

## Analysis of GLCM Feature Extraction for Choosing Appropriate Angle Relative to BP Classifier

Dr. Tawfiq A. Alasadi<sup>1</sup>, Wadhah R. Baiee<sup>2</sup>

<sup>1</sup>(Dean of Information Technology College/ Babylon University, Iraq)

<sup>2</sup>(Software Department, Information Technology College/ Babylon University, Iraq)

---

**Abstract :** GIS can manage remotely sensed images, users must have an appropriate digital map that represents lands each one has information according to its owner, status, and some other data. The classification of such lands is a great problem which take long time depending on human efforts. Many kinds of classifications had been used, one of them is the use of supervised multi-layer perceptron with backpropagation neural network classifier and using second order statistics Gray Level Co-occurrence Matrix (GLCM) to calculate eight textural features for each one of three visible bands (RGB) for each land sample. In this research we analyzed the GLCM feature extraction algorithm to detect the appropriate angle that can be chosen, relatively with the training of BP classifier had been used according to the number of hidden nodes inside the hidden layer of ANN. As a result the system produce high accuracy with the best angle choosing of GLCM, these results are achieved by comparing the classification results from system test trials with desired user predefined classification dataset.

**Keywords:** classification, features extraction, GIS, GLCM, neural networks, texture.

---

### I. INTRODUCTION

Geographic Information System (GIS) are commonly defined as an information system that manages, manipulates, and analyzes spatial data [1]. Geographical-Spatial data has both spatial and thematic components. 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 [2]. The classification aim is to appoint each object in the study area to one or more elements of a defined label set, so that the radiometric information contained in the image is converted to thematic information, The process can be regarded as a mapping function that constructs a linkage between the raw data and the user-defined label set [3]. There are two types of classification; supervised classification methods which are based on prior knowledge of certain aspects of the statistical nature of the spectral classes of an image pixels are to be identified, and unsupervised classification methods which are performed by using a classification algorithm without any predefinition of spectral classes of interest [4][5]. Texture is an appropriate property of objects. It contains important information about the structural arrangement of surfaces. The use of texture in addition to spectral features for image classification might be expected to result in some level of accuracy improvement, depending on the spatial resolution and the size of the area being classified [6].

Several methods had been improved to use the texture with many kinds of classification to produce new features, [7] 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. [8][9] proposed a hybrid classifier for polar metric SAR images. The feature sets consist of span image, the H/A/ $\alpha$  decomposition algorithm, and the GLCM-based texture features. Then, a probabilistic neural network (PNN) was adopted for classification. [10] used GIS techniques to improve the classification capabilities of a feature extraction algorithm for land use/cover change detection in a deciduous forest environment. [11] used various techniques, grey level co-occurrence matrices (GLCM), Gabor filters, combined GLCM and Gabor filters. Here the real scene images are classified in four classes such as near natural, near manmade, far natural and far manmade. [6] used the GLCM with artificial neural networks (ANN) Backpropagation (BP) classifier technique and used features that extracted from GIS images objects according to their positions to be the input to the ANN. Our research analyzes GLCM feature extraction method to choose the appropriate angle of texture in relative to BP classifier that used in [6].

### II. INFORMATION GATHERING FROM GIS DATA

The system will clip a colored sampled images with three bands (Red, Green, Blue) from underneath the polygon to be the raster image sample of the polygon (region) of land use layer and will be moved to the textural feature extraction stage [6], we propose in this research to clip an image window of size equal to (24\*24) pixels for each studied land instead of the size had been proposed in [6]. The sample is clipped from a random position of the related polygon. Information from database are required for the classification learning

reason, so this information is located in attribute database of ArcMap GIS. These steps is programed and developed using ArcMap Objects and C#.NET environment [12]. Collected values represent the agriculture status for each land, this value will become the classifier learning desired output for each related extracted land object. Now we have gathered information for each studied land represented by (Land ID, Image Sample, and Current Status).

### III. GLCM AND ANN CLASSIFIER

Texture feature extraction technique based on the gray-level co-occurrence matrix (GLCM), sometimes called the gray-tone spatial-dependency matrix [13]. The principal concept of GLCM is that the texture information contained in an image is defined by the adjacency relationships that the gray tones in an image have to one another. 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  $\theta$ . 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. The Eight textural features [6][14][15][16][17][18] that is extracted in proposed system for each visible band (R, G, B), which are : Angular Second Moment (1) , Contrast (2), Inverse Difference Moment (3), Entropy (4), Correlation (5), Cluster Shade (6), Cluster Prominence (7), Haralick Correlation (8) .

