Polymetallic Nodules Abundance Estimation using Sidescan Sonar: A Quantitative Approach using Artificial Neural Network

Wong Liang Jie, Bharath Kalyan, Mandar Chitre and Hari Vishnu

Keppel-NUS Corporate Laboratory, Acoustic Research Laboratory
National University of Singapore
E-mail: {ljwong, bharath, mandar, hari}@arl.nus.edu.sg

Abstract*—There is a high abundance of polymetallic nodules (PMN) scattered across the vast Clarion and Clipperton Fracture Zone (CCFZ) in the Pacific Ocean. These nodules possess high economic potential as they are rich in minerals such as manganese, nickel, copper and rare earth elements. Quantification of nodule coverage is important for economic feasibility studies and planning of effective exploitation strategies. Traditional methods for nodule quantification are highly labour and time intensive as they rely on freefall box corer measurements and/or image processing of seabed photographs. Using sidescan sonar data and geotagged photographs collected from an autonomous underwater vehicle (AUV) in our region of interest at CCFZ, we propose a novel technique based on artificial neural network (ANN) to estimate PMN abundance using texture variations from sidescan sonar data. Compared to an optical camera, the sidescan sonar provides a much larger area of coverage, which in effect can drastically increase the area surveyed by an AUV in a given amount of time. Till date, this is the first known published work to elaborate on a data-driven approach in estimating PMN abundance using sidescan sonar backscatter data. Our network yielded a test accuracy of 84%, which shows that it can be used as an effective tool in estimating nodule abundance from sidescan sonar. This approach allows faster evaluation of nodule abundance for future exploration without the need for an underwater camera.

***Index Terms*—Deep seabed mining; Polymetallic nodules; Sidescan sonar image processing; Artificial neural network**

I. Introduction

The occurrence and high abundance of PMN on the abyssal seabed of the CCFZ has been well documented [1], [2]. These PMN possess economic potential as they are rich in minerals such as manganese, nickel, copper and rare-earth elements that are commonly used in many industrial applications [3].

Quantification of nodule coverage is important for economic feasibility studies and planning of effective exploitation strategies. However, these PMN are unevenly distributed across CCFZ with higher abundance in the central and north-eastern region [2], [4]. Furthermore, it has been reported that abundance of PMN exhibits large variability within a span of kilometres [5]. Thus it is important to have a more detailed quantification of nodule abundance and its variation within the region of interest.

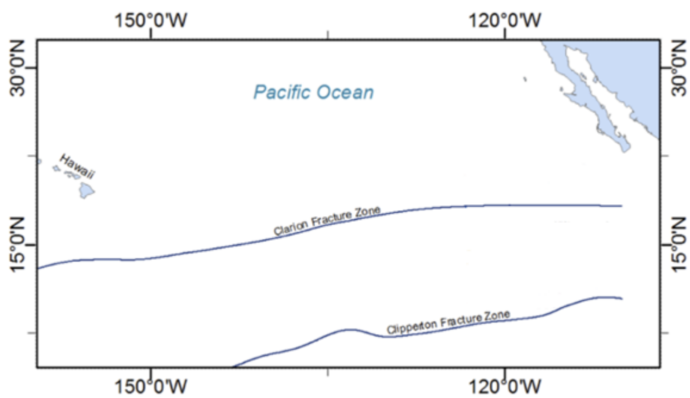
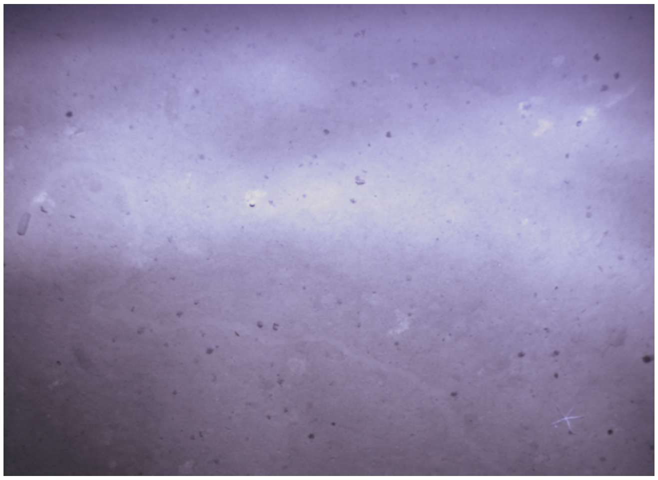


Fig. 1: Data in this study is collected from a sampling area of $4.8\,\mathrm{km}^2$ in the eastern region of the CCFZ.

