This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

Acoustic Assessment of Polymetallic Nodule Abundance Using Sidescan Sonar and Altimeter

Liang Jie Wong, Bharath Kalyan, *Member, IEEE*, Mandar Chitre, *Senior Member, IEEE*, and Hari Vishnu, *Member, IEEE*

***Abstract*—Polymetallic nodules (PMN) are potato-sized concretions containing metals, such as manganese, copper, nickel, cobalt, and rare earth elements, and are a potential valuable resource of minerals. They occur in high abundance and are unevenly distributed across the Clarion Clipperton Fracture Zone. Existing PMN abundance estimation methods using box corers, and manual assessment through seabed photographs are labor and time intensive, and can only survey small sections of seabed at a time. Compared to an underwater camera, acoustic sensors are able to survey the PMN abundance across larger tracts of seabed at a time. In this article, we present a method for PMN abundance assessment using heterogeneous acoustic data, which is a combination of bathymetry information and sidescan sonar measurements of seabed backscatter. We achieve this using an artificial neural network model that classifies a given region into a low or high PMN density region using these features. Our model will enable faster estimation of PMN abundance for future deep seabed site surveys without the need for underwater cameras. To date, our proposed method yields an average accuracy of 85.36% on a testing data set, demonstrating our method's effectiveness in estimating PMN abundance.**

***Index Terms*—Artificial neural network (ANN), Clarion and Clipperton Fracture Zone, deep seabed mining, polymetallic nodule estimation.**

I. Introduction

LOCATED in the subequatorial region in the northeast of the Pacific Ocean, the Clarion Clipperton Fracture Zone (CCFZ) is a submarine region spanning between 5° N to 20° N and 120° W to 160° W. The prevalence of polymetallic nodules (PMN) in high abundance on the abyssal seabed of the CCFZ is well-documented [1]–[3]. These potato-sized concretions may come in several shapes such as spherical, discoidal, or irregular. They are formed around the nuclei found in indurated sediments through diagenetic and/or hydrogenous growth [4]. It is estimated that the CCFZ holds 21 billion tons of PMN which represents the world's largest concentration [3]. These PMN can be quantified from seabed photographs and/or core sampling as past explorations have found that PMN reside mostly on or just beneath the surface of the abyssal seabed [4].

The PMN possess high economic value as they consist of metals such as manganese, copper, nickel, cobalt, and rare earth elements which are greatly in demand as raw materials in many industries [3], [5]. As such, numerous commercial mining consortia are increasingly looking at these as an alternative mineral resource to meet the growing needs of the industries [3].

Although there is a high abundance of PMN scattered across the CCFZ, the distribution of PMN across the CCFZ is uneven, with higher abundance in the central and north-eastern regions [2], [6]. Studies have shown that the distribution of PMN exhibits considerable variability even within a span of kilometers [6], [7]. As a result, a more extensive exploration method is required to better assess the quantity of PMN in the CCFZ. Accurate assessment of PMN distribution in these regions is crucial for evaluating the economic feasibility of exploiting these resources, and forming strategies for doing so effectively. Therefore, prospecting consortia would require an accurate PMN abundance estimation method to aid their exploration strategies.

In the early days of PMN exploration, Murray and Renard obtained coarse estimations of PMN abundance on small seabed areas using coring methods [8]. Such manual PMN abundance estimation methods are laborious and time-intensive as they rely on planimeters or point-counting of PMN collected from various sampling devices such as free-fall grab and box corers [9], [10]. In addition, these methods only sample a small seabed area per deployment with sampling points potentially spaced kilometers apart. Thus, they yield only pockets of sparse sampling points, and there would be uncertainty on the PMN abundance in regions where very less sampling was done. Although such sparse sampling may be adequate in assessing PMN abundance across a small seabed area, interpolating these measurements across a vast prospecting seabed area could result in a poor estimate of PMN abundance [11].

A comparatively faster method of underwater photography which involved lowering the underwater camera to a preset altitude above the seabed was used by Glassy and Singleton [12] to perform *in situ* estimation of PMN abundance. Compared to the limited seabed area covered by each sampling grab, underwater cameras are able to provide PMN data from a larger number of seabed locations, and are thus a more effective means of estimating PMN abundance. Sharma estimated PMN abundance through the use of a vessel's tow-frame-mounted underwater

Manuscript received February 18, 2019; revised August 13, 2019 and December 2, 2019; accepted January 9, 2020. This article was presented in part at the MTS/IEEE OCEANS Conference, Aberdeen, Scotland, June 19–22, 2017. *(Corresponding author: Liang Jie Wong.)*

Associate Editor: M. Atmanand.

The authors are with the Keppel-NUS Corporate Laboratory and Acoustic Research Laboratory, Tropical Marine Science Institute, National University of Singapore, Singapore 119077 (e-mail: tmswong@nus.edu.sg; bharath@nus.edu.sg; mandar@nus.edu.sg; tmshv@nus.edu.sg).

Digital Object Identifier 10.1109/JOE.2020.2967108

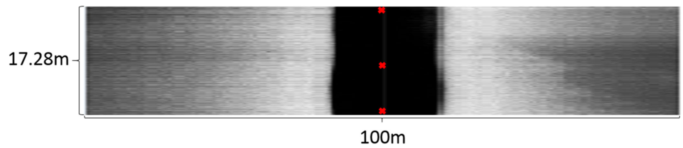


Fig. 1. Segment of the SSS seabed backscatter data. Over a timespan of 12 s, the AUV covered a distance of 17.28 m with an average speed of 2.8 kn. The AUV-mounted SSS imaged a seabed area of 1728 m^2 while the AUV-mounted camera photographed a seabed area of 36 m^2. Due to the time required for flash recharging between each photograph, only three seabed photographs (each capturing a seabed area of 12 m^2) depicted by the red "X"s were taken during this period.

camera which allows seabed photographs to be taken at more frequent intervals [13]. To expedite the analysis of these seabed photographs, Sharma *et al.* [14] digitized these photographs and developed a machine-learning (ML)-based image-processing software to estimate PMN abundance from these seabed photographs. Later, Okazaki and Tsune [15] used AUV-mounted underwater camera and acoustic equipment to survey the seabed area within the Japanese licensed zone in the CCFZ. Using this AUV configuration, they collected more than 10 000 seabed photographs and also built a detailed bathymetry map of the area. Using these, they studied the correlation between bathymetry features and PMN distribution.

AUV-mounted and deep-towed cameras have gained significant traction among many researchers for quantifying deep sea PMN abundance [6], [15], [16]. Image processing algorithms that enable accurate PMN abundance estimation based on seabed photographs have been explored [17]. Even though the efficiency of estimating PMN abundance can be increased through this method, the total seabed area photographed by AUV-mounted and deep-towed cameras is still too limited to enable rapid assessment of the seabed as compared to an underwater acoustic survey, as depicted in Fig. 1. Advancements in underwater sensing technology have enabled the use of underwater acoustic equipment in various seabed applications, such as seabed terrain classification and study of PMN abundance based on its correlation with sidescan sonar (SSS) backscatter [18], [19]. Although studies have suggested a qualitative relationship between PMN abundance and acoustic backscatter returns of SSS [19], [20], there is no work in the current literature that details how this relationship can be exploited to assess PMN abundance.

In our earlier work, we conducted a preliminary investigation on assessing PMN abundance using SSS seabed backscatter data [21]. Taking a step further, we present an in-depth discussion on our acoustic-based PMN abundance assessment method. Our method aims to classify whether a seabed area is of high or low abundance. Many works in the literature have pointed out that the nodule abundance at a location shows correlation with the seabed bathymetry [6], [19]. This indicates that some information on PMN abundance may be inferred from an independent measurement of the bathymetry. Taking a cue from this, we aim to enhance the accuracy of our PMN abundance assessment technique by using terrain variations along with the SSS backscatter. Fig. 2 illustrates our approach to assess PMN abundance.

