NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

THESIS

TEXTURAL ANALYSIS AND STATISTICAL INVESTIGATION OF PATTERNS IN SYNTHETIC APERTURE SONAR IMAGES

by

Suluck P. Klumprasert

September 2022

Thesis Advisor: Derek Olson
Second Reader: Mara S. Orescanin

THIS PAGE INTENTIONALLY LEFT BLANK

REPORT DOCUMENTATION PAGE			*Form Approved OMB No. 0704-0188*

1. AGENCY USE ONLY *(Leave blank)*	**2. REPORT DATE** September 2022	**3. REPORT TYPE AND DATES COVERED** Master's thesis

4. TITLE AND SUBTITLE TEXTURAL ANALYSIS AND STATISTICAL INVESTIGATION OF PATTERNS IN SYNTHETIC APERTURE SONAR IMAGES	**5. FUNDING NUMBERS** RQQ89
6. AUTHOR(S) Suluck P. Klumprasert	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Naval Postgraduate School Monterey, CA 93943-5000	**8. PERFORMING ORGANIZATION REPORT NUMBER**
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) Office of Naval Research, Arlington, VA , 22217	**10. SPONSORING / MONITORING AGENCY REPORT NUMBER**

11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.

12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release. Distribution is unlimited.	**12b. DISTRIBUTION CODE** A

13. ABSTRACT (maximum 200 words)

Textural analysis and statistical investigation of patterns in synthetic aperture sonar (SAS) images is useful for oceanographic purposes such as biological habitat mapping or bottom type identification for offshore construction. Seafloor classification also has many tactical benefits for the U.S. Navy in terms of mine identification and undersea warfare. Common methods of texture analysis rely on statistical moments of image intensity, or more generally, the probability density function of the scene. One of the most common techniques uses Haralick's Grey Level Co-occurrence Matrix (GLCM) to calculate image features used in the applications listed above. Although widely used, seafloor classification and segmentation are difficult using Haralick features. Typically, these features are calculated at a single scale. Improvements based on the understanding that patterns are multiscale was compared with this baseline, with a goal of improving seafloor classification. Synthetic aperture sonar (SAS) data was provided by the Norwegian Research Defense Establishment for this work, and was labeled into six distinct seafloor classes, with 757 total examples. We analyze the feature importance determined by neighborhood component analysis as a function of scale and direction to determine which spatial scale and azimuthal direction is most informative for good classification performance.

14. SUBJECT TERMS textural analysis, SAS, sonar, identification, seabed, classification	**15. NUMBER OF PAGES** 73
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	**18. SECURITY CLASSIFICATION OF THIS PAGE**	**19. SECURITY CLASSIFICATION OF ABSTRACT**	**20. LIMITATION OF ABSTRACT**
Unclassified	Unclassified	Unclassified	UU

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. 239-18

THIS PAGE INTENTIONALLY LEFT BLANK

Approved for public release. Distribution is unlimited.

TEXTURAL ANALYSIS AND STATISTICAL INVESTIGATION OF PATTERNS IN SYNTHETIC APERTURE SONAR IMAGES

Suluck P. Klumprasert
Lieutenant, United States Navy
BS, University of California, Davis, 2015

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN PHYSICAL OCEANOGRAPHY

from the

NAVAL POSTGRADUATE SCHOOL
September 2022

Approved by: Derek Olson
Advisor

Mara S. Orescanin
Second Reader

Peter C. Chu
Chair, Department of Oceanography

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

Textural analysis and statistical investigation of patterns in synthetic aperture sonar (SAS) images is useful for oceanographic purposes such as biological habitat mapping or bottom type identification for offshore construction. Seafloor classification also has many tactical benefits for the U.S. Navy in terms of mine identification and undersea warfare. Common methods of texture analysis rely on statistical moments of image intensity, or more generally, the probability density function of the scene. One of the most common techniques uses Haralick's Grey Level Co-occurrence Matrix (GLCM) to calculate image features used in the applications listed above. Although widely used, seafloor classification and segmentation are difficult using Haralick features. Typically, these features are calculated at a single scale. Improvements based on the understanding that patterns are multiscale was compared with this baseline, with a goal of improving seafloor classification. Synthetic aperture sonar (SAS) data was provided by the Norwegian Research Defense Establishment for this work, and was labeled into six distinct seafloor classes, with 757 total examples. We analyze the feature importance determined by neighborhood component analysis as a function of scale and direction to determine which spatial scale and azimuthal direction is most informative for good classification performance.

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF FIGURES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF TABLES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF ACRONYMS AND ABBREVIATIONS

ASM	angular second moment
FFI	Norwegian Defense Research Establishment
GLCM	gray level co-occurrence matrix
IEEE	Institute of Electrical and Electronic Engineers
MSE	mean squared error
NCA	neighborhood component analysis
PDF	probability density function
SAS	synthetic aperture sonar
SOFAR	sound fixing and ranging
SONAR	sound navigation and ranging
T-SNE	t-distributed stochastic neighbor embedding
UUV	unmanned underwater vehicle

THIS PAGE INTENTIONALLY LEFT BLANK

ACKNOWLEDGMENTS

I would like to thank Marc Geihufe from the Norwegian Defence Research Establishment for sharing their SAS data, images of the HUGIN 1000 UUV, and his helpful review of this thesis. I would also like to thank Marko Orescanin for his help with data processing and analysis. Thanks also to Mara Orescanin for her patience and assistance in helping with editing this paper. Most of all, thanks to Derek Olson, who has spent the past year helping me with answering all my questions and guiding me to the conclusion of this research.

THIS PAGE INTENTIONALLY LEFT BLANK

I. INTRODUCTION

Sonar data is the main means of information gathering in underwater applications, more so than optical data. Entire platforms are dependent on sonar data collection, such as submarines and minesweepers. Most of the antisubmarine warfare and mine countermeasure disciplines rely almost exclusively on the ability to collect and analyze sonar information. The primary method of detection of submerged objects has been and continues to be through active sonar. The significance of this underwater acoustic technology was used to find the HMS Titanic, detection of U-boats during WWII, and continues to prove useful in modern applications from environmental mapping to mine detection, as shown by D'Amico and Richard [1]. The core concept of active sonar remains the same – a transmitter sends a short pulse of sound in the environment. The sound scatters across the environment and is received by a hydrophone.

Traditional side scan sonar typically works by capturing high resolution images that decrease in resolution the greater the distance from the transducer array [2]. The issue with side-scan sonar is that the resolution of the image tends to degrade the further from the track of the vehicle housing the sonar. In contrast, synthetic aperture sonar or SAS combines several acoustic pings to form a higher resolution image when compared to traditional methods of sonar gathering. Synthetic aperture sonar or SAS is an improvement on traditional methods of standard, narrow beamwidth side-scan sonar. Traditional side-scan sonar images are processed with each ping echo being independently built into a coherent image. SAS instead compiles multiple pings over an area and builds a single, coherent picture with less distortion than an image that is built using only traditional side-scan sonar methods [3]. While the SAS method is more costly and time-consuming, the enhanced image that it provides is more useful for analyzing important features such as texture and object recognition.

Traditionally, target detection performance has been based on relative strength of the signal scattered by a target, and the received energy scattered by the environment [4], [5]. More generally, the probability density functions (PDF) of each intensity can be

compared to obtain an estimate of the false alarm rate and missed detection rate [5]. For areas with statistically homogeneous roughness features, like uniform sand, the energy metric, or PDF-based metrics are sufficient [5],[6], although the PDF of the scattered field may vary as a function of range from the sonar [6], [7], and [8].

High resolution SAS can resolve small features on the seafloor (on the order of centimeters) can resolve fine textures that can also be mistaken for targets [9], [10], [11], [12]. Modern detection algorithms [13], [14] look for patterns in images, and can be fooled when seafloor texture resembles the patterns sought. As compared with the traditional methods of determining system performance, high resolution systems require some knowledge of how these patterns translate into false alarm. These patterns are broadly termed "texture" and represent the patchiness of large-scale intensity changes. Quantification of texture is useful both as a means of predicting false alarms in sonar data (e.g., using humans or machines) [14], and as a way of characterizing seafloors, and classifying them into different types [2].

Texture is a paradoxical feature, being intuitive for understanding but hard to quantify. Objective methods used to calculate texture vary widely since there are no widely accepted quantitative metrics for measuring texture [15], [16], [17]. Humans are generally good at recognizing features such as smooth, rough, or even fuzzy, even without a tactile response to verify it. However, there is a lack of a standardized metric when it comes to having a machine automatically quantify image textures. Different disciplines have different methods of analyzing textural data. For this study, we will focus on gray level co-occurrence matrices or GLCMs. GLCMs have had relative success in the past in the realm of seafloor classification [15]. Previous scientific work has only used few features at specific spatial scales. Seafloor texture is multiscale, and is anisotropic, which may not be captured adequately by a few scales. The focus of this paper is to study how different GLCM features and scales can be used to evaluate sonar images. Further study is needed to understand the precise nature of geographical patterns, spatial clustering, and the spatial scales of intensity. GLCMs can potentially offer an increased level of understanding regarding these lesser studied features. What precisely makes a feature useful for classification is one of the primary objectives of this paper.

Underwater target detection is a field of study that would benefit greatly with an improved method of detection and textural analysis. A particular interest is the practice of detecting targets and the methods used to reduce the frequency of false identifications or false positives. False identifications or false positives are instances in which a system, machine, or operator incorrectly identifies an object that is not a mine as a mine. Prior studies by Williams [13] and Galusha [14] have focused on analyzing sonar images by clustering areas of high intensity and low intensity to separate seafloor features from objects.

Contemporary target detection methods generally focus on the positive identification of either a mine-like object or the "shadow" of a mine-like object within a sonar image [16]. This positive identification is contingent on that the mine-like object being proud of the surrounding seafloor environment on which it sits. As sonar, or primarily side scan sonar, collects data, any feature or object that is proud of the seafloor tends to return a stronger or brighter image return. Subsequently, the object prevents the area behind it from returning a signal, creating what is known as a "shadow" [16].

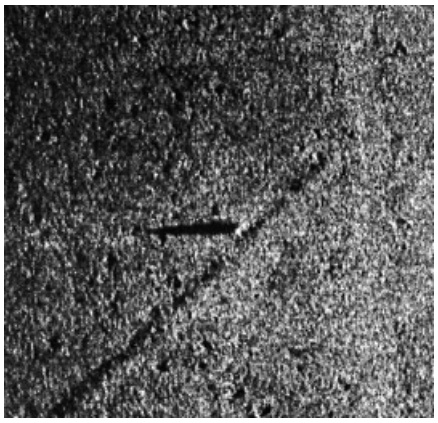


Figure 1. Side scan sonar imagery of a mine-like object and its shadow. Source: [16].

Some researchers like Galusha [14] have developed algorithms that aim to maximize the amount of accurate positive identifications are returned from SAS images. In these algorithms, pixels with high intensity returns are classified as mine-like objects,

while the dark pixels immediately adjacent to the mine-like objects are classified as shadows. Using multiple algorithms, they can test fast and accurate detection methods for basic target isolation. The downside of these types of algorithms is the prerequisite the mine-like object is proud of the seafloor environment in which it is located. The less homogeneous the seafloor environment, the harder it is to accurately identify mine-like objects, and the increase of false positives occurs.

Being able to quantify spatial patterns, particularly among sonar images, would vastly increase the accuracy of classification of bottom types and even bottom objects. Particularly, if the patterns or textures could be incorporated into classification software or machine learning algorithms then it would greatly increase the precision of positive identification and remove potential guesswork that is usually associated with an operator determining environments in a manual fashion. This could potentially be used by a sonar operator tasked with evaluating sonar images. Whether it is for classification of bottom types or identification of mine-like objects, any ability to quantify a feature that does not require a human operator can be used either as an additional measurement or even as the primary means of identification by itself. Being able to incorporate these measurements in tandem with machine learning algorithms would greatly increase the volume of images able to be processed, as well as, potentially increasing the accuracy of the classification or identification themselves.

In this paper, GLCMs will be used to access various features and parameters surrounding the textural properties in multiple SAS images. Upon analysis of GLCM calculations, the most descriptive features and parameters will be selected to test whether they can prove to be more helpful than traditional metrics of textural analysis. The remainder of this paper is organized as follows: Chapter II will provide an overview of previous studies and a review of the state of modern methods of textural analysis. Chapter III will describe the origin of the dataset being used in this thesis. Chapter IV will explain the methodology used in this experiment and how the data was processed. Chapter V will be an analysis of the results gathered and what information can be generated from it. Chapter VI will cover the mistakes made in this process and what further questions and

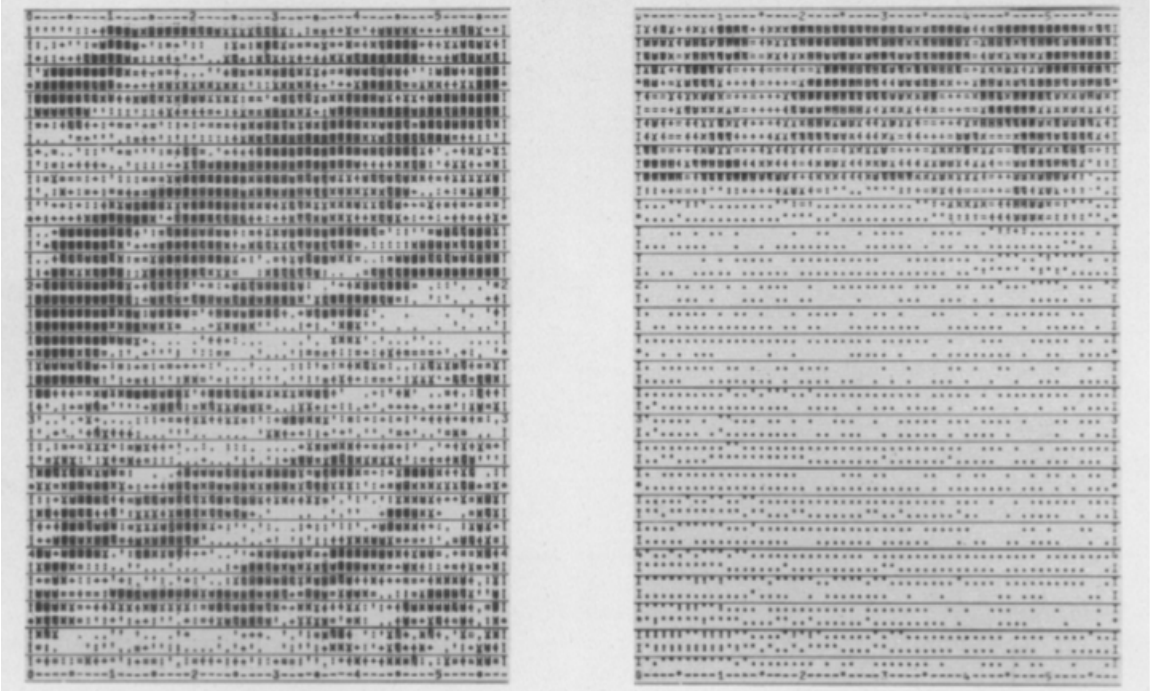
studies can be conducted in the future to broaden the understanding of textural and statistical analysis of seafloor features.

