Journal of Marine Science and Technology

Volume 21 | Issue 4 Article 9

THE APPLICATION OF THE GENETIC ADAPTIVE NEURAL NETWORK IN LANDSLIDE DISASTER ASSESSMENT

Jing-Wen Chen
Department of Civil Engineering, National Cheng Kung University, Tainan, Taiwan, R.O.C., geochen@mail.ncku.edu.tw

Yung-Sheng Chue
Department of Civil Engineering, National Cheng Kung University, Tainan, Taiwan, R.O.C.

Yie-Ruey Chen
Department of Land Management and Development, Chang Jung Christian University, Tainan, Taiwan, R.O.C.

Recommended Citation

Chen, Jing-Wen; Chue, Yung-Sheng; and Chen, Yie-Ruey (2013) "THE APPLICATION OF THE GENETIC ADAPTIVE NEURAL NETWORK IN LANDSLIDE DISASTER ASSESSMENT," *Journal of Marine Science and Technology*: Vol. 21: Iss. 4, Article 9.
DOI: 10.6119/JMST-012-0709-2
Available at: https://jmstt.ntou.edu.tw/journal/vol21/iss4/9

Journal of Marine Science and Technology, Vol. 21, No. 4, pp. 442-452 (2013)
DOI: 10.6119/JMST-012-0709-2

THE APPLICATION OF THE GENETIC ADAPTIVE NEURAL NETWORK IN LANDSLIDE DISASTER ASSESSMENT

Jing-Wen Chen[1], Yung-Sheng Chue[1], and Yie-Ruey Chen[2]

Key words: landslide, Genetic Algorithms, Artificial Neural Network, Geographic Information System.

ABSTRACT

This study applied the Genetic Adaptive Neural Network (GANN) structure to satellite image classification and the assessment of landslide disaster. First, the study conducted quantitative analysis of the various factors of slope development and natural environmental hazards in some parts of the catchment areas of the Laonong River in Southern Taiwan. Meanwhile, using the weighting ratios of various disaster causing factors inferred from the best structure of GANN, this study explored the degree of slope land disturbance. Then, this study incorporated the relationship between rainfall and landslides to draw a landslide potential map using the discriminant analysis approach combined with the GIS platform. The findings of this research will be a valuable reference in the follow-up drafting of slope development and treatment policies, and the academic and engineering assessment of landslide disasters caused by slope development.

I. INTRODUCTION

In terms of plate tectonics, Taiwan is a major island formed by the collision of the Philippine Sea Plate and the Eurasian Plate. Due to plate movement, Taiwan is young in geologic age with many faults and earthquakes. Along with broken rocks, the terrain of Taiwan features rolling hills and mountains. Taiwan is located on the west side of the Pacific Ocean close to the Tropic of Cancer and is prone to torrential rain brought by typhoons in summer and autumn. In area of high altitude and narrow space, landslides and derived debris flows are common, resulting in debris. In recent years, many scholars have investigated and discussed hazard factors relating to landslide disasters in [8, 23, 25]. It can be been that each area prone to landslide has specific cause factors, often a number of possible influential factors. The influential factors can be roughly divided into potential factors (e.g., factors of geology, geological structure, topography, hydrology, etc.) and incentives (e.g., rainfall, earthquake, human developmental factors, etc.). With economic development and social changes, the land use types are increasingly complex, and the damage to the natural environment and misuse of land resources are becoming increasingly serious. Previous studies [4, 22] have conducted the correlation analysis of land use types and elevation, gradient, slope aspect, etc. Although changes in land use type have accumulated over a long time, the possible use of machinery for development or the damage to the local natural environment may rapidly change the environment. In recent years, the interpretations of large-area landslide in [12, 24] often use aerial photographs or satellite images to determine landslide damage and assess risks. Since the data acquisition period of satellite images is short, the land surface changes can be rapidly identified. Due to features such as wide image data coverage and low cost, in particular in mountainous areas or inaccessible areas, with the help of computer analysis and the GIS (Geographic Information System), Liu [20] was able to quickly determine land features. Satellite images are a tool for timely and large scale monitoring changes in land use. In recent years, several scholars [1, 17, 19] have applied the ANN (Artificial Neural Network) of AI (Artificial Intelligence) in landslide potential analysis. Some studies in [1, 16, 32] have explored the applications of ANN in satellite image classification, and landslide area judgment. Based on the above, the application of ANN in satellite image classification, and landslide area judgment has advantages and is economic as it can extract more information and it has capabilities of learning, integration and interaction. As a kind of AI algorithm, GA was proposed by Holland [14]. It is an optimized search mechanism constructed by computer simulation of biological "natural selection" and "genetic law" in [11]. GAs search desired parameters by natural evolution processes including reproduction, mating and mutation. The difference between GAs and traditional numerical search

Paper submitted 11/03/11; revised 03/21/12; accepted 07/09/12. Author for correspondence: Jing-Wen Chen (e-mail: geochen@mail.ncku.edu.tw).
[1] *Department of Civil Engineering, National Cheng Kung University, Tainan, Taiwan, R.O.C.*
[2] *Department of Land Management and Development, Chang Jung Christian University, Tainan, Taiwan, R.O.C.*

methods lies in the fact that GAs convert possible solutions into chromosomes by encoding techniques and directly search in the appropriate search space. Meanwhile, with the help of the specific search operators, GAs exchange information during the search process, and thus, have the function of searching for the global optimal solution. The Jianshan Area in the Chiahsien District of Kaohsiung, Taiwan provided the scope for research, and this study applied the GANN (Genetic Adaptive Neural Network) in the determination of the satellite image classifications and quantitatively explored the factors of influence regarding rainfall-triggered landslide disasters. In addition, by discussing the relationship between slope development, rainfall and landslides, this study employed the discriminant analysis integrated with the landslide potential maps on the GIS platform to establish a set of assessment procedures with incorporate the impacts of influential factors of slope development on landslide triggered by rainfall in Southern Taiwan.

