2014 IEEE Congress on Evolutionary Computation (CEC)
July 6-11, 2014, Beijing, China

Automatic Target Recognition Using Multiple-Aspect Sonar Images

Xiaoguang Wang, Xuan Liu and Nathalie Japkowicz
School of Electrical Engineering & Computer Science,
University of Ottawa, Ottawa, Canada
{bwang009, nat}@ eecs.uottawa.ca,
xliu107@uottawa.ca

Stan Matwin
Faculty of Computer Science,
Dalhousie University, Halifax,
Canada
Stan@cs.dal.ca

Bao Nguyen
Defence R&D Canada Centre for
Operational Research and Analysis,
Ottawa, Canada
Bao.Nguyen@drdc-rddc.gc.ca

***Abstract*—Automatic Target Recognition (ATR) methods have been successfully applied to detect possible objects or regions of interest in sonar imagery. It is anticipated that the additional information obtained from additional views of an object should improve the classification performance over single-aspect classification. In this paper the detection of mine-like objects (MLO) on the seabed from multiple side-scan sonar views is considered. We transform the multiple-aspect classification problem into a multiple-instance learning problem and present a framework based upon the concepts of multiple-instance classifiers. Moreover, we present another framework based upon the Dempster–Shafer (DS) concept of fusion from single-view classifiers. Our experimental results indicate that both the presented frameworks can be successfully used in mine-like object classification.**

Keywords—Automatic Target Recognition; multiple-instance learning

I. INTRODUCTION

To acquire high-resolution sonar imagery for the detection of mine-like objects (MLO) and other objects of interest on the seabed, side-scan sonar equipped vehicles such as Autonomous Underwater Vehicles (AUVs) are frequently used by military forces or commercial organizations. For this purpose, Automatic Target Recognition (ATR) methods have been successfully applied to detect possible objects or regions of interest in sonar imagery [1]-[12]. Since many of the sonar images of the same object are taken from different sonar passes, there are multiple views of the same object at different ranges and aspects of the sonar. It is anticipated that the additional information obtained from additional views of an object should improve the classification performance over single-aspect classification. Recent research [4] [7] [8] proved this hypothesis by experimental results and found that it was possible to obtain more accurate classifications if the detection was based upon multiple views of the object instead of a single view of the object. Since each view represents a different aspect of the object, here the multiple-aspect classification is also regarded as the multiple views classification.

In this paper, we present two frameworks for multiple-aspect classification on side scan sonar images. The first one uses multiple-instance learning (MIL) method [5] [16] for a combination of the information of multiple views of a target. As far as we know we are the first to apply MIL method on multiple-aspect classification of MLOs. The second one uses the Dempster-Shafer (DS) theory [21] on multiple views of a target [4] [7] [8]. Although there are some applications of DS theory [21] on fusing single views of a target for MLOs classification, as far as we know we are the first to extend this method to multiple-instance classifications.

The remainder of the paper is organized as follows: Section II discusses previous work that has been done in multiple-aspect based classification. Section III presents the multiple-instance learning [16] methodologies and the fusion methodology used for solving the multiple aspects classification problem. Section IV presents the data preprocessing method used in our research. Section V illustrates the efficiency of our algorithm as determined by experimentation, and offers some final remarks. Finally, Section VI presents the conclusion, followed by the references.

II. RELATED WORKS

In previous works, B. Zerr et al. [1] [6] were the first to describe a method to estimate the three-dimensional aspects of underwater objects using a sequence of sonar images. The sonar images are segmented into three kinds of regions: echo, shadow and background. A study [4] conducted using sonar images of various objects and height profiles as features showed that the highest classification performance when taking an object's image twice can be achieved with an angular increment of 90 degrees between the two images. M. Couillard et al. [7] extended this study and considered the classification performance of all admissible secondary aspects for different types of mine-like objects by studying the correct classification rates and the error rates as a function of the angular difference between aspects. In their work, two different approaches were used to combine multiple images of an object. The first approach creates a new object for classification by combining the features of the two images into a single vector. The second approach is simply to fuse the single aspect classification probabilities obtained from the classifier according to the desired angular increment between the images.

J. Fawcett et al. [8] investigated two approaches for fusing multiple views: fuse-feature and fuse-classification. In the first approach the two feature sets taken at different aspects were combined to form a large feature vector. Then a kernel based classifier was trained with this feature vector. In the second approach, two individual-aspect classifications of two feature vectors were fused using the Dempster-Shafer (DS) theory

[21], which has frequently been used as an alternative to Bayesian theory and fuzzy logic for data fusion.

S. Reed et al. [9] [10] have also investigated the classification of a target by fusing several views using DS theory [21]. They present a model to extend the standard mine/not-mine classification procedure to provide both shape and size information on the object. The difference between their work and the work of others is that they generated the mass functions using a fuzzy functions membership algorithm based on fuzzy logic.