THIS PAGE INTENTIONALLY LEFT BLANK

II. PREVIOUS LITERATURE

The overall aim of this thesis is to use GLCMs to find features and parameters that work best for textural and statistical quantification of seafloor textures, and use these parameters to determine which scales are the most informative for classifying different types of features. While others have sought to find the best performance for classification tasks, we further seek to uses these highly informative features to determine what perceptually, about the images is driving this performance. Few have conducted research on what exactly the best method of quantification is. The three main methods that will be discussed, as well as, improved upon were methods developed by Haralick [15], Blondel [16], and Zare [17].

One of the earliest works in texture quantification was in 1973 by Haralick [15]. In this work, he proposed a method for computing textural features using gray-tone spatial relationships [15], [18]. Haralick stated that texture was an innate property of all surfaces and that the structural arrangement of surfaces and their relationship to surrounding surfaces provided meaningful information in terms of classification. Prior studies focused on power spectra, Markov meshes, and Rayleigh probability density functions [15], [19]. While not entirely unfeasible, he found that these various methods left much to be desired in terms of classification and largely ignored the wide range of complexities found in the real world. By using gray-tone spatial dependence matrices, Haralick was able to extrapolate spatial features such as angular second moment (ASM), contrast, and correlation. Using the features generated from the matrices, he was able to classify features found in a variety of aerial photographs – from woodlands to urban areas [15].



	Grassland			Water Body		
Angle	ASM	Contrast	Correlation	ASM	Contrast	Correlation
0°	.0128	3.048	.8075	.1016	2.153	.7254
45°	.0080	4.011	.6366	.0771	3.057	.4768
90°	.0077	4.014	.5987	.0762	3.113	.4646
135°	.0064	4.709	.4610	.0741	3.129	.4650
Avg.	.0087	3.945	.6259	.0822	2.863	.5327

Figure 2. Haralick's GLCM calculations of Grassland and Water Body photographs using the parameters: Angle, Angular Second Moment, Contrast, and Correlation. Source: [15].

In 2000, Blondel stated that side scan sonar imagery of seafloor surveys still tends to be evaluated visually rather than qualitatively [16]. The geophysical and environmental features that make up the seafloor are interpreted easily using visual patterns discernable to someone evaluating an image, but a quantitative approach would be helpful for further understanding and assessment of side scan sonar data. This would be particularly useful for evaluating images in bulk or for the purposes of machine learning data acquisition. Blondel used GLCMs to extract textural features using two factors: homogeneity and entropy. His research proved to have limited success as seen in Figure 4. While not entirely

useful, Blondel was able to develop segmentation for one type of rock from all other types of environmental features.

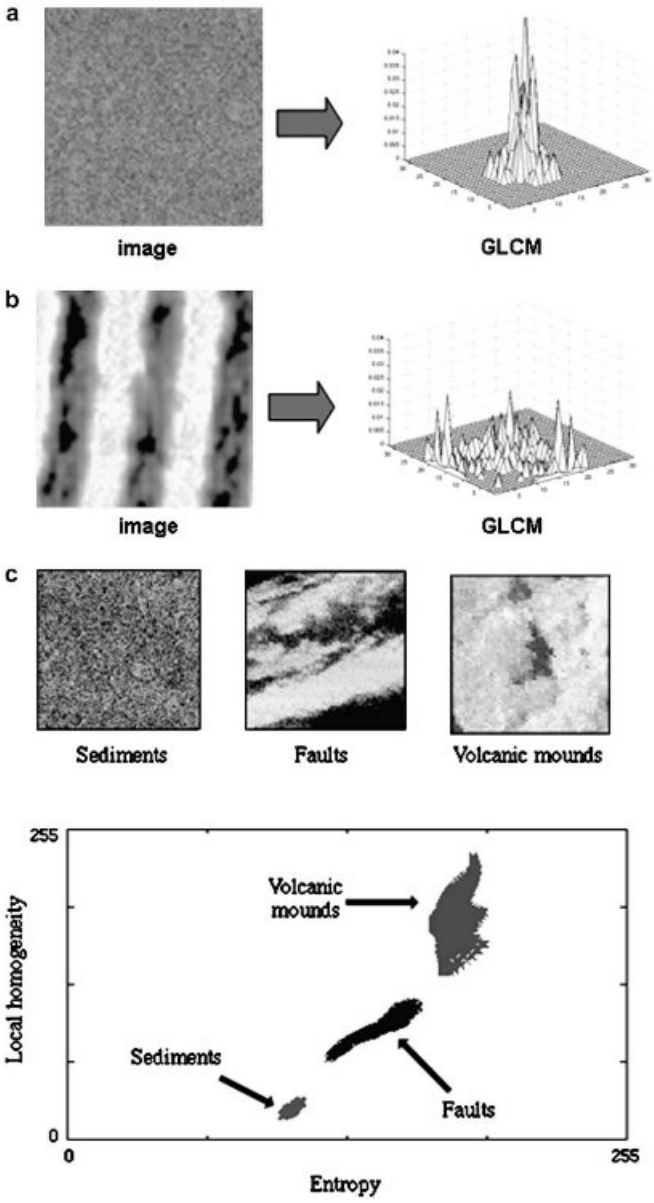


Figure 3. 3D representation of how an image's GLCM calculations are represented and segmented using the features Local Homogeneity and Entropy. Bottom graph shows an optimal graph using homogeneity and entropy for good feature segmentation. The actual graph of the data is shown in Figure 4. Source: [20].

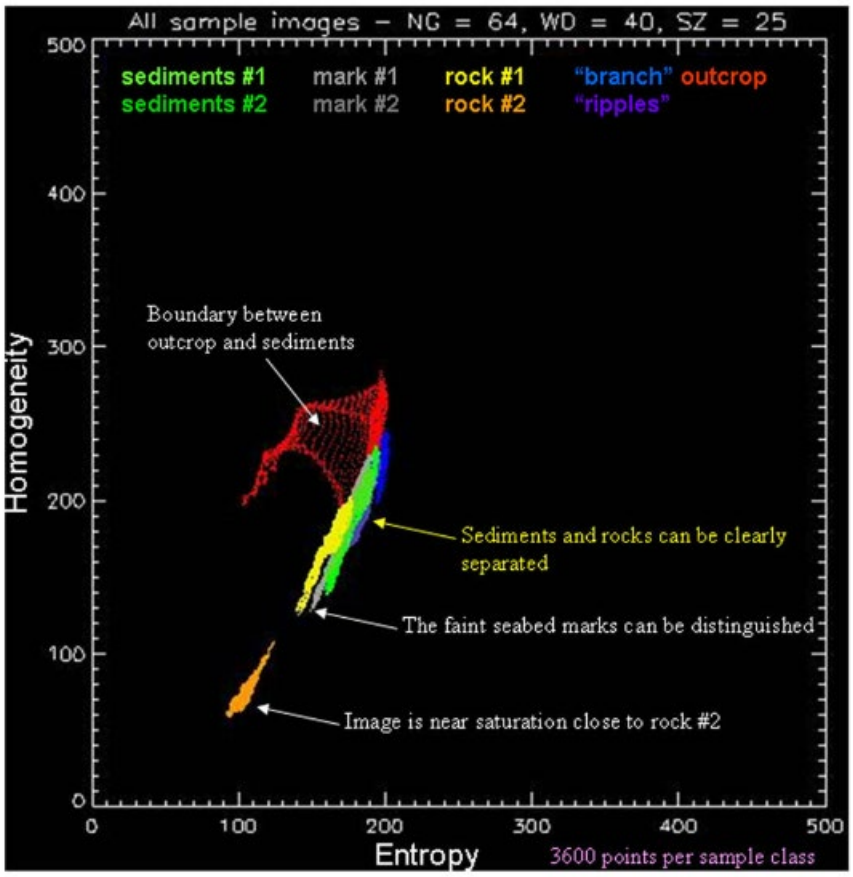


Figure 4. Philip Blondel took the photographs of the different features and separated them using GLCMs and the features Homogeneity and Entropy. Source: [20].

In 2017, Zare [17] developed a method of analyzing side-look synthetic aperture sonar images using a method coined Possibilistic Fuzzy Local Information C-Means (PFLICM). PFLICM combines existing techniques of data clustering to incorporate fuzzy and possibilistic clustering with spatial information to perform soft segmentation on an SAS image. Soft segmentation differs from hard segmentation in which hard segmentation draws a concrete line between areas that are different. Soft segmentation calculates the probability of a spatial location to be certain features or textures. PFLICM combines the two methods: Fuzzy Local Information C-Means (FLICM) and Possibilistic Fuzzy Clustering algorithm (PFCM) [17]. Using PFLICM, Zare attempted to improve the contemporary approaches to segmenting SAS images and accurately classify spatial areas

of SAS images as different textures. Zare also had an overall goal to create an algorithm with similar rate of successes in identifying various sea floor textures comparable to the ability of a human observer.

Previous studies have hinted at the implications of using textural analysis techniques for seafloor segmentation. However, much more research is needed to explain the parameters and features that are most useful in this endeavor. In this paper, the question that will be answered is precisely which features and under what parameters are most informative when it comes to processing the tonal and spatial relationship between distinct patterns found in varying seafloor textures.

THIS PAGE INTENTIONALLY LEFT BLANK

III. DATA

The high-resolution SAS images used in this paper were collected by the Norwegian Defense Research Establishment (FFI) in 2017. The unmanned underwater vehicle used to collect the data was the HUGIN HUS AUV carrying the HISAS 1032 [19]. The location where the data was collected was in the vicinity of the North Sea, west of the city of Bergen.

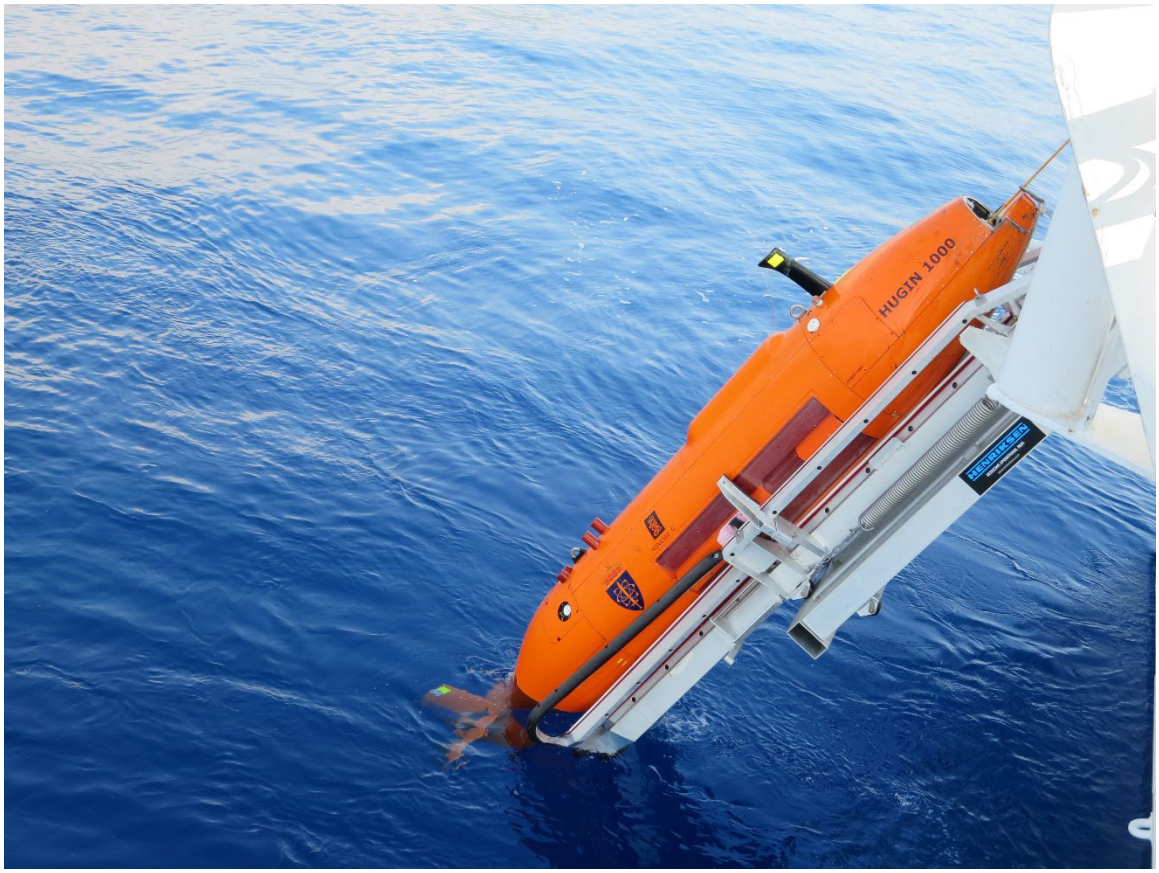


Figure 5. A photo of the HUGIN 1000 UUV. Photo Credit: Norwegian Defence Research Establishment (FFI)

At a vehicle altitude of 30m, the HUGIN 1032 collected sonar imagery at a frequency of 100 kHz and a bandwidth of 30 kHz [19]. A total of seventeen SAS images were used, each image with varying bottom types and textures. A total of 757 snippets were extracted from the seventeen images as part of this thesis. Similar labeling was performed in [21] using the same dataset, but this labeling was performed independently of that work. The snippets were labeled with mud, rock – gradient, rock - plateau, sand – homogeneous, sand – rocks, and sand – ripples.

