Generating Rules from Trained Neural Network using FCM for Satellite Images Classification

Dr. Nabeel Hashem Kaghed 1*, Samaher Hussein Ali 2**

* Department of Computer Science, University of Babylon, Iraq

nhkaghed@yahoo.com

** *Department of Computer Science, University of Babylon, Iraq*

Samaher_hussein@yahoo.com

Abstract: This paper presents a methodology for image classification of both visual and spectral data with the use of hybrid techniques represented by soft computing to classify objects from satellite images. In this article the searching capability of a Fuzzy c-Means Model (FCM) has been exploited for automatically evolving the number of clusters as well as proper clustering of data set. This paper concerns with classifying five kinds of objects (Residential area, Agriculture area, Road, River and Tennis stadium). Accordingly. The database which describes that objects depending on their attributes were built, then, c-means clustering algorithm to the Fuzzy c-Means Algorithm (FCMA) was used. After that, two types of features for each cluster were extracted. Then, unsupervised classification algorithm (Kohonen winner-take-all network) determines the class under which each feature vector belongs to was used. At the last stage, (IF-Then rule) to form several rules that govern each class attributes were used. The proposed methodology provides fast and adaptive learning for image classification.

Key words: Kohonen Winner-Take-All Network, Satellite Image Classification, Spectral Imaging, Visual Description.

INTRODUCTION

The remote sensing technology is currently being offered a wide variety of digital imagery that covers most of the Earth's surface. This up-to-date image data is a promising tool for producing accurate land cover maps. To maximize the benefit of such data, automatic and efficient classification methods are needed. To achieve this objective, pixel-based classification has been extensively used for the past years. Currently the prospects of a new classification concept, object-based classification, are being investigated. Recent studies have proven the superiority of the new concept over traditional classifiers [1]. The new concept's basic principle is to make use of important information (shape, texture and contextual information) that is present only in meaningful image object and their mutual relationships.

In order to obtain image objects, classification software is developed by ours. It gives convenient environment to nonspecialists, because operated automatically. Features database is constructing for automatic land cover classification. Features database has information of five class (Residential area, Agriculture area, Road, River and Tennis stadium) features in Landsat images. Proposed method will be higher classification accuracy than

That of traditional pixel-based supervised classification and gives convenient environment to non-specialist users. The object classification concept is that important semantic information necessary to interpret an image is not represented in single pixels, but in meaningful image objects and their mutual relations. Image analysis is based on contiguous, homogeneous image regions that are generated by initial image segmentation. Connecting all the regions, the image content is represented as a network of image objects. These image objects act as the building blocks for the subsequent image analysis. In comparison to pixels, image objects carry much more useful information. Thus, they can be characterized by far more properties such as shape, texture, shadow, association, pattern [2].

The reminder of this paper is organized as follows. Section 2 reviews soft computing. Section 3 describes the proposed system for discovering classification rules. Section 4 reports computational results. Finally section 5 presents the conclusions.

1. Soft Computing

In traditional hard computing, the prime desiderata are precision, certainty, and rigor. By contrast, in soft computing the principal notion is that precision and certainty carry a cost and that computation, reasoning, and decision-making should exploit (wherever possible) the tolerance for imprecision, uncertainty, approximate reasoning, and partial truth for obtaining low-cost solutions. This leads to the remarkable human ability of understanding distorted speech, deciphering sloppy handwriting, comprehending the nuances of natural language, summarizing text, recognizing and classifying images, driving a vehicle in dense traffic and, more generally, making rational decisions in an environment of uncertainty and imprecision. The challenge, then, is to exploit the tolerance for imprecision by devising methods of computation that lead to an acceptable solution at low cost. This, in essence, is the guiding principle of soft computing [3].

Soft computing is a consortium of methodologies that works synergetically and provides in one form or another flexible information processing capability for handling real life ambiguous situations. Its aim is to exploit the tolerance for imprecision, uncertainty, approximate reasoning, and partial truth in order to achieve tractability, robustness, and low-cost solutions. The guiding principle is to devise methods of computation that lead to an acceptable solution at low cost by seeking for an approximate solution to an imprecisely/precisely formulated problem.

The term Soft Computing (SC) refers to a family of computing techniques that originally comprise five different partners: fuzzy logic, evolutionary computation, neural networks, probabilistic reasoning and hybrid system [4].

