

## 基于双目视觉的海参体积测量方法

镇 帅<sup>1,2,3</sup>, 林远山<sup>1,2,3</sup>, 盛亦凡<sup>1</sup>, 洪胜呈<sup>1</sup>, 王文良<sup>1</sup>,  
陈启俊<sup>4</sup>, 杨志庆<sup>5</sup>, 李智军<sup>1\*</sup>

(1. 大连海洋大学信息工程学院, 大连 116023; 2. 大连市智慧渔业重点实验室, 大连 116023; 3. 设施渔业教育部重点实验室(大连海洋大学), 大连 116023; 4. 大连鑫玉龙海洋生物种业科技股份有限公司, 大连 116007;  
5. 贵港市荣创木业有限公司, 贵港 537000)

**摘要:** 海洋珍品的体积测量可以为水产养殖过程中海洋珍品的生长状态观测以及价值评估提供科学的数据支撑。针对目前海参养殖过程中体积测量耗时耗力、效率低下、精度参差不齐等问题, 该研究提出了一种基于双目视觉的海参体积测量方法, 以真实养殖场景下的原位海参作为研究对象, 包含海参目标检测、海参实体分割、海参体积估计 3 个模块。其中, 目标检测模块针对水下光线多变、环境复杂的问题, 设计了融合多头自注意力机制的 YOLOv8 目标检测器, 提高水下环境下的检测精度; 实例分割模块构建了加入 adapter 机制的 SAM (segment anything model) 海参分割模型, 以此获取有效精准的掩码信息; 体积估计模块将掩码信息映射到三维空间获取海参的点云数据, 通过泊松表面重建和体素化处理的方法克服点云数据稀疏、噪声干扰的障碍, 重构海参的三维表征。结果表明, 改进后的检测模型展现了综合性能的最佳表现, 在精度和帧率方面分别达到了 92.5% 和 31 帧/s, 相较于 Faster RCNN、Cascade RCNN、YOLOv7、YOLOv8 等算法, 均保持一定的领先优势; 泊松表面重建的重建效果明显优于 Alpha shapes、Ball pivoting 算法, 更贴近于真实海参的表征信息, 可以达到较好的体积测量效果, 较于其他海参体积测量算法, 该研究提出的算法在不同深度下均有更好的表现, 在距离为 100 cm 处表现最佳, 测量精度达到了 94%, 比最小包围框测量法和 Ball pivoting 测量法分别高出 22 和 14 个百分点。可见, 所提方法能较为准确地测量出海参的体积, 基本能够满足养殖户的评估需求, 可为海参的科学养殖提供更为便捷精准的数据支持。

**关键词:** 水产养殖; 双目视觉; 体积测量; 泊松表面重建; 深度学习; 海参

doi: 10.11975/j.issn.1002-6819.202407228

中图分类号: TP391.41 文献标志码: A 文章编号: 1002-6819(2024)-21-0165-10

镇帅, 林远山, 盛亦凡, 等. 基于双目视觉的海参体积测量方法[J]. 农业工程学报, 2024, 40(21): 165-174. doi: 10.11975/j.issn.1002-6819.202407228 <http://www.tcsae.org>

ZHEN Shuai, LIN Yuanshan, SHENG Yifan, et al. Measurement of sea cucumber volume using binocular vision[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2024, 40(21): 165-174. (in Chinese with English abstract) doi: 10.11975/j.issn.1002-6819.202407228 <http://www.tcsae.org>

### 0 引言

随着公众健康意识的普遍提高以及生活水平的不断提升, 海参作为一种高营养价值的海洋珍品, 在当前市场中获得了广泛的关注和重视。在海洋珍品养殖过程中, 养殖户需要各类重要的信息来评估海洋珍品的生长状态以及它们对养殖环境的适应性<sup>[1-3]</sup>。因此, 进行阶段性的数据观测具有极其重要的意义。然而, 传统的观测方法存在诸多局限性。以海参为例, 养殖户通常需要派遣潜

水员潜入水下, 近距离观测和接触海参, 以此来近似估计海参的尺寸、体积等信息。首先, 这种方法存在明显的劳动密集和时间成本高的问题, 测量精度得不到保证, 效率低下。其次, 人为的干预会对自然生长环境下的海参造成一定程度的干扰和损害。因此, 为了提高海参养殖过程中生长状态观测的效率和精确度, 有必要探索和引入更为先进便捷的测量技术。机器视觉的不断发展为农业观测评估任务带来了全新的可能性<sup>[4-5]</sup>。相比传统的人工测量方法, 机器视觉测量具有更高的效率和准确性。通过使用高分辨率的摄像头和先进的图像处理算法, 即可实现非接触式的测量, 无需人工干预, 避免了传统方法中可能存在的测量误差和干扰, 从而为农业评估提供了更为可靠和及时的生产数据支持。

国内外学者纷纷进行了机器视觉辅助农业生产的探索<sup>[6-9]</sup>, 目前在二维图像方面的研究已经卓有成效。MIRBOD 等<sup>[10]</sup>开发了一种苹果尺寸测量的机器视觉系统, 通过训练两个深度神经网络模型来检测候选苹果的尺寸, 然后外推被遮挡的水果区域以改进尺寸估计测量, 有效测量苹果的外圈直径。不过由于图像像素和镜头畸变的

收稿日期: 2024-07-28 修订日期: 2024-09-14

基金项目: 广西重点研发计划 (桂科 AB23075150); 辽宁省属本科高校基本科研业务费专项资金资助 (2024JBZDZ004); 辽宁省重点研发计划 (2023JH26/10200015); 2023 中央财政对辽宁渔业补助项目; 辽宁省自然基金资助计划 (2020-KF-12-09); 辽宁省教育厅基本科研项目 (LJKZ0730, QL202016); 设施渔业教育部重点实验室开放课题 (202219); 辽宁省应用基础计划项目 (2022JH2/101300187)

作者简介: 镇帅, 研究方向为目标检测、目标分割、点云处理。

Email: 1817340986@qq.com

\*通信作者: 李智军, 研究员, 研究方向为机器视觉、机器人。

Email: lizhijun@dlou.edu.cn

影响, 位于靠近图像边缘水果的直径测量仍存在较大的不准确性。董鹏等<sup>[1]</sup>利用水下双目标定方法, 消除水下光线折射造成的图像失真, 并在此基础上, 利用凸包与旋转卡壳算法, 提出了一种水下海参自动检测与尺寸测量方法, 实现了水下环境中的海参长度的测量。但这种依赖于特征点选取的测量方法仍停留在二维图像层面, 不仅存在一定的测量误差, 更重要的是, 无法进行海参体积的测量, 而单一的长度信息无法满足养殖户对海参各项生长数据的评估需求。

尽管图像的二维特征可以提供一部分有价值的数据, 但生物体本质上是三维的, 体积和身体形态等表型信息很难在二维成像的过程中被捕捉到。随着计算机视觉技术的不断发展, 三维重建的相关方法越来越多地应用于农业<sup>[2-15]</sup>。GEN'E-MOLA 等<sup>[16]</sup>提出了一种基于运动结构和多视角立体的点云生成技术, 借助点云信息进行果实大小的估算。分别使用基于最小二乘法和估计样本一致性算法将球体模型拟合到苹果的特征点, 最后应用模板匹配对苹果点进行三维模型拟合, 准确地测量出球体果实的体积信息。虽然该方法能够在一定程度上应对目标物体遮挡的挑战, 但在求解特征点以及数据的采集处理方面仍有待进一步研究, 以提高方法的效率和实用性。LIAO 等<sup>[17]</sup>开发了一款 3DPheoFish 软件, 从三维点云数据中提取鱼类形态表型, 为区分不同物种样本提供了信息特征。但这样的方法依赖于高精度的工业级点云扫描仪, 数据采集时需要将鱼类捕捞并麻醉进行工作, 数据获取难度大, 且会对生物体造成不可挽回的损伤。因此, 需要一种保持海参原位的体积测量方法, 在不损害海参的前提下进行高准确率的体积估计。

