Superpixel-based interactive classification of very high resolution images

John E. Vargas*, Priscila T. M. Saito*†, Alexandre X. Falcão*, Pedro J. de Rezende* and Jefersson A. dos Santos‡
*Institute of Computing, University of Campinas, Campinas, Brazil
Email: jvargas@liv.ic.unicamp.br, {maeda, afalcao, rezende}@ic.unicamp.br
†Department of Computer Engineering, Federal Technological University of Parana, Cornelio Procopio, Brazil
‡Department of Computer Science, Universidade Federal de Minas Gerais, Belo Horizonte, Brazil
Email: jefersson@dcc.ufmg.br

***Abstract*—Very high resolution (VHR) images are large datasets for pixel annotation — a process that has depended on the supervised training of an effective pixel classifier. Active learning techniques have mitigated this problem, but pixel descriptors are limited to local image information and the large number of pixels makes the response time to the user's actions impractical, during active learning. To circumvent the problem, we present an active learning strategy that relies on superpixel descriptors and *a priori* dataset reduction. Firstly, we compare VHR image annotation using superpixel- and pixel-based classifiers, as designed by the same state-of-the-art active learning technique — *Multi-Class Level Uncertainty* (MCLU). Even with the dataset reduction provided by the superpixel representation, MCLU remains unfeasible for user interaction. Therefore, we propose a technique to considerably reduce the superpixel dataset for active learning. Moreover, we subdivide the reduced dataset into a list of subsets with random sample rearrangement to gain both speed and sample diversity during the active learning process.**

***Keywords*-Very high resolution image processing; active learning; superpixel representation; dataset reduction; supervised classification.**

I. Introduction

Pixel annotation in hyperspectral and very high resolution (VHR) images has relied on supervised classifiers so far [1]–[4]. However, manual selection and annotation of a sufficiently large number of training pixels is unfeasible for very large datasets. In order to handle large datasets, active learning techniques have been proposed for selecting a small training set that represents well not only all classes under annotation but also discriminative samples on the boundary between classes [1], [2], [5]–[8].

Such active learning process starts off with a small training set for manual annotation. The labeled samples are then used to train a preliminary classifier that selects and labels new samples for user validation. The user may confirm/correct the labels of the selected samples, increasing the size of the training set to generate an improved classifier for the next iteration. The process may be halted by the user once the accuracy of the classifier reaches a certain level, which can be verified either by cross validation on the training set or, regarding images, by visualizing the classification of the entire image.

The effectiveness of active learning techniques has been substantiated in the literature as superior to a simple random selection of samples [1], [2] even when much larger amounts of randomly selected samples are considered. Nonetheless, the performance of an active learning technique depends on a clever strategy for selection of representative and discriminative samples from the large unlabeled dataset. The Multi-Class Level Uncertainty (MCLU) method is considered a state-of-the-art approach for pixel annotation in VHR and hyperspectral images [2]. However, works in this line evaluate their methods restricted to a subset of the image domain for which the pixel labels can be easily determined by the user. Moreover, the pixel datasets are labeled and the methods are assessed without considering the mean response time for the user's actions. Indeed, they are impractical from the user's point of view. Besides, pixel descriptors are also limited to local image information and susceptible to noise.

We circumvent these problems by presenting an active learning strategy based on the MCLU method that relies on superpixel descriptors and *a priori* dataset reduction. The image segmentation into a superpixel representation is done by the Simple Linear Iterative Clustering algorithm (SLIC), described in [9]. Other effective algorithms [10]–[13] could also be used, but we found SLIC very accurate when it comes to obtaining superpixels for annotation that respect the boundaries between distinct classes. The superpixel representation initially provides a considerable dataset reduction and also allows the combination of higher-level descriptors within regions that respect the boundaries between classes.

Our contribution. Firstly, we demonstrate the effective gain of the superpixel representation over the pixel representation using the MCLU active learning technique over an entire VHR image. Secondly, given that, as such, the method is still impractical for interactive response times, we propose an *a priori* superpixel dataset reduction using the Optimum-Path Forest (OPF) clustering technique [14]. The choice for OPF clustering is justified by previous works on active learning [8], [15] and its ability to find natural clusters with no shape constraints or need to specify a number of clusters. Under OPF, each cluster is an optimum-path tree rooted at a maximum of the dataset's probability density function. The roots are likely to cover all classes, and the samples on the boundaries between clusters are likely to be the most discriminative ones for class separation. Therefore, our approach selects the roots and a few random samples from the boundary set to compose the first training set for manual annotation. The MCLU technique is based on Support Vector Machine (SVM) classification. Hence, the first instance of the SVM classifier is applied to select a reduced set from the entire dataset with the most uncertainty samples. Moreover, we subdivide the reduced dataset into m subsets and include them in a queue to improve efficiency. The active learning process uses one subset per iteration until the queue is empty. Then, the unlabeled samples of these subsets are merged and randomly rearranged to compose a new queue of subsets, in order to attain sample diversity in the subsets for the next m iterations. The result is a feasible and very effective active learning approach, as demonstrated here.

