

Semantic segmentation of major macroalgae in coastal environments using high-resolution ground imagery and deep learning

Jesús Balado , Celia Olabarria , Joaquín Martínez-Sánchez , José R. Rodríguez-Pérez & Arias Pedro

To cite this article: Jesús Balado , Celia Olabarria , Joaquín Martínez-Sánchez , José R. Rodríguez-Pérez & Arias Pedro (2021) Semantic segmentation of major macroalgae in coastal environments using high-resolution ground imagery and deep learning, International Journal of Remote Sensing, 42:5, 1785-1800, DOI: [10.1080/01431161.2020.1842543](https://doi.org/10.1080/01431161.2020.1842543)

To link to this article: <https://doi.org/10.1080/01431161.2020.1842543>



Published online: 20 Dec 2020.



Submit your article to this journal [↗](#)






View related articles [↗](#)



View Crossmark data [↗](#)



Semantic segmentation of major macroalgae in coastal environments using high-resolution ground imagery and deep learning

Jesús Balado ^a, Celia Olabarria ^{b,c}, Joaquín Martínez-Sánchez ^a, José R. Rodríguez-Pérez^d and Arias Pedro^a

^aCINTECX, Universidade de Vigo, GeoTECH Group, Vigo, 36310, Spain; ^bFacultade de Ciencias do Mar, Universidade de Vigo, Vigo, 36310, Spain; ^cCIM, Centro de Investigación Mariña, Universidade de Vigo, Vigo, 36331, Spain; ^dUniversidad de León, GEOINCA Group, Ponferrada, 24071, Spain

ABSTRACT

Macroalgae are a fundamental component of coastal ecosystems and play a key role in shaping community structure and functioning. Macroalgae are currently threatened by diverse stressors, particularly climate change and invasive species, but they do not all respond in the same way to the stressors. Effective methods of collecting qualitative and quantitative information are essential to enable better, more efficient management of macroalgae. Acquisition of high-resolution images, in which macroalgae can be distinguished on the basis of their texture and colour, and the automated processing of these images are thus essential. Although ground images are useful, labelling is tedious. This study focuses on the semantic segmentation of five macroalgal species in high-resolution ground images taken in 0.5×0.5 m quadrats placed along an intertidal rocky shore at low tide. The target species, *Bifurcaria bifurcata*, *Cystoseira tamariscifolia*, *Sargassum muticum*, *Sacchoriza polyschides* and *Codium* spp., which predominate on intertidal shores, belong to different morpho-functional groups. An explanation of how to convert vector-labelled data to raster-labelled data for adaptation to Convolutional Neural Network (CNN) input is provided. Three CNNs (MobileNetV2, Resnet18, Xception) were compared, and ResNet18 yielded the highest accuracy (91.9%). The macroalgae were correctly segmented, and the main confusion occurred at the borders between different macroalgal species, a problem derived from labelling errors. In addition, the interior and exterior of the quadrats were correctly delimited by the CNNs. The results were obtained from only one hundred labelled images and the method can be performed on personal computers, without the need to use external servers. The proposed method helps automation of the labelling process.

ARTICLE HISTORY

Received 6 July 2020

Accepted 8 October 2020

1. Introduction

Macroalgae are important primary producers on subtidal and intertidal rocky shores worldwide (Jenkins et al. 2008) and make a substantial contribution to carbon

sequestration, nutrient cycling and global oxygen production (Macreadie et al. 2017; Bañolas et al. 2020). As ecosystem engineers, macroalgae modify habitat conditions, facilitating the existence and survival of other intertidal species, thus strongly influencing the structure and functioning of coastal ecosystems (Purvaja et al. 2018). Intertidal species of macroalgae are vulnerable to stressors, including climate change, habitat loss, eutrophication, overfishing, pollution and the introduction of non-native species (Hawkins et al. 2009; Griffiths, Connolly, and Brown 2020). Shifts in the distributional range of diverse intertidal macroalgae due to increased air and Sea Surface Temperatures (SSTs) on the Atlantic shores of the Iberian Peninsula have been documented (Lamela-Silvarrey et al. 2012; Duarte et al. 2013; Lima et al. 2007).

Obtaining information about the distribution and abundance of macroalgae is therefore important for monitoring, managing and understanding coastal ecosystems, particularly in the context of global change, in which multiple stressors act together (Floor, van Koppen, and van Tatenhove 2018). Video and photographic monitoring have proven valuable ground-based and remote-sensing techniques for evaluating the cover and distribution of coastal organisms with high spatial and temporal resolution. This type of monitoring has been conducted using satellites (Sagawa et al. 2012; Wang et al. 2018; Topouzelis et al. 2016; Wilson, Skinner, and Lotze 2019; Li et al. 2012), Unmanned Aerial Vehicles (UAVs) (Tamondong et al. 2018; Wang et al. 2019; Ventura et al. 2018; Duffy et al. 2018; Taddia et al. 2019) and underwater drones (Moniruzzaman et al. 2019; Rahnemoonfar and Dobbs 2019; Kellaris et al. 2019; Martin-Abadal et al. 2018).

Image processing techniques in Red Green Blue (RGB) photographic traditionally used to detect and classify objects, including macroalgae are being displaced by machine learning and deep learning methods, which provide more accurate results (O'Byrne et al. 2018). Classic image processing techniques such as Local Binary Patterns (LBP) lack the ability to learn complex features, and the results are thus less accurate than those obtained with Convolutional Neural Networks (CNNs). The accuracy of segmentation of single seagrass species is only 85.0% with LBP techniques, but reaches 93.4% with CNN (Wang et al. 2018; Reus et al. 2018). In order to improve detection and classification, several authors continue to employ image processing techniques using data acquired at various wavelengths in the spectrum of non-visible light (Xing et al. 2019; Qi et al. 2020), but multi-spectral information is not always available. Techniques based on deep learning are limited by the large numbers of samples needed to train the classifier. To differentiate species of macroalgae, samples must be acquired using traditional approaches based on field data, such as diving or intertidal sampling with in situ quadrats or line transects, which provide high accuracy and resolution, but which are time consuming and limited to small areas (Casal et al. 2013).

The aim of this study was to automate the process of labelling high-resolution images to differentiate five macroalgae: *Bifurcaria bifurcata* Linnaeus, *Cystoseira tamariscifolia* (Hudson) Papenfuss, *Sargassum muticum* (Yendo) Fensholt, *Sacchoriza polyschides* (Lightfoot) Batters, and *Codium* spp. The process was automated using semantic segmentation and CNNs. We believe that this is the first study addressing semantic segmentation of five macroalgal species simultaneously in high-resolution ground photographic images through deep learning techniques. This paper reports a new method of converting labels (from polygons to raster images) and compares the results obtained with three different CNNs (MobileNetV2, Resnet18 and Xception). The research reported is part of the

ALGANAT2000 project, which aimed to monitor the spatio-temporal distribution of macroalgae in an intertidal coastal area within a marine protected area in Galicia (NW Spain) during 2019.

