

DOI:10.19651/j.cnki.emt.2517837

# 基于 YOLO-Crab 与改进的 DeepSORT 的水下河蟹检测与计数方法<sup>\*</sup>

吴佳文 姬伟 翟科龙 许波

(江苏大学电气信息工程学院 镇江 212013)

**摘要:**为实现淡水池塘中无人水产养殖船的精准投饵,提出了 YOLO-Crab+改进的 DeepSORT 的河蟹计数方法。首先,针对水下河蟹图像模糊、对比度低等问题,提出了一种基于 CLAHE 预处理下的 YOLOv8 的河蟹检测模型 YOLO-Crab。YOLO-Crab 在主干中增加坐标注意力机制,提高检测精度,同时,通过 SimSPPF 池化和 GSConv+ Slim Neck 设计减轻模型量级。改进的 DeepSORT 算法用 DIOU 匹配替代 IOU 匹配来解决水草遮挡导致的河蟹 ID 跳变问题。实验表明,YOLO-Crab 模型检测精度和 F1 分别达到了 97.3% 和 94%,计数方法平均精度为 81%。同时,将模型移植到 Jeston AGX Orin 上,检测精度达到 95%,检测速度为 60 fps,提升了 50%,计数精度为 78%,能够为无人水产养殖船精准投饵提供可靠依据。

**关键词:**无人水产养殖;YOLOv8;目标检测;河蟹计数;DeepSORT

**中图分类号:** TP391.4;TN919.8 **文献标识码:** A **国家标准学科分类代码:** 520.604

## Detection and counting method for underwater crabs based on YOLO-Crab and the improved DeepSORT

Wu Jiawen Ji Wei Zhai Kelong Xu Bo

(School of Electrical and Information Engineering, Jiangsu University, Zhenjiang 212013, China)

**Abstract:** To realize accurate feeding of unmanned aquaculture vessels in freshwater ponds, a river crab counting method with YOLO-Crab + improved DeepSORT is developed. First, to address the problems of blurring and low contrast of underwater river crab images, a river crab detection model YOLO-Crab based on YOLOv8 under the preprocessing of CLAHE is proposed. YOLO-Crab adds the coordinate attention mechanism in the backbone to improve the detection precision, and, at the same time, reduces the model magnitude by SimSPPF pooling and GSConv+ Slim Neck design to mitigate the model magnitude. The improved DeepSORT algorithm replaces IOU matching with DIOU matching to solve the problem of river crab ID jumping caused by aquatic grass occlusion. Experiments show that the detection precision and F1 of YOLO-Crab model reach 97.3% and 94%, respectively, and the average precision of counting methods is 81%. At the same time, the model was transplanted to Jeston AGX Orin, and the detection accuracy reached 95%, the detection speed was 60 fps, an increase of 50%, and the counting accuracy was 78%, which can provide a reliable basis for accurate feeding of unmanned aquaculture vessels.

**Keywords:** unmanned aquaculture; YOLOv8; target detection; river crab counting; DeepSORT

## 0 引言

河蟹因其美味、丰富的营养和极高的经济价值而深受人们的喜爱。近年来,河蟹养殖规模不断扩大,在我国淡水渔业产量排行中名列前茅<sup>[1]</sup>。目前,河蟹养殖主要依赖渔民经验进行喂养。这种方法具有很强的随机性,因为河蟹

所需的饵料受季节、天气、水温、分布区域等因素的影响。河蟹在养殖过程中,饵料成本约占总成本的 50%。在传统方法中饵料过少时,河蟹会因饥饿而抢食,甚至蚕食蜕壳的新蟹,降低养殖成活率;饵料过多时,会因长时间未被食用而污染水质<sup>[2]</sup>。因此,对于可变的饲喂需求非常迫切<sup>[3]</sup>。随着计算机视觉的发展,实时定位和检测复杂环境下分布

收稿日期:2025-01-07

<sup>\*</sup> 基金项目:国家自然科学基金(61973141)项目资助

不均匀的河蟹成为可能<sup>[4]</sup>。此外,河蟹领地意识强、活动范围小<sup>[5]</sup>,因此可以利用安装在无人养殖船上的计算机视觉系统来计算特定区域的河蟹数量,从而为精准投饵提供可靠依据。

水下图像因光散射而常显低对比度、亮度不均及模糊。为改善此状况,已有许多方法被提出。Iqbal 等<sup>[6]</sup>采用变换域法,经拉普拉斯分解后处理低频去雾、归一化及白平衡,高频放大,最终合成增强图像。Cao 等<sup>[7]</sup>提出深度学习端到端暗光增强方法,有效提升亮度,解决水下弱光问题。

快速有效检测水下河蟹是实现河蟹数量统计的前提。过去,检测算法主要依靠目标的颜色<sup>[8]</sup>、纹理<sup>[9]</sup>和形状<sup>[10]</sup>等特征,并利用形态学操作进行图像分割<sup>[11]</sup>以检测目标区域。但是传统的算法难以兼顾实时性和准确性,且研究结果不具有通用性。基于深度学习的目标检测算法能够更高效地表达数据集的特征,具有更强的泛化能力,更容易应用于实际场景。其中 YOLO 等的一阶段检测算法与 Faster RCNN 等的二阶段检测算法比较具有代表性,被广泛的应用于河蟹的检测研究<sup>[12-15]</sup>。Ji 等<sup>[16]</sup>改进了 CenterNet 模型用于水下目标检测,提高了模型对小目标的检测能力;周志耀等<sup>[17]</sup>提出 RCE-YOLO 模型,有效解决水下检测模型参数量和计算量过大的问题。

近年来,面向水下目标计数方法得到广泛研究。安志强等<sup>[18]</sup>提出了一种基于视频多目标跟踪的多类别海产品计数方法。利用 YOLOv7 和 BYTE 算法,解决了成本高、效率低的问题;Wu 等<sup>[19]</sup>基于改进的 DeepSORT 算法,利用 YOLOv5 模型实现了对鱼道过鱼种类的动态识别和自动计数;Connolly 等<sup>[20]</sup>通过调整置信度、使用顺序非极大值抑制等方法对现有的深度学习方法进行改进,提高了鱼类计数准确性。

目前,水下河蟹的检测的研究已经比较成熟,检测精度也能达到很高,但大多数只是放在安卓设备进行测试实验,没有放到边缘设备上实验,此外,目前水下目标计数的研究大多集中于鱼类,关于水下河蟹的计数研究很少。为了指导无人水产养殖船自动准确的投放饵料,本文采用基于

深度学习的目标检测技术和多目标跟踪技术实现对视频中河蟹的计数。本文主要贡献如下:

1) 提出了一种名为 YOLO-Crab+改进的 DeepSORT 的水下河蟹检测方法,该方法可以为无人水产养殖船的精准投饵提供可靠的依据。

2) 为了解决水下图像模糊、对比度低等复杂环境特征,以满足在边缘设备上快速准确检测河蟹的实时要求,本文在基于限制对比度自适应直方图均衡化(contrast limited adaptive histogram equalization,CLAHE)预处理的主干网络中加入了坐标注意力机制,以增强网络中河蟹特征的提取,随后使用 SimSPPF(simplified spatial pyramid pooling-fast)和 GSConv(group shuffle convolution)模块在减少参数的同时加速模型的收敛,并在颈部采用 Slim Neck 操作以降低成本。

3) 用本文的 YOLO-Crab 替换 DeepSORT 检测,并用传统的交并比(intersection over union,IOU)匹配替换距离交并比(distance intersection over union,DIOU)匹配,以减少由水生植物遮挡引起的 ID 跳跃问题。

## 1 材料和方法

本文的总体框架如图 1 所示。首先,投饵船将采集的视频序列以帧的形式送入河蟹目标检测器中,采用所提出的 YOLO-Crab 河蟹目标检测模型得到帧的检测结果,该模型应用 SimSPPF 模块和 GSConv+Slim Neck 模块在保持精度同时减轻量级,并在主干中加入坐标注意力模块可以提高对河蟹特征的识别精度;然后,将此检测结果与第  $i-1$  帧河蟹轨迹经卡尔曼预测得到的预测框使用匈牙利算法进行目标匹配,得到第  $i$  帧的匹配结果,针对其中目标匹配时由于水草遮挡引起的 ID 跳变问题,使用距离交并比替换普通交并比可以有效解决;最后,根据匹配结果更新河蟹的轨迹状态,实现对同一物体的定位。以此反复,直到整个视频序列处理完成,得到最大的 ID 值,即检测范围内河蟹的数量,将此数据作为无人水产养殖船进行饵料精准投放的依据。

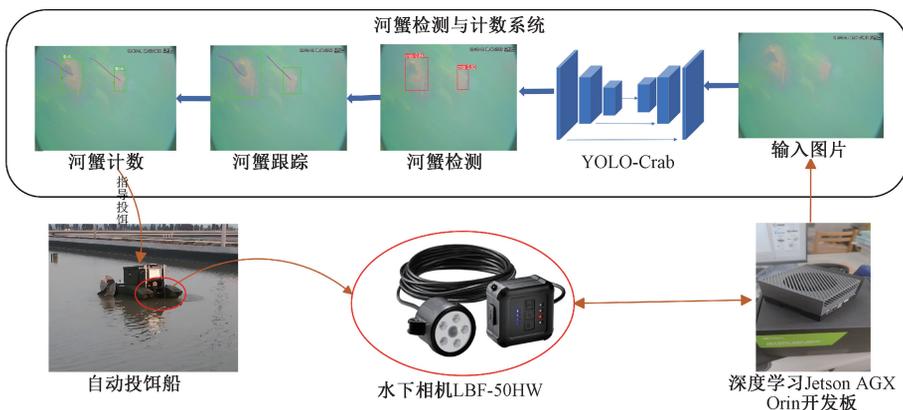
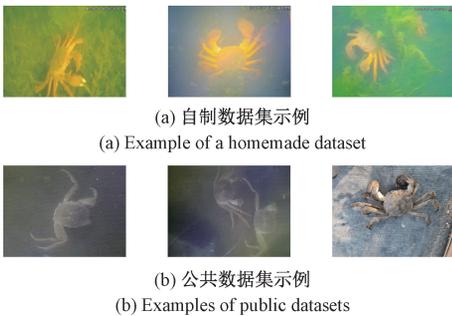


图 1 河蟹计数方法总体框架

Fig. 1 General framework for river crab counting method

### 1.1 图像数据集的获取

水下河蟹视频是在江苏省常州市长荡湖养殖基地 10 个蟹塘里收集的,总占地约  $1.68 \times 10^4 \text{ m}^2$ 。每个蟹塘深约 0.5 m,占地面积约  $50 \text{ m} \times 30 \text{ m}$  并且每个蟹塘约有 2 500 只河蟹。图像采集设备是安装在自动投饵作业船底部由青岛博飞海洋有限公司开发的包含 5 个辅助 LED 照明灯的可伸缩便捷式水下摄像机 LBF-50HW,分别在傍晚和清晨对 10 个池塘进行数据采集,共获得 10 个时长约 30 min 的视频。通过对采集到的视频进行逐帧提取,剔除部分相似和模糊的图像,最终获得 1 877 幅河蟹清晰的图像。同时,为进一步丰富数据集并且提高模型的泛化能力,从公共数据集 Aquatic Species Dataset 以及百度上共挑选了 200 张图片。本实验所使用的图像示例如图 2 (a) 和 (b) 所示。



(a) 自制数据集示例  
(a) Example of a homemade dataset  
(b) 公共数据集示例  
(b) Examples of public datasets

图 2 数据集示例  
Fig. 2 Example of datasets

为增加数据集数量以提高深度学习训练效果,本文采用上下翻转、镜像翻转、缩放、亮度调节、旋转等数据增强方法对数据集进行增强,增强后获得数据集图像 14 566 张,增强结果如图 3 (b) 和 (c) 所示。



(a) 原图  
(a) Original picture  
(b) 上下翻转的图像  
(b) Flip up/down image  
(c) 镜像翻转的图像  
(c) Mirror flipped image

图 3 数据集增强举例

Fig. 3 Example of dataset enhancements

### 1.2 基于 YOLO-Crab 模型的河蟹检测

针对水下图像模糊、对比度低等复杂环境特点以及在边缘设备上快速准确检测河蟹的实时性要求,本文在 YOLOv8 网络模型基础上,提出了河蟹检测模型 YOLO-Crab,其结构图如图 4 所示。首先,为解决图像模糊、对比度低的问题,对图像进行 CLAHE 的预处理;接着,在骨干网络结构中添加坐标注意力机制,增强了网络对河蟹相关特征的提取;并替换 SPPF 模块为 SimSPPF 模块和将普通卷积替换为 GSCnv,减少参数的同时增加模型收敛速度;接着对颈部进行 Slim Neck 操作来降低计算成本。

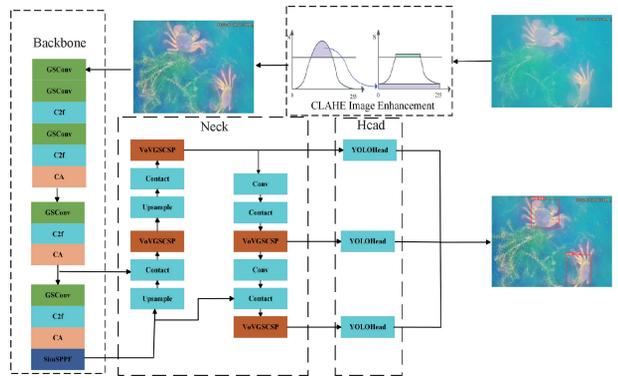


图 4 YOLO-Crab 结构  
Fig. 4 YOLO-Crab structure

#### 1) 限制对比度自适应直方图均衡化

针对水下环境光照不均匀以及水介质对光的吸收等引起水下图像模糊和对比度低的特点,采用 CLAHE 的预处理方法对水下图像进行增强<sup>[21]</sup>。CLAHE 算法是一种局部对比度增强技术,通过将图像划分为小区域并独立应用直方图均衡化,同时限制对比度以增强细节并减少噪声,能够有效的改善图片的局部视觉效果。

