

翟先一, 魏鸿磊, 韩美奇, 等. 基于改进 YOLO 卷积神经网络的水下海参检测[J]. 江苏农业学报, 2023, 39(7): 1543-1553.
doi:10.3969/j.issn.1000-4440.2023.07.011

基于改进 YOLO 卷积神经网络的水下海参检测

翟先一¹, 魏鸿磊¹, 韩美奇², 黄 萌¹

(1. 大连工业大学机械工程与自动化学院, 辽宁 大连 116034; 2. 中国科学院空天信息创新研究院/传感技术国家重点实验室, 北京 100190)

摘要: 为了实现水下海参的自动化捕捞, 需要利用机器视觉方法实现水下海参的实时检测与定位。本研究提出一种基于改进 YOLOv5s 的水下海参检测定位方法。针对海参与水下环境对比度较低的问题, 引入多尺度视觉恢复算法对图像进行处理, 增强图像对比度; 为了提高模型特征提取能力, 加入了注意力机制模块; 原始模型对 YOLOv5s 小目标的检测效果不佳, 改进后的 YOLOv5s 模型替换了原有的激活函数, 并在 Head 网络中加入了新的针对小目标的 Detect 层。使用改进的 YOLOv5s 模型与 YOLOv5s、YOLOv4 和 Faster-RCNN 在相同的图像数据集上进行试验, 结果表明, 改进的 YOLOv5s 模型的检测精度和置信度, 尤其是对小目标的检测效果优于其他模型。与 YOLOv5s 模型相比, 改进后的 YOLOv5s 模型的精度和召回率分别提高了 9.6 个百分点和 12.4 个百分点, 能够满足水下海参的实时检测要求。

关键词: YOLO; 目标检测; 深度学习; 机器视觉; 卷积神经网络

中图分类号: S126 **文献标识码:** A **文章编号:** 1000-4440(2023)07-1543-11

Underwater sea cucumber identification based on improved YOLO convolutional neural network

ZHAI Xian-yi¹, WEI Hong-lei¹, HAN Mei-qi², HUANG Meng¹

(1. School of Mechanical Engineering and Automation, Dalian Polytechnic University, Dalian 116034, China; 2. Aerospace Information Research Institute, Chinese Academy of Sciences/State Key Laboratory of Transducer Technology, Beijing 100190, China)

Abstract: To realize automatic fishing of sea cucumber underwater, it is necessary to use machine vision method to realize real-time detection and positioning of underwater sea cucumber. In this study, a detection and localization method for underwater sea cucumber based on improved YOLOv5s was proposed. Aiming at the problem of low contrast between the sea cucumber and the underwater environment, a multi-scale vision restoration algorithm was introduced to process the images to enhance the contrast of the images. The attention mechanism module was added to improve the feature extraction ability of the model. The original model didn't show good detection effect on small object of YOLOv5s. The improved YOLOv5s model replaced the original activation function and added a new Detect layer into the Head network which aimed

at small object. The improved YOLOv5s model was used to conduct experiments with YOLOv5s, YOLOv4 and Faster-RCNN on the same image data set. The results showed that, the improved YOLOv5s model showed better detection accuracy and degree of confidence compared with other models, especially for small target detection. Compared with the YOLOv5s model, the precision and recall rate of the improved YOLOv5s model increased by 9.6 percentage points and 12.4 percentage points respectively,

收稿日期: 2022-09-14

基金项目: 辽宁省教育厅 2021 年度科学研究经费面上项目 (LJKZ0535、LJKZ0526); 2021 年度大连工业大学本科教育教学综合改革项目 (JGLX2021020、JCLX2021008); 大连工业大学研究生创新基金项目 (2023CXYJ13)

作者简介: 翟先一 (1998-), 男, 天津人, 硕士研究生, 研究方向为机器视觉和目标检测。 (E-mail) zhaixianyi9421@foxmail.com

通讯作者: 魏鸿磊, (E-mail) weihl2005@163.com

which could meet the requirement of real-time detection of underwater sea cucumber.

Key words: YOLO (You only look once); object identification; deep learning; computer vision; convolutional neural network

随着生活水平的提高,人们对海参等海珍品的需求量越来越大。底播养殖是海参的主要养殖方式,即将海参苗播撒在海底,待长成后进行捕捞。由于海参生长环境特殊,主要靠人工潜水进行捕捞,捕捞过程危险、收益低,因此养殖企业需要自动化的装备来取代人工作业^[1]。海参自身颜色和环境颜色较为接近,区分度较低。为了能够在水下环境更加准确地捕捉海参,研究海参的精准快速检测,对研究海参采捕机器人,从而实现水下捕捞的自动化有较大的价值。

在海参检测的研究中,众多学者已经做出了研究和探索,李娟等^[2]基于计算机视觉技术进行研究,对水下海参进行边缘检测,提取海参刺的质心,并通过椭圆拟合进行水下海参检测,检测准确率为 93.33%;随着机器学习等相关学科的飞速发展,其在水下海参检测领域也逐渐被使用。王伟等^[3]提出一种基于改进的单激发多框探测器(SSD)算法进行水下目标检测,此方法检测精度达到 95.06%;林宇等^[4]提出了基于 Cascade RCNN 的水下目标检测,通过调节超参数,有效提高了对小目标的检测效果,平均准确率为 89.10%;马孔伟^[5]在水下机器人中应用了基于卷积神经网络(CNN)的海参检测技术,将轻量化模型 MobileNet 与 SSD 模型相结合,平均精度达到 90.06%,且采用卷积核剪枝对特征提取网络进行修剪,模型尺寸缩减 75%;王璐等^[6]提出的基于 Faster-RCNN 的水下目标检测算法研究中,将更深层次的神经网络 ResNet-101 替换原始的 VGG-16 网络,对增强数据集的平均测试精度达到 63.03%。

上述方法利用海参的形态特征、外观特点和背景颜色差异等单个特征或多个特征组合进行海参的检测,但是绝大多数算法都是面向实验室环境,海参所处的实际水下环境中,因为水草、其他海洋动物以及人类垃圾的影响,存在过多不确定因素;海参习惯群居,一旦出现重叠和遮挡等情况,检测准确率会大幅降低。

2015 年,Joseph Redmon 提出了单阶段检测(One-stage)算法(You only look once version 1, YOLOv1)。与 YOLO 算法同样经典的目标检测方法

