测试与故障诊断

文章编号:1671-4598(2025)06-0076-10 DOI:10.16526/j. cnki.11-4762/tp.2025.06.010 中图分类号:TP18; TP242

文献标识码:A

基于改进 YOLOv5s 的水下自主机器人 垃圾检测技术研究

祛雅清、丝 勇、 朱

(南昌工学院信息与人工智能学院,南昌 330013)

摘要:针对水下环境垃圾污染的检测问题,研究提出一种基于改进 YOLOv5s 的水下自主机器人垃圾检测方法,该 方法采用 MSCA 模块结合轻量化网络 MobileNet V3,对原模型的主干网络进行优化,以增强模型对复杂水下图像特征的 提取能力;对于颈部网络,引入了 FCS 模块,并融合 FPN、CAM 和 SPP 网络对其进行改进,以提升模型对多尺度目标 的特征融合能力;为了进一步提高模型的检测精度,研究加入了 SIoU 损失函数,全面考虑预测框与真实抗之间的差异; 在水下图像的预处理中,研究通过高斯滤波技术的双边滤波器进行降噪,以提升图像质量;实验结果表明,处理后的图 像色彩度和图像清晰度评估值分别为 2.65 和 0.59; 在复杂水环境下,改进后的 YOLOv5s 模型平均精度达到 0.948; 结 果证明所提水下垃圾检测技术能够有效提升模型的检测精度,且在实际应用中具有较优的性能。

关键词:YOLOv5s;水下机器人;垃圾检测;检测模型;检测精度

Underwater Autonomous Robot Garbage Detection Technology **Based on Improved YOLOv5s**

ZHANG Yaqing, XIONG Yong, ZHU Qing

(School of Information and Artificial Intelligence, Nanchang Institute of Science &

Technology, Nanchang 330013, China)

Abstract: Aiming at the detection problem of underwater environmental garbage pollution, an underwater autonomous robot garbage detection method based on improved YOLOv5s is proposed. This method utilizes the multi-spectral channel attention (MSCA) module and lightweight network MobileNet V3 to optimize the backbone network of the original model to enhance the feature extraction capability of the model in complex underwater images. For the neck network, an FCS module is introduced and integrated with the feature pyramid network (FPN), channel attention mechanism (CAM) and space pyramid pool (SPP) network to improve the feature fusion ability of the model on the multi-scale targets. By fully considering the difference between the prediction frame and real reactance, an SIoU loss function is added to further improve the detection accuracy of the model. In the pre-processing of underwater images, the double-sided Gaussian filter is used to reduce noise, so as to improve the image quality. Experimental results show that the evaluation values of image color and image sharpness after processing are 2.65 and 0.59 respectively. In complex water environments, the average accuracy of the improved YOLOv5s model reaches up to 0.948. The results show that the proposed underwater garbage detection technology can effectively improve the detection accuracy of the model and has a better performance in practical applications.

Keywords: YOLOv5s; underwater robot; garbage detection; detection model; detection accuracy

0 引言

水下自主机器人因其高效、安全和能够覆盖广阔水 域的优势,已成为当下解决水下垃圾问题的关键技术之 一。该技术能够携带高精度的传感器与摄像头,深入水 下进行长时间的监测和数据收集,为水下垃圾的检测和 定位提供了技术支持[1-2]。然而,由于水下环境较为复 杂,且光条件较差,现有的水下机器人拍摄的垃圾图像 存在图像模糊、色彩有偏差等问题,导致对水下垃圾检 测的精度较差^[3-4]。因此,还需要进一步探究具有高精

收稿日期:2025-01-10; 修回日期:2025-02-18。

作者简介:张雅清(1977-),女,硕士,副教授。

引用格式:张雅清, 熊 勇,朱 清.基于改进 YOLOv5s 的水下自主机器人垃圾检测技术研究[J]. 计算机测量与控制, 2025, 33 (6):76-85.

