

doi: 10.6046/zrzyg.2022425

引用格式:王振华,谭智联,李静,等. Re-YOLOX:利用 Resizer 改进的 YOLOX 近岸海域监测目标识别模型[J]. 自然资源遥感,2023,35(3):10-16. (Wang Z H,Tan Z L,Li J,et al. Re-YOLOX: A YOLOX model for identifying nearshore monitoring targets improved based on the Resizer model[J]. Remote Sensing for Natural Resources,2023,35(3):10-16.)

Re-YOLOX: 利用 Resizer 改进的 YOLOX 近岸海域监测目标识别模型

王振华^{1,2}, 谭智联^{1,2}, 李 静^{1,2}, 常英立¹

(1. 上海海洋大学信息学院,上海 201306; 2. 自然资源部海洋环境探测
技术与应用重点实验室,广州 510000)

摘要: 近岸海域监测包括自然环境监测和人类活动监测,其监测目标的高精准识别对海洋经济的健康发展、海洋环境的生态保护以及海洋防灾减灾等都有重要的作用。近岸海域监测目标具有多类型、多尺寸和不确定性等特征,现有识别模型在对近岸海域监测目标识别时,存在精度和效率欠佳、小目标漏检现象严重等问题。针对上述问题,利用可学习的图像调整模型(Resizer model)改进 YOLOX,提出了面向近岸海域监测目标的识别模型(Re-YOLOX),包括:①利用 Resizer model 加强模型训练,提升模型的特征学习能力和表达能力,提高模型的召回率;②改进 YOLOX 的特征金字塔融合结构,减少小目标识别的漏检问题。用无人机监测的近岸海域视频数据作数据集,以车辆、船只和堆砌物为监测目标,将提出的 Re-YOLOX 模型与 CenterNet, Faster R-CNN, YOLOv3 和 YOLOX 等模型进行比较。结果表明,Re-YOLOX 模型的平均预测精准率 mAP 可达 94.23%, 平均召回率 mR 可达 91.99%, 平均 $F1$ 值 $mF1$ 可达 89.67%, 均高于对比模型。综上所述,文章提出 Re-YOLOX 在保证目标识别效率的前提下提高了目标识别的精度,可为近岸海域管理提供技术支撑。

关键词: 近岸海域; 目标识别; YOLOX 算法; 无人机监测数据

中图法分类号: TP 79; TP 183 **文献标志码:** A **文章编号:** 2097-034X(2023)03-0010-07

0 引言

近岸海域是指 20 m 等深线至海岸线的海域^[1]。近岸海域监测目标既包括自然环境监测要素(如海、浪、流、潮和水边线等),又包括人类活动监测要素(如船只、车辆、垃圾、人和大坝等),近岸海域监测目标的高精准识别对海洋经济的健康发展、海洋环境的生态保护以及海洋防灾减灾等都具有重要的作用。受自然环境和人类活动的影响,近岸海域监测目标具有多类型、多尺寸和不确定性等特征,如何提高近岸海域监测目标的识别效率是制约海域管理智能化水平的关键问题之一。

深度学习是机器学习算法中的新技术,其动机在于创建一个神经网络并模拟人脑进行分析学习^[2]。Hinton 于 20 世纪 80 年代提出了适用于多层

感知机(multi-layer perceptron,MLP)的反向传播算法,使深度学习进入人们的视野^[3]。其后,LeCun 等^[4]于 1998 年提出了卷积神经网络(convolutional neural networks,CNN); Girshick 等^[5]于 2014 年提出了区域卷积神经网络(regions with CNN, R-CNN)目标识别模型; Redmon 等^[6]于 2015 年提出了 YOLO(you only look once)目标识别模型。深度学习逐渐被应用于目标识别领域,其目标识别的精度和效率逐渐优化。目前,已有学者将深度学习应用于海洋监测领域,关克平等^[7]提出将目标检测算法 CenterNet 与目标跟踪算法 DeepSORT 相结合的船舶交通流视觉图像统计方法;许延雷等^[8]提出了基于自适应阈值的改进 CenterNet 航拍图像目标检测算法;岳邦铮等^[9]设计了适用于(synthetic aperture Radar,SAR)影像的船舶目标检测特征提取模型;聂鑫等^[10]通过改进 YOLOv3 提出了一种在复杂

