

文章编号 1004-924X(2026)03-0481-16

基于动态特征增强的水下鱼类目标检测

朱晓龙^{1,2*}, 陈郁隈¹, 王嘉宇¹, 郭海涛¹, 陈祥子^{1,2}

(1. 海南热带海洋学院 崖州湾创新研究院, 海南 三亚 572022;

2. 海南热带海洋学院 海洋科学技术学院, 海南 三亚 572022)

摘要: 水下鱼类的高效监测是海洋生态系统保护、生物多样性评估和水产资源可持续管理的基础。为应对水下复杂环境因素对检测稳定性和检测效率的影响, 基于 YOLOv8n 框架提出一种动态特征增强模型 FDN-YOLO (Fish Detection Network YOLO)。首先, 构建多尺度可变形感受野模块 (Multi-scale Deformable Receptive Field, MDRF), 通过自适应调节有效感受野, 使主干网络更充分地表征不同形态与尺度的鱼类目标。其次, 设计下采样轻量化空间重排深度可分离模块 (Lite Space-to-Depth Depthwise Separable, Lite SPD-DS), 在控制计算开销的同时, 保留下采样阶段的细粒度空间线索。最后, 提出将自适应 IoU 与 Varifocal Loss 融合的损失函数 (Adaptive IoU-aware Varifocal Loss, AIVF Loss), 以强化对高质量定位样本的学习能力, 并缓解类别与样本分布不均带来的训练偏置。基于 TF-DET 数据集的实验结果表明, FDN-YOLO 的 mAP50 与 mAP50:95 分别提升 2.8% 与 2.1%, 参数量与计算量分别降低 13.3% 与 16.0%。通过对比泛化实验进一步表明, FDN-YOLO 在准确性、效率与稳定性之间的优越平衡性能, 具备在生态调查与数据驱动的海洋资源管理中的应用潜力。

关键词: 水下检测; 动态增强; 深度学习; 卷积神经网络; YOLOv8n

中图分类号: TN911.7; TP391 **文献标识码:** A

doi: 10.37188/OPE.20263403.0481

CSTR: 32169.14.OPE.20263403.0481

Dynamic feature-augmented for detection of underwater fish targets

ZHU Xiaolong^{1,2*}, CHEN Yuwei¹, WANG Jiayu¹, GUO Haitao¹, CHEN Xiangzi^{1,2}

(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)

* Corresponding author, E-mail: zhuxiaolonglong@sina.cn

Abstract: Efficient monitoring of underwater fish species is essential for marine ecosystem conservation, biodiversity assessment, and the sustainable management of aquatic resources. To address reductions in detection robustness and efficiency under complex underwater conditions, a dynamic feature enhancement model, termed Fish Detection Network YOLO (FDN-YOLO), is proposed based on the YOLOv8n

收稿日期: 2025-10-10; **修订日期:** 2025-10-30.

基金项目: 国家自然科学基金资助项目 (No. 42266005); 海南省院士创新平台科研项目 (No. YSPTZX202507); 三亚市科技创新专项资助项目 (No. 2022KJCX83); 海南热带海洋学院崖州湾创新研究院重大科技计划资助项目 (No. 2022CXYZD003, No. 2023CXYZD001)

framework. First, a Multi-scale Deformable Receptive Field (MDRF) module is incorporated to adaptively regulate the effective receptive field, thereby improving backbone representations of fish targets with diverse shapes and scales. Second, a lightweight down-sampling module, Lite Space-to-Depth Depthwise Separable (Lite SPD-DS), is designed to preserve fine-grained spatial cues during subsampling while maintaining low computational cost. Third, an Adaptive IoU-aware Varifocal Loss (AIVF Loss) is introduced by integrating adaptive IoU weighting with Varifocal Loss to strengthen the learning of high-quality localization samples and mitigate training bias caused by class and sample imbalance. Experiments on the TF-DET dataset show that FDN-YOLO increases mAP50 and mAP50:95 by 2.8% and 2.1%, respectively, while reducing parameters and computational complexity by 13.3% and 16.0%. Additional comparative and generalization experiments further confirm that FDN-YOLO achieves a favorable trade-off among accuracy, efficiency, and robustness, highlighting its potential for ecological monitoring and data-driven marine resource management.

Key words: underwater detection; dynamic enhancement; deep learning; convolutional neural network; YOLOv8n

1 引 言

在智能化海洋观测系统中,水下鱼类检测作为核心任务之一,广泛应用于生态调查、水产养殖管理及环境评估等领域^[1-2]。随着水下观测对自动化与高精度要求的不断提升,基于计算机视觉的目标检测技术正逐渐成为关键支撑手段。得益于深度卷积神经网络的发展,近年来地面场景中的目标检测性能已显著提升^[3-4]。然而,将这些方法直接迁移至水下环境应用仍面临诸多挑战,其主要原因在于水下环境独特的物理限制。相较于陆地目标检测,水下图像感知严重受限于光吸收、散射以及颜色衰减等因素^[5],导致图像普遍存在纹理模糊、对比度低和色彩失真等问题。此外,悬浮颗粒、珊瑚礁等复杂背景以及动态水体条件进一步引入大量视觉噪声^[6],阻碍了特征信息的稳定提取。同时,水下鱼类目标通常具有尺度小、密集成群和姿态变化大的特点^[7-8],使传统检测模型在定位精度和类别判别方面面临更大挑战。

为缓解上述问题,研究人员相继提出针对水下场景的网络结构改进、损失函数优化以及注意力机制引入等策略,均取得了一定进展。Li等^[9]构建了基于YOLOv5的改进模型,引入可变形卷积以增强特征提取能力,并采用自适应阈值模块降低水下鱼类的漏检率。相较于原始结构,该方

法在检测精度和稳定性方面均有所提升。苗力恒等^[10]提出将多尺度注意力机制引入YOLOv8,并构建多尺度卷积模块增强水下目标的特征提取能力。Li等^[11]在YOLOv5基础上融合了改进的C3CA模块提升特征提取效果,扩展了检测头结构以增强多尺度检测能力,并引入EIOU损失函数以优化定位精度并加速收敛过程。Zhu等^[12]提出将EfficientDet-D3模块嵌入YOLOv4,通过优化模型主干与颈部的连接方式,实现对鱼类的有效检测。Lu等^[13]以VGG16为基础,构建了融合多全连接层与新激活函数的检测网络,实现了对六种鱼类的准确分类。Xu等^[14]提出基于SE-ResNet152的改进模型,通过引入Focal Loss算法和迁移学习机制,解决了样本不足与模型过拟合问题,实现了对三种鱼类的有效识别。Iqbal等^[15]通过调整AlexNet的层数、训练轮数及参数,构建检测模型,采用无背景干扰的样本图像进行训练,最终达到了在简单环境下对六种鱼类的准确判别。Smadi等^[16]整合了AlexNet,VGG-Net等的网络特征,重构全连接网络,并结合RM-Sprop,Adam和Adamax等优化器,改善了对鱼类的检测性能。由此可知,现有方法在理想或简化条件下对个体目标及其种类的检测和识别已取得较好成果,然而在复杂水下环境中仍面临显著挑战。水下小目标重叠、姿态多变及背景复杂等现象,常导致模型出现误检、漏检与分类不佳等问题。此外,在检测精度与模型复杂度之间取得

有效平衡依然是当前方法较难克服的瓶颈,进而影响其稳定性与实用性。因此,设计一种兼顾轻量化和稳定性的水下鱼类检测算法,对于推进智能化海洋观测技术的发展具有重要的应用价值。为此,本文提出一种新型改进模型FDN-YOLO (Fish Detection Network YOLO),即在有限计算资源条件下提升特征建模能力与判别可靠性。该模型从特征提取机制、下采样结构以及损失函数构建3个关键环节进行针对性的优化设计。

针对水下鱼类目标尺度变化明显且姿态不规则的问题,提出一种多尺度可变形感受野模块(Multi-scale Deformable Receptive Field, MDRF)。该模块通过多分支可变形卷积与共享偏移建模,增强网络对不同尺度与形态目标的自适应感知能力。

针对下采样过程中特征信息丢失对检测性能造成的不利影响,提出一种轻量化空间重排深度可分离模块(Lite Space-to-Depth Depthwise Separable, Lite SPD-DS)。该模块通过空间至通道重排、深度可分离卷积以及轻量级通道注意力,实现高效特征重构和动态特征增强。

为提升模型对目标定位质量的敏感性,提出一种自适应IoU感知的损失函数(Adaptive IoU-aware Varifocal Loss, AIVF Loss)。通过引入定位质量驱动的动态权重调节机制,引导模型更加关注高质量预测结果,提升检测稳定性。

