喻聪. 基于 YOLOv8n 改进的水下生物检测算法[J]. 智能计算机与应用,2026,16(7):107-114. DOI:10.20169/j.issn.2095-2163.24111105

基于 YOLOv8n 改进的水下生物检测算法

喻 聪
(南昌市卫生学校 信息中心, 南昌 330006)

摘 要: 针对水下环境复杂多变,如光线不足、物体体积小、生物密集分布等问题使得水下图像质量受限,由此带来了水下生物目标检测挑战。为了解决这些问题,提出了一种基于 YOLOv8n 改进的水下生物检测算法 YOLOv8n-MSC。在 YOLOv8n 中,添加了多头自注意力模块(MHSA)与空间深度转换卷积模块(SPD-Conv)以提升模型性能。MHSA 模块通过多个注意力头来捕捉输入数据中不同部分之间的关系,增强模型的注意力机制,从而提升模型的表示能力。SPD-Conv 模块通过将图像空间信息转换为深度信息来提高卷积神经网络的特征提取能力,从而可以捕获更丰富的特征信息。实验结果表明,与基线模型 YOLOv8n 相比,mAP@50 提高了 2.2%,验证了该方法的可行性和有效性。

关键词: YOLOv8n; 水下生物检测; MHSA; SPD-Conv
中图分类号: TP391.4 **文献标志码**: A **文章编号**: 2095-2163(2026)07-0107-08

Improved underwater biological detection algorithm based on YOLOv8n

YU Cong
(Information Center, Nanchang Health School, Nanchang 330006, China)

Abstract: In view of the complex and changeable underwater environment, such as insufficient light, small size of objects and dense distribution of organisms, the quality of underwater images is limited. In order to solve these problems, an improved underwater biological detection algorithm YOLOv8n-MSC based on YOLOv8n is proposed. In YOLOv8n, a Multi-Head Self-Attention module (MHSA) and a SPatial Depth Conversion convolution module (SPD-Conv) are added to improve model performance. The MHSA module captures the relationship between different parts of the input data through multiple attention heads, enhances the attention mechanism of the model, and improves the representation ability of the model. The SPD-Conv module improves the feature extraction ability of the convolutional neural network by converting the spatial information of the image into depth information, so that it can capture richer feature information. The experimental results demonstrates that mAP@50 is improved by 2.2% compared with the baseline model YOLOv8n, which verifies the feasibility and effectiveness of the method.

Key words: YOLOv8n; underwater biological detection; MHSA; SPD-Conv

0 引 言

近年来,随着海洋经济与水产养殖领域的蓬勃发展,水下目标检测技术[1]在海洋科学研究中逐渐占据重要地位。水下图像目标检测在保护海洋生态系统及其脆弱环节方面发挥着关键作用。随着传感器技术[2]和图像处理手段的日益精进,水下计算机视觉领域吸引了越来越多的关注。凭借其独特的优势,水下图像处理已成为水下机器人视觉系统[3]、海洋生物跟踪[4]与监测[5]等多个领域的关键技术。

然而,水下环境复杂多变,光线不足、物体体积小、生物密集分布等问题使得水下图像质量受限,进而对目标检测任务构成了严峻挑战。相较于地面目标检测[6],水下检测面临更为复杂的环境条件,图像采集难度大,且高质量的水下目标数据集相对稀缺。此外,由于水下拍摄距离较远,图像中小目标占比显著提升,这对水下小目标检测技术的准确性提出了更高要求。尽管已有部分目标检测算法应用于水下环境,但其性能仍有待提升,尤其是检测精度方面。

当前,基于深度学习的水下目标检测算法[7]主要分为 2 类:两阶段算法和单阶段算法。其中,两阶段算法首先确定潜在的目标候选区域,随后进行分类和位置调整。而单阶段算法则直接通过端到端网络预测所有目标的位置,更适合实时检测任务。以 Faster R-

作者简介: 喻 聪(1996—),男,硕士。主要研究方向:计算机视觉研究。Email:yucong12101210@qq.com。

收稿日期: 2024-11-11

CNN[8-9]为代表的两阶段算法已成功应用于水下鱼类检测[10]和底栖生物识别[11]。为了提升检测精度,Li 等学者[12]提出了基于 YOLOv5 的单阶段检测算法,

结合了 SE 注意力机制[13]和加权双向特征金字塔网络(BiFPN)[14],在平均精度均值(mAP)和 $F1$ 分数上显著优于 Faster R-CNN。葛锡云等学者[15]则基于 YOLOv5 提出了多尺度水下目标检测算法,通过稳定的底层特征提取和自适应特征融合机制,增强了算法的多尺度检测能力和鲁棒性。梁秀满等学者[16]则基于 YOLOv7 设计了多信息流融合注意力机制,解决了卷积过程中特征丢失的问题,并结合暗通道先验增强了水下图像,进一步提升了检测性能。此外,通过引入 C2f 结构和无锚框设计,YOLOv8 在多种检测场景下表现出色。然而,研究者也发现了 YOLOv8 在复杂场景下小目标检测方面的局限性[17-21]。

