doi:10.6041/j.issn.1000-1298.2025.11.058

基于改进 YOLO v8 的南美白对虾虾苗跟踪方法

陈　明[1,2]　甘冬梅[1,2]　卢　鹏[1,3]　顾　浩[1,2]　栗　征[2,4]

(1. 上海海洋大学信息学院, 上海 201306; 2. 农业农村部渔业信息重点实验室, 上海 201306;
3. 海洋测绘与智能分析实验室, 上海 201306; 4. 上海时鳌信息科技有限公司, 上海 201306)

摘要: 虾苗养殖的智能化在提高养殖效率和降低人工成本方面展现出巨大潜力,虾苗跟踪技术对于深入研究虾苗的活力和质量具有重要价值。由于虾苗体积微小、身体半透明且密度高,传统方法面临许多挑战:小尺寸目标易于被漏检,高密度下检测精度低,以及因虾苗外表相似、移动速度快导致跟踪误差大。针对此问题,选取 P15 阶段的南美白对虾虾苗作为研究对象,提出一种基于改进 YOLO v8 的南美白对虾虾苗跟踪方法。选取改进 YOLO v8 网络作为检测器,使用 CARAFE 代替 YOLO v8 的 Upsample 上采样层,CARAFE 自适应上采样保留虾苗的半透明边缘和形态特征,减少失真;结合 BiFormer 注意力机制和 C2f 提出 C2f - BFB 模块优化骨干网络,在保持模型轻量化的同时,提升模型在处理小目标时的精度;将原模型的 CIoU 损失替换成 Wise - Focaler - ShapeIoU 损失函数,进一步提升了检测精度。提出了一种融合距离匹配算法和角度匹配算法的匹配策略,并引入金字塔 LK 光流法,从而构建一个准确又稳定的跟踪器,有效解决了因虾苗移动快、外表相似导致的跟踪误差问题。实验结果表明,改进的 YOLO v8 的平均精度均值(mAP@0.5)比原模型提高 1.1 个百分点,召回率提升 2.9 个百分点,显著减少了漏检情况并提高了检测精度。本文所提跟踪器的 MOTA 为 74.52%、MOTP 为 75.62%、IDF1 为 76.42%,与 DeepSORT、SORT、ByteTrack 算法相比,MOTA 显著提高,跟踪效果良好。本研究为自动化虾苗养殖提供了有力的技术支持,也为相关领域的研究提供了思路和方法。

关键词: 虾苗; YOLO v8; 多目标跟踪; 金字塔 Lucas - Kanade 光流法

中图分类号: TP391.41; S951.2　　**文献标识码**: A　　**文章编号**: 1000-1298(2025)11-0599-13　　**OSID**:

Tracking Method for *Litopenaeus vannamei* Shrimp Larvae Based on Improved YOLO v8

CHEN Ming[1,2]　GAN Dongmei[1,2]　LU Peng[1,3]　GU Hao[1,2]　LI Zheng[2,4]

(1. *College of Information Technology*, *Shanghai Ocean University*, *Shanghai* 201306, *China*
2. *Key Laboratory of Fisheries Information*, *Ministry of Agriculture and Rural Affairs*, *Shanghai* 201306, *China*
3. *Laboratory for Marine Surveying and Mapping with Intelligent Analysis*, *Shanghai* 201306, *China*
4. *Shanghai Shilin Information Technology Co.*, *Ltd.*, *Shanghai* 201306, *China*)

Abstract: The intelligent automation of shrimp larvae farming demonstrates great potential in improving farming efficiency and reducing labor costs. Tracking technology for shrimp larvae is of significant value for in-depth research on their vitality and quality. Due to the small size, semi-transparency, and high density of shrimp larvae, traditional methods often face many challenges: small targets are easily missed, detection accuracy is low under high density, and tracking errors are high due to the similar appearance and fast movement of shrimp larvae. To address these issues, focusing on P15 stage Pacific white shrimp larvae, a tracking method was proposed based on an improved YOLO v8. Firstly, an improved YOLO v8 network was selected as the detector, with CARAFE replacing YOLO v8's Upsample layer. CARAFE's adaptive upsampling preserved the semi-transparent edges and morphological features of shrimp larvae, reducing distortion. The backbone network was optimized by using the BiFormer attention mechanism and the C2f - BFB module, which enhanced precision in handling small targets while keeping the model

收稿日期: 2024-07-17　修回日期: 2024-09-29

基金项目: 广东省重点领域研发计划项目(2021B0202070001)

作者简介: 陈明(1966—),男,教授,博士,主要从事水产物联网和数据挖掘研究,E-mail: mchen@shou.edu.cn

通信作者: 卢鹏(1981—),男,副教授,博士,主要从事深度学习算法以及图像处理应用研究,E-mail: plu@shou.edu.cn

lightweight. Finally, the original model's CIoU loss was replaced with the Wise - Focaler - ShapeIoU loss function, further improving detection accuracy. A matching strategy that combined distance and angle matching algorithms was also proposed, and the pyramid LK optical flow method was introduced, constructing an accurate and stable tracker that effectively addressed tracking errors caused by the fast movement and similar appearance of shrimp larvae. Experimental results showed that the improved YOLO v8 achieved a 1.1 percentage points increase in mAP@0.5 and a 2.9 percentage points increase in recall rate compared with that of the original model, significantly reducing missed detections and improving detection accuracy. The proposed tracker achieved a MOTA of 74.52%, MOTP of 75.62%, and IDF1 of 76.42%. Compared with DeepSORT, SORT, and ByteTrack algorithms, MOTA was significantly improved, demonstrating excellent tracking performance. The research result can provide strong technical support for the automation of shrimp larvae farming and offer ideas and methods for related research fields.

Key words: shrimp larvae; YOLO v8; multi target tracking; pyramid Lucas - Kanade optical flow method

0 引言

南美白对虾(*Litopenaeus vannamei*)是全球养殖量最大的虾种之一,在中国及世界均占有重要地位[1],据《2023 中国渔业统计年鉴》,2022 年我国南美白对虾总产量约 2.1×10^6 t,约占全球三分之一[2]。科学的虾苗养殖是保证产品质量与产量的关键,虾苗活动情况可反映其健康状况[3],目前多数养殖厂仍依赖人工肉眼观察评估虾苗活力和质量[4],该方法主观且耗时。通过对虾苗进行跟踪,可获得其游动轨迹和速度等信息[5],是研究虾苗活动的重要方式。

传统机器学习在水产动物跟踪中具有一定优势[6]。MORAIS 等[7]使用基于贝叶斯滤波算法的方法实现水下鱼类跟踪。沈蔚等[8]利用回声阈值结合 $\alpha-\beta$ 轨迹跟踪算法,在声学数据处理上实现了对鱼类的精确追踪。但复杂的特征提取和小目标识别难度限制了其应用。

深度学习多目标跟踪算法(Multi-object tracking, MOT)因其非接触、高精度、可量化的优点,逐渐应用于水产动物跟踪研究领域[9],其中,基于检测的跟踪(Detection-based tracking, DBT)因逐帧检测目标,具有更高稳定性,成为主流方法;而基于无检测的跟踪(Detection-free tracking, DFT)需人工标注初始帧,应用较少。针对水产动物的目标跟踪方法主要采用 DBT,SORT[10]是较早的 DBT 算法,ZHANG 等[11]采用 YOLO v5 - Nano 和改进 SORT 算法跟踪鱼苗,但 SORT 算法在目标运动状态不稳定时存在跟丢的情况。针对该问题,DeepSORT[12]在 SORT 基础上引入运动与外观信息,洪宇等[13]利用 YOLO v4 和 DeepSORT 跟踪海中的鱼类,但在复杂背景和长时视频中表现不足。XU 等[14]提出一种 Faster R - CNN 和 KCF 算法跟踪海参和扇贝,然而 Faster R - CNN 模型具有很高的计算复杂度,无法对目标实时跟踪。赵海翔等[15]验证了基于 YOLO v8 和 Bytetrack 的多目标跟踪方法在斑马鱼苗视频中的有效性,但其使用的数据集数量少,无法验证该方法的有效性。

大多数算法会考虑目标外观特征以提升准确率[16-17],但当目标外观相似时,依赖外观特征反而会干扰结果。安志强等[18]针对海珍品外观高度相似,借鉴 BYTE 算法,舍弃外观信息,仅用运动与位置信息,并结合 YOLO v7,实现多类别海珍品精确跟踪。申阳等[6]也同样考虑鱼苗相似性,删除 BoT - SORT 的外观关联部分,结合改进 YOLO v8,有效实现鱼苗跟踪。由于海珍品移动缓慢、鱼苗在滑道上匀速运动,上述方法能保证较高准确率,但未充分考虑目标快速运动的情况。

虾苗因体型小、半透明,易在检测中遗漏;其外观相似、姿态差异小且运动迅速,也易引发跟踪错误。本文基于 DBT 提出改进 YOLO v8 方法:在检测器中优化主干网络、上采样与回归损失;在跟踪器中结合运动与位置信息,设计距离与角度融合匹配,并用金字塔 LK 光流预测轨迹;最后在不同光照下采集视频进行对比实验。

