Semi-Supervised Classification of Hyperspectral Data Using Spectral Unmixing Concepts

Inmaculada Dópido, Jun Li, Antonio Plaza
Hyperspectral Computing Laboratory
University of Extremadura
Avda. de la Universidad s/n, 10003 Cáceres, Spain
{inmaculadadopido, junli, aplaza}@unex.es

Paolo Gamba
Telecommunications and Remote Sensing Laboratory
University of Pavia
Via Ferrata 1, 27100 Pavia, Italy
Email: paolo.gamba@unipv.it

***Abstract*— Spectral unmixing and classification have been widely used in the recent literature to analyze remotely sensed hyperspectral data. However, possible connections between semi-supervised classification and spectral unmixing concepts have been rarely investigated. In this work, we propose a new method to perform semi-supervised classification of hyperspectral images by exploiting the information retrieved with spectral unmixing. The proposed method integrates a well-established discriminative classifier (multinomial logistic regression) with different spectral unmixing chains, thus bridging the gap between unmixing and classification. Furthermore, the proposed method uses active learning when generating new unlabeled samples for classification. The proposed method is experimentally validated using real hyperspectral data sets, indicating that the combination of spectral unmixing and semi-supervised classification can lead to powerful new algorithms for hyperspectral data interpretation.**

I. Introduction

Spectral unmixing and classification are two active areas of research in hyperspectral data interpretation. On the one hand, spectral unmixing is a fast growing area in which many algorithms have been recently developed to retrieve pure spectral components (*endmembers*) and determine their abundance fractions in mixed pixels, which dominate hyperspectral images [1]. On the other hand, supervised hyperspectral image classification is a difficult task due to the unbalance between the high dimensionality of the data and the limited availability of labeled training samples in real analysis scenarios. While the collection of labeled samples is generally difficult, expensive and time-consuming, unlabeled samples can be generated in a much easier way [2]. This observation has fostered the idea of adopting semi-supervised learning techniques in hyperspectral image classification.

The area of semi-supervised learning has experienced a significant evolution in terms of the adopted models, which comprise complex generative models [3]–[6], self learning models [7], [8], multi-view learning models [9], [10], transductive support vector machines (SVMs) [11], [12], and graph-based methods [13]. A survey of semi-supervised learning algorithms is available in [14]. The aforementioned semi-supervised algorithms generally assume that a limited number of labeled samples are available *a priori*, and then enlarge the training set using unlabeled samples, thus allowing these approaches to address ill-posed problems. However, in order for this strategy to work, several requirements need to be met. First and foremost, the new (unlabeled) samples should be generated without significant cost/effort. Second, the number of unlabeled samples required in order for the semi-supervised classifier to perform properly should not be too high in order to avoid increasing computational complexity in the classification stage. In other words, as the number of unlabeled samples increases, it may be unbearable for the classifier to properly exploit all the available training samples due to computational issues. Further, if the unlabeled samples are not properly selected, these may confuse the classifier, thus introducing significant divergence or even reducing the classification accuracy obtained with the initial set of labeled samples. In order to address these issues, it is very important that the most highly informative unlabeled samples are identified in computationally efficient fashion, so that significant improvements in classification performance can be observed without the need to use a high number of unlabeled samples.

In this work, we develop a new approach to perform semi-supervised classification of hyperspectral images by exploiting the information retrieved with spectral unmixing. Our main goal is to synergize two of the most widely used approaches to interpret hyperspectral data into a unified framework which uses active learning techniques for automatically selecting unlabeled samples in semi-supervised fashion. Specifically, we use active learning to select highly informative unlabeled training samples in order to enlarge the initial (possibly very limited) set of labeled samples and perform semi-supervised classification based on the information provided by a well-established discriminative classifier (multinomial logistic regression [15]) and different spectral unmixing chains [16], [17]. Experimental results are conducted using the well-known Airborne Visible Infra-Red Imaging Spectrometer (AVIRIS [18]) scene over the Indian Pines test site, Indiana.

II. Proposed Approach

The proposed approach consists of three main ingredients: semi-supervised learning, spectral unmixing and active learning, which are described before introducing our method.

