融合 YOLOv7 和 BYTE 多目标跟踪的多类别海珍品计数方法

安志强[1,2,3]，李智军[1]，刘 硕[1,2,3]，赵永刚[4]，陈启俊[4]，左然涛[5]，林远山[1,2,3]※

（1. 大连海洋大学信息工程学院, 大连 116023；2. 辽宁省海洋信息技术重点实验室, 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. 大连海洋大学水产与生命学院, 大连 116023 ）

摘 要：针对目前养殖过程中海珍品计数方法成本高、效率低、计数精度难以保障等问题，该研究以真实底播养殖环境下的海珍品为研究对象，以水下拍摄的海珍品视频为数据源，提出一种基于视频多目标跟踪的多类别海珍品计数方法。首先，采用性能优异的 YOLOv7 算法实现海珍品目标检测器，为多目标跟踪提供输入；然后，结合真实养殖环境下同类别海珍品外观相似性高、不清晰等特点，借鉴 BYTE 算法的多目标跟踪思想，设计多类别轨迹生成策略和基于轨迹 ID 号的计数策略，提出一种多类别海珍品跟踪与计数方法。并提出一套更适用于基于轨迹 ID 号计数方法的评估指标。试验结果表明，改进平均计数精度、改进平均绝对误差、改进均方根误差及帧率分别为 91.62%、5.75、6.38 和 32 帧/s，各项指标多优于 YOLOv5+DeepSORT、YOLOv7+DeepSORT、YOLOv5+BYTE、YOLOv7+BYTE 等算法，尤其改进平均计数精度、帧率指标比 YOLOv5+DeepSORT 高了 29.51 个百分点和 8 帧/s，且在改进平均绝对误差、改进均方根误差指标上分别降低 19.50 和 12.08。该研究方法可有效帮助水产养殖企业掌握水下海珍品数量，为现代化渔业的测产研究提供技术参考，为水产养殖的智慧管理提供科学决策依据。

关键词：机器视觉；深度学习；海珍品计数；水产养殖；多目标跟踪

doi：10.11975/j.issn.1002-6819.202301076

中图分类号：TP391.41 **文献标志码：**A **文章编号：**1002-6819(2023)09-0183-07

安志强，李智军，刘硕，等. 融合 YOLOv7 和 BYTE 多目标跟踪的多类别海珍品计数方法[J]. 农业工程学报，2023，39(9)：183-189. doi: 10.11975/j.issn.1002-6819.202301076 http://www.tcsae.org

AN Zhiqiang, LI Zhijun, LIU Shuo, et al. Multi-category sea foods counting method integrating YOLOv7 and BYTE multi-target tracking[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2023, 39(9): 183-189. (in Chinese with English abstract) doi: 10.11975/j.issn.1002-6819.202301076 http://www.tcsae.org

0 引 言

海珍品库存是水产养殖企业在海珍品放苗、养殖状态监测、精准投喂和经济效益估算等方面的重要参考数据[1]。传统的海珍品库存估算主要采用人工抽样方法，例如称质量和框选计数等。这类方法劳动力成本高、效率低、且难以保证结果的准确性和全面性。因此，自动、准确、快速地统计完整养殖圈内海珍品的数量已成为当前迫切需要解决的问题之一[2]。近年来，无人船[3]、水下机器人[4] 和深度学习[5-6] 等新技术的发展为真正解决实际养殖环境下的海珍品数量估算问题提供了可能。

国内外学者在水下拍摄图像的基础上，尝试采用传统图像处理[7-8] 和机器学习[9-12] 技术开展海珍品自动计数的研究。传统图像处理方法的基本思路是首先在 RGB、HSV、Lab 或 YCbCr 等[13-16] 颜色空间对海珍品对象进行分割，然后采用连通区域统计、端点细化、曲线演化、模板匹配和分类检测等方法统计数量，并有效应用在鱼群、虾苗[17]、扇贝苗[18] 等的计数上。但是传统图像处理方法受水下图像模糊、对比度差、噪声和复杂背景等因素干扰[19-20]，导致海珍品数量统计的准确率普遍较低。

机器学习中的深度学习具有自动逐层提取语义特征、泛化能力强、鲁棒性好等优点[21]，为此国内外学者纷纷尝试将深度学习技术应用到海珍品计数上，涌现出分割计数、检测计数等方法。基于图像分割的计数方法，一般利用 U-Net[22]、Mask R-CNN[23] 等对图像进行像素级的背景检测和目标分割，识别出图像中的虾体[24-25]、鱼体[26] 等目标，最后在此基础上使用时间阈值法与空间阈值法确定海珍品数量。由于基于图像分割的计数方法需要对目标进行像素级的定位，计算量较大。为此，部分学者提出了基于目标检测的海珍品计数方法，其仅需要预测出目标边界框，便可得到相关计数依据，更加方便快捷。这种方法的基本思路是首先使用 Faster R-CNN[27]、YOLO[28] 等目标检测算法对图像中海珍品对象进行检测，然后根据目标边界框的数量确定海珍品的数量，已被应用于鱼群[29-30] 和扇贝[31] 的计数上。

然而，上述利用深度学习技术的海珍品计数方法大多只能统计单张图像内的海珍品数量，并未实现对完整养殖圈内海珍品的数量统计，而且只针对单一品种进行检测计数。为此，本文以真实底播养殖环境下的海珍品

收稿日期：2023-01-17 修订日期：2023-03-07

基金项目：辽宁省教育厅基本科研项目 (LJKZ0730，QL202016)；辽宁省自然基金资助计划 (2020-KF-12-09)；辽宁省重点研发计划 (2020JH2/10100043)；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)；2023 中央财政对辽宁渔业补助项目

