

# 基于改进 YOLO v7 的鲑鱼检测模型轻量化研究

郑荣才<sup>1</sup> 谭鼎文<sup>1,2</sup> 徐青<sup>2</sup> 陈大勇<sup>1</sup> 元轶新<sup>1</sup>

(1. 南方海洋科学与工程广东省实验室(湛江), 湛江 524013;

2. 广东海洋大学机械工程学院, 湛江 524088)

**摘要:** 为实现水下复杂环境下鲑鱼的快速准确识别, 提出一种基于 YOLO v7 轻量化的鲑鱼检测模型 YOLO v7 - CSMRep。首先, 采用 Stem 模块合并 Backbone 层的前 4 个卷积操作, 有效降低了模型计算量。其次, 使用多尺度重参数化 (Multi-directional reparameterization, MRep) 模块替代 YOLO v7 的 ELAN 和 ELAN - H 模块, 增强了单向特征提取能力, 同时大幅减少参数数量和计算量。最后, 在 Backbone 层末端集成卷积块注意力模块 (Convolutional block attention module, CBAM), 提升网络空间和通道特征提取能力。试验结果表明, 改进后模型内存占用量、参数数量和计算量分别降低 4.28%、5.29%、31.30%, F1 值、mAP<sub>0.5</sub> 分别提高 0.5、0.7 个百分点, 分别达到 93.1%、97.1%, 帧率提高 15.41%, 达到 140.8 f/s。对比 YOLO v5s、YOLO v6s、YOLO v7、YOLO v7 - tiny、YOLO v8s 模型, mAP<sub>0.5</sub> 分别提高 1.0、2.0、0.7、0.8、1.2 个百分点。因此, 本文提出的方法能够快速而准确地识别鲑鱼, 可为深远海养殖生物量监测提供技术支持。

**关键词:** 深远海养殖; 鲑鱼检测; YOLO v7; Stem 模块; 多尺度重参数化; 卷积块注意力模块

**中图分类号:** S951.2; TP391.4 **文献标识码:** A **文章编号:** 1000-1298(2024)11-0132-08

**OSID:**



## Lightweight Salmon Detection Model Based on Improved YOLO v7

ZHENG Rongcai<sup>1</sup> TAN Dingwen<sup>1,2</sup> XU Qing<sup>2</sup> CHEN Dayong<sup>1</sup> YUAN Kexin<sup>1</sup>

(1. Southern Marine Science and Engineering Guangdong Laboratory (Zhanjiang), Zhanjiang 524013, China

2. College of Mechanical Engineering, Guangdong Ocean University, Zhanjiang 524088, China)

**Abstract:** In order to achieve rapid and accurate identification of salmon in complex underwater environments, a lightweight salmon detection model, YOLO v7 - CSMRep, was proposed based on YOLO v7. Firstly, by adopting the Stem module, the first four convolutional operations in the backbone layer were merged into an efficient convolutional operation, reducing the computational load of the model. Secondly, the ELAN and ELAN - H modules of the YOLO v7 network were replaced with the multi-directional reparameterization (MRep) module, which enhanced the one-way feature extraction capability while greatly reducing parameters and calculations. Finally, at the end of the backbone layer, the convolutional block attention module (CBAM) was integrated to enhance the network's spatial and channel feature extraction capabilities. The experimental results showed that the improved model's volume, parameter count, and computational load were reduced by 4.28%, 5.29% and 31.30%, respectively. The F1 score and mAP<sub>0.5</sub> were increased by 0.5 and 0.7 percentage points, and reached 93.1% and 97.1%, respectively. Additionally, the frame rate was increased by 15.41%, and reached 140.8 f/s. Compared with that of YOLO v5s, YOLO v6s, YOLO v7, YOLO v7 - tiny, and YOLO v8s models, the mAP<sub>0.5</sub> was improved by 1.0, 2.0, 0.7, 0.8, and 1.2 percentage points, respectively. Therefore, the method proposed can rapidly and accurately identify salmon and provide technical support for biomass monitoring in deep-sea aquaculture.

**Key words:** deep-sea aquaculture; salmon detection; YOLO v7; Stem module; multi-directional reparameterization; convolutional block attention module

收稿日期: 2024-02-28 修回日期: 2024-04-26

**基金项目:** 国家重点研发计划项目(2022YFD2401201)、广东省海洋经济发展(海洋六大产业)专项资金项目(GDNRC[2023]33)、南方海洋科学与工程广东省实验室(湛江)项目(011Z23002)和湛江湾实验室人才团队引进科研项目(ZJW-2023-05)

**作者简介:** 郑荣才(1973—),男,研究员,博士,主要从事海洋装备和智能机器人控制系统研究,E-mail: zrc618@126.com

**通信作者:** 徐青(1980—),男,教授,博士,主要从事海洋能源利用研究,E-mail: xuqing@gdou.edu.cn

## 0 引言

随着我国水产养殖的多样性不断扩大,渔业资源衰退问题日益凸显。为此,中央一号文件强调了现代海洋牧场和深远海养殖的建设<sup>[1]</sup>。水下图像目标检测技术<sup>[2]</sup>在深远海养殖中可用于生物量估算、存活率评估和养殖管理,目前的检测模型复杂且对硬件设备要求高,这限制了其在实际深远海环境中的应用。因此,在保持较高检测精度的前提下,轻量化研究以适应边缘设备的资源限制变得尤为重要。

