

基于 GS-YOLOv8 的轻量化水下生物目标检测算法

周梦雯, 李海涛, 张俊虎

(青岛科技大学 信息科学技术学院, 山东 青岛 266061)

摘要: 为解决现有的水下生物目标检测模型参数过多, 难以部署到资源有限的移动端的问题, 提出了一种基于 GS-YOLOv8 的轻量化水下生物目标检测模型; 该模型基于 YOLOv8s 模型进行改进, 设计一种轻量化的 RepHGNetV2 网络作为 YOLOv8s 的主干网络, 以降低模型的计算复杂度和参数量; 使用轻量化卷积 GSConv 替换颈部网络中所有的标准卷积, 进一步减少模型参数, 提高检测性能; 引入设计的 C2fAK 模块, 使模型能够更好地适应不同形状和大小水下生物目标, 从而提高检测精度; 实验结果显示, 在 URPC2020 数据集上, 与原模型 YOLOv8s 相比, 改进后的 GS-YOLOv8 网络模型的参数量降低了 37.7%, 计算量降低了 27.8%, mAP@0.5 提高了 0.9%; 此外, 与目前较为先进的 YOLOv10 模型相比, 改进后的 GS-YOLOv8 模型在检测精度和轻量化方面更有优势。

关键词: YOLOv8; 水下目标检测; GSConv; 轻量化; HGNetV2; AKConv

A Lightweight Underwater Biological Target Detection Algorithm Based on GS-YOLOv8

ZHOU Mengwen, LI Haitao, ZHANG Junhu

(School of Information Science and Technology, Qingdao University of Science and Technology, Qingdao 266061, China)

Abstract: In order to solve the problems of too many parameters and difficulty in deploying to mobile terminals with limited resources in existing underwater biological target detection models, a lightweight underwater biological target detection model based on GS-YOLOv8 is proposed. The model is improved based on YOLOv8s model, and a lightweight RepHGNetV2 network is designed as the backbone network of YOLOv8s to reduce the computational complexity and the number of parameters of the model. The lightweight convolutional GSConv is used to replace all standard convolution in the neck network, further reducing model parameters and improving detection performance. The C2fAK module is introduced to make the model better adapt to different shapes and sizes of underwater biological targets, so as to improve the detection accuracy. Experimental results show that on the URPC2020 dataset, compared with the original model YOLOv8s, the improved GS-YOLOv8 network model reduces the parameters by 37.7%, the computation complexity is reduced by 27.8%, and the mAP@0.5 is increased by 0.9%. In addition, compared with the currently advanced YOLOv10 model, the improved GS-YOLOv8 model has notable advantages in detection accuracy and lightweight.

Keywords: YOLOv8; underwater object detection; GSConv; lightweight; HGNetV2; AKConv

收稿日期:2024-09-30; 修回日期:2024-11-11。

基金项目:山东省重点研发计划(科技示范工程)课题(2021SFGC0701);青岛市海洋科技创新专项(22-3-3-hygg-3-hy)。

作者简介:周梦雯(1999-),女,硕士研究生。

张俊虎(1974-),男,博士,副教授,硕士生导师。

通讯作者:李海涛(1978-),男,博士,副教授,硕士生导师。

引用格式:周梦雯,李海涛,张俊虎. 基于 GS-YOLOv8 的轻量化水下生物目标检测算法[J]. 计算机测量与控制, 2025, 33(11): 65-72.

0 引言

海洋是生命之源, 储藏着大量尚未开发的生物和矿产资源。目前, 与海洋资源开发利用相关的设备和技术种类繁多, 其中水下机器人占有重要地位, 也是近年来的发展热点。可控水下机器人进行的水下作业和海底勘探可应用于渔业、海洋矿产勘探、海洋考古等领域。它不仅可以降低人工操作的风险, 提高工作效率, 还可以降低成本, 有利于水产养殖业的发展。

水下目标的检测与识别是水下机器人的关键技术。水下生物目标检测技术主要分为传统的目标检测算法和基于深度学习的目标检测算法。传统的目标检测算法一般包括 3 个主要阶段: 区域选择、特征提取和特征分类^[1]。在区域选择阶段, 其目的是定位目标, 因为目标在图像中的位置和宽高比可能会有所变化。通常情况下, 这一阶段通过使用滑动窗口策略在整个图像上执行检测, 考虑了不同的尺度和宽高比。接下来, 采用多种特征提取技术, 如尺度不变特征变换^[2]和梯度直方图^[3]等, 以获取相关特征。最后, 通过支持向量机^[4]分类器对这些提取的特征进行分类。然而, 传统目标检测方法存在一些局限性。一方面, 区域选择策略缺乏针对性, 这导致了时间复杂度较高且存在窗口冗余的问题; 另一方面, 在复杂的水下环境中, 由于难以设计出适用于所有情况的统一特征模型, 人工提取的特征使得检测鲁棒性较差。近年来, 随着深度学习的飞速发展, 相对于传统的检测方法而言, 基于深度学习的目标检测在鲁棒性和准确率上表现更为优越。因此, 越来越多研究者将基于深度学习的目标检测算法应用到水下检测, 并取得了显著的效果^[5]。这类算法主要分为单阶段和两阶段两大类。其中, 两阶段目标检测算法如 R-CNN^[6]、Faster R-CNN^[7]等具有较高准确率, 但其速度通常较慢, 因此可能不适合实时应用场景。文献 [8] 将图像增强技术集成到扩展的 VGG16 特征提取网络中, 并使用带有特征映射的 Faster R-CNN 网络, 使用 URPC 数据集对水下生物目标进行检测和识别。单阶段目标检测的代表性算法有 SSD^[9]和 YOLO^[10]系列等。这些算法提供了更快的检测速度, 使其成为水下目标检测领域的热门研究领域。目前, 许多目标检测算法通过优化网络结构来提高模型的检测准确率, 但这往往会增加其参数和复杂度, 难以将其部署到资源有限的移动端。因此, 目标检测的轻量化研究逐渐受到重视。文献 [11] 提出了一种基于改进 YOLOv5 的水下目标检测算法, 该算法在 YOLOv5 算法的基础上添加了 Ghost 模块和 SE 注意力机制, 提高了目标检测的精度和速度, 减少了模型参数量。文献 [12] 提出了一种改进的基于 YOLOv4 的轻量级水下目标检测网络, 模型显著减小了模型尺寸, 但

