DEVELOPING SPECTRAL LIBRARIES USING MULTIPLE TARGET MULTIPLE INSTANCE ADAPTIVE COSINE/COHERENCE ESTIMATOR

Susan Meerdink [1], James Bocinsky [1], Erin Wetherley [2], Alina Zare [1], Connor McCurley [1], & Paul Gader [1]

[1] University of Florida, Gainesville, FL, USA, [2] University of California, Santa Barbara

ABSTRACT

Traditional methods of developing spectral libraries for unmixing hyperspectral images tend to require domain knowledge of the study area and the material's spectra. In this paper, we propose using the Multiple Target Multiple Instance Adaptive Cosine/Coherence Estimator (Multi-Target MI-ACE) algorithm to develop spectral libraries that will capture the same spectral variability as traditional methods but require less processing time and domain knowledge. We compared traditional and Multi-Target MI-ACE generated spectral libraries' ability to accurately predict sub-pixel composition using Multiple Endmember Spectral Mixture Analysis (MESMA). Multi-Target MI-ACE spectral libraries maintained the same sub-pixel composition accuracy compared to traditional libraries, while significantly reducing model complexity. Additionally, the Multi-Target MI-ACE confidence values could be used to constrain MESMA model complexity and considerably reduce the number of endmember permutations needed. In summary, Multi-Target MI-ACE has been found to successfully develop spectral libraries that capture the full spectral variability compared to traditional approaches, while reducing MESMA model complexity and the need for domain knowledge.

***Index Terms*—** endmember extraction, endmember variability, hyperspectral, unmixing, urban mapping

1. INTRODUCTION

In remote sensing, each pixel measures the interaction of electromagnetic radiation with multiple surface constituents, regardless of spatial resolution [1]. The presence of surface mixtures, independent of scale, therefore requires decomposition of measured signals in order to map surface variability. Methods for doing so primarily rely on linear spectral mixture analysis (SMA), which assumes that the measured reflected signal of a mixed-composition pixel is a linear combination of reflectance from all sub-pixel surfaces, proportional to their pixel fraction [2].

Accurate SMA requires appropriate endmember selection, which involves identifying both the number and type of endmembers [3, 4]. The collective group of endmembers used to unmix an image is called a spectral library, and ideally captures the full spectral variability of materials present in the image. For imagery that contain high levels of material diversity, such as urban scenes, identifying representative endmembers can be particularly challenging [5]. Too many endmembers, or endmembers not representative of image materials, lead to physically inaccurate proportion estimates [6]. Many solutions to select the optimal number and type of endmembers have been developed, but vary in the range of user input necessary [7, 8].

The first technique developed to address endmember variability was Multiple Endmember Spectral Mixture Analysis (MESMA) [9]. In this method, endmembers are allowed to vary on a per pixel basis and multiple endmembers can represent each class, which removes the fixed endmember restriction of SMA. MESMA is a computationally complex process, and therefore requires a small spectral library. Producing this small library typically requires beginning with a spectral library of hundreds and iteratively finding the best-fit model of endmember combinations to assign to each pixel. Although iterative mixture analysis cycles have been shown to produce good results, the computational complexity of the method is a significant drawback when applied on hyperspectral data [10]. Additionally, MESMA requires domain knowledge of the study area for the final development of spectral libraries [5, 11].

Alternatively, we propose a method for developing spectral libraries using Multiple Target Multiple Instance Adaptive Cosine/Coherence Estimator (Multi-Target MI-ACE), an algorithm originally designed for hyperspectral target detection. The first step of Multi-Target MI-ACE is to identify representative endmembers that characterize the spectral variability seen in the imagery. This step is leveraged to develop Multi-Target MI-ACE spectral libraries by choosing endmembers from a large spectral library. In this paper, we compare Multi-Target MI-ACE and Iterative Endmember Selection/Iterative Classification Reduction (IES/ICR) spectral libraries to assess their ability to retrieve sub-pixel composition.