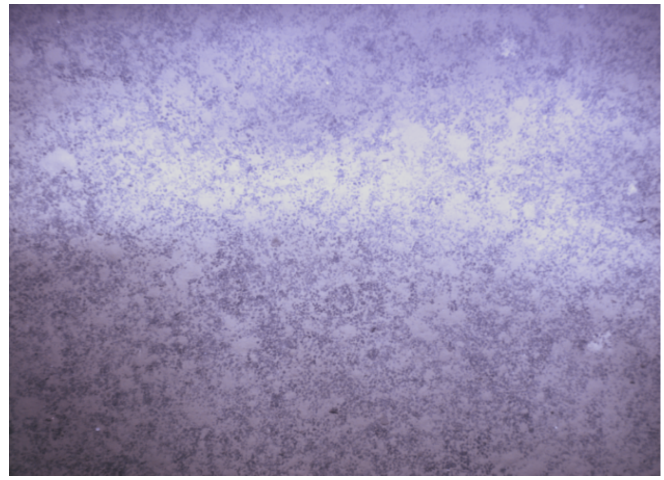
Traditional methods are highly labour and time intensive and rely on planimeter and point counting of PMN collected from various forms of sampling devices such as freefall grab and box corer [6], [7]. A more recent method uses image processing of seabed photographs captured with a camera mounted on an AUV. Due to its greater speed in recording data from the seabed, this method has gained significant traction as the preferred method for quantifying deep-sea PMN [8]–[11]. In order to assess the large quantities of photographs collected, high-performance computing, with efficient image processing algorithms running on graphic-processing units has been used [12]. However, the total seabed area photographed by the AUV-mounted camera is still too small to allow for a more precise large-area assessment of seabed needed for feasibility studies and exploitation.

With the advancements in underwater sensing technology, there have been studies on the use of acoustic sensors such as the multibeam and sidescan sonar to perform classification of seabed terrain [13]. Studies have also suggested a qualitative relationship between the acoustic returns of sidescan sonar and the PMN abundance [14], [15]. However, there has been no published work that details a data-driven approach to use the patterns found in sidescan sonar for PMN abundance

978-1-5090-5278-3/17/$31.00 ©2017 IEEE



(a) Photograph with low PMN coverage of $5.674\,\%$ translating to an area of $0.682\,\mathrm{m}^2$.



(b) Photograph with high PMN coverage of $61.351\,\%$ translating to an area of $7.585\,\mathrm{m}^2$.

Fig. 2: Identification and quantification of PMN from seabed photographs

estimation.

We propose a novel technique to estimate PMN abundance using texture variations from sidescan sonar backscatter. We do this by using an ANN to interpret the sidescan backscatter, by training it against ground truth data consisting of seabed photographs from the same location. During training, the ANN models the relationship between the patterns in the sidescan sonar data and the amount of PMN indicated from the photographs obtained from the camera. Once trained, the ANN can be used to infer nodule abundance at any other site using only sidescan sonar data. Thus, we would be able to survey large areas using this ANN model to interpret the sidescan sonar data.

In the following section, we describe the geographical area of study, the data collection method, and the preprocessing techniques used. Section III discusses the training, validation and testing of the ANN model, and section IV presents results demonstrating the accuracy of the model in quantifying sidescan sonar data. Finally we conclude the paper in section V.

II. Data collection and processing

A. Study site

The CCFZ is a geological submarine region of approximately 15.5 million km^2 situated between $120\,^\circ$ to $120\,^\circ\mathrm{W}$ and $0\,^\circ$ to $20\,^\circ\mathrm{N}$, in the Pacific Ocean as illustrated in Fig. 1 [1]. Regions within CCFZ lie mostly within depths of $3\,\mathrm{km}$ to $6\,\mathrm{km}$. The data used in this paper was collected as part of an environmental baseline survey cruise, where an AUV was deployed at specific region of interest along the north-east region of CCFZ in 2015.

B. Equipment

The AUV utilized during this data collection run was equipped with an inertial navigation system, doppler velocity log, camera, lighting and laser scaling system and sidescan sonar. In addition, a long baseline system was also used for positioning and navigation of the AUV.

