Semi-automatic Classification of Remote Sensing Images

Jefersson Alex dos Santos[1,2], Alexandre Xavier Falcão[1], Ricardo da Silva Torres[1]
[1]Institute of Computing, University of Campinas
13083-852, Campinas, SP - Brazil
Email: {jsantos,rtorres,afalcao}@ic.unicamp.br

Philippe-Henri Gosselin[2], Sylvie Philipp-Foliguet[2]
[2]ETIS, CNRS, ENSEA, University of Cergy-Pontoise
95018 Cergy-Pontoise Cedex, France
Email: {philipp,gosselin}@ensea.fr

***Abstract*—With the increasing accessibility to new technologies, the main problems in region recognition of remote sensing images are: (1) classification methods are dependent on the segmentation quality; and (2) the selection of representative samples for training. The major challenge is that the samples indicated by the user are not always enough to define the best segmentation scale. Furthermore, the indication of samples can be costly, since it often requires to visit studied places in loco. This PhD thesis[1] addressed the segmentation-dependence problem by introducing two new approaches that rely on multiple scales instead of using only one segmentation result. The selection of representative samples, on the other hand, was supported in this work by the development of a new interactive classification approach based on active learning. Significant contributions were also obtained concerning the description of regions in remote sensing images by means of: an evaluation study of 19 descriptors; and two new strategies for speeding up feature extraction from a hierarchy of segmented regions.**

***Keywords*-remote sensing; multiscale classification; active learning; feature propagation; bag of visual words**

I. Introduction

Since the satellite imagery information became available to the civil community in the 1970s, a huge effort has made on the creation of high quality thematic maps to establish precise inventories about land cover use [1]. Although remote sensing images (RSIs) are often used as reference data in many applications (such as agricultural monitoring [2], urban planning [3], and deforestation detection [4]), their peculiarities combined with the traditional image classification challenges have turned RSI classification into a hard task.

The research challenges in remote sensing image classification can be arranged into three main axes as illustrated in Figure 1. These axes are based on the following aspects: data representation, target recognition, and user interaction.

The *data representation* axis concerns the kind of data which are considered as the samples in the classification process (e.g., pixels [5], blocks of pixels [6], regions [2], and hierarchy of regions [7]).

Many region-based classification approaches, also called geographic object-based analysis (GEOBIA) [8], have been proposed by exploiting segmented regions in contrast with the traditional pixel-based approaches [1]. The main drawback of GEOBIA is its dependence on the employed segmentation algorithms, which delineate the objects of interest in the images. In this context, a common challenge is the selection of a suitable segmentation scale [9]. A suitable segmentation scale relies on the semantics and its association with the studied targets.

[1]This work relates to a Ph.D. thesis.

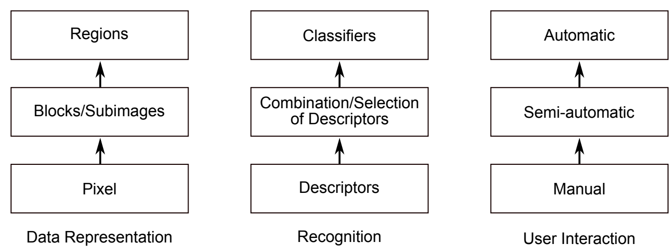


Fig. 1. The remote sensing image classification research axes.

The *recognition* axis comprises the research challenges related to feature extraction and classification of samples. The feature extraction provides a mathematical description for each object (by taking into account for example, spectral characteristics, texture, or shape). The classification module is in charge of separating objects from distinct classes based on machine learning techniques.

With the increasing amount of data provided by different kinds of sensors, there is a constant need for extraction algorithms able to produce good description for the targets of interest [10]. Thus, there is a need for machine learning techniques able to take advantage of those features to produce effective classifiers [11]. Finally, it also requires strategies both to select features and to improve classification results by creating ensemble of classifiers [8], [12].

The *user interaction* axis refers to the challenges that are related to user interactivity over the classification process: manual, automatic, and semi-automatic. In a manual classification, the recognition is completely dependent on users' perceptions and decisions. This process typically consists of drawing the areas of interest in the RSI by using some software. It often requires visits to the studied place to confirm obtained results. In automatic approaches, the user indicates the training set samples and some supervised method is used to classify the remaining samples given a learning process.