2. MULTI-TARGET MI-ACE

The Multi-Target MI-ACE algorithm is an extension of MI-ACE for the purpose of identifying multiple targets [12]. The objective of the Multi-Target MI-ACE algorithm is to learn a dictionary of class representations with the focus of maximizing detection of those classes against a background [13, 14]. This algorithm fits the Multiple Instance Learning framework and assumes the data is grouped into bags with bag-level labels [15]. With this, let $\mathbf{X} = \{\mathbf{x}_1, ..., \mathbf{x}_N\}$ be training data with each sample, $\mathbf{x}_i$ being a vector with dimensionality D. The data is grouped into J bags $\mathbf{B} = \{\mathbf{B}_1, ..., \mathbf{B}_J\}$ with labels, $L = \{L_1, ..., L_J\}$, 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, $\mathbf{x_{ji}}$, in bag j that is from the target class, $l_{ji} = 1$, seen in Equation (1). Additionally, a bag is considered negative, $\mathbf{B_j^-}$, with label $L_j = 0$, if all instances in bag j are from the background class, $l_{ji} = 0$, seen in Equation (2). The number of instances in both positive and negative bags is variable.

$$\text{if} \quad L_j = 1, \quad \exists \mathbf{x}_{ji} \in \mathbf{B}_j^+ \quad \text{s.t.} \quad l_{ji} = 1 \tag{1}$$

$$\text{if} \quad L_j = 0, \quad \forall \mathbf{x}_{ji} \in \mathbf{B}_j^-, \quad \text{s.t.} \quad l_{ji} = 0 \tag{2}$$

After data is grouped into bags, the algorithm learns a dictionary of class endmembers, $\mathbf{S}$, that maximizes the objective function shown in Equation (3).

$$\begin{aligned} \max_{\mathbf{S}} \quad & \frac{1}{N^+} \sum_{j:L_j=1} \max_{s_k \in \mathbf{S}} (D(\mathbf{x}^*_{j,k}, \mathbf{s}_k)) - \\ & \frac{1}{N^-} \sum_{j:L_j=0} \frac{1}{N_j^-} \sum_{x_i \in B_j^-} D(\mathbf{x}_i, \mathbf{s}_k) - \\ & \frac{\alpha}{\binom{K}{2}} \sum_{k=1}^{K-1} \sum_{l=k+1}^{K} D(\mathbf{s}_k, \mathbf{s}_l) \quad \text{s.t.} \quad D(\mathbf{s}_k, \mathbf{s}_k) = 1 \end{aligned} \tag{3}$$

Here N^+ and N^- are the number of positive and negative bags respectively, N_j^- is the number of instances in negative bag j, and $\mathbf{s}_k$ is the k^{th} class endmember in the dictionary. $\mathbf{x}^*_{j,k}$ is known as the positive bag representative and is the instance in the j^{th} positive bag that is the most similar to the k^{th} estimated class endmember, $\mathbf{s}_k$. This is shown in Equation (4).

$$\mathbf{x}^*_{j,k} = \arg \max_{x_i \in B_j^+} D(\mathbf{x}_i, \mathbf{s}_k) \tag{4}$$

$D(\mathbf{x}, \mathbf{s})$ represents the ACE detection statistic between an unknown instance, $\mathbf{x}$ and a class endmember, $\mathbf{s}$. The ACE detection statistic is shown in (5), where $\hat{\mathbf{x}} = \mathbf{D}^{-\frac{1}{2}} \mathbf{U}^T (\mathbf{x} - \boldsymbol{\mu_b})$, $\hat{\mathbf{s}} = \mathbf{D}^{-\frac{1}{2}} \mathbf{U}^T \mathbf{s}$, $\mathbf{U}$ and $\mathbf{D}$ are the eignenvectors and eigenvalues of the background covariance.

$$D_{ACE}(\mathbf{x}_n, \mathbf{s}) = \hat{\hat{\mathbf{s}}}^T \hat{\hat{\mathbf{x}}}, \quad \hat{\hat{\mathbf{x}}} = \frac{\hat{\mathbf{x}}}{||\hat{\mathbf{x}}||}, \quad \hat{\hat{\mathbf{s}}} = \frac{\hat{\mathbf{s}}}{||\hat{\mathbf{s}}||} \tag{5}$$

There are two steps to the Multi-Target MI-ACE algorithm: initialization and optimization. Initialization maximizes the objective function (Eq. (3)) using the user defined max number of endmembers and greedily selects a set of positive instances. In this paper, the set of targets derived during the initialization process are called the MT MI-ACE (init). The second step optimizes the objective function, which is posed as a lagrangian problem, to the update equation and derive the k^{th} target signature. During optimization, endmembers are removed if found to have no positive bag representatives with a higher ACE confidence relative to that of all other endmembers. In other words, endmembers are removed if they do not describe a target class better than other endmembers. The optimized library is a subset of the initialized library. Multi-Target MI-ACE projects spectra into a whitened space to determine the best representative targets. In order to use the original reflectance values instead of projecting the image into this space, indices of selected endmembers were tracked and used to pull the matching endmember from the spectral library. In this paper, the set of targets derived during the optimization process are called the MT MI-ACE (opt). For more information on algorithm are found in [13, 14].

