



Proposed Classification System by Using Artificial Neural Network

Esraa Z. Mohammed

State Company for Internet Services / Ministry of Communication

Isra_mohammed2@yahoo.com

Received date: 12 / 11 / 2014

Accepted date: 8 / 3 / 2015

ABSTRACT

The research presented in this paper was aimed to develop a recognition system for microscopic images of human tissues samples. The system should classify different types of tissues (i.e., Breast, Liver and blood cells). In this paper, co-occurrence matrix, run length matrix features combined with developed method to measure the roughness were used to extract a set of textural features in order to perform texture analysis for tissues samples. A feed forward neural network was used to classify different types of tissues according to the extracted feature vectors. For ANN training purpose the back-propagation training algorithm was used. Evaluation tests were carried on 550 tissues images. The test results indicated that the best attained success rate was around 93%. The proposed system was implemented using “visual basic.net” and all tests be done on windows operating system environment.

Keywords: GLCM, RLM, neural network, texture analysis

نظام تصنيف مقترح باستخدام الشبكات العصبية الاصطناعية

اسراء زكي محمد

وزارة الاتصالات / الشركة العامة لخدمات الشبكة الدولية للمعلومات

Isra_mohammed2@yahoo.com

تاريخ قبول البحث: 2015 / 3 / 8

تاريخ استلام البحث: 2014 / 11 / 12

المخلص

البحث المقدم في هذه الورقة يهدف إلى تطوير نظام التعرف على الصور المجهرية لعينات مختلفة من الأنسجة البشرية. حيث يقوم النظام بتصنيف أنواع مختلفة من الأنسجة البشرية بالاعتماد على الخصائص المستخرجة من الصورة نفسها، و لوصف القوام في الانسجة تم استخدام الخصائص المستخرجة من مصفوفات Co-occurrence و Run length مع استخدام طريقة جديدة لحساب خصائص الخشونة في الصورة. واستخدمت الشبكة العصبية، من النوع التغذية الامامية، في تصنيف الانواع المختلفة من الانسجة. ولغرض تدريب الشبكة العصبية استخدمت طريقة التغذية العكسية. ولغرض اختبارات التقييم تم استخدام (550) صورة نسيج. وأشارت نتائج الاختبار الى ان معدل النجاح الذي تم بلوغه هي بحدود 93% . وقد تم استخدام لغة فيجوال بيسك دوت نت في برمجة النظام واختبار النتائج في بيئة نظام التشغيل ويندوز.

الكلمات الدالة: تصنيف الصور، الشبكات العصبية الاصطناعية، حساب الخشونة، تحليل القوام.

1.INTRODUCTION

Intelligent systems cover a major application area providing significant assistance in medical diagnosis. In most cases, the development of these systems leads to valuable diagnostic tools that may largely assist in the identification of tumors or malignant formations, enhancing the capability to make accurate diagnosis, provide rapid identification of abnormalities and enable diagnosis in real time [1].

Texture analysis on images are native and complex visual patterns that reproduce the data



of gray level statistics, anatomical intensity variations, texture, spatial relationships, shape, and structure. Image texture analysis aims to interpret and understand these real-world visual patterns, which involves the study of methods broadly used in image filtering, normalization, classification, segmentation, labeling, synthesis and shape from texture. Texture classification involves extracting features from different texture images to build a classifier. It determines to which of a finite number of physically defined classes (such as normal and abnormal tissue) a homogeneous texture region belongs. The classifier is then used to classify new instances of texture images. The textural properties of spatial patterns on digital images have been successfully applied to many practical vision systems, such as the classification of images to analyze diagnosis tissues for dementia, tumors, hyper spectral satellite images for remote sensing, content based retrieval, detection of defects in industrial surface inspection, and so on [2].