目标检测技术分为双阶段的 R-CNN<sup>[3]</sup> 系列(包括 Fast R-CNN<sup>[4]</sup>、Faster R-CNN<sup>[5]</sup>、Mask R-CNN<sup>[6]</sup> 等)和单阶段的 SSD<sup>[7]</sup>、YOLO 系列<sup>[8-13]</sup>。单阶段算法通过端到端的检测方式直接输出预测结果,因其结构简洁高效、快速检测的特点而被广泛使用。在深度学习网络轻量化研究中,采用了多种策略,包括设计轻量网络架构<sup>[14]</sup>、网络剪枝<sup>[15]</sup>、知识蒸馏<sup>[16]</sup> 和模型压缩<sup>[17]</sup>。例如, MobileNet<sup>[18-21]</sup>、ShuffleNet<sup>[22]</sup>、GhostNet<sup>[23]</sup> 和 EfficientNet<sup>[24]</sup> 等轻量化模型,通过高效特征提取结构降低计算和参数量,实现了在维持性能的同时减小模型规模。

近年来,深度学习在复杂环境中的应用越来越受到关注<sup>[25-30]</sup>。在水下鱼类检测应用中, SUNG 等<sup>[31]</sup>将地面目标检测的 YOLO 模型应用于水下鱼类识别,虽然在精度和实时性方面取得了成就,但在适应水下机器人硬件限制和光照变化方面仍有局限。AI MUKSIT 等<sup>[32]</sup>提出了 YOLO-Fish-1 和 YOLO-Fish-2 两种轻量模型,有效改进了 YOLO v3 在小型鱼类和复杂海洋环境下的检测能力,但这些模型的实际海洋环境检测准确率尚需提高。此外,李庆忠等<sup>[33]</sup>提出的基于改进 YOLO 模型和迁移学习的水下鱼类实时检测算法,在处理小目标和重叠目标方面表现出色,但在应对水下环境的多变性方面还有改进空间。汪沛洁等<sup>[34]</sup>对 YOLO v5 模型的改进虽然提升了小鱼群检测的精确度,但在不同环境下的泛化性和对多种大小目标的适应性方面仍有待优化。涂万等<sup>[35]</sup>对 YOLO v8 模型改进,增加大尺寸检测头和引入 ECAM 注意力机制,在水下鱼群检测精度上有所提升,但在处理复杂环境,如水面反光的检测方面,效果仍需优化。这些研究展示了深度学习在应对水下鱼类检测等复杂场景的潜力,突出了提升模型效能和适应性的重要性。

本文以深远海网箱养殖的鲑鱼为研究对象,在基于 YOLO v7 算法的基础上,提出一种改进轻量化目标识别模型 YOLO v7-CSMRep。采用轻量 Stem 模块,通过深度可分离卷积和  $1 \times 1$  卷积等轻量化技

术,将骨干网络 0~3 层的 4 个卷积操作合并成 1 个高效卷积操作,以降低模型内存占用量和计算量;引入 MRep 模块代替原有的 ELAN 和 ELAN-H 模块,以提高 X 和 Y 方向特征提取能力,加快网络特征提取,并通过重参数化减少参数量;同时,结合 CBAM 注意力模块,以提高模型检测精度,使模型具备更强的抗干扰能力。

## 1 材料与方法

### 1.1 数据集构建

本试验数据来自山东某公司的养殖网箱,该网箱的水下高清摄像机由青岛杰瑞工控技术有限公司研制,型号为 JARI-HC-100-A1,水下 LED 补光灯型号为 JARI-HL-100-A1。于 2022 年 12 月 12 日采集水下鲑鱼监测视频,利用每 15 帧抽取 1 帧的方法,共获取约 4 600 幅  $1\,920$  像素  $\times$   $1\,080$  像素的初始图像。为确保数据质量,经过人工清洗,剔除了因鲑鱼运动模糊、遮挡摄像头以及水下补光灯过度曝光导致的低质量图像,筛选出 1 200 幅高质量的图像。由于鲑鱼数据集样本的单一性,为了增强模型的泛化能力和减少过拟合,保留 400 幅用于模型测试,对剩余的 800 幅图像进行高斯噪声、亮度调整、翻转和镜像等数据增强处理,从增强后的 4 800 幅图像中精选 3 600 幅用于模型的训练和验证,最终构建了包含 4 000 幅图像的数据集。部分数据增强图像如图 1 所示。

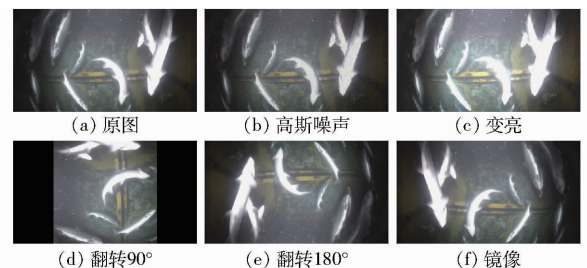


图 1 鲑鱼数据增强图像

Fig. 1 Data enhanced images of salmon

利用 LabelImg 标注软件对这些图像中的鲑鱼进行人工标注,标注样例如图 2 所示。标注标签为“salmon”,生成包含目标类别和坐标信息的 txt 文件。最后,按照比例 7:2:1 将数据集随机划分为训练集、验证集和测试集,对应的图像数量分别为 2 800、800、400 幅。

### 1.2 YOLO v7 检测模型

YOLO v7<sup>[13]</sup>是原 YOLO v4<sup>[11]</sup>团队在 2022 年提出的目标检测算法,其网络结构主要包括 4 部分: Input、Backbone、Neck 和 Head。在 Input 层,图像经过预处理后被输入到 Backbone 层;Backbone 层采用

