doi:10.6041/j.issn.1000-1298.2025.10.009

多目标鱼类实例分割 LIS－YOLO 轻量化模型研究

徐文凯[1,2]　郎　平[1,2]　李道亮[1,2]

（1.中国农业大学国家数字渔业创新中心，北京 100083；2.中国农业大学信息与电气工程学院，北京 100083）

摘要：水下目标的准确分割为研究水生动物的行为、生物量等奠定了基础。然而，现有的水下实例分割算法面对水下特有的环境干扰（如悬浮颗粒、颜色衰减、背景噪声）时，往往缺乏足够的鲁棒性，并且提升轻量化模型在非结构化、动态水下场景中的泛化能力是一项关键挑战。为此，本文提出了一种轻量化的实例分割模型 LIS－YOLO（Lightweight instance segmentation YOLO v8），可以从不同的拍摄角度对鲟鱼、幼年鲈鱼和成年鲈鱼进行有效分割，并基于 PyQt5 构建了水下目标实时分割系统。首先设计了一个轻量化、高精度的 C2f－Faster－EMA 模块，通过更轻量化的 C2f－Faster 模块取代了原模型中复杂的 C2f 模块，并集成了高效的多尺度关注，提高了对小目标鱼类的特征提取能力。其次，在改进模型中引入 Wise－IoU，减少了低质量样本产生的有害梯度，提高了模型在复杂环境下的分割能力。最后，利用图形用户界面和 PyQt5 接口开发了水下多目标实例分割的实时分割系统，能够可视化不同种类鱼。实验结果表明，LIS－YOLO 模型的精度、平均精度均值、浮点运算量和帧率分别为 97.2%、95.9%、3.60×10^{10} 和 127 f/s。模型参数量压缩至 9.0×10^{6}，占原模型的 76.3%。本研究不仅为水下目标识别提供了一种准确、轻量级的实例分割模型，而且探讨了不同拍摄角度下鱼类分割的有效性，对提高水产养殖智能化水平具有实际应用价值。

关键词：深度学习；实例分割；多目标鱼类；YOLO v8；PyQt5；轻量化模型

中图分类号：TP391.41　　**文献标识码**：A　　**文章编号**：1000-1298(2025)10-0094-08　　**OSID**：

Lightweight LIS－YOLO Model for Multi-target Fish Instance Segmentation

XU Wenkai[1,2]　LANG Ping[1,2]　LI Daoliang[1,2]

(1. *National Innovation Center for Digital Fishery*, *China Agricultural University*, *Beijing* 100083, *China*
2. *College of Information and Electrical Engineering*, *China Agricultural University*, *Beijing* 100083, *China*)

Abstract: Accurate segmentation of underwater objects laid the foundation for studying aspects such as the behaviour and biomass of aquatic animals. However, existing underwater instance segmentation algorithms often lacked sufficient robustness when facing underwater environment-specific interferences, such as suspended particles, colour attenuation, and background noise, and enhancing the generalization ability of lightweight models in unstructured and dynamic underwater scenes remained a key challenge. To this end, a lightweight instance segmentation lightweight instance segmentation YOLO v8 (LIS－YOLO) model was proposed, which can effectively segment sturgeon, juvenile bass, and adult bass from different shooting angles, and a real-time underwater object segmentation system was developed based on PyQt5. Firstly, a lightweight and high-precision C2f－Faster－EMA module was designed, in which the original complex C2f module was replaced with a more lightweight C2f－Faster module, and an efficient multi-scale attention mechanism was integrated to improve the feature extraction capability for small-object fish. Secondly, Wise－IoU was introduced into the improved model to reduce harmful gradients caused by low-quality samples, thereby enhancing the model's segmentation capability in complex environments. Finally, a real-time multi-object instance segmentation system for underwater objects was developed by using a graphical user interface and the PyQt5 framework, enabling the visualization of different fish species. The experimental results showed that the LIS－YOLO model achieved precision, mean average precision, floating-point operations, and frame rate of 97.2%, 95.9%, 3.60×10^{10}, and 127 f/s,

收稿日期：2025－08－01　修回日期：2025－08－15

基金项目：国家自然科学基金项目(32373186)

作者简介：徐文凯(1996—)，男，博士生，主要从事机器视觉在水产养殖中的应用研究，E-mail：wenkai.xu@cau.edu.cn

通信作者：李道亮(1971—)，男，教授、博士生导师，主要从事渔业先进感知与智能信息处理研究，E-mail：dliangl@cau.edu.cn

respectively. The number of model parameters was compressed to 9.0×10^6, accounting for 76.3% of the original model. This research result not only provided an accurate and lightweight instance segmentation model for underwater object recognition but also explored the effectiveness of fish segmentation from different shooting angles, offering practical application value for improving the level of intelligent aquaculture.

Key words: deep learning; instance segmentation; multi-object fish; YOLO v8; PyQt5; lightweight model

0　引言

水生动物的生存环境不同于陆生动物，生存条件更为复杂，这在一定程度上阻碍了人工智能等现有新技术在水产养殖领域进行更深入的研究。因此，如何将现有的先进技术应用到水产养殖领域，对水生动物进行快速检测和分析，是亟待解决的问题之一。

精准检测是获取水生动物长度估计[1]、行为量化[2-3]等信息的前提。目前，对目标检测可以分为两种方法：基于深度学习的目标检测方法和传统的图像处理方法[4]。传统的图像处理方法通常是根据物体的形状和颜色从含有背景的图像中提取物体区域。在这种情况下，YANG 等[5]提出了一种结合 Canny 边缘检测器的监督多阈值分割模型来识别马铃薯的发芽情况。LIN 等[6]提出了一种基于颜色理论的概率图像分割方法，旨在将 RGB 图像划分为二值掩模，以检测球形或圆柱形水果。SUN 等[7]提出了一种自适应快速聚类图像分割算法来检测鱼群，虽然能够实现对小目标鱼体的有效分割，但是对噪声敏感（例如气泡、悬浮物等），导致分割不准确。尽管传统图像处理方法在特定场景下表现良好，但面对复杂多变的水下环境（如低光照、鱼类形态相似等挑战），易将相邻鱼体误判为同一区域导致难以分离重叠或遮挡个体。因此，传统图像处理方法的鲁棒性和可移植性存在明显局限性，而基于深度学习的目标实例分割方法通过数据驱动特征学习，显著提升了模型的鲁棒性和可移植性。

