

# 基于改进YOLOv8的南极磷虾群智能识别技术研究

杨浩东<sup>1\*</sup>, 郑汉丰<sup>1</sup>, 郭爱<sup>1</sup>, 李灵智<sup>1</sup>, 周纪东<sup>2</sup>, 戴阳<sup>1#</sup>

<sup>1</sup>中国水产科学研究院东海水产研究所, 农业农村部渔业遥感重点试验室, 上海

<sup>2</sup>中兵北斗应用研究院有限公司, 上海

收稿日期: 2025年12月16日; 录用日期: 2026年1月15日; 发布日期: 2026年1月21日

## 摘要

南极磷虾作为全球渔业中资源量最大的生物, 关系着极地渔业的捕捞效率和经济效益。南极磷虾的捕捞正逐步成为极地渔业发展的热点问题。随着水下声学探测技术与机器视觉的快速发展, 精准且富有效率的声呐图像识别成为南极磷虾智能化捕捞的前沿技术。鉴于目前人工声呐图像识别工作时长较长、占用捕捞船船员资源和识别准确率及效率较低的缺点, 本文提出YOLOv8s-CBAM模型架构。通过在YOLOv8s的主干网络和颈部嵌入CBAM(卷积块注意力模块), 实现了对复杂声呐背景的鲁棒性提升。利用平均/最大池化和通道-空间双维度动态加权, 并结合C2f结构优化特征聚合效率。强化了模型对南极磷虾的特征响应能力。实验结果表明, YOLOv8s-CBAM在保持轻量级架构的同时, 实现了mAP50达到0.92, 识别速度超过30 FPS的性能平衡, 在定量分析与定性分析优于YOLOv8及YOLOv8-BiFPN等常用模型。

## 关键词

南极磷虾, 智能捕捞, 声学探测技术, YOLOv8, 资源密度, 极地海域

# Research on Intelligent Identification Technology of Antarctic Krill Swarms Based on Improved YOLOv8

Haodong Yang<sup>1\*</sup>, Hanfeng Zheng<sup>1</sup>, Ai Guo<sup>1</sup>, Lingzhi Li<sup>1</sup>, Jidong Zhou<sup>2</sup>, Yang Dai<sup>1#</sup>

<sup>1</sup>East China Sea Fisheries Research Institute, Chinese Academy of Fishery Sciences, Key Laboratory of Fisheries Remote Sensing, Ministry of Agriculture and Rural Affairs, Shanghai

<sup>2</sup>China National Armament Beidou Application Research Institute Co., Ltd., Shanghai

Received: December 16, 2025; accepted: January 15, 2026; published: January 21, 2026

\*第一作者。

#通讯作者。

文章引用: 杨浩东, 郑汉丰, 郭爱, 李灵智, 周纪东, 戴阳. 基于改进 YOLOv8 的南极磷虾群智能识别技术研究[J]. 计算机科学与应用, 2026, 16(1): 230-241. DOI: 10.12677/csa.2026.161019

## Abstract

Antarctic krill, as the organism with the largest biomass in global fisheries, is intrinsically linked to the fishing efficiency and economic viability of polar fisheries. The harvesting of Antarctic krill is increasingly emerging as a focal issue in the development of polar fisheries. With the rapid advancements in underwater acoustic detection technology and machine vision, precise and efficient sonar image recognition has become a cutting-edge technology for intelligent krill harvesting. In light of the limitations of current manual sonar image recognition—such as prolonged processing times, substantial consumption of fishing vessel crew resources, and sub optimal accuracy and efficiency—this paper proposes the YOLOv8-CBAM model architecture. By integrating the CBAM into the backbone and neck of the YOLOv8s architecture, achieve enhanced robustness against complex sonar backgrounds. This integration utilizes average/max pooling along with channel-spatial dynamic weighting to reinforce the model's feature response capabilities, specifically for Antarctic krill. The architecture maintains optimized feature aggregation efficiency by incorporating the C2f structure. Experimental results demonstrate that YOLOv8s-CBAM strikes an excellent balance between performance and lightweight architecture, achieving an mAP50 of 0.92 and a detection speed exceeding 30 FPS. Furthermore, both quantitative and qualitative analyses confirm its superior performance compared to popular baseline models such as standard YOLOv8 and YOLOv8-BiFPN.

## Keywords

Antarctic Krill, Intelligent Fishing, Acoustics Detection Technology, YOLOv8, Resource Density, Polar Water

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

南极磷虾富含蛋白质及多种人体必需的矿物质元素以及活性酶、虾青素等多种活性成分, 具有极高经济价值, 已成为极地地区重要的渔业资源[1]。高效、精准地对南极磷虾进行捕捞对极地渔业资源开发具有重要意义[2]-[4]。目前声学探测已成为最常用且高效的海洋生物监测手段, 依托声学设备可实现对南极磷虾的种群分布进行广域的实时观测[5][6], 通过分析声呐回波信号的强度、频率响应等特征, 可推断磷虾群体的密度与垂直空间分布以获得最佳捕捞深度, 但此类数据通常高度复杂, 从中精准提取关键信息的技术难度较高。传统声呐图像识别方法依赖人工目视解读声呐回波图像, 存在精度不足与实时性差的问题。以一般南极磷虾捕捞船为例, 其船员配额通常为 40~70 人, 除船舶运行值班人员以及轮班休息人员外, 参与渔捞作业的船员仅为 15~19 人[7], 且人工目视解读声呐回波图像为连续性工作, 将极大增加南极磷虾捕捞船船员工作时长, 增加船员工作负担。