II. RESEARCH METHOD

1. ANN

ANN (Artificial Neural Network) attempts to simulate the information system of biological brain cell signal processing and the computing units are simple processors similar to biological neurons. The ANN system generally has three layers: input layer, hidden layer, output layer. The system receives external signals via the input layer and transmits the signals into the network for processing by the interconnections between nodes in the numerical way. The network can also store the signals. The final processing results are output by the output layer. As a type of ANN, BPN (Back Propagation Neural Network) is a feed-forwarded ANN, as each layer of it only receives the output data of the previous layer, as the new input data. Its hidden layer contains a few artificial neurons and the layers are connected by weights and bias. Based on the minimal squared error (MSE), the network learning method applies the supervised learning approach. The function of BPN is to multiply the input values with the total addition of weights to get output values by the transfer function. Then, it compares the output values and target values, and modifies the weighted vales to minimize the error functions by using the algorithm. The function of BPN activation function is to convert the range of output values superposed by the input information. This study used the Tansig Function as the activation function and the set normalized output values in the range of -1~1. The distribution of the types of Tansig activation function is as shown in Fig. 1.

2. GAs

When finding solutions to the optimized problems, GAs (Genetic Algorithms) generally set up a string program by encoding related to the search parameters for the problem to simulate the chromosome composed of genes and generate a certain number of initial species by random mechanism.

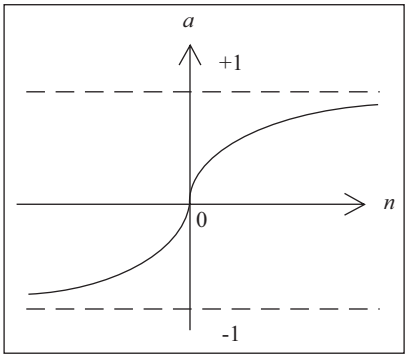

Fig. 1. Tansig function.

Then, according to the solution-finding conditions, GAs designed the fitness function, which represents the competitive advantages of the genes in the survival contest. The value of fitness function determines whether the species has been selected into the mating pool. Next, according to the operations of the mating and mutation process in [11], a generation of genetic computation has been completed. Repeating the computation process of the "survival of the fittest" until conditions of fitness function are satisfied, the final solution is the best parameter solution. GAs operation parameters include population size, selection rate, mating rate, mutation rate and fitness function. Each chromosome consists of a couple of genes representing groups of possible solutions, and the number of chromosomes in a genetic algorithm operating process is known as population size. The mating and mutation in the genetic algorithm process can result in variability of chromosomes as well as loss of excellent combinations of genes. Setting the selection rate by reproducing the program to select excellent gene groups can preserve the genes in the next generation. During the mating process, in a given p(cross), if the random number is selected once in case of chromosome groups of two, the random number will be smaller than p(cross), and the group of chromosomes will exchange genes, otherwise the group of chromosomes will not exchange genes. In the calculation process, the purpose of mutation is to avoid the falling of search process into local optimal solution by increasing variables to avoid limitations on the search. The fitness function of this study is the objective function, which can reflect the fitness of various chromosomes. The value is the basis for the elimination of chromosomes of poor adaptability. When the fitness function value gets to the goal and there is no longer any progress in adaptability by generation, the search is terminated after a certain number of calculations.

3. Genetic Adaptive Neural Network (GANN)

In general, the common problems of ANN in application include the local minimum value rather than the overall minimum value, inadequate training or overtraining and non-convergence etc. To improve the possible disadvantages of ANN, this study optimized parameters by using GA-assisted

ANN. The basic theory is to set the network structural parameters, including weighting matrix of ANN as the chromosomes of GAs, namely, the desired target solutions of the searching process. This study used the difference between the network prediction value and actual value as the fitness function, that is, using MSE (Mean Squared Error) as the benchmark for the assessment of the fitness of chromosomes. The actual value refers to the classification type of the factor to be interpreted; the network prediction value is the output classification prediction value of the spectral value of various sample areas after the simulation of GANN. By the capability of GAs to determine the optimal solution, this study can determine the network structure with minimum error between the prediction value and the actual value. As mentioned above, many researchers have proposed different encoding, exchange, mutation and selection to make the genetic algorithms suitable for applications in various research fields. The proposed encoding method in this study used the binary encoding, and employed the operation rules of elitism selection and uniform mating. During the operational process, the population size (number of chromosome groups) is usually in the range of 50~500. Taking into consideration the operating time and solution variability, this study selected 200 as the number of each generation, which is also used by other researchers [5]. The commonly used selection rate is about 0.10. The mating probability is mostly set in the range of 0.5~0.8. A higher mating rate means faster speed when searching for the optimal solution and a very low mating rate may easily stall the search process. As a result in [1], this study selected 0.6 as the mating rate. In general, the mutation probability is not high. If the mutation probability is set too high, it will be similar to a random search. Heng *et al*. [13] proposed that the mutation rate should be in the range from 0.1 to 0.001. To avoid the loss of best search capability, this study selected 0.1 as the mutation rate. The parameters of ANN include the number of hidden layers, number of neurons, learning rate and number of learn steps. The optimal convergence can be achieved when the number of hidden layers is in the range of 1 to 2. One hidden layer is sufficient for general problems. Two hidden layers can generally satisfy most of the problems. Therefore, this study set the upper limit for the number of hidden layers as two. The computer power and cases determine the upper limit of the number of neurons. When the number of neurons is larger than 30, the computer operational speed will be reduced considerably. In fact, cases of applying more than 30 neurons are rare. Therefore, according to the bit limit of the computer operation (32 bits), this study set the upper limit of the number of neurons as 32. This study set the upper limit of learning rate as 3.2. Too many learning steps can result in over-training and too few learning steps may result in difficulty to determine the optimal solution. Based on previous testing experience, the network will converge when the number of learning stepss is 15,000. Hence, this study set the upper limit of the learning steps as 15,000 to enable the algorithm to implement the program of searching for the optimal solution in this range. The rule for stopping the operation of the optimal search of GANN is set as satisfying one of the following conditions:

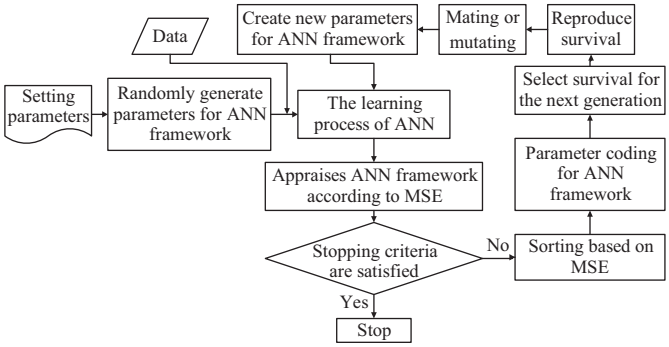

Fig. 2. GANN computation flow chart [2].

