

基于改进 YOLO v8 模型的水下小型生物目标检测算法研究

周焕银¹ 李明桂¹ 刘怡佳² 朱睿鹏¹

(1. 东华理工大学电子与电气工程学院, 南昌 330000; 2. 燕山大学电气工程学院, 秦皇岛 066000)

摘要: 针对传统检测算法在水下小型生物目标检测中存在的定位精度低以及难以识别微小目标和重叠目标的问题,提出了一种改进的 YOLO v8 小目标检测算法(WLDA-YOLO v8)。设计了轻量级多尺度卷积模块 LEM,替换原网络中 C2f 模块的标准卷积,降低了计算复杂度,提高了模型在复杂环境下的特征提取能力;融合了 ATSS 样本分配策略和动态检测头 DyHead,实现动态自适应的训练样本选择;集成改进的 WIoU v3s 边界框回归损失函数,以提高模型对水下小目标的定位精度。在自建数据集上的综合测试中,WLDA-YOLO v8 的 mAP50-95 比原始 YOLO v8s 提高 5.5 个百分点,同时参数量降低至 7.9×10^6 ,浮点运算量由 1.98×10^{10} 降至 1.59×10^{10} ,大幅降低了模型复杂度。研究结果为提高水下小型目标检测的实时性和准确性提供了有效途径。

关键词: 水下目标检测; 小型生物; YOLO v8; LEM; ATSS; WIoU v3s

中图分类号: TP39

文献标识码: A

文章编号: 1000-1298(2026)19-0114-12

OSID:



Target Detection Algorithm for Underwater Small Organisms Based on Improved YOLO v8 Model

Zhou Huanyin¹ Li Minggui¹ Liu Yijia² Zhu Ruipeng¹

(1. School of Electronic and Electrical Engineering, East China University of Technology, Nanchang 330000, China

2. School of Electrical Engineering, Yanshan University, Qinhuangdao 066000, China)

Abstract: Under water environment is complex and dynamic. When light propagates in water, it undergoes severe attenuation, scattering and color cast, leading to prevalent defects in underwater images, including low contrast, strong noise and blurred details, which restricted the application of underwater target detection technology in marine resource exploration, ecological monitoring and underwater engineering. To tackle the challenges of low localization accuracy and difficulties in detecting small and overlapping targets in traditional detection algorithms for underwater biological targets, WLDA-YOLO v8, an improved version of the YOLO v8 algorithm optimized for small-object detection was proposed. A lightweight multi-scale convolution module (LEM) was designed to replace the original C2f module in the network, reduce computational complexity and enhance feature extraction in complex environments. To enable dynamic and adaptive training sample selection, the ATSS sample assignment strategy and the dynamic detection head (DyHead) were integrated. Additionally, the incorporation of an improved WIoU v3s bounding box regression loss function improved the model's localization accuracy for small underwater objects. Compared with the original YOLO v8s, some comprehensive tests on the custom data set showed that mAP50-95 of WLDA-YOLO v8 was improved by 5.5 percentage points, while the number of parameters was reduced to 7.9×10^6 and GFLOPs was decreased from 1.98×10^{10} to 1.59×10^{10} , significantly reducing model complexity. The experimental results verified that the proposed algorithm significantly elevated the accuracy and efficiency of the model in underwater target detection tasks, as well as its convergence speed and training stability.

Key words: underwater object detection; small organisms; YOLO v8; LEM; ATSS; WIoU v3s

收稿日期: 2026-06-23 修回日期: 2026-06-26

基金项目: 江西省科技厅重点基金项目(2022ACB204022)、国家自然科学基金项目(62063001)和燕山大学创新创业基金项目(CXXL20250206)

作者简介: 周焕银(1975—),女,教授,博士,主要从事水下机器人应用及控制研究,E-mail: hyzhou@ecut.edu.cn

0 引言

海洋生态系统的复杂性和多样性对于地球环境的平衡以及生物资源的可持续利用至关重要^[1]。高效、准确地检测和识别水下生物,对促进海洋资源的合理利用和生物多样性保护具有重要意义^[2]。然而,水下图像通常受到光线衰减、色彩偏移、低对比度和模糊等因素的影响,导致图像质量大幅下降^[3]。此外,水下生物的目标通常体积较小、容易聚集,且在图像中的占比极低,进一步增加了检测难度^[4]。这些问题使得传统目标检测方法在水下环境中表现不佳。

随着深度学习的迅速发展,基于卷积神经网络(CNN)的目标检测算法在水下环境中展现了巨大的潜力。两阶段检测算法如 Faster R-CNN^[5] 和 Mask R-CNN^[6] 尽管在检测精度上表现优异,但由于其计算复杂度较高,不适用于实时检测任务。相比之下,YOLO 等单阶段检测算法因其高效、实时的检测能力,成为水下目标检测的理想选择。特别是 YOLO v8,作为 YOLO^[7] 系列的经典版本,在网络结构和训练策略上有了显著提升,显著提高了检测精度和速度。然而,YOLO v8 在处理水下复杂环境中的小目标和模糊目标时,仍然存在一定的局限性。例如,其标准卷积模块(C2f)在多尺度特征提取上表现不足,难以应对复杂场景中的目标检测需求,Song 等^[8] 利用 Mask R-CNN 对水下生物进行识别,达到了 97.30% 的准确率,Mask R-CNN 精度较高,但增加了计算资源的开销。此外,传统的标签分配策略和损失函数在面对水下环境中的遮挡和复杂背景时,可能导致模型收敛速度减慢及检测精度降低。Guo 等^[9] 利用 FasterNet 模块取代 YOLO v8 骨干网络,轻量化了模型,但计算效率不高,泛化能力有限。

为了解决计算资源开销大、计算效率低,泛化能力受限等问题,提高水下生物检测任务效率,本文提出一种改进的 WLDA-YOLO v8 模型。本研究设计一种多尺度卷积模块 LEM,以替换 YOLO v8 中的标准卷积模块 C2f,该模块具有高效提取模型特征的能力^[10];融合 ATSS 标签分配策略^[11] 结合动态检测头 DyHead^[12],实现动态自适应训练样本选择,以提升检测精度与增强模型在复杂场景中的泛化能力;在目标边界框回归损失中采用改进的 WIoU v3s 损失函数^[13],以提高模型对水下小目标的检测性能与高定位精度。

1 YOLO v8 算法

YOLO v8 沿袭了 YOLO 系列的设计理念,注重

速度和效率,将架构上的重要创新适应于各类应用场景^[14]。采用 YOLO v8 进行水下小目标识别的主要优势为:

(1) 具有提高特征提取能力、减少计算量等优势,因其主干网络是基于 CSPDarknet53 设计,并引入改进的 C2f 模块,C2f^[15] 模块采用阶段间双向切换和部分连接,增加网络处理各种尺度特征的能力,可减少样本训练过程的计算工作量。这种设计通过优化信息流增加了特征提取的深度和广度,同时保持了响应能力。

(2) 具有超强的微小目标检测能力,颈部采用特征金字塔网络(FPN)与路径聚合网络(PAN)相结合的方式,有效地结合多种尺度的特征,使得模型能够从细到粗捕获各种细节信息,从而提升了对小目标的检测能力。

(3) 具有精确定位与很好的目标识别能力,检测头采用了解耦的设计,分类和回归任务同时分开处理,以提升各自的效率和精确度,这种解耦的检测头允许模型在减少复杂度的同时,更加精确地定位和识别目标。

(4) YOLO v8 为进一步提升损失函数的性能,引入了 Varifocal Loss 和 CIoU Loss,这些改进有助于处理类别不平衡和复杂重叠场景下的检测问题。同时模型也支持 Mosaic 和 Mixup 在内的多种数据增强技术,有效提升模型能力和复杂环境下整体的鲁棒性。

2 WLDA-YOLO v8 模型构建

针对经典检测模型在复杂水下环境中存在的定位偏差较大,小目标易漏检、误检等突出问题,提出了 WLDA-YOLO v8,对 YOLO v8 进行三大优化:设计一种高效的轻量级卷积模块 LEM (Lightweight efficient multi-scale module),代替 YOLO v8 网络中原有的标准卷积结构,以提升模型对小型和模糊目标的特征感知能力;引入一种智能的样本分配方法——Dynamic ATSS (Adaptive training sample selection),以帮助模型在训练过程中动态自适应地调整正负样本的分离阈值;引入先进的 WIoU v3s 边界框回归损失函数,以进一步优化目标定位的准确性和检测头的泛化能力。网络整体结构如图 1 所示。