作者简介：安志强，研究方向为目标检测、目标跟踪。Email：zhiqiang.an@qq.com。

※通信作者：林远山，博士，副教授，研究方向为机器视觉、机器人学习。Email：linyuanshan@dlou.edu.cn。

为研究对象，以水下拍摄的海珍品视频为数据源，将基于深度学习的检测模型与多目标跟踪算法相结合，尝试利用视频目标跟踪的方法同时统计多种海珍品（海参、扇贝、海胆）的数量。为保证海珍品检测跟踪算法的实时性与准确性，采用新近发布的 YOLOv7[32] 作为为目标检测网络，以期在保证检测精度的前提下实现更快的处理速度。在此基础上，借鉴多目标跟踪 BYTE[33] 跟踪算法的思想，提出一种面向多类别海珍品的跟踪与计数方法，为水产养殖的智慧管理提供科学决策依据。本文所提方法已在大连鑫玉龙海洋生物种业科技股份有限公司海参库存评估中得到应用。

1 材料与方法

本节首先介绍了多类别海珍品技术方法的总体框架，使用到的算法技术为基于检测的跟踪算法，因此又分别介绍了检测器与跟踪器的设计，最后介绍了用来评价本文所提算法的评价指标。

1.1 多类别的海珍品计数方法总体框架

本文旨在解决统计视频中各类海珍品数量的问题。鉴于多目标跟踪技术能够检测并跟踪视频中的每个物体，因此采用多目标跟踪技术实现视频中各类海珍品的计数。其基本思路是使用目标检测器检测新的海珍品，并利用跟踪器对其进行跟踪，获取视频中所有海珍品的轨迹信息，最后根据轨迹数量计算各类海珍品的数量，总体框架如图 1 所示。首先，将视频中第 i 帧的图像送入基于深度学习的海珍品目标检测器中，得到第 i 帧的检测结果；然后，将此检测结果与第 i-1 帧海珍品轨迹经跟踪轨迹预测算法得到的预测框进行目标匹配，得到第 i 帧的匹配结果；最后，根据匹配结果更新海珍品的轨迹状态；以此反复，直到整个视频处理完成。最终得到所有海珍品的轨迹，通过分析这些轨迹即可估算出各类海珍品的数量。

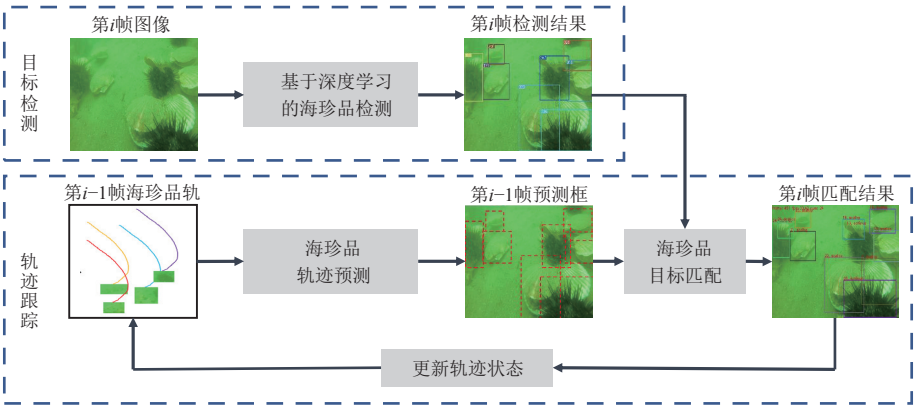

图 1 多类别的海珍品计数方法总体框架

Fig.1 General framework of multi-category sea foods counting method

1.2 基于 YOLOv7 的海珍品检测器设计

本文所采用的多目标跟踪方法为基于检测的跟踪技术，其跟踪性能严重依赖目标检测器的性能。为此，选用 YOLO 系列最新版本 YOLOv7 作为检测器基础算法。YOLOv7 是一种一阶段的目标检测算法，算法通过直接回归的方式一次处理并获得对应物体的目标区域、位置及类别，其检测速度通常比两阶段目标检测算法快，能在精度和效率上进行很好的平衡，为实现实时海珍品的跟踪计数奠定了基础。

YOLOv7 针对不同的场景和应用提供了 YOLOv7-tiny、YOLOv7、YOLOv7-X 等 7 种不同的模型，综合海珍品计数的实际需求，在此选用标准版的 YOLOv7 模型权重实现快速、高精度的海珍品检测器并使用“2017 首届水下机器人目标抓取大赛（2017 Underwater Robot Picking Contest，URPC2017）”提供的真实水下图像数据集[34] 训练 YOLOv7 模型。首先将包含海参、海胆、扇贝三类海珍品的 17 655 张图像按 8:1:1 的比例随机划分成训练集 14 055 张图片用于训练，校验集 1 800 张图片用于验证和测试集 1 800 张图片用于测试。

YOLOv7 检测器训练的设置为图片自适应大小 640×640，初始学习率为 0.01，动量为 0.9，权重衰减为 0.000 5，批量大小为 16，训练回合数为 300，每训练 10 个回合进行 1 个测试回合。海珍品检测器训练所采用的操作系统为 Ubuntu18.04，深度学习框架为 PyTorch，试验处理器为 AMD Ryzen Threadripper 1920X 12 核，显卡为 NVIDIA GeForce RTX2080。

