

DOI:10.19651/j.cnki.emt.2416588

基于轻量级改进的 YOLOv8 水下目标检测模型^{*}

周志耀 马常霞 杨丽莎 仲兆满 胡文彬

(江苏海洋大学计算机工程学院 连云港 222006)

摘要: 在恶劣和多变的水下环境中工作的设备是进行水下研究和开发的基本保障。现阶段的水下目标检测模型参数量和计算量过大,在资源有限的水下设备上部署受限。为解决水下检测模型参数量和计算量过大问题,提出一种轻量级的水下目标检测模型 RCE-YOLO。首先,利用 RFACnv 的空间注意力权重来改进 CBS 处理接受域信息的能力和提升 C2f 对空间特征信息融合的能力,增强模型对小密集目标的检出能力。其次,融合 CCFM 与 Dysample 模块,该融合模块能够更有效的利用不同尺度信息并通过内部的点采样方法减少原先采样产生的模糊和失真。最后,在 SPPF 前向传播过程中融合高效多尺度注意力机制,该机制使得模型重点关注水下目标关键信息,降低误检率和错检率。实验结果表明,改进的轻量级模型在数据集 DUO 上进行验证,mAP50、mAP50:90 值分别达到 83.6%、64.2%,相较于 YOLOv8 基准模型 mAP50、mAP50:90 值分别提升了 1.4%、1.2%,参数量和计算量分别下降了 32.3%、0.9 G。相较于其他目标检测模型满足了恶劣多变环境下的水下目标检测需求,为水下设备轻量级部署奠定基础。

关键词: 水下目标检测;轻量化;RFA 卷积模块;CCFM 模块;注意力机制

中图分类号: TP391.41;TN919.8 **文献标识码:** A **国家标准学科分类代码:** 520.6040

Lightweight improved YOLOv8 based underwater target detection model

Zhou Zhiyao Ma Changxia Yang Lisha Zhong Zhaoman Hu Wenbin

(College of Computer Engineering,Jiangsu Ocean University,Lianyungang 222006, China)

Abstract: Equipment that operates in harsh and variable underwater environments is essential for conducting underwater research and development. The current underwater target detection models are too large in parameter count and computation, which limits the deployment of underwater equipment with limited resources. In order to solve the problem of excessive parameter count and computational volume of underwater detection models, a lightweight underwater target detection model RCE-YOLO is proposed. Firstly, the spatial attention weights of RFACnv are utilized to improve the ability of CBS to process the information in the receptive domain and to enhance the ability of C2f to fuse spatial feature information, so as to enhance the model's ability of detecting small and dense targets. Second, the CCFM is fused with the Dysample module, which is able to utilize the different scale information more effectively and reduce the blurring and distortion produced by the original sampling through the internal point sampling method. Finally, the Efficient multi-scale attention mechanism is fused in the SPPF forward propagation process, which makes the model focus on the key information of the underwater target and reduces the false detection rate and misdetection rate. The experimental results show that the improved lightweight model is validated on the dataset DUO, and the mAP50 and mAP50:90 values reach 83.6% and 64.2%, respectively, which are 1.4% and 1.2% higher compared to the mAP50 and mAP50:90 values of the benchmark model of YOLOv8, and the number of parameters and the amount of computation drop by 32.3% and 0.9 G, respectively. compared to other The target detection model meets the needs of underwater target detection in harsh and variable environments, and lays the foundation for lightweight deployment of underwater equipment.

Keywords: underwater target detection;lightweighting;RFACnv module;CCFM module;attention mechanism

0 引 言

随着计算机视觉中目标检测的不断发展,水下设备在

海洋领域具有广泛的经济前景^[1]。取得巨大收益的同时伴随着严峻的挑战:首先,特定养殖区域检测数值的略微波动,容易对海洋经济生物的日常活动产生剧烈影响。其次,

收稿日期:2024-08-01

^{*} 基金项目:国家自然科学基金(72174079)项目资助

当海水温度超过 33 ℃ 海参便会大面积死亡,造成难以估量的经济损失。最后,受到潮汐、风浪等因素影响,导致工作人员水下作业时间受限。这意味着水下检测模型要做到轻量且高效。目前,基于深度学习的水下目标检测模型取得巨大的成功。水下目标检测模型分成双阶段模型和单阶段模型。双阶段检测模型如:Faster-Rcnn^[2]等存在计算复杂度高、计算量大等问题,不适合当前复杂多样的水下环境。单阶段检测模型如:YOLO^[3]系列,则直接使用固定的锚框(anchor)进行识别定位更加轻量。而 YOLO 系列能保证水下目标检测精度成为了大多数学者研究的对象一些学者在水下目标检测研究中取得一些成就,但在一定欠缺。为解决网络重点特征信息提取效果差问题,周新等^[4]在 C3 模块引入交叉注意力机制改善小目标检测的精度,但带来了一定的参数量。受到 ATFHead 探测头启发,Lyu 等^[5]将其引入到 YOLOX 的检测层,加强多尺度特征的提取的同时也带来了巨大的计算量。Yu 等^[6]提出了新的 BoT-SORT 算法改良 YOLOv7 检测头提高了一定的检测精度,但增加了模型的内存访问次数。针对水下小目标特征提取不足问题,Zhang 等^[7]在 YOLOv8 检测层添加了小目标检测头,小目标错检率下降的同时也额外增加了一定的复杂度。施克权等^[8]将 Vision Transformer (ViT) 与 YOLOv8 主干网络相结合提高了检测精度,但模型复杂度有着进一步降低空间。为改善 YOLOv8 空间维度信息保留不足问题,Qin 等^[9]在 SPPF 前添加 GAM 注意力机制模块。提高了水下中小目标的检测精度,但忽略了该机制空间计算带来的计算量。Zhou 等^[10]设计 CSMB 模块引入到 YOLOv8 的 Neck 中,增加了模型的鲁棒性,但其未考虑复杂的网络结构带来很大的计算量。