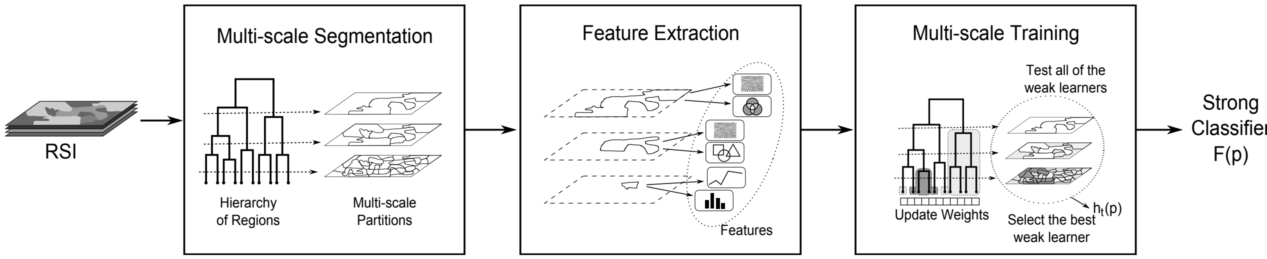


Fig. 2. Steps of the multiscale classification approach. At the begining, several partitions P_λ of hierarchy H at various scales λ are selected. Then, at each scale λ, a set of features is computed for each region $R \in P_\lambda$. Finally, a classifier $F(p)$ is built by using the Multiscale Training or the Hierarchical Multiscale Training.

The semi-automatic classification strategy does not only use supervised classification but also allows the user to refine the classification process along iterations.

The work developed in this PhD thesis [13] has contributed to address challenges in all those axes. We mainly focus on the data representation and feature extraction problems in order to develop effective solutions for interactive classification of remote sensing images. The main contributions are:

1) **Two inovative approaches for multiscale classification of remote sensing images [7]**. The objective is to combine features extracted from multiple scales instead of using only one segmentation scale. The first approach, Multiscale Classifier (MSC), builds a strong classifier that combines features extracted from multiple scales of segmentation. The other, Hierachical Multiscale Classifier (HMSC), exploits the hierarchical topology of segmented regions to improve training efficiency without accuracy loss when compared to the MSC.
2) **A novel approach for interactive multiscale classification of remote sensing images [14].** The proposed system allows the classification of regions based on features extracted from multiple segmentation scales with a reduced training set by exploiting an active learning strategy. It not only reduces the redundance of the training data but also enables user to refine the classification results.
3) **A descriptor evaluation study in the context of agricultural targets classification [15], [16].** Twelve descriptors that enconde spectral/color and seven texture descriptors that have never been used in remote sensing classification tasks were tested for the recognition of pasture and coffee crops [15]. We have also analysed the correlation among descriptors at different segmentation scales [16].
4) **Two approaches for feature extraction from a hierarchy of regions**. The BoW-Propagation [17] and the H-Propagation [18] allow the extraction of features faster than low-level descriptors at all segmentation scales. The features present similar quality of representation.

Besides those contributions, many others related to this PhD work were published in [2], [5], [6], [19]–[26], [26]–[28]. The following sections detail each of the contributions obtained from the developed research.

II. Multiscale Classification

Regardless of the data representation model adopted in supervised classification of RSIs, both the training input and the result of the classifier can be expressed as sets of pixels. In spite of that, data representation cannot only rely on pixels, because their image characteristics are not usually enough to capture the patterns of the classes (regions of interest). In order to bridge that semantic gap, multiscale image segmentation can play an important role. As pointed out by Trias-Sanz et al. [29], most of the image segmentation methods use threshold parameters to create a partition of the image. These methods usually create a single-scale representation of the image: small thresholds give segmentation with small regions and many details, while large thresholds preserve only the most salient regions. The problem is that various structures can appear at different scales and this segmentation result can be difficult to obtain without prior knowledge about the data or by using only empirical parameters. It is difficult to define the optimal scale for segmentation. Some parts of an image may need a fine segmentation, since the plots are small, whereas, in other parts, a coarse segmentation is sufficient. For this reason, the main drawback of classification methods based on regions is that they depend on the segmentation method used. Bearing this in mind, many researchers have exploited multiple scales of data [4], [8], [9].

The contribution published in [7] introduces two innovative multiscale training approaches based on boosting of weak classifiers for multiscale classification of RSIs. The first approach, *Multiscale Classifier (MSC)*, builds a strong classifier that combines features extracted from multiple scales of segmentation. Figure 2 illustrates the general idea.

