

于俊, 陈辉, 朱大明, 等. 基于软硬协作决策的半监督珊瑚礁底质分类方法[J]. 海洋学报, 2023, 45(4): 154–164. doi:10.12284/hyxb2023049
Yu Jun, Chen Hui, Zhu Daming, et al. A semi-supervised coral reef substrate classification method based on soft and hard collaborative decision making[J]. Haiyang Xuebao, 2023, 45(4): 154–164. doi:10.12284/hyxb2023049

基于软硬协作决策的半监督珊瑚礁底质分类方法

于俊¹, 陈辉², 朱大明^{1*}, 程亮², 段志鑫², 庄启智², 楚森森², 杨伟¹, 杜思雨¹

(1. 昆明理工大学 国土资源工程学院, 云南 昆明 650093; 2. 南京大学 地理与海洋科学学院, 江苏 南京 210023)

摘要: 珊瑚礁底质分类对海洋资源开发和海洋生态环境保护起到至关重要的作用。目前, 深度学习语义分割方法在遥感图像分类领域应用广泛, 但在底质分类方面的研究较少。由于基于全监督深度学习的方法中逐像素标注标签的成本较高, 不适用于大规模、高频次的底质分类工作, 基于半监督的深度学习方法能够有效利用已标注标签为无标签数据产生伪标签, 从而有效降低人工成本, 然而现有半监督方法的性能易受伪标签噪声的干扰。针对以上问题, 本文提出了一种基于软硬协作决策的半监督底质分类方法。首先, 利用多模型联合决策生成高质量的伪标签; 然后, 提出了一种能够顾及伪标签像素置信度的损失函数来指导模型进行训练; 最后, 采用软硬协作的决策方式得到精确的底质分类结果。在美属维尔京群岛圣克罗伊岛北部的巴克岛礁和夏威夷群岛的中途岛东南约 400 km 处的珍珠与爱马仕环礁的浅层底栖生物栖息地地图数据集上评估了本文方法的精度, 实验结果表明, 本文提出的方法与全监督学习方法精度相当, 比主流的语义分割方法精度平均高 3.08%, 能够有效服务于珊瑚礁底质调查工作。

关键词: 珊瑚礁底质分类; 软硬协作; 语义分割; 半监督学习; 全监督学习; 遥感; 珊瑚礁; 卷积神经网络

中图分类号: P737.22

文献标志码: A

文章编号: 0253-4193(2023)04-0154-11

1 引言

海洋拥有非常丰富的资源和地质信息, 具有巨大的发展潜力, 随着对海洋资源需求的增加和珊瑚礁科学研究的发展, 海洋研究越来越受到世界各国的重视。珊瑚礁底质分类是开发利用各种海洋资源和获取珊瑚礁信息的重要手段, 研究准确的、大规模的珊瑚礁底质分类方法具有重要意义^[1-2]。

珊瑚礁底质分类数据一般通过遥感影像和声学、水下摄影采集。声学、水下摄影数据能够提供高质量的珊瑚礁底质分布数据, 但是在进行大规模珊瑚礁底质分类时可能会耗费大量的时间和人力。由于遥感数据覆盖范围广、更新速度快, 基于遥感的方式能够进行大面积、高频次的底质分类^[3]。目前, 基于遥感影像的底质分类方法一般可分为两种: (1) 基于传

统机器学习的方法; (2) 基于深度学习的分类方法。基于传统机器学习的方法主要依赖于图像的纹理、颜色和几何形状等浅层特征来进行图像分割。如主成分分析^[4]、最大似然估计^[5]、K-means 聚类等无监督分类方法^[6]。随着高分辨率图像的发展, 基于对象的图像分析方法受到研究人员青睐^[7]。此类方法, 一般使用支持向量机^[8]、随机森林^[9]等机器学习模型进行分类。如万佳馨等^[10]、逢今朝等^[11]和董娟等^[12]基于 WordView-2(WV-2)和高分二号(GF-2)卫星影像数据对赵述岛和永乐群礁分别采用基于决策树、随机森林进行底质变化检测和分类体系构建。李晓敏等^[13]、Wan 和 Ma^[14]在 GF-2 和 WV-2 影像数据上利用支持向量机对中国西沙群岛分别进行底质分类体系的构建和相关底质的变化检测。Huang 等^[15]利用 Modis 数据对广西涠洲岛周围的活珊瑚覆盖范围进行分析。然

收稿日期: 2022-08-16; 修订日期: 2022-10-22。

基金项目: 国家自然科学基金(42001401)。

作者简介: 于俊(1994—), 男, 江西省九江市人, 研究方向为遥感、地理信息系统。E-mail: 2571445820@qq.com

* 通信作者: 朱大明(1970—), 博士, 副教授, 研究方向为地理信息系统、空间分析、3S 集成。E-mail: 634617255@qq.com

而,由于珊瑚礁底质在不同的海洋环境下,颜色、纹理等特征可能不同,经验性的特征设计只能用特定的数据解决特定的问题。

深度学习方法提供一种端到端的学习模式,将原始数据学习过程与整体工作结果相结合,最终学习到的模型是原始数据和预期结果的映射^[16]。基于深度学习的珊瑚礁底质调查工作虽然起步较晚,但也取得了一定的进展。例如,King等^[17]测试了两种深度神经网络使用水下珊瑚礁图像进行语义分割的效果,结果表明,深度学习架构确实能够超越水下珊瑚礁图像中传统的语义分割和对象分类方法。Li等^[18]将Planet Dove卫星图像与千年珊瑚礁测绘项目的珊瑚礁范围相结合,利用卷积神经网络与随机森林生成了一个全球珊瑚礁概率图。Wang与Hu^[19]基于多源遥感影像利用深度卷积神经网络监测和跟踪马尾藻水华。然而,深度学习相关网络严重依赖标注数据,逐像素标注底质样本需要花费大量时间和高昂的成本^[20]。考虑到半监督语义分割能够在有标签数据的基础上为无标签数据生成伪标签。因此,将半监督语义分割应用到珊瑚礁底质分类中,可以有效地降低人工成本,但是,半监督语义分割方法的性能易受伪标签噪声的干扰。

针对以上问题,本文提出了一种基于软硬协作决策的半监督珊瑚礁底质分类方法,该方法综合利用主流的语义分割模型协作决策生成高精度伪标签数据与伪标签像素的置信度,然后在一种协同决策置信损失函数(Collaboration Choice of Decision Confidence Loss Function, 3CLoss)的指导下,将生成的伪标签与有标签数据融合后进行训练,经过软硬协作决策的方式得到底质分类结果。据我们所知,本文首次将半监督学习引入底质分类领域,主要贡献如下:

(1) 提出了一个多模型协作决策的伪标签数据生成方法,该方法能有效降低伪标签数据中噪声对网络模型的干扰;

(2) 提出了一种能够有效结合伪标签置信度进行训练的3CLoss损失函数,在该损失函数的指导下,在获得高精度的底质分类模型的情况下与软硬协作决策共同作用,得到高精度的底质分类结果;