2 基础模型

YOLOv8n凭借其在架构设计与算法层面的优化,已成为当前应用最为广泛的目标检测模型之一。这些改进不仅促使其派生出多个高效变体,也显著提升了模型在多种应用场景中的检测精度与效率。YOLOv8n整体结构由3个主要功能模块组成:特征提取的主干、特征融合的颈部和用于目标预测的检测头^[17]。在主干部分,该模型融合了传统卷积模块、C2f模块和SPPF模块^[18]。C2f模块通过引入梯度分流连接与残差子模块堆叠,强化特征表达的同时保持网络轻量化。SPPF模块依次采用3组不同的最大池化操作,并将特征图拼接后传递至后续层。在颈部部分,YOLOv8n引入的PAN-FPN^[19]结构实现了底层空间特征与高层语义信息之间的双向流动,增

强了对不同尺度目标的定位能力。在检测头部分,其采用解耦设计,将目标分类与位置回归两个子任务分离,有助于各自特征更精准的学习。

3 优化方法

3.1 MDRF 模块

为提升在复杂水下环境中的感知能力,提出具备多尺度可变形感受野的MDRF模块。该模块融合了多尺度可变形采样、动态特征增强与自适应特征融合3种机制,旨在提升对小目标和低对比度目标的检测性能。

MDRF模块结构如图1所示,它由多个不同尺度的可变形卷积分支构成,卷积核尺寸分别为 1×1 、 3×3 和 5×5 。通过引入可变形卷积Deformable convolution^[20]机制,模型能够灵活调整感受野形状,使其更好地贴合输入图像中的关键结构,从而提升对形态多变和边缘模糊目标的适应能力与检测精度。各分支分别从不同尺度提取局部细节与全局上下文信息,多尺度感知能力显著增强。同时,采用共享偏移量预测机制,使各分支在不显著增加计算开销的前提下动态调整感受野,增强对目标几何结构与形变的敏感性^[21]。为了进一步实现尺度感知的特征增强与动态权重分配,所有分支输出特征首先进行通道拼接,随后引入基于压缩-激励机制(Squeeze-and-Excitation, SE)的轻量级通道注意力模块,实现不同尺度语义信息的自适应重标定。最后,通过一个 1×1 卷积层完成通道压缩,使输出特征在保持信息表达能力的同时与主干网络特征维度对齐,从而确保模块结构紧凑、高效。

假设输入特征图为 $\mathbf{X} \mathbf{R}^{C \times H \times W}$,其中 C, H 和 W 分别表示通道数、高度和宽度。 S 表示每个分支中可变形卷积的采样点数量, Q 表示分支组的数量,则多尺度可变形卷积分支提取的特征可表示为:

$$y_q = \sum_{s=1}^S m_{qs} x_q(p_0 + p_s + \Delta p_{qs}), \quad (1)$$

$$F_{\text{cat}} = \text{concat}(\left| y_1, y_2, \dots, y_Q \right|, \text{axis} = -1), \quad (2)$$

其中: P_0, P_s 分别表示当前像素点、第 s 个预定义采样点的位置, Δp_{qs} 表示基于第 q 个子组中采样点 P_s 空间坐标计算得到的偏移量, x_q 和 y_q 分别为子组处理前后的输入与输出特征。 m_{qs} 表示第 q

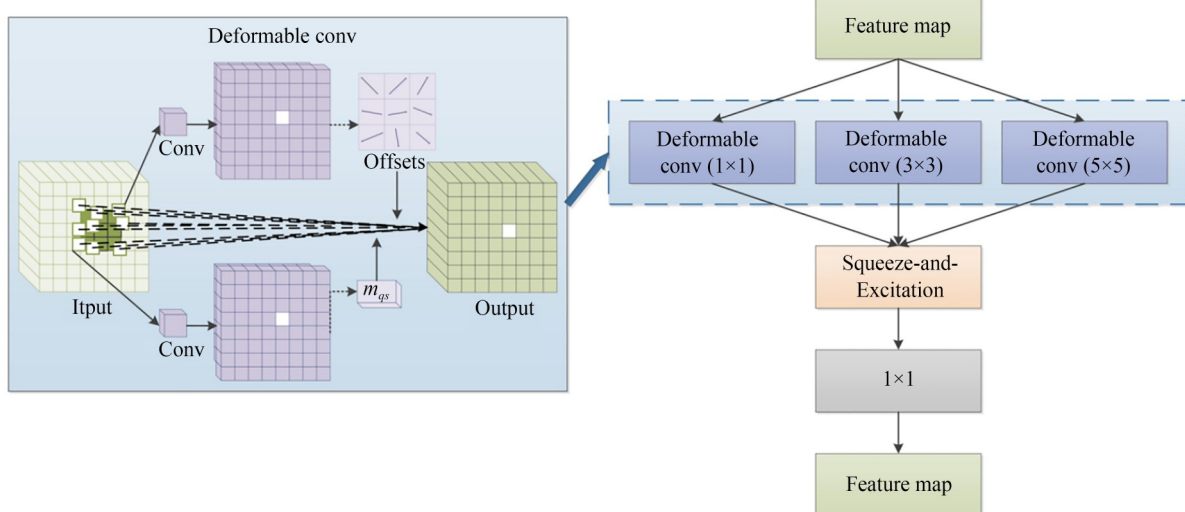


图1 MDRF 模块结构

Fig. 1 Structure of MDRF module

个子组中第 s 个采样点对应的调整权重。通过多尺度可变形卷积提取得到的所有特征拼接后记作为 F_{cat} 。为了进一步实现动态特征增强与通道重要性的自适应重标定,引入了SE注意力机制,如下所示:

$$z(c) = \frac{1}{H \cdot W} \sum_{i=1}^H \sum_{j=1}^W F_{cat}(c, i, j), \quad (3)$$

$$s = \sigma(W_2 \cdot \delta(W_1 \cdot z)), \quad (4)$$

$$F_{out}(c, i, j) = s(c) \cdot F_{cat}(c, i, j), \quad (5)$$

其中: σ 和 δ 分别表示 ReLU 和 Sigmoid 函数, W_1 和 W_2 表示可训练权重。最后,融合后的特征图通过一次 1×1 逐点卷积投影回原始的通道维度,以保持特征结构的一致性并控制模型复杂度。

3.2 Lite SPD-DS 模块

为提升特征下采样的效率并在模型中保留语义完整性,提出了 Lite SPD-DS 模块。如图 2 所示,该模块采用紧凑的结构设计,融合了空间重组、动态特征增强与轻量级语义融合机制,旨在降低计算开销的同时尽可能保留关键信息。

该模块首先通过 Space-to-Depth^[22] 算法对输入特征图进行空间压缩,同时将局部精细上下文信息编码至通道维度,实现空间与语义信息的有效重组。这种空间与通道混合设计能够在尽量减少信息损失的前提下,实现高效的下采样操作。随后,模块引入了 Depthwise separable convolution^[23],用于分离空间滤波与通道混合,从而

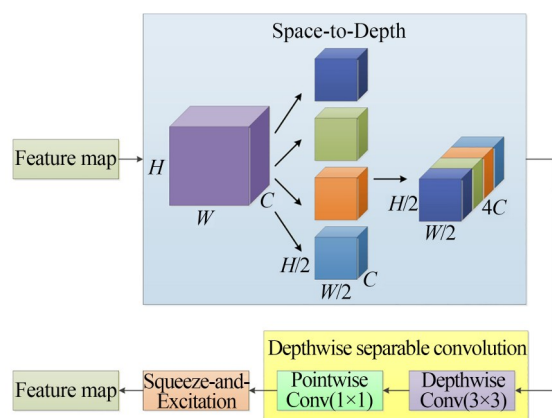


图2 Lite SPD-DS 模块结构

Fig. 2 Structure of Lite SPD-DS module

大幅降低参数数量与浮点运算开销。通过逐点卷积实现通道间特征的融合,以低复杂度实现语义整合。为进一步实现自适应的特征增强,模块引入了SE注意力机制,动态提升具有判别性的通道响应,同时抑制冗余信息,从而实现特征的重点强化与语义选择。

假设下采样因子为 r ,通过 Space-to-Depth 算法将每个 $r \times r$ 空间块重排为多个通道,其表达如下:

$$X_{SPD} \in \mathbb{R}^{C \cdot r^2 \times H/r \times W/r} \in \mathbb{R}^{4C \times H/2 \times W/2}, \quad (6)$$

$$X_{SPD}(c \cdot r^2 + i \cdot r + j, h, w) = X(c, r \cdot h + i, r \cdot w + j), \quad (7)$$

