We are IntechOpen,
the world's leading publisher of
Open Access books
Built by scientists, for scientists

6,900
Open access books available

186,000
International authors and editors

200M
Downloads

154
Countries delivered to

Our authors are among the

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

Chapter

Land Cover/Land Use Mapping Using Soft Computing Techniques with Optimized Features

Selvaraj Rajesh and Gladima Nisia T.

Abstract

The chapter discusses soft computing techniques for solving complex computational tasks. It highlights some of the soft computing techniques like fuzzy logic, genetic algorithm, artificial neural network, and machine learning. The classification of the remotely sensed images is always a tedious task. So, here we explain how these soft computing techniques could be used for image classification. Image classification mainly concentrates on the feature's extraction process. The features extracted in an efficient manner improve classification accuracy. Hence, the different kinds of features and different methods for these extractions are explained. The best extracted features are selected using genetic algorithm. Various algorithms are shown and comparisons are made. Finally, the results are verified using a hypothetical case study.

Keywords: land cover/land use mapping, soft computing techniques, feature extraction, artificial neural network, wavelet transforms, feature classification

1. Introduction

The remote sensing (RS) image has millions and millions of details hidden into it. The interpretation of RS images thus leads to a variety of new improvements in our daily life. Since RS image coils a lot of areas in a single image, intensive care has to be taken while handling each and every pixel [25, 41, 42, 61]. Also, extraction of features plays an important role. Using those features, a particular pixel can be classified easily [32, 33, 46, 55, 56]. Deciding which features we are going to extract is important, and it has to be done based on the application and type of image.

The classified output has several uses in civil engineering. It is also useful in planning for large airports, industrial estates, and harbors and the construction of dams, bridges, and pipelines. It also provides valuable data for the process and design of roads and highways. The application areas also extend to extracting building footprints, detecting roads, and outlining urban changes from a pair of images taken at different dates. It also extends to the field of forest investigation, water management, and disaster management.

Similarly, the interpretation of RS images has many applications [34]. They include the study of forest where investigating the landscape of forest area can avoid deforestation and degradation processes. Forest land cover describes the physiographical characteristics of the environment from bare rock to tropical forest.

So, classifying these will result in the understanding of the variety and type of land cover. Another important advantage with forest land cover is identification of very specific habitats and distribution of both individual species and species assemblies. In the case of urban planning, the year-wise RS images are analyzed to find whether the occupation is growing in the right place. While planning the urban area utilization, the government may plan with the RS image, so that the road construction plan, water pipeline construction plan, and power supply connection plan can be made easy. If in case our urban occupation is happening in the vegetation area, then it should be taken care of and constructions are to be made in other areas.

RS images are also used in water management system to clearly display sediment pollution and oil spills over water bodies and help to monitor the quality of water resources. They are also used in disaster management. In case of natural disaster, risk-prone areas are detected, and risk management is undertaken. When sudden natural disaster happens, it is difficult for humans to collect data at that moment, and so using RS technology, we can handle the situation.

The application area also covers the hazard management. As water-related natural hazards occur due to a number of factors, such as structure, drainage, slope, land use, road network, etc., they must be taken into account when assessing the region's instability and potential hazard risks. It is essential because proper hazard management can help us take timely measures to prevent flooding and following landslides.

The chapter is organized in the following way. Section 2 explains the feature extraction process, Section 3 explains the feature subset selection, Section 4 explains about feature classification, and Section 5 concludes the chapter.

2. Feature extraction

To classify/segment the different objects in a digital image, features are of much important. Texture feature is one such important feature. Texture is more useful as it is expressed in terms of smoothness, coarseness, fineness, linearization, granularity, and randomness. Analysis of texture requires the identification of features that will differentiate the textures for classification, segmentation, and recognition [17–19, 22, 23, 26, 35–37, 43]. Scale is another important property of texture. The appearance of texture changes when it is viewed at different resolutions. Remotely sensed images are analyzed using gray-level co-occurrence features, features extracted from Gabor filter. There are many methods for extracting features.

2.1 Extraction of features using wavelet packet transform

The main reason for the usage of such wavelet-based multi-resolution analysis [7–10, 12, 27, 29, 30, 39] in remote sensing is that the resolution of the remotely sensed imagery may be different in many cases and it is important to understand how information changes over different scales of imagery.

The work in [1] proposed a system in which statistical and co-occurrence features of the input patterns are first extracted, and those features are used for classification [11–13, 20, 38, 48]. The continuous wavelet transform of a 1-D signal f(x) is defined using Eq. (1):

$$W(a,b) = \int f(x)\Psi_{a,b}(x)dx \qquad (1)$$

where $\Psi_{a,b}(x) = \frac{1}{\sqrt{a}}\Psi\left(\frac{x-a}{b}\right)$.

The mother wavelet y has to satisfy the admissibility criterion to ensure that it is a localized zero mean function [39]. Typically, some more constraints are imposed on y to ensure that the transform is non-redundant and complete and constitutes a multi-resolution representation of the original signal. This results in a good real-space transform implementation using quadrature mirror filters. The convolution is performed, and the results with the low-pass filter are called approximation image, and the results with the high-pass filter in specific directions are called detail images. In earlier processes, the image is split into an approximation and detail images. The approximation is then split itself into a second level of approximation and details. For a n-level, the signal decomposition can be represented using Eq. (3):

$$
\begin{aligned}
A_n &= \left[H_x * \left[H_y * A_{n-1} \right] \downarrow_{2,1} \right] \downarrow_{1,2} \\
D_{n1} &= \left[H_x * \left[G_y * A_{n-1} \right] \downarrow_{2,1} \right] \downarrow_{1,2} \\
D_{n2} &= \left[G_x * \left[H_y * A_{n-1} \right] \downarrow_{2,1} \right] \downarrow_{1,2} \\
D_{n3} &= \left[G_x * \left[G_y * A_{n-1} \right] \downarrow_{2,1} \right] \downarrow_{1,2}
\end{aligned}
\tag{2}
$$

where "*" denotes the convolution operator, "$\downarrow 2,1$" denotes the downsampling along the rows (columns), A0 = I is the original image, and H and G are low-pass and high-pass filters, respectively. I(x, y) is the original image. A_n is obtained by low-pass filtering and is the approximation image at scale n. The detail images D_{ni} are obtained by band-pass filtering in a specific direction (i = 1, 2, 3 for vertical, horizontal, and diagonal directions, respectively) and thus contain directional detail information at scale n. The original image, I, is thus represented by a set of sub-images at several scales: $\{A_n, D_{ni}\}$.