尽管上述研究取得了显著成果,但水下目标检测仍面临精度不足等问题,亟待进一步优化。因此,本文提出了一种改进的水下生物检测算法—YOLOv8n-MSC。该算法基于 YOLOv8n 模型,旨在解决复杂水下海洋生物检测任务中的挑战。具体改进措施包括:

(1)在模型的主干网络中引入了多头自注意力(Multi-Head Self-Attention,MHSA)模块[22],旨在增强模型对水下生物上下文信息的深度理解能力。

(2)同时,在主干网络中嵌入了 SPD-Conv 模块[23],以进一步捕获图像的细粒度信息,从而提高检测精度。该模块能够提取图像中的丰富空间信息和深度信息,在处理不同尺度和背景的目标时展现出更强的泛化能力。

1 YOLOv8 算法简介

YOLOv8[24]作为当前领先的目标检测模型,凭借其卓越的检测速度和精度,在目标检测领域广受好评。该模型结构(见图 1)由四大核心组件构成:输入处理模块(Input)、特征提取主干(Backbone)[25]、特征融合颈部(Neck)[26]以及目标检测头部(Head)。在输入端,集成了马赛克数据增强、图像预处理及自适应锚框计算等功能,以优化输入数据。主干网络则采用了高效的 CSPDarknet 架构,融合了 Conv 卷积、C2f 模块及 SPPF 等组件[27],精准提取图像特征。

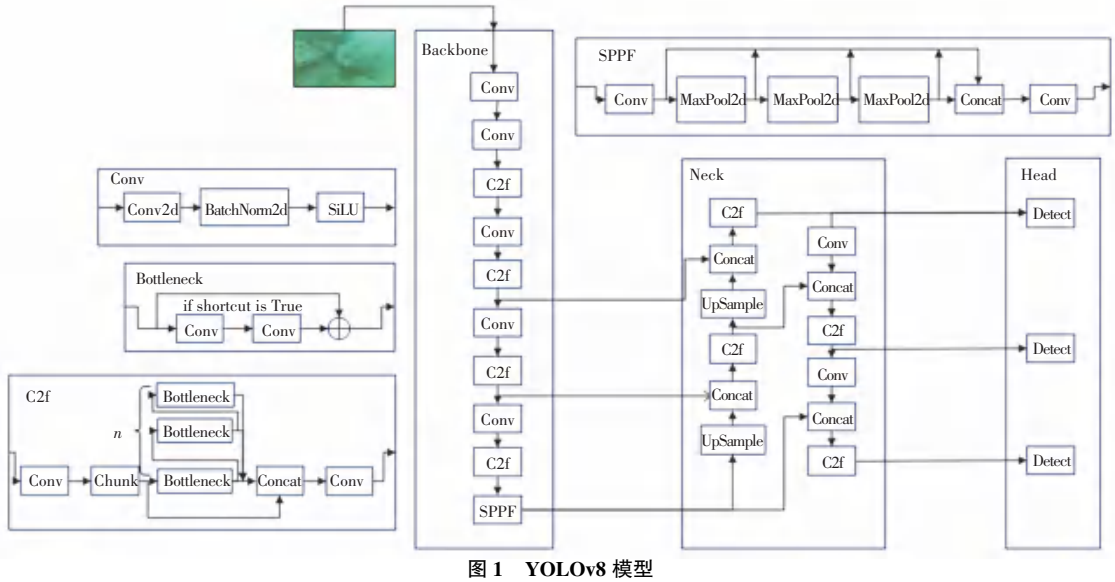


图 1 YOLOv8 模型

Fig. 1 Model of YOLOv8

颈部网络部分,由特征金字塔网络(Feature Pyramid Networks, FPN)[28]与路径聚合网络(Path Aggregation Network, PAN)[29]协同工作,实现了高层与底层特征的深度融合,增强了特征表达能力。检测头部分,则创新性地设计了 3 个尺寸各异的解耦头,通过分离检测与分类任务,显著提升了检测的精确度。同时,模型还摒弃了传统的 Anchor-Based 方法[30],转而采用 Anchor-Free 策略[31],以实现更精确的目标定位与分类。

在损失函数设计上,YOLOv8 引入了 Task-Aligned Assigner 正负样本匹配策略,并结合 DFL(Distribution Focal Loss)损失函数,进一步优化了检测性能。根据模型的复杂度和计算需求,YOLOv8 提供了从 YOLOv8n 到 YOLOv8x 五个不同版本的配置。鉴于 YOLOv8n 版本在保持高性能的同时,具有最小的参数量,且完全满足本文的计算需求,因此被

