2001

A hierarchical classifier for multispectral satellite imagery

Abdesselam Bouzerdoum
University of Wollongong, bouzer@uow.edu.au

A hierarchical classifier for multispectral satellite imagery

Abstract

In this article, a hierarchical classifier is proposed for classification of ground-cover types of a satellite image of Kangaroo Island, South Australia. The image contains seven ground-cover types, which are categorized into three groups using principal component analysis. The first group contains clouds only, the second consists of sea and cloud shadow over land, and the third contains land and three types of forest. The sea and shadow over land classes are classified with 99% accuracy using a network of threshold logic units. The land and forest classes are classified by multilayer perceptrons (MLPs) using texture features and intensity values. The average performance achieved by six trained MLPs is 91%. In order to improve the classification accuracy even further, the outputs of the six MLPs were combined using several committee machines. All committee machines achieved significant improvement in performance over the multilayer perceptron classifiers, with the best machine achieving over 92% correct classification.

Keywords

hierarchical, classifier, satellite, multispectral, imagery

Disciplines

Engineering | Science and Technology Studies

Publication Details

A. Bouzerdoum, "A hierarchical classifier for multispectral satellite imagery," IEICE Transactions on Electronics, vol. E84-C, (12) pp. 1952-1958, 2001.

PAPER Special Issue on New Technologies in Signal Processing for Electromagnetic-wave Sensing and Imaging

A Hierarchical Classifier for Multispectral Satellite Imagery

Abdesselam BOUZERDOUM†a), *Nonmember*

SUMMARY In this article, a hierarchical classifier is proposed for classification of ground-cover types of a satellite image of Kangaroo Island, South Australia. The image contains seven ground-cover types, which are categorized into three groups using principal component analysis. The first group contains clouds only, the second consists of sea and cloud shadow over land, and the third contains land and three types of forest. The sea and shadow over land classes are classified with 99% accuracy using a network of threshold logic units. The land and forest classes are classified by *multilayer perceptrons* (MLPs) using texture features and intensity values. The average performance achieved by six trained MLPs is 91%. In order to improve the classification accuracy even further, the outputs of the six MLPs were combined using several committee machines. All committee machines achieved significant improvement in performance over the multilayer perceptron classifiers, with the best machine achieving over 92% correct classification.

key words: *neural networks, multilayer perceptrons, image classification, satellite imagery, committee machines*

1. Introduction

The classification of satellite imagery has important military and civilian applications. The challenging task of generating discrete classes of known identity for land-cover types is very useful in remote sensing, for example, for general resource assessment, inventory and management. This article presents the results of an investigation into the classification of ground-cover types in a satellite image. The image, which consists of the three visible spectral bands, is of a portion of Kangaroo Island, South Australia. It contains sea, land, forest and clouds, which produce shadows on the ground and obscure vision of the ground in some areas.

Multi-spectral classification of satellite images has been extensively investigated and several approaches for classification have been proposed [9]. However, classification based on the intensity values of a single pixel can lead to an increased misclassification error if the number of classes is large. Texture, on the other hand, is characterized by pixel values of a small region surrounding the pixel. Useful texture features convey information about the pixel itself and the region surrounding it. Texture analysis has been used in several applications. One of the most actively studied applications is in image segmentation where the image is divided into similar "meaningful" regions. Here we use texture features to improve the classification of ground-cover types of a satellite image.

In 1996 we reported the results of a preliminary investigation into the classification of a satellite image using texture features and neural networks [2]. In that study almost a 13% increase in classification rate was achieved by including texture features, in addition to pixel intensity values. In this article, we extend the study by using more accurate neural network models and committee machines. The average correct classification rate achieved by six multilayer perceptrons (MLPs) is 91% compared to 88% accuracy achieved in the previous study. To improve the classification accuracy even further, the outputs of the six MLPs are combined using several committee machines. Thus, the correct classification rate is increased to over 92%.

The paper is organized as follows. In the next section the image data used in this analysis is described. The third section describes the mechanism used to extract the features used in the classification process. Section 4 presents the hierarchical classification scheme used to identify the ground-cover types in the image. Finally, the conclusion is presented in Sect. 5.