Organization. Section II presents background on active learning, explains the MCLU method, and describes the SLIC algorithm for superpixel representation. Section III explains our proposal for data reduction and efficiency gain in active learning. Section IV describes the experiments that compare pixel- versus superpixel-based classifiers and shows the efficiency gain in the selection process of the MCLU technique with the superpixel representation. Section V states our conclusion and discusses future work.

II. Active learning for image classification

Active learning techniques aim to iteratively select the most informative samples from a given unlabeled dataset U for training a supervised classifier. In the first step, a few unlabeled samples are selected for manual annotation, forming an initial training set T. A classifier C is trained with the labeled set T and a query function Q uses C to classify and select a set X of a few more informative unlabeled samples from U. Then the user confirms or corrects the labels of the selected samples. The labeled samples in X are included into T and the classifier C is retrained with the new training set. This process is repeated until the user is satisfied with the accuracy of the classifier C. Algorithm 1 describes a general procedure for active learning.

Algorithm 1 A general procedure for active learning.

1: Select a set $T \subset U$ of unlabeled samples for user annotation.
2: Train an initial classifier C with the annotated set T.
3: Classify the unlabeled samples in U.
4: **repeat**
5: The query function Q uses C to label and select a new set X of samples from U.
6: The user confirms/corrects the labels of X and T is increased by X.
7: Retrain the classifier C with the new set T.
8: **until** the user is satisfied with the accuracy of C, as observed in Line 6.

Generally, active learning techniques only differ on the query function Q. In [2], the method Multi-Class Level Uncertainty (MCLU) is proposed for the problem of remote-sensing image classification, demonstrating high accuracy rates for the problem.

A. Multi-Class Level Uncertainty

The Multi-Class Level Uncertainty (MCLU) technique is based on Support Vector Machine (SVM) classification [16] — one of the most successful methods in pattern recognition, which has inspired many other active learning techniques [7], [17]–[20]. The MCLU technique selects the unlabeled samples with higher uncertainty in classification among all samples in the dataset U. The uncertainty criterion is defined by a function $c(x)$, $x \in U$, which is calculated on the basis of the signed distances $d_i(x)$, $i = 1, 2, \ldots, n$, from sample x to each of the n decision boundaries (n classes) of the one-versus-all SVM classifier [21]. The distance set $\{d_1(x), \ldots, d_n(x)\}$ for every $x \in U$ is computed and the confidence value $c(x)$ can be computed in two ways, as explained in [2]: 1) The smallest distance to the hyperplane given by

$$c_{\text{min}}(x) = \min_{i=1,2,\ldots,n}\{abs(d_i(x))\} \tag{1}$$

and 2) The difference between the first and second largest distance values to the hyperplanes given by

$$\max{}_1 = \arg\max_{i=1,2,\ldots,n}\{d_i(x)\} \tag{2}$$

$$\max{}_2 = \arg\max_{i=1,2,\ldots,n, i\neq \max_1}\{d_i(x)\} \tag{3}$$

$$c_{\text{diff}}(x) = d_{\max_1}(x) - d_{\max_2}(x). \tag{4}$$

The function $c_{\text{min}}(x)$ models the uncertainty taking into account the minimum distance to the hyperplanes. Moreover, the function $c_{\text{diff}}(x)$ models the uncertainty using the two most likely classes. If $c_{\text{diff}}(x)$ is small, the label attribution is more uncertain. In this work, we use MCLU with $c_{\text{diff}}(x)$ function to model uncertainty. The $|X|$ (cardinality of X) samples with lower $c(x)$ values are selected to be displayed to the user. The MCLU technique is validated in [2] on one VHR image and two hyperspectral images. In both cases, they used pixels as samples and their values in each band to compose the pixel descriptors. In this work, we propose the use of superpixels as samples for VHR image classification.