1 虾苗跟踪算法

采用改进 YOLO v8 算法作为检测器,并结合虾苗的特点设计了一种跟踪算法作为跟踪器。图 1 为虾苗跟踪技术路线,检测器对视频序列进行目标检测,并将检测结果作为跟踪器的输入。在跟踪器中,先提取虾苗位置信息和光流信息,光流信息对虾苗运动状态的详细描述可称为运动信息,再根据这两种信息计算前后 2 帧目标的匹配程度,以此实现对虾苗的跟踪。

1.1 检测器

在水产养殖领域,YOLO 系列因其出色性能被

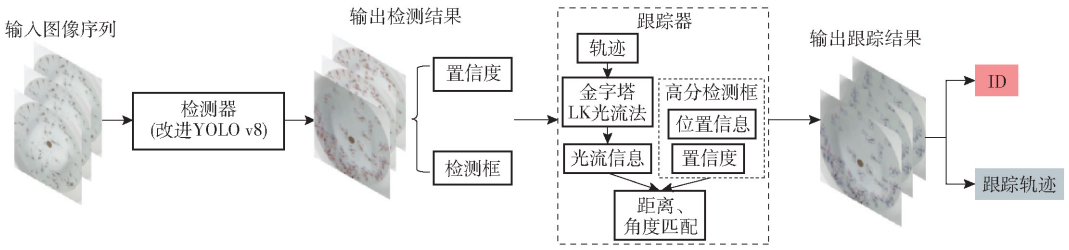


图 1　虾苗跟踪技术路线

Fig. 1　Technical route for tracking shrimp larvae

广泛应用于跟踪算法的检测器，例如 XU 等[19]提出了 DT－YOLO v5 算法用于海参识别，并基于自动帧匹配坐标实现多目标行为跟踪。SAAD 等[20]将 StereoYolo 与 DeepSORT 结合，提出了鱼类跟踪框架。LIU 等[21]将 YOLO v3 与并行相关滤波器结合，实现了鱼类实时跟踪。YOLO 系列检测器不仅在鱼类识别中表现出色，在虾的识别任务中也有良好效果，LU 等[22]利用 YOLO 算法识别和跟踪包括虾在内的海洋生物，取得了显著效果。因此，本文选择 YOLO 系列的 YOLO v8 作为虾苗检测器。

YOLO v8 的骨干网使用的是 CSPDarknet53 结构，其中 C2f 和 SPPF 结构通过卷积、池化等方式提取图像特征。颈部（Neck）部分使用的特征融合方法为 FPN（Feature pyramid networks）＋PAN（Path aggregation network），FPN 旨在解决同时检测不同尺寸物体的问题，并在此基础上添加自下而上的 PAN，FPN 自上而下传达了强烈的语义特征，而 PAN 自下而上传达了强烈的定位特征。头部（Head）部分使用 YOLOX 的解耦头部操作进行分类和定位分离，也从 Anchor－Based 转变为 Anchor－Free。损失计算涵盖分类损失与回归损失，采用 BCE（Binary cross-entropy）计算分类损失，使用 DFL（Distribution focal loss）[23]和 CIoU[24]损失函数计算回归损失。

本文对 YOLO v8 进行了针对性的改进以适应虾苗检测的特定需求。改进模型结构如图 2 所示，改进部分通过虚线边框标注。本文在网络的主干部分使用提出的 C2f－BFB 模块替换 C2f 模块；在 Neck 部分，将原模型 Upsample 上采样层替换成 CARAFE；最后采用 Wise－Focaler－ShapeIoU 损失函数替换原模型的 CIoU 损失函数。

1.1.1　主干网络改进

虾苗目标较小，在图像中占据的像素往往很少，易被高层次特征忽略。BiFormer[25]通过双层路由注意力机制（Bi-level routing attention，BRA）能够在不同尺度上捕捉图像中特征，这意味着其可以同时关注到虾苗的细微特征和更大范围的背景信息，从而减少漏检的可能性。BRA 机制在处理特征时，不仅仅依赖于高层次的语义信息，还能够从低层次的视

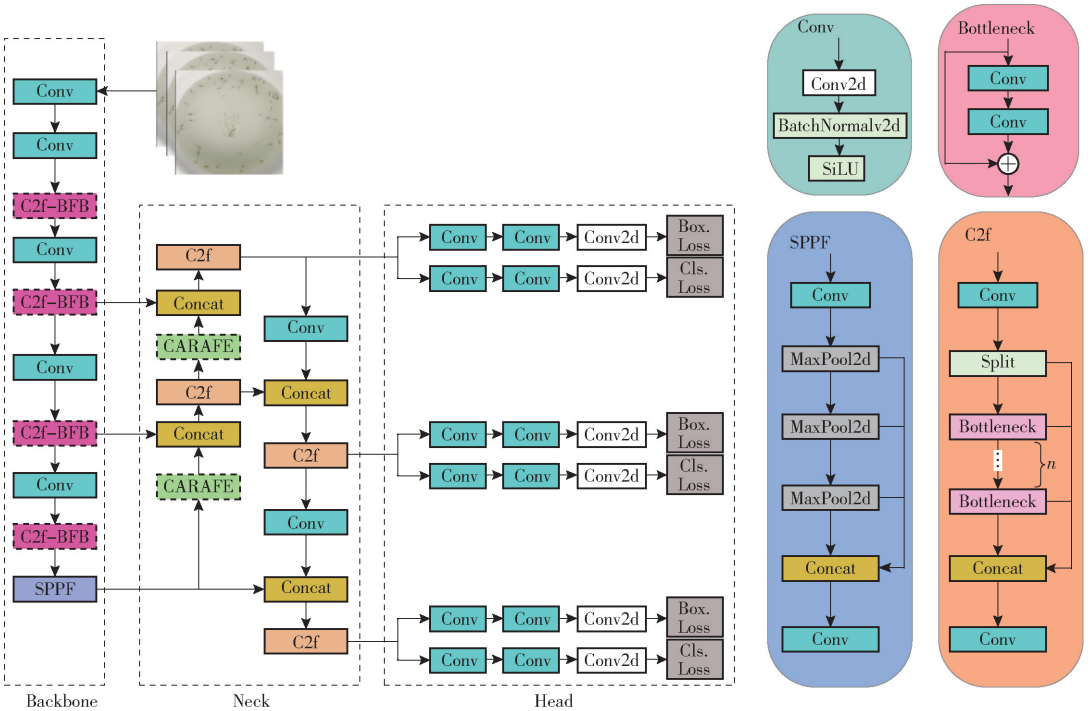


图 2　改进 YOLO v8 模型结构图

Fig. 2　Improved YOLO v8 network model structure diagram

觉特征过渡到高层次特征。这种过渡使得模型在检测小目标时,能够更好地保留关键细节,提升模型对虾苗的识别精度。

如图 3a 所示,H、W 表示输入特征的高度和宽度,C 表示特征图的通道数,X 表示特征图,S 表示划分区域的步长,$\boldsymbol{Q}$、$\boldsymbol{K}$、$\boldsymbol{V}$ 表示由全连接层投影得到的查询(Query)、键(Key)、值(Value)矩阵。双层路由注意力机制首先将输入特征图 $X \in \mathbf{R}^{H\times W\times C}$ 划分成 $S\times S$ 个非重叠区域,使得 X 转化为 $X^r \in \mathbf{R}^{S^2\times HWC/S^r}$;然后采用全连接层对 X^r 进行线性投影,得到 $\boldsymbol{Q}$、$\boldsymbol{K}$ 和 $\boldsymbol{V}$,并计算 $\boldsymbol{Q}$、$\boldsymbol{K}$ 的均值分别得到 $\boldsymbol{Q}^r$、$\boldsymbol{K}^r$;再计算邻接矩阵 $\boldsymbol{A}^r = \boldsymbol{Q}^r(\boldsymbol{K}^r)^{\mathrm{T}}$,衡量不同区域之间的语义相似度。