深度学习技术的迅速发展促使国内外学者一致研究精确、快速的目标检测方法，从而解决传统图像检测方法固有的局限性。基于深度学习的物体检测方法依赖于物体的纹理、颜色和形状特征进行学习。这些方法在不同的领域得到了广泛的应用，包括但不限于医学、工业、动物学和其他跨学科领域[8-13]。研究人员在高精度和轻量化中实现均衡，这也是目前研究的主要方向之一。

在检测过程中，模型经常以矩形框的形式检测目标，导致在边界框中不仅存在目标，且存在背景干扰。基于实例分割的方法可以根据目标的轮廓分割出精确的像素级掩码，有效地解决了这一问题。随着实例分割方法的不断优化，在作物分割[14]、细胞分割[15]、农场边界分割[16]等领域[17]得到了广泛的应用。与普通目标检测生成的边界框相比，基于实例分割的方法可以实现精确的目标分割，从而进一步获得目标的周长、面积等信息。然而，由于鱼类的不停歇运动和遮挡，再加上水下反射和折射的影响，给实例分割带来了一些挑战。目前，也有一些将实例分割应用于水生动物检测的研究。例如，LIU 等[18]使用 YOLO v5 作为主干结合语义来分割鱼类，准确率达到 98.5%。YU 等[19]提出了 CAM－Decoupled－SOLO 来提取鱼类轮廓特征，实现了鱼类尺寸的估算。然而，目前的研究主要集中在同一规格、同一类型的鱼分割上，还未见针对多类型鱼的轻量化实例分割模型的研究报道。

为提高模型在鱼类实例分割任务中的精度与边界质量，本文拟在 YOLO v8－Seg 框架基础上，引入 C2f－Faster 结构增强主干网络的特征提取能力，并采用 EMA 机制提高小目标鱼类的特征提取能力。虽然已有部分研究（如文献[20－22]）已将 C2f－Faster 与 EMA 应用于 YOLO v8 中，但其主要面向目标检测任务，且未涉及实例分割。本文进一步融合 Wise－IoU 损失函数，以改善传统 IoU 在分割任务中对目标边界的拟合能力，特别适用于轮廓复杂、边界模糊的鱼类外形分割。

1　材料与方法

1.1　实验材料

1.1.1　实验鱼驯化

鲈鱼和鲟鱼是本次实验中使用的两种鱼类，它们均饲养在国家数字渔业创新中心（中国农业大学）循环水系统中。使用的鲈鱼分为体长约 12 cm 的幼年鲈鱼和体长约 24 cm 的成年鲈鱼两类，分别饲养在两个圆柱形水箱中。由于体型的原因，只选择体长在 30 cm 左右的鲟鱼，单独饲养在另一个圆柱形水箱中。鲈鱼和鲟鱼分别于每天 08:00 和 18:00 投喂，投喂量约为其质量的 1%。水温和溶解氧浓度分别维持在 20℃和 8.0 mg/L。

1.1.2 图像数据采集与标注

鲈鱼和鲟鱼的原始图像由水下摄像机拍摄获取。考虑到模型对小目标的泛化能力较差,加入幼年鲈鱼的数据以提高模型对小物体的分割能力。为了获取不同拍摄角度的各种鱼类图像以丰富数据集,本文将摄像机设置为 3 个角度进行拍摄,不同拍摄角度的图像如图 1 所示(第 1 行图像为摄像机从顶部以 45°角拍摄的俯视图,第 2 行图像为相机水平拍摄的正视图,第 3 行图像为摄像机从底部以 45°角拍摄的底部视图)。为了减少水的折射和反射影响,采用水下摄像机在水下采集图像数据。最终得到 450 幅图像,按照 7∶2∶1的比例划分为训练集、验证集和测试集。使用 Labelme 软件对图像进行标注,标注后的标签文件以 Json 格式保存。

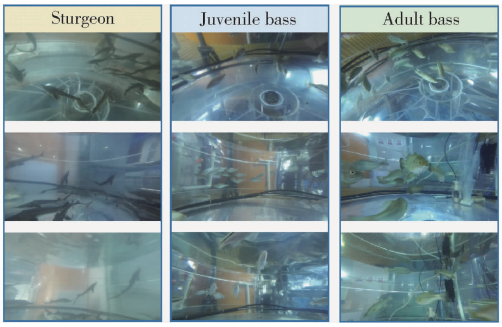


图 1 不同拍摄角度的图像

Fig. 1 Images from different shooting angles

1.2 LIS-YOLO v8 模型

YOLO v8 为目前主流的一阶段检测器,与 Mask R-CNN[23] 和 Faster R-CNN[24] 等二阶段检测器相比,一阶段检测器省略了 ROI 池化,从而在一定程度上加快了检测速度。因此,本文对原来的 YOLO v8 进行了改进。将原来的 C2f 模块替换为 C2f-Faster-EMA 模块,在模块压缩过程中引入了注意力机制。进一步将损失函数改进为 Wise-IoU,以提供更准确的目标分割评估。图 2 为 LIS-YOLO v8 模型总体框架。

1.2.1 C2f-Faster 模块

为了从高质量图像中提取特征,YOLO v8 模型没有使用 YOLO v5 模型中的 C3 模块,而是使用了 C2f 模块。C2f 模块的设计灵感来源于 ELAN 和 C3 模块的基本原理,包括由 2 个 1×1 卷积层和 1 个 3×3 卷积层,然后通过多次瓶颈运算获得更丰富的梯度信息。然而,这些操作会导致模型复杂性和参数量的增加。为此,在保持模型精度的前提下,集成 FasterNet 模块[25]重构 C2f,即 C2f-Faster 模块。这种调整旨在减少参数量和总体复杂性,从而实现模型的轻量化。