#### 2) 坐标注意力模块设计

压缩和激励模块 (squeeze-and-excitation networks, SE)<sup>[22]</sup> 仅使用了通道建模,在河蟹检测通道数较少时,其检测结果并不理想;虽然卷积块注意力模块 (convolutional block attention module, CBAM)<sup>[23]</sup> 考虑到了通道、空间两个维度,但将两者进行串联,缺少了长程依赖。为了提高河蟹精确位置信号的捕获能力以及重要特征关注度,本文在 YOLOv8 特征提取网络中增添坐标注意力模块 (coordinate attention, CA)<sup>[24]</sup>,将河蟹图像中的空间信息建模到通道中,少量增加参数与计算量,提升模型对重要特征区域的关注程度。

CA 具体的结构如图 5 所示。在 CA 模块中,河蟹特征图经过宽度和高度方向的全局平均池化后,分别得到  $[C, H, 1]$  和  $[C, 1, W]$  尺寸的特征图。这些特征图在合并后变为  $[C, 1, H+W]$ ,再经过卷积、标准化、激活函数处理。之后再次分开为两个并行阶段,分别处理宽度和高度,最终获得河蟹  $[C, H, 1]$  和  $[C, 1, W]$  尺寸的特征层。通过  $1 \times 1$  卷积和 sigmoid 函数调整通道数,将其与原始特征相乘,就是 CA 注意力机制。

#### 3) SimSPPF 模块设计

SimSPPF 的具体结构如图 6 所示。SimSPPF<sup>[25]</sup> 使用多个串联的小型池化核,而不是空间金字塔池化 (spatial pyramid pooling, SPP) 模块中的单个大型池化核,同时 SimSPPF 模块增大了感受野。SimSPPF 结构是将输入串行通过多个  $5 \times 5$  大小的 MaxPool 层,串行两个  $5 \times 5$  大小的 MaxPool 层是和一个  $9 \times 9$  大小的 MaxPool 层计算结果是一样的,串行 3 个  $5 \times 5$  大小的 MaxPool 层是和一个

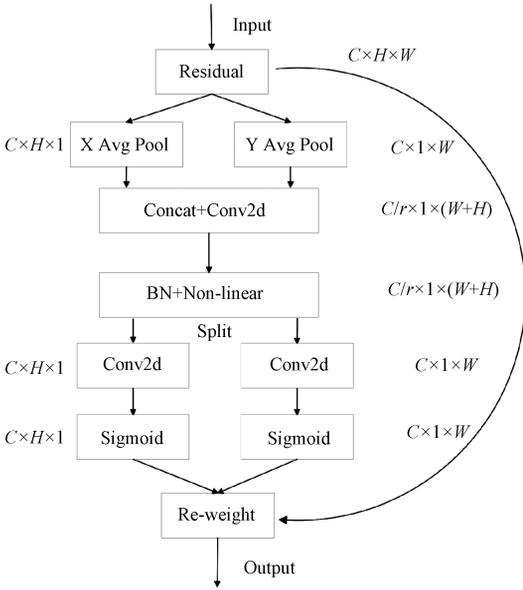


图5 坐标注意力机制结构

Fig. 5 Structure of coordinate attention mechanism

13×13 大小的 MaxPool 层计算结果是一样的。这种设计既保留了原有的功能,又减少了运算量,提高了运行速度,使 SimSPPF 结构更加高效。

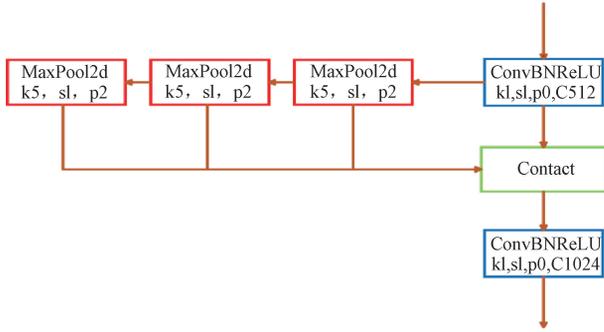


图6 SimSPPF 结构

Fig. 6 Structure of SimSPPF

#### 4) GSConv+Slim Neck 设计

GSConv<sup>[26]</sup>是一种轻量级卷积方法,用于替代普通卷积来获取模型骨干的特征映射。GSConv 的结构如图 7 所示。它通过将特征提取操作分割,减少计算量,并使用多通道密集卷积来保留通道间的隐藏连接,从而减少语义信息的损失。最后,通过通道融合完成特征提取。

图 8 是 VoVGSCSP 模块的结构图。VoVGSCSP 模块降低了计算和网络结构的复杂性,但保持了足够的精度,且冗余重复信息较少。VoVGSCSP 骨干利用两个 GSConv 卷积迅速传递强语义特征,完成上采样和下采样,最终减少模型信息处理时间。

### 1.3 水下河蟹的跟踪和计数

通过 YOLO-Crab 获得河蟹的位置信息后,采用 DeepSORT 算法<sup>[27]</sup>对河蟹进行跟踪计数。与单帧图像的

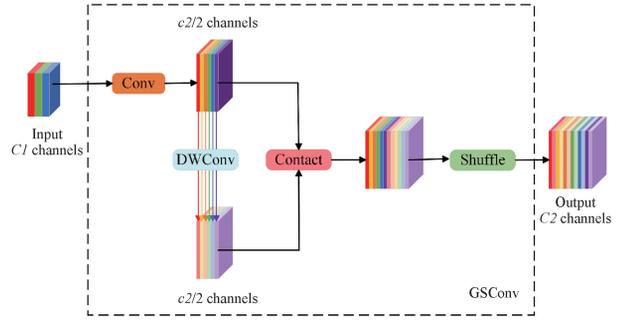


图7 GSConv 模块结构

Fig. 7 Structure of the GSConv module

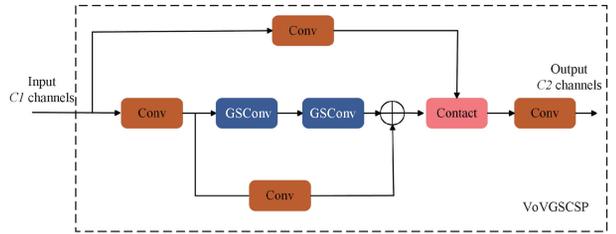


图8 VoVGSCSP 模块结构

Fig. 8 Structure of the VoVGSCSP module

计数不同,DeepSORT 算法是对视频序列进行跨帧计数,获得视频段内所有河蟹数量。改进后的 DeepSORT 的流程图如图 9 所示。本文将检测部分由原始的 Faster RCNN 替换为 YOLO-Crab 对河蟹进行逐帧检测,得到河蟹的位置信息;然后,根据卡尔曼滤波对河蟹位置进行预测,用匈牙利算法匹配连续帧中的相同河蟹;同时,使用 DIOU 替换 IOU,有效解决水草遮挡引起 ID 跳跃的问题;用关联度量来衡量预测河蟹和检测河蟹之间的相似性,实现对同一河蟹的定位。