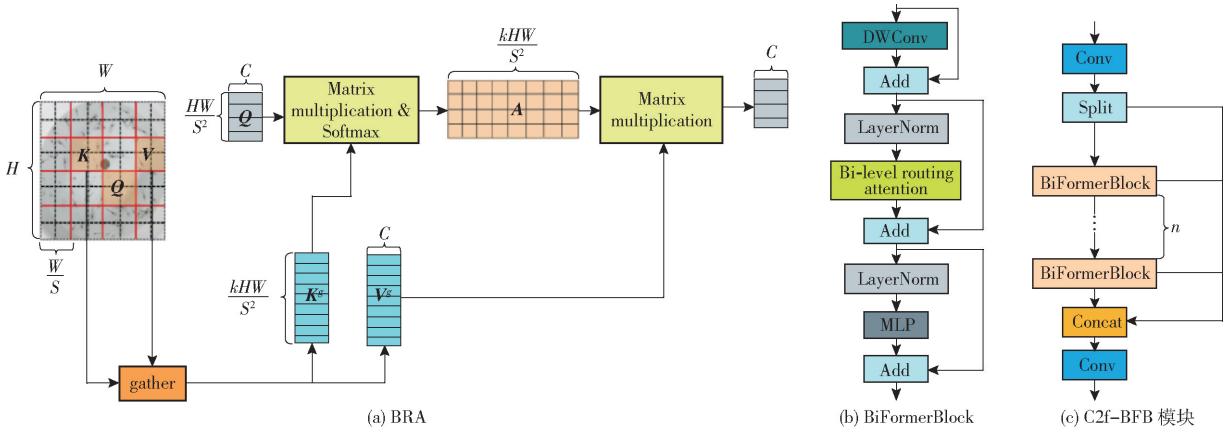


图 3 C2f－BFB 模块及结构细节

Fig. 3 C2f－BFB module and structural details

BRA 通过在粗糙区域首先过滤掉大部分不相关的键值对,有效去除了冗余信息,只保留少部分关键的路由区域,在初步筛选后,再在保留的路由区域中应用更精细的注意力机制。这种逐层精炼过程,使得 BRA 能够在保持检测精度的同时,显著降低计算量。相比直接在全局范围内使用自注意力机制,这种方法避免了处理大量无关信息的开销。BiFormer 结合 BRA 构建的 BiFormerBlock 结构如图 3b 所示,在开始时使用深度可分离卷积(DWConv)隐式编码输入的相对位置信息,然后依次使用 BRA 和多层感知机(MLP)模块对输入的信息进行跨位置关系建模和逐位置嵌入。

将 C2f 模块中的 Bottleneck 替换为 BiFormerBlock,从而提出了新模块 C2f－BFB(图 3c),这一替换目的是利用 BiFormerBlock 的 BRA 和高效的特征建模能力,增强主干网络对多尺度特征的捕捉和处理能力。通过在 C2f 模块中引入 BiFormerBlock,在保持模型轻量的同时,提升模型在处理小目标时的精度。

1.1.2 Neck 改进

虾苗的半透明性会导致其在图像中与其他虾苗或背景难以区分,这种半透明性使得虾苗在与背景颜色相似时,边缘模糊,难以准确分辨和跟踪。YOLO v8 的上采样(Upsample)层以最近邻插值方法为主,仅简单复制临近像素值,易导致特征失真和信息损失,特别是对于细节丰富的图像及小目标的表达能力有限。而 CARAFE[26] 可以根据图像的内容自适应地生成上采样核,这使得其可以更好地保留图像中的重要特征,减少因上采样而可能引入的失真或模糊,这对于半透明的虾苗图像特别重要,因为这些图像中的细节通常较为复杂和微小。CARAFE 能够在大范围内聚合信息,这意味着其可以捕捉到更广泛的上下文信息,有助于识别和保留细微的图像细节,如虾苗的边缘和形态特征。

CARAFE 流程如图 4 所示,包括上采样核预测模块和内容感知重组模块。首先压缩特征图以减少计算量,然后进行内容编码和上采样核预测,生成标准化的上采样核。特征重组模块将每个位置映射回输入特征图,并利用预测的上采样核进行点积操作,以生成输出特征图。

1.1.3 优化损失函数

YOLO v8 的回归损失 CIoU 在某些情况下无法充分应对数据集中的低质量示例和样本不平衡问题。为此本文提出了 Wise－Focaler－ShapeIoU 损失函数替换 CIoU,进一步优化了损失函数。

(1) Focaler－IoU

现有 IoU 损失函数通过利用边界框之间的几何关系来提高回归性能,而忽略了难易样本分布对边界框回归的影响,Focaler－IoU[27] 关注不同难度的回归样本,以提高检测器在各种检测任务中的性能。Focaler－IoU 公式为

$$IoU_{focaler} = \begin{cases} 0 & (IoU < a) \\ \dfrac{IoU - a}{u - a} & (a \leqslant IoU \leqslant u) \\ 1 & (IoU > u) \end{cases} \tag{1}$$

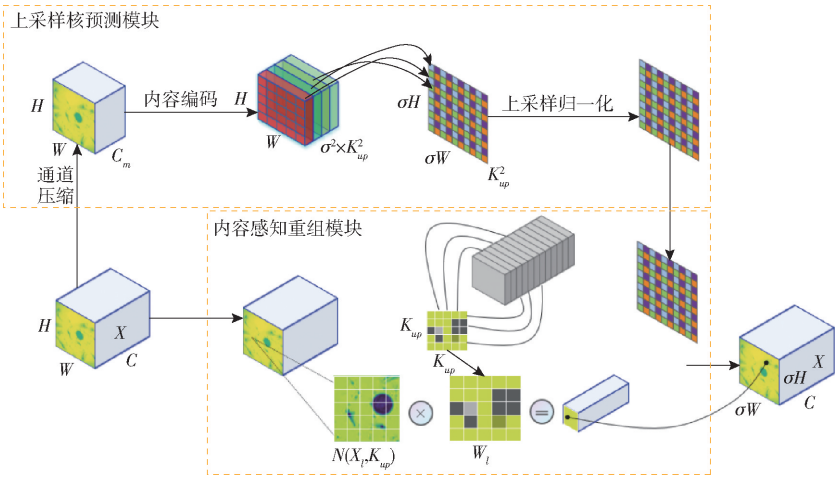


图 4　CARAFE 流程图

Fig. 4　CARAFE process diagram

式中　*IoU*——真实框和预测框的交并比

a——下限阈值　　*u*——上限阈值

该设计使损失函数对 IoU 值的某些区间段敏感，从而更专注于中等重叠度的预测边界框与真实边界框的样本。这有助于模型从中等难度的样本中更好地学习特征，而不是仅专注于最易或最难的样本。

(2) Focaler－ShapeIoU

现有的边界框回归方法通常考虑 GT 框和预测框之间的几何关系，使用相对位置和形状计算损失，但忽略了边界框本身的形状、尺度对回归结果的影响。小尺度真实框对 IoU 值变化更敏感，特别是在虾苗这样的小目标中，不同形状的真实框 IoU 值的偏差较大。Shape－IoU[28] 方法可以关注目标框本身的形状和尺度来计算损失，从而提高精度。如图 5 所示，Shape－IoU 的计算过程为

$$ww=\frac{2(w^{gt})^{scale}}{(w^{gt})^{scale}+(h^{gt})^{scale}} \tag{2}$$

$$hh=\frac{2(h^{gt})^{scale}}{(w^{gt})^{scale}+(h^{gt})^{scale}} \tag{3}$$

$$distance^{shape}=hh(x_c-x_c^{gt})^2/c^2+ww(y_c-y_c^{gt})^2/c^2 \tag{4}$$

$$\Omega^{shape}=\sum_{t=w,h}(1-e^{w_t})^{\theta}\quad(\theta=4) \tag{5}$$

$$\begin{cases}w_w=hh\dfrac{|w-w^{gt}|}{\max(w,w^{gt})}\\ w_h=ww\dfrac{|h-h^{gt}|}{\max(h,h^{gt})}\end{cases} \tag{6}$$

式中　w^{gt}、h^{gt}——真实框宽、高

x_c^{gt}、y_c^{gt}——真实框中心点水平坐标、垂直坐标

w、h——先验框(也称锚框)宽、高

scale——缩放因子，与数据集中目标的尺度有关

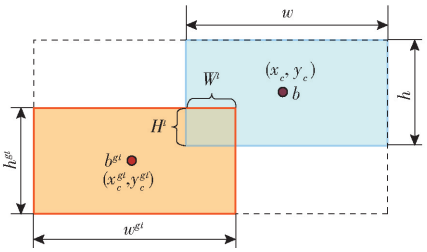


图 5　预测框和真实框交并比示意图

Fig. 5　Illustration of intersection over union between predicted boxes and ground truth boxes

w_w、w_h——宽度方向、高度方向的差异权重

x_c、y_c——先验框中心点横向、纵向位置

ww、*hh*——水平方向、竖直方向的权重系数，其取值与真实框形状有关

$distance^{shape}$——用于衡量预测边界框与真实边界框之间几何差异的加权距离度量

Ω^{shape}——对宽高误差进行非线性压缩后的总和

为了融合 Focaler－IoU 和 Shape－IoU 的优势，本文将 Focaler－IoU 的思想引入 Shape－IoU 的 IoU 计算中，优化新的损失函数以同时关注中等难度的样本和边界框的形状尺度特性，对应的边框回归损失为

$$L_{Focaler-ShapeIoU}=1-IoU_{focaler}+distance^{shape}+0.5\Omega^{shape} \tag{7}$$

(3) Wise－Focaler－ShapeIoU

WIoU[29] 能有效解决样本不平衡问题并加快模型收敛速度，提高检测精度。为了进一步优化模型性能，定义离群度 β 描述锚框质量，其定义为

$$\beta=\frac{L_{IoU}^{*}}{\overline{L}_{IoU}}\in[0,+\infty) \tag{8}$$

式中　L_{IoU}^{*}——单调聚焦系数

$\overline{L}_{IoU}$——动量为 m 的滑动平均值

利用 β 构建非单调聚焦系数 r，并与 Focaler－

ShapeIoU 损失函数相乘形成 Wise - Focaler - ShapeIoU,其定义为

$$L_{Wise-Focaler-ShapeIoU}=rL_{Focaler-ShapeIoU} \tag{9}$$

$$r=\frac{\beta}{\delta\alpha^{\beta-\delta}} \tag{10}$$

式中 δ、α——超参数,通过设置不同的值以拟合不同的模型

综上所述,Wise - Focaler - ShapeIoU 损失函数结合了 Focaler - IoU 的样本关注度、Shape - IoU 的形状尺度敏感性和 WIoU 的样本质量评估,为边界框回归问题提供了更全面高效的解决方案。

