A Multiclass Approach for Land-Cover Mapping by Using Multiple Data Sensors

Edemir Ferreira de Andrade Jr.(✉), Arnaldo de Albuquerque Araújo, and Jefersson A. dos Santos

Department of Computer Science, Universidade Federal de Minas Gerais (UFMG), Av. Antônio Carlos, 6627, Pampulha, Belo Horizonte, MG 31270-901, Brazil
{edemirm,arnaldo,jefersson}@dcc.ufmg.br

Abstract. An usual way to acquire information about monitored objects or areas in earth surface is by using remote sensing images. These images can be obtained by different types of sensors (e.g., active and passive) and according to the sensor, distinct properties can be observed from the specified data. Typically, these sensors are specialized to encode one or few properties from the object (e.g. spectral and spatial properties), which makes necessary the use of diverse and different sensors to obtain complementary information. Given the amount of information collected, it is essential to use a suitable technique to combine the different features. In this work, we propose a new late fusion technique, a majority voting scheme, which is able to exploit the diversity of different types of features, extracted from different sensors. The new approach is evaluated in an urban classification scenario, achieving statistically better results in comparison with the proposed baselines.

Keywords: Data fusion · Remote sensing · Late fusion · Land-cover

1 Introduction

Over the years, there has been a growing demand for remotely-sensed data. Specific objects of interest are being monitored with earth observation data, for the most varied applications. Some examples include ecological science [1], hydrological science [2], agriculture [3], military [4], and many other applications. Remote sensing images (RSIs) have been used as a major source of data, particularly with respect to the creation of thematic maps. This process is usually modeled as a supervised classification problem where the system needs to learn the patterns of interest provided by the user and assign a class to the rest of the image regions. In the last few decades, the technological evolution of sensors has provided remote sensing analysis with countless distinct information, e.g., spatial, spectral, temporal, thermal.

Typically, these sensors are designed to be specialists in obtaining one or few properties from the earth surface. Therefore, it is necessary the utilization of diverse and different sensors to gather the most complementary information

A. Pardo and J. Kittler (Eds.): CIARP 2015, LNCS 9423, pp. 59–66, 2015.
DOI: 10.1007/978-3-319-25751-8_8

as possible. In this scenario, it is essential to use a more suitable technique to combine the different features in a effective way. Mura et al. [5] confirmed the benefit of the use of data fusion in the challenges associated with RSI analysis in competitions. They pointed out that it is difficult to conclude which method has the best performance, since it depends on the foundation of the problem and the nature of the data used.

In this work, we propose a new late fusion technique, able to exploit the diversity of these different types of features, extracted from various sensors. Our approach, called *Dynamic Majority Vote*, uses different learning techniques on the extracted features to create base classifiers. Then, it assigns weights to each classifier according to their ability in identifying individual classes. The weights are calculated regarding the confusion matrix of a classifier in a validation set. Our method exploits the specialty of each classifier to solve multiclass problems.

2 Related Work

Despite the recent advances in feature extraction and representation for RSIs, the combination/fusion of these features, especially when they are extracted by different sensors, requires the development of new techniques.

In this context, Li et al. [6] developed a classification technique based on active learning to combine spatial and spectral information. Petitjean et al. [7] proposed an extraction approach to explore the spatiotemporal characteristics for classification in RSIs. Yang et al. [8] presented a system for evaluating the growth of crops using high resolution images from satellites and airplanes.

Ouma et al. [9] and Wang et al. [10] showed approaches that use multi-scale data to identify land use changes. In Ouma et al. [9], the authors presented a multi-scale segmentation technique with a neural network (unsupervised) for analysis of vegetation. Wang et al.[10] in the other hand, proposed an approach to change detection in urban areas. That method is based on the fusion of characteristics from multiple scales through the average pixel of each scale. The result is a new image corresponding to the combination of scales.

More recently, Gharbia et al. [11] made an analysis of fusion techniques images (Intensity-Hue-Saturation (IHS), Brovey Transform (BT), Principle Component Analysis (PCA)) for remote sensing tasks, at pixel level, showing that all techniques have limitations when used individually. They encourage the use of hybrid systems as a solution. Mura et al. [5] analyzed the approaches used in the past nine years of data fusion competition (Data Fusion Contest). The approaches are separated into three main categories: the level of information/pixels, where the data are combined in the way they were extracted; feature level, where the data are extracted and used as entries for a classification model; and the decision level, which uses a combination of different outputs from various sources, to increase the robustness of final decision (using, e.g., a majority vote). After investigated the last challenges, Mura et al. confirmed the benefits of the use of data fusion in the challenges associated with RSI analysis in competitions. In the majority of cases, the frameworks proposed in the literature are projected to deal with a specific scenario or a particular region, using techniques apart of each domain and object, e.g., roofs are checked with shape features, tree and vegetation are discriminated using a vegetation index. However, it is very difficult to conclude what is the best approach, since it depends on the foundation of the problem, the nature of the data used and the source of information utilized.