The second training approach, *Hierarchical Multiscale Classifier (HMSC)*, exploits the hierarchical topology of segmented regions to improve training efficiency without accuracy loss when compared to the MSC. Figure 3 illustrates this process. It consists of individually selecting the weak classifiers for each scale, starting from the coarsest one to the finest one. Thereby, each scale provides a different stage of training. At the end of each stage, only the most difficult samples are selected, limiting the training set used in the next stage.

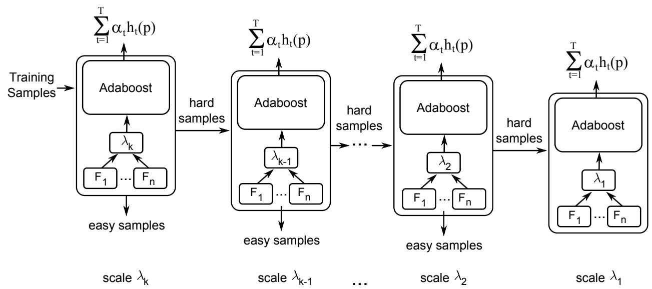


Fig. 3. The hierarchical multiscale training strategy.

The MSC and HMSC approaches differ from the other studies found in the literature in several aspects. The main novelty is the combination of multiple segmentation scales, which increases the power of the final classifier. Moreover, they also build classifiers that are able to combine different features by weighting the ones more suitable for each application. Experiments show that it is better to use multiple scales than use only one segmentation scale result. For a more detailed discussion, please refer to [7].

III. Interactive Classification

The size and the redundancy of the training set are the most challenging issues concerning the indication of samples for supervised classification [30]. These issues have a direct impact on the execution time needed for training and on the final result of the classification. In addition, labeling of samples often requires visits to the study site, which can add extra costs to the analysis. The training set must, thus, be carefully chosen, avoiding redundancy patterns, but also ensuring a good representation of the considered classes.

The contribution published in [14] presents a novel approach for interactive multiscale classification of remote sensing images. We proposed an active learning strategy to allow the refinement of classification results by the user along iterations.

Figure 4 gives an overview of the architecture used in our approach for interactive classification. The framework is composed of three main processing modules: segmentation, feature extraction, and classification. Segmentation and feature extraction are offline steps. When an image is inserted into the system, the segmentation is performed, building a hierarchical representation of regions. Feature vectors from these regions are then computed and stored.

The interactive classification starts with the user's annotation. He/she selects a small set of relevant and non-relevant pixels. Using these pixels as training set, the method builds a classifier to label the remaining pixels. Although the training set is at the pixel level, the training is performed by using features extracted from the segmented regions for each considered scale. At the end of the classification step, the method selects regions for possible feedback. When the result of the classification is displayed, the user feeds the system by labeling the region with the correct class. These steps are repeated until the user finishes the process. The final classification is a multiscale result combining all scales of segmentation.

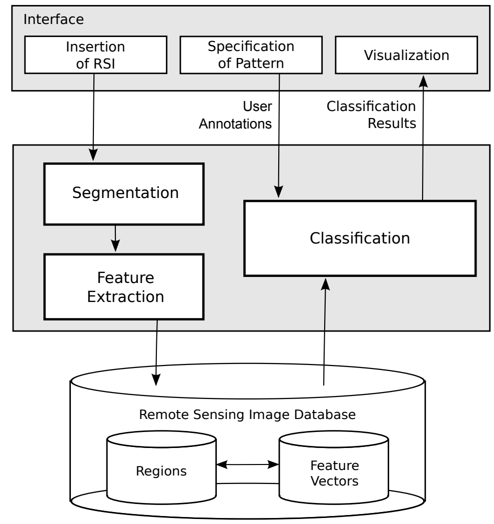


Fig. 4. Architecture of the interactive classification system.

The proposed method is the first one in the literature that considers multiple scales instead of pixel-based information. The experimental results also showed that the combination of scales produces better results than isolated scales in a relevance feedback process. Furthermore, the interactive method achieves good results with few user interactions. The proposed method needs only a small portion of the training set to build classifiers that are as strong as the ones generated by a supervised method that uses the whole training set. Please refer to [14] for a more detailed discussion.

IV. Descriptor Evaluation

Many image descriptors proposed in the literature achieve good results in various applications, but many of them have never been used in remote sensing classification tasks.