Scientific interest in artificial neural networks mainly arises from their potential ability to perform interesting computational tasks. In principle, ANNs are mostly used for pattern matching capabilities. Their human-like characteristics are utilized to assist medical decision making. Neural networks are extremely useful, because they are not only capable of recognizing patterns with the aid of expert, but also of generalizing the information contained in the input data. They can represent relations which are complex combinations of digital image processes, and carry out the required recognition and classification. ANN with ability to learn by example is very flexible and powerful in medical diagnosis [1]. The objective of ANNs is to support doctors and not to replace them. In addition, ANNs can provide a valuable tool that could minimize the disagreement and inconsistencies in interpretation, and handling uncertainty and ambiguity in distorted or noisy images. Thus, such methods provide human experts with significant assistance in medical diagnosis [3]. Many studies have applied the concepts of texture feature extraction and neural network to analyze and classify cells, tumors, and other region of interest in biomedical images for example:

Ribar, et al. [4] applied self-organizing mapping neural networks in the analysis of breast cancer luminescence data. Data consist of three dimensional vectors presenting normal and malignant human tissue. The possibility of such data classification in two groups (normal and malignant tissue) is analyzed. The network performed successful classification. George and Mohammed [5], developed a recognition system for microscopic images of breast tissues

samples. The system classified breast tissues as malignant or not, or identifying their malignancy types. The multi-scale fractal dimension concept was used to extract a set of textural features for breast tissues samples. The box counting method was used to estimate the multi fractal dimensions. A feed forward neural network was used to classify different types of breast tissues according to the extracted fractal dimension vectors. Evaluation tests were carried on 368 breast tissues images. The test results indicated that the best attained success rate was around 97%. Arunadevi, et al [2] improved classifier for brain tumor tissue characterization. The classifier obtained 98.25% accuracy. They extended the computation of gray level co-occurrence matrix (GLCM) and Run length matrix (RLM) to a three-dimensional form for feature extraction. An improved Extreme Learning Machine (ELM) classifier algorithm was explored, for training single hidden layer artificial neural network, integrating an enhanced swarm-based method in optimization of the best parameters (input-weights, bias, norm and hidden neurons), enhancing generalization and conditioning of the algorithm. Padma and Sukanesh [6] classify and segment the brain soft tissues from computed tomography images using the wavelet based dominant gray level run length feature extraction method with Support Vector machine (SVM) classifier. An average accuracy rate of above 98% was obtained using this classification and segmentation algorithm.

2.CONCEPTS AND METHODS

2.1.CO-OCCUERRENCE MATRIX

In a statistical texture analysis, texture features were computed on the basis of statistical distribution of pixel intensity at a given position relative to others in a matrix of pixels representing image. Depending on the number of taken pixels in each combination, there is first-order statistics, second-order statistics or higher-order statistics. Feature extraction based on Co-occurrence matrix is the second-order statistics that can be used to analyze image content as a texture. **Figure (1)** below presents an example about the formation of the Co-occurrence matrix of the gray image (4 levels) image at the distance $d = 1$ and the direction of 0° [7].

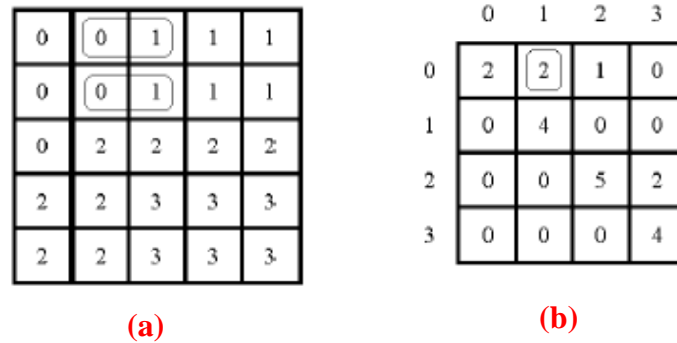


Figure (1): An example of GLCM formation

In this work in addition to the horizontal direction (0°), GLCM can also be formed for the direction of 45°, 90° and 135°. Co-occurrence matrix is a matrix of frequencies at which two pixels, separated by a certain vector, occur in the image. The contents of the GLCM matrix depend on the scan direction and the distance relationship between pixels. By varying the separation distance it allows to capture different texture characteristics, which will reflect important information's about the nature and extent of existing local correlation between pixels (i.e., several values were tested in this work 1,2, and 3). After counting the frequency of each possible transition between pixels values, there is still one step to take before texture measures can be calculated. The measures require that each Co-occurrence matrix cell contain not a count, but rather a probability. It is defined by $P(a,b|d,\theta)$ which expresses the probability of the couple pixels at θ direction and d interval. When θ and d is determined, $P(a,b|d,\theta)$ is showed by $P(a,b)$. Once the normalized Co-occurrence matrix has been created, various features can be computed from it. Haralick and his colleagues extracted 14 features from the Co-occurrence matrix, although in many applications only eight features are widely used, and in this work only these 8 features have been used, they are: Contrast, Energy, Norm Entropy, Homogeneity, Cluster Shade, Cluster Prominence, Inverse Difference Moment, and Maximum Probability, which are obtained using the following equations (1 to 8) respectively [8, 9]:

