POLYMETALLIC NODULE ABUNDANCE ESTIMATION USING SIDESCAN SONAR: A QUANTITATIVE APPROACH USING ARTIFICIAL NEURAL NETWORK

Wong Liang Jie

(B. Tech. (Hons.), NUS)

A THESIS SUBMITTED FOR THE DEGREE OF MASTER OF ENGINEERING DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING NATIONAL UNIVERSITY OF SINGAPORE

 $\mathbf{2019}$

SUPERVISOR: ASSOCIATE PROFESSOR MANDAR ANIL CHITRE

EXAMINERS: DR FENG JIASHI ASSOCIATE PROFESSOR HONG GEOK SOON

DECLARATION

I hereby declare that the thesis is my original work and it has been written by me in its entirety.

I have duly acknowledged all the sources of information which have been used in the thesis.

This thesis has also not been submitted for any degree in any university previously.

1

Wong Liang Jie 25th March 2019

"You know, I couldn't do it. I couldn't reduce it to the freshman level. That means we really don't understand it."

Richard P. Feyman

Acknowledgment

The research done in this thesis is made possible under the kind guidance of Prof. Mandar Chitre, Dr. Bharath Kalyan and Dr. Hari Vishnu. Without their invaluable suggestions and comments, the completion of this thesis would not be possible.

Contents

Abstract viii			viii
Li	st of	Tables	ix
Li	st of	Figures	x
1	Intr	roduction	1
	1.1	Motivation	2
	1.2	Approach	4
	1.3	Thesis Layout	5
2	Bac	kground	7
	2.1	Existing PMN Abundance Estimation	
		Methods	8
	2.2	Study Site	10
	2.3	Equipment	10
	2.4	Data Collection	11
	2.5	Processing of Sidescan Sonar Image	13
	2.6	Processing of Seabed Photographs	14
3	Met	thodology	16
	3.1	Artificial Neural Network Architecture	16
	3.2	ANN Dataset Preparation	18
	3.3	Training Algorithm	20

4	Dat	Dataset Features Analysis 2	
	4.1	4.1 Contrast Limited Adaptive Histogram	
		Equalisation (CLAHE)	29
	4.2	Enhanced Dataset Field of View	31
	4.3	Enhanced Dataset Representation	33
	4.4	Heterogeneous Dataset	35
	4.5	Results	39
	4.6	Summary	45
5	Conclusion		48
	5.1	Summary of Contributions	48
	5.2	Future Work	49
	5.3	Final Thoughts	50
Bi	Bibliography 51		

Abstract

Scattered abundantly across the vast regions of the Clarion and Clipperton Fracture Zone (CCFZ) are pockets of Polymetallic Nodules (PMN). These PMN possess high economic potential as they are rich in minerals such as manganese, nickel, copper and rare earth elements. Quantification of such PMN coverage is important for economic feasibility studies and planning of exploitation strategies. Traditional methods for PMN quantification are labour and time intensive as they rely on freefall box corer measurements and/or image processing of seabed photographs. This research thesis explores PMN abundance estimation using a data-driven method based on Artificial Neural Network (ANN). Data used are geotagged Sidescan Sonar (SSS) seabed backscatter images and seabed photographs collected using an Autonomous Underwater Vehicle (AUV) within the CCFZ. Compared to an underwater camera, the SSS provides a much larger area of coverage, effectively increasing the AUV's efficiency in the task of seabed surveying within the limited dive-time. This is the first known published work to elaborate on a data-driven approach in estimating PMN abundance using SSS seabed backscatter data. The trained ANN model yielded an average accuracy performance of 85.36%, demonstrating that it can be an effective tool in estimating PMN abundance from SSS seabed backscatter images. This approach enables faster evaluation of PMN abundance for future deep seabed exploration without the need for underwater cameras.

List of Tables

3.1	Accuracy performance of the ANN model with the training	
	and testing datasets. Improvement in accuracy performance	
	with the training dataset can be seen with a larger num-	
	ber of training iterations. Overfitting of the ANN model to-	
	wards the training dataset is evident from the decrease in the	
	model's accuracy performance towards the testing dataset	
	beyond 16,000 training iterations	26
11	Accuracy performance of two different ANN architectures	
4.1	Accuracy performance of two different Arviv arenitectures,	
	trained from ten configurations of training, validation and	
	testing datasets.	44

List of Figures

1.1	Photograph of naturally occurring PMN on the seabed surface.	1
1.2	$50\times50~\mathrm{cm}$ box corer grab (Inset: Sample of recovered sedi-	
	ments and PMN from the box corer) $\ldots \ldots \ldots \ldots$	3
1.3	A segment of the SSS seabed backscatter data. Over a times-	
	pan of 12 seconds, the AUV covered a distance of $17.28\mathrm{m}$	
	with an average speed of $2.8\mathrm{knots}.$ The AUV-mounted SSS	
	collected seabed backscatter data from a seabed area of $1728\mathrm{m}^2$	
	while the AUV-mounted camera only photographed a seabed	
	area of $36 \mathrm{m}^2$. Due to the time required for flash recharging	
	between each photograph, only three seabed photographs	
	(each capturing a seabed area of $12\mathrm{m}^2)$ depicted by red 'X's	
	were taken during this period.	4
1.4	A simple depiction of the problem statement, which is to	
	correlate PMN abundance seen in seabed photographs to	
	SSS seabed backscatter data	5
2.1	The extent of the CCFZ is demarcated by the blue dashed	
	lines. The Singapore licensed zone for PMN exploration is	
	shown within the red circle.	10
		10

2.2	A pre-programmed lawnmower pattern route taken by the	
	AUV during the environmental baseline seabed survey within	
	a CCFZ's seabed of interest. An AUV-mounted SSS col-	
	lected seabed backscatter data from a seabed area of ap-	
	proximately $5 \mathrm{km}^2$.	11
2.3	Symbol ' \mathbf{x} ' indicates position where each seabed photograph	
	was taken. (a) Segment of dive where seabed photograph	
	was taken at approximately three-second interval. (b) Due	
	to unforeseen variability in time taken for the camera's flash	
	recharging, there were instances where seabed photograph	
	was taken at intervals that exceeded three seconds. \ldots .	12
2.4	Swath-width of the SSS seabed backscatter image with nadir	
	portion removed.	14
2.5	PMN identified as region of interest on both seabed pho-	
	tographs. (a) Photograph depicting a seabed with a low	
	PMN coverage of 5.674% translating to an area of $0.682\mathrm{m^2}.$	
	Photograph depicting a seabed area with a high PMN cov-	
	erage of 61.351% translating to an area of $7.585\mathrm{m}^2.$	15
3.1	A two-hidden-layer ANN architecture where ' I ' represents	
	the number of features for each training sample, 'S' repre-	
	sents the number of output neurons and $`N"$ represents the	
	number of training samples. The number of neurons used	
	for hidden layer 1 and 2 respectively are denoted as 'J' and	
	'K' respectively	17

3.2	(a) Pre-programmed lawnmower path taken by the AUV.
	(b) A segment of the SSS seabed backscatter data, the AUV-
	mounted camera took a seabed photograph every 3 seconds
	at every location marked by symbol 'x'. PMN density seen at
	each geotagged seabed photograph is assumed to be uniform
	within the $4\mathrm{m}$ along-track and $75\mathrm{m}$ across-track of the SSS
	seabed backscatter image
3.3	Multi-row data samples are sub-divided into single-row data
	samples to increase the total number of data samples 20 $$
3.4	Figure depicting the problem of overfitting as the ANN train-
	ing iteration number increases. Although the training dataset
	accuracy increases with the increasing number of iterations,
	there is also a gradual drop in validation dataset accuracy
	after 2250 iterations. Thus, the ANN is overfitting on the
	training dataset after 2250 iterations. The ANN's intercon-
	necting weights between the neurons is taken at this iteration
	where it is evaluated on the unseen testing dataset 27
4.1	CLAHE enhancement of a generic image (right image) has
	a more prominent contrast between its darker and lighter
	shade segments when compared to the original image (left
	image)
4.2	(a) Comparison of a portion of SSS seabed backscatter im-
	ages processed using Histogram Equalisation and CLAHE.
	(b) Enlarged view of SSS seabed backscatter images pro-
	cessed using Histogram Equalisation. (c) Enlarged view of
	SSS seabed backscatter images processed using CLAHE. Vi-
	sual contrast between high and low PMN abundance area is
	more prominent after CLAHE application

xii

- 4.4 Overlapping region from the left and right adjacent runlength paths increases the column size of the dataset from 750 to 1250 pixels This dataset is then further separated into single-row data samples as previously discussed in Section 3.2. 32

- 4.7 Bathymetry GeoTIFF image from $5\,{\rm km^2}$ region of interest. . 35

- 4.8 (a) Data collection path of the AUV. (b) SSS seabed backscatter image and (c) bathymetry image from the same geotagged location. Similar to the assumption made with the SSS seabed backscatter dataset, PMN density seen in each geotagged photograph is assumed to be uniform to its surrounding bathymetry pixels as indicated by the red boundary. (For illustration purpose, not drawn to scale) 36

4.10 Method for combining two datasets of different resolutionrepresenting the same seabed area into a heterogeneous dataset. 38

- using a toy example dataset with ten data samples. All data samples are randomly split into training (80%), validation (10%) and testing (10%) datasets for each fold. 41
- 4.14 Training iteration graphs of a dual-input ANN architecture using ten variations of randomly distributed training, validation and testing data samples.