(3) 使用公开数据集进行实验,实验结果表明,本文方法精度明显优于其他主流半监督语义分割方法,并且与全监督方法精度相当。

2 研究区域与数据集

为了评估本文提出方法,我们采用美属维尔京群

岛圣克罗伊岛北部的巴克岛礁(Benthic Habitats of Buck Island Reef National Monument, BIRNM)和西北夏威夷群岛中的珍珠与爱马仕环礁(Pearl and Hermes Atoll, PHA)的浅层底栖生物栖息地地图数据集进行了实验,两个数据都由美国国家海洋和大气管理局沿海监测和评估中心年代生物地理学分支与美国国家公园管理局,通过半自动化分类和视觉解释技术相结合生成。图1a和图1c分别为BIRNM岛礁的0.5 m分辨率影像和标签数据,如表1所示,BIRNM数据集底质类型被分为4类:珊瑚、碎屑、岩石、砂,共包含675张256×256像素大小的RGB图像,其中训练、测试、无标签数据各为405、135、135张;图1b和图1d分别为PHA的10 m分辨率影像数据集和标签数据,数据集分为3类:珊瑚、碎屑和砂。PHA数据集共包含728张256×256大小的RGB图像,其中训练、验证、无标签数据分别为437、146、145张。

3 研究方法

本文提出的方法主要包括3个部分:基于多模型联合决策的伪标签生成、软硬协作底质分类与迭代训练。

3.1 基于多模型联合决策的伪标签生成

本文提出的基于多模型联合决策的伪标签生成方法如图2所示,因非重叠的子数据集可能会有不同的数据分布,所以,训练得到的网络映射也不尽相同。不同的网络模型在同一区域的预测结果是不同的,因此,多个深度语义分割网络可以互相纠正错误、共同优化标签^[21]。基于以上理论,伪标签生成方法的流程如下:

(1) 将有标签的数据集 D 分为 n 份互不重叠的子数据集 D_1, D_2, \dots, D_n ;

(2) 使用每个子数据集 D_j (j 表示子数据集序号)分别训练深度语义分割网络 $\text{Net } i$ (i 表示模型序号),得到深度模型 $M_{i,j}^p$ 。

(3) 利用得到的模型 $M_{i,j}^p$ 分别对无标签数据进行预测得到一组伪标签 $L_{i,j}^p$,将伪标签 $L_{i,j}^p$ 联合在一起得到分类矩阵 L 。

$$L = \left[\begin{array}{cccc} \left[\begin{array}{ccc} a_{0,0}^{0,0} & a_{0,1}^{0,0} & \dots & a_{0,v}^{0,0} \\ \vdots & \vdots & & \vdots \\ a_{u,0}^{0,0} & a_{u,1}^{0,0} & \dots & a_{u,v}^{0,0} \end{array} \right] & \dots & \left[\begin{array}{ccc} a_{0,0}^{i,j} & a_{0,1}^{i,j} & \dots & a_{0,v}^{i,j} \\ \vdots & \vdots & & \vdots \\ a_{u,0}^{i,j} & a_{u,1}^{i,j} & \dots & a_{u,v}^{i,j} \end{array} \right] \\ \vdots & & \vdots & \end{array} \right], \quad (1)$$

式中, $a_{u,v}^{i,j}$ 表示在第 (i,j) 组伪标签图中像素点 $I_{(u,v)}^{i,j}$ 的类别。

由于各个模型对同一区域预测结果不同,多模型

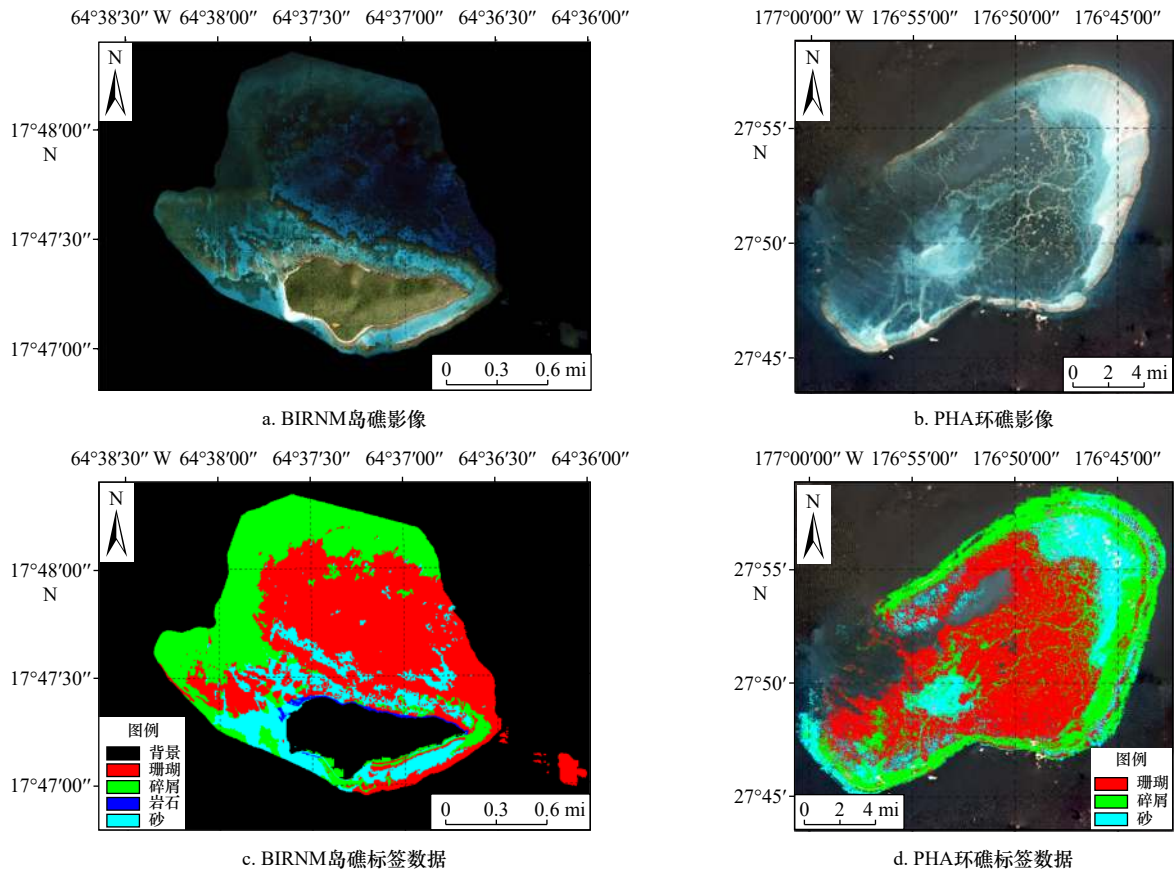


图 1 数据集影像 (a, b) 与标签数据 (c, d) (1 mi ≈ 1.609 km)

Fig. 1 Data set image (a, b) with label data (c, d) (1 mi ≈ 1.609 km)

