

# Comparison of Prescreening Algorithms for Target Detection in Synthetic Aperture Sonar Imagery

Princess Lyons<sup>a</sup>, Daniel Suen<sup>a</sup>, Aquila Galusha<sup>b</sup>, Alina Zare<sup>a</sup>, and James Keller<sup>b</sup>

<sup>a</sup>University of Florida, Gainesville, Florida

<sup>b</sup>University of Missouri, Columbia, Missouri

## ABSTRACT

Automated anomaly and target detection are commonly used as a prescreening step within a larger target detection and target classification framework to find regions of interest for further analysis. A number of anomaly and target detection algorithms have been developed in the literature for application to target detection in Synthetic Aperture Sonar (SAS) imagery. In this paper, a comparison of two anomaly and one target detection algorithm for target detection in synthetic aperture sonar is presented. In the comparison, each method is tested on a large set of real sonar imagery and results are evaluated using receiver operating characteristic curves. The results are compiled and quantitatively shown to highlight the strengths and weakness of the variety of approaches within various sea-floor environments and on particular target shapes and types.

**Keywords:** Automated Anomaly Detection, Receiver Operating Characterist Curves, RX Algorithm, Synthetic Aperture Sonar, Target Detection

## 1. INTRODUCTION

Often target detection systems employ an initial prescreening stage where regions of interest that may contain targets are identified for further investigation. One aim of the prescreening stage is to reduce overall computation by applying any more computationally complex analysis only on the identified regions of interest (as opposed to the entire data set). Thus, we aim for a prescreening stage that is very fast in implementation and can identify regions of interest that include all targets in the scene under test with a minimal number of false alarms. In this work, two anomaly detectors and one target detection algorithm are considered as prescreeners for application to target detection in Synthetic Aperture Sonar data. Anomaly detection refers to algorithms where there are no *a priori* target information on which to make a detection decision.<sup>1</sup> They are based solely on nonconformity to a background model, since no other information exists when used in an unsupervised manner.<sup>1</sup> In contrast, a target detector leverages assumed target information to search for regions that appear target-like in the scene.<sup>2</sup>

In many applications of SAS imagery, a side-scan imaging system is attached to an autonomous underwater vehicle (AUV) where measurements of a reflected pulse that is emitted toward an object is recorded and used to form high-resolution intensity images.<sup>3</sup> Objects of interest within a SAS image are usually accompanied by a highlight-shadow pattern due to side-scan image acquisition. These patterns occur in the SAS images when objects have a high reflectivity and occlude the regions behind them. The corresponding shadows and surrounding background have little to no reflected beam measurements.<sup>3</sup> Both high and low frequency images (based on the frequency of the emitted pulse) can be measured with SAS systems. Generally, high frequency SAS imagery contain detailed texture information over an imaged scene whereas low frequency imagery contains less fine detail but provides the ability to penetrate and look through some materials in the scene.

Variability of imaging systems, target types, and underwater environmental conditions and differences are key factors that are taken into account when creating systems that can accurately detect or identify targets.<sup>4</sup> For

---

Further author information: (Send correspondence to P. L.)

P. L.: E-mail: plyons@ufl.edu

D. S.: E-mail: dsuen1@ufl.edu

A. G.: Email: apgqpc@mail.missouri.edu

A. Z.: Email: azare@ece.ufl.edu

J. K.: Email: KellerJ@missouri.edu

example, basic detection systems focus on detecting anomalies or specific pixels in the scene that differ from their surroundings.<sup>5</sup> A difficulty in underwater target detection is the reverberation level which is a strong function of the seafloor composition.<sup>6</sup> If a seafloor contains various textures such as sand ripple, hard-packed sand or rocky areas, care must be taken to account for this type of background variation. The seafloor information can then be used to distinguish and characterize a target within an image. The use of prescreeners in these environments can enable detection systems to reduce the search area where more computationally intensive algorithms are applied.

