ENHANCED DETECTION AND CLASSIFICATION OF MINE-LIKE OBJECTS USING SITUATIONAL AWARENESS AND DEEP LEARNING

Philip Chapple[a], Timothy Dell[a] and Daniel Bongiorno[b]

[a]DST Eveleigh, Locked Bag 7005, Liverpool NSW 1871, Australia
[b]DST Canberra, PO Box 7931, Canberra BC 2610, Australia

Contact author: Dr Phil Chapple, philip.chapple@dst.defence.gov.au,
fax +61 2 9381 0030.

Abstract: *An unsupervised approach to Automatic Target Recognition (ATR) has proven effective for detecting mine-like objects from high-resolution sidescan and synthetic aperture sonar images collected from autonomous underwater vehicles. The software uses statistics of local sections of an image to identify highlight and shadow outliers with dimensions comparable to mine-like objects. This software operates with high efficiency and acceptably low false alarm rates in the survey areas that we have most frequently encountered, with sandy seabeds and few natural features of sizes comparable to mines. In areas where there are numerous rocks or other debris on the seabed, the number of false detections is significantly higher. In this work, we enhance our existing software by incorporating knowledge of typical features in the locality, to reduce the number of false detections. Moreover, Deep Learning techniques are investigated to improve the classification of detected objects, scoring images according to their resemblance to images of mine-like objects. In this way, the overall ATR performance is improved.*

Keywords: *automatic target recognition, ATR, CAD/CAC, sonar image processing*

1. INTRODUCTION

Automatic target recognition (ATR) for processing high-frequency sonar imagery to detect mine-like objects (MLOs) on the seabed has been investigated by many researchers in the last two decades, [1]-[4]. Both supervised and unsupervised algorithms have been investigated and demonstrated. Supervised techniques involve the training of the detector based on available seabed imagery and known mine-like objects. These techniques have the advantage of being able to optimise detection performance for the conditions in which the training dataset was collected. Their main disadvantage is their reliance on large volumes of ground-truth training data, including many images of MLOs, which are not easy to acquire and label. Unsupervised techniques [5] involve an algorithmic approach to ATR, without such a strong reliance on training data.

Autonomous underwater vehicles (AUVs) operating within several metres of the seabed consistently collect high resolution sonar imagery that is suitable for ATR. Targets of interest in sidescan sonar and synthetic aperture sonar (SAS) imagery are often characterised by a bright highlight and an adjacent shadow, creating opportunities for an algorithmic approach. Complications arise because the appearance of an object will generally depend strongly on its aspect with respect to the sonar, and on how far the object is from the nadir (beneath the sonar), presenting various challenges for the image processing. If reliable ATR capability can be achieved, it provides not only the capability for rapid post-processing of sonar data, but also enhanced autonomy for AUVs which are able to conduct ATR processing of data in real-time during a mission.

While ATR processing can be effective for flat sandy areas where there are few rocks and other objects on the seabed, it becomes more challenging in highly cluttered areas – particularly when there are numerous objects of similar dimensions to mine-like objects. A human observer can recognise a rock field and appropriately discount many of the objects as being typical of the surrounding environment. It is more difficult to program a machine to do so, particularly if the clutter areas are not pervasive but occupy only patches of the imagery, as is often the case.

Defence Science & Technology (DST) Group has developed SonarDetect software [6] for visualising and ATR processing sonar imagery, using an unsupervised detection approach. This software has been successful in detecting mine-like objects with acceptably low false alarm rates in the areas where most of our trials have taken place. This paper describes the main features of the software, including visualisation features, ATR performance and analysis of results. The paper also describes recent work to improve ATR detection by scoring a detected object according to the occurrence of nearby similar objects. Finally, initial investigations with Deep Learning are reported to assist in the classification of detected objects.