为了减少浮点运算量,将部分卷积(Partial convolution,PConv)集成到 FasterNet 模块中,以优化内存访问并减少计算冗余。PConv 应用规则卷积从部分输入通道提取空间特征,同时保证其余部分不变。

C2f-Faster 由 4 部分组成,每个部分之前都有一个合并层(2×2 卷积,步长为 2)或一个嵌入层(4×4 卷积,步长为 4),用于通道扩展和空间下采样。此外,C2f-Faster 还包括一个 PConv 层,随后是 2 个 PWConv 层,使用 ReLU 和批处理归一化作为激活函数和归一化来降低浮点运算量。图 3 为 C2f-

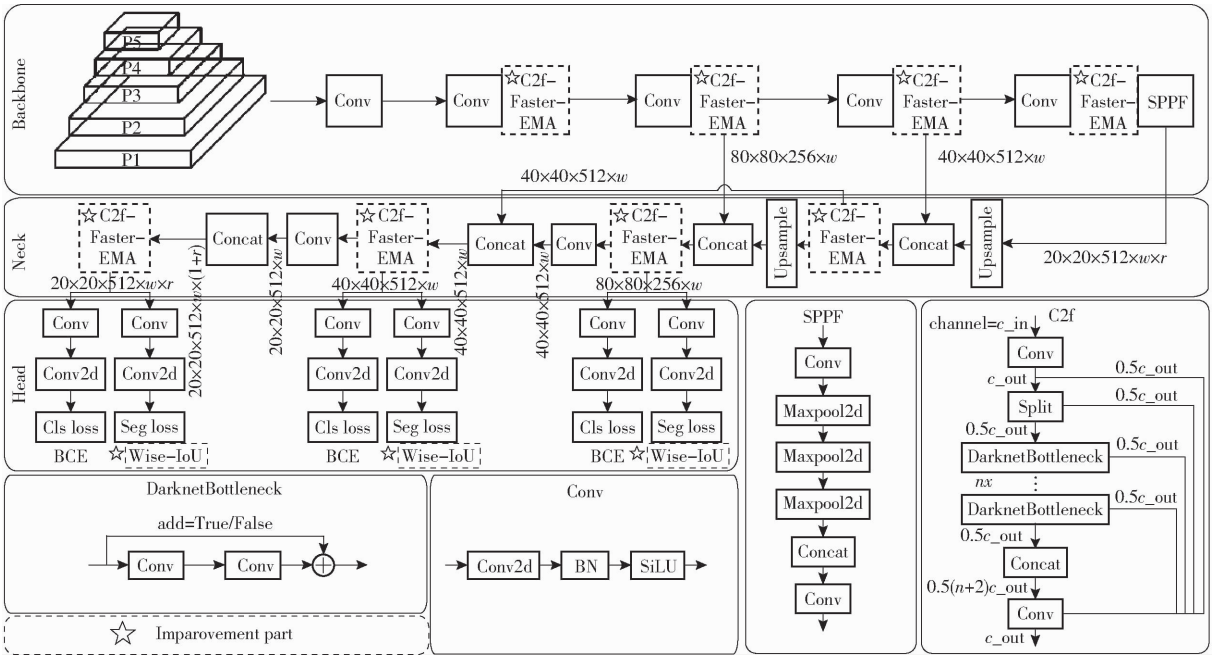


图 2 LIS-YOLO v8 模型结构图

Fig. 2 Structure diagram of LIS-YOLO v8 model

Faster 结构图。

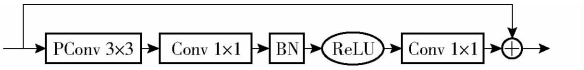


图 3　C2f－Faster 结构图

Fig. 3　Structure diagram of C2f－Faster

1.2.2　EMA 模块

由于注意力机制具有灵活的结构特点，研究人员经常将其引入到目标检测模型中。现有的注意力机制可分为 3 种类型：空间注意力机制、通道注意力机制和两者的结合，如 CBAM[26] 和 SE[27]。然而，这些模块大多基于手工设计，结构复杂、依赖大量池化操作，导致计算开销显著上升，难以兼顾分割精度与推理效率。因此，为了提升模型对鱼类复杂轮廓和局部纹理的识别能力，同时控制计算成本，本文将 EMA 模块集成至 C2f－Faster 主干结构中，用以增强特征表达能力与感受域适应性，特别适用于处理鱼体重叠、姿态变化大的实例分割场景。

为了减少浮点运算的次数，原始 YOLO v8 的 C2f 模块被 C2f－Faster－EMA 所替代。此外，在前向传播过程中加入了注意力机制，提高了模型的准确性。虽然在模型中引入注意力机制增加了计算复杂度，但改进后的轻量化模型与原来的 YOLO v8 模型相比仍然更小。因此，注意力机制被整合到模型的主干和颈部部分。

1.2.3　Wise－IoU

在实例分割任务中，IoU（Intersection over Union）损失函数是用于评估真实边界框与预测边界框之间重叠程度的常用度量。但是，IoU 损失函数在一些小物体的检测任务中，特别是在高度压缩的边界盒的情况下，并不能有效地解决这类问题。因此，TONG 等[28]提出了以 IoU 为基础的损失函数，具有动态非单调聚焦机制，称为 Wise－IoU，该损失函数结合了 IoU 和 GIoU[29]。通过引入新的权重参数，可以有效地测量目标盒的重叠程度，能够更好地适应不同大小和形状的鱼类目标。此外，Wise－IoU 采用离群度代替 IoU 来评估锚定质量，为梯度收益分配提供了一种改进的策略。

值得注意的是，本研究所构建的鱼类实例分割数据集中普遍存在鱼体边缘模糊、轮廓重叠、部分标注不确定等问题，这对掩码区域的边界回归提出了更高要求。与强调前景-背景不均衡的 Focal－IoU 等其他 IoU 不同，Wise－IoU 更适合处理此类标注质量不稳定与边界定位困难的任务。因此，本文在分割头中引入 Wise－IoU 作为边界损失项，以提升模型对鱼类复杂形态的拟合能力，改善分割掩码的边界精度。