The contribution published in [15] presents an evaluation of twelve descriptors that encode spectral/color properties and seven texture descriptors for classification and retrieval tasks of coffee and pasture targets. To evaluate descriptors in classification tasks, we also proposed a methodology based on the KNN (k-nearest neighbors) classifier. Figure 5 presents the results for the COFFEE dataset. Among other interesting conclusions, we could pointed out Joint Autocorrelograms [31] as an effective option to describe coffee and pasture targets.

Another contribution in this context, which was published in [17], is a correlation analysis among a set of descriptors in different segmentation scales. The experiments, illustrated in Figure 6, carried out correlation analysis that confirms different segmentation scales can improve classification results as observed in other works in the literature [9], [32], [33]. However, the experiments also show that not all scales contribute the same way. Coarser scales offer great power of description while the finer ones can improve the classification by detailing the segmentation. Note that the overall correlation between

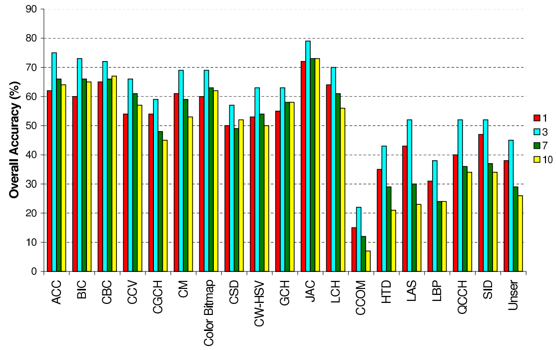


Fig. 5. Overall accuracy classification of each descriptor for the COFFEE dataset, using KNN with k equal to 1, 3, 7 and 10.

scales with regions of different sizes is low. This suggests that the use of different scales improves the classification of RSI according to what have been reported in the literature.

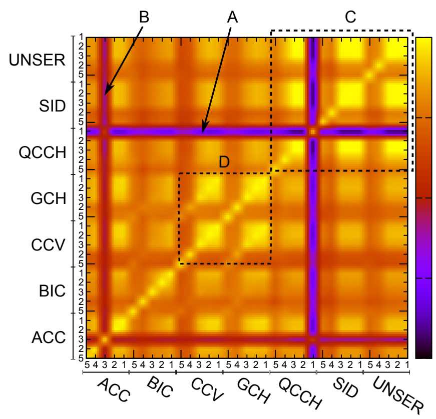


Fig. 6. Complete correlation coefficients for each descriptors at the segmentation scales $\lambda_1, \ldots, \lambda_5$.

It is worth mentioning that there is no work in the literature that applies more descriptors than ours for remote sensing image classification. A more detailed discussion about the evaluation of descriptors and its correlation can be found at [15], [16].

V. Hierarchical Feature Propagation

Several approaches have been proposed for remote sensing image applications to address the segmentation scale problem, by exploiting multiscale analysis considering [7], [9], [14], [32], [34]. In these approaches, the feature extraction at various segmentation scales is an essential step. However, depending on the strategy, the extraction can be a very costly process. If we apply the same feature extraction algorithm for all regions of different segmentation scales, for example, the pixels in the image would need to be accessed at least once for each scale.

The contribution published in [17] presents an strategy called *BoW-Propagation*, which exploits the bag-of-word concept to iteratively propagate texture features along the hierarchy from the finest regions to the coarsest ones. The features are quickly propagated to the upper scales by exploiting the hierarchical association among regions at different scales. The strategy starts by creating a visual dictionary based on low-level features extracted from the pixel level (the base of the hierarchy). The low-level feature space is quantized, creating the visual words, and each region in the base of the hierarchy is described according to that dictionary. The features are then propagated to the other scales. At the end, all regions in the hierarchy will be represented by a bag of visual words. Figure 7 illustrates each step of the proposed approach in an example using three scales.