选定为本次研究的基准模型。

2　YOLOv8n-MSC 算法简介

针对水下生物检测中因特征不足导致的漏检与误检问题，本文对 YOLOv8n 模型进行了优化，创新性地引入了 CNN 模块 SPD-Conv[22]。该模块旨在替代主干网络中的跨步卷积层和池化层，通过摒弃传统步长和池化操作来实现特征图的降采样。这种方法不仅有效减少了计算量，还保留了更多具有区分性的特征信息，展现出显著优势。此外，SPD-Conv 作为一种通用且统一的方法，能够轻松融入任何 CNN 架构中，以相同方式替换跨步卷积和池化层。

此外，本文在主干网络的 SPPF 层与 C2f 层中融入了 MHSA（Multi-Head Self-Attention，多头自注意力）[32]机制。这一举措使得模型能够学习到更为全面且丰富的特征表示，显著增强了模型对水下生物图像中复杂空间关系和上下文信息的理解能力。通过这种方式，模型能够更深入地捕捉水下生物的语义信息和特征表示，进而提升目标检测的鲁棒性。优化后的 YOLOv8n-MSC 模型结构如图 2 所示。

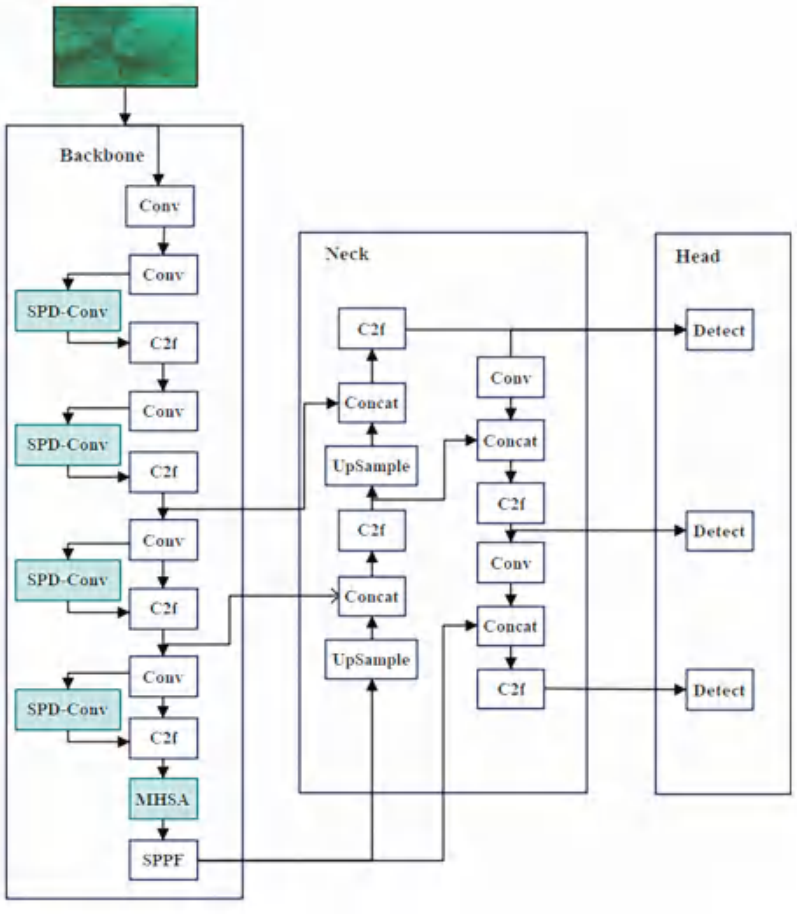


图 2　YOLOv8n-MSC 模型

Fig. 2　Model of YOLOv8n-MSC

2.1　SPD-Conv 模块

SPD-Conv（SPace-to-Depth Convolution）[31]由 2 层构成，如图 3 所示，分别是空间到深度层和非跨步卷积层，其中 c 表示特征图的通道数，s 表示特征图的宽和高。

（1）空间到深度层：SPD 组件对 CNN 的特征图进行下采样。对于任意的特征图 X，将其切分为子特征图序列：

$$\begin{cases} f_{0,0} = X[0:S:r,0:S:r], f_{1,0} = X[1:S:r,0:S:r],\cdots, \\ \qquad f_{r-1,0} = X[r-1:S:r,0:S:r] \\ f_{0,1} = X[0:S:r,1:S:r], f_{1,1} = X[1:S:r,1:S:r],\cdots, \\ \qquad f_{r-1,1} = X[r-1:S:r,1:S:r] \\ \cdots \\ f_{0,r-1} = X[0:S:r,r-1:S:r] \\ f_{1,r-1} = X[1:S:r,r-1:S:r],\cdots, \\ f_{r-1,r-1} = X[r-1:S:r,r-1:S:r],\cdots, \end{cases} \tag{1}$$

