electronics

MDPI

Article

Enhanced Satellite Analytics for Mussel Platform Census Using a Machine-Learning Based Approach

Fernando Martín-Rodríguez *, Luis M. Álvarez-Sabucedo, Juan M. Santos-Gago and Mónica Fernández-Barciela

atlanTTic Research Center for Telecommunication Technologies, University of Vigo, C/Maxwell SN, 36310 Vigo, Spain; luis.sabucedo@uvigo.gal (L.M.Á.-S.); jsgago@det.uvigo.es (J.M.S.-G.); monica.barciela@uvigo.es (M.F.-B.)
* Correspondence: fmartin@uvigo.es; Tel.: +34-986-812151

Abstract: Mussel platforms are big floating structures made of wood (normally about 20 m × 20 m or even a bit larger) that are used for aquaculture. They are used for supporting the growth of mussels in suitable marine waters. These structures are very common near the Galician coastline. For their maintenance and tracking, it is quite convenient to be able to produce a periodic census of these structures, including their current count and position. Images from Earth observation satellites are, a priori, a convenient choice for this purpose. This paper describes an application capable of automatically supporting such a census using optical images taken at different wavelength intervals. The images are captured by the two Sentinel 2 satellites (Sentinel 2A and Sentinel 2B, both from the Copernicus Project). The Copernicus satellites are run by the European Space Agency, and the produced images are freely distributed on the Internet. Sentinel 2 images include thirteen frequency bands and are updated every five days. In our proposal, remote-sensing normalized (differential) indexes are used, and machine-learning techniques are applied to multiband data. Different methods are described and tested. The results obtained in this paper are satisfactory and prove the approach is suitable for the intended purpose. In conclusion, it is worth noting that artificial neural networks turn out to be particularly good for this problem, even with a moderate level of complexity in their design. The developed methodology can be easily re-used and adapted for similar marine environments.

Keywords: remote sensing; satellite imaging; Sentinel 2; aquaculture; mussel rafts; machine-learning

check for updates

Citation: Martín-Rodríguez, F.; Álvarez-Sabucedo, L.M.; Santos-Gago, J.M.; Fernández-Barciela, M. Enhanced Satellite Analytics for Mussel Platform Census Using a Machine-Learning Based Approach. *Electronics* **2024**, *13*, 2782. https://doi.org/10.3390/electronics13142782

Academic Editors: Yue Wu, Songyang Zhang and Shuai Zhang

Received: 14 May 2024
Revised: 2 July 2024
Accepted: 11 July 2024
Published: 15 July 2024

Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

The mussel is a highly valued culinary product in several countries due to its high nutritional value, flavor, and versatility in cooking. The cultivation of this product is practiced in coastal areas that meet suitable conditions for its growth, mainly in bays, estuaries, and rias. These mussels are mainly cultured on floating wooden platforms, from which thick ropes are hung where the mollusk seeds are attached. Local authorities manage licenses that include the allowed position for each platform. For enforcing this regulation, it is interesting to elaborate a periodic census of these structures (a periodical position check) that would allow the detection of changes. It is normal to find missing platforms because they have been taken to be maintained, but there should not be ones that are unregistered, moved into illegal positions, or even unlicensed. Satellites that obtain periodic images for Earth observation can be a convenient choice for this issue. Note that the satellite can check positions but not verify the license numbers due to the limitations of the available resolution. The discovery of irregularities should trigger an "on-site" inspection using other means, like drones or inspection boats.

Among the different sets of images from existing satellites, in this approach, Sentinel 2 was the chosen option. This is a satellite deployed within the Copernicus project [1], operated by the European Space Agency (ESA). The produced images are freely obtained on the Internet [2]. These are multispectral images of thirteen bands [3] that are updated every five days (see Figure 1 for a band listing). In our application, we use normalized

differential indices (very typical in remote sensing) and machine-learning techniques applied to multiband data. Different methods are described and tested, and the obtained results are presented.

Sentinel-2 Bands	Central Wavelength (µm)	Resolution (m)
Band 1 - Coastal aerosol	0.443	60
Band 2 - Blue	0.490	10
Band 3 - Green	0.560	10
Band 4 - Red	0.665	10
Band 5 - Vegetation Red Edge	0.705	20
Band 6 - Vegetation Red Edge	0.740	20
Band 7 - Vegetation Red Edge	0.783	20
Band 8 - NIR	0.842	10
Band 8A - Vegetation Red Edge	0.865	20
Band 9 - Water vapour	0.945	60
Band 10 - SWIR - Cirrus	1.375	60
Band 11 - SWIR	1.610	20
Band 12 - SWIR	2.190	20

Figure 1. Sentinel 2 bands.

The main contribution of this paper is the creation of a new application for raft position monitoring (a novel detection method), which can track changes and help local authorities to apply regulations and maintain and update their data. The machine-learning methods used are not new, but the application is completely novel.

1.1. Related Work

The use of machine-learning techniques on remote-sensing aerial or satellite images is a common procedure nowadays. Many examples can be found in the literature, with different purposes. For example, in [4], the authors use convolutional neural networks (CNNs) to estimate canopy height in forest areas. In [5], the authors demonstrate the superior performance of machine learning over traditional methods in the problem of river water segmentation. In [6], the authors compare various deep-learning techniques, such as "You Only Look Once" (YOLO) and faster recurrent convolutional neural networks (RCNNs) for the detection of shape-defined targets.

Focusing on the problem of rafts and wooden floating structures, there are other systems with the same purpose using other kinds of satellite information. They use SAR (Synthetic Aperture Radar) data. These data can be obtained from Sentinel 1 or other older satellites. For example, in [7], Marino et al. studied the Vigo estuary using SAR data from COSMO-SkyMed. In [8], the authors use the same kind of information for a similar purpose in Matsushima Bay, Japan. Other examples using SAR data are [9] in the Philippines, [10] in China, and [11] in the Philippines. Generally, the results from SAR data have much less precision due to the minor resolution of SAR data (30 × 30 m in COSMO-SkyMed) and less spectral resolution (bands HH and HV). The aforementioned publications detect groups of rafts (polygons) instead of individual ones. Nevertheless, the use of new SAR sensors, like those on Sentinel 1, can yield interesting results and is regarded as a main future line of work.

Another application found in [12] is interesting. This reference is not a paper but a Python notebook. The authors use ANNs over visible data (RGB), but they count with high-resolution information from the Arousa estuary. This means (according to the authors) 30 cm pixels. The origin of the images is not declared, but it is probably a (high-cost) high-resolution satellite image or even a photomosaic from aerial imagery. This is a kind of

image that is very expensive to obtain for periodic monitoring. Conversely, Sentinel images are free. Furthermore, the authors do not document tests with more than this individual image, which necessarily corresponds to a single date.

In [13], the authors also use visible (RGB) high-resolution images for detecting buoys that signal underwater mussel farms in New Zealand. Here, the authors use "in situ" data from cameras located at floating platforms and traversing vessels. In [14], Zeng et al. conduct a very similar work detecting buoys from "in situ" RGB images using YOLO.

Compared to these precedents, the work in this paper uses an intermediate resolution between SAR-based imagery and "in situ" photographs. Availability and periodicity are assured compared to "in situ" solutions, thereby achieving better coverage of the regions of interest. Per pixel processing using multiple wavelength bands yields the detection of individual rafts, achieving the objective of an automated periodic census.

1.2. Remainder of Paper

The remainder of this paper is organized as follows: in Section 2, data acquisition is described and the proposed methods are developed; in Section 3, the proposed system is tested with real-world data, comparing preliminary (dataset only) results with real-world application. In Section 4, an election of final methods is presented and justified, whereas in Section 5 conclusions and possible future lines are presented. Appendix A is simply informative and describes Sentinel 2 bands with graphical examples for interested readers.

2. Materials and Methods

2.1. About the Images Used

The images obtained from Sentinel 2 have a spatial resolution of 10 m per pixel (really, only some bands have that pixel size, and there exist bands of 20 m and 60 m) which would result in a size of 2×2 or 3×3 pixels for the rafts. In addition, as the structure of a platform is not a continuous wooden platform but, rather, a lattice of planks (Figure 2), the rafts appear in the visible bands only as small squares within the water with a color a little less saturated than their surroundings (Figure 3). Therefore, it will be necessary to use the non-visible bands of the image to be able to make reliable detections. Note that detection will be performed with a pixel classification basis, searching to characterize the special nature of a platform pixel (a lattice of wood with seawater in the background that yields an average pixel with the influence of both materials on the reflection). It is not possible to detect any shape, a fact that discards methods like YOLO, Detectron, or even the CNN's.



