

# SPECTRAL UNMIXING-BASED POST-PROCESSING FOR HYPERSPECTRAL IMAGE CLASSIFICATION

Inmaculada Dópido<sup>1</sup>, Paolo Gamba<sup>2</sup> and Antonio Plaza<sup>1</sup>

<sup>1</sup>Hyperspectral Computing Laboratory  
Department of Technology of Computers and Communications  
University of Extremadura, Caceres, Spain  
<sup>2</sup>Telecommunications and Remote Sensing Laboratory  
University of Pavia, Italy

## ABSTRACT

Spectral unmixing and classification have been widely used in the recent literature to analyze remotely sensed hyperspectral data. However, possible connections between classification and spectral unmixing concepts have been rarely investigated. In this work, we propose a simple spectral unmixing-based post-processing method to improve the classification accuracies provided by supervised and semi-supervised techniques for hyperspectral image classification. The proposed approach exploits the information retrieved with spectral unmixing in order to complement the results obtained after the classification stage (which can be either supervised or semi-supervised), thus bridging the gap between unmixing and classification and exploiting both techniques in synergistic fashion for hyperspectral data interpretation. The proposed method is experimentally validated using a real hyperspectral data set collected by the Reflective Optics Spectrographic Imaging System (ROSIS). Our experimental results indicate that the proposed unmixing-based preprocessing can improve the classification results for some of the classes, particularly the most highly mixed ones, in supervised and semi-supervised scenarios using limited training samples.

**Index Terms**— Hyperspectral imaging, classification, spectral unmixing, semi-supervised learning.

## 1. 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 [2]. 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 has fostered the idea of adopting semi-supervised learning techniques in hyperspectral image classification [3].

---

This work has been supported by the CARIPO project “Azioni di internazionalizzazione per il post-laurea nell’ambito delle tecnologie dell’ICT e biomediche”.

Although not extensively, the integration of spectral unmixing and classification has been already explored in previous works. For instance, in [4] spectral unmixing was used as a feature extraction strategy prior to classification. It was found that the features obtained by unmixing can be often associated to physical features in the scene as opposed to statistical approaches for feature extraction such as the principal component analysis (PCA) [5] or the minimum noise fraction (MNF) [6], in which the physical meaning of the features is generally lost. This idea was expanded in [7] by analyzing additional unmixing chains for feature extraction prior to classification, including chains which perform the unmixing based on the available training samples (used as endmembers for mixture characterization) and chains which integrate spatial and spectral information in spectral unmixing via clustering techniques. More recently, the synergistic nature of spectral unmixing and classification has been further explored in the context of a semi-supervised self learning framework [8]. This strategy provides a joint approach for hyperspectral data interpretation that considers simultaneously the output provided by both unmixing and classification, where the weight given to either technique can be controlled by the end-user to adjust the classification output. As a follow-up to the work in [8], here we explore the possibility of using spectral unmixing as a post-processing technique in order to refine the output given by a certain classifier. Our introspection is that spectral unmixing can be useful as a post-processing technique when the classification scenario is dominated by highly mixed pixels, which may prevent the classifier from providing a correct classification output (particularly, in situations in which pixels are given by several constituent materials residing at sub-pixel levels). In those cases, spectral unmixing may provide a useful tool in order to refine the obtained classification by conducting a further interpretation of the mixing properties in those pixels and possibly changing the classification labels accordingly.