表 1 BIRNM 数据集底质分类类型描述

Table 1 Description of substrate classification types in the BIRNM dataset

类别	描述	卫星	水下
珊瑚	连续的、高浮雕式的珊瑚形成, 形状各异, 包括平行于大陆架边缘的线性珊瑚		
碎屑	死去的、不稳定的珊瑚瓦砾, 经常被丝状物或其他大型藻类所占据。这种底质经常出现在礁顶, 珊瑚礁碎石可以在宽阔的近海沙地上以低密度聚集的方式出现		
岩石	从岛屿基岩延伸到海上的固体碳酸盐块的聚集, 或从原生床剥离和运输的松散碳酸盐碎片, 根据温特沃斯标准, 单个巨石的直径在 0.25~3 m 之间		
砂	粗糙的沉积物, 通常存在于海流或波浪能量影响的区域。颗粒大小在 1/16~256 mm 不等		

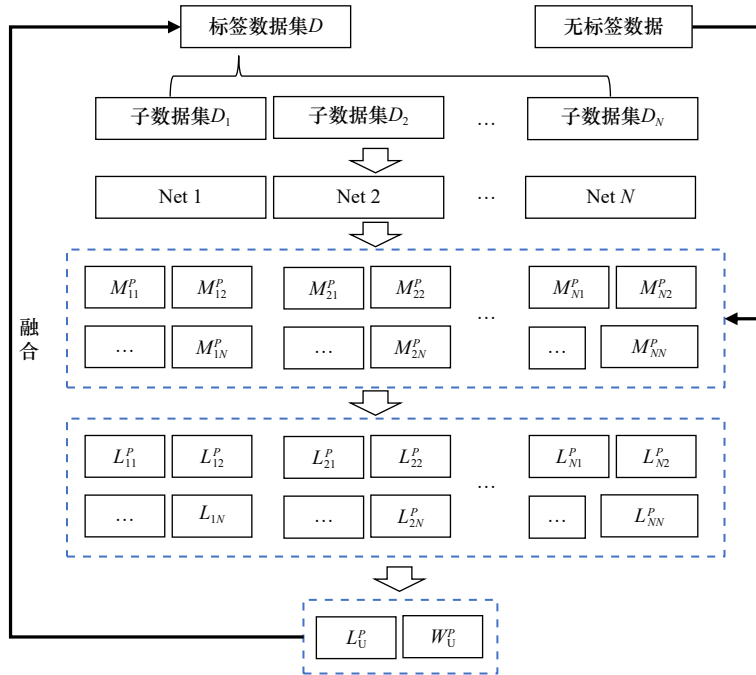


图2 基于多模型联合决策的伪标签生成

Fig. 2 Pseudo-label generation based on joint multi-model decision making

预测结果中存在不同区域的噪声,为了减少伪标签数据中噪声的影响,我们将多模型预测的伪标签 L_{ij}^p 使用硬投票的方式生成最终伪标签 L_{ij}^l ,同时记录其置信度 W 参与下一阶段的训练。

硬投票方法如下:在分类矩阵 L 中,如果某一像素点 $L_{(u,v)}^j$ 属于某一类别 c 的个数超过阈值 σ ,则保留该像素点的类别,同时记录该像素点属于类别 c 的个数,作为其置信度 $W_{u,v}$,否则,该像素不参与训练,且其置信度 $W_{u,v}$ 权重为0。其中, $\sigma = \frac{(iXj)}{2}$ 。

3.2 软硬协作底质分类

得到高质量伪标签数据之后,将其与有标签数据进行融合。由于难以避免伪标签中存在一定的噪声,为了尽可能减轻伪标签噪声对底质分类模型性能的影响,提出了一种能够顾及伪标签置信度的3CLoss损失函数指导模型进行训练。如图3所示,利用3CLoss进行训练得到高精度的分割模型之后,采用软硬协作决策的方式得到最终底质分类结果。

3.2.1 3CLoss 损失函数

设交叉熵损失(Cross Entropy, CE)表示为

$$CE = - \sum_x q(x) \log_2 p(x), \quad (2)$$

式中, q 代表真实类别值; p 表示预测类别概率值。

则本文提出的3CLoss损失函数可以由式(3)表示为

$$3CLoss = - \sum_x q(x) \mathbf{W} \times \log_2 p(x), \quad (3)$$

式中, \mathbf{W} 表示伪标签置信度权重矩阵。

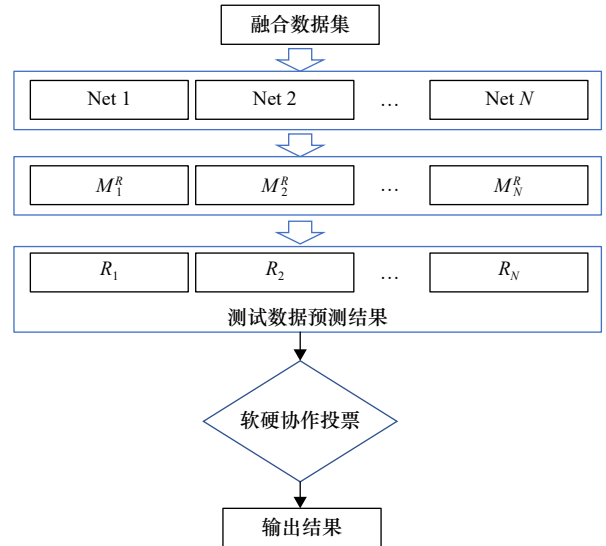


图3 3CLoss 指导下的软硬协作底质分类

Fig. 3 3CLoss-guided classification of soft and hard collaborative substrates

3.2.2 软硬协作底质分类

软硬协作的底质分类方法具体步骤如下:首先,融合后数据集在3CLoss指导下进行训练得到模型 M_i^r ,利用 M_i^r 对待分类数据进行预测,得到 R_i ;然后将 R_i 联合在一起得到每类概率矩阵 \mathbf{P} 与分类矩阵 \mathbf{L}_R 。由于每个模型预测结果 R_i 并不是一致的,若仅适用硬投票方式产生分割结果,难免会产生不包含任何类

别信息的空白像素,因此,为了解决以上问题,本文使用一种软硬协作的分类方式。具体如下:

设分类矩阵 L_R 为

$$L_R = \left[\begin{array}{cccc} \left[\begin{array}{cccc} s_{0,0}^0 & s_{0,1}^0 & \cdots & s_{0,v}^0 \\ \vdots & \vdots & & \vdots \\ s_{u,0}^0 & s_{u,1}^0 & \cdots & s_{u,v}^0 \end{array} \right] & \cdots & \left[\begin{array}{cccc} s_{0,0}^n & s_{0,1}^n & \cdots & s_{0,v}^n \\ \vdots & \vdots & & \vdots \\ s_{u,0}^n & s_{u,1}^n & \cdots & s_{u,v}^n \end{array} \right] \end{array} \right]. \quad (4)$$

