Proceedings of the 5th WSEAS International Conference on Applied Computer Science, Hangzhou, China, April 16-18, 2006 (pp503-511)

Satellite-Derived Land-Cover Classification Using Immune Based Mining Approach

Ta-Cheng Chen, Chao-Yuan Chen
Department of Information Management, Optimization and Logistics LAB
National Formosa University
64 Wenhua Rd., Huwei, Yunlin
TAIWAN

Abstract: - With the epidemic application of geographic information systems retaining remotely sensed data as layers, the accuracy assessment of the map generated from any remotely sensed data has been even more crucial. In recent, data mining has been widely applied in many areas. It means the methodologies and tools for the efficient new knowledge discovery from databases. In this paper, a immune algorithms based mining approach to accuracy assessment is proposed for extracting the decision rules including the predictors, the corresponding inequality and threshold values simultaneously so as to building a decision-making model with maximum prediction accuracy. Early many studies of handling the satellite-derived land-cover classification problems used the statistical related techniques. As the land-cover classification is highly nonlinear in nature, it is hard to develop a comprehensive model using conventional statistical approaches. Recently, numerous studies have demonstrated that neural networks (*NNs*) are more reliable than the traditional statistical approaches. The usefulness of using *NNs* have been reported in literatures but the most obstacle is the in the building and using the model in which the classification rules are hard to be realized. We compared our results against commercial GIS software, and we show experimentally that the proposed rule extraction approach is promising for improving land-cover classification accuracy and enhancing the modeling simplicity. In particular, our approach is capable of extracting rules which can be developed as a computer model for classification of satellite-derived land-cover potential like expert systems.
Key-Words: - Immune algorithms, Decision rules, Remotely sensed data, Data mining, Land-cover

1 Introduction

Remote sensing land cover images mostly are with some unique texture patterns so that the texture feature extraction and classification methods become in remote sensing land cover image processing. Many studies of handling the satellite-derived land-cover classification problems used the statistical related techniques [1] [7], such as maximum likelihood classification (MLC) based on the Bayesian theorem, have been successfully used in remote sensing land cover classification. As the land-cover classification is highly nonlinear in nature, it is hard to develop a comprehensive model using conventional statistical approaches. In order to overcome this problem, which is inherent in statistical approaches, non-parametric classification techniques such as neural networks and rule-based classifiers are increasingly being applied [16]. Especially, numerous studies have demonstrated that neural networks (*NNs*) are more reliable than the traditional statistical approaches [15]. The usefulness of using *NNs* have been reported in literatures but the most obstacle is the in the building and using the model in which the classification rules are hard to be realized. In this study, a data mining approach is proposed to solve the above problem.

Nowadays, the *LC* image data can be captured and stored hugely and easily in database. However, such kind of raw data is rarely of direct benefit. So, the value of these data is predicated on the ability to extract information useful for decision support or exploration, and understanding the phenomenon governing the LC data source. Traditionally, data analysis to retrieve the knowledge by the analyst(s) was a visual interpreted process instead of the automatic process. However, those visual interpreted processes easily break down while the size of the data grows and the number of dimensions increases. For dealing with the scale of data manipulation, exploration going beyond human capacities, the computing technologies for automating the process is desired and need to be developed. Data mining usually means the methodologies and tools for the efficient new knowledge discovery from databases. It is also a form of knowledge discovery essential for solving problems in a specific domain. In this study, we consider the land cover classification as a general multi-class classification problem and developed a

data mining based approach for solving the problems. While the classification rules are minded, the accuracy assessment is needed because of the ability to compare method/algorithms quantitatively, and the desire to use the resulting maps/spatial information in some decision-making process [13].

This paper is arranged as follows: in the next section the proposed IAs based mining approach is introduced. Section 3 provides a description of the data used for this study and the results of our experiments are also discussed. Finally, the conclusion of the paper is summarized and the directions for future research are described.