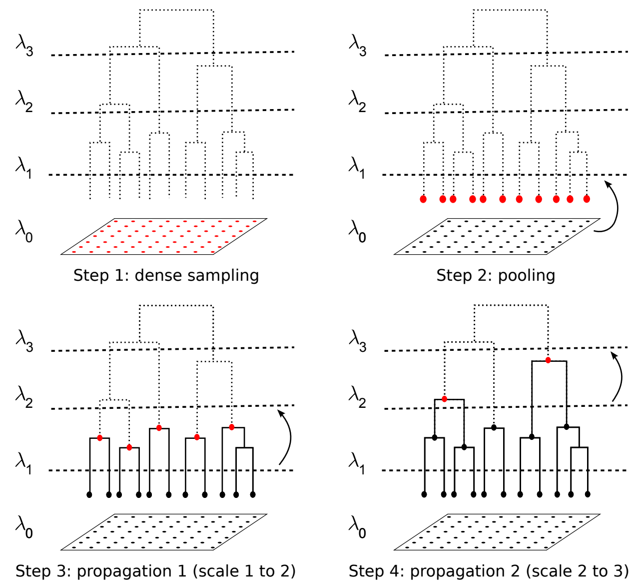


Fig. 7. The BoW-propagation main steps. The process starts with the dense sampling in the pixel level (scale λ_0). Low-level features are extracted from each interest point. Then, in the second step, a feature histogram is created for each region $R \in P_{\lambda_1}$ by pooling the features from the internal interest points. In the third step, the features are propagated from scale λ_1 to scale λ_2. In the fourth step, the features are propagated from scale λ_2 to the coarsest considered scale (λ_3). To obtain the BoWs of a given scale, the propagation is performed by considering the BoWs of the previous scale.

Figure 8 illustrates a schema to represent a segmented region by using dense sampling through a bag of words. The low-level features extracted from the internal points are assigned to visual words and combined by a pooling function.

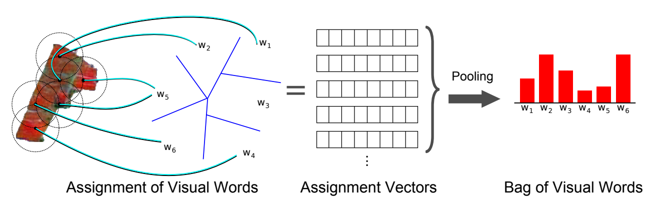
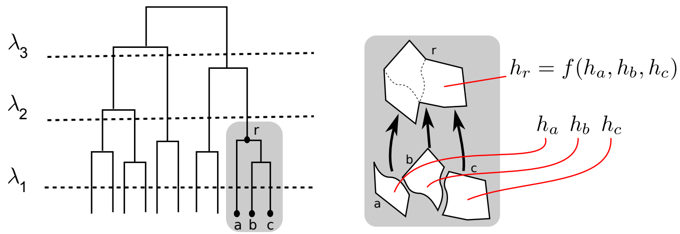


Fig. 8. Schema to represent a segmented region based on a visual dictionary with dense sampling feature extraction.

A contribution published in [18] presents another strategy for hierarchical feature propagation called *H-Propagation*. It consists in estimating the feature histogram representation of a region R, given the low-level histograms extracted from the R subregions. It is an extension of the *BoW-propagation* in



(a) (b)

Fig. 9. Computing a histogram h_r of region r by combining the histograms h_a, h_b, and h_c from the subregions a, b, and c.

the sense it propagates any kind of low-level features based on histograms from fine scales to the coarsest ones. Figure 9 illustrates an example by using the combination function f to compute the histogram h_r of a region r.

H-propagation does not quantize the low-level feature space to create a visual dictionary. Another difference, when compared with the BoW-propagation, is that H-propagation propagates histogram bins instead of the probabilities of visual words. BoW-propagation is suitable for propagating low-level local features. H-propagation, on the other hand, is designed only for global descriptors based on histogram representations.

VI. Conclusion

In this PhD thesis were proposed solutions that address important challenges related to classification of remote sensing images such as data representation, interactivity, and feature extraction. It was completed in four years (from March 2009 to March 2013) and has resulted in six international journal papers [7], [14], [19]–[22], and thirteen international conference papers [2], [5], [6], [15]–[18], [23]–[26], [26]–[28]. Figure 10 summarizes the main contributions of the thesis with the directly related publications.

Future work includes processing of hyperspectral images, multitemporal data, and combination of data from different sensors:

- *Spatio-temporal feature extraction*. This is a topic of great interest not only for the remote sensing community [33], but also in research areas such as Phenology [24]. Some challenges are: how to extract representative features? How to deal with the high dimensionality of the data?
- *Feature extraction from hyperspectral images for object-based classification*. Some challenges are: how to create effective descriptors? How to deal with both spectral and spatial aspects? How to avoid the curse of dimensionality?
- *Combination of features from multiple sensors*. It may involve selection of spectral bands from each sensor. A challenge consists in the adjustment and the maintainance of the georeference among different spatial resolutions.

From the point of view of user interactivity, possible extensions include: new active learning techniques for multiscale classification; and improvements on the visualization and annotation of regions by the user. Furthermore, we wonder if the results may be improved by changing the hierarchy structure along the interactions. This would allow not only the multiscale interactive classification, but also interactive multiscale segmentation.