$$Contrast = \sum_{i=1}^N \sum_{j=1}^N (i - j)^2 * Co(i, j) \quad (1)$$

Where, $Co(i, j)$ is the probability element of the Co-occurrence matrix of the source pixel having value i and the target pixel having value j .

Contrast is the main diagonal near the moment of inertia, it measures the distribution status of the matrix elements and if there is local changes in number, also, it reflects the image clarity and texture of shadow depth. Therefore, the Contrast feature is a measure of the image contrast or the amount of local variations present in an image.

$$Energy = \sum_{i=1}^N \sum_{j=1}^N Co^2(i, j) \quad (2)$$

Energy is a gray-scale image texture measure of homogeneity changing, it reflects the distribution of image gray-scale uniformity of weight and texture. Hence it is a suitable measure for detection of disorders in textures.

$$NormEntropy = \frac{\sum_{i,j=1}^N |co(i, j)|^{1.5}}{N} \quad (3)$$

Entropy measures image texture randomness, when the Co-occurrence matrix elements have similar values, then the Norm Entropy achieves small value; on the other hand, if the values of Co-occurrence matrix elements are very uneven, the Norm Entropy value will be high. Therefore, the Entropy gives a measure of complexity of the image. Complex textures tend to have higher entropy.

$$LocalHomogeneity = \sum_{i,j=1}^N \frac{1}{1+(i-j)^2} * co(i, j) \quad (4)$$

Homogeneity is measure for uniformity of Co-occurrence matrix, and if the large valued elements lie on the main diagonal, the Homogeneity value will be large, compared to other cases.

$$ClusterShade = \sum_{i,j=1}^N (i - p_x + j - p_y)^3 * co(i, j) \quad (5)$$

Dissimilarity measures how different elements of the Co-occurrence matrix are from each other.

$$Clusterpro\ min\ ence = \sum_{i,j=1}^N (i - p_x + j - p_y)^4 * co(i, j) \quad (6)$$

Cluster shade and cluster prominence are measures of the skewness of the matrix, in other words the lack of symmetry. When cluster shade and cluster prominence are high, the image is considered not symmetric.

$$\text{InverseDifferrentMoment} = \sum_{i=1}^N \sum_{j=1}^N \frac{Co(i, j)}{|i - j|^2} \text{ where, } i \neq j \quad (7)$$

Inverse Difference Moment measures the local changes in texture image elements.

$$\text{Maximum Pr obability} = \text{Max}[Co(i, j)] \quad (8)$$

Where,

$$P_x = \sum_{i,j=1}^N i * co(i, j) \quad \text{and} \quad P_y = \sum_{i,j=1}^N j * co(i, j)$$

Provides the pixel pair that is most predominant in the image; the maximum probability is expected high if the occurrence of the most predominant pixel pair is high.

2.2.Run Length Matrix

Run-length statistics capture the coarseness of a texture in specified directions. A run is defined as a string of consecutive pixels which have the same intensity along a specific linear orientation. Fine textures tend to contain more short runs with similar intensities, while coarse textures have more long runs with significantly different intensities.

Run length is the number of adjacent pixels that have the same intensity in a particular direction. Run-length matrix is a two-dimensional matrix where each element is the number of elements j with the intensity i , in a given direction. For example, **Figure (2.a)** below shows a matrix of size 4x4 pixel image with 4 levels. **Figure (2.b)** the corresponding Run-length matrix in the direction of 0° [7].