V. Myers and D. P. Williams [11] [12] introduced a model for classifying targets in sonar images from multiple views by using a partially observable Markov decision process (POMDP). This POMDP model allows one to adaptively determine which additional views of an object would be most beneficial in reducing the classification uncertainty.

In other related work, G. Dobeck fused multiple images from different frequency bands [13], J. Tucker et al. [14] fused multiple images from different platforms, and M. Azimi-Sadjadi et al. [15] fused multiple images from multiple-aspect target echo classification.

III. FRAMEWORKS FOR MULTIPLE-ASPECT CLASSIFICATION IN MLOS CLASSIFICATION

In sonar images taken by AUVs, we can get multiple views of one object based on its position. The number of sonar images of an object may vary depending on the conditions of AUVs. Fig.1. gives an example of multiple views of one object. These sonar images are taken from different AUVs or the same AUV from different angles or distances. In training datasets of MLO classification, each object has more than one view and each view is saved as an instance in the dataset. Therefore each object has a group of instances which has the same label.

In our research, the MLOs classification on multiple views of an object is defined as a binary learning problem.

Input: Given a set of instance groups $\chi_i, i = 1, \dots, N$, where each group represents an object which may or may not be an MLO. One group consists of an arbitrary number of instances containing information about the sonar images of multiple views of this object:

$$\chi_i = \{x_i^1, x_i^2, \dots, x_i^{n_i}; y_i\}, i = 1, \dots, N, y_i \in \{-1, +1\} \qquad (1)$$

where each instance $x_i^{n_i}$ is an M-tuple of attribute values belonging to a certain domain or instance space $\mathbb{R}$. There is a function c that classifies individual bags as $+1$ (MLO) or -1(non-MLO).

Output: a model that predicts whether or not a new example which consists of an arbitrary number of sonar images of multiple views of this object falls into the MLO category.

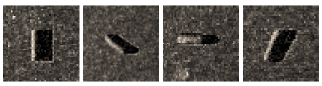

Fig. 1. Example of multiple views of one target in sonar imagery for the detection of mine-like objects (MLO).

In previous works [3]-[11] and [13]-[15] , the one common point is that fusion methods are used to combine different views for classification. Although using fusion methods such as Dempster-Shafer fusion of single aspect classification results was shown to be effective in some cases [4] [7] [8], we still anticipate a number of challenges and limitations in some ATR applications using fusion methods [18]. Thus it is necessary to develop other methods to combine different information from multiple views in this research. In this paper, two frameworks are provided for this kind of classification. In the first framework, we use a multiple-instance method for this purpose and in the second framework we choose Dempster-Shafer fusion method to combine single aspect classification results. Moreover, we will compare these two frameworks in our experiments.

A. *Multiple-instance methodologies*

Multiple-instance learning (MIL) [16] is a framework choice for multiple-aspect classification. MIL is concerned with supervised learning but differs from normal supervised learning in two points: (1) it has multiple instances in an example, and (2) only one class label is observable for all the instances in an example.

The standard multiple-instance learning problem [16] can be defined as:

Given: a set of bags B_i, $i = 1, \dots, N$, their label $c(B_i) \in \{0,1\}$, and the instances $e_{ij} (j = 1, \dots, n_i)$ belonging to each bag.

Output: the existence of an unknown function f that classifies individual instances as 1 or 0, and for which it holds that $c(B_i) = 1$ if and only if there exists $e_{ij} \in B_i : f(e_{ij}) = 1$ (multiple-instance constraint, MIC).

From this definition we can find that the standard assumption of multiple-instance learning problem [16] is based on the multiple-instance constraint. If we extend this assumption to a general assumption which has no multiple-instance constraint, the binary multiple views learning problem in (1) can be transformed into a multiple-instance problem. After this transformation, each bag in MIL represents an object in multiple views learning and each view of the object can be represented as an instance in this bag.

Here we choose two algorithms for multiple-instance learning [16]: the decision tree [20] and the logistic regression methods [22].

1) Tree method

Similar to a single-instance decision tree (like C4.5 [20]), the multiple-instance tree method is based on the information gain of a feature of the instance. The difference between the multiple-instance decision tree and the single-decision tree is that instead of using one instance to develop the information gain, the growing of a multiple-instance tree is based on the information gain of a set of instances. The concept of information gain and entropy are extended to bags of instances in the MIL framework. Suppose S is a collection of instances which belongs to $p(S)$ positive bags and $n(S)$ negative bags, F is the feature being considered as the splitting criterion and S_n

is the collection of instances whose value of feature F is n. The extended information gain and entropy are defined as:

$$Entropy_{multi}(S) = -\frac{p(S)}{p(S)+n(S)} \times log_2\left(\frac{p(S)}{p(S)+n(S)}\right) - \frac{n(S)}{p(S)+n(S)} \times log_2\left(\frac{n(S)}{p(S)+n(S)}\right) \quad (2)$$

$$InfoGain_{multi}(S,F) = Entropy_{multi}(S) - \sum_{n \in Values(F)} \frac{p(S_n)+n(S_n)}{p(S)+n(S)} \times Entropy_{multi}(S_n) \quad (3)$$

In this paper the multiple-instance tree method implements the top-down decision tree learning approach known from propositional tree inducers such as C4.5 [20], with two key modifications: (a) nodes are expanded in best-first order guided by a heuristic that aims to identify pure positive leaf nodes as quickly as possible, and (b) whenever a pure positive leaf node is created, all positive bags containing instances in this leaf node are deactivated.

