Semantic Segmentation of the Intertidal Zone of an Estuary—In Search of the Best Solution

Maria da Conceição Proença[1,2*], Ricardo Nogueira Mendes[1,3], Ricardo Melo[1,4]

[1]Marine and Environmental Sciences Centre (MARE-ULisboa) & ArNet, Lisbon, Portugal
[2]Department of Physics, Faculty of Sciences, University of Lisbon, Lisbon, Portugal
[3]Department of Animal Biology, Faculty of Sciences, University of Lisbon, Lisbon, Portugal
[4]Department of Plant Biology, Faculty of Sciences, University of Lisbon, Lisbon, Portugal
Email: *mcproenca@fc.ul.pt

Abstract

An aerial photographic coverage acquired on two consecutive days in October 2021 with a ground resolution of 20 cm and a spectral resolution of 4 bands (red, green, blue and near infrared), allowed to distinguish most of the classes of interest present in the intertidal zone of the Sado estuary. We explored the possibilities of thematic classification in the powerful and complex software ArcGIS Pro; we presented the methodology used in a detailed way that allows others with minimal knowledge of GIS to reproduce the classification process without having to decipher the specifics of the software. The classification implemented used ground truth from four classes related to the macro-occupations of the area. In a first phase we explore the standard algorithms with object-based capabilities, like K-Nearest Neighbor, Random Trees Forest and Support Vector Machine, and in a second phase we proceed to test three deep learning classifiers that provide semantic segmentation: a U-Net configuration, a Pyramid Scene Parsing Network and DeepLabV3. The resulting classifications were quantitatively evaluated with a set of 500 control points in a test area of 37,500 × 12,500 pixels, using confusion matrices and resorting to Cohen's kappa statistic and the concept of global accuracy, achieving a Kappa in the range [0.72, 0.81] and a global accuracy between 88.9% and 92.9%; the option U-Net had the most interesting results. This work establishes a methodology to provide a baseline for assessing future changes in the distribution of Sado estuarine habitats, which can be replicated in other wetland ecosystems for conservation and management purposes.

Keywords

Estuary, Intertidal Zone, ArcGIS Pro, Segmentation, Global Changes

1. Introduction

Estuarine areas play a crucial role in coastal ecosystems, being transitional zones where freshwater from rivers meets and mixes with saltwater from the open sea.

These areas are dynamic and experience tidal fluctuations, allowing for a free exchange of water between land and sea. Estuarine ecosystems are characterized by a mix of fresh and saltwater that provides abundant nutrients, making estuaries highly productive habitats to support a diverse range of species, including fish, invertebrates, and birds. In the case of river Sado, seagrass meadows and marshes found in nearshore estuarine and marine ecosystems contribute to this high productivity (Beck et al., 2001).

From the human and social perspective, governance of estuaries is a complex subject in Portugal (Fidélis & Carvalho, 2013), with multiple interests and multiple jurisdictions that do not contribute to a holistic approach of such a rich and fragile environment.

This work analyses the efficacy of several classification methods available in ArcGIS Pro for mapping estuarine habitats exposed to different conditions of tides, using aerial photographic imagery at 20 cm ground resolution. These flights were the result of a two-day flight plan, covering the area of interest as well as possible and considering the variations and heights of the tides, and the wave delay (Khojasteh et al, 2021).

Several authors have worked with deep learning methods and high-resolution images with similar goals; (Zhang et al., 2020) is a recent review of land cover classification and object detection approaches, in which traditional standard approaches are compared with deep learning models. The better performance of the latter is attributed to the simultaneous use of spectral and spatial information in object-based methods, while older approaches are based on pixel-by-pixel methods, which result in maps with the typical salt-and-pepper noise incorporated. In the last 20 years deep learning methods began to appear applied to land cover classification (Audebert et al., 2016; Huang et al., 2018; Kemker et al., 2018) with promising results and can be found in more applications in remote sensing and Earth sciences (Reichstein et al., 2019), mainly in land use and land cover (LULC) classification, to which (Vali et al., 2020) provides a complete framework. A Joint Deep Learning model (Zhang et al., 2019) provides novelty using spatial and hierarchical relationships between land cover probabilities and land use classifications, applied to an urban/suburban environment. The deep learning approach is so promising to handle large amounts of data in time series that large datasets such as EuroSAT are already publicly available for benchmarking (Helber et al., 2019), using Sentinel-2 images and 10 classes for LULC (27,000 georeferenced sub images at 10 m ground resolution in 13 spectral bands). More recently, the use of images from unoccupied aerial vehicles (UAVs) paired with deep learning algorithms (Gonzalez-Perez et al., 2022) has emerged as a tool with great potential for the study of coastal systems, both in terms of the results obtained and the associated costs (Durgan et al., 2020, Prentice et al.,

2021). The UAVs technology being stable for the last decade, they become a resourceful tool for coastal surveys (Turner et al., 2016, Liu et al., 2018).

