

# A SURVEY ON CORAL SPECIES CLASSIFICATION USING VARIOUS LEARNING ALGORITHMS

<sup>1</sup>Ms. MD Zaheda Parveen\*, <sup>2</sup>Dr. L.R. Sudha, <sup>3</sup>Dr. R. Santhoshkumar,

<sup>1</sup>Research Scholar, <sup>2,3</sup>Associate Professor,

Department of Computer Science and Engineering

<sup>1,2</sup>Annamalai University, Chidambaram, India

<sup>3</sup>St. Martin's Engineering College, Secunderabad, India

\*Corresponding Author

Email: parveenprofessional841@gmail.com

**Abstract**—Seagrass is an important component of the marine ecosystem and plays a vital role in preserving the water quality. The traditional approaches for sea grass identification are either manual or semi-automated, resulting in costlier, time consuming and tedious solutions. There has been an increasing interest in the automatic identification of seagrasses and this article provides a survey of automatic classification techniques that are based on machine learning, fuzzy synthetic evaluation model and maximum likelihood classifier along with their performance. The article classifies the existing approaches on the basis of image types (i.e. aerial, satellite, and underwater digital), outlines the current challenges and provides future research directions.

**Keywords**—Seagrass, automated detection, underwater imagery.

## I. INTRODUCTION

Seagrasses are flowering plants which are found in near shore environment of the continents [1]. They play an important role in providing food and shelter to other marine plants and animals which include tiny worms, shellfish, seastars and crustaceans. But in reality, seagrasses have been provided with less recognition for their importance to our world [2].

Western Australia is a rich habitate of a diverse species of seagrasses including *Halophila ovalis*, *Halodule uninervis*, *Halophila spinulosa*, *Halophila decipiens*, *Cymodocea serrulata*, *Serrulata*, *C. angustata*, *Syringodium isoetifolium* and *Thalassia hemprichii* [3]. Fig.1 gives a glimpse of seagrass coverage at Shark Bay in Western Australia. However, it has been shown in a number of studies that the abundance of the seagrass is declining worldwide due to storms, diseases, dredging, changes in water quality, pollutions, effects of development of seashores, overgrazing and sedimentation. To better understand seagrass' health, their growth and diversity in any area, efficient and automatic data analysis is necessary. Besides, standardizing remote sensing and tracking of seagrass species and their habitat along with the monitoring of vast seabed area is also important.

In the recent years, by the use of digital cameras along with the development of autonomous under water vehicles (AUV) and unmanned under water vehicles (UUV), there is a drastic exponential increase of availability of underwater imagery [4].

This availability motivated researchers to look closely to the issue and apply different techniques to detect and map seagrass. This article provides the survey of existing techniques applied to monitor seagrasses and their patterns, reproduction rate and ecology system.

The rest of the paper is organized as follows. Classification of existing approaches of detecting seagrass automatically are discussed in Section II. Associated challenges are highlighted in Section III. Prospects for future possible works have been mentioned in Section IV and finally, conclusion is drawn in Section V.



Fig. 1. Seagrass coverage of Shark Bay, WA (around 4500 square KM) [4]

## II. AUTOMATED SEAGRASS IDENTIFICATION TECHNIQUES

We classify the current literature into three broad categories as follows:

1. Satellite Image Based techniques
2. Aerial Image Based Techniques
3. Underwater Digital Image Based Techniques

Table 1 provides the summary of the each category, techniques used for the identification of the seagrass along with, their performances, dataset characteristics, image types used, features selected and other additional parameters. In the following section, we will discuss each of this category in detail.

### A. Satellite Image Based techniques

This section further classify the existing techniques applied on the basis of different types of satellite images used to identify seagrass as follows:

1) *Generalised Linear Model*: Saunders et al. [5] developed a seagrass distribution model using benthic radiance and wave height as predictors on five benthic images. Digital terrain model, water clarity, significant wave light and Benthic substrates were included in data set. A Generalized Linear model, also called as a habitat distribution model, was fitted to check the presence vs. absence of the seagrass. In this approach, seagrass was counted as present if the probability was greater than a threshold level (0.16). The presence of seagrass turned out to be positively proportional to high light

correct atmospheric correction. Near Infrared was applied to reduce light effects from images. Preprocessed images were classified using Maximum Likelihood Classifier. Data from 60 transects was used as training/test purposes to classify SAV. The percentage of cover retrieved from this step can be classified into five categories as shown in Table 2. To evaluate results, overall accuracy and Kappa, were performed [6].

