Seafloor Classification based on an AUV Based Sub-bottom Acoustic Probe Data for Mn-crust survey

Umesh Neettiyath[1], Harumi Sugimatsu[1], Blair Thornton[2,1]

[1]Institute of Industrial Science, The University of Tokyo, Komaba, 153-8505, Japan
(E-mail: umesh@iis.u-tokyo.ac.jp)

[2]Southampton Marine and Maritime Institute, University of Southampton, Southampton SO16 7QF, UK

Abstract—**The possibility of automatically classifying high frequency sub-bottom acoustic reflections collected from an Autonomous Underwater Robot is investigated in this paper. In field surveys of Cobalt-rich Manganese Crusts (Mn-crusts), existing methods relies on visual confirmation of seafloor from images and thickness measurements using the sub-bottom probe. Using these visual classification results as ground truth, an autoencoder is trained to extract latent features from bundled acoustic reflections. A Support Vector Machine classifier is then trained to classify the latent space to idetify seafloor classes. Results from data collected from seafloor at 1500m deep regions of Mn-crust showed an accuracy of about 70%.**

I. INTRODUCTION

Deep sea mineral deposits, in particular, Cobalt-rich manganese crusts (Mn-crust) are gaining popularity as a potential source of valuable minerals including Cobalt, Nickel, and rare earth elements. These are formed mainly on the slopes and shoulders of seamounts ranging from 800m to 2400m [1], [2]. Stakeholders have attempted different methods of surveying such as core drills for detailed analysis, towed camera surveys for visual inspection and ship based multibeam backscatter studies for large area surveys. While sampling gives high accuracy thickness measurements with very low spatial resolution, multibeam and camera studies cannot measure thickness values. Thickness and coverage are important results necessary to estimate the volume of Mn-crust; and an integrated approach is necessary to accurately estimate the resource potential.

Towards this goal, an acoustic probe for in-situ contactless thickness measurement, and an Autonomous Underwater Vehicle (AUV), called Boss-A, to conduct continuous surveys were developed by the Institute of Industrial Science of the University of Tokyo [3], [4], an Illustration of which is given in Fig. 1. A method for calculating the volumetric distribution of Mn-crust was developed by combining the thickness measured from the acoustic probe with the coverage of exposed Mn-crust calculated from 3D visual maps [5]. The system is being successfully deployed for Mn-crust volumetric surveys regularly.

This was achieved by classifying the 3D colour point cloud of the seafloor generated using a Support Vector Machine

classifier (SVM) [6]. Typically occurring seafloor types - Mn-crust, nodules, and sediment deposits were the target classes, as shown in Fig. 2. In areas detected as Mn-crust, a thickness was calculated from the acoustic reflections. In the existing volumetric estimation method, although the seafloor characterization was done from the visual data [5], it was observed that the seafloor reflections from different areas appear different. The proposed paper intends to create a new classifier for classifying acoustic data without relying on visual information. Also, it is possible that features not visible in the visual images can be observed in sub-bottom scans, leading to further insights during seafloor characterization.
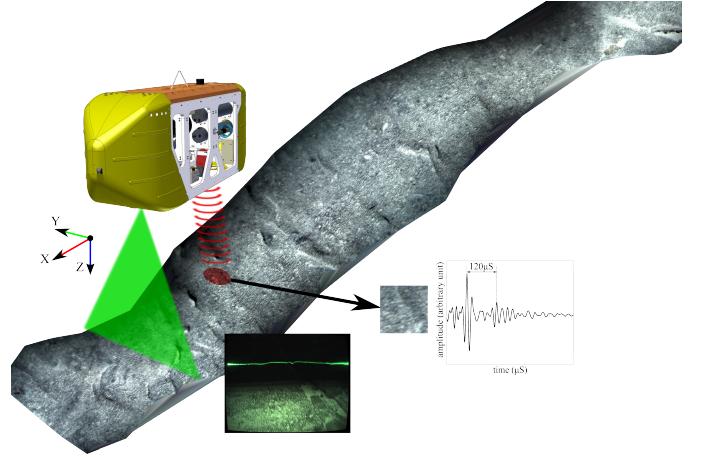

Fig. 1. Illustration of the mapping scenario using AUV Boss-A. The robot surveys the seafloor using an acoustic subbottom probe and a visual 3D mapping system. Examples of images collected and acoustic reflections from Mn-crust areas are shown.

Rest of this paper is organized as follows. A description of the acoustic and visual systems are described in section II. Section III details the algorithms, and the data processing methods. Results from field surveys and discussion are gievn in section IV. Section V contains the concluding remarks and future directions.

