

文章编号: 1674-5566(2023)05-1015-09

DOI: 10.12024/jsou.20230604238

基于 YOLO 的贻贝检测算法

董兆鹏¹, 岳晓雪¹, 田中旭¹, 侯思凡¹, 姜立盛²

(1. 上海海洋大学 工程学院, 上海 201306; 2. 大连南成修船有限公司, 辽宁 大连 116011)

摘要: 基于 YOLO 的贻贝 (*Mytilus edulis*) 识别与检测技术, 是实现贻贝分级、分苗等作业环节机械化和智能化的关键。然而, 贻贝因外部特征不够清晰明确, 给识别准确率的提高带来了挑战。本文提出一种基于改进 YOLOv5 算法的贻贝目标检测模型 (CST-YOLO)。该算法融合 CoordAttention 注意力机制, 以增强特征表达能力; 采用 SIoU 作为边界框回归损失函数, 以减少边界框回归损失, 提高模型的检测速度; 将 Head 替换为改进的解耦头 TSCODE Head 来提高检测准确率。并在自制的贻贝数据集上进行算法测试, 实验结果显示: 相比 YOLOv5 算法, CST-YOLO 算法的准确率 P 提高了 0.428%, mAP_{0.5}: 0.95 达到 92.221%, 提高了 1.583%。实验表明 CST-YOLO 算法在保证检测速度的前提下, 有效提高了贻贝目标的检测精度。本研究有助于机器视觉技术在贻贝养殖业自动化与智能化生产加工中的应用。

关键词: 贻贝; 目标检测; 注意力机制; 损失函数; 解耦头

中图分类号: S 968.3; TP 391.4 **文献标志码:** A

贻贝 (*Mytilus edulis*) 是一种营养价值高、产量大、易于人工养殖的海产品^[1]。在贻贝的分苗、分级等机械化作业中, 需要对贻贝进行识别, 同时提取数量与尺寸的信息^[2]。关于贻贝机器视觉技术的研究与应用, 成为贻贝作业装备向机械化、智能化发展的关键。因贻贝本身外部特征不清晰, 易受环境干扰, 为提高贻贝机器视觉识别准确率带来了挑战。

在生物识别机器视觉研究中, 基于卷积神经网络 (Convolutional neural network, CNN) 的目标检测算法因其出色的鲁棒性和高度自学习性而被广泛应用。在算法应用中, 研究者在研究对象的特征方面进行了改进, 取得了较好效果。如 LIU 等^[3] 基于改进的 LeNet-5 结构, 提出了一种新型深度卷积神经网络, 来区分混合有缺陷或病害的虾与健全的虾, 以提升虾养殖的整体品质。HAN 等^[4] 基于深度卷积神经网络, 提出一种通过水下视觉进行目标识别的方法, 提高水下目标的检测精度。陈旭等^[5] 提出一种以

YOLOv5s 算法为基础, 添加小目标检测层 M, 引入 BoT 模块, 并添加 CBAM 注意力机制, 替换解耦头的方法有效提高了检测精度。黄英来等^[6] 提出在 ResNet-50 模型中替换卷积核, 使用 LeakyReLU 激活函数, 改变残差块中批标准化层、激活函数与卷积层的排列顺序, 大幅度提升玉米叶片病害的识别正确率。陈金广等^[7] 提出一种以 ShufflenetV2 网络作为 YOLOv5 骨干网络模型, 加入 SimAM 注意力机制, 引入 Alpha-IoU 边界框回归损失函数的方法, 减少算法参数量、加快识别速度。

目前, 在贻贝检测识别问题上, 一方面, 国内外有针对性的研究较少; 另一方面, 由于贻贝的外部特征不太明显, 易受环境因素干扰, 导致常规方法的检测识别准确率低。为解决前述问题, 本文在提升特征表达能力和定位信息精确性方面采用了融合注意力机制、优化损失函数以及改进解耦头的方法, 以提高贻贝目标检测的准确率。

收稿日期: 2023-06-15 修回日期: 2023-08-10

基金项目: 十三五“蓝色粮仓科技创新”国家重点研发计划(2019YFD0900803); 上海海洋大学科技发展专项(A2-2006-22-200209)

作者简介: 董兆鹏(1985—), 男, 讲师, 博士, 研究方向为图像识别技术和精密仪器。E-mail: zpdong@shou.edu.cn

通信作者: 田中旭, E-mail: zxtian@shou.edu.cn

版权所有 ©《上海海洋大学学报》编辑部(CC BY-NC-ND 4.0)

Copyright © Editorial Office of Journal of Shanghai Ocean University (CC BY-NC-ND 4.0)

<http://www.shhydx.com>

1 材料与方法

1.1 数据集制作

图像采集设备为1 200万像素摄像头,拍摄环境为白天自然光照条件,采集贻贝不同角度的



图1 自制的贻贝图像数据集展示

Fig. 1 Display of self-made mussel image dataset

选取1 631张贻贝图像作为初始数据集,利用Python代码采取随机抽样的方法划分10%作为验证集验证模型训练情况,10%作为测试集评价模型训练情况,其余1 304张图像作为训练集训练模型。本文采用LabelImg制作贻贝数据集标签。