1.3　评价指标

将精度、召回率、平均精度均值、参数量、浮点运算量和帧率作为模型的性能评价指标[30]。平均精度均值表示模型整体分割性能的评价指标，将每个类别的平均精度求和再除以类别的数量，从而得到平均精度均值。除了上述评价指标外，还使用浮点运算量和模型参数量来评价模型的复杂性。浮点运算量是模型所需的总计算量，模型参数量表示模型卷积层所学习到的权重个数。

2　结果与分析

2.1　不同模型对比

对比实验和消融实验在配备有 12 GB 内存的 NVIDIA RTX3060 显卡、英特尔酷睿 i5－12400F CPU 和 Windows 11 操作系统中进行，使用 Python 3.8 版本以及 CUDA 11.7 作为开发环境。选取了几种主流的分割模型来验证该方法的有效性。Mask R－CNN、DeepLabv3、FCN、LR－ASPP 和 YOLO v5 模型使用相同的参数设置和数据集进行训练和测试。LIS－YOLO 与这些模型的比较结果如表 1 所示。

表 1　主流模型性能对比

Tab. 1　Comparisons of mainstream models

模型	平均精度均值/%	参数量	浮点运算量	帧率/(f·s^{-1})
Mask R－CNN	92.3	4.44×10^{7}	2.01×10^{11}	10
DeepLabv3	79.6	4.20×10^{7}	3.32×10^{10}	18
FCN	80.0	3.53×10^{7}	2.84×10^{10}	23
LR－ASPP	60.3	3.20×10^{6}	3.90×10^{8}	111
YOLO v5[31]	89.0	7.40×10^{6}	2.57×10^{10}	196
LIS－YOLO	95.9	9.00×10^{6}	3.60×10^{10}	127

从表 1 可以看出，所提出方法的平均精度均值高于其他 5 个模型。虽然 LIS－YOLO 的参数量、浮点运算量和帧率没有达到最优值，但总体上达到了预期的效果。Mask R－CNN 作为二阶段检测器的代表之一，其浮点运算量最高，且帧率最低，在实际应用场景中难以完成实时分割任务。DeepLabv3 和 FCN 指标相似，分割精度不高，且帧率太低。LR－ASPP 的模型参数量和浮点运算量最低，但平均精度均值过小，无法满足精度的要求。YOLO v5 作为 YOLO 系列模型的较早版本，其帧率最高，模型参数量较少，但分割精度不理想。总体来说，LIS－YOLO 不仅保证了较高的分割精度，而且具有较高的帧率和较少的模型参数量，可以满足实时分割的精度和速度要求。

2.2　消融实验

为了进一步验证所提模型的有效性，对改进模

型进行了消融实验,实验结果如表 2 所示。在原有 YOLO v8 模型的基础上,将 C2f 模块替换为 C2f - Faster 模块,可以有效减少模型参数量,但平均精度均值降低 0.1 个百分点。因此,引入 EMA 注意力机制来弥补这一精度低的缺陷。从表 2 可以看出,引入注意力机制后,模型参数量并没有明显增加,帧率下降,但也能满足实时分割的要求。为了提高模型的识别平均精度,引入了 Wise - IoU,在不显著降低精度的情况下实现了轻量化。总体来说,改进的 LIS - YOLO 模型具有更好的整体分割性能。

表 2 消融实验结果

Tab. 2 Results of ablation experiments

模型	平均精度均值/%	浮点运算量	参数量	帧率/(f·s^{-1})
YOLO v8s	96.0	4.24×10^{10}	1.18×10^{7}	159
YOLO v8s + EMA	95.1	4.27×10^{10}	1.18×10^{7}	143
YOLO v8s + Wise - IoU	94.9	4.24×10^{10}	1.18×10^{7}	145
YOLO v8s + C2f - Faster	94.5	3.54×10^{10}	8.96×10^{6}	139
YOLO v8s + C2f - Faster + EMA	95.1	3.60×10^{10}	8.99×10^{6}	115
YOLO v8s + EMA + Wise - IoU	95.9	4.27×10^{10}	1.18×10^{7}	150
YOLOv8s + C2f - Faster + Wise - IoU	95.5	3.54×10^{10}	8.96×10^{6}	159
LIS - YOLO	95.9	3.60×10^{10}	9.00×10^{6}	127

精度曲线和损失曲线是评价模型性能的常用指标。模型的精度作为最佳分割能力的决定因素,而模型的鲁棒性可以通过损失曲线的变化间接判断。图 4 为 LIS - YOLO 添加不同改进模块的精度曲线和损失曲线。由图 4a 可以看出,当迭代次数小于 171 时,LIS - YOLO 的分割精度相比于其他模型并不是最优的;当迭代次数大于 171 时,LIS - YOLO 的分割精度开始出现最优值,虽然在随后几个迭代中仍出现低于其他模型精度的现象,但最终 LIS - YOLO 的精度回升且优于其他模型。从图 4b 可以直观看出,改进的 LIS - YOLO 与原始的 YOLO v8 模型展现出高度相似的损失收敛轨迹,表明所改进模