2.1 LEM 模块

在 YOLO v8 的基础上设计了一个轻量且高效的多尺度卷积模块——LEM,替代 C2f 模块中的标准卷积。LEM 模块的设计目的是降低模型参数和计算复杂度,同时增强捕获多尺度空间特征的能力,

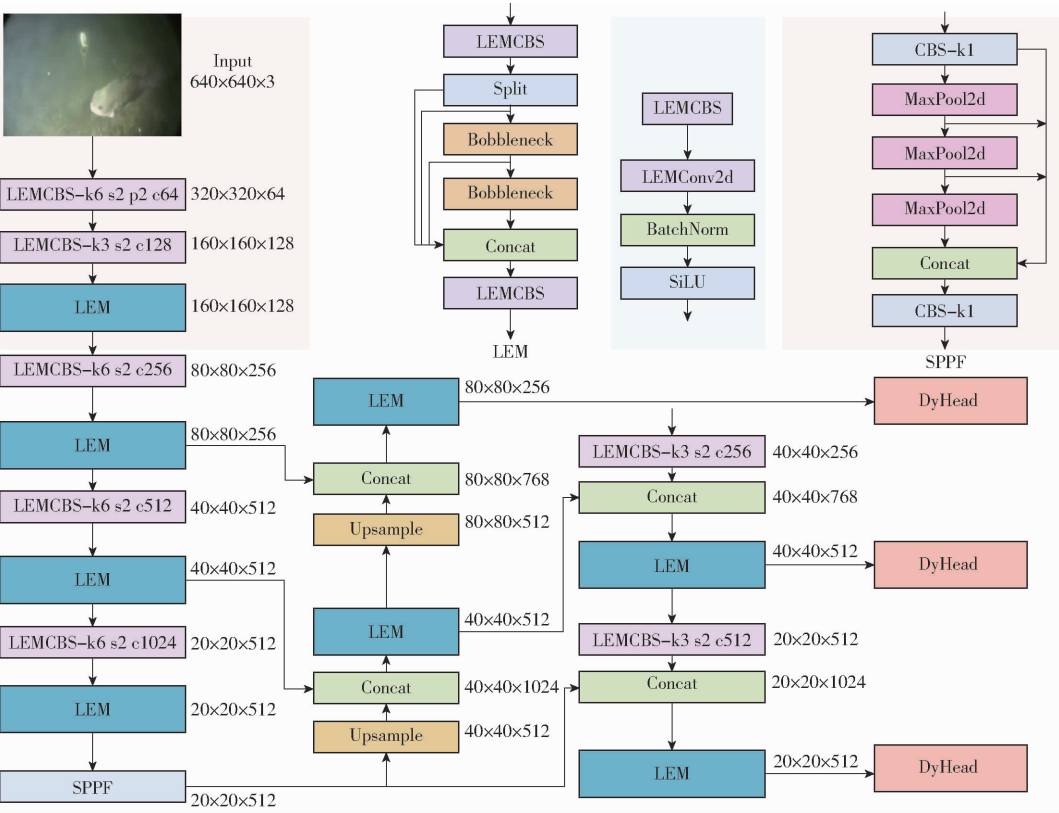


图 1 WLDA - YOLO v8 整体结构图

Fig. 1 Overall structure of WLDA - YOLO v8

易于处理复杂水下环境中的目标。LEM 模块使用具有不同核大小的卷积层处理多尺度特征,因此可以有效捕获图像中各个尺度的信息,如图 2 所示。

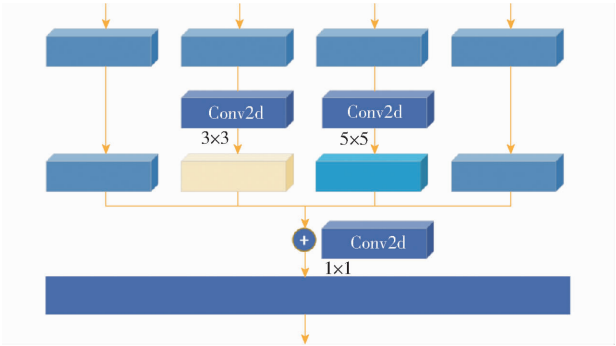


图 2 LEMConv2d 模块结构图

Fig. 2 LEMConv2d module structure

LEM 模块中的 LEMConv2d 操作将输入通道划分为 4 部分,每个部分处理不同的卷积操作。这 4 部分即为卷积通道,其中第 1 个和第 3 个通道保持不变,第 2 个和第 4 个通道分别进行 3×3 和 5×5 卷积。后将所有通道通过卷积操作整合生成最终的输出特征映射。这种多卷积核设计不但扩大了感受野,而且增强了模型处理长距离依赖特征的能力,使得模型在处理复杂场景中的小物体时表现更好。图 2 详细展示了 LEMConv2d 模块中每个操作流程,包括全局平均池化、全连接层和 Sigmoid 函数。用生

成的权重系数调整每个通道的权重,以确保模型可以动态地突出重要特征,使其更加关注重要特征并减少冗余信息的计算。

通过这种设计,LEM 模块在减少参数数量和避免特征冗余的同时,保留了关键特征信息。主体的一半特征采用了卷积与原特征融合,LEM 模块可在维持模型性能的同时生成丰富的特征表达。LEM 模块能够捕捉局部细节,并且随着网络的加深,逐步保留和增强目标物体的细微特征和语义信息,这使得模型在检测小尺度物体方面优于检测大规模物体。LEM 模块的整体结构如图 3 所示。通过上述改进,LEM 模块在减少网络参数和加快检测速度的同时捕获目标的多尺度特征和局部细节,对 YOLO v8 整体性能的提升做出了显著贡献。

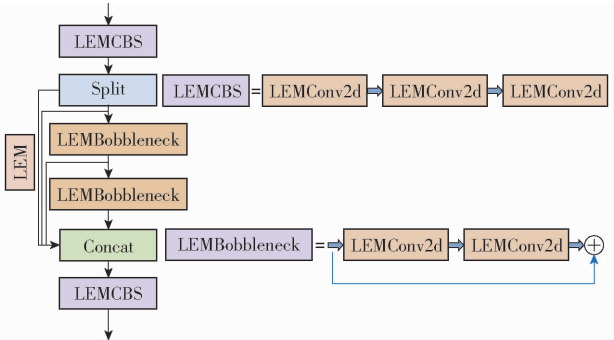


图 3 LEM 模块结构图

Fig. 3 LEM module structure

2.2 动态自适应训练样本选择 (Dynamic ATSS)

传统特征分配方法通常基于固定的 IoU 阈值,该方法在面对多维目标和复杂背景时虽然能够取得最佳的特征分配效果,但易影响模型的检测性能。为了解决这个问题,本文提出了一种创新的 ATSS 标签分配策略,引入一种智能机制来动态调整正负样本的选择,评估每个候选框与目标对齐,并动态确定合适的 IoU 阈值以自适应地筛选最具代表性的训练样本。该策略不仅明显提高模型的检测精度,而且提高模型对复杂干扰环境的泛化能力。

为进一步提升目标检测性能,DyHead (Dynamic Head) 模块^[16]在 YOLO v8 中原有的分段式注意力机制之前引入。DyHead 模块通过融合尺度感知、空间感知和通道感知 3 种不同类型的注意力机制,帮助模型更准确地生成多维特征图,从而在各类检测任务中取得更好的性能。DyHead 模块的结构包括多个 DyHead 块,每个 DyHead 块均包含 3 个对应注意力子模块。图 4a 清晰地呈现了 DyHead 整体结构,其中 L 、 S 和 C 分别代表特征图的尺度、空间位置和通道维度。图 4b 详细展示了 DyHead 各注意力机制的实现细节。 π_L 先通过全局平均池化和 1×1 卷积操作,提取特征图中不同尺度的全局信息,从而自适应地调整特征图的尺度响应;然后 π_S 采用了可变形卷积机制 (Deformable convolution),动态学习特征图空间维度上的重要区域,从而实现空间注意力的精细化调度。最后 π_C 自适应全局平均池化消除冗余维度,使用两个全连接层和动态 ReLU 激活函数自适应地增强通道信息,并突出对任务性

