Next Generation Class

Introduction

Remote sensing as a field of study has reached its adulthood; computer-assisted classifiers have been in development for more than two decades. The complexity of remote sensing classification has led to a variety of methods, some of them based on artificial intelligence (AI), and provides motivation for this special issue. AI techniques range from simple out-of-the-box implementations to algorithms tailored to the specifics of remote sensing classification.

Recently, we have also observed a significant increase in parallel processing capabilities. Computer workstations with multiple processors are becoming the mainstream in research laboratories. Physical limitations in processor design indicate that future computational power improvements will result from parallel processing rather than single processor advances. Parallel computing presents a unique opportunity and challenge for image classifiers. The question arises: how can we harvest this new power to improve classification results?

One solution is to implement hybrid classifiers, i.e. methods that merge multiple approaches together. In the field of machine learning several works have shown the potential of a hybrid approach (Hansen and Salamon, 1990; Perrone, 1992; Wolpert, 1992). In the simplest implementation of the hybrid concept predictions of different classifiers are averaged together (Krogh and Vedelsby, 1995; Breiman, 1996;). More advanced methods implement rules to optimally merge multiple methods (Steele, 2000). Recently, Coe *et al.* (2005) developed a hybrid model combining an object-oriented and a pixel-based approach. Also, Liu *et al.* (2004) presented a hybrid classification approach using decision trees and ARTMAP neural networks followed by a winner-takes-all methodology using a fuzzy merging of multiple classifier outputs.

Integration Characteristics

Building on existing hybrid classifiers still poses important challenges, mostly related to the selection, interoperability and scalability of different classifiers. As the classifiers themselves are approaching their limits, the next natural challenge is to establish frameworks that integrate multiple classifiers into a single unified approach (Figure 1). There are several characteristics that such frameworks should exhibit:

- Multi-method support. Frameworks should be open enough to support classifiers from various AI methods, for example different neural networks (e.g. self-organizing maps, backpropagation and support vector machines), decision trees, and genetic algorithms among others.
- Complexity on-demand. Simple classification tasks may not require complex algorithms or multi-dimensional inputs. A successful framework should be able to match algorithms to problem complexity. By doing so, such systems typically exhibit better generalization in results because they do not suffer from overfitting.
- Transparency. The ability to backtrack errors to the original source (i.e. classifier) is critical to assess each classifier's performance. Framework architecture should be simple enough to allow identification of problematic classifiers.
- Classifier independency. Within the framework, each classifier should work independently of another. This is important to move towards "plug-in" classifiers where one may be replaced without causing a ripple effect requiring replacement of others.
- Scalability. Framework architecture should allow adjustments as problem requirements change, and/or additional groundtruth data become available.

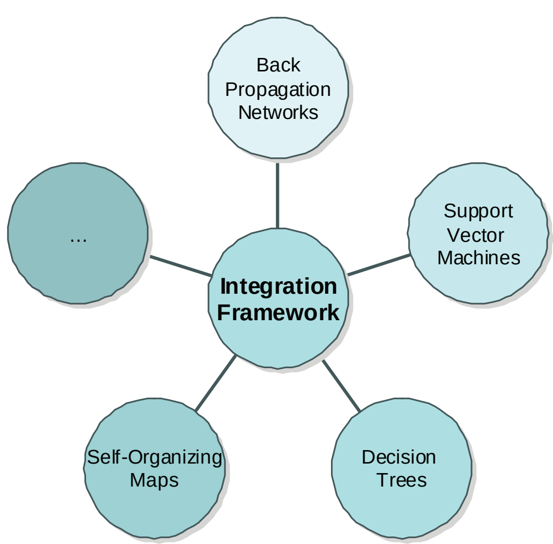

Figure 1. Establishing an integration framework supporting multiple classifiers.

ifiers: Focusing on Integration Frameworks

by Giorgos Mountrakis

Looking beyond issues related to framework architecture, a challenging task remains: establishing a methodology for optimal selection among several competing classifiers, while preserving the desired characteristics mentioned above. Work being performed in our Intelligent Geocomputing Lab at the State University of New York College of Environmental Science and Forestry has begun to show the necessity and advantages of such a task. We have successfully established an expert-based system that segments a binary multispectral classification (e.g. urban vs. non-urban areas) into context-specific sub-problems (e.g. urban areas of high brightness vs. soil of high brightness). In a classification of a Landsat scene (Figure 2) algorithmic complexity adjusts to problem specifics (Figure 3). We have also automated a process using eight different neural networks for urban sprawl modeling. Both works are currently under review (for updates and paper availability please visit www.aboutgis.com). The two aforementioned efforts are small steps towards unified frameworks, with substantial work still remaining.