1.2 实验环境与参数配置

本研究使用的计算机操作系统是Windows10,配备了Intel Xeon E5-2689 v4@3.10GHz的中央处理器(CPU)以及RTX 3090的图形处理器(GPU),显存容量为512 GB。CUDA版本为11.6,Cudnn版本为8.4。实验使用了Python 3.7和PyTorch 13.1作为主要的编程语言和深度学习框架。

在模型训练过程中,输入640像素×640像素的RGB图片,采用随机梯度下降(Stochastic gradient descent,SGD)优化器,将网络训练的学习率初始化为0.01,动量因子设为0.937,权重衰减设为0.000 5,超参数使用hyp.scratch-low,批次大小batch-size设为8,总Epoch设为50轮。

1.3 模型选择

YOLOv5是一种基于深度学习的目标检测模型,相较于之前的YOLO版本(YOLOv1-YOLOv4^[8]),它具备更高的检测精度和更快的检测速度。YOLOv5包含了4个版本:YOLOv5s、YOLOv5m、YOLOv5l和YOLOv5x。由于YOLOv5s检测效率高并对硬件要求低,因此本研究基于

正面、侧面和反面图像。在采集中,也改变了光照条件,保证数据来源广泛且拍摄背景多样化,以符合自然场景的应用需求。自制的贻贝图像数据集展示如图1所示。

YOLOv5s算法对贻贝检测任务进行改进,提出CST-YOLO算法。

2 CST-YOLO算法

2.1 注意力机制融合

贻贝图像采集中可能存在图像模糊的情况,从而导致信息的缺失。为了增强目标检测精度,并提高对检测目标的关注度,在YOLOv5模型中引入了坐标注意力机制(Coordinate attention,CA)^[9]。CA是一种专注于空间坐标的注意力机制,旨在增强特征表达的性能。它能够根据输入特征的空间坐标位置,自动调整不同位置的权重,从而更加准确地关注图像中的目标区域。相比于传统的注意力机制,CA注意力机制具有更高的计算效率和更好的自适应性。引入CA注意力机制,能够提升模型的特征表达能力,其结构示意图见图2。

CA注意力机制分为坐标信息嵌入和协同注意力生成两个步骤。在坐标信息嵌入部分,将输入特征图分别在宽度和高度方向上使用池化核($H,1$)和($1,W$)进行平均池化,高度为 h 的第 c 通道输出公式如下:

$$Z_c^h(h) = \frac{1}{W} \sum_{0 \leq j < W} x_c(h, j) \quad (1)$$

式中: c 为输入特征图的通道数; h 为输入特征图的高度; W 为输入特征图的宽度; x_c 为坐标信息的特定分量。

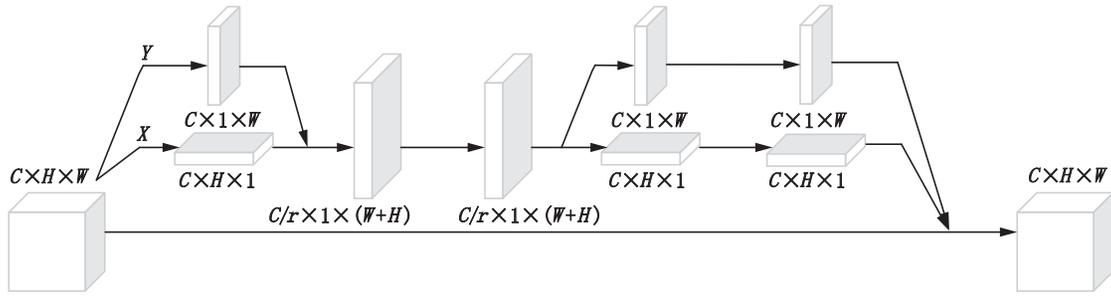


图2 CA注意力机制结构

Fig. 2 CA attention mechanism structure

宽度为 w 的第 c 通道输出公式如下:

$$Z_c^w(w) = \frac{1}{H} \sum_{0 \leq j < H} x_c(j, w) \quad (2)$$

式中: c 为输入特征图的通道数; w 为输入特征图的宽度; H 为输入特征图的高度。

在协同注意力部分,将公式(1)和公式(2)融合,通过 1×1 卷积操作,再进行批量归一化和非线性激活函数处理,以提升特征表达能力,公式如下:

$$f = \delta(F_1[z^h, z^w]) \quad (3)$$

式中: $[\]$ 为融合操作; F_1 为卷积操作; δ 为非线性激活函数; f 为中间特征图, $f \in \mathbb{R}^{C/r \times (H+W)}$ 。

再进行 1×1 卷积操作,公式如下:

$$g^h = \sigma[F_h(f^h)] \quad (4)$$

$$g^w = \sigma[F_w(f^w)] \quad (5)$$

式中: f^h, f^w 为特征图, $f^h \in \mathbb{R}^{C/r \times H}, f^w \in \mathbb{R}^{C/r \times W}$; F_h, F_w 为卷积操作; σ 为 Sigmoid 激活函数; g^h, g^w 为注意力向量, $g^h \in \mathbb{R}^{C \times H \times 1}, g^w \in \mathbb{R}^{C \times W \times 1}$ 。