其中: i 和 j 分别表示每个 $r \times r$ 空间块内的局部空

间偏移。为在保留局部特征表示的同时降低计算开销,引入深度可分离卷积。该结构首先通过一个深度卷积 Depthwise Conv 在每个通道上独立执行空间滤波。随后,通过逐点卷积 Pointwise Conv 进行通道融合,并将输出特征映射至 C' 维度。该过程可表示为:

$$X_{PW} = PWConv(DWConv(X_{SPD})), \quad (8)$$

其中: $DWConv$ 和 $PWConv$ 分别表示 Depthwise Conv 和 Pointwise Conv 运算。随后,采用 SE 注意力机制提取全局上下文信息,并将其压缩至目标通道。该通道信息随后被送入一个包含两层全连接层的神经网络中生成 s 。最终,经过动态增强后的输出 X'' 可表示如下:

$$Z_c = \frac{1}{H' \cdot W'} \sum_{i=1}^{H'} \sum_{j=1}^{W'} X_{PW}(c, i, j), \quad (9)$$

$$s = \sigma(W_2 \cdot \delta(W_1 \cdot z)), \quad (10)$$

$$X'' = s \cdot X_{PW}(c), \quad (11)$$

其中: $H' = H/2$, $W' = W/2$, σ 和 δ 分别表示 ReLU 和 Sigmoid 函数。

3.3 AIVF Loss 函数

在复杂的水下环境中,检测模型常常需要面临类别不平衡和小目标判别的问题。YOLOv8n 中原始的分类损失函数将所有正负样本视为二元标签,未能考虑预测框的定位质量,导致分类置信度无法准确反映候选框的空间精度,从而限制了检测置信度分数的判别性与可靠性。为克服这一局限,Varifocal Loss^[24]被提出。该算法通过分配连续标签,使模型在分类学习过程中能够区分预测结果的空间质量。此外,Varifocal Loss 能够为正负样本生成更为可靠的梯度信号,有效缓解严重类别不平衡所带来的训练困难。其数学表达形式如下:

$$L = -q \cdot \log(p) - (1-q)p^\gamma \cdot \log(1-p), \quad (12)$$

其中: p 表示目标类别的预测值, q 表示目标的 IoU 得分。调制因子 γ 用于抑制易分负样本对梯度的影响,从而提升训练稳定性。尽管具有上述优势,但其对所有正样本采用统一加权策略,未能根据 IoU 水平区分处理,从而限制了模型对高质量定位样本的敏感性。

为进一步增强模型对高精度定位正样本的响应能力,提出 AIVF Loss,其通过引入非

线性调制因子 $\alpha(\text{IoU})$, 根据定位质量动态调整正样本损失的权重。AIVF Loss 的计算公式如下:

$$\alpha(\text{IoU}) = \delta + (1-\delta) \cdot \text{IoU}^\beta, \quad (13)$$

$$L_{\text{AIVF}} = -\alpha(\text{IoU}) \cdot q \cdot \log(p) - (1-q)p^\gamma \cdot \log(1-p), \quad (14)$$

其中: $\delta(0, 1)$ 为防止梯度消失的响应权重, β 控制对 IoU 预测的调节因子。该损失设计强化了对高精度定位样本的训练贡献,同时抑制了噪声正样本的影响。

为进一步说明 AIVF Loss 在优化机制上的效果,对其梯度特性进行推导,表达形式如下:

$$\frac{\partial L_{\text{AIVF}}}{\partial p} = -(1-q) \left[\gamma p^{\gamma-1} \log(1-p) + \frac{p^\gamma}{1-p} \right] - \frac{\alpha(\text{IoU}) \cdot q}{p}. \quad (15)$$

当 q 趋近于 1 时,表明当前为高质量样本。此时 $\alpha(\text{IoU})$ 近似为 $(\delta + 1)$, 可得:

$$\frac{\partial L_{\text{AIVF}}}{\partial p} \approx -\frac{(\delta + 1)}{p}. \quad (16)$$

同理,当 q 趋近于 0 时,表明当前为低质量样本。此时 $\alpha(\text{IoU}) \cdot q$ 近似为 0, 可得:

$$\frac{\partial L_{\text{AIVF}}}{\partial p} \approx -\left[\gamma p^{\gamma-1} \log(1-p) + \frac{p^\gamma}{1-p} \right]. \quad (17)$$

由上述表达式可知, AIVF Loss 在高 IoU 区域显著放大梯度幅值,使得高质量样本获得更强的置信度提升驱动,从而提高预测得分与定位精度的一致性。相反, AIVF Loss 在低 IoU 区域会自动降低学习强度,抑制噪声样本对训练过程的干扰。

相比于 Varifocal Loss 和 Quality Focal Loss, AIVF Loss 通过 $\alpha(\text{IoU})$ 引入非线性 IoU 感知调制项,实现质量优先梯度分配策略,使高质量样本被强化、低质量样本被抑制。然而, Varifocal Loss 的 IoU 得分和权重为线性关系,无法根据样本质量高低形成明显的梯度分离。Quality Focal Loss 则是基于预测误差驱动的损失函数,其梯度幅值受预测误差影响,而非真实质量驱动,无法根据 IoU 质量分配梯度的损失。

当 q 趋近于 1 时, AIVF Loss 的梯度幅值明

显大于 Varifocal Loss 的 $-q/p$ 和 Quality Focal Loss 的误差型梯度,使高质量样本获得更强学习驱动力;当 q 趋近于 0 时, AIVF Loss 随 $\alpha(\text{IoU}) \cdot q$ 趋近于 0 而自动衰减,学习力度被极大抑制。但 Varifocal Loss 仍保留缓慢衰减的梯度, Quality Focal Loss 甚至会因误差放大导致部分伪正样本被学习。综上所述, AIVF Loss 不依赖单一加权因子或置信度误差调节,而是基于 IoU 质量分布对梯度进行重构,实现动态调整的作用。

3.4 FDN-YOLO 检测模型

FDN-YOLO 的整体框架如图 3 所示。MDRF 被应用于主干网络,用以弥补传统卷积在处理空间形变与多尺度特征方面的不足。该模

块结合了权重共享的多尺度可变形卷积与轻量级通道注意力机制,能够实现空间语义的动态增强,从而在不增加计算复杂度的前提下,有效提升模型对姿态变化大或局部遮挡的水下目标的表征能力。Lite SPD-DS 通过空间到通道的重排列操作压缩空间维度,并采用深度可分离卷积实现轻量化计算。同时引入通道注意力机制,对显著区域进行动态增强,从而实现紧凑而有效的特征融合,有利于对小目标和遮挡目标的检测。AIVF Loss 嵌入检测头中,以应对水下场景中类别不均衡和置信度不佳的问题,使模型能够为定位精度高的候选框分配更强的学习信号,从而实现分类置信度的一致性提升,并增强模型的空间感知能力。