2. SONAR DATA COLLECTION AND PROCESSING

DST Group has operated AUVs for approximately ten years, including a REMUS 100 operating the Marine Sonic Technology (MST) sonar (900 kHz and 1800 kHz) and a Gavia also operating an MST sonar (600 kHz and 1200 kHz). SonarDetect software is able to process MST files from these sonars, which contain sonar imagery and also vehicle state information. DST Group also possesses a REMUS 600 vehicle operating a Kraken AquaPix Interferometric Synthetic Aperture Sonar (InSAS). With in-house software, the data can be converted to image files together with time and position data, which SonarDetect can also process. Through data exchanges with local and international colleagues we have also acquired Klein data as SDF files, from several sidescan sonars including the Klein 3500. SonarDetect software can also read and process this data.

An important feature of the software is its ability to display sonar imagery alongside the chart, as shown in Fig. 1. The sonar image on the left-hand side can be easily related to the geographical location of the swath, shown in the chart image on the right-hand side (the chart is not evident here because the view is zoomed in to a small area). The masked regions of the sonar image are regions where the image is distorted, such as close to the nadir and where there are vehicle turns. These areas are not suitable for ATR processing and are excluded in order to minimise false alarms. In the sonar image the white rectangle indicates an object which has been detected by the ATR. It is possible to zoom in to inspect the object and to use measuring tools to measure its characteristics, such as its length, width and the height indicated by its shadow. The user has chosen to mark this object and label it 'MLO'. This label appears in both the sonar image and the chart image. All detections and object markings are saved in the database for the mission, which is a comma-separated variable (CSV) file, so that this information is read the next time the mission is loaded into the software.

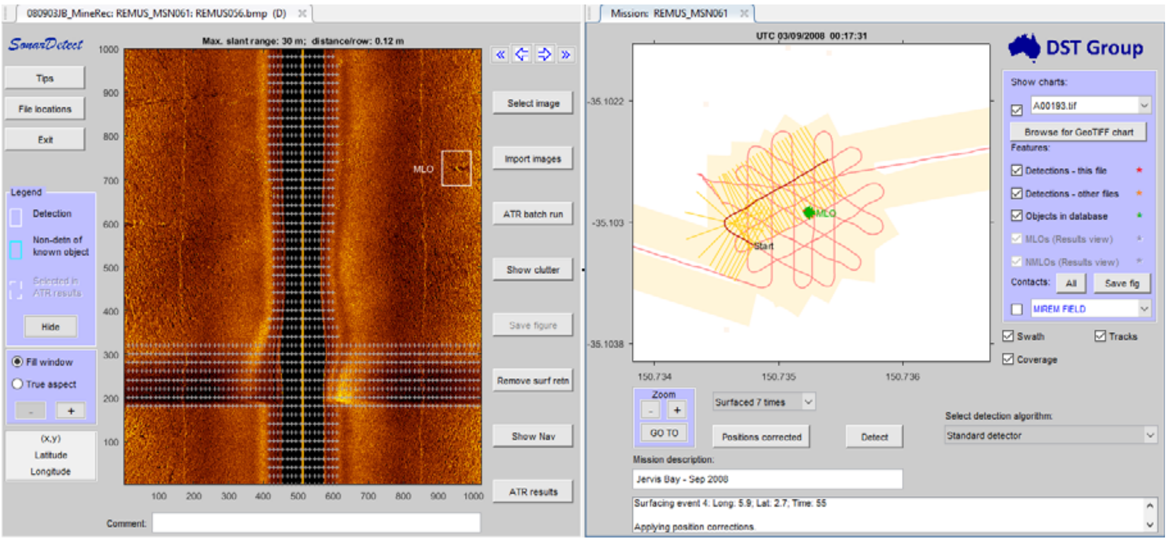


Fig. 1: SonarDetect display of a sidescan sonar image, showing a detected object, together with the chart view (zoomed in to display the current swath). In the sonar image, the nadir region (extending vertically in the middle) and a vehicle turn are masked out.