则其某一像素位置 (u,v) 的分类矩阵为 $L_R^{(u,v)} = [s_{(u,v)}^0, \cdots, s_{(u,v)}^n]$, 若 $L_R^{(u,v)}$ 中出现频次最高的类别 c 大于 $n/2$, 则采用硬投票方式直接将该位置像素类别赋为 c 。

若 c 出现的频次小于 $n/2$, 则采用软投票。软投票流程如下:

设概率矩阵 P 表示为

$$P = \left[\begin{array}{cccc} \left[\begin{array}{cccc} b_{0,0,1}^1 & b_{0,0,2}^1 & \cdots & b_{0,0,c}^1, \cdots \\ \vdots & \vdots & & \vdots \\ b_{u,0,1}^1 & b_{u,0,2}^1 & \cdots & b_{u,0,c}^1, \cdots \end{array} \right] & \cdots & \left[\begin{array}{cccc} b_{0,v,1}^1 & b_{0,v,2}^1 & \cdots & b_{0,v,c}^1, \cdots \\ \vdots & \vdots & & \vdots \\ b_{u,v,1}^1 & b_{u,v,2}^1 & \cdots & b_{u,v,c}^1, \cdots \end{array} \right] \\ \left[\begin{array}{cccc} b_{0,0,1}^n & b_{0,0,2}^n & \cdots & b_{0,0,c}^n, \cdots \\ \vdots & \vdots & & \vdots \\ b_{u,0,1}^n & b_{u,0,2}^n & \cdots & b_{u,0,c}^n, \cdots \end{array} \right] & \cdots & \left[\begin{array}{cccc} b_{0,v,1}^n & b_{0,v,2}^n & \cdots & b_{0,v,c}^n, \cdots \\ \vdots & \vdots & & \vdots \\ b_{u,v,1}^n & b_{u,v,2}^n & \cdots & b_{u,v,c}^n, \cdots \end{array} \right] \end{array} \right], \quad (5)$$

式中, $b_{u,v,c}^n$ 为第 n 组分割图中像素 $Q_{(u,v)}$ 属于类别 c 的概率。则 $Q_{(u,v)}$ 属于某一类 c 的总体概率 $P_{(u,v)}^c$ 为

$$P_{(u,v)}^c = \text{SUM} \left[\left(b_{u,v,c}^1, b_{u,v,c}^2, \cdots, b_{u,v,c}^n \right) \right]. \quad (6)$$

该像素 $Q_{(u,v)}$ 的最大类别概率可由式(7)得到

$$M_{(u,v)}^p = \text{MAX} \left(P_{(u,v)}^0, \cdots, P_{(u,v)}^c, \cdots, P_{(u,v)}^n \right), \quad (7)$$

式中, C 为类别总数。得到最大类别概率之后, 与之对应的类别 c 便是像素 $Q_{(u,v)}$ 的类别。

3.3 迭代训练

上述利用多模型协同决策生成伪标签之后, 伪标签中噪声有效减少。但是单次优化后底质分类伪标签精度可能还有提升空间。因此采用多次迭代的方式持续优化标签精度, 进一步提高分割精度。迭代过程如下:

- (1) 按照 3.1 节与 3.2 节中方法完成第一次迭代;
- (2) 将 3.1 节生成的伪标签及权重平均分配到各子数据集 D_i 中, 重复 3.1 节与 3.2 节中步骤;
- (3) 重复步骤(2)。

4 实验结果与分析

4.1 实验环境与参数

本文使用实验平台搭载 64 位 Windows 10 操作系统, 平台硬件条件: 英特尔 i7-CPU, 英伟达 GeForce

RTX 2 060 SUPER 显卡(内存 8 G)。深度学习框架为 pytorch, 采用 SGD 算法进行训练调优, 所有实验网络初始学习率为 0.001, 更新梯度中使用的样本数量 (batchsize) 为 20, 每个模型训练 100 次。

在训练阶段, 将 BIRNM 和 PHA 训练数据集分别分为 3 个子数据集, 每个子数据集图片分别为 135 张和 145 张, 在多模型联合决策的伪标签生成和软硬协作底质分类阶段中, 模型数量 $N=3$, Net1、Net2、Net3 分别代表 SegFormer^[22], PanopticDeep^[23] 和 Lawin^[24] 3 种主流语义分割网。

Lawin: 允许本地窗口以很少的计算开销查询更大区域的上下文窗口。通过调节上下文区域与查询区域的比例, 我们使大窗口注意力能够在多个尺度上捕获上下文信息。

SegFormer: 一种简单、有效且鲁棒性强的语义分割方法, 改进了编码器和解码器的设计, 产生了非常高效的语义分割。缺点为仅仅依靠增加编码器的模型容量来逐步提高性能, 这可能会降低效率上限。

PanopticDeep: 一种简单、高效的自底向上的全景分割方法, 基于 DeeplabV3+ 的语义分割和基于中心点回归的实例分割的结合。缺点为固定大小的卷积核, 而且超出物体边缘范围的像素点被分割给其他类别。

3 种模型能够很好的优势互补从而更加准确而高效地进行珊瑚礁底质语义分割。

4.2 评价指标

实验所采取的定量评价指标为图像语义分割问题中常用的平均交并比 (mIoU), 计算方式如下:

$$\text{mIoU} = \frac{1}{k+1} \sum_{i=0}^k \frac{\text{TP}}{\text{FN} + \text{FP} + \text{TP}}, \quad (8)$$

式中, TP、FP 和 FN 分别代表对阳性样本的正确预测、对阳性样本的错误预测和对阴性样本的错误预测的总像素个数。

4.3 实验结果

为证明本文方法的有效性, 我们分别采用主流的半监督语义分割的方法 (Lawin, SegFormer 和 PanopticDeep) 在两个数据集上与本文方法进行了比较。

如表 2 和表 3, 本文方法相较于其他主流语义分割方法都有一定程度的提高。从表 2 可以看出, 在 BIRNM 数据集上, 本文方法分别与 Lawin 相比, mIoU 提高了 4.8%; 比 SegFormer 提高了 1.8%; 比 PanopticDeep 提高了 5.3%。如表 3 所示, 在 PHA 数据集上本文方法相比较于 Lawin、SegFormer、PanopticDeep 其 mIoU 分别提高了 1.48%、2.38%、2.74%。如图 4 所示, 本文所提方法的整体效果图与标签数据相当; 从

图5和图6可以看出, 本文方法对比其他3种半监督语义分割方法更加接近真实标签。因此, 本文方法能够较好地地区分底质类型, 有效地减少误检并且保证了同一底质类型像素的完整性。

4.4 讨论

为了验证本文提出的伪标签生成方法与3CLoss的有效性和稳定性, 本文分别进行包含3个实验的消融实验。

表2 巴克岛礁数据集半监督实验结果

Table 2 Results of semi-supervised experiments on the Buck Island dataset

方法	背景	珊瑚	碎屑	岩石	砂	mIoU
Lawin	0.930 0	0.853 2	0.729 8	0.367 0	0.740 0	0.723 3
SegFormer	0.942 8	0.863 1	0.689 5	0.513 1	0.762 5	0.754 2
PanopticDeep	0.930 4	0.851 2	0.679 0	0.442 5	0.713 8	0.718 3
本文方法	0.939 5	0.869 1	0.726 1	0.539 9	0.784 7	0.771 9

