

中图法分类号: TP751 文献标识码: A 文章编号: 1006-8961(2024)08-2162-13

论文引用格式: Wang S K, Wang X D, Qu L, Yao F Q, Liu Y Y, Li C H, Wang Y Z and Zhong G Q. 2024. Semantic segmentation benchmark dataset for coastal ecosystem monitoring based on unmanned aerial vehicle (UAV). Journal of Image and Graphics, 29(08):2162-2174(王胜科, 王贤栋, 曲亮, 姚凤芹, 刘莹莹, 李聪慧, 王玉琪, 仲国强. 2024. 面向无人机海岸带生态系统监测的语义分割基准数据集. 中国图象图形学报, 29(08):2162-2174)[DOI:10.11834/jig.230670]

面向无人机海岸带生态系统监测的 语义分割基准数据集

王胜科^{1*}, 王贤栋¹, 曲亮², 姚凤芹¹, 刘莹莹¹, 李聪慧¹, 王玉琪¹, 仲国强¹

1. 中国海洋大学信息科学与工程学部, 青岛 266100; 2. 国家海洋局北海环境监测中心, 青岛 266033

摘要: 目的 海岸带生态系统的监测是维护生态平衡和可持续发展的重要任务。语义分割技术在遥感影像中的应用为海岸带生态系统的精准监测提供了有效手段。然而, 目前该领域仍面临一个挑战, 即缺乏公开的专门面向海岸带生态系统的语义分割基准数据集。方法 受到自然环境和人为因素的共同影响, 海岸带外部形态变化快, 目前的卫星遥感和常规的人工测量、船载作业等监测方式已难以满足实时观测的要求。因此本文首先利用无人机对海岸带典型生态群落进行拍摄、采集和标注, 构建了面向海岸带生态系统的基准数据集, 名为 OUC-UAV-SEG。接着, 通过统计的方法, 对 OUC-UAV-SEG 进行定量分析, 并对数据集中存在的挑战进行了深入探讨。最后, 使用典型视觉语义分割算法对 OUC-UAV-SEG 展开评估。结果 对 OUC-UAV-SEG 数据集进行了基准测试, 评估了当前可用的语义分割算法在该数据集上的性能, 并揭示了这些算法的局限性。结论 OUC-UAV-SEG 数据集的建立为海岸带生态系统的监测提供了新的资源, 有助于科学家、生态学家和决策者们更好地了解海岸带生态系统的现状和变化趋势, 为海洋生态保护和管理提供更精确的信息。该数据集链接: <https://github.com/OucCVLab/OUC-UAV-SEG>。

关键词: 海岸带生态系统; 遥感; 无人机(UAV); 基准数据集; 语义分割

Semantic segmentation benchmark dataset for coastal ecosystem monitoring based on unmanned aerial vehicle (UAV)

Wang Shengke^{1*}, Wang Xiandong¹, Qu Liang², Yao Fengqin¹, Liu Yingying¹,

Li Conghui¹, Wang Yuzhen¹, Zhong Guoqiang¹

1. Faculty of Information Science and Engineering, Ocean University of China, Qingdao 266100, China;

2. North China Sea Environmental Monitoring Center, State Oceanic Administration, Qingdao 266033, China

Abstract: Objective The coastal ecosystem is a natural system composed of biological communities and their interactions with the environment, including typical coastal ecosystems such as mangroves, salt marshes, coral reefs, seagrass beds, oyster reefs, sandy shores, as well as complex ecosystems like estuaries and bays. These ecosystems play a crucial role in maintaining high-quality ecological environments and fostering rich marine biodiversity. A healthy coastal ecosystem is not

收稿日期: 2023-09-28; 修回日期: 2024-01-10; 预印本日期: 2024-01-17

* 通信作者: 王胜科 neverme@ouc.edu.cn

基金项目: 国家重点研发计划资助(2018AAA0100400); 山东省自然科学基金项目(ZR2020MF131, ZR2021ZD19); 海洋科技协同创新中心项目(22-05-CXZX-04-03-17); 青岛市科技惠民项目(21-1-4-ny-19-ns; 24-1-8-cspz-22-nsh); 中国海洋大学联合培育项目(202265007)

Supported by: National Key R&D Program of China (2018AAA0100400); Natural Science Foundation of Shandong Province, China (ZR2020MF131, ZR2021ZD19)

only a crucial support for the sustainable economic development of China but also an essential component of the ecological security of the country. The application of semantic segmentation techniques in remote sensing imagery has provided an effective means for the precise monitoring of coastal ecosystems, which offers scientists, ecologists, and decision makers clear and highly comprehensive information to understand the current state and changing trends of coastal ecosystems. However, a significant challenge persists, that is, a specialized, comprehensive, and fine-grained data support system for coastal ecosystems is lacking, which causes difficulty in accurately understanding the distribution, area, and changes in ecosystems such as salt marshes, seagrass beds, and reed beds. This challenge has become a pressing issue in the current national marine ecological conservation efforts. Currently, the monitoring of coastal ecosystems relies primarily on satellite remote sensing and traditional surveying methods. Satellite remote sensing, with its unique advantages of all-weather, all-day, large-scale, and long-time observation, is widely used for monitoring marine ecology and resources through the analysis of satellite data. However, the spatial resolution of satellite remote sensing images has limitations, which introduces errors in cases that require fine-scale monitoring. For example, in narrow rivers, small wetland areas, or islands, the limited spatial resolution may result in unclear visibility of small features in the images. Traditional surveying methods often require professionals to conduct field surveys, which poses safety hazards in complex and risky environments. In addition, these methods are susceptible to human and natural factors, which leads to challenges in precise positioning and depiction, coupled with extended and inefficient monitoring cycles. Field surveys and mapping work for coastal ecosystems face significant challenges using traditional methods. **Method** This study utilized unmanned aerial vehicles (UAVs) to capture, collect, and annotate data from typical coastal ecosystems in real time for addressing urgent issues in marine ecological conservation. This effort led to the establishment of the OUC-UAV-SEG dataset, which includes various typical vegetation types found in coastal ecosystems, such as reed, seagrass beds, spartan, and crucially, addresses a key marine event—oil spills. **Result** In contrast to previous research that predominantly focused on individual dataset scenarios, such as spartan or oil spills, using traditional remote sensing methods for data analysis, this study stands out by using statistical methods to conduct a detailed quantitative analysis of OUC-UAV-SEG, which covers various categories and their respective quantities. In discussing the challenges posed by the dataset, including the discrete clustered morphology of eelgrass, interlaced spiderweb-like oil spills, fragile tubular Sargassum, and mottled tufted seagrass beds, existing segmentation algorithms exhibited suboptimal performance in handling these distinctive features. Finally, classical visual semantic segmentation algorithms were applied to evaluate OUC-UAV-SEG by conducting benchmark tests to assess the performance of currently available semantic segmentation algorithms on this dataset and revealing their limitations. **Conclusion** The establishment of the OUC-UAV-SEG dataset marks a significant step forward in providing a novel resource for monitoring coastal ecosystems. This dataset empowers scientists, ecologists, and decision makers with a more comprehensive understanding of the current conditions and evolving trends in coastal ecosystems. These insights are crucial for furnishing more accurate information to support marine ecological conservation and management efforts, which holds positive significance. Future efforts will focus on expanding the diversity of the dataset by collecting data on marine ecological environments under more extensive geographic and meteorological conditions to comprehensively reflect ecosystem diversity. Furthermore, by regularly organizing workshops and challenges, the aim is to advance cutting-edge research in the field of coastal ecosystem monitoring. Finally, active use of social media and other channels will be employed to engage with the public for raising awareness of marine environmental issues and inspiring more people to participate in the monitoring and protection of coastal ecosystems. For comprehensive details about this dataset, please refer to the project link: <https://github.com/OucCVLab/OUC-UAV-SEG>. **Key words:** coastal ecosystem; remote sensing; unmanned aerial vehicle (UAV); benchmark dataset; semantic segmentation