(1) The fitness function value is zero or the allowed value;
(2) All generations of percentages excluding the mutation rate satisfy the optimal fitness value without the occurrence of evolutionary changes more than 1000 times or above;
(3) When the above two termination conditions cannot be achieved, the operational times reach the set cumulative number of times.

The relevant GANN operational process shown in Fig. 2 was discussed by [2].

III. CONSTRUCTION OF THE ASSESSMENT MODEL FOR THE IMPACT OF SLOPE DEVELOPMENTAL FACTORS ON LANDSLIDE DISASTERS

This study on the assessment procedure for the impact of slope developmental factors on landslide disasters triggered by rainfall in Southern Taiwan is mainly divided into 7 stages. The content and steps of the various stages are explained below:

1. The Selected Study Area

From August 6 to August 10, 2009, Typhoon Morakot struck the central and southern parts of Taiwan, causing severe damage and heavy rainfall in the mountainous areas of Chiayi and Kaohsiung. Considering factors such as the existence of multiple slope developmental factors and effective satellite images of the typhoon attacked areas, this study selected the water gathering area of the Laonong River in the mountainous region of Kaohsiung as the study subject by selecting the rectangular area formed by the 2°UTM coordinates (214958, 2556344) and (219798, 2552520) as the study area. The study area is mainly located in the Liouguei District and Taoyuan District in Kaohsiung City. The satellite image of the study area is as shown in Fig. 3.

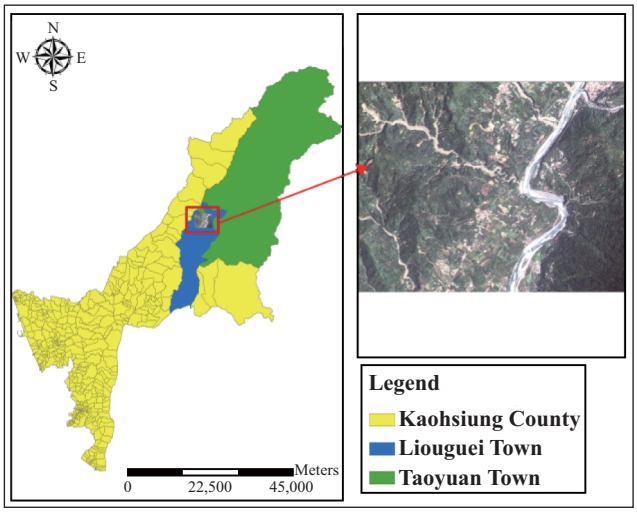

Fig. 3. Study area and corresponding satellite image.

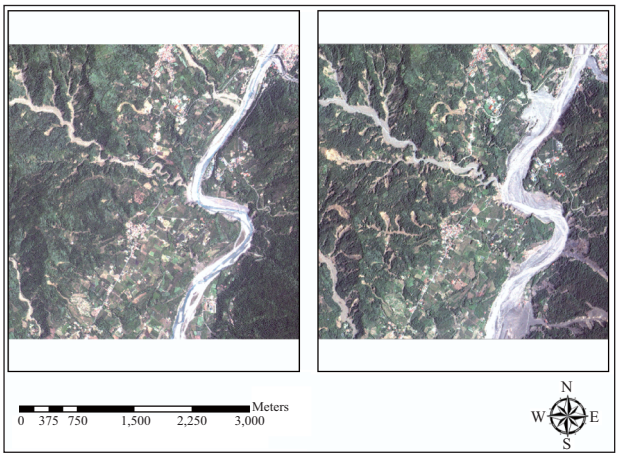

Fig. 4. Satellite images of the study area before and after Typhoon Morakot (left: May 9th, 2009, right: August 24th, 2009).

2. Building of Original Map Data

The original map data in this study included the FORMOSAT-2 (FM2) satellite images (Fig. 4), DEM and digital geological images of the study area in the Laonong River before and after the damage triggered by the heavy rainfall brought by Typhoon Morakot (2009.05.09 and 2009.08.24). This study combined these map data with GIS to build up the relevant database.

3. Establishment of Basic Grid

The spatial resolution of the integrated FM2 multi-spectral images used in this study can be up to 2 M while the DEM resolution is 40 M × 40 M. Hence, this study used the GIS software ArcGIS [7] to establish a grid of 40 M × 40 M as the basic grid of this study.

4. Selection of Factors of Influence

1) Natural Factors

Table 1. Geological strength level classification and encoding of the study area [1, 15].

Geological term	Characteristics	Strength Level	Level Encoding
Terrace Accumulation	gravel, clay, soil, sand	extremely weak	1
Liji layer, Kenting Layer	Foreign rocks in mudstone (badlands terrain)	extremely weak	1
Alluvial Layer	Soil, sand, gravel	very weak	2
Takangkou Formation	shale, sandstone, conglomerate rock	weak	3
Lushan Layer, Sule Layer	Hard shale, slate, phylite, hard sandstone	Medium strong	4
Bilushan Layer	Shale, quartize sandstone in phylite	Medium strong	4
Tananao Schists	Black schist , green schist, siliceous schist	strong	5
Duran Mountain	Agglomerate, tuffaceous sandstone, limestone, convex mirror body	Very strong	6

Natural factors are relevant to the geological factors and slope potential hazards of the study area. According to literature reviewed, this study selected the most widely used, most significant and easiest-to-get four factors of gradient, elevation, geology and slope roughness as the natural factors of the landslide disaster. Next, this study applied the Spatial Analysis Model of ArcGIS to get the average values of gradient and elevation in various unit grids and encode them by 6 levels of values. This study classified the geological factors' strength according to various geological terms and corresponding features. According to the corresponding compression strengths of the geological lithological properties, this study classified the geological features with such lithological properties by referring to the relationship between compression strength and strength level proposed by ISRM [15] and conducted level encoding as shown in Table 1. The slope roughness is the gradient undulations of each unit grid. Wilson and Gallant [31] proposed to measure the gradient changes by using the gradient standard deviations in a round window, which has an index meaning in the measurement of gradient undulations in an area of specific radius. The slope roughness is encoded by six grades.

2) Rainfall Factor

According to the typhoon observation records of the Central Weather Bureau in Taiwan, Typhoon Morakot was born in the sea about 1,000 km northeast of the Philippines on August 4, 2009. At 2 am on August 7, it moved slightly to a southern course and slowed down at about 5 am, before moving towards a west northwest course. Then, the heavy rain generated by the typhoon gradually affected East Taiwan.