最后进行归一化和加权运算,得到最终的输出特征图,输出公式如下:

$$y_c(i, j) = x_c(i, j) \times g_c^h(i) \times g_c^w(j) \quad (6)$$

式中: i, j 为像素坐标值。

2.2 损失函数优化

YOLOv5 算法采用 CIoU (Complete intersection over union)^[10] 损失函数优化模型的检测性能。该算法在 IoU 损失函数的基础上,引入对预测框和真实框长宽比、中心点距离的惩罚项,对 IoU 进行修正。CIoU 损失函数的计算量较大,主要依赖于边界框回归指标的聚合,未考虑到期望的真实框和预测框之间的不匹配,从而导致收敛速度慢。该函数定义如公式(7)所示:

$$L_{\text{CIoU}} = 1 - P_{\text{IoU}} + \frac{\rho^2(b, b^{\text{gt}})}{c^2} + \alpha v \quad (7)$$

式中: P_{IoU} 为预测框和真实框的重叠部分与并集

的比值; α 为权重参数; v 为长宽比的相似性; c 为包含预测框和真实框的最小闭包区域的对角线距离; b 为预测框中心坐标; $b_{\text{cx}}^{\text{gt}}$ 为真实框中心坐标; ρ 为两点之间欧式距离。

为解决这些问题,本文引入了 SIoU (Smoothed intersection over union) 损失函数。该函数重新定义了惩罚度量,考虑了期望回归之间的向量夹角,不限制预测框和真实框的长宽比,以适应不同形状的物体,引入面积比例惩罚项,提高检测结果准确性,且计算量较小,训练速度和效率高。SIoU 损失函数由角度损失、距离损失、形状损失和 IoU 损失 4 个损失函数组成。

角度损失函数如公式(8)所示:

$$\Lambda = 1 - 2 \times \sin^2[\arcsin(\frac{c_h}{\sigma}) - 4\pi] \quad (8)$$

式中: σ 为真实框和预测框中心点的距离; c_h 为真实框和预测框中心点的高度差。

距离损失函数如公式(9)、(10)所示:

$$\Delta = \sum_{t=x/y} [1 - e^{-(A-2)\rho_t}] = 2 - e^{-(A-2)\rho_x} - e^{-(A-2)\rho_y} \quad (9)$$

$$\rho_x = (\frac{b_{\text{cx}}^{\text{gt}} - b_{\text{cx}}}{c_w})^2, \rho_y = (\frac{b_{\text{cy}}^{\text{gt}} - b_{\text{cy}}}{c_h})^2 \quad (10)$$

式中: c_w, c_h 为真实框和预测框最小外接矩形的宽和高; $b_{\text{cx}}^{\text{gt}}, b_{\text{cy}}^{\text{gt}}$ 为真实框中心坐标; $b_{\text{cx}}, b_{\text{cy}}$ 为预测框中心坐标。

形状损失函数如公式(11)、(12)所示:

$$\Omega = \sum_{t=w/h} (1 - e^{-\omega_t})^\theta = (1 - e^{-w_\omega})^\theta + (1 - e^{-w_h})^\theta \quad (11)$$

$$w_\omega = \frac{|w - w^{\text{gt}}|}{\max(w, w^{\text{gt}})}, w_h = \frac{|h - h^{\text{gt}}|}{\max(h, h^{\text{gt}})} \quad (12)$$

式中: $w, h, w^{\text{gt}}, h^{\text{gt}}$ 分别为预测框和真实框的宽和高; θ 为对形状损失的关注程度。

IoU 损失函数如公式(13)所示:

$$L_{\text{IoU}} = 1 - P_{\text{IoU}} \quad (13)$$

因此,SIoU 损失函数定义如公式(14)所示:

$$L_{\text{IoU}} = 1 - P_{\text{IoU}} + \frac{\Delta + \Omega}{2} \quad (14)$$

2.3 解耦头改进

在贻贝检测过程中,可能会出现目标互相遮挡的情况,因此,需要准确的定位信息来确定各个目标的位置。为了提高定位和识别的精确度,采用了一种改进的上下文解耦检测头(Task-specific context decoupling head, TSCODE Head)来进行目标检测。TSCODE Head 能够为分类任务生成空间粗糙但语义信息更强的特征编码,并为定位任务提供包含更多边缘信息的高分辨率特征图,从而有效地提高模型的检测准确率。

