Efficient Multi-temporal Hyperspectral Signatures Classification Using a Gaussian-Bernoulli RBM Based Approach[1,2]

S. Hemissi and Imed Riadh Farah

Laboratoire RIADI, ENSI,Telecom Bretagne, Brest, France

e-mail: selim.hemissi@ensi.rnu.tn; imed.farah@ensi.rnu.tn

Abstract—This paper presents an efficient Gaussian-Bernoulli Restricted Boltzmann Machines (GB-RBM) framework in order to better address the classification challenge of remotely sensed images. The proposed approach relies on generating well-designed features for a new 3D modality of spectral signature. For this purpose, mesh smoothing is introduced to reduce noise while conserving the main geometric features of the multi-temporal spectral signature. Then, we propose the use of an RBM (Restricted Boltzmann Machine) framework as stand-alone non-linear classifier. The adapted framework focuses on a cooperative integrated generative-discriminative objective allowing the integration of modeling input features and their classification process in one-pass algorithm. The main benefit of the proposed approach is the ability to learn more discriminative features. We evaluated our approach within different scenarios and we demonstrated its usefulness for noisy high dimensional hyperspectral images.

Keywords: efficient multi-temporal hyperspectral.

DOI: 10.1134/S1054661816010211

1. INTRODUCTION

The few last years have seen a significant increase in the amount and the availability of multi-temporal images which brings a great challenge to the traditional pattern recognition approaches [1]. The great spectral resolution of hyperspectral images offers an exceptional characterization of land-cover types with unrivaled accuracy. Nevertheless, the processing of multi-temporal images turns out to be more challenging than common grayscale images. This is, mainly, due to of the high dimensionality of data, their probable non-linear nature and the singular noise and uncertainty level. Such uncertainties and non-linearities may result from several factors, including temporal variability, heterogeneity of mixed pixels, as well as acquisition distortions [2]. Such facts lead to a distinct non-linear aspect of the classifier.

In hyperspectral imaging, different features (e.g., spectral, texture, or shape features) are used to describe pixels from diverse kind of perspectives [3]. In fact, due to the increasingly impressive computational potential and the expeditious improvement of feature extraction techniques, images (i.e., objects) are described by features extracted from multifarious points of view. Often, a data object described by many features can be naturally decomposed into multiple "views," each modeling a subset of features (e.g., an image may have a color view, a shape view, a textural view, etc.). Following, we investigate the idea of properly combining different features in a 3D modeling which may results in better classification accuracy [4].

Formally, each view represents an object in a definite feature space with some particular statistical characteristics of processed data and an appropriate physical significance [3]. Since that each view reflects a limited aspect and has distinctive intrinsic discriminative potential, a curious idea is to combine them into a unified model of the spectral signature to boost the classification accuracy. Nevertheless, most existing techniques are designed for the single-view scenario. This approach neglects the physical aspect and ignores the underlying complementarity and correlation between views.

The update of remotely sensed data to the multiview scenario should be done thoroughly and carefully. To reach that goal, we introduce a 3D model including various facets of the hyperspectral images: temporal, spectral, and spatial. Moreover, an intuitive solution is to combine multi-view features and to design a multifeatures algorithm which performs a low-level combination of complementary features in a supervised context. The major purpose is to be able to cope with dimentionality and lack of knowledge present in the data, while following a soft-classification model considered for hyperspectral data interpretation. Nowadays, generative-discriminative learning research is a growing area of research. A considerable improvement begins to be shown when reviewing associated recent

[1] This paper uses the materials of the report submitted at the 9th Open German-Russian Workshop on Pattern Recognition and Image Understanding, held in Koblenz, December 1–5, 2014 (OGRW-9-2014).

[2] The article is published in the original.