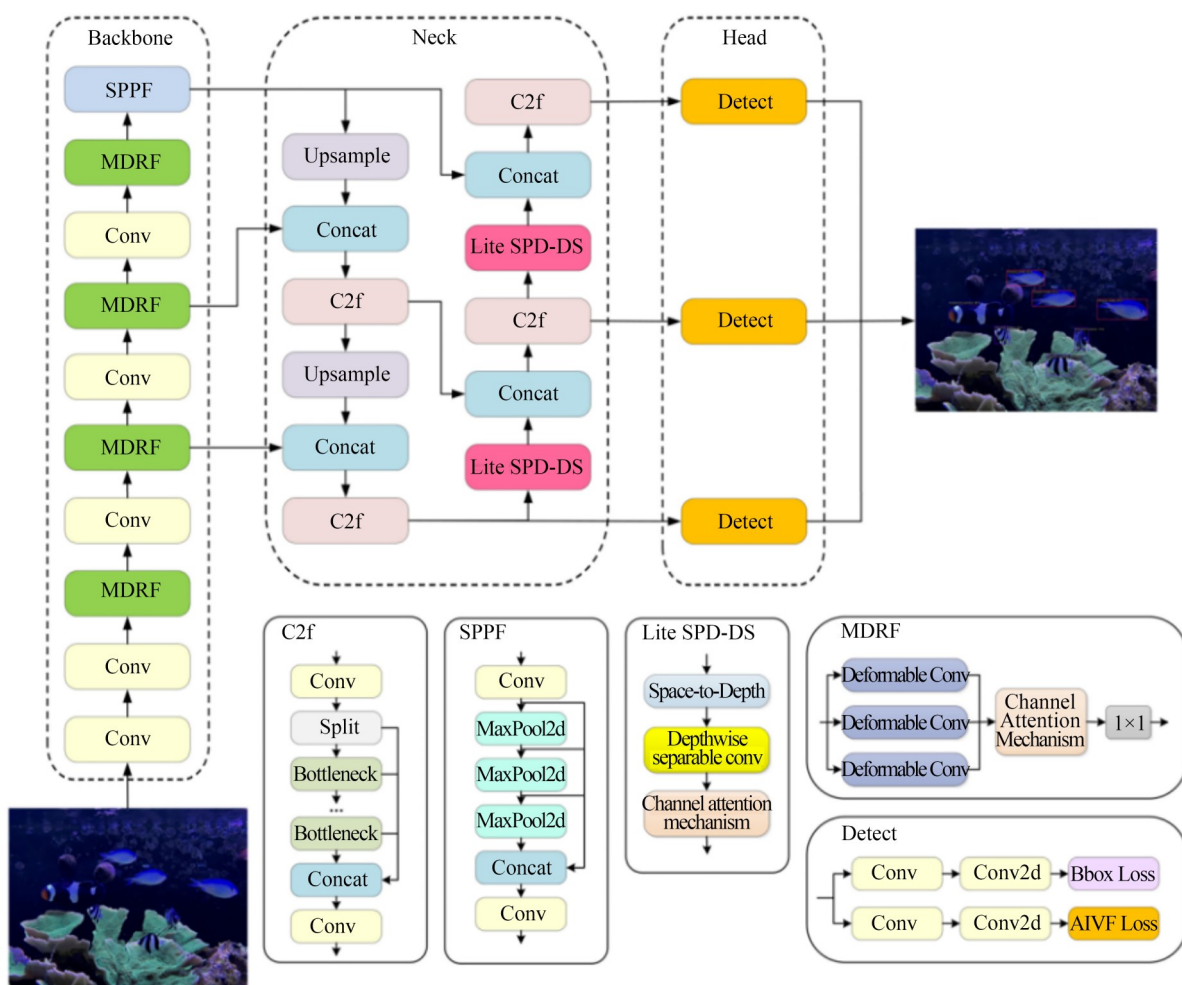


图 3 FDN-YOLO 的整体框架

Fig. 3 Overall framework of FDN-YOLO

4 实验与分析

4.1 数据集

为验证模型性能,采用 TF-DET^[25]数据集进行实验评估。该数据集涵盖了南海三亚珊瑚礁区域常见的10种代表性热带鱼类。TF-DET数据集图像样本如图4所示。

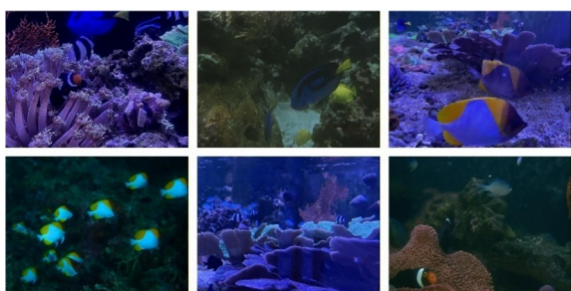


图4 TF-DET数据集的图像样本

Fig. 4 Image samples in TF-DET dataset

不同于其他在简单、单一环境中采集的鱼类数据集,TF-DET中的所有图像均来源于真实的水下生境,真实反映了复杂的场景因素,如光照不均、目标遮挡以及背景杂乱等。此外,该数据集中的大多图像中均包含多类别和多尺寸的目标个体,为复杂环境下目标检测性能评估提供了更具挑战性的要求。为提升小目标的数量,采用裁剪、缩放等增强策略,确保对小于 32×32 像素

的小目标的充分覆盖,使其符合COCO数据集对小目标尺寸划分标准。最终数据集包含4 000张图像,按8:1:1比例划分为训练集、验证集与测试集。所有增强策略仅应用于训练集,以避免潜在的数据泄露问题。

图5展示了数据集中目标在空间位置与比例上的分布特征。经过归一化处理后可知,大多数目标中心点集中分布于图像中部,并且呈现中心聚集趋势。目标高度与宽度之间具有明显的正相关性,并且高度和宽度主要分布在0.1至0.2之间,反映出小尺寸鱼类目标在数据集中占据主导地位。

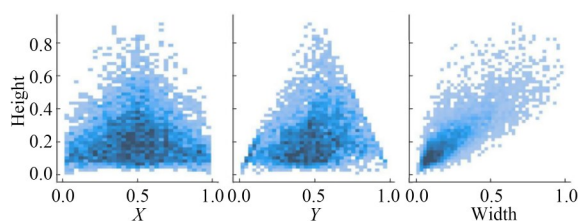


图5 数据集中目标的空间和比例分布

Fig. 5 Spatial and scale distributions of objects in dataset

图6展示了数据集中鱼类尺寸的统计可视化结果,采用小提琴图形式分别表示10种鱼类的边界框宽度(蓝色)和高度(橙色)分布(彩图见期刊电子版)。每对小提琴图对应一种鱼类,小提琴内的中心点表示该分布的中位数值。由图可知,

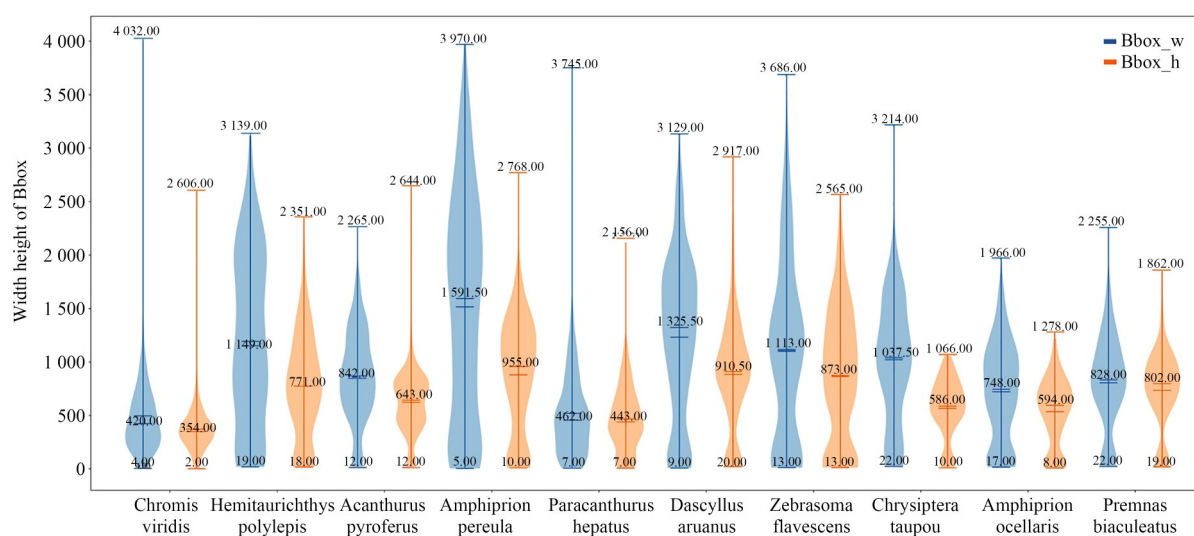


图6 数据集中水下鱼类的小提琴图分布

Fig. 6 Violin plot distributions of underwater fish in dataset

一些鱼类的目标边界框尺寸分布范围较广,表明它们尺寸存在较大变化性;而另一些的分布更为集中且更窄,表明这些鱼类目标的变化范围较小且聚焦于小尺寸。

4.2 实验配置与评估指标

本实验软件配置包括:Linux 5.4.0 操作系统, Python 3.7 编程语言, PyTorch v1.10 深度学习框架, PyCharm 开发环境。硬件配置包括:12 GB NVIDIA GeForce RTX3060 GPU 和 27 GB Intel Xeon E5-2680 v4 CPU。所有模型的训练参数设置:批量大小为 16, 训练轮数为 300 Epoch, 权重衰减系数为 0.000 5, 初始学习率设为 0.01, 动量参数设为 0.937。

在评估指标方面,目标检测算法的性能主要从检测精度与模型复杂度两个维度进行衡量。检测精度采用准确率(Precision, P)、召回率(Recall, R)以及平均精度均值(mean Average Precision, mAP)进行量化。模型的效率与可扩展性则通过参数量(Params)与浮点运算量(FLOPs)进行评估。

4.3 MDRF 对目标检测的影响

为更直观体现 MDRF 在多尺度目标检测中的贡献,表 1 对比了引入 MDRF 前后模型在不同尺寸目标的平均精度。尺寸划分标准依据 COCO 数据集分为小目标、中目标和大目标三类,其对应的平均精度分别由 AP-s、AP-m 和 AP-l 表示。

表 1 引入 MDRF 前后不同尺寸目标的平均精度
Tab.1 AP of targets at different scales before and after introducing MDRF (%)

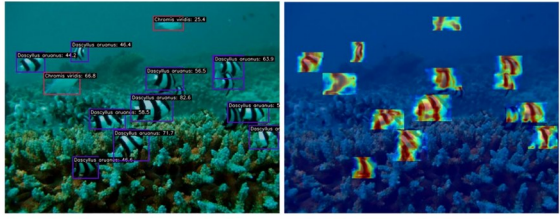
Model	AP-s	AP-m	AP-l
YOLOv8n	48.2	75.1	83.7
YOLOv8n+MDRF	50.5	76.6	83.8

由表 1 可知,引入 MDRF 后,模型的检测平均精度在小目标和中目标上分别提升了 2.3% 和 1.5%,而在大目标上几乎保持一致。图 7(a)与图 7(b)进一步给出了引入 MDRF 前后,模型对水下目标的检测结果与热力图。由图可知,在背景杂乱、目标纹理模糊的水下退化场景中,未引入 MDRF 的模型存在漏检与误检现象,而引入 MDRF 后的模型能够增强对小目标的感知能力,