#### 1) 卡尔曼滤波

卡尔曼滤波利用先验估计,得到后验估计,其主要作用是通过检测器测量获得预测数据,再根据预测和更新的公式,可以基于河蟹前一时刻的位置来预测当前时刻的位置。

预测公式:

$$\hat{\mathbf{x}}_k^- = \mathbf{A}\hat{\mathbf{x}}_{k-1} + \mathbf{B}\mathbf{u}_k \quad (1)$$

$$\mathbf{P}_k^- = \mathbf{A}\mathbf{P}_{k-1}\mathbf{A}^T + \mathbf{Q} \quad (2)$$

更新公式:

$$\mathbf{K}_k = \frac{\mathbf{P}_k^- \mathbf{C}^T}{\mathbf{C}\mathbf{P}_k^- \mathbf{C}^T + \mathbf{R}} \quad (3)$$

$$\hat{\mathbf{x}}_k = \hat{\mathbf{x}}_k^- + \mathbf{K}_k(\mathbf{y}_k - \mathbf{C}\hat{\mathbf{x}}_k^-) \quad (4)$$

$$\mathbf{P}_k = (\mathbf{I} - \mathbf{K}_k\mathbf{C})\mathbf{P}_k^- \quad (5)$$

其中,  $\hat{\mathbf{x}}_k^-$  和  $\mathbf{P}_k^-$  为先验状态估计值和协方差。 $\hat{\mathbf{x}}_{k-1}$ 、 $\mathbf{K}_k$ 、 $\hat{\mathbf{x}}_k$  和  $\mathbf{P}_k$  分别代表后验状态估计值、卡尔曼增益、更新的观测值和误差协方差。 $\mathbf{A}$ 、 $\mathbf{B}$ 、 $\mathbf{Q}$  和  $\mathbf{u}_k$  分别代表状态转移矩阵、输入状态转移矩阵、系统协方差和输入值。 $\mathbf{C}$ 、 $\mathbf{R}$ 、 $\mathbf{I}$  和

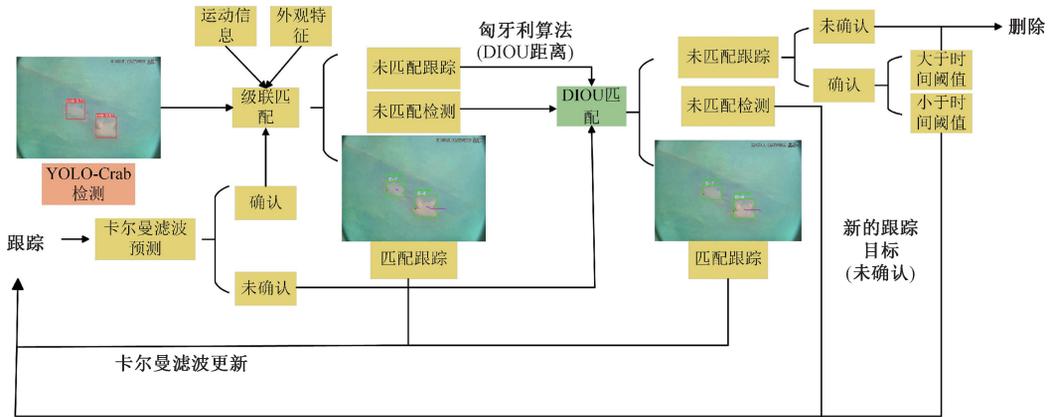


图 9 DeepSORT 算法改进结构

Fig. 9 Flowchart of the improved DeepSORT algorithm

$y_k$  代表观测转换矩阵、噪声协方差、单位矩阵和观测值。

### 2) 匈牙利算法

匈牙利算法解决指派问题,即当前帧检测到的河蟹该匹配到前面哪个跟踪信息,被用来建立河蟹检测和河蟹预测之间的关系。其作用是解决分配问题,把一群检测框和卡尔曼预测的框做分配,让卡尔曼预测的框找到和自己最匹配的检测框,达到追踪效果。

### 3) 关联度量

DeepSORT 通过结合目标框马氏距离和外观特征余弦距离两个度量来整合运动和外观信息。为了获取河蟹运动信息,DeepSORT 使用马氏距离计算预测目标框和检测框之间的匹配情况,可以有效计算两个未知样本之间的相似度,公式如下:

$$d^{(1)}(i, j) = (d_j - y_i)^T S_i^{-1} (d_j - y_i) \quad (6)$$

其中,  $d_j$  表示第  $j$  个候选目标框,  $y_i$  表示对当前帧图像进行河蟹位置预测的第  $i$  种运动状态,  $S_i$  表示检测目标框与预测框之间的协方差矩阵。

考虑到单条运动轨迹不间断特性,利用马氏距离对目标状态轨迹进行预测,筛选候选目标框,可以用式(7)阈值函数来表示:

$$b_{i,j}^{(1)} = 1[d^{(1)}(i, j) \leq t^{(1)}] \quad (7)$$

其中,  $t^{(1)}$  表示相关阈值,  $b_{i,j}^{(1)}$  表示阈值指标。

当出现水草遮挡的问题时,马氏距离不能很好地进行度量,ID 跳跃会导致不准确计数,为了解决这类问题,引入外观特征信息作为关联信息。外观信息关联可利用检测的特征描述与跟踪的特征描述之间的最小余弦距离来度量。公式如下:

$$d^{(2)}(i, j) = \min\{1 - r_j^T r_k^{(i)} \mid r_k^{(i)} \in R_i\} \quad (8)$$

其中,  $r_j$  表示第  $j$  个检测河蟹结果的特征向量,  $r_k^{(i)}$  表示第  $k$  次跟踪成功的特征向量,  $r_j^T r_k^{(i)}$  表示第  $j$  个检测和第  $k$  个跟踪成功两个向量之间的余弦相似度,  $1 - r_j^T r_k^{(i)}$  表示第  $j$  个检测和第  $k$  个跟踪成功两个向量之间的余弦距离。