Around 5 pm, it turned to a northwest course at a slower pace and the heavy rainfall affected various places around Taiwan. Typhoon Morakot landed at around 11:50 am near the

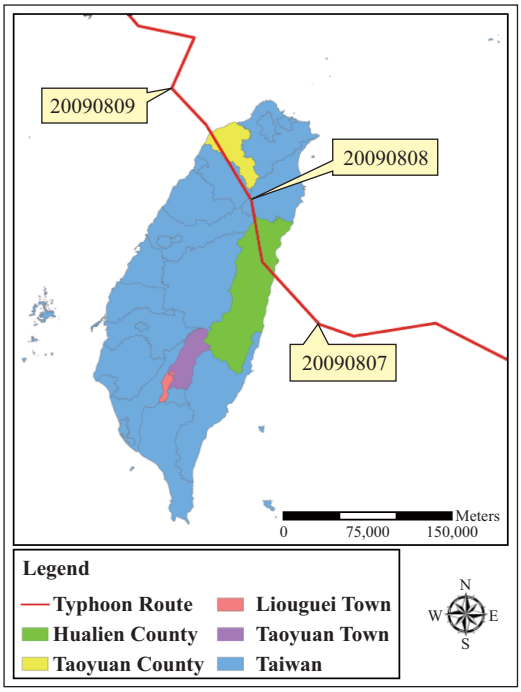

Fig. 5. Typhoon Morakot's roadmap.

city of Hualien in eastern Taiwan and its strength gradually became weaker at 2 am on August 8, before it moved northward at around 5 o'clock and left Taiwan near Taoyuan in northwest Taiwan at around 14.00 as shown in Fig. 5. The typhoon zone of Typhoon Morakot had very high water content, in particular in the area southwest area of the epicenter of the typhoon. Moreover, the typhoon zone lingered in the Island of Taiwan for more than 60 hours. This study referred to the EAR (Effective Accumulative Rainfall) defined in previous studies [27] as the addition of the direct rainfall and the previous indirect rainfall. The direct rainfall was the total rainfall until the occurrence of landslide disaster measured from the first rain of the major rainfall field to the occurrence of landslide. The indirect rainfall of the previous period was the total rainfall in the 7 days before the major rainfall, namely,

$$\sum_{k=1}^{7} k^n P_n = P_b \tag{1}$$

where, P_n is the amount of rainfall in n days before the major rainfall field (mm); k was the degression coefficient set as 0.9 in this study. As direct rainfall is directly related to the occurrence of a landslide, its contribution to the disaster is relatively direct and effective; it was not discounted in this study. Hence, the EAR can be inferred by the following equation:

$$EAR = P_r + P_b \tag{2}$$

where, P_r is the direct rainfall. After the calculation of the EARs of various rainfall observation stations in the study area, this study used the IDW (Inverse Distance Weighted) method to infer the average EAR of various grids in the study area.

Table 2. Parameters of network training structure.

Image Date	No. of Hidden Layers	1st hidden layer No. of Neurons	2nd hidden layer No. of Neurons	Learning Rate	No. of Learning steps	OA (%)
2009 05.09	2	22	29	1.5	15000	91.9
2009 08.24	2	31	24	1.5	15000	92.1

3) Slope Development Factor

According to previous studies, improper slope development can be categorized into deforestation, change of terrain and building of community facilities. Therefore, when classifying the interpretation of satellite images, this study referred to 10 interpretation factors proposed by Chen *et al*. [1], including water area, fruit tree, buildings, forest, grassland, bare land, paddy fields, dry land, roads and river channels and selected 6 factors including green coverage (including grassland, forest), farmland (including paddy fields, dry land) planting rate, fruit tree planting rate, bare land, building density, and road density as the slope developmental factors affecting the occurrence of landslide disasters in the study area. Quantitative indicators relating to various slope developmental factors were defined as the area percentage of each factor in the grid area.

5. Judgment and Interpretation of Satellite Images

1) Preprocessing of Satellite Images

In order to avoid the impact of cloud on the interpretation of satellite images, this study first applied the ArcGIS software to digitalize the areas covered by clouds and shadows in the images, in addition to selecting areas of relatively less cloud shadowing. Then, this study applied the function of ERDAS IMAGINE [6] to cut out the selected areas, and acquired the four types of spectral values of R, G, NIR and SWIR in the images for the training and classification of image interpretations.

2) Sample Area Drawing and Model Training

This study selected the representative sample areas from satellite images by circling and wrote the GANN model on the platform of MATLAB [20] to train samples. The trained structural parameters of the satellite images of the study area before and after the heavy rainfall are shown in Table 2. The training results of the satellite images before the heavy rainfall (2009.05.09) were: the number of hidden layers was 2, the number of neurons of the first layer was 22, the number of neurons of the second layer was 29, learning rate was 1.5, when the number of learning steps was 15000, the OA (overall accuracy) can be up to 91.9%. The training results of the (2009.08.24) satellite images after the heavy rainfall were: the number of hidden layers was 2, the number of neurons of the

Table 3. The relationship of classification factors of the error matrix [30].

		Actual Surface		Total
		Classification A	Classification B	
Classification Results	Classification A	E_{11}	E_{12}	E_{+i}
	Classification B	E_{21}	E_{22}	E_{+i}
Total		E_{i+}	E_{i+}	E_{++}

Table 4. Error matrix of classification results of the images before heavy rainfall (2009.05.09).

	Road	Farmland	Bare land	Water Area	Forest	Fruit Tree	Building	Total	User Accuracy (%)
Road	16	0	3	2	0	0	0	21	76
Farmland	1	40	3	7	0	7	1	59	67
Bare Land	7	0	43	0	0	0	5	55	78
Water Area	1	0	0	14	0	0	0	15	93
Forest	0	6	1	2	47	4	0	60	78
Fruit Tree	0	4	0	0	3	14	0	21	66
Building	0	0	0	0	0	0	19	19	100
Total	25	50	50	25	50	25	25	250	
Producer Accuracy (%)	64	80	86	56	94	56	76		Kappa = 0.75 OA = 77%

first layer was 31, the number of neurons of the second layer was 24, learning rate was 1.5, when the number of learning steps was 15000, the OA can be up to 92.1%.