注进使得一一 检测方法,旨在解决

度的垃圾检测方法。近年来,深度学习的不断推进使得 水下机器人对垃圾的高效和高精度检测技术迎来了新的 机遇^[5]。

深度学习因其具有较强的学习能力和分层特征学习 模式近年来广泛应用于计算机视觉、图像识别和目标检 测等领域[6]。文献[7]针对水环境污染问题提出一种基 于改进 YOLOv4 (You Only Look Once Version 4) 的垃圾 检测方法,旨在实现高速和高精度的目标检测。研究通 讨将 YOLOv4 转化为四尺度检测方法对目标进行检测。 结果表明,研究所提方法的检测速度高达 66.67 f/s,平 均精度高达 95.099%。文献「8]提出一种通道空间深 度感知的轻量化水下目标检测方法,旨在解决水下场景 复杂导致的误检和漏检问题。该方法采用双注意门控策 略对特征的敏感度进行增强。结果表明,该目标检测方 法的精度提高了 2.4%。文献「9〕针对海洋环境的复 杂性和水下探测器拍摄图像的低分辨率问题,提出一种 基于掩模区域的卷积神经网络的改进算法。结果表明, 研究所提方法在垃圾检测和实例分割两个任务上的精度 分别提高了 9.6% 和 5.0%。文献 [10] 提出了一种改 进 YOLOv8 模型的轻量化水下目标检测算法。研究通 过引入一种自适应注意力下采样模块和大可分离核注意 力机制,以此提升模型的精度。结果表明,该模型检测 精度提升了 1.4%, 计算复杂度降低了 15.9%。文献 [11] 提出一种改良的 YOLOv5 算法,旨在解决水下垃 圾检测设备储存与计算资源受限的问题。该方法采用卷 积块注意模块提高特征的提取能力。结果表明,该方法 的检测准确率可达到 97.5%。文献 [12] 提出一种水 下垃圾检测和实时分割方法,研究主要利用基于实时实 例分割模型和掩模区域的卷积神经网络对水下垃圾进行 检测和分割。结果表明,分割和检测模型的平均精度分 别为 0.377 和 0.365, 该方法具有较快的检测速度。文 献[13]针对水下目标检测效率低下的问题,提出一种 改进"You Only Look Once version 7"(YOLOv7)的水下 小目标检测技术。研究通过 YOLOv7 网络融合 SENet 注意力机制等策略提高检测模型的精度。结果表明,研 究所提模型能够有效提高水下目标的检测精度。文献 [14] 针对水下目标检测存在漏检和尺度不一的问题, 提出一种基于改进 YOLOv8 的水下目标检测算法。研 究通过空洞卷积空间金字塔模块,旨在提高模型对水下 多尺度目标的感知能力。结果指出,改进后的目标检测 模型精度提高了 3.4%。文献「15]提出一种用于现实 生活垃圾检测的 Skip-YOLO 模型,旨在快速、智能地 识别多种类型的生活垃圾。该模型通过大尺度卷积核扩 大模型的接受域,以此增强图像的浅层信息。结果表明, 研究所提模型的总体检测准确率提高了 22.5%。文献 「16〕提出一种融合 RetinaNet 和注意力机制的水下目标

检测方法,旨在解决目前水下机器人目标检测算法存在 目标识别率低和图像退化的问题。结果表明,研究所提 方法的检测精度高达 81.9%,检测速度为 16.8 f/s,具 有较优的检测性能。

综上,众多学者针对水下机器人对水下垃圾的检测 和处理进行了相关的研究,且取得较好的效果。然而, 现有的水下机器人在检测水下垃圾时,存在成像效果 差、亮度低和雾化等问题,导致水下垃圾检测的效率和 准确性降低。针对这些问题,研究提出一种基于深度学 习和图像融合的水下机器人图像增强技术,旨在提高水 下垃圾图像的清晰度和质量。为了进一步提高水下机器 人对水下垃圾的检测精度,研究提出一种基于改进的 YOLOv5s的水下自主机器人垃圾检测技术。此次研究 的创新性在于:1)提出一种结合 MobileNet V3 与多光 谱通道注意 (MSCA, multi-spectral channel attention) 模 块的主干结构,并引入特特征金字塔网络 (FPN, feature pyramid network)、通道注意力机制(CAM, channel attention mechanism)及空间金字塔池化(SPP, spatial pyramid pooling)网络,以增强模型对不同尺度目标的特征 融合能力,提升检测精度;2)引入了 SloU (SCYLLA Intersection over Union) 损失函数, 通过全面考虑预测框 与真实框之间的差异,有效提高目标检测模型的精度。

1 YOLOv5s 网络

YOLOv5s 网络是 YOLO 系列的最新版本之一,其 被广泛应用于实时目标检测任务。该网络通过在单一网 络中同时进行目标识别和边界框回归,从而实现快速且 准确的目标检测。YOLOv5s 网络由 3 个主要部分组成: 骨干网络、颈部网络和检测头部。骨干网络负责提取图 像特征,颈部网络对特征进行进一步处理和融合,检测 头部则负责最终的目标分类和边界框回归^[17-18]。

YOLOv5s的骨干网络主要任是提取目标图像的特征。该网络作为一种高效的卷积神经网络结构,通过通 道拆分和深度卷积以减少模型的计算量,同时保持特征 提取的有效性,并增强网络的学习能力。颈部网络位于 主干网路和检测头部之间,其主要负责对特征图进行多 尺度特征融合,并将这些特征传递给预测层。 YOLOv5s的颈部网络通过路径聚合网络(PANet,path aggregation network)来实现特征图的融合,以此增强 模型对不同尺度目标图像的检测能力。检测头部主要负 责最红的目标分类与边界框回归,其包括上采样、链接 与Conv块,其主要用于对特征图的进一步处理,并输 出最终的目标检测结果。

2 适用于水下条件的 YOLOv5s 模型改进

为了适应复杂的水下检测环境,研究对 YOLOv5s 模型进行了针对性的改进,主要从主干网络、颈部网络 和损失函数等方面进行优化,以提升模型对水下垃圾的 检测性能。改进后 YOLOv5s 模型的结构如图 1 所示。

2.1 MobileNet V3-MSCA 主干模块

水下光线穿透性差,导致图像采集时色彩失真和对 比度降低,为了解决这一问题,研究设计了一种结合 MobileNet V3 与 MSCA 模块的主干结构,以更好地适 应复杂的水下环境。MobileNet V3 网络作为一种创新 的轻量化深度学习构架,其在保持高性能的同时,能够 显著降低模型的计算复杂度。MobileNet V3 网络通过 引入一种非线性设计函数,以显著提升模型的非线性处 理能力。这种设计不仅可以提供更加稳定的梯度信息, 还能够加快网络的训练速度,同时保持模型的高准确度 监测能力。