随着深度学习技术的快速发展, 为声呐数据处理和人工解读图像困难的问题提供了解决方案[8]-[10]。随着 YOLO 系列算法的迭代, 不仅提升了模型鲁棒性, 也显著改善目标检测与分类性能, 为深度学习在图像识别中的应用奠定基础[11]-[13]。并因其端到端的检测结构与高检测速度被广泛应用于声呐图像等目标检测场景[14]-[18]。如 YOLOv3、YOLOv5、YOLOv7 等已经在声呐图像检测中实现了显著性能提升, 在不同声学数据集上的检测精度达到 89%以上[19]-[21]。

研究表明, 构建一种兼具注意力机制、特征提取的轻量化声呐图像识别网络是面向南极磷虾声呐图

像识别的关键方向[22]-[26]。对此，本文提出用于南极磷虾群识别的改进 YOLOv8s-CBAM 检测模型。该模型在 YOLOv8s 的特征融合网络(Neck)关键节点嵌入 CBAM (Convolutional Block Attention Module)注意力模块，构建了一种注意力引导的特征增强机制。通过集成通道与空间双维度的多路池化策略(平均池化与最大池化)，模型能够自适应地重校准特征权重，有效剔除海底混响与假底干扰，同时显著增强对磷虾群目标的特征响应。实验结果表明，该改进策略在保留 C2f 模块高效特征聚合能力的同时，大幅提升了模型在复杂水声环境下的鲁棒性与识别精度。

## 2. 基于 CBAM 模块的 YOLOv8s 模型

由于渔船上工作状况较差，船载计算机一般采用工控机，其抵抗恶劣环境能力较强，性能稳定，但性能特别是算力资源受限；而声呐图像中南极磷虾群目标特征易受噪声、藻类回声及海底及假底噪声干扰影响，传统检测模型难以在效率与精度之间取得平衡，需要改进一般的 YOLO 检测算法，以适应这种特殊要求。本文基于 YOLOv8s 架构，提出了一种深度优化的轻量化注意力检测模型——YOLOv8s-CBAM。该模型在特征融合网络(Neck)的关键路径(P3/P4)中战略性地嵌入 CBAM 注意力模块，利用其串行通道 - 空间注意力机制，有效抑制海底回波与环境杂波，并显著增强对微弱磷虾群信号的特征提取能力。配合噪声自适应训练策略，该模型在维持低计算开销的同时，显著提升了复杂海底环境下目标检测的鲁棒性与精度。为南极磷虾资源监测与智能捕捞系统提供了可靠的技术支撑。该方法的网络模型架构图与主干网络如图 1 和图 2 所示。

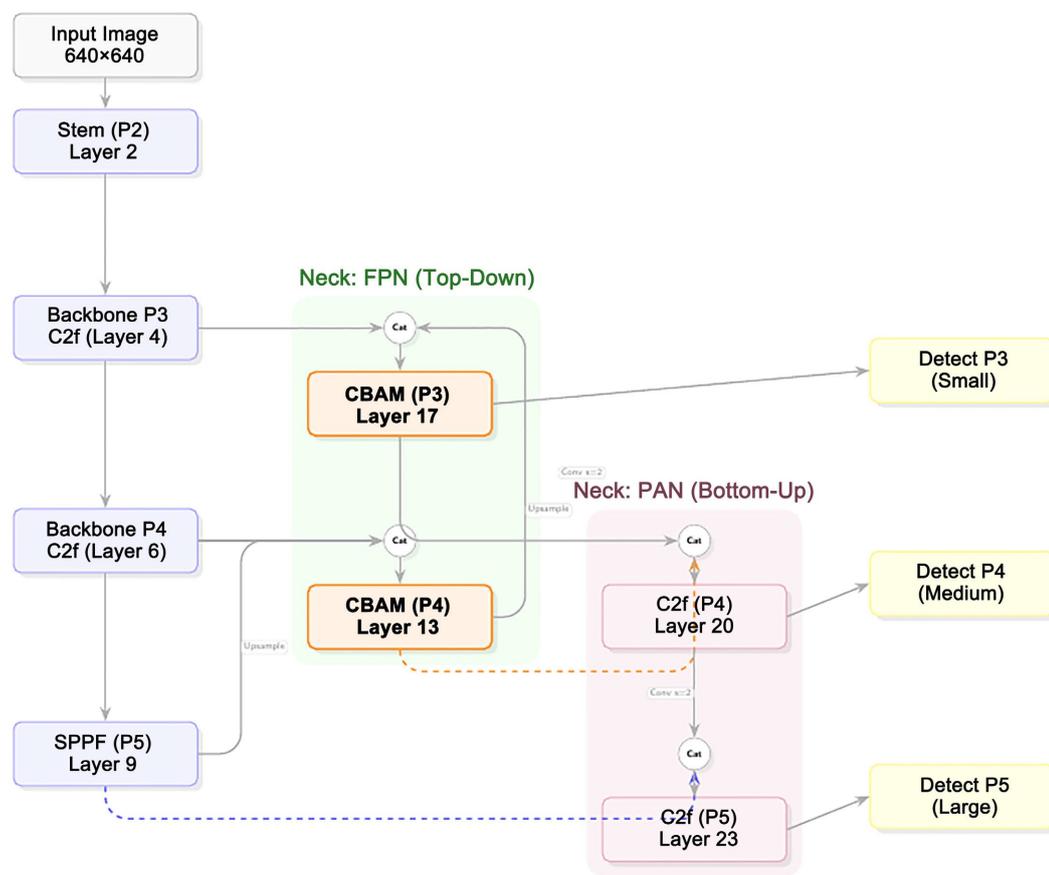


Figure 1. Framework diagram of the YOLOv8s-CBAM network model  
图 1. YOLOv8s-CBAM 网络模型框架图

