Multitarget Multiple-Instance Learning for Hyperspectral Target Detection

Susan Meerdink, James Bocinsky, Alina Zare, *Senior Member, IEEE*, Nicholas Kroeger, Connor McCurley, *Student Member, IEEE*, Daniel Shats, and Paul Gader, *Fellow, IEEE*

***Abstract*— In remote sensing, it is often challenging to acquire or collect a large data set that is accurately labeled. This difficulty is usually due to several issues, including but not limited to the study site's spatial area and accessibility, errors in the global positioning system (GPS), and mixed pixels caused by an image's spatial resolution. We propose an approach, with two variations, that estimates multiple-target signatures from training samples with imprecise labels: multitarget multiple-instance adaptive cosine estimator (MTMI-ACE) and multitarget multiple-instance spectral match filter (MTMI-SMF). The proposed methods address the abovementioned problems by directly considering the multiple-instance, imprecisely labeled data set. They learn a dictionary of target signatures that optimizes detection against a background using the adaptive cosine estimator (ACE) and spectral match filter (SMF). Experiments were conducted to test the proposed algorithms using a simulated hyperspectral data set, the MUUFL Gulfport hyperspectral data set collected over the University of Southern Mississippi–Gulfpark Campus, and the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) hyperspectral data set collected over Santa Barbara County, CA, USA. Both simulated and real hyperspectral target detection experiments show that the proposed algorithms are effective at learning target signatures and performing target detection.**

***Index Terms*— Adaptive cosine estimator (ACE), hyperspectral, multiple instance, multiple target, spectral matched filter, target characterization, target detection.**

Manuscript received March 5, 2020; revised July 10, 2020, November 10, 2020, and January 8, 2021; accepted February 14, 2021. This work was supported in part by the Army Research Office through the U.S. Army RDECOM CERDEC NVESD under Grant W911NF-17-1-0213, in part by the DARPA Advanced Plant Technologies Grant (SENTINEL: SENsing Threats In Natural Environments Using Ligand-Receptor Modules), and in part by Harris Corporation. *(Susan Meerdink and James Bocinsky contributed equally to this work.) (Corresponding author: Susan Meerdink.)*

Susan Meerdink is with the Department of Geographical and Sustainability Sciences, The University of Iowa, Iowa, IA 52242 USA, also with the Department of Electrical and Computer Engineering, University of Florida, Gainesville, FL 32611 USA, and also with the Engineering School of Sustainable Infrastructure and Environment, University of Florida, Gainesville, FL 32611 USA (e-mail: susan-meerdink@uiowa.edu).

James Bocinsky, Alina Zare, Connor McCurley, and Daniel Shats are with the Department of Electrical and Computer Engineering, University of Florida, Gainesville, FL 32611 USA (e-mail: azare@ece.ufl.edu).

Nicholas Kroeger is with the Department of Computer and Information Science Engineering, University of Florida, Gainesville, FL 32611 USA.

Paul Gader is with the Engineering School of Sustainable Infrastructure and Environment, University of Florida, Gainesville, FL 32611 USA, and also with the Department of Computer and Information Science Engineering, University of Florida, Gainesville, FL 32611 USA.

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TGRS.2021.3060966.

Digital Object Identifier 10.1109/TGRS.2021.3060966

I. Introduction

HYPERSPECTRAL data are well suited for discrimination between a target and a background because the hundreds of narrowbands can be used to identify subtle spectral shifts caused by materials' differences in chemistry, physiology, and structure. However, each pixel measures the interaction of electromagnetic radiation with multiple surface constituents, regardless of spatial resolution [1]. Often targets of interest do not comprise a whole pixel resulting in a mixed-signal and require subpixel target detection. However, nearly, all hyperspectral target detectors rely on having an accurate target spectral signature in advance. Usually, target signatures are obtained from spectral libraries collected in either controlled laboratory settings, outdoor handheld spectrometer measurements, or pulled manually from a hyperspectral image. Spectra collected in controlled laboratory settings often do not match atmospheric or lighting conditions present in the hyperspectral imagery. Outdoor handheld spectrometer measurements can have similar issues if not collected during hyperspectral image collection. Both of these methods require knowing and measuring the majority of materials present in the imagery.

Often pulling spectra manually from a hyperspectral image is ideal because spectra capture the current environmental and lighting conditions. However, this method of developing spectral libraries can be complicated from many sources but can be mostly summarized into three categories. First, precise training labels for targets are often difficult to obtain for multiple scenarios. For example, a target's global positioning system (GPS) coordinates can have errors of several meters or pixels. As a result, the target's real location might be several meters from the measured coordinates. In addition, "pure" target pixels, or pixels that contain 100% of the target, are challenging to find on the landscape. Second, the number of training pixels for a target class is small compared to the nontarget training pixels. There are often only a few pixels for target training, especially for hyperspectral images covering a large spatial area. Finally, because each pixel is an interaction of multiple surface constituents, many targets are subpixel, and the targets' proportion is unknown. These complications can make obtaining the best-suited target signature difficult, which will ultimately drive the target detection algorithm's success. Spectral libraries developed using an outdoor handheld spectrometer would also inherently contain these complications.

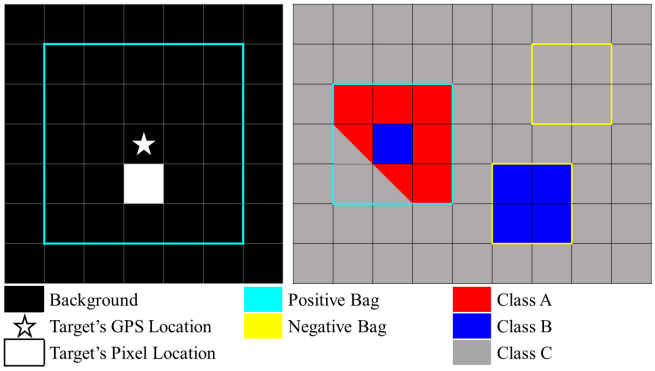

Fig. 1. Simplified positive and negative bagging examples with each square representing a pixel in an image. (Left) When a target's GPS location does not match the target's pixel location in the imagery. This example represents the bagging methodology used on the MUUFL Gulfport data set. (Right) When reference polygons are used for a class and may not have all pixels in a polygon representing the same class. This example describes the bagging methodology used on the AVIRIS Santa Barbara data set.

Multiple-instance learning (MIL) can overcome the need to have precise training labels [2], [3]. MIL only requires the labeling of positive and negative bags, which are groupings of pixels. Each bag may contain many pixels, but a bag is labeled positive if at least one of the pixels belongs to the target class (see Fig. 1). This framework alleviates the need to have accurate labels that are inherently challenging to collect. Since the introduction of MIL, numerous MIL algorithms have been proposed [4]–[7]. A pair of MIL algorithms known as the multiple-instance adaptive cosine estimator (MI-ACE) and the multiple-instance spectral match filter (MI-SMF) have been used with hyperspectral data and have shown competitive results with other algorithms in terms of single-target concept estimation and detection [5]. However, these algorithms only determine one signature for a target. Most targets contain enough within-class spectral variability that it is difficult to capture that variability with a single signature. For example, if the target is a tree species, then the spectral signatures can vary significantly between individuals due to differences in structure, biochemistry, or phenology. An urban target can contain a lot of spectral variability due to material diversity (e.g., concrete, asphalt, and paint) that also changes with age.

In this article, we propose the multitarget multiple-instance adaptive cosine/coherence estimator (MTMI-ACE) and multitarget multiple-instance spectral match filter (MTMI-SMF) algorithms to extend the MIL framework and learn multiple-target signatures compared to a single-target signature [8]. The objective of the MTMI-ACE and MTMI-SMF algorithms is to learn a dictionary of target representations, focusing on maximizing the detection of those targets against a background. Our overarching aim is to demonstrate the improvements and advantages of these algorithms for hyperspectral target detection.

A. Related Work

Two of the earliest proposed MIL algorithms are the diverse density algorithm [3] and the expectation–maximization with the diverse density (EM-DD) algorithm [9]. These algorithms learn a target concept closest to the intersection of the largest number of positive bag instances while being as far from any negative bag instances. These algorithms learn a single-target signature and use the Euclidean distance to measure the similarity between instances. These two methods introduced the noisy-OR model that many later MIL algorithms relied upon [10]. Although foundational, these two methods were not designed with hyperspectral detection in mind, and other methods soon outperformed these for hyperspectral applications.

Another family of algorithms that belong to the MIL framework is the functions of multiple-instance (FUMI) algorithms. The original FUMI algorithm [11] extends the approach of the sparsity promoting iterated constrained endmember (SPICE) algorithm [12]. SPICE is an unsupervised algorithm that learns the subpixel proportions and endmembers of an unlabeled data set for unmixing imagery. FUMI extends the SPICE algorithm by using the labeled data to learn a target's signature as well as nontarget endmembers. Variations of FUMI have been developed, such as convex FUMI (cFUMI) [13] and extended FUMI (eFUMI) [13]. While cFUMI assumes that the exact target locations in a training image are known, the eFUMI algorithm needs only an approximate knowledge of target locations in training data. eFUMI learns multiple-target signatures by determining the convex combinations of target and nontarget signatures using an expectation–maximization approach. Through this approach, the algorithm focuses on learning the discriminative features between different target types, resulting in better target characterization and discrimination. However, eFUMI estimates signatures and does not discriminate prototypes. In addition, the number of targets and background signatures is needed to find target signatures that require domain knowledge of the data set.

The MI-ACE algorithm estimates a single-target signature that optimizes the widely used adaptive cosine estimator (ACE) subpixel target detector on a training data set with multiple-instance-style imprecise labels [5]. The MI-SMF algorithm does the same but using the spectral match filter (SMF) subpixel target detector. However, these algorithms assume that a target's spectral variability can be captured with a single-target signature [5].

Another recent MIL algorithm is the multiple-instance hybrid estimator (MI-HE) algorithm. The MI-HE algorithm learns multiple targets and background signatures that maximize the probability that positive bags are labeled positive and negative bags are labeled negative [6]. The objective function is simplified by only needing to maximize a single instance from each positive bag. MI-HE makes the noisy-OR model more flexible by implementing a hyperparameter adjustable generalized mean to vary the operation between a min and max operation. In addition, the algorithm solves for a sparsity vector to support dictionary element diversity. The MI-HE algorithm also determines multiple-target signatures by using a data mixing model and optimizing the hybrid subpixel detector's response within an MIL framework. The algorithm iterates between estimating a set of representative target and nontarget signatures and solving a sparse unmixing problem.

Although this algorithm performs highly competitively with other MIL algorithms, the algorithm has a complex number of hyperparameters and takes significantly longer to train than other MIL algorithms.

