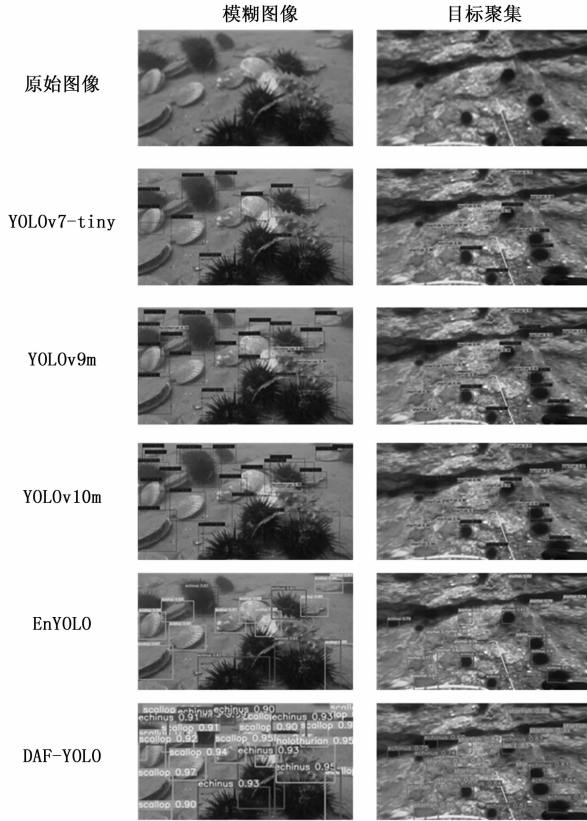


图 6 不同算法检测结果

DAF-YOLO 在检测精度上超越了其他主流算法, 同时在水下目标检测中颇具竞争, 并能处理水下场景中的图像模糊、背景复杂和目标聚集等问题。它的检测精度和推理速度表现出色, 此外, 它还加强了模型对特征提取能力, 改善了模型特征提取困难和目标漏检的情况, 解决了水下场景中常见的难题。

3.5 稳定性评估实验

水下环境复杂性对目标检测算法的稳定性提出了严峻挑战。为验证 DAF-YOLO 模型在不同极端条件下的鲁棒性, 本文基于 URPC 测试数据集模拟了低光、强光、浑浊水质以及目标遮挡场景, 并进行了实验测试, 实验设计如下。通过调整图像亮度, 生成低光与强光数据子集, 模拟光照分布不均的场景; 通过添加高斯噪声与均值模糊模拟浑浊水质条件; 通过随机遮挡目标区域, 来模拟目标遮挡场景。表 3 展示了 DAF-YOLO 在不同水下场景中的性能表现。

表 3 DAF-YOLO 在不同水下环境中的性能表现

Model	$mAP/\%$	FPS
原始数据	86.1	115.6
低光	84.3	112.4
强光	83.9	113.0
浑浊水质	83.5	110.2
目标遮挡	83.2	109.8

实验结果表明, DAF-YOLO 在极端水下环境中展现出较强的稳定性与适用性。在光照变化条件下, 低光与强光场景的 mAP 仅下降 1.8% 和 2.2%, 同时 FPS 波动小于 3 帧, 得益于多尺度注意力机制对跨通道信息的有效整合, 显著缓解了光照不均引起的特征丢失问题。面对浑浊水质干扰, 模型通过可变形卷积动态调整感受野, 抑制模糊效应的影响, mAP 降幅控制在 2.6%, 使得性能波动处于合理范围。在目标遮挡场景中, 尽管部分目标被遮挡, 仍能利用局部可见特征实现目标定位, 使得 mAP 保持在 83.2%。整体而言, DAF-YOLO 在复杂环境下的 mAP 波动小于 3%, 检测速度稳定于 109 FPS 以上, 验证了其在水下复杂场景中的鲁棒性。

4 结束语

目前的水下目标检测算法优先考虑精度, 需要使用大量复杂参数, 这样会导致模型实时检测性能大幅度下降, 不适合水下移动设备和嵌入式系统。因此, 本文介绍了一种轻量级水下目标检测算法 DAF-YOLO, 旨在优化精度和速度, 使其适用于此类应用场景。首先, 通过重构骨干网络, 设计了 D2F 特征提取模块, 将可变形卷积、标准卷积和 EMA 多尺度注意力机制结合, 并应用于主干网络中, 增强特征表征能力, 捕捉复杂细节并处理形变。其次, 采用 AFPN 改进特征融合网络, 促进非邻近层特征的有效融合, 减少语义冲突和信息丢失, 丰富特征信息, 提升对水下目标的检测能力。最后, 借助 FasterNet 中的轻量级卷积 PConv 构建 C2f-Faster 模块, 优化颈部网络, 降低模型参数量和计算量, 实现轻量化。实验结果表明, 本文算法 DAF-YOLO 的 mAP 达到 86.1%, 处理速度为 115.6 FPS, 兼具高精度和高速性能, 在复杂水下环境下优于其他现有模型。总体而言, 本文算法当前检测效率和处理速度已经达到较高水平。尽管如此, 网络结构的优化对推进水下目标探测方面仍具有巨大潜力。在水下目标探测方面, 实时处理和算法轻量化是未来的研究方向, 我们将持续探索这些领域的应用潜力。

