Proceedings of the 2007 IEEE Symposium on
Computational Intelligence and Data Mining (CIDM 2007)

Evolutionary Neural Networks Applied to Land-cover Classification in Zhaoyuan, China

Yan Guo, Lishan Kang, Fujiang Liu, Huashan Sun and Linlu Mei

Abstract—**This paper proposes a method for the classification of land cover in remote sensing imagery using evolutionary artificial neural networks (EANN) compared against multilayer perceptrons (MLP) with backpropagation algorithm. Evolutionary neural networks have combined the features of artificial neural networks (ANN) and evolutionary algorithms (EA) in the way that simultaneously evolving ANN architecture and weights. The parsimony of evolved ANN is encouraged by preferring node mutation and connection mutation. This enables consistent reductions of mean square errors of spectral classification with respect to sample pixels. Land-cover classification experiments were carried out by EANN-based classifiers and MLP-based classifiers in a 300×300 pixels Landsat-7 Enhanced Thematic Mapper plus (ETM+) high-resolution image of Zhaoyuan in Shandong province in eastern China. We found that the use of evolutionary algorithms for finding the optimal ANN results mainly in improvements in overall accuracy of an ANN with backpropagation algorithm and produce more compact ANN with good generalization ability in comparison with MLP. It is observed that classification accuracy of up to 90% is achievable for Landsat data produced by EANN.**

I. INTRODUCTION

Both artificial neural networks (ANN) and Evolutionary algorithms (EA) have long been used for the land-cover classification in remotely-sensed images, such as [1]-[3] with artificial neural networks and [4]-[6] with evolutionary algorithms. While each methodology has unique properties, these have been used separately. Evolutionary neural networks (EANN) [7]-[8] are the combination of artificial neural networks and evolutionary algorithms. This merge enabled these two methods to complement the disadvantages of the other [7]. For example, a contribution by artificial neural networks was the flexibility of nonlinear function approximation, which cannot be easily implemented with prototype evolutionary algorithm. On the other hand, evolutionary algorithm has freed artificial neural networks from simple gradient descent approaches of optimization. Indeed, traditional artificial neural networks based on backpropagation algorithms have some limitations. At first, the architecture of the artificial neural networks is fixed and a designer needs much knowledge to determine it. Also, error function of the learning algorithm must have a derivative. Finally, it frequently gets stuck in local optima because it is based on gradient-based search without stochastic property. The combination of evolutionary algorithm and artificial neural networks can overcome these shortcomings and is particularly useful when the activation function of the neurons is non-differentiable and traditional gradient-based training algorithms cannot be used. Because EA can treat nondifferentiable and multimodal spaces, which are the typical case in the classification of remotely-sensed imagery, there must be a great motivation to apply EANN to classification of remotely-sensed imagery [8].

A lot of works have been made on EANN. Evolutionary algorithms have been used to help to obtain more accurate ANN with better generalization abilities. For example, searching the optimal weight set of a ANN, designing its architecture, finding it's most adequate parameter set (number of neurons in the hidden layer, learning rate, etc.) among others tasks.

The paper aims to search for the best ANN among evolving populations of potential solutions, regarding their ability to classification of land cover. To do so, EA was used for evolving simultaneously the architecture and connection weights (including biases) for an optimal ANN. The parsimony of evolved ANN is encouraged by preferring node mutation and connection mutation. Moreover, land-cover classification experiments were carried out by EANN-based classifiers in a 300×300 pixels Landsat-7 Enhanced Thematic Mapper plus (ETM+) high-resolution image of Zhaoyuan in Shandong province in eastern China. For comparing, multilayer perceptrons classifiers with backpropagation were employed.

The rest of this document is organized as follows. In the second section, the proposed EANN algorithm is described. In the third section, the data sets we selected for this study are described and the experiments and the main results of this work are discussed. Finally, in the fourth section the

Manuscript received October 30, 2006. This work was supported in part by the National Natural Science Foundation of China under Grant 60133010, the Research Foundation for Outstanding Young Teachers of China University of Geosciences (Wuhan) under Grant CUGQNL0640 and CUGQNL0628 and the Postdoctoral Foundation of the Shandong Zhaojin Group Co. under Grant 2005026212.