Aerial photography seems the best solution for estuary mapping and monitorization, catching simultaneously the detail and the context (Bendell & Wan, 2011), with the advantage that nowadays the GIS software has built-in deep learning tools, although coverage is expensive and difficult to achieve at the best of times in terms of low tides and wave effects—perhaps these are the reasons that have made this approach rare.

The case study presented in this article focused on finding the most consistent methodology among those available in the software, to classify high-resolution images and produce thematic maps, which will form a reference base for monitoring the evolution of the most relevant habitats in the estuarine zone, allowing future assessments of the local ecosystem, as well as the identification of natural and anthropogenic changes that have occurred in the meantime.

2. Materials and Methods

2.1. Study Area

The study concerns 18,776 ha of the Sado estuary, located in the center of Portugal mainland (Figure 1), bounded by the line between estuarine bed and fringe that corresponds to the highest astronomical tide (Rilo et al., 2014). This is a region with some fieldwork carried out, so there was information available to be used as ground truth.
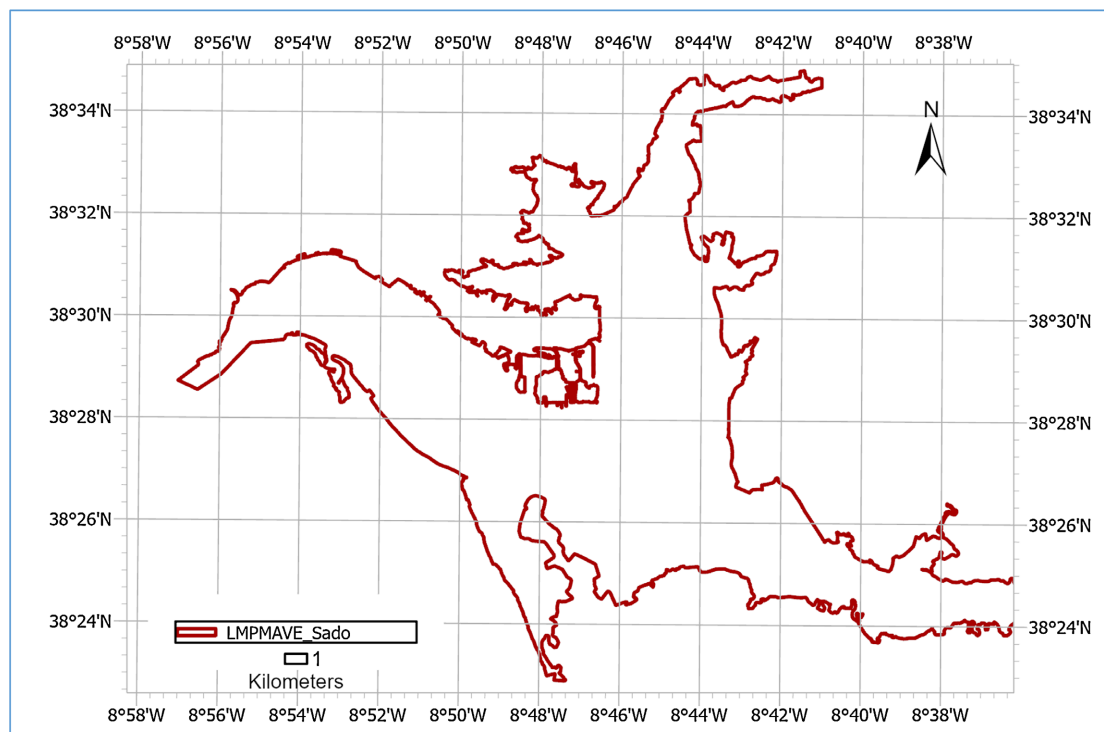

Figure 1. Sado estuarine area delimited by the "Linha de Máxima Preia-Mar de Águas Vivas Equinociais" (LMPMAVE), the limit corresponding to the equinoctial high tide maxima line.

The estuary circulation is driven mainly by the tides and the freshwater inputs of river Sado. The anthropogenic pressures in the region are spread differently by the estuarine margins. Northside includes a solid naval activity (with the 4th National Port), ship maintenance and repair industries, and a growing oyster farming that has, in many places, replaced the previous aquaculture farms that developed on traditional salt plants. The southside includes agriculture and forestry industries and significant tourism developments at the Troia península.

2.2. Materials

The image data set is in the form of orthoimages (ETRS 1989 TM06) acquired during a two-days aerial survey, 7 and 8 October 2021, at the best time to maximize the area observed. The images were acquired in 4 spectral bands, red (R), green (G), blue (B) and near infrared (NIR), with a ground resolution of 0.20 m. Geometric and radiometric corrections were previously made at the supplier's premises. The 41 images were mosaicked in larger tiles to be processed in a regular laptop, an HP Pavilion with Intel Core i7, RAM 16 GB, 512 GB disk, and an NVIDIA GeForce RTX 2060, with 6 GB.