能有重大贡献的通道特征。通过结合 ATSS 策略和 DyHead 模块,模型不仅提高了检测精度,也提高了在复杂场景中的模型性能。

2.3 改进的损失函数

在水下目标检测任务中,由于细长物体的比例较高,合理设计损失函数对提高模型的检测性能至关重要。YOLO v8 使用了 CIOU 和 DFL 作为边界框回归损失的主要计算方法。然而,CIOU 在处理复杂场景时仍存在一些局限性:① 难易样本的平衡。CIOU 在样本难度的平衡方面表现有限,无法自适应调整难样本与易样本之间的权重分配^[17]。这意味着模型在面对具有不同难度的样本时,可能会表现出不一致的学习效果。② 长宽比的局限性。CIOU 以长宽比差异作为惩罚因子之一,但当预测框与真实框长宽比相同而绝对尺寸不同情况下,CIOU 的惩罚机制无法有效体现二者之间的差异,可能导致某些情况下预测框定位不够准确。③ 计算复杂性。CIOU 的计算公式中包含函数,在计算资源有限的环境中增加了额外计算量,可能影响模型的推理效率。

尽管 CIOU 具有弥补传统 IoU 不足的优势,但其在处理复杂背景和困难样本时仍有一定的提升空间。为了解决上述问题,以改进的 WIoU v3s 作为新的边界框回归损失函数。所改进的 WIoU v3s 综合了 CIOU、EIoU 和 SIOU 的优点,引入了动态非单调聚焦机制,可更有效地评估锚框的质量并优化损失分配策略^[18]。其设计理念主要考虑 3 方面:角度惩罚因子 (SIOU)、动态非单调聚焦机制与多场景适应性。

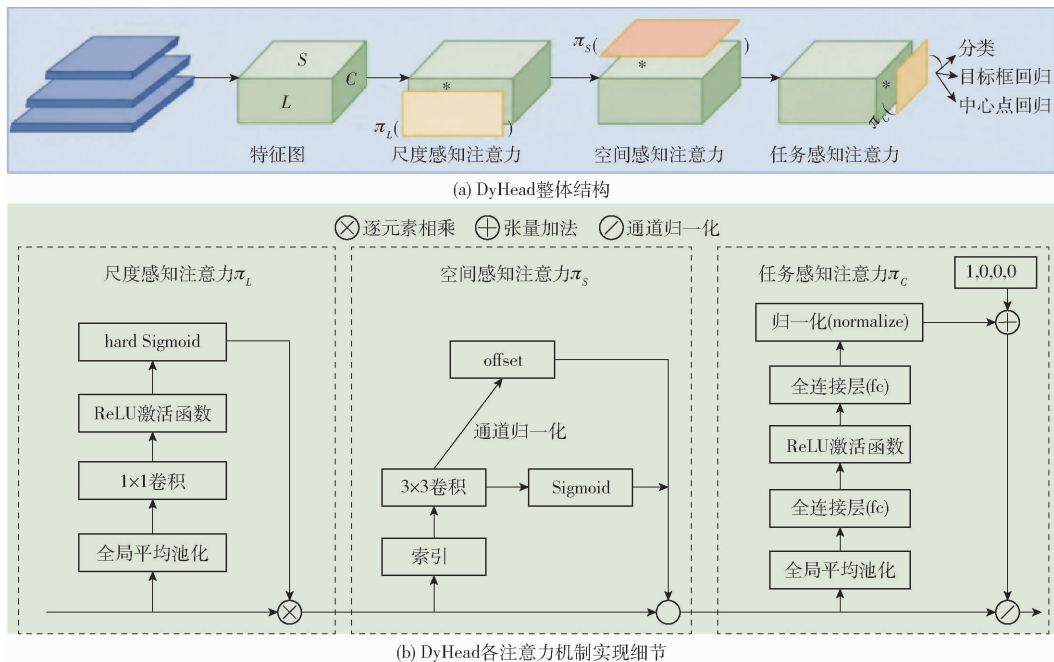


图 4 DyHead 结构图

Fig. 4 DyHead structure diagram

角度惩罚因子(SIoU):SIoU 在损失函数中首次引入了预测框与真实框中心连线的夹角作为惩罚因子。通过动态调整角度,SIoU 使预测框更快地接近目标真实框,降低回归过程的自由度,从而加快模型的收敛。CIoU 计算公式给出了 SIoU 损失中与角度相关的计算,基于角度和的大小变化,预测框会更快地向实际框移动并最终与其交互。通过引入角度惩罚项,SIoU 可有效降低目标尺度相同但位置不同情况下的定位误差。图 5 为损失函数相关参数示意图。

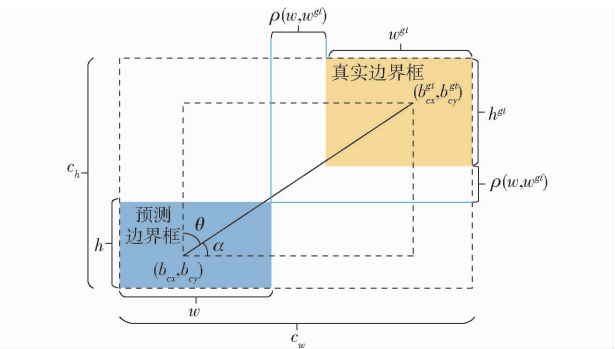


图 5 损失函数相关参数示意图

Fig.5 Schematic of parameters related to loss function

动态非单调聚焦机制:WIoU v3s 借鉴了 Focal-EIoU 的思想,采用动态非单调的梯度聚焦因子来自适应调整不同质量锚框对损失的贡献^[19]。对于预测准确、IoU 较高者的锚框,赋予较小的梯度增益,降低其在总损失中的权重;对于 IoU 较低的困难样本的锚框,赋予较大的梯度增益,促使模型更加关注易出错样本^[20]。WIoU v3s 的计算公式为

$$L_{WIoUv3} = \gamma L_{WIoUv1} \tag{2}$$

其中
$$\gamma = \frac{\beta}{\delta \alpha^{\beta - \delta}} \tag{3}$$

$$\beta = \frac{L_{lou}^*}{L_{lou}} \in [0, +\infty) \tag{4}$$

式中 γ ——权重系数,用于缩放基础损失 L_{WIoUv1}
 β ——衡量锚框质量的指标
 L_{lou}^* ——训练得到的 l_{ou} 损失最优值
 L_{lou} ——训练得到的实际 IoU 损失值

其中 δ, α 为可调超参数,通过基于 β 构造的非单调聚焦因子 γ 并将其应用于 WIoU v3s,模型能动态平衡高、低质量样本的损失权重分配——减小高质量样本的梯度干扰,同时适当提升对低质量样本的关注。

多场景适应性:WIoU v3s 在小物体检测和复杂背景场景下表现出色。通过动态提升小物体在损失函数中的权重,WIoU v3s 能够显著增强模型对这些场景的适应能力。在细小目标众多或目标与背景难以区分的情况下,该损失函数有助于模型更好地收敛于最优解。此外,对于水下细长物体的检测任务,

WIoU v3s 所引入的角度惩罚和动态聚焦机制可以缓解定位难度,提升检测的准确性。

3 实验设计与结果分析

3.1 数据集构建与实验环境

所构建的数据集命名为 UMDD (Underwater multi-target detection dataset),其主要数据来源包括 RUOD 数据集、2021 年水下机器人专业竞赛官方数据集(URPC2021),以及奥尔堡大学公开的 Brackish Underwater 数据集(BUD)^[21]。从这 3 个数据集中共收集了 11 741 幅图像用于模型训练与评估。按照比例 7:1:2 将数据集划分为训练集、验证集和测试集,共 8 221 幅训练图像、1 173 幅验证图像和 2 345 幅测试图像,自建水下目标检测数据集的详细信息。如此划分确保了模型训练的准确性,并在验证和确认阶段提供了可靠的性能估计。数据收集过程中分析了来自不同水下环境(例如浅海、深海、珊瑚礁等)的样本。在模型训练过程中,选择 batch size 为 8,并运行了 200 个 epoch,初始学习率设为 0.01,使用权重衰减 0.000 5,动量设置为 0.827。这些超参数的设置考虑了显存容量、训练速度和防止过拟合的需求。为了确保标注的准确性,使用了 Labelme 工具对所有图像进行了细致标注,生成了相应的 XML 文件,这些文件包括每幅图像的尺寸、标记框的位置信息以及对应的海洋生物类别。