1	2	3	4
1	3	4	4
3	2	2	2
4	1	4	1

Level	Run Length (j)			
	1	2	3	4
i				
1	4	0	0	0
2	1	0	1	0
3	3	0	0	0
4	3	1	0	0

(a)

(b)

Figure (2): An example of RLM formation

In addition to the 0° direction, run length matrix can also be formed in the other directions, in this work run length matrix was calculated in 0° direction (horizontally) and 90° direction (vertically), because most of the medical images are concerned with complication tissues which doesn't show fine isotropic symmetries along specific direction. After determination the frequency of occurrence for each possible run, then the probability is calculated for each direction and for both directions at same time, which in turn is used to calculate the run length features values. From each run-length matrix p(i, j), many numerical texture measures can be computed. In this work thirty features of run length statistics have been used (i.e., ten for horizontal, ten for vertical and ten for both directions), and these features are obtained using equations (9 to 18) respectively [10]:

Short Run Emphasis (SRE): it measures the distribution of short runs. It is highly dependent on the occurrence of short runs and it is expected large for fine textures. It is measured using the following equation:

$$SRE = \frac{1}{n_r} \sum_{i=1}^M \sum_{j=1}^N \frac{P(i, j)}{j^2} \tag{9}$$

Where, p(i, j) is the run length matrix, M be the number of gray levels and N be the maximum run length, nr is the total number of runs.

Long Run Emphasis (LRE): it measures distribution of long runs. It is highly dependent on the occurrence of long runs and is expected large for coarse structural textures. It is measured using the following equation:

$$LRE = \frac{1}{n_r} \sum_{i=1}^M \sum_{j=1}^N P(i, j) * j^2 \quad (10)$$

Gray-Level Non uniformity (GLN): it measures the similarity of gray level values throughout the image. It is expected small if the gray level values are alike throughout the image. It is determined using:

$$GLN = \frac{1}{n_r} \sum_{i=1}^M \left(\sum_{j=1}^N P(i, j) \right)^2 \quad (11)$$

Run Length Non uniformity (RLN): it measures the similarity of the length of runs throughout the image. It is expected small if the run lengths are alike throughout the image. It is determined using:

$$RLN = \frac{1}{n_r} \sum_{j=1}^N \left(\sum_{i=1}^M P(i, j) \right)^2 \quad (12)$$

Low Gray-Level Run Emphasis (LGRE): it measures the distribution of low gray level values. It is expected large for the image with low gray level values. It is determined using following equation:

$$LGRE = \frac{1}{n_r} \sum_{i=1}^M \sum_{j=1}^N \frac{P(i, j)}{i^2} \quad (13)$$

High Gray-Level Run Emphasis (HGRE): it measures the distribution of high gray level values. It is expected large for the image with high gray level values. It is determined using the following equation:

$$HGLRE = \frac{1}{n_r} \sum_{i=1}^M \sum_{j=1}^N P(i, j) * i^2 \quad (14)$$

Short Run Low Gray-Level Emphasis (SRLGE): it measures the joint distribution of short runs and low gray level values. It is expected large for the image with many short runs and lower gray level values. It is determined using the following equation:

$$SRLGE = \frac{1}{n_r} \sum_{i=1}^M \sum_{j=1}^N \frac{P(i, j)}{i^2 * j^2} \quad (15)$$

Short Run High Gray-Level Emphasis (SRHGE): it measures the joint distribution of short runs and high gray level values. It is expected large for the image with many short runs and high gray level values. It is determined using the following equation:

$$SRHGE = \frac{1}{n_r} \sum_{i=1}^M \sum_{j=1}^N \frac{P(i, j) * i^2}{j^2} \quad (16)$$

Long Run Low Gray-Level Emphasis (LRLGE): it measures the joint distribution of long runs and low gray level values. It is expected large for the image with many long runs and low gray level values, and it is determined using the following equation:

$$LRLGE = \frac{1}{n_r} \sum_{i=1}^M \sum_{j=1}^N \frac{P(i, j) * j^2}{i^2} \quad (17)$$

Long Run High Gray-Level Emphasis (LRHGE): it measures the joint distribution of long runs and high gray level values. It is expected large for images with many long runs and high gray level values. It is determined using the following equation:

$$LRHGE = \frac{1}{n_r} \sum_{i=1}^M \sum_{j=1}^N P(i, j) * i^2 * j^2 \quad (18)$$

2.3. Roughness Feature

Roughness is a measure of the texture of a surface. It is quantified by the vertical deviations of a real surface from its ideal form. If these deviations are large, the surface is rough; if they are small the surface is smooth. Roughness is typically considered to be the indication for the degree of existing high frequency, short wavelength component of a measured surface. The surface of a rough texture presents a high number of asperities. In an image, roughness can be described as a set of quick spatial transitions with varying amplitude. From a frequential point of view, the image asperities in the spatial domain correspond to the presence of high frequencies.