1.3 基于 BYTE 的多类别海珍品跟踪计数算法设计

1.3.1 跟踪器设计的总体思路

实际养殖环境下的海珍品视频具有同类海珍品的相似性高、水下图像模糊缺色等特点，这可能会严重影响使用外观信息进行数据关联的多目标检测器的性能。为此，本文借鉴 BYTE 的基本思想，仅利用视频中目标的运动信息及其他启发信息实现多类别海珍品跟踪器，其总体思路如图 2 所示。

首先将检测的结果进行分级，分为高分目标 D_h 和低分目标 D_l，高分目标 D_h 与信任目标预测框 T_c 通过匈牙利算法进行第一次匹配。匹配成功的目标保存到信任轨迹中并更新，剩余的目标预测框 T_{c_r} 与低分目标 D_l 通过匈牙利算法进行第二次匹配。剩余的的高分目标 D_{h_r} 与不信任目标预测框 T_u 通过匈牙利算法进行第三次匹配。

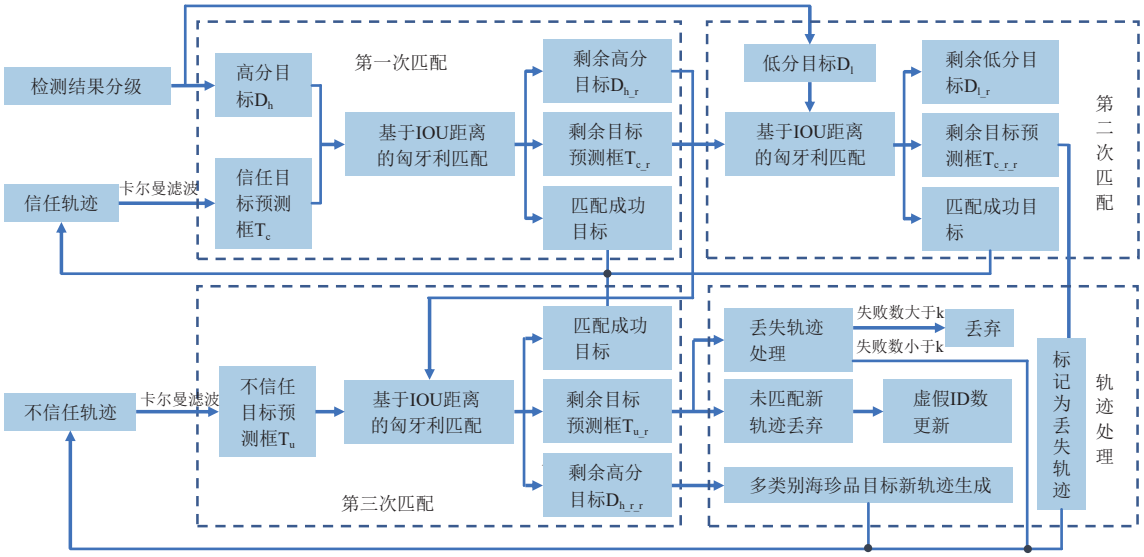

图 2　基于 BYTE 的多类别海珍品跟踪计数算法设计的总体思路

Fig.2　The general idea of multi-category sea foods tracking and counting algorithm design based on BYTE

第二次匹配成功的目标同样保存到信任轨迹中并更新，剩余的目标预测框 $T_{c_r_r}$ 标记为丢失轨迹并存放在不信任轨迹中，剩余的低分目标 D_{l_r} 将不再对其进行处理。第三次匹配成功的目标与前两次进行相同的操作，剩余的目标预测框 T_{u_r} 对于已匹配上的轨迹出现丢失的情况对其进行丢失轨迹处理，并为其设置一个阈值 k，当匹配失败的次数大于阈值 k 将其丢弃否则将其标记为丢失轨迹存放到不信任轨迹中。其他的剩余目标预测框 T_{u_r} 对其进行未匹配新轨迹丢弃处理，之后对虚假身份 (identity,ID) 数进行更新，剩余的高分目标 $D_{h_r_r}$ 将对其进行多类别海珍品目标新轨迹生成处理并将其存放到不信任轨迹中，最后更新轨迹状态，以此反复。

1.3.2　多类别海珍品新轨迹生成策略

多目标跟踪器为了更好地跟踪每个物体，需要在每个物体刚出现时为其创建一条轨迹记录并在后续帧中进行更新、维护。现有大多数基于多目标跟踪的计数方法对不同类别的物体用同一套 ID 系统无差别地创建、维护轨迹，这样需要对所有轨迹进行分析才能实现多类别物体计数，较为耗时、费事。为了充分利用 ID 本身所蕴含的数量信息，实现多类别海珍品的快速计数，为此设计一种多类别海珍品新轨迹生成策略，按类别为每类海珍品分配独立的 ID，具体如图 3 所示。

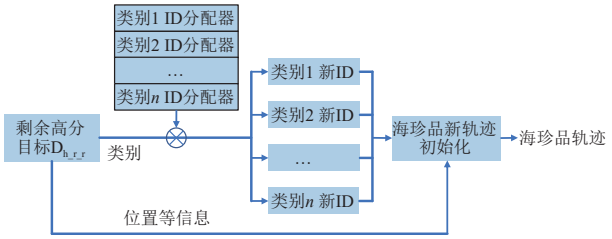

图 3　多类别 ID 分配策略流程图

Fig.3　Flowchart of the multi-class ID allocation policy

经过两次匹配仍未匹配成功的高分目标 $D_{h_r_r}$，认为其是一个新的海珍品的轨迹，因此将对其进行激活并对其进行 ID 分配，其包含的主要信息有类别、位置等。首先将高分目标 $D_{h_r_r}$ 的类别传入 ID 分配模块中，在 ID 分配模块中创建了 n 个 ID 分配器用来存放每个类别的海珍品 ID。然后根据各类别为其分配 ID，如当前类别为海参，其对应的数字类别为 1，所以类别 1ID 分配器将会为其分配一个 ID，并将原有的 ID 数量进行加一操作，其他类别以此类推，分配好的 ID 将会与其对应的位置等信息结合到一起形成最终输出的海珍品轨迹。