B. Superpixel generation

Recently, superpixels have received special attention to speed up graph-based image segmentation [22]. Image segmentation into superpixels aims at grouping pixels into homogeneous regions that capture the image redundancy. The superpixel representation allows higher-level image descriptors as compared to the pixel representation. In [9], many state-of-the-art techniques are compared to each other under two metrics: boundary recall and under-segmentation error. The Simple Linear Interactive Clustering (SLIC) technique was found to be effective for both metrics. Therefore, we use the SLIC technique in order to divide the VHR image into homogeneous regions (superpixels).

SLIC is based on the k-means algorithm and relies on two parameters: the number of desired superpixels $|U|$ and a compactness parameter t. Each pixel is described by a vector $\vec{v}$ of its Lab color bands and the pixel location (horizontal and vertical coordinates): $\vec{v} = [L, a, b, x, y]$. The procedure begins with the initialization of the $|U|$ cluster center vectors: $C_1, C_2, ..., C_{|U|}$. Those clusters are sampled on a regular grid spaced S pixels apart, $S = \sqrt{N/|U|}$, where N is the number of pixels of the image. To avoid seeding a superpixel in a noisy location, the centers are moved to the lowest gradient location in a 3×3 neighborhood. Then, each pixel is assigned to the nearest cluster center in a region of $2S \times 2S$ window around it, which reduces the number of distance calculations. Equation 5 is used to compute the distance between a pixel p and a cluster center C_i:

$$D = \sqrt{d_c^2 + \left(\frac{d_s}{S}\right)^2 t^2}, \tag{5}$$

where d_c is the Euclidean distance between the color of p and the cluster center C_i, and d_s is the Euclidean distance between the location of p and the location of the cluster center C_i. The compactness parameter t weighs the relative importance of color similarity and spatial proximity.

The update step moves the cluster centers to the mean vector of all the pixels that belong to the cluster. The assignment and update steps are repeated until convergence or a limit of iterations is achieved. Finally, a post-processing step assigns the disjoint pixels to one of the $|U|$ cluster centers, reinforcing the connectivity and oversegmenting the image into $|U|$ superpixels. Figure 1 shows an image fragmented into superpixels as obtained by SLIC.

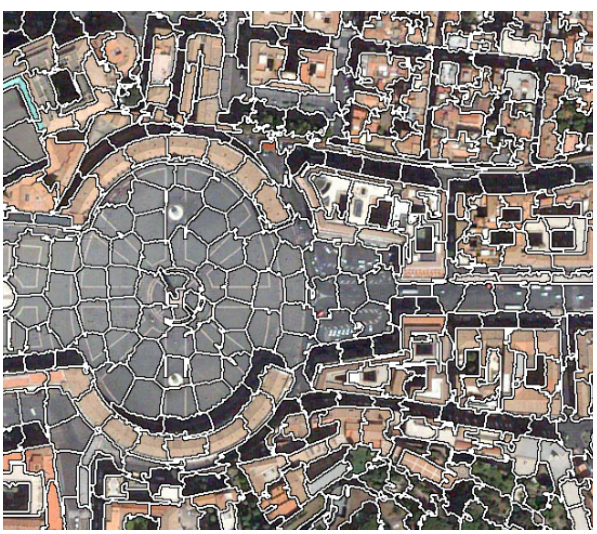

Fig. 1. Superpixels superimposed onto the original image

In typical VHR image classification tasks, the noise of the color/intensity features of the pixels may cause misclassification. In contrast, the superpixel texture features provide a more robust and higher-level image description. The feature vector used in the pixel-based classification of VHR images is the color bands of the pixel. In the case of superpixels, many color and texture descriptors can be used. We chose to use the concatenation of four region descriptors used in image segmentation and classification tasks [22], [23]:

1) Mean color: Mean color of the pixels contained in the superpixel.

2) Color histogram [24]: It is a widely used descriptor, which quantizes the color space in an uniform way and counts the number of pixels that belong to each bin. The size of the descriptor depends on the number of bins. In this work, we split the color space into 512 bins.

3) Local binary patterns (LBP): This method encodes the texture features from grayscale intensity images by comparing each pixel intensity with the intensities of its neighboring pixels [25]. An eight bit code is assigned to every pixel, based on its eight neighboring intensity values (3×3 neighborhood of the pixel). A bit i of the 8-bit code is set to 1 when the i-th neighbor of a pixel p has an intensity value higher than or equal to the one of p. The descriptor is a 256-bin histogram of the pixel codes.