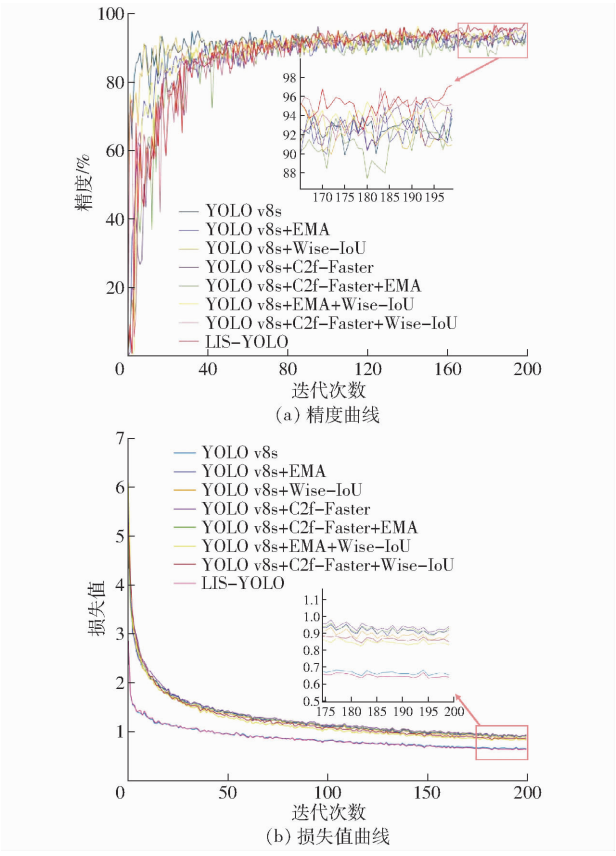


图 4 消融实验性能曲线

Fig. 4 Performance curves of ablation experiment

型保持了原始模型的优化稳定性,且在后期阶段 LIS - YOLO 的损失曲线始终低于原始模型的损失曲线。然而,其他消融实验中模型的损失并未表现出较优的收敛效果。出现以上现象可能是因为单一或两两相融的改进并不能协同优化模型的损失,而 LIS - YOLO 可以实现训练稳定性、损失函数敏感度和特征提取效率这 3 种不同维度的协同优化。

2.3 LIS - YOLO v8 模型结果及分析

深度卷积神经网络的特征提取过程往往是抽象的。因此,为了更好地理解模型对图像中关键信息的特征提取过程,图 5 给出了 LIS - YOLO 模型特征图的可视化结果。Conv、C2f - Faster、SPPF、Upsample、EMA、C2f - Faster - EMA 等不同模块的特征图以热力图的形式表示。从图中可以看出,在没有引入注意力机制的情况下,C2f - Faster - EMA 模

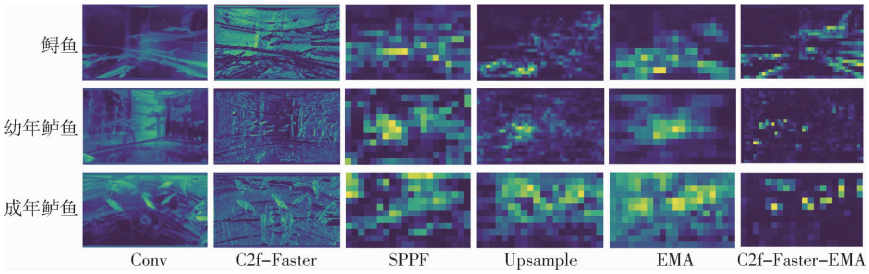


图 5 LIS - YOLO 中不同模块特征映射可视化结果

Fig. 5 Visualization results of feature maps for different modules in LIS - YOLO

块比 C2f－Faster 更能专注于复杂背景下鱼类的特征信息。这归功于引入的注意力机制收集多尺度空间信息的能力。相比之下，EMA 模块要分割的对象区域过于宽泛，而 C2f－Faster－EMA 更关注目标本身，使得检测到的目标区域更加具体。

由于数据集中包含正视、俯视和仰视 3 种不同视角的图像，而且存在明显的尺度变化、遮挡和水面高光的影响，因此为了保障细粒度轮廓质量和跨视角稳健性，在选择模型权重时需要考虑其具备适度容量以捕捉边界和形变。为了更加直观地看出不同权重模型的分割效果，表 3 给出了在相同训练配置下的 5 种不同权重的 YOLO v8 性能参数。通常来说，模型权重越大，其分割精度越高，但模型的参数量和分割时间也会增加。因此，为了平衡模型的分割精度和速度，选择 YOLO v8s 作为本研究的基准模型进行改进。

表 3　不同权重的 YOLO v8 模型性能

Tab. 3　Comparison of YOLO v8 models with different weights

模型	精度/%	召回率/%	mAP_{50}/%	浮点运算量	参数量	帧率/(f·s^{-1})
YOLO v8n	90.8	90.7	95.2	1.20×10^{10}	3.30×10^{6}	208
YOLO v8s	93.3	90.4	96.0	4.24×10^{10}	1.18×10^{7}	159
YOLO v8m	93.8	92.9	96.9	1.10×10^{11}	2.72×10^{7}	105
YOLO v8l	94.3	93.3	97.0	2.20×10^{11}	4.59×10^{7}	74
YOLO v8x	93.0	95.3	96.7	3.44×10^{11}	7.17×10^{7}	54

鱼类个体间存在遮挡现象，阻碍了模型分割精度的提高。一个良好的实例分割模型可以解决这个问题。因此，本文引入了注意力机制，从物体中提取更多的空间信息。图 6 为原始 YOLO v8s 模型、仅引入注意力机制的 YOLO v8s 模型（YOLO v8s－EMA）和所提出的 LIS－YOLO 模型的部分分割结果示例。图 6a 为各种模型对鲟鱼的识别结果，从原始图像的椭圆标注区域可以看出，LIS－YOLO 具有优越的分割性能。当面对两条纠缠在一起的鱼时，

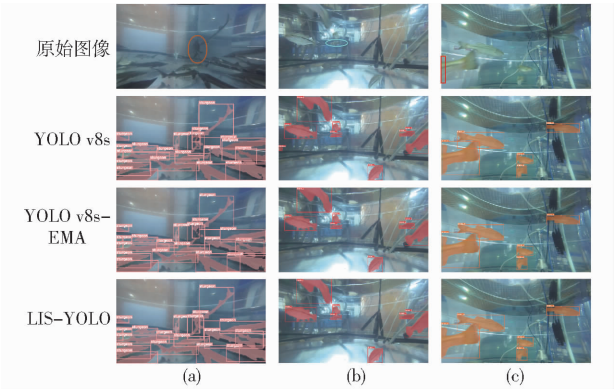


图 6　分割结果示例图

Fig. 6　Sample images of detection results

YOLO v8s 模型和 YOLO v8s－EMA 模型都难以有效地分割两条鱼。从图 6b 可以看出，LIS－YOLO 模型对鱼类的整体分割性能较好，而另外两种模型在处理距离相机较远的鱼类时，对鱼类轮廓的分割效果并不理想。YOLO v8s 和 YOLO v8s－EMA 模型丢失了一些图像边缘信息，而 LIS－YOLO 模型缓解了这一问题，如图 6c 中原始图像的红框区域所示。