2) Logistic Regression method

For single-instance classification, Logistic Regression [22] assumes a parametric form for the distribution $Pr(Y|X)$, then directly estimates its parameters from the training data. The parametric model assumed by Logistic Regression in the case where Y is a boolean is:

$$Pr(Y=1|X) = \frac{1}{1+exp(\omega_0+\sum_{i=1}^{n}\omega_i X_i)} \quad (4)$$

and

$$Pr(Y=0|X) = \frac{exp(\omega_0+\sum_{i=1}^{n}\omega_i X_i)}{1+exp(\omega_0+\sum_{i=1}^{n}\omega_i X_i)} \quad (5)$$

However, the standard logistic regression model [22] does not apply to multiple-instance data because the instances' class labels are masked by the "collective" class label of a bag. X. Xu and E. Frank [17] use a two-stage framework to upgrade linear logistic regression for multiple-instance data.

The instance-level class probabilities are given by $Pr(y=1|x) = \frac{1}{1+exp(-\beta x)}$ and $Pr(y=0|x) = \frac{1}{1+exp(\beta x)}$ respectively, where β is the parameter vector to be estimated.

Given a bag b with n instances $x_i \in b$, we assume that the bag-level class probability is either given by

$$Pr(Y|b) = \frac{1}{n}\sum_{i=1}^{n} Pr(y|\, x_i) \quad (6)$$

or by

$$log\frac{Pr(y=1|b)}{Pr(y=0|b)} = \frac{1}{n}\sum_{i=1}^{n} log\frac{Pr(y=1|\, x_i)}{Pr(y=0|\, x_i)} \quad (7)$$

From (7) we can get

$$Pr(y=1|b) = \frac{[\prod_i^n Pr(y=1|\, x_i)]^{\frac{1}{n}}}{[\prod_i^n Pr(y=1|\, x_i)]^{\frac{1}{n}} + [\prod_i^n Pr(y=0|\, x_i)]^{\frac{1}{n}}} = \frac{exp(\frac{1}{n}\beta\sum_i x_i)}{1+exp(\frac{1}{n}\beta\sum_i x_i)} \quad (8)$$

and

$$Pr(y=0|b) = \frac{[\prod_i^n Pr(y=0|\, x_i)]^{\frac{1}{n}}}{[\prod_i^n Pr(y=1|\, x_i)]^{\frac{1}{n}} + [\prod_i^n Pr(y=0|\, x_i)]^{\frac{1}{n}}} = \frac{1}{1+exp(\frac{1}{n}\beta\sum_i x_i)} \quad (9)$$

Based on (8) and (9) we can estimate the parameter vector β by maximizing the bag-level binomial log-likelihood function:

$$LL = \sum_{i=1}^{N}[y_i log Pr(y=1|b) + (1-y_i) log Pr(y=0|b)] \quad (10)$$

where N is the number of bags.

As usual, the maximization of the log-likelihood function is carried out via numeric optimization because there is no direct analytical solution. The optimization problem can be solved very efficiently because we are working with a linear model.

B. Fusion methodologies

Data fusion is a technology which collates information from different sources. It considers the same scene in an attempt to provide a more complete description. When we try to combine multiple-aspect sonar images for classification, the most common numerical fusion techniques used are Bayesian probability theory, Fuzzy systems and Dempster-Shafer (DS) theory [21].

Fuzzy systems contain many possible fusion operators while some of the operators are non-associative and the choice of operators may be case dependent. This means that the order in which the information is fused has an impact on the final result. Bayesian and DS models [21] have both been successfully applied but DS theory [21] provides some features that Bayesian theory does not. One of the most significant features of DS theory [21] is that it can consider the union of classes. This feature is used to improve the separability of different classes. The DS method [21] is a popular data fusion method which has been used by other authors for side scan sonar image classification [3]-[11], [13]-[15].

The DS method [21] is based on two ideas: obtaining degrees of belief for one question from subjective probabilities for a related question, and Dempster's rule [21] for combining such degrees of belief when they are based on independent items of evidence.

The Dempster's rule [21] of combination is a purely conjunctive operation (AND). The combination rule results in a belief function based on conjunctive pooled evidence and it can also be used for multiple-aspect classification.

In DS theory, each unique class makes up a set called the frame of discernment $\Theta = \{\omega_1, \omega_2, \ldots, \omega_M\}$. Belief is attributed to hypotheses within the power set through a basic probability assignment, called the mass function $m(A)$.