The wavelet packet decomposition offers a richer signal analysis. Here, the split happens for both detail image and approximation image. This results in a wavelet decomposition tree. The details present in detail images are helpful in analyzing texture and discrimination. To characterize a texture, the features derived from detail images are used. The following section discusses the way in which the features from wavelet transformed image to be used for classification.

The filter choice and its order may vary for each application. Here, two levels of wavelet packet decomposition with different wavelet families are done and shown in **Figure 1**. There is no need to perform a deeper decomposition because, after the second level, the size of images become too small and no more valuable information is obtained. Sixteen wavelet coefficient matrices containing texture information are produced from the second level of decomposition.

In texture training, the known texture images are decomposed using DWPD. To create feature database, a set of WPSF, such as mean and standard deviation, is
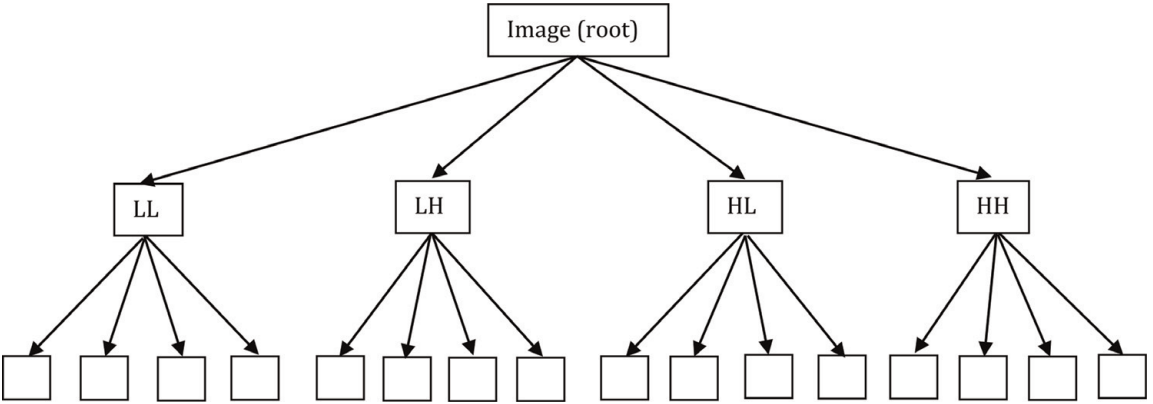

Figure 1.
A wavelet packet tree.

calculated to form the original image, and a set of wavelet packet co-occurrence features and spectral feature NDVI is calculated using Eqs. 4–11 and Eq. (12), respectively. These features are saved for further use in texture classification.

$$Mean \quad \overline{x} = \frac{1}{N^2} \sum_{i=1}^{N} \sum_{j=1}^{N} x_{i,j} \tag{3}$$

$$Variance \quad V = \frac{1}{N^2} \sum_{i=0}^{N} \sum_{j=1}^{N} (x_{i,j} - \overline{x})^2 \tag{4}$$

$$Entropy = -\sum_{i=1}^{N} \sum_{j=1}^{N} (C(i,j)) \log (C(i,j)). \tag{5}$$

$$Contrast = \sum_{i,j=0}^{N} (i-j)^2 C(i,j) \tag{6}$$

$$Energy = \sum_{i=1}^{N} \sum_{i=1}^{N} C(i,j)^2 \tag{7}$$

$$Local\ homogeneity = \sum_{i,j=0}^{n} 1 / \left(1 + (i-j)^2\right) C(i,j) \tag{8}$$

$$Cluster\ shade = \sum_{i,j=0}^{n} \left(i - M_x + j - M_y\right)^3 C(i,j) \tag{9}$$

$$Cluster\ prominence = \sum_{i,j=0}^{n} \left(i - M_x + j - M_y\right)^4 C(i,j) \tag{10}$$

where $M_x = \sum_{i,j=0}^{n} iC(i,j)$ and $M_y = \sum_{i,j=0}^{n} jC(i,j)$.

$$Correlation = \frac{\sum_{i=1}^{N} \sum_{j=1}^{N} [ijC(i,j)] - \mu_x \mu_y}{\sigma_x \sigma_y}. \tag{11}$$

where $\mu_x = \sum_{i}^{N} i \sum_{j}^{N} C(i,j)$, $\mu_y = \sum_{j}^{N} j \sum_{i}^{N} C(i,j)$, $\sigma_x^2 = \sum_{i}^{N} (a - \mu_x)^2 \sum_{j}^{N} C(i,j)$, $\sigma_y^2 = \sum_{j}^{N} \left(b - \mu_y\right)^2 \sum_{i}^{N} C(i,j)$.

$$NDVI = (near\ IR\ band - red\ band)/(near\ IR\ band + red\ band) \tag{12}$$

The input Madurai LISS IV image is shown in **Figure 2**. The procedure for classification is explained in the later content, but the results are presented here for better understanding. The classification of LISS IV Madurai image is done with wavelet filters such as Daubechies (DB2), symlet (Sym2), Coiflet (Coif2), and bi-orthogonal (Bi-or2.2) and is shown in **Figure 3(a)–(e)**.