该函数的表达式如式(1)所示:

$$\begin{cases} h - swish(x) = x \frac{ReLU6(x+3)}{6} \\ ReLU = \max(0, x) \end{cases}$$
(1)

式中, h - swish(x) 表示非线性激活函数; ReLU6 表示 激活函数。其中 h - swish(x) 函数不会存在梯度消失的 问题,且具有加快的计算速度和收敛速度。该函数在输

入为负值时输出为 0, 而在输入为正值时, 其导数不为 0。这种设计使得激活函数在正区间内保持非零梯度, 避免了传统 *ReLU* 函数在输入为负时梯度为零的问题。 此外, *h* - *swish*(*x*) 函数的导数在输入为正时逐渐趋于 稳定,进一步减少了梯度消失的可能性。该函数的计算 复杂度较低,使得网络在训练和推理过程中能够更快地 完成前向和反向传播,从而显著提高了计算效率。有研 究表明,使用 *h* - *swish*(*x*) 函数的网络在训练初期能够 更快地收敛,这是由于其非零梯度特性使得网络在反向 传播时能够更有效地更新权重,从而加速了模型的收敛 速度^[19]。

主干网络的设计通过非线性设计函数的引入,使得 MobileNet V3 网络能够在保持轻量化的同时,更好地 对复杂的水下图像特征进行捕捉和处理,这对于提高模 型在水下环境中的性能至关重要。该网络采用深度可分 离卷积技术,先对输入数据进行通道分离,然后利用逐 通道进行处理。其倒残差模块通过对网络的深度进行增 加,并使用非线性操作,以保持计算成本不变。最后, 将两个卷积模块进行连接,以此实现网络信息之间的跨 层传递,降低模型的计算量,提升效率。



图 1 改进后 YOLOv5s 模型的结构示意图

为了进一步增强模型的并行处理能力,研究在 MobileNet V3 网络中引入了 MSCA 模块。MSCA 模块通 过考虑水下环境中光的散射和吸收特性,对不同通道的 特征图进行加权,以此提高模型对水下图像中目标特征 的识别能力。这种处理不仅可以增强模型对关键特征的 关注,还能够提升模型对不同尺度目标的泛化能力。 MSCA模块的引入,可以在光线条件变换和图像质量 首先的情况下对水下图像进行高效处理。该模块的优势 在于并行处理机制。在分组卷积核通道重组之前,该模 块的通道分割技术将输入的特征分解为两个独立的部分 进行处理。这种设计允许模型同时处理不同通道的特 征,从而显著提升模型的检测速率。此外, MSCA 模 块通过融合全局平均池化和额外的频率信息,以增强特 征图的表示能力,从而实现对目标特征的高效率识别和 学习。MSCA 模块的引入,使得模型能够更加关注于 图像中的关键信息,尤其是在水下环境中由于光线条件 变化而难以捕捉的目标特征。这种设计旨在提高模型对 关键特征的识别度、并提高模型对不同尺度目标的泛化 能力。为了进一步融合全面的频率信息,研究采用二维 离散余弦比变换的方法,以有效地捕捉图像中的频率成 分,为模型提供更加丰富的特征表示。其表达式如式 (2) 所示:

$$\mathbf{F}_{\text{req}} = cat \left[(F_{\text{req}}^{0}, F_{\text{req}}^{1}, \cdots, F_{\text{req}}^{n-1}) \right]$$

$$Ms_{\text{att}} = \text{sigmoid} \left[fc(F_{\text{req}}) \right]$$
(2)

式中, F_{req} 为多光谱向量, $M_{s_{att}}$ 为多光谱通道注意力。 MobileNet V3-MSCA 整体结构如图 2 所示。

2.2 FCS 颈部模块

在对 YOLOv5s 模型的颈部网络进行改进的过程 中,研究设计了一种融合 FPN、CAM 与 SPP 的 FCS 模 块,旨在增强模型对水下图像中多尺度目标的特征融合 能力,提高检测精度。该方法通过自顶向下的路径和横 向连接,有效地对来自不同层次的特征信息进行融合, 以此保持特征的空间分辨率,从而对目标进行更准确的 定位和分类^[20-21]。FPN 的核心优势在于其能够从不同 层次的特征图中提取丰富的语义信息,并将其有效地融 合,显著增强了模型对不同尺度目标的检测能力。在水 下环境中,由于光线条件的复杂性,FPN 能够帮助模 型更好地捕捉到不同深度和光照条件下的目标特征表达 式如式(3) 所示:

$$\begin{cases} F_{l}^{\text{top}} = Conv_{1\times 1}(F_{l}) \\ F_{l}^{\text{fuse}} = \text{Upsample}(F_{l}^{\text{top}}) + F_{l-1} \end{cases}$$
(3)

