FUSION OF GENETIC-PROGRAMMING-BASED INDICES IN HYPERSPECTRAL IMAGE CLASSIFICATION TASKS

Juan F. Hernández Albarracín[1], *Edemir Ferreira Jr*[2], *Jefersson A. dos Santos*[2], *and Ricardo da S. Torres*[1]

[1]Institute of Computing, University of Campinas, Campinas, SP 13083–852
Emails: juan.albarracin@ic.unicamp.br, rtorres@ic.unicamp.br
[2]Department of Computer Science, Universidade Federal de Minas Gerais, Belo Horizonte, MG
Emails: edemirm@dcc.ufmg.br, jefersson@dcc.ufmg.br

ABSTRACT

This paper introduces a two-step hyper- and multi-spectral image classification approach. The first step relies on the use of a genetic programming (GP) framework to both select and combine appropriate bands. The second step is concerned with the image classification itself. We present two strategies for multi-class classification problems based on the combination of GP-based indices defined in binary classification scenarios. Performed experiments involving well-known and widely-used datasets demonstrate that the proposed approach yields comparable or better effectiveness performance when compared to several traditional baselines.

***Index Terms*—** hyper- and multi-spectral images, genetic programming, fusion of classifiers.

1. INTRODUCTION

The use of hyper- and multi-spectral remote sensing image classification systems plays an important role for land cover change understanding. Typical applications range from urban planning activities to complex ecological global-scale change analysis. Usually, the creation of effective classification systems relies on three main steps: (i) selection of appropriate bands, i.e., those that best encode features of objects of interest; (ii) combination of selected bands in order to define useful indices (e.g., NDVI, EVI, and ExG); and (iii) use of machine learning apparatuses for classifying regions/pixels based on their indices.

In this paper, we propose a two-step approach for hyper- and multi-spectral image pixel classification. In the first step, we take advantage of a genetic-programming approach [1] for both the selection of appropriate bands and their combination with the objective of better discriminating pixel values observed for different classes. In a second step, we combine indices defined by the genetic programming framework in a classifier-fusion system which is able to assign pixels to their classes very effectively.

Many successful approaches have been proposed aiming at supporting both band selection and/or combination. Most of them, however, do not address both problems at the same time. Some examples include methods targeting the identification of the most informative bands [2–5], clustering techniques to find the most representative bands based on constructed subsets [6–10], rank-based methods [11–13], and search strategies [14–16]. In another research venue, ensembles of classifiers have also been developed to support hyper-spectral image understanding in different contexts such as band selection [17] and classification tasks [18, 19]. None of them, however, has considered the use of GP-generated indices.

2. MULTI-CLASS CLASSIFICATION SCHEMES

Figures 1 and 2 present the two classification systems considered in our study, which are introduced in Sec. 2.2 and 2.3, respectively. Both depend on indices defined by a GP-based framework proposed recently (Sec. 2.1).

2.1. Learning GP-based indices

We use the GP framework introduced in [1] to select and combine bands. The GP framework basically evolves formulas that arithmetically combine image bands. Each formula is seen as an individual of a population which evolves along generations subject to genetic operators (e.g., cross-over, mutation, and reproduction). Trees are the typical representation used for individuals, such that leaves encode band values, while internal nodes are used to encode math operators. A formula (i.e., an individual) is considered a good one depending on the inter- and intra-class distance of the resulting pixels (*fitness* function). Pixels belonging to the same class should be clustered together in a compact way, being far away from pixels of different classes. The GP learning process is based on determining the best formula considering a binary classi-

Thanks to CAPES, CNPq (grant 449638/2014-6), FAPEMIG (APQ-00768-14), Microsoft Azure for Research, FAPESP (grant #2014/12236-1) and FAPESP-Microsoft Virtual Institute (grants #2014/50715-9, #2013/50155-0, and #2013/50169-1) for funding.