2.2 Extraction of deep features

Deep feature learning plays an important role in image classification. In order to extract different features automatically, the convolution neural network (CNN) is
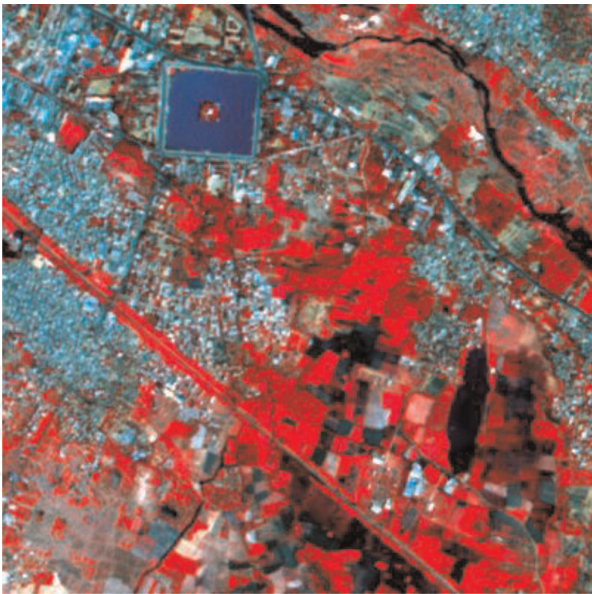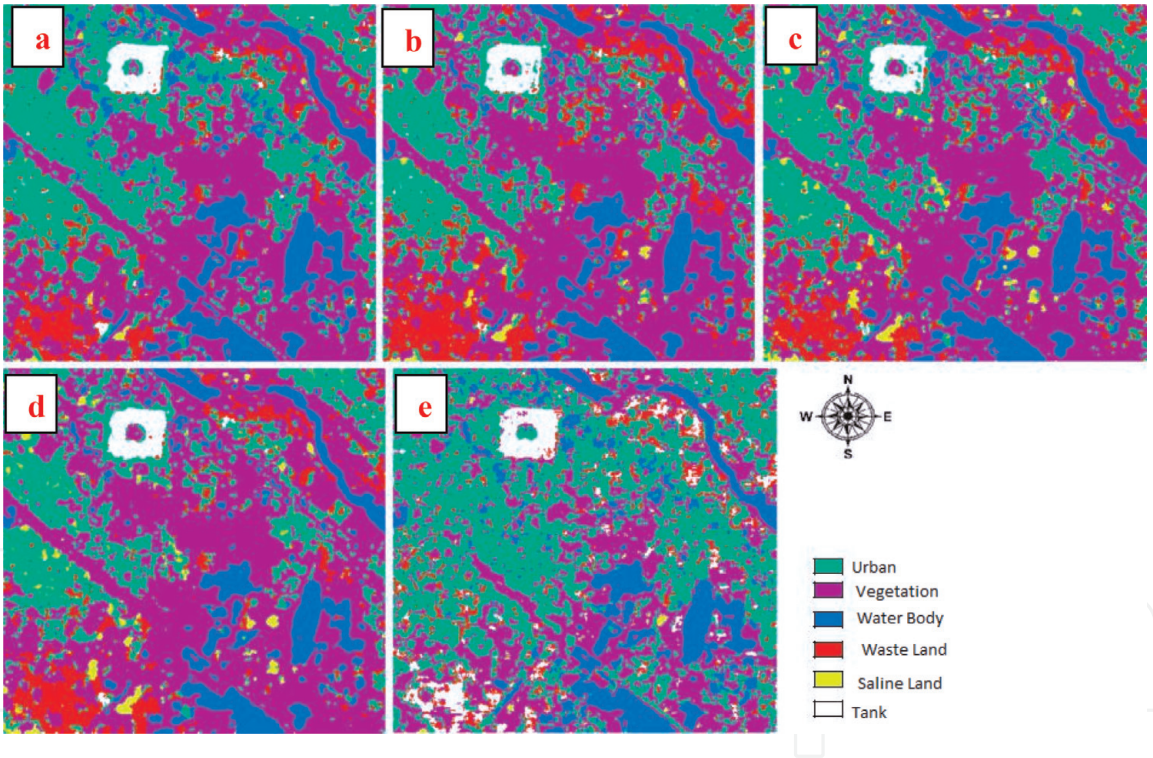
Figure 2.
Madurai city (size 400 × 400).

Figure 3.
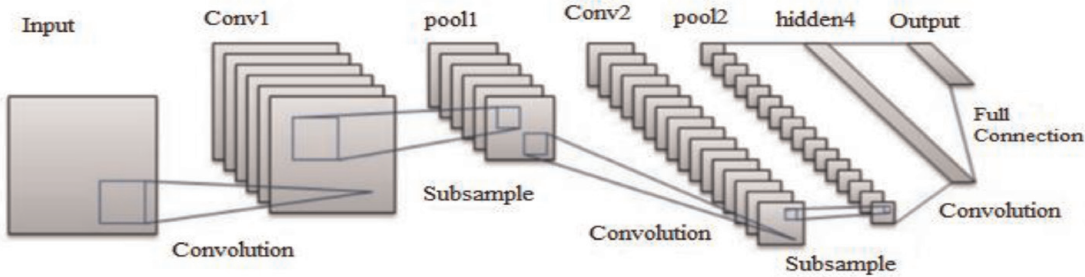*Classified output images using (a) DB2, (b) symlet 2, (c) Coiflet 2, (d) Bi-or 2.2, and (e) DB2 without NDVI.*

Figure 4.
Architecture of convolutional neural network. [source: https://images.app.goo.gl/YcBQH2Y4ZXPyMhVr8].