尽管上述研究取得一定的成就,但面对复杂多变的水下目标检测任务时,无法兼顾轻量且高效。首先,现有的方法在面对昏暗和模糊环境时,其漏检率和错检率需进一步降低。其次,许多现有的改进方案都是基于高计算复杂度和大参数量的主干网络,检测头等,模型的计算复杂度上升。因此,针对 YOLOv8 基准水下检测模型参数量和计算量过大与特征提取能力有限问题,本文提出了一种轻量级 RCE-YOLO 模型。与现有的改进方案相比,本文提出的模型保证精度的同时不会带来太大的参数量和计算负担。本文的主要贡献包含 3 个方面:首先,针对 CBS 特征提取能力下降和 C2f 特征融合效果欠佳,使用带有感受场注意力(receptive field attention, RFA)卷积^[11]模块替换主干网络部分 CBS 和改进 C2f。RFA 卷积模块中通过空间注意力矩阵调整卷积核权重,以便更好关注重点信息,以提高水下物体检测精度。其次,设计带有高效多尺度注意力(efficient multi scale attention, EMA)的 CE 模块,增加水下关键信息保留,提高检测精度。CE 模块在模型前向传播过程中,将高效多尺度注意力机制融合到卷积操作之后。很好地重塑维度,以便重点关注小密集的水下生物。最后,

传统的 PAN+FPN 结构、Unsample 上采样造成了计算量、参数量过大和检测精度下降问题。将跨尺度融合模块(cross scale feature fusion, CCFM)与 Dysample^[12]上采样器相结合。将 CCFM-D 引入到 YOLOv8 的 Neck 层中,增加模型的表现能力的同时实现轻量化。

1 YOLOv8 与 RCE-YOLO 模型

YOLOv8 于 2023 年提出,具有较高的精度和稳定性。YOLOv8 的 Backbone 中全新的 C2f 模块延续了 CSP 思想,加深了梯度信息的融合程度能学习到更多的特征信息并实现更低的计算复杂度。YOLOv8 在 Head 部分增加最核心的改动解耦头(decoupled head),将原先的检测头分解成两个部分,这样更关注于各自的任务,提高检测的准确度。

虽然,YOLOv8 能够实现较高的精确度。但在水下世界仍存在若干挑战:首先,光线受到水吸收、反射和散射的影响。水下图像大多昏暗模糊,影响检测效果;其次,模型内部使用的 CBS、C2f 和 SPPF 模块,这些模块在小密集、昏暗模糊环境易出现漏检和错检情况;最后,YOLOv8 的 PAN+FPN 结构随着模块的叠加,易造成参数量和计算量过大问题,这样不利于水下设备的部署。因此,提出了一种轻量级 RCE-YOLO 模型,模型总体结构图如图 1 所示。

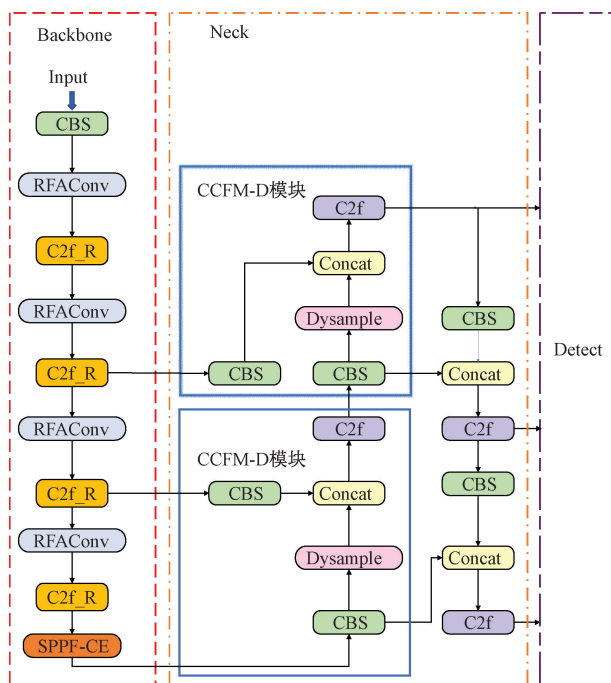


图 1 RCE-YOLO 模型结构

Fig. 1 Structure of RCE-YOLO

2 基于 YOLOv8 的轻量级改进方法

2.1 RFAconv 改进主干网络

Liu 等^[13]引入 EMA 注意力机制只对 C2f 模块的最终

输出进行不同空间尺度的特征加权操作,提升了C2f模块对额外的信息捕捉,但其忽略了前后模块的关联性。由于水下生物存在小尺度、多密集、多伪装等特性,传统的CBS模块会导致小密集目标信息丢失,同时上一层的特征信息映射到BottleNeck中又会加剧信息丢失,以降低整体改进C2f的空间信息整合能力。为了解决上述问题,本文引入RFAConv对主干网络中的CBS、C2f模块进行改进。RFAConv如图2所示。

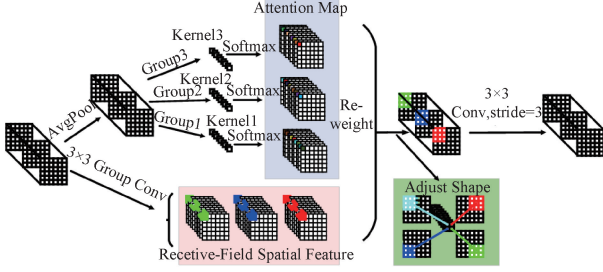


图2 RFA卷积结构

Fig. 2 Structure of RFAConv

RFAConv通过以下4个步骤,实现对水下目标的精准检测。首先,为了增加输入数据全局信息的捕捉和独立特征的学习。依次使用AvgPool操作和 1×1 组卷积操作。其中, X 代表初始输入信息。AvgPool代表平均池化操作,可以有效地聚合全局信息。GroupConv代表组卷积操作,组卷积的卷积核为1可以学习单组中独有特征,如式(1)所示。