RCNN(Region with CNN features)与 YOLO 的本质区别是,RCNN 检测时分为分类问题和回归问题 2 步运算,YOLO 只进行一步回归问题求解,因此 YOLO 算法处理图像速度更快;YOLO 利用全图信息进行训练和预测以降低背景预测错误率,由于 RCNN 无法利用全局图像,所以在检测过程中可能会将背景中的斑块错误地检测为目标。徐建华等^[7]基于 YOLOV3 算法提出的水下目标检测跟踪方法实现了水中目标的检测定位与跟踪,平均准确率为 75.10%,检测速度为 15 FPS;张聪辉^[8]基于轻量化的 tiny-YOLOv3 网络提出了海参目标检测算法,算法对原始 YOLO 进行了修改优化,检测精度达到 90%以上,在低性能运算平台上仍然能达到 47 FPS;朱世伟等^[9]提出的基于类加权 YOLO 网络的水下目标检测,平均准确率为 78.30%,处理速度为 12 FPS,对多目标检测效果提升明显。因此 YOLO 算法兼具较高的准确率和内存占用小的优点。

针对海参与其所处水下环境在自然条件下区分度低和目前算法小目标检测不精确的问题,本研究提出一种基于改进 YOLOv5 的水下海参检测算法。采用 MSRCR 算法对图像进行预处理,以增强海参与环境的区分度,增加模型中 Neck 网络的上采样次数,在 Head 网络中增加新的 Detect 层,引入更适合目标检测的激活函数,以提高模型检测较小海参目标的能力;压缩和激励网络(Squeeze and excitation networks, SENet)可提高模型训练效率,增强特征提取能力,作为即插即用模块集成到现有网络架构中。本研究对改进的 YOLOv5s 模型进行消融研究,验证改进的可行性,与 YOLOv5s、YOLOv4 和 Faster-RCNN 相比,改进后的 YOLOv5s 具有更高的准确率和召回率,对小目标海参检测效果优异。

1 材料与方法

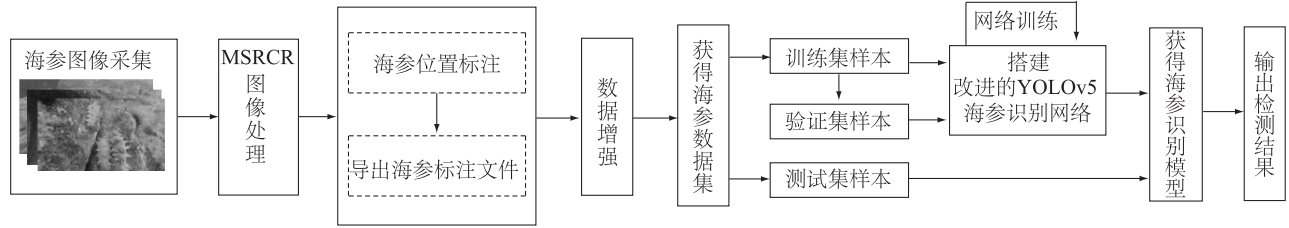
水下海参检测模型的训练与检测过程如图 1 所示。首先对图像进行 MSRCR 算法处理,然后使用标签工具 Make Sense 对图像进行标注,完成后导出 txt 格式的标签。使用数据扩增对图像数据集进行扩充,最后使用改进的 YOLOv5s 模型对数据集进行

训练,得到模型并测试分析模型性能。

1.1 数据集的获取

本研究所使用的海参图为潜水员在獐子岛海参养殖水域拍摄的海参图像,拍摄时间为 2021 年 5 月

19 日 9:00–11:00,自然光条件下采集,采集时,潜水员使用高清数码相机在距离海参 30~70 cm 的地方拍摄,图像分辨率为 640 像素×360 像素,共获得 892 幅图像,其中部分海参图像如图 2 所示。



MSRCR:带颜色恢复的多尺度视网膜增强方法。

图 1 海参检测流程

Fig.1 Sea cucumber identification flowchart



图 2 部分海参图像

Fig.2 Part of the sea cucumber images

1.2 图像预处理

由于水下采集到的图像亮度过低,存在失真、难以明显区分海参与背景的问题,给模型的训练造成困难,故引入带颜色恢复的多尺度视网膜增强方法(Multi-scale retinex with color restoration, MSRCR)对图像进行预处理,以增强海参与背景的区分度,MSRCR 具体实现如公式(1)^[10-11]所示:

$$R_{\text{MSRCR}}(x, y) = C_i(x, y) R_{\text{MSR}_i}(x, y) \quad (1)$$

式中, $R_{\text{MSRCR}}(x, y)$ 表示使用 MSRCR 算法后第 i 个通道恢复后的图像; $C_i(x, y)$ 为 MSRCR 加入的色彩恢复因子,如公式(2)所示,用来调节原始图像颜色通道的比例,突出图像暗部区域信息,消除颜色失真,公式(2)中 α 是调节因子, β 是增益常数, $I(x, y)$ 表示原图像^[12]。 $R_{\text{MSR}_i}(x, y)$ 表示使用多尺度滤波后第 i 个通道的恢复图像[公式(3)], $G_n(x, y)$ 是单尺度的高斯滤波, λ_n 表示权值, m 为尺度数目,取值一般为 $3^{[13-14]}$ 。 n 为从 1 到 3 的尺度数目。

$$C_i(x, y) = \beta \{ \lg[\alpha \cdot I_i(x, y)] - \lg[\sum_{i=1}^N I_i(x, y)] \} \quad (2)$$

$$R_{\text{MSR}_i}(x, y) = \sum_{n=1}^m \lambda_n \{ \lg I(x, y) - \lg [G_n(x, y) \cdot I(x, y)] \} \quad (3)$$

MSRCR 处理前后图像对比如图 3 所示,左侧为原始图像,右侧为处理后图像,左上角显示算法运行时长。可以看出,MSRCR 算法在对水下海参图像进行处理时,平均处理时间为 26.26 ms,兼顾了处理效果和速度,每秒可处理约 38 张图像,能够满足实时视频检测要求。

1.3 数据增强

在深度学习训练过程中,数据集图像的数量对模型有一定的影响。由于采集的海参图像数量有限,在训练之前需要进行数据增强^[15]。在实际使用海参捕捞机器人的过程中,由于光线、距离、角度、水流等引起的抖动,经常会出现角度、位置、大小、亮度的变化,以及图像模糊等情况。为了提供更全面的训练图像,利用 YOLOv5 自带的数据增强模块,结合 Pytorch 框架下的 OpenCV,对采集到的海参图像进行镜像、正负角度旋转 45°、平移、随机缩放、随机裁剪、亮度变化等操作,共得到 5 017 幅训练集图像。

26.26 ms (算法运行时长)