3) Classification Results of Images

This study applied the optimal network architecture obtained from the above-mentioned GANN training to carry out the global satellite image classification. To confirm the classification accuracy, this study randomly selected 25 points in the satellite image as the verification point regarding various factors of interpretation. At present, the most commonly used method to assess the image accuracy is the error matrix method presented in [30]. The error matrix is a rectangle representing the difference between classification results and reference data of different land surface types such as the four types shown in Table 3. The column represents the category of the reference data, and the row represents the classification while various elements represent the amount of data under corresponding categories. As shown in Table 3, E_{12} represents the amount of data of category B as interpreted as category A; E_{11} represents the amount of data correctly interpreted as category A. This study adopted the coefficient of agreement Kappa [3] and OA as the basis of the classification accuracy as illustrated below:

A. OA (Overall Accuracy)

OA is the simplest overall description method. Of all the methods, this value represents the probability of correct classification of any point in the area, namely, the weighted average of the classification results of various classification accuracies against amounts. The OA is calculated as:

$$\mathrm{OA} = \left(\frac{1}{N} \sum_{i=1}^{r} E_{ii} \right) \times 100\% \tag{3}$$

where, N is the total number of verified points, r is the order of the matrix, E_{ii} is the number of correctly classified points.

B. Kappa $(\hat{\mathrm{K}})$

The Kappa indicator represents the degree of classification results which were better than the random classification. Kappa is calculated as below:

$$\hat{\mathrm{K}} = \frac{N \sum_{i=1}^{r} E_{ii} - \sum_{i=1}^{r} \left(E_{i+} \times E_{+i} \right)}{N^2 - \sum_{i=1}^{r} \left(E_{i+} \times E_{+i} \right)} \tag{4}$$

Kappa indicators take into account two types of consistent differences including the consistency between automatic classification and the reference data and the consistent probability between sampling and reference classifications. In general, Kappa value is in the range of 0~1, and a larger Kappa value represents higher classification accuracy. With images before heavy rainfall as an example (as shown in Table 4), the overall accuracy of image classification was 77%, and the Kappa value was 0.75. Similarly, the overall accuracy of image classification after the heavy rainfall was 75%, and the Kappa value was 0.72. According to relevant studies in [18], the interpretation results of applying the GANN in image classification in this study were acceptable and of medium and high accuracy level.

Table 5. Co-relational examination results of slope developmental factors.

	Green Coverage Rate	Fruit Tree Planting Rate	Farmland Planting Rate	Bareness	Road Density	Building Density
Green Coverage Rate	1	0.18	-0.24	-0.12	-0.29	-0.14
Fruit Tree Planting Rate	0.18	1	0.28	-0.11	-0.20	-0.08
Farmland Planting Rate	-0.24	0.28	1	-0.07	0.06	0.01
Bareness	-0.12	-0.11	-0.07	1	0.32	0.25
Road Density	-0.29	-0.20	0.06	0.32	1	0.25
Building Density	-0.14	-0.08	0.01	0.25	0.25	1

Table 6. Correlation examination results of natural factors.

	Geology	Gradient	Elevation	Slope Roughness
Geology	1	-0.16	-0.31	-0.05
Gradient	-0.16	1	0.45	0.31
Elevation	-0.31	0.45	1	-0.04
Slope Roughness	-0.05	0.31	-0.04	1

Table 7. The optimal network architectural parameters of the landslide assessment model and the training and testing results of the model.

Type of Factors of Influence	No. of Hidden Layers	No. of Neurons		Learning Rate	No. of Learning Steps	Training Group			Testing Group		
		1st	2nd			Total No. of Batches	No. of Erroneous Batches	Accuracy Rate	Total No. of Batches	No. of Erroneous Batches	Accuracy Rate
Slope Developmental Factors	2	29	27	0.4	5000	126	9	92.9%	36	9	75%
Natural Environmental Factors	2	25	14	0.3	5000	98	9	90.8%	64	20	68.8%

6. Selection of Landslide Area

In this study, when the difference between bare land in the images of a given grid before and after Typhoon Morakot accounted for more than 1/2 of the grid area and the NDVI (Normalized Difference Vegetation Index) gap of the grid images was smaller than 0, the grid would be defined as the place of landslide occurrence. The NDVI indicator is inferred according to the following relationship equation in [26, 29]:

$$NDVI = \frac{NIR - RED}{NIR + RED} \tag{5}$$

7. The Building of the Assessment Model

This study applied the GANN in the quantitative assessment of the impact of the slope developmental factors and natural factors on the landslide disaster, and the building procedure is detailed below:

1) Factor Correlation Examination

To confirm whether slope developmental and natural factors were correlated, this study used the Pearson correlation tool of the SPSS (2005) [28] to conduct the factor correlation examination. The correlation examination results of slope developmental factors and natural factors were as shown in Tables 5 and 6. The findings suggested that both slope developmental factors and natural factors had low or no correlation in between them.

2) Building of the Landslide Assessment Model

In both cases of slope developmental factors or natural factors, the training group and the testing group for the landslide assessment model in this study were selected in the ratio of 7:3 in terms of number of samples. Meanwhile, the ratio of sample areas of landslide and non-landslide in various groups of samples was 1:1. This study adopted various factor data of 162 grids for the training and testing of the GANN model. The optimal network architectural parameters obtained after the training and testing were as shown in Table 7. The comparison of the GANN prediction results and the status revealed satellite images showed that there were nine batches of data indicating inconsistency between the actual situation and GANN prediction results in respect of slope developmental factors and the overall rate of successful prediction of disasters was 92.9%. In addition, there were nine batches of data of incon-

sistency between the actual situation and GANN prediction results in respect of natural factors, and the overall rate of successful prediction of disasters was 90.8%.

3) Testing of the Landslide Assessment Model

Besides the training sets (126) for the building of the assessment model, the remaining 36 sets of data were for the testing of the model. The testing results indicated that there were nine erroneous batches of data in respect of the slope developmental factors with a success rate of 75%. The total success rate of the training sets was up to 88.8%. As shown in Table 7, there were 20 erroneous batches of data in respect of natural factors with successful rate up to 68.8%. The total success rate of the training sets was up to 82.1%.