1.2 跟踪器

1.2.1 跟踪器设计

本文的跟踪算法核心思想是使用金字塔 LK 光流法预测下一帧的轨迹,并使用距离匹配算法和角度匹配算法逐帧进行数据关联。本文跟踪算法的一大优势在于,其无需依赖虾苗的外观特征(如颜色、纹理等),仅利用光流矢量(即运动信息),即可有效解决因虾苗外表相似、快速移动而难以准确跟踪的问题。本文跟踪算法具体流程如图 6 所示。

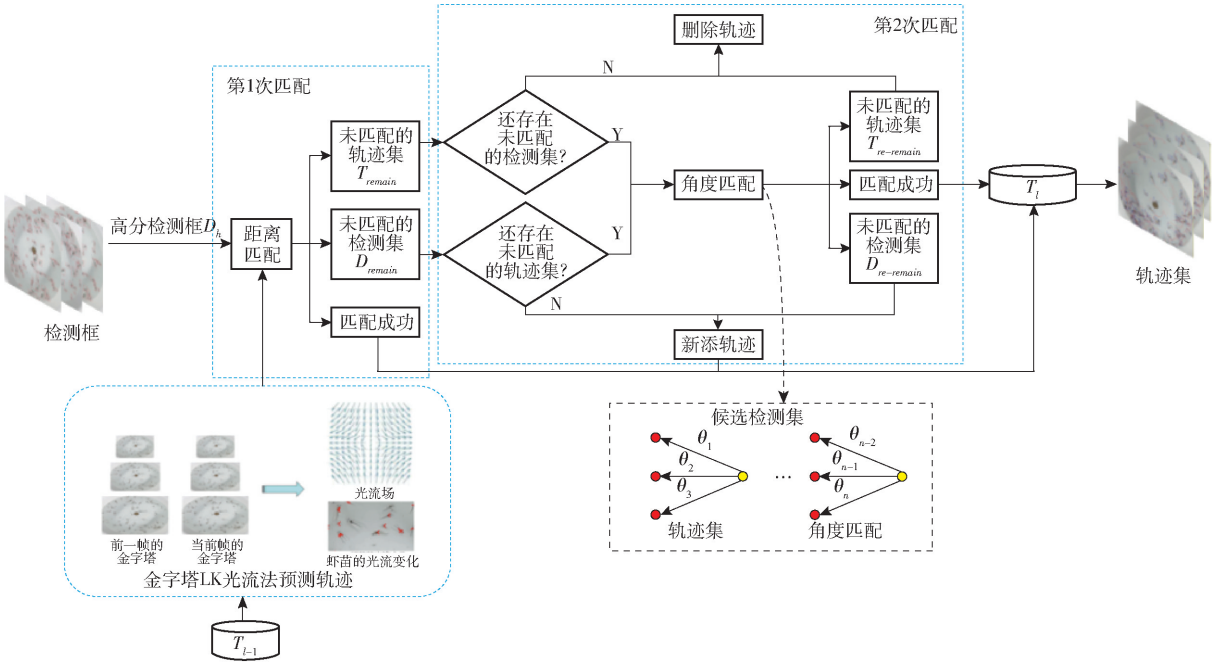


图 6 跟踪流程图

Fig. 6 Tracking process diagram

本文算法的伪代码如下:

输入:视频序列 V;检测器 Det

输出:跟踪轨迹 T

初始化:$T\leftarrow\phi$

for f_k in V do

  $D_k\leftarrow Det(f_k)$

  $D_h\leftarrow\phi$

  $T_k\leftarrow\phi$

  for d in D_k do //获取高分检测框

    if $d.score>0.25$ then

      $D_h\leftarrow D_h\cup\{d\}$

    end if

  end for

  for t in T do //预测轨迹的新位置

    $t\leftarrow OpticalFlowPyrLK(t)$

    $T_k\leftarrow T_k\cup\{t\}$

  end for

  /* 第 1 次匹配 */

  T_k 和 D_h 进行距离匹配

  $D_{best-match}\leftarrow D_h$ 中匹配成功的框

  $D_{remain}\leftarrow D_h$ 中未匹配成功的框

  $T_{remain}\leftarrow T$ 中未匹配成功的轨迹

  $T\leftarrow update(D_{best-match})$ //更新轨迹

  /* 第 2 次匹配 */

  if $D_{remain}\neq None$ and $T_{remain}\neq None$ then

    for t in T_{remain} do

      /* 依次选取 3 个距离最小的候选框 */

      $D_{c3}\leftarrow selT3Cds(t,D_{remain})$

      D_{c3} 和 t 进行角度匹配

    End for

    $D_{re-best-match}\leftarrow D_{remain}$ 中匹配成功的框

    $D_{re-remain}\leftarrow D_{remain}$ 中未匹配成功的框

    $T_{re-remain}\leftarrow T_{remain}$ 中未匹配成功的轨迹

    $T\leftarrow update(D_{re-best-match})$

    $T\leftarrow T\backslash T_{re-remain}$ //删除未成功匹配轨迹

    for d in $D_{re-remain}$ do //初始化为新的轨迹

$T \leftarrow T \cup \{d\}$

end for

elif $D_{remain} = None$ then

$T \leftarrow T \backslash T_{remain}$　　//删除未成功匹配轨迹

elif $T_{remain} = None$ then

for d in D_{remain} do　　//初始化为新的轨迹

$T \leftarrow T \cup \{d\}$

end for

end if

end for

Return T

算法首先选取置信度大于 0.25 的检测框作为高分检测框 D_h。同时，采用金字塔 LK 光流法预测 T 中每个轨迹的当前帧中的新位置。

第 1 次匹配中，将预测轨迹 T_k 和当前帧高分检测框 D_h 进行距离匹配。匹配结果只存在不匹配的轨迹 T_{remain} 或只存在未匹配的检测集 D_{remain} 时直接执行伪代码 33 ~ 38 行。若两者同时存在，表明可能存在快速移动的目标导致匹配失败，此时需要进行第 2 次匹配。

第 2 次匹配中，从 D_{remain} 中为 T_{remain} 中的每个预测框筛选出 3 个距离最近的候选检测框，并进行角度匹配。对于仅剩余的未匹配轨迹集 $T_{re-remain}$，视为失去跟踪目标并删除；对于仅剩余的未匹配检测集 $D_{re-rematch}$，则将其视为新出现的目标并初始化为新的轨迹。

1.2.2　数据匹配算法

(1)距离匹配

本文的距离匹配算法具体思路是：通过计算上一帧预测轨迹中的目标框中心点和当前帧检测集中的目标框中心点的欧氏距离，两者按照欧氏距离最小原则进行匹配，计算公式为

$$D(p,q) = \sqrt{(x_i^p - x_j^q)^2 + (y_i^p - y_j^q)^2} \tag{11}$$

式中　p——第 $t-1$ 帧中第 i 个预测目标

q——第 t 帧中第 j 个检测目标

其中，设置一个距离阈值，只有当欧氏距离小于这个阈值时，才认为该匹配是有效的。

为节省计算资源，引入了基于 $k-D$ 树的最近邻搜索算法。$k-D$ 树是一种空间划分数据结构，通过递归地将数据点划分到不同的维度上，从而构建一个 k 维空间的二叉树。在本文中，具体实现为：首先计算当前帧的检测目标中心点 x 和 y 维度的方差，选择方差最大的维度特征作为根节点构建成一个 $k-D$ 树，再对于每个预测目标使用 $k-D$ 树进行最近邻搜索，找到距离最近的检测目标中心点。

经观察发现，虾苗在运动过程中可能出现突然的快速移动现象，如图 7 所示，当第 $t-1$ 帧中虾苗 I 快速移动到第 t 帧中的位置后(图中 d_1、d_2 为两虾苗之间的距离，θ_{deg} 为游动角度)，根据距离匹配原则，第 $t-1$ 帧中的虾苗 I 会匹配到第 t 帧的虾苗 J，出现匹配错误，因此需要设置适当小的距离阈值以避免这种匹配错误，然而，设置过小的阈值导致第 $t-1$ 帧中的虾苗 I 和第 t 帧中的虾苗 I 无法正确匹配，此时则出现伪代码第 20 行的情况，需要进行第 2 次匹配。

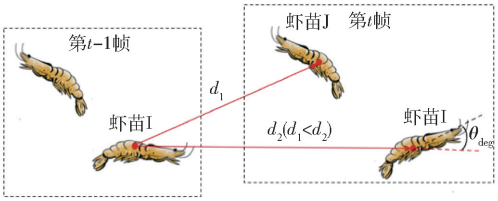


图 7　虾苗运动示意图

Fig. 7　Schematic of shrimp larvae movement

(2)角度匹配

针对虾苗快速移动可能导致匹配失败的问题，本文提出了一种角度匹配算法。图 8 为虾苗在连续 10 帧内的跟踪情况。观察发现，当虾苗快速移动时，其游动方向与轨迹几乎重合为同一直线，即快速移动时虾苗的游动角度变化不大。