同样设置阈值函数公式如下:

$$b_{i,j}^{(2)} = 1[d^{(2)}(i, j) \leq t^{(2)}] \quad (9)$$

运动信息在确定河蟹位置信息时,对于长时间遮挡的河蟹,运动信息的匹配能力就会减弱,这时距离度量会考虑外观信息,外观匹配对长期遮挡丢失的河蟹轨迹具有一定的有效性。因此选择将两种度量方式通过超参数  $\lambda$  的线性加权结合的方式进行结果判断,公式为:

$$c_{i,j} = \lambda d^{(1)}(i, j) + (1 - \lambda) d^{(2)}(i, j) \quad (10)$$

### 4) DeepSORT 算法优化

IOU 只关注预测框和检测框的重叠区域,而无法有效处理遮挡的情况。DIOU 通过引入距离度量因子来解决遮挡问题。除了计算预测框和检测框之间的重叠区域, DIOU 还考虑了预测框和检测框的中心点之间的距离以及宽高之间的差异。这个距离度量因子可以量化预测框和检测框之间的位置和尺寸差异,使得 DIOU 能够更好地处理预测框和检测框之间的遮挡情况。具体计算步骤如下:

- 1) 计算预测框和检测框之间的距离,使用两个框中心点之间的欧氏距离  $\rho(\mathbf{A}, \mathbf{B})$ ;
- 2) 计算预测框和检测框之间的最小包围框的对角线距离  $c$ ;
- 3) 根据欧式距离和最小包围框的对角线距离,计算 DIOU。

$$DIOU = IOU - \frac{\rho^2(\mathbf{A}, \mathbf{B})}{c^2} \quad (11)$$

其中,  $\mathbf{A}$  代表预测框,  $\mathbf{B}$  代表检测框。

## 2 实验

### 2.1 实验环境

本研究实验在云服务器基于 Ubuntu18.04 操作系统上进行的,硬件配置包括 12 vCPU Intel(R) Xeon(R) Platinum 8255C CPU @ 2.50 GHz 和 NVIDIA GeForce RTX 3080(10 GB) GPU。主要操作环境包括 Pytorch 2.0.0, Python 3.8 和 CUDA 11.8。图像的大小调整为

640 pixel×640 pixel,总的迭代次数为150轮,批大小设置为16,采用SGD作为优化器,算法学习率设置为0.01。

## 2.2 实验评价指标

### 1) 水下图像评价指标

水下彩色图像质量评估(underwater color image quality evaluation, UCIQE)<sup>[28]</sup>和 underwater 图像质量测量(underwater image quality measure, UIQM)<sup>[29]</sup>是评价图像质量两个重要指标,它们得分越高,说明图像复原的效果越好。

$$UCIQE = c_1 \times \sigma_c + c_2 \times con_l + c_3 \times \mu_s \quad (12)$$

$$UIQM = c_1 \times UICM + c_2 \times UISM + c_3 \times UIConM \quad (13)$$

其中,  $\sigma_c$  为色度的标准差,  $con_l$  为亮度的对比度,  $\mu_s$  为饱和度的平均值,  $c_1, c_2$  和  $c_3$  分别为线性组合的权重值。UICM 为水下图像色彩测量, UISM 为水下图像清晰度测量, 和 UIConM 为水下图像对比度测量。

### 2) 河蟹检测性能评价指标

本文使用总体评估指标(F1 score, F1)、平均精度均值(mean average precision, mAP)来评价训练得到的网络模型精度, FPS(每秒帧)、参数和模型大小评价速度, 计算公式如下:

$$P = \frac{TP}{TP + FP} \quad (14)$$

$$R = \frac{TP}{TP + FN} \quad (15)$$

$$mAP = \frac{1}{n} \int_0^1 P(R) dR \quad (16)$$

$$F1 = \frac{2TP}{2TP + FN + FP} \quad (17)$$

其中,  $TP$  指正样本被预测为正样本。  $FP$  指负样本被预测为正样本。  $FN$  指正样本被预测为负样本。  $P$  表示预测为正样本中实际为正样本的占比。  $R$  表示所有正样本中被预测为正样本的占比,  $n$  表示类别数。

### 3) 河蟹计数性能评价指标

使用平均计数精度(average counting precision, ACP)来评价基于多目标跟踪的计数方法的性能。

$$ACP = \left( 1 - \frac{\sum_{i=1}^n |S_i - N_i|}{\sum_{i=1}^n N_i} \right) \times 100\% \quad (18)$$

其中,  $S$  表示算法统计河蟹的数量,  $N$  表示人工统计河蟹的数量,  $i$  表示视频序号,  $n$  表示计数实验的视频总数量。

## 2.3 图像质量比较

通过结合主观视觉和客观视觉指标来判断增强前后的图像质量。增强前后图像对比及参数指标比较如图10和表1所示。从图10可以看出,图10(b)对河蟹目标的识别效果比图10(a),不需要关注的背景区域也没有过度增

强。从表1中可以看出,与原始图像相比,CLAHE图像增强算法在UIQM和UCIQE上都有很大的提高。然后,提供更高质量的图像以进行后续检测。为验证算法性能,对100张图像进行了测试,发现UCIQE和UIQM指标平均分别提升3.68%和90.96%,UIQM改善尤为显著,表明算法在图像颜色、对比度、饱和度和锐度上表现良好。进一步验证显示,采用该图像增强方法后,YOLOv8河蟹检测模型的mAP提高了1%。因此,它是一种适用于水下河蟹图像增强的算法。

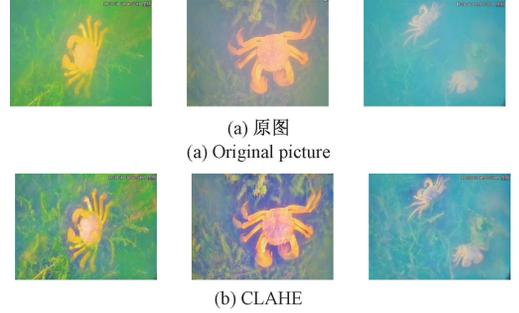


图10 水下图像增强前后效果

Fig. 10 Before and after enhancement of underwater images

表1 增强前后的参数对比

Table 1 Comparison of parameters before and after enhancement

评价指标	UCIQE	UIQM	mAP/%
原图	0.456 6	1.283	93.8
CLAHE	0.473 4	2.450	94.8

## 2.4 河蟹的目标检测实验

所提出的河蟹检测算法的关键部分是改进CA模块、SimSPPF模块和GSConv+Slim Neck设计范式。消融试验的结果如表2所示。可以看出,CA的引入导致mAP提高了1.7%,这是由于CA模块作为一个注意力机制可以突出重要的特征区域。SimSPPF模块的引入使mAP又提高了1.2%,这是因为SimSPPF可以增大感受野,同时由于SimSPPF的结构减少了计算量,模型的参数量减少了5%。GSConv+Slim Neck设计范式的加入,虽然mAP下降了0.4%,但是参数量减少了3.5%并且FPS增加了25.2,这对于轻量化和快速性是可以接受的。采用YOLO-Crab的目标检测模型mAP、F1和FPS分别可以达到97.3%、94%和141.1 fps,单张图像检测速度约为9 ms,这能够很好的满足水产养殖系统的实时要求。