C. Data collection

The photographs and sidescan sonar images used in this paper were collected by the AUV at an average depth of $4125\,\mathrm{m}$. During the run, the AUV travelled at an average speed of $2.8\,\mathrm{knots}$ at an altitude of $8\,\mathrm{m}$ above the seabed in a lawnmower pattern across the seabed. Photographs of the seabed were taken at approximately 3-second intervals while the sidescan sonar data was collected continuously for the entire AUV run. Around 3500 photographs, each depicting a seabed area of approximately $12\,\mathrm{m}^2$, and sidescan sonar data spanning $4.8\,\mathrm{km}^2$, were used for the training, validating and testing of the ANN.

D. Processing of seabed photographs

The collected photographs were processed to correct variations in illumination conditions. Then, a feature-based image processing technique for quantifying nodule distribution from photographs was used to identify the nodules and quantify their coverage area within each photograph as illustrated in Fig. 2. We classified the photographs into two categories, taking into consideration that the economically acceptable range for mining is between $5\,\mathrm{kg/m}^2$ to $20\,\mathrm{kg/m}^2$ [3]. A threshold of 40% translates to a nodule density of around $23\,\mathrm{kg/m}^2$. Based on this threshold, the photographs were classified into high and low nodule coverage regions. $45\,\%$ of the photographs were labelled as high nodule coverage while the remaining $55\,\%$ were labelled as low nodule coverage category. However, the ANN can be specially trained to separate seabed with a specific PMN abundance coverage requirement.

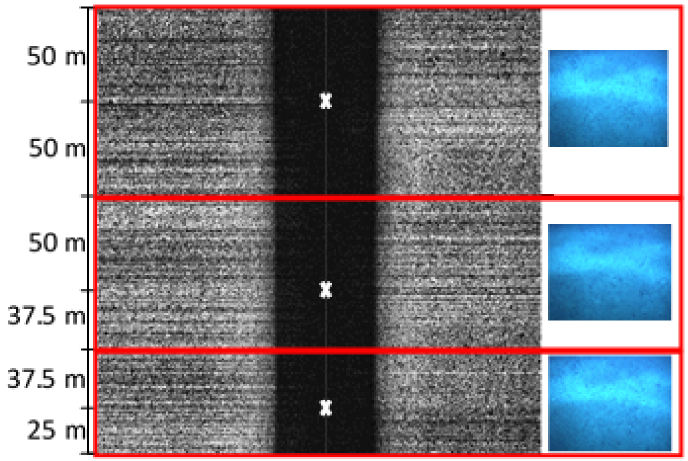


Fig. 3: Illustration of back scatter from sidescan sonar. Symbol 'x' indicates position of geotagged photograph seen on the right. Nodule density seen at each geotagged photograph is uniform up to $50\,\mathrm{m}$ along-track and across-track the sidescan image as shown by the red border.

III. Implementation

A. Methodology

The geotagged photographs are superimposed onto the sidescan sonar image allowing us to correlate nodule abundance shown in the photographs with the sidescan sonar image. A visual comparison reveals that the variations in the two are somewhat correlated. Based on this observation, we aim to capture this correlation using ANN, thereby allowing us to use sidescan sonar for nodule quantification. We use ANN, which is known to be good at learning features or patterns from a given labelled training dataset [16]. ANN is able to capture the unknown, complex and nonlinear relationships between the features and the labels. Thus, it is an ideal tool to learn the nonlinear functions required to interpret sidescan sonar patterns in terms of nodule density estimates.

From our training dataset, we use the ANN training algorithm (details to be discussed in section III-C) to learn the best interconnecting weight parameters between the neurons in each layer. This is done by minimizing the cost function which is the mean cross-entropy error between the labelled values and the ANN predicted values. During the training phase, the ANN weights are iteratively modified to best represent the relationship between the sidescan sonar data and the nodule density obtained via photographs.

Ample labelled training data samples allow an ANN to have better insights on the underlying patterns of the dataset, enabling the ANN to be sufficiently trained in making meaningful predictions. If the number of training data samples is too small, the network would not have enough information to learn adequately the dependencies between the labels and features.