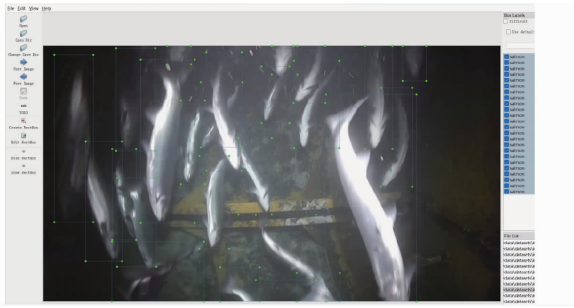


图2 鲑鱼人工标注图

Fig.2 Manual marking of salmon

高效的 E-ELAN 网络以提高网络的特征提取能力,同时引入 MPCnv 模块代替普通的池化层,实现更有效的下采样操作,以避免特征丢失和图像失真; Neck 层使用 PANet 结构实现多尺度特征的有效融合; Head 层采用 RepConv 模块,通过重参数化对不同尺寸的 PANet 特征进行通道数整合,同时保持精

度并提高推理速度。

本文基于 YOLO v7 基础架构提出了一种改进的轻量化识别模型,改进后的 YOLO v7 网络结构如图 3 所示。首先,模型引入了轻量 Stem 模块,利用深度可分离卷积和  $1 \times 1$  卷积技术来合并前几层的卷积操作,形成一个更高效的卷积单元,显著减小了模型内存占用量和计算量。其次,采用 MRep 模块替代原有的 ELAN 和 ELAN-H 模块,该模块基于 RepVGG 架构增加了  $X$  和  $Y$  方向的特征提取,并利用重参数化技术减少参数量,加快特征提取过程,提升了模型的特征提取能力,使模型在处理复杂图像时更加准确和快速。在骨干网络末端的 MRep 模块之后加入了 3 个 CBAM 模块,通过通道和空间注意力机制关注重要特征,提高模型对重要信息的关注能力,增强了检测精度和对环境干扰的抗性。

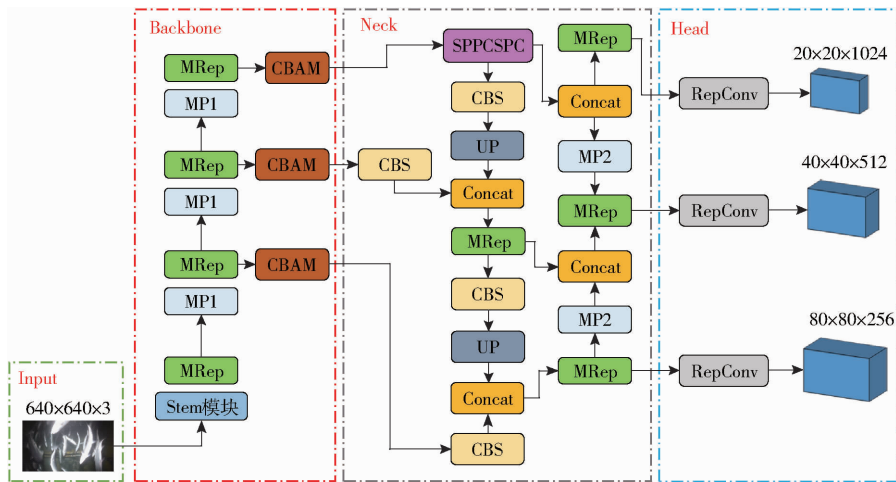


图3 改进的 YOLO v7 网络结构图

Fig.3 Improved YOLO v7 network structure diagram

### 1.3 添加轻量 Stem 模块

轻量 Stem 模块通过优化 YOLO v7 网络起始部分的卷积层,有效减少了模型内存占用量和计算量。如图 4 所示,Stem 模块通过一个分支网络的多个卷积层、批量归一化和 ReLU 激活函数进行特征提取,以确保非线性和稳定性。另一个分支使用最大池化操作减小特征图尺寸。这两个分支处理后的特征图通过 Concat 连接整合,从而在保持模型轻量的同时提高网络特征识别能力。

### 1.4 引入 MRep 模块

RepVGG<sup>[36]</sup> 以传统 VGG<sup>[37]</sup> 为骨干,同时融入残差结构,根据模型训练和推理阶段的特性差异,设计出不同的训练策略。其核心思想是将深度卷积操作转换为一系列由  $1 \times 1$  卷积和  $3 \times 3$  卷积构成的基本卷积模块,并引入重参数化来简化网络结构和加速推理,RepVGG 分为 5 个阶段,部分结构如图 5

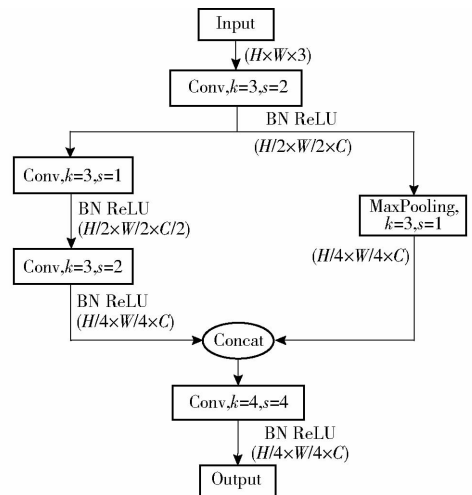


图4 Stem 模块

Fig.4 Stem module

所示。

本文改进的 MRep 模块在 RepVGG 的基础上引入  $X$  和  $Y$  方向的卷积操作,可增强对单方向特征的

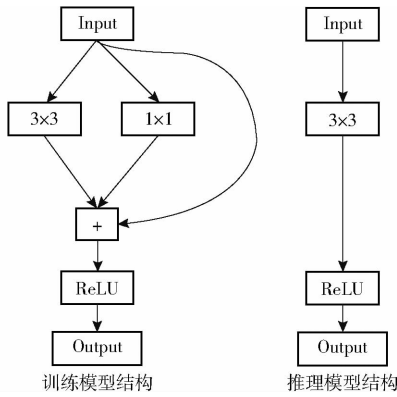


图 5 RepVGG 部分结构示意图

Fig. 5 Schematic of RepVGG part structure

建模能力,再通过重参数化,实现了使用少量卷积块就能够达到和 ELAN 相近性能的目标,如图 6 所示。同时,通过快速特征提取模块,显著减少了参数量和卷积算子,从而提高了推理速度。