Suppose that we have two views of target S_1 and S_2 and the mass functions $m_1(S_1)$ and $m_2(S_2)$. Based on the Dempster's rule, the mass after fusion for the set A is:

$$m_{12}(A) = \frac{\sum_{S_1 \cap S_2 = A} m_1(S_1)m_2(S_2)}{1-\sum_{S_1 \cap S_2 = \emptyset} m_1(S_1)m_2(S_2)} \quad (11)$$

The classification rule for this case is:

$$g(x_1, x_2) = argmax_i \, m_{12}(\omega_i) \quad (12)$$

Like many authors, we use DS theory [21] as a choice for multiple-aspect classification in our research. The difference between our algorithm and other people's work ([3]-[11] and [13]-[15]) is that we use the same format of MIL dataset to represent the multiple views of objects, which means each object has a group of instances which has the same label, and we still call this group of instances a "bag". The group number has been represented as the "bag ID" in datasets which is the same as in the MIL. First we transform the multiple-instance training and testing dataset into a single-instance training dataset and a single-instance testing dataset. The bag IDs remain for each instance in this transformation and the label of each instance is set as the bag label. A single-aspect classifier is used on this training dataset. Then the predicted class labels from the testing data are saved. Using T cross validation we can get a $T \times M$ output matrix. Let $\beta_i(k), k = 1,2,\dots,T$, correspond to the ith column of the prediction vector for the kth testing feature vector. For n output vectors $\beta_i(k),\ i = 1,2,\dots n$ obtained from n single-aspect classifications which has the same bag ID, the n sets of masses, which are also n bags, are finally fused using Dempster's rule and the final decision is given by the classification rule $g(x_1, x_{2,\dots,}x_n)$.

IV. DATA PREPROCESSING

In mine countermeasure missions (MCM), sonar images collected by AUVs will convey important information about underwater conditions. Raw sonar images must be properly pre-processed before machine learning algorithms are applied. The processing of the sonar images will have a significant impact on the subsequent MLOs detection and classification stages.

In the MCMs, a large part of the sonar images collected by AUVs represents the background—seabed. In MLOs detection and classification, we are more interested in the object that lies on the seabed rather than the background. That is to say, we are only interested in the foreground objects. Therefore, the areas from the images with only background information can be simply discarded.

The first step of this classification task is the segmentation of the sonar images. Image segmentation is a widely used image processing technique to detect target objects and segment the original images into small pieces that contain the target objects.

In this step, our goal is to delete image data that contains only background information and reduce the total amount of data to be processed. Therefore whether or not the size, shape and location of the target object are accurately found is not a main concern in this step. The foreground objects are assumed to have a more complex texture than the seabed. Thus, the foreground object areas are obtained by using local range and standard deviation filters.

A. Image filtering

Before performing image segmentation, it is necessary to apply filters to remove noise from the original sonar image. The operation of image filtering is to calculate the convolution or correlation between the original image $I(x, y)$ and the filter $F(x, y)$. For grayscale images, the summation of all the elements in the filter should be equal to 1. The pixels of the image after filtering $I'(x, y)$ are a weighted sum of the original pixel values. The processes are expressed as follows:

$$I'(x, y) = \sum_{i=-m}^{m} \sum_{j=-n}^{n} I(x - i, y - j) F(i, j) \quad (13)$$

$$I'(x, y) = \sum_{i=-m}^{m} \sum_{j=-n}^{n} I(x + i, y + j) F(i, j) \quad (14)$$

$$\sum_{i=-m}^{m} \sum_{j=-n}^{n} F(i, j) = 1 \quad (15)$$

Here, m and n represent constants that define the size of the filter. Fig.2. shows the images before and after we apply the Gaussian filter. The target in the square box in this figure is the mine.

B. Image segmentation

In sonar image segmentation, edges are significant local changes of intensity in an image. Edge detection is often used in order to extract the pieces of MLO objects from the background of the seabed. Discrete differencing is a local edge enhancement technique. It is used to sharpen edge elements in an image by discrete differencing in either the vertical or horizontal direction, or in a combined fashion. In 2D image, edges are expressed using partial derivatives. Points which lie on an edge can be found by detecting the maximum of gradient.