The proposed method aim at exploiting multi-sensor data in a more general way. We propose a framework based on a supervised learning scheme, dealing with different scenarios, regions and objects, on the creation of thematic maps for the classification task. For that we propose a new approach, at decision level, to handle with an amount of decisions from different classifiers, and combine them for a final decision for each pixel in the thematic map. Contrary to approaches from the literature, our method uses the kappa index [12] as effectiveness measure to compare two classifiers. This fact brings some advantages since kappa index is more robust in dealing with unbalanced training sets.

3 Proposed Method

The proposed method is projected to receive two images from the same place with different domains as input: an image with very high spatial (VHS) resolution and another one with hyperspectral (HS) resolution. Our method is developed for a multiclass mapping scenario. Its main characteristic is to exploit the expertise of each learning approach over each class in order to find the most specialized classifiers. The result of this process is a *dynamic weight matrix*.

Our approach is divided into five main steps: object representation, feature extraction, training, dynamic weight matrix construction, and predicting. Figure 1 illustrates the proposed framework. We detail each step next.

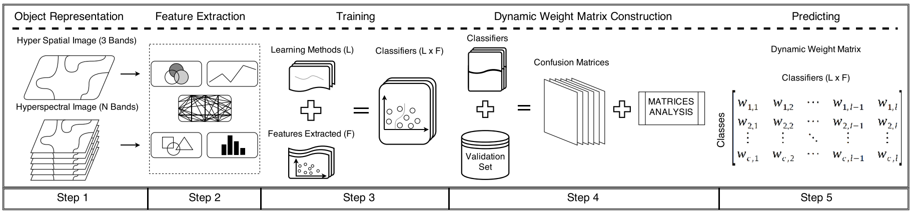
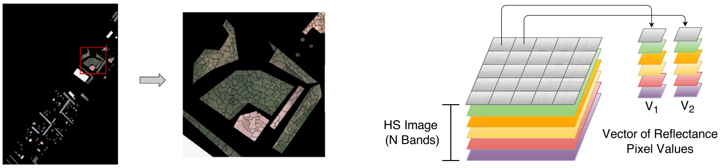


Fig. 1. The Proposed *Dynamic Weight Matrix* (DWM)-based framework

Object Representation. Let Y_R, the set of labels of regions R, be the input dataset, Y_R^t is the training set. In an experimental scenario, $Y_R = Y_R^t \cup Y_R^{t'}$, where $Y_R^{t'}$ is the test set. Let I_{VHS} and I_{HS} be the input images, the first step is to define the objects to be described by the feature extraction algorithms. For the I_{VHS} image, we perform a segmentation process over the regions of Y_R^t in order to split the entire image into more spatially homogeneous objects. It allows the codification of suitable texture features for each part of the image. Due to the low spatial resolution, we consider the pixel as the unique spatial unit for the I_{HS} image.



(a) (b)

Fig. 2. (a) VHS input image segmented. (b) HS input image with reflectance values.

Anyway, we are more interested in exploiting the spectral signature of each pixel. Figure 2 illustrates the object representation phase for each input image.

Feature Extraction. We use the descriptor definition proposed by Torres et al. [13]. Concerning I_{VHS} image, we have used image descriptors based on visible color and texture information to encode complementary features. For the I_{HS} image, we exploit dimensionality reduction/projection properties from the spectral signature in order to obtain diversity.

Training. Let $Y_R^v \subset Y_R^t$ be a validation set split from the training set. We use the features extracted by each descriptors over the remaining training samples and a set of learning methods to create an amount of classifiers (tuples of descriptor/learning method). We use the obtained classifiers to learn the probability distribution of the training set. Notice that training process requires a mapping between spatial and spectral resolutions, using an interpolation method, since I_{VHS} and I_{HS} images are from different domains.