4) Weight Proportion

After the generation of the optimal network architecture by GANN, the GANN-optimized weight proportion can be used to deduce the impact of each input factor on the overall network architecture. This study referred to the weight proportion calculation method proposed by [9] and [10] to deduce the weight of each factor. With the GANN output weights as examples, the inference process is illustrated below:

No. of Neurons (Hidden layer)	Weight						
	Input No. 1	Input No. 2	Input No. 3	Input No. 4	Input No. 5	Input No. 6	Output
1	IW_{11}	IW_{12}	IW_{13}	IW_{14}	IW_{15}	IW_{16}	OW_1
2	IW_{21}	IW_{22}	IW_{23}	IW_{24}	IW_{25}	IW_{26}	OW_2
3	IW_{31}	IW_{32}	IW_{33}	IW_{34}	IW_{35}	IW_{36}	OW_3

The Input in the above table indicates network input factors; and the Output in the above table indicates the network output target values; IW_{23} represents the weight value of the No. 3 input factor of the second neuron of the hidden layer.

Step 1: The multiplication of the absolute values of the weighted values o of various input variables j of i number neurons of the hidden layer number of neurons and the absolute values of various output values can get P_{ij}

No. of Neurons (Hidden Layer)	Weight						
	Input No. 1	Input No. 2	Input No. 3	Input No. 4	Input No. 5	Input No. 6	Output
1	$P_{11} = IW_{11} \times OW_1$	$P_{12} = IW_{12} \times OW_1$	$P_{13} = IW_{13} \times OW_1$	$P_{14} = IW_{14} \times OW_1$	$P_{15} = IW_{15} \times OW_1$	$P_{16} = IW_{16} \times OW_1$	$P_{11} = IW_{11} \times OW_1$
2	$P_{21} = IW_{21} \times OW_2$	$P_{22} = IW_{22} \times OW_2$	$P_{23} = IW_{23} \times OW_2$	$P_{24} = IW_{24} \times OW_2$	$P_{25} = IW_{25} \times OW_2$	$P_{26} = IW_{26} \times OW_2$	$P_{21} = IW_{21} \times OW_2$
3	$P_{31} = IW_{31} \times OW_3$	$P_{32} = IW_{32} \times OW_3$	$P_{33} = IW_{33} \times OW_3$	$P_{34} = IW_{34} \times OW_3$	$P_{35} = IW_{35} \times OW_3$	$P_{36} = IW_{36} \times OW_3$	$P_{31} = IW_{31} \times OW_3$

Step 2: Regarding each neuron of the hidden layer, the division of P_{ij} by the multiplication of the input parameters of the same layer can get Q_{ij}. For example:

$$Q_{12} = P_{12}/(P_{11} + P_{12} + P_{13} + P_{14} + P_{15} + P_{16})$$

Step 3: Regarding each input parameter, the addition of the Q_{ij} of various parameters can get S_j. For example:

$$S_1 = Q_{11} + Q_{21} + Q_{31}$$

No. of Neurons (Hidden Layer)	Weight					
	Input No. 1	Input No. 2	Input No. 3	Input No. 4	Input No. 5	Input No. 6
1	Q_{11}	Q_{12}	Q_{13}	Q_{14}	Q_{15}	Q_{16}
2	Q_{21}	Q_{22}	Q_{23}	Q_{24}	Q_{25}	Q_{26}
3	Q_{31}	Q_{32}	Q_{33}	Q_{34}	Q_{35}	Q_{36}
Sum	S_1	S_2	S_3	S_4	S_5	S_6

Step 4: The division of Sj by ΣSj can get the weight proportion of various parameters. For example:

$$\text{weight proportion} = (S_1 \times 100\%)/(S_1 + S_2 + S_3 + S_4 + S_5 + S_6)$$

When the hidden layer is set as one layer, the weight proportion deduction of the optimized ANN architecture model can be proceeded as the above Steps 1 to 4. When the hidden layer is set as two, the above Steps 1 to 4 can infer the weight proportion of the second hidden layer, which is the target output value of the first hidden layer. Then, the procedure of Steps 1 to 4 was repeated to calculate the weights of various factors. The weight proportions of factors of slope development and natural environment affecting landslides were as shown in Tables 8 and 9.

IV. THE DRAWING OF THE LANDSLIDE POTENTIAL DIAGRAM

1. Deduction of Slope Development Degree

To explore the relationship between slope development, rainfall and landslide occurrence, this study referred to the concept of the slope DLD (Degree of Land Disturbance) proposed by Chen *et al*. [1]. The DLD of the slope in the grid can be defined as:

$$DLD = \frac{DC}{EC} \tag{6}$$

where, DC (Disturbance Condition) refers to the slope developmental factors in the grid of 40 M × 40 M, namely, the land surface development type in the grid. EC (Environment

Table 8. Weights W_{DC} of slope developmental factors of landslide.

Slope Develop. Factor	Fruit Tree Planting	Green Coverage Rate	Farmland Planting	Bareness	Road Density	Building Density
Weights	0.100	0.107	0.120	0.121	0.193	0.359

Table 9. Weights W_{EC} of slope natural factors of landslide.

Natural Conditional Factor	Geology	Gradient	Elevation	Slope Roughness
Weight	0.46	0.20	0.19	0.15

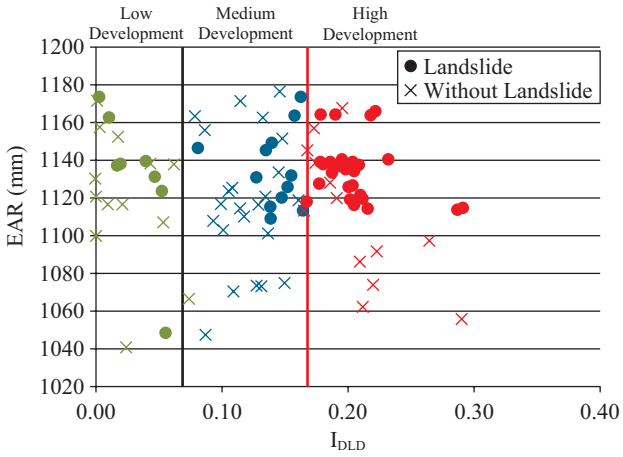

Fig. 6. Relationship between slope development degree, rainfall and the occurrence of landslides.

