

YOLOESD-基于改进 YOLOv8 的海洋垃圾图像检测

李翠, 王姣[†]

(大连交通大学软件学院, 辽宁 大连 116021)

摘要: 海洋垃圾不仅严重威胁海洋动物的健康及其栖息地, 由其释放的有毒物质进入食物链后对人类身体健康同样造成消极影响。由于海洋图像受到光照投影的影响, 并且垃圾的尺寸通常较小, 以往的目标检测算法对海洋垃圾的检测性能并不理想, 因此提出了一种基于 YOLOv8 网络模型的改进算法(YOLOESD), 该算法共有三个改进点: 首先, 采用 Stemblock 模块替换了模型的初始卷积, 在减少模型参数量的同时, 提高模型检测的精确度; 其次, 融合高效多尺度注意力模块(EMA), 有效减少了模型的漏检和误检问题; 最后, 在原模型的头部额外增加一个小目标检测头, 提高模型对小尺度目标的敏感度。实验结果表明, 改进后的 YOLOv8 网络模型与原网络模型相比, 漏检情况得到明显改善, mAP@0.5 达到 90.8%, 精度提高了 3.6 个百分点; YOLOESD 网络模型的检测效果优于原网络模型及经典的网络模型。

关键词: 目标检测; 小目标检测; 海洋垃圾检测; EMA 注意力机制; Stem 模块; YOLOESD

中图分类号: TP391

文献标识码: A

YOLOESD- Marine Litter Image Detection Based on Improved YOLOv8

LI Cui, WANG Jiao[†]

(School of Software, Dalian Jiaotong University, Dalian, Liaoning 116021, China)

Abstract: Marine litter is a serious threat to the health of marine animals and their habitats, and it also has a negative impact on human health through the release of toxic substances that enter the food chain. The performance of previous target detection algorithms for marine litter is not satisfactory due to the fact that marine images are affected by light projection and the size of the litter is usually small. Therefore, this paper proposes an improved algorithm (YOLOESD) based on the YOLOv8 network model, which has three improvement points. Firstly, the initial convolution of the model is replaced by the Stemblock module. While reducing the number of model parameters, improve the detection accuracy of the model. Secondly, we have integrated the efficient multi-scale attention module (EMA). Effectively reducing the problem of missed and false alarms in the model. Finally, an additional small target detection head is added to the head of the original model to improve the model's sensitivity to small-scale targets. The experiments results show that the improved YOLOv8 network model has significantly improved leakage detection compared to the original network model, the mAP@0.5 reached 90.8%, and the accuracy is improved by 3.6 percentage points; the YOLOESD network model outperforms both the original network model and the classical network model in terms of detection.

Key words: target detection; small target detection; marine litter detection; EMA attention mechanism; Stem module; YOLOESD

目前,海水资源受污染程度越来越重,自 1950 年以来,塑料垃圾在全球范围内存在且呈指数级增长,据估计,每年仅陆地地区就有 13 万吨塑料垃圾流入海洋^[1]。塑料垃圾可作为运输其他污染物的载体,天然毒素可以吸附在塑料上,海洋生物通过摄入毒素正在影响器官的健康。此外,垃圾还存在养分循环恶化、濒危物种灭绝等潜在的影响,这不仅污染海洋环境,还波及人类健康,甚至如今的海水中检测出了直径小于 5 mm 的微塑料,这种微塑料已经渗透到食物链中,最终流向人体内部^[2]。

海洋垃圾检测方面仍然有许多挑战需要解决。Ma 等^[3]将 RetinaNet 算法与数据增强、损失函数优化以及迁移学习相结合,有效提升了海洋垃圾检测的精度值。Zaaboub 等^[4]将无人机与机器学习技术、K 近邻算法相结合,在垃圾识别的四次测试中误差平均值仅为 6.3%。Winans 等^[5]使用单次多盒探测器结合 MobileNet-V2 特征提取器对夏威夷沿海海洋垃圾进行了检测,最终达到了 71.8%的精度值。虽然以上的研究都表明,新技术的使用,可以在有效地检测海洋垃圾的同时节省大量的人力物力,但详细查看结果时,可以发现,实验结果总体精度偏低,特别是在小目标物体增加时,漏检率很高,检测精度削弱。

为了解决上述问题,提升目标检测的精度,有效解决海洋垃圾中小目标漏检及误检问题,文中提出了一种改进的 YOLOv8 目标检测模型(YOLOESD)。本文的主要改进如下:

- (1)融合新型高效多尺度注意力(EMA)^[6]机制,通过将先前帧的预测结果纳入考虑范围,将当前帧的预测结果与先前帧的结果相结合,有助于模型降低漏检率与误检率。
- (2)模型轻量化,用 Stemblock 模块^[7]替换 YOLOv8 网络模型^[8]中的起始卷积,在精度稍有提高的情况下,减少模型的参数量。
- (3)在 YOLOv8 网络模型的头部额外添加一个小目标检测头,有效解决小目标漏检率高的问题,进一步提升检测精度。

1 改进 YOLOv8 的方法

本文采用近年来较先进的单阶段目标检测算法 YOLOv8 模型作为本次实验的基本模型,修改 YOLOv8 模型来定位整个图像中的垃圾点,并进行有效分类。模型整体的框架图如图 1 所示,将在下面的小节中进行详细解释。

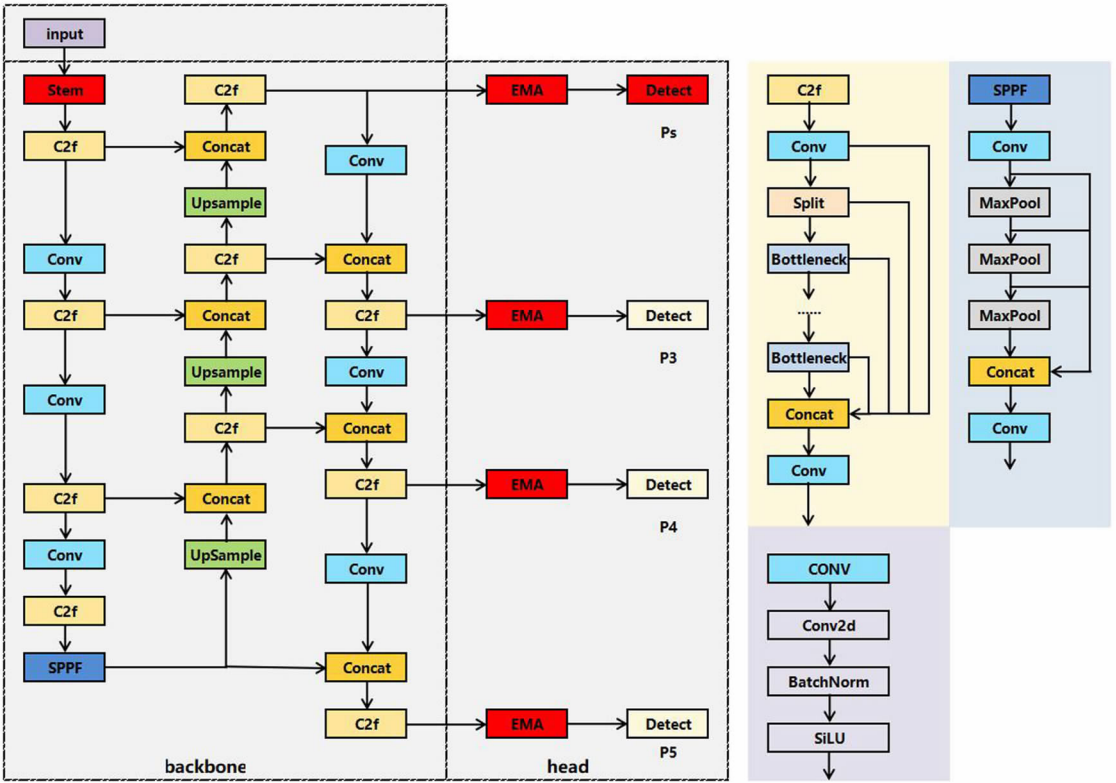


图 1 改进版 YOLOv8 网络模型框架图

1.1 融合 EMA 注意力模块

跨通道关系建模提取深度视觉表示时,通道降维会带来副作用,为了解决这一问题,高效多尺度注意力模块(EMA)横空出世。EMA 专注于保留每个通道的信息,并尽可能减少计算开销,一方面对全局信息进行编码,校准每个并行分支中的通道权重;另一方面,通过跨维度交互来进一步组合两个并行分支的输出特征。EMA 的总体结构如图 2 所示。