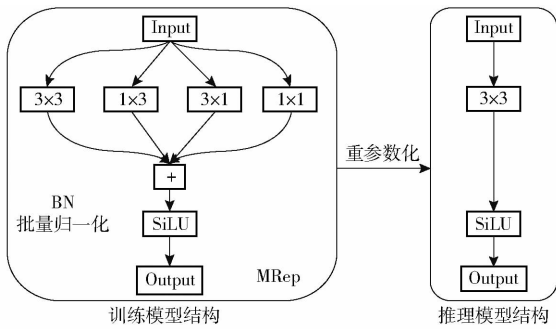


图 6 MRep 结构示意图

Fig. 6 Schematic of MRep structure

### 1.5 加入 CBAM 注意力模块

由于鲑鱼的形状极其相似,同时游动时存在严重的遮挡,CBAM 模块可以使模型更聚焦于重要的特征,从而提高检测准确率。CBAM 模块由通道注意力模块和空间注意力模块组成。为了平衡模型精度和轻量化,在骨干网络末端 MRep 模块后加入 3 个 CBAM 模块,CBAM 模块如图 7 所示。

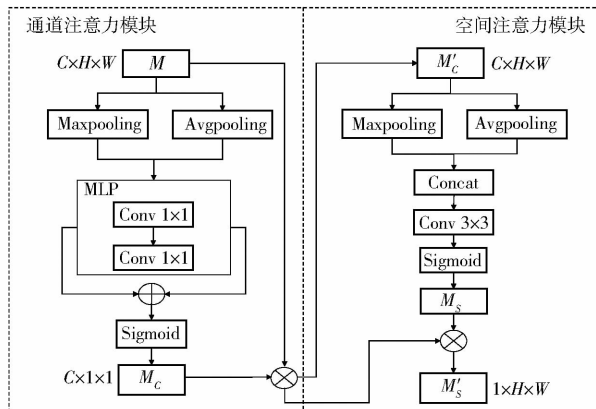


图 7 CBAM 模块

Fig. 7 CBAM module

其中通道注意力模块先对输入特征图  $M$  进行

全局池化操作,随后通过多层感知机 MLP 和 Sigmoid 激活函数生成通道注意力特征图  $M_c$ ,以增大有效通道权重;空间注意力模块对通道注意力特征图进行全局池化,经过全连接层和 Sigmoid 函数生成空间注意力特征矩阵  $M_s$ ,最后通过逐元素相乘得到缩放后的新特征图  $M'_s$ ,使得模型能够关注输入特征图中重要的空间位置,从而提高特征表征能力。 $M_c$  和  $M_s$  计算公式为

$$M_c = \sigma_{\text{sigmoid}}(\text{MLP}(\text{AvgPool}(M))) + \text{MLP}(\text{MaxPool}(M)) \quad (1)$$

$$M_s = \sigma_{\text{sigmoid}}(f^{(7 \times 7)} \text{Con}(\text{AvgPool}(M_c), \text{MaxPool}(M_c))) = S(f^{(7 \times 7)} \text{Con}(M_{\text{cavg}}^s, M_{\text{cmax}}^s)) \quad (2)$$

- 式中  $M_c$ ——通道注意力特征矩阵
- $\sigma_{\text{sigmoid}}$ ——Sigmoid 激活函数
- $\text{AvgPool}(M)$ ——对  $M$  应用平均池化
- $\text{MaxPool}(M)$ ——对  $M$  应用最大池化
- $M_s$ ——空间注意力特征矩阵
- $f^{(7 \times 7)}$ —— $7 \times 7$  的卷积
- $\text{Con}$ ——Concat 操作
- $M_{\text{cavg}}^s$ ——平均池化操作得到的特征
- $M_{\text{cmax}}^s$ ——最大池化操作得到的特征

### 1.6 试验环境与参数设置

本试验使用的操作系统为 Linux (Ubuntu 20.04), CPU 为 Gold 6330 @ 2.00 GHz, GPU 为 RTX 3090(24 GB)。编程语言选择 Python 3.8,深度学习框架为 PyTorch 2.0, GPU 加速库为 CUDA 11.8。在训练时,迭代次数为 300 次,批处理量设为 16,初始学习率设为 0.01,动量设为 0.937。

### 1.7 评价指标

采用精确率 (Precision,  $P$ )、召回率 (Recall,  $R$ )、F1 值 (F1-score)、平均精度均值 (Mean average precision, mAP)、浮点计算量 (Floating point operations, FLOPs)、模型参数量 (Params)、帧率 (Frames per second) 和模型内存占用量作为模型性能评价指标。

## 2 试验结果分析

### 2.1 消融试验

为验证改进模型 YOLO v7 - CSMRep 的性能有效性,以 YOLO v7 为基线模型,在自制的鲑鱼数据集上进行验证,评估各个改进方法对基线模型性能的影响。消融试验结果如表 1 所示,表中“ $\checkmark$ ”表示引入此方法,“ $\times$ ”表示不加入此方法。

