

基于改进YOLOv7的海滩垃圾检测方法

朱道恒¹, 朱滄灵², 李志强¹, 刘润³

(1.广东海洋大学 电子与信息工程学院, 广东 湛江 524088; 2.贵州大学 大数据与信息工程学院, 贵州 贵阳 550025; 3.广东海洋大学 化学与环境学院, 广东 湛江 524088)

摘要: 海滩垃圾污染对海洋生态系统和人类健康构成巨大威胁, 海滩垃圾监测和清理是一项繁重且复杂的任务。传统的人工调查存在监测效率低、检测范围小和时效性差等问题, 为了解决这些问题, 本文提出一种基于改进 YOLOv7 的海滩垃圾检测方法。首先, YOLOv7 与量化感知 RepVGG (QARepVGG) 相结合, 实现快速计算并降低模型参数量。其次, 加入简单注意力机制 (simple attention mechanism, SimAM), 增强网络对图像感兴趣区域的特征提取能力。最后, 结合双向特征金字塔网络 (bi-directional feature pyramid network, BiFPN) 结构, 改进原有路径聚合网络 (path aggregation network, PAN), 提高网络学习垃圾特征的效率, 增强对不同尺寸垃圾的识别能力。在自建数据集上的实验结果表明: (1) 改进模型对 8 类海滩垃圾有良好的检测能力; (2) 与 YOLOv7 相比, 改进模型的总体平均精度均值 (mean average precision, *mAP*) 提升 5.8%, 每秒传输帧数 (frames per second, *FPS*) 提高 17, 改进模型对泡沫、塑料类和纸制品垃圾的识别精度最高; (3) 与几种流行的检测模型相比, 改进模型的识别精度和效率最高。实际场景中的检测结果表明, 改进模型能满足海滩垃圾实时性检测需求。

关键词: 海滩垃圾; 环境污染; 目标检测; QARepVGG; 注意力机制

中图分类号: X830.2

文献标识码: A

文章编号: 1007-6336(2024)05-0684-12

A beach litter detection method based on improved YOLOv7

ZHU Daoheng¹, ZHU Hanling², LI Zhiqiang¹, LIU Run³

(1.School of Electronics and Information Engineering, Guangdong Ocean University, Zhanjiang 524088, China; 2.School of Big Data and Information Engineering, Guizhou University, Guiyang 550025, China; 3.School of Chemistry and Environment, Guangdong Ocean University, Zhanjiang 524088, China)

Abstract: Beach litter monitoring and cleaning is a difficult and involved undertaking since beach litter contamination poses a serious hazard to human health and marine ecosystems. Due to the low monitoring efficiency, limited detection range, and unreliable timeliness of typical human surveys, this work proposes a beach litter detection method based on improved YOLOv7. First, YOLOv7 is initially merged with Quantitative awareness RepVGG (QARepVGG) to achieve rapid computation and reduce the number of model parameters. Second, the SimAM attention mechanism is included to improve the network's capacity for feature extraction from the targeted region of the picture. To further enhance the original path aggregation network (PAN), the Bi-directional feature pyramid network (BiFPN) structure is merged with it. As a result, the network is more effectively able to identify targets of various sizes and recognize the characteristics of different litter sizes.

收稿日期: 2023-08-31, 修订日期: 2023-11-14

基金项目: 国家自然科学基金项目 (42176167); 广东海洋大学科研经费启动项目 (060302112317)

作者简介: 朱道恒 (1992—), 男, 湖北孝感人, 博士, 主要研究方向为海岸带与海洋监测技术, E-mail: dhzhu911@163.com

通信作者: 李志强 (1974—), 男, 广东湛江人, 博士, 主要研究方向为海滩过程与海洋环境, E-mail: qiangzli1974@163.com

Experimental findings on the self-constructed dataset demonstrate that the modified model has strong identification capability for eight different categories of beach litter. The revised model outperforms YOLOv7 in terms of mean average precision (*mAP*) improvement by 5.8%, frame per second (*FPS*) improvement by 17, and the improved model has the highest recognition accuracy for styrofoam, plastic, and paper. When compared to a number of widely used detection models, our improved model has the best performance. These detection results in real scenarios show that the improved model is capable of detecting beach litter in real time.

Key words: beach litter; environmental pollution; object detection; QARepVGG; attention mechanism

海洋环境污染问题持续影响着人类和海洋生物的生存和健康,其中海洋垃圾是重要的污染源。海洋垃圾的来源主要分为陆源和海源^[1],陆源垃圾来自沿海地区的生产生活、海岸开发和河流汇入等^[2],海源垃圾来自海洋资源开发、渔业捕捞和海上货运等^[3]。海岸带作为陆海连接的重要地带,是人类生存和发展的重要场所。海滩垃圾影响海岸带环境和海洋生态平衡,如何有效治理海滩垃圾是许多国家关注的焦点。海岸垃圾具有分布广、持久存在、累积性污染等特点,包括塑料、泡沫、金属、木材、纸张、橡胶和纺织物等^[4]。海岸垃圾含有多种难降解的有害物质,留存在海滩上会污染海岸环境、降低海滩吸引力等,而被海风、洋流和潮汐带入海洋后则可能阻碍海上交通、破坏船只、导致海洋生物受伤或死亡等^[5]。我国高度重视垃圾污染问题,已制定了相关法律禁止进口大多数塑料垃圾^[6],同时积极推动海岸清洁和垃圾监测技术相关的研究,以改善垃圾对海岸环境的影响,提升海滩质量^[7]。

国内外学者针对海岸垃圾的监测、分布和识别分类开展了许多探索工作。在垃圾监测方面, Martin 等^[8]结合人工调查和无人机远程调查的方式量化了海滩垃圾的丰度,但海滩覆盖率不足导致分类结果不理想。Fallati 等^[9]基于无人机图像和深度学习技术实现了马尔代夫旅游海滩垃圾的自动检测和量化,与人工调查结果对比有较高的准确率,但检测效率不高且图像质量受环境因素影响很大。Nakashima 等^[10]使用气球辅助航空摄影结合原位测量来估计海滩上大型垃圾的数量,由于不同位置的垃圾密度差异较大,因此垃圾数量在评估上存在很多误差。在垃圾分布调查方面, Kako 等^[11]远程控制悬挂在气球上的数码相机,建立了一个低空遥感系统,结合投影变换法和均匀颜色空间中的色差提取不同

