

DOI:10.19651/j.cnki.emt.2519020

基于动态自适应特征学习的水下鱼类识别方法^{*}

朱晓龙^{1,2} 李超凡² 陈郁隈² 陈祥子^{1,2} 褚文敬²

(1.海南热带海洋学院崖州湾创新研究院 三亚 572022; 2.海南热带海洋学院海洋科学技术学院 三亚 572022)

摘要: 针对水下复杂环境中鱼类识别易出现误检与漏检的问题,提出了一种基于 YOLOv8n 的动态自适应特征学习的识别方法 FD-YOLO。首先,在主干中设计将多尺度并行卷积与具备区域自适应调节的 RFA 结合,构建 MRFA 模块增强对鱼类细微特征差异的学习能力。其次,在颈部中提出将双通道融合结构与 CARAFE 结合,构建 ECARU 模块作为上采样模块动态调整特征权重,提升鱼类局部细节重构。最后,为抑制样本分布不均导致的识别偏差,引入具有动态调节因子的 Varifocal Loss 作为损失函数,提高对水下鱼类位置的判定能力。实验结果表明,与 YOLOv8n 相比,FD-YOLO 的精确率、召回率、mAP50-95 分别提升了 3.6%、4.9% 和 3.7%,参数量和计算量分别降低至 2.5 MB 和 6.8 GFLOPs。该研究结果能够为水下目标自动检测和监测提供技术支持和参考。

关键词: 深度学习;水下图像;鱼类识别;动态自适应;YOLOv8n

中图分类号: TN919.8;TP391 **文献标识码:** A **国家标准学科分类代码:** 510.10

Underwater fish recognition method based on dynamic adaptive feature learning

Zhu Xiaolong^{1,2} Li Chaofan² Chen Yuwei² Chen Xiangzi^{1,2} Chu Wenjing²

(1. Yazhou Bay Innovation Institute, Hainan Tropical Ocean University, Sanya 572022, China;

2. College of Marine Science and Technology, Hainan Tropical Ocean University, Sanya 572022, China)

Abstract: To address the issues of false detections and missed detections in underwater fish recognition under complex environmental conditions, this paper proposes FD-YOLO, a recognition method based on YOLOv8 incorporating dynamic adaptive feature learning. First, in the backbone, we design a MRFA module that combines parallel multi-scale convolution with RFA to enhance the network's ability to capture fine-grained differences in fish features. Second, in the neck, we introduce the ECARU module, which integrates a dual-channel fusion structure with the CARAFE mechanism. This upsampling module adaptively reweights features to improve the reconstruction of local fish details. Finally, to mitigate recognition bias caused by class imbalance, we adopt Varifocal Loss, which includes a dynamic adjustment factor, to improve the accuracy of underwater fish localization. Experimental results demonstrate that, compared with YOLOv8n, the proposed FD-YOLO achieves improvements of 3.6%, 4.9% and 3.7% in Precision, Recall and mAP50-95, respectively. Moreover, the parameter size and computational cost are reduced to 2.5 MB and 6.8 GFLOPs. These findings demonstrate that the proposed method can provide technical support and reference for automated detection and monitoring of underwater targets.

Keywords: deep learning; underwater image; fish recognition; dynamic adaptation; YOLOv8n

0 引言

精准识别水下鱼类对于获取其种类、数量及形态等信息具有重要意义,可为水域生态状况评估与渔业资源监测

提供直观依据^[1-3]。然而,水下常伴有复杂光照和环境干扰的问题,导致采集图像质量较低,进而影响识别效果。此外,图像在类别分布上的不均衡性也加剧了识别任务的难度^[4]。近年来,随着计算机视觉技术与图像处理算法的持

收稿日期:2025-06-05

^{*} 基金项目:国家自然科学基金(42266005)、海南省院士创新平台科研项目(YSP TZX202507)、海南省科技计划三亚崖州湾科技城科技创新联合项目(2021CX LH0002)、三亚市科技创新专项项目(2022KJ CX83)、海南热带海洋学院崖州湾创新研究院重大科技计划项目(2022CX YZD003、2023CX YZD001)资助

续发展,基于深度卷积神经网络的鱼类识别方法逐步取代了传统人工特征提取方式^[5-8]。

Monkman等^[9]提出采用NASNet, ResNet-101和MobileNet三种R-CNN模型识别图像中的鱼类,并提取鱼的特征点用于判定鱼体轮廓和位置,但该方法仅适用于背景单一的场景,且计算资源需求较大。Cai等^[10]采用深度可分离卷积进行特征提取,并将MobileNetV1用作YOLO的主干网络,构建轻量级识别模型实现了水下鱼类检测。但该模型在小目标检测及复杂环境下的鲁棒性仍显不足。袁红春等^[11]采用多尺度Retinex与颜色恢复MSRCR方法增强水下图像质量。以VGG16和ResNet-101作为特征提取网络,并基于Faster R-CNN完成鱼类识别任务,但整体模型复杂度较高。Zhao等^[12]在残差网络基础上构建复合主干网络CBResNet,并结合增强路径聚合结构EPANet,实现鱼类的检测与定位,但其在应对复杂水下场景时识别性能仍有限。Liu等^[13]提出一种基于YOLOv4的双分支深度卷积神经网络,用于海洋牧场中多类鱼群识别。尽管该方法鱼类统计准确率达到95.6%,但对小目标与遮挡场景的识别精度仍待提升。Li等^[14]针对池塘鱼类数量的检测问题,提出将可变形卷积模(deformable convolution module, DCM)与自适应阈值模块(adaptive threshold module, ATM)集成至YOLOv5。但该方法在密集鱼群与小体型鱼类检测中效果有限。刘雨青等^[15]在CenterNet基础上引入特征融合模块,实现不同尺度鱼类特征的融合。该模型在背景较为简单、目标尺寸较大的场景中平均精度均值(mean average precision, mAP)达到了85.45%,但在复杂水下环境中的识别能力仍较为有限。王伟芳等^[16]在YOLOv5框架下引入MobileNet作为主干网络,Leaky ReLU作为激活函数,提升了鱼类个体识别能力。尽管模型的mAP为92.46%,但仍主要局限于单一鱼类与简单背景场景。吕俊霖等^[17]采用Ghost模块、GELU激活函数及标签平滑损失等技术,构建多阶段鱼类识别模型。尽管提升了识别效果,但模型整体参数量与计算开销较大,限制了其应用场景。

综上所述,现有改进算法在简单背景下的鱼类个体识别任务中已展现出较高的准确性,但在面对水下复杂因素干扰时,仍存在识别精度不足、对小目标适应性差等问题。为此,本文选取兼具检测效率与精度优势的YOLOv8n作为基础架构,提出一种适用于复杂水下场景的鱼类识别方法FD-YOLO,主要工作如下:

1)在主干中构建MRFA模块增强对鱼类细小特征的感知与表达能力。MRFA模块融合了多尺度并行卷积与具备区域自适应调节的感受野注意力(receptive field attention, RFA),其侧重于感受野内部的空间特征,能够更精准识别和处理局部区域,提高模型对细粒度鱼类信息的提取精度。

2)在颈部中设计ECARU上采样模块,通过将双通道

融合结构与具备内容感知重组能力的模块(content-aware reassembly of features, CARAFE)结合,使ECARU模块能够自适应的生成内核,有效保留边缘与细节信息,从而提升对模糊或小鱼目标的识别能力。