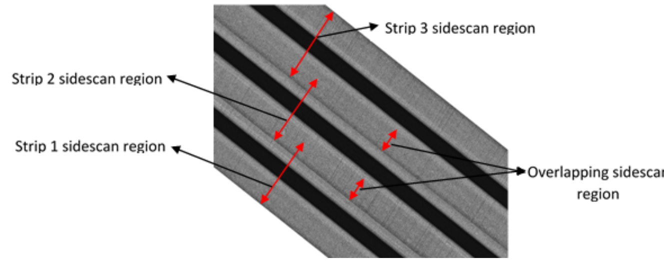


Fig. 4: Overlapping regions between each strip of sidescan data.

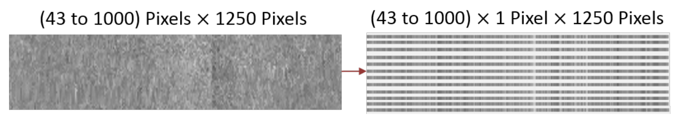


Fig. 5: Multi row data samples are sub-divided into single row data samples to increase the number of data samples.

B. Data set preparation

During each mission, the AUV traversed in a lawnmower fashion collecting data from both the camera and the sidescan sonar. The swath-width of the sidescan sonar was set to $100\,\mathrm{m}$. At an altitude of $8\,\mathrm{m}$, the area covered by a photo was $12\,\mathrm{m}^2$, which was much smaller compared to the image generated by sidescan sonar backscatter. Thus, in order to fully utilized the data gathered from the sidescan sonar, we assumed that the nodule density is uniform for distances of up to $50\,\mathrm{m}$ (or in pixel co-ordinates of $500\,\mathrm{px}$) from the geotagged position of each photo, both in along-track and across-track directions, as shown in Fig. 3.

Even though the optical imaging system was programmed to take photographs every three seconds, due to some variability in time taken for flash recharging, the photographs were taken at irregular intervals. Hence at instances where the distance between two consecutive photographs is less than $50\,\mathrm{m}$, the along-track segment of the sidescan sonar image was divided equally as illustrated in Fig. 3. Each segment was then labelled in accordance with the closest geotagged image. This process split the sidescan sonar image along-track into tiny segments ranging from $4.3\,\mathrm{m}$ to $100\,\mathrm{m}$ (translates to $43\,\mathrm{px}$ to $1000\,\mathrm{px}$ in pixel coordinates) in length. Lastly, the sidescan sonar's nadir of $250\,\mathrm{px}$ as illustrated by the central black strip in Fig. 3 and Fig. 4 was removed, as it did not contain any useful textural information. Thus, the size of each data packet consisted of segments varying from $43\,\mathrm{px}$ to $1000\,\mathrm{px}$ along-track and $750\,\mathrm{px}$ across-track, depending on the distance between consecutive photos.

In addition, the AUV's lawnmower path was programmed to have at least $25\,\%$ sidescan overlap between run-lengths, resulting in overlapping regions as illustrated in Fig. 4. These overlapping segments of sidescan data of approximately $25\,\mathrm{m}$ ($250\,\mathrm{px}$) on either side of any run-length would be appended to the data packet increasing the across-track width from $750\,\mathrm{px}$ to $1250\,\mathrm{px}$. Thus, the size of each data packet consisted

Fold	1	2	3	4	5	6	7	8	9	10	**Tenfold Average**
Accuracy %	84.87	82.23	86.77	83.5	83.03	82.28	84.45	86.97	83.06	84.88	**84.20**

TABLE I: Accuracy of trained ANN model on test data based on 10 training/validation/testing datasets.

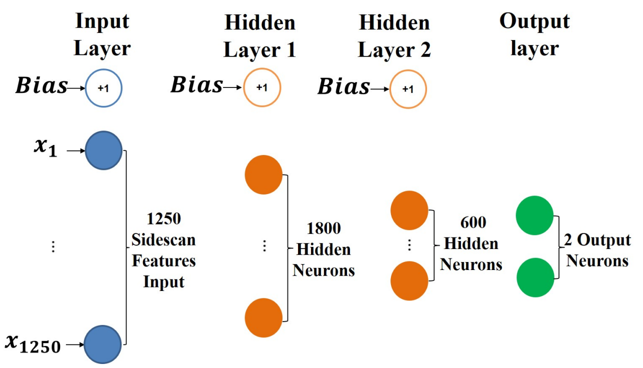


Fig. 6: ANN architecture used in PMN abundance estimation.

of segments varying from 43 px to 1000 px along-track and 1250 px across-track, depending on the distance between consecutive photographs. This data packet was further separated into single strips (henceforth, referred to as data samples) of 1 px × 1250 px, as illustrated in Fig. 5.