参数量仍为 49.2 M, 需要进一步优化。文献 [13] 提出了一种基于 YOLOv5 的轻量级海洋生物目标检测算法, 用深度可分离卷积代替传统的卷积, 并引入了一个基于 Ghost 卷积的 C3 模块, 降低参数量和计算复杂度, 提高推理速度, 最后通过引入 RepVGG 模块, 提高模型的特征提取能力。文献 [14] 提出了一种轻量化的水下底栖生物检测算法 YOLOv7-RFPCW, 重新设计 YOLOv7 的网络结构, 降低主干网络的参数量和计算量, 设计了 P-ELAN 和 P-ELAN-W 模块, 进一步减少模型参数, 引用 CBAM 注意力模块增强特征融合, YOLOv7-RFPCW 模型显著减小了参数量和计算量。

综上所述, 现有的目标检测在水下目标检测领域取得了一定的成功, 但还存在着一些不足之处。大部分轻量化后的网络模型参数仍然过多, 对于硬件要求较高, 不利于移动端部署的任务。为了解决上述问题, 本文提出了一种轻量化的水下生物目标检测模型 GS-YOLOv8。该模型基于 YOLOv8s 进行改进, 首先使用轻量化的 RepHGNNetV2 网络重构 YOLOv8s 的主干网络, 使得主干网络的参数量和计算量大大减少; 将颈部网络中所有的 Conv 替换为轻量化的 GSconv (Gfriend Simulator Conversation)^[15]卷积, 进一步减少模型的参数量和计算量; 最后使用设计的全新模块 C2fAK 替换颈部网络中的 C2f 模块, 使得网络充分融合特征信息, 提高模型的性能。

1 算法原理介绍

YOLOv8 支持多种计算机视觉任务, 特别是在目标检测领域, YOLOv8 目标检测模型是性能最好的模型之一。YOLOv8 有 5 个版本, 即 YOLOv8n, YOLOv8s, YOLOv8m, YOLOv8l, YOLOv8x。这些版本在网络深度、宽度、参数大小和计算需求上各不相同, 最小的是 YOLOv8n, 最大的是 YOLOv8x。本文选用轻量级的 YOLOv8s, 选择这个版本是因为它在有限的计算成本下提供了更快的检测速度和更高的检测精度。如图 1 所示, YOLOv8s 由 3 个部分组成: Backbone, Neck, 及 Head。

主干网络 (Backbone) 包括 Conv、C2f 和 SPPF 模块, 使用两个堆叠的 Conv 模块提取初始特征。C2f 模块是残差特征学习模块, 通过跨层连接生成特征表示能力更强的神经网络模块。在 Backbone 网络的末端, 仍然使用最流行的 SPPF 模块, 由 3 个 5×5 大小的最大池化进行序列化, 然后拼接每一层, 在保证轻量化的同时保证各种尺度下目标的检测准确性。

颈部网络 (Neck) 采用路径聚合特征金字塔网络 (PAFPN) 结构, 增强了不同尺度下特征层信息的融合和利用。它首先通过特征金字塔网络 (FPN)^[16]以自顶

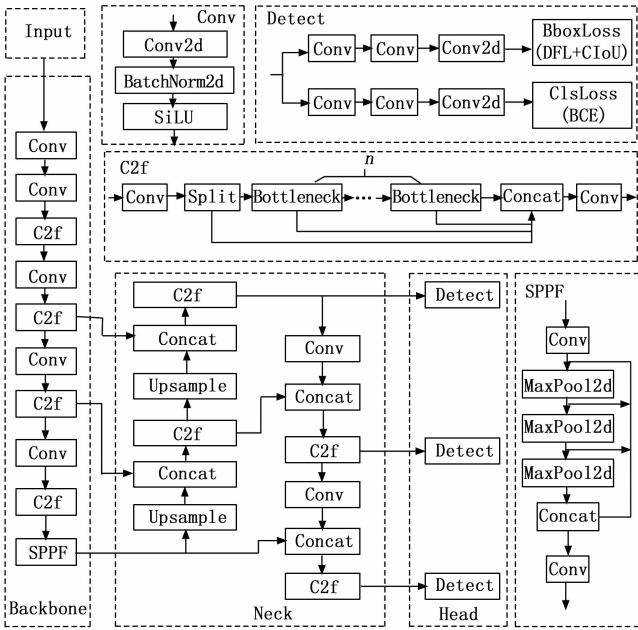


图 1 YOLOv8s 模型结构图

向下的方式传播深度特征图的语义信息。然后通过路径聚合网络 (PANet) 将纹理和细节信息从低级特征映射传播到高级特征映射。最后, 向 Head 模块输入 3 个不同尺度的特征图。

