A NEW MULTIPLE CLASSIFIER SYSTEM FOR SEMI-SUPERVISED ANALYSIS OF HYPERSPECTRAL IMAGES

Jun Li[1], Prashanth Reddy Marpu[2], Antonio Plaza[1], Jose Manuel Bioucas Dias[3] and Jon Atli Benediktsson[2]

[1]*Hyperspectral Computing Laboratory, University of Extremadura, Avda. de la Universidad s/n, 10003 Caceres, Spain*

[2]*Faculty of Electrical and Computer Engineering, University of Iceland, 101 Reykjavik, Iceland*

[3]*Instituto de Telecomunicações, Instituto Superior Técnico, Av. Rovisco Pais, 1049-001 Lisbon, Portugal*

Keywords: Multiple classifier system, Remotely sensed hyperspectral images, Semi-supervised learning.

Abstract: In this work, we propose a new semi-supervised algorithm for remotely sensed hyperspectral image classification which belongs to the family of multiple classifier systems. The proposed approach combines the output of two well-established discriminative classifiers: sparse multinomial logistic regression (SMLR) and quadratic discriminant analysis (QDA). Our approach follows a two-step strategy. First, both SMLR and QDA are trained from the same set of labeled training samples and make predictions for the unlabeled samples in the image. Second, the set of unlabeled training samples is expanded by combining the estimates obtained by both classifiers in the previous step. The effectiveness of the proposed method is evaluated via experiments with a widely used hyperspectral image, collected by the Airborne Visible Infra-Red Imaging Spectrometer (AVIRIS) over the Indian Pines region in Indiana. Our results indicate that the proposed multiple classifier method provides state-of-the-art performance when compared to other methods.

1 INTRODUCTION

The recent availability of remotely sensed hyperspectral images in different application domains has fostered the development of techniques able to interpret this kind of high-dimensional data in many different contexts (Landgrebe, 2003). Specifically, many techniques for hyperspectral image classification have been proposed in recent years, with the ultimate goal of taking advantage of the detailed information contained in hyperspectral pixel vectors (spectral signatures) to generate thematic maps (Plaza et al., 2009). A relevant challenge for supervised classification techniques (which assume prior knowledge in the form of class labels for some spectral signatures) is the limited size of training sets, since their collection generally involves expensive ground campaigns. As a result, there is an unbalance between the number of available training samples and the high dimensionality of the hyperspectral data, which potentially results in the Hughes phenomenon (Landgrebe, 2003). The development of techniques intended to address this phenomenon constitutes a very active area of research in hyperspectral image classification (Camps-Valls and Bruzzone, 2005; Fauvel et al., 2008).

While the collection of labeled samples is generally difficult, expensive and time-consuming, unlabeled samples can be generated in a much easier way. This observation fostered the idea of semi-supervised learning (Zhu, 2005), which has recently become a very active area of research in hyperspectral image classification (Camps-Valls et al., 2006; Tuia and Camps-Valls, 2009; Camps-Valls et al., 2007; Li et al., 2009; Velasco-Forero and Manian, 2009; Li et al., 010c; Li et al., 010a). The main assumption of these techniques is that new (unlabeled) training samples can be obtained from the (limited) set of available labeled samples without significant effort/cost, and without the need to design a ground campaign (Krishnapuram et al., 2004). Most available semi-supervised learning algorithms use some type of regularization which encourages that "similar" features belong to the same class. The effect of this regularization is to push the boundaries between classes towards regions with low data density (Chapelle et al., 2006), where a rather usual way of building such regularizers is to associate the vertices of a graph with the complete set of samples and then build the regularizer depending on the variables defined on such vertices.