Figure 2. Platforms from the air.

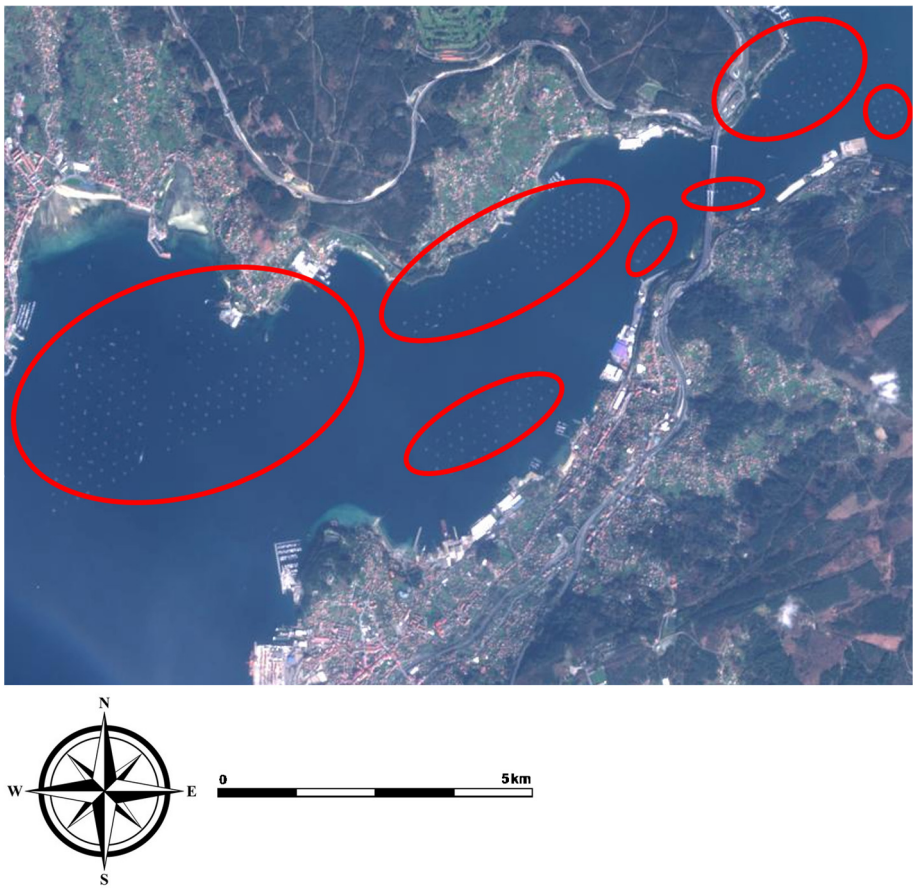


Figure 3. Polygons of mussel rafts in a Sentinel 2 image (Vigo estuary, Spain). Rafts detected inside the red ellipses.

Sentinel 2 only has coverage in near-shore waters and inland seas. In our case, this is more than enough.

In Appendix A, an informational explanation of each band utility is provided, including real example images.

Sentinel's public repository contains images of 100 km × 100 km (100 Mpx with 10 m pixels) that comprise all bands and are updated every five days. For each image, we have two versions: the TOA correction that contains the thirteen bands and the BOA correction that only contains twelve, since the tenth band is used within the correction process to estimate the atmospheric state [15].

As shown later on, within our proposal, both types of images were used for testing purposes. In both cases, we have discarded the 60 m bands due to the significant scaling required to combine them with the others and because they provide information very dependent on the atmosphere.

2.2. Detection Method

The process for addressing the planned issue consists of several steps. Initially, data are collected from the repository after satellite acquisition. These data are then analyzed to segregate water pixels. For water body detection, two methods are tested and compared (NDWI: Normalized Differential Water Index, and MLP: Multi-Layer Perceptron). Following this, machine-learning algorithms are tested to detect pixels associated with rafts. Three methods have been tested and compared here (MLP: Multi-Layer Perceptron, SVM: Support Vector Machine, and BT: Bagged Tree). The final stage involves a post-processing procedure to eliminate false positives. The entire workflow is depicted in Figure 4.

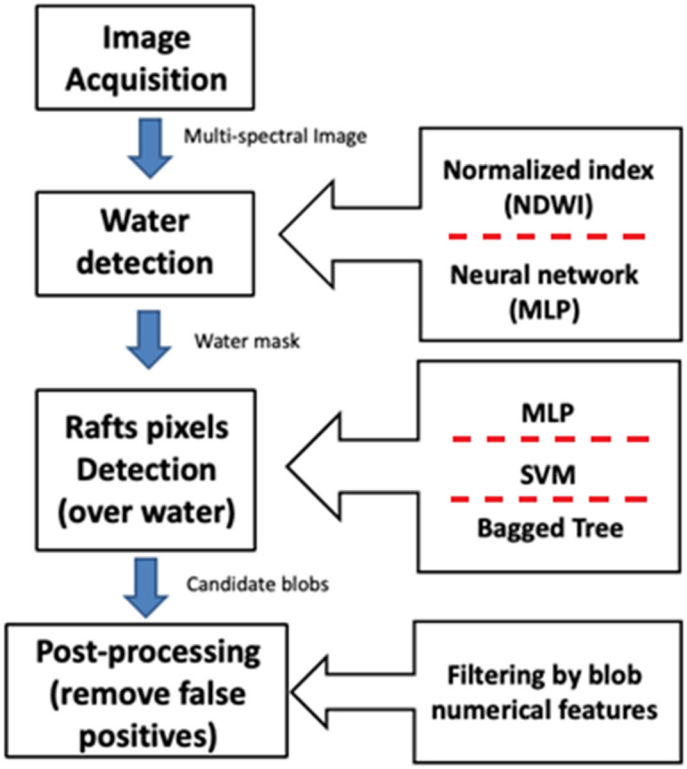


Figure 4. General system flow diagram. Red dotted lines indicate alternate methods.

2.3. Water Detection

The first purpose is to detect an area of interest in which to apply a detector that can distinguish the points belonging to platforms. It would be possible to use an already existing map to work over sea points, but a more general approach is preferable as that information is not always available. In addition, a water detection method will exclude cloud-covered areas and will also consider the effect of tides or other variations. For this issue, we have tested two different approaches: first, the use of a normalized index and, second, using a neural network directly on pixel information to classify the terrain type (detecting water points in this case).

2.3.1. Calculation of Normalized Indexes

In remote sensing, the so-called normalized indexes are used very often. Indexes are calculated from pairs of components [16,17]. In particular, the *NDWI* (Normalized Differential Water Index) is defined as follows:

$$NDWI = \frac{GREEN - NIR}{GREEN + NIR} \quad (1)$$

This value is calculated from Bands 3 (*GREEN*) and 8 (*NIR*). *NDWI* will always be in the range $[-1, +1]$, and the positive values will tend to correspond to bodies of water, while the negative ones will be dry areas. *NDWI* was chosen due to its ease of computation and, as will be demonstrated in the following paragraphs, with appropriate post-processing it can solve the problem very efficiently.

As it can be seen in Figure 5, the brighter (numerically larger) values correspond to water. Nevertheless, the value obtained for water differs between images captured on different days. By making all negative pixels equal to zero (a non-linear operation equivalent to a ReLU function), a bimodal histogram is achieved with a strong peak at zero (from non-water pixels) and another one in the gray level corresponding to water points. ReLU is a nonlinear function extensively used in neural network applications, normally in the first stages of convolutional neural networks (feature extraction convolutional stages).

Although a convolutional network is not used here, ReLU is used for completing feature extraction after NDWI computation. ReLU is defined as $(x + |x|)/2$ (the identity for $x > 0$, 0 otherwise). See the histogram in Figure 6, where a logarithmic scale is necessary to avoid that peak at zero making other values go unnoticed.