2. Material and methods

2.1. Study area

The study was conducted in the Atlantic Islands National Park (Galicia, NW Spain), a terrestrial and marine reserve formed by four main archipelagos. The exposed intertidal area of Bufardo on the Illa de Monteagudo (area surrounding coordinates 42.23551° N, 8.89956° W) belonging to the Illas Cíes archipelago was selected as the sampling location (Figure 1). The location is a gently sloping rocky platform where the upper intertidal zone is dominated by *Pelvetia canaliculata* Decaisne and Thuret, and the mid and low intertidal zones are dominated by conspicuous red, green and brown macroalgae, such as *Asparagopsis armata* Harvey, *B. bifurcata*, *C. tamariscifolia*, *S. polyschides* and *Codium* spp.

The images of the five species (Figure 2) were acquired from 0.5 × 0.5 m quadrats placed on the low intertidal shore during low spring tides in July, August and September 2019. The area is emerged, and thus exposed to the air, for about 3 to 3.5 hours during low spring tides. Depending on the altitude at which the macroalgae occur on the shore, they experience different conditions of solar radiation and

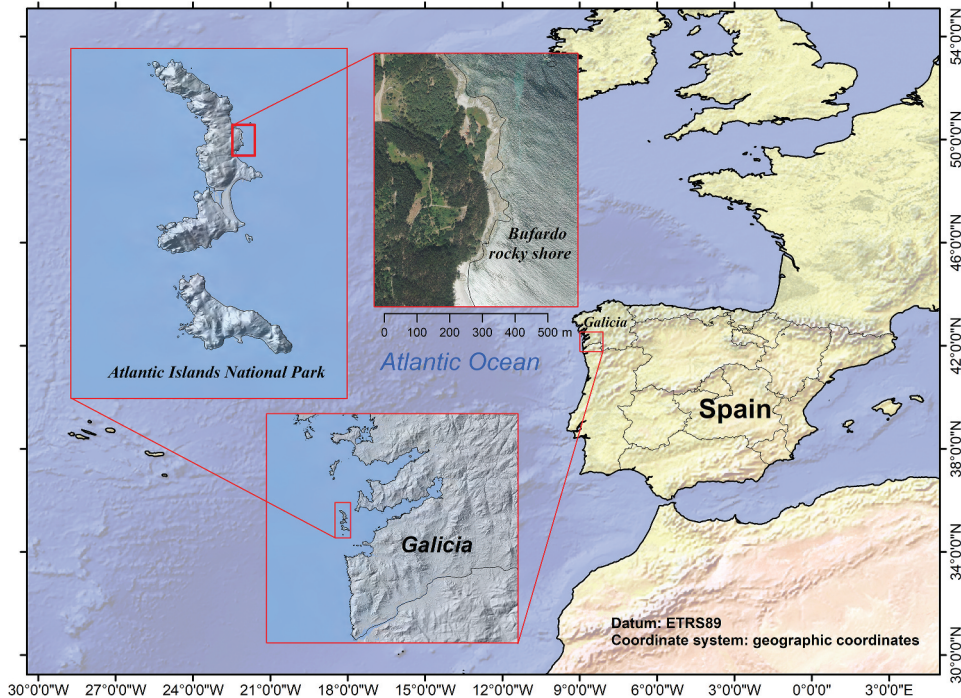


Figure 1. Location of the case study.

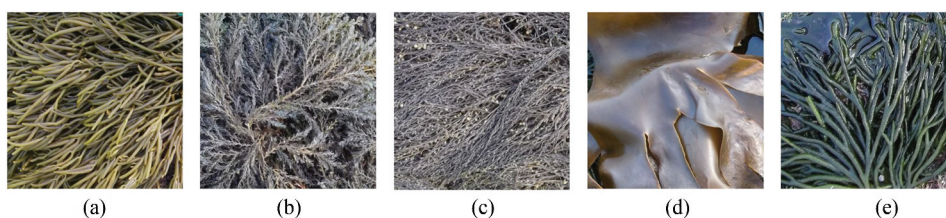


Figure 2. Target macroalgae used in the semantic segmentation: (a) *B. bifurcata*, (b) *C. tamariscifolia*, (c) *S. muticum*, (d) *S. polyschides*, (e) *Codium* spp.



Figure 3. Quadrat used for image acquisition.

desiccation. Thus, macroalgae living on the upper shore are drier and absorb more heat than macroalgae inhabiting the lower shore.

The images were acquired with a Fujifilm FinePix JV200 camera mounted on a tripod, with a top view perspective 0.7 m above the ground (see Figure 3). Because of the shape of the tripod, the base was also captured in each image. The square base delimited the labelling area, and the area outside of the base was labelled as the 'out' class.

2.2. Methods

Labelled data (in the form of manually digitized georeferenced vector polygons) and high-resolution images of macroalgae were used for semantic segmentation. Vector labelled data were adapted to raster labelled data following CNN standards. The choice of CNN was justified by the higher success rate of the method compared to other traditional methods based on texture analysis (O'Byrne et al. 2018), Histogram of Oriented Gradients (HOG) or LBP (Reus et al. 2018). In addition, pre-trained CNNs also enable more efficient feature extraction, with a faster design process than traditional techniques (assembly of successive masks with successive tests). The analysis was approached from the perspective of data analysis, and the system improved as new CNN architectures became available, without the entire work process having to be redesigned. The workflow of the method is represented in Figure 4.

2.2.1. Label adaptation for CNN

Conventional labelling of vector objects with a geographic information system (GIS) is not suitable for CNN-based semantic segmentation. The conventional method consists of creating a vector file layer with polygons labelled with the class. These polygons are delimited, by an expert, on the background image collected in the field. Apart from their topological attributes, the only condition for labelling is that the polygons must contain only one class of pixels. The expert selects which pixels to polygonise (for training the subsequent class). The expert classification is considered ground truth and consists of a labelled vector layer.

CNN for semantic segmentation only allows images (raster data) as input for both ground truth and labelled data. In addition to the data type, the labelling content is distinct and must obey the following rules:

- All pixels in the image must be labelled. Unlabelled pixels are assigned as 'others'. The 'others' class may include macroalgae that are not of interest for the study, e.g. sand, rocks and unidentified objects.
- All pixels of the objects must be labelled in their corresponding class. The 'others' class must not include pixels representing objects belonging to any class.