Some current methods in the literature for target detection in SAS imagery use matched filter techniques. In these techniques, a template of the target’s highlight-shadow pattern is used to scan the image and produce correlation measures for detection. One improvement on the simple matched filter method was done using a frequency-domain adaptive matched filter.<sup>7</sup> This implementation showed significant improvement of the matched filter method alone.<sup>7</sup> Another method presented in the literature that uses the geometry of the target and highlight-shadow pattern to generate regions of interest to be processed later in a cascade architecture had success in detecting targets in SAS imagery.<sup>6</sup> It also benefits from its range-independent nature by accounting for the nonlinear nature of the sonar process. It acknowledges the fact that at different ranges and orientations, objects will produce completely different shadow regions<sup>6,8</sup>.

In this paper, we compare two anomaly detection methods, the Reed-Xiaoli and the Reed-Xiaoli Left-Right algorithm, and one target detection approach, the adaptive cosine estimator, that searches for highlight-shadow patterns. Section 2 details each method. Section 3 shows the results of the methods after testing on two underwater sites, Site A and Site B. Site A is characterized by its varied seafloor texture, many scattered non-target objects, and linear shaped targets. Site B is characterized by its relatively smooth seafloor and circular targets. Each method is tested on both sites. Three-fold cross-validation of the real sonar imagery is evaluated using receiver operating characteristic (ROC) curves. Section 5 discusses the results and performance of the methods.

## 2. METHODOLOGY

### 2.1 Reed-Xiaoli (RX) Algorithm

Anomaly detection is a common prescreening first step in target detection applications. The goal of anomaly detection is to locate anything that looks distinct from its surroundings.<sup>9</sup> In this paper, we apply the Reed-Xiaoli (RX) algorithm on our dataset of SAS imagery.<sup>9</sup> We recognize that the targets present in our data appear as highlight-shadow patterns surrounded by a majority of dark background pixels. The RX Algorithm therefore stands as a good candidate for anomaly detection in SAS imagery.

The RX detector<sup>9</sup> uses the statistics surrounding a center pixel to detect anomalies within the SAS image. The window that generates the output of the RX detector is described by two kernels. The center pixel is surrounded by an inner and outer kernel that calculates the variance and mean of the pixels within them. The inner and outer kernels are separated by a predefined number of pixels, making the outer kernel an annulus.<sup>9</sup> The parameters of the RX detector implementation used in this paper are 41 x 41 pixels for the inner kernel and 201 x 201  $\oplus$  151 x 151 pixels for the outer kernel ( $\oplus$ , symmetric difference).<sup>10</sup> All of the dimensions are square. The output of the RX detector is defined as

$$RX = \frac{(\mu_{in} - \mu_{out})^2}{\sigma_{out}^2}.$$

The variables  $\mu_{in}$  and  $\mu_{out}$  represent the means of the pixel intensity values of the pixels in the inner and outer kernel, respectively.  $\sigma_{out}^2$  is the variance of the outer kernel pixels.

The RX detector quantifies the difference between the pixels in the regions of the two kernels. If the detector’s output is high, the regions differ greatly. If the detector’s value is low, then the two regions have little difference. A visual of this two-kernel window is shown in figure 1(a).<sup>10</sup>

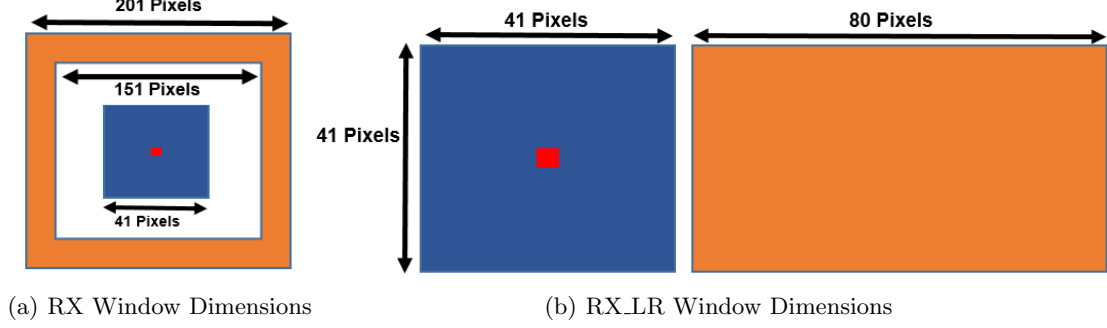


Figure 1. Kernel Dimensions for RX Methods

## 2.2 Reed-Xiaoli Left-Right (RX\_LR) Algorithm

We also employ a modified version of the RX algorithm which exploits the characteristics of side-scan SAS imagery.<sup>10</sup> Since targets cast a shadow on the far side of the SAS camera, the outer kernel of the RX detector is moved to use the statistics of the region of pixels to the right of the target. These pixels should be darker due to the shadow. The new kernel now becomes a rectangle that is basically surrounding the shadow cast by the target. This allows the detector to capture the entire highlight-shadow pattern, prompting the name RX Left-Right (RX\_LR). The RX\_LR detector is defined as:

$$RX\_LR = \frac{(\mu_{left} - \mu_{right})^2}{\sigma_{right}^2}.$$

A visual of this modified RX\_LR window along with its dimensions is shown in Figure 1(b).<sup>10</sup>

## 2.3 Adaptive Cosine Estimator (ACE)

A matched filter is an inner product detector that has commonly been used to detect the presence of a target within a given signal. The matched filter can be viewed as a supervised target detector (as compared to the unsupervised anomaly detectors previously described) since a target signature is required to be either known in advance or estimated from training data. ACE is a version of the matched filter that computes the inner product between the target signature and a test sample in a whitened coordinate space.<sup>2</sup> The ACE detector is defined as follows:

Let  $\mathbf{x}$  be a test data point and  $\mathbf{s}$  be a known target signature, then

$$D_{ACE}(\mathbf{x}, \mathbf{s}) = \frac{(\mathbf{s} - \boldsymbol{\mu}_b)^T \boldsymbol{\Sigma}_b^{-1} (\mathbf{x} - \boldsymbol{\mu}_b)}{\sqrt{(\mathbf{s} - \boldsymbol{\mu}_b)^T \boldsymbol{\Sigma}_b^{-1} (\mathbf{s} - \boldsymbol{\mu}_b)} \cdot \sqrt{(\mathbf{x} - \boldsymbol{\mu}_b)^T \boldsymbol{\Sigma}_b^{-1} (\mathbf{x} - \boldsymbol{\mu}_b)}},$$

where  $\boldsymbol{\mu}_b$  and  $\boldsymbol{\Sigma}_b$  are the image background mean and covariance, respectively. The above value computes the angle between whitened versions of the test data point and known target signatures. If the angle between the two points is small, the detector value will be high indicating similarity between the test data point and the target. If the angle between the two points is large, the detector value will be low indicating dissimilarity. In our implementation, the ACE detector value is squared and regarded as a confidence value that measures how likely it is that a point is or is not a target. In our application, each data point is a vectorized image chip (i.e., a small image window is reshaped into vector format).

Targets in general take up a minimal amount of image area and a given SAS image is composed overwhelmingly of background. Therefore, the background mean and variance to be used within ACE can be estimated using the sample mean and variance of the intensity values of each individual SAS image.

Here, ACE target signature is learned from the available ground truth data of SAS imagery. The dataset for each site is separated into three folds to facilitate three-fold cross validation. To learn target signatures for

use within ACE, target windows centered around each ground truth point, which we will call mugshots, are averaged. The mugshots consists of the 30 rows of pixels above and below the ground truth coordinate as well as 20 columns of pixels to the left and 90 columns to the right of it. Thus, the dimensions of the target window is 61 x 111 pixels, as shown in figure 2 along with example targets from Site A and B. The rectangular shape of the window along with its center being shifted to the left allows both the highlight and shadow pattern of the target to be extracted. Data points used for testing are constructed in the same way. Each pixel within the SAS image is enclosed by a 61 x 111 window, provided that the entire window fits within the boundaries of the image. Our choice for window dimensions is done in a ad hoc manner. For a given training fold, the mugshots are manually observed and grouped for similarity in appearance. Each group is averaged to create one target signature. With this technique, the number of target signatures per fold depends on the appearance of the mugshots in the fold. It is possible for one fold to have one or two target signatures while another may have five or six target signatures.