Yan Guo is with the School of Computer, China University of Geosciences, Wuhan, 430074 China (phone: +86-27-67884684; fax: +86-27-67883716; e-mail: guoyanwuhan@yahoo.com.cn).

Lishan Kang was with Wuhan University, Wuhan, 430074 China. He is now with School of Computer, China University of Geosciences, Wuhan, 430074 China (e-mail: kang_whu@yahoo.com).

Fujiang Liu is with the Faculty of Information Engineering, China University of Geosciences, Wuhan, 430074 China (e-mail: felixwuhan@yahoo.com.cn).

Shanhua Sun is with the Faculty of Resources, China University of Geosciences, Wuhan, 430074 China (e-mail: sunhsh@126.com).

Linlu Mei is with the Faculty of Information Engineering, China University of Geosciences, Wuhan, 430074 China (e-mail: meilinlu@163.com).

conclusions are presented and some comments about future work are also made.

II. METHODOLOGY

A. Evolutionary Neural Networks for land-cover classification

The evolutionary neural networks for land cover classification method are inspired by the evolutionary approach as proposed by [9]-[12]. This method aims to search for the best ANN among evolving populations of potential solutions, regarding their ability to classification of land cover. To do so, we evaluate the strength of the correlation between each individual and the ground truth samples. Consider a three layer feed-forward neural network of n input units (features), h hidden units in the only hidden layers, and m output units.

Each sample is used to train or validate each individual of the population (a set of individuals). A data separation procedure divides the data into three distinct sample sets without overlapping: training set, validation set, and test set. The training set is used for partial training and the validation set is used for fitness calculation.

The designed algorithm is specified in the following pseudocode:

--1. Generate a random initial population of potential solutions with random initial weights and random number of hidden nodes within certain ranges according to a uniform distribution.

--2. Each individual in the population is trained partially by backpropagation algorithms for a certain number of epochs on the training set to help the evolution search the optimal architecture of ANN.

--3. Evaluate the individuals in the population on the validation set based on their error.

--4. Pick the V best individuals in the population as elite individual and select W individuals based on rank-based select mechanism.

--5. Apply tailored genetic operations to selected W individuals and obtain W offspring.

--6. Replace the worst individuals with the new ones only when the offspring is better than the current worst and form a new generation.

--7. Loop to Step 2 until an individual shows better performance than predefined accuracy or iteration number exceeds predefined maximum number of generations.

--8. Use the best ANN on an unseen testing set and evaluate the testing error.

As end-products of this process, an optimal ANN for classifying land cover is obtained, which in this case, are products of the evolutionary neural networks approaches that are proposed.

B. Encoding Scheme for Feed-forward ANN

In order to evolve the ANN's architecture and connection weights (including biases) simultaneously, a $N\times N$ matrix $C=(c_{ij})_{N\times N}$ is used to represent an ANN individual with N nodes, where c_{ij} ($i \leqslant j$) indicates presence or absence of the connection from node i to node j (We can use c_{ij} = 1 to indicate a connection and c_{ij} = 0 to indicate no connection) and c_{ij} ($i > j$) is the corresponding real valued weight.

The maximum number of hidden nodes must be predefined in this representation, but it is not necessary that all hidden nodes are used. It is obvious that such an encoding scheme is straightforward to implement and easy to apply tailored genetic operators [13].

C. Fitness Evaluation, selection mechanism and replacement strategy

As mentioned above, the goal is to evolve ANN so that to minimize the mean square errors (*MSE*) of each ANN, as

$$min\ MSE = \frac{1}{|G|\cdot|S|}\sum_{g\in G}\sum_{s\in S}(Y_g(s)-O_g(s))^2 \qquad (1)$$

where G is the output set of the EANN, S is the set of validation data, $Y_g(s)$ and $O_g(s)$ is the ideal and actual outputs in [0, 1] of validation pixel s for class g. Equation (1) makes the error measure less dependent on the size of the validation set and the number of output nodes. It is determined through a validation set which does not overlap with the training set. Such use of a validation set in an evolutionary learning system improves the generalization ability of evolved ANN but introduces little overhead in computation time [13].