Head 中 YOLOv8 的检测部分采用解耦头结构。解耦的头结构使用两个独立的分支进行对象分类和预测边界盒回归, 并对两类任务使用不同的损失函数。对于分类任务, 使用 BCE Loss。对于回归任务, 使用 DFL Loss+CIoU Loss。该检测结构可以提高检测精度, 加快模型收敛速度。YOLOv8 是一种无锚检测模型, 可以简洁地指定阳性和阴性样本。利用任务对齐分配器对样本进行动态分配, 提高了模型的检测精度和鲁棒性。

2 GS-YOLOv8 模型

针对现有的目标检测网络模型参数过多、计算量大等问题, 本文提出了一种轻量化水下生物目标检测模型 GS-YOLOv8。其结构如图 2 所示, 主要从 3 个方面对 YOLOv8s 进行改进, 粗体框表示的地方即为改进的部分。首先使用 RepHGNetV2 网络替换 YOLOv8s 主干网络, 显著减少模型的参数量和计算复杂度, 实现模型的轻量化; 其次在颈部网络引入 GSConv 模块, 减少信息损失, 加快计算速度, 使模型既能获得较低的计算复杂度, 又能获得较高的检测准确度; 最后引入 C2fAK 模块, 使模型能够更好地适应不同形状和大小水下生物目标, 增强对目标的检测能力。

2.1 改进的 RepHGNetV2 网络

HGNetV2 (High-Performance GPU Net V2)^[17] 是由百度飞桨视觉团队研发的一种轻量级网络, 是 RT-

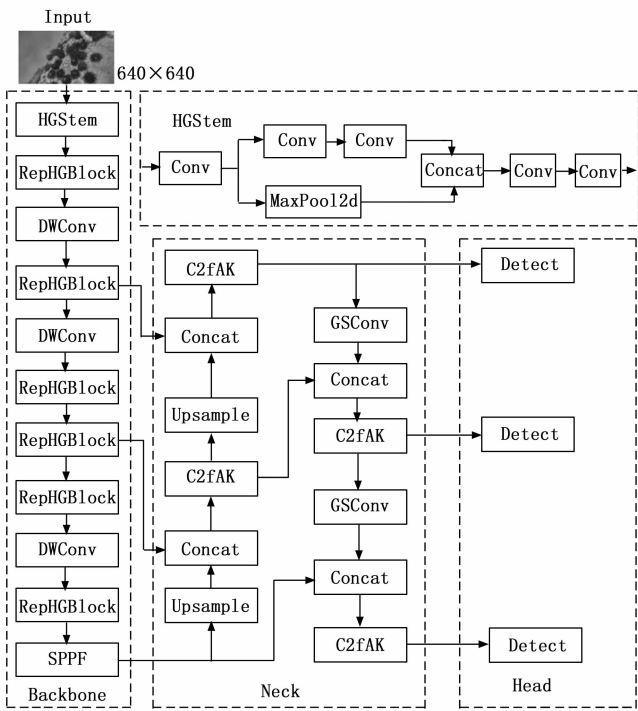


图 2 GS-YOLOv8 模型结构图

DETR 模型的主干网络。HGNetV2 网络的设计理念是采用分层的方式来提取特征, 在不同的尺度和抽象层次上学习, 从而增强其处理复杂图像数据的能力。这种分层和高效的特征提取方式特别适用于图像分类等任务。HGNetV2 的网络结构如图 3 所示, 采用多阶段设计, 主要包括一个 HGStem 模块和 4 个阶段模块。HGStem 模块作为网络的初始预处理层, 由 Conv 模块、Concat 模块和池化模块三部分组成。其主要功能是降低输入图片的分辨率, 减少后续的计算量, 同时增加通道数, 初步提取低层特征。每个阶段都包含 HGBlock 模块, 其中阶段 1 仅包含一个 HGBlock 模块, 而阶段 2 至 4 则包含一个可学习下采样 (LDS) 层和多个 HGBlock 模块。HGBlock 模块有 5 个 3×3 的卷积, 能够捕获不同维度的特征。随后, 通过 1×1 卷积核的卷积层来压缩和处理级联的特征, 并执行连接操作以合并和输出特征信息。LDS 层采用深度可分离卷积 (DWConv) 进行卷积操作, 从而有效降低计算量和参数数量。阶段模块的主要功能是深度提取特征, 当前阶段的输出将作为下一阶段的输入, 逐步提取更高层次的特征。此外, 不同阶段输出的特征图可以作为多尺度特征, 有助于更有效地检测不同大小的目标。

考虑到 HGBlock 中大量的 3×3 卷积会显著增加参数量, 本文提出使用 RepConv^[18] 替换 HGBlock 模块中的所有 3×3 卷积。RepConv 是一种由华中科技大学研究团队提出的卷积神经网络结构, 旨在改进目标检测模型的性能。该模块通过引入重复卷积 (Repetition