根据图3所示的改进的 TSCODE Head 结构,

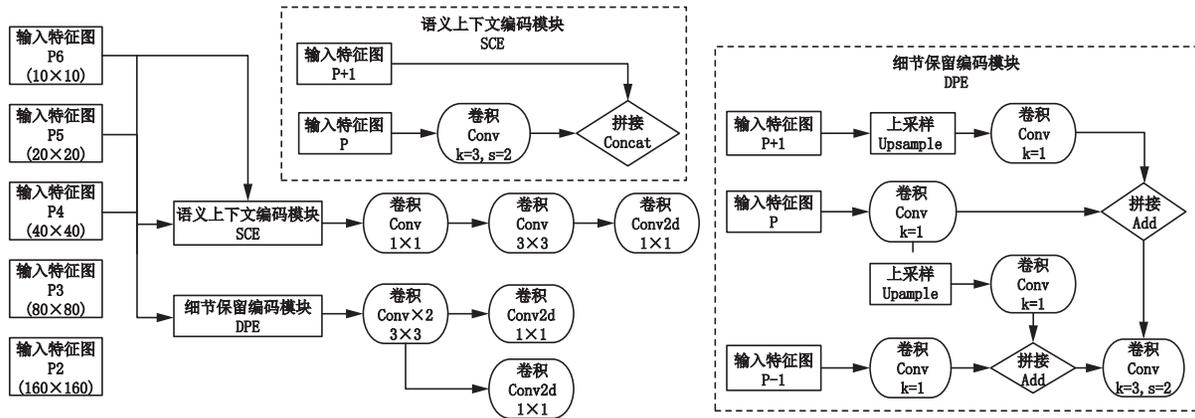


图3 改进的 TSCODE Head
Fig. 3 Improved TSCODE Head

2.4 CST-YOLO算法模型

YOLOv5算法框架由骨干网络(Backbone)、颈网络(Neck)和头网络(Head)等3个部分组成。骨干网络负责提取特征并进行下采样,以供后续的检测任务使用。主要使用CSP Darknet,它采用CSP(Cross-stage partial network)^[11]的思想将输入的特征图分成两部分处理,以降低计算量和内存占用,并提高特征表达能力。颈网络由特征金字塔网络(Feature pyramid networks, FPN)^[12]和路径聚合网络(Path aggregation network, PAN)^[13]组成,用于融合不同尺度的特征图,生成更具表达能力的多尺度特征表示。头网络是模型的检测头部分,将颈网络输出的特征图转换为检测结果。

在此基础上,CST-YOLO算法融合CA注意力机制,以增强特征表达能力。同时,采用SIoU作为边界框回归损失函数,并在Head部分采用改进的TSCODE Head。CST-YOLO算法模型如图4所示。

该部分输出包括分类分支和定位分支两部分。在分类分支中,首先使用语义上下文编码(Semantic context encoding, SCE)生成语义丰富的特征图,以提取目标的局部信息;其次通过1×1卷积降低通道数,再经过3×3卷积提取目标的特征信息;最后使用1×1卷积进一步压缩特征图的通道数,并完成分类任务。在定位分支中,首先使用细节保留编码(Detail preserving encoding, DPE)实现精确定位;其次采用2个3×3卷积提取目标的特征信息。将得到的特征图分为两部分,分别经过1×1卷积进行通道数的压缩,一部分获取边界框的置信度信息,另一部分获取边界框的回归预测,即边界框的位置和尺寸信息。

3 结果与讨论

3.1 评价指标

为评估YOLOv5模型在贻贝图像识别中的准确性,本文采用了目标检测中常用的性能指标,包括:准确率(Precision, P)、召回率(Recall, R)、平均精度(Mean average precision, mAP)和检测时间(Time, t),具体计算公式如下:

