

# 基于改进 YOLOv5 的激光雷达水下障碍物检测

董帮林<sup>1,2</sup> 尚建华<sup>1\*</sup> 贺岩<sup>2\*</sup> 吕德亮<sup>2</sup> 侯春鹤<sup>2</sup> 阮英杰<sup>2</sup>

1 东华大学 信息科学与技术学院

2 中国科学院上海光学精密机械研究所空天激光技术与系统部王之江创新中心

DOI:10.12238/acair.v3i2.13502

**[摘要]** 机载激光雷达已被广泛应用于水下地形测绘等领域,针对水下障碍物探测及识别精度低、速度慢等问题,本文在YOLOv5算法的基础上设计了目标分类和目标检测两个网络,实现了对水下障碍物的探测和识别。借助ConvNeXt Block模块对分类和检测两个模型的主干网络进行了优化,加强了网络对激光雷达图像中小目标特征的提取能力,提高了目标分类和检测的精度;选用EIoU损失函数作为检测网络的边框回归损失函数,提升了模型的收敛速度和目标预测精度。实验结果表明,改进后的分类模型的分类准确率可达96.8%,比原分类模型的准确率提高了3.2%;改进后的检测网络的精确度和平均精度均值比原有检测网络分别提高了3%和2.8%。

**[关键词]** YOLOv5算法; 损失函数; 图像处理; 激光雷达

**中图分类号:** TP751; TN958.98 **文献标识码:** A

## Research on Underwater Obstacle Detection with LiDAR Based on Improved YOLOv5

Banglin Dong<sup>1,2</sup> Jianhua Shang<sup>1\*</sup> Yan He<sup>2\*</sup> Deliang Lv<sup>2</sup> Chunhe Hou<sup>2</sup> Yingjie Ruan<sup>2</sup>

1 College of Information Science and Technology, Donghua University

2 Wangzhijiang Innovation Center for Laser, Aerospace Laser Technology and System Department, Shanghai Institute of Optics and Fine Mechanics

**[Abstract]** Airborne Lidar has been widely used in underwater terrain mapping and other fields. Aiming at the problems of low accuracy and slow speed of underwater obstacle detection and identification, two networks of target classification and target detection based on YOLOv5 algorithm were designed in order to realize the detection and identification of underwater obstacles. With the ConvNeXt Block module, the backbone networks of both the classification and detection models were optimized, enhancing the networks' ability to extract features of small targets in LiDAR images and improving the accuracy of target classification and detection. The EIoU loss function was selected as the Bounding Box Regression loss function for the detection network, which improved the model's convergence speed and target prediction accuracy. Experimental results show that the classification accuracy of the improved classification model can reach 96.8%, representing a 3.2% increase compared to the original classification model. Additionally, the precision and mean average precision of the improved detection network increased by 3% and 2.8%, respectively, compared to the original detection network.

**[Key words]** YOLOv5 algorithm; loss function; image processing; laser radar

## 引言

随着我国港口与水上货物运输业的迅猛发展,水上交通运输领域面临的突发事件数量正在持续攀升。其中,水下障碍物已成为水上及水下活动中不可忽视的潜在安全风险,迫切需要采取高效且精准的探测和识别手段应对这些航行障碍<sup>[1]</sup>。无论是港口设施的建设,还是航道的维护和治理,对水下障碍物实施精准探测与识别都有着迫切的需求。因此,发展水下障碍物的探测

和识别具有十分重要的意义。

目前,扫描式激光雷达系统普遍采用波形处理作为其数据处理手段。尽管波形处理方法能够较准确的拟合波形并确定水面及障碍物的位置,但通常需要根据实际情况不断修正和优化拟合波形。2024年,阮英杰等人<sup>[2]</sup>提出一种针对激光雷达水下障碍物探测的剖面图像处理方法,可使有关水下障碍物的分析从原始的波形处理转换为图像处理。然而,传统的图像处理方法在