II. SYSTEM OVERVIEW

The AUV Boss-A system consist of visual and acoustic survey subsystems, as shown in Fig. 1. These systems are

TABLE I
SPECIFICATIONS OF THE SYSTEMS USED IN THE SURVEY

Visual mapping system	
Camera resolution	1328×1048
Horizontal opening angle	$65°$
Vertical opening angle	$53°$
Laser to camera baseline	1.22 m
Frame rate	15 fps
Acoustic Probe	
Frequency	2 MHz (carrier), 200kHz (signal)
-3 dB footprint	< 2 cm (dynamic focusing)
Ping rate	20 Hz
Sampling rate	2 MHz

(a) Top view of a 21m long seafloor 3D map

(b) Close-up views of each classes

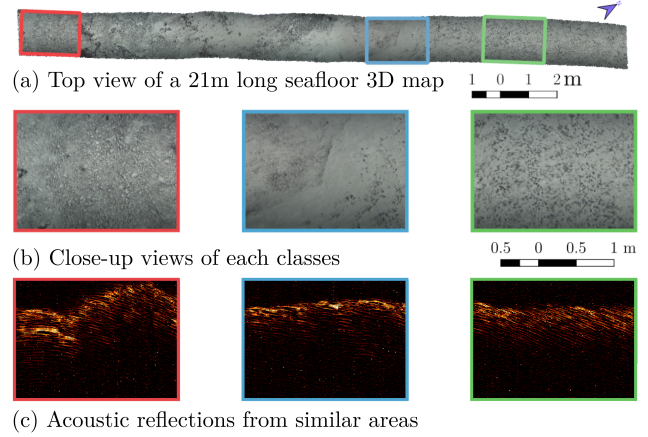(c) Acoustic reflections from similar areas

Fig. 2. Examples of typical seafloor in Takuyo Daigo seamount at 1500m depth. Flat Mn-crust (red), sediments (blue), and Nodules (green) have distinct acoustic sub-bottom and visual footprints.

supported by the navigational system on the AUV - Doppler Velocity Log, AHRS systems, and depth sensor. The specifications of the systems are given in Table I.

A. Visual mapping system

The visual mapping system is based on the light-sectioning based seafloor 3D mapping system developed in [7]. It consists of a sheet laser, LEDs for illumination, and a camera which records the laser projection on the seafloor. From each recorded image, the laser deformation is calculated and used to calculate the seafloor bathymetry. From the LED illuminated regions, RGB colour values of each point on the seafloor is recognised and combined with the bathymetry information to generate 3D color reconstructions of the seafloor. The resolution of the generated 3D map is ~1.4 mm.

B. Acoustic sub-bottom probe

The acoustic probe is a parametric subsurface sonar that records subsurface reflections of the seafloor for estimating the Mn-crust thickness [3]. The probe consists of a five-channel annular array of 2-MHz piezoelectric transducers for transmission and a 200-kHz piezoelectric transducer to record reflections. It is dynamically focused on the seafloor at ranges from 0.5 to 2.5 m. The recorded signal typically consists of reflections from the top of the seafloor and the bottom of the Mn-crust layer in crust covered areas. The thickness can then be calculated by multiplying the time delay between the reflections with the velocity of sound in Mn-crust. In sediment and nodule covered areas, no secondary reflection can be seen, with the exception of buried Mn-crust layers as studied in [8]. As shown in Fig. 2(c), while only Mn-crust areas have secondary reflections, the primary reflections are also characteristic of the seafloor type.

III. ALGORITHM AND PROCESSING WORKFLOW

A flowchart of the proposed algorithm is given in Fig. 3. The visual 3D colour reconstruction of the seafloor is divided into 10cm x 10cm cells and each cell is classified into Mn-crust, sediment or nodules by an SVM classifier, developed in [6]. Each sample is georeferenced to identify cells in which acoustic sub-surace reflections are also made. Since the generated 3D map is ~2 m in width, whereas the acoustic data

is collected along a line in the middle, only a fraction of the cells would be selected.

The acoustic reflections collected are stacked into time-series images, sections of which are shown in Fig. 2(c). This image is sampled to select the reflections around the point of incidence of the acoustic pulse on the seafloor. The size of each sample(called an acoustic tile) is 30x30 pixels. By using speed of sound in Mn-crust (2932 ± 179 m/s), the selected region corresponds to an approximate physical size of 44 mm in the vertical direction, which is large enough to capture the top reflection, while excluding the unwanted reflections. In the horizontal direction, the acoustic tile would have a size of 15 cm assuming a constant AUV velocity. Fig. 4 shows some of the acoustic tiles generated. The Mn-crust thickness values are also calculated from the acoustic data using the method demonstrated in [9], but is not included in the current paper.