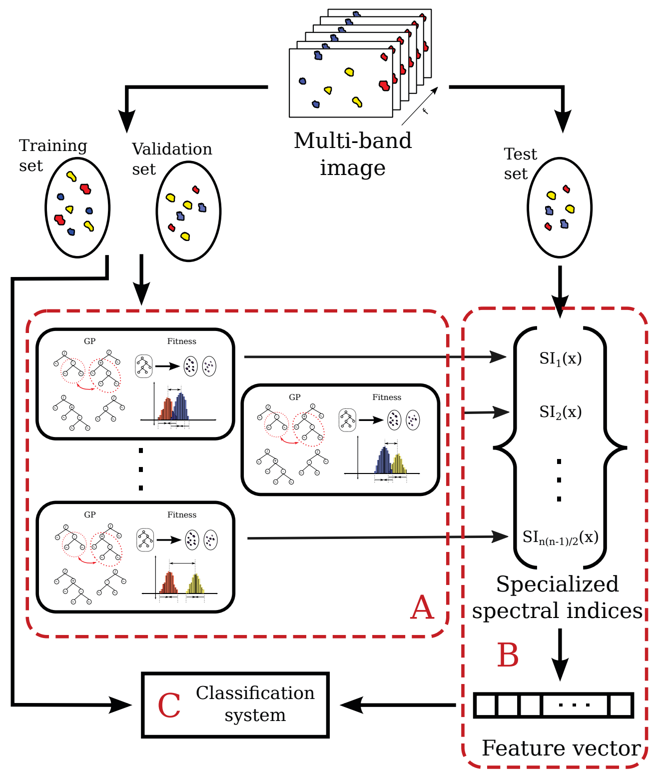


Fig. 1. Vector-based Fusion. The learned indices from training and validation data map samples into a new feature vector that will be used in another classification system.

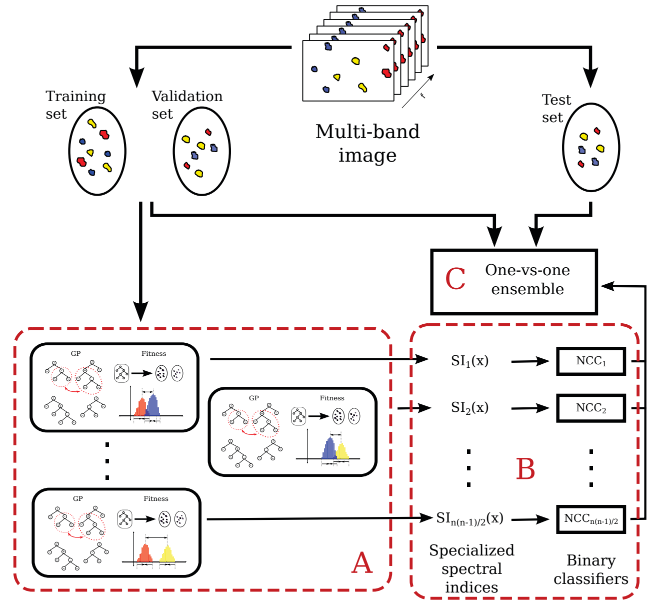


Fig. 2. One-vs-One Classifier Fusion. A NCC classifier will be built for each learned index and later used in an OvO ensemble strategy.

fication setting. This step of our classification approaches is labeled as A in Figs. 1 and 2. See [1] for more details.

2.2. Vector-based Fusion (GP-VBF)

For each pair of (n) classes, a specialized spectral index is learned, aiming at maximizing the separation between the training samples of both classes. A total of $\frac{n(n-1)}{2}$ indices learned will be used to map each multi-band pixel into a feature space of the same number of dimensions (Step B in Figure 1). The resulting feature vectors are then used in a classification system (Step C). In our work, we evaluate the use of the Random Forest classifier (RF) in this last stage.

2.3. One-vs-One Classifier Fusion (GP-OVO)

The $\frac{n(n-1)}{2}$ learned indices will be used as separate feature spaces where the same number of Nearest Centroid Classifier algorithms (NCC) will be trained (Step B in Figure 2), then they will compose a one-vs-one ensemble method for classification (Step C).