表3 珍珠与爱马仕环礁半监督实验结果

Table 3 Results of semi-supervised experiments on the Pearl and Hermes Atoll

方法	背景	珊瑚	碎屑	砂	mIoU
Lawin	0.874 1	0.675 5	0.528 5	0.452 1	0.632 5
SegFormer	0.872 5	0.661 2	0.523 1	0.437 1	0.623 5
PanopticDeep	0.866 6	0.672 6	0.497 5	0.443 1	0.619 9
本文方法	0.875 3	0.684 2	0.557 7	0.472 1	0.647 3

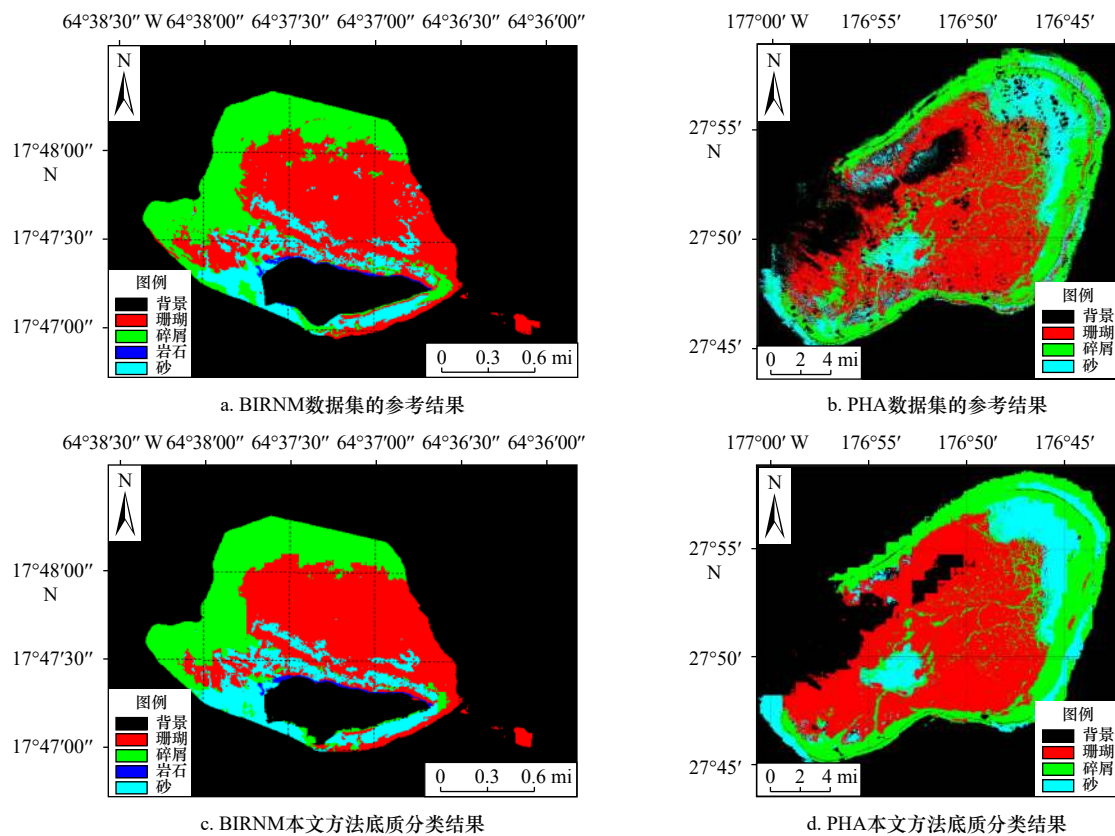


图4 BIRNM、PHA数据集的参考结果(a, c)和本文方法底质分类结果(b, d)(1 mi≈1.609 km)

Fig. 4 Reference results of BIRNM, PHA datasets (a, c) and substrate classification results of the method in this paper (b, d) (1 mi≈1.609 km)

