

# AUTOMATIC FUSION OF REGION-BASED CLASSIFIERS FOR COFFEE CROP RECOGNITION

*Fabio A. Faria, Jefersson A. dos Santos, Ricardo da S. Torres, Anderson Rocha, Alexandre X. Falcão*

RECOD Lab – Institute of Computing, University of Campinas, Campinas, SP, Brazil  
{ffaria,jsantos,rtorres,anderson.rocha,afalcao}@ic.unicamp.br

## ABSTRACT

Coffee crop recognition in remote sensing images is a complex task. It poses several challenges due to different spectral responses and texture patterns that can be extracted from coffee regions. This paper presents a novel framework for combining different classifiers using support vector machine technique (SVM), which try to learn with each one of classifiers previews experiences (meta-learning). We investigate the combination of seven learning methods and seven image descriptors aiming at creating low-cost classifiers for coffee crops recognition. The objective is to provide an effective mechanism for coffee crop recognition by fusion of region-based classifiers in remote sensing images. The experiments showed that the proposed framework for fusion of classifiers produces better results than the traditional majority voting fusion approach and all base classifiers tested.

**Index Terms**— Support vector machines (SVMs), fusion of classifiers, coffee crops.

## 1. INTRODUCTION

Recognizing coffee crop regions in remote sensing images is not a trivial task. The place or the age of the crop may hinder the recognition process, as different spectral response and texture patterns can be observed for different regions with coffee.

Existing techniques to address this problem can be grouped into two categories: pixel-based [1] and region-based [2]. To the extent that high-resolution remote sensing images became available to the civil community, new approaches of representation and feature extraction have been proposed in order to make a better use of data. Blaschke et al. [3] show that the more high-resolution images were available, the more approaches for RSI representation were proposed. One natural consequence was the development of alternative techniques to the classification based on pixels.

---

This work was supported by CAPES (PDSE 1260-12-0), FAPESP (Grants 2008/58528-2, 2007/52015-0, 2008/57428-4, 2009/18438-7, and 2010/14910-0), and CNPq (grants 306587/2009-2 and 303673/2010-9).

In fact, several methods using region-based analysis have presented better results when compared with traditional methods based on pixels [4, 5].

The performance of coffee crop recognition systems depends on several factors: image descriptors, segmentation methods, machine learning approaches, among others. Many of those factors have been evaluated in several works. For example, Santos et al. [6] evaluated the accuracy performance of several image descriptors in coffee image classification tasks. In [7], Santos et al. addressed the same problem by proposing a genetic programming approach for combining image descriptors. Another venue is concerned with the evaluation/proposal of interactive methods [8] that exploit different machine learning approaches.

This paper follows the steps of those previous studies by providing new insights regarding the combination of image descriptor and learning methods. We investigate the combination of seven learning methods and seven image descriptors aiming at creating low-cost classifiers for recognizing coffee crops. Furthermore, we propose the framework for combining those base classifiers using support vector machine technique (SVM). Performed experiments demonstrate that the proposed framework for fusion of classifiers yields better results than the traditional majority voting fusion approach.

## 2. FRAMEWORK FOR FUSION OF CLASSIFIERS

Let  $\mathcal{L}$  be a set of learning methods (e.g., Decision Tree, Naïve Bayes, and kNN) and  $\mathcal{F}$  be a set of image descriptors (e.g., Color Histogram). Suppose that base classifiers are created by combining each available learning method with each image descriptor. For example, three classifiers could be created by combining the learning methods Decision Tree, Naïve Bayes and kNN with the Color Histogram descriptor. Let  $\mathcal{C}$  be the set of classifiers created by that combination, where  $|\mathcal{C}| = |\mathcal{L}| \times |\mathcal{F}|$ . Let  $\mathcal{S}$  be a set of images, where the class of  $s_i \in \mathcal{S}$  ( $1 < i \leq |\mathcal{S}|$ ) is known. The set  $\mathcal{S}$  is used to construct both the training ( $T$ ) and validation ( $V$ ) sets, where  $T \cup V = \mathcal{S}$  and  $T \cap V = \emptyset$ .