颜色垃圾的像素值,从而估计海滩或海洋垃圾的分布,该方法没有定位和识别垃圾,检测效率不高且耗费人力物力。Bao 等^[12]将阈值滤波法用于无人机拍摄的图像来监测海滩垃圾的分布情况,虽然无人机遥感技术能提高垃圾监测效率,但不同颜色物体的光谱部分重叠,导致较高的误分率。邱琰皓等^[13]采用无人机低空摄影的方法监测青岛雀石滩的垃圾的空间分布,该方法有操作简单、效率高、成本低的优势,但主要依赖垃圾的大小和颜色来区分类别,在背景复杂海滩上的适用性可能不足。无人机遥感和机器学习方法相结合是海岸垃圾分类和监测的新方法,对探究大型垃圾的分布有很好的前景^[14],但对于大范围实施和小目标检测仍存在局限^[15]。在垃圾的识别方面,激光雷达技术不同于成像技术和投影变换的方法,能够实时获取垃圾的三维空间信息和特性,而且能全天候、大范围地作业,为海滩垃圾的快速识别提供新途径^[16]。深度学习的方法应用到垃圾检测研究中,能够实现各类海滩垃圾的精确和快速识别。例如, Pfeiffer 等^[17]对比了 YOLOv5 和 Faster R-CNN 两种深度学习算法在无人机航拍的垃圾图像上的检测性能,发现 YOLOv5 的 *mAP* 值(54.2%)高于 Faster R-CNN (32.8%),虽然该研究表明单阶段算法比双阶段算法在海滩垃圾检测方面更具优势,但该数据集全是无人机图像,且将所有垃圾归为一类而没有细分类型。Song 等^[18]基于 YOLOv5 检测相机拍摄的海滩垃圾图像,识别精度总体达到 87%,并将垃圾分为 7 类,每个类别都取得良好检测效果。虽然其测试结果中存在漏检和误检,但验证了模型的性能受垃圾形态特征、参数设置和训练数据量影响较大,同时表明 YOLO 模型在海岸垃圾检测方面具有巨大潜力和应用价值。

基于前人的研究,传统的海滩垃圾监测和清

洁方式是人工调查、识检和分类回收,消耗大量人力、物力且效率较低,因此改变监测和清洁方式以提升效率非常重要。自动化技术和深度学习方法是解决海岸垃圾监测和污染问题的新途径^[19]。当前在实际场景中的海滩垃圾检测存在两个主要问题:一是检测范围广,多数深度学习

模型的参数量较大导致检测速度慢而无法满足应用需求;二是在复杂的海滩背景下存在较多干扰物和目标遮挡的问题(图1),同时不同类型垃圾的尺寸差距较大,导致模型的检测精度难以提升。为了解决上述问题,本文提出改进的YOLOv7模型用于识别海滩垃圾并进行分类。

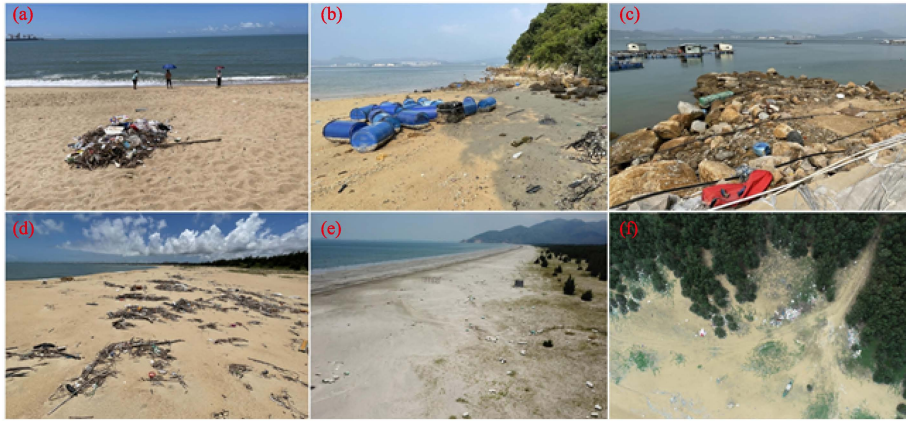


图1 华南沿海不同类型海滩上存在大量的垃圾

Fig. 1 Large amount of litter on different types of beaches along the coast of South China

1 模型与方法

1.1 QARepVGG 结构

RepVGG^[20]中自定义的权重衰减设计构建了权重分布更强的模型,但激活分布的方差被放大,下一层的输入依赖于激活,标准偏差随网络层深度的增加逐层递增,从而降低模型的精度。量化感知 RepVGG(quantization-aware RepVGG, QARepVGG)^[21]改进了 RepVGG,基于结构重参数化思想,在训练时尽量用多分支结构来提升网络性能,而推理时将其变为单路结构,不仅减小模型的参数量,而且提升了模型的推理速度。为了提升YOLOv7模型检测海滩垃圾的效率,本研究将QARepVGG引入该模型。RepVGG和QARepVGG的网络结构如图2所示。

QARepVGG引入均方误差(mean-square error, MSE)进行误差量化,见式(1):

$$MSE[Q(w, t, n_b), w] = \frac{1}{n} \sum_i [Q(w_i, t, n_b) - w]^2 \quad (1)$$

式中: Q 为量化过程; $w \in R^n$ 为权值的第 n 通道; t 为阶段阈值; n_b 为量化比特数。输出见式(2):

$$M_{(2)} = BN(3 \times 3) + BN(1 \times 1) + BN(Identity) \quad (2)$$

式中: $M_{(2)}$ 表示输出的特征图; $BN(3 \times 3)$ 表示通道数为 3×3 的批量样本归一化(batch normalization, BN)层,最后输出见式(3):

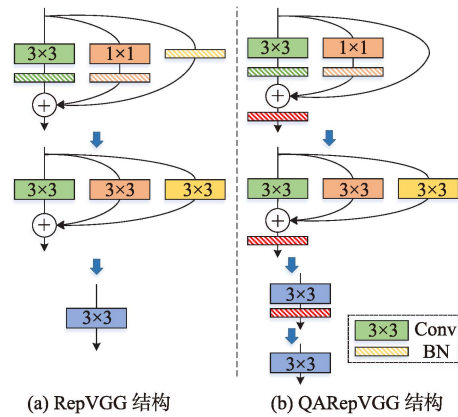


图2 RepVGG 和 QARepVGG 的网络结构

Fig. 2 Network Architecture of RepVGG and QARepVGG

$$Y_{(3)} = \gamma_{(3)} \times \frac{M_{(1)} \times W_{(3)} - \mu_{(3)}}{\sqrt{\varepsilon + \sigma_{(3)} \times \sigma_{(3)}}} + \beta_{(3)} \quad (3)$$

式中: $Y_{(3)}$ 为 3×3 分支的输出; $\gamma_{(3)}$ 、 $\mu_{(3)}$ 、 $\sigma_{(3)}$ 和 $\beta_{(3)}$ 分别是 3×3 卷积后 BN 层的比例因子、均值、标准差和偏差; $M_{(1)}$ 为输入; $W_{(3)}$ 为 3×3 的卷积核; ε 是定值(默认取 10^{-5})。引入随机变量

X 和标量 λ , 由 $D(\lambda X) = \lambda^2 D(X)$, 令 $X_{(3)} = M_{(1)} W_{(3)}$, 则输出方差见式(4):

$$D(Y_{(3)}) = \frac{\gamma_{(3)} \times \gamma_{(3)}}{\varepsilon + \sigma_{(3)} \times \sigma_{(3)}} \times D(X_{(3)}) \quad (4)$$