0 引言

海岸带生态系统是指由生物群落及其与环境相

互作用构成的自然系统,包括红树林、盐沼、珊瑚礁、海草床、牡蛎礁、砂质海岸等典型海岸带生态系统,以及河口、海湾等复合型生态系统。这些生态系统在维护优质生态环境和丰富海洋生物多样性方面发

挥着重要作用(刘志媛和刘青松, 2023; Liu 等, 2022)。

《“十四五”海洋生态环境保护规划》提到,“蓝色海湾”整治行动、海岸带保护修复工程等深入实施,海洋生态环境总体改善,局部海域生态系统服务功能明显提升。但我国海洋环境污染和生态退化等问题仍然突出,治理体系和治理能力建设亟待加强。重点工作中强调了恢复修复典型海洋生态系统的重要性,同时加强海洋生态监测监管,以提升海洋生态系统的质量和稳定性。2021年10月20日,习近平总书记在考察黄河入海口时(董亚宁等, 2022)指出,黄河流域生态保护和高质量发展是我国重大的国家战略。健康的海岸带生态系统是我国经济可持续发展的重要支撑,也是国家生态安全的重要组成部分。目前,仍然面临一个问题,即缺乏分类分区、全面和精细化的海岸带生态系统数据支撑系统,无法准确了解盐沼、海草床、芦苇等生态系统的分布范围、面积以及海岸带区域的变化情况。这一问题已成为当前国家海洋生态保护中亟需解决的痛点和难点问题。

目前,对海岸带生态系统的监测主要采用卫星遥感和传统测绘方法。卫星遥感(蒋兴伟等, 2019; 刘启超等, 2021)具有全天时、全天候、大范围和长时序观测的独特优势,主要通过对卫星数据的分析处理来获取信息(徐科杰等, 2021; 汪顺清等, 2023; 刘志媛和刘青松, 2023),广泛应用于海洋生态(郑嘉豪等, 2023)与资源监测(张磊等, 2023)等领域。闵钰魁等人(2023)提出融合 Sentinel-2 和 GF-1 时序影像进行互花米草的清除。朱小波等人(2023)提出基于耀光反射差异的海面溢油遥感识别提取。然而,卫星遥感图像的空间分辨率存在一定的限制,对于一些需要精细监测的情况存在误差,因此无法满足某些海岸带生态系统面积统计的精细化需求。例如,在面对狭窄的河流、小型湿地区域或者岛屿,由于卫星空间分辨率有限,细小的特征可能无法在图像中清晰可见。传统测绘方法通常需要专业人员进行实地勘测,在复杂危险的环境下存在安全隐患。此外,传统测绘容易受到人力因素和自然因素的影响,导致准确定位和描绘困难,且周期长、时效性差。对于海岸带生态系统的野外调查和制图等工作,传统方法面临较大困难。

因此,为了应对海洋生态保护领域中亟需应对

的问题,本文借助无人机,建立了面向海岸带生态系统监测的数据集,即 OUC-UAV-SEG。该数据集包含了海岸带生态系统中一系列典型的植被,如互花米草、海草床和浒苔等,同时还涵盖一个关键的突发海洋事件——溢油。相较之前的研究工作(闵钰魁等, 2023; 朱小波等, 2023),这些研究聚焦于单一的数据集场景,主要涉及互花米草或者溢油的研究。这些研究采用了传统的遥感手段进行数据分析。然而,值得注意的是,它们并未提供关于数据集的详细信息,特别是有关数据集规模和地面真实标注的信息。

相比传统遥感方法通常受限于较低的空间分辨率,难以捕捉生态系统内部细微结构,本文采用了基于深度学习的语义分割技术,通过像素级别的分类,能够更深入地理解海岸带生态系统的结构和功能。这将为科学家、生态保护者和政策制定者提供重要的决策支持和管理工具,更有效地保护和管理海岸带生态系统,实现可持续发展的目标。

1 数据集属性

为了建立面向无人机的海岸带生态系统环境监测的高质量基准数据集,考虑了以下数据集属性:

1) 图像数据。数据集应包含大量的图像数据,这些数据由无人机在海岸带区域进行高空拍摄而得。图像应具有高分辨率和良好的视角,以捕捉海岸带生态系统的细节和特征。

2) 地理信息。每幅图像都应具备准确的地理信息,包括经纬度坐标或与特定参考框架相关的位置信息,这将有助于对图像进行地理定位和空间分析。

3) 标注信息。数据集需要准确的标注信息,以对海岸带生态系统图像中的不同要素进行分类和分割。例如,沙滩、岩石、水域、红树林等生态要素需要被标注并与相应的类别关联。

4) 多样性。数据集应覆盖不同类型的海岸带生态系统,包括红树林、盐沼、珊瑚礁、海草床、牡蛎礁、砂质海岸等。这样可以确保数据集的多样性和代表性,以满足各种研究和应用需求。

5) 数据更新。由于海岸带生态系统可能会发生变化,数据集应具备一定的更新机制,以确保数据的时效性和反映实际情况。定期采集新的图像数据和

更新标注信息将使数据集保持最新状态。

综上所述,一个高质量的基准数据集应包含高分辨率图像数据、准确的地理和标注信息,涵盖多样的海岸带生态系统类型,并具备良好的数据质量和足够的规模,以支持面向无人机的海岸带生态环境监测的研究和应用。同时,数据集也应具备更新机制,以反映实际环境的变化。

2 数据集描述

2.1 区域的选择

确定采集区域是在数据集采集过程中最重要的一步,尤其对于海岸带生态这一领域更是至关重要。海岸带生态是一个复杂多样的生态系统,涉及海洋、陆地和空气等多个要素的相互作用。在海岸带生态系统中,植被季节性的依赖性极为显著,并且不同区域的植被出现时间也存在着明显差异。以浒苔为例,它在江浙一带从5月份开始出现,而在青岛沿海

则是在6—7月份。海草床虽然也有季节性,但由于其生长处于水下,受潮水变化的影响较大,低潮时露出水面,高潮时难以观测。而6—8月则是互花米草生长旺盛的时期。面对如此复杂多样的生态现象,选择合适的采集区域显得尤为重要。鉴于这些季节性变化特征,本文主要集中在渤海地区进行数据采集。这一区域拥有丰富的海岸带生态资源,包括浒苔、海草床和互花米草等重要植被类型。通过在该区域的采集工作,可以更全面地了解渤海海岸带生态系统的动态特征和相互关系,为生态环境的保护和管理提供科学依据。图 1 展示了 OUC-UAV-SEG 数据采集的主要区域,包括曹妃甸、辽河口和黄河口。

2.2 数据集采集

在数据集的采集过程中,本研究选用了来自大疆公司的8种不同型号的无人机,相关参数如表 1 所示。这些无人机型号分别为:DJI M100、DJI M210、DJI M300、DJI Mavic 30T、DJI Mavic 3T、DJI MINI 3、

