Foreword to the Special Issue on Machine Learning for Remote Sensing Data Processing

I. Introduction

REMOTE sensing is nowadays a part of society, industry, science, and engineering. As a prominent example of its societal value, remote sensing images are integrated in navigation systems, and an ever-increasing number of images taken from space or from airborne sensors is continually used by the press and in the monitoring of public events. Long before its recreative use began, remote sensing had become a valuable source of data for scientists, entrepreneurs, public institutions and manufacturers to aid decision making in resource exploration, environmental protection, ecology, agriculture, urban development, quality control, and for discoveries about uncharted territories (including other planets). The easy access and comprehensive coverage that remote sensing provides to hard-to-reach parts of the world has opened wide possibilities, for example, for disaster forecasting and proactive mitigation [1] or for observing the evolution of dynamic processes at the global scale [2], [3].

Remote sensing is a polyvalent and accurate source of data recording the processes at work at the surface. These data are often massive, high-dimensional, and evolving in time; hence, mining such rich datasets has become a priority in order to provide actionable information to various stakeholders including the general public. This much-needed information is hidden in huge archives of undistilled digital data, from which simple queries may not produce interpretation on the level required for decision making. Therefore, machine learning algorithms have become a natural choice to facilitate the translation from raw data to useful information, as introduced in [4] and detailed in [5].

Machine learning algorithms [6]–[9] offer solutions that can generalize well to unseen situations, implement tracking of space/time processes, or discover uncommon events. However, remote sensing data also carry some unique characteristics such as geographic consistency, spatial context, multiscale behaviors, scattering geometry, or spectral correlations, which require specific knowledge and dedicated approaches different from those used in machine learning models applied to other types of data [10], [11]. For this reason, synergies between machine learning and remote sensing science are increasingly emerging to help tackle the problems specific to data dimensionality [12]–[14], complex data manifolds [15]–[17], small sample scenarios [18], [19], spatial dependencies [20], signal mixing [21], content-based retrieval from databases of high-dimensional data [22], and presence/discovery of novelties [23], [24]. Such fruitful synergies allow to look at the future—and the explosion of the number of sensors and the deluge of data to be treated—in an optimistic way.

This special issue of the IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing is a follow-up to special sessions organized at WHISPERS conferences with the involvement of two of the guest editors (M.G. and E.M.). Such sessions drew unexpectedly large attendance, signaling the interest and need for a focused platform to exchange knowledge at the intersection of machine learning and remote sensing. The collection of papers in this special issue presents a comprehensive sample of the latest trends in the design of machine learning algorithms for geospatial data. It covers a wide spectrum of remote sensing applications and presents new solutions to answer to the call of the new generation of sensors, covering the electromagnetic range from optical to microwave data.

We have selected 27 papers for this special issue. In this foreword, we introduce the contributions and provide an overview of the related topics. The guest editors would like to thank all authors for their excellent work and all the reviewers who made this special issue possible.

II. Overview of the Special Issue

This section presents the salient features of the contributions under several nonexclusive categories.

A. Machine Learning Methods: Supervised Classification Is the Main Topic

Either through discriminative or generative approaches, supervised classification has been the most investigated topic in the special issue. Out of the accepted papers, 20 propose supervised classification techniques and related methodologies for remote sensing data [25]–[41] or target detection (classification with a single class of interest) [42]–[44]. In contrast, unsupervised classification is considered in [45] via the use of evolutionary algorithms.

This is in-line with the trends observed in the remote sensing community, whose main products are thematic maps describing the spatial distribution of land cover types and land use. Support vector machines (SVMs) [7] are considered as the state-of-the-art supervised classifiers and are often used as a *de facto* standard. SVMs are therefore challenged by the proposed innovative and established approaches, especially in small sample scenarios (see, e.g., [28]–[30], [36], [39], [40]). In these cases, active learning [18] is often proposed as a way to solve exploratory problems [27], [35], [41].