The most recent algorithm added to the MIL framework and hyperspectral target detection is the multiple-instance learning for multiple diverse characterizations adaptive cosine estimator (MILMD-ACE) and multiple-instance learning for multiple diverse characterizations spectral match filter (MILMD-SMF) algorithms [4]. This algorithm learns multiple-target concepts by maximizing the collective dictionary's detection statistic across the positive bags while minimizing the detection across negative instances. A unique aspect of this algorithm is the assumption that each positive bag is constructed with multiple-target types, deviating from the traditional MIL framework. This approach is useful if two targets exist in every positive bag but not in negative bags.

The algorithms mentioned above have shown success at determining subpixel target signatures for target detection. However, every algorithm has limitations. Our proposed algorithms address four of these common limitations. First, the MTMI algorithms maximize the dictionary's detection statistic with respect to expected target signatures provided from the training data. Second, the MTMI algorithms do not assume a set number of target signatures, but instead, learn the appropriate number of target signatures. The algorithms mentioned above require that the number of targets, and often the number of background signatures, is known, which can be difficult without domain knowledge. Third, the uniqueness of returned target signatures is adjustable through a hyperparameter. The uniqueness hyperparameter allows for the multiple-target signatures to be more similar or more distinct depending on the data set's characteristics. Finally, the introduced algorithms are trained using a dot product in a transformed data space, which is fundamentally quick to compute, resulting in an efficient algorithm.

B. Adaptive Cosine Estimator and Spectral Match Filter

The ACE and SMF are detection statistics often used for hyperspectral subpixel target detection [14]–[17]. ACE performs detection by solely considering spectral shape, whereas SMF also considers magnitude. Assuming target signature $\mathbf{s}$ and unknown instance $\mathbf{x}$, the ACE and SMF detectors [5] can be written as

$$D_{ACE}(\mathbf{x},\mathbf{s}) = \frac{\mathbf{s}^T\mathbf{W}^T}{||\mathbf{s}^T\mathbf{W}^T||}\frac{\mathbf{W}(\mathbf{x}-\boldsymbol{\mu_b})}{||\mathbf{W}(\mathbf{x}-\boldsymbol{\mu_b})||} \tag{1}$$

$$D_{SMF}(\mathbf{x},\mathbf{s}) = \frac{\mathbf{s}^T\mathbf{W}^T}{||\mathbf{s}^T\mathbf{W}^T||}\mathbf{W}(\mathbf{x}-\boldsymbol{\mu_b}) \tag{2}$$

$$\mathbf{W} = \mathbf{E}^{-\frac{1}{2}}\mathbf{U}^T \tag{3}$$

where D is the detection response of the given detector, either ACE or SMF. In (1) and (2), the detection statistics are shown as the dot product between the target signature and the unknown sample in a whitened coordinate space. Here, a translation is done using the mean of the background, $\boldsymbol{\mu_b}$, followed by a transformation using the eigenvectors, $\mathbf{U}$, and eigenvalues, $\mathbf{E}$, of the background covariance. When choosing which statistic is appropriate for a data set, it is essential to determine whether the spectral magnitude is necessary for a target's discrimination from the background. When a target has a similar spectral shape to the background but is brighter or darker (e.g., lower or higher reflectance), the SMF detection statistic would be appropriate.

II. Methods

A. Multiple Target MIL

Following the MIL framework, this algorithm assumes that the data are grouped into bags with bag level labels [2]. With this, let $\mathbf{X} = \{\mathbf{x}_1, \ldots, \mathbf{x}_N\}$ be training data with each sample and $\mathbf{x}_i$ be a vector with dimensionality D. The data are grouped into J bags $\mathbf{B} = \{\mathbf{B}_1, \ldots, \mathbf{B}_J\}$ with labels, $L = \{L_1, \ldots, L_J\}$, where $L_j \in \{0, 1\}$. A bag is considered positive, $\mathbf{B_j^+}$, with label $L_j = 1$, if there exists at least one instance, $\mathbf{x}_i$, in bag j that is from the target class. In addition, a bag is considered negative, $\mathbf{B_j^-}$, with label $L_j = 0$, if all instances in bag j are from the background class. The number of instances in both positive and negative bags is variable. In this algorithm, it is assumed that each positive bag is only required to contain one instance of a target and is not guaranteed to contain multiple-target types in each bag. With this assumption, only a single instance from each positive bag should contribute to the objective function calculation. The formulation of the objective function assumes that different target types will exist across different positive bags.

In this work, the original MI-ACE algorithm's objective function has been extended to include multiple-target signatures. K target signatures are in the dictionary, $\mathbf{S}$ (4). $\mathbf{S}$ is a matrix containing vectors s_i, where each s_i is a target signature being estimated

$$\mathbf{S} = \begin{bmatrix}\mathbf{s}_1, & \mathbf{s}_2, & \ldots, & \mathbf{s}_K\end{bmatrix}. \tag{4}$$

The goal of MTMI-ACE and MTMI-SMF is to estimate the set of target signatures that maximize the detection statistic for the target instances in each positive bag and minimize the detection statistic overall negative instances. This is accomplished by maximizing the following objective function:

$$\begin{aligned}&\max_{\mathbf{S}} \quad M_+(\mathbf{B},\mathbf{S}) - M_-(\mathbf{B},\mathbf{S}) - M_u(\mathbf{B},\mathbf{S})\\ &\text{s.t. } D(\mathbf{s}_k,\mathbf{s}_k) = 1.\end{aligned} \tag{5}$$

The objective function is comprised of three terms and a constraint (5). The first term (6) is the average detection statistic of the selected instances from each positive bag. This term also contains the target signatures to be learned. By including the max operation in the first term, each target signature will learn a particular target type. This operation is introduced because it is assumed that every positive bag does not contain every target type. The dictionary of learned target signatures will maximize the objective function by individually maximizing a subset of positive bags

$$M_+(\mathbf{B},\mathbf{S}) = \quad \frac{1}{N^+}\sum_{j:L_j=1}\max_{s_k\in\mathbf{S}}(D(\mathbf{x}^*_{j,k},\mathbf{s}_k)) \tag{6}$$

where N^+ is the number of positive bags and $\mathbf{x}^*_{j,k}$ is the selected instance from the positive bag $\mathbf{B_j^+}$ that is most likely

a target instance in the bag. The selected instance $\mathbf{x}^*_{j,k}$ is identified as the point in bag j with the maximum detection statistic given a target signature, $\mathbf{s}_k$

$$\mathbf{x}^*_{j,k} = \underset{x_j \in \mathbf{B}_\mathbf{j}^+}{\arg\max}\, D(\mathbf{x}_j, \mathbf{s}_k). \tag{7}$$

The second term (8) is the average detection statistic of the negative bag's instances and the target signatures to be learned. This term discourages learning any target signature that is similar to the background. By including the second summation, each negative bag has an equal weight in the objective function. Here, N^- is the number of negative bags, and N_j^- is the number of instances in negative bag j. Finally, an outer sum across K targets is included to ensure that each target concept contributes equally to the objective function

$$M_-(\mathbf{B}, \mathbf{S}) = \frac{1}{K}\sum_{k=1}^{K}\frac{1}{N^-}\sum_{j:L_j=0}\frac{1}{N_j^-}\sum_{x_i \in B_j^-} D(\mathbf{x}_i, \mathbf{s}_k). \tag{8}$$

The third term, known as the uniqueness term (9), is introduced to encourage the algorithm to learn distinct target signatures by penalizing the algorithm for similar target signatures. The uniqueness term has a hyperparameter weight, α, which changes how similar the target signatures can be. The larger the weight, the more the algorithm will be encouraged to learn different signatures

$$M_u(\mathbf{B}, \mathbf{S}) = \frac{\alpha}{\binom{K}{2}}\sum_{k=1}^{K-1}\sum_{l=k+1}^{K} D(\mathbf{s}_k, \mathbf{s}_l). \tag{9}$$

Finally, the constraint, $D(\mathbf{s}_k, \mathbf{s}_k) = 1$, is included to restrict the algorithm from maximizing the objective function by learning target signatures of erroneously large magnitude.

The MTMI-ACE and MTMI-SMF algorithms have two primary steps: initialization and optimization. The initialization process uses a greedy approach along with clustering to aid in computation complexity and representative target diversification. The optimization process learns the number of needed signatures to describe the target class optimally while generalizing the signatures considering all of the positive bags. These steps are described in the following. The algorithm pseudocode is provided in Algorithm 1.

B. Target Dictionary Initialization

To reduce the initialization process's computation complexity, the K-means clustering algorithm [18] is used to aid in target concept initialization. K-means is used by clustering all of the data, regardless of bag structure, into C clusters. The cluster centers that maximize the objective function in (5) are iteratively selected until K signatures have been added to the target dictionary. In this process, the algorithm is greedily selecting the next best target concept for initialization. As long as the number of clusters, C, and the number of iterations i, remains small, the K-means approach will have a lower computational cost than searching through all of the positive instances. Using K-means, the algorithm only needs to search through C candidates instead of N^+ candidates to initialize a target signature.

Algorithm 1 MTMI-ACE and MTMI-SMF

```
Compute μ_b, E, and U
Subtract μ_b and whiten X, Equation (3)
if ACE then
    Normalize X
end if
K-Means cluster instances in positive bags
Greedily initialize S, as the cluster centers that maximize
    the objective function, Equation (5)
repeat
    Update the set of x*_{j,k} for each s_k, Equation (7)
    Determine indicators δ+_{j,k} for each ŝ_k, Equation (14)
    for k = 1 to K do
        if sum(δ+_k) = 0 then
            Remove s_k from S
        else
            Update ŝ_k, Equation (10)
        end if
    end for
until Stopping criterion reached
Normalize and de-whiten ŝ_k, Equation (15)
return all optimized s_k as S
```