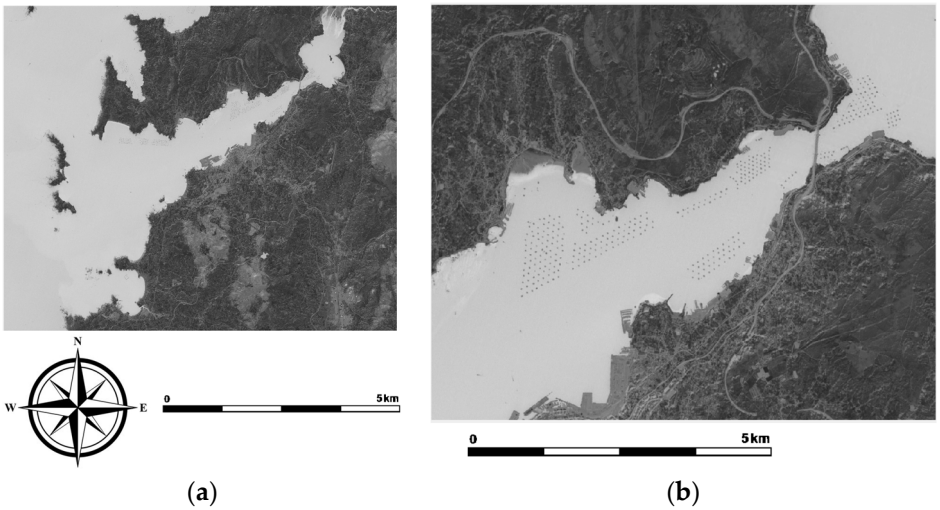


Figure 5. NDWI is represented in grayscale. (**a**) Entire study area, (**b**) zoom of the rafts.

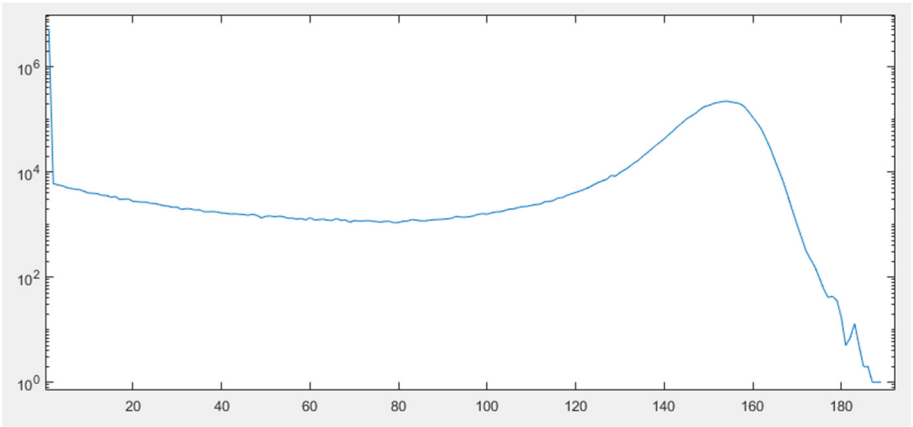


Figure 6. Histogram of the NDWI image (logarithmic scale). *X*-axis: NDWI image value, *Y*-axis: value repetition count.

At this stage, the well-known Otsu method [18] can calculate an adequate threshold to distinguish water. The Otsu method has its foundations in statistics and it is able to find the optimum threshold to split the histogram into two partial distributions with minimum partial variance; graphically, this corresponds to the minimum between the two maxima. Despite its simplicity, the application of this index proved to be extremely relevant for the final detection, according to the tests conducted. No similar post-processing was found in the reviewed literature.

2.3.2. Detection Using Neural Networks

Using the methodology designed in [19], a feature vector for each pixel is defined, consisting of the values of each band at that point. Therefore, this will become a numerical vector of size 10 once the lower-resolution bands (Bands 1, 9, and 10) are removed. Note that for the bands of resolution equal to 20 m, we will have to perform an interpolation, for which we choose the filter "Lanczos3" [20]. To classify these vectors, we train a simple neural network of the MLP (Multi-Layer-Perceptron) type [21]. The use of MLP with this

feature vector size turned out to be quite convenient. The application of, for example, a convolutional neural network (CNN) in this context is not convenient as there is no discernible pattern (shape) to be identified. In other words, with only 10 numerical features there is no need to apply convolutional stages.

In this case (as supported by [16]), the network is trained to distinguish 5 types of surfaces: 1 is empty (part of the image without information), 2 is solid ground, 3 is water, 4 is cloud, and 5 is foam (foam points on the coastline, very typical in the Atlantic).

The structure of the network fits on an MLP (multi-layer perceptron) architecture:

- 10 inputs, i.e., the size of the characteristic vector.
- 5 outputs, 1 for each class identified.
- 10 hidden neurons. This decision is the result of fine-tuning of the system.
- The activation functions are a hyperbolic tangent in the hidden layer and softmax in the output layer.

The training has been carried out with labeled points obtained from real images. The number of samples per class has been balanced by preserving the water samples (class of interest) and randomly eliminating samples from the majority classes. The training method has been "Backpropagation with conjugate gradient" [22] and the computer tool MATLAB [23].

The outcomes have been favorable for all classes, except foam, which is not a significant factor in this particular context, as shown in the confusion matrix (see Figure 7). In this work, 70% of the samples were used for training, 15% to validate and stop the training, and the remaining 15% for the final test. The total number of samples is greater than 19 million.

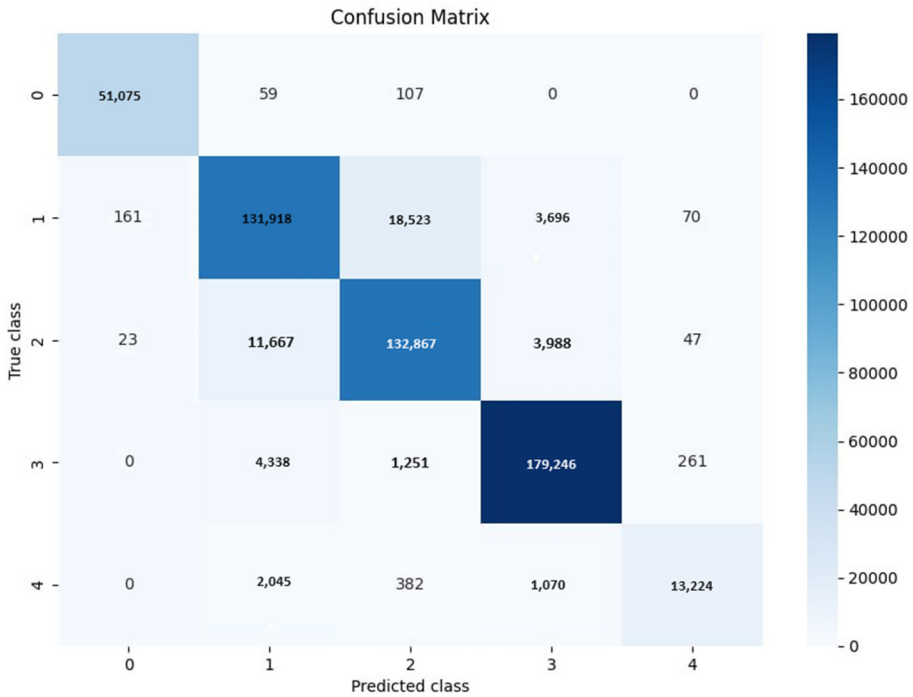


Figure 7. Confusion matrix (only for the 15% test samples).

In Figure 8, the result obtained for a sub-image containing the Vigo estuary is shown. Output 3 (water detection) of the neural network has been represented as a grayscale image. Values close to 1.0 (white in the image) mean positive water detection, and other values close to 0.0 (dark in the image) mean negative detection. This image is binarized with a high threshold (0.90) to obtain a water mask. This obtained mask is processed using mathematical morphology [24] to obtain a cleaner and more compact result.