Dynamic Weight Matrix Construction. Algorithm 1 outlines the proposed steps for the construction of the dynamic weight matrix (W_{dyn}). Let $C = \{c_i \in C, 1 < i \leq |C|, i \in \mathbb{N}^*\}$, be a set of trained classifiers c_i over different features from spatial and spectral domains, and evaluated in the Y_R^v. Let $M_C = \{M_i \in M_C, 1 < i \leq |C|, i \in \mathbb{N}^*\}$, the set of confusion matrices M_i computed from c_i, be the input of the algorithm. Let $L = \{l_i \in L, 2 < i \leq |L|, i \in \mathbb{N}^*\}$, be a set of all classes l_i in the problem. For each l_i, the hits at the class l_i (h_{l_i}) are extracted from M_C, and created a list of pairs (h_{l_i}/c_i) sorted by the h_{l_i} (Line 2), and for every pair (h_{l_i}/c_i) a initial weight is assigned in W_{dyn}, regarding with the position j of the pair (h_{l_i}/c_i) in the sorted list (Line 3-4). In Lines 5-6, the column i of M_i is used to compute: (1) the sparsity S, which indicates the degree of importance of c_i at l_i, given by the ratio of the highest miss value at column i (max_{miss}) and the sum of all predicts, (hits and misses); and (2) the uniform misses expected for each class (m_{exp}), given by the the percentage of misses (p_{miss}) uniformly distributed to the other classes. Finally, at Lines 7-11 the weights of W_{dyn} are updated when the kappa index of c_i (κ_{c_i}) is greater than the mean of all classifiers kappa's index ($\bar{\kappa}$). When S is less than m_{exp}, the weight in W_{dyn} for c_i in l_i is increased and decreased otherwise, regarding with the ratio between κ_{c_i} and $\bar{\kappa}$. The reweight in the W_{dyn} aims to explore the specialty of each classifier in every class, given a gain (or penalty) for those classifiers which

Algorithm 1. Construction of the Dynamic Weight Matrix.

Input: Stack of Confusion Matrices (M_C)
Initializing: Set $W_{dyn} \leftarrow 0$, individual κ_{c_i} and mean $\bar{\kappa}$ kappa index.
for each class l_i in L **do**
    Creating of a sorted list of pairs (h_{l_i}/c_i)
    for every pair (h_{l_i}/c_i) at position j **do**
        Initial weight in $W_{dyn} \leftarrow (j + 1.0)/|L|$
        Compute the sparsity $S \leftarrow max_{miss}/(hits + misses)$
        Compute $m_{exp} \leftarrow p_{miss}/(|L| - 1)$
        if $\kappa_{c_i} > \bar{\kappa}$ **then**
            if $S < 2 * m_{exp}$ **then**
                Gain at $W_{dyn} \leftarrow (\kappa_{c_i}/\bar{\kappa})^* W_{dyn}$
            else
                Penalty at $W_{dyn} \leftarrow (\bar{\kappa}/\kappa_{c_i})^* W_{dyn}$
    end for
end for

show a sparsity (or density) in the predicts by class at the confusion matrix, regarding to the validation set.

Predicting. Once the W_{dyn} is built, the same method of segmentation is used in $Y_R^{t'}$, and the segmented objects are labeled by the classifiers as regions (spatial tuple) or pixel by pixel (spectral tuple) creating a thematic map for each classifier. Once more, since the thematic maps from the I_{HS} image have a different resolution, we apply the same interpolation method as in training phase, to map its outcomes to the spatial resolution domain. Finally, the thematic maps are used as input of the dynamic majority vote technique. We used the weight of each classifier in their respective predicted classes for each labeled pixel in thematic maps with the dynamic weight matrix previously built, and taking the final decision according to the highest final weight class for that pixel in specific. An example of how to use the W_{dyn} is showed in Figure 3.

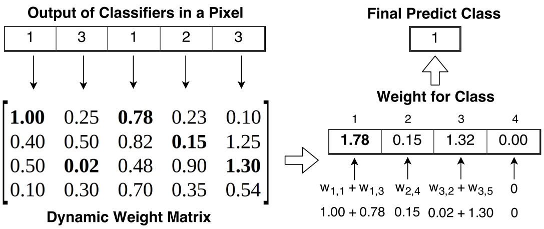


Fig. 3. Given the output of the classifiers in a pixel, the relevance of each prediction is given by the dynamic weight matrix, and chosen the class with the highest final weight.