式中: $\frac{\gamma_{(3)} * \gamma_{(3)}}{\varepsilon + \sigma_{(3)} * \sigma_{(3)}}$ 控制 $X_{(3)}$ 的方差大小, 基于重新参数化的架构有权重分布和激活分布两个组件需要量化。自定义权重衰减 (L_2) 的计算见式(5):

$$L_2 = \frac{|W_{eq}|_2^2}{\left| \frac{\gamma_{(3)}}{\sqrt{\varepsilon + \sigma_{(3)} \times \sigma_{(3)}}} \right|_2^2 + \left| \frac{\gamma_{(1)}}{\sqrt{\varepsilon + \sigma_{(1)} \times \sigma_{(1)}}} \right|_2^2} \quad (5)$$

式中: W_{eq} 为多分支权值。RepVGG 通过放大分母减少权重损失, 但扩大了标准偏差分布, 放大激活分布偏差。QARepVGG 用标准权重衰减设计替换 RepVGG 中的自定义权重衰减设计, 同时删除 Identity 分支中的 BN 和 1×1 分支中的 BN, 使模块成功进行权重量化。为稳定训练过程, QARepVGG 在 3 个分支后添加批量归一化, 最后输出见式(6):

$$M_{(2)} = BN(BN(3 \times 3) + (1 \times 1) + Identity) \quad (6)$$

训练时, 模型主干网络的 QARepVGG 结构

和 Head 中的部分 E-ELAN 模块被 QARepVGG 结构替换。由于 QARepVGG 较 RepVGG 具有更多的层和参数, 若 Backbone 和 Head 中的 E-ELAN 全部替换为 QARepVGG, 则需要更强大的计算设备来训练模型。权衡模型的计算效率和性能, 考虑到本文设备和资源有限, 因此保留了 Head 中其他的 E-ELAN 模块。当无 BN 和有 BN 时, 卷积的参数量计算分别见式(7)和式(8):

$$params = C_o \times C_{in} \times k_h \times k_w + C_o \quad (7)$$

$$params = C_o \times C_{in} \times k_h \times k_w + C_o \times 2 \quad (8)$$

式中: C_o 和 C_{in} 分别为输出和输入通道数; k_h 和 k_w 分别是卷积核的高度和宽度。主干网络中 E-ELAN 的参数量设为 $params_1$ 。替换后主干网络中 QARepVGG 的参数量设为 $params_2$, QARepVGG 中 3×3 卷积带有 BN 操作, 且在末端总体新增一个 BN 层。Head 部分被 QARepVGG 替换的 E-ELAN 的参数量设为 $params_3$, 替换的 QARepVGG 参数量为 $params_4$ 。YOLOv7 中 3 个卷积的参数在改进前后记为 $params_5$ 和 $params_6$, 另两个卷积的参数在改进前后记为 $params_7$ 和 $params_8$ 。由表 1 可知, 修改后的参数量明显降低。

表 1 模型修改前后的参数量比较

Tab.1 Comparison of parameter counts before and after model modification

模型	参数	值
YOLOv7	$params_1 + params_3 + params_5 \times 3 + params_7 \times 2$	1857024
YOLOv7+QARepVGG	$params_2 + params_4 + params_6 \times 3 + params_8 \times 2$	1067776

1.2 S-MP 结构

一些图像存在干扰物造成背景复杂、遮挡目标和目标模糊的问题, 从而产生漏检和误检, 本文加入注意力机制来减弱这些影响。

传统注意力机制沿通道维度或空间维度生成一维或二维权重, 只关注本通道或空间, 限制了学习注意力权重在通道或空间中变化的灵活性。简单注意力机制 (simple attention mechanism, SimAM) 是一种三维注意力机制^[22], 在不增加参数量的同时给每一个空间或通道中的神经元分配一个唯一的权重, 并使用三维权重对特征进行增强。在 YOLOv7 模型 Head 部分的下采样模块 (MP) 引入 SimAM 替换第二条分支的 BConv,

组成 S-MP。S-MP 的输入在第一条分支经 Maxpool 进行下采样, 再通过 BConv 调整通道数。第二条分支在经过 SimAM 调整通道数的同时, 对特征进行增强, 再通过 BConv 进行下采样。最后融合两个分支得到增强后的下采样结果。

SimAM 为每个神经元定义了能量函数, 见式(9):

$$e_i(w_t, b_t, y, x_i) = (y_i - \hat{t})^2 + \frac{1}{M-1} \sum_{i=1}^{M-1} (y_0 - \hat{x}_i)^2 \quad (9)$$

式中: t 为目标神经元; x_i 是与输入特征同一通道的其他神经元; $\hat{t} = w_t t + b_t$, $\hat{x}_i = w_t x_i + b_t$, 分别是 t 和 x_i 的线性变换, i 是空间维度上的指数;

$M = H \times W$ 是通道上的神经元数量; w_t 和 b_t 分别是变换的权值和偏置, 计算公式见式(10)、式(11)。当式(9)达到最小值时, $\hat{t} = y_t$, 任意的 $\hat{x}_i = y_0$ 。将 y_t 和 y_0 二进制标签化, 并在式(9)加入正则化得到式(12):

$$e_t(w_t, b_t, y, x_i) = \frac{1}{M-1} \sum_{i=1}^{M-1} [-1 - (w_t x_i + b_t)]^2 + [1 - (w_t t + b_t)]^2 + \lambda \omega_t^2 \quad (10)$$

$$w_t = -\frac{2(t - \mu_t)}{(t - \mu_t)^2 + 2\sigma_t^2 + 2\lambda} \quad (11)$$

$$b_t = -\frac{1}{2}(t + \mu_t)w_t \quad (12)$$

式中: $\mu_t = \frac{1}{M-1} \sum_{i=1}^{M-1} x_i$ 和 $\sigma_t^2 = \frac{1}{M-1} \sum_{i=1}^{M-1} (x_i - \mu_t)^2$ 分别指计算该通道除 t 以外的所有神经元的均值与方差。最小能量的计算见式(13)至式(15):

$$e_t^* = \frac{4(\hat{\sigma}^2 + \lambda)}{(t - \hat{\mu})^2 + 2\hat{\sigma}^2 + 2\lambda} \quad (13)$$

$$\hat{\mu} = \frac{1}{M} \sum_{i=1}^M x_i \quad (14)$$

$$\hat{\sigma}^2 = \frac{1}{M} \sum_{i=1}^M (x_i - \hat{\mu})^2 \quad (15)$$

式中: $\frac{1}{e_t^*}$ 表示每个神经元的重要性, 即 e_t^* 值越小, 目标神经元与其他神经元的区别越大。根据注意力机制进行特征增强, 见式(16):