2 The discover of Decision Rules Using Immune Algorithms

The natural immune system of all animals is a very complex system for defense against pathogenic organisms. A two-tier line of defense is in the system including the innate immune system and the adaptive immune system. The basic components are lymphocytes and antibodies [8]. The cells of the innate immune system are immediately available to combat against a wide variety of antigen without previous exposure to them. The antibody production in response to a determined infectious agent (antigen) is the adaptive immune response mediated by lymphocytes which are responsible for recognition and elimination of the pathogenic agents [6]. The cells in the adaptive system are able to develop an immune memory so that they can recognize the same antigenic stimulus when it is presented to the organism again. Also, all the antibodies are produced only in response to specific infections. There are two main types of lymphocytes: B-lymphocytes (B-cell) and T-lymphocytes (C-cell). B-cell and T-cell carry surface receptor molecules capable of recognizing antigens. The antigens will only bind to these receptors with which it makes a good fit.

To distinguish and eliminate the intruders of the organism is the main task of the immune system so that it must has the capability of self/non-self discrimination. As mentioned previously, various antibodies can be produced and then can recognize the specific antigens. The portion of antigen recognized by antibody is called epitope which acts as an antigen determinant. Every type of antibody has its own specific antigen determinant which is called idiotope. Moreover, In order to produce enough specific effector cells to against an infection, and activated lymphocyte has to proliferate and then differentiate into these effector cells. This process is called clonal selection [20] and followed by the genetic operations such that a large clone of plasma cell is formed. Therefore, the antibodies can be secreted and ready to bind antigens. According to above facts, [10] proposed an idiotype network hypothesis which is based on the clonal selection theory. In his hypothesis, some types of recognizing sets are activated by some antigens and produce an antibody which will then activate other types of recognizing sets. By this way, the activation is propagated through entire network of recognizing sets via antigen-antibody reactions. It is noted that the antigen identification is not done by a single or multiple recognizing sets but by antigen-antibody interactions. From this point of view, for solving the combinatory optimization problems, the antibody and antigen can be looked as the solution and objection function respectively.

2.1 The IA based mining approach

Before extracting the decision rule(s), the most significant predictors of the best subset have to be decided; otherwise, the insignificant predictors become the noise which may worsen the genetic learning or even mislead the wrong learning and cause an unreasonable classification. In addition to the best subset of the significant predictors considered in this study, multiple rules are explored for increasing the prediction accuracy. Unlike those approaches used in literature [2] [11] [17], the proposed approach is not to select the first *n* best-fit rules within a genetic searching process as the mined rules for building the prediction model. As mentioned previously, those best *n* rules are the rules not converged to be the best one. In our proposed approach, additional new decision rule is to be explored when the previous rule(s) failed to classify all the sample data correctly. In other words, if one rule is generated but not with good enough prediction accuracy, an additional new rule is going to be extracted using those data which can not be classified correctly. The computation procedures of the proposed rule mining approach contain three major processes, which are data preprocess, rule mining process and data modification process respectively (see Fig.1). The discussion of the experimental results comes in sequence.

2.1.1 Data preprocess

The training data are selected from the whole dataset randomly and directly fed into the proposed mining approach. By beginning the mining procedure, the attribute of rules to be extracted should be decided too.

2.1.2 Rule mining process

Step_1. Generate an initial population of strings randomly.

Step_2. Convert each individual of the current population into *If-Then* rule.

Step_3. Using the training dataset from the procedure of data preprocess to evaluate each of *If-Then* rules, i.e., calculate the corresponding fitness value (prediction accuracy) for each individual.

Ste_4. Check the stopping criterion, if not stop then go to *Step_5*. Otherwise go to *Step 6*.

Step_5. The individuals in *Step_4* will suffer the genetic operation process, i.e., selection, crossover and mutation [14]. After new generation is generated, then go to *Step 2*.

Step_6. Check whether the predefined accuracy can be met or not? If not met then go to *Step 7*, otherwise go to *Step 8*.

Step_7. Start the procedure to extract an additional ***If-Then*** rule. Go to the *Data modification process* and then update the training dataset. After updating the training dataset, go back to *Step 1*.