1.1. Fuzzy c-Means Algorithm

Partitional clustering essential deals with the task of partitioning a set of entities into a number of homogeneous clusters, with respect to a suitable similarity measure. Due to the fuzzy clustering methods have been developed following the general fuzzy set theory strategies outlined by Zadeh [5]. The main difference between the traditional hard clustering and fuzzy clustering can be stated as follows. While in hard clustering an entity belongs only to one clusters, in fuzzy clustering entities are allowed to belong to many clusters with different degrees of membership.

The fuzzy c-means algorithm [6] is one of the most widely used methods in fuzzy clustering. It is based on the concept of fuzzy c-partition, introduced by Ruspini [7], summarized as follows.

Let $X=\{x_{1,\dots,}x_n\}$ be a set of given data, where each data point x_k (k=1,…,n) is a vector in $R^{P,}$ U_{cn} be a set of real c×n matrices, and c be an integer, 2≤c<n. Then, the fuzzy c-partition space for X is the set

$$M_{fcn}=\{U \in U_{cn} : u_{ik} \in [0,1] \quad \sum_{i=1}^{c} u_{ik} = 1,\ 0 < \sum_{k=1}^{n} u_{ik} < n\} \tag{1}$$

Where u_{ik} is the membership value of x_k in cluster i (i=1,…,c).

The aim of the FCM algorithm is to find an optimal fuzzy c-partition and corresponding prototypes minimizing the objective function

$$J_m(U,V;X)=\sum_{k=1}^{n}\sum_{i=1}^{c}(u_{ik})^m \left\|x_k - v_i\right\|^2 \tag{2}$$

In (2), $V=(v_1,v2,\dots,vc)$ is a matrix of unknown cluster centers(prototypes) $v_i \in \mathscr{R}^p$, $\|.\|$ is the Euclidean norm, and the weighting exponent m in [1,∞) is a constant that influences the membership values.

To minimize criterion J_m, under the fuzzy constraints define in (1), the FCM algorithm is define as an alternating minimization algorithm (cf.[3] for the derivations), as follows.

Choose a value for c,m and ε , a small positive constant; then, generate randomly a fuzzy c-partition U^0 and set iteration number t=0. A two-step iterative process works as follows. Given the membership values $u_{ik}^{(t)}$, the cluster centers $v_i^{(t)}$,(i=1,…,c) are calculated by

$$V_i^{(t)} = \frac{\sum_{k=1}^{n}(u_{ik}^{(t)})^m . x_k}{\sum_{k=1}^{n}(u_{ik}^{(t)})^m} \tag{3}$$

Given the new cluster centers $v_i^{(t)}$, update membership values $u_{ik}^{(t)}$:

$$u_{ik}^{(t+1)} = \left[\sum_{j=1}^{c}\left(\frac{\left\|x_k - v_i^{(t)}\right\|^2}{\left\|x_k - v_j^{(t)}\right\|^2}\right)^{\frac{2}{m-1}}\right]^{-1} \tag{4}$$

The process stops when $\left|U^{(t+1)} - U^{(t)}\right| \le \varepsilon$, or a predefined number of iterations is reached.

Several algorithms for clustering data when the number of clusters is known a priori are available in the literature viz., K-means, branch and bound procedure, maximum likelihood estimate technique, graph theoretic approaches.

In most real life situations the number of clusters in a data set is not known a priori. The real challenge in this situation is to be able to automatically evolve a proper value of C as well as providing the appropriate clustering. In this article, we propose a FCM based clustering technique which can automatically evolve the appropriate clustering of a data set.

1.2. Kohonen Winner-Take-All Network

Artificial neural networks are processing systems analogous to biological neural networks, presenting neurons, axons, dendrites, neural layers, transfer functions, and so on. Their paradigms fall in three main categories: supervised, reinforced and self-organized.

This classification takes into account the amount of data needed for the training phase. Supervised networks use a previous knowledge about the desired outputs, in such a way that the error between the actual input and expected output is a suitable parameter. Reinforced networks rely on the measure of the overall error but do not need their exact output. Self-organizing networks determine by themselves the internal weight representation for the presented input data and do not need supervision [8].

Kohonen winner-take-all network [Kohonen (1989)], also known as self-organizing neural networks, are networks which incorporate a topology scheme, i.e., take into account the topological structure among units.