4) Border interior pixel classification (BIC) [26]: This descriptor has been successfully applied to content-based image retrieval [27] and classification of remote sensing images [23]. In this method, the pixels in the image are classified as border or interior pixels. The pixel colors are quantitized and a pixel is classified as interior when it is in the same level of its four neighbors. Otherwise, it is classified as border pixel. Then, a 512-bin histogram computed for interior pixels is concatenated with a 512-bin histogram computed for border pixels.

III. Dataset reduction method

Active learning methods usually randomly select the initial training set and classify and/or sort the entire dataset at every iteration of the process. For large datasets, the latter makes the response time for the user's actions impractical [8].

In this work, we propose a data reduction method to circumvent the problem. Inspired on a recent work [8], our approach also explores the Optimum-Path Forest (OPF) clustering algorithm [14] for dataset reduction and initial training sample selection. However, differently from [8], we select the initial training samples as the representative cluster samples and randomly selected samples from the boundary between clusters; and then reduce the dataset by using the resulting SVM classifier from the initial training set.

The OPF clustering algorithm represents a dataset as a graph whose nodes are samples and each node is connected with its k-nearest neighbors in the feature space. The probability density function (pdf) in every node is estimated from the distance between adjacent samples and a connectivity function is designed so that the maximization of a connectivity map defines an optimum-path forest, and each maximum of the pdf becomes root of an optimum-path tree (cluster), composed by the most strongly connected samples to the maximum than to any other root.

The initial classifier should be as effective as possible because the proposed reduction method is based on the classification results of the first classifier. Clustering is performed in order to find representative samples that are likely to cover all classes, as well as, uncertainty samples between classes, which are likely to be on the boundary between clusters. The roots of the clusters (maxima of the pdf) and a small set of randomly selected boundary samples are then annotated by the user to form a small initial training set T.

Figure 2 illustrates the pipeline of the proposed active learning method based on dataset reduction. Before the learning cycle, the samples of the labeled set T are used to train the classifier C. Then, the query function Q uses C to classify the entire set of unlabeled samples U. These unlabeled samples are sorted according to the uncertainty criterion in Equation 4 — lower is the confidence value $c(x)$, higher is the uncertainty value of the sample x. The $|U|/r$ most uncertainty samples form the reduced set R. Note that the value of r controls the size of the reduced set.

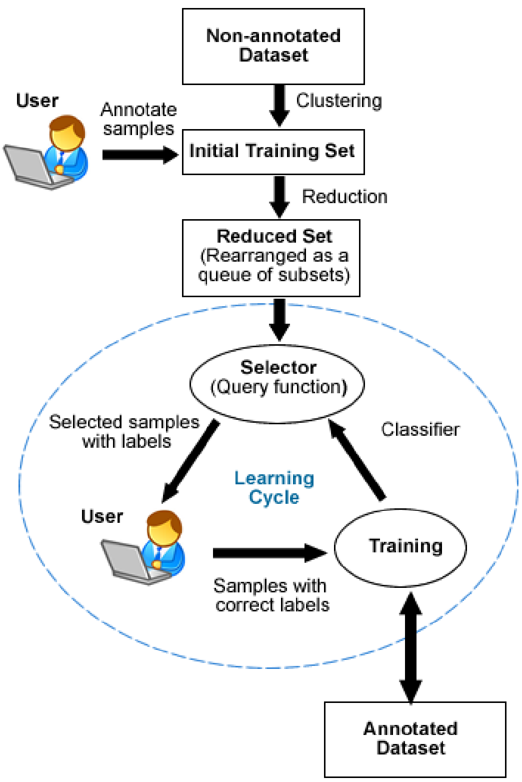

Fig. 2. Pipeline of the proposed data reduction active learning method

In order to further reduce the number of unlabeled samples that have to be classified in the query function Q at every iteration, we propose to split the reduced set R into m equal-sized subsets $\{R_1, \ldots, R_m\}$ of random samples, and include these subsets in a queue L for processing, one per iteration, until L is empty. After the m-th iteration, the unlabeled samples in R are merged, randomly divided into new m subsets $\{R_1, \ldots, R_m\}$, and included again in L for the next m iterations. This strategy also aims to attain sample diversity in these subsets along the iterations. Algorithm 2 describes the proposed method.

Algorithm 2 Active learning algorithm with dataset reduction.

