EXTENDED FUNCTIONS OF MULTIPLE INSTANCES FOR TARGET CHARACTERIZATION

Alina Zare and Changzhe Jiao

Electrical and Computer Engineering, University of Missouri

ABSTRACT

An extension of the Function of Multiple Instances (FUMI) algorithm for target characterization is presented. FUMI is a generalization of Multiple Instance Learning (MIL). However, FUMI differs significantly from standard MIL and supervised learning approaches because only data points which are functions of class concepts are available. For applicability to hyperspectral data, this paper addresses the problem in which data points are convex combinations of target and non-target concepts. The presented method, eFUMI, extends previous methods to allow for further unspecificity in training labels while estimating target and non-target concepts, the number of non-target concepts, and the weight associating each concept to each data point. For eFUMI, training data need only binary labels indicating whether a spatial area in an input image contains or does not contain some proportion of target material; the specific locations or target proportions for training data are not needed. After learning the target concept, target detection can be performed on test data. Results showing sub-pixel target detection on simulated and real Hyperspectral data are provided.

1. INTRODUCTION

Multiple-instance learning (MIL) is a variation on supervised learning for problems with incomplete labels for training samples [1–3]. In MIL, training data are divided into positive and negative "bags." A bag is defined to be a multi-set of data points. A positive bag includes at least one target point. The exact number of points belonging to the target class in each positive bag is unknown. Negative bags are composed entirely of non-target data. MIL methods are effective for learning target *concepts* and developing classifiers for cases where accurate sample-level labeled training data is unavailable. Here, *concepts* refer to generalized class prototypes in the feature space; in hyperspectral image unmixing, a FUMI concept is equivalent to an endmember.

Functions of Multiple Instance (FUMI) approaches [4–6] are a generalization of MIL. FUMI treats each data point as a function of a positive or a negative bag. Given some functional form, FUMI learns the target and non-target concepts and the function parameters defining the relationship between each data point and the concepts. Specifically, suppose there is a given data set $\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_N\}$ where each data point is some unknown function of concepts, $\mathbf{x}_i = f(\mathbf{E}_i, \mathbf{p}_i)$ where $\mathbf{p}_i$ are the set of parameters for $\mathbf{x}_i$ and $\mathbf{E}_i$ is the "bag" of concepts that contribute in a non-negligible way to the data point $\mathbf{x}_i$. Each training point $\mathbf{x}_i$ is given a binary label $l(\mathbf{x}_i)$ where $l(\mathbf{x}_i) = 1$ if $\mathbf{e}_T \in \mathbf{E}_i$ and $l(\mathbf{x}_i) = 0$ if $\mathbf{e}_T \notin \mathbf{E}_i$. Using these binary labels, FUMI estimates the functional parameters, $\mathbf{p}_i$, for each data point and the target and non-target concepts for the data set, $\mathbf{E} = \cup_i \mathbf{E}_i$, where $\mathbf{E}_i$ is a subset of $\mathbf{E}$ for data point $\mathbf{x}_i$, and $\mathbf{E}_i$ and $\mathbf{E}_j$ are not mutually exclusive.

The proposed Extended Functions of Multiple Instance, eFUMI, approach allows for an additional layer of unspecificity in the training labels. In remote sensing applications, groundtruth information is often uncertain (with respect to UTM- or pixel-coordinates). For example, UTM-coordinates of several targets of interest may have been collected with an accuracy ranging across several pixels. Thus, accurate pixel-level groundtruth is unknown; only an approximate locations are available. Furthermore, the proportion of each pixel associated with a particular target is often unknown as it depends on the relationship between spatial resolution and the (unknown) field of view associated with each pixel. Previous FUMI approaches were able to address the uncertainty associated with unknown proportion values. However, these previous methods require data point-specific binary labels and, thus, cannot address location uncertainty. The proposed eFUMI algorithm addresses both unspecificity due to both unknown target proportion and location in training data.

