

周若彤, 谭凯, 杨建儒, 等. 结合 SAM 视觉分割模型与随机森林机器学习的无人机影像盐沼植被“精灵圈”提取[J]. 海洋学报, 2024, 46(5): 116–126, doi:10.12284/hyxb2024048

Zhou Ruotong, Tan Kai, Yang Jianru, et al. Extraction of salt-marsh vegetation “fairy circles” from UAV images by the combination of SAM visual segmentation model and random forest machine learning algorithm[J]. Haiyang Xuebao, 2024, 46(5): 116–126, doi:10.12284/hyxb2024048

结合 SAM 视觉分割模型与随机森林机器学习的 无人机影像盐沼植被“精灵圈”提取

周若彤¹, 谭凯^{1*}, 杨建儒¹, 韩江涛¹, 张卫国¹

(1. 华东师范大学 河口海岸学国家重点实验室, 上海 200241)

摘要:“精灵圈”是海岸带盐沼植被生态系统中的一种“空间自组织”结构, 对盐沼湿地的生产力、稳定性和恢复力有重要影响。无人机影像是实现“精灵圈”空间位置高精度识别及解译其时空演化趋势与规律的重要数据源, 但“精灵圈”像素与背景像素在色彩信息和外形特征上差异较小, 如何从二维影像中智能精准地识别“精灵圈”像素并对识别的单个像素形成个体“精灵圈”是目前的技术难点。本文提出了一种结合分割万物模型 (Segment Anything Model, SAM) 视觉分割模型与随机森林机器学习的无人机影像“精灵圈”分割及分类方法, 实现了单个“精灵圈”的识别和提取。首先, 通过构造索伦森-骰子系数 (Sørensen-Dice coefficient, Dice) 和交并比 (Intersection over Union, IOU) 评价指标, 从 SAM 中筛选预训练模型并对其参数进行优化, 实现全自动影像分割, 得到无属性信息的分割掩码/分割类; 然后, 利用红、绿、蓝 (RGB) 三通道信息及空间二维坐标将分割掩码与原图像进行信息匹配, 构造分割掩码的特征指标, 并根据袋外数据 (Out of Bag, OOB) 误差减小及特征分布规律对特征进行分析和筛选; 最后, 利用筛选的特征对随机森林模型进行训练, 实现“精灵圈”植被、普通植被和光滩的自动识别与分类。实验结果表明: 本文方法“精灵圈”平均正确提取率 96.1%, 平均错误提取率为 9.5%, 为精准刻画“精灵圈”时空格局及海岸带无人机遥感图像处理提供了方法和技术支撑。

关键词: 盐沼植被; 精灵圈; segment anything model (SAM); 无人机影像; 机器学习

中图分类号: P237

文献标志码: A

文章编号: 0253-4193(2024)05-0116-11

1 引言

作为滨海三大“蓝碳”生态系统之一, 盐沼湿地具有极高的固碳和储碳能力、丰富的动植物资源、独特的水文与生物地球化学循环过程以及多重生态和经济价值^[1-2]。为了规避干旱、入侵、退化等诸多即将面临的生存风险, 盐沼湿地植被生态系统会自然地发生

“空间自组织现象”^[3], 其个体单元会通过短距离内的相互作用自发地形成比个体尺度大很多倍的时空有序圆状结构, 即“精灵圈”(fairy circles)。“精灵圈”对于研究盐沼植被的生长、演替、扰动、退化、恢复等过程具有重要科学价值^[4-6]。因此, 为了科学地解释其形成机理、发育过程及生态学机制, 必须快速、准确获取“精灵圈”的空间位置、尺寸大小、整体数量及

收稿日期: 2023-09-29; 修订日期: 2023-12-28。

基金项目: 国家自然科学基金项目(4217010220, 41901399); 上海市自然科学基金项目(22ZR1420900); 重庆市自然科学基金项目(CSTB2022NSCQ-MSX1254); 测绘遥感信息工程湖南省重点实验室开放基金项目(E22335); 上海市科委社发研究项目(20DZ1204700)。

作者简介: 周若彤(2000—), 女, 山东省菏泽市人, 主要从事海岸带遥感研究。E-mail: 51253904045@stu.ecnu.edu.cn

* 通信作者: 谭凯(1987—), 男, 湖南省娄底市人, 副研究员, 主要从事海岸带多源遥感及应用研究。E-mail: ktan@sklec.ecnu.edu.cn

分布格局等信息。

盐沼湿地位于潮涨潮落的近海前缘地带, 滩面泥泞, 潮沟纵横, 水文气象多变, 生境复杂, 对台站设立、现场监测及设备仪器等构成多重挑战。因此, 实地调查和样方估计等传统植被调查方法难以实现大范围“精灵圈”的时空监测^[7]。卫星遥感技术的发展为大区域、长时序“精灵圈”监测提供了新的技术思路。如 Liu 等^[8]利用高分一号卫星(GF-1)、高分二号(GF-2)卫星和中国-巴西地球资源卫星(CBERS)4星图像对黄河三角洲的“精灵圈”进行提取制图, 比较了K均值和支持向量机等方法在这3种遥感影像中“精灵圈”的检测能力。Shi 等^[9]使用航空照片与高空间分辨率卫星图像, 研究了跨越22年的7个时间段的黄河三角洲“精灵圈”时间变化序列。Zhang 等^[10]利用法国斯波特5号卫星(SPOT-5)、日本先进陆地观测卫星(ALOS)和中国资源三号卫星(ZY-3)影像分别对“精灵圈”进行制图, 并对比分析了这3类影像的提取能力。然而, 常规光学卫星遥感影像分辨率最高只能达到米级, 对尺寸较小的“精灵圈”仍存在提取能力弱、提取精度低等问题。无人机可见光遥感技术具有图像分辨率高、采集速度快、机动灵活、成本低等优点, 提供了一种在区域空间尺度上实现精准对地观测的技术手段, 可对“精灵圈”进行时空精细尺度监测及关键理化信息获取^[11]。然而, 无人机影像只有红(R)、绿(G)、蓝(B)三通道色彩信息, 如何从二维无人机影像中剔除其他地物以及常规植被的干扰, 快速、准确、自动地识别“精灵圈”个体是目前的技术难点。

近年来, 人工智能大模型发展迅速, 在自然图像分割领域, Facebook公司的Meta AI实验室发布了分割万物模型(Segment Anything Model, SAM)^[12]。SAM

可真正做到“零提示”, 对于任何陌生目标, 都不需要再进行额外的训练, 直接使用预训练模型进行分割就能达到甚至优于以往完全监督类模型的分割结果^[13-16]。因此, 将SAM应用到“精灵圈”提取, 不需要构造大量样本去训练复杂的大模型, 就可以直接针对数据进行分割, 自动生成结果。SAM针对目标对象进行无差别分割, 也就是分割所得的不仅仅是“精灵圈”掩码/类, 还包括普通植被、裸地等多种掩码/类。因此, 还需对SAM分割所得的掩码进行进一步分类。随机森林(Random Forest, RF)作为一种集成学习的机器学习方法, 具备很好的遥感图像分类性能^[17-22], 因此可采用RF对SAM分割得到的掩码进行进一步分类。