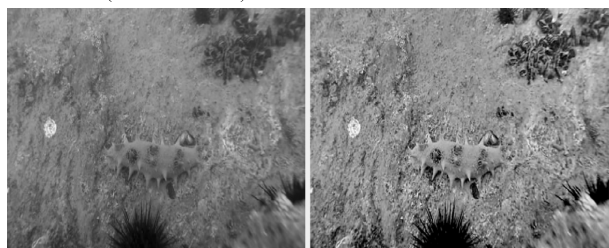


图3 图像处理前(左)后(右)对比

Fig.3 Comparison of images before (left) and after (right) processing

1.4 数据集准备

为训练图片添加标签时使用线上标注网站 Make Sense (<https://www.makesense.ai/>), 标注规则为: 对于形态完整的海参使用矩形框标注; 对于有遮挡或有重叠的海参, 对显示在图像中的部分进行标注; 对位于图像边框以及显示区域小于 15% 的海参不进行标记。标记完成后导出视觉目标分类 (VOC) 类型数据集。选取数据集中图像的 80% 作为训练集, 10% 作为验证集, 10% 作为测试集。

2 海参检测网络

在最新版本的 YOLOv5 官方代码中, 共提供了 YOLOv5n、YOLOv5s、YOLOv5m、YOLOv5l、YOLOv5x 5 种不同深度的模型^[16]。本次研究由于检测的海参数量较少, 且对速度要求较高, 所以选择了兼顾速度和精度的 YOLOv5s 作为基础模型。

2.1 YOLOv5s 网络模型

YOLOv5s 网络模型如图 4 所示, 包括 Backbone、Neck 和 Head 等。YOLOv5s 的 Backbone 网络采用 CSPDarknet53 结构作为基础网络, Backbone 的主要作用是提取输入信息的基本特征^[16]。Neck 网络位于 Head 网络和 Backbone 网络之间, 其主要功能是进一步增强待提取特征的多样性和网络的鲁棒性。Head 网络的作用是检测目标的输出结果。YOLOv5s 模型有 3 个 Detect 层, Head 网络接收 3 个不同尺度的特征图, 然后对这 3 个特征图进行网格预测, 通过卷积运算得到特征输出, 主要通过大目标、中目标和小目标 3 个维度进行检测^[17]。

与 YOLOv4 相比, YOLOv5 增加了 Focus 结构, 用来防止下采样时信息丢失。图像进入 Backbone 网络之前, Focus 模块对其进行切片。具体操作是

在图片中每隔一个像素取 1 个值, 得到 4 张互补的图片, 且没有信息丢失^[18]。此操作将水平和垂直的信息集中在通道空间, 输入通道扩大了 3 倍, 即与原来的红绿蓝 (RGB) 三通道模式相比, 拼接后的图片变成了 12 通道^[19-20]。最后对得到的新图像进行卷积, 得到无信息丢失的 2 倍下采样特征图。

以 YOLOv5s 为例, 原 $640 \times 640 \times 3$ 图像输入 Focus 结构, 通过切片操作, 先生成 $320 \times 320 \times 12$ 的特征图, 再经过卷积运算, 最终生成 $320 \times 320 \times 32$ 的特征图^[21]。切片操作如图 5 所示。

2.2 损失函数

交并比 (Intersection over union, IoU) 是目标检测中常用的评价指标, 一般用来区分正负样本, 反映预测框和真实框的距离^[22-24]。计算方法如公式 (4) 所示。

$$IoU = \frac{A \cap B}{A \cup B} \quad (4)$$

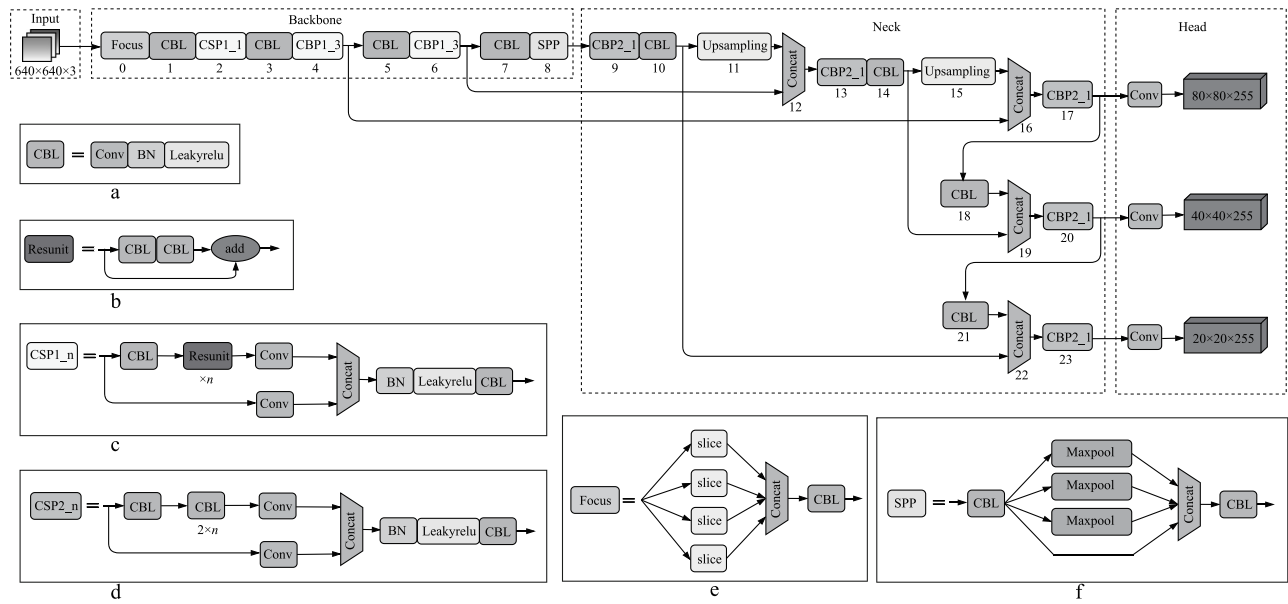
IoU 算法应用广泛, 对尺度变化不敏感, 具有非负性等。但是由于 IoU 并没有考虑真实框与预测框之间的距离, 所以直接将 IoU 作为 Loss 函数时会暴露各种问题: 若预测框与真实框之间没有重合部分, 根据定义 IoU 为 0, 此时不能计算出 2 个框之间的距离; 由于不同方向、不同角度的对齐方式不同, 会导致 IoU 完全相同, 此时不能判断预测框与真实框的重合情况。YOLOv5 算法采用的是 Generalized intersection over union_loss ($GIoU_Loss$) 作为边界框回归的损失函数^[25-26]。 $GIoU$ 在关注重叠部分的同时, 也关注其他非重叠部分, 可以更精准地描述预测框和真实框的各类重叠情况。假设 A 为预测框, B 为真实框, 令 C 表示包含 A 与 B 的最小闭合框。 $GIoU_Loss$ 的计算公式如下:

$$GIoU_Loss = 1 - GIoU = IoU - \frac{|C/(A \cup B)|}{|C|} \quad (5)$$

2.3 YOLOv5s 网络的改进

水下海参个体的姿态和大小差异较大, YOLOv5 对于小目标的检测并不理想, 在实际测试中, 一些海参不能被检测或检测不精确。为了提高对小目标的检测能力, 对 YOLOv5s 结构进行了改进。