ACE is applied in a sliding window fashion across a SAS image. The confidence value is computed at every pixel using the square of the detection statistic,  $D_{ACE}^2(\mathbf{x}, \mathbf{s})$ , producing a confidence map for each learned target signature. For example, if the fold learns 2 target signatures, 2 confidence maps will be produced for that image. The maximum value among the multiple maps at each pixel is used to form one combined confidence map. Next, non-maximum suppression is performed on the combined confidence map to locate regions of interest with high confidence values. Peaks that fall within a predefined radius of a ground truth location are marked as detected targets. An example of this method is shown in figure 3 where a target is located in the bottom left region of the images.

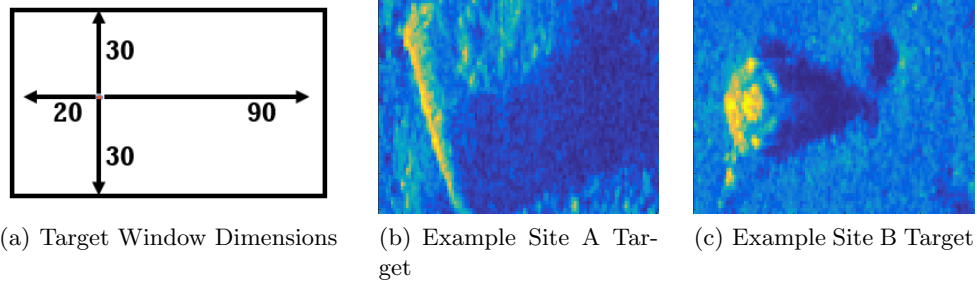


Figure 2. Window dimensions and examples of targets from Site A and Site B

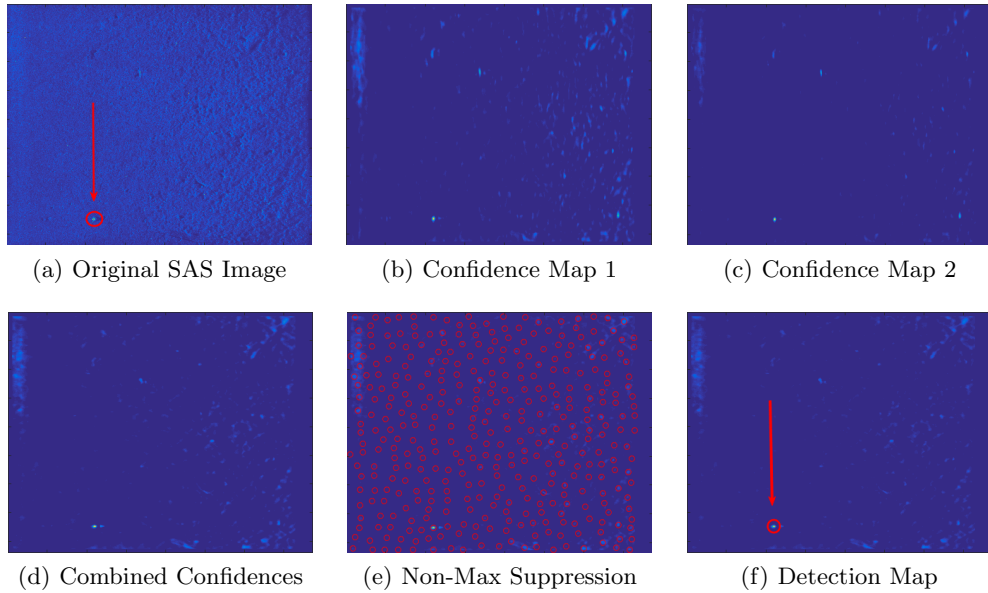


Figure 3. Example of ACE/MIL-ACE method with 2 learned target signatures. To aid with the visualization, the figures (b) - (f) have been transformed via the monotonic natural log function.

### 3. EXPERIMENTAL RESULTS

#### 3.1 ACE Target Signatures

The results in figures 4 and 5 show the learned target signatures of ACE for testing on fold 1 of both Site A and Site B, respectively. The top row of each figure shows the learned signatures using only the high frequency SAS images of the site. The bottom row shows the learned target signatures using only the low frequency SAS images of the site. Since mugshots are grouped by their appearance in each fold and for each frequency, it is possible for some folds to have more learned target signatures than others. For example, figures 4(a) - 4(d) and 5(a) - 5(d) are only 4 of the 7 learned target signatures for testing on fold 1 in both sites. Only 4 are shown as an example however, all 7 signatures are used for testing on fold 1.

