

◆机器学习模型◆

基于注意力机制的 SK-YOLOv5 海洋目标检测分类算法*

李斌^{1,2}, 文莉莉^{2,3**}, 邬满², 刘画宁⁴, 许贵林²

(1. 广西自然资源职业技术学院商贸管理系, 广西崇左 532100; 2. 广西科学院, 广西近海海洋环境科学重点实验室, 广西人机交互与智能决策重点实验室, 数字孪生新技术研究院, 广西南宁 530007; 3. 广西壮族自治区药用植物园, 信息产业办, 广西南宁 530023; 4. 广西壮族自治区机构编制和绩效管理数据中心, 业务部, 广西南宁 530012)

摘要: 基于遥感影像的海洋目标图像具有多尺度、形状变化大、颜色暗淡、目标边界不清、图像模糊等特点, 需要在现有的目标检测算法上进行改进, 以满足遥感影像海洋目标检测及分类需要。针对这些问题, 在 You Only Look Once version 5 (YOLOv5) 的网络架构中引入 Selective Kernel Networks (SKNet) 注意力模块, 提出一种新的 SK-YOLOv5 网络, 增强网络对多尺度复杂海洋目标的特征提取和自适应能力。经对比实验测试, 在相同的海洋目标数据集上, 改进后的网络比原网络整体检测及分类准确率提升了约 9%。

关键词: 注意力机制; SKNet; YOLOv5; 海洋目标检测; 特征提取

中图分类号: TP399, TP181 文献标识码: A 文章编号: 1005-9164(2023)01-0132-07

DOI: 10.13656/j.cnki.gxkx.20230308.015

2017年以来, 国家海洋督察组、中央生态环境保护督察组针对海域使用及海洋环境问题开展了轮番督察督办, 加大了相关部门对海洋资源、海洋环境监测监管的压力与需求。在海域监管方面, 通常是利用卫星遥感、无人机航拍、现场拍照和远程视频监控等多种手段, 对海岸带现状和海洋工程建设运营情况进行监视监测, 及时发现和遏制非法用海、破坏生态、污染海洋环境等活动, 提升海域监管和海洋保护能力。

但随着无人机和卫星遥感影像数据的暴发式增长, 监测图像的审核工作量日益增加, 基于人工智能的海洋目标检测与识别技术对提高海洋监管效率显得尤为重要。

图像目标检测技术在交通、安防、金融、国土等领域早已成熟应用^[1], 但在海洋领域的应用还处在比较初级的阶段。一方面, 海洋监测应用场景复杂, 监测目标有着多样化、多尺度、时空关联的复杂特征; 另一方面, 目前交通、安防、自然场景等领域均有大量的公

收稿日期: 2022-03-15

修回日期: 2022-04-16

* 广西科技重大专项“空天地一体协同重大灾害应急智慧服务平台研发与应用示范”(桂科 AA22068072), 国家自然科学基金区域创新发展联合基金重点支持项目“台风影响下北部湾海浪与风暴潮演变特征、预测预报及灾变关系研究”(U20A20105)和自然资源部海洋信息技术创新中心2019年度开放基金项目“多源时空数据海洋目标智能提取与行为分析预警”资助。

【第一作者简介】

李斌(1976-), 男, 高级讲师, 主要从事深度学习、职业教育研究, E-mail: 3640708@qq.com。

【**通信作者】

文莉莉(1988-), 女, 高级工程师, 主要从事数据挖掘、深度学习研究, E-mail: 704663138@qq.com。

【引用本文】

李斌, 文莉莉, 邬满, 等. 基于注意力机制的 SK-YOLOv5 海洋目标检测分类算法[J]. 广西科学, 2023, 30(1): 132-138.

LI B, WEN L L, WU M, et al. SK-YOLOv5 Ocean Target Detection and Classification Algorithm Based on Attention Mechanism [J]. Guangxi Sciences, 2023, 30(1): 132-138.