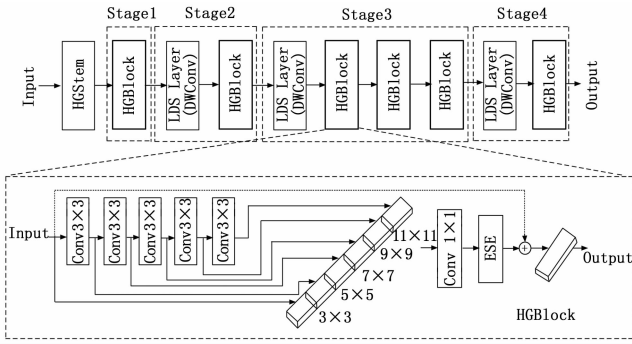


图 3 基于 HGBlock 的 HGNetV2 网络

Convolution), 在每个卷积层中多次重复应用相同的卷积核, 从而减少参数量、提高特征表达能力, 并增强模型的鲁棒性。与传统卷积相比, RepConv 具有更少的参数, 可以有效地增加模型的感受野, 提升特征提取的能力。如图 4 所示, RepHGBlock 模块利用多个 RepConv 模块从不同维度提取特征信息。随后, 这些多维特征信息会被拼接在一起, 并通过 1×1 卷积层进行压缩处理。这种压缩操作有效地重新校准特征映射中每个通道的权重, 从而整合来自低层和高层的特征。改进后的 RepHGNetV2 网络结构如图 4 所示, 利用多个 RepHG-

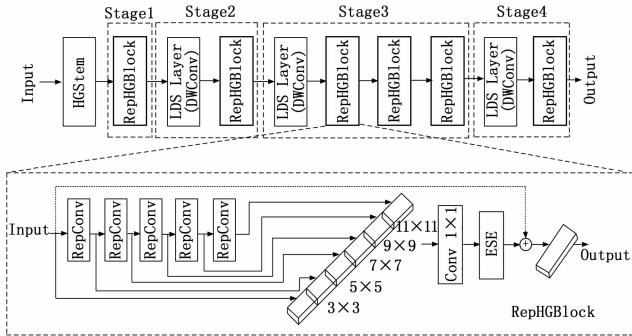


图 4 基于 RepHGBlock 的 RepHGNetV2 网络

Block 模块捕获不同抽象层次的特征, 该结构有效地集成了较低和较高级别的特征, 增强了对上下文信息的理解。将 RepHGNetV2 网络作为 YOLOv8s 的主干网络, 减少了模型的参数量和计算量, 使得该模型更易于在移动端进行部署。

2.2 GSConv 模块

在水下生物目标检测任务中, 对于检测精度有着较高的要求, 而推理速度则成为至关重要的因素。一般来说, 模型的参数越多, 检测精度越高。然而, 仅仅追求更高的准确率已经不再是水下生物目标检测任务的完美解决方案。GSConv+Slim-Neck 是针对车载边缘自动驾驶计算平台提出的轻量级网络。此网络设计的目的是为了更方便高效的目标检测, 满足实时应用需求。GSConv 在模型精度和速度之间取得平衡, 在保持精度的同时实现模型轻量化。因此, 本文提出的 GS-YOLOv8 模型将

YOLOv8s 颈部网络中 Conv 卷积替换为轻量化的 GSConv 卷积, 为的是降低模型复杂度, 提高检测速度。GSConv 模块包括 DWConv 模块、Conv 模块、Concat 模块和 Shuffle 模块。如图 5 所示, 假设输入特征映射中的通道数为 $C1$ 。GSConv 模块首先用标准卷积对输入进行下采样; 然后, 使用 DWConv 深度卷积, 将两个卷积的输出连接起来进行特征融合; 随后, 它执行 Shuffle 操作来改变通道的数量, 以便将前两个卷积的相应数量的通道连接起来。最后, 特征图中的输出通道数为 $C2$ 。通过 Shuffle 操作通道信息被均匀打乱, 保证了多通道信息的有效保存, 增强了对语义信息的提取, 充分融合特征信息。与标准卷积相比, GSConv 具有最小化信道间隐藏连接的优点, 同时保持较低的时间复杂度, 减少信息损失, 加快计算速度。但如果在模型的所有阶段都使用 GSConv, 则模型的网络层将更深, 从而显著增加推理时间。当特征映射到达颈部时, 通道尺寸被最大化, 宽度和高度尺寸被最小化, 并且不需要进一步的转换。因此, 选择仅在 Neck 网络中使用 GSConv。在此阶段, 使用 GSConv 对注意机制中的特征图进行处理, 冗余和重复信息较少, 不需要压缩, 注意机制更加有效, 从而减少了模型参数和推理时间。

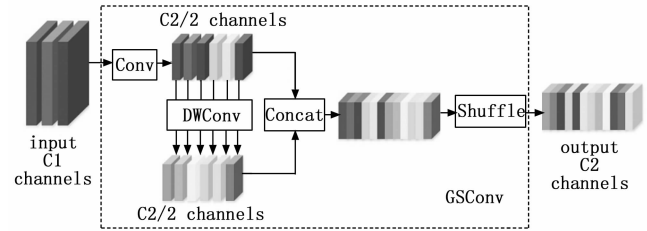


图 5 GSConv 模块

2.3 改进的 C2fAK 模块

标准卷积有 2 个缺陷: 首先, 卷积运算在固定大小的窗口运行, 无法捕获其他窗口的信息, 并且窗口的形状是固定的; 其次, 卷积核的尺寸固定为 $k \times k$, 这是一个固定的正方形, 随着 k 增加, 参数会快速增加。这种固定的窗口形状和卷积核尺寸使得标准卷积在处理水下生物数据集时表现不佳, 因为水下生物通常体型较小, 而且形状、大小在不同位置变化多端。为了提高对这些变化多样目标的检测精度, 本文提出将 AKConv (Adaptive Kernel Convolution)^[19] 引入到 YOLOv8 模型的 C2f 结构中。

AKConv 是一种创新的变核卷积, 它提供了一种灵活的卷积机制, 允许卷积核具有任意数量的参数和任意采样形状, 能够使用任意数量的参数来提取特征。这种方法突破了传统卷积在局部窗口和采样形状上的限制, 使卷积运算能够更精确地适应不同的数据集和目标位置, 从而显著提升了卷积神经网络的性能。如图 6 所