无人机影像在分辨率、精度、质量等方面都不同于一般日常数字影像, 其数据大小也存在量级差异, 从无人机复杂场景海量影像数据中进行高精度分割比常规影像要困难很多, 对SAM的通用性和鲁棒性也提出了很大挑战。截至目前, SAM在遥感领域还鲜有应用, SAM在无人机遥感图像处理和解译中的性能和精度尚不清楚。本文提出了一种结合SAM视觉分割模型和RF机器学习的无人机影像“精灵圈”提取方法, 证实了SAM在无人机遥感图像处理中的巨大潜在价值, 为“精灵圈”时空格局智能反演及无人机遥感图像智能解译提供了方法和技术支撑。

2 研究区域概况与数据

2.1 研究区域

本文选取上海市崇明岛东滩鸟类自然保护区北部的某潮间带作为研究区域, 该区域长约2.0 km, 宽约1.3 km(图1白色矩形框)。该研究区域覆盖有多种类型盐沼植被, 主要有芦苇(*Phragmites australis*)、

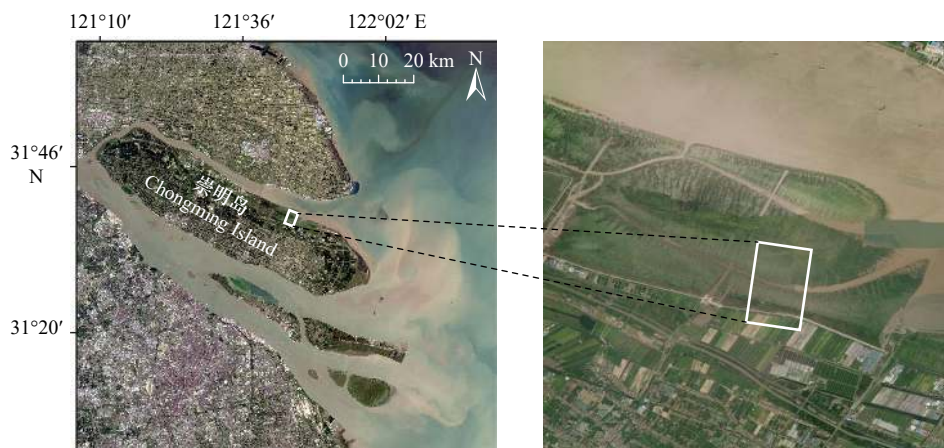


图1 研究区域概况

Fig. 1 Overview of the study area

互花米草 (*Spartina alterniflora*) 和海三棱蔗草 (*Bolboschoenoplectus mariqueter*) 等^[8]。在研究区中部横跨一条东西走向的潮沟, 在该潮沟两侧连接了数条南北走向的狭窄潮沟分支。靠近潮沟两侧的区域为光滩裸地, 其余地方均分布有茂密的植被。

由于“精灵圈”主要出现在潮沟两侧的光滩或植被稀疏的区域, 大片且茂盛的植被区域并没有“精灵

圈”存在^[8], 因此本文共选取研究区域中 I、II 两个“精灵圈”密集分布的典型区域展开实验分析 (图 2a, 图 2c)。I 区域大小约 $0.24 \text{ km} \times 0.20 \text{ km}$, II 区域大小约 $0.37 \text{ km} \times 0.15 \text{ km}$ 。两个区域的“精灵圈”分布特征不同: I 区域的“精灵圈”分布较为分散, 没有明显位置规律; 而 II 区域的“精灵圈”分布较为集中, 主要分布在研究区域中部潮沟的北侧。

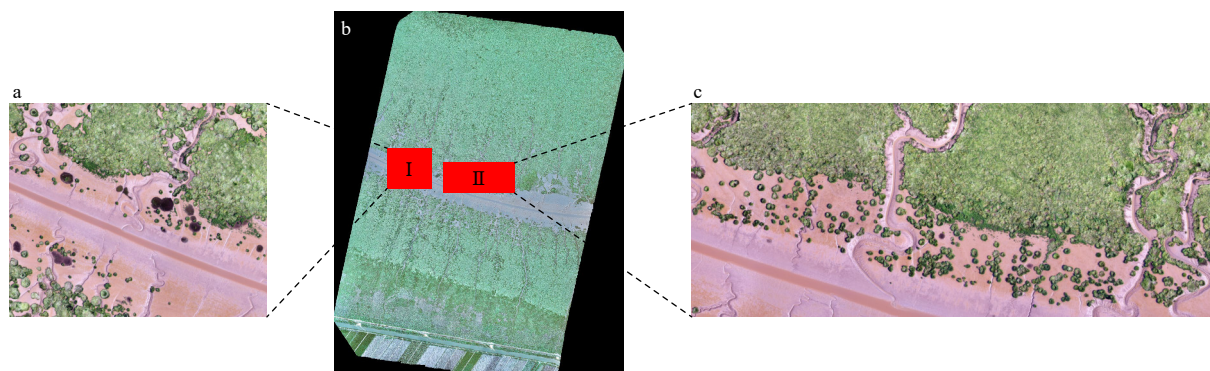


图 2 I 区域正射影像 (a)、整个研究区域正射影像 (b) 和 II 区域正射影像 (c)

Fig. 2 Orthophoto of Region I (a), orthophoto of the entire study area (b), and orthophoto of Region II (c)

2.2 实验数据采集与处理

本文实验数据采集时间为 2019 年 8 月 17 日, 利用 ZR-M66 多旋翼无人机平台搭载索尼 RX1R II 相机拍摄所得 (图 3a)。该相机的有效像素为 4 240 万, 镜头焦距为 35 mm。无人机飞行高度为 200 m, 飞行速度 12 m/s, 影像共采集了 16 个航带 (图 3b), 旁向和航向重叠率分别为 60% 和 80%, 共计拍摄相片 1 566 张。利用 PIX4D 软件进行姿态数据解析、航空三角测量

等处理, 最终得到整个研究区的正射影像 (图 2b), 影像的空间分辨率约为 2 cm, 整体像素为 $67\,584 \times 96\,325$ 。其中, I 区域像素为 $11\,707 \times 9\,769$, II 区域像素为 $17\,660 \times 7\,202$ 。I 区域作为模型训练区, 将其 SAM 分割后的掩码集合按照 7:3 分为训练集和测试集, II 区域作为模型验证区。通过目视解译人工手动在正射影像上进行逐个标注, 获取“精灵圈”真实数据。