水下原位海参的体积测量还存在一定的挑战。其一, 水下光线多变、环境复杂, 如何获取有效精准的三维数据是一个难题; 其二, 获取的点云数据疏密不齐且存在噪声, 对后续的体积计算有很大影响。为此, 本文提出一种基于双目视觉<sup>[18-19]</sup>的体积测量方法, 在不惊动海参、得到最真实数据的前提下, 保持较高的检测精度和良好的分割效果, 且不借助特征点的选取, 而是利用三维点云信息, 进行海参整体的体积测量。该方法以真实养殖

场景下的原位海参作为研究对象, 首先借助双目摄像头搭载无人船在海参养殖圈采取完整的数据视频; 其次, 通过改进的 YOLOv8 算法对海参进行目标检测, 加强水下环境的检测效果; 接着, 用调整过后的 SAM (segment anything model) 分割模型进行海参的分割; 最后, 结合分割掩码和三维信息的映射结果, 用泊松重建和体素化的方法进行海参体积的测量。以期满足养殖户的评估需求, 为海参的科学养殖提供更为精准的数据支持。

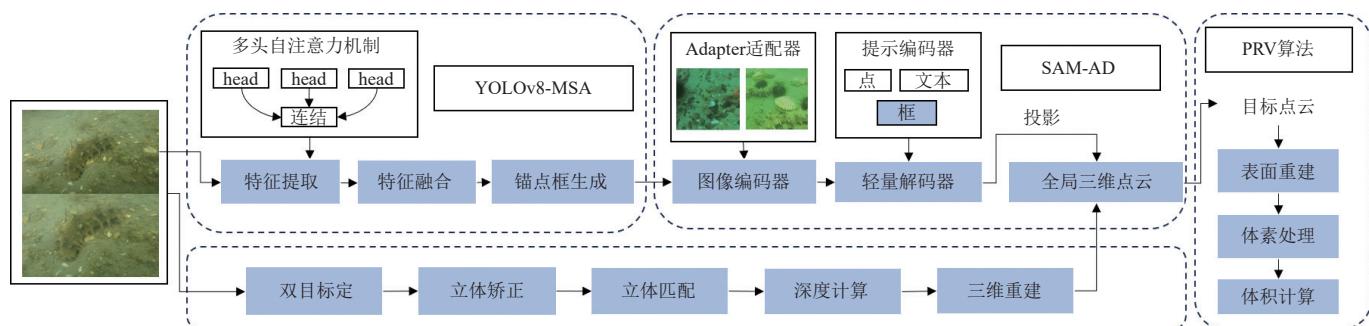
## 1 材料与方法

### 1.1 海参数据集构建

本研究利用双目相机设备, 在大连鑫玉龙海参养殖圈内采集海参视频, 制作了包含 3 192 张图片的数据集, 囊括了不同季节下海参的各种生长形态。拍摄获取的视频为 30 帧/s, 使用抽帧脚本以 30 帧抽一次的手段抽取 3 192 张图像, 并使用 labelimg 标注工具以 yolo 格式进行标注。数据集未使用离线数据增强, 在训练过程中使用在线数据增强, 数据集以 8:1:1 的比例随机划分为训练集 (2 554 张)、测试集 (319 张)、验证集 (319 张), 从而确保试验数据的真实有效。

### 1.2 海参体积测量方法总体框架

双目视觉是机器视觉的一种重要形式, 它是一种基于视差原理并利用成像设备从不同的位置获取被测物体的两幅图像, 通过计算图像对应点间的位置偏差, 来获取物体三维几何信息的方法。该框架如图 1 所示, 主要包括以下几个步骤: 海参目标检测、海参目标分割、海参点云体积测量。首先通过搭载双目相机<sup>[20]</sup>的水下图像采集设备拍摄海参养殖圈的相关视频, 然后采用 YOLOv8-MSA (YOLOv8 multi-head self-attention) 算法对图像中的海参进行检测, 接着采用 SAM-AD (SAM adapter) 算法借助检测框信息, 得到只属于海参的 mask 掩码信息, 利用双目技术获取到的全局点云信息和海参的掩码信息进行映射投影, 从而得到只属于海参的点云信息, 最后利用泊松重建和体素化处理算法 PRV (poisson reconstruction and voxelization) 求解海参的体积。



注: YOLOv8-MSA 表示 YOLOv8 multi-head self-attention 模型, 用于生成海参检测框; SAM-AD 表示 segment anything model adapter 模型, 用于生成海参掩码信息; PRV 表示 poisson reconstruction and voxelization 算法, 用于海参体积计算。

Note: YOLOv8-MSA represents YOLOv8 multi-head self-attention model, which is used for generating sea cucumber detection boxes; SAM-AD represents segment anything model adapter model, responsible for generating sea cucumber mask information; PRV presents poisson reconstruction and voxelization algorithm, which is applied to calculate the volume of sea cucumbers.

图 1 海参体积测量总体框架  
Fig.1 Overall framework for measuring the volume of sea cucumbers

### 1.3 融合多头自注意力的海参检测器的设计

YOLOv8 检测模型在速度和精度之间实现了优良的平衡, 具备高效的特点提取和轻量化设计, 适用于实时检测应用<sup>[21-22]</sup>。它采用先进的损失函数, 提高了检测精度和鲁棒性。结合对于速度和精度两方面的考量, 本研究选用 YOLOv8s 作为本文的基础检测器。但是, 水下环境具有其独特性。由于水下的光线条件与陆地上大不相同, 水的吸收和散射作用会导致光线减弱, 色温出现偏差。YOLOv8s 有时难以适应这种光线条件的变化, 从而影响检测精度。此外, 水下图像往往由于水体颗粒物和运动模糊导致图像质量降低。YOLOv8s 的性能在处理这些噪声和模糊时会出现下降, 特别是在小目标和复杂背景下。

对此, 自注意力机制可以提供良好的解决方案<sup>[23-24]</sup>。自注意力机制可以动态分配特征图上的权重, 模型能够关注到重要的区域和特征, 忽略无关的背景噪声。这对于水下环境中复杂背景和噪声干扰显著的情况尤为重要。

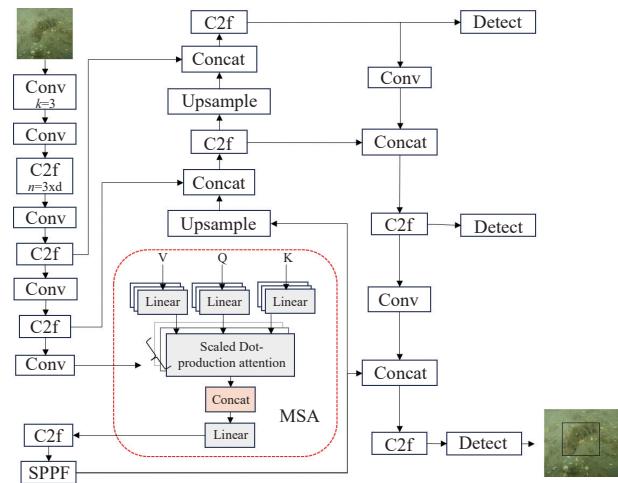
多头自注意力机制是一种有效的注意力机制, 它通过将输入序列分割成多个头部, 并对每个头部的信息进行并行处理, 从而提取更丰富的特征表示。多头自注意力机制允许模型在不同的子空间中学习到不同的特征表示。每个注意力头可以关注输入序列中的不同部分或不同的特征维度, 从而更好地捕捉输入数据的细节。水下环境通常具有复杂的光照条件、较高的噪声水平以及遮挡等问题。多头自注意力机制能够从多个视角分析和理解水下目标的特征, 帮助模型在这些复杂条件下更准确地识别和检测海参。此外, 通过多个头的并行计算, 多头自注意力机制能够在不同的头之间产生更多样化的表示, 使得模型对噪声和输入变化更加鲁棒。最后, 多头自注意力机制能够有效地融合来自不同注意力头的多样化信息, 这种信息融合有助于提升模型对复杂场景的理解能力。水下目标检测通常需要处理大量的感知信息, 包括颜色、形状、纹理等。多头自注意力机制可以有效地整合这些多维度信息, 从而更准确地识别和定位目标。综上所述, 多头自注意力机制在水下目标检测任务中具有显著的优势。