开数据集供研究人员使用,便于研究人员测试各种算法模型,但海洋领域目前几乎没有公开的图像数据集,而自己建立数据集是一项工作量巨大、成本高昂的工作。因此,探索利用人工智能技术,自动对海量的海洋监测图像进行目标自动提取及识别,对提高海洋的综合监管能力和开展人工智能技术在海洋中的应用研究具有重要意义。在目前的目标检测算法中,快速区域卷积神经网络(Faster Region-Convolutional Neural Networks, Faster R-CNN)和 You Only Look Once version 5 (YOLOv5)算法在检测率上有着出色的表现^[2],但是对于背景复杂、目标尺度跨度大的海洋场景来说,其算法性能尚有较大的提升空间。其中,Faster R-CNN 属于 two-stage 算法,比较耗时,而 YOLOv5 属于 one-stage 算法,效率更高。因此,本研究利用注意力机制对复杂特征出色的提取能力,将其引入 YOLOv5 网络结构中,提升网络对复杂海洋场景的目标检测能力。

1 相关算法

1.1 注意力机制

注意力机制是模拟人脑快速获取图像关键信息、过滤无用信息的一种视觉焦点机制,即快速提取图像的关键特征^[3]。在机器视觉领域,注意力机制常用来增强神经网络的特征提取能力,通过对输入权重分配的关注,增强或放大图像的有效特征信息,可以看作是一种含参数的池化方法^[4]。目前常用的注意力机制模块有 Convolutional Block Attention Module (CBAM)、Squeeze - and - Excitation Networks (SENet)、Selective Kernel Networks (SKNet)等^[5]。

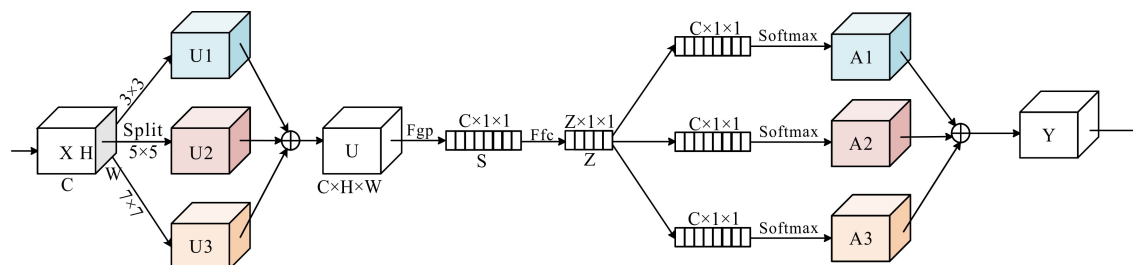


图1 SKNet 结构

Fig.1 Structure of SKNet

1.2 YOLOv5 网络

YOLO 系列目标检测算法是基于深度学习的回归方法,目前已经经历了 YOLOv1、YOLOv2、YOLOv3、YOLOv4、YOLOv5 共 5 代的发展^[10]。YOLO 没有显示求取 region proposal 的过程,相对

其中,CBAM 将注意力过程分为通道注意力和空间注意力两个独立的模块,减少了参数也节约了计算力,但需要手动设计池化、多层感知器等复杂操作^[6];SENet 是针对特征图的通道注意力机制,通过结合全局上下文对不同通道权值的重标定,实现对通道依赖的调整^[7];SKNet 是针对卷积核的注意力机制,根据不同大小的目标对不同大小的卷积核敏感度不同,SKNet 对不同的图像动态生成不同的卷积核,尝试找到最适合大小的卷积核^[8]。因此,SKNet 具有目标尺度自适应的特点,对多尺度目标检测较其他注意力机制模块具有更好的效果^[9]。SKNet 模块的处理主要包括以下 3 步。

①Split:使用两个大小不一样的卷积核对输入的特征图进行 Group Convolution (efficient grouped、depthwise convolutions、batch normalization、ReLU function) 等操作,如使用 3×3 、 5×5 的卷积核同时进行处理,得到两个不同的特征图,主要测试不同卷积核对目标的敏感度,以提高检测精度。

②Fuse:对上一步得到的两个不同的特征图进行融合,其思路是通过门来控制多个分支的信息流。这里使用 element-wise summation 来融合两个分支的结果,其中 Fgp 为全局平均池化操作,Ffc 为先降维再升维的两层全连接层。

③Select:使用两个不同的权重矩阵分别对两路结果进行加权处理,然后求和得到输出向量。由于整个过程中采用了不同的卷积和,因此该模块具有自适应地调整自身感受野的能力,对多尺度目标具有更高的检测能力和精度(图 1)。