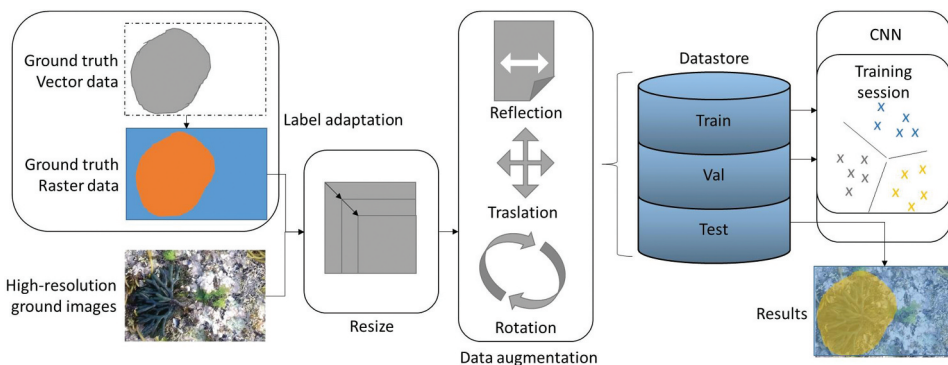


Figure 4. Workflow.

In order to fulfil these requirements for CNN training, each image was re-labelled accordingly. The re-labelling procedure depended on the following scenarios that can occur in each image-labelled data:

- The expert only polygonised the largest or most relevant macroalgae. In this option, the other pixels must be manually analysed and assigned to the corresponding class (whether macroalgae or 'others').
- The expert polygonised both large and small objects. In this case, only the data corresponding to ground and no relevant species need to be labelled 'others'. In this option, the process could be performed automatically by rasterizing the polygons.

Considering that the option used for each image was not known, the re-labelling process was performed manually. In addition, due to the image acquisition method used, macroalgae were labelled exclusively in a Region of Interest (ROI) in the images. In this case, the ROI in each image was the area enclosed by the quadrat, and the area outside of the quadrat was labelled 'out', regardless of whether it included macroalgae, acquired data or empty pixels caused by image rotation in the metadata orientation. In the example, from the picture acquired ([Figure 5\(a\)](#)), only four polygons corresponding to three different classes were labelled when the expert polygonised the largest relevant macroalgae ([Figure 5\(b\)](#)). In this case the polygons did not cover all pixels corresponding to each macroalga. A label was then assigned to each pixel ([Figure 5\(c\)](#)). The contours of the macroalgae were more detailed than those of the respective polygons.

2.2.2. Semantic segmentation

Semantic segmentation and object detection are classification methods that can be applied to image segmentation and labelling (Ruiz-Santaquiteria et al. 2020). Although both methods are based on deep learning, semantic segmentation aims to assign classes to each pixel of the image while the detector frames the detected objects in a bounding box. This bounding box is defined by a fixed number of vertices that frequently cover several pixels that do not correspond to the class detected. Such misclassification is usual in complex scenarios with contiguous classes, as in the case of distribution of macroalgae.

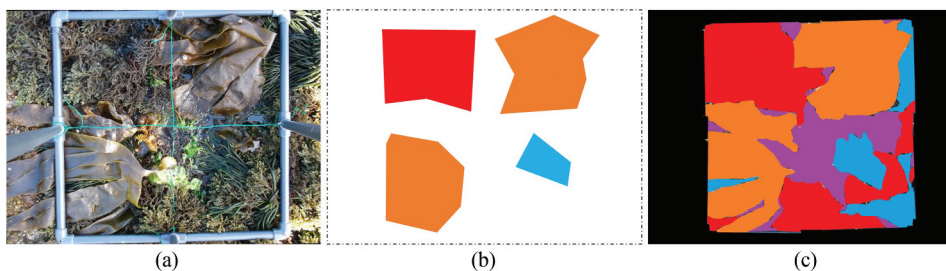


Figure 5. Types of labelled data: (a) geo-referenced and acquired data, (b) polygons labelled and selected for training in GIS software, (c) raster data for training a semantic segmentation CNN. Classes are represented by different colours.

Semantic segmentation delineates the classes more precisely, because it is a pixel-based classification, and it was therefore selected as the classification method for this research.

In this paper, we compared the performance of three CNNs in relation to semantic segmentation: MobileNetV2, Resnet18 and Xception. These networks each represent different architectures and perform well in segmentation/classification problems. In addition, the training cost of all three CNN is low, both in terms of computation and the number of labelled samples, as indicated by the number of hidden layers and adjustable parameters. Labelling a large number of samples is a tedious manual task that requires time from biologists familiar with differentiation of macroalgal species. In addition, many laboratories and professionals do not have access to expensive servers to train more complex neural networks, and they are limited to using personal computers. The characteristics of the different CNNs are summarized below:

- MobilNetv2. This CNN is specifically designed for operating on mobile devices, and the ratio between accuracy and cost of training is therefore particularly high. It consists of 53 layers and only 3.5 million adjustable parameters and is based on an inverted residual structure in which the shortcut connections are between the thin bottleneck layers (Sandler et al. 2018).
- ResNet18. This is the shallowest of the Deep Residual Networks. It has 18 layers and 11.7 million adjustable parameters. The most important aspect of this CNN is that, during training, it can skip layers if it considers that feature extraction does not contribute relevant information (He et al. 2016).
- Xception. This is an evolution of Inception architecture. It has 71 layers and 22.9 million adjustable parameters, and is thus the deepest of the networks used in this study. This CNN is based entirely on depth-wise separable convolution layers (Chollet 2017).

Images for semantic segmentation with CNN must have minimum dimensions according to the feature extractor ($224 \times 224 \times 3$ pixels for MobilNetv2 and ResNet18, and $299 \times 299 \times 3$ pixels for Xception). In the present study, the dimensions of the acquired images were $4288 \times 3216 \times 3$ pixels. Given this high resolution, the images included a great deal of detail, facilitating manual labelling by experts. Unfortunately, the amount of computer resources that must be allocated for network training increases with the image size. In order to train the networks on a conventional computer, the images were re-sized maintaining the aspect ratio of $1000 \times 750 \times 3$ pixels. This resolution still retained a high level of detail in the images. CNNs were adapted from image classification for semantic segmentation using DeepLabV3 (Chen et al. 2018) in Matlab.

2.2.3. Data augmentation and distribution

Data augmentation allows the training set to be extended to automatically generate new samples through small modifications of the original set. The data set was extended by applying reflections over the x and y axes, 20 pixels translations on both axes and rotations with angles less than 25° .

One aim of the data acquisition process was to obtain a representative number of images of each species. Nevertheless, the percentage occupation of each image was unbalanced (Table 1), which is a typical problem in semantic segmentation. Although

Table 1. Number of pixels per class.