C. Target Dictionary Optimization

The target concept update equation is calculated for each target concept. Each target concept is updated in each optimization iteration until the stopping criteria are reached. In this manner, the target dictionary is optimized by optimizing the set of target concepts through each iteration (see Algorithm 1 for details). This solution is closed form. The optimization will produce the same result given a fixed set of positive bags, negative bags, and initialized target signatures [5]. This assumes that no cycles occur during optimization, and the max number of iterations is not reached. The update equation to perform target concept optimization is derived by maximizing the objective function for the target signatures, $\hat{\hat{\mathbf{s}}}_k$. The derivation of the update equation is included in Appendix A. The resulting update equation is

$$\hat{\hat{\mathbf{s}}}_k = \frac{\hat{\hat{\mathbf{t}}}}{\left\|\hat{\hat{\mathbf{t}}}\right\|} \tag{10}$$

where

$$\hat{\hat{\mathbf{t}}} = \frac{1}{N_k^+}\sum_{j:L_j=1}\delta^+_{j,k}\hat{\hat{\mathbf{x}}}^*_{j,k} \tag{11}$$

$$-\frac{1}{N^-}\sum_{j:L_j=0}\frac{1}{N_j^-}\sum_{\hat{\hat{x}}_i \in B_j^-}\hat{\hat{\mathbf{x}}}_i - \frac{\alpha}{(K-1)}\sum_{l,l\neq k}\hat{\hat{\mathbf{s}}}_l \tag{12}$$

$$\hat{\hat{\mathbf{x}}} = \frac{\mathbf{W}(\mathbf{x}-\boldsymbol{\mu}_b)}{||\mathbf{W}(\mathbf{x}-\boldsymbol{\mu}_b)||} \tag{13}$$

where

$$\delta^+_{j,k} = \begin{cases} 1 & \text{if } D\left(\mathbf{x}^*_{j,k}, \mathbf{s}_k\right) > D\left(\mathbf{x}^*_{j,l}, \mathbf{s}_l\right) \quad \forall l \neq k \\ 0 & \text{otherwise.} \end{cases} \tag{14}$$

This update equation is interpretable. The first term is the average of the selected positive instances assigned to the same target type as $\hat{\mathbf{s}}_k$. The selected positive instances are determined by computing the positive bag identifiers δ^+ for each target signature. The second term is the average of all negative instances and can be precomputed. The third term is the average of all other target signatures. This result is also intuitive. A learned target signature will be an average of the target type being optimized, dissimilar from the background, and pushed away from the other target signatures to encourage target signature uniqueness. Depending on the application, α is configurable to allow for more distinct or similar target signatures. If α is increased, more distinct target signatures are allowed. Finally, all target signatures are optimized simultaneously until the bag identifiers, δ^+, and each target signature's bag representatives, $\hat{\mathbf{x}}_k^*$, remain the same across subsequent iterations or the max number of iterations is reached. Finally, the target concepts are normalized and dewhitened using

$$\mathbf{s}_k = \frac{\mathbf{t}_k}{||\mathbf{t}_k||}, \mathbf{t}_k = \mathbf{W}^{-1}\hat{\mathbf{s}}_k. \tag{15}$$

In practice, this optimization is fast and per iteration has a computational complexity on the order of $\mathcal{O}(KDN^+)$, where K is the number of target concepts, D is the dimensionality of the data, and N^+ is the number of positive instances in the training data.

D. Learning Number of Target Concepts

During optimization, the number of target concepts is estimated iteratively by removing unnecessary target signatures. Target signatures are removed by observing the bag identifiers, in (14), for each of the k target signatures during the iterations of optimization. If all j positive identifiers for the kth target are 0, then the kth target signature will be dropped from the set of target signatures $\mathbf{S}$. Namely, the kth signature will be dropped when the detection similarity between the kth target signature and all j positive bag representatives, $\hat{\mathbf{x}}_k^*$, is smaller than all other target signatures' detection similarities to their corresponding j bag representatives. MTMI-ACE and MTMI-SMF remove the need for domain-specific knowledge for how many target signatures may exist while still being adjustable by changing the value of α to encourage more or less target signature uniqueness, where a lower α value can potentially lead to more target signatures to be estimated since signatures are allowed to be more similar.

III. Experiments

In the following, MTMI-ACE and MTMI-SMF are evaluated and compared to several MIL framework methods using simulated data and two real hyperspectral data sets. The simulated data experiments illustrate the properties of MTMI-ACE and MTMI-SMF, providing insight into how and when the methods are effective. The hyperspectral data sets are included to illustrate how MTMI-ACE and MTMI-SMF perform in real-world scenarios with two different bagging methods (see Fig. 1). In this section, we will compare our proposed algorithms with other MIL algorithms from the literature, including multiple-instance learning for multiple diverse characterizations (MILMD) [4], MI-ACE [5], MI-SMF [5], MI-HE [6], and eFUMI [13]. These algorithms were selected due to their relevance, prevalence in the literature, or their recent development.

A. Simulated Data

1) Experimental Data Set: Simulated data were generated from five spectra selected from the ECOSTRESS Spectral Library [19], formally known as the ASTER Spectral Library [20]. Those five spectra were from the rock class (basalt, pyroxenite, verde antique, phyllite, and slate) and had 211 bands ranging from 400 to 2500 nm. The simulated data set was generated following steps and code detailed in [13] and [21]. In this experiment, the simulated data set was created using two target signatures: basalt and verde antique. All other spectra (three classes) were used as background. The parameters used to develop the simulated data set were 10 positive bags, 20 negative bags, 500 points in each bag, 250 target points in each positive bag, 0.3 mean target proportion, and 20 signal-to-noise ratio. Two simulated data sets were generated using those parameters, with one designated for training and the other for testing. This experimental design was repeated for ten iterations. For these simulated experiments, algorithms were evaluated on these data using the area under the receiver operating characteristic curve (AUC) in which the value reported is the AUC up to false alarm rate (FAR) of 1×10^{-3}. The results shown are the average AUC and standard deviation from these ten iterations.

2) Effects of K: The K parameter in the multitarget multiple-instance algorithm controls the number of initial target signatures. Too few targets may not capture the spectral variability in the target's pixels. A large K parameter slows down the computation. Often the user does not know what the appropriate K value is for their data set. In this experiment, all MTMI-ACE and MTMI-SMF parameters were kept constant ($\alpha = 0.5$), whereas the K value was changed from 2, 4, and 8 to 16.

With changing K, the developed target signatures are similar, and the total returned number of targets was consistent (see Fig. 2). Some iterations did show a larger number of returned targets with K equal to 8 and 16, but for most iterations, the total returned number of targets was 2, indicating the stability in estimating the number of target signatures. The target detection performance shown in the receiver operating characteristic (ROC) curves and AUC values does not show much variability with changing K.

Another benefit of the MTMI algorithms is the fact that the returned target signatures are interpretable. The returned signatures are similar to the original signatures, but with a few key differences (see Fig. 2). The MTMI algorithms maximize differences between the targets and the background, so spectral features that are found beneficial for separation are exaggerated. For example, the peak in the verde antique spectra between 1.5 and 2.3 μm is a unique feature of this class. The returned MTMI target signature reflects this difference showing elevated reflectance values in this spectral range. These returned target signatures could link back to

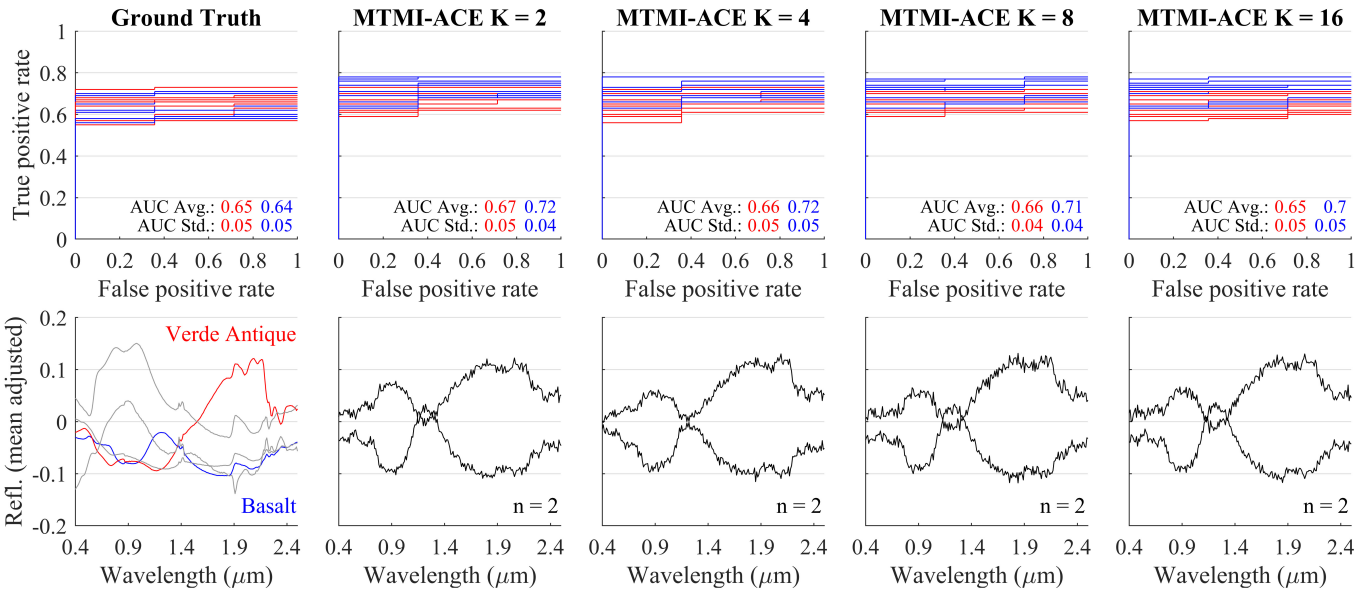

Fig. 2. MTMI-ACE ROC curves (Top) across the ten data set simulations and changing K values and target signatures from (bottom) highest performing iteration. The first column shows the target detection results using the original signatures, after whitening, with the two targets in red and blue with background signatures in gray. The false-positive rate scale is shown as 1×10^{-3}.

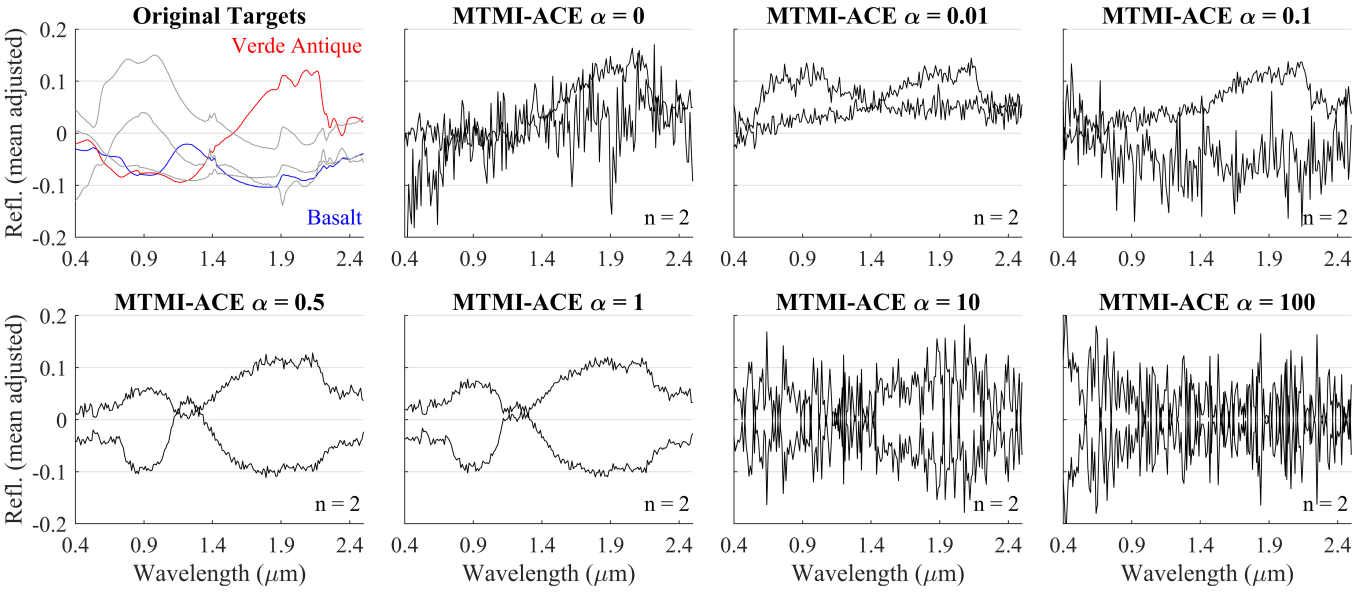

Fig. 3. MTMI-ACE target signatures from the highest performing iteration across different α values. The first panel shows the original signatures, after whitening, with the two targets in red and blue with background signatures in gray.

specific biochemistry or physical properties of the target class. It is often useful in remote sensing applications to know what wavelengths make a target class different from background classes.