本研究在主干网络中引入多头自注意力机制<sup>[25]</sup> (multi-head self-attention, MSA), 构建了 YOLOv8-MSA 算法模型, 模型结构如图 2 所示。本文在 Backbone 的最后一层卷积神经网络中引入多头自注意力机制, 从而更好地捕捉图像中的细节, 增强模型的鲁棒性, 提高复杂场景下对目标的识别能力。通过这种注意力机制, 模型能够更好地理解图像内容, 并从中提取更丰富的特征表示, 从而更准确地识别目标。

自注意力机制是 Transformer 神经网络中的核心组件之一, 对输入序列中的元素进行加权表示, 以捕捉元素之间的相关性。自注意力机制通过计算元素之间的相似度来找到相关的信息。为此, 使用点积注意力机制 (dot-product attention) 来计算每对元素之间的相似度。对于向量  $X_i$  和  $X_j$ , 它们之间的相似度表示见式 (1)。

$$S(X_i, X_j) = \frac{X_i \cdot X_j}{\sqrt{d}} \quad (1)$$

式中  $S$  为  $X_i$  和  $X_j$  的相似度,  $d$  为向量的维数,  $X_i \cdot X_j$  为向量  $X_i$  和  $X_j$  的点积。



注: Conv 为卷积操作, C2f 为特征提取操作,  $k$  表示卷积核的大小,  $n$  表示模块内部堆叠的瓶颈块的数量, SPPF 为空间金字塔池化结构, Unsample 为上采样操作, Concat 表示特征融合, Linear 为线性变换, Scaled Dot-production attention 为缩放点积注意力机制, Detect 为检测头。  
Note: Conv refers to the convolution operation, C2f denotes the feature extraction operation,  $k$  represents the size of convolutional kernels,  $n$  denotes the number of stacked bottleneck blocks within the module. SPPF represents the spatial pyramid pooling structure, Unsample indicates the upsampling operation, Concat signifies feature fusion, Linear refers to linear transformation, Scaled dot-product attention denotes the scaled dot-product attention mechanism, and Detect refers to the detection head.

图 2 YOLOv8-MSA 网络结构示意图  
Fig.2 YOLOv8-MSA(multi-head self-attention) network architecture diagram

接下来, 将相似度转换为权重, 以便在计算上下文表示时使用。为此, 需要对相似度进行归一化, 使它们加起来为 1。具体来说, 使用 Softmax 函数将相似度作为指数, 以得到每个元素的权重。对于向量  $X_i$  和  $X_j$ , 它们之间的权重见式 (2)。

$$W(X_i, X_j) = \frac{\exp(S(X_i, X_j))}{\sum_1^n (S(X_i, X_j))} \quad (2)$$

式中  $n$  为输入序列的长度,  $W(X_i, X_j)$  为第  $i$  个向量对于第  $j$  个向量的权重。此时, 对于每个向量  $X_i$ , 都得到了一个权重向量, 最后使用权重向量  $W(X_i)$  来计算每个向量的上下文表示 (结合了全局信息之后新的特征表示)。

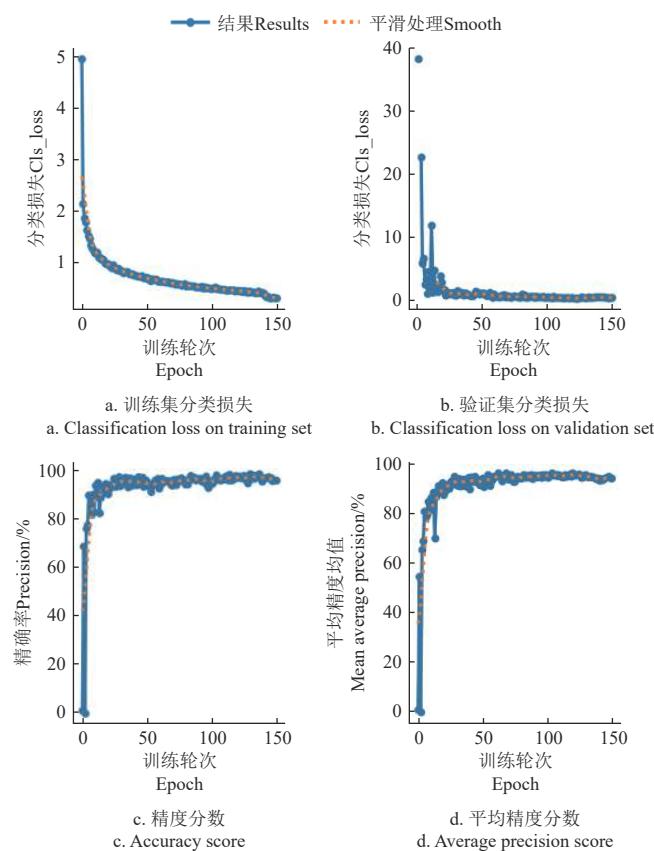
多头注意力机制允许网络在不同的空间位置或特征表示中同时执行多个不同的 SA (self-attention)。在传统的 SA 中, 神经网络通过对输入进行加权平均来获取注意力加权和, 但是这种方法仅仅考虑了一种特征表示, 而忽略了其他可能与当前任务相关的特征。为了解决这个问题, 多头注意力机制引入了多个并行的注意力机制, 从不同的角度观察输入数据, 然后将它们的结果合并起来。输入数据会通过多个线性变换, 将变换的结果分成多个头部, 每个头部独立地计算注意力权重, 并使用它们来对输入进行加权求和, 具体表示见式 (3)

$$M_h(Q, K, V) = \text{Concat}(h_1, h_2, h_3, \dots, h_n) W_0 \quad (3)$$

式中  $Q$ 、 $K$  和  $V$  为输入的查询、键和值,  $h$  为头部的数量,  $W_0$  为一个线性变换, Concat 将多个头部连接起来。以此实现关注不同的特征表示, 从而提高模型的泛化能力和表现能力。

海参目标检测器训练时, 图片大小为 640 像素×640 像素, 初始学习率为 0.01, 动量为 0.937, 权重衰减为 0.0005, 批量大小为 32, 训练回合数为 150, 每次训练回合都进行 1 次测试回合。训练所采用的操作系统为 Windows11, 深度学习框架为 PyTorch, 训练的处理器为 16 核 i7-13700KF, 显卡为 NVIDIA GeForce RTX 2080 ti。

图 3 展示了改进后的海参检测器在训练过程中的损失函数和测试指标的变化趋势。由图 3a 和 3b 可见, 无论是在训练集还是验证集中, 分类损失 (cls\_loss) 都在稳定下降, 并在训练后期趋近于零, 验证集的损失曲线也与训练集保持一致。此外, 测试指标见图 3c 和 3d, 可以看出精度和 mAP 随着训练轮次的增加, 在 25 轮左右快速达到较高水平, 均达到 0.9 左右。随后, 这些指标继续随着训练轮次的增加而稳步提升, 并最终收敛。这表明模型在训练过程中不存在验证集损失与训练集损失显著偏离的情况, 即未出现明显的过拟合现象。



注: 图 3a 和 3b 的纵坐标为分类损失的数值化表示, 用来衡量模型预测与实际标签之间的差距。

Note: The vertical axis in figure 3a and 3b represents the numerical value of the loss, which is used to measure the difference between the model's predictions and the actual labels.