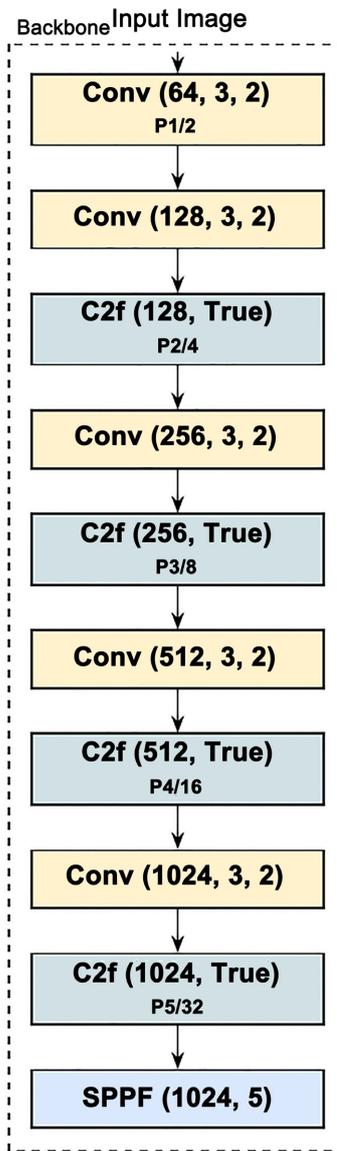


Figure 2. Main network diagram of the YOLOv8s-CBAM model

图 2. YOLOv8s-CBAM 模型主干网络图

在模型架构层面，通过在特征金字塔网络(Neck)的特征融合与精炼阶段的 P4 (16 倍下采样，对应中等目标)和 P3 (8 倍下采样，对应小目标)特征层集成了 CBAM 注意力模块保证模型处理机制，采用通道注意力通过自适应重标定特征通道权重，有效增强模型对磷虾回波与背景噪声的语义辨识度；随后采用空间注意力利用  $7 \times 7$  卷积核进一步细化特征，强化了网络对小规模磷虾群目标空间分布的感知能力。这种设计确保了特征去噪，在低计算增量下显著提升了特征质量。

为进一步抑制背景噪声导致的误检，增加准确率，模型重构了损失函数权重体系，将分类损失权重 ( $\lambda_{cls}$ )提升至 1.5，同时将定位损失权重下调至 5.0。确立了“磷虾群分类准确性优先”的导向，使网络专注于学习目标本质声学特征而非单纯的边界回归。其次，针对声呐数据的低信噪比特性，采用 AdamW 优化器配合低初始学习率以确收敛稳定性。从而保证了模型在真实声呐环境下的泛化性能。CBAM 模块结构图如图 3 所示。

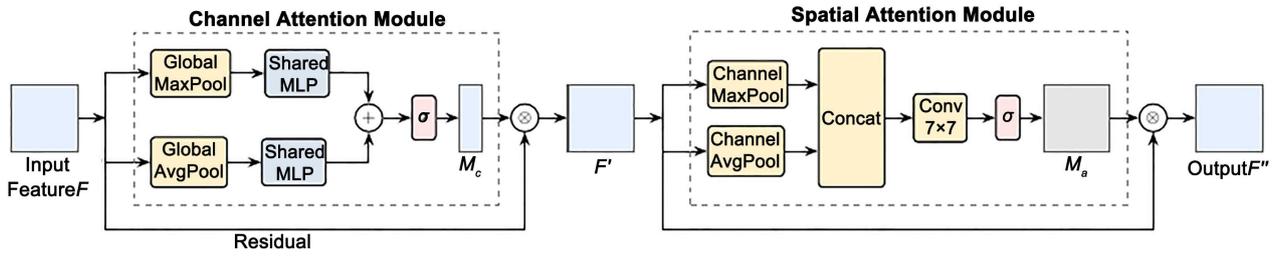


Figure 3. Schematic diagram of the CBAM module structure  
图 3. CBAM 模块结构示意图

### 3. 实验与分析

#### 3.1. 训练数据与参数

为验证本文所提出方法的可行性和实用性，本文使用南极磷虾生产船“深蓝号”采集的声呐数据进行研究。“深蓝号”作为我国第一艘自主研发建造的渔业捕捞加工船，主要用于远洋捕捞，采集的声学数据均来自科学回声探测仪设备 EK80。EK80 作为现代渔业声学调查的重要工具，具有高精度、多频率的数据采集能力，在渔业资源探测和捕捞生产中得到了广泛应用。本文使用的频率 EK80 设备工作频率为 38 kHz。数据采集时间为 2023 年 1 月 23 日到 2023 年 6 月 16 日，数据采集地点在南极磷虾 48 渔区。声学数据探测范围为 0~1100 m，探测数据包括磷虾信息，还有海底深度信息，这些信息为进一步研究海底识别及深度计算方法提供了支撑。采用的 EK80 数据单 ping 数据声学数据由 75,765 个回声点组成，探测精度达到  $1.5 \times 10^{-2}$  m，满足磷虾识别的数据精度需求。

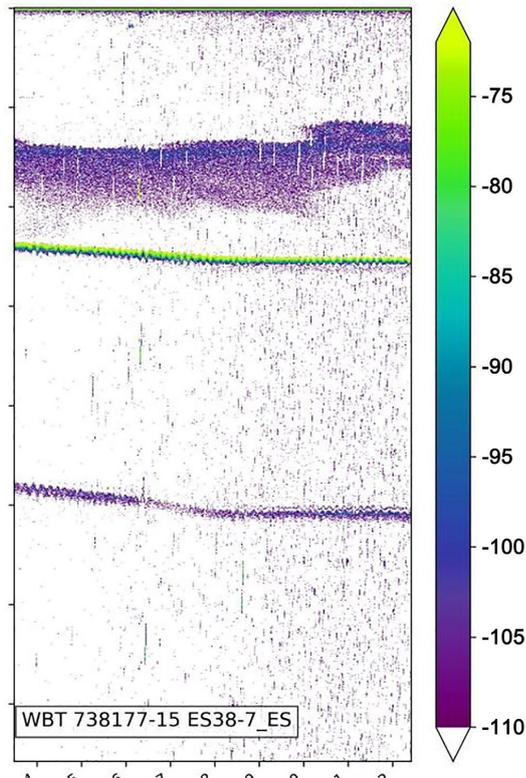


Figure 4. Sonar image of Antarctic krill after preprocessing  
图 4. 预处理后的南极磷虾声呐图像