$$\tilde{X} = \text{sigmoid}\left(\frac{1}{E}\right) \times X \quad (16)$$

在空间和通道维度上, E 对所有的 e_t^* 分组。

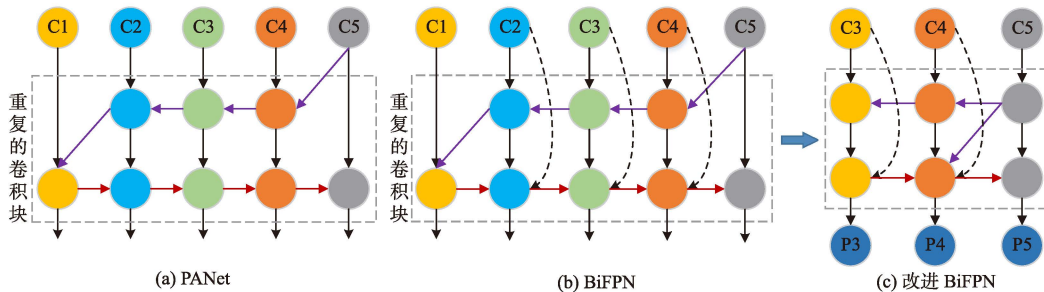


图3 简化的 PANet 结构、原始 BiFPN 结构和改进后的 BiFPN 结构

Fig. 3 Simplified PANet, original BiFPN and modified BiFPN structure

1.4 改进的 YOLOv7 模型

综合前面提出的改进方法, 得到改进 YOLOv7 模型的整体结构, 如图 4 所示。改进的

Sigmoid 能避免 E 值过大, 而不影响每个神经元的相关性。

1.3 改进 BiFPN 结构

传统的特征金字塔网络 (feature pyramid network, FPN)^[23] 以自上而下的方式融合多尺度特征, 下采样过程会损失最高层的特征信息, 降低多尺度的表示能力。YOLOv7^[24] 在网络中加入了一个自底向上的路径聚合网络, 组成高效聚合网络 (efficient long-range attention network, ELAN)。在卷积过程中, 卷积核只建模领域像素间的关系, 不易捕获长距离的依赖关系, 导致小目标特征丢失而产生漏检现象。

BiFPN 是一种特征融合方法, 包括双向跨尺度连接和加权特征融合, 能在不增加太多成本的情况下融合更多特征, 同时将每个双向路径视为一个特征网络层, 并多次重复同一层, 以实现更高层次的特征融合。在自建数据集中, 不同类型的海滩垃圾尺寸大小不一, 为加强不同尺度特征的提取能力, 提升小目标的检测效率, 将路径聚合的特征金字塔网络 (path aggregation network with feature pyramid network, PAFPN) 结构替换为改进的 BiFPN, 从而实现不同网络特征层的双向融合, 增强不同网络特征层之间的信息传递。如图 3 所示, 主干网络提取到的 3 个有效特征层分别为 C3、C4 和 C5, 将它们分别传入改进后的 BiFPN, 实现有效的双向跨尺度连接和加权特征融合, 最终得到 3 种不同尺度的输出特征 (P3、P4 和 P5)。

模型将 YOLOv7 主干部分的第一个 ELAN 模块和 Head 部分的 E-ELAN 模块替换为量化感知的 QARepVGG 结构, 将 Head 中 MP 模块替换为

融入了 SimAM 注意力机制的 S-MP 模块, 并基于 *BiFPN* 的思想改进了具有特征金字塔网络的路径聚合网络。

首先通过 input 将垃圾图像大小调整为 640×640 , 然后输入调整后的图像到 Backbone 中

提取特征, 得到 3 个有效特征层 (C3、C4 和 C5), 再输入 Head 中加强网络的特征提取, 输出结果经过重参数卷积 (re-parameterized convolution, *RepConv*) 调整通道数, 预测海滩垃圾的锚框、置信度和类别, 最后输出检测结果。