- 4.15 Confusion matrix for the single-input ANN's 4th fold test result. The two green cells represent the number and percentage of correct classification prediction made, while the red cells represent the number and percentage of incorrect classification prediction made. Lastly, the grey cell represents the ANN overall test accuracy for the 4th fold. 45

List of Notations

b	Scalar
b	Vector
В	Matrix
\mathbf{B}^{T}	Transpose of matrix \mathbf{B}
$\exp(\mathbf{B})$	Exponential of matrix \mathbf{B}
$\ln(\mathbf{B})$	Natural logarithm of matrix ${f B}$
\odot	Element-wise multiplication
\Leftrightarrow	Element-wise equivalence comparison with logical output
N	Number of training samples
n	n^{th} training sample
Ι	Number of features in each training sample
i	i^{th} feature of the dataset
S	Number of output neurons
S	Output value of s^{th} neuron at output layer
X	Input dataset
x_{ni}	i^{th} feature from n^{th} training sample
Y	Labelled output
y_{nOPs}	\boldsymbol{n}^{th} training sample's labelled output value at \boldsymbol{s}^{th} neuron
\mathbf{W}^p	Collective weight matrix originating from layer 'p'
J	Number of neurons in hidden layer '1'
K	Number of neurons in hidden layer '2'

List of Abbreviations

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
ISA	International Seabed Authority
PMN	Polymetallic Nodules
SSS	Sidescan Sonar
UNCLOS	United Nations Convention on the Law of the Sea

Chapter 1

Introduction

The International Seabed Authority (ISA) is an inter-governmental organisation established under the United Nations Convention on the Law Of the Sea (UNCLOS). It serves as the authority for regulating all seabed mineral-related activities in international waters.

In the year 2015, ISA licensed to Singapore, a seabed area of $58,000 \text{ km}^2$ within the CCFZ for mineral deposits exploration. These mineral deposits take the form of potato-sized concretions called PMN. Past explorations have found these PMN to reside mostly on the surface or just beneath the abyssal seabed of all major oceans as shown in Figure 1.1 [1].



Fig. 1.1. Photograph of naturally occurring PMN on the seabed surface.

Studies have documented the occurrence and high abundance of such PMN on the CCFZ's abyssal seabed, and it is estimated that there are 21 billion tons of PMN within the CCFZ [1, 2, 3]. These PMN are an alternative source of minerals from their land-based counterparts, offering potentially high economic value from their rich contents of manganese, copper, nickel, and cobalt [4]. This development is exciting for a city-state like Singapore, which has no natural resources to speak of and serves to offer an excellent opportunity for the country to further diversify and expand its economy through these potential metal markets.

Due to the vastness of the seabed area licensed to Singapore, there are no efficient sampling and estimation methods which can accurately quantify PMN abundance on such an extensive scale. The methods discussed in this thesis improves the efficiency and accuracy of existing methods by using a labelled SSS seabed backscatter data-driven approach. Based on this proposed method, PMN abundance estimation over large seabed area can be more efficiently and accurately assessed.

1.1 Motivation

A prospecting agency's ability to effectively and accurately quantify PMN abundance over a large seabed area is crucial for the purpose of economic feasibility studies and development of effective PMN exploitation strategies. Studies have reported PMN distribution exhibiting considerable variability within a span of kilometres across the CCFZ with higher abundance in the central and north-eastern regions [3, 5, 6]. Thus, it is vital to have an estimation method which gives an accurate portrayal of PMN abundance and its variation within any licensed prospecting seabed to ensure a proper economic evaluation of these minerals.

Traditional PMN estimation methods are highly laborious and timeconsuming as they rely on planimeter and point-counting of PMN collected from various forms of seabed sampling devices such as freefall grab and box corer as shown in Figure 1.2 [7, 8]. Unless conducted extensively over large seabed area, these methods produced pockets of sparse sampling points which may suffice for estimating PMN abundance within small seabed area, but performing interpolation using these sparse measurements across the vast prospecting seabed area could result in a poor estimate of PMN abundance [9].

A more recent method uses image processing of seabed photographs captured with an AUV-mounted camera. Due to the comparatively higher efficiency in capturing seabed data, this technique has gained significant traction as the preferred method for quantifying PMN abundance [6, 10, 11, 12]. High-performance computers with efficient image processing algorithms running on graphics-processing units are also used to estimate PMN abundance from the large number of seabed photographs collected. [13]. However, the total seabed area photographed using an AUV-mounted camera during each dive is still relatively small compared to the extent of the entire licensed seabed area to be surveyed as shown in Figure 1.3.



Fig. 1.2. 50×50 cm box corer grab (Inset: Sample of recovered sediments and PMN from the box corer)



Fig. 1.3. A segment of the SSS seabed backscatter data. Over a timespan of 12 seconds, the AUV covered a distance of 17.28 m with an average speed of 2.8 knots. The AUV-mounted SSS collected seabed backscatter data from a seabed area of 1728 m^2 while the AUV-mounted camera only 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 red 'X's were taken during this period.

Advancements in underwater sensing technology have led to the use of underwater acoustic equipment in various seabed applications, such as performing seabed terrain classification using a multibeam sonar and a relationship study between seabed characteristics and PMN abundance using a SSS [14, 15]. Although studies have suggested the existence of a qualitative relationship between acoustic seabed backscatter returns and PMN abundance [15, 16, 17], there is no published work that details a datadriven approach in using SSS seabed backscatter data for PMN abundance estimation thus far.

1.2 Approach

A visual comparison between the seabed photographs and SSS seabed backscatter data collected from the Singapore licensed seabed area show that PMN abundance variations from the two datasets are indeed correlated as reported by [15, 16, 17]. Based on this observation, this thesis explores the use of texture variations observed in SSS seabed backscatter data to estimate seabed PMN abundance as shown in Figure 1.4.



Fig. 1.4. A simple depiction of the problem statement, which is to correlate PMN abundance seen in seabed photographs to SSS seabed backscatter data.

This is achieved by training the ANN, which is a supervised machine learning technique for solving classification problems to model the relationship between the patterns found in the SSS seabed backscatter data and the PMN seabed coverage seen from its corresponding geotagged photographs. Once trained, the ANN can be used to estimate future sites' PMN abundance with greater efficiency and accuracy using only SSS seabed backscatter data.

1.3 Thesis Layout

The results from this thesis, descriptions of all methods applied, along with their respective accuracy improvements are discussed in an order based on a sequential workflow methodology. Conceptual illustrations of the various methods used are presented alongside the discussions to help the readers understand the rationale of these methods.

Chapter 2 provides a review on existing PMN data collection and abundance estimation methods. The thesis's background information such as the geographical area of study, data collection, and data processing methods are also discussed in this chapter. Chapter 3 discusses the general ANN training procedures together with the trained model's accuracy performance on unseen SSS seabed backscatter dataset. Chapter 4 discusses the implementation of various additional methods which further improves the model accuracy performance and the methods used in assessing the overall reliability and the generalisation capability of the trained model. Lastly, Chapter 5 presents the conclusion drawn from this thesis as well as the possible area of future research.