图 3 训练过程中的损失函数和测试矩阵

Fig.3 Loss function and test matrix in training process

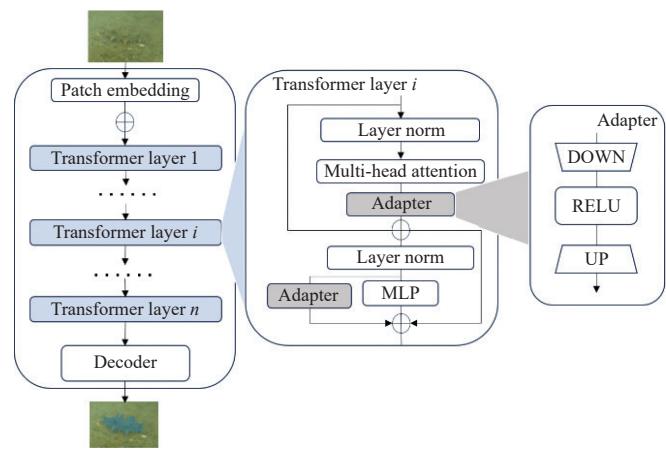
可见, 改进后的海参检测器不仅在训练过程中表现出色, 而且在测试阶段也具有较高的稳定性和精度。模

型能够在较短的训练时间内迅速达到较高的检测性能, 具有良好的稳定性和泛化能力, 证明了其在水下应用场景中的有效性。

#### 1.4 融合 adapter 机制的海参分割模型的改进

分割的目的是从左视图中分割出海参, 并给出其在图像平面上的掩码信息, 本研究选用 SAM<sup>[26-29]</sup> 作为图像分割模型。SAM 具备高度通用性和灵活性, 能够处理不同类型的分割任务, 它可以在不需要额外训练的情况下对不熟悉的对象和图像进行零样本泛化, 从而“剪切”任何图像中的任何对象。SAM 利用卷积神经网络对图像进行特征提取, 并利用轻量级编码器将任何提示转换为嵌入向量, 然后在轻量级解码器中组合这两个信息源, 预测分割掩码。然而, 在实际养殖环境下, 海参常常具有与周围环境高度相似的颜色和纹理, 使得难以准确区分和分割。由于海参的伪装特性, SAM 在分割任务中容易受到干扰, 导致分割精度下降, 限制了模型在水下任务中的有效性。

为了提升 SAM 在水下场景中的分割性能, 增强其在下游任务中的表现能力, 本研究在 SAM 模型中引入 adapter 机制, 提出 SAM-AD 模型。调整后的模型框架如图 4 所示。adaptor 机制是一种轻量级的模块, 可以插入到预训练模型的中间层中, 实现对新任务或新领域的快速适应。



注: Patch Embedding 为图像嵌入; Transformer Layer 为转换层, 用于捕捉图像全局依赖关系和上下文信息; MLP 为多层感知机; decoder 为解码器。Note: Patch Embedding refers to image embedding; Transformer Layer refers to the transformation layer, which is used to capture the global dependencies and contextual information of the image; MLP represents multilayer perceptron; decoder refers to the decoding unit.

图 4 SAM adapter 模型结构图

Fig.4 SAM(segment anything model) adapter model architecture diagram

具体而言, 所有适配器的架构是较为统一的: 一个低阶全连接层 DOWN, 一个 RELU 激活函数和一个高阶全连接层 UP, 它可以保留预训练模型的参数不变, 只需训练少量的参数即可达到良好的效果, 同时减少了计算资源和存储空间的需求。每个 Transformer 块由两个子块组成, 分别是多头自注意力层和 MLP 层。每块的自注意力层会搭配一个 adapter 适配器, 在第二个 MLP 子块之前进行残差连接。在第二个 MLP 层中, 另一个全新的适

配器会与原始的 MLP 层并行, 从而生成当前块的输出特征。使用更多样化和高质量的水下图像来预训练 SAM 模型, 借此增强其在水下图像中的泛化能力和准确性, 使其能够适应水下环境的复杂性和多样性。

此外, 针对水下任务的特殊环境, 该模型选用了 YOLOv8 的识别框作为提示方式来指导 SAM 进行分割。改进后的 YOLOv8 在水下图像中能够快速且准确地定位出物体的位置和大小, 从而为 SAM-AD 提供了有效的先验信息。将 YOLOv8 检测到的海洋珍品的边界框作为输入, 传递给 SAM 模型。SAM 模型将其作为一个提示, 根据边界框的位置和大小, 生成属于目标物体的分割掩码, 从而实现海洋珍品的分割。这样的补充操作有效减少了分割过程中的误差和噪声, 提高了分割的精度和效率。

## 1.5 基于泊松重建和体素化处理的点云体积计算方法

将 mask 掩码映射到由双目信息构建的全局点云, 据此即可得到只属于目标物体的点云信息<sup>[30-32]</sup>。但这样的海参点云数据存在一个问题。经过分割后获得的海参点云并不是完全闭合的。环境的复杂以及相机的精度都可

能带来一部分的点云噪声。此外, 点云模型的下部呈镂空状, 有大部分的数据缺失。同时, 体积计算时还可能涉及高度值的确定, 这都会对体积计算带来影响。因此, 在进行体积计算之前, 首先需要对点云数据进行表面重建以达到封闭图形的效果。

传统的表面重建方法各有优劣。Alpha shapes<sup>[33]</sup> 算法是一种非常灵活的表面重建方法, 通过调整参数  $\alpha$  进而控制重建结果的细节水平, 但会产生过度光滑或不完整的表面, 也容易出现拓扑错误, 从而出现不连通的表面或孔洞; Ball pivoting<sup>[34]</sup> 法通过模拟球在点云表面滚动, 逐步连接点云形成表面模型来计算体积, 尽管该方法在处理点云数据方面具有一定优势, 但当点云数据中存在较多的噪声点时, 滚动球容易受到干扰, 生成不准确的表面模型, 从而影响体积计算结果, 产生较大的波动偏差; 泊松重建<sup>[35]</sup> 法通过对点云数据进行密集采样并利用法线信息推断表面连续性, 生成连续且无漏洞的网格模型。综上所述, 本文提出了一种基于泊松重建和体素化处理的 PRV 点云体积计算方法, 总体流程如图 5 所示。

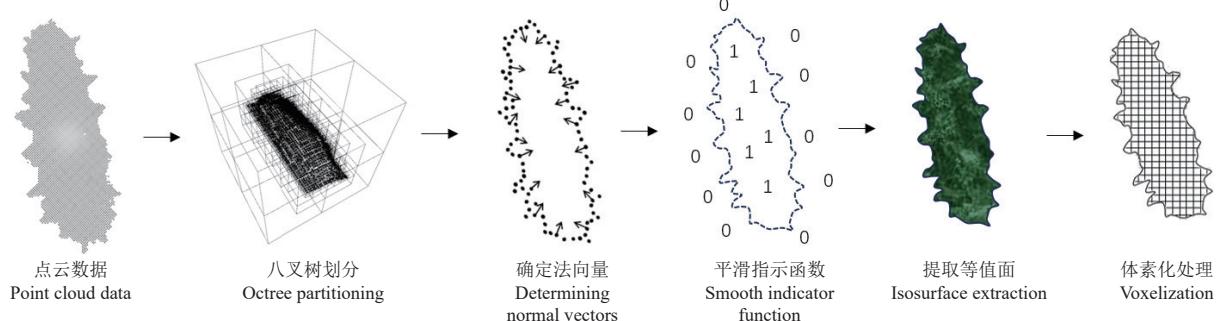


图 5 PRV 算法体积计算流程图

Fig.5 PRV (poisson reconstruction and voxelization) algorithm volume calculation flowchart