$$F_1 = \text{GroupConv}_{1 \times 1}(\text{AvgPool}(X)) \quad (1)$$

其次,对特征 F_1 进行Softmax操作,以生成包含不同组权重的空间注意力矩阵 A_{rf} 。其中,Softmax使得组卷积后的有效特征分布更加均匀。空间注意力矩阵内部的注意力权重提升了卷积核处理接受域信息的能力。 F_1 为式(1)操作后输出的特征,如式(2)所示。

$$A_{rf} = \text{Softmax}(F_1) \quad (2)$$

然后,进行式(1)、(2)的同时,对原始水下特征进行 3×3 组卷积操作。 3×3 组卷积可以获取到更多的水下空间信息。经过归一化(Norm)处理和ReLU激活后生成空间特征图 F_{rf} 。其中,归一化操作和ReLU激活操作是避免空间特征信息异常以保证后续特征输出稳定。如式(3)所示。

$$F_{rf} = \text{ReLU}(\text{Norm}(\text{Group}_{3 \times 3}(X))) \quad (3)$$

最后,对式(2)、(3)进行乘积操作以生成由空间注意力矩阵调整后的空间特征图,如式(4)所示。

$$F = A_{rf} \times F_{rf} \quad (4)$$

针对C2f中BottleNeck的空间信息整合能力下降问题,对C2f当中的BottleNeck模块进行改进。将原先BottleNeck中 3×3 CBS模块使用RFAConv卷积进行替换以增强空间信息整合能力,具体改进如图3所示。

改进后的BottleNeck_R通过2个步骤实现水下空间

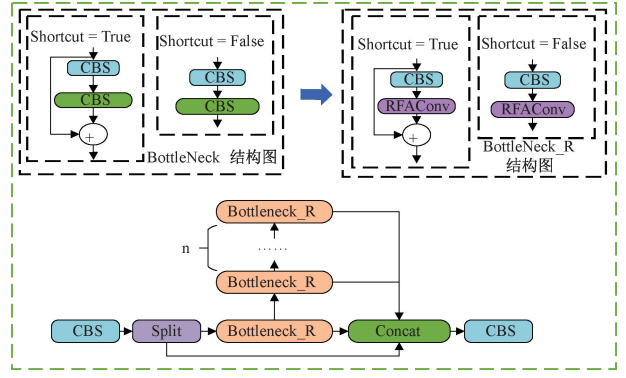


图3 C2f_R结构

Fig. 3 Structure of C2f_R

特征信息的进一步聚合。

第一步,对来自上一层RFAConv输出的空间特征图进行 1×1 卷积操作,以进一步降低计算复杂度。其中, X_{rf} 代表上一层RFAConv输出的空间特征图,如式(5)所示。

$$F' = (\text{Conv}_{1 \times 1}(X_{rf})) \quad (5)$$

第二步,进行RFAConv操作,以进一步强化水下目标信息和空间特征信息的融合能力,如式(6)所示。

$$F_{rf} = \text{RFAConv}(F') \quad (6)$$

综上,使用RFAConv模块改进的CBS模块和C2f单元,可以很好的处理水下检测任务。

2.2 融合EMA注意力机制的CE模块

Song等^[14]在SPPF基础上,通过引出额外的CBS模块分支进行全局信息保留并在最后与SPPF最终融合信息进行连接,有助于不同尺度的特征信息的编码。但在复杂多变的水下场景中,受到水体中的海洋积雪、光源衰弱等背景干扰,原先的CBS模块会将无关的水下背景信息当做关键特征保留,进而干扰最大池化对显著特征的提取能力。因此,针对传统下采样对背景无关特征的错误提取,对CBS前向传播的过程中使用高效多尺度注意力机制规避错误背景信息。具体的SPPF-CE模块结构如图4所示。

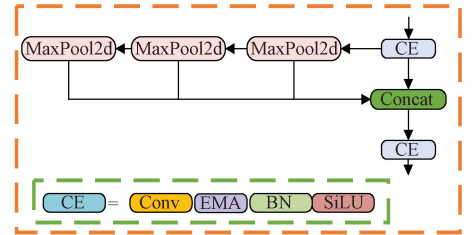


图4 SPPF-CE结构

Fig. 4 Structure of SPPF-CE

高效多尺度注意力机制通过内部的跨空间学习方法,提高重点信息的关注度。高效多尺度注意力机制结构如图5所示。

CE模块通过以下4个步骤实现对水下目标检测的高准确检测。首先,经过 1×1 卷积降维后开始执行EMA

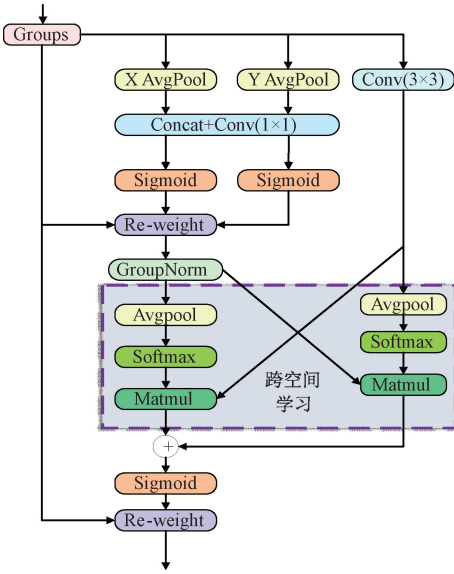


图 5 高效多尺度注意力机制结构

Fig. 5 Structure of EMA

注意力机制,为了学习到更多的通道之间的交互信息,对原始特征 F_k 进行组卷积操作并将其分成 3 部分: F_h, F_w, F_c 再对分组好的特征 F_h, F_w 进行全局平均池化操作,以进一步获取全局信息。如式(8)、(9)所示。特征 F_c 可以很好的保存最初的水下特征信息,避免信息丢失,如式(10)所示。其中, GAP 代表全局平均池化操作, Group 代表组卷积操作,先进行组卷积再进行全局平均池化可以进一步减少参数数量和计算量,提取丰富的水下全局特征。