3) Effects of α: The α parameter controls the similarity or diversity of target signatures obtained from MTMI-ACE and MTMI-SMF. A smaller α will allow for more similar target signatures, whereas a larger α will force target signatures to be more diverse. Changing the α parameter can significantly affect returned target signatures depending on the spectral variability in the data set. In this experiment, all MTMI-ACE and MTMI-SMF parameters are kept constant ($K = 2$), whereas the α value is changed from 0, 0.01, 0.1, 0.5, 1, and 10 to 100.

For this data set, the best α for returning interpretable target signatures was 0.5 and 1 (see Fig. 3). After a value of 1, the signatures become noisy and do not reflect any features from the original signatures. As mentioned above, the α parameter controls the amount of diversity between developed signatures. α of 0 does not use the third term of the objective function and does not encourage diversity among the multiple targets. As α increases, the target signatures become

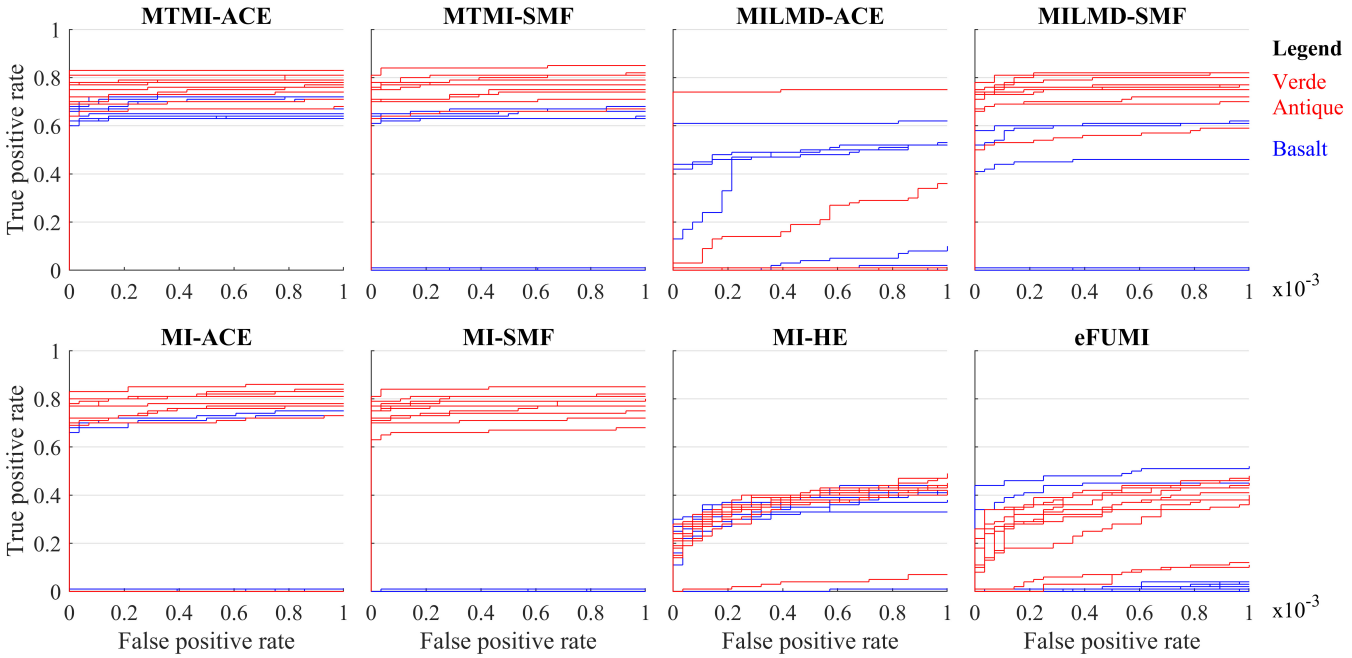

Fig. 4. Two target simulated data set ROC curves for each of the ten data set simulations across the nine target detection algorithms. The false-positive rate scale is shown as 1×10^{-3}. At this range, some of the iterations had a zero true-positive rate.

increasingly different until a threshold is met, and signatures resemble noise. However, the target detection performance shown in the ROC curves and AUC values does not show much variability with changing α. It is essential to examine the target signatures to ensure that they are making physical sense and retain confidence in detection results.

In this experiment, α values less than 1 and greater than 0 resulted in optimal performance and interpretability. The appropriate α is dependent on the spectral differences between target and background classes. A smaller α would ensure that developed target signatures would reflect patterns in the data set for targets that are similar spectrally or exhibit significant overlap between targets. While not seen in this data set, larger α can reduce the number of developed target signatures to maximize diversity. If multiple targets are needed for accurate detection, smaller α will ensure that multiple-target signatures will be developed because the signatures are allowed to be more similar.

4) Target Detection Results: Using the two-target simulated data set, we compared MTMI with six other multiple-instance target detection algorithms in the literature. For this experiment, MTMI-ACE and MTMI-SMF parameters were fixed with $\alpha = 1$ and $K = 4$. Although we know that this data set only has two targets, four potential targets were allowed to demonstrate that MTMI algorithms return the appropriate number of targets. This flexibility eases the restrictions on the user-defined K parameter. Other algorithm parameters are found in Appendix B.

Fig. 4 shows the ROC curves for the two classes across algorithms, and Table I shows the average normalized AUC (NAUC) and standard deviations from the ten data set simulations. This experiment's results are more understood after considering the original target signatures spectral separability, as shown in panel 1 of Fig. 2. Compared to the other spectra, verde antique has a unique spectrum specifically in the 1.5–2.3-μm spectral range, whereas basalt shares more similar features. The more distinct the target spectrum, the easier it is for the target detection algorithms to determine the appropriate target signature. All the algorithms were able to determine an appropriate signature for verde antique, resulting in higher accuracy for that target class. However, the main deviation among algorithm performance happened when the algorithms tried to determine the appropriate signature for basalt and could not capture the spectral signature due to the spectral similarity to the other materials.

MTMI-ACE performs well for detecting basalt and verde antique compared to many of the other algorithms. Notably, the NAUC values are good for both target types compared to MI-ACE, where one target performs significantly better.

TABLE I

AVERAGE AUC RESULTS, WITH STANDARD DEVIATION IN PARENTHESES, FOR THE SIMULATED DATA SET WITH TWO TARGETS (BASALT AND VERDE ANTIQUE). THE BEST RESULTS ARE SHOWN IN BOLD, AND THE SECOND-BEST RESULTS ARE UNDERLINED

Algorithm	Basalt	Verde Antique
MTMI-ACE	**0.652 (0.019)**	**0.784 (0.056)**
MTMI-SMF	0.318 (0.333)	0.741 (0.054
MILMD-ACE	0.164 (0.237)	0.078 (0.233)
MILMD-SMF	0.155 (0.250)	0.712 (0.077)
MI-ACE	0.138 (0.287)	0.608 (0.323)
MI-SMF	0.001 (0.002)	0.745 (0.052)
MI-HE	0.223 (0.087)	0.211 (0.077)
eFUMI	0.080 (0.167)	0.167 (0.105)

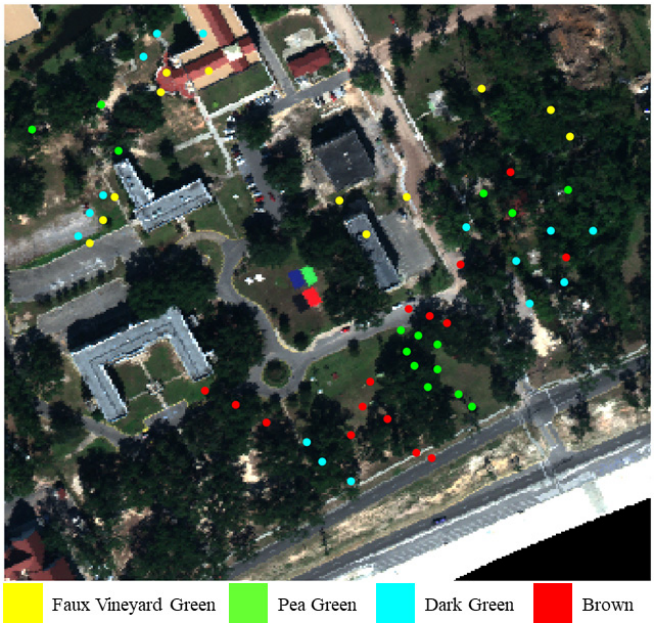

Fig. 5. MUUFL Gulfport data set red/green/blue image and the 57 target locations.

Looking at the NAUC, it becomes clear that the ability to determine multiple targets leads to an increase in detection compared to MI-ACE and MI-SMF, which can only return one target signature. These algorithms could detect the verde antique target with high accuracy, but not the basalt target because only a verde antique target was developed. MI-ACE returned target signatures similar to verde antique because it was distinct from the other targets, while basalt was too similar. Another recently released multiple-target MIL algorithm, MILMD-SMF, also detected the two targets well but returned target signatures that favored the verde antique target leading to a decrease mean NAUC for basalt target detection. MILMD-ACE shows large variability across iterations. These algorithms were developed with the assumption that both targets exist in each positive bag instead of the traditional approach of a single target, which is how this data set was created. The MI-HE algorithm showed high consistency across random iterations, whereas many other algorithms showed high variability in results and learned target signatures.