2. Image Data

The data used in this study was obtained from a satellite image of a part of Kangaroo Island, South Australia. The image consists of the three visible spectral bands, each of size 500 × 500. This image, which is shown in Fig. 1, contains six ground-cover types, three types of forest, beach, sea, and land. In addition, there are clouds which produce shadows on the ground and obscure vision of the ground in some areas.

Figure 2 shows the ground truth image containing seven distinct classes: forest types 1, 2, and 3, sea, beach, land, and clouds. The class shadow on the ground does not appear on this image, but it can easily be seen in the original image shown in Fig. 1: it's the dark areas over the land region. The ground truth image was obtained by image registration using GIS (Geographical Information Systems) data provided by the Defence Science and Technology Organisation (DSTO). In this study, we attempt to classify the pixels of the satellite image into one of seven identifiable classes: forest types 1, 2, and 3, sea, land, clouds, and shadow over the ground. Since the beach area is too small and is very similar to the clouds, it is considered a part of the cloud class in this study.

Manuscript received April 2, 2001.
Manuscript revised May 29, 2001.

†The author is with the School of Engineering and Mathematics, Edith Cowan University, 100 Joondalup Drive, Perth, WA 6027, Australia.
a) E-mail: a.bouzerdoum@ecu.edu.au

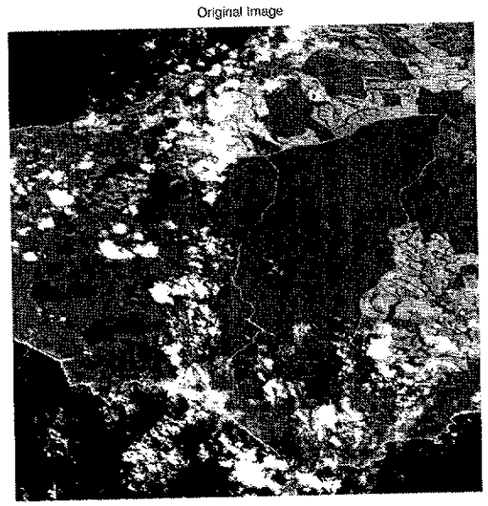

Fig. 1 The satellite image of a portion of Kangaroo Island, SA. The image consists of the three visible spectral bands.

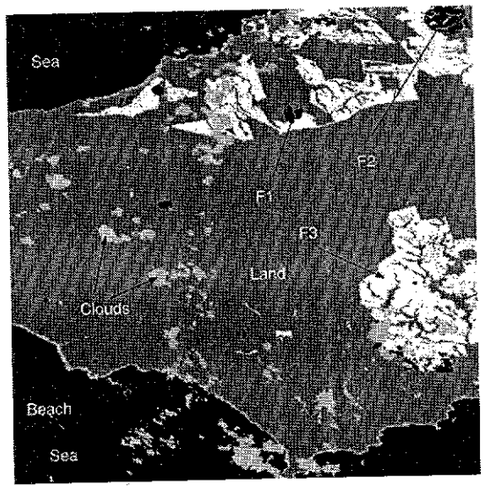

Fig. 2 The discrete classes of ground-cover types: forest 1 (F1), forest 2 (F2), forest 3 (F3), sea, beach, land, and clouds.

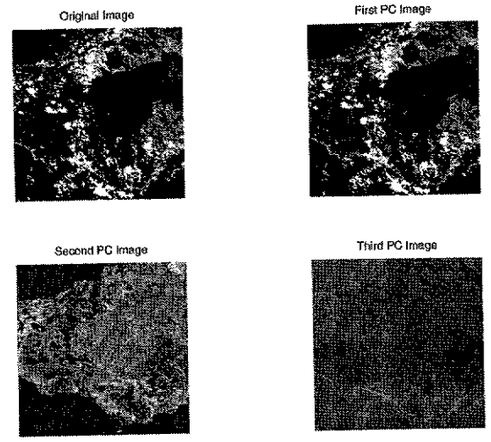

Fig. 3 The principal component images.