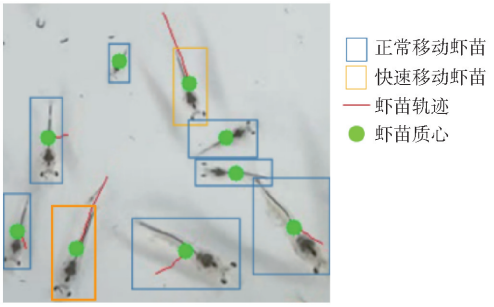


图 8　虾苗跟踪示意图

Fig. 8　Tracking diagram of shrimp larvae

如图 9 所示，统计了每只虾苗的游动角度和速度情况。其中，速度定义为前后 2 帧虾苗中心点像素点位移。根据图 9 可知，当速度增大时，虾苗游动角度变小，这进一步证实了快速游动的虾苗具有较小的游动角度变化特性。基于这一发现，本文提出了角度匹配算法，旨在通过识别虾苗游动角度的稳定模式来优化匹配过程，从而解决虾苗快速移动可能导致的匹配失败问题。

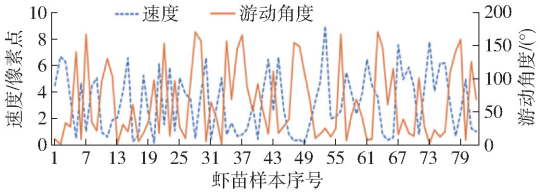


图 9　游动角度和速度统计图

Fig. 9　Statistics of swimming angle and speed

角度匹配算法的具体思路如下:首先为目标选取 3 个距离最近的候选匹配目标,计算它们的游动角度;然后根据游动角度的最小变化原则进行匹配,游动角度公式为

$$\theta_{\deg}=\frac{180}{\pi}\arccos\left(\frac{\boldsymbol{v}_1\boldsymbol{v}_2}{|\boldsymbol{v}_1||\boldsymbol{v}_2|}\right) \tag{12}$$

式中 $\boldsymbol{v}_1$——虾苗方向向量

$\boldsymbol{v}_2$——虾苗移动方向向量

1.2.3 金字塔 Lucas – Kanade 光流法

金字塔 Lucas – Kanade 光流法(Pyramidal Lucas – Kanade optical flow)[30]是一种广泛应用于计算机视觉中的光流估计算法,特别适用于存在大位移情况时进行运动估计。虾苗的快速移动导致的大位移是跟踪过程中的一大挑战,金字塔 Lucas – Kanade 光流法在高层图像中,快速移动的目标可表现为较小位移,这使得跟踪算法更易跟上目标运动轨迹。这个方法在基本的 Lucas – Kanade 光流法[31]的基础上引入了图像金字塔结构,使其能够处理更大范围的运动。

Lucas – Kanade 光流法涉及 3 个假设:①亮度恒定。②时间连续或运动是一个小运动。③空间一致。从假设①可以推导出光流限制方程为

$$I(x,y,t)=I(x+\mathrm{d}x,y+\mathrm{d}y,t+\mathrm{d}t) \tag{13}$$

式中 $I(x,y,t)$——点(x,y)在时间 t 的亮度

根据假设②,即假设像素移动足够小,可以将泰勒公式应用于公式(13),得到

$$I_xv_x+I_yv_y+I_t=0 \tag{14}$$

其中 $v_x=\mathrm{d}x/\mathrm{d}t$ $v_y=\mathrm{d}y/\mathrm{d}t$

式中 v_x、v_y——水平、垂直方向上的光流速度

变换为矩阵形式

$$\begin{bmatrix}I_{x_1} & I_{y_1}\\ I_{x_2} & I_{y_2}\\ \vdots & \vdots\\ I_{x_n} & I_{y_n}\end{bmatrix}\begin{bmatrix}v_x\\ v_y\end{bmatrix}=-\begin{bmatrix}I_{t_1}\\ I_{t_2}\\ \vdots\\ I_{t_n}\end{bmatrix} \tag{15}$$

使用最小二乘法求解可得

$$\begin{bmatrix}v_x\\ v_y\end{bmatrix}=\begin{bmatrix}\sum I_x^2 & \sum I_xI_y\\ \sum I_xI_y & \sum I_y^2\end{bmatrix}\begin{bmatrix}I_xI_t\\ I_yI_t\end{bmatrix} \tag{16}$$

当图像中的位移较大时,不满足假设②,Lucas – Kanade 方法无法准确估计运动,为解决这个问题,引入图像金字塔概念,如图 10 所示。金字塔通过逐层缩小原始图像,直至顶层图像分辨率足够小,像素移动足够小,满足 LK 光流法的假设②。从金字塔顶层开始估计光流,使用其作为顶层迭代的初始值计算顶层的运动矢量。然后,前一层的跟踪结果用作下一层光流迭代的初始值,依此类推,直至达到金字塔最低层(原始图像)。这一过程获取最终的运动矢量,实现相邻图像之间的关键点匹配。

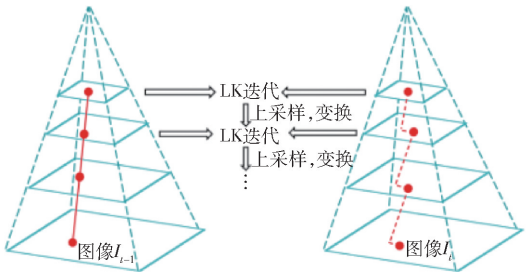


图 10 图像金字塔

Fig. 10 Image pyramid

2 实验

2.1 数据集和评价指标

2.1.1 数据集构建

南美白对虾的生长发育阶段包括无节幼体、蚤状幼体、糠虾幼体和仔虾等多个阶段,其中仔虾阶段可进一步细分为 P1 ~ P20 等多个阶段。在养殖业中,工人通常在仔虾阶段通过观察虾苗的行为、活力等指标来判断虾苗的健康状态,此时的虾苗已具备一定形态特征和辨识性。因此,本文选用 P15 阶段的南美白对虾虾苗作为实验对象,该阶段的虾苗尺寸约 1.5 cm。

在上海海洋大学滨海养殖基地建立了对虾虾苗视频采集平台。如图 11 所示,该平台主要包括水箱(半径 37 cm,水深 40 cm)、摄像机、计算机和光源等设备。所有视频均采用俯拍视角,摄像机距水面 50 cm。实验开始前,先向水箱中加入养殖场备用的养殖水,再放入虾苗。光源设备具有 3 个可调节挡位,一挡为相对较低的光照强度,为水箱内提供了弱光环境;二挡偏暖色调,提供中等光照强度;三挡为较高光照强度,提供明亮且偏冷光照环境,在拍摄期间,通过调节光源设备的挡位获取不同强度光照下的视频。

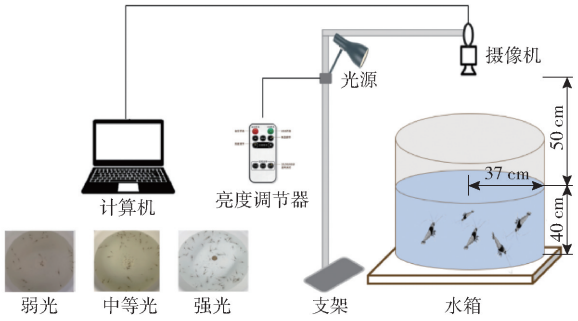


图 11 数据采集示意图

Fig. 11 Data collection diagram

虾苗的活动通常有正常游动、逆水游动和应激反应。当健康虾苗受到水流刺激时,其会绕着水箱边缘进行逆时针或顺时针的逆水游动。对外界刺激

十分敏感的虾苗在容器受到敲击时，会迅速散开，表现出明显应激反应。在视频拍摄过程中，不仅收集了虾苗正常游动的视频，还通过手动制造水流来记录其逆水游动行为，同时通过敲击水箱获取其应激反应视频。

对视频进行抽帧并剔除不合格图像后，得到 786 幅图像用于检测，为消除不必要的背景干扰，通过裁剪获得分辨率为 1 080 像素 ×1 080 像素。通过 LabelImg 软件进行虾苗图像标注，将数据集按比例 7∶2∶1划分为训练集、验证集、测试集。测试数据集如表 1 所示，为评估多目标跟踪算法，选取 5 段不同光照和不同活动状态下的虾苗视频，使用 DarkLabel 软件逐帧标注虾苗 ID。

表 1　测试数据集参数

Tab. 1　Test dataset

视频编号	光照强度	活动情况	虾苗个数	视频时长/s
01	弱	正常	30	79
02	中等	正常	50	76
03	强	正常	68	38
04	强	逆水	105	102
05	强	应激	79	93

2.1.2　评价指标

(1)检测评价指标

采用精确率(P)、召回率(R)、平均精度均值(mAP@0.5)、参数量、帧率(FPS)作为评价指标。其中，mAP 表示平均精度均值，采用 mAP@0.5，即只有当检测帧和真实目标之间的 IoU 大于 0.5 时，检测框才是正确的。