- (1),  $\sum_{i=0}^{G-1} \sum_{j=0}^{G-1} \{P(i, j)\}^2$
- (2),  $\sum_{n=0}^{G-1} n^2 \{ \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} P(i, j) \}$ ,  $|i - j| = n$
- (3),  $\sum_{i=0}^{G-1} \sum_{j=0}^{G-1} \frac{1}{1+(i-j)^2} P(i, j)$
- (4),  $-\sum_{i=0}^{G-1} \sum_{j=0}^{G-1} P(i, j) \times \log(P(i, j))$
- (5),  $\frac{\sum_{i=0}^{G-1} \sum_{j=0}^{G-1} (i-\mu_i)(j-\mu_j)}{\sigma_i \sigma_j}$
- (6),  $\sum_{i=0}^{G-1} \sum_{j=0}^{G-1} \{i + j - \mu_i - \mu_j\}^3 \times P(i, j)$
- (7),  $\sum_{i=0}^{G-1} \sum_{j=0}^{G-1} \{i + j - \mu_i - \mu_j\}^4 \times P(i, j)$
- (8),  $\frac{\sum_{i=0}^{G-1} \sum_{j=0}^{G-1} (ij)P(i, j) - \mu_x \mu_y}{\sigma_x \sigma_y}$

These features become the input to the ANN classifier. The neural network which be used in the system is multilayer perceptron with supervised backpropagation algorithm. In [6] and in the proposed research ANN classifier consists of three main layers (Input layer, one Hidden layer, and Output layer), it is completely connected net.

### IV. ANALYSIS OF GLCM ANGLES EFFECTS

In this research we analyze the GLCM algorithm relative to hidden nodes of ANN classifier that be used to choose and proof the right angle of feature extracting must be used in such system. Instead of choosing the angle randomly and train the ANN [6][19], we put a systematic way to apply the angle for each defined number of hidden nodes in the hidden layer of the used ANN classifier . The distance that used in our system is  $d=1$  of GLCM , and for each angle of that algorithm  $\{0,45,90,135\}$  of adjacent pixels we trained the BP classifier of the system and modify the hidden nodes number to achieve high accuracy test of the classification system , results will be compared with user predefined classification dataset . Fig.1 illustrates the block diagram of the main stages of proposed system.

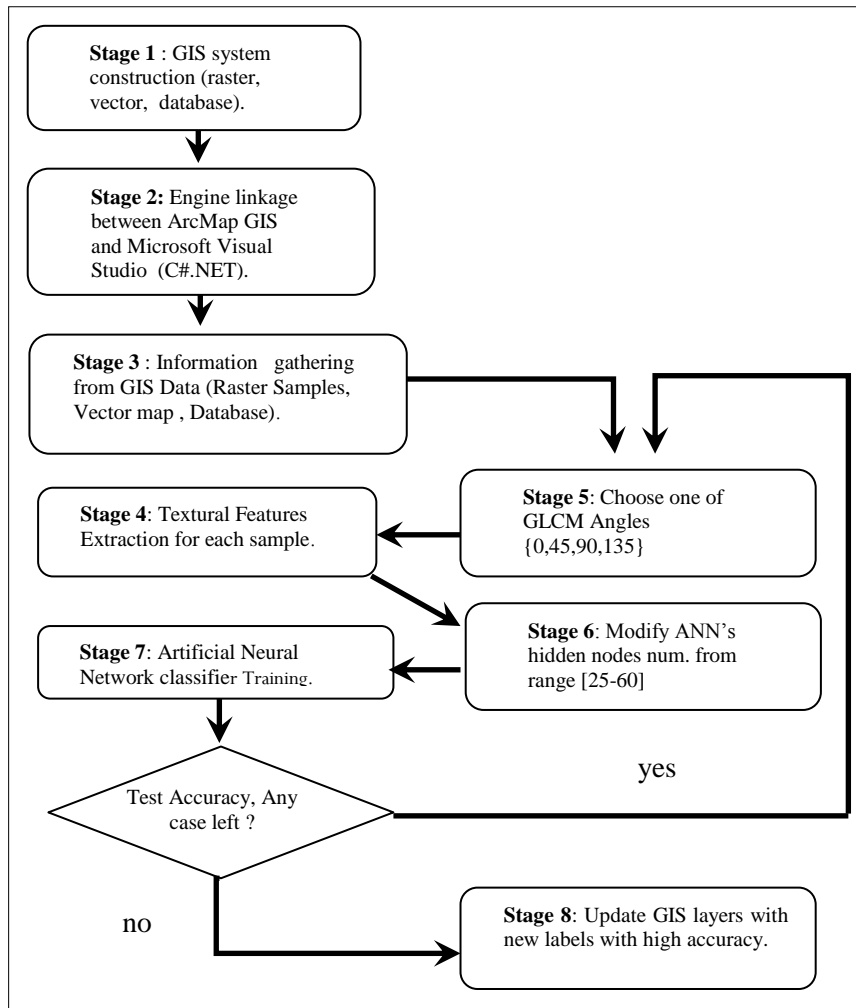


Figure 1 : block diagram of the proposed system.