实现更稳定的目标判别。



(a) 引入 MDRF 前
(a) Before introducing MDRF



(b) 引入 MDRF 后
(b) After introducing MDRF

图 7 引入 MDRF 前后的检测结果与热力图
Fig. 7 Detection results and heatmaps before and after introducing MDRF

该变化的原因在于 MDRF 通过可变形多尺度感受野融合结构,使网络能够自适应调整特征采样位置并扩大局部上下文建模范围,从而缓解下采样过程中对小目标特征的语义压缩与信息丢失问题。

4.4 下采样因子对比实验

为评估 Lite SPD-DS 中下采样因子 r 对检测性能的影响,在保持网络其余结构不变的条件下,分别设置 r 为 2, 3 和 4 进行对比实验,结果如图 8 所示。实验表明,当 r 取值为 2 时,模型在 mAP50:95 和 mAP50 上均获得最优,相较基准模

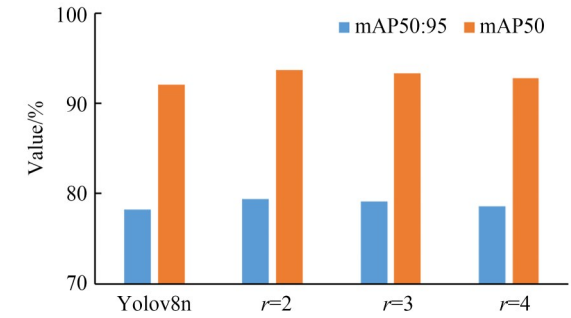


图 8 不同下采样因子的 mAP 数值对比
Fig. 8 Comparison of mAP values under different down-sampling factors

型分别提升了 1.1% 和 1.6%。当 r 进一步增大时,检测性能产生一定程度的减弱,但整体仍优于基准模型。

这表明过大的重排比例会削弱目标特征表达能力。采用较小的 r 值能够在空间分辨率保留与通道扩展能力之间取得更优平衡,而偏大的 r 值会导致特征信息在通道维度被过度分散,使模型对小目标与局部纹理的感知能力下降。因此,将 $r=2$ 设定为模型的默认配置,并在后续实验中保持一致。

4.5 损失函数对比实验

为了分析提出的 AIVF Loss 在水下检测的效果,表 2 展示了响应权重 δ 在不同取值下的对比结果。其中,调节因子 β 取值为 2。这是由于当 β 取值为 1 时,调制因子 $\alpha(\text{IoU})$ 仅对 IoU 进行线性加权,无法实现对正样本的非线性强化;而当 β 取值过大时, $\alpha(\text{IoU})$ 的曲率过陡,易导致梯度震荡,从而影响检测性能。

由表 2 可知,当 δ 取值为 0.2 时,等分较低的正样本被过度抑制,导致召回不足。当 δ 取值为 0.8 时, $\alpha(\text{IoU})$ 的动态范围收窄,样本权重趋同,降低了对高质量预测框的区分度,导致检测性能下降。当 δ 取值为 1 时,该函数退化为 Varifocal Loss,调节因子 β 缺失致使性能下滑。综合来看,当 δ 取值为 0.4 时,检测性能达到最优。

表 2 损失函数参数对检测效果的影响

Tab. 2 Influence of loss function parameters on detection performance

δ	β	$P/\%$	$R/\%$	mAP50:95/%
0.2	2	85.0	78.1	79.2
0.4	2	85.1	80.8	79.5
0.6	2	84.7	78.6	79.3
0.8	2	84.3	79.4	79.0
1	2	84.6	77.7	78.8

4.6 消融实验

为了验证所提出改进模块在水下鱼类检测任务中的有效性,表 3 展示了各组消融实验的性能指标。相较基础模型 YOLOv8n,引入 MDRF 模块可使模型的准确率提升至 85.3%,召回率提升至 81.1%。将下采样模块替换为 Lite SPD-DS 后,在保持精度略有提升的同时,模型复杂度下降至 2.4 M 和 6.6 GFLOPs。采用 AIVF Loss 可以在不增加计算开销的同时提升检测精度。当同时引入 MDRF 与 Lite SPD-DS 时,模型的计算资源为 2.6 M 和 6.8 GFLOPs,并且仍然取得较高的准确率与召回率。最终提出的 FDN-YOLO 在精度与效率方面均优于基础模型,验证了 3 种改进模块在提升水下目标检测方面的有效性。

表 3 消融实验的性能指标

Tab. 3 Performance metrics for ablation experiments

MDRF	Lite SPD-DS	AIVF Loss	$P/\%$	$R/\%$	Params/M	FLOPS/G
			84.0	79.5	3.0	8.1
✓			85.3	81.1	3.0	8.2
	✓		84.8	80.5	2.4	6.6
		✓	85.1	80.8	3.0	8.1
✓	✓		86.3	82.2	2.6	6.8
✓	✓	✓	86.7	82.8	2.6	6.8

图 9 对比了消融实验各模型的 mAP 曲线。由图可知,基础模型 YOLOv8n 的 mAP50 和 mAP50:95 分别为 92.1% 和 78.2%。引入 MDRF、Lite SPD-DS 和 AIVF Loss 之后,模型均可在不同程度上提升检测能力。最终提出的

FDN-YOLO 的提升幅度最大,该模型的 mAP50 和 mAP50:95 与基础模型相比,分别提高了 2.8% 和 2.1%。综上所述,各改进模块在准确率与资源消耗之间实现了良好平衡,显著提升了整体检测性能,使 FDN-YOLO 在各项指标中表现最优。

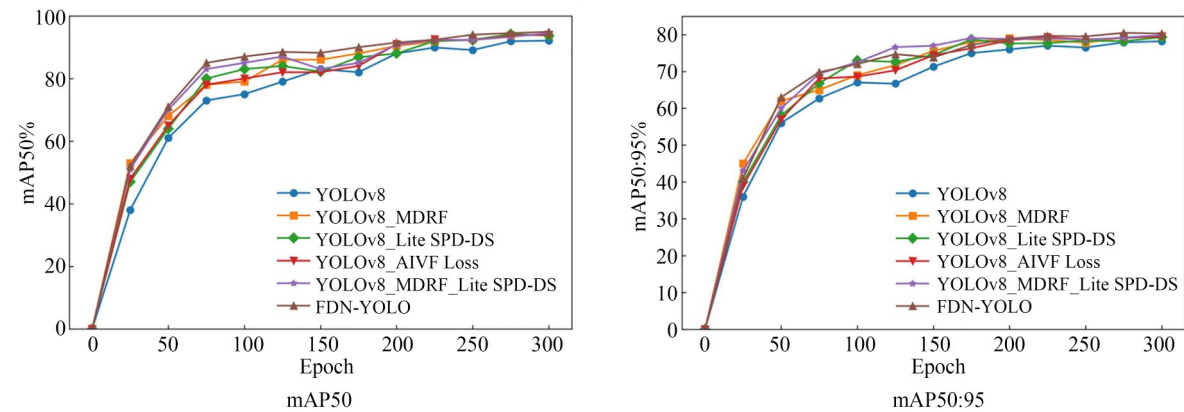


图9 消融实验的mAP曲线
Fig. 9 mAP curves of ablation experiments

4.7 不同模型对比实验

为了进一步验证所提出改进方法在水下鱼类检测任务中的优越性能,FDN-YOLO与其他主流检测模型进行了对比实验。所有对比模型

均基于TF-DET数据集,并且采用与FDN-YOLO相同的训练参数重新训练,以确保实验结果的公平性。检测性能的对比结果如表4所示。

表4 不同模型的检测性能
Tab. 4 Detection performance among different models

Model	mAP50/%	mAP50:95/%	P/%	R/%
YOLOv3-tiny	88.8	72.4	83.3	78.2
YOLOv5n	89.6	74.5	81.6	76.2
YOLOv7-tiny ^[26]	90.4	76.2	82.9	77.1
YOLOv8n ^[27]	92.1	78.2	84.0	79.5
YOLOv10n ^[28]	91.3	77.5	83.7	80.4
YOLOv11n	92.7	78.8	85.2	81.3
YOLOv12n ^[29]	92.3	78.6	84.5	80.6
FDN-YOLO	94.9	80.3	86.7	82.8

由表4可知,提出的FDN-YOLO在水下检测方面表现出较好的性能,体现了其在不同IoU阈值下对目标的精准定位与分类能力。YOLOv8n, YOLOv11n和YOLOv12n同样表现出良好性能,其检测精度略低于FDN-YOLO。YOLOv7-tiny和YOLOv5n的检测数值则相对较低。相比之下,YOLOv3-tiny表现较差,显示其在复杂水下环境中存在明显的泛化能力不足。综上所述,各模型性能的差异验证了FDN-YOLO结