The implements roughness sub module have two stages, the first one aims to determine the deflections (i.e., residue) of pixels values of the original image, from the corresponding values which belong the image after applying smoothing. The smoothing task is performed using the average filter. The residue values will be considered as indicators for the existing local roughness in the image. The second stage is feature extraction, where a set of 15 statistical moments are determined for the determined histogram of residue component.

A. Smoothing Image Using Average Filter

In this stage, the average filter is applied on image. The average (mean) filter smooth the image data and it thus eliminates the noise. This filter performs spatial filtering on each individual pixel in an image in a square window surrounding each pixel, then the average filter computes the sum of all pixels in the filter window and then divides the sum by the number of pixels in the filter window, that is

$$\overline{\text{Im } g(x, y)} = \frac{1}{L^2} \sum_{-L/2}^{L/2} \sum_{-L/2}^{L/2} \text{Im } g(x+i, y+j) \quad (19)$$

Where $\text{Im } g()$ is the mean value, L is the length of the square window. The degree of the smoothing depends on number of iterations (i.e., in this work the number of iteration set 2). Also, different window size values were tested (i.e., 5, 7, and 9).

B. Moments Determination

In this stage, the central moments (moments about the mean) were used because they are more interesting than the moments about zero. The expression for the n th order moments about the mean is given by:

$$\mu_n = \sum_{i=0}^{L-1} (z_i - m)^n \rho(z_i) \quad (20)$$

$$m = \sum_{i=0}^{L-1} z_i * \rho(z_i) \quad (21)$$

Where, z_i is a random variable indicating intensity, $P(z_i)$ is the histogram of the intensity levels in a region, L is the number of possible intensity levels and m is the mean (average) intensity.

The first step towards extracting the feature vectors of the roughness attribute is the determination of histogram of residue which represents the difference between the original image and smoothed image. Since the using of histogram as it is as a feature vector; then the feature size will be high such that it will increase the computation cost of the next steps in the retrieval system. So to overcome this problem and reduce the size of the feature vector then instead of histogram elements values the moments (up to five orders) are calculated. Also, as smoothing component the mean, median and maximum were used. So, the size of roughness feature vector was taken 15, five features for each one.

2.4. Neural Network

Artificial neural network models have been studied for many years in the hope of achieving human-like performance in several fields such as speech and image understanding. The networks are composed of many nonlinear computational elements operating in parallel and arranged in patterns reminiscent of biological neural networks. The network nodes, belong to adjacent layers, are connected, and their weights are typically adapted during the training phase to achieve high network performance [11].

Generally, the classification and recognition problems have been solved by traditional mathematical or statistical techniques. However, when there is large amount of data had to be processed, and it has wide dynamic range of variations, then the neural nets are capable to successfully process this data with reasonable amount of calculations.

In this research the multilayer feed forward artificial neural network had been trained to classify different kinds of tissues. For training purpose, the back-propagation algorithm was used. The architecture of the applied neural network consists of four layers: an input layer, two hidden layers and an output layer. The ANN input is 85 extracted texture features.

2.5. Proposed System

The system is composed of the following main processes as shown in **Figure (3)**:

A) Image loading: The type of image format used in this research is 24BMP format. Then the loaded RGB color image was transformed into gray images using the following equations:

$$Gr = 0.3R + 0.59G + 0.11B \quad (22)$$

Where, R is red, G is green, and B is blue color component, Gr is the gray.

Also, the quantization level was taken (20) to quantize the intensity component, because fewer number of grey levels faster the computation when the statistics are applied.

B) Feature extraction and analysis: Eighty Five features were extracted, divided into three categories: forty features extracted from Co-occurrence matrix, eight features for each direction (horizontally, vertically, diagonal, inverse diagonal, all directions) and thirty features based on run-length matrix, ten features for each direction (horizontal, vertical and both direction) and fifteen features for roughness measure using five order moment (five for mean, five for median and five for maximum probability).