Received September 8, 2015

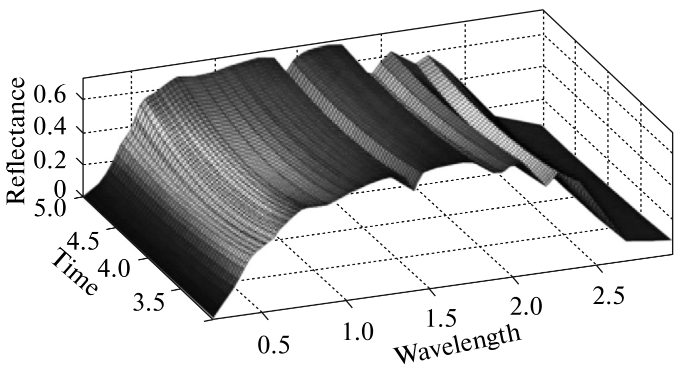


Fig. 1. 3D model of the spectral signature.

literature covering a wide variety of applications, such as clustering [5], classification [6], and dimensionality reduction [7]. Moreover, recent researches shows that the adequate exploitation of information contained in multiple views can substantially improve the recognition performance [8]. One typical case showing it's soundness is that a video clip can be represented by different heterogeneous features, such as language, shape, etc. Thereby, human perception is regularly associated with various visual features, which can be either appearance features or motion features (e.g., optical ow).

The remainder of the paper is organized as follows. Section 2 is dedicated to problem formulation and literature review, Section 3 presents in greater details the proposed approach. Section 4 shows the experiment on a synthetic data set, while Section 5 illustrates the experimental results on real hyperspectral data. Section 6 finally draws the conclusions and the future directions.

2. PROBLEM STATEMENT

When tackling hyperspectral data processing from a 3D perspective, various refinements can be envisaged by considering hyperspectral images as a rich source of information. In fact, combining information from various data sources is become a prominent research topic in machine learning [9, 10]. For this purpose, some approaches have exploited the original spectral information or other features linearly derived from spectral signatures in order to separate classes which are assumed to be linearly separable. In conjunction, other methods, aiming to reduce data dimensionality, were interested to extract features, obtained through nonlinear transformations. This allows either a suitable modeling of the inherent nonlinearity of hyperspectral images (e.g., kernels [11]) or an optimal exploitation of the spatial information (e.g., using morphological analysis [12]).

Pioneer works adopt the idea stipulating that if we increase the number of features, the accuracy will increase unquestionably. Perhaps strangely, the classifier performance was not automatically improved (reduced in some cases) by the profusion of multibands images [13]. Concurrently, this phenomenon is the after effect of the increase of spectral class variability, Hughes showed that classification performance decreased, as further features were involved. This phenomena is designated by "the curse of dimensionality" [14].

Moreover, one of the prominent issues of hyperspectral images is the relatively low spatial resolution which is quite common with sensors covering wide areas [15]. This can lead naturally to the challenging problem of mixed pixels affecting seriously the classifier accuracy. Conventional machine learning algorithms, such as SVM or composite kernel machines integrate all multiple views into one single view to standardize the learning setting, Nevertheless, this investigation is not physically meaningful since that each view has a particular statistical characteristics and induces an over-fitting when only few training samples are available. Otherwise to single view case, 3D modeling of the spectral signature is a new paradigm introduces the idea of modeling each specific view and jointly optimizes them to exploit the redundant informations and improve the learning achievement. Hence, Restricted Boltzmann Machines (GR-RBM) learning has been receiving increased thinking.

3. A GAUSSIAN-BERNOULLI RBM BASED APPROACH FOR HYPERSPECTRAL CLASSIFICATION

The main objective is to study the effectiveness of multi-temporal hyperspectral images classification by using a Gaussian-Bernoulli RBM based approach. Given a set of training data objects extracted from the 3D model of the spectral signature, the proposed classifier tries to learn a classifier by incorporating the complementary information from features.