5

utilized [2]. The architecture of CNN is shown in **Figure 4**. In convolution layer, the features are extracted using different filters to input image. The ReLU layer handles the output from convolutional layer by figuring out the negative pixel value into zero, retaining the dimensionality of the matrix unchanged. Pooling helps in retaining the most important information while reducing the size of feature map. Each training sample is applied with the same processes and thus resulting in different feature sets.

3. Feature subset selection

Feature subset selection is the process of selecting those features that are most useful to a particular classification problem from all those available. The most popular methods for feature reduction in remote sensing are the use of the principal components transform [6]. The principal components (PC) transformation transforms the original data into a new smaller set, which are less correlated to the first data set. Therefore, a reduced number of new variables represent the information content of the original set. However, although frequently used, the PC transform is not appropriate for feature extraction in classification, because it does not consider the classes of interest, but only the data set. Therefore, it may not produce the optimum subspace for the classification. So, we are utilizing genetic algorithm (GA) for feature subset selection [3, 49–54, 57–60, 63–66].

3.1 Genetic algorithms

Computational studies of Darwinian evolution and natural selection have led to numerous models for computer optimization. GAs comprise a subset of these evolution-based optimization techniques focusing on the application of selection, mutation, and recombination to a population of competing problem solutions. Being a directed search rather than an exhaustive search, population members cluster near good solutions; however, the GA's stochastic component does not rule out wildly different solutions, which may turn out to be better. This has the benefit that, given enough time and a well-bounded problem, the algorithm can find a global optimum. This makes them well suited to feature selection problems, and they can find near optimum solutions using little or no prior knowledge.

There are three major design decisions to consider when implementing GA to solve a particular problem. A representation for candidate solutions must be chosen and encoded on the GA chromosome, an objective (fitness) function must be specified to evaluate the quality of each candidate solution, and finally the GA run parameters must be specified, including which genetic operators to use, such as crossover, mutation, selection, and their possibilities of occurrence. Until a satisfactory solution is found, the fitness-dependent selection and application of genetic operators to generate successive generations of individuals are repeated many times.

In the problem of feature selection, feature subsets are represented as binary strings where a value of 1 will represent the inclusion of a particular feature in the training process and a value of 1 will represent its absence. Since a chromosome is represented through a binary string, genetic algorithm will operate on a pool of binary strings. The mutation and crossover operators operate in the following way: mutation operates on a single string and generally changes a bit at random. Thus, a string 10,010 may, as a consequence of random mutation, get changed to 10,110. Crossover on two parent strings produces two offsprings. With a randomly chosen crossover position 2, the two strings 01101 and 11,000 yield the offspring 01000
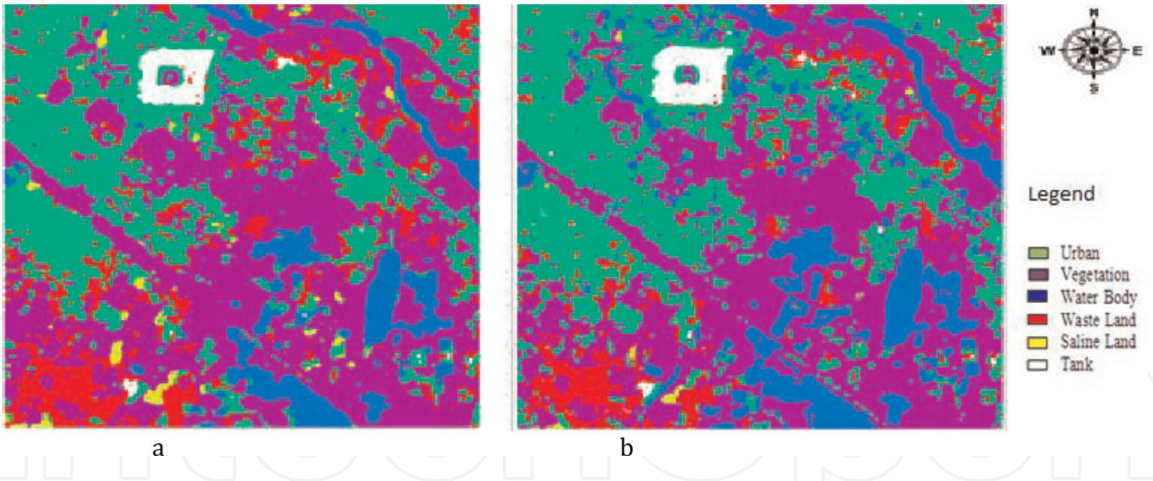
Figure 5.
Classified output using DB2 with (a) full feature set and (b) optimal feature set.

Number of features	Accuracy indices			
	Overall	Kappa	Producer	User
10	84.2437	0.7717	81.8685	78.8469
13	84.534	0.7802	82.2341	79.8645
15	85.5042	0.7898	83.1250	80.5573
16	85.2941	0.7875	82.8616	80.1797
17	**86.528**	**0.8125**	**84.5967**	**81.8991**
18	85.2941	0.7855	81.7592	79.8890
19	84.6639	0.7772	79.9678	79.6327
21	85.5042	0.7900	81.0822	80.2417
23	85.0840	0.7855	82.3767	80.3079
25	85.9244	0.7958	82.0482	80.2008

Bold: 17 Features is giving Max Overall Kappa Producer and User's Accuracy.

Table 1.
Accuracy indices for various feature sets.

and 11,100 as a result of crossover. If the obtained feature set X using wavelet-based technique contains 45 features for each pixel of the image of size 400×400 pixels, then the feature set X of the data is of dimension $160{,}000 \times 45$, where each column represents the features of the respective pixel in the data. Using GA, the feature set X of size $45 \times 400 \times 400$ is mapped into new feature set denoted by Y of size $17 \times 400 \times 400$. This reduction in feature set improves the overall execution speed and the classification accuracy [52]. The classification results (for both full feature set and optimal feature set) are shown in **Figure 5**.

