

Creates a New GIS Tool for Classifying Brain Tumor

Samira A. Hussain , Samar A. Kassir , Israa M. Hassoon

Abstract— The medical image processing technique for diagnosing diseases, is considered very important for human life. Pattern recognition is an essential part of any high-level image analysis systems. There are many practical applications of pattern recognition such as the recognition of handwritten letters, signature, fingerprints, face and others. A recognition system for tumor type in human brain is tackled in this work. This work displays a system that can diagnose some type(s) of tumor in the human brain by using the Magnetic Resonance Imaging (MRI) images.

Typical structure of a pattern recognition in this system consists of three steps: the first step is the construction of GIS system to represent the data in GIS as raster, vector and attribute database; the second step is feature extraction that extracts a vector of textural features for each image using second order statistics Gray Level Co-occurrence Matrix (GLCM) and calculate eight textural feature.; and the third step is classification that will recognize the input pattern with respect to the features extracted in previous step, and then the pattern is assigned to an appropriate class using a classifier, supervised K- Nearest Neighbor classification is implemented in proposed system to classify this type of tumor.

Index Terms— Gray-Level Co-occurrence Matrix (GLCM), Texture features, Geographic Information System (GIS).

1 INTRODUCTION

Developing a system for classification of images to any class with the use of their textures features in different direction of GLCM matrix . This association between local trained features values and recognized class of tumor as an example led to obtain a good results using this method. Another direction is extracting the texture feature for unknown image and let the system detect the type of this image using K-Nearest Neighbor. Conceptually, geographic data can be broken up in two elements: observation or entity and attribute or variable. GIS have to be able to manage both elements. Spatial component, the observations have two aspects in its localization, absolute localization based in a coordinates system and topological relationship referred. A GIS is able to manage both while computer assisted cartography packages only manage the absolute one. R. Methre et al. [17] investigated the texture retrieval using combination of local features of Haralick derived from one level discrete wavelet transform coefficients and global statistical features computed from three level wavelet transformed images. Y. Zhang et al. [20] proposed a hybrid classifier for polarimetric SAR images, the feature sets consist of span image, the H/A/ α decomposition algorithm, and the GLCM-based texture features. Then, a probabilistic neural

network (PNN) was adopted for classification. E. S. Flores et al [3] used GIS techniques to improve the classification capabilities of a feature extraction algorithm for land use/cover change detection in a deciduous forest environment.

2 METHODS AND MATERIALS

2.1 GEOGRAPHIC INFORMATION SYSTEMS (GIS)

A key characteristic of GIS is the linkage between geographic features represented on a map with attribute data that describe the geographic feature [14].

Within the GIS database a user can enter, analyze and manipulate data that is associated with some spatial element in the real world. The cartographic software of the GIS enables one to display the geographic information at any scale or projection and as a variety of layers which can be turned on or off. Each layer would show some different aspect of a place on the Earth. These layers could show things like a road network, topography, vegetation cover, streams and water bodies, and the distribution of annual precipitation received [2] .

2-2- GLCM AND TEXTURAL FEATURES

Gray-level co-occurrence matrix (GLCM) is the statistical method of examining the textures that considers the spatial relationship of the pixels. The GLCM functions characterize

the texture of an image by calculating how often pairs of pixel with specific values and in a specified spatial relationship occur in an image, creating a GLCM, and then extracting statistical measures from this matrix [15].

The matrix element $P(i, j | d, \theta)$ contains the second order statistical probability values for changes between gray levels i and j at a particular displacement distance d and at a particular angle θ . Instead of using the frequency values in a GLCM directly, it is common practice to normalize them to the range $[0, 1]$ to avoid scaling effects. Eight textural features are extracted in proposed system for each visible band from each land sample to be the vector of features which represented the entire land sample [7][16].