4 Experiments

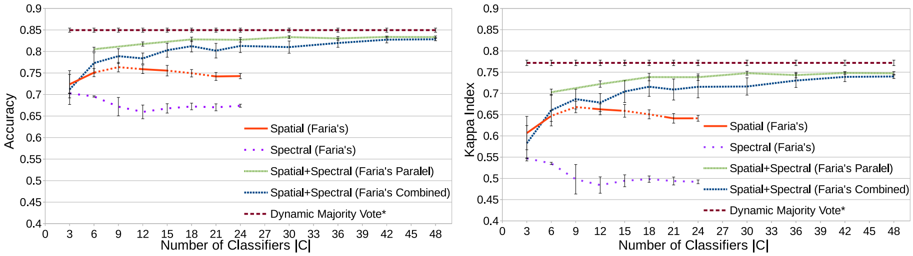
4.1 Setup

Dataset. We have used the grss_dfc_2014 [14]. It is an urban classification scenario with two sensors information: (a) Very High Spatial (VHS) resolution; and (2) Hyper Spectral (HS) resolution. **Measures.** We used Overall Accuracy and Kappa index. For the statistical test of significance, we used paired Student t-test (confidence of 95%), with 10 samples for each experiment, since in a statistical way that will still provide the desired confidence for our experiments. **Segmentation.** We used the IFT-Watershed [15] with spatial radius 5 and volume threshold equal to 100. **Feature Extraction.** We used four image descriptors to encode spatial information [16]: BIC, CCV, GCH, and Unser. To extract spectral information, we have used four different approaches: (1) the raw data of HS image (84 Bands), (3) the Fisher Linear Discriminant (FLD) [17] components, (3) the first 3 principal components of PCA [18], and (4) the first 4 PCA components. **Training.** A validation set is split from training set and trained in a Stratified ShuffleSplit cross validation scheme, using a group of 6 weak learners: Gaussian Naive Bayes, k-Nearest Neighbors (3, 5 and 10-Nearest Neighbors), Decision Tree, and a Support Vector Machine with linear kernel, using the features extracted by each descriptors, resulting in the total of 48 classifiers (24 from each domain). We have used the implementation of those learning methods available in the Scikit-Learn Python library All learning methods were used with default parameters which means we did not optimize them whatsoever. The management of HS data is made using the Spectral Python (SPy) Library, including the extraction of features from spectral domain. We used the Nearest-neighbor interpolation to the mapping in training and predicting phases. **Baselines.** We have implemented a diversity-based fusion framework as proposed by Faria et al. [19], varying the number of classifiers selected, and using the majority vote at the meta learning phase. We setup the framework with 4 different ways: using only the spatial and spectral images, the spatial and spectral images in parallel and fusioning the results, and combining the spatial and spectral domains at the construction of the validation matrix. Refer to [19] for further details about the framework.

4.2 Results and Discussion

The results obtained by the proposed method (*Dynamic Majority Vote*) against the baselines, with the confidence intervals (95%), are presented in Figure 4.

The comparison shows a statistical significant difference among our approach and the baseline proposed at the confidence of 95%, regarding with the t-student test. Since our method is based on the simple majority vote, a special case where all weights in the dynamic weight matrix are equal to 1, was already expected the outperforms results. Our approach has the ability to handle with the issue of instead give to a classifier a fix weight (when used a weight majority vote), assign to each classifier a separated weight for each class predicted. In this way



(a) (b)

Fig. 4. The results of the proposed approach in comparison with the baselines based on [19], regarding of the accuracy measure in (a) and kappa index in (b).

we exploit these classifiers who have a specialty in some specifics classes, but would be suppressed by the others classifiers in an equal weight scheme.

Another good point in our approach, is the capacity to deal with information from different domains, without being unfair in the weighting, just because the initial weight is assigned without seeing the general performance of the classifier, but only at the specific class. As a drawback, our method do is not handle with the binary classification problem (since the analysis of the sparsity of a classifier with two classes does not make sense), and the initial weight might not be enough to deal with an amount of bad classifiers.

5 Conclusion

In this paper, we proposed a framework called *Dynamic Majority Vote* for remote sensing image classification with data from multiple sensors. Our approach extracts features from different domains, which are trained with different learning techniques. This process creates a set of classifiers with different expertise. Our method assigns a weigh for each classifier according to their expertise in each specific class. The creation of the final thematic maps consists in classifying each non-labeled region by fusioning the predicted output of each classifier according to their weights. We conducted a series of experiments in the grss_dfc_2014 [14] dataset (IEEE GRSS Data Fusion Contest 2014) that demonstrate a significant improvement in comparison with the proposed baselines. For the future work, we intend to extend this framework exploring the use of more descriptors, classifiers, and other late fusions methods. We also plan to test our method with other real scenarios, such as agriculture and environmental monitoring.