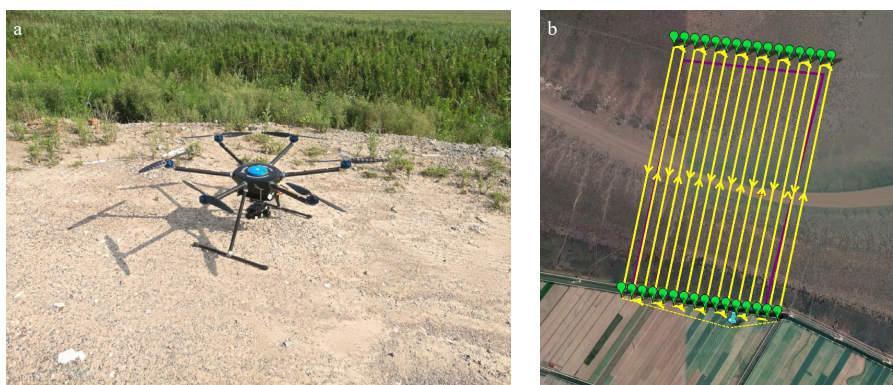


图 3 无人机摄像系统 (a) 和数据采集飞行路线 (b)

Fig. 3 UAV camera system (a) and data collection flight path (b)

3 研究方法

本文方法总体技术路线如图 4 所示。首先, 对研究区影像进行裁剪、重采样等预处理; 其次, 从 SAM 中筛选预训练模型结构并对模型的相关参数进

行调整, 实现全自动影像分割, 得到无属性信息的系列分割掩码/分割类; 然后利用空间二维坐标信息以及 RGB 色彩信息, 将分割掩码与无人机影像进行信息匹配, 构造并筛选分割掩码的系列 RGB 指数特征以及空间位置与形态指标特征; 最后, 利用筛选的特

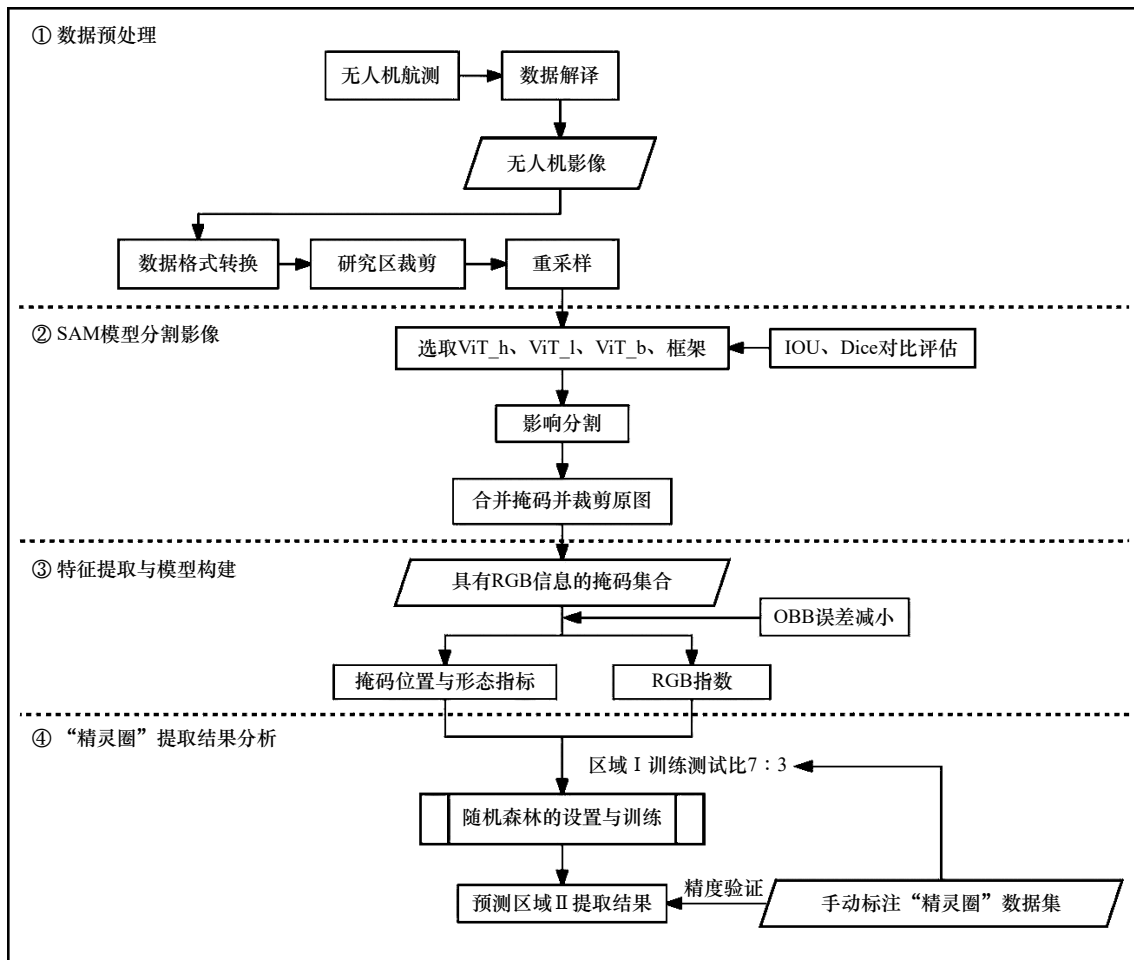


图4 研究方法流程

Fig. 4 Research method flow

征对 RF 模型进行训练, 实现“精灵圈”植被、普通植被和光滩的识别与分类。

3.1 SAM 模型

SAM 主要包括 3 个部分: 图像编码器、提示编码器和轻量级的掩码解码器^[13]。其中, 图像编码器用于计算图像嵌入, 提示编码器用于嵌入提示, 而掩码解码器用于将以上两个信息源组合在一起以达到预测分割掩码的目的。SAM 共有 3 个交互模式: 点击模式 (Click)、框选模式 (Box) 和任意模式 (Everything), 其中前两个模式需要人机互动, 而 Everything 模式可在整个检测范围内进行自动图像分割^[13]。遥感图像涉及范围广、数量大, 快速且不依靠人工的全自动识别至关重要。因此, 本文选用 Everything 模式来完成“精灵圈”图像的分割。

SAM 使用计算机视觉中的算法视觉变压器 (Vision Transformer, ViT) 作为其预训练模型, ViT 有 3 种模型变体: ViT_h (High resolution)、ViT_l (Low resolution) 和 ViT_b (Base resolution), 每种模型对应不同的分割效果和分割效率。索伦森-骰子系数 (Sørensen-

Dice Coefficient, Dice) 和交并比 (Intersection Over Union, IOU) 是分割网络中最常用的评价指标, 能够衡量分割结果与标准掩码的相似性, 客观且直接的评价分割模型的性能^[15-16]。因此, 本文使用 Dice 和 IOU 来评估 ViT_h、ViT_l、ViT_b “精灵圈”提取的精度, 并最终确定适用的模型。Dice 和 IOU 的计算如式 (1) 所示。

$$\begin{cases} \text{Dice} = \frac{2|X \cap X'|}{|X| + |X'|} \\ \text{IOU} = \frac{|X \cap X'|}{|X| + |X'| - |X \cap X'|} \end{cases}, \quad (1)$$

式中, X 为真实掩码; X' 为对应模型预测所得的掩码。

此外, 预训练模型中自动掩码生成器的默认参数不能有效用于“精灵圈”的分割, 本文对其中的关键参数进行调整, 包括: points_per_batch, pred_iou_thresh 和 stability_score_thresh。其中, points_per_batch 代表模型同时运行的点数, 代表了运行速度和效率, 合适的数值能同时保证 GPU (计算机图形处理器) 运行的畅通以及模型分割的最大效率, 本文将其设置为 8。pred_iou_thresh 代表了分割所得的每个掩码的质量, 取