首先, 将物体点云数据利用八叉树划分空间区域, 将每个点放到正方形的空间里, 设定某一阈值, 如果该正方形内的数据点超过这个阈值, 就将该正方体再次划分, 直到每个小正方体包含的点数小于或等于这个阈值, 每个小正方体就是一个叶节点, 深度为  $D$ , 调整  $D$  的参数大小即可改变重建的精密程度。其次, 泊松重建的过程很大程度上依赖于法线向量的确定。因此, 需要估计每个点的法线向量, 以便了解表面的方向信息。法线向量是垂直于表面的向量, 可以指示表面的朝向和曲率。本文采用了基于最近邻搜索的方法来估计每个点的法线向量。具体而言, 使用 KD 树 (K-dimension tree) 搜索方法, 通过搜索每个点周围的最近邻点来估计法线的方向。基此, 通过隐式地拟合一个由物体派生的指示函数, 可以给出一个平滑的物体表面的估计。给定一个区域  $S$  及其边界  $\partial S$ , 指示函数  $X_s$  定义见式 (4)。

$$X_s = \begin{cases} 1, & x \in S \\ 0, & x \notin S \end{cases} \quad (4)$$

这个指示函数定义模型内部的值大于 0, 外部的值小于 0, 而为 0 的部分即为等值面  $\partial S$ , 提取出来就是几何目标模型的表面。

指示函数作为一个分段函数, 不具有连续性, 因此需要用滤波器来平滑卷积指示函数, 然后再求解平滑函数的梯度场。根据高斯散度理论, 平滑指示函数的梯度等于平滑表面法向场得到的向量场。对于任意点  $p \in \partial S$ , 定义  $\vec{N}_{as}(p)$  表示向内的表面法向量,  $q$  为表面的任意一点,  $F(q)$  为是一个平滑滤波器,  $F_p(q) = F(q - p)$  是  $F$  沿  $p$  方向的平移。 $X_s$  一般意义上是不好求导的, 这里用  $X_s \times F$  的导数来近似, 记作  $\tilde{F}$ , 具体见式 (5)

$$\nabla(X_s \times \tilde{F})(q) = \int_{\partial S} \tilde{F}_p(q) \vec{N}_{as}(p) dp \quad (5)$$

式中  $\nabla$  为散度算子;  $\times$  是卷积符号, 这里通过卷积实现平滑操作。 $\tilde{F}$  表示平滑函数,  $q$  点为表面的任意一点, 用于确定平滑处理的方向,  $\tilde{F}_p$  表示  $p$  点邻域内的平滑处理,  $\tilde{F}_p(q)$  表示沿方向的平移。

由于物体表面点云数据是离散的,  $\vec{N}_{as}(p)$  对于表面每个点  $q$  的分布是未知的, 这就需要通过观测  $p = (p_i, n_i)$  来近似。将初始离散样本点集记为  $T$ ,  $t$  为  $T$  中的一个点 ( $t \in T$ ),  $t$  包含位置信息 ( $t.p$ ) 和法向量信息 ( $t.\vec{N}$ )。将  $\partial S$  按照空间划分成不同的表面区域  $\partial T$ , 即  $t \in T$ ,  $\partial T \subset \partial S$ 。将上式转化成积分求和, 其中每个小的积分

可以近似为常函数，可以用  $(t,p)$  对应的函数和  $\partial T$  的面积的积分代替，见式 (6)。

$$\begin{aligned} \nabla \left( X_s \times \tilde{F} \right)(q) &= \sum_{t \in T} \int_{\partial T} \tilde{F}_p(q) \vec{N}_{\partial S}(p) dp \\ &\approx \sum_{t \in T} |\partial T| \tilde{F}(q - t.p) t \cdot \vec{N} = \vec{V}(q) \end{aligned} \quad (6)$$

式中  $\vec{V}(q)$  表示划分表面后点云的向量空间。

这样，由梯度关系得到采样点和指示函数的积分关系，根据积分关系利用划分块的方法获得点集的向量场，计算指示函数梯度场的逼近，从而构成泊松方程。向量空间  $\vec{V}$  和  $X_s$  满足式 (7)。

$$\nabla \vec{X} = \vec{V} \quad (7)$$

式中  $\vec{V}$  表示整体点云的向量空间。  $\vec{X}$  表示平滑后的指示函数。

将两边同时求导，得到式 (8)。

$$\Delta \vec{X} = \nabla \cdot \nabla \vec{X} = \nabla \cdot \vec{V} \quad (8)$$

式中  $\Delta$  为拉普拉斯算子，方程的解可以用拉普拉斯方程基本解与函数卷积求出，得到指示函数。

然后，根据泊松方程使用矩阵迭代求出近似解，采用移动立方体算法提取等值面，对所测数据点集重构出被测物体的模型，泊松方程在边界处的误差为零，因此得到的模型不存在假的表面框。

最后，一旦获得了连续的三角形网格，就可以将其转换为体素表示。体素化是将连续的几何表面离散化为规则网格的过程。在本研究中，使用体素化的方法，将连续的三角网格转换为规则的立方体网格。体素化过程中，通过指定体素的大小，即体素边长，以决定体素的密度和最终体积估算的精度。通过统计体素网格中的非零元素数量即可估算网格的体积。每个非零元素代表了体素网格中的一个小体积单元，对应于原始网格中的一部分。因此，通过计算非零元素的数量并乘以体素的体积，可以得到对原始网格体积的估算。

## 2 结果与分析

### 2.1 海参体积测量方法可行性验证

为了验证本文所提出的方法的可行性，在实际测量过程中，本研究使用高分辨率的双目摄像头拍摄海参，从不同角度获取两张图像，通过图像处理技术提取海参的三维点云数据。随后，程序对点云数据进行处理，计算得到海参的体积。[图 6](#) 为海参体积测量程序的实例展示。

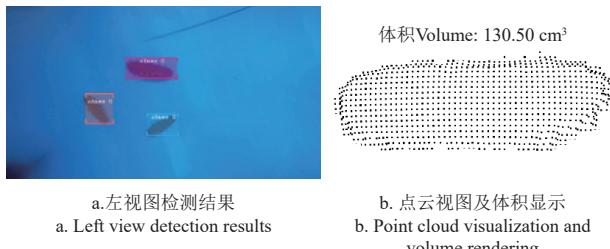


图 6 海参体积测量程序实例图

Fig.6 Example of sea cucumber volume measurement program

### 2.2 海参体积测量方法有效性验证

#### 2.2.1 检测模型对比试验

为了验证 YOLOv8-MSA 模型的有效性，本研究进行了检测精度的对比试验。所有试验均在配备 NVIDIA RTX 2080 ti GPU 的计算机上进行，系统为 Windows 11，使用的深度学习框架为 PyTorch1.8.1，CUDA 版本为 11.7。试验使用自制的水下原位海参图像的数据集，对比了 Faster RCNN、Cascade RCNN、YOLOv7、YOLOv8s 和 YOLOv8-MSA5 种模型的检测性能。

