S. Cope, V. Zetterlind and B. Tougher, "Integrating marine radar in a multi-sensor platform for remote, unsupervised vessel tracking in the nearshore environment," *Global Oceans 2020: Singapore – U.S. Gulf Coast*, 2020, pp. 1-5, doi: 10.1109/IEEECONF38699.2020.9389460.
https://ieeexplore.ieee.org/document/9389460

Integrating marine radar in a multi-sensor platform for remote, unsupervised vessel tracking in the nearshore environment

Samantha Cope
ProtectedSeas
Anthropocene Institute
Palo Alto, USA
samantha@anthinst.org

Virgil Zetterlind
ProtectedSeas
Anthropocene Institute
Palo Alto, USA
virgil@anthinst.org

Brendan Tougher
ProtectedSeas
Anthropocene Institute
Palo Alto, USA
brendan@anthinst.org

***Abstract*—The Marine Monitor (M2) is a surveillance tool that integrates data from marine radar, a high-definition camera, and the Automatic Identification System (AIS) to monitor all vessel activity in nearshore areas regardless of participation in vessel monitoring systems. To overcome the inherent issue of false targets in radar data, the M2 system utilizes feature selection and machine learning to predict the likelihood that a radar-detected trajectory is that of a true vessel. By classifying greater than 90% of records accurately, the M2 system provides live situational awareness and a robust historical record of true vessel activity.**

***Index Terms*—sensor systems, radar applications, machine learning, environmental management**

I. Introduction

Marine protected areas (MPAs) conserve ecosystems by restricting activities within their boundaries and are a key component of global efforts to improve ocean health. Vessel-related resource exploitation, such as fishing, is a primary stressor in marine habitats, and successful enforcement of fishing regulations is required to achieve the desired ecological benefits of MPAs, which include fish stock replenishment, habitat protection, and increased biodiversity [1].

Tracking vessel activity helps MPA managers evaluate compliance with restrictions and enforce regulations [2], but there are challenges in practice. Remote locations can be difficult to monitor due to physical distance from human resources [3] while still experiencing fishing pressure when commercially valuable species are present [4]. Illegal fishing activity may also be purposefully clandestine [5]. Ultimately, limited budget capacity is a barrier to any monitoring effort and prevents effective management of many MPAs [6] especially in developing countries [7].

Vessel monitoring, typically done via on-water patrolling, can be supplemented by cooperative vessel tracking technologies like the Automatic Identification System (AIS) and vessel monitoring systems (VMS) which provide records of vessel identification and position over time. These technologies have been widely used to analyze spatio-temporal patterns of commercial fishing effort [8], [9], but less than 5% of the global fishing fleet transmits these data [10]. Further, small-scale fishing and recreational vessels that are not mandated by law to transmit AIS data also participate in illegal fishing activity [11].

Given these obstacles, a cost-effective method for integrating multiple tracking technologies is required to capture the full range of vessel activities in many MPAs [12]. Marine radar has been used to supplement AIS data in nearshore environments where recreational and other small vessels are common [13]. As a traditional tool in maritime surveillance, radar tracking provides an unbiased, non-cooperative method for tracking vessels ideal for independent monitoring [14].

The ProtectedSeas Marine Monitor (M2) is a shore-based vessel monitoring platform that integrates commercial off the shelf (COTS) marine radar paired with a high-definition camera and AIS receiver to autonomously and continuously capture all vessel activity regardless of cooperation [15]. Assisted by an Automatic Radar Plotting Aid (ARPA), vessels within five nautical miles of the system are detected using radar, and the camera is directed to their positions for capturing photos. All sensors are collocated in a single platform, and solar panels and a mobile trailer give the newest Mobile Marine Monitor (M3) system complete off-grid capability (Fig. 1).

Fig. 1. Mobile Marine Monitor (M3) equipped with marine radar system, high-definition camera, AIS receiver, solar panels, and mobile trailer monitors vessel traffic.

Cloud-based data management provides both live and historical situational awareness to M2 users which can be especially valuable for management of MPAs in remote locations. M2 software captures, processes, and integrates all tracking data in a web-based user interface, the M2 Viewer (Fig. 2), where users access vessel information, trajectories, and photos. Geofencing around areas of concern, such as MPA boundaries, alerts users via text message or email when trigger activities occur, and a mobile-friendly version of the viewer is also available. M2 can facilitate data-driven patrol efforts and more targeted use of resources.