Our approach involves assessing PMN abundance by capturing its correlation with SSS backscatter and bathymetry using an artificial neural network (ANN). An ANN is an ML method that learns the underlying pattern residing within a given data set and its relationship to certain features. This is achieved by training an ANN with labeled data. In our case, the labels correspond to ground-truths available on the quantity we want to assess, namely PMN abundance estimates from seabed photographs. The effectiveness of the model can be characterized by testing its performance on a testing data set that was not used during training. This would indicate whether the PMN abundance at a location is indeed correlated with the features used by us, namely SSS backscatter and bathymetry at the same location.

The objectives are summarized as follows.

1) To formulate a data-driven approach to assess PMN abundance using a heterogeneous feature set of seabed backscatter and bathymetry data, and labels obtained from photograph-based estimates.
2) To highlight the importance of appropriate data processing methods in achieving an ANN with good assessment performance.
3) To demonstrate a methodology to obtain a wider understanding on the PMN abundance in a region, by expanding the information from a limited quantity of seabed photographs (smaller coverage) to interpret SSS measurements (larger coverage). Our method allows for larger scale PMN abundance assessment using acoustic measurements, which are calibrated against visual ground-truth measurements made at a smaller scale.

So far in Section I, we have covered the limitations posed by existing PMN abundance data collection and estimation methods, and what we aim to achieve using our proposed data-driven method. The remaining sections of this article are organized as follows. Section II presents the geographical area of study, data collection and processing methods, and the ML algorithm used in the modeling of PMN abundance. Section III presents the performance of our trained models and the methods used in ensuring its overall reliability and generalization capability. Section IV concludes this article and describes future research directions. A list of acronyms used in this article is provided in Table I.

II. Methodology

A. Equipment

Data used in this article were collected from a region of interest (ROI) spanning approximately 5 km^2 located within the eastern part of the CCFZ as shown in Fig. 3. The water depth within this ROI ranges from approximately 4.1 to 4.24 km with a gentle sloping variation of 140 m over a distance of 4 km. A REMUS 6000 AUV [shown in Fig. 4(a)] was launched from a research vessel *Thomas G. Thompson* [shown in Fig. 4(b)] to collect the data. The AUV was equipped with an inertial navigation system, Doppler velocity log, camera coupled with lighting and laser scaling system for seabed photography, and an SSS. In addition, the AUV utilized a long baseline system for its positioning and navigation.

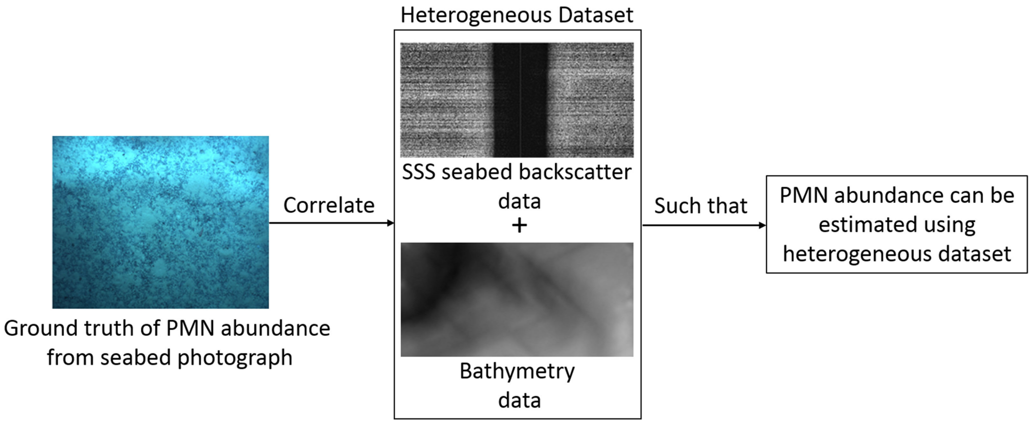


Fig. 2. Correlating PMN abundance seen in seabed photographs to SSS backscatter and bathymetry data.

TABLE I
LIST OF ACRONYMS

Acronym	Definition
ANN	Artificial Neural Network
AUV	Autonomous Underwater Vehicle
CCFZ	Clarion and Clipperton Fracture Zone
CLAHE	Contrast Limited Adaptive Histogram Equalisation
GeoTIFF	Georeferenced Tagged Image File Format
PMN	Polymetallic Nodules
ML	Machine learning
ROI	Region of Interest
SSS	Sidescan Sonar

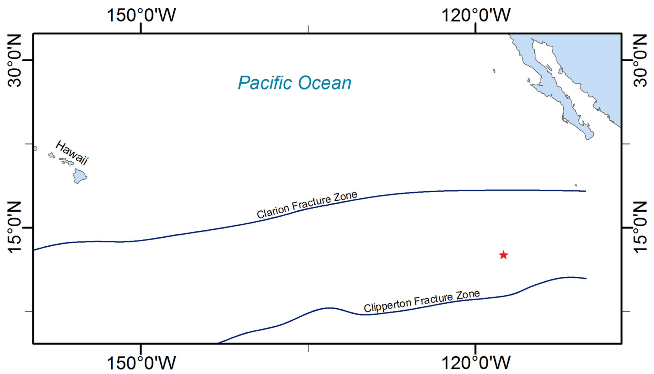


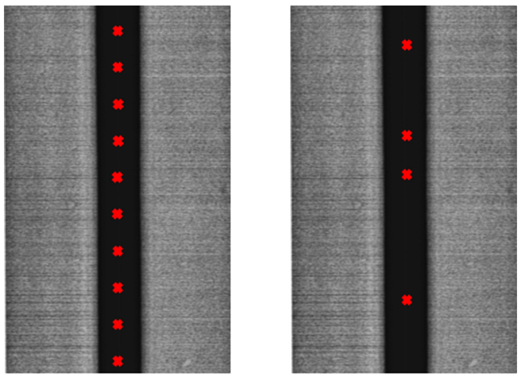
Fig. 3. Map of northeastern Pacific (reproduced from [21]), showing the extent of the CCFZ by the blue lines. Data used in this article were collected from a sampling area of approximately 5 km^2 within the Singapore license area for PMN exploration, marked by the red star.

B. Data Collection

During the data collection dive, the AUV traveled at an average speed of 2.8 kn while maintaining an average altitude of 8 m. The AUV-mounted SSS collected the backscatter data, and a low-resolution bathymetric map of the area was produced by combining data collected using the AUV-mounted depth, altimeter, and navigation sensors. In addition, the AUV-mounted camera was programmed to take a seabed photograph every three seconds. However, due to the variability in the time taken for flash recharging, the seabed photographs were occasionally taken at irregular intervals resulting in patches of along track seabed with no photograph as shown in Fig. 5. During the dive time of 15 h, the AUV traversed a preprogrammed

(a) (b)

Fig. 4. (a) REMUS 6000 AUV. (b) Research vessel *Thomas G. Thompson*.



(a) (b)

Fig. 5. Red symbol "**x**" indicates the geotagged position where each seabed photograph was taken. (a) Segment of AUV dive where seabed photograph was taken at an approximately 3-s interval. (b) Segment of AUV dive with irregular spaced seabed photographs due to variability in time taken for the camera's flash recharging.