3)采用Varifocal Loss作为损失函数,通过动态调整权重,抑制样本不均对模型训练的影响,从而提升水下鱼类的分类准确性与空间定位性能。

1 YOLOv8n

本文选用YOLOv8n作为基础模型,它是由Ultralytics公司于2023年发布的最新版目标检测算法,其网络结构延续了主干网络、颈部网络与检测头三部分组成形式^[18],整体结构如图1所示。在主干网络中,YOLOv8n集成了普通卷积Conv模块、C2f模块以及SPPF模块^[19-20],并通过引入梯度并联连接与残差连接机制增强特征的融合与传递能力。YOLOv8n的颈部网络将C2f模块与主干网络生成的特征图进行上、下采样操作,并以级联特征融合策略增强对特征的表征能力^[21]。在检测头部分,YOLOv8n综合使用CIoU边界框损失、Focal Loss分类损失与二元交叉熵(binary cross-entropy, BCE)置信度损失,以共同优化目标在图像中的类别判定与定位能力。

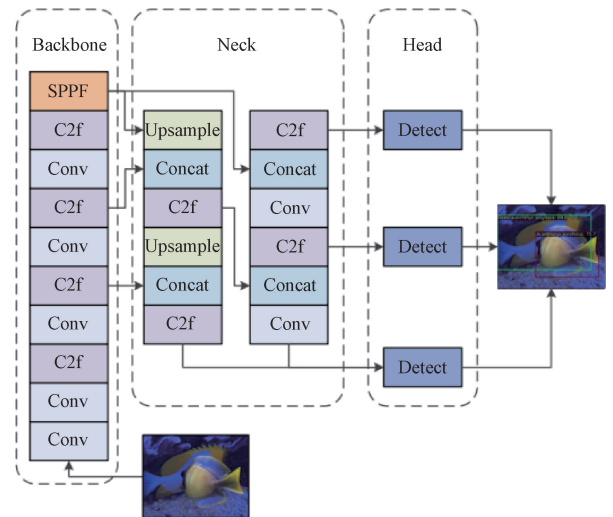


图1 YOLOv8n的结构图

Fig. 1 Structure diagram of YOLOv8n

2 改进的FD-YOLO

2.1 MRFA模块

为了增强对水下图像中鱼类特征提取能力,提出构建MRFA模块用于特征提取。如图2所示,该模块首先对输入特征图采用多尺度并行卷积操作,以提取不同感受野范围内的上下文信息,增强局部与全局特征融合^[22]。然后,将输出的特征图传入RFA。RFA能够为卷积核动态生成权重,促使卷积核更加聚焦在不同鱼类的细微特征差异。

同时, RFA 能够根据不同的感受野区域自适应调整参数, 允许模型根据不同目标信息提供更合适的学习策略^[23], 从提升模型对水下复杂环境中鱼类的判别能力。最后, 加权后的特征图通过卷积操作整合并输出。

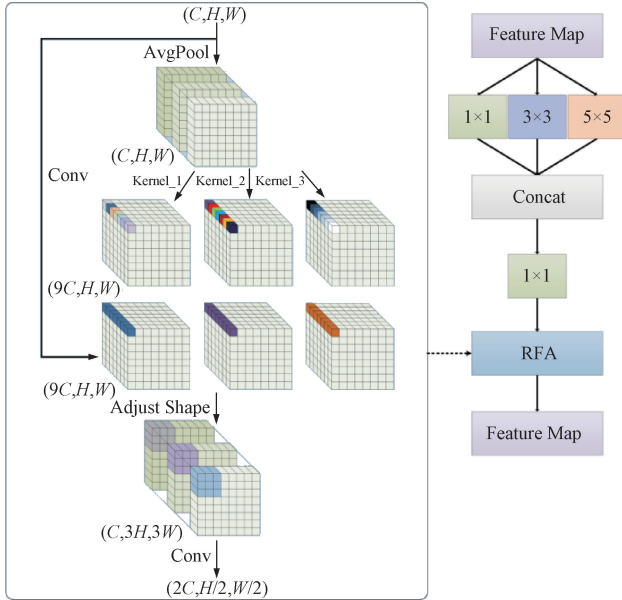


图 2 MRFA 模块的结构图

Fig. 2 Structure diagram of MRFA module

假设输入鱼类特征图为 (C, H, W) , 其中 C, H 和 W 分别表示特征图的通道数、高度和宽度。首先该特征图通过卷积核为 $1 \times 1, 3 \times 3$ 和 5×5 的并行卷积操作, 在保持特征尺寸不变的情况下, 通过通道调整使输出特征图为 (C, H, W) 。在 RFA 中, 通过平均池化操作 (AvgPool) 减少特征图冗余计算, 并且采用分组的卷积核 (Kernel_1, Kernel_2, Kernel_3) 进行动态特征提取, 生成尺寸为 $(9C, H, W)$ 的注意力特征图。同时, 原特征图采用卷积操作

(Conv) 增强空间特征表示能力, 输出 $(9C, H, W)$ 的空间特征图。在此基础上, 通过逐通道加权重构特征图, 以增强各通道在空间上的作用。并且对特征图形状调整形成 $(C, 3H, 3W)$ 的特征图。最后再次通过卷积操作 (Conv) 将特征图调整为 $(2C, H/2, W/2)$, 完成特征图的空间压缩和语义增强。MRFA 模块的算法如下:

$$F_k = W_k * X_k \in \{1, 3, 5\} \quad (1)$$

$$F_g = Up(AvgPool(X)) \quad (2)$$

$$F = [F_g; F_1; F_3; F_5] \quad (3)$$

$$a = softmax(W_2 * ReLU(W_1 * AvgPool(F))) \quad (4)$$

$$Y = reshape(a_g F_g + a_1 F + a_2 F_3 + a_3 F_5) \quad (5)$$

其中, X 和 Y 分别表示输入和输出的特征图; W_k 表示分组卷积核尺寸为 $k \times k$; F_k 和 F_g 分别表示不同感受野的卷积特征和全局上下文特征; Up 表示双线性上采样; F 表示多分支特征拼接结果; W_1 和 W_2 分别表示注意力全连接权重; a 表示分支的自适应权重。

2.2 ECARU 模块

为了提高对空间结构与细节信息的学习能力, 进而提升对模糊或较小的鱼类的识别能力, 提出构建 ECARU 模块作为上采样结构。如图 3 所示, 首先设计双通道融合结构对输入特征图进行处理。一方面, 采用 1×1 卷积实现线性映射以压缩通道维度; 另一方面, 引入轻量级的压缩与激励机制 (squeeze-and-excitation, SE)^[24] 增强通道维度上的特征表达能力。两个分支的输出通过拼接进行融合, 并将其作为 CARAFE 的输入特征图。在上采样过程中, CARAFE 能够利用内容感知机制, 依据特征图的局部上下文信息动态地重构特征^[25], 进而更好地保留和恢复鱼类的结构和细节。此外, 通过利用可学习的权重分配机制, 能够更准确地重构目标的轮廓和特征^[26]。相比之下, 传统上采样算法无法充分考虑特征图的特点, 从而导致细节丢失或重构效果不佳。