Chapter 2

Background

In recent years, PMN containing commercially viable minerals such as manganese, nickel, copper and cobalt have increasingly drawn economic interest from the mining industry [3, 18]. Although Murray and Renard discovered the first CCFZ's PMN deposits during their scientific expedition on board H.M.S Challenger in the early 1870s [19]. It was only in the 1960s, that reports on the abundance of such PMN deposits within the CCFZ spurred the economic interest towards these deposits [4, 20]. Since then, numerous country-led consortia have conducted exploration studies to determine the extent of these PMN abundance and the viability of harvesting minerals from these PMN deposits [21]. To show the rationale of using the proposed data-driven approach in inferring PMN abundance through SSS seabed backscatter data, selected literatures with an emphasis on existing PMN sampling and abundance estimation methods are discussed in this chapter.

2.1 Existing PMN Abundance Estimation Methods

In the early days of PMN exploration, Murray and Renard [19] obtained coarse estimations of PMN abundance on small seabed areas using an extensive range of labour intensive seabed collection methods such as clamshell sampler, dredge haul and gravity core. Using these methods, chunks of seabed are brought up to the surface, and the sediments are sieved to extract PMN which are then weighed to determine their abundance. However, such methods only sampled a small seabed area per deployment with sampling points potentially spaced kilometres apart. This adds uncertainty on the variations in PMN abundance around each sampled point, making PMN abundance estimation over the vast extent of the prospecting seabed area rather inaccurate. In addition, it is difficult to determine the dredger's exact covered distance, as it may not be in contact with the seabed while being towed, further casting a doubt on the accuracy on the samples collected.

Glassy and Singleton [22] used a comparatively faster method of underwater photography to perform in situ estimation of PMN abundance using an underwater camera lowered onto the seabed. These seabed photographs, apart from providing PMN abundance estimation, also provided additional seabed information such as seabed conditions, sediment characteristics, and the directional flow of water currents through scour marks seen on the seabed. The use of underwater cameras proved to be a more effective means of estimating PMN abundance as each seabed photograph depicted an average seabed area of 7.5 m^2 compared to an average seabed area of 0.25 m^2 covered by each sampling grab.

Sharma [23], achieved a comparatively quicker way of PMN abundance estimation through the use of a vessel's tow-frame-mounted underwater camera. To expedite the analysis of these seabed photographs, Sharma et al. [24] digitised these photographs and developed a machine learning based image processing software to estimate PMN abundance from these seabed photographs. To train this software, an operator creates polygons around the PMN images in order to teach the software to recognise these relevant features. Once taught, the software can automatically calculate the total number of PMN or sediment pixels and tabulate them accordingly for each digitised photograph.

With the democratisation of AUV and underwater acoustic equipment, Okazaki and Tsune [11] used these equipment to conduct a survey on an approximate 40 km² of seabed area within the Japanese licensed area in the CCFZ. Using an AUV equipped with an underwater camera and a multibeam echo sounder, they collected more than 10,000 seabed photographs and built a detailed bathymetry map of the area from which they studied the correlation between bathymetry features and PMN distribution.

Similarly, Lee and Kim [15] used a ship-towed SSS to collect seabed backscatter images and freefall grab samples within their Korean licensed area in the CCFZ. Based on visual comparison from both datasets, the authors observed that low SSS seabed backscattering returns correspond to freefall grab areas with low PMN abundance while medium-to-high SSS seabed backscattering returns correspond to freefall grab samples from areas with high PMN abundance, suggesting a correlation between SSS seabed backscatter returns and PMN abundance.

In summary, mechanical devices such as samplers, dredgers and corers are limited to small seabed area sampling where collected samples are insufficient in providing an accurate estimation of PMN abundance over the vast extent of the entire licensed area. On the contrary, the comparatively larger seabed coverage imaging capability of acoustic-based equipment such as multibeam sonar and SSS have shown promising results in correlating PMN abundance based on their acoustic seabed backscatter returns. This thesis builds on this observation by exploring the feasibility of modelling this correlation through the use of a data-driven supervised machine learning method.

2.2 Study Site

The CCFZ is a submarine region between $120 \,^{\circ}\text{W}$ to $160 \,^{\circ}\text{W}$ and $5 \,^{\circ}\text{N}$ to $20 \,^{\circ}\text{N}$ within the equatorial northeast Pacific Ocean with the typical depth ranging from 3 km to 6 km [2]. The 58,000 km² seabed area licensed to Singapore is situated within the CCFZ between Hawaii and Mexico as shown in Figure 2.1. The data used in this thesis were collected using an AUV during an environmental baseline seabed survey from an approximate $5 \,\text{km}^2$ region of interest within this licensed area in 2015.



Fig. 2.1. The extent of the CCFZ is demarcated by the blue dashed lines. The Singapore licensed zone for PMN exploration is shown within the red circle.

2.3 Equipment

The AUV used in the seabed survey is equipped with an inertial navigation system for position tracking, a doppler velocity log for vehicle navigation, a camera coupled with lighting and laser scaling system for seabed photography, and an SSS for seabed backscatter imaging. In addition, the AUV utilised a long baseline system for its positioning and navigation.

2.4 Data Collection

During the environmental survey, an AUV-mounted camera and an AUVmounted SSS collected the seabed photographs and seabed backscatter data respectively. The AUV travelled at an average speed of 2.8 knots while maintaining an average altitude of 8 m over the surveyed area where the average seabed depth is 4125 m. It took the AUV a dive time of 15 hours to cover a seabed area of approximately 5 km² in a pre-programmed lawnmower pattern as shown in Figure 2.2.



Fig. 2.2. A pre-programmed lawnmower pattern route taken by the AUV during the environmental baseline seabed survey within a CCFZ's seabed of interest. An AUV-mounted SSS collected seabed backscatter data from a seabed area of approximately 5 km^2 .

Within this 15-hours period, the AUV-mounted SSS collected seabed backscatter data continuously, while 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, seabed photographs were occasionally taken at irregular intervals resulting in alongtrack patches of seabed with no seabed photograph as shown in Figure 2.3.



Fig. 2.3. Symbol ' \mathbf{x} ' indicates position where each seabed photograph was taken. (a) Segment of dive where seabed photograph was taken at approximately three-second interval. (b) Due to unforeseen variability in time taken for the camera's flash recharging, there were instances where seabed photograph was taken at intervals that exceeded three seconds.

The SSS seabed backscatter data collected covered a seabed area of approximately 5 km^2 and consists of 11 Georeferenced Tagged Image File Format (GeoTIFF) images, each of size 1000×48000 pixels (translating to a seabed area of $0.1 \times 4.8 \text{ km}^2$). Georeferencing information such as longitude and latitude of the data collection points are embedded into each image file. The AUV-mounted camera collected around 3500 seabed GeoTIFF photographs. The size of each photographed seabed image is 1015×811 pixels, depicting an average seabed area of 12 m^2 .

2.5 Processing of Sidescan Sonar Image

SSS backscatter data collected are of binary file format. We process all SSS seabed 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. 2.2. The size of each image strip is 1000-by-48,000 pixels which corresponds to an approximate seabed area of 100-by-4800 m.

From the visual patterns observed in our SSS seabed backscatter dataset, we observe that the dark and light contrasting segments depict seabed areas of low and high PMN abundance respectively. Thus, to enhance these contrasting segments, histogram equalisation, an image processing method for enhancing contrast variation, is applied to the SSS seabed backscatter images for contrast equalisation. The nadir (seabed directly underneath the AUV) portion of the SSS seabed backscatter images which contains no useful textural information is removed, resulting in SSS seabed backscatter images with a swath-width of 75 m (Image swath of 750 pixels) as seen in Figure 2.4.



Fig. 2.4. Swath-width of the SSS seabed backscatter image with nadir portion removed.

2.6 Processing of Seabed Photographs

A feature-based image processing technique for quantifying PMN distribution in seabed photographs is used to identify and quantify PMN seabed coverage area within each seabed photograph as shown in Figure 2.5. The PMN coverage area for each seabed photograph is defined as the seabed percentage occupied by the visible PMN.

Taking into consideration that the economically acceptable range for mining is between 5 kg/m^2 to 20 kg/m^2 [4], the threshold separating high and low PMN seabed coverage photographs is set at 40%. This translates to a PMN density of around 23 kg/m^2 which is above the required economically acceptable range. Thus, of the 3500 photographs used in this thesis, 45% of them are labelled as photographs showing high PMN abundance seabed while the remaining 55% are labelled as photographs showing low PMN abundance seabed.



