

文章编号 1004-924X(2025)11-1803-15

面向自主水下机器人的深海污染物检测

张彪*, 朱振洋, 许家忠

(哈尔滨理工大学 黑龙江省复杂智能系统与集成重点实验室, 黑龙江 哈尔滨 150080)

摘要:海洋垃圾是全球性环境问题的主要组成部分,自主水下机器人可以在一定程度上清理深海垃圾。为了能够快速准确地检测深海垃圾,利用深度学习的方法建立了一个轻量级检测模型(Deep-sea debris yolo, Debris-yolo)。基于全球海洋学数据中心(GODAC)推出的 Deep-sea debris 数据库创建了深海垃圾数据集。引入改进的 BiFPN 特征融合网络,减少模型参数数量,提高模型区分背景的能力。然后,设计了轻量级检测头,降低了计算复杂度,提高了深海垃圾检测模型的实用性和可部署性。最后,提出 Wise-DIoU (Wise Distance Intersection over Union) 损失函数,抑制低质量的样本对模型的干扰,使模型能够更准确地定位深海垃圾。在训练过程中,采用数据增强和自适应色彩均衡水下复原对训练数据集进行增强。实验结果表明,与 YOLOv8n 相比,本文提出模型 Debris-yolo 的 $mAP_{0.5}$ 和 $mAP_{0.5:0.95}$ 分别提高了 1.3% 和 1.6%,参数数量和 GFLOPS 分别减少了 48.2% 和 36%。

关键词:计算机视觉;目标检测;水下机器人;深海垃圾

中图分类号:TP394.1;TH691.9 **文献标识码:**A

doi:10.37188/OPE.20253311.1803 **CSTR:**32169.14.OPE.20253311.1803

Deep-sea pollutant detection for autonomous underwater robots

ZHANG Biao*, ZHU Zhenyang, XU Jiazhong

(Heilongjiang Provincial Key Laboratory of Complex Intelligent System and Integration,
Harbin University of Science and Technology, Harbin 150080, China)

* Corresponding author, E-mail: zhangbiao@hrbust.edu.cn.

Abstract: Marine debris constitutes a significant global environmental challenge. Autonomous underwater robots offer a potential solution for the removal of deep-sea debris. To enable rapid and accurate detection of such debris, this study presents a lightweight detection model, termed Deep-sea Debris YOLO (Debris-YOLO), developed using deep learning techniques. An initial deep-sea debris dataset was constructed based on the Deep-sea Debris Database provided by the Global Oceanographic Data Centre (GODAC). Subsequently, an enhanced BiFPN feature fusion network was employed to reduce model parameters while enhancing background discrimination capabilities. A lightweight detection head was designed to decrease computational complexity, thereby improving the model's practicality and deployability for deep-sea debris detection. Furthermore, the Wise-DIoU (Wise Distance Intersection over Union) loss function was introduced to mitigate the influence of low-quality samples, enabling more precise localization of deep-sea debris. Data augmentation and adaptive color-balanced underwater restoration were applied to enrich the training dataset. Experimental results demonstrate that the proposed Debris-YOLO model achieves im-

收稿日期:2025-01-18;修订日期:2025-02-25.

基金项目:国家自然科学基金资助项目(No. 61806060)

provements of 1.3% and 1.6% in $mAP_{0.5}$ and $mAP_{0.5:0.95}$, respectively, compared to YOLOv8n, while reducing the number of parameters and GFLOPS by 48.2% and 36%, respectively.

Key words: computer vision; target detection; underwater robot; deep-sea debris

1 引 言

海洋垃圾的累积对海洋生态构成了相当大的威胁,已经成为了全球性的问题^[1-2]。在 20 世纪 80 年代初,随着海洋垃圾的日益剧增,人们开始研究这一问题的有效解决方法^[3]。尽管如此,近几十年流入海洋的垃圾仍不断增加^[4]。人类对塑料的严重依赖导致它在海洋中无处不在^[5],据估计,每年有 4.8 到 12.7 百万公吨的塑料从陆地流入海洋^[6],并且这个数量在 2025 年可能会增加一个数量级。因此,实现海洋生态系统可持续发展的关键在于解决海洋废弃物污染问题,这其中包括垃圾排放的管理和清除工作^[7]。事实上,深海是绝大多数海洋垃圾的最终归宿^[8]。它们一部分会漂浮在海洋表面,绝大多数已经进入了海里,甚至沉入海底^[9]。据统计,海平面和海岸上的垃圾分布密度为 50 个/ km^2 ,而深海中垃圾分布密度高达 6 633 个/ km^2 ^[10]。这些数据表明,相比于海洋表面上的垃圾,深海垃圾对海洋环境及生物有着更严重的威胁。

目前,主要采用人工操纵的水下机器人机器臂的方式清理深海垃圾,这对技术人员的要求很高,会消耗巨大的人力物力。随着深海探测技术的发展,水下机器人已应用到深海科学研究、海底资源勘探和海洋污染检测中^[11-12]。自主水下机器人可以凭借深海垃圾检测模型和机器臂的自主操作完成深海垃圾采集、打捞和清理工作。这种方式大大提高了深海垃圾的清理效率^[13]。

为了有效地清除深海垃圾,自主水下机器人在深海中检测模型的准确性和实时性至关重要。Ding 等^[14]以 YOLOv5n 为基础提出了一种轻量化深海生物检测方法,通过设计一种结合知识提取的迁移学习策略,降低了模型对数据量的依赖,并提高了模型的检测性能。光的本质是电荷产生的电场和磁场,光与水介质相互作用的极化机制会导致散射、反射、折射和吸收等现象。这些现象大大增加了机器人水下作业的难度。

Khan 等^[15]提出了一种互补损失函数,以优化水下作业中极化引起的误差,提升了水下目标检测的精确度。2017 年 3 月,日本海洋地球科学技术机构(JAMSTEC)的全球海洋学数据中心(GODAC)推出了深海废弃物数据库,该数据库包含 JAMSTEC 的潜水器“SHINKAI6500”、“HYPER-DOLPHIN”等在深海调查中发现的海洋废弃物的视频和照片。Fulton 等^[16]自制了一个海洋塑料垃圾数据集,并用它来训练 YOLOv2, Tiny-YOLO, Faster R-CNN 和 SSD 4 个模型,旨在证明深度学习方法检测海洋塑料垃圾的可行性。王普威^[17]使用来自不同样本的微塑料训练 Mask R-CNN 网络,并在模型中加入侵蚀和扩张操作。实验结果表明,深度卷积神经网络在微塑料自动分割和形状分类方面具有潜力。Xue 等^[18]构建了几乎涵盖深海垃圾全部类别的深海垃圾检测数据集,并采用包括 YOLOv3, Faster RCNN, Tiny-YOLO 和 SSD 4 种深度学习架构,结合 MobileNetv2, VGG16, ResNet50 3 种分类模型,分析了不同分类模型对检测性能和速度的影响。Rahul 等^[19]以 InceptionResNetv2 架构对图像中垃圾区域的检测和特征提取进行研究。以上研究均以提高深海垃圾检测模型的精度为目的,但模型实现实时检测的研究也极为重要^[20-22]。Huang 等人^[23]考虑深海环境中图像质量存在的噪声、模糊和失真等问题对模型检测性能的影响,通过引入一种混合损失函数,提出了一种轻量级的神经网络 DSDebrisNet。该轻量化模型能够有效支持自主水下机器人在海洋环境中的运行,为海洋保护提供可行的技术支持。

深度学习方法能够快速准确地检测深海垃圾,但深海垃圾目标检测仍存在精度较低、模型参数量大等问题。针对这些问题,本文基于 YOLOv8n 构架提出一种轻量化深海垃圾检测模型 Debris-yolo,确保精准和快速地检测深海垃圾的同时能够有效地区分深海生物与深海垃圾。

2 YOLOv8 模型

2.1 YOLOv8 结构

YOLOv8 是一种先进的实时目标检测算法,

根据不同应用场景的需求分为 n,s,m,l,x 五个版本。它们保持一致的整体架构,但在深度和宽度上有所不同^[24]。如图 1 所示,YOLOv8 模型主要包括骨干网络、颈部网络和检测头三个部分。