## 2. UNMIXING-BASED POST-PROCESSING

The strategy used for unmixing-based post-processing is based on the well-known linear mixture model [1]. Let us denote a remotely sensed hyperspectral scene with  $n$  bands by  $\mathbf{I}$ , in which the pixel at the discrete spatial coordinates  $(i, j)$  of the scene is represented by a vector  $\mathbf{X}(i, j) = [x_1(i, j), x_2(i, j), \dots, x_n(i, j)] \in \mathbb{R}^n$ , where  $\mathbb{R}$  denotes the set of real numbers to which the pixel’s spectral response  $x_k(i, j)$  at sensor channels  $k = 1, \dots, n$  belongs. Under the linear mixture model assumption, each pixel vector in the original scene

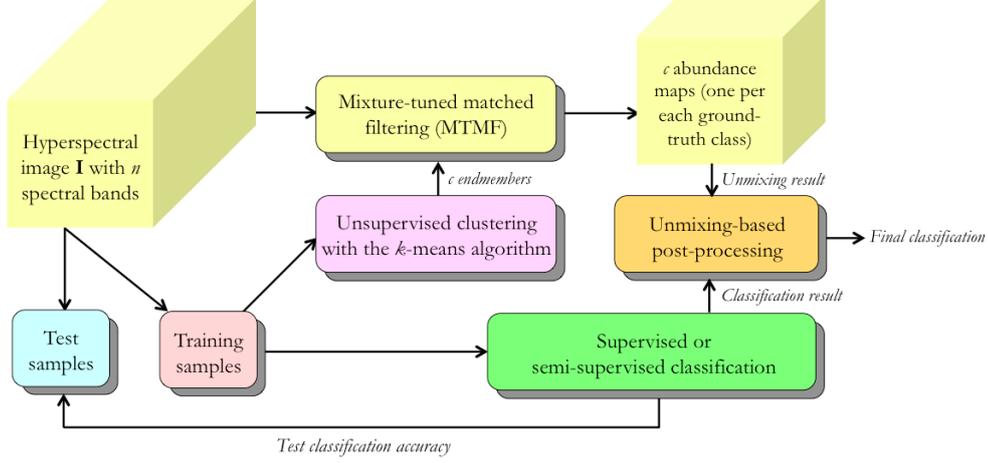


Fig. 1. Block diagram illustrating the proposed unmixing-based post-processing approach.

can be modeled using the following expression:

$$\mathbf{X}(i, j) = \sum_{z=1}^p \Phi_z(i, j) \cdot \mathbf{E}_z + \mathbf{n}(i, j), \quad (1)$$

where  $\mathbf{E}_z$  denotes the spectral response of endmember  $z$ ,  $\Phi_z(i, j)$  is a scalar value designating the fractional abundance of the endmember  $z$  at the pixel  $\mathbf{X}(i, j)$ ,  $p$  is the total number of endmembers, and  $\mathbf{n}(i, j)$  is a noise vector. Two physical constraints are generally imposed into the model described in (1), these are the abundance non-negativity constraint (ANC), i.e.,  $\Phi_z(i, j) \geq 0$ , and the abundance sum-to-one constraint (ASC), i.e.,  $\sum_{z=1}^p \Phi_z(i, j) = 1$  [9].

With the aforementioned notation in mind, the proposed post-processing strategy can be summarized by the flowchart given in Fig. 1. As shown by Fig. 1, first we obtain a set of endmembers after clustering the training set (using the popular  $k$ -means algorithm [10]). The number of endmembers to be extracted is given by the total number of different classes,  $c$ , in the labeled samples available in the training set. In this scenario, it is likely that the actual number of endmembers in the original image,  $p$ , is larger than the number of different classes comprised by available labeled training samples,  $c$ . Therefore, in order to unmix the original image we need to address a partial unmixing problem.

A successful technique to estimate abundance fractions in such partial unmixing scenarios is mixture-tuned matched filtering (MTMF) [11] –also known in the literature as constrained energy minimization (CEM) [12, 13]– which combines the best parts of the linear spectral unmixing model and the statistical matched filter model while avoiding some drawbacks of each parent method. From matched filtering, it inherits the ability to map a single known target without knowing the other background endmember signatures, unlike the standard linear unmixing model. From spectral mixture modeling, it inherits the leverage arising from the mixed pixel model and the constraints on feasibility including the ASC and ANC requirements. It is essentially a target detection algorithm designed to identify the presence (or absence) of a specified material by producing a score of 1 for pixels wholly covered by the material of interest, while keeping the average score over an image as small as possible. It uses just one endmember spectrum (that of the target of interest) and therefore behaves as a partial unmixing method that suppresses background noise and estimates the sub-pixel abundance