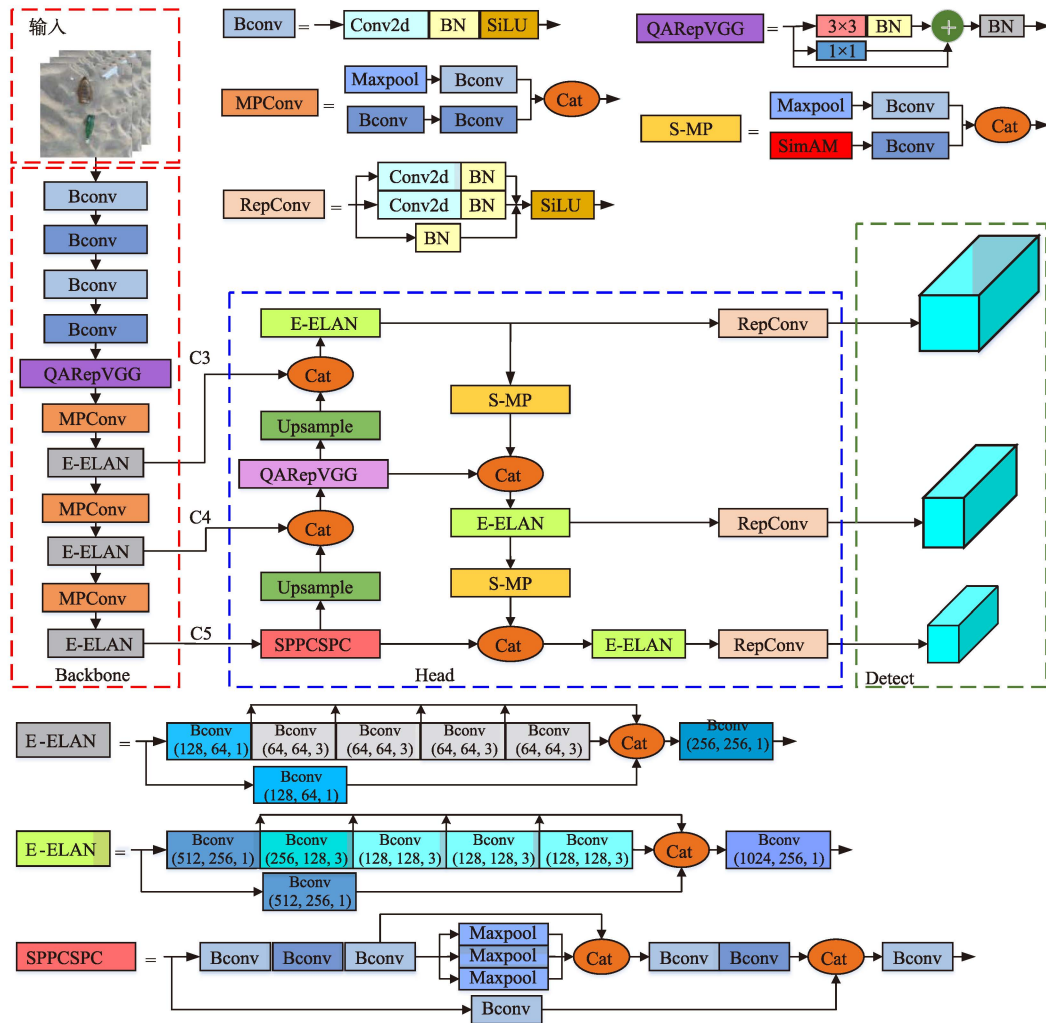


图4 改进 YOLOv7 的网络结构

Fig. 4 Improved YOLOv7 Network Architecture

2 数据和实验环境

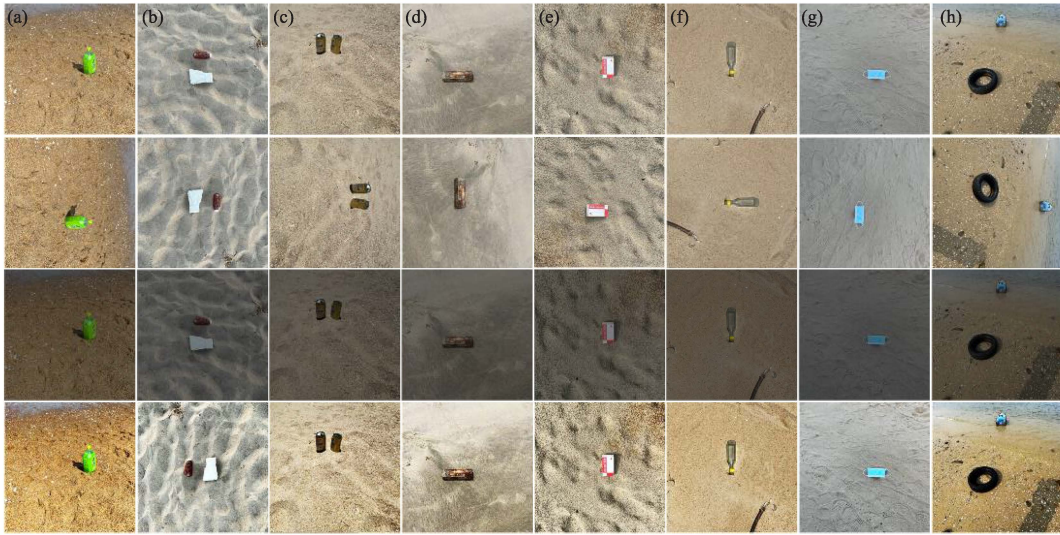
2.1 数据集

本文使用的海滩垃圾图像来自华南沿海多个海滩, 其中汕头海门湾、龙虎滩; 汕尾碣石湾、红海湾和鲗门湾; 惠州双月湾; 阳江十里银滩和湛江龙海天沙滩是砂质海岸, 深圳龙岐湾是人工海岸, 深圳大鹏湾是基岩海岸。在海滩调查过程

中采用相机拍照和无人机航拍的方式获得 1587 张含有垃圾的海滩图像, 分辨率从 4000×2250 像素到 480×360 像素不等。根据图像中垃圾的特征将所有垃圾划分为 8 个类别: 塑料、金属、纸制品、木材、泡沫、玻璃、织物类和橡胶。采用 Labelimg 软件标注出图像中垃圾的位置。为防止在训练过程中因样本量不足而产生过拟合现象, 另外对每张图像进行数据增强操作(包括图

片旋转、颜色增强和对比度增强等)。图片旋转是将图像绕中心点旋转一定角度,本文选择统一顺时针旋转 90° 。颜色增强是通过改变图像的颜色属性来增加多样性,包括亮度、饱和度、对比度等方面的调整,本文统一调整亮度为原始的20%。对比度增强是增加图像中颜色之间的差

异程度,可通过调整像素值来实现,本文统一调整图像对比度至原来的2倍。不同类型的垃圾经数据增强后的图像示例如图5所示。经数据增强后的数据集共有6348张,再经resize操作将图像大小调整为 640×640 ,并按9:1的比例划分为训练集(5713张)和测试集(635张)。



(a)塑料, (b)泡沫和塑料, (c)金属, (d)木材, (e)纸制品, (f)玻璃和木材, (g)织物, (h)橡胶和塑料;第一行是原始图像,第二行至第四行是经旋转、亮度变暗、对比度增强操作后的图像

图5 不同垃圾类型和数据增强后的图像示例

Fig. 5 Examples of different types of beach litter and data enhanced images

2.2 实验环境和参数

实验使用的硬件环境为 Intel(R) Core(TM) i5-10400KF CPU@ 2.90 GHz, 内存 16 G, GPU 采用 NVIDIA GeForce RTX 3090Ti 的 12 G 显卡。软件环境为 Ubuntu18.04 操作系统, 深度学习框架为 PyTorch 1.8.0, Python 版本为 3.7, CUDA 版本为 10.2, Torchvision 版本为 0.7.0。模型输入图像尺寸为 640×640 , 通道数为 3。模型训练时设置动量系数为 0.937, 权重衰减系数为 0.0005, 学习率为 0.01, 批尺寸为 8, 训练的迭代次数为 300。

2.3 评价指标

使用平均精确度 (average precision, AP)、 mAP 、模型参数量 (params)、FPS 和 $F1$ 值来评估模型的检测效果^[25]。 AP 用来衡量单个海滩垃圾类别的平均检测精度, 其值为精确率-召回率 (P - R) 曲线的面积。 P - R 曲线的纵坐标为精确率

(precision, P), 横坐标为召回率 (recall, R), P 和 R 的计算分别采用式(17)和式(18)。 mAP 用于衡量模型所有类别的平均检测精度, 模型参数量用来衡量模型的大小。 FPS 用于衡量模型的检测速度。 $F1$ 是 P 和 R 的调和均值, 用于表征被召回的真阳性类别的比例。 AP 、 mAP 和 $F1$ 值的计算分别采用式(19)、式(20)和式(21)。

$$P = \frac{TP}{TP + FP} \quad (17)$$

$$R = \frac{TP}{TP + FN} \quad (18)$$

$$AP = \int_0^1 P(R) dR \quad (19)$$

$$mAP = \frac{\sum_{i=1}^N AP_i}{N} \quad (20)$$

$$F1 = \frac{2PR}{(P + R)} \quad (21)$$

式中: TP 表示正确检测出目标的数量; FP 表示错误检测为目标的数量; FN 表示误检或

漏检目标的数量; N 表示目标的类别数。通常 mAP 值越高代表模型检测效果越好, FPS 越大代表检测速度越快, $F1$ 值越高说明模型性能越好。

3 实验结果

3.1 结合 QARepVGG 对比实验

分别用 YOLOv7、YOLOv7+RepVGG 和 YOLOv7+QARepVGG 模型对自建的海滩垃圾数据集进行训练和测试, 对比结果列于表 2。

RepVGG 结构重参数化架构在训练时的

1×1 分支比 QARepVGG 结构的 1×1 分支多一个 BN 操作, 造成前者的总体参数量比后者大 256。在 3 种不同模型中, YOLOv7+QARepVGG 的推理时间最短 (8.0 s), 检测速度最快。虽然 RepVGG 结构能缩短推理时间, 提升检测速度, 但牺牲了检测精度。QARepVGG 结构不仅缩短了推理时间, 而且提升了检测速度, 同时降低了模型的参数量, 最终在检测精度上提升了 1.7%。综合比较后选择 QARepVGG 模块加入 YOLOv7 网络。

表 2 不同模型的实验结果比较

Tab.2 Comparison of experimental results of different models

模型	参数量	推理时间/ms	mAP@0.5/(%)	F1	FPS
YOLOv7	37223526	11.8	79.0	0.73	40
YOLOv7+RepVGG	36434534	8.4	77.7	0.71	56
YOLOv7+QARepVGG	36434278	8.0	80.7	0.78	58