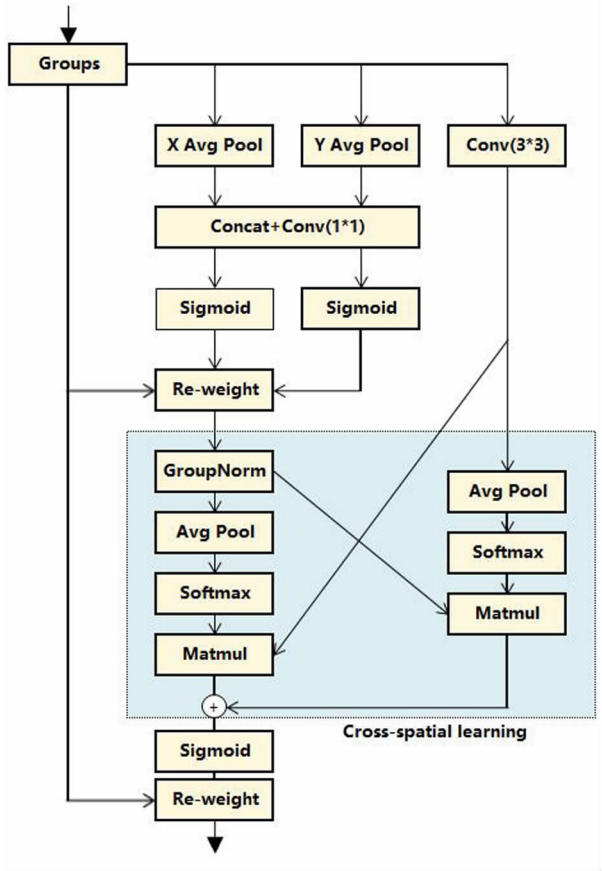


图 2 EMA 注意力机制的总体结构图

EMA 注意力机制的跨空间信息聚合方法:由图 2 阴影部分可知,1×1 分支的输出和 3×3 分支的输出作为阴影部分的输入。然后,在 1×1 分支中,利用 2D 全局平均池化对全局空间信息进行编码,使用非线性函数 Softmax 来拟合线性变换。将上述输出与矩阵点积运算相乘,导出了第一个空间注意力图。文中类似的利用 2D 全局平均池化来编码 3×3 分支中的全局空间信息,导出保留了整个精确空间位置信息的第二空间注意力图。

EMA 在卷积运算中不降低通道维度的情况下学习有效的通道描述,并为高级特征图产生更好的像素级关注。2D 全局池化操作式如下:

$$Z_c = \frac{1}{H \times W} \sum_{0 \leq j \leq H} \sum_{0 \leq i \leq W} X_c(i, j) \tag{1}$$

式中, H 代表特征图的高, W 代表特征图的宽, X_c 表示在第 c 个通道处的输入特征。

1.2 集成 Stem block 模块

Stem block 结构是用于下采样的方法,该模块能够在保持较强特征能力的同时减少模型的参数,基于多次实验后分析 YOLOv8 网络模型的结构发现,backbone 部分起始的两个 3×3 卷积是为了进行下采样操作,较少提取小物体的空间信息。因此,如图 1 所示,在原 YOLOv8 网络模型的基础上,本文使用 Stem block 模块替换起始卷积,修改原模型的主干网络,在精度稍有提高的同时降低模型的整体参数。

Stem block 模块的结构图如图 3 所示。

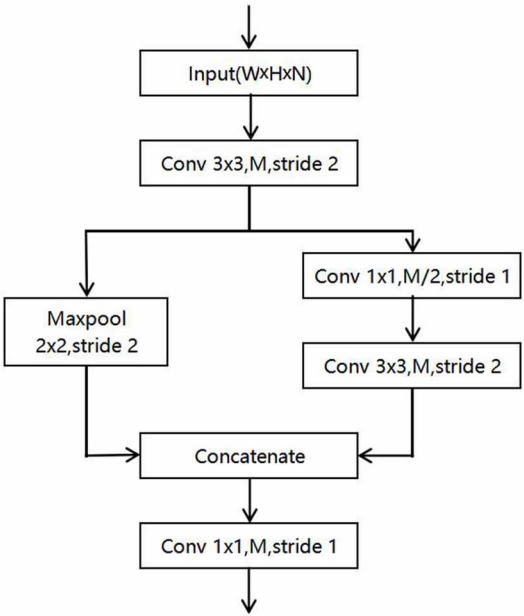


图 3 Stemblock 模块结构图

从图 3 可以看出,Stem block 结构有左右两个分支,右侧分支先将通道数量减少,再进行下采样;左侧分支将原始输入进行最大值池化;之后两个分支的结果进行拼接。目的是将输入中的部分信息进行传递,确保最终的结果既减少了参数量又具备足够的语义信息,不会造成信息的过度损失。