收稿日期:2022-11-02; 修订日期:2023-03-20

基金项目: 自然资源部海洋环境探测技术与应用重点实验室开放基金项目“基于深度学习的海岛与海岸带典型要素智能监测关键技术研究及试点应用”(编号:MESTA-2021-B007)和上海市地方院校能力建设项目“复杂潮汐环境影响下海岛(礁)地物信息提取与精度验证方法及其示范应用”(编号:19050502100)共同资助。

第一作者: 王振华(1982-),女,博士,教授,研究方向为海洋大数据处理及分析。Email: zh-wang@shou.edu.cn。

通信作者: 常英立(1977-),女,博士,副教授,研究方向为图像处理。Email: ylchang@shou.edu.cn。

场景下的船舶检测方法;齐亮等^[11]提出了一种改进的基于区域的快速卷积神经网络(faster region with convolutional neural network, Faster R-CNN)的船舶目标检测方法;盛明伟等^[12]通过在YOLOv3中引入注意力机制,提出了一种改进YOLOv3的船舶目标检测算法。

上述研究将深度学习应用于海域监测,提高了海域监测目标识别的精度和效率。但现有的识别模型多针对某一特定的监测目标,直接将其应用于多类型近岸海域监测目标的同步识别,存在精度和效率欠佳现象,特别是面对小目标的识别其漏检或错检现象严重。本文针对近岸海域监测目标的特点,提出了一种近岸海域监测目标识别模型(Re-YOLOX),提高多类型近岸海域监测目标同步识别

的精度和效率,特别是提高小目标的识别精度。

1 近岸海域监测目标识别模型(Re-YOLOX)

图1所示为利用可学习的图像调整模型(Resizer model)改进的YOLOX近岸海域监测目标识别模型(Re-YOLOX),其中A和B分别为Resizer model和YOLOX,包括:①利用Resizer model^[13]加强模型的训练,提高模型的召回率,如图1中红色虚线框A所示;②改进YOLOX^[14]模型特征金字塔融合结构,提高Re-YOLOX模型多尺度特征的融合能力,如图1中黑色虚线框B所示。

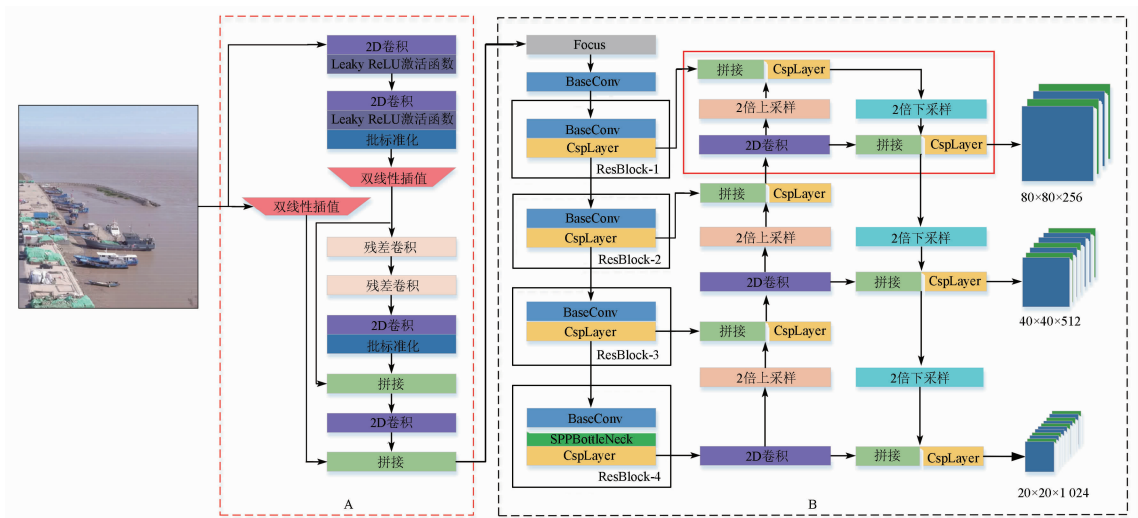


图1 近岸海域监测目标识别模型(Re-YOLOX)结构框架

Fig.1 Structure diagram of target recognition model (Re-YOLOX)