Integration Benefits

The underlying objective behind integration is not to present yet another single-thread classifier; instead we strive to establish a framework for collaborative algorithms. The appropriate merging of multiple algorithms offers the following advantages:

Support for algorithmic evaluation by non-experts.

Remote sensing products often act as an additional input layer for numerous environmental studies (e.g. hydrology, biology, urban planning). It is often the case that non-experts have high expectations from remote sensing products without realizing potential sensor, acquisition and classification limitations. Therefore, there is a clear need to incorporate advanced accuracy metrics associated with remote sensing products that express usefulness and limitations of incorporated methodologies. Various works already have realized the benefits of spatially-explicit accuracy metrics (Foody *et al.*, 1992; Canters, 1997; Steele *et al.*, 1998; Carpenter *et al.*, 1999; Pontius, 2000; Alimohammadi *et al.*, 2004; Liu *et al.*, 2004; Aires *et al.*, 2004). Integrated frameworks naturally support variable

Figure 2. Landsat scene from Las Vegas, NV (April, 2000 - natural color).

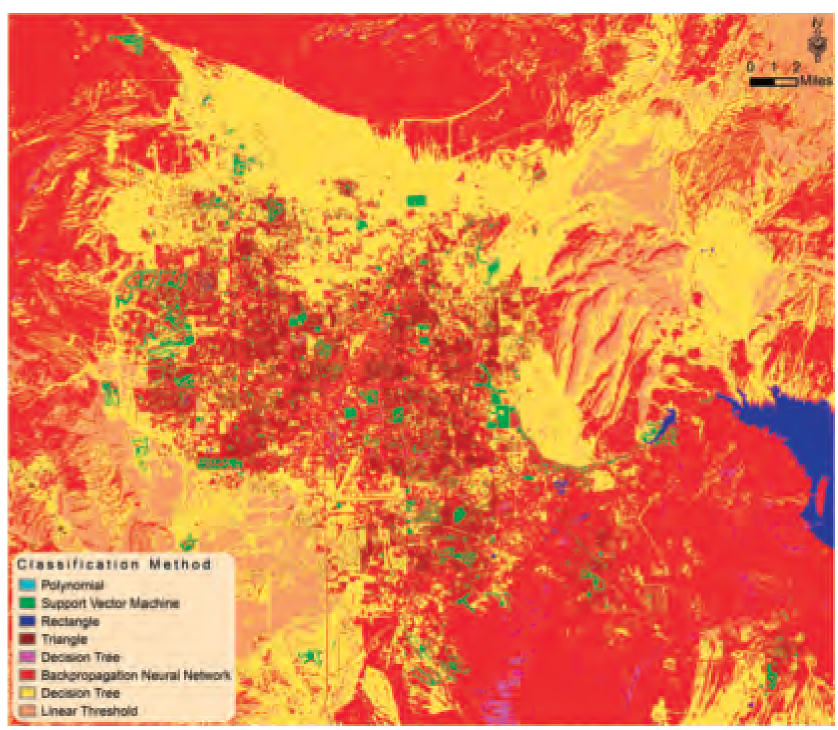

Figure 3. Spatial footprint of each selected algorithm within the framework.

continued on page 1180

continued from page 1179

accuracy metrics, each associated with a specific algorithm within the framework. For example, in Figure 3 there is an accuracy metric linked to each of the eight algorithms.

Error correction capabilities. Assuming a successful integration framework where each implemented algorithm is independent of another (i.e. independent plug-ins), future algorithmic revisions should target algorithms with lower accuracy. Ancillary datasets may be acquired in targeted areas (e.g. high-resolution imagery, lidar or census data) and as new scientific methods arise incremental algorithmic improvements can be achieved without sacrificing existing accurate works. For example, as shown in Figure 3, the decision tree (in yellow) is ~82% accurate while the whole scene is ~92% accurate, making the decision tree a prime candidate for revision. We should emphasize that this is due to both the relatively low accuracy *and* the large spatial footprint.

Support for scientific collaboration. Frameworks may separate a classification task in multiple sub-tasks as discussed earlier. There are no restrictions forcing each sub-task to be tackled by the same algorithm or scientist. Collaborative environments can be created where scientist specialization is at the foreground leading to shared instead of competitive efforts. For example, a scientist in one university could establish vegetation extraction algorithms while another focuses on urban build up identification, with their individual efforts later joined together.