1.3.3　多类别海珍品计数策略

若使用上述多类别海珍品新轨迹生成策略对海珍品轨迹进行初始化，原理上每类的最新 ID 号（当前最大值）即是该类别海珍品的数量，因为最新 ID 号为此类的轨迹数。然而，经分析发现，受目标检测器精度的影响，检测结果会出现虚假的目标，进而产生虚假的轨迹，造成数量的虚报，这会对最终的计数产生重要的影响。

为了解决此问题，本文引入信任机制，即将跟踪轨迹分成信任轨迹和不信任轨迹，其中信任轨迹为连续多帧均能很好匹配的轨迹，而不信任轨迹主要包括跟丢的轨迹、新轨迹等。为统计虚假轨迹的数量，在新轨迹生成之初先将其标记为新轨迹，并加入不信任轨迹中待进一步考察，若下一帧能很好与高分目标匹配成功，则将其加入信任轨迹；否则，对应类别的虚假轨迹数加一，将对应轨迹删除。在匹配过程中优先对信任轨迹进行匹配，不信任轨迹最后匹配。为此，多类别海珍品的数量采用式 (1) 计算。

$$N^c = I^c_{max} - N^c_f \tag{1}$$

式中 N^c 为类别 c 的海珍品数量；I^c_{max} 为类别 c 海珍品最大 ID 值，可从 ID 分配器获得；N^c_f 为类别 c 海珍品虚假轨迹数量，在跟踪过程中用上述的策略计数得到。

1.3.4　多类别海珍品计数算法阐述

在上述各策略的基础上，可得到多类别海珍品跟踪计数算法，算法的输入是海珍品视频（图像序列）V，目标检测器为 Det, 跟踪分数阈值。算法的输出是海珍品

各类别的数量 N^c 以及对应个体的轨迹 T。

算法将当前帧的目标检测结果分为高分目标框集合 D_h 和低分目标框集合 D_l，并且利用卡尔曼滤波器预测上时刻轨迹的预测框集合 T_c，为后续的多次关联奠定基础。

算法经过三次关联后对未匹配的轨迹 T_{u_r} 进行删除，将仍未匹配的轨迹标记为 T_{lost}，如果 T_{lost} 中轨迹存在超过一定时间（30 帧），则从 T 中删除，否则继续在 T 中保存，对于从 T_{lost} 中删除的轨迹将其统计起来，存放到 N_f^c 中。然后初始化新的轨迹以及获得海珍品的真实数量，对于未匹配的高分目标框集合 D_{h_r} 中的检测框，如果得分高于 ϵ 且存活超过两帧，则初始化为新的轨迹。在为新轨迹分配 ID 时，首先判断其类别，根据其类别为其分配与之对应的 ID，并保存到 ID 计数器当中，然后用类别 c 海珍品最大 ID 值 I_m^c 减去类别 c 海珍品虚假轨迹数量 N_f^c，从而得到准确的 ID 值，也就是类别 c 的海珍品数量 N^c。

1.4 评价指标设计

平均计数精度 (average counting precision,ACP)、平均绝对误差 (mean absolute error,MAE)、均方根误差 (root mean square error,RMSE) 等指标通常被用来评价基于多目标跟踪的计数方法的性能。这些指标更适用于基于过线模式的计数方法，而并不适用于基于 ID 号模式的计数方法，因为基于 ID 号方法的结果受目标的误检、ID 切换和漏检影响较为严重。针对误检这一指标，一种情况是将某一类误检为另一类，但由于本文研究的 3 种海珍品之间差异较大，这种情况很少出现。另一种情况是原本此处没有海珍品，但算法错误检测了几帧，此时便为其分配了一个 ID，造成海珍品数量统计较真实情况增多，但本文通过改进算法解决了这个问题，因此没有将其考虑进评价指标。而 ID 切换和漏检是十分常见的，ID 切换会造成海珍品统计数量较真实值偏大的结果，而漏检又会使得海珍品统计数量较真实值偏小，如果出现这样的情况按照传统的指标去评价就会出现测试结果与真实值结果相近，但真实情况相差较大的问题。为此，本文提出一套考虑 ID 切换和漏检的计数平均指标，分别为改进平均计数精度 (modified average counting precision, M_{ACP})、改进平均绝对误差 (modified mean absolute error, M_{MAE})、改进均方根误差 (modified root mean square error, M_{RMSE})，具体计算见式（2）～（4）。这套指标能较客观地评价基于 ID 号模式的计数方法的性能。

$$M_{ACP} = \frac{\sum_1^n (1 - \frac{|S-G| + N_{IDs} + N_L}{G})}{N_V} \tag{2}$$

$$M_{MAE} = \frac{1}{N_V} \sum_{i=1}^{N_v} |S-G| + N_{IDs} + N_L \tag{3}$$

$$M_{RMSE} = \sqrt{\frac{1}{N_V} \sum_{i=1}^{N_V} (|S-G| + N_{IDs} + N_L)^2} \tag{4}$$