Section 2 presents the proposed eFUMI. Section 3 includes results on simulated and measured hyperspectral data.

2. EXTENDED FUMI

eFUMI extends FUMI to be able to learn the target concept without the need of instance-level labels. Specifically, the goal of eFUMI is to estimate a target concept, $\mathbf{e}_T$, non-target concepts, $\mathbf{e}_k, \quad \forall k = 1, \ldots M$, the number of needed non-target concepts, M, and the function parameters, $\mathbf{p}_i$, which define the relationship between each data point, $\mathbf{x}_i$, and concept. These are estimated given a set of input training data, $\{\mathbf{x}_i\}_{i=1}^N \in \mathbb{R}^D$ which have been partitioned into K "bags," $\mathbf{B} = \{B_1, \ldots, B_K\}$, with associated bag-level labels, $L = \{L_1, \ldots, L_K\}$. Each training point is assumed to be a function of target and non-target concepts, $\mathbf{x}_i = f(\mathbf{E}_i, \mathbf{p}_i)$. In the results shown here, we consider the case of a convex combination for the functional form, $\mathbf{x}_i = p_{iT}\mathbf{e}_T + \sum_{k=1}^M p_{ik}\mathbf{e}_k$ subject to the constraints that $p_{iT} + \sum_{k=1}^M p_{ik} = 1$, $p_{iT} \geq 0$,

$p_{ik} \geq 0$, $\|\mathbf{e}_k\|^2 = 1$, and $\|\mathbf{e}_T\|^2 = 1$. Furthermore, the bag-level labels are *unspecific* since, if for bag B_j with $L_j = 1$ (thus, B_j is a positive bag), this indicates that there is at least one data point in B_j with a positive p_{iT} indicating some presence of target, as in (1), where $\boldsymbol{\varepsilon}_i$ is an error term.

$$\text{if } L_j = 1, \exists \mathbf{x}_i \in B_j \text{ s.t. } \mathbf{x}_i = p_{iT}\mathbf{e}_T + \sum_{k=1}^{M} p_{ik}\mathbf{e}_k + \boldsymbol{\varepsilon}_i, p_{iT} > 0 \tag{1}$$

However, the exact number of data points in a positive bag with a target contribution (i.e., $p_{iT} > 0$), is unknown, also, the target proportion is unknown. Furthermore, if B_j is a negative bag (i.e., $L_j = 0$), then this indicates that none of the data in B_j contain any target, as in (2).

$$\text{if } L_j = 0, \forall \mathbf{x}_i \in B_j, \mathbf{x}_i = \sum_{k=1}^{M} p_{ik}\mathbf{e}_k + \boldsymbol{\varepsilon}_i \tag{2}$$

Given training data of this form, instance-level labels are unknown. eFUMI addresses this problem using an Expectation Maximization (EM) approach in which the instance-level labels are the hidden, latent variables in the EM model. The assumed *complete* data log-likelihood is proportional to (3), where the first term is the approximation error between the given data and the product of estimated endmembers and corresponding proportions; the second and third terms are to drive the endmembers have a tight fit around the global data mean; the forth term is the sparse promoting term used to determine M, the number of constituent endmembers. In (3), z_i is the unknown instance-level labels, μ_0 is the global data mean, and u is a regularization parameter; the weight $w_i = \frac{\alpha N^-}{N^+}$ for points in positively labeled bags and $w_i = 1$ for points in negatively labeled bags where N^- and N^+ are the number of points in negative and positive bags, respectively, as described in [5]; finally, $\gamma_k = \frac{\Gamma}{\sum_{i=1}^{N} p_{ik}^{(t-1)}}$ is used in a sparsity promoting term to drive proportion values with respect to unnecessary endmembers to zero and prune the unnecessary endmembers with Γ set to a fixed value as in [7]. Following EM optimization methods, the desired parameters are estimated by performing the *E-step* by computing the expected value over the latent variables resulting in (4). In (4), $\boldsymbol{\theta}^t$ is the set of parameters estimated at iteration t and $p(z_i|\boldsymbol{x}_i, \boldsymbol{\theta}^{(t-1)})$ is determined given the parameter set estimated in the previous iteration and the constraints of the bag-level labels, L_j, as shown in (5). Then, the *M-step* is performed by optimizing (4) for each of the desired parameters. The algorithm is summarized in Alg. 1.