水下障碍物的识别精度和效率方面存在一定的局限性。为了解决这一问题,本文采用深度学习目标检测算法YOLOv5,以实现激光雷达水下障碍物的高效、高精度探测和识别。

## 1 YOLOv5算法的改进

1.1 引入ConvNeXt Block模块优化分类和检测模型的主干网络

针对原主干网络在提取激光雷达图像小目标特征能力不足且检测精度偏低的问题,本文提出采用ConvNeXt Block模块全面替换YOLOv5分类与检测原主干网络中的C3模块,进而实现对主干网络的优化,提升其性能。

2022年, Facebook AI Research在ResNet的基础上,借鉴Swin Transformer网络,提出了ConvNeXt网络结构。作为一种纯卷积神经网络,ConvNeXt展现出卓越的网络学习能力,在同等计算量下,它拥有更高的准确率及更快的推理速度<sup>[3]</sup>。目前,ConvNeXt在多种视觉任务中表现出色,已被广泛应用于图像分类、目标检测、语义分割以及实例分割等领域。

ConvNeXt Block模块是ConvNeXt网络的主要构成模块,融合了逆瓶颈结构和7×7大卷积核。原始的残差结构采用“两端大,中间小”的卷积块设计,产生通道数减少时图像特征损失的问题。为了解决这一问题,ConvNeXt Block模块采用“两端小,中间大”的逆瓶颈结构,该结构不仅能有效提取高维及深层特征信息,还能有效解决激光雷达图像信息在降采样过程中的特征丢失问题,进而提高了对激光雷达图像中小目标的检测精度。

### 1.2 检测网络损失函数的改进

神经网络的损失函数是衡量真实值和预测值之间差异的关键指标。在YOLOv5中,损失函数包括置信度损失、分类损失和边框回归损失三部分。对于边框回归损失,YOLOv5默认选用完全交并比损失函数(LCIoU)进行预测,其表达式如式(1)所示。

$$L_{CIoU} = 1 - R_{IoU} + \frac{\rho^2(\mathbf{b}, \mathbf{b}^{gt})}{c^2} + \alpha v, \quad \text{其中} \quad \begin{cases} \alpha = \frac{v}{(1 - R_{IoU}) + v} \\ v = \frac{4}{\pi^2} \left( \arctan \frac{\omega^{gt}}{h^{gt}} - \arctan \frac{\omega}{h} \right)^2 \\ R_{IoU} = \frac{A \cap B}{A \cup B} \end{cases} \quad (1)$$

式中, A, B分别为预测框和真实框的面积,  $R_{IoU}$  为预测框与真实框的交集区域和并集区域之比,  $\rho$  为两个中心点间的欧氏距离,  $\alpha$  为权重系数,  $c$  表示含预测框和真实框的最小封闭区域的对角线长度,  $v$  是长宽比一致性参数,  $\mathbf{b}^{gt}$  为真实框的中心点,  $\mathbf{b}$  为预测框的中心点,  $h^{gt}$  为真实框的高度,  $h$  为预测框的高度,  $\omega^{gt}$  为真实框的宽度、 $\omega$  为预测框的宽度。

CIoU损失函数由于未充分平衡难易样本的影响,并且忽视了宽高比,这不利于模型的优化。为了解决这些问题,本文选择有效交并比损失函数(EIoU)来预测边框回归损失。EIoU在CIoU的基础上进行了改进,同时还引入了焦点损失机制,进而进一步

提升模型的训练效果和性能。EIoU损失函数的表达式如式(2)所示。

$$L_{EIoU} = L_{IoU} + L_{CDL} + L_{WHL} = 1 - R_{IoU} + \frac{\rho^2(\mathbf{b}, \mathbf{b}^{gt})}{c^2} + \frac{\rho^2(\omega, \omega^{gt})}{c_\omega^2} + \frac{\rho^2(h, h^{gt})}{c_h^2} \quad (2)$$