The main steps for processing a mission of sonar data files are:

1. Import the raw data files from MST or SDF files, which are converted and saved as image files and text files with position and time information.
2. Automatically create a CSV database file for the mission.
3. Conduct an ATR batch run to process all the files in the mission, using the standard detector or another detection algorithm selected by the user. The detections are automatically saved to the database.
4. If desired, examine all the images and mark any objects of interest, whether they have been detected or not (as for the marked object labelled 'MLO' in Fig. 1).
5. Analyse the ATR detections and collate results. In this step, geographical sites where detections have occurred and the sites of any marked objects can be designated as MLO, NMLO (non-mine-like object) or false alarm. The software lists for each site

all the image files where an ATR detection has occurred and also the files where the site has been missed, enabling the user to examine this location in each image, if desired.

6. Generate a report on the ATR detection results.

The ATR process for each image also consists of a number of steps:

1. Remove or mitigate the surface return in each image [7]. The surface return (evident in the sonar images in Figs 1 and 2) is a broken and bifurcated line, on both the port and starboard sides of the image, which can cause false alarms.
2. Pre-process the image for detecting highlights. This process includes setting masked regions to a constant value and normalising each image, so that the average image brightness for each angular displacement from the nadir is constant [8].
3. Conduct spatial filtering of the image to detect shadows, blurring features which are smaller than the target shadows of interest.
4. Divide the image into local areas and determine significant highlight and shadow pixels in each area by binning pixels according to their brightness.
5. Join connected groups of highlight and shadow pixels. Clusters of pixels satisfying certain size and shape requirements are retained.
6. Paired highlight and shadow clusters of pixels which are sufficiently close to one another and in a suitable geometrical relationship are counted as valid detections.

This general algorithm is governed by many algorithm parameters, which can be set and tested within the research version of the software. Different user-selectable algorithms are available for different situations, but there is a standard algorithm which is suitable for most of the sonar data tested.

Optimising the performance of the algorithm has involved maximising the detection probability for MLOs while constraining the probability of false alarms. Settings that work well for one kind of mine and seabed type can fail under different conditions.

Fig. 2 shows the ATR results analysis display, where the user can see a listing of all the files corresponding to a particular geographical site (the location of a detected object or user-marked object), and view each of those images to understand the reasons than an object has been detected or missed. In some cases, an object is not detected because it is outside the reliably observable range, i.e. in the nadir region (where the highlight may be obscure and the shadow minimal) or near the edge of the image (where the shadow area may extend beyond the edge of the image). Detections and non-detections in these areas are not counted in the reported statistics.

The user can mark each site as MLO, NMLO or false alarm. In the present context, an NMLO is a significant object which is atypical of the surrounding environment, or which an ATR detector could reasonably be expected to detect, that has been assessed as being non-mine-like. A false alarm is a detection of a feature typical of the surrounding environment (including fish) which preferably should not be detected. For this particular mission of 210 images, the analysis indicates that of 7 MLOs on the seabed, all were detected. There were 44 possible sightings of the MLOs in the observable range (most MLOs on the seabed were imaged multiple times), and they were detected 34 times, indicating a nominal probability of detection (Pd) of 77%. There were also 4 detections of NMLOs. There were 12 false alarms – a rate (Pfa) of 0.06 per image.