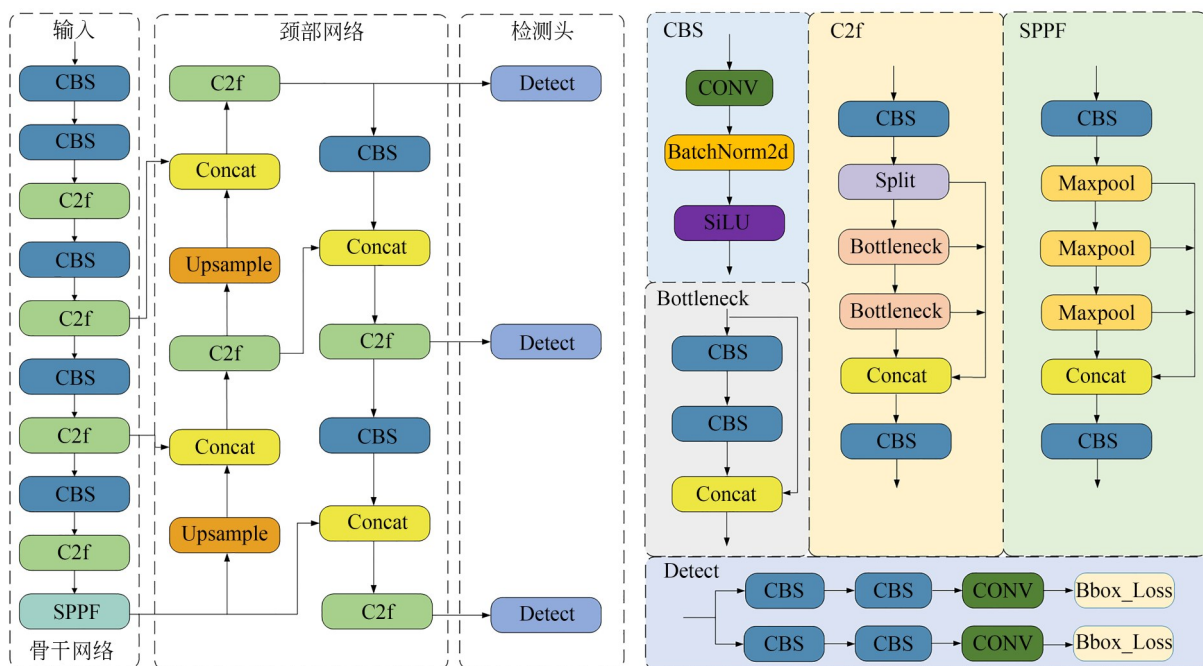


图 1 YOLOv8 模型

Fig. 1 YOLOv8 model

YOLOv8 的骨干网络由卷积层, C2f 和 SPPF 构成。C2f 模块借助分割梯度流的思想将特征图分为两部分, 一个分支直接传递给输出, 旨在保留图像的细节和低级特征; 另一分支经过多个 Bottleneck 模块的处理, 用于捕捉图像中的主要信息和特征。这样的分支设计实现了更丰富的梯度组合, 大大提高了卷积网络的学习能力和效率。Bottleneck 参考 ELAN 并行更多梯度流分支的思想, 在保持轻量化的同时获得更加丰富的梯度流信息。

YOLOv8 的颈部网络采用路径聚合网络和特征金字塔网络 (Path Aggregation Network with Feature Pyramid Network, FPN+PAN), 通过自顶向下的 FPN 增加了低层特征图的特征信息量, 获得强语义特征。PAN 通过自下而上的上采样增加上层特征图的位置信息量获取强定位特征, 有效地提高特征的提取和融合效率。

YOLOv8 使用不同分支处理分类任务和回

归任务。分类分支负责预测目标的类别, 回归分支负责预测目标的位置。这种结构虽然提高了模型的预测性能, 但却带来了巨大的参数量和计算量^[25]。

2.2 工作原理

YOLOv8 将目标检测任务描述为一个回归问题, 从图像像素直接预测目标的预测边界框和类别。将输入图像分为 $m \times n$ 的网格单元, m 和 n 为输入图像的行和列。如图 2 所示, 如果检测物体的中心位置位于某个网格单元, 则该单元负责检测物体, 并生成相应的位置得分和类别得分。

针对所有单元生成的类别得分和位置得分, 通过加权的方式得到最终的加权分数。选择分数最高的 Topk 作为正样本, 公式如下:

$$t = s^{\alpha} \times u^{\beta}, \quad (1)$$

其中: s 为预测类别得分; u 为预测框与真实框的 CIoU 值; α 和 β 为权重。两者相乘可以衡量预测

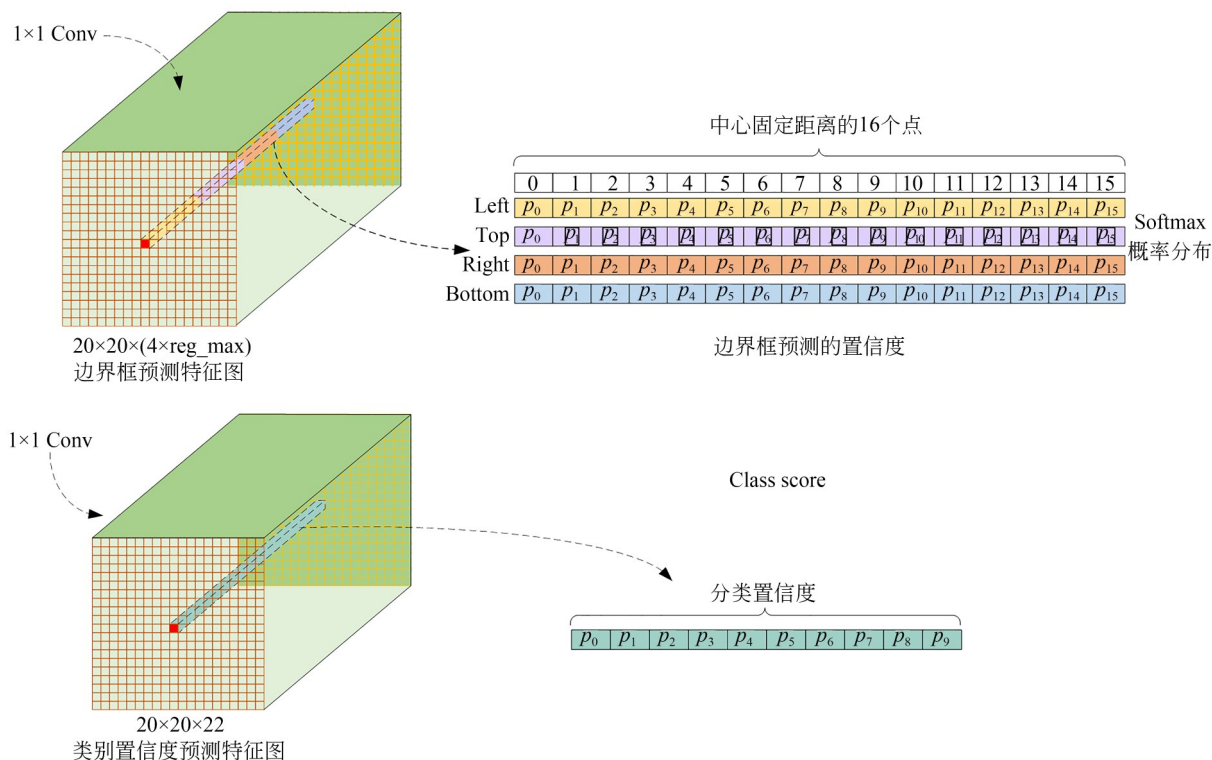


图2 网格单元对边界框和类别的预测值

Fig. 2 Predicted values of grid cells for bounding boxes and categories

框与真实框的匹配程度。 s 和 u 越高, t 越接近于1,预测框与真实框也就越匹配。

为了解决网格单元多余分配的情况,模型会将得到的正样本预测框与mask矩阵进行求和,通过留下CIoU值最大的预测框,从而得到网络的预测值。

3 改进的YOLOv8模型

为了在深海污染物检测过程中实现出色的检测性能,本文基于YOLOv8的架构提出一种Debris-yolo算法。该算法通过引入改进的双向特征金字塔网络(Bidirectional Feature Pyramid Network, BiFPN)提高模型前景与后景的区分能力。其次,设计一个轻量化检测头Efficient-Head,降低模型的计算量。最后,提出一种Wise-DIoU损失函数,采用动态梯度增益分配策略使得模型更加关注普通示例,提高模型的检测性能。

3.1 改进的双向特征金字塔网络

在深海环境中,主动光照下的深海图像会出

现颜色失真、深海垃圾外观模糊等问题。因此,不明显的视觉特征成为深海垃圾检测任务中的一个难题。针对该问题,本文采用BiFPN,其结构如图3所示。BiFPN采用双向跨尺度连接的结构获取低层次特征图的强语义特征和高层次特征图的强定位特征,确保模型能够准确预测不同尺寸的图像。为了融合更多特征,额外添加一条路径连接同级网络的输入与输出节点,并删除只有一个输入的节点。由于此节点未进行特征融合,对特征融合网络的贡献较小从而简化网络的结构。