式中 S 表示算法统计的目标数量，G 表示真实数量，N_{IDs} 表示跟踪 ID 切换的次数，N_L 表示漏检海珍品的数量，N_V 表示海珍品视频的个数。

2 结果与分析

2.1 多类别海珍品计数方法有效性验证

为了验证本文方法的有效性和可用性，使用测试视频（其中的图像未用来训练）中的一个片段对其进行评价，并从该视频中随机挑选一帧用来展示，被挑选出的视频帧如图 4 所示。图中 holothurian、echinus、scallop 分别表示海参、海胆和扇贝，其对应的检测框颜色为红色、黄色和蓝色。

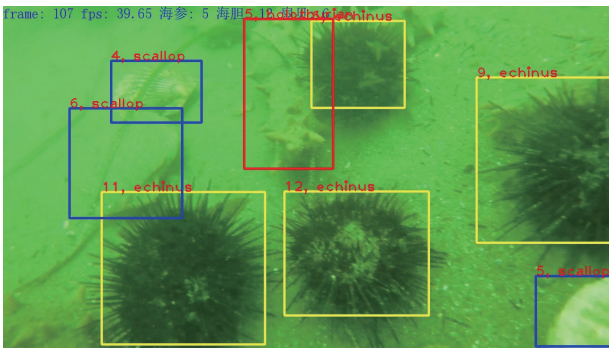

图 4 测试视频 4 的前 107 帧统计结果

Fig.4 Statistical results of the first 107 frames of test video 4

其数量为 5、12 和 6 个，这些数量为被测海珍品视频前 107 帧各类别的总和，统计结果显示在该帧的正上方。海珍品的数量是根据其真实跟踪轨迹 ID 得到的，真实跟踪轨迹 ID 为当前类别海珍品最大 ID 值减去该类别海珍品虚假轨迹数量得到的。结果表明，本文基于视频多目标跟踪的多类别海珍品计数方法能够多类别、快速准确统计水下海珍品视频中的海珍品数量。

2.2 多类别海珍品计数方法对比试验

为了充分验证本文算法的性能，以 4 个不同于训练集的海珍品视频中的海珍品为试验数据，以目标检测领域应用广泛、性能良好的 YOLOv5[35] 和 YOLOv7 检测模型为检测器，分别结合 DeepSORT[36] 和 BYTE 算法进行海珍品的数量统计。其中，测试视频的真实海珍品数量由 5 名研究人员进行统计。首先记录视频的第一帧出现的海珍品数量，然后，逐帧播放并记录新出现的海珍品数量，播放结束后每名人员分别得到各类别的数量，最后将 5 名研究人员得到海珍品数量进行取均值，用作试验的真实数值，视频 1-4 中海珍品的统计结果分别为 66、85、59 和 61。

表 1 为各算法性能对比结果。从表 1 中可以看出，DeepSORT 类的计数算法得到的结果普遍虚高，其原因是在跟踪过程中 ID 切换次数过多，比如 YOLOv5+DeepSORT 在视频 1 中产生了 7 次 ID 切换。而 YOLOv5+BYTE 和 YOLOv7+BYTE 的计数算法 ID 切换次数较少，因此得到的海珍品数量统计结果比较接近真实值，但是其漏检数量偏高，偏高的原因是跟踪过程中轨迹丢失较多引起的，将这项指标考虑进去便容易得出其结果是偏低的。本文方法无论是 ID 切换的数量还是漏检的数量都是相对较低的，得出的结果也是比较可靠的并且还实现了多类别同时计数的功能。

表 1　不同海珍品计数方法的数量统计结果和定量评价

Table 1　Quantitative statistical results and quantitative evaluation of different counting methods for sea foods

方法 Method	视频 1 Video 1			视频 2 Video 2			视频 3 Video 3			视频 4 Video 4			改进平均计数精度 Modified average counting precision /%	改进平均绝对误差 Modified mean absolute error	改进均方根误差 Modified root mean square error	帧率 Frame rate /(帧·s^{-1})
	计数结果	ID 切换	漏检	计数结果	ID 切换	漏检	计数结果	ID 切换	漏检	计数结果	ID 切换	漏检				
YOLOv5+DeepSORT	104	7	2	101	3	2	83	3	1	63	2	1	62.11	25.25	29.41	25
YOLOv5+BYTE	61	2	8	78	1	5	54	2	3	57	1	3	82.98	11.50	11.81	30
YOLOv7+DeepSORT	98	6	1	96	3	2	72	2	1	73	2	1	67.59	21.50	23.76	28
YOLOv7+BYTE	67	2	8	83	1	5	54	1	3	67	0	2	86.39	9.00	9.08	33
YOLOv7+BYTEFit(ours)	36,15,11	2	1	30,25,23	0	1	9,30,13	0	0	11,25,24	0	0	91.62	5.75	6.38	32

注：YOLOv7+BYTEFit(ours) 的视频计数结果的 3 个值分别对应的是海参、海胆和扇贝的数量。

Note: The three values of the video count result of YOLOv7+BYTEFit(ours) correspond to the number of holothurian, echinus and scallops, respectively.