一般来说,对于任意的特征映射 X, 子映射 $f_{x,y}$ 由所有特征映射组成特征图 $X(i,j)$,$i+x$ 和 $j+x$ 可以被比例整除。每个子图按照比例因子 r 向下采样 X。例如,当比例因子为 2 时,子映射为 $f_{0,0}$, $f_{0,1}$, $f_{1,0}$, $f_{1,1}$,其形状都是 $\frac{S}{2}\times\frac{S}{2}\times C_1$ 的。再将这些子特征映射按照通道维度连接,就得到了一个新的特征映射,并具有以比例因子减少的空间维度和以比例因子的平方增加的通道维度。

经过空间到深度层,特征映射 X 的空间维度减少,通道维度增加,得到新的特征映射 X', 这个输出值代表了原始特征图区域的概括信息。

(2)非跨步卷积层:即步长为 1 的卷积层。SPD 在空间到深度层之后添加了一个带有 C_2 过滤器的非跨步卷积层,其中 $C_2 < r^2 < C_1$,将上层 $X'\left(\frac{S}{r},\frac{S}{r},r^2C_1\right)$ 继续进行变换为 $X''\left(\frac{S}{r},\frac{S}{r},r^2C_1\right)$。使用非跨步卷积是为了尽可能地保留所有判别特征的信息。通过多次 SPD 组件,可以逐渐减少特征图的空间规模,并提取出更加抽象、更丰富的特征信息,有助于提高模型的性能和泛化能力。

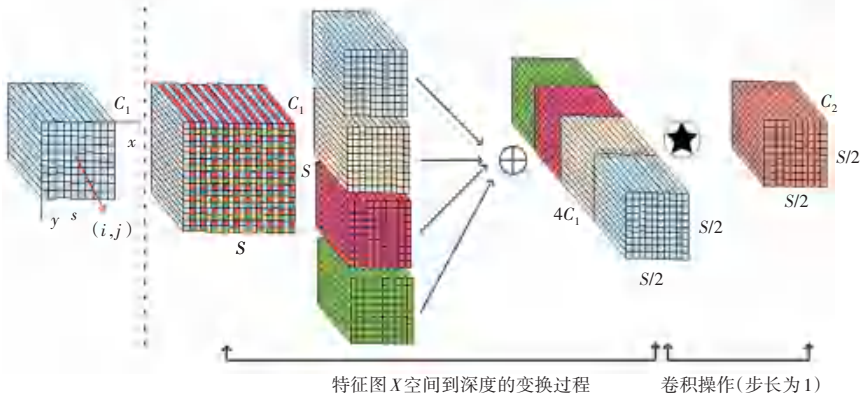


图 3 SPD-Conv 变换过程

Fig. 3 SPD-Conv transformation process

2.2 多头自注意力机制

为了增强输出端检测头的性能,并更专注于待检测的目标物体,在输出端的普通卷积层之间引入了瓶颈转换器 BoT 模块中的多头自注意力机制(Multi-Head Self-Attention, MHSA)[22],具体结构如图 4 所示。在此机制中,$H=4$ 表示 MHSA 拥有 4 个注意力头。Rh 和 Rw 分别表示在高度和宽度方向上的相对位置编码,而 q、k、r 则分别表示查询(query)、键(key)和位置(position)编码。“⊕”和“⊗”分别表示元素相加与矩阵乘法操作,而 1×1 则表示点卷积处理。

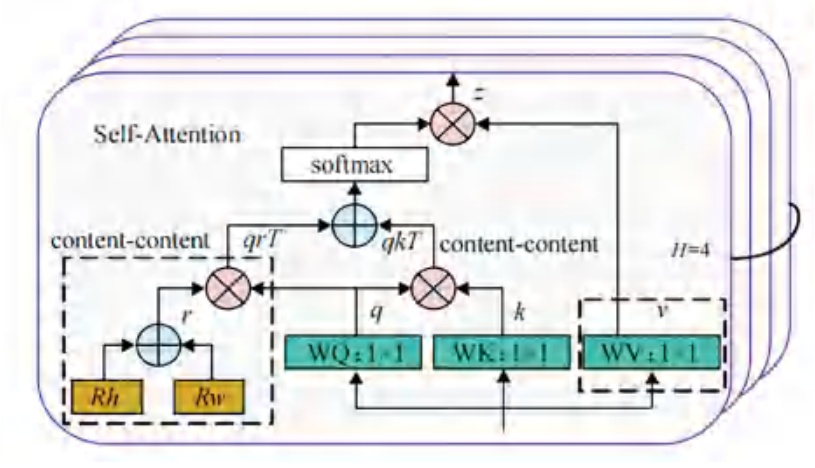


图 4 多头自注意力机制

Fig. 4 Multi-Headed Self-Attention mechanism

MHSA 首先通过全连接层实现特征的线性升维,随后对 4 个注意力头并行执行自注意力计算。每个自注意力头都会独立地学习多组权重,并与特征权重进行矩阵乘法,从而生成相应的特征输出。