在下采样的过程中,随着像素数量的减少,

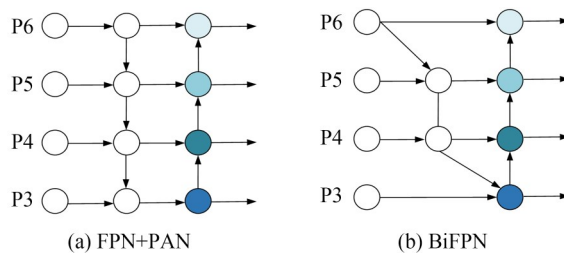


图3 特征融合网络对比

Fig. 3 Comparison of feature fusion networks

图像细节和纹理信息会丢失,这会影模型最终的处理结果和准确性。本文采用Adown结构对特征图进行下采样,如图4所示。输入特征图经过MaxPool处理后分成两个部分,一部分通过卷积对特征进行提取,另一部分通过MaxPool提取关键特征信息,最后将两部分合并,以尽可能地保留图像信息。

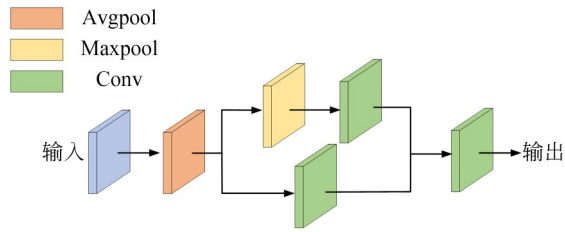


图4 Adown模块结构

Fig. 4 Structure of adown module

特征融合过程中,不同的输入特征图的分辨率不同,对输出的特征图的贡献也不相同。深海环境中采集的多尺度深海垃圾目标较多,针对这一现象,BiFPN采用加权融合的方式,对不同输入特征进行有区分的融合。网络能够根据输入特征的重要程度进行学习,挖掘深海检测目标的深层信息,解决因环境复杂造成的误检和漏检问题。其输入与输出的关系式如下:

$$o = \sum_i \frac{\omega_i \cdot I_i}{\epsilon + \sum_j \omega_j}, \quad (2)$$

式中: I_i 为输入特征; ω_i 为输入特征对应的权重,通过ReLU函数确保 $\omega_i \geq 0$ 。为避免数值不稳定,设置初始学习率 $\epsilon = 0.0001$,使得权重在 $[0,1]$ 之间下降。

3.2 轻量化检测头

潜水器捕捉的图像由于距离、种类、光线等因素干扰具有复杂性,这需要模型具备较强的多尺度特征提取能力。如图5所示,检测输出层通过卷积运算将输入图像分成 $20 \times 20, 40 \times 40, 80 \times 80$ 三个不同尺度的特征图,用于检测不同尺寸的物体。具体来说,较大的特征图感受野较大,用来检测较小的对象;较小的特征图则负责检测较大的对象,这有助于复合尺度的深海垃圾检测。

本文设计了一种轻量化检测头 Efficient Head,如图5所示。该检测头将两个预测分支改

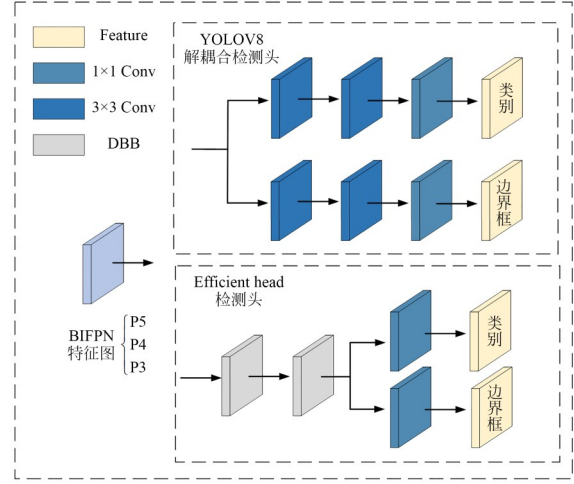


图5 轻量化检测头结构

Fig. 5 Structure of efficient head

为共享参数结构,并利用多样化分支模块解决检测头参数数量和计算量冗余问题。

受深海环境的影响,塑料与金属之间具有相似特征,容易出现误检现象。多分支结构能够从多尺度提取特征,从而有效解决垃圾类间相似性的问题,但是使用多分支并行结构会导致模型的计算量和参数数量变大,从而降低模型的检测效率。单分支结构简单,具有良好的检测效率,但它对特征的提取能力较差。多样化分支模块 DBB 融合了多分支结构和单分支结构,在训练阶段通过多分支并行结构丰富卷积块空间,获得不同感受野和复杂度的特征信息。在验证阶段, DBB 等价变换成一个分支,提高单个卷积的表现水平,实现模型的快速推理,其结构如图6所示。

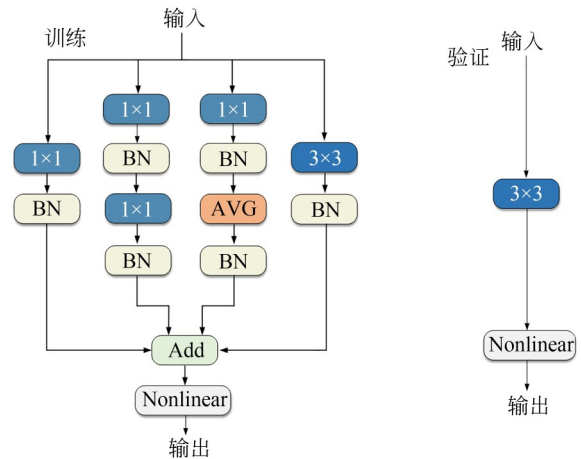


图6 多样化分支区块结构

Fig. 6 Structure of diversified branching block

在推理阶段, DBB 将卷积层与 BN 层融合, 以达到加快模型推理速度的目的, 融合过程如下:

$$\begin{cases} BN(x) = \gamma \frac{x - g_{av}(x)}{\sqrt{\sigma^2(x)}} + \beta \\ Conv = W(x) + b \\ BN(Conv(x)) = \gamma \frac{W(x)}{\sqrt{\sigma^2(x)}} + \gamma \frac{b - g_{av}(x)}{\sqrt{\sigma^2(x)}} + \beta \end{cases} \quad (3)$$

多个并行的相同卷积可以通过式(4)进行合并。对于卷积和平均池化层的融合, 可以将 kernel size 为 k , 步长为 s 的平均池化层等价于相同大小和步长的卷积, 其中 $k \times k$ 卷积的每个权重为 $1/(k \times k)$, 融合过程如下:

$$\begin{cases} W' = W_1 + W_2 \\ b' = b_1 + b_2 \end{cases}, \quad (4)$$

$$\begin{cases} W' = W_2 \otimes TRANS(W_1) \\ b' = b_1 \otimes W_2 + b_2 \end{cases}, \quad (5)$$

式中: x 为输入值; W 和 b 分别为卷积模块权重和偏置; W_1 和 b_1 代表 1×1 卷积的权重和偏置; W_2 和 b_2 代表 3×3 卷积的权重和偏置; W' 和 b' 为融合后卷积模块的权重和偏置; g_{av} 为输入的均值, σ 为方差; γ 和 β 为 BN 的权重。

3.3 Wise-DIoU 损失函数

检测过程中, 较大的位置预测偏差会导致无法精准定位目标。边界框回归损失函数对于优化待检测深海垃圾与真实深海垃圾之间的位置误差有着重要作用。实际的深海检测图像中存在一些低质量的垃圾图像, 即标注框与垃圾位置产生偏移。这会对模型的泛化能力造成不利影响。针对这一问题, 本文设计了 Wise-DIoU, 可以动态调整对边界框的惩罚。

DIoU 将两个边界框中心点之间的归一化距离作为惩罚因子。这使得网络更倾向于移动边界框的位置来减少损失函数, 解决预测框和目标框不相交时 IoU 无法优化模型的问题。其表达式为:

$$L_{DIoU} = L_{IoU} + R_{DIoU}, \quad (6)$$

$$L_{IoU} = 1 - IoU = 1 - \frac{w^i h^i}{s^u}, \quad (7)$$

$$R_{DIoU} = \frac{\rho^2(b^{gt}, b^{prd})}{c^2}, \quad (8)$$

式中: R_{DIoU} 为 DIoU 的惩罚因子; w_i 和 h_i 分别表

示目标框和预测框之间的宽和高; s_u 表示预测框与真实框并集的面积; ρ 表示预测框与真实框之间的距离; b^{gt} 为真实框, b^{prd} 为预测框; c 为这两个框最小外接矩阵的对角线长度。图 7 为边界框回归的计算因子。