Fuzzy synthetic evaluation model was used to get the abundance of the seagrass. Three biometric factors: SAV cover, leaf area index (LAI) and biomass were used as features

TABLE 1. DIFFERENT TECHNIQUES FOR SEAGRASS IDENTIFICATION.

| Type of images            | Author (Year)                       | Dataset  | Technique   | Classification algorithm  |
|---------------------------|-------------------------------------|--|---|---|
| Satellite Images          | Saunders et al. (2013) [5]          | Landsat Thematic mapper sensor image   | Seagrass distribution model   | Generalized linear model  |
|                           | Pu and Bell (2013) [6]              | Three images from 3 satellite sensors and 60 transects                               | Fuzzy Synthetic Evaluation Model  | Fuzzy Synthetic evaluation technique, Maximum Likelihood Classifier                       |
|                           | Baumstark et al. (2016) [8]         | One multispectral image  | Object based image analysis   | Unsupervised classification, logistic regression model                                    |
|                           | Ressom et al. (2016) [14]           | Spectral reflectance values using spectral sensors(Not a digital photographic Image) | Spectral reflectance analysis   | Neural network classifier along with PCA  |
| Aerial photographs        | Uhrin and Townsend (2015) [13] [11] | Total 242 individual images 1524.39x1524.39  | Unmixing of spectral components of pixels                                 | Comparison of Linear Spectral Unmixing (LSU) classifier & manually digitalized classifier |
|                           | Agostini et al. (2003) [12]         | Two aerial photographs and field data collected using transects                      | MULTI-SCOPE software  | Hypercube classification  |
| Underwater digital images | Massot-Campos et al.(2013) [9]      | Digital images taken underwater  | Laws' Texture Energy Measurements and the Grey Level Co-occurrence matrix | Logistic Model Tree (LMT)   |

penetration and inversely proportional to great wave height. The performance came out to be 83% for predicting seagrass.

2) *Fuzzy Synthetic Evaluation and Maximum Likelihood Classifier*: This method was used to evaluate the abundance of seagrass in Pinellas County, FL, USA. There are three seagrass species which includes *Syringodium filiforme*, *Thalassia testudinum* and *Halodule wrightii*. But there is abundance of the rhizophytic algae in that area which is 80% in few locations mixed with seagrass. So to include this category as well it is named submerged aquatic vegetation (SAV). This study had one more step after image processing which is called image optimization algorithms for seagrass classification. Atmospheric and sun light corrections were done to the sensor of the satellites [Landsat 5 Thematic Mapper (TM), Earth Observing-1 (EO-1) Advanced Land Imager (ALI) and Hyperion (HYP)] as image preprocessing rules. Three scenes from each satellite were acquired.

In this work, two operational image classification algorithms were applied called VRadCor and SRSSH. First one was for destripe and second one was for denoise. These algorithms helped improving quality of the images and did not affect the spectral components present in the images [18]. After that Fast Line-of-sight Atmospheric Analysis of Spectral Hypercube (FLAASH) was applied to optimized images to

for monitoring seagrass health. Values of all three features were calculated from 60 transects and then regression model was developed. These values were used in multiple regression models for getting biometric of each pixel satellite images. Then those retrieved features were used to make membership maps for all three biometrics. Five membership functions were created for all three biometrics and then seagrasses were mapped by using equations of synthetic evaluation model [6]. The performance of the three sensors was: HYP (OA=87%, K=0.83), ALI (OV=82%, K=0.77) and TM (OA=79%, K=0.77).