3.1. 3D Hyperspectral Signature Smoothing for Noise Reduction

The few years have seen significant increase in the amount and the availability of multi-temporal images, and it brings a great challenge to the traditional pattern recognition approaches. To overcome these drawbacks, the authors proposed in [10] a novel model baptized the "multi-temporal spectral signature" (Fig. 1). This model is defined by putting the classical spectral signature of the same pixel at different times in the same 3D Cartesian coordinate system. In this way, we obtain a bounded domain Ω and we denote by $E = \{p_1, ..., p_n\}$ the set of points in $\mathbb{R}^3$ called sites. The Delaunay triangulation of E, noted $Del(E)$, is the geometric dual of $Vor(E)$. We denote by $Vor(E)$ the Voronoi diagram of E which is the subdivision of space instigated by the Voronoi cells $V(p_1), ..., V(p_n)$.

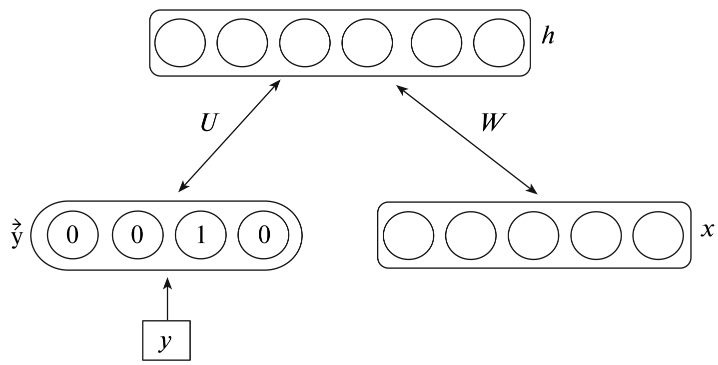


Fig. 2. RBM adopted architecture.

Therefore, we obtain for each pixel a 3D multi-temporal signatures defined by Eq. (1). This new model incorporates the time (*T*) and the spectral (λ) dimensions. So, we can express the reflectance (*Ref*) at each pixel using the Eq. (1).

$$Ref_{Pixel_{i,j}} = f(T, \lambda). \quad (1)$$

In this stage, the atmospheric distortions affecting the reflectance values of pixels alter the position of points in the 3D reconstruction space. Subsequently, this noise turns into a geometric noise which requiring a specific treatment in order to reduce its effects on the classification accuracy.

The first contribution of this paper is the use of Laplacian smoothing in order to limit the impact of noise. Our choice was guided by the aim to establish a simple and robust optimization of 3D signatures, while maintaining both sampling rate and connectivity. The choice of Laplacian smoothing technique is motivated by the fact that it changes the position of nodes without modifying the topology of the multi-temporal signatures.

Laplacian smoothing is a well studied method for polygonal mesh [16]. We denoted by S a triangular mesh surface, X is the vertices of the mesh, L is the Laplacian, and λ is a scalar that controls the diffusion speed. Our goal is to produce a new mesh surface S' with an enhanced quality ratio. This quality reflects a critical factor in the accuracy and stability of pattern recognition tasks. The key idea is to, incrementally, move the vertices of the mesh S in the direction of the Laplacian. Since that the Laplacian operator is linear, the smoothing equation is:

$$X(n+1) = (I + \lambda dtL)X(n). \quad (2)$$

This investigation allows us to generate a smoothed 3D signature which is close as possible to S and preserves the features of S. Then we generate for each pixel a feature vector denoted X_i using the same features as in our previous work [10].

