doi:10.6041/j.issn.1000-1298.2025.09.057

基于轻量化 YOLO v8 和 BoT－SORT 的石斑鱼跟踪方法

段青玲[1,2]　乔雅琪[1,3]　刘怡然[1,2]　冯晓晓[1,3]　冉　逊[1,2]　刘春红[1,2]

（1. 中国农业大学国家数字渔业创新中心，北京 100083；2. 中国农业大学信息与电气工程学院，北京 100083；
3. 中国农业大学烟台研究院，烟台 264670）

摘要：水产养殖中，鱼类跟踪是实现鱼类行为监测、水质异常报警、鱼类生长状况评估的基础，但现有方法存在计算耗时长、模型占用空间大、在边缘端设备部署困难等问题。针对上述问题，本文以石斑鱼为研究对象，提出一种基于轻量化 YOLO v8 与 BoT－SORT 的石斑鱼跟踪方法，该方法包括目标检测和目标跟踪两个阶段。在目标检测中，采用 YOLO v8m 作为基线网络，引入卷积模块 FasterConv 以减少参数量；加入 EMA（Excitation and modulation attention）机制以保持模型精度；使用多尺度特征融合模块 Fusion 并调整 Neck 网络结构以提高模型的特征融合能力。在目标跟踪部分，BoT－SORT 算法简化了鱼体的运动状态变量，加入相机运动补偿（Camera motion compensation，CMC）以应对鱼体外观剧烈变化，最后利用 ResNeST50 网络提取较高置信度检测框内鱼体的外观特征，实现了鱼体跟踪。在自建的石斑鱼数据集上进行了训练和验证，目标检测模型 mAP@0.5 为 95.80%；其模型内存占用量为 23.7 MB，相较原始 YOLO v8m 模型降低 54.42%；将本文的轻量化目标检测模型应用到 BoT－SORT 算法，MOTA 为 78.774%，FPS 达到 28.20 f/s，在对比实验中综合性能大幅超过 SORT、DeepMoT 等算法。本方法可以实现石斑鱼的检测与跟踪，为石斑鱼的养殖提供技术支撑。

关键词：石斑鱼；目标检测；目标跟踪；YOLO v8；BoT－SORT；轻量化

中图分类号：TP391.41；S965.334　　**文献标识码**：A　　**文章编号**：1000-1298(2025)09-0667-10　　**OSID**：

Grouper Fish Tracking Method Based on Lightweight YOLO v8 and BoT－SORT

DUAN Qingling[1,2]　QIAO Yaqi[1,3]　LIU Yiran[1,2]　FENG Xiaoxiao[1,3]　RAN Xun[1,2]　LIU Chunhong[1,2]

（1. *National Innovation Center for Digital Fishery*, *China Agricultural University*, *Beijing* 100083, *China*
2. *College of Information and Electrical Engineering*, *China Agricultural University*, *Beijing* 100083, *China*
3. *Yantai Institute*, *China Agricultural University*, *Yantai* 264670, *China*）

Abstract: In aquaculture, fish tracking is fundamental for monitoring fish behavior, detecting water quality anomalies, and assessing the growth conditions of fish. However, existing methods suffer from issues such as computational time consumption, large model size, and difficulties in deploying on edge devices. To address these challenges, taking grouper fish as the research subject, a tracking method was proposed based on a lightweight YOLO v8 and BoT－SORT. This method consisted of two stages: target detection and target tracking. In the target detection phase, YOLO v8m was used as the baseline network. Firstly, a convolutional module, FasterConv, was introduced to reduce the number of parameters. Then the excitation and modulation attention (EMA) mechanism was incorporated to maintain model accuracy. Finally, a multi-scale feature fusion module, Fusion, was employed and the structure of the Neck network was adjusted to enhance the model's capability for feature integration. For the target tracking part, the BoT－SORT algorithm simplified the motion state variables of the fish body and included camera motion compensation (CMC) to deal with drastic changes in fish appearance. Subsequently, the ResNeST50 network was utilized to extract the appearance features of the fish within high-confidence detection bounding boxes, achieving fish tracking. The method was trained and validated on a self-built grouper dataset, achieving a mAP@0.5 of 95.80% for the object detection model, with a model size of 23.7 MB—a 54.42% reduction compared to that of the original YOLO v8m model. When

收稿日期：2024－03－09　修回日期：2024－05－14

基金项目：国家重点研发计划项目（2022YFD2001701）和重庆市技术创新与应用发展专项重点项目（CSTB2022TIAD－ZXX0053）

作者简介：段青玲（1967—），女，教授，博士生导师，主要从事智能信息处理研究，E-mail：dqling@cau.edu.cn

applying the lightweight object detection model to the BoT - SORT algorithm, the MOTA reached 78.774%, with an FPS of 28.20 f/s, significantly outperforming SORT, DeepMoT, and other algorithms in comparative experiments. The results demonstrated that this method can achieve detection and tracking of grouper fish, providing technical support for the cultivation of groupers.

Key words: grouper fish; object detection; object tracking; YOLO v8; BoT - SORT; lightweight

0 引言

石斑鱼因其美味和营养备受消费者喜爱,然而,养殖过程中会出现鱼体游动迟缓、胀气失衡以及低温贴壁等异常现象,难以被工人及时发现,这影响了养殖效率[1-2]。在鱼类健康监测中,研究鱼类行为特征,能够有效预警鱼类异常行为或疾病行为[3-5];另外,通过分析鱼体游动模式评估其适应性,可为水环境管理提供科学依据[6-8]。基于计算机视觉的鱼类跟踪技术能够实现连续无损监测,从而及时发现问题、降低风险并提升养殖效益。因此,研究鱼类跟踪方法,实时监测其行为状况和游泳模式,对提升水产养殖效率具有重要意义。

计算机视觉技术具有高效、无损和可扩充性等优势,被广泛应用在水产养殖的目标跟踪中。目标跟踪是指持续监测视频或图像序列并分析监测目标位置和运动的过程,其能够获取监测目标的速度、加速度及轨迹等关键参数[9]。目前,基于目标检测的跟踪方法是多目标跟踪的主流算法。该方法首先使用目标检测技术分析图像特征,并根据这些特征确定目标种类和位置,接着将检测到鱼体检测框传递给跟踪算法,跟踪算法利用视频帧之间鱼体的位置与特征相似性对目标鱼体进行跟踪[10]。

目标检测是目标跟踪的基础,按照前向传播次数分为单阶段和双阶段两种。单阶段目标检测算法主要包括 YOLO、SSD 与 RetinaNet 等,在同一次前向传播中同时完成目标检测和分类;双阶段目标检测算法主要包括 Faster R - CNN、Cascade R - CNN 等,通常在第一阶段生成候选目标区域,第二阶段对候选区域进行分类[11]。

基于 YOLO 系列和 Faster R - CNN 的水下目标检测方法,通过结合图像增强技术、改进网络结构或优化特征映射,有效提升了检测精度和效率,但两种目标检测算法均需要较大的计算资源和内存占用量,但水产养殖检测模型常部署于资源受限的边缘端设备,这导致目标检测算法难以应用于生产实践[12-16]。