978-1-4673-2443-4/12/$31.00 ©2012 IEEE

A. Semi-Supervised Learning

For the semi-supervised part of our approach, we use the multinomial logistic regression (MRL) classifier [15]. In this work, we use the MLR classifier as it provides probabilistic outputs which play an essential role in our active learning process. Furthermore, we adopt a sparsity inducing prior on the regressors in order to obtain sparse estimates. As a result, most of the components of the regressors are zero. This allows us to control the complexity of our proposed techniques and their generalization capacity. Finally, we use LORSAL algorithm [21] to learn the MLR classifier as it is able to learn directly the posterior class distributions and deal with high dimensionality of hyperspectral data in a very effective way. This is very important for semi-supervised learning since in the end of the process we would like to include as much unlabeled samples as possible which brings difficulty for normal algorithms from the viewpoint of computational complexity.

B. Spectral Unmixing

Several spectral unmixing chains are explored in [16]. These chains are based on the well-known linear mixture model [1] and fully constrained least-squares linear spectral unmixing (FCLSU) [23]. However, the unmixing-based chains discussed in [16] do not include spatial information, which is an important source of information since hyperspectral images exhibit spatial correlation between image features. Also, the study in [16] suggested that partial unmixing using mixture-tuned matched filtering (MTMF) [24], [25] could be an effective solution to deal with the likely fact that not all pure spectral constituents in the scene (needed for spectral unmixing purposes) are known *a priori*, but a more exhaustive investigation of partial unmixing (in combination with spatial information) is desirable. In this work, we consider several unmixing strategies addressing the aforementioned issues as discussed in [17]:

1) *FLSU-Based Unmixing* (ST.1), which first assumes that the labeled samples are made up of spectrally pure constituents (endmmebers) and then calculates their abundances by means of the FCLSU method and provides a set of fractional abundance maps (one per labeled class).
2) *MTMF-Based Unmixing* (ST.2), which also assumes that the labeled samples are made up of spectrally pure constituents (endmmebers) but now calculates their abundances by means of the MTMF method, thus providing a set of fractional abundance maps (one per labeled class). In [16] it is shown that MTMF can outperform other techniques for abundance estimation such as FCLSU since it can provide meaningful abundance maps by means of partial unmixing in case not all endmembers are available *a priori*.
3) *Unsupervised Clustering followed by FCLSU* (ST.3), which is intended to solve the problems highlighted by endmember extraction algorithms which are sensitive to outliers and pixels with extreme values of reflectance. By using an unsupervised clustering method such as the k-means on the available labeled samples, the endmembers extracted (from class centers) are expected to be more spatially significant. Then, FLCSU-based unmixing is conducted using the resulting endmembers.
4) *Unsupervised Clustering followed by MTMF* (ST.4), which is exactly the same as the previous strategy (ST.3) but this time MTMF-based unmixing is conducted using the resulting endmembers after k-means clustering.

C. Active Learning

The third ingredient of our proposed method consists of using active learning to improve the selection of unlabeled samples for semi-supervised learning. In our proposed strategy, the candidate set for the active learning process (based on the available labeled and unlabeled samples) is inferred using spatial information (specifically, by applying a first-order spatial neighborhood on available samples) so that high confidence can be expected in the class labels of the obtained candidate set. This is similar to human interaction in supervised active learning, where the class labels are known and given by an expert. In a second step, we run active learning to select the most informative samples from the candidate set. This is similar to the machine interaction level in supervised active learning, where in both cases the goal is to find the samples with higher uncertainty. Due to the fact that we use a discriminative classifier (MLR) and spectral unmixing techniques, active learning algorithms which focus on the boundaries between the classes (which are often dominated by mixed pixels) are preferred. In this way, we can combine the properties of the probabilistic MLR classifier and spectral unmixing concepts to find the most suitable (complex) unlabeled samples for improving the classification results through the selected active learning strategy. It should be noted that many active learning techniques are available in the literature [26]. In this work, we use the well-known breaking ties (BT) [27] to evaluate the proposed approach. This algorithm finds the samples minimizing the distance between the first two most probable classes.