2.3.1 增加注意力机制模块 为了提高 YOLO 训练和检测的效率, 引入了 SENet 注意力机制。SENet 首先将输入的特征图进行全局池化, 将特征图变成 $1 \times 1 \times C$ 通道数, 然后使用全连接层和激活函数对 $1 \times$



Input: 输入; Backbone: 主干网络; Leakyrelu: 修正线性单元; Focus: 切片操作; Slice: 切片; Concat: 融合; Upsampling: 上采样; Neck: 颈部网络; Head: 头部网络。

图4 YOLOv5s 网络模型

Fig.4 YOLOv5s network model

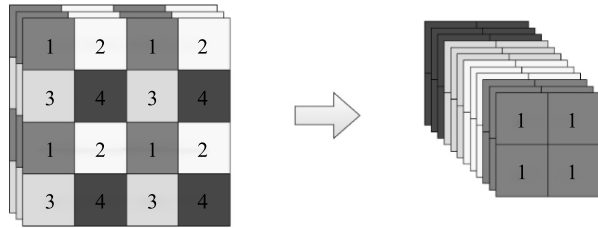


图5 Focus 结构切片操作

Fig.5 Slice operation of focus structure

$1 \times C$ 通道数的特征图进行调整,得到每一个特征图的权重,最后与输入的特征进行相乘。SENet 结构如图 6 所示,它能够显著提高检测性能,同时保持较小的开销。此次试验中,将 SENet 模块添加在 CBL 结构的 Sigmoid 结构中后,如图 7a 所示。

2.3.2 替换激活函数 YOLOv5s 使用的是 Leakyrelu 激活函数,在回归任务中效果更为明显,故而将 Leakyrelu 替换为更适合二分类的激活函数 Sigmoid。更改后的激活函数如图 7a 所示。

2.3.3 增加小目标 Detect 层 原始 YOLOv5s 共进行 2 次上采样,本研究方法在第 17 层后,对特征图进行上采样处理,使得特征图进一步增大;在第 23 层时,将骨干网络中第 2 层特征图与上采样获取的特征图进行特征融合,得到更大的特征图来改善小目标检测不精准的问题。在第 30 层以后添

加一个新的小目标 Detect 层,综合以上,改进后的模型如图 7 所示。

3 模型训练和测试

3.1 试验平台

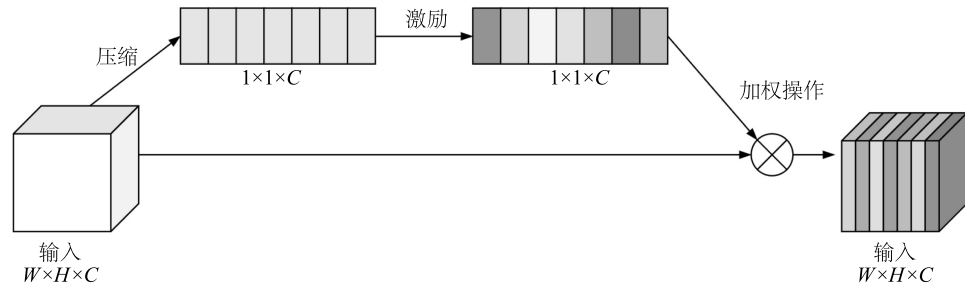
本试验环境如表 1 所示。

3.2 消融试验

为了验证模型改进的可行性,进行消融试验。

3.2.1 图像预处理的消融试验 首先使用未经过处理的水下图像作为训练集进行模型训练,得到模型 NONE-YOLO。再对相同的图像,使用 MSRCR 水下图像处理算法进行图像预处理,并作为训练集训练相同周期,得到模型 MSRCR-YOLO。模型对比结果如表 2 所示。由表 2 可以看出,MSRCR-YOLO 相比于 NONE-YOLO 模型,精确率提升 2.1 个百分点,训练时长、权重大小和检测速度几乎无变化。由于加入了 MSRCR 图像预处理,图像质量有了一定提高,YOLO 模型更加容易提取出海参的特征,准确率有一定提升。

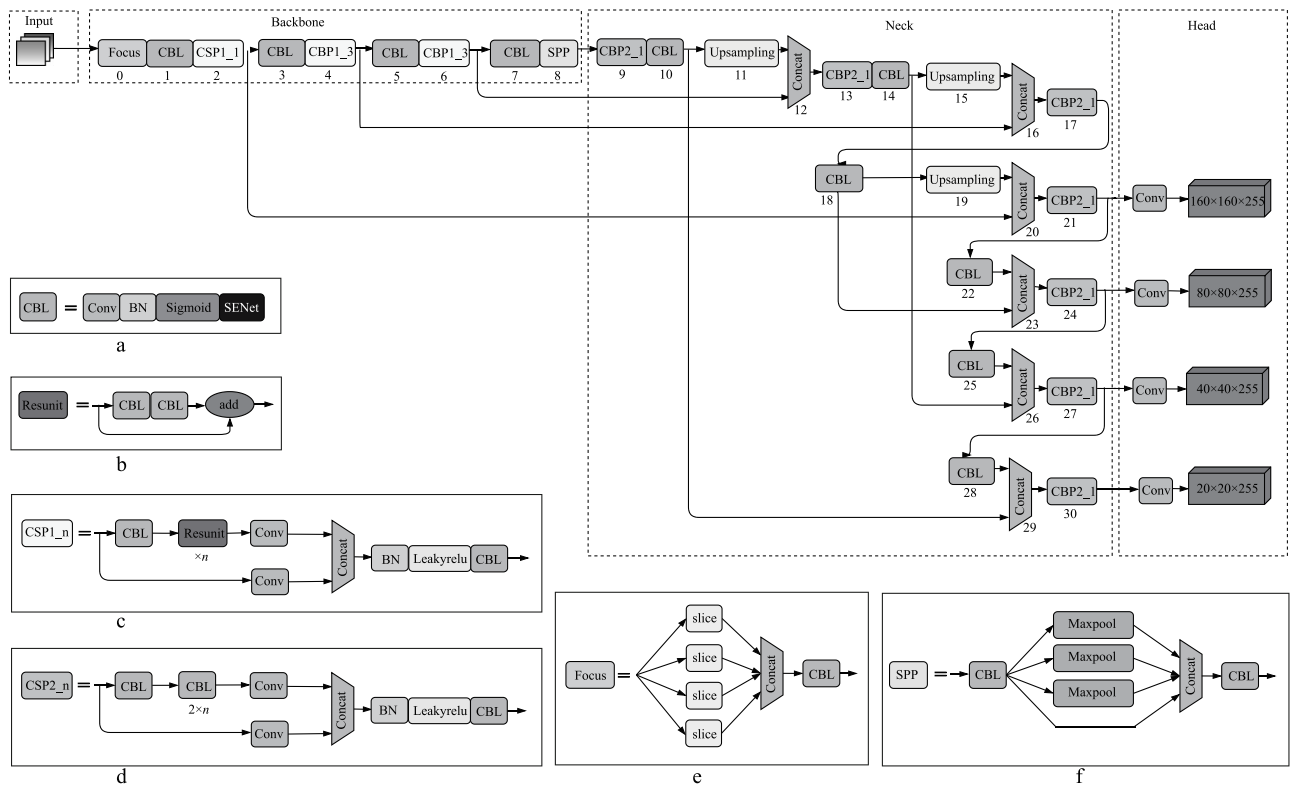
为了显示 2 个模型的区别,使用 2 个模型对相同图像进行测试,测试结果如图 8 所示。可以看出,MSRCR-YOLO 模型相比于 NONE-YOLO 模型在海参检测的置信度方面有一定提升。