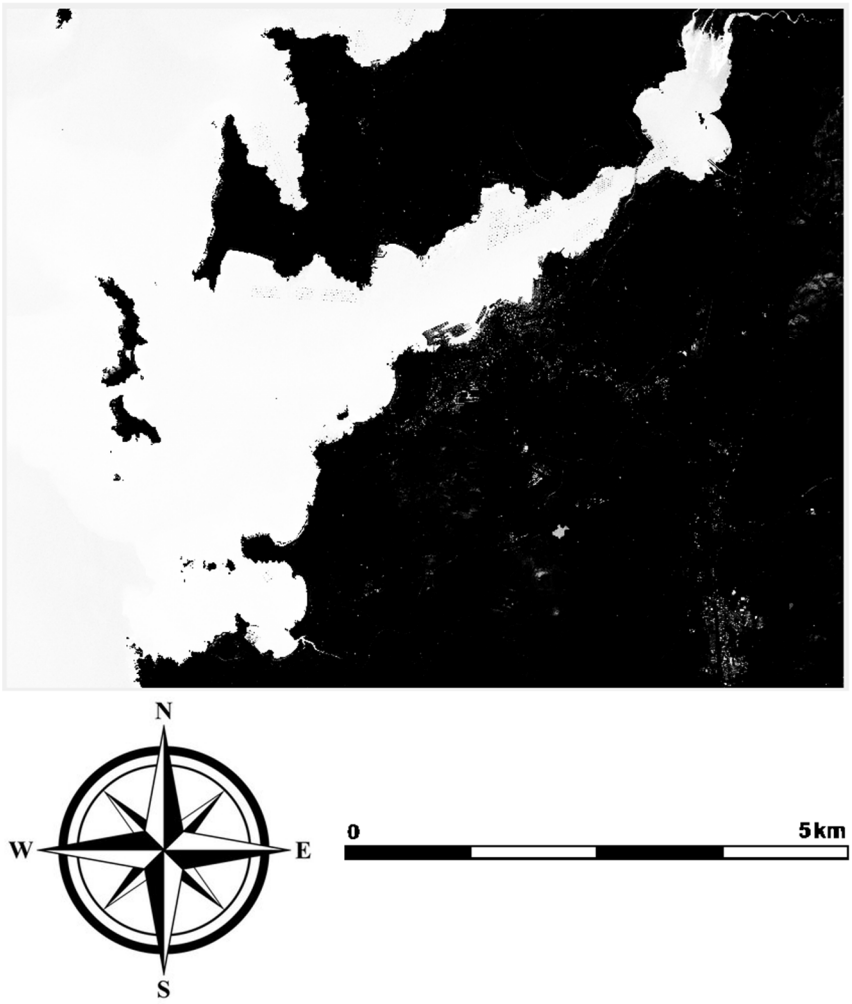


Figure 8. Detection with neural networks.

This process can be expressed in mathematical morphology terms:

1. Closing, erasing isolated non-water points.
2. Opening, erasing isolated water points.
3. Erosion, used to eliminate points very close to the coastline.

Note that these same operations are also performed with the mask obtained by the alternate method (NDWI).

2.4. Detection of Platforms

The next issue in this task addressed the classification of all the pixels previously detected as water, i.e., those that have a positive value in the masks obtained in the previous section. The result of this classifier will be binary: "platform" or "not platform". This classifier is based on extracting a feature vector for each pixel consisting of the different reflection values for each band (again excluding the "atmospheric" 60 m bands). These vectors will be used to train classifiers; we have tested three kinds of classifiers: a second neural network (also an MLP), a support vector machine (SVM), and a classifier called "Bootstrap Aggregated Trees" (Bagged Tree). The results obtained are treated as an image with a black background and white "active" points that define "connected components" (blobs) that are possible platforms. This image is processed with classical computer vision techniques in order to eliminate false positives that would reduce the final success rate.

In the next sections, classification methods will be described including a preliminary assessment of each based on accuracy and F_1-score (these are common figures of merit for classifiers and we provide a detailed description of them in the next section).

2.4.1. Neural Networks

For our new MLP, we have ten inputs again (ten bands of sufficient resolution) and a single output neuron (the value obtained will be close to 1.0 when we are detecting a platform). For this second case, we can use fewer neurons at the intermediate level: in particular, we have achieved training convergence with only two hidden neurons. The activation functions are a hyperbolic tangent (hidden layer) and a sigmoid (output layer).

As we can see in Figures 4 and 8, water masks usually present dark holes in the platform points. Obviously, this is a negative detection; that is, "it happens because those points are NOT water". When processing the mask, the closing operation makes those and other holes (holes due to boats or other floating objects) disappear. A morphological operation known as "Bottom Hat" (or negative "Top Hat") would allow us to obtain those points as "active" ones (white) on a black background:

$$\text{BottomHat(Im)} = \text{Close(Im)} - \text{Im} \tag{2}$$

That would not be a detection of enough reliability. Nevertheless, this method is used (with manual corrections) to find training samples.

The training has been carried out with the same method that we explained in the previous section. The total number of samples is 12,976. It has been based on 6488 samples (pixels) of platforms in sub-images of the estuaries of Pontevedra and Vigo. Afterward, the same number of water samples was obtained extracting random water points from the same images.

In Figure 9, the confusion matrix for this new network is presented where it is shown that the error rate is below 2%. Note that it is the matrix for the test set only.

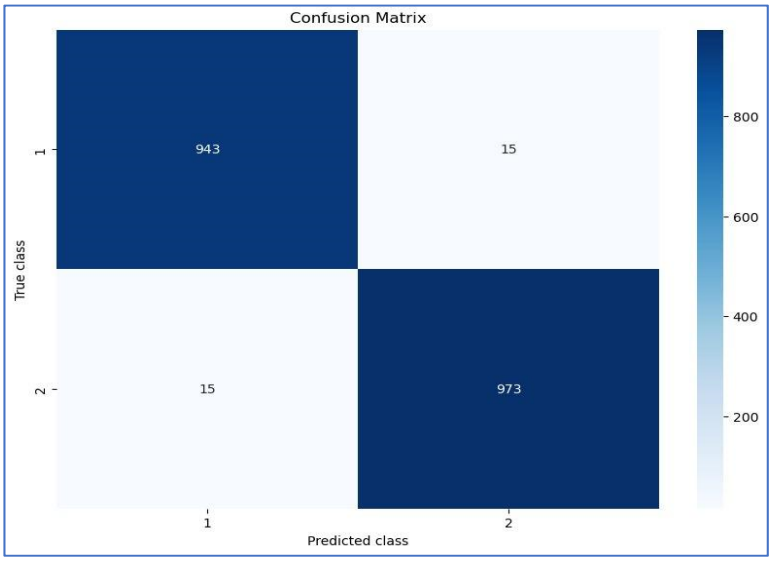


Figure 9. Confusion matrix (only for 15% of test samples).

With these data, the usual metrics of precision, recall, and, F_1-score can be computed. The obtained value for the F_1-score is 0.98, a result that is quite interesting (very close to 1.0). See Section 3 for a detailed description of F_1-score computation.

2.4.2. Support Vector Machines

A Support Vector Machine (SVM) [25] is a kind of classifier that is highly recommended for two-class problems like the one in this section. For an nth-dimensional vector space, SVM searches for a transformation into a higher dimensionality space where classes can be separated linearly so that a linear function can be used (a hyperplane) as a discriminant function.

Training an SVM with the same data used for the neural network, we obtain a 0.97% recognition error and an F_1-score of 0.99. At this point, it is even better than the former method (MLP).

2.4.3. Bagged Tree

Bootstrap Aggregated Trees (Bagged Tree [26]) belongs to a different category of classifiers. In this case, several decision trees are trained with random (and overlapping) subsets of the training set. The final decision is derived by running in parallel all the constructed trees and prorating the results.

We have trained a Bagged Tree with 5 trees obtaining a 0.57% recognition error, and an F_1-score of 0.99. This is an excellent result as well.

2.5. Post-Processing of Results

The results on other images of the same estuaries and also on other estuaries were good at least for some of the classifiers (see the next section); however, certain false positives were detected on other artificial structures. As an example (Figure 10), we see a false positive on a bridge in the estuary of Noia (besides the bridge, two ancient stone structures result in another, line-shaped, false positive blob).



(**a**) (**b**)

Figure 10. Example of a false positive (bridge of "Ría de Noia"). (**a**) Result of detection, (**b**) aerial in situ photo showing the structure.

These types of errors can be easily eliminated according to their irregular shape and their size much larger than a platform.

Therefore, the output of the classifier (only active on the water mask) is post-processed. For each connected object (blob), conditions are imposed on its geometry: "area less than a maximum", "equivalent diameter less than a maximum", "Euler number equal to 1" (Euler number: number of components minus number of holes), and "solidity greater than a minimum" (solidity: percentage of blob points versus the area of its convex hull). With this filtering, highly satisfactory results are obtained.

3. Results

Up to this point, the results obtained have been based on an "extracted dataset" (made from real samples from real images, but this is still not a real-world test). In these conditions, all of the machine-learning methods have yielded good results. The results on the dataset are summarized in Table 1. Note that these results were already mentioned in the previous section.

Table 1. Results obtained on a separate dataset.

Classifier	F_1-Score	Error Rate (%)
MLP	0.98	2.5%
SVM	0.99	1.0%
Bagged Tree	0.99	<1.0%

To assess the final results for all the methods, we have applied the study to Sentinel image clippings corresponding to the estuaries of Vigo (635 platforms), Pontevedra (321 platforms), Arousa (2307 platforms), Noia (126 platforms), and Corcubión. The latter does not contain any raft, so it has been added as a control input. For this final test, images from different dates have been used. As is required in supervised learning, the images used for training have not been included in the tests.