由于与本文对比的算法并没有实现多类别 ID 分配功能，得出的指标均是按照总数计算得来的，其中表现较差的是 YOLOv5、YOLOv7+DeepSORT 算法，最差的是 YOLOv5+DeepSORT，其 M_{ACP}、M_{MAE}、M_{RMSE}、和帧率（FR）均最低，分别为62.11%、25.25、29.41 和25 帧/s，证明 DeepSORT 算法不适用于根据 ID 统计海珍品数量这种方法；BYTE 算法结合 YOLOv5、YOLOv7 目标检测算法表现效果更好，较好的是 YOLOv7+BYTE，其 M_{ACP}、M_{MAE}、M_{RMSE}、和 FR 分别为 86.39%、9.00、9.08 和 33 帧/s，但其计数精度不能满足实际工作需要，不适合直接应用到实际工作中，仍需加以改进；本文的 YOLOv7+BYTEFit 表现最好，其 M_{ACP}、M_{MAE}、M_{RMSE}、和 FR 分别为 91.62%、5.75、6.38 和 32 帧/s，均优于以上各算法，尤其在 M_{ACP} 和 FR 指标上比 YOLOv5+DeepSORT 高了 29.51 个百分点和 8 帧/s，在 M_{MAE}、M_{RMSE} 指标上分别降低 19.50 和 12.08。

2.3　实际应用

为了将本文所提方法更好地应用到实际的水产养殖生产中，基于所提方法实现了简单的海珍品计数系统，如图 5 所示。

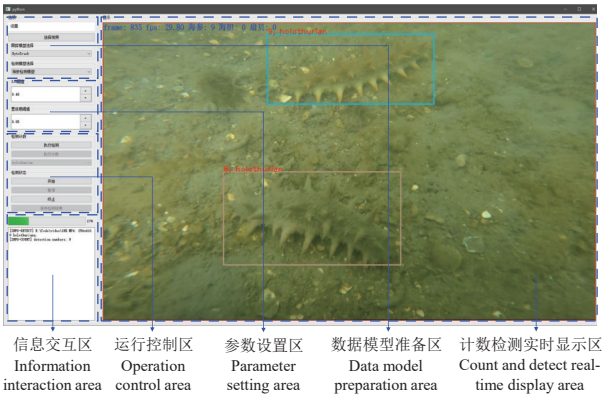

图 5　计数系统软件界面图

Fig.5　Software interface diagram of Counting system

软件界面包括海珍品计数检测信息的实时显示、跟踪检测模型的选择、运行参数的设置以及运行控制和显示运行信息等功能。计数系统工作时在计数检测实时显示区显示检测到的海珍品画面，在其顶部显示类别数量等信息；数据模型准备区用来选择跟踪检测模型和待检测视频；参数设置区可调整检测海珍品的 IoU 和置信度的阈值；运行控制区用来控制计数系统需要执行的内容以及运行结果的处理等功能；信息交互区用来显示计数系统运行的阶段和已完成检测的海珍品信息等内容。本文软件已在大连鑫玉龙海洋生物种业科技股份有限公司海参库存评估中得到实际应用。

3　结　论

本文以水下海珍品视频为研究对象，针对统计海珍品数量的问题，提出一种基于视频多目标跟踪的多类别海珍品计数方法。主要工作为设计改进 BYTE 目标跟踪方法，以及将该方法封装成系统并应用到实际生产工作中。试验结果得到如下结论：

1）测试结果显示本文方法的改进平均计数精度为 91.62%、改进平均绝对误差为 5.75、改进均方根误差为 6.38、帧率为 32 帧/s，其计数精度和计数效率满足实际的需求可有效解决传统海珍品数量统计费时费力、不精准以及安全性不高等问题。

2）本文方法与 YOLOv5+DeepSORT、YOLOv5+BYTE、YOLOv7+DeepSORT 以及 YOLOv7+BYTE 等算法进行对比，所提方法在改进平均计数精度、改进平均绝对误差和改进均方根误差这 3 个指标上优于其他算法，尤其在改进平均计数精度和帧率指标比 YOLOv5+DeepSORT 高了 29.51 个百分点和 8 帧/s，在改进平均绝对误差、改进均方根误差指标上分别降低 19.50 和 12.08，是实现多类别海珍品数量快速准确统计的有效方法。

需要指出的是，本方法对于不同水下环境的海珍品计数存在误检的情况，虽然使用相同环境的水下数据集训练检测模型可以很好地解决由于不同环境导致的误检等问题，但是水下环境因受光照等因素影响，变化较大，若每次采集海珍品视频都需先使用相同环境的数据集训练检测模型后，再进行统计计数，这将耗费巨大的人力物

力。此外，本方法还会出现 ID 切换和漏检的问题，这两个问题也是影响计数精度的关键所在。因此未来还需就水下环境与 ID 切换和漏检这两方面做进一步的改进与探索。

[参 考 文 献]

[1] 于红. 水产动物目标探测与追踪技术及应用研究进展[J]. 大连海洋大学学报，2020，35(6)：793-804.
YU Hong. Research progress on object detection and tracking techniques utilization in aquaculture: A review[J]. Journal of Dalian Ocean University, 2020, 35(6): 793-804. (in Chinese with English abstract)

[2] LI D, MIAO Z, PENG F, et al. Automatic counting methods in aquaculture: A review[J]. Journal of the World Aquaculture Society, 2021, 52(2): 269-83.

[3] BAI X, LI B, XU X, et al. A review of current research and advances in unmanned surface vehicles[J]. Journal of Marine Science and Application, 2022, 21(2): 47-58.

[4] 李道亮，包建华. 水产养殖水下作业机器人关键技术研究进展[J]. 农业工程学报，2018，34(16)：1-9.
LI Daoliang, BAO Jianhua. Research progress on key technologies of underwater operation robot for aquaculture[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2018, 34(16): 1-9. (in Chinese with English abstract)

[5] SUN M, YANG X, XIE Y. Deep learning in aquaculture: A review[J]. J Comput, 2020, 31(1): 294-319.

[6] DONG S, WANG P, ABBAS K. A survey on deep learning and its applications[J]. Computer Science Review, 2021, 40: 100379.