C) Training the neural network: In this stage the multilayer feed forward artificial neural network had been trained, using back-propagation algorithm, a set of feature vectors are extracted from known tissues images and saved in a feature vector database, then these vectors are used to train a feed forward neural network by adjusting its nodes weights and bias values using back-propagation algorithm. The collection of training samples consists of 350 tissues, and same number of feature vectors have been registered and saved in feature vector database. The computed weights and bias values of the trained network are also registered in the dedicated database called weight vector database.

In the training stage the network starts with a random set of weights and the training pattern is presented at the input layer. Then, the outputs of the network are evaluated and compared with the expected output vector, the error is calculated and the results are fed back from output layer to adjust weights. These steps are repeated for all the training set, and at each time the weights are adjusted. The training continues until the overall mean square error (MSE) between the desired and actual output becomes less than (or equal to) a predefined threshold value. When reaching this error level then the network is considered well trained, and can be used to recognize the types on any unknown input tissue image.

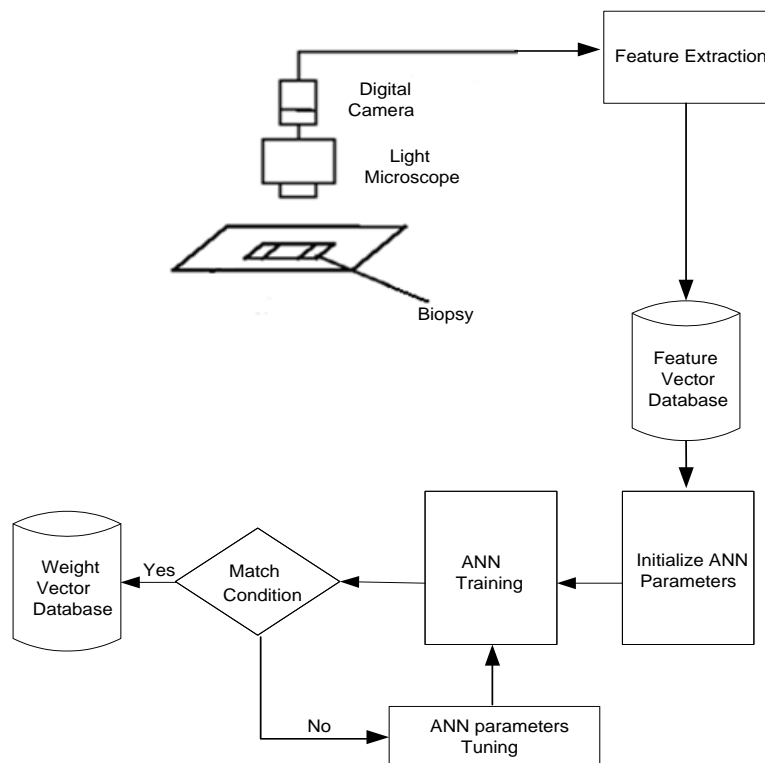


Figure (3): The Diagram of Proposed system

In the established system, the number of input nodes is set equal to the size of extracted texture features (i.e., 85). Two hidden layer were used, the number of hidden nodes were varied to find out the best smallest number of hidden nodes required to get best classification rate. Also, the best value of learning rate was investigated during the learning phase, taking into consideration that this parameter has significant effect on the training time and accuracy. The number of output nodes was taken 2, to represent the tissue class index in binary form. The sigmoid function was taken:

$$out = \frac{1}{1 + e^{-inp}} \quad (23)$$

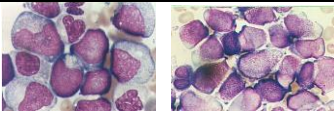
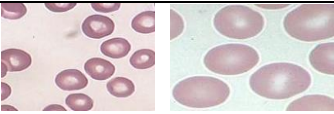
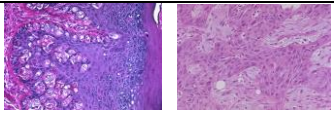
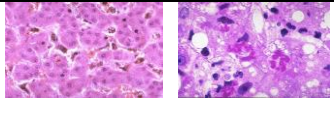
In general, to train the ANN many of the available data should be used, although it is not necessary to use them all. From the available training data a sufficient number of patterns are needed to be included in the training data set. The remaining data (i.e., 200) can be used to test the network to verify that the network can perform the desired mappings on the input vectors; which they have never been encountered during training.