图1中Focus由4种不同的空间压缩方式和一个2D卷积层组成的压缩卷积层;BaseConv是由2D卷积层、批归一化层(batch normalization, BN)以及SiLU激活函数组成的卷积单元;CspLayer由残差卷积模块组成的卷积层;SPPBottleNeck由2个BaseConv单元和4个大小不同的池化层组成的金字塔池化层。

YOLOX模型由旷视科技于2021年提出,该模型保持了YOLOv5的网络结构,其主干特征提取网络(Backbone)采用交叉阶段部分连接(cross stage partial network, CSPNet)^[15]并结合Darknet-53的CSPDarknet结构。YOLOX模型将空间金字塔池化网络(spatial pyramid pooling networks, SPPNet)^[16]应用到主干特征提取网络中,代替了原有池化层,提高了目标检测的准确率。同时,在特征层融合部分,将上采样和下采样结合,实现了多尺度特征的融合^[17]。YOLOX应用双解耦头结构,对来自主干特

征提取网络的特征层分别进行分类操作和定位回归操作,加快了模型的收敛速度,提高了检测精度。与单阶段目标识别模型比较,YOLOX模型无须生成检测目标先验框,可直接对特征点进行回归分析,精简了模型参数并提高了检测精度。与双阶段目标识别模型相比,YOLOX模型无须先生成候选区域,可直接对图像进行特征提取,亦提高了检测速度。因此,面向多类型、多尺寸和不确定性等特征的近岸海域监测目标识别,本文将YOLOX模型作为基础模型。

1.1 利用 Resizer model 加强模型训练

YOLOX模型采用固定尺度的图像输入,通过对图像进行缩放、镜像和对比度变化等操作提高模型的泛化性。此过程易导致图像细节特征丢失,当图像存在重叠面积的2个或多个目标,或存在较多小面积目标时,细节特征丢失现象会加剧,进而引起漏检测或错检测等问题。为解决该问题,本文引入Resizer model加强模型的训练。

Resizer model 是谷歌研究院于 2021 年提出的图像预处理模型,其借鉴 Resnet^[18] 的跳跃连接思想,将图像经过原始线性插值方法与学习后线性插值方法融合处理后的特征信息,整合到了 CNN 中。在进行模型训练时,Resizer model 可根据识别模型每轮训练的反向传播结构优化其图像调整参数,使 Resizer model 成为一种可学习的图像调整模型。在 Re-YOLOX 模型中,图像直接输入 Resizer model,通过卷积运算和双线性插值方法实现图像预处理,将预处理的结果直接输入到图 1(B) 所示的识别模型中进行特征提取。Resizer model 根据其识别模型每次的反向传播结果,优化调整卷积核的参数,进一步调整双线性插值方法对图像的预处理。经过 Resizer model 处理后的图像尺寸保持一致,且保留了图像的细节特征,加强了模型的特征提取能力。

1.2 改进 YOLOX 模型的特征金字塔融合结构

YOLOX 模型的主干特征提取网络较深,多达上百层。小目标的特征信息会在多次卷积操作后减弱甚至消失。为提升 Re-YOLOX 模型对小目标物体的识别能力,在 YOLOX 模型的特征金字塔融合结构中新增一个高分辨率特征层,新增特征层融合结构如图 1(B) 中红色实线框所示。

YOLOX 模型的主干特征提取网络中 ResBlock - n 表示第 n 个组合卷积结构,每个组合卷积结构由 BaseConv 单元和 CspLayer 层组成。其中 BaseConv 单元由 2D 卷积层、BN 层以及 SiLU 激活函数组成; CspLayer 层由残差卷积模块组成。如在输入图像宽高为 640 像素 × 640 像素时,主干网络中的第一个组合卷积结构 ResBlock - 1 的输出特征层的宽高为 160 像素 × 160 像素,其中拥有大量小目标物体的特征信息。Re-YOLOX 模型将 160 像素 × 160 像素的特征层与来自特征金字塔底层经过上采样得到的特征层,在通道维度上进行拼接,并将拼接得到的特征层用于密集卷积操作,后再一次进行下采样特征提取。在得到经下采样特征提取的特征层后,Re-YOLOX 模型将其与特征金字塔上采样过程中具有相同宽高尺寸的特征层再次拼接,实现多尺度特征融合。改进后的特征金字塔融合结构充分利用了主干特征提取网络对图像提取的浅层特征信息,对小目标更加敏感。

