

Random Forest for Seagrass and Seaweed Habitat Mapping in Suo Nada, the Seto Inland Sea

Student ID	47-186833
Name	Xiaoyu Qin
Supervisor	Prof. Jun Sasaki

1 Introduction

1.1 background

Seagrass and seaweed ecosystems act as primary producers in coastal waters, providing essential functions by producing and exporting organic carbon, regulating carbon dioxide, and storing them inside the water[1]. Although it is important to the environment, seagrass and seaweed are under human pressure. Taking precautions is urgent and important and mapping is the first and urgent step to assist management and conservation.

To map coastal ecosystems, satellite-based remote sensors have been successfully used for benthic mapping with high to moderate resolution. With advances in remote sensing technology, commercial sensors can provide fine resolution and frequent revisit time which enables prompt analysis after big events.

One of the newest commercial satellite imagery resources is PlanetScope. The strengths lie in its 3 m fine resolution, 1-day short revisit period and broad covering area[2]. However, although with huge potential to assist with seagrass and seaweed mapping using remote sensing technique, there hasn't been enough satisfactory investigation on this important topic using PlanetScope.

1.2 objective and significance

Only few studies tried to map seagrass habitat with PlanetScope imagery and the accuracy was not satisfactory[3]. However, by reviewing the methods used in those studies, improvements are expected to be achieved in two ways. Hence, this study aims to improve the accuracy of classification of seagrass and seaweed by PlanetScope and discuss how we can make the best use of this new resource.

2 materials and methodology

2.1 study site description

Suo Nada, as shown in fig 2.1 located in the western part of the Seto Inland sea, has abundant tidal flats, seagrass and seaweeds[4]. The study site is selected according to bathymetry and data availability.

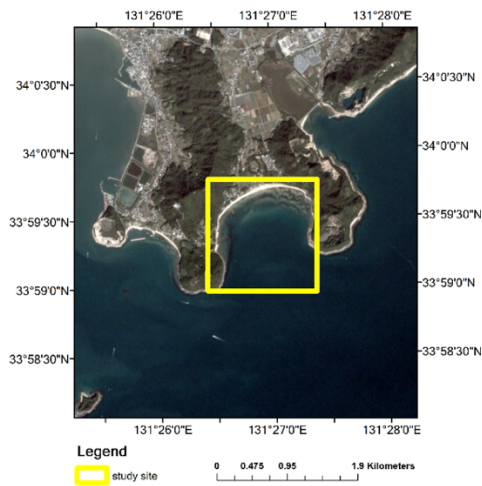


Fig. 2.1 study site description

2.2 data

In terms of satellite data, surface reflectance product of PlanetScope was used in this study. In total four images were used which were recorded in October 27 and 30, November 2 and 5, 2017. In terms of training data that is necessary for supervised leaning, the actual distribution obtained from a large-scale investigation conducted by the Ministry of Environment was used. Bathymetry data *Tsunami fault model (5)* *Topographical data* with 10 m mesh water

depth was used. It was obtained from an open access website, geo-spatial data center.

2.3 method

Random forest was adopted as the classification method. As an ensemble method, the essence of the method is to build multiple trees in randomly selected subspaces of the feature space and generalize their classification in complementary ways[5].

The workflow of this study is shown below in fig 2.2. Follow the flow, two different classifiers were constructed: RF-1 and RF-2. RF-1 took only one image recorded on October 31, 2017 as input while RF-2 took all four images described in section 2.2 as input.

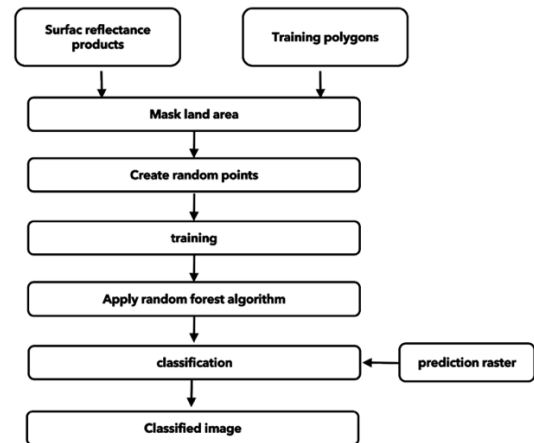


Fig. 2.2 workflow

As a result, three classes are predicted: seagrass and seaweed, tidal flats and water.

3 result

3.1 tuning

In terms of tuning, the best tree number and training size for RF-1 is 200 trees and 3250 input data. For RF-2, the numbers are 100 trees and 13000, respectively.

3.2 feature importance

In terms of feature importance, it was found that the most contributing features are distance to coastline, bathymetry and green, blue band. This shows that the newly added environmental features have been crucial and useful in classification.