The selection mechanism used here is rank based. The probability $p(k)$ for the k^{th} individual to be selected among the M-size population is given by:

$$p(k) = (M-K)\Big/\sum\nolimits_{i=1}^{M} i \qquad (2)$$

where M individuals are sorted and numbered as 0, 1, ……, $M-1$, with the zeroth being the fittest.

Since it is risky to apply mutations to all the chromosomes, it is a convention to leave some best chromosomes (elite individuals) that generate the less *MSE* free from mutation. At every generation, the population of chromosomes are ranked and sorted according to their *MSE* or fitness. As the natural selection the worse of the population die and are replaced by the clones of the better.

D. Genetic Operators

Because permutation makes crossover operator very inefficient and ineffective in producing good offspring [8] [13], crossover operation is avoided here and only mutation operation is adopted.

The mutation operator includes architecture mutation and modification of existing connection weights. The modification of weights in an ANN here is only carried out by backpropagation algorithms. The number of epochs used by backpropagation algorithms to train each ANN in a population is defined by user-specified parameter. This training process is called partial training because there is no guarantee that an ANN will converge to even a local optimum after those epochs.

The architecture mutation operator is used to grow and prune hidden nodes and connections, including connection mutation and node mutation. Connection mutation operator selects an ANN from the population randomly and then chooses one connection from at random. If the connection does not exist and the connection entry of the ANN matrix is 0, the new connection is added with random weights. Otherwise, if the connection already exists, the connection and weight information are deleted. Node mutation includes addition of a new hidden node and deletion of an existing hidden node and its related connections. Node mutation is easy to perform by flipping a bit in the ANN matrix.

III. EXPERIMENTS AND RESULTS

The proposed EANN algorithm was applied to real world classification tasks of remotely-sensed images. The obtained optimal solutions are compared with multilayer perceptron (MLP) classifiers with backpropagation algorithm. All the programs were written in MATLAB Version 7.2.0.232 (R2006a), and executed on a PC with a single 2.4 GHz Intel Pentium IV processor with 1 GB of memory. The following subsections explain the experimentation and results of the study.

A. Data Set

The study area selected for this study was an agricultural area located in Zhaoyuan in Shandong province in eastern China. Zhaoyuan is located in northwest of Shandong Peninsular with a latitude 37°05′ – 37°33′ and a longitude 120°08′ – 120°38′.

A Landsat-7 Enhanced Thematic Mapper plus (ETM+) high resolution image acquired on June 12, 2000 was employed to identify land-cover classes. The intensity of a pixel is resolved on the electromagnetic spectrum into seven bands, which are taken to be seven features. Since the features are highly correlated, principal component analysis was done to reduce the seven features to three principal features. The three bands are: Green band of wavelength (0.52–0.60 um), Red band of wavelength (0.63–0.68 um), and near infrared band of wavelength (0.76–0.90 um). The classification problem involved the identification of four land cover types, namely, buildings, pond and river, forest, and farmland. A 300 by 300 pixel sub-scene Landsat-7 ETM+ image of the area in Zhaoyuan, Shandong, China after pre-processing is shown in Fig.1.(a). Evolutionary artificial neural and required analyses including accuracy assessment and kappa statistic were carried out using MATLAB.

In order to present the spectral variation of each land cover type to the artificial neural networks, sample sets for each class were selected from the 300 by 300 pixels image and a total of 2400 samples were generated. The input attributes used in this work were rescaled in the range [0, 1] and divided into ten non-overlapping splits, each one with 60% of the data for training while 20% is used for validation and the remaining 20% for testing.

TABLE I
LIST OF PARAMETERS USED TO EXECUTE THE EANN ON THE LAND-COVER CLASSIFICATION

Parameter	Value
population size	20
stopping criterion	error = 0.01; maximum generation=5000
selection operator	rank-based selection
elitism	elite number= 2
maximum number of hidden units	50

BP partial train	learning rate=0.05;
	maximum epochs=50;
	momentum coefficient=0.9

B. Experiments and Results

A ten-fold crossvalidation trial was performed; that is, the EANN algorithm was executed ten times, each time using a different split on the data with 60% of the total dataset for training while 20% is used for validation and the remaining 20% for testing.