式中, *F*_{*i*} 为输入的特征图, 1×1 为卷积压缩通道数, *F*_{*i*}^{op} 表示压缩后的特征图, *F*_{*i*}^{isse} 为融合后的特征图。在 FPN 的基础上, 引入了 CAM, 以增强模型对关键特征的关 注度。CAM 通过全局平均池化和全局最大池化, 对通 道间的依赖关系进行捕捉, 以强化特征图中有用的信 息, 并抑制无关的特征, 从而提高模型的检测精度。为 了进一步提升模型对不同尺度目标的适应性, 研究在颈 部网络中加入了 SPP 网络。SPP 通过在不同尺度上池 化特征图, 生成了具有多尺度不变性的特征表示, 使得 模型能够更好地处理水下环境中尺度变化较大的目标。 该方法的表达式如式(4)所示:



图 2 MobileNet V3-MSCA 整体结构图

投稿网址:www.jsjclykz.com

$$F_{spp} = SPP(F) =$$

$$ax\{max[max(F, k = 1), k = 3], k = 5\}$$
(4)

式中, F_{spp} 为经过 SPP 处理后的特征图, k 为不同尺寸的核。输入处理后特征图后, 研究通过多层感知机 (MLP, ultilayer perceptron)和 Sigmoid 函数生成最终的通道注意力图^[22]。FCS 颈部模块如图 3 所示。

2.3 模型对象分类及边界框回归方法

真实水下环境中的垃圾目标大小不一, 因此,为了有效处理不同尺寸的垃圾目标, 并提高模型的检测精度,研究在对象分类 及边界框回归方法上进行了优化。对于对 象的分类,研究通过将全连接层(FC,fully connected layers)与全局平均池化层 (GAP,global average pooling)进行结合, 以实现水下图像中目标特征的有效提取和 分类。GAP将每个特征图的高和宽维度上 的特征平均,输出固定长度的特征向量, 使得网络对输入图像的尺寸具有不变性。 FC 通过将全局平均池化后的特征向量映射

到类别概率分布上,以此对目标进行自动化识别。 GAP与FC的表达式如式(5)所示:

$$\begin{cases} \mathbf{P} = \frac{1}{H \times W} \sum_{i=1}^{H} \sum_{j=1}^{W} F(i,j) \\ C = FC(\mathbf{P}) \end{cases}$$
(5)

式中, F(i,j)为特征图在位置(i,j)的值; P为全局平 均池化后的特征向量; H和W分别为特征图的高与宽。 为了准确预测目标的位置和尺寸,研究采用边界框回归 方法,通过几个卷积层后接一个线性层来预测调整锚框 的偏移量,以覆盖不同尺寸的目标。在水下环境中,由 于目标可能会因为水流的作用而产生形变,因此,研究 通过结合对象分类和边界框回归,以此同时预测图像中 的目标类别和位置,实现快速准确的目标检测。

2.4 基于 SIoU 损失函数优化的非极大值抑制

在目标检测过程中,同一个目标可能会被多个锚框 检测到,导致产生多个重叠的边界框。为了解决这一问题,研究采用了非极大值抑制(NMS,non-maximum suppression)技术,以优化检测结果并减少冗余。改进 的YOLOv5s模型在进行对象分类及边界框回归后,采 用了基于 SIoU损失函数优化的非极大值抑制策略,以 进一步提升检测的准确性和效率^[23]。对于每个类别, 按照置信度对所有边界框进行排序,然后选取置信度最 高的边界框,抑制与其 IoU高于特定阈值的所有其他 边界框。此次研究考虑到水下环境的特殊性,对传统的 NMS进行了改进,引入了基于 SIoU 的 NMS 策略。



图 3 FCS 颈部模块

SIoU损失函数不仅考虑了边界框的位置和尺寸,还考虑了预测框与真实框之间的形状差异,使得 NMS 过程更加精细和适应水下环境。目标检测模型的损失函数由3个部分组成:置信度损失、分类损失和定位损失。其中定位损失中的边界框误差占据了较大的比例^[24]。最终的预测框去顶是基于候选框中 IoU 值进行计算。 SIoU损失函数的定义如式(6)所示:

$$L_{SloU} = 1 - IoU + \frac{\Delta + \Omega}{2} \tag{6}$$

式中, L_{SloU} 为 SloU 损失函数; Δ 为距离损失; Ω 为形状 损失。SloU 损失函数主要由角度损失 (Angle Cost)、 距离损失 (Distance Cost)、形状损失 (Shape Cost) 和 IoU损失 (IoU Cost) 四部分组成。其中,角度损失、 距离损失、形状损失与 IoU损失的表达式如式 (7) 所示:

$$\begin{cases} Ac = |\theta| \\ Dc = \sqrt{(x_{p} - x_{g})^{2} + (y_{p} - y_{g})^{2}} \\ Sc = \frac{(w_{p} - w_{g})^{2}}{w_{g}^{2}} + \frac{(h_{p} - h_{g})^{2}}{h_{g}^{2}} \\ IoUc = 1 - \frac{A_{i}}{A_{p} + A_{g} - A_{i}} \end{cases}$$
(7)