图 6 展示了训练集的基本信息,涵盖了每个类别的样本数量、目标框尺寸、目标中心点的分布以及目标的高宽比。所研究数据集中不同类别的目标大小变化显著,尤其是小目标所占比例较高。模型在训练过程中需处理目标差异大和背景复杂问题,增加了检测难度。图表的分析还表明,小目标在训练集中占据了较大的比例,尤其是螃蟹、小鱼和虾类目标的尺寸较小,这对模型的检测精度提出了更高要求。此外,目标的中心点分布显示出这些目标并不总是集中在图像中心,而是随机分布在图像的各个角落,这也增加了目标定位的难度。模型在训练过程中需要应对目标大小变化显著且小目标占比大、目标随机分布定位难等多种挑战,从而提高检测的泛化能力和精度。

3.2 模型评估指标

通过多个评估指标对改进的 YOLO v8 模型进行了详细评估,包括精确率、召回率、平均检测时间、平均精确率均值(mAP)、参数量和浮点运算量。这些指标全面反映了模型在目标检测任务中的性能。

mAP 是评估模型在多类别检测中的整体性能的关键指标,特别是 mAP50-95,反映了模型在不

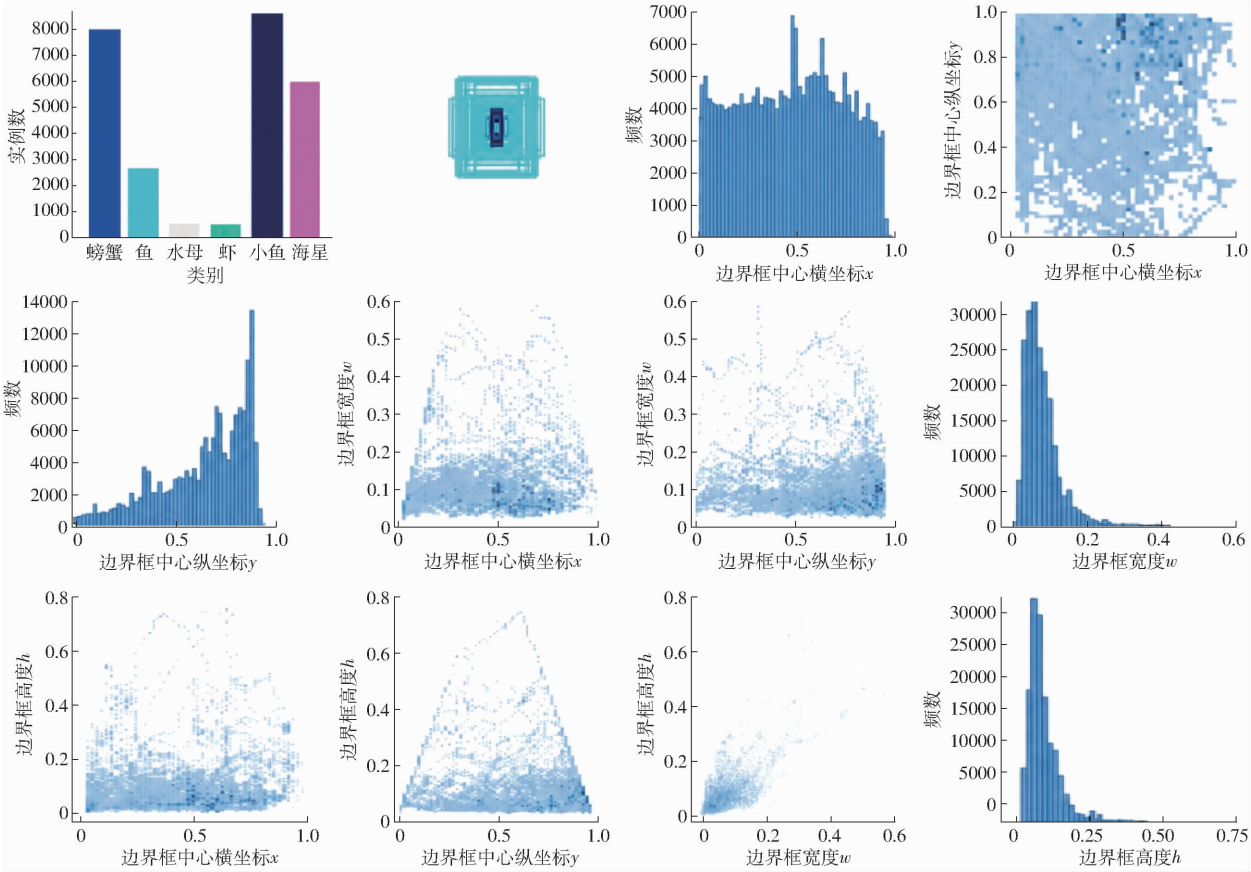


图 6 训练集基本信息图

Fig. 6 Basic information diagram of training set

同 IoU 阈值下的表现。计算方法包括将每个类别的 AP 值相加,然后除以类别总数。

图 7 对改进的 YOLO v8 模型进行了详细的图形对比分析。评估指标包括召回率、平均精确率均值、边界框回归损失 (box_Loss)、分类损失 (cls_Loss)、交叉熵损失函数 (dfl_Loss) 的变化,以及训练

过程中精确率和召回率的曲线表现,这些指标可以全面反映模型在目标检测任务中的性能表现。

从图 7 中 mAP50 和 mAP50-95 曲线可以看出,WLDA-YOLO v8 模型在大部分训练阶段都取得了更高的精度。尤其是在严格的 mAP50-95 指标下,WLDA-YOLO v8 明显优于对比模型 WSSOD-

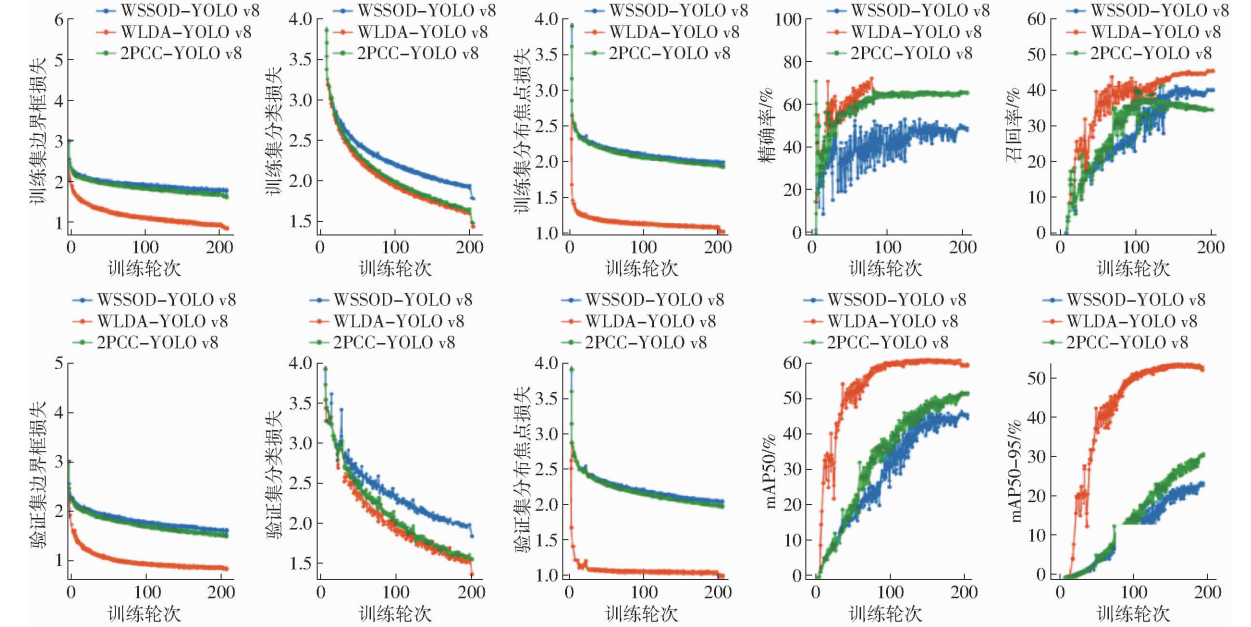


图 7 WLDA-YOLO v8 性能评估对比图

Fig. 7 WLDA-YOLO v8 performance evaluation comparison charts

YOLO v8 和 2PCC - YOLO v8。引入多尺度卷积模块的 WLDA - YOLO v8 策略有效增强了模型对小目标的检测能力。mAP50 - 95 曲线也显示出模型精度随着训练迭代迅速提升并最终趋于稳定。