示, AKConv 的具体工作原理如下: 首先, 输入图像的维度为 (C, H, W) , 其中 C 是通道数, H 和 W 分别是图像的高度和宽度, 接着给出了 AKConv 特有的卷积核的初始采样形状, 对输入图像执行卷积操作, 通过学习得到的偏移量来调整初始采样形状, 这一步是 AKConv 的关键, 允许卷积核形状动态调整以适应图像的特征, 然后根据调整后的采样形状对特征图进行重采样, 重采样后的特征图经过重塑、再次卷积、标准化, 最后通过激活函数 SiLU 输出最终结果。通过这种动态调整采样形状的方式, AKConv 能够根据图像内容灵活改变卷积操作, 显著提高特征提取的精度和效率。此外, 这种灵活的卷积核设计不仅提升了模型性能, 还降低了模型的参数量和计算开销, 尤其在轻量级模型的应用中展现出明显优势。综上所述, AKConv 增强了卷积操作的灵活性和适应性, 在不牺牲性能的前提下, 提供了显著的资源效率, 使其特别适用于处理具有多样形状和大小的目标的图像任务。

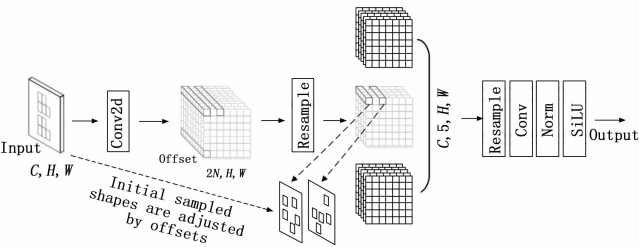


图 6 AKConv 模块

改进后的 C2fAK 模块如图 7 所示, 它通过将 C2f 模块中的 Bottleneck 模块替换为 AKBLOCK 模块来提升性能。在 AKBLOCK 中, 依然采用两个卷积模块, 但与 C2f 中的标准卷积模块不同, 这里使用了 AKConv 模块。这一设计使得 C2fAK 模块能够自适应地选择不同的卷积核, 以处理不同位置上的特征信息。这种灵活性不仅增强了模型的感受野和感知范围, 还使得模型能够捕捉到更多有用的特征。此外, C2fAK 模块还有效地改善了不同层次特征之间的融合效果, 从而增强了低级特征与高级特征之间的结合。通过这种紧密结合, 模型能够更全面地理解输入数据, 尤其是在处理复杂的场景时表现更为出色。在将 C2fAK 模块引入 YOLOv8s 模型后, 该模型能够更好地适应不同形状和大小的水下生物

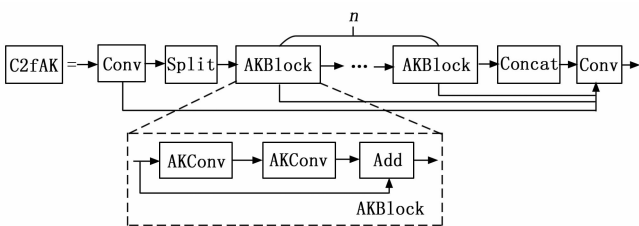


图 7 基于 AKBLOCK 的 C2fAK 模块

目标。这种改进不仅提高了目标检测的检测精度, 还在一定程度上减少了模型的参数量和计算复杂度。

3 实验和实验结果与分析

3.1 数据集构建与分析

本文使用的数据集来自 2020 年全国水下机器人目标检测算法竞赛中的公开数据集 URPC2020。该数据集中的所有图像都是由水下机器人拍摄的, 比赛提供了四类注释数据: 海星 (starfish)、海参 (holothurian)、海胆 (echinus) 和扇贝 (scallop)。URPC2020 数据集包含 5 543 幅图像, 这些图像以 7 : 1 : 2 的比例被随机分成训练集, 验证集和测试集, 使用 3 879 幅图像进行训练, 使用 555 幅图像进行验证, 使用 1 109 幅图像进行测试。该数据集呈现了各种复杂的情况, 如水下生物聚集遮挡、光照不均匀、运动镜头模糊等, 这使其能够真实地表征水下环境, 因此将提高模型的泛化能力。

3.2 实验环境与参数配置

本文中的所有实验都是在 Linux 系统上进行的, 具体的实验环境如表 1 所示。

表 1 实验环境配置

名称	版本
操作系统	Ubuntu18.04
CPU	Intel(R) Xeon(R) Silver 4210R CPU @ 2.40 GHz
GPU	NVIDIA GeForce RTX 3080 Ti
Python	3.8.18
Pytorch	1.12.1
CUDA	11.6
CuDNN	8.4

实验参数配置如表 2 所示。

表 2 实验参数配置

参数名称	参数设置
迭代次数	300
批次大小	16
初始学习率	0.01
最终学习率	0.01
动量因子	0.937
优化器	SGD
权重衰减系数	0.000 5
输入图片尺寸	640×640

3.3 评价指标

本文使用平均精度均值 (mAP, mean average precision)、参数量 (Params)、每秒浮点运算次数 (GFLOPs) 和帧率 (FPS, frames per second) 作为评估模型性能的指标。mAP 是通过精确率 (P, precision) 和召回率 (R, recall) 计算得出的, 用于衡量模型的检测准确性。GFLOPs 用于评估模型或算法的复杂