式中, Ac 为角度损失; Dc 为距离损失; θ 为真实框与预测框之间的夹角; (x_{p}, y_{p}) 和 (x_{g}, y_{g}) 分别表示预测框 与真实框的中心点坐标。Sc 为形状损失; IoUc 为 IoU 损 失; w_{p} 和 h_{p} 分别为预测框的宽与高; w_{g} 和 h_{g} 分别为真 实框的宽与高; A_{p} 和 A_{g} 分别为预测框与真实框的面积;

m





图 4 SIoU 损失函数示意图

2.5 水下图像处理方法

对于水下图像的预处理,研究利用基于高斯滤波技 术的双边滤波器,并考虑图像像素的相似性对水下垃圾 图像进行高效处理,以此对图像进行降噪,改善图像的 色彩偏差。为了增强水下图像的对比度,研究通过对输 入图像的多尺度权重图进行提取,获取水下图像的最优 特征。该方法通过对局部对比度权重和显著性进行计 算,以提高水下图像的局部区域对比度,并突出图像的 目标区域的显著性特征。为了对水下图像的纹理和细节 信息进行高效提取,研究采用多尺度融合的方法将全 局、局部、显著性与曝光权重进行归一化处理。对多尺 度图像进行融合后,研究通过对背景光和透射率进行估 计和优化,以实现对水下的图像的复原。对于背景光的 估计,研究采用二值化方法对亮通道图进行处理,获取 高亮度区域和背景光平坦区域[25]。为了保证水下图像 的细节与边缘信息,研究引入了导向滤波,对水下的透 射率进行计算,并采用逆向求解的方法对水下图像进行 复原,如式(8)所示:

式中, $t_{G}(x)$ 为透射率, ω 为调节参数,G为导向滤波, A^{ϵ} 为估计的背景光值,J(x)为复原图像的输出结果, $t_{G}(x)$ 为优化后所获得的透射率, t_{o} 为门限值,通常取 0.1。研究采用图像色彩与质量两种评价指标对水下图 像的清晰度与色彩度进行评价。两种指标的计算如公式 (9)所示:

 $\left\{egin{aligned} UIQM &= c_1 imes UICM + c_2 imes UISM + c_3 imes UIConM \ UCIQE &= d_1 imes Chr + d_2 imes Sat + d_3 imes Con \end{aligned}
ight.$

式中, UISM 和 UICM 分别表示图像的清晰度与与色彩度; Chr 表示色度标准差; Con 表示亮度对比度; Sat 表示饱和度均值; UIConM 表示图像的对比度。其中 c_1, c_1, c_3 分别设置为 0.280、0.295、3.572, d_1, d_2, d_3 分别设置为 0.465、0.273、0.258。

3 实验方法与结论

3.1 改进 YOLOv5s 的模型训练方法

模型训练采用了分阶段训练策略。研究使用的数据 集包括在简单和复杂水下环境中采集的400张水下垃圾 图像。为了确保模型的泛化能力和训练效果,数据集被 划分为训练集和验证集,比例为7:3。其中,70%的 数据(280张图像)用于模型训练,剩余30%的数据 (120张图像)用于验证模型性能。

为了加速模型的收敛并提高训练效率,研究使用预 训练的 MobileNet V3 权重作为骨干网络的初始化参数。 预训练权重在大规模图像数据集上训练得到,能够为模 型提供丰富的特征提取能力。模型训练采用随机梯度下 降 (SGD, stochastic gradient descent) 优化器,动量设 置为 0.9, 以加速梯度下降过程并避免局部最优。初始 学习率设置为 0.001,并在训练过程中采用学习率衰减 策略。学习率在每50个迭代周期后衰减为原来的0.1 倍,以确保模型在训练后期能够稳定收敛。为了优化模 型性能,研究对多个超参数进行了细致调整,包括学习 率、批大小、迭代次数等。通过交叉验证和网格搜索方 法,确定了最优的超参数组合。最终,批大小设置为 16, 迭代次数设置为 200 次, 以确保模型在训练过程中 既能快速收敛,又能达到较高的检测精度。在模型训练 过程中,通过监控验证集上的损失函数值和准确率,实 时调整模型参数。如果验证集上的损失函数值连续多个 迭代周期不再下降,则提前终止训练,以避免过拟合。 同时,根据验证集的性能表现,动态调整学习率等超参 数,以进一步优化模型性能。模型训练采用 PyTorch 1.7.1 深度学习框架,数据处理使用 Python 3.6 编程语 言。硬件环境为 NVIDIA GeForce 930M GPU。

3.2 实验结果与分析

研究在两种不同的水下环境中进行了实验:简单场 景和复杂场景。简单场景指的是水下环境较为清晰,光 线条件较好,垃圾分布较为单一的情况;而复杂场景则 是指水下环境光线条件差,垃圾种类繁多,背景杂乱的 情况。研究选取了某海域最为实际实验场地,分别在简 单环境和复杂环境分别拍摄 200 张水下垃圾图像作为数 据集,以此验证研究所提图像处理技术的实际应用效 果。为了验证研究所提方法的有效性,研究对图像处理 方法进行了定量分析。在实验中,数据集被划分为训练

(9)

集和验证集,旨在确保模型在训练过程中能够学习到足 够的特征,同时通过验证集评估模型的泛化能力,训练 集和验证集的占比为7:3。

此次研究将水下垃圾定义为对海洋生态环境造成潜 在危害的人造废弃物。主要包括塑料制品(如塑料瓶、 塑料袋、塑料碎片)、金属制品(如易拉罐、金属罐)、 渔网、玻璃制品以及其他非自然材料的废弃物。简单环 境与复杂环境的区分如下:简单环境指水下光线条件较 好、背景单一且垃圾分布较为稀疏的场景。在这种环境 下,图像中的垃圾目标清晰可见,背景干扰较少。例如, 浅水区域的清澈水域,水下垃圾主要为塑料瓶、塑料袋 等大型漂浮物。复杂环境指下光线昏暗、背景杂乱且垃 圾种类繁多的场景。在这种环境下,图像中可能存在多 种类型的垃圾(如塑料碎片、渔网、金属罐等),且垃圾 目标可能被遮挡或与背景融合,增加了检测难度。