2.3. Methods

The software explored is ArcGIS Pro, versão 3.0.2. Although complex and computationally heavy, it has several options for classification, covering the spectrum from standard machine learning methods such as K-Nearest Neigbors (KNN) (Cover & Hart, 1967), some more elaborated as Support Vector Machine (SVM) (Mountrakis et al., 2011), Random Tree Forest (RT) (Breiman, 2001) already using object based segmentation, to more recent deep learning (DL) approaches using Convolution Neural Networks (CNNs), such as U-Net (Ronneberger et al., 2015), PSP-Net (Zhao et al., 2017) and DeeplabV3 (Chen et al., 2016). All classifications were done within the official estuarine limits defined by the LMPMAVE.

Object based methods are well-suited for analysis of very high-resolution images, as its sequence of two phases (segmentation and classification) contributes to avoid the heterogeneity inherent to sub-meter pixels that could raise very noisy pixel-based classifications (Belgiu & Thomas, 2013). The segmentation aggregates semantically similar pixels in groups (segments) based on radiometric and geometric properties, and the object classification follow the rules of the supervised classification, allocating each segment to one of the pre-defined classes (Diesing et al., 2016; Lang et al., 2018).

The classes are defined during a train phase, common to all the algorithms used, based in areas known to belong to each class—it's the *ground truth*, from which all the parameters and models will be generated.

2.4. Pre-Processing Methodology

The first action to prepare the working images consisted in clipping the area of

interest (AOI) bounded by the line of maximum tide. The diversity of water bodies included in the estuary, in addition to the river itself, such as salt pans, active or abandoned, and aquaculture ponds with different degrees of filling, leads to a first segmentation to isolate land and water. A chlorophyll index calculated as the ratio between the NIR and green bands plus one made it possible to obtain a segmentation mask that only needs to be "cleaned"—a procedure that gives consolidation to large and thin areas and eliminates small, isolated spots with a few pixels (Figure 2).

The working area was isolated by application of this mask water/land to the four original bands (Figure 3).
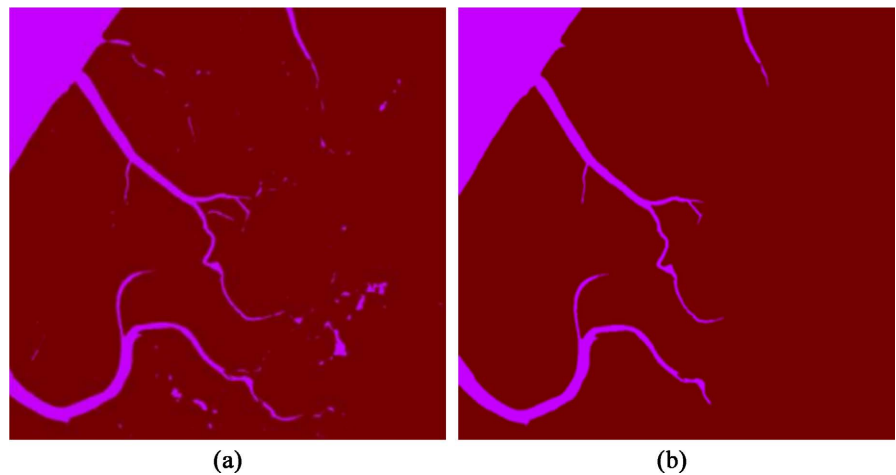

(a)   (b)

Figure 2. Water mask (a) before and (b) after a procedure to eliminate small spots and consolidate thin structures.
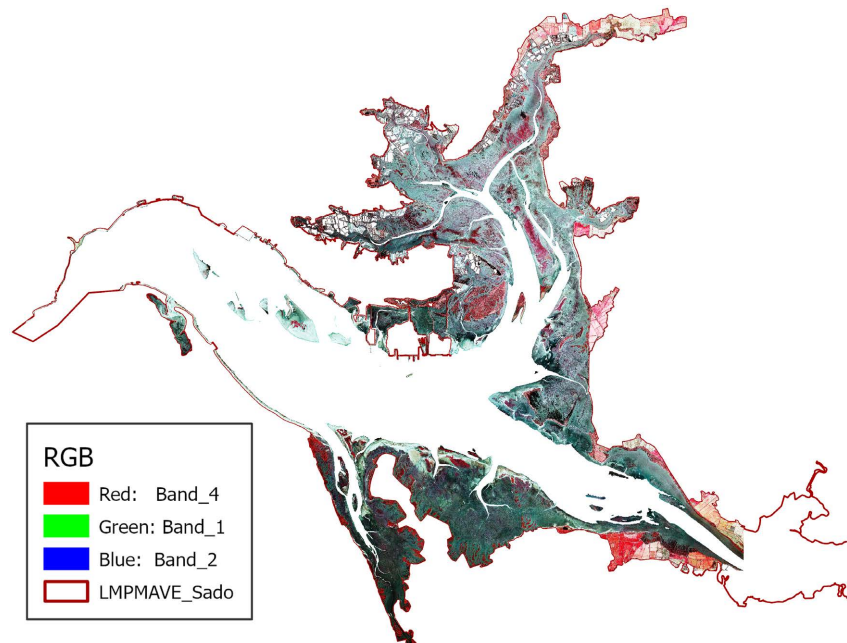