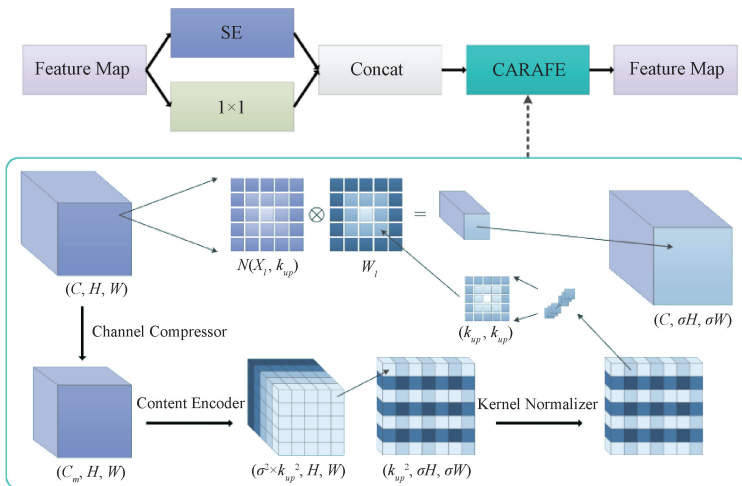


图 3 ECARU 的结构图

Fig. 3 Structure diagram of ECARU

假设输入的水下鱼类特征图 (C, H, W) ,上采样率为 σ ,经过预测和重构两个流程后,CARAFE最终生成的特征图 $(C, \sigma H, \sigma W)$ 。在预测流程中,输入的特征图先被压缩通道降低参数量至 C_m ,然后通过引入编码参数 $k_{encoder}$ 获取更加丰富的感受野,形成 $(k_{up}^2, \sigma H, \sigma W)$ 的特征图,并将它们归一化处理。在重构流程中,采用加权求和的方法将输入特征图与归一化后的特征图融合,确保每个像素在上采样区域中均有贡献,且重构后的特征信息依赖于内容信息而非距离信息,从而使得重构后的特征图具备更丰富的局部细节和语义表达。 $N(X_l, k)$ 表示以输入特征图上某点 l 为中心的 $k \times k$ 子区域,输出特征图对应位置 l' 的重构采样核和重构领域分别为 $W_{l'}$ 和 $X_{l'}$ 。ECARU模块的主要算法如下:

$$X_{SE} = \sigma(W_2 \delta(W_1(GAP(X)))) \quad (6)$$

$$F = [X_{SE}; X_1] \quad (7)$$

$$K_{raw} = Conv(W_c * F) \quad (8)$$

$$K_{norm}(h, w) = softmax(K_{raw}(:, h, w)) \quad (9)$$

$$W_{l'} = \varphi(N(X_l, k_{encoder})) \quad (10)$$

$$X_{l'} = \phi(N(X_l, k_{up}), W_{l'}) \quad (11)$$

其中, X_{SE} 表示SE加权后特征; GAP 为全局平均池化; F 表示通道拼接特征; K_{raw} 和 K_{norm} 分别表示未归一化和归一化权重向量; φ 和 ϕ 分别表示预测模块和重构模块。

2.3 Varifocal Loss

为了进一步提升水下鱼类分类的精度,提出采样Varifocal Loss作为损失函数。该算法通过引入动态加权机制,对正负样本赋予不同的权重,降低正负样本不均对鱼类判别的影响。即避免训练过程中因负样本较多产生的偏差,促使模型更关注正样本的特征信息^[27]。同时,Varifocal Loss对IoU较高的样本赋予更高的权重,这使得模型在训练过程中能够更关注那些位置准确的预测框,从而提升鱼类目标定位的精确度。VariFocal Loss可以表示为:

$$VFL(p, q) = \begin{cases} -q(q \log(p) + (1-q) \log(1-p)), & q > 0 \\ -\alpha p^\gamma \log(1-p), & q = 0 \end{cases} \quad (12)$$

其中, p 表示鱼类目标的预测值, q 表示鱼类目标的IoU分数。 α 表示可调比例因子, γ 表示调制因子。

2.4 FD-YOLO识别网络

改进后的识别模型FD-YOLO的结构如图4所示。该模型在YOLOv8n框架基础上进行优化,分别在主干网络、颈部网络和检测头中引入针对性模块以提升识别性能。在主干网络部分,保留前两个Conv模块以完成鱼类特征的粗特征提取,将后续3个Conv模块替换为所提出的MRFA模块。MRFA模块通过引入具备动态调节能力的感受野机制,能够更有效地感知和学习多尺度鱼类目标的特征信息,从而显著提升模型在多类别与多尺寸鱼类识

别任务中的表现。在颈部结构中,采用ECARU模块替代传统上采样操作。该模块融合双通道融合结构与内容感知重组机制,能够根据上下文信息自适应生成重构核,从而更精确地还原细节与边缘信息,有效提升对模糊及小目标鱼类的识别能力。此外,在检测头中引入Varifocal Loss作为损失函数,以动态加权方式联合优化分类置信度与边界框定位精度。该损失函数可有效缓解由于样本不平衡带来的训练偏差,增强模型在复杂水下环境中对小目标鱼类的鲁棒性与检测精度。

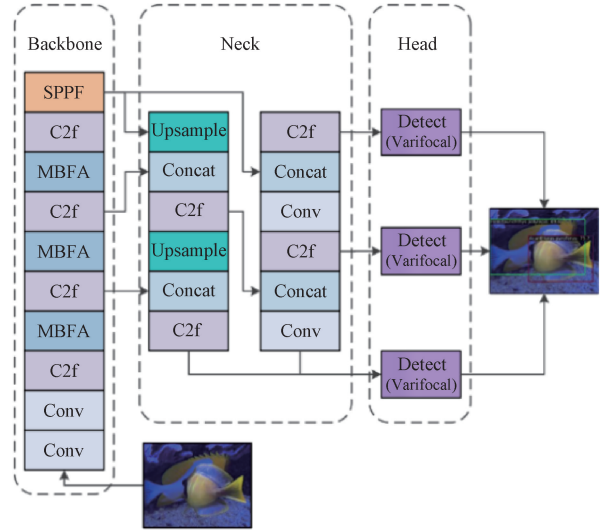


图4 FD-YOLO的结构图

Fig. 4 Structure diagram of FD-YOLO

3 实验与分析

3.1 数据集

本实验采用的水下鱼类数据集为TF-DET^[28],涵盖了10种典型的海洋热带鱼类。该数据集中的所有图像均采集于水下真实环境,涵盖了光照不均、背景杂乱及相互遮挡等干扰因素。鱼类目标的尺寸划分标准依据COCO数据集的定义,目标像素面积被分为小尺寸目标(小于 $32 \text{ pixel} \times 32 \text{ pixel}$)、中尺寸目标($32 \text{ pixel} \times 32 \text{ pixel} \sim 96 \text{ pixel} \times 96 \text{ pixel}$)和大尺寸目标(大于 $96 \text{ pixel} \times 96 \text{ pixel}$)。为测试模型对小尺寸、多类别目标的鲁棒性,引入随机旋转、图像压缩等数据增强策略,以提升数据的多样性与训练难度。最终,数据集扩展至4500张水下图像,并按照8:2的比例随机划分为训练集与测试集。图5为数据集中特征相关性的成对图可视化,用以直观揭示数据结构特性及目标框的位置与尺寸的分布规律。从图5中可知,鱼类目标的中心位置主要分布于图像中部区域,呈现出近似正态分布的特征。目标框的宽度与高度分布整体偏小,也为后续模型在小目标检测方面提出了更高的识别要求。

