MULTIPLE INSTANCE LEARNING FOR HYPERSPECTRAL IMAGE ANALYSIS

Jeremy Bolton and Paul Gader

University of Florida

1. INTRODUCTION

For the purposes of target classification in remotely sensed hyperspectral imagery (HSI), an algorithm will typically analyze spatial and spectral information from an observed image and compare this information to a previously learned (or known) target concept or model. Typically during the learning process of a target model, model parameters are optimized using machine learning techniques, given target and non-target signatures. In standard machine learning techniques, it is imperative that the learning algorithms know the class labels of the training data. However, in remotely sensed hyperspectral imagery, images are typically sensed from a distance which allows for the introduction of corrupting signals, image formation errors, spectral mixing, and geo-registration (mapping from ground location to image location) errors. The result is data with uncertain class labels; an issue that cannot be confronted using standard machine learning techniques. Furthermore (during testing), many classification algorithms make assumptions about target and background areas which may also be confounded by this uncertainty.

Multiple instance learning (MIL) is a learning framework that attempts to combat these issues [1, 2], explicitly. Only recently has MIL been used for HSI analysis [3]. In this work, Torrione *et. al* applied a unique interpretation of MIL to HSI analysis, which differed from the standard MIL problem statement. In the following, RSF-MIL and MI-RVM are used to learn a target emissivity signal for the purposes of landmine detection, which has not previously been done. These two MIL approaches have the highest classification performance given benchmark data sets [4] and will help identify the benefits of using MIL approaches for the purposes of HSI image analysis.

The remainder of this paper is organized as follows. First, MIL is discussed including Maron's statistical development of Diverse Density. Next, the HSI under test is briefly described and an intuitive reasoning behind an MIL solution is developed. Then, the algorithms under test are detailed. Finally, experimental design, results and conclusions are detailed.

2. MULTIPLE INSTANCE LEARNING: ANALYSIS OF BAGS

In MIL, a learner is presented with sets of samples (refered to as bags); whereas in standard techniques, a learner is presented with individual samples. Bags are labeled positive if there exists at least one sample that induces a target concept and are labeled negative if every sample does not induce a target

Generally speaking, MIL approaches have two benefits over standard machine learning techniques. First, it has the ability to learn target concepts when the labeling of the training data is uncertain. Second, it typically provides a set-based classifier, thus a set of data can be classified rather than a single sample at a time, which is very useful in image analysis. In the following, DD and RSF-MIL are reviewed and benefits of Multiple Instance (MI) approaches for image analysis are listed. The reader is directed to the literature for a comprehensive review of MIL [2, 4, 1].

Research was sponsored by the U. S. Army Research Office and U. S. Army Research Laboratory and was accomplished under Cooperative Agreement Number DAAD19-02-2-0012. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Army Research Office, Army Research Laboratory, or the U. S. Government. The U. S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation hereon. Research was also partially supported by NSF program Optimized Multi-algorithm Systems for Detecting Explosive Objects Using Robust Clustering and Choquet Integration (CBET-0730484). The authors thank William Clark and Russell Harmon of ARO, Miranda Silvious of NVESD and Michael Cathcart of Georgia Tech for their support of this research.

2.1. Diverse Density

In standard DD approaches, a target concept, which is characterized by a feature vector t, is learned given positively and negatively labeled bags, B_i^+ and B_i^-, respectively.

$$\hat{t} = \text{argmax}_t \left[\prod_{i=1}^{n^+} P(t|B_i^+) \prod_{i=1}^{n^-} P(\neg t|B_i^-) \right] \tag{1}$$

Assuming a Noisy OR-Gate model [5], the terms in (1) can be calculated in terms of the bags constituent samples B_{ij} using (2) and (3).

$$P(t|B_i^+) = P(t|B_{i1}^+, B_{i2}^+, ..., B_{im^+}^+) = 1 - \prod_{j=1}^{m^+} \left(1 - P(t|B_{ij}^+\right) \tag{2}$$

$$P(\neg t|B_i^-) = P(t|B_{i1}^-, B_{i2}^-, ..., B_{im^-}^-) = \prod_{j=1}^{m^-} \left(1 - P(t|B_{ij}^-\right) \tag{3}$$

Note (2) states that *the probability a bag induces a target concept is equal to the probability that it is not the case that each sample does not induce a target concept*. Equation (3) states that *the probability that a bag does not induce a target concept is equal to the probability that each samples does not induce a target concept*.

2.2. RSF-MIL

A more general MIL solution, Random Set Framework for Multiple Instance Learning (RSF-MIL) has recently been introduced by the authors and used for feature learning in GPR image analysis [4]. RSF-MIL provides a random set-model to perform analysis on bags. A random set Ξ is completely characterized by its capacity functional $T_\Xi(X) \equiv P(\Xi \cap X \neq \phi)$. The capacity functional permits analysis on set values (bags), thus providing a solid mathematical framework for MIL.