$$F_k = \text{Conv}_{1 \times 1}(F) \quad (7)$$

$$F_h = \text{GAP}(\text{Group}(F_k)) \quad (8)$$

$$F_w = \text{GAP}(\text{Group}(F_k)) \quad (9)$$

$$F_c = \text{Conv}_{3 \times 3}(\text{Group}(F_k)) \quad (10)$$

其次,对 F_h, F_w 依次进行 Concat 操作和 1×1 卷积操作,再依次把特征进行分割、Sigmoid 后相乘得到特征 F_s , 如式(11)、(12)所示。其中,Concat 的作用是连接两种分支的特征信息,从而丰富了特征的表达能力。再通过 1×1 卷积锁定关键信息。最后的分割操作目的是获得更加丰富的水下梯度信息。

$$X_h, X_w = \text{Split}(\text{Conv}_{1 \times 1}(\text{Concat}(F_h, F_w))) \quad (11)$$

$$F_s = \text{Sigmoid}(X_h) \times \text{Sigmoid}(X_w) \quad (12)$$

接着,得到 F_{ms}, F_{mc} 如式(13)、(14)所示。再对 F_{ms}, F_{mc} 进行 Add 操作输出交叉特征,如式(15)所示。其中,相较于批归一化(batch norm, BN),组归一化(group norm, GN)操作可以很好的处理不同分组之间子特征之间的独有联系。提高了处理特征的能力,有利于后续的水下交叉特征融合。Matmul 操作是快速匹配来自不同尺度下的维度信息,有利于水下交叉信息的表达。

$$F_{ms} = \text{Matmul}(\text{Softmax}(\text{AP}(\text{GN}(F_s))), F_c) \quad (13)$$

$$F_{mc} = \text{Matmul}(\text{Softmax}(\text{AP}(F_c)), \text{GN}(F_s)) \quad (14)$$

$$F_{cross} = F'_{ms} + F'_{mc} \quad (15)$$

然后,交叉空间特征与组卷积的初始特征进行权重调整,如式(16)所示。

$$F_{EMA} = \text{Sigmoid}(F_{cross}) \times \text{Groups}(F) \quad (16)$$

最后,将 EMA 输出的特征依次进行 BN、SiLU 操作,如式(17)所示。

$$F_{CE} = \text{SiLU}(\text{BN}(F_{EMA})) \quad (17)$$

综上,SPPF-CE 模块可以重点关注小密集的水下生物。

2.3 改进的 CCFM-D 模块

Jiang 等^[15]使用 Lark block 模块进行对主干网络进行改进,该模块使用大核卷积的方式对水下特征进行提取,增加了感受野的同时提升了全局特征的捕捉能力。大核卷积的方式会带来大量的参数量,增加了计算开销。Yi 等^[16]依靠 Carafe 内部的动态卷积进行采样,动态卷积同样会带来一定的复杂度。上述改进都是基于 PAN+FPN 网络结构,该结构在网络的最深处易造成参数量和计算量堆积。Carafe 与 Unsample 上采样方式易造成复杂度增加与采集有效信息减少。为解决上述问题,将 RT-DETR^[17] 的 CCFM 结构和上采样模块 Dysample 相结合。CCFM 模块通过结合不同尺度的特征信息来提高模型对图像细节信息的提取能力。Dysample 通过点采样克服 Unsample 插值结果偏离真实值的问题。Dysample 结构图如图 6 所示,图 6 中 C, W, H 为采样前特征图的尺寸, $2gs^2$ 表示经过线性操作后的通道数, g, sW, sH 代表采样后的特征图的尺寸。

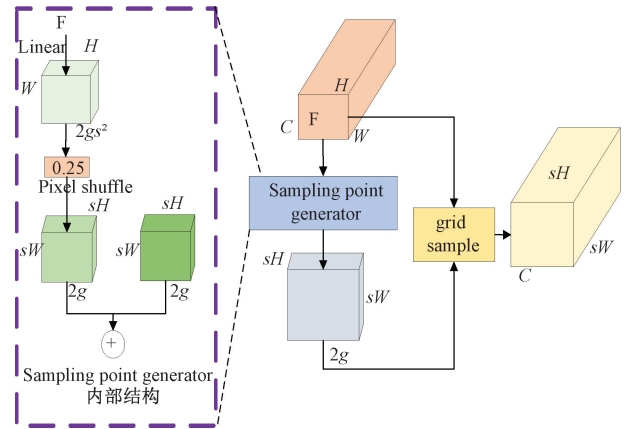


图 6 Dysample 结构

Fig. 6 Structure of dysample

CCFM-D 模块通过以下 5 个步骤,实现对水下目标的精准检测。首先,在 Dysample 上采样之前,先通过 1×1 卷积进行降维处理,可以进一步将水下关键信息保留并降低多余计算量。对 1×1 卷积输出特征进行线性变换(Linear)操作以产生偏移量,偏移量由 Linear 操作中的偏置项产生,增加了模型的表现能力。其中, F 为 SPPE-CE 输出特征,如式(18)所示。

$$F_1 = Linear(Conv_{1\times 1}(F)) \tag{18}$$

其次,对高维投影进行 0.25 倍像素洗牌(Pixelshuffle)操作。并对上述偏移量与原始特征进行相加操作以产生采样集。其中,像素洗牌操作可以最大程度维持水下特征图的信息完整性,减少上采样过程中可能引入的模糊或失真。 O 为像素洗牌后的偏移量, F_1 为线性变换后的偏移量,如式(19)、(20)所示。

$$O = 0.25(Pixelshuffle(F_1)) \tag{19}$$

$$S = O + F \tag{20}$$

接着,对采样集与原始输入特征进行网格采样(Gridsample)操作生成最终特征。其中,网格采样是点采样的一种方式,它可以学习到复杂的空间变换关系,增加水下采样信息采集,如式(21)所示。