此外,各类鱼在不同尺寸区间内的数量分布如表1所示,各种鱼类的样本数量存在明显差异,并且相同种类在

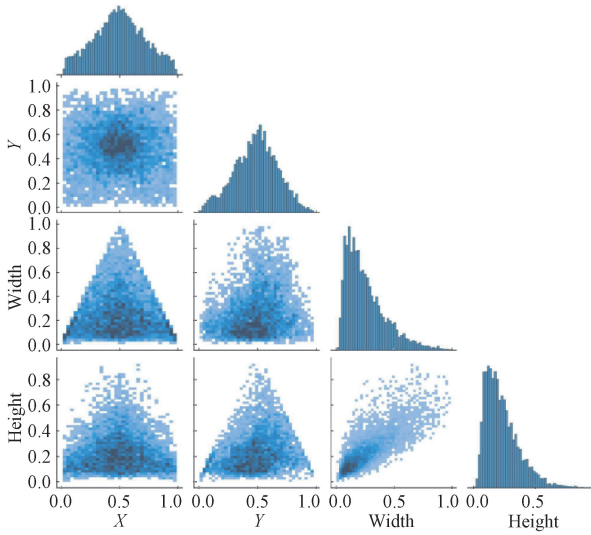


图 5 数据集中特征相关性的成对图可视化

Fig. 5 Pair plot visualization of feature correlations in the dataset

不同尺寸上的样本数量也不均衡。这种类别与尺寸分布的不平衡性,增加了模型训练与识别的难度。

表 1 数据集中不同尺寸和类别的鱼数量

Table 1 The number of fish of different sizes and categories in the dataset

Number	Name	Small	Medium	Large
Fish1	Acanthurus pyroferus	16	42	747
Fish2	Amphiprion ocellaris	21	70	371
Fish3	Amphiprion pereula	14	75	523
Fish4	Chromis viridis	42	55	1166
Fish5	Chrysiptera taupou	10	21	317
Fish6	Dascyllus aruanus	11	97	582
Fish7	Hemitaurichthys polylepis	7	41	275
Fish8	Paracanthurus hepatus	45	71	837
Fish9	Premnas biaculeatus	5	40	230
Fish10	Zebrasoma flavescens	23	87	351

3.2 环境配置

本实验的硬件配置: 2.4 GHz Intel Xeon E5-2680 v4 CPU, 12 GB NVIDIA GeForce RTX 3060 GPU, 32 GB RAM; 软件配置: PyTorch v2.0.1, Python 3.10.12, PyCharm Community 2023, anaconda-navigator v2.4.0 等。批量训练系数为 16, 动量系数为 0.937, 迭代周期系数为 300, 学习率初始系数为 0.01 权重衰减系数为 0.0005。

3.3 评价指标

识别准确度和模型复杂度是衡量识别模型优劣最常用的评价指标。其中, 识别准确度可由精确率 (precision, P)、召回率 (recall, R)、平均精度 (average precision, AP) 和 mAP 评价。上述指标的表达公式如下:

$$P = \frac{TP}{TP + FP} \quad (13)$$

$$R = \frac{TP}{TP + FN} \quad (14)$$

$$AP = \int_0^1 PR dR \quad (15)$$

$$mAP = \frac{1}{n} \sum_{i=1}^n \int_0^1 P(R) dR \quad (16)$$

其中, TP 、 FP 、 FN 分别表示正确识别目标、错误将非目标判断为目标、错误将目标判断为非目标。此外, 模型复杂度可由参数量 (Params) 和浮点运算量 (floating point operations, FLOPs) 评价, 二者数值越小表示模型复杂度越低。

3.4 改进模块的对比实验

为了验证每个改进模块的有效性, 本研究以 YOLOv8n 为基础模型, YOLO-MRFA、YOLO-ECARU 和 YOLO-Varifocal 分别表示引入 MRFA 模块、ECARU 模块和 Varifocal Loss 的改进模型, FD-YOLO 为本文提出的最终模型。图 6 展示了改进前后模型平均精度分布的箱线图, 直观反映了各模型识别性能的集中趋势与离散程度。图中红色虚线表示中位数, 反映当前模型识别数据的中心趋势; 黄色实线箱体的长度代表数据的离散性, 反映模型对不同鱼类识别结果的稳定性; 离群点以独立圆点表示, 反应模型对某些鱼类的识别精确性偏低。YOLO-MRFA 与 YOLO-ECARU 的中位数较高, 而 FD-YOLO 整体表现最佳。FD-YOLO 中位数最高达到 85.5%, 比 YOLOv8n 提升了 2.3%。尽管所有模型均有离群点, 但 YOLOv8n 的离群点最低, 仅为 60.4%。

图 7 展示了改进前后模型的归一化混淆矩阵。其中, 图 7(a) 为 YOLOv8n 的混淆矩阵, 图 7(b)、(c) 和 (d) 分别为改进后 YOLO-MRFA、YOLO-ECARU 和 YOLO-Varifocal 的混淆矩阵。由此可知, 除个别鱼类目标外, 改进后的模型在混淆矩阵对角线位置的归一化数值均高于原始 YOLOv8n。这表明引入的 MRFA 模块、ECARU 模块和 Varifocal Loss 能够有效增强模型对水下各类鱼种的判别和分类准确性。综上所述, 引入改进模块的模型在识别精度与泛化能力均表现出更优的性能。

3.5 消融实验

识别模型的消融实验如表 2 所示, 评价指标从准确性、存储需求和计算成本等方面评估。从整体上看, 各改进模型均实现不同程度的提升。FD-YOLO 的 Precision 虽未达到最高, 但也比基础模型提高了 3.6%。在 Recall 和 mAP50-95 方面, FD-YOLO 表现最优, 分别为 79.4% 和 82.7%, 表明动态特征聚合和增强感受野机制在提升检测准确性和完整性方面具有显著效果, 尤其在应对水下复杂环境中鱼类目标的检测中展现出较强的鲁棒性。在计算资源方面, FD-YOLO 的参数量与运算量相较于 YOLOv8n 分别减少了 0.5 MB 和 1.3 G。