1.3 增加小目标检测头

在进行海洋垃圾检测时,由于海洋垃圾体积较小,易出现小目标漏检问题。在 YOLOv8 中,检测头由 P3、P4 和 P5 三个输出特征映射组成,P3 对应的特征图大小为 80×80,用于检测大小在 8×8 以上的目标,下采样 8 倍;P4 对应的特征图大小为 40×40,用于检测大小在 16×16 以上的目标,下采样 16 倍;P5 对应的特征图大小为 20×20,

用于检测大小在 32×32 以上的目标,下采样 32 倍。虽然上述三个检测头能够对多尺度目标进行检测,但较大的下采样倍数使模型对小目标及微小目标的检测能力不佳。本文提出的模型中,增加了一个额外的检测头 Ps,如图 1 所示,其特征图大小为 160×160 ,用于检测大小在 4×4 以上的目标,下采样 4 倍。较小的下采样倍数包含更多的低层次信息,可以有效地检测出小目标,提高了模型在给定图像中有效检测物体的能力,进一步提高了模型的检测精度。

2 实验

2.1 数据集

在实验部分,使用 Okahublot 公开的 Flow-Img^[9]数据集来验证模型的准确性,Flow-Img 数据集是 Okahublot 发布的无人船视角下的漂浮垃圾数据集,共包括 2000 张图片。

2.2 实验细节

网络实验环境为 Windows10、Python3. 9. 13 和 PyTorch1. 13. 0,相关硬件配置和模型参数如表 1 所示,数据的训练轮数为 200。

表 1 实验相关硬件配置和模型参数

名称	配置	参数	参数值
GPU	RTX3060Ti	图片大小/(像素×像素)	640×640
CPU	CoreTMi5-12400	学习率	0. 01
CUDA	11. 8	RAM	24GB
CuDNN	8. 7. 0	批量大小	8

实验指标主要选取平均精度值 mAP (mean Average Precision)。mAP@0. 5 代表 IoU 设置为 0. 5 时,所有图片的平均准确率;mAP@0. 5—0. 95 代表 IoU 从 0. 5 到 0. 95 步长为 0. 05 时取得的平均准确率;公式如下所示:

$$P=\frac{\text{True Positive}}{\text{True Positive}+\text{False Positive}} \tag{2}$$

$$R=\frac{\text{True Positive}}{\text{True Positive}+\text{False Negative}} \tag{3}$$

$$\text{mAP}=\int_0^1 P(R)d(R) \tag{4}$$

其中:式(2)代表准确率 P (Precision),式(3)代表召回率 R (Recall)。True Positive 表示预测正确;False Positive 表示预测错误,包括目标检测类别错误和漏检两种情况。

2.3 实验对比

在本节中,将在 Flow-Img 数据集上比较 YOLOESD 模型与其他经典网络模型的精度值,主要

包括 Faster-RCNN 网络模型 (MobileNetv2 骨干网络)^[10]、SSD 网络模型 (ResNet50 骨干网络)^[11]、YOLOv7 网络模型^[12]、YOLOv8 网络模型。实验结果见表 2。

从表 2 中可以看出,改进版 YOLOESD 网络模型,相较于以往的经典模型,检测精度遥遥领先,mAP@0. 5 达到了 90. 8%,mAP@0. 5—0. 95 达到了 49. 1%,与基线模型 Yolov8 相比,mAP@0. 5 提升了 3. 6 个百分点,mAP@0. 5—0. 95 提升了 1. 8 个百分点。

表 2 Flow-Img 数据集测试结果比较

网络模型	Input size/ (像素×像素)	mAP@0. 5	mAP@0. 5—0. 95
SSD	640×640	0. 593	0. 256
Faster-RCNN	640×640	0. 625	0. 263
YOLOv7	640×640	0. 737	0. 345
YOLOv8	640×640	0. 872	0. 473
YOLOESD	640×640	0. 908	0. 491