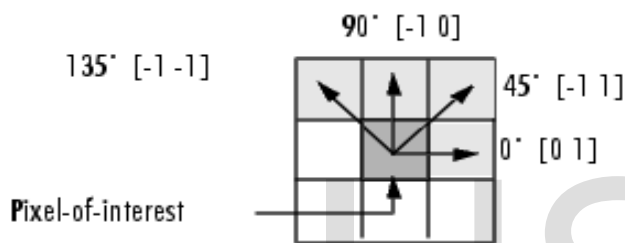


Fig.(1): Directions for generation of GLCM

2.3. K-NEAREST NEIGHBOR CLASSIFIERS

The intuition underlying Nearest Neighbor Classification is quite straightforward, examples are classified based on the class of their nearest neighbors. It is often useful to take more than one neighbor into account so the technique is more commonly referred to as k-Nearest Neighbor (k-NN) Classification where k nearest neighbors are used in determining the class. Since the training examples are needed at run-time, i.e. they need to be in memory at run-time, it is sometimes also called Memory-Based Classification. Because induction is delayed to run time, it is considered a Lazy Learning technique. Because classification is based directly on the training examples it is also called Example-Based Classification or Case-Based Classification [5].

2.4 TUMOR TYPES OF THE BRAIN

Body is made of many cells. Each cell has specific duty. The cells grow in the body and are divided to reproduce other cells. These divisions are very vital for correct functions

of the body. When each cell loses the ability of controlling its growth, these divisions are done without any limitations, and tumor emerges [8]. A tumor is a mass of tissue that grows out of control of the normal forces that regulates growth [9].

Brain is the kernel part of the body. Brain has a very complex structure. And the brain can be affected by a problem which causes change in its normal structure and its normal behavior. This problem is known as brain tumor. It is one of the major causes for the increase in mortality among children and adults [22].

The Brain Tumors can be classified as follows [6]:

1. Benign Tumor i.e. Non Cancerous Tumor

It is a type of tumor, which is noncancerous, meaning it does not spread or invade the surrounding tissue.

2. Malignant Tumor i.e. Cancerous Tumor

It is a type of tumor which is cancerous, meaning it spreads and invades the surrounding tissue. It is categorized as primary and secondary tumor.

- a. Primary Tumors

A primary tumors refer to a tumor or mass that is growing in the location where cancer originated. Most of them are usually successfully treated with techniques such as surgery.

- b. Secondary Tumor (Metastatic)

A secondary (Metastatic) brain tumor occurs when cancer cells spread to the brain from a primary cancer in another part of the body. Secondary tumors are about three times more common than primary tumors of the brain.

The number of collected images are 140 (20 image for each type of the six tumors and 20 image for the error case), they are gray level image with 8 bit (pixel value 0..255) and the type of them are: BMP, JPEG, and PNG. With various image sizes.

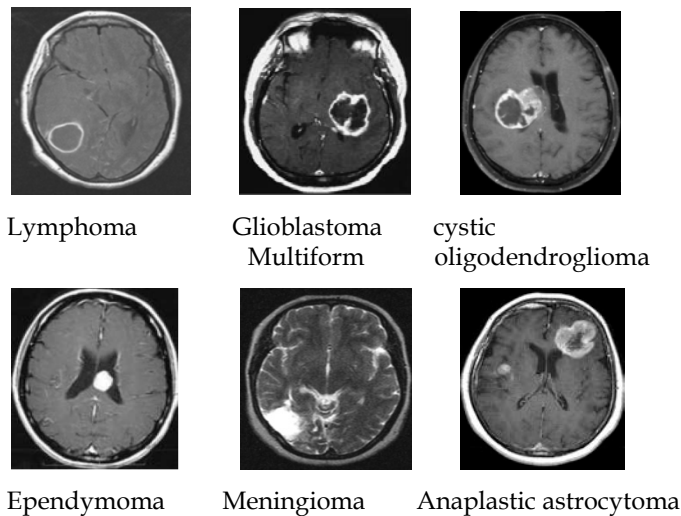


Figure (2): Samples of different types of brain MRI images.

3- THEORETICAL BASICS OF THE PROPOSED CLASSIFIED SYSTEM

This section shows a theoretical basic of the techniques that can be used in this paper as a proposed system classifier.

3-1- PROPOSED SYSTEM ENVIRONMENT:

The proposed system application is programmed using visual studio 2008 (visual basic.Net). Therefore, we will use the Arc Engine (. NET Developer KIT) to give the language Visual basic.Net possibility of dealing with the ESRI ArcMap GIS. The lands classification tool which is the proposed system application tool, it was added to ArcMap GIS as a new GIS dll this file contains the application classes code for the proposed system and also, the application interface windows codes to be run into ArcMap GIS software, it become a portable application; it can be used by any machine which has installed ArcMap GIS software.

The selected tumor has covered by polygon vector layer previously. The block diagram of the proposed system which explains the practical work stages, and figure (4) shows the combination of raster and vector layers in ArcMap GIS.