Fig. 2.5. PMN identified as region of interest on both seabed photographs. (a) Photograph depicting a seabed with a low PMN coverage of 5.674% translating to an area of $0.682 \,\mathrm{m^2}$. Photograph depicting a seabed area with a high PMN coverage of 61.351% translating to an area of $7.585 \,\mathrm{m^2}$.

Chapter 3

Methodology

Based on the literature survey conducted in Section 2.1, there exist shortcomings in the current PMN estimation methods. Although seabed data collected through underwater acoustics sensors cover a larger area when compared to existing methods, there is still a need to reliably in estimating PMN abundance from these data. This chapter investigates the possibility and effectiveness estimating PMN abundance using pattern variations from a dataset of SSS seabed backscatter images through data-driven method.

This chapter discusses the initial implementation of the ANN model. Section 3.1 discusses the rationale for using ANN and the architecture used. Section 3.2 discusses the various processing methods used in ANN dataset preparation. Section 3.3 discusses the ANN training algorithm and the method used in preventing dataset overfitting.

3.1 Artificial Neural Network Architecture

ANN is a computational model of interconnected nodes loosely based on the biological neural network structure. It is a supervised machine learning method known for its ability to learn mapping functions of underlying features from a given labelled training dataset [25]. In addition, a trained ANN model is capable of capturing unknown, complex and non-linear relationships for a wide variety of different application areas, ranging from image classification to speech recognition [26, 27]. Thus, ANN is selected as the tool to model the non-linear function between the SSS seabed backscatter data and seabed PMN abundance.

The ANN architecture selected for this thesis has a topological network with two hidden layers. For this selected ANN architecture, the number of neurons for hidden layer '1' and '2' is dependent on the size of the training dataset as shown by the formula in Figure 3.1 [28]. This particular architecture is chosen for its ability to learn underlying patterns from a large number of distinct data samples using a comparatively small number of hidden neurons.



Fig. 3.1. A two-hidden-layer ANN architecture where 'I' represents the number of features for each training sample, 'S' represents the number of output neurons and 'N' represents the number of training samples. The number of neurons used for hidden layer 1 and 2 respectively are denoted as 'J' and 'K' respectively.

The aim of the ANN is to minimise the error between the given labelled dataset and the trained model's predicted values by adjusting the interconnecting weight parameters between all layers iteratively (details to be discussed in Section 3.3).

Ample labelled training data samples are necessary for the ANN to have better insight into the underlying patterns of the training dataset as this allows the ANN to be sufficiently trained in making meaningful predictions. If the number of training data samples is too small, the ANN will not have sufficient relevant information to adequately learn these dependencies resulting in a trained model with relatively lower accuracy performance.

3.2 ANN Dataset Preparation

To correlate the collected seabed photographs and SSS seabed backscatter images, all seabed photographs are superimposed onto their corresponding geotagged locations on the SSS seabed backscatter images. However, only a total approximate seabed area of 0.042 km^2 is covered by all 3500 seabed photographs (each photograph covered an approximate seabed area of 12 m^2) compared to the total approximate seabed area of 5 km^2 covered by SSS seabed backscatter images.

Due to this vast difference in seabed coverage area, PMN seabed coverage seen in each seabed photograph is assumed to be uniform 2 m alongtrack and 50 m across-track (translates to 20×500 pixels) from the geotagged position of each seabed photograph resulting in a 4×100 m² SSS seabed backscatter segment. This size is subsequently reduced to 4×75 m² with the removal of across-track nadir (25 m) as shown in Figure 3.2. The assumption made here is reasonable as it has been shown that, on average, only a 10% PMN density change is expected every 450 m [5]. In addition, this assumption prevents overlapping of along-track neighbouring sliced SSS seabed backscatter data and also covers the entire across-track image width which minimises unused SSS seabed backscatter data.

To increase the number of data samples available to the ANN, these



Fig. 3.2. (a) Pre-programmed lawnmower path taken by the AUV. (b) A segment of the SSS seabed backscatter data, the AUV-mounted camera took a seabed photograph every 3 seconds at every location marked by symbol 'x'. PMN density seen at each geotagged seabed photograph is assumed to be uniform within the 4 m along-track and 75 m across-track of the SSS seabed backscatter image.

sliced SSS seabed backscatter segments are further separated into single strips of 1×750 pixels resulting in a total data sample size of 146,903 as shown in Figure 3.3. These data samples are normalised and labelled as either '1' or '2' (denoting high or low PMN coverage) based on the PMN seabed coverage seen from their respective geotagged photograph.

These labelled data samples are randomly separated into training (80% of total labelled data samples), validation (10%) and testing (10%) datasets. The training dataset is used for the training of the ANN and is the only dataset exposed to the ANN during the training phase. The validation dataset is used as a mechanism to prevent training dataset overfitting while the testing datasets is used in the evaluation of the generalising performance of the trained ANN. To achieve an unbiased estimation of the trained ANN's accuracy performance, both the validation and testing datasets are not exposed to the ANN in the training phase.



Fig. 3.3. Multi-row data samples are sub-divided into single-row data samples to increase the total number of data samples.

3.3 Training Algorithm

The SSS training dataset takes the form of a $N' \times I+1'$ matrix '**X**' as shown in Equation 3.1. The first column of value '1's denotes the bias, notation 'N' denotes the sample size of the training dataset, and notation 'I' denotes the number of features (number of image pixel columns) in each data sample. The labelled output for the training dataset is denoted by a matrix '**Y**', where element ' y_{n1} ' equates to value '1' and element ' y_{n2} ' equates to value '0' when the n^{th} training data is labelled as value '1' (high
PMN abundance) and so forth.

$$\mathbf{X} = \begin{bmatrix} 1 & x_{11} & \dots & x_{1I} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{N1} & \dots & x_{NI} \end{bmatrix}, \qquad \mathbf{Y} = \begin{bmatrix} y_{n1} & y_{n2} \\ \vdots & \vdots \\ y_{N1} & y_{N2} \end{bmatrix}$$
(3.1)

From Equation 3.2, element ' w_{ij}^{1} ' denotes the inter-connecting weight between the input layer and hidden layer '1', where notations 'i' and 'j' represent the ' i^{th} ' neuron in the input layer and ' j^{th} ' neuron in hidden layer '1' respectively. Likewise, element ' w_{jk}^{2} ' denotes the inter-connecting weight between hidden layer '1' and '2', where notations 'j' and 'k' represent the ' j^{th} ' neuron in hidden layer '1' and '2', where notations 'j' and 'k' represent the ' j^{th} ' neuron in hidden layer '1' and ' k^{th} ' node in hidden layer '2' respectively and so forth. The collective weight between neighbouring neuron layers is denoted by matrix ' \mathbf{W}^{p} ', where notation 'p' denotes the weight's starting layer. These weights are randomly initialised close to zero to ensure that the neurons do not perform the same computation.

$$\mathbf{W}^{1} = \begin{bmatrix} w_{bias1}^{1} & \dots & w_{biasJ}^{1} \\ w_{11}^{1} & \dots & w_{1J}^{1} \\ \vdots & \ddots & \vdots \\ w_{I1}^{1} & \dots & w_{IJ}^{1} \end{bmatrix}, \\ \mathbf{W}^{2} = \begin{bmatrix} w_{bias1}^{2} & \dots & w_{biasJ}^{2} \\ w_{11}^{2} & \dots & w_{1K}^{2} \\ \vdots & \ddots & \vdots \\ w_{21}^{2} & \dots & w_{2K}^{2} \end{bmatrix}, \\ \\ \mathbf{W}^{3} = \begin{bmatrix} w_{bias1}^{3} & w_{biasJ}^{3} \\ w_{11}^{3} & w_{12}^{3} \\ \vdots & \vdots \\ w_{K1}^{3} & w_{K2}^{3} \end{bmatrix}$$
(3.2)

The ANN's training algorithm iterates between the feed-forward and the back propagation process. In the feed-forward propagation process, the data samples are propagated through the model's weights and neurons' activation function, generating a matrix of interim predictions at the neuron outputs. To effectively classify the non-linear relationship between PMN abundance and SSS seabed backscatter training dataset, a sigmoid activation function is applied to the weighted input at every neuron in hidden layer '1', hidden layer '2', and the output layer. In addition, the inputs (SSS seabed backscatter dataset) to the ANN are normalised to avoid saturating the neurons' sigmoid activation function. The mathematical descriptions of the feed-forward propagation process are shown in the equations below.