表1 消融试验结果

Tab.1 Ablation test results

模型编号	Stem	MRep	CBAM	参数量	模型内存占用量/MB	浮点计算量	F1 值/%	mAP <sub>0.5</sub> /%	帧率/(f·s <sup>-1</sup> )
1	×	×	×	3.648 × 10 <sup>7</sup>	74.7	1.032 × 10 <sup>11</sup>	92.6	96.4	122.0
2	√	×	×	3.634 × 10 <sup>7</sup>	74.4	8.680 × 10 <sup>10</sup>	92.9	96.8	135.1
3	×	√	×	3.289 × 10 <sup>7</sup>	68.9	7.960 × 10 <sup>10</sup>	92.9	96.8	151.5
4	×	×	√	3.664 × 10 <sup>7</sup>	75.1	1.034 × 10 <sup>11</sup>	93.1	96.8	120.5
5	√	√	×	3.275 × 10 <sup>7</sup>	68.6	6.350 × 10 <sup>10</sup>	93.0	96.8	192.3
6	√	√	√	3.455 × 10 <sup>7</sup>	71.5	7.090 × 10 <sup>10</sup>	93.1	97.1	140.8

由表1可知,模型1为原始YOLO v7算法,模型内存占用量为74.7 MB,浮点计算量为1.032 × 10<sup>11</sup>,F1值、mAP<sub>0.5</sub>分别为92.6%、96.4%,帧率为122.0 f/s,以此作为评价优化效果的基准。

在模型优化方面,模型2采用Stem模块优化初级网络层结构,通过精简参数量同时保持关键特征的提取能力,大部分指标都有提升。模型3引入MRep模块替代原先的ELAN和ELAN-H模块,强化了单方向特征提取能力,并通过加速网络特征处理和减少冗余计算,与模型1对比,性能指标有较大的提升。模型4通过在骨干网络末端集合CBAM注意力机制,增强了网络末端对图像关键区域的聚焦效果,尽管略增计算量和模型内存占用量,但提高了网络的特征识别能力和抗干扰性。模型5结合Stem模块和MRep模块的优势,显著提升检测速度并减少了计算量,在6组消融试验中表现出色,特别是除F1值和mAP之外的指标取得了最佳效果。模型2~5中的mAP<sub>0.5</sub>都为96.8%,表明单独引入Stem模块、MRep模块、CBAM模块及其部分组合虽然各自针对模型的不同改进方向进行了优化,但这些改进对综合性能指标mAP<sub>0.5</sub>的提升效果相对有限。

模型6为本文方法,综合了前述所有策略,与模型1相比,它的性能全面提升,F1值、mAP<sub>0.5</sub>分别提高约0.5、0.7个百分点,达到93.1%、97.1%,帧率提高15.41%,达到140.8 f/s,同时模型内存占用量、参数量、计算量分别减少4.28%、5.29%、31.30%。尽管模型6大多数指标未能超过模型5,但它的精确率和F1值有小幅提升,并且其他指标表现稳定。因此,综合考虑鲑鱼游动产生的遮挡及其他因素,最终确定模型6作为本文方法。

## 2.2 不同注意力机制对比试验

为验证本文选取CBAM模块作为特征提取的有效性,将其与其他主流的注意力机制进行多组对比试验,试验结果如表2所示。

由表2可知,引入不同类型的注意力机制后,检

表2 不同注意力机制试验结果

Tab.2 Results of experiments with different attention mechanisms

注意力机制	P/%	R/%	F1值/%	mAP <sub>0.5</sub> /%	mAP <sub>0.5:0.95</sub> /%	模型内存占用量/MB
无	92.9	92.0	92.4	96.4	73.7	74.7
ECA	94.1	91.9	93.0	96.8	74.5	74.7
CA	94.1	91.8	92.9	96.8	74.7	75.2
CBAM	94.2	92.1	93.1	96.9	74.8	75.1

测模型的精度都略有提升。相比之下,ECA模块在不增加模型尺寸的情况下,F1值提升0.6个百分点;而CA模块使模型增加0.5 MB,但也使F1值提高0.5个百分点;引入CBAM模块后,尽管模型增加0.4 MB,但各项指标均优于其他注意力机制,相比原模型,P、R、F1值、mAP<sub>0.5</sub>、mAP<sub>0.5:0.95</sub>分别提升1.3、0.1、0.7、0.5、1.4个百分点。综合对比模型各项性能的结果,选择融入CBAM注意力模块作为特征提取的优化方案更为合理。

## 2.3 不同检测模型综合对比试验

为了进一步验证本文算法YOLO v7-CSMRep的有效性,与当前主流的单阶段目标检测算法进行对比试验,包括YOLO v5s、YOLO v6s、YOLO v7、YOLO v7-tiny和YOLO v8s,不同模型综合对比结果如表3所示。

由表3可以明显看出,与YOLO v5s、YOLO v6s、YOLO v7、YOLO v7-tiny、YOLO v8s模型相比,本文方法F1值和mAP<sub>0.5</sub>均展现最佳性能,分别达到93.1%、97.1%。具体而言,与这些模型相比,F1值提升1.3、1.9、0.7、1.2、1.8个百分点,mAP<sub>0.5</sub>则提升1.0、2.0、0.7、0.8、1.2个百分点,原因为网络结构简化和特征提取能力增强。尽管模型内存占用量、参数量和计算量上不占优势,与YOLO v7-tiny相比,帧率下降15.5%,这是因为集成了CBAM注意力机制导致计算量略增,进而影响了帧率,但其帧率相比YOLO v5s、YOLO v6s、YOLO v7、YOLO v8s提升11.2%、19.2%、15.4%、39.4%。总体而言,本文方法在自建鲑鱼数据集上相比其他模型展示了