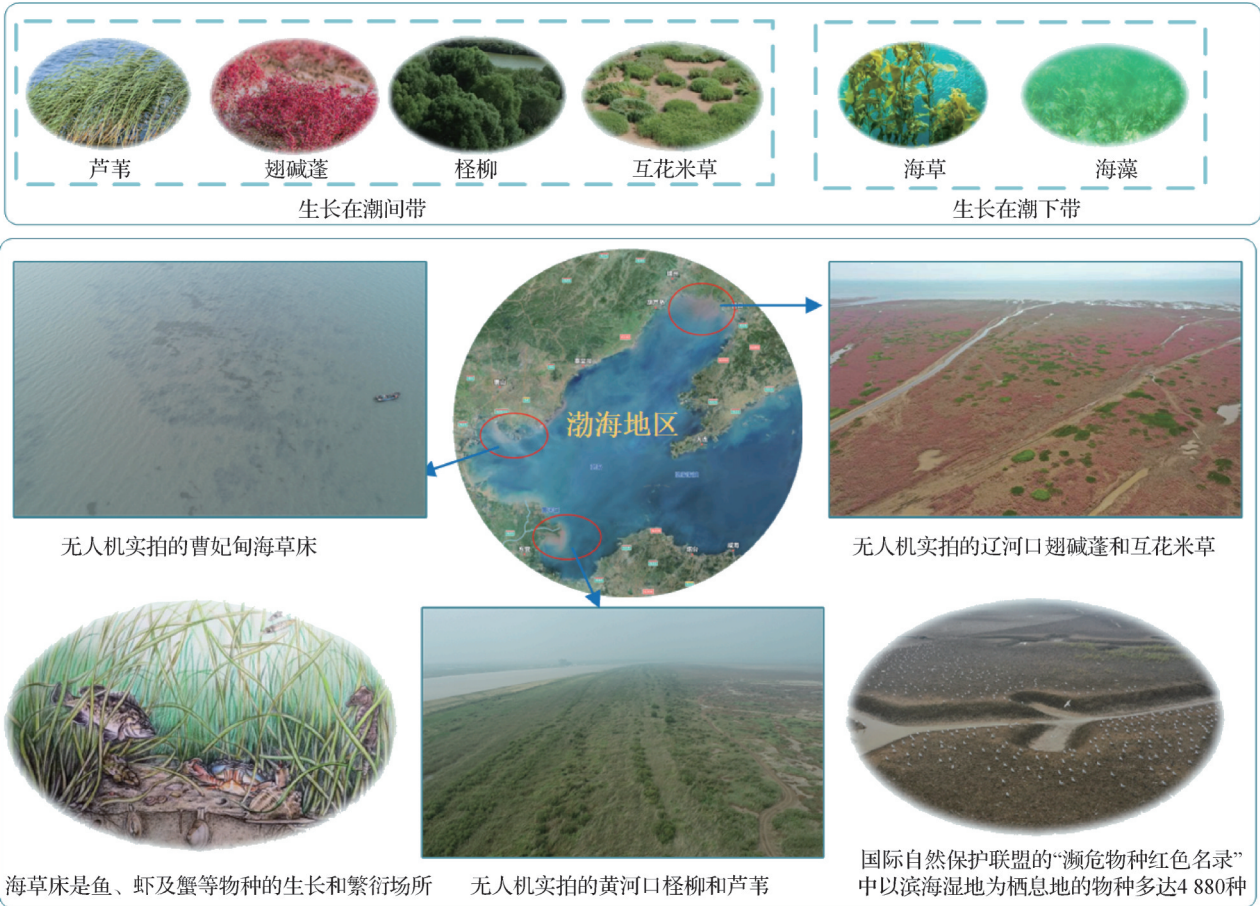


图 1 渤海海岸带生态系统
Fig. 1 The Bohai Sea coastal zone ecosystem

DJI 精灵 2 以及御 Mavic mini。每种无人机都具有独特的特性和功能,例如,DJI MINI3 的最高分辨率可达到 $8\,064 \times 6\,048$ 像素;而 DJI M300 则配备了红外感知系统,可以捕获物体的红外信息。通过使用多种型号的无人机,能够适应不同的场景需求,并获得更全面、详细的海岸带生态系统图像数据。本文研究团队经过专业培训,熟练掌握无人机的操作和控制,以确保采集到的图像质量和数据的准确性,表 2 展示了数据集采集时的相关信息。通过利用这些先进的无人机技术,能够覆盖广泛的海岸带区域,捕捉到细节丰富的生态系统特征,为后续的分类、分割和分析工作提供了可靠的基础。

表 1 无人机传感器相关信息
Table 1 Unmanned aerial vehicle sensor-related information

无人机型号	有效像素/万	最大照片尺寸/像素	最大视频分辨率/像素
DJI M300	2 000	$5\,184 \times 3\,888$	$3\,840 \times 2\,160$
DJI M210	2 000	$5\,184 \times 3\,888$	$3\,840 \times 2\,160$
DJI M100	1 200	$4\,000 \times 3\,000$	$1\,920 \times 1\,080$
DJI Mavic 30T	4 800	$8\,000 \times 6\,000$	$3\,840 \times 2\,160$
DJI Mavic 3T	4 800	$8\,000 \times 6\,000$	$3\,840 \times 2\,160$
DJI MINI 3	4 800	$8\,064 \times 6\,048$	$3\,840 \times 2\,160$
DJI 精灵 2	1 400	$4\,384 \times 3\,288$	$1\,920 \times 1\,080$

表 2 数据采集相关信息
Table 2 Data collection-related information

场景	时间	光照	飞行角度/(°)	飞行高度/m	潮位/cm
溢油	白天	良好	30 ~ 90	20 ~ 300	*
海草床	白天	良好	30 ~ 90	20 ~ 300	100 ~ 200
浒苔	白天	良好	30 ~ 90	20 ~ 300	*
互花米草	白天	良好	30 ~ 90	20 ~ 300	*

注:“*”表示采集过程不受潮位影响。

2.3 数据集标注

数据集标注是创建高质量基准数据集的关键步骤,对于模型后续的学习和模型性能具有直接影响。在 OUC-UAV-SEG 中,本文特别关注了海岸带生态系统中典型的植被,同时也涵盖了一个海洋突发事件——溢油。通过采用 Labelme 标注工具(Wada, 2016),进行了仔细标注,如图 2 所示。

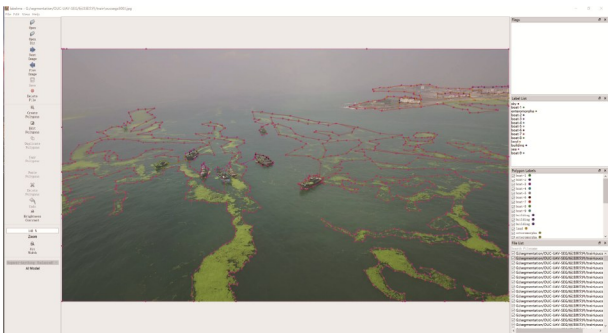


图 2 使用 Labelme 工具进行数据集的标注

Fig. 2 Annotation of dataset using Labelme tool

OUC-UAV-SEG 中包含了 15 个类别,分别是海草(seaweed)、浒苔(enteromorpha)、溢油(oil)、海洋(sea)、河流(river)、土地(land)、互花米草(spartin)、碱蓬(suaeda)、怪柳(tamarix)、芦苇(reed)、植被(vegetation)、建筑物(building)、天空(sky)、道路(road)、船(boat)。在标注过程中,对每个类别进行了精确的标注,保证了标签的准确性和一致性。对于典型的海岸带植被如互花米草、海草床、浒苔等,本文特别关注其生长位置和形态特征,例如互花米草只生长在河道,因此在标注过程中本文借助地理先验知识将其与其他标签区分开。这种精确标注为后续模型提供了准确的学习样本。

质量控制是确保数据集质量的另一个重要环节。本文进行了复核和验证,及时发现并修正标注中的错误和不一致性。只有高质量的标注数据才能为后续模型学习提供准确的训练基础,从而提高监测效果和准确率。图 3 是 OUC-UAV-SEG 数据集的标注信息展示。从上到下场景依次为怪柳碱蓬混生植被、互花米草、浒苔、溢油、海草床,图 3(b)标注图中不同颜色代表不同的标签。

2.4 数据集分析

OUC-UAV-SEG 数据集共包含 726 幅已经标注好的数据和 1 750 幅待标注的数据,每幅图像分辨率为 $1\,920 \times 1\,080$ 像素。这个数据集的规模和丰富性为海岸带生态系统监测和研究提供了重要的支持。数据集划分为 3 个子集:训练集、验证集和测试集。训练集包含 477 幅图像,验证集包含 140 幅图像,测试集包含 109 幅图像。这样的划分合理地考虑了训练、验证和测试样本之间的平衡和代表性,为模型的学习和评估提供了可靠的基础。