Table 1 summarizes the number of learned target signatures from each fold for high and low frequency images, per site. To aid with the visualization, the figures have been transformed via the monotonic natural log function.

Table 1. Number of Learned Target Signatures. Target signatures are learned from 2 of 3 folds and used to test on the third fold.

Frequency	Site A		Site B	
	High	Low	High	Low
<b>Fold 1</b>	7	5	7	3
<b>Fold 2</b>	7	4	4	3
<b>Fold 3</b>	6	4	5	2

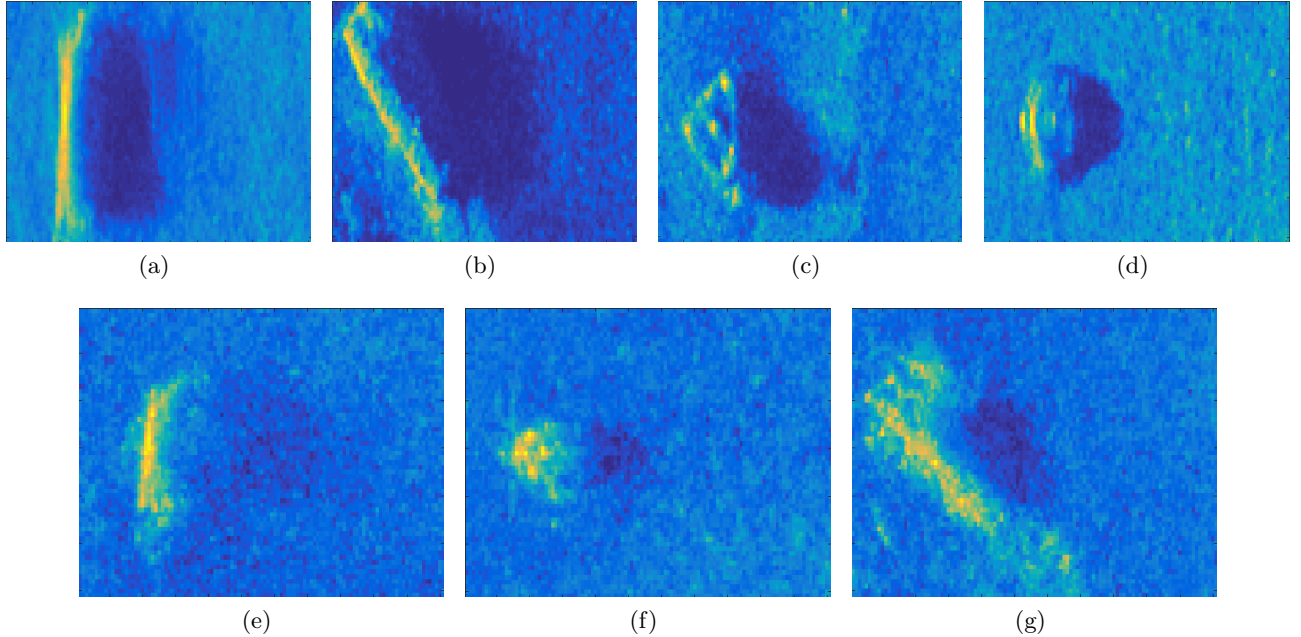


Figure 4. Site A ACE 3-Fold Cross Validated testing target signatures for Fold 1: (a) - (d) Learned high frequency target signatures; (e) - (f) Learned low frequency target signatures;