式中,  $L_{IoU}$  为重叠损失、 $L_{WHL}$  表示宽高损失、 $L_{CDL}$  代表中心距离损失。 $C_h$  与  $C_\omega$  分别代表预测框和真实框的最小外接矩形框的高度与宽度。EIoU损失函数通过引入“宽高损失”的措施,更加精细地调整了锚框和目标框的尺寸,进而将宽度与高度的差异值最小化<sup>[5]</sup>。这一改进不仅有利于模型收敛速度的加快,还能显著提高模型的预测精度。

## 2 实验和结果分析

### 2.1 实验环境

本文基于Python 3.10和PyTorch 1.12.1深度学习库构建网络模型,在配备NVIDIA GeForce RTX 2080Ti GPU的硬件平台上进行训练和测试。为了加快模型在训练过程中的收敛速度,训练时采用随机梯度下降优化器更新模型参数。有关检测网络的超参数配置如下: Epoch设为300, Batch-size为16, Patience设为100, Img-size设为256256, Learning-rate设为0.01。

### 2.2 数据集

本文使用的激光雷达水下障碍物图像数据集是从机载海洋激光雷达外场实验获得的,其获取成本较高。本次外场实验在杭州市淳安县千岛湖进行,实验采用的是由中国科学院上海光学精密机械研究所研制的无人机载海洋激光雷达系统。实验中,选用高度为80cm、直径为60cm的圆柱形桶作为激光雷达水下障碍物探测的目标对象,并尽量保持该圆桶完全没入水中且其侧面和湖面保持水平状态,以确保实验数据的准确性和可靠性。

激光雷达扫描水下障碍物时,其收集到的原始数据为二进制格式。为了直观展现水下障碍物的相关信息,首先将这些二进制数据进行图像拼接并做水面校正处理,进而转化为水下障碍物回波能量剖面图像数据<sup>[2]</sup>。根据上述回波能量剖面图,可以清晰看出激光雷达是否成功探测到障碍物。此外,本次实验中共安排了8个架次的飞行实验,经过数据处理之后,共筛选出6540张图像用于目标分类和检测任务。然而,受水体后向散射的影响,激光雷达回波信号较弱,所得到的含有障碍物信息的图像数量相对较少。为了训练出性能优异的分类和检测网络模型,除了采用迁移学习<sup>[4]</sup>的方法外,还需要在模型训练和测试过程中采取一定的策略和方法。

由于该系统包含分类和检测两个独立网络,需要分别对分类和检测两个模型进行训练,因此需分别制作分类数据集和检测数据集。为了提高分类网络和检测网络在学习含障碍物和不含障碍物图像特征时的效率,并考虑激光雷达图像中目标区域相对较小的实际情况,本文在单独训练分类和检测模型时,采用中心剪裁方法将分辨率为913×761的原始图像裁剪为分辨率为

256×256的图像,以便两个网络专注学习相关的图像特征,进而能够训练出性能更加出色的模型。

### 2.3评价指标

为了全面且深入地评估改进后的检测模型性能,本文采用模型参数量(Params)、浮点运算次数(FLOPs)、精确率(P)、召回率(R)、平均精度(AP)以及平均精度均值(mAP)等关键评价指标评价改进后模型的性能表现。

$$P = \frac{TP}{TP + FP} \quad (3)$$

$$R = \frac{TP}{TP + FN} \quad (4)$$

$$AP = \int_0^1 P(R) dR \quad (5)$$

$$mAP = \frac{1}{N} \sum_{i=0}^N AP(i) \quad (6)$$

式中,真正例TP代表实际包含目标且被正确识别为包含目标的图像数量;假正例FP代表实际不含目标但被错误识别为包含目标的图像数量;假反例FN代表实际含目标但被判为不包含目标的图像数量。

### 2.4消融实验

为了分析改进模块对YOLOv5模型在目标检测与目标分类任务中的性能影响,本文针对目标检测与目标分类任务分别设计了消融实验。由于本文对激光雷达水下障碍物探测精度要求较高,所以选择较为复杂的YOLOv51检测模型和YOLOv51-c1s分类模型作为基准。