Figure 3. Area of work, circumscribed by the LMPMAVE and with the water zones removed, displayed in a combination of the NIR-R-G bands.

Object-based image analysis (OBIA) needs segmented images, which in Arc-GIS Pro are produced via the mean shift segmentation algorithm (Comaniciu & Meer, 2002), that requires three parameters, the first two referred as spatial and spectral details, consisting in a spatial radius and a radiometric range in a 0-20 scale, and the third being the minimum size accepted for each segment/object in pixels. The spatial detail concerns the distance from the analyzed pixel used to homogenize the neighborhood, the spectral detail defines the maximum distance allowed in radiometric space, and the third defines the minimum size of the final segments (Teodoro & Araujo, 2016). The introduction of segmented images makes it possible to reduce the local spectral variability inherent to high resolution, which can inevitably have various origins: shadows, different textures, terrain roughness, etc., and which strongly influences classification. **Figure 4** illustrates a segmentation of an area in the vicinity of a salt pan.
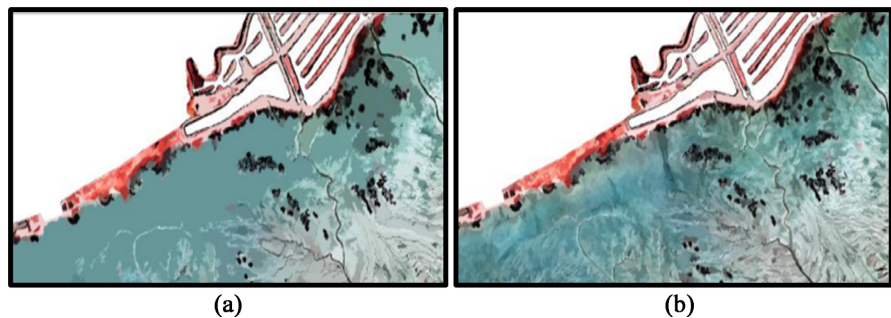

Figure 4. Segmented images at a detail level of (a) 14 and (b) 20 for both spectral and spatial detail. The range allowed is between 0 and 20, with the detail increasing as the parameter increases.

After some tests, the parameters for segmentation were chosen to be 16 for both spatial and spectral detail, and 2000 for the minimum size segment.

Object-based image analysis allows the use of six attributes computed from these segments: *Active chromaticity colour*, *Mean digital number*, *Standard deviation*, *Count of pixels*, *Compactness* and *Rectangularity*, the last one being more relevant in urban applications but with some positive influence whenever man-made structures are present, as it is the case.

2.5. Ground Truth

The delimitation of the ground truth areas was carried out with the *Training samples manager*, considering all the *a priori* information available and the experience gained from past fieldwork. The four classes considered reflects the macro-occupations characteristic of the estuarine zone, and are *Saltmarsh*, which represents vegetated areas dominated by halophytic plants that tolerate saltwater inundation, typically found in intertidal zones, *Seagrass*, aquatic vegetation found in permanently submerged areas, deeper than the intertidal zones, *Bare soil*, including sand covered areas and other bare surfaces inland, such as mudflats—expansive areas of fine sediment exposed during low tide, character-

ized by very low or no vegetation cover, and *Shallows*, encompassing all intertidal flat areas, with or without filamentous plants, that are alternately exposed and submerged by tidal action, often characterized by a mixture of sediment types and vegetation cover. The colour codes for the four classes are purple for Saltmarsh, pink for Seagrass, green for Bare soil and orange for Shallows.

3. Results and Discussion

A quantitative evaluation of results was carried out using 500 control points in the test area, with the overall kappa index (Cohen's Kappa statistic) and the overall accuracy (OA), both based on the confusion matrix, considered to be an indicator of the ability of the algorithm to identify all classes simultaneously. In short, *User's accuracy* concerns false positives or errors of commission: points incorrectly classified as belonging to one class when they belong to another. *Producer's accuracy* reflects false negatives or errors of omission: points in a class that have not been identified as such.