The accuracy assessments are made using accuracy indices, namely, overall accuracy, producer accuracy, user accuracy, and kappa coefficient and are listed in **Table 1**.

4. Classification

Using the features obtained, so far the classification is done using the obtained features. We use different classifiers for the classification. The classifier is an algorithm that maps the input data to a specified category.

4.1 Classification based on Mahalanobis distance

In this method, the decomposition for test texture image is done using DWPD. In the same manner, another set of features are obtained and compared with the obtained feature values. The class of textures is represented as C, the mean signature of class C is represented with m_c, and the Mahalanobis distance is given by

$$D^2(x_{i,} C) = (x_i - m_c) \sum (x_i - m_c) \qquad (13)$$

If the distance D(i) is smallest, then the test texture image is classified as ith texture [15, 21, 47]. Features are obtained using many wavelet filters, and it is followed using classification process [14, 28, 44]. The overall, user, producer, and kappa accuracy indices obtained for the different wavelet filters presented in **Table 2** show that the DB2 wavelet filter gives superior results than the other wavelet filters. Thus, the DB2 wavelet filter will be more useful for land cover/land use mapping.

4.2 Classification based on adaptive neuro-fuzzy inference system (ANFIS)

The adaptive network-based fuzzy inference system (ANFIS) is a useful neural network approach for the solution of function approximation problems [4, 31, 40, 45, 62]. To determine the optimal distribution of membership functions, the ANFIS gives the mapping relation between the input and output data. ANFIS architecture consists of both artificial neural network (ANN) and fuzzy logic (FL). The system includes five layers. The node function describes several nodes, which are to be included in the ANFIS layer. The inputs of present layers are obtained from the previous layers. For example, consider two inputs (x, y) and one output (fi) are used in this system. The rule base contains fuzzy if-then rules. Thus, the two rules are:

- Rule 1: If x is A1 and y is B1, then z is f1(x, y).

- Rule 2: If x is A2 and y is B2, then z is f2(x, y).

where x and y are the inputs and A and B are the fuzzy sets fi (x, y). The feature extraction is done using DB2 wavelet filter, and the optimum features are obtained using GA [16]. Then the classification is done using GA with neural network and GA with ANFIS, and the results are shown in **Figure 6**. Based on the classified output, it is clearly understood that the GA and ANFIS shows the better classification.

4.3 Classification using CNN

The classification using CNN is done using the deep features obtained from the training phase of CNN [2, 5, 24]. In training, it carries out the predefined process with one or multiple layers. In a fully connected layer, every neuron is connected to

	Overall	User	Producer	Kappa
DB2	87.60	82.02	89.57	0.82
Symlet 2	78.78	76.45	76.43	0.69
Coiflet 2	77.3	73.3	70.6	0.67
Bi-or 2.2	79.2	79.76	76.69	0.69

Table 2.
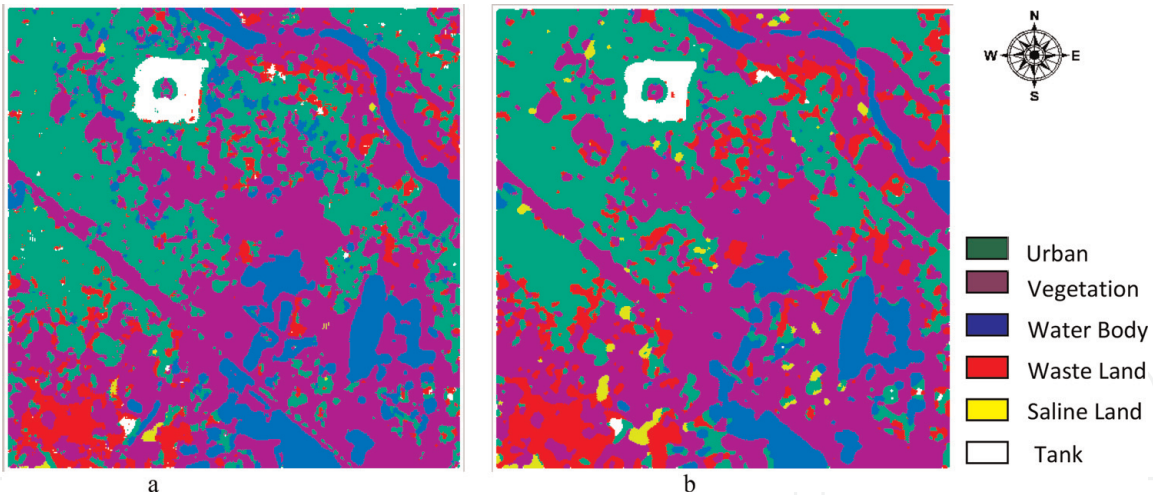*Classification results of Madurai city for different wavelet packet transforms.*

Figure 6.
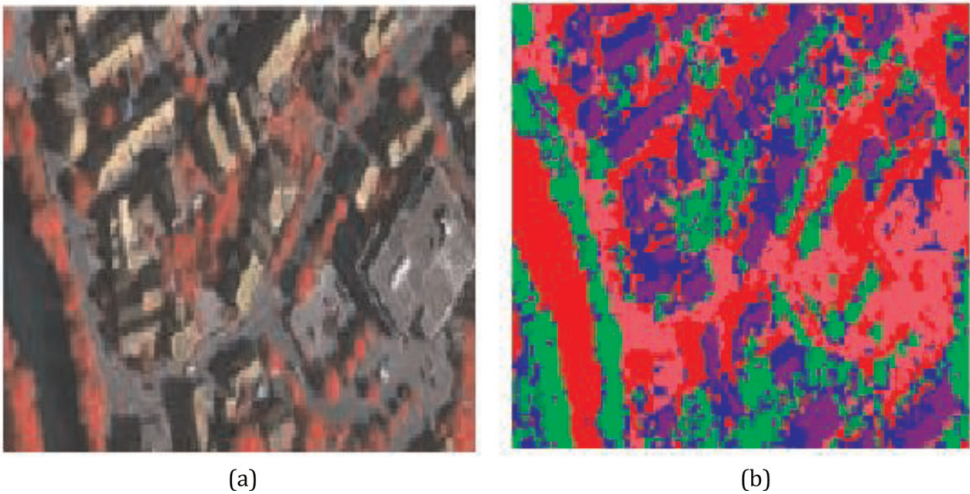*Classified output using DB2 with (a) GA and neural network and (b) GA and ANFIS.*