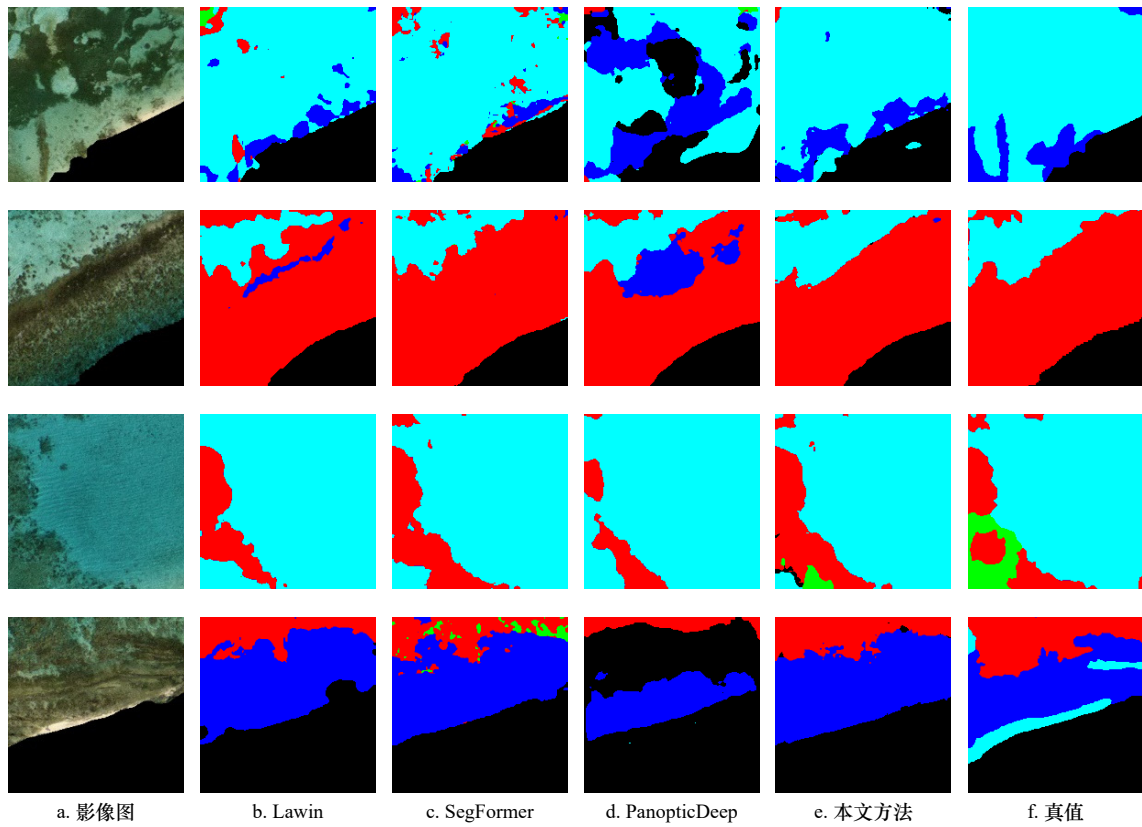


图 5 不同网络在 BIRNM 数据上的语义分割结果

Fig. 5 Semantic segmentation results of different networks on BIRNM data

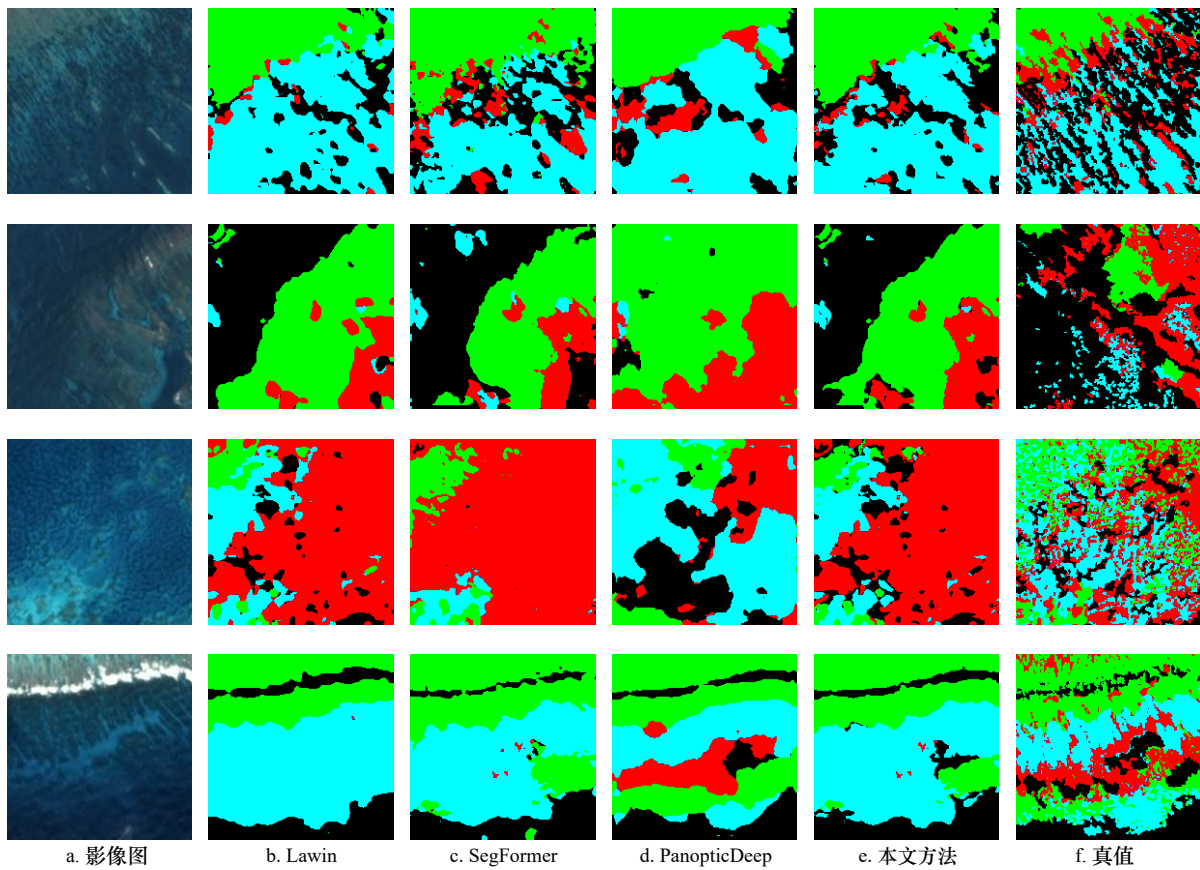


图 6 不同方法在 PHA 数据集上的语义分割结果

Fig. 6 Semantic segmentation results of different networks on PHA dataset

4.4.1 3CLoss 对实验精度的影响

为证明 3CLoss 对实验结果的改进,我们在 BIRNM 数据集上使用软硬协作半监督语义分割方法对不同

损失函数进行实验,如表 4 所示,在 3CLoss 与 Cross Entropy Loss 的指导下,本文提出的方法精度从 74.81% 提高到 77.19%,从而证明了 3CLoss 的有效性。

表 4 3CLoss 对实验精度的影响

Table 4 3CLoss effect on experimental accuracy

方法	Loss	背景	珊瑚	碎屑	岩石	砂	mIoU
本文方法	3CLoss	0.939 5	0.869 1	0.726 1	0.539 9	0.784 7	0.771 9
本文方法	Cross Entropy Loss	0.942 7	0.859 8	0.713 7	0.477 9	0.746 4	0.748 1

4.4.2 不同迭代次数对实验精度的影响

由于每次迭代训练后,为无标签数据生成的伪标签不断变化,为选择实验迭代最佳次数,本文进行多次迭代实验。如表 5 所示,迭代次数为 2 次时精度达

到最高,相较于第 1 次提高了 0.16%,第 3 次开始性能有所下降,造成以上结果的原因,可能是过多的迭代造成了模型的过拟合。以上结果说明,本文方法不需要过多的迭代训练成本就可以达到最佳性能。

表 5 迭代次数对平均交并比的影响

Table 5 Effect of number of iterations on mIoU

迭代次数	背景	珊瑚	碎屑	岩石	砂	mIoU
1	0.935 3	0.867 0	0.729 8	0.545 0	0.774 2	0.770 3
2	0.939 5	0.869 1	0.726 1	0.539 9	0.784 7	0.771 9
3	0.937 8	0.866 9	0.733 8	0.531 8	0.772 1	0.768 4

4.4.3 与全监督方法比较

为了进一步证明本文方法的有效性,将本文方法与全监督语义分割方法进行实验比对,结果如表 6 所示,在 BIRNM 数据集上本文方法相比于 Lawin、SegFormer、PanopticDeep 3 种全监督方法,精度分别提高了 1.23%、0.13%、2.81%;使用 PHA 数据训练时,本文方法与上述 3 种语义分割方法相比分别提高了 1.38%、

0.92%、2.26%。与亚米级分辨率的 BIRNM 数据相比,本文方法对米级分辨率的 PHA 数据实验结果精度较低。从图 7 可以看出,本文方法的整体分类效果与全监督方法相近,以上实验说明了本文方法底质分类效果都达到了全监督语义分割的水平,从而充分证明了所提出方法的有效性。

表 6 本文方法与全监督语义分割方法对比结果

Table 6 Comparison results between the method in this paper and the fully supervised semantic segmentation method

方法	监督方法	mIoU(BIRNM)	mIoU(PHA)
Lawin	全监督	0.759 6	0.633 5
SegFormer	全监督	0.770 6	0.638 1
PanopticDeep	全监督	0.743 8	0.624 7
本文方法	半监督	0.771 9	0.647 3