The following parameters are chosen for the training of BP training algorithm , they produced system accuracy reaches  $\cong 90\%$  , these parameters are the best among other training parameters that used in our analysis by trial and error : momentum = 0.1, learning rate = 0.01 , sigmoid alpha value = 2 , error limit = 0.01 , threshold value = 0.8 , but the number of hidden layer's nodes and the angel of the GLCM are varied . According to Table 1, we have 32 training cases each one with different hidden nodes number and different GLCM angle and fixing the above parameters . Table 1 shows the training trials each with its parameters and (GLCM , Hidden Nodes) variations. And Fig.2 shows the chart of the relations among different angles and hidden nodes .

Table1 : Trainings with different GLCM angles and Hidden Nodes numbers of ANN

Trial No.	Iteration	Hidden Nodes	GLCM angle	Accuracy%
1	21567	20	0	59.2968
2	45312	25	0	63.2968
3	4334	30	0	66.03653
4	8009	35	0	69.40639
5	31140	40	0	74.86301
6	38754	50	0	76.6895
7	19982	55	0	79.23288
8	42066	60	0	75.77626

9	11117	20	45	52.2968
10	45312	25	45	61.2968
11	4334	30	45	56.03653
12	6009	35	45	68.40639
13	21689	40	45	71.86301
14	38754	50	45	70.6895
15	19982	55	45	72.23288
16	42066	60	45	68.77626
17	40556	20	90	76.60274
18	37811	25	90	78.05936
19	37887	30	90	83.05936
20	38963	35	90	84.51598
21	31511	40	90	87.51598
22	37877	50	90	88.93772
23	37425	55	90	90.42922
24	37091	60	90	87.9726
25	40321	20	135	61.99087
26	11128	25	135	66.10046
27	40378	30	135	67.47032
28	6530	35	135	77.77626
29	41724	40	135	75.77626
30	17928	50	135	79.6895
31	37677	55	135	82.6895
32	37121	60	135	78.6895

Another consideration has been done in this research which is the relation of the number of ANN classifier iterations in each case of training with different angle of GLCM and hidden nodes of the net to reach the error limit has been defined previously. We considered that the worst case (more iterations) is when the system has (GLCM angle 0, 25 hidden nodes) it was trained with 45312 iterations and the best case of training (little iterations) to reach the limit is when the system has (GLCM angle 45, 30 hidden nodes), it was trained with 4334 iterations . the rest cases can be shown in Fig.3 which illustrates a chart of the relation of iterations among different hidden nodes and GLCM angles .

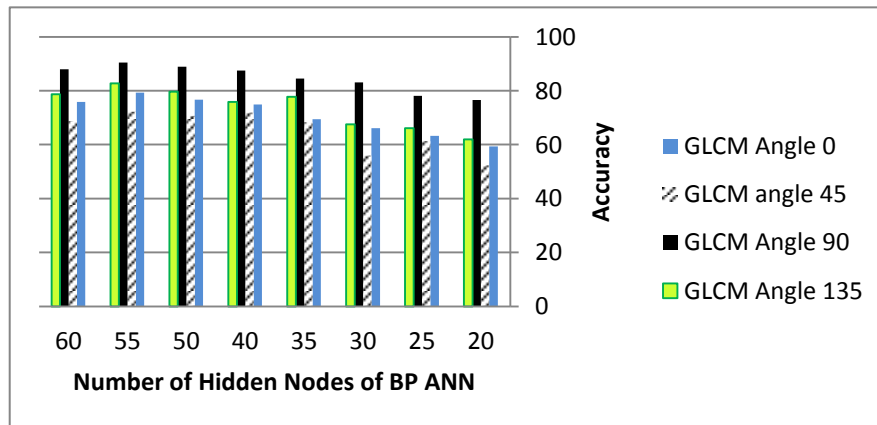


Figure 2 : chart of the relation of GLCM angles- hidden nodes that affect classification accuracy.