These 400,000 data samples were normalized and labelled as either '1' or '2' denoting high and low percentage nodule abundance. The labelled data samples were then separated into training (80 % of total labelled data samples), validation (10 %) and testing (10 %) dataset. The validation dataset was used for selecting the model hyperparameters and the final ANN model to use. The testing dataset was used to evaluate the performance of the model on data which it had not been trained or chosen with, and thus represented a somewhat objective measure of its performance in generalizing its estimates.

C. *Training Algorithm*

Our ANN is a feedforward network with two hidden layers. This network architecture is able to learn underlying patterns from a large number of distinct data samples with a comparatively small number of hidden neurons. Based on [17], we choose the number of neurons in the ANN model to be 1800 and 600 neurons for hidden layer 1 and 2 respectively as illustrated in Fig. 6.

In our training method, the sidescan training dataset is collectively treated as a single 1250 × n matrix and feedforward propagated through the neural network, where 'n' is the number of data samples in the training dataset. A sigmoid activation function whose role is to generate a non-linear decision output based on the weighted input is applied to the output of every neuron in hidden layer 1 and 2, and the output layer. The input data is normalized before applying the ANN weights, so that it does not saturate the nonlinearity.

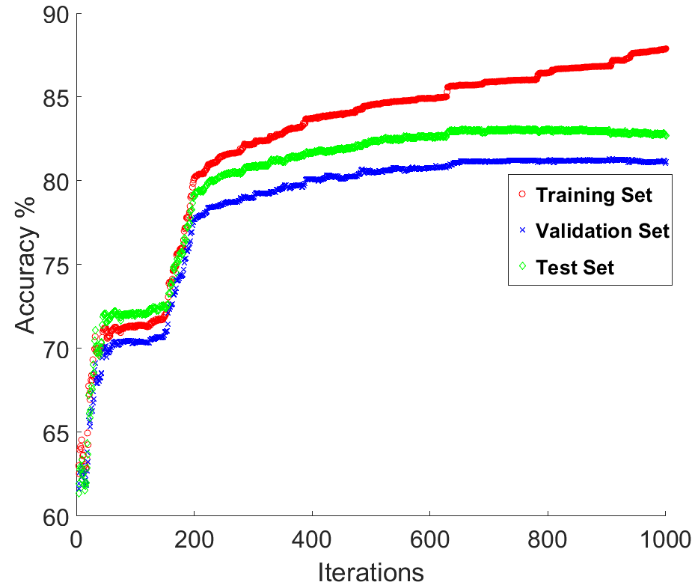


Fig. 7: An example demonstrating overfitting occuring on the 4^{th} training dataset after a certain number of iterations are over. Observe that there is an increase in training accuracy, but a drop in validation and test accuracy after about 904 iterations. Thus, the model seems to be overfitting after about 904 iterations.

These randomly generated weights of the ANN are trained using the feedforward and backpropagation methods through function minimization by conjugate gradient, [18], [19]. The feedforward and backpropagation processes are repeated and with each iteration the ANN weights would be automatically re-adjusted to further minimize the cost function which is the mean cross-entropy error between the predicted and actual labelled outputs.

Repeating the feedforward and backpropagation processes indefinitely will increase the ANN's accuracy rate towards the training dataset. However, doing so would also lead to overfitting whereby the trained weight parameters are so specifically tuned towards the training dataset that they begin to erroneously treat its underlying noise as features. Having a trained ANN overfitting on a particular training dataset will result in a low accuracy for subsequent unseen datasets. Thus, along with the training process, it is important to ensure that the trained network is able to generalize to any future datasets. To achieve this, the generalization ability of the ANN is monitored by checking the model against the validation dataset after every iteration process of the training phase. The final set of interconnecting weights chosen for the ANN will be the