Step_8. Stop. The optimal or near optimal decision rule(s), also called prediction rule(s) can be obtained from the training data set.

2.1.3 Data Modification Process

In this process, both the selected attribute of training data (malignant data) not being classified correctly and all the unselected data (benign data) are preserved for mining an additional rule. The additional mined rule is expected to enhance the prediction accuracy so as to correctly classify those data which were incorrectly classified by previous mined rule(s).

After the above three processes are completed, the rules output can serve as meta knowledge concerning the given dataset. The validation dataset are then used to verify the accuracy of the prediction rules mined. Again, unlike the GAs based rule generation approaches proposed by [17] and [11], any additional rule extracted is based on the data not been classified correctly. Obviously, the prediction accuracy of mined rules is on the cumulative way to enhance the prediction accuracy. In our implementation, the decision rules are represented by strings of binary digits. Each string consisting of substring includes the number of significant predictors (decision variables) and the corresponding rule containing variables, equality/inequality and threshold. The stopping criterion is the maximum iterations in this article.

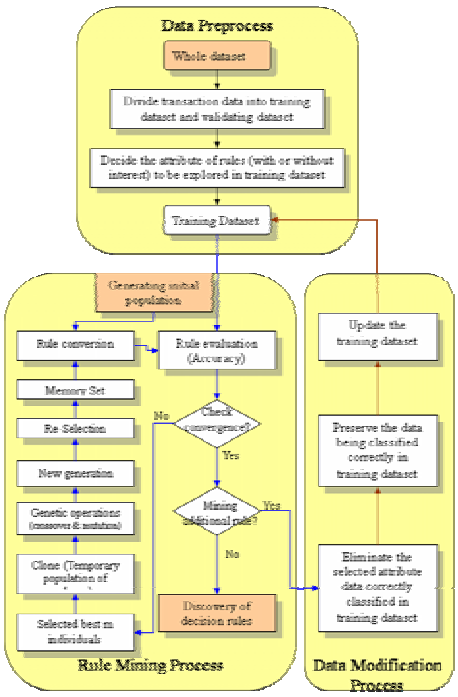

Fig.1. The IA based mining approach

2.2 The Rule Representation Mechanism

Before forming a prediction rule, the most significant predictors have to be decided; otherwise, the insignificant predictors become the noise which may worsen the prediction accuracy or even mislead the wrong evolution of GA and cause an unreasonable prediction. However, the number of prediction variables in many literatures are decided in advance but not determined by using the statistical concept of the best subset [11] [17]. Without considering the most significant predictors for building the prediction model, the insignificant variables in the decision model worsen the prediction accuracy. For this reason, a special rule representation mechanism of chromosome is proposed for deciding the best subset from the decision variables and decision rule simultaneously. The rule representation mechanism is illustrated in Fig.2.

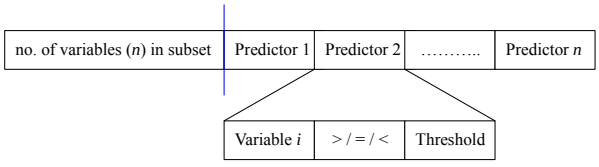

Fig.2. The general rule structure representation.

The general form of the rule generated by the IAs based approach is like an ***If-Then*** rule: ***If*** [variable i is ≥/< threshold i] ***Then*** prediction is belong to the specified category, where $i \in I$ (best subset of

variables). The range of the threshold value is based on the data source's range. Above all, the key issue of the proposed rule representation is to determine the best subset of the prediction variables for each of mining rules.

In our IAs implementation, ***If-Then*** rule is represented by a binary string (see Fig.2). Each string firstly divided in to two main substrings: the first binary substring (*integer number string*) representing the number of predictors in the best subset will be decoded into an integer (*n*). Any binary string can be decoded into real numbers [14] and then they can be rounded up to be the integer numbers. While the value of *n* is obtained, the second binary substring (*real number substring*) is then divided into *n* substrings. For each of *n* substrings, again, is to be divided into three sections: the three sections are the significant predictor, inequality/equality symbols (≥/<) and the value of threshold respectively. So, the general form of the *If-Then* rule can be represented by the proposed binary string.

