Impacts of Neural Network Architecture upon Image Classification: A Preliminary Study

Xiaojun Yang*, Libin Zhou**

* Department of Geography, Florida State University, Tallahassee, FL32317, USA; Email: xyang@fsu.edu
** Thomson Reuters Lanworth, 135 S LaSalle St, Ste 3050, Chicago, IL 60603, USA; Email:lzhou@lanworth.com

Abstract. The paper reports part of our research efforts with the aim to investigate whether neural network architectures can affect the performance of image classification. Here we initially consider multi-layer perceptron (MLP) neural networks and adaptive-resonance-theory (ART) neural networks in order to evaluate the performances of the two basic neural network structures: the simplest feed-forward structure (MLP as an example) and the recurrent structure (ART as an example). We carefully configure a set of neural network models with different internal parameters for each architecture type. Then, we use these models to classify a Landsat Enhanced Thematic Mapper Plus (ETM+) image, and the accuracy of each classified map is assessed. The optimal parameter settings for each of the two architectures are identified, and their performance is further assessed. Our initial findings indicate that the neural network architecture can significantly affect the performance of image classification.

Keywords: Neural Networks, Architecture, Image Classification, Accuracy Assessment

1. Introduction

Many studies have demonstrated that neural networks can produce identical or improved classification accuracies when compared to the outcome from conventional classifiers (e.g. Benediktsson et al. 1990; Civco 1993; Atkinson & Tatnall 1997; Mannan et al. 1998; Seto & Liu 2003; Petropoulos et al. 2010). Nevertheless, the performance of neural networks is contingent upon a wide range of algorithmic and non-algorithmic parameters (Mas & Flores 2008). Handling these diverse parameters present a challenge to beginners and even some experienced users as an inappropriate treatment can lead to suboptimal or unacceptable performance.

Several studies have investigated the sensitivity of some of the aforementioned factors affecting image classification. For example, Foody et al. (1995) examined how training set size and composition can affect neural network classification. Zhou and Yang (2011) evaluated several internal neural network parameters affecting image classification accuracy. Zhou and Yang (2010) investigated the performance of some commonly used training algorithms for image classification by the multi-layer-perceptron neural networks. Despite these efforts, some further research is needed to help better understand how various internal and external factors could affect the performance of neural networks in image classification.

The objective of this research was to investigate whether neural network architectures can affect image classification performance. We considered a representative architecture for each of two basic neural network structures: the feed-forward structure and the recurrent structure. Some discussions on these two structures and specific architectures are given in the next section. Specifically, our research comprised several major components. Firstly, we carefully configured a set of neural network models with different internal parameters for each architecture type. Then, we used these models to classify a Landsat Enhanced Thematic Mapper Plus (ETM+) image, and the accuracy of each classified map is assessed. The optimal parameter settings for each of the two architectures were identified, and their performance was further assessed. The following sections will introduce specific neural network architectures considered in our study, document the research procedural route adopted, and discuss the results and implications.

2. Neural Network Architectures Considered

Various types of neural networks can fall into either of the two basic categories of architectures: feed-forward networks and recurrent networks. The former includes single-layer networks comprising an input layer that projects onto an output layer as well as multilayer networks having at least one hidden layer that allows the networks to extract high-order statistics. A recurrent network distinguishes itself from feed-forward networks by having at least one feedback loop whose presence can greatly affect the training capability and performance (Haykin 1999).

In this study, we considered a representative architecture for each of two basic neural network structures. For the feed-forward structure, the multi-layer-perceptron (MLP) neural network architecture was targeted because of its overwhelming popularity. And for the recurrent structure, the adaptive-resonance-theory (ART) neural network architecture was considered for a similar reason.

The MLP neural networks are relatively easy to understand and implement. As the workhorse of neural networks, they have been increasingly used in remote sensing (Mas & Flores 2008). They comprise distributed neurons and weighted links. While the MLP structure and the concept of back-propagation algorithm are relatively simple, network topology and training parameter settings can complicate the overall performance of neural networks for image classification (Zhou & Yang 2010, 2011).