D. Proposed Hybrid Strategy

Let $\mathbf{x} \equiv (\mathbf{x}_1, \ldots, \mathbf{x}_n) \in \mathbb{R}^{d \times n}$ an image of d-dimensional feature vectors, $\mathbf{y} \equiv (y_1, \ldots, y_n)$ an image of labels. With this notation in mind, the proposed method can be describe as follows:

$$\widehat{p}_i(y_i = k|\mathbf{x}_i) = \alpha f_1(y_i = k|\mathbf{x}_i) + (1-\alpha) f_2(y_i = k|\mathbf{x}_i), \quad (1)$$

where $\widehat{p}_i(\cdot)$ is the joint estimate for the kth class, *i.e.*, $y_i = k$, obtained by the classification and unmixing methods given observation $\mathbf{x}_i$, where $\widehat{p}_i(\cdot)$ will serve as the indicator, *i.e.*, probability, for the semi-supervised active learning. In this work, function $f_1(\cdot)$ is the probability obtained by the classification algorithm, *i.e.*, MLR classifier; and function $f_2(\cdot)$ is the abundance fraction obtained from the spectral unmixing chains presented in subsection II-B. The degree between the classification probabilities and abundance fractions is controlled by parameter α, where $0 \leq \alpha \leq 1$. As shown in expression (1), if $\alpha = 1$, only classification probabilities are

considered by the proposed strategy. On the other hand, if $\alpha = 0$, only spectral unmixing is taken into account for the proposed strategy. Therefore, by tuning α from 0 to 1, we can adjust the impact between classification and unmixing methods. More, by introducing parameter α, the proposed hybrid strategy takes advantage from both classification and unmixing such that the new unlabeled samples selected are more informative in comparison with those samples selected only from classification or unmixing methods.

III. Experimental Results

In this section, we use real hyperspectral data to evaluate the proposed approach. In our experiments with the MLR classifier, we apply the Gaussian RBF kernel to a normalized version of the considered hyperspectral data set. In all cases, the reported figures of overall accuracy (OA), average accuracy (AA), κ statistic, and class individual accuracies are obtained by averaging the results obtained after conducting 10 independent Monte Carlo runs with respect to the labeled training set from the ground truth image, where the remaining samples are used for validation purposes.

In order to illustrate the good performance of the proposed approach, we use very small labeled training sets on purpose. As a result, the main difficulties that our proposed approach should circumvent can be summarized as follows. First and foremost, it is very difficult for supervised algorithms to provide good classification results as very little information is available about the class distribution. Poor generalization is also a risk when estimating class boundaries in scenarios dominated by limited training samples. Since our approach is semi-supervised, we take advantage of unlabeled samples in order to improve classification accuracy. However, if the number of labeled samples is very small, increasing the number of unlabeled samples could bias the learning process. This effect is explored in the remainder of this section, which is organized as follows. In subsection III-A we introduce the dataset used for evaluation purposes in this work. In subsection III-B, we describe the experiments conducted with this scene. In all cases, the results obtained by the supervised version of the considered classifier are also reported for comparative purposes.

A. Hyperspectral data set

The hyperspectral image used in experiments was collected by the AVIRIS sensor over the Indian Pines region in Northwestern Indiana in 1992. This scene, with a size of 145 lines by 145 samples, was acquired over a mixed agricultural/forest area, early in the growing season. The scene comprises 220 spectral channels in the wavelength range from 0.4 to 2.5 μm, nominal spectral resolution of 10 nm, moderate spatial resolution of 20 meters by pixel, and 16-bit radiometric resolution. After an initial screening, several spectral bands were removed from the data set due to noise and water absorption phenomena, leaving a total of 200 radiance channels to be used in the experiments. For illustrative purposes, Fig. 1(a) shows a false color composition of the AVIRIS Indian Pines scene, while Fig. 1(b) shows the ground-truth map available for the scene, displayed in the form of a class assignment for each labeled pixel, with 16 mutually exclusive ground-truth classes, in total, 10366 samples. These data, including ground-truth information, are available online[1], a fact which has made this scene a widely used benchmark for testing the accuracy of hyperspectral data classification algorithms. This scene constitutes a challenging classification problem due to the presence of mixed pixels in all available classes, and because of the unbalanced number of available labeled pixels per class.

B. Experiments with AVIRIS Indian Pines Data Set

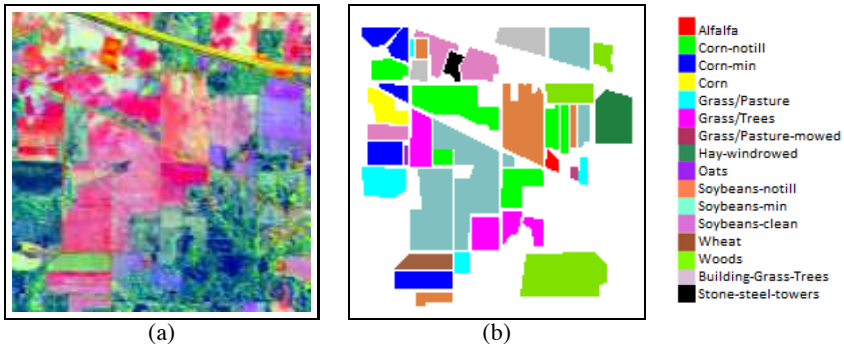
In this set of experiments, we evaluated the classification accuracy of the proposed approach using the AVIRIS Indian Pines data set in Fig. 1(a). The first step in our experiments was to carefully optimize parameter α controlling the impact between classification and unmixing.