To preprocess the images, the first step was to normalize the images using a CFAR split window normalizer. The window sizes were 500 pixels in range and 3000 pixels across the range. This would remove large-scale intensity changes created by the sonar equation (e.g., spherical spreading, vertical directivity patterns). The snippets were log-transformed by converting the magnitude square of the normalized pixel level to decibels using an arbitrary reference. For each snippet, overall mean level (i.e., tone in the terminology of Haralick [15]) was removed by subtracting the decibel version of mean intensity. This would assist in normalizing the tone inconsistencies found in varying areas of the pictures. Retaining this mean for use in classification may provide valuable information and could be explored in future work. These images were uncalibrated, meaning that the overall conversion between image units and complex pressure is unknown. Use of calibrated images may provide significant information for use in classification, although the focus of this work is on spatial texture patterns, rather than mean intensity. Equal probability quantization was performed to obtain the binned intensity levels as uniform quantization has produced historically poor results, as referenced in Haralick's 1973 study [15]. Once the data is processed in this format, the GLCMs were calculated using Matlab's graycomatrix function.

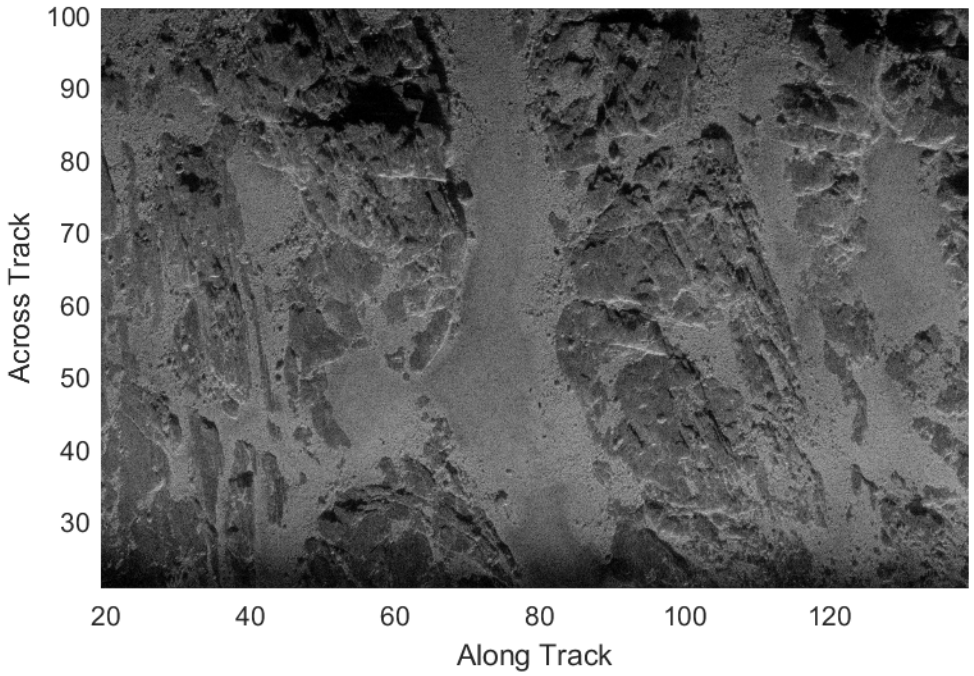


Figure 6. High resolution SAS image of North Sea bottom environment. Axes in meters.

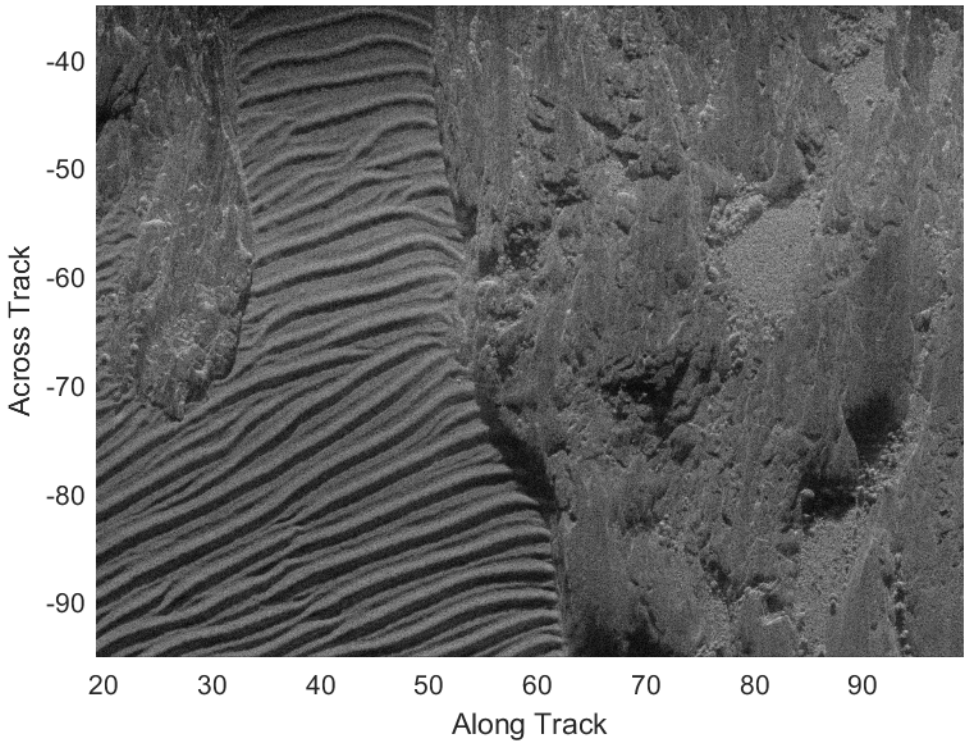


Figure 7. High resolution SAS image of North Sea bottom environment. Axes in meters.

IV. METHODS

Haralick stated that texture is an important characteristic of evaluating an image, particularly regarding geophysical and environmental features [8]. He explored different computational techniques to discern an image's features using its gray-tone spatial dependencies. By using statistics to explain textures, he was able to describe how one gray tone compared to another gray tone in a specified spatial relationship. Using these comparisons, categorization, and identifications can be made. However, image classification is not always a clear-cut prospect. Image classification is composed of several steps to include data processing, feature extraction, feature selection, and classification method selection [22]. Objective accuracy assessment for image classification can be difficult, since methods will vary depending on the type of data being analyzed and what purpose is classification being done. Plainly, the image classification process to categorize different species of tropical birds in photographs will differ greatly from the process needed to perform segmentation for seafloor features inside SAS images. Since trying to determine which features will prove most descriptive in seafloor texture classification is the goal of this thesis, GLCMs will be used as a tool to objectively measure the relationship between multiple features and the parameters in which they are measured.

However, first one must define texture. Texture can sometimes be defined as the relationship between high and low points on a topographical surface [15]. For instance, a texture can be described as rough if it has a large difference between high and low points with a relatively small difference in the space between those high and low points. Conversely, a smooth surface would have a small difference between high and low points in a similar spatial scale. When pictures are represented in a grayscale, certain textures can be quantified using the gray level differences between pixels (contrast), the size or area where the contrast occurs, and the directionality of the change relative to a selected pixel. Using these relationships, Haralick was able to propose 14 different measurements that could be taken such as the variance of the difference between adjacent pixels and the maximum correlation coefficient [15].

Grey level co-occurrence matrices (GLCMs) are tools that can be used to objectively quantify textures. According to Singh and Srivastava [23], "a GLCM is a matrix that represents the relative frequencies of a pair of grey levels present at a certain distance *d* apart and at a particular angle Θ. As shown in Figure 8 (adapted from [23), GLCM's values are calculated for a selected pair of distance and angle. The final matrix is then populated based on the sum of all the frequencies in order. Multiple GLCMs can be calculated based on the number offsets used in the calculations as shown in Figure 10. In this paper, 4 Θ values will be used: 0, 45, 90, and 135.

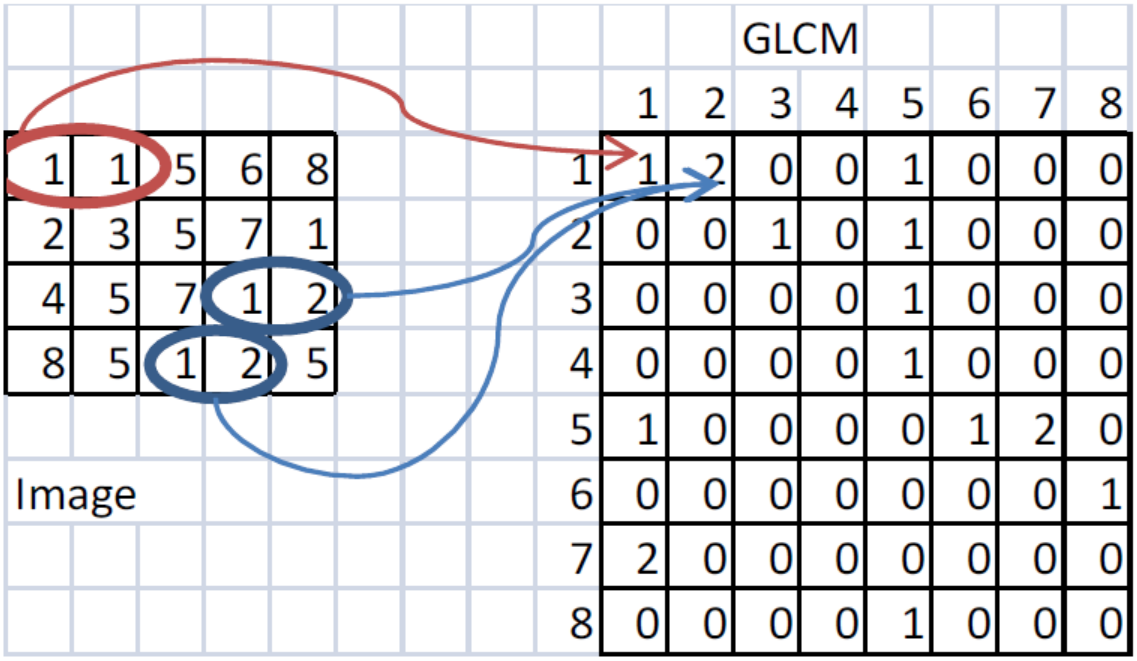


Figure 8. An example of how a GLCM is calculated. The image (left) is assigned values based on the brightness of the pixel. The GLCM (right) is then compiled using the spatial relationship of each pixel in relation to another adjacent pixel. Source: [23].

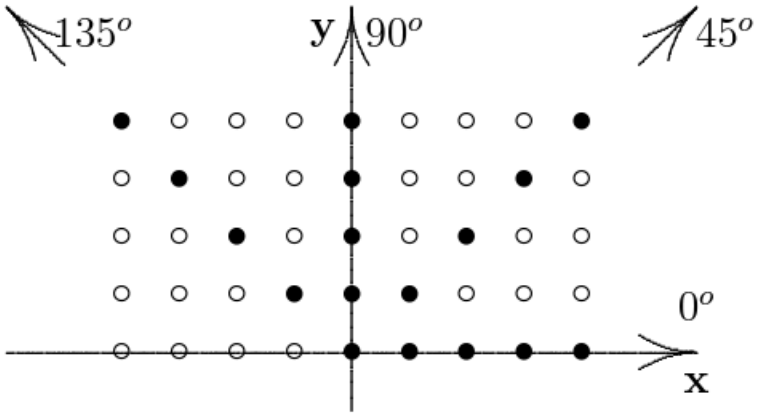


Figure 9. GLCM values can be expanded not only by calculating the relationship of the pixel directly to its right. Multiple angles of offset can be measured such as: 45, 90, and 135 degree offsets. Source: [24].

Texture features are single values that represent characteristics such as contrast, homogeneity, or entropy. The GLCMs that were calculated are converted to these features for each snippet using the 21 equations below.

1) Angular Second Moment (Energy) (Source: [15]):

$$f_1 = \sum_i \sum_j \{p(i,j)\}^2 \tag{1}$$

2) Contrast (Source: [15]):

$$f_2 = \sum_{n=0}^{N_g-1} n^2 \left\{ \sum_{i=1}^{N_g} \sum_{\substack{j=1 \\ |i-j|=n}}^{N_g} p(i,j) \right\} \tag{2}$$