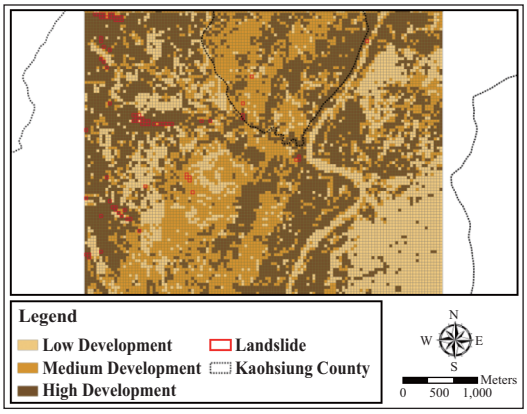

Fig. 7. Slope DLD classification.

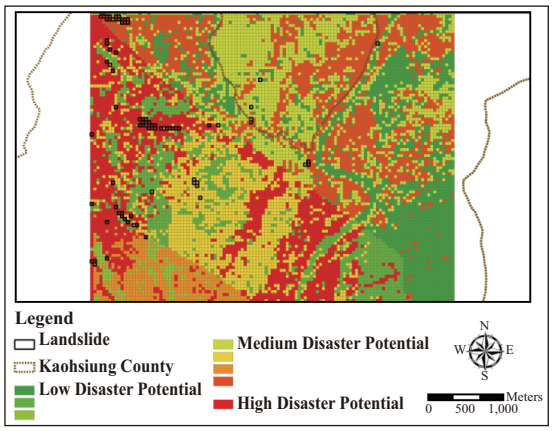

Fig. 8. Landslide potential map.

Condition) refers to the natural factors (average gradient, average elevation, geological strength, slope roughness) of the study area after classification. Hence, the Index for DLD (I_{DLD}) can be defined as:

$$I_{DLD} = \frac{G_{DC}}{G_{EC}} = \frac{\sum W_{DC} \times R_{DC}}{\sum W_{EC} \times R_{EC}} \quad (7)$$

where, G_{DC} (Grading of DC) refers to the grading indicator of slope development in various grids. G_{EC} (Grading of EC) is the grading indicator of natural factors of influence. W_{DC} and W_{EC} are the weight proportions of the above mentioned slope developmental factors and natural environmental factors. R_{DC} is the proportion of the area of the slope developmental factors in the grid area. The slope DLDs of the 162 grids constructed and verified by the model proposed in this study can be deduced. With the grid slope DLD, EAR and the interpretation data of landslides, a relational map can be described as shown in Fig. 6. As seen, X-axis is the Index for DLD, and Y-axis is EAR. By cluster analysis, the Index for DLD can be classified into high, medium and low grades. As shown in the diagram, regardless of the slope development degree, the occurrence of landslides is generally proportional to EAR. In addition, according to previous studies, a rainfall-triggering landslide can be predicted by the strength or accumulated amount of rainfall. The strength of rainfall gradually rose to the peak value and the accumulated amount of rainfall increased rapidly. Next, when the rainfall strength exceeded the peak value and started to decline, namely, the increasing rate of the accumulated rainfall curve started to slow down, it was the time for slope landslides to occur. When the accumulated rainfall amount was above 1110 mm, regardless of high, medium or low slope development degrees, there would be many landslide spots in the study area.

2. Drawing of Landslide Potential Map

According to the inference results by Eq. (7) and Fig. 6, the map of slope DLD classification can be drawn (see Fig. 7). This study used the discriminant analysis function of SPSS statistical analysis software to determine the discriminant equations of the sample data as shown in Fig. 6 (namely, the EAR, Index for DLD and the occurrence of landslides). The input independent variables included the Index for DLD and EAR, and the dependent variable is the occurrence of landslides. The discriminant equation is as shown in Eq. (8).

$$ODL = 30.976 - (I_{DLD} \times 11.108) - (EAR \times 0.026) \quad (8)$$

where, ODL (The Occurrence Degree of Landslides) is the level of landslides; I_{DLD} and EAR represent the above mentioned Index for DLD and EAR. Based on the discriminant equation of landslide occurrence and data of slope DLD and EAR, this study described the potential map of landslides due to slope development (see Fig. 8). The study first determined the threshold value of the landslide occurrence for the discriminant equation (the judgment of landslides), then input the degree values of I_{DLD} (namely, the discriminant values of low and medium slope development, medium and high slope development) to determine the corresponding EAR values to describe the landslide disaster potential map. The red, orange, light green and green colors in the map represent the high potential danger zone to low potential danger zone. From statistics, there were 23 landslide disaster spots in the study area. Three spots were in the low potential danger zone of this study, while 20 spots were in the high potential danger zone. Therefore, the historical occurrences of landslide disasters were roughly consistent with the descriptions of the landslide disaster potential map.

V. CONCLUSIONS

This study applied the Genetic Adaptive Neural Network (GANN) structure in satellite image classification and the assessment of landslide disasters. In addition, this study explored the relationships between slope development degree, rainfall and landslide occurrence and described the disaster potential map by integration with the GIS. The study reached the following conclusions:

(1) This study used the MATLAB as the numerical simulation software platform to develop the GANN model. OA and coefficient of agreement of image interpretation results indicated that the proposed satellite image interpretation procedure study can be successfully applied in the assessment of the impact of the slope developmental factors and natural environmental factors on landslides disaster to reduce calculation complexity and calculation time.

(2) This study applied the optimal network architecture obtained from the training of GANN in the global interpretation and classification of satellite image and randomly selected 25 spots in the image as verification spots. The overall accuracy rate of interpretation of the image before heavy rainfall can be up to 77%, and the Kappa value is 0.75; the overall accuracy rate of interpretation of the image after heavy rainfall can be up to 75%, and the Kappa value is 0.72. In both cases, the result is of medium and high accuracy level.

(3) This study used the catchment area of the Laonong River in Kaohsiung, Southern Taiwan as the study area using the GANN in the assessment of the impact of slope developmental factors and natural factors on landslides. The assessment results demonstrated that the overall prediction rate of landslide occurrence can be 88.8% and 82.1% in case of slope developmental factors and natural environmental factors respectively.

(4) After establishing the optimal network structure by the GANN program, the weights of various slope developmental factors and natural environment factors can be obtained to assess their impact on the slope redevelopment and establish the Index for DLD as the basis for the quantitative assessment of the impact of slope development on landslides.

(5) The results of discussion on the relationships between the Index for DLD, rainfall and landslide occurrence revealed that the occurrence of landslide is proportional to EAR value regardless of the Index for DLD. When the accumulated rainfall of the study area was more than 1110 mm during the attack of Typhoon Morakot, landslides occurred in many spots.