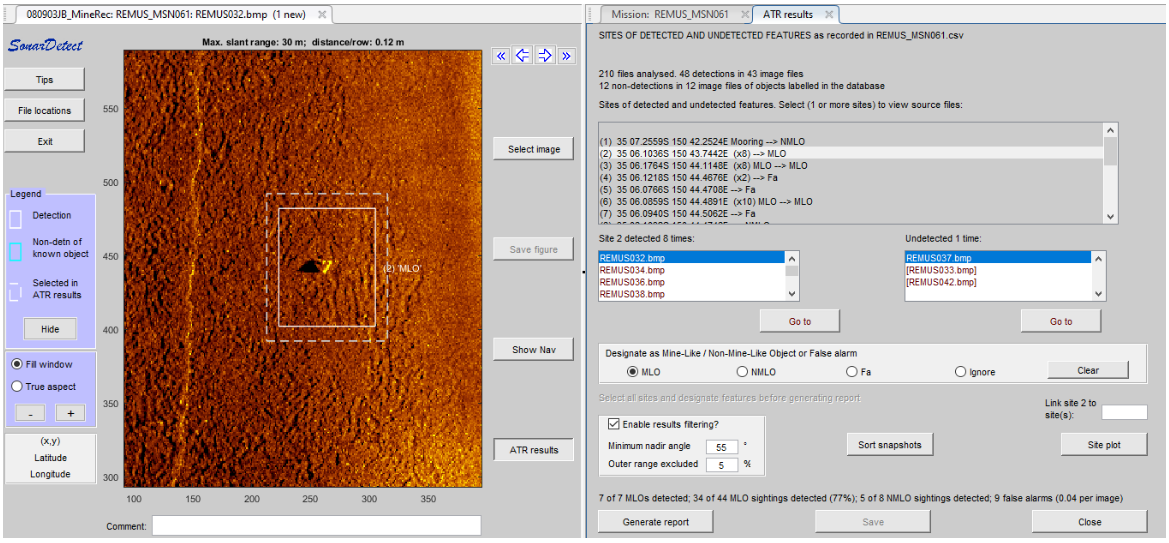


Fig. 2: ATR results analysis showing an object detected at site 2.

ATR processing of InSAS imagery is similar to processing of conventional sidescan data, although there are differences:

1. The images are larger and at higher resolution than is required for ATR.
2. The SAS data contains considerably more speckle, fragmenting highlights and shadows.
3. The SAS images are projected in horizontal displacement on the seabed rather than slant range, and the nadir region is not necessarily in the middle of the image.

Modifications to the software were necessary to allow for these differences and increase processing speed.

Fig. 3 shows a mine-like object detected in imagery from the AquaPix InSAS sonar. For this particular mission, there were 14 images provided, covering 3 MLOs on the seabed. Of the 8 possible MLO sightings, all were detected by the ATR. There was one false alarm.

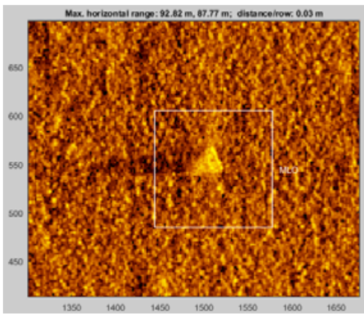


Fig. 3: Detected object from SAS data.

3. OPTIMISING ATR IN CLUTTERED AREAS USING SITUATIONAL AWARENESS

The above methods work well for prominent mine-like objects in areas where the seabed is benign, such as flat sandy seabeds in the absence of rocks and strong sand ripples/waves. When there is clutter or the objects are partially buried, the performance is poorer. Humans processing such imagery easily recognise certain areas as being rocky or covered by sand waves. Features in these areas are more likely to belong to the background, unless they stand out in some way. It is more difficult to program a computer to make this distinction reliably. Attempts to modify the ATR algorithm to make MLOs more detectable in some areas cause unacceptably high false alarm rates in cluttered areas.

In the literature, several methods have been investigated to allow the software to recognise environmental clutter, such as modelling the background as a Markov random field [9]. In the case of sand ripples or waves, some researchers have used Fourier analysis to recognise and mitigate the ripples [10]. In DST Group's research, we found that this approach works better in some situations than in others, depending on the spatial frequency, straightness and uniformity of the sand ripples or waves.

Recent improvements in DST's ATR have been achieved by scoring detected objects according to the number of similar or larger objects around them, in a local neighbourhood in the image. Firstly, the highlight of a detected object is scored by sorting nearby highlights, according to their number of pixels. The object highlight will have a score of 1 if it is the largest highlight, ½ if it is the second largest and so on. Secondly, the shadows are scored by a comparison with nearby shadows. Then we combine these two scores to derive an overall score for the object. By rejecting detections with scores below a certain threshold, insignificant objects are removed.