Class	Total number of pixels ($\times 10^6$)	Training number of pixels ($\times 10^6$)	Validation number of pixels ($\times 10^6$)	Testing number of pixels ($\times 10^6$)
'Out'	58.93	41.16	4.47	13.29
<i>B. bifurcata</i>	6.28	4.46	0.47	1.34
<i>C. tamariscifolia</i>	4.96	3.50	0.68	0.77
<i>S. muticum</i>	1.36	0.59	0.30	0.47
<i>S. polyschides</i>	8.33	6.10	0.59	1.63
<i>Codium spp</i>	5.89	3.43	0.46	2.00
'Other'	11.76	8.25	0.52	2.99

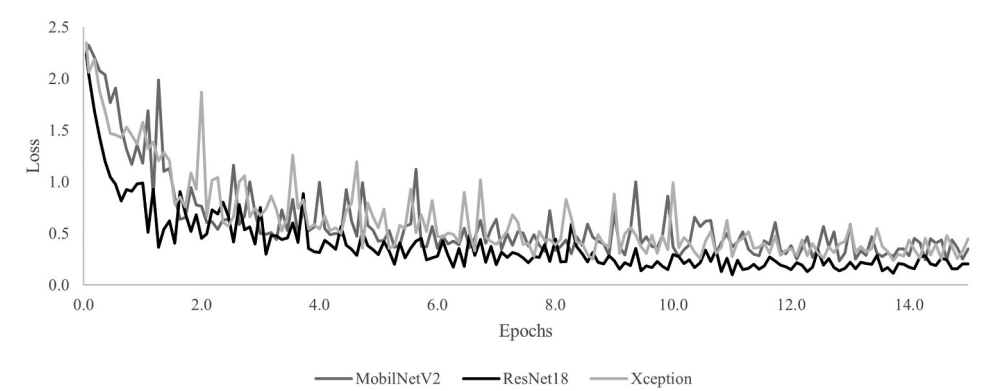


Figure 6. Variation in the loss during the training process.

almost all classes were of the same order of magnitude, relative to other species, there were very few samples of *S. muticum*. The 'out' class included a larger number of pixels, as it appeared in all images outside the ROI. The imbalance between classes can be minimized by assigning weights to the pixels according to the quantity in the training set. For the validation and testing sets, balanced sample sets were chosen to maximize the equivalence of the results. In total, 130 images were labelled and distributed as follows: 90 images for training, 10 images for validation and 30 images for testing.

2.2.4. Training

The network was trained on a laptop computer (GPU NVIDIA GTX1050 4GB GDDR5, CPU i7-7700HQ 2.8 Ghz and 16GB RAM DDR4). The hyperparameters were chosen experimentally after several tests, maximizing the performance and minimizing the overfitting. The hyperparameters that obtained the best result for training were as follows: optimization method, *sgdm*; learning rate, 0.001; momentum, 0.9; L2 regularization, 0.005; and max epochs, 15. The mini batch size was set at 4 and was limited by the amount of memory of the graphic card. The time consumed by each training was around 150 min. The programming language used was *Matlab*. All training sessions converged satisfactorily (Figure 6).

3. Results

The overall accuracy values obtained on training, validation and test sets are shown in Table 2. Overfitting was detected among the training and validation sets, although it was

Table 2. Overall accuracy obtained on training, validation and test sets.

	Train (%)	Validation (%)	Test (%)
ResNet18	93.7	90.6	91.9
MobileNetV2	91.4	85.0	88.4
Xception	90.9	84.7	87.3

Table 3. Confusion matrix for ResNet18.

Reference\predicted	Out	<i>B. bifurcata</i>	<i>C. tamariscifolia</i>	<i>S. muticum</i>	<i>S. polyschides</i>	<i>Codium spp</i>	Other
Out	0.962	0.004	0.001	0.001	0.012	0.010	0.009
<i>B. bifurcata</i>	0.003	0.921	0.002	0.009	0.001	0.007	0.058
<i>C. tamariscifolia</i>	0.004	0.020	0.590	0.042	0.017	0.032	0.296
<i>S. muticum</i>	0.015	0.008	0.197	0.618	0.018	0.004	0.141
<i>S. polyschides</i>	0.008	0.003	0.004	0.000	0.939	0.019	0.026
<i>Codium spp</i>	0.003	0.010	0.012	0.002	0.026	0.912	0.034
Other	0.008	0.025	0.032	0.004	0.025	0.052	0.854

Table 4. Confusion matrix for MobileNetV2.

Reference\predicted	out	<i>B. bifurcata</i>	<i>C. tamariscifolia</i>	<i>S. muticum</i>	<i>S. polyschides</i>	<i>Codium spp</i>	Other
out	0.902	0.008	0.005	0.004	0.023	0.027	0.032
<i>B. bifurcata</i>	0.007	0.907	0.006	0.007	0.002	0.014	0.058
<i>C. tamariscifolia</i>	0.003	0.016	0.606	0.039	0.009	0.013	0.314
<i>S. muticum</i>	0.031	0.006	0.121	0.601	0.005	0.008	0.229
<i>S. polyschides</i>	0.016	0.003	0.007	0.000	0.899	0.027	0.048
<i>Codium spp</i>	0.007	0.007	0.009	0.003	0.020	0.926	0.028
Other	0.012	0.016	0.034	0.004	0.018	0.044	0.871

Table 5. Confusion matrix for Xception.

Reference\predicted	Out	<i>B. bifurcata</i>	<i>C. tamariscifolia</i>	<i>S. muticum</i>	<i>S. polyschides</i>	<i>Codium spp</i>	Other
Out	0.908	0.012	0.008	0.001	0.019	0.021	0.031
<i>B. bifurcata</i>	0.004	0.884	0.020	0.000	0.001	0.014	0.077
<i>C. tamariscifolia</i>	0.006	0.019	0.725	0.003	0.009	0.031	0.206
<i>S. muticum</i>	0.017	0.014	0.500	0.266	0.003	0.010	0.190
<i>S. polyschides</i>	0.016	0.011	0.008	0.000	0.878	0.026	0.060
<i>Codium spp</i>	0.009	0.010	0.027	0.001	0.019	0.899	0.036
Other	0.011	0.023	0.076	0.002	0.017	0.042	0.829

reduced in the testing set. The difference between the validation and test sets is due to the difference in the number of pixels per class. Although the overfitting was reduced by adjusting the hyperparameters, it was not completely eliminated. The remaining overfitting was considered acceptable in view of the results, both qualitative and quantitative. The best result was obtained with Resnet18, although the performance was not the same for all classes. Confusion matrices for test data of the three CNNs are shown in [Tables 3–5](#). ResNet18 produced better segmentation of the *B. bifurcata*, *S. muticum*, *S. polyschides* and 'out' classes. Accurate identification of the 'out' class led to good delimitation of the ROI. MobileNetV2 produced better segmentation of the *Codium spp.* and 'others' (mainly composed of sand) classes, but produced very similar results to ResNet18. Xception produced by far the best results for the *C. tamariscifolia* class. In the confusion matrices, the success rates were lowest for the *C. tamariscifolia* and *S. muticum* classes, which yielded more confusion than the other classes. Specifically, *C. tamariscifolia* was confused

with the 'others' classes by 0.296, and *S. muticum* was confused with the 'others' by 0.197, and with *C. tamariscifolia* by 0.141. The colours of these classes were similar; in addition, very few samples of the *S. muticum* class were available for training. Good success rates were obtained for the remaining classes, and the confusion between them was minimal.