研究通过所提算法对实际海域的简单水环境和复杂 水环境下所拍摄的图像进行了图像处理,分析了 RGB 通道直方图,以此探究研究所提体水下图像处理方法的 效果,结果如图5所示。图5(a)中表明,原水下图像 的 RGB 通道颜色分布集中,当灰度级为 40 时, R 通道 像素值达到峰值,为2458。G通道和B通道在灰度级 分别为 82 和 156 时达到峰值,像素值分别为 1 624 和 1805。图5(b)中显示,在简单水环境下,灰度级为 100 时,研究所提算法处理后的图像 RGB 通道的像素 点分别为 424、287 和 658。当灰度级增加到 200 时,图 像 RGB 通道的像素点分别降低为 238、112 和 98。结果 可以看出 RGB 通道的像素值均匀分布,其像素值分布 在 0~658 范围内。图 5 (c) 中可以看出,复杂水环境 下, 灰度级为 100 时, 研究所提算法处理后的图像 RGB 通道的像素点分别为 378、413 和 925, 其通道的 像素点均在1000以内。结果表明,研究所提方法能够 对实际海域的简单水环境和复杂水环境下图像的色彩偏 差进行有效处理。

研究通过所提图像处理算法对实际海域的简单环境 与复杂环境图像处理的熵值、图像色彩度和图像清晰度 的结果进行了分析,以此探究该算法在实际水下环境中 的图像处理效果,结果如图6所示。图6(a)中,在简 单水环境和复杂水环境下,昏暗图像的熵值分别为 7.80和7.75。简单水环境下,色偏图像与浑浊图像的 熵值分别为7.82和7.93,复杂环境下,两种图像的熵 值分别为7.71和7.89。结果表明,简单环境与复杂环 境下的3种图像的熵值均在7.7以上,且两种环境下的 图像处理误差较小,均在0.15内。图6(b)中显示, 简单水环境与复杂水环境的图像色彩度分别为2.79和 2.65,图像清晰度分别为0.61和0.59。结果可以看出,



图 6 不同水环境下的图像处理熵值和评价指标结果

简单水环境与复杂水环境之间的 UIQM 相差 0.14,图 像清晰度相差 0.02。结果表明,两种水环境下的图像 色彩度和图像清晰度均较高,证明研究所提水下图像处 理方法在实际水环境中能够较好地均衡图像的清晰度、 饱和度和对比度,图像质量较高。

为了探究水下图像处理对垃圾检测精度的影响, 研究对实际谁环境下下处理前后的水下图像的准确率 和召回率进行了分析,结果如图7所示。图7(a)中 表明,当迭代次数为30时,实际水环境下原始图像的 准确率为85.4%,采用研究所提算法处理后的图像准 确率为 87.25。当迭代次数达到 120 时, 原始图像和 处理后的图像准确率分别为 92.4%和 88.7%。当迭代 次数达到 200 时,实际海域环境下原始图像和处理后 的图像准确率分别为94.6%和92.8%。结果表明,使 用研究所提算法对图像进行处理后具有较高的准确率。 图 7 (b) 中显示, 迭代次数为 30 时, 原始图像与处 理后的图像的召回率分别为 65.0%和 72.1%。当迭代 次数为120时,原始图像的召回率为83.6%,处理后 图像的召回率为 90.8%。当迭代次数达到 200 时,原 始图像与处理后的图像的召回率分别为 84.2% 和 92.2%。结果表明,研究所提算法对实际海域水环境 下的图像处理后具有较高的召回率,证明了该算法的 有效性。



图 7 水下图像处理前后的检测结果

通过对水下图像处理效果的详细分析,研究验证了 所提算法在改善推向质量和提高检测精度方面的有效 性。实验进一步采用处理后的水下图像对改进的 YOLOv5s模型的性能进行验证。实验中,模型 70%的 数据作为训练集,剩余 30%的数据为验证集。模型训练时,其迭代次数设置为 200,图像尺寸为 320×320。 研究采用的深度学习框架为 Pytorch 1.7.1,数据处理 软件为 Python 3.6, GPU 为 NVIDIA GeForce 930 M。

研究对不同水环境下基于改进 YOLOv5s 的垃圾检 测模型的精度和损失函数进行了分析,以此验证该模型 的可行性,结果如图8所示。图8(a)中表明,模型迭 代次数为 30 时,简单水环境和复杂水环境的 YOLOv5s 模型的平均精度分别为 0.418 和 0.375。当迭代次数为 90 时, 简单环境下 YOLOv5s 模型的平均精度为 0.857, 复杂环境下 YOLOv5s 模型的平均精度为 0.825。当迭 代次数达到 200 时,两种水环下 YOLOv5s 模型的平均 精度分别为 0.952 和 0.948。结果显示,研究所提垃圾 检测模型在简单水环境与复杂水环境下具有较优的检测 精度。图 8(b) 中,模型迭代次数为 30 时,简单环境 下 YOLOv5s 模型的损失函数为 0.019 2, 复杂水环境 下 YOLOv5s 模型的损失函数为 0.019 5。当迭代次数 为90时,两种水环境下 YOLOv5s 模型的损失函数分 别为 0.0020 和 0.001 8。当迭代次数达到 200 时, 简单 水环境下 YOLOv5s 模型的损失函数为 0.001 2, 复杂 水环境下 YOLOv5s 模型的损失函数为 0.000 8。结果 表明,研究所提水下垃圾检测模型在两种不同的水环境 中具有较优的性能。