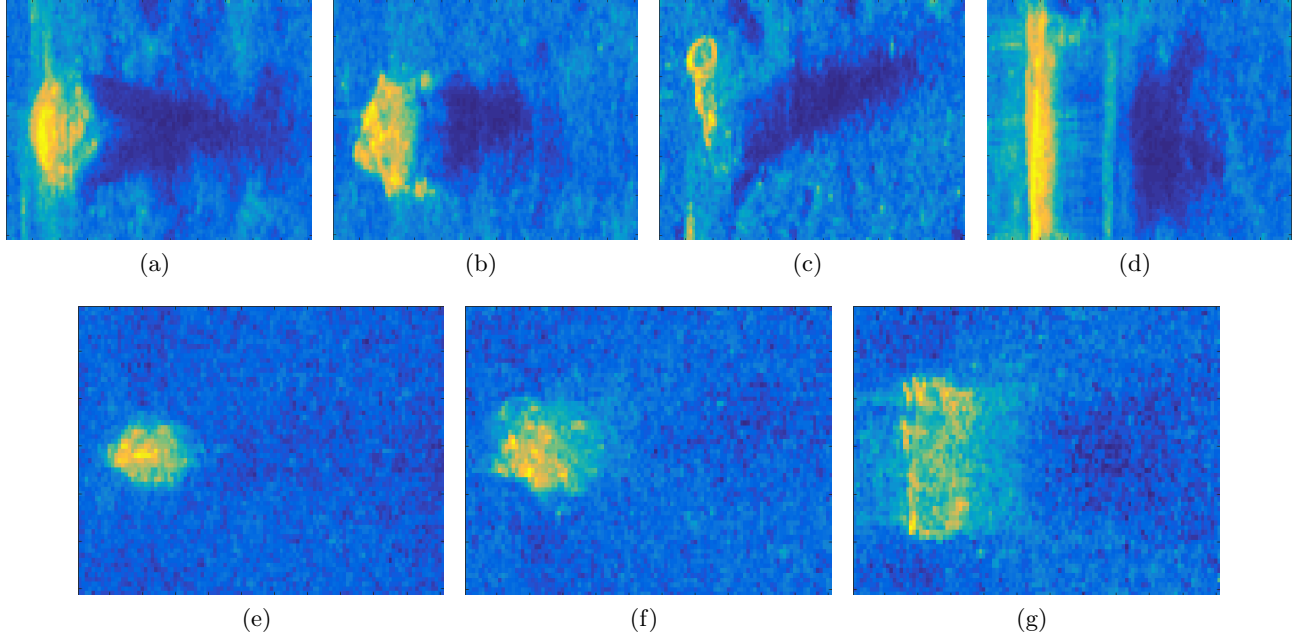


Figure 5. Site B ACE 3-Fold Cross Validated testing target signatures for Fold 1: (a) - (d) Learned high frequency target signatures; (e) - (f) Learned low frequency target signatures;

### 3.2 Quantitative Results

Figures 6 and 7 show the detection results for both Site A and B, respectfully. Each method's performance is displayed as an individual receiver operator characteristic (ROC) curve with detection being done on the high frequency SAS images alone and the low frequency SAS images alone.

Figure 6(c) and Figure 7(c) show the results of combining high and low frequency confidence maps to perform detection. The curve labeled 'Combined RX' was created by multiplying the high and low frequency confidence maps of the RX and RX\_LR results then performing detection. The multiplication was done pixel by pixel. The curve labeled 'ACE (High/Low)' was created by combining the high and low frequency confidence maps of the ACE method. The maximum value at each pixel is taken to create the combined confidence map and detection is then done using this map.