2.4　不同拍摄角度下的分割结果及分析

不同的鱼有不同的生活习惯，比如有些鱼喜欢栖息在鱼缸的底部，而有些鱼喜欢上下游动。因此，图 7 给出了 3 种拍摄角度下不同鱼类的模型识别精度，以测试拍摄角度对模型识别精度的影响。

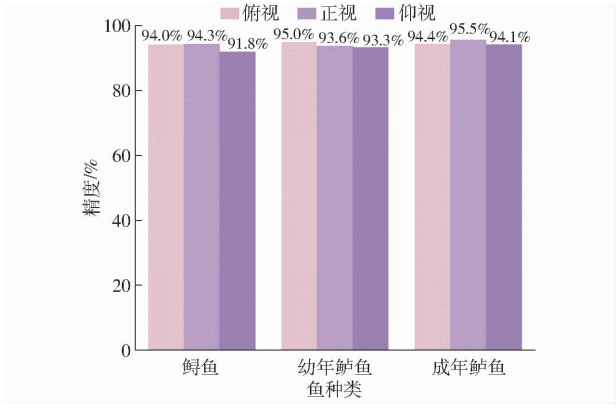


图 7　不同拍摄角度下的分割结果

Fig. 7　Detection results at different shooting angles

从图 7 中可以看出，相比于其他两种拍摄角度，正视拍摄的鱼类识别精度更高。然而，这种情况只适用于鲟鱼和成年鲈鱼，而不适用于幼年鲈鱼的识别结果。在幼年鲈鱼的分割结果中，俯视识别精度略高于正视识别精度。出现这种现象的原因可能是在正视图下鲟鱼和成年鲈鱼这两种鱼的鱼体轮廓特征明显，能够完整展示其关键特征，因此正视图下的分割效果较优。幼年鲈鱼由于其体型相比于其他两种鱼更小，正视图下可能因其侧向厚度不足导致特征稀疏，而俯视图能较为清晰展示其整体轮廓，再加上其游动更活跃，因此俯视图下的分割效果更优。值得注意的是，幼年鲈鱼和成年鲈鱼有相似的颜色和纹理，主要是尺寸不同，可能导致错误的识别。总的来说，3 种拍摄角度下的识别精度相似，都能满足实时分割的要求。

2.5　水下多目标鱼类实时分割系统

本文基于所提出的轻量化模型并结合 PyQt5 模块实现一个高度兼容且用户友好的 GUI 软件，该软件可以在各种硬件平台和操作系统环境中稳定运行，如图 8 所示。其优势具体表现为：可以上传用户自己的模型权重，并根据实际情况可以选择摄像头实时地对水下鱼类进行实例分割，或通过上传视频、

图像的形式进行分割。此外,预测结果会立即反馈到用户界面,相关的日志信息也会同步更新至日志区域,确保用户可实时获取准确且全面的信息,提高了软件的实用性。该软件可实现对目标计数的功能,在实时检测目标的同时能够将目标的数量显示在用户界面的下方,以清楚地告知用户相关的信息。



图 8 基于 PyQt5 开发的 GUI 可视化界面

Fig. 8 GUI visualization interface diagram based on PyQt5 development

3 讨论

3.1 LIS - YOLO 更轻量化

本文提出的 LIS - YOLO 模型是对原始 YOLO v8s 模型的改进。由表 2 可以看出,LIS - YOLO 模型的参数量较少,为原 YOLO v8s 模型的 76%。改进后的模型更轻量化的前提是保证识别精度不显著降低[32],LIS - YOLO 实现了这一目标。此外,C2f - Faster 模块取代原有的 C2f 模块是模型参数量减少的主要原因。但为了保证识别精度,引入了 EMA 注意力机制和 Wise - IoU,通过 C2f - Faster 模块与两者之间的协同作用最终获得了较好的识别精度。相比于 YOLO v8n,LIS - YOLO 的参数量和帧率并不是最优的,这是因为考虑到最终的分割精度和平均精度均值,不能过度压缩模型的参数量,因此这是平衡模型精度和帧率后的选择。

3.2 模型性能的影响因素

图像丰富度的限制:虽然使用不同的拍摄角度获得各种形态的水下鱼类图像数据,但由于鱼体是非刚性的,不可能获得所有鱼类姿态的图像数据,从而影响了模型的泛化能力[33]。图 7 为不同拍摄角度下的分割精度,可以看出,不同的拍摄角度对模型的识别效果影响不大,识别精度可以满足实例分割任务。

图像质量的限制:水下相机在拍摄鱼类时经常受到水的折射和反射的影响,导致图像质量较低,这给模型识别带来了另一个挑战[34]。此外,由于鱼类游动速度快,相机在拍摄过程中可能会捕捉到模糊的图像,导致模型的分割能力下降。

模型的局限性:LIS - YOLO 模型实现了原 YOLO v8 模型的轻量化,但模型的平均精度均值比原 YOLO v8 模型降低了 0.1 个百分点(见表 2)。模型识别精度的提高和轻量化是研究人员追求的目标之一,但两者很难同时实现[35]。因为模型识别精度的提高往往需要在原模型中引入适当的模块,而一些模块的引入会增加模型参数量[36]。因此,在保证识别精度不变或降低不多的情况下,实现模型的轻量化是可以接受的。

4 结束语

在利用边界框进行鱼类检测的传统目标检测模型中,通常在感兴趣区域内包含大量不相关信息,例如背景元素。实例分割模型可以根据鱼的目标轮廓进行分割,保证分割区域内不存在无关信息。因此,本文提出了一种基于 YOLO v8 模型的轻量化实例分割模型 LIS - YOLO,并基于 PyQt5 开发了水下多目标鱼类实例分割系统。该模型创新性地将原有的 C2f 模块修改为 C2f - Faster - EMA 模块,在保持模型精度的同时有效地减少了模型参数量。与原始的 YOLO v8 模型相比,本文提出的 LIS - YOLO 模型在保证平均精度均值基本不变的情况下,将模型参数量压缩了 23.7%。虽然 LIS - YOLO 的帧率比原始模型低 32 f/s,但仍然可以满足实时分割的要求。总体而言,所提出的 LIS - YOLO 模型既能准确分割不同形状的鱼类,又能有效分割小目标鱼类,具有良好的泛化性能。

参 考 文 献

[1] LI H, MA X, LIU H. Cross-modal feature fusion Mask R - CNN and point cloud normalization segmentation transformation for fish length estimation[J]. Aquaculture International, 2024, 32(7): 9201 - 9224.