Furthermore, above general rule structure is shown in Table 1. Subset regression has been applied so as to decide the significant predictors. The maximum number of predictors in the best subset for each rule can be up to total number of predictors.

2.3 Evaluation

An individual with a higher fitness function value has a higher probability to be selected to propagate a new generation. With the use of crossover and mutation operations, the parents with the best solutions from among the previous generation are selected to breed the next generation. The fitness measure is based on the prediction accuracy of the individual.

2.4 Genetic operators

The implementation of genetic operators including the crossover operator and mutation operator requires the selection of the crossover point(s) and mutation point(s) for each string under a predetermined crossover probability and mutation probability. The crossover operator provides a thorough search of the sample space to produce good solutions. The mutation operator performs random perturbations to selected solutions to avoid the local optimum. The crossover point is randomly selected and the parent strings are broken into substrings at that point. The offspring are then generated by swapping the substrings of parent 1 and parent 2. In the mutation operation, a mutation point is randomly selected and the binary digit is mutated with a specified probability.

2.4 Accuracy Assessment

With the epidemic application of geographic information systems retaining remotely sensed data as layers, the accuracy assessment of the map generated from any remotely sensed data has been even more crucial. One of the most important issues of the accuracy assessment is the ability to compare algorithms/approaches quantitatively. Particularly, the best approach is used to extract rules which can be developed as a computer model for prediction or classification of map data in decision-process like expert systems. Generally, the history of accuracy assessment can be divided into four developmental ages [13]. In the beginning, no real accuracy assessment was performed; rather, an "*it-looks-good*" mentality prevailed. The second age of accuracy assessment could be called the epoch of nonsite-specific assessment. The following age is called site-specific assessment in which actual places on the ground were compared to the same place on the map and a measure of overall accuracy was computed. The last epoch and current epoch of accuracy assessment could be called the age of the error matrix. The review of error matrix is described in more detail in [5]. In this study, surely, the error matrix will be used for verifying the feasibility and effectiveness of the proposed immune algorithm based approach for Land-Cover classification.

Once the error matrix is created correctly, it can be used as the beginning for a series of descriptive and analytical statistical means. For example, the most ordinary and easiest descriptive statistic is overall accuracy, which is obtained by dividing the total correct by the total number of sample units in the error matrix. Additionally, the total number of correct sample units in a category is divided by the total number of sample units of that category from the reference data. This accuracy measure is called the producer's accuracy. Moreover, if the total number of correct sample units in a category is divided by the total number of sample units that were classified into that category called user's accuracy [13] [18]. Besides the above descriptive ways, a discrete multivariate approach of use in accuracy assessment is Kappa [3], which is another used as another measure of agreement or accuracy. Although the value of Kappa is from +1 to -1, the positive value is expected since that represents the positive correlation between the remotely sensed classification and the reference data. The range for Kappa can be divided into three intervals [12]: Kappa > 0.8 represents strong correlation; 0.4 ≤ Kappa ≤ 0.8 represents moderate correlation; and kappa < 0.4 represents poor correlation.

2.5 Sampling Size

Sample size is a significant consideration when assessing the accuracy of remotely sensed data. The sample point collected is costly so that the sample size must be preserved to a minimum. However, oppositely, it is also critical to keep a large enough sample size that makes any analysis performed is statistically valid. Many researchers tried to provide the guidelines for determining the appropriate sample size [4] [9] [19]. A generally accepted rule of thumb is to use a minimum of 50 samples for each LC category in the error matrix [13]. If the number of categories is more than twelve, the minimum number of samples should be raised to 75 ~ 100 samples per category. In this paper, the number of LC categories is seven so that 50 samples for each category will be taken for purpose of accuracy measurement.

3 Experiments and Results

To verify the feasibility and effectiveness of the proposed IA based approach for Land-Cover classification, a remotely sensed data has been used in this study. For evaluating the performance of the proposed approach, the commercialized GIS software called ***ERDAS IMAGINE***® is used to make a comparison with the proposed approach.