3.RESULTS AND DISCUSSION

The data sets used in this study are sets of medical images taken from different sources. The first group of images used in this research has been collected from well-known medical atlas, and the second group images were taken from Kirkuk Educational Hospital. The second group was collected by capturing pre-diagnosed images using a digital camera connected to the microscope. The images are in true color (RGB) images and of varying sizes.

The samples listed in [Table \(1\)](#) present examples of various image classes. Short remarks are given with each class.

Table (1): Examples of image classes

Classes of images	Example of images	Remarks
Blood		Blood Cancer (Leukemia)
Blood		Normal Red Blood Cell
Breast		Breast Tissues
Liver		Liver Tissues

The main stages of the established system are: feature extraction and recognition using artificial neural network. The feature extraction unit has two parameters, namely; Co-occurrence jump step and roughness window size. The parameters of this stage have considerable effects on the discriminating power of extracted feature vector. In the recognition unit the artificial neural network has several parameters, namely; learning rate, number of hidden layer, and number of hidden nodes each one plays important role to achieve good recognition rate.

The Effect of Co-occurrence Matrix Jump Step

When using Co-occurrence matrix not only one jump step is adopted, several jump steps are tested to analyze each area. It is important to find the suitable value for it. In this work, the tested values are (1, 2, and 3). The assignment of jump step value is very important to get more accurate texture analysis for the image. **Table (2)** shows the effect of using different jump steps on system success rate.

Table (2): The effect of Co-occurrence jump step on system success rate

Jump step	Success rate
1	93%
2	92%
3	90%

A. The Effect of Roughness Window Size

In this set of tests the system success rate was determined for different window sizes to estimate the image surface roughness. Table III illustrates the effect of window size parameter on system success rate.

Table (3): The effect of window size on system success rate

Window size	Success rate
5	93%
7	91%
9	90%

B. The Effect of Number of Hidden Layers

When using ANN, not only single architecture is adopted, several architectures were tested. In this research project, the effect of number of hidden layers had been investigated to define the effect of number of hidden layers on system efficiency and ANN learning time. When using single hidden layer, the best attained system success rate was (85%), while for multi hidden layers neural network the attained efficiency was increased. Also, the effect of number of hidden layers on learning time was tested, taking into consideration when adding new additional hidden layer additional time is required to train the neural network, see **Table (4)** which illustrates the effect of this parameter on system success rate and learning time.

Table (4): Effect of No. of hidden layers on system success rate and learning time

No. of layers	Success rate	Learning time
1	85%	21 Minuet
2	93%	63 Minuet
3	93%	185 Minuet

C. The Effect of Number of Hidden Nodes on Success Rate and Learning Time

Deciding the number of neurons in the hidden layers is a very important part of deciding

your overall neural network architecture. In both layers the number of neurons in each of these hidden layers must be carefully considered. Table (5) shows the effect of the number of hidden nodes on system success rate and learning time of the ANN respectively.

Table (5): The effect of Number of hidden nodes on system success rate and learning time

No. of first layer hidden nodes	No. of second layer hidden nodes	Success Rate	Learning Time
45	15	87%	40 Minuet
55	25	90%	52 Minuet
65	35	93%	63 Minuet
70	40	93%	69 Minuet
75	43	93%	76 Minuet

D. The Effect of Learning Rate

One of the important parameters that affect the accuracy of multi-layer feed forward network is the learning rate; it is used to control the rate of weights adjustments. If the value of learning rate is too small then the learning process takes longer time; and if it is too large then the learning rate may disrupt all previous knowledge. There is no analytical method for finding the optimal learning rate; it is usually optimized empirically, just by trying different values. Table (6) shows the effect of learning rate on system success rate and learning time.

Table (6): The effect of learning rate on system success rate and learning time

Learning Rate	Success rate	Learning Time
0.1	93%	70 Minute
0.2	93%	66 Minute
0.3	92%	63 Minute
0.5	90%	59 Minute
0.8	89%	54 Minute
1	87%	49 Minute

As shown in above table several values of learning rate (i.e., 0.1, 0.2, 0.3, 0.5, 0.8, and 1) were tested, the test results indicated that less value give higher success rate and long learning time in comparison with those obtained when using larger values.