W 、 H 、 C 分别表示输入图像的长、宽、通道数。

图 6 注意力机制模块

Fig.6 Module of attention mechanism



Input:输入;Backbone:主干网络;Leakyrelu:修正线性单元;Focus:切片操作;Slice:切片;Concat:融合;Upsampling:上采样;Neck:颈部网络;Head:头部网络。

图 7 改进后的 YOLOv5s

Fig.7 Improved YOLOv5s

表 1 试验环境

Table 1 Experimental environment

配置	参数
CPU	Intel Core i7-8750H
GPU	NVIDIA GeForce GTX1660Ti 6 G
操作系统	Windows11
加速环境	CUDA11.6
开发平台	PyCharm2022.2.1

CPU:中央处理器;GPU:图像处理器。

表 2 NONE-YOLO 和 MSRCR-YOLO 模型性能对比

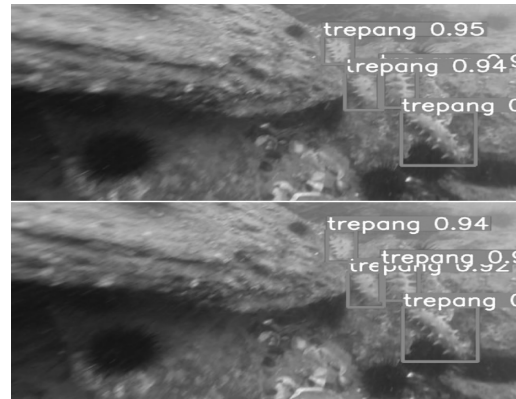
Table 2 Comparison of performance between NONE-YOLO model and MSRCR-YOLO model

模型	精确率 (%)	训练时长 (h)	权重 (Mb)	检测时间 (ms)
NONE-YOLO	88.4	5.7	18.4	19.2
MSRCR-YOLO	90.5	5.7	18.5	19.4

3.2.2 模型改进的消融试验 使用改进的 YOLOv5s

与原始 YOLOv5s、YOLOv5s + SENet 和 YOLOv5s + SENet+Detect 进行比较。选取1 035张海参图像作为数据集,训练集、测试集和验证集的比例为8 : 1 : 1,使用同一平台进行训练,训练迭代次数(*Epoch*)设置为200,最终结果如表3所示。结果表明,与 YOLOv5s 原始模型相比,在 YOLOv5s 中只添加 SENet 模块、只添加 Detect 层、同时添加 SENet 模块和 Detect 层3种情况都可以提高精确率,但由于 YOLOv5s 模型的复杂度增加,会导致资源占用增加,训练时间会变长,训练产生的权重会增加,检测图像的时间会略有增加。

为了更加直观地对比4种算法的区别,使用相同图像进行对比测试。对比结果如图9所示。从图9可以看出,小目标检测效果算法4,即本研究改进的 YOLOv5s 模型的效果相对更优,置信度相对更高。故本试验对于模型的改进具有可行性。



图片未能完整显示预测框置信度,NONE-YOLO 预测框置信度为0.94、0.93、0.92、0.92;MSRCR-YOLO 预测框置信度为0.95、0.95、0.93、0.92。

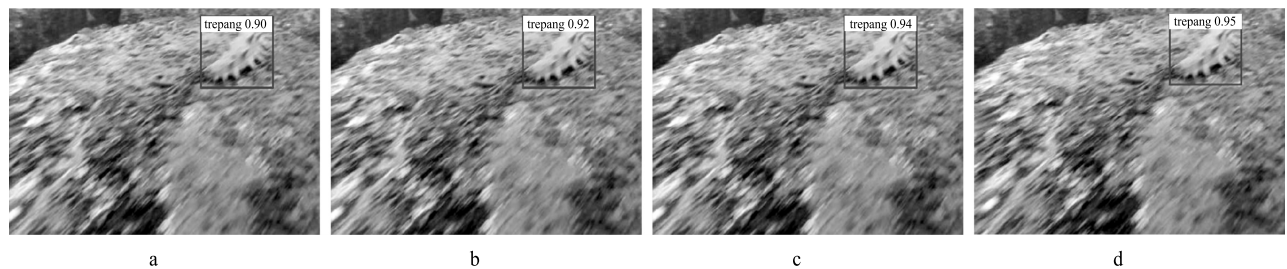
图8 使用水下图像处理算法前后模型检测效果对比

Fig.8 Comparison of model detection effect before and after using underwater image processing algorithm

表3 改进 YOLOv5s 对比原始 YOLOv5s、YOLOv5s+SENet、YOLOv5s+ SENet+Detect

Table 3 Improved YOLOv5s model compared with YOLOv5s, YOLOv5s+SENet and YOLOv5s+SENet+Detect

算法序号	模型	精确率 (%)	训练时长 (h)	权重 (Mb)	每张图像检测时间 (ms)
1	YOLOv5s	84.1	4.3	14.6	16
2	YOLOv5s+SENet	87.7	4.7	15.1	20
3	YOLOv5s+Detect	88.2	5.1	15.7	22
4	YOLOv5s+SENet+Detect	91.8	5.6	16.3	24



a 为 YOLOv5s 检测效果;b 为 YOLOv5s+SENet 检测效果;c 为 YOLOv5s+Detect 检测效果;d 为 YOLOv5s+SENet+Detect 检测效果。