准确率是指模型预测的所有目标中,预测正确的目标比例,可以称之为查准率。

$$P = \frac{T_p}{T_p + F_p} \quad (15)$$

式中: T_p 为IoU>0.5的检测框数量,即模型识别贻贝为正确的数量; F_p 为IoU≤0.5的检测框数量,即模型识别贻贝为错误的数量。

召回率是指所有真实目标中,模型预测正确的目标比例,可以称之为查全率。

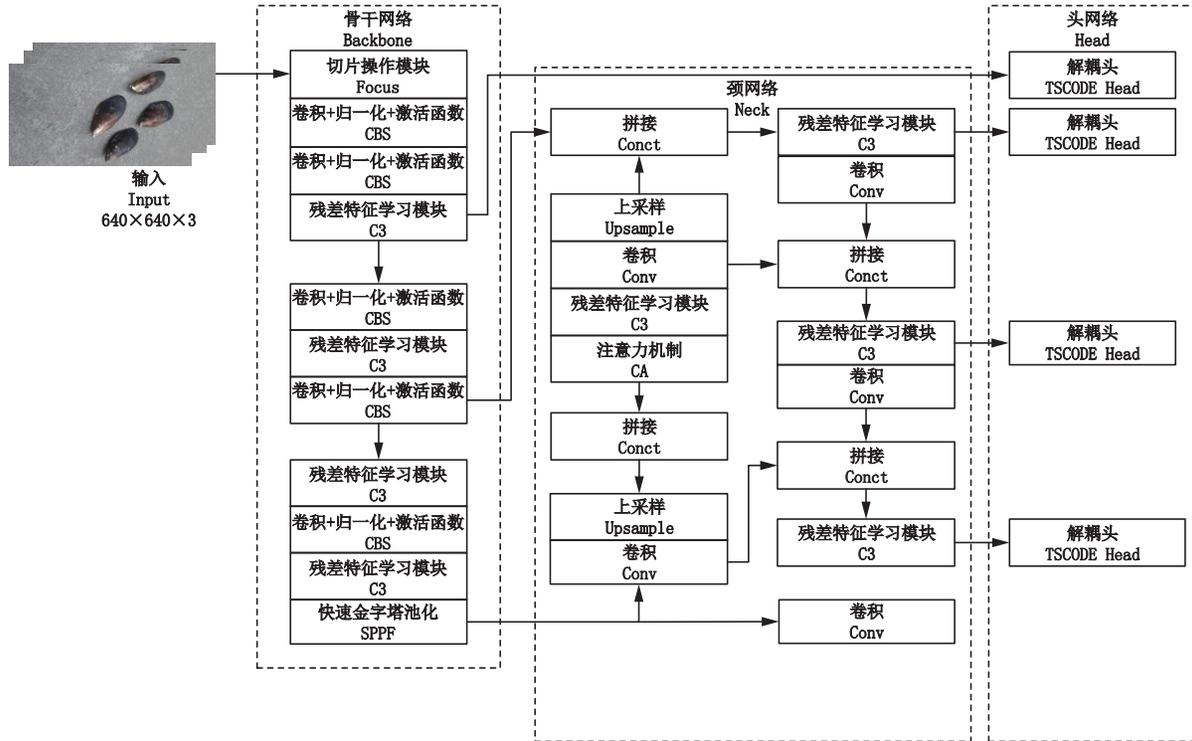


图4 CST-YOLO算法模型

Fig. 4 CST-YOLO algorithm model

$$R = \frac{T_p}{T_p + F_N} \quad (16)$$

式中: F_N 为模型未识别到的贻贝的数量。

AP(Average precision)是指单个类别的平均精度, mAP是指各类别AP的平均值。本文使用 $mAP_{0.5}$ 和 $mAP_{0.5-0.95}$ 对模型进行评估。 $mAP_{0.5}$ 表示IoU设为0.5时的平均精确度, $mAP_{0.5-0.95}$ 表示IoU从0.5到0.95, 步长为0.05时的平均精确度。

$$I_{mAP} = \frac{\sum I_{AP}}{N_c} \times 100\% \quad (17)$$

式中: I_{mAP} 为各类别平均精度的平均值; I_{AP} 为贻贝的平均精度; N_c 为模型能检测出目标的类别数量。

此外, 检测时间 t 为单张贻贝图像平均处理时间(s)。

3.2 注意力机制实验

为了研究CA注意力机制在YOLOv5s模型中的效果, 将CA注意力机制与SE、SimAM注意力机制在贻贝数据集上进行实验对比。实验结果如表1所示, 不同注意力机制 $mAP_{0.5-0.95}$ 如图5所示。

实验结果显示, CA注意力机制效果最佳, 提升了平均精度, 验证了其在特征提取中的有效性。

表1 注意力机制实验结果

Tab. 1 Experimental results of attention mechanism

网络模型 Network model	注意力机制 Attention mechanism	$mAP_{0.5}/\%$	$mAP_{0.5-0.95}/\%$
Yolov5s	Baseline	99.497	90.628
Yolov5s	CA	99.500	90.633
Yolov5s	SE	99.500	90.327
Yolov5s	SimAM	99.497	90.510

3.3 损失函数实验

为验证SIoU边界框回归损失函数在YOLOv5s模型中的高效性, 进行其与CIoU、EIoU边界框回归损失函数的对比实验。损失函数实验结果如表2所示, 不同损失函数的 $mAP_{0.5-0.95}$ 和 Box_loss 情况如图6和图7所示。

实验结果显示, SIoU损失函数 Box_loss 最低, $mAP_{0.5-0.95}$ 提升至90.681%, 相较于YOLOv5s模型提高了0.053%。这表明SIoU损失函数能够提高贻贝目标检测的准确性。

3.4 消融实验

为验证CST-YOLO算法对模型性能提高的有效性, 进行一系列消融实验, 对比不同的模块及其组合对目标检测算法性能的影响。选择YOLOv5s作为基准算法, 在自制的贻贝数据集上

进行实验。消融实验结果如表 3 所示, $mAP_{0.5-0.95}$ 情况如图 8 所示。

实验结果显示, 最佳的目标检测性能是在同时采用融合注意力机制 CA、引入 SIoU 边界框回归损失函数以及替换为 TSCODE Head 的情况下

实现。相对于 YOLOv5s, CST-YOLO 算法的精确率提高了 0.428%, 召回率下降了 0.131%, $mAP_{0.5}$ 提高了 0.003%, $mAP_{0.5-0.95}$ 提高了 1.583%。结果表明 CST-YOLO 算法能够提升目标检测效果, 并证明了其在贻贝目标检测性能上的优势。

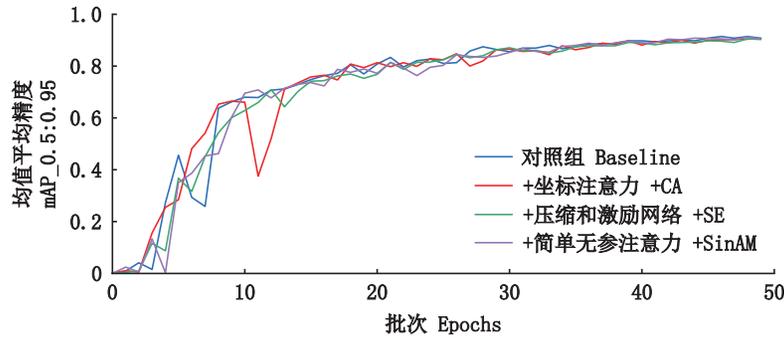


图 5 不同注意力机制 $mAP_{0.5:0.95}$ 曲线图

Fig. 5 Different attention mechanisms $mAP_{0.5:0.95}$ curve graph

表 2 损失函数实验结果

Tab. 2 Experimental results of loss function

网络模型 Network model	损失函数 Loss function	边界框损失 Box_loss	$mAP_{0.5}/\%$	$mAP_{0.5-0.95}/\%$
Yolov5s	CIoU	0.015	99.497	90.628
Yolov5s	SIoU	0.015	99.493	90.681
Yolov5s	EIoU	0.017	99.497	90.284