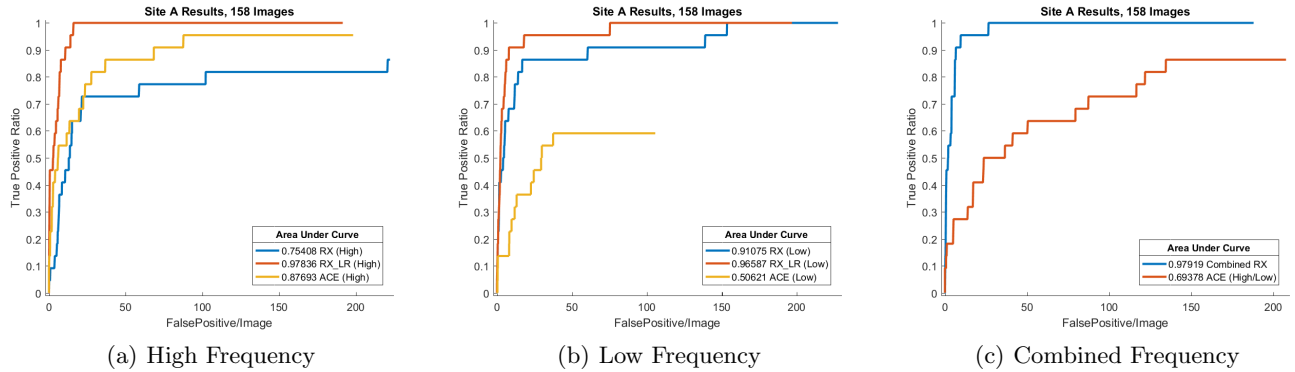


Figure 6. Site A ROC Curves



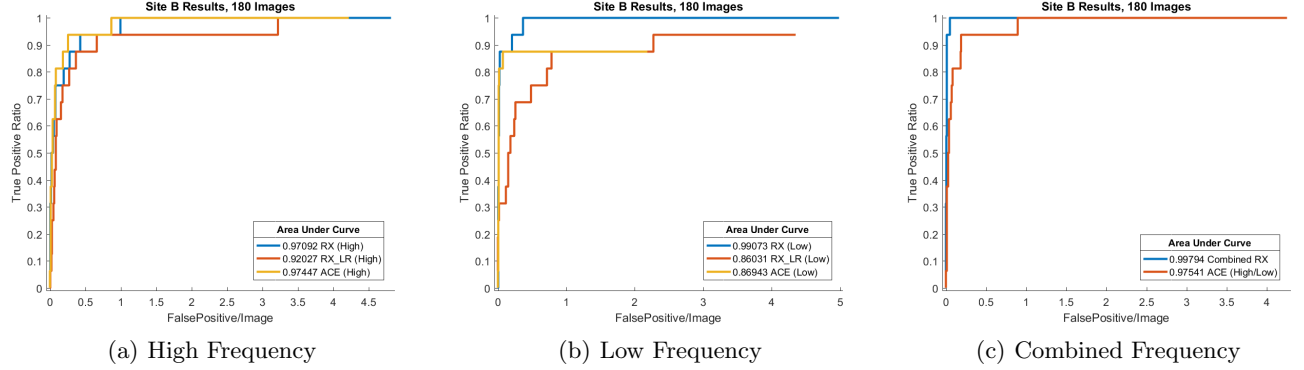


Figure 7. Site B ROC Curves

## 4. DISCUSSION

According to the ROC curves of figures 6 and 7, the unsupervised RX methods perform better as region of interest detectors. This could largely be due to the fact that they find any type of anomaly within the image whereas the supervised ACE method performs detection using learned signatures of targets. The supervised method is not only looking for anomalies, but essentially rejects those that do not look like the learned target signatures. A pitfall of these learned signatures is that they are not robust to shifts or rotations.

### 4.1 Site A

The site conditions also contribute to the detection performance of the methods. Site A is characterized by its varied and uneven seafloor texture that is scattered with non-target objects. These non-target objects produce highlight-shadow patterns that skew results. Transitions in the seafloor texture also generate responses from the detectors. Most of the targets in this site are also linearly shaped with some slight tilts to the left and right. Site A contains 16 targets.

For Site A, the RX\_LR algorithm has the best performance from both high and low frequency detection, finding all of the targets with the area under the curve (AUC) being 97% and 96%, respectively. It tends to perform better when the shadow is stronger since the right kernel is dedicated to that part of the highlight-shadow pattern. In the high frequency images, the shadow is more prominent. The AUC for the RX algorithm is 75% and 91% for high and low frequency detection, respectively. The drop in performance for the standard RX could be due to the site conditions and detector window. The site's background (seafloor characteristics) varies more so the detector's response to non-target objects or seafloor changes affects the performance. However, in the low frequency images, the image quality is reduced a bit, effectively smoothing textures and/or removing variance in the seafloor. This explains the boost in performance for the low frequency images.

The AUC for the ACE algorithm is 87% and 50% for high and low frequency, respectively. ACE benefits from prominent highlight-shadow patterns since the learned target signatures are able to characterize them. This is why ACE performs better on the high frequency data rather than the low frequency data where the shadows are lessened. ACE also learns a variety of target signatures so it can recognize more of them (figure 4). The low frequency targets are of less quality than the high which is why they are detected at a lower rate.

### 4.2 Site B

The performance of the methods on Site B data benefits from the relatively uniform seafloor. There are not many different seafloor textures or transitions. Additionally, there aren't many non-target objects in the data which reduces the number of objects mistaken as a target. It also prevents the learned targets of the supervised methods from being skewed. This makes detection of Site B less difficult than Site A. Site B contains 22 targets which vary in shape and size.