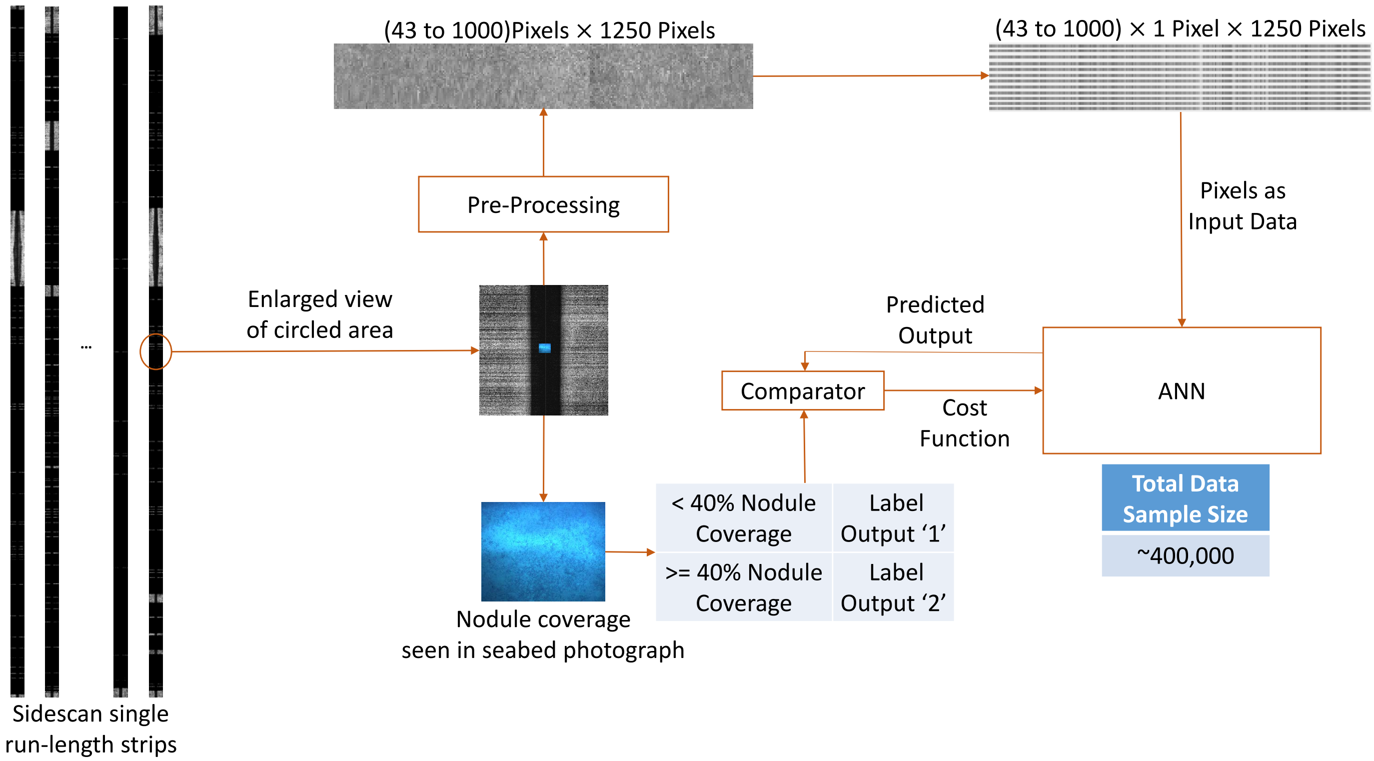


Fig. 8: Schematic showing various stages of proposed method. Black lines across-track of single run-length strips are indicative of locations where areas surrounding each geotagged photographs are extracted in preparation for input into ANN.

one which yields maximum performance with the validation dataset, which in this case would be the weights at iteration 904 for the example illustrated in Fig. 7.

The performance metric used to gauge the ANN's performance was its accuracy with the test dataset. This test accuracy is a measure of the ANN's ability to make generalized predictions with new datasets.

IV. Results

Our results show an average accuracy rate of 84% in the ANN's ability to classify sidescan images between high and low PMN coverage. This entire process is repeated tenfold where each fold will have a randomly chosen configuration of samples from the training, validation and testing dataset. The accuracy results for all the ten trained ANNs are tabulated in Table 1, and the average accuracy is computed. This tenfold method of computing accuracy is more reliable, as it is generated based on not just one, but ten configurations of training, validation and test datasets averaged out. The training process iteration is stopped at 1000 iterations when signs of overfitting appears, after which the ANN's weight matrices will be based on the iteration where the maximum accuracy for validation dataset occurs to ensure that the ANN does not overfit on the training dataset as illustrated in Fig. 7.