最后，将这些来自不同注意力头的输出进行拼接，以实现多角度的特征融合。

此外，MHSA 还具有在二维特征图上实现全局注意力的能力。能够将网络在特征提取和融合两个阶段所生成的特征进行有机关联，并整合全局注意力下的信息，从而进一步提升检测效果。通过这种设计，将致改力于实现更精确、更鲁棒的水下生物目标检测。

3　实验结果与分析

3.1　实验环境与数据

本实验平台为 Windows 10 系统，CPU Intel(R) Core(TM) i5-10400F CPU @ 2.90 GHz，GPU 为 NVIDIA GeForce RTX 2060 (12 G)，运行内存为 32 GB，Python 版本为 3.8，深度学习框架为 Pytorch 1.13.0，CUDA 版本为 11.7。实验过程均采用相同的超参数，具体参数见表 1。

表 1　实验训练参数

Table 1　Experimental training parameters

参数	设置
训练轮次(Epoch)	200
优化器(Optimizer)	SGD[32]
动量(Momentum)	0.937
学习率(Learning Rate)	0.01
权重衰减(Weight Decay)	0.000 5
图像大小(Image Size)	640×640
批量大小(Batch Size)	32

在本次实验中，采用了水下机器人目标抓取大赛(URPC)中的水下数据集，该数据集由自主无人水下机器人(AUV)[33]在大连獐子岛海洋牧场采集的海底图像构成，准确地还原了真实的海洋环境。该数据集包含了 6 760 张图像，并细分为海星、海参、海胆和扇贝这 4 个类别。数据集的样本实例如图 5 所示。为了适配 YOLO 模型，对数据进行了预处理操作，具体是将原有的 VOC 格式转换为 YOLO 所需的格式。此外，还按照 8∶1∶1 的比例，将数据集划分为训练集、验证集、测试集，以确保模型训练和评估的有效性。

3.2　评价指标

本文采用以下指标来全面评估模型性能：精度 P (Precision)、召回率 R (Recall)、平均精度均值(mean Average Precision，mAP)。这些指标的具体计算公式如下：

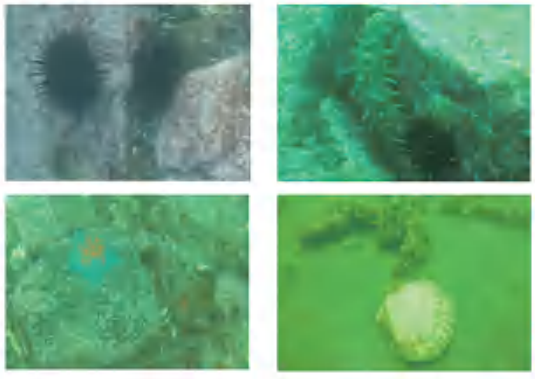


图 5　数据集实例

Fig. 5　Examples of dataset

$$P = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}} \tag{2}$$

$$R = \frac{\mathrm{TP}}{\mathrm{FP} + \mathrm{FN}} \tag{3}$$

$$\mathrm{AP} = \int_0^1 P(R)\,\mathrm{d}R \tag{4}$$

$$\mathrm{mAP} = \frac{1}{k}\sum_{i=1}^{k} \mathrm{AP}_i \tag{5}$$

其中，TP(真正例)表示被模型正确预测为正样本的数量；FP(假正例)表示被错误预测为正样本的其他类别样本数量；FN(假负例)表示被错误预测为其他类别的正样本数量。精度 P (Precision)反映了模型区分负样本的能力，P 值越高，模型预测为正样本的准确性就越高。召回率 R (Recall)则体现了模型识别正样本的能力，R 值越高，正样本被模型正确识别的比例就越大。在式(5)中，k 表示实验中需要检测的类别总数，本实验中为 4 类。在计算平均精度均值(mAP)时，设定了交并比(IoU)的阈值为 0.5，即当预测框与真实框的交并比大于 0.5 时，认为模型的预测是正确的。

3.3　不同注意力机制效果对比

为了深入探究本次改进中引入的 MHSA 注意力机制对网络性能的具体提升效果，设计了一系列对比实验。以 YOLOv8n 作为基础模型(基线)，分别向基线网络中融入了 EMA[34]、SE[13]、ECA[35]、SimAM[36]以及 MHSH[22]等多种注意力机制。通过对比这些实验的结果，可以清晰地观察到 MHSA 注意力机制对改进网络性能的显著影响，具体实验结果请参见表 2。根据表 2 的数据分析结果，可以得出以下结论：