1: Cluster the set of unlabeled samples U with OPF and select root and boundary samples, which are annotated by the user to form the initial training set T.
2: Train the classifier C with the labeled set T.
3: The query function Q uses C to classify and sort the samples in the unlabeled set U according to the uncertainty criterion.
4: Select the $|U|/r$ most uncertainty samples from U to form the reduction set R.
5: Divide the set R into m subsets $\{R_1, ..., R_m\}$ and put them in the queue of subsets L.
6: **repeat**
7:    The query function Q uses C to classify and select a set of samples X from the next subset of unlabeled samples in the queue of subsets L.
8:    The user confirms/corrects labels and T is increased by X.
9:    Retrain the classifier C with the new set T.
10:    If the m subsets are processed, then merge the subsets $\{R_1, ..., R_m\}$ and divide them again into m subsets for L.
11: **until** the user is satisfied with the accuracy of C.

In the learning cycle of the active learning process (Line 6), the query function Q uses C to classify and select a set of samples X, which are the most uncertain samples in the next subset of the queue of subsets L. In our implementation, the query function Q used is the MCLU. The user confirms or corrects the labels of the samples selected by Q and the set T is increased by X (Line 8). Then, the classifier C is trained with the new training set T. If the m subsets in the queue L have been processed, the subsets are combined and randomly divided again into m subsets and rearranged in a queue of subsets L (Line 10). These steps are repeated until the user is satisfied with the accuracy of the classifier C. Note that any other query function could be used instead of the MCLU technique.

IV. Experiments

In this section, we describe the experiments conducted to compare superpixel- and pixel-based classification, and to evaluate the proposed data reduction method. In all experiments, each method was run ten times, so each graph reports the mean of these runs.

A. Dataset

Given that annotated VHR images are presently scarce, we opted for manually annotating a VHR image acquired on the Vatican City in April 2004. It is a 2847×2817 natural color image obtained from the *Mapmart QuickBird satellite imagery samples* website. The image was labeled with seven classes of interest (Road, Shadow, Tree, Water, Building, Grass, and Bare soil). In many remote sensing image classification works [1],

[2], [28]–[30], the labeled dataset is restricted to some part of the complete image and this set of samples is often partitioned into three sets: training, learning, and test. In the present work, we focus on a real-world application where the labeling process should be applied to all pixels in the image. In all experiments, we selected an initial small set comprised of 64 samples. For the proposed approach, the 64 samples include the roots of the clusters (in order to guarantee samples from all classes) and samples from the boundary between clusters. The rest of the samples form the learning set. In every iteration of the active learning process, the training set grows and the number of samples in the learning set decreases. The VHR image and the ground truth image are shown in Figure 3.