Li J., Reddy Marpu P., Plaza A., Manuel Bioucas Dias J. and Atli Benediktsson J..
A NEW MULTIPLE CLASSIFIER SYSTEM FOR SEMI-SUPERVISED ANALYSIS OF HYPERSPECTRAL IMAGES.
DOI: 10.5220/0003849504060411
In *Proceedings of the 1st International Conference on Pattern Recognition Applications and Methods* (PRARSHIA-2012), pages 406-411
ISBN: 978-989-8425-98-0

Although a good performance can in general be expected from these methods, difficulties arise from the viewpoint of the complexity of the model and its high computational cost, which call for new developments in this area.

In this paper, we present a new multiple classifier system (Xu et al., 1992; Kittler and Roli, 2000) intended to specifically cope with the limited availability of training samples. The proposed semi-supervised algorithm is based on two discriminative classifiers: sparse multinomial logistic regression (SMLR) (Krishnapuram et al., 2005) and quadratic discriminant analysis (QDA) (Duda et al., 200). Owing to their rather different structures, these two classifiers exhibit complementary discriminatory capabilities, a necessary condition for the success of any multiple classifier system. In this work, we adopt a decision-directed semi-supervision strategy (Roli, 2005). First, SMLR and QDA are trained using labeled samples. Second, a set of unlabeled training samples is created and expanded by taking advantage of the estimates obtained in the previous step. Specifically, we implement a simple strategy to enlarge the available training set; our assumption is that, if both classifiers obtain the same estimate of the class label for a given pixel, then this sample is included into the set of unlabeled samples and the procedure is repeated again, thus growing the unlabeled training set without significant effort. The effectiveness of the proposed method is evaluated via experiments with a well-known hyperspectral image collected by the Airborne Visible Infra-Red Imaging Spectrometer (AVIRIS) (Green et al., 1998) over the Indian Pines region in Indiana.

The remainder of the paper is organized as follows. Section 2 introduces the two considered discriminative classifiers. Section 3 presents the proposed multiple classifier system for semi-supervised learning based on the aforementioned classifiers. Section 4 reports classification results on the considered AVIRIS hyperspectral data set. Section 5 concludes with some remarks and future lines.

2 BASE CLASSIFIERS

First, we briefly define the mathematical notations adopted hereinafter. Let $\mathcal{K} \equiv \{1,\ldots,K\}$ denote a set of K class labels, $\mathcal{S} \equiv \{1,\ldots,n\}$ a set of integers indexing the n pixels of an image, $\mathbf{x} \equiv (\mathbf{x}_1,\ldots,\mathbf{x}_n)$ an image of d-dimensional feature vectors, $\mathbf{y} \equiv (y_1,\ldots,y_n)$ an image of labels, $\mathcal{D}_L \equiv \{(y_1,\mathbf{x}_1),\ldots,(y_L,\mathbf{x}_L)\}$ a set of labeled samples, $\mathcal{Y}_L \equiv \{y_1,\ldots,y_L\}$ the set of labels in $\mathcal{D}_L$, $\mathcal{X}_L \equiv \{\mathbf{x}_1,\ldots,\mathbf{x}_L\}$ the set of feature vectors in $\mathcal{D}_L$, $\mathcal{D}_U^* \equiv \{(y_{L+1}^*,\mathbf{x}_{L+1}),\ldots,(y_{L+U}^*,\mathbf{x}_{L+U})\}$ a set of pairs made up of unlabeled feature vectors and pseudo-labels obtained from the base classifiers, as described below. Next, we revisit the SMLR (Krishnapuram et al., 2005) and QDA classifiers with the objective of developing a new multiple classifier system.

2.1 Sparse Multinomial Logistic Regression (SMLR)

The multinomial logistic regression (MLR) (Böhning, 1992) models the posterior class probabilities as

$$p(y_i = k|\mathbf{x}_i,\boldsymbol{\omega}) \equiv \frac{\exp(\boldsymbol{\omega}^{(k)}\mathbf{h}(\mathbf{x}_i))}{\sum_{k=1}^{K}\exp(\boldsymbol{\omega}^{(k)}\mathbf{h}(\mathbf{x}_i))}, \qquad (1)$$