1.3 训练 Re-YOLOX 模型及设置参数

Re-YOLOX 模型通过 2 个阶段进行训练。在第一阶段,使用 YOLOX 原有的图像预处理方法对模型进行训练,在模型收敛后,保存权重文件,作为第二训练阶段的初始化模型。在第二阶段,用第一训练阶段保存的权重文件对模型参数初始化后,使

用 Resizer model 对模型进行加强训练,使模型在第一训练阶段的基础上,继续学习目标物体的更多细节特征。第一和第二训练阶段的 Epoch 都为 100,为加强对 Resizer model 和添加高分辨率特征层的特征金字塔融合结构的训练。前 50 Epoch 冻结模型的主干特征提取网络,第一阶段仅对模型的多尺度特征融合结构进行训练;第二阶段对多尺度特征融合结构和 Resizer model 进行训练。2 个训练阶段其余参数设置相同,具体包括 Momentum = 0.9, Decay = 0.005, batch size = 4, 学习率 LR 设置为 0.001。后 50 Epoch 对全部参数进行训练,为提升模型的稳定性,后 50 Epoch 的学习率 LR 设置为 0.000 1。优化器为 SGD,使用余弦退火学习机制。

2 实验设计

2.1 数据集

选取无人机航拍的上海市近岸海域视频数据,采用视频抽帧方式构建训练和测试数据集(图 2)。实验数据为抽取的 1 010 张图像,大小为 960 像素 × 1 080 像素,随机选择 80% 作为训练集,10% 为验证集,10% 为测试集。以车辆、船只和堆砌物为近岸海域监测目标,用开源标注工具 LabelImg 对监测目标进行标注,共标注监测目标 15 425 个。图 3 所示为 15 425 个目标物体的尺寸统计,横、纵坐标分别

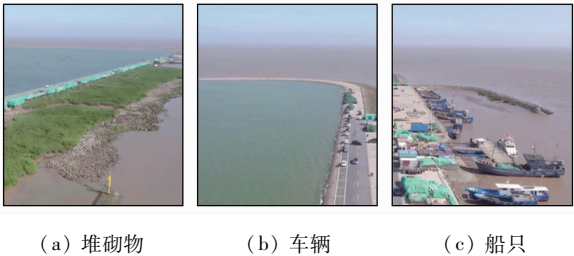


图 2 近岸海域监测数据示意图

Fig. 2 Schematic diagram monitoring data in offshore area

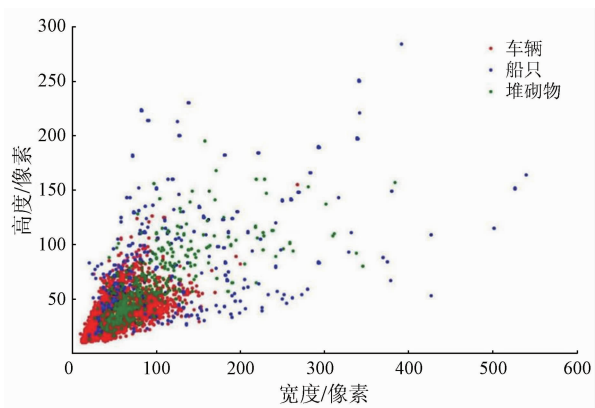


图 3 监测目标尺寸统计

Fig. 3 Distribution of detection target sizes

为目标的宽度和高度,单位为像素,根据 COCO 数据集标准,将像素面积小于 32 像素 × 32 像素的目标划分为小目标,像素面积大于 96 像素 × 96 像素的目标划分为大目标^[19]。由图 3 可知,数据集上的中小型目标占比较大,大多数目标集中分布于 0 ~ 100 像素之间。

2.2 实验环境

实验硬件配置为 Inter Core i7-9700 CPU 处理器, GeForce GTX 2060 显卡的服务器。操作系统为 64 位 Windows 10,深度学习框架为 pytorch,可视化工具为 Matplotlib。