图9 算法序号1~4 海参检测效果对比

Fig.9 Comparison of sea cucumber detection effect by Algorithms 1-4

3.3 模型训练及测试

为了避免过拟合,试验中采用5-Folder交叉验证方法,如图10所示。

图11为训练结果。图11a为精度图,即找对的正类(*TP*)/所有找到的正类(*TP+FP*);精度(*precision*)计算方法如式(6)所示,在*IoU*设置为0.5时,经过改进的模型精度呈现稳定上升趋势,100次迭代以后,训练结果精度达到99.0%。

$$precision = \frac{TP}{TP+FP} \quad (6)$$

图11b为召回率随训练次数变化曲线,即找对的正类(*TP*)/所有本应该被找对的正类(*TP+FN*);召回率(*recall*)计算公式如式(7)所示,在*IoU*设置为0.5时,经过改进的模型呈现稳定上升趋势,100次迭代以后,训练结果精度达到99.6%。

$$recall = \frac{TP}{TP+FN} \quad (7)$$

在目标检测中,测量检测精度的方法是平均精度(mAP)。在多类目标检测中,每一类都可以根据召回率和精度绘制一条曲线, AP 为曲线下的面积, mAP 为多类 AP 的平均值。 AP 值越高,模型的平均精度越高。 $mAP_{0.50}$ 和 $mAP_{0.50:0.95}$ 是 2 个最常用的指标。 $mAP_{0.50}$ 表示当 IoU 设置为 0.50 时,计算每一类所有图像的 $AP^{[27]}$,然后对所有类别取平均值; $mAP_{0.50:0.95}$ 为不同 IoU 阈值(0.50~0.95,步长 0.05,即 0.50、0.55、0.60、0.65、0.70、0.80、0.85、0.90、0.95)下的平均 mAP 。图 11c 表示 $mAP_{0.50}$ 随迭代次数变化曲线,图 11d 表示 $mAP_{0.50:0.95}$ 随迭代次数变化曲线。

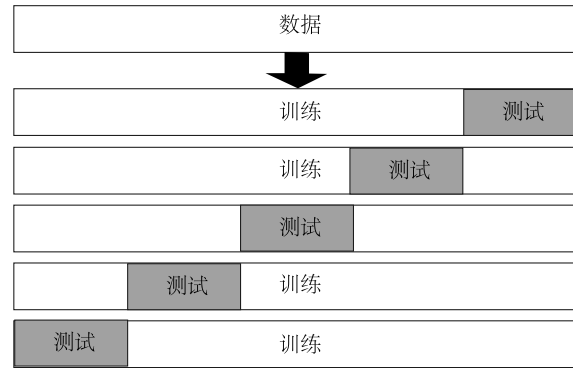
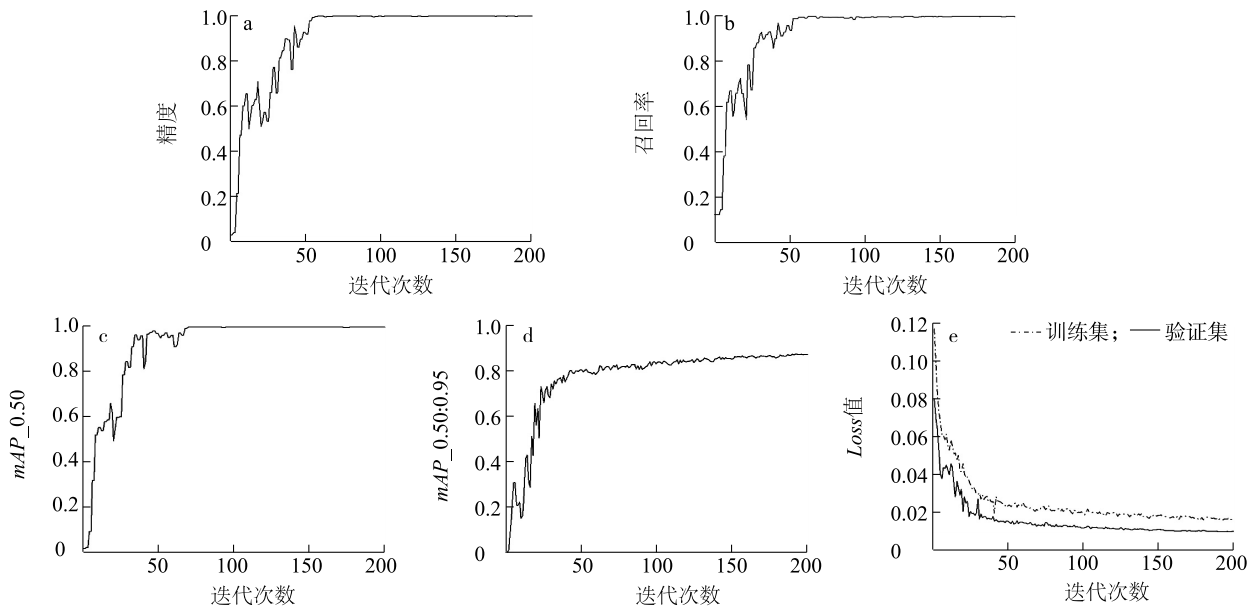


图 10 5-Folder 交叉验证

Fig.10 Five-folder cross validation



a:精度随迭代次数变化曲线;b:召回率随迭代次数变化曲线;c: $mAP_{0.50}$ 随迭代次数变化曲线;d: $mAP_{0.50:0.95}$ 随迭代次数变化曲线;e:损失函数随迭代次数变化曲线。 $mAP_{0.50}$ 表示当 IoU 设置为 0.50 时,计算每一类所有图像的 AP (曲线下的面积),然后对所有类别取平均值; $mAP_{0.50:0.95}$ 为不同 IoU 阈值(0.50~0.95,步长 0.05,即 0.50、0.55、0.60、0.65、0.70、0.80、0.85、0.90、0.95)下的平均精度。

图 11 各参数随迭代次数变化的曲线

Fig.11 Curve of parameters varied with iteration times

图 11e 显示了训练集和验证集的损失曲线,数值越小表示预测框越精确^[28-29]。当 IoU 设置为 0.5 时,随着训练周期的增加,训练集和验证集的损失曲线呈下降趋势,最终趋于稳定;100 次迭代后,训练集和验证集的损失函数值达到最小值(0.02 左右)。