Computational speed. By design, integration frameworks shine in large-scale applications, because the benefits (as mentioned above) outperform the initial cost of establishing the framework. In such environments training, error-correction and simulation speeds are important - think of a yearly update of the National Land Cover Dataset. The ability to train and simulate algorithms in a parallel fashion will utilize the latest hardware developments and in the future will allow us to analyze much higher data volumes – the majority of which is already waiting to be converted into useful products.

Summary

As algorithmic improvements in remote sensing classifiers reach their limits, the next natural frontier is the integration of multiple approaches into a unified framework. In this highlight article, characteristics for integrated frameworks are discussed, along with a demonstration of a classification process and associated benefits. Considering that image availability is expected to rapidly increase with the recent announcement from the USGS to allow free access to the Landsat archive, integrated approaches offer a unique opportunity for collaborative systems and science within our field.

Acknowledgements

The author is grateful to the following programs supporting this work: NASA New Investigator Program, NSF Geography and Regional Science Program, and Syracuse Center of Excellence Collaborative Activities for Research and Technology Innovation (CARTI) Program.

References

Aires, F., C. Prigent, and W.B. Rossow, 2004. Neural network uncertainty assessment using Bayesian statistics: A remote sensing application. *Neural Computation*, 16:2415-2458.

Alimohammadi, A., H.R., Rabiei, and P.Z. Firouzabadi, 2004. A new approach for modeling uncertainty in remote sensing change detection process, *Proc. 12th Int. Conf. on Geoinformatics – Geospatial Information Research: Bridging the Pacific and Atlantic*, 7-9 June, 2004, University of Gavle, Sweden, pp. 503-508.

Breiman, L., 1996. Bagging predictors, *Machine Learning*, 24(2):123–140.

Canters, F., 1997. Evaluating the uncertainty of area estimates derived from fuzzy land-cover classification, *Photogrammetric Engineering and Remote Sensing*, 63(4):403-414.

Carpenter, G. A., S. Gopal, S. Macomber, S. Martens, C.E. Woodcock, and J. Franklin, 1999. A neural network method for efficient vegetation mapping, *Remote Sensing of Environment*, 70:326-338.

Coe., Stefan E., M. Alberti, J.A. Hepinstall, and R. Coburn, 2005. A Hybrid approach to detecting impervious surface at multiple-scale. Proceedings of the ISPRS WG VII/1 'Human Settlements and Impact Analysis' 3rd International Symposium Remote Sensing and Data Fusion Over Urban Areas (URBAN 2005), 14-16 March 2005, Tempe, AZ.

Foody, G.M., N.A. Campbell, N.M. Trodd, and T.F. Wood, 1992. Derivation and applications of probabilistic measures of class membership from maximum-likelihood classification, *Photogrammetric Engineering and Remote Sensing*, 58(9):1335-1341.

Hansen, L.K., and P. Salamon, 1990. Neural network ensembles, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12:993–1001.

Krogh, A., and J. Vedelsby, 1995. Neural network ensembles, cross validation, and active learning, within Tesauro, G., Touretzky, D., Leen, T., *Advances in Neural Information Processing Systems*, 7, Cambridge, MA, pp. 238, MIT Press.

Liu,W.G., S. Gopal, and C.E. Woodcock, 2004. Uncertainty and confidence in land cover classification using a hybrid classifier approach, *Photogrammetric Engineering & Remote Sensing*, 70(8):963-971.

Perrone, M., 1992. A soft-competitive splitting rule for adaptive treestructured neural networks, *Proceedings of the International Joint Conference on Neural Networks*, Baltimore, MD, pp. 689–693.

Pontius, R.G., 2000. Quantification error versus location error in comparison of categorical maps, *Photogrammetric Engineering and Remote Sensing*, 66(8):1011-1016.

Steele, B.M., 2000. Combining multiple classifiers: An application using spatial and remotely sensed information for land cover type mapping, *Remote Sensing of Environment*, 74:545–556.

Steele, B.M., J.C. Winne, and R.L. Redmond, 1998. Estimation and mapping of misclassification probabilities for thematic land cover maps, *Remote Sensing of Environment*, 66(2):192-202.

Wolpert, D.H., 1992. Stacked generalization, *Neural Networks*, 5:241–259.

Author

Giorgos Mountrakis
Assistant Professor of GIS/Remote Sensing
Director of Intelligent Geocomputing Lab
Department of Environmental Resources and Forest Engineering
State Uni. of New York College of Environmental Science and Forestry
Syracuse, NY
gmountrakis@esf.edu - www.aboutgis.com