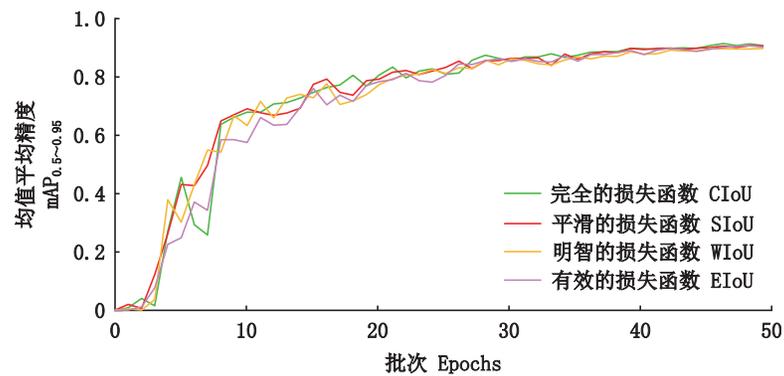


图 6 损失函数 $mAP_{0.5:0.95}$ 曲线图

Fig. 6 Loss function $mAP_{0.5:0.95}$ curve graph

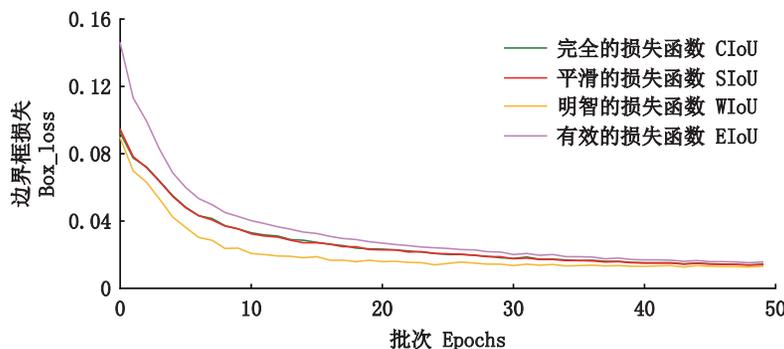


图 7 损失函数 box_loss 曲线图

Fig. 7 Loss function box_loss curve graph

表3 消融实验结果
Tab. 3 Results of ablation experiment

网络模型 Network model	精确率 Precision /%	召回率 Recall/%	mAP _{0.5} /%	mAP _{0.5-0.95} /%
YOLOv5s	99.518	99.660	99.497	90.628
YOLOv5s+CA	99.323	99.809	99.500	90.633
Yolov5s+SIOU	98.919	100.000	99.493	90.681
Yolov5s+TSCODE Head	99.475	99.660	99.497	92.136
Yolov5s+CA+SIOU+TSCODE Head	100.000	99.529	99.500	92.211

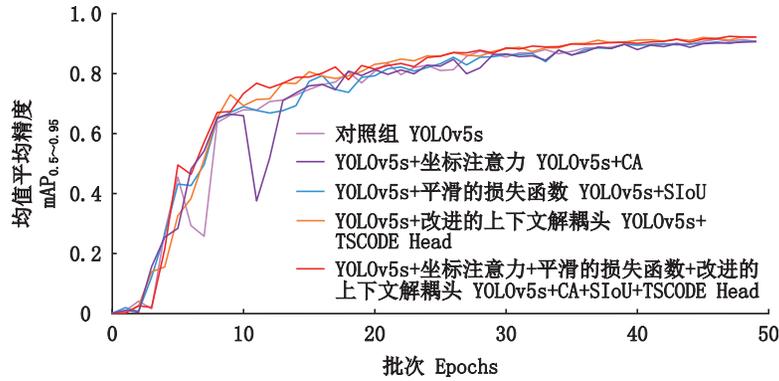


图8 消融实验 mAP_{0.5-0.95} 曲线图
Fig. 8 Ablation experiment mAP_{0.5-0.95} curve graph

3.5 对比实验

为了证明本文所提 CST-YOLO 算法对贻贝目标检测的有效性,将其与 YOLOv5s 算法、Faster R-CNN 算法在自制贻贝数据集上进行实验对比,实验结果如表 4 所示,改进前后的检测效果对比图如图 9 所示。

实验结果显示, Faster R-CNN^[14] 作为以 ResNet50 为主干网络的二阶段目标检测算法,模

型结构大,检测速度慢,mAP_{0.5-0.95}为 91.3%。CST-YOLO 算法的检测速度比 Faster R-CNN 算法提高了 83 倍,虽然相比 YOLOv5s 算法有所下降,但仍能满足实时性检测需求。CST-YOLO 算法的 mAP_{0.5-0.95} 为 92.221%,比 YOLOv5s 算法提升了 1.583%。实验结果表明,CST-YOLO 算法在保证检测速度的同时提高了检测精度,在贻贝目标检测方面表现出优越性。