B. Variety of Other Machine Learning Approaches

Closely related to classification, a large body of research in the special issue is devoted to feature extraction and feature selection. These strategies are used to help increase performance through the use of more expressive features than those in the raw data or to reduce data dimensionality in order to apply available classifiers. The majority of the above-cited papers include some

Digital Object Identifier 10.1109/JSTARS.2014.2311915

feature extraction/selection approach. Also fundamental to classification is the similarity metric used. Exploring the effects of Mahalonobis metrics is pursued in [31].

Regression of biophyisical parameters is also represented in the special issue. In [46], a series of multivariate regression methods (including Gaussian process, kernel ridge regression, multilayer perceptron, and others) are extensively compared for retrieval of land biophysical parameters. In [47], neural networks deliver sophisticated retrieval of atmospheric parameters from combined hyperspectral and microwave data. Two papers tackle the problems of super-resolution and pansharpening. In [48], authors propose the use of support vector regression to achieve super-resolution of land cover maps from mid-resolution satellite data. Pansharpening of medium-resolution images is proposed in [49], with a contribution rooted in dictionary learning and sparse coding [50].

C. Data Types: Hyperspectral as a Driving Force

Out of the accepted contributions, 16 propose (although not always exclusively) methodologies for hyperspectral images (HSI) [25]–[32], [39], [43]–[48], [51]. Due to their inherent high dimensionality and complexity, HSI have traditionally been a driver for the synergies between machine learning and remote sensing [52]. Works in HSI sometimes have contributed to the advances of machine learning in remote sensing, such as the recognition of specific nonintuitive properties of high-dimensional data [53], the effect of class noise in classifier performance [54], or the very early introduction of semi-supervised classification ideas [55]. The contributions to this special issue indicate that hyperspectral imagery continues to be one of the main reasons for methods development.

Very-high-resolution (VHR) passive optical sensors, providing submetric passive optical images, are also well represented. Contributions focus on problems specific to the given resolution, as in object retrieval by [35] and human–machine interaction by [27]. Medium–resolution sensors with long temporal series such as MODIS [36] or AVHRR [37] are preferred in studies exploiting temporal trends at large scales (see Section III-B) or change detection [41].

While passive optical measurements are the dominant data source, papers dealing with other types of sensors such as synthetic aperture radar (SAR) [34], [35], polarimetric SAR [38], [42], and their joint use with HSI [47] introduce new research objectives and raise novel challenges such as the extraction of features specific to these type of signals [38] or the use of the acquisition geometry in the algorithms. The potential of machine learning in these fields is still underexploited, but the results of the studies above promise definite advances in the near future.

D. Wide Range of Applications

A wide range of applications is represented in the special issue. In addition to the construction of thematic maps, a number of interesting applications show the potential of techniques and methodologies in new areas: in [43], authors detect chemical plumes, while [47] studies the reconstruction of atmospheric profiles. Authors in [46] present a toolbox including a set of methodologies to retrieve parameters such as Chlorophyll and Leaf Area Index (LAI) in fields, while authors in [37] present a method to map and survey the evolution of burned areas. Authors in [40] compare features and classificaton methods specifically for land use and land cover mapping of tropical regions from PALSAR data. The search for patterns in large databases in [35] poses interesting teaching/learning problems that are common to diverse image modalities such as optical and SAR. Additionally, [56] proposes a method for the estimation of time of flight in sonar and radars, while [31] turns the focus far from our planet by studying the presence of minerals on Mars.

E. Validation Methodology

Acceptance of these new machine learning approaches by the community requires a rigorous validation methodology, to ensure a set of algorithmic standards and best practices to be followed. The first aspect of this methodological standard is the comparison of algorithms on widely accepted benchmark datasets (e.g., the Indian Pines, University of Pavia or Salinas datasets in HSI). It is an accepted standard for papers (e.g., [25]–[30], [39]) to exploit such datasets as a proof of concept and in order to demonstrate the effectiveness of the proposed solution in specific scenarios. However, the need for the creation and maintenance of public repositories of datasets is real and made urgent by the fast pace of innovation in the computational algorithms as well as in the sensors. New data sources may render obsolete previous computational approaches, or validation results obtained on old datasets may be meaningless for the new data stream. New challenging repositories of data benchmarks (along with ground samples) are therefore much needed by the community.