3. EXPERIMENTAL DESIGN

To determine how Multi-Target MI-ACE derived spectral libraries compared to traditional approaches, we used a dataset previously developed for urban fractional cover studies in Santa Barbara, CA [11]. Hyperspectral image data were collected with the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) sensor at 18m spatial resolution on August 29, 2014. More details on the collection, sensor, and image processing are found in [11]. As part of this dataset, 67 fraction validation polygons (180 x 180 m size) were developed to assess the accuracy of the MESMA products. The fractions of these polygons were prepared using E-Cognition to classify 1m NAIP imagery and were then manually corrected through comparison to August 2014 Google Earth imagery. Additionally, this dataset has a spectral library that was extracted from imagery using polygons that contained pure spectra of materials. A total of 3288 unique endmembers were obtained from the 237 pure polygons to form the reference library that contained endmembers for six classes: turfgrass, tree, paved, roof, non-photosynthetic vegetation (NPV), and soil. Since MESMA runs all permutations of class endmembers, further reduction of this library is necessary for computational efficiency.

One approach for reducing library size is a combination of Iterative Endmember Selection (IES) [16] and Iterative Classification Reduction (ICR) [5, 11]. IES is an automated algorithm that selects representative spectra of the larger reference library by comparing all pairs of endmembers and selecting those that have the highest kappa value for classifying the entire reference library. To optimally capture the variability of each reference library, ICR is run to maximize class separability by classifying the image using MESMA constrained to one endmember per pixel. Results are visually inspected and endmembers that over or under-map are removed. This process is repeated until no discernable improvement is found. The unmixing results from this library are referred to as IES/ICR in this paper.

In this paper, we propose the use of Multi-Target MI-ACE to select the best representative endmembers. We used the same endmember reference library that contains the six classes and 3288 endmembers. Iterating through each class, all polygons matching that class label were selected as positive bags, while all other polygons were chosen as negative bags. The background mean and covariance were calculated from all endmembers in the reference library, the initial number of class endmembers was 20, and α was 1. We developed two Multi-Target MI-ACE libraries: one using initialized targets referred to as MT MI-ACE (init) and one using optimized targets referred to as MT MI-ACE (opt).

We used MESMA to calculate class sub-pixel proportions with each spectral library. MESMA selects the best fitting model based on maximum and threshold RMSE values, which we set to 2.5% and 0.7%, respectively. In other words, a pixel could not be modeled with an RMSE below 2.5% reflectance, and a more complex model would be used if it improved the RMSE by at least 0.7% [3]. Fractions were constrained between 0 and 1, no pixel could contain $>80\%$ shade, and the number of endmembers per pixel was limited to a maximum of three plus shade. MESMA restricts the overall endmember combination to one class representative per pixel and will not evaluate a possible mixture of two soil endmembers or two tree endmembers. Lastly, we shade normalized the MESMA proportions. We used validation polygon boundaries to extract the total proportion coverage from MESMA products and compared against validation polygon proportions.

4. RESULTS

Table 1 shows the number of endmembers selected for each class across the three libraries. Each library started with the 3288 endmember reference library. The IES/ICR library contained 226 endmembers after IES, then 61 endmembers after ICR. The MT MI-ACE (init) library yielded a total of 82 endmembers, while the MT MI-ACE (opt) library yielded fewer than five endmembers per class resulting in a spectral library of 19 endmembers. The MT MI-ACE (init) spectral library found a slightly higher number of endmembers for each class compared to the IES/ICR spectral library, which resulted in a larger endmember library. Even a library with 20 more endmembers can result in a significant increase in computation time because MESMA finds all permutations of a class's endmembers for 2, 3, and 4 endmember models. However, the MT MI-ACE (opt) library found significantly fewer endmembers compared to the other libraries with only three endmembers per class. The user has some control over the size of MT MI-ACE libraries. They control the number of initialized endmembers and alpha which allows for less endmembers with distinct signatures or more endmembers with similar signatures.