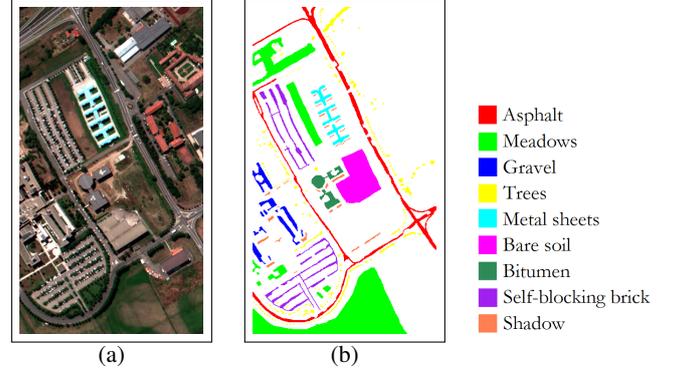


Fig. 2. (a) False color composition of the ROSIS Pavia scene. (b) Ground truth-map containing 9 mutually exclusive land-cover classes.

of a single endmember material without assuming the presence of all endmembers in the scene, as it is the case with fully constrained linear spectral unmixing (FCLSU) [9]. If we assume that  $\mathbf{E}_z$  is the endmember to be characterized, MTMF estimates the abundance fraction  $\Phi(i, j)$  of  $\mathbf{E}_z$  in a specific pixel vector  $\mathbf{X}(i, j)$  of the scene as follows:

$$\hat{\Phi}_{\text{MTMF}}(i, j) = ((\mathbf{E}_z^T \mathbf{R}^{-1} \mathbf{E}_z)^{-1} \mathbf{R}^{-1} \mathbf{E}_z)^T \mathbf{X}(i, j), \quad (2)$$

where  $\mathbf{R}$  is the matrix:

$$\mathbf{R} = \frac{1}{s \times l} \sum_{i=1}^s \sum_{j=1}^l \mathbf{X}(i, j) \mathbf{X}^T(i, j), \quad (3)$$

with  $s$  and  $l$  respectively denoting the number of samples and the number of lines in the original hyperspectral image.

Once a set of fractional abundance maps have been obtained by the MTMF technique, as shown in Fig. 1 the next step is to use the derived abundance fractions to refine the output of a supervised or semi-supervised classifier (based on the initial set of training samples available for the scene). In order to achieve such post-processing, we assume that the considered classifier provides a

probabilistic output:  $f_c[y(i, j) = k|\mathbf{X}(i, j)]$ , where  $\mathbf{X}(i, j)$  is the pixel to be classified,  $y(i, j)$  is the label assigned to this given pixel, and  $k$  is the class label. In other words, the classifier provides, for each pixel  $\mathbf{X}(i, j)$ , the probability that its associated label belongs to a certain class  $k$ , where  $k \in \{1, \dots, c\}$ , and  $c$  is the total number of classes. Similarly, the abundances provided by the unmixing process can be interpreted as probabilistic function  $f_u[y(i, j) = k|\mathbf{X}(i, j)]$ , in which the abundance of a given class endmember is interpreted as a probability that the pixel belongs to that given (pure) class. With those assumptions in mind, the proposed post-processing strategy is given by the following expression:

$$\hat{p}(i, j)[y(i, j) = k|\mathbf{X}(i, j)] = \alpha f_c[y(i, j) = k|\mathbf{X}(i, j)] + (1 - \alpha)f_u[y(i, j) = k|\mathbf{X}(i, j)], \quad (4)$$