构改进方案的有效性。

图10对比了不同模型的mAP、参数量与浮点运算量之间的三维热力图。在图10(a)中,FDN-YOLO的mAP50最高,参数量和浮点运算量仅为2.6 M和6.8 G。YOLOv11n与YOLOv12n展现出较强的综合性能,在保持结构紧凑的同时实现了较高精度。YOLOv3-tiny虽然具有最大的参数规模和计算量,但其mAP50最低。在图10(b)中,FDN-YOLO也展现出最优的

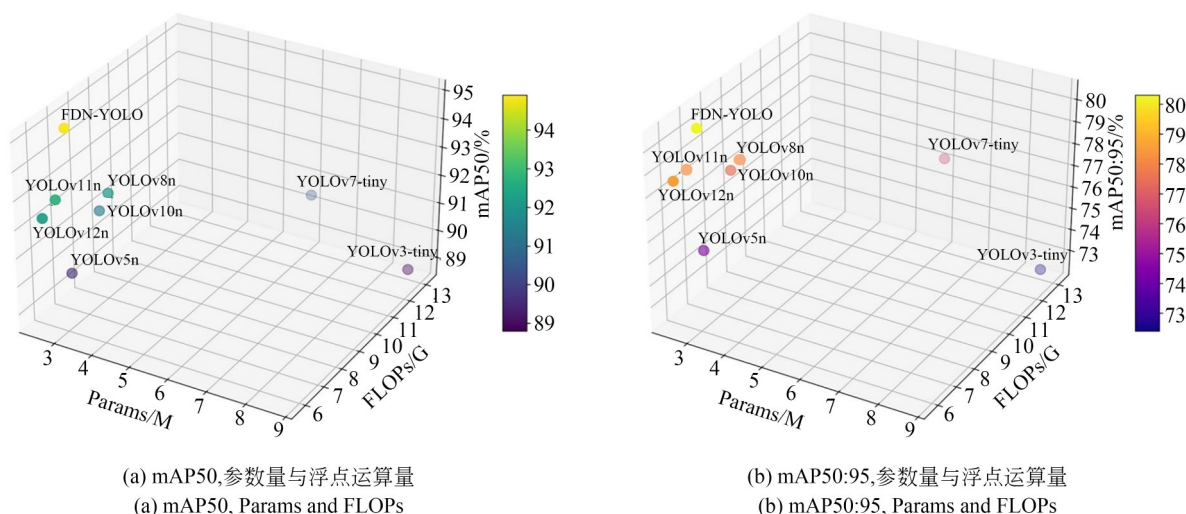


图10 不同模型的3D热图

Fig. 10 3D heatmaps of different models

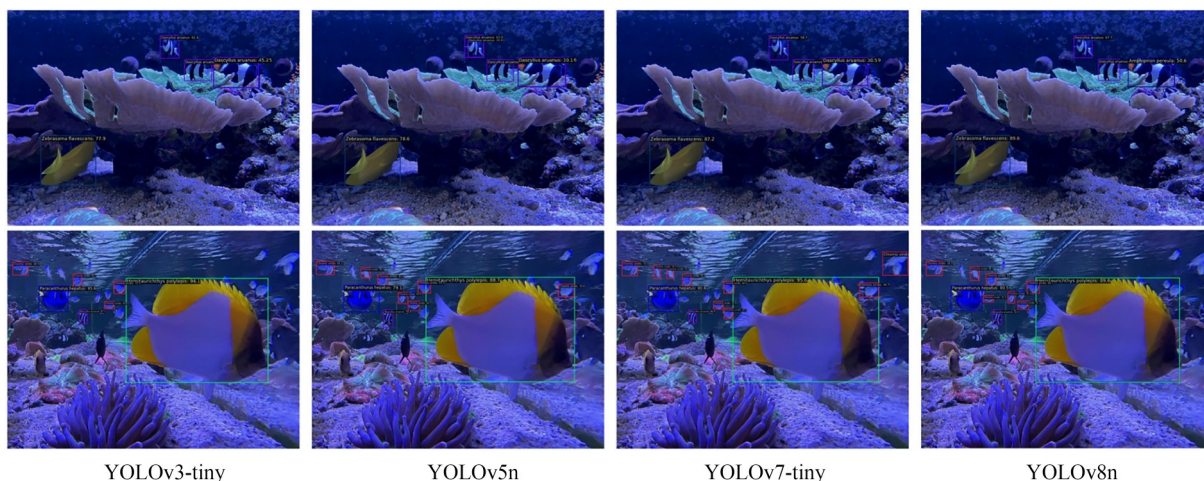
效果,体现了其在水下检测中的稳健性。

FDN-YOLO 在参数量和运算量方面的降低,主要得益于 Lite SPD-DS 的结构优化。该模块利用空间至通道重排与深度可分离卷积实现空间压缩和语义重组,大幅降低卷积核数量。此外,MDRF 具备共享偏移预测与通道压缩机制,能够在提升特征多尺度感知的同时控制参数增长。这使得模型在精准检测与结构紧凑之间实现了良好平衡。

4.8 检测结果与对比分析

图 11 展示了不同模型对水下鱼类的检测结果。对比可知, YOLOv3-tiny、YOLOv5n 和 YOLOv7-tiny 在检测处于背景中具有相似纹理特征

的目标时存在明显的误检现象。YOLOv5n 与 YOLOv10n 在此过程中产生了冗余的检测框。同时, YOLOv3-tiny、YOLOv5n 和 YOLOv8n 等在处理背景模糊或特征不明显的目标时,无法判别出目标导致检测效果不佳。相比之下, FDN-YOLO 在复杂水下环境中展现出更高的识别准确性。该模型能够以更高的置信度区分鱼类,得益于 MDRF 与 Lite SPD-DS 对细粒度、多样化特征的提取能力,同时有效减少了特征信息的损失。此外, FDN-YOLO 生成的检测框更精准贴近鱼体轮廓,得益于 AIVF Loss 有效调节目标与非目标区域的权重分布,从而提升了目标定位的精度。



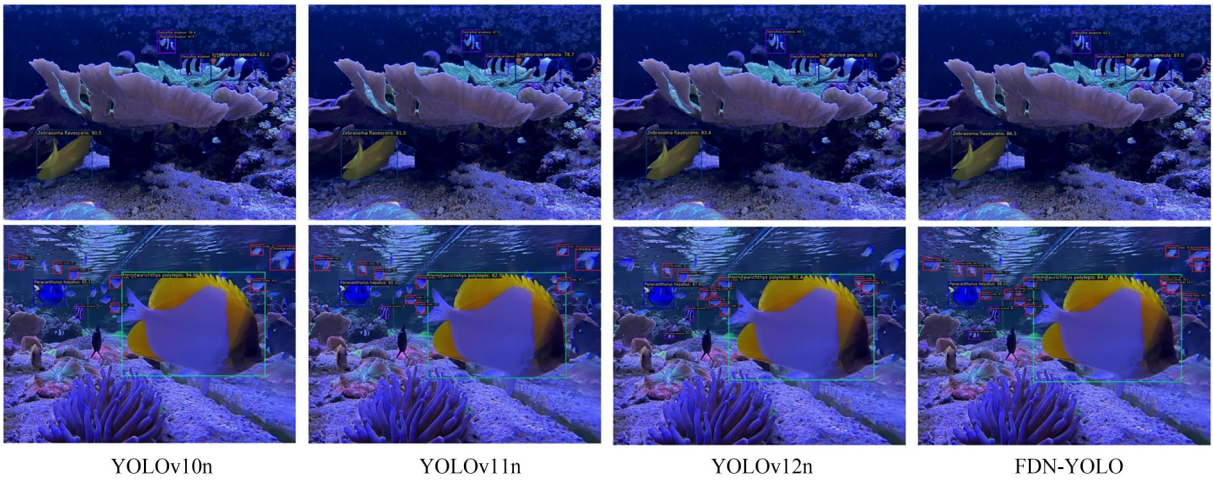


图 11 水下复杂环境的检测结果

Fig. 11 Detection results under complex underwater scenarios

4.9 多数据集对比泛化实验

为验证所提出模型在不同水下场景中的适应性与稳定性,选取 URPC^[30]、RUOD^[31]和 DUO^[32] 3个典型水下目标检测数据集进行实验。URPC 数据集由 7 600 张真实海域图像组成,包含 4 类水下物种,拍摄场景位于大连獐子岛近海,具有光照衰减、浑浊水体等多种水下成像条件。RUOD 数据集包含 14 000 张图像,覆盖 10 类水下目标,具有显著的尺度变化、光照差异与背景复杂度,可充分反映水下成像条件的多样性。DUO 数据集由 7 782 张图像构成,涵盖 4 类典型海洋生物,其小目标比例高且多数图像包含多个实例,为密集分布目标检测带来更大挑战。为保