3) Correlation (Source: [15]):

$$f_3 = \frac{\sum_i \sum_j (ij)\, p(i,j) - \mu_x \mu_y}{\sigma_x \sigma_y} \tag{3}$$

where μ_x, μ_y, σ_x, and σ_y are the means of standard deviation of p_x and p_y

4) Sum of Squares: Variance (Source: [15]):

$$f_4=\sum_i \sum_j (i-\mu)^2 p(i,j) \tag{4}$$

5) Inverse Difference Moment (Source: [15]):

$$f_5=\sum_i \sum_j \frac{1}{1+(i-j)^2} p(i,j) \tag{5}$$

6) Sum Average (Source: [15]):

$$f_6=\sum_{i=2}^{2N_g} i p_{x+y}(i) \tag{6}$$

7) Sum Variance (Source: [15]):

$$f_7=\sum_{i=2}^{2N_g} (i-f_8)^2 p_{x+y}(i) \tag{7}$$

8) Sum Entropy (Source: [15]):

$$f_8=\sum_{i=2}^{2N_g} p_{x+y}(i) \log\{p_{x+y}(i)\} \tag{8}$$

9) Entropy (Source: [15]):

$$f_9 = -\sum_i \sum_j p(i,j)\log(p(i,j)) \tag{9}$$

10) Difference Variance (Source: [15]):

$$f_{10} \text{ variance of } p_{x-y} \tag{10}$$

11) Difference Entropy (Source: [15]):

$$f_{11} = -\sum_{i=1}^{N_{g-1}} p_{x-y}(i)\log\{p_{x-y}(i)\} \tag{11}$$

12) Information Measures of Correlation (Source: [15]):

$$f_{12} = \frac{HXY - HXY1}{\max\{HX, HY\}} \tag{12}$$

13) Information Measures of Correlation (Continued) (Source: [15]):

$$f_{13} = (1 - \exp[-2.0(HXY2 - HXY)]^{1/2} \tag{13}$$

Where HX and HY are entropies of p_x and p_y, and

$$HXY = -\sum_i \sum_j p(i,j)\log(p(i,j)) \tag{14}$$

$$HXY1 = -\sum_i \sum_j p(i,j)\log\{p_x(i)p_y(j)\} \tag{15}$$

$$HXY2 = -\sum_i \sum_j p_x(i) p_y(j) \log\{p_x(i) p_y(j)\} \tag{16}$$

14) Maximal Correlation Coefficient (Source: [15]):

$$f_{14} = (\text{Second largest eigenvalue of Q})^{1/2} \tag{17}$$

Where

$$Q(i,j) = \sum_k \frac{p(i,k)p(j,k)}{p_x(i)p_y(k)} \tag{18}$$

15) Autocorrelation (Source: [25]):

$$f_{15} = \sum_i \sum_j (ij) p(i,j) \tag{19}$$

16) Dissimilarity (Source: [25]):

$$f_{16} = \sum_i \sum_j |i-j| \cdot p(i,j) \tag{20}$$

17) Cluster Shade (Source: [25]):

$$f_{17} = \sum_i \sum_j (i+j-\mu_x-\mu_y)^3 p(i,j) \tag{21}$$

18) Cluster Prominence (Source: [25]):

$$f_{18} = \sum_i \sum_j (i+j-\mu_x-\mu_y)^4 p(i,j) \tag{22}$$

19) Maximum Probability (Source: [25]):

$$f_{19} = \underset{i,j}{MAX}\ p(i,j) \tag{23}$$

20) Homogeneity (Source: [26]):

$$f_{20} = -\sum_{i}\ \sum_{j} \frac{1}{1+(i-j)^2}\, p(i,j) \tag{24}$$

21) Inverse Difference (Source: [26]):

$$f_{21} = \sum_{i}\ \sum_{j} \frac{1}{1+\left|(i-j)\right|} p(i,j) \tag{25}$$

It is worth noting that the equations and definitions for these features will sometimes vary depending on the study. For instance, Equation 20 for Homogeneity is called Inverse Difference Moment by Clausi [26] but Homogeneity by Singh [23]. All texture features were calculated from input GLCMs using Matlab function GLCM_Features1 written by Avinash Uppuluri [27]. These equations were taken from Haralick's 1973 study [15], Soh's 1999 study [25], and Clausi's 2002 study [26].

As previously discussed, the labels for each snippet were labeled mud, rock – gradient, rock - plateau, sand – homogeneous, sand – rocks, and sand – ripples. This labeling procedure is shown in Figures 10 and 11. There was a distinction made between snippets that appeared to be smooth and largely flat rock surfaces, and snippets that appeared to have mostly rocky but jagged features. The former was labeled as Rock – Plateau, while the latter was labeled as Rock – Gradient. There was greater variation in the patterns found in sandy textures, so three distinct classification labels for sand. If the sand was mostly homogeneous and smooth, it was labeled as Sand – Homogeneous. If ripples were the dominant features, it was labeled as Sand – Ripple. If the snippet consisted of gravel or larger rock, but sand was still the dominant feature, the snippet was labeled as Sand – Rock. Mud snippets were dominated by relatively dark tonal features (i.e., low

scattering strength) with a mostly homogeneous tonal makeup (similar to Sand – Homogeneous). Since some of the snippets were comprised of two types of seafloors, multilabel classification techniques may need to be used in future research.

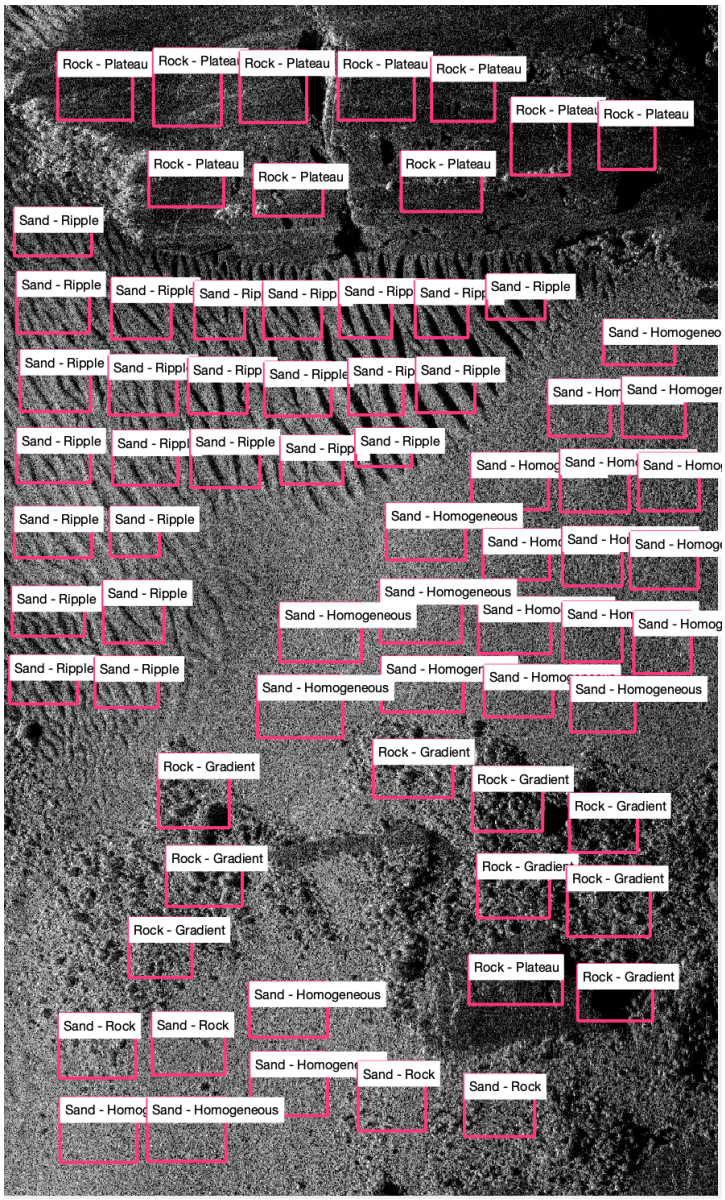


Figure 10. A visual representation of how different textures in the environment were labeled. These are example images of how the snippets of different features were selected. Each individual snippet represents a single and distinct geological feature. The six labels that were defined were: Rock – Plateau, Rock – Gradient, Mud, Sand – Homogeneous, Sand – Rock, and Sand – Ripple.

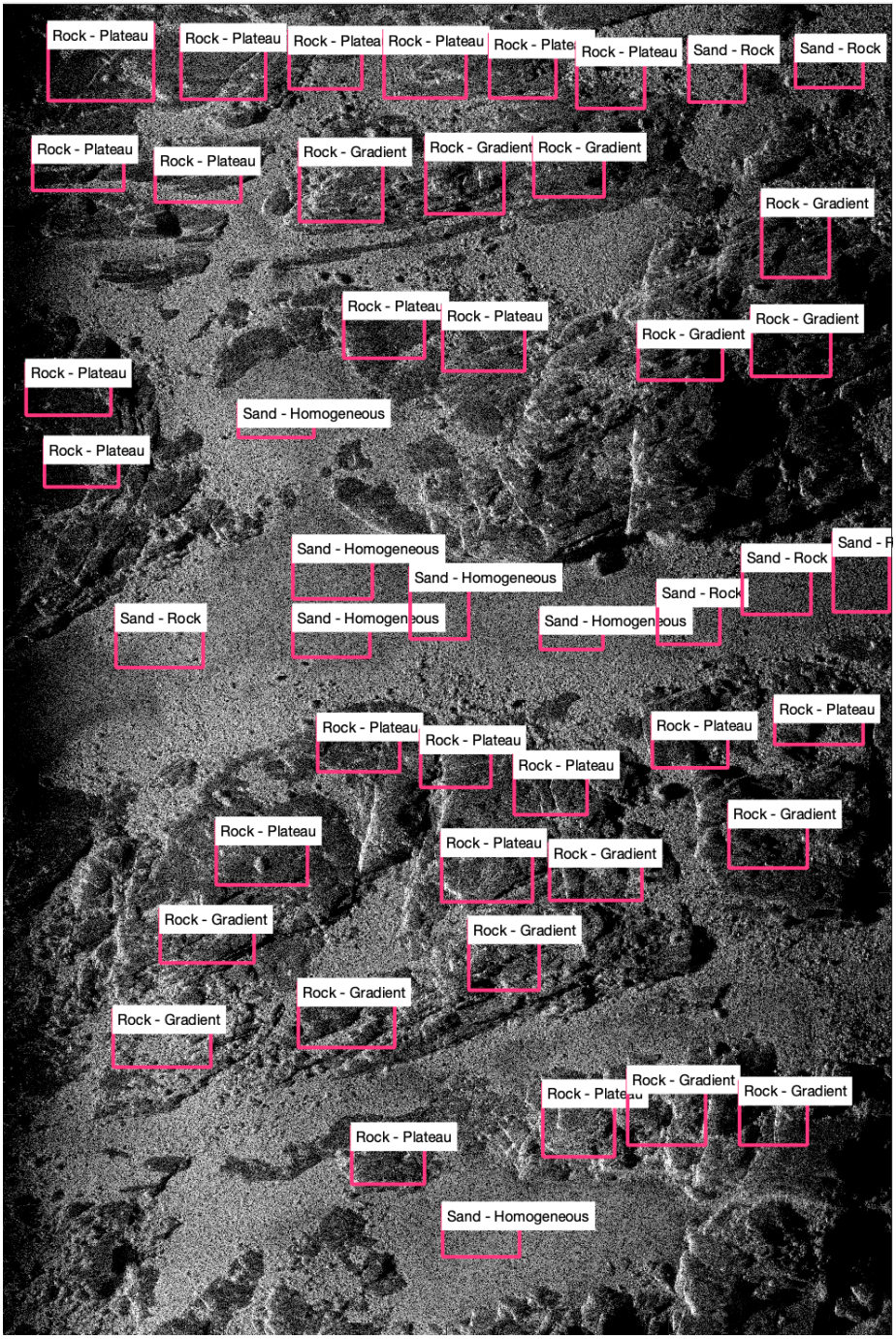


Figure 11. Another visual representation of how different textures in the environment were labeled.

Once this was done, the snippets had to be analyzed using gray level co-occurrence matrices or GLCMs. Each snippet was analyzed using four parameters: gray levels, offset

values, window sizes, and angle. In addition, a total of 21 features were calculated. The grey level selected was 16. The offset values chosen were 1 to 19 at an interval of 1. The window sizes selected were 20 through 40 at an interval of 1. The selected snippets are all various sizes in terms of pixel-by-pixel lengths, meaning that the dataset collected is not uniform in terms of picture size. Most snippets are approximately 200–300 pixels in each dimension. Having these selected window sizes helps to standardize the measurements. A percentile cutoff limit of the entire dataset was also used to control the signal to noise ratio of the data. The percentile cutoff on the lower limit was 5% and the percentile cutoff on the upper limit was 100%. This resulted in a signal to noise ratio of about 30 dB.

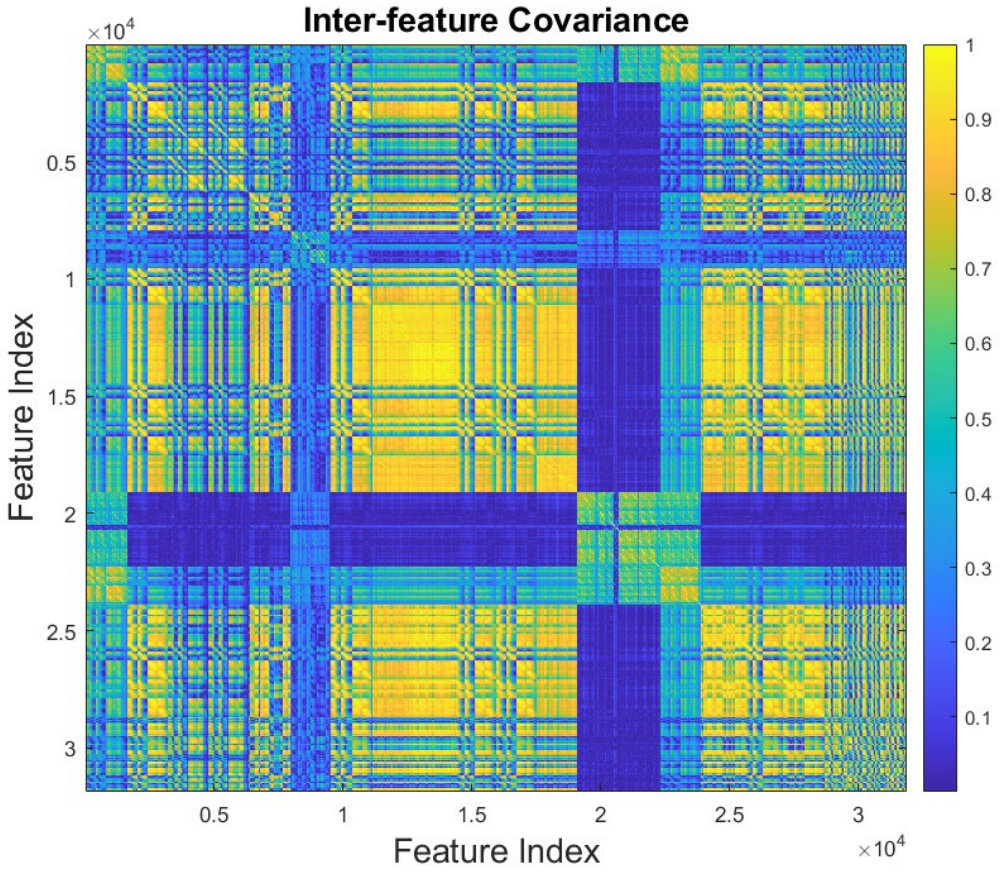


Figure 12. Covariance plot illustrating the relationship between Feature Index by Feature Index.