表 3 不同模型综合对比结果

Tab. 3 Comprehensive comparison results of different models

算法	F1 值/%	mAP <sub>0.5</sub> /%	模型内存占用量/MB	参数量	浮点计算量	帧率/(f·s <sup>-1</sup> )
YOLO v5s	91.8	96.1	14.3	7.010 × 10 <sup>6</sup>	1.580 × 10 <sup>10</sup>	126.6
YOLO v6s	91.2	95.1	36.76	1.850 × 10 <sup>7</sup>	4.520 × 10 <sup>10</sup>	118.1
YOLO v7	92.4	96.4	74.7	3.648 × 10 <sup>7</sup>	1.032 × 10 <sup>11</sup>	122.0
YOLO v7-tiny	91.9	96.3	12.2	6.010 × 10 <sup>6</sup>	1.300 × 10 <sup>10</sup>	166.7
YOLO v8s	91.3	95.9	6.2	3.010 × 10 <sup>6</sup>	8.100 × 10 <sup>9</sup>	101.0
YOLO v7-CSMRep	93.1	97.1	71.5	3.456 × 10 <sup>7</sup>	7.090 × 10 <sup>10</sup>	140.8

更优越的综合性能。

此外,虽然 YOLO v8s 的模型内存占用量相对较小,但其设计可能更侧重于减少参数和计算量,这在广泛的公共数据集测试中表现出色,展现了其优于 YOLO v7 的性能。然而,在自制数据集上,YOLO v8s 的 F1 值、mAP<sub>0.5</sub> 和帧率等性能指标均低于 YOLO v7 和本文方法。这一差异表明,尽管 YOLO v8s 在泛化能力上有显著优势,但针对特定任务,如鲑鱼图像目标检测,YOLO v7 原模型的复杂网络架构能更好地平

衡准确性与推理速度。本文方法在 YOLO v7 的基础上引入轻量化 Stem 模块、MRep 模块和 CBAM 模块,有效减少了模型计算量和内存占用量,同时提高了模型特征提取能力和处理图像速度。

## 2.4 检测结果对比分析

图 8 展示了原模型与改进模型在目标检测任务中的效果对比。图 8a~8d 表现了原模型中的漏检现象,其中目标因严重重叠未能被成功识别,相应的漏检区域通过椭圆形状进行了高亮标注。

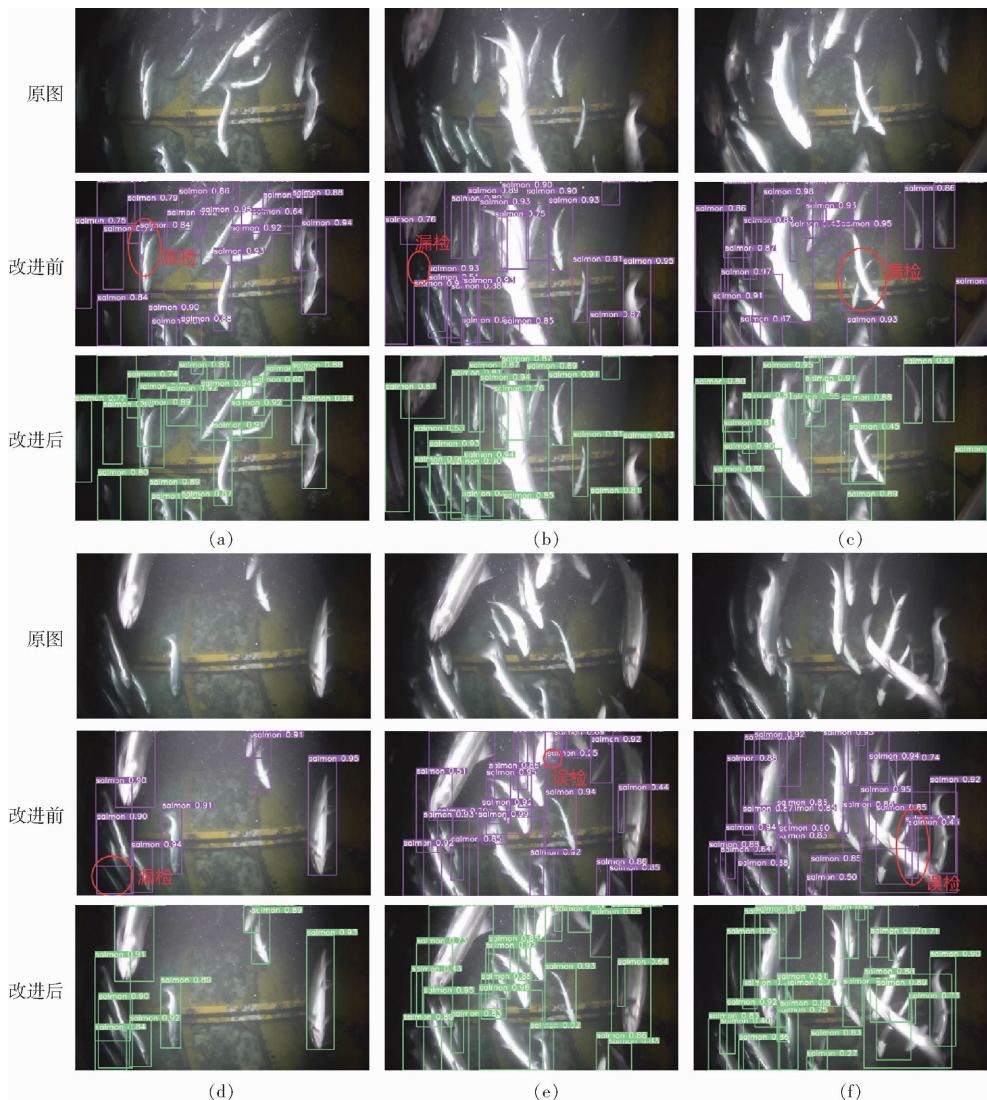


图 8 改进前后检测效果对比

Fig. 8 Comparison of detection effect before and after improvement

此外,图 8e、8f 展现了误检情况,即非目标对象被错误地识别为目标。改进后模型显著降低了这两类错误的发生,体现了引入的 Stem 模块在提升底层特征准确性方面的贡献,以及 MRep 和 CBAM 模块在增强特征提取能力上的作用。这 3 个模块综合作用不仅提升了目标检测准确率,而且有效减少了目标重叠导致的漏检问题,并且在一定程度上抑制了误检现象。