(2)跟踪评价指标

在多目标跟踪评估中，采用 4 个关键指标：跟踪准确度(MOTA)、跟踪精度(MOTP)、ID 切换数(IDSW)和综合指标(IDF1)。MOTA 反映跟踪算法在检测物体和保持轨迹时的性能，值越大性能越好；MOTP 量化检测器的定位精度，值越大精度越高；IDSW 评估同一只目标 ID 的连续性，值越小越好；IDF1 评估跟踪算法的稳定性，值越大越稳定。

2.2　结果与分析

2.2.1　虾苗检测实验与结果分析

(1)检测器的对比实验

为进一步验证改进 YOLO v8 对于虾苗识别的可靠性，分别和 Faster R－CNN[32]、SSD[33]、FCOS[34]、EfficientDet[35]、YOLO 系列的 YOLO v9 算法[36]进行对比实验。结果如表 2 所示。

由表 2 可知，本文基准模型(YOLO v8)的精确率、召回率和 mAP@0.5 等指标均优于其他模型。在参数量最少的同时，YOLO v8 的检测速度也最快，综合表现最佳。YOLO v8 的各项指标均高于 YOLO 系列的 YOLO v9，表明在虾苗检测任务上 YOLO v8 为最优选。

表 2　检测对比结果

Tab. 2　Comparison of detection results

模型	精确率/%	召回率/%	mAP@0.5/%	参数量	帧率/(f·s^{-1})
SSD	46.5	69.4	45.7	2.37×10^7	51
EfficientDet	88.3	96.3	91.8	5.23×10^7	51
Faster R－CNN	90.0	95.2	93.0	4.13×10^7	46
FCOS	81.1	88.3	86.7	3.23×10^7	46
YOLO v9	92.7	91.1	96.3	3.34×10^6	106
YOLO v8	91.9	95.6	97.7	3.08×10^6	131
改进 YOLO v8	93.0	96.0	98.1	2.49×10^6	160

(2)消融实验

为了验证本文改进 YOLO v8 模型的有效性，以 YOLO v8 模型为基础，通过使用不同模块来验证模型性能，结果见表 3(表中，①表示 CARAFE，②表示 C2f－BFB，③表示 Wise－Focaler－ShapeIoU)。

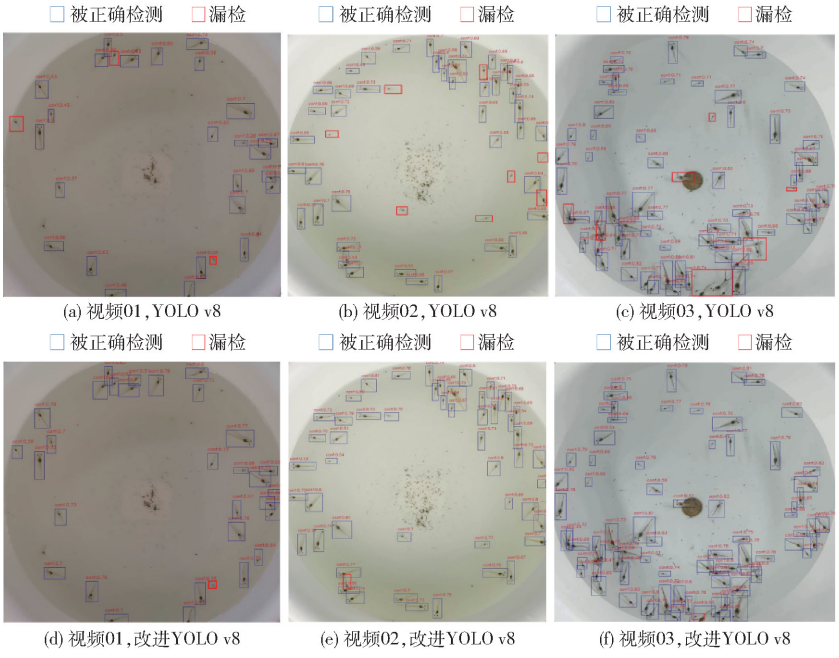
表 3　消融实验对比结果

Tab. 3　Comparison results of ablation experiments

模型	精确率/%	召回率/%	mAP@0.5/%	参数量	帧率/(f·s^{-1})
YOLO v8	92.1	93.1	97.0	3.078×10^6	131
YOLO v8－①	92.5	94.1	97.1	2.700×10^6	145
YOLO v8－②	92.5	94.6	97.5	2.030×10^6	147
YOLO v8－③	92.8	95.2	97.4	3.080×10^6	152
YOLO v8－①－②－③	93.0	96.0	98.1	2.490×10^6	160

由表 3 可知，改进 YOLO v8 模型相较原模型，平均精度均值(mAP@0.5)提升 1.1 个百分点。值得注意的是，其召回率的提升尤为显著，达到了 2.9 个百分点，这意味着改进 YOLO v8 模型有效减少了虾苗漏检情况。此外，改进模型在参数量和检测速度上也表现出色。改进的 YOLO v8 模型不仅将参数量从 3.078×10^6 减少到 2.49×10^6，而且推理速度显著提升，帧率从 131 f/s 增加到 160 f/s。这些改进确保了模型在计算资源有限情况下，依然能够以更快的速度进行高精度检测，充分展示了改进的 YOLO v8 模型在虾苗检测任务中具备更高的检测性能。

将 YOLO v8 和改进 YOLO v8 在验证集上的检测效果进行对比，来验证改进模型的有效性。图 12 分别展示了弱光、中等光、强光状态下的检测效果，对比发现改进的 YOLO v8 检测出的目标框置信度大部分比原模型高，充分展示了改进的 YOLO v8 模型具有更高的精确率。在不同光照强度状态



(a) 视频01,YOLO v8 (b) 视频02,YOLO v8 (c) 视频03,YOLO v8

(d) 视频01,改进YOLO v8 (e) 视频02,改进YOLO v8 (f) 视频03,改进YOLO v8

图 12 检测结果对比

Fig. 12 Comparison of test results

下,原模型 YOLO v8 漏检情况多,特别是在虾苗体积小和虾苗密集情况下,而改进的 YOLO v8 比原模型漏检情况少。由此表明,改进的 YOLO v8 在不同光照强度下都能准确地识别虾苗,并且能够更好地适应小目标检测,有效提高虾苗检测精度,改善漏检情况。

为了进一步分析改进的 YOLO v8 模型的泛化能力,还收集了 P5、P8、P12 阶段的虾苗图像进行验证,结果如表 4 所示。

表 4 不同阶段虾苗的验证结果

Tab. 4 Validation results for shrimp larvae at different stages

虾苗阶段	虾苗长度/cm	图像数量	精确率/%	召回率/%
P5	0.5	68	80.4	81.6
P8	0.8	74	90.6	91.5
P12	1.2	103	91.5	93.4
P15	1.5	89	93.0	96.0

由表 4 可知,由于 P5 阶段虾苗特征不明显且体积较小,召回率相对较低,这表明存在一定漏检情况。然而,对于 P10、P12 和 P15 阶段的虾苗,模型检测指标均超过 90%。这表明该模型整体上具有良好的泛化能力,能够有效应对不同发育阶段的虾苗检测任务。

2.2.2 虾苗跟踪实验与结果分析

为了对本文的跟踪算法做一个整体性能评估,选取了 DeepSORT、SORT、ByteTrack[37] 作为跟踪器进行对比实验,结果如表 5 所示。由表 5 可知,使用改进 YOLO v8 模型作为检测器,与原始 YOLO v8 模型相比,同一个跟踪器下的各项跟踪评价指标都有明显改善,这也证明了改进 YOLO v8 模型有效性。本文跟踪算法各项指标均优于当前主流的 DeepSORT 和 ByteTrack 算法,这两大算法在设计中均引入了目标的外观信息作为考量因素,但在虾苗养殖场景中,由于虾苗外表特征的相似性,反而导致外表特征成为干扰因素,影响了跟踪算法的准确性。DeepSORT 算法是在 SORT 基础上的改进,但其各项跟踪指标均不如 SORT,这也进一步验证了在虾苗外表特征相似的情况下引入外观信息反而使跟踪性能变差。本文跟踪算法中引入了金字塔 Lucas - Kanade 光流法,同时提出角度匹配算法,解决因虾苗快速移动而带来的跟踪难题,因此在表 5 数据中,本文的跟踪算法各项指标均优于 SORT 算法。

表 5 本文跟踪算法和不同跟踪器的性能对比结果

Tab. 5 Tracking algorithm and performance comparison results of different trackers

检测器	跟踪器	MOTA/%	MOTP/%	IDF1/%	IDSW
YOLO v8	DeepSORT	48.817	69.911	65.573	94
	SORT	62.307	69.378	74.733	88
	ByteTrack	61.145	73.498	72.813	78
	本文跟踪算法	69.011	73.881	75.211	71
改进 YOLO v8	DeepSORT	52.728	70.314	65.002	123
	SORT	65.495	69.996	75.657	80
	ByteTrack	63.837	74.27	71.331	61
	本文跟踪算法	74.520	75.624	76.420	43