The second aspect of the methodological best practices is the statistical soundness of the design of the experiments. For example, we observed that the strict separation of test and training data is enforced less often than desired. The imperative to report improving results, instead of well reasoned and described processes often seems to suppress the mandate for the best methodology. In this special issue we strove to strongly encourage methodology standards (including clear description of separate training and test data). Authors in [31], [39] bring this concern to an exemplary level by specifying further clear separation of data used for training and model selection.

The final aspect to be considered is the publication of the code, which allows for reproducibility of results and remains the best way to disseminate the proposed methods. A paradigmatic example is the LIBSVM [57], which has been influential in the dissemination of SVMs as a *de facto* standard. Following this reasoning, the contribution of [46] is exemplar, since the authors provide a fully functional and open toolbox for regression algorithms along with their paper.

III. Contribution to the Development of Machine Learning in the Remote Sensing Community

This section is a subjective taxonomy of the papers of the special issue, driven by some common methodological trends observed in the contributions.

A. Enforcing Spatial Consistency

Spatial consistency is an assumption that allowed the main advances in remote sensing image processing [20], [58]: the images are spatially smooth and so are the objects they represent. In other words, data (i.e., pixels in the images) are highly correlated spatially. Approaches proposed in this issue enforce spatial smoothness via the use of spatial filters [25], [26], [39], spatial clustering [45], segmentation based on Markov Random Fields [34], or by hypothesis generation from sets of neighboring pixels [28]. In the latter reference, a linear correction of the pixel spectra is computed by regularized least squares based on the neighboring pixels. In addition, the approach also profits from spectral redundancies by performing a block decomposition based on the correlation between bands. The approach in [39] involves spatial preprocessing in order to obtain improved features for classification, by the combined use of Wavelet transforms and morphological profiles, features that are also used in [29] for spectral–spatial feature extraction.

B. *Exploiting Temporal Information*

Analysis of data structured in time series is one of the challenging issues in remote sensing [59]. The availability of time series of images is a primary advantage of medium-resolution sensors such as MODIS or AVHRR. Mining such time series and using the observed spectral trends is the basis of the proposal of [36] and [37], where the authors exploit the temporal structures to improve the quality of the classifications of vegetation and burned areas, respectively. Another traditional application in bitemporal remote sensing is change detection, represented by [41], where authors propose the use of active learning to improve the detection of changes.

C. *Increasing Robustness Under Insufficient Ground Truth*

The acquisition of remote sensing data remains much faster than the generation of ground truth. The imbalance between the presence of labeled information and the increasing size (and dimensionality) of the data acquired is among the main problems faced in training classification algorithms or in searching large databases (see [35]). In many contributions to this issue, authors of [28], [30], [31] acknowledge this situation by testing their algorithms on a collection of small-sample cases. To increase robustness, some authors explore unsupervised algorithms [45], while others use the information contained in unlabeled pixels: such information either leads to the selection of new sampling sites via active learning and relevance feedback [27], [35], [41] or it allows to increase the robustness of models using the unlabeled data as an opposing class in one-class problems [42]. In all cases, these methodologies lead to drastic improvements in performance.

D. *The Strength of Collective Decision*

Ensemble methods [60], [61]—offering the strength of collective decision—are probably the most represented family of innovation algorithms of this special issue. Some well known approaches, i.e., Random Forests, are used as standards comparable to SVMs [36]. A combination of different classifiers is proposed in [33], while [28] proposes to evaluate classifiers based on different hypotheses about the location of training samples and spectral subsets. Reference [29] applies successful bagging and boosting methodologies to increase the robustness of an extreme learning machine. Authors in [51] use ensembles of random spectral subspaces selected with a genetic optimization to improve SVM classification. Authors in [35] propose a cascade of coarse to fine classifiers trained with active learning for object recognition in databases. Finally, authors in [32] analyze ensemble methods, feature selection and post-classification in terms of bias-variance decomposition.