### 3 结论

(1) 提出了一种改进轻量化的 YOLO v7 - CSMRep 模型,以 YOLO v7 为基础模型。通过 Stem 模块将 Backbone 层的前 4 个卷积操作合并成 1 个高效的卷积操作。同时,将原网络的 ELAN 模块替换为 MRep 模块,以提高  $X$  和  $Y$  方向的特征提取能力,从而加快网络的特征提取。这两个模块在提高检测精度的同时降低了模型内存占用量和计算量,实现了模型轻量化。此外,通过在骨干网络末端集成 CBAM 模块,提高网络的特征提取能力,有效解

决了鲑鱼游动中相互遮挡的问题,使模型具备更强的抗干扰能力。

(2) 为了验证 YOLO v7 - CSMRep 模型的性能,设计了 6 组网络进行消融试验。包括 YOLO v7、引入轻量 Stem 模块的骨干网络、将 MRep 模块替换 ELAN 和 ELAN - H 模块的网络、骨干网络末端集成 CBAM 模块、同时集成 Stem 模块和 MRep 模块,以及本文算法。试验结果表明,改进后模型在检测精度方面优于其他 5 组模型,同时检测速度、浮点计算量、模型内存占用量等性能指标均优于原模型。

(3) 在相同试验条件下,将本文算法 YOLO v7 - CSMRep 与 YOLO v5s、YOLO v6s、YOLO v7、YOLO v7 - tiny、YOLO v8s 模型进行比较。试验结果表明,本文算法在自制鲑鱼数据集上取得了较优效果,mAP<sub>0.5</sub>分别提升 1.0、2.0、0.7、0.8、1.2 个百分点,达到 97.1%,且帧率也表现出色,达到 140.8 f/s。综上,YOLO v7 - CSMRep 模型在兼顾精度和实时性方面表现出色,满足快速、准确地对深远海养殖水下鲑鱼图像进行检测的需求。

### 参 考 文 献

- [1] 徐乐俊,赵蕾,李雪,等.从中央一号文件看渔业强国建设基本思路——2023 年中央一号文件解读兼谈渔业发展战略[J].中国渔业经济,2023,41(2):1-13.  
XU Lejun, ZHAO Lei, LI Xue, et al. The basic thoughts of building a powerful fishery country from the perspective of No. 1 Central Document; Interpretation of the No. 1 Central Document for 2023 and the fishery development strategy [J]. Chinese Fisheries Economics, 2023, 41(2): 1-13. (in Chinese)
- [2] 罗逸豪,刘奇佩,张吟,等.基于深度学习的水下图像目标检测综述[J].电子与信息学报,2023,45(10):3468-3482.  
LUO Yihao, LIU Qipei, ZHANG Yin, et al. Review of underwater image object detection based on deep learning [J]. Journal of Electronics & Information Technology, 2023, 45(10): 3468-3482. (in Chinese)
- [3] 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.
- [4] GIRSHICK R. Fast R - CNN [C] // Proceedings of the IEEE International Conference on Computer Vision, 2015: 1440-1448.
- [5] REN S, HE K, GIRSHICK R, et al. Faster R - CNN: towards real-time object detection with region proposal networks [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2016, 39(6): 1137-1149.
- [6] HE K, GKIOXARI G, DOLLAR P, et al. Mask R - CNN [C] // Proceedings of the IEEE International Conference on Computer Vision, 2017: 2961-2969.
- [7] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot multibox detector [C] // European Computer Vision - ECCV 2016: 14th European Conference. Amsterdam: Springer International Publishing, 2016: 21-37.
- [8] 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. IEEE, 2016: 779-788.
- [9] REDMON J, FARHADI A. YOLO9000: better, faster, stronger [C] // IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2017: 7263-7271.
- [10] REDMON J, FARHADI A. YOLO v3: an incremental improvement [C] // Computer Vision and Pattern Recognition, 2018: 1-6.
- [11] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. YOLO v4: optimal speed and accuracy of object detection [C] // IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2020: 1-17.
- [12] SUN Z, CHEN B. Research on pedestrian detection and recognition based on improved YOLO v6 algorithm [C] // International Conference on Artificial Intelligence in China, 2022: 281-289.
- [13] SOEB M J A, JUBAYER M F, TARIN T A, et al. Tea leaf disease detection and identification based on YOLO v7 (YOLO - T) [J]. Scientific Reports, 2023, 13(1): 6078.
- [14] 武星,齐泽宇,王龙军,等.基于轻量化 YOLO v3 卷积神经网络的苹果检测方法[J].农业机械学报,2020,51(8):17-25.  
WU Xing, QI Zeyu, WANG Longjun, et al. Apple detection method based on light - YOLO v3 convolutional neural network [J]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(8): 17-25. (in Chinese)
- [15] MENG F, CHENG H, LI K, et al. Pruning filter in filter [J]. Advances in Neural Information Processing Systems, 2020, 33: 17629-17640.
- [16] 方鹏,郝宏运,王红英.基于知识蒸馏的叠层笼养蛋鸡行为识别模型研究[J].农业机械学报,2021,52(10):300-306.