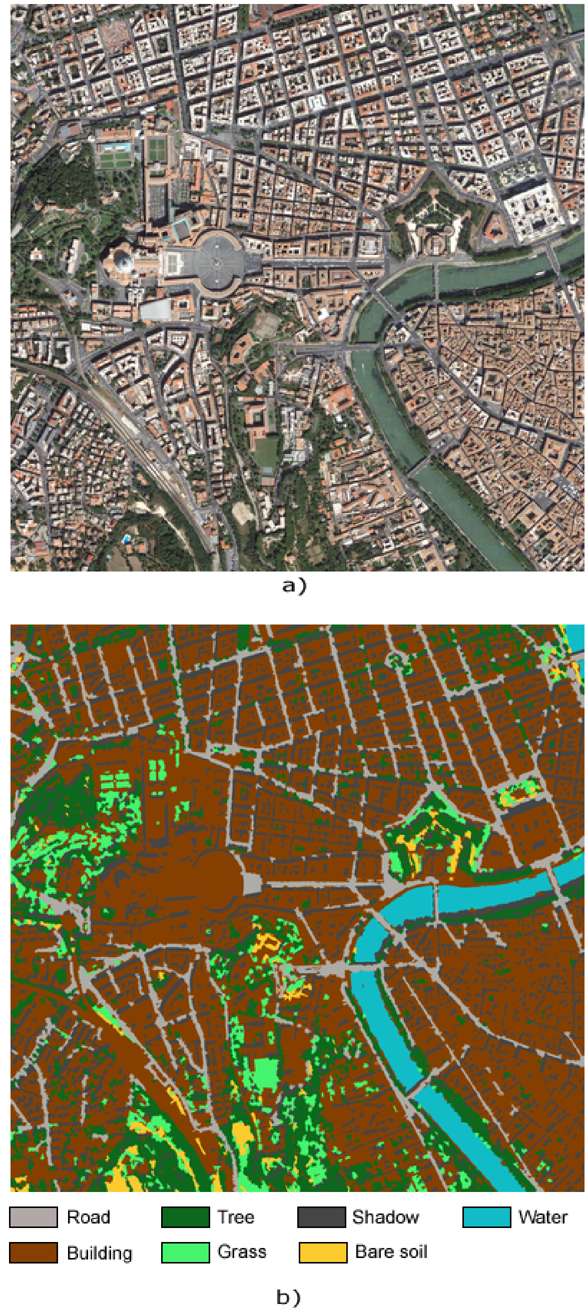

Fig. 3. a) Vatican City, VHR image b) Ground truth image

B. Superpixel vs Pixel

To do a comparison between the pixel- and superpixel-based approaches for the task of VHR image classification, we used the active learning method MCLU with the same number of initial training samples and an equal number of samples labeled per iteration. Note that retrieving N samples per iteration, in the case of pixel representation, also limits the size of the region covered in the image with respect to the superpixel representation, in each iteration. Retrieving more pixels per iteration than superpixels would result in a considerably higher user effort in verifying the classes suggested for the returned samples. In all experiments, the active learning methods select 14 samples, twice the number of classes (as suggested in [31]), to be validated by the user in each iteration. As explained in Section II, the descriptor used for superpixels is a combination of four descriptors: Mean color, Color histogram, Local binary patterns (LBP), and Border/ Interior Classification (BIC). The compactness parameter used to generate the superpixels by the SLIC algorithm was $t = 15$, and the color of the pixel in the Lab color space was chosen as pixel descriptor. The overall accuracy of the two approaches is presented in Figure 4. The superpixel-based classification is consistently superior in comparison with the pixel-based classification. In both approaches, the initial training sets are composed of 64 randomly selected samples. In order to set the parameters of the SVM-RBF classifier of the MCLU technique used in the experiments, we performed a grid-search model selection in the first and fifth iterations for the pixel and superpixel-based approaches.

C. Dataset reduction

The VHR image used in the experiments is comprised of $|U| = 24968$ superpixels obtained by the SLIC algorithm. The time in the selection process of the MCLU method is not suitable for an application where interactive response times are expected (we measured 87 seconds on average). As baseline for the comparison of the superpixel-based methods, we used a random sampling method that selects samples from the learning set to be validated by the user. The proposed MCLU method of data reduction was validated with the same configurations as the MCLU method. Thirty percent of the entire set of unlabeled samples was selected to form the reduced set ($1/r = 0.3$), which was divided into $m = 4$ subsets. Note that these two parameters controls the number of samples selected in each iteration. In Figure 5, we present the overall accuracy for the proposed MCLU method with data reduction in comparison with the original MCLU method and a Random Sampling selection strategy. In Figure 6, we present the time spent per iteration for the both MCLU selection strategies.

As illustrated in Figure 5, the accuracy of the MCLU with data reduction is comparable to the one achieved by

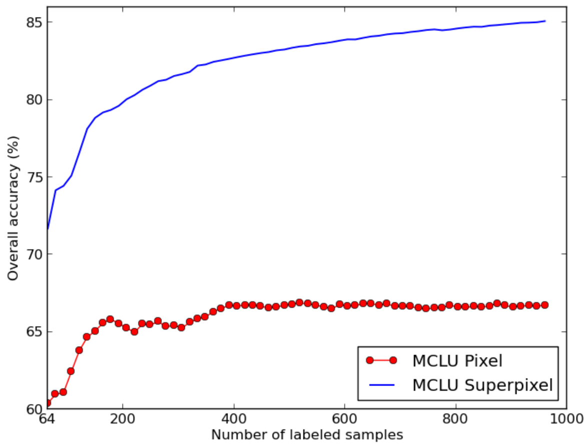

Fig. 4. Overall accuracy Superpixel vs Pixel classification.

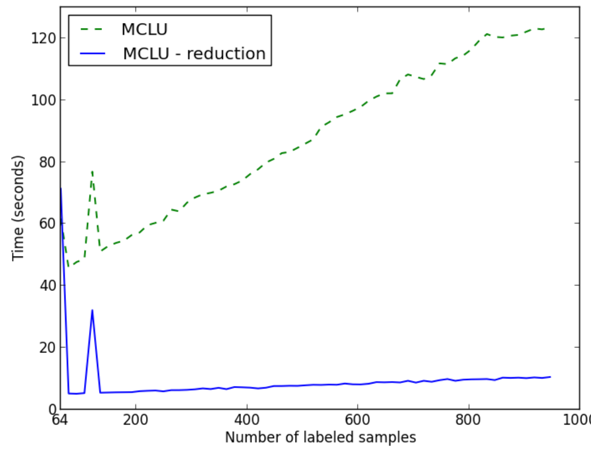

Fig. 6. Time by iteration MCLU and proposed MCLU-data reduction.