E. *Alternative Representations*

The papers grouped here consider alternative data representations [62] to retrieve either compact or nonlinear descriptions of the data space. Authors in [30] propose to model HSI as a series of class-manifolds and evaluate classification of a test sample in terms of the perturbation that such sample provides on the manifold characterization (if the perturbation is small, the pixel is likely to belong to that manifold). The use of sparse representations is proposed in [49], where pansharpening is performed by learning coupled dictionaries at each resolution, along with sparse coefficients for optimal reconstruction. In [43], authors perform sparse feature selection in the feature space spanned by the empirical kernel map [63]. The exploration of the diverse Mahalanobis metric learning algorithms in [31] is another way to perform alternative representation trough the use of adaptive distances for similarity-based classification algorithms.

IV. Concluding Remarks

This special issue is a snapshot of the open problems and challenges that bring machine learning and remote sensing closer. The spectrum of topics covered raises a variety of lively questions, and calls for further debate about the future of synergy between the two disciplines. We hope that you will enjoy reading it and that it will foster discussion and future developments.

Devis Tuia, *Guest Editor*
Laboratoire des Systémes d'Information
Géographique (LaSIG), Ecole Polytechnique
Fédérale de Lausanne (EPFL), Lausanne 1015,
Switzerland

Erzsébet Merényi, *Guest Editor*
Department of Statistics and the Department
of Electrical and Computer Engineering,
Rice University, Houston, TX 77005-1827, USA

Xiuping Jia, *Guest Editor*
School of Engineering and Information
Technology, University of New South Wales,
Canberra Campus, Canberra ACT2600, Australia

Manuel Graña-Romay, *Guest Editor*
Computational Intelligence Group,
University of the Basque Country,
Guipúzcoa 20018, Spain

References

[1] G. Singh, Y. Yamaguchi, W.-M. Boerner, and S. E. Park, "Monitoring of the March 11, 2011 off-tohoku 9.0 earthquake with super-Tsunami disaster by implementing fully polarimetric high-resolution POLSAR techniques," *Proc. IEEE*, vol. 101, no. 3, pp. 831–846, Mar. 2013.

[2] S. J. Goetz, A. G. Bunn, G. J. Fiske, and R. A. Houghton, "Satellite-observed photosynthetic trends across boreal North America associated with climate and fire disturbance," *Proc. Nat. Acad. Sci.*, vol. 102, no. 38, pp. 13521–13525, 2005.

[3] G. P. Asner, E. N. Broadbent, P. J. C. Oliveira, M. Keller, D. E. Knapp, and J. N. M. Silva, "Condition and fate of logged forests in the Brazilian Amazon," in *Proc. Nat. Acad. Sci.*, 2006, vol. 103, no. 34, pp. 12947–12950.

[4] S. M. Davis, D. Landgrebe, T. Phillips, P. Swain, R. Hoffer, J. Lindenlaub, and L. Silva, *Remote Sensing: The Quantitative Approach.* New York, NY, USA: McGraw-Hill, 1978.

[5] J. A. Richards and X. Jia, *Remote Sensing Digital Image Analysis*, 4th ed. New York, NY, USA: Springer-Verlag, 2006.

[6] K. Fukunaga, *Introduction to Statistical Pattern Recognition*, 2nd ed. Boston, MA, USA: Academic, 1990.

[7] V. Vapnik, *Statistical Learning Theory.* Hoboken, NJ, USA: Wiley-Interscience, 1998.

[8] S. Haykin, *Neural Networks—A Comprehensive Foundation*, 2nd ed. Upper Saddle River, NJ, USA: Prentice Hall, 1999.

[9] C. M. Bishop, *Pattern Recognition and Machine Learning.* New York, NY, USA: Springer-Verlag, 2006.

[10] G. Camps-Valls, D. Tuia, L. Gómez-Chova, S. Jimenez, and J. Malo, *Remote Sensing Image Processing, Synthesis Lectures on Image, Video, and Multimedia Processing.* New York, NY, USA: Morgan and Claypool, 2011.