where  $\hat{p}(i, j)[y(i, j) = k|\mathbf{X}(i, j)]$  is the joint estimate for the  $k$ -th class, i.e.,  $y(i, j) = k$ , obtained by the classification and unmixing methods given the pixel  $\mathbf{X}(i, j)$ . As a result, function  $f_c[\cdot]$  is the probability obtained by the classification algorithm and function  $f_u[\cdot]$  is the abundance fraction obtained from the spectral unmixing chain in Fig. 1, i.e.,  $f_u[y(i, j) = k|\mathbf{X}(i, j)] = \hat{\Phi}_{\text{MTMF}}(i, j)$ . The relationship between the classification probabilities and abundance fractions is controlled by parameter  $\alpha$ , where  $0 \leq \alpha \leq 1$ . As implicit in (4), if  $\alpha = 1$ , only classification probabilities are considered by the proposed strategy and no unmixing-based post-processing is applied. On the other hand, if  $\alpha = 0$ , only spectral unmixing is taken into account. By tuning  $\alpha$  from 0 to 1, we can adjust the impact of the unmixing-based post-processing on the classification output.

To conclude this section, we would like to emphasize that the presented post-processing is exploited in this work by assuming that a probabilistic classifier is available. However, the post-processing could be used with other types of classifiers, provided that they offer some form of class membership which can be interpreted as a probability. This is the case for many classification techniques available for hyperspectral image data. Hence, we believe that the post-processing presented in this contribution is quite general, although in the experiments reported in the following section we will only conduct the evaluation using probabilistic classifiers.

### 3. EXPERIMENTAL RESULTS

The hyperspectral data set used in the experiments was collected by the ROSIS optical sensor over the urban area of the University of Pavia, Italy. The flight was operated by the *Deutschen Zentrum for Luftund Raumfahrt* (DLR, the German Aerospace Agency) in the framework of the HySens project, managed and sponsored by the European Union. The image size in pixels is  $610 \times 340$ , with very high spatial resolution of 1.3 meters per pixel. The number of data channels in the acquired image is 103 (with spectral range from 0.43 to  $0.86 \mu\text{m}$ ). Fig. 2(a) shows a false color composite of the image, while Fig. 2(b) shows nine ground-truth classes of interest, which comprise urban features, as well as soil and vegetation features.

In our experiments, we use the multinomial logistic regression (MLR) probabilistic classifier [14] with the Gaussian RBF kernel [15], which is applied to a normalized version of the considered hyperspectral data set. The semi-supervised classifier strategy adopted in experiments is the one described in [3], also based on the MLR classifier. In all cases, we report the overall accuracy (OA), average accuracy (AA) and  $\kappa$  statistic, which are obtained by averaging the results obtained after conducting 10 independent Monte Carlo runs on a subset randomly selected from the ground-truth image in Fig. 2(b), where the remaining samples are used for validation purposes.

**Table 1.** Overall and average classification accuracies, and  $\kappa$  statistic obtained using different classification strategies (based on 5, 10 and 15 labeled samples) for the ROSIS Pavia University hyperspectral data set. In all cases, the results correspond to the mean values obtained after 10 Monte Carlo runs, and the standard deviations are also reported.