Once GLCMs were computed using the 757 snippets and selected parameters, there were 31,864 total feature combinations for each snippet. To trim down this number, a covariance matrix was used to remove features with low correlation with the labels. The overall inter-feature covariance matrix is shown in Figure 12. Using a covariance technique and a correlation cutoff at 0.35, only the top 1,918 features remained. The goal was to retain only features that possessed high correlation with the labels (Mud, Rock, Sand – Ripple, etc.). The feature-label covariance and threshold used is shown in Figure 13.

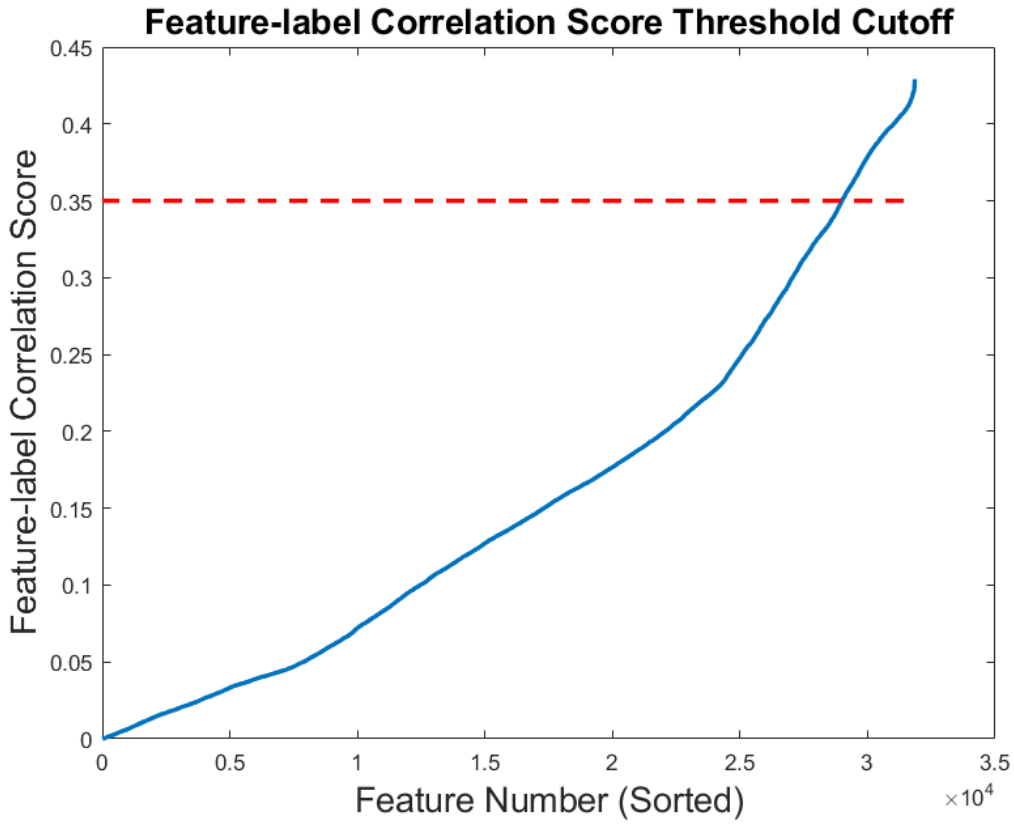


Figure 13. Reorganized covariance plot showing an optimal cutoff of features selection occurs around 0.35. Features above the threshold were considered to have high correlation with labels.

To find the features with the most informative values, neighborhood component analysis (NCA) [28] was used on the reduced set of 1918 features (using the MATLAB function fscnca [29]). NCA is an algorithm that allows the analysis of a linear

transformation of a dataset by optimizing k-nearest neighbor performance and can result in an estimate of feature weights in addition to training a classifier. In this respect, it is very similar to principal component analysis (PCA) because it calculates a few important principal components while discarding the rest [30]. The advantage of using NCA over PCA is that NCA offers a more accurate representation of the dataset based on dimensionality while PCA takes a matrix and reduces it linearly. FSCNCA gives us feature importance weights that tell us the most descriptive features. Since NCA is searching for the values of a large matrix (about 2000x2000 in this case) and we only have 757 data samples, the problem is underdetermined. Therefore, we use a regularization penalty to ensure that fewer than the full number of features is used. The parameter lambda (λ) is used to set the degree of regularization. The λ value used was 0.0126 and we used a threshold of 0.01 to determine a final list of the most important features. This was chosen by varying λ and examining the training loss (a measure of classification accuracy). This plot is shown in Figures 15 and 16. The NCA model results in feature weights that allow us to assess which features and parameters are more important for classification, as well as a classifier.

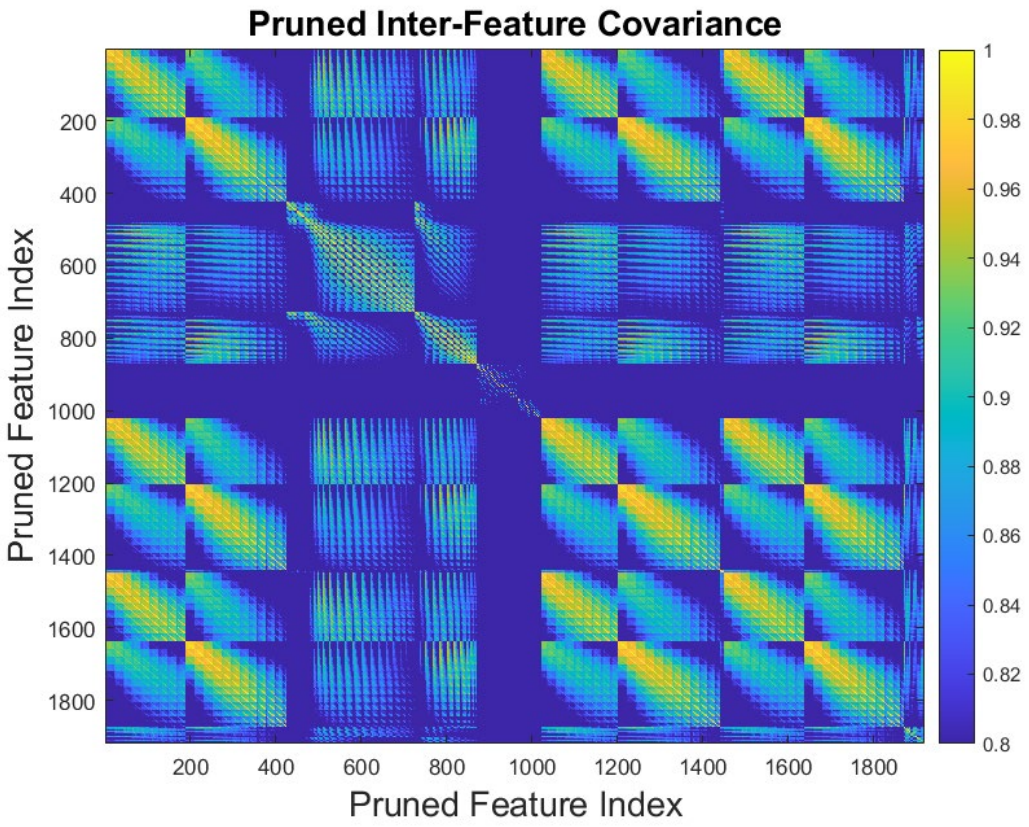


Figure 14. A pruned version of the covariance plot, but only using the features that were above the 0.35 correlation threshold.

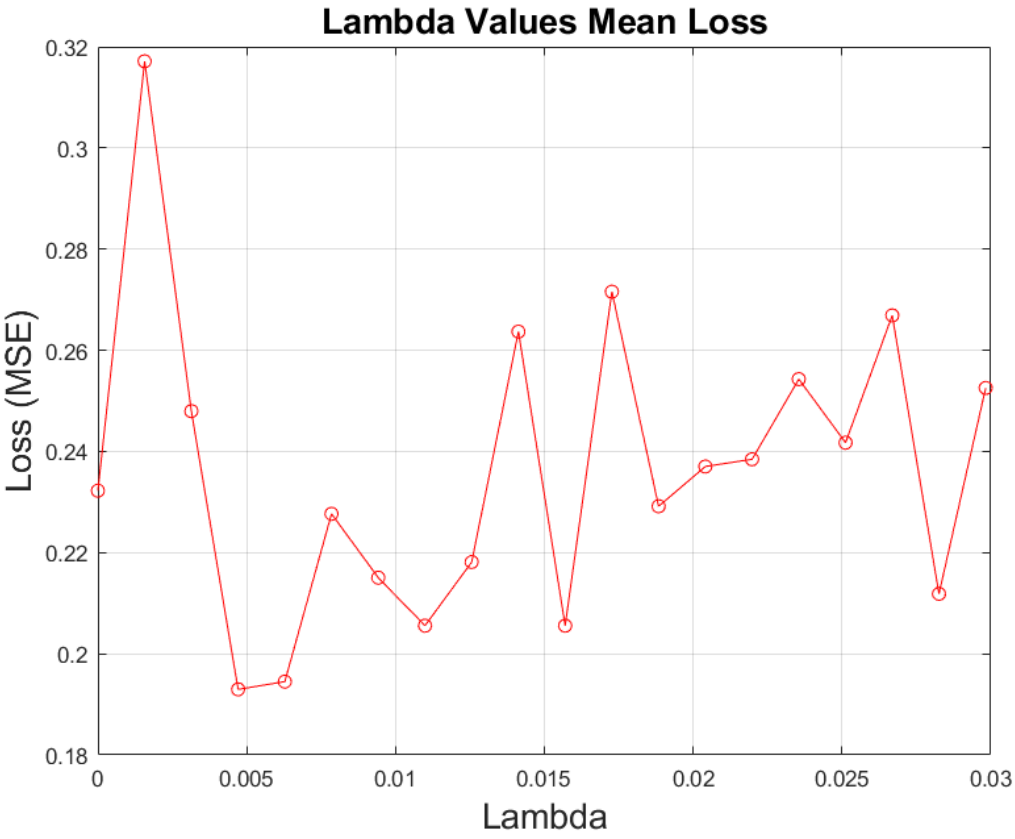


Figure 15. Graph of Lambda values compared to its calculated Loss value. Lambda values with low loss were preferable to values with higher loss values.

V. RESULTS AND DISCUSSION

Once the GLCMs are processed and the results of which features and what parameters are the most descriptive is found, the main question is what is driving the successful features? What patterns are the features capturing?

Using t-SNE, or t-distributed stochastic neighbor embedding, the data was able to be organized in both two and three-dimensional maps for visualization of classification clustering [18].

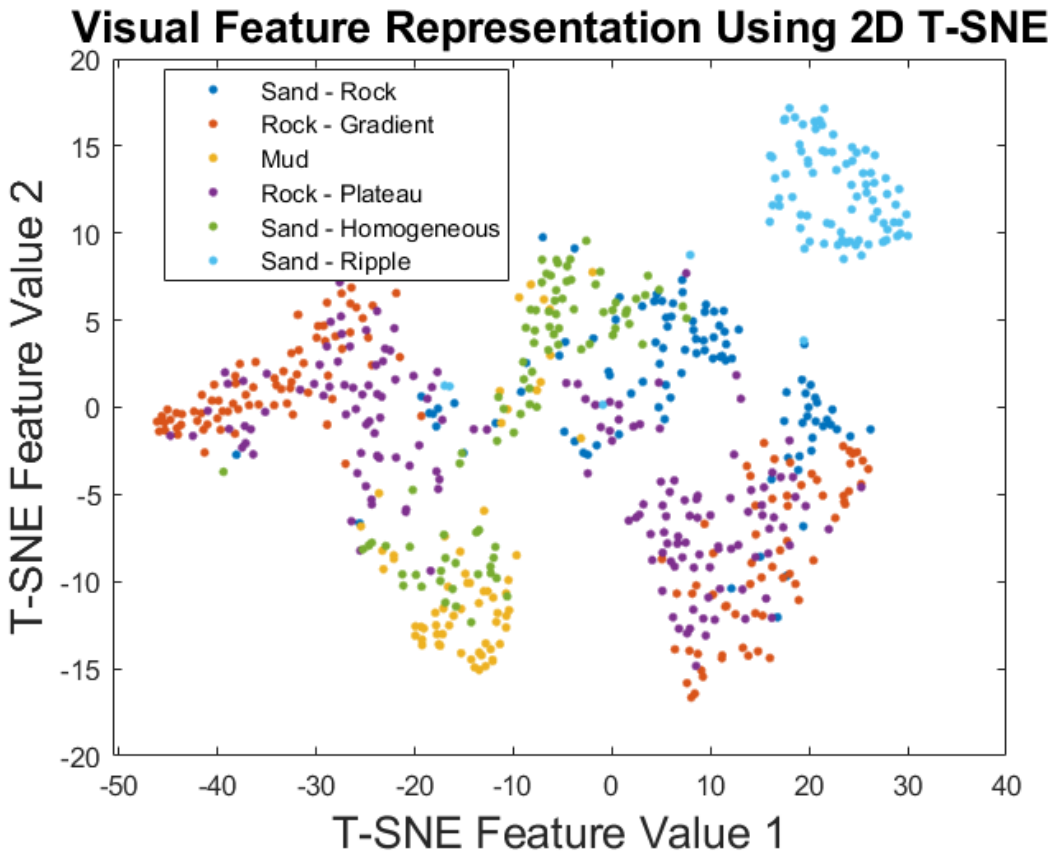


Figure 16. A scatter plot showing the clustering of different classifications

This classification can also be seen using t-SNE mapping to determine which features and labels have the highest levels of clustering and segmentation. Sand – Ripple

examples are well-defined and can be seen clearly separated from the other feature types. Using three-dimensional t-SNE mapping, different angles can be observed that show the distinction between three clusters of classifications.