The acoustic tiles generated are used as the input into a feature extractor in order to generate a minimal vector which can accurately represent the data. This was achieved using an autoencoder based on Alexnet [10]. Alexnet is a deeply stacked artificial neural network with convolutional layers and has been used in a variety of image data extraction applications [11], [12]. In [10], Alexnet was combined with an inverted version of Alexnet in order to reconstruct the images from the latent vectors. In addition, physical proximity of the samples is considered into a loss function, bearing the idea that images (in this case, acoustic tiles) collected from nearby areas are more likely to be similar. An overall loss function is calculated by combining reconstruction loss with proximity loss. The autoencoder is then trained to minimize the loss during reconstruction of the images.

Once the training is completed, the Alexnet classifier is used to generate latent vectors for each acoustic tile. The extracted feature vectors are classified using a second Support Vector Machine (SVM) classifier. SVM is an unsupervised classification method widely used for image classification and
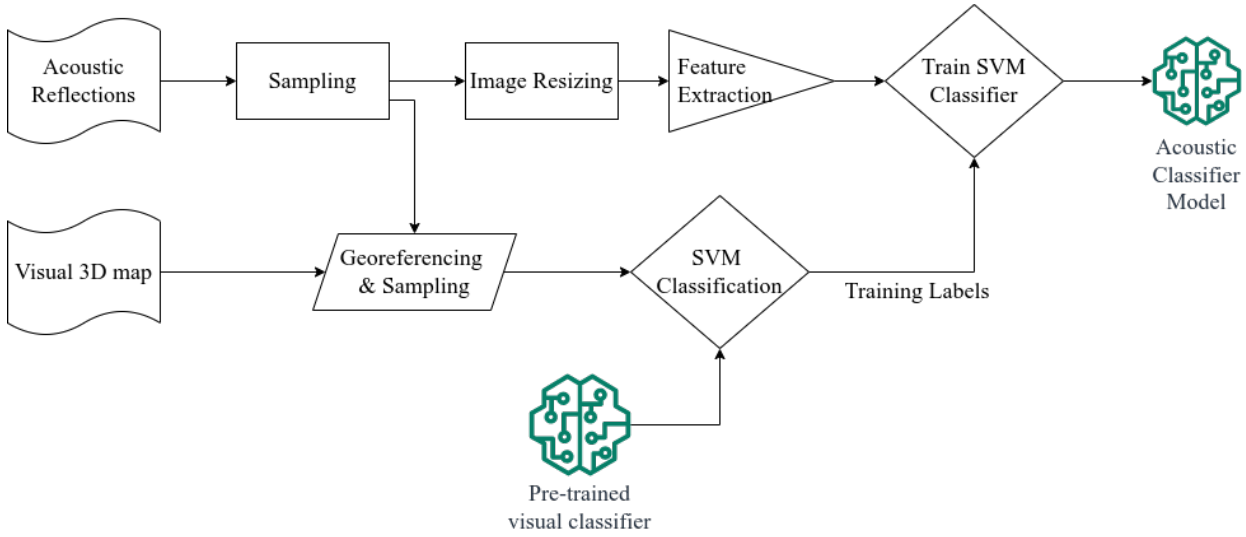
Fig. 3. Flowchart of the proposed classification method. Using previously developed classifier for visual 3D data as ground truth, a classifier for acoustic subbottom reflections from co-located areas are trained.
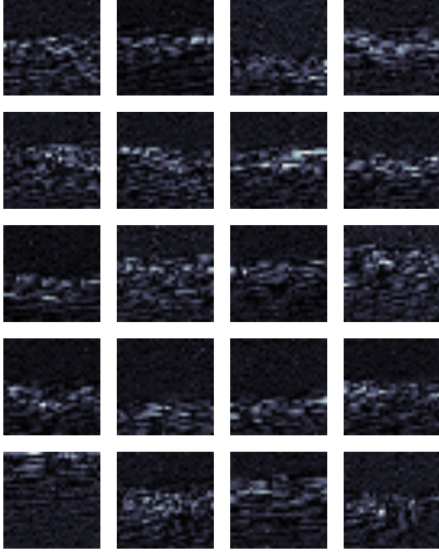

Fig. 4. Acoustic tiles used for feature extraction

Transect	Area (m^2)	No. of samples			
		Mn-crust	Sediment	Nodules	Total
A	970	730	189	377	1296
B	462	335	231	106	672

IV. IMPLEMENTATION AND RESULTS

AUV Boss-A was deployed in field experiments at Takuyo Daigo seamount in the northwestern Pacific Ocean at depths ranging from 1350 m to 1650 m [5] during the KR16-01 cruise of R/V Kairei of the Japan Agency for Marine Earth Science and Technology. The AUV surveyed the seafloor at a velocity of 0.2 kn (0.1 m/s) at a target altitude of 1.5 m.