## 2.5 不同检测模型的消融实验

为了证明所提出YOLO-Crab水下河蟹目标检测的性能,与Faster RCNN、SSD、YOLOv5和YOLOv7进行了比较。图11和12分别为不同检测模型的雷达图和PR曲线。表3显示不同模型对河蟹检测的性能指标。

表 2 基于 YOLOv8 的消融实验

Table 2 Ablation experiments based on YOLOv8

YOLOv8	CA	SimSPPF	GSCov+ Slim Neck	mAP/%	F1/%	FPS/fps	Params/M
✓	—	—	—	94.8	88	89.6	6.0
✓	✓	—	—	96.5	91	90.4	6.0
✓	✓	✓	—	97.7	94	115.9	5.7
✓	✓	✓	✓	97.3	94	141.1	5.5

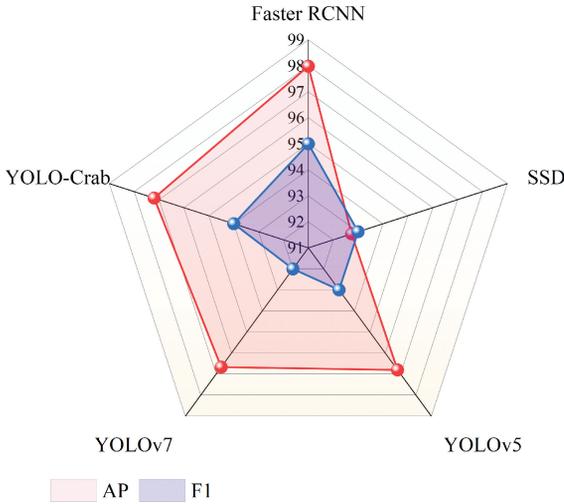


图 11 消融实验雷达图

Fig. 11 Radagram of ablation experiments

从雷达图中能够看到 SSD 的检测性能最差,检测性能表现较好的是本文的模型 YOLO-Crab 和 Faster RCNN。从 PR 曲线和表 3 可以看出, YOLO-Crab 的河蟹目标检测

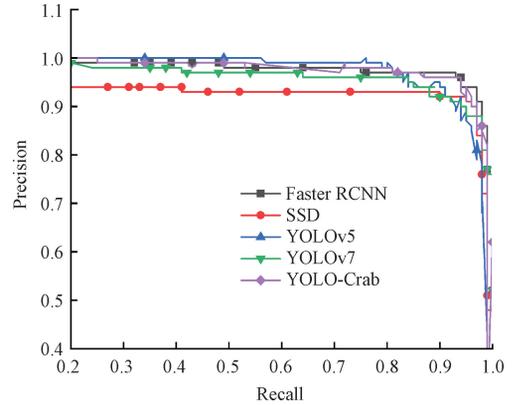


图 12 不同检测模型的 PR 曲线

Fig. 12 PR curves for different detection models

精度第二,检测速度最快。YOLO-Crab 的检测精度比两阶段的 Faster RCNN 稍微差了点,但是模型大小和速度明显优于 Faster RCNN,模型大小减少了 102.5 MB,识别速度约为 5 倍。YOLO-Crab 与单阶段的 SSD、YOLOv5 和 YOLOv7 相比,检测速度和模型大小都是最优的,并且与 YOLOv5 比其参数量减少了 22%。

表 3 河蟹检测任务中不同模型的性能指标

Table 3 Performance metrics of different models for river crab detection tasks

模型	mAP/%	F1/%	FPS/fps	Params/M	模型大小/MB
Faster RCNN	98.00	95	26.3	136.69	108
SSD	92.54	93	135.7	23.61	90.6
YOLOv5	96.80	93	52.1	7.06	27.1
YOLOv7	96.67	92	70.8	37.20	87.5
YOLO-Crab	97.30	94	141.1	5.5	5.5

为了研究检测器对计数性能的影响,本文将以 YOLOv5 和 YOLO-Crab 作为检测器进行计数实验,实验结果如图 13 和 14 所示。首先,从图中不难发现 YOLO-Crab 的检测结果整体比 YOLOv5 的检测结果要好;然后在图 13 中,河蟹从第 101 帧的 ID 值 2 到第 150 帧时跳变为 3,而在图 14 中就不会发生 ID 跳变的情况;其次,从第 101、150、206 帧和第 281 帧处图 13 存在河蟹漏检的情况,而在图 14 中可以有效的检测出漏检的河蟹。因此,检测器的结果决定了 DeepSORT 追踪的好坏,一个好的检测器

能够准确地检测出目标物体,并提供准确的边界框位置和特征向量。这样,DeepSORT 才能更好的追踪目标,减少 ID 跳变,提高计数精度。

2.6 河蟹计数结果分析

本研究旨在对河蟹计数从而可以对无人水产养殖船的自动投饵提供科学的指导,由于池塘水下生态环境的差异性导致河蟹在池塘的分布情况差异很大,所以计数时间不宜过长,因此实验设定为每段 1 min。

关于水草遮挡的计数问题如图 15 和 16 所示,第 4 帧

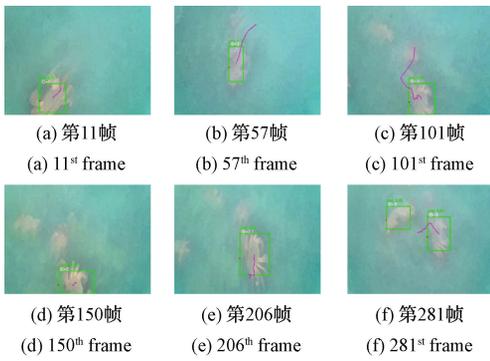


图 13 YOLOv5+DeepSORT 河蟹跟踪检测  
Fig. 13 YOLOv5+DeepSORT river crab detection

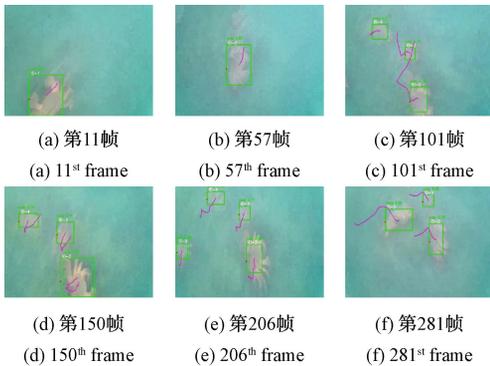


图 14 YOLO-Crab+DeepSORT 河蟹跟踪检测  
Fig. 14 YOLO-Crab+DeepSORT river crab detection

时没有水草遮挡,都能够准确检测出河蟹并初始化 ID,第 60 帧时虽然被水草遮挡,但是并不是全部被遮挡住还能提取特征,所以 ID 值保持不变,在第 183 帧时,由于河蟹全部被水草遮挡,虽然一直被跟踪,但是跟踪框中并不存在该河蟹,使得卡尔曼滤波预测的结果不确定性逐步增加,导致该河蟹的轨迹与检测结果不能成功匹配,所以,当它再次出现时,算法错误的把它认为新目标对 ID 值进行了更新。而本文改进的算法考虑了目标框的中心点之间的