3.3 ATSS 标签分配策略效果

为了验证 ATSS 策略的有效性,对比了原始模型与引入 ATSS 策略的模型在 UMDD 数据集上的检测表现。ATSS 通过动态选择训练样本,使标签分配更加精准,从而提高模型对不同类别目标的识别能力。实验结果如表 1 所示,应用 ATSS 后模型整体

检测性能得到提升,除鱼类 AP 略有下降外,其余类别 AP 均有所提高。提升幅度虽有限,但在复杂背景或小目标检测场景下具有重要意义。图 8a 为基线模型的混淆矩阵,图 8b 为加入 ATSS 策略后的混淆矩阵。对比可见,引入 ATSS 策略后,明显减少了类别间的误识别情况,螃蟹、鱼、海星等类别的识别准确性显著提高,尤其是海星类别的误识别数量明显减少。这说明 ATSS 策略有效提升了模型的分类精准度,增强了模型在复杂水下环境中的鲁棒性和泛化能力。

表 1 加入 ATSS 标签分配策略效果

Tab.1 Effect of adding ATSS label allocation strategy %

类别	AP						mAP50	mAP50 - 95
	螃蟹	鱼	水母	虾	小鱼	海星		
UMDD	69.7	90.6	82.4	63.4	69.4	83.6	76.5	39.9
UMDD + ATSS	70.6	89.9	83.2	65.1	70.5	85.1	77.7	42.9

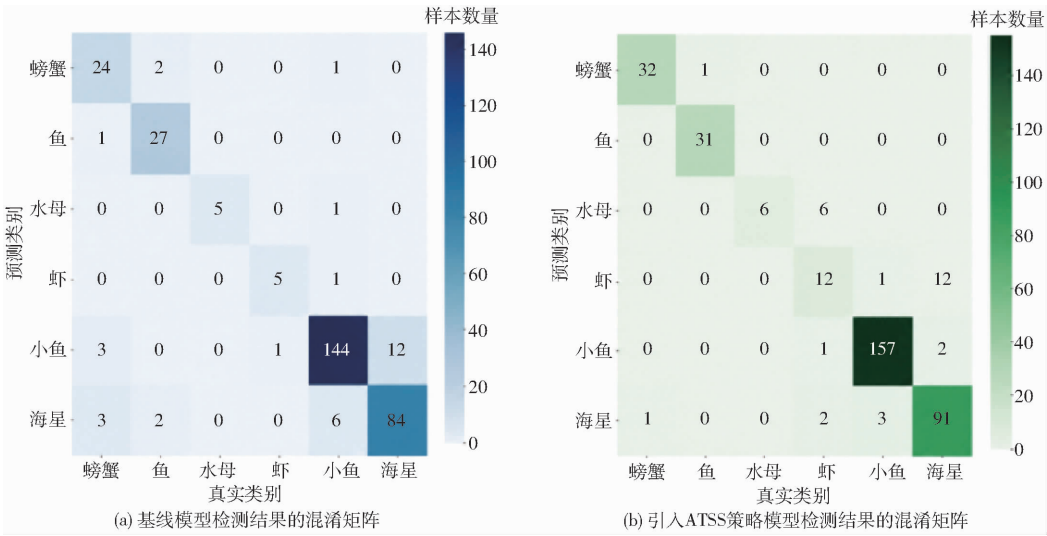


图 8 检测结果的混淆矩阵

Fig. 8 Confusion matrix of detection results

3.4 不同检测头的对比实验

为了验证 DyHead 以及其他检测头的有效性,分别在 BUD 和 UMDD 数据集上进行了实验。实验结果如表 2 所示。DyHead 相比 AsDDet^[22]、LADH^[23] 和 TransHead^[24], DyHead 的 mAP 表现更高,在 UMDD 数据集上达到了 82.2%,展现了其在

复杂场景下优秀的目标检测能力;同时其计算复杂度也更低,BUD 数据集上的浮点运算量仅为 8.6×10^9 ,UMDD 数据集上为 6.2×10^9 。表 2 表明 DyHead 不仅保持了高精度,还能有效控制计算成本。此外,DyHead 的参数量为 2.640×10^7 ,显著低于 Detect_FRM 的 2.293×10^8 ,在资源消耗方面同样存在优

表 2 YOLO v8s 改进不同检测头的检测结果比较

检测头	BUD		UMDD		参数量
	mAP50/%	浮点运算量	mAP50/%	浮点运算量	
AsDDet	65.3	3.28×10^{10}	53.6	3.42×10^{10}	8.050×10^7
LADH	58.8	1.98×10^{10}	66.9	4.24×10^{10}	2.970×10^7
TransHead	61.1	4.56×10^{10}	77.0	1.67×10^{10}	3.550×10^7
Detect_FRM	60.8	2.21×10^{10}	81.8	1.65×10^{10}	2.293×10^8
DyHead	70.6	8.60×10^9	82.2	6.20×10^9	2.640×10^7

势。DyHead 在 mAP50、计算复杂度和模型参数量 3 方面均有较好的表现,适合对检测性能和资源消耗都要求较高的应用场景。

3.5 损失函数的对比实验

表 3 对不同的损失函数进行了详细的实验对

比。WIoU v3s 具有更准确地捕捉对象真实边界、并降低复杂边界回归任务中回归误差的能力。在计算复杂度方面,WIoU v3s 的浮点运算量在 BUD 和 UMDD 数据集上低于 CIoU 和 GIoU。WIoU v3s 不仅提高了检测精度,还优化了计算资源的利用,提升了

表 3 YOLO v8s 引入不同损失函数后的检测结果比较

Tab.3 Comparison of detection results of YOLO v8s introducing different loss functions

损失函数	BUD			UMDD		
	mAP50/%	浮点运算量	精确率/%	mAP50/%	浮点运算量	精确率/%
+ CIoU	68.9	3.28×10^{10}	70.5	85.7	3.92×10^{10}	60.5
+ EIoU	69.2	3.98×10^{10}	69.1	85.8	3.36×10^{10}	59.1
+ GIoU	70.3	3.56×10^{10}	70.9	85.8	3.21×10^{10}	61.6
+ WIoU v1	71.6	2.21×10^{10}	73.1	86.1	2.57×10^{10}	65.1
+ WIoU v2	71.9	2.86×10^{10}	75.6	86.4	2.46×10^{10}	68.9
+ WIoU v3	72.2	2.61×10^{10}	75.9	86.6	2.06×10^{10}	69.7
+ WIoU v3s	72.4	2.06×10^{10}	76.1	86.7	1.98×10^{10}	70.1

模型的整体效率。在精确率指标上,WIoU v3s 在 BUD 和 UMDD 数据集上分别达到了 76.1% 和 70.1%。

表 3 表明,WIoU v3s 具有强大的准确检测目标物体的能力,且可在复杂场景中保持有效的分辨率。WIoU v3s 通过优化边界框的回归损失函数,弥补了改进前版本在处理小物体和遮挡问题的局限。通过考虑物体的形状和位置,使得模型在复杂的场景中仍能保持高效的检测能力。

3.6 消融实验

为了深入探讨各改进模块对 WLDA – YOLO v8

模型性能的贡献,在 UMDD 数据集上以 YOLO v8s 为基准进行了消融实验,各模块的性能提升效果如表 4 所示。

(1)引入 WIoU v3 模块后,模型的边界框回归精度得到显著提升。WIoU v3 专注于优化边界框的回归损失,尤其对形态复杂或边界模糊的目标表现突出。例如,在难以检测的小鱼等小目标类别上表现更为明显,mAP50 – 95 提高了 13.1 个百分点。模型的 mAP50 – 95 提升了 2.4 个百分点,这清晰地表明 WIoU v3 模块对于精确定位边界框与提升小目标检测准确性具有重要作用。