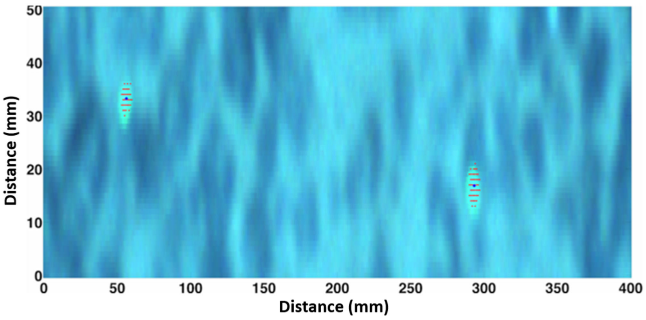


Fig. 6. Green laser detection in AUV image.

lawnmower pattern and collected approximately 5-km^2 area of seabed backscatter and bathymetry data.

C. Processing of Seabed Photographs

The primary objective of processing AUVs seabed photographs is to quantitatively estimate the coverage of the polymetallic nodules from high-resolution photographs captured by the AUV. During the dive, the AUV-mounted camera collected 5185 seabed photographs, of which approximately 3500 served as the ground-truth labels for our data set. We briefly present the details of the photograph analysis used to quantify PMN distribution.

The laser scaling system provides a visual scaling reference for the photographic imagery acquired. These lasers were mounted at distance of 200 mm and aimed into the field-of-view of the camera providing a scale bar for acquired images. Mounted in parallel 200 mm apart, these two lasers add scale to still images in the form of two green dots (as shown in Fig. 6) separated by a known distance. The two green dots in the image are extracted using an iterative *k*-means algorithm [22] and the distance between the cluster centroids is calculated. This process is repeated over a large set of images captured at various altitudes from the seabed and a relationship between the altitude of the AUV and the area covered by the image is obtained through polynomial curve fitting.

The photos were captured in Bayer-16 TIFF format, and then demosaicked into RGB48 TIFF. A sample of underwater photographs generated by an optical imaging suite after demosaicking is shown in Fig. 7(a). It is clearly evident that there is a significant falloff in light intensity in the images in Fig. 7(a) toward the corners. This is primarily attributed to the lighting setup. To enable an automated computational analysis of all images with the same setup (i.e., with one set of image processing parameters), we preprocessed the images to correct the varying illumination conditions. As all images feature a lightness falloff toward the corners, we subtracted a Gaussian filtered version of the image. To equalize the color contrast across images of the transect, histogram equalization techniques were applied. Thereby the peak of the histogram of each image was shifted toward the center of the intensity scale. The resulting contrast and illumination corrected grayscale photos are shown in Fig. 7(b).

The preprocessed grayscale seabed photographs are further processed for quantifying PMN distribution from optical images. This is done by identifying the PMN outlines and coverage area in the photographs using image processing, as shown in Fig. 7(c). In the photo illustrated in Fig. 7(c), the nodule coverage was 18.007% of the photographed image area. At the altitude of 8.05 m where the photo was captured, this translates to a nodule coverage area of 2.297 m^2.

Mero [5] states that the average oceanic coverage of economically acceptable PMN deposits is between 5 and 20 kg/m^2, whereas some other authors state that the minimum cutoff abundance for feasibility is 10 kg/m^2 [23], [24]. We set a threshold of 40% PMN coverage area in classifying all seabed photographs into two classes—high (seabed photographs depicting 40% or more PMN coverage area) or low (seabed photographs with PMN coverage area that are less than 40%) PMN abundance. This leads to a near-equal number of photographs for both classes and translates to a PMN density of around 23 kg/m^2 which is above the minimum range for economic viability. With this setting, 45% of the photographs are labeled as 1 (high PMN abundance) and the remaining 55% of the photographs are labeled as 2 (low PMN abundance).

D. Processing of SS Backscatter Data Set

SSS data collected are of binary file format. We process all SSS backscatter data into thirteen strips of waterfall GeoTIFF images. Each strip represents the along track length of the lawnmower path taken by the AUV as shown in Fig. 8(a). The size of each image strip is 1000-by-48 000 pixels which corresponds to an approximate seabed area of 100-by-4800 m.

Applying an appropriate image equalization technique [25] on our SSS data set is vital for achieving good performance using our data-driven approach. An ideal training image should possess features that can help our ANN model characterize regions of higher or lower PMN abundance. Insufficient contrast between regions of varying abundance can hinder the ability of the ANN to do so. This can result in a trained ANN with a comparatively lower performance.

From the visual patterns observed in our SSS data set, we observe that the dark and light contrasting segments depict seabed areas of low and high PMN abundance, respectively. To better enhance these contrasting segments, we applied contrast-limited adaptive histogram equalization (CLAHE) to our SSS data set [26].

E. ANN Data Set Preparation

ANNs have found widespread application in fields ranging from image classification to speech recognition [27]–[29]. The ANN's architecture comprises interconnected nodes inspired by biological structures of the animal brain. This method is able to distinguish unknown, nonlinear, and complex dependencies from a labeled data set. We investigate the effectiveness of using ANN as a modeling tool to learn the nonlinear mapping required to interpret PMN abundance from our data set features.

1) Preprocessing of SSS Data Set: As with any ML algorithm, a key factor affecting the performance of the ANN is the

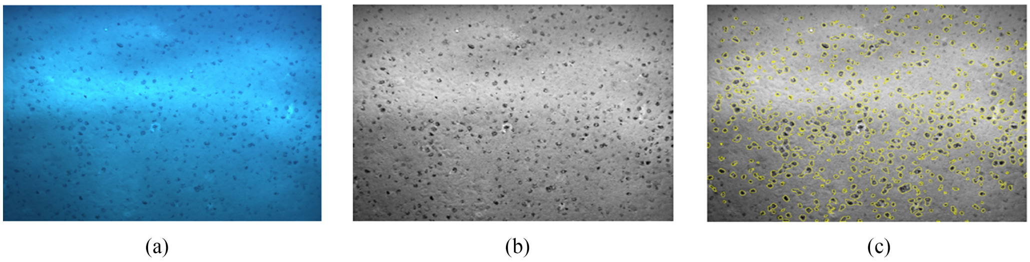


Fig. 7. (a) Seabed photograph with significant falloff in light intensity toward the corners. (b) Contrast and illumination-corrected photograph. (c) PMN outlines identified in a seabed photograph, which shows a PMN coverage of 18.07% that translates into a seabed area of 2.297 m^2.