3. Feature Extraction

In image classification, usually a small image region is reduced into a feature vector which is then assigned, using some algorithm, to one of the known classes [9]. Here we use supervized classification to assign the feature vectors to one of the different land-cover types. In supervised classification, a data subset representative of all the existing classes is extracted from the original data. Then, feature vectors are formed from this subset and used to train the classifier system. In our classification scheme we used two types of features: intensity features and texture features. The intensity features are extracted by performing a principal component analysis transformation on the original data.

3.1 Principal Component Analysis

Multispectral image data usually has extensive inter-band correlation; that is, images generated by different channels appear similar and convey essentially the same information. Principal component analysis (PCA) is a technique that is often used to reduce redundancy and dimensionality of the data [9]. The aim of PCA is to find a set of m orthogonal vectors in the n-dimensional (n-D) data space that can account for as much as possible of the data's variance. Therefore, projecting the data vectors from their original n-D space to the m-D space spanned by the principal component vectors then performs a dimensionality reduction that often retains most of the intrinsic information in the data.

For classification purposes, the aim is to find the directions which provide maximum class separability. These directions may not necessarily be the directions with the largest eigenvalues. Principal component analysis of the image under investigation revealed that the first principal component accounts for 88.36% of the total variance of the original data, the second component accounts for 11.04%, and the remaining 0.6% is due to the third principal component. Figure 3 shows the three principal component images. Clearly, the first principal component image is very similar to the original image.

Figure 4 presents the estimated probability density functions (pdfs) of the seven classes, using Parzen windows. This figure clearly shows the separability of the

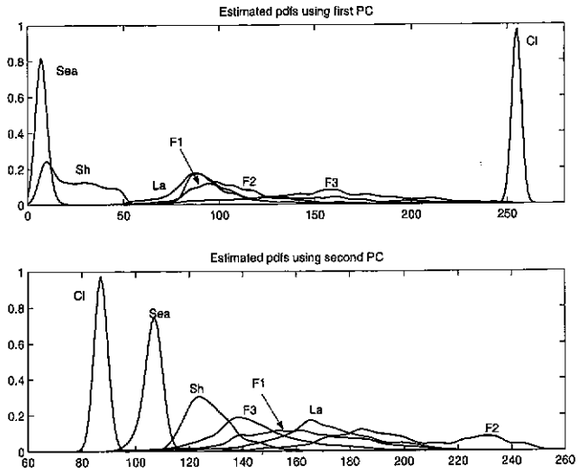

Fig. 4 Estimated pdfs of the classes: sea (Sea), shadow over land (Sh), land (La), clouds (Cl), forest type-1 (F1), forest type-2 (F2), and forest type-3 (F3).

three classes Cl, Sea and Sh. However, the other four classes have considerable overlap in the first and second principal components. The situation is rather worse in the third component as this accouts only for 0.6% of the total variance of the original data; due to space limitations, the class pdfs associated with the third principal component are not shown. From the pdfs in Fig. 4, we can conclude that classification based on intensity values of single pixels may not achieve high accuracy. Therefore, we consider texture features as well as intensity values to improve the classification accuracy of the classes F1, F2, F3 and La.

3.2 Texture Features

In order to determine the texture features which are useful in the classification process, it is necessary to have an intuition as to what constitutes a texture. Texture is a blanket term used to describe the brightness variations observed in almost all surfaces; it conveys information about a pixel and its surrounding. Textures may be regular, for example a smooth surface with a grid painted on it. Alternatively textures may be highly irregular in some respects, for example salt scattered in a random pattern over a surface. There is however often some defining quality which characterizes a particular texture, and can be used to separate it from other textures.

The basic texture primitive is known as a *texel* (texture element). The texel consists of several pixels, whose placement could be periodic, quasi-periodic or random. The nature of texture is determined by two aspects: structure and tone. Tone is related directly to the intensities of the pixels without regard to their spatial relationships. Structure, on the other hand, is related to the spatial relationship of the pixels. When a small area of the image (3×3 say) has a wide variation in brightness, the dominant property of that area is texture. However, when the small area has little brightness variation, the dominant property of that area is tone.