where $\mathbf{h}(\mathbf{x}_i) \equiv [h_1(\mathbf{x}_i),...,h_l(\mathbf{x}_i)]^T$ is a vector of l fixed functions of the input, often termed features; $\boldsymbol{\omega}$ denotes the regressors and $\boldsymbol{\omega} \equiv [\boldsymbol{\omega}^{(1)^T},...,\boldsymbol{\omega}^{(K-1)^T}]^T$. Since the density (1) does not depend on translations on the regressors $\boldsymbol{\omega}^{(k)}$, in this work we have $\boldsymbol{\omega}^{(K)} = \mathbf{0}$. It should be noted that the function $\mathbf{h}$ may be linear (*i.e.*, $\mathbf{h}(\mathbf{x}_i) = [1,x_{i,1},...,x_{i,d}]^T$, where $x_{i,j}$ is the j-th component of $\mathbf{x}_i$) or nonlinear. A kernel is some symmetric function which offers a mechanism to deal with the nonlinear case, *i.e.*, $\mathbf{h}(\mathbf{x}_i) = [1,K_{\mathbf{x},\mathbf{x}_1},...,K_{\mathbf{x},\mathbf{x}_l}]^T$, where $K_{\mathbf{x}_i,\mathbf{x}_j} = K(\mathbf{x}_i,\mathbf{x}_j)$ and $K(\cdot,\cdot)$. Kernels have been largely used in this context since they tend to improve data separability in a transformed space. In this work, we use the Gaussian Radial Basis Function (RBF) kernel: $K(\mathbf{x},\mathbf{z}) \equiv -\exp(\|\mathbf{x}-\mathbf{z}\|^2/(2\sigma^2))$, which has been widely used in hyperspectral image classification (Camps-Valls and Bruzzone, 2005). In order to control the machine complexity and, thus, its generalization capacity, the SMLR algorithm introduced in (Krishnapuram et al., 2005) models $\boldsymbol{\omega}$ as a random vector with Laplacian density and computes its maximum a posteriori (MAP) estimate.

If the features are given by kernels, then the original SMLR algorithm is limited to data sets with products $L \times K$ not larger than, say, 1000. Therefore, most hyperspectral data sets are beyond the reach of this algorithm. This difficulty was removed by the introduction of the LORSAL algorithm (Bioucas-Dias and Figueiredo, 2009), which is able to deal with training set sizes in the order of a few thousands, regardless of the number of classes. LORSAL plays a central role, for example, in (Li et al., 010c; Li et al., 010b).

2.2 Quadratic Discriminant Analysis (QDA)

QDA has been widely used in pattern recognition applications (Bishop, 2007). The concept can be simply explained as follows. Let $\mu^{(k)}$ and $\sigma^{(k)}$ be the mean vector and covariance matrix of a given class k, then the decision function for the QDA classifier is given by:

$$\begin{array}{l}\log p(y_i = k|\mathbf{x}_i, \mu^{(k)}, \sigma^{(k)}) \equiv \\ -\frac{1}{2}(\mathbf{x}_i - \mu^{(k)})^T (\sigma^{(k)})^{-1}(\mathbf{x}_i - \mu^{(k)}) \\ -\frac{1}{2}\log|\sigma^{(k)}| + \log p(y_i = k) + C,\end{array} \quad (2)$$

where $p(y_i = k)$ is the prior probability of class k, and C is an additive constant.

3 MULTIPLE CLASSIFIER SYSTEM

In this section, we present a new multiple classifier system which exploits unlabeled training samples generated using both SMLR and QDA classifiers. In our setup, we run the classifiers in iterative fashion. First, SMLR and QDA are trained by using the same set of labeled samples. Second, we form and increase a set of unlabeled samples by fusing the results obtained (in consensus) by both classifiers. This procedure has similarities with active data selection (Krishnapuram et al., 2004; Mackay, 1992), in which unlabeled samples are sequentially selected according to a given criterion. In the proposed approach, we actively increase the unlabeled training set based on fusing the results which are obtained in agreement by both SMLR and QDA.