(6) This study applied discriminant analysis to analyze the relationships between occurrence of landslide, EAR and Index for DLD, and determine the discriminant equation of the landslide occurrence. Combining the discriminant equation with the Index for DLD and EAR data, this study drew the map of potential landslides triggered by slope development for the reference of academia and engineers in the assessment of landslide disasters caused by slope development.

REFERENCES

1. Chen, Y. R., Chen, J. W., Hsieh, S. C., and Ni, P. N., "The application of remote sensing technology to the interpretation of land use for rainfall-induced landslides based on genetic algorithms and artificial neural networks," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, Vol. 2, No. 2, pp. 87-95 (2009).
2. Chen, Y. R., Hsieh, S. C., and Liu, C. H., "Simulation of stress-strain behavior of saturated sand in undrained triaxial tests based on genetic adaptive neural networks," *Electronic Journal of Geotechnical Engineering*, Vol. 15, pp. 1815-1834 (2010).
3. Cohen, J., "A coefficient of agreement for nominal scales," *Educational and Psychological Measurement*, Vol. 20, pp. 37-46 (1960).
4. Coppin, N. J. and Richards, I. G., *Use of Vegetation in Civil Engineering*, CIRIA Butterworths, London, pp. 23-36 (1990).
5. D'Ambrosio, D., Spataro, W., and Iovine, G., "Parallel genetic algorithms for optimizing cellular automata models of natural complex phenomena: An application to debris flows," *Computer and Geosciences*, Vol. 32, pp. 861-875 (2006).
6. ERDAS, *ERDAS IMAGE Tour Guide*, ERDAS World Headquarter, Atlanta (2011).
7. ESRI, *ArcGIS 9 - Using ArcGIS Desktop*, Environmental Systems Research Institute (2006).
8. Fernandez, C. I., Castillo, T. F., Hamdouni, R. E., and Momtero, J. C., "Verification of landslide susceptibility mapping: a case study," *Earth Surface Processes and Landforms*, Vol. 24, pp. 537-544 (1999).
9. Garson, G. D., "Interpreting neural-network connection weights," *AI Expert*, Vol. 6, No. 7, pp. 47-51 (1991).
10. Goh, A., "Seismic liquefaction potential assessed by neural networks," *Journal of Geotechnical Engineering ASCE*, Vol. 120, No. 9, pp. 1467-1480 (1994).
11. Goldberg, D. E., *Genetic Algorithms in Search, Optimization, and Ma-*

chine Learning, Addison-Wesley, Reading (2007).
12. Guillande, G., Pascale, G., Jacques-Marie, B., Robert, B., Jean, C., Benoît, D., and Jean-François, P., "Automated mapping of the landslide hazard on the island of Tahiti based on digital satellite data," *Mapping Sciences and Remote Sensing*, Vol. 32, No. 1, pp. 59-70 (1995).
13. Heng, L., Cao, J. N., and Love, P. E., "Using machine learning and GA to solve time-cost trade-off problem," *Journal of Construction Engineering and Management*, Vol. 125, pp. 347-353 (1999).
14. Holland, J. H., *Adaptation in Natural and Artificial Systems*, The University of Michigan Press (1975).
15. ISRM, *Rock Characterization Testing and Monitoring: ISRM Suggested Method*, Pergamon Press, London (1981).
16. Jarvis, C. H. and Stuart, N., "The sensitivity of a neural networks for classifying remotely sensed imagery," *Computers and Geosciences*, Vol. 22, No. 9, pp. 959-967 (1996).
17. Kanungo, D. P., Arora, M. K., Sarkar, S., and Gupta, R. P., "A comparative study of conventional, ANN black box, fuzzy and combined neural and fuzzy weighting procedures for landslide susceptibility zonation in Darjeeling Himalayas," *Engineering Geology*, Vol. 85, pp. 347-366 (2006).
18. Landis, J. R. and Koch, G. G., "The measurement of observer agreement for categorical data," *Biometrics*, Vol. 33, No. 1, pp. 159-174 (1977).
19. Lee, S., Ryu, J. H., Won, J. S., and Park, H. J., "Determination and application of the weights for landslide susceptibility mapping using an artificial neural network," *Engineering Geology*, Vol. 71, pp. 289-302 (2004).
20. Liu, H. Y., Gao, J. X., and Li, Z. G., "The advances in the application of remote sensing technology to the study of land covering and land utilization," *Remote Sensing for Land and Resources*, Vol. 4, pp. 7-12 (2001).
21. MATLAB, *User Guide*, The Mathworks, Inc. (2010).
22. Miller, D. J. and Sias, J., "Deciphering large landslides: linking hydrological, groundwater and slope stability models through GIS," *Hydrological Processes*, Vol. 12, pp. 923-941 (1998).
23. Mora, S. and Vahrson, W., "Macrozonation methodology for landslide hazard Determination," *Bulletin of the Association of Engineering Geologists*, Vol. 31, pp. 49-58 (1994).
24. Nikolakopoulos, K. G., Vaiopoulos, D. A., Skianis, G. A., Sarantinos, P., and Tsitsikas, A., "Combined use of remote sensing, GIS and GPS data for landslide mapping," *Proceedings of the 2005 IEEE International Geoscience and Remote Sensing Symposium*, Vol. 7, pp. 5196-5199 (2005).
25. Popescu, M. E., "Landslide causal factors and landslide remedial options," Keynote Lecture, *Proceedings of the 3rd International Conference on Landslides, Slope Stability and Safety of Infra-Structures*, Singapore, pp. 61-81 (2002).
26. Schott, J. R., *Remote Sensing*, Oxford University Press (1997).
27. Seo, K. and Funasaki, M., "Relationship between sediment disaster (mainly debris flow damage) and rainfall," *International Journal of Erosion Control Engineering*, Vol. 26, No. 2, pp. 22-28 (1973).
28. SPSS, *Brief Guide*, SPSS Inc. (2005).
29. Tucker, C. J., "Red and photographic infrared linear combinations for monitoring vegetation," *Remote Sensing of Environment*, Vol. 8, pp. 127-150 (1979).
30. Verbyla, D. L., *Satellite Remote Sensing of Natural Resources*, CRC Press, New York (1995).
31. Wilson, J. P. and Gallant, J. C., *Terrain Analysis*, John Wiley & Sons, New York (2000).
32. Yoshida, T. and Omatu, S., "Neural networks approach to land cover mapping," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 32, No. 5, pp. 1103-1109 (1994).