目标跟踪根据跟踪对象的数量分为单目标跟踪(Single-object tracking, SOT)和多目标跟踪(Multi-object tracking, MOT),多目标跟踪对视频中多个目标进行连续识别和跟踪,算法复杂性较高、鲁棒性强且应用场景广泛。在基于目标检测的跟踪方法中,目标检测模型的精确度和轻量化程度对模型整体性能有重要影响,优质的检测算法可显著提升跟踪方法性能[17-21]。

针对基于检测的目标跟踪模型计算耗时长、所占空间大、难以部署边缘端等问题,本文以石斑鱼为对象,提出基于轻量化 YOLO v8 和 BoT - SORT 的石斑鱼跟踪方法。在目标检测方法中,对基线网络 YOLO v8m 进行轻量化改进,将改进后的目标检测方法应用于多目标跟踪算法 BoT - SORT[22],以期在模型轻量化的同时保证精度。

1 数据采集与预处理

1.1 数据采集

实验平台在明波水产养殖基地搭建,该平台包括养殖系统和数据采集装置,平台结构和实景如图 1 所示。养殖系统拥有一个半径为 0.7 m、深约为 0.4 m 的养殖池,并配有外海水供应装置、水质参数检测装置和换水排便装置。摄像头安装在养殖池正上方 1.2 m 处,拍摄的视频通过路由器传输到便携式计算机进行存储。采集视频所使用的摄像头型号为 Hikvision 2CD3T86FWDV2 - I3S,所采集的视频分辨率为 1 920 像素 × 1 080 像素,帧率为 30 f/s。

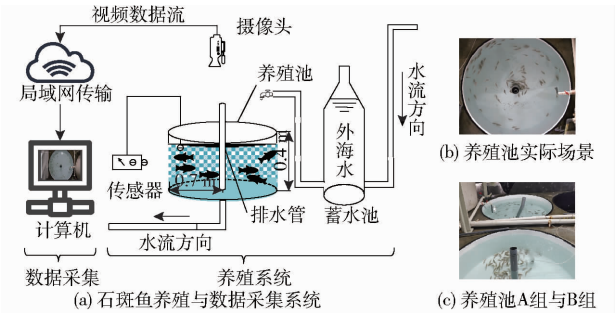


图 1 数据采集系统与实际养殖场景

Fig. 1 Data collection system and actual breeding scenarios

为了模拟实际养殖车间的多种场景,设计包括正常养殖、改变养殖密度、改变溶解氧浓度和改变光照的 4 组实验。通过改变曝气石、水池上方的养殖灯和遮光布以及捞出鱼体等调整养殖条件。使用 200 条平均质量约为 10 g 的石斑鱼进行实验,将其分为 A 组和 B 组,每组 100 条。每组实验结束后,对石斑鱼进行 3 d 以上正常饲养以消除上次实验的影响。数据采集说明如表 1 所示。

表 1　数据集采集说明

Tab. 1　Dataset collection instructions

组别	溶解氧质量浓度/(mg·L^{-1})	养殖密度/(条·池$^{-1}$)	光照	具体胁迫措施	采集视频时长/s
1	6.0～7.0	100	自然光	无胁迫,定时定量投喂	30
2	6.0～7.0	25、50、75、100、125、150、175、200	自然光	在 A 组和 B 组养殖不同数量鱼体	190
3	2.7～7.0	100	自然光	从养殖池取出和放入曝气石	50
4	6.0～7.0	100	黑暗、强光、自然光	添加养殖灯和搭建遮光布	80

1.2　数据处理

数据集制作分为目标检测数据集和目标跟踪验证集两部分。目标检测数据集包含 4 种实验场景，将视频拆解成帧后选取原始图像 523 幅，使用 Labelimg 工具标注为 YOLO 格式，并将所有图像尺寸由1 920 像素×1 080 像素处理为 640 像素×640 像素。为了防止数据量不足导致的过拟合，使用 OpenCV 工具对目标检测的训练集进行增强，将图像按照向右旋转 15°、向左旋转 15°、调整亮度和饱和度、增减曝光度以及添加噪点的方法进行处理。完成训练集扩充后，目标检测数据集图像共 1 810 幅，其中训练集1 448 幅，测试集 181 幅，验证集 181 幅，三者比例为 8∶1∶1，测试集和训练集为未经增强的原始图像。目标跟踪验证集仅用于评估算法性能，选择一段 4 s 视频，使用 DarkLabel 工具按照 MOT 格式进行标注，该验证集共 120 幅。

2　石斑鱼检测与跟踪方法

2.1　总体流程

在工厂化鱼类养殖中，为降低成本和远程监控而将模型部署在资源受限的边缘端，但大型跟踪模型难以适配。为此，本文提出一种基于轻量化 YOLO v8 与 BoT－SORT 的石斑鱼跟踪方法，该方法分为目标检测和目标跟踪两部分，流程如图 2 所示。YOLO v8 原模型按照规模分为 YOLO v8n、YOLO v8s、YOLO v8m 等，网络规模越大，训练后模型的检测精度越高。为保证对石斑鱼的检测精度，以

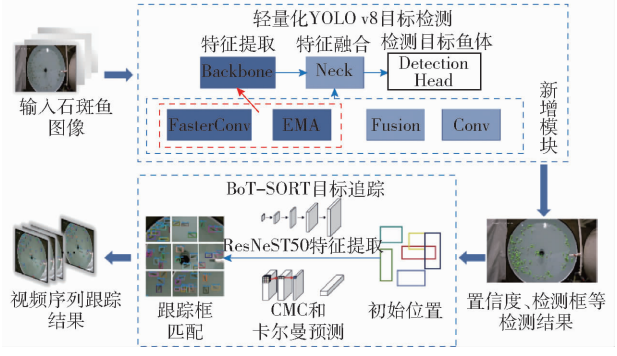


图 2　检测与跟踪方法整体流程图

Fig. 2　Overall process diagram of detection and tracking methods

YOLO v8m 为目标检测基线网络，首先对其进行轻量化改进，训练完成后，目标检测模型能够对输入图像序列进行检测，接着带有置信度等的检测结果被传递给 BoT－SORT 算法，从而逐帧对石斑鱼视频序列进行跟踪。在目标检测部分，主要目的是获得视频帧中的待检测鱼体。Backbone 网络引入 FasterConv 模块和 EMA 机制以完成对鱼体特征提取；Neck 网络除了加入 FasterConv 和 EMA，还增加特征融合模块 Fusion 和卷积模块 Conv，轻量化 Neck 网络在特征融合时对细小特征更加敏感；最后检测头将检测结果输出。在目标跟踪部分，逐帧预测出石斑鱼的位置进行跟踪与匹配。BoT－SORT 跟踪算法得到目标检测结果后，使用相运动补偿(CMC)和 ResNeST50 网络等得到预测框，完成位置匹配并输出结果。