2.3 评价指标

选用精准率 (precision, P)、召回率 (recall, R)、平均均值准确率 (mean average precision, mAP) 和平均召回率 (mean recall, mR)、平均 F1 值 (mean F1, mF1) 和模型的检测速度 (frame per second, FPS) 对识别模型的性能进行评价^[20]。IOU 表示真实标注框与检测框的交并比,本文所有算法模型的 IOU 均为 0.5。P 和 R 的公式分别为:

P = TP / (TP + FP) , (1)

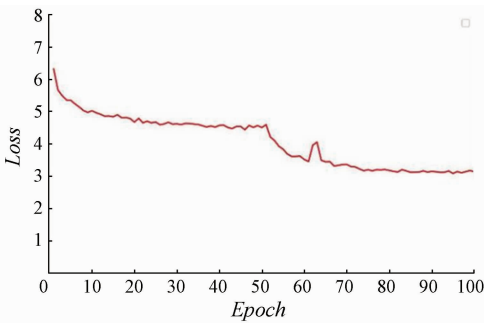
R = TP / (TP + FN) , (2)

式中: TP 为模型将正样本预测为正样本的个体数量; FP 为模型将负样本预测为正样本的个体数量; FN 为模型将正样本预测为负样本的数量。

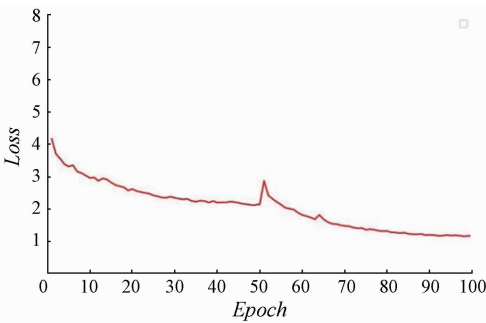
mAP 的计算公式为:

AP = ∫₀¹ P(R) dR , (3)

mAP = (∑_{k=1}^C AP(k)) / C , (4)



(a) YOLOX 损失函数变化曲线



(b) Re-YOLOX 损失函数变化曲线

图 4 损失函数变化曲线对比
Fig. 4 Comparison of Loss curve

表 1 所示为 Re-YOLOX 模型中各改进模型的性能比较,其中, YOLOX + resizer 表示将 Resizer 加

mR = (∑_{k=1}^C R(k)) / C , (5)

式中: C 为数据集上的检测类别数; AP(k) 和 R(k) 分别为第 k 个类别的检测精度和召回率。

F1 是对识别模型预测精准率和召回率的综合考量,用来评价识别模型的识别精度、避免精准率和召回率之间的相互影响, mF1 的计算公式为:

F1 = 2 * (P * R) / (P + R) , (6)

mF1 = (∑_{k=1}^C F1(k)) / D , (7)

式中: D 为数据集中样本类别个数; F1(k) 为第 k 个类别的 F1 值。

FPS 是指识别模型每秒处理的图像数,用于评价模型的检测速度。

3 实验对比与分析

本节分别设计了消融实验和对比实验对提出的 Re-YOLOX 模型进行可行性和性能验证。

3.1 消融实验

图 4 给出了 YOLOX 模型与 Re-YOLOX 模型损失函数变化曲线比较。由图 4(a)可看出, YOLOX 模型在训练过程中损失函数值随着迭代次数的增加逐渐降低,在 Epoch = 80 时,损失函数变化曲线趋于稳定,表明模型逐渐收敛。由图 4(b)可看出, Re-YOLOX 模型在原有数据集上继续训练,损失函数值继续下降,在 Epoch = 95 时,损失函数值最低,此时得到的训练模型为最佳网络模型。结果表明,利用 Resizer model 对 YOLOX 模型进行加强训练,提升了模型的学习能力。

入 YOLOX 模型的图像预处理中; YOLOX + add 表示在 YOLOX 模型的特征金字塔融合结构中添加高

分辨率特征层。由表 1 中可知,YOLOX + resizer 模型和 YOLOX + add 模型均提升了原有 YOLOX 模型的 mAP 值、mR 值和 mF1 值,但其检测速度略有下降。YOLOX + resizer 模型通过增设 Resizer 模块,强化了 YOLOX 模型的特征提取能力,船只的精准率和召回率提升明显,分别增长 4.32 和 4.94 百分点;同时其 mAP 值提高 2.14 百分点,mR 值提高 4.26 百分点。YOLOX + add 模型通过在特征金字塔融合