5 结论

针对基于半监督学习的珊瑚礁底质分类易受到标签噪声影响的问题,提出一种基于软硬协作决策的半监督底质分类方法。该方法使用多模型协作决策

的方式生成高质量的伪标签与伪标签置信度,之后在能够顾及伪标签像素置信度的 3CLoss 指导下进行训练得到高精度底质分类模型,最后基于分类模型采用软硬决策的方式得到高精度底质分类结果。为了验证本文方法的有效性,分别使用米级和亚米级分辨率

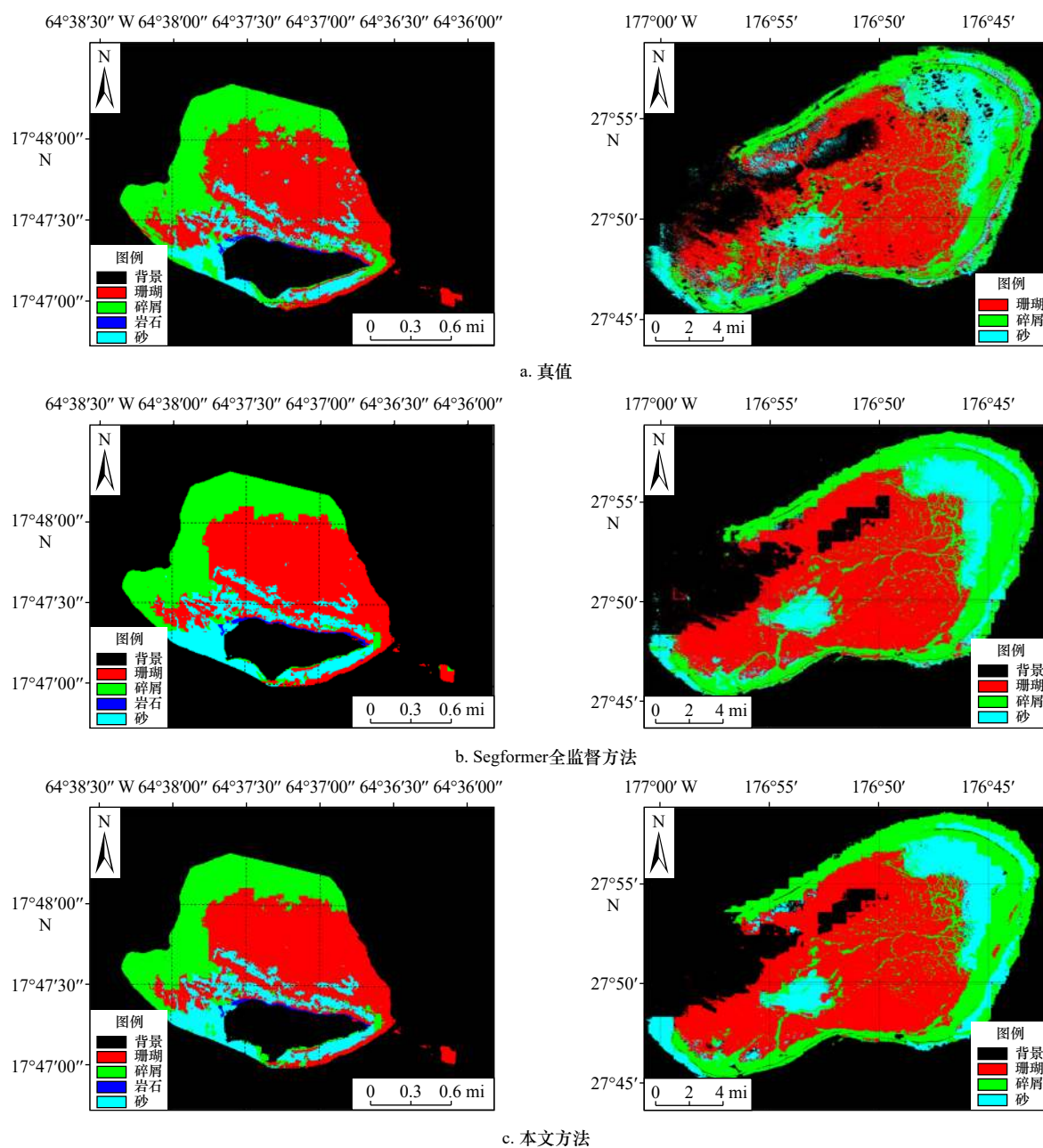


图 7 不同底质分类方法整体效果图(左列为 BIRNM, 右列为 PAH)(1 mi \approx 1.609 km)

Fig. 7 Overall effect of different substrate classification methods (the left column are BIRNM, right column are PAH) (1 mi \approx 1.609 km)

的数据集进行了详尽的实验, 结果表明:(1) 本文提出的软硬协作半监督语义分割方法在珊瑚礁底质分类方面, 准确率高于目前主流半监督方法;(2) 面对基

于不同分辨率的遥感影像底质分类任务, 提出方法的精度与全监督方法相当, 从而有效减少了实验人员对标签数据的标注工作。

参考文献:

- [1] 逢岩, 许枫, 刘佳. 基于Gammatone滤波器组时频谱和卷积神经网络的海底底质分类[J]. *应用声学*, 2021, 40(4): 510–517.
Pang Yan, Xu Feng, Liu Jia. Seabed sediment classification based on Gammatone filter banks time-frequency spectrum and convolutional neural networks[J]. *Journal of Applied Acoustics*, 2021, 40(4): 510–517.
- [2] Gregr E J, Haggarty D R, Davies S C, et al. Comprehensive marine substrate classification applied to Canada's Pacific shelf[J]. *PLoS One*, 2021, 16(10): e0259156.
- [3] Reshitnyk L, Costa M, Robinson C, et al. Evaluation of WorldView-2 and acoustic remote sensing for mapping benthic habitats in temperate coastal Pacific waters[J]. *Remote Sensing of Environment*, 2014, 153: 7–23.