[11] J. Bioucas-Dias, A. Plaza, G. Camps-Valls, P. Scheunders, N. Nasrabadi, and J. Chanussot, "Hyperspectral remote sensing data analysis and future challenges," *IEEE Remote Sens. Mag.*, vol. 1, no. 2, pp. 6–36, Jun. 2013.

[12] G. Camps-Valls, D. Tuia, L. Bruzzone, and J. A. Benediktsson, "Advances in hyperspectral image classification," *IEEE Signal Proc. Mag.*, vol. 31, no. 1, pp. 45–54, Jan. 2014.

[13] X. Jia, B.-C. Kuo, and M. Crawford, "Feature mining for hyperspectral image classification," *Proc. IEEE*, vol. 101, no. 3, pp. 676–697, Mar. 2013.

[14] M. J. Mendenhall and E. Merényi, "Relevance-based feature extraction for hyperspectral images," *IEEE Trans. Neural Netw.*, vol. 19, no. 4, pp. 658–672, Apr. 2008.

[15] D. Lunga, S. Prasad, M. Crawford, and O. Ersoy, "Manifold-learning-based feature extraction for classification of hyperspectral data," *IEEE Signal Proc. Mag.*, vol. 31, no. 1, pp. 55–66, Jan. 2014.

[16] T. Villmann, E. Merényi, and B. Hammer, "Neural maps in remote sensing image analysis," *Neural Netw.*, vol. 16, no. 3–4, pp. 389–403, 2003.

[17] E. Merényi, K. Tasdemir, and L. Zhang, "Learning highly structured manifolds: Harnessing the power of SOMs," in *Similarity Based Clustering* (Lecture Notes in Computer Science, LNAI 5400), M. Biehl, B. Hammer, M. Verleysen, and T. Villmann, Eds., New York, NY, USA: Springer-Verlag, 2009, pp. 138–168.

[18] M. M. Crawford, D. Tuia, and L. H. Hyang, "Active learning: Any value for classification of remotely sensed data?," *Proc. IEEE*, vol. 101, no. 3, pp. 593–608, Mar. 2013.

[19] E. Merényi, W. H. Farrand, J. V. Taranik, and T. B. Minor, "Classification of hyperspectral imagery with neural networks: Comparison to conventional tools," in *Machine Learning Reports*, vol. 5, T. Villmann and F.-M. Schleif, Eds., ISSN: 1865-3960, 2011, pp. 1–15[Online]. Available: http://www.techfak.uni-bielefeld.de/fschleif/mlr/mlr_04_2011.pdf.

[20] M. Fauvel, Y. Tarabalka, J. A. Benediktsson, J. Chanussot, and J. C. Tilton, "Advances in spectral-spatial classification of hyperspectral images," *Proc. IEEE*, vol. 101, no. 3, pp. 652–675, Mar. 2013.

[21] W. Ma, J. Bioucas-Dias, T. H. Chan, N. Gillis, P. Gader, A. Plaza *et al.*, "A signal processing perspective on hyperspectral unmixing," *IEEE Signal Proc. Mag.*, vol. 31, no. 1, pp. 67–81, Jan. 2014.

[22] M. Grana and M. A. Veganzones, "An end member-based distance for content based hyperspectral image retrieval," *Pattern Recognit.*, vol. 49, no. 9, pp. 3472–3489, 2012.

[23] N. Nasrabadi, "Hyperspectral target detection," *IEEE Signal Proc. Mag.*, vol. 31, no. 1, pp. 34–45, 2014.

[24] E. Merényi, W. H. Farrand, R. H. Brown, Th. Villmann, and C. Fyfe, "Information extraction and knowledge discovery from high-dimensional and high-volume complex data sets through precision manifold learning," in *Proc. NASA Sci. Technol. Conf.* (NSTC), College Park, MD, USA, Jun. 19–21, 2007, p. 11, ISBN 0-9785223-2-X.

[25] W. Li and Q. Du, "Gabor-filtering based nearest regularized subspace for hyperspectral image classification," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 4, pp. 1012–1022, Apr. 2014.