Acknowledgments. This work was partially financed by CNPq, CAPES, Fapemig. The authors would like to thank Telops Inc. (Québec, Canada) for acquiring and providing the data used in this study, the IEEE GRSS Image Analysis and Data Fusion Technical Committee and Dr. Michal Shimoni (Signal and Image Centre, Royal Military Academy, Belgium) for organizing the 2014 Data Fusion Contest, the Centre de

Recherche Public Gabriel Lippmann (CRPGL, Luxembourg) and Dr. Martin Schlerf (CRPGL) for their contribution of the Hyper-Cam LWIR sensor, and Dr. Michaela De Martino (University of Genoa, Italy) for her contribution to data preparation.

References

1. Ghiyamat, A., Shafri, H.Z.: A review on hyperspectral remote sensing for homogeneous and heterogeneous forest biodiversity assessment. IJRS **31**(7), 1837–1856 (2010)
2. Schmid, T., Koch, M., Gumuzzio, J.: Multisensor approach to determine changes of wetland characteristics in semiarid environments (central spain). IEEE Transactions on Geoscience and Remote Sensing **43**(11), 2516–2525 (2005)
3. Lanthier, Y., Bannari, A., Haboudane, D., Miller, J.R., Tremblay, N.: Hyperspectral data segmentation and classification in precision agriculture: a multi-scale analysis. In: IGARSS 2008, vol. 2, p. II-585. IEEE (2008)
4. Briottet, X., Boucher, Y., Dimmeler, A., et al.: Military applications of hyperspectral imagery. In: DSS, ISPIE, p. 62390B (2006)
5. Mura, M.D., Prasad, S., Pacifici, F., Gamba, P., Chanussot, J.: Challenges and opportunities of multimodality and data fusion in remote sensing. In: 2013 Proceedings of the 22nd European, EUSIPCO, pp. 106–110. IEEE (2014)
6. Li, J., et al.: Spectral-spatial classification of hyperspectral data using loopy belief propagation and active learning. TGRS on **51**(2), 844–856 (2013)
7. Petitjean, F., Kurtz, C., Passat, N., Gançarski, P.: Spatio-temporal reasoning for the classification of satellite image time series. PRL **33**(13), 1805–1815 (2012)
8. Yang, C., Everitt, J.H., Du, Q., Luo, B., Chanussot, J.: Using high-resolution airborne and satellite imagery to assess crop growth and yield variability for precision agriculture. Proceedings of the IEEE **101**(3), 582–592 (2013)
9. Ouma, Y.O., Josaphat, S., Tateishi, R.: 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 & Geosciences **34**(7), 715–737 (2008)
10. Wang, W.J., Zhao, Z.M., Zhu, H.Q.: Object-oriented change detection method based on multi-scale and multi-feature fusion. In: JURSE, pp. 1–5. IEEE (2009)
11. Gharbia, R., Azar, A.T., Baz, A.E., Hassanien, A.E.: Image fusion techniques in remote sensing. arXiv preprint arXiv:1403.5473 (2014)
12. Congalton, R.G., Green, K.: Assessing the accuracy of remotely sensed data: principles and practices. CRC Press (2008)
13. da Silva Torres, R., Falcão, A.X.: Content-based image retrieval: Theory and applications. RITA **13**(2), 161–185 (2006)
14. 2014 IEEE GRSS Data Fusion Contest. http://www.grss-ieee.org/community/technical-committees/data-fusion/
15. Lotufo, R., Falcao, A., Zampirolli, F.: Ift-watershed from gray-scale marker. In: SIBGRAPI, pp. 146–152 (2002)
16. dos Santos, J.A., Penatti, O.A.B., da Silva Torres, R.: Evaluating the potential of texture and color descriptors for remote sensing image retrieval and classification. In: VISAPP, Angers, France, May 2010
17. Fisher, R.A.: The use of multiple measurements in taxonomic problems. Annals of Eugenics **7**(2), 179–188 (1936)
18. Jolliffe, I.: Principal component analysis. Wiley Online Library (2002)
19. Faria, F.A., Dos Santos, J.A., Rocha, A., Torres, R.D.S.: A framework for selection and fusion of pattern classifiers in multimedia recognition. PRL **39**, 52–64 (2014)