表4 对比实验结果
Tab. 4 Comparison of experimental results

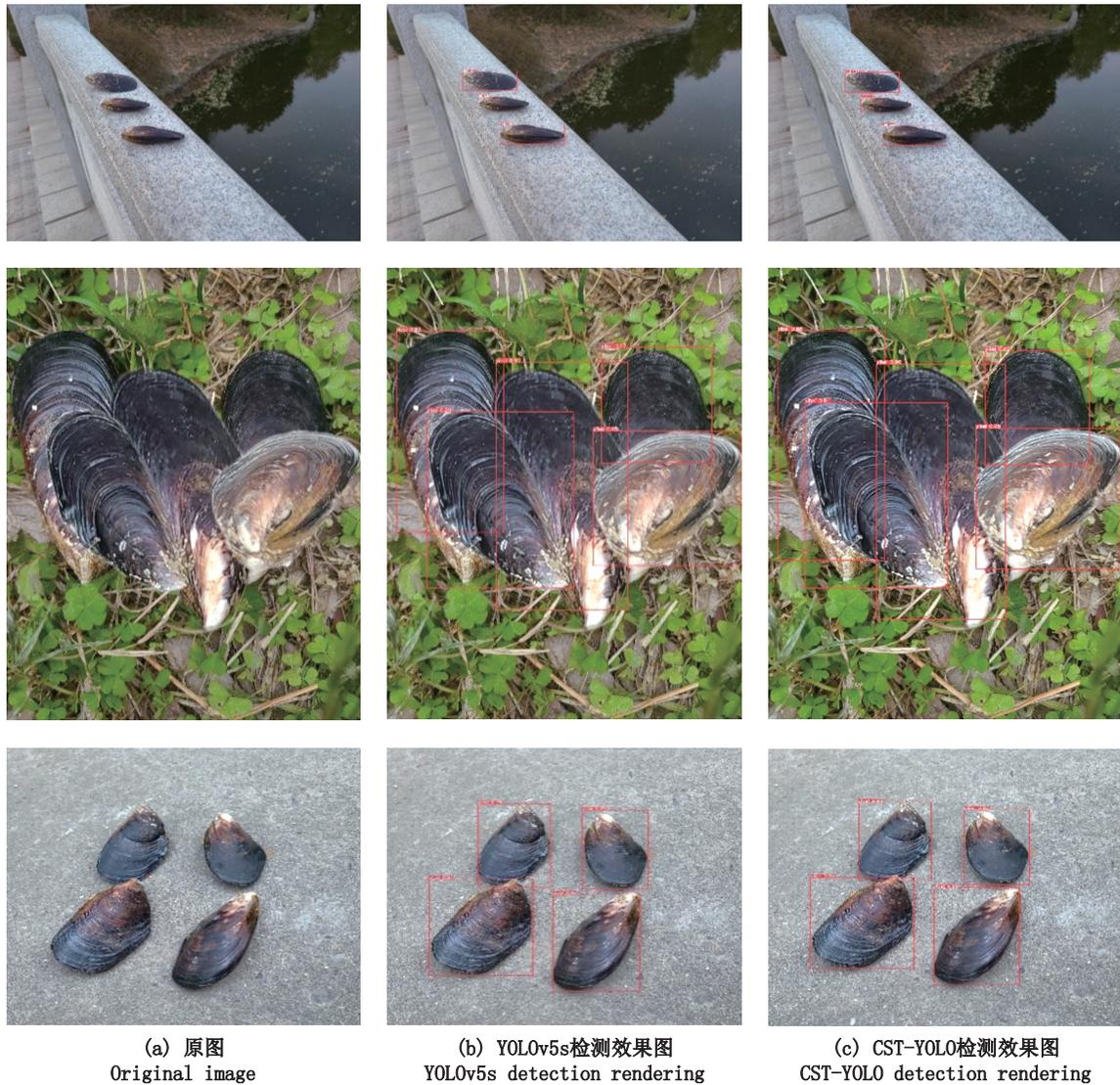
网络模型 Network model	主干网络 Backbone	检测时间 Detection time/s	mAP _{0.5} /%	mAP _{0.5-0.95} /%
Yolov5s	CSPDarknet	0.025	99.497	90.628
Faster R-CNN	ResNet50	4.968	100.000	91.300
CST-YOLO	CSPDarknet	0.059	99.500	92.221

4 结论

本文提出了一种基于改进 YOLOv5 算法的贻贝目标检测方法 (CST-YOLO),旨在有助于贻贝分级、分苗等作业环节机械化和智能化地实现。主要改进在:引入 CA 注意力机制,以提升特征表达能力,使模型更加关注目标区域的重要特征,提高目标的检测准确性;采用 SIOU 损失函数,减少边界框回归损失,使其在复杂背景下的检测表

现更加鲁棒;使用改进的 TSCODE Head,以增强定位信息的精确性,特别是在目标存在尺度变化和相互遮挡的情况下。实验结果验证了 CST-YOLO 算法在贻贝特征识别方面的适用性。

然而,在复杂背景、小目标检测、重叠遮挡等方面的研究仍存在不足,在未来的研究中,将进一步探索和优化目标检测算法,以应对更复杂的环境和更多样化的目标场景,推动机器视觉技术在实际应用中的发展。



(a) 原图
Original image

(b) YOLOv5s检测效果图
YOLOv5s detection rendering

(c) CST-YOLO检测效果图
CST-YOLO detection rendering

图9 YOLOv5和CST-YOLO检测效果对比

Fig. 9 Comparison of YOLOv5 and CST-YOLO detection effects

参考文献:

- [1] 徐林通, 郝俊, 郑艳坤, 等. 中国贝类资源现状及管理问题探讨[J]. 科技创新导报, 2018, 15(16): 201, 203.
XU L T, HE J, ZHEN Y K, et al. Discussion on the status quo and management of shellfish resources in China [J]. Science and Technology Innovation Herald, 2018, 15(16): 201, 203.
- [2] 林艾光, 孙宝元, 矢田贞美. 基于机器视觉的虾夷扇贝分级检测方法研究[J]. 水产学报, 2006, 30(3): 397-403.
LIN A G, SUN B Y, YADA S. Studies on the detecting method of scallop grading based on machine vision [J]. Journal of Fisheries of China, 2006, 30(3): 397-403.
- [3] LIU Z H, JIA X J, XU X S. Study of shrimp recognition methods using smart networks [J]. Computers and Electronics in Agriculture, 2019, 165: 104926.
- [4] HAN F L, YAO J Z, ZHU H T, et al. Marine Organism Detection and Classification from Underwater Vision Based on the Deep CNN Method [J]. Mathematical Problems in Engineering, 2020, 2020: 3937580.
- [5] 陈旭, 彭冬亮, 谷雨. 基于改进YOLOv5s的无人机图像实时目标检测[J]. 光电工程, 2022, 49(3): 210372.
CHEN X, PENG D L, GU Y. Real-time object detection for UAV images based on improved YOLOv5s [J]. Opto-Electronic Engineering, 2022, 49(3): 210372.
- [6] 黄英来, 艾昕. 改进残差网络在玉米叶片病害图像的分类研究[J]. 计算机工程与应用, 2021, 57(23): 178-184.
HUANG Y L, AI X. Research on classification of corn leaf disease image by improved residual network [J]. Computer Engineering and Applications, 2021, 57(23): 178-184.
- [7] 陈金广, 李雪, 邵景峰, 等. 改进YOLOv5网络的轻量级服装目标检测方法[J]. 纺织学报, 2022, 43(10): 155-160.

- CHEN J G, LI X, SHAO J F, et al. Lightweight clothing detection method based on an improved YOLOv5 network [J]. *Journal of Textile Research*, 2022, 43 (10) : 155-160.
- [8] REDMON J, FARHADI A. YOLO9000: Better, Faster, Stronger [C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu: IEEE, 2017: 7263-7271.
- [9] HOU Q B, ZHOU D Q, FENG J S. Coordinate Attention for Efficient Mobile Network Design [C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Nashville: IEEE, 2021: 13713-13722.
- [10] ZHENG Z H, WANG P, LIU W, et al. Distance-IoU Loss: Faster and Better Learning for Bounding Box Regression [C]//Proceedings of the 34th AAAI Conference on Artificial Intelligence. New York: AAAI, 2020: 12993-13000.
- [11] WANG J Y, LIAO H Y M, WU Y H, et al. CSPNet: A New Backbone that can Enhance Learning Capability of CNN [C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). Seattle: IEEE, 2020: 390-391.
- [12] LIN T Y, DOLLÁR P, GIRSHICK R, et al. Feature pyramid networks for object detection [C]// 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu: IEEE, 2017: 2117-2125.
- [13] LIU S, QI L, QIN H F, et al. Path Aggregation Network for Instance Segmentation [C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018: 8759-8768.
- [14] REN S Q, HE K M, GIRSHICK R, et al. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 39(6): 1137-1149.

Mussel detection algorithm based on YOLO

DONG Zhaopeng¹, YUE Xiaoxue¹, TIAN Zhongxu¹, HOU Sifan¹, JIANG Lisheng²

(1. *College of Engineering Science and Technology, Shanghai Ocean University, Shanghai 201306, China*; 2. *Dalian Nancheng Ship Repair Co., Ltd., Dalian 116011, Liaoning, China*)

Abstract: YOLO-based mussel (*Mytilus edulis*) recognition and detection technology is the key to achieving mechanization and intelligence in the grading, seedling separation, and other operational links of mussels. However, the unclear and indefinite external features of mussels present a challenge to improving recognition accuracy. This paper proposes a mussel object detection model based on the improved YOLOv5 algorithm (CST-YOLO), which integrates the CoordAttention attention mechanism to enhance feature expression ability, uses the SIoU boundary box regression Loss function to reduce the boundary box regression loss and improve the detection speed of the model, and designs an improved decoupled head TSCODE to enhance the detection accuracy. The algorithm testing was performed on a self-constructed dataset of mussels, and the experimental results demonstrated that the CST-YOLO algorithm exhibited an improvement in accuracy (P) by 0.428% compared to the YOLOv5 algorithm, and mAP (mean Average Precision) at IoU thresholds of 0.5 and 0.95 reached 92.221%, indicating a significant increase of 1.583%. These findings highlight the effective enhancement of mussel object detection accuracy while ensuring optimal detection speed achieved by the CST-YOLO algorithm. The present study contributes to the advancement of machine vision technology in facilitating the automation and intelligent production and processing within the mussel farming industry.

Key words: mussels; target detection; attention mechanism; loss function; decoupling head