于 Faster R-CNN 等算法,具有更高效的检测效率^[11]。YOLOv4 是 YOLO 家族的集大成者,在 COCO 数据集上已经具有非常优异的表现^[12]。YOLOv5 是在 YOLOv4 的基础上改进^[13],其模型更加轻量化,且速度与精度都得到极大提升,其主要改

进思路如下:

①在模型训练阶段实现了 Mosaic 数据增强、自适应锚框计算、自适应图片缩放等功能;

②在基准网络中引入 Focus 结构与 CSP 结构等其他算法的思路;

③在 Backbone 与最后的 Head 输出层之间添加了 FPN+PAN 结构;

④在 Head 输出端改进了训练时的损失函数

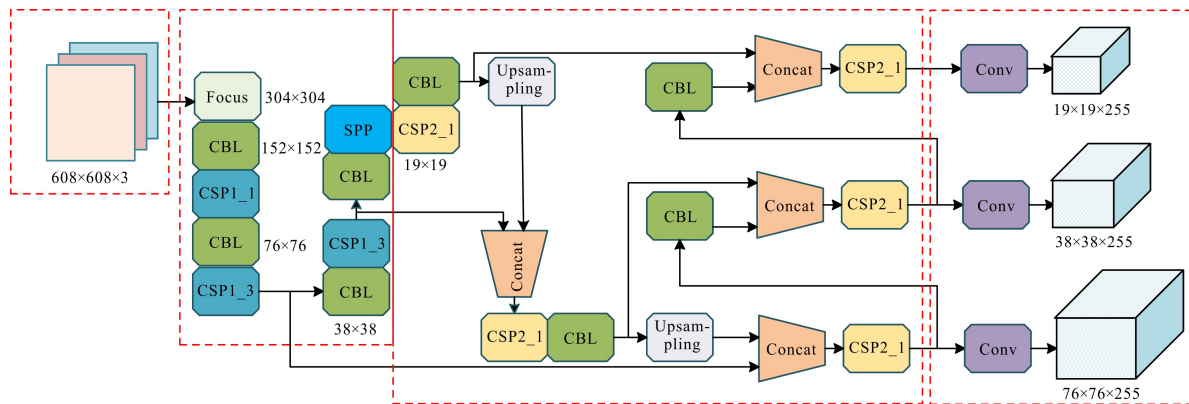


图2 YOLOv5网络结构

Fig. 2 Structure of YOLOv5 network

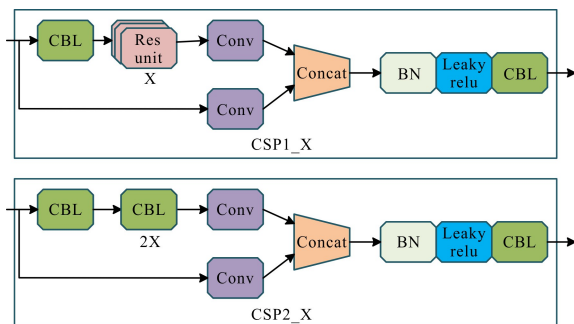


图3 CSP1_X和CSP2_X模块结构

Fig. 3 Structure of CSP1_X and CSP2_X module

①输入端:输入图片的预处理。该阶段首先将输入的图片缩放到该网络的输入大小,即 608×608 ,并进行归一化、Mosaic数据增强、自适应锚框计算等操作^[15]。

②基准网络:该阶段通常是用一些性能优异的分类器网络,如 CSPDarknet53 结构,提取一些通用的特征表示。在 YOLOv5 中使用 Focus 结构作为基准网络,该结构的主要思想是通过 slice 操作来对输入图片进行裁剪^[16]。

③Neck 网络:YOLOv5 中使用了 SPP 模块、FPN+PAN 模块来进一步提升特征的多样性和鲁棒性^[17]。该模块位于基准网络和头网络中间。该模块采用借鉴 CSPnet 设计的 CSP2 结构,从而加强网络特征融合能力。

GIOU_Loss,以及预测框筛选的 DIOU_nms。

YOLOv5 算法包括 YOLOv5s、YOLOv5m、YOLOv5l、YOLOv5x 4 个版本^[14],下面以 YOLOv5s 为例进行分析。该网络结构分为输入端、基准网络、Neck 网络、Head 输出端 4 部分(图 2)。重点改进的模块 CSP1_X、CSP2_X 的结构图如图 3 所示。