值范围为 0~1, 分割所得的掩码质量参差不齐, 因此需要舍弃质量较低的掩码, 本文设置为 0.8。stability_score_thresh 取值范围为 0~1, 代表了每个掩码的稳定性, 即对于整个分割影像中背景区域的划分, 该参数的大小决定了整个无人机影像对于分割类别的精细程度, 因此需要舍弃稳定性较差的掩码, 本文将其设置为 0.88。

3.2 RF 算法及分类特征指标构建

RF 是一个集成分类器, 具有随机性、高效性、鲁棒性和可解释性等优势^[23-25]。RF 由多个决策树组成, 每个决策树都独立且随机地选择特征子集和样本子集, 然后通过投票或平均值的方式进行决策, 这大大增加了模型的泛化能力, 有效减小过拟合的风险^[24]。RF 可以通过对决策树数量、最小叶子数等一些参数的调整从而不断优化模型, 使其更加贴合指定任务。本文在 RF 模型训练中参数设置值如下: 决策树数目 300, 最小叶子树 8。

“精灵圈”与“非精灵圈”在色彩、形态、空间位置等存在差异, 因此, 可以构建相应的特征训练 RF 模型, 对二者进行识别。无人机影像中可供选取的特征主要有 RGB 指数特征以及空间位置与形态指标特征。其中, 常用 RGB 指数特征包括: 超绿指数 (Excess Green Index, EXG)、绿色比值指数 (Green Ratio Index, GCC)、绿红植被指数 (Green-Red Vegetation Index, GRVI)、kawashima 指数 (kawashima Index, IKAW)、改进型绿红植被指数 (Modify Green-Red Vegetation Index, MGRVI)、改进型可见光大气阻抗指数 (Modify Visible Atmospherically Resistant Index, MVARI)、红绿蓝植被指数 (RGB Vegetation Index, RGBVI)、三角绿度指数 (Triangular Greenness Index, TGI)、可见光大气阻抗指数 (Visible Atmospherically Resistant Index, VARI) 及可见光差值植被指数 (Visible-band Difference Vegetation Index, VDVI)^[26-27](表 1)。由于“精灵圈”主要表现为圆形或准原形特征, 且“精灵圈”掩码区域的面积明显小于与大块植被和裸地掩码的面积, 因此可采用掩码圆度 C 、面积 Area、宽度 W 、高度 H 、纵横比 W/H 、横坐标 bbox-x0 及纵坐标 bbox-y0 作为 RF“精灵圈”提取时的空间位置与形态特征。其中, 掩码圆度计算公式如下^[7]:

$$C = \frac{|X_{size} - Y_{size}|}{\max(X_{size}, Y_{size})}, \quad (2)$$

式中, X_{size} 和 Y_{size} 分别为掩码 X 轴和 Y 轴方向上的大小。

RF 分类任务中特征值的选取十分关键, 过多的特征值不仅会大大提高模型的复杂度, 也可能会面临

表 1 RGB 植被指数
Table 1 RGB vegetation index

指数	公式
EXG	$2 \times G - R - B$
GCC	$G/B + G + R$
GRVI	$(G - R)/(G + R)$
IKAW	$(R - B)/(R + B)$
MGRVI	$(G^2 - R^2)/(G^2 + R^2)$
MVARI	$(G - B)/(G + R - B)$
RGBVI	$(G^2 - B \times R)/(G^2 + B \times R)$
TGI	$G - (0.39 \times R) - (0.61 \times B)$
VARI	$(G - R)/(G + R - B)$
VDVI	$(2 \times G - R - B)/(2 \times G + R + B)$

过拟合的风险^[25]。因此, 根据每个特征的分布规律以及每个特征在不同类型掩码之间的差异性, 利用袋外数据 (Out of Bag, OOB) 误差减小值可评估每个特征的重要性^[23], 对上述 17 个特征进行分析和筛选, 剔除不重要的特征, 减少冗余性和相关性, 确定最终选定的特征, 用以训练 RF 模型。其中, 基于 OOB 数据置换的指标的特征重要性计算过程如下:

(1) 假设 RF 共有 N 棵决策树 $K = \{k_1, k_2, \dots, k_n\}$, 首先利用每棵决策树 k_n 对袋外数据 OOB 数据进行预测, 预测值 Y_n 与真实值 Y 的误差为 OOB_n 。

(2) 在保证 OOB 其他特征值不变的情况下, 将其第 i 个特征值的顺序进行随机打乱, 然后利用决策树 k_n 再对 OOB 数据及逆行预测, 所得到的预测结果为 Y_n^i , 这时预测值 Y_n 与真实值 Y 的误差为 OOB_n^i 。

(3) 第 i 个特征值的顺序打乱前后的均方根误差之差则可以表示该特征值对决策树 k_n 预测精度的影响, 表示为 $errOOB_n^i = OOB_n^i - OOB_n$ 。

(4) 重复(1)~(3)的步骤, 可以得到第 i 个特征值对所有 n 棵决策树的影响, 则该特征指标对 RF 准确率的影响, 即重要性为

$$FIM = \frac{1}{N} \sum_{n=1}^N errOOB_n^i, \quad (3)$$

式中, N 为 RF 决策树的数量; $errOOB_n^i$ 表示第 i 个特征值对决策树 k_n 预测精度的影响, FIM 即为第 i 个特征指标的重要性。

4 实验结果与分析

4.1 SAM 模型性能评估结果

在研究区 I 中, SAM 的 3 个预训练模型表现如下: ViT_b 模型的平均 IOU 为 0.701 7, 平均 Dice 为 0.826 6; ViT_l 模型的平均 IOU 为 0.752 3, 平均 Dice 为 0.883 5; ViT_h 模型的平均 IOU 为 0.762 5, 平均 Dice 为 0.883 7。

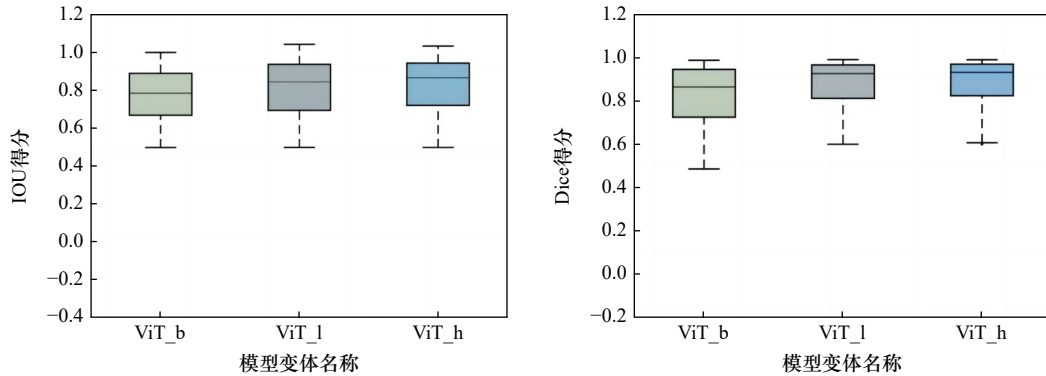


图 5 ViT 3 种模型变体的 IOU 与 Dice 箱式图

Fig. 5 Box chart of IOU and Dice for the three variants of ViT model

4.2 RF 样本特征值的选取

本文选用研究区 I 的数据进行特征重要性的计算, 根据 OOB 误差减小得到 17 个特征的重要性(图 6)。在植被指数中, EXG 的重要性最高, MVARI 和 IKAW 的重要性最低。位置形态指标中, 面积 Area、圆度 C、纵横比 W/H 表现良好, 而掩码横坐标 bbox-x0 和纵坐标 bbox-y0 重要性最低。