The proposed IA based rule mining approach is implemented in MATLAB® on the Pentium-4 2.0 GHz PC with the following parameters: mutation rate = 0.01, crossover rate = 0.86 and the number of generations for evolving each rule was specified to be 150. The determination of IA's parameters is a significant problem for the IAs implementation. However, there is no formal way which can be used to solve the problem because various value-combinations of the parameters result to different characteristics as well as different performance of IAs. Therefore, the best values for the IAs' parameters are case-dependent and based upon the experience from preliminary runs.

The study area comprised the Huwei town (located in Yunlin, Taiwan), where the central location is at latitude:23.7051792, longitude: 120.4263219 (see Fig.3). The area covered approximately 11×14 km^2. This numerous geospatial data is provided by the Center for Space and Remote Sensing Research (CSRSR) at the National Central University, Taiwan. The original data was received and process by CSRSR from SPOT-5 (http://www.spot.com). Seven categories are included in the land cover map: grass, forest, road, bare, built-up water and shade. Due to there are seven categories in the land cover, fifty samples for each category are used for evaluating the accuracy. A generally accepted rule of thumb is to use a minimum of 50 samples for each LC category in the error matrix [13]. If the number of categories is more than twelve, the minimum number of samples should be raised to 75 ~ 100 samples per category.

Fig.3 Satellite-derived LC image

Based on the data samples for each category, the detailed decision rules mined for each category by the proposed approach are shown in Table 2. Each category with different number of rules, and each rule may contain different variables. For instance, there are three rules for the first category (Water) while there is only one rule extracted for the second category (Shade). Moreover, as shown the rules in Category 1 (water), two variables contained in rule 1 and three variables contained in rule 2. Compared with the previous methodologies in literature, the proposed approach is with two very distinctive features, which are: (1) the significant variables with the corresponding thresholds and inequality (≥/<) are decided simultaneously. (2) New rule is to be explored based on the misclassified data set. So, the classification accuracy can be enhanced by multiple rules. Table 3 is the transformed from Table 2 and explicated in the semantic way. As shown in Table3, while multiple rules are extracted from a category, these rules can be compounded as one general rule for categorizing the specific category. Each category is with its own general classification rule for categorizing the LC data.

The results of accuracies assessment for the proposed approach and the ERDAS IMAGINE® based on the maximum likelihood approach are presented in Table 4 and Table 5 respectively. As shown in the two tables, overall map accuracies are 88.86% and 85.14% for the proposed IA based mining approach and the maximum likelihood approach (ERDAS IMAGINE®). It indicates that the proposed have better classification accuracy. Moreover, the Kappa value obtained by using proposed approach is 0.87 which is also greater than 0.8236 provided by using ERDAS IMAGINE®. It means the proposed approach can provide much higher positive correlation between the remotely sensed classification and the reference data.

Additional two more tests are made and the three tests are concluded in Table 6. By the comparison of results in Table 6, it illustrates that the proposed IA based mining approach has higher correct classification accuracies in comparison with the uses of ERDAS IMAGINE® software. It is noted that the proposed approach incorporate non-spectral and collateral knowledge.

4 Conclusion

This paper provides an alternative approach by using Immune based mining approach to discover the useful decision rules automatically from the LC digital, remotely sensed data. By using the proposed IA based approach, for each LC class the significant predictors with the corresponding equality/inequality and threshold values are decided simultaneously, so as to generate the decision rules. Based on the extracted rules, a prediction model is then built to discriminating the all the categories LC image data with great precision. The results of the experiment show that the rules obtained by proposed method have higher accuracy than those by ERDAS IMAGINE®. Also, unlike the neural networks based classifier, the rules extracted by the proposed approach are explicit and easier to be understood. It indicates that the proposed mining approach is suitable approach for eliciting and representing experts' decision rules and thus it provides effective decision supports for solving the multiclass LC image classification problem. Further improvements may be obtained by incorporating the fuzzy inference approach in the proposed method, so as to have higher classification accuracy with fewer prediction rules.