4.CONCLUSIONS

The use of 85 features extracted from Co-occurrence, run length matrixes, and roughness measure can be utilized to describe the textural content of various tissues.

A new idea based on taking advantage from using the histogram of residue between the original image and smoothed image to be used as indicator for roughness existence in an image. Using developed method in roughness feature extraction for textured images lead to more accurate results when combined with other traditional methods (i.e., Co-occurrence and run length) to overcome the weakness of these methods.

The established system gave better success rate (93%), when Co-occurrence jump taken 1, roughness window size taken 5, value of ANN hidden layers is set 2, the number of input nodes was 85, value of ANN first hidden nodes equal 65, value of ANN second hidden nodes equal 35, value of learning rate is set 0.2, the number of output nodes was 2, and the time required for training the neural network was 63 minutes.

REFERENCES

- [1] S. Karkanis, G. D. Magoulas, and N. Theofanous, “*Image Recognition and Neuronal Networks: Intelligent Systems for the Improvement of Imaging Information*”, Minimal Invasive and Applied Technologies, vol. 9, no. 3, pp. 225-230, 2000.
- [2] B. Arunadevi and S. N. Deepa, “*Brain Tumor Tissue Categorization in 3D Magnetic Resonance Images Using Improved PSO for Extreme Learning Machine*”, Progress In Electromagnetics Research, vol. 49, pp. 31-54, 2013.
- [3] H. A. Abbass, “*An Evolutionary Artificial Neural Networks Approach for Breast Cancer Diagnosis*”, Artificial Intelligence in Medicine, Published by Scientific Literature Digital Library (CiteSeer.IST), vol. 25, no. 3, pp. 265-281, 2002.



- [4] S Ribar, M. Dramićanin, T. Dramićanin and L. Matija, “*Classification of Breast Cancer Luminescence Data Using Self-Organizing Mapping Neural Network*”, Faculty of Mechanical Engineering, Belgrade (FME) Transactions, vol. 34, no. 2, pp. 87-91, 2006
- [5] L. E. George and E. Z. Mohammed, “*Cancer Tissues Recognition System Using Box Counting method and Artificial Neural Network*”, Proceedings of IEEE International Conference on Soft Computing and Pattern Recognition, ISBN. 978-1-4577-11947- ,pp. 5-9,Dalian, China, 2011.
- [6] A. Padma and R. Sukanesh, “*SVM Based Classification of Soft Tissues in Brain CT Images Using Wavelet Based Dominant Gray Level Run Length Texture Features*”, Middle-East Journal of Scientific Research, vol. 13, pp. 883-888, 2013.
- [7] A. Mohanty, S. Beberta, and S. Lenka, “*Classifying Benign and Malignant Mass using GLCM and GLRLM based Texture Features from Mammogram*”, International Journal of Engineering Research and Applications (IJERA), ISSN: 2248-9622, vol. 1, issue 3, pp. 687-693, 2001.
- [8] C. S. Kotapuri, and M. H. Rivjee, “*Image Search Engine Based on Combined Features of Image Sub-blocks*“, International Journal of Image Processing and Vision Sciences, ISSN: 2278 – 1110, vol. 1, issue 2, pp. 10-16, 2012.
- [9] S. M. Mukane, D. S. Bormane, and S. R. Gengaje, “*Wavelet and Co-occurrence Matrix Based Rotation Invariant Features for Texture Image Retrieval Using Fuzzy Logic*”, International Journal of Computer Applications, vol. 24, no. 7, 2011.
- [10] X. Tang, “*Texture Information in Run Length Matrices*”, IEEE Transactions on Image Processing, vol. 7, no. 11, 1998.
- [11] M. L. Antonie, O. R. Zaiane and A. Coman, "*Application of data mining techniques for medical image classification*", Proc. Of second international workshop on multimedia data mining, pp. 94-100, san Francisco, USA, 2001.



AUTHOR



Esraa Zeki Mohammed received B.Sc. in Computer Science from Mosul University / Mosul-Iraq in 2001, M.Sc and Ph.D. from Sulaimani University /Sulaimani-Iraq in 2009 and 2013, respectively. From 2002 till now she worked as a senior programmer and then head of advisory office in Ministry of Communication/State Company for Internet Services. Also She worked as lecturer in Kirkuk Technical Institute and Kirkuk University.