性,而 Params 则表示模型的规模。通常,参数数量和计算量越小,意味着模型对计算能力的需求越低,对硬件性能的要求也相应降低,这使得在资源有限的设备上部署变得更加可行。FPS 则表示每秒能够处理的帧数,用于评估推理速度。

P 、 R 、 mAP 、 FPS 的计算公式如下:

$$P = \frac{TP}{TP + FP} \quad (1)$$

$$R = \frac{TP}{TP + FN} \quad (2)$$

$$mAP = \frac{1}{K} \sum_{i=1}^K AP_i \quad (3)$$

$$FPS = \frac{FrameNum}{ElapsedTime} \quad (4)$$

其中: TP 为正确检测水下生物某个类的个数, FP 为错误检测为该类的个数, FN 为漏检该类的个数。 AP_i 为第 i 类的平均精确率, K 为待检测水下生物目标的类别数量。 $FrameNum$ 表示模型在测试过程中处理的总帧数,而 $ElapsedTime$ 则表示所需的总时间,以秒为单位。

3.4 不同主干网络的对比实验

为了能够在移动端资源有限的条件下进行水下生物的检测与识别,本文使用一些轻量化网络替换为 YOLOv8s 的主干网络,如 HGNetV2、RepHGNetV2、EfficientViT^[20]、FasterNet^[21] 和 EfficientFormerV2^[22] 网络,并在 URPC2020 数据集上对替换后的各个模型进行了实验。实验结果显示在表 3 中。

表 3 多种主干网络对比结果

Backbone	$mAP@0.5/\%$	Params/M	GFLOPs	FPS
EfficientViT	83.4	8.2	20.4	25.8
FasterNet	82.3	8.1	21.7	67.6
EfficientFormerV2	84.6	9.0	22.6	42.1
HGNetV2	84.5	8.2	23.3	83.7
RepHGNetV2(ours)	85.0	8.1	23.3	87.3

由表 3 可知,相比于其他轻量化网络,如 EfficientViT、FasterNet 和 EfficientFormerV2,虽然 RepHGNetV2 网络在计算量方面稍有逊色,但在 $mAP@0.5$ 和推理速度方面具有极大的优势。相比于 HGNetV2 网络,基于 RepConv 改进的 RepHGNetV2 网络在 $mAP@0.5$ 、参数数量和推理速度方面具有优势。综合来看,将 RepHGNetV2 网络作为 YOLOv8s 的主干网络,在保持模型检测精度的情况下,最大限度的减少参数量,使得该模型更易于在资源有限的移动端进行部署。

3.5 消融实验

为了验证各改进模块对水下生物目标检测的有效

性,本文以 YOLOv8s 为基线,并以平均精度均值、参数量、计算量、FPS 作为评价指标,共设计了多组消融实验,这些实验使用相同的设备、参数设置和数据集进行。实验结果如表 4 所示。

表 4 消融实验结果

模型名称	$mAP@0.5/\%$	Params/M	GFLOPs	FPS
YOLOv8s	85.4	10.6	28.4	94.1
YOLOv8s + RepHGNetV2	85.0	8.1	23.3	87.3
YOLOv8s + GSConv	85.7	10.2	27.9	96.4
YOLOv8s + C2fAK	86.1	9.5	26.1	87.6
YOLOv8s + RepHGNetV2 + GSConv	85.4	7.7	22.8	89.1
GS-YOLOv8(本文)	86.3	6.6	20.5	81.3

如表 4 所示, YOLOv8s + RepHGNetV2 表示通过将 RepHGNetV2 网络替换为 YOLOv8s 模型的主干网络而获得的模型。可以观察到,将 RepHGNetV2 作为 YOLOv8s 的主干网络后可以有效降低模型参数量和计算量,模型参数量和计算量分别比 YOLOv8s 降低了 2.5 M 和 5.1 G,有效地促进后续模型的移动端部署。然而, $mAP@0.5$ 和 FPS 略有降低。 YOLOv8s + GSConv 表示通过将轻量化卷积 GSConv 替换 YOLOv8s 颈部网络中的 Conv 卷积而获得的模型。GSConv 提供了与标准卷积相似的计算效率,同时减少了模型的参数。与 YOLOv8s 相比, YOLOv8s + GSConv 的模型参数量、计算量分别减少了 0.4 M、0.5 G,并将 $mAP@0.5$ 提高了 0.3 个百分点。 YOLOv8s + C2fAK 指将 C2fAK 引入到 YOLOv8s 的颈部。可以看出,加入 C2fAK 模块可以有效提高模型精度, $mAP@0.5$ 比 YOLOv8s 提高了 0.7 个百分点,使模型更加关注图像中的生物目标。模型参数量、计算量分别比 YOLOv8s 减少了 1.1 M、2.3 G。将 RepHGNetV2 作为 YOLOv8s 的主干网络并将轻量化卷积 GSConv 引入到颈部网络后,与 YOLOv8s 相比, $mAP@0.5$ 提高 0.4 个百分点,并且模型参数量、计算量分别减少了 2.9 M、5.6 G。GS-YOLOv8 模型集成了上述改进,在水下生物检测任务中取得了优异的性能。与原始的 YOLOv8s 模型相比,本文改进后的 GS-YOLOv8 模型,虽然 FPS 有略微下降,但是在 $mAP@0.5$ 上提高了 0.9 个百分点。在模型轻量化方面,GS-YOLOv8 模型参数量仅为 6.6 M,计算量仅为 20.5 G,分别下降了 37.7%、27.8%,便于实际应用中的轻量化部署。总的来说,GS-YOLOv8 最大限度地减少了模型参数量和计算量,同时提高了检测精度,证明了本文所提出的算法改进的有效性。