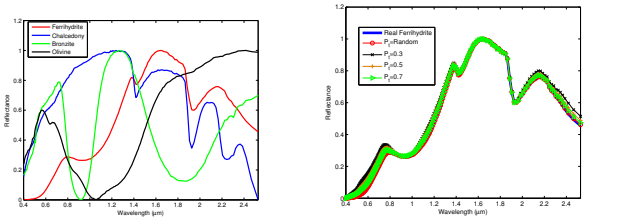
$$p(z_i|\mathbf{x}_i, \theta^{(t-1)}) = \tag{5}$$
$$\begin{cases} p(z_i = 0|\mathbf{x}_i \in B_j^+, \boldsymbol{\theta}^{(t-1)}) = e^{-\beta\left\|\mathbf{x}_i - \sum_{k=1}^{M} p_{ik}\mathbf{e}_k\right\|_2^2} \\ p(z_i = 1|\mathbf{x}_i \in B_j^+, \boldsymbol{\theta}^{(t-1)}) = 1 - e^{-\beta\left\|\mathbf{x}_i - \sum_{k=1}^{M} p_{ik}\mathbf{e}_k\right\|_2^2} \\ p(z_i = 0|\mathbf{x}_i \in B_j^-, \boldsymbol{\theta}^{(t-1)}) = 1 \\ p(z_i = 1|\mathbf{x}_i \in B_j^-, \boldsymbol{\theta}^{(t-1)}) = 0 \end{cases}$$

Algorithm 1 eFUMI EM algorithm

Require:

1: Initialize $\theta^0 = \{\mathbf{e}_T, \mathbf{E}, \mathbf{P}\}$, normalize input data, $t = 1$
2: **repeat**
3: *E-step*: Compute (5) given θ^{t-1}
4: *M-Step*:
5: Update $\mathbf{e}_T$ by minimizing (4) w.r.t. $\mathbf{e}_T$[1]
6: Update $\mathbf{E}$ by minimizing (4) w.r.t. $\mathbf{E}$
7: Update $\mathbf{P}$ by minimizing (4) w.r.t. $\mathbf{P}$
8: $t \leftarrow t + 1$
9: **until** Convergence
10: **return** $\mathbf{e}_T, \mathbf{E}, \mathbf{P}$

Due to space constraints, the update equations used on lines (5)-(7) and their corresponding derivations are posted here: `http://engineers.missouri.edu/zarea/tigersense/publications/`



(a) Endmembers used in Synthetic Data Experiment

(b) Comparison of true and estimated target spectra in simulated data experiment

Fig. 1. Simulated Data True and Estimated Target Spectra

3. EXPERIMENTAL RESULTS

eFUMI was tested both on synthetic data and real hyperspectral data. First, eFUMI was run on simulated data generated using random mixtures of Ferrihydrite, Chalcedony, Bronzite and Olivine spectra selected from the USGS spectral library [8]. The spectra are shown in Fig. 1(a). The spectra have 209 bands ranging from 0.4μm to 2.5μm in wavelength. Ferrihydrite was labeled as the target endmember. Five bags were generated with 1000 points in each bag, the first two bags were labeled as positive with 250 containing some proportion of target. The number of background endmembers found in each point was drawn from a uniform multinomial distribution. Then, proportions were generated using a uniform Dirichlet distribution. The mean square error (MSE) and mean spectral angle distance (MSAD) between estimated and real Ferrihydrite spectrum are calculated over 10 randomly initialized runs of eFUMI and shown in Table 1. Fig. 1(b) plots estimated vs. real target spectra, where P_t denotes the mean target proportion value in highly mixed target points.