TABLE 2: CLASSIFYING SAV INTO FIVE CLASSES [6]

| Class | Description of SAV Cover | Training samples (pixels) | Test samples (pixels) |
|-------|--------------------------|---------------------------|-----------------------|
| 5     | >75                      | 747                       | 616                   |
| 4     | 51-75                    | 335                       | 302                   |
| 3     | 26-50                    | 471                       | 585                   |
| 2     | 1-25                     | 478                       | 447                   |
| 1     | <1                       | 661                       | 581                   |

3) *Unsupervised Machine Learning and Logistic Regression Model*: Baumstark et al. [8] proposed object based classification method for the identification of the seagrass. This

method used unsupervised classification analysis and logistic regression model for classification which is a statistical model. Results were compared with traditional photo-interpretation method. The project was carried out in Florida. Worldview-2 satellite images were used. Worldview-2 satellite image takes 8 different multispectral bands at 2m pixel resolution. The image contains spectral values for colors. To get the same spectral features for seagrass and hard bottom, image has to be preprocessed to increase accuracy. For that purpose, light attenuation was applied and correction factors were adjusted for every spectral band in the image. The extracted information of spectral bands is used in classification to from a benthic image to give the view of seagrass along the shore.

Preprocessing also includes noise reduction. There were three main feature classes which were then subdivided into six classes based on their percentage cover and their mixture with each other. After image processing and deciding on feature classes, object based image analysis classification process was performed which consists of three steps: 1) image segmentation 2) classification of pixels into three main class types, and 3) calculation of the percentage of the cover by any class from main three classes to categorize them into 6 subdivided classes. In the first step, during image segmentation, the size was kept 0.5 acre MMU. Segmentation also checks pixel value, objects sharing the same properties such as shape, size and orientation. Unsupervised classification was applied for classification of the three main types based on the spectral values of the pixels. Logistic regression model is used to differentiate between the same spectral cover types. Regression logistic model used distance and water depth as independent parameters for presence of the seagrass (dependent variable). Unsupervised classification and the regression model used together to classify spectral values. After classifying, percentage cover of the spectral values in segmented image is calculated. Segments with less than 10% seagrass were classified as sand, and with percentage of 14% were classified as seagrass medium. For validation, random sites were chosen to check whether results satisfied the method. Overall accuracy of 71% was achieved with this method by using error matrix by comparing user's and producer's accuracies.

4) *Neural Networks*: This method is based on the spectral reflectance of seagrass and other objects. All objects on the earth absorb and reflect electromagnetic energy in different days because of their different physical and nonphysical features. Those features are color, structure and texture. Sensors measure the reflected energy and if all energy is reflected it is called 100% reflectance and all is absorbed and none is reflected then it is called reflectance is 0%. Electromagnetic spectrum can be developed using reflectance values for all the objects of the world and can be compared for detection [28].

A technique based on spectral reflectance was developed by Resson et al. (2016) [5] for the identification of the seagrass to monitor the health of the seagrass. A neural network classifier

was used for detecting seagrass called *Zostera Capricorni* and distinguish it from other three species and then neural network architecture was used to monitor its health by estimating photosynthetic efficiency. The advantage of neural network is that they can be adjusted according to input, can tolerate noise factors and work better with nonlinear relationships very effectively. But the spectral reflectance data is quiet high dimensional. Therefore, correction analysis was performed to choose inputs which affect output data only. Spectral component analysis was performed on input data to make input vector components uncorrelated. Data used for this project contained spectral reflectance values of the three seagrass species *Zostera capricorni*, *Posidonia australis* and *Halophila ovalis*. Data was collected between 1999 and 2000 with values ranging 430nm to 900nm. Dataset contained 139 total samples. Those samples were divided into three categories: training set containing 69 samples, validation set containing 36 samples and test set containing 34 samples. After collection, data normalization was done. After performing spectral component analysis / principle component analysis, data was given to neural network. Selected neural network was a multi-layer feed forward, back propagation in nature. There were five inputs, and two hidden layers, first layer having seven neurons and second layer having two neurons and three outputs. Data fed to neural network contained both preprocessed data and raw data and the spectral values of three species had strong differences in wavelengths ranging from 530-580nm which is green color spectral value. Confusion matrix (see Table 3) was used to check classification results against prediction results. [14].