④Head 输出端:用来完成目标检测结果的输出,通常包含一个分类分支和一个回归分支。YOLOv5 中采用 GIoU_Loss 做 Bounding box 的损失函数^[18]。

另外,图 2 还包括一些基本计算单元,如 CBL、CSP1_X(X 代表有 X 个此模块,如 CSP1_1、CSP1_3)、CSP2_X(同 CSP1_X)、Focus、SPP 等。其中,CBL 模块由 Conv+BN+Leaky_relu 激活函数组成;CSP1_X 是借鉴 CSPNet 网络结构,由 CBL 模块、Res unit 模块及卷积层 Concat 组合而成;CSP2_X 也是借鉴 CSPNet 网络结构,由卷积层和 2X 个 CBL 模块 Concat 组合而成;Focus 结构首先将多个 slice 结果 Concat 起来,然后将其送入 CBL 模块中;SPP 单元采用 1×1 、 5×5 、 9×9 和 13×13 的最大池化方式,进行多尺度特征融合。

2 引入注意力机制的 SK-YOLOv5 网络

2.1 改进的思路及理由

遥感影像由于年份、数据来源、季节、分辨率、拍摄角度不同等多种原因,会导致同类目标颜色明暗、图像清晰度、目标尺寸等差异大的问题。同时不同类型的海洋目标本身也具有多尺度、形状变化大、目标边界不清、特征复杂等问题,如养殖池塘和红树林具有形状多样、特征复杂、多尺度等特性。由于上述原

因,基于遥感影像的海洋目标检测具有其独特的复杂性。SKNet 可对不同的图像动态生成不同的卷积核,对复杂的多尺度目标检测较其他注意力网络具有更强的适应性和特征提取能力;另外,当前在目标检测领域,YOLOv5 网络在检测精度和训练收敛性上均具有优异的表现。因此,本研究将 SKNet 引入 YOLOv5 网络中,提出一种可进一步增强网络多尺度自适应能力及检测精度的 SK-YOLOv5 网络。

2.2 改进的 CSP 结构

为了抑制网络中的无用特征,SK-YOLOv5 在 CSP1_X 和 CSP2_X 中引入 SKNet 模块,以增加网络深度并提升特征提取能力。网络在保持轻量化的同时,将 CSP1_X、CSP2_X 与 SKNet 结合,形成新的 SCSP1_X、SCSP2_X 单元(图 4)。

由图 4 可知,改进后的 CSP 结构利用新引入的 SKNet 模块增强特征图的有效特征,弱化无用特征,使网络对复杂、多尺度目标图像的适应能力明显提升。网络的深度有所增加,但基本可以忽略。

2.3 SK-YOLOv5 结构

该结构仍然是由输入端、基准网络、Neck 网络、

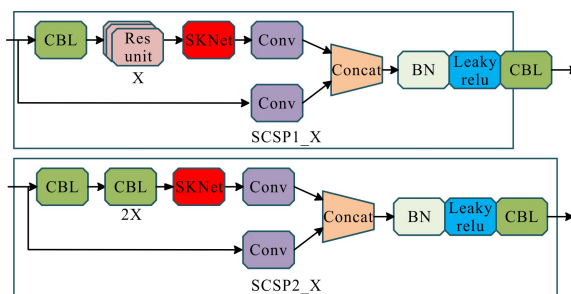


图 4 SCSP1_X 和 SCSP2_X 模块结构

Fig. 4 Structure of SCSP1_X and SCSP2_X module

Head 输出端 4 部分组成。其中,输入端和输出端跟 YOLOv5 完全一样,在基准网络和 Neck 网络中,使用改进的 SCSP1_X、SCSP2_X 单元分别替代原网络中的 CSP1_X、CSP2_X 单元,其他处理单元及网络结构不变(图 5)。因此,改进后的 SK-YOLOv5 网络整体结构与原 YOLOv5 网络基本保持一致,只是改进了 CSP 模块的内部结构,提升了网络的整体特征提取和自适应能力。SK-YOLOv5 网络既保持了 YOLOv5 网络的轻量、计算快速的优点,又具备其他 one-stage 目标检测算法不具备的多尺度目标自适应能力。