In the Table I the overall classification accuracies [%] obtained when it is changed the parameter in AVIRIS Indian Pines hyperspectral data set with 5 labeled samples per class are reported. The total number of iterations are given in the parentheses.After an extensive set of experiments, we empirically set $\alpha = 0.75$ meaning that classification should have more weight than unmixing in order to achieve optimal results(Table I).

With this parameter setting in mind, Table II shows the overall and average classification accuracies (in percentage) and the κ statistic obtained by the supervised strategy –based on the MLR classifier trained using only $\{5, 10, 15\}$ labeled samples per class– and by the proposed semi-supervised approach, using two different strategies for active learning: random sampling (RS) and the BT algorithm (both executed using 300 iterations to actively select 300 unlabeled samples). Finally, we also report the classification results obtained using the hybrid classification-unmixing approach with the MLR classifier and four spectral unmixing chains (with $\alpha = 0.75$). It should be noted that the selected classification scenario represents a very challenging one. For instance, when 5 labeled samples are used per class, only 80 labeled samples in total are assumed to be available as the initial condition for the considered classifier, which is much lower than the number of spectral bands available in the scene.

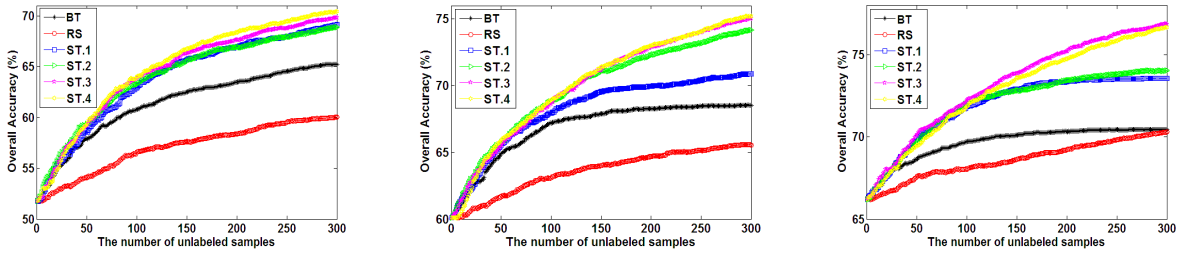
Several conclusions can be obtained from Table II. First of all, it is remarkable that the inclusion of unlabeled samples significantly improved the classification results in all cases. For the supervised classifier with only 10 labeled samples per class (see Table II), an OA of 60.12% was obtained. When the proposed strategy (combining classification and spectral unmixing) was used, the classification accuracies improved to an OA of 75.29% (ST.4) which represents a significant improvement over the supervised case. As shown by Table II, the classification accuracies increase as the number of labeled training samples increases. This is expected, since