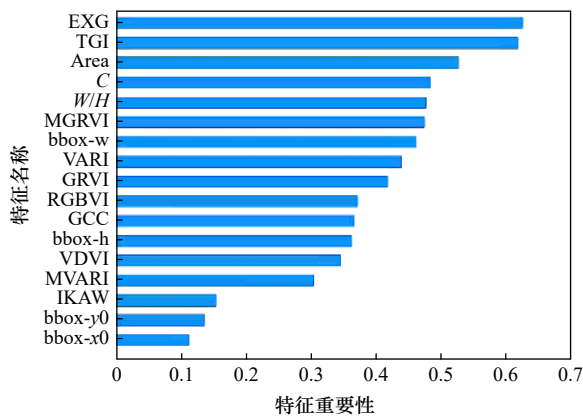


图 6 特征指标重要性情况

Fig. 6 Importance of the selected indicators

将特征值归一化, 得到特征值分布(图 7)。由图 7a 可知, 掩码圆度 C、面积 Area、宽度 bbox-w、高度 bbox-h 和纵横比 W/H 这 5 个掩码位置形态指标都随不同类别变化存在较为明显的差距与规律性, 主要表现为:“精灵圈”掩码的宽度、高度、圆度、纵横比以及面积大小都恒定在一定的范围内, 而普通植被和光滩的这些位置形态指标表现不稳定, 会有较大波动。对应的, 由图 7b 可知, 掩码横坐标 bbox-x0 和纵坐标 bbox-y0 在 3 种类别掩码中并未有明显区别, 整体分

3 个模型的 IOU 和 Dice 得分整体表现如图 5 所示, ViT_h 模型在 3 个模型变体中表现最佳。因此尽管 ViT_h 中具有较多的参数、模型更为复杂, 本研究最终选择其作为无人机影像分割预训练模型变体。

布杂乱无章。由图 7c 可知, EXG、GCC、GRVI、MGRVI、RGBVI、TGI、VARI、VDVI 这 8 个 RGB 植被指数均能很好地区分“精灵圈”、光滩和普通植被。对应的, 由图 7d 可知, IKAW 与 MVARI 这两个植被指数分布杂乱无章, 在 3 类掩码中并未有明显区分度。

结合上述对特征值分布的定量评估与定性分析, 本文将 bbox-x0、bbox-y0、IKAW 和 MVARI 这 4 个特征进行删除, 最终利用 EXG、GCC、GRVI、MGRVI、RGBVI、TGI、VARI、VDVI、掩码圆度 C、掩码面积 Area、掩码宽度 W、掩码高度 H 和掩码纵横比 W/H 这 13 个特征来进行基于 RF 的训练和分类提取。

4.3 “精灵圈”提取结果

本文使用影像叠加显示进行视觉层面的“精灵圈”提取直观定性展示, 使用混淆矩阵进行数据层面的定量评价, 计算“精灵圈”的正确识别率和错误识别率评估^[28]。其中, 正确识别率为使用本文方法提取所得的正确“精灵圈”数量占研究区域内所有“精灵圈”数量的比值, 错误识别率为使用本文方法提取所得的被误认为“精灵圈”的“非精灵圈”数量占研究区域内共计识别出“精灵圈”数量的比值。理论上, 区域 I 在本文中为训练集, 仅用于 RF 模型的训练和测试, 训练出的 RF 模型不需再预测区域 I 的“精灵圈”。考虑到区域 I 和 II“精灵圈”分布特征差异较大, 为了验证模型的普适性和鲁棒性, 同时对比分析两个区域“精灵圈”提取的差异性, 本文将由区域 I 训练好的模型同时对区域 I 和 II 进行了“精灵圈”预测。

研究区 I 的 SAM 模型分割结果以及“精灵圈”植

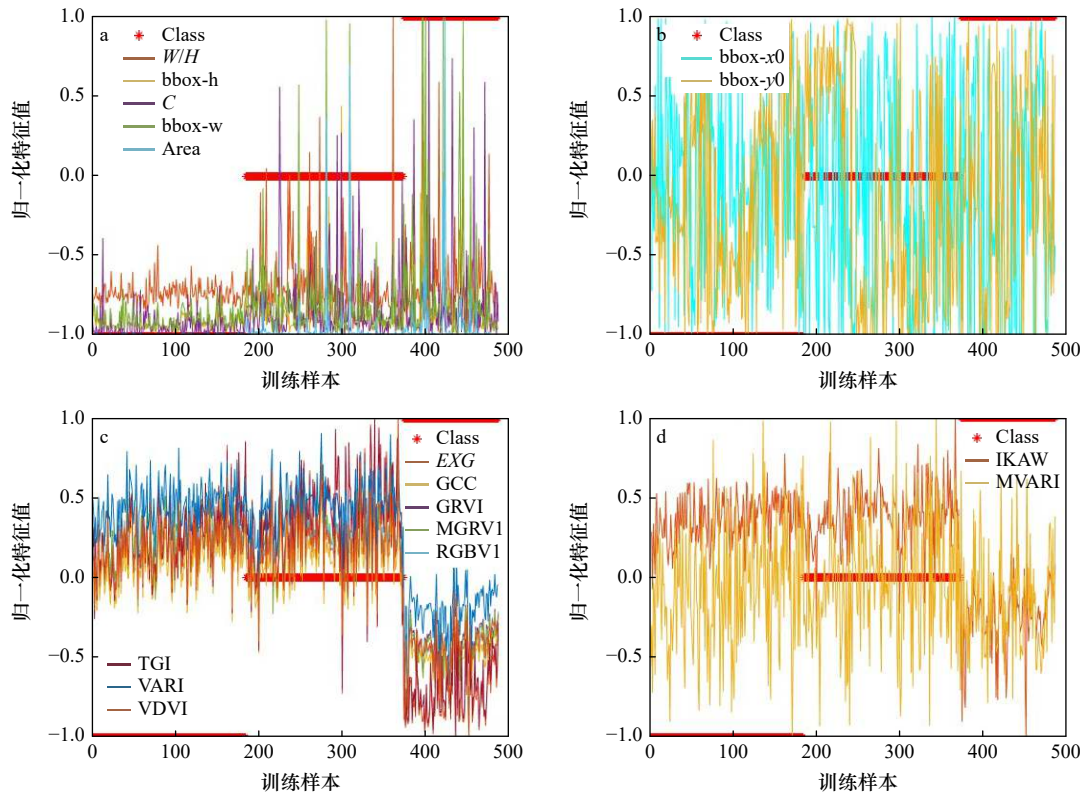


图 7 位置与形态指标分布(a, b)和 RGB 指数指标分布(c, d)