[2] LI D, WANG Z, WU S, et al. Automatic recognition methods of fish feeding behavior in aquaculture: a review[J]. Aquaculture, 2020, 528: 735508.

[3] 聂鹏程, 钱程, 汪清平, 等. 赤点石斑鱼氨氮应激行为嵌入式表征研究[J]. 农业机械学报, 2025, 56(2): 503 - 510.
NIE Pengcheng, QIAN Cheng, WANG Qingping, et al. Behavioral representation for ammonia-nitrogen stress of *Epinephelus akaara* for embedded system[J]. Transactions of the Chinese Society for Agricultural Machinery, 2025, 56(2): 503 - 510. (in Chinese)

[4] LI D, YANG Y, ZHAO S, et al. A fish image segmentation methodology in aquaculture environment based on multi-feature

fusion model[J]. Marine Environmental Research, 2023, 190: 106085.
[5] YANG Y, ZHAO X, HUANG M, et al. Multispectral image based germination detection of potato by using supervised multiple threshold segmentation model and Canny edge detector[J]. Computers and Electronics in Agriculture, 2021, 182: 106041.
[6] LIN G, TANG Y, ZOU X, et al. Color-, depth-, and shape-based 3D fruit detection[J]. Precision Agriculture, 2019, 21(1): 1－17.
[7] SUN L, LUO B, LIU T, et al. Algorithm of adaptive fast clustering for fish swarm color image segmentation[J]. IEEE Access, 2019, 7: 178753－178762.
[8] XUAN K, DENG L, XIAO Y, et al. SO－YOLOv5: small object recognition algorithm for sea cucumber in complex seabed environment[J]. Fisheries Research, 2023, 264: 106710.
[9] 吴佳文, 姬伟, 翟科龙, 等. 基于 YOLO－Crab 与改进的 DeepSORT 的水下河蟹检测与计数方法[J/OL]. 电子测量技术, 1－10.[2025－07－02]. https://link.cnki.net/urlid/11.2175.TN.20250701.1727.043.
WU Jiawen, JI Wei, ZHAI Kelong, et al. A detection and counting method for underwater crabs based on YOLO－Crab and the improved DeepSORT[J/OL]. Electronic Measurement Technology, 1－10. [2025－07－02]. https://link.cnki.net/urlid/11.2175.TN.20250701.1727.043. (in Chinese)
[10] 孙鹏麒, 胡家祯, 黄小华, 等. 基于 YOLO 的养殖鱼群全向声呐实时监测方法研究与应用[J]. 中国水产科学, 2025, 32(2): 409－419.
SUN Pengqi, HU Jiazhen, HUANG Xiaohua, et al. Research and application of real-time monitoring method for cultured fish based on YOLO[J]. Journal of Fishery Sciences of China, 2025, 32(2): 409－419. (in Chinese)
[11] 徐龙琴, 郑钦月, 高学凯, 等. 基于轻量级 CDW－YOLO_v7 的鱼类排便行为自动检测方法[J]. 农业机械学报, 2025, 56(6): 554－564.
XU Longqin, ZHENG Qinyue, GAO Xuekai, et al. Automatic detection of fish defecation behavior based on lightweight CDW－YOLO_v7[J]. Transactions of the Chinese Society for Agricultural Machinery, 2025, 56(6): 554－564. (in Chinese)
[12] WAN Z, LI M, WANG Z, et al. CellT－Net: a composite transformer method for 2－D cell instance segmentation[J]. IEEE Journal of Biomedical and Health Informatics, 2024, 28(2): 730－741.
[13] XU W, ZHAO L, LI J, et al. Detection and classification of tea buds based on deep learning[J]. Computers and Electronics in Agriculture, 2022, 192: 106547.
[14] 王纪章, 姚承志, 周静, 等. 基于三维点云的黄瓜叶片分割与表型参数提取方法[J]. 农业机械学报, 2025, 56(3): 354－362.
WANG Jizhang, YAO Chengzhi, ZHOU Jing, et al. Cucumber leaf segmentation and phenotype extraction method based on three-dimensional point cloud[J]. Transactions of the Chinese Society for Agricultural Machinery, 2025, 56(3): 354－362. (in Chinese)
[15] SALVI M, MICHIELLI N, MEIBURGER K M, et al. cyto－Knet: an instance segmentation approach for multiple myeloma plasma cells using conditional kernels[J]. International Journal of Imaging Systems and Technology, 2023, 34(1): 1－14.
[16] WANG Y, GU L, JIANG T, et al. MDE－UNet: a multitask deformable UNet combined enhancement network for farmland boundary segmentation[J]. IEEE Geoscience and Remote Sensing Letters, 2023, 20: 1－5.
[17] YE Z, ZHOU J, JI B, et al. Feature fusion of body surface and motion-based instance segmentation for high-density fish in industrial aquaculture[J]. Aquaculture International, 2024, 32(6): 8361－8381.
[18] LIU Q, GONG X, LI J, et al. A multitask model for realtime fish detection and segmentation based on YOLOv5[J]. PeerJ Computer Science, 2023, 9: e1262.
[19] YU X, WANG Y, LIU J, et al. Non-contact weight estimation system for fish based on instance segmentation[J]. Expert Systems with Applications, 2022, 210: 118403.
[20] ZHU J, HU T, ZHENG L, et al. YOLOv8－C2f－Faster－EMA: an improved underwater trash detection model based on YOLOv8[J]. Sensors, 2024, 24(8): 2483.
[21] XU D, XIONG H, LIAO Y, et al. EMA－YOLO: a novel target-detection algorithm for immature yellow peach based on YOLOv8[J]. Sensors, 2024, 24(12): 3783.
[22] YUE Y R, CUI S H, SHAN W. Apple detection in complex environment based on improved YOLOv8n[J]. Engineering Research Express, 2024, 6(4): 045259.
[23] HE K, GKIOXARI G, DOLLÁR P, et al. Mask R－CNN[C]// Proceedings of the the IEEE International Conference on Computer Vision (ICCV). IEEE, 2017: 2980－2988.
[24] 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, 2017, 39(6): 1137－1149.
[25] CHEN J, KAO S H, HE H, et al. Run, don't walk: chasing higher FLOPS for faster neural networks[C]// Proceedings of the 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2023: 12021－12031.
[26] WOO S, PARK J, LEE J Y, et al. CBAM: convolutional block attention module[C]// Proceedings of the Proceedings of the European Conference on Computer Vision (ECCV). Munich: Springer, 2018: 3－19.