Acknowledgements

The research is supported by National Science Council, Taiwan, under contract NSC 91-2213-E-150-005.

References:

[1] Azencott, R., Graffigne, C., and Labourdelte, C., Edge Detection and Segmentation of Textured Plane Images, *Proceedings of the International Conference on Analysis of Image and Patterns*, Berlin, 1995, pp. 75-88.

[2] Bauer, R. J., *Genetic Algorithms and Investment Strategies*, Wiley, New York, 1994.

[3] Cohen, J., A coefficient of agreement for nominal scales, *Educ. Pshchol. Meas.*, Vol.20, 1960, pp. 37-46.

[4] Congalton, R.G., A comparison of sampling schemes used in generating error matrices for assessing the accuracy of maps generated from remotely sensed data, *Photogram. Eng. Remote Sens.*, Vol.54, 1988, pp. 593-600.

[5] Congalton, R. and Green, K., Assessing the Accuracy of Remotely Sensed Data: Principles and Practices, Lewis, Boca Raton, FL., 1999.

[6] De Castro, L.N. and Timmis J., Artificial Immune Systems: A new Computational Intelligence Approach, Springer, New York, 2002.

[7] Derin, H., and Cole, W. S., Segmentation of Textured Images Using Gibbs Random Fields, *Computer Vision, Graphics and Image Processing*, Vol.35, 1986, pp. 72-98.

[8] Farmer, J.D., Packard, N.H., and Perelson, A.S., The immune system, adaptation, and machine learning, *Physica*, Vol.22D, 1986, pp. 187-204.

[9] Hord, R.M. and Brooner, W., Land use map accuracy criteria, *Photogram. Eng. Remote Sens.*, Vol.42, 1976, pp. 671-677.

[10] Jerne, N.K., The immune system, *Scientific America*, Vol.229, No.1, 1973, pp. 52-60.

[11] Kim, M.-J., & Han, I., The discovery of experts' decision rules from qualitative bankruptcy data using genetic algorithms, *Expert Systems with Applications*, Vol.25, No.4, 2003, pp. 637–646

[12] Landis, J. and Koch, G., The measurement of observer agreement for categorical data, *Biometrics*, Vol. 33, 1977, pp. 159-174.

[13] Lunetta, R.S., and Lyon, J.G., *Remote Sensing and GIS Accuracy Assessment*, *CRC Press*, NY., 2004.

[14] Michalewicz, Z., *Genetic algorithm =Data structures + Evolution programs.* Springer, Berlin, 1994.

[15] Miller, D.M., Kaminsky, E.J., and Rana, S., Neural network classification of remote-sensing data, *Computers and Geosciences*, Vol.21, No.3, 1995, pp. 377-386.

[16] Pal, M. and Mather, M., Decision tree based classification of remotely sensed data, *Proceedings of the 22nd Asia Conference on Remote Sensing*, Singapore, 2001.

[17] Shin, K.-S., & Lee, Y.-J., A genetic algorithm application in bankruptcy prediction modeling, *Expert Systems with Applications*, Vol.23, No.3, 2002, pp. 321–328.

[18] Story, M. and Congalton, R., Accuracy assessment: a user's perspective, *Photogram, Eng. Remote Sens.*, Vol.52, 1986, pp. 397-399.

[19] Tortora, R.A., A note on sample size estimation for multinomial populations, *Am. Stat.*, Vol.32, 1978, pp 100-102.

[20] Weissman, I.L. and Cooper, M.D., How the immune system develops, *Scientific American*, Vol.269, No.3, 1993, pp. 33-40.

Table 1. A general rule structure

Best Subset	**Variable 1**		**Variable 2**		...		Variable_*n*		Category
Significant **V**ariable	SP_{1i}		SP_{2i}		...		SP_{ni}		
Greater or **E**qual/ **L**ess than (≥ / <)	GEL_{1j}	***AND***	GEL_{2j}	***AND***	...	***AND***	GEL_{nj}	***THEN***	***Category K***
Threshold	T_1		T_2		...		T_n		

Note: i = variable number ; j =1: less than(<); 2: grater than or equal to (≥)

(NOTE: Table 2 is behind Table 3 and Table 4 on next page)

Table 3. The summarized rules for LC classification.