The adaptive-resonance-theory (ART) neural networks are powerful to recognize unexpected patterns and remember them for future use through feedback connections (Duda et al. 2001). The simplest ART network structure includes an input layer, a hidden layer and an output layer. Fuzzy ARTMAP neural networks are a supervised ART classification method. This method includes two ART modules that create stable pattern recognition based upon arbitrary sequences of training samples (Carpenter et al. 1997). Three parameters may influence the performance of fuzzy ARTMAP neural networks: the vigilance parameter ρ ranging from 0 to 1, the learning rate parameter β varying between 0 to 1, and the choice parameter α (Carpenter et al. 1997).

3. Research Methods

A scene of Landsat Enhanced Thematic Mapper Plus (ETM+) image covering Gwinnett County, Georgia, USA was used to evaluate the performance of the two supervised neural networks architectures discussed in Section 2. The test site has been a rapidly suburbanizing area, and is characterized by a variety of landscape types (Yang 2002; Yang & Lo 2002). The adopted land use/cover classification scheme included six major classes: 1) high-density urban use, mostly large commercial, educational, industrial, and transportation buildings or facilities, also including high-density residential areas in the city cores; 2) Low-density urban, mostly single/multiple family houses and public rental housing areas as well as local roads and small open spaces; 3) Exposed land, mainly non-impervious areas with sparse vegetation; 4) Cropland/grassland, including golf courses, lawns, pasture, and cropland; 5) Forest, including coniferous, deciduous, and mixed forest; and 6) Water, typically streams, rivers, lakes, and reservoirs (Figure 1).

To apply supervised classifiers for image classification, it is essential to collect training/test samples. In this case, we carefully collected a data set, including 500 pixels for each land cover class. This data set was evenly divided into two. One was used for networks training. The other was used to assess classification accuracies by neural networks. Neural networks with various network architectures and different parameters settings were trained by the selected training samples. The resultant neural network models were applied to classify the experimental image into six land cover classes.

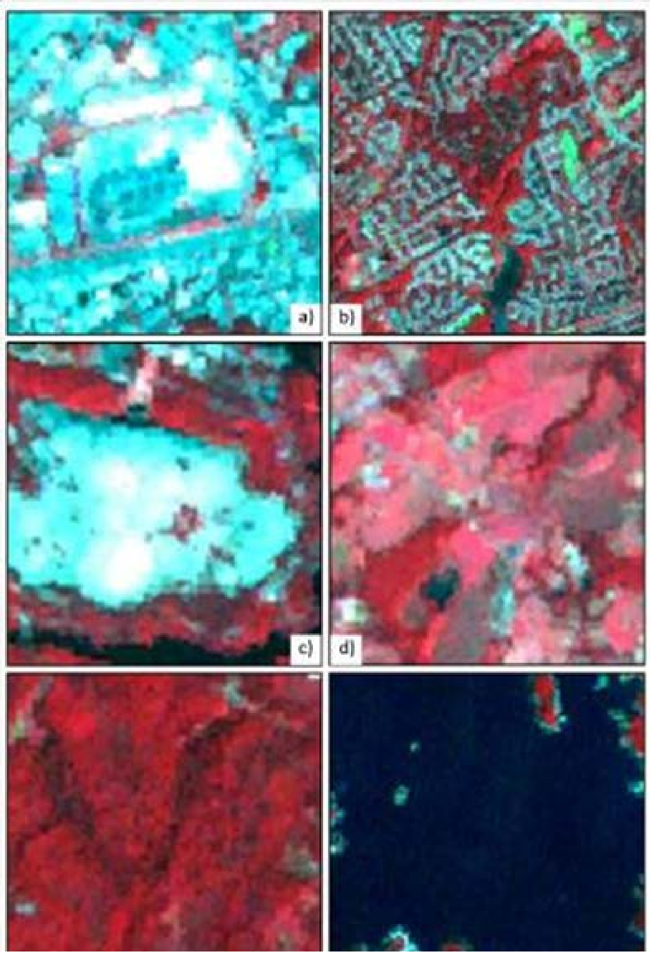

Figure 1 Six major land-cover classes in the Landsat ETM+ image to be classified (Red/Green/Blue: bands 4, 3, 2). (a) High-density urban use: a commercial centre; (b) Low-density urban use: single-family residential communities; (c) Exposed land: bare soil with sparse vegetation cover; (d) Cropland/grassland: cropland with plants on; (e) Forest: forest land; (f) Water: a lake.