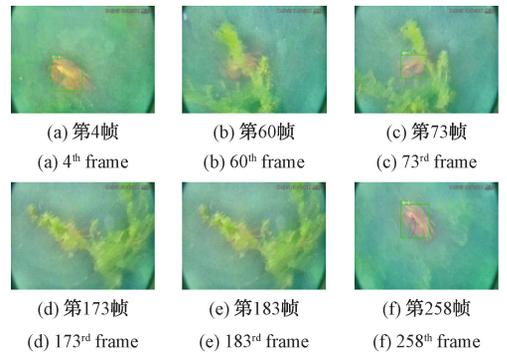


图 15 YOLO-Crab+DeepSORT 河蟹跟踪结果  
Fig. 15 YOLO-Crab+DeepSORT river crab tracking result

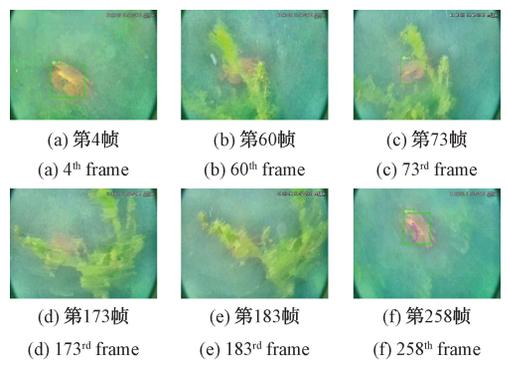


图 16 YOLO-Crab+改进的 DeepSORT 河蟹跟踪结果  
Fig. 16 YOLO-Crab+ Improved DeepSORT river crab tracking result

距离以及宽高之间的差异,并且引入距离度量因子使其量化目标框之间的位置和尺寸差异,很好的处理水草遮挡的问题,所以当河蟹再次出现时,它的 ID 值保持不变。

从数据集中随机选取 10 个 1 min 的河蟹视频,人工统计 10 个视频段的河蟹数量,并将这 10 个视频输入 YOLO-Crab+DeepSORT 和 YOLO-Crab+改进的 DeepSORT 算法中,对河蟹的数量进行统计,得到如表 4 所示的计数信息。基于式(18)计算 10 个河蟹视频的计数精度为 81%。

表 4 河蟹计数试验结果统计

Table 4 Statistics on the results of the river crab counting test

视频序号	YOLO-Crab+DeepSORT			YOLO-Crab+改进的 DeepSORT		
	人工计数 S	模型计数 N	差值   S-N	人工计数 S	模型计数 N	差值   S-N
01	7	9	2	7	7	0
02	15	21	6	15	18	3
03	6	5	1	6	6	0
04	11	14	3	11	12	1
05	10	13	3	10	10	0
06	4	4	0	4	4	0
07	11	21	10	11	17	6
08	12	18	6	12	14	2
09	19	34	15	19	25	6
10	5	7	2	5	6	1
总计 Sum	100	146	48	100	119	19
准确率 ACP		52%			81%	

## 2.7 投饵船上部署与验证

为了验证模型在边缘设备上的运行有效性,将 YOLO-Crab 模型训练后的 best.pt 权重文件简化为 onnx 模型,再使用 TensorRT 的 API 将其转换为高度优化的 TensorRT 模型,最后将 TensorRT 模型加载到 TensorRT 的引擎中,在 Jetson AGX Orin 上执行模型推理,完成模型的移植。随机选取 300 张图片和 5 个视频,将 YOLO-Crab 与 YOLOv8 两

个模型都部署在边缘设备上,随后分别在改进的 DeepSORT 上进行计数,相应的检测结果对比情况如表 5 所示。由于边缘设备算力有限,但改进后模型的检测精度相较于原始模型提升了 2.8%,并且检测速度提升了 50%,证明了对 YOLOv8 模型进行轻量化改进同时精度保持不变在边缘设备上具有较好的适应性。改进后模型的河蟹计数精度提高了 8%,为河蟹养殖提供强有力的技术支持。

表 5 模型在投饵船上测试性能比较

Table 5 Comparison of test performance of the model on unmanned aquaculture vessels

模型	输入图片大小	平均 GPU 利用率/%	mAP/%	FPS/fps	ACP/%
YOLOv8	640×640	50	92.3	40	70
YOLO-Crab	640×640	36	95.0	60	78

## 3 结 论

本文提出一种基于 YOLO-Crab+改进的 DeepSORT 算法的池塘河蟹检测、计数方法,为无人水产养殖船精准投饵提供可靠依据。针对水下河蟹检测任务中存在图像模糊、对比度低的问题,使用 CLAHE 图像预处理的方法;针对检测效果差以及部署边缘设备的要求,在 YOLOv8 模型中结合了 CA、SimSPPF 和 GSConv+ Slim Neck 机制,对河蟹进行了更好的特征提取,提高了检测精度和速度。在目标跟踪方面,在 DeepSORT 算法的基础上,将 IOU 替换为 DIOU,通过引入距离度量因子很好的解决遮挡的问题。对本文提出的方法进行验证,结果表明:增强后的图像具有更高的 UCIQE 和 UIQM 值,能够提高后续检测模型的性能。与 YOLOv8 相比,YOLO-Crab 的 mAP 提高了 2.5%,参数量下降了 3.5%。改进的 DeepSORT 的 ACP 为 81%。此外,将模型移植到 Jetson AGX Orin 开发板上,检测精度达到了 95%,检测速度达到 60 fps,计数精度达到 78%,表明了算法的有效性,研究结果为池塘河蟹养殖投饵系统精准投喂提供了可靠的依据。

## 参考文献

- [1] 戴红君,孙艺伟,任妮,等. 我国中华绒螯蟹产业现状调查及发展对策分析[J]. 江苏农业科学, 2021, 49(18): 248-252.
- [2] LI D H, WANG Z H, WU S Y, et al. Automatic recognition methods of fish feeding behavior in aquaculture: A review [J]. Aquaculture, 2020, 528: 735508.
- [3] 韩文峰,孙云飞,刘健,等. 不同饵料投喂模式对中华绒螯蟹幼蟹生长和生理生化的影响[J]. 中国水产科学, 2021, 28(3): 314-325.