Management with limited budget capacity benefits from the low cost of the M2 system compared to similar solutions. More complex radar systems used to monitor MPAs have a cost on the order of millions [16], [17], but the COTS radar keeps the cost of M2 at less than \$100,000 for hardware, software, and deployment with \$2,500 to \$5,000/year for ongoing maintenance and support [15]. M2 systems have been deployed to over twenty locations in five countries.

While M2 solves a number of issues that hinder effective management of MPAs, false detection of vessels in background sea clutter is a challenge related to a non-cooperative tracking tool like marine radar [18]. But detection points of physical objects, like vessels, exhibit spatial geometric correlation as they move through a coordinate system while false targets typically do not, providing a potential method for identifying true vessel trajectories regardless of geolocation [19]–[21].

This research highlights feature selection and machine learning techniques employed to identify trajectory attributes useful in discriminating between vessels and false targets in radar tracking data and ultimately feature true vessel records for M2 users. Machine learning is a common tool for analyzing vessel trajectories through two-dimensional space, useful in route prediction [19] and detection of illegal fishing [9], [22]. By identifying true vessels versus false targets, the M2 system can extract meaningful vessel activity from radar data for local MPA management.

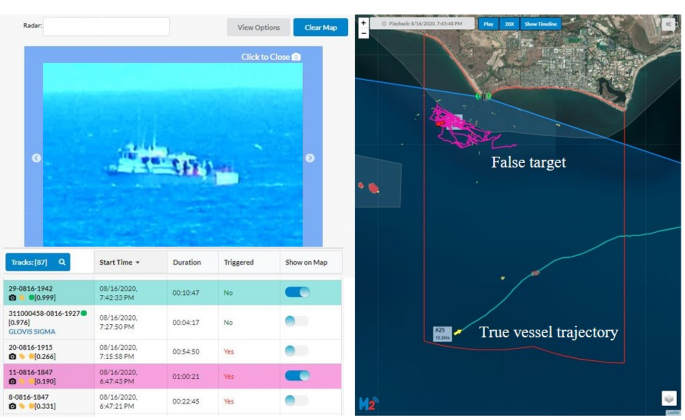

Fig. 2. Web-based user interface provides users with trajectory information and target photos (left) and visualization on a map with managed area boundaries (right) for both live and historical trajectory records.

II. Methods

What follows is a description of the steps employed to assign unique M2 records with a target confidence score as a measure of the likelihood that the record represents the trajectory of a true vessel. Steps include the creation of training and testing datasets, feature selection, and trajectory identification. To create a generalizable predictive model reliant only on the trajectories themselves, and thus non-location-specific [19], data from all M2 sites are used in model development. Cross-validation results are presented from an M2 site located in Cabo Pulmo, Baja California Sur, Mexico which monitors vessel traffic in the Cabo Pulmo National Park (CPNP), an area of increased biodiversity and fish biomass [23].

A. Dataset

Through the M2 Viewer, users tag trajectory records with observable information, such as vessel type and activity, whether a target was valid or false, and other notes. The result is an evolving dataset of known vessel and false target trajectories. All records tagged as a true vessel or false target between Jan. 1, 2020 and Jun. 2, 2020 were used in analysis with target validity classified as 1 or 0, respectively.

Trajectory features based on concise and interpretable properties of the path travelled are constructed for each record [24]. Each trajectory is comprised of $n + 1$ detection points received from the radar system $(x_0, v_0, t_0), (x_1, v_1, t_1), \ldots, (x_n, v_n, t_n)$, where x_i is the position, and v is the instantaneous velocity at time t of point i from which features are calculated (Table 1).

B. Feature Selection

First, each unique feature was tested for near zero variance [25]. Features with low variance should not be used as model predictor variables [26]. Next, the Shapiro-Wilk test determined that variables were not normally distributed, therefore the non-parametric Spearman correlation coefficients were calculated for each feature pair. Coefficients with absolute values greater than 0.7 indicated bivariate correlation [27].