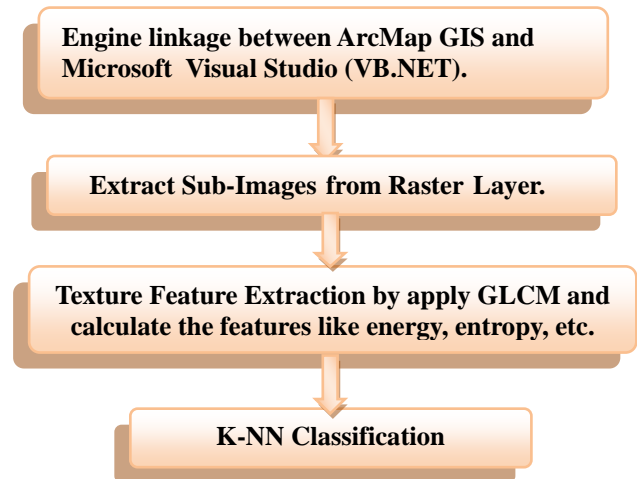
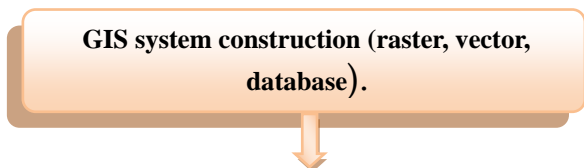


Figure (3) : Block diagram of the proposed system.

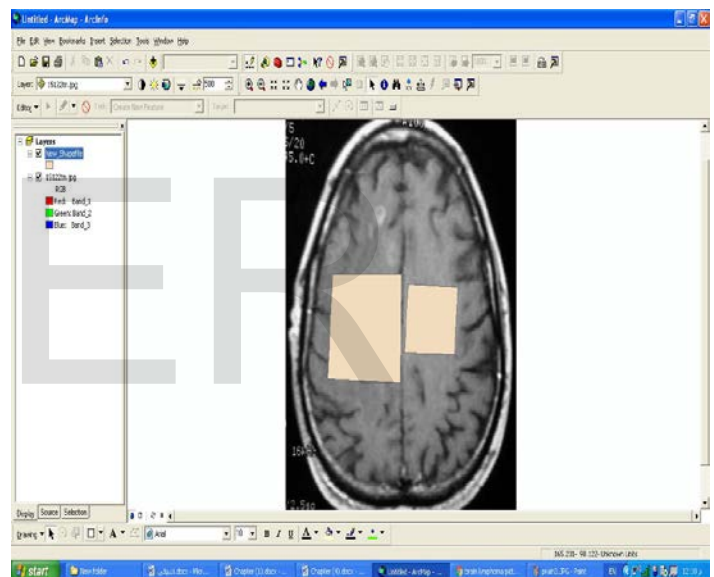


Figure 4: The combination of raster and vector layers in ArcMap GIS.

3-2- EXTRACT SUB-IMAGE FROM RASTER LAYER

One polygon (tumor) has been selected, the system need to investigate the sub-image under this polygon , the system use some techniques to do that. while the raster and vector layer are identical with their spatial reference UTM zone, the system gets the polygon angles points where its edge line are crossed to extract the centroid of the polygon which represent the centroid of the sub-image in raster image . The system the system needs to clip an image window of proposed size equal to (256

256) pixels. The sample is clipped from the center of the related polygon. The center of the buffer window is equal to centroid point of a desired polygon as illustrated in figure

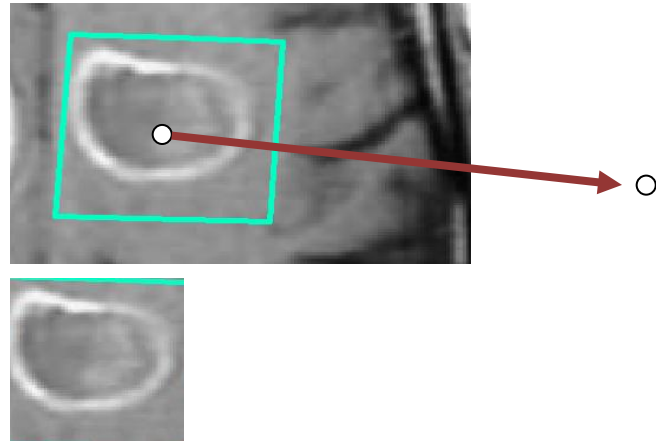


figure (5): polygon and centroid of the polygon

3-4- TEXTURE FEATURE EXTRACTION

The proposed system used Gray Level Co-occurrence Matrix GLCM technique to find the statistic textural feature for each image that has been clipped from raster layer. The output from the previous stage are used in the current stage, Mathematically, those features can be represented as follows[12][21].