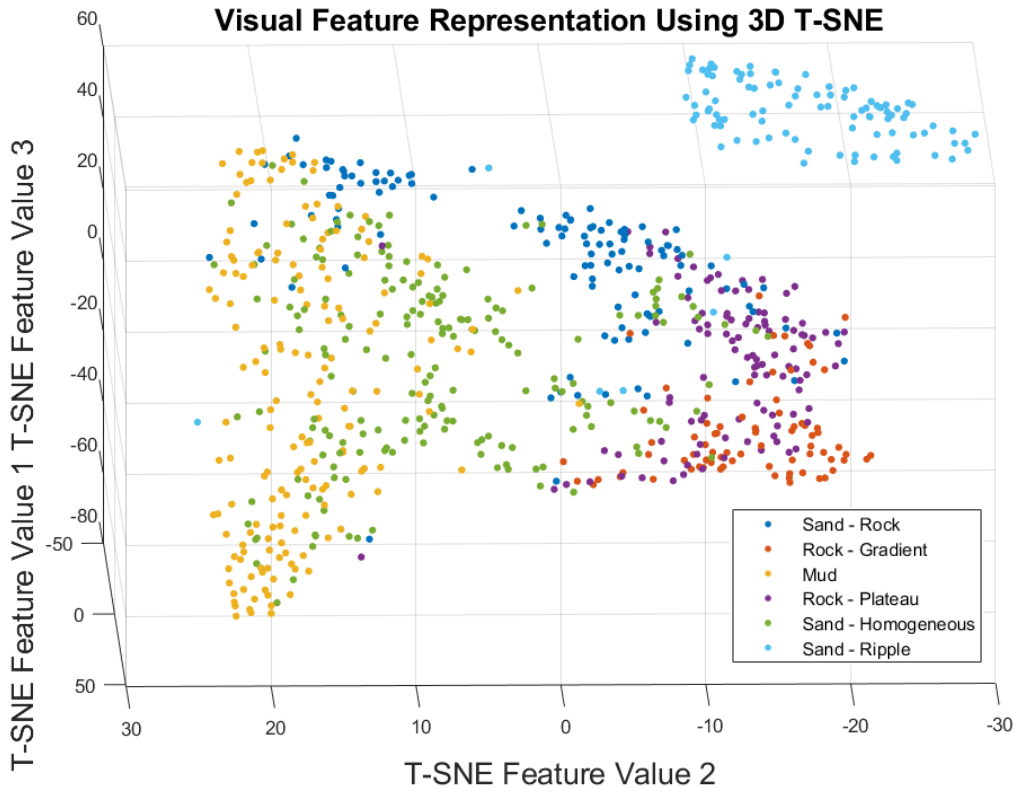


Figure 17. Three-dimensional t-SNE map showing three distinct clusters of identification. The color map for these observations is consistent with the previous two-dimensional t-SNE map in Figure 16.

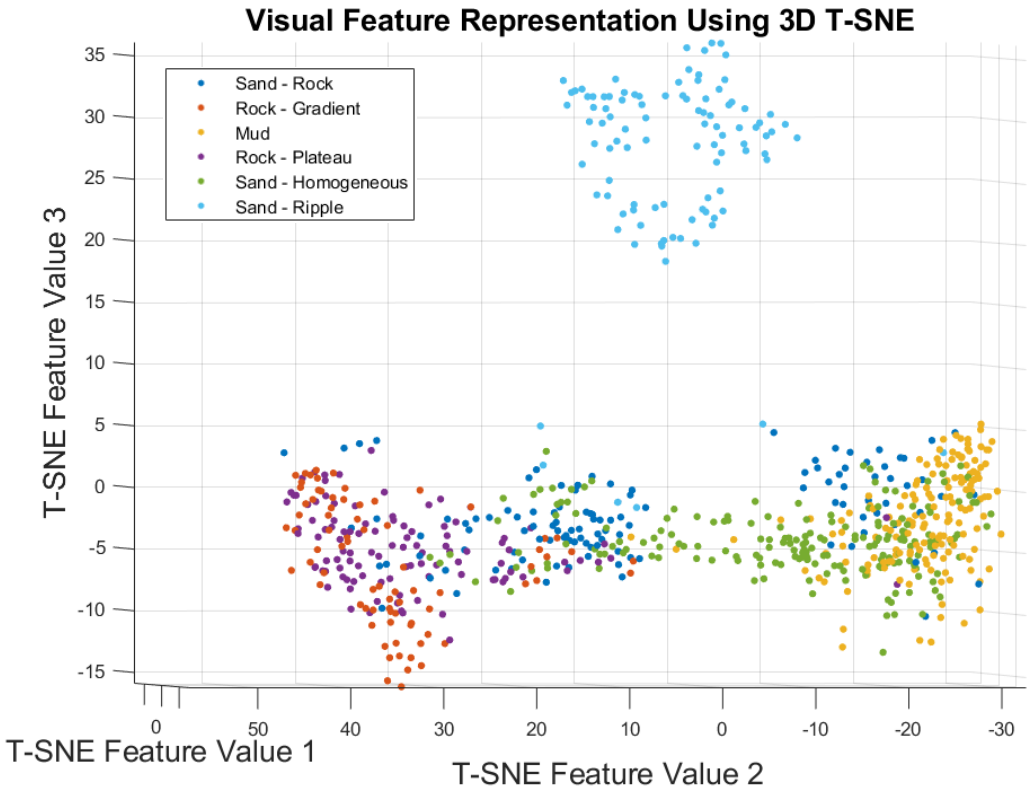


Figure 18. The same three-dimensional t-SNE map showing three distinct clusters of identification but rotated to illustrate different orientation.

It is worth noting that it still needs to be answered on what visual or perceptual characteristics are being picked up and why those characteristics are being captured by these good feature and parameters. In one cluster, just like in the two-dimensional t-SNE map, Sand – Ripple is clearly defined and separated from the other features. Another cluster is seen composing of predominantly Mud, Sand – Homogeneous, and Sand – Rock. The last cluster is composed predominantly of Sand – Rock, Rock – Plateau, and Rock – Gradient. Sand – Rock seemed to be the most inconsistent feature, being heavily represented by two distinct clusters. However, this is also intuitively logical since this classification is a hybrid classification. Sometimes the rocks in the Sand – Rock category caused the features to be grouped with the Rock – Gradient and Rock – Plateau observations. Other times, Sand – Rock was clustered with the Mud and Sand – Homogeneous observations since the sand in the snippet was the dominant feature. Mud and Sand – Homogenous snippets being clustered together also makes sense because of the

homogenous spatial patterns found in both, even if Mud was more tonally intense. The mean values taken out of these calculations meant that tonal intensity did not factor into classification. Features in the Rock – Plateau, Rock – Gradient, and Sand – Rock cluster often had sharply defined shapes and patterns of high and low intensity.

To further understand what patterns are being captured with the calculated GLCM features, example snippets of typical and marginal observations were chosen for illustrative purposes. In the Figure 19 and Figure 20, example snippet types are shown in two classes. Typical observations were observations that were centered in the cluster of their respective class. Marginal observations were observations that were found on the fringes of their respective clusters or were outliers. Figure 21, Figure 22, Figure 23, and Figure 24 are example GLCMs calculated from their respective snippets.

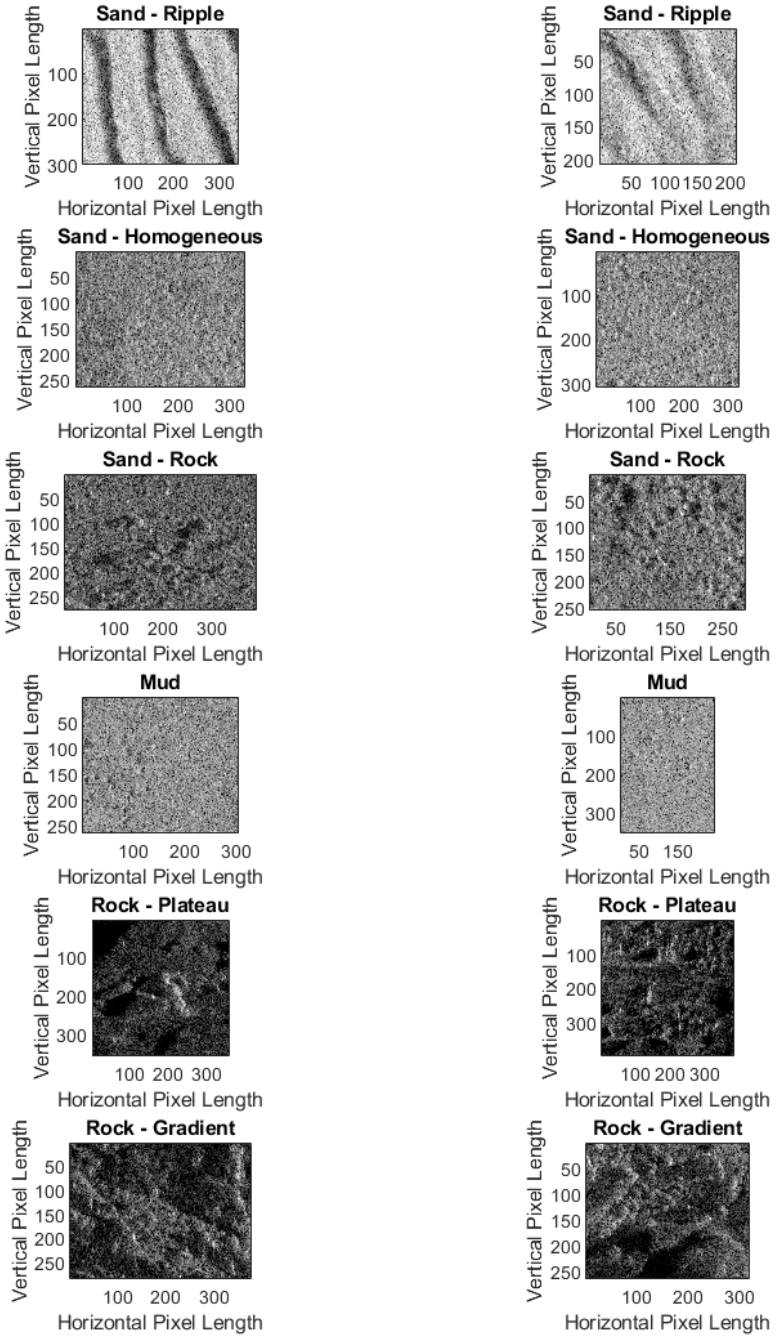


Figure 19. Example snippets of typical observations in their respective classification.

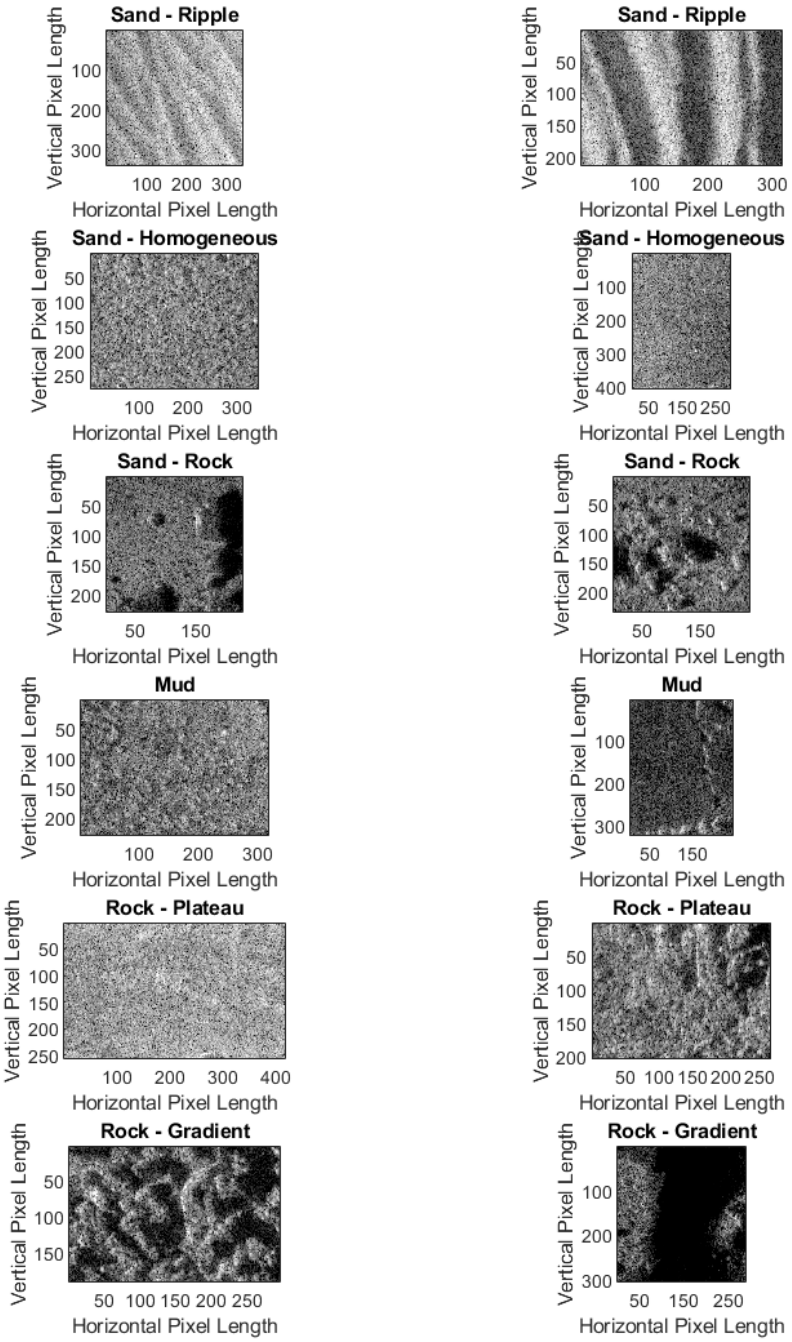


Figure 20. Example snippets of marginal observations in their respective classification.