The gradient is a vector which has a certain magnitude and direction. The gradient of a 2-D image can be expressed as:

$$\nabla I = \begin{pmatrix} \frac{\partial I}{\partial x} \\ \frac{\partial I}{\partial y} \end{pmatrix} \quad (16)$$

$$magn(\nabla I) = \sqrt{(\frac{\partial I}{\partial x})^2 + (\frac{\partial I}{\partial y})^2} = \sqrt{u^2 + v^2} \quad (17)$$

$$dir(\nabla I) = tan^{-1}(v/u) \quad (18)$$

The magnitude of a gradient is usually approximated by:

$$magn(\nabla I) \approx |u| + |v| \quad (19)$$

The gradient can be approximated by finite differences:

$$\frac{\partial I}{\partial x} = \frac{I(x+h_x,y)-I(x,y)}{h_x} = I(x + 1, y) - I(x, y), (h_x = 1) \quad (20)$$

$$\frac{\partial I}{\partial y} = \frac{I(x,y+h_y)-I(x,y)}{h_y} = I(x, y + 1) - I(x, y), (h_y = 1) \quad (21)$$

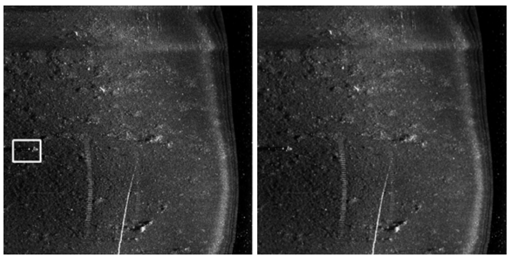

Fig. 2. Example of using Gaussian filter to reduce noise: left is the original image; right is the image after filtering.

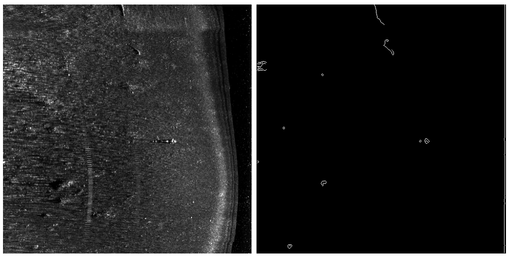

Fig. 3. Example of an edge detection result on an image provided by the Defence Research and Development Canada Atlantic.

Here we choose the Canny method [2] as the edge detection method. The Canny method will first smooth the image with a Gaussian filter. Then it will calculate the gradient of $I(x, y)$ according to (16). An edge pixel is defined to be a local maximum in the gradient direction and results in a ridge which will be reduced to the width of one pixel by the Canny method. Then two thresholds T1 and T2 (T1<T2) will be applied to deal with the ridge. The stronger threshold T2 will find a contour that contains more true edges, fewer false edges, but this contour is very likely to be discontinuous because of a high threshold. When the contour reaches its end, Canny method [2] will switch to the weaker threshold T1 to grow the contour with faint edges.

Fig. 3 illustrates the extraction of foreground objects from a sonar image which was provided by the Defence Research and Development Canada Atlantic. Areas that do not have a reasonable size will be ignored.

The objective of the image processing procedures at this point is data reduction rather than MLOs detection. Thus, a relatively high false alarm rate is acceptable.

The next step of this image processing task is the segmentation of the sonar images into three distinct regions: highlight or target echo (sound scattered by the target by active sonar), shadow (regions of low acoustic energy created by an object or seabed feature blocking the sound propagation) and background or seabed.

For object detection tasks, an object should be detected through a single view, no matter where and how it lies on the seabed. Therefore, the features used should not be sensitive to the location and orientation of the object. The grayscale histogram, a simple but informative statistical feature, is considered. In many image recognition systems, many complex features are used, but such features will inevitably increase the computational complexity, impeding the real time detection. The histogram is easy to calculate and the distribution of the grayscale value can be well described by this feature.

The size of each sonar image used in our study is 500×1024, so there are 512,000 pixel values in each raw image. In order to translate this image information into data (which can be processed by machine learning algorithms), we draw the histogram for the final image. As in the unit 8 image, the brightness of a pixel falls into the region 0-255. This region is divided into 16 bins. The number of pixels falling into each bin is counted and normalized by the total number of the pixels. These 16 numbers are taken as the features of the data. The grayscale histograms are normalized to the frequency that a pixel value falls into each bin.

V. Experimental Results Of Multiple-aspect Images Classification

In J. Fawcett et al. [8]'s work, objects representing targets (cylinder, manta, and rockan shapes) and non-targets (rocks) were deployed to test multiple-aspect classification. Differing from their method, we divided the MLO classification into two steps in our experiment. In the first step, we apply the classifier to make a classification between MLOs and non-MLOs. In the second step, the classifiers are used again on MLOs with different shapes.

In this section, we will demonstrate three experiments. The first experiment studies the classification performances as a function of the number of aspects using multiple views classification frameworks presented in this paper with different classifiers. The second experiment is studies classification performances on different shapes of MLOs using presented multiple views classification frameworks with different classifiers. The third experiment studies classification performances using all presented multiple views classification frameworks with different classifiers.

A. Experimental Setup

The sonar images used in our study are provided by the Defence Research and Development Canada Atlantic (DRDC-Atlantic, Dartmouth, NS, Canada).