Let $\mathcal{D}_{L+U} \equiv \{\mathcal{D}_L, \mathcal{D}_U^*\}$ be a joint training set made up of labeled and unlabeled samples. Similarly, let $\mathcal{S}_U$ be the set of unlabeled image pixels, and let $\widehat{\mathbf{y}}_{\text{SMLR}}$ and $\widehat{\mathbf{y}}_{\text{QDA}}$ be the classification results obtained by the SMLR and QDA classifiers, respectively. The basic strategy of the proposed method can be simply described as follows. For any given pixel $\{\mathbf{x}_i, i \in \mathcal{S}_U\}$, if both SMLR and QDA obtain the same class label k, *i.e.*, $\widehat{y}_{i\text{SMLR}} = \widehat{y}_{i\text{QDA}} = k$, then we increment the unlabeled set $\mathcal{D}_U$ by assigning k to y_i, *i.e.*, $y_i = k$, and $\mathcal{D}_U^* \equiv \{\mathcal{D}_U^*, (\mathbf{x}_i, y_i)\}$.

A pseudo-code for the proposed semi-supervised algorithm is shown in Algorithm 1, where u denotes the number of unlabeled samples selected in each iteration and *stopping criterion* denotes the criterion used to terminate the semi-supervised algorithm, *e.g*, a maximum number of unlabeled samples. Line 2 combines the labeled and unlabeled samples as a joint training set. Lines 3 and 4 compute the classification estimates using the SMLR and QDA classifiers, respectively. In line 5, function $\mathcal{F}(\cdot)$ selects u unlabeled samples from the unlabeled set $\mathcal{S}_U$ based on the classification results obtained by the SMLR and QDA, according to the proposed semi-supervised strategy. Let U_{all} denote the number of samples in which a consensus was achieved by both classifiers, *i.e.*, $U_{\text{all}} = \sum(\widehat{\mathbf{y}}_{\text{SMLR}} = \widehat{\mathbf{y}}_{\text{QDA}})$. In general, at each iteration the size of U_{all} is quite large. In this work, other than using all of the available U_{all} unlabeled samples, we resort to an iterative scheme mainly due to two reasons: (i) computational complexity: the cost of the learning stage is related to the number of training samples), and (ii) balance: there is no parameter intended to control the trade-off between the number of labeled and unlabeled samples, hence an iterative scheme can balance the impact of using unlabeled samples in case of poor generalization ability. This often happens when very few labeled samples are used. In practice, we empirically set $u \leq L$, which leads to good performance results as it will be shown in the following section.

Algorithm 1: The proposed multiple classifier-based semi-supervised algorithm.

Require: $\mathbf{x}$, $\mathcal{D}_L$, $\mathcal{D}_U^*$, $\mathcal{S}_U$, u

1: **repeat**
2: $\quad\mathcal{D}_{L+U} \equiv \{\mathcal{D}_L, \mathcal{D}_U^*\}$
3: $\quad\widehat{\mathbf{y}}_{\text{MLR}} := \text{MLR classifier}(\mathbf{x}, \mathcal{D}_{L+U})$
4: $\quad\widehat{\mathbf{y}}_{\text{QDA}} := \text{QDA classifier}(\mathbf{x}, \mathcal{D}_{L+U})$
5: $\quad\mathcal{D}_u \equiv \mathcal{F}(\widehat{\mathbf{y}}_{\text{MLR}}, \widehat{\mathbf{y}}_{\text{QDA}}, \mathcal{S}_U), u$
6: $\quad\mathcal{D}_U^* = \mathcal{D}_U^* + \mathcal{D}_u$
7: $\quad\mathcal{S}_U = \mathcal{S}_U - \{1, \ldots, u\}$
8: **until** some *stopping criterion* is met