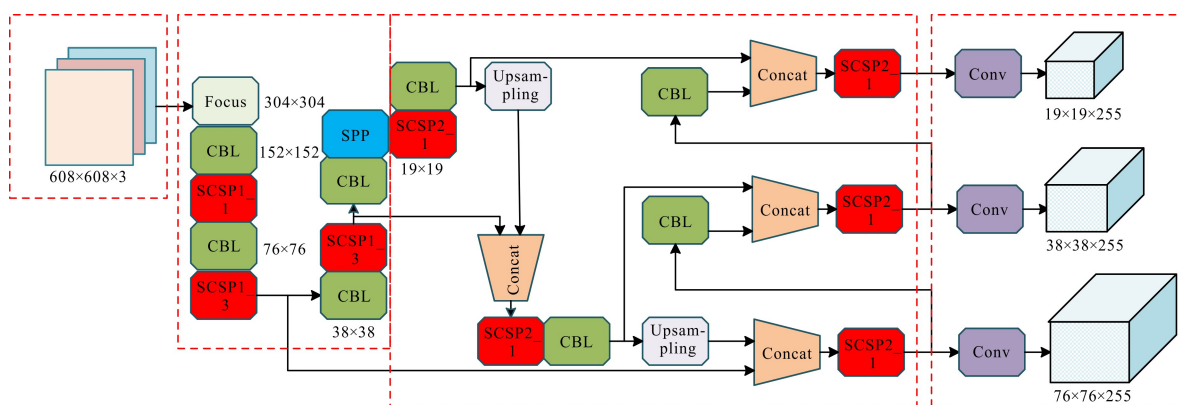


图 5 SK-YOLOv5 网络结构

Fig. 5 Structure of SK-YOLOv5 network

2.4 损失函数

SK-YOLOv5 网络中采用 BECLogits 损失函数计算目标性评分(Objectness score)的损失,采用 BCEclsloss 损失函数计算类别概率评分(Class probability score),以及使用 *GIoU* Loss 来评价预测框与真实框的位置损失。假设 p_1 、 p_2 为预测框的左上和右下两个顶点, l_1 、 l_2 为真实框的左上和右下两个顶点,预测框与真实框的位置如图 6 所示。

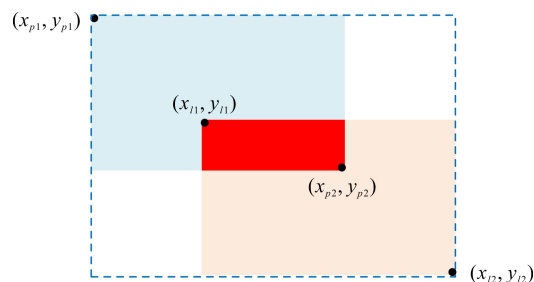


图 6 预测框与真实框的位置示意图

Fig. 6 Location schematic diagram of prediction frame and real frame

则 $GIoU$ Loss 的计算公式如下:

$$S_1 = (\min(x_{p2}, x_{l2}) - \max(x_{p1}, x_{l1})) \times (\min(y_{p2}, y_{l2}) - \max(y_{p1}, y_{l1})), \quad (1)$$

$$S_2 = (x_{p2} - x_{p1}) \times (y_{p2} - y_{p1}) + (x_{l2} - x_{l1}) \times (y_{l2} - y_{l1}) - S_1, \quad (2)$$

$$S_3 = (\max(x_{p2}, x_{l2}) - \min(x_{p1}, x_{l1})) \times (\max(y_{p2}, y_{l2}) - \min(y_{p1}, y_{l1})), \quad (3)$$

$$IoU = \frac{|A \cap B|}{|A \cup B|} = \frac{S_1}{S_2}, \quad (4)$$

$$GIoU = IoU - \frac{|C \setminus (A \cup B)|}{|C|} = IoU - \frac{S_3 - S_2}{S_3}, \quad (5)$$

$$GIoU \text{ Loss} = 1 - GIoU, \quad (6)$$

其中, IoU 表示预测框与真实框的交集与并集面积的比值, 也称交并比; A 表示预测框, B 表示真实框, C 表示 A 和 B 的最小包围框; S_1 表示 A 和 B 重叠区域的面积(图 6 红色部分), S_2 表示 A 和 B 求相并后的面积(图 6 蓝色 + 红色 + 黄色部分), S_3 为包围 A 、 B 的最小矩形框的面积(图 6 蓝色虚线框包含的面积); $GIoU$ Loss 表示最终求得的位置损失。