表 4 WLDA – YOLO v8 的消融实验结果

Tab.4 Results of ablation test

方法	mAP50 – 95/%						mAP50 – 95 平均值/%	内存占用 量/MB	参数量	浮点运算量
	螃蟹	鱼	水母	虾	小鱼	海星				
YOLO v8s	61.9	77.5	47.7	72.6	45.7	75.6	65.7	21.46	1.11×10^7	2.84×10^{10}
+ WIoU v3s	64.0	78.8	51.8	79.7	58.8	76.7	68.1	21.68	1.12×10^7	3.00×10^{10}
+ Dynamic ATSS	55.2	79.1	52.5	73.3	56.5	77.3	68.6	14.51	7.40×10^6	2.74×10^{10}
+ LEM	56.3	79.1	53.4	79.8	59.4	77.8	69.2	15.56	7.90×10^6	2.84×10^{10}

(2)通过加入 Dynamic ATSS 模块,模型获得了更加智能的样本选择机制。Dynamic ATSS 通过动态调整样本分配权重,有效地处理复杂场景与目标形态的多样性。尤其在鱼和海星这两个类别上表现出明显的性能提升,mAP50 – 95 分别提高到 79.1% 和 77.3%。

(3)通过引入 LEM 模块,模型的特征提取能力得到进一步优化,尤其在小目标检测方面表现显著。例如,小鱼类别的 mAP50 – 95 从 56.5% 提升到 59.4%,体现出 LEM 模块在捕获小目标细节和背景信息的有效性。此外,LEM 模块在其他类别,如水

母和虾类别的表现也非常出色,mAP50 – 95 分别提高至 53.4% 和 79.8%。

WLDA – YOLO v8 模型在真实水体场景中目标检测的性能如图 9 所示,所有类别目标都可以准确识别,特别是在复杂的环境背景和边界条件下,展现出稳定高效的检测能力。图 10 显示消融实验中的 F1 分数曲线、召回率曲线,各个模块的加入对模型的性能有不同的影响,WIoU v3 模块显著提高了 F1 分数,体现出精准定位的优势。Dynamic ATSS 模块主要提高了模型在高置信度区域的精确率。LEM 模块则明显改善了低置信度区域的召回率,体

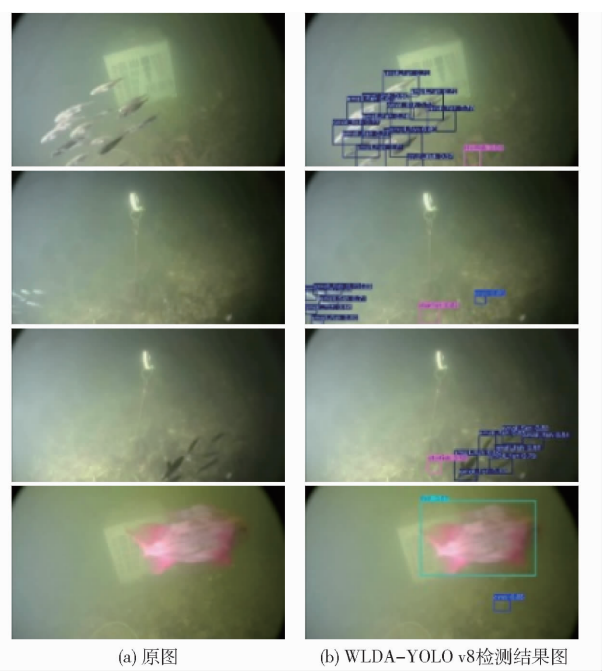


图 9 WLDA - YOLO v8 目标类别检测效果

Fig.9 WLDA - YOLO v8 target category detection effect diagrams

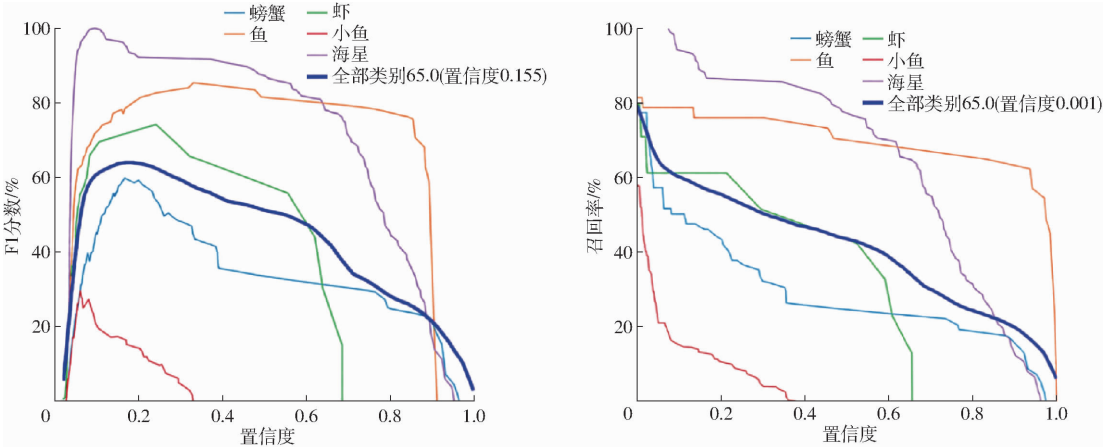


图 10 F1 分数曲线和召回率曲线

Fig.10 F1-score curves and recall curves

小型目标,由于其像素占比较小、特征信息有限且容易受到背景噪声干扰,检测难度仍然较高。总体来看,WLDA - YOLO v8 对不同类别的水下目标均具有一定的检测能力,但不同类别之间仍存在性能差异。

现了其对于细节特征捕捉的提升能力。

综上所述,消融实验验证了每个模块的引入都对模型性能有累计的正向促进作用。WIoU v3s 提升了定位精度,Dynamic ATSS 改善了训练样本选择策略以增强复杂场景下的表现,LEM 模块强化了特征提取以更好地检测小目标和复杂背景中的目标。最终构成的 WLDA - YOLO v8 模型相较基线共提高约 3.5 个百分点 mAP50 - 95,参数量仍控制在 8×10^6 级别,证明了各改进策略在保持模型轻量化的同时具有显著的性能增益。

3.7 不同检测模型对比

在对比多种检测模型时,采用自建 UMDD 数据集对各个模型进行了训练和测试,以验证其在水下目标检测中的有效性。图 10 给出了 WLDA - YOLO v8 在不同类别上的 F1 分数曲线和召回率曲线,反映了不同类别目标在不同置信度阈值下的检测表现。由图 10 可见,不同类别的检测性能存在一定差异,其中鱼类和海星类别的 F1 分数和召回率整体较高,而小鱼类别的检测性能相对较低。这表明对于

使用自建的 UMDD 数据集,对多个检测模型进行训练与测试,以验证它们在水下目标检测任务中的有效性。通过与 Faster R - CNN、YOLO v5s、YOLO v6^[25]、YOLO v7 等模型的对比实验,对 mAP、参数量、浮点运算量等指标进行分析,如表 5 所示。

表 5 不同方法在 UMDD 数据集上的检测结果比较

Tab.5 Comparison of detection results of different methods on UMDD dataset				
检测网络	主干网络	mAP50 - 95/%	参数量	浮点运算量
Faster R - CNN	ResNet50 - FPN	62.3	4.14×10^6	8.85×10^{10}
YOLO v5s	CSPDarkNet	65.1	9.1×10^6	1.38×10^{10}
YOLO v6	CSPDarkNet	64.8	1.72×10^7	2.41×10^{10}
YOLO v7	CSPDarkNet	75.2	1.15×10^7	3.10×10^{10}
YOLO v8s	CSPDarkNet	72.4	1.03×10^7	1.98×10^{10}
Boosting R - CNN	ResNet50	62.0	—	8.23×10^{10}
EFP - YOLO	CSPDarkNet	57.7	—	3.30×10^{10}
WLDA - YOLO v8	CSPDarkNet 变体	77.9	7.9×10^6	1.59×10^{10}