需要特别注意的是,考虑到海草床容易受到潮

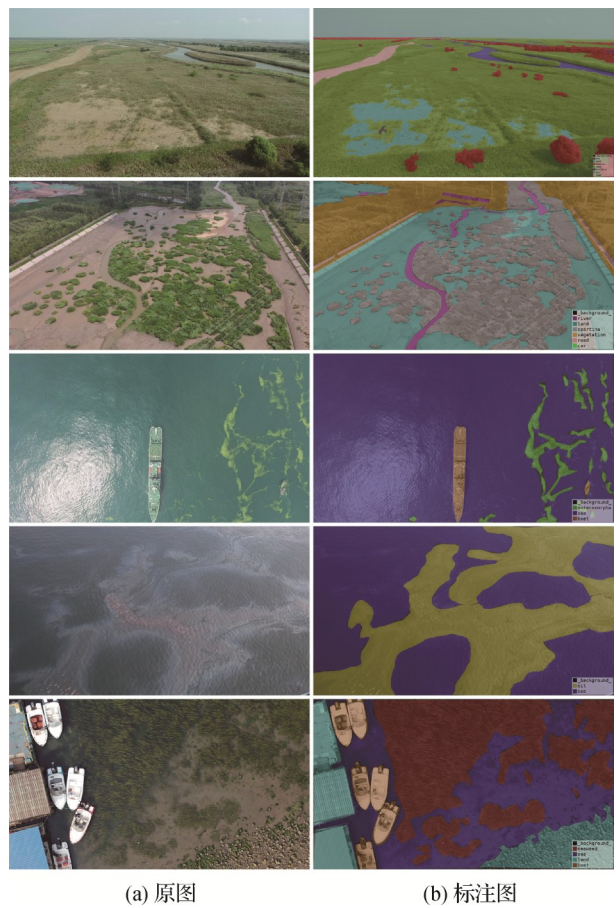


图 3 OUC-UAV-SEG 数据集展示
Fig. 3 Presentation of the OUC-UAV-SEG dataset
((a) original images; (b) annotation information)

汐的影响,并且观测期相对较短,关于海草床研究的数据采集工作还在持续进行中。截至 2023 年 9 月 1 日,本文已在青岛市西海岸新区设置了 25 个采集

点。在每个采集点,利用无人机在不同的潮位条件下进行了 70 次拍摄,共获得 1 750 张照片,旨在全面捕捉海草床在不同潮汐条件下的特征变化。然而,由于海草床所在的生态环境较为复杂,场景特殊性导致了一定的标注难题。例如,海草床的形态在不同潮汐和水深条件下可能发生显著变化,同时受到潮汐的周期性影响。这使得在不同时期和位置标注海草床的边界和分布变得更具挑战性,需要更加仔细和精确的处理,如图 4 所示。鉴于这种特殊性,将这部分数据集在后续阶段与 OUC-UAV-SEG 数据集进行整合,以确保数据集的质量和一致性。

图 5 展示了 OUC-UAV-SEG 数据集中 15 个类别的像素占比以及图像数量,按照像素数量由高到低排列。图 5 清晰地呈现了不同类别之间像素数量的不平衡分布情况。主要集中在海洋、陆地、芦苇等类别上,而像浒苔、道路和船只等类别的像素数相对较少,它们甚至不足总像素数的 1%。然而,这种较低的像素数并不一定是由于实例数量较少,而很可能是因为每个实例的大小不同。举例来说,一个建筑物可能需要超过 10 000 个像素才能完整表示,而图像中的一个船类实例可能只需要少于 100 个像素。这种大小差异导致了不同类别在像素数上的悬殊差距,从而使得数据集呈现出明显的像素数分布不平衡的现象。

在 OUC-UAV-SEG 数据集中,如图 6 所示,存在 4 个独特而复杂的挑战问题,这些问题与一般的图

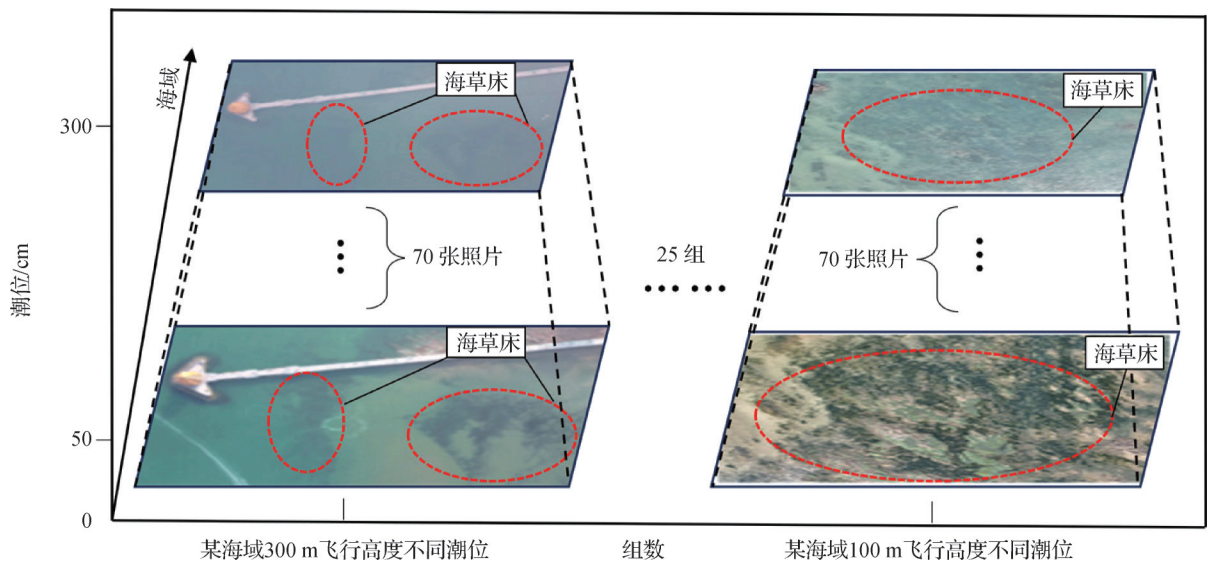


图 4 不同飞行高度不同潮位下的海草床
Fig. 4 Sea grass beds at different tidal levels and flight altitudes

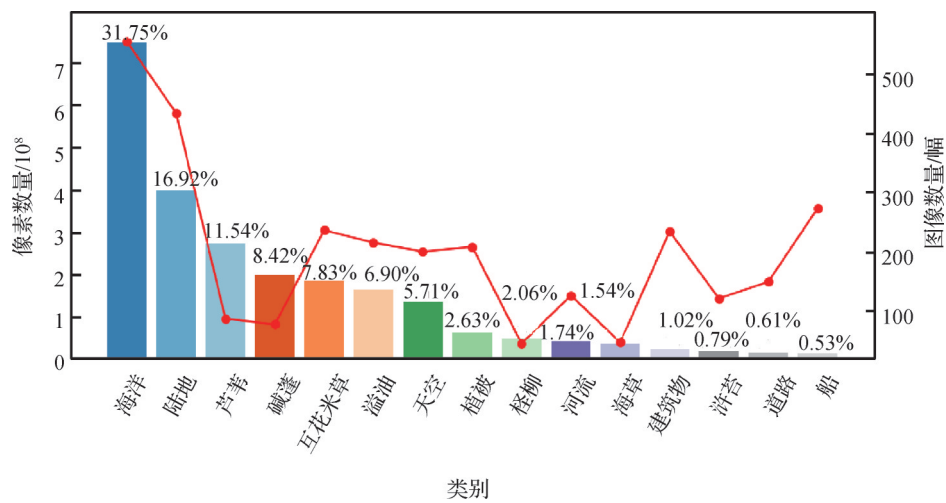


图 5 OUC-UAV-SEG 数据集中类别像素数量和图像数量占比

Fig. 5 The category pixel distribution and image count in the OUC-UAV-SEG dataset