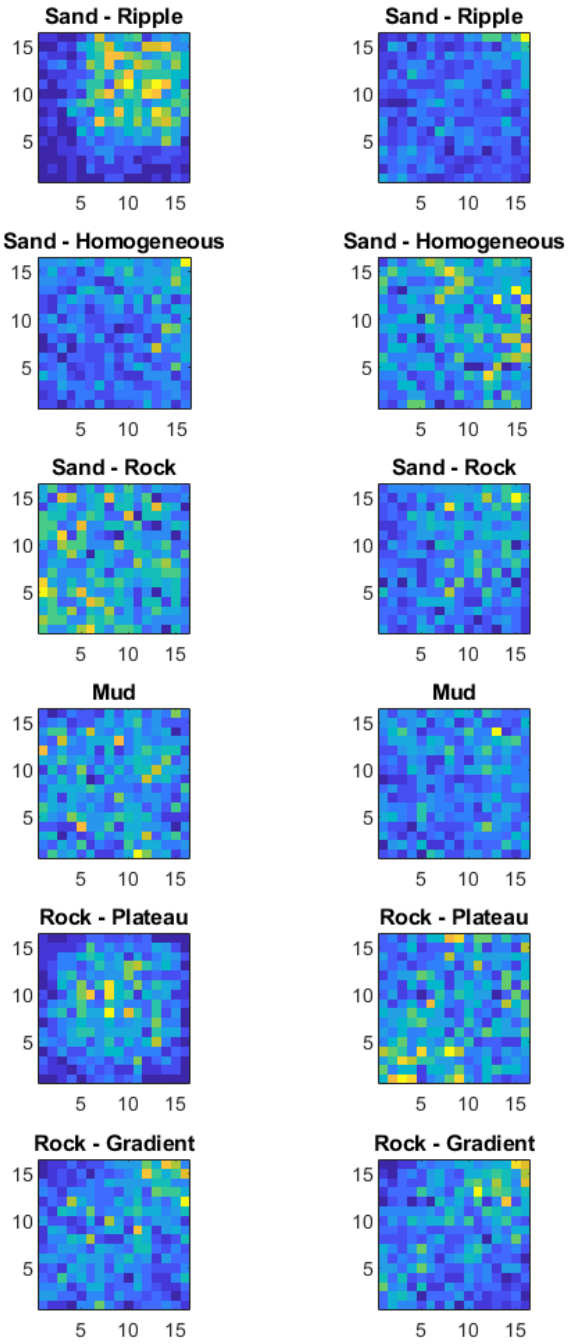


Figure 21. GLCM of typical observations with the feature/parameter combination with the highest weight.

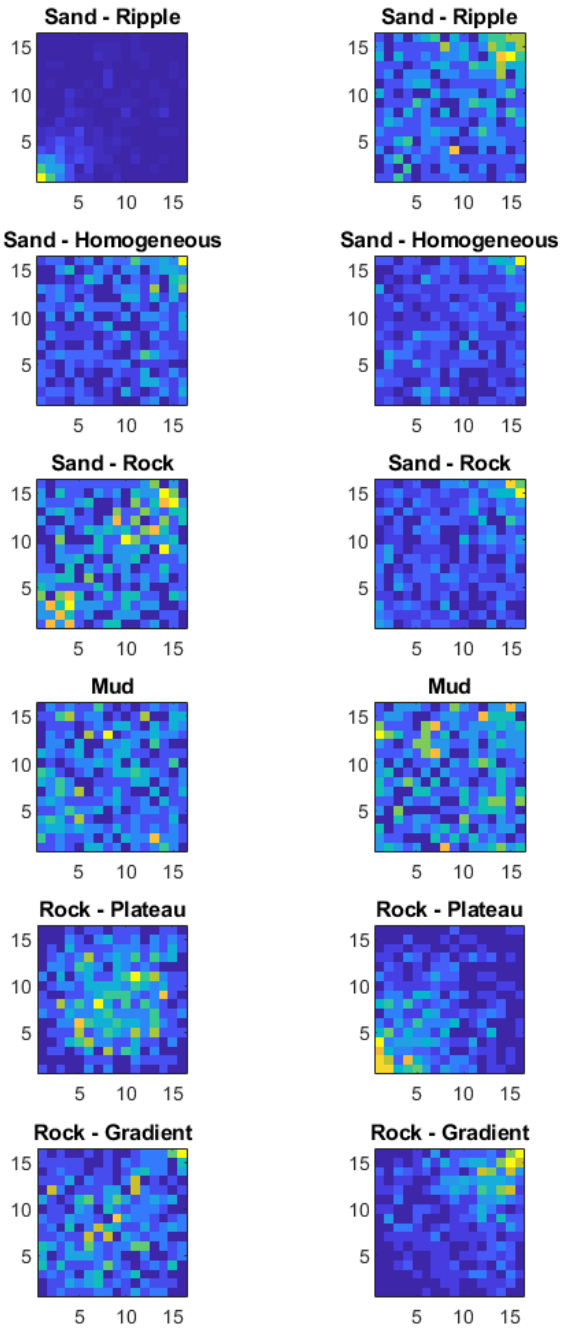


Figure 22. GLCM of typical observations using the feature/parameter combination with the second highest weight.

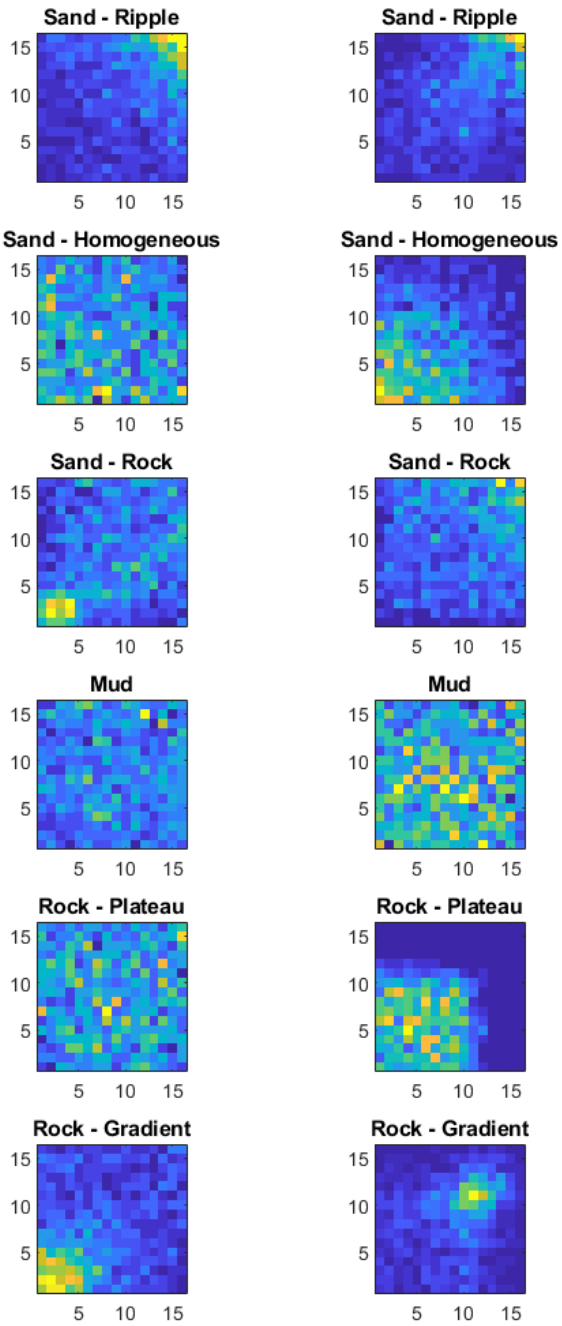


Figure 23. GLCM of marginal observations using the feature/parameter combination with the highest weight.

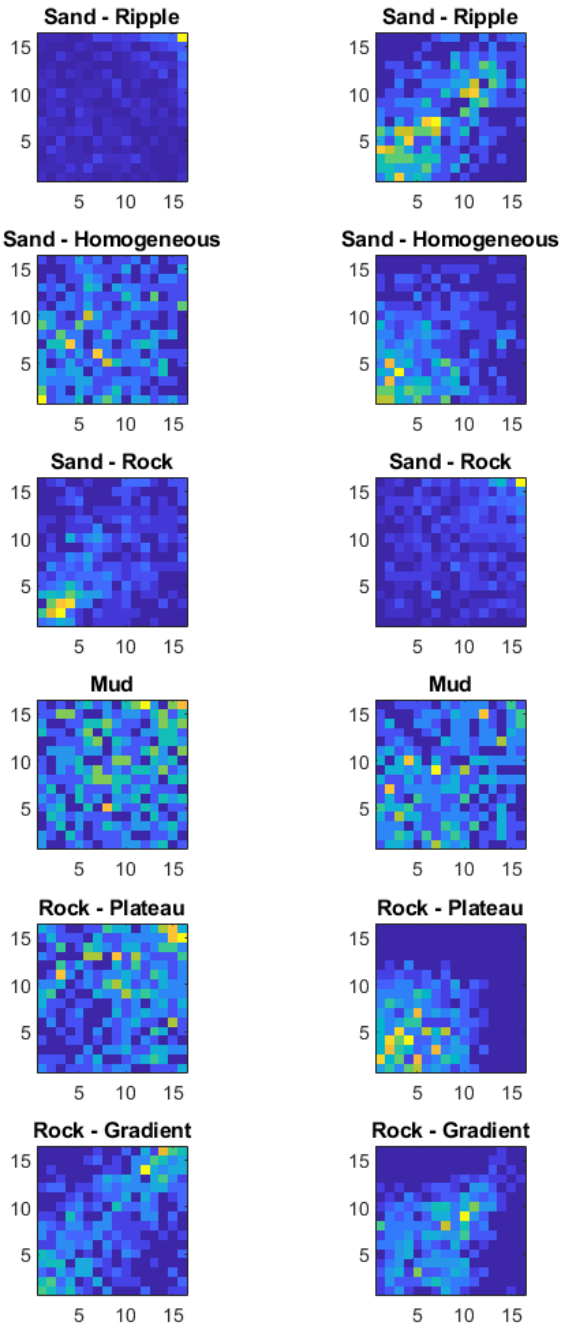


Figure 24. GLCM of same marginal observations but using the feature/ parameter combination with the second highest weight.

Looking at the differences between the typical and marginal observations, the typical observations in Figure 19 were more spatially consistent throughout their snippets. This made it so the features were easily recognized and classified. Marginal observations in Figure 20 had more variation in the patterns found within the snippets. For example, marginal rock cases had mostly bright pixels with few dark ones. As opposed to the typical rock snippets which had consistently high levels of dark pixels and very little bright ones. In both typical and marginal observations, mud and homogeneous sand were very similar due to their consistency of pixels. In the GLCMs above, it can be observed that in the GLCMs for snippets labeled Sand – Ripples, the GLCMs tend to have bright and dark pixels that are clustered together. In contrast, Sand – Ripple snippets, snippets labeled Sand – Homogenous and Mud have a more distributed layout in both dark and bright pixels.

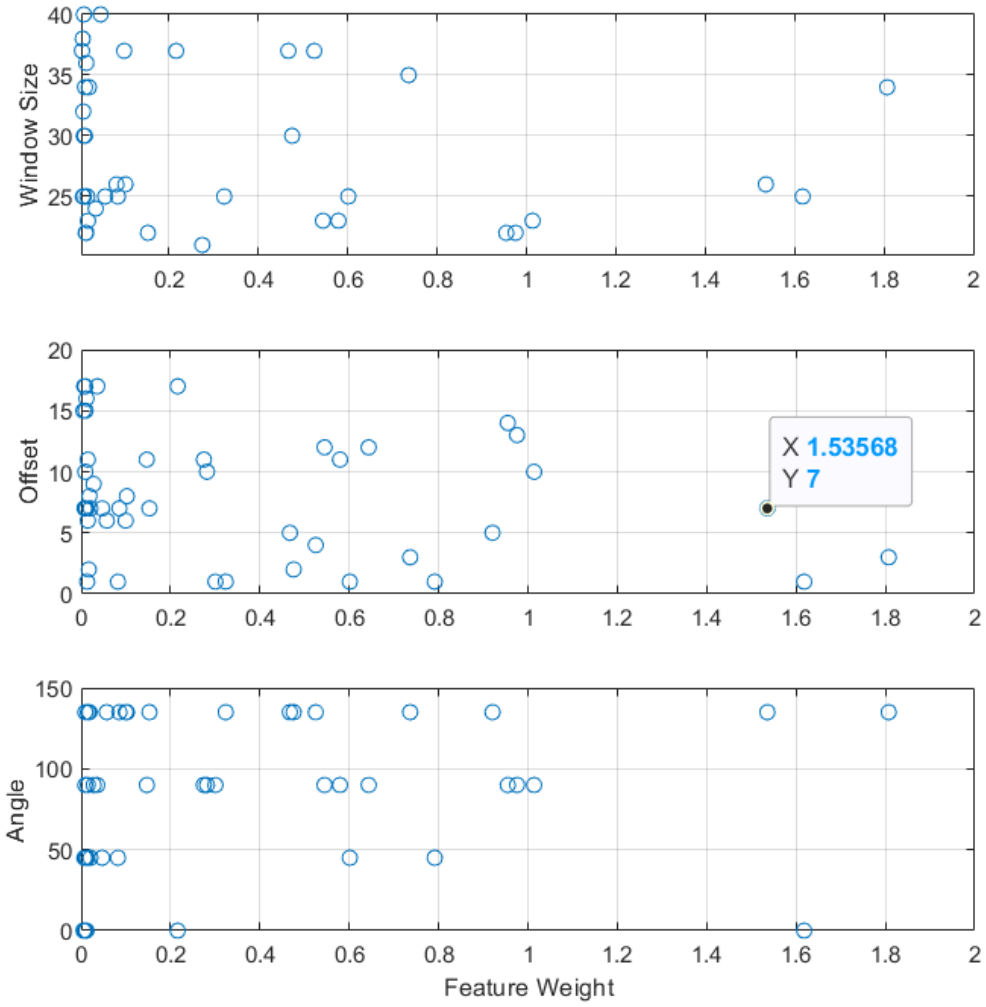


Figure 25. Top 50 features compared by Feature weight and three parameters.