The Kappa statistic (Cohen, 1960) is a metric that provides an overall assessment of the accuracy of the classification, comparing it with a random classification. Another useful number is the global or overall accuracy, which indicates the percentage of well-identified points (sum of the diagonal of the confusion matrix) in the total number of control points used. Of the three standard algorithms tested, Random Trees provided the highest overall accuracy, with a Kappa of 0.811 and correctly classifying 92.9% of the control points (Table 1).

Table 1. Quantitative evaluation for the results obtained with SVM, RT and KNN algorithms.

Method	Cohen's Kappa	Global accuracy
SVM	0.739	90.3%
RT	0.811	92.9%
KNN	0.716	88.9%

The RT classifier offers the best results both quantitatively and qualitatively, by visual inspection (Figure 5(b)); it has the disadvantage of having a random component, which makes the realization of a good model more difficult to ensure because it is not just a function of the chosen parameters.

Three deep-learning models have results with enough quality to be explored, and among the results obtained with this type of image, the U-Net model showed slightly better results (Table 2).

The segmented images resulting from the three deep-learning options are compared in Figure 6. The lower resolution of DLabV3 is obvious (Figure 6(b)), although version 3 of the algorithm has already been mentioned as an improvement in this area (Li & Dong, 2022).

The results with the U-Net model are clearly more homogeneous and with more precise contours (Figure 6(d)), showing slightly better quantitative results
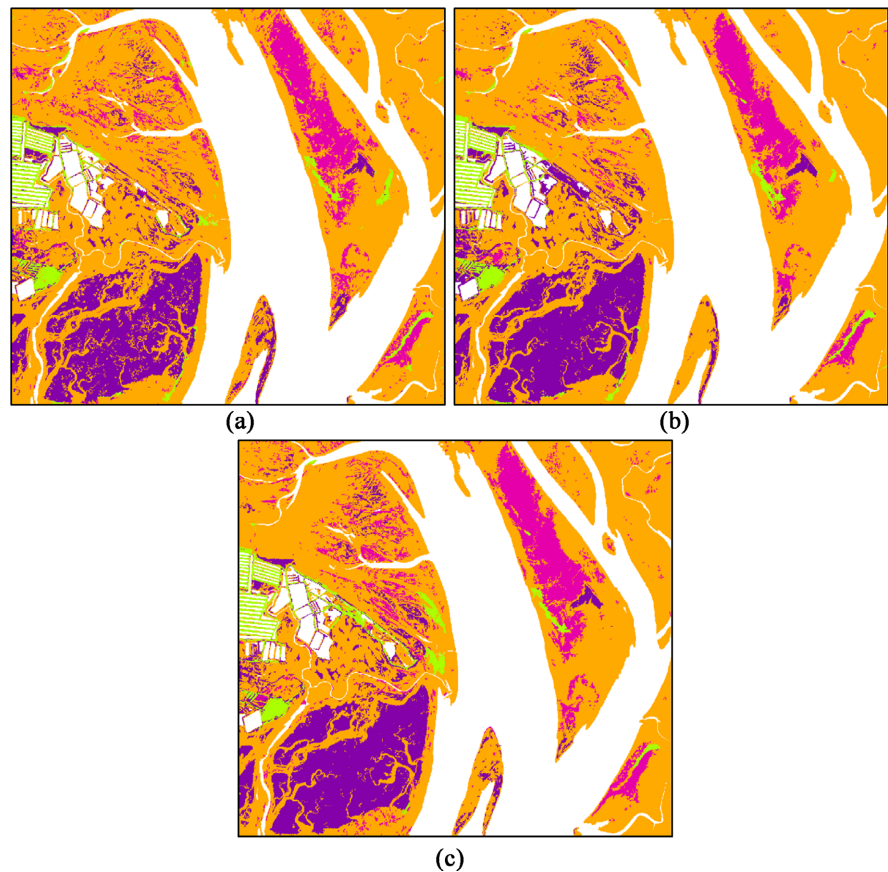
Figure 5. Classification of an area in the test image into 4 classes with different methods and different parameterizations: (a) K-Nearest Neighbours, considering 8 neighbours, with the segmented level 16 and 4 attributes (Kappa = 0.716, 88.9%), (b) Random Trees, same segmented level and 6 attributes (Kappa = 0.811, 92.9%) and (c) Support Vector Machine also with the segmented level 16 and 6 attributes (Kappa = 0.739, 90.3%).

Table 2. Quantitative evaluation of the results obtained with the models U-Net, PSPnet and DLabV3.

Method	Cohen's Kappa	Global accuracy
U-Net	0.781	91.5%
PSPnet	0.722	89.5%
DLabV3	0.716	90.1%

than the others (Table 2); it's also more accurate and coherent with the photo-interpretation of the image in the reference area, with the Bare soil areas being observed in the expected configuration, as well as a more correct identification of the boundaries of Seagrass patches available as ground truth (Figure 7).