The probability that an observed set, *X*, has a non-empty intersection with a random set,Ξ, is equal to the probability that it is not the case that each sample, x_p, has an empty intersection with the random set. Using this operator, similarities between bags can be analyzed, models can be learned, and posterior probabilities can be calculated. Note formulations of random sets that can be used in (4) provides a more general mathematical framework and solution than (2) [4].

$$T_\Xi(X) = 1 - \prod_{p=1}^{P} (1 - T_\Xi(x_p)), \text{where } X = x_1, x_2, ..., x_P \tag{4}$$

2.3. MI-RVM

MI-RVM as developed by Raykar et al [6] uses a noisy-OR model with a logistic sigmoid, σ to calculate probabilities of observed bags.

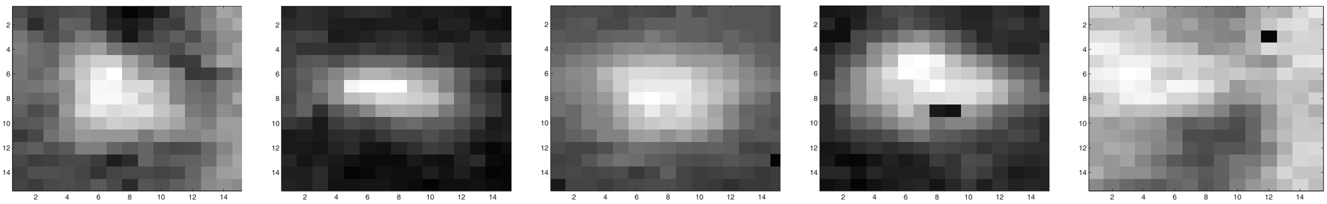
$$P(y = 1|B_i^+) = P(t|B_{i1}^+, B_{i2}^+, ..., B_{im^+}^+) = 1 - \prod_{j=1}^{m^+} \left(1 - \sigma(w^T B_{ij}^+)\right) \tag{5}$$

In this model, the weight vector is modeled using a Gaussian prior. Optimization is performed using the Newton-Raphson update. Feature selection is optimized using type II maximum likelihood method.

3. HYPERSPECTRAL DATA SET AND MIL

The classifiers under test were applied to AHI (Airborne-Hyperspectral-Imager) imagery [7]. AHI was flown over an arid site at altitudes of 300 m and 600 m resulting in a spatial resolution of 10 cm and 15 cm, respectively. Each image contains 70 spectral bands after trimming and binning, ranging over LWIR wavelengths 7.88um 11.02um. Each image contains millions of spatial pixels. Experiments are performed using the apparent emissivity of the observed spectra, which is calculated using the well-known Emissivity Normalization method. Classification is done in the emissivity space to mitigate the effects of varying spectra intensities, but may not account for other context-dependent characteristics.

The well-known RX algorithm was applied as a prescreener to identify areas of interest (AOIs), which were used to construct HSI chips. The resulting HSI chips contained various anomalies which included targets and non-targets. Targets consisted of areas of disturbed soil and various types of mines buried 10.2 cm deep or flush with the soil. Examples are shown in Figure 1. There are a approximately 900 targets and 1,200 non-targets.



(1) HSI Chip 1 (2) HSI Chip 2 (3) HSI Chip 3 (4) HSI Chip 4 (5) HSI Chip 5

Fig. 1. Examples of HSI chips showing variations of location, number, and shape of target signatures within an AOI. Note each spatial pixel in each chip has a corresponding spectral vector; the spectral response for 9.4um is shown within the image.

In remotely sensed HSI analysis, a prescreener (simple anomaly detector) is typically run to identify AOIs. This lessens the amount of imagery that a more complex detection algorithm will need to process, thus lessening overall computation time. In many standard remote sensing image classifiers, target and background statistics are calculated and used for overall confidence calculation. However, in these approaches target and background areas are assumed to be known and constant (as shown in Figure 1), which may not be the case (see Figure 2).

In an AOI the observable evidence of a target may be highly variable. Errors in geolocation information and centering the AOIs may lead to testing and training issues since the target may reside in variable spaces within the AOI, see Figure 1. Furthermore, the number and shape of the exemplar target signatures may vary due to variable thermal characteristics, variable target sizes, spectra mixing, and errors in image formation, see Figure 1.

These factors can confound learning and classification algorithms which rely on knowing the target and background areas, *a priori*. However, these learning and testing conditions exactly fit the MI problem, where the shape, and number of target and background signatures are explicitly assumed to be variable since analysis is performed using bags.