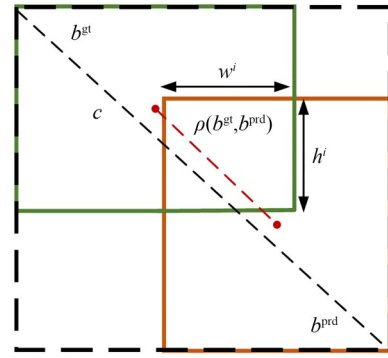


图 7 边界框回归的计算因子

Fig. 7 Calculation factor for bounding box regression

DIoU 通过引入中心点距离强化对边界框损失的拟合能力, 但同时会加剧对低质量示例的惩罚, 导致对深海垃圾检测能力的下降。Wise-DIoU 损失函数, 通过赋予低质量样本更低的梯度增益, 从而抑制模型对该示例的学习, 模型能更加专注普通质量的示例, 使得边界框回归聚焦到普通质量的锚框上。其公式如下:

$$L_{Wise-DIoU} = \gamma L_{DIoU}, \quad (9)$$

$$\gamma = \frac{\beta}{\delta \alpha^{\beta - \delta}}, \quad (10)$$

$$\beta = \frac{L_{IoU}^*}{L_{IoU}} \in [0, +\infty), \quad (11)$$

$$m = 1 - \sqrt[n]{0.5}, \quad (12)$$

式中: β 表示离群度; γ 为锚框的量化指标; 上标*表示分离操作, 用于防止产生阻碍收敛的梯度; $\overline{L_{IoU}}$ 为动量 m 的运行平均值; t 为总的训练次数; n 为批量大小。

4 数据预处理

4.1 数据集描述

本文数据来自日本海洋地球科学技术机构 (JAMSTEC) 的全球海洋学数据中心 (GODAC) 的 Deep-sea debris 数据库。该数据库中的深海废弃物集中在西太平洋地区。这些废弃物照片和视频可以追溯到 1983 年, 最大深度可达 10 900 m。

Deep-sea debris 数据库简单地对视频和照片进行分类,涵盖了渔具、布料、玻璃、金属、塑料、橡胶和自然垃圾等类别,还提供了深海废弃物沉入海底的经纬度、海洋深度和拍摄时间等。该数据库中图像少,视频多,不适合直接进行图像检测。因此,这里从视频中提取适当的帧,并与原始深海废弃物合并。对采集到的帧图像进行标注处理,形成可以用于检测任务的深海废弃物数据集。然而,该数据集采集的深海污染物图像除了废弃物以外,还存在潜水器、鱼、鳗鱼等物体。当下大多数深海污染物检测只关注垃圾本身,忽略了深海生物和检测设备在采集过程中对污染物检测的干扰。因此,在数据集制作过程中,将深海生物进行标注,创建更加真实的深海污染物检测数据集 Debris,使得模型在检测过程中能够有效地区分深海生物与深海污染物,避免在实际应用中出现错检。

Debris 数据集分为 10 类:潜水器、鱼、鳗鱼、渔具、布料、玻璃、金属、塑料、橡胶和自然垃圾。为了提升模型的稳定性和泛化能力,采用一系列数据增强技术处理选定图像。具体增强方法包括随机旋转、缩放、翻转等方式,通过这些处理将数据量扩充至 7 832 张。

图 8 展示了 Debris 中的部分图像,从左到右依次为塑料垃圾、金属垃圾以及橡胶垃圾。可以看出,该数据集塑料垃圾和金属垃圾之间存在着较多相似的特征。图像中整个场景偏暗沉,亮度

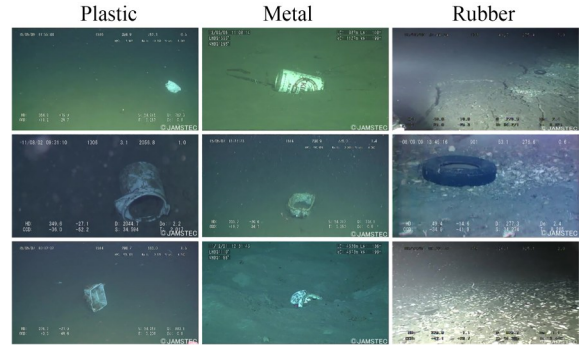


图 8 Debris 数据集中部分图像

Fig. 8 Selected images from Debris dataset

较低;此外,部分目标物与背景区别不是很明显,特别是当水下机器人离检测物体距离较远时橡胶垃圾与背景颜色较为接近,即对比度偏低。

由于 Debris 数据集集中的图像采取主动光照辅助拍摄,对于部分光源直射的场景,图像具有一定的亮度分布不均匀现象;部分图片的拍摄场景较远,图像的颜色存在明显的失真问题;对于光照较弱的场景,图像偏暗沉、亮度低;对于部分拍摄场景,图像存在明显的模糊和散射现象。

4.2 基于自适应色彩均衡水下图像复原

光线在水中的传播特性使得水下图像呈现出模糊、细节丢失、颜色偏移等现象^[26]。这些现象会使图像丢失一些特征信息,从而降低水下目标检测的准确率^[27]。本文基于自适应色彩均衡对水下图像复原算法 IBLA 进行改进,改进结构如图 9 所示。

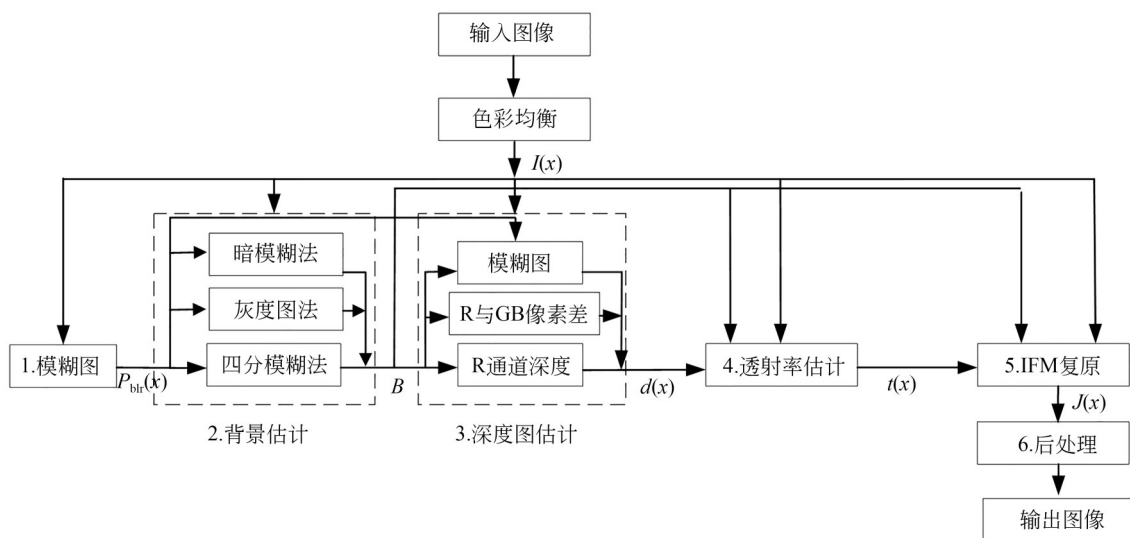


图 9 水下图像复原算法结构

Fig. 9 Structure of underwater image restoration algorithm

本文在 IBLA 算法的基础上添加色彩均衡算法用于突出图像纹理细节,以提升图像的复原效果。该算法的背景估计、深度图估计和透射率估计沿用 IBLA 中的方式。本文对衰减补偿进行改进。颜色通道补偿如下:

$$I_{rc}(\theta) = I_r(\theta) + \min(I_{sumg}/I_{sumr}, 1) \times (\bar{I}_g - \bar{I}_r) \times (1 - I_r(\theta)) \times I_g(\theta), \quad (13)$$

$$I_{bc}(\theta) = I_b(\theta) + \min(I_{sumg}/I_{sumb}, 1) \times (\bar{I}_g - \bar{I}_b) \times (1 - I_b(\theta)) \times I_g(\theta), \quad (14)$$

式中: $I_{rc}(\theta)$ 和 $I_{bc}(\theta)$ 为补偿后 x 点红通道和蓝通道的数值; $I_r(\theta)$, $I_b(\theta)$ 和 $I_g(\theta)$ 为原图像 x 点出红、蓝、绿通道的值; $\bar{I}_r$, $\bar{I}_b$ 和 $\bar{I}_g$ 为原图像在 x 点处红、蓝、绿通道均值; I_{sumr} , I_{sumb} 和 I_{sumg} 为原图像红蓝绿通道值的总和。

对每个颜色通道进行动态拉伸,得到:

$$L_{out} = \frac{L_{in}^c - L_{mean} + \lambda L_{std}}{2\lambda L_{std}}, \quad (15)$$

式中: L_{mean} 和 L_{std} 表示各颜色通道的均值和标准差; λ 表示范围调整参数; L_{in}^c 和 L_{out} 为拉伸前和拉伸后的图像。