B. MUUFL Gulfport Data—Single Target

1) Experimental Data Set: The MUUFL Gulfport Hyperspectral data set was used to perform experiments on hyperspectral data that contained subpixel targets (see Fig. 1). This data set was collected over the University of Southern Mississippi–Gulfpark Campus with 1-m spatial resolution and 72 bands ranging between 367.7 and 1043.4 nm. This experiment uses two images (flights 1 and 3) that cover the same spatial area but were flown approximately 10 min apart. These images contain 57 human-made targets made of cloth panels in four different colors: brown (15 panels), dark green (15 panels), faux vineyard green (12 panels), and pea green (15 panels). The targets' spatial location is shown as scattered points over a red/green/blue image of the scene in Fig. 5. This data set is a very challenging target detection task as trees partially or fully occlude many of the targets. Furthermore, the targets vary in size and could be 0.25, 1, and 9 m^2. Thus, a target that has 0.25 m^2 covers at most a 0.25 proportion of the pixel signature if the pixel falls directly on the target. However, many of these targets straddle multiple pixels and are occluded, resulting in a highly mixed, subpixel target detection task. A bag included all pixels in a 5×5 rectangular region around each ground-truth point. The GPS device used to record the ground-truth locations had a maximum of 5-m accuracy, which would result in a 10×10 rectangular region. However, the GPS accuracy on the day of collection was around 2–3 m, so the size of 5×5 was chosen. The remaining area that did not contain the target class was grouped into one big negative bag. Two iterations were run in which flight 1 was selected for training and flight 3 for testing, and vice versa. The target types were iterated through, and in each iteration, a single-target type was selected as the positive bags and all other image pixels were selected as a negative bag. Thus, there are 12–15 positive bags in each training set in this experiment. For this experiment, algorithms were evaluated on this data using the NAUC in which the area was normalized out to an FAR of 1×10^{-3} false alarms/m^2 [21].

2) Target Detection Results: As mentioned above, the MUUFL Gulfport data set represents a challenging subpixel detection environment due to many trees covering targets. In addition, these targets are spectrally similar to the background because they are cotton fabric (which exhibit vegetation spectral features) and are similar colors (e.g., shades of green and brown). All these data set characteristics result in lower detection results compared to other data sets. However, it does provide the opportunity to test the capabilities of the detection algorithms in a challenging scenario. The MUUFL Gulfport data set NAUC results are shown in Table II for the two training and testing splits across flight lines, and Fig. 6 shows the four target ROC curves for a subset of algorithms.

The MI-ACE algorithm generally performed the best, demonstrating that a single target was sufficient for accurate detection in this experiment designed to detect a single target. However, MTMI-ACE ranked among the highest performing across the training/testing split, demonstrating that the algorithm's multiple-target version maintains similar accuracy. The MILMD-SMF algorithm, most similar to our proposed algorithm, did have comparable results to the MTMI algorithm. However, the MILMD-ACE algorithm could not determine an appropriate target signature for many of the classes. MI-HE and eFUMI algorithms also performed comparably, but slightly decreased accuracies. However, these algorithms took much longer to computationally execute and have more user-defined parameters that can impact the final results.

C. MUUFL Gulfport Data—Multiple Targets

1) Experimental Data Set: This experiment also leveraged the MUUFL Gulfport data set, but instead of running target detection for a single class, it was run with all classes (see Table III). A bag included all pixels in a 5×5 rectangular

TABLE II

RESULTS FOR THE MUUFL GULFPORT DATA SET FOR THE SINGLE TARGET EXPERIMENT SHOWING THE NAUC IN WHICH THE AREA WAS NORMALIZED OUT TO AN FAR OF 1×10^{-3} FALSE ALARMS/M^2. THE BEST RESULTS ARE SHOWN IN BOLD, AND THE SECOND-BEST RESULTS ARE UNDERLINED

	Train on flight 1; Test on flight 3				Train on flight 3; Test on flight 1			
	Brown	Dark Green	Faux V. Green	Pea Green	Brown	Dark Green	Faux V. Green	Pea Green
MTMI-ACE	**0.496**	0.391	**0.659**	0.300	0.772	0.518	0.634	**0.416**
MTMI-SMF	0.452	0.365	0.472	0.267	0.667	0.417	0.492	0.396
MILMD-ACE	0.391	0.013	0.000	0.000	0.742	0.000	0.000	0.267
MILMD-SMF	0.430	0.343	0.246	0.264	0.638	0.291	0.307	0.350
MI-ACE	0.486	**0.392**	0.643	**0.301**	**0.777**	**0.519**	**0.652**	0.398
MI-SMF	0.452	0.364	0.468	0.267	0.668	0.406	0.510	0.396
MI-HE	0.433	0.379	0.104	0.267	0.710	0.360	0.111	0.266
eFUMI	0.383	0.360	0.106	0.238	0.404	0.497	0.185	0.387

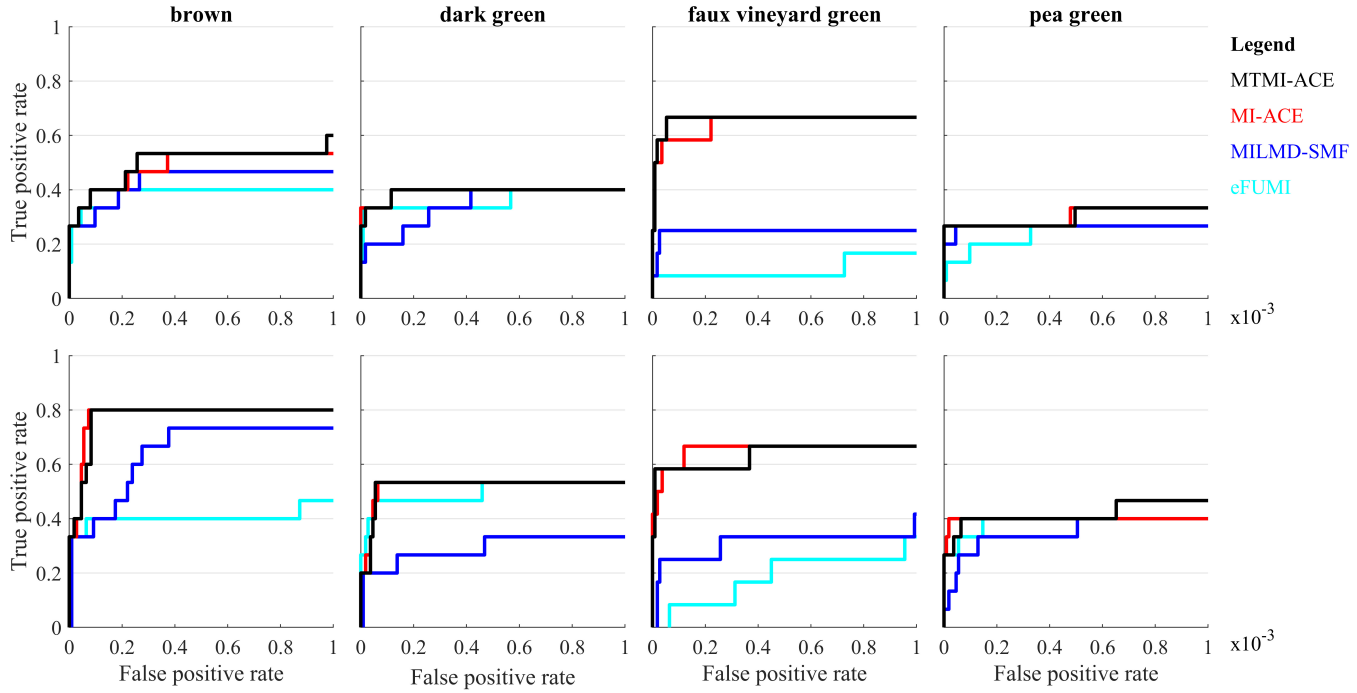

Fig. 6. MUUFL Gulfport data set ROC curves for a subset of algorithms. (Top row) Results from training with flight line 1 and testing on flight line 3. (Bottom row) Results from training with flight line 3 and testing on flight line 1. The false-positive rate scale is shown as 1×10^{-3}.

region around each ground-truth point. All classes were designated as positive bags, while the remaining area that did not contain the target class was grouped into one big negative bag. Thus, there are 57 positive bags in each training set and one negative bag in this experiment. Two iterations were run in which flight 1 was selected for training and flight 3 for testing, and vice versa. For this experiment, algorithms were evaluated on this data using the NAUC in which the area was normalized out to an FAR of 1×10^{-3} false alarms/m^2 [21].

2) Target Detection Results: Single-target algorithms, such as MI-ACE, had higher accuracy for specific targets such as brown and dark green but suffered on detection of the other two targets. Multiple-target algorithms, such as MTMI-SMF, had more consistent accuracy across all targets because they can return multiple targets capturing more of the spectral variability present. MILMD-ACE is another multiple-target algorithm that performed better across all targets compared to single-target algorithms, but it did determine a larger number of target signatures compared to MTMI algorithms.

D. AVIRIS Santa Barbara Data

1) Experimental Data Set: The Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) Santa Barbara data set was used to perform experiments on real hyperspectral data that contained training data in the form of polygons that may not contain all pure pixels (see Fig. 1). The imagery was collected with the AVIRIS sensor as part of the HyspIRI Airborne Preparatory Campaign on April 16, 2014 [22]. AVIRIS measures 224 bands of radiance between 360 and 2500 nm with a full-width at half-maximum of 10 nm [23]. This experiment uses a spatial subset of imagery from the Santa Barbara flight box, which includes ten of the 11 flight lines acquired with a 35° northeast–southwest orientation and 18-m spatial resolution (see Fig. 7). These ten flight lines cover a

TABLE III

RESULTS FOR THE MUUFL GULFPORT DATA SET FOR THE MULTIPLE TARGET EXPERIMENT SHOWING THE NAUC IN WHICH THE AREA WAS NORMALIZED OUT TO AN FAR OF 1×10^{-3} FALSE ALARMS/M^2. THE BEST RESULTS ARE SHOWN IN BOLD, AND THE SECOND-BEST RESULTS ARE UNDERLINED. THE S COLUMN DESIGNATES HOW MANY TARGET SIGNATURES WERE DETERMINED BY AN ALGORITHM