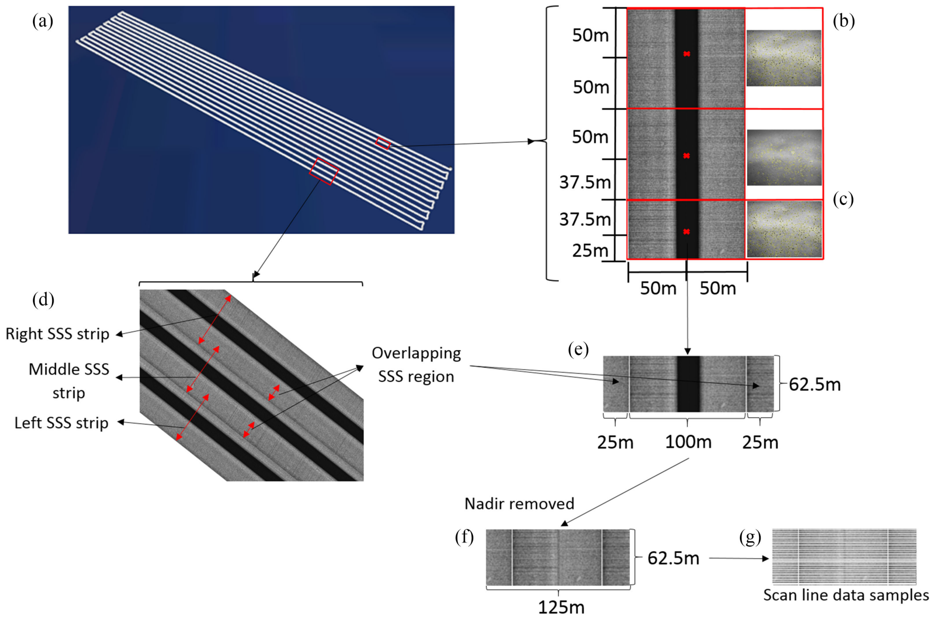


Fig. 8. Overview of data set preparation for ANN. (a) Preprogrammed lawnmower pattern route taken by the AUV during the environmental baseline seabed survey within our CCFZ's area of interest. SSS collected backscatter data from a seabed area of 5 km^2. (b) Segment of SSS backscatter data. Red symbol "x" indicates the position where each geotagged seabed photograph as shown on the right is taken. Depending on the location where each photograph was taken, PMN density seen at each geotagged photograph are uniform up to 50 m along track and 50 m across track the SSS image as shown by the red border. (c) At instances where photographs were taken at shorter length interval, the assumed PMN abundance for along track length would be less than 50 m and is set at equidistant between the two photographs. (d) Segment of three neighboring SSS backscatter image. Each SSS backscatter strip has a 25% overlapping SSS regions from the left and right adjacent run-length strips. (e) SSS backscatter image with extended features from neighboring strips, thus increasing the data set field of view. (f) Resultant SSS backscatter image after removal of nadir. (g) Multirow data samples are subdivided into single-row scan line data samples to increase the total number of data samples. (subfigures (b)–(d) reproduced from [21]).

quality of the data set used in training it. It is crucial that the training data set used be of the correct scale with meaningful features that are beneficial toward solving our formulated classification problem.

To investigate this, we synchronized all seabed photographs with the corresponding SSS images from the same locations. The seabed photographs collected within our ROI only covered a seabed area of approximately 0.042 km^2 (with each photograph depicting a seabed area of approximately 12 m^2) compared to the SSS imaged area of 5 km^2. In between locations where photographs were taken, no visual data were available. Such regions span up to 50 m along track and 50 m across track from the position where the photographs were taken. To be able to use the SSS data from these regions in training our method, we enhance our data set by assuming that the PMN abundance in these regions stays the same as the quantity estimated in the nearest photograph, up to a distance of 50 m. This is illustrated in Fig. 8(b). This is a reasonable assumption as studies have shown that on an average, only a 10% variation in PMN abundance is expected over every 450 m [7]. Using this assumption, we can now obtain photograph-based PMN abundance ground truths corresponding to the whole length of SSS backscatter, thus allowing us to maximize the use of this data.

2) Enhanced Data Set Field of View: The AUV's lawnmower path is programmed to have at least 25% overlap of seabed covered between each track-run as shown in Fig. 8(d). In our previous work, these overlapping regions were not used as part of our data preparation method. However, these overlapping regions covered by neighboring track-runs contain relevant PMN abundance information that can help improve the accuracy of our

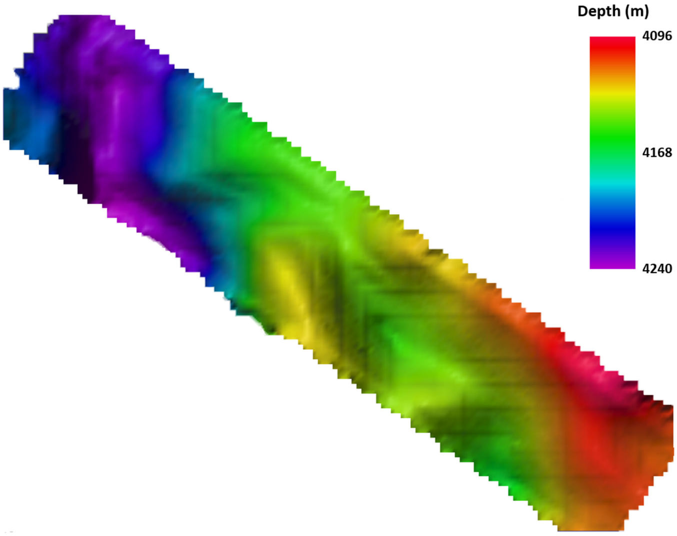


Fig. 9. Bathymetry GeoTIFF image from the 5-km^2 region of interest.

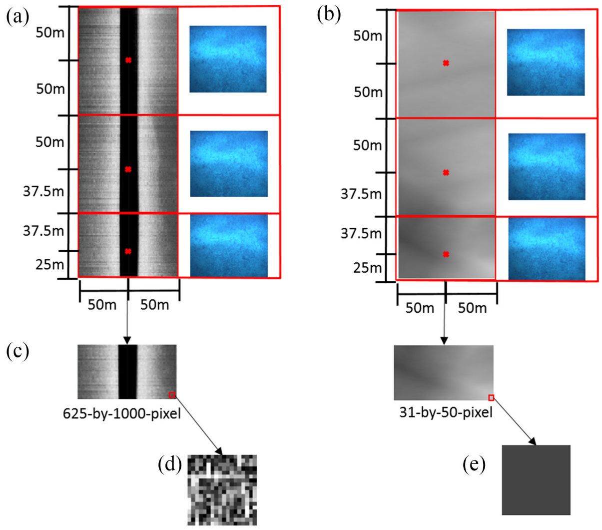


Fig. 10. (a) SSS seabed backscatter image and (b) bathymetry image from the same geotagged location. Similar to the assumption made with the SSS seabed backscatter data set, PMN density seen in each geotagged photograph is assumed to be uniform to its surrounding bathymetry pixels as indicated by the red boundary. (c) Although both SSS seabed backscatter and bathymetry images depict a seabed area of 6250 m^2, the number of pixel representation for both images are different as the seabed backscatter and bathymetry data were collected using different acoustic equipment. (d) 18-by-18-pixel SSS seabed backscatter image and (e) 1-by-1-pixel bathymetry image representing the same seabed area of 3.24 m^2.

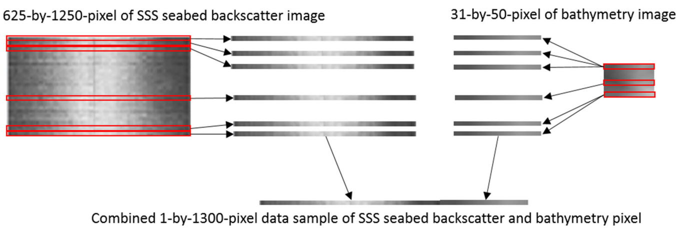


Fig. 11. Method for combining two data sets of different resolution representing the same seabed area into a heterogeneous data set.

trained ANN. To use this additional information, we appended the overlapping regions to the width of each run and used it as part of the feature set as shown in Fig. 8(e). The width of the run-overlap region of SSS backscatter appended on both sides is approximately 25 m (250 pixels). This increased the across track width from 100 to 150 m.

The nadir portion of the data which does not contain any useful information is removed as it would degrade the performance of the trained ANN. After this, the resultant data set yields a swath width of 125 m (1250 pixels) as shown in Fig. 8(f).