The most notable results for semantic segmentation with ResNet18 were the areas classified as macroalgae outside the ROI (Figure 7). However, these areas corresponded to macroalgae that were well classified and with continuous macroalgae within the ROI. In addition, although the centres of the macroalgae were well defined, the borders were quite irregular and not well defined. The borders did not fit properly, mainly in dark areas, overlapping areas between macroalgae or when a small macroalga was surrounded by another macroalga.

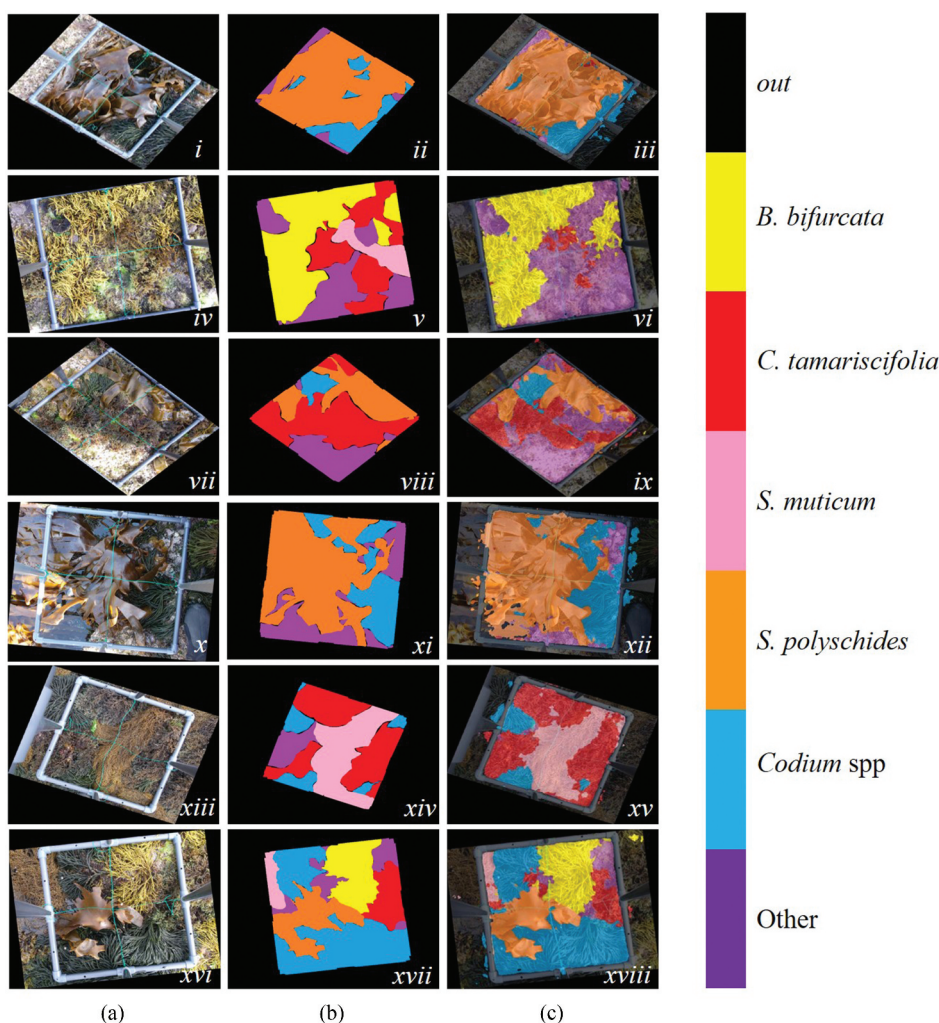


Figure 7. Acquired images (a), labelled images (b) and results (c) of semantic segmentation with ResNet18. Note: the acquired image was superimposed on colours representing each class.

Resnet18 produced the best segmentation of macroalgae. It was expected that MobileNetv2 would not perform particularly well, given the fewer configurable parameters. However, Xception did not produce better results, despite being a much deeper CNN with the capacity to extract more complex features. The Xception network only outperformed the other CNNs in the accuracy of segmenting the *C. tamariscifolia* class (one of the classes for which ResNet18 produced the least accurate results), but at the cost of increasing confusion about the *S. muticum* class, for which relatively poor results were obtained.

Very high success rates were obtained for the segmentation of most classes (including three different macroalgal species). ResNet18 learned the texture and colour patterns of different species, regardless of factors that led to changes, such as the time out of water between acquisitions. Although the *B. bifurcata* and *Codium* spp. classes were of similar texture, they were easily distinguished by their colour. The *S. polyschides* class did not coincide in colour or texture with any of the other classes. Low success rates (of around 60%) were only obtained for the *C. tamariscifolia* and *S. muticum* classes, possibly due to the similar colour and texture of these species.

4. Discussion

In this paper, we report a CNN-based segmentation procedure for macroalgae, which yielded a success rate > 90% for all three CNNs tested. One of the key reasons for the high success rate in classifying the species was the use of high-resolution ground images collected in the field. The images were rescaled in order to save time and computational resources, to a final resolution of 1000 × 750 pixels, which was high relative to the examples reported in the literature. By contrast, underwater images used to segment seagrass coverage were reduced to 512 × 256 pixels in previous studies (Weidmann et al. 2019). The study findings show that the proposed resolution satisfactory differentiated the five species and the interior/exterior zones of each quadrat. In addition, a laptop workstation was adequate for training the CNNs at this resolution, and computational resources from external servers were not required.

Although the accuracy rate was similar to that obtained in other studies with satellite, aerial and submarine sources, the present study aimed to differentiate five different macroalgae and it is, therefore, not generally comparable to other studies concerning the detection of single species. The accuracy achieved in number of CNN-based studies is very variable: 99.4% (Zhou et al. 2019), 97.0% (S. Wang et al. 2019), 95.8% (Rahnemoonfar and Dobbs 2019), 95.0% (Martin-Abadal et al. 2018) and 90.1% (Arellano-Verdejo, Lazcano, and Cabanillas-Terán 2018). These studies, generally based on satellite and airborne data of lower resolution than ground data, have only focused on detecting the predominant macroalgal species. The technical complexity of these studies is considerably lower than that presented here. The previous studies segmented one macroalgae class from the bottom, often sand or water, without differentiating between macroalgal species. Differentiating between macroalgal species is feasible when broad taxonomic groups are considered, e.g. green, red and brown algae (Andrefouet et al. 2004; Kotta et al. 2018).