- Angular Second Moment

$$ASM = \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} \{P(i, j)\}^2 \quad (1)$$

- Entropy

$$ENTROPY = - \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} P(i, j) \times \log(p(i, j)) \quad (2)$$

- Contract

$$CONSTRAST = \frac{1}{(N-1)^2} \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} (i - j)^2 P(i, j) \quad (3)$$

- Homogenous

$$IDM = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \frac{1}{1+(i-j)^2} P(i, j) \quad (4)$$

- Correlation

$$\text{Correlation} = \frac{\sum_{i=0}^{N-1} \sum_{j=0}^{N-1} (i-\mu_i)(j-\mu_j) P(i, j)}{\sigma_i \sigma_j} \quad (5)$$

- Dissimilarity

$$\text{Dissimilarity} = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} p(i, j) |i - j| \quad (6)$$

- Variance

$$\text{Variance} = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} p(i, j) (i - u_i)^2 \quad (7)$$

- Standard deviation

$$\sigma = \sqrt{\frac{\sum_{i=1}^N (\sum_{j=1}^N P(i, j) - u)^2}{N-1}} \quad (8)$$

Then each image has vectors of eight feature extracted by the system and saved in database , this calculated vector of textural features will be saved temporally in the system database buffer to be used in the next stage to classify the image to which class it belongs.

3-5- CLASSIFICATION

The proposed classification system used the k-Nearest Neighbor classification to classify the new extracted sub-image . This can be done by gets the vector of features for sub-image extracted by the system and compared with the vectors of objects in database to get the smallest distance between extracted sub-image and that images in database .

Nearest-neighbor classifiers are based on learning by analogy, that is, by comparing a given test tuple with training tuples that are similar to it.

The training tuples are described by n attributes. In this way, all of the training tuples are stored in an n-dimensional space. When given an unknown tuple, a k-nearest-neighbor classifier searches for the k training tuples are closest to the unknown tuple. These k training tuples are the k “nearest neighbors” of the unknown tuple. “Closeness” is defined in terms of a distance metric, such as Euclidean distance.

The Euclidean distance between two points or tuples, say, X1 = (x11, x12, , x1n) and X2 = (x21, x22, , x2n).

The minimum value will be taken and then the system can repeat these processes until it recognize the flat area near the selected object.

4 EXPERIMENTAL RESULT

This section will present the results that have been achieved from executing the proposed system's phases and its stages that described in previous sections. The system have polygon vector map.

4.1 TEXTURE FEATURE EXTRACTION RESULTS:

This section shows the results of the features extraction stage for sub-image of selected object extracted from the raster

image and the features of sub-image in database, these features calculated from GLCM with angles (0o, 45o, 90o and 135o) and the result are tabulated in table (2)

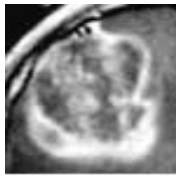


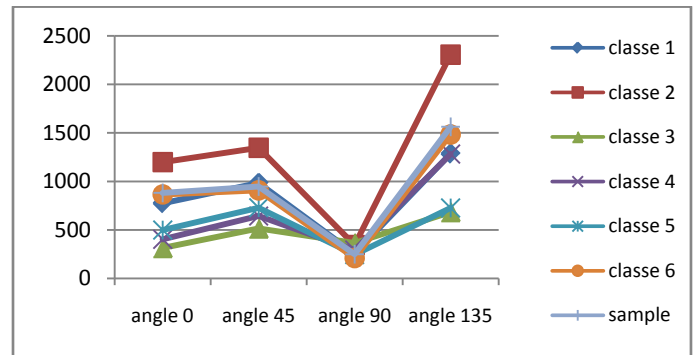
Figure (6) : Anaplastic astrocytoma sample

Table (2): GLCM textural features for sample Anaplastic astrocytoma.