试验数据具体如表 1，结果表明，Faster RCNN 和 Cascade RCNN 作为目前表现很好的两阶段目标检测算法，在检测精度方面表现突出，mAP@0.5 分别达到了 88.9% 和 89.4%，显著高于 YOLOv7 约 10 个百分点。然而，这两种模型的参数量和 FLOPs 都较为庞大，尤其是 Cascade RCNN，其参数量达到 69.4 M，FLOPs 为 256.3 G，大大降低了检测速度。这说明，尽管 Faster RCNN 和 Cascade RCNN 在精度上占优，但它们不具备足够的效率。相比之下，YOLOv8s 在大幅减少参数量和计算量的同时，依然保持了较高的精度。本文方法在其基础上融合了 MSA 模块之后，其 mAP@0.5 达到了 92.5%，相较于原始 YOLOv8s 提高了 1.3 个百分点，同时参数量和 FLOPs 仅有 11.4 M 和 28.9 G。这表明，YOLOv8-MSA 检测模型在水下海参的检测方面具有一定的优势。此外，YOLOv7 的 R 指标为 0.641，显著低于其他模型。这表明，尽管 YOLOv7 在 P 指标上表现良好，但在召回率方面存在不足，水中的复杂环境导致部分目标未被检测到。与此相比，YOLOv8s 和 YOLOv8-MSA 在保持高精确率的同时，R 指标也较高，分别为 0.873 和 0.885，使得它们的 F1 得分显著高于其他检测模型。可见，YOLOv8-MSA 引入 MSA 模块，有效提升了模型的整体检测性能，展现了综合性能的最佳表现。

表 1 检测模型性能对比分析

Table 1 Performance comparison and analysis of detection models

模型 Models	mAP@ 0.5%	FPS/ (帧·s <sup>-1</sup> )	P/%	R/%	F1 评分 F1 score/%	Flops/G	参数量 Params/M
Faster RCNN	88.9	11	0.769	0.901	0.829	208.1	40.7
Cascade RCNN	89.4	13	0.789	0.914	0.847	256.3	69.4
YOLOv7	78.2	24	0.910	0.641	0.752	104.7	37.2
YOLOv8s	91.2	33	0.937	0.873	0.904	28.5	11.3
YOLOv8-MSA	92.5	31	0.963	0.885	0.922	28.9	11.4

注：mAP@0.5 为交并比为 0.5 时的平均精度均值，FPS 为帧率，P 为精确率，R 为召回率，Flops 为浮点计算量。

Note: mAP@0.5 is the mean average precision when the IoU threshold is 0.5, FPS stands for frames per second, P is precision, R is recall, Flops is floating point operations per second.

#### 2.2.2 表面重建效果对比试验

为了验证表面重建算法的有效性，本研究进行了 3 种表面重建算法的效果对比试验。试验随机选取了一帧包含海参的图像并获取其点云数据，分别运用 Alpha shapes、Ball pivoting 以及泊松重建 3 种不同的表面重建算法对海参进行表面重建。[图 7](#) 为 3 种不同方法的海参表面重建效果图。可见，Alpha shapes 表面重建的效果不如人意，海参的形态特征没有得到较好的体现，并且还

存在较大的空洞区域; Ball pivoting 重建的效果有所提高, 但仍然存在许多零碎的空洞区域; 泊松重建的效果最佳, 不仅能体现出海参的模型表征, 且重建后的表面具有连续性, 贴合海参的真实形态。这给后续的体素化处理以及体积测量提供了更有利的数据支持。

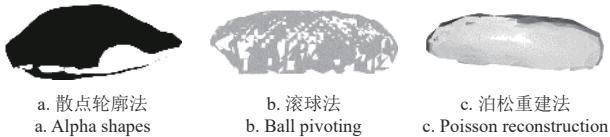


图 7 海参表面重建效果图

Fig.7 Surface reconstruction effect diagram of sea cucumber

### 2.2.3 体积测量算法对比试验

为了验证体积测量的有效性和准确性, 本研究进行了体积准确度测量的相关试验。排水法是传统的体积测量方法之一, 具有较高的精度和可靠性。本试验选取了大小不一的 10 个海参模型, 使用排水法分别测量其实际体积, 随后, 利用本文提出的体积测量算法, 在 100 cm 的深度下, 对相同的海参进行测量。计算结果如表 2 所示。

将实际体积和 PRV 算法测量的体积结果进行比较, 发现程序测量体积的平均准确率在 94% 左右。这一结果表明, 基于双目视觉的体积测量程序具有较高的准确性。

表 2 不同方法下海参体积测量准确率比较

Table 2 Comparison of accuracy in sea cucumber volume measurement using different methods

算法 Algorithm	体积 Volume/cm <sup>3</sup>										平均准确率 Average accuracy/%
	1	2	3	4	5	6	7	8	9	10	
实际体积	193	213	175	226	290	358	223	272	146	213	
最小包围框法	253	300	236	332	446	465	323	373	180	296	72
Alpha shapes	112	136	124	383	136	240	134	446	73	292	62
Ball pivoting	164	172	154	167	367	393	187	185	113	168	80
PRV (本研究)	176	198	177	199	278	365	207	256	143	194	94

为了验证本文算法在不同的检测距离下的可靠性, 本研究进行了不同深度下的测量精度对比。试验设定在 70、100、120、150 cm 4 个不同距离下, 分别运用上述 4 种方法进行了海参体积测量准确率的计算。试验结果如表 3 所示。

表 3 不同深度海参体积测量准确率比较

Table 3 Comparison of accuracy in sea cucumber volume measurement at different depths

深度 Depth/cm	准确率 Accuracy/%			
	最小包围框法	Alpha shapes	Ball pivoting	PRV (本研究)
70	74	46	75	91
100	46	62	80	94
120	75	68	82	88
150	91	53	77	83

可见, 无论在哪一个深度情况下, PRV 算法测量的准确率均优于其他 3 种方法。结合双目相机的最佳工作距离, 在距离为 100 cm 左右时达到最高的体积测量精度 94%, 相比最小包围框法、Alpha shapes、Ball pivoting 方法分别提高了 22、32、14 个百分点。上述试验结果表明, 泊松重建法的体积测量结果偏差在可控范围内, 具有更高的精度和稳定性。

和可靠性, 操作简便, 测量速度快, 能够有效替代传统的排水法进行海参体积测量, 适合大规模养殖场景中的快速检测需求。

此外, 采用不同的测量算法对海参的点云模型进行了体积测量, 分别使用 PRV 算法、最小包围框测量法、Alpha shapes 法和 Ball pivoting 法 4 种不同的方法进行对比分析, 以确定最适合的测量方法。以第一组数据为例, 最小包围框法通过构建一个最小的长方体将整个点云模型包围起来, 以此计算体积。然而, 由于包围框包含了海参之外的部分空间, 这导致体积估算的上限较高, 计算得到的体积普遍偏大 (253 cm<sup>3</sup>), 无法精确反映海参的真实体积。尤其对于形态不规则的海参, 包围框的体积会明显大于实际体积, 导致测量结果不准确。Alpha shapes 法构建了过于光滑的表面, 且出现了一部分空洞, 对重建的效果和体积的测量也存在较大影响。Ball pivoting 法在实际测量中由于点云数据的离散性和噪声, 导致体积测量误差较大, 稳定性较差。泊松重建法通过对点云数据进行密集采样并利用法线信息推断表面连续性, 生成连续且无漏洞的网格模型。该方法能够较好地贴合海参的真实形态, 提供更为准确的体积测量结果 (176 cm<sup>3</sup>), 与实际体积 (193 cm<sup>3</sup>) 最为接近。

### 3 结论

本文提出了一种基于双目视觉和点云处理的海参体积计算方法, 通过双目相机获取到目标物体的全局信息, 然后利用二维信息进行检测和分割, 接着通过二维和三维的映射关系, 得到海参的三维点云数据, 并通过表面重建算法对数据进行过滤和优化, 最后求解海参的体积。

1) 本文在 YOLOv8 的基础上, 加入多头自注意力机制, 构建 YOLOv8-MSA (YOLOv8 multi-head self-attention) 检测模型, 平均精度均值达到了 92.5%, 相较于原始模型提高了 1.3 个百分点。此外, 运用融合 adapter 机制的 SAM-AD (segment anything model adapter) 分割模型进行海参的目标分割, 为海参的体积测量提供精准可靠的掩码数据。