We have computed three easily interpretable parameters in all tests: P (precision), "the reliability of true detection for true cases"; R (recall), "the probability of the detection of true cases"; and F_1-score, "the harmonic mean of P and R". Obviously, the greater these quantities are, the better the performance we achieve. The strict definitions are as follows:

$$P = \frac{TrueDetections}{TrueDetections + FalseDetections}$$

$$R = \frac{TrueDetections}{TrueDetections + MissedDetections}$$

$$F_1 = \frac{2 \cdot P \cdot R}{P + R} \tag{3}$$

Note the process followed to obtain the results. For each clipping (10 bands clipped), a resulting image is computed. This means detecting the water body and running the machine-learning method on the obtained water mask. Each pixel will be classified as belonging to a raft or not. After applying post-processing (see Section 2.5), the resulting blobs are considered as detected platforms. In Figure 11, a real Sentinel clipping is represented with the detected blobs superimposed (in yellow). The blob image is compared with another blob image (the truth table). Blobs that overlap define a true positive; if a raft is detected that is not present in the truth table, it will be a false positive; false negatives are defined by blobs in the truth table that have not been detected. The truth table is created manually. The image obtained in Figure 5b is used as a preliminary version. As explained before, that image is obtained by detecting the holes in the water mask. Correcting manually this image means comparing with a magnified RGB image to detect missing platforms and delete false positives; the truth table is constructed.

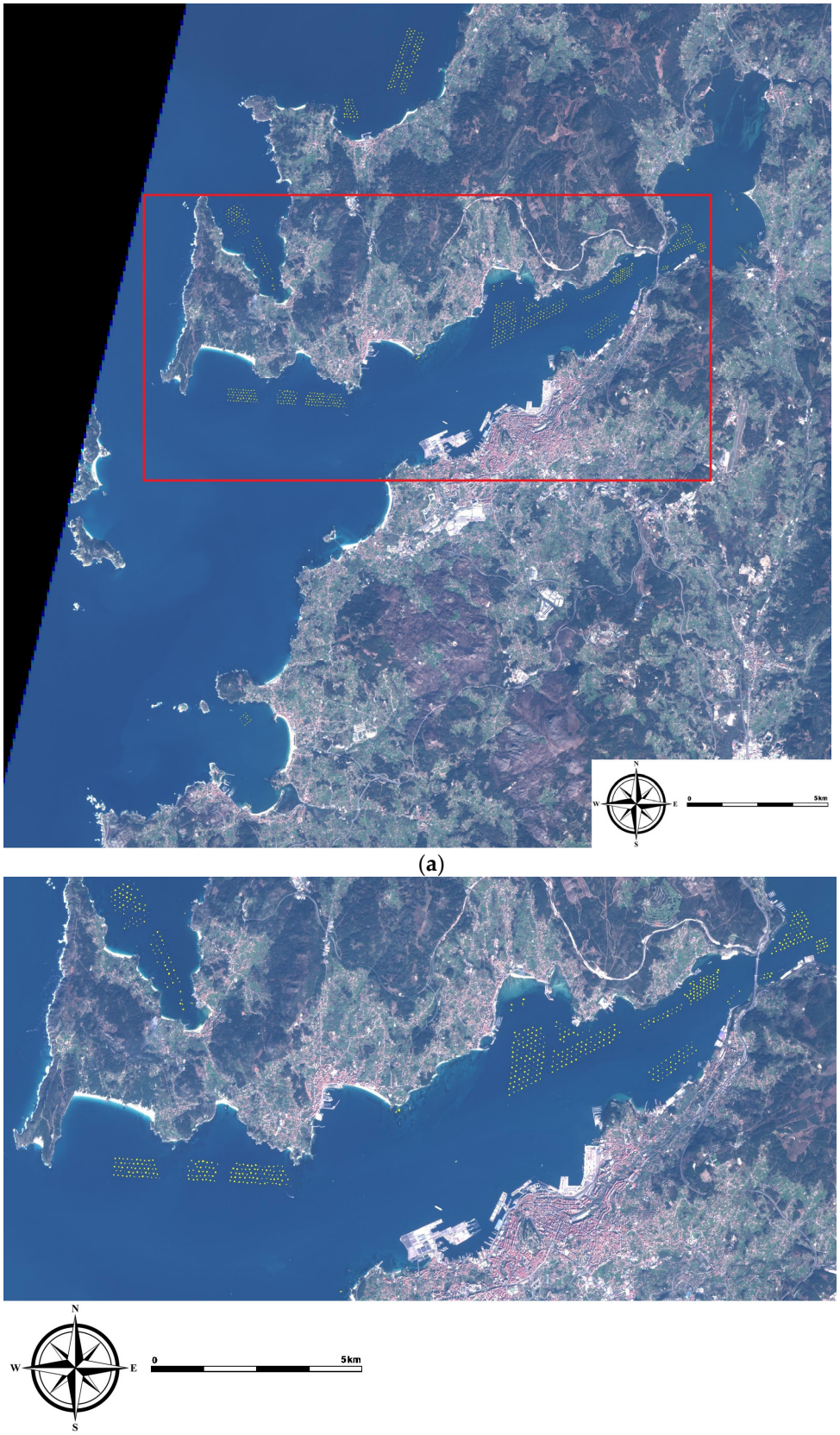
The results are summarized in Tables 2 and 3 (TOA corrected images). See Section 4 for appropriate discussion.

Table 2. TOA correction (without 60 M bands), NDWI for water detection.

Classifier	Precision	Recall	F_1-Score
MLP	0.9146	0.9918	0.9516
SVM	0.7003	0.2783	0.3983
Bagged Tree	0.7530	0.2631	0.3900

Table 3. TOA correction (without 60 M bands), MLP for water detection.

Classifier	Precision	Recall	F_1-Score
MLP	0.9103	0.9912	0.9490
SVM	0.7278	0.2775	0.4018
Bagged Tree	0.7587	0.2631	0.3907



(**a**)

(**b**)

Figure 11. Example of result (crop of Vigo estuary). (**a**) Entire area of study, (**b**) area of the rafts.

4. Discussion

Starting from the result tables above, it is seen that for final detection MLP outperforms the other two methods. In fact, SVM and Bagged Tree are revealed as merely unusable. We think this is due to less robustness against changes in atmospheric conditions that have not been properly normalized by the correction algorithms. SVM and Bagged Tree performed relatively well with conditions similar to those of training but the results went down when those conditions changed. See Figure 12 for an example of MLP use.

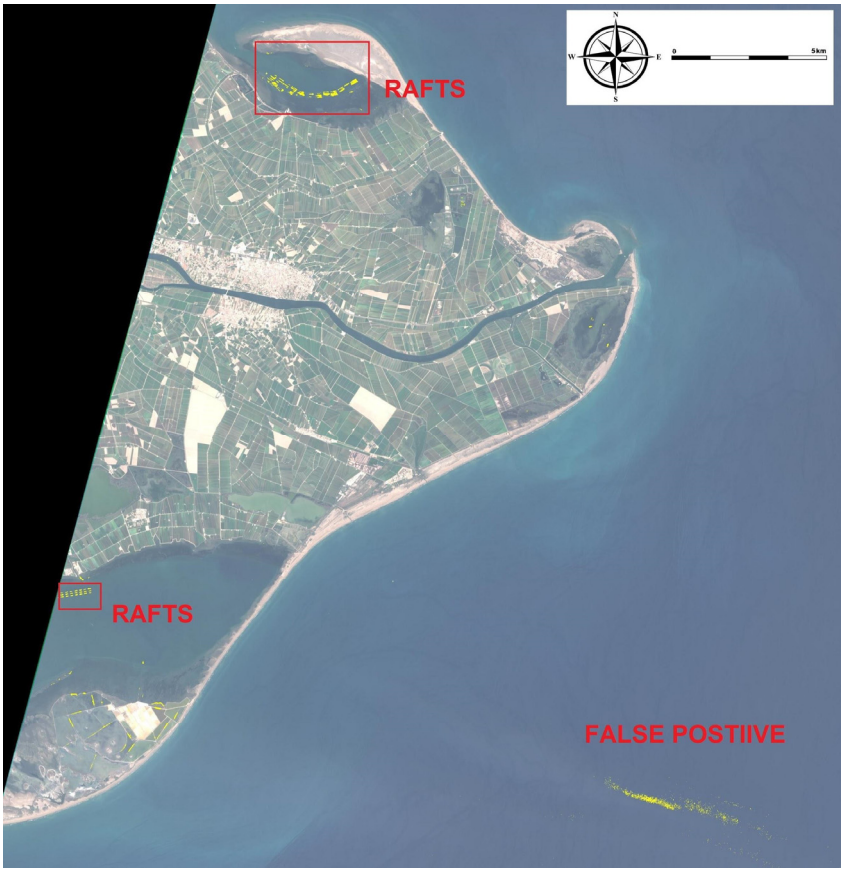


Figure 12. Platforms in the delta of the river Ebro.