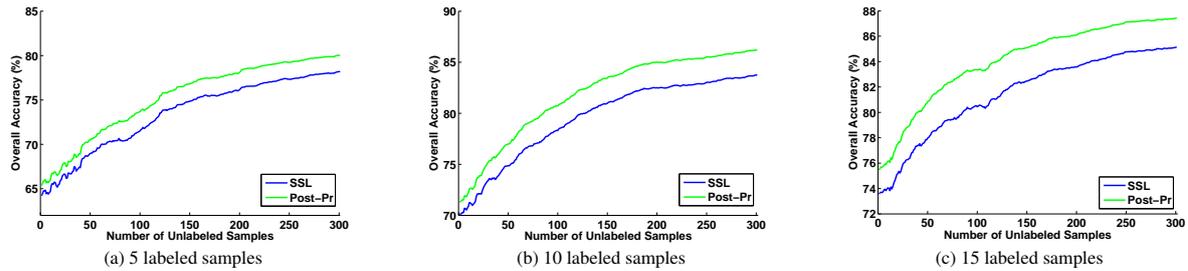
	5 labeled samples per class		
	Overall accuracy %	Average accuracy %	Kappa%
Supervised	64.23 ± 3.78	72.92 ± 1.67	55.38 ± 3.63
Semi-Supervised	78.22 ± 2.55	79.81 ± 3.41	71.55 ± 3.61
Post-processing	80.05 ± 2.58	81.77 ± 3.36	73.97 ± 3.60
	10 labeled samples per class		
	Overall accuracy %	Average accuracy %	Kappa%
Supervised	70.14 ± 2.50	78.91 ± 1.78	62.64 ± 2.67
Semi-Supervised	83.72 ± 2.37	83.62 ± 1.14	78.56 ± 2.87
Post-processing	86.17 ± 2.65	86.52 ± 1.37	81.83 ± 3.26
	15 labeled samples per class		
	Overall accuracy %	Average accuracy %	Kappa%
Original	73.59 ± 3.05	81.32 ± 1.05	66.79 ± 3.27
Semi-Supervised	85.14 ± 2.26	84.49 ± 1.36	80.47 ± 2.68
Post-processing	87.45 ± 2.26	87.29 ± 0.97	83.52 ± 2.70

In order to illustrate the good performance of the proposed approach, we use on purpose very small training sets (5, 10 and 15 pixels per class).

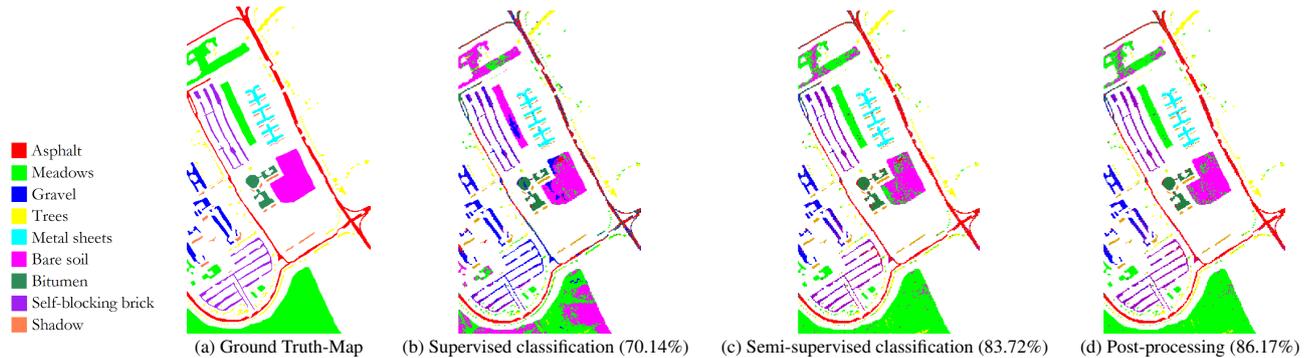
Table 1 reports the classification accuracies obtained by using supervised, semi-supervised and post-processing strategies with different number of labeled samples per class. The semi-supervised classifier used a total of 300 unlabeled samples, and the post-processing was applied after semi-supervised classification. For the post-processing we conducted an extensive optimization of the parameter  $\alpha$ , which controls the relative weight between unmixing and classification. We empirically observed that the best classification accuracies were obtained after setting  $\alpha = 0.2$  (this means, classification has a relative weight of 20% while unmixing has a relative weight of 80% in the post-processing stage). It should be noted that the results reported in Table 1 correspond to the mean values obtained after 10 Monte Carlo runs (the standard deviations are also reported in the table). For clarity, Fig. 3 graphically illustrates the advantages of using the proposed post-processing for supervised classification (which means no unlabeled samples in Fig. 3) and for semi-supervised classification using different numbers of unlabeled samples (from 1 to 300). As shown by Fig. 3, the improvements obtained after applying the proposed post-processing are more significant as the number of unlabeled samples is increased (while being also relevant for supervised classification based on labeled samples only). Finally, Fig. 4 shows some of the classification maps obtained for the ROSIS Pavia University 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 1. The improvements obtained by the unmixing-based post-processing strategy can be appreciated in classes such as *bare soil* or *meadows*, which are quite mixed in nature. Specifically, the *meadows* class is actually a mixture of vegetation and soil, as it is also the case for the *bare soil* class, as it can be visually appreciated in the false color composition of the ROSIS Pavia University scene in Fig. 2(a).