Two transects of the robot are selected to study the proposed method, as shown in Table II. Non-overlapping samples were selected at nearly uniform intervals to form two datasets. It can be seen that most of the samples are Mn-crust, since the target area in which the surveys were conducted was Mn-crust. Efforts were made to remove overfitting the classifier by normalising the input data.

The two datasets were combined and sampled randomly to generate training and testing datasets. 70% of the points were assigned to the training set and 30% of the points were assigned to the test set. Training test was used to train the SVM classifier and the performance of the classifier was examined using the test set.

The results of the test set are shown in Fig. 5. It can be seen that Mn-crust is classified accurately (93% accuracy), whereas performance is lower in other classes, some of sediment and Nodules were classified wrongly as Mn-crust. While it could be an issue due to the imbalance of the training set, efforts

segmentation tasks [13]. It works by generating a hyperplane in an N-dimensional space which maximizes the separation between the nearest training input data samples. The input data consists of a feature vector, which is the latent space representation of generated by the autoencoder in the previous step and training labels, which are generated by classifying the co-located visual 3D maps using the first SVM classifier. The training labels belong to the 3 classes of seafloor typically found in Mn-crust covered seamounts - Mn-crust, sediment and nodules, as shown in Fig. 2. The training and testing of the classifier using field collected data is done using machine learning workflows and is described in the following section.

were made to reduce overfitting. Some areas were seen which contained Mn-crusts with a very thin sediment layer of a few millimeters or less. In such areas, while the visual images indicate sediments, acoustic reflections could indicate Mn-crusts. Further analysis is required into these specific areas to fully understand this scenario.
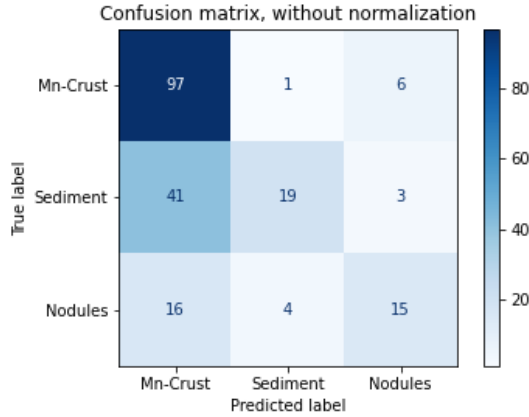

Fig. 5. Confusion matrix of the classifier on the test set. It can be seen that the prediction is biased towards Mn-crust for all classes.

The classification results of the complete dataset is shown in Fig. 6. The results are overlaid on the top view of the 3D reconstruction for visual confirmation reference. Two selected areas are shown below and their locations are indicated in the overall results. In Fig. 6(b), a nodule covered area is shown. It can be seen that most of the area is classified accurately, except for two points towards the left. The classification results from the visual 3D classification using the existing classifier is shown in Fig. 6(d), where some misclassification of nodules as sediments is visible in the upper portions. This is caused by the limitations of the colour correction algorithm as part of the 3D map generation. This could cause some of the training labels to be incorrect, creating classification errors for nodule tiles. Efforts to correct this are underway. Fig. 6(c) and (d) shows a Mn-crust covered section. The left side of the section has flat Mn-crusts whereas towards the middle, pillowy Mn-crust is visible. In the crevices between the Mn-crusts, sediment deposits are visible. This section has been accurately classified in both visual and acoustic classification schemes. Since different topographies of Mn-crust are visible, the impact of this in the acoustic reflections can be studied further.

If the acoustic probe can reliably characterize Mn-crust covered seafloors and calculate thickness measurements, it is possible to conduct surveys without a visual mapping system, thereby simplifying the systems and reducing the payload requirements. Another potential research direction is identifying the substrate types using these methods. Since substrates are buried under several centimeters of Mn-crust, visual or sonar surveys cannot provide information regarding them; at present, sampling is the only available method.