目前解析声呐数据主要使用 Echopype Python 工具包。其作为一个为实现声学数据可扩展性而构建的工具包[27], 在海洋声呐数据处理中发挥了重要作用, 这些数据广泛用于获取包括南极磷虾在内的各种海洋生物的有关分布和资源情况。但由于 Echopype Python 工具包的声呐图像处理功能对计算平台要求较高, 要求 CPU 算力较高, 船载工控机平台运行困难, 因此, 本文设计基于 python 的声呐图像预处理功能, 将 raw 声呐数据转为 nc 格式, 然后转换图像并进行声呐图像降噪预处理后输入 YOLOv8s-CBAM 模型进行训练。经过预处理后的声呐图像如图 4 所示。

通过预处理后的声呐图像共 1949 张, 采用随机原则按照 8:1:1 的比例划分数据集, 其中训练集 1560 张, 测试集 194 张, 验证集 195 张。

模型训练使用的环境为 python3.12, 使用的 IDE 为 pycharm2025.1.1。通过实验表明, 该模型进行 70 轮训练后, 其准确率和召回率均趋于稳定。因此, 为了进一步优化模型性能, 最终决定将训练轮次设定为 100 轮。考虑到本研究提出的模型 GFLOPs 较低, 理论上在 NVIDIA Jetson TX2 或 Xavier 等边缘计算设备上可满足实时检测需求。未来的工作将在实际船载硬件平台上进一步验证其推理速度。目前为了进一步模拟船载工控机算力, 选用的网络模型训练环境参数如表 1 所示。

**Table 1.** Model training environment parameters

**表 1.** 模型训练环境参数

Pytorch	2.5.1
Python	3.12
Cuda	12.1
GPU	NVIDIA-RTX3050 Laptop
CPU	12th Gen Intel(R)Core(TM) i5-12400
操作系统	Windows 11

### 3.2. 定量分析

为进一步定量评价本文方法的可行性和实用性, 本文选择常用的精确度(Precision)、召回率(Recall)和 mAP, fps, 单张处理时长五个指标进行定量评估。将制作的数据集输入神经网络模型进行训练, 记录训练时的参数变化, 计算模型的查准率和召回率。计算公式如下

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

式中: TP 为识别正确, 模型认为是正样本; FP 为识别错误, 模型认为是正样本; FN 为识别错误, 模型认为是负样本。

平均精度(Average precision, AP)表示查准率 - 召回率(Precision-Recall)曲线下的面积, 对该图每一类的平均精度求均值即平均精度均值(Mean average precision, mAP), 其计算公式如下

$$\text{mAP} = \frac{\sum \text{AP}}{N}$$

式中: N 为该图不同类别目标的总和, 本文  $N = 1$ 。

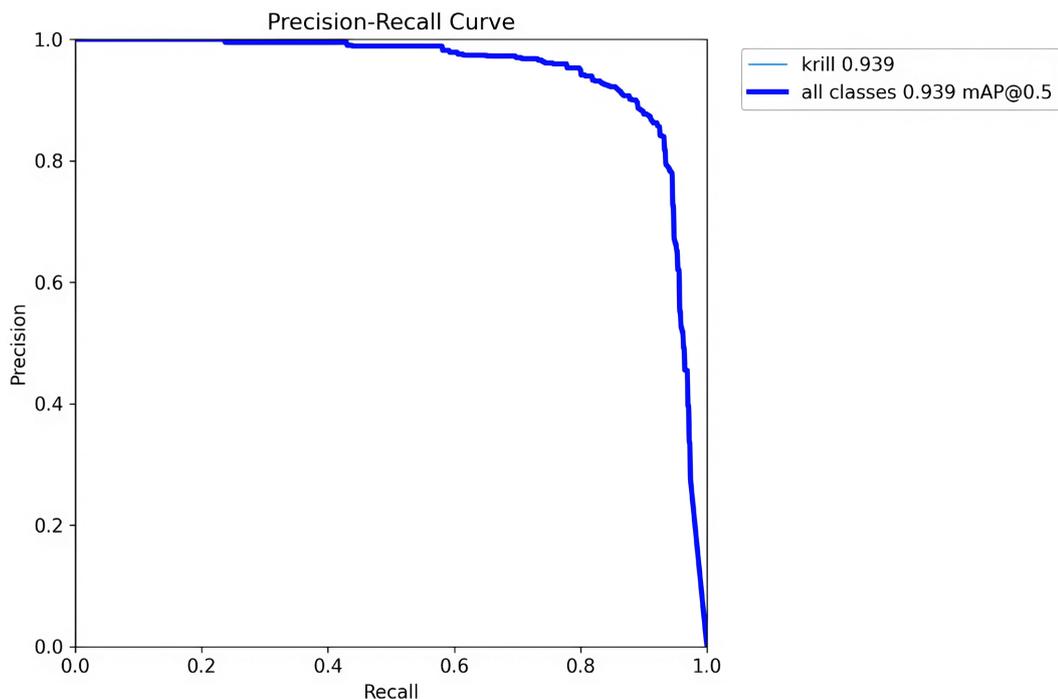
每秒帧数(FPS), 表示一秒内能够处理和识别(推理)多少张图像, 计算公式如下

$$FPS = \frac{1000}{\text{单张图像推理耗时 (ms)}}$$

网络训练过程中设置 batchsize = 16, 图像尺寸为 640 × 640。其余参数均相同。模型训练结果如表 2, YOLOv8s-CBAM 模型 P-R 曲线图如图 5 所示:

**Table 2.** Comparison of experimental results of different network models  
**表 2.** 不同网络模型实验对照效果

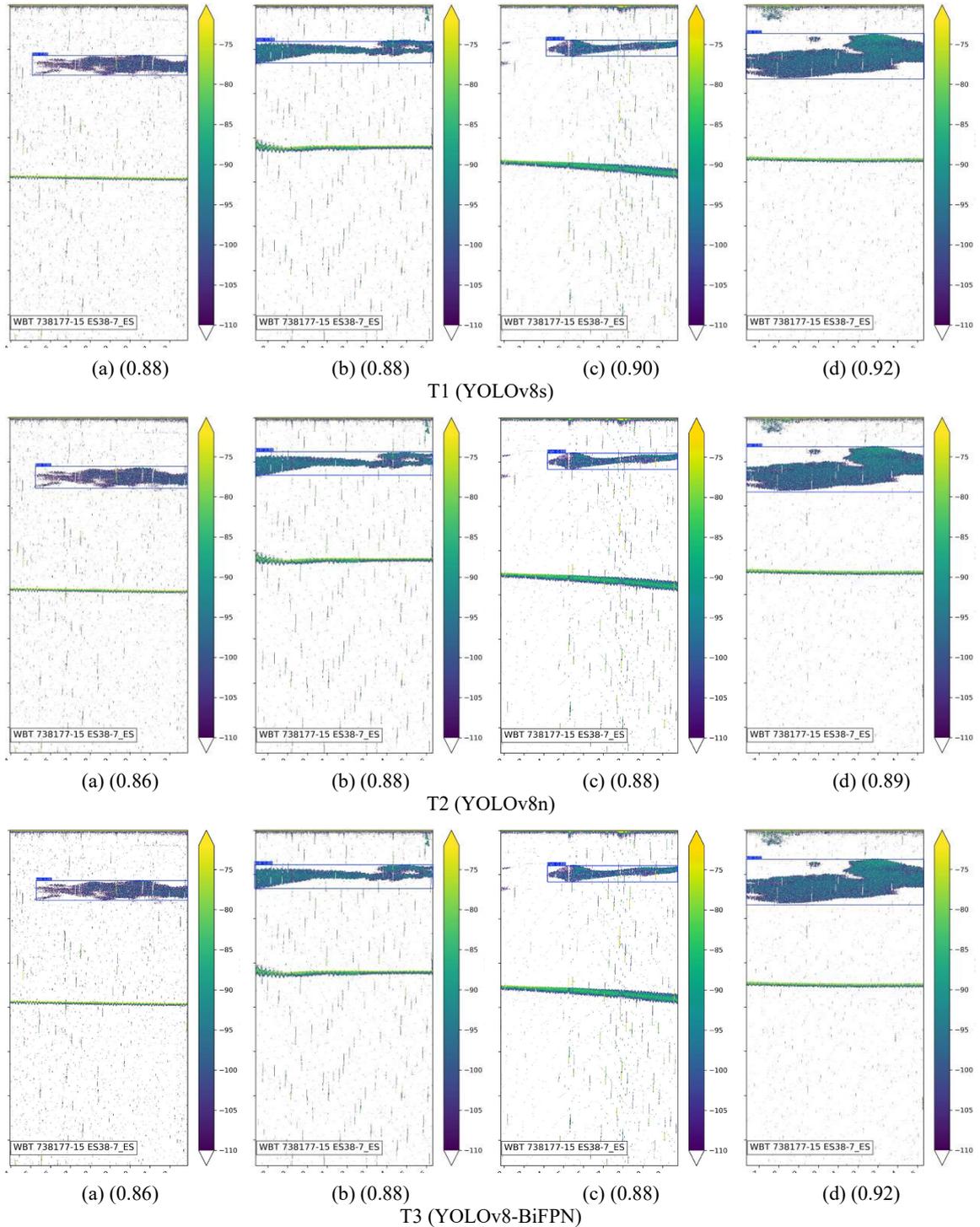
组别	模型	Precision	Recall	mAP	fps	Parameters	GFLOPs
1	YOLOv8s	83.56%	83.80%	85.22%	35	11.2	28.6
2	YOLOv8n	74.15%	80.31%	80.55%	40	3.2	8.7
3	YOLOv8-BiFPN	85.31%	82.51%	85.23%	28	12.5	32.1
4	YOLOv8-BiFPN-ghostv2	83.90%	86.56%	86.04%	25	10.8	25.4
5	YOLOv8m	86.05%	85.51%	86.55%	22	25.9	78.9
6	YOLOv8s-CBAM	91.70%	87.20%	91.61%	37	11.5	29.2



**Figure 5.** P-R curve of the YOLOv8s-CBAM model  
**图 5.** YOLOv8s-CBAM 模型 P-R 曲线图

对照试验以 YOLOv8s 模型为基准, 采用了 YOLOv8n、YOLOv8-BiFPN 等常用算法进行算法的定量分析。YOLOv8n 通过更小的模型尺寸, 虽损失了一定的准确性但提升检测速度; YOLOv8-BiFPN 作为引入 BiFPN 模块之后的检测模型, 利用双向特征融合和加权融合机制, 更有效地结合不同尺度的特征信息, 相较于 YOLOv8s 的准确性将有所增加; YOLOv8-BiFPN-ghostv2 为进一步优化 BiFPN 模块的 YOLOv8-BiFPN 检测模型, 通过 Ghost 卷积引入轻量化设计, 减少计算量并提升特征表达效率, 各项定量评估指标相较于 YOLOv8-BiFPN 均有所提升; YOLOv8m 沿用 YOLOv8 的核心架构, 包括 C2f 模块、PANet 和