To identify those features most useful in predicting target validity, the model-based ranking method, Extreme Gradient Boosting (XGBoost), provided the gain value [28], a measure

TABLE I
Trajectory features calculated by M2

Feature	*Definition*
f_1^*	Time elapsed between t_0 and t_n
f_2^*	Minimum velocity detected
f_3^*	Maximum velocity detected
f_4	Average velocity detected
f_5	Sum of distances between x_{i-1} and x_i
f_6^*	Maximum distance of x_i from x_0
f_7^*	Ratio of f_5 to f_6
f_8^*	Average heading (circular)
f_9	Standard deviation of heading (circular)
f_{10}^*	Average heading change between x_{i-1} and x_i
f_{11}	Standard deviation of heading change between x_{i-1} and x_i

*Feature retained for analysis.

of feature importance, by evaluating the presence of a feature when constructing classification trees in model training [29]. XGBoost has previously been used to successfully identify fishing vessel trajectories using trajectory features [22]. A bootstrap method was applied where all outputs henceforth reflect the average across ten unique XGBoost models. Within each collinear pair, the feature with the lower gain was removed from further analysis.

Finally, the Wrap method was used to ensure all remaining features positively influenced prediction accuracy [22]. Features were added one by one in order of decreasing gain to an XGBoost model to verify prediction accuracy increased at each step.

Model prediction accuracy is henceforth evaluated using the following process. With each iteration, 4-fold cross-validation is performed on the dataset of tagged vessel and false target trajectories using input features and known target validity. Hyperparameter tuning is conducted in model training [30]. Records retained for testing receive continuous prediction values between 0 and 1, deemed the target confidence score, reflecting the likelihood a trajectory was that of a true vessel. To evaluate accuracy, scores are converted to binary format using the threshold p associated with the highest true skill statistic value in testing data [31], [32]. Predicted validity is then compared to known validity, and accuracy is reported as the percent of correctly classified trajectories.

C. Trajectory Identification

Selected features were used to train and test an XGBoost model using the methods described in the previous section. Cross-validation results for the Cabo Pulmo M2 site are presented and are an average with standard deviation across ten iterations.

III. Results

A. Dataset

A total of 9580 tagged trajectory records from nine unique M2 sites were included in analysis with 49.7% of trajectories classified as true vessels and 50.3% classified as false targets. Prediction results were evaluated on 854 records from the Cabo Pulmo M2 site.

B. Feature Selection

All eleven features did not have near zero variance and thus were tested for collinearity of which four feature pairs were positively correlated (Fig. 3). The stepwise addition of each of the remaining seven features improved model prediction accuracy (Fig. 4), so all were ultimately retained for analysis.

C. Trajectory Identification

Overall prediction accuracy of Cabo Pulmo trajectories was $94.1 \pm 1.6\%$ when $p = 0.53 \pm 0.16$ (Table 2). True vessels and false targets were accurately classified at comparable rates, both greater than 90%.

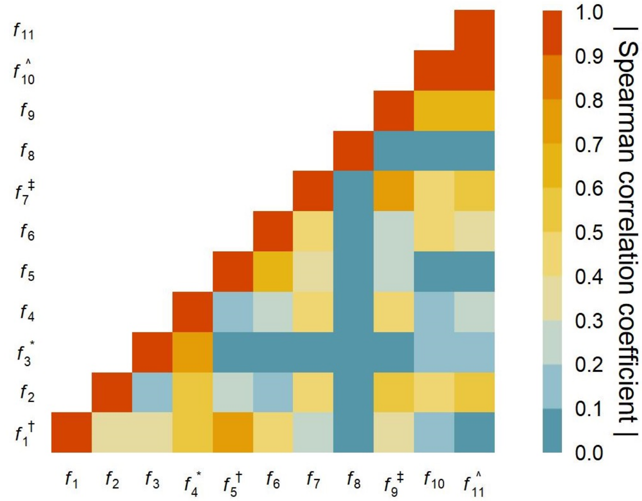

Fig. 3. Spearman correlation coefficients (absolute value) for each feature pair. Symbols indicate collinear pairs.

IV. Discussion

The trajectory features used to train and test the final model were generally consistent with those employed to discriminate between trajectories of fishing vessels using different gears, such as speed [8], [9], [22] and directional changes [33]. By employing a probability model, a high level of classification accuracy was achieved using these features [32].