3.6 多种不同算法的对比实验

为了验证本文提出的 GS-YOLOv8 模型的优越性,

以平均精度均值、参数量、计算量、FPS 作为评价指标, 将 GS-YOLOv8 算法与 YOLOv10s、YOLOv8s、YOLOv6s、YOLOv5s、YOLOv4、YOLOv3、YOLOv3-tiny、SSD 和 Faster R-CNN 等经典目标检测算法进行了比较, 这些实验使用相同的设备、参数设置和数据集进行。结果如表 5 所示。

表 5 不同算法性能对比

算法	$mAP@0.5/\%$	Params/M	GFLOPs	FPS
Faster R-CNN	72.9	137.1	370.0	11
SSD	64.4	26.2	62.5	24
YOLOv3	83.6	98.9	282.2	58.2
YOLOv3-tiny	81.6	11.6	20.9	150.9
YOLOv4	75.2	61.4	143.2	68.6
YOLOv5s	84.3	8.7	23.8	86.9
YOLOv6s	84.2	15.5	44.0	123.1
YOLOv8s	85.4	10.6	28.4	94.1
YOLOv10s	85.9	8.0	24.5	116.5
GS-YOLOv8	86.3	6.6	20.5	81.3

根据表 5 的实验结果, 与 Faster R-CNN 算法相比, 本文提出的 GS-YOLOv8 算法在模型参数、计算复杂度、检测精度和推理速度方面具有显著优势。与 SSD 单阶段目标检测算法相比, GS-YOLOv8 算法在推理速度和精度上都有显著的提高。YOLOv3-tiny 虽然实现了轻量级模型, 推理速度较快, 但丢失了大部分检测精度。早期的 YOLO 系列算法, 如 YOLOv3 和 YOLOv4, 结构复杂, 参数数量大, 检测精度相对较低。与之前的 YOLO 版本相比, YOLOv5s 和 YOLOv8s 的模型尺寸更小, 参数更少, 检测性能更好。然而, 这两种模型不能满足不同形状和大小的目标的检测需求, 因此模型检测精度低于本文所提出的 GS-YOLOv8 算法。与 YOLOv6s 检测模型相比, 虽然 GS-YOLOv8 算法的 FPS 有所下降, 但是参数减少了 8.9 M, 计算量降低了 23.5 G, $mAP@0.5$ 提高了 2.1%。与 YOLOv10s 检测模型相比, 虽然 GS-YOLOv8 算法在 FPS 方面略有逊色, 但在模型参数、计算复杂度、检测精度方面具有显著优势, 更易部署在移动端。综上所述, 本文提出的改进算法 GS-YOLOv8 不仅最大限度地减少了模型参数量和计算量, 而且提高了检测精度, 具有较高的实用价值。

3.7 检测结果可视化

为了更加直观地展示出本文提出的改进算法 GS-YOLOv8 在水下场景中的检测效果, 在 URPC2020 数据集的测试集中随机选择了两张不同环境下的图片对模型进行验证。实验中与具有代表性的 YOLOv8s、YOLOv6s、YOLOv5s、YOLOv3、SSD 和 Faster R-CNN 检测算法进行可视化结果比较, 可视化结果如图 8

所示。

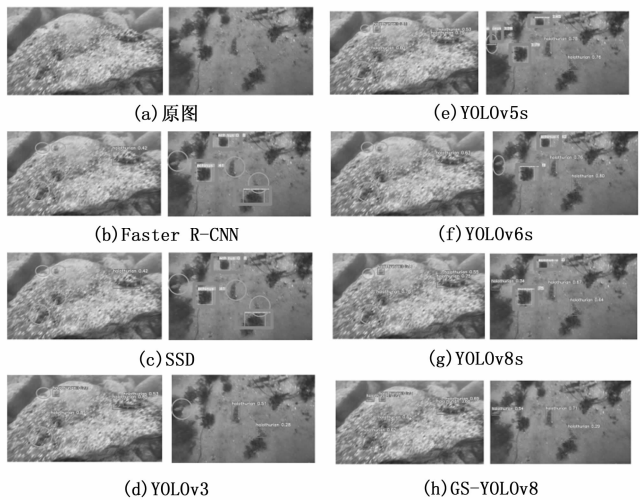


图 8 不同算法的可视化结果

图 8 第一组图片是图像像素模糊情况下进行检测的结果, 图中椭圆表示的地方为漏检的目标。与本文提出的 GS-YOLOv8 网络模型相比, YOLOv8s、YOLOv6s、YOLOv5s、YOLOv3、SSD 和 Faster R-CNN 模型均存在没有将所有海参检测出来的漏检现象。通过观察图 8 第一组图 (b) ~ (g), 发现由于图像模糊, 岩石和海参的颜色相近, 导致海参有未被检测出来的情况。图 8 的第二组图片检测的是背景比较复杂的场景, 图中椭圆表示的地方为漏检的目标, 加粗框表示的地方为误检的目标。Faster R-CNN、SSD、YOLOv5s、YOLOv6s 模型均存在将海草误以为海参的误检现象, 未检测出海草丛中的海参的漏检现象。YOLOv3 存在漏检现象, 未检测出海草丛中的海参。YOLOv8s 将海草误以为海参产生了误检。海草和海参在某些情况下可能具有相似的形状或颜色, 尤其是在复杂的水下环境中, 不容易辨别, 导致模型在特征提取时产生混淆。海草丛中的海参部分被海草挡住了, 且在图中较远的位置, 未被检测出来。本文提出的 GS-YOLOv8 算法通过引入 GSConv 卷积和 C2fAK 模块, 大大降低了漏检和误检的发生。总的来说, 本文提出的 GS-YOLOv8 算法相对于主流的目标检测算法表现出更好的性能, 一定程度上改善了漏检和误检问题, 能够在移动端资源有限情况下应用于复杂的水下环境。