3.2. 3D Hyperspectral Signatures Classification Using a GB-RBM Algorithm

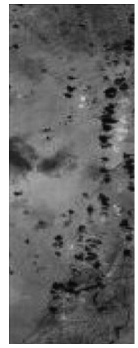
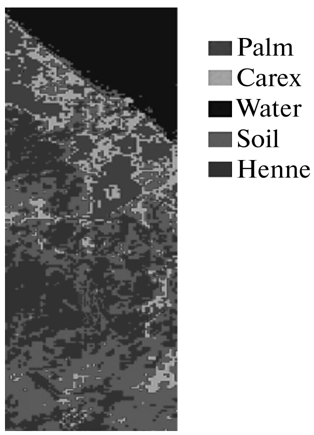
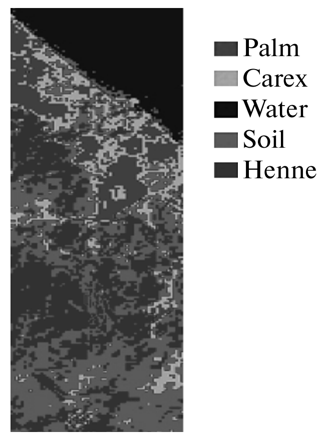
The data produced by the previous stage are characterized mainly by their high dimensionality and their probable non-linear nature. To improve their classification, we will need more sophisticated and adapted framework.

The Gaussian-Bernoulli Restricted Boltzmann Machines (GB-RBM) is a stochastic neural network including a non-linear generative model. The theoretical proofs of this model correspond to the requirements of our data. Recently, RBMs have attracted much attention to address a wide range of pattern recognition problems. But, they are mainly employed to initialize deep neural networks [17]. Briefly, the RBM model mainly includes m visible units $v = (v_1, \ldots, v_m)$ and n hidden units $h = (h_1, \ldots, h_n)$, with fully connecting between them (Fig. 2). The visible units are combined with the feature vectors extracted in the previous section.

In our case, we assume given a training set $Train = f\{(x_i, y_i)\}$ involving for the ith pixel an input feature vector x_i and a target class $y_i \in \{1, \ldots, C\}$. As known, the generative learning objective of standard RBMs aims to minimize the energy of the inputs and to induce the most truthful representation of the input. Nevertheless, this representation, which aims to model the intra-class variation of the learning samples, is not automatically useful for a classifier. Therefore, the goal of this part is to adapt the well-known learning objective to the classification of multi-temporal hyperspectral images.

Recent researches such as Larochelle et al. [18] and Louradour et al. [19] prove that RBMs offers a self-contained framework to implement an efficient non-linear classifier. Consequently, we propose to use a learning objective that promotes inter-class separability. To achieve this goal, the Gaussian-Bernoulli RBM (GB-RBM) framework was adopted including v_m visible units with real-value and h_n binary hidden units.

Table 1. Sequence of real images bands [6, 19, 33]) and their corresponding true classification maps (first time serie: 2009/2010)

	03/06/2009	23/09/2009	27/12/2009	09/01/2010	30/04/2010
Real Hyperion images					
Thematic map			True map		

Following the same theory, the energy function of the GB-RBM is expressed as:

$$E(v, h|\theta) = -\sum_{i=1}^{m}\sum_{k=1}^{n}\omega_{ij}h_j\frac{v_i}{\sigma_i^2} - \sum_{i=1}^{m}\frac{(v_i - b_i)^2}{2\sigma_i^2} - \sum_{j=1}^{n}c_i h_j. \quad (3)$$

Under this configuration, the computation of conditional probabilities is done by using the following expressions:

$$p(v_i = v|h) = N\left(v|b_i + \sum_j h_j\omega_{ij}\sigma_i^2\right), \quad (4)$$

$$p(h_i = 1|v) = sigmoid\left(c_i + \sum_i \omega_{ij}\frac{v_i}{\sigma_i^2}\right), \quad (5)$$

where N denotes the Gaussian probability density function with mean μ and variance σ^2. As shown, the proposed framework is a derived RBM model that models the conditional distribution $p(y|x)$ instead of $p(y, x)$. To train the GB-RBM model, we use the contrastive divergence algorithm [20].

4. EXPERIMENTAL RESULTS

This section illustrates the performance of the proposed method in a challenging multi-temporal classification case. The proposed approach was tested on two different data sets. These datasets involve several types of data, and with dimensions ranging from 176 to 183 bands. The first dataset, Hyperion, contains vegetation type data, is divided into five classes, has 183 spectral bands and has a pixel size of 30*m*. The second set is from an airborne sensor (AVIRIS), divided into 7 closes, has 176 spectral bands and a pixel size of 18*m*.