[26] S. Jia, Z. Zhu, and Q. Li, "A two-stage feature selection framework for hyperspectral image classification using few labeled samples," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 4, pp. 1023–1035, Apr. 2014.

[27] L. Z. Huo and P. Tang, "Batch-mode active learning algorithm using region-partitioning diversity for SVM classifier," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 4, pp. 1036–1046, Apr. 2014.

[28] C. Chen, E. Tramel, M. Cui, S. Prasad, and J. Fowler, "Spectral-spatial preprocessing using multihypothesis prediction for noise-robust hyperspectral image classification," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 4, pp. 1047–1059, Apr. 2014.

[29] A. Samat, P. Du, S. Liu, J. Li, and L. Cheng, "E2LMs: Ensemble extreme learning machines for hyperspectral image classification," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 4, pp. 1060–1069, Apr. 2014.

[30] L. Chapel, T. Burger, N. Courty, and S. Lefèvre, "PerTurbo manifold learning algorithm for weakly labelled hyperspectral image classification," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 4, pp. 1070–1078, Apr. 2014.

[31] B. Bue, "An evaluation of low-rank Mahalanobis metric learning techniques for hyperspectral image classification," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 4, pp. 1079–1088, Apr. 2014.

[32] A. Merentitis, C. Debes, and R. Heremans, "Ensemble learning in hyperspectral image classification: Towards selecting a favorable bias-variance tradeoff," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 4, pp. 1089–1102, Apr. 2014.

[33] F. Faria, D. Pedronette, J. dos Santos, A. Rocha, and R. Torres, "Rank aggregation for pattern classifier selection in remote sensing images," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 4, pp. 1103–1115, Apr. 2014.

[34] C. D'Elia, S. Ruscino, M. Abbate, B. Aiazzi, S. Baronti, and L. Alparone, "SAR image classification through information-theoretic textural features, MRF segmentation, and object-oriented learning vector quantization," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 4, pp. 1116–1126, Apr. 2014.

[35] P. Blanchart, M. Ferecatu, S. Cui, and M. Datcu, "Pattern retrieval in large image databases using multiscale coarse-to-fine cascaded active learning," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 4, pp. 1127–1141, Apr. 2014.

[36] Z. Xue, P. Du, and F. Li, "Phenology-driven land cover classification and trend analysis based on long-term remote sensing image series," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 4, pp. 1142–1156, Apr. 2014.

[37] J. Moreno-Riuz, D. Riaño, and S. Kefauver, "The synergy of the 0.05° (~ 5 km) AVHRR Long Term Data Record (LTDR) and Landsat TM archive to map large fires in the North American boreal region from 1984 to 1998," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 4, pp. 1157–1166, Apr. 2014.

[38] B. Banerjee, A. Bhattacharya, and K. M. Buddhiraju, "A generic land-cover classification framework for polarimetric SAR images using the optimum Touzi decomposition parameter subset—An insight on mutual information based feature selection techniques," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 4, pp. 1167–1176, Apr. 2014.

[39] P. Quesada Barriuso, F. Argüello, and D. B. Heras, "Spectral-spatial classification of hyperspectral images using wavelets and extended morphological profiles," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 4, pp. 1177–1185, Apr. 2014.

[40] T. Shiraishi, T. Motohka, R. Thapa, M. Watanabe, and M. Shimada, "Example-based super-resolution land cover mapping using support vector regression," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 4, pp. 1271–1283, Apr. 2014.

[41] M. Roy, S. Ghosh, and A. Ghosh, "A neural approach under active learning mode for change detection in remotely sensed images," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 4, pp. 1200–1206, Apr. 2014.

[42] W. Yang, X. Yin, H. Song, Y. Liu, and X. Xu, "Extraction of built-up areas from fully polarimetric SAR imagery via PU learning," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 4, pp. 1207–1216, Apr. 2014.

[43] P. Gurram and H. Kwon, "Optimal sparse kernel learning in the empirical kernel feature space for hyperspectral classification," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 4, pp. 1217–1226, Apr. 2014.