证比较的公平性,所有模型在各数据集上均采用与主实验一致的参数设置、训练策略和评估指标。

表 5 展示了模型在 URPC 数据集上的实验结果。尽管 YOLOv3-tiny 和 YOLOv7-tiny 具有更高的参数量与浮点运算量,其检测精度仍明显落后于其他轻量化模型。YOLOv8n 与 YOLOv11n 在精度表现上有所提升,但整体增幅有限。YOLOv12n 虽进一步压缩了模型规模,却带来一定程度的精度下降。相比之下,FDN-YOLO 在保持相同量级模型复杂度的前提下,在 mAP50 和 mAP50:95 等检测指标上均取得最优表现。

表 5 在 URPC 数据集的实验结果
Tab. 5 Experimental results on URPC dataset

Model	mAP50/%	mAP50:95/%	P/%	R/%	Params/M	FLOP/G
YOLOv3-tiny	80.7	45.1	79.2	73.3	8.7	12.9
YOLOv5n	82.5	47.2	80.4	75.4	2.5	7.1
YOLOv7-tiny	81.4	45.9	80.6	72.8	6.0	13.1
YOLOv8n	82.8	47.6	80.8	75.1	3.0	8.1
YOLOv10n	81.6	46.4	79.7	75.5	2.7	8.2
YOLOv11n	83.3	47.9	80.9	76.1	2.6	6.3
YOLOv12n	81.0	45.8	79.6	73.7	2.5	6.0
FDN-YOLO	84.3	49.1	81.7	75.5	2.6	6.8

由于同一模型在不同数据集上的参数量与浮点运算量保持一致,表 6 和表 7 主要对 RUOD

和 DUO 数据集的 mAP50,mAP50:95、准确率和召回率进行比较。

表 6 在 RUOD 数据集的实验结果

Tab. 6 Experimental results on RUOD dataset (%)

Model	mAP50	mAP 50:95	P	R
YOLOv3-tiny	77.7	43.8	80.8	71.2
YOLOv5n	81.2	56.0	82.6	73.2
YOLOv7-tiny	80.3	55.4	81.2	72.3
YOLOv8n	81.5	56.7	82.6	73.8
YOLOv10n	80.7	56.2	81.1	73.7
YOLOv11n	81.6	57.2	81.5	74.4
YOLOv12n	81.3	57.3	82.4	74.1
FDN-YOLO	82.6	59.0	82.9	76.0

表 7 在 DUO 数据集的实验结果

Tab. 7 Experimental results on DUO dataset (%)

Model	mAP50	mAP 50:95	P	R
YOLOv3-tiny	80.0	57.5	82.9	70.8
YOLOv5n	82.7	62.6	84.7	73.8
YOLOv7-tiny	82.4	61.3	84.2	73.4
YOLOv8n	83.6	63.7	82.1	76.2
YOLOv10n	82.8	63.0	84.4	74.6
YOLOv11n	84.0	64.2	84.6	75.6
YOLOv12n	83.7	64.1	83.7	76.8
FDN-YOLO	84.8	65.7	84.1	78.4

由表 6 可知, YOLOv3-tiny 的整体检测性能仍处于较低水平, 与其他模型存在明显差距。YOLOv11n 与 YOLOv12n 的 mAP50:95 分别提升至 57.2% 与 57.3%, 但优势仍不突出。FDN-YOLO 的 mAP50 和 mAP50:95 分别达到 82.6% 和 59.0%, 均为各模型最高, 表明其在复杂背景下的特征建模能力更为稳定。

类似的趋势同样体现在表 7 中。YOLOv3-tiny 在各项检测指标上依然处于明显劣势, 而 YOLOv5n 与 YOLOv8n 的综合检测性能居于中等水平, 表明它们在精细特征刻画方面仍显不足。YOLOv11n 与 YOLOv12n 的精度进一步提升, 但整体表现仍逊于 FDN-YOLO。上述实验结果表明, 所提出的改进策略能够在不同数据分布与水下环境中保持稳定的特征表达能力, 并具有良好的跨数据集检测稳定性。

为了更直观的展示改进模型的优越性, 图 12 对比展示了基础模型 YOLOv8n 和 FDN-YOLO 在不同水下数据集的检测效果与热力图。可以看出, YOLOv8n 在低可见度、强光照衰减或背景纹理复杂的场景中, 其注意力响应呈现明显的离散化特征, 热力图激活区域分布零散, 且部分目标仅获得弱激活。这导致小目标遗漏、边缘模糊目标定位不准, 以及在背景区域出现误检或冗余检测框。相比之下, FDN-YOLO 在 3 个数据集

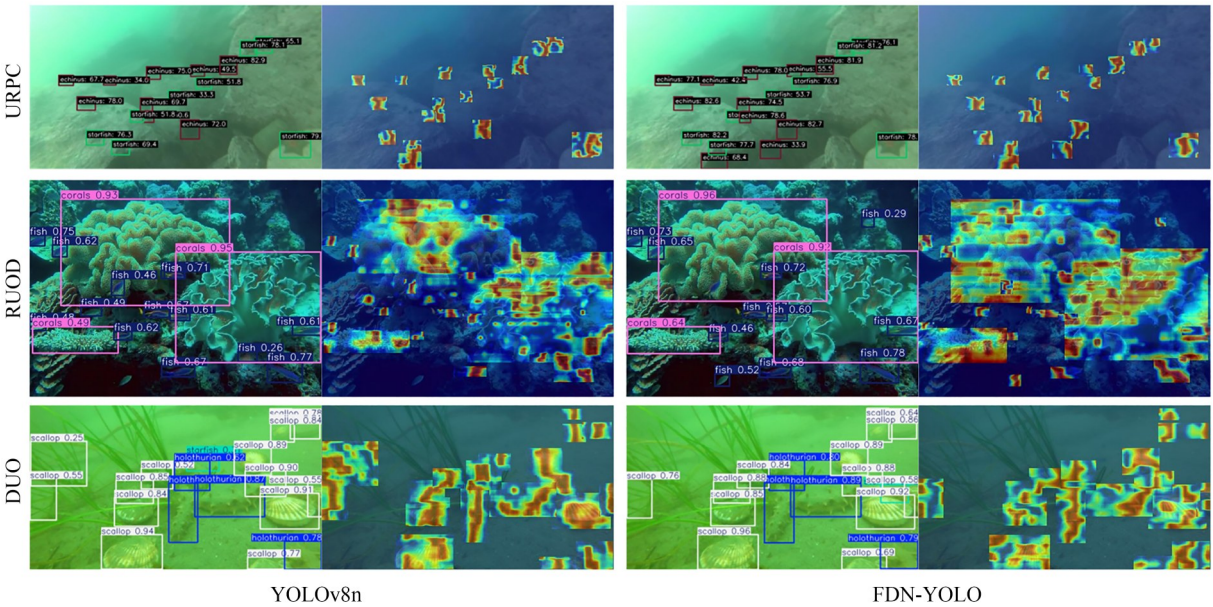


图 12 YOLOv8n 和 FDN-YOLO 在多数数据集的检测对比

Fig. 12 Detection comparison of YOLOv8n and FDN-YOLO across multiple datasets

均表现出更集中的特征关注与更稳定的空间响应,其检测框数量更合理、形状更贴合目标轮廓,并能在低对比度区域保持较高置信度。

综合分析可知,FDN-YOLO性能提升源于模型结构的多重改进。其中,MDRF模块通过可变形多尺度感受野增强,使网络能够在模糊、遮挡或姿态变化显著的区域自适应调整特征采样位置,因此在RUOD和DUO等复杂背景中仍能保持稳定检测,热力图呈现更清晰的目标轮廓与跨尺度语义聚焦。Lite SPD-DS模块有效减轻了传统下采样导致的小目标纹理丢失问题,使得FDN-YOLO对远距离小目标仍保持高强度响应。同时,AIVF Loss通过引入目标质量感知的非线性梯度调节,使模型在学习过程中更加关注定位精度高的高质量样本,从而减少了低置信度框与冗余检测,使FDN-YOLO在三类数据集上均具备更好的边界框一致性与类别可靠性。综合来看,FDN-YOLO在不同水下成像条件下均表现出更稳健的特征聚合能力,体现了所提出模块在跨数据集场景中的有效性与泛化优势。

5 结 论

本研究提出了一种动态特征增强的水下检测模型FDN-YOLO。该模型在YOLOv8n的基础上提出3种新模块:MDRF,Lite SPD-DS和AIVF Loss,旨在改善因背景复杂、姿态多变和目