图 8 YOLOv5s 垃圾检测模型载不同水环境下的训练结果

为了进一步验证改进后的 YOLOv5s 垃圾检测模型 的优越性,研究将该模型与当前流行的目标检测模型的 检测精度与每秒传输帧数(FPS, frames per second)进 行了分析,对比模型包括快速区域卷积神经网络(Faster R-CNN, faster region-based convolutional neural network) 模型、单次多框检测器 (SSD, single shot multibox detector) 模型与 YOLOv5s-ShuffleNet 模型, 结果如 图 9 所示。图 9 (a) 中表明, Faster R-CNN 与 SSD 模型 的平均检测精度分别为 0.978 和 0.957, YOLOv5s-ShuffleNet模型与研究所提模型的平均检测精度分别为 0.924 和 0.989。与其他 3 种模型相比,研究所提模型的精度 分别提高了 1.11%、3.24% 和 6.57%。图 9 (b) 中显 示, Faster R-CNN、SSD 与 YOLOv5s-ShuffleNet 模型 的平均 FPS 分别为 14、26 和 32, 研究所提模型的平均 FPS为34。结果表明,研究所提改进的YOLOv5s垃圾 检测模型具有较优的检测精度,且计算效率较高。



图 9 不同检测模型的检测性能结果

为了验证研究所提方法的可行性,研究在该海域的 实际水下环境中对垃圾进行了检测,检测效果如图 10 所示。图 10 (a)中表明,研究所提改进后的 YOLOv5s 模型对塑料袋的检测精度高达 0.95,对塑料瓶的检测 精度为 0.90。图 10 (b)中,该模型对水下易拉罐的检 测精度高达 0.96。结果表明改进后 YOLOv5s 模型能够 更加准确地对水下垃圾进行检测,且该方法比改进前 YOLOv5s 模型的检测效果更好,证明了研究所提方法 的可靠性。





(a)改进后YOLOv5s模型对 简单环境下的垃圾检测效果

(b)改进后Y0L0v5s模型对 复杂环境下的垃圾检测效果

图 10 改进后 YOLOv5s 模型对不同水环境下的 垃圾检测效果

3.3 消融实验

为来进一步验证改进后 YOLOv5s 垃圾检测模型的 性能,研究进行了消融实验。其中原始的 YOLOv5s 模 型为基线模型,改进的组件包括轻量化网络 MobileNet V3、MSCA、颈部网络。消融实验结果如表1所示。表 1中显示,研究所提模型具有较优的性能。

实验 编号	组件变更	平均精 度 AP/ %	损失函 数值	FPS /%
1	基线模型(YOLOv5s)	91.2	0.001 9	27
2	基线模型+MobileNet V3	88.5	0.001 5	30
3	基线模型+MSCA	90.2	0.001 6	28
4	基线模型+颈部网络	89.7	0.001 4	26
5	基线模型+MobileNet V3+ MSCA	91.5	0.001 3	32
6	基线模型+MobileNet V3+ 颈部网络	91.4	0.001 2	31
7	基线模型+MSCA+颈部网络	92.2	0.001 1	29
8	改进后的 YOLOv5s 模型	94.8	0.000 8	34

表1 YOLOv5s 垃圾检测模型的消融实验结果

4 结束语

研究针对水下垃圾污染问题,提出一种基于改进 YOLOv5s的水下自主机器人垃圾检测方法。该方法采 用 MobileNet V3 作为骨干网络,引入 MSCA 模块、 FPN和 CAM,优化了 YOLOv5s 模型的特征提取和融 合能力,并通过 *SIoU*损失函数,进一步提升了模型的 检测精度。实验结果表明,改进后的 YOLOv5s 模型平 均精度分别达到 0.952 和 0.948,损失函数值分别为 0.001 2 和 0.000 8,证明了所提方法的有效性。与Faster R-CNN、SSD 与 YOLOv5s-ShuffleNet 模型相比,研 究所提模型的平均 *FPS* 分别提高了 58.82%、23.53% 和 5.88%。结果表明,研究所提垃圾检测模型在实际 水环境中具有较优的性能和检测效果。 此次研究的局限性在于仅对环境条件较差的水下 环境构建的数据集进行了分析,数据集的多样性和规 模有限,导致模型的泛化能力提升有限。因此,在未 来的研究中,可以构建多数据量和多种水下场景的护 士。进一步提升模型的泛化能力。同时,可以探索融 合更多先进的深度学习技术和多模态信息,以提高水 下垃圾检测的准确性和鲁棒性。研究所提模型能够为 水下环境保护和资源管理提供技术支持,具有重要的 应用前景和价值。