3. EXPERIMENTAL PROTOCOL

3.1. Datasets

We considered two scenes – see Fig. 3-(a) and Fig. 4-(a):
Salinas: Collected over Salinas Valley, California, by the AVIRIS sensor in 224 bands. 512×217 image, 3.7 meters/pixel,[1] and wavelengths between 400 and 2500 nanometers, containing 16 classes, as shown in Table 1.
Thetford Mines: Collected over an urban area near Québec, Canada by a 84-channel sensor. 795×564 image, about 1 meter/pixel, and wavelengths between 7.8 and 11.5 micrometers. Dataset of the 2014 IEEE GRSS Data Fusion Contest,[2] containing seven classes, as shown in Table 2.

3.2. GP Parameters

Some parameters are the normally recommended in general evolutionary algorithms, such as a crossover rate of 0.9 and a tournament selection method of 3 individuals. The mutation rate was set to 0.2 to avoid local maxima, and an elitism of 10 individuals was considered. The population size was set to 200 and the number of generations, to 300. The stop condition is the maximum number of generations reached. The

[1]Data retrieved from `http://www.ehu.eus/ccwintco/index.php?title=Hyperspectral_Remote_Sensing_Scenes`. Last accessed on Jan, 2017.

[2]`http://www.grss-ieee.org/community/technical-committees/data-fusion/2014-ieee-grss-data-fusion-contest/`. Last accessed on Jan, 2017.

Table 1. Salinas.

Class	# examples
Brocoli_green_weeds_1	2009
Brocoli_green_weeds_2	3726
Fallow	1976
Fallow_rough_plow	1394
Fallow_smooth	2678
Stubble	3959
Celery	3579
Grapes_untrained	11271
Soil_vinyard_develop	6203
Corn_senesced_green_weeds	3278
Lettuce_romaine_4wk	1068
Lettuce_romaine_5wk	1927
Lettuce_romaine_6wk	916
Lettuce_romaine_7wk	1070
Vinyard_untrained	7268
Vinyard_vertical_trellis	1807

Table 2. Thetford Mines.

Class	# examples
road	4293
trees	1027
red roof	1739
grey roof	1973
concrete roof	3797
vegetation	7167
bare soil	1711

arithmetic operations encoded in the internal nodes are addition, subtraction, multiplication, protected division, protected square root, and protected natural logarithm as suggested by Koza [20], to avoid division by zero or imaginary roots.

3.3. Evaluation protocol

The 5-fold cross validation approach was used. Three folds were used for training, one for validation, and one for testing. Five runs were executed in such a way that folds were shifted to be used for training, validation, and testing.

3.4. Baselines

Our method was compared to traditional dimensionality reduction techniques such as Uni-variate feature selection (UFS), Principal component analysis (PCA), Linear discriminant analysis (LDA) and selection with Random forest (RFS), as well as no selection (NS). The classifiers used for those methods were Nearest centroid classifier (NCC) and Random forest (RF).

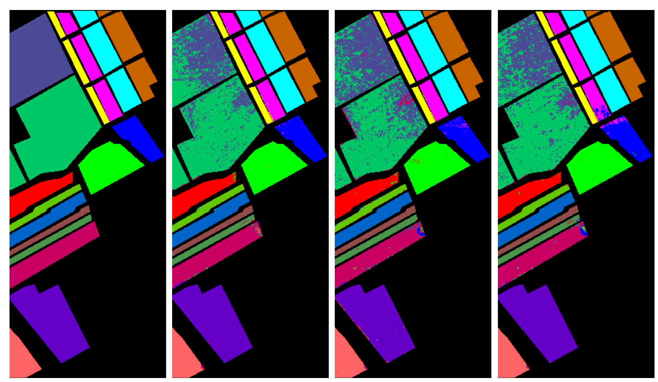
4. RESULTS

Table 3 shows the mean of the normalized accuracies obtained for each fold in the cross validation approach comparing the baselines and the proposed methods. As we can observe, the proposed fusion approaches are very effective in both collections, yielding better or comparable results with respect to the best baselines (confirmed with paired Wilcoxon statistical test, with significance level of 0.05). A ▲ symbol indicates that the fusion approach was statistically superior to the alternative and • means that both methods are statistically tied. The first and the second symbol refer to the comparison with the GP-OVO+NCC and GP-VBF+RF methods, respectively.