图 10 展示了改进 IBLA 的效果,对比可以看出,基于自适应色彩均衡化改进的 IBLA 算法在视觉效果上更好,提高了检测目标特征细节。

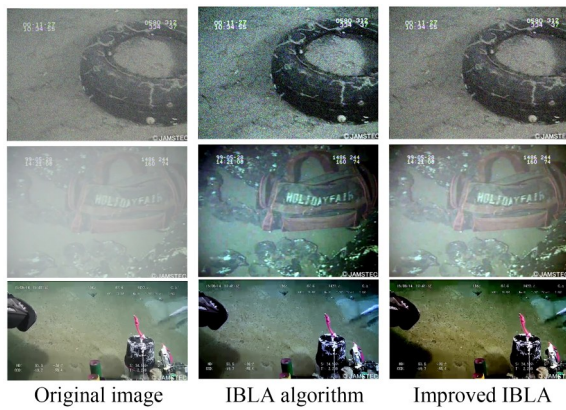


图 10 不同算法的水下图像复原结果

Fig. 10 Restoration result of underwater image by different algorithms

5 实验结果分析

5.1 评价指标

本研究分别从模型的精度和轻量化两个维度对模型进行评估。模型精度采用精确度、召回率和 mAP 分数作为评价指标。轻量化评价指标

包括检测时间、参数量、浮点运算数、模型大小和 FPS。精确度,召回率的定义如下:

$$P = \frac{TP}{TP + FP}, \quad (16)$$

$$R = \frac{TP}{TP + FN}, \quad (17)$$

其中: TP 表示模型正确判定为垃圾的真阳性数目,即 IoU 大于设定阈值 k 的边界框数量; FP 表示检测出错误的深海垃圾数量; FN 指出被错误检测为深海垃圾的假阴性数量,即 IoU 小于或等于阈值 k 的预测边界框数量。

mAP 是一个综合测量指标,它计算了在不同 IoU 阈值下的平均精度 (Average Precision, AP)。这里选择 $mAP_{0.5}$ 和 $mAP_{0.5:0.95}$ 作为度量, $mAP_{0.5}$ 表示 IoU 阈值为 0.5 时的 AP, $mAP_{0.5:0.95}$ 指的是 IoU 阈值为 0.50~0.95 的 10 个 AP 的平均值。这些特定阈值下的 mAP 值能够在特定标准下评估模型的性能,适用于不同应用环境中。mAP 和 AP 的公式如下:

$$AP = \frac{1}{m} \sum_i^m P_i = \frac{1}{m} P_1 + \frac{1}{m} P_2 + \dots + \frac{1}{m} P_m = \int P(R) dR, \quad (18)$$

$$mAP = \frac{1}{k} \sum_i^k AP_i, \quad (19)$$

式中: P 和 R 分别代表精确度和召回率; k 为本研究中所有类别的数量。

5.2 实验平台

本文所提出的 Debris-yolo 网络由 PyTorch 框架实现。实验平台的操作系统为 Ubuntu18.04; CPU 为 Intel (R) Xeon (R) Platinum 8255C CPU @ 2.50 GHz; GPU 为 NVIDIA GeForce RTX 3090 24 GB; 深度学习框架采用 python3.9+cuda12.1。

Debris 数据集的 85% 用于训练,其余 15% 用作测试。采用数据增强、水下图像复原对训练数据集进行增强。

超参数设置: 网络优化器采用 SGD, 动量因子设为 0.937, 初始学习率为 0.01; 非最大值抑制设为 0.7; cpu 线程数为 8; 模型每次训练 128 张图像, 经过 300 个 epochs 训练完成收敛。

5.3 实验结果分析

Debris-yolo 模型在验证集的 mAP_{50} 为

0.851, $mAP_{50:95}$ 为 0.645, 能准确地检测深海垃圾, 如图 11(e) 和 11(f) 所示。图 11(a) 可视化了每个训练周期的边界框损失、分类损失和分布焦点损失。可以看出, 该算法的损失函数在 100 个训练周期后逐渐收敛, 具有较快的收敛速度。损失函数在训练和验证过程中都被最小化, 从而表

明模型识别深海垃圾的能力逐步提升。此外, 图 11 还展示了每个训练周期的精确度、召回率、 mAP_{50} 和 $mAP_{50:95}$ 。在前 50 个训练周期, 精准度、召回率、和平均精度都有明显的提升, 直到第 100 个周期收敛。这一趋势表明, 模型的性能随着每个周期的训练而提高。

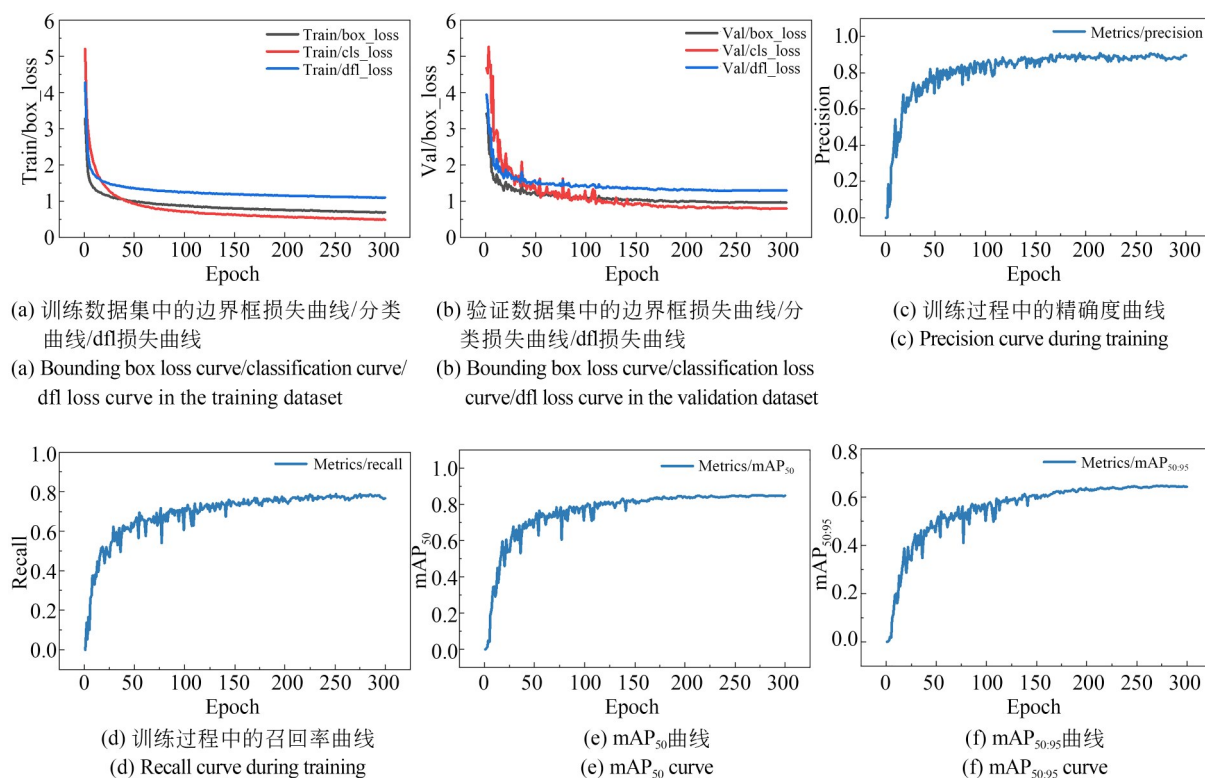


图 11 300 个 epochs 训练结果

Fig. 11 Training results for 300 epochs

混淆矩阵提供了每个类别真正例、假正例、真负例、假负例的数量。其中, 每一列为预测类别, 每一行为实际类别, 每一列的值和为 1。所提算法的混淆矩阵如图 12 所示 (彩图见期刊电子版)。可以看出, 模型对织物识别正确率最高为 91%。红色框圈住的为被误识别较多的类, 这些类别之间的形状、纹理和其他特征具有相似性, 因此模型无法较好地捕捉它们之间特征边界。对于具有高可区分性特征的类别 (如轮胎、玻璃), 模型可以较好地找到特征边界, 因此识别正确率较高。

本文选取 Debris 数据集的部分图像, 依据光学与物理特性, 探讨模型如何对目标特征进行识别与检测。如图 13 所示, 金属垃圾通常具有较高的反射率, 会导致特征图中目标区域呈现出较亮