Identifying false target records at the Cabo Pulmo M2 site allows users to focus on likely true vessels in the vicinity of the CPNP. Since the tagging of vessel trajectory records as true or false is somewhat opportunistic in practice, the dataset used may not indicate the realized ratio of true vessel and false target records. But results indicate that over 90% of tagged false targets are identified correctly using this modelling process, so by flagging these records, local management can view a dataset that better reflects the true vessels within the local MPA. This leads to fewer false alarms sent to managers and provides a more precise depiction of vessel activity hot

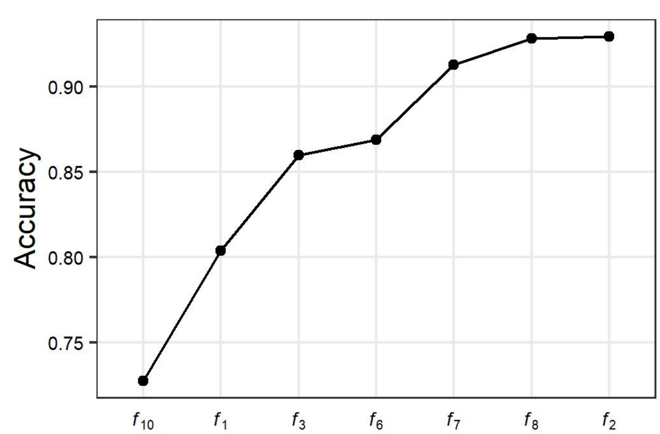

Fig. 4. Prediction accuracy with the addition of each feature.

TABLE II
CONFUSION MATRIX OF CROSS-VALIDATION RESULTS

		Predicted		
		True vessel	False target	Percent correct
Actual	True vessel	156 ± 8	10 ± 4	$93.8 \pm 2.6\%$
	False target	2 ± 2	42 ± 6	$95.4 \pm 3.7\%$
			Overall	$94.1 \pm 1.6\%$

spots (Fig. 5). While CPNP is regarded as a successful MPA, enforcement of fishing regulations is an ongoing concern [34].

The cloud-based data management flow from local M2 systems to the M2 Viewer interface enables prediction scores to be calculated on incoming records based on a pre-trained machine learning model. Trajectory features are continuously calculated for active tracks and updated values are integrated into the model at five-minute intervals thereby providing target confidence scores for both live and historical records in the M2 Viewer (Fig. 2). The process of selecting features, tuning and cross-validation, and ultimately training a model for deployment is performed bi-monthly to evaluate prediction accuracy and incorporate newly tagged records for model improvement.

Target confidence score integration in the M2 Viewer provides a simple solution to users for discriminating between true vessels and false targets. In this way, M2 overcomes the challenge of handling false targets inherent in marine radar data. Providing users with the opportunity to easily contribute to the process through tagging records in the M2 Viewer integrates local human knowledge facilitating meaningful user engagement in the technology [35].

When combined with AIS records and target photos, true vessel radar-detected records enhance a picture of vessel traffic in a given area. With more information on vessel activities occurring in the vicinity of an MPA, including activities of

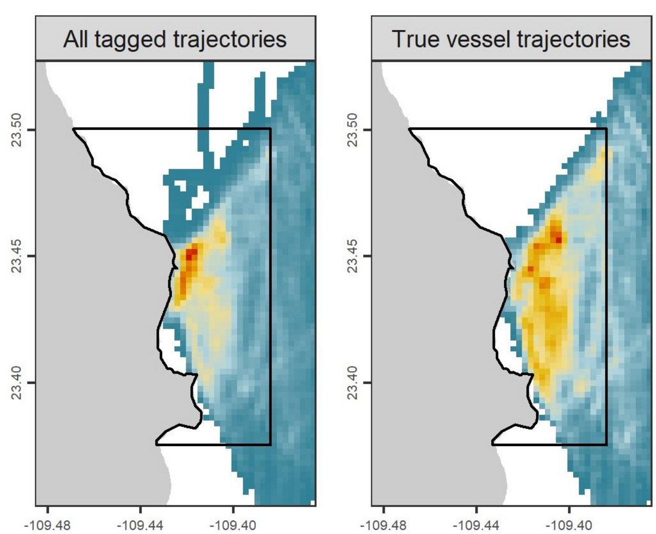

Fig. 5. Heat map of relative presence (red = high; blue = low) for all tagged trajectories (left) and true vessel trajectories (right) in the vicinity of Cabo Pulmo National Park (black outline).

those vessels not participating via AIS or VMS, managers can more effectively evaluate compliance with established regulations and more finely design enforcement efforts. Future improvements to the M2 platform include thermal cameras to provide information on vessels detected at night, drone integration for extending the range of surveillance, and additional modeling efforts to further identify activities of interest.