The confusion matrix based on validation data is the most widely used approach to systematically measure the classification accuracy (Congalton 1991). In this study, a confusion matrix for each of the land use/cover maps was generated based on the selected test samples. Overall accuracy, overall producer's accuracy regarding two urban land classes, and producer's accuracies for high density urban use and residential urban use were contained in the error matrices.

4. Results and Discussion

4.1. MLP neural networks

To appropriately parameterize MLP neural networks, the effect of network topology and training parameters settings was assessed in terms of classification accuracy. We found that the performance of MLP neural networks varied with different parameters setting. In particular, three parameters, namely, number of hidden layers, training rate, and number of iterations had an important influence on classification accuracy. The selection of momentum and threshold slightly affected classification results.

To investigate the impact of number of hidden layer, neural networks were configured with zero, one, two, and three hidden layers while other parameters unchanged. We found that the single hidden layer neural network model performed the best in terms of the classification accuracy. And as the number of hidden layers increases, classification accuracies decrease.

The selection of momentum also affected classification accuracies. Our experiment indicates that the MLP neural networks with momentum values larger than 0.5 had higher classification accuracies. Training threshold had very small influence on the performance of land cover classification. The MLP neural networks with threshold values smaller than 0.5 provided relatively higher accuracies. Learning rate and number of iteration had an important influence on land cover classification by MLP neural networks. The neural networks with learning rate 0.01 had the best performance (greater than 95%). Classification accuracies increased as the number of iterations increased. But when the number of iterations is up to 1600, increasing iterations had limited impact on classification accuracies.

Based on the above experiment, the best performance of land use/cover classification was obtained by using the single-layer neural network model in which learning rate is set as 0.01, iterations are 1900, momentum value is 0.6, and training threshold is 0.2. And the highest overall classification accuracy achieved by MLP neural networks is 96.67%. The overall producer’s accuracy regarding two urban classes is 93%. The producer’s accuracies for high density urban use and residential urban use are 92.4% and 93.6%, respectively.

4.2. Fuzzy ARTMAP Neural Networks

To optimize ARTMAP neural networks, a set of ARTMAP neural networks were parameterized with the choice value ranging from 0 to 0.9 while keeping other parameters unchanged. Another set of neural networks were configured with the vigilance value ranging from 0.1 to 1 while fixing other parameters. Ten more neural networks were parameterized by changing the learning rate. These networks were used to classify the ETM+ image. The optimal parameter setting for ARTMAP neural networks was evaluated by comparing the classification accuracies by these models.

The performance of ARTMAP neural networks was very sensitive to the vigilance value but less sensitive to the choice value. Classification accuracies declined rapidly as the vigilance value decreased. In addition, all the experiments with the learning rate rather than 1 failed to classify the image. The highest classification accuracy by the Fuzzy ARTMAP neural networks (92.13%) was achieved when the choice value was 0.7 or 0.8, the vigilance value was 0.98, and the learning rate was 1. By using this configuration, we obtained the overall classification accuracy of 0.921 and conditional Kappa coefficients of 0.775 for the high-density urban use and 0.858 for the low-density urban use, respectively.

Based on the above experiments, the MLP neural networks outperformed the ARTMAP networks in terms of the overall classification accuracy and the accuracy for each of the two urban land uses (Table 1).

<table>
<tr><th rowspan="3">Neural Networks</th><th colspan="3">Best Performance</th></tr>
<tr><th rowspan="2">Overall Classification Accuracy</th><th colspan="2">Conditional Kappa Coefficients</th></tr>
<tr><th>High-density urban</th><th>Low-density urban</th></tr>
<tr><td>MLP</td><td>0.962</td><td>0.897</td><td>0.908</td></tr>
<tr><td>ARTMAP</td><td>0.921</td><td>0.775</td><td>0.858</td></tr>
</table>

Table 1. Comparisons of the best performances of two neural network architectures for land use/cover classification.