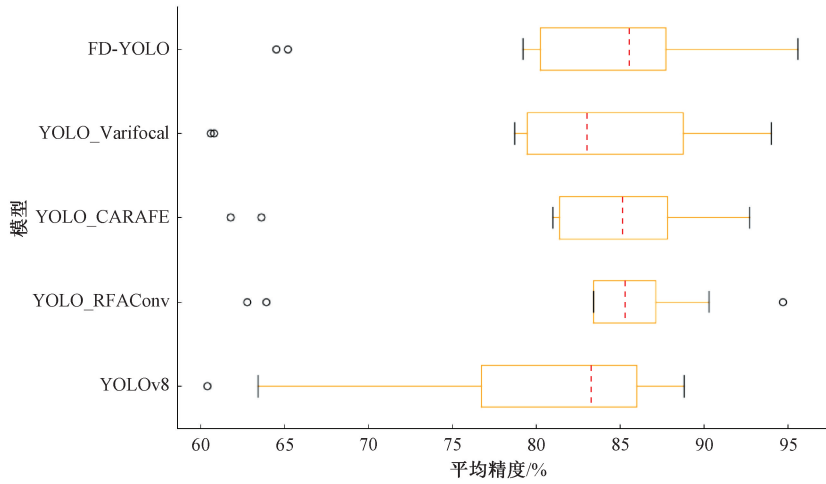


图 6 改进前后平均精度分布箱线图

Fig. 6 Boxplot of AP distribution before and after improvement

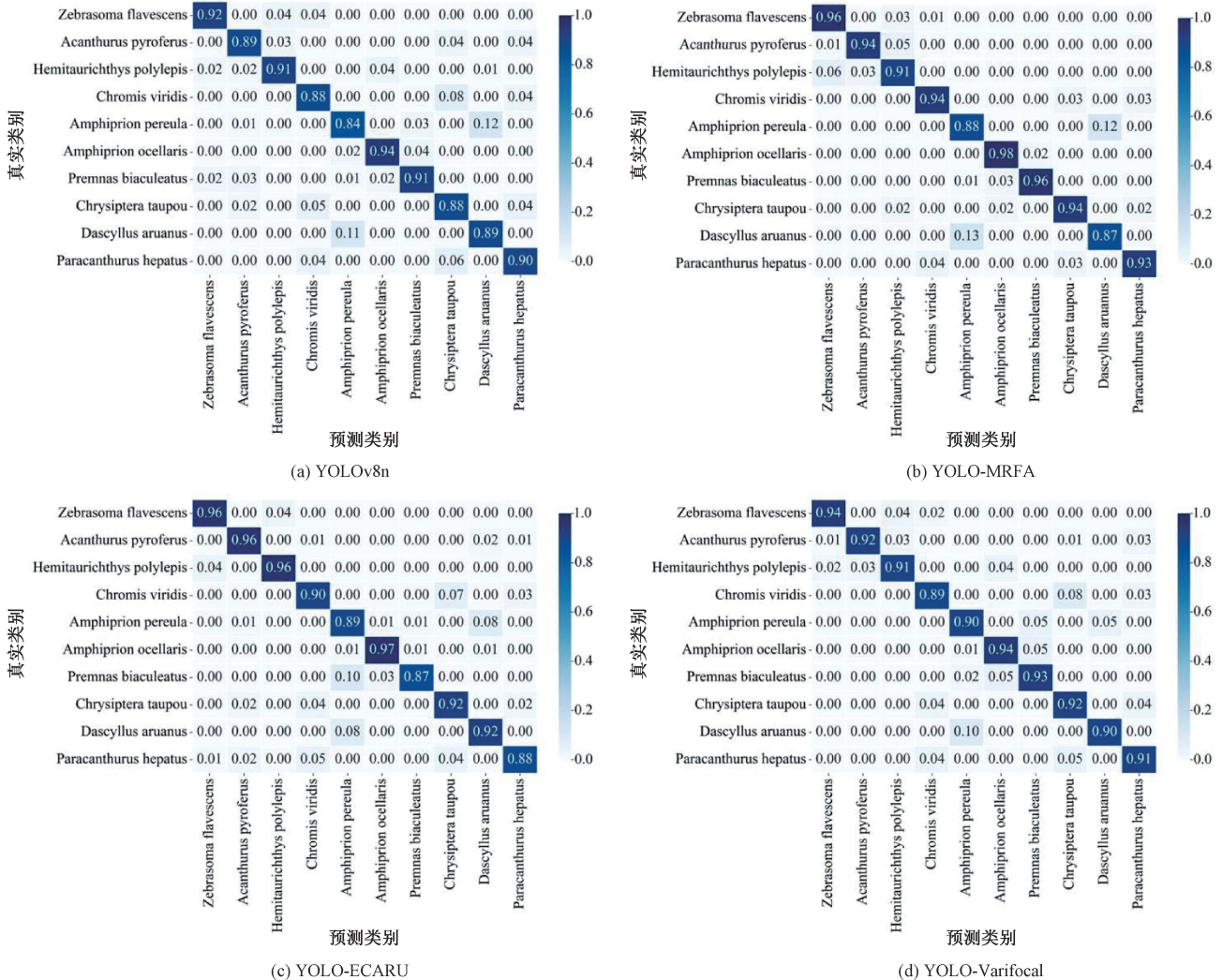


图 7 归一化混淆矩阵

Fig. 7 Normalized confusion matrix

表 2 消融实验对比

Table 2 Ablation experiment comparison

YOLOv8n	MRFA	ECARU	Varifocal	Precision/%	Recall/%	mAP50-95/%	Params/MB	FLOPs/G
✓				84.6	74.5	79.0	3.0	8.1
✓	✓			88.3	77.3	81.6	2.8	7.3
✓		✓		87.3	77.6	81.2	2.6	7.2
✓			✓	86.5	75.2	80.1	3.0	7.8
✓	✓	✓		88.4	78.6	82.5	2.4	6.8
✓	✓		✓	87.8	78.4	82.3	2.8	7.1
✓		✓	✓	87.3	77.9	82.2	2.6	6.8
✓	✓	✓	✓	88.2	79.4	82.7	2.5	6.8

通过对比可知,具备自适应特征学习能力的 MRFA 模块增强了模型对鱼类特征的提取与表达能力,从而有效提升了水下鱼类的识别准确性。ECARU 模块具有更高的轻量化特性,其采用动态特征重组机制进行上采样,不仅有效保留了关键细节信息,也在识别准确性方面表现出明显提升。Varifocal Loss 对网络结构的参数量和推理效率影响较小,同时有助于优化正负样本不平衡问题,提高模型在水下复杂场景中的鲁棒性。综上所述,本文提出的各改进方法在不同性能维度上均展现出良好效果,最终提出的 FD-YOLO 在保持较低计算需求的同时显著提升识别能力,验证了各改进策略在水下识别任务中的有效性。

3.6 不同模型的对比实验

为了更全面评估所提出的 FD-YOLO 在水下鱼类识别任务中的性能表现,将其与多种主流目标检测模型进行

对比实验,包括 Faster R-CNN^[29]、YOLOv5n、YOLOv6n、YOLOv7-tiny^[30]、YOLOv8、YOLOv10n^[31] 和 YOLOv11n^[32]。不同识别模型的性能对比如表 3 所示。FD-YOLO 在 Precision、Recall 以及 mAP50-95 中均取得最优结果。Faster RCNN 的识别性能与 FD-YOLO 相近,但其参数量与运算量明显偏高,难以满足高效的应用场景。YOLOv5n 在参数量和计算量方面表现更为轻量,但其 mAP50-95 比 FD-YOLO 降低了 8.4%,水下识别能力较弱。与 YOLOv6n 和 YOLOv7-tiny 相比,FD-YOLO 的检测性能明显提升,mAP50-95 方面分别提升了 4.0% 和 4.2%。尽管 FD-YOLO 在推理效率方面略低于 YOLOv10n 和 YOLOv11n,但 mAP50-95 方面较后两者分别提高了 1.5% 和 0.9%。相比于 YOLOv8n,FD-YOLO 在 mAP50-95 方面提高了 3.7%,在参数量和计算量方面分别降低了 16.7% 和 16.0%。

表 3 不同识别模型的性能对比

Table 3 Performance comparison of different recognition models

模型	Precision/%	Recall/%	mAP50-95/%	Params/MB	FLOPs/G
Faster RCNN	86.4	77.1	81.6	41.8	134.4
YOLOv5n	80.53	71.01	74.3	1.8	4.2
YOLOv6n	84.1	74.6	78.7	4.2	11.6
YOLOv7-tiny	83.2	74.4	78.5	6.0	13.0
YOLOv8n	84.6	74.5	79.0	3.0	8.1
YOLOv10n	86.6	77.2	81.2	2.3	6.5
YOLOv11n	86.3	77.6	81.8	2.6	6.3
FD-YOLO	88.2	79.4	82.7	2.5	6.8