与 YOLO v8s 相比, WLDA - YOLO v8 在 mAP50 - 95 指标上提升了 5.5 个百分点。同时参数量降低至 7.9×10^6 , 浮点运算量由 1.98×10^{10} 降至 1.59×10^{10} , 大幅降低了计算复杂度。与 Faster R - CNN 相比, 由于采用双阶段检测方式, 其 mAP50 - 95 仅 62.3%, 且参数量高达 4.14×10^6 , 计算复杂度 (浮点运算量) 高达 8.85×10^{10} , 明显不适用于高效实时检测任务。此外, YOLO v5s 和 YOLO v6 尽管计算量相对较低, 但 mAP50 - 95 相对较低, 精度要求难以达到。综上, WLDA - YOLO v8 在保持轻量化的同时, 仍能提供最高精度, 充分证明其有效性和先进性。

图 11 展示了 WLDA - YOLO v8 与 WSSOD - YOLO v8^[26] 和 2PCC - YOLO v8^[27] 在损失值和 mAP50 - 95 曲线上的对比。从 mAP 变化曲线来看, WLDA - YOLO v8 在整个训练过程中始终保持较快的精度增长, 特别是在训练初期, 表明模型能迅速学习到有效特征。在约 25 轮次之后, mAP 逐渐趋于稳定, 最终达到 77.9%, 显著高于 WSSOD - YOLO v8 和 2PCC - YOLO v8。大约 25 轮次之前, WLDA - YOLO v8 的损失值下降更快, 在较短时间内达到较低的损失值, 说明训练稳定性好, 有更快的收敛速度。

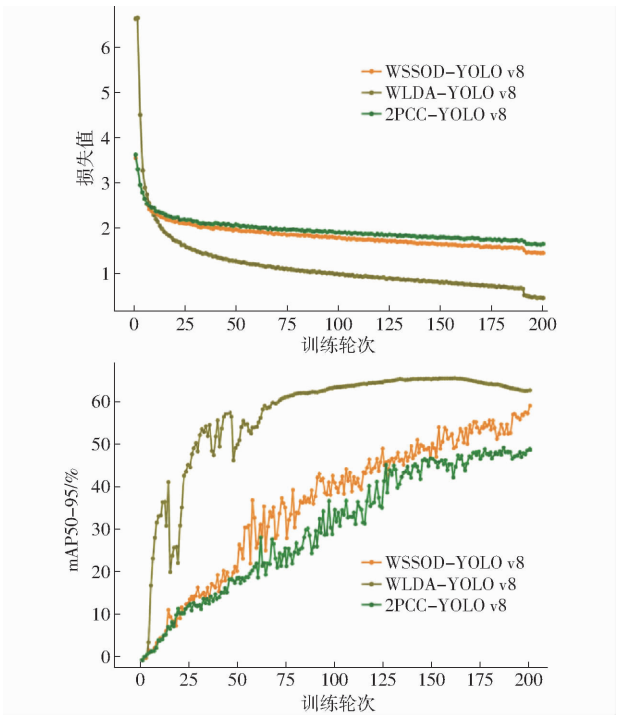


图 11 改进的损失曲线和改进的 mAP 曲线比较

Fig. 11 Loss and improved mAP curves comparison

从图 12 可以明显看出, WLDA - YOLO v8 在不同延迟下均取得了显著高于其他模型的 mAP50 - 95。这表明该模型在检测精度和推理速度之间达到了更优的平衡。在延迟为 1.0 ms 时, 其 mAP50 - 95 已经接近 50%, 远超其他对比模型。当延迟提高到

3.0 ms 时, WLDA - YOLO v8 的 mAP 进一步攀升至 52% 以上, 仍然保持最高性能。这一趋势证明了其更高的计算效率和检测性能。DSW - YOLO v8n^[28] 和 YOLO v5 等模型虽然在低延迟下表现优异, 但在延迟超过 2.5 ms 时, 其 mAP 增幅明显减缓。

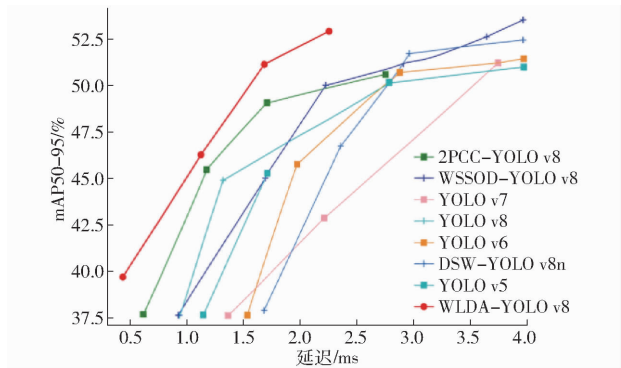


图 12 WLDA - YOLO v8 与其他 YOLO 模型 mAP 对比

Fig. 12 WLDA - YOLO v8 and other YOLO models mAP comparison

为了定性对比检测结果, 研究了 3 组信号条件不佳或目标与背景颜色相近复杂图像, 包含小型和重叠目标, 进行模型性能测试。蓝色标记表示漏检, 红色标记表示误检。图 13 展示了 3 种方法的结果预测。WLDA - YOLO v8 在检测小型目标和复杂重叠目标时表现优异, 尤其在检测鱼类和螃蟹类等小型水下生物时, WLDA - YOLO v8 明显减少了漏检和误检现象。

综上所述, WLDA - YOLO v8 不仅提升了小目标和复杂背景下的检测精度, 还有效解决了 2PCC - YOLO v8 和 WSSOD - YOLO v8 在检测重叠目标时的局限性。这些实验结果验证了该模型在水下目标检测中的性能优势, 无论是小目标、重叠目标还是复杂背景环境下, 改进模型均表现出了卓越的检测效果, 证明了其实际应用中的适用性。

4 结论

(1) 本文提出了一种基于改进 YOLO v8 模型的海洋生物智能检测算法, 针对水下环境中的复杂挑战进行了深入研究并提出了创新性解决方案。通过设计轻量级的多尺度卷积模块 LEM、动态自适应训练样本选择策略 ATSS, 以及优化后的损失函数 WIoU v3s, 显著提升了模型在水下目标检测任务中的精度和效率。

(2) 首先, LEM 模块的设计使得模型能够更有效地捕捉多尺度特征, 特别是在小目标检测中表现出色。这一模块不仅减少了模型的计算复杂度, 还提高了对细节的敏感性, 从而更好地适应了水下环境中的复杂背景和小目标特征。其次, 动态自适应

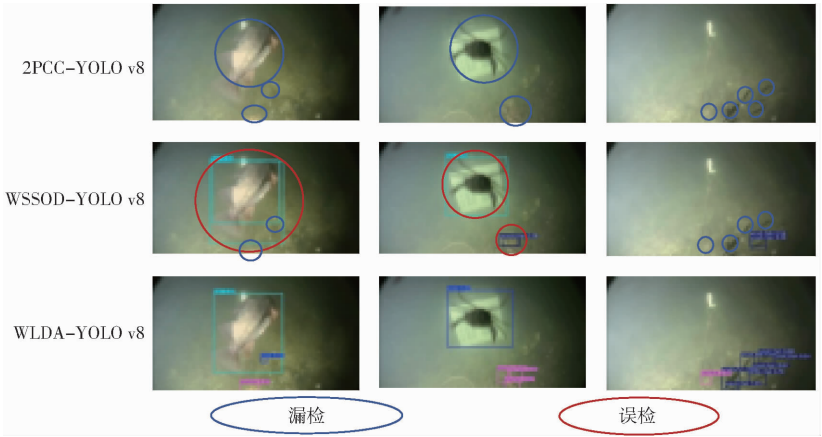


图 13 不同方法的检测结果对比

Fig. 13 Comparison of detection results of different methods

训练样本选择策略 ATSS 的融合,使模型能够根据目标的复杂性灵活调整训练样本,提升了模型的检测精度和泛化能力。这一策略在处理多尺度和复杂背景目标时,显著增强了模型的适应性,确保了模型

在不同环境下的鲁棒性。最后,改进的损失函数 WIoU v3s 针对水下检测任务的特殊需求进行了优化,使得边界框的回归更加精确,模型的收敛速度和稳定性得到了显著提升。

参 考 文 献

[1] Sun Y, Zheng W, Du X, et al. Underwater small target detection based on yolox combined with mobilevit and double coordinate attention[J]. Journal of Marine Science and Engineering, 2023, 11(6): 1178.

[2] Kovacova M, Oláh J, Popp J, et al. The algorithmic governance of autonomous driving behaviors: multi-sensor data fusion, spatial computing technologies, and movement tracking tools[J]. Contemporary Readings in Law and Social Justice, 2022, 14(2): 27–45.