The MLP still performs well with images of other geographical areas (like platforms in the Mediterranean at Tarragona, Spain); see Figure 13 where the same process described in Section 3 has been tested. Here, the platforms are larger and post-processing has been deactivated causing false positives.

From the results of Tables 2 and 3, the election for water detection is not so critical; nevertheless, NDWI is slightly better and easier to use.

For images with BOA correction, many false positives have been observed that are very difficult to eliminate, which, at least for the moment, makes this option a bad choice (Figure 13). Perhaps BOA correction is good for land applications but not so good for the marine part.

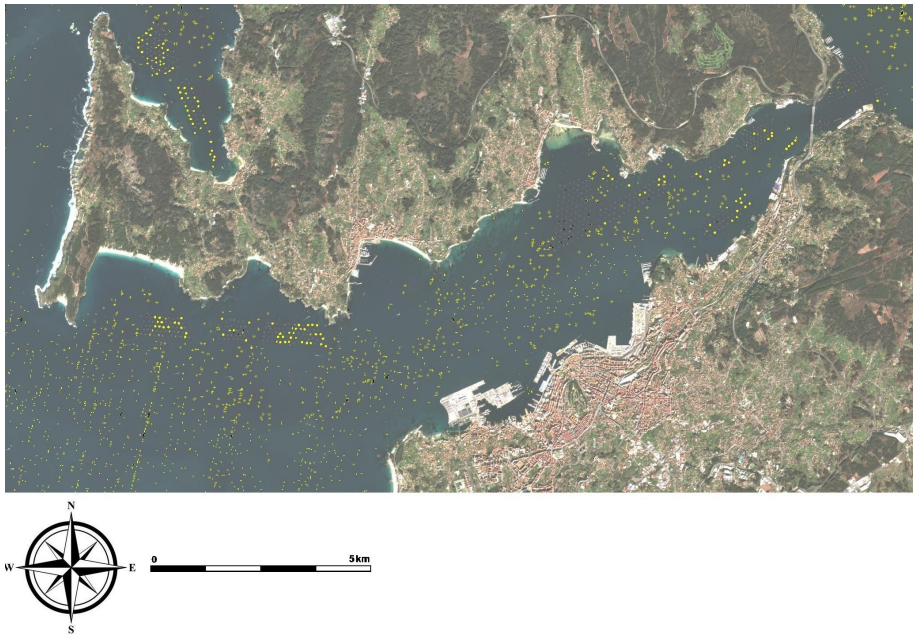


Figure 13. Use of BOA correction (Pontevedra estuary).

5. Conclusions

We have developed a method capable of locating the mussel platforms of the Galician estuaries (that can be used anywhere else), using Sentinel 2 images and MATLAB processing (which, of course, can be implemented on other platforms). This development can be used as the basis for creating an automated census application that produces results compatible with GIS environments. This final application could be used to monitor rafts' movements and activity, reinforcing the application of legal regulations.

For this particular problem, it seems better to use images with TOA correction (L1C) than those with BOA correction (L2A).

Between the two methods used to detect water bodies (NDWI and MLP), the results of Tables 2 and 3 recommend the NDWI-based method, although the difference in performance in this case is not very significant.

The possible use of traditional image processing techniques; for example, finding holes in the water mask can be useful for some tasks like creating truth tables but inaccuracies (mainly false positives) make this a weaker approach than machine-learning techniques.

For the detection of platforms within the water portions, the MLP is the most usable method (at least up to date).

As future lines, we would highlight the following:

- Process automation, implementing the model in an environment more suitable for an end-user application (C++ or Python), performing the automatic download and cropping of the images.
- Obtaining an output compatible with GIS tools, as nowadays the proposed application only obtains raft coordinates on the Sentinel images.
- Further study on the failure for SVM and Bagged Tree and research on other machine-learning techniques.
- Study of the reasons for the poor results with BOA correction.
- Explore the use of high-resolution SAR data (5 m $\times$ 5 m) from Sentinel 1.

Author Contributions: Initial concept, data acquisition, and software development: F.M.-R., J.M.S.-G., M.F.-B. and L.M.Á.-S. Writing and revising: F.M.-R., J.M.S.-G., M.F.-B. and L.M.Á.-S. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Sentinel data are publicly available through the ESA web services.

Acknowledgments: The authors wish to thank all personnel in the AtlantTIC research center for their support.

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A. Bands of Sentinel 2

This appendix describes the different bands of a Sentinel 2 image, showing a real example.

The Sentinel 2 sensor is a push broom scanner. This means that it obtains 1 pixel height images (lines) and full images have to be synthesized stacking individual lines. Each of these lines already has information on each of the frequency (or wavelength) bands. Note that the Sentinel is called a multi-spectral sensor because it receives information in multiple spectrum bands and those bands are chosen to be sources of different kinds of information.

The ESA server receives original unprocessed elementary images (lines) and they process information to create square images that comprise regions of 100 km $\times$ 100 km (for the biggest resolution of 10 m per pixel, this means 10,980 $\times$ 10,980, 100 Mpixel images). Images are corrected to take into account some atmospheric distortions and also for geometric distortions due to the Earth's curvature.

As an example, we will use a real Sentinel image of the northwest of the Iberian Peninsula. That means a part (south) of the Spanish region of Galicia and another part of northern Portugal. The top left corner of the images corresponds to the estuaries of Pontevedra and Vigo (part of this study), whereas the bottom left corner corresponds to the town of Esposende (Portugal) in the estuary of the river Cávado. Other major rivers present in the image are Miño and Limia/Lima.

Sentinel bands are ordered by wavelength. Band 1 (443 nm) is the only band in the ultraviolet region. It is called "coastal aerosol" and its function is helping to detect the atmosphere status. This is a "low resolution" band with 60 m per pixel (Figure A1). The cardinal points and scale of the following images are not presented as they are original Sentinel images.

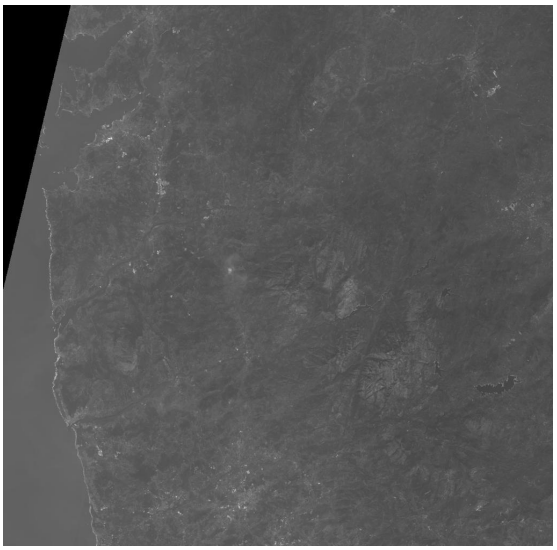


Figure A1. Band 1 (UV, coastal aerosol).

The black triangle in the top left corner is a region with no available information, i.e., no image was captured of this part in this satellite scan. Therefore, as the reader may note, an image with the same geographical limits (and with the same black triangle) will be available every five days.

Bands 2, 3, and 4 are the "visible bands", i.e., they include color information: blue, green, and red, in the order of the wavelength, opposite to the usual RGB order of almost all image formats. The aim of these images is to see visible objects. This is the more intuitive part of the information. The resolution of these bands is high (10 m per pixel). In Figure A2, an RGB composition of these bands is presented.