In the laptop described in 2.2, processing times ranged from around 1.5 hours to extract the data using the ground truth previously defined, 2 to 5 hours to train the model, depending on the model chosen, and 3 to 7 hours to classify each image block, depending on the model chosen and the size of the image.

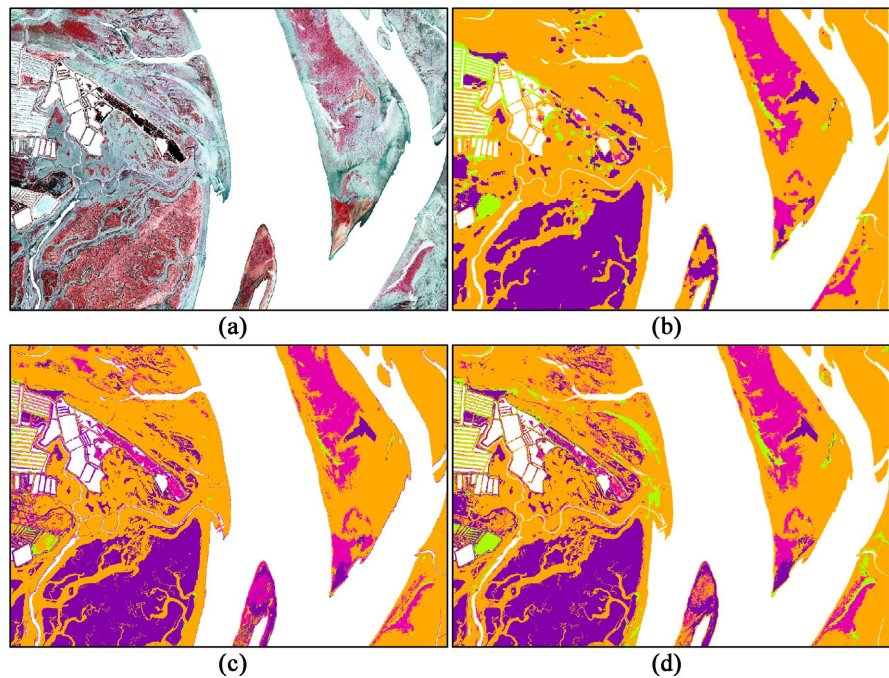As no other similar approach was found with this kind of data, we can only

Figure 6. Detail of deep learning classifications compared to (a) the original multispectral image of the area; models (b) DLabV3, (c) PSPnet and (d) U-Net.
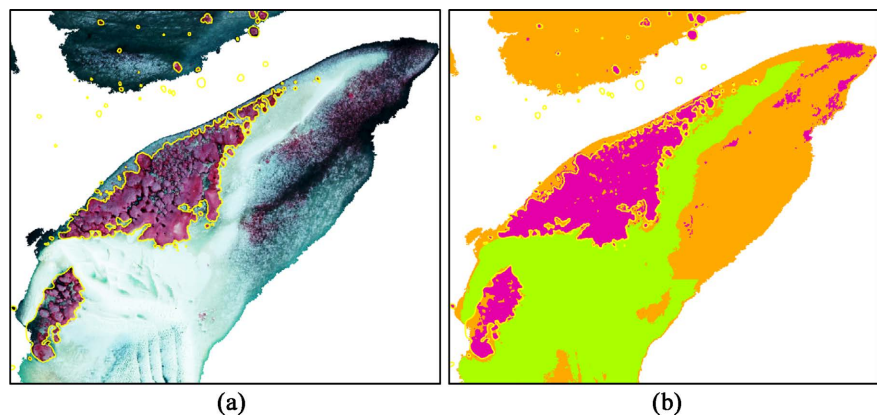

Figure 7. Detail of deep learning classification of a Seagrass patch: (a) multispectral image with ground truth contoured in yellow and (b) U-Net result, with backbone ResNet-34.

discuss the results against each other, as presented above. The U-Net model gives high-quality results with aerial photography, making all the post-processing steps previously required by conventional classifications unnecessary; ArcGIS Pro provides all the tools, with a learning curve feasible for a user with some background in classification methods, without the necessity of the informatic means and skills to implement complex deep learning procedures.

4. Conclusion

From the results illustrated and many others that we have explored with less success, but which have also contributed to guide the choice of the parameters of

the models tested, we conclude that the deep learning results obtained with the U-Net model with ResNet-34 as the backbone are superior to the standard machine learning methods for this type of high-resolution multispectral images, using the four bands R, G, B and Near-infrared. More compact patches and better definitions of contours are obtained in the most intricated areas, and fine elements are preserved and correctly identified. Even in the definition of Salt-marsh patches, which is a class correctly identified in general in both options, U-Net's performance is superior, simultaneously presenting well-defined contours and homogeneous patches, and managing to identify even the most problematic areas, such as the presence of salt marsh on the walls separating the tanks from the salt pans. Like (Gonzalez-Perez et al., 2022) we use machine learning algorithms and deep learning models trained with the same training set and tested with the same control points and we found a clear advantage of the U-Net model in the classification of the estuarine zone under study.