为了说明改进版模型的性能,文中以 Flow-Img 数据集上 YOLOv8 网络模型和 YOLOESD 网络模型的检测结果为例进行展示。在图 4 中,第一行为 YOLOv8 网络模型的检测结果,第二行为 YOLOESD 网络模型的检测结果。具体来看,图 4 中的第一组图片,YOLOv8 网络模型检测出目标的置信度分数为 0. 3,改进版模型检测出的分数为 0. 4;第三组图片 YOLOv8 网络模型检测出的分数为 0. 5 和 0. 8,改进版模型检测出的分数为 0. 6 和 0. 9,均高于原网络模型;第四组图片,改进版模型在分数提高的基础上,额外检测出了一个原模型漏检的小目标,由此可见,改进版模型的精度更高、性能更好(注:同一列的一对图片称为一组,由左向右,分别为第一组至第四组)。

2.4 消融实验

表 3 是 YOLOESD 消融实验的结果。使用 YOLOv8 网络模型作为基线模型。为了验证不同模块的性能,本文进行了 8 个不同的实验,首先,单独增加小目标检测头、融合 EMA 注意力机制、集成 Stem block 模块,检测精度均有所上升;这表明三个改进点对 YOLOv8 网络模型的精度提高均是有益的。为了进一步分析模型性能,将三个改进点两两组合进行实验,结果显示,除了 EMA 与 Stem block 的组合外,其余组合实验精度提升比单独的改进更明显。EMA 与 Stem block 的结合精度无增加,分析认为 Stem block 结构的引进是为了降低模型的复杂程度,进一步降低模型的通道数,而 EMA 注意力机制善于融合通道信息,通道减少,

融合信息相应缩减,因此模型的精度没有上升。最后,本文融合三个改进点,对提出的 YOLOESD 网络模型进行实验,如表 2 和表 3 所示,该模型取得

了最优的结果,检测精度相比基线模型提升了 3.6 个百分点。



图 4 YOLOv8 与 YOLOESD 检测结果对比

表 3 消融实验结果显示

YOLOv8	EMA	Stem block	增加检测头	mAP@0.5/%	mAP@0.5—0.95/%
✓				87.2	47.3
✓	✓			88.3 (+1.1)	47.5 (+0.2)
✓		✓		88.9 (+1.7)	49.0 (+1.7)
✓			✓	89.1 (+1.9)	48.1 (+0.8)
✓	✓	✓		87.2 (+0)	47.1 (−0.2)
✓	✓		✓	90.2 (+3)	48.7 (+1.4)
✓		✓	✓	90.5 (+3.3)	48.7 (+1.4)
✓	✓	✓	✓	90.8 (+3.6)	49.1 (+1.8)

模型的改进会增加模型的参数,进而增加模型的复杂度,降低模型的运行时间。为了验证模型的运行速度,本文进行了一系列实验。结果显示,YOLOv8 模型的运行速度为 118.89 FPS,YOLOv7 模型的运行速度为 103.22FPS,YOLOESD 模型的运行速度为 99.4 FPS。运行速度相比原模型稍有下降,但仍然能够满足实时检测的需求。本文认为,较小的速度下降换来较大的精度提升是值得的,YOLOESD 在达到高精度的同时也满足了实时检测的需求。

3 结 论

近年来,海洋污染日益严重,检测与清理海洋垃圾成为现阶段的热议话题,海洋垃圾不仅污染水源、危害水中的动植物,更为病毒传播提供了有效途径,加快其传播速度,危害陆地动植物,进一步危害人类。基于海洋垃圾图像中小目标多且难辨别的问题,

本文提出了一种基于改进 YOLOv8 网络模型的海洋垃圾检测方法-YOLOESD。首先,文中将 EMA 注意力机制融合进 YOLOv8 网络模型,通过 EMA 注意力机制关联通道信息,显著提升了目标检测的精度;其次,将 Stem block 模块与 YOLOv8 网络模型集成,在提升精度的同时有效地降低参数量;最后对 YOLOv8 网络模型的整体结构进行更改,添加小目标检测头,使模型更精准地抓住小目标。实验结果表明,文中提出的三个改进点可以有效地提升目标检测的精度,mAP@0.5 与 mAP@0.5—0.95 分别为 90.8%与 49.1%,相较于原模型分别提升了 3.6个百分点与 1.8个百分点。

海洋垃圾检测是一个值得研究的问题,现阶段海洋垃圾图像较少且种类单一,极端情况的图片较少,例如大雾天气、大雨天气、雷暴天气等。这种极端天气会给目标检测带来极大挑战。未来,将对当前工作进行扩展,应用到不同的情景中。