像分割场景,如 ADE20k、CityScapes 数据集有着明显区别。

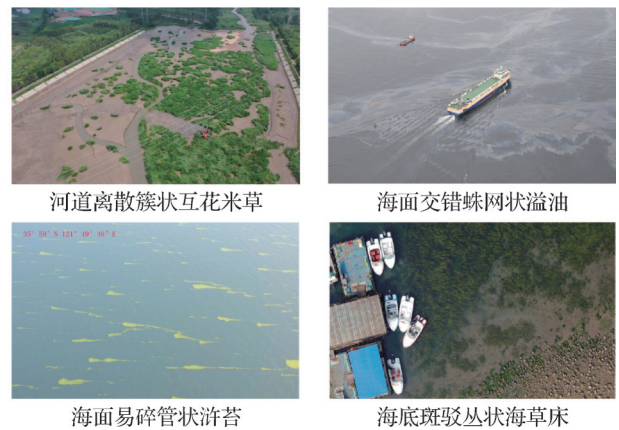


图 6 OUC-UAV-SEG 数据集存在的挑战

Fig. 6 Challenges in the OUC-UAV-SEG dataset

首先是河道入海口离散簇状形态的互花米草场景,互花米草分布不均匀且形态尺度变化大,这在一定程度上给常规的分割方法带来了挑战。这个问题需要研究如何准确地分割出散布在河道中的互花米草簇群,同时克服光照变化、水流干扰等因素对分割结果的影响,以实现高精度的分割效果。

其次是海面交错蛛网状溢油场景,由于溢油的特殊性,光线、反射和浪涌等因素增加了分割的难度。这个挑战需要寻求有效的深度学习方法,以捕捉溢油的复杂纹理和形态变化,从而实现对蛛网状溢油的精准分割。

另外,海面上易碎的管状浒苔生长细长且脆弱,受到海流和风浪的影响,使得特征表现复杂多变。

这一问题需要研究如何区分易碎管状浒苔与海面其他细小纹理,以准确地定位和分割出这些浒苔区域。

最后,海底斑驳丛状海草床的分割面临着海底环境的昏暗和复杂性,以及海草床的生长状态受到海底地形和潮汐的影响,观测期短。这个挑战需要考虑如何利用多模态信息,包括水深数据和其他光谱信息,以提高分割的准确度和稳健性。

综上所述,OUC-UAV-SEG 数据集集中的 4 个独特而复杂的挑战问题代表了计算机视觉领域在特定领域应用中的前沿挑战。这些问题不仅与一般的图像分割场景有明显区别,还受到海洋环境的特殊性质和条件的影响,使得它们更加复杂和具有挑战性。

3 数据集验证与评估

3.1 评价指标

与 U-Net(Ronneberger 等,2015)一致,本文选择平均交并比(mean intersection over union, mIoU)和平均精确度(mean accuracy, mAcc)作为模型的评价指标。具体而言,平均交并比通过计算每个类别预测的区域与真实标签的交并比,并对所有类别的交并比取平均值,来衡量模型的分割准确度。平均精确度则通过计算每个类别被正确分类的像素数与该类别总像素数之比的平均值,来评估模型的分类准确性。

3.2 模型评估

本文选取了 10 个代表性的分割模型,它们分别在 2015 年至 2023 年之间提出,在图像分割领域取得了显著的突破。为了对上述模型进行训练,本文

选择了由 OpenMMLab 维护的语义分割仓库, OpenMMLab/MMSegmentation(<https://github.com/open-mmlab/mms Segmentation>)仓库是一个用于模型训练的开源项目,旨在为语义分割任务提供支持。它提供了丰富的工具和资源,帮助研究人员和开发者训练、评估和部署各种语义分割模型,是计算机视觉领域中重要的开源资源之一。

U-Net(Ronneberger 等, 2015)使用编码器—解码器结构和跳跃连接在医学图像等领域表现出色。FCN (fully convolutional networks) (Shelhamer 等, 2017)采用完全卷积的架构使其能够进行任意大小图像的语义分割,为图像分割带来了新的思路。PSPNet (pyramid scene parsing network) (Zhao 等, 2017)通过金字塔池化策略捕获多尺度上下文信息,提高了分割模型对于物体不同尺寸的理解能力。DeepLabV3+(Chen 等, 2018)结合空洞卷积和解码器模块,有效地处理物体边界和细节,适用于高分辨率图像。DANet (dual attention network) (Fu 等, 2019)引入了自适应注意力机制,使网络能够动态地分配注意力以改善特征的代表能力。CGNet (light-weight context guided network) (Wu 等, 2020)专注于轻量级分割,通过组合全局和局部信息,在资源受限的情况下提供高效的性能。BiSeNetV2 (bilateral network with guided aggregation for real-time semantic segmentation) (Yu 等, 2021)强调了快速实时分割,通过利用不同分辨率的特征图来平衡速度和精度。SegFormer (Xie 等, 2021)结合 Transformer 架构和分割任务,将自注意机制引入图像分割,取得了优异的结果。Mask2Former (masked-attention mask Transformer) (Cheng 等, 2022)延续了 Transformer 架构,通过引入分割特定的结构改进,进一步提升分割性能。PIDNet (real-time semantic segmentation network inspired from pid controller) (Xu 等, 2023)采用了潜在层次的特征交互设计,以提高图像分割的精确性和鲁棒性。为了保证实验的公平性,上述 10 个模型实验均未使用预训练模型进行迁移训练。

3.3 定量分析

对 OUC-UAV-SEG 数据集进行了全面测试,并使用当前主流的分割模型来评估。测试结果如表 1 所示,在测试的 10 个模型中,PIDNet (Xu 等, 2023)在该数据集上取得了最佳的平均交并比(mIoU)和平均准确度(mAcc),这在一定程度上得益于它新颖的

三支网络架构,通过解析详细信息、上下文信息和边界信息,并应用边界注意机制来引导融合,以解决双分支网络中细节特征被上下文信息淹没的问题。

值得关注的是,虽然 PSPNet 的发表时间可以追溯到 2017 年,但其在 OUC-UAV-SEG 数据集上的表现依然引人注目,甚至超越了 2022 年提出的 Mask2Former,取得了较优的结果。这些发现强调了 OUC-UAV-SEG 场景与普通语义分割场景之间的差异。OUC-UAV-SEG 是一种特殊的语义分割任务,因此其数据具有独特的特征和挑战。

本文进一步分析了模型在 OUC-UAV-SEG 数据集中不同类别上的性能,结果如表 3—表 5 所示。本文观察到不同模型在不同类别上取得的最优值是不一样的。以海草类别为例,FCN 在这个类别上取得了最佳的交并比(IoU)值,但在准确度(Acc)这个评价指标上,PSPNet 表现最优。这种现象进一步凸显了 OUC-UAV-SEG 数据集的复杂性和多样性。在这个特殊场景下,不同类别的图像具有各自独特的特征和挑战,导致模型在不同类别上表现出差异。对于海草类别而言,FCN 可能在处理细粒度结构和形状方面具有一定优势,从而取得了较高的 IoU 值。而 PSPNet 在处理上下文信息和多尺度特征时表现优异,因此在准确度上更出色。

图 7 清晰地展示了模型性能与每秒帧数(frames per second, FPS)之间的关系。BiSeNetV2 以其独特的轻量设计和双路网络结构脱颖而出,表现

表 3 不同模型在 OUC-UAV-SEG 数据集中的表现
Table 3 Performance of different models on the OUC-UAV-SEG dataset

模型	mIoU/%	mAcc/%
U-Net(Ronneberger等,2015)	57.31	68.93
FCN(Shelhamer等,2017)	57.66	69.04
PSPNet(Zhao等,2017)	58.45	70.28
DeepLabV3+(Chen等,2018)	57.98	70.07
DANet(Fu等,2019)	56.97	69.51
CGNet(Wu等,2020)	58.63	70.46
BiSeNetV2(Yu等,2021)	56.68	68.80
SegFormer(Xie等,2021)	58.38	70.28
Mask2Former(Cheng等,2022)	58.37	70.44
PIDNet(Xu等,2023)	59.00	71.15