2.2　基于轻量化 YOLO v8 的目标检测方法

在实际养殖中，石斑鱼幼鱼外观相似度高、体积小且游动中鱼体常发生重叠和遮挡现象，难以被识别与跟踪。为此，本文提出轻量化 YOLO v8 石斑鱼目标检测模型，使得模型体积更小并改进了其对细小鱼体特征的提取能力，模型结构如图 3 所示。在该模型中，采用 CFE(C2f_FasterConv_EMA)替代原模型中的 C2f(Channel-to-Pixel Feature)完成特征提取；引入 Fusion 模块完成特征融合并调整 Neck 网络结构以增强其对小尺度特征敏感度。轻量化目标检测模型的改进目标是体积更小且保持较高精度，使模型更加适用于实际养殖。

2.2.1　CFE 轻量化特征提取

在原 YOLO v8m 网络中，完成图像特征提取的 C2f 模块使用较多卷积，在池化操作后通道数翻倍，所以该模型参数量多、计算耗时长。本文使用轻量化模块 CFE 替换 C2f 进行图像特征提取，CFE 模块由模块 C2f、FasterConv[23] 和 EMA[24] 等构成。

FasterConv 能够轻量化的原因是其中的 PConv(Partial Convolution)模块仅用 1/4 输入进行卷积和过滤计算，提取特征后的结果与原来未进行计算的 3/4 输入连接作为输出。FasterConv 的引进降低了模型的参数量和内存访问次数，但 PConv

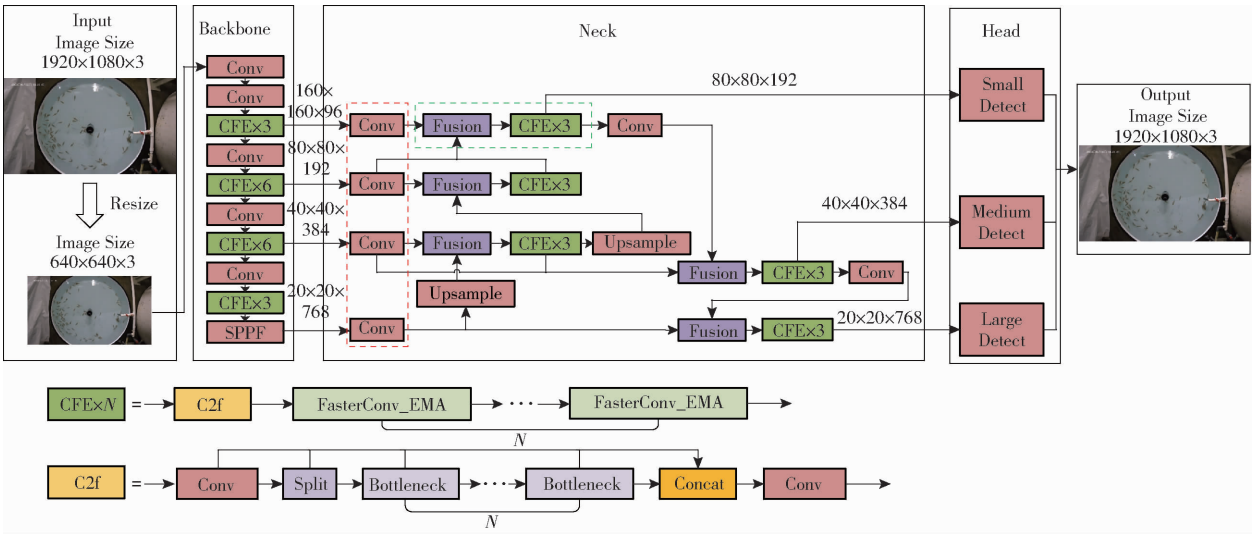


图 3 轻量化 YOLO v8 模型结构图

Fig. 3 Structure diagram of lightweight YOLO v8 model

减少了提取特征的通道,造成模型检测精度下降。为了收集多尺度的特征信息,在 FasterConv 模块中引入 EMA 机制,EMA 模块的并行子网络结构和分组策略可以高效处理图像细节,通过融合全局信息完成对多尺度特征的学习,EMA 工作流程如图 4 所示。

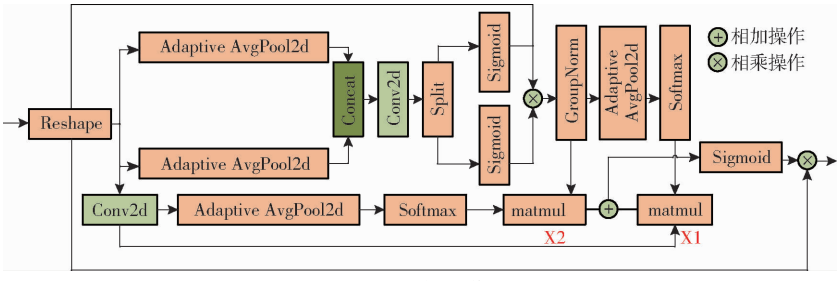


图 4 EMA 机制工作流程图

Fig. 4 EMA attention mechanism workflow

基于 PConv 和 EMA 构建 FasterConv_EMA 模块,其工作流程如图 5 所示。首先 PConv 模块提取特征,接着通过由一维卷积和二维卷积构成的多层感知层进行计算,再通过 EMA 模块感知层的输出结果进一步提取特征,将最开始的输入与 EMA 模块的输出相加得到最终输出。

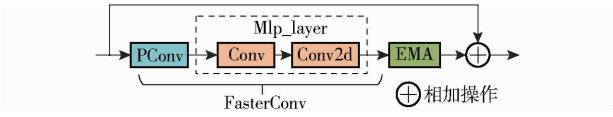


图 5 FasterConv_EMA 工作流程图

Fig. 5 Schematic of FasterConv_EMA workflow

本文提出 CFE 模块,代替 C2f 进行特征提取。根据 C2f 迭代次数在其后添加 N 个 FasterConv_EMA 模块,从而得到 CFE 模块。YOLO v8m 的 Backbone 网络和 Neck 网络通过引入 CFE 模块,可以降低参数量,同时保证了模型对图像特征的提取能力。