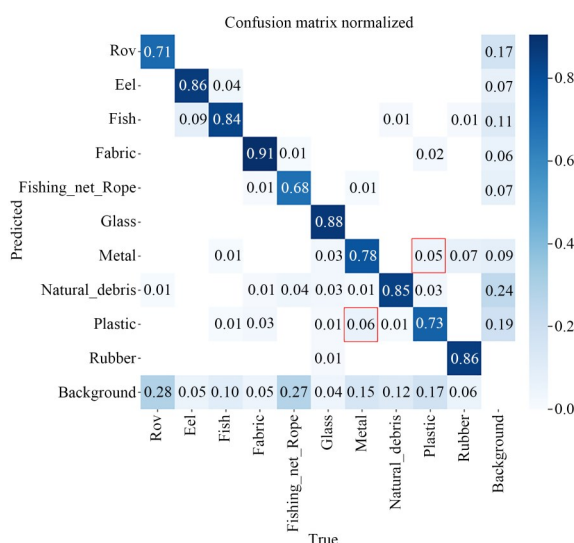


图 12 Debris-yolo 模型的混淆矩阵

Fig. 12 Confusion matrix for Debris-yolo model

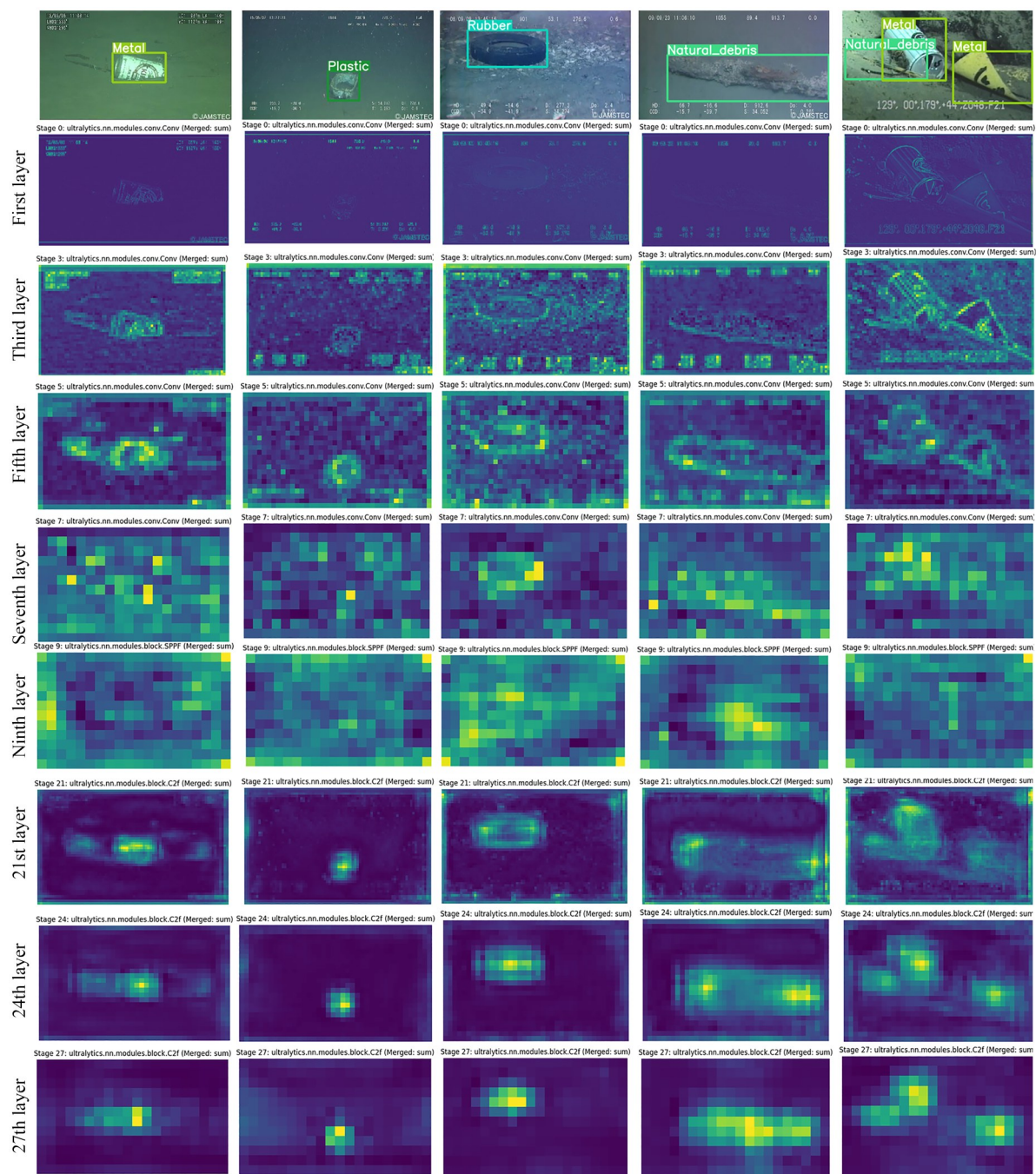


图 13 Debris-yolo模型的特征图
Fig. 13 Feature images of Debris-yolo model

的色调或者高光点。从第0层的特征图可以看出,金属区域有着较为明显的亮色区域。然而,塑料和橡胶等非金属材料通常具有较低的反射率,在图像中的颜色分布会更加均匀。因此,在第0层特征图中可以看出塑料和橡胶的颜色相对较深且均匀。除此之外,不同类型的垃圾具有不同的形状和结构。随着网络层数的增加,特征图

逐渐突出这些形状上的差异,使得模型能够更好地识别和分类不同类型的碎片。长时间暴露在水下环境中会导致材料发生不同程度的腐蚀和磨损。这种变化在物体外观上形成独特的纹理特征。在更高层次的特征图中,这些细节特征会不断得到加强,帮助模型提高识别精度。

表1为消融实验结果。A,B和C分别表示改

进的 BiFPN 模块、Efficient Head 轻量化检测头和 Wise-DIoU。引入改进 BiFPN 模块的模型 1 与 YOLOv8n 模型相比,参数量和 GFLOPS 分别减少了 41.9% 和 15.8%, $mAP_{0.5}$ 提高了 1.2%, $mAP_{0.5:0.95}$ 提高了 0.3%;引入 Wise-DIoU 后,模型 3 的 $mAP_{0.5}$ 和 $mAP_{0.5:0.95}$ 分别提高了 1.1% 和 1.1%。在不同模块的组合实验中,3 种模块的组合模型 7(Debris-Yolo)保留了模型 4 的轻量级优势。这些数据表明,模块组合能显著提高模型的整体性能。

表 1 消融实验结果
Tab. 1 Results of ablation experiments

Model	A	B	C	<i>P</i>	<i>R</i>	$mAP_{0.5}$	$mAP_{0.5:0.95}$	Parameter	GFLOPS
0	—	—	—	88.6	75.2	83.8	62.9	3.15	8.20
1	✓	—	—	90	76.2	85.0	63.2	1.83	6.90
2	—	✓	—	88.3	75.9	84.7	63.8	3.66	8.10
3	—	—	✓	88.2	78.0	84.9	64.0	2.88	8.20
4	✓	✓	—	88.3	76.9	84.6	63.1	1.63	5.60
5	✓	—	✓	87.9	77.3	84.8	62.9	1.83	6.90
6	—	✓	✓	88.9	77.6	84.9	64.5	3.66	8.10
7	✓	✓	✓	89.0	78.2	85.1	64.5	1.63	5.60

本文采用水下图像增强和数据增强对 Debris 数据集进行预处理,表 2 为数据预处理对模型检测效果的影响。模型 0 是未使用数据预处理

表 2 应用数据预处理效果对比
Tab. 2 Comparison of application data preprocessing effects

Model	<i>P</i>	<i>R</i>	$mAP_{0.5}$	$mAP_{0.5:0.95}$
0	86.2	72.2	80.7	61.7
1	89.0	78.2	85.1	66.4

的检测效果,模型 1 是使用数据预处理后的结果。可以看出,经过数据预处理后大大提高了模型的性能。

5.4 实验对比结果分析

本文选择两阶段检测网络 ResNet50-Faster-RCNN、深海垃圾监测检测方法 DSDebrisNet、单阶段网络 YOLOv5s, YOLOv10n 以及基准模型 YOLOv8n 进行比较。表 3 显示了不同模型的检测结果。

表 3 不同模型的检测结果对比
Tab. 3 Comparison of detection results of different models

Model	<i>P</i>	<i>R</i>	$mAP_{0.5}$	$mAP_{0.5:0.95}$
ResNet50-Faster R-CNN	82.5	72.3	78.9	58.2
YOLOv5s	88.2	76.8	83.5	60.9
DSDebrisNet	88.3	77.0	83.9	61.4
YOLOv8n	88.6	75.2	83.8	62.9
YOLOv10n	87.7	77.1	83.6	63.8
Debris-yolo	89.0	78.2	85.1	64.5

实验结果表明, ResNet50-Faster-RCNN 网络在模型精准率的检测指标上无法与本文提出的模型相比, ResNet50 虽然提供了更深的网络结构,但简单的双分支的残差结构无法更好地学习目标的特征信息,因此导致模型具有较差的检测