The main disadvantage is the processing time involved, but we were working with a fairly large area (18,776 ha) with a high resolution (0.2 m) on a laptop with a normal configuration, so this could probably be improved with an upgraded working configuration. The learning curve is variable but is quickly mastered with some method and a prior knowledge of supervised classification that facilitates familiarization with the various requirements of the models and the successive steps needed to complete the procedures.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

Audebert, N., Saux, B. L., & Lefevre, S. (2016). *Semantic Segmentation of Earth Observation Data Using Multimodal and Multi-Scale Deep Networks*.
https://arxiv.org/abs/1609.06846.33

Beck, M. W., Heck, K. L., Able, K. W., Childers, D. L., Eggleston, D. B., Gillanders, B. M. et al. (2001). The Identification, Conservation, and Management of Estuarine and Marine Nurseries for Fish and Invertebrates: A Better Understanding of the Habitats That Serve as Nurseries for Marine Species and the Factors That Create Site-Specific Variability in Nursery Quality Will Improve Conservation and Management of These

Areas. *BioScience, 51,* 633-641.
https://doi.org/10.1641/0006-3568(2001)051[0633:ticamo]2.0.co;2

Belgiu, M., & Thomas, J. (2013). Ontology Based Interpretation of Very High Resolution Imageries-Grounding Ontologies on Visual Interpretation Keys. In *16th AGILE Conference on Geographic Information Science* (pp. 1-5). Association of Geographic Information Laboratories for Europe (AGILE).

Bendell, L. I., & Wan, P. C. Y. (2011). Application of Aerial Photography in Combination with GIS for Coastal Management at Small Spatial Scales: A Case Study of Shellfish Aquaculture. *Journal of Coastal Conservation, 15,* 417-431.
https://doi.org/10.1007/s11852-010-0101-8

Breiman, L. (2001). Random Forests. *Machine Learning, 45,* 5-32.
https://doi.org/10.1023/a:1010933404324

Chen, L., Papandreou, G., Kokkinos, I., Murphy, K., & Yuille, A. L. (2018). Deeplab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFS. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 40,* 834-848. https://doi.org/10.1109/tpami.2017.2699184

Cohen, J. (1960). A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement, 20,* 37-46. https://doi.org/10.1177/001316446002000104

Comaniciu, D., & Meer, P. (2002). Mean Shift: A Robust Approach toward Feature Space Analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 24,* 603-619.
https://doi.org/10.1109/34.1000236

Cover, T., & Hart, P. (1967). Nearest Neighbor Pattern Classification. *IEEE Transactions on Information Theory, 13,* 21-27. https://doi.org/10.1109/tit.1967.1053964

Diesing, M., Mitchell, P., & Stephens, D. (2016). Image-Based Seabed Classification: What Can We Learn from Terrestrial Remote Sensing? *ICES Journal of Marine Science: Journal du Conseil, 73,* 2425-2441. https://doi.org/10.1093/icesjms/fsw118

Durgan, S. D., Zhang, C., Duecaster, A., Fourney, F., & Su, H. (2020). Unmanned Aircraft System Photogrammetry for Mapping Diverse Vegetation Species in a Heterogeneous Coastal Wetland. *Wetlands, 40,* 2621-2633.
https://doi.org/10.1007/s13157-020-01373-7

Fidélis, T., & Carvalho, T. (2013). Estuary Planning and Management in Portugal—Sharing Local Boundaries for Water Resources Management. In *Proceedings of the TWAM2013 International Conference & Workshops* (pp. 1-5). Springer Nature.
https://api.semanticscholar.org/CorpusID:199340944

Gonzalez-Perez, A., Abd-Elrahman, A., Wilkinson, B., Johnson, D. J., & Carthy, R. R. (2022). Deep and Machine Learning Image Classification of Coastal Wetlands Using Unpiloted Aircraft System Multispectral Images and Lidar Datasets. *Remote Sensing, 14,* Article No. 3937. https://doi.org/10.3390/rs14163937

Helber, P., Bischke, B., Dengel, A., & Borth, D. (2019). Eurosat: A Novel Dataset and Deep Learning Benchmark for Land Use and Land Cover Classification. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 12,* 2217-2226.
https://doi.org/10.1109/jstars.2019.2918242

Huang, B., Zhao, B., & Song, Y. (2018). Urban Land-Use Mapping Using a Deep Convolutional Neural Network with High Spatial Resolution Multispectral Remote Sensing Imagery. *Remote Sensing of Environment, 214,* 73-86.
https://doi.org/10.1016/j.rse.2018.04.050