2.2.2 Neck 网络优化

(1) Fusion 特征融合

为使模型更大程度轻量化并提升其特征融合能力,本文引入 Fusion 模块来替换 Neck 部分的 Concat 模块。BiFPN 网络[25]的 Fusion 模块使用加权的方式进行计算,这种方式能够智能选择重要特征,避免直接拼接特征造成的信息冗余和额外计算。Fusion 模块的主要作用是计算融合的特征并输出,其加权特征融合机制使得 Neck 网络进一步轻量化,并且特征融合性能得到进一步改进。计算公式为

$$W = \frac{\mathrm{ReLU}(W_{init}\mathrm{clone}())}{\sum W_{init}} \tag{1}$$

$$F = \sum (W_i \boldsymbol{X}_i) \tag{2}$$

式中 ReLU——激活函数

W_{init}——原始值

clone()——Pytorch 框架中基础函数,在本文起张量复制作用

W——需要计算的融合权重

$\boldsymbol{X}_i$——输入特征向量

F——最终的特征融合结果

i——权重计算次数

(2)Neck 网络结构调整

养殖池中目标鱼体积较小，可以提取的特征有限，但原 Neck 网络特征融合模块少且对细小特征不敏感。本文对 Neck 网络结构进行调整，改进前后对比如图 6 所示。在轻量化改进上，用 CFE 模块替换 C2f 进行特征融合，接着引入 Fusion 模块替代 Concat 进行特征融合。达到轻量化目标后，为了改善原网络对细小特征的识别能力，首先在图 6b 的红色虚线框内加入 4 个卷积层以调整通道数，因为第 1 个卷积层位于较浅的网络层次，所以在绿色虚线框中加入 Fusion 模块和 CFE 模块进行特征提取和融合的成本并不高，通过引入绿色框内的模块，模型对 8×8 细微特征的识别能力得到了增强。

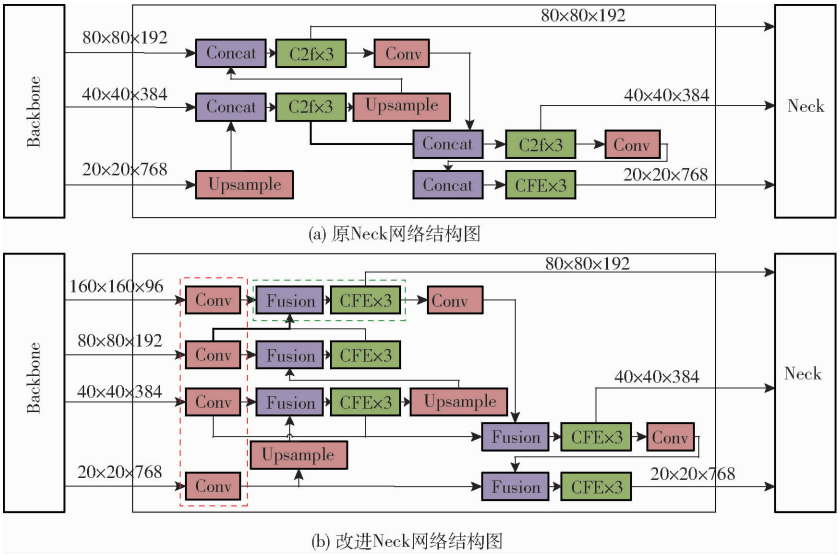


图 6　Neck 网络改进前后对比

Fig. 6　Comparison of Neck network before and after improvement

改进后的 Backbone 网络对图像分别进行 64、32、16、8 次下采样，通过步长为 2 的卷积运算提取图像的特征，提取的特征图宽×高×通道数为160×160×96、80×80×192、40×40×384 和 20×20×768。绿色虚线框通过 Fusion 模块与 3 次迭代的 CFE 模块后，输出作为 8×8 小目标检测头的输入，与其他模块输出配合完成不同尺寸特征的融合。

2.3　BoT－SORT 多目标跟踪

BoT－SORT 与 SORT 算法均是对卡尔曼滤波的改进[26－27]，该算法能够提取目标鱼体之外的背景关键点，从而实现更加精确的跟踪，算法流程如图 7 所示。

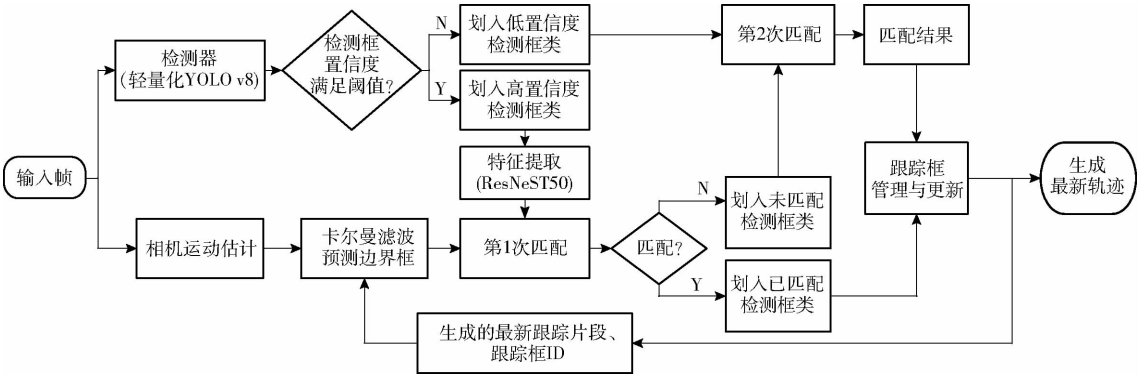


图 7　BoT－SORT 算法流程图

Fig. 7　Algorithm flowchart of BoT－SORT

在每一帧的跟踪中，首先将目标检测模型获取的检测框按置信度分为高、低两组，对高置信度框提取特征并与经过相机运动估计和卡尔曼预测的边界框匹配，再次分组得到匹配和未匹配跟踪框。未匹配跟踪框再与低置信度框匹配，输出最终匹配结果和最新轨迹。BoT－SORT 相较 SORT 等算法主要有如下改进：

(1)简化了运动状态变量。在轨迹状态预测方面，SORT 算法中跟踪目标状态变量的 7 个维度信息是$(x,y,s,a,\dot{x},\dot{y},\dot{s})$，分别表示检测框的中心点 x 坐标、中心点 y 坐标、面积、高宽比、x 轴速度、y 轴速度和面积变化速率。在 BoT－SORT 中，将状态变量

简化为检测框的中心点横纵坐标 x、y 和检测框长宽 h、w，以及其对时间的导数，即(x, y, w, h, $\dot{x}$, $\dot{y}$, $\dot{w}$, $\dot{h}$)。

(2)针对跟踪目标外观剧烈改变的问题加入相机运动补偿(CMC)。在石斑鱼运动过程中，短时间内鱼体形状可能发生剧烈改变，仅依靠颜色等外观信息无法有效进行跟踪，BoT-SORT 算法通过加入相机运动补偿(CMC)来投射仿射矩阵，根据相邻两帧背景图像的相似度来提取目标之外的画面关键点，从而完成更加精确的跟踪。仿射矩阵被定义为

$$\boldsymbol{A} = [\boldsymbol{M}, \boldsymbol{T}] \in \mathbf{R}^{2\times 3} \tag{3}$$

式中 $\boldsymbol{A}$——仿射矩阵

$\boldsymbol{M}$——描述二维平面缩放和旋转变化的矩阵

$\boldsymbol{T}$——平移变化向量

为了用一次矩阵乘法完成对卡尔曼滤波检测后状态量中心点和高宽的仿射变换，构造公式进行卡尔曼滤波计算，即

$$\boldsymbol{M}'_{k|k-1} = \mathrm{diag}\{\boldsymbol{M}, \boldsymbol{M}, \boldsymbol{M}, \boldsymbol{M}\} \in \mathbf{R}^{8\times 8} \tag{4}$$

$$\boldsymbol{T}'_{k|k-1} = \mathrm{diag}\{\boldsymbol{T}, 0, 0, 0, 0, 0, 0, 0\} \in \mathbf{R}^{8} \tag{5}$$