[44] R. Zhao, B. Du, and L. Zhang, "A robust nonlinear hyperspectral anomaly detection approach," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 4, pp. 1227–1234, Apr. 2014.

[45] Y. Zhong, A. Ma, and L. Zhang, "An adaptive memetic fuzzy clustering algorithm with spatial information for remote sensing imagery," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 4, pp. 1235–1248, Apr. 2014.

[46] J. P. Rivera, J. Verrelst, J. Muñoz-Marí, J. Moreno, and G. Camps-Valls, "Toward a semiautomatic machine learning retrieval of biophysical parameters," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 4, pp. 1249–1259, Apr. 2014.
[47] W. J. Blackwell and A. B. Milstein, "A neural network technique for high-resolution profiling of cloudy atmosphere," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 4, pp. 1260–1270, Apr. 2014.
[48] Y. Zhang, Y. Du, L. Feng, S. Fang, and X. Li, "Example-based super-resolution land cover mapping using support vector regression," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 4, pp. 1271–1283, Apr. 2014.
[49] M. Guo, H. Zhang, J. Li, L. Zhang, and H. Shen, "An online coupled dictionary learning approach for remote sensing image fusion," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 4, pp. 1284–1294, Apr. 2014.
[50] A. S. Charles, B. A. Olshausen, and C. J. Rozell, "Learning sparse codes for hyperspectral imagery," *IEEE J. Sel. Top. Signal Proc.*, vol. 5, no. 5, pp. 963–978, Sep. 2011.
[51] Y. Chen, X. Zhao, and Z. Lin, "Optimizing subspace SVM ensemble for hyperspectral imagery classification," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 4, pp. 1295–1305, Apr. 2014.
[52] D. A. Landgrebe, *Signal Theory Methods in Multispectral Remote Sensing.* Hoboken, NJ, USA: Wiley, 2003.
[53] L. O. Jimenez and D. A. Landgrebe, "Supervised classification in high-dimensional space: Geometrical, statistical, and asymptotical properties of multivariate data," *IEEE Trans. Syst. Man. Cybern. C*, vol. 28, no. 1, pp. 39–54, Feb. 1998.
[54] D. A. Landgrebe and E. Malaret, "Noise in remote-sensing systems: The effect on classification error," *IEEE Trans. Geosci. Remote Sens.*, vol. GE-24, no. 2, pp. 294–300, Mar. 1986.
[55] B. M. Shahshahani and D. A. Landgrebe, "The effect of unlabeled samples in reducing the small sample size problem and mitigating the Hughes phenomenon," *IEEE Trans. Geosci. Remote Sens.*, vol. 32, no. 5, pp. 1087–1095, Sep. 1994.
[56] A. Apartsin, L. Cooper, and N. Intrator, "Energy-efficient time of flight estimation in the presence of outliers: A machine learning approach," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 4, pp. 1306–1313, Apr. 2014.
[57] C. C. Chang and C. J. Lin, "LIBSVM: A library for support vector machines," *ACM Trans. Intell. Syst. Technol.*, vol. 2, pp. 1–27, 2011.
[58] K. Schindler, "An overview and comparison of smooth labeling methods for land-cover classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 11, pp. 4534–4545, Nov. 2012.
[59] L. Bruzzone, M. Marconcini, U. Wegmuller, and A. Wiesmann, "An advanced system for the automatic classification of multitemporal SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 6, pp. 1321–1334, Jun. 2004.
[60] M. Wozniak, M. Grana, and E. Corchado, "A survey of multiple classifier system as hybrid systems," *Inf. Fusion*, vol. 16, pp. 3–17, 2014.
[61] P. O. Gislason, J. A. Benediktsson, and J. R. Sveinsson, "Random forests for land cover classification," *Pattern Recognit. Lett.*, vol. 25, pp. 294–300, 2006.
[62] X. Jia and J. A. Richards, "Cluster-space representation for hyperspectral data classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 40, no. 3, pp. 593–598, Mar. 2002.
[63] B. Schölkopf and A. Smola, *Learning With Kernels.* Cambridge, MA, USA: MIT Press, 2002.