- FANG Peng, HAO Hongyun, WANG Hongying. Behavior recognition model of stacked-cage layers based on knowledge distillation[J]. Transactions of the Chinese Society for Agricultural Machinery,2021,52(10):300-306. (in Chinese)
- [17] TAN K, WANG D. Compressing deep neural networks for efficient speech enhancement[C]//IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP),2021:8358-8362.
- [18] HOWARD A G, ZHU M, CHEN B, et al. Mobilenets: efficient convolutional neural networks for mobile vision applications [J]. arXiv Preprint, arXiv:1704.04861,2017.
- [19] SANDLER M, HOWARD A, ZHU M, et al. Mobilenetv2: inverted residuals and linear bottlenecks[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition,2018:4510-4520.
- [20] HOWARD A, SANDLER M, CHU G, et al. Searching for Mobilenetv3 [C]//Proceedings of the IEEE/CVF International Conference on Computer Vision,2019:1314-1324.
- [21] FRAN C. Deep learning with depth wise separable convolutions [C]//IEEE Conference on Computer Vision and Pattern Recognition (CVPR),2017.
- [22] ZHANG X, ZHOU X, LIN M, et al. Shufflenet: an extremely efficient convolutional neural network for mobile devices [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition,2018:6848-6856.
- [23] HAN K, WANG Y, TIAN Q, et al. Ghostnet: more features from cheap operations [C]//Proceedings of the IEEE/CVF International Conference on Computer Vision and Pattern Recognition,2020:1580-1589.
- [24] TAN M, LE Q. Efficientnet: rethinking model scaling for convolutional neural networks [C]//International Conference on Machine Learning. PMLR, 2019: 6105-6114.
- [25] 罗志聪,李鹏博,宋飞宇,等. 嵌入式设备的轻量化百香果检测模型[J]. 农业机械学报,2022,53(11):262-269,322.  
LUO Zhicong, LI Pengbo, SONG Feiyu, et al. Lightweight passion fruit detection model based on embedded device [J]. Transactions of the Chinese Society for Agricultural Machinery,2022,53(11):262-269,322. (in Chinese)
- [26] 苗荣慧,李志伟,武锦龙. 基于改进 YOLO v7 的轻量化樱桃番茄成熟度检测方法[J]. 农业机械学报,2023,54(10):225-233.  
MIAO Ronghui, LI Zhiwei, WU Jinlong. Lightweight maturity detection of cherry tomato based on improved YOLO v7 [J]. Transactions of the Chinese Society for Agricultural Machinery,2023,54(10):225-233. (in Chinese)
- [27] 李丽,卢世博,任浩,等. 基于改进 YOLO v5 的复杂环境下桑树枝干识别定位方法[J]. 农业机械学报,2024,55(2):249-257.  
LI Li, LU Shibao, REN Hao, et al. Mulberry branch identification and location method based on improved YOLO v5 in complex environment [J]. Transactions of the Chinese Society for Agricultural Machinery,2024,55(2):249-257. (in Chinese)
- [28] 李健源,柳春娜,卢晓春,等. 基于改进 YOLO v5s 和 TensorRT 部署的鱼道过鱼监测 [J]. 农业机械学报,2022,53(12):314-322.  
LI Jianyuan, LIU Chunna, LU Xiaochun, et al. Fish passage monitoring based on improved YOLO v5s and TensorRT deployment [J]. Transactions of the Chinese Society for Agricultural Machinery,2022,53(12):314-322. (in Chinese)
- [29] 覃学标,黄冬梅,宋巍,等. 基于目标检测及边缘支持的鱼类图像分割方法[J]. 农业机械学报,2023,54(1):280-286.  
QIN Xuebiao, HUANG Dongmei, SONG Wei, et al. Fish image segmentation method based on object detection and edge support [J]. Transactions of the Chinese Society for Agricultural Machinery,2023,54(1):280-286. (in Chinese)
- [30] 李道亮,姜国旗,杨建安,等. 基于计算机视觉的鱼类低氧胁迫行为检测与跟踪算法 [J]. 农业机械学报,2023,54(10):399-406.  
LI Daoliang, JIANG Guoqi, YANG Jian'an, et al. Detection and tracking algorithm of fish hypoxia stress behavior based on computer vision [J]. Transactions of the Chinese Society for Agricultural Machinery,2023,54(10):399-406. (in Chinese)
- [31] SUNG M, YU S C, GIRDHAR Y. Vision based real-time fish detection using convolutional neural network [C]//OCEANS 2017-Aberdeen. IEEE,2017:1-6.
- [32] AI MUKSIT A, HASAN F, EMON M F H B, et al. YOLO - Fish: a robust fish detection model to detect fish in realistic underwater environment [J]. Ecological Informatics,2022,72:101847.
- [33] 李庆忠,李宜兵,牛炯. 基于改进 YOLO 和迁移学习的水下鱼类目标实时检测 [J]. 模式识别与人工智能,2019,32(3):193-203.  
LI Qingzhong, LI Yibing, NIU Jiong. Real-time detection of underwater fish based on improved YOLO and transfer learning [J]. Pattern Recognition and Artificial Intelligence,2019,32(3):193-203. (in Chinese)
- [34] 汪沛洁,湛雨章,王诗琦,等. 基于改进 YOLO v5 的鱼群小目标检测优化算法 [J]. 湖北大学学报(自然科学版),2024,46(1):14-24.  
WANG Peijie, CHEN Yuzhang, WANG Shiqi, et al. Optimization algorithm of fish shoal small target detection based on improved YOLO v5 [J]. Journal of Hubei University (Natural Science),2024,46(1):14-24. (in Chinese)
- [35] 涂万,于红,张鹏,等. 基于通道非降维与空间协调注意力的改进 YOLO v8 养殖鱼群检测 [J]. 大连海洋大学学报,2023,38(4):717-725.  
TU Wan, YU Hong, ZHANG Peng, et al. Farmed fish detection by improved YOLO v8 based on channel non-degradation with spatially coordinated attention [J]. Journal of Dalian Ocean University,2023,38(4):717-725. (in Chinese)
- [36] DING X, ZHANG X, MA N, et al. Repvgg: making vgg-style convnets great again [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition,2021:13733-13742.
- [37] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition [J]. arXiv Preprint, arXiv:1409.1556,2014.