To perform the simulations, the EANN was executed using the parameters shown in Table I. In the EANN-based classifiers used to evolve land-cover classifiers, the population size is kept equal to 20 individuals. The target training performance of EANN was set to 0.0100 and the maximum generation number is 5000. Fitness function of EANN is defined as Equation (1) on validation data. Each ANN is feed-forward ANN with one hidden layer and the transfer function of every unit is the sigmoid hyperbolic tangent function. Backpropagation algorithm is used as partial training algorithm with the learning rate of 0.05, the momentum coefficient of 0.9 and 50 epochs.

After some experiments using the evolutionary neural networks methodology described in the previous section, a best evolutionary neural network structure of 3-10-4 (3 indicates the number of inputs, 10 is the number of nodes in the hidden layer, and the number of output classes is 4) was found to be appropriate to learn the characteristics of the training data and validation data. In the optimal neural network ,there are 14 connections from input nodes to hidden nodes, 3 connections from input nodes to output nodes, 3 connections from hidden nodes to hidden nodes, and 5 connections that from hidden nodes to output nodes, which is more compact than multilayer perceptrons.

In order to compare the performance of the EANN classifiers with other well know ones, a 3-layered multilayer perceptron (MLP) was used with 5–50 hidden nodes, 3 input nodes, 4 output nodes, 0.05 of learning rate and 0.9 of momentum. For the MLP-based classifiers, the target training performance is 0.0100 and maximum epoch is 5000.

Ten runs of experiments were conducted and the means of the overall performance and kappa statistic were reported.

Here, the classification results were compared with our interpretation results. These results are shown in Table II. Different land- cover classification images are illustrated in Fig. 1 (b)–(c), respectively.

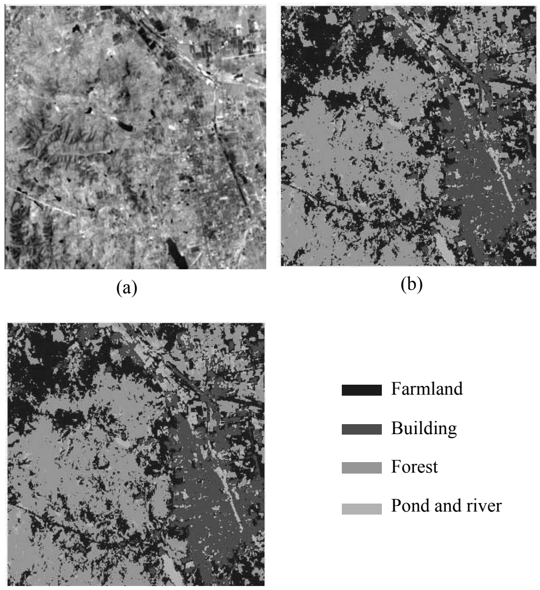

Fig. 1. Classification results of Landsat-7 Enhanced Thematic Mapper (ETM+) high resolution image of Zhaoyuan, Shandong, China. (a) A sub-scene Landsat-7 ETM+ image of the area in Zhaoyuan, Shandong, China. (b) Classified image of the sub-scene in Fig.1.(a) using a MLP-based classifier. (c) Classified image of the sub-scene in Fig.1.(a) using a EANN-based classifier.

Table II shows the above measures for different classifiers. Here, all parameters for the classifiers were kept the same as above. It is observed from Table II the evolutionary neural networks classifiers performed better (in terms of user's accuracy and kappa value) than the simple backpropagation method for the classification of land cover. The overall user's accuracies for EANN-based classifiers and MLP-based classifiers are 90.12% and 88.06% respectively.

The reason for better accuracy with EANN classifier than BP classifier is possibly because EA could be used effectively in the training of connection weights to evolve and find a near-optimal ANN globally, which avoid to be trained several times with different ANN to prevent the networks becoming stuck into a local minimum but make the influence of ANN's architecture and the initial values of the connection weights on the final classification accuracy least.

IV. CONCLUSION AND FUTURE WORK

This paper present a comparison of multilayer perceptrons with backpropagation training and a method for combining EA and ANN (evolving simultaneously the architecture and connection weights for an optimal ANN) applied to the classification of land cover from the Landsat-7 Enhanced Thematic Mapper plus high resolution image. Our experimentation suggests that evolutionary neural networks methods are better than the simple backpropagation method for the classification of land cover in terms of overall accuracy and produce more compact ANN with good generalization ability in comparison with MLP.