表 6 为本文跟踪算法在 5 段不同测试视频上的跟踪结果。其中,05 号视频展现出了最优的跟踪

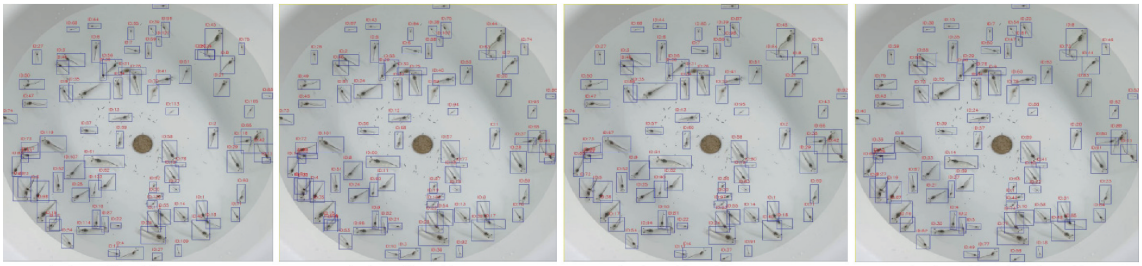
表 6　测试视频的跟踪结果

Tab. 6　Tracking results of test video

视频段	MOTA/%	MOTP/%	IDF1/%	IDSW
01	75. 76	76. 36	74. 78	26
02	69. 76	71. 32	70. 65	52
03	73. 79	74. 28	76. 86	38
04	76. 11	77. 35	79. 67	59
05	77. 18	77. 44	80. 14	40
平均值	74. 52	75. 62	76. 42	43

性能。这一优越表现主要归因于该视频场景中的特定条件：应激状态下的虾苗分布相对均匀，并未形成过于密集的聚集，且是在强光环境下。这种环境特点有助于减少跟踪过程中的遮挡和混淆现象，进而显著提升了跟踪算法的准确性和鲁棒性。

图 13 展示了 DeepSORT、SORT、ByteTrack 和本文算法的跟踪效果对比，4 种算法在虾苗分配上的最大 ID 依次为 124、105、98、91，而实际图中虾苗数量为 79 只。这一对比反映了随着跟踪时间延长，这 4 种算法在跟踪虾苗过程中均会出现不同程度的丢失现象，从而导致分配给虾苗的 ID 逐渐增大。本文跟踪算法在相同条件下的最大虾苗 ID 最小，表明与其他 3 种算法相比，本文算法在跟踪过程中发生 ID 切换的次数最少，保持了较好的跟踪稳定性。



(a) 视频05第50帧，DeepSORT　(b) 视频05第50帧，SORT　(c) 视频05第50帧，ByteTrack　(d) 视频05第50帧，本文算法

图 13　跟踪结果对比

Fig. 13　Comparison of tracking results

如图 14 所示，为了方便观察，本文选取了不同光照强度下的 4 只最靠近图像四角的虾苗作为跟踪对象。在跟踪至第 10 帧时，4 种算法在初始阶段均展现出了相似的跟踪效果。然而，随着帧数增加，各算法在持续跟踪能力上出现了显著差异。DeepSORT 算法在跟踪过程中，仅在强光条件下能够维持对 1 只虾苗的持续跟踪，而在其他光照条件下虾苗全部跟丢。SORT 和 ByteTrack 算法仅在强光条件下才能保持 3 只虾苗的跟踪，SORT 算法在中等光照条件下甚至出现了对 1 只虾苗的跟踪错

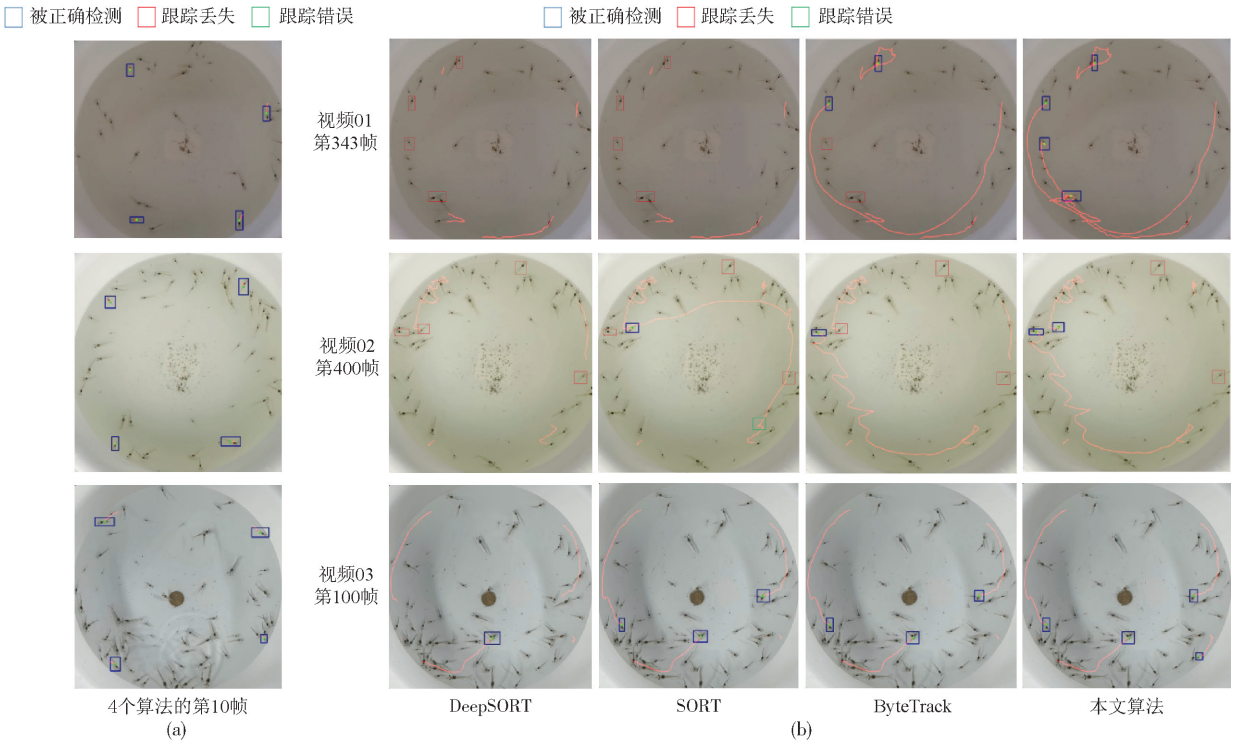


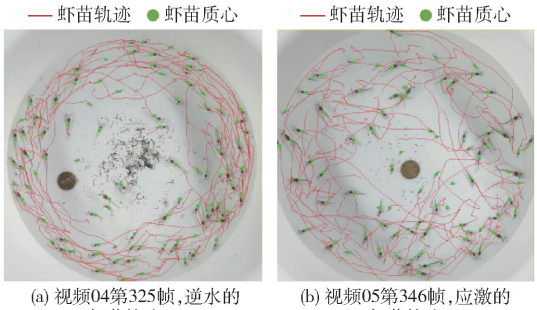
图 14　跟踪轨迹结果对比

Fig. 14　Comparison of tracking trajectories

误,这导致了其轨迹绘制与其他算法存在差异。相比之下,本文算法在不同光照条件下均展现出优异的跟踪稳定性。具体而言,在中等光照条件下,尽管出现了 2 只虾苗的跟踪丢失,但在其他光照条件下,该算法均能够实现对 4 只虾苗的稳定跟踪。这一结果表明,本文算法在虾苗跟踪任务上相较于其他算法具有更高的鲁棒性和稳定性。

2.2.3 虾苗轨迹可视化

工人在养殖场常用白盆评估虾苗质量:制造水流观察虾苗是否逆流游动,MIRZAEI 等[38]以逆流游动比例作为指标;同时通过敲击白盆或伸手入水,检验虾苗对刺激的敏捷反应,以判断活力。为直观地观察虾苗情况,利用本文算法对虾苗进行跟踪,并绘制虾苗轨迹如图 15 所示,逆水虾苗轨迹主要集中在白盆边缘,应激虾苗轨迹则呈现出无序分布状态。



(a) 视频04第325帧,逆水的虾苗轨迹　(b) 视频05第346帧,应激的虾苗轨迹

图 15 不同活动情况下的虾苗跟踪轨迹

Fig. 15 Tracking trajectories of shrimp larvae under different activity conditions

3 结论

(1)提出了一种基于改进 YOLO v8 的南美白对虾虾苗跟踪方法。采用改进 YOLO v8 算法作为检测器,获取虾苗在图像中位置信息。提出了一种融合距离匹配算法和角度匹配算法的匹配策略,并引入金字塔 LK 光流法,从而构建跟踪器。将检测器检测到的位置信息输入到跟踪器中,实现了对虾苗的稳定跟踪。为虾苗养殖管理与研究提供了有力技术支持。

(2)改进 YOLO v8 比原模型的 mAP@0.5 提高 1.1 个百分点,召回率提升 2.9 个百分点,拥有较高的检测性能。本文所提跟踪算法的 MOTA 为 74.52%、MOTP 为 75.62%、IDF1 为 76.42%,与 DeepSORT、SORT、ByteTrack 算法相比,MOTA 显著提高,跟踪效果良好。

参 考 文 献

[1] CHEN Y, MITRA A, RAHIMNEJAD S, et al. Retrospect of fish meal substitution in Pacific white shrimp (*Litopenaeus vannamei*) feed: alternatives, limitations and future prospects[J]. Reviews in Aquaculture, 2023, 16(1): 382 – 409.