$$F' = Gridsample(F, S) \tag{21}$$

然后,对 Backbone 的 C2f 输出特征进行 1×1 卷积处理。把上述特征与 Dysample 输出特征进行 Concat 操作,如式(22)所示。其中, F_b 是来自 Backbone 的特征, F' 代表 Dysample 输出的特征。

$$F_2 = Concat(F', Conv_{1\times 1}(C2f(F_b))) \tag{22}$$

最后,进一步融合处理,如式(23)所示。其中 F_2 代表 Concat 输出的特征。

$$F_{CCFM_D} = C2f(F_2) \tag{23}$$

综上,使用 CCFM-D 模块,对多个维度信息进行降维、融合处理,实现轻量化并提高检测精度。

3 实 验

3.1 数据集

为了更好地贴近海洋真实环境,本文采用 DUO^[18]数据集。水下目标检测数据集 DUO 包含全国水下机器人专业大赛挑战赛(URPC)多年数据集的整合。DUO 数据集删除 URPC 的冗余数据并把错误标签进行改正、提高了数据质量,具体划分如表 1 所示。

表 1 DUO 数据集划分情况
Table 1 Division of the duo dataset

类别	图像数量
训练集	6 004
验证集	667
测试集	1 111

DUO 内包含训练集 6 671 张和测试集 1 111 张。该数据集主要包含 4 类海洋生物:海参、海胆、海星、扇贝。取训练集的 10% 作为验证集,DUO 数据集的图像划分数量和标签数据如表 2 所示。

3.2 评价指标

精度、召回率、AP 和 mAP 是目标检测常用的精度评估指标,它们反映了模型检测的性能。AP 是针对单个类

表 2 数据集标签数据

Table 2 Label data of the dataset

类别	训练集标签	验证集标签	测试集标签
总数	57 351	6 647	10 517
海参	6 080	728	1 079
海胆	38 518	4 437	7 201
扇贝	1 473	234	217
海星	11 280	1 248	2 020

别的性能评估指标,mAP 是在多个类别上的平均检测精度指标。mAP50 和 mAP50:90 越大,表示模型检测效果越好。其中,mAP50:90 更能体现模型在严苛的水下环境下的检测效果。参数量和计算量是目标检测常用的轻量化指标。参数量越大模型的存储和运行时的内存需求越大。计算量 FLOPs 指每秒浮点运算次数,FLOPs^[19] 越大代表消耗的计算资源越多、计算时间越长。具体评价指标如式(24)~(28)所示。

$$P = \frac{TP}{TP + FP} \tag{24}$$

$$R = \frac{TP}{TP + FN} \tag{25}$$

$$AP = \int_0^1 P(R) dR \tag{26}$$

$$mAP = \frac{\sum_{i=1}^k AP_i}{k} \tag{27}$$

$$FLOPs = K^2 \times W \times H \times C_o \times C_i \tag{28}$$

3.3 参数设置

硬件环境采用处理器为 13th Gen Intel(R)Core i5-13490F 显卡为 RTX 4060Ti 显卡,16 GB 显存。软件环境为 Windows 10 操作系统,开发环境为 Pytorch,训练轮次为 200,训练时使用早停机避免过拟合。输入图像为 640×640 ,训练的动量设置为 0.837,迭代批次为 4。使用 SGD 优化器进行优化模型,初始学习率为 0.01。权重衰减设为 0.000 5。

3.4 经典模型对比对比实验

通过 DUO 数据集来评价 RCE-YOLO 的优越性,与其他主流 YOLO 模型进行比较,分别是:v3-tiny^[20]、v5、v5-p6、v6^[21]、YOLOX-tiny^[22]、v8、v10^[23]。其中,基准模型为 YOLOv8。加粗部分为该类别在所有模型中的最优值。RCE-YOLO 的 mAP50、mAP50:90、参数量、计算量分别达到了 83.6%、64.2%、0.034 988 和 7.2 G。尽管在计算量方面与 YOLOv5 相比略微高 0.1 G,但其参数量比 YOLOv5 下降了 18.7%、mAP50 和 mAP50:90 分别提升了 2% 和 2.2%。两种精度指标代表着 RCE-YOLO 在复杂的水下目标检测检测任务中表现良好。对比结果如表 3 所示。

3.5 RCE-YOLO 与其他改进方式对比实验

本部分实验采用与其他最新的改进方法进行对比,其

表 3 经典模型性能对比

模型	mAP50	mAP50:90	参数量	计算量
V3-Tiny	79.6	57.5	12 129 720	18.9
V5	81.8	62.0	2 503 724	7.1
V5-p6	81.6	61.7	4 126 576	7.2
V6	81.0	60.8	4 234 140	11.8
YOLOX-tiny	80.1	54.5	5 055 855	15.4
V8	82.2	63.0	3 006 428	8.1
V10	80.8	61.2	2 695 976	8.2
本文	83.6	64.2	2 034 988	7.2

中两种基准模型分别为 YOLOX 和 YOLOv8,参数量和计算量括号内部的加减代表与基准模型对比,参数量和计算量分别增加或减少了多少。文献[25]的改进方法 mAP 与本文相差 1%。文献[26]的计算量虽与本文方法相近,但是 mAP 还有一段距离。对比结果如表 4 所示。

表 4 文献模型性能对比

模型	基准	参数量	计算量	mAP
文献[24]	X	—	—	62.5
文献[25]	v8s	11.17(+0.03)	16.21(+1.94)	63.2
文献[26]	v8n	—	7.4(−1.3)	58.1
本文	v8n	2.0(−1.0)	7.2(−0.9)	64.2

3.6 融合层改进方法与上采样模块对比实验

通过和其他主流融合层策略进行比较,分别是 PAN+FPN、BiFPN、SlimNeck 以突显出 CCFM-D 在参数量和计算量上的优越性。RCE-YOLO 在 mAP50:90 方面略逊色于基准模型,但在参数量和计算量上分别下降了 34.5% 和 1.5 G。可见 RCE-YOLO 展现出更少的内存需求和计算资源消耗。对比结果如表 5 所示。