TABLE I. TABLE 3: CONFUSION MATRIX FOR CORRECTED SPECTRAL REFLECTANCE DATA [14]

| Class Number (Number of samples) | Class 1 (11) | Class 2 (9) | Class 3 (14) | Prediction error |
|----------------------------------|--------------|-------------|--------------|------------------|
| Class 1(11)                      | 11           | 0           | 0            | 0%               |
| Class 2(9)                       | 0            | 9           | 0            | 0%               |
| Class 3(14)                      | 0            | 0           | 14           | 0%               |
| Classification error             | 0%           | 0%          | 0%           |                  |

#### B. Aerial Image Based Techniques

1) *MULTI-SCOPE Software*: A software called MULTI-SCOPE from Matra Cap Systems was used to identify seagrass from aerial images. Image processing of this technique was applied to two aerial photographs which were taken in 1997. Photographs were digitized by IMAGE-IN scan and Paint software. A total of 317 points were allowed per inch and photographs contained 16.8 million colors. Each point represented a vector which had density of the spectral bands containing base colors (red, blue, green). A mask was applied to the land in black and contrast of the image was increased [26]. Seagrass contains green and blue colors, so principle component analysis was applied to make the patterns clear enough. Reference polygons were digitized based on ground truth data. Hypercube classification was applied to colored composition. After that, all the colors in polygons

were applied to all the images. Main types were followings: 1) sand 2) mud 3) pebbles 4) mixed meadows of seagrass, and 5) the litter (dead leaves). Field data was collected by divers. Seagrasses were identified and specified using transect method. This data was collected from different points of the area under observation. By combining image processing and ground truth data, 76% reliability was found for the experiment.

2) *Linear Spectral Unmixing (LSU)*: This technique was used by Uhrin et al. [6] for determining the abundance of an object in multi spectral images. Multi-spectral images have mixed pixels and they reduce but not eliminate errors in classification [27]. Authors used LSU on seagrass data collected from shallow waters of Albemarle-Pamlico Sound Estuary System in North Carolina. Images were collected in three different bands. The first band was blue, ranging from 410 to 490nm, the second was green, ranging from 510 to 590 and the third band was red, ranging from 610 to 690nm. Seagrass images were digitized at minimum mapping unit of 15m.

Before applying LSU, image preprocessing was performed which included a forward Minimum Noise Fraction transformation to reduce spectral noise of the images for better output. After that, end members were identified from all images. End members are main components in this method, because LSU considers each end member as known. LSU is done on clipped images to retrieve the maps of seagrass which gets information about seagrass presence from image pixel. LSU takes an image or proportion of image with mixed pixel, and then divides that spectrum to individual spectra of each component or end member present in that mixed pixel proportion of image.

The performance of the LSU was measured using two criteria: the Kappa (K) statistic and Area Under Curve (AUC). Kappa was evaluated for each error matrix, where K is the value which tells how well classification has performed according to reference data. As a rule of thumb, if  $K > 0.80$  the relationship between reference data and classification is considered strong. The value of K came out to be 0.72 to 0.98 which indicates positive strong results [13].

### C. Underwater Digital Image Based Techniques

Very limited automatic approaches were found using underwater digital images [30]. Massot-Campos et al. [9] quantified the presence of *Posidonia oceanica*, a variant of seagrass, on analogic RGB data collected at Palma Bay. They used Logistic Model Tree (LMT) as a classifier. They also used Law's energy measurements and grey level co-occurrence matrix to identify the differences in texture.

### III. CURRENT CHALLENGES

The first challenge for seagrass identification is lack of ground truth dataset. Many of the proposed works did not report

enough ground truth data. Besides, due to the similarities of the spectral components of seagrass varieties, it is difficult to distinguish different classes. Moreover, the large and high quality digital data processing capacity of the existing is also concerning.

While considering the image source, satellite images have few limitations: 1) Narrow coverage of spectral bands in hyper spectral remote sensing, 2) Limited temporal resolution, 3) high photographic distortion, 4) Low radiometric resolution, 5) Cloud contamination, 6) Interpretation difficulty in deep and shallow water, 7) Errors due to converting analogue airborne photos to digital images and 8) high cost when high spatial and spectral resolution is required. If image preprocessing for removing effects of water column is not performed in satellite images, then accuracy is affected by 17%. With other techniques such as principle component analysis, normalized difference vegetation index and leaf area index, the accuracy could be affected in a range of 22% [29].

### IV. FUTURE WORK

Classification of seagrass was mostly carried out using spectral components. However, they have different spatial patterns which can be analyzed using texture pattern analysis such as spatial frequency, a measure of change in pixel brightness value per unit distance. Neural networks can be used with larger dataset to include more spatial scales to improve performance. So far, the result of neural network was much better than the rest of the techniques.