First, we present experiments that assess the classification accuracy of the proposed approach (PA). We also included an SVM classifier and a Multilayer Pereeptron (MP) classifier in our comparison as a baseline. Figure 1 and Table 2 summarizes the obtained results.

The results indicate that the proposed approach is highly capable of classifying simulated data in a noisy/free environment.

The Hyperion dataset benchmark constitutes a challenging problem due to the significant presence of noisy pixels in all available classes. Figures 3 and 4 report the classification accuracies and standard deviations obtained as functions of the number of the size of learning dataset.

It can be observed that very good results are obtained by using more samples. For instance, with only 400 learning samples, the proposed classifier obtained an *OA* of 67.87% and a κ of 0.64. However, by including 600 more samples (which come at more cost), the proposed approach increased the *OA* to 85.16% (with κ of 0.76).

5. DISCUSSIONS AND FUTURE DIRECTIONS

In this paper, we have introduced an Gaussian-Bernoulli Restricted Boltzmann Machines (GB-RBM) framework offering a direct solution to the soft classification issue for hyperspectral images processing. Taken as a whole, the obtained results show the effectiveness and the widespread applicability of this idea for multi-temporal hyperspectral images classification. The strength of the proposed approach lies in enjoying the complementary of different facets of the spectral signature and their integration in an united 3D model. Several additional challenges confronting our perspective concerning mainly the lack of prior knowledge and the heterogeneous nature of extracted features. So beyond the classical schema of RBM classification, we introduced several useful and complementary enhancements to deal with multi-temporal hyperspectral data. These included: globally integrated learning of both labeled and unlabeled samples, adapted distance metric based on kernel function. These enhancements can be adapted and concatenated to meet the demands of challenging issue in hyperspectral imagery.

Table 2. Evaluation of the proposed approach compared to several conventional approaches

Algorithm	AVIRIS dataset	Hyperion dataset
PA	88.24	85.16
SVM	85.67	83.47
MP	84.29	79.54

In our case, more than one facet can be generated to characterize the spectral signature. Instead of elect one view from the original data, we were interested to propose a model considering the divergence and the complementarity of different facets. Through analyzing the obtained results, we notice that we happened to study the problems with respect to how to generate 3D model and how to smooth it. The experimental results show the promising accuracy compared to previous works. We can conclude that an enhanced RBM classifier is practical and promising when introduced for hyperspectral data processing, but it has not been well-addressed until now.

Although promising study has been carried out in this direction, several crucial research issues need to be addressed in the future. Since the process of 3D model generation, smoothing and classification largely influence the performance of learning, it is intuitive to place more emphasis on this direction. It will be also valuable to develop a general framework of multi-view learning which includes the merits of die rent multi-view learning methods.

Future work will focuses on several open issues. For many real-world applications in phenomena monitoring, risk management or information retrieval, the data sets can be much bigger. For very large data sets, our current implementation will be more effective if it will be implemented in a parallel manner. It remains

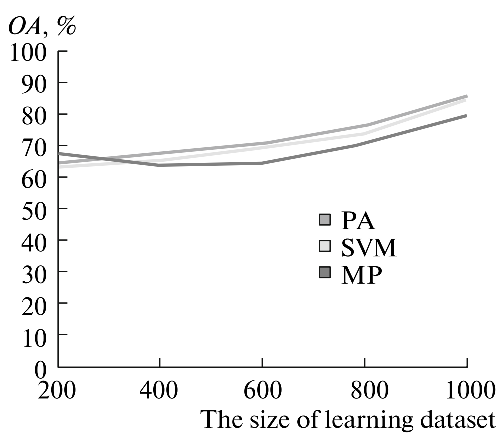


Fig. 3. *OA* (%) accuracy results as function of the size of learning dataset.