Unmixing results using MESMA show that MT MI-ACE spectral libraries perform similar to the IES/ICR spectral library and none were found to be significantly different (Table 2; Figure 1). Both MT MI-ACE libraries performed better than the IES/ICR library on NPV, paved, roof, and tree classes. However, the IES/ICR library significantly outperformed for the turfgrass and soil classes. All three libraries had similar issues such as over predicting turfgrass and soil proportions at low validation proportions, which are classes that exhibit brighter reflectance. For proportions less than 0.2, the algorithms cannot accurately retrieve proportions. MT MI-ACE (opt) library had only three endmembers per class but appeared to have captured the spectral variability present in the imagery due to the comparable unmixing results. Having a library of 19 endmembers compared to 61 or 82 significantly decreases MESMA processing time.

Class	MT MI-ACE (opt)	MT MI-ACE (init)	IES/ICR
NPV	2	15	14
Paved	4	20	6
Roof	5	20	17
Soil	2	9	3
Tree	4	7	11
Turfgrass	2	11	10
Total	**19**	**82**	**61**

Table 1. Number of endmembers in each class that were selected for each spectral library.

Class	MT MI-ACE (opt)	MT MI-ACE (init)	IES/ICR
NPV	**0.080**	**0.081**	0.095
Paved	**0.064**	**0.094**	0.120
Roof	**0.091**	**0.052**	0.095
Soil	0.073	0.057	**0.048**
Tree	**0.073**	**0.079**	0.081
Turfgrass	0.126	0.128	**0.057**

Table 2. Root mean squared error (RMSE) for proportion predictions across the three spectral libraries. Bold values designate MT MI-ACE library classes with lower RMSE than the IES/ICR library.

In the current implementation of MESMA, unmixing an image requires the algorithm to iterate through all permutations of class endmembers in combinations of 2, 3, and 4 endmember models to find the best fit model. This can result in a large number of iterations depending on spectral library size.