(3)在第 1 次相似度关联匹配时，借助 ResNeST50 网络得到外观特征并进行高置信度框的检测。由于高置信度检测框内遮挡、模糊等干扰通常较少，便于提取鱼体特征，因此，BoT-SORT 跟踪算法在第 1 次相似度匹配时，只匹配高置信度检测框。匹配时使用 ResNeST50 网络提取外观特征，并用指数移动平均(Exponential moving average)机制来平衡过去和当前的外观特征，从而提高第 1 次相似度匹配的准确性。

2.4 实验设计与模型评估

2.4.1 实验设计

为了验证检测模型的性能，本文为轻量化 YOLO v8 目标检测模型设计了 3 组实验。第 1 组为目标检测消融实验，目的是逐项验证算法改进的有效性；第 2 组为目标检测对比实验，目的是验证本文方法与其他目标检测模型在同一数据集上的检测效果；第 3 组为目标跟踪的对比实验，目的是对比验证 BoT-SORT 算法与其他跟踪算法使用本文的轻量化 YOLO v8 目标检测模型进行目标跟踪的性能。

在本实验中，搭建了基于 Ubuntu 20.04、Python 3.8 和 Pytorch 1.11.0 的深度学习框架。为该框架配备型号为 Intel(R) Xeon(R) Platinum 8255C 的 CPU、型号为 Nvidia GeForce RTX 3080(10 GB)的 GPU 以及计算统一设备架构 CUDA 11.3。

2.4.2 目标检测评价指标

在目标检测实验中，对不同模型的超参数进行统一设置：训练次数为 300，批大小为 8，动量为 0.937，权值衰减为 0.000 5，初始学习率为 0.01，采用余弦退火策略来降低学习率。采用精确率(P)、召回率(R)、平均精度均值(mAP@0.5)、模型内存占用量、模型参数量、浮点运算量等指标来评价石斑鱼目标检测模型性能。

2.4.3 目标跟踪评价指标

在多目标跟踪评估方面，采用 TrackEval 工具对 BoT-SORT 进行性能评估，指标分别为跟踪准确度(MOTA)、跟踪精度(MOTP)、ID 切换数(IDSW)和跟踪调和平均数(IDF1)、跟踪视频帧率(FPS)、轨迹命中数(MT)、轨迹被打断总数量(Frag)。其中，MOTA 是对模型检测与跟踪性能的整体评估，MOTP 是衡量跟踪目标位置误差的一个相对指标，IDF1 是对跟踪模型稳定性的衡量。

3 实验结果和分析

3.1 目标检测模型消融实验

为了验证每项改进的有效性，使用石斑鱼游泳数据集进行消融实验。以 YOLO v8m 为基线网络，分别在基准网络中加入 FasterConv 模块、EMA 机制以及对 Neck 网络进行优化，得到 3 个改进模型。各改进模型实验结果见表 3。

表 3 目标检测模型消融实验结果

Tab.3 Ablation experimental results of target detection model

模型	P/%	R/%	mAP@0.5/%	参数量	浮点运算量	模型内存占用量/MB
YOLO v8m	93.4	93.0	96.5	2.58×10^{7}	7.87×10^{10}	52.0
YOLO v8m + FasterConv	91.9	91.6	95.7	1.66×10^{7}	4.98×10^{10}	33.6
YOLO v8m + FasterConv + EMA	92.4	91.1	95.7	1.67×10^{7}	5.23×10^{10}	33.9
YOLO v8m + FasterConv + EMA + Neck 网络优化	92.1	92.2	95.8	1.16×10^{7}	4.74×10^{10}	23.7

如表 3 所示，YOLO v8m 引入 FasterConv 模块后，训练得到的模型精确率、召回率以及 mAP@0.5 均有所下降，但参数量、浮点运算量和模型内存占用量也均有明显下降，轻量化效果明显；在 YOLO v8m 中引入 FasterConv 的基础上加入 EMA 机制，训练后模型精确率得到提升，相比仅引入 FasterConv 上升 0.5 个百分点；最后，在 YOLO v8m 中同时使用 FasterConv、EMA 和 Neck 网络优化，相比 YOLO v8m

＋FasterConv＋EMA 召回率提升 1.1 个百分点，同时参数量、浮点运算量以及模型内存占用量再次有明显下降，相比原 YOLO v8m 模型分别下降 1.42×10^{7}、3.13×10^{10}和 28.3 MB（比原模型下降 55.04%、39.77%和 54.42%）。本文方法对模型的轻量化有较明显效果，整体模型内存占用量仅为 23.7 MB，且该模型相比 YOLO v8m 精确率、召回率和 mAP@0.5 分别只下降 1.3、0.8、0.7 个百分点，说明该模型在基本满足精度要求的情况下，也达到了显著压缩模型的效果。

选择简单与复杂两种场景图像进行检测比较，复杂场景鱼体密度较大且存在鱼体倒影，各模型的测试输出图像如图 8 所示，其中黄框内为漏检鱼体、红框内为被误检的鱼体倒影。在简单场景的检测中，4 种模型均能正确检测大部分鱼体，只有 YOLO v8m＋FasterConv 和 YOLO v8m＋FasterConv＋EMA 分别在黄框处出现了 1 次漏检现象。在复杂场景检测中，YOLO v8m 分别出现了 2 处漏检和 5 处误检；在 YOLO v8m＋FasterConv 中，多处黄框内鱼体被漏检且红框内有 5 处倒影被误检为鱼体；在 YOLO v8m＋FasterConv＋EMA 中，误检鱼体数量下降但漏检鱼体数量上升，这符合消融实验中模型精度上升但召回率下降的结果；在 YOLO v8m＋FasterConv＋EMA＋Neck 网络优化中，黄框内漏检的鱼体明显减少且倒影被误检成鱼的红框数量也少于 YOLO v8m。说明通过逐步改进，模型的检测效果逐渐提升并接近原水平。实验结果表明，本文的改进促进了轻量级模型的实现，在保证模型精度的同时降低了对设备计算和存储的要求。