3) Improving Ratio of Data Sample Size to Feature Size: The pixels of the SSS data are the features used in our ANN training process. However, if our ANN are trained on SSS features of large dimension as seen in Fig. 8(f), an overfitting of the model would occur as the ratio of data samples to features size would be low. To avoid this, we separate each data sample into individual scan lines of 1-by-1250 pixels. This increases the number of data samples to 4 00 000 which increases our ratio of data samples to feature size, and allows us to identify high/low PMN abundance from individual scan lines of the SSS data.

The labels 1 or 2, indicating whether the PMN coverage corresponding to these data samples were low or high, are applied based on the photographs as previously discussed in Section II-E1.

4) Incorporating Bathymetry Information Into Feature Set: ANN performance can be enhanced by using diverse information available from heterogeneous features. Previous studies have shown that bathymetric variations are key considerations in determining the PMN presence and variability [6], [19]. Based on this, we decided to enhance the diversity of features used in our prediction by incorporating bathymetry information. A bathymetric map for the ROI is shown in Fig. 9. The bathymetry surrounding the location of each data sample is used as a feature along with the corresponding SSS backscatter features, as shown in Fig. 10(b).

The SSS and bathymetry data sets have different resolutions as shown in Fig. 10(c). Each sample used in modeling uses one scan line of SSS backscatter and the nearest row of bathymetry as features. One row of bathymetry spans 50 pixels (100 m) across track. Thus, the feature size of each data sample is increased to 1300 pixels after the inclusion of bathymetry data set as shown in Fig. 11.

The 4 00 000 labeled data samples obtained are separated into training (80% of total labeled data samples), validation (10%), and testing (10%) data sets. We use the training data set to tune the weights of the ANN model. This involves minimizing a cost function which represents the error between the labels and the network predictions. This is done by iteratively adjusting the interconnecting weights between all layers. The ANN requires sample training data to obtain insights into the underlying patterns and thus make meaningful predictions. If the number of training samples is too small, the ANN would not be exposed to sufficient information to adequately learn the dependencies between the features and their corresponding labels. The validation data set is used to prevent overfitting during training while the testing data set is used to evaluate the performance of

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

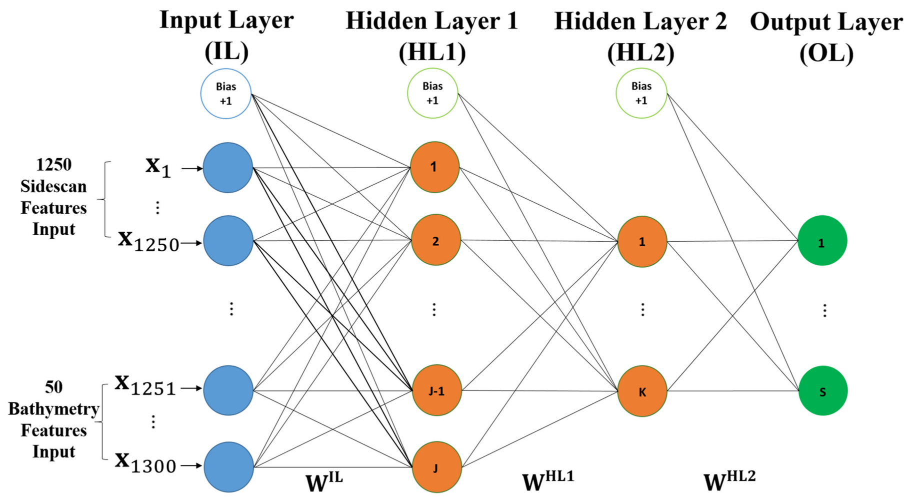


Fig. 12. Illustration of the used ANN architecture with two hidden layers. The notation x represents the input features, J represents the number of neurons in hidden layer 1, K represents the number of neurons in hidden layer 2, and S represents the number of output neurons.

our trained model. Note that the training, validation, and testing data sets are nonoverlapping. This ensures that the validation and test procedures yield unbiased evaluations of the generalization capability of the trained network.

5) ANN Architecture: We selected an ANN architecture with a two-hidden-layer feedforward topology network for our classification problem. The number of neurons in each layer affects ANN performance in a similar way as the training data size. Utilizing too few neurons will impede the ANN's ability to gain adequate insight from our training data set while utilizing too many neurons will result in the ANN overfitting. We choose the number of neurons used for hidden layers 1 and 2 (denoted as J and K, respectively) using the expression suggested by Huang [30], reproduced as follows:

$$J = \sqrt{(S+2)N} + 2\sqrt{N/(S+2)} \tag{1}$$

and

$$K = S\sqrt{N/(S+2)} \tag{2}$$

where N denotes the training data set size and S denotes the output layer size. Based on our training data set size, the number of neurons allocated are $J = 1800$ and $K = 600$. The ANN architecture is depicted in the schematic in Fig. 12. We chose the feedforward network architecture for its ability to learn the dependence of labels on features from a large number of data samples using a comparatively small number of hidden neurons [30].

In our mathematical computation, we denote matrices with bold upper case notations and vectors with bold lower case notations. For the training process, the heterogeneous data set is collectively treated as an N by $(I+1)$ matrix $\mathbf{X}$, where the $+1$ is for the bias, N denotes the number of data samples, and I denotes the number of features (image pixel columns) in each sample.

Let matrix $\mathbf{W}^p$ denote the collective weights between neighboring neuron layers, with notation p denoting the weight's starting layer. These are randomly initialized with values close to zero before the training process.

The ANN's training algorithm iterates through the data set matrix $\mathbf{X}$ by alternately applying the feedforward and back propagation passes on the ANN architecture, and a sigmoid activation function is used to generate a nonlinear decision output based on the weighted input. This is applied to the output of every neuron in hidden layer 1, hidden layer 2, and the output layer. The input data is normalized before applying the ANN weights, so that it does not saturate the activation function as shown in Fig. 12.

The randomly initialized weights of the ANN are trained using the backpropagation method by minimizing the cost function using conjugate gradient descent [31], [32]. The feedforward and backpropagation processes are repeated, and with each iteration the ANN weights are automatically readjusted to further minimize the cost function. The cost function used is the mean cross-entropy error between the predicted and actual labeled outputs [33].

The performance metric used to gauge the ANN's performance is its accuracy with the test data set. The effectiveness of the trained model is assessed in terms of accuracy between the predicted values and the given data set labels. Test accuracy is a measure of the ANN's ability to make generalized predictions with new data sets.

F. Preventing Overfitting

Iterating the ANN's feedforward and backpropagation training processes improves its performance on the training data set. However, training beyond a certain number of iterations can lead to an overfitted model. Overfitting occurs when the weight parameters are so explicitly tuned toward the training data set that they cannot be applied to any other data set of the same type. It is vital to avoid this and ensure that the trained ANN can generalize well to data outside the training data set. We monitor the generalization ability of the ANN by evaluating the

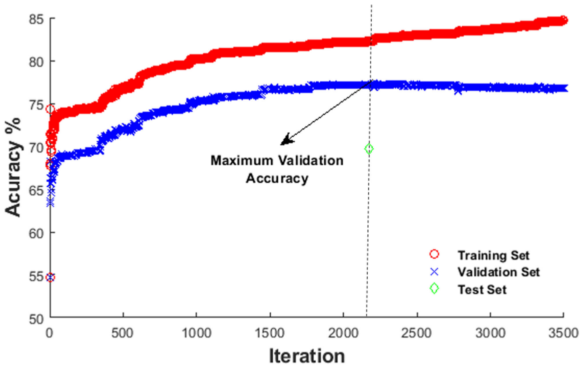


Fig. 13. Trained model can be seen overfitting on the training data set as the training iterations progress. With an increase in number of iterations, although there is an increase in training accuracy, a drop in validation data set accuracy can be seen after 2210th iterations. Thus, the ANN is likely overfitting on the training data set after the 2210th iteration. The ANN network weights chosen for testing are those obtained at the end of this iteration.