参考文献:

- [1] 李加林, 沈满洪, 马仁锋, 等. 海洋生态文明建设背景下的海洋资源经济与海洋战略 [J]. 自然资源学报, 2022,

- 37 (4): 829–949.
- [2] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation [C] //Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2014: 580–587.
- [3] REN S, HE K, GIRSHICK R, et al. FasterR-CNN: towards real-time object detection with region proposal networks [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2016, 39 (6): 1137–1149.
- [4] HE K, GKIOXARI G, DOLLÁR P, et al. 2017 IEEE international conference on computer vision (ICCV) [J]. *IEEE*, Venice, Italy, 2017: 2980–2988.
- [5] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 779–788.
- [6] 叶赵兵, 段先华, 赵楚. 改进 YOLOv3-SPP 水下目标检测研究 [J]. *计算机工程与应用*, 2023, 59 (6): 231–240.
- [7] LEI F, TANG F, LI S. Underwater target detection algorithm based on improved YOLOv5 [J]. *Journal of Marine Science and Engineering*, 2022, 10 (3): 310.
- [8] LIANG H, SONG T. Lightweight marine biological target detection algorithm based on YOLOv5 [J]. *Frontiers in Marine Science*, 2023, 10: 1219155.
- [9] 辛世澳, 葛海波, 袁昊, 等. 改进 YOLOv7 的轻量化水下目标检测算法 [J]. *Journal of Computer Engineering & Applications*, 2024, 60 (3): 88–99.
- [10] LOU H, DUAN X, GUO J, et al. DC-YOLOv8: small-size object detection algorithm based on camera sensor [J]. *Electronics*, 2023, 12 (10): 2323.
- [11] 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] //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023: 7464–7475.
- [12] LIN T Y, DOLLÁR P, GIRSHICK R, et al. Feature pyramid networks for object detection [C] //Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017: 2117–2125.
- [13] LIU S, QI L, QIN H, et al. Path aggregation network for instance segmentation [C] //Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018: 8759–8768.
- [14] HU J, LIU B, PENG S. Forecasting salinity time series using RF and ELM approaches coupled with decomposition techniques [J]. *Stochastic Environmental Research and Risk Assessment*, 2019, 33: 1117–1135.
- [15] ZHU X, HU H, LIN S, et al. Deformable convnets v2: More deformable, better results [C] //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019: 9308–9316.
- [16] OUYANG D, HE S, ZHANG G, et al. Efficient multi-scale attention module with cross-spatial learning [C] // 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). *IEEE*, 2023: 1–5.
- [17] IOFFE S, SZEGEDY C. Batchnormalization: Accelerating deep network training by reducing internal covariate shift [C] //International conference on Machine Learning. *PMLR*, 2015: 448–456.
- [18] ELFWING S, UCHIBE E, DOYA K. Sigmoid-weighted linear units for neural network function approximation in reinforcement learning [J]. *Neural Networks*, 2018, 107: 3–11.
- [19] YANG G, LEI J, ZHU Z, et al. AFPN: Asymptotic feature pyramid network for object detection [C] //2023 IEEE International Conference on Systems, Man, and Cybernetics (SMC). *IEEE*, 2023: 2184–2189.
- [20] LIU S, HUANG D, WANG Y. Learning spatial fusion for single-shot object detection [J]. *Arxiv Preprint Arxiv*: 1911.09516, 2019.
- [21] CHEN J, KAO S, HE H, et al. Run, don'twalk: chasing higher FLOPS for faster neural networks [C] //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023: 12021–12031.
- [22] CAO S, WANG T, LI T, et al. UAV small target detection algorithm based on an improved YOLOv5s model [J]. *Journal of Visual Communication and Image Representation*, 2023, 97: 103936.
- [23] WANG C Y, YEH I H, MARK L H Y. Yolov9: learning what you want to learn using programmable gradient information [C] //European Conference on Computer Vision. Cham: Springer Nature Switzerland, 2024: 1–21.
- [24] WANG A, CHEN H, LIU L, et al. Yolov10: Real-time end-to-end object detection [J]. *Advances in Neural Information Processing Systems*, 2025, 37: 107984–108011.
- [25] WEN J, CUI J, ZHAO B, et al. EnYOLO: a real-time framework for domain-adaptive underwater object detection with image enhancement [C] //2024 IEEE International Conference on Robotics and Automation (ICRA), *IEEE*, 2024: 12613–12619.