Figures 3 and 4 show the thematic maps of the ground truth and the classification results of the best baselines and the proposed fusion schemes for the Salinas and Thetford Mines datasets, respectively. We can also observe that the results related to the proposed methods are quite similar to the ones of the best baselines.

Table 3. Results (% normalized accuracy)

	Salinas		**T. Mines**	
Method	μ	σ	μ	σ
NS + NCC	78.86▲▲	3.63	39.17▲•	6.68
NS + RF	93.70••	3.77	49.95••	9.76
UFS + NCC	72.35▲▲	6.35	33.39▲▲	2.74
UFS + RF	89.11▲•	6.89	36.49▲▲	4.40
PCA + NCC	78.06▲▲	3.78	39.11▲•	6.66
PCA + RF	90.55▲•	4.78	53.22▲•	10.24
LDA + NCC	93.77▲•	3.65	54.14••	10.74
LDA + RF	95.56••	1.86	53.93••	8.83
RFS + NCC	79.38▲▲	4.08	36.50▲•	5.29
RFS + RF	93.59••	3.99	49.21••	9.35
GP-OVO + NCC	94.46	2.23	54.08	10.40
GP-VBF + RF	95.09	3.07	54.10	9.63

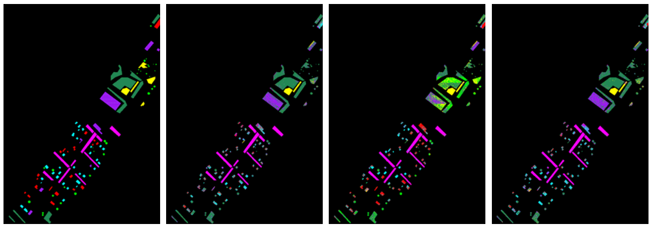


(a) Ground truth (b) LDA + RF (c) GP-OVO (d) GP-VBF

Fig. 3. Thematic map of the ground truth and classification of the two proposed methods and the best baseline for the Salinas dataset.

5. CONCLUSIONS AND FUTURE WORK

This paper has introduced novel approaches for combining Genetic-Programming-based indices in multiclass hyperspectral image classification problems. A genetic programming framework is employed to determine both the most suitable bands and their combination in binary problems. Later, these indices/classifiers are used in two different fusion methods. Performed experiments using well-known datasets demonstrate that the proposed methods are effective, yielding better or comparable results with respect to traditional



(a) Ground truth (b) LDA + NCC (c) GP-OVO (d) GP-VBF

Fig. 4. Thematic map of the ground truth and classification of the two proposed methods and the best baseline for the Thetford Mines dataset.

methods in the literature. Future work will be concerned with the use of GP indices in other land-cover change studies. We also plan to investigate the use of recently proposed fusion approaches [18, 19, 21] to combine GP indices.

6. REFERENCES

[1] J. Hernández, J. A. dos Santos, and R. Torres, "Learning to combine spectral indices with genetic programming," in *29th Conference on Graphics, Patterns and Images (SIBGRAPI)*, 2016.

[2] R. Huang and M. He, "Band selection based on feature weighting for classification of hyperspectral data," *IEEE Geosci. Remote Sens. Lett.*, vol. 2, pp. 156–159, Apr. 2005.

[3] H. Yang, Q. Du, H. Su, and Y. Sheng, "An efficient method for supervised hyperspectral band selection," *IEEE Geosci. Remote Sens. Lett.*, vol. 8, pp. 138–142, Jan. 2011.