First, all base classifiers  $c_j \in \mathcal{C}$  ( $1 < j \leq |\mathcal{C}|$ ) are trained on set  $T$ . Next, the performance of each classifier on the validation set  $V$  is computed and stored into a matrix  $M_V$ , where  $|M_V| = |V| \times |\mathcal{C}|$ . Given a new image  $I$ , we use each classifiers  $c_j \in \mathcal{C}$  ( $1 < k \leq |\mathcal{C}|$ ) to determine the class to which  $I$  belongs. Those  $j$  outcomes are used as input of a fusion technique (e.g., majority voting, support vector machine – SVM) that takes the final decision regarding the definition of the class of  $I$ . In the case of a fusion technique that requires prior training (e.g., SVM),  $M_V$  is used for that. Figure 1 illustrates the proposed framework for combining classifiers.

### 3. EXPERIMENTS AND RESULTS

This section presents performed experiments to validate the proposed framework.

#### 3.1. Experimental Methodology

##### 3.1.1. Dataset

The used image was captured by the SPOT satellite and corresponds to the Monte Santo de Minas county, in the State of Minas Gerais, Brazil, a traditional place of coffee cultivation. The region where this image was captured is mountainous. Therefore, the spectral patterns tend to be affected by the topographical differences and interference generated by the shadows. Another problem is that coffee is not a seasonal crop. Thus, in the same area, there may be crops of different ages. Concerning classification aspects, we have several completely different patterns representing the same class while some of these patterns are much closer to other classes.

To evaluate the accuracy, we use a ground truth that indicates all coffee region in the image. The subimage used in the experiments is composed by  $1000 \times 1000$  pixels. It was divided into 4, 885 regions by using the method for multi-scale segmentation proposed by Guigues et al. [9]. As the experiments were performed with region level image and the ground truth is in pixel level, it was necessary to define a rule to label each region: if more than 80% a region contains pixels of coffee, that region was labelled as “coffee”; otherwise it is a non-coffee region. The dataset was divided into three parts: 60% of regions were used for training, 20% for building the validation set, and 20% for testing.

##### 3.1.2. Image Descriptor

We have used color and texture descriptors in our experiments. The used color descriptors include ACC, BIC, CCV, and GCH. The used texture descriptors in-

clude QCCH, SID and UNSER. More details about those image descriptors can be found in [10].

##### 3.1.3. Learning Methods and Baselines

We have used seven learning methods in our framework: Naïve Bayes (NB), Decision Tree (DT), Simple Logistic (SL), Naïve Bayes Tree (NBT), and k-Nearest Neighbourhood (kNN) using  $k = 1$ ,  $k = 3$ , and  $k = 5$ . Those methods are simple and fast, being suitable to be combined in a recognition system. We have used the implementation of those learning methods available in the WEKA<sup>1</sup> data mining library. All learning techniques were used with default parameters. We have used as baselines the following approaches: bagging using  $k = 3$  (BAGG-3) and  $k = 5$  (BAGG-5) iterations; support vector machine using polynomial (SVM-PK) and RBF (SVM-RBF) kernels.

##### 3.1.4. Evaluation Measures

In our experiments, we have used evaluation measures from the confusion matrix. The three evaluation measures are: accuracy, kappa [11], and tau [12] indexes. We also assess the results quality for region-based classification at pixel level. That measure computes the percentage of pixels of region correctly classified.

#### 3.2. Results

We report the average results considering a 5-fold cross-validation protocol.

##### 3.2.1. The Best Learning Methods for Each Image Descriptors

Table 1 presents the best results for each image descriptor. As it can be observed, BAGG-5 is the learning method with the best accuracy performance for five descriptors. BAGG-5 is an ensemble of trees with 5 iterations that is more stable and yields better results than the sample decision tree (DT). SL and kNN-5 are the best learning methods for the ACC and BIC descriptors, respectively. In fact, kNN-5 with BIC descriptor (kNN-5-BIC) is the best classifier.

##### 3.2.2. Fusion of Classifiers

Table 2 presents the results obtained for each fusion technique and for the best single classifier (kNN-5), considering four different measures.

In the experiments, we implemented the proposed framework considering two different fusion techniques: SVM with RBF kernel (FSVM-RBF-49) and majority voting (MV-49) both with 49 classifiers

<sup>1</sup><http://www.cs.waikato.ac.nz/~ml/weka> (As of 01/10/2012).