5. Conclusions

In this study, we have evaluated the performances of the two basic neural network structures: the simplest feed-forward structure (with the multilayer perceptron neural networks MLP as an example) and the recurrent structure (with the adaptive-resonance-theory neural networks as an example). The entire work has gone through several procedures. Firstly, we carefully configured a set of neural network models with different internal parameters for each architecture. Then, we used these models to classify a Landsat Enhanced Thematic Mapper Plus (ETM+) image covering an urban area, and the accuracy of each classified map was assessed. The optimal parameter settings for each of the two architectures were identified, and their performance was further assessed. We targeted urban classes because their heterogeneous nature normally makes difficult to classify. We found that the MLP neural networks clearly outperformed the ART networks. Although only two architecture types have been considered, our research suggests that the neural network architecture can significantly affect the performance of land use/cover classification from remotely sensed imagery. Our further research intends to consider some more architecture types in order to better understand neural network architectures affecting image classification accuracy. This can help select and design efficient neural network models for improved performance of remote sensor image classification.

Acknowledgments

The authors would like to thank the Florida State University for the time release in conducting this work. The research was partially supported by the Florida State University Council on Research and Creativity, China's State Key Laboratory of Urban and Regional Ecology, and the Chinese Academy of Sciences through the International Partnership Project "Ecosystem Processes and Ecological Services".

References

Atkinson PM, Tatnall ARL (1997) Introduction: Neural Networks in Remote Sensing. International Journal of Remote Sensing 18(4):699-709.

Benediktsson JA, Swain PH, Ersoy OK (1990) Neural Network Approaches versus Statistical Methods in Classification of Multisource Remote Sensing Data. IEEE Transactions on Geosciences and Remote Sensing 28(4):540-551.

Carpenter GA, Gjaja MN, Gopal S, Woodcock CE (1997) ART Neural Networks for Remote Sensing: Vegetation Classification from Landsat TM and Terrain Data. IEEE Transactions on Geoscience and Remote Sensing 35(2):310-324.

Civco DL (1993) Artificial Neural Networks for Land-Cover Classification and Mapping. International Journal of Geographical Information Systems, 7(2):173-186.

Congalton RG (1991) A Review of Assessing the Accuracy of Classification of Remotely Sensed Data. Remote Sensing of Environment 37(1): 35-46.

Duda RO, Hart PE, Stork DG (2001) Pattern Classification. John Wiley & Sons.

Foody GM, McCulloch MB, Yates WB (1995) The Effect of Training Set Size and Composition on Artificial Neural-Network Classification. International Journal of Remote Sensing 16(9):1707-1723.

Haykin S (1999) Neural Networks: A Comprehensive Foundation (2nd). Prentice Hall.

Mannan B, Roy J, Ray AK (1998) Fuzzy ARTMAP Supervised Classification of Multi-Spectral Remotely-Sensed Images. International Journal of Remote Sensing 19(4): 767-774.

Mas JF, Flores JJ (2008) The Application of Artificial Neural Networks to the Analysis Of Remotely Sensed Data. International Journal of Remote Sensing 29(3): 617-663.

Petropoulos GP, Vadrevu KP, Xanthopoulos G, Karantounias G, Scholze M (2010) A Comparison of Spectral Angle Mapper and Artificial Neural Network Classifiers Combined with Landsat TM Imagery Analysis for Obtaining Burnt Area Mapping. Sensors 10(3):1967-1985.

Seto KC, Liu WG (2003) Comparing ARTMAP Neural Network with the Maximum-Likelihood Classifier for Detecting Urban Change. Photogrammetric Engineering and Remote Sensing 69(9):981-990.

Yang X (2002) Satellite Monitoring of Urban Spatial Growth in the Atlanta Metropolitan Region. Photogrammetric Engineering and Remote Sensing 68(7):725-734.

Yang X, Lo CP (2002) Using a Time Series of Normalized Satellite Imagery to Detect Land Use/Cover Change in the Atlanta, Georgia Metropolitan Area. International Journal of Remote Sensing 23(9):1775-1798.

Zhou L, Yang X (2010) Training Algorithm Performance for Image Classification by Artificial Neural Networks. Photogrammetric Engineering and Remote Sensing 76(8): 945-951.

Zhou L, Yang X (2011) An Assessment of Internal Neural Network Parameters Affecting Image Classification Accuracy. Photogrammetric Engineering and Remote Sensing 77(12): 1233-1240.