Figure 7.
(a) Vaihingen city and (b) classified output of Vaihingen city.

every other neurons of previous layer. Softmax is the final layer and it calculates the probability value. The higher probability becomes the output. After training, the system will be able to classify the image automatically without human intervention. The classification is done for Vaihingen city and the results are displayed in **Figure 7**.

4.4 Classification using multilayer perceptron layer

The multilayer perceptron (MLP) layer realizes intelligent classification using features from the wavelet layer. The training parameters of the MLP are shown in **Table 3**. These parameters were selected to give best performance, after several experiments, such as the number of hidden layers, size of the hidden layers, value of the moment constant and learning rate, and type of the activation functions.

4.5 Limitations of different methodologies

4.5.1 Genetic algorithm

In GA, the selection of wrong fitness value may affect the solution of the problem. Other parameters like population size, mutation and crossover also plays

Architecture	
The number of layers	4
The number of neuron on the layers	
Input	17
Hidden1	25
Hidden2	25
Output	6
The initial weights and biases	Random
Activation functions	Tangent sigmoid
Training parameters	
Learning rule	Back propagation + Levenberg-Marquardt
Learning rate	0.01
Momentum constant	0.8
Mean-squared error	1e_07

Table 3.
MLP architecture and training parameters.

an important role in providing solution. GA belongs to a non-deterministic class of algorithms. The optimal solution we get from GA may vary each time we run our algorithm for the very same input data.

4.5.2 Convolutional neural network

CNN requires a lot of training. Also, it requires a lot of data sets for training. A convolution is always a slower operation. Deeper the network, the longer is it's processing time.

4.5.3 ANFIS

Defining the membership function remains a difficult task.

5. Field survey

The results of the entire work are verified with the help of the ground truth. Ground truth is the process done onsite, in which a "pixel" on a satellite image is compared to what is there in reality. It is done in order to verify the contents of the "pixel" on the image. For an image of 400 × 400 size, we have taken 500 points as ground truth data. By performing field visit, these data are collected. The outcome of each method is verified with those points.

6. A hypothetical case study

A hypothetical case study is presented to show the application of land cover/land use mapping in real-life scenario. Assume the XXX company wants to plan their production center construction in Madurai city. For the production centre to be established, large area is needed and thus unoccupied areas in the city have to be

investigated. Also, it sends out waste material that is toxic and should not be present in the urban areas. The products, which the company produces, are sent to other parts of the country and some are exported. So, road routes also have to be checked.

So, initially a place has to be selected and a plan to be made accordingly. In order to plan the construction, it acquires the satellite image of Madurai city. Then the features are obtained using wavelet feature extraction method, and the classified output is obtained using adaptive neuro-fuzzy inference system classification. The classified image can be clearly understood and can be given to the construction planning team for their further processing. Here, also in addition if the PAN image of the Madurai city and MS image of the Madurai city are fused and then if classification is performed, it would yield still better classification results.

7. Conclusion

The chapter focused on the methods used to obtain the perfect classification of the RS images. It discusses various methods used for feature extraction. Different feature extraction methods are discussed. After feature extraction, the number of obtained features is reduced using the feature subset selection methods. The best features are considered and the features which contribute less are neglected. The optimal features are thus taken into account for the classification process. The classification also discusses different techniques through which efficient results are obtained. The methods are implemented using LISS IV image of Madurai city. The classified outputs are shown wherever necessary, and accuracy assessments are also calculated. Thus, the chapter gives overall idea for handling RS image using optimal features.

Author details

Selvaraj Rajesh[1]* and Gladima Nisia T.[2]

1 Department of Information Technology, Mepco Schlenk Engineering, Sivakasi, India

2 Department of Computer Science Engineering, AAA College of Engineering and Technology, Sivakasi, India

*Address all correspondence to: srajesh@mepcoeng.ac.in

IntechOpen

References

[1] Rajesh S, Arivazhagan S, Pratheep Moses K, Abisekaraj R. Land cover/land use mapping using different wavelet packet transforms for LISS IV Madurai imagery. Journal of the Indian Society of Remote Sensing. Jun. 2012;**40**(2): 313-324

[2] LeCun Y, Bengio Y, Hinton G. Deep learning. International Journal of Science. May 2015;**521**:436-444

[3] Rajesh S, Arivazhagan S, Pratheep Moses K, Abisekaraj R. Genetic algorithm based feature subset selection for land cover/land use mapping using wavelet packet transforms. Journal of the Indian Society of Remote Sensing. Jul. 2013;**41**(2):237-248

[4] Rajesh S, Arivazhagan S, Pratheep Moses K, Abisekaraj R. ANFIS based land cover/land use mapping of LISS IV imagery using optimized wavelet packet features. Journal of the Indian Society of Remote Sensing. Jul. 2014;**42**(2): 267-277

[5] Priya KP, Rajesh S, Nisia TG. Object-based convolutional neural network for remote sensed imagery classification. International Journal of Pure and Applied Mathematics. 2018;**119**(12): 13837-13844