4 EXPERIMENTAL RESULTS

The well-known AVIRIS Indian Pines scene was used in our experiments. The data were collected over Northwestern Indiana in June of 1992 (Landgrebe, 2003), and contains 145×145 pixels and 220 spectral bands. A total of 20 bands were removed prior to experiments due to noise and water absorption in those channels. The ground-truth data contains 16 mutually exclusive classes, and a total of 10366 labeled pixels. This image is a classic benchmark to validate the accuracy of hyperspectral image analysis algorithms and constitutes a challenging problem due to the significant presence of mixed pixels in all available classes, and also because of the unbalanced number of available labeled pixels per class. In this

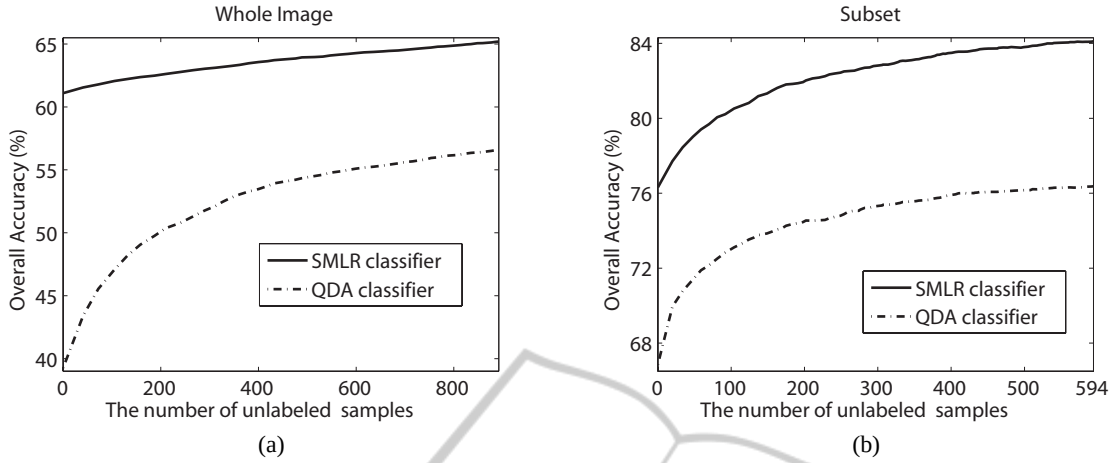


Figure 1: (a) OA results (as a function of the number of unlabeled samples) obtained after 100 Monte Carlo runs using the proposed multiple classifier strategy on the whole AVIRIS Indian Pines hyperspectral image; (b) The same OA results for the subset.