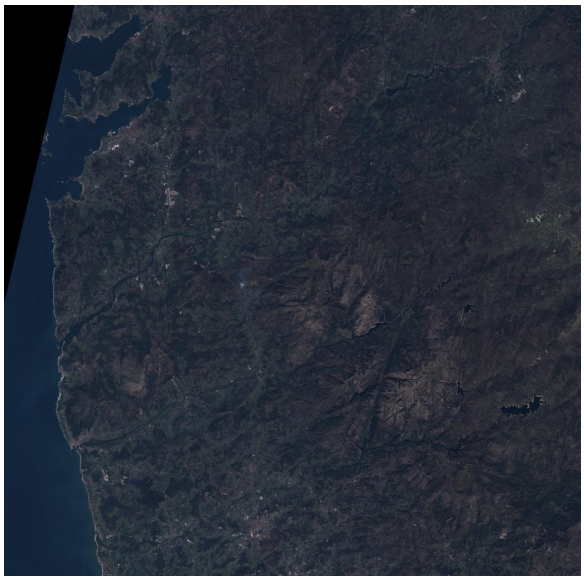


Figure A2. RGB composition of visible bands.

Bands 5, 6, and 7 are medium-resolution (20 m per pixel) bands called "Vegetation Red Edge" (Figure A3). These wavelengths (705, 740, and 783 nm) are in the frontier between the visible spectrum (red) and infrared. These bands play a main role in detecting the maturation of several cultivations.

Band 8 is an NIR band (near infrared) at 842 nm. It is of high resolution (10 m per pixel) due to its importance (Figure A4). NIR is important for detecting water (water bodies present no reflection in this wavelength) and also for detecting vegetation health (using the strong reflection in this wavelength on objects rich in chlorophyll). For vegetation, it is usual to use the index NDVI ($NDVI = \frac{NIR-RED}{NIR+RED}$), and for water the preferred index is NDWI ($NDWI = \frac{GREEN-NIR}{GREEN+NIR}$).

Band 8A is a medium-resolution (20 m band) that is called Vegetation Red Edge as with Bands 5, 6, and 7 (Figure A5). Despite being in the infrared region (865 nm), this band is also useful for detecting maturation of some vegetables.

Bands 9 and 10 are low-resolution (60 m), atmospheric bands; they are normally used for additional corrections (Figure A6). Band 9 is also called "water vapor" and Band 10 is called "Cirrus"; both will have mainly information about clouds. In fact, in Band 10, many times it is difficult to recognize the Earth region below.

Bands 11 and 12 are SWIR bands (short wave infrared), located at 1610 and 2190 nm (Figure A7). SWIR is still below thermal infrared. Its use is important for detecting human-built structures. The index NDBI is common for this application ($NDBI = \frac{SWIR-NIR}{SWIR+NIR}$).

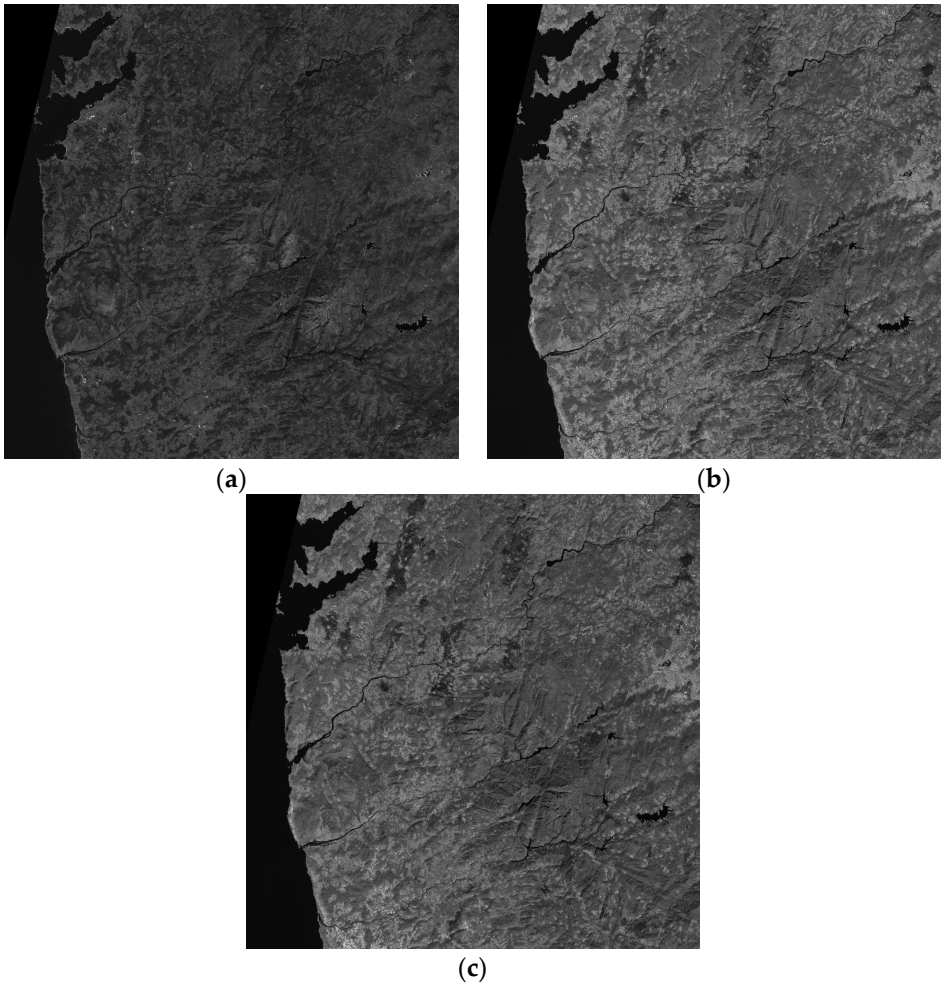


Figure A3. (**a**) 705 nm Red Edge, (**b**) 740 nm Red Edge, (**c**) 783 nm Red Edge.

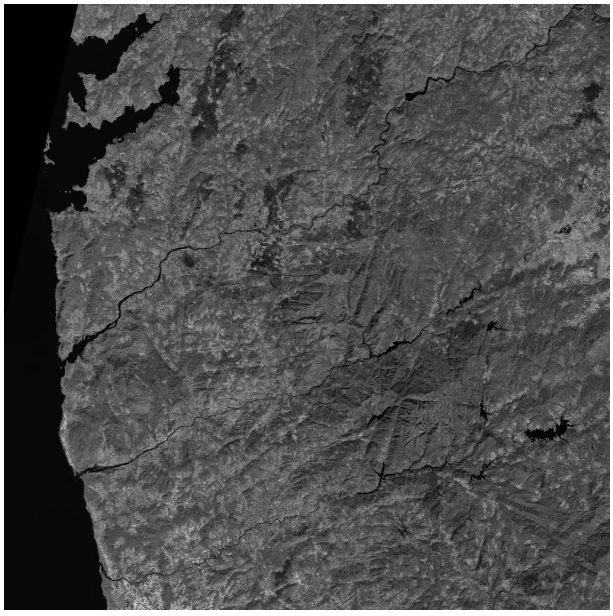


Figure A4. NIR.