图 8 对比展示了不同模型在水下场景中的识别结果。由图 8(a)和(d)可知,在复杂的水下环境中,Faster R-CNN 与 YOLOv7-tiny 在处理鱼类间相互遮挡时,存在明显的冗余检测框。由图 8(b)和(c)可知,YOLOv5n 和 YOLOv6n 对水下鱼群的分类准确度效果较差。由图 8(f)和(g)可知,虽然 YOLOv10n 和 YOLOv11n 的识别结果相对较好,但在鱼类目标的定位精度方面仍存在少量偏差。此外,部

分模型未能识别出隐藏于昏暗背景中的模糊小目标,并且存在冗余识别的问题。与图 8(e)的 YOLOv8n 相比,图 8(h)的 FD-YOLO 能够更准确检测出鱼类的空间位置,并实现种类的精准分类。这种性能提升主要得益于 FD-YOLO 引入的动态自适应特征学习机制,在特征提取与表达方面具备更强的适应性,不仅提升了对小目标和遮挡目标的识别能力,也增强了对空间上下文的理解和特征

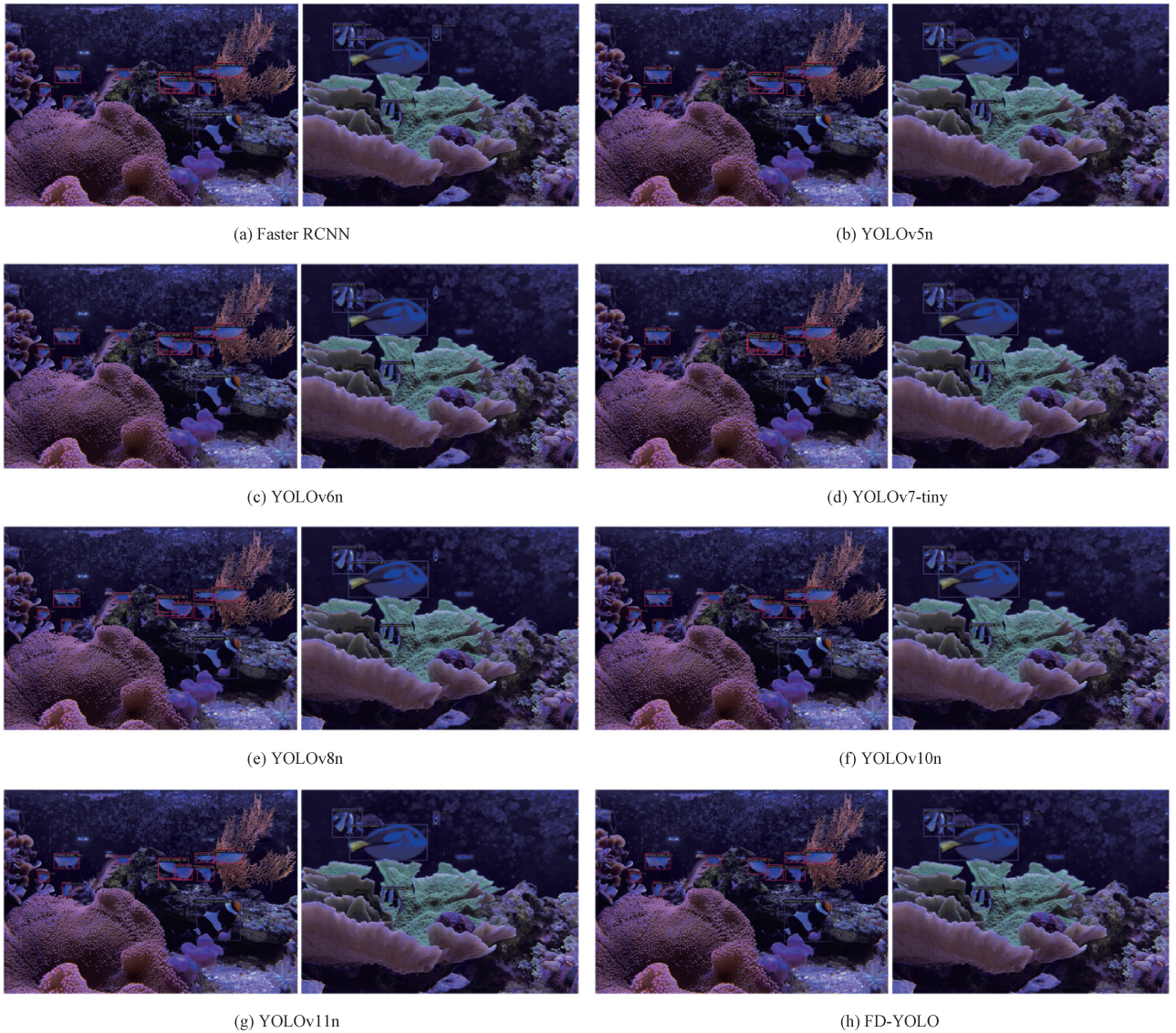


图8 不同识别模型的识别结果

Fig. 8 Recognition results across different detection models

融合能力,从而实现了水下鱼的种类与位置的更精准判定。

图9展示了模型改进前后的热力图对比结果。其中,图9(a)为水下原图,图9(b)和(c)分别为YOLOv8n和FD-YOLO的热力图。热力图以颜色梯度形式反映模型对特征区域的激活强度,颜色越趋于红色表示该区域的关注度或响应强度越高,反之则说明该区域对决策的影响较小。与YOLOv8n相比,改进后的FD-YOLO在目标鱼类区域的响应更为集中且边缘响应更为明确,表明其具备更强的特征提取能力和空间注意力聚焦能力。由此可见,FD-YOLO在小目标检测与复杂环境下具备更高的鲁棒性,更适用于水下场景下的目标识别任务。

为进一步评估模型在真实海域中的识别性能,选取水下较浑浊水体的图像进行测试,识别结果如图10所示。由图10(a)可知,Faster RCNN在正确定位鱼类目标的同时,将背景区域误判为其他鱼类,存在误识别现象。由图10(b)、(c)和(d)可知,在模糊的水下环境中,YOLOv5n与YOLOv6n的检测精度明显偏低,而YOLOv7-tiny虽能检测出目标,但产生了明显的冗余框。由图10(e)、(f)和(g)可知,YOLOv8n、YOLOv10n和YOLOv11n在检测准确性和鲁棒性方面表现相对较好,但整体性能仍略低于本文提出的FD-YOLO。图10(h)的FD-YOLO在精确检测与误检抑制方面均取得最佳综合表现,进一步证明了其在水下识别中的适用性与优势。