结构中添加高分辨率特征层,降低了小目标在深度卷积过程中信息损失对模型的影响,车辆和船只的召回率分别提高了 7.23 和 5.97 百分点,同时其 mR 值提高了 5.69 百分点。而本文提出的 Re - YOLOX 对小目标更加敏感,特别是对尺寸较小的车辆和船只,相较于 YOLOX,召回率分别提高了 8.07 百分点和 8.42 百分点。

表 1 消融实验测试结果
Tab. 1 Test results of ablation

模型	P/%			R/%			mAP/%	mR/%	mF1/%	FPS/ (张·s ⁻¹)
	车辆	堆砌物	船只	车辆	堆砌物	船只				
YOLOX	94.09	86.58	75.24	83.09	87.16	83.16	89.79	84.47	84.67	16.36
YOLOX + resizer	88.69	88.27	79.56	89.87	88.23	88.10	91.93	88.73	87.04	15.44
YOLOX + add	89.65	89.49	80.34	90.32	91.03	89.13	93.28	90.16	88.24	14.32
Re - YOLOX	88.91	92.00	81.31	91.16	93.24	91.58	94.23	91.99	89.67	13.91

3.2 对比实验

将提出的 Re - YOLOX 模型与 CenterNet, Faster R - CNN 和 YOLOv3 模型进行比较,其中 Faster R - CNN 与 YOLOv3 是基于 anchor 设计的代表性模型,CenterNet 基于 anchor - free 设计的代表性模型。实

验以 4 幅不同的近岸海域监测数据为例,表 2 给出了不同模型的识别结果,其中红色框、蓝色框和绿色框分别表示车辆、船只和堆砌物的识别。表 3 给出了不同模型的识别结果评价指标。

表 2 不同模型的目标识别结果
Tab. 2 Target recognition results of different models

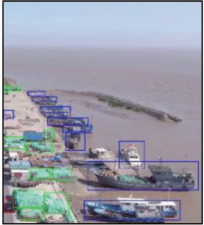
图像序号	标签	CenterNet	Faster R - CNN	YOLOv3	Re - YOLOX
a					
					
					
					

表 3 不同模型的评价指标

Tab.3 Comparison of evaluation indicators of different models

模型	P/%			R/%			mAP/%	mR/%	mF1/%	FPS/ (张·s ⁻¹)
	车辆	堆砌物	船只	车辆	堆砌物	船只				
CenterNet	95.38	97.56	96.72	51.07	54.05	62.11	80.48	55.74	71.00	14.38
Faster R-CNN	46.42	74.84	77.00	36.35	78.38	81.05	59.02	65.26	65.67	7.41
YOLOv3	77.38	73.87	73.63	72.71	61.19	52.34	68.28	58.04	67.67	21.44
Re-YOLOX	88.91	92.00	81.31	91.16	93.24	91.58	94.23	91.99	89.67	13.91

由表 2 和表 3 可看出,Re-YOLOX 模型对于位置邻近且有重叠遮挡的目标和小尺寸目标具有更好的识别效果,如面向图像 a 的目标识别,只有 Faster R-CNN 和 Re-YOLOX 模型将右下方的 2 只船只识别开来,但 Re-YOLOX 预测框的回归效果明显优于 Faster R-CNN 预测框;面向图像 c 的目标识别,其他模型均存在明显漏检现象,尤其是 CenterNet 模型漏检现象最严重,Re-YOLOX 模型识别出了所有船只和车辆。同时,Re-YOLOX 模型具有最高的 mAP, mR 和 mF1 值,分别为 94.23%,91.99%和 89.67%。但 Re-YOLOX 模型的 FPS 略低于 CenterNet 和 YOLOv3,高于 Faster R-CNN 模型。总体而言,Re-YOLOX 模型在保证目标识别效率的前提下提高了目标识别精度。