与 YOLOv8n 基线模型相比，引入 EMA 注意力机制后，模型的 mAP@50 性能下降了 0.1%，显示出

EMA 在该数据集上的应用效果并不理想。而当尝试添加 SE 和 ECA 这 2 种通道注意力机制时,虽然 mAP@ 50 有所提升,但提升幅度较为有限,仅为 0.3%。这主要是因为,在提取通道注意力的过程中,SE 和 ECA 方法会将空间维度信息进行压缩,从而容易导致空间位置信息的损失。

表 2 不同注意力机制对比实验

Table 2 Comparative experiments of different attention mechanisms %

Models	AP				Precision	Recall	mAP@ 50
	echinus	holothurian	starfish	scallop			
YOLOv8n	90.9	73.2	89.6	83.9	83.7	77.0	84.4
+EMA[34]	90.6	73.1	89.0	84.5	84.9	76.8	84.3
+SE[13]	90.8	74.5	89.0	84.8	85.0	76.5	84.7
+ECA[35]	91.4	73.0	89.7	84.8	84.9	76.7	84.7
+SimAM[36]	91.7	72.8	89.3	84.9	84.7	76.8	84.8
+MHSH[22]	**91.8**	**76.8**	**89.8**	**87.0**	**85.1**	**80.5**	**86.4**

相比之下,SimAM 作为一种轻量级且无参数的注意力机制,在嵌入网络后能够为模型带来 0.4% 的 mAP@ 50 提升。而 MHSH 注意力机制则展现出更为出色的性能,能够完整地保留每个通道的信息,有效地避免了降维操作所带来的问题,进而显著提升了目标检测的性能。

特别值得一提的是,当在 YOLOv8n 的 Backbone 中融合了 MHSH 注意力机制后,模型在各个类别上的检测精度均实现了显著提升:海胆类的检测精度上涨了 0.9%,海参类上涨了 4.6%,海星类上涨了 0.2%,扇贝类上涨了 3.1%。同时,模型在所有类别上的 mAP@ 50 也实现了 2%的增长。这些数据充分证明了 MHSH 注意力机制在提升模型检测精度方面的有效性。

3.4 消融实验

为了全面验证本文提出的算法改进模块在提升水下生物检测模型性能方面的有效性,设计了一系列消融实验。这些实验采用基础 YOLOv8n 模型作为对比基准,并逐步整合各个改进模块,以分析其对模型整体性能的贡献。实验结果已汇总见表 3。

表 3 消融实验结果

Table 3 Ablation experimental results

Models	MHSA	SPD-Conv	AP/%				Precision/%	Recall/%	mAP@ 50/%
			echinus	holothurian	starfish	scallop			
1			90.9	73.2	89.6	83.9	83.7	77.0	84.4
2	√		91.8	76.8	89.8	87.0	85.1	80.5	86.4
3		√	91.6	77.3	89.7	85.4	85.1	78.5	86.1
4	√	√	**92.0**	**77.7**	**90.0**	**86.5**	**85.4**	79.9	**86.6**

表 3 中的“√”指示 YOLOv8n 网络已采纳了特定的改进措施。基于这些实验结果,可以总结出以下结论:

(1)模型 2 对骨干网络进行了优化,采纳了 MHSA,从而实现了精度的明显提升,mAP@ 50 指标增长了 2 个百分点。

(2)在模型 2 的基础上,模型 3 又增加了 SPD-Conv,相较于原始的 YOLOv8n,mAP@ 50 指标提高了 1.7 个百分点。

(3)模型 4 则在模型 3 的基础上再次引入了 MHSA,与 YOLOv8n 相比,mAP@ 50 指标增加了 2.2 个百分点。

综上所述,在结合 MHSA 与 SPD-Conv 后,YOLOv8n 网络在水下生物检测任务中展现出了出色的性能,这验证了各改进措施的有效性。

3.5 对比实验

为评估改进算法相较于其他算法的优势,实验在相同数据集上对比了改进算法与当前主流的目标检测算法,具体实验结果已列于表 4 中。

与传统 SSD 和 Faster R-CNN 网络相比,YOLOv8n-MSC 的 mAP@ 50 值实现了超过 30%的显著提升。与 YOLOv8n 模型相比,YOLOv8n-MSC 在海

胆类上的 mAP@50 提高了 1.1%，在海参类上提升了 3.9%，在海星类上增长了 0.4%，在扇贝类上则增加了 2.6%。同时，在所有检测类别上，YOLOv8n-MSC 的 mAP@50 综合提升了 2.2%。综上所述，YOLOv8n-MSC 在近年来的水下生物检测任务中展现出了卓越的性能，这充分验证了该算法的有效性和优越性。