[2] 农业农村部渔业渔政管理局. 中国渔业统计年鉴[M]. 北京: 中国农业出版社, 2023: 21.

[3] 周超. 基于卷积神经网络的虾苗计数、定位、测量及追踪的研究[D]. 杭州,杭州电子科技大学,2024.
ZHOU Chao. Research on shrimp larvae counting, locating, sizing and tracking based on convolutional neural network[D]. Hangzhou: Hangzhou Dianzi University, 2024. (in Chinese)

[4] 杨其彬,李晓兰,孙苗苗,等. 斑节对虾虾苗活力检测方法[J]. 南方水产,2010,6(1):12 – 18.
YANG Qibin, LI Xiaolan, SUN Miaomiao, et al. Methods of testing vitality of postlarvae of *Penaeus monodon*[J]. South China Fisheries Science, 2010, 6(1): 12 – 18. (in Chinese)

[5] 李鹏龙,张胜茂,沈烈,等. 基于深度学习的鱼类跟踪技术研究进展[J]. 渔业现代化,2024,51(2):1 – 13.
LI Penglong, ZHANG Shengmao, SHEN Lie, et al. Research progress in fish tracking technology based on deep learning[J]. Fishery Modernization, 2024, 51(2): 1 – 13. (in Chinese)

[6] 申阳,王冲宇,赵佳怡,等. 基于改进 YOLOv8 和多目标跟踪的鱼苗计数方法[J]. 农业工程学报,2024,40(16):163 – 170.
SHEN Yang, WANG Chongyu, ZHAO Jiayi, et al. Fry counting method using improved YOLOv8 and multi-target tracking[J]. Transactions of the CSAE, 2024, 40(16): 163 – 170. (in Chinese)

[7] MORAIS E F, CAMPOS M F M, PADUA F L C, et al. Particle filter-based predictive tracking for robust fish counting[J/OL]. IEEE, 2006. DOI: 10.1109/SIBGRAPI.2005.36.

[8] 沈蔚,朱振宏,张进,等. 基于双频识别声呐的鱼类目标识别与计数[J]. 渔业现代化,2020,47(6):81 – 87.
SHEN Wei, ZHU Zhenhong, ZHANG Jin, et al. Fish target recognition and counting based on dual-frequency identification sonar [J]. Fishery Modernization, 2020, 47(6): 81 – 87. (in Chinese)

[9] 范良忠,刘鹰,余心杰,等. 基于计算机视觉技术的运动鱼检测算法[J]. 农业工程学报,2011,27(7):226 – 230,394.
FAN Liangzhong, LIU Ying, YU Xinjie, et al. Fish motion detecting algorithms based on computer vision technologies[J]. Transactions of the CSAE, 2011, 27(7): 226 – 230, 394. (in Chinese)

[10] 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.

[11] ZHANG H, LI W, QI Y, et al. Dynamic fry counting based on multi-object tracking and one-stage detection[J]. Computers and Electronics in Agriculture, 2023, 209: 107871.

[12] VEERAMANI B, RAYMOND J W, CHANDA P. DeepSort: deep convolutional networks for sorting haploid maize seeds[J]. BMC Bioinformatics, 2018, 19: 1 – 9.

[13] 洪宇,叶瑞娟,冯国富. 一种基于改进 DeepSORT 的淡水环境下鱼类计数方法[J/OL]. 海洋渔业,1 – 12. https://doi.org/10.13233/j.cnki.mar.fish.20230703.001.
HONG Yu, YE Ruijuan, FENG Guofu. A fish counting method based on improved DeepSORT in freshwater environment[J/OL]. Marine Fisheries, 1 – 12. https://doi.org/10.13233/j.cnki.mar.fish.20230703.001. (in Chinese)

[14] XU F, DING X, PENG J, et al. Real-time detecting method of marine small object with underwater robot vision[C]//2018 OCEANS-MTS/IEEE Kobe Techno-Oceans (OTO). IEEE, 2018: 1 – 4.

[15] 赵海翔,崔鸿武,黄桢铭,等. 基于 Bytetrack 的多目标跟踪算法在斑马鱼毒性行为识别中的应用[J]. 渔业科学进展,2024,45(2):136 – 149.
ZHAO Haixiang, CUI Hongwu, HUANG Zhenming, et al. Application of a Bytetrack-based, multi-target tracking algorithm for zebrafish toxicity-response behavior recognition[J]. Progress in Fishery Sciences, 2024, 45(2): 136 – 149. (in Chinese)

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

[17] DU Y, ZHAO Z, SONG Y, et al. Strongsort: make deepsort great again[J]. IEEE Transactions on Multimedia, 2023, 25: 8725 – 8737.

[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 Zhiqiang, LI Zhijun, LIU Shuo, et al. Multi-category sea foods counting method integrating YOLOv7 and BYTE multi-target tracking[J]. Transactions of the CSAE, 2023,39(9): 183 – 189. (in Chinese)

[19] XU W, WANG P, JIANG L, et al. Intelligent recognition and behavior tracking of sea cucumber infected with *Vibrio alginolyticus* based on machine vision[J]. Aquacultural Engineering, 2023, 103: 102368.

[20] SAAD A, JAKOBSEN S, BONDØ M, et al. StereoYolo + DeepSORT: a framework to track fish from underwater stereo camera in situ[C]//Sixteenth International Conference on Machine Vision (ICMV 2023). SPIE, 2024, 13072: 321 – 329.

[21] LIU S, LI X, GAO M, et al. Embedded online fish detection and tracking system via YOLOv3 and parallel correlation filter[C]//Oceans 2018 Mts/Ieee Charleston. IEEE, 2018: 1 – 6.

[22] LU H, UEMURA T, WANG D, et al. Deep-sea organisms tracking using dehazing and deep learning[J]. Mobile Networks and Applications, 2020, 25: 1008 – 1015.

[23] LI X, WANG W, WU L, et al. Generalized focal loss: learning qualified and distributed bounding boxes for dense object detection[J]. Advances in Neural Information Processing Systems, 2020, 33: 21002 – 21012.

[24] ZHENG Z, WANG P, REN D, et al. Enhancing geometric factors in model learning and inference for object detection and instance segmentation[J]. IEEE Transactions on Cybernetics, 2021, 52(8): 8574 – 8586.

[25] ZHU L, WANG X, KE Z, et al. Biformer: vision transformer with bi-level routing attention[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023: 10323 – 10333.

[26] WANG J, CHEN K, XU R, et al. Carafe: content-aware reassembly of features[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision, 2019: 3007 – 3016.

[27] ZHANG H, ZHANG S. Focaler-IoU: more focused intersection over union loss[J]. arXiv Preprint, arXiv:2401.10525, 2024.

[28] HANG H, ZHANG S. Shape-IoU: more accurate metric considering bounding box shape and scale[J]. arXiv Preprint, arXiv: 2312.17663, 2023.

[29] TONG Z, CHEN Y, XU Z, et al. Wise-IoU: bounding box regression loss with dynamic focusing mechanism[J]. arXiv Preprint, arXiv:2301.10051, 2023.

[30] WANG Z, YANG X. Moving target detection and tracking based on Pyramid Lucas-Kanade optical flow[C]//2018 IEEE 3rd International Conference on Image, Vision and Computing (ICIVC). IEEE, 2018: 66 – 69.

[31] LUCAS B D, KANADE T. An iterative image registration technique with an application to stereo vision[C]//IJCAI'81: 7th International Joint Conference on Artificial Intelligence, 1981: 674 – 679.

[32] LI X, SHANG M, HAO J, et al. Accelerating fish detection and recognition by sharing CNNs with objectness learning[C]//OCEANS 2016-Shanghai, 2016: 1 – 5.

[33] LIU W, ANGUELOV D, ERHAN D, et al. Ssd: single shot multibox detector[C]//Computer Vision – ECCV 2016: 14th European Conference, 2016: 21 – 37.

[34] TIAN Z, SHEN C, CHEN H, et al. FCOS: fully convolutional one-stage object detection[J]. arXiv Preprint, arXiv:1904.01355, 1904.

[35] 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.

[36] WANG C Y, YEH I H, LIAO H Y M. Yolov9: learning what you want to learn using programmable gradient information[J]. arXiv Preprint, arXiv:2402.13616, 2024.

[37] ZHANG Y, SUN P, JIANG Y, et al. Bytetrack: multi-object tracking by associating every detection box[C]//European Conference on Computer Vision, 2022: 1 – 21.

[38] MIRZAEI N, MOUSAVI S M, YAVARI V, et al. Quality assessment of *Litopenaeus vannamei* postlarvae produced in some commercial shrimp hatcheries of Choubdeh Abadan, Iran[J]. Aquaculture, 2021, 530: 735708.