eFUMI was also run on the RIT SHARE 2012 data set collected near Rochester, NY [9]. A subset of the AVON hyperspectral imagery from this collection containing 300×320 pixels with 360 bands correspond wavelengths from 400.3nm to 2452.8nm was used. The spatial resolution was 1 meter.

$$F = -\frac{1}{2}(1-u)\sum_{i=1}^{N} w_i \left\| (\mathbf{x}_i - z_i p_{iT}\mathbf{e}_T - \sum_{k=1}^{M} p_{ik}\mathbf{e}_k) \right\|_2^2 - \frac{u}{2}\sum_{k=1}^{M}\left\|(\mathbf{e}_k - \boldsymbol{\mu}_0)\right\|_2^2 - \frac{u}{2}\left\|(\mathbf{e}_T - \boldsymbol{\mu}_0)\right\|_2^2 - \sum_{k=1}^{M}\gamma_k \sum_{i=1}^{N} p_{ik} \quad (3)$$

$$E[F] = \sum_{z_i \in \{0,1\}} \left[-\frac{1}{2}(1-u)\sum_{i=1}^{N} w_i P(z_i|\mathbf{x}_i, \boldsymbol{\theta}^{(t-1)}) \left\| \mathbf{x}_i - z_i p_{iT}\mathbf{e}_T - \sum_{k=1}^{M} p_{ik}\mathbf{e}_k \right\|_2^2 \right] - \frac{u}{2}\sum_{k=1}^{M}\|\mathbf{e}_k - \boldsymbol{\mu}_0\|_2^2 - \frac{u}{2}\|\mathbf{e}_T - \boldsymbol{\mu}_0\|_2^2 - \sum_{k=1}^{M}\gamma_k \sum_{i=1}^{N} p_{ik} \quad (4)$$

Mean P_t	0.3	0.5	0.7
MSE	0.8362 ± 0.0321	0.5766 ± 0.0163	0.3067 ± 0.0148
Mean SAD	0.0959 ± 0.0032	0.0669 ± 0.0023	0.0367 ± 0.0017

Table 1. USGS simulated data error and standard deviation (over 10 runs) between the estimated and true endmembers. The values in the table are the total error summed over all wavelengths and averaged over 10 runs with random initialization.

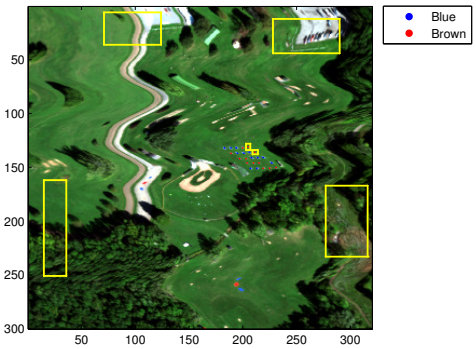


Fig. 2. RGB image of a subset of the RIT SHARE 2012 AVON AM Data containing 24 Scattered Target Points

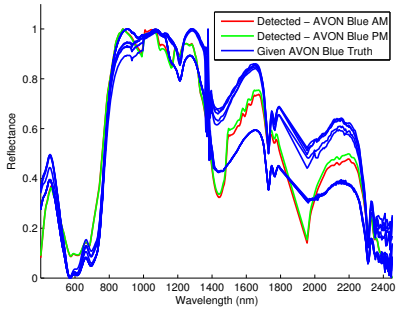


Fig. 3. Estimated vs. Ground Truth Blue Target Spectra

Two sets of this data were provided: (1) AVON AM collected in the morning of 20-Sept-2012 and (2) AVON PM collected in the afternoon of the same day. We considered 12 examples of two target types: man-made blue and brown tarps. As shown in Fig. 2, a 5×5 rectangular region centered around each approximate target location were labeled as positive bags and large background regions (shown as large yellow blocks in Fig. 2) were labeled as negative bags.