	Train on flight 1; Test on flight 3					Train on flight 3; Test on flight 1				
	Brown	Dark Green	Faux V. Green	Pea Green	S	Brown	Dark Green	Faux V. Green	Pea Green	S
MTMI-ACE	0.389	0.371	0.040	0.047	5	<u>0.772</u>	0.450	0.000	0.032	5
MTMI-SMF	0.374	0.344	**0.323**	**0.251**	5	0.580	0.341	0.000	<u>0.262</u>	5
MILMD-ACE	0.355	0.241	<u>0.246</u>	<u>0.224</u>	8	0.613	0.386	0.000	0.056	8
MILMD-SMF	0.000	0.047	0.110	0.021	8	0.000	0.000	0.000	0.000	8
MI-ACE	**0.518**	**0.422**	0.000	0.000	1	**0.773**	<u>0.495</u>	0.000	0.000	1
MI-SMF	<u>0.457</u>	0.367	0.000	0.000	1	0.665	0.415	0.000	0.000	1
MI-HE	0.000	0.000	0.004	0.057	1	0.723	**0.515**	0.000	0.049	1
eFUMI	0.388	<u>0.373</u>	0.000	0.000	3	0.397	0.410	0.000	**0.290**	3

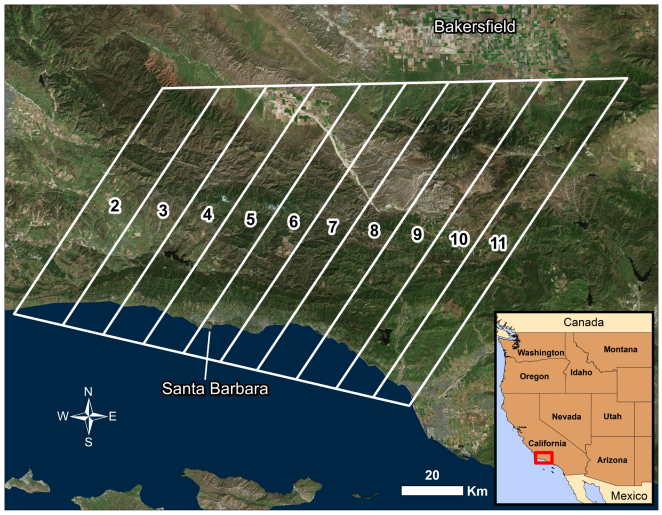

Fig. 7. Santa Barbara flight box and the HyspIRI Airborne Preparatory campaign flight lines used [19].

diverse landscape that is approximately 12 980 km^2. For more information about imagery preprocessing and development of the training data set, please refer to [19]. The original training data were collected to classify plant species, but this study grouped plant species into their plant functional types (PFTs). This resulted in nine classes of PFTs: annual herb (AH), deciduous broadleaf tree (DBT), deciduous shrub (DS), evergreen broadleaf tree (EBT), evergreen broadleaf shrub (EBS), evergreen needleleaf shrub (ENS), evergreen needleleaf tree (ENT), rock/soil (RS), and urban (URB). The training data set comprised of spatial polygons designating where on the landscape 'pure' patches of species existed. These locations were identified in the field and using AVIRIS and National Agriculture Imagery Program (NAIP) imagery. However, it is often challenging to find 100% pure patches of species on the landscape, so patches having greater than 75% single species composition were recorded.

Each polygon was treated as a bag, and often, not all of the polygon's pixels belonged to the target class. Classes had a different number of bags, and each bag size was variable (Appendix B-D). Often in traditional classifiers, these nontarget pixels will add too much variability and confusion. Data were split into training and testing using fivefold cross validation. Iterating through each class, all polygons matching that class label were selected as positive bags, while all other polygons were chosen as negative bags. For MTMI-ACE/MTMI-SMF parameters, the background mean and covariance were calculated from all pixels in the reference library, the K was 15, and α was 1. For this experiment, algorithms were evaluated on these data using the AUC in which the value reported is the AUC up to FAR of 1×10^{-2}. The results shown are the average AUC and standard deviation from these iterations.

2) Target Detection Results: Overall, PFTs were detected with high accuracy, considering how much spectral variability is in each class (see Table IV and Fig. 8). Each of these classes contains multiple plant species spread across the ten flight lines, which covered approximately 12 980 km^2. This data set includes spectral variability that is inherent in all plant spectral data sets, which is caused by differences in properties such as plant structure, biochemistry, and water status. However, additional spectral variability is added due to the large spatial extent of this data set. Classes that are spectrally more homogenous due to the plant species sharing similar plant properties (e.g., AH and EBT) performed the best. Classes with more spectral variability due to significant differences in species (e.g., EBS and ENS) or PFTs are challenging to map due to open canopies (e.g., ENT) had lower performance. For example, ENS and ENT classes can have significant overlap between classes, but MTMI-ACE algorithms could distinguish the positive and negative bags from each other and yield appropriate target signatures. On the other hand, the URB class was easily detected by most algorithms because it is so different from the negative bags, which in this case, were mostly vegetation. In this particular data set, spectral magnitude did not improve target detection, as shown by MTMI-ACE outperforming MTMI-SMF.

The ability to determine multiple targets for target detection gave MTMI-ACE and MTMI-SMF a boost in performance compared to the single-target detection algorithms (MI-ACE and MI-SMF). The exception to this is the DS and EBT classes, which had better AUC results using MI-ACE and MI-SMF. These results demonstrate that even fewer targets could have been returned using MTMI-ACE and MTMI-SMF for these classes. In general, MTMI performs consistently well across all PFT classes, whereas other algorithms had more variability results across the classes.

TABLE IV

AVERAGED AUC WITH STANDARD DEVIATION IN PARENTHESES FOR EACH OF THE NINE CLASSES IN THE AVIRIS SANTA BARBARA DATA SET ACROSS THE DIFFERENT METHODS. THE BEST RESULTS (BASED ON AVERAGE AUC) ARE SHOWN IN BOLD; THE SECOND-BEST RESULTS ARE UNDERLINED

Methods	AH	DBT	DS	EBS	EBT	ENS	ENT	RS	URB
MTMI-ACE	0.440(0.078)	**0.219(0.082)**	0.213(0.050)	**0.121(0.041)**	**0.506(0.069)**	**0.090(0.024)**	0.249(0.098)	**0.664(0.121)**	0.881(0.101)
MTMI-SMF	0.419(0.102)	0.095(0.085)	0.079(0.070)	0.033(0.035)	0.299(0.058)	0.037(0.043)	0.061(0.028)	0.550(0.267)	0.879(0.077)
MILMD-ACE	0.543(0.107)	0.203(0.064)	0.220(0.039)	0.075(0.050)	0.381(0.073)	0.069(0.025)	0.248(0.091)	0.658(0.205)	**0.895(0.081)**
MILMD-SMF	0.008(0.006)	0.000(0.000)	0.000(0.001)	0.001(0.001)	0.091(0.098)	0.002(0.004)	0.000(0.000)	0.129(0.111)	0.838(0.088)
MI-ACE	**0.545(0.116)**	0.128(0.062)	**0.234(0.052)**	0.099(0.037)	0.376(0.049)	0.069(0.013)	**0.262(0.092)**	0.614(0.157)	0.859(0.069)
MI-SMF	0.494(0.088)	0.061(0.046)	0.136(0.054)	0.102(0.049)	0.362(0.042)	0.066(0.021)	0.129(0.070)	0.628(0.150)	0.836(0.053)
MI-HE	0.149(0.040)	0.095(0.033)	0.036(0.030)	0.029(0.034)	0.134(0.068)	0.041(0.020)	0.080(0.029)	0.232(0.080)	0.478(0.186)
eFUMI	0.051(0.044)	0.054(0.023)	0.014(0.003)	0.018(0.019)	0.048(0.022)	0.008(0.005)	0.056(0.020)	0.233(0.105)	0.189(0.094)

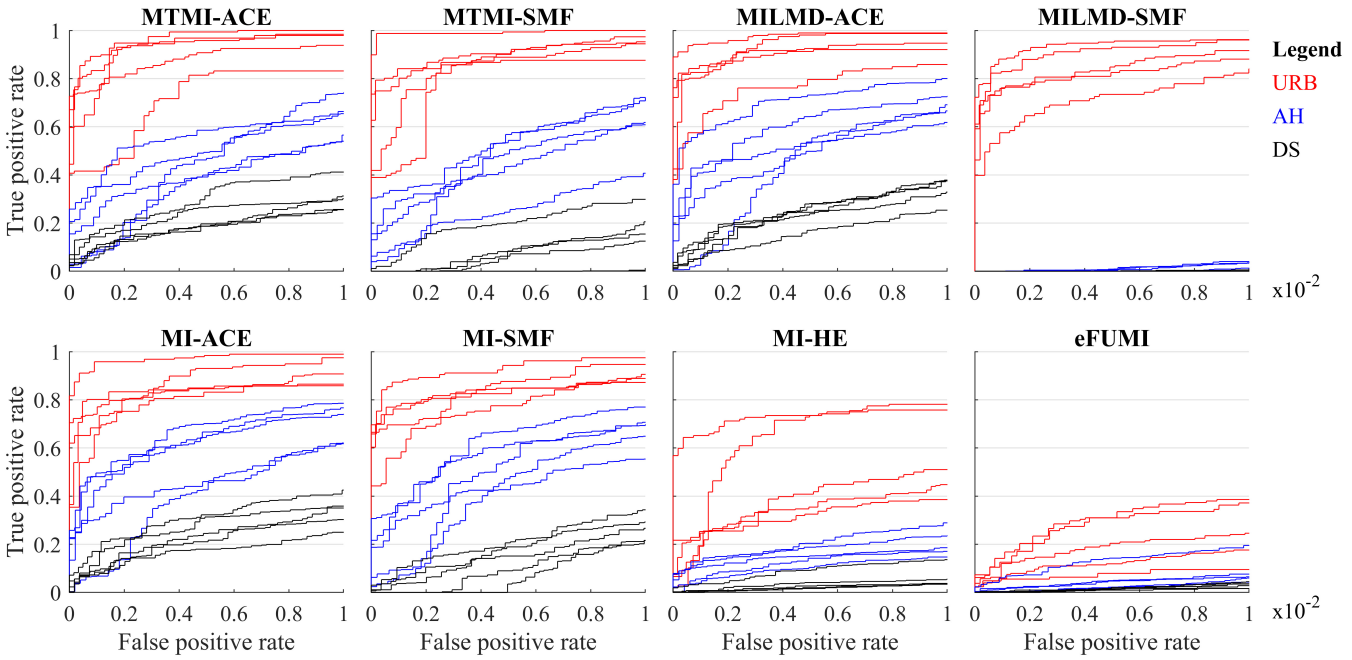

Fig. 8. AVIRIS Santa Barbara data set ROC curves for three classes (URB: urban, AH: annual herbs, and DS: deciduous shrubs) across the fivefold iterations. The false-positive rate scale is shown as 1×10^{-2}.

The MILMD-ACE algorithm, another multiple-target multiple-instance algorithms, also performed well when detecting PFT targets. Average AUC values between MTMI-ACE, MTMI-SMF, and MILMD-ACE were often very comparable with the standard deviations. For example, with the RS class, the MILMD algorithm performed better based on average AUC, but once the standard deviation was accounted for, the results were not different.