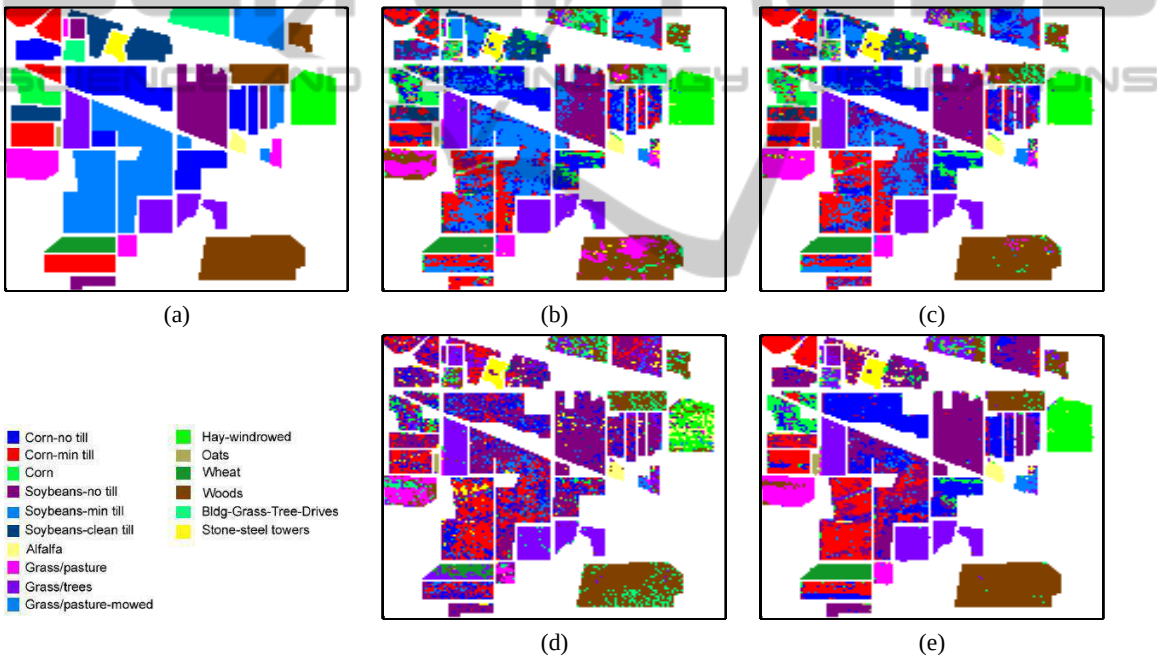
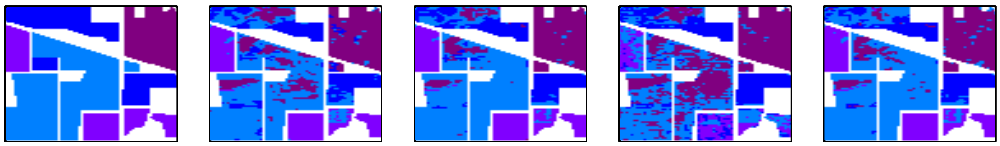


Figure 2: Classification maps obtained for the whole AVIRIS Indian Pines image by using 160 labeled samples and 892 unlabeled samples. (a) Ground-truth image. (b) Supervised SMLR classifier: OA = 61.2%. (c) Semi-supervised SMLR classifier with the proposed multiple classifier strategy: OA = 65.1%. (d) Supervised QDA classifier: OA = 40.7%. (e) Semi-supervised QDA classifier with the proposed multiple classifier strategy: OA = 58.2%.

work, we consider two scenes: (a) the full image with 16 classes, and (b) a subset (consisting of pixels in columns [27-94] and rows [31-116]) with total size of 68×86 pixels and 4 ground-truth classes. In our experiments, labeled training samples are randomly selected from the ground-truth data, whereas the remaining samples are used as the validation set. In order to increase the statistical significance of the results, each value of overall accuracy (OA) reported in this work is obtained as the average of 100 Monte Carlo runs.

Prior to the analysis of our results, we would like to emphasize that, for the SMLR classifier, we use all available spectral bands. However, for the QDA classifier feature extraction needs to be applied. This is because the QDA requires that the number of labeled



(a) (b) (c) (d) (e)

Figure 3: Classification maps obtained for the subset by using 20 labeled samples and 594 unlabeled samples. (a) Ground-truth image. (b) Supervised SMLR classifier: OA = 76.1%. (c) Semi-supervised SMLR classifier with the proposed multiple classifier strategy: OA = 84.9%. (d) Supervised QDA classifier: OA = 63.1%. (e) Semi-supervised QDA classifier with the proposed multiple classifier strategy: OA = 82.3%.

samples per class be larger than the dimensionality of the feature space [see (3)]. For dimensionality reduction, we use the hyperspectral signal identification by minimum error (HySime) algorithm (Dias and Nascimento, 2008). However, other feature extraction techniques can also be used.