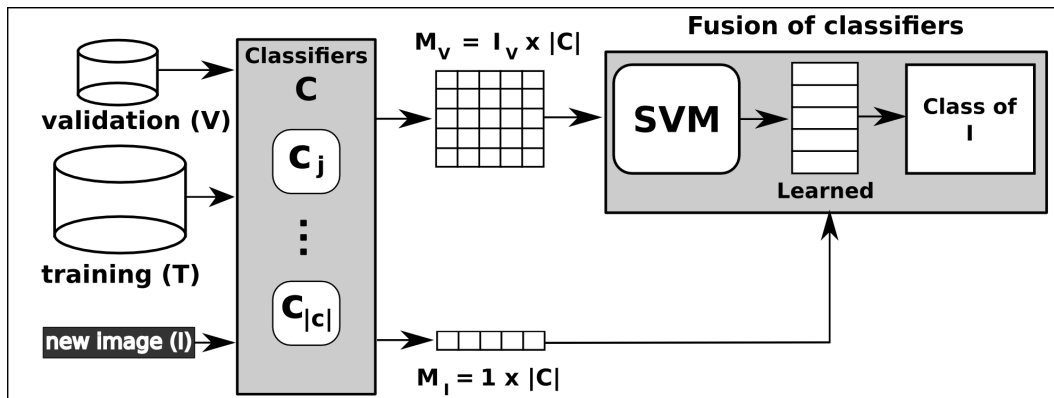


Fig. 1. Framework for fusion of classifiers.

Descriptor	Learning Method	Measures - Region Level		
		Accuracy	Kappa	TAU
ACC	SL	86.06%±1.60	0.47±0.06	0.62±0.03
BIC	kNN-5	87.29%±1.03	0.59±0.03	0.67±0.02
CCV	BAGG-5	86.10%±1.36	0.52±0.04	0.63±0.02
GCH	BAGG-5	86.14%±1.40	0.52±0.04	0.63±0.03
QCCH	BAGG-5	83.95%±1.08	0.45±0.05	0.59±0.03
SID	BAGG-5	79.00%±1.58	0.21±0.06	0.48±0.03
UNSER	BAGG-5	85.12%±1.19	0.49±0.04	0.61±0.02

Table 1. The performance of the best learning method for each image descriptor with their respective standard deviations.

(7 learning methods  $\times$  7 image descriptors). Note that using 49 classifiers means to employ all available classifiers in the fusion process.

In this experiments, FSVM-RBF-49 achieved 89.09% of accuracy against 88.50% and 87.29% achieved by MV-49 and kNN-5-BIC, respectively. FSVM-RBF-49 obtained the best results among all four measures, which suggests that the use of learning techniques for combining different classifiers and descriptors is appropriate.

The superiority of the RBF-FSVM-49 is also observed for the *pixel level* as well. FSVM-RBF-49 has achieved 83.63% of matching, while MV-49 and kNN-5-BIC yield 81.75% and 82.67%, respectively. Note that despite of the best results of MV with regard to three measures (Accuracy, Kappa, and TAU), when we consider the *Pixel Level*, MV gets worse rate than kNN-5-BIC.

Figures 2 (a) and (b) show the original image and the ground truth, respectively. Figures 2 (c–e) illustrate the results for FSVM-RBF-49, MV-49 and kNN-5-BIC classification approaches.

According to the results presented in Figure 2, we can note that the three approaches visually achieves almost the same coffee crops represented in the ground truth. Observe that there are a considerable number of false positives. This is due the presence of a high number of new coffee crops, which are the most difficult

regions to classify in this dataset. The spectral pattern extracted from these targets are quite similar to bare soil and every kind of agricultural regions with small plants.

#### 4. CONCLUSION

This paper presented a framework for fusion of low-cost classifiers using support vector machine (FSVM). We also compared several different learning methods and image descriptors in the region-based RSI classification task. The performed experiments demonstrate that good accuracy performance can be observed for low-cost used classifiers (mainly, BAGG and kNN). The proposed framework overcomes both the best used classifier and the well-known majority voting fusion approach.