The possibility of using a semi-supervised classification on the extracted feature vector is a potential future research theme. The aim of this will to detect different sub-classes within each class - such as different surface topographies of Mn-crust deposits. It is useful in cross-validating visual classification schemes and for error correction. It can be used when visual data is unreliable or not available. Further applications in scientific studies into Mn-crust topography and formations are possible.

V. CONCLUSION

This paper proposed a method for automatically classifying the seafloor sub-surface reflections collected by an acoustic sub-bottom sonar probe. The acoustic data is bundled into image tiles around the seafloor incidence point, latent features were extracted using an autoencoder, and was classified into Mn-crust, sediment or nodules. Using a field collected dataset from an AUV at 1500m depth,the method was validated and found to have an accuracy of 65% overall and 93% for Mn-crusts. The method can assist the volumetric estimations of Mn-crust deposits, by cross-validating visual classifier results and reduce errors in areas such as thin sediment covered Mn-crusts. It also introduces the possibility to simplify the survey setup and improve scalability by eliminating the 3D mapping system; albeit losing some useful data.

REFERENCES

[1] M. R. Clark, R. Heydon, J. R. Hein, S. Petersen, A. Rowden, S. Smith, E. Baker, and Y. Beaudoin, *Deep Sea Minerals: Cobalt-rich Ferromanganese Crusts, a physical, biological, environmental, and technical review*, E. Baker and Y. Beaudoin, Eds. Secretariat of the Pacific Community, 2013, publication Title: Deep Sea Minerals.

[2] A. Usui, K. Nishi, H. Sato, Y. Nakasato, B. Thornton, T. Kashiwabara, A. Tokumaru, A. Sakaguchi, K. Yamaoka, S. Kato, S. Nitahara, K. Suzuki, K. Iijima, and T. Urabe, "Continuous growth of hydrogenetic ferromanganese crusts since 17 Myr ago on Takuyo-Daigo Seamount, NW Pacific, at water depths of 800–5500 m," *Ore Geology Reviews*, vol. 87, pp. 71–87, Jul. 2017, publisher: Elsevier B.V.

[3] B. Thornton, A. Asada, A. Bodenmann, M. Sangekar, and T. Ura, "Instruments and methods for acoustic and visual survey of manganese crusts," *IEEE Journal of Oceanic Engineering*, vol. 38, no. 1, pp. 186–203, Jan. 2013, iSBN: 0364-9059 VO - 38.

[4] Y. Nishida, K. Nagahashi, T. Sato, A. Bodenmann, B. Thornton, A. Asada, and T. Ura, "Autonomous Underwater Vehicle "BOSS-A" for Acoustic and Visual Survey of Manganese Crusts," *Journal of Robotics and Mechatronics*, vol. 28, no. 1, pp. 91–94, Feb. 2016.