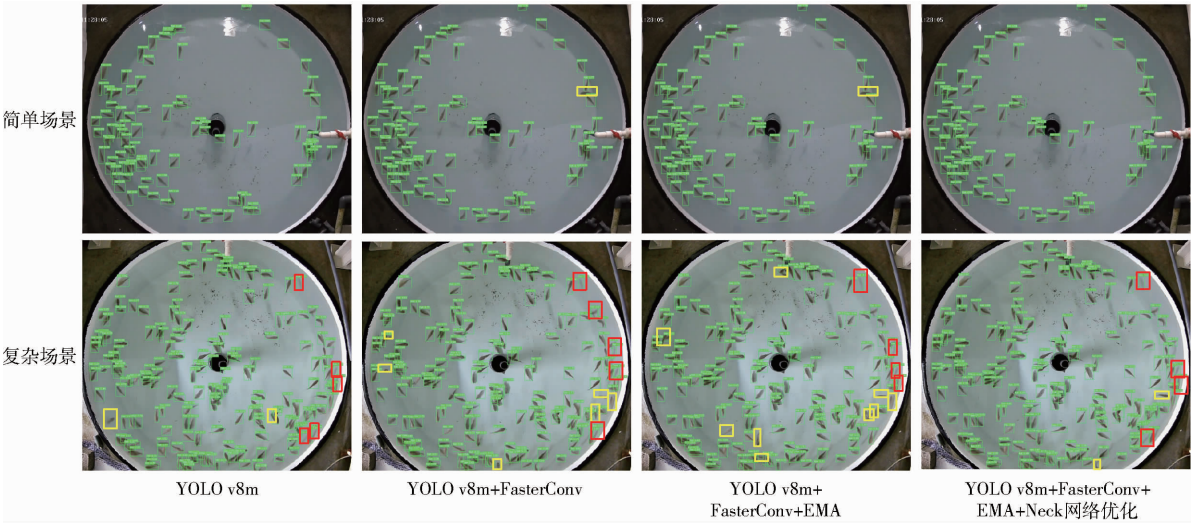


图 8　消融实验检测结果

Fig. 8　Results of ablation experiment detection

3.2　目标检测模型对比实验

为比较本文目标检测模型的性能，本文将该模型在石斑鱼目标检测数据集中进行训练和测试，并在使用同样数据集情况下，选用 YOLO v5m、DAMO－YOLO－M、YOLO v7x 与本文改进后的 YOLO v8m 进行对比，得到的实验结果如表 4 所示。

表 4　目标检测模型对比实验结果

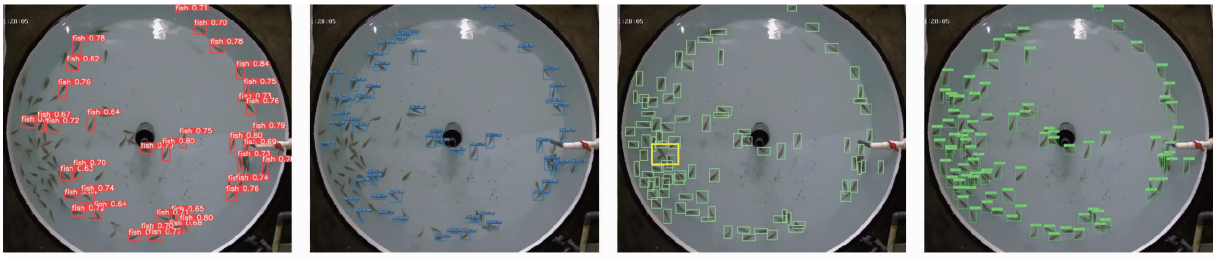
Tab. 4　Comparison experiment results of target detection models

模型	P/%	R/%	mAP@0.5/%	参数量	浮点运算量	模型内存占用量/MB
YOLO v5m	95.2	82.6	89.3	2.09×10^{7}	4.790×10^{10}	42.1
DAMO－YOLO－M			94.1	2.75×10^{7}	5.985×10^{10}	65.6
YOLO v7x	95.5	93.5	94.8	7.08×10^{7}	1.880×10^{11}	142.1
本文方法	92.1	92.2	95.8	1.16×10^{7}	4.740×10^{10}	23.7

如表 4 所示，YOLO v5m 的 mAP@0.5 低于本文方法 6.5 个百分点，同时本文模型内存占用量只有 YOLO v5m 的 56.29%，说明本文方法在精度和轻量化程度上均优于 YOLO v5m；DAMO－YOLO－M 的 mAP@0.5 低于本文方法 1.7 个百分点，但参数量、浮点运算量和模型内存占用量均超过本文方法；相比 YOLO v7x，本文方法在 mAP@0.5 高于 YOLO v7x 的情况下，参数量、浮点运算量和模型内存占用量只占 YOLO v7x 的 16.38%、25.21%和 16.68%；与 YOLO v7x、DAMO－YOLO－M 和 YOLO v5m 相

比,本文模型在石斑鱼目标检测数据集上的 mAP@0.5 提高 1.0、1.7、6.5 个百分点,同时模型内存占用量分别减少 118.4、41.9、18.4 MB。结果表明,本文方法在满足精度的情况下可以部署在边缘端,在实际生产中应用可能性较大。

选择同 1 幅测试图像,使用对比实验中训练好的模型进行检测,观察输出结果以对比不同模型的检测效果,各模型的测试输出图像如图 9 所示,黄框代表漏检鱼体。在 YOLO v5m 的检测结果中,在样例图像左下、中心以及右下鱼体交叠部分均出现漏检现象,且图像左下有大量漏检鱼体,这与其模型低召回率、高精度的特点相符合;在 DAMO-YOLO-M 的检测结果中,样例图像中心与右下鱼体交叠处的漏检现象有所改善,但图像左下角仍存在漏检现象;在 YOLO v7x 的测试样例中,检测效果显著改善,但在黄框内两条重叠的鱼体被漏检,检测精度仍待提高;在本文方法的测试结果中,未出现漏检和误检的现象,在 4 组实验中检测效果最好。



(a)YOLO v5m (b) DAMO-YOLO-M (c) YOLO v7x (d) 本文方法

图 9 对比实验检测结果

Fig. 9 Comparison experiment detection results

3.3 目标跟踪模型对比实验

从采集的数据中截取一段 4 s 的视频,在使用本文目标检测方法的基础上,分别用 BoT-SORT、UavMot、C_Biou、ByteTrack、DeepMoT、DeepSORT、SORT、StrongSORT 8 种方法进行跟踪,结果如表 5 所示。

表 5 目标跟踪模型对比实验结果

Tab. 5 Comparison experiment results of target tracking models

跟踪模型	MOTA/%	MOTP/%	IDF1/%	FPS/(f·s^{-1})	MT	IDSW	Frag
BoT-SORT	78.774	71.972	83.634	28.20	22	6	48
StrongSORT	59.245	73.555	55.225	9.97	8	51	73
ByteTrack	72.138	70.752	81.437	27.93	16	5	64
UavMot	69.057	70.753	79.347	18.23	16	7	112
DeepMoT	52.579	71.088	60.402	17.16	8	29	149
DeepSORT	59.906	71.881	49.008	14.29	9	70	248
SORT	63.836	72.061	66.839	12.49	11	29	87
C_Biou	68.270	73.149	71.372	14.66	10	19	81