表 5 不同融合层方法性能对比

方法	mAP50	mAP50:90	参数量	计算量
PAN+FPN	82.2	63.0	3 006 428	8.1
BiFPN ^[27]	81.9	60.0	2 783 400	8.1
Slim_Neck ^[28]	82.4	60.2	2 798 940	7.3
CCFM-D	82.4	62.8	1 969 436	6.6

通过和其他主流上采样模块进行比较,分别是 Carafe 和 Unsample,以突显出 Dysample 在检测性能上的优越性。改进方法与 Unsample 相比虽然增加了少量的参数,但在检测精度方面优于 Unsample。代表着 Dysample 在复杂多

变的水下目标检测检测任务中表现良好。对比结果如表 6 所示。

表 6 不同上采样模块性能对比

模块	mAP50	mAP50:90	参数量	计算量
Unsample	82.0	62.2	1 965 276	6.6
Carafe ^[29]	81.5	61.8	2 088 996	6.9
Dysample	82.4	62.8	1 969 436	6.6

3.7 SPPF 改进对比实验

通过和其他注意力机制进行比较,分别是 ACMix、CBAM、IRMB,以突显出 EMA 注意力机制在检测精度上和计算量上的优越性。通过对传统 CBS 模块融合不同的注意力机制进行对比。所有指标均优于其他注意力机制。可见模型展现出更少的内存需求和更低的计算消耗。对比结果如表 7 所示。

表 7 不同注意力机制性能对比

方法	mAP50	mAP50:90	参数量	计算量
ACmix ^[30]	83.1	63.8	2 312 664	7.4
CBAM ^[31]	83.0	63.6	2 116 640	7.2
IRMB ^[32]	82.9	63.9	2 366 812	16.9
EMA ^[33]	83.6	64.2	2 034 988	7.2

3.8 热力图分析

为了直观了解 RCE-YOLO 的改进效果,选取密集遮挡、曝光模糊的水下真实场景图片作为热力图分析对象。通过热力图分析对比 YOLOv8 和 RCE-YOLO 在上述场景中的检测效果,以突显出 RCE-YOLO 的优越性。密集遮挡和曝光模糊热力图分析如图 7 和 8 所示。

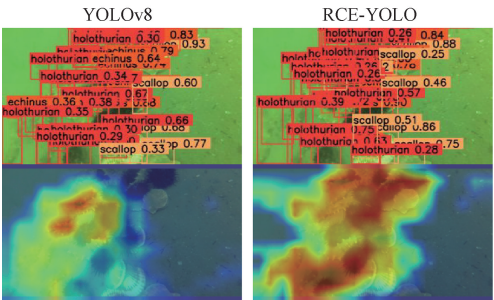


图 7 密集遮挡目标热力结果

Fig. 7 Thermal results of densely occluded targets

图 7 和 8 中展示了密集遮挡场景和曝光模糊场景的目标检测结果图及相应热力图。展示的两种情况能够更加贴

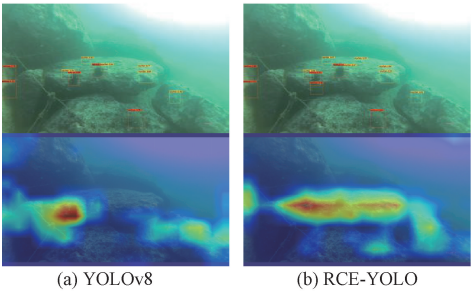


图 8 模糊目标检测、热力结果

Fig. 8 Fuzzy object detection and thermal results

切真实的海洋环境。从图 7 观察到大范围密集的海参、海胆、扇贝互相遮挡。尤其是在热力图片的上部和中心密集

区域,在覆盖范围和关注程度方面,YOLOv8 效果逊色于 RCE-YOLO。由于水下生物的互相遮挡,在检测效果上也弱于 RCE-YOLO。尤其是中心区域的海参和右边区域的扇贝均出现了不同程度的漏检。总的来说,RCE-YOLO 漏检率要小于基准模型。从图 8 热力图可以观察到,越靠近光源位置 YOLOv8 对水下生物关注程度越差,在中间的位置几乎没有热力覆盖。在没有热力覆盖的地方出现了海星漏检。总体来看,RCE-YOLO 对模糊目标和密集遮挡目标的检测能力较强,适用于水下真实场景的目标检测。

3.9 消融实验

本节通过消融实验,验证不同改进方法对本文模型性能的影响。采用 DUO 数据集对基准模型进行消融实验,结果如表 8 所示。

表 8 消融实验性能分析

Table 8 Performance analysis of ablation experiment

方法	RFACnv	C2f_RFA	CCFM	CCFM-D	SPPF-CE	mAP50	mAP50:90	参数量	计算量
A	—	—	—	—	—	82.2	63.0	3 006 428	8.1
B	✓	—	—	—	—	83.6	64.2	3 032 828	8.3
C	—	✓	—	—	—	83.1	63.7	3 044 348	8.4
D	—	—	✓	—	—	82.0	62.2	1 965 276	6.6
E	—	—	—	—	✓	82.4	63.1	3 007 660	8.1
F	—	—	✓	✓	—	82.4	62.8	1 969 436	6.6
G	✓	✓	—	—	—	83.8	64.2	3 070 748	8.7
H	✓	✓	✓	✓	—	83.6	63.8	2 033 756	7.1
I	✓	✓	✓	✓	✓	83.6	64.2	2 034 988	7.2