表 4　不同算法对比结果

Table 4　Comparison results of different algorithms　　%

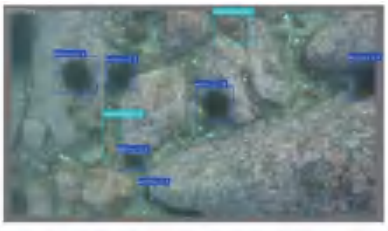
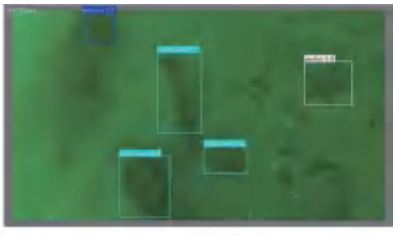
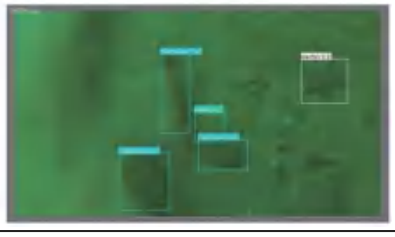
Models	AP				mAP@50
	echinus	holothurian	starfish	scallop	
Faster R-CNN	75.8	24.6	47.3	71.6	54.8
SSD	77.0	38.2	69.8	39.3	56.0
YOLOv5	91.0	74.6	89.1	85.0	84.9
YOLOv6	90.3	73.8	88.7	83.1	84.0
YOLOv8n	90.9	73.2	89.6	83.9	84.4
YOLOv10n	91.1	66.7	88.6	82.9	82.3
YOLOv8n-MSC	**92.0**	**77.7**	**90.0**	**86.5**	**86.6**

3.6　可视化分析

为了更直观地体现改进算法的检测性能，分别采用改进后的网络与 YOLOv8n 网络在数据集上进行了检测实验，实验结果的可视化展示见表 5。

表 5　实验效果图对比

Table 5　Comparison of experimental results

模型	数据集(URPC2020)	
	第 1 组	第 2 组
YOLOv8n		
YOLOv8n-MSC		

分析上述结果后，可以发现 YOLOv8n 在进行水下生物检测时表现欠佳，存在误检情况，错误地将水下生物类似背景识别为海参与海胆。相比之下，YOLOv8n-MSC 则未将类似背景误判为海参与海胆，且能够准确识别出相应的水下生物。这表明 YOLOv8n-MSC 的性能相较于 YOLOv8n 更为优越，在水下生物检测方面展现出了更高的准确性。

4　结束语

本文提出了一种水下生物检测算法 YOLOv8n-MSC，该算法基于 YOLOv8n 进行了改进，并通过一系列实验验证了其有效性。实验结果显示，YOLOv8n-MSC 在 URPC2020 水下数据集上显著提升了检测性能，特别是在多尺度目标的检测精度方面表现出色。

在算法设计中，本文引入了新的 CNN 模块 SPD-Conv，该模块能够在减少计算量的同时，保留更多的区分性特征信息，从而展现出明显的优势。此外，鉴于水下目标检测场景中显著的尺度变化和大量小尺度目标的存在，本文还引入了 MHSA 注意力机制。这一机制使模型能够学习到更全面的特征表示，增强了模型对水下生物图像中复杂的空间关系和上下文信息的理解能力。这不仅提升了水下生物的语义信息和特征表示能力，还提高了目标检测的鲁棒性。

然而,当前模型在计算复杂度方面仍面临挑战。未来的研究可以聚焦于优化网络模型,以降低计算复杂度并实现轻量化,从而更好地适应实际应用中的水下目标检测任务。通过持续的改进和优化,YOLOv8n-MSC 有望在更多实际场景中发挥更大的作用。

参考文献

[1] 黄建兴. 基于光学图像的水下小目标检测技术研究[D]. 南京:东南大学, 2023.

[2] 杜鹏. 无线拖网监测系统水下传感器节点关键技术研究与实现[D]. 济南:齐鲁工业大学, 2024.

[3] 琚玲, 周星群, 胡志强, 等. 基于合成数据的水下机器人视觉定位方法[J]. 信息与控制, 2023, 52(2):129-141.

[4] 付晓雪. 基于深度学习的海洋底栖生物定量识别算法研究[D]. 青岛:青岛科技大学, 2022.

[5] 张俊恒, 司亚超, 郑孟然. CO_YOLO:基于改进 YOLOv5 的海洋生物目标检测[J]. 长江信息通信, 2024, 37(10):69-71.

[6] 刘晶, 祁文哲, 吕德芳, 等. 改进 YOLOv5 的轨面伤损检测方法[J]. 智能计算机与应用, 2024, 14(6):71-78.

[7] 陈晓, 杨琪, 姚海洋, 等. 基于细粒化特征感知的水下目标检测算法[J]. 陕西科技大学学报, 2024, 42(4):177-183.

[8] 齐兴斌, 赵丽, 耿海军, 等. 基于改进 Faster R-CNN 的域自适应红外目标检测方法[J]. 计算机工程与设计, 2024, 45(10):2994-3001.

[9] 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, 2016, 39(6): 1137-1149.

[10] 张翔宇, 朱立军. 改进的 Faster R-CNN 海洋鱼类检测模型[J]. 沈阳化工大学学报, 2022, 36(6):562-568.