解耦检测头，牺牲了一部分推理速度，换取更高的平均精度；本文提出的 YOLOv8s-CBAM 模型通过在通道和空间维度上施加自适应权重，增强了模型对关键特征信息的聚焦能力，提升了网络的特征表示质量，从而优化了目标检测性能。相较于 YOLOv8s 以及其他模型，各项评估指标总体取得了明显提升。以 YOLOv8s 模型为基准，Precision 提升 9.7%，Recall 提升 1.6%，mAP 提升 3.7%，fps 提升 5.7%，验证了本文所提出方法的可行性和有效性。



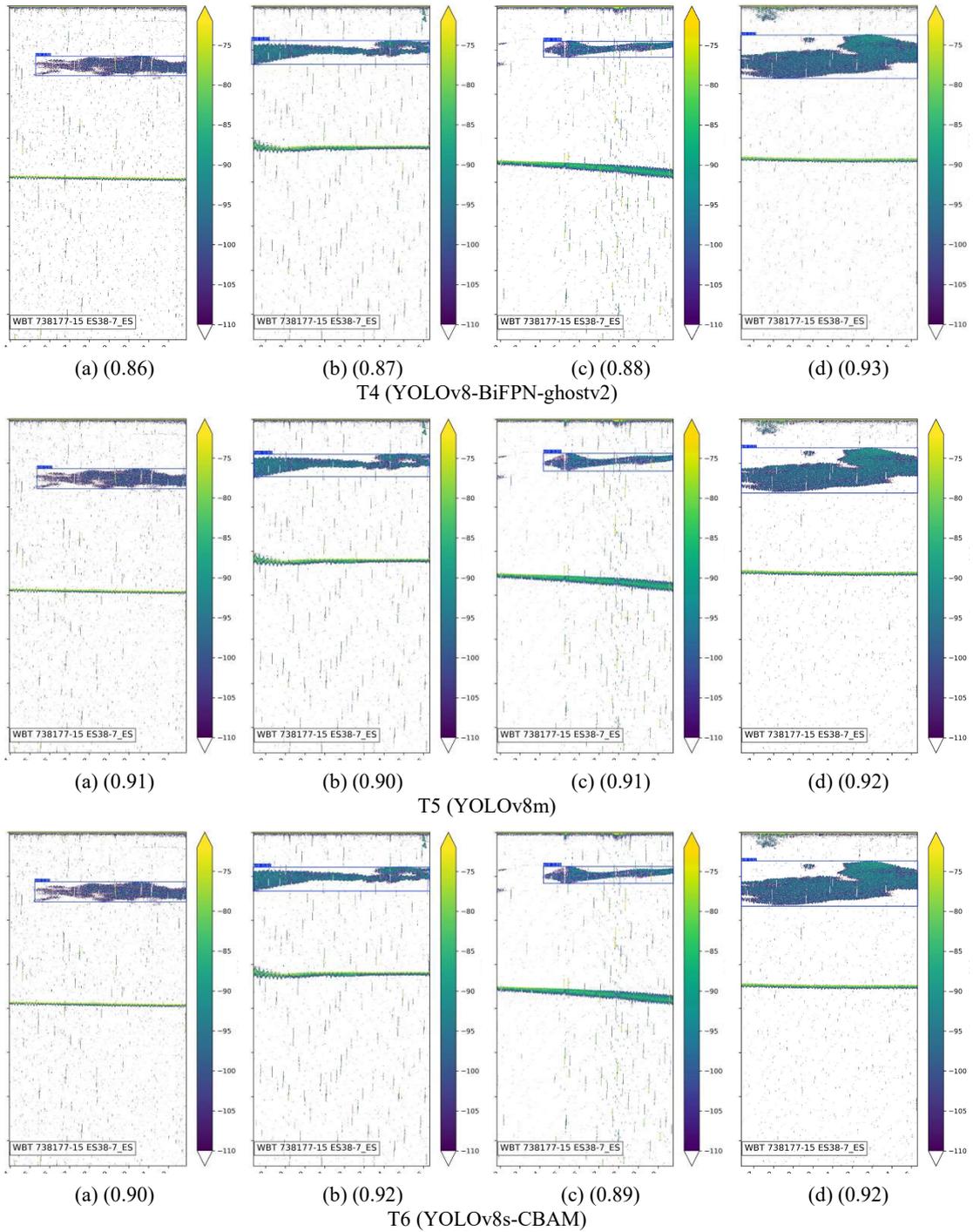


Figure 6. Comparison of detection results of different models  
 图 6. 不同模型检测结果比较图

### 3.3. 定性分析

为进一步验证本文方法南极磷虾识别的有效性，我们选择了不同时间尺度、磷虾群密度不同的声呐图像对不同架构的模型进行测试，分别选用(a) 理想水声环境下的单体高密度磷虾群，目标回波强度显著高于背景噪声，边界清晰；(b) 复杂的贴底磷虾生物群落场景，由于南极磷虾具有昼夜垂直移动习性，常

贴近海床分布。声呐回波中，高强度的海底反射信号与磷虾群信号在空间上高度重叠，极易造成误检；(c) 小尺度的密集磷虾群场景，目标尺寸跨度大，且群体垂直间距狭窄。(d) 大尺度生物量，同时也未漏检边缘的离散小群体。这四张声呐图像构建了各种工况下的梯度验证体系，全面量化模型在复杂背景抑制、微弱特征捕捉及尺度泛化方面的鲁棒性，从而证明算法在真实南极捕捞作业中的可靠性。下图分别为各模型的测试结果图，图中蓝色实线框表示识别出的南极磷虾群，边界框上方的数值表示预测的置信度分数(Confidence Score)。测试结果如图 6 所示。