方法 B 使用 RFACnv 替换 YOLOv8 主干网络中的部分 CBS,提高了目标检测精度。实验 C 在 C2f 的 BottleNeck 中替换 3×3 卷积,进一步增加 C2f 的水下信息整合能力。实验 B 和实验 C 在提升检测精度上有着良好的效果,但是参数量和计算量也有所增加。实验 D 和实验 F 都对 Neck 进行一定的改动。其中,实验 F 在实验 D 的基础上将 Unsample 替换成了 Dysample。方法 F 能够增加水下采样信息的丰富度。相较于实验 D,mAP50 与 mAP50:90 分别提升了 0.2% 和 0.6%。实验 E 将 CBS 结构与 EMA 注意力机制相融合,模型能学习到不同通道之间的交互信息。相较于基准模型 A,最终的改进模型 I 在各方面都有一定提升,mAP50 提升了 1.4%、mAP50:90 提升了 1.2%,参数量、计算量分别下降 32.3% 和 0.9 G。说明改进的模型实现了精度更高的检测效果的同时降低了内存需求和计算资源的消耗。

3.10 不同模型的检测结果分析

图 9 展示了在密集遮挡环境、模糊环境、昏暗环境下主流模型与 RCE-YOLO 的检测效果对比,并把漏检、错检情况在图中标注了出来。在昏暗环境下,大多数模型出现错误检测。而 RCE-YOLO 表现出更加良好的防错检、漏检效果。在密集遮挡环境下所有模型出现错检和漏检,RCE-YOLO 对比其他模型只出现少量漏检。在模糊场景

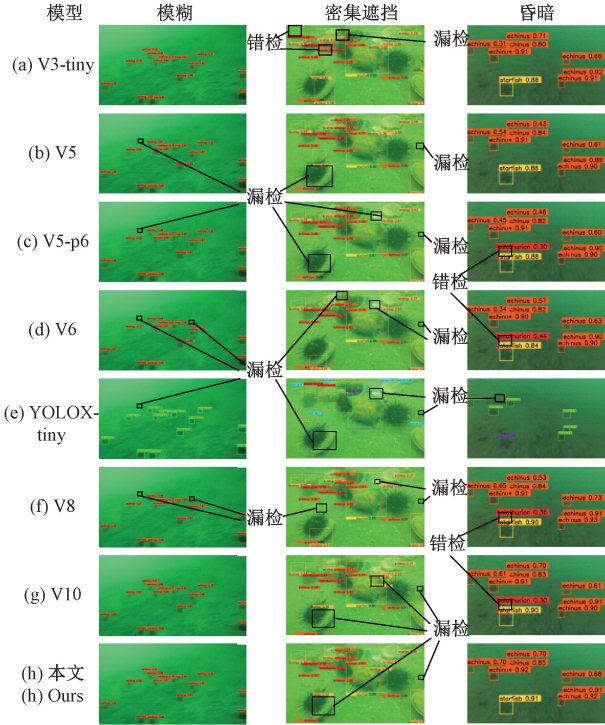


图 9 不同模型的检测结果

Fig. 9 Detection results of different models

下, YOLOv3-tiny、YOLOv10、RCE-YOLO 没有出现错检情况。YOLOv5、YOLOv5-p6、YOLOv6、YOLOv8 漏检了最远处的海胆。总体来看, 本文对于 YOLOv8 的整体改进提高了模型对昏暗、模糊和密集遮挡目标的检测能力, 使其适应于水下目标检测场景。

4 结 论

本文提出了一种基于 YOLOv8 轻量化水下目标检测模型。首先, 引入 RFACnv 模块对 CBS 模块和 C2f 模块进行替换和改进, 以强化水下目标信息和感受场特征信息的聚合能力。其次, 在 SPPF 模块的基础上, 设计 CE 模块。将其替换到 SPPF 模块的首尾两处, 使得模型更加关注水下目标并减少特征丢失。最后, 将 CCFM 模块与 Dysample 模块进行结合, 减少模型的参数量和计算量。在 DUO 数据集上的测试结果表明, 该模型在检测精度、参数量和计算量方面超越基准方法。与其他经典模型相比, RCE-YOLO 具有很好的综合性能。但改进后的 RCE-YOLO 与其他大参数量模型相比, 在检测精度上有待改进, 但总体在复杂的水下场景中依然能保持良好的检测效果。未来将针对这一问题设计一个系统, 并进行改进。