4 结论

本文针对近岸海域监测目标识别精度要求高和实时性要求强等现实需求,兼顾近岸海域监测目标的多尺度及小目标多等特性,提出了一种近岸海域监测目标识别模型(Re-YOLOX 模型),包括:利用可学习的图像调整模型加强 Re-YOLOX 模型训练,提高模型的特征学习能力和表达能力;改进 Re-YOLOX 模型的特征金字塔融合结构,缓解小目标识别的漏检问题。

用无人机监测的近岸海域视频数据作数据集,以车辆、船只和堆砌物为监测目标,通过消融实验和对比实验对提出的 Re-YOLOX 模型进行可行性和性能验证。消融实验结果表明,在 Re-YOLOX 模型中 Resizer 模块的增设,强化了 YOLOX 模型的特征提取能力;在特征金字塔融合结构中添加高分辨率特征层,提高了小目标的识别能力。通过与 CenterNet, Faster R-CNN, YOLOv3 和 YOLOX 模型的对比实验结果表明,Re-YOLOX 模型可满足近岸海域监测目标的识别需求,其 mAP, mR 和 mF1 值均优于其他识别模型。

参考文献(References):

[1] 何 荣,石海明,屠建波,等. 海洋工程对天津近岸海域环境的

影响研究[J]. 海洋开发与管理,2019,36(5):63-66.

He R, Shi H M, Tu J B, et al. The influence of ocean engineering in the coastal area of Tianjin[J]. Ocean Development and Management, 2019, 36(5): 63-66.

[2] 周 扬,基于深度学习的目标跟踪算法研究[D]. 扬州: 扬州大学,2019.

Zhou Y. Research on target tracking algorithm based on deep learning[D]. Yangzhou: Yangzhou University, 2019.

[3] 范航恺,基于卷积神经网络的序列特异性预测研究[D]. 昆明: 云南大学,2016.

Fan H K. Prediction of sequence specificity based on convolutional neural network[D]. Kunming: Yunnan University, 2016.

[4] LeCun Y, Bottou L. Gradient-based learning applied to document recognition[J]. Proceedings of the IEEE, 1998, 86(11): 2278-2324.

[5] Girshick R, Donahue J, Darrell T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]// 2014 IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA, 2014: 580-587.

[6] Redmon J, Divvala S, Girshick R, et al. You only look once: Unified, real-time object detection[C]// 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 2016: 779-788.

[7] 关克平,韩 笑,蒋 宇. 基于 tbd 策略的船舶交通流视觉图像统计方法[J]. 上海海事大学学报, 2021, 42(2): 40-44, 95.

Guan K P, Han X, Jiang Y. Visual image statistics method of ship traffic flow based on tbd strategy[J]. Journal of Shanghai Maritime University, 2021, 42(2): 40-44, 95.

[8] 许延雷,梁继然,董国军,等. 基于改进 CenterNet 的航拍图像目标检测算法[J]. 激光与光电子学进展, 2021, 58(20): 192-201.

Xu Y L, Liang J R, Dong G J, et al. Target detection algorithm for aerial images based on improved CenterNet[J]. Laser and Optoelectronics Progress, 2021, 58(20): 192-201.

[9] 岳邦铮,韩 松. 基于改进 Faster R-CNN 的 SAR 船舶目标检测方法[J]. 计算机与现代化, 2019, 7(9): 90-95.

Yue B Z, Han S. A SAR ship detection method based on improved Faster R-CNN[J]. Computer and Modernization, 2019, 7(9): 90-95.

[10] 聂 鑫,刘 文,吴 巍. 复杂场景下基于增强 YOLOv3 的船舶目标检测[J]. 计算机应用, 2020, 40(9): 2561-2570.

Nie X, Liu W, Wu W. Ship target detection based on enhanced YOLOv3 in complex scenes[J]. Journal of Computer Applications, 2020, 40(9): 2561-2570.

[11] 齐 亮,李邦昱,陈连凯. 基于改进的 Faster R-CNN 船舶目标检测算法[J]. 中国造船, 2020, 61(s1): 40-51.

Qi L, Li B Y, Chen L K. Ship target detection algorithm based on improved Fast R – CNN [J]. Shipbuilding of China, 2020, 61 (s1) : 40 – 51.

[12] 盛明伟, 李 俊, 秦洪德, 等. 基于改进 YOLOv3 的船舶目标检测算法 [J]. 导航与控制, 2021, 20 (2) : 95 – 109.