The findings of the present study showed that colour attributes alone were not sufficient for correct classification, as *C. tamariscifolia* and *S. muticum* are very similar in colour and only differ in texture. Depending on the classification scale, the texture feature

is not extractable from satellite and aerial images due to the lower resolution of these. In addition, the same species can display notable differences in colour depending on the morphology, thickness of thalli and cellular architecture, which determine pigment density, absorption and thus reflectance spectra (Vogelmann and Björn 1986). Environmental conditions, such as the emersion time and intensity of solar radiation, also contribute to differences in pigmentation within and between macroalgal species (Dieter, Wiencke, and Bischof 2004). The resolution of images obtained by underwater drones is higher than that of airborne data, and the colour is modified relative to images taken outside the water; however, the modification affects all species equally (O'Byrne et al. 2018). Macroalgal species should be able to be differentiated by these characteristics (resolution and colour), and therefore texture, in underwater images. Nevertheless, most studies based on underwater images and also studies based on satellite and aerial images have only focused on detecting single macroalgal species (Gonzalez-Cid et al. 2017; Moniruzzaman et al. 2019).

The features extracted from one species were easier to learn with a classifier based on artificial intelligence, as the macroalgal classes shared more features with each other than with non-macroalgal classes such as rock, sand and seawater. However, when the macroalgal class no longer corresponded to one species and was divided into five classes, as in this study, it became more difficult for the algorithm to both find and extract distinctive features.

The CNNs under study proved very useful for segmentation of the five macroalgae considered. Although, in theory, a large number of labelled images was required, in practice the number of pixels was more important for semantic segmentation. Given the high resolution of the images used, the number of pixels was sufficient to train a CNN with only 100 images, which can be obtained quickly. From these 100 labelled images, and after training, infinite images can be labelled without further human intervention. However, the labelling process for training must be conducted carefully, as CNNs can learn labelling errors. Confusion at the borders of macroalgae (Figure 7) was due to labelling errors in the images (Vogelmann and Björn 1986). For human observers, the centre of the algae is easy to segment and label manually, as with CNN segmentation. However, the borders of many algae overlap and it is not easy to define outlines to separate them. Because of these errors in the labelled images, the errors were also learnt by the CNN and replicated in the segmentation. These types of errors tend to be minimized when the data are tagged by different people.

The acquisition time is slower with ground imaging than in satellite and drone-based methods because the quadrat has constantly to be moved to a new location for each new image. In addition, the area covered by each image was only 0.25 m². Nevertheless, ground images are required in order to provide reference data to train models and map data obtained with other automated instruments. This study focused on the exclusive use of RGB photographic images to minimize pre-processing time by fusing information and acquisition efforts. Because of the high success rates obtained, other types of data, such as spectroradiometer data (Hu 2009), multispectral-hyperspectral images (Li et al. 2012; Taddia et al. 2019; Fauzan et al. 2017; Zacharias, Niemann, and Borstad 1992) and environmental data (De Oliveira and Brigitte 2006), were not included.

5. Conclusion and future work

This study involved CNN-based semantic segmentation of high-resolution ground images of five different macroalgae inhabiting rocky shores. The study findings demonstrate that vector-labelled samples can be adapted for use with CNNs. Of the three different CNNs compared, ResNet18 produced the best results, i.e. 91.9% accuracy. Most of the samples were correctly labelled, although there was a tendency for some macroalgae outside the ROI to be labelled, and the borders between macroalgal species were diffuse. Although theoretically considered an error, in practice segmentation of macroalgae outside the ROI is not problematical, as long as the classification is correct, as in this case. Definition of borders is also a problem experienced by human observers. The proposed method is therefore considered a suitable alternative for the automation of sample labelling.

The study findings demonstrated that automation of the labelling process is possible with only 100 high-resolution images obtained in the field, without the need for other types of data. A further step will be to apply the method to UAV-acquired data. Nevertheless, further research is required in ground resolution effects to guarantee correct results and for transfer to learning between UAV and ground images, which differ in resolution and, therefore, in texture and colour. The use of UAVs, together with the findings presented here, will facilitate the rapid acquisition and mapping of macroalgal cover on intertidal rocky shores, with a high degree of automation. Use of these methods could greatly improve the management of coastal areas.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the Fundación Biodiversidad, the Ministerio para la Transición Ecológica y el Reto Demográfico through the Pleamar program, co-funded by the European Maritime and Fisheries Fund (EMFF) [call 2018]. It was also partly funded through grants awarded by the Xunta de Galicia for human resources and competitive reference groups [ED481B-2019-061 and ED431C 2016-038]; and by the Ministerio de Ciencia, Innovación y Universidades -Gobierno de España [RTI2018-095893-B-C21]. This document only reflects the views of the authors, and the statements made herein are solely the responsibility of the authors.

ORCID

Jesús Balado  <http://orcid.org/0000-0002-3758-3102>

Joaquín Martínez-Sánchez  <http://orcid.org/0000-0003-0320-4191>

References

Andrefouet, S., C. Payri, E. Hochberg, C. Hu, M. Atkinson, and F. Muller-Karger. 2004. "Use of in Situ and Airborne Reflectance for Scaling-Up Spectral Discrimination of Coral Reef Macroalgae from Species to Communities." *Marine Ecology-Progress Series* 283: 161–177. November. doi:[10.3354/meps283161](https://doi.org/10.3354/meps283161).