3.3 prediction result evaluation

Kappa index and f1 score were shown in fig. 3.1 and 3.2. Horizontal axis cases 1 to 4 mean four different satellite images recorded on October 27 and 30, November 2 and 5, respectively.

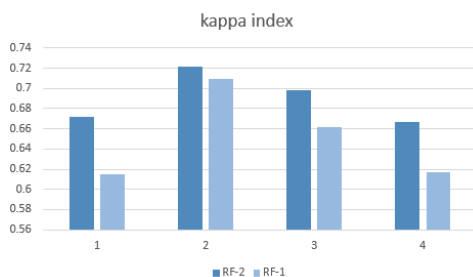


Fig. 3.1 kappa index of RF-1 and RF-2

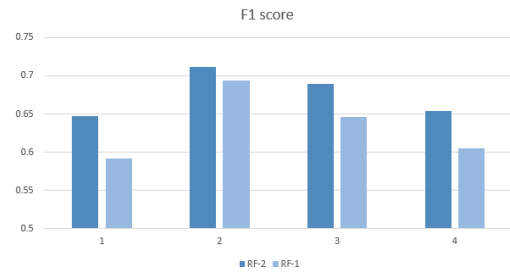


Fig. 3.2 f1 score index of RF-1 and RF-2

Both kappa and f1 score have similar tendency when RF-2 has higher performance than RF-1. This reveals that RF-2, trained with four images input has performed better compared with RF-1 that trained with only one imagery.

Another conclusion is that, case 2 has highest kappa and f1 score compared with other scenes. Possible reasons behind this phenomenon will be provided in discussion section.

Compared with previous study[3] conducted in Indonesia, the overall accuracy and kappa have been improved since both kappa and overall accuracy are higher, as can be observed from table 3.1

Table 3.1 comparison between current study and previous literature

	Current study	Previous study
Kappa (highest)	0.7219	0.54
Overall accuracy (highest)	0.874773	72.09

Discussion

4.1 reason for different performance on four cases

One apparent difference among the four cases is the tide level. The tide levels for four cases are 175, 111, 148, 297 cm, respectively. As can be observed in fig 4.1, with higher tide level, both OA and f1 score decreases accordingly.

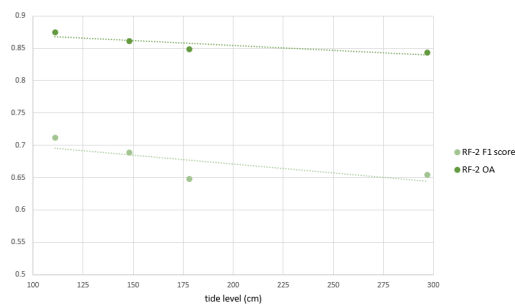


Fig. 4.1 correlation between tide level and mode performance

4.2 Strength and weak point of this study

As feature importance results show, newly added features have been crucial in classification which is considered as one of the strengths of this classifier. However, this can also become weak point since the classifier is too rely on distance to coastline feature.

The predicted map of RF-2 on case 2 shown in fig 4.2 reveals that, with smaller distance, the classifier tended to recognize the actual water

pixel(dark blue color) into seagrass and seaweed (light blue color), the false positive problem.

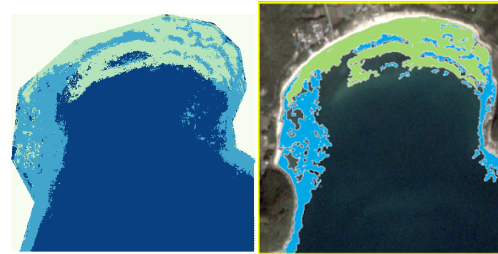


Fig 4.2 predicted map and actual distribution

Hence, new features that can balance the weight of distance to coastline should be considered in future work.

Reference

- [1] G. Nellemann, C., Corcoran, E., Duarte, C. M., Valdés, L., De Young, C., Fonseca, L., Grimsditch, *Blue Carbon. A Rapid Response Assessment*. 2009.
- [2] Planet, "Planet Imagery : Product Specification," Planet, no. January, p. 96, 2019, [Online]. Available: <https://www.planet.com/products/planet-imagery/>.
- [3] A. Ariasari, . Hartono, and P. Wicaksono, "Random forest classification and regression for seagrass mapping using PlanetScope image in Labuan Bajo, East Nusa Tenggara," no. December 2019, p. 77, 2019, doi: 10.1117/12.2541718.
- [4] M. of the Environment, "Survey on Distribution of Seagrass and Seaweed Beds and Tidal Flats in the Seto Inland Sea (Western Sea)," pp. 0–19, 2018.
- [5] Tin Kam Ho, "Random Decision Forests," *Proc. 3rd Int. Conf. Doc. Anal. Recognit.*, pp. 278–282, 1995, doi: 10.1109/ICDAR.1995.598994.