BoT-SORT 的 MOTA 达到 78.774%,超过本组最低 MOTA 值 26.195%,在 8 种方法中精确度最高,综合性能最好。其他指标中,BoT-SORT 有 4 项达到最优,BoT-SORT 的 IDF1 评估结果为 83.634%,在本组跟踪算法中稳定性最好;BoT-SORT 的轨迹命中数 MT 和轨迹被打断次数 Frag 分别为 22 和 48,在 8 种算法中轨迹命中数最多和轨迹被打断次数最少;在跟踪速度上,BoT-SORT 检测速度最快,可达到 28.20 f/s。在 BoT-SORT 没有达到最优的指标中,其跟踪精度 MOTP 只比本组最高 MOTP 低 1.583 个百分点,跟踪精度不相上下;最后,在轨迹更换次数 IDSW 的评估结果对比中,BoT-SORT 跟踪的视频中所有鱼只发生了 6 次 ID 更换,本组最少 ID 更换 5 次,最多更换 70 次。结果表明,BoT-SORT 是本组对比实验中跟踪准确度、精度、速度和稳定性等方面效果最好的跟踪算法。

3.4 目标跟踪模型实验结果

采用跟踪对比实验中前 4 种结果较优的方法进行跟踪,对跟踪结果进行分析,对比跟踪结果如图 10 所示,其中红框表示跟踪丢失的鱼体,黄框表示检测遗漏的鱼体。在跟踪结果中,BoT-SORT 方法未出现检测遗漏和跟踪丢失的鱼体;ByteTrack 方法在第 T 帧左上角鱼体重叠处出现一条跟踪丢失的鱼体,并出现一条检测遗漏鱼体;但在 UavMot 方法中,第 T 帧出现两条跟踪丢失鱼体且出现一条检测遗漏的鱼体;StrongSORT 方法检测遗漏鱼体和跟踪丢失的鱼体在 4 种方法中最多,跟踪效果最差。结果表明,BoT-SORT 方法跟踪效果最好。

4 跟踪模型应用

基于本文方法,开发基于 Web 的鱼类跟踪与行为量化原型系统,界面如图 11 所示。该 Web 端应用程序前端使用 Ajax(Asynchronous JavaScript and XML)工具进行数据传输,后端程序基于 Flask 框架使用 PyCharm 工具编写,整个系统部署在一台 CPU 型号为 Intel(R)Xeon(R)Platinum 8255C 的计算机上。目前,可以通过上传视频到应用程序进行跟踪。

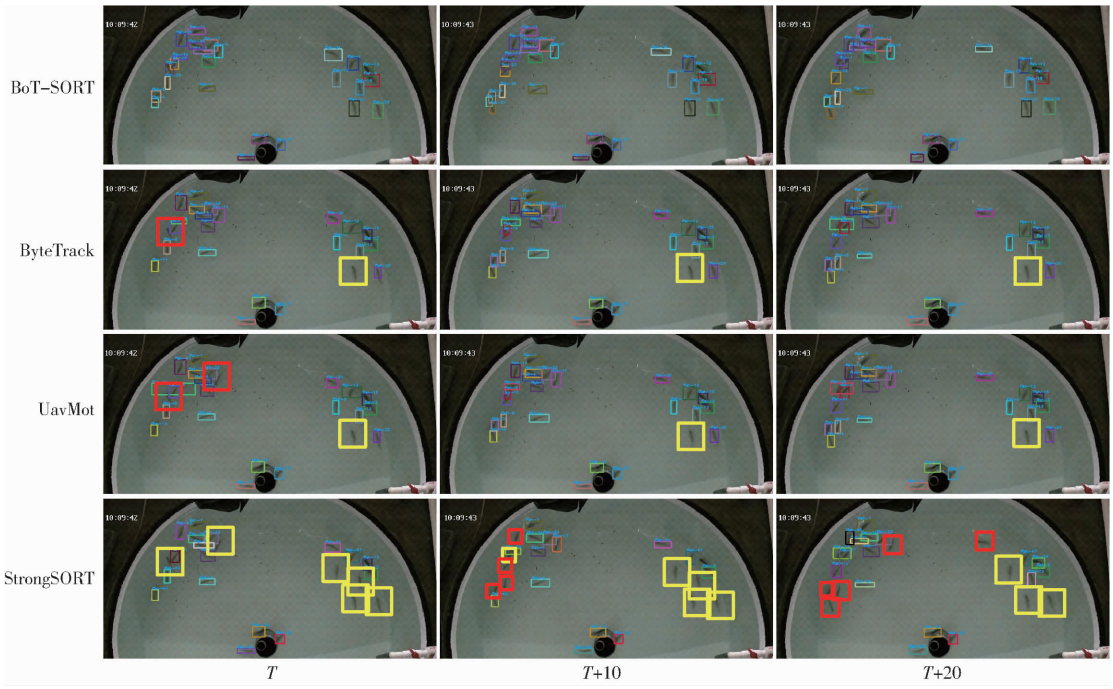


图 10　跟踪模型对比实验结果

Fig. 10　Comparison experiment results of tracking models

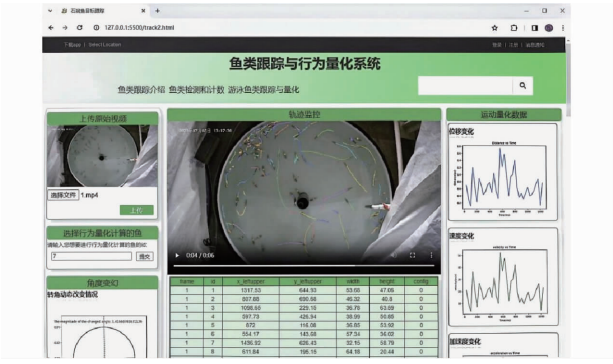


图 11　基于 Web 的鱼类跟踪与行为量化原型系统界面

Fig. 11　Homepage of a Web-based fish tracking and behavior quantification prototype system

系统在完成目标跟踪后，首先得到位置和轨迹等结果，接着计算得到单条鱼体的位移、速度、加速度、转角等行为量化信息的可视化结果。本文轻量化后的石斑鱼跟踪算法可以根据实际需求进行部署，以仅有 CPU 的计算机终端作为测试，对一段石斑鱼游泳视频进行目标跟踪，视频的处理速度（FPS）为 2.87 f/s，结果表明，本文方法能够应用于计算资源有限的边缘端设备并对石斑鱼进行有效跟踪。

5　结论

（1）针对工厂化养殖石斑鱼多种场景的准确性和效率要求，本文提出基于轻量化 YOLO v8 和 BoT－SORT的石斑鱼跟踪方法。在目标检测部分，用引入 EMA 机制的 FasterConv 模块对 C2f 卷积进行改造，并使用特征融合模块 Fusion 优化 Neck 网络。在目标跟踪部分，将本文的目标检测模型应用到多目标跟踪算法 BoT－SORT 中，实现了石斑鱼的高效检测与跟踪。

（2）为了验证模型的有效性，本文搭建石斑鱼养殖实验平台，采集视频数据并制作了石斑鱼数据集，在该数据集上进行实验。实验结果表明，目标检测模型 mAP@0.5 达到 95.80%，模型内存占用量为 23.7 MB，参数量为 1.16×10^7，模型规模较小且检测精度较高；目标跟踪模型 MOTA 达到 78.774%，FPS 达到 28.20 f/s，跟踪效果较好。

（3）将本文方法部署在计算机端的 Web 原型系统中，在边缘端算力受限情况下，跟踪速度下降到 2.87 f/s，但仍能完成对石斑鱼的高精度检测与快速跟踪，满足实际生产需求，能够为石斑鱼养殖提供技术支持。

参　考　文　献

[1] DAS S K, XIANG T W, NOOR N M, et al. Temperature physiology in grouper (Epinephelinae: Serranidae) aquaculture: a brief review[J]. Aquaculture Reports, 2021, 20: 100682.