Very limited work has been done using digital images to detect seagrass meadows. The cases where they were used, they were interpreted using different software to get the contrast of the image and make decisions. Due to the advancements in camera technology to capture digital underwater images and computation power such as Graphics Processing Units (GPU), Neural Networks Architectures specially Deep Neural Networks (DNN) are a better candidate solution for automatic seagrass detection, classification and mapping problem. Therefore, in near future, we will use DNN for automatic sea grass detection and classification.

### V. CONCLUSION

This paper presents recent approaches that have been used to automatically estimate the abundance of seagrass. The purpose of the survey is to identify what type of image is being used and type of algorithms are being used for classification purposes and find their performance. It is found that, most of the methods are based on the satellite image which gives image of the whole area where detection needs to be applied. Manual detection takes months, so scientists have developed automated techniques. In satellite imagery, there are some parameters which affect the accuracy of the algorithm. Few of them are atmospheric such as distance from water, cloud contamination, water quality are not very clear from such a distance. Digital images can be taken underwater and the image preprocessing can be applied. Digital image can improve the performance as their resolution and clarity are much better than satellite

images. Using digital image with deep artificial neural networks would be a good choice for automatic seagrass detection as the performance of neural networks demonstrated very high accuracy in many object detection applications.

#### REFERENCES

1. R. Coles, L. McKenzie, M. Rashed, J. Mellors, H. Taylor, K. Dew, S. McKenna, T. Sankey, A. Carter and A. Grech, "Status and trend of seagrass habitats in the Great Barrier Reef world heritage area," Report to the Marine and Tropical Sciences Research Facility, Reef and Rainforest Research Centre on behalf of the Australian Government's Marine and Tropical Sciences Research Facility, Cairns, Australia, 2007.
2. F. T. Shorta, B. Polidoro, S. R. Livingstone, K. E. Carpenter, S. Bandeira, J. S. Bujang, H. P. Calumpang and T. J. Carruthers, "Extinction risk assessment of the world's seagrass species," *Biological Conservation*, vol. 144, pp. 1961-1971, 2011.
3. D. I. Walker and R. I. T. Prince, "Distribution and biography of seagrass species on the northwest coast of Australia," *Aquatic Botany*, vol. 29(1), pp.19-32, 1987.
4. M. Vanderklift, D. Bearham, M. Haywood, H. Lozano-Montes, R. McCallum, J. McLaughlin, K. McMahon, N. Mortimer and P. Lavery, "Natural dynamics: understanding natural dynamics of seagrasses in north-western Australia," WAMSI Dredging Science Node Report Perth, Oct 2016.
5. M. I. SAUNDERS et al., "Coastal retreat and improved water quality mitigate losses of seagrass from sea level rise," *global change Biology*, vol. 19, p. 2569–2583, 2013.
6. R. Pu and S. Bell, "A protocol for improving mapping and assessing of seagrass abundance along the West Central Coast of Florida using Landsat TM and EO-1 ALI/Hyperion image," *ISPRS Journal of Photogrammetry and Remote Sensing*, pp. 116-129, Jul 2013.
7. C.-F. Chen, V.-K. Lau, N.-B. Chang, N.-T. Son, P.-H.-S. Tong and S.-H. Chiang, "Multi-temporal change detection of seagrass beds using integrated Landsat TM/ETM+/OLI imageries in Cam Ranh Bay, Vietnam," *Ecological Informatics*, pp. 43-54, Jul 2016.
8. R. Baumstark, R. Duffey and R. Pu, "Mapping seagrass and colonized hardbottom in Spring Coast, Florida using Worldview 2 satellite imagery," *Estuarine, Coastal and Shelf Science*, pp. 83-92, 2016.
9. M. Massot-Campos, G. Oliver-Codina, L. Ruano-Amengual and M. Miró-Juliá, "Texture Analysis of seabed Images: quantifying the presence of *Posidonia oceanica* at Palma Bay," *OCEANS-Bergen, MTS/IEEE*, pp. 1-6, 2013.