Fig. 7 Distributions of position and shape indexes (a, b), and distributions of RGB indexes (c, d)

横坐标为样本序号, 纵坐标为特征值, 红色星号代表掩码的类型(从左到右分别为“精灵圈”植被、普通植被和光滩), 彩色折线代表特征值的分布情况

Where the horizontal coordinate is the sample serial number, the vertical coordinate is the feature index value, the red asterisk represents the type of mask (from left to right are "fairy circle" vegetation, common vegetation, and bare mudflat), and the color line represents the distribution of feature value

被掩码提取结果如图 8 所示。可以看出, SAM 在识别“精灵圈”的同时, 会将一些光滩和大块背景植被无差别分割出来。背景植被和光滩掩码明显区别于较为规律的“精灵圈”掩码, “精灵圈”掩码呈近圆形且大小形态参数较为稳定, 因此进一步利用 RF 的机器学习方法, 根据本文选取的 13 个特征值进行训练, 可将图 8b 所示的掩码集合最终将分成“精灵圈”植被、背景植被、光滩植被这 3 种类型。最终提取的“精灵圈”如图 8c 所示。研究区 I 中“精灵圈”实际为 184 个, 识别出的“精灵圈”为 200 个。其中, 有 172 个被正确识别, 正确识别率 93.5%; 有 28 个“非精灵圈”被误识别为“精灵圈”, 误识别率 14%(表 2)。

研究区 II 的 SAM 模型分割结果以及“精灵圈”植被掩码提取结果如图 9 所示。研究区 II 中“精灵圈”实际为 364 个, 识别出的“精灵圈”为 378 个。其中, 有 359 个被正确识别, 正确识别率 98.6%; 有 19 个“非精灵圈”被误识别为“精灵圈”, 误识别率 5%(表 2)。虽然 RF 模型由研究区 I 训练得到, 但是当训练出的模型用于两个区域时, 研究区 II 的“精灵圈”正确识别

率反而明显高于研究区 I, 误识别率也较小。主要原因是研究区 I 中的部分“精灵圈”发育不完全, 并未离群, 背景植被不呈大块分布, 导致部分背景植被的 SAM 掩码与“精灵圈”掩码极为相似, 从而出现误提。研究区 I 训练的模型应用于研究区 II 得到了十分精确的分类结果, 这表明了本文方法具有很好的可迁移性和稳健性。值得注意的是, 本文是基于无人机正射影像进行“精灵圈”提取的, 但本文方法是可以适用于常规无人机影像的。常规无人机影像存在几何畸变和失真, 提取出来的“精灵圈”不利于后期的进一步成图和分析。

本文基于“精灵圈”的 RGB 特征以及空间位置与形态特征, 共选取了 17 个特征(详见 3.2 小节), 经特征重要性及分布分析, 最终共保留了 13 个主要的特征(详见 4.2 小节)。为了验证剔除的 4 个参数对提取结果并无实际影响, 本文基于同样的参数设置使用 17 个特征对区域 I 进行训练, 利用得到的模型对区域 I 和 II“精灵圈”进行提取(图 8e 和图 9e)。如表 2 所示, 删除 bbox-x0、bbox-y0、IKAW 和 MVARI 这 4 个

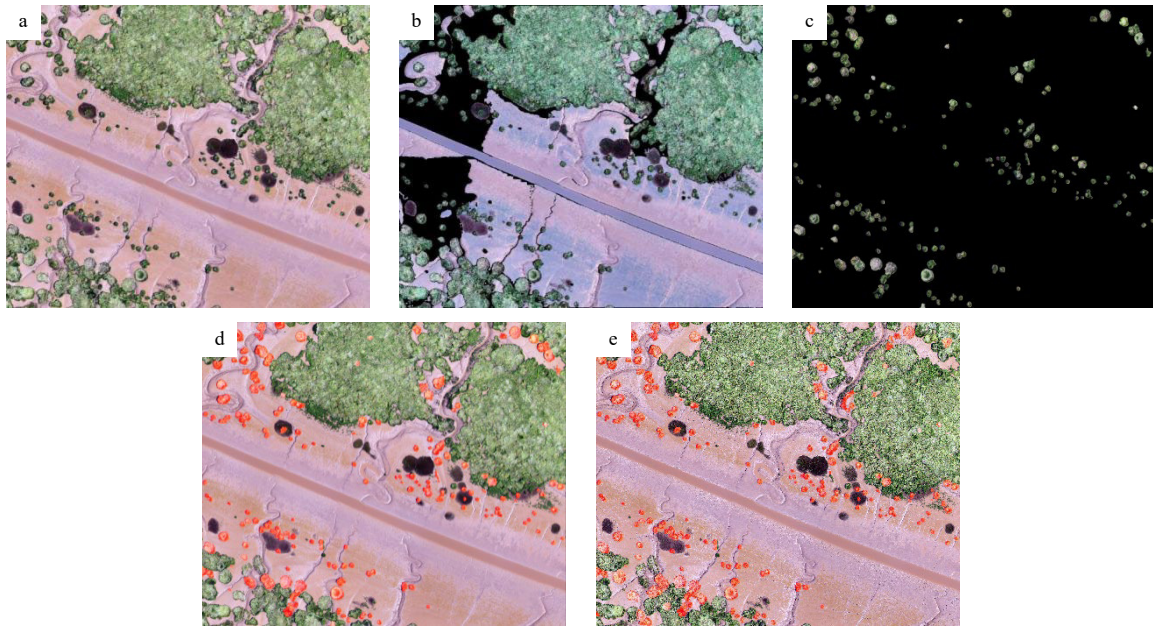


图 8 研究区 I 提取结果

Fig. 8 Extraction results of Region I

- a. 原始正射影像; b. SAM 分割结果 RGB 显示; c. RF 分类结果; d. “精灵圈”提取结果叠加原始正射影像; e. 删除特征 bbox-x0、bbox-y0、IKAW 和 MVARI 4 个特征前“精灵圈”提取结果
- a. Original orthophoto; b. RGB display of SAM segmentation results; c. random forest classification result; d. "fairy circle" extraction result superimposed on original orthophoto; e. "fairy circle" extraction result superimposed on original orthophoto before removing the features of bbox-x0, bbox-y0, IKAW, and MVARI

表 2 “精灵圈”提取混淆矩阵, 类别 1、2、3 分别代表“精灵圈”、背景植被和光滩

Table 2 “Fairy circle” extraction confusion matrix, and categories 1, 2 and 3 represent fairy circle, background vegetation, and bare flat, respectively

研究区I(13个特征)				研究区II(13个特征)			
实际类别	预测类别			实际类别	预测类别		
	1	2	3		1	2	3
1	172	11	1	1	359	4	1
2	27	157	5	2	18	35	0
3	1	1	112	3	1	1	39
正确识别率	93.5%			正确识别率	98.6%		
错误识别率	14%			错误识别率	5.0%		