Acknowledgment

This research was partially supported by FAPESP (grants 2008/58528-2, 2007/52015-0, 2008/57428-4, and 2009/18438-7), CAPES/COFECUB (592/08 and BEX 5224101), and CNPq (grants 306587/2009-2 and 303673/2010-9).

References

[1] G. G. Wilkinson, "Results and implications of a study of fifteen years of satellite image classification experiments," *Geoscience and Remote Sensing, IEEE Trans. on*, vol. 43, pp. 433–440, March 2005.

[2] F. Faria, J. A. dos Santos, R. da S. Torres, A. Rocha, and A. X. Falcão, "Automatic fusion of region-based classifiers for coffee crop recognition," in *IEEE Int. Geosc. and Remote Sensing Symposium*, Munique, Germany, July 2012.

[3] T. Hermosilla, L. Ruiz, J. A. Recio, and M. López, "Assessing contextual descriptive features for plot-based classification of urban areas," *Landscape and Urban Planning*, vol. 106, no. 1, pp. 124 – 137, 2012.

[4] Y. O. Ouma, S. S. Josaphat, and R. Tateishi, "Multiscale remote sensing data segmentation and post-segmentation change detection based on logical modeling: Theoretical exposition and experimental results for forestland cover change analysis," *Computers and Geosciences*, vol. 34, no. 7, pp. 715 – 737, 2008.

[5] R. Nakamura, J. P. Papa, L. Fonseca, J. A. dos Santos, and R. da S. Torres, "Hyperspectral band selection through optimum-path forest and evolutionary-based algorithms," in *IEEE Int. Geosc. and Remote Sensing Symposium*, Munich, Germany, July 2012.

[6] J. A. dos Santos, A. T. da Silva, R. da S. Torres, A. X. Falcão, L. P. Magalhães, and R. A. C. Lamparelli, "Interactive classification of remote sensing images by using optimum-path forest and genetic programming," in *CAIP*, 2011, pp. 300–307.

[7] J. A. dos Santos, P. Gosselin, S. Philipp-Foliguet, R. da S. Torres, and A. X. Falcão, "Multiscale classification of remote sensing images," *Geoscience and Remote Sensing, IEEE Trans. on*, vol. 50, no. 10, pp. 3764–3775, 2012.

[8] M. Kim, T. A. Warner, M. Madden, and D. S. Atkinson, "Multi-scale geobia with very high spatial resolution digital aerial imagery: scale, texture and image objects," *Int. Journal of Remote Sensing*, vol. 32, no. 10, pp. 2825–2850, 2011.

[9] Y. Tarabalka, J. Tilton, J. Benediktsson, and J. Chanussot, "A marker-based approach for the automated selection of a single segmentation from a hierarchical set of image segmentations," *Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of*, vol. 5, no. 1, pp. 262 –272, feb. 2012.

[10] D. C. Duro, S. E. Franklin, and M. G. Dubé, "Multi-scale object-based image analysis and feature selection of multi-sensor earth observation imagery using random forests," *Int. Journal of Remote Sensing*, vol. 33, no. 14, pp. 4502–4526, 2012.

[11] D. Lu and Q. Weng, "A survey of image classification methods and techniques for improving classification performance," *Int. Journal of Remote Sensing*, vol. 28, no. 5, pp. 823–870, 2007.

[12] A. B. Santos, A. de A Araujo, and D. Menotti, "Combiner of classifiers using genetic algorithm for classification of remote sensed hyperspectral images," in *IEEE Int. Geosc. and Remote Sensing Symposium*, 2012, pp. 4146–4149.

[13] J. A. dos Santos, "Semi-automatic classification of remote sensing images," Ph.D. dissertation, University of Campinas, Campinas, Brazil, March 2013, in cotute with University of Cergy-Pontoise, France.

[14] J. A. dos Santos, P. Gosselin, S. Philipp-Foliguet, R. da S. Torres, and A. X. Falcão, "Interactive multiscale classification of high-resolution remote sensing images," *Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of*, 2013, to appear.

[15] J. A. dos Santos, O. A. B. Penatti, and R. da S. Torres, "Evaluating the potential of texture and color descriptors for remote sensing image retrieval and classification," in *The Int. Conference on Computer Vision Theory and Applications*, Angers, France, May 2010, pp. 203–208.