ACKNOWLEDGMENT

The authors would like to acknowledge Pronatura Noroeste A.C. for their role in managing the Cabo Pulmo M2 system and Global Conservation for providing funding for the site.

REFERENCES

[1] G. J. Edgar et al., "Global conservation outcomes depend on marine protected areas with five key features," Nature, vol. 506, pp. 216–220, Feb. 2014.

[2] R. Elahi et al., "Leveraging vessel traffic data and a temporary fishing closure to inform marine management," Front. Ecol. Environ., vol. 16, pp. 440–446, Aug. 2018.

[3] D. J. McCauley et al., "Ending hide and seek at sea," Science, vol. 351, pp. 1148–1150, Mar. 2016.

[4] N. A. J. Graham and T. R. McClanahan, "The last call for marine wilderness?" BioScience, vol. 63, no. 5, pp. 397–402, May 2013.

[5] A. Arias, R. L. Pressey, R. E. Jones, J. G. Álvarez-Romero, and J. E. Cinner, "Optimizing enforcement and compliance in offshore marine protected areas: a case study from Cocos Island, Costa Rica," Oryx, vol. 50, no. 1, pp. 18–26, Jan. 2016.

[6] D. A. Gill et al., "Capacity shortfalls hinder the performance of marine protected areas globally," Nature, vol. 543, pp. 665–669, Mar. 2017.

[7] A. P. Guarderas, S. D. Hacker, and J. Lubchenco, "Ecological effects of marine reserves in Latin America and the Caribbean," Mar Ecol Prog Ser., vol. 429, pp. 219–225, May 2011.

[8] J. Lee, A. B. South, and S. Jennings, "Developing reliable, repeatable, and accessible methods to provide high-resolution estimates of fishing-effort distributions from vessel monitoring system (VMS) data," ICES J. Mar. Sci., vol. 67, no. 6, pp. 1260–1271, Sep. 2010.

[9] E. N. de Souza, K. Boerder, S. Matwin, and B. Worm, "Improving fishing pattern detections from satellite AIS using data mining and machine learning," PLoS One, vol. 11, no. 9, e0158248, Sep. 2016.

[10] M. Taconet, D. Kroodsma, and J. A. Fernandes, "Global atlas of AIS-based fishing activity - challenges and opportunities," FAO, Rome, Italy, 2019. Accessed Aug. 18, 2020. [Online]. Available: http://www.fao.org/3/ca7012en/CA7012EN.pdf

[11] D. K. Moutopoulos, G. Katselis, G. Prodromitis, and C. Koutsikopoulos, "Mapping fisheries hot-spot and high-violated fishing areas in professional and recreational small-scale fisheries," Aquac. Fish., in press.

[12] D. D. Bloisi, L. Iocchi, D. Nardi, and M. Fiorini, "Integrated visual information for maritime surveillance," in Clean Mobility and Intelligent Transport Systems, M. Fiorini and J. Lin, Eds. London, UK: IET, 2015, pp. 237–264.

[13] S. Cope, E. Hines, R. Bland, J. D. Davis, B. Tougher, and V. Zetterlind, "Application of a new shore-based vessel traffic monitoring system within San Francisco Bay," Front. Mar. Sci., vol. 7, no. 86, Feb. 2020.

[14] W. Kazimierski and A. Stateczny, "Radar and Automatic Identification System Track Fusion in an Electronic Chart Display and Information System," J. Navig., vol. 1, no. 6, pp. 1–14, Jun. 2015.

[15] "Marine Monitor (M2)," ProtectedSeas, Palo Alto, CA, USA, 2020. Accessed: Aug. 18, 2020. [Online]. Available: https://protectedseas.net/

[16] "Leonardo DiCaprio Foundation Donates Towards Costa Rica Radar Facility on Cocos Island," The Costa Rican Times, New York, NY, USA, 2014. Accessed: Aug. 18, 2020. [Online]. Available: https://www.costaricantimes.com/leonardo-dicaprio-foundation-donates-towards-costa-rica-radar-facility-on-cocos-island/32481

[17] T. Prudente, "Military-Grade Radar Network Watching For Oyster Poachers," Capital Gazette, Annapolis, MD, USA, 2014. Accessed: Aug. 18, 2020. [Online]. Available: https://www.capitalgazette.com/news/ph-ac-cn-oyster-radar-0927-20140925-story.html

[18] A. G. Bole, A. D. Wall, and A. Norris, Radar and ARPA Manual: Radar, AIS and Target Tracking for Marine Radar Users. Oxford, UK: Butterworth-Heinemann, 2013.