[3] Zhao L, Yun Q, Yuan F, et al. YOLO v7 – CHS: an emerging model for underwater object detection[J]. Journal of Marine Science and Engineering, 2023, 11(10): 1949.

[4] 李路,孙超奇,周玉凡,等. 基于 FAS – YOLO v8n 的爬岸上草小龙虾多源图像融合识别方法[J]. 农业机械学报,2025, 56(8):526–534.

Li Lu, Sun Chaoqi, Zhou Yufan, et al. Multi-source image fusion recognition of cra wling bank grass crayfish based on FAS – YOLO v8n[J]. Transactions of the Chinese Society for Agricultural Machinery,2025,56(8):526–534. (in Chinese)

[5] Zeng L, Sun B, Zhu D. Underwater target detection based on Faster R – CNN and adversarial occlusion network [J]. Engineering Applications of Artificial Intelligence, 2021, 100: 104190.

[6] Song S, Zhu J, Li X, et al. Integrate MSRCR and Mask R – CNN to recognize underwater creatures on small sample datasets [J]. IEEE Access, 2020, 8: 172848–172858.

[7] Yu Y, Zhao J, Gong Q, et al. Real-time underwater maritime object detection in side-scan sonar images based on transformer – YOLO v5[J]. Remote Sensing, 2021, 13(18): 3555.

[8] Song S J, Zhu J X. Research on underwater biological target recognition based on Mask R – CNN and transfer learning[J]. Application Research of Computers,2020,37(S2):386–388,391.

[9] Guo A, Sun K, Zhang Z. A lightweight YOLO v8 integrating FasterNet for real-time underwater object detection[J]. Journal of Real-Time Image Processing, 2024, 21(2): 49.

[10] Wang C, Liao H, Wu Y, et al. CSPNet: a new backbone that can enhance learning capability of CNN[C] //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2020;14–19.

[11] Wu J, Chen J, Lu Q, et al. U – ATSS: a lightweight and accurate one-stage underwater object detection network[J]. Signal Processing: Image Communication, 2024, 126: 117137.

[12] Yang J, Sun L, Wang G. Improving the underwater object detection algorithm for YOLO v5 [C] // Third International Conference on Electronic Information Engineering and Data Processing (EIEDP 2024), 2024: 1812–1818.

[13] Feng J, Jin T. CEH – YOLO: a composite enhanced YOLO-based model for underwater object detection[J]. Ecological Informatics, 2024, 82: 102758.

[14] Terven J, Córdova-esparza D M, Romero-gonzález J A. A comprehensive review of YOLO architectures in computer vision: from YOLO v1 to YOLO v8 and YOLO – NAS[J]. Machine Learning and Knowledge Extraction, 2023, 5: 1680–1716.

[15] 王旺,王福顺,张伟进,等. 基于改进 YOLO v8s 的羊只行为识别方法[J]. 农业机械学报, 2024, 55(7):325–335, 344.

Wang Wang,Wang Fushun, Zhang Weijin, et al. Sheep behavior recognition method based on improved YOLO v8s[J].

Transactions of the Chinese Society for Agricultural Machinery,2024,55(7):325 – 335,344. (in Chinese)

[16] Zhang M, Xu S, Song W, et al. Lightweight underwater object detection based on YOLO v4 and multi-scale attentional feature fusion[J]. Remote Sensing, 2021, 13(22): 4706.

[17] Yifan L, Yiping S, Xuezhi L, et al. Research on the improvement of garbage detection and classification algorithm based on YOLO v7 – tiny [C] // 2024 2nd International Conference on Intelligent Perception and Computer Vision (CIPCV). IEEE, 2024: 121 – 127.

[18] 袁红春,李春桥. 基于改进可变形卷积的 FDC – YOLO v8 水下生物目标检测方法研究[J]. 农业机械学报,2024,55(11):140 – 146.

Yuan Hongchun, Li Chunqiao. Research on FDC – YOLO v8 underwater biological object detection method improved by deformable convolution[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(11): 140 – 146. (in Chinese)

[19] Yu Q, Han Y, Han Y, et al. Enhancing YOLO v5 performance for small-scale corrosion detection in coastal environments using iou-based loss functions[J]. Journal of Marine Science and Engineering, 2024, 12(12): 2295.

[20] 刘进一,晏伏山,董赫,等. 基于改进 MBS – YOLO v8 的火龙果目标检测[J]. 农业机械学报,2025,56(5):425 – 432.

Liu Jinyi, Yan Fushan, Dong He, et al. Pitaya fruit target detection and localization method based on improved MBS – YOLO v8[J]. Transactions of the Chinese Society for Agricultural Machinery,2025,56(5):425 – 432. (in Chinese)

[21] Chen W, Zhuang T, Zhang Y, et al. YOLO – UOD: an underwater small object detector via improved efficient layer aggregation network[J]. IET Image Processing, 2024, 18(9): 2490 – 2505.

[22] 吴立栋,彭宗举,李欣,等. 基于样本迭代融合的海洋生物检测[J]. 激光与光电子学进展, 2023, 60(2): 0220001.

Wu Lidong, Peng Zongju, Li Xin, et al. Marine organism detection based on sample iterative fusion[J]. Laser & Optoelectronics Progress, 2023, 60(2): 0220001. (in Chinese)

[23] Zhang T, Zhang J, Pan P, et al. YOLO – RRL: a lightweight algorithm for PCB surface defect detection[J]. Applied Sciences, 2024, 14(17): 7460.

[24] Cai H, Xu L, Guo L. Feature fused detector for small objects [C] // 2022 4th International Conference on Advances in Computer Technology, Information Science and Communications (CTISC). IEEE, 2022: 1 – 6.

[25] Ouyang W, Wei Y, Liu G. A lightweight object detector with deformable upsampling for marine organism detection[J]. IEEE Transactions on Instrumentation and Measurement, 2024, 73: 1 – 9.

[26] Yang X, She H, Lou M, et al. A joint ship detection and waterway segmentation method for environment-aware of USVs in canal waterways[J]. IEEE Transactions on Automation Science and Engineering, 2025, 22: 2132 – 2144.

[27] Miao L, Tian Y. Detection of small underwater organisms based on improved YOLO v8[J]. IAENG International Journal of Computer Science, 2024, 51(8):1020 – 1026.

[28] Liu Q, Huang W, Duan X, et al. DSW – YOLO v8n: a new underwater target detection algorithm based on improved YOLO v8n[J]. Electronics, 2023, 12(18): 3892.

(上接第 31 页)

[111] Cui Y, Yuh J. A unified adaptive force control of underwater vehicle-manipulator systems (UVMS) [C] // Proceedings of the 2003 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS); IEEE, 2003: 553 – 558.

[112] Olguin-Diaz E, Arechavaleta G, Jarquin G, et al. A passivity-based model-free force-motion control of underwater vehicle-manipulator systems[J]. IEEE Transactions on Robotics, 2013, 29(6): 1469 – 1484.

[113] Barbalata C, Dunnigan M W, Petillot Y. Position/force operational space control for underwater manipulation[J]. Robotics and Autonomous Systems, 2018, 100: 150 – 159.

[114] Liu H, Yu M, Zhang W, et al. Jamming-free impedance control of UVMS based on contact mechanisms for peg-in-hole tasks [C] // Proceedings of the 2024 36th Chinese Control and Decision Conference (CCDC); IEEE, 2024: 2450 – 2455.

[115] Taira Y, Sagara S, Oya M. Impedance control based on error feedback for the manipulator of an underwater vehicle-manipulator system[J]. Artificial Life and Robotics, 2023, 28(4): 830 – 849.

[116] Wu B, Liu X, Hui N, et al. Fuzzy adaptive impedance control method for underwater manipulators based on bayesian recursive least squares and displacement correction[J]. Machines, 2025, 14(1): 39.

[117] Carlucho I, De Paula M, Wang S, et al. Adaptive low-level control of autonomous underwater vehicles using deep reinforcement learning[J]. Robotics and Autonomous Systems, 2018, 107: 71 – 86.

[118] Carlucho I, Stephens D W, Barbalata C. An adaptive data-driven controller for underwater manipulators with variable payload [J]. Applied Ocean Research, 2021, 113: 102726.