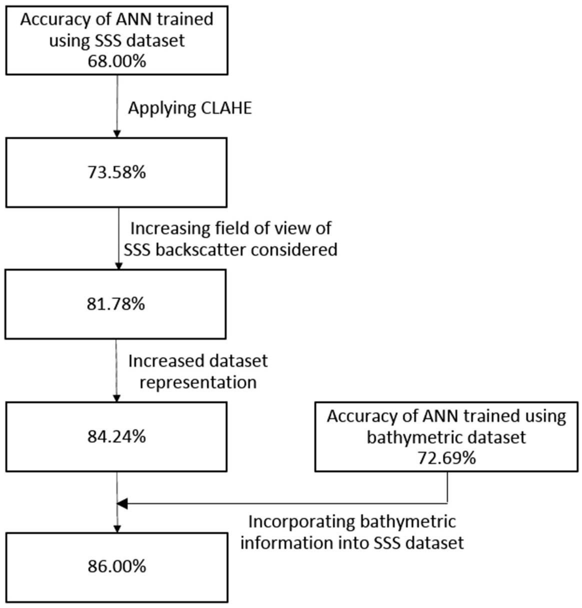


Fig. 14. Improvements in ANN accuracy due to each step of data preparation implemented.

trained model against the validation data set after every iteration of training. The final set of ANN weights chosen for testing is the one which yields the maximum performance with the validation data set. For the example shown in Fig. 13, the ANN weights chosen for testing are those obtained at training iteration 2210.

III. Results and Discussion

The key challenge in obtaining an accurate model is to prevent overfitting while ensuring predictive capability. This requires measures to ensure generalization of the network. These are rooted in domain expertise, a good understanding of the data set, and preprocessing to ensure that the trained model interprets this data correctly. To demonstrate the performance improvement from the various measures adopted by us, we present these comparative results in Fig. 14. The steps which contributed toward improving our ANN's performance are as follows.

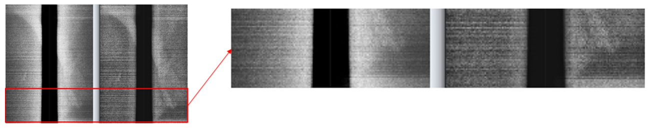


Fig. 15. (a) Comparison of a portion of SSS seabed backscatter images processed using conventional histogram equalization and CLAHE. (b) Enlarged view of SSS seabed backscatter images processed using conventional histogram equalization. (c) Enlarged view of SSS seabed backscatter images processed using CLAHE. Visual contrast between high and low PMN abundance areas is comparatively more prominent after CLAHE application.

1) Applying CLAHE for contrast adjustment: We choose CLAHE [26] over conventional histogram equalization as it adjusts individual pixels based on the histogram of their surrounding pixels. This works well on images with significant dark and light regions as it amplifies their contrasting features and enhances regional features. In Fig. 15, we compare conventional histogram equalization and CLAHE by applying them on a portion of the SSS seabed backscatter image. The figure shows that the CLAHE image exhibits greater contrast between high and low PMN abundance seabed areas when compared to conventional histogram equalization. Applying an appropriate image equalization method to the SSS images is vital for the accuracy of the ANN. An ideal training image should possess features with distinct contrasting segments depicting the traits of each class. Insufficiently contrasting segments or the absence of contrast can hinder the ANN's ability to discern features pertaining to each class during the training process, resulting in a lower accuracy.
2) Increasing field of view of SSS backscatter considered: As discussed earlier, for any sample in the data, we append to the SSS feature set a small portion of the backscatter data from the edges overlapping with the adjacent runs. This additional region also contains some information on the seabed area being assessed, namely a fresh perspective of the same area from a different viewing angle. Thus, appending it supplements our existing feature set with additional information, which further boosts the classification performance of our ANN.
3) Increased data set representation: Instead of using a large SSS backscatter image as a feature set in the modeling, we use a single scan line. This increases the size of the data set available for modeling and also reduces the feature space size, allowing faster training. It also reduces overfitting to some degree as the ratio of data to features increases.
4) Incorporating bathymetric information into the data set: In addition to the SSS images which we explored as features in a previous work [21], we also use depth information for each location being assessed. This was inspired from studies indicating the correlation of bathymetry to PMN presence [6], [19]. Using diverse features makes a richer set of information available to the ANN. Thus, adding bathymetric information into our feature set helps our trained ANN model make better abundance assessments. Our performance evaluation on the test data set showed that an ANN model trained using bathymetry features

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

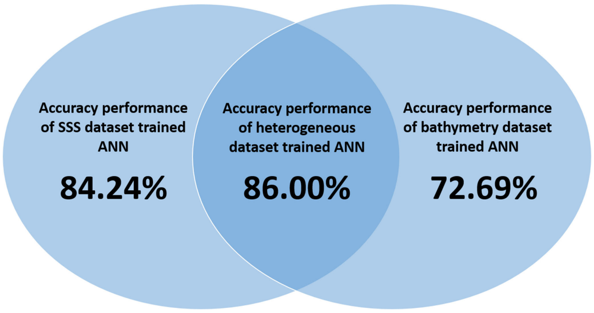


Fig. 16. Performance of ANN trained using SSS backscatter features, bathymetry, and a data set consisting of both.

alone yields an accuracy of 72.69%, and one trained using SSS backscatter features alone yields an accuracy of 84.24%. However, when both bathymetry and SSS backscatter are together used as features, the ANN's classification accuracy improves to 86.00%. This indicates that the SSS and bathymetry contain diverse information, and when these are used together the ANN is able to tap into this richer set of information and yield improved performance. This improvement is illustrated in the form of a Venn diagram in Fig. 16.

A. *Performance Evaluation*

The k-fold cross validation is a technique used to obtain a better evaluation of a trained model based on available data, especially in cases where the amount of data available is limited [34]. Here, we use the same approach for test performance evaluation, and refer to it as k-fold performance evaluation to reflect the purpose more accurately. k-fold performance evaluation involves training k different ANN models with the same hyperparameters. The available data samples are partitioned into k subsets. During the training of each ANN model, one subset is used as the testing data set while the remaining $k-1$ subsets serve as the training data set. This process is iterated for all k ANN models until every data subset is used once as the testing data set and $k-1$ times as part of the training data set. The overall accuracy of the ANN is computed as the average testing accuracy of the k trained ANN models. This provides a more reliable performance assessment as it reduces biases in the evaluation due to the random choice of the splitting of the data set alone. The bias is reduced because in the k-fold approach, the evaluation is based on not just one configuration, but by averaging across k different configurations of training, validation, and testing data sets.

After preprocessing, we evaluate the performance of the ANN architecture by iterating the training, validation, and testing discussed in Section II for tenfolds (i.e, $k = 10$). Each fold is based on a different randomly selected distribution of training (80%), validation (10%), and testing (10%) data sets. An example of one such fold is illustrated in Fig. 17. The accuracies of the ANN models selected in the tenfolds are tabulated in Table II. The accuracies ranged from 82.81% to 87.89%, yielding an average accuracy of 85.36%.