[7] 梁习卉子，陈兵旗，李民赞，等. 质心跟踪视频棉花行数动态计数方法[J]. 农业工程学报，2019，35(2)：175-182.
LIANG Xihuizi, CHEN Bingqi, LI Minzan, et al. Dynamic counting method of cotton rows in video based on centroid tracking[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2019, 35(2): 175-182. (in Chinese with English abstract)

[8] BARBEDO J G. Using digital image processing for counting whiteflies on soybean leaves[J]. Journal of Asia-Pacific Entomology, 2014, 17(4): 685-694.

[9] ZHAO S, ZHANG S, LIU J, et al. Application of machine learning in intelligent fish aquaculture: A review[J]. Aquaculture, 2021, 540: 736-724.

[10] LE J, XU L. An automated fish counting algorithm in aquaculture based on image processing[C]//2016 International Forum on Mechanical, Control and Automation (IFMCA 2016), Shenzhen, China, 2017:358-366.

[11] 祁洋，李亚楠，孙明，等. 基于特征融合的棉花幼苗计数算法[J]. 农业工程学报，2022，38(9)：180-186.
QI Yang, LI Ya' nan, SUN Ming, et al. Cotton seedling counting algorithm using feature fusion[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2022, 38(9): 180-186. (in Chinese with English abstract)

[12] KAKEHI S, SEKIUCHI T, ITO H, et al. Identification and counting of Pacific oyster Crassostrea gigas larvae by object detection using deep learning[J]. Aquacultural Engineering, 2021, 95: 102197.

[13] 郭常有，曹广斌，韩世成，等. 基于图像处理技术的扇贝尺寸及位置自动识别[J]. 大连海洋大学学报，2012，27(6)：578-582.
GUO Changyou, CAO Guangbin, HAN Shicheng, et al. Automatic measurement of size and location in scallop based on image processing[J]. Journal of Dalian Ocean University, 2012, 27(6): 578-582. (in Chinese with English abstract)

[14] ATMAJA RD, MURTI MA, HALOMOAN J, et al. An image processing method to convert RGB image into binary[J]. Indonesian Journal of Electrical Engineering and Computer Science, 2016, 3(2): 377-82.

[15] SALUNKHE RP, PATIL AA. Image processing for mango ripening stage detection: RGB and HSV method[C]//In 2015 Third International Conference on Image Information Processing (ICIIP).IEEE, HP Solanwakna, India, 2015:362-365.

[16] PREMAL C E, VINSLEY S S. Image processing based forest fire detection using YCbCr colour model[C]//In 2014 International Conference on Circuits, Power and Computing Technologies (ICCPCT-2014).IEEE, Nagercoil, India, 2014:1229-1237.

[17] 季玉瑶，魏伟波，赵增芳，等. 基于改进 TV-L1 模型的虾苗计数方法[J]. 青岛大学学报（自然科学版），2018，31(4)：62-68,82.
JI Yuyao, WEI Weibo, ZHAO Zengfang, et al. A Method on shrimp counting based on improved TV-L1 model[J]. Journal of Qingdao University (Natural Science Edition), 2018, 31(4): 62-68,82. (in Chinese with English abstract)

[18] 李明智，陈海泉，刘鹰，等. 扇贝苗规格识别与计数装置优化设计与试验[J]. 农业工程学报，2021，37(3)：37-46.
LI Mingzhi, CHEN Haiquan, LIU Ying, et al. Optimization design and experiments of specification identification and counting device for scallop seedlings[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2021, 37(3): 37-46. (in Chinese with English abstract)

[19] SAHU P, GUPTA N, SHARMA N. A survey on underwater image enhancement techniques[J]. International Journal of Computer Applications. 2014, 87(13): 19-23.

[20] YAN X, WANG G, WANG G, et al. A novel biologically-inspired method for underwater image enhancement[J]. Signal Processing: Image Communication. 2022, 104:116670.

[21] LECUN Y, BENGIO Y, HINTON G. Deep learning[J]. Nature, 2015,521(7553):436-444.

[22] ZHOU Z, RAHMAN Siddiquee MM, TAJBAKHSH N, et al. Unet++: A nested u-net architecture for medical image segmentation[C]//Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support, Granada, Spain, 2018: 3-11.

[23] HE K, GKIOXARI G, DOLLAR P, et al. Mask r-cnn[C]// Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 2017: 2961-2969.

[24] THAI T T N, NGUYEN T S, PHAM V C. Computer vision based estimation of shrimp population density and size[C]// Proceedings of the 2021 International Symposium on Electrical and Electronics Engineering. IEEE, Ho Chi Minh City, Vietnam, 2021: 145-148.

[25] NGUYEN K T, NGUYEN C N, WANG C Y, et al. Two-phase instance segmentation for whiteleg shrimp larvae counting[C]// Proceedings of the 2020 International Conference on Consumer Electronics. IEEE, 2020: 1-3.

[26] TSENG C H, KUO Y F. Detecting and counting harvested fish and identifying fish types in electronic monitoring system videos using deep convolutional neural networks[J]. ICES Journal of Marine Science, 2020, 77(4): 1367-1378.

[27] GIRSHICK R. Fast r-cnn[C]//Proceedings of the IEEE International Conference on Computer Vision, Santiago, Chile, 2015:1440-1448.

[28] 洪亮，王芳，蔡克卫，等. 面向海洋牧场智能化建设的海珍品实时检测方法[J]. 农业工程学报，2021，37(9)：304-311.
HONG Liang, WANG Fang, CAI Kewei, et al. Real-time detection method of seafood for intelligent construction of marine ranch[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2021, 37(9): 304-311. (in Chinese with English abstract)

[29] LAINEZ S M D，GONZALES D B. Automated fingerlings counting using convolutional neural network [C] // Proceedings of the 2019 4th International Conference on Computer and Communication Systems. IEEE, Singapore, 2019：67-72.