[5] U. Neettiyath, B. Thornton, M. Sangekar, Y. Nishida, K. Ishii, A. Bodenmann, T. Sato, T. Ura, and A. Asada, "Deep-Sea Robotic Survey and Data Processing Methods for Regional-Scale Estimation of Manganese Crust Distribution," *IEEE Journal of Oceanic Engineering*, vol. 46, no. 1, pp. 102–114, Jan. 2021.
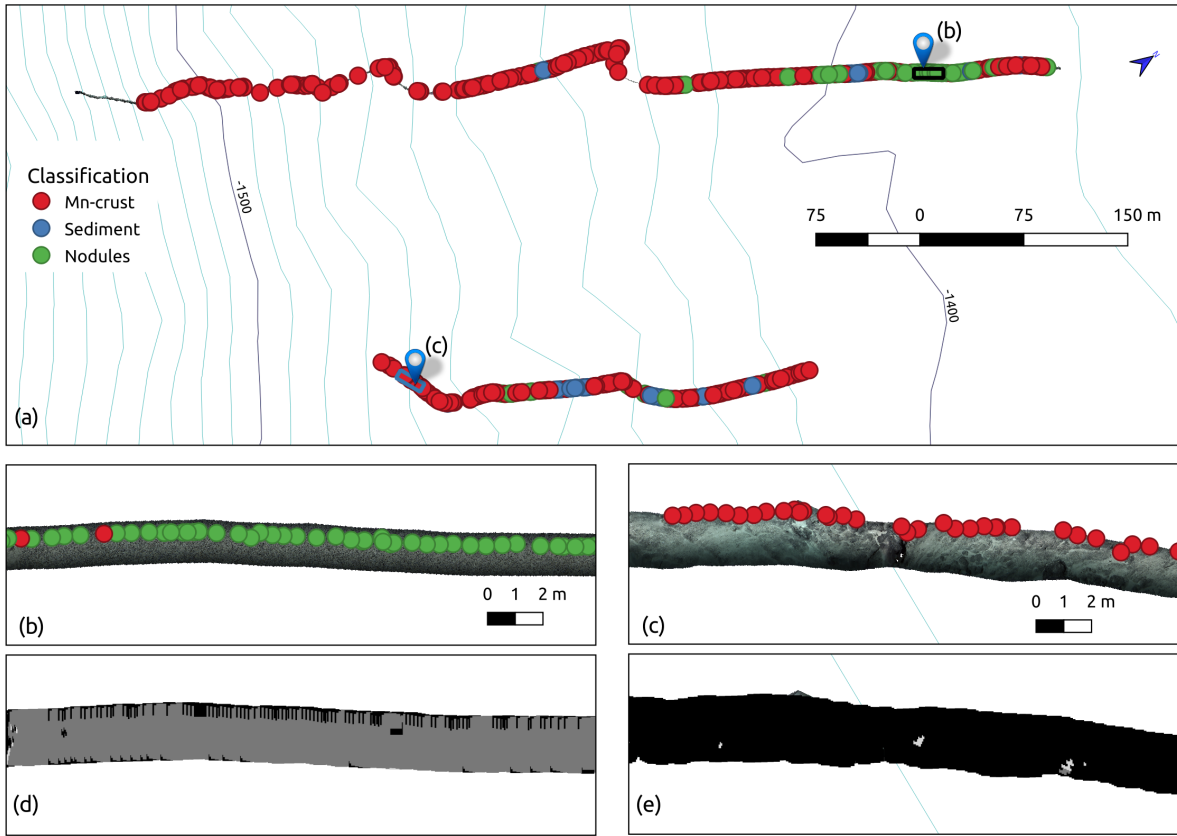
Fig. 6. Results of classification of two transects surveyed by the AUV at 1500 m depth. Most of the area is covered in Mn-crust, with occasional sections of sediment and nodules. (a) Overall results. Two selected sections containing (b)sediment and (c)Mn-crust are also shown. (d) and (e) shows the visual classification results for the sections (b) and (c) respectively.

[6] U. Neettiyath, T. Sato, M. Sangekar, A. Bodenmann, B. Thornton, T. Ura, and A. Asada, "Identification of manganese crusts in 3D visual reconstructions to filter geo-registered acoustic sub-surface measurements," in *OCEANS 2015 - MTS/IEEE Washington*. IEEE, Oct. 2015, pp. 1–6.

[7] A. Bodenmann, B. Thornton, and T. Ura, "Generation of High-resolution Three-dimensional Reconstructions of the Seafloor in Color using a Single Camera and Structured Light," *Journal of Field Robotics*, vol. 34, no. 5, pp. 833–851, Dec. 2017, arXiv: 10.1.1.91.5767 ISBN: 9783902661623.

[8] U. Neettiyath, B. Thornton, H. Sugimatsu, T. Sunaga, J. Sakamoto, and H. Hino, "Automatic Detection of Buried Mn-crust Layers Using a Sub-bottom Acoustic Probe from AUV Based Surveys," in *Oceans 2022*, Chennai, 2022.

[9] U. Neettiyath, B. Thornton, M. Sangekar, K. Ishii, T. Sato, A. Bodenmann, and T. Ura, "Automatic Extraction of Thickness Information from Sub-Surface Acoustic Measurements of Manganese Crusts," in *Oceans 2017 - Aberdeen*. IEEE, Jun. 2017, pp. 1–7.

[10] T. Yamada, A. Prügel-Bennett, and B. Thornton, "Learning features from georeferenced seafloor imagery with location guided autoencoders," *Journal of Field Robotics*, no. October 2019, pp. 1–16, 2020.

[11] G. Cheng, P. Zhou, and J. Han, "Learning Rotation-Invariant Convolutional Neural Networks for Object Detection in VHR Optical Remote Sensing Images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 12, pp. 7405–7415, Dec. 2016.

[12] T. Yamada and A. Pr, "GeoCLR : Georeference Contrastive Learning for Efficient Seafloor Image Interpretation," *ArXiv*, 2021, arXiv: 2108.06421v1.

[13] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, A. Müller, J. Nothman, G. Louppe, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine Learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, Jan. 2011, arXiv: 1201.0490 ISBN: 1532-4435.