The input signals are *n*-tuples and there is a set of *c* cluster units (we use fuzzy c-means clustering algorithm to determine c as illustrated a above). Each input is fully connected to all units, which respond differently to the input pattern. At each step in the training phase, the cluster unit with weights that best match the input pattern is elected the winner (usually in a minimum Euclidean distance sense).

This winning unit and a neighborhood around it are then updated in such a way that their internal weights be closer to the presented input. The adopted updating factor is not equal for all neurons, but stronger near the winning unit, decreasing for more distant units. Figure 1 shows the basic structure of kohonen winner-take-all. There are input components connected to all cluster units. The cluster units can assume any spatial distribution, which are usually linear or planar arrays.

Weights are associated to each connection. With time, the gain factor must be reduced and also the neighborhood decreases in size.

During the learning phase the node weights are changed in an ordered manner, in such a way that the main image features tend to be organized according to a topological distribution in the network. Adjacent nodes respond similarly, while distant nodes respond diversely.

The convergence of the features in the winner-take-all occurs considering some limitations on the gain factor while updating the weights.

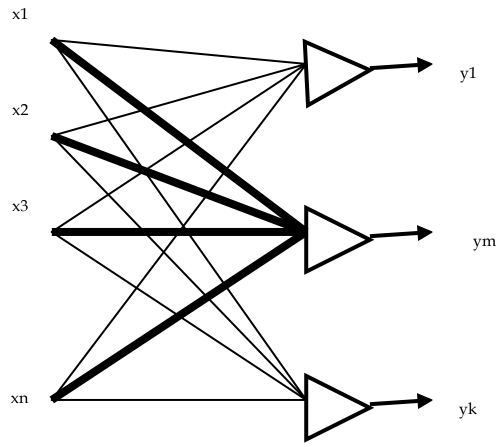

Figure (1): Structure of Kohonen Winner-Take All

2. The Structure of Proposed System

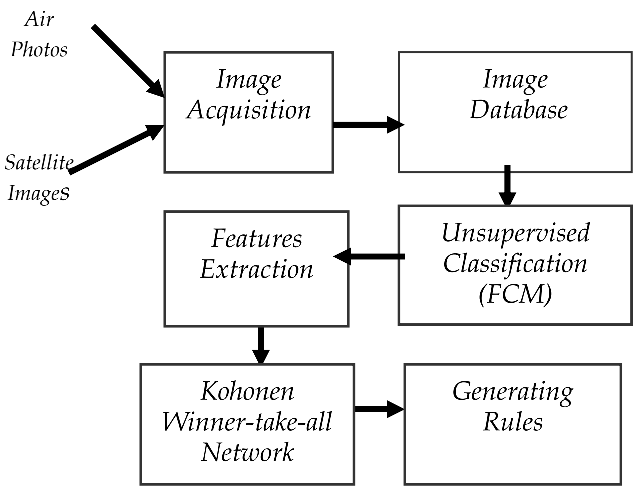

Figure (2): A Block Diagram of overall System

2.1. Image Acquisition

In this stage image (air photo or satellite image) acquisition that require classification their objects by proposed system that can be performed by using one of the remote sensing devices such as satellite or camera carried by air-planes.

2.2. Image Database

According to the proposed system, one needs to describe each object in the image depending on their features (attributes), where digital images were segmented into different objects, which are labeled with appropriate category names. As explained in Table 1.

The majority of inductive learning algorithms for the classification task discover rules (or another kind of knowledge representation) involving only the original attributes of the data being mined. In addition, the majority of rule induction methods analyze the data on a one-attribute-at-a-time basis. Hence, the performance of these methods is considerably limited by the predictive power of individual attributes, so that these methods do not cope very well with the problem of attribute interaction.

The goal of an attribute construction method is to construct new attributes out of the original ones, transforming the original data representation into a new one where regularities in the data are more easily detected by the classification algorithm, which tends to improve the predictive accuracy of the latter.

Attribute construction methods can be roughly divided into two groups, with respect to the construction strategy: hypothesis-driven methods and data-driven methods.

Hypothesis-driven methods construct new attributes out of previously-generated hypotheses (discovered rules or another kind of knowledge representation). In general they start by constructing a hypothesis, for instance a decision tree, and then examine that hypothesis to construct new attributes. These new attributes are then added to the set of original attributes, and a new hypothesis is constructed out of this extended set of attributes. The new hypothesis is used to generate new attributes, and so on. This process is repeated until a given stopping criterion is satisfied, such as a satisfactory extended set of attributes has been found. Note that the performance of this strategy is strongly dependent on the quality of the previously-discovered hypotheses.