Devis Tuia (S'07–M'09) was born in Mendrisio, Switzerland, in 1980. He received the Ph.D. degree in environmental sciences at the University of Lausanne, Lausanne, Switzerland, in 2009.

He was a Postdoc Researcher with both the University of València, Spain, and the University of Colorado, Boulder, CO, USA. He is now a Senior Research Associate with the LaSIG laboratory, EPFL, Lausanne, Switzerland. His research interests include the development of algorithms for information extraction and classification of very-high-resolution remote sensing images using machine learning algorithms.

He serves as an Associate Editor for IEEE Journal of Selected Topics in Applied Earth Observation and Remote Sensing and as a Co-Chair of the *Image Analysis and Data Fusion* Technical Committee of the GRSS.

Erzsébet Merényi received the M.Sc. degree in mathematics in 1975 and the Ph.D. degree in computational science in 1980 from Szeged University, Hungary.

She holds a joint appointment as a Research Professor in the Departments of Statistics and Electrical and Computer Engineering, Rice University, Houston, TX, USA. She teaches artificial neural networks, statistics, probability, and remote sensing courses. Previously, she worked with the Lunar and Planetary Laboratory, University of Arizona, Tucson, AZ, USA, and at the Central Research Institute for Physics, Budapest, Hungary. Her research focuses on neural computation, self-organized learning, high-dimensional complex manifolds, pattern discovery and classification, data fusion, data mining, mapping planetary surfaces and biological tissues from hyperspectral imagery, and brain mapping from fMRI. She coauthored over a hundred refereed papers, three book chapters, organized machine learning/hyperpsectral special sessions at international conferences (WHISPERS 2010, 2011, ESANN 2012, 2008, 1999), organized Dagstuhl perspective seminar on learning high-dimensional data (2011) and served as guest editor for Neurocomputing, EURASIP-JASP.

Dr. Merényi is a member of the Technical Committee on Data Mining, IEEE Computational Intelligence Society.

Xiuping Jia (M'93–SM'03) received the B.Eng. degree from the Beijing University of Posts and Telecommunications, Beijing, China, in 1982, and the Ph.D. degree in electrical engineering from the University of New South Wales, Canberra Campus, Canberra, Australia, in 1996.

Since 1988, she has been with the School of Information Technology and Electrical Engineering, the University of New South Wales, where she is currently a Senior Lecturer. She is also a Guest Professor with Harbin Engineering University, China, and an Adjunct Researcher with the National Engineering Research Center for Information Technology in Agriculture, Beijing, China. Her research interests include remote sensing and imaging spectrometry. She is the coauthor of the remote sensing textbook titled *Remote Sensing Digital Image Analysis* [Springer-Verlag, 3rd ed. (1999) and 4th ed. (2006)].

Dr. Jia was a Guest Editor of IEEE Transactions on Geoscience and Remote Sensing Special Issue on Spectral Unmixing of Remotely Sensed Data. She is an Editor of the *Annals of GIS*. She is an Associate Editor for the IEEE Transactions on Geoscience and Remote Sensing.

Manuel Graña-Romay (M'94) received the M.Sc. and Ph.D. degrees from the Universidad del Pais Vasco (UPV/EHU), Donostia, Spain, in 1982 and 1989, respectively, both in computer science.

He is a Full Professor (Catedrático de Universidad) with the Computer Science and Artificial Intelligence Department, the Universidad del Pais Vasco (UPV/EHU). He is the Head of the Computational Intelligence Group (Grupo de Inteligencia Computational). His current research interests are in applications of computational intelligence to multicomponent robotic systems, specifically linked multicomponent robotic systems, medical image in the neurosciences, multimodal human computer interaction, remote sensing image processing, content-based image retrieval, lattice computing, semantic modeling, data processing, classification, and data mining. He has coauthored more than a hundred journal research papers and over 200 conference contributions, has advised over 20 Ph.D. students, and has organized international conferences IWANN 2007, HAIS 2010, KES 2012, and InMed 2014.