eFUMI estimated Blue target spectra are shown in Fig. 3. These are compared to truth spectra measured using an Analytical Spectral Devices FieldSpec Pro spectrometer and provided in the RIT dataset. There are multiple sets of given ground truth spectra due to differences in target shape (circle, equilateral triangle and square) and target opaqueness (due to varying ply values). Fig. 5 displays unmixing results using eFUMI estimated spectra. Fig.5(a) is the proportion map for Blue targets. Also, eFUMI estimates background proportions as shown in Fig.5(b), Fig.5(c) and Fig.5(d) for bare soil, vegetation and manmade stuff, respectively.

In order to investigate the quality of the estimated target spectra, the adaptive coherence/cosine estimator (ACE) [10, 11] was applied for target detection for both target types. Fig. 4(a) shows the detection map for the Blue targets. This illustrates that eFUMI is capable of detecting a quality target signature from inaccurately labeled points. For quantitative evaluation, receiver operating curves (ROC) were computed on cross-validation results between AVON AM and AVON PM both for Blue and Brown targets [12]. The detailed value of PD VS. FAR of Fig. 6 is shown in Table 2. As shown in Fig. 6 and Table 2, eFUMI cross-validation results outperform the ASD-measured spectra as well as hand-selected target spectra from the scene.

4. REFERENCES

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

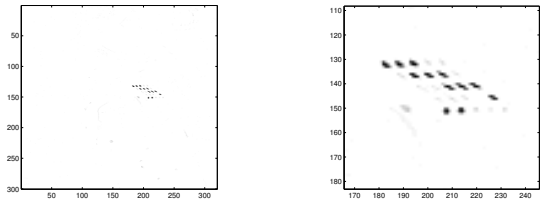
[2] O. Maron and T. Lozano-Perez, "A framework for multiple-instance learning." *Neural Inform. Process. Syst.*, vol. 10, 1998.

[3] J. Bolton, P. Gader, H. Frigui, and P. Torrione, "Random set framework for multiple instance learning," *Inform. Sci.*, vol. 181, no. 11, pp. 2061–2070, June 2011.

[4] A. Zare and P. Gader, "Pattern recognition using functions of multiple instances," *Proc. of the Int. Conf. on Pattern Recognition*, pp. 1092–1095, Aug. 2010.

FAR	1.02×10^{-5}	1.02×10^{-4}	5.12×10^{-3}	FAR	1.02×10^{-5}	1.02×10^{-4}	5.12×10^{-3}
eFUMI Blue PM	100%	100%	100%	eFUMI Brown PM	100%	100%	100%
ASD Blue PM	0%	0%	16.67%	ASD Brown PM	0%	0%	16.67%
Picked Blue PM	0%	0%	8.33%	Picked Brown PM	0%	0%	25%
eFUMI Blue AM	100%	100%	100%	eFUMI Brown AM	83.33%	100%	100%
ASD Blue AM	0%	0%	25%	ASD Brown AM	0%	0%	8.33%
Picked Blue AM	0%	8.33%	25%	Picked Brown AM	0%	0%	25%

Table 2. Probability of Detection (PD) VS. False Alarm Rate (FAR) by ACE Detector Using Different Target Spectra



(a) ACE Output (b) ACE Output Zoomed

Fig. 4. ACE with eFUMI Target Spectrum Detection Results of Blue Targets on AVON PM