[1] Available online: http://dynamo.ecn.purdue.edu/biehl/MultiSpec



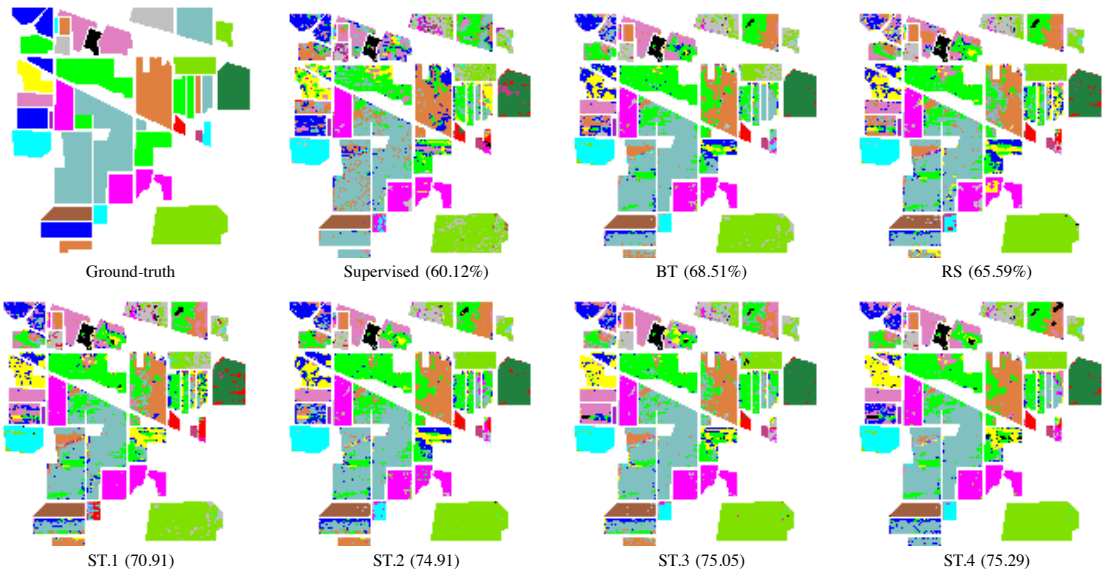
(a) (b)

Fig. 1. (a) False color composition of the AVIRIS Indian Pines scene. (b) Ground truth-map containing 16 mutually exclusive land-cover classes (right).



(a) 5 labeled samples (b) 10 labeled samples (c) 15 labeled samples

Fig. 2. Overall classification accuracies (as a function of the number of unlabeled samples) obtained for the AVIRIS Indian Pines data set by different classifiers. BT is the semi-supervised classifier where unlabeled samples are selected using breaking ties. RS is the semi-supervised classifier where unlabeled samples are selected using random sampling. Finally, ST.1 to ST.4 denote the semi-supervised hybrid classifier integrating classification and spectral unmixing (with $\alpha = 0.75$), where unlabeled samples are selected using BT.



Ground-truth Supervised (60.12%) BT (68.51%) RS (65.59%)

ST.1 (70.91) ST.2 (74.91) ST.3 (75.05) ST.4 (75.29)

Fig. 3. Classification maps and overall classification accuracies (in the parentheses) obtained after applying different classifiers to the AVIRIS Indian Pines data set. In all cases the number of labeled samples was 10, and the number of unlabeled samples (used in the semi-supervised strategies: BT, RS, ST.1, ST.2, ST.3 and ST.4) was set to 300.

TABLE I

OVERALL CLASSIFICATION ACCURACIES [%] OBTAINED WHEN IT IS CHANGED THE PARAMETER IN AVIRIS INDIAN PINES HYPERSPECTRAL DATA SET WITH 5 LABELED SAMPLES PER CLASS. THE TOTAL NUMBER OF ITERATIONS ARE GIVEN IN THE PARENTHESES.

AVIRIS Indian Pines					
Number of labeled samples per class ($l = 5$)					
Parameter (ite.)	1.0	0.75	0.50	0.25	0.00
ST. 1 (100)	60.82	62.99	55.25	54.24	53.80
ST. 2 (100)	60.82	63.29	63.15	63.95	55.87
ST. 3 (100)	60.82	63.95	52.37	56.21	55.29
ST. 4 (100)	60.82	64.04	62.75	58.55	55.57

TABLE II

OVERALL, AVERAGE CLASSIFICATION ACCURACIES [%], AND κ STATISTIC OBTAINED USING DIFFERENT CLASSIFIERS WHEN APPLIED TO THE AVIRIS INDIAN PINES HYPERSPECTRAL DATA SET.

	5 labeled samples per class						
	Supervised	BT	RS	ST.1	ST.2	ST.3	ST.4
OA	51.78	65.25	60.03	69.10	68.89	69.83	70.45
AA	63.82	69.98	66.31	72.73	70.81	71.37	72.32
κ	46.26	60.65	54.81	64.82	64.54	65.58	66.25
	10 labeled samples per class						
	Supervised	BT	RS	ST.1	ST.2	ST.3	ST.4
OA	60.12	68.51	65.59	70.91	74.91	75.05	75.29
AA	71.74	75.66	73.49	77.13	78.21	78.44	79.05
κ	55.43	64.40	61.17	67.08	70.69	71.60	71.84
	15 labeled samples per class						
	Supervised	BT	RS	ST.1	ST.2	ST.3	ST.4
OA	66.20	70.46	70.29	73.57	74.04	76.89	76.68
AA	77.39	79.47	78.76	80.86	79.76	81.90	81.94
κ	62.09	66.65	66.41	70.16	70.51	73.71	73.53

the uncertainty of the classifier boundaries decreases as more labeled samples are used in the supervised case.