A confusion matrix on the 4th fold test result shows a visualization of the ANN's classification performance as illustrated by Fig. 9. It can be seen that the ANN achieved high accuracy in predicting high nodule coverage and low nodule coverage samples. This demonstrates the ANN's ability to correctly differentiate majority of the seabed photographs between high and low nodule presence. Note that accuracy is a good performance metric in this case because of the nearly balanced number of samples between the two labelled classes. Training an ANN with a heavily skewed dataset can result in an over representation of one class to the ANN, which will severely affect the prediction capability of the ANN towards the least represented class.

The trained ANN can be further improved upon by increasing the number of different training samples and adding more relevant features to the dataset. However, this will also increase the time needed to train our ANN. Currently it takes around 20 hours to train on approximately $360,000$ ($80\,\%$) training data samples using MATLAB software on a workstation with a dual-processor Intel Xeon E5-2630 V3 CPU@2.4 GHz processor.

V. Conclusion

The total number of data samples is about 400,000, with each sample size occurring in the form of $1\,\mathrm{px} \times 1250\,\mathrm{px}$ sidescan sonar image. Of these, $80\,\%$ are used for training, $10\,\%$ for validation and the remaining $10\,\%$ for testing. These data samples are labelled '1' or '2' denoting 'high' or 'low' percentage nodules abundance based on the geotagged photograph corresponding to the sidescan sonar location. The nodule

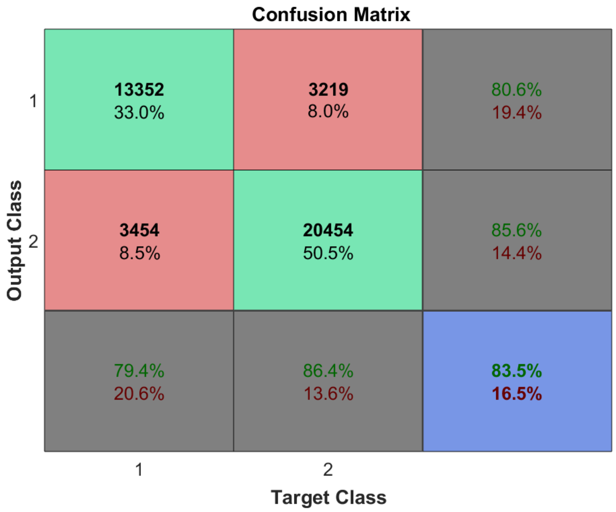


Fig. 9: Confusion matrix for 4th fold test result, the 2 diagonal green cells show the number and percentage of correct classification by our trained ANN. The grey cells in the 3rd column reveal that out of $16,571$ low nodule coverage predictions, $80.6\,\%$ of them are predicted correctly and out of the $23,908$ high nodule coverage predictions, $86\,\%$ of them are predicted correctly. The grey cells in the 3rd row reveal that out of $16,806$ low nodule coverage samples, $97.4\,\%$ are correctly predicted and out of the $23,673$ high nodule samples, $86.4\,\%$ are correctly predicted. Lastly the blue cell shows the overall accuracy of the ANN.

abundance threshold for separating these two output labels is set at $40\,\%$ as illustrated in Fig. 8. The two hidden layer neural network model used in this paper consists of 1800 and 600 neurons for hidden layer 1 and 2 respectively.

The ANN discussed in this paper is shown to be capable of approximating PMN abundance considering the relatively small number of photographs we obtained, in comparison to the vastness of the CCFZ area.

Till date, this is the first known published work to make use of a data-driven approach to perform PMN abundance estimation using backscatter pattern from the sidescan sonar. Our network yielded an average test performance of $84\,\%$ accuracy, which shows that it can be used as an effective tool in estimating nodule abundance using only sidescan sonar. This approach allows faster evaluation of nodules abundance for future deep seabed sites without the need for an underwater camera.

In addition, we can potentially utilize this model in different environmental conditions due to the neuroplasticity property of the ANN which means there is no need to redesign a new algorithm to cater to any new specific features discovered from a dataset as any of these new features will be automatically learned from the dataset.