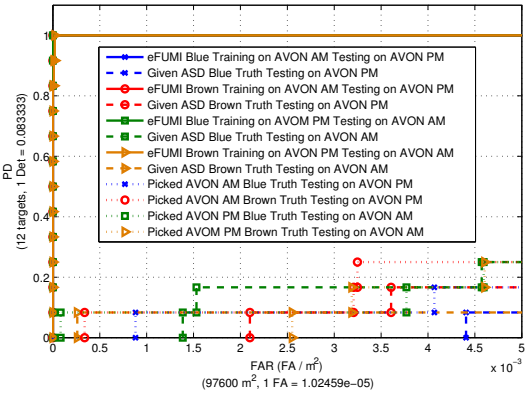
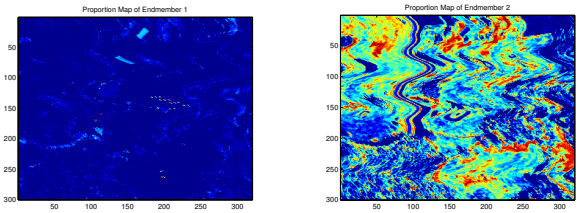
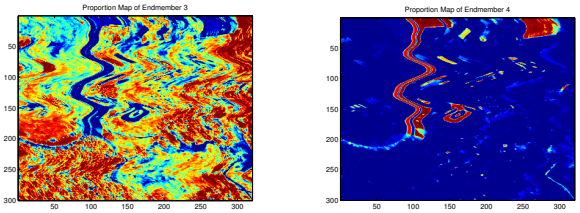


Fig. 6. ROC Results of eFUMI-estimated Target Spectra vs ASD Truth Spectra and hand-selected Target Spectra on AVON Data



(a) Proportion Map of Blue Target Endmember (b) Proportion Map of Background Endmember 1-Bare Soil



(c) Proportion Map of Background Endmember 2-Vegetation (d) Proportion Map of Background Endmember 3-Manmade Stuff

Fig. 5. Proportion Maps of Detection Results for Blue Targets on AVON AM Data by eFUMI

[5] A. Zare, P. Gader, J. Bolton, S. Yuksel, T. Dubroca, and R. Close, "Subpixel target spectra estimation using functions of multiple instances," in *3rd IEEE Workshop on Hyperspectral Image and Signal Process.: Evolution in Remote Sensing*, 2011.

[6] A. Zare and P. Gader, "Multiclass subpixel target detection using functions of multiple instances," in *SPIE Defense, Security, and Sensing*, 2011, pp. 804 811–804 811–5.

[7] ——, "Sparsity promoting iterated constrained endmember detection for hyperspectral imagery," *IEEE Geoscience and Remote Sensing Letters*, vol. 4, no. 3, pp. 446–450, July 2007.

[8] R. N. Clark, A. Swayze, R. Wise, K. E. Livo, T. M. Heofen, R. F. Kokaly, and S. J. Sutley. (2014, Mar.) USGS digital spectral library. Spectroscopy Lab, U.S. Geological Surey. [Online]. Available: http://speclab.cr.usgs.gov/spectral-lib.html

[9] A. Giannandrea, N. Raqueno, D. W. Messinger, J. Faulring, J. P. Kerekes, J. van Aardt, K. Canham, S. Hagstrom, E. Ontiveros, A. Gerace *et al.*, "The share 2012 data campaign," in *SPIE Defense, Security, and Sensing*. International Society for Optics and Photonics, 2013, pp. 87 430F–87 430F.

[10] S. Kraut and L. Scharf, "The CFAR adaptive subspace detector is a scale-invariant GLRT," *IEEE Trans. on Signal Process.*, vol. 47, no. 9, pp. 2538 –2541, Sept. 1999.

[11] S. Kraut, L. L. Scharf, and L. T. McWhorter, "Adaptive subspace detectors," *Signal Process., IEEE Trans. on*, vol. 49, no. 1, pp. 1–16, 2001.

[12] T. Glenn, A. Zare, P. Gader, and D. Dranishnikov, "Bullwinkle: Scoring code for sub-pixel targets (version 1.0) [software]," 2013, http://engineers.missouri.edu/zarea/code/.