3.2 不同注意力机制对比实验

在 Head 部分的 MP 模块中分别结合卷积块注意力模块 (Convolutional block attention module, CBAM)^[26]、高效通道注意力模块 (Efficient channel attention module, ECA)^[27] 和 SimAM 3 种注意力机制, 并与原 YOLOv7 模型在训练过程中的参数量和检测效果进行对比, 结果见表 3。其中 C-MP 和 E-MP 分别表示使用 CBAM 和 ECA 替换原始 Head 中 MP 第二条分支的第一个卷积。

表 3 加入 3 种不同注意力机制的检测结果比较

Tab.3 Comparison of detection results with the addition of three different attention mechanisms

模型	参数量	mAP@0.5/(%)	F1
YOLOv7	37223526	79.0	0.73
YOLOv7+C-MP	37151684	74.8	0.70
YOLOv7+E-MP	37140838	68.5	0.65
YOLOv7+S-MP	37140838	82.9	0.77

CBAM 为混合注意力机制, YOLOv7+C-MP 模型的检测精度比 YOLOv7 降低 3.2%, $F1$ 降低 0.03, 但参数量降低较少。ECA 和 SimAM 为无参数注意力机制, 替换相同卷积后模型的参数量不变, 但加入 ECA 后, mAP 值比 YOLOv7 减小 10.5%。CBAM 与 ECA 均降低了垃圾的检测精

度, 而加入 SimAM 后, 检测精度提升 3.9%, $F1$ 提升 0.04, 因此选择引入 SimAM 模块。

3.3 加强特征提取网络对比实验

为增强网络在实际场景下对不同尺寸目标的检测能力, 本研究用改进的 BiFPN 替换 Head 部分的 PAFPN。改进的 BiFPN 结构增强了不同尺寸目标特征的提取能力, 最终得到优于原 YOLOv7 模型的检测效果。从表 4 中的结果可知, 加入改进的 BiFPN 结构后, mAP 值提高 4.1%, $F1$ 值增大 0.05。虽然模型的参数量略有增长, 但检测速度没有变化, 因此证明改进的 BiFPN 结构能明显提升模型的检测精度而不影响检测速度。

表 4 加强特征提取网络的检测结果对比

Tab.4 Comparison of detection results of enhanced feature extraction networks

模型	参数量	mAP@0.5/(%)	F1	FPS
YOLOv7	37223526	79.0	0.73	40
YOLOv7+BiFPN	37368452	81.6	0.75	40
YOLOv7+改进BiFPN	37354598	83.1	0.78	40

3.4 垃圾分类结果

在不同检测模型中, 检测同一数据集得到不同类别垃圾的平均精度列于表 5。由表 5 中

YOLOv7 模型的检测结果可知, 泡沫、纸制品、玻璃都具有较高的检测精度, 塑料、金属和木材的平均精度相对较低。在 YOLOv7+QARepVGG 中, 各类别的平均精度相对均衡, 金属制品的 *AP* 降低较为明显, 减少了 3.9%。在 MP 模块中引入 SimAM 注意力机制后, 明显看出每一个类别的 *AP* 均获得了提升, 且相对本文的其他模型, 金属和玻璃都达到最好的平均精度。将 YOLOv7+

改进 *BiFPN* 用于海滩垃圾检测, 同样为各类别检测精度的提升做出了有效的贡献, 其中尺寸变化较大的木材的 *AP* 提升了 12.4%, 提升最为明显。在改进模型中, 相对于其他模型, 泡沫、塑料类、纸制品都达到了最好的检测精度, 且玻璃和木材的 *AP* 分别提升了 3.9% 和 6.6%, 仅金属类的 *AP* 略微下降。总体上, 改进算法明显优于 YOLOv7 算法。

表 5 不同模型检测各类海滩垃圾的平均精度
Tab.5 Average precision of different models for detecting various types of beach litter

模型	泡沫/(%)	塑料/(%)	金属/(%)	纸制品/(%)	玻璃/(%)	木材/(%)	织物类/(%)	橡胶/(%)
YOLOv4	82.1	64.2	63.5	78.3	78.6	68.7	69.1	80.2
YOLOv5s	85.7	67.4	68.2	81.2	86.4	71.6	73.2	82.5
YOLOv7	88.6	69.6	69.1	83.1	92.7	70.9	75.1	82.2
YOLOv7+QARepVGG	90.2	77.3	65.3	86.6	92.2	77.5	74.2	79.8
YOLOv7+S-MP	92.2	71.4	71.4	90.4	98.7	73.4	77.5	89.2
YOLOv7+改进BiFPN	93.2	77.9	68.1	87.4	88.9	83.3	79.3	86.7
改进YOLOv7	94.3	80.7	68.4	92.5	96.6	77.5	78.6	89.6

3.5 消融实验结果

为了验证每处改进模块检测垃圾的效果, 本研究进行了多组消融实验, 结果列于表 6。由表 6 中的实验 2 可知, 在 YOLOv7 模型中加入量化感知的 QARepVGG 模块后, 模型参数量降低, 推理时间减少, 检测速度加快, *mAP* 值增大。加入 SimAM 模块和改进 *BiFPN* 模块后, 推理速度不变, 但 *F1* 值和 *mAP* 值均增大。同时, 加入 QARepVGG 模块和 SimAM 模块后, 推理速度、

F1 值和 *mAP* 值均有所提升。加入改进 *BiFPN* 模块后, *F1* 值和 *mAP* 值进一步增大, 其中 *F1* 值和 *mAP* 值分别增大至 0.81、84.8%, 较原始 YOLOv7 分别提高 0.08 和 5.8%, 同时参数量减小 281600, 推理时间缩短 3.5 ms, 模型的检测速度提升 43%。图 6 展示了每个模块单独加入到 YOLOv7 进行改进后与原 YOLOv7 模型检测性能之间的差异, 充分体现了改进后的模型在海滩垃圾检测中的高效性和准确性。

表 6 消融实验结果
Tab.6 Results of ablation experiments

实验	QARepVGG	S-MP	改进BiFPN	参数量	推理时间/ms	<i>F1</i>	<i>mAP</i> @0.5/(%)	FPS
1				37223526	11.8	0.73	79.0	40
2	√			36434278	8.0	0.75	80.7	58
3		√		37140838	11.8	0.77	82.9	40
4			√	37354598	11.8	0.78	83.1	40
5	√	√		36382514	8.3	0.79	83.6	52
6	√	√	√	36941926	8.3	0.81	84.8	57

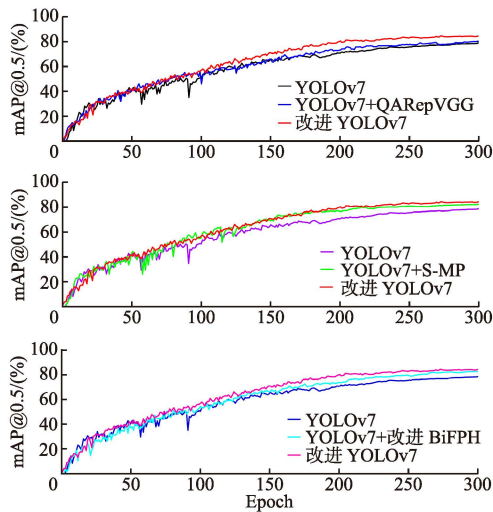


图6 在不同模型上的消融实验结果

Fig. 6 Results of ablation experiments on different models.