	Rule(s) for each category	Finalized Decision Rule
$Rule_{w-1}$	***If*** B ≥ 54 ***And*** G < 55	***IF*** $Rule_{w-1}$ ***OR*** $Rule_{w-2}$ ***OR*** $Rule_{w-3}$ ***THEN*** Water
$Rule_{w-2}$	***If*** R < 63 ***And*** G ≥ 47 ***And*** B≥ 63	
$Rule_{w-3}$	***If*** R < 79 ***And*** B ≥ 63	
$Rule_{s-1}$	***If*** B < 51 ***And*** G < 52	***IF*** $Rule_{s-1}$ ***THEN*** Shade
$Rule_{r-1}$	***If*** B ≥ 63 ***And*** G ≥ 74 ***And*** R < 95	***IF*** $Rule_{r-1}$ ***OR*** $Rule_{r-2}$ ***OR*** $Rule_{r-3}$ ***THEN*** Road
$Rule_{r-2}$	***If*** G < 83 ***And*** B ≥ 73	
$Rule_{r-3}$	***If*** B ≥ 68 ***And*** R < 103	
$Rule_{g-1}$	***If*** R < 87 ***And*** R ≥ 69	***IF*** $Rule_{g-1}$ ***THEN*** Grass
$Rule_{f-1}$	***If*** G < 206 ***And*** G ≥ 47 ***And*** R < 63	***IF*** $Rule_{f-1}$ ***OR*** $Rule_{f-2}$ ***THEN*** Forest
$Rule_{f-2}$	***If*** R < 68 ***And*** G ≥ 58	
$Rule_{b-1}$	***If*** B< 66 ***And*** R ≥ 92	***IF*** $Rule_{b-1}$ ***OR*** $Rule_{b-2}$ ***OR*** $Rule_{b-3}$ ***THEN*** Bare
$Rule_{b-2}$	***If*** G < 73 ***And*** R ≥ 100	
$Rule_{b-3}$	***If*** R ≥ 101 ***And*** G < 75	
$Rule_{bd-1}$	***If*** R ≥ 109	***IF*** $Rule_{bd-1}$ ***OR*** $Rule_{bd-2}$ ***OR*** $Rule_{bd-3}$ ***THEN*** Build-up
$Rule_{bd-2}$	***If*** B < 61 ***And*** R ≥ 100	
$Rule_{bd-3}$	***If*** G < 206 ***And*** B< 63 ***And*** G ≥ 74	

Table 4. Error Matrix by using proposed immune based mining approach.

Land-Cover Class	water	shade	grass	forest	road	bare	built-up	Totals
Water	49	0	0	1	0	0	0	50
Shade	0	50	0	0	0	0	0	50
Grass	0	0	48	0	1	0	1	50
Forest	0	0	0	50	0	0	0	50
Road	0	0	0	0	32	11	7	50
Bare	0	0	0	0	2	47	1	50
Built-up	0	0	6	6	0	3	35	50
Totals	49	50	54	57	35	61	44	350

Land-Cover Class	Reference Totals	Classified Totals	Number Correct	Producer Accuracy	User Accuracy
Water	49	50	49	100.00%	98.00%
Shade	50	50	50	100.00%	100.00%
Grass	54	50	48	88.89%	96.00%
Forest	57	50	50	87.72%	100.00%
Road	35	50	32	91.43%	64.00%
Bare	61	50	47	77.05%	94.00%
Built-up	44	50	35	79.55%	70.00%
Totals	350	350	311		

Note: Overall Classification Accuracy = 88.86%; Overall Kappa Statistics = 0.8700