[6] Du H, Qi H. An FPGA implementation of parallel ICA for dimensionality reduction in hyper spectral images. In: Proceedings IEEE International Geosciences and Remote Sensing Symposium (IGARSS). Vol. 5. Sep. 2004. pp. 3257-3260

[7] Rajesh S, Arivazhagan S. Land cover/land use mapping using different wavelet packet transforms for LISS IV imagery. In: Proceedings of the IEEE International Conference on Computer, Communication & Electrical Technology (ICCCET). Vol. **2011**. 2011. pp. 103-108

[8] Acharyya M, Kundu MK. Adaptive basis selection for multi texture segmentation by m-band wavelet packet frame. In: Proceedings of the 2001 International Conference on Image Processing. 2001. pp. 622-625

[9] Acharyya M, Kundu MK. Wavelet-based texture segmentation of remotely sensed images. In: Proceedings of the 2001 International Conference on Image Analysis and Processing. 2001. pp. 69-74

[10] Acharyya M, Kundu MK. Image segmentation using wavelet packet frames and Neuro-fuzzy tools. International Journal of Computational Cognition. 2007;**5**(4):27-43

[11] Acharyya M, De RK, Kundu MK. Extraction of features using m-band wavelet packet frame and their neuro-fuzzy feature evaluation for multi texture segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2003;**25**(12): 1639-1644

[12] Acharyya M, De RK, Kundu MK. Segmentation of remotely sensed images using wavelet features and their evaluation in soft computing framework. IEEE Transactions on Geoscience and Remote Sensing. 2003; **41**(12):2900-2905

[13] Alagu Raja S, Anand V, Maithani S, Senthil Kumar A. Wavelet frame based feature extraction technique for improving classification accuracy. Journal of the Indian Society of Remote Sensing. 2009;**37**:423-431

[14] Arivazhagan S, Ganesan L. Texture classification using wavelet transform. Pattern Recognition Letters. 2003;**24**: 1513-1521

[15] Arivazhagan S, Ganesan L. Performance analysis of texture classification techniques using MRMRF

and WSFS & WCFS. In: Proceedings of the Sixth International Conference on Computational Intelligence and Multimedia Applications (ICCIMA'05). 2005

[16] Bandyopadhyay S, Pal SK. Pixel classification using variable string genetic algorithms with chromosome differentiation. IEEE Transactions on Geo science and Remote Sensing. 2001; **39**(2):303-308

[17] Baraldi A, Parmiggiani F. An investigation of the textural characteristics associated with gray level cooccurrence matrix statistical parameters. IEEE Transaction on Geoscience and Remote Sensing. 1995; **33**(2):293-304

[18] Chen C, Chen DC. Multiresolution Gabor filters in texture analysis. Pattern Recognition Letters. 1996;**17**(10): 1069-1076

[19] Du LJ. Texture segmentation of SAR images using localized spatial filtering. Proceedings of the International Geoscience and Remote Sensing Symposium. 1990:1983-1986

[20] Fukuda S, Hirosawa H. A wavelet-based texture feature set applied to classification of multi frequency polarimetric SAR images. IEEE Transactions on Geoscience and Remote Sensing. 1999;**7**(5):2282-2286

[21] Fung T. An assessment of TM imagery for land-cover change detection. IEEE Transactions on Geoscience and Remote Sensing. 1990; **28**(4):681-684

[22] Haralick RM. Statistical and structural approaches to texture. Proceedings of the IEEE. 1979;**67**(5): 786-804

[23] Haralick RM, Shanmugam K, Dinstein I. Textural features for image classification. IEEE Transactions on Systems, Man, and Cybemetics. 1973; **SMC-3**(6):610-621

[24] Shah H, Mitra SK, Banerjee A. Information slicing: An application to object classification in satellite images. CVGIP: Image Understanding. 2008; **57**(3):458-465

[25] Knipling EB. Physical and physiological basis for the reflectance of visible and near-infrared radiation from vegetation. Remote Sensing Environment. 1970;**1**:155-159

[26] Lillesand MT, Ralph Kiefer W, Jonathan Chipman W. Remote Sensing and Image Interpretation. 5th ed. Wiley International Edition; 2004. pp. 586-592

[27] Lindsay RW, Percival DB, Rothrock DA. The discrete wavelet transform and the scale analysis of the surface properties of sea ice. IEEE Transactions on Geoscience and Remote Sensing. 1996;**34**(3):771-787

[28] Mahalanobis PC. On the generalised distance in statistics. 1936. Available from: http://www.insa.ac.in/insa_pdf/20005b8c_49.pdf

[29] Mallat S. A theory for multi resolution signal decomposition: The wavelet representation. IEEE Transactions on Pattern Analysis and Machine Intelligence. 1989;**11**(7): 674-693

[30] Niedermeier A, Romaneesen E, Lehner S. Detection of coastline SAR images using wavelet methods. IEEE Transactions on Geoscience and Remote Sensing. 2000;**38**(5):2270-2281

[31] Pal SK, Ghosh A, Uma Shankar B. Segmentation with remotely sensed images with fuzzy thresholding, and quantitative evaluation. International Journal of Remote Sensing. 2000;**21**(11): 2269-2300

[32] Perumal K, Bhaskaran R. SVM-based effective land use classification system for multispectral remote sensing images. International Journal of Computer Science and Information Security. 2009;**6**:2

[33] Perumal K, Bhaskaran R. Supervised classification performance of multispectral images. Journal of Computing. 2010;**2**:124-129

[34] Quackenbush LJ, Hopkins PF, Kinn GJ. Developing forestry products from high resolution digital aerial imagery. Photogrammetric Engineering and Remote Sensing. 2000;**66**(11): 1337-1346

[35] Reed TR, Du Buf JMH. A review of recent texture segmentation and feature extraction techniques. CVGIP: Image Understanding. 1993;**57**(3):359-372