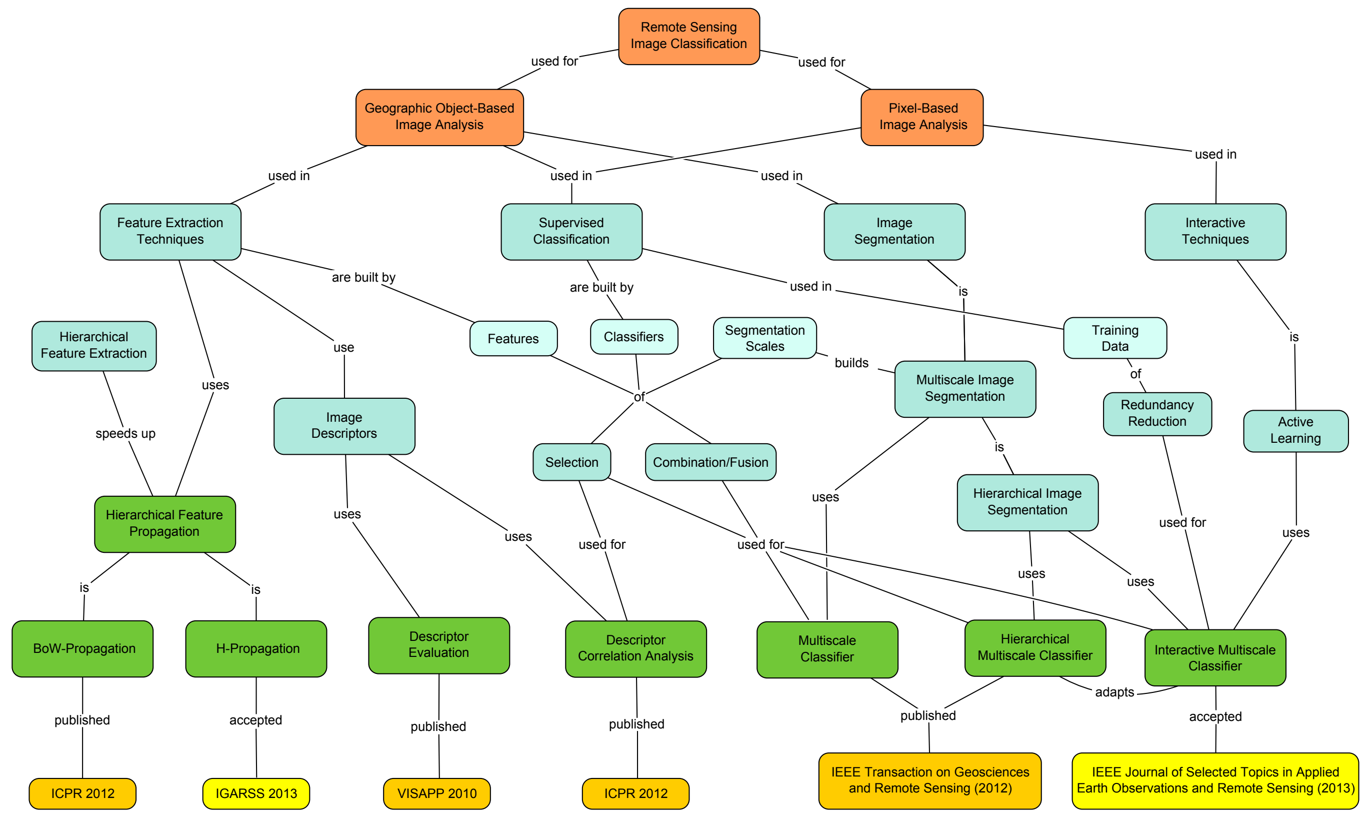


Fig. 10. Conceptual map of the thesis. In green, the main contributions. In Yellow, the directly related publications associated to this work. In orange and ligth blue, the main concepts.

[16] J. A. dos Santos, F. A. Faria, R. da S. Torres, A. Rocha, P.-H. Gosselin, S. Philipp-Foliguet, and A. X. Falcão, "Descriptor correlation analysis for remote sensing image multi-scale classification," in *ICPR*, Tsukuba, Japan, November 2012.

[17] J. A. dos Santos, O. A. B. Penatti, R. da S. Torres, P.-H. Gosselin, S. Philipp-Foliguet, and A. X. Falcão, "Improving texture description in remote sensing image multi-scale classification tasks by using visual words," in *ICPR*, Tsukuba, Japan, November 2012.