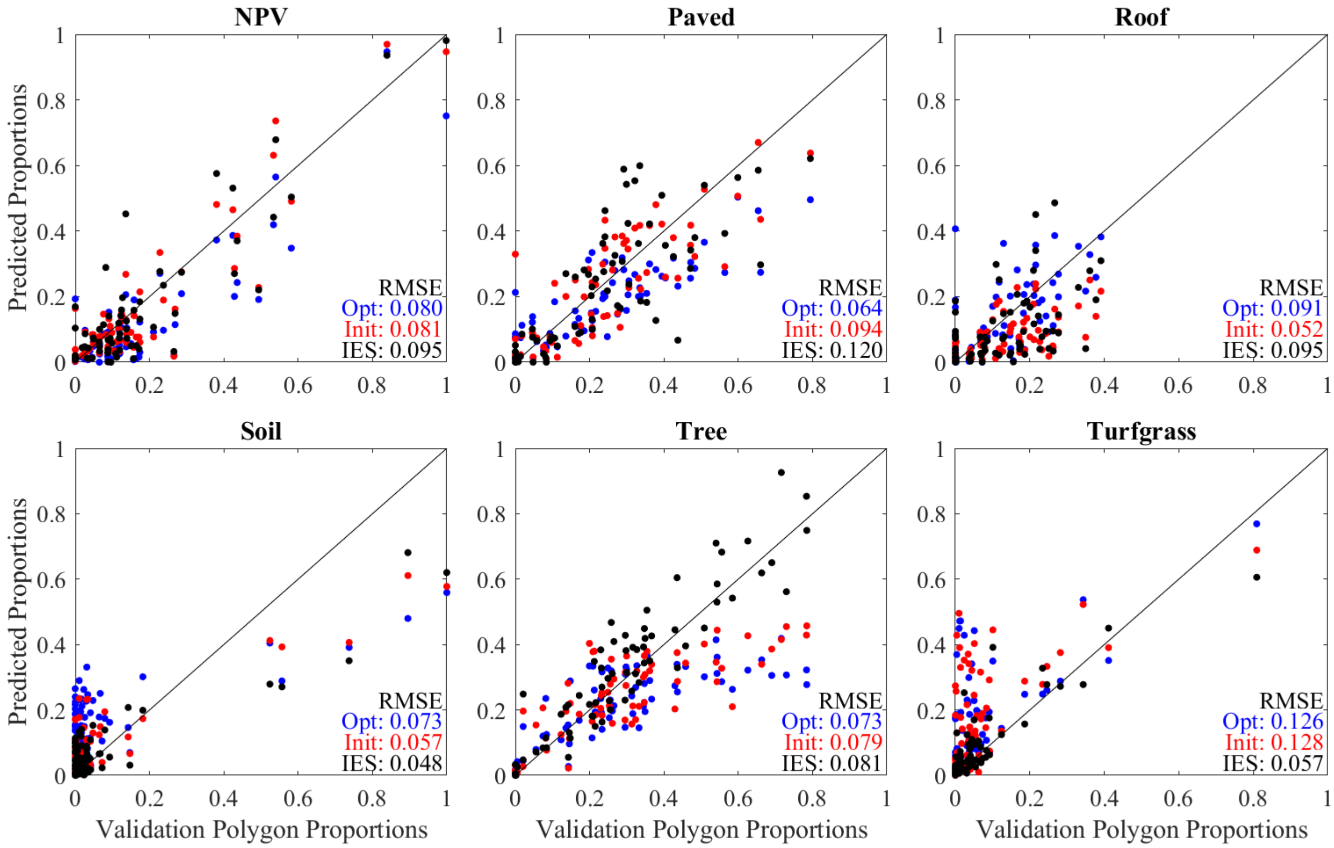

Fig. 1. Scatterplots comparing validation polygon proportions (n = 67) to MESMA predicted proportions with the root mean squared error (RMSE) using the three spectral libraries.

Class	MT MI-ACE (opt)	MT MI-ACE (init)	IES/ICR
NPV	88.1	**89.6**	88.1
Paved	**92.5**	**92.5**	91.0
Roof	79.1	80.6	80.6
Soil	76.1	76.1	76.1
Tree	95.5	95.5	95.5
Turfgrass	**83.6**	**80.6**	77.6
Iterations	$1.25x10^6$	$1.29x10^8$	$2.87x10^{10}$

Table 3. The percent of pixels that had the classes correctly classified using confidence values from Multi-Target MI-ACE. Bold values are classes that were predicted with higher accuracy than the IES/ICR library.Final row shows the number of MESMA models using ACE confidence values to constrain pixel classes.

For example, with this paper's dataset, the IES/ICR library would run 19881 models to cover all endmember combinations across 1,445,964 pixels for a total of 2.87x10^{10} iterations. The Multi-Target MI-ACE algorithm can not only be used to develop spectral libraries but also constrain the number of iterations necessary to unmix with MESMA. For each pixel in an image, the Multi-Target MI-ACE algorithm generates a confidence value (between -1 and 1) for each endmember, indicating the likelihood that a pixel contains an endmember or combination of endmembers. We selected all endmembers with a positive confidence value and used them to determine which classes are present in a pixel. The confidence values predicted which classes each pixel contains as accurately as MESMA (Table 3). By applying these confidence values, MEMSA iterations can be constrained only to run class endmember combinations that are present in a pixel rather than all endmember combinations, which dramatically reduces the number of permutations for an image. Using these constraints and the IES/ICR library, the number of MESMA iterations would decrease from 2.87x10^{10} to 1.29x10^8 using confidence values from the initialized library and 1.25x10^6 using the optimized library confidence values.

5. CONCLUSIONS

In this paper, we explored the potential for Multi-Target MI-ACE to generate spectral libraries capable of quantifying sub-pixel fractional cover with accuracy similar to, or better than, manually intensive IES/ICR developed spectral libraries. The optimized MT MI-ACE library contained only 18 endmembers, yet captured the spectral variability across the image with comparable performance to the IES/ICR library with 61 endmembers. IES/ICR libraries require significant user effort and domain knowledge. On the other hand, Multi-Target MI-ACE does not require domain knowledge and quickly develops a spectral library that yields comparable results. The MT MI-ACE spectral libraries did not perform as well with turfgrass and soil classes but did perform better with materially-variable classes (e.g., paved, roof) that are notoriously difficult to map at sub-pixel scales. In addition to developing spectral libraries, Multi-Target MI-ACE could be used to constrain MESMA models by specifying which classes are present in pixels, significantly decreasing processing time. In summary, Multi-Target MI-ACE can be used to efficiently develop spectral libraries that capture spectral variability and retrieve accurate sub-pixel proportions.