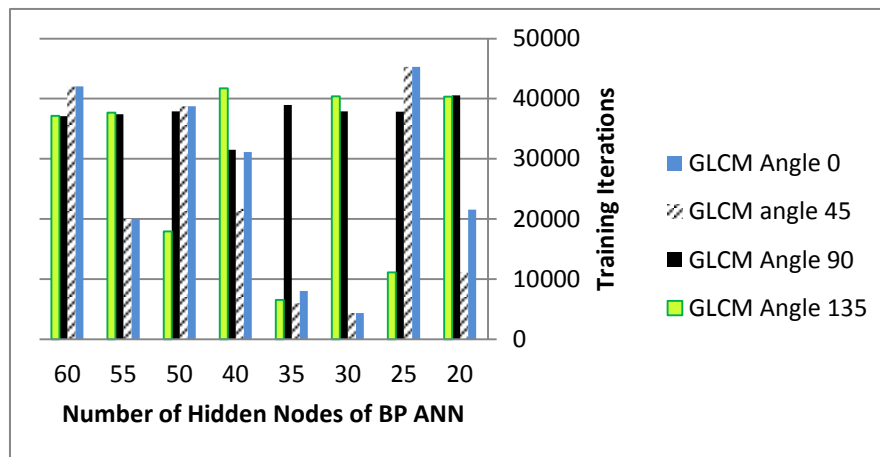


Figure 3 : chart of the relation of GLCM angles- hidden nodes that affect training iterations.

## V. CONCLUSIONS

According to the use of GIS data (raster, vector, and database) in textural features extracting from grey level co-occurrence matrix GLCM for each land sample to be the input to back propagation ANN classifier, the proposed system has analyzed the relationship between GLCM angles and the number of hidden layer nodes then it proves that angle of degree 90 among {0,45,90,135} with 55 hidden nodes with an input of 24 features of GLCM is the best for classification using this hybrid system and produce an accuracy near 90% of classification. Furthermore, the research concludes that there is a relation between the number of hidden nodes of the ANN classifier and the GLCM angle and the number of the training iterations of system training phase to get the best case of fastest training. For future works in this field of GIS image classifications, we recommend to use support vector machine SVM instead of ANN classifier and the extracted features can be refined with some mechanism like principle component analysis PCA to get more accurate classifications with less system error.

## REFERENCES

- [1] K. C. Clarke et al., *Geographic Information System and Environmental Modeling*, (Prentice Hall, 2009).
- [2] F. Harvey, *A primer of GIS fundamental geographic and cartographic concepts*, (The Guilford Press, 2008).
- [3] I. T. Service, *Introduction to GIS using ArcGIS*, (Durham University, 2006).
- [4] A. M. Coleman, *An Adaptive Landscape Classification Procedure Using Geoinformatics and Artificial Neural Networks*, MSc. Thesis; Faculty of Earth and Life Sciences Vrije University, Amsterdam The Netherlands, MA, 2008.
- [5] S. Haykin, *Neural Networks and Learning Machines*, (PHI Learning Private Limited, 2009).
- [6] W. R. Baiee et. Al., Classification of GIS Image using GLCM and Neural Network, *Basra Journal of Science VOL Vol.31(2)*, 110-119, 2013.
- [7] 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.
- [8] Y. C. Chen et al., Texture Features for Classification of Ultrasonic Liver Images, *IEEE transactions on Medical; Vol.11; No. 2*; Pages 141-151, 1992.
- [9] Y. Zhang et al., Remote-Sensing Image Classification Based on an Improved Probabilistic Neural Network, *Sensors 9*; pages 7516-7539, 2009.
- [10] E. S. Flores et al., GIS Improved Object-Based Classification For Land Use/Cover Change Detection In Human Altered Deciduous Forest, *The American Society for Photogrammetry & Remote Sensing 2009 Annual Conference*; Baltimore; Maryland, 2009.
- [11] Ranjita Mishra, Scene Image Analysis using GLCM & Gabor Filter, *International Journal of Engineering Research and Applications*, Vol. 3, Issue 2, March -April 2013.
- [12] K. T. Chang, *Programming ArcObjects with VBA*, Taylor & Francis Group, 2008.
- [13] B. Tso et al., *Classification Methods for Remotely Sensed Data*, Taylor & Francis Group, 2009.
- [14] R. M. Haralick et al., Textural Features for Image Classification, *IEEE Transaction on Systems Vol. SMC-3 No. 6*, 1973.
- [15] M. S. Nixon et al., *Feature Extraction and Image Processing*, Elsevier, 2008.
- [16] F. Albrechtsen, Statistical Texture Measures Computed from Gray Level Co-occurrence Matrices, *University of Oslo*, 1995.
- [17] F. Aguera et al., Using Texture Analysis to Improve Per-pixel Classification of Very High Resolution Images for Mapping Plastic Greenhouses, *ISPRS Journal of International Society for Photogrammetry & Remote Sensing (63)*, P 635-646, 2008.
- [18] M. Cheong et al., An Approach to Texture-Based Image Recognition by Deconstructing Multispectral Co-occurrence Matrices using Tchebichef Orthogonal Polynomials, *IEEE*, 2008.
- [19] D. I. Verbil, *Practical GIS analysis*, Taylor & Francis Group, 2003.