- HAN W F, SUN Y F, LIU J, et al. Effects of different bait feeding patterns on growth and physiological biochemistry of juvenile Chinese mitten crab *eriocheir japonica sinensis*[J]. Journal of Fishery Sciences of China, 2021, 28(3): 314-325.
- [4] TANG C, ZHANG G, HU H G, et al. An improved YOLOv3 algorithm to detect molting in swimming crabs against a complex background[J]. Aquacultural Engineering, 2020, 91: 102115.
- [5] 孙月平,赵德安,洪剑青,等. 河蟹养殖全覆盖自动均匀投饵的轨迹规划与试验[J]. 农业工程学报, 2016, 32(18): 190-200.
- SUN Y P, ZHAO D AN, HONG J Q, et al. Trajectory planning and experimentation of automatic uniform feeding for full-coverage river crab culture[J]. Transactions of the Chinese Society of Agricultural Engineering, 2016, 32(18): 190-200.
- [6] IQBAL M, RIAZ M M, SOHAIB ALI S, et al. Underwater image enhancement using laplace decomposition [J]. IEEE Geoscience and Remote Sensing Letters, 2022, 19: 1500105.
- [7] CAO S, ZHAO D, SUN Y P, et al. Learning-based low-illumination image enhancer for underwater live crab detection[J]. ICES Journal of Marine Science, 2021, 78(3): 979-993.
- [8] QIAO X, BAO J H, ZHANG H, et al. Underwater sea cucumber identification based on principal component analysis and support vector machine[J]. Measurement, 2019, 133: 444-455.
- [9] BAI Y X, ZHANG S H, FAN Z, et al. Automatic multiple zebrafish tracking based on improved HOG features[J]. Scientific Reports, 2018, 8(1): 1-14.
- [10] KANNAN S. Intelligent object recognition in underwater images using evolutionary-based Gaussian mixture model and shape matching[J]. Signal, Image

- and Video Processing, 2020, 14(5): 877-885.
- [11] 郭传鑫, 李振波, 乔曦, 等. 基于融合显著图与GrabCut算法的水下海参图像分割[J]. 农业机械学报, 2015, 46(S1): 147-152.  
GUO CH X, LI ZH B, QIAO X, et al. Segmentation of underwater sea cucumber images based on fused saliency map and GrabCut algorithm[J]. Transactions of the Chinese Society of Agricultural Machinery, 2015, 46(S1): 147-152.
- [12] JI W, ZHANG T, XU B, et al. Apple recognition and picking sequence planning for harvesting robot in the complex environment [J]. Journal of Agricultural Engineering, 2024, DOI: 10.4081/jae.2023.1549.
- [13] YU Y, ZHAO J, GONG Q, et al. Real-time underwater maritime object detection in side-scan sonar images based on transformer-YOLOv5 [J]. Remote Sensing, 2021, 13: 3555.
- [14] WEN G, LI S, LIU F, et al. YOLOv5s-CA: A modified YOLOv5s network with coordinate attention for underwater target detection[J]. Sensors, 2023, 23: 3367.
- [15] CAI K, MIAO X, WANG W, et al. A modified YOLOv3 model for fish detection based on MobileNetv1 as backbone [J]. Aquacultural Engineering, 2020, 91: 102117.
- [16] JI W, PENG J Q, XU B, et al. Real-time detection of underwater river crab based on multiscale pyramid fusion image enhancement and mobile CenterNet model[J]. Computers and Electronics in Agriculture, 2023, 204: 107522.
- [17] 周志耀, 马常霞, 杨丽莎, 等. 基于轻量级改进的YOLOv8水下目标检测模型[J]. 电子测量技术, 2024, 47(19): 181-189.  
ZHOU ZH Y, MA CH X, YANG L SH, et al. Lightweight and improved YOLOv8 underwater target detection model [J]. Electronic Measurement Technology, 2024, 47(19): 181-189.
- [18] 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.  
AN ZH Q, LI ZH J, LIU SH, et al. A multi-category seafood counting method incorporating YOLOv7 and BYTE multi-target tracking[J]. Transactions of the Chinese Society of Agricultural Engineering, 2023, 39(9): 183-189.
- [19] WU B, LIU C, JIANG F, et al. Dynamic identification and automatic counting of the number of passing fish species based on the improved DeepSORT algorithm [J]. Frontiers in Environmental Science, 2023, 11: 1059217.
- [20] CONNOLLY R M, FAIRCLOUGH D V, JINKS E L, et al. Improved accuracy for automated counting of a fish in baited underwater videos for stock assessment [J]. Frontiers in Marine Science, 2021, 8: 658135.
- [21] YUAN Z, ZENG J, JIN L, et al. CLAHE-based low-light image enhancement for robust object detection in overhead power transmission system [J]. IEEE Transactions on Power Delivery, 2023, 38 (3): 2240-2243.
- [22] HU J, SHEN L, ALBANIE S, et al. Squeeze-and-excitation networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 42 (8): 2011-2023.
- [23] WOO S H, PARK J, LEE J Y, et al. CBAM: Convolutional block attention module[C]. 15th European Conference on Computer Vision, 2018: 3-19.
- [24] HOU Q, ZHOU D, FENG J, et al. Coordinate attention for efficient mobile network design[C]. 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021:13708-13717.
- [25] LI C, LI L, JIANG H, et al. YOLOv6: A single-stage object detection framework for industrial applications[J]. ArXiv preprint arXiv: 2209.02976, 2022.
- [26] LI H, LI J, WEI H, et al. Slim-neck by GSConv: A better design paradigm of detector architectures for autonomous vehicles[J]. ArXiv preprint arXiv: 2206.02424, 2022.
- [27] WOJKE N, BEWLEY A, PAULUS D. Simple online and real time tracking with a deep association metric[C]. 2017 IEEE International Conference on Image Processing (ICIP). IEEE, 2017: 3645-3649.
- [28] YANG M, SOWMYA A. An underwater color image quality evaluation metric[J]. IEEE Transactions on Image Processing, 2015, 24(12): 6062-6071.
- [29] PANETTA K, GAO C, AGAIAN S. Human-visual-system-inspired underwater image quality measures[J]. IEEE Journal of Oceanic Engineering, 2016, 41 (3): 541-551.

## 作者简介

吴佳文, 硕士研究生, 主要研究方向为图像处理。

E-mail: wujiawen1712@163.com

姬伟(通信作者), 博士, 教授, 主要研究方向为农业机器人关键控制技术研究、农业机械装备自动控制与导航技术、生产过程的综合自动化控制系统等。

E-mail: jiwei@ujs.edu.cn

翟科龙, 硕士研究生, 主要研究方向为图像处理。

E-mail: zkl18015259276@163.com

许波, 博士, 副教授, 主要研究方向为信号检测与处理、图像识别与处理等。

E-mail: xubo@ujs.edu.cn