There are three principal approaches to describing the texture of an image region; they are statistical, structural and spectral [6]. Statistical techniques attempt to describe texels by attributing them with statistical distributions of pixels, while structural techniques use the arrangement of image primitives and shapes to classify and recognize texels. Spectral techniques, based on the Fourier transform, are used mainly to detect global periodicities. The spectral methods are generally regarded as inferior to the statistical ones in terms of terrain classification.

In our study, we used the first-order statistical texture measures: mean, standard deviation and range. Here the range is defined as the difference between the intensities of the brightest pixel and the darkest pixel in a 3×3 neighborhood. Likewise, the mean and standard deviation are computed from the same 3×3 region. The three texture features are computed from the first and second principal component images, see Fig. 3. Although these features do not take into account the spatial relationships of pixels, they are rotation invariant and easy to compute. Higher order statistical texture measures, based on the gray-level co-occurrence matrices, provide information about the spatial relationships among pixels [6].

4. The Classification Scheme

We used a hierarchical classification scheme to determine the ground-cover type of a pixel in the image. First a thresholding process is applied to the intensity values of the first principal component image. If the intensity value is greater than 250, the pixel is classed as clouds. Using this thresholding mechanism, we found only 112 cloud pixels, from a total of 12841, have intensity values less than 250. This means that the accuracy of classifying clouds is 99.12%. If, however, the intensity value is less than 50, then the pixel is said to belong to the classes of sea and shadow over land. This pixel is then passed to a simple classifier which discriminates between the two classes shadow over land and sea. The pixels whose intensity values range between 50 and 250 belong to one of four classes: the three forest types or land. These pixels are then processed by a neural network classifier system to determine to which of the four classes the pixel belongs.

4.1 Sea-Shadow Classifier

The sea-shadow classifier, shown in Fig. 5, is a simple network of *threshold logic units* [10], which consists of two inputs, two hidden units and one output. The inputs x_1 and x_2 are the intensity values of the pixel in the first and second principal component images, respectively. The weights of the network were designed from a training set containing 4000 samples from the class "sea" and 409 samples from the class "shadow-over-land," see Fig. 6.

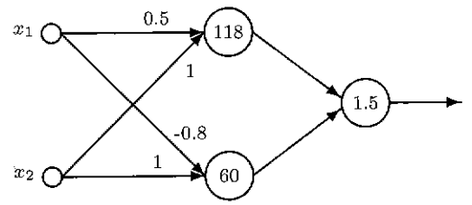

Fig. 5 The sea-shadow classifier consists of a network of threshold logic units. The output of each unit is 1 if the net sum of its weighted inputs is greater than or equal to the threshold (number inside the circle); otherwise, it is zero. If the output is 1, the pixel class is shadow; otherwise, the pixel belongs to the class of sea.

Table 1 Confusion matrix for the training set.

Ground Truth	Class Label		Total No. Samples	Accuracy %
	Sea	Shadow		
Sea	4000	0	4000	100
Shadow	1	408	409	99.76
Total	4001	408	4409	99.98

Table 2 Confusion matrix for the test set.

Ground Truth	Class Label		Total No. Samples	Accuracy %
	Sea	Shadow		
Sea	14284	95	14379	99.34
Shadow	142	11206	11348	98.75
Total	14426	11301	25727	99.08

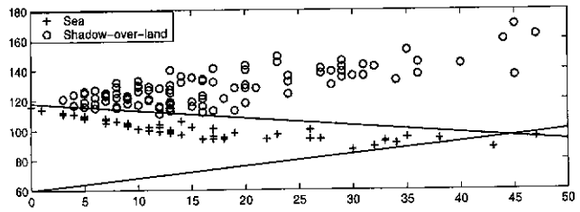

Fig. 6 Training set used to design the sea-shadow classifier. The two lines represent the decision surfaces (lines) of the two hidden units of the network of Fig. 5.