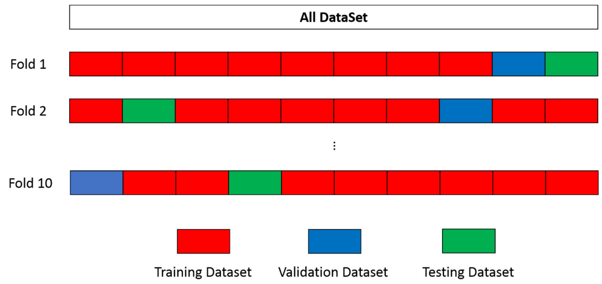


Fig. 17. Overview of the tenfold performance validation method using a toy example data set with ten data samples. All data samples are randomly split into training (80%), validation (10%), and testing (10%) data sets for each fold.

TABLE II
ACCURACY OF ANN MODELS TRAINED FROM TEN DIFFERENT FOLD CONFIGURATIONS OF TRAINING, VALIDATION, AND TESTING DATA SETS

Fold	**Accuracy(%) of ANN model**
1	86.00
2	85.19
3	87.89
4	85.41
5	84.70
6	82.81
7	84.02
8	86.88
9	85.21
10	85.52
Ten-fold Average	**85.36**

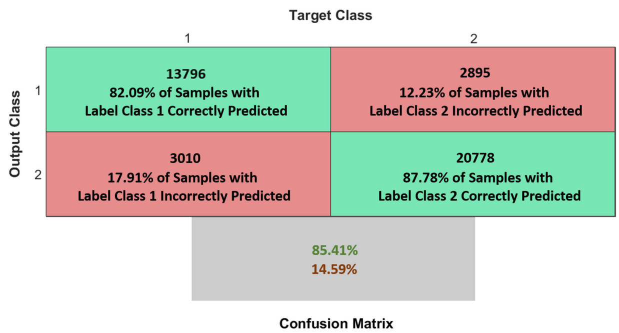


Fig. 18. Confusion matrix for the single-input ANN's fourth fold test result. The two green cells represent the number and percentage of correct classification prediction made, whereas the red cells represent the number and percentage of incorrect classification prediction made. Finally, the gray cell represents the ANN overall test accuracy for the fourth fold.

The classification performance of our ANN approach can be visualized in terms of a confusion matrix on the test data. It presents a visual summary of the ANN's correct and wrong classifications for each class. The confusion matrix for the 4th fold test result is illustrated in Fig. 18. It can be seen that our ANN model correctly classifies the majority of the test data samples into high/low PMN abundance categories with no significant bias toward a particular class.

B. Summary

The ANN training algorithms implemented in this article are coded in MATLAB R2017b. Currently, it takes around 30 h for the ANN to train approximately 360 000 data samples (80% of total data set) on a workstation with a dual-processor Intel Xeon E5-2630 V3 CPU@2.4 GHz processor.

The inputs consist of a heterogeneous feature set of SSS backscatter measurements and bathymetry data. The data set contains 400 000 samples, corresponding to a seabed area of approximately 5 km^2. Each sample consists of 1300 features containing both SSS backscatter and bathymetry, and is labeled with either 1 or 2 denoting high or low PMN abundance. The label is determined from PMN abundance estimates from seabed photographs as discussed earlier in Section II-C. The available data samples are randomly split into 80% training, 10% validation, and 10% testing data set for each fold of a tenfold performance evaluation. The two-hidden-layer ANN model used in this article comprised 1800 and 600 neurons for hidden layers 1 and 2, respectively. The number of neurons for each hidden layer was selected based on the number of training data samples and classification outputs. The PMN abundance ground-truth data are taken from 3500 seabed photographs, covering a total seabed area of 0.42 km^2. The SSS backscatter data used in the 400 000 samples encompasses a total seabed area of 5 km^2. To aid the training process of the ANN, various techniques such as CLAHE which enhanced the contrast between high and low PMN coverage area are applied.

IV. Conclusion

This article described a methodology to assess PMN abundance over a seabed area using an ANN model trained using a combined SSS and altimeter data sets. The model presented here improves upon what we presented in our previous work [21], yielding a test accuracy of 85.36% in assessing PMN abundance. This improvement was achieved by expanding the feature set to include bathymetric information that had been collected in the same data collection run.

The technique presented herein successfully demonstrates that the use of an ANN model to interpret a combined data set from different sensors enables an efficient assessment of large seabed areas. The model's result was also verified through k-fold cross validation to assess its reliability.

The ANN model trained in this work has not been tested with data from other nodule-bearing regions. However, pending future deployment opportunities, the methodology discussed in this article could enable faster assessments of larger seabed areas without the need for underwater cameras. A model trained using one particular data set may also be helpful in making assessments using other data sets if we can effectively tap into the power of transfer learning or domain adaptation [35]. This may be possible even if the other data sets were made at differing experimental conditions, provided that similar information at comparable scales is available in the features. Undertaking similar quantifications using data of a very different scale—say, from a ship-mounted sonar, might require significant modifications in the approach. This is because ship-mounted sonar would offer data of worse resolution as compared to an AUV-mounted sonar. This resolution would not be sufficient to adjudge the small-scale variations in nodule density as we have done in this study. Furthermore, we would not be able to validate them against seabed photographs as we have done in this study, as the photographs are also at small scales comparable to the AUV-mounted SSS.

To make our ANN technique more universal in assessment, a modeler would need to incorporate the difference in operating conditions of the AUV, and the acoustic recording equipment used and its settings. In the current study, we have not delved into what specific features or details of the SSS scatter data and bathymetry are being picked up by the ANN, or how these features are being interpreted. This in-depth study could make for an interesting future work.

Acknowledgment

The authors would like to thank the National Research Foundation, Keppel Corporation, and the 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," *Mar. Geol.*, vol. 98, no. 2, pp. 411–423, 1991.

[2] International Seabed Authority, "A geological model of polymetallic nodule deposits in the Clarion Clipperton fracture zone," Int. Seabed Authority, Kingston, Jamaica, Tech. Rep. 6, 2010.

[3] E. Lange, S. Petersen, L. Rüpke, E. Söding, and K. Wallmann, "Marine resources—Opportunities and risks," *World Ocean Rev.*, vol. 3, ch. 2, pp. 54–93, 2014.

[4] T. Kuhn, A. Wegorzewski, C. Rühlemann, and A. Vink, "Composition, formation, and occurrence of polymetallic nodules," in *Deep-Sea Mining: Resource Potential, Technical and Environmental Considerations*, R. Sharma, Ed. Berlin, Germany: Springer, 2017, ch. 2, pp. 23–63.

[5] J. L. Mero, "Economic aspects of nodule mining," in *Marine Manganese Deposits*, G. Glasby, Ed. Amsterdam, The Netherlands: Elsevier, 1977, ch. 11, pp. 327–355.

[6] 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," *Mar. Georesources Geotechnol.*, vol. 19, no. 3, pp. 167–180, 2001.

[7] R. Kaufman, "The selection and sizing of tracts comprising a manganese nodule ore body," in *Proc. Offshore Technol. Conf.*, Houston, TX, USA, 1974, pp. 283–293.

[8] M. John and A. F. Renard, "Report on deep-sea deposits based on the specimens collected during the voyage of H.M.S. challenger in the years 1873 to 1876," 1891, ch. 1, pp. 1–32.

[9] M. Atmanand and G. Ramadass, "Concepts of deep-sea mining technologies," in *Deep-Sea Mining: Resource Potential, Technical and Environmental Considerations*, R. Sharma, Ed. Berlin, Germany: Springer, 2017, ch. 10, pp. 305–343.

[10] P. Stoffers, A. Sioulas, G. Glasby, and T. Thijssen, "Geochemical and sedimentological studies of a box core from the western sector of the Peru basin," *Mar. Geol.*, vol. 49, no. 3/4, pp. 225–240, 1982.