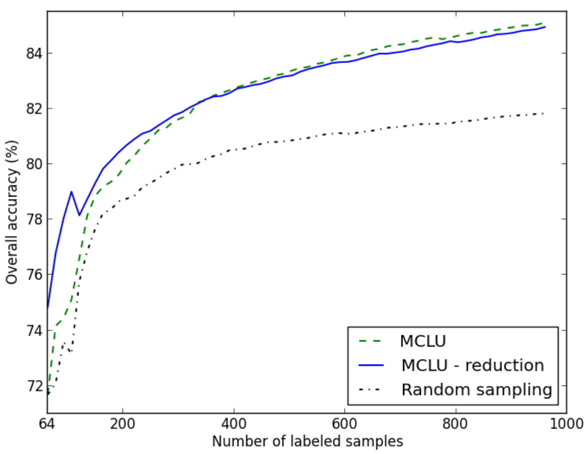

Fig. 5. Overall accuracy MCLU and proposed MCLU-data reduction.

the original MCLU method. However, the computing time is remarkably smaller as one can observed in Figure 6.

The data reduction process and the grid-search performed for parameter selection of the SVM-RBF classifier causes the peak to be in the first and fifth iterations as seen in Figure 6.

V. Conclusion

In this paper, we have shown that interactive classification of very high resolution images, through active learning sessions, benefits more from a superpixel-based approach than from a pixel-based one.

We have also proposed a method that reduces and rearranges the dataset, lowering the computing time of the selection process (from 87 to 9 seconds on average) and keeping similar accuracy in comparison with the active learning method that processes the entire set of unlabeled samples. This cutback in time makes user interaction feasible in active learning.

One of the main reasons for the success of our method is the effectiveness of the procedure that selects samples for the initial training set, which allows us to design an effective first classifier for data reduction.

As future work, we plan to extend the proposed data reduction method by including a diversity criterion in the selection process. We also intend to carry out experiments to evaluate our method in an ample range of hyperspectral images.

Acknowledgment

This research was supported by grants from Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) #477692/2012-5, #303673/2010-9, #479070/2013-0, #131832/2013-1, #141795/2010-7, #552559/2010-5, Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) #2007/52015-0, Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG), PRPq/Universidade Federal de Minas Gerais, and FAEPEX/UNICAMP.

We thank Mapmart for the QuickBird satellite imagery.

References

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

[2] B. Demir, C. Persello, and L. Bruzzone, "Batch-mode active-learning methods for the interactive classification of remote sensing images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 3, pp. 1014–1031, March 2011.

[3] S. Rajan, J. Ghosh, and M. Crawford, "An active learning approach to hyperspectral data classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 46, no. 4, pp. 1231–1242, April 2008.

[4] D. Tuia, E. Pasolli, and W. Emery, "Using active learning to adapt remote sensing image classifiers," *Remote Sensing of Environment*, vol. 115, no. 9, pp. 2232 – 2242, 2011.

[5] J. A. dos Santos, P.-H. Gosselin, S. Philipp-Foliguet, R. Da S Torres, and A. X. Falcão, "Interactive multiscale classification of high-resolution remote sensing images," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 6, no. 4, pp. 2020–2034, Aug 2013.

[6] M. Li and I. Sethi, "Confidence-based active learning," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 8, pp. 1251–1261, Aug 2006.