标较小而导致的水下鱼类检测效果不佳的现象。首先,MDRF采用多尺度特征提取与可变形卷积运算动态扩展感受野,增强对多尺寸和多姿态鱼类目标的特征学习能力。其次,Lite SPD-DS将轻量级空间压缩结构和深度可分离卷积运算结合作为下采样策略,有效保留模糊小目标的关键细节特征。最后,AIVF Loss通过增加响应权重和调节因子自适应的增强对水下目标的定位和判别能力。结果表明,FDN-YOLO在TF-DET数据集的mAP50和mAP50:95分别提升了2.8%和2.1%,参数量和计算量分别减少了13.3%和16.0%。多数据集对比泛化实验进一步验证了该模型在检测精度与计算效率之间的良好平衡。

尽管FDN-YOLO在水下检测任务中展现出良好的泛化能力,但仍难以完全应对海洋环境中鱼类种类多样性及水下环境复杂多变所带来的挑战。因此,未来的研究工作将致力于模型的轻量化与检测算法优化,使其适合于部署在水下观测平台和智能化水下生态监测设备当中。

作者贡献声明:

朱晓龙:测量方法提出,论文构思,测量实验设计,可视化呈现及初稿写作;

陈郁隈:论文修改,审核与编辑写作;

王嘉宇:测量实验数据分析;

郭海涛:测量实验数据管理;

陈祥子:测量实验数据验证。

参考文献:

- [1] 姚婷婷,李宁,张煜. 感知增强混合网络的水下目标检测[J]. 光学精密工程, 2025, 33(8): 1303-1312.
YAO T T, LI N, ZHANG Y. Perception enhanced hybrid network for underwater object detection[J]. *Optics and Precision Engineering*, 2025, 33(8): 1303-1312. (in Chinese)
- [2] LI J, CONVERTINO M. Temperature increase drives critical slowing down of fish ecosystems[J]. *PLoS One*, 2021, 16(10): e0246222.
- [3] POUR F B, PARRA L, LLORET J, et al. Measuring and evaluating the speed and the physical characteristics of fishes based on video processing[J]. *Water*, 2023, 15(11): 2138.
- [4] HU J, ZHAO D D, ZHANG Y F, et al. Real-time nondestructive fish behavior detecting in mixed polyculture system using deep-learning and low-cost devices[J]. *Expert Systems with Applications*, 2021, 178: 115051.
- [5] 彭晏飞,张添洪,安彤. 基于像素级水平的通道自适应水下图像增强算法[J]. 液晶与显示, 2024, 39(8): 1037-1045.
PENG Y F, ZHANG T Q, AN T. Channel adaptive underwater image enhancement algorithm based on pixel level[J]. *Chinese Journal of Liquid Crystals and Displays*, 2024, 39(8): 1037-1045. (in Chinese)

- [6] LIU H Y, LIU T, GU Y Z, *et al.* A high-density fish school segmentation framework for biomass statistics in a deep-sea cage[J]. *Ecological Informatics*, 2021, 64: 101367.
- [7] DONG C, WANG Z S, XING C D, *et al.* Occlusion detection of dense correspondence fields using motion statistics and triangulation[J]. *IEEE Access*, 2019, 7: 83569-83580.
- [8] CARRE M, MOTTA P J. Tooth length and occlusion in four species of piscivorous fishes: getting a grip on prey[J]. *Environmental Biology of Fishes*, 2020, 103(8): 903-912.
- [9] 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.
- [10] 苗力恒, 田莹. 改进 YOLOv8n 的多尺度轻量化水下目标检测[J]. 电子测量与仪器学报, 2025, 39(4): 141-151.
MIAO L H, TIAN Y. Improved YOLOv8n multi-scale and lightweight underwater target detection[J]. *Journal of Electronic Measurement and Instrumentation*, 2025, 39(4): 141-151. (in Chinese)
- [11] LI J Y, LIU C N, LU X C, *et al.* CME-YOLOv5: an efficient object detection network for densely spaced fish and small targets[J]. *Water*, 2022, 14(15): 2412.
- [12] ZHU J D, HE W W, WENG W D, *et al.* An embedding skeleton for fish detection and marine organisms recognition[J]. *Symmetry*, 2022, 14(6): 1082.
- [13] LU Y C, TUNG C, KUO Y F. Identifying the species of harvested tuna and billfish using deep convolutional neural networks[J]. *ICES Journal of Marine Science*, 2020, 77(4): 1318-1329.
- [14] XU X L, LI W S, DUAN Q L. Transfer learning and SE-ResNet152 networks-based for small-scale unbalanced fish species identification[J]. *Computers and Electronics in Agriculture*, 2021, 180: 105878.
- [15] IQBAL M A, WANG Z J, ALI Z A, *et al.* Automatic fish species classification using deep convolutional neural networks[J]. *Wireless Personal Communications*, 2021, 116(2): 1043-1053.
- [16] SMADI AAL, MEHMOOD A, ABUGABAH A, *et al.* Deep convolutional neural network-based system for fish classification[J]. *International Journal of Electrical and Computer Engineering (IJECE)*, 2022, 12(2): 2026.
- [17] SAYDIRASULOVICH S N, MUKHIDDINOV M, DJURAEV O, *et al.* An improved wildfire smoke detection based on YOLOv8 and UAV images[J]. *Sensors*, 2023, 23(20): 8374.
- [18] 张彪, 朱振洋, 许家忠. 面向自主水下机器人的深海污染物检测[J]. 光学精密工程, 2025, 33(11): 1803-1817.
ZHANG B, ZHU Z Y, XU J Z. Deep-sea pollutant detection for autonomous underwater robots[J]. *Optics and Precision Engineering*, 2025, 33(11): 1803-1817. (in Chinese)
- [19] SONG D K, HUO H, SONG D K, *et al.* Lightweight underwater target detection algorithm based on YOLOv8n[J]. *Electronics*, 2025, 14(9): 1810.
- [20] DAI J F, QI H Z, XIONG Y W, *et al.* Deformable convolutional networks[C]. 2017 *IEEE International Conference on Computer Vision (ICCV)*. October 22-29, 2017, Venice, Italy. IEEE, 2017: 764-773.
- [21] XIONG Y W, LI Z Q, CHEN Y T, *et al.* Efficient deformable ConvNets: rethinking dynamic and sparse operator for vision applications[C]. 2024 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. June 16-22, 2024, Seattle, WA, USA. IEEE, 2024: 5652-5661.
- [22] SUNKARA R, LUO T. No more strided convolutions or Pooling: a new CNN building block for Low-Resolution images and Small objects[C]. *Machine Learning and Knowledge Discovery in Databases*. Cham: Springer, 2023: 443-459.
- [23] HOWARD A G, ZHU M L, CHEN B, *et al.* MobileNets: efficient convolutional neural networks for mobile vision applications[EB/OL]. 2017: *arXiv*: 1704.04861. <https://arxiv.org/abs/1704.04861>.
- [24] ZHANG H Y, WANG Y, DAYOUB F, *et al.* VarifocalNet: an IoU-aware dense object detector[C]. 2021 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. June 20-25, 2021, Nashville, TN, USA. IEEE, 2021: 8510-8519.
- [25] LU Z 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.
- [26] WANG C Y, BOCHKOVSKIY A, LIAO H M. YOLOv7: trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[C]. 2023 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. June 17-24, 2023, Vancouver, BC, Canada. IEEE, 2023: 7464-7475.
- [27] SOHAN M, SAI RAM T, RAMI REDDY C V. A review on YOLOv8 and its advancements[C]. *Data Intelligence and Cognitive Informatics*. Singapore: Springer, 2024: 529-545.
- [28] CHEN H, CHEN K, DING G G, *et al.* YOLOv10: real-time end-to-end object detection[C]. *Advances in Neural Information Processing Systems* 37. December 10-15, 2024. Vancouver, BC, Canada. *Neural Information Processing Systems Foundation, Inc. (NeurIPS)*, 2024: 107984-108011.
- [29] TIAN Y J, YE Q X, DOERMANN D. YOLOv12: attention-centric real-time object detectors [EB/OL]. 2025: *arXiv*: 2502.12524. <https://arxiv.org/abs/2502.12524>.
- [30] SHEN X, YUAN G L, WANG H B, *et al.* Un-supervised clustering optimization-based efficient attention in YOLO for underwater object detection [J]. *Artificial Intelligence Review*, 2025, 58(7): 219.
- [31] FU C P, LIU R S, FAN X, *et al.* Rethinking general underwater object detection: Datasets, challenges, and solutions[J]. *Neurocomputing*, 2023, 517(C): 243-256.
- [32] LIU C W, LI H J, WANG S C, *et al.* A dataset and benchmark of underwater object detection for robot picking[C]. 2021 *IEEE International Conference on Multimedia & Expo Workshops (ICMEW)*. July 5-9, 2021, Shenzhen, China. IEEE, 2021: 1-6.

作者简介:



朱晓龙(1990—),男,黑龙江哈尔滨人,博士,讲师,硕士生导师,2013年、2016年和2020年于东北林业大学分别获得学士、硕士和博士学位,主要从事计算机视觉、深度学习及其在水下目标检测应用方面的研究。E-mail: zhuxiaolonglong@sina.cn