TABLE I. Multiple views MLOs Dataset I

Datasets	#objects	#attribute	#positive objects	#negative objects
MLO_0	360	16	180	180

TABLE II. Multiple views MLOs Dataset II With Different Shapes

Datasets	#objects	#attribute	#cylinder	#manta	#wedding _cake
MLO_1	279	16	93	93	93

TABLE III. Multiple views MLOs Dataset III

Datasets	#attributes	#positive objects	#positive Instances	#negative objects	#negative instances
MLO_2	16	57	114	58	116
MLO_3	16	64	132	64	144
MLO_4	16	65	132	64	156
MLO_5	16	63	132	70	146

The binary datasets used in study one, two, and three are described in Table I, Table II and Table III respectively. The negative examples denote the non-MLOs and the positive examples denote the MLOs. In Table I, dataset MLO_0 has 360 objects; the number of MLOs (positive) and non-MLOs (negative) are both 180. Each object has three views so we can study the classification performances as a function of the number of views. In Table II, dataset MLO_1 has 279 objects. We have three different shapes of MLOs which are cylinder, manta and wedding_cake shapes and each shape has 93 objects. Dataset MLO_1 has binary classes and more than one "view" on an object. Table III lists four datasets and each object has more than one "view" which is represented as an instance; therefore, each object has a group of instances.

B. Experimental Results

1) Classification of multiple views of an object: in this experiment, we study the classification performances as a function of the number of aspects and compare the experimental results using DS [21] and multiple-instance learning [16] frameworks.

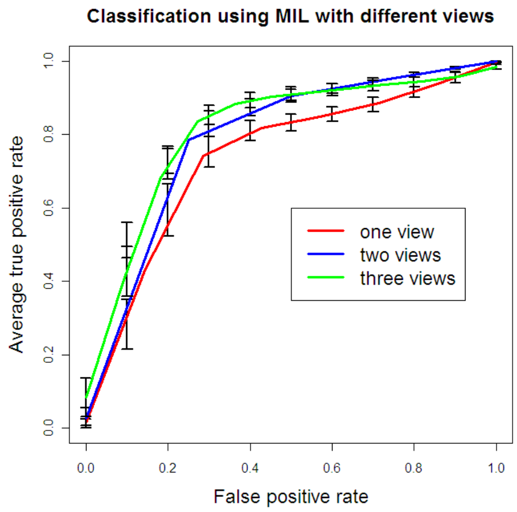

Fig. 4. Classification performances (ROC curves) as a function of the number of views using Multiple-instance decision tree as the classifier

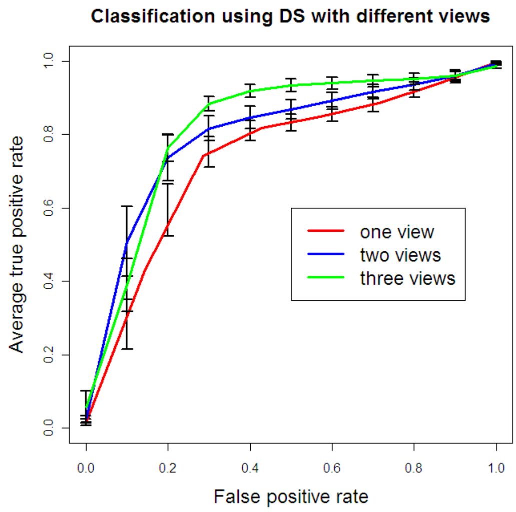

Fig. 5. Classification performances (ROC curves [1]) as a function of the number of aspects using DS on decision tree as the classifier

We choose MLO_0 in Table I as the training dataset for this experiment. ROC curves [1] are chosen as the measurement technique for the classification performance. The experimental results are shown in fig. 4 and fig. 5. Fig. 4 shows the ROC curves as a function of the number of aspects using Multiple-instance decision tree as the classifier and fig. 5 shows the ROC curves using DS [21] with decision tree as the classifier. We find that for both frameworks, the more views used for classification, the better the performance.

2) Classification of different shape of MLOs with multiple views: After making a classification of MLOs and non-MLOs, we can make a classification of the shape of the MLO. MLO_1 in Table II is chosen as the dataset for this experiment.

This study includes two parts. The first uses the tree method as a classification algorithm and the second uses Logistic Regression instead. For each method, we study single-aspect, multiple-instance as the multiple-aspect classification framework, and DS [21] as the multiple-aspect classification. For single-aspect classification, each object has one view. For multiple-aspect classification, each object has more than one view and we choose the number of views randomly between two and three.

We use the confusion matrix to display the performances of the classifiers. The definition is expressed in Table IV.

TABLE IV. CONFUSION MATRIX

	Predicted Positive	Predicted Negative
Actual Positive	TP (True Positive)	FN (False Negative)
Actual Negative	FP (False Positive)	TN (True Negative)

- Tree method

Table V to Table VII show the confusion matrices resulting from single-aspect classification using decision tree, multiple-aspect classification using the multiple-instance tree method and multiple-aspect classification using DS [21] with decision tree respectively.