- Arellano-Verdejo, J., H. Lazcano, and N. Cabanillas-Terán. 2018. In *ERISNet: Deep Learning Network for Sargassum Detection along the Coastline of the Mexican Caribbean* edited by PeerJ, 7: e6842. <https://doi.org/10.7717/peerj.6842>
- Bañolas, G., S. Fernández, F. Espino, R. Haroun, and F. Tuya. 2020. "Evaluation of Carbon Sinks by the Seagrass *Cymodocea Nodosa* at an Oceanic Island: Spatial Variation and Economic Valuation." *Ocean & Coastal Management* 187: 105112. doi:10.1016/j.ocecoaman.2020.105112.
- Casal, G., T. Kutser, J. Gómez, N. Sánchez-Carnero, and J. Freire. 2013. "Assessment of the Hyperspectral Sensor CASI-2 for Macroalgal Discrimination on the Ría De Vigo Coast (NW Spain) Using Field Spectroscopy and Modelled Spectral Libraries." *Continental Shelf Research* 55: 129–140. March. doi:10.1016/j.csr.2013.01.010.
- Chen, L.-C., Y. Zhu, G. Papandreou, F. Schroff, and H. Adam. 2018. "Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation." *Proceedings of the European Conference on Computer Vision (ECCV)*, 801–818. 15th European Conference, September 8–14, Munich, Germany.
- Chollet, F. 2017. "Xception: Deep Learning with Depthwise Separable Convolutions." 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 1800–1807. doi:10.1109/CVPR.2017.195.
- De Oliveira, E. P. J., and G. Brigitte. 2006. "Predictive Modelling of Coastal Habitats Using Remote Sensing Data and Fuzzy Logic: A Case for Seaweed in Brittany (France)." In *EARSeL EProceedings*, Edited by European Association of Remote Sensing Laboratories. <https://archimer.ifremer.fr/doc/00000/2279/>
- Dieter, H., C. Wiencke, and K. Bischof. 2004. "Photosynthesis in Marine Macroalgae." *Photosynthesis in Algae* 14: 413–435. doi:10.1007/978-94-007-1038-2_18.
- Duarte, L., R. M. Viejo, B. Martínez, M. deCastro, M. Gómez-Gesteira, and T. Gallardo. 2013. "Recent and Historical Range Shifts of Two Canopy-Forming Seaweeds in North Spain and the Link with Trends in Sea Surface Temperature." *Acta Oecologica* 51: 1–10. doi:10.1016/j.actao.2013.05.002.
- Duffy, J. P., L. Pratt, K. Anderson, P. E. Land, and J. D. Shutler. 2018. "Spatial Assessment of Intertidal Seagrass Meadows Using Optical Imaging Systems and a Lightweight Drone." *Estuarine, Coastal and Shelf Science* 200: 169–180. doi:10.1016/j.ecss.2017.11.001.
- Fauzan, M. A., I. S. W. Kumara, R. Yogyantoro, S. Suwardana, N. Fadhillah, I. Nurmallasari, S. Apriyani, and P. Wicaksono. 2017. "Assessing the Capability of Sentinel-2A Data for Mapping Seagrass Percent Cover in Jerowaru, East Lombok." *The Indonesian Journal of Geography* 49 (2): 195–203. Universitas Gadjah Mada, Faculty of Geography. doi:10.22146/ijg.28407.
- Floor, J. R., C. S. A. (Kris) van Koppen, and J. P. M. van Tatenhove. 2018. "Science, Uncertainty and Changing Storylines in Nature Restoration: The Case of Seagrass Restoration in the Dutch Wadden Sea." *Ocean & Coastal Management* 157: 227–236. doi:10.1016/j.ocecoaman.2018.02.016.
- Gonzalez-Cid, Y., A. Burguera, F. Bonin-Font, and A. Matamoros. 2017. "Machine Learning and Deep Learning Strategies to Identify Posidonia Meadows in Underwater Images." *OCEANS 2017 - Aberdeen*, 1–5. doi:10.1109/OCEANSE.2017.8084991.
- Griffiths, L. L., R. M. Connolly, and C. J. Brown. 2020. "Critical Gaps in Seagrass Protection Reveal the Need to Address Multiple Pressures and Cumulative Impacts." *Ocean & Coastal Management* 183: 104946. doi:10.1016/j.ocecoaman.2019.104946.
- Hawkins, S., H. Sugden, N. Mieszkowska, P. Moore, E. Poloczanska, R. Leaper, R. Herbert, et al. 2009. "Consequences of Climate-Driven Biodiversity Changes for Ecosystem Functioning of North European Rocky Shores." *Marine Ecology Progress Series* 396: 245–259. January. doi:10.3354/meps08378.
- He, K., X. Zhang, S. Ren, and J. Sun. 2016. "Deep Residual Learning for Image Recognition." *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 770–778. doi:10.1109/CVPR.2016.90.
- Hu, C. 2009. "A Novel Ocean Color Index to Detect Floating Algae in the Global Oceans." *Remote Sensing of Environment* 113 (10): 2118–2129. doi:10.1016/j.rse.2009.05.012.

- Jenkins, S. R., P. Moore, M. T. Burrows, D. J. Garbary, S. J. Hawkins, A. Ingólfsson, K. P. Sebens, P. V. R. Snelgrove, D. S. Wetthey, and S. A. Woodin. 2008. "Comparative Ecology of North Atlantic Shores: Do Differences in Players Matter for Process?" *Ecology* 89 (sp11): S3–S23. John Wiley & Sons, Ltd. doi:10.1890/07-1155.1.
- Kellaris, A., A. Gil, J. Faria, R. Amaral, I. Moreu-Badia, A. Neto, and C. Yesson. 2019. "Using Low-Cost Drones to Monitor Heterogeneous Submerged Seaweed Habitats: A Case Study in the Azores." *Aquatic Conservation: Marine and Freshwater Ecosystems* 29 (11): 1909–1922. John Wiley & Sons, Ltd. doi:10.1002/aqc.3189.
- Kotta, J., N. Valdivia, T. Kutser, K. Toming, M. Rätsep, and H. Orav-Kotta. 2018. "Predicting the Cover and Richness of Intertidal Macroalgae in Remote Areas: A Case Study in the Antarctic Peninsula." *Ecology and Evolution* 8 (17): 9086–9094. John Wiley & Sons, Ltd. doi:10.1002/ece3.4463.
- Lamela-Silvarrey, C., C. Fernández, R. Anadón, and J. Arrontes. 2012. "Furoid Assemblages on the North Coast of Spain: Past and Present (1977–2007)." *Botanica Marina* 55 (3). doi:10.1515/bot-2011-0081.
- Li, R., J.-K. Liu, A. Sukcharoenpong, J. Yuan, H. Zhu, and S. Zhang. 2012. "A Systematic Approach toward Detection of Seagrass Patches from Hyperspectral Imagery." *Marine Geodesy* 35 (3): 271–286. Taylor & Francis. doi:10.1080/01490419.2012.699019.
- Lima, F. P., P. A. Ribeiro, N. Queiroz, S. J. Hawkings, and A. M. Santos. 2007. "Do Distributional Shifts of Northern and Southern Species of Algae Match the Warming Pattern?" *Global Change Biology* 13 (12): 2592–2604. John Wiley & Sons, Ltd. doi:10.1111/j.1365-2486.2007.01451.x.
- Macreadie, P. I., J. Jarvis, S. M. Trevathan-Tackett, and A. Bellgrove. 2017. "Seagrasses and Macroalgae: Importance, Vulnerability and Impacts." *Climate Change Impacts on Fisheries and Aquaculture*. Wiley Online Books. doi:10.1002/9781119154051.ch22.
- Martin-Abadal, M., E. Guerrero-Font, F. Bonin-Font, and Y. Gonzalez-Cid. 2018. "Deep Semantic Segmentation in an AUV for Online Posidonia Oceanica Meadows Identification." *IEEE Access* 6: 60956–60967. doi:10.1109/ACCESS.2018.2875412.
- Moniruzzaman, M., S. M. S. Islam, P. Lavery, and M. Bennamoun. 2019. "Faster R-CNN Based Deep Learning for Seagrass Detection from Underwater Digital Images." 2019 Digital Image Computing: Techniques and Applications (DICTA), 1–7. doi:10.1109/DICTA47822.2019.8946048.
- O'Byrne, M., V. Pakrashi, F. Schoefs, and B. Ghosh. 2018. "Semantic Segmentation of Underwater Imagery Using Deep Networks Trained on Synthetic Imagery." *Journal of Marine Science and Engineering* 6 (3): 93. doi:10.3390/jmse6030093.
- Purvaja, R., R. S. Robin, D. Ganguly, G. Hariharan, G. Singh, R. Raghuraman, and R. Ramesh. 2018. "Seagrass Meadows as Proxy for Assessment of Ecosystem Health." *Ocean & Coastal Management* 159: 34–45. doi:10.1016/j.ocecoaman.2017.11.026.
- Qi, L., C. Hu, K. Mikelsons, M. Wang, V. Lance, S. Sun, B. B. Barnes, J. Zhao, and D. Van der Zande. 2020. "In Search of Floating Algae and Other Organisms in Global Oceans and Lakes." *Remote Sensing of Environment* 239: 111659. doi:10.1016/j.rse.2020.111659.
- Rahnemoonfar, M., and D. Dobbs. 2019. "Semantic Segmentation of Underwater Sonar Imagery with Deep Learning." IGARSS 2019 IEEE International Geoscience and Remote Sensing Symposium, 9455–9458. doi:10.13140/RG.2.2.32343.52644.
- Reus, G., T. Moller, J. Jager, S. Schultz, C. Kruschel, J. Hasenauer, V. Wolff, and K. Fricke-Neuderth. 2018. *Looking for Seagrass: Deep Learning for Visual Coverage Estimation*. In 2018 OCEANS-MTS/ IEEE Kobe Techno-Oceans (OTO) (pp. 1–6). IEEE. doi:10.1109/OCEANSKOB.2018.8559302.
- Ruiz-Santaquiteria, J., G. Bueno, O. Deniz, N. Vallez, and G. Cristobal. 2020. "Semantic versus Instance Segmentation in Microscopic Algae Detection." *Engineering Applications of Artificial Intelligence* 87: 103271. doi:10.1016/j.engappai.2019.103271.
- Sagawa, T., A. Mikami, M. Aoki, and T. Komatsu. 2012. "Mapping Seaweed Forests with IKONOS Image Based on Bottom Surface Reflectance." Proceedings of SPIE - The International Society for Optical Engineering. 8525 vols. doi:10.1117/12.975678.
- Sandler, M., A. Howard, M. Zhu, A. Zhmoginov, and L. Chen. 2018. "MobileNetV2: Inverted Residuals and Linear Bottlenecks." 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 4510–4520. doi:10.1109/CVPR.2018.00474.