注:加粗字体表示各列最优结果。

表4 不同模型不同类别IoU值
Table 4 IoU values of different models for different classes

	/%										
模型	海草	浒苔	溢油	海洋	河流	土地	互花米草	碱蓬	怪柳	芦苇	植被
U-Net	71.12	54.58	55.77	77.19	53.68	68.31	76.94	79.36	22.41	65.03	72.56
FCN	71.19	52.81	58.72	77.69	58.79	68.64	77.67	80	23.12	65.32	73.39
PSPNet	71.15	52.28	58.28	78.59	56.71	69.53	77.93	78.64	19.4	64.26	74.09
DeepLabV3+	70.91	52.69	59.91	78.35	56.93	68.61	78.14	78.42	19.17	64.17	73.83
DANet	70.32	55.52	57.22	76.69	52.75	67.92	77.57	76.6	14.87	62.6	74.44
CGNet	70.37	52.05	58.2	78.1	52.95	70.85	79.02	81.24	36.8	67.78	74.51
BiSeNetV2	70.32	52.21	59.34	78.69	56.05	65.44	77.8	75.05	13.27	62.45	72.89
SegFormer	70.64	54.22	58.41	78.83	55.32	69.94	76.82	80	21.6	64.59	73.65
Mask2Former	70.25	54.89	56.58	74.97	44.26	73.07	75.92	79.96	30.64	65.3	75.21
PIDNet	70.81	57.44	59.86	78.39	56.89	66.54	76.72	77.75	36.04	67.03	73.33

注:加粗字体表示各列最优结果。

表5 不同模型不同类别Acc值
Table 5 Acc values of different models for different classes

	/%										
模型	海草	浒苔	溢油	海洋	河流	土地	互花米草	碱蓬	怪柳	芦苇	植被
U-Net	95.5	59.79	75.27	89.52	65.39	80.53	91.88	91.95	23.66	89.82	80.62
FCN	95.75	58.54	76.4	91.09	72.4	80.6	91.37	90.57	24.34	91.49	80.82
PSPNet	96.17	57.72	77.89	89.34	73.96	81.53	91.88	91.47	20.27	90.74	81.89
DeepLabV3+	95.81	58.82	78.42	89.65	76.17	80.13	90.97	90.03	20.49	90.86	82.18
DANet	96.15	61.62	77.25	87.67	71.87	79.69	91.55	90.33	16.3	88.89	81.42
CGNet	95.63	56.57	77.66	88.52	74.91	84.97	91.6	95.27	38.99	87.86	83.42
BiSeNetV2	94.85	56.24	77.37	89.55	74.39	77.43	92.59	93.04	15.6	88.44	80.17
SegFormer	95.8	59.55	75.8	89.82	74.12	82.25	89.38	92.01	22.06	90.49	81.57
Mask2Former	95.19	59.77	75.21	87.34	64.59	82.25	88.05	90.99	31.66	90.29	82.4
PIDNet	96.02	64.37	78.92	88.26	76.98	78.76	92.15	92.54	41.62	88.86	82.39

注:加粗字体表示各列最优结果。

出卓越的速度性能,尽管在精度方面做出了一定的妥协。相比之下,PIDNet在精度方面表现出色,但相对于BiSeNetV2而言速度略慢。未来,可以汲取这两个模型的优点,设计更快、更强大的神经网络,并将其成功地部署到资源受限的无人机设备上,以实现海岸带生态系统的实时分析。

为了更深入地分析当前模型在OUC-UAV-SEG数据集中的问题,本文采用混淆矩阵可视化方法对目前表现最佳的模型PIDNet的分类能力进行了分

析。混淆矩阵的可视化结果如图8所示,数字越大,颜色越深,代表预测的类别概率越大。可以看到,模型在识别海草、互花米草和赤碱蓬方面表现出色,分别达到了96.02%、92.15%和95.13%的高准确率。

然而,在怪柳类别的识别中,却出现了高达54.72%的误分类率。这是因为怪柳和芦苇在外形上非常相似,网络难以准确区分它们。在道路类别上,同样也出现了类似的情况。这表明网络在区分形态相近的目标时,鲁棒性受到了影响,未来研究可

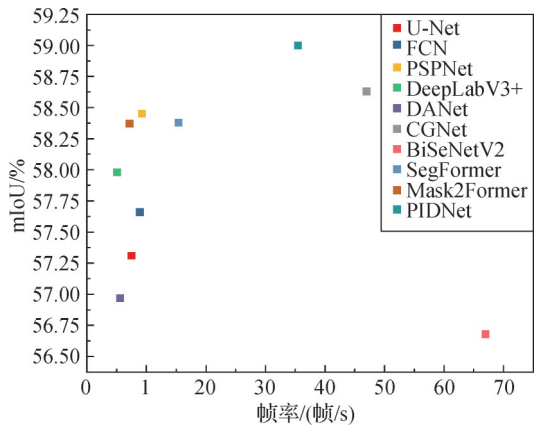


图 7 模型性能与 FPS 的关系
Fig. 7 The relationship between model performance and FPS

以重点关注如何改进这一方面。

图 9 展示了 OUC-UAV-SEG 数据集中每个类别

在 PIDNet 网络中的准确率和召回率,以探讨长尾现象对模型性能的影响。尽管船和浒苔的像素数量较少,但由于它们具有良好的外观特征,并没有影响模型的准确率。相反,对于道路类别来说,根据图 8 的混淆矩阵可知,由于其与土地类别的特征相近,模型容易出现误分类。这说明像素数量的多少对模型的分类能力影响不是很大,更关键的是类别是否具有可区分的特征。

3.4 定性分析

本文使用 10 个模型中效果最好的 PIDNet(Xu 等, 2023)对 OUC-UAV-SEG 数据集中具有挑战性的部分进行了可视化,结果如图 10 所示。图 10(a)是原图,图 10(b)是标注图,图 10(c)是预测图。不同颜色代表不同的类别标签。