研究区I(17个特征)				研究区II(17个特征)			
实际类别	预测类别			实际类别	预测类别		
	1	2	1		1	2	3
1	170	14	0	1	354	4	1
2	35	151	3	2	31	35	0
3	3	1	110	3	3	1	39
正确识别率	92.4%			正确识别率	97.3%		
错误识别率	18.3%			错误识别率	8.8%		

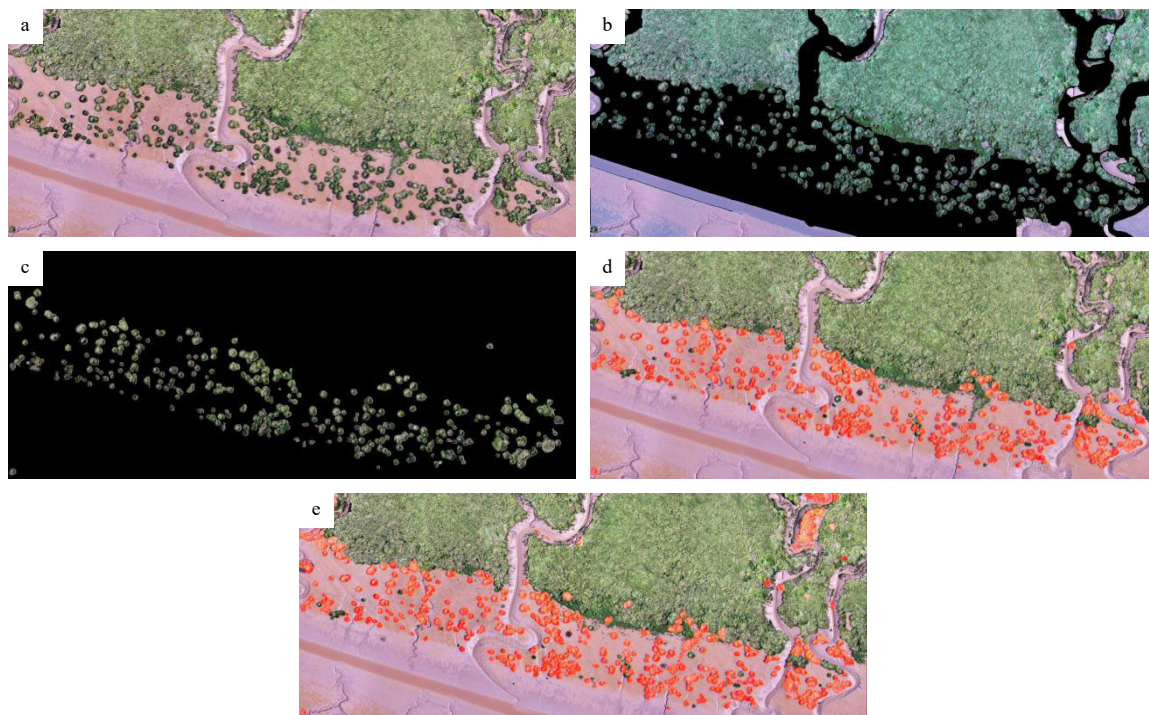


图 9 研究区 II 提取结果

Fig. 9 Extraction results of Region II

a. 原始正射影像; b. SAM 分割结果 RGB 显示; c. RF 分类结果; d. “精灵圈”提取结果叠加原始正射影像; e. 删除特征 bbox-x0、bbox-y0、IKAW 和 MVARI 4 个特征前“精灵圈”提取结果

a. Original orthophoto; b. RGB display of SAM segmentation results; c. random forest classification result; d. "fairy circle" extraction result superimposed on original orthophoto; e. "fairy circle" extraction result superimposed on original orthophoto before removing the features of bbox-x0, bbox-y0, IKAW, and MVARI

特征解决了数据冗余问题,同时提升了“精灵圈”正确识别率约 1.2%,降低了错误识别率约 4.1%。

5 结论

本文提出了一种结合 SAM 计算机视觉分割模型与 RF 机器学习的无人机光学影像“精灵圈”识别和提取方法,利用 2019 年 8 月 17 日上海市崇明岛东滩北部的两景潮间带无人机影像示例数据开展试验,其结果为:“精灵圈”正确识别率平均为 96.1%,错误识别率

平均为 9.5%。本文分析和证明了 SAM 在无人机影像中的良好分割性能,拓展了 SAM 在遥感图像处理和解释中的潜在应用,为研究“精灵圈”形成、分布、时空演变,以及潮间带生态系统和非生物因素之间复杂的相互作用等提供重要支撑。然而, SAM 目前还是一个通用视觉大模型,对于执行不同的下游任务时可进一步根据场景进行预训练和改进,使其可以更好的针对特定任务场景或特殊数据进行图像分割。此外,“精灵圈”植被与普通植被的精细区分还需进一步优化。

参考文献:

- [1] 韩倩倩,牛振国,吴孟泉,等. 基于潮位校正的中国潮间带遥感监测及变化[J]. *科学通报*, 2019, 64(4): 456-473.
Han Qianqian, Niu Zhenguo, Wu Mengquan, et al. Remote-sensing monitoring and analysis of China intertidal zone changes based on tidal correction[J]. *Chinese Science Bulletin*, 2019, 64(4): 456-473.
- [2] 陈一宁,陈鹭真. 滨海蓝碳生态系统的碳库间相互作用研究进展及展望[J]. *海洋学研究*, 2023, 41(1): 3-13.
Chen Yining, Chen Luzhen. Interactions between vegetation and sediment carbon pools within coastal blue carbon ecosystems: A review and perspective[J]. *Journal of Marine Sciences*, 2023, 41(1): 3-13.
- [3] Zhao Lixia, Zhang Kang, Siteur K, et al. Fairy circles reveal the resilience of self-organized salt marshes[J]. *Science Advances*, 2021, 7(6): eabe1100.
- [4] Rietkerk M, Bastiaansen R, Banerjee S, et al. Evasion of tipping in complex systems through spatial pattern formation[J]. *Science*, 2021, 374(6564): eabj0359.
- [5] Ruiz-Reynés D, Gomila D, Sintés T, et al. Fairy circle landscapes under the sea[J]. *Science Advances*, 2017, 3(8): e1603262.