The performances of this simple classifier on the training and test sets are summarized in Tables 1 and 2, respectively. The overall classification accuracy of the sea-shadow classifier is 99.08%.

4.2 Land-Forest Classifier

Classification of land and forest types is significantly more difficult than the sea-shadow classification. The reason is that the forest and land classes are not as easily separable by their intensity values, see Fig. 4. A pixel whose first principal component value is in the range 50 to 250 is considered to be of land or forest types. A neural network classifier is then used to classify the pixel into one of four classes: forest 1, forest 2, forest 3, or land. The neural network classifiers, used to classify land and forest types, are introduced in the next subsection. In the following subsection, committee machines are described, which are used to improve the performance of the land-forest classifier. In the third subsection, the issue of classifier accuracy is discussed.

4.2.1 The Neural Network Classifier

The popular feedforward neural network architecture, known as the *multilayer perceptron* (MLP), was chosen to carry out the land-forest classification task. Before a multilayer perceptron can be used for classification, it must be trained to determine the free parameter values (synaptic weights) which achieve the best performance. Therefore, a training set which contains 683 samples from forest 1, 486 samples from forest 3 and 1000 samples from land was selected. Five attributes were chosen as inputs from the first and second principal component images, Fig. 3; they are the two intensity values and three texture features: mean, standard deviation and range, computed over a 3×3 region. Note that the inputs are first scaled down to the range 0 and 1 before they are applied to the MLP network. The training set was then used to train several multilayer perceptron networks to classify a given pixel into one of four classes: forest 1, forest 2, forest 3, or land. The trained MLPs were tested on a second set containing 683 samples from forest 1, 486 samples from forest 2, 24,965 samples from forest 3 and 24,896 samples from land. All MLPs were trained with *error backpropagation* using the Levenberg-Marquardt algorithm [7]. The best network architecture found has five inputs, eight units in the first hidden layer, four in the second hidden layer, and four outputs.

The error rates of six MLPs, trained with different initial values of the synaptic weights, are shown in Table 3; all six MLPs performed reasonably well. Among the six, MLP6 is deemed the best classifier, i.e., the classifier with the lowest test error; it achieves an error rate of 6.41% on the training set, 8.7% on the test set and an overall error rate of 8.56%. The *confusion* or *error matrix* of this classifier is presented in Table 4.

Table 3 also shows the 0.95 confidence interval on the error rates of the individual MLPs. It is computed using the total error $\hat{E}$ as [5]

$$\left[\hat{E}-1.96\sqrt{\frac{\hat{E}(100-\hat{E})}{N}},\ \hat{E}+1.96\sqrt{\frac{\hat{E}(100-\hat{E})}{N}}\right],$$

where N is the size of the data set. Using a data set of size $N = 54198$ and a total error $\hat{E} = 8.56\,\%$, we can say with 95% confidence that the accuracy of MLP6 is somewhere between 91.2% and 91.67%. Therefore, we

Table 3 Error rates (%) of six MLPs trained using texture features and intensity values.

Neural Network	Training Error	Test Error	Total Error	Confidence Inter.(95%)
MLP1	6.47	9.50	9.33	[9.08, 9.57]
MLP2	5.90	9.63	9.42	[9.17, 9.66]
MLP3	6.60	9.12	8.97	[8.73, 9.22]
MLP4	6.00	8.80	8.64	[8.40, 8.87]
MLP5	6.63	9.11	8.96	[8.72, 9.20]
MLP6	6.41	8.70	8.56	[8.33, 8.80]
Average	6.33	9.14	8.98	[8.74, 9.22]

Table 4 Apparent confusion matrix of MLP6. The " * " indicates the overall accuracy of the network, and FA is the false alarm rate.

Ground Truth	Class Label				Accuracy (%)
	F1	F2	F3	Land	
F1	1363	0	2	0	99.85
F2	0	829	99	44	85.29
F3	182	945	23997	841	92.42
Land	18	1235	1274	23369	90.24
FA (%)	12.80	72.45	5.42	3.65	91.44*

can conclude from Table 3 that the accuracy we can obtain using individual MLPs is about 91%.