Table 2. Rules extracted by using proposed approach

Classification Rules	Rule Operators	Variable 1	Variable 2	Variable 3	Category
$Rule_{w-1}$	Significant Variable(s)	3	2		Water
	≥ / <	2	1		
	Threshold	54	55		
$Rule_{w-2}$	Significant Variable(s)	1	2	3	
	≥ / <	1	2	2	
	Threshold	63	47	63	
$Rule_{w-3}$	Significant Variable(s)	1	3		
	≥ / <	1	2		
	Threshold	79	63		
$Rule_{s-1}$	Significant Variable(s)	2	3		Shade
	≥ / <	1	1		
	Threshold	53	51		
$Rule_{r-1}$	Significant Variable(s)	1	2	3	Road
	≥ / <	1	2	2	
	Threshold	95	74	63	
$Rule_{r-2}$	Significant Variable(s)	2	3		
	≥ / <	2	3		
	Threshold	83	73		
$Rule_{r-3}$	Significant Variable(s)	1	3		
	≥ / <	1	2		
	Threshold	68	103		
$Rule_{g-1}$	Significant Variable(s)	1	1		Grass
	≥ / <	1	2		
	Threshold	87	69		
$Rule_{f-1}$	Significant Variable(s)	1	2	2	Forest
	≥ / <	1	1	2	
	Threshold	63	206	47	
$Rule_{f-2}$	Significant Variable(s)	1	2		
	≥ / <	1	2		
	Threshold	68	58		
$Rule_{b-1}$	Significant Variable(s)	1	2		Bare
	≥ / <	2	1		
	Threshold	92	66		
$Rule_{b-2}$	Significant Variable(s)	1	2		
	≥ / <	2	1		
	Threshold	100	73		
$Rule_{b-3}$	Significant Variable(s)	1	2		
	≥ / <	2	1		
	Threshold	101	75		
$Rule_{bd-1}$	Significant Variable(s)	1			Build-up
	≥ / <	2			
	Threshold	109			
$Rule_{bd-2}$	Significant Variable(s)	1	3		
	≥ / <	2	1		
	Threshold	100	61		
$Rule_{bd-3}$	Significant Variable(s)	2	2	3	
	≥ / <	2	1	1	
	Threshold	74	206	63	

NOTE: **Significant Variable**：1~3 represents the three waves (R,G.,B) of image fusion respectively.

≥ / < : 1~2; 1 represents (<), 2 represents (≥);

Category: 1~7; 1 (Water), 2 (Shade), 3 (Road), 4 (Grass), 5 (Forest), 6 (Built-up), 7 (Bare).

Table5. Error Matrix by using ERDAS IMAGINE®.

Land-Cover Class	water	shade	grass	forest	road	bare	built-up	Totals
Water	38	0	4	1	0	0	0	50
Shade	1	77	9	4	0	1	2	50
Grass	1	0	33	0	0	0	0	50
Forest	0	0	3	54	0	0	0	50
Road	0	0	2	0	42	12	7	50
Bare	0	0	0	0	1	36	1	50
Built-up	0	0	1	0	1	1	18	50
Totals	40	77	52	59	44	50	28	350

Land-Cover Class	Reference Totals	Classified Totals	Number Correct	Producer Accuracy	User Accuracy
Water	40	43	38	95.00%	88.37%
Shade	77	94	77	100.00%	81.91%
Grass	52	34	33	63.46%	97.06%
Forest	59	57	54	91.53%	94.74%
Road	44	63	42	95.45%	66.67%
Bare	50	38	36	72.00%	94.74%
Built-up	28	21	18	64.29%	85.71%
Totals	350	350	298		

Note: Overall Classification Accuracy = 85.14%; Overall Kappa Statistics = 0.8236

Table 6. The comparisons of map accuracies.

	TEST 1		TEST 2		TEST 3	
	Overall Classification Accuracy	Overall Kappa Statistics	Overall Classification Accuracy	Overall Kappa Statistics	Overall Classification Accuracy	Overall Kappa Statistics
IA based mining approach	88.86%	0.8700	88.29%	0.8633	88.86%	0.8700
Maximum likelihood approach by using ERDAS IMAGINE®	85.14%	0.8236	85.71%	0.8317	86.86%	0.8467