2) 本文构建融合泊松重建和体素化处理的 PRV (poisson reconstruction and voxelization) 算法进行海参体积的测量。试验表明, 该方法在深度为 100 cm 时, 测量精度达到 94%, 比表现较好的 Ball pivoting 方法高出 14 个百分点, 具有较高的精确性, 能够在实际的生产养殖过程中提供较大的帮助。

本研究采用的方法并不依赖于特征点的选取和目标

物体模型的构建，而是借助双目相机的视差分析结合图像像素级别的处理，完成整体的体积测量。该方法不仅能有效测量海参的体积，加以改进，迁移到扇贝、海胆等海洋珍品的数据测量任务上也会有不错的表现。整体的体积测量算法具有较强的鲁棒性和迁移性。

本研究采用的分割模型为 SAM，作为一款分割大模型，由于其庞大的参数量和复杂的计算过程，对计算资源的需求较高，难以满足高实时性要求的任务，导致算法整体的运行速率欠佳。此外，泊松表面重建的方法也存在对边界处理能力差、精度有限的问题，有待改进。接下来的研究中，将致力于优化分割任务，在实现高精度的同时提升运算效率。同时，优化表面重建的效果，提高体积测量算法在各个不同深度水域中的鲁棒性，使其更加智能化、通用化，能够适应各种不同的水下环境。

### [参考文献]

- [1] 段延娥，李道亮，李振波，等. 基于计算机视觉的水产动物视觉特征测量研究综述[J]. 农业工程学报，2015, 31(15): 1-11.  
DUAN Yan'e, LI Daoliang, LI Zhenbo, et al. Review on visual characteristic measurement research of aquatic animals based on computer vision[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2015, 31(15): 1-11. (in Chinese with English abstract)
- [2] 郭旺，杨雨森，吴华瑞，等. 农业大模型：关键技术、应用分析与发展方向[J]. 智慧农业（中英文），2024, 6(2): 1-13.  
GUO Wang, YANG Yusen, WU Huarui, et al. Big Models in Agriculture: Key technologies, application and future directions[J]. Smart Agriculture, 2024, 6(2): 1-13. (in Chinese with English abstract)
- [3] 张胜茂，李佳康，唐峰华，等. 基于深度学习的鱼类养殖监测研究进展[J]. 农业工程学报，2024, 40(5): 1-13.  
ZHANG Shengmao, LI Jiakang, TANG Fenghua, et al. Research progress on fish farming monitoring based on deep learning technology[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2024, 40(5): 1-13. (in Chinese with English abstract)
- [4] NEUPANE C, PEREIRA M, KOIRALA A, et al. Fruit sizing in orchard: A review from caliper to machine vision with deep learning[J]. *Sensors*, 2023, 23(8): 3868.
- [5] MIRANDA J C, GEN'E -MOLA J, ZUDE-SASSE M, et al. Fruit sizing using AI: A review of methods and challenges[J]. *Postharvest Biology and Technology*, 2023, 206: 112587.
- [6] ISLAM M S, SCALISI A, O'CONNELL M G, et al. A ground-based platform for reliable estimates of fruit number, size, and color in stone fruit orchards[J]. *HortTechnology*, 2022, 32(6): 510-522.
- [7] GUTIERREZ S, WENDEL A, UNDERWOOD J. Ground based hyperspectral imaging for extensive mango yield estimation[J]. *Computers and Electronics in Agriculture*, 2019, 157: 126-135.
- [8] NEUPANE C, KOIRALA A, WALSH K B. In-orchard sizing of mango fruit: 1. Comparison of machine vision based methods for on-the-go estimation[J]. *Horticulturae*, 2022, 8(12): 1223.
- [9] AMARAL M H, WALSH K B. In-orchard sizing of mango fruit: 2. Forward estimation of size at harvest[J]. *Horticulturae*, 2023, 9(1): 54.
- [10] MIRBOD O, CHOI D, HEINEMANN P H, et al. On-tree apple fruit size estimation using stereo vision with deep learning-based occlusion handling[J]. *Biosystems Engineering*, 2023, 226: 27-42.
- [11] 董鹏，周烽，赵惊惊，等. 基于双目视觉的水下海参尺寸自动测量方法[J]. *计算机工程与应用*，2021, 57(8): 271-278.  
DONG Peng, ZHOU Feng, ZHAO Congcong, et al. Automatic measurement of underwater sea cucumber size based on binocular vision[J]. *Computer Engineering and Applications*, 2021, 57(8): 271-278. (in Chinese with English abstract)
- [12] GONGAL A, KARKEE M, AMATYA S. Apple fruit size estimation using a 3D machine vision system[J]. *Information Processing in Agriculture*, 2018, 5(4): 498-503.
- [13] FERRER-FERRER M, RUIZ-HIDALGO J, GREGORIO E, et al. Simultaneous fruit detection and size estimation using multitask deep neural networks[J]. *Biosystems Engineering*, 2023, 233: 63-75.
- [14] 李艳君，黄康伟，项基. 基于立体视觉的动态鱼体尺寸测量方法[J]. 农业工程学报，2020, 36(21): 220-226.  
LI Yanjun, HUANG Kangwei, XIANG Ji. Measurement of dynamic fish dimension based on stereoscopic vision[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2020, 36(21): 220-226. (in Chinese with English abstract)
- [15] MORENO H, AND'UJAR D. Proximal sensing for geometric characterization of vines: A review of the latest advances[J]. *Computers and Electronics in Agriculture*, 2023, 210: 107901.
- [16] GEN'E-MOLA J, SANZ-CORTIELLA R, ROSELL-POLO J R, et al. In-field apple size estimation using photogrammetry-derived 3D point clouds: Comparison of 4 different methods considering fruit occlusions[J]. *Computers and Electronics in Agriculture*, 2021, 188: 106343.
- [17] LIAO Y H, ZHOU C W, LIU W Z, et al. 3DPhenoFish: Application for two-and three-dimensional fish morphological phenotype extraction from point cloud analysis[J]. *Zoological Research*, 2021, 42(4): 492-502.
- [18] YANI R A, MINARNI M, SAKTIOTO S, et al. Volumetric prediction of symmetrical-shaped fruits by computer vision[J]. *Science, Technology and Communication Journal*, 2020, 1(1): 20-26.
- [19] NEUPANE C, KOIRALA A, WANG Z, et al. Evaluation of