TABLE V. THE CONFUSION MATRICES RESULTING FROM SINGLE-ASPECT CLASSIFICATION USING DECISION TREE

(a)	Single-Aspect		
	Cylinder	Manta	Wedding_Cake
Cylinder	58.9	24.9	16.2
Manta	11.9	73.5	14.6
Wedding_Cake	11.4	16.2	72.4

TABLE VI. THE CONFUSION MATRICES RESULTING FROM MULTIPLE-ASPECT CLASSIFICATION USING MULTIPLE-INSTANCE TREE METHOD

(b)	Multiple-Aspect		
	Cylinder	Manta	Wedding_Cake
Cylinder	80.6	11.8	7.5
Manta	8.6	83.3	8.1
Wedding_Cake	10.2	3.2	86.6

TABLE VII. THE CONFUSION MATRICES RESULTING FROM MULTIPLE-ASPECT CLASSIFICATION USING DS FUSION WITH DECISION TREE

(c)	Multiple-Aspect		
	Cylinder	Manta	Wedding_Cake
Cylinder	73.6	17.2	9.2
Manta	4.8	84.4	10.8
Wedding_Cake	8.6	7.0	84.4

- Logistic Regression method

Table VIII to Table X give the confusion matrices resulting from single-aspect classification using Logistic Regression, multiple-aspect classification using multiple-instance logistic regression and multiple-aspect classification using DS [21] with logistic regression respectively.

TABLE VIII. THE CONFUSION MATRICES RESULTING FROM SINGLE-ASPECT CLASSIFICATION USING LOGISITIC REGREESION

(a)	Single-Aspect		
	Cylinder	Manta	Wedding_Cake
Cylinder	58.9	22.7	18.4
Manta	30.8	55.1	14.1
Wedding_Cake	4.9	15.1	80.0

TABLE IX. THE CONFUSION MATRICES RESULTING FROM MULTIPLE-ASPECT CLASSIFICATION USING MULTIPLE-INSTANCE LOGISITIC REGREESION

(b)	Multiple-Aspect		
	Cylinder	Manta	Wedding_Cake
Cylinder	72.6	16.7	10.7
Manta	10.3	76.3	5.4
Wedding_Cake	8.1	3.8	88.1

TABLE X. THE CONFUSION MATRICES RESULTING FROM MULTIPLE-ASPECT CLASSIFICATION USING DS FUSION WITH LOGISITIC REGREESION

(c)	Multiple-Aspect		
	Cylinder	Manta	Wedding_Cake
Cylinder	69.4	19.9	10.7
Manta	16.7	75.3	8.0
Wedding_Cake	7.0	8.1	84.9

From these classification results we can see that the classification performance, both using the multiple-instance framework and data fusion framework, was improved by using more "views" in the classification.

3) Comparison of different multiple-aspect classification methods: In this experiment, all presented multiple-aspects classification methods are compared. They include multiple-instance tree method (MI_Tree), multiple-instance logistic regression method (MI_LR), Dempster-Shafer with decision tree method (DS_tree) and Dempster-Shafer with logistic regression method (DS_LR). We choose MLO_2 to MLO_5 in Table III as the training datasets for this experiment. Area under a Receiver Operating Characteristic (AUC) [1] is chosen as the measurement technique for the classification performance.

TABLE XI. THE COMPARISON MATRICES RESULTING FROM PRESENTED MULTIPLE-ASPECT CLASSIFICATION METHODS BY AUC

	MI_Tree	***MI_LR***	***DS_Tree***	***DS_LR***
MLO_2	70.66±1.42	66.70±2.01	65.14±1.01	65.62±3.53
MLO_3	88.12±1.51	90.70±0.73	87.32±1.25	92.42±0.36
MLO_4	85.44±2.37	90.10±1.96	84.86±2.97	91.32±0.78
MLO_5	86.30±1.67	91.94±0.73	91.64±1.63	93.64±0.70
Average	82.63±1.71	84.86±1.39	82.24±1.74	85.75±1.31

Table XI shows the experimental results in this experiment. From this table we find that both the multiple-instance method and Dempster-Shafer method are effective in multiple-aspect classification. Compared with the decision tree method, logistic regression method shows better performance when combining with multiple-instance method and Dempster-Shafer method.

VI. CONCLUSIONS

In this paper, we have demonstrated the performance improvements in the classification of side scan sonar images obtained by using feature sets corresponding to multiple sonar views of the same object. There are two basic ways in which the multiple feature sets can be utilized. The first approach uses multiple-instance classification methods to classify multiple feature vectors. The second approach consists of fusing the multiple individual classifications of the multiple feature vectors with the DS method [21]. Tree methods and Logistic Regression methods were chosen as the base learners for these two approaches in our experiments. We have transformed the multiple aspects classification problem into a multiple-instance learning problem [22] and we have found that the DS method [21] can be an alternative and effective framework for the multiple-instance learning problem [22].

We have also studied the classification performances as a function of the number of aspects. From the experimental results we find that for both frameworks, the performance is enhanced when more views are used for classification.