RX\_LR and RX methods again performed well on Site B imagery detecting all targets. RX\_LR had AUC of 91% and 84% for high and low frequency, respectively. The drop in performance for low frequency can be

attributed to the fact that the shadows are less prominent in the low frequency images. RX had AUCs of 96% for both high and low frequency imagery. This can be due to the uniformity of the site conditions.

The AUC for ACE is 97% and 86% for high and low frequency, respectively. ACE was able to detect all of the targets only for the high frequency imagery. Again, high frequency ACE learns a variety of target signatures that vary in shape and size so the detection of targets is more robust.

## 5. CONCLUSION

From the discussion, it is reasonable to conclude that the combined RX detector serves as the best prescreener for target detection in SAS imagery. However, many tweaks and improvements can be done to make it, and even the supervised method, more robust. One thing to consider for improving these methods is making the detection radius bigger. For these experiments, a peak in the confidence map was marked as a target after non-maximum suppression only if it fell within a 1 meter radius of the ground truth. Making this radius larger should increase the number of true positives since ground truth coordinates are typically marked on one pixel of the target in the image and does not cover the entire target. This makes it possible for a target to be missed simply because the ground truth coordinate was out of range of the radius although a peak was found on the target itself. Region growing could possibly be used after non-maximum suppression to make a potential target more prominent rather than relying on one pixel to represent a potential target.

It may also be beneficial to account for the image boundaries to increase the detection area of the image to the full image. In these experiments the detection area is reduced due to the constraint of the detector windows. Only pixels that allowed the entire detection window to fit in the image were considered. Because of this, some ground truth targets were omitted from training or not able to be detected in testing since they were either partially occluded or not visible within the detection area. Accounting for the boundary would allow all pixels to receive a detector value.

In conclusion, there are many factors to be considered when attempting to perform target detection on SAS imagery. Domain knowledge plays an important part in knowing what factor is or isn't useful. Target shapes, sizes and seafloor characteristics give vital information that aids in the development of detection algorithms.

## REFERENCES

- [1] Eismann, M. T., [*Hyperspectral remote sensing*], SPIE (2012).
- [2] Zare, A., Jiao, C., and Glenn, T., "Discriminative multiple instance hyperspectral target characterization," *IEEE Trans. Pattern Anal. Mach. Inteli.* (To Appear).
- [3] Kim, B. and Yu, S. C., "Imaging sonar based real-time underwater object detection utilizing adaboost method," in [*2017 IEEE Underwater Technology (UT)*], 1–5 (Feb 2017).
- [4] Ho, K. C., Harris, S., Zare, A., and Cook, M., "Anomaly detection of subsurface objects using handheld ground-penetrating radar," in [*Proc. SPIE 9454, Detection and Sensing of Mines, Explosive Objects, and Obscured Targets XX*], **9454** (May 2015).
- [5] Kk, S. and Yksel, S. E., "Comparison of rx-based anomaly detectors on synthetic and real hyperspectral data," in [*2015 7th Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS)*], 1–4 (June 2015).
- [6] Williams, D. P. and Groen, J., "A fast physics-based, environmentally adaptive underwater object detection algorithm," in [*OCEANS 2011 IEEE - Spain*], 1–7 (June 2011).
- [7] Zhao, Z., Zhao, A., Hui, J., Hou, B., Sotudeh, R., and Niu, F., "A frequency-domain adaptive matched filter for active sonar detection," *Sensors* **17**(7) (2017).
- [8] Reed, S., Petillot, Y., and Bell, J., "An automatic approach to the detection and extraction of mine features in sidescan sonar," *IEEE Journal of Oceanic Engineering* **28**, 90–105 (Jan 2003).
- [9] Kwon, H. and Nasrabadi, N. M., "Kernel rx-algorithm: a nonlinear anomaly detector for hyperspectral imagery," *IEEE Transactions on Geoscience and Remote Sensing* **43**, 388–397 (Feb 2005).
- [10] Galusha, A., Galusha, G., Keller, J. M., and Zare, A., "A fast target detection algorithm for underwater synthetic aperture sonar imagery," (2018).