Because this approach has been successful in mitigating false alarms in cluttered areas, it has been possible to make the detector more sensitive. We can now detect MLOs with smaller highlights, without paying too high a price in terms of false alarms.

The increased sensitivity of the new standard detector is illustrated in Table 1, for three missions (two in Australia and one in the UK) including cluttered areas.

Mission #	*# views of MLOs*	*Algorithm*	*Pd*	*Pfa*
1 210 images	44	Previous detector	0.57	0.08
		Standard detector	0.77	0.06
2 80 images	32	Previous detector	0.78	0.00
		Standard detector	0.88	0.04
3 60 images	7	Previous detector	0.57	0.03
		Standard detector	0.71	0.08

Table 1: Comparison between the standard detector (incorporating situational awareness scoring) and the previous detector. Again, Pd is the proportion of MLOs in the observable range detected and Pfa is the number of false alarms per image.

4. AUTOMATION AND CLASSIFICATION OF DETECTED OBJECTS

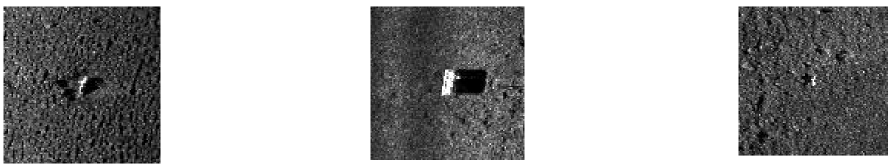
One goal in the development of ATR technology is to automate AUV operations so that underwater vehicles can detect and respond to seabed features of interest without human intervention. DST Group has run the ATR software on its REMUS 100 vehicle, and has demonstrated autonomous responses to targets detected in real-time.

For autonomous operations to work reliably, it is necessary to ensure that the false alarm rate is very low. This may be achieved by classifying snapshot images generated during the ATR process according to their resemblance to mine-like objects.

Recently we began investigating the application of Deep Learning [11] for classifying automatically generated snapshot images of the detected objects. This approach uses Google's Inception multi-layered convolutional neural network (CNN) developed by Szegedy *et al.* [12]-[13]. The network was trained on the ImageNet database [14] of more than 10 million labelled images from more than 10,000 image classes, including plants, cars and buildings. Most of the 27 layers of the network apply to many kinds of images, accounting for corners, edges and similar features. It was necessary only to retrain the last layer for the ensemble of sonar images. We used the TensorFlow open-source software library to do so, following the methodology described by Warden [15]. Using this method, the retraining of the CNN takes less than an hour on a desktop computer (on a CPU), rather than many weeks, which would have been required to train the whole network.

Hamilton and Cleary [16] used an unsupervised clustering method to separate an ensemble of 256 x 256 pixel sonar images of the seabed into different classes, for automated seabed classification. There were 13 different classes with 50 to 100 images in each class. The class labels were MLO, sands, flats, sea grass, rock, outcrops, drags, OSO (other significant object), pocks, blacks, mega, ripples and roughs. We used this set of images to set to train the CNN to classify a larger set of approximately 1,000 images per class. Following inspection of these classes, we deleted mislabelled images or transferred them to their correct classes. We then used the larger set of images to retrain the CNN.

Next, we applied the trained CNN to classify the 101 x 101 pixel snapshot images generated by SonarDetect during ATR processing, for the images from Mission 1 of Table 1. Because there were many classes and the CNN was not trained on these snapshot images, the classifier did not always label the human-labelled MLOs as 'MLO', but in 74% of cases, 'MLO' was one of the top 3 labels. For false alarms from this mission, there were none with 'MLO' in the top 3 labels (Fig. 4). This method provides a suitable way of classifying detections in real time processing. There is scope for improving this process considerably by retraining the CNN with labelled snapshot images and with fewer classes.



MLO 0.51; OSO 0.30; Flat 0.12

OSO 0.55; MLO 0.42; Outcrop 0.02

OSO 0.28; Sand 0.28; Drag 0.17

Fig. 4: CNN classification results for two MLOs and a false alarm.