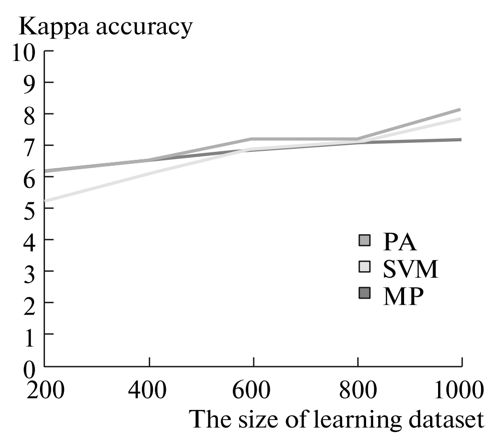


Fig. 4. Kappa statistics results as function of the size of learning dataset.

an also interesting challenge to additional strategies for the generation of soft labels (e.g., obtaining the label estimates from the whole image) will be used to fully substantiate our findings.

CONCLUSIONS

Recent advances in multi-temporal hyperspectral imaging offer the opportunity to easily integrate temporal and spatial facets by using a 3D model of the spectral signature. While improving objects recognition, some challenges such as the curse of dimensionality and noisy data are still difficult to address and need further improvements. To overcome these problems, this paper introduces two main innovations: I, the noise affecting the multi-temporal signatures is modeled and reduced by a post-processing Laplacian smoothing method and II, the concept of Restricted Boltzmann Machines (RBM) is introduced for hyperspectral data classification and the Gaussian-Bernoulli RBM framework is adapted to handle this challenge. The experiments prove the competitive results of the proposed approach.