[19] A. Valsamis, K. Tserpes, D. Zissis, D. Anagnostopoulos, and T. Varvarigou, “Employing traditional machine learning algorithms for big data streams analysis: the case of object trajectory prediction,” J. Syst. Softw., vol. 127, pp. 249–257, May 2017.

[20] D. Huang, G. Cui, X. Yu, M. Ge, and L. Kong, “Joint range-velocity deception jamming suppression for SIMO radar,” IET Radar, Sonar Navig., vol. 13, no. 1, pp. 113–122, 2019.

[21] Q. Li, L. Zhang, Y. Zhou, S. Zhao, N. Liu, and J. Zhang, “Discrimination of active false targets based on Hermitian distance for distributed multiple-radar architectures,” IEEE Access, vol. 7, pp. 71872–71883, Jun. 2019.

[22] H. Huang, F. Hong, J. Liu, C. Liu, Y. Feng, and Z. Guo, “FVID: Fishing Vessel Type Identification Based on VMS Trajectories,” J. Ocean Univ. China, vol. 17, no. 5, pp. 1–10, May 2018.

[23] J. P. Egerton, A. F. Johnson, J. Turner, L. LeVay, I. Mascareñas-Osorio, and O. Aburto-Oropeza, “Hydroacoustics as a tool to examine the effects of Marine Protected Areas and habitat type on marine fish communities,” Sci. Rep., vol. 8, no. 47, Jan. 2018.

[24] M. D. Rintoul and A. T. Wilson, “Trajectory analysis via a geometric feature space approach,” Stat. Anal. Data Min., vol. 8, no. 5–6, pp. 287–301, Oct. 2015.

[25] M. Kuhn, “Caret: classification and regression training,” R Package Version 6.0–86, 2020. [Online]. Available: https://cran.r-project.org/web/packages/caret/index.html

[26] V. Clarkson, P. J. Kootsookos, and B. G. Quinn, “Analysis of the variance threshold of Kay’s weighted linear predictor frequency estimator,” IEEE Trans. Signal Process., vol. 42, no. 9, pp. 2370–2379, Sep. 1994.

[27] B. G. Tabachnick and L. S. Fidell, Using Multivariate Statistics, 3rd ed. New York, NY, USA: Harper Collins, 1996.

[28] T. Chen et al., “Xgboost: Extreme Gradient Boosting,” R Package Version 1.1.1.1, 2020. [Online]. Available: https://cran.r-project.org/web/packages/xgboost/index.html

[29] T. Chen and C. Guestrin, “Xgboost: a scalable tree boosting system,” in Proc. 22nd ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining, Aug. 2016, pp. 785–794.

[30] B. Bischl et al., “Mlr: machine learning in R,” J. Mach. Learn. Res., vol. 17, no. 170, pp. 1–5, 2016.

[31] G. Collell, D. Prelec, and K. R. Patil, “A simple plug-in bagging ensemble based on threshold-moving for classifying binary and multiclass imbalanced data,” Neurocomputing, vol. 275, no. 31, pp. 330–340, Jan. 2018.

[32] A. Muench, G. S. DePiper, and C. Demarest, “On the precision of predicting fishing location using data from the Vessel Monitoring System (VMS),” Can. J. Fish. Aquat. Sci., vol. 75, no. 7, pp. 1036–1047, 2018.

[33] M. I. Marzuki, P. Gaspar, R. Garello, V. Kerbaol, and R. Fablet, “Fishing gear identification from vessel-monitoring-system-based fishing vessel trajectories,” IEEE J. Ocean. Eng., 2017.

[34] R. B. Anderson, “Beyond “success”: community, governance, and the future of Cabo Pulmo National Park,” Hum. Organ., vol. 78, no. 2, pp. 147–157, 2019.

[35] A. Vandecasteele, R. Devillers, and A. Napoli, “A semi-supervised learning framework based on spatio-temporal semantic events for maritime anomaly detection and behavior analysis,” CoastGIS 2013 – The 11th Int. Symp. for GIS and Computer Cartography for Coastal Zone Management, Aug. 2013.