4 测试和分析

为了验证水下海参目标检测算法的性能,采用改进的 YOLOv5s、YOLOv5s、YOLOv4 和 Fast-RCNN 模型对单图像库和混合图像库进行训练和测试。

4.1 试验结果对比

表 4 比较了 IoU 设置为 0.5 时改进后的 YOLOv5s 与其他 3 种模型的测试结果。与 YOLOv5s 模型相比,改进后的 YOLOv5s 模型准确率提高约 9.6 个百分点,召回率提高约 12.4 个百分点。由于改进的方法增加了计算量,检测速度相比于原始 YOLOv5s 算法略有下降,但是相比于其他算法仍有一定优势。

4.2 海参图像检测

使用改进的 YOLOv5s、YOLOv5s、YOLOv4 和

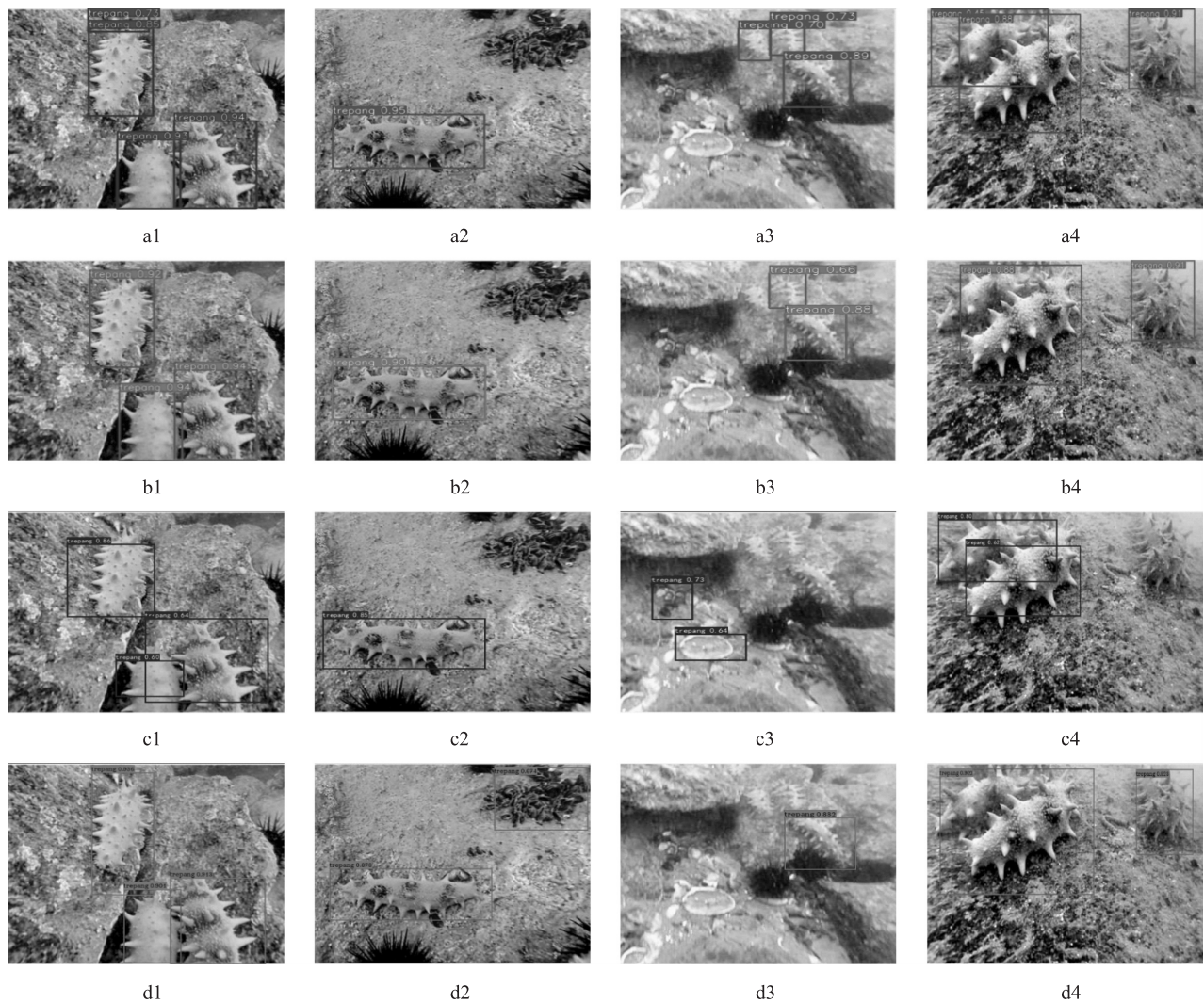
Faster-RCNN 对相同的海参图像进行检测,试验结果如图 12 所示。

从图 12 可以看出,改进的 YOLOv5s 模型对图像中小目标的检测效果高于其他 3 种模型,预测框的选择范围更精确,对同一目标的置信度更高。例如,在 YOLOv5s 的检测结果中,在图 12b1、图 12b3 和图 12b4 中存在漏检;在 YOLO4 的检测结果中,图 12c1 存在预测框不精确的问题,图 12c3 存在错误检测,图 12c4 存在漏检;在 Faster-RCNN 结果中,图 12d1、图 12d3、图 12d4 相比于 YOLOv5s 检测到的海参数量,存在漏检,图 12d2 有错误检测。

表 4 改进的 YOLOv5s 与 YOLOv5s、YOLOv4、Faster-RCNN 性能对比

Table 4 Comparison of properties between improved YOLOv5s model and YOLOv5s, YOLOv4 and Faster-RCNN

模型	训练时间 (h)	准确率 (%)	召回率 (%)	检测时间 (ms)
改进的 YOLOv5s	6.7	96.6	95.8	27
YOLOv5s	5.9	87.0	83.4	22
YOLOv4	7.1	77.2	76.3	32
Faster-RCNN	7.5	81.3	79.5	34



a1、a2、a3、a4 为改进的 YOLOv5s 的检测结果; b1、b2、b3、b4 为 YOLOv5s 的检测结果; c1、c2、c3、c4 为 YOLOv4 的检测结果, d1、d2、d3、d4 为 Faster-RCNN 的检测结果。

图 12 改进的 YOLOv5s、YOLOv5s、YOLOv4、Faster-RCNN 对海参图像的检测结果

Fig.12 Results of improved YOLOv5s, YOLOv5s, YOLOv4 and Faster-RCNN on sea cucumber image detection