4.2.2 Committee Machines

In order to improve on the classification performance of individual MLPs, their outputs can be combined together to form a *committee machine* [1], [7]. The committee machine fuses the knowledge acquired by individual classifiers to arrive at a decision that is supposedly superior to any decision attainable by an individual classifier acting alone. Mathematically, the output of the committee machine can be expressed as

$$\mathbf{z} = f(\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_n), \tag{1}$$

where $\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_n$ are m-dimensional vectors representing the outputs of individual classifiers, f is the combining function, and $\mathbf{z}$ is the committee machine or combiner output, also an m-dimensional vector.

There are a number of ways by which the individual classifier outputs can be combined together to improve the overall performance of the system; they can broadly be categorized into four groups:

- unconditional mixture of experts,
- conditional mixture of experts,
- stacked generalization, and
- boosting.

In this study, we employ the first approach, i.e., unconditional mixture of experts. In this approach, the overall decision is made based solely on the outputs of the individual classifiers, which are combined together in a committee-like manner without any knowledge of the input pattern. By contrast, in a conditional mixture of experts, a *gating network* is employed to combine the individual classifier outputs. The gating network plays a role of mediator among the expert networks. It places on the output of each expert a weight which is determined by a nonlinear function of the input pattern, hence the name conditional mixture of experts.

A neural network classifier can be trained to approximate the *posterior probabilities* of class membership, provided the error function is chosen appropriately [8]. For a 1-of-c target coding scheme, where each class is represented by an output unit, the outputs of a network, trained by minimizing the sum-of-squares error function, can be arranged in such a way that each output unit models the corresponding posterior probability $p(C_k|\mathbf{x})$, where C_k is the kth class, and $\mathbf{x}$ is the input pattern. In this case the outputs of the network must sum to one.

The MLPs presented in the previous section were trained and their outputs normalized so that they can be interpreted as posterior probabilities of class memberships. The aim is to combine the outputs together to obtain better estimates of posterior probabilities, and hence better classification performance. There are numerous ways to do this, but here we considered only five methods:

- **Majority Voting Combiner.** Here each network makes a decision as to which class the input pattern belongs by taking the maximum posterior probability. Then the 0-1 decisions from different networks are summed together, and the class with the highest score is voted the winning class.
- **Maximum Value Combiner.** In this method, the posterior probability of each class is taken as the maximum of all corresponding network outputs.

$$p(C_k|\mathbf{x}) = \max_j (\mathbf{y}_{jk})$$

The class with the maximum posterior probability is declared the winner.

- **Median Value Combiner.** Instead of taking the maximum, as in the previous method, the median value is used as the estimate of the posterior probability.

$$p(C_k|\mathbf{x}) = \text{median}_j (\mathbf{y}_{jk})$$

- **Mean Value Combiner.** Here the *arithmetic mean* of the network outputs is used as the estimate for the posterior probability.

$$p(C_k|\mathbf{x}) = \frac{1}{n}\sum_{j=1}^{n} \mathbf{y}_{jk}$$

This is also known as the *simple ensemble averaging* [1], [7].

- **Weighted Average Combiner.** This method is also known as the *generalized ensemble averaging*. As the name implies, the posterior probability is approximated by a weighted average of the network outputs.

Table 5 Error rates of individual network combiners.

Network Combiner	Overall Error (%)	Confidence Interval (95%)
Majority Voting	8.45	[8.22, 8.68]
Maximum Value	8.26	[8.03, 8.49]
Median Value	8.08	[7.85, 8.31]
Mean Value	8.04	[7.81, 8.27]
Weighted Average	7.91	[7.68, 8.14]

Table 6 Apparent confusion matrix of weighted average network combiner. The " * " indicates the overall accuracy, and FA is the false alarm rate.