- Taddia, Y., P. Russo, S. Lovo, and A. Pellegrinelli. 2019. "Multispectral UAV Monitoring of Submerged Seaweed in Shallow Water." *Applied Geomatics*. May. doi:10.1007/s12518-019-00270-x.
- Tamondong, A., C. Cruz, J. Guihawan, M. Garcia, R. R. Quides, J. Cruz, and A. Blanco. 2018. *Remote Sensing-Based Estimation of Seagrass Percent Cover and LAI for above Ground Carbon Sequestration Mapping*. In *Remote Sensing of the Open and Coastal Ocean and Inland Waters* (Vol. 10778, p. 1077803). International Society for Optics and Photonics. doi:10.1117/12.2324695.
- Topouzelis, K., S. C. Spondylidis, A. Papakonstantinou, and N. Soulakellis. 2016. "The Use of Sentinel-2 Imagery for Seagrass Mapping: Kalloni Gulf (Lesvos Island, Greece) Case Study." *Proceeding SPIE* 9688. doi:10.1117/12.2242887.
- Ventura, D., A. Bonifazi, F. Maria Gravina, A. Belluscio, and G. Ardizzone. 2018. "Mapping and Classification of Ecologically Sensitive Marine Habitats Using Unmanned Aerial Vehicle (UAV) Imagery and Object-Based Image Analysis (OBIA)." *Remote Sensing* 10 (9): 1331. doi:10.3390/rs10091331.
- Vogelmann, T. C., and L. O. Björn. 1986. "Plants as Light Traps." *Physiologia Plantarum* 68 (4): 704–708. John Wiley & Sons, Ltd. doi:10.1111/j.1399-3054.1986.tb03421.x.
- Wang, M., X. Fei, Y. Zhang, Z. Chen, X. Wang, Y. Jin Tsou, D. Liu, and X. Lu. 2018. "Assessing Texture Features to Classify Coastal Wetland Vegetation from High Spatial Resolution Imagery Using Completed Local Binary Patterns (CLBP)." *Remote Sensing*. doi:10.3390/rs10050778.
- Wang, S., L. Liu, L. Qu, C. Yu, Y. Sun, F. Gao, and J. Dong. 2019. "Accurate Ulva Prolifera Regions Extraction of UAV Images with Superpixel and CNNs for Ocean Environment Monitoring." *Neurocomputing* 348: 158–168. doi:10.1016/j.neucom.2018.06.088.
- Weidmann, F., J. Jäger, G. Reus, S. T. Schultz, C. Kruschel, V. Wolff, and K. Fricke-Neuderth. 2019. "A Closer Look at Seagrass Meadows: Semantic Segmentation for Visual Coverage Estimation." *OCEANS 2019 - Marseille*, 1–6. doi:10.1109/OCEANSE.2019.8867064.
- Wilson, K. L., M. A. Skinner, and H. K. Lotze. 2019. "Eelgrass (*Zostera Marina*) and Benthic Habitat Mapping in Atlantic Canada Using High-Resolution SPOT 6/7 Satellite Imagery." *Estuarine, Coastal and Shelf Science* 226: 106292. doi:10.1016/j.ecss.2019.106292.
- Xing, Q., D. An, X. Zheng, Z. Wei, X. Wang, L. Li, L. Tian, and J. Chen. 2019. "Monitoring Seaweed Aquaculture in the Yellow Sea with Multiple Sensors for Managing the Disaster of Macroalgal Blooms." *Remote Sensing of Environment* 231: 111279. doi:10.1016/j.rse.2019.111279.
- Zacharias, M., O. Niemann, and G. Borstad. 1992. "An Assessment and Classification of a Multispectral Bandset for the Remote Sensing of Intertidal Seaweeds." *Canadian Journal of Remote Sensing* 18 (4): 263–274. Taylor & Francis. doi:10.1080/07038992.1992.10855331.
- Zhou, Y., J. Wang, B. Li, Q. Meng, E. Rocco, and A. Saiani. 2019. "Underwater Scene Segmentation by Deep Neural Network." https://repository.lboro.ac.uk/articles/Underwater_scene_segmentation_by_deep_neural_network/9405458