5 结 论

本研究提出了一种基于改进 YOLOv5s 的水下海参检测方法,旨在解决小目标海参检测效果较差的问题,提高检测精度,实现自然条件下水下海参的快速精确检测。研究主要利用 MSRCR 算法对水下海参图像进行预处理,减少了图像失真,增强了对比度,提高了海参与环境的对比度,降低了模型训练的难度。根据水下海参的实际情况,选择 YOLOv5s 算法作为算法实现的基础,并对模型进行了改进:增加了 SENet 注意力机制模块,以增加少量检测时间为代价,换来了训练的效率 and 准确率的极大提高。替换了原有的激活函数,改进了 Head 网络,增加了一层新的针对小目标海参的 Detect 层,实现了对输入图像进行更多次的上采样,使得改进的 YOLOv5s 模型对小目标的检测精度进一步提高。对改进模型进行了消融试验,验证其可行性。与 YOLOv5s、YOLOv4 和 Faster-RCNN 等主流目标检测算法进行对比试验,结果表明,本研究方法对小目标的检测效果明显提高。与 YOLOv5s 相比,能检测到 YOLOv5s 无法检测的小目标海参,精确率提高了 9.6 个百分点,召回率提高了 12.4 个百分点,改进模型能够以较高的置信度检测到海参目标,小目标检测效果明显提高,具有较明显的优势。但是本算法训练时间较长,且需要额外进行图像预处理,在后续研究中将优化算法,尽量使用轻量级网络实现,并部署在移动设备或边缘设备上,同时对海参进行高精度追踪。

参考文献:

- [1] 刘吉伟,魏鸿磊,裴起潮,等. 采用相关滤波的水下海参目标跟踪[J]. 智能系统学报, 2019, 77(3): 525-532.
- [2] 李 娟,朱学岩,葛凤丽,等. 基于计算机视觉的水下海参检测方法研究[J]. 中国农机化学报, 2020, 41(7): 171-177.
- [3] 强 伟,贺显曜,郭玉锦,等. 基于改进 SSD 的水下目标检测算法研究[J]. 西北工业大学学报, 2020, 38(4): 747-754.
- [4] 林 宇,何水原. 改进 Cascade RCNN 的水下目标检测[J]. 电子世界, 2022(1): 105-108.
- [5] 马孔伟. 基于 CNN 的海参检测技术及其在水下机器人中的应用[D]. 哈尔滨:哈尔滨工业大学, 2019: 56.
- [6] 王 璐,王雷欧,王东辉. 基于 Faster-rcnn 的水下目标检测算法研究[J]. 网络新媒体技术, 2021, 10(5): 43-51,58.
- [7] 徐建华,豆毅庚,郑亚山. 一种基于 YOLO-V3 算法的水下目标检测跟踪方法[J]. 中国惯性技术学报, 2020, 28(1): 129-133.
- [8] 张聪辉. 基于 YOLO-v3 的海参目标检测系统的设计与实现[D]. 哈尔滨:哈尔滨工程大学,2020.
- [9] 朱世伟,杭仁龙,刘青山. 基于类加权 YOLO 网络的水下目标检测[J]. 南京师大学报(自然科学版), 2020, 43(1): 129-135.
- [10] WANG J, LU K, XUE J, et al. Single image dehazing based on the physical model and MSRCR algorithm[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2017, 28(9): 2190-2199.
- [11] LIU Y, YAN H, GAO S, et al. Criteria to evaluate the fidelity of image enhancement by MSRCR[J]. IET Image Processing, 2018, 12(6): 880-887.
- [12] 许伟栋,赵忠盖. 基于卷积神经网络和支持向量机算法的马铃薯表面缺陷检测[J]. 江苏农业学报, 2018, 34(6): 1378-1385.
- [13] KUMARI N, RUF V, MUKHAMETOV S, et al. J mobile eye-tracking data analysis using object detection via YOLO v4[J]. Sensors, 2021, 21(22): 7668.
- [14] 张开兴,吕高龙,贾 浩,等. 基于图像处理和 BP 神经网络的玉米叶部病害检测[J]. 中国农机化学报, 2019, 40(8): 122-126.
- [15] 尚志军,张 华,曾 成,等. 基于机器视觉的枳壳自动定向方法与试验[J]. 中国农机化学报, 2019, 40(7): 119-124.
- [16] LIU S, XU Y, GUO L M, et al. Multi-scale personnel deep feature detection algorithm based on Extended-YOLOv3[J]. Journal of Intelligent & Fuzzy Systems, 2021, 40(1): 773-786.
- [17] JOBSON D J, RAHMAN Z, WOODDELL G A. A multiscale retinex for bridging the gap between color images and the human observation of scenes[J]. IEEE Transactions on Image Processing, 1997, 6(7): 965-976.
- [18] WANG Q Z, LI S, QIN H, et al. Super-resolution of multi-observed RGB-D images based on nonlocal regression and total variation[J]. IEEE Transactions on Image Processing: A Publication of the IEEE Signal Processing Society, 2016, 25(3): 1425-1440.
- [19] ZHANG M H, XU S B, SONG W, et al. Lightweight underwater object detection based on YOLOv4 and multi-scale attentional feature fusion[J]. Remote Sensing, 2021, 13(22): 4706.
- [20] TAN L, HUANGFU T R, WU L, et al. Comparison of RetinaNet, SSD, and YOLOv3 for real-time pill identification[J]. BMC Medical Informatics and Decision Making, 2021, 21(1): 1-11.
- [21] ZHANG M, XU S, SONG W, et al. Lightweight underwater object detection based on YOLO v4 and multi-scale attentional feature fu-

- sion[J]. Remote Sensing, 2021, 13(22): 4706.
- [22] 朱 磊. Retinex 图像增强算法的研究与 FPGA 实现[D]. 北京:清华大学, 2012.
- [23] 朱秋旭,李俊山,朱英宏,等.Retinex 理论下的自适应红外图像增强[J]. 微电子学与计算机, 2013, 30(4): 22-25.
- [24] 陈旭君,王承祥,朱德泉,等. 基于 YOLO 卷积神经网络的水稻秧苗行线检测[J].江苏农业学报,2020,36(4):930-935.
- [25] LEI F, TANG F, LI S. Underwater target detection algorithm based on improved YOLOv5[J]. Journal of Marine Science and Engineering, 2022, 10:310.
- [26] QIAO X, BAO J, ZHANG H, et al. fvUnderwater sea cucumber identification based on principal component analysis and support vector machine[J]. Measurement, 2019, 133:444-455.
- [27] ZHANG L, XING B, WANG W et al. Sea cucumber detection algorithm based on deep learning[J]. Sensors, 2022, 22:5717.
- [28] XUE L, ZENG X, JIN A. A novel deep-learning method with channel attention mechanism for underwater target recognition [J]. Sensors, 2022, 22:5492.
- [29] 孟 亮,郭小燕,杜佳举,等. 一种轻量级 CNN 农作物病害图像识别模型[J].江苏农业学报,2021,37(5):1143-1150.

(责任编辑:陈海霞)