5. CONCLUSION

We used an unsupervised approach to automatic target recognition to process high-resolution sonar images from AUVs, with acceptably high detection rates and low levels of false alarms. The software has proved effective for processing sonar imagery from many different AUV missions. To improve the results in cluttered areas, we developed a technique for scoring detections according to similar-sized or larger nearby objects, excluding objects typical of the background. This technique allowed us to improve the sensitivity of the ATR without a large increase in false alarms. Finally, we demonstrated classification of ATR snapshot images with a Deep Learning classifier, to recognise detected objects with a high probability of being mine-like objects.

6. ACKNOWLEDGEMENTS

We wish to acknowledge the work of the team at DST Group's Unmanned Systems & Autonomy Group in conducting numerous trials to collect sidescan sonar and SAS data. We also acknowledge the work of Mr Les Hamilton and Ms Jane Cleary in preparing the ensemble of classified sonar images used for training the CNN classifier.

REFERENCES

[1] **G.J. Dobeck, J.C. Hyland, and L. Smedley**, Automated detection/classification of sea mines in sonar imagery, *Proc. SPIE* 3079, 90-110, 1997.

[2] **J.A. Fawcett**, Image-based classification of sidescan sonar detections, *Proc. CAD/CAC 2001*, Halifax, Canada, 2001.

[3] **E. Coiras, P.-Y. Mignotte, Y. Petillot, J. Bell and K. Lebart**, Supervised target detection and classification by training on augmented reality data, *IET Radar, Sonar & Navigation* 1, 83-90, 2007.

[4] **P. Chapple**, Automated detection and classification in high-resolution sonar imagery for autonomous underwater vehicle operations, *DSTO-GD-0537*, 2008; http://dspace.dsto.defence.gov.au/dspace/handle/1947/9897 .

[5] **P.B. Chapple**, Unsupervised detection of mine-like objects in seabed imagery from autonomous underwater vehicles, *Proc. OCEANS'09 MTS/IEEE Biloxi*, 2009.

[6] **P. Chapple**, SonarDetect User Manual, in preparation.

[7] **P. Chapple, A. Gong, W. Zhou and S. Anstee**, Detection of mine-like objects and change detection using sonar imagery from autonomous underwater vehicles, *Proc. 2012 International Conf. on Detection and Classification of Underwater Targets*, Cambridge Scholars Publishing, 2014.

[8] **J.A. Fawcett**, Automatic Target Recognition Methods for Sidescan Sonar Images: The Advances and the Challenges, *Proc. 2012 International Conf. on Detection and Classification of Underwater Targets*, Cambridge Scholars Publishing, 2014.

[9] **A.T. Çelebi, M.K. Güllü and S Ertürk**, Mine detection in side scan sonar images using Markov Random Fields with brightness compensation, *IEEE Conf. Sig. Proc. Commun. Applications*, Antalya, Turkey, 2011.

[10] **R. Quintal, P. Dysart and R. Greene**, Automated Side-scan Data Analysis, *Hydro International* 11 (No. 9), 22-25, October 2007.

[11] **I. Goodfellow, Y. Bengio and A. Courville**, *Deep Learning*, MIT Press, 2016.

[12] **C. Szegedy *et al.***, Going Deeper with Convolutions, *IEEE Conf. Computer Vision and Pattern Recognition*, 1-9, 2015.

[13] **C. Szegedy *et al.***, Rethinking the Inception Architecture for Computer Vision, Cornell Uni Library, USA, https://arxiv.org/abs/1512.00567 .

[14] Stanford Vision Lab, *ImageNet*, Stanford University, USA, www.image-net.org .

[15] **P. Warden**, https://codelabs.developers.google.com/codelabs/tensorflow-for-poets

[16] **L.J. Hamilton and J. Cleary**, Autonomous Processing of Sidescan Sonar Imagery for Unmanned Underwater Vehicles, *Acoustics 2015 Hunter Valley*, Acoustics Australia, 2015.