Kemker, R., Salvaggio, C., & Kanan, C. (2018). Algorithms for Semantic Segmentation of Multispectral Remote Sensing Imagery Using Deep Learning. *ISPRS Journal of Photo-*

grammetry and Remote Sensing, 145, 60-77.
https://doi.org/10.1016/j.isprsjprs.2018.04.014

Khojasteh, D., Glamore, W., Heimhuber, V., & Felder, S. (2021). Sea Level Rise Impacts on Estuarine Dynamics: A Review. *Science of the Total Environment, 780,* Article ID: 146470. https://doi.org/10.1016/j.scitotenv.2021.146470

Lang, S., Baraldi, A., Tiede, D., Hay, G., & Blaschke, T. (2018). Towards a (GE). OBIA 2.0 Manifesto-Achievements and Open Challenges in Information & Knowledge Extraction from Big Earth Data. In *Proceedings of the GEOBIA* (pp. 1-15). Springer Nature.

Li, Z., & Dong, J. (2022). A Framework Integrating Deeplabv3+, Transfer Learning, Active Learning, and Incremental Learning for Mapping Building Footprints. *Remote Sensing, 14,* Article No. 4738. https://doi.org/10.3390/rs14194738

Liu, T., Abd-Elrahman, A., Morton, J., & Wilhelm, V. L. (2018). Comparing Fully Convolutional Networks, Random Forest, Support Vector Machine, and Patch-Based Deep Convolutional Neural Networks for Object-Based Wetland Mapping Using Images from Small Unmanned Aircraft System. *GIScience & Remote Sensing, 55,* 243-264. https://doi.org/10.1080/15481603.2018.1426091

Martínez Prentice, R., Villoslada Peciña, M., Ward, R. D., Bergamo, T. F., Joyce, C. B., & Sepp, K. (2021). Machine Learning Classification and Accuracy Assessment from High-Resolution Images of Coastal Wetlands. *Remote Sensing, 13,* Article No. 3669. https://doi.org/10.3390/rs13183669

Mountrakis, G., Im, J., & Ogole, C. (2011). Support Vector Machines in Remote Sensing: A Review. *ISPRS Journal of Photogrammetry and Remote Sensing, 66,* 247-259. https://doi.org/10.1016/j.isprsjprs.2010.11.001

Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N., & Prabhat (2019) Deep Learning and Process Understanding for Data-Driven Earth System Science. *Nature, 566,* 195-204. https://doi.org/10.1038/s41586-019-0912-1

Rilo, A., Freire, P., Mendes, R. N., Ceia, R., Catalão, J., Taborda, R. et al. (2014). Metodologia para o traçado da Linha de Máxima Preia-Mar de Águas Vivas Equinociais em ambientes de transição: Aplicação ao estuário do Tejo (Portugal). *Revista de Gestão Costeira Integrada, 14,* 95-107. https://doi.org/10.5894/rgci450

Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. In *MICCAI 2015, Part III, LNCS 9351* (pp. 234-241). Springer. https://arxiv.org/abs/1505.04597

Teodoro, A. C., & Araujo, R. (2016). Comparison of Performance of Object-Based Image Analysis Techniques Available in Open Source Software (Spring and Orfeo Toolbox/Monteverdi) Considering Very High Spatial Resolution Data. *Journal of Applied Remote Sensing, 10,* Article ID: 016011. https://doi.org/10.1117/1.jrs.10.016011

Turner, I. L., Harley, M. D., & Drummond, C. D. (2016). Uavs for Coastal Surveying. *Coastal Engineering, 114,* 19-24. https://doi.org/10.1016/j.coastaleng.2016.03.011

Vali, A., Comai, S., & Matteucci, M. (2020). Deep Learning for Land Use and Land Cover Classification Based on Hyperspectral and Multispectral Earth Observation Data: A Review. *Remote Sensing, 12,* Article No. 2495. https://doi.org/10.3390/rs12152495

Zhang, C., Sargent, I., Pan, X., Li, H., Gardiner, A., Hare, J. et al. (2019). Joint Deep Learning for Land Cover and Land Use Classification. *Remote Sensing of Environment, 221,* 173-187. https://doi.org/10.1016/j.rse.2018.11.014

Zhang, X., Han, L., Han, L., & Zhu, L. (2020). How Well Do Deep Learning-Based Methods for Land Cover Classification and Object Detection Perform on High Resolution Remote Sensing Imagery? *Remote Sensing, 12,* Article No. 417.

https://doi.org/10.3390/rs12030417

Zhao, H., Shi, J., Qi, X., Wang, X., & Jia, J. (2017). Pyramid Scene Parsing Network. In *2017 IEEE Conference on Computer Vision and Pattern Recognition* (pp. 2881-2890). Institute of Electrical and Electronics Engineers. https://doi.org/10.1109/cvpr.2017.660