参考文献:

- [1] ZHANG Y, HUANG Z, CHEN C, et al. A spiral-propulsion amphibious intelligent robot for land garbage cleaning and sea garbage cleaning [J]. Journal of Marine Science and Engineering, 2023, 11 (8): 1482 - 1494.
- [2] WEI L, KONG S, WU Y, et al. Image semantic segmentation of underwater garbage with modified U-Net architecture model [J]. Sensors, 2022, 22 (17): 6546-6559.
- [3] CHRISTENSEN L, DE GEA FERNANDEZ J, HILDE-BRANDT M, et al. Recent advances in AI for navigation and control of underwater robots [J]. Current Robotics Reports, 2022, 3 (4): 165-175.
- [4]张 天,温显斌,薛彦兵,等.基于频域注意力的水下目标 检测算法研究 [J].光电子·激光,2024,35 (6):604 -611.
- [5] REN Z, ZHANG H, LI Z. Improved YOLOv5 network for real-time object detection in vehicle-mounted camera capture scenarios [J]. Sensors, 2023, 23 (10): 4589 - 4604.
- [6] 李海波, 付家兴, 颛俊尧, 等. 配电线路表面外破缺陷小目 标双通道 YOLO 识别算法 [J]. 环境技术, 2024, 42 (7): 111-117.
- [7] TIAN M, LI X, KONG S, et al. A modified YOLOv4 detection method for a vision-based underwater garbage cleaning robot [J]. Frontiers of Information Technology & Electronic Engineering, 2022, 23 (8): 1217 - 1228.
- [8] 赵瑞金,李海涛,陆光豪.通道空间深度感知的轻量化水 下目标检测 [J].计算机测量与控制,2024,32 (9):86-93.
- [9] DENG H, ERGU D, LIU F, et al. Anembeddable algorithm for automatic garbage detection based on complex marine environment [J]. Sensors, 2021, 21 (19): 6391 - 6404.
- [10] 梁秀满,赵佳阳,于海峰.基于 YOLOv8 的轻量化水下目 标检测 算法 [J]. 红外技术,2024,46 (9):1015-1024.
- [11] WU C, SUN Y, WANG T, et al. Underwater trash detection algorithm based on improved YOLOv5s [J]. Journal of Real-Time Image Processing, 2022, 19 (5): 911-920.

- [12] CORRIGAN B C, TAY Z Y, KONOVESSIS D. Realtime instance segmentation for detection of underwater litter as a plastic source [J]. Journal of Marine Science and Engineering, 2023, 11 (8): 1532-1547.
- [13] 杜 锋. 基于改进 YOLOv7 的水下小目标检测算法研究 [J]. 计算机测量与控制, 2024, 32 (9): 108-117.
- [14] 周 昕, 李远禄, 吴明轩, 等. 改进 YOLOv8 的水下目标 检测 [J]. 计算机系统应用, 2024, 33 (11): 177-185.
- [15] LUN Z, PAN Y, WANG S, et al. Skip-YOLO: Domestic garbage detection usingdeep learning method in complex multi-scenes [J]. International Journal of Computational Intelligence Systems, 2023, 16 (1): 139-154.
- [16] 黄珍伟,陈 伟,王文杰,等. 基于改进 RetinaNet 网络的水下机器人目标检测与实验 [J]. 计算机工程与科学, 2024,46 (2): 264-271.
- [17] GAO R, LI Q, LI S. Multi-scale behavior recognition method for dairy cows based on improved YOLOV5s network [J]. Transactions of the Chinese Society of Agricultural Engineering, 2022, 38 (12): 163-172.
- [18] AMRANI M, BEY A, AMAMRA A. New SAR target recognition based on YOLO and very deep multi-canonical correlation analysis [J]. International Journal of Remote Sensing, 2022, 43 (15/16): 5800 - 5819.
- [19] JIANG L. Applied mathematics and nonlinear sciences
 [J]. Applied Mathematics and Nonlinear Sciences, 2024,
 9 (1): 1-18.
- [20] QIAOLI Z, LI M, LI Y C, et al. Identification of tomato leaf diseases based on improved lightweight convolutional neural networks MobileNetV3 [J]. Smart Agriculture, 2022, 4 (1): 47-56.
- [21] 朱正林, 钱予阳, 马辰宇, 等. 基于改进 YOLOv5s 的跌 倒行为检测 [J]. 计算机测量与控制, 2024, 32 (10): 26-31.
- [22] SIRISHA U, PRAVEEN S P, SRINIVASU P N, et al. Statistical analysis of design aspects of various YOLObased deep learning models for object detection [J]. International Journal of Computational Intelligence Systems, 2023, 16 (1): 126-154.
- [23] WANG Q, LIU H, PENG W, et al. A vision-based approach for detecting occluded objects in construction sites [J]. Neural Computing and Applications, 2024, 36 (18): 10825 - 10837.
- [24] 张绍文,史卫亚,张世强,等. 基于加权感受野和跨层融合的遥感小目标检测 [J]. 电子测量技术,2023,46 (18):129-138.
- [25] BERA SK, GHOSH S, BHOWMIK S, et al. A non-parametric binarization method based on ensemble of clustering algorithms [J]. Multimedia Tools and Applications, 2021, 80 (5): 7653 - 7673.