- depth cameras for use in fruit localization and sizing: Finding a successor to kinect v2[J]. *Agronomy*, 2021, 11(9): 1780.
- [20] TRAN T M, TA K D, HOANG M, et al. A study on determination of simple objects volume using ZED stereo camera based on 3D-points and segmentation images[J]. *International Journal*, 2020, 8(5): 1990.
- [21] 张玉玉, 郎树营, 纪元浩, 等. 基于改进 YOLOv8s 的玫瑰鲜切花分级方法[J]. 智慧农业 (中英文), 2024, 6(2): 118-127.  
ZHANG Yuyu, BING Shuying, JI Yuanhao, et al. Grading method of fresh cut rose flowers based on improved YOLOv8s[J]. *Smart Agriculture*. 2024, 6(2): 118-127. (in Chinese with English abstract)
- [22] KALA A, WALSH K B, WANG Z, et al. Deep learning for real-time fruit detection and orchard fruit load estimation: Benchmarking of ‘MangoYOLO’[J]. *Precision Agriculture*, 2019, 20(6): 1107-1135.
- [23] 齐咏生, 焦杰, 鲍腾飞, 等. 基于自适应注意力机制的复杂场景下牛脸检测算法[J]. *农业工程学报*, 2023, 39(14): 173-183.  
QI Yongsheng, JIAO Jie, BAO Tengfei, et al. Cattle face detection algorithm in complex scenes using adaptive attention mechanism[J]. *Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE)*, 2023, 39(14): 173-183. (in Chinese with English abstract)
- [24] 周佳龙, 季柏民, 倪伟强, 等. 基于三维姿态拟合的非接触式红鳍东方鲀全长精准估算方法[J]. *农业工程学报*, 2023, 39(4): 154-161.  
ZHOU Jialong, JI Baimin, NI Weiqiang, et al. Non-contact method for the accurate estimation of the full-length of Takifugu rubripes based on 3D pose fitting[J]. *Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE)*, 2023, 39(4): 154-161. (in Chinese with English abstract)
- [25] 任媛, 于红, 杨鹤, 等. 融合注意力机制与 BERT+BiLSTM+CRF 模型的渔业标准定量指标识别[J]. *农业工程学报*, 2021, 37(10): 135-141.  
REN Yuan, YU Hong, YANG He, et al. Recognition of quantitative indicator of fishery standard using attention mechanism and the BERT+BiLSTM+CRF model[J]. *Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE)*, 2021, 37(10): 135-141. (in Chinese with English abstract)
- [26] KIRILLOV A, MINTUN E, RAVI N, et al. Segment anything[C]// Proceedings of the IEEE/CVF International Conference on Computer Vision. Paris, France: IEEE, 2023: 4015-4026.
- [27] 王淼, 黄智忠, 何晖光, 等. 分割一切模型 SAM 的潜力与展望: 综述[J]. *中国图象图形学报*, 2024, 29(6): 1479-1509.  
WANG Miao, HUANG Zhizhong, HE Huiguang, et al. Potential and prospects of segment anything model: A survey[J]. *Journal of Image and Graphics*, 2024, 29(6): 1479-1509. (in Chinese with English abstract)
- [28] MAZUROWSKI M A, DONG H, GU H, et al. Segment anything model for medical image analysis: An experimental study[J]. *Medical Image Analysis*, 2023, 89: 102918.
- [29] CHEN T, ZHU L, DENG C, et al. Sam-adapter: Adapting segment anything in underperformed scenes[C]// Proceedings of the IEEE/CVF International Conference on Computer Vision. Paris, France: IEEE, 2023: 3367-3375.
- [30] GEN'E -MOLA J, FERRER-FERRER M, GREGORIO E, et al. Looking behind occlusions: A study on amodal segmentation for robust on-tree apple fruit size estimation[J]. *Computers and Electronics in Agriculture*, 2023, 209: 107854.
- [31] 郭奕, 黄佳芯, 邓博奇, 等. 改进 Mask R-CNN 的真实环境下鱼体语义分割[J]. *农业工程学报*, 2022, 38(23): 162-169.  
GUO Yi, HUANG Jiaxin, DENG Boqi, et al. Semantic segmentation of the fish bodies in real environment using improved Mask-RCNN model[J]. *Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE)*, 2022, 38(23): 162-169. (in Chinese with English abstract)
- [32] LI H, FU B, WANG R, et al. Point2Real: Bridging the gap between point cloud and realistic image for open-world 3D recognition[C]// Proceedings of the AAAI Conference on Artificial Intelligence. Washington, DC, USA: AAAI, 2024, 38(4): 3055-3063.
- [33] GARDINER J D, BEHNSEN J, BRASSEY C A. Alpha shapes: Determining 3D shape complexity across morphologically diverse structures[J]. *BMC Evolutionary Biology*, 2018, 18:184.
- [34] OVCHINNIKOV D A, MILEVICH A A, KRUTOV T Y. Creating 3D models using segmented point clouds[C]// 2024 6th International Youth Conference on Radio Electronics, Electrical and Power Engineering (REEPE). Moscow, Russian Federation: IEEE, 2024: 1-6.
- [35] KAZHDAN M, CHUANG M, RUSINKIEWICZ S, et al. Poisson surface reconstruction with envelope constraints[J]. *Computer Graphics Forum*, 2020, 39(5): 173-182.

## Measurement of sea cucumber volume using binocular vision

ZHEN Shuai<sup>1,2,3</sup>, LIN Yuanshan<sup>1,2,3</sup>, SHENG Yifan<sup>1</sup>, HONG Shengcheng<sup>1</sup>, WANG Wenliang<sup>1</sup>,  
CHEN Qijun<sup>4</sup>, YANG Zhiqing<sup>5</sup>, LI Zhijun<sup>1\*</sup>

(1. School of Information Engineering, Dalian Ocean University, Dalian 116023, China; 2. Dalian Key Laboratory of Smart Fisheries, Dalian Ocean University, Dalian 116023, China; 3. Key Laboratory of Environment Controlled Aquaculture (Dalian Ocean University) Ministry of Education, Dalian 116023, China; 4. Dalian Xinyulong Marine Biological Seed Industry Technology Co., Ltd, Dalian 116007, China; 5. Guigang Rongchuang Wood Industry Co., Ltd, Guigang 537000, China)

**Abstract:** Accurate volume measurement is essential to the marine treasures in aquaculture. However, existing approaches cannot fully meet the growth and evaluation in the volume measurement of sea cucumbers. In this study, the binocular vision was introduced to efficiently and precisely in-situ measure the volume of sea cucumbers. Three modules also included target detection, segmentation, and volume estimation. In the first module, the target detection was used in the variable lighting and environmental conditions under the typical underwater. The reason was that object detection previously failed to monitor underwater environments, due to the fluctuation of light conditions, reflections, and debris in the water. YOLOv8 target detection model was introduced to integrate the multi-head self-attention mechanism, in order to enhance the detection accuracy in these unpredictable conditions. This attention mechanism was significantly improved to detect the sea cucumbers under the complex underwater landscape. A high-precision system was obtained to more effectively process information from different parts of images under these challenging circumstances. In the second module, instance segmentation was used to focus on the precise identification and segmentation of sea cucumber entities. A segmentation model of sea cucumber was constructed using the segment anything model (SAM) with an adapter mechanism. The SAM model was used to more accurately isolate the sea cucumber from its background, even in the presence of noise and other marine organisms in the water. The segmentation was successfully essential for the accuracy of subsequent steps, as the generated mask served as the input for the volume estimation. High-quality mask information was provided to ensure that the shape of the sea cucumber was captured with the necessary details for accurate volume reconstruction. In the third module of volume estimation, the 2D mask maps were input into 3D space to obtain a point cloud representation of sea cucumbers. Poisson surface reconstruction and voxelization were employed to reduce the distortion of the 3D model because point cloud data often suffered from sparsity and noise. Specifically, the sparse data and noise interference were avoided for the accurate 3D model. As such, the actual volume of sea cucumber was faithfully represented for precise measurement. The experimental results demonstrate that the superior performance of the improved detection model was achieved, compared with the existing algorithms. The 92.5% accuracy and a frame rate of 31 frames per second (fps) outperformed the rest, such as Faster RCNN, Cascade RCNN, YOLOv7, and YOLOv8. Additionally, the Poisson surface reconstruction also produced a closer shape approximation of the true sea cucumber, compared with the Alpha Shapes and Ball Pivoting. More accurate volume measurements were better performed at various depths. Specifically, the accuracy of volume measurement reached 94% at a distance of 100 cm, which was 22 percentage points higher than the minimum bounding box and 14 points higher than the Ball Pivoting. Reliable data was then provided to accurately measure the sea cucumber in aquaculture. In conclusion, high efficiency and precision were obtained to measure the sea cucumber volume after 3D reconstruction in underwater conditions. The finding can also offer reliable data to monitor the growth and assess value for more accurate and efficient aquaculture practices.

**Keywords:** aquaculture; binocular vision; volume measurement; poisson surface reconstruction; deep learning; sea cucumber