3.6 不同模型的检测性能

为了验证改进模型在自建数据集中检测海滩垃圾的效果,本研究在几种流行的目标检测模型(包括 Faster R-CNN^[28]、EfficientDet^[29]、SSD^[30]、YOLOv3^[31]、YOLOv4^[32]、YOLOv5^[33]、YOLOX^[34]和 YOLOv7^[24])上进行了多组实验,实验结果列于表 7。由表 7 可知,相同的数据集上,改进的 YOLOv7 算法对海滩垃圾的检测平均精度达到 84.8%,FPS 达到 57,检测效果和效率优于其他

模型,说明针对 YOLOv7 模型进行的改进能明显提升海滩垃圾的识别效果。

3.7 实际场景下的检测效果

为了验证改进模型对海滩垃圾的实际检测性能,本研究对比了 YOLOv7 和改进 YOLOv7 模型在实际海滩场景下的检测效果(图 7)。根据图 7a、图 7b、图 7d、图 7e 的检测结果,绝大部分类型的垃圾都能被识别出来,但图 7a 存在漏检的情况,经改进之后,模型成功检测出漏检和误检的垃圾。图 7b 的背景较复杂,部分目标被遮挡导致 YOLOv7 出现误检,将树叶识别为纸制品和塑料类垃圾,而且漏检了小尺寸垃圾。改进模型成功检测出这些垃圾,且预测框置信度高于 YOLOv7 模型置信度。图 7c 中垃圾的尺寸较小,YOLOv7 无法识别出图中的塑料垃圾,但改进模型成功识别到该垃圾。无人机图像中的目标尺寸普遍偏小,导致 YOLOv7 模型出现漏检(图 7g)和误检(图 7h)的情况,但改进模型成功识别出这些小目标,并且得到了更高的预测置信度。因此,改进的模型比原 YOLOv7 模型在海滩垃圾的识别和分类上更具优势,不但减少了漏检和误检情况,而且取得了更高的检测精度,验证了改进模型检测海滩垃圾的高效性和准确性。

表 7 不同模型检测结果对比

Tab.7 Comparison of detection results of different models

模型	骨干网络	参数量/M	mAP@0.5/(%)	F1	FPS
Faster R-CNN	VGG16	108.2	50.94	0.61	13
EfficientDet-D0	EfficientNet	6.6	51.67	0.63	23
SSD	VGG16	24.28	62.51	0.67	43
YOLOv3	Darknet53	123.5	70.46	0.69	35
YOLOv4	CSPDarknet19	64.3	75.23	0.71	31
YOLOv5s	CSPDarknet53	7.0	76.81	0.72	36
YOLOX	Darknet53	63.7	77.15	0.72	37
YOLOv7	CSPDarknet53	37.2	79.04	0.73	40
改进YOLOv7	改进CSPDarknet53	36.9	84.80	0.81	57

4 结论

(1)本研究在华南沿岸收集了大量海滩垃圾图像,提出一种改进 YOLOv7 算法用于从图像中检测并进行海滩垃圾分类,通过数据增强方法扩

充数据集,扩大海滩垃圾的检测范围,能有效避免因训练样本不足而产生的过拟合现象。

(2)本文提出用 QARepVGG 结构替换 YOLOv7 模型的部分 E-ELAN 结构以减小模型参数量。针对数据集中存在背景复杂和目标尺



YOLOv7 模型(第 1、3 行)与改进 YOLOv7 模型(第 2、4 行)检测结果可视化, 其中(a)至(f)是相机照片, (g)至(l)是无人机图片

图 7 不同场景下的模型检测结果可视化

Fig. 7 Visualization of model detection results of different scenes

寸范围大等问题, 将 SimAM 结构引入 Head 中的 MP 模块, 使模型更专注于目标, 同时用改进的 BiFPN 结构替换 PAFPN 结构, 提高了网络对不同尺寸目标的特征学习能力和检测精度。

(3)在不同模型上的对比结果表明, 改进的 YOLOv7 算法针对海滩垃圾的检测效果优于 YOLOv7 模型和其他模型, 且在不同海滩场景中, 模型能满足实时性检测需求。

(4)在未来的研究中, 将加强海滩垃圾数据的收集, 建立一个海滩垃圾基准数据集, 并开发

出完整的检测系统, 部署到无人机以降低大范围海滩监测的成本。同时针对不同类型海滩的复杂环境, 进一步增强模型的检测性能和泛化能力, 提高模型的实用价值。

参考文献:

[1] PHAM C K, RAMIREZ L S, ALT C H S, et al. Marine litter distribution and density in European seas, from the shelves to deep basins[J]. PloS one, 2014, 9(4): 95839.
[2] 赵 肖, 綦世斌, 廖 岩, 等. 我国海滩垃圾污染现状及控制