在目标检测任务中,采用自制的激光雷达水下障碍物图像检测数据集对以下四种模型进行训练:原始的YOLOv51模型、以ConvNeXt Block模块优化的模型(YOLOv51-CN)、以EIoU损失函数优化的模型(YOLOv51-EU),以及同时融合了ConvNeXt Block模块和EIoU损失函数的优化模型(YOLOv51-CN-EU)。通过对比上述四种模型在测试集上的表现,评估了不同改进措施对原始YOLOv51模型性能的提升效果,测试结果如表1所示。

表1 目标检测消融实验

Model	Parameters/10 <sup>6</sup>	FLOPs/10 <sup>9</sup>	P/%	R/%	mAP/%
YOLOv51 (base)	46.5	109.0	93.2	84.6	92.1
YOLOv51-CN	53.0	130.1	95.3	88.3	93.6
YOLOv51-EU	46.5	109.0	94.7	87.5	93.2
YOLOv51-CN-EU	53.0	130.1	96.2	89.7	94.9

根据表1可知,相比于YOLOv51基准模型,主干网络改进模型(YOLOv51-CN)的精确率提高了2.1%,召回率提高了3.7%,平均精度均值也提高了1.5%,然而,模型参数量和浮点运算次数均有所增加;损失函数改进模型(YOLOv51-EU)的精确率提高了1.5%,

召回率提高了2.9%,平均精度均值提高了1.1%,表明损失函数的改进可有效提高模型的检测精度;整体改进模型(YOLOv51-CN-EU)的性能提升明显,精确率提高了3%,召回率提高了5.1%,平均精度均值提高了2.8%。

在目标分类任务中,将基准分类模型(YOLOv51-c1s)和经过ConvNeXt Block模块优化主干网络的模型(YOLOv51-c1s-CN)分别在激光雷达水下障碍物图像分类数据集上进行训练。训练完成后,两组模型的对比测试结果如表2所示。从分类消融实验的结果可知,采用ConvNeXt Block模块优化主干网络的模型(YOLOv51-c1s-CN),其分类准确率(Accuracy)达到了96.8%,与基准分类模型相比,其分类准确率提高了3.2%

表2 目标分类消融实验

Model	Parameters/10 <sup>6</sup>	FLOPs/10 <sup>9</sup>	Accuracy/%
Yolo51-c1s(base)	27.9	70.4	93.6
Yolo51-c1s-CN	34.3	91.5	96.8

## 3 结论

针对无人机载激光雷达在水下障碍物探测及识别中的精度低、速度慢等问题,本文以YOLOv5算法为基础,对分类和检测两个网络进行了改进和优化。在激光雷达水下障碍物图像数据集上的测试结果表明,改进后的分类网络在分类准确率上得到了显著提升,同时改进后的检测网络在精确率、召回率和平均精度均值等方面也均有明显改善,这表明改进后的分类与检测模型的性能均得到了有效增强。

本文系国家自然科学基金项目《基于多点阵列式激光多普勒振动测量的碳纤维复合材料无损检测研究》(52173219)和上海市协同创新项目支持(项目编号24xtcx00100)联合研究成果。

### [参考文献]

- [1]董希贵,黄东武.特定环境下水下地形和碍航物的复合探测[J].海洋测绘,2004,(02):32-36.
- [2]阮英杰,贺岩,吕德亮,等.激光雷达水下障碍物剖面图像处理[J].红外与激光工程,2024,53(07):128-139.
- [3]Sun L, Pan J, Tang J. Shufflemixer: An efficient convnet for image super-resolution[J].Advances in Neural Information Processing Systems,2022,35:17314-17326.
- [4]杜俊.基于YOLOv8和迁移学习的垃圾分类方法[J].智能计算机与应用,2024,14(09):63-69.
- [5]Zhang Y F, Ren W, Zhang Z, et al. Focal and efficient IOU loss for accurate bounding box regression[J].Neurocomputing, 2022,506:146-157.

### 作者简介:

董帮林(1998--),男,汉族,河南信阳人,东华大学信息科学与技术学院,硕士生,研究方向:激光雷达图像处理。