- [4] Wicaksono P, Aryaguna P A. Analyses of inter-class spectral separability and classification accuracy of benthic habitat mapping using multispectral image[J]. *Remote Sensing Applications: Society and Environment*, 2020, 19: 100335.
- [5] Dempster A P, Laird N M, Rubin D B. Maximum likelihood from incomplete data via the EM algorithm[J]. *Journal of the Royal Statistical Society: Series B (Methodological)*, 1977, 39(1): 1–22.
- [6] Pillay T, Cawthra H C, Lombard A T. Characterisation of seafloor substrate using advanced processing of multibeam bathymetry, backscatter, and sidescan sonar in Table Bay, South Africa[J]. *Marine Geology*, 2020, 429: 106332.
- [7] Blaschke T. Object based image analysis for remote sensing[J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2010, 65(1): 2–16.
- [8] Furey T S, Cristianini N, Duffy N, et al. Support vector machine classification and validation of cancer tissue samples using microarray expression data[J]. *Bioinformatics*, 2000, 16(10): 906–914.
- [9] Breiman L. Random forests[J]. *Machine Learning*, 2001, 45(1): 5–32.
- [10] 万佳馨, 任广波, 马毅. 基于WorldView-2和GF-2遥感影像的赵述岛礁坪底质变化研究[J]. *海洋科学*, 2019, 43(10): 43–54.
Wan Jiaxin, Ren Guangbo, Ma Yi. Study on substrate changes of Zhaoshu reef flat based on WorldView-2 and GF-2 remote sensing images[J]. *Marine Sciences*, 2019, 43(10): 43–54.
- [11] 逢今朝, 任广波, 施祺, 等. 基于底质类型变化监测的2005–2018年西沙永乐群岛珊瑚礁白化分析[J]. *海洋科学*, 2021, 45(6): 92–106.
Pang Jinzhao, Ren Guangbo, Shi Qi, et al. Analysis of coral reef bleaching in Yongle Islands of Xisha from 2005 to 2018 based on sediment types change monitoring[J]. *Marine Sciences*, 2021, 45(6): 92–106.
- [12] 董娟, 任广波, 胡亚斌, 等. 基于高分辨率遥感的珊瑚礁地貌单元体系构建和分类方法——以8波段Worldview-2影像为例[J]. *热带海洋学报*, 2020, 39(4): 116–129.
Dong Juan, Ren Guangbo, Hu Yabin, et al. Construction and classification of coral reef geomorphic unit system based on high-resolution remote sensing: using 8-band Worldview-2 Image as an example[J]. *Journal of Tropical Oceanography*, 2020, 39(4): 116–129.
- [13] 李晓敏, 马毅, 吕喜玺. 南海珊瑚岛礁遥感分类体系和解译标志[J]. *海洋科学*, 2021, 45(5): 23–30.
Li Xiaomin, Ma Yi, Lü Xixi. Establishing a remote sensing classification system and interpretation marks for the coral islands and reefs in the South China Sea[J]. *Marine Sciences*, 2021, 45(5): 23–30.
- [14] Wan Jiaxin, Ma Yi. Multi-scale spectral-spatial remote sensing classification of coral reef habitats using CNN-SVM[J]. *Journal of Coastal Research*, 2020, 102(S1): 11–20.
- [15] Huang Rongyong, Zhang Huiya, Yu Kefu. Analysis on the live coral cover around Weizhou Island using MODIS data[J]. *Sensors*, 2019, 19(19): 4309.
- [16] 谭琨, 王雪, 杜培军. 结合深度学习和半监督学习的遥感影像分类进展[J]. *中国图象图形学报*, 2019, 24(11): 1823–1841.
Tan Kun, Wang Xue, Du Peijun. Research progress of the remote sensing classification combining deep learning and semi-supervised learning[J]. *Journal of Image and Graphics*, 2019, 24(11): 1823–1841.
- [17] King A, Bhandarkar S M, Hopkinson B M. A comparison of deep learning methods for semantic segmentation of coral reef survey images[C]//Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. Salt Lake City: IEEE, 2018: 1394–1402.
- [18] Li Jiwei, Knapp D E, Fabina N S, et al. A global coral reef probability map generated using convolutional neural networks[J]. *Coral Reefs*, 2020, 39(6): 1805–1815.
- [19] Wang Mengqiu, Hu Chuanmin. Satellite remote sensing of pelagic *Sargassum* macroalgae: the power of high resolution and deep learning[J]. *Remote Sensing of Environment*, 2021, 264: 112631.
- [20] 耿艳磊, 陶超, 沈靖, 等. 高分辨率遥感影像语义分割的半监督全卷积网络法[J]. *测绘学报*, 2020, 49(4): 499–508.
Geng Yanlei, Tao Chao, Shen Jing, et al. High-resolution remote sensing image semantic segmentation based on semi-supervised full convolution network method[J]. *Acta Geodaetica et Cartographica Sinica*, 2020, 49(4): 499–508.
- [21] 李鑫伟, 李彦胜, 张永军. 弱监督深度语义分割网络的多源遥感影像水体检测[J]. *中国图象图形学报*, 2021, 26(12): 3015–3026.
Li Xinwei, Li Yansheng, Zhang Yongjun. Weakly supervised deep semantic segmentation network for water body extraction based on multi-source remote sensing imagery[J]. *Journal of Image and Graphics*, 2021, 26(12): 3015–3026.
- [22] Xie Enze, Wang Wenhai, Yu Zhiding, et al. SegFormer: simple and efficient design for semantic segmentation with transformers[J]. *Advances in Neural Information Processing Systems*, 2021, 34: 12077–12090.
- [23] Cheng Bowen, Collins M D, Zhu Yukun, et al. Panoptic-deeplab: A simple, strong, and fast baseline for bottom-up panoptic segmentation[C]//Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. New York: IEEE, 2020: 12475–12485.
- [24] Yan Haotian, Zhang Chuang, Wu Ming. Lawin transformer: improving semantic segmentation transformer with multi-scale representations via large window attention[EB/OL]. (2022-01-05)[2022-08-15]. <https://arxiv.org/abs/2201.01615>.

A semi-supervised coral reef substrate classification method based on soft and hard collaborative decision making

Yu Jun¹, Chen Hui², Zhu Daming¹, Cheng Liang², Duan Zhixin², Zhuang Qizhi²,
Chu Sensen², Yang Wei¹, Du Siyu¹

(1. Faculty of Land and Resources Engineering, Kunming University of Technology, Kunming 650093, China; 2. School of Geography and Marine Science, Nanjing University, Nanjing 210023, China)

Abstract: Coral reef substrate classification plays a crucial role in marine resource development and marine ecological protection. At present, deep learning semantic segmentation methods are widely used in the field of remote sensing image classification, but less research has been conducted in substrate classification. Due to the high cost of pixel-by-pixel labeling in the fully supervised deep learning-based method, it is not suitable for large-scale and high-frequency substrate classification work. The semi-supervised deep learning-based method can effectively use the labeled labels to generate pseudo-labels for unlabeled data, thus effectively reducing the labor cost, however, the performance of the existing semi-supervised method is vulnerable to the interference of pseudo-label noise. To address the above problems, this paper proposes a semi-supervised substrate classification method based on soft and hard collaborative decision making. First, a high quality Pseudo tag is generated using joint decision making of multiple models; then, a loss function (Collaboration Choice of decision Confidence Loss function, 3C Loss) is proposed to take into account the confidence of Pseudo tag pixels and guide the model for training; finally, a soft and hard collaborative decision making approach is used to obtain accurate substrate classification results. The accuracy of this paper was evaluated on the shallow benthic habitat atlas datasets of Buck Island Reef in the northern part of St. Croix, U.S. Virgin Islands, and Pearl and Hermes Atolls, about 400 km southeast of Midway Island, Hawaiian Islands, and the experimental results show that the accuracy of the proposed method is comparable to that of the fully supervised learning method, and 3.08% higher than that of the mainstream semantic segmentation methods on average, which can effectively serve the coral reef substrate survey.

Key words: seabed substrate classification; soft and hard collaboration; semantic segmentation; semi-supervised learning; fully supervised learning; remote sensing; coral reefs; convolutional neural networks