[7] K. Brinker, "Incorporating diversity in active learning with support vector machines," in *Proceedings of the Twentieth International Conference on Machine Learning*, 2003.
[8] P. T. M. Saito, P. J. de Rezende, A. X. Falcão, C. T. N. Suzuki, and J. F. Gomes, "Improving active learning with sharp data reduction," in *WSCG International Conference on Computer Graphics, Visualization and Computer Vision*, 2012, pp. 27–34.
[9] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Susstrunk, "Slic superpixels compared to state-of-the-art superpixel methods," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 11, pp. 2274–2282, Nov 2012.
[10] P. F. Felzenszwalb and D. P. Huttenlocher, "Efficient graph-based image segmentation," *International Journal of Computer Vision*, vol. 59, no. 2, pp. 167–181, Sep. 2004.
[11] A. Moore, S. Prince, J. Warrell, U. Mohammed, and G. Jones, "Superpixel lattices," in *IEEE Conference on Computer Vision and Pattern Recognition*, June 2008, pp. 1–8.
[12] A. Levinshtein, A. Stere, K. N. Kutulakos, D. J. Fleet, S. J. Dickinson, and K. Siddiqi, "Turbopixels: Fast superpixels using geometric flows," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, no. 12, pp. 2290–2297, Dec. 2009.
[13] O. Veksler, Y. Boykov, and P. Mehrani, "Superpixels and supervoxels in an energy optimization framework," in *European Conference on Computer Vision*, 2010.
[14] L. M. Rocha, F. A. M. Cappabianco, and A. X. Falcão, "Data clustering as an optimum-path forest problem with applications in image analysis," *Int. J. Imaging Syst. Technol.*, vol. 19, no. 2, pp. 50–68, Jun. 2009.
[15] P. T. M. Saito, P. J. de Rezende, A. X. Falcão, C. T. N. Suzuki, and J. F. Gomes, "A data reduction and organization approach for efficient image annotation," in *Proceedings of the 28th Annual ACM Symposium on Applied Computing*, ser. SAC '13. New York, NY, USA: ACM, 2013, pp. 53–57.
[16] V. N. Vapnik, *The Nature of Statistical Learning Theory*. New York, NY, USA: Springer-Verlag New York, Inc., 1995.
[17] G. Schohn and D. Cohn, "Less is more: Active learning with support vector machines," in *Proceedings of the Seventeenth International Conference on Machine Learning*, ser. ICML '00. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 2000, pp. 839–846.
[18] S. Tong and D. Koller, "Support vector machine active learning with applications to text classification," *Journal of Machine Learning Research*, vol. 2, pp. 45–66, Mar. 2002.
[19] P. Mitra, B. U. Shankar, and S. K. Pal, "Segmentation of multispectral remote sensing images using active support vector machines," *Pattern Recognition Letters*, vol. 25, no. 9, pp. 1067 – 1074, 2004.
[20] L. Bruzzone and C. Persello, "Active learning for classification of remote sensing images," in *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, vol. 3, July 2009, pp. III–693–III–696.
[21] F. Melgani and L. Bruzzone, "Classification of hyperspectral remote sensing images with support vector machines," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 42, no. 8, pp. 1778–1790, Aug 2004.
[22] P. E. Rauber, A. X. Falcão, T. V. Spina, and P. J. de Rezende, "Interactive segmentation by image foresting transform on superpixel graphs," in *Conference on Graphics, Patterns and Images (SIBGRAPI), 2013 26th SIBGRAPI*, Aug 2013, pp. 131–138.
[23] J. A. dos Santos, P.-H. Gosselin, S. Philipp-Foliguet, R. Torres, and A. X. Falcão, "Multiscale classification of remote sensing images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, no. 10, pp. 3764–3775, Oct 2012.
[24] M. J. Swain and D. H. Ballard, "Color indexing," *International Journal of Computer Vision*, vol. 7, no. 1, pp. 11–32, Nov. 1991.
[25] T. Ojala, M. Pietikinen, and D. Harwood, "A comparative study of texture measures with classification based on featured distributions," *Pattern Recognition*, vol. 29, no. 1, pp. 51 – 59, 1996.
[26] R. O. Stehling, M. A. Nascimento, and A. X. Falcão, "A compact and efficient image retrieval approach based on border/interior pixel classification," in *Proceedings of the Eleventh International Conference on Information and Knowledge Management*, ser. CIKM '02. New York, NY, USA: ACM, 2002, pp. 102–109.
[27] 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 Recogn. Lett.*, vol. 32, no. 1, pp. 27–37, Jan. 2011.
[28] D. Tuia, F. Pacifici, M. Kanevski, and W. Emery, "Classification of very high spatial resolution imagery using mathematical morphology and support vector machines," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 47, no. 11, pp. 3866–3879, Nov 2009.
[29] J. Benediktsson, J. Palmason, and J. Sveinsson, "Classification of hyperspectral data from urban areas based on extended morphological profiles," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, no. 3, pp. 480–491, March 2005.
[30] J. Munoz-Mari, D. Tuia, and G. Camps-Valls, "Semisupervised classification of remote sensing images with active queries," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, no. 10, pp. 3751–3763, Oct 2012.
[31] P. T. M. Saito, P. J. de Rezende, A. X. Falcão, C. T. N. Suzuki, and J. F. Gomes, "An active learning paradigm based on a priori data reduction and organization," *Expert Systems with Applications (ESwA)*, vol. 41, no. 14, pp. 6086–6097, 2014.