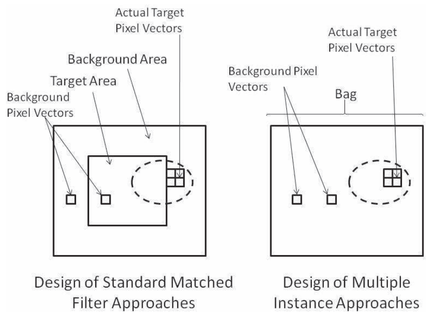


Fig. 2. Spatial assumptions made in standard matched filter, image processing algorithms as compared to MI approaches.

4. EXPERIMENTAL DESIGN, RESULTS AND CONCLUSIONS

Multiple Instance (MI) algorithms under test include RSF-MIL and MI-RVM. RSF-MIL and MI-RVM have the best reported [4] classification results on Musk data sets (MI benchmarks)[8]. Results are compared to the Whiten-Dewhiten (WDW) transform, developed my Mayer *et al.* [9], that uses target and background statistics of the training and testing data to calculate target confidence.

Experiments were performed using crossvalidation at the image level, *i.e.*, when classifying HSI chips from image i, optimization of the classifiers' parameters (2) is done using only those HSI chips from images other than the image i. Note, this decreases classification results since each image was observed in somewhat different contexts; however, this mimics the real-world testing situation.

Each of the MI classifiers under test do not account for spatial information in their present form. Fourteen models were constructed for the WDW algorithm and treated as different contexts. Rather than combining target and background statistics into one model, multiple models were used and confidence was averaged over each model (this improved the classification results). Also note that in WDW there are some spatial assumptions about target and background areas, so confidence values were computed by averaging the confidence within the assumed target area.

In summary, WDW requires 1) prior knowledge during training of target and background models, 2) it requires a large amount of training data and training parameters to model target and background space in high dimensions, 3) does not supply feature selection and 4) does not supply a set-based classifier. Conversely, RSF-MIL and MI-RVM 1)do not require prior knowledge during training target and background data, 2) do not require a large amount of training data and training parameters to model target and background space in high dimensions, 3) supply a feature selection capability and 4) supply a set-based classifier. Also, note that WDW incorporated spatial and contextual information, whereas, the MI approaches under test did not make use of this information and outperformed WDW.

Future work includes the incorporation of spatial information into the MI framework, which should improve overall classification results of MI approaches.

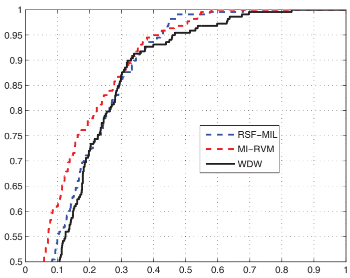


Fig. 3. ROC curve of classification rates.

5. REFERENCES

[1] T.G. Dietterich, R.H. Lathrop, and T. Lozano-Pérez, "Solving the multiple instance problem with axis-parallel rectangles," *Artificial Intelligence*, vol. 89, no. 1-2, pp. 31–71, 1997.

[2] O. Maron and T. Lozano-Pérez, "A framework for multiple-instance learning," *Advances in neural information processing systems*, pp. 570–576, 1998.

[3] Ratto C. Torrione, P. and L. Collins, "Multiple Instance and Context Dependent Learning in Hyperspectral Data," *IEEE Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing*, 2009.

[4] J. Bolton and P. Gader, "Random Set Framework for Multiple Instance Learning," *Information Sciences*, submitted.

[5] J. Pearl, *Probabilistic reasoning in intelligent systems: networks of plausible inference*, Morgan Kaufmann, 1988.

[6] V.C. Raykar, B. Krishnapuram, J. Bi, M. Dundar, and R.B. Rao, "Bayesian multiple instance learning: automatic feature selection and inductive transfer," in *Proceedings of the 25th international conference on Machine learning*. ACM New York, NY, USA, 2008, pp. 808–815.

[7] P.G. Lucey, T.J. Williams, M. Mignard, J. Julian, D. Kobubun, G. Allen, D. Hampton, W. Schaff, M.J. Schlangen, E.M. Winter, et al., "AHI: an airborne long-wave infrared hyperspectral imager," in *Proceedings of SPIE*. SPIE, 1998, vol. 3431, p. 36.

[8] A. Asuncion and D.J. Newman, "University of california, irvine, school of information and computer sciences, machine learning repository: http://mlearn.ics.uci.edu/mlrepository.html," 2007.

[9] R. Mayer, F. Bucholtz, and D. Scribner, "Object detection by using" whitening/dewhitening" to transform target signatures in multitemporal hyperspectral and multispectral imagery," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 41, no. 5, pp. 1136–1142, 2003.