By contrast, data-driven methods do not suffer from the problem of depending on the quality of previous hypotheses. They construct new attributes by directly detecting relationships in the data.

The process of attribute construction in data base can also be roughly divided into two approaches, namely the interleaving approach and the preprocessing approach.

In the preprocessing approach the process of attribute construction is independent of the inductive learning algorithm that will be used to extract knowledge from the data. In other words, the quality of a candidate new attribute is evaluated by directly accessing the data, without running any inductive learning algorithm. In this approach the attribute construction method performs a preprocessing of the data, and the new constructed attributes can be given to different kinds of inductive learning methods.

By contrast, in the interleaving approach the process of attribute construction is intertwined with the inductive learning algorithm. The quality of a candidate new attribute is evaluated by running the inductive learning algorithm used to extract knowledge from the data, so that in principle the constructed attributes' usefulness tends to be limited to that inductive learning algorithm.

In this paper we follow the data-driven strategy and the preprocessing approach, mainly for two reasons.

First, using this combination of strategy/approach the constructed attributes have a more generic usefulness, since they can help to improve the predictive accuracy of any kind of inductive learning algorithm. Second, an attribute construction method following the preprocessing approach tends to be more efficient than its interleaving counterpart, since the latter requires many executions of an inductive learning algorithm.

Table (1): Explain image database

Object	Residential	Road	Agriculture	River	Tennis- Stadium
Pattern	Uniform	Non-uniform	Non-uniform	Near-uniform	Uniform
Shadow	Yes	No	Yes	No	No
Texture	Rough	Rough	Rough	Smooth	Rough
Shape	Known , Unknown	Unknown	Unknown	Unknown	Known
Associative	School, Road, Playgrounds	Car	Water , Vegetation	Vegetation, Bridges	Scrolls, Car-attitude
Hue	$[H_{min},H_{max}]$	$[H_{min},H_{max}]$	$[H_{min},H_{max}]$	$[H_{min},H_{max}]$	$[H_{min},H_{max}]$
Saturation	$[S_{min},S_{max}]$	$[S_{min},S_{max}]$	$[S_{min},S_{max}]$	$[S_{min},S_{max}]$	$[S_{min},S_{max}]$
Intensity	$[I_{min},I_{max}]$	$[I_{min},I_{max}]$	$[I_{min},I_{max}]$	$[I_{min},I_{max}]$	$[I_{min},I_{max}]$

2.3. Unsupervised Classification (FCM)

In this section, an attempt has been made to use fuzzy c-means clustering algorithms for automatically clustering an image data set. This includes determination of number of clusters as well as appropriate clustering of the data. The methodology is explained first followed by the description of the implementation results. The main befit from using fuzzy c-means clustering technique is to find a number of clusters and to provide the best seed for each cluster. We can explain the Hierarchical clustering as figure (3).

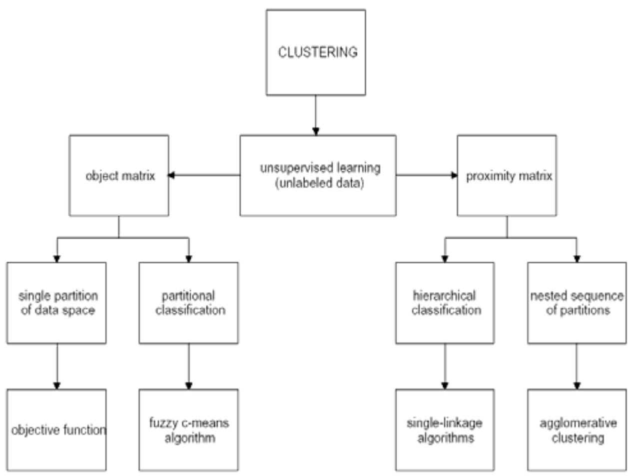

Figure (3): Hierarchical Clustering

Example: Butterfly Data Set

The data set **M** consists of 15 points in the plane:

j	1	2	3	4	5	6	7	8
m_j	(0,0)	(0,2)	(0,4)	(1,1)	(1,2)	(1,3)	(2,2)	(3,2)

j	9	10	11	12	13	14	15	
m_j	(4,2)	(5,1)	(5,2)	(5,3)	(6,0)	(6,2)	(6,4)	