4 结束语

本文提出了一种轻量化的 GS-YOLOv8 目标检测算法, 旨在解决现有的水下生物目标检测模型参数过多, 难以部署到资源有限的移动端的问题。该模型将设计的轻量化网络 RepHGNetV2 作为主干网络, 有效降低了模型参数量和计算量; 同时用轻量化的 GSConv 代替颈部网络中 Conv, 来保持整体精度, 进一步减少模型的

参数;引入基于 AKConv 模块的 C2fAK 模块,提高了目标检测的检测精度。本文在 URPC2020 数据集上进行了实验,改进的 GS-YOLOv8 模型与原模型 YOLOv8s 相比,参数量降低了 37.7%,计算量降低了 27.8%, $mAP@0.5$ 提高了 0.9%,大大减少了漏检和误检。为了验证其有效性,与该领域领先的目标检测算法进行了比较分析,证实了其优越的检测精度、轻量化和整体性能,更能适应复杂的水下环境。尽管本文提出的算法取得了令人满意的结果,但仍存在改进的空间。由于数据集的目标非常分散,使得模型无法使所有目标的检测精度同时达到峰值。该模型还可以通过数据集改进得到进一步的改进。因此,为了进一步提高水下生物目标检测的精度,需要对改进数据集进行深入的研究。

参考文献:

- [1] FU H, SONG G, WANG Y. Improved YOLOv4 marine target detection combined with CBAM [J]. Symmetry, 2021, 13 (4): 623.
- [2] LOWE D G. Distinctive Image features from scaleinvariant key points [J]. International Journal of Computer Vision, 2004, 60 (2): 91–110.
- [3] KUANG H L, CHAN L H, YAN H. Multiclass fruit detection based on multiple color channels [C] //International Conference on Wavelet Analysis and Pattern Recognition, 2015: 1–7.
- [4] KOKLU M, NLEREN M F, OZKAN I A, et al. A CNN-SVM Study based on selected deep features for grapevine leaves classification [J]. Measurement, 2022, 188 (110425): 1–10.
- [5] 袁明阳, 宋亚林, 张 潮, 等. 基于 GA-RetinaNet 的水下目标检测 [J]. 计算机系统应用, 2023, 32 (6): 80–90.
- [6] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation [C] //Proceedings of the IEEE Conference on Computervision and Pattern Recognition, 2014: 580–587.
- [7] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: towards realtime object detection with region proposal networks [J]. Advances in Neural Information Processing Systems, 2015, 28: 91–99.
- [8] ZHOU H, HUANG H, YANG X, et al. Faster R-CNN for marine organism detection and recognition using data augmentation [C] //Proceedings of the International Conference on Video and Image Processing, 2017: 56–62.
- [9] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot multibox detector [C] //European Conference on Computer Vision, 2016: 21–37.
- [10] LIU Y, CHU H, SONG L, et al. An improved Tuna-YOLO model based on YOLO v3 for Real-Time tuna detection considering lightweight deployment [J]. Journal of Marine Science and Engineering, 2023, 11 (3): 542.
- [11] LI Y, LIU W, LI L, et al. Vision-based target detection and positioning approach for underwater robots [J]. IEEE Photonics Journal, 2022, 15 (1): 1–12.
- [12] SHI X, WANG H. Improved lightweight underwater target detection network based on YOLOV4 (you only look once v4) [J]. J. Harbin Eng. Univ, 2023, 44: 154–160.
- [13] LIANG H, SONG T. Lightweight marine biological target detection algorithm based on YOLOv5 [J]. Frontiers in Marine Science, 2023, 10: 1219155.
- [14] 赵 晨, 陈 明. 水下底栖生物轻量化目标检测算法 YOLOv7-RFPCW [J]. 农业工程学报, 2024, 40 (11): 168–177.
- [15] LI H, LI J, WEI H, et al. Slim-neck by GSConv: a lightweight-design for real-time detector architectures [J]. Journal of Real-Time Image Processing, 2024, 21 (3): 62.
- [16] LIN T Y, DOLLÁR P, GIRSHICK R, et al. Feature pyramid networks for object detection [C] //Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017: 2117–2125.
- [17] ZHAO Y, LV W, XU S, et al. Detrs beat yolos on real-time object detection [C] //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2024: 16965–16974.
- [18] SOUDY M, AFIFY Y, BADR N. RepConv: a novel architecture for image scene classification on Intel scenes dataset [J]. International Journal of Intelligent Computing and Information Sciences, 2022, 22 (2): 63–73.
- [19] HANG X, SONG Y, SONG T, et al. AKConv: convolutional kernel with arbitrary sampled shapes and arbitrary number of parameters [J]. ArXiv Preprint ArXiv: 2311.11587, 2023.
- [20] LIU X, PENG H, ZHENG N, et al. EfficientViT: memory efficient vision transformer with cascaded group attention [C] //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023: 14420–14430.
- [21] CHEN J, KAO S, HE H, et al. Run, don't walk: chasing higher FLOPS for faster neural networks [C] //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023: 12021–12031.
- [22] LI Y, HU J, WEN Y, et al. Rethinking vision transformers for mobilenet size and speed [C] //Proceedings of the IEEE/CVF International Conference on Computer Vision, 2023: 16889–16900.