REFERENCES

1. Yuan Yuan, Haobo Lv, and Xiaoqiang Lu, "Semi-supervised change detection method for multi-temporal hyperspectral images," Neurocomputing **148**, 363–375 (2015).
2. Junhwa Chi and M. M. Crawford, "Selection of landmark points on nonlinear manifolds for spectral unmixing using local homogeneity," IEEE Trans. Geosci. Remote Sensing Lett. **10** (4), 711–715 (2013).
3. Lefei Zhang, Liangpei Zhang, Dacheng Tao, and Xin Huang, "On combining multiple features for hyperspectral remote sensing image classification," IEEE Trans. Geosci. Remote Sensing **50** (3), 879–893 (2012).
4. Hongjun Su, Yehua Sheng, Peijun Du, Chen Chen, and Kui Liu, "Hyperspectral image classification based on volumetric texture and dimensionality reduction," Frontiers Earth Sci. **9** (2), 225–236 (2015).
5. Qiyue Yin, ShuWu, Ran He, and Liang Wang, "Multiview clustering via pairwise sparse subspace representation," Neurocomputing **156**, 12–21 (2015).
6. M. Volpi, G. Matasci, M. Kanevski, and D. Tuia, "Semi-supervised multiview embedding for hyperspectral data classification," Neurocomputing **145**, 427–437 (2014).
7. M. Gnen, "Coupled dimensionality reduction and classification for supervised and semi-supervised multilabel learning," Pattern Recogn. Lett. **38**, 132–141 (2014).
8. Jun Yu, Dacheng Tao, Yong Rui, and Jun Cheng, "Pairwise constraints based multiview features fusion for scene classification," Pattern Recogn. **46** (2), 483–496 (2013).
9. Shuhan Chen, Weiren Shi, and Xiao Lv, "Feature coding for image classification combining global saliency and local difference," Pattern Recogn. Lett. **51**, 44–49 (2015).
10. S. Hemissi, I. R. Farah, K. Saheb Ettabaa, and B. Solaiman, "Multi-spectro-temporal analysis of hyperspectral imagery based on 3D spectral modeling and multilinear algebra," IEEE Trans. Geosci. Remote Sensing **51** (1), 199–216 (2013).
11. Bor-Chen Kuo and Cheng-Hsuan Li, "Kernel nonparametric weighted feature extraction for classification," in *AI 2005: Advances in Artificial Intelligence,* Ed. by Shichao Zhang and Ray Jarvis (Springer, Berlin, Heidelberg, 2005), pp. 567–576.
12. M. Dalla Mura, A. Villa, J. A. Benediktsson, J. Chanussot, and L. Bruzzone, "Classification of hyperspectral images by using extended morphological attribute profiles and independent component analysis," IEEE Trans. Geosci. Remote Sensing Lett. **8** (3), 542–546 (2011).
13. L. Journaux, M.-F. Destain, J. Miteran, A. Piron, and F. Cointault, "Texture classification with generalized Fourier descriptors in dimensionality reduction context: an overview exploration," in *Artificial Neural Networks in Pattern Recognition,* Ed. by L. Prevost, S. Marinai, and F. Schwenker (Springer, Berlin, Heidelberg, 2008), pp. 280–291.
14. L. Yan and D. P. Roy, "Improved time series land cover classification by missing-observation-adaptive nonlinear dimensionality reduction," Remote Sensing Environ. **158**, 478–491 (2015).
15. Jun Li, J. M. Bioucas-Dias, and A. Plaza, "Semisupervised hyperspectral image classification using soft sparse multinomial logistic regression," IEEE Trans. Geosci. Remote Sensing Lett. **10** (2), 318–322 (2013).
16. A. Nealen, T. Igarashi, O. Sorkine, and M. Alexa, "Laplacian mesh optimization," in *Proc. 4th ACM Int. Conf. on Computer Graphics and Interactive Techniques in Australasia and Southeast Asia, GRAPHITE'06* (New York, 2006), pp. 381–389.
17. Takashi Kuremoto, Shinsuke Kimura, Kunikazu Kobayashi, and Masanao Obayashi, "Time series forecasting using a deep belief network with restricted Boltzmann machines," Neurocomputing **137**, 47–56 (2014).
18. H. Larochelle, M. Mandel, R. Pascanu, and Y. Bengio, "Learning algorithms for the classification restricted Boltzmann machine," J. Mach. Learn. Res. **13** (1), 643–669 (2012).
19. J. Louradour and H. Larochelle, "Classification of sets using restricted Boltzmann machines," CoRR, abs/1103.4896, 2011.
20. Liu Jian-wei, Chi Guang-hui, and Luo Xiong-lin, "Contrastive divergence learning of restricted Boltzmann machine," in *Proc. 2nd* IEEE *Computer Soc. Int. Conf. on Electric Technology and Civil Engineering, ICETCE'12* (Washington, 2012), pp. 712–715.

Selim Hemissi received the PHD degree in computer sciences and signal processing jointly from Telecom Bretagne, Brest, France and the Ecole Nationale des Sciences de Informatique (ENSI), Manouba, Tunisia, in 2014. He is a Permanent Researcher at Laboratory RIADI, University of Manouba, since 2008. His work is mainly related with pattern recognition, signal processing, and machine learning applied to remote sensing hyperspectral images. He also enjoyed a research partnership with the Department ITI in Telecom Bretagne where he is currently an Associate Researcher. Mr. Hemissi is a member of Arts-Pi Tunisia and IEEE Student Brunch of Telecom Bretagne.

Imed Riadh Farah received the MD degree from the ISG Institute of Computer Sciences, Tunis, Tunisia, in 1995, and the Dr. Eng. degree from the Ecole Nationale des Sciences de l'Informatique (ENSI), Manouba, Tunisia, in 2003. After working as a Research Assistant (from 1996) and a Permanent Researcher at Laboratory RIADI, EXSI National School of Computer Sciences engineering (since 1995), he has been an Associate Professor at the University of Manouba, since 2010. He has been an Associate Researcher in the Department ITI-Telecom Bretagne, Brest, France, since January 2009. His research interests include image processing, pattern recognition, artificial intelligence, data mining, and their application to remote sensing. Currently, he is the Director of the Higher Institute of Arts and Multimedia. Dr. Farah is a member of Arts-Pi Tunisia.