In a second experiment we evaluated the impact of the number of unlabeled samples on the classification performance achieved by the two considered probabilistic classifiers. Fig. 2 shows the OAs in classification accuracy obtained by the supervised MLR (trained using only 5, 10 and 15 labeled samples per class) and by the proposed approach (based on the same classifier plus spectral unmixing) using the four considered strategies for including unmixing information and as a function of the number of unlabeled samples. For the active learning part we considered again two strategies: RS and BT. The plots in Figs. 2 reveal clear advantages of using unlabeled samples for the proposed approach when compared with the supervised algorithm alone. In all cases, the proposed unmixing strategies significantly outperform the corresponding supervised algorithm, and the increase in performance is more relevant as the number of unlabeled samples increases. These unlabeled samples are automatically selected by the proposed approach, and represent no cost in terms of data collection or human supervision. In Fig. 2 it can also be seen that using an intelligent active learning algorithm such BT greatly improved the obtained accuracies in comparison with simple random selection (RS).

For illustrative purposes, Fig. 3 shows some of the classification maps obtained for the AVIRIS Indian Pines scene. These classification maps correspond to one of the 10 Monte-Carlo runs that were averaged in order to generate the classification scores reported in Table II. The advantages obtained by adopting a semi-supervised learning approach which combines classification and unmixing concepts can be clearly appreciated in the classification maps displayed in Figs. 3, which also report the classification OAs obtained for each method in the parentheses.

IV. CONCLUSIONS AND FUTURE RESEARCH LINES

In this paper, we have developed a new semi-supervised method which integrates the probabilistic results provided by a state-of-the-art classifier such as the MLR and the fractional abundances provided by different spectral unmixing chains. This represents an innovative contribution since classification and spectral unmixing have been widely used in the hyperspectral analysis literature, but have been rarely exploited in combined fashion. In this regard, our synergistic approach integrates both sources of information (unmixing and classification) and allows controlling the importance of each source in the final results obtained. Furthermore, the proposed method uses active learning when generating new unlabeled samples for classification which allows the exploitation of the proposed classifier in semi-supervised fashion. Our experimental results, conduced using the well-known AVIRIS Indian Pines scene, indicate that the proposed hybrid strategy exhibits significant potential as it can greatly increase the classification accuracies obtained in the supervised case through the incorporation of information about the sub-pixel composition of pixels in the hyperspectral image, and also with the adoption of unlabeled samples which can be obtained with very little cost and effort.

In the future, we will explore additional strategies to generate unlabeled samples through active learning and also consider additional spectral unmixing chains and probabilistic classifiers that can be easily integrated in the proposed framework. Also, we will explore in more detail the optimal setting of parameter α (which controls the weight of classification to spectral unmixing) in different application domains. Finally, we will also include experiments with additional synthetic and real scenes to confirm the very encouraging results obtained by the proposed methodology in this work.

REFERENCES

[1] J. M. Bioucas-Dias, A. Plaza, N. Dobigeon, M. Parente, Q. Du, P. Gader, and J. Chanussot, "Hyperspectral unmixing overview: Geometrical, statistical and sparse regression-based approaches," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 5, no. 2, pp. 354–379, 2012.

[2] F. Bovolo, L. Bruzzone, and L. Carline, "A novel technique for subpixel image classification based on support vection machine," *IEEE Transactions on Image Processing*, vol. 19, pp. 2983–2999, 2010.

[3] S. Baluja, "Probabilistic modeling for face orientation discrimination: Learning from labeled and unlabeled data," in *Neural Information Procesing systems (NIPS '98)*, 1998.

[4] A. Blum and T. Mitchell, "Combining labeled and unlabeled data with co-training." Morgan Kaufmann Publishers, 1998, pp. 92–100.

[5] T. Mitchell, "The role of unlabeled data in supervised learning," in *Proceedings of the Sixth International Colloquium on Cognitive Science*, 1999.