The MTMI algorithms outperformed MI-HE and eFUMI. This data sets difficulty is knowing how many targets are necessary to capture the spectral variability for high detection accuracy. In these algorithms, the user specifies how many targets or background targets are present in the data set. That exact number is returned, which may not be the optimal number of targets. The MTMI algorithm's benefit is that the user enters a max number of targets, but the algorithm determines the appropriate quantity for the class using the α parameter.

IV. DISCUSSION

As demonstrated in the experiments above, MIL for target detection has shown to be highly suitable for hyperspectral data applications. However, as many who have worked with these data sets know, the development of training labels even at the bag level can be challenging. For example, it is often difficult to guarantee that a negative bag will not have any target pixels. In MTMI-ACE and MTMI-SMF's objective function, we minimize the effect that these positive instances may have with the inclusion of the mean in (8), as long as the total number of positive instances are the minority. While the contribution of positive instances in negative bags was not explicitly tested, the AVIRIS Santa Barbara experiment had a high likelihood of target pixels in negative bags due to the nature of PFTs across this very diverse landscape with 18-m spatial resolution imagery. In that experiment, classes were still detected with sufficient accuracy and did not appear to reduce detection ability significantly. An experiment testing

this effect has been done in [6] where the MI-HE, MI-ACE, and MI-SMF algorithms were found not to be highly influenced by incomplete background knowledge and the addition of positive instances in negative bags. However, eFUMI was confused by these positive instances, and accuracy was decreased.

Future work with this algorithm should be to thoroughly test the ability to generate learned target signatures from one scene and use them for target detection on a new scene or sensor. Those experiments would increase the applicability and functionality of the algorithm. When testing this ability, especially to a new scene, the background statistics must be consistent with the original data sets background statistics since the target signatures were developed for maximum separability from the original background. In addition, these experiments do not test what effect the size of the training data has on accuracy. Experiments were compiled with real hyperspectral data sets, but other applications may not have as much training, especially when even bag-level labels are difficult to obtain.

V. Conclusion

In this work, the MTMI-ACE and MTMI-SMF algorithms for MIL problems are proposed and investigated. Both algorithms can learn multiple discriminative target concepts from ambiguously labeled data. Comprehensive experiments show that the proposed MTMI-ACE and MTMI-SMF algorithms are effective in learning discriminative target concepts. These two algorithms achieved superior performance over other state-of-the-art MIL algorithms in several experiments that tested different target detection scenarios.

In addition, MTMI-ACE and MTMI-SMF present a few advantages over comparison algorithms. First, the appropriate number of target signatures for a target's detection is returned, reducing the need for a user's knowledge of the target's spectral variability. Second, the MTMI-ACE and MTMI-SMF algorithms do not require that more than one target be present in each positive bag, increasing the number of applications. Finally, these algorithms efficiently determine target signatures compared to other sampling algorithms. Although this article focuses on hyperspectral target detection, the MTMI-ACE and MTMI-SMF algorithms follow the general MIL framework, which can be applied to any domain containing mixed and ambiguously labeled training data.

Appendix A
Optimization Update Equation Derivation

The objective function is written in (17). In (17), the detection statistic function, D, is expanded out for the ACE statistic showing the whitened data and inner product. The derivation in this appendix may be done using the SMF statistic following the same format except that $\hat{\hat{\mathbf{x}}}$ would be replaced with $\hat{\mathbf{x}}$:

$$\begin{aligned}\max_{\mathbf{S}} \quad & \frac{1}{N^+}\sum_{j:L_j=1}\max_{s_k\in\mathbf{S}}(D(\mathbf{x}^*_{j,k},\mathbf{s}_k)) \\ & -\frac{1}{N^-}\sum_{j:L_j=0}\frac{1}{N_j^-}\sum_{x_i\in B_j^-}D(\mathbf{x}_i,\mathbf{s}_k)-\frac{\alpha}{\binom{K}{2}}\sum_{k,l,l\neq k}D(\mathbf{s}_k,\mathbf{s}_l) \\ \text{s.t.} \quad & D(\mathbf{s}_k,\mathbf{s}_k)=1 \end{aligned} \tag{16}$$

TABLE V
Model Parameters for the Simulated Data Experiment

Algorithm	Parameters
MTMI-ACE	K = 4, α = 1
MTMI-SMF	K = 4, α = 1
MILMD-ACE	K = 2, α = 0.5
MILMD-SMF	K = 2, α = 0.5
MI-ACE	K = 1, α = 0
MI-SMF	K = 1, α = 0
MI-HE	T = 2, M = 3, ρ = 0.8, b = 5, β = 5, $\lambda = 1x10^{-3}$
eFUMI	M = 3, α = 1.2, β = 60, Γ = 10, u = 0.05

TABLE VI
Model Parameters for the MUUFL Single Target Experiment

Algorithm	Parameters
MTMI-ACE	K = 2, α = 0.1
MTMI-SMF	K = 2, α = 0.1
MILMD-ACE	K = 2, α = 0.1
MILMD-SMF	K = 2, α = 0.1
MI-ACE	K = 1, α = 0
MI-SMF	K = 1, α = 0
MI-HE	T = 1, M = 9, ρ = 0.3, b = 5, β = 1, $\lambda = 5x10^{-3}$
eFUMI	M = 7, α = 2, β = 60, Γ = 5, u = 0.05

TABLE VII
Model Parameters for the MUUFL Multiple Target Experiment

Algorithm	Parameters
MTMI-ACE	K = 8, α = 0.5
MTMI-SMF	K = 8, α = 0.5
MILMD-ACE	K = 8, α = 0.5
MILMD-SMF	K = 8, α = 0.5
MI-ACE	K = 1, α = 0
MI-SMF	K = 1, α = 0
MI-HE	T = 4, M = 2, ρ = 0.3, b = 5, β = 1, $\lambda = 5x10^{-3}$
eFUMI	M = 2, α = 2, β = 60, Γ = 5, u = 0.05

$$\begin{aligned}\max_{\mathbf{S}} \quad & \frac{1}{N^+}\sum_{j:L_j=1}\max_{\hat{\hat{\mathbf{s}}}_k\in\mathbf{S}}\left(\hat{\hat{\mathbf{x}}}^{*T}_{j,k}\hat{\hat{\mathbf{s}}}_k\right) \\ & -\frac{1}{N^-}\sum_{j:L_j=0}\frac{1}{N_j^-}\sum_{\hat{\hat{x}}_i\in B_j^-}\left(\hat{\hat{\mathbf{x}}}_i^T\hat{\hat{\mathbf{s}}}_k\right)-\frac{\alpha}{\binom{K}{2}}\sum_{k,l,l\neq k}\left(\hat{\hat{\mathbf{s}}}_k^T\hat{\hat{\mathbf{s}}}_l\right) \\ \text{s.t.} \quad & \hat{\hat{\mathbf{s}}}_k^T\hat{\hat{\mathbf{s}}}_k=1. \end{aligned} \tag{17}$$

The optimal update equation for each target signature, $\hat{\hat{\mathbf{s}}}_k$, can be solved for using the associated Lagrangian, written in the following equation:

$$\begin{aligned}\mathcal{L} = & \frac{1}{N^+}\sum_{j:L_j=1}\max_{\hat{\hat{\mathbf{s}}}_k\in\mathbf{S}}\left(\hat{\hat{\mathbf{x}}}^{*T}_{j,k}\hat{\hat{\mathbf{s}}}_k\right) \\ & -\frac{1}{N^-}\sum_{j:L_j=0}\frac{1}{N_j^-}\sum_{\hat{\hat{x}}_i\in B_j^-}\left(\hat{\hat{\mathbf{x}}}_i^T\hat{\hat{\mathbf{s}}}_k\right) \\ & -\frac{\alpha}{\binom{K}{2}}\sum_{k,l,l\neq k}\left(\hat{\hat{\mathbf{s}}}_k^T\hat{\hat{\mathbf{s}}}_l\right)-\lambda\left(\hat{\hat{\mathbf{s}}}_k^T\hat{\hat{\mathbf{s}}}_k-1\right). \end{aligned} \tag{18}$$

The derivative of the Lagrangian with respect to the target signature is taken and shown in (19). Here, the max operation

TABLE VIII
MODEL PARAMETERS FOR THE AVIRIS SANTA BARBARA EXPERIMENT

Class	Number of Bags	Total Number of Pixels	Algorithm	Parameters
AH	199	6329	MTMI-ACE	$K = 10$, $\alpha = 0.05$
DBT	75	2213	MTMI-SMF	$K = 10$, $\alpha = 0.05$
DS	97	2361	MILMD-ACE	$K = 10$, $\alpha = 0.05$
EBS	182	4472	MILMD-SMF	$K = 10$, $\alpha = 0.05$
EBT	239	6467	MI-ACE	$K = 1$, $\alpha = 0$
ENS	105	3052	MI-SMF	$K = 1$, $\alpha = 0$
ENT	112	2050	MI-HE	$T = 10$, $M = 20$, $\rho = 0.8$, $b = 5$, $\beta = 1$, $\lambda = 5x10^{-3}$
RS	66	1298	eFUMI	$M = 20$, $\alpha = 2$, $\beta = 60$, $\Gamma = 5$, $u = 0.05$
URB	68	686		

on the first term is expanded out using an indicator function

$$\frac{\partial \mathcal{L}}{\partial \hat{\hat{\mathbf{s}}}_k} = \frac{1}{N_k^+} \sum_{j:L_j=1} \delta_{j,k}^+ \hat{\hat{\mathbf{x}}}_{j,k}^* - \frac{1}{N^-} \sum_{j:L_j=0} \frac{1}{N_j^-} \sum_{\hat{\hat{x}}_i \in B_j^-} \hat{\hat{\mathbf{x}}}_i - \frac{\alpha}{(K-1)} \sum_{l,l\neq k} \hat{\hat{\mathbf{s}}}_l - 2\lambda \hat{\hat{\mathbf{s}}}_k. \tag{19}$$

The positive bag indicator for target signature $\hat{\hat{\mathbf{s}}}_k$ is defined as

$$\delta_{j,k}^+ = \begin{cases} 1 & \text{if } \hat{\hat{\mathbf{x}}}_{j,k}^{*T} \hat{\hat{\mathbf{s}}}_k > \hat{\hat{\mathbf{x}}}_{j,l}^{*T} \hat{\hat{\mathbf{s}}}_l, \quad \forall l \neq k \\ 0 & \text{otherwise.} \end{cases} \tag{20}$$