Fig. 1 illustrates the obtained OA results as a function of the number of unlabeled training samples in the following cases: (a) 160 labeled samples (10 samples per class, which constitutes a very low number) for the whole AVIRIS Indian Pines image; and (b) 20 labeled samples (5 samples per class, which is even lower) for the subset. Notice that both (a) and (b) constitute very difficult problems, both for supervised techniques (as very few labeled samples are available and poor generalization capability is expected) and for semi-supervised techniques (since the trade-off between a large number of unlabeled samples versus a small number of labeled samples could bias the learning process). In both cases, we adopted a multiple classification strategy in which one of the classifiers (either SMLR or QDA) was used as a baseline and then the proposed multiple classifier-based strategy was used to generate additional unlabeled samples so that the classifiers can benefit from additional unlabeled samples. Several conclusions can be drawn from the experimental results reported in Fig. 1:

- First and foremost, it can be seen that the proposed multiple classifier-based strategy increases the OA for both classifiers. This is particularly the case when the QDA classifier is used as a baseline. For this classifier, the incorporation of additional unlabeled samples by means of the proposed strategy greatly improves the final classification results.
- Second, it can be seen that the SMLR classifier can significantly increase the OA results by including unlabeled training samples. For example, an OA higher than 65% was obtained for the whole image by using only 160 labeled samples and around 900 unlabeled samples. For the subset, only 5 labeled samples per class and around 600 unlabeled samples in total allowed increasing the OA to around 84%.
- Another important observation is that, with the considered number of unlabeled samples, none of the two considered classifiers converged in terms of the achieved OAs. This indicates that both methods can still benefit from the inclusion of additional unlabeled samples, thus leaving an open path for future developments of the method.

For illustrative purposes, Figs. 2 and 3 show the obtained classification maps along with the respective ground-truth images in the considered experiments. The improvements of the proposed multiple classifier-based strategy to each of the baseline methods can be observed in the results reported in these figures.

5 CONCLUSIONS

In this work, we have proposed a simple strategy for incorporating additional unlabeled samples in semi-supervised hyperspectral image classification by means of the consensus of multiple classifiers. The proposed system has been validated using two discriminative classifiers: (i) sparse multinomial logistic regression (SMLR), and (ii) quadratic discriminant analysis (QDA). The proposed approach is simple yet highly effective, as illustrated by our experimental results conducted with the famous AVIRIS Indian Pines dataset. Compared to the baseline supervised classifiers, the proposed semi-supervised method has the potential to greatly improve classification accuracies with very little effort, by simply including additional unlabeled samples after the consensus of the considered classifiers. This strategy is applicable to other classifiers but has been tested in this work only with the SMLR and QDA to illustrate its potential. In the future, additional classifiers and data sets will be used in the experimental validation of our proposed multiple classifier system. Also, we will target efficient mechanisms for exploiting the unlabeled information in a more efficiently way, *e.g.* by means of active learning. A more detailed evaluation of the computational complexity of the proposed approach (includ-

ing potential mechanisms able to reduce such cost) will be also explored in future developments of the method.

REFERENCES

Bioucas-Dias, J. and Figueiredo, M. (2009). Logistic regression via variable splitting and augmented lagrangian tools. Technical report, Instituto Superior Tecnico, TULisbon.

Bishop, C. M. (2007). *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Springer, 1st edition.

Böhning, D. (1992). Multinomial logistic regression algorithm. *Annals of the Institute of Statistical Mathematics*, 44:197–200.

Camps-Valls, G., Bandos, T., and Zhou, D. (2007). Semi-supervised graph-based hyperspectral image classification. *IEEE T. Geoscience and Remote Sensing*, 45(10):3044–3054.

Camps-Valls, G. and Bruzzone, L. (2005). Kernel-based methods for hyperspectral image classification. *IEEE T. Geoscience and Remote Sensing*, 43(6):1351–1362.

Camps-Valls, G., Gomez-Chova, L., Muoz-Mar, J., Vila-Francs, J., and Calpe-Maravilla, J. (2006). Composite kernels for hyperspectral image classification. *IEEE Geoscience and Remote Sensing Letters*.

Chapelle, O., Chi, M., and Zien, A. (2006). A continuation method for semi-supervised svms. In *Proceedings of the 23rd International Conference on Machine Learning*, pages 185–192. ACM Press.