(下转第 139 页)

[16] ZHANG Jie, LIU-Henke X. Model-based design of the vehicle dynamics control for an omnidirectional automated guided vehicle (AGV)[C]//2020 International Conference Mechatronic Systems and Materials (MSM): IEEE, 2020.

[17] BADGUJAR Chetan, FLIPPO Daniel. Artificial neural network to predict traction performance of autonomous ground vehicle on a sloped soil bin and uncertainty analysis[J]. Computers and Electronics in Agriculture, 2022, 196: 106867.

[18] 张伟荣,陈学庚,齐江涛,等. 基于 IPO-VMD-GRNN 的田间四足机器人摔倒状态预测方法[J]. 农业机械学报, 2025, 56(2): 175-186.
ZHANG Weirong, CHEN Xuegeng, QI Jiangtao, et al. Field quadruped robot fall state prediction method based on IPO VMD GRNN[J]. Transactions of the Chinese Society for Agricultural Machinery, 2025, 56(2): 175-186. (in Chinese)

[19] 陈耀晖,李家一,鲍泽韩,等. 自主导航柑橘表型巡检机器人设计与试验[J]. 农业机械学报, 2025, 56(3): 49-57.
CHEN Yaohui, LI Jiayi, BAO Zehan, et al. Design and testing of autonomous navigation citrus phenotypic inspection robot [J]. Transactions of the Chinese Society for Agricultural Machinery, 2025, 56(3): 49-57. (in Chinese)

[20] LÓPEZ J, ZALAMAB E, GÓMEZ-GARCÍA-BERMEJO J, et al. A simulation and control framework for AGV based transport systems[J]. Simulation Modelling Practice and Theory, 2022, 116: 102430.

[21] 王宇,王文浩,徐凡,等. 基于改进蚁群算法的植保无人机路径规划方法[J]. 农业机械学报, 2020, 51(11): 103-112, 92.
WANG Yu, WANG Wenhao, XU Fan, et al. Path planning method for plant-protection UAV based on improved ant colony algorithm[J]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(11): 103-112. (in Chinese)

[22] CHEN J C, DU C L, ZHANG Y, et al. A clustering-based coverage path planning method for autonomous heterogeneous UAVs [J]. Ieee Transactions on Intelligent Transportation Systems, 2022, 23(12): 25546-25556.

[23] KODAGODA K R S, WIJESOMA W S, ZHANG Y, et al. Fuzzy speed and steering control of an AGV[J]. IEEE Transactions on Control Systems Technology, 2002, 10(1): 112-120.

[24] 李云伍,徐俊杰,王铭枫,等. 丘陵山区田间道路自主行驶转运车及其视觉导航系统研制[J]. 农业工程学报, 2019, 35(1): 52-61.
LI Yunwu, XU Junjie, WANG Mingfeng, et al. Development of an autonomous transport vehicle with vision navigation for field roads in hilly areas[J]. Transactions of the CSAE, 2019, 35(1): 52-61. (in Chinese)

[25] LI Sen, ZHANG Ran, DING Yuanming, et al. Multi-UAV path planning algorithm based on BINN-HHO[J]. Sensors, 2022, 22(24): 9786.

[26] PACEJKA H B, BAKKER E. The magic formula tyre model[J]. Vehicle System Dynamics, 1992, 21(Supp. 1): 1-18.

(上接第 101 页)

[27] HU J, SHEN L, SUN G. Squeeze-and-excitation networks[C]//Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. IEEE, 2018: 7132-7141.

[28] TONG Z, CHEN Y, XU Z, et al. Wise-IoU: bounding box regression loss with dynamic focusing mechanism[C]// Proceedings of the Computer Vision and Pattern Recognition. IEEE, 2023.

[29] REZATOFIGHI H, TSOI N, GWAK J, et al. Generalized intersection over union: a metric and a loss for bounding box regression[C]//Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2019: 658-666.

[30] XU W, LIU C, WANG G, et al. Behavioral response of fish under ammonia nitrogen stress based on machine vision[J]. Engineering Applications of Artificial Intelligence, 2024, 128: 107442.

[31] XI D, QIN Y, WANG S. YDRSNet: an integrated Yolov5-Deeplabv3+real-time segmentation network for gear pitting measurement[J]. Journal of Intelligent Manufacturing, 2021, 34(4): 1585-1599.

[32] YOU X, WANG Y, ZHAO X. A lightweight monocular 3D face reconstruction method based on improved 3D morphing models [J]. Sensors, 2023, 23(15): 6713.

[33] XU W, ZHU Z, GE F, et al. Analysis of behavior trajectory based on deep learning in ammonia environment for fish[J]. Sensors, 2020, 20(16): 4425.

[34] SONG W, WANG Y, HUANG D, et al. Enhancement of underwater images with statistical model of background light and optimization of transmission map[J]. IEEE Transactions on Broadcasting, 2020, 66(1): 153-169.

[35] LIU Y, ZHANG C, WU W, et al. MiniYOLO: a lightweight object detection algorithm that realizes the trade-off between model size and detection accuracy[J]. International Journal of Intelligent Systems, 2022, 37(12): 12135-12151.

[36] SHI C, ZHANG X, SUN J, et al. A lightweight convolutional neural network based on group-wise hybrid attention for remote sensing scene classification[J]. Remote Sensing, 2021, 14(1): 161.