Ground Truth	Class Label				Accuracy (%)
	F1	F2	F3	Land	
F1	1361	0	4	0	99.71
F2	2	856	84	30	88.07
F3	185	1007	24087	686	92.77
Land	19	1075	1194	23608	91.16
FA (%)	13.15	70.86	5.05	2.94	92.09*

$$p(C_k|\mathbf{x}) = \frac{1}{n}\sum_{j=1}^{n} \alpha_j \mathbf{y}_{jk}$$

where the weights α_j sum to one. The optimal combination weights, α_j, can be obtained by minimizing the sum-of-squares error of the committee machine subject to $\sum_j \alpha_j = 1$, which is a function of the error correlation matrix [1].

Table 5 presents the average error rates and the 95% confidence intervals of the committee machines described above. Clearly, all the machines achieve an improvement over the individual MLPs, and the improvements are significant. In particular the weighted average committee machine achieves the best performance; its average error rate is 7.91%, compared to 8.56% for the best MLP (MLP6). Upon comparison of the confidence intervals in Tables 3 and 5, we can conclude that the expected worst error rate of the weighted average combiner (8.14%) is lower than the best error rate (8.74%) expected of an MLP. For comparison purposes, the confusion matrix of the weighted average committee machine is presented in Table 6. It shows that classification accuracies are increased and false alarm rates are decreased for all classes, except forest 1 (F1); for this class MLP6 achieved the best performance of any individual or committee machine classifier.

Classification of type 2 forest seems to be the most challenging task; this class has the lowest classification accuracy (88.07%) and the worst false alarm rate (70.86%). The reason is that there exists significant overlap between this class and the two classes of land and forest 3 (see Fig. 4). An attempt to reduce the false alarm rate and improve the classification accuracy of type 2 forest by including texture features from the third principal component didn't help at all. We think the problem is that type 2 forest consists of shrubs that resemble type 3 forest in some regions, but do not provide dense coverage of the ground in other regions, hence the significant overlap between this class and the classes of forest 3 and land. If type 2 and type 3 forests are merged together into a single class F2-3, the classification accuracy will be 96.64% and the false alarm rate will be 8.02%. Therefore, the major difficulty is in distinguishing between type 2 and type 3 forests. A separate classifier was designed to distinguish between these two classes, but the results (not shown here) didn't improve the situation very much. The problem is that type 2 forest has a small number of samples (972 compared to 25,965 for forest 3) and a large portion of them overlaps with samples from type 3 forest. Consequently, the classifier couldn't improve the classification accuracy of forest 2 without increasing the false alarm rate; that is, the only way to improve the classification accuracy of forest 2 is by shifting the decision boundary towards forest 3, which gives a high false alarm rate by misclassifying samples from forest 3 as forest 2.

4.2.3 Classifier Accuracy

The κ coefficient of agreement is usually used to assess the classifier accuracy compared to chance classification. It is a measure of the difference between the actual agreement between reference data and the classifier and the chance agreement between the reference data and a random classifier. The κ statistic is computed as

$$\kappa = \frac{P_o - P_e}{1 - P_e}$$

where P_o is the observed accuracy and P_e is the chance agreement [3], [4]. For a chance agreement, it is expected that $\kappa = 0$, whereas for a true agreement $\kappa = 1$. The value of the κ coefficient computed from the apparent confusion matrices shown in Tables 4 and 6 are, respectively, $\kappa = 0.847$ and $\kappa = 0.856$. Clearly, both classifiers, MLP6 and the weighted average combiner, are much better than a chance classifier.

4.3 The Hierarchical Classifier

After an indepth analysis of the performances of all designed classifiers, the final classification scheme is a hierarchical classifier comprised of a simple thresholding mechanism to identify the cloud and beach classes, a thresholding mechanism followed by a thresholding logic unit to classify sea and shadow-over-land classes, six MLPs followed by the weighted average committee machine to classify the three forest types and land. The final algorithm is summarized below, where x represents the intensity value from the first principal component image.