### 4. CONCLUSIONS AND FUTURE WORK

Classification and spectral unmixing are two techniques that have been traditionally used in separate contexts, but which exhibit potential to be exploited in synergistic fashion. In this work, we have developed a simple unmixing-based post-processing strategy which



**Fig. 3.** Overall accuracies (as a function of the number of unlabeled samples) for the ROSIS Pavia data set (with and without post-processing).



**Fig. 4.** Classification maps and overall classification accuracies (in the parentheses) obtained using 10 labeled samples per class. The semi-supervised classifier used a total of 300 unlabeled samples, and the post-processing was applied after semi-supervised classification. Notice the improvements introduced by semi-supervised classification with unlabeled training samples and with the proposed post-processing, especially in classes which may contain mixed pixels such as *meadows* or *bare soil* [as indicated by the false color representation of the scene displayed in Fig. 2(a)].

can be used to refine the results obtained by supervised and semi-supervised classification strategies for hyperspectral image data. The proposed approach has been tested using the probabilistic MLR classifier and a semi-supervised strategy also based on the MLR. Our experimental results, conducted using a hyperspectral scene collected by the ROSIS instrument over the city of Pavia, Italy, indicate that a post-processing strategy based on unmixing concepts can be particularly useful to improve the classification results obtained in classes dominated by mixed pixels. In future works we will test the proposed post-processing with other supervised and semi-supervised classifiers (e.g., those based on the probabilistic support vector machine) and extend our analysis to additional hyperspectral scenes with coarser spatial resolution, in order to fully calibrate the advantages that unmixing-based post-processing can provide in classification scenarios dominated by mixed pixels.

## 5. 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] D. A. Landgrebe, *Signal Theory Methods in Multispectral Remote Sensing*. John Wiley & Sons: New York, 2003.
- [3] I. Dopido, J. Li, P. R. Marpu, A. Plaza, J. M. Bioucas-Dias, and J. A. Benediktsson, "Semi-supervised self-learning for hyperspectral image classification," *IEEE Transactions on Geoscience and Remote Sensing*, 2013.
- [4] 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.
- [5] J. A. Richards and X. Jia, *Remote Sensing Digital Image Analysis: An Introduction*. Springer, 2006.
- [6] A. A. Green, M. Berman, P. Switzer, and M. D. Craig, "A transformation for ordering multispectral data in terms of image quality with implications for noise removal," *IEEE Trans. Geosci. Remote Sens.*, vol. 26, pp. 65–74, 1988.
- [7] 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.
- [8] I. Dopido, J. Li, A. Plaza, and P. Gamba, "Semi-supervised classification of hyperspectral data using spectral unmixing concepts," *Proceedings of the 2012 Tyrrhenian Workshop on Advances in Radar and Remote Sensing (TyWRRS)*, vol. 1, pp. 353–358, 2012.
- [9] 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.
- [10] J. A. Hartigan and M. A. Wong, "Algorithm as 136: A k-means clustering algorithm," *Journal of the Royal Statistical Society, Series C (Applied Statistics)*, vol. 28, pp. 100–108, 1979.
- [11] 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.
- [12] 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.
- [13] C.-I. Chang, *Hyperspectral Imaging: Techniques for Spectral Detection and Classification*. Kluwer Academic/Plenum Publishers: New York, 2003.
- [14] D. Böhning, "Multinomial logistic regression algorithm," *Annals of the Institute of Statistical Mathematics*, vol. 44, pp. 197–200, 1992.
- [15] 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.