Sheng M W, Li J, Qin H D, et al. Ship target detection algorithm based on improved YOLOv3 [J]. Navigation and Control, 2021, 20 (2) : 95 – 109.

[13] Talebi H, Milanfar P. Learning to resize images for computer vision tasks [C] // 2021 IEEE/CVF International Conference on Computer Vision (ICCV), Montreal, QC, Canada, 2021 : 487 – 496.

[14] Ge Z, Liu S, Wang F, et al. YOLOX: Exceeding YOLO series in 2021 [EB/OL]. [2021 – 08 – 06]. <https://arxiv.org/abs/2107.08430>.

[15] Wang C Y, Liao H Y M, Wu Y H, et al. CSPNet: A new backbone that can enhance learning capability of CNN [C] // in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, 2020 : 390 – 391.

[16] He K, Zhang X, Ren S, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37 (9) : 1904 – 1916.

[17] 刘 鑫, 陈思溢, 陈小龙, 等. 基于深度学习的深层次多尺度特征融合目标检测算法 [J]. 激光与光电子学进展, 2021, 58 (12) : 304 – 312.

Liu X, Chen S Y, Chen X L, et al. Deep multi – scale feature fusion target detection algorithm based on deep learning [J]. Laser and Optoelectronics Progress, 2021, 58 (12) : 304 – 312.

[18] Szegedy C, Ioffe S, Vanhoucke V, et al. Inception – v4, inception – resnet and the impact of residual connections on learning [C] // Thirty – first AAAI Conference on Artificial Intelligence, 2017 : 4278 – 4284.

[19] Lin T Y, Maire M, Belongie S, et al. Microsoft COCO: Common objects in context [C] // Proceedings of the European Conference on Computer Vision, 2014 : 740 – 755.

[20] 陈朋弟, 黄 亮, 夏 炎, 等. 基于 Mask R – CNN 的无人机影像路面交通标志检测与识别 [J]. 国土资源遥感, 2020, 32 (4) : 61 – 67. doi: 10. 6046/gtzyyg. 2020. 04. 09.

Chen P D, Huang L, Xia Y, et al. Detection and recognition of road traffic signs in UAV images based on Mask R – CNN [J]. Remote Sensing for Land and Resources, 2020, 32 (4) : 61 – 67. doi: 10. 6046/gtzyyg. 2020. 04. 09.

Re – YOLOX: A YOLOX model for identifying nearshore monitoring targets improved based on the Resizer model

WANG Zhenhua^{1,2}, TAN Zhilian^{1,2}, LI Jing^{1,2}, CHANG Yingli¹

(1. College of Information Technology, Shanghai Ocean University, Shanghai 201306, China; 2. Key Laboratory of Marine Environmental Survey Technology and Application, Ministry of Natural Resources, Guangzhou 510000, China)

Abstract: Nearshore monitoring covers natural environments and human activities. High – accuracy identification of nearshore monitoring targets significantly influences the healthy development of the marine economy, the ecological protection of marine environments, and the prevention and mitigation of marine disasters. The nearshore monitoring targets feature multiple types, diverse sizes, and uncertainty. The existing identification models suffer low accuracy, low efficiency, and severe omission of small targets. This study proposed an identification model (Re – YOLOX) for nearshore monitoring targets by improving YOLOX using a learnable image resizer model (the Resizer model). First, the model training was intensified using the Resizer model to improve the feature learning and expression abilities and the recall rate of the Re – YOLOX model. Then, the feature pyramid fusion structure of the YOLOX algorithm was improved to reduce the omission of small targets in the identification. With the nearshore video data from UAV monitoring as the data set and cars, ships, and piles as monitoring targets, this study compared the Re – YOLOX model with other models, including CenterNet, Faster R – CNN, YOLOv3, and YOLOX. The results show that the Re – YOLOX model yielded a mean average precision of 94. 23% , a mean recall of 91. 99% , and a mean F1 score of 89. 67% , all of which were higher than those of the other models. In summary, the Re – YOLOX model can improve the target identification accuracy while ensuring target identification efficiency, thus providing technical support for managing nearshore seas.

Keywords: nearshore sea; target identification; YOLOX algorithm; UAV monitoring data

(责任编辑: 张 仙)