- [6] de Paoli H, van der Heide T, van den Berg A, et al. Behavioral self-organization underlies the resilience of a coastal ecosystem[J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2017, 114(30): 8035–8040.
- [7] Tao Pengjie, Tan Kai, Ke Tao, et al. Recognition of ecological vegetation fairy circles in intertidal salt marshes from UAV LiDAR point clouds[J]. *International Journal of Applied Earth Observation and Geoinformation*, 2022, 114: 103029.
- [8] Liu Qingsheng, Huang Chong, Liu Gaohuan, et al. Comparison of CBERS-04, GF-1, and GF-2 satellite panchromatic images for mapping quasi-circular vegetation patches in the Yellow River Delta, China[J]. *Sensors*, 2018, 18(8): 2733.
- [9] Shi Lei, Liu Qingsheng, Huang Chong, et al. Mapping quasi-circular vegetation patch dynamics in the Yellow River Delta, China, between 1994 and 2016[J]. *Ecological Indicators*, 2021, 126: 107656.
- [10] Zhang Yunjie, Liu Qingsheng, Liu Gaohuan, et al. Mapping of circular or elliptical vegetation community patches: A comparative use of SPOT-5, ALOS and ZY-3 imagery[C]//Proceedings of the 8th International Congress on Image and Signal Processing. Shenyang: IEEE, 2015.
- [11] Zhang Xianlong, Zhang Fei, Qi Yaxiao, et al. New research methods for vegetation information extraction based on visible light remote sensing images from an unmanned aerial vehicle (UAV)[J]. *International Journal of Applied Earth Observation and Geoinformation*, 2019, 78: 215–226.
- [12] Kirillov A, Mintun E, Ravi N, et al. Segment Anything[C]//2023 IEEE/CVF International Conference on Computer Vision (ICCV). 0[2024-02-29]. DOI:10.1109/ICCV51070.2023.00371.
- [13] Ning Guochen, Liang Hanyin, Jiang Zhongliang, et al. The potential of 'Segment Anything' (SAM) for universal intelligent ultrasound image guidance[J]. *Bioscience Trends*, 2023, 17(3): 230–233.
- [14] Chen Fang, Chen Lingyu, Han Haojie, et al. The ability of Segmenting Anything Model (SAM) to segment ultrasound images[J]. *Bioscience Trends*, 2023, 17(3): 211–218.
- [15] Shi Peilun, Qiu Jianing, Abaxi S M D, et al. Generalist vision foundation models for medical imaging: A case study of segment anything model on zero-shot medical segmentation[J]. *Diagnostics*, 2023, 13(11): 1947.
- [16] Maxwell A E, Warner T A, Fang Fang. Implementation of machine-learning classification in remote sensing: an applied review[J]. *International Journal of Remote Sensing*, 2018, 39(9): 2784–2817.
- [17] Pádua L, Adão T, Hruška J, et al. Vineyard classification using machine learning techniques applied to RGB-UAV imagery[C]//Proceedings of 2020 IEEE International Geoscience and Remote Sensing Symposium. Waikoloa: IEEE, 2020.
- [18] Juel A, Groom G B, Svenning J C, et al. Spatial application of Random Forest models for fine-scale coastal vegetation classification using object based analysis of aerial orthophoto and DEM data[J]. *International Journal of Applied Earth Observation and Geoinformation*, 2015, 42: 106–114.
- [19] 周小成, 郑磊, 黄洪宇. 基于多特征优选的无人机可见光遥感林分类型分类 [J]. *林业科学*, 2021, 57(6): 24–36.
Zhou Xiaocheng, Zheng Lei, Huang Hongyu. Classification of forest stand based on multi-feature optimization of UAV visible light remote sensing[J]. *Scientia Silvae Sinicae*, 2021, 57(6): 24–36.
- [20] Yang Shuting, Gu Lingjia, Li Xiaofeng, et al. Crop classification method based on optimal feature selection and hybrid CNN-RF networks for multi-temporal remote sensing imagery[J]. *Remote Sensing*, 2020, 12(19): 3119.
- [21] Fu Bolin, Liu Man, He Hongchang, et al. Comparison of optimized object-based RF-DT algorithm and SegNet algorithm for classifying Karst wetland vegetation communities using ultra-high spatial resolution UAV data[J]. *International Journal of Applied Earth Observation and Geoinformation*, 2021, 104: 102553.
- [22] Han Kai, Wang Yunhe, Chen Hanting, et al. A survey on vision transformer[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2023, 45(1): 87–110.
- [23] Liaw A, Wiener M. Classification and regression by randomForest[J]. *R News*, 2002, 2(3): 18–22.
- [24] Breiman L. Random forests[J]. *Machine Learning*, 2001, 45(1): 5–32.
- [25] Rodriguez-Galiano V F, Ghimire B, Rogan J, et al. An assessment of the effectiveness of a random forest classifier for land-cover classification[J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2012, 67: 93–104.
- [26] Nguyen M H, de la Torre F. Optimal feature selection for support vector machines[J]. *Pattern Recognition*, 2010, 43(3): 584–591.
- [27] Morgan G R, Wang Cuizhen, Morris J T. RGB indices and canopy height modelling for mapping tidal marsh biomass from a small unmanned aerial system[J]. *Remote Sensing*, 2021, 13(17): 3406.
- [28] Congalton R G, Green K. Assessing the Accuracy of Remotely Sensed Data: Principles and Practices[M]. 3rd ed. Boca Raton: CRC Press, 2019.

Extraction of salt-marsh vegetation “fairy circles” from UAV images by the combination of SAM visual segmentation model and random forest machine learning algorithm

Zhou Ruotong¹, Tan Kai¹, Yang Jianru¹, Han Jiangtao¹, Zhang Weiguo¹

(1. State Key Laboratory of Estuarine and Coastal Research, East China Normal University, Shanghai 200241, China)

Abstract: The “fairy circle” represents a unique form of spatial self-organization found within coastal salt marsh ecosystems, profoundly influencing the productivity, stability, and resilience of these wetlands. Unmanned Aerial Vehicle (UAV) imagery plays a pivotal role in precisely pinpointing the “fairy circle” locations and deciphering their temporal and spatial development trends. However, identifying “fairy circle” pixels within two-dimensional images poses a considerable technical challenge due to the subtle differences in color and shape characteristics between these pixels and their surroundings. Therefore, intelligently and accurately identify “fairy circle” pixels from two-dimensional images and form individual “fairy circle” for the identified pixels were the current technical difficulties. This paper introduced an innovative approach to extract “fairy circle” from UAV images by integrating the SAM (Segment Anything Model) visual segmentation model with random forest machine learning. This novel method accomplished the recognition and extraction of individual “fairy circle” through a two-step process: segmentation followed by classification. Initially, we established Dice (Sørensen-Dice coefficient) and IOU (Intersection Over Union) evaluation metrics, and optimize SAM’s pre-trained model parameters, which produced segmentation mask devoid of attribute information by fully automated image segmentation. Subsequently, we aligned the segmentation mask with the original image, and utilized RGB (red, green, and blue) color channels and spatial coordinates to construct a feature index for the segmentation mask. These features underwent analysis and selection based on Out-of-Bag (OOB) error reduction and feature distribution patterns. Ultimately, the refined features were employed to train a random forest model, enabling the automatic identification and classification of “fairy circle” vegetation, common vegetation, and bare flat areas. The experimental results show that the average correct extraction rate of “fairy circle” is 96.1%, and the average wrong extraction rate is 9.5%, which provides methodological and technological support for the accurate depiction of the spatial and temporal pattern of “fairy circle” as well as the processing of coastal remote sensing images by UAVs.

Key words: salt marsh vegetation; fairy circle; segment anything model (SAM); unmanned Aerial Vehicle images; machine learning