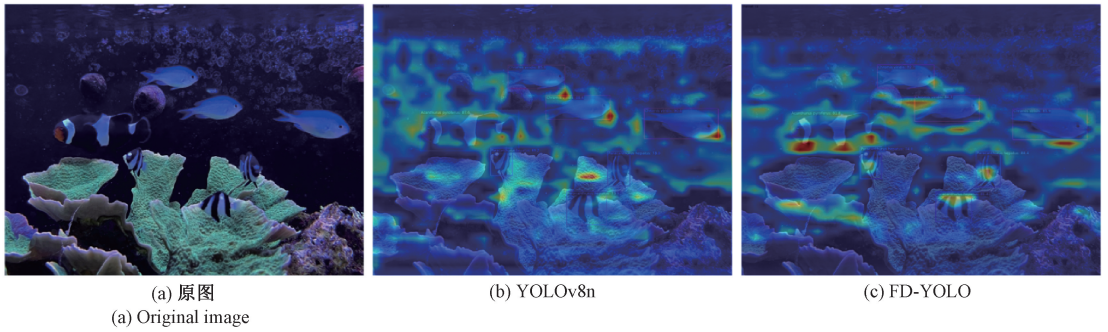


图 9 热力图对比结果
Fig.9 Heatmap comparison results

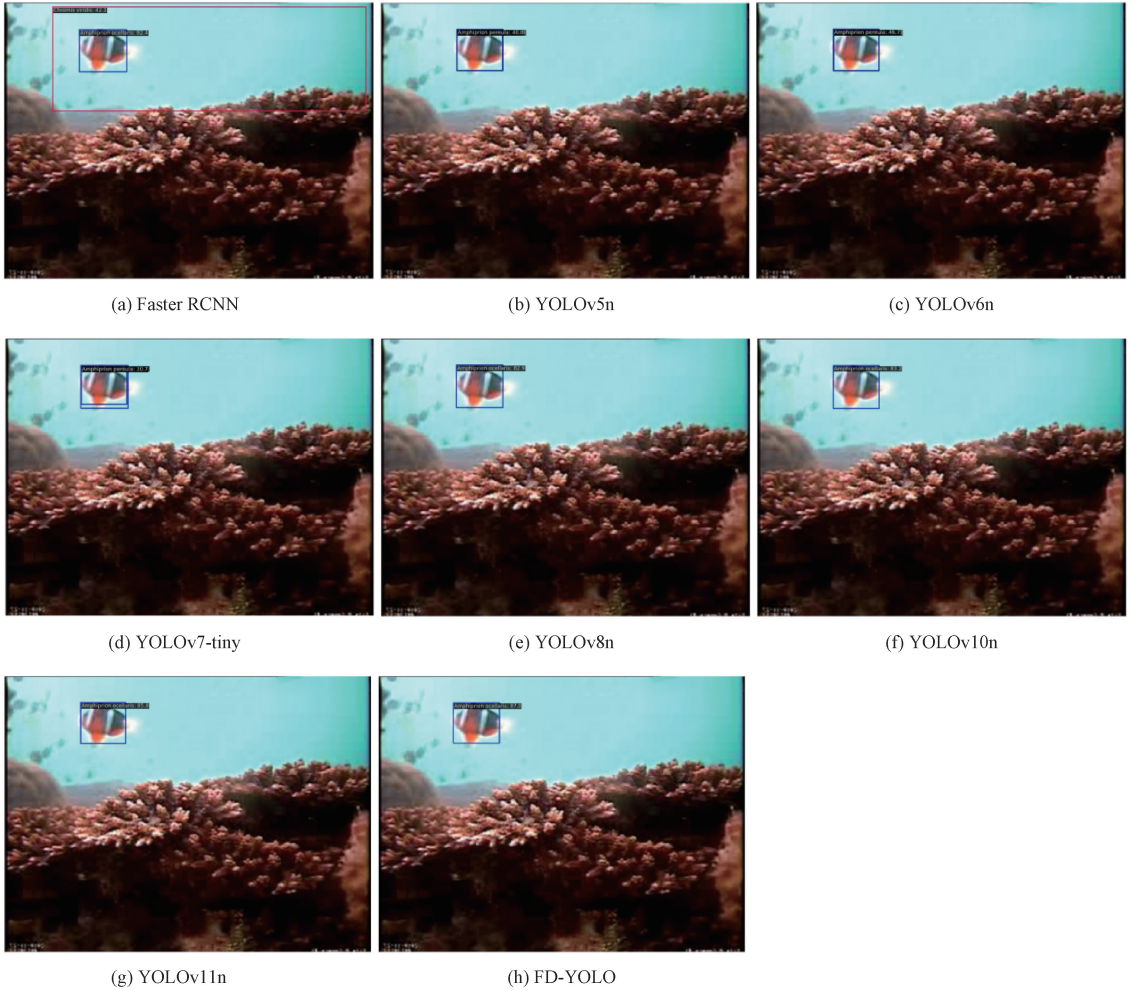


图 10 浑浊水质下的识别结果
Fig.10 Detection results under turbid water conditions

4 结 论

为提升在水下复杂环境中对鱼类的识别精度,本文提出了一种改进的目标识别方法 FD-YOLO。该方法以 YOLOv8n 为基础模型,在其主干中构建多尺度动态感知的 MRFA 模块,在颈部中设计双通道自适应特征重构的

ECARU 模块,并在检测头中融入 Varifocal Loss,以实现水下不同尺度、不同形态鱼类目标的精准识别。这些具备动态自适应特征学习能力的改进方法有效增强了对水下边缘模糊、小目标的特征提取与表达能力。通过多个主流模型对比表明,FD-YOLO 在保持参数量与计算需求较低的同时,能够显著提升水下识别性能。与 YOLOv8n 相