Then, solving for the target signature the update equation and the Lagrangian multiplier is solved for in the following equations:

$$\hat{\hat{\mathbf{s}}}_k = \frac{1}{2\lambda} \left(\frac{1}{N_k^+} \sum_{j:L_j=1} \delta_{j,k}^+ \hat{\hat{\mathbf{x}}}_{j,k}^* - \frac{1}{N^-} \sum_{j:L_j=0} \frac{1}{N_j^-} \sum_{\hat{\hat{x}}_i \in B_j^-} \hat{\hat{\mathbf{x}}}_i - \frac{\alpha}{(K-1)} \sum_{l,l\neq k} \hat{\hat{\mathbf{s}}}_l \right) \tag{21}$$

$$\lambda = \frac{||\hat{\hat{\mathbf{t}}}||}{2}. \tag{22}$$

Finally, the update equation for the kth target signature with the Lagrangian multiplier is shown in the following equation:

$$\hat{\hat{\mathbf{s}}}_k = \frac{\hat{\hat{\mathbf{t}}}}{\left\|\hat{\hat{\mathbf{t}}}\right\|} \text{ where } \hat{\hat{\mathbf{t}}} = \frac{1}{N_k^+} \sum_{j:L_j=1} \delta_{j,k}^+ \hat{\hat{\mathbf{x}}}_{j,k}^* - \frac{1}{N^-} \sum_{j:L_j=0} \frac{1}{N_j^-} \sum_{\hat{\hat{x}}_i \in B_j^-} \hat{\hat{\mathbf{x}}}_i - \frac{\alpha}{(K-1)} \sum_{l,l\neq k} \hat{\hat{\mathbf{s}}}_l. \tag{23}$$

APPENDIX B
PARAMETERS FOR EXPERIMENTS

A. Simulated Data Set

See Table V.

B. MUUFL Gulfport Data Set—Single Target

See Table VI.

C. MUUFL Gulfport Data Set—Multiple Targets

See Table VII.

D. AVIRIS Santa Barbara Data Set

See Table VIII.

ACKNOWLEDGMENT

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.

REFERENCES

[1] N. Keshava and J. F. Mustard, "Spectral unmixing," *IEEE Signal Process. Mag.*, vol. 19, no. 1, pp. 44–57, Jan. 2002.
[2] T. G. Dietterich, R. H. Lathrop, and T. Lozano-Pérez, "Solving the multiple instance problem with axis-parallel rectangles," *Artif. Intell.*, vol. 89, nos. 1–2, pp. 31–71, Jan. 1997.
[3] O. Maron and T. Lozano-Pérez, "A framework for multiple-instance learning," in *Proc. Adv. Neural Inf. Process. Syst.*, 1998, pp. 570–576.
[4] P. Zhong, Z. Gong, and J. Shan, "Multiple instance learning for multiple diverse hyperspectral target characterizations," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 31, no. 1, pp. 246–258, Jan. 2019.
[5] A. Zare, C. Jiao, and T. Glenn, "Discriminative multiple instance hyperspectral target characterization," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 40, no. 10, pp. 2342–2354, Oct. 2018.
[6] C. Jiao, C. Chen, R. G. McGarvey, S. Bohlman, L. Jiao, and A. Zare, "Multiple instance hybrid estimator for hyperspectral target characterization and sub-pixel target detection," *ISPRS J. Photogramm. Remote Sens.*, vol. 146, pp. 235–250, Dec. 2018.
[7] C. Jiao and A. Zare, "Multiple instance hybrid estimator for learning target signatures," 2017, *arXiv:1701.02258*. [Online]. Available: http://arxiv.org/abs/1701.02258
[8] J. Bocinsky, "Learning multiple target concepts from uncertain, ambiguous data using the adaptive cosine estimator and spectral match filter," M.S. thesis, Dept. Elect. Comput. Eng., University of Florida, Gainesville, FL, USA, 2019. [Online]. Available: https://ufdc.ufl.edu/UFE0054471/00001
[9] Q. Zhang and S. A. Goldman, "EM-DD: An improved multiple-instance learning technique," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 2, 2002, pp. 1073–1080.
[10] S. Srinivas, "A generalization of the noisy-or model," in *Proc. 9th Int. Conf. Uncertainty Artif. Intell.* Burlington, MA, USA: Morgan Kaufmann Publishers, 1993, pp. 208–215.
[11] A. Zare and P. Gader, "Pattern recognition using functions of multiple instances," in *Proc. 20th Int. Conf. Pattern Recognit. (ICPR)*, Aug. 2010, pp. 1092–1095.

[12] A. Zare and P. Gader, "Sparsity promoting iterated constrained endmember detection in hyperspectral imagery," *IEEE Geosci. Remote Sens. Lett.*, vol. 4, no. 3, pp. 446–450, Jul. 2007.
[13] C. Jiao and A. Zare, "Functions of multiple instances for learning target signatures," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 8, pp. 4670–4686, Aug. 2015.
[14] S. Kraut and L. L. Scharf, "The CFAR adaptive subspace detector is a scale-invariant GLRT," *IEEE Trans. Signal Process.*, vol. 47, no. 9, pp. 2538–2541, Sep. 1999.
[15] S. Kraut, L. L. Scharf, and L. T. McWhorter, "Adaptive subspace detectors," *IEEE Trans. Signal Process.*, vol. 49, no. 1, pp. 1–16, Jan. 2001.
[16] W. F. Basener, "Clutter and anomaly removal for enhanced target detection," *Proc. SPIE*, vol. 7695, May 2010, Art. no. 769525.
[17] N. M. Nasrabadi, "Regularized spectral matched filter for target recognition in hyperspectral imagery," *IEEE Signal Process. Lett.*, vol. 15, pp. 317–320, 2008.
[18] J. MacQueen *et al.*, "Some methods for classification and analysis of multivariate observations," in *Proc. 5th Berkeley Symp. Math. Statist. Probab.*, Oakland, CA, USA, vol. 1, 1967, pp. 281–297.
[19] S. K. Meerdink, D. A. Roberts, K. L. Roth, J. Y. King, P. D. Gader, and A. Koltunov, "Classifying California plant species temporally using airborne hyperspectral imagery," *Remote Sens. Environ.*, vol. 232, Oct. 2019, Art. no. 111308, doi: 10.1016/j.rse.2019.111308.
[20] A. M. Baldridge, S. J. Hook, C. I. Grove, and G. Rivera, "The ASTER spectral library version 2.0," *Remote Sens. Environ.*, vol. 113, no. 4, pp. 711–715, Apr. 2009, doi: 10.1016/j.rse.2008.11.007.
[21] T. Glenn, A. Zare, P. Gader, and D. Dranishnikov. (2016). *Bullwinkle: Scoring Code for Sub-Pixel Targets (Version 1.0) [Software]*. [Online]. Available: https://github.com/GatorSense/MUUFLGulfport/
[22] C. M. Lee *et al.*, "An introduction to the NASA hyperspectral InfraRed imager (HyspIRI) mission and preparatory activities," *Remote Sens. Environ.*, vol. 167, pp. 6–19, Sep. 2015. [Online]. Available: http://linkinghub.elsevier.com/retrieve/pii/S0034425715300419
[23] R. O. Green, B. Pavri, J. Faust, and O. Williams, "AVIRIS radiometric laboratory calibration, inflight validation and a focused sensitivity analysis in 1998," NASA Jet Propulsion Laboratory, Pasadena, CA, USA, Tech. Rep., 1998. [Online]. Available: https://trs.jpl.nasa.gov/handle/2014/13692

Susan Meerdink received the Ph.D. degree in geography from the University of California at Santa Barbara, Santa Barbara, CA, USA, in 2018.

She is an Assistant Professor with the Department of Geographical and Sustainability Sciences, The University of Iowa, Iowa, IA, USA. Her research interest includes remote sensing analysis with an emphasis on hyperspectral and thermal data sets.

James Bocinsky received the bachelor's degree in computer engineering from the University of Florida, Gainesville, FL, USA, in 2017, and the Master of Science degree from the Department of Electrical and Computer Engineering, University of Florida, in 2019.

While a Researcher at the Machine Learning and Sensing Laboratory, University of Florida, he focused his research on developing algorithms for subsurface explosive hazard detection, especially focusing on algorithms that can handle ambiguously labeled data. After disseminating his work at conferences and completing his Master of Science degree, he joined American Express, Sunrise, FL, USA, as a Machine Learning Engineer in 2019. He has found success at American Express by winning the 2019 American Express internal Kaggle competition and developing proactive servicing machine learning models. His primary areas of interest are classical machine learning, time series prediction, embedded systems programming, remote sensing, and pattern recognition.

Alina Zare (Senior Member, IEEE) received the Ph.D. degree in computer and information science and engineering from the University of Florida, Gainesville, FL, USA, in 2008.

She is a Professor with the Department of Electrical and Computer Engineering and the Director of the Machine Learning and Sensing Laboratory, University of Florida. Her research interests include analysis of a variety of remote sensing data and the development of machine learning and pattern recognition algorithms for automated analysis of remote sensing data sets.

Nicholas Kroeger received the B.S. degree in computer science from the University of Florida, Gainesville, FL, USA, in 2018, where he is pursuing the Ph.D. degree in computer science and machine learning.

His research interests include machine learning interpretability, outlier detection for neural networks, and underwater acoustic modeling for coral reef environments.

Connor McCurley (Student Member, IEEE) received the B.S. degree in electrical engineering from Oklahoma State University, Stillwater, OK, USA, in 2017. He is pursuing the Ph.D. degree with the Department of Electrical and Computer Engineering, University of Florida, Gainesville, FL, USA.

His research interests include machine learning and pattern recognition, manifold learning, metric embedding and learning from imprecise, and uncertain training data.

Daniel Shats received the B.Sc. degree in computer science engineering and the B.Sc. degree in mathematics from the University of Florida, Gainesville, FL, USA, in 2019. He is pursuing the master's degree in computer science with Technion–Israel Institute of Technology, Haifa, Israel.

His research interest includes deep learning for computer vision.

Paul Gader (Fellow, IEEE) received the Ph.D. degree in mathematics for image-processing-related research from the University of Florida, Gainesville, FL, USA, in 1986.

He is a Professor with the Department of Computer and Information Science and Engineering and the Engineering School of Sustainable Infrastructure and Environment, University of Florida. He performed his first research in image processing in 1984, working on algorithms for the detection of bridges in forward-looking infrared imagery as a Summer Student Fellow at Eglin Air Force Base. He has authored or coauthored hundreds of refereed journal articles and conference papers. He has since worked on a wide variety of theoretical and applied research problems, including fast computing with linear algebra, mathematical morphology, fuzzy sets, Bayesian methods, handwriting recognition, automatic target recognition, biomedical image analysis, landmine detection, human geography, and hyperspectral and light detection, and ranging image analysis projects.