[11] 吴睿, 毕晓君. 基于改进 YOLOv5 算法的珊瑚礁底栖生物识别方法[J]. 哈尔滨工程大学学报, 2022, 43(4):580-586.

[12] Li W, Zhang Z, Jin B, et al. A real-time fish target detection algorithm based on improved YOLOV5[J]. Journal of Marine Science and Engineering, 2023, 11(3): 572.

[13] Hu J, Shen L, Sun G. Squeeze-and-excitation networks[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ:IEEE, 2018: 7132-7141.

[14] Tan M, Pang R, Le Q V. Efficientdet: Scalable and efficient object detection[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2020: 10781-10790.

[15] 葛锡云, 张崇丙, 李晓伟, 等. 水下复杂环境高效鲁棒目标检测方法[J]. 舰船科学技术, 2023, 45(22):148-154.

[16] 梁秀满, 李然, 于海峰, 等. 改进 YOLOv7 的水下目标检测算法[J]. 计算机工程与应用, 2024, 60(6):89-99.

[17] Zheng X, Zou J, Du S, et al. Small target detection in refractive panorama surveillance based on improved YOLOv8[J]. Sensors, 2024, 24(3): 819.

[18] Shen L, Lang B, Song Z. DS-YOLOv8-based object detection method for remote sensing images[J]. IEEE Access, 2023, 11: 125122-125137.

[19] Wang X, Gao H, Jia Z, et al. BL-YOLOv8: An improved road defect detection model based on YOLOv8[J]. Sensors, 2023, 23 (20): 8361.

[20] Ma M, Pang H. SP-YOLOv8s: An improved YOLOv8s model for remote sensing image tiny objectdetection [J]. Applied Sciences, 2023, 13(14): 8161.

[21] Li Y, Fan Q, Huang H, et al. A modified YOLOv8 detection network for UAV aerial imagerecognition[J]. Drones, 2023, 7 (5): 304.

[22] Tan H, Liu X, Yin B, et al. MHSA-Net: Multihead self-attention network for occluded person re-identification[J]. IEEE Transactions on Neural Networks and Learning Systems, 2022, 34 (11): 8210-8224.

[23] Wang G, Chen Y, An P, et al. UAV-YOLOv8: A small-object-detection model based on improved YOLOv8 for UAV aerial photographyscenarios[J]. Sensors, 2023, 23(16): 7190.

[24] Gao S H, Cheng M M, Zhao K, et al. Res2net: A new multi-scale backbone architecture [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2019, 43(2): 652-662.

[25] Yuan Z, Shao P, Li J, et al. YOLOv8-ACU: improved YOLOv8-pose for facialacupoint detection [J]. Frontiers in Neurorobotics, 2024, 18: 1355857.

[26] Wang W, Meng Y, Li S, et al. HV-YOLOv8 by HDPconv: Better lightweight detectors for small object detection[J]. Image and Vision Computing, 2024, 147: 105052.

[27] Lin T Y, Dollár P, Girshick R, et al. Feature pyramid networks for object detection[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ:IEEE, 2017: 2117-2125.

[28] Liu S, Qi L, Qin H, et al. Path aggregation network for instance segmentation [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ:IEEE, 2018: 8759-8768.

[29] Liu Y C, Ma C Y, Kira Z. Unbiased teacher v2: Semi-supervised object detection for anchor-free and anchor-based detectors [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway, NJ:IEEE, 2022: 9819-9828.

[30] Cheng G, Wang J, Li K, et al. Anchor-free oriented proposal generator for object detection [J]. IEEE Transactions on Geoscience and Remote Sensing, 2022, 60: 5625411.

[31] Yang Z, Wu Q, Zhang F, et al. A new semantic segmentation method for remote sensing images integrating coordinate attention and SPD-Conv[J]. Symmetry, 2023, 15(5): 1037.

[32] Chandra K, Xie A, Ragan-Kelley J, et al. Gradient descent: The ultimate optimizer[J]. Advances in Neural Information Processing Systems, 2022, 35: 8214-8225.

[33] Sahoo A, Dwivedy S K, Robi P S. Advancements in the field of autonomous underwater vehicle[J]. Ocean Engineering, 2019, 181: 145-160.

[34] 江结林, 史明月, 杨海东, 等. 多头注意力机制的全频图像去噪算法[J]. 计算机工程与应用, 2024, 60(16):236-247.

[35] Wang Q, Wu B, Zhu P, et al. ECA-Net: Efficient channel attention for deep convolutional neural networks[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway, NJ:IEEE, 2020: 11534-11542.

[36] Zhang Y, Sun Z. An advanced YOLOv5s approach for vehicle detection integrating swin Transformer and SimAM in dense traffic surveillance[J]. Journal of Industrial Intelligence, 2024, 2(1): 31-41.