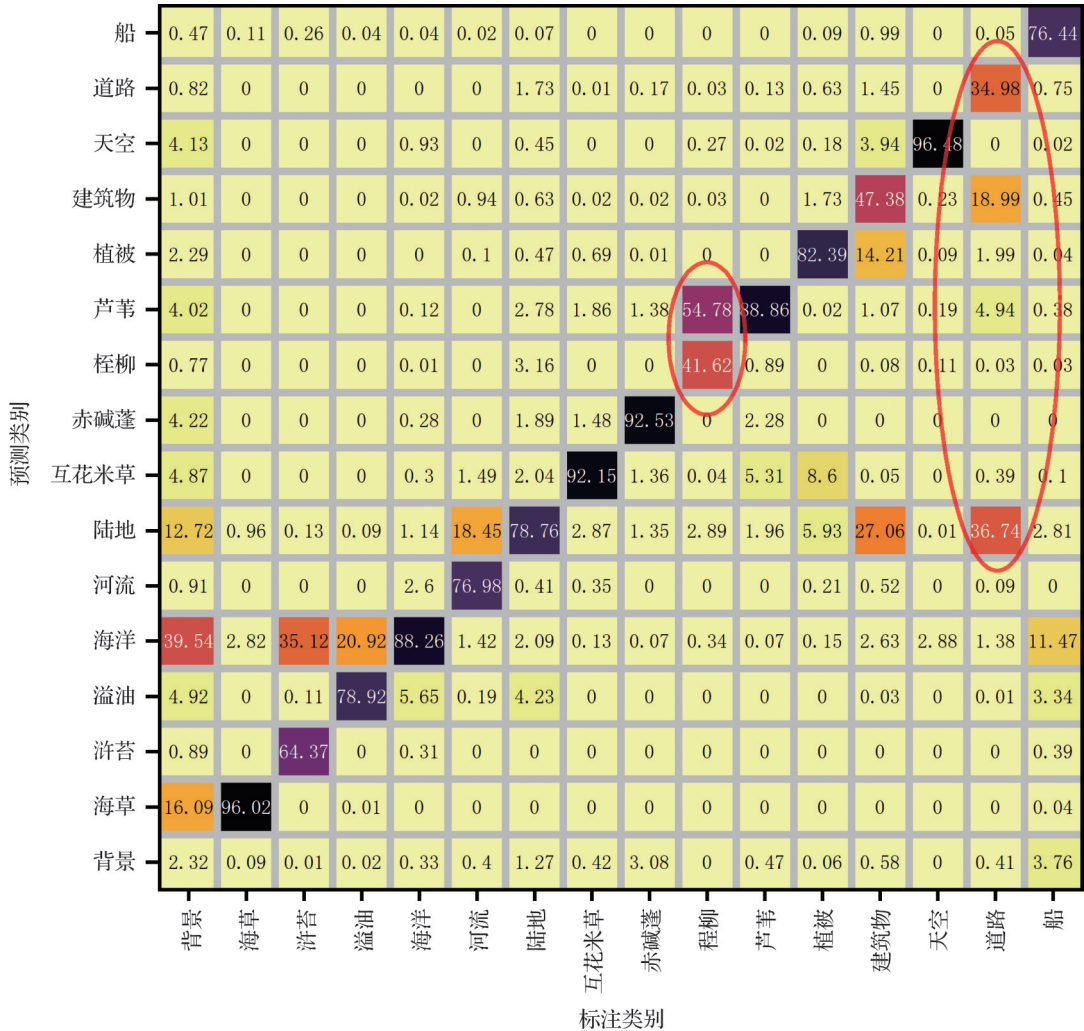


图 8 混淆矩阵可视化
Fig. 8 Confusion matrix visualization

对于互花米草的场景,由于其独特的团簇状形态,导致分布范围较近的离散的簇状目标被模型错

误地识别为一个整体。这是因为目标之间的连接性和相似性在图像中产生了模糊性,模型产生幻觉,从

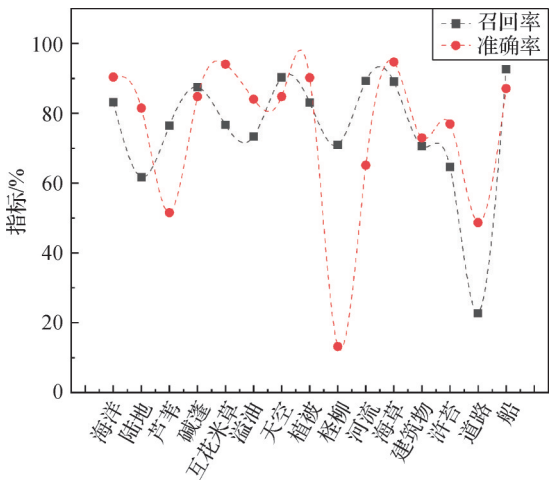


图9 各类别在PIDNet模型上的召回率和准确率
Fig. 9 The recall and accuracy rates for each category on the PIDNet model

而导致误分类。对于海面溢油的场景,由于其与海洋相比,溢油区域呈现低对比度特征,纹理性相对较弱,容易与海洋表面融为一体。这种轻微的差异在图像中不够明显,模型容易错误地将溢油区域识别为背景。对于浒苔的场景,其特有的细长易破碎的特征,导致目标形态复杂,对网络模型来说,其特征

的提取是不友好的。对于海草床的场景,由于潮水的淹没作用,海草床的关键特征可能变得模糊或难以分辨。
未来的研究可以着重于优化网络结构,引入更丰富的上下文信息和跨场景特征表示,以提高模型对多样化海洋场景下目标的准确识别能力。

4 结 论

本文借助无人机构建了面向海洋生态环境监测的数据集 OUC-UAV-SEG,旨在解决海岸带生态系统缺乏分类分区、全面和精细化数据支撑的问题,为海岸带生态系统的保护和管理提供可靠的基础。除此之外,本文还对当前现有模型在 OUC-UAV-SEG 数据集上进行了全面测试,并建立了相应的基准,以评估其性能。在测试过程中,本文发现了现有模型的一些不足,这为后续的改进工作提供了有价值的参考。未来,将继续扩充数据集的多样性,收集更广泛的地理和气象条件下的海洋生态环境数据,以更全面地反映生态系统的多样性。同时,通过定期举办

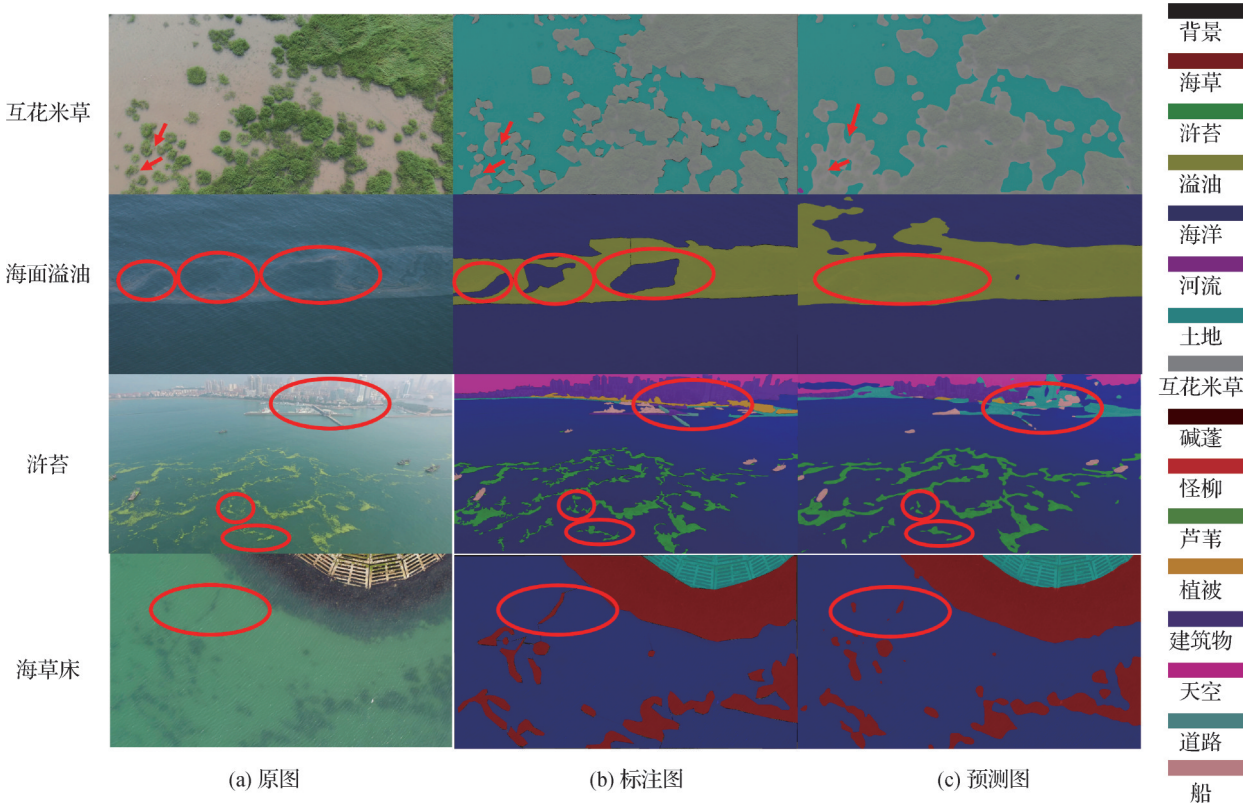


图10 PIDNet在OUC-UAV-SEG数据集的推理结果
Fig. 10 The inference results of PIDNet on the OUC-UAV-SEG dataset
((a) original images; (b) annotated images; (c) predicted images)

研讨会和挑战赛,致力于推动海岸带生态系统监测领域的前沿研究。最后,将积极利用社交媒体等渠道与社会大众互动,提高公众对海洋环境问题的关注,激发更多人参与到海岸带生态系统的监测和保护中。

参考文献(References)