精度。Debris-yolo 模型相较于 YOLOv8n 基准模型 $mAP_{0.5}$ 提高了 1.3%, $mAP_{0.5:0.95}$ 提高了 1.6%。与 DSDebrisNet 深海垃圾检测模型相比,在 Recall, $mAP_{0.5}$, $mAP_{0.5:0.95}$ 指标上具有更好的检测性能,分别提升了 1.2%, 1.2%, 3.1%。由此可知,

本文提出的模型拥有不错的检测能力,可以更准确地定位和识别深海垃圾。

此外,本文总结了不同模型对每个类别的检测效果,结果如表 4 所示。可以看出,Debris-yolo 在大多数类别中达到了最高的 AP 值。Fabric 类别 AP 值最高,达到了 96%。这是因为 Fabric 具有较为单一的特征,且有充足的训练样本,因此很容易被模型检测到。相反 Fishing net Rope 的 AP 值最小,为 73.7%,这是因为深海附生植物附着在上面导致其形状和规模较为复杂,且训练样本较少,导致模型对 Fishing net Rope 的学习不够充足,使检测更加困难。Natural_debris, Glass 和 Rubber 通常具有较为明显的纹理和轮廓特征,且有较多的训练样本,因此具有不错的检测效果,

精度绝大部分超过 80%。对于具有相似特征的 Plastic, Metal 类别, AP 值分别为 76.5%, 83.9%, 比 DSDebrisNet 提高了 3.2%, 0.4%, 与 YOLOv8n 相比分别提升了 3.2%, 3.2%。这说明对可区分性差的类别的检测, Debris-yolo 能够取得令人满意的结果。

为了直观感受不同模型的检测效果,图 14 显示了 Debris-yolo 模型与其他模型的检测结果。从 Natural debris 的检测结果可以看出,通过生成候选框的 ResNet50-Faster-RCNN 生成了大量的预测框以尽可能框住检测对象,这虽然赋予了模型较高的召回率,然而更多的预测框会使模型的准确率下降。除此之外, ResNet50-Faster-RCNN 存在着漏检问题,其他模型同样存在着漏检、预

表 4 不同模型对每个类别检测结果的对比

Tab. 4 Comparison of results of different models for each category of detection

Modal	Rov	Eel	Fish	Fabric	Fishing_net_ Rope	Glass	Metal	Natural_ debris	Plastic	Rubber
ResNet50-Faster R-CNN	71.6	84.6	85.7	94.2	73.6	93.6	80.7	79.9	72.5	91.5
YOLOv5s	75.5	86.5	86.2	93.6	72.2	92.1	81.5	81.6	72.3	93.2
DSDebrisNet	74.7	85.5	86.4	93.4	71.8	93.8	83.5	79.9	73.3	96.7
YOLOv8n	74.6	84.4	87.6	94.6	72.5	95.1	80.7	81.6	73.3	93.7
YOLOv10n	70.7	90.2	87.8	93.9	72.6	94.4	82.0	81.0	74.0	89.8
Debris-yolo	78.3	87.6	88.3	96.0	73.7	92.4	83.9	82.1	76.5	91.8

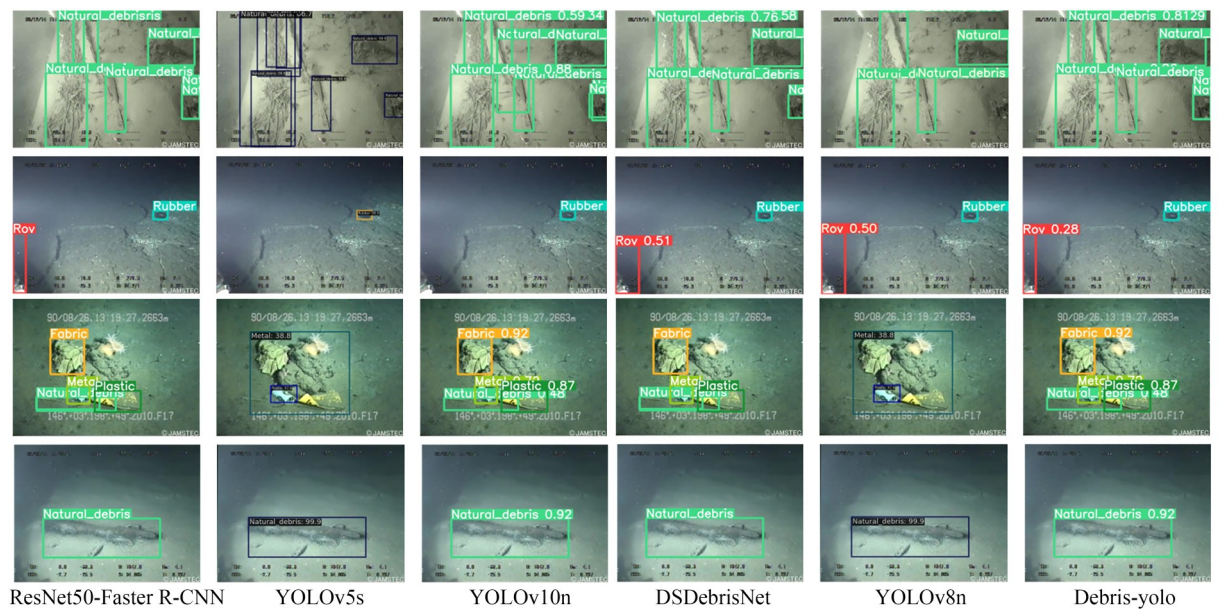


图 14 不同模型部分检测结果的可视化

Fig. 14 Visualisation of detection results from different model sections

测框冗余的问题。

为了展示本文提出算法的轻量级优势,将 Debris-yolo 算法与其他算法进行了轻量级度量的比较,结果如表 5 所示。实验结果表明,Debris-yolo 在轻量级方面优于 YOLOv8n 算法,参数量减少了 48.2%,GFLOPS 减少了 36%,模型大小也减少了 28%;Debris-yolo 的参数量、GFLOPS 和模型大小分别为轻量级 DSDebrisNet 的 0.24, 0.35, 0.31 倍。因此,Debris-yolo 在轻量化方面有着出色的表现。

本文从 Deep-sea debris 数据库中选取 3 种光线衰减不同场景下的深海探测视频,对比不同模型在不同情形下的检测速度和检测时间,结果如表 6 所示。其中,Scene1 为视频呈蓝色调的橡胶垃圾视频,Scene2 为视频呈黄绿调的自然垃圾视频,Scene3 为视频呈黄色调的塑料检测。从表 6

表 5 不同模型轻量级指标对比
Tab. 5 Comparison of different model lightweight metrics

Model	Detection	Parame- ters/MB	GFLOPS	Size/ MB
	speed/ (frame• s ⁻¹)			
Faster R-CNN	19.11	41.39	208.00	478
YOLOv5s	102.04	6.71	15.80	14.5
DSDebrisNet	95.24	6.71	15.80	13.8
YOLOv8n	88.49	3.15	8.70	6.00
YOLOv10n	98.45	2.30	66.70	5.50
Debris-yolo	90.91	1.63	5.60	4.33

可以看出,Debris-yolo 在 3 个测试场景中具有较高的 FPS 和较低的处理时间,满足水下机器人实时检测的需求。

表 6 不同情形下的检测速度和检测时间
Tab. 6 Detection speed and inspection time in different situations

Model	Scene 1		Scene 2		Scene 3	
	FPS	t/ms	FPS	t/ms	FPS	t/ms
Faster R-CNN	39.6	25.27	43.1	23.20	42.9	23.31
YOLOv5s	193.6	5.17	208.5	4.80	191.3	5.23
DSDebrisNet	187.3	5.34	201.3	4.97	186.2	5.37
YOLOv8n	181.9	5.49	209.9	4.76	191.6	5.20
YOLOv10n	176.8	5.08	208.1	4.80	195.7	5.11
Debris-yolo	185.7	5.38	210.2	4.76	194.0	5.16

6 结 论

根据深海污染物检测要求高精度、易部署的要求,本文提出深海垃圾检测模型。构建 Debris 数据集,该数据集包括深海垃圾、深海生物、水下机器人在内共有十种类别。以确保训练数据更接近真实场景。为了提高模型对前景和背景的区分能力,引入改进的双向特征金字塔网模块,提高模型特征提取能力。通过采用多分支处理结构减少下采样过程中信息丢失问题,并大大降低模型的参数量。基于共享参数思想和 DBB (Diverse Branch Block)重参数化模块设计一种轻量化检测

