RANDOM ENSEMBLE FEATURE SELECTION FOR LAND COVER MAPPING

[§]Anthony Gidudu, [§]Abe T. Bolanle, [§]Tshilidzi Marwala

[§]School of Electrical and Information Engineering, University of the Witwatersrand, Private Bag X3, 2050 Wits, Johannesburg, South Africa

Anthony.Gidudu@wits.ac.za, Bolanle.Abe@students.wits.ac.za, Tshilidzi.Marwala@wits.ac.za

1. INTRODUCTION

The advent of ever improving spectral resolution in satellite data has led to the interest in ensemble classification systems for land cover mapping. Whereas previously research focused on the development and adoption of modern classification techniques, there has of recent been an interest in ensemble approaches whereby the final output is based on a 'consensus' of a group of classifiers. Ensemble classification systems are premised on the need for diversity between the constituent base classifiers as well as determining the 'consensus' rules by which the base classifier outputs are to be combined [1], [2]. Ensemble feature selection is a unique approach to ensemble classification whereby diversity is imposed by using base classifiers with different band combinations [1]. Traditionally, diversity in ensemble systems is ensured using different classifiers or constituting base classifiers by training a classifier with different training data e.g. in boosting and bagging [1]. Varying the feature subsets used to generate the ensemble classifier similarly ensures diversity since the base classifiers tend to err in different subspaces of the feature space [3]. Of the available 'consensus' rules, majority voting was used in this research in all classification outputs, principally due to its simplicity [4]. In random ensemble feature selection, the base classifiers are constituted by randomly selecting the bands [5]. Random ensemble feature selection has been seen to yield comparable results with standard feature selection techniques such as genetic algorithms and exhaustive search methods [1], [5]. This research seeks to further the understanding of random ensemble feature selection by investigating if and/or how the size of ensemble influences the consequent classification accuracies.

2. METHODOLOGY

The online hyperspectral Indiana Pines dataset sourced from the AVIRIS sensor was used in this research. It is a freely accessible online dataset which comes with accompaigning ground truth data. 180 of the 224 bands were used, 4 were discarded because they contained zeros and the remaining 20 bands were left out due to atmospheric distortion [6]. The classes in this dataset included; alfalfa, corn-notill, corn-minimum till, corn, grass/pasture, grass/trees, grass/pasture-mowed, hay-windrowed, oats, soybeans-notill, soybeans-minimum till, soybean-clean, wheat, woods, building-grass-tree-drives, stone-steel towers. Training data for the mentioned classes were selected in reference to the provided ground truth data.

Image classification was effected through the use of Support Vector Machines (SVMs). SVMs are a supervised classification technique that has its roots in Statistical Learning Theory. They basically operate by placing a linear discriminant between two given classes of interest. Remote sensing data is rarely linearly separable and SVMs counter this by nonlinearly projecting the data to a very high dimension space where the previously nonlinearly separable data becomes separable. Placing a linear discriminant in this very high dimension has the same effect of placing a nonlinear discriminant in the previous space. Polynomial, Gaussian and Sigmoid functions may be used to project the data to the high dimension space. In this research, Gaussian SVMs were used. SVMs are primarily binary classifiers, however there exist techniques to adopt them to the multiple classes encountered in remote sensing. In this research, the one against one approach was adopted. A more detailed treatise of SVMs can be found in [7].

Two approaches were used to investigate the effect of ensemble size on classification accuracy. In the first case classification accuracy was related to increasing numbers of features per base classifier and in the second case accuracy was related to increasing number of base classifiers per ensemble. A total of seven ensembles was constituted with increasing numbers of features per base classifier. The first ensemble had two features per base classifier while the last one had 14 features per base classifier. Similarly, for each ensemble, base classifiers were cumulatively added and the ensemble classification consequently derived.

3. DISCUSSION OF RESULTS

The results showed that in general, there was a significant improvement in classification accuracy as the number of bands per base classifier increased. This could be due to the fact that given the high number of classes, 16 classes in this case, more features were needed to appropriately separate the classes. On the other hand, whereas there were few instances where incrementally adding base classifiers to the ensemble significantly improved the classification accuracy, the general trend was that there was no improvement in accuracy as the number of base classifiers increased. These results principally inform us that in the design of ensemble feature selection classification systems, increasing the number of base classifiers may not necessarily translate into improved ensemble classification accuracies. In which case, the minimum possible number of base classifiers will suffice. In conclusion, size in ensemble systems does matter but only if the number of bands per base classifier is increased. Increasing the number of base classifiers evidently doesn't improve classification accuracy.

4. REFERENCES

[1] Chan, J., C., and Paelinckx, D. Evaluation of Random Forest and Adaboost tree-based ensemble classification and spectral band selection for ecotope mapping using airborne hyperspectral imagery. *Remote Sensing of Environment,* 112, pp 2999 – 3011. 2008

[2] Oza, C. N. and Tumer, K. Classifier ensemble: Select real – world applications. *Information Fusion*, vol. 9, pp 4 – 20. 2008

[3] Polikar, R. Ensemble based systems in decision making. *IEEE Circuits and Systems Magazine*, pp 21 – 44. 2006

[4] Foody, G.M., Boyd, D.S. and Sanchez-Hernandez, C. Mapping a specific class with an ensemble of classifiers. *International Journal of Remote Sensing*, 28(8), pp 1733 – 1746. 2007

[5] Mahesh Pal. Ensemble learning with decision tree for remote sensing classification. *Proceedings of World Academy of Science, Engineering and Technology*, Vol 26, ISSN 1307-6884. Dec 2007

[6] Bazi, Y., and Melgani, F. Toward an optimal SVM classification system for hyperspectral remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing,* vol. 44, Issue 11, pp 3374 – 3385. 2006

[7] Christianini, N., and Shawe-Taylor, J. *An introduction to support vector machines: and other kernel-based learning methods*, (Cambridge and New York: Cambridge University Press). 2000