Feature/angles	Angles 0	Angles 45	A
Energy	0.2150530264	0.201757888017	0.2084054122
Entropy	4.149220073	4.33125270326	3.4600961538
Contrast	1198.95	1347.039423077	3460.0961538
Homogeneity	0.28821657933	0.27198973853	0.2795000000
Correlation	0.6107348268	0.56600070137	0.910066869
Dissimilarity	13.0951923077	14.79038461538	3.4600961538
Variance	5409.9336538	5359.0778846	5900.37043
Standard deviation	73.55225118	73.55225118	76.8096075

4.3 THE CLASSIFICATION RESULTS OF THE SYSTEM

When the Euclidean distance measure is implemented at the sub-image features and the six classes features to determine to which class is belong to as shown in figures below which illustrate that sample belongs to class of the Anaplastic astrocytoma that classification system has searched for and then specified, the figure (7) to (10) shows the Energy, the Homogeneity, the Contrast and the Correlation Features comparison,



Figure(7): Contrast Features comparison

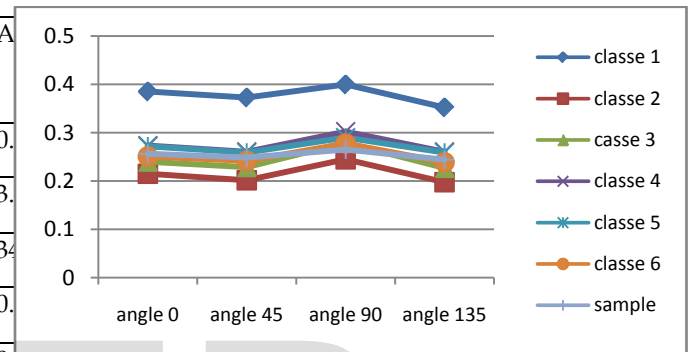
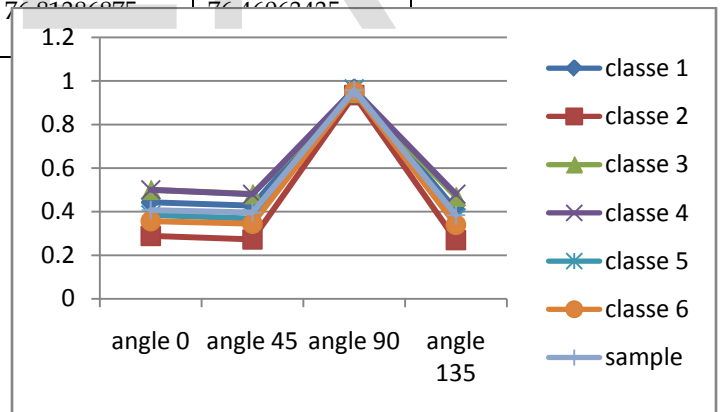
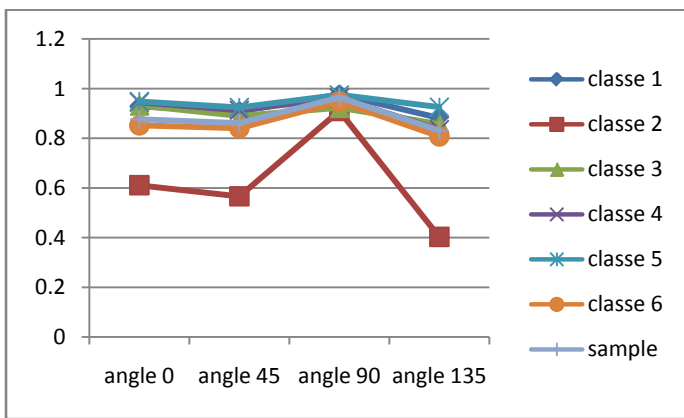


Figure (8): Energy Features comparison



Figure(9): Homogeneity Features comparison



Figure(10): Correlation Features comparison

Then the system will choose the sample that has a minimum distance to class 6

5 CONCLUSIONS

There is a way to create custom commands with VBA in ArcGIS. we can create a new button, tool, combo box, or edit box (collectively named UIControls), then attach code to the control's events, such as what happens when we click a button. After we have created it, we can drag this new control onto a toolbar. The advantages of GIS include its database manner of data management, its capability to overlay and integrate different spatial information, and its powerful tools and functions for image analysis, spatial analysis and modeling Thus, it well

suited for image classification The results indicate that co-occurrence matrix approach is also an effective method in angle 90o evaluation of textures.