Table 1. Top 20 most important features in descending order (top features at the top)

Type of Feature	Window Size	Offset	Grey Level	Angle
Cluster Shade	35	3	16	135
Cluster Shade	21	1	16	0
Cluster Shade	24	7	16	135
Dissimilarity	23	11	16	90
Cluster Shade	31	1	16	135
Cluster Shade	26	8	16	135
Cluster Shade	22	13	16	90
Dissimilarity	20	5	16	135
Cluster Shade	21	11	16	90
Information Correlation	25	25	16	45
Contrast	25	16	16	135
Cluster Shade	21	1	16	45
Cluster Shade	34	3	16	135
Cluster Shade	26	7	16	135
Cluster Shade	20	15	16	90
Cluster Shade	31	1	16	0
Cluster Shade	22	14	16	90
Cluster Shade	20	1	16	90
Cluster Shade	20	12	16	90
Cluster Shade	20	1	16	0

In the table above, the top twenty most useful features were ranked in descending order (most useful at the top). The two features that found that had the highest level of correlation with their labels were Cluster Shade and Dissimilarity. This list is provided with the spatial scales used in Table 1. The grey level scale is consistent because only one grey level was used. According to Soh [25], grey levels past 16 produced similar results and there was an exponential decrease in useful information. This was consistent with the findings in this study, so grey levels were limited at 16. Figure 25 shows the feature weight with the three variable parameters. The window sizes, offsets, and angles were all widely varied. This is useful because it means that the patterns of the pixels all have informative features across a varied spatial scale. The offset values with the highest feature weights tended to be on the lower scale of the values measured. This implies that smaller offset, and a finer resolution increases the amount of detail able to be gathered from an image.

Using a confusion matrix (Figure 26) to test the accuracy of the classifications, it can be seen to have a high prediction rate, especially for snippets labeled for rock. The confusion matrix was created using 120 of the initial 757 snippets as test data (637 snippets used for training). Overall, the current parameters possess an accuracy rate of 74.17%. The precision for each class varied, but the lowest performer was the Mud and Sand – Homogeneous categories at 57.14% and 56.25% respectively. It was often the case that snippets from both categories were mislabeled as each other, as they appear similar on a GLCM. Another area of incorrectly classified snippets was those that were labeled Rock – Gradient but were mislabeled as Rock – Plateau. The current parameters are good at identifying rocks, but not between different types of rocks. This implies that the characteristics being picked up on GLCMs are similar on both types of classes of rock. Table 1 and Table 2 show the difference in precision rates when the rock categories are combined. Sacrificing the granularity in rock classifications, combining the rock categories also increases accuracy from 74.17% to an overall 82.50%.

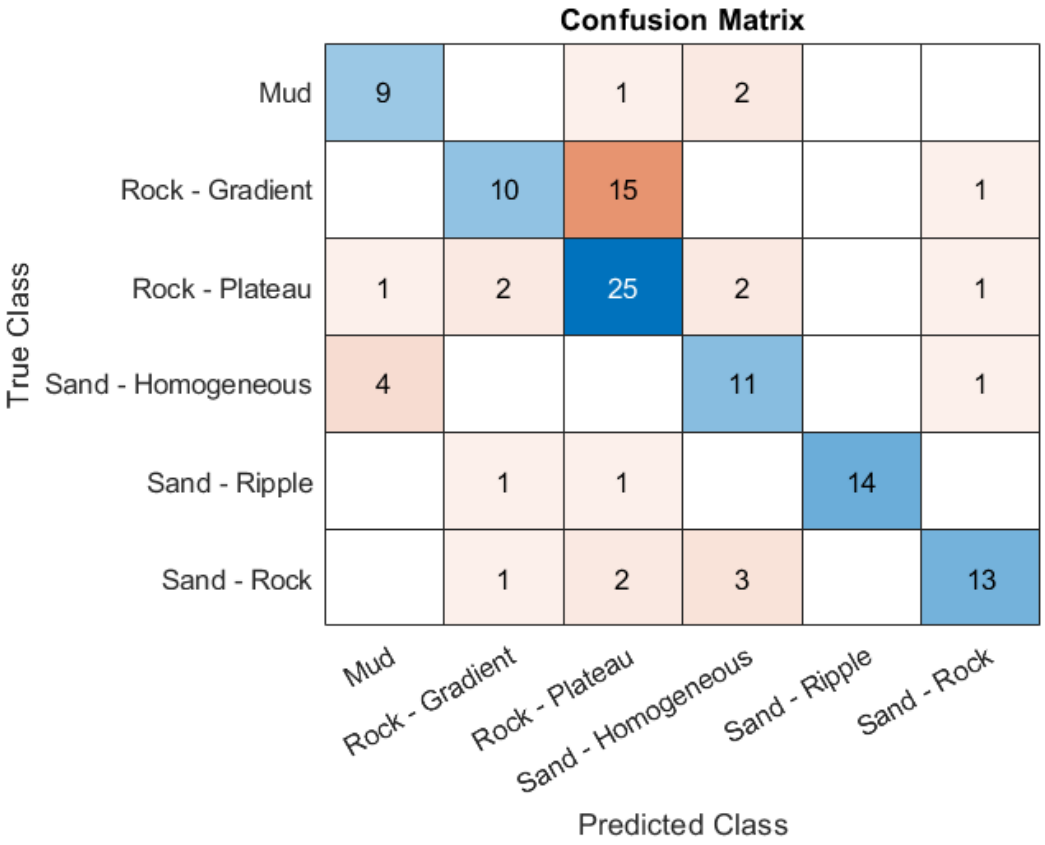


Figure 26. A confusion matrix showing how accurate an algorithm is at determining the correct label for a snippet based on GLCM calculated values. The blue values are correct classifications while the red values are incorrect classifications.

Table 2. Precision rates with all six category labels

Label Class	Precision Rate
Mud	57.14%
Rock – Gradient	81.82%
Rock – Plateau	66.67%
Sand – Homogeneous	56.25%
Sand - Ripple	100%
Sand – Rock	93.75%

Table 3. Precision rates with rock categories combined

Label Class	Precision Rate
Mud	57.14%
Rock – Gradient and Plateau Combined	88.52%
Sand – Homogeneous	56.25%
Sand - Ripple	100%
Sand – Rock	93.75%

VI. CONCLUSION

Textural analysis and statistical investigation of patterns in SAS images has potential in deepening the current understanding of the underwater environment. The tactical implications of applying textural analysis can further strengthen the U.S. Navy's understanding in warfare areas such as mine identification and undersea warfare. GLCMs can be an effective tool for classification and segmentation of underwater data collected via sonar, but its application can be applied to any type of image that can be tonally or spatially evaluated.

Future research should focus on using multiple locations and more variety in environmental features. All the SAS data used in this paper were collected in the North Sea, off the coast of Bergen, Norway. Other locations with image textures resulting from differing geomorphologies could offer a deeper understanding of how different environmental features affect GLCM calculations and subsequently, how that data can be used in classification. This data can further be expanded by integrating modern methods of target identification algorithms.

Naval research can always benefit from an increased understanding of the oceanographic environment. Textural analysis is still poorly understood in terms of quantitative metrics. There are potential research opportunities to be found by integrating textural analysis techniques to the current understanding of oceanography.

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF REFERENCES

[1] A. D'Amico and R. Pittenger, "A brief history of active sonar," *Aquatic Mammals*, vol. 35, no. 4, pp. 426–434, Dec. 2009, doi: 10.1578/am.35.4.2009.426.

[2] C. Chen, A. Zare, and J. Cobb, "Invariant parameter estimation across varying seabeds in synthetic aperture sonar imagery," Proceedings of the Institute of Acoustics, 2014.

[3] M. P. Hayes and P. T. Gough, "Synthetic aperture sonar: A review of current status," *IEEE Journal of Oceanic Engineering*, vol. 34, no. 3, pp. 207–224, Jul. 2009, doi: 10.1109/joe.2009.2020853.

[4] R. Urick, "The Backscattering of Sound from a Harbor Bottom," *Citation: The Journal of the Acoustical Society of America*, vol. 26, no. 26, p. 231, 1954.

[5] D. A. Abraham, *Underwater acoustic signal processing: Modeling, detection, and estimation*. Cham, Switzerland: Springer, 2019.

[6] A. Lyons and D. Abraham, "Statistical characterization of high-frequency shallow-water seafloor backscatter," The Journal of the Acoustical Society of America, Aug. 1999. Accessed: Aug. 21, 2022. [Online]. Available: https://asa.scitation.org/doi/abs/10.1121/1.428034.

[7] D. Olson and A. Lyons, "Resolution dependence of rough surface scattering using a power law roughness spectrum," *The Journal of the Acoustical Society of America*, vol. 149, no. 28, Jan. 2021, doi: 10.1121/10.0002974.

[8] A. Lyons, D. Olson, and R. Hansen, "Modeling the effect of random roughness on synthetic aperture sonar image statistics," *The Journal of the Acoustical Society of America*, vol. 152, no. 3, p. 1363, 2022, doi: 10.1121/10.0013837.

[9] D. Olson and A. Lyons, "Measurements of high-frequency acoustic scattering from glacially eroded rock outcrops," *The Journal of the Acoustical Society of America*, vol. 139, no. 1833, Apr. 2016, doi: 10.1121/1.4945589.

[10] D. Olson, A. Lyons, D. Abraham, and T. Sæbø, "Scattering statistics of rock outcrops: Model-data comparisons and Bayesian inference using mixture distributions," *The Journal of the Acoustical Society of America*, vol. 145, no. 761, Feb. 2019, doi: 10.1121/1.5089892.

[11] L. Mayer, R. Raymond, G. Glang, M. Richardson, P. Traykovski, and A. Trembanis, "High-resolution mapping of mines and ripples at the Martha's Vineyard coastal observatory," *IEEE Journal Of Oceanic Engineering*, vol. 32, no. 1, 2007, doi: 10.1109/JOE.2007.890953.

[12] A. Lyons and S. Johnson, "Modeling the effect of seafloor ripples on synthetic aperture sonar speckle statistics," *IEEE Journal Of Oceanic Engineering*, vol. 35, no. 2, 2010, doi: 10.1109/JOE.2009.2039656.

[13] D. P. Williams, "The Mondrian detection algorithm for sonar imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, no. 2, pp. 1091–1102, Feb. 2018, doi: 10.1109/tgrs.2017.2758808.

[14] A. P. Galusha, J. M. Keller, A. Zare, and G. Galusha, "A fast target detection algorithm for underwater synthetic aperture sonar imagery," *Detection and Sensing of Mines, Explosive Objects, and Obscured Targets XXIII*, vol. 10628, Apr. 2018, doi: 10.1117/12.2304976.

[15] R. M. Haralick, K. Shanmugam, and I. Dinstein, "Textural Features for Image Classification," *IEEE*, vol. 3, no. 6, pp. 610–621, Nov. 1973, doi: 10.1109/ TSMC.1973.4309314.

[16] Ph. Blondel, "Automatic mine detection by textural analysis of COTS sidescan sonar imagery," *International Journal of Remote Sensing*, vol. 21, no. 16, pp. 3115–3128, Jan. 2000, doi: 10.1080/01431160050144983.

[17] A. Zare, N. Young, D. Suen, T. Nabelek, A. Galusha, and J. Keller, "Possibilistic fuzzy local information C-Means for sonar image segmentation," *2017 IEEE Symposium Series on Computational Intelligence (SSCI)*, Nov. 2017, doi: 10.1109/ssci.2017.8285358.

[18] R. Haralick, "Statistical and Structural Approaches to Texture," *Proceedings Of The IEEE*, vol. 67, no. 5, 1979.

[19] T. O. Sæbø, H. J. Callow, R. E. Hansen, B. Langli, and E. O. Hammerstad, "Bathymetric capabilities of the HISAS interferometric synthetic aperture sonar," *OCEANS 2007*, Sep. 2007, doi: 10.1109/oceans.2007.4449357.

[20] Ph. Blondel and O. Gómez Sichi, "Textural analyses of multibeam sonar imagery from Stanton Banks, Northern Ireland continental shelf," *Applied Acoustics*, vol. 70, no. 10, pp. 1288–1297, Oct. 2009, doi: 10.1016/j.apacoust.2008.07.015.

[21] A. Lehman, "Resolution dependence of acoustic scattering statistics for complex seafloors," *Naval Postgraduate School*, 2022.

[22] D. Lu and Q. Weng, “A survey of image classification methods and techniques for improving classification performance,” *International Journal of Remote Sensing*, vol. 28, no. 5, pp. 823–870, Mar. 2007, doi: 10.1080/01431160600746456.

[23] S. Singh, D. Srivastava, and S. Agarwal, “GLCM and Its Application in Pattern Recognition,” IEEE, Oct. 2017. Accessed: Aug. 12, 2022. [Online]. Available: https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8053537

[24] F. Albregtsen, “Statistical Texture Measures Computed from Gray Level Coocurrence Matrices,” University of Oslo, Nov. 2008. Accessed: Sep. 02, 2022. [Online]. Available: https://www.uio.no/studier/emner/matnat/ifi/INF4300/h08/undervisningsmateriale/glcm.pdf

[25] L.-K. Soh and C. Tsatsoulis, “Texture analysis of SAR sea ice imagery using gray level co-occurrence matrices,” *IEEE Transactions On Geoscience And Remote Sensing*, vol. 37, no. 2, 1999.

[26] D. A. Clausi, “An analysis of co-occurrence texture statistics as a function of grey level quantization,” *Canadian Journal of Remote Sensing*, vol. 28, no. 1, pp. 45–62, Jan. 2002, doi: 10.5589/m02-004.

[27] A. Uppuluri, “GLCM texture features,” *www.mathworks.com*, Nov. 25, 2008. https://www.mathworks.com/matlabcentral/fileexchange/22187-glcm-texture-features (accessed Sep. 02, 2022).

[28] J. Goldberger, S. Roweis, G. Hinton, and R. Salakhutdinov, “Neighbourhood Components Analysis,” Association for Computing Machinery, Department of Computer Science, University of Toronto, Dec. 2004.

[29] “Neighborhood Component Analysis (NCA) Feature Selection - MATLAB & Simulink,” *www.mathworks.com*. https://www.mathworks.com/help/stats/neighborhood-component-analysis.html

[30] K. Zakka, “kNN classification using Neighbourhood Components Analysis,” *Kevin Zakka’s Blog*, Feb. 10, 2020. https://kevinzakka.github.io/2020/02/10/nca/ (accessed Aug. 12, 2022).

[31] “Visualize High-Dimensional Data Using t-SNE - MATLAB & Simulink,” *www.mathworks.com*. https://www.mathworks.com/help/stats/visualize-high-dimensional-data-using-t-sne.html (accessed Aug. 22, 2022).

THIS PAGE INTENTIONALLY LEFT BLANK

INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center
   Ft. Belvoir, Virginia

2. Dudley Knox Library
   Naval Postgraduate School
   Monterey, California