[18] ——, "Remote sensing image representation based on hierarchical histogram propagation." in *IEEE Int. Geosc. and Remote Sensing Symposium*, Melbourne, Australia, 2013, to apear.

[19] R. Nakamura, L. Fonseca, J. A. dos Santos, Y. X.-S. Torres, R.da S., and J. Papa, "A nature-inspired framework for hyperspectral band selection," *Geoscience and Remote Sensing, IEEE Trans. on*, 2013, to appear.

[20] A. T. da Silva, J. A. dos Santos, A. X. Falcão, R. da S. Torres, and L. P. Magalhães, "Incorporating multiple distance spaces in optimum-path forest classification to improve feedback-based learning," *Computer Vision and Image Understanding*, vol. 116, no. 4, pp. 510–523, 2012.

[21] J. A. dos Santos, C. D. Ferreira, R. da S.Torres, M. A. Gonçalves, and R. A. C. Lamparelli, "A relevance feedback method based on genetic programming for classification of remote sensing images," *Information Sciences*, vol. 181, no. 13, pp. 2671–2684, 2011.

[22] C. D. Ferreira, J. A. dos Santos, R. da S. Torres, M. A. Gonçalves, R. C. Rezende, and W. Fan, "Relevance feedback based on genetic programming for image retrieval," *Pattern Recognition Letters*, vol. 32, no. 1, pp. 27–37, 2011.

[23] R. .da S. Torres, M. Hasegawa, S. Tabbone, J. Almeida, J. A. dos Santos, B. Alberton, and P. Morellato, "Shape-based time series analysis for remote phenology studies," in *IEEE Int. Geosc. and Remote Sensing Symposium*, Melbourne, Australia, 2013, to appear.

[24] J. Almeida, J. A. dos Santos, B. Alberton, R. da S. Torres, and L. P. C. Morellato, "Remote phenology: Applying machine learning to detect phenological patterns in a cerrado savanna," in *8th IEEE Int. Conference on eScience*, Chicago, USA, October 2012.

[25] F. Faria, J. A. dos Santos, R. da S. Torres, and A. Rocha, "Automatic classifier fusion for produce recognition," in *SIBGRAPI 2012*, Ouro Preto-MG, Brazil, August 2012.

[26] G. de S. Fedel, C. B. Medeiros, and J. A. dos Santos, "Sinimbu - multimodal queries to support biodiversity studies," in *ICCSA (1)*, 2012, pp. 620–634.

[27] C. G. N. Macário, J. A. dos Santos, C. M. B. Medeiros, and R. da S. Torres, "Annotating data to support decision-making: a case study," in *6th Workshop on Geog. Inf. Retrieval*, Zurich, Switzerland, 2010, pp. 1–7.

[28] J. A. dos Santos, F. A. Faria, R. T. Calumby, R. da S. Torres, and R. A. C. Lamparelli, "A genetic programming approach for coffee crop recognition," in *IEEE Int. Geosc. and Remote Sensing Symposium*, Honolulu, USA, July 2010, pp. 3418–3421.

[29] R. Trias-Sanz, G. Stamon, and J. Louchet, "Using colour, texture, and hierarchial segmentation for high-resolution remote sensing," *Journal of Photog. and Remote Sensing*, vol. 63, no. 2, pp. 156 – 168, 2008.

[30] D. Tuia, M. Volpi, L. Copa, M. Kanevski, and J. Munoz-Mari, "A survey of active learning algorithms for supervised remote sensing image classification," *Selected Topics in Signal Processing, IEEE Journal of*, vol. 5, no. 3, pp. 606 –617, june 2011.

[31] A. Williams and P. Yoon, "Content-based image retrieval using joint correlograms," *MTAP*, vol. 34, no. 2, pp. 239–248, 2007.

[32] A. Alonso-González, S. Valero, J. Chanussot, C. López-Martínez, and P. Salembier, "Processing multidimensional sar and hyperspectral images with binary partition tree," *Proceedings of the IEEE*, vol. PP, no. 99, pp. 1 –25, 2012.

[33] F. Petitjean, C. Kurtz, N. Passat, and P. Gançarski, "Spatio-temporal reasoning for the classification of satellite image time series," *Pattern Recognition Letters*, vol. 33, no. 13, pp. 1805 – 1815, 2012.

[34] J. Munoz-Mari, D. Tuia, and G. Camps-Valls, "Semisupervised classification of remote sensing images with active queries," *Geoscience and Remote Sensing, IEEE Trans. on*, vol. 50, no. 10, pp. 3751 –3763, oct. 2012.