[2] CONDINI M V, GARCÍA-CHARTON J A, GARCIA A M. A review of the biology, ecology, behavior and conservation status of the dusky grouper, *Epinephelus marginatus* (Lowe 1834)[J]. Reviews in Fish Biology and Fisheries, 2018, 28: 301－330.

[3] ABANGAN A S, KOPP D, FAILLETTAZ R. Artificial intelligence for fish behavior recognition may unlock fishing gear selectivity[J]. Frontiers in Marine Science, 2023, 10: 1010761.

[4] 李道亮, 姜国旗, 杨建安,等. 基于计算机视觉的鱼类低氧胁迫行为检测与跟踪算法[J]. 农业机械学报, 2023, 54(10):399 - 406.
LI Daoliang, JIANG Guoqi, YANG Jian'an, et al. Behavior detection and tracking algorithm for fish under hypoxic stress based on computer vision[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(10): 399 - 406. (in Chinese)

[5] YANG L, LIU Y, YU H, et al. Computer vision models in intelligent aquaculture with emphasis on fish detection and behavior analysis: a review[J]. Archives of Computational Methods in Engineering, 2021, 28: 2785 - 2816.

[6] LATZ N, SEEHAUSEN O, BAKKER T C M. The influence of environmental factors on fish behaviour: implications for aquaculture and fish welfare[J]. Aquaculture, 2019, 507: 674853.

[7] SÁNCHEZ-HERNÁNDEZ J M, MARTIN DEL CAMPO D, GARCÍA-GÁMEZ P. Computer vision systems for fish behaviour analysis in aquaculture: a review[J]. Aquaculture, 2018, 502: 100 - 110.

[8] LI Y, ZHOU G, MENG Q. Using computer vision to assess aquaculture water environments: a review[J]. Aquaculture, 2019, 508: 134 - 143.

[9] BOLTES A, SEBASTIAN T. Fish tracking in video sequences using a multi-camera setup: evaluation and comparison of different tracking algorithms[J]. Computers and Electronics in Agriculture, 2017, 145: 238 - 252.

[10] LI Z, YUAN Y. A review on deep learning for object detection[J]. Journal of Electronic Imaging, 2020, 31(2): 023006.

[11] 张阳婷, 黄德启, 王东伟,等. 基于深度学习的目标检测算法研究与应用综述[J]. 计算机工程与应用,2023,59(18): 1 - 13.
ZHANG Yangting, HUANG Deqi, WANG Dongwei, et al. Review on research and application of deep learning-based target detection algorithms[J]. Computer Engineering and Applications, 2023, 59(18): 1 - 13. (in Chinese)

[12] MOHAMED H E D, FADL A, ANAS O, et al. Msr-yolo: method to enhance fish detection and tracking in fish farms[J]. Procedia Computer Science, 2020, 170: 539 - 546.

[13] LI X, HAO Y, AKHTER M, et al. A novel automatic detection method for abnormal behavior of single fish using image fusion [J]. Computers and Electronics in Agriculture, 2022, 203: 107435.

[14] HU X, LIU Y, ZHAO Z, et al. Real-time detection of uneaten feed pellets in underwater images for aquaculture using an improved YOLO-V4 network[J]. Computers and Electronics in Agriculture, 2021, 185: 106135.

[15] ROSALES M A, PALCONIT M G B, ALMERO V J D, et al. Faster R - CNN based fish detector for smart aquaculture system [C]// 2021 IEEE 13th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment, and Management (HNICEM). IEEE, 2021: 1 - 6.

[16] 王法安,何忠平,张兆国,等. 基于 PN - YOLO v8s - Pruned 的轻量化三七收获目标检测方法[J]. 农业机械学报,2024, 55(11):171 - 183.
WANG Faan, HE Zhongping, ZHANG Zhaoguo, et al. Lightweight object detection method for *Panax notoginseng* based on PN - YOLO v8s - Pruned[J]. Transactions of the Chinese Society for Agricultural Machinery,2024,55(11):171 - 183. (in Chinese)

[17] 刘月峰,刘博,暴祥,等. 基于超轻量化孪生网络的自然场景奶牛单目标跟踪方法[J]. 农业机械学报,2023,54(10): 282 - 293.
LIU Yuefeng, LIU Bo, BAO Xiang, et al. Single target tracking method for dairy cows in natural scenes[J]. Transactions of the Chinese Society for Agricultural Machinery,2023,54(10):282 - 293. (in Chinese)

[18] DU K, BOBKOV A. An overview of object detection and tracking algorithms[J]. Engineering Proceedings, 2023, 33(1): 22.

[19] BARREIROS M O, DANTAS D O, SILVA L C O, et al. Zebrafish tracking using YOLO v2 and Kalman filter[J]. Scientific Reports, 2021, 11(1): 3219.

[20] WANG H, ZHANG S, ZHAO S, et al. Real-time detection and tracking of fish abnormal behavior based on improved YOLO V5 and SiamRPN + +[J]. Computers and Electronics in Agriculture, 2022, 192: 106512.

[21] 黄一凡, 陈欣, 袁飞. 鱼类应激行为作用下的水质视频监测分析系统[J]. 厦门大学学报(自然科学版),2017,56(4): 584 - 589.
HUANG Yifan,CHEN Xin,YUAN Fei. The video monitoring system of water quality based on stress behavior analysis of fish [J]. Journal of Xiamen University(Natural Science),2017,56(4):584 - 589. (in Chinese)

[22] AHARON N, ORFAIG R, BOBROVSKY B Z. BoT - SORT: robust associations multi-pedestrian tracking[J]. arXiv Preprint, arXiv:2206.14651, 2022.

[23] CHEN J, KAO S, HE H, et al. Run, don't walk: chasing higher FLOPS for faster neural networks[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023: 12021 - 12031.

[24] OUYANG D, HE S, ZHANG G, et al. Efficient multi-scale attention module with cross-spatial learning[C]// ICASSP 2023—2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2023: 1 - 5.

[25] 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, 2020: 10781 - 10790.

[26] BEWLEY A, GE Z, OTT L, et al. Simple online and realtime tracking[C]// 2016 IEEE International Conference on Image Processing (ICIP). IEEE, 2016: 3464 - 3468.

[27] 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, 2017: 3645 - 3649.