10. J. I. McDonald, G. T. Coupland and G. A. Kendrick, "Underwater video as a monitoring tool to detect change in seagrass cover," *Journal of Environmental Management*, vol. 80 (2), pp. 148-155, Jul 2006.
11. S. Rende, A. Irving, T. Bacci, L. Parlagreco, F. Bruno, F. D. Filippo, M. Montefalcone, M. Penna, B. Trabucco, R. D. Mento and A. Cicero, "Advances in micro-cartography: A two-dimensional photo mosaicing technique for seagrass monitoring," *Estuarine, Coastal and Shelf Science*, vol. 167, pp. 475-486, 2015.
12. S. Agostini, A. Capiomont, B. Marchand and G. Pergent, "Distribution and estimation of basal area coverage of subtidal seagrass meadows in a Mediterranean coastal lagoon," *Estuarine, Coastal and Shelf Science*, vol. 56, pp. 1021–1028, 2003.
13. A. V. Uhrin and P. A. Townsend, "Improved seagrass mapping using linear spectral unmixing of aerial photographs," *Estuarine, Coastal and Shelf Science*, pp. 11-22, 2016.
14. H. Resson, S. K. Fyfe, P. Natarajan and S. Srirangam, "Monitoring Seagrass Health using Neural Networks," *Proc. of the Int. Joint Conf. on Neural Networks*, 2003.
15. D. M. Alongi, "Present state and future of the world's mangrove forests," *Environmental Conservation*, vol. 29, no. 3, pp. 331-349, Sept 2002.
16. F. T. Short and H. A. Neckles, "The effects of global climate change on seagrasses," *Aquatic Botany*, vol. 63, no. 3-4, p. 169–196, April 1999.
17. C. M. Duarte, "The future of seagrass meadows," *Environmental Conservation*, vol. 29 (2), pp. 192-206, 2002.
18. Y. Zhao, R. Pu and S. S. Bell, "Hyperion image optimization in coastal waters," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 51 (2), pp. 1025-1036, Feb 2013.
19. M. Teng, M. Mehrubeoglu, S. King, K. Cammarata and J. Simons, "Investigation of epifauna coverage on seagrass blades using spatial and spectral analysis of hyperspectral images." In: 5th Workshop on Hyperspectral Image and Signal Processing: pp. 25-28, 2013.
20. E. Oguslu, J. Li, S. Erkanli, V. Hill, R. Zimmerman and W. Bissett, "Detection of seagrass scars using sparse coding and morphological filter," *Remote Sensing of the Ocean, Sea Ice, Coastal Waters, and Large Water Regions 2014*.
21. P. Fairweather, "Statistical power & design requirements for environmental monitoring," *Australian Journal of Marine and Freshwater Research*, vol. 42, pp. 555-567, 1991.
22. A. Underwood, "Experiments in ecology," in Cambridge University Press, Cambridge, 1997.
23. E. Graham, M. Dethier, S. Cohen and L. Tear, "Visual versus random-point percent cover estimations: objective is not always better," *Marine Ecology Progress Series*, vol. 96, p. 93–100, 1993.
24. A. Wojcicka and Z. Wrobel, "The panoramic visualization of metallic materials in macro- and microstructure of surface analysis using Microsoft Image Composite Editor (ICE)," pp. 358-368, 2012.
25. C. N. Rooper and M. Zimmermann, "A bottom-up methodology for integrating underwater video and acoustic mapping for seafloor substrate classification," *Continental Shelf Research*, vol. 27 (7), pp. 947–957, Apr 2007.
26. R. Hummel, "Image enhancement by histogram transformation, computer graphics and image processing," vol. 6, pp. 184-195, 1997.
27. C. E. Woodcock, "The factor of scale in remote sensing," *Remote Sensing of Environment*, vol. 21, no. 3, pp. 311-332, April 1987.
28. "What is spectral signature in remote sensing?," 2013. <https://articles.extension.org/pages/40071/what-is-a-spectral-signature-in-remote-sensing> (last accessed: Nov 3, 2017)
29. P. J. Mumby, C. D. Clark, E. P. Green and A. J. Edwards, "Benefits of water column correction and contextual editing for mapping coral reefs," *International Journal of Remote Sensing*, vol. 19 (1), 1998.
30. H. Lu, Y. Li, Y. Zhang, M. Chen, S. Serikawa and H. Kim, "Underwater optical image processing: a comprehensive review," *Proc. CVPR*, 2017.