[36] Reed TR, Wechsler H. Segmentation of textured images and gestalt organization using spatial/spatial frequency representation. IEEE Transactions on Pattern Analysis and Machine Intelligence. 1990;**12**:1-12

[37] Rignot E, Kwok R. Extraction of textural features in SAR images: Statistical model and sensitivity. In: Proceedings of IEEE Geoscience and Remote Sensing Symposium. 1990

[38] Sklansky J. Image segmentation and feature extraction. IEEE Transactions on System Man and Cybernetics. 1978;**8**: 237-247

[39] Soman KP, Ramachandran KI. Insight into Wavelets from Theory to Practice. 2nd ed. Prentice-Hall of India Pvt. Ltd.; 2005. p. 44

[40] Thitimajshima P. Multi resolution fuzzy clustering for SAR image segmentation. In: Proceedings of the International Geoscience and Remote Sensing Symposium IGARSS'99; vol. 5. 1999. pp. 2507-2509

[41] Tucker CJ. Red and photographic infrared linear combinations for monitoring vegetation. Remote Sensing of Environment. 1979;**8**:127-150

[42] Tucker CJ, Sellers C. Satellite remote sensing of primary production. International Journal of Remote Sensing. 1986;**7**:1395-1416

[43] Unser M, Eden M. Multi resolution feature extraction and selection for texture segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence. 1989;**11**:717-728

[44] Unser M. Texture classification and segmentation using wavelet frames. IEEE Transactions on Image Processing. 1995;**4**(11):1549-1560

[45] Blum AL, Langley P. Selection of relevant features and examples in machine learning. Artificial Intelligence. 1997;**97**:245-271

[46] Bittencourt HR, Clarke RT. Feature selection by using classification and regression trees (CART). International Archives of Photogrammetry Remote Sensing and Spatial Information Sciences: Natural Resources Canada. 2004;**35**:66-70

[47] Bruske J, Merényi E. Estimating the intrinsic dimensionality of hyper spectral images. In: ESANN. 1999. pp. 105-110

[48] Bon FJ. Selection y Extraction de Characteristics. 2005. Available at: h ttp://www.etsi2.ugr.es/depar/ccia/rf/ www/tema5_00-01_www/node/node 16.html

[49] Huang C-L, Wang C-J. A GA-based feature selection and parameters optimization for support vector machines. Expert Systems with Applications. 2006;**31**:231-240

[50] Du H, Qi H. An FPGA implementation of parallel ICA for

dimensionality reduction in hyper spectral images. In: Proceedings of the IEEE International Geosciences and Remote Sensing Symposium (IGARSS), vol. 5. 2004. pp. 3257-3260

[51] Van Coillie FMB, Verbeke LPC, De Wulf RR. Feature selection by genetic algorithms in object-based classification of IKONOS imagery for forest mapping in Flanders, Belgium. Remote Sensing of Environment. 2007;**110**(4):76-487

[52] Foody GM, Arora MK. An evaluation of some factors affecting the accuracy of classification by an artificial network. International Journal of Remote Sensing. 1997;**18**(4):799-810

[53] Goldberg DE. Genetic Algorithms in Search Optimization and Machine Learning. Addison Wesley; 1989. p. 41

[54] Uğuz H. A two-stage feature selection method for text categorization by using information gain, principal component analysis and genetic algorithm. Knowledge-Based Systems. 2011;**24**(7):1024-1032

[55] Kavzoglu T, Mather PM. The use of back propagating artificial neural networks in land cover classification. International Journal of Remote Sensing. 2003;**24**(23):4907-4938

[56] Kumar S, Ghosh J, Crawford MM. Best-bases feature extraction algorithms for classification of hyper spectral data. IEEE Transactions on Geo science and Remote Sensing. 2001;**39**(7)

[57] Lennon G, Mercier M, Mouchot C, Hubert-Moy L. Independent component analysis as a tool for the dimensionality reduction and the representation of hyper spectral images. In: Presented at IGARSS 2001. Australia: Sydney; 2001

[58] Lin H, Bruce LM. Parametric projection pursuit for dimensionality reduction of hyper spectral data. Proceedings of the IEEE International Geoscience and Remote Sensing Symposium (IGARSS). 2003;**6**:3483-3485

[59] Mather PM. Computer Processing of Remotely-Sensed Images: An Introduction. 3rd ed. Chichester: John Wiley and Sons; 2004

[60] Tseng M-H, Chen S-J, Hwang G-H, Shen M-Y. A genetic algorithm rule-based approach for land-cover classification. ISPRS Journal of Photogrammetry and Remote Sensing. 2008;**63**(2):202-212

[61] Pal M. Support vector machine-based feature selection for land cover classification: A case study with DAIS hyper spectral data. International Journal of Remote Sensing. 2006;**27**:2877-2894

[62] Roger Jang JS, Sun CT. Functional equivalence between radial basis function networks and fuzzy inference systems. IEEE Transactions on Neural Network. 1993;**4**:156-159

[63] Serpico SB, D'inca M, Melgani F, Moser G. A comparison of feature reduction techniques for classification of hyper spectral remote-sensing data. Transactions of Image and Signal Processing of Remote Sensing. 2002;**8**:4885

[64] Wang J, Chang C. Independent component analysis-based dimensionality reduction with applications in hyper spectral image analysis. IEEE Transactions on Geo science and Remote Sensing. 2006;**44**(6)

[65] Zhang Y, Desai MD, Zhang J, Jin M. Adaptive subspace decomposition for hyper spectral data dimensionality reduction. In: IEEE International Conference on Image Processing, vol. 2. 1999. pp. 326-329

[66] Liu Z, Liu A, Wang C, Niu Z. Evolving neural network using real coded genetic algorithm (GA) for multispectral image classification. International Journal of Future Generation Computer Systems. 2004; **20**:1119-1129