Algorithm:

If $x > 250$,

pixel belongs to clouds

elseif $x < 50$,

the pixel is from Shadow-over-land or Sea classes, apply the sea-shadow classifier

else $(50 < x < 250)$

the pixel belongs to one of four classes: forest 1, forest 2, forest 3, or land

apply six MLPs and combine the outputs using the weighted average combiner

end

5. Conclusion

In this paper, a hierarchical classification scheme for classifying the ground-cover types of a satellite image was proposed. The image studied contains clouds, sea, shadow on the ground, three types of forest and land. Principal component analysis is first used to reduce the number of classes. The first principal component image can separate ground-cover types into three categories: the first is clouds/beach; the second is sea and shadow-over-land; and the third is the three forest types with land. A classifier, which achieves 99% accuracy, was designed to distinguish between the classes sea and shadow-over-land. The land and forest classes are classified using a committee machine which consists of six MLPs and a weighted average combiner. Several other committee machines were also considered. They were all able to improve on the performance of the individual MLPs, but the best committee machine was the weighted average combiner. This committee machine achieved an overall error rate of 7.91%, whereas the best MLP achieved an error rate of 8.56%.

Acknowledgments

The Satellite image and the ground truth data were provided by the Defence Science and Technology Organisation (DSTO) in Salisbury, SA. Special thanks to Peter Deer and Richard Cole for the initial discussions which lead to this work.

References

[1] C.M. Bishop, Neural Networks for Pattern Recognition, Clarendon Press, Oxford, 1995.
[2] A. Bouzerdoum, R.J. Cole, and P. Deer, "Classification of satellite imagery based on texture features using neural networks," Proc. Fourth Int. Conf. on Control, Automation, Robotics and Vision, pp.2257–2261, Singapore, 3–6 Dec. 1996.
[3] R. Congalton, "A review of assessing the accuracy of classifications of remotely sensed data," Remote Sensing of Environment, vol.37, pp.35–46, 1991.
[4] G.M. Foody, "On the compensation for chance agreement in image classification accuracy assessment," Photogrammetric Eng. & Remote Sensing, vol.58, pp.1459–1460, 1992.
[5] J.E. Freund and R.E. Walpole, Mathematical Statistics (4th ed.), Prentice-Hall, Englewood Cliffs, N.J., 1987.
[6] R.M. Haralick, "Statistical and structural approaches to texture," Proc. IEEE, vol.67, pp.786–804, 1979.
[7] S. Haykin, Neural Networks: A Comprehensive Foundation (2nd ed.), Prentice-Hall, Upper Saddle River, N.J., 1999.
[8] M.D. Richard and R.P. Lippmann, "Neural network classifiers estimate Bayesian a-posteriori probabilities," Neural Computation, vol.3, pp.461-483, 1991.
[9] J.A. Richards, Remote Sensing Digital Image Processing, (2nd ed.), Springer Verlag, Berlin, 1993.
[10] J.M. Zurada, Introduction to Artificial Neural Systems, West Publishing Company, New York, 1992.

Abdesselam Bouzerdoum received the Masters and Ph.D. degrees, both in electrical engineering, from the University of Washington, Seattle, in 1986 and 1991, respectively. From 1991 to 1997 he was with the University of Adelaide, South Australia, first as a Postdoctoral fellow, then as Lecturer (Assistant Professor) and senior lecturer. In 1998 he was appointed Associate Professor at Edith Cowan University. He is currently the Deputy Director of the Centre for Very High Speed Microelectronic System. He is the founder and group leader of the Visual Information Research Group. He has received several distinguished awards, amongst them the Vice Chancellor's Distinguished Researcher Award in 1998 and 1999, and the awards for Excellence in Research Leadership and Excellence in Postgraduate Supervision in 2000. In 2001 he was awarded a Distinguished Researcher Fellowship (Chercheur de Haut Niveau) from the French Ministry of Research to spend three months at the National Research Centre LASS - CNRS in Toulouse. Dr. Bouzerdoum is a member of IEEE and INNS. He served as Chairperson of the IEEE Neural Network Council Region Interest Group (RIG) of the South Australian Section from 1995 to 1997, and he is currently serving as Associate Editor for the IEEE Transactions on Systems, Man and Cybernetics. He authored over 120 technical articles in Journals, book chapters and conference proceedings. His research interests include machine vision and pattern recognition, neural networks and their applications to signal and image processing problems, and VLSI implementation of smart micro-sensors.