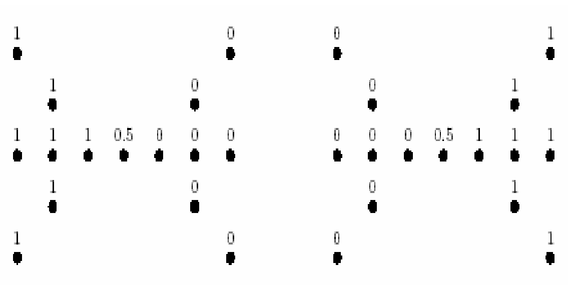

Figure (4): Fuzzy c-means clustering of the Butterfly data set. w = 1.25

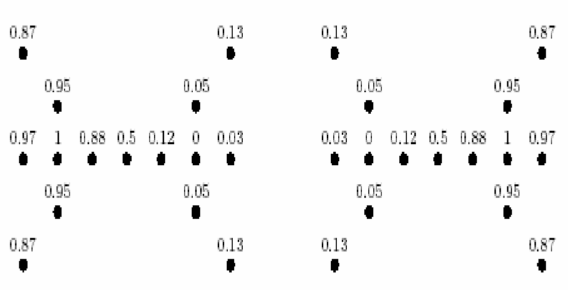

Figure (5): Fuzzy c-means clustering of the Butterfly data set. w = 2.

2.4. Features Extraction

In order to make an automated classification between different objects, some classifying features have to be extracted from the objects. In this part of this work we introduce two feature types, visual features and spectrum –based features.

2.4.1. Visual Description

Recognizing objects is the key to interpretation and information extraction. Observing the differences between objects and their backgrounds involves comparing different targets based on any, or all, of the visual elements of (shape, pattern, texture, shadow, and association). Visual interpretation using these elements is often a part of our daily lives, whether we are conscious of it or not Identifying targets in remotely sensed images based on these visual elements allows us to further interpret and analyze. The nature of each of these interpretation elements is described below [9].

- **Shape** refers to the general form, structure, or outline of individual objects.
- **Pattern** refers to the spatial arrangement of visibly discernible objects. Typically an orderly repetition of similar tones and textures will produce a distinctive and ultimately recognizable pattern.
- **Texture** refers to the impression of "smoothness" or "roughness" of image features is caused by the frequency of change of tone in photographs.
- **Shadow** is also helpful in interpretation as it may provide an idea of the profile and relative height of a target or targets which may make identification easier.
- **Association** takes into account the relationship between other recognizable objects or features in proximity to the target of interest. The identification of features that one would expect to associate with other features may provide information to facilitate identification.

2.4.2. Features based on the spectrum

The light, which is reflected by the object, forms a spectrum. The visible part of the spectrum of light is located between 0.4 and 0.7 um. Characterizations of the light are related to science of color. All colors are seen as variable combination of the three primary colors, red(R), green (G), and blue (B).

Combination of three primary colors is useful in spectrum measurement, when the visible part of the spectrum is considered. However to extract the spectrum information from the object, the consideration should be done in HSI(Hue, Saturation, Intensity)model. In the HSI-model hue (H) describes pure color in terms of the dominant wavelength (e.g., red, orange, yellow, ect.),

$$H = COS^{-1}\left\{\frac{\frac{1}{2}\left[(R-G)+(R-B)\right]}{\left[(R-G)^2+(R-B)(G-B)\right]^{1/2}}\right\} \tag{5}$$

Whereas the saturation(S) gives the measure of degree to which a pure color is diluted by white light (e.g., pink is diluted red),

$$S = 1 - \frac{3 * \min(R, G, B)}{(R + G + B)} \tag{6}$$

Intensity (I) is decoupled from the color information of the object [10].

$$I = \frac{1}{3}(R + G + B) \tag{7}$$

2. 5. Training Kohonen Network

After we get the number of clusters expected in the image data set and the feature vectors that represent each cluster we can train kohonen winner-take all network.