Comparing the single aspect classification rates with two multiple-aspect approaches on different shapes, we see that collecting multiple views produces a significant increase in hit rate and a significant decrease in error rate for all mine shapes. Moreover, the multiple-instance learning [22] method results in better classification performance than the Dempster-Shafer (DS) method [21] with the single aspect classifier on all shapes on the same number of aspects combined.

Moreover, we have compared the multiple aspect classification frameworks with different classifiers presented in this paper. The experimental results show that both the multiple-instance method and Dempster-Shafer method [21] are effective in multiple-aspect classification. Compared with the tree method, logistic regression method shows better performance in both of the presented frameworks. Our future works will involve applying these two frameworks on other multiple aspect learning applications and using more classifiers on these two frameworks.

REFERENCES

[1] James A. Hanley, Barbara J. McNeil. The Meaning and Use of the Area under a Receiver Operating Characteristic (ROC) Curve. Radiology 143 (1): 29–36. (1982).

[2] Canny, J., A Computational Approach to Edge Detection, IEEE Trans. Pattern Analysis and Machine Intelligence, 8(6):679–698, (1986).

[3] B. Zerr, B. Stage. Three-dimensional reconstruction of underwater objects from a sequence of sonar images, Proceedings of the IEEE International Conference on Image Processing, pp. 927–930, (1996).

[4] B. Zerr, B. Stage and A. Guerrero, Automatic Target Classification Using Multiple Sidescan Sonar Images of Different Orientations, SACLANT CEN Memorandum SM-309 (1997).

[5] Dietterich, T., Lathrop, R. and Lozano-P´erez, T., "Solving the multiple instance problem with the axis-parallel rectangles," Artificial Intelligence, 89(1-2), 31–71 (1997).

[6] B. Zerr, E. Bovio, B. Stage, Automatic mine classification approach based on AUV manoeuverability and cots side scan sonar, Proceedings of Goats 2001 Conference, La Spezia, Italy, (2001).

[7] M. Couillard, J. Fawcett, M. Davison and V. Myers, Optimizing time-limited multi-aspect classification, Proceedings of the Institute of Acoustics 29(6), 89-96 (2007).

[8] J. Fawcett, V. Myers, D. Hopkin, A. Crawford, M. Couillard, B. Zerr. Multiaspect classification of sidescan sonar images: Four different approaches to fusing single-aspect information, Oceanic Engineering, IEEE Journal of 35(4): 863 –876 (2010).

[9] S. Reed, Y. Petillot, J. Bell, Model-based approach to the detection and classification of mines in side scan sonar, Applied Optics 43(2): 237–246. (2004).

[10] S. Reed, Y. Petillot, J. Bell, Automated approach to classification of mine-like features in sidescan sonar using highlight and shadow information, IEE Proc. Radar, Sonar & Navigation 151 (No.1), 48-56, (2004).

[11] V. Myers, D. P. Williams, A POMDP for multi-view target classification with an autonomous underwater vehicle, OCEANS, pp. 1-5, (2010).

[12] V. Myers, D. P. Williams, Adaptive Multiview Target Classification in Synthetic Aperture Sonar Images Using a Partially Observable Markov Decision Process, Oceanic Engineering, IEEE Journal of, On page(s): 45 - 55, Volume: 37 Issue: 1, Jan. (2012)

[13] G. Dobeck, "Fusing sonar images for mine detection and classification," Proc. SPIE—Int. Soc. Opt. Eng., vol. 3710, 1999, DOI: 10.1117/12.357082.

[14] J. Tucker, N. Klausner, and M. Azimi-Sadjadi, "Target detection in M-disparate sonar platforms using multichannel hypothesis testing," in Proc. OCEANS Conf., Quebec City, QC, Canada, 2008, DOI: 10.1109/ OCEANS.2008.5151818.

[15] M. Azimi-Sadjadi, A. Jamshidi, and G. Dobeck, "Adaptive underwater target classification with multi-aspect decision feedback," Proc. SPIE—Int. Soc. Opt. Eng., vol. 4394, 2001, DOI: 10.1117/12.445444.

[16] X. Xu, Statistical learning in multiple instance problems, PhD diss., the University of Waikato, (2003).

[17] X. Xu and E. Frank. Logistic regression and boosting for labeled bags of instances. In Lecture Notes in Computer Science, volume 3056, pages 272–281, April 2004.

[18] Hall, D. L. and Steinberg, A.. Dirty Secrets in Multisensor Data Fusion, http://www.dtic.mil. (2001).

[19] Blockeel, H., Page, D., Srinivasan, A.: Multi-instance tree learning. In: ICML (2005).

[20] Quinlan, J. R. C4.5: Programs for Machine Learning. Morgan Kaufmann Publishers, (1993).

[21] Shafer, Glenn; Dempster–Shafer theory, (2002).

[22] Hosmer, David W.; Lemeshow, Stanley. Applied Logistic Regression (2nd ed.). Wiley. (2000).