[4] R.Y.M. Nakamura, L.M. Garcia Fonseca, J. A. dos Santos, R. Torres, X. Yang, and J. P. Papa, "Nature-inspired framework for hyperspectral band selection," *IEEE TGRS*, vol. 52, pp. 2126–2137, Apr. 2014.

[5] S. Patra, P. Modi, and L. Bruzzone, "Hyperspectral band selection based on rough set," *IEEE TGRS*, vol. 53, no. 10, pp. 5495–5503, 2015.

[6] S. Jia, Y. Qian, and L. Shen, "Unsupervised band selection for hyperspectral imagery classification without manual band removal," *IEEE J. Sel. Topics Appl. Earth Observations Remote Sens.*, vol. 5, no. 2, pp. 531–543, Apr. 2012.

[7] R. Hedjam and M. Cheriet, "Hyperspectral band selection based on graph clustering," in *ISSPA*, July 2012, pp. 813 –817.

[8] J. C. Wu and G. C. Tsuei, "Unsupervised cluster-based band selection for hyperspectral image classification," in *ICACSEI*, July 2013.

[9] L. C. B. dos Santos, S. J. F. Guimaraes, and J. A. dos Santos, "Efficient unsupervised band selection through spectral rhythms," *IEEE Journal of Sel. Top. in Sig. Proc.*, vol. 9, no. 6, pp. 1016–1025, Sept 2015.

[10] G. Zhu, Y. Huang, J. Lei, Z. Bi, and F. Xu, "Unsupervised hyperspectral band selection by dominant set extraction," *IEEE TGRS*, vol. 54, no. 1, pp. 227–239, 2016.

[11] P. Bajcsy and P. Groves, "Methodology for hyperspectral band selection," *Photogrammetric Engineering and Remote Sensing Journal*, vol. 70, pp. 793–802, 2004.

[12] Q. Du and H. Yang, "Similarity-based unsupervised band selection for hyperspectral image analysis," *IEEE Geosci. Remote Sens. Lett.*, vol. 5, no. 4, pp. 564–568, Oct. 2008.

[13] S. Jia, G. Tang, J. Zhu, and Q. Li, "A novel ranking-based clustering approach for hyperspectral band selection," *IEEE TGRS*, vol. 54, no. 1, pp. 88–102, 2016.

[14] J. Kittler, "Feature set search algorithm," in *Pattern Recognition and Signal Processing*, 1978, pp. 41–60.

[15] X. Geng, K. Sun, L. Ji, and Y. Zhao, "A fast volume-gradient-based band selection method for hyperspectral image," *IEEE TGRS*, vol. 52, no. 11, pp. 7111–7119, Nov 2014.

[16] Y. Yuan, G. Zhu, and Q. Wang, "Hyperspectral band selection by multitask sparsity pursuit," *IEEE TGRS*, vol. 53, no. 2, pp. 631–644, Feb 2015.

[17] S. Li, Z. Zheng, Y. Wang, C. Chang, and Y. Yu, "A new hyperspectral band selection and classification framework based on combining multiple classifiers," *PRL*, vol. 83, Part 2, pp. 152 – 159, 2016.

[18] N. Alajlan, Y. Bazi, F. Melgani, and R. R. Yager, "Fusion of supervised and unsupervised learning for improved classification of hyperspectral images," *Inf. Scie.*, vol. 217, pp. 39 – 55, 2012.

[19] J. Xia, J. Chanussot, P. Du, and X. He, "Rotation-based support vector machine ensemble in classification of hyperspectral data with limited training samples," *IEEE TGRS*, vol. 54, no. 3, pp. 1519–1531, March 2016.

[20] J. R. Koza, *Genetic Programming: On the Programming of Computers by Means of Natural Selection*, MIT Press, Cambridge, MA, USA, 1992.

[21] F. A. Faria, J. A. dos Santos, A. Rocha, and R. Torres, "A framework for selection and fusion of pattern classifiers in multimedia recognition," *PRL*, vol. 39, no. 0, pp. 52 – 64, 2014.