The network is composed of an orthogonal grid of cluster units, each associated to several internal weights for the features data. The initial values for weights are set to small random values before the learning phase. When an input sample is presented to the network, a search is made to choose the winning unit *c*, as stated in (8). The input vector **x** at time *t* is compared to each of the weight vectors $\mathbf{m}_i \in R^n$ and the minimum Euclidean distance between the input signal and the neuron weights determines the closest match.

$$\|\mathbf{x}(t) - \mathbf{m}_c(t)\| = \min_i \{\|\mathbf{x}(t) - \mathbf{m}_i(t)\|\} \qquad (8)$$

Then, the weights are updated considering a circular neighborhood around the winning unit, as showed in (9). A scalar gain function $\alpha = \alpha$ (t) is employed to establish how the updating will be done inside the neighborhood *Nc*. Outside this neighborhood the weights remain unchanged. The gain function values are $0 < \alpha < 1$ and decreases with time.

$$\mathbf{m}_i(t+1) = \mathbf{m}_i(t) + \alpha(t)[\mathbf{x}_i(t) - \mathbf{m}_i(t)] \quad \forall i \in N_c(t) \qquad (9)$$

2.6. Generating Rules

There are several methods to specify a certain class. One of these methods depends on use IF-Then Rule, such as images are described in terms of many characteristics and a rule is given which specifies the attributes that determine membership of the target class [11]. This paper proposes a soft computing (SC) system for discovering simple classification rules in the following format: IF (a-certain combination -of-attribute values-is-satisfied) THEN (predict-a-certain-class). Each pattern represents a set of these IF-THEN rules. This rule format has the advantage of being intuitively comprehensible for the user. Hence, he/she can combine the knowledge contained in the discovered rules with his/her own knowledge, in order to make intelligent decisions about the target classification problem.

3. Experiment

3.1. A Sample of Image Data for the Experiment

A System Pour l'Observation de la Terre (SPOT), data 2002, satellite image of Baghdad, Iraq was used for the classification. The SPOT image has 3 spectral bands sensing the red, green and blue portions of the electromagnetic spectrum.Typically; remote sensing data provides a large number of examples for each class. Figure (6) explains a study area.

Figure (6): The study area

3.2. Result of Test:

Building database that describes each object in image and the database used in our experiments consisted of 312 records and 17 attributes.

Apply fuzzy c-means clustering algorithm (FCM) to clustering image data set (i.e. find best seed for each cluster). Before this, we need to determine some of parameters relate to FCM as explains in procedure fuzzy c-means (in this experiment, w=2, ε*=0.01, n=(width*height), l=20).*

From the definition of fuzzy clustering and traditional hard clustering methods we found main difference between them can be stated as follows. While in hard clustering an entity belongs only to one clusters, in fuzzy clustering entities are allowed to belong to many clusters with different degrees of membership. And this point is consider a very important in this work because the complex natural of satellite images and overlapping the objects of these images therefore when applying the fuzzy c-means clustering model in this work we obtain of good clustering result through benefiting from the membership value for each cluster.

Calculate features for each cluster. Then compare these features with image database.

Train Kohonen Winner-take-all neural network on the feature vectors for each cluster to determine the correct class for each feature vector. Before this, we need to determine some of parameters relate to kohonen network (maximum Number of epochs=1000, Learning rate=0.7, Momentum rate=0.01. As a result the network enables from classification of all image objects.

Use IF-Then Rule to form several rules that govern each class attributes. This step makes the system more precise, as follows

IF(hue=[0.89429,0.54728])and(saturation=[0.01835,1])and(intensity=[89.66667,172.6666])and(pattern_uniform=1)and(pattern_near_uniform=0)and(pattern_non_uniform=1)and(shape_known=1)or(shape_unknown=1)and(texture_smooth=0)and(texture_rough=1)and(shadow_yes=1)and(shadow_no=0)and(associative_school=1)or(associative_road=1)or(associative_playground=1)and(associative_car=0)and(associative_water=0)and(associative_vegetation=0)and(associative_bridg es=0)and(associative_scrolls=0)and(associative_car_attitude=0) **Then class Residential Area**.

IF(hue=[0.46273,0.53991])and(saturation=[0.04294,1])and(intensity=[78,130])and(pattern_uniform=0) and(pattern_near_uniform=0)and(pattern_non_uniform=1)and(shape_known=0)and(shape_unknown=1)and (texture_smooth=0)and(texture_rough=1)and(shadow_yes=0)and(shadow_no=1)and(associative_school=0) and(associative_road=0)and(associative_playground=0)and(associative_car=1)and(associative_water=0)and (associative_vegetation=0)and(associative_bridges=0) and(associative_scrolls=0)and(associative_car_attitude=0) **Then class Road**.