6. REFERENCES

[1] Nirmal Keshava and John F. Mustard, "Spectral unmixing," *IEEE Signal Processing Magazine*, vol. January, pp. 44–57, 2002.

[2] J J Settle and N A Drake, "Linear mixing and the estimation of ground cover proportions," *International Journal of Remote Sensing*, vol. 14, no. 6, pp. 1159–1177, 1993.

[3] Stefanie Tompkins, John F. Mustard, Carlé M. Pieters, and Donald W. Forsyth, "Optimization of endmembers for spectral mixture analysis," *Remote Sensing of Environment*, vol. 59, no. 3, pp. 472–489, 1997.

[4] Andrew J. Elmore, John F. Mustard, Sara J. Manning, and David B. Lobell, "Quantifying Vegetation Change in Semiarid Environments : Precision and Accuracy of Spectral Mixture Analysis and the Normalized Difference Vegetation Index," *Remote Sensing of Environment*, vol. 73, pp. 87–102, 2000.

[5] Dar A. Roberts, Dale A. Quattrochi, Glynn C. Hulley, Simon J. Hook, and Robert O. Green, "Synergies between VSWIR and TIR data for the urban environment: An evaluation of the potential for the Hyperspectral Infrared Imager (HyspIRI) Decadal Survey mission," *Remote Sensing of Environment*, vol. 117, no. 2012, pp. 83–101, 2012.

[6] Conghe Song, "Spectral mixture analysis for subpixel vegetation fractions in the urban environment: How to incorporate endmember variability?," *Remote Sensing of Environment*, vol. 95, no. 2, pp. 248–263, 2005.

[7] Ben Somers, Gregory P. Asner, Laurent Tits, and Pol Coppin, "Endmember variability in Spectral Mixture Analysis: A review," *Remote Sensing of Environment*, vol. 115, no. 7, pp. 1603–1616, 2011.

[8] Alina Zare and K C Ho, "Endmember variability in hyperspectral analysis: Addressing spectral variability during spectral unmixing," *IEEE Signal Processing Magazine*, vol. 31, no. 1, pp. 95–104, 2014.

[9] Dar A. Roberts, M. Gardner, R. Church, Susan L. Ustin, and G. Scheer, "Mapping Chaparral in the Santa Monica Mountains Using Multiple Endmember Spectral Mixture Models," *Remote Sensing of Environment*, vol. 65, pp. 267–279, 1998.

[10] B. Somers, J. Verbesselt, E.M. Ampe, N. Sims, W.W. Verstraeten, and P. Coppin, "Spectral mixture analysis to monitor defoliation in mixed-aged Eucalyptus globulus Labill plantations in southern Australia using Landsat 5-TM and EO-1 Hyperion data," *International Journal of Applied Earth Observation and Geoinformation*, vol. 12, no. 4, pp. 270–277, 2010.

[11] E.B. Wetherley, D.A. Roberts, and J. McFadden, "Mapping spectrally similar urban materials at sub-pixel scales," *Remote Sensing of Environment*, vol. 195, pp. 170–183, 2017.

[12] Alina Zare, Changzhe Jiao, and Taylor Glenn, "Discriminative Multiple Instance Hyperspectral Target Characterization," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 40, no. 10, pp. 2342–2354, 2018.

[13] Susan Meerdink, James Bocinsky, Alina Zare, Nicholas Kroeger, Connor McCurley, Daniel Shats, and Paul Gader, "Multi-Target Multiple Instance Learning for Hyperspectral Target Detection," *arXiv e-prints*, p. arXiv:1909.03316, Sep 2019.

[14] James M. Bocinsky, "Learning multiple target concepts from uncertain, ambiguous data using the adaptive cosine estimator and spectral match filter," M.S. thesis, University of Florida, 2019.

[15] Oded Maron and Tomas Lozano-perez, "A Framework for Multiple-Instance Learning," in *Advances in neural information processing systems*, pp. 570–576. 1998.

[16] Abigail N. Schaaf, Philip E. Dennison, Gregory K. Fryer, Keely L. Roth, and Dar a. Roberts, "Mapping Plant Functional Types at Multiple Spatial Resolutions Using Imaging Spectrometer Data," *GIScience & Remote Sensing*, vol. 48, no. 3, pp. 324–344, 2011.