头 Efficient-Head 解决了检测模型检测头参数量、计算量过大的问题。提出一种 Wise-DIoU 损失函数,以中心点距离作为惩罚项,采用动态梯度增益分配策略使得模型更加关注普通示例,避免低质量示例对模型泛化性能的下降。采用数据增强和自适应水下图像复原技术对训练数据进行预处理,提高图像中目标的特征细节提高模型的准确性。实验结果表明,本文提出的模型在精确度,召回率, $mAP_{0.5}$, $mAP_{0.5:0.95}$ 精度评估指标上取得了出色的结果,分别为 89%, 78.2%, 85.1%, 64.5%。Debris-yolo 在 Parameters, GFLOPS, Size 轻量化评估指标上同样有着卓越的表现,参数量,GFLOPS 和模型大小分别减少了 48.2%,

36%, 28%。其紧凑的结构能够快速、经济地部署在检测设备中, 总体而言, 该模型提高了检测深海垃圾的性能, 同时降低了计算的复杂性, 基本满足可操作水下机器人的作业要求。

参考文献:

- [1] GALIMANY E, MARCO-HERRERO E, SOTO S, *et al.* Benthic marine litter in shallow fishing grounds in the NW Mediterranean Sea [J]. *Waste Management*, 2019, 95: 620-627.
- [2] HIDALGO-RUZ V, THIEL M. Distribution and abundance of small plastic debris on beaches in the SE Pacific (Chile): a study supported by a citizen science project [J]. *Marine Environmental Research*, 2013, 87: 12-18.
- [3] TOPOUZELIS K, PAPAGEORGIOU D, SUARIA G, *et al.* Floating marine litter detection algorithms and techniques using optical remote sensing data: a review [J]. *Marine Pollution Bulletin*, 2021, 170: 112675.
- [4] SHI H H, FRIAS J, SAYED A E H, *et al.* Small plastic fragments: a bridge between large plastic debris and micro- & nano-plastics [J]. *TrAC Trends in Analytical Chemistry*, 2023, 168: 117308.
- [5] WU D, LIU J C, CORDOVA M, *et al.* The PlastoPol system for marine litter monitoring by citizen scientists [J]. *Environmental Modelling & Software*, 2023, 169: 105784.
- [6] PENG X, CHEN M, CHEN S, *et al.* Microplastics contaminate the deepest part of the world's ocean [J]. *Geochemical Perspectives Letters*, 2018: 1-5.
- [7] RANGEL-BUITRAGO N, VERGARA-CORTÉS H, BARRÍA-HERRERA J, *et al.* Marine debris occurrence along Las Salinas beach, viña del mar (Chile): magnitudes, impacts and management [J]. *Ocean & Coastal Management*, 2019, 178: 104842.
- [8] SCHLINING K, VON THUN S, KUHNZ L, *et al.* Debris in the deep: Using a 22-year video annotation database to survey marine litter in Monterey Canyon, central California, USA [J]. *Deep Sea Research Part I: Oceanographic Research Papers*, 2013, 79: 96-105.
- [9] JONES E S, ROSS S W, ROBERTSON C M, *et al.* Distributions of microplastics and larger anthro-

作者贡献声明:

张彪: 测量方法的提出, 论文构思和撰写;
朱振洋: 实验设计及数据分析;
许家忠: 论文审核与编辑。

- pogenic debris in Norfolk Canyon, Baltimore Canyon, and the adjacent continental slope (Western North Atlantic Margin, U. S. A.) [J]. *Marine Pollution Bulletin*, 2022, 174: 113047.
- [10] MIKKELSEN N, PLANQUE B, ARNEBERG P, *et al.* Multiple stakeholders' perspectives of marine social ecological systems, a case study on the Barents Sea [J]. *Ocean & Coastal Management*, 2023, 242: 106724.
- [11] JIA T L, KAPELAN Z, DE VRIES R, *et al.* Deep learning for detecting macroplastic litter in water bodies: a review [J]. *Water Research*, 2023, 231: 119632.
- [12] 王立玲, 单忠宇, 马东, 等. 基于改进 Camshift 算法的 NAO 机器人目标跟踪 [J]. 半导体光电, 2020, 41(6): 896-901.
WANG L L, SHAN ZH Y, MA D, *et al.* NAO robot target tracking based on improved camshift algorithm [J]. *Semiconductor Optoelectronics*, 2020, 41(6): 896-901. (in Chinese)
- [13] WANG J S, DONG J H, TANG M R, *et al.* Identification and detection of microplastic particles in marine environment by using improved faster R-CNN model [J]. *Journal of Environmental Management*, 2023, 345: 118802.
- [14] DING Z J, LIU C, LI D W, *et al.* Deep-sea biological detection method based on lightweight YOLOv5n [J]. *Sensors*, 2023, 23(20): 8600.
- [15] KHAN A, FOUDA M M, DO D T, *et al.* Underwater target detection using deep learning: methodologies, challenges, applications, and future evolution [J]. *IEEE Access*, 2024, 12: 12618-12635.
- [16] FULTON M, HONG J, ISLAM M J, *et al.* Robotic detection of marine litter using deep visual detection models [C]. 2019 *International Conference on Robotics and Automation (ICRA)*. May 20-24, 2019. Montreal, QC, Canada. IEEE, 2019: 5752-5758.
- [17] 王普威. 基于改进 Mask R-CNN 的水下污染物目标检测与分割研究 [D]. 哈尔滨: 哈尔滨工程大

- 学, 2022: 53-70.
- WANG P W. *Research on Underwater Pollutant Target Detection and Segmentation Based on Improved Mask R-CNN* [D]. Harbin: Harbin Engineering University, 2022: 53-70 (in Chinese).
- [18] XUE B, HUANG B X, WEI W B, *et al.* An efficient deep-sea debris detection method using deep neural networks [J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2021, 14: 12348-12360.
- [19] BAJAJ R, GARG S, KULKARNI N, *et al.* Sea debris detection using deep learning: diving deep into the sea [C]. 2021 *IEEE 4th International Conference on Computing, Power and Communication Technologies (GUCON)*. September 24-26, 2021. Kuala Lumpur, Malaysia. IEEE, 2021: 1-6.
- [20] CHEN G, HUANG B X, CHEN X Y, *et al.* Deep blue AI: a new bridge from data to knowledge for the ocean science [J]. *Deep Sea Research Part I: Oceanographic Research Papers*, 2022, 190: 103886.
- [21] HAN G Q, CEN H B, JIANG J H, *et al.* Applying machine learning in devising a parsimonious ocean mixing parameterization scheme [J]. *Deep Sea Research Part II: Topical Studies in Oceanography*, 2022, 203: 105163.
- [22] REN Y B, LI X F, ZHANG W H. A data-driven deep learning model for weekly sea ice concentration prediction of the pan-Arctic during the melting season [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2022, 60: 3177600.
- [23] HUANG H, CAI H W, QURESHI J U, *et al.* Proceeding the categorization of microplastics through deep learning-based image segmentation [J]. *Science of The Total Environment*, 2023, 896: 165308.
- [24] VARGHESE R, SAMBATH M. YOLOv8: a novel object detection algorithm with enhanced performance and robustness [C]. 2024 *International Conference on Advances in Data Engineering and Intelligent Computing Systems (ADICS)*. April 18-19, 2024. Chennai, India. IEEE, 2024: 1-6.
- [25] 王燕, 王宏辉, 刘树东, 等. 多级特征筛选和任务动态对齐的声呐图像小目标检测 [J]. *光电工程*, 2024, 51(10): 128-141.
- WANG Y, WANG H H, LIU SH D, *et al.* Small target detection in sonar images with multi-level feature screening and task dynamic alignment [J]. *Opto-Electronic Engineering*, 2024, 51(10): 128-141. (in Chinese)
- [26] 林森, 查子月. 结合色彩补偿与双背景光融合的水下图像复原 [J]. *光学精密工程*, 2024, 32(7): 1059-1074.
- LIN S, ZHA Z Y. Underwater image restoration method combining color compensation and dual background light fusion [J]. *Opt. Precision Eng.*, 2024, 32(7): 1059-1074. (in Chinese)
- [27] 江晋剑, 王远志, 钱萌. 基于小波的水下图像对比度增强方法研究 [J]. *计算机技术与发展*, 2008, 18(4): 35-37, 41.
- JIANG J J, WANG Y ZH, QIAN M. Research of underwater image processing based on wavelet transform [J]. *Computer Technology and Development*, 2008, 18(4): 35-37, 41. (in Chinese)

作者简介:



张 彪(1986—),男,黑龙江哈尔滨人,博士,副教授,2011年、2020年于哈尔滨工程大学分别获得硕士和博士学位,主要从事船舶运动预测与控制、数字图像处理和深度学习算法方面的研究。E-mail: zhangbiao@hrbust.edu.cn



许家忠(1977—),男,黑龙江哈尔滨人,教授,2007年、2011年于哈尔滨理工大学分别获得硕士和博士学位,主要从事机器人智能控制、工业机器人技术、智能装备及机电一体化技术、复合材料成型工艺及智能装备方面的研究。E-mail: comlwxtg0022@163.com