IF(hue=[0,0.51468])and(saturation=[0.02016,1])

and(intensity=[82.66667,255])and(pattern_uniform=0) and(pattern_near_uniform=0)and(pattern_non_uniform=1)and(shape_known=0)and(shape_unknown=1)and (texture_smooth=0)and(texture_rough=1)and(shadow_yes=1)and(shadow_no=0)and(associative_school=0) and(associative_road=0)and(associative_playground=0)and(associative_car=0)and(associative_water=1)or(associative_vegetation=1)and(associative_bridges=0)and(associative_scrolls=0)and(associative_car_attitude=0) **Then class Agriculture Area.**

IF(hue=[0.50550,0.53982])and(saturation=[0.08911, 0.67416])and(intensity=[89, 143]) and(pattern_uniform=0)and(pattern_near_uniform=1)and(pattern_non_uniform=0)and(shape_known=0)and(shape_unknown=1)and(texture_smooth=1)and(texture_rough=0) and(shadow_yes=0)and(shadow_no=1)and(associative_school=0)and(associative_road=0)and(associative_playground=0)and(associative_car=0)and(associative_water=0)and(associative_vegetation=1)and(associative_bridges=0)and(associative_scrolls=0)and(associative_car_attitude=0) **Then class River.**

By looking of these rules we can understand the physicality of each region and how we can benefit from it. Also these rules explains some of features for each region therefore if any person needs to use these regions in any project that not require from him /her to study the nature of region but need to look of these rules. As a result we can consider these rules are very simple and advantageous.

4. Conclusion

This paper introduces a fully automatic approach for the generating rules of remote sensing imagery and the result explains: The proposed method will be higher classification accuracy than that of traditional pixel-based supervised classification and gives convenient environment to non-specialist users. To receive an acceptable classification result, the training areas need to be spectrally separable. This can be done with clustering or expert knowledge, and the proposed methodology verifies this by applying Fuzzy c-means clustering. The randomly estimation of the number of the clusters that are found in the image may lead to error in classification process. Therefore, the proposed methodology can solve this problem by determining it automatically. Forming several rules that govern each class attributes by using IF-Then Rule format makes the system more precise because the features extracted from each class which are used to train Kohonen winner-take-all network are congruent to the conditions of this rule, while the resultant classes from Kohonen network are congruent to the actions of this rule. In other words the extracted rules are used to justify the inferred decisions.

REFERENCES

[1] Esch, T., Roth, A., Strunz, G. and Dech, S., 2003. *Object oriented classification of Landsat-7 data a for regional planning purpose*, The international archives of the photo grammetry, remote sensing and spatial information sciences, vol. XXXIV-7/W9, Regenburg.

[2] Baatz, M., Benz, U., Dehghani, S., Heynen, M., Holtje, A.,Hofmann, P., Lingenfelder, I.,Mimler, M., Sohlbach, M.,Weber, M. and Willhauck, G., 2002. *eCognition 3.0 USER GUID*, DEFINIENS imaging,

[3] Lotfi ,Z.,1994.*Fuzzy logic neural network ,and soft computing*,Commun.ACM,vol.37,pp.77-84.

[4] Ocordon.O and Herrera-Viedma. E.,2003. *Special issue on Soft Computing Applications to Intelligent Information Retrieval on the Internet*, International Journal of Approximate Reasoning ,vol.34,PP. 89-95.

[5] Zadeh.L, 1965. *Fuzzy Set" Information and control"*, 8, pp.338-352.

[6] Bezdek.J,1981.Pattern *Recognition with Fuzzy Objective Function Algorithms*, Plenum Press, New Youk.

[7] Ruspini.E, 1969.*A New Approach to Clustering, Information and Control*, 15, pp.22-32.

[8] Jander. M, Luciano. F,2002**.** *Neural-based color image segmentation and classification using self-organizing maps.*

[9] Andrew.M.,2004. *Fundamental of Remote Sensing*, Canada, Canada Central for Remote Sensing.

[10] Paolo.C., 2004.*Image Databases*.

[11] Sushmita.M and Yoichi.H.,2000. *Neuro-Fuzzy Rule Generation: Survery in Soft Computing Framework*, IEEE Trnasaction on Neural Network,Vol. 11,No. 3.