[6] A. Fujino, N. Ueda, and K. Saito, "A hybrid generative/discriminative approach to semi-supervised classifer degin," in *AAAI'05 Proceedings of the 20th national coference on Artificial intelligence*, vol. 2, 2005.
[7] D. Yarowsky, "Unsupervised word sense disambiguation rivaling supervised methods," in *Proceedings of the 33rd annual meeting on Association for Computational Linguistics*, ser. ACL'95, 1995, pp. 189–196.
[8] C. Rosenberg, M. Hebert, and H. Schneiderman, "Semi-supervised self-training of object detection models," in *Seventh IEEE Workshop on Applications of Computer Vision*, January 2005.
[9] I. Cowan, T. G, and V. R. D. Sa, "Learning classification with unlabeled data," 1994.
[10] U. Brefeld, T. Gärtner, T. Scheffer, and S. Wrobel, "Efficient co-regularised least squares regression," in *Proceedings of the 23rd international conference on Machine learning*, ser. ICML'06, 2006, pp. 137–144.
[11] V. Vapnik, *Statistical Learning Theory*. New York: Wiley, 1998.
[12] T. Joachims, "Transductive inference for text classification using support vector machines," in *Proceedings of the Sixteenth International Conference on Machine Learning*, ser. ICML'99, 1999, pp. 200–209.
[13] A. Blum and S. Chawla, "Learning from labeled and unlabeled data using graph mincuts," in *Proceedings of the Eighteenth International Conference on Machine Learning*, ser. ICML'01, 2001, pp. 19–26.
[14] X. Zhu, "Semi-supervised learning literature survey," Computer Sciences, University of Wisconsin-Madison, Tech. Rep. 1530, 2005.
[15] D. Böhning, "Multinomial logistic regression algorithm," *Annals of the Institute of Statistical Mathematics*, vol. 44, pp. 197–200, 1992.
[16] I. Dopido, M. Zortea, A. Villa, A. Plaza, and P. Gamba, "Unmixing prior to supervised classification of remotely sensed hyperspectral images," *IEEE Geoscience and Remote Sensing Letters*, vol. 8, pp. 760–764, 2011.
[17] I. Dopido, A. Villa, A. Plaza, and P. Gamba, "A quantitative and comparative assessment of unmixing-based feature extraction techniques for hyperspectral image classification," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 5, no. 2, pp. 421–435, 2012.
[18] R. O. Green, "Imaging spectroscopy and the airborne visible-infrared imaging spectrometer (AVIRIS)," *Remote Sensing of Environment*, vol. 65, pp. 227–248, 1998.
[19] G. Camps-Valls and L. Bruzzone, "Kernel-based methods for hyperspectral image classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, pp. 1351–1362, 2005.
[20] B. Krishnapuram, L. Carin, M. Figueiredo, and A. Hartemink, "Sparse multinomial logistic regression: Fast algorithms and generalization bounds," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 6, pp. 957–968, 2005.
[21] J. Bioucas-Dias and M. Figueiredo, "Logistic regression via variable splitting and augmented Lagrangian tools," Instituto Superior Técnico, TULisbon, Tech. Rep., 2009.
[22] J. Li, J. Bioucas-Dias, and A. Plaza, "Hyperspectral image segmentation using a new Bayesian approach with active learning," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 19, pp. 3947–3960, 2011.
[23] D. Heinz and C.-I. Chang, "Fully constrained least squares linear mixture analysis for material quantification in hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, pp. 529–545, 2001.
[24] C.-I. Chang, J.-M. Liu, B.-C. Chieu, C.-M. Wang, C. S. Lo, P.-C. Chung, H. Ren, C.-W. Yang, and D.-J. Ma, "Generalized constrained energy minimization approach to subpixel target detection for multispectral imagery," *Optical Engineering*, vol. 39, pp. 1275–1281, 2000.
[25] J. Boardman, "Leveraging the high dimensionality of AVIRIS data for improved subpixel target unmixing and rejection of false positives: mixture tuned matched filtering," *Proceedings of the 5th JPL Geoscience Workshop*, pp. 55–56, 1998.
[26] D. Tuia and G. Camps-Valls, "Urban image classification with semisupervised multiscale cluster kernels," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 4, no. 1, pp. 65 –74, march 2011.
[27] T. Luo, K. Kramer, D. B. Goldgof, S. Samson, A. Remsen, T. Hopkins, and D. Cohn, "Active learning to recognize multiple types of plankton," *Journal of Machine Learning Research*, pp. 589–613, 2005.