图 6 中，T1 为 YOLOv8s 基准模型的检测结果，作为性能参考；T2 为 YOLOv8n 模型的检测结果，其通过浅层窄通道轻量化设计显著提升处理速度，但因特征表示能力减弱导致整体精度下降，在声呐噪声场景中鲁棒性不足；T3 为 YOLOv8-BiFPN 模型的检测结果，该模块通过加权双向融合强化多尺度特征交互，提升小目标定位效能，Recall 与 Precision 均获平衡优化；T4 为 YOLOv8-BiFPN-Ghostv2 模型的检测结果，其集成 Ghost 卷积廉价变换压缩冗余，增强特征表达效率并提升各项指标，置信度整体稳定，检测框精度未见显著下降；T5 为 YOLOv8m 模型的检测结果，沿用 YOLOv8 的核心架构，包括 C2f 模块、PANet 和解耦检测头，牺牲了一部分推理速度，换取更高的平均精度；T6 为本文 YOLOv8s-CBAM 模型的检测结果，针对相同目标，其检测精度相较 T1~T4 均有提升，目标框更精确且置信度更高。与 T5 相比在检测精度上近似相同，但检测速度更快。验证了算法的有效性。

## 4. 总结与展望

### 4.1. 总结

针对南极磷虾声学图像识别中存在的设备依赖强、实时处理延迟高以及噪声干扰导致的精度不足问题，本文基于 YOLOv8s 架构，提出了一种集成 CBAM 注意力机制与噪声自适应训练策略的改进模型——YOLOv8s-CBAM。通过嵌入通道-空间注意力模块与优化损失-增强机制，有效强化了在复杂水下声学环境下的目标辨识鲁棒性与计算效率，适用于极地渔业资源监测与智能捕捞系统。

首先，在主干网络(Backbone)构建阶段，模型采用 YOLOv8s 版本的复合缩放系数重构特征提取路径。相较于 YOLOv8n 版本，该设计显著增加了卷积通道数与模块堆叠深度，提升了模型对声学图像特征的表达容量；同时，末端的 SPPF 空间金字塔池化模块进一步强化了对磷虾群深度尺度分布的全局捕捉能力，缓解了声学图像中低信噪比目标的特征稀疏问题。

其次，在 Neck 特征融合阶段嵌入 CBAM (Convolutional Block Attention Module)注意力模块。CBAM 模块被置入于特征金字塔的 C2f 模块，针对中尺度通道与小尺度通道特征进行后处理增强。采用通道注意力分支通过全局平均/最大池化与 MLP 捕捉跨通道依赖；空间注意力分支则利用  $7 \times 7$  卷积压缩通道信息以生成空间权重。两者通过乘法门控机制协同工作，有效抑制背景噪声，显著提升模型对声呐图像中磷虾、海底与假底三类目标的区分能力。

最后，结合噪声导向的训练优化策略替换原始参数配置。将分类损失权重提升至并适当降低定位损失权重，以优先保证类别判别的准确性，抑制海底和假底导致的假阳性误检，确保模型在包含大量噪点的声学数据集上实现稳定收敛与特征学习。实验结果表明，YOLOv8s-CBAM 在南极磷虾声学数据集上实现了性能增益，同时推理速度与基准 YOLOv8s 保持近似，在维持轻量化优势的前提下，定量与定性分析均优于 YOLOv8-BiFPN 等对比模型。该方法验证了在极地声学探测场景下的实用可行性，为南极磷虾种群评估与智能化渔业管理提供了可靠技术支撑。

### 4.2. 展望

本文提出的 YOLOv8s-CBAM 实现了通道注意力与空间双维度的动态加权，有效强化了深层语义特

征的提取,并在上采样阶段对海底与假底进行过滤,准确识别南极磷虾。利用 CBAM 的平均/最大池化机制,显著增强了对南极磷虾等关键目标的特征响应能力,提升了对复杂背景的鲁棒性,能够应对水下噪声干扰与目标稀疏的挑战,为极地渔业智能化捕捞提出了思路。尽管该模型在特征提取与感知优化方面取得了理论进展,但仍需面对实际部署中的多变环境适应性问题,如多频段声学数据融合与船载平台的实时性约束。此外,目前的识别工作主要基于单帧声呐图像。考虑到实际捕捞作业的连续性,未来的研究将致力于结合 DeepSORT 或 ByteTrack 等多目标跟踪算法,将模型从短时间单帧检测扩展到视频流下的磷虾群持续跟踪与轨迹预测,以提供更具动态价值的捕捞辅助信息。

未来研究将聚焦多模态集成策略,将声学图像与水下摄像头、物理海洋数据相结合,利用多种数据共同提升磷虾密度动态估算的准确性;同时,引入知识蒸馏与神经网络量化技术,进一步压缩模型体积,适应南极磷虾捕捞船的低算力平台的应用需求。随着新的网络架构设计的出现,南极磷虾声呐图像的识别能力也将出现更优的方案,推动声学目标检测向更智能的方向演进,促进南极渔业从传统经验向数据驱动转型,最终助力全球海洋资源可持续利用与生态保护。