Future work includes exploring the possibility of employing what we have learned onto multibeam sonar images in estimating the abundance of PMN on an even larger scale seabed area.

ACKNOWLEDGMENT

The authors thank the National Research Foundation, Keppel Corporation and National University of Singapore for supporting this work done in the Keppel-NUS Corporate Laboratory. The conclusions put forward reflect the views of the authors alone, and not necessarily those of the institutions within the Corporate Laboratory.

REFERENCES

[1] U. Von Stackelberg and H. Beiersdorf, “The formation of manganese nodules between the clarion and clipperton fracture zones southeast of hawaii,” *Marine Geology*, vol. 98, no. 2, pp. 411–423, 1991.

[2] I. S. Authority, “A geological model of polymetallic nodule deposits in the Clarion Clipperton Fracture zone,” Tech. Rep. 1, 2010.

[3] J. Mero, “Economic aspects of nodule mining,” *Marine manganese deposits*, vol. 15, p. 327, 1977.

[4] “A geological model of polymetallic nodule depostis in the clarion-clipperton fracture zone,” ISA Technical Study: No. 6, 2010.

[5] R. KAUFMAN, “The selection and sizing of tracts comprising a manganese nodule ore body.” *Offshore Technology Conference*, pp. 283–293, 1974.

[6] A. H. Bouma, *Methods for the study of sedimentary structures*. Wiley Interscience, 1969.

[7] R. Fewkes, W. McFarland, and W. Reinhart, “Development of a reliable method for evaluation of deep sea manganese nodule deposits,” 1979.

[8] H. S. Jung, Y. T. Ko, S. B. Chi, and J. W. Moon, “Characteristics of seafloor morphology and ferromanganese nodule occurrence in the Korea Deep-Sea Environmental Study (KODES) area, NE equatorial Pacific.” *Marine Georesources And Geotechnology*, vol. 19, no. 3, pp. 167–180, 2001.

[9] R. Sharma, N. Khadge, and S. Jai Sankar, “Assessing the distribution and abundance of seabed minerals from seafloor photographic data in the Central Indian Ocean Basin,” *International Journal of Remote Sensing*, vol. 34, no. 5, pp. 1691–1706, 2013.

[10] M. Okazaki and A. Tsune, “Exploration of polymetallic nodule using auv in the central equatorial pacific,” *Tenth ISOPE Ocean Mining and Gas Hydrates Symposium*, pp. 32–38, 2013.

[11] Tsune, “Effects of size distribution of deep-sea polymetallic nodules on the estimation of abundance obtained from seafloor photographs using conventional formulae,” in *Eleventh Ocean Mining and Gas Hydrates Symposium*, 2015.

[12] T. Schoening, D. Langenkämper, B. Steinbrink, D. Brün, and T. W. Nattkemper, “Rapid image processing and classification in underwater exploration using advanced high performance computing,” *OCEANS 2015 - MTS/IEEE Washington*, pp. 1–5, 2015.

[13] N. C. Mitchell and J. E. H. Clarke, “Classification of seafloor geology using multibeam sonar data from the scotian shelf,” *Marine Geology*, vol. 121, no. 3-4, pp. 143–160, 1994.

[14] S. H. Lee and K.-H. Kim, “Side-scan sonar characteristics and manganese nodule abundance in the clarion and clipperton fracture zones, NE equatorial Pacific,” *Marine Georesources and Geotechnology*, vol. 22, no. 1-2, pp. 103–114, 2004.

[15] M. M. P. Weydert, “Measurments of the deep seafloor and some implications for the assessment of manganese nodule resources,” 2016.

[16] J. Schmidhuber, “Deep Learning in neural networks: An overview,” *Neural Networks*, vol. 61, pp. 85–117, 2015.

[17] G.-B. Huang, “Learning capability and storage capacity of two-hidden-layer feedforward networks.” *IEEE transactions on neural networks / a publication of the IEEE Neural Networks Council*, vol. 14, no. 2, pp. 274–81, jan 2003.

[18] R. Rumelhart, D.E., Hinton, G.E., Williams, “Learning representations by back-propagating errors,” *Nature*, vol. 319, no. 30, pp. 402–403, 1986.

[19] R. Fletcher and C. M. Reeves, “Function minimization by conjugate gradients,” *Journal of Optimization Theory and Applications*, vol. 7, no. 2, pp. 149–154, 1964.