参考文献

- [1] XU SH B, ZHANG M H, SONG W, et al. A systematic review and analysis of deep learning-based underwater object detection [J]. *Neurocomputing*, 2023, 527: 204-232.
- [2] REN SH Q, HE K M, GIRSHICK R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks [C]. *Advances in Neural Information Processing Systems*, 2015: 91-99.
- [3] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: Unified, real-time object detection[C]. *IEEE Conference on Computer Vision and Pattern Recognition*, 2016: 779-788.
- [4] 周新, 张春堂, 樊春玲. 基于 YOLOv5_PGS 的轻量级水下生物识别目标检测 [J]. *电子测量术*, 2023, 46(21): 168-175.
ZHOU X, ZHANG CH T, FAN CH L. Lightweight YOLOv5 _ PGS based objective detection for underwater biological identification [J]. *Electronic Measure Technology*, 2023, 46(21): 168-175.
- [5] LYU L, LIU Y, XU X CH, et al. EFP-YOLO: A quantitative detection algorithm for marine benthic organisms[J]. *Ocean & Coastal Management*, 2023, 243: 106770.
- [6] YU G Y, CAI R L, SU J P, et al. U-YOLOv7: A network for underwater organism detection [J]. *Ecological Informatics*, 2023, 75: 102108.
- [7] ZHANG M H, WANG ZH H, SONG W, et al. Efficient small-object detection in underwater images using the enhanced YOLOv8 network [J]. *Applied Sciences*, 2024, 14(3): 1095.
- [8] 施克权, 李祺, 隋皓, 等. IEMAYOLOViT: 基于改进 YOLOv8 的水下目标检测算法 [J]. *电讯技术*, 2024: 1-10. DOI:10.20079/j.issn.1001-893x.231206005.
SHI K Q, LI Q, SUI H, et al. IEMAYOLOViT: An underwater target detection algorithm based on improved YOLOv8 [J]. *Telecommunication Engineering*, 2024: 1-10. DOI:10.20079/j.issn.1001-893x.231206005.
- [9] QIN X R, YU CH D, LIU B SH, et al. YOLO8-FASG: A high-accuracy fish identification method for underwater robotic system [J]. *IEEE Access*, 2024, 12: 73354-73362.
- [10] ZHOU H, KONG M W, YUAN H X, et al. Real-time underwater object detection technology for complex underwater environments based on deep learning [J]. *Ecological Informatics*, 2024, 82: 102680.
- [11] ZHANG X, LIU C, YANG D G, et al. RFACnv: Innovating spatial attention and standard convolutional operation [J]. *ArXiv preprint arXiv: 2304.03198*, 2023.
- [12] LIU W Z, LU H, FU H T, et al. Learning to upsample by learning to sample [C]. *IEEE/CVF International Conference on Computer Vision*, 2023.
- [13] LIU L L, CHU CH X, CHEN CH CH, et al. MarineYOLO: Innovative deep learning method for small target detection in underwater environments [J]. *Alexandria Engineering Journal*, 2024, 104: 423-433.
- [14] SONG G W, CHEN W, ZHOU Q L, et al. Underwater robot target detection algorithm based on YOLOv8 [J]. *Electronics*, 2024, 13(17): 3374.
- [15] JIANG X, ZHUANG X T, CHEN J SH, et al. YOLOv8-MU: An improved YOLOv8 underwater detector based on a large kernel block and a multi-branch reparameterization module [J]. *Sensors*, 2024, 24(9): 2905.
- [16] YI W G, YANG J W, YAN L W. Research on underwater small target detection technology based on single-stage USSTD-YOLOv8n [J]. *IEEE Access*, 2024, 12: 69633-69641.
- [17] ZHAO Y AN, LYU W Y, XU SH L, et al. DETRs beat YOLOs on real-time object detection [C]. *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2024: 16965-16974.
- [18] LIU CH W, LI H J, WANG SH CH, et al. A dataset

- and benchmark of underwater object detection for robot picking[C]. 2021 IEEE International Conference on Multimedia & Expo Workshops(ICMEW), IEEE, 2021; 1-6.
- [19] 沈学利,王灵超. 基于 YOLOv8n 的无人机航拍目标检测[J]. 计算机系统应用, 2024, 33(7): 139-148.
- SHEN X L, WANG L CH. UAV aerial photography target detection based on YOLOv8n[J]. Computer Systems Applications, 2024, 33(7): 139-148.
- [20] REDMON J, FARHADI A. YOLOv3: An incremental improvement[J]. ArXiv preprint arXiv:1804.02767, 2018.
- [21] LI CH Y, LI L L, JIANG H L, et al. YOLOv6: A single-stage object detection framework for industrial applications [J]. ArXiv preprint arXiv: 2209.02976, 2022.
- [22] GE ZH, LIU S T, WANG F, et al. YOLOx: Exceeding YOLO series in 2021[J]. ArXiv preprint arXiv:2107.08430, 2021.
- [23] WANG AO, CHEN H, LIU L H, et al. YOLOv10: Real-time end-to-end object detection [J]. ArXiv preprint arXiv:2405.14458, 2024.
- [24] CHEN J, ER M J, ZHANG Y N, et al. Novel dynamic feature fusion strategy for detection of small underwater marine object[C]. 2022 5th International Conference on Intelligent Autonomous Systems(ICoIAS), IEEE, 2022; 24-30.
- [25] LIU ZH Y, WANG B, LI Y, et al. UnitModule: A lightweight joint image enhancement module for underwater object detection[J]. Pattern Recognition, 2024, 151: 110435.
- [26] FAN Y B, ZHANG L Y, LI P. A lightweight model of underwater object detection based on YOLOv8n for an edge computing platform[J]. Journal of Marine Science and Engineering, 2024, 12(5): 697.
- [27] TAN M X, PANG R M, LE Q V. Efficientdet: Scalable and efficient object detection[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020; 10781-10790.
- [28] LI H L, LI J, WEI H B, et al. Slim-neck by GSConv: A lightweight-design for real-time detector architectures [J]. Journal of Real-Time Image Processing, 2024, 21(3): 62.
- [29] WANG J Q, CHEN K, XU R, et al. Carafe: Content-aware reassembly of features[C]. IEEE/CVF International Conference on Computer Vision, 2019: 3007-3016.
- [30] PAN X R, GE CH J, LU R, et al. On the integration of self-attention and convolution [C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022; 815-825.
- [31] 王子昭,景明利,史金钢,等. 一种改进 CBAM 机制和细节恢复的单幅图像去雾算法[J]. 电子测量技术, 2023, 46(2): 161-168.
- WANG Z ZH, JING M L, SHI J G, et al. A single image defogging algorithm based on improved CBAM mechanism and detail recovery [J]. Electronic Measurement Technology, 2023, 46(2): 161-168.
- [32] ZHANG J N, LI X T, LI J, et al. Rethinking mobile block for efficient attention-based models[C]. IEEE/CVF International Conference on Computer Vision (ICCV), IEEE Computer Society, 2023: 1389-1400.
- [33] 卢开喜,段先华,陶宇诚,等. KThin-YOLOV7:轻量级的焊接件表面缺陷检测[J]. 电子测量术, 2024, 47(7): 9-18.
- LU K X, DUAN X H, TAO Y CH, et al. KThin-YOLOV7: Lightweight inspection of surface defects on welded parts [J]. Electronic Measurement Technology, 2024, 47(7): 9-18.

作者简介

周志耀, 硕士研究生, 主要研究方向为目标检测。

马常霞(通信作者), 博士, 主要研究方向为计算机视觉、模式识别。

E-mail: 1997000062@jou.edu.cn

杨丽莎, 硕士研究生, 主要研究方向为目标检测、多模态。

仲兆满, 博士, 主要研究方向为海洋大数据采集与分析。

胡文彬, 博士, 主要研究方向为模式识别。