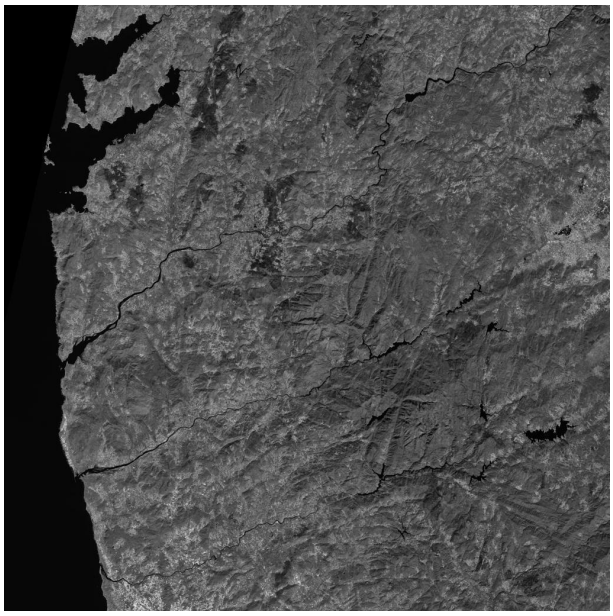
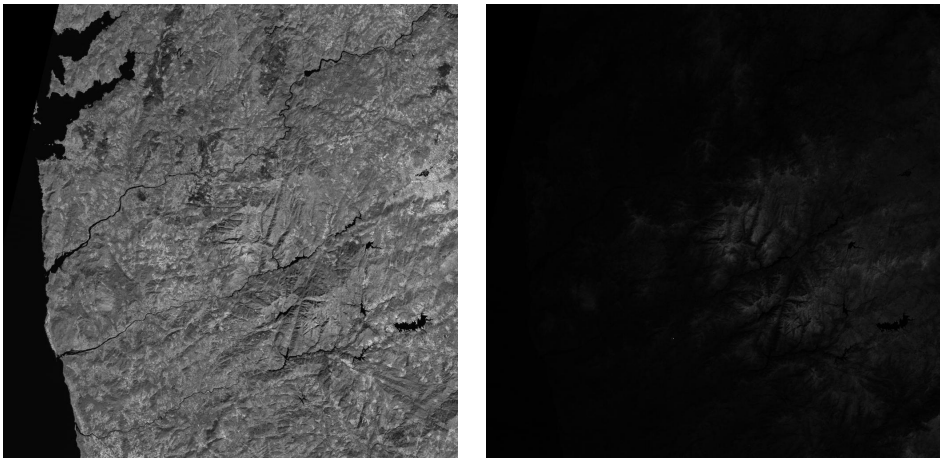
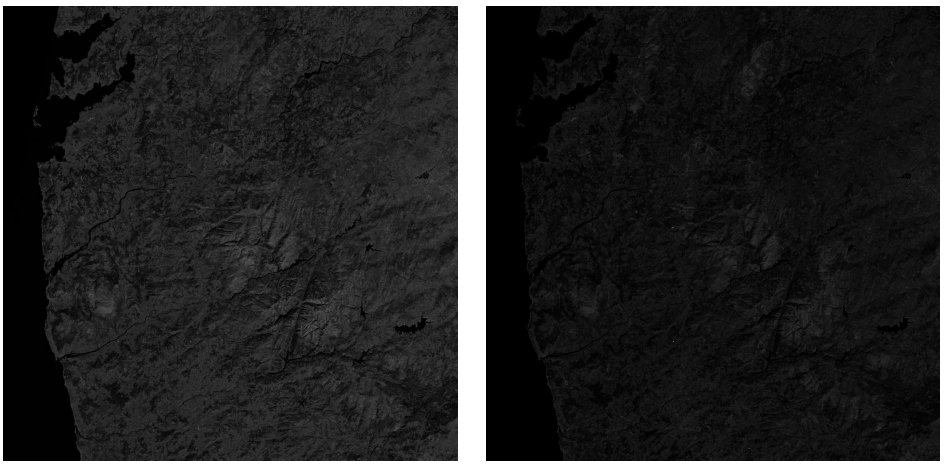


Figure A5. 865 nm Red Edge.



(**a**) (**b**)

Figure A6. (**a**) Water vapor, (**b**) Cirrus.



(**a**) (**b**)

Figure A7. (**a**) SWIR at 1610 nm, (**b**) SWIR at 2190 nm.

References

1. Available online: https://www.copernicus.eu/en/about-copernicus/infrastructure/discover-our-satellites (accessed on 1 January 2024).
2. Available online: https://scihub.copernicus.eu/ (accessed on 1 January 2024).
3. Available online: https://sentinel.esa.int/web/sentinel/user-guides/sentinel-2-msi/resolutions/spatial (accessed on 1 January 2024).
4. Illarionova, S.; Shadrin, D.; Ignatiev, V.; Shayakhmetov, S.; Trekin, A.; Oseledets, I. Estimation of the Canopy Height Model from Multispectral Satellite Imagery with Convolutional Neural Networks. *IEEE Access* **2022**, *10*, 34116–34132. [CrossRef]
5. Moghimi, A.; Welzel, M.; Celik, T.; Schlurmann, T. A Comparative Performance Analysis of Popular Deep Learning Models and Segment Anything Model (SAM) for River Water Segmentation in Close-Range Remote Sensing Imagery. *IEEE Access* **2024**, *12*, 52067–52085. [CrossRef]
6. Tahir, A.; Munawar, H.S.; Akram, J.; Adil, M.; Ali, S.; Kouzani, A.Z.; Mahmud, M.A.P. Automatic Target Detection from Satellite Imagery Using Machine Learning. *Sensors* **2022**, *22*, 1147. [CrossRef] [PubMed]
7. Marino, A.; Nunziata, F.; Vilas, L.G. Detecting aquaculture platforms using COSMO SkyMed. In Proceedings of the 13th European Conference on Synthetic Aperture Radar, EUSAR 2021, Online, 29 March–1 April 2021; pp. 1–6. Available online: https://ieeexplore.ieee.org/document/9472734 (accessed on 10 July 2024).
8. Murata, H.; Komatsu, T.; Yonezawa, C. Detection and discrimination of aquacultural facilities in Matsushima Bay, Japan, for integrated coastal zone management and marine spatial planning using full polarimetric L-band airborne synthetic aperture radar. *Int. J. Remote Sens.* **2019**, *40*, 5141–5157. [CrossRef]
9. Travaglia, C.; Profeti, G.; Aguilar-Manjarrez, J.; Lopez, N.; Marçal, A. *Mapping Coastal Aquaculture and Fisheries Structures by Satellite Imaging Radar. Case Study of the lingayen Gulf, the Philippines*; FAO Fisheries Technical Paper; FAO: Rome, Italy, 2004; ISBN 9251051143. Available online: https://www.fao.org/fishery/docs/DOCUMENT/gisfish/studycase11travagliaetal/studycase11travagliaetal_00.pdf (accessed on 10 July 2024).
10. Fan, J.; Chu, J.; Geng, J.; Zhang, F. Floating raft aquaculture information automatic extraction based on high resolution SAR images. In Proceedings of the 2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Milan, Italy, 26–31 July 2015; pp. 3898–3901. [CrossRef]
11. Kurekin, A.A.; Miller, P.I.; Avillanosa, A.L.; Sumeldan, J.D.C. Monitoring of Coastal Aquaculture Sites in the Philippines through Automated Time Series Analysis of Sentinel-1 SAR Images. *Remote Sens.* **2022**, *14*, 2862. [CrossRef]
12. Available online: https://developers.arcgis.com/python/samples/detecting-mussel-farms-using-deep-learning/ (accessed on 10 July 2024).
13. McMillan, C. Improving Buoy Detection with Deep Transfer Learning for Mussel Farm Automation. *arXiv* **2023**, arXiv:2308.09238v2. [CrossRef]
14. Zeng, D.; Liu, I.; Bi, Y.; Vennell, R.; Briscoe, D.; Xue, B.; Zhang, M. A new multi-object tracking pipeline based on computer vision techniques for mussel farms. *J. R. Soc. N. Z.* **2023**, 1–20. [CrossRef]
15. Available online: https://inta.es/INTA/es/blogs/copernicus/BlogEntry_1509095468013# (accessed on 1 January 2024).
16. Gao, B.-C. NDWI—A Normalized Difference Water Index for Remote Sensing of Vegetation Liquid Water from Space. *Remote Sens. Environ.* **1996**, *58*, 257–266. [CrossRef]
17. McFeeters, S.K. The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. *Int. J. Remote Sens.* **1996**, *17*, 1425–1432. [CrossRef]
18. Otsu, N. A threshold selection method from gray-level histograms. *IEEE Trans. Syst. Man Cybern.* **1979**, *9*, 62–66. [CrossRef]
19. Martín-Rodríguez, F.; Mojón-Ojea, O. *Big Plastic Masses Detection Using Satellite Images & Machine Learning*; Instrumentation Viewpoint; Upcommons (Universitat Politècnica de Catalunya): Barcelona, Spain, 2021; pp. 30–31.
20. Getreuer, P. Linear Methods for Image Interpolation. *Image Process. Line* **2011**, *1*, 238–259. [CrossRef]
21. Haykin, S. *Neural Networks: A Comprehensive Foundation*, 2nd ed.; Prentice Hall: Upper Saddle River, NJ, USA, 1998.
22. Johansson, E.; Dowla, F.; Goodman, D. Backpropagation Learning for Multilayer Feed-Forward Neural Networks Using the Conjugate Gradient Method. *Int. J. Neural Syst.* **1991**, *02*, 291–301. [CrossRef]
23. MATLAB. Available online: https://es.mathworks.com/products/matlab.html (accessed on 28 February 2023).
24. González, R.C.; Woods, R.E.; Eddins, S.L. *Digital Image Processing Using MATLAB*; Gatesmark Publishing, Cop.: Upper Saddle River, NJ, USA, 2009.
25. Noble, W.S. What is a support vector machine? *Nat. Biotechnol.* **2006**, *24*, 1565–1567. [CrossRef] [PubMed]
26. Hothorn, T.; Lausen, B. Bagging tree classifiers for laser scanning images: A data- and simulation-based strategy. *Artif. Intell. Med.* **2002**, *27*, 65–79. [CrossRef] [PubMed]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.