Dias, J. and Nascimento, J. M. P. (2008). Hyperspectral subspace identification. *IEEE Trans. on Geoscience and Remote Sensing*, 46(8):2435–2445.

Duda, R. O., Hart, P. E., and Stork, D. G. (200). *Pattern Classification*. Wiley-Interscience, 2nd edition.

Fauvel, M., Benediktsson, J. A., Chanussot, J., and Sveinsson, J. R. (2008). Spectral and spatial classification of hyperspectral data using svms and morphological profiles. *IEEE Transactions on Geoscience and Remote Sensing*, 46(11):3804–3814.

Green, R. O., Eastwood, M. L., Sarture, C. M., Chrien, T. G., Aronsson, M., Chippendale, B. J., Faust, J. A., Pavri, B. E., Chovit, C. J., Solis, M., and et al (1998). Imaging spectroscopy and the airborne visible/ infrared imaging spectrometer (aviris). *Remote Sensing of Environment*, 65(3):227–248.

Kittler, J. and Roli, F. (2000). Multiple classifier systems. In *In Proc. of the First Int. Workshop MCS 2000. Lecture Notes in Computer Science*, volume 1857.

Krishnapuram, B., Carin, L., Figueiredo, M., and Hartemink, A. (2005). Sparse multinomial logistic regression: Fast algorithms and generalization bounds. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(6):957968.

Krishnapuram, B., Williams, D., Xue, Y., Hartemink, A., Carin, L., and Figueiredo, M. (2004). On semisupervised classification. In *Proc. 18th Annual Conference on Neural Information Processing Systems*, Vancouver, Canada.

Landgrebe, D. A. (2003). *Signal Theory Methods in Multispectral Remote Sensing*. John Wiley, Hoboken, NJ.

Li, J., Bioucas-Dias, J., and Plaza, A. (2009). Semisupervised hyperspectral image classification based on a markov random field and sparse multinomial logistic regression. In *Proc. IEEE International Geoscience and Remote sensing Symposium*.

Li, J., Bioucas-Dias, J., and Plaza, A. (2010a). Exploiting spatial information in semi-supervised hyperspectral image segmentation. In *IEEE GRSS Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS10)*.

Li, J., Bioucas-Dias, J., and Plaza, A. (2010b). Hyperspectral image segmentation using a new bayesian approach with active learning. In *IEEE Transactions on Geoscience and Remote Sensing (submitted)*.

Li, J., Bioucas-Dias, J., and Plaza, A. (2010c). Semisupervised hyperspectral image segmentation using multinomial logistic regression with active learning. In *IEEE Transactions on Geoscience and Remote Sensing*, volume 48, page 40854098.

Mackay, D. (1992). Information-based objective functions for active data selection. *Neural Computation*, 4:590–604.

Plaza, A., Benediktsson, J. A., Boardman, J. W., Brazile, J., Bruzzone, L., Camps-Valls, G., Chanussot, J., Fauvel, M., Gamba, P., Gualtieri, A., Marconcini, M., Tilton, J. C., and Trianni, G. (2009). Recent advances in techniques for hyperspectral image processing. *Remote Sensing of Environment*, 113(110-122).

Roli, F. (2005). Semi-supervised multiple classifier systems: Background and research directions. *Multiple Classifier Systems*, pages 1–11.

Tuia, D. and Camps-Valls, G. (2009). Semi-supervised hyperspectral image classification with cluster kernels. *IEEE Geoscience and Remote Sensing Letters*.

Velasco-Forero, S. and Manian, V. (2009). Improving hyperspectral image classification using spatial preprocessing. *IEEE Geoscience and Remote Sensing Letters*, 6:297–301.

Xu, L., Krzyzak, A., , and Suen, C. Y. (1992). Methods for combining multiple classifiers and their applications to handwriting recognition. *IEEE Trans. on Systems, Man and cyb.*, 22:418435.

Zhu, X. (2005). Semi-supervised learning literature survey. technical report 1530, computer sciences. Technical report, University of Wisconsin-Madison.