比,FD-YOLO的mAP50-95增长了3.7%,参数量和计算量分别缩减了16.7%和16.0%。该研究能够为水下目标检测的自动化与智能化发展提供技术支撑。

参考文献

- [1] 汤永华,张志鹏,林森,等. 基于多标签补偿的改进YOLOv8鱼体病害检测方法[J]. 农业工程学报, 2024, 40(23):227-234.
TANG Y H, ZHANG ZH P, LIN S, et al. An improved YOLOv8 method for fish disease detection based on multi-label compensation[J]. Transactions of the Chinese Society of Agricultural Engineering, 2024, 40(23): 227-234.
- [2] 汤永华, 张志鹏, 林森, 等. 基于深度学习的鱼类识别相关技术研究现状及展望[J]. 海洋渔业, 2024, 46(2):246-256.
TANG Y H, ZHANG ZH P, LIN S, et al. Research status and prospects of fish recognition technology based on deep learning[J]. Marine Fisheries, 2024, 46(2): 246-256.
- [3] 崔海朋,秦朝旭,马志宇. 基于深度学习的鱼类特征点检测与体征识别方法[J]. 中国农机化学报, 2024, 45(6):201-207.
CUI H P, QIN CH X, MA ZH Y. A method for fish feature point detection and body recognition based on deep learning[J]. Transactions of the Chinese Society of Agricultural Machinery, 2024, 45(6): 201-207.
- [4] 葛艳,张亚婷,李海涛. 基于特征融合的FL-BCNN鱼类识别算法研究[J]. 计算机应用与软件, 2023, 40(8):174-179.
GE Y, ZHANG Y T, LI H T. Research on fish recognition algorithm based on feature fusion of FL-BCNN [J]. Computer Applications and Software, 2023, 40(8): 174-179.
- [5] 李艳君,黄康为,项基. 基于立体视觉的动态鱼体尺寸测量[J]. 农业工程学报, 2020, 36(21):220-226.
LI Y J, HUANG K W, XIANG J. Dynamic fish body size measurement based on stereo vision [J]. Transactions of the Chinese Society of Agricultural Engineering, 2020, 36(21): 220-226.
- [6] 陈子文,李卓璐,杨志鹏,等. 基于YOLOv5的工厂化养殖虾目标检测方法研究[J]. 海洋渔业, 2022, 44(5):610-620.
CHEN Z W, LI ZH L, YANG ZH P, et al. Research on target detection method for factory-cultured shrimp based on YOLOv5 [J]. Marine Fisheries, 2022, 44(5): 610-620.
- [7] 涂雪滢,刘世晶,钱程. 基于ResNet的典型养殖鱼类识别方法研究[J]. 渔业现代化, 2022(3):81-88.
TU X Y, LIU SH J, QIAN CH. Research on typical aquaculture fish recognition method based on ResNet[J]. Fisheries Modernization, 2022(3): 81-88.
- [8] 张丽娟,游浩海,李芝贻,等. 基于改进YOLOv8n轻量化的人参外观质量精准识别[J]. 农业工程学报, 2024, 40(24): 274-282.
ZHANG L J, YOU H H, LI ZH Y, et al. Precise recognition of ginseng appearance quality based on improved lightweight YOLOv8n[J]. Transactions of the Chinese Society of Agricultural Engineering, 2024, 40(24): 274-282.
- [9] MONKMAN G G, HYDER K, KAISER M J, et al. Using machine vision to estimate fish length from images using regional convolutional neural networks [J]. Methods in Ecology and Evolution, 2019, 10(12): 2045-2056.
- [10] CAI K W, MIAO X Y, WANG W, et al. A modified YOLOv3 model for fish detection based on MobileNetv1 as backbone[J]. Aquacultural Engineering, 2020, 91: 102117.
- [11] 袁红春,张硕. 基于Faster R-CNN和图像增强的水下鱼类目标检测方法[J]. 大连海洋大学学报, 2020, 35(4): 612-619.
YUAN H CH, ZHANG SH. Underwater fish target detection method based on Faster R-CNN and image enhancement[J]. Journal of Dalian Ocean University, 2020, 35(4): 612-619.
- [12] ZHAO ZH X, LIU Y, SUN X D, et al. Composited FishNet: Fish detection and species recognition from low-quality underwater videos[J]. IEEE Transactions on Image Processing, 2021, 30: 4719-4734.
- [13] LIU T, LI P L, LIU H Y, et al. Multi-class fish stock statistics technology based on object classification and tracking algorithm [J]. Ecological Informatics, 2021, 63: 101240.
- [14] LI H Q, YU H, GAO H T, et al. Robust detection of farmed fish by fusing YOLOv5 with DCM and ATM[J]. Aquacultural Engineering, 2022, 99: 102301.
- [15] 刘雨青,王亚茹,黄璐瑶. 基于FML-Centernet算法的鱼类识别检测[J]. 激光与光电子学进展, 2022, 59(16):317-324.
LIU Y Q, WANG Y R, HUANG L Y. Fish recognition and detection based on the FML-Centernet algorithm [J]. Laser & Optoelectronics Progress, 2022, 59(16): 317-324.
- [16] 王伟芳,殷健豪,高春奇,等. 增强细节信息特征提取的鱼类个体识别算法[J]. 现代电子技术, 2024, 47(2):183-186.
WANG W F, YIN J H, GAO CH Q, et al. Fish individual recognition algorithm for enhanced feature extraction of detailed information [J]. Modern Electronics Technology, 2024, 47(2): 183-186.

- [17] 吕俊霖, 陈作志, 李碧龙, 等. 基于多阶段特征提取的鱼类识别研究[J]. 南方水产科学, 2024, 20(1):99-109.
LYU J L, CHEN Z ZH, LI B L, et al. Research on fish recognition based on multi-stage feature extraction[J]. Southern Fisheries Science, 2024, 20(1): 99-109.
- [18] 胡峻峰, 李柏聪, 朱昊, 等. 改进 YOLOv8 的轻量化无人机目标检测算法[J]. 计算机工程与应用, 2024, 60(8):182-191.
HU J F, LI B C, ZHU H, et al. Lightweight UAV target detection algorithm based on improved YOLOv8 [J]. Computer Engineering and Applications, 2024, 60(8): 182-191.
- [19] 王海群, 武泽锴, 晁帅. 改进 YOLOv8n 的遥感图像目标检测算法[J]. 电子测量与仪器学报, 2025, 39(4):84-94.
WANG H Q, WU Z K, CHAO SH. Improve the YOLOv8n object detection algorithm for remote sensing images[J]. Journal of Electronic Measurement and Instrumentation, 2025,39(4):84-94.
- [20] 李泽胤, 李栋, 房建东, 等. 改进的 YOLOv8n 遥感图像轻量化检测模型[J]. 电子测量技术, 2025, 48(6):130-142.
LI Z Y, LI D, FANG J D, et al. Improved YOLOv8n light weight detection model for remote sensing images [J]. Electronic Measurement Technology, 2025, 48(6): 130-142.
- [21] 赵会鹏, 曹景胜, 潘迪敬, 等. 改进 YOLOv8 算法的交通标志小目标检测[J]. 现代电子技术, 2024, 47(20):141-147.
ZHAO H P, CAO J SH, PAN D J, et al. Small target detection of traffic signs using the improved YOLOv8 algorithm [J]. Modern Electronics Technology, 2024, 47(20): 141-147.
- [22] CHEN L C, PAPANDREOU G, KOKKINOS I, et al. Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 40(4): 834-848.
- [23] ZHANG X, LIU C, YANG D G, et al. RFAConv: Innovating spatial attention and standard convolutional operation [J]. ArXiv preprint arXiv: 2304.03198, 2023.
- [24] HU J, SHEN L, SUN G. Squeeze-and-excitation networks[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2018: 7132-7141.
- [25] 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.
- [26] 刘翰林, 刘凌云, 李超凡, 等. 基于 YOLOv5s-CBC 的钢材表面缺陷检测[J]. 组合机床与自动化加工技术, 2024(10): 155-159.
LIU H L, LIU L Y, LI CH F, et al. Steel surface defect detection based on YOLOv5s-CBC[J]. Modular Machine Tool & Automatic Manufacturing Technology, 2024(10): 155-159.
- [27] ZHANG H Y, WANG Y, DAYOUB F, et al. Varifocalnet: An iou-aware dense object detector[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021: 8514-8523.
- [28] LU ZH X, ZHU X L, GUO H T, et al. FishFocusNet: An improved method based on YOLOv8 for underwater tropical fish identification[J]. IET Image Processing, 2024, 18(12): 3634-3649.
- [29] REN SH Q, HE K M, 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.
- [30] WANG CH Y, BOCHKOVSKIY A, LIAO H Y M. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023: 7464-7475.
- [31] WANG A, CHEN H, LIU L H, et al. YOLOv10: Real-time end-to-end object detection[J]. Advances in Neural Information Processing Systems, 2024, 37: 107984-108011.
- [32] KHANAM R, HUSSAIN M. YOLOv11: An overview of the key architectural enhancements [J]. ArXiv preprint arXiv:2410.17725, 2024.

作者简介

朱晓龙(通信作者), 博士, 讲师, 硕士生导师, 主要研究方向为计算机视觉与目标识别。

E-mail: zhuxiaolonglong@sina.cn

李超凡, 硕士研究生, 主要研究方向为海洋目标检测算法。

E-mail: 13117153692@163.com

陈郁隈, 硕士研究生, 主要研究方向为机器视觉与海洋探测。

E-mail: yuweichen@163.com

陈祥子, 博士, 副教授, 硕士生导师, 主要研究方向为海洋光学和光学成像。

E-mail: xzchen913@163.com

褚文敬, 本科, 主要研究方向为计算机视觉与海洋技术。

E-mail: uadwn123@163.com