由上述公式可知, $S_3 - S_2$ 表示蓝色虚线框中既不属于 A 也不属于 B 的白色区域, $(S_3 - S_2)/S_3$ 则表示这个区域占虚线框区域的面积比, 这个比值越小, 说明预测框与真实框距离越近、重叠度越高, 反之则表示距离越远, 即误差越大。 $GIoU$ 是为克服 IoU 作为损失函数时梯度无法计算的缺点而提出的, 当 A 与 B 完全无重叠时, IoU 的值为 0, 则 $GIoU$ 为负值; 当 A 与 B 无重叠且距离无限远时, $GIoU$ 的值为 -1; 当 A 与 B 完全重叠时, 则 $GIoU$ 的值为 1。这种损失计算方式解决了当 A 与 B 无重叠时 IoU 恒为 0 的问题。

3 实验与结果分析

3.1 数据来源及数据集说明

本次实验采用的数据集是基于本团队利用大疆精灵 4 航飞的无人机高清影像, 以及 91 卫图助手遥感影像制作的海洋目标样本库, 影像分辨率为 0.2 - 5.0 m。选取船舶、蚝排、红树林、养殖池塘这 4 种常见且尺度变化大、特征复杂的海洋目标作为测试对象。每张图像分辨率为 400×400 像素, 共建立训练样本库 8 000 张, 其中各类目标(船舶、蚝排、红树林、养殖池塘)各 2 000 张, 测试样本库 2 000 张。数据分

布如表 1 所示。

表 1 不同类别海洋目标数据统计

Table 1 Statistics of different types of marine target data

目标类别 Target category	训练集 Training set	验证集 Validation set	测试集 Test set
Vessel	1 200	800	500
Mangrove	1 200	800	500
Oyster raft	1 200	800	500
Pond	1 200	800	500

由于数据集的像素维度各不相同, 需要进一步对图像进行变换处理。实验采用 YOLOv5 自带的 Mosaic 数据增强与图片缩放功能对图片进行预处理。

3.2 实验环境及过程设计

本次实验基于 Pytorch 深度学习框架, 集成开发及测试工具采用 PyCharm, 系统环境为 deepin 20.4, GPU 硬件环境为 Tesla T4。针对预处理后的归一化数据, 采用上文设计的 SK-YOLOv5 深度学习网络模型对样本进行训练与测试。本实验包括两个主要步骤: ①利用原 YOLOv5 网络对样本进行训练及测试; ②利用 2.3 节提出的 SK-YOLOv5 网络对同样的样本进行训练及测试; ③分别统计两种网络对同一样本的检测准确率。本实验选用相同的网络参数, 预加载模型为 yolov5x.pt, epochs 设置为 300, batch-size 为 32, img-size 为 [400 400], 除数据集和目标类别外, 其他参数均保持 YOLOv5 的默认参数不变。

3.3 结果与分析

本实验针对 4 类目标首先用两种网络进行测试, 测试数据集从 4 类目标中随机抽取各 500 张, 两组测试均使用一组相同的图片, 测试结果如表 2 所示。由测试结果可知, 改进后的网络整体检测准确率有明显提升, 针对上述数据集整体可提升 8.6% 左右。

表 2 原网络与改进后网络的对比测试结果

Table 2 Comparative test results of the original network and the improved network

网络模型 Network model	正确数量 Correct quantity	错误数量 Number of errors	准确率(%) Accuracy (%)
YOLOv5	418	82	83.6
SK-YOLOv5	461	39	92.2

然后再用改进后的网络分别对 4 类目标的测试数据集进行测试, 每类 500 张测试图片, 测试结果如表 3 所示。在这 4 类目标的单独测试中, 船舶由于特征明显, 检测准确率最高; 蚝排次之; 如图 7 所示, 红树林由于边界模糊、特征不明显、形状变化大等原因,