[11] V. N. Hari, B. Kalyan, M. Chitre, and V. Ganesan, "Spatial modeling of deep-sea ferromanganese nodules with limited data using neural networks," *IEEE J. Ocean. Eng.*, vol. 43, no. 4, pp. 997–1014, Oct. 2018.

[12] G. Glassy and R. Singleton, "Underwater photographs of manganese nodules from the southwestern pacific basin," *New Zealand J. Geol. Geophys.*, vol. 18, no. 4, pp. 597–604, 1975.

[13] R. Sharma, "Quantitative estimation of seafloor features from photographs and their application to nodule mining," *Mar. Georesources Geotechnol.*, vol. 11, no. 4, pp. 311–331, 1993.

[14] R. Sharma, S. J. Sankar, S. Samanta, A. Sardar, and D. Gracious, "Image analysis of seafloor photographs for estimation of deep-sea minerals," *Geo-Mar. Lett.*, vol. 30, no. 6, pp. 617–626, 2010.

[15] M. Okazaki and A. Tsune, "Exploration of polymetallic nodule using AUV in the central equatorial pacific," in *Proc. 10th ISOPE Ocean Mining Gas Hydrates Symp.*, 2013, pp. 32–38.

[16] 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," *Int. J. Remote Sens.*, vol. 34, no. 5, pp. 1691–1706, 2013.

[17] 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," in *Proc. MTS/IEEE OCEANS Conf.*, 2015, pp. 1–5.

[18] N. C. Mitchell and J. E. H. Clarke, "Classification of seafloor geology using multibeam sonar data from the Scotian Shelf," *Mar. Geol.*, vol. 121, no. 3/4, pp. 143–160, 1994.

[19] 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," *Mar. Georesources Geotechnol.*, vol. 22, no. 1/2, pp. 103–114, 2004.

[20] M. Weydert, "Measurements of the acoustic backscatter of selected areas of the deep seafloor and some implications for the assessment of manganese nodule resources," *J. Acoust. Soc. Amer.*, vol. 88, no. 1, pp. 350–366, 1990.

[21] W. L. Jie, B. Kalyan, M. Chitre, and H. Vishnu, "Polymetallic nodules abundance estimation using sidescan sonar: A quantitative approach using artificial neural network," in *Proc. OCEANS Conf.*, Aberdeen, Scotland, 2017, pp. 1–6.

[22] J. MacQueen, "Some methods for classification and analysis of multivariate observations," in *Proc. 5th Berkeley Symp. Math. Stat. Probab.*, vol. 1, no. 14, 1967, pp. 281–297.

[23] G. P. Glasby, "Deep seabed mining: Past failures and future prospects," *Mar. Georesources Geotechnol.*, vol. 20, no. 2, pp. 161–176, Apr. 2002.

[24] T. Marjoram, H. Cameron, G. Ford, A. Garner, and M. Gibbons, "Manganese nodules and marine technology," *Resources Policy*, vol. 7, no. 1, pp. 45–57, Mar. 1981.

[25] Y. Jin, L. M. Fayad, and A. F. Laine, "Contrast enhancement by multiscale adaptive histogram equalization," in *Wavelets: Applications in Signal and Image Processing IX*, vol. 4478, 2001, pp. 206–214.

[26] S. M. Pizer *et al.*, "Adaptive histogram equalization and its variations," *Comput. Vis., Graph., Image Process.*, vol. 39, no. 3, pp. 355–368, 1987.

[27] J. Schmidhuber, "Deep learning in neural networks: An overview," *Neural Netw.*, vol. 61, pp. 85–117, 2015.

[28] D. Cireşan, U. Meier, J. Masci, and J. Schmidhuber, "Multi-column deep neural network for traffic sign classification," *Neural Netw.*, vol. 32, pp. 333–338, 2012.

[29] A. Hannun *et al.*, "Deep speech: Scaling up end-to-end speech recognition," 2014, *arXiv:1412.5567*.

[30] G.-B. Huang, "Learning capability and storage capacity of two-hidden-layer feedforward networks," *IEEE Trans. Neural Netw.*, vol. 14, no. 2, pp. 274–81, Mar. 2003.

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

[32] R. Fletcher and C. M. Reeves, "Function minimization by conjugate gradients," *J. Optim. Theory Appl.*, vol. 7, no. 2, pp. 149–154, 1964.

[33] P.-T. De Boer, D. P. Kroese, S. Mannor, and R. Y. Rubinstein, "A tutorial on the cross-entropy method," *Ann. Oper. Res.*, vol. 134, pp. 19–67, 2005.

[34] M. Stone, "Cross-validatory choice and assessment of statistical predictions," *J. Roy. Stat. Soc. B*, vol. 36, pp. 111–147, 1974.

[35] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Trans. Knowl. Data Eng.*, vol. 22, no. 10, pp. 1345–1359, Oct. 2010.

Liang Jie Wong received the B.Tech. degree in electronics from the National University of Singapore, Singapore, in 2011.

From 2011 to 2015, he was a Project Engineer with the Defence Science and Technology Agency, where he was involved in acquisition management activities for naval sensor related projects. In late 2015, he joined the Acoustic Research Laboratory, National University of Singapore, as a Research Engineer. His research interests include image processing, data analysis, deep seabed mining and tailoring of ML methods toindustrial needs.

Bharath Kalyan (Member, IEEE) received the Ph.D. degree in autonomous underwater vehicle navigation from Nanyang Technological University (NTU), Singapore, in 2010.

He is currently with the Acoustic Research Laboratory (ARL), Tropical Marine Science Institute, National University of Singapore, Singapore, where he manages myriad projects on deep sea mining and marine robotics. Prior to joining ARL, he was a Research Staff member with the Intelligent Robotics Lab, Department of Electrical and Electronic Engineering, NTU, where he worked on various underwater projects, primarily focusing on underwater sensing and navigation for surface and undersea vehicles. His research interests include ML, deepsea mining, co-operative marine robotics, AUV navigation, target tracking, and random finite sets.

Mandar Chitre (Senior Member, IEEE) received the B.Eng. and M.Eng. degrees in electrical engineering from the National University of Singapore (NUS), Singapore, in 1997 and 2000, respectively, the M.Sc. degree in bioinformatics from Nanyang Technological University, Singapore, in 2004, and the Ph.D. degree in underwater communications from NUS, in 2006.

From 1997 to 1998, he worked with the Acoustic Research Laboratory, NUS. From 1998 to 2002, he was the Head of the Technology Division of a Regional Telecommunications Solutions Company. In 2003, he rejoined ARL, initially as the Deputy Head (Research) and is now the Head of the Laboratory. He also holds a joint appointment with the Department of Electrical and Computer Engineering, NUS, as an Associate Professor. His research interests include underwater communications, autonomous underwater vehicles, model-based signal processing, and modeling of complex dynamic systems.

Hari Vishnu (Member, IEEE) received the Ph.D. degree in computer engineering from Nanyang Technological University, Singapore, for research on improved underwater signal processing algorithms for noisy waters.

He has been a Research Fellow with the Acoustic Research Laboratory, National University of Singapore, since 2013. He has been a reviewer for several reputed international journals and conferences. His research interests and project involvements include machine-learning for marine applications, sonar, and signal processing, especially for cryoacoustic and bioacoustic applications.

Dr. Vishnu is a recipient of the IEEE OES young-professionals-BOOST award, and currently an editor of the science outreach magazine *Earthzine*, the Chair of the IEEE Oceanic Engineering Society, Singapore chapter and serves on the IEEE Singapore AUV Challenge Committee.