REFERENCES

- [1] B. Tso et al., "Classification Methods for Remotely Sensed Data, Taylor & Francis Group, 2009.
- [2] D.Hyam; " The GIS files" , Corporate Publishing, 2003.
- [3] E. S. Flores et al., "GIS Improved Object-Based Classification For Land Use/Cover Change Detection In Human Altere Deciduouss Forest", The American Society for Photogrammetry & Remote Sensing 2009 Annual Conference; Baltimore; Maryland, 2009.
- [4] F. R. Renzetti et al., "Use of a Gray Level Co-occurrence Matrix to Characterize Duplex Stainless steel Phases Microstructure", University of Roma, 2011.
- [5] H. Jiawei , "Data Mining: Concepts and Techniques", Elsevier

- Inc, 2006"
- [6] Kadam D. B., Gade S. S., Uplane M. and Prasad R., "Neural Network Based Brain Tumor Detection Using MR Images", International Journal of Computer Science and Communication (IJCSC), Vol. 2, No. 2, pp. 325-331, July-December, 2011.
- [7] L.A. Al-ani, "Classification of Digital Satellite Images", PhD Thesis; college of science; Al-nahrain University, 1996.
- [8] Lashkari A., "A Neural Network-Based Method for Brain Abnormality Detection in MR Images Using Zernike Moments and Geometric Moments", International Journal of Computer Applications, Vol. 4, No. 7, pp. 1-8, July, 2010.
- [9] Logeswari. T., Karnan M., "An Enhanced Implementation of Brain Tumor Detection Using Segmentation Based on Soft Computing", International Journal of Computer Theory and Engineering, Vol. 2, No. 4, pp. 586-590, August, 2010.
- [10] L. Semler et al., "A comparison Wavelet-Based and Ridgelet-based Texture Classification of Tissues in Computed Tomography", Advanced in Computer Graphics and Computer Vision 4 (2007).
- [11] M. Neteler et al., " Open Source GIS a Grass GIS Approach", Springer Science and Business Media, 2008.
- [12] M. Nixon et al., "Feature Extraction and Image Processing", Elsevier, 2008.
- [13] M. Tuceryan et al., "Texture Analysis. The handbook of Pattern Recognition and Computer Vision (2nd Edition)' / C.H. Chen, L.F. Pau, P.S.P. Wang (eds), World Scientific Publishing Co.. - PP. 207-248,1998.
- [14] M. Roumi, "Implementing Texture Feature Extraction Algorithms on FPGA", MSc Thesis; Delft University of Technology, 2009.
- [15] M. Partio, et al., "Rock Texture Retrieval using Gray Level Co-occurrence Matrix", Tampere University of Technology.
- [16] M. Seetha, et al., "Artificial Neural Networks and other Methods of Image Classification", Journal of Theoretical and Applied Information Technology, 2008.
- [17] N. R. Chrisman, "Exploring Geographic Information Systems", John Wiley and Sons, 1997.
- [18] N. Zulpe et al., " GLCM Textural Features for Brain Tumor Classification", IJCSI International Journal of Computer Science Issues, Vol. 9, Issue 3, No 3, May 2012.
- [19] R.M. Haralick et al., 'Textural Features for Image Classification', IEEE Transactions on Systems, Man and Cybernetics.No 6. - PP. 610-621, Nov 1973..
- [20] R. Methre et al., "Exploring Spatial Information in Spectral Features for Texture Image Retrieval", International Journal of Computer and Network Security Vol. 1; No. 3, 2009.
- [21] R. S. Choras, "Image Feature Extraction Techniques and Their Applications for CBIR and Biometrics Systems", International

Journal of Biology and Biomedical Engineering.

- [22] Ul-Ain Q., Latif G., Kazmi S., Jaffar M. and Mirza A., "Classification and Segmentation of Brain Tumor using Texture Analysis", proceedings of the 9th WSEAS International Conference on Artificial Intelligence, Knowledge Engineering and Databases (AIKED), pp. 147-155, 2010.
- [23] W. Razoqi, "Agricultural Lands Classification ", MSc Thesis; Science College; University of Babylon, 2010.
- [24] Y. Zhang et al., "Remote-Sensing Image Classification Based on an Improved Probabilistic Neural Network", Sensors 9; pages 7516- 7539, 2009.

IJSER