表 3 改进后算法分类测试结果

Table 3 Classification test results of improved algorithm

类别 Category	正确数量 Correct quantity	错误数量 Number of errors	准确率(%) Accuracy (%)
Vessel	483	17	96.6
Mangrove	436	64	87.2
Oyster raft	471	29	94.2
Pond	440	60	88.0

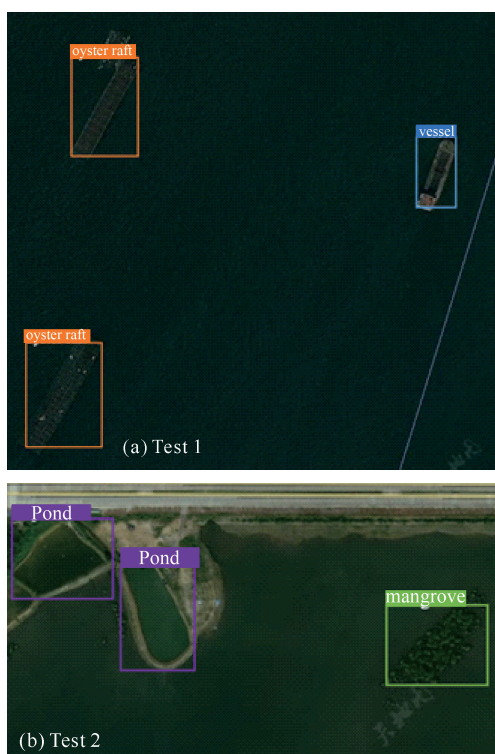


图 7 基于遥感影像的目标识别效果图

Fig. 7 Effect picture of target recognition based on remote sensing image

检测准确率相对最低,只有 87.2%。

在海域管理中,通过对遥感影像进行分块识别,提取关注的海洋目标、类别及其地理坐标,并结合空间规划数据,可以实现对疑似非法用海的目标进行预警分析。目标提取测试结果如图 7 所示,另外,还可以结合空间规划数据对海洋目标进行预警分析,通过该方法可以快速发现在空间规划禁止区进行养殖的蚝排。由此可见,基于深度学习的海洋目标智能提取方法可以应用于大范围海域的快速管理中,为智慧海洋建设提供技术支持。

4 结论

本研究通过引入 SKNet 注意力模块,改进了 YOLOv5 算法的网络结构,在不影响其轻量级和便

捷性的同时,增强了其针对特征复杂、边界模糊目标图像特征的提取能力和多尺度目标自适应能力,提升了其对复杂海洋目标的检测及分类准确率。经与原算法对比测试表明,改进后的算法总体检测及分类准确率和鲁棒性更强。该算法可实现对大范围海域的快速、智能监管,为智慧海洋建设提供技术基础。另外,自然资源领域监管中的图像(影像)审查工作量巨大,基于目标检测技术,结合海洋空间规划数据,实现对区域内海洋目标的用海合法性进行预警分析,对提高自然资源的综合监管能力和效率有着至关重要的作用。

参考文献

- [1] 张鹏,孔韦韦,滕金保. 基于多尺度特征注意力机制的人脸表情识别[J]. 计算机工程与应用,2022,58(1):182-189.
- [2] 赖振意,陈人和,钱育蓉. 结合空洞卷积的 CNN 实时微表情识别算法[J]. 计算机应用研究,2020,37(12):3777-3780.
- [3] WU H,ZOU Z,GUI J,et al. Multi-grained attention networks for single image super-resolution [J]. IEEE Transactions on Circuits and Systems for Video Technology,2020,31(2):512-522.
- [4] 王书献,张胜茂,朱文斌,等. 基于深度学习 YOLOV5 网络模型的金枪鱼延绳钓电子监控系统目标检测应用[J]. 大连海洋大学学报,2021,36(5):842-850.
- [5] 贺琳,张雨,巴韩飞. 基于注意力机制和深度学习模型的外来海洋生物命名实体识别[J]. 大连海洋大学学报,2021,36(3):503-509.
- [6] 袁景凌,丁远远,盛德明,等. 基于视觉方面注意力的图像文本情感分析模型[J]. 计算机科学,2022,49(1):219-224.
- [7] 金雨芳,吴祥,董辉,等. 基于改进 YOLO v4 的安全帽佩戴检测算法[J]. 计算机科学,2021,48(11):268-275.
- [8] 周治平,张威. 结合视觉属性注意力和残差连接的图像描述生成模型[J]. 计算机辅助设计与图形学学报,2018(8):1536-1542,1553.
- [9] TRUONG Q T,LAUW H W. VistaNet:visual aspect attention network for multimodal sentiment analysis [C]//The Thirty-Third AAAI Conference on Artificial Intelligence (AAAI - 2019). Honolulu, Hawaii state, USA:AAAI,2019:306-312.
- [10] YANG C,WANG Z,SHEN H,et al. Multi-modality image manipulation detection [C]//IEEE International Conference on Multimedia and Expo (ICME). Shenzhen, China: IEEE, 2021: 1 - 6. DOI: 10. 1109/IC-