## 基金项目

中国水产科学研究院院级基本科研业务费专项: 2025CG03 南极磷虾无人船探测关键技术研究。

## 参考文献

- [1] 李琳, 等. 基于近红外光谱技术的南极磷虾品质快速评定[J]. 食品工业科技, 2026, 47(1): 318-325.
- [2] Cox, M.J., Smith, A.J.R., Brierley, A.S., Potts, J.M., Wotherspoon, S. and Terauds, A. (2023) Scientific Echosounder Data Provide a Predator's View of Antarctic Krill (*Euphausia superba*). *Scientific Data*, **10**, Article No. 284. <https://doi.org/10.1038/s41597-023-02187-y>
- [3] Bairstow, F., Gastauer, S., Wotherspoon, S., Brown, C.T.A., Kawaguchi, S., Edwards, T., et al. (2022) Krill Biomass Estimation: Sampling and Measurement Variability. *Frontiers in Marine Science*, **9**, Article ID: 903035. <https://doi.org/10.3389/fmars.2022.903035>
- [4] 屈泰春, 陈帅, 黄洪亮, 等. 南极大磷虾声学调查数据噪声处理与资源密度评估[J]. 极地研究, 2014, 26(4): 451-459.
- [5] 姚宇青, 戴阳, 王鲁民, 等. 基于层次分析法的南极磷虾瞄准捕捞网口路径规划[J]. 海洋渔业, 2022, 44(5): 598-609.
- [6] Zhao, Y., Wang, X., Zhao, X. and Ying, Y. (2022) A Statistical Assessment of the Density of Antarctic Krill Based on "Chaotic" Acoustic Data Collected by a Commercial Fishing Vessel. *Frontiers in Marine Science*, **9**, Article ID: 934504. <https://doi.org/10.3389/fmars.2022.934504>
- [7] <https://www.akerbiomarine.com/antarctic-endurance>
- [8] 刘凡诚, 邢传玺, 魏光春, 等. 一种改进 YOLOv8 的水下声呐图像目标检测方法[J]. 应用科技, 2025, 52(1): 34-40.
- [9] 刘彪. 基于声呐图像处理的水下目标探测关键技术研究[D]: [硕士学位论文]. 大理: 大理大学, 2024.
- [10] Lee, S., Park, B. and Kim, A. (2018) Deep Learning from Shallow Dives: Sonar Image Generation and Training for Underwater Object Detection.
- [11] Huang, H., Li, B., Liu, J., et al. (2024) Sonar Image Underwater Target Recognition: A Comprehensive Overview and Prospects. *Journal of Electronics & Information Technology*, **46**, 1742-1760.
- [12] 罗逸豪, 刘奇佩, 张吟, 等. 基于深度学习的水下图像目标检测综述[J]. 电子与信息学报, 2023, 45(10): 3468-3482.
- [13] 马启星. 基于深度学习声呐图像目标识别与跟踪研究[D]: [硕士学位论文]. 镇江: 江苏科技大学, 2022.
- [14] 陆彬, 毛义萱, 王露. 基于改进 YOLOv8 的侧扫声呐图像目标检测方法研究[J]. 水利水电快报, 2025, 46(1): 36-42.
- [15] 葛慧林, 戴跃伟, 朱志宇, 等. 基于改进 YOLOv7 声光融合水下目标检测方法[J]. 舰船科学技术, 2023, 45(12): 122-127.
- [16] 赵冬冬, 谢墩翰, 陈朋, 等. 基于 ZYNQ 的轻量化 YOLOv5 声呐图像目标检测算法及实现[J]. 光电工程, 2024,

---

51(1): 60-72.

- [17] 郭亚龙. 基于深度学习的水下声呐图像目标检测方法研究[D]: [硕士学位论文]. 济南: 山东交通学院, 2023.
- [18] Zhang, H., Tian, M., Shao, G., Cheng, J. and Liu, J. (2022) Target Detection of Forward-Looking Sonar Image Based on Improved YOLOv5. *IEEE Access*, **10**, 18023-18034. <https://doi.org/10.1109/access.2022.3150339>
- [19] Chen, Z., Xie, G., Deng, X., Peng, J. and Qiu, H. (2024) DA-YOLOv7: A Deep Learning-Driven High-Performance Underwater Sonar Image Target Recognition Model. *Journal of Marine Science and Engineering*, **12**, Article No. 1606. <https://doi.org/10.3390/jmse12091606>
- [20] Yu, S. (2022) Sonar Image Target Detection Based on Deep Learning. *Mathematical Problems in Engineering*, **2022**, Article ID: 5294151. <https://doi.org/10.1155/2022/5294151>
- [21] Ma, C., Wang, H., Zhang, K., Shen, G. and Chen, Z. (2025) LMFNet: Lightweight Multiscale Feature Fusion Network for Underwater Structural Defect Detection. *Journal of Field Robotics*, **43**, 376-386. <https://doi.org/10.1002/rob.70041>
- [22] 王维捷, 樊钢洲, 赵宪勇, 等. 基于渔业水平扫描声呐数据的南极磷虾集群回波信号识别及特征分析[J]. *中国水产科学*, 2025, 32(10): 1478-1488.
- [23] Ji, H., Zhu, D. and Chen, M. (2025) A Side-Scan Sonar Seabed Target Detection Algorithm Based on YOLOv8-RDE. *International Journal of Distributed Sensor Networks*, **2025**, Article ID: 6543345. <https://doi.org/10.1155/dsn/6543345>
- [24] Zhang, D. and Ma, H. (2024) YOLODCC: Improved Yolov8 Combined with Dynamic Confidence Compensation for Lightweight Moving Object Detection. *IET Image Processing*, **18**, 3699-3715. <https://doi.org/10.1049/ipr2.13207>
- [25] 罗逸豪. 基于深度学习的声呐图像目标检测系统[J]. *数字海洋与水下攻防*, 2023, 6(4): 423-428.
- [26] 朱俊, 封磊. 基于声呐图像的鱼群识别与计数方法[J]. *南京理工大学学报*, 2023, 47(6): 782-789.
- [27] Echopype-echopype v0.7.1 Manual. <https://github.com/OSOceanAcoustics/echopype>