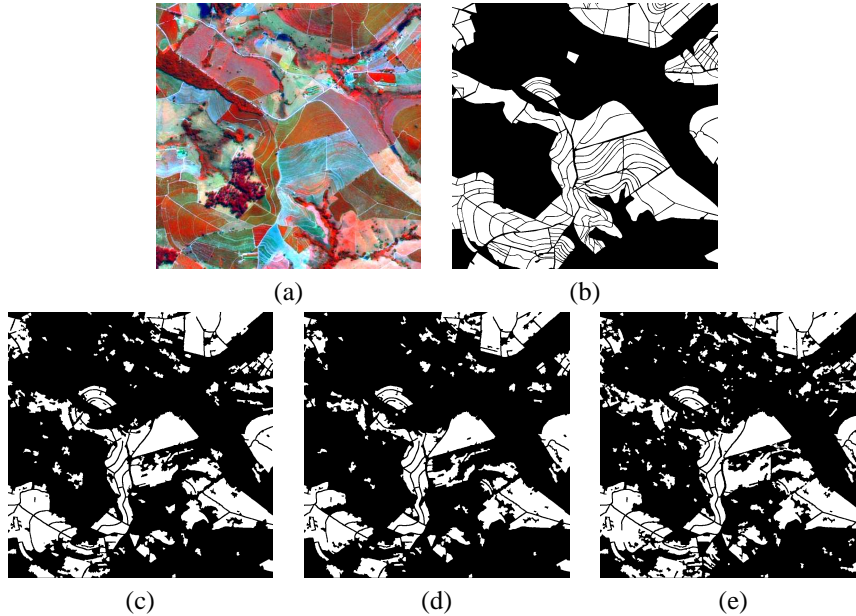
Future work considers the use of diversity measures [13] to select base classifiers to be combined by the proposed fusion approach. We also plan to investigate the use of the proposed framework in other applications.

#### 5. REFERENCES

- [1] R. Pisani, P. Riedel, A. Gomes, R. Mizobe, and J. Papa, "Is it possible to make pixel-based radar image classification user-friendly?," in *IEEE IGARSS*, 2011.

Classifiers	Measures			
	Region Level			Pixel Level
	Accuracy	Kappa	TAU	
FSVM-RBF-49	89.09%±1.09	0.62±0.02	0.70±0.02	83.63%
MV-49	88.50%±1.34	0.59±0.04	0.68±0.03	81.75%
kNN-5-BIC	87.29%±1.03	0.59±0.03	0.67±0.02	82.67%

**Table 2.** Performance results for the best single classifier and for the fusion approaches based on SVM and Majority Voting with their respective standard deviations.



**Fig. 2.** (a) Original image, (b) Ground truth, (c) FSVM-RBF-49, (d) MV-49, and (e) kNN-5-BIC.

- [2] G. Moser and S.B. Serpico, "Multitemporal region-based classification of high-resolution images by markov random fields and multiscale segmentation," in *IEEE IGARSS*, 2011.
- [3] T. Blaschke, "Object based image analysis for remote sensing," *ISPRS Journal of Photogrammetry and Remote Sensing*, 2010.
- [4] Castillejo-González, López-Granados, García-Ferrer, Peña-Barragán, Jurado-Expósito, de la Orden, and González-Audicana, "Object- and pixel-based analysis for mapping crops and their agro-environmental associated measures using quickbird imagery," *Elsevier COMPAG*, 2009.
- [5] Soe W. Myint, Patricia Gober, Anthony Brazel, Susanne Grossman-Clarke, and Qihao Weng, "Per-pixel vs. object-based classification of urban land cover extraction using high spatial resolution imagery," *Remote Sensing of Environment*, 2011.
- [6] 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 *VISAPP*, 2010.
- [7] J.A. dos Santos, F. Faria, R. Calumby, R. da S Torres, and R.A.C. Lamparelli, "A genetic programming approach for coffee crop recognition," in *IEEE IGARSS*, 2010.
- [8] 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," *Inf. Sci.*, 2011.
- [9] L. Guigues, J. Cocquerez, and H. Le Men, "Scale-sets image analysis," *International Journal of Computer Vision*, 2006.
- [10] O. A .B. Penatti, E. Valle, and R. da S. Torres, "Comparative study of global color and texture descriptors for web image retrieval," *Journal of Visual Communication and Image Representation*, 2012.
- [11] R. L. Brennan and D. J. Prediger, "Coefficient kappa: Some uses, misuses, and alternatives," *Educational and psychological measurement*, 1981.
- [12] Z. Ma and R. L. Redmond, "Tau coefficients for accuracy assessment of classification of remote sensing data.," *Photogrammetric Engineering and Remote Sensing*, 1995.
- [13] L. I. Kuncheva and C. J. Whitaker, "Measures of diversity in classifier ensembles and their relationship with the ensemble accuracy," *Machine Learning*, 2003.