[30] 张璐，黄琳，李备备，等. 基于多尺度融合与无锚点

YOLO v3 的鱼群计数方法[J]. 农业机械学报，2021，52(S1)：237-244.
ZHANG Lu, HUANG Lin, LI Beibei, et al. Fish school counting method based on multi-scale fusion and no anchor YOLO v3[J]. Transactions of the Chinese Society of Agricultural Machinery, 2021, 52(S1): 237-244. (in Chinese with English abstract)
[31] RASMUSSEN C, ZHAO J, FERRARO D, et al. Deep census: AUV-based scallop population monitoring[C]//Proceedings of the 2017 International Conference on Computer Vision Workshops. IEEE, Venice, Italy, 2017:2865-2873.
[32] WANG CY, BOCHKOVSKIY A, LIAO HY. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[J/OL]. (2022-07-06)[2022-10-12]. http://arXiv.ORG/ABS/2207.02696.
[33] ZHANG Y, SUN P, JIANG Y, et al. Bytetrack: Multi-object tracking by associating every detection box[C]//European Conference on Computer Vision, Tel Aviv, Israel, 2022: 1-21.
[34] DALIAN UNIVERSITY OF TECHNOLOGY. 2017 Underwater Robot Picking Contest, URPC2017[EB/OL]. (2017-08-21) [2022-10-12]. http://2017.cnurpc.org/a/js/2017/0829/66.html.
[35] ULTRALYTICS. YOLOv5[EB/OL]. (2020-06-26)[2021]. https://github.com/ultralytics/yolov5.
[36] WOJKE N, BEWLEY A, PAULUS D. Simple online and realtime tracking with a deep association metric[C]//2017 IEEE International Conference on Image Processing (ICIP). IEEE, Beijing, China, 2017: 3645-3649.

Multi-category sea foods counting method integrating YOLOv7 and BYTE multi-target tracking

AN Zhiqiang[1,2,3], LI Zhijun[1], LIU Shuo[1,2,3], ZHAO Yonggang[4], CHEN Qijun[4], ZUO Rantao[5], LIN Yuanshan[1,2,3]※

(1. *School of Information Engineering*, *Dalian Ocean University*, *Dalian* 116023, *China*;　2. *Key Laboratory of Marine Information Technology of Liaoning Province*, *Dalian* 116023, *China*;　3. *Key Laboratory of Environment Controlled Aquaculture (KLECA)*, *Ministry of Education*, *Dalian Ocean University*, *Dalian* 116023, *China*;　4. *Dalian Xinyulong Marine Biological Seed Industry Technology Co.*, *Ltd*, *Dalian* 116007, *China*;　5. *School of Fisheries and Life Sciences*, *Dalian Ocean University*, *Dalian* 116023, *China*)

Abstract: Accurate and rapid counting has been one of the most important steps on sea foods in modern aquaculture. Taking the sea foods in a real bottom-sowing aquaculture environment as the research object, this study aims to guarantee the cost-saving, highly efficient and easy counting on the multi-category sea foods using video multi-target tracking. The underwater videos of sea foods were taken as the data source. Firstly, the YOLOv7 algorithm with excellent performance was used to realize the target detector of sea treasure, providing the input for the multi-target tracking. Specifically, the image data was captured under real underwater using the "2017 Underwater Robot Picking Contest, URPC2017", and then used to train the YOLOv7 model. 17,655 images contained the holothurians, echinus, and scallops. The images were randomly divided into the training set of 14055 images in the ratio of 8:1:1, the verification set of 1800 images, and the test set of 1800 images. The training setting of the YOLOv7 detector was set as the image adaptive size of 640*640, the initial learning rate of 0.01, the momentum of 0.9, the weight decay of 0.0005, the batch size of 16, the number of training rounds of 300. One test was performed every 10 rounds of training round. The operating system was selected the Ubuntu18.04 for Haizhenpin detector training, the deep learning framework was PyTorch, the experimental processor was AMD Ryzen Threadripper 1920X 12 cores, and the graphics card was NVIDIA GeForce RTX2080. Secondly, the characteristics of high similarity were combined with the unclear appearance of the same type of sea foods in the real breeding environment. The multi-target tracking of the BYTE algorithm was used for the reference. A multi-category trajectory generation and a counting strategy were then designed for the sea treasure tracking using the trajectory ID number. Finally, an optimal combination of the indicators was also evaluated on the performance of the improved model. The test results show that the average counting precision, mean absolute error (MAE), root mean square error (RMSE), and frame rate were 91.62%, 5.75, 6.38, and 32 frames/s, respectively. All indicators were better than those of the current YOLOv5+DeepSORT, YOLOv7+DeepSORT, YOLOv5+BYTE, and YOLOv7+BYTE. Especially, the average counting accuracy and frame rate index ratio of YOLOv5+DeepSORT were improved by 29.51, and 28 percentage points than before. The MAE and RMSE of the improved model were reduced by 19.50 and 12.08, respectively. The quantity of underwater sea foods was effectively measured in the modern fishery. The finding can provide the technical reference for the production measurement and the scientific decision-making on the intelligent management of aquaculture. In addition, the underwater dataset was used to train the detection model under the same environment, in order to reduce the false detections caused by different conditions. Nevertheless, the underwater environment changes greatly, due to various influencing factors, such as lighting. Statistical counting was then performed to collect the video of marine treasures under the same environment every time during training. Therefore, the current model can be expected to improve on the different underwater environments in the future.

Keywords: machine vision; deep learning; sea foods counting; aquaculture; multi-target tracking