- ME51207. 2021. 9428232.
- [11] 袁红春, 张硕. 基于 Faster R-CNN 和图像增强的水下鱼类目标检测方法[J]. 大连海洋大学学报, 2020, 35(4): 612-619.
- [12] 许虞俊, 李晨. 基于 YOLO 优化的轻量级目标检测网络[J]. 计算机科学, 2021, 48(11A): 265-269.
- [13] 李婕, 周顺, 朱鑫潮, 等. 结合多通道注意力的遥感图像飞机目标检测[J]. 计算机工程与应用, 2022, 58(1): 209-217.
- [14] ZHANG T, WANG G, ZHUANG Y, et al. Feature enhanced centernet for object detection in remote sensing images [C]//Proceedings of IEEE International Geoscience and Remote Sensing Symposium. Waikoloa, HI, USA: IEEE, 2020: 1639 - 1642. DOI: 10. 1109/IGARSS39084. 2020. 9323116.
- [15] MANDAL V, MUSSAH A R, ADU-GYAMFI Y. Deep learning frameworks for pavement distress classification; a comparative analysis [C]//IEEE International Conference on Big Data. Atlanta, GA, USA: IEEE, 2020: 5577-5583. DOI: 10. 1109/BigData50022. 2020. 9378047.
- [16] HU J, SHEN L, ALBANIE S, et al. Squeeze-and-excitation networks [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 42(8): 2011-2023.
- [17] 覃晓, 黄呈铖, 施宇, 等. 基于卷积神经网络的图像分类研究进展[J]. 广西科学, 2020, 27(6): 587-599.
- [18] 王亚朝, 赵伟, 徐海洋, 等. 基于多阶段注意力机制的多种导航传感器故障识别研究[J]. 自动化学报, 2021, 47(12): 2784-2790.

SK-YOLOv5 Ocean Target Detection and Classification Algorithm Based on Attention Mechanism

LI Bin^{1,2}, WEN Lili^{2,3**}, WU Man², LIU Huaning⁴, XU Guilin²

(1. Department of Business Management, Guangxi Vocational and Technical College of Natural Resources, Chongzuo, Guangxi, 532100, China; 2. Guangxi Key Laboratory of Marine Environmental Science, Guangxi Key Laboratory of Human-machine Interaction and Intelligent Decision, New Technology Research Institute on Digital Twin, Guangxi Academy of Sciences, Nanning, Guangxi, 530007, China; 3. Information Industry Office, Guangxi Botanical Garden of Medicinal Plants, Nanning, Guangxi, 530023, China; 4. Business Department, Guangxi Zhuang Autonomous Region Organization Establishment and Performance Management Data Center, Nanning, Guangxi, 530012, China)

Abstract: The ocean target image based on remote sensing image has the characteristics of multi-scale, large shape change, dim color, unclear target boundary and fuzzy image. Therefore, the existing target detection algorithm needs to be improved to meet the needs of ocean target detection and classification of remote sensing image. Aiming at this problem, the Selective Kernel Networks (SKNet) attention module is introduced into the network architecture of You Only Look Once version 5 (YOLOv5), and a new SKYOLOv5 network is proposed to enhance the feature extraction and adaptive ability of the network to multi-scale complex ocean targets. Through comparative experimental tests, on the same ocean target data set, the overall detection and classification accuracy of the improved network is improved by about 9% compared with the original network.

Key words: attention mechanism; SKNet; YOLOv5; ocean target detection; feature extraction

责任编辑: 陆雁