TABLE II
COMPARISON OF AVERAGE TEST ACCURACY: USER ACCURACY AND KAPPA VALUES IN % CORRESPONDING TO THE DIFFERENT CLASSES (TEN RUNS)

Land cover	User accuracy (%)		Kappa statistic	
	MLP	EANN	MLP	EANN
Pond and river	79.09	88.78	0.6812	0.8477
Building	88.67	90.16	0.8665	0.8867
Forest	90.54	93.43	0.9367	0.9244
Farmland	91.25	71.33	0.9018	0.6890
Overall	88.06	90.12	0.8635	0.8879

In this study, EA and ANN were combined, and more sophisticated versions of these methods could produce better results. For example, the use of indirect representation encoding chromosomes, use of EANN for feature selection, etc. could bring different results. The inclusion of backpropagation training in the EANN have consequences of longer computation times, so alternatives to backpropagation should be tested in order to reduce time costs.

As future work, it would be useful to include and process other remotely sensed images, in order to have more examples to test how different resolutions could affect system effectiveness. It would get better generalization and classification accuracy to evolve the architecture and other learning parameters of ANN such as learning rate and momentum coefficient by using EA which is another work in the future. Although the evolutionary neural networks method has several unique capabilities, more works are needed in overcoming the slow convergence of the evolutionary approach to ANN and evaluating the performance of the proposed method on larger remotely-sensed images.

V. ACKNOWLEDGEMENTS

Yan Guo thanks Prof. Yong Liu for his insightful comments while discussing the initial idea with him. We would also like to show our appreciation to the anonymous reviewers for their invaluable comments and suggestions.

REFERENCES

[1] J. D. Paola, R. A. Schowengerdt, "A review and analysis of backpropagation neural networks for classification of remotely-sensed multi-spectral imager", International Journal of Remote Sensing, 16(16), 3033-3058(1995).

[2] P. M. Atkinson, A. R. L. Tatnall, Neural networks in remote sensing, Int. J. Remote Sensing, 18, 699–709(1997).

[3] I. Kanellopoulos, G. G. Wilkinson, Strategies and best practice for neural network image classification, Int. J. Remote Sensing, 18, 711–725(1997).

[4] Steven P. Brumby, James Theiler, Jeffrey J. Bloch, Neal R. Harvey, “Evolving land cover classification algorithms for multi-spectral and multi-temporal imagery”, Proc. SPIE 4480, 120-129(2002).
[5] Sankar K. Pal, Sanghamitra Bandyopadhyay, C.A. Murthy, "Genetic classifiers for remotely sensed images: comparison with standard methods," International Journal of Remote Sensing, 22(1), 2545-2569 (2001).
[6] S. Perkins, J. Theiler, S. P. Brumby, N. R. Harvey, R. B. Porter, J. J. Szymanski, J. J. Bloch, “GENIE: a hybrid genetic algorithm for feature classification in multi-spectral images,” Proc. SPIE 4120, 52-62(2000).
[7] D. B. Fogel, L. J. Fogel, V. W. Porto, “Evolving neural networks,” Biological Cybernetics, 63, 487–493(1990).
[8] Xin Yao, “Evolving artificial neural networks,” in Proceedings of the IEEE, 87(9), 1423–1447(1999).
[9] Rolando R. Hernandez-Cisneros, Hugo Terashima-Marın, “Evolutionary neural networks applied to the classification of microcalcification clusters in digital mammograms”, IEEE Congress on Evolutionary Computation, 8628-8635(2006).
[10] Yuya Sasaki, Hugo de Garis, Paul W. Box, “Genetic Neural Networks for Image Classification,” 0-7803-7929-2/03 2003 IEEE
[11] Tatiana M. Kuplich, “Classifying regenerating forest stages in Amazonia using remotely sensed images and a neural network”, Forest Ecology and Management 234, 1-9(2006)
[12] Kyung-Joong Kim, Sung-Bae Cho, “Prediction of colon cancer using an evolutionary neural network,” Neurocomputing 61, 361-379(2004)
[13] Xin Yao, Yong Liu, “A New Evolutionary System for Evolving Artificial Neural Networks”, IEEE Transaction on neural networks, 8(3), 694-713(1997).