- Chen L C, Zhu Y K, Papandreou G, Schroff F and Adam H. 2018. Encoder-decoder with atrous separable convolution for semantic image segmentation//Proceedings of the 15th European Conference on Computer Vision. Munich, Germany: Springer: 801-818 [DOI: 10.1007/978-3-030-01234-2_49]
- Cheng B M, Misra I, Schwing A G, Kirillov A and Girdhar R. 2022. Masked-attention mask Transformer for universal image segmentation//Proceedings of 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition. New Orleans, USA: IEEE: 1290-1299 [DOI: 10.1109/CVPR52688.2022.00135]
- Dong Y N, Fan B K, Li S P and Shan J J. 2022. Research on ecological protection and high-quality development of the yellow river basins from the perspective of ecological civilization. *Ecological Economy*, 38(2): 217-222 (董亚宁, 范博凯, 李少鹏, 单菁菁. 2022. 生态文明视角下黄河流域生态保护和高质量发展研究. *生态经济*, 38(2): 217-222)
- Fu J, Liu J, Tian H J, Li Y, Bao Y J, Fang Z W and Lu H Q. 2019. Dual attention network for scene segmentation//Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach, USA: IEEE: 3146-3154 [DOI: 10.1109/CVPR.2019.00326]
- Jiang X W, He X Q, Lin M S, Gong F, Ye X M and Pan D L. 2019. Progresses on ocean satellite remote sensing application in China. *Haiyang Xuebao*, 41(10): 113-124 (蒋兴伟, 何贤强, 林明森, 龚芳, 叶小敏, 潘德炉. 2019. 中国海洋卫星遥感应用进展. *海洋学报*, 41(10): 113-124) [DOI: 10.3969/j.issn.0253-4193.2019.10.008]
- Liu Q C, Xiao L and Yang J X. 2021. Parallel implementation of content-guided deep convolutional network for hyperspectral image classification. *Journal of Image and Graphics*, 26(8): 1926-1939 (刘启超, 肖亮, 杨劲翔. 2021. 面向高光谱图像分类的内容引导卷积深度网络并行实现. *中国图象图形学报*, 26(8): 1926-1939) [DOI: 10.11834/jig.200411]
- Liu T S, Lam K M, Zhao R and Qiu G P. 2022. Deep cross-modal representation learning and distillation for illumination-invariant pedestrian detection. *IEEE Transactions on Circuits and Systems for Video Technology*, 32(1): 315-329 [DOI: 10.1109/TCSVT.2021.3060162]
- Liu Z Y and Liu Q S. 2023. Study on the effect and mode of ecological civilization construction in coastal areas of China. *Environmental Protection*, 51(9/10): 55-60 (刘志媛, 刘青松. 2023. 我国滨海地区生态文明建设成效及模式研究. *环境保护*, 51(9/10): 55-60) [DOI: 10.14026/j.cnki.0253-9705.2023.z2.014]
- Min Y K, Ke Y H, Han Y, Yin X L and Zhou D M. 2023. Dynamic monitoring of invasive *Spartina alterniflora* clearance via fusion of Sentinel-2 and GF-1 time series images. *National Remote Sensing Bulletin*, 27(6): 1467-1479 (闵钰魁, 柯樱海, 韩月, 尹小岚, 周德民. 2023. 融合 Sentinel-2 和 GF-1 时序影像的入侵植物互花米草清除动态监测. *遥感学报*, 27(6): 1467-1479 [DOI: 10.11834/jrs.20232279]
- Ronneberger O, Fischer P and Brox T. 2015. U-Net: convolutional networks for biomedical image segmentation//Proceedings of the 18th International Conference on Medical Image Computing and Computer-Assisted Intervention. Munich, Germany: Springer: 234-241 [DOI: 10.1007/978-3-319-24574-4_28]
- Shelhamer E, Long J and Darrell T. 2017. Fully convolutional networks for semantic segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(4): 640-651 [DOI: 10.1109/TPAMI.2016.2572683]
- Wada K. 2016. Labelme: image polygonal annotation with python [CP/OL]. [2023-09-28]. <https://doi.org/10.5281/zenodo.5711226>
- Wang S Q, Yang J X, Shao Y T and Xiao L. 2023. Non-negative sparse component decomposition based modeling and robust unmixing for hyperspectral images. *Journal of Image and Graphics*, 28(2): 613-627 (汪顺清, 杨劲翔, 邵远天, 肖亮. 2023. 高光谱图像非负稀疏分量分解建模与鲁棒性解混方法. *中国图象图形学报*, 28(2): 613-627) [DOI: 10.11834/jig.211054]
- Wu T Y, Tang S, Zhang R, Cao J and Zhang Y D. 2020. CGNet: a light-weight context guided network for semantic segmentation. *IEEE Transactions on Image Processing*, 30: 1169-1179 [DOI: 10.1109/TIP.2020.3042065]
- Xie E Z, Wang W H, Yu Z D, Anandkumar A, Álvarez J M and Luo P. 2021. SegFormer: simple and efficient design for semantic segmentation with Transformers//Proceedings of the 35th International Conference on Neural Information Processing Systems. [s.l.]: [s.n.]
- Xu J C, Xiong Z X and Bhattacharyya S P. 2023. PIDNet: a real-time semantic segmentation network inspired by PID controllers//Proceedings of 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Vancouver, Canada: IEEE: 19529-19539 [DOI: 10.1109/CVPR52729.2023.01871]
- Xu K J, Deng P F and Huang H. 2021. HSRS-SC: a hyperspectral image dataset for remote sensing scene classification. *Journal of Image and Graphics*, 26(8): 1809-1822 (徐科杰, 邓培芳, 黄鸿. 2021. HSRS-SC: 面向遥感场景分类的高光谱图像数据集. *中国图象图形学报*, 26(8): 1809-1822) [DOI: 10.11834/jig.200835]
- Yu C Q, Gao C X, Wang J B, Yu G, Shen C H and Sang N. 2021. BiSeNet V2: bilateral network with guided aggregation for real-time semantic segmentation. *International Journal of Computer Vision*,

129(11): 3051-3068 [DOI: 10.1007/s11263-021-01515-2]

Zhang L, Chen W and Wang Y H. 2023. Key sub-region feature fusion network for fine-grained ship detection and recognition in remote sensing images. *Journal of Image and Graphics*, 28(9): 2940-2955 (张磊, 陈文, 王岳环. 2023. 用于遥感舰船细粒度检测与识别的关键子区域融合网络. *中国图象图形学报*, 28(9): 2940-2955) [DOI: 10.11834/jig.220671]

Zhao H S, Shi J P, Qi X J, Wang X G and Jia J Y. 2017. Pyramid scene parsing network//*Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition*. Honolulu, USA: IEEE: 2881-2890 [DOI: 10.1109/CVPR.2017.660]

Zheng J H, Sun C, Lin Y, Li L and Liu Y C. 2023. Classification of salt marsh vegetation based on pixel-level time series from Landsat images. *National Remote Sensing Bulletin*, 27(6): 1400-1413 (郑嘉豪, 孙超, 林昀, 李璐, 刘永超. 2023. 基于 Landsat 像元级时间序列的海岸带盐沼植被分类. *遥感学报*, 27(6): 1400-1413 [DOI: 10.11834/jrs.20232461]

Zhu X B, Shen Y F, Liu J Q, Ding J, Jiao J N, Ju W M and Lu Y C. 2023. Optical extraction of oil spills based on sunglint reflection difference in HY-1C CZI images. *National Remote Sensing Bulletin*, 27(1): 197-208 (朱小波, 沈亚峰, 刘建强, 丁静, 焦俊男,

居为民, 陆应诚. 2023. 基于耀光反射差异的海面溢油遥感识别提取. *遥感学报*, 27(1): 197-208 [DOI: 10.11834/jrs.20221688]

作者简介

王胜科, 男, 副教授, 主要研究方向为计算机视觉、数字图像处理以及模式识别。E-mail: neverme@ouc.edu.cn

王贤栋, 男, 硕士研究生, 主要研究方向为目标检测和图像分割。E-mail: wangxiandong@stu.ouc.edu.cn

曲亮, 男, 硕士研究生, 主要研究方向为海洋生态保护修复和海洋信息化。E-mail: qingdaoquliang@aliyun.com

姚凤芹, 女, 博士研究生, 主要研究方向为遥感图像处理和模型轻量化。E-mail: yaofengqin8312@stu.ouc.edu.cn

刘莹莹, 女, 硕士研究生, 主要研究方向为遥感图像分割。E-mail: 445284353@qq.com

李聪慧, 女, 硕士研究生, 主要研究方向为零样本图像分割。E-mail: liconghui@stu.ouc.edu.cn

王玉琪, 男, 硕士研究生, 主要研究方向为图像复原和图像分割。E-mail: wangyutian2099@stu.ouc.edu.cn

仲国强, 男, 教授, 主要研究方向为人工智能、机器学习和模式识别。E-mail: gqzhong@ouc.edu.cn