- 对策[J]. 环境科学研究, 2016, 29(10): 1560-1566.
- [3] MURRAY F, COWIE P R. Plastic contamination in the decapod crustacean *Nephrops norvegicus* (Linnaeus, 1758)[J]. Marine Pollution Bulletin, 2011, 62(6): 1207-1217.
- [4] LANDY M K, ROBERTS M J, THOMAS S R, et al. The environmental protection agency: asking the wrong questions[J]. Journal of Public Health Policy, 1992, 13(4): 518-519.
- [5] 潘永珩, 高郭平, 张春玲, 等. 中国典型海滩垃圾组成和时空变化特征[J]. 海洋湖沼通报, 2021 (1): 142-148.
- [6] 张相君. 东亚海区域海洋环境保护法律制度的构建[J]. 广东海洋大学学报, 2011, 31(2): 7-12.
- [7] PERVEZ R, LAI Z P. Spatio-temporal variations of litter on Qingdao tourist beaches in China[J]. Environmental Pollution, 2022, 303: 119060.
- [8] MARTIN C, PARKES S, ZHANG Q N, et al. Use of unmanned aerial vehicles for efficient beach litter monitoring[J]. Marine Pollution Bulletin, 2018, 131: 662-673.
- [9] FALLATI L, POLIDORI A, SALVATORE C, et al. Anthropogenic marine debris assessment with unmanned aerial vehicle imagery and deep learning: a case study along the beaches of the Republic of Maldives[J]. Science of the Total Environment, 2019, 693: 133581.
- [10] NAKASHIME, ISOBE A, MAGOME S, et al. Using aerial photography and in situ measurements to estimate the quantity of macro-litter on beaches[J]. Marine Pollution Bulletin, 2011, 62(4): 762-769.
- [11] KAKO S, ISOBE A, MAGOME S. Low altitude remote-sensing method to monitor marine and beach litter of various colors using a balloon equipped with a digital camera[J]. Marine Pollution Bulletin, 2012, 64(6): 1156-1162.
- [12] BAO Z, SHA J, LI X, et al. Monitoring of beach litter by automatic interpretation of unmanned aerial vehicle images using the segmentation threshold method[J]. Marine Pollution Bulletin, 2018, 137: 388-398.
- [13] 邱琰皓, 张 芳. 基于无人机低空拍摄图像的垃圾监测——以青岛石雀滩为例[J]. 海洋与湖沼, 2022, 53(6): 1349-1359.
- [14] GRUNDLEHNER A, DIEPENS N J, LINDERS T, et al. Towards continuous mass and size distributions for beach plastic litter: Spatiotemporal analyses of abundance and composition[J]. Journal of Hazardous Materials, 2023, 458: 131984.
- [15] 周志伟, 邓 化, 施华宏. 基于无人机遥感与机器学习的岸滩大型塑料垃圾监测方法[J]. 海洋环境科学, 2023, 42(1): 141-150.
- [16] 何钰湛, 葛振鹏, 李道季, 等. 基于激光雷达的海滩垃圾快速识别[J]. 海洋学报, 2019, 41(11): 156-162.
- [17] PFEIFFER R, VALENTINO G, FARRUGIA R A, et al. Detecting beach litter in drone images using deep learning[C]//Proceedings of 2022 IEEE International Workshop on Metrology for the Sea; Learning to Measure Sea Health Parameters. Milazzo: IEEE, 2022: 28-32.
- [18] SONG K, JUNG J Y, LEE S H, et al. A comparative study of deep learning-based network model and conventional method to assess beach debris standing-stock[J]. Marine Pollution Bulletin, 2021, 168: 112466.
- [19] CESARANO C, AULICINO G, CERRANO C, et al. Marine beach litter monitoring strategies along Mediterranean coasts. A methodological review[J]. Marine Pollution Bulletin, 2023, 186: 114401.
- [20] DING X H, ZHANG X U, MA N M, et al. RepVGG: making VGG-style ConvNets great again[C]//Proceedings of 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville: IEEE, 2021: 13728-13737.
- [21] CHU X X, LI L, ZHANG B. Make RepVGG greater again: a quantization-aware approach[J]. arXiv: 2212.01593, 2022.
- [22] YANG L X, ZHANG R Y, LI L D, et al. SimAM: a simple, parameter-free attention module for convolutional neural networks[C]//Proceedings of the 38th International Conference on Machine Learning. PMLR, 2021: 11863-11874.
- [23] LIN T Y, DOLLÁR P, GIRSHICK R, et al. Feature pyramid networks for object detection[C]//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu: IEEE, 2017: 2117-2125.
- [24] WANG C Y, BOCHKOVSKIY A, LIAO H Y M. YOLOv7: trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[C]//Proceedings of 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Vancouver: IEEE, 2023.
- [25] LIN J, LIN H F, WANG F. A semi-supervised method for real-time forest fire detection algorithm based on adaptively spatial feature fusion[J]. Forests, 2023, 14(2): 361.
- [26] WOO S, PARK J, LEE J Y, et al. CBAM: convolutional block attention module[C]//Proceedings of the 15th European Conference on Computer Vision. Munich: Springer, 2018.
- [27] WANG Q L, WU B G, ZHU P F, et al. ECA-Net: efficient channel attention for deep convolutional neural networks [C]//Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle: IEEE, 2020: 11531-11539.

- tion of marine litter on the seafloor of the gulf of Sant Jordi (western Mediterranean Sea)[J]. *Environments*, 2021, 8(10): 106.
- [4] 许林之. 我国海洋垃圾监测与评价[J]. *环境保护*, 2008(19): 67-68.
- [5] UNEP. From Pollution to Solution: a global assessment of marine litter and plastic pollution[R]. Nairobi: United Nations Environment Programme, 2021.
- [6] 孙承君, 蒋凤华, 李景喜, 等. 海洋中微塑料的来源、分布及生态环境影响研究进展[J]. *海洋科学进展*, 2016, 34(4): 449-461.
- [7] 国峰, 周鹏, 李志恩, 等. 2011年东中国海沿岸海域海洋垃圾分布、组成与来源分析[J]. *海洋湖沼通报*, 2014, 36(3): 193-200.
- [8] 中华人民共和国生态环境部. 中国海洋生态环境状况公报[EB/OL]. [2024-02-02]. <https://www.mee.gov.cn/hjzl/sthjzk/jagb/>.
- [9] 广东省生态环境厅. 广东省生态环境状况公报[EB/OL]. [2024-02-02]. <https://gdee.gd.gov.cn/hjzkgb/>.
- [10] 赵蓓, 唐伟, 周艳荣, 等. 黄河入海主要污染物特征研究[J]. *海洋环境科学*, 2008, 27(增刊2): 60-62.
- [11] MUGILARASAN M, KARTHIK R, ROBIN R S, et al. Anthropogenic marine litter: an approach to environmental quality for India's southeastern Arabian Sea coast[J]. *Science of the Total Environment*, 2023, 866: 161363.
- [12] RANGEL-BUITRAGO N, VELEZ-MENDOZA A, GRACIA C A, et al. The impact of anthropogenic litter on Colombia's central Caribbean beaches[J]. *Marine Pollution Bulletin*, 2020, 152: 110909.
- [13] SCHULZ M, NEUMANN D, FLEET D M, et al. A multi-criteria evaluation system for marine litter pollution based on statistical analyses of OSPAR beach litter monitoring time series[J]. *Marine Environmental Research*, 2013, 92: 61-70.
- [14] 赵肖, 綦世斌, 廖岩, 等. 我国海滩垃圾污染现状及控制对策[J]. *环境科学研究*, 2016, 29(10): 1560-1566.
- [15] ZHOU P, HUANG C G, FANG H D, et al. The abundance, composition and sources of marine debris in coastal seawaters or beaches around the northern South China Sea (China)[J]. *Marine Pollution Bulletin*, 2011, 62(9): 1998-2007.
- [16] KANG B, LIN L, LI Y, et al. Facing marine debris in China[J]. *Marine Pollution Bulletin*, 2022, 184: 114158.
- [17] MUGILARASAN M, KARTHIK R, PURVAJA R, et al. Spatiotemporal variations in anthropogenic marine litter pollution along the northeast beaches of India[J]. *Environmental Pollution*, 2021, 280: 116954.
- (本文编辑: 胡莹莹)

(上接第 695 页)

- [28] 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.
- [29] TAN M X, PANG R M, LE Q V. EfficientDet: scalable and efficient object detection[C]//Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle: IEEE, 2020: 10781-10790.
- [30] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot MultiBox detector[C]//Proceedings of the 14th European Conference on Computer Vision. Amsterdam: Springer, 2016: 21-37.
- [31] REDMON J, FARHADI A. YOLOv3: an incremental improvement[J]. *arXiv*: 1804.02767, 2018.
- [32] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. YOLOv4: optimal speed and accuracy of object detection[J]. *arXiv*: 2004.10934, 2020.
- [33] WANG H B, WANG Y X. Improved glove defect detection algorithm based on YOLOv5 framework[C]//Proceedings of the IEEE 6th Advanced Information Technology, Electronic and Automation Control Conference. Beijing: IEEE, 2022: 1192-1197.
- [34] ZHENG G, LIU S T, WANG F, et al. YOLOX: exceeding YOLO series in 2021[J]. *arXiv*: 2107.08430, 2021.
- (本文编辑: 胡莹莹)