$$\mathbf{HL1_{IP}} = \mathbf{XW^{1}} = \begin{bmatrix} 1 & x_{11} & \dots & x_{1I} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{N1} & \dots & x_{NI} \end{bmatrix} \begin{bmatrix} w_{bias1}^{1} & \dots & w_{biasJ}^{1} \\ w_{11}^{1} & \dots & w_{1J}^{1} \\ \vdots & \ddots & \vdots \\ w_{I1}^{1} & \dots & w_{MJ}^{1} \end{bmatrix}$$
(3.3)

$$\mathbf{HL1_{OP}} = \frac{1}{1 + \exp(-\mathbf{HL1_{IP}})}$$
(3.4)

$$\mathbf{HL2_{IP}} = \mathbf{HL1_{OP}}\mathbf{W^2} \tag{3.5}$$

$$\mathbf{HL2_{OP}} = \frac{1}{1 + \exp(-\mathbf{HL2_{IP}})} \tag{3.6}$$

$$OP_{IP} = HL2_{OP}W^3$$
(3.7)

$$\mathbf{OP_{op}} = \frac{1}{1 + \exp(-\mathbf{OP_{IP}})} \tag{3.8}$$

After the completion of each forward propagation process, the back propagation process begins with a calculation of the cross-entropy error (denoted by 'E') between the given dataset labelled and the model's predicted values. Cross-entropy is used here as it has a faster weight learning rate when the resultant cross-entropy error is large, thus any slowdown in weight optimisation is avoided as the cross-entropy error minimises with each training iteration [29, 30].

$$E = -\frac{1}{N} \sum \sum \left[\mathbf{Y} \odot ln(\mathbf{OP_{op}}) + (1 - \mathbf{Y}) \odot ln(1 - \mathbf{OP_{op}}) \right]$$
(3.9)

Next, partial derivatives calculations are made throughout the ANN to determine the cross-entropy error value with respect to each preceding weight in the ANN [31]. The mathematical descriptions of using partial derivation to calculate the error function with respect to each individual weight are shown in the equations below.

$$\alpha^{3} = \frac{\delta E}{\delta \mathbf{OP_{OP}}} \odot \frac{\delta \mathbf{OP_{OP}}}{\delta \mathbf{OP_{IP}}}$$
(3.10)

$$\frac{\delta E}{\delta \mathbf{W^3}} = \alpha^{\mathbf{3}T} \frac{\delta \mathbf{OP_{IP}}}{\delta \mathbf{W^3}} \tag{3.11}$$

$$\alpha^{2} = \alpha^{3} \left[\frac{\delta \mathbf{OP_{IP}}}{\delta \mathbf{HL2_{OP}}} \right]^{T} \odot \frac{\delta \mathbf{HL2_{OP}}}{\delta \mathbf{HL2_{IP}}}$$
(3.12)

$$\frac{\delta E}{\delta \mathbf{W}^2} = \alpha^{2^T} \frac{\delta \mathbf{HL2_{IP}}}{\delta \mathbf{W}^2} \tag{3.13}$$

$$\alpha^{1} = \alpha^{2} \left[\frac{\delta \mathbf{HL2_{IP}}}{\delta \mathbf{HL1_{OP}}} \right]^{T} \odot \frac{\delta \mathbf{HL1_{OP}}}{\delta \mathbf{HL1_{IP}}}$$
(3.14)

$$\frac{\delta E}{\delta \mathbf{W}^{\mathbf{1}}} = \alpha^{\mathbf{1}T} \frac{\delta \mathbf{HL} \mathbf{1}_{\mathbf{IP}}}{\delta \mathbf{W}^{\mathbf{1}}}$$
(3.15)

All calculated partial derivatives are used by the conjugate gradient descent algorithm to iteratively locate the next path to the minimal of the cross-entropy error function. This method is used as it convergences the cross-entropy error function using less iterations when compared to the typical gradient descent method [32].

The feed-forward and back propagation processes are repeated, and with each iteration, the ANN's weights are automatically re-adjusted to further lower the resultant cross-entropy error between the given labelled answer and the predicted output. This is continued till convergence of the cross-entropy error function or a maximum preset training iteration number is met. To ensure the effectiveness of the training process, the accuracy performance of the trained model is assessed through a comparison between the given dataset labelled and the final predicted values as shown by the mathematical descriptions below.

$$\mathbf{p} = \operatorname{argmax}(\mathbf{OP_{op}}) \tag{3.16}$$

$$\mathbf{y} = \operatorname{argmax}(\mathbf{Y}) \tag{3.17}$$

Accuracy of Model (%) =
$$\frac{1}{N} \sum (\mathbf{p} \Leftrightarrow \mathbf{y}) \times 100$$
 (3.18)

Iterating the ANNâĂŹs feedforward and backpropagation training processes indefinitely increases its accuracy performance on the training dataset. However, doing so would also lead to an overfitted model where the weight parameters are so explicitly tuned towards the given training dataset that they begin to erroneously fit underlying noise in the training dataset as features. On the other hand, having insufficient iterations undermine the ANN's ability to adequately capture the underlying features from the given labelled training dataset as can be seen in the drop in accuracy perfromance against the testing dataset after iteration 8000 in Table 3.1.

Having an ANN overfitting on a training dataset resulted in the model performing poorly with subsequent unseen datasets. Thus, to mitigate overfitting and to ensure that a trained model can generalise well to any data outside the training datasets, the generalisation ability of the trained model is monitored by evaluating the accuracy performance of the model against the validation dataset at the end of every forward and back propagation iteration during the training phase. The final set of ANN weights chosen is the one which yields the maximum accuracy performance with the validation dataset, which in this case would be the value of the weights

Table 3.1

Accuracy performance of the ANN model with the training and testing datasets. Improvement in accuracy performance with the training dataset can be seen with a larger number of training iterations. Overfitting of the ANN model towards the training dataset is evident from the decrease in

the model's accuracy performance towards the testing dataset beyond 16,000 training iterations.

Iteration	Training Dataset Accuracy $\%$	Testing Dataset Accuracy $\%$
80	54.67	50
800	82.58	60.14
1600	86.02	61.59
8000	89.87	61.59
16000	92.01	63.04
80000	95.95	57.25

at iteration 2250 as shown in Figure 3.4.

The ANN's accuracy performance metric is gauged based on the ANN's ability to make generalised predictions towards the unseen testing dataset. Using the above-mentioned methods, the trained ANN model achieved an accuracy performance of 68.00% in classifying SSS seabed backscatter images into high and low PMN abundance seabed areas.



Fig. 3.4. Figure depicting the problem of overfitting as the ANN training iteration number increases. Although the training dataset accuracy increases with the increasing number of iterations, there is also a gradual drop in validation dataset accuracy after 2250 iterations. Thus, the ANN is overfitting on the training dataset after 2250 iterations. The ANN's interconnecting weights between the neurons is taken at this iteration where it is evaluated on the unseen testing dataset.

Chapter 4

Dataset Features Analysis

The current trained ANN model achieved an accuracy performance of 68 % in PMN abundance estimation. To further increase the accuracy performance of the ANN, a proper understanding on the dataset features is crucial towards implementing suitable data processing methods.

This chapter conducts an in-depth analysis into the dataset features and discusses the methods used from our analysis findings to improve further from the ANN's current accuracy performance of 68%. The rationale for using these methods along with its resultant accuracy improvements are discussed in the following four sections. Section 4.1 discusses the contrast enhancement of relevant SSS seabed backscatter feature through the use of Contrast Limited Adaptive Histogram Equalisation (CLAHE). Sections 4.2 and 4.3 discuss the methods used in increasing the sample and feature size of the SSS seabed backscatter dataset respectively. Section 4.4 discusses the method of creating additional relevant dataset features by combining the bathymetry dataset to the existing SSS seabed backscatter dataset.

4.1 Contrast Limited Adaptive Histogram Equalisation (CLAHE)

Compared to the initial Histogram Equalisation method (Section 2.5) which does contrast enhancement based on the entire image's histogram, CLAHE method adjusts individual pixels based on the histogram of their surrounding pixel values [33]. This method works well on images with significant light and dark regions as it amplifies their contrasting features and enhances their regional features as shown in Figure 4.1.



Source:www.mathworks.com/help/examples/images/win64/ApplyCLAHEToAColorImageExample_01.png

Fig. 4.1. CLAHE enhancement of a generic image (right image) has a more prominent contrast between its darker and lighter shade segments when compared to the original image (left image).

Applying an appropriate choice of image equalisation method to the SSS seabed backscatter dataset images is vital towards improving the existing accuracy performance of the ANN. An ideal training image should possess features with distinct contrasting segments depicting the traits of each class. Insufficiently distinct contrasting segments or the absence thereof can hinder the ANN's ability to discern features pertaining to each class during the training process, resulting in a lower accuracy performance.

As observed previously in Section 2.5, the dark and light contrasting segments depict seabed areas of low and high PMN abundance respectively. Thus, to better enhance these contrasting segments, CLAHE method is used in preference over the initial Histogram Equalisation method to process the SSS seabed backscatter images.

To show the resultant visual difference, both Histogram Equalisation and CLAHE methods are applied separately on an identical portion of SSS seabed dataset image. A visual comparison showed that the CLAHE method processed image has a more distinct contrasting segment between high and low PMN abundance seabed areas when compared to the Histogram Equalisation processed image as shown in Figure 4.2.

The accuracy performance of the ANN showed an improvement when trained with a CLAHE method processed SSS seabed backscatter images. The accuracy performance of the trained model also showed an improvement from 68.00 % to 73.58 %, demonstrating the effectiveness of enhancing features that are relatable PMN abundance. Thus, it is important to first identify dataset features that are relatable to each individual class before applying a suitable processing method to best highlight these features.



Fig. 4.2. (a) Comparison of a portion of SSS seabed backscatter images processed using Histogram Equalisation and CLAHE. (b) Enlarged view of SSS seabed backscatter images processed using Histogram Equalisation. (c) Enlarged view of SSS seabed backscatter images processed using CLAHE. Visual contrast between high and low PMN abundance area is more prominent after CLAHE application.

4.2 Enhanced Dataset Field of View

The AUV's lawnmower path was programmed to have at least 25 % sidescan overlap for each run-length, resulting in overlapping regions from both adjacent run-length paths as shown in Figure 4.3.



Fig. 4.3. A segment of three neighbouring SSS seabed backscatter images. Each SSS seabed backscatter image has a 25% overlapping SSS region, coming from both the left and right adjacent run-length paths.

As the neighbouring left and right overlapping regions are within the demarcation zone of the centre run-length strip, we appended these overlapping SSS segments to the centre run-length strip dataset. Each overlapping SSS seabed backscatter region from the left and right neighbouring run-length strips is approximately 25 m (250 pixels). Both overlapping regions are appended to the existing dataset, increasing the across-track width from 750 to 1250 pixels as shown in Figure 4.4.

Appending the centre run-length strip with neighbouring left and right overlapping SSS regions increases the overall seabed feature information in the dataset. This helps in the training process of the ANN, as a training dataset with additional relevant features provides additional underlying patterns that are relatable to PMN abundance. The effectiveness of this method is shown by a further improved accuracy performance from 73.58 % to 81.78 %.



Fig. 4.4. Overlapping region from the left and right adjacent run-length paths increases the column size of the dataset from 750 to 1250 pixels This dataset is then further separated into single-row data samples as previously discussed in Section 3.2.

4.3 Enhanced Dataset Representation

Due to variability in the time taken for the AUV-mounted camera's flash recharging, seabed photographs were not taken at some instances. This resulted in numerous patches of seabed area not used as part of the ANN dataset due to the absence of ground-truth photograph as seen in Figure 4.5.



Fig. 4.5. Symbol ' \mathbf{x} ' shows the geotagged location where seabed photograph was taken. PMN seabed coverage from each geotagged photograph is assumed to be uniform to its surrounding SSS pixels and is processed into data samples for the ANN (shown by the white segments). The remaining SSS image represents seabed area not used as data sample for the ANN.

To fully utilise these unused patches of seabed areas, the PMN seabed coverage seen in each photograph is assumed to be uniform 50 m along-track of the SSS seabed backscatter image (identical to the 50 m across-track assumption made in Section 3.2). In the event where distance between two consecutive photographs is less than 50 m, the along-track segment of the SSS seabed backscatter image is spaced such that the segment boundary is equidistant from two consecutive photographs as shown by the red border in Figure 4.6.



Fig. 4.6. Symbol ' \mathbf{x} ' indicates the position where seabed photographs shown on the right were taken. PMN seabed coverage seen in each geotagged photograph is uniform up to a distance of 50 m along-track and acrosstrack of the SSS seabed backscatter image as shown by the red border. This distance is shorter at instances where consecutive photographs were taken in close proximity of one another as shown by the bottom two symbol ' \mathbf{x} ' where two consecutive seabed photographs were taken at 75 m from each other.

This method utilises previously unused segments of the SSS seabed backscatter image which increases the total SSS seabed backscatter data sample size from 146,903 to 400,000. With this increase, the ANN is subjected to a more diverse set of training data samples during its training process, allowing additional variability between the two different classes to be 'learned'. The effectiveness of this method is shown by the improved accuracy performance of the trained ANN which increased from 81.78% to 84.24%.

4.4 Heterogeneous Dataset

In addition to SSS seabed backscatter images and seabed photographs, the AUV-mounted depth, altimeter and navigation sensors provided bathymetry data of the seabed area during the environmental baseline survey. A Geo-TIFF image generated from these bathymetry data is shown in Figure 4.7.



Fig. 4.7. Bathymetry GeoTIFF image from 5 km^2 region of interest.

As there had been studies indicating the correlation of bathymetry to PMN presence [6, 15], an opportunistic investigative study is conducted to gauge the ANN's accuracy performance using the bathymetry dataset.

Due to its format similarity to the SSS seabed backscatter dataset, bathymetry data collected are subjected to the same data preparation as the SSS seabed backscatter dataset (Section 3.2), where all seabed photographs are superimposed onto their corresponding geotagged positions on the bathymetry GeoTIFF image. PMN seabed coverage seen in each geotagged photograph is also assumed to be uniform to its surrounding bathymetry pixels as shown in Figure 4.8. These labelled bathymetry images are further separated into single-row pixel slices and used as data samples for the ANN training process (identical method previously shown in Figure 3.3).



Fig. 4.8. (a) Data collection path of the AUV. (b) SSS seabed backscatter image and (c) bathymetry image from the same geotagged location. Similar to the assumption made with the SSS seabed backscatter dataset, PMN density seen in each geotagged photograph is assumed to be uniform to its surrounding bathymetry pixels as indicated by the red boundary. (For illustration purpose, not drawn to scale)

ANN trained using the bathymetry dataset resulted in a relatively lower accuracy performance of 72.69%, when compared to the accuracy performance of 84.24% achieved by the ANN trained using the SSS seabed backscatter dataset. However, the lower accuracy performance is within expectation as the bathymetry dataset images are of lower resolution (1.78 m

per pixel), when compared to the SSS seabed backscatter image resolution (0.1 m per pixel), resulting in a less detailed set of seabed information as shown in Figure 4.9.



Fig. 4.9. 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. (a) A 18 × 18 pixels SSS seabed backscatter image and (b) a single pixel bathymetry image representing the same seabed area of 3.24 m^2 .

Even though the bathymetry dataset trained ANN resulted in a lower accuracy performance, its accuracy result of 72.69% suggested that the seabed bathymetry dataset contains useful features that can be correlated to PMN abundance. Taking a cue from this result, the total features of the dataset is increased by complementing the existing SSS dataset with the bathymetry dataset. This is done by appending geo-tagged single-pixel-row data samples of the bathymetry dataset to the end of each corresponding geo-tagged SSS seabed backscatter single-pixel-row data sample. However, due to the difference in pixel resolution between these two datasets, multiple single-pixel-row SSS seabed backscatter data samples are appended to the same single-row bathymetry data sample as shown in Figure 4.10.



Combined 1-by-1300-pixel data sample of SSS seabed backscatter and bathymetry pixel

Fig. 4.10. Method for combining two datasets of different resolution representing the same seabed area into a heterogeneous dataset.

Appending the bathymetry dataset to the existing SSS seabed backscatter dataset increases the column size (features) of each data sample by 50 pixels to 1 by 1300 pixels. With this increase in feature size using a heterogeneous dataset of both SSS and bathymetry data, the accuracy performance of the ANN is further improved to 86.00 %.

With the implementation of heterogeneous dataset, the ANN's current input comprises data collected from two different acoustic equipment. The feasibility of further improving the ANN's accuracy performance is explored through the implementation of an alternative ANN architecture with two separate input and hidden '1' layers as shown in Figure 4.11. The number of hidden layer '1' neurons from the single-input ANN is proportionally split based on the feature size from the SSS seabed backscatter and bathymetry dataset into neurons for hidden layer '1' and '1.5' respectively. The intent of this ANN architecture is to allow the neurons from hidden layer '1' and '1.5' to first learn the respective underlying patterns from each equipment's dataset before combining them into layer '2' to learn the overall underlying feature. The accuracy performance achieved through this ANN architecture is 85.59 % which is of similar performance to the single layer '1' and '2' ANN architecture previously discussed in Section 3.1.

The random nature by which the dataset is split into training, validation and testing dataset may create a performance bias towards a particular ANN architecture, resulting in an exceptionally good or bad accuracy performance. This may occur due to the training dataset having particularly dominant underlying pattern that performs particularly well or badly towards a particular ANN architecture. The following chapter discusses a method that both mitigates the issue of performance bias and ensures that the chosen ANN architecture has the best generalisation capability.



Fig. 4.11. Dual-input ANN architecture.

4.5 Results

This section discusses the use of a k-fold cross-validation based method in validating the generalised classification ability and accuracy performance of both ANN architectures implemented in Sections 3.1 and 4.4.

An ANN's accuracy performance is optimised through an ideal configuration of hyperparameter values. In its original use, k-fold cross-validation method is used to determine the best performing combination of ANN hyperparameter values, such as the number of hidden layers and the number of neurons per hidden layer. This method involves training 'k' different ANN models with a set of predefined hyperparameter values. All 'k' ANN models are trained using 'k' equal data subsets partitioned from all available data samples. During each ANN model training process, one data subset is used as the testing dataset while the remaining 'k-1' data subsets served as the training dataset. This entire process is iterated for 'k' ANN models till every data subset is sequentially rotated through as the testing dataset. The accuracy performance of the ANN based on this particular set of predefined hyperparameter values is cross-validated using the average accuracy result from all 'k' trained ANN models. The algorithm is then repeated with a different set of predefined hyperparameter values until the ANN achieved a targeted accuracy result.

However, a non-iterative method for determining the number of hidden layers and neurons was established in Sections 3.1. As such, the k-fold crossvalidation is instead used as a method to verify the ANN architecture's generalisation capability and thus renamed as k-fold performance validation method to reflect this purpose.

The generalisation capability of the ANN architecture is achieved by iterating the preprocessing, training, validation and testing methods discussed in Chapters 3 and 4 for ten folds. Each fold possessed its own randomly distributed training (80%), validation (10%) and testing (10%) datasets as shown by the toy example in Figure 4.12. Through this method, the two ANN architectures each produced ten different accuracy performance results as shown in Figures. 4.13 and 4.14.

The average accuracy performance calculated using ten-fold performance validation method provides extra reliability in computing accuracy and removes any dataset performance bias towards any ANN architecture, as this overall result was generated, based on not just one, but ten different configurations of training, validation and testing datasets averaged out.

Based on the overfitting preventive method discussed in Section 3.3, the accuracy results produced from both ANN architectures are tabulated in



Fig. 4.12. An overview of the ten-fold performance validation method using a toy example dataset with ten data samples. All data samples are randomly split into training (80%), validation (10%) and testing (10%) datasets for each fold.

Table. 4.1. The tabulated accuracy performance from all twenty ANN models ranged from 82.66 % to 87.89 % giving an average accuracy of 85.36 % and 84.83 % for single-input and dual-input ANN respectively. The small, varying range from both sets of results showed the reliability and stability of both ANN architectures when trained and tested using different configurations of datasets.

As shown in Table 4.1, the single-input ANN architecture produced a higher average accuracy performance when compared to the dual-input ANN architecture. An analysis of both ANN architectures showed that this is due to the difference in the number of non-linear mapping process which occurred at the first hidden layer. For the dual-input ANN architecture, there are two separate mapping processes occurring in hidden layer '1' and hidden layer '1.5'. This resulted in two separate feature spaces, one of which can only linearly separate the SSS seabed backscatter dataset, while the other is only capable of linearly separating the bathymetry dataset. Although hidden layer '2' non-linearly mapped both of these outputs from hidden layer '1' and '1.5' onto a new feature space which linearly separates the majority of the dataset, this method proved to be disadvantageous



Fig. 4.13. Training iteration graphs of a single-input ANN architecture using ten variations of randomly distributed training, validation and testing data samples.



Fig. 4.14. Training iteration graphs of a dual-input ANN architecture using ten variations of randomly distributed training, validation and testing data samples.

when compared to the single-input ANN architecture. This is because the single-input ANN architecture only utilises one mapping function for the heterogeneous dataset at the first hidden layer, mapping the input dataset onto a feature space which linearly separates the majority of the dataset as a whole instead of two different entities. This resultant output from hidden layer '1' is further subjected to another round of non-linear mapping function at hidden layer '2', resulting in another feature space where even more of the heterogeneous dataset are linearly separated, thus producing a higher average accuracy performance.

Fold	Accuracy(%) of ANN architecture with single-input	Accuracy(%) of ANN architecture with dual-input
1	86.00	85.57
2	85.19	84.19
3	87.89	86.98
4	85.41	84.82
5	84.70	83.57
6	82.81	82.66
7	84.02	83.80
8	86.88	87.20
9	85.21	83.81
10	85.52	85.78
Ten-fold Average	85.36	84.83

Table 4.1 Accuracy performance of two different ANN architectures, trained from ten configurations of training, validation and testing datasets.

The classification performance of the single-input ANN architecture is shown by applying confusion matrix on a randomly selected 4th fold test result as shown in Figure 4.15. Confusion matrix presents a visual summary of the ANN's accuracy performance for each class. From the confusion matrix of the 4th fold test result, it can be seen that single-input ANN has the ability to accurately classify the majority of the data samples into high and low PMN abundance classification with no skewed performance bias towards a particular class.



Confusion Matrix

Fig. 4.15. Confusion matrix for the single-input ANN's 4th fold test result. The two green cells represent the number and percentage of correct classification prediction made, while the red cells represent the number and percentage of incorrect classification prediction made. Lastly, the grey cell represents the ANN overall test accuracy for the 4th fold.

The trained ANN can be further improved upon by increasing the number of different training samples and adding more relevant features to the existing heterogeneous dataset. However, these will also increase the time needed to train the ANN. Currently, it takes around 30 hours for the ANN to train on approximately 360,000 (80% of total dataset) data samples using MATLAB software installed on a workstation with a dual-processor Intel Xeon E5-2630 V3 CPU@2.4 GHz processor. Once trained, the time taken for the ANN to classify future heterogeneous datasets will be almost instantaneous.

4.6 Summary

Singapore completed its first-ever biological and environmental baseline study within its CCFZ licensed area in 2015. Data used in this thesis were collected from a specific seabed area of approximately 5 km^2 within the CCFZ licensed area. These data comprised 3500 GeoTIFF photographs, each depicting a seabed area of 12 m^2 , SSS seabed backscatter GeoTIFF dataset, and a GeoTIFF bathymetry dataset. The seabed backscatter and the bathymetry data were combined into a heterogeneous dataset and processed into 400,000 data samples, each individual sample size being of the form of 1×1300 pixels. These data samples are split into 80% training dataset, 10% validation dataset, and 10% testing dataset.

These data samples were labelled with either '1' or '2' denoting 'high' or 'low' percentage PMN abundance. The PMN seabed abundance captured in each geotagged seabed photograph determined the classification label for all corresponding geotagged data samples. Based on economic feasibility for seabed mining, the threshold for PMN abundance seabed coverage was set at 40 % which is equivalent to 23 kg/m^2 . The two hidden layer ANN models used in this thesis comprised 1800 and 600 neurons for hidden layer 1 and 2, respectively. The number of neurons for each hidden layer is dependent on the number of training data samples and classification outputs.

The PMN abundance ground truth data were taken from 3500 seabed photographs, covering a total seabed area of 0.42 km^2 . The 400,000 heterogeneous data samples collected encompasses a total seabed area of 5 km^2 . A summary showing the methods used in this research thesis is shown in Figure 4.16.





Chapter 5

Conclusion

5.1 Summary of Contributions

This thesis proposed a data-driven method that addresses the efficiency and accuracy of existing PMN abundance estimation methods. The main contributions are summarised as follows:

- First known work of a data-driven approach in performing PMN abundance estimation using a heterogeneous dataset of seabed backscatter and bathymetry data.
- Demonstrated the possibility of correlating PMN ground truth data from a limited quantity of small seabed coverage photographs to a large seabed coverage heterogeneous dataset.
- Demonstrated the ANN architecture's reliability with a ten-fold average accuracy performance of 85.36% in estimating PMN abundance from unseen heterogeneous dataset.
- Under similar trained conditions, the trained ANN enables faster estimation of PMN abundance for future deep seabed sites without the need for underwater cameras.

The strategy used in approaching this classification problem started with an accuracy performance benchmark from an ANN trained with a dataset prepared using common data preparation methods. This was followed by several focus methods to improve upon this accuracy performance benchmark as presented in the various sections of this thesis. These accuracy performance improving and verification methods are as follow:

- Prevented ANN overfitting through the use of validating dataset instead of the usual iterative regularisation method.
- Increased dataset field of view through the utilisation of overlapping SSS seabed backscatter regions.
- Increased dataset representation by maximising the usage of our heterogeneous dataset and preprocessing them from scan images into single-pixel-row scan line data samples.
- Opportunistic use of bathymetry data as additional dataset features to improve the accuracy performance of the ANN.
- A k-fold performance validation method providing a reliable verification of the ANN architecture's accuracy performance.

5.2 Future Work

Future work includes:

- Estimating PMN abundance on an even larger seabed area by extending existing methods discussed in this thesis towards exploring the possibility of using acoustic data collected from a surface vessel.
- Exploring the feasibility of using different machine learning models in estimating PMN abundance.

5.3 Final Thoughts

Due to the neuroplasticity property of the ANN, there is no need to redesign a new ANN architecture to accommodate future features that will be added to the dataset as these new features will be automatically learned in the training phase of the existing ANN architecture. As such, researchers working on PMN abundance classification problem with a dataset from another acoustic equipment or those operating in different environmental conditions can potentially use this ANN architecture together with the various dataset preparation and performance verification methods discussed in this thesis.

Lastly, from this thesis, it can be learnt that given adequate and accurately labelled dataset, implementing a data-driven approach model such as an ANN in quantifying PMN abundance is not difficult. However, achieving a high accuracy performing ANN requires a domain expert who has a good understanding on the dataset features together with the methods used in preparing the dataset.

Bibliography

- 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., ch. 2, pp. 23–63. Springer, 2017.
- [2] 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.
- [3] International Seabed Authority, "A geological model of polymetallic nodule depostis in the clarion-clipperton fracture zone," Technical Study: No. 6, 2010.
- [4] M. JL, "Economic aspects of nodule mining," in *Marine Manganese Deposits*, G. Glasby, Ed., ch. 11, pp. 327–355. Elsevier, 1977.
- [5] R. Kaufman, "The selection and sizing of tracts comprising a manganese nodule ore body." Offshore Technology Conference, pp. 283– 293, 1974.
- [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." *Marine Georesources And Geotechnology*, vol. 19, no. 3, pp. 167–180, 2001.

- M. Atmanand and G. Ramadass, "Concepts of deep-sea mining technologies," in *Deep-Sea Mining: Resource Potential, Technical and En*vironmental Considerations, R. Sharma, Ed., ch. 10, pp. 305–343.
 Springer, 2017.
- [8] 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," *Marine Geology*, vol. 49, no. 3-4, pp. 225–240, 1982.
- [9] V. N. Hari, B. Kalyan, M. Chitre, and V. Ganesan, "Spatial modeling of deep-sea ferromanganese nodules with limited data using neural networks," *IEEE Journal of Oceanic Engineering*, 2017.
- [10] 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.
- [11] 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.
- [12] A. Tsune, "Effects of size distribution of deep-sea polymetallic nodules on the estimation of abundance obtained from seafloor photographs using conventional formulae," *Eleventh Ocean Mining and Gas Hydrates Symposium*, pp. 153–159, 2015.
- 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 MTS/IEEE Washington*, pp. 1–5, 2015.

- [14] 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.
- [15] 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 Geotechnol*ogy, vol. 22, no. 1-2, pp. 103–114, 2004.
- [16] M. Weydert, "Measurements of the acoustic backscatter of selected areas of the deep seafloor and some implications for the assessment of manganese nodule resources," *The Journal of the Acoustical Society* of America, vol. 88, no. 1, pp. 350–366, 1990.
- [17] C. de Moustier, "Inference of manganese nodule coverage from seabeam acoustic backscattering data," *Geophysics*, vol. 50, no. 6, pp. 989–1001, 1985.
- [18] E. Lange, S. Petersen, L. Rüpke, E. Söding, and K. Wallmann, "Marine resources - opportunities and risks," World Ocean Review, vol. 3, pp. 54–193, 2014.
- [19] 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 years 1873 to 1876," Tech. Rep., ch. 2, pp. 33–183, 1891.
- [20] J. Mero, "The deep-sea floor," in *The mineral resources of the sea*,J. Mero, Ed., ch. 6, pp. 103–241. Elsevier, 1965.
- [21] G. Glasby, "Deep seabed mining: past failures and future prospects," *Marine Georesources and Geotechnology*, vol. 20, no. 2, pp. 161–176, 2002.

- [22] G. Glassy and R. Singleton, "Underwater photographs of manganese nodules from the southwestern pacific basin," New Zealand Journal of Geology and Geophysics, vol. 18, no. 4, pp. 597–604, 1975.
- [23] R. Sharma, "Quantitative estimation of seafloor features from photographs and their application to nodule mining," *Marine Georesources and Geotechnology*, vol. 11, no. 4, pp. 311–331, 1993.
- [24] R. Sharma, S. J. Sankar, S. Samanta, A. Sardar, and D. Gracious, "Image analysis of seafloor photographs for estimation of deep-sea minerals," *Geo-marine letters*, vol. 30, no. 6, pp. 617–626, 2010.
- [25] J. Schmidhuber, "Deep Learning in neural networks: An overview," *Neural Networks*, vol. 61, pp. 85–117, 2015.
- [26] D. CiresAn, U. Meier, J. Masci, and J. Schmidhuber, "Multi-column deep neural network for traffic sign classification," *Neural Networks*, vol. 32, pp. 333–338, 2012.
- [27] A. Hannun, C. Case, J. Casper, B. Catanzaro, G. Diamos, E. Elsen,
 R. Prenger, S. Satheesh, S. Sengupta, A. Coates, and A. Y. Ng, "Deep Speech: Scaling up end-to-end speech recognition," arXiv:1412.5567, 2014.
- [28] G. B. Huang, "Learning capability and storage capacity of two-hiddenlayer feedforward networks." *IEEE transactions on neural networks*, vol. 14, no. 2, pp. 274–81, 2003.
- [29] P. T. de Boer, D. P. Kroese, S. Mannor, and R. Y. Rubinstein, "A tutorial on the cross-entropy method," Annals of Operations Research, vol. 134, pp. 19–67, 2005.
- [30] M. Stone, "Cross-validatory choice and assessment of statistical predictions," Journal of the royal statistical society. Series B (Methodological), vol. 36, no. 2, pp. 111–147, 1974.

- [31] R. Rumelhart, D.E., Hinton, G.E., 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," *Journal of Optimization Theory and Applications*, vol. 7, no. 2, pp. 149–154, 1964.
- [33] S. M. Pizer, E. P. Amburn, J. D. Austin, R. Cromartie, A. Geselowitz, T. Greer, B. Ter Haar Romeny, J. B. Zimmerman, and K. Zuiderveld, "Adaptive histogram equalization and its variations," *Computer vision, graphics, and image processing*, vol. 39, no. 3, pp. 355–368, 1987.

Publication

 L.J. Wong, B. Kalyan, M.Chitre, and H. Vishnu, "Polymetallic